

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE INGENIERÍA

MAESTRÍA Y DOCTORADO EN CIENCIAS E INGENIERÍA



Valoración de registros de señales electromiográficas de antebrazo para el control de un sistema antebrazo/mano virtual.

T E S I S

Que presenta para obtener el grado de MAESTRO EN CIENCIAS

AMEZQUITA GARCIA JOSE ALEJANDRO

**DIRECTOR DE TESIS:
MIGUEL ENRIQUE BRAVO ZANOQUERA
CODIRECTOR DE TESIS
FERNANDO GONZALEZ NAVARRO**

MEXICALI, B. C.

AGOSTO DE 2017

AGRADECIMIENTOS

A mi familia específicamente a mi madre que supo enseñarme el amor por el aprendizaje.

A mis maestros específicamente a mi director de tesis y tutor por su dirección y su manera tan especial de guiarme, por ser un pionero en UABC en nuevas áreas de la tecnología y su compromiso por brindar más conocimiento.

A la UABC, a la Facultad de Ingeniería por tener programas tan innovadores y buenos, con maestros de calidad.

Al Consejo Nacional de Ciencia y Tecnología que permite desarrollar talento.

A México que tiene programas que permiten la investigación.

A mis compañeros y amigos que luchan por sus sueños.

¡Al mundo entero que nos ve crecer! Y por el que luchamos día con día.

Resumen

En esta tesis se desarrolló un modelo capaz de clasificar movimientos de los dedos a partir de señales electromiográficas (EMG) de antebrazo para sujetos individuales o para un grupo muestral de 7 sujetos, utilizando técnicas del aprendizaje automático, así como el procesamiento digital de señales. Se utilizó una base de datos de libre acceso de señales EMG de antebrazo que representan 15 tipos de movimientos de un grupo de 8 sujetos. Cada uno de los registros muestreados de la base de datos se evaluaron con métodos de minería de datos, para tratar de extraer la información necesaria para descifrar el tipo de movimiento realizado. Se confirmó que era de vital importancia procesar la señal antes de clasificarla porque la señal cruda del EMG posee información redundante. Se utilizaron segmentos de EMG de 62.5 ms para extraer características típicas que revelan información de esta señal. Se probaron distintas metodologías para tratar de reducir el costo computacional y que el modelo sea aplicable a un dispositivo para el control de una prótesis multifuncional, un exoesqueleto o para aplicaciones de tiempo-real basada en reconocer gestos con la mano. Este trabajo ha propuesto un modelo de clasificación con un alto porcentaje de reconocimiento de 91.37% para un grupo muestral de 7 sujetos o un promedio de 94.24% para cada uno de los sujetos por separado, ambos con una ventana de datos de 62.5 ms; lo que comparado con otras investigaciones presenta un avance de exactitud, con potencial a ser de tiempo-real. Los resultados de predicción del movimiento han sido aplicados a un programa de simulación musculoesquelético, Opensim (el cual es aceptado científicamente), transformados a un formato de ángulos de cada una de las articulaciones del modelo musculoesquelético para visualizar los movimientos.

ÍNDICE TEMATICO

AGRADECIMIENTOS.....	ii
RESUMEN.....	iii
ÍNDICE TEMATICO	iv
ÍNDICE DE FIGURAS	x
ÍNDICE DE TABLAS	xii
INTRODUCCIÓN	1
PLANTEAMIENTO DEL PROBLEMA	3
HIPÓTESIS.....	3
OBJETIVOS	3
METAS.....	4
JUSTIFICACIÓN.....	4
CAPITULO 1 ESTADO DEL ARTE	5
1.1 Introducción	5
1.2 Historia: instituciones y dispositivos	6
1.2.1 Empresas (dedicadas al desarrollo de prótesis).....	7
1.3 Prótesis	8
1.3.1 Prótesis de miembro superior.....	8
1.3.1.1 Tipos de prótesis	9
1.3.1.2 Prótesis y sistemas de control	10
1.3.1.3 Investigaciones más relevantes en minería de datos y aprendizaje automático para señales EMG.....	12
1.3.1.3.1 Características encontradas en la literatura	14
1.3.2 Tipos de amputaciones.....	15
1.3.3 Estadísticas de Amputación en el mundo.....	16
1.3.4 Exoesqueletos	17
1.4 Terapias de rehabilitación músculo esqueléticas y su evaluación.....	17
1.4.1 Distrofia muscular.....	17
CAPITULO 2 MARCO TEORICO.....	21
2.1 Localización anatómica (cuerpo humano)	21
2.1.1 Definición: planos y ejes de un modelo	22
2.1.1.1 Ejes del cuerpo	22

2.1.1.2 Planos de sección	22
2.1.2 Miembro superior.....	23
2.1.2.1 Formas exteriores del brazo	24
2.1.2.2 Secciones de interés (antebrazo-mano).....	24
2.1.2.2.1 Huesos.....	24
2.1.2.2.2 Músculos	25
2.2 Señal electromiográfica.....	27
2.2.1 Adquisición	28
2.2.1.1 Electroodos	28
2.2.1.1.1 Sensor EMG De-2.1 (Delsys)	29
2.2.1.1.2 Posición.....	29
2.3 Base de datos de sEMG encontrados en la red.....	29
2.3.1 Base de datos del Dr. Rami N. Khushaba para 15 tipos de movimientos	30
2.4 Minería de datos	32
2.4.1 Valores atípicos.....	33
2.4.1.1 Distancia Euclidiana	33
2.4.1.2 Distancia Mahalanobis.....	36
2.4.2 Ventaneo	37
2.4.2.1 Extracción de características	38
2.5 Aprendizaje automático.....	41
2.5.1 Descifrando la información.....	41
2.5.1.1 Aprendizaje supervisado	41
2.5.1.1.1 Clasificación.....	42
2.5.1.1.1.2 La necesidad de un enfoque probabilístico	42
2.5.1.1.2 Regresión	43
2.5.1.2 Aprendizaje no supervisado	43
2.5.2 Algunos conceptos básicos en aprendizaje automático.....	43
2.5.2.1 Modelos paramétricos contra no paramétricos	43
2.5.2.3 La maldición de la dimensionalidad	44
2.5.2.4 Modelos paramétricos para clasificación y regresión	44
2.5.2.4 Sobre ajuste	44
2.5.2.5 Métodos de visualización.....	44
2.6 Probabilidad	45

2.6.1 Teoría de probabilidad: Variables aleatorias discretas.....	45
2.6.2 Reglas fundamentales	46
2.6.2.1 Probabilidad de unión de dos eventos.....	46
2.6.2.2 Probabilidades conjuntas.....	46
2.6.2.3 Probabilidad condicional.....	46
2.6.2.4 Regla de Bayes.....	47
2.6.2.5 Media y varianza.....	48
2.6.3 Distribuciones discretas	48
2.6.3.1 Distribución binomial	48
2.6.3.2 Distribución multinomial	49
2.6.4 Distribuciones continuas Distribución normal (gaussiana).....	49
2.6.5 Distribuciones de probabilidades conjuntas.....	50
2.6.5.1 Covarianza y correlación	50
2.6.6 La multivariada normal o multivariada gaussiana	51
2.6.5.2 Distribución t student multivariada.....	51
2.6.6 Transformaciones de variables aleatorias	52
2.6.6.1 Transformaciones generales.....	53
2.6.7 Teorema del límite central	54
2.6.8 Modelos generativos	54
2.6.8.1 Concepto bayesiano	54
2.6.8.2 Probabilidad vs Verosimilitud (likelihood).....	55
2.6.8.3 Clasificador Naive Bayes.....	55
2.6.8.3.1 Ajuste del modelo	56
2.6.9 Modelos gaussianos: Análisis discriminante gaussiano.....	56
2.6.9.1 Análisis discriminante cuadrático (QDA).....	57
2.6.9.2 Análisis discriminante lineal (LDA).....	57
2.6.10 Transformaciones a un nuevo espacio para visualizar	58
2.6.11 Normalización.....	59
2.6.12 Selección secuencial de atributos.....	59
2.6.13 Trazo de las matrices dispersas y su proyección en el espacio de medias	61
2.6.13.1 Algoritmo ortogonalización de Gram-Schmidt.....	63
2.6.13.2 Blanqueamiento de los datos.....	63

2.6.13.2.1 Decorrelación: transformación de datos para tener una matriz de covarianzas diagonal.....	64
2.6.13.2.2 Blanqueamiento	65
2.6.13.2.3 Interpretación del blanqueamiento.....	66
2.6.13.3 Calculo de las matrices dispersas en el espacio de las medias de los grupos	66
2.7 Comunicación Matlab-Opensim	67
2.7.1 Matlab/Simulink	67
2.7.1.1 Procesos en tiempo real.....	67
2.7.2 Programas de simulación evaluados	67
2.7.2.1 Opensim	68
2.7.2.1.1 Crear y editar modelos:	69
2.7.2.1.2 Librería de modelos musculoesqueléticos	69
2.7.2.1.3 Analizar y simular modelos y movimientos.....	69
2.7.2.1.3 Modelo seleccionado.....	71
2.7.2.1.4 Programas utilizados para configurar el modelo.....	71
2.7.2.1.4.1 Rhinoceros 5	71
2.7.2.1.4.2 Notepad++.....	71
2.7.2.1.4.3 Paraview 5.0.1.....	72
2.7.3 Comunicación entre Opensim y Matlab.....	72
CAPÍTULO 3 EXPERIMENTACIÓN	73
3.1 Estructura de la base de datos evaluada	73
3.2 Valores atípicos	75
3.3 Ventaneo.....	76
3.4 Clasificación.....	77
3.4.1 Naive Bayes gaussiano.....	78
3.4.2 LDA	79
3.4.3 QDA.....	79
3.5 Normalización	80
3.6 Selección de características hacia adelante	81
3.7 Clasificación en la secuencia de experimentación 2	82
3.8 Algoritmo de matrices dispersas	82
3.9 Clasificación en la secuencia de experimentación 3	84
3.10 Secuencia de procesamiento cuando ya se tiene el modelo de clasificación creado.....	84

3.11 Técnica para aumentar el porcentaje de reconocimiento en un grupo muestral	85
3.11.1 Secuencia del procesamiento con las clases modificadas con el modelo de clasificación ya creado en las experimentaciones anteriores	86
3.12 Opensim	86
3.12.1 Configuración del modelo seleccionado en Opensim	87
3.12.1.2 Modelo Opensim.....	87
3.11.1.3 Configuración del modelo musculoesquelético, cambiando el origen de las figuras geométricas	89
3.12.2 Desarrollo de los movimientos	90
3.12.3 Estimación del movimiento en Matlab	92
3.12.4 Diagrama general de procesamiento	93
CAPÍTULO 4 RESULTADOS Y DISCUSIÓN	95
Resultados.....	95
4.1 Eliminación de valores atípicos.....	95
4.1.1 Ruta 1: Eliminación de valores atípicos por movimientos separados	95
4.1.2 Ruta 2: Eliminación de valores atípicos por movimientos en general	96
4.2 Ventaneo.....	99
4.2.1 Ventaneo Ruta 1.....	99
4.2.2 Ventaneo a través de la Ruta 2.....	100
4.3 Clasificación.....	102
4.3.2 Resultados de la experimentación inicial a través de la Ruta 2	102
4.4 Selección de atributos hacia adelante.....	103
4.4.1 Resultados de la Experimentación 2 a través de la Ruta 2.....	103
4.5 Algoritmo de matrices dispersas	108
4.6 Opensim	110
4.6.1 Articulaciones del modelo creado.....	110
4.6.3 Movimientos capaces de realizar	111
4.6.4 Movimientos reproducidos	112
4.6.5 Matrices configuradas para cada movimiento en Matlab.....	114
4.6.6 Operaciones y tiempo de procesamiento.....	118
DISCUSIÓN	119
CONCLUSIONES	126
ANEXOS	128

REFERENCIAS	144
-------------------	-----

ÍNDICE DE FIGURAS

Figura 1. 1 Diagrama de los primeros sistemas de control.	10
Figura 1. 2 Diagrama de control con retroalimentación.	11
Figura 1. 3 Diagrama de minería datos para controlar procesos con señales de entrada.	11
Figura 1. 4 Tipos de amputaciones a nivel de miembro superior.	16
Figura 1. 5 SIMM.	19
Figura 1. 6 Modelo desarrollado en SIMM y exportado a Opensim	20
Figura 2. 1 Modelo anatómico y sus planos imagen tomada de [26].	21
Figura 2. 2 Planos de un modelo anatómico tomado de [27].	23
Figura 2. 3 Secciones de interés: antebrazo y mano plano anterior.	24
Figura 2. 4 Electromiograma de Mano cerrada y sus 8 electrodos.	28
Figura 2. 5 Arreglo de electrodos utilizado en la base de datos con posición acorde a [13].	31
Figura 2. 6 Los 15 Movimientos evaluados, imagen tomada de Rami N. Khushaba en [13].	32
Figura 2. 7 Distancia entre dos puntos (dE) en el espacio determinado por E1 y E2.	34
Figura 2. 8 Nuevo par de puntos después de restar las medias del conjunto de datos.	35
Figura 3. 1 Secuencia de experimentación inicial, involucra procesos desde los datos crudos hasta la normalización para conocer y crear un primer modelo de clasificación referente.	80
Figura 3. 2 Secuencia de experimentación 2. Se agrega un algoritmo de selección de características (SFS) para mejorar el modelo de clasificación de referencia.	82
Figura 3. 3 secuencia de procesamiento para el algoritmo de matrices dispersas.	83
Figura 3. 4 Secuencia de experimentación 3. Adición de una conversión a un nuevo espacio para simplificar el modelo de clasificación de la experimentación 2.	84
Figura 3. 5 Diagrama para el procesamiento y evaluación de nuevos datos, en la experimentación 1 2 y 3.	85
Figura 3. 6 Archivo en formato .osim del modelo utilizado visto en notepad++, este es el formato como representa el código este programa se puede observar como el archivo del modelo se encuentra dividido en secciones que definen al modelo.	87
Figura 3. 7 Sección específica BodySet del archivo donde encontramos definida una figura geométrica con formato '.vtp'.	88
Figura 3. 8 Radius.vtp leído con paraview donde se le aplicara un filtro y se guardara con formato '.sto'	89
Figura 3. 9 Modelo terminado de mano con todos sus movimientos.	90
Figura 3. 10 Posición final de 10 de los 15 movimientos evaluados.	91
Figura 3. 11 Diagrama para la reproducción del movimiento a partir de las sEMG hasta Opensim.	93
Figura 3. 12 Proceso general para la creación de la matriz final de movimiento en Matlab. El único movimiento que posee dos matrices es mano cerrada para poder controlar el movimiento del dedo pulgar contra los otros dedos. Es por eso que posee un contador, cuando el movimiento mano cerrada ha tenido un número mayor de 18 empieza a activarse el movimiento 2 donde el pulgar empieza su movimiento.	94
Figura 4. 1 Valores atípicos por medio de la Ruta 1, donde M son los valores mantenidos y E los eliminados.	97

Figura 4. 2 Valores atípicos por medio de la Ruta 2, donde M son los valores mantenidos y E los eliminados.	98
Figura 4. 3 Visualización de los movimientos evaluados con solo 3 características en el nuevo espacio de las 14 originales para el sujeto número 6 después de SFS. Orientación donde es más visible la separación de las clases.	108
Figura 4. 4 Visualización de 8 movimientos evaluados con solo las primeras 3 características en el nuevo espacio de las 14 originales para el sujeto 6 después de SFS. Orientación donde es más visible la separación de las clases.	108
Figura 4. 5 Grafica resultante de la Tabla 4.22.	110
Figura 4. 6 Modelo Original de la muñeca.	111
Figura 4. 7 Modelo modificado de la muñeca, los dedos de la mano modificados poseen un color distinto, característica que puede ser modificada al gusto.	111
Figura 4. 8 Movimiento Flexión Índice	113
Figura 4. 9 Movimiento Flexión índice-medio	113
Figura 4. 10 Movimiento Flexión índice-medio-anular.	113
Figura 4. 11 Movimiento Mano cerrada.	113

ÍNDICE DE TABLAS

Tabla 1. 1 Investigaciones más relevantes en la interpretación del EMG con minería de datos y aprendizaje automático.	15
Tabla 2. 1 Músculos de antebrazo y función que realizan.	25
Tabla 2. 2 Músculos de la mano y función que realizan.	26
Tabla 2. 3 Coordenadas en el espacio con 8 dimensiones.	33
Tabla 3. 1 Matriz de datos para una repetición de un movimiento.	73
Tabla 3. 2 Matriz de datos para un movimiento con 12 repeticiones.....	74
Tabla 3. 3 Representación completa de los datos de un sujeto de prueba.....	74
Tabla 3. 4 Datos mantenidos de un sujeto de prueba después del procesamiento de valores atípicos.	75
Tabla 3. 5 Matriz para cada movimiento.	76
Tabla 3. 6 Ventana de 250 datos.	76
Tabla 3. 7 Nuevas características para un número de N ventanas (Nx96) por cada movimiento.	77
Tabla 3. 8 Matriz / movimiento	77
Tabla 3. 9 Matriz con los 15 movimientos y el número total de ventanas.....	77
Tabla 3. 10 pseudocódigo de SFS.....	81
Tabla 3. 11 A continuación se representa una sección del formato de un archivo ‘. mot’ para reproducir el movimiento en un modelo Opensim. Normal_close representa el nombre del modelo al que se le aplicará el movimiento, versión es la versión del archivo, nRows y nColumns son tanto las filas como las columnas del archivo completo, inDegrees indica si las mediciones están en grados, endheader indica el fin de los datos del encabezado para dar inicio a los datos en grados de cada una de las articulaciones, donde la primer columna representa el tiempo y las siguientes cada uno de los grados de libertad del modelo al que se le aplicará este archivo de movimiento (elbow_flexion, MCP2_flex, PIP:flex y DIP_flex).	91
Tabla 3. 12 Matriz para el movimiento predicho Donde cada fila corresponde a un vector de avance.....	93
Tabla 4. 1 Valores atípicos eliminados y mantenidos por sujetos por separado (Ruta 1).....	95
Tabla 4. 2 Valores atípicos eliminados y mantenidos por sujetos por separado (Ruta 2).....	96
Tabla 4. 3 Valores mantenidos después de la eliminación de valores atípicos por la Ruta 1	99
Tabla 4. 4 Número de ventanas de cada uno de los 8 sujetos con cada uno de sus 15 movimientos. Los datos que se encuentran en amarillo muestran los datos que hubieran pasado desapercibidos si no se hubiese llevado a cabo la ruta 1.	100
Tabla 4. 5 sEMG sujetos movimiento mano cerrada (HC) para los 8 sujetos y sus 12 repeticiones matriz de Mov. completos ruta 2.....	101
Tabla 4. 6 Resultados obtenidos a partir de la experimentación 1 mostrada en la Figura 3.1.	102
Tabla 4. 7 Resultados obtenidos a partir de la experimentación 1 en conjunto con el postprocesamiento donde las clases son ajustadas de 15 a 105 (técnica de extensión de clases). El modelo inicial para los 7 sujetos (fila 1) discrimina para las 105 clases.	103
Tabla 4. 8 Porcentajes de los modelos creados con cada uno de los algoritmos utilizados.	104

Tabla 4. 9 Número de características mantenidas en cada uno de los algoritmos para cada sujeto.....	104
Tabla 4. 10 Comparación entre el modelo normal y el modelo con clases ajustadas.	104
Tabla 4. 11 LDA Características Eliminadas de cada uno de los sujetos y posición de electrodo a la que pertenecen (Ruta 2).	106
Tabla 4. 12 QDA Características Eliminadas de cada uno de los sujetos y posición de electrodo a la que pertenecen (Ruta 2).	107
Tabla 4. 13 Tabla comparativa del porcentaje de reconocimiento del espacio original contra el nuevo espacio con únicamente 3 características.....	109
Tabla 4. 14 Tabla comparativa del grupo muestral 1-7 donde se evalúa el porcentaje de reconocimiento frente al número de características en el nuevo espacio, la columna de prueba es el resultado obtenido con la evaluación de las 105 clases, la columna de prueba con clases modificadas es después de haber reducido a 15 clases las 105 y haber calculado el error.	109
Tabla 4. 15 Comparación entre movimientos-Grados de libertad (Gdl) entre el modelo original y el modificado	112
Tabla 4. 16 Movimientos utilizados.....	112
Tabla 4. 17 vector de suma para mano cerrada cuando $\text{cont} < 18$	115
Tabla 4. 18 Vector de suma para mano cerrada cuando $\text{cont} > 18$	115
Tabla 4. 19 Vector de suma para flexión del dedo índice.	115
Tabla 4. 20 Vector de suma para flexión del dedo medio.....	115
Tabla 4. 21 Vector de suma para flexión del dedo anular.....	115
Tabla 4. 22 Vector de suma para flexión del dedo meñique.	116
Tabla 4. 23 Vector de suma para flexión del dedo pulgar-índice.	116
Tabla 4. 24 Vector de suma para flexión del dedo pulgar-medio.	116
Tabla 4. 25 Vector de suma para flexión del dedo pulgar-anular.	116
Tabla 4. 26 Vector de suma para flexión del dedo pulgar-meñique.	116
Tabla 4. 27 Vector de suma para flexión del dedo Índice-medio.	116
Tabla 4. 28 Vector de suma para flexión del dedo medio-anular.	117
Tabla 4. 29 Vector de suma para flexión del dedo anular-meñique.....	117
Tabla 4. 30 Vector de suma para flexión del dedo índice-medio-anular.	117
Tabla 4. 31 Vector de suma para flexión del dedo medio-anular-meñique.	117
Tabla 4. 32 Operaciones contra tiempo de procesamiento, la sección marcada en rojo indica el tiempo requerido para formar una ventana de 250 datos para iniciar el procesamiento, la amarilla es el tiempo de cada operación realizada a la ventana de 250 datos, las franjas negras indican las operaciones suprimidas porque no ayudan al modelo, la naranja el procesamiento de la transformación al nuevo espacio, el azul a la evaluación del modelo creado y por ultimo a la asignación de la matriz según el resultado del modelo de clasificación para el movimiento del modelo virtual de mano. Donde VM=Vector de medias de datos crudos, VC=Valor crítico, SC=subconjunto, VN=Vector nuevo de datos crudos, M=Vector medias de nuevas características, W=Matriz de blanqueo, bhat=Matriz de ortogonalización, Matriz de movimiento final=ceros.	118

Introducción

La ciencia y la tecnología avanzan a pasos apresurados, nuevas áreas nacen y otras van quedando en el olvido, esta tesis trabaja con áreas como la medicina, la biomecánica, la estadística, la probabilidad y con un área prácticamente nueva y explotada en las últimas décadas la minería de datos y el aprendizaje automático. Todas ellas con un propósito en esta investigación: el desarrollo de un sistema de control que utiliza las señales electromiográficas superficiales (sEMG) para el control de prótesis. Las investigaciones en prótesis nos muestran lo mismo: se busca comprender como funciona el sistema nervioso-musculoesquelético para reproducirlo y obtener un movimiento optimo, mejor que el que nosotros realizamos. Como mencionamos este propósito involucra áreas especializadas que pareciera que nuestro trabajo es muy ambicioso, si tomamos en cuenta el proceso desde la adquisición de las señales sEMG, la interface maquina usuario, la compresión del EMG ya que como veremos hasta el momento es indescifrable, su aplicación en el control de movimientos, y su implementación en una prótesis multifuncional, por esta razón decidimos enfocarnos únicamente en la comprensión de las señales, su procesamiento, el desarrollo de un modelo para un grupo muestral de n cantidad de sujetos que nos permita identificar movimientos. Además, pondremos a prueba el modelo creado en un software de simulación seleccionado entre los aceptados por la comunidad científica.

El presente trabajo realizará un estudio minucioso sobre el análisis de las investigaciones actuales de bases de datos de arreglos sEMG (electromiografía superficial) en el desarrollo de prótesis, planteará un método para analizar los movimientos de un miembro natural (específicamente los dedos de una mano), prediciendo movimientos y aplicando esta predicción a un software de simulación del cuerpo humano conocido en el ámbito científico como Opensim. Se creará una comunicación entre dos softwares de gran impacto en el ámbito científico: Opensim (programa de simulación del cuerpo humano) y Matlab (programa matemático), para el análisis de sistemas musculo esqueléticos que puede ser aplicado en terapias de rehabilitación y en el desarrollo de prótesis, control de exoesqueletos e incluso muchas de las nuevas tecnologías están empezando a utilizar un tipo de control más personalizada dejando de lado los teclados y los controles físicos por ejemplo el control gestual.

Los sistemas de análisis musculo esquelético comerciales son actualmente caros, el software va incluido con la venta de equipos, lo que demanda de una gran inversión. Por esta razón muy pocas veces son utilizados en hospitales incluso en los especializados.

Así encontramos que en el ambiente clínico no hay una herramienta que valore cuantitativamente la deficiencia muscular directamente sin la utilización de equipos comerciales, actualmente se utiliza valoración física directa además de ciertas pruebas para evaluar el estado muscular de un paciente. Este trabajo permitirá el análisis de movimientos de un brazo, a partir de una base de datos, que podría ser aplicado a terapias de rehabilitación y podrá ser utilizado como el ante trabajo para aplicación de estos modelos a una prótesis biónica. Aunque la pérdida de fuerza muscular no es una enfermedad como tal es más bien un signo de distintas enfermedades entre las que se pueden mencionar: la distrofia muscular de Duchenne, la distrofia miotónica, el Síndrome de Kugelberg-Welander y en la enfermedad de Pompe, que podrían ser evaluadas con la comunicación propuesta y sin la adquisición de un equipo costoso.

Planteamiento del problema

Aunque existen estudios sobre técnicas que valoran distintos métodos eficaces en la clasificación de movimientos definidos, existen pocos estudios que reportan completar el problema desde la valoración de las señales, la predicción de movimientos para el análisis de fuerzas biomecánicas en un sistema virtual de un sujeto en particular, para posibles aplicaciones en prótesis multifuncionales, exoesqueletos, terapias de rehabilitación y nuevas tecnologías como el actual control gestual de mano.

Hipótesis

La Electromiografía en arreglo de electrodos contiene la información necesaria para el control de una prótesis multifuncional en tiempo real, el desarrollo de un método de decodificación de información de sEMG en antebrazo y su aplicación a un sistema virtual permitirá realizar análisis musculo esqueléticos que apoyen el desarrollo de prótesis además apoyará como un sistema de análisis para la rehabilitación musculo esquelética a un determinado paciente o a un grupo de ellos.

Objetivos

Objetivo general

Crear un procedimiento específico aplicable al control de una prótesis multifuncional en tiempo real.

Objetivos específicos

Investigar nuevos enfoques para el problema de control mioelectrico y su aplicación en prótesis multifuncionales.

Crear un modelo capaz de descifrar el significado de las señales EMG

Idear una comunicación entre las plataformas de Opensim y Matlab para modelar los movimientos que se generan en un brazo humano a partir de arreglos de sEMG y puedan aplicarse en Opensim.

Metas

Evaluación de diferentes bases de datos sEMG, su control, su codificación y su decodificación para la implementación de una comunicación entre Matlab y Opensim.

Revisión de trabajos en control mioeléctrico y su aplicación en dispositivos reales.

Evaluación del software Opensim sus capacidades de simulación y procesamiento de datos.

Evaluación de artículos relacionados con herramientas de Opensim como CMC, dinámica inversa, dinámica hacia adelante, y optimización estática para definir activaciones musculares con registros analógicos de electromiogramas.

Justificación

El estudio minucioso de arreglos de EMG permitirá el uso de estas señales como controladores para una prótesis multifuncional. Las técnicas del aprendizaje automático y de minería de datos permiten generar un modelo de clasificación de movimientos que puede ser aplicado en el control de prótesis. Aunque en este trabajo no se implementará en una prótesis física, el modelo creado permitirá clasificar movimientos que pueda ser controlado por estas señales y visualizarlos biomecánicamente con el programa abierto de Opensim. También permitirá realizar análisis de movimientos que pueden ser utilizados como técnicas de evaluación muscular.

CAPITULO 1 ESTADO DEL ARTE

1.1 Introducción

Las enfermedades, así como la pérdida de algún miembro del cuerpo son circunstancias que el hombre ha padecido siempre, por lo tanto, ha tenido que lidiar con estos problemas desde el inicio de su existencia, aunque el enfrentarse con la pérdida de algún miembro del cuerpo ha evolucionado desde apoyos para caminar a ciertos tipos de prótesis cosméticas de brazo o pierna para brindar una mejor calidad de vida. Aún en la actualidad el remplazo del miembro perdido no es sustituido funcionalmente.

En las últimas décadas han surgido prótesis multifuncionales que incorporan grados de libertad que en décadas pasadas era imposible desarrollar, agregando funciones específicas enfocadas a cubrir las necesidades en las actividades de la vida diaria con una estética aceptable.

En este capítulo se mencionan los estudios de mayor impacto en el ámbito científico, se analizan los métodos de control para prótesis multifuncionales desarrolladas en instituciones y las encontradas comercialmente. La investigación, en el desarrollo de prótesis también ha sido objeto de competencia, entre los países más destacados encontramos EE. UU., Japón, Brasil y Alemania [3]. Entre las universidades más introducidas en el campo encontramos a la universidad de John Hopkins en el laboratorio de física aplicada donde poseen un proyecto que intenta recrear el control neuronal a miembro superior para aplicarlo a una prótesis de mano y la universidad de Dallas con el desarrollo de una prótesis activa de miembro inferior. Aunque son muchos los métodos utilizados, las estrategias de adquisición de datos, preprocesamiento, procesamiento, desarrollo de los modelos, evaluación de los sistemas de control, aún queda un gran hueco de información por descubrir para desarrollar una prótesis biónica, una que sea capaz de reproducir la fisiología humana natural, en este capítulo se retomará todo aquel desarrollo importante en esta área de investigación en las últimas décadas.

1.2 Historia: instituciones y dispositivos

El cuerpo humano esta tan especializado para realizar una infinidad de funciones. Con la pérdida de alguna parte de nuestro cuerpo el hombre ha ideado métodos alternativos para suplir la parte dañada y recuperar parte de la función perdida [1]. La evolución en el desarrollo de prótesis es compleja no hay una fecha exacta donde inicio, sabemos cuáles son los descubrimientos en la actualidad gracias a la difusión de la información y sabemos a dónde podemos llegar o hacia dónde vamos.

El largo y complejo desarrollo de las prótesis biónicas que encontramos en el presente son el resultado de esta difusión y cooperación entre investigaciones fallidas y exitosas, el avance en el diseño de prótesis ha estado ligado al progreso en los materiales utilizados, en el desarrollo tecnológico y en el entendimiento de la biomecánica del cuerpo humano [1].

En 1858, se identificó en Italia una pierna artificial que data aproximadamente del 300 a.C. Pertenecía a un amputado por debajo de la rodilla, fabricada con hierro y cobre con un núcleo de madera. Mas adelante entre el 23-79 d.C. Plinio el viejo escribió sobre un general romano amputado del brazo derecho que utilizaba una mano de hierro para sostener el escudo y así volver al campo de batalla [2]. En la Alta Edad Media hubo pocos avances, a un caballero se le colocaba una prótesis diseñada solamente para cumplir una función cosmética o para cumplir la función más esencial por ejemplo en amputados de pierna obtener un apoyo, por lo tanto, se prestaba poca atención a la funcionalidad. Los relojeros eran particularmente buenos para diseñar funciones extras a estos dispositivos con engranajes y resortes [2].

Se pudo determinar que en 1508 se crearon un par de manos de hierro para el mercenario alemán Gotz von Berlichingen. Contaba con un mecanismo que le permitía fijar la mano mecánica con la mano natural a base de resortes y mecanismos de liberación. [2].

En esa época existía Ambroise Paré barbero y cirujano francés considerado el padre de la cirugía de la amputación y del diseño protésico moderno. Modernizó procedimientos de amputación (1529) en la medicina y elaboro prótesis para amputados de miembro superior e inferior (1536). [2]. Paré crea la prótesis llamada Le petit Loraine brazo artificial a nivel de codo, con un mecanismo sencillo donde los dedos podían abrirse y cerrarse, contaba con una palanca para la flexión o extensión del antebrazo a nivel del codo. Un mecanismo relativamente novedoso además creó la primera mano estética de cuero, iniciado de esta forma una nueva utilización de materiales en el diseño de prótesis en miembro superior [2]. En estas épocas se empezaban a

desarrollar prótesis enfocadas a la función las direcciones en la investigación empezaban a diversificarse, en este punto predominaban las prótesis controladas por mecanismos pasivos que eran necesarios activar mecánicamente.

En Alemania inician las prótesis controladas por músculos del muñón gracias a *Sauerbruch*, de cierta forma logra idear como conectar los músculos de antebrazo con el mecanismo de la prótesis haciendo posible que la prótesis fuera de cierta forma activa. [2]

En 1960 en Rusia empieza a surgir el control mioeléctrico. Este tipo de prótesis funcionaba con impulsos eléctricos obtenidos de los músculos del muñón con funciones prensoras inicialmente se utilizó para amputados de antebrazo y su función era prensión de la mano con 2 kilos de fuerza [2].

El control en las prótesis empezó a ser un foco para la investigación, era necesario mejorar los resultados obtenidos hasta llegar a crear un miembro biónico, se empezó a investigar el control proporcional, a buscar patrones de reconocimiento en la señal EMG, a aplicar máquinas de estados, a aplicar algoritmos más avanzados como modelos probabilísticos.

Los países con mayor avance tecnológico e investigación en prótesis, según [2] son Alemania, Estados Unidos, Francia, Inglaterra y Japón.

El área de control mioeléctrico y su aplicación en prótesis multifuncionales toma un impulso con la llegada de la minería de datos donde los modelos de clasificación y la gran cantidad de algoritmos utilizados en la minería de datos empiezan a resolver o descifrar la información contenida en las señales electromiográficas.

1.2.1 Empresas (dedicadas al desarrollo de prótesis)

Dependiendo de la región geográfica, el mercado de dispositivos prostéticos y ortopédicos esta segmentado en 7 regiones clave: América del norte, América del sur, Europa oriental y Europa occidental, el pacífico de Asia Japón y medio oriente y África.

Norte América tiene mayor participación en el mercado mundial de dispositivos de prótesis ortopédicas seguida de Europa y Asia del pacifico. Las naciones en desarrollo de Asia Pacífico, Oriente Medio y África tienen un enorme potencial para el crecimiento en el mercado mundial de dispositivos de prótesis ortopédica, debido a las economías en desarrollo y los avances tecnológicos [3]

Los principales proveedores de prótesis a nivel mundial son: Hanger Inc., Otto Bock HealthCare GmbH, Endolite Ltd., Ossur hf., The Ohio Willow Wood Co. Fillauer LLC. Otto Bock

HealthCare GmbH en 2007 lanza su primer prototipo de prótesis de brazo con control mioeléctrico-neuronal, aunque las prótesis de mano no son su principal producto se está realizando investigación. Touch Bionics Inc, empresa recientemente adquirida por otra empresa dedicada al desarrollo de prótesis Ossur hf., con un amplio espectro de prótesis de miembro superior, uno de sus productos es i-limb quantum esta posee un sistema de control que permite cambiar los patrones de agarre con control por gestos, posee tipos predefinidos de agarre y es cosméticamente agradable, también desarrollaron la i-digits quantum es una prótesis para amputaciones parciales de mano, la amputación debe ser más distal a la muñeca también agrega el control patentado por gesto, y también entre sus productos encontramos la: living skin una prótesis funcional formada por silicona de alta definición y pintada para que coincida con el tono del paciente. Otra empresa es Fillauer LLC., aquí, aunque no precisamente su ramo es la producción y desarrollo de prótesis de miembro superior posee productos como la prótesis: nexo prótesis tipo gancho para realizar funciones específicas y prótesis pasivas cosméticas. Muchas de estas empresas están apoyadas o asesoradas por institutos de salud, por universidades y hasta por el gobierno. En México en particular el sistema médico del país no está relacionado con la provisión de estos dispositivos sin embargo existen grupos de ayuda para pacientes de bajos recursos que no pueden costearlo.

1.3 Prótesis

Una prótesis es una extensión artificial que reemplaza una parte perdida del cuerpo tal como un miembro superior o inferior, el desarrollo de prótesis es parte del campo de la ciencia de la biomecatronica, la ciencia que usa dispositivos mecánicos ayudándose con músculos, esqueletos y nervios humanos para asistir o mejorar el control motor perdido por trauma, enfermedad o defecto. Un miembro artificial es un tipo de prótesis que reemplaza una extremidad perdida, tales como brazos o piernas. El tipo de miembro artificial usado es determinado por la parte que ha sido perdida, los miembros artificiales son necesarios por una gran cantidad de razones, incluyendo enfermedades, accidentes y defectos congénitos [4].

1.3.1 Prótesis de miembro superior

La cantidad de prótesis existentes para miembro superior es igual al número de secciones que posee el miembro, por lo tanto existe una diversidad de prótesis para este miembro, las

configuraciones de estos dispositivos depende del nivel de amputación que existente en el paciente, por ejemplo amputaciones de mano parcial (ausencia de dedos), de antebrazo, de brazo, hombro, este tema más adelante será abordado, sin embargo, también existe una clasificación de las prótesis a partir de la función que son capaces de suplir, prótesis activas, pasivas, cosméticas, más específicamente con N grados de libertad. Actualmente están en desarrollo prótesis que son capaces de proveer sensaciones al paciente y una estabilidad prácticamente perfecta en los movimientos que son capaces de ejecutar.

Como se estableció en secciones anteriores el desarrollo de estos dispositivos a veces está ligado a empresas, a instituciones educativas o a asociaciones entre estas dos ramas.

1.3.1.1 Tipos de prótesis

Esta sección empezará con las prótesis comerciales, touchbionics una empresa que compró una patente que ha agregado a sus dispositivos: control por gestos (tecnología i-mo) posee prótesis de mano parciales llamada i-digits quantum este dispositivo es para amputaciones parciales de mano el individuo debe tener la amputación por encima de la muñeca, tiene también la i-limb quantum una prótesis de mano completa también con tecnología i-mo, la i-limb revolution, donde el movimiento del pulgar también es automático, además trae 36 tipos de agarre que le da una amplia gama de actividades que puede realizar, i-limb ultra que aunque posee menos agarres que la anterior (catorce tipos de agarre) y la activación del pulgar es mecánico es una prótesis cosméticamente aceptable con una amplia gama de movimientos. Supro wrist es una prótesis que permite el control de movimientos de la muñeca también con tecnología i-mo, muchos de estos equipos traen integrado un control de selección de agarres por medio de aplicaciones en celulares. También posee prótesis pasivas donde la función es cosmética prótesis formadas con silicona pintadas al tono de piel del paciente [5].

Fillauer otra empresa dedicada al desarrollo de prótesis posee una sección para amputaciones de miembro superior dedicada principalmente a dispositivos tipo gancho y prótesis pasivas cosméticas, su equipo tipo gancho: nexo es su producto más desarrollado para miembro superior [6]. En el ámbito académico encontramos a la universidad de Jonh Hopkins específicamente el laboratorio de física aplicada está ayudando a un programa de gobierno de los Estados Unidos (DARPA) para crear un miembro artificial controlado neuronalmente que restaurara la capacidad motora y sensorial para pacientes amputados de miembro superior, esta universidad está

liderando un equipo interdisciplinario constituyéndolo las instituciones de John Hopkins , agencias de gobierno, universidades y firmas privadas para implementar la visión de DARPA y proveer la más avanzada prótesis de extremidad superior. Otra empresa de desarrollo es DEKA esta posee una prótesis desarrollada (LUKE arm) principalmente para este programa gubernamental DARPA, el dispositivo aún no está a la venta. Así como la Universidad del Estado de Idaho USA está trabajando con sistemas de control para prótesis muchas universidades en Estados Unidos, y en muchos otros países también se encuentran tratando de abrir un camino al control biónico de las prótesis de miembro superior, sin embargo, como ya se ha concluido en investigaciones aún se necesita conocer la forma en la que funciona el cerebro para obtener un control más natural y eficiente de estos dispositivos.

1.3.1.2 Prótesis y sistemas de control

Los sistemas de control existentes para las prótesis de miembro superior son tan diversos que sería imposible enumerarlos todos, se podría decir que son tan diversos como lo son los temas de la minería de datos, que engloban temas del aprendizaje automático, la estadística, los enfoques probabilísticos, aunado a los sistemas de control moderno.

Inicialmente como se describió en secciones anteriores los dispositivos protésicos tenían una actuación meramente mecánica activada manualmente, con el avance de la tecnología estos dispositivos empezaron a activarse por medio de sistemas eléctricos/electrónicos utilizando sistemas de control donde el diagrama general puede ser expresado en la Figura 1.1.



Figura 1. 1 Diagrama de los primeros sistemas de control.

En un sistema de control para prótesis la señal de entrada pudiera ser una señal eléctrica (un botón para un movimiento en particular), o una señal de músculos, el controlador permite activar todos los mecanismos para un movimiento en particular y el proceso al movimiento final de la prótesis. Este tipo de dispositivos aún existen en el mercado sin embargo se sabe que son obsoletos, posteriormente este tipo de control empezó a ser desplazado por controles con retroalimentación, proporcionada por los sensores con los cuales era posible un movimiento más específico en función del grado de avance de un movimiento en particular, ver Figura 1.2.

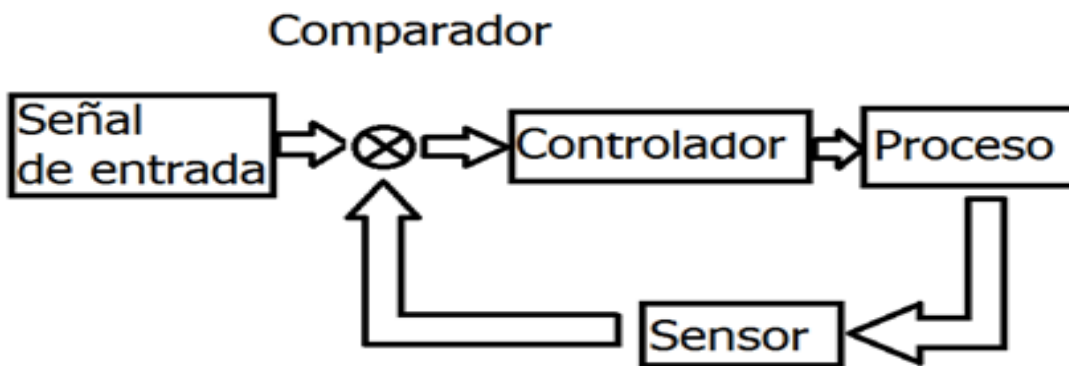


Figura 1. 2 Diagrama de control con retroalimentación.

En las últimas décadas ha crecido el interés por descifrar las señales de control naturales para los músculos para que sean aplicables a prótesis multifuncionales donde áreas como la minería de datos y el aprendizaje automático se están aplicando satisfactoriamente, el diagrama más simple que representa estas áreas, ver Figura 1.3.

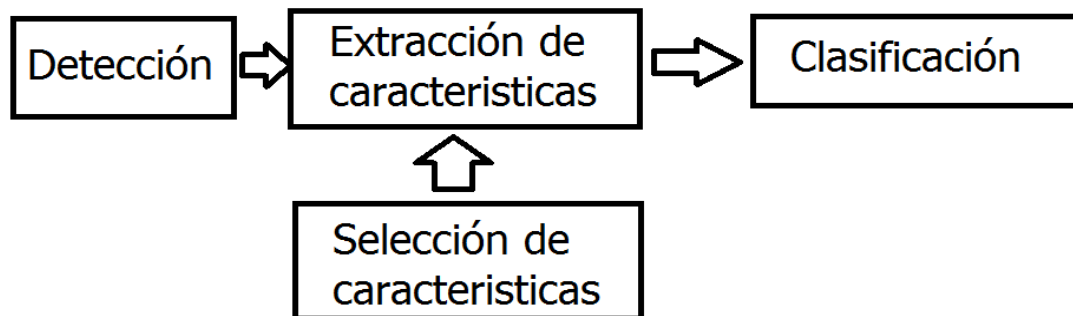


Figura 1. 3 Diagrama de minería de datos para controlar procesos con señales de entrada.

- Detección

La detección es el inicio del proceso, la adquisición de la señal de control en este caso la muscular adquirida por medio de electrodos de superficie, amplificada y muestreada en el tiempo para una cantidad definida de movimientos, es la traducción de procesos fisiológicos a una interpretación digital.

- Extracción de características

Dentro del EMG existe información contenida que no es apreciable directamente sobre la señal digital cruda de los músculos, por eso es necesario descifrarla para tratar de encontrar patrones al momento de clasificar los datos.

- Selección de características

Después de elegir las características que podrían representar la información que deseamos es momento de seleccionar dentro de nuestro conjunto, cuales son las que más ayudan a diferenciar las clases existentes.

Y por último la clasificación donde a partir de las características seleccionadas y las etiquetas de clase para cada medición se crea un modelo con técnicas de aprendizaje automático, para resolver problemas de clasificación como Naive Bayes, discriminantes lineales, discriminantes cuadráticos, redes neuronales etc.

Esta es una de las rutas típicas seguidas por las investigaciones en minería de datos y aprendizaje automático al momento de querer desarrollar un modelo capaz de clasificar señales EMG para establecer comandos de activación y desactivación para una prótesis.

1.3.1.3 Investigaciones más relevantes en minería de datos y aprendizaje automático para señales EMG.

El EMG es un conglomerado de señales individuales de fibras musculares todas superpuestas en determinando momento ante x movimiento. Su comprensión aún está en investigación ya que la relación que existe entre sistema nervioso central-músculos-movimiento aún no está definida.

Una señal mioeléctrica se ha definido que comprende dos estados: (i) un estado de transición generado por una activación de fibras musculares, cuando un musculo va del reposo a un nivel de contracción voluntaria y (ii) un estado estacionario emanado durante una constante contracción mantenida en un musculo [7].

Hudgins et al. [8] fueron los primeros en considerar el contenido de información de una señal transitoria que viene con el inicio de una contracción. Aunque las características extraídas de un estado transitorio, aproximadamente 100 ms después del inicio muestran una alta capacidad de clasificación, no está claro si esto es debido al determinismo electrofisiológico u otras razones, tales como potenciales de estiramiento de la piel o movimiento del electrodo. También se menciona que la principal debilidad de utilizar señales de transición es que el movimiento debe iniciar del reposo. Esto limita el cambio o la transición de clase a clase en la clasificación o en cierta forma limita la coordinación de complejas tareas envolviendo múltiples grados de libertad

[7]. Englehart [9] mostró que los datos de estado estacionario son clasificados con más exactitud que los datos de transición y la clasificación sufre menos degradación con longitudes de segmentos más cortos. La razón de clasificación decrece más rápidamente como la longitud de segmento de los datos de transición decrecen que con los datos estacionarios [7].

Una señal mioeléctrica tiene estados indeterminados durante la transición entre diferentes niveles de contracción, por lo tanto, la mayoría de los errores en clasificación ocurren cuando hay una transición entre clases [7].

Después de segmentar la longitud y el estado de los datos un tercer punto importante en la segmentación de los datos es la técnica de ventaneo de los datos. Existen dos técnicas principales: ventanas adyacentes y ventanas superpuestas [9].

En el ventaneo adyacente los segmentos disjuntos con una longitud definida se utilizan para extraer características y un movimiento clasificado surge después de cierto retardo de procesamiento, el procesamiento se hace en el tiempo del siguiente segmento evaluado. En el ventaneo superpuesto se toma también una longitud definida para todos los segmentos y después de saber cuánto es lo que dura el procesamiento para dar una respuesta de movimiento inicia la nueva adquisición de datos.

Englehart y Hudgins investigaron los efectos de un incremento de segmento más pequeño produce un flujo más denso, pero semi-redundante de decisiones de clase que podría mejorar el tiempo respuesta y la precisión. En [7] mencionan que la segmentación superpuesta aumenta el tiempo de procesamiento sin un aumento significativo en la precisión de características espectrales.

Segmentación Continua

La segmentación continua se basa en datos tanto de transición como de estado estacionario

En el análisis de sEMG para el control de prótesis, el aprendizaje automático logra la clasificación de movimientos a través de los métodos de extracción y selección de características donde genera nuevos dominios espaciales para aumentar la separación entre las clases evaluadas y de esta forma aumentar el porcentaje de reconocimiento.

Hoy en día aún se busca un modelo lo suficientemente discriminatorio para seleccionar decisiones de un gran conjunto de movimientos, y en base a este objetivo se ha empezado a enfocar la atención en los errores de los modelos generados y el porqué de ellos.

Por ejemplo, en [11], no solo se dedica a crear y a aplicar modelos de clasificación, sino que se centra en evaluar la clasificación errónea y corregirla. En [12], [13], trabajaron con diversas metodologías para generar un modelo que puede ser utilizado por una prótesis multifuncional, y crearon un repositorio en línea de la base de datos sEMG utilizada en su trabajo, que consideramos conveniente utilizar. También hay revisiones completas de prótesis multifuncionales y su control, tales como [7], [9], y [14]. Incluso algunas de estas investigaciones evalúan señales sEMG en sujetos amputados y sujetos sanos [11], [15].

1.3.1.3.1 Características encontradas en la literatura

La selección de características que serán evaluadas para generar el modelo es un tema relevante. Es deseable tener características que generen una diferenciación entre cada una de las clases evaluadas y una disminución de la dispersión entre las mediciones de una clase. Las características de sEMG se dividen generalmente en tres grupos: el dominio del tiempo, el dominio de la frecuencia y la representación tiempo-frecuencia, ([7]; [16]; [17]).

Características del dominio de tiempo

Características del dominio del tiempo se han utilizado ampliamente tanto en la medicina como en la ingeniería [16]. Una desventaja importante de las características de este grupo proviene de una propiedad no estacionaria de la señal EMG, las propiedades estadísticas de la señal a lo largo del tiempo cambian, para trabajar con las características del dominio del tiempo se asume que los datos provienen de una señal estacionaria [19].

El aprendizaje automático proporciona la base técnica de la minería de datos, algoritmos, métodos, herramientas que se utilizan para extraer información de las bases de datos. Los problemas de clasificación es común dividirlos en 2 etapas separadas, el paso de inferencia en el cual se usan los datos de entrenamiento para que el modelo aprenda las probabilidades de cada una de las clases dadas las variables predictoras $p(C_k | x)$ y el paso siguiente de decisión donde las probabilidades posteriori se usan para predecir las clases. Una posible alternativa sería resolver ambos problemas juntos y simplemente crear una función que mapea las entradas x directamente a decisiones. Esta función se denomina función discriminante [20].

En la Tabla 1.1 podemos ver las investigaciones más relevantes en el descifrado de la información contenida en el EMG se puede observar que la mayoría de las investigaciones superan el 90% de reconocimiento también se observa que ninguna tiene un grupo de estudio lo

suficientemente grande para hacer conclusiones estadísticas de una población, aunque algunas aplican sus modelos a prótesis o sistemas de simulación. Otras trabajan con amputados, incluso hacen comparaciones estadísticas sobre si hay diferencias significativas entre mediciones de sujetos sanos contra sujetos amputados.

Tabla 1. 1 Investigaciones más relevantes en la interpretación del EMG con minería de datos y aprendizaje automático.

Año, Autor	Num. Electr.	Num. mov.	Num. de sujetos	Tipos de movimientos	Características usadas	Sanos- amput.	Porcentaje obtenido	Porc. del grupo
Wuang 2006	2	4, 4 rep.	10	Pron. antebrazo, Hc, Ho, supin. antebrazo	Wavelet packed	no	93.75%, 86.25, 87.5	si
Huang 1999	2	8	2	Tipos de agarre	Tiempo, frecuencia	no	85% y 71%	si
Jiang 2005	2, 4	6	10	Ext flex 3 dedos I,M T	Wavelet packed	no	70,80%	si
Andrew 2009	2345678	4, 2 Rep.	12	ejercicios	Tiempo frecuencia	no	92%	no
Englehart 2003	4	4, 20 Rep.	12	Mano y muñeca por 5s	tiempo	no	94%	no
Cipriani 2011	8 pares elect bip	7, 3 rep trn, 3 val,	10 5-amp 5 sanos	7 mov dedos 5s		si	89% sanos 79% amp.	
Khushaba 2014	8	15	8	Dedos indiv. Y combinados 12 rep	Tiempo, frecuencia	no	95%	no
Tenore 2009	32,19	12, 30 Rep.	6	10 mov indiv. Y 2 comb. 25-30 rep	1 caract tiempo	si	90%	no

En secciones siguientes veremos el desglose completo de estas áreas en particular y las rutas clásicas seguidas por los investigadores para descifrar información.

1.3.2 Tipos de amputaciones

Existen amputaciones de miembro superior a nivel dedos, mano, muñeca, antebrazo, codo, brazo, hombro, cada una de ellas proporciona propiedades que son necesarias definir al momento de utilizar alguna prótesis adecuada al tipo de amputación, ver Figura 1.4.

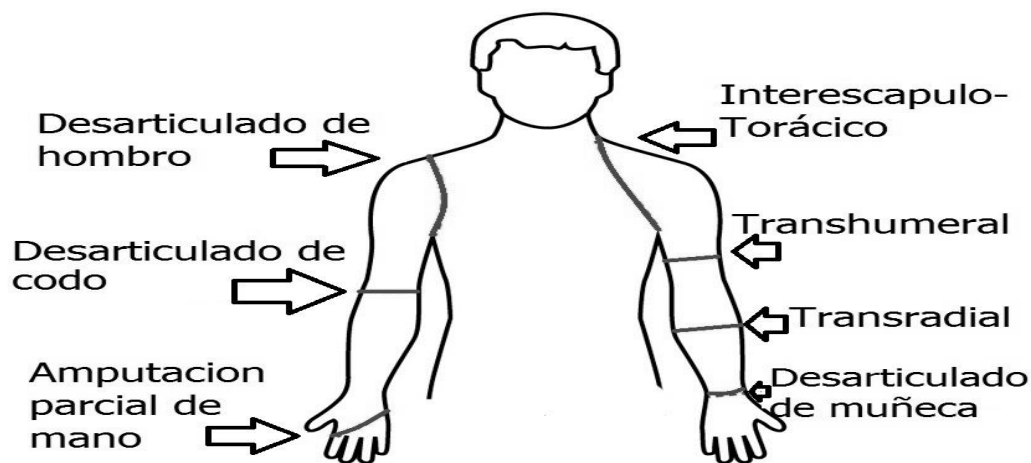


Figura 1. 4 Tipos de amputaciones a nivel de miembro superior.

1.3.3 Estadísticas de Amputación en el mundo

El primer reporte del mundo sobre discapacidad producido por la organización mundial de la salud y el banco mundial sugieren que más de mil millones de personas en el mundo hoy experimentan algún tipo de discapacidad. Estadísticamente en USA cada año existen 50,000 nuevas amputaciones basados en la información del centro nacional salud estadística (National Center for Health Statistics 1970) la razón entre miembros superiores amputados contra miembros inferiores es de 1:4, las amputaciones más comunes son amputaciones parciales de mano con pérdidas de 1 o más dedos, la siguiente más común es perdida de brazo existen 350,000 personas en los estados unidos 30% con amputaciones perdida de miembro superior, en 1998 se publicó que existían alrededor de dos millones de personas amputadas en los estados unidos [21], reportaron que las principales causas de estas cifras eran: enfermedades vasculares (54%) incluyendo diabetes y enfermedades arteriales periféricas, trauma(45%) y cáncer (menos del 2%) [21]. En 2009 el costo total de hospitales fue de 8.3 billones de dólares relacionados a problemas con amputaciones, mencionaron que aproximadamente 185,000 amputados ocurren en los estados unidos cada año (National Center for Health S. 1998) las personas que tienen enfermedades vasculares mueren dentro de 5 años, y las relacionadas con diabetes que sufrieron amputación de una pierna pierden la otra de dos a tres años. En una publicación [22], en 2008 establecen que existen alrededor 1.6 millones de personas viviendo con alguna perdida de algún miembro concluye diciendo que según la estadística actualmente existe una persona amputada cada 190 habitantes americanos y que para el 2050 este número se duplicara.

En México la estadística no es tan exacta hay estimaciones sobre discapacidad en general Al año 2010 según INEGI, las personas que tienen algún tipo de discapacidad son cinco millones setecientos treintainueve mil doscientos setenta, lo que representa 5.1% de la población total. Dentro de las cuales el 58.3% de esta población tiene un problema motriz o de movilidad sin especificar qué tipo de problema en específico se posee.

1.3.4 Exoesqueletos

Este tipo de dispositivos se utiliza más que nada para enfermedades donde se pierde el control muscular parcial o total del cuerpo humano como por ejemplo las enfermedades cerebrovasculares (ECV) donde el paciente sufre de una distrofia muscular en alguna sección del cuerpo, este tipo de herramientas son utilizadas como métodos para controlar el cuerpo las partes inmóviles y métodos de rehabilitación donde el paciente puede generar cierta fuerza en sus miembros y el aparato es capaz de adquirir datos de los movimientos del paciente y reproducirlos o generar fuerzas necesarias para ayudar al paciente a continuar con el movimiento. De igual forma los exoesqueletos están ligados a aumentar las capacidades en términos de fuerza para los usuarios aprovechándose de esto ha tenido oportunidad en el ámbito militar.

1.4 Terapias de rehabilitación músculo esqueléticas y su evaluación

Existen dos posibles problemas con el aparato locomotor: pérdida del miembro o pérdida de la función (movimiento), en países desarrollados existen laboratorios para análisis muscular esquelético los cuales están equipados con una gran cantidad de equipo por ejemplo cámaras de captura de movimiento, placas de fuerza donde se evalúa la marcha de los sujetos, equipo electromiográfico, para análisis de fuerzas, máquinas de control de movimiento para evaluación de fuerza y grado de movimiento en los brazos, además del software con el que viene equipado todo este laboratorio, en países en vías de desarrollo o países subdesarrollado el análisis de estos problemas es básicamente visual con alguna que otra técnica para valorar el grado de fuerza que posee un individuo al realizar movimientos.

1.4.1 Distrofia muscular

La distrofia muscular técnicamente es la pérdida de fuerza en los músculos para generar movimientos, puede ser causada por varias enfermedades, se les otorgo el nombre del que descubrió tal enfermedad distrofia muscular de Duchene, de Becker, de Emery-Dreifuss, o de la

localización anatómica de la afección como la distrofia muscular facioescapulohumeral, o la distrofia muscular de la cintura escapulohumeral. Las distrofias musculares son un grupo de padecimientos hereditarios. Pueden presentarse en la niñez o en la adultez [23].

1.4.2 Programas utilizados (SIMM, OPENSIM)

Los software de simulación empezaron su desarrollo al momento de que las capacidades graficas de los sistemas computacionales empezaron a desarrollarse, al contar con estas herramientas se creó una apertura en las investigaciones biomecánicas del cuerpo humano, la biomecánica como ciencia fue capaz de reproducir movimientos interactivamente para analizar todo tipo de estructuras anatómicas, análisis estáticos como dinámicos del sistema locomotor del ser humano algunos de los programas que se han considerado con este propósito son SIMM (software Interactivo para modelación musculo esquelética) y Opensim.

MusculoGraphics, Inc., una división de Motion Analysis Corporation, es una empresa de desarrollo de software especializada en la visualización, simulación y tecnologías de realidad virtual para aplicaciones [24]. La compañía fue creada en 1992 para ayudar a satisfacer los requerimientos de modelos que ayudasen a representar el cuerpo humano con alta precisión, su anatomía y su funcionamiento, productos de MusculoGraphics proporcionan representaciones dinámicas, visualmente atractivas de las estructuras anatómicas que dan una idea de la función subyacente del cuerpo [24]. Su producto principal es SIMM es un conjunto de herramientas de simulación y análisis mecánicos (dinámicos y estáticos) que facilitan el modelado, animación, y el análisis de los sistemas musculoesqueléticos en 3D además de agregarle la reproducción del movimiento en tiempo real [24].

En SIMM, un modelo musculo esquelético consiste en representaciones de los huesos, músculos, ligamentos y otras estructuras. Los músculos se extienden desde las articulaciones y son capaces de desarrollar fuerzas, generando así momentos en las articulaciones. SIMM permite un análisis de un modelo musculo esquelético mediante el cálculo de los momentos conjuntos que cada músculo puede generar en cualquier posición del cuerpo. Mediante la manipulación de un modelo mediante la interfaz gráfica, el usuario puede explorar rápidamente los efectos de los cambios en la geometría musculo esquelético y otros parámetros del modelo de fuerzas musculares y momentos articulares [24].

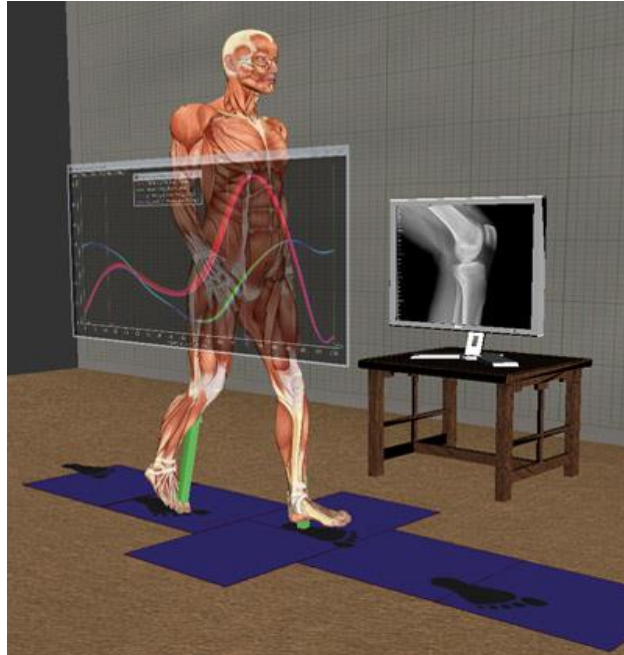


Figura 1. 5 SIMM.

Opensim es software para modelado de humanos, animales, robots, y el ambiente simulando su interacción y su movimiento. Opensim tiene una interfaz gráfica de usuario para la visualización de modelos generando y analizando simulaciones, fue desarrollado como una plataforma con código abierto pudiendo extender el software, también incluye una interfaz de programación de aplicaciones para desarrolladores pueden usar para extender el software [25]. Dentro de las capacidades que tiene Opensim está la creación y desarrollo de modelos genéricos o específicos, el análisis y simulación de movimientos, la personalización del flujo de trabajo y la extensión del código fuente del programa. Los creadores del proyecto poseen un objetivo muy ambicioso pretenden crear una comunidad mundial donde los investigadores de todo el mundo utilicen este software, lo completen, le agreguen nuevas funcionalidades, además poseen expertos en el software donde dan asesorías y crean seminarios.

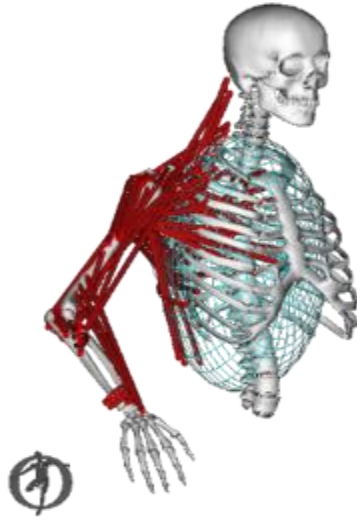


Figura 1. 6 Modelo desarrollado en SIMM y exportado a Opensim

CAPITULO 2 MARCO TEORICO

2.1 Localización anatómica (cuerpo humano)

El cuerpo humano es el medio que nos conecta al ambiente, nos desenvolvemos en él, todas nuestras actividades las realizamos a través de aquel, por lo cual existen ciencias para su estudio como la anatomía que es la ciencia que facilita su estudio, define, ubica y orienta cada parte del cuerpo, ayuda a otras ciencias a ampliar sus conocimientos como la biomecánica, la fisiología, la histología, entre otras.

El cuerpo humano consta de dos pares de miembros uno superior que se utiliza principalmente para la prensión y otro inferior para la locomoción, el miembro superior está constituido principalmente por dos segmentos, la cintura escapular y la porción del miembro libre.

Existen estructuras anatómicas que se explican mejor de una forma visual son tantas las posiciones para referirnos a determinada parte del cuerpo anterior, posterior, ventral, dorsal, lateral, proximal, distal entre otras como se ilustra en la imagen 2.1.

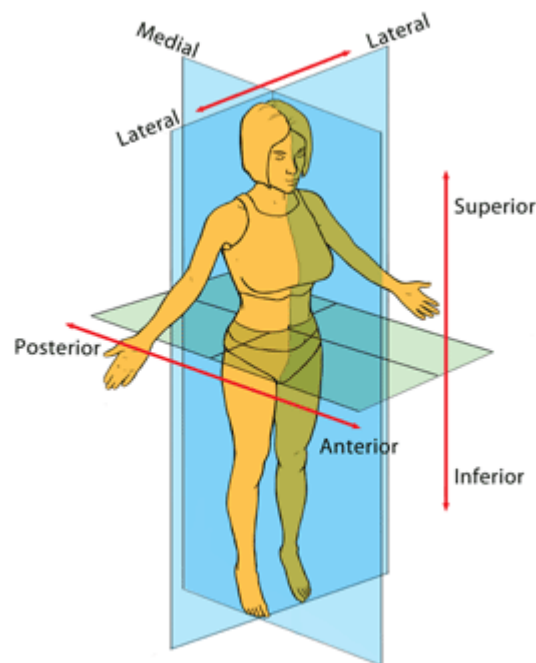


Figura 2. 1 Modelo anatómico y sus planos imagen tomada de [26].

2.1.1 Definición: planos y ejes de un modelo

En cualquier libro de anatomía encontramos que la descripción anatómica emplea comparaciones geométricas. Utilizando cilindros, cubos, vértices, caras, para la comparación de cuerpos y volúmenes. Se describen todo tipo de elementos geométricos como ángulos, circunferencias, áreas, diámetros y perímetros. Dentro de las comparaciones geométricas se emplean como referencias ejes y planos. [27]

El cuerpo para su estudio es dividido en planos y ejes cada una de las partes que lo conforman pueden ser localizadas en un eje o en un plano específico. Los ejes propuestos por Latarjet en [27] son el eje sagital, el longitudinal y el transversal, y los planos son planos sagitales, planos coronales y planos horizontales. Cada uno de estos ejes o planos son importantes al evaluar matemáticamente los modelos, un movimiento posee dirección, magnitud, punto de aplicación, etc. Los planos nos permiten localizar los músculos involucrados en determinada sección del modelo, músculos que a la vez están involucrados en la ejecución de los movimientos evaluados, Latarjet define varios planos y ejes en su tratado de anatomía.

2.1.1.1 Ejes del cuerpo

El primer eje por determinar es el eje sagital, se encuentra dentro del plano sagital (plano medial), también llamado anteroposterior o ventrodorsal se ubica como una flecha que atraviesa el cuerpo de adelante hacia atrás, aunque el eje es horizontal, el plano se podría decir que es vertical ver Fig. 2.2. Eje longitudinal, craneocaudal o supero inferior es de dirección vertical. Se dirige hacia abajo desde el punto más alto del cráneo, pasa por el centro de gravedad del cuerpo, con el eje sagital forma al plano sagital o medial. Eje transversal, laterolateral, es de dirección horizontal y perpendicular a los planos sagitales ver Fig. 2.1 y 2.2 [27].

2.1.1.2 Planos de sección

Para referirnos a estructuras anatómicas es necesario también definir planos algunos de ellos ya los hemos mencionado con la definición de los ejes y muchos de ellos comparten sus nombres. Los principales planos anatómicos son plano sagital cuando un plano sagital se encuentra juntamente en el centro del cuerpo analizado se le llama plano mediano, plano frontal, y plano transversal Fig. 2.2:

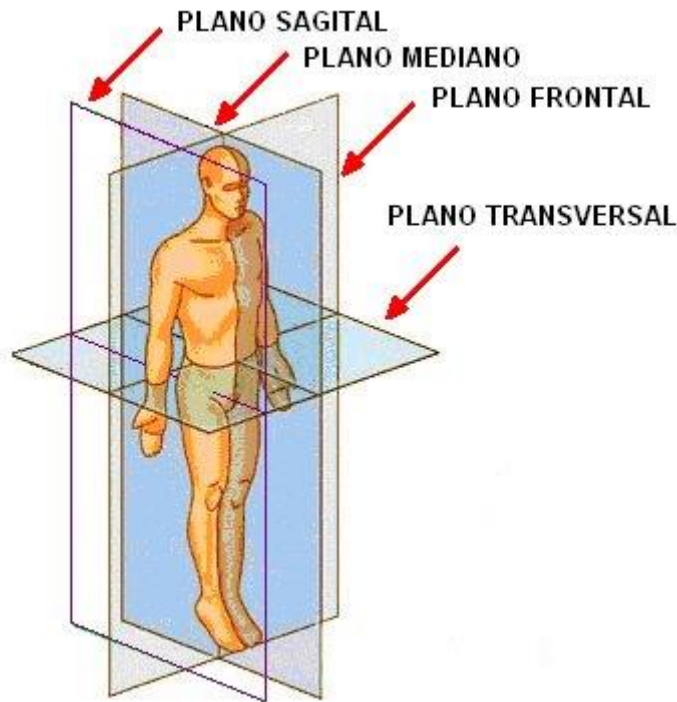


Figura 2. 2 Planos de un modelo anatómico tomado de [27].

Planos sagitales: son planos verticales, orientados en sentido anteroposterior conformado por los ejes sagital y longitudinal. Cuando se trata del plano medial el cuerpo se encuentra dividido en una mitad derecha y otra izquierda. Los planos sagitales que no se encuentran justamente en el centro del modelo son llamados paramedianos o parasagitales. Planos coroneales o frontales: son planos verticales que dividen al cuerpo en una parte anterior y una posterior, los ejes que lo componen son el eje longitudinal y el latero lateral. Planos horizontales o transversales: dividen al cuerpo en una parte inferior y una superior los ejes que lo componen son el eje sagital y el latero lateral.

2.1.2 Miembro superior

El miembro superior está constituido principalmente por dos secciones: La cintura escapular compuesta por escapula y clavícula y la porción libre del miembro superior conformada por humero, radio, cubito huesos del carpo, metacarpos, falanges, falangetas y falanginas. Estas secciones además de los huesos poseen una gran cantidad de músculos estriados separados por pleuras musculares las que permiten la contracción independiente de cada uno de ellos o de un

conjunto específico de ellos, también dentro del tejido encontramos una gran cantidad de vasos sanguíneos que son los encargados de transportar la sangre oxigenada (arterias), como la sangre desoxigenada (venas) además también encontramos vasos linfáticos, nódulos, nervios encargados de transportar las señales del Sistema Nervioso Central (SNC).

2.1.2.1 Formas exteriores del brazo

Además de la división por componentes funcionales en hombro, brazo, codo, antebrazo, muñeca, mano y dedos.

2.1.2.2 Secciones de interés (antebrazo-mano)

Hay 28 músculos en antebrazo-mano localizados en diferentes planos anatómicos: Plano profundo, Plano intermedio, Plano superficial, Fig. 2.3. Cada músculo posee capacidad de contracción. Dependiendo de la posición anatómica donde se encuentra adquiere su funcionalidad: músculos flexores, músculos extensores, músculos abductores, músculos aductores, etc. Los músculos propios de la mano poseen funciones más específicas flexión de las falanges más distales, extensión de las mismas, y gran parte de la movilidad del pulgar, además de los músculos abductores y aductores de los dedos.



Figura 2. 3 Secciones de interés: antebrazo y mano plano anterior.

2.1.2.2.1 Huesos

Antebrazo consta de dos huesos que lo conforma radio y cubito, mano consta de 27 huesos: 8 huesos de carpo, 5 metacarpos, 5 falanges proximales, 4 falanges medias y 5 falanges distales. Cada uno de los huesos que componen antebrazo y mano consta de inserciones musculares que controlan esa parte del cuerpo (proporcionan movimiento).

2.1.2.2.2 Músculos

Antebrazo consta con 19 músculos, mano consta de 11 de los cuales en conjunto están encargados de los movimientos de supinación y pronación de mano, encargados de flexión/extensión de las falanges proximales, medias y distales, así como el movimiento independiente del pulgar: flexión, extensión, aducción y abducción, los músculos propios de la mano poseen la especialización de los movimientos de esta.

El desarrollar la función en la que están involucrados cada uno de los músculos y la magnitud en la que intervienen en cada uno de los movimientos podría convertirse textualmente en un sacrificio para el lector por lo que se anexara en una tabla donde se involucre musculo y su función tanto para antebrazo como para mano Tabla 2.1 y 2.2 respectivamente (ver anexo 1 para localización de músculos).

Tabla 2. 1 Músculos de antebrazo y función que realizan.

Músculos de antebrazo	Función realizada
Pronador cuadrado	Sitúa el antebrazo y la mano en pronación.
Flexor profundo de los dedos	Flexiona la falange distal sobre la falange media proximal y esta sobre el hueso metacarpiano y la mano sobre el antebrazo.
Flexor largo del pulgar	Flexiona la falange distal sobre la proximal y esta sobre el hueso metacarpiano.
Músculo flexor superficial de los dedos	Flexiona la falange media sobre la falange proximal y esta sobre el metacarpiano y la mano sobre el antebrazo.
Músculo pronador redondo	Imprime al antebrazo un movimiento de rotación que desplaza el dedo pulgar medialmente y la palma de la mano posteriormente.
Músculo flexor radial del carpo	Flexor de la mano sobre el antebrazo.
Músculo palmar largo	Flexor de la mano sobre el antebrazo.
Músculo flexor cubital del carpo	Flexor y aductor de la mano.
Músculo supinador	Imprime al antebrazo un movimiento de rotación que desplaza el dedo pulgar lateralmente y pone la palma de la

	mano anteriormente.
Músculo extensor radial corto del carpo	Extensor y abductor de la mano.
Músculo extensor radial largo del carpo	Extensor y abductor de la mano.
Músculo braquiorraquial	Flexor del antebrazo sobre el brazo.
Músculo abductor largo del pulgar	Desplaza el pulgar anterior y lateralmente.
Músculo extensor corto del pulgar	Es extensor y abductor del dedo pulgar y de su metacarpiano.
Músculo extensor largo del pulgar	Extiende la falange distal sobre la proximal, la proximal sobre el metacarpiano y este sobre el carpo.
Músculo extensor del índice	Extensor del índice.
Músculo extensor de los dedos	Extiende la falange media y distal sobre la proximal y esta sobre los metacarpos y estos sobre el antebrazo.
Músculo extensor del meñique	Extensor del meñique.
Ancóneo	Extensor del antebrazo.

Tabla 2. 2 Músculos de la mano y función que realizan.

Músculos de la mano	Función realizada
Músculos interóseos dorsales	Flexionan la falange proximal y extienden las otras dos. separan del eje de la mano los dedos donde se insertan.
Músculos interóseos palmares	Flexiona la falange proximal y extienden las otras dos. Aproximan al eje de la mano los dedos donde se insertan.
Aductor del pulgar	Aductor del dedo pulgar.
Flexor corto del pulgar	Desplaza el dedo pulgar medial y anteriormente (sobre todo aductor).
Oponente del pulgar	Desplaza el primer metacarpiano anterior y medialmente.

Abductor corto del pulgar	Desplaza lateral y anteriormente el primer metacarpiano y la falange proximal del dedo pulgar (abductor).
Oponente del menique	Desplaza el dedo menique anterior y lateralmente lo opone al dedo pulgar.
Flexor corto del menique	Flexor del dedo menique.
Abductor del menique	Flexor y abductor del dedo menique.
Palmar corto	Pliega la piel de la eminencia hipotenar.
Lumbricales	Flexionar la primera falange y extender el resto.

2.2 Señal electromiográfica

La señal electromiográfica es aquella señal eléctrica que se genera en músculo a partir de una contracción de este, una definición particular de Carlo J. De Luca menciona que *“la electromiografía provee fácil acceso a procesos fisiológicos que causa el musculo para generar fuerza, producir movimiento, y generar las múltiples funciones que nos permiten interactuar con el mundo que nos rodea”*, coincidiendo con su idea la señal registrada superficialmente es la generada por los músculos al momento de contraerse o activarse, no se sabe en si cual es la señal de control de entrada, si los músculos están controlados independientemente o estos están controlados en sincronía donde una señal de control activa varios para un determinado movimiento, además de este detalle es muy difícil manejar la señal separadamente de activaciones musculares aisladas, el fenómeno de comunicación cruzada es un efecto que de cualquier forma que tomes la muestra estará presente. Como se puede observar en la Figura 2.4 el total del movimiento registrado es un conglomerado de señales interpuestas en un mismo electrodo además de la interposición con los otros electrodos en el tiempo. Aún falta entender al sistema nervioso central para entender la fisiología del sistema muscular.

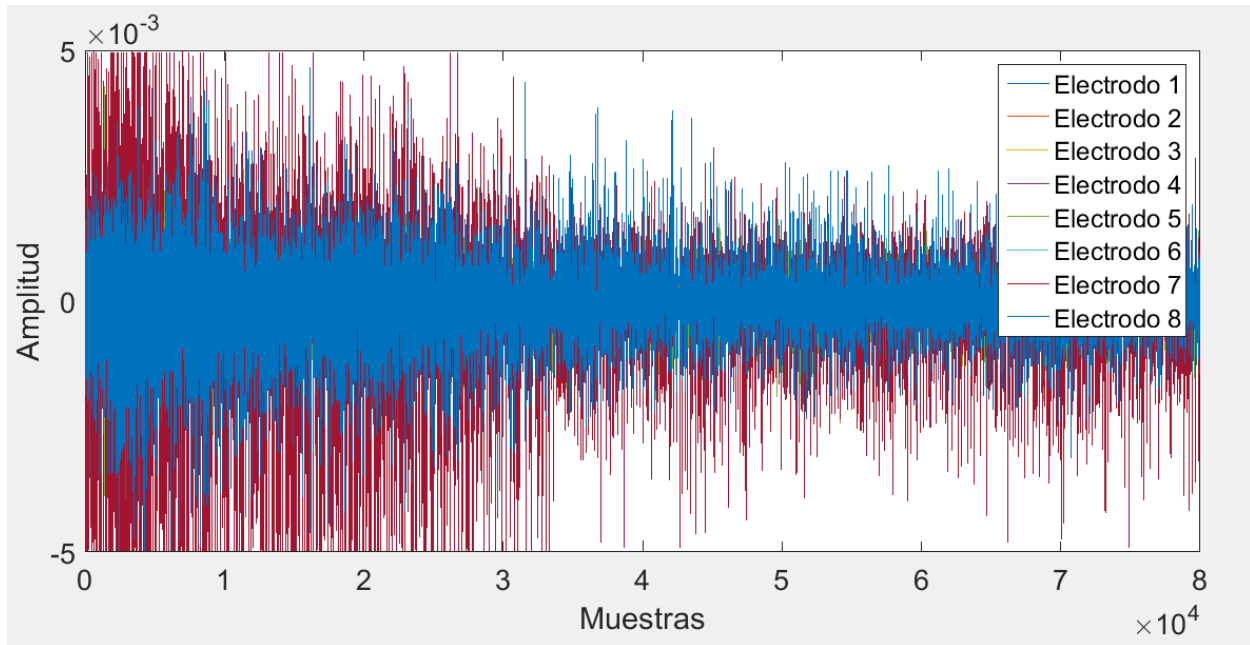


Figura 2. 4 Electromiograma de Mano cerrada y sus 8 electrodos.

2.2.1 Adquisición

Existen diversos sistemas de adquisición de datos para este tipo de señales, desde equipos sencillos con una frecuencia de muestro normal hasta para registrar esas señales tan altas que la cantidad de datos es desorbitante, se sabe que la frecuencia muscular se encuentra localizada en una región entre 5 y 500 Hertz por lo que de acuerdo con el teorema de Nyquist con una frecuencia de 1000 Hertz es suficiente.

2.2.1.1 Electrodo

Es posible encontrar en el mercado gran cantidad de formas y tipos de electrodos, principalmente electrodos superficiales comúnmente utilizados para evitar los problemas que conlleva utilizar electrodos invasivos como agujas, donde las infecciones y los daños a tejidos muchas veces es imposible evitar, dentro de los electrodos superficiales encontramos tipos palanca, pulsera, chupón, hasta guantes y adhesivos.

Un electrodo es un sensor capaz de registrar señales eléctricas, en sistemas vivos podrían considerarse ejemplos de estas señales las electrocardiográficas, encefalografías y en nuestro caso señales mioeléctricas, técnicamente está definido por la real academia española como *el extremo de un conductor que registra una señal eléctrica en contacto con el medio*. Existen una gran cantidad de materiales y configuraciones para los electrodos, el Doctor Rami N. Kushaba y

su equipo que realizaron la base de datos que utilizamos utilizaron electrodos bipolares que pertenecen a la empresa Delsys, serie 2.x sensores EMG. Estos sensores fueron diseñados para la adquisición de señales EMG reales y sin complicaciones. Dichos sensores activos son diseñados para detectar y optimizar las señales EMG registradas en la superficie de la piel con la posibilidad de rechazar señales de ruido comunes tales como movimiento y artefactos de los cables cumpliendo con una excelente razón señal a ruido [28].

2.2.1.1.1 Sensor EMG De-2.1 (Delsys)

EL sensor EMG DE-2.1 (Delsys) sustrae los potenciales detectados en dos distintas localizaciones sobre la superficie de la piel directamente sobre un musculo activo. Los potenciales EMG son siempre medidos con respecto al potencial eléctrico de un sitio neutral localizado a un lado de la fuente de la señal EMG. Este potencial es detectado por el electrodo de referencia. El sensor es diseñado usando una geometría de contacto de barras paralelas para asegurar la estabilidad de la señal, repetitividad entre registros y una óptima representación del contenido de la frecuencia. Diseñado para cumplir con el registro tanto para músculos cortos y largos [28].

2.2.1.1.2 Posición

La cantidad de músculos distribuidos en una región tan pequeña como antebrazo es un problema al cual nos enfrentamos. El lugar óptimo para obtener la señal es en el centro del cuerpo muscular donde la mayoría del trabajo se lleva a cabo. Encontrar un lugar optimo donde el centro muscular de todos los músculos coincidan es una tarea difícil, la tarea es encontrar un punto de localización más cercano a la mayoría de los cuerpos musculares útiles para los movimientos definidos.

2.3 Base de datos de sEMG encontrados en la red

Actualmente la investigación se está centrando en áreas específicas de la ciencia, la electrónica en microelectrónica, la minería de datos en aprendizaje profundo, en sí la especialización de la ciencia es un hecho y lo vivimos día a día. Debido a esto, como es parte de nuestros objetivos el poder analizar y desarrollar un sistema de control con sEMG se decidió buscar en la red bases de datos utilizadas por otros investigadores que estuvieran disponibles.

La investigación para poder avanzar de una forma más rápida está requiriendo por parte de los investigadores ahora más que nunca más claridad en sus trabajos, que suban todos aquellos materiales o elementos que utilizaron para reproducir sus experimentos, bases de datos, algoritmos, softwares, aplicaciones, todo esto por parte de la diversificación que está afligiendo a la ciencia, por ejemplo, un área prácticamente nueva en la ciencia es la minería de datos, el buscar información contenida en los datos que está oculta, para esto existen infinidad de métodos tanto clásicos, como nuevos, ha habido una variante de esta tarea que es el aprendizaje profundo en la que áreas como adquisición de datos, como el procesamiento de estos ya pasa a segundo plano, por eso mismo ahora el investigador puede especializarse en áreas como el pre procesamiento de la información o la adquisición de estos y construir su carrera en ellos sin embargo, para lograr esto es requerimiento poner a disposición el material resultante de las investigaciones para que la ciencia avance con más rapidez.

Aún pese a estas circunstancias la disponibilidad de los datos es un factor importante en el desarrollo del conocimiento. De las bases que se pudieron encontrar para este trabajo fueron solamente tres, dos de un mismo doctor (Dr. Rami N. Khushaba) y otra perteneciente al repositorio de la UCI (Universidad de California Irvine). De las tres bases de datos mencionadas se determinó que la base de datos numero 2 (15 movimientos) del Dr. Rami N. Khushaba era la que utilizaríamos ya que cuenta con un mayor número de movimientos, por lo tanto, son más específicos, la forma de adquisición nos pareció conveniente porque es la que utilizan en varias investigaciones, siendo un arreglo circular colocado en una posición definida en antebrazo y con un mayor número de electrodos.

2.3.1 Base de datos del Dr. Rami N. Khushaba para 15 tipos de movimientos

En el artículo “*Electromyogram (EMG) Feature Reduction Using Mutual Components Analysis for Multifunction Prosthetic Fingers Control*”. El Doctor Rami N. Khushaba recolectó datos de ocho sujetos, seis hombres y dos mujeres de edades comprendidas entre los 25 y los 30 años, las personas fueron seleccionadas aleatoriamente en la ciudad de Sídney Australia, fueron reclutados para realizar movimientos seleccionados de dedos. Los sujetos tenían sus miembros en estado normal sin desordenes neuronales o musculares. Todos los participantes estaban de acuerdo en participar en el estudio, los sujetos fueron sentados en una silla con su brazo soportado y flexionado en una posición. La base de datos fue registrada usando 8 canales montados a través

de la circunferencia del antebrazo y procesado por un sistema Bagnoli EMG de Delsys Inc., los electrodos fueron posicionados con una pegatina en la piel, el electrodo de referencia fue colocado en la muñeca ver Figura 2.5. Las señales EMG colectadas fueron amplificadas usando un Delsys Bagnoli amplificador de 8 canales a una ganancia total de 1000. Un convertidor analógico digital de 12 bit (National Instruments, BNC-2090) fue usado para muestrear la señal a 4000 Hz; los datos de la señal fueron adquiridos usando un programa de adquisición Delsys EMG Works. La señal fue filtrada con un filtro pasa banda entre 20 y 450 Hz con un filtro implementado para remover la línea de interferencia de 50Hz [13].

Quince clases de movimientos fueron colectados durante este experimento incluyendo la flexión de cada uno de los dedos individuales: pulgar (T), índice(I), medio (M), Anular (A),meñique (L), y la combinación pulgar-índice (T-I), pulgar-medio (T-M), pulgar-anular(T-R), pulgar-meñique (T-L), índice-medio(I-M), medio-anular(M-R), anular-meñique (RL), Índice-medio-anular (I M R), Medio-Anular-Meñique (MRL) y finalmente la mano cerrada (HC) ver Figura 2.6 [13].

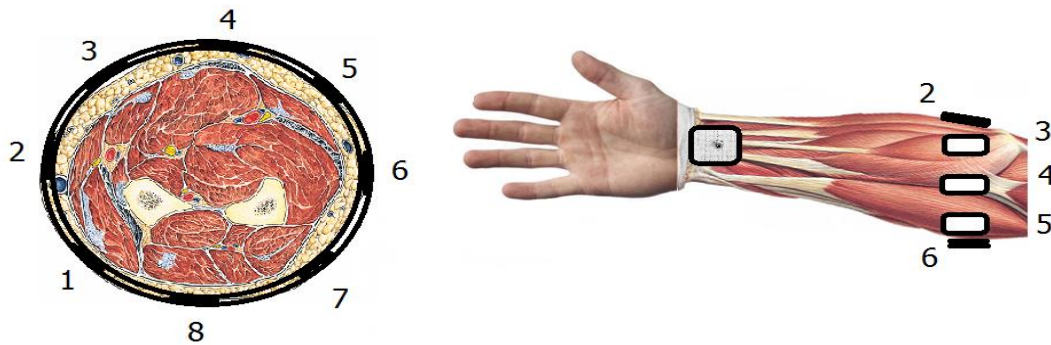


Figura 2. 5 Arreglo de electrodos utilizado en la base de datos con posición acorde a [13].

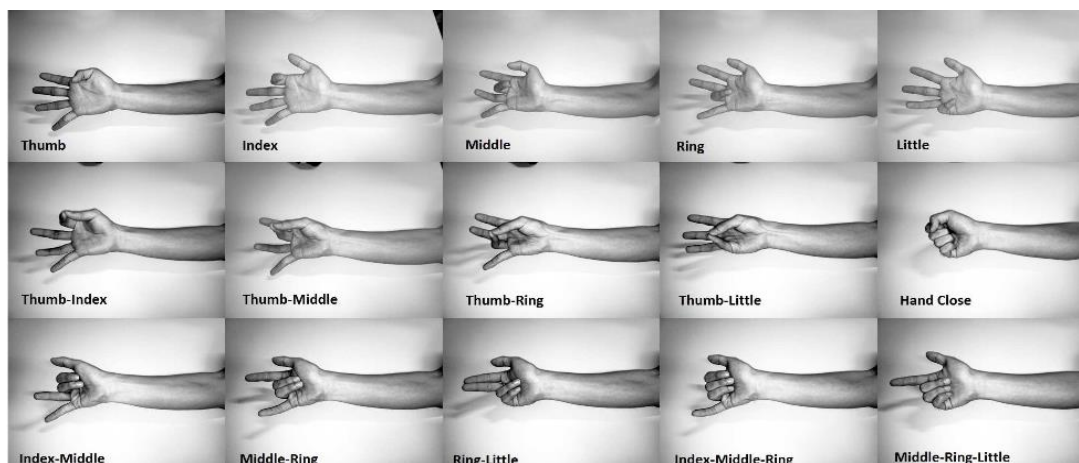


Figura 2. 6 Los 15 Movimientos evaluados, imagen tomada de Rami N. Khushaba en [13].

Cuando se colectaron los datos, los sujetos fueron llamados a realizar cada movimiento por separado de los quince que se evaluaron, se realizó cada movimiento por un periodo de cinco segundos en cada prueba, doce repeticiones en cada movimiento fueron realizadas.

2.4 Minería de datos

Las nuevas rutas que ha tomado la investigación para crear conocimiento han necesitado la unificación de ciencias para dar nacimiento a otras nuevas, de esta forma nace la minería de datos, un área nueva creada a partir de ciencias como la comunicación, la computación y la estadística para el procesamiento de datos, donde se busca descubrir información. Vivimos en una sociedad que se alimenta de información, sin embargo, la mayor parte de la información se adquiere de una forma cruda: datos. Y si consideramos a los datos como hechos registrados en el tiempo, entonces la información es el conjunto de patrones que subyacen a los datos [29]. Hay una gran cantidad de información contenida en los datos y es el objetivo de la minería de datos revelarla.

Una definición de minería de datos es “La minería de datos es la extracción de la información implícita previamente desconocida y potencialmente usable de los datos” [29]. La idea es crear programas que sean capaces de extraer estos patrones automáticamente, en nuestro caso la utilizaremos para descubrir la información que nos es realmente útil en la señal electromiográfica, nos apoyaremos en el aprendizaje automático para revelar ciertos patrones importantes en esta señal y modelarlos para diferenciar clases de movimientos para poder crear

un modelo y hacer predicciones con datos futuros. Unos de los problemas con que se enfrenta la minería de datos es la existencia de datos que no nos dicen nada acerca de la información que necesitamos, algunas señales poseen demasiada información redundante, en algunos casos no sabemos cuál de las características evaluadas nos está siendo útil y cual no, en esta parte la minería de datos también te ofrece algoritmos que te permiten visualizar tus datos para de esta forma descubrir el comportamiento de estos, también proporciona métodos que te permiten despreciar o mantener variables útiles, en particular la señal sEMG por la forma de adquisición, y la fuente de esta mucha de la información contenida es redundante por lo tanto es necesario visualizar y seleccionar los datos

2.4.1 Valores atípicos

Técnicamente un valor atípico es un dato, una observación dentro un conjunto de datos u observaciones que es numéricamente distante al resto del conjunto. Una de las métricas más utilizadas para medir las distancias entre los datos es la distancia mahalanobis que mide la distancia de una observación en particular al centro de masa del conjunto de los datos y por ello es necesario describir lo que es la distancia mahalanobis.

Para describir la distancia Mahalanobis es necesario hacer una comparación entre una distancia ya conocida como la distancia euclidiana.

2.4.1.1 Distancia Euclidiana

La distancia euclidiana se puede representar como la distancia entre dos puntos en el espacio euclídeo, usualmente es la distancia que se mide con una regla, esta distancia se obtiene a partir del teorema de Pitágoras. De forma práctica podemos representar una distancia entre dos puntos P y Q. Donde P y Q pueden tener n dimensiones. Para representar los datos que evaluaremos en nuestro trabajo cada punto en el espacio posee 8 dimensiones que representarían el arreglo utilizado por 8 electrodos.

Tabla 2. 3 Coordenadas en el espacio con 8 dimensiones.

Punto	E1	E2	E3	E4	E5	E6	E7	E8
P=	(p1,	p2,	p3,	p4,	p5,	p6,	p7,	p8)

Q=	(q1,	q2,	q3,	q4,	q5,	q6,	q7,	q8)
----	------	-----	-----	-----	-----	-----	-----	-----

La Tabla 2.1 muestra dos puntos en el espacio de 8 dimensiones, técnicamente no podemos representar gráficamente estos puntos, para realizar el análisis solo tomaremos las primeras dos dimensiones E1, y E2 que representan las mediciones de los electrodos 1 y 2.

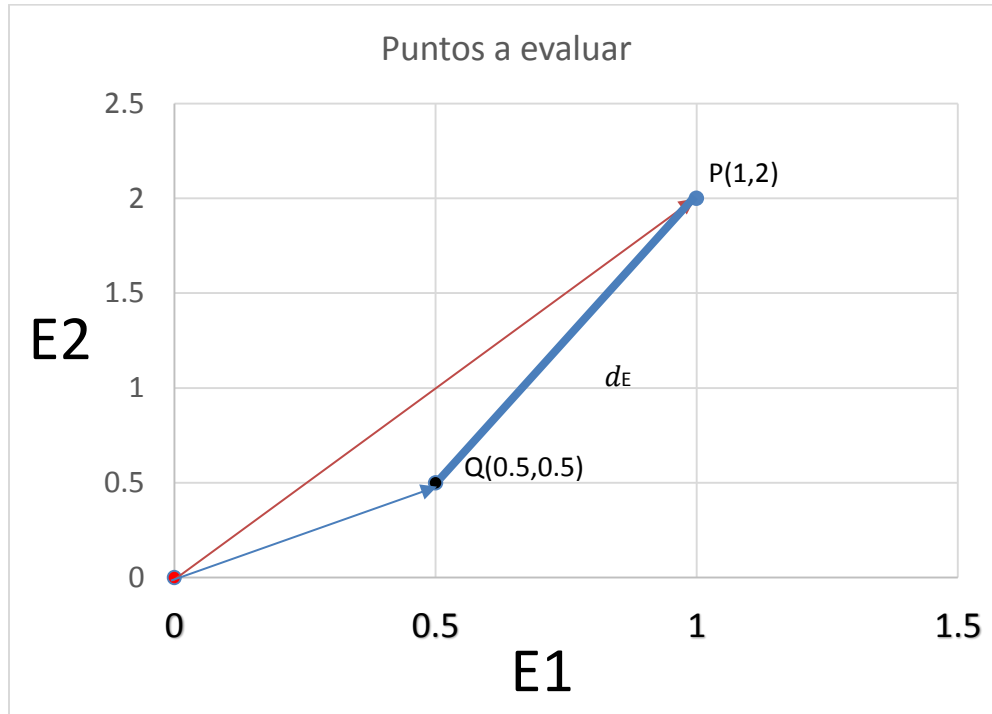


Figura 2. 7 Distancia entre dos puntos (d_E) en el espacio determinado por E1 y E2.

$$d_E = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \quad (1)$$

O en notación vectorial

$$d_E = \sqrt{(\vec{P} - \vec{Q})^T (\vec{P} - \vec{Q})} \quad (2)$$

Como se mencionó en la definición de valor atípico buscamos la distancia entre un punto y el centro de masa del conjunto de datos al que pertenece, por lo tanto, no queremos sacar la distancia entre puntos sino la distancia entre cada uno de los puntos y el centro de masa de todos

los datos, procedemos a calcular la media de cada eje evaluado ($E_1, E_2, \dots, E_n$) y restársela a cada uno de los datos.

$$\overline{X_{E1}} = \frac{1 + 0.5}{2} = 0.75$$

$$\overline{X_{E2}} = \frac{2 + 0.5}{2} = 1.25$$

Se restan las medias obtenidas a cada uno de los puntos evaluados y obtenemos la siguiente grafica

$$P = (p1 - \overline{X_{E1}}, p2 - \overline{X_{E2}})$$

$$P = (1 - 0.75, 2 - 1.25) = (0.25, 0.75)$$

$$Q = (q1 - \overline{X_{E1}}, q2 - \overline{X_{E2}})$$

$$Q = (0.5 - 0.75, 0.5 - 1.25) = (-0.25, -0.75)$$

Ahora P y Q representan nuevos puntos en el espacio E_1 y E_2 donde ya han sido restado las medias del conjunto total (E_1 y E_2)

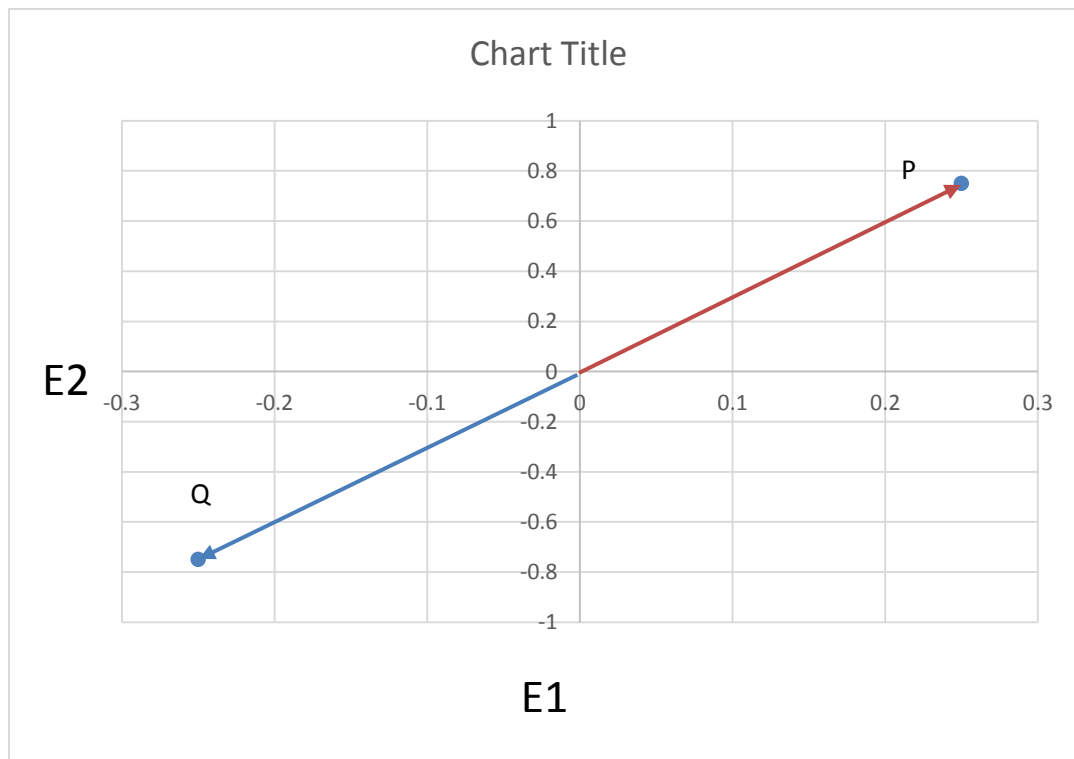


Figura 2. 8 Nuevo par de puntos después de restar las medias del conjunto de datos.

Como ahora las distancias se valoran con respecto al origen (centro de masa del conjunto de datos), el punto dos es el origen (0,0) y por la tanto se puede representar la distancia como sigue:

$$\begin{aligned} d_E P &= \sqrt{(\vec{P} - \vec{0})^T (\vec{P} - \vec{0})} \\ d_E P &= \sqrt{(\vec{P})^T (\vec{P})} \end{aligned} \quad (3)$$

$$\begin{aligned} d_E Q &= \sqrt{(\vec{Q} - \vec{0})^T (\vec{Q} - \vec{0})} \\ d_E Q &= \sqrt{(Q)^T (\vec{Q})} \end{aligned} \quad (4)$$

Donde $d_E P$ y $d_E Q$ representan las distancias Euclidianas de los puntos P y Q al origen. De esta forma es como obtenemos esta distancia, ahora pasemos a la distancia mahalanobis.

2.4.1.2 Distancia Mahalanobis

La distancia mahalanobis es la distancia entre un punto y un centroide de un espacio multivariado, a diferencia de la distancia euclidiana esta utiliza la correlación entre las variables multidimensionales. Volviendo con los puntos utilizados en la distancia Euclidiana tenemos:

$$d_M = \sqrt{\frac{(p_1 - q_1)^2}{\sigma_1} + \frac{(p_2 - q_2)^2}{\sigma_2}} \quad (5)$$

O en notación vectorial

$$d_M = \sqrt{(\vec{P} - \vec{Q})^T S^{-1} (\vec{P} - \vec{Q})} \quad (6)$$

Donde d_M representa la distancia Mahalanobis del punto P al Q, σ_1 y σ_2 representan las desviaciones estándar de cada una de las variables evaluadas, E1 y E2 en este caso, y S la matriz de varianzas.

Debido a que no podemos afirmar que las variables son independientes unas de las otras (no podemos afirmar que los valores de E1 y E2 son independiente) cambiamos la matriz de varianzas S por la matriz de covarianzas Σ .

$$d_M = \sqrt{(\vec{P} - \vec{Q})^T \Sigma^{-1} (\vec{P} - \vec{Q})} \quad (7)$$

Aplicando el mismo procedimiento que en la distancia euclidiana restamos la media a cada uno de los puntos en nuestro análisis para quedarnos con la fórmula:

$$d_M P = \sqrt{(\vec{P})^T \Sigma^{-1} (\vec{P})} \quad (8)$$

$$d_M Q = \sqrt{(\vec{Q})^T \Sigma^{-1} (\vec{Q})} \quad (9)$$

Ya con nuestra matriz de distancias pasamos a evaluar con el valor crítico de una distribución chi cuadrada cuáles son valores atípicos y cuáles no.

La teoría de la distancia de Mahalanobis asume que los datos multivariados están normalmente distribuidos, por lo tanto, la distancia de Mahalanobis de las muestras sigue una distribución Chi cuadrada con d grados de libertad (Por definición la suma de variables aleatorias normales estándar sigue una distribución chi cuadrada con d grados de libertad) [37]. Donde los grados de libertad están dados por el número de valores que pueden ser asignados de forma arbitraria, después de haber impuesto alguna restricción a los mismos, técnicamente podríamos decir que los grados de libertad son una función del tamaño de la muestra y del número de variables n , donde $n-1$ representan los grados de libertad. En nuestro caso seleccionamos 7 grados de libertad ya que el total de nuestras variables son 8 [36].

2.4.2 Ventaneo

El ventanear o segmentar los datos es una actividad comúnmente utilizada en el análisis de la señal electromiográfica para extraer información útil de la señal. Un segmento es una sección en el tiempo de los datos mioeléctricos adquiridos para la extracción de características. Como el objetivo de nuestro trabajo es desarrollar un control en tiempo real con algún retroceso en el tiempo asegurándonos que sea el menor posible, uno de los puntos a considerar es la longitud del segmento adyacente más el tiempo de procesamiento para generar los comandos de clasificación para el control debe ser igual o menor a 300 ms según [7]. Además, la longitud de un segmento debe ser lo suficientemente largo ya que la varianza de las características aumenta cuando la

longitud del segmento decrece y consecuentemente se reduce el rendimiento de la clasificación. Por lo tanto, un compromiso entre respuesta en el tiempo y precisión existe. Sin embargo, algunos autores [7] resaltaron que, adoptando una segmentación continua sobre una señal estacionaria, la longitud del segmento puede ser reducido a 128 ms o incluso a 32 ms sin una considerable disminución en la precisión.

2.4.2.1 Extracción de características

Las características en el dominio del tiempo están siendo usadas en investigaciones y en la práctica ingenieril y médica [16], la principal desventaja de las características de este grupo viene de la propiedad no estacionaria de la señal EMG, cambiando las propiedades estadísticas sobre el tiempo; en las características en el dominio del tiempo se asumen los datos como una señal estacionaria. [19]. Diez características en el dominio del tiempo están siendo propuestas en esta investigación de las cuales encontramos a (suponiendo que x es la señal de entrada los datos de un electrodo, i es el número de muestras $i=1$ hasta $i=N$, y N es el número total de muestras):

Valor absoluto medio (MAV)

El valor absoluto medio es uno de los más utilizados en análisis de señales EMG [16]. MAV es un promedio del valor absoluto de la amplitud de la señal en un segmento el cual puede ser definido como:

$$MAV = 1/N \sum_{i=1}^N |x_i| \quad (10)$$

Integral cuadrada simple (SSI)

Integral cuadrada simple (SSI) o integral cuadrada usa la energía de la señal EMG como una característica [16]. Generalmente este parámetro es definido como un índice de energía en la señal y puede ser expresado como:

$$SSI = \sum_{i=1}^N x_i^2 \quad (11)$$

La varianza del EMG

La varianza del EMG (VAR) es otro índice de energía en la señal [16]:

$$VAR = \frac{1}{N-1} \sum_{i=1}^N x_i^2 \quad (12)$$

Momentos temporales (TM3, TM4, TM5)

Momentos temporales (TM3, TM4, TM5) es un análisis estadístico para ser usado en prótesis de mano TM3, TM4, TM5, por [16] en su estudio sobre reducción y selección de características para clasificación de señales EMG:

$$TM3 = \left| \frac{1}{N} \sum_{i=1}^N x_i^3 \right|; TM4 = \left| \frac{1}{N} \sum_{i=1}^N x_i^4 \right|; TM5 = \left| \frac{1}{N} \sum_{i=1}^N x_i^5 \right| \quad (13)$$

Raíz cuadrática media (RMS)

Raíz cuadrática media (RMS) es otra característica popular en el análisis de señales electromiográficas, esto es modulado como la amplitud de un proceso aleatorio modulado gaussiano es similar al método de la desviación estándar [16]

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (14)$$

Longitud de forma de onda (WL)

Longitud de forma de onda (WL) es una medida de la complejidad de la señal EMG [16]. Esta es definida como longitud acumulativa de la forma de onda en un segmento de tiempo. Puede ser calculado como

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (15)$$

Cruces por cero (ZC)

Cruces por cero (ZC) es una medida de la frecuencia de la señal definida en el dominio del tiempo [16]. Es el número de veces que la amplitud pasa por el límite establecido como valor

cero. Puede establecerse a través de $x_i > 0$ y $x_{i+1} < 0$, o $x_i < 0$ y $x_{i+1} > 0$ y $|x_i - x_{i+1}| \geq umbral$. Dicho matemáticamente de otra forma tenemos:

$$ZC = \sum_{i=1}^{N-1} [\text{signo}(x_i * x_{i+1}) \cap |x_i - x_{i+1}| \geq umbral] \quad (16)$$

$$\text{signo}(x) = \begin{cases} 1 & \text{if } x \geq umbral \\ 0 & \text{en otro caso} \end{cases}$$

Cambio en el promedio de la amplitud (AAC)

Cambio en el promedio de la amplitud (AAC) es cercanamente equivalente a WL excepto que la longitud de onda se promedia [16]. Y puede expresarse como:

$$ACC = \frac{1}{N} \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (17)$$

Porcentaje de miopulso (MYOP)

Porcentaje de miopulso (MYOP) es un valor promedio de la salida del miopulso el cual es definido como el valor absoluto cuando el valor de la amplitud excede cierto umbral [16] y puede formularse así:

$$MYOP = \frac{1}{N} \sum_{i=1}^N [f(x_i)] \quad f(x) = \begin{cases} 1 & \text{if } x \geq umbral \\ 0 & \text{en otro caso} \end{cases} \quad (18)$$

Cambio de signo de pendiente (SSC)

Cambio de signo de pendiente (SSC) está relacionada con características como ZC, MYOP [16]. Este es otro método que representa información en frecuencia de la señal EMG:

$$SSC = \sum_{i=1}^{N-1} [f[(x_i - x_{i-1}) * (x_i - x_{i+1})]] \quad (19)$$

$$f(x) = \begin{cases} 1 & \text{if } x \geq \text{umbral} \\ 0 & \text{en otro caso} \end{cases}$$

En donde el umbral usado fue igual a 70 mV en estas características es normal seleccionar un umbral entre 50 μ V y 100 mV [16].

2.5 Aprendizaje automático

El aprendizaje automático (*Machine Learning*) provee las bases técnicas de la minería de datos, los algoritmos, los métodos, las herramientas que son usadas para extraer la información de las bases de datos. Generalmente se divide en dos partes el aprendizaje predictivo o supervisado y el descriptivo o no supervisado.

2.5.1 Descifrando la información

2.5.1.1 Aprendizaje supervisado

Existen dos corrientes principales en la evaluación de los datos para la solución de problemas con aprendizaje supervisado: clasificación y regresión, en las dos corrientes si no se conoce la naturaleza de los datos es muy difícil establecer la mejor estrategia para tratarlos, por lo tanto, es conveniente visualizar la estructura de los datos, la visualización de la información se abordará más adelante. Los problemas de clasificación por ejemplo son problemas en los que se involucran la separación de los datos con etiquetas de clase, donde a partir de las nuevas variables predictoras (variables independientes) se crea un modelo que es capaz de categorizar de acuerdo a las nuevas variables independientes la clase a la que pertenece. En los problemas de regresión un conjunto de datos (variables independientes) está relacionado con otro conjunto de mediciones (variable a predecir) generalmente continua en el tiempo, por ejemplo, el precio promedio de una casa en específico dadas ciertas características, posición, tamaño, metros cuadrados, centralidad de la colonia, etc., el valor de un auto con determinadas características, las señales fisiológicas de determinada especie evaluadas, entre muchos otros casos [30].

2.5.1.1.1 Clasificación

Aquí el objetivo es encontrar una relación entre las entradas x y las salidas y donde $y \in \{1, \dots, C\}$, con C siendo el número de clases. Si $C=2$, este problema es llamado clasificación binaria (en este caso nosotros asumimos que $y \in \{0,1\}$); si $C>2$, este problema es llamado clasificación multiclase. Si las etiquetas de clase no son mutuamente exclusivas (por ejemplo, tal vez sea clasificado como alto y fuerte), podemos llamarlo clasificación multietiqueta, cuando utilicemos el termino clasificación nos referiremos a una clasificación multiclase con una sola salida a menos que se indique otro tipo de clasificación.

Una forma de formalizar el problema es como una función de aproximación. Se asume $y = f(x)$ para alguna función f desconocida y el objetivo del aprendizaje es estimar la función f dado un set de entrenamiento y luego hacer predicciones usando $\hat{y} = \hat{f}(x)$ una función estimada. El objetivo es hacer predicciones con nuevas entradas que jamás hemos visto (esto es la generalización) [30].

2.5.1.1.1.2 La necesidad de un enfoque probabilístico

En ciertos casos encontramos que una muestra a clasificar es ambigua, podría pertenecer a dos clases evaluadas, en este punto es deseable poseer una probabilidad para definir a que clase pertenece.

Se denotará la distribución de probabilidad sobre todas las posibles etiquetas, dado un vector de entrada x y un conjunto de entrenamiento D por $p(y = 1|x, D)$, ya que $p(y = 1|x, D) + p(y = 0|x, D) = 1$. En nuestra notación hacemos explicito que la probabilidad es condicional sobre el vector x de prueba tan bien como el set de entrenamiento D .

Dada una salida probabilística, nosotros podemos computar siempre la mejor opción como la etiqueta correcta usando:

$$\hat{y} = \hat{f}(x) = \operatorname{argmax}_{c=1}^C p(y = c|x, D) \quad (20)$$

Esto corresponde a la clase más probable y es llamado la moda de una distribución $p(y|x, D)$; esta también es conocida como el máximo a posteriori.

2.5.1.1.2 Regresión

La regresión es igual a la clasificación solo que la variable respuesta es continua. Considerando una entrada como un valor real $x_i \in R$ y un único valor real como respuesta $y_i \in R$.

2.5.1.2 Aprendizaje no supervisado

En este enfoque solo poseemos entradas y el objetivo es meramente encontrar patrones interesantes en los datos. A esto algunas veces se le llama descubriendo la información estos problemas están mucho menos definidos ya que nadie menciona que tipo de patrones debes encontrar. A diferencia del aprendizaje supervisado aquí no hay nada que te diga cuál es la salida deseada para cada entrada. En lugar de eso formalizamos la tarea como una estimación de densidad, con lo cual queremos construir modelos de la forma $p(x_i|\theta)$. Hay dos diferencias con el aprendizaje supervisado. Primero, escribimos $p(x_i|\theta)$ en lugar de $p(y_i|x_i, D)$; que es, el aprendizaje supervisado es la estimación condicional de la densidad en lugar de estimación incondicional de la densidad. Segundo x_i es un vector de características entonces necesitamos crear modelos de probabilidad multivariada. En contraste en el aprendizaje supervisado, y_i es usualmente una sola variable que nosotros intentamos predecir. Esto quiere decir que para la mayoría de los problemas supervisados nosotros podemos usar modelos de probabilidad univariada lo cual significativamente simplifica el problema [30].

2.5.2 Algunos conceptos básicos en aprendizaje automático

2.5.2.1 Modelos paramétricos contra no paramétricos

Los modelos probabilísticos evaluados serán de la forma $p(y|x)$. Hay muchas formas para definir tales modelos, pero la más importante distinción es esta: ¿los modelos tienen un número fijo de parámetros o el número de parámetros crece con la cantidad de datos de entrenamiento? El primero se llama un modelo paramétrico y el segundo es llamado un modelo no paramétrico. Los modelos paramétricos tienen la ventaja en algunas ocasiones de ser más rápidos al usar, pero la desventaja de hacer fuertes suposiciones acerca de la naturaleza de las distribuciones de los datos. Los modelos no paramétricos son más flexibles, pero regularmente computacionalmente intratables por los grandes conjuntos de datos [30].

2.5.2.3 La maldición de la dimensionalidad

El principal problema del clasificador KNN es que este no trabaja bien con entradas dimensionalmente grandes. La pobre eficiencia con altas dimensiones es debido a la maldición de la dimensionalidad.

Para explicar la maldición se darán algunos ejemplos, si consideramos aplicar un clasificador KNN a datos donde las entradas están uniformemente distribuidas en un cubo con una D-dimensión. Supongamos que estimamos la densidad de las etiquetas de clase alrededor del punto de prueba x creciendo un hiper cubo alrededor de x hasta que este contenga una fracción deseada f de los puntos de datos. La longitud esperada del cubo borde del cubo esperado será $e_D(f) = f^{\frac{1}{D}}$. Si $D=10$, y queremos basar nuestra estimación en el 10% de los datos, tenemos $e_{10}(0.1) = 0.8$, por lo que necesitamos extender el cubo 80 % a lo largo de cada dimensión de x . incluso si usamos el 1% de los datos, encontramos $e_{10}(0.01) = 0.63$. a que el rango entero de los datos es únicamente 1 a lo largo de cada dimensión, vemos que el método ya no es muy local, a pesar del nombre vecinos cercanos [30].

2.5.2.4 Modelos paramétricos para clasificación y regresión

La principal ruta para combatir la maldición de la dimensionalidad es hacer algunas aseveraciones acerca de la naturaleza de la distribución de los datos. Estas aseveraciones conocidas como sesgo inductivo son algunas veces embebidos en la forma de modelo paramétrico el cual es un modelo estadístico con un número fijo de parámetros [30].

2.5.2.4 Sobre ajuste

Cuando sobre ajustamos un modelo a un set de datos en específico, estamos cayendo un error común en la minería de datos, el sobre ajuste de los modelos cosa que deberíamos evitar a toda costa. Esto es, deberíamos evitar intentar modelar cada pequeña variación en la entrada, ya que es más probable que sea el ruido de la señal verdadera.

2.5.2.5 Métodos de visualización

Muchas de las veces desconocemos cómo se comportan los datos, ¿poseen un límite separable?, ¿poseen un patrón repetible?, ¿los datos se separan en grupos diferenciables en el espacio?, ¿su distribución en el espacio es indiferenciable? y para saber con qué técnica es preferible tratarlos

es necesario visualizarlos, a lo largo del tiempo se ha trabajado con diferentes técnicas de visualización, más adelante se revisará una de las técnicas de visualización para datos con etiqueta de clase.

2.6 Probabilidad

Como hemos revisado la probabilidad juega un papel importante dentro del aprendizaje automático, aunque las teorías de probabilidad pueden llegar a ser muy extensas se describirá solo lo necesario para entender los algoritmos implementados en esta tesis.

¿Para dar inicio a este tema podemos plantearnos la pregunta que es probabilidad? Estamos todos familiarizados con la frase la probabilidad de una moneda caerá Águila es de 0.5. pero qué significa esto? Actualmente hay dos definiciones de probabilidad una de ellas es la llamada interpretación frecuentista. Viéndolo así, la probabilidad representa la frecuencia de un evento por ejemplo el enunciado anterior diría si nosotros lanzamos una moneda al aire muchas veces esperaremos caiga águila la mitad de las veces. La otra interpretación es la llamada interpretación bayesiana de probabilidad. En esta visión la probabilidad es usada para cuantificar nuestra incertidumbre acerca de algo, por lo tanto, está relacionado con información en lugar de ensayos repetidos [30].

En la opinión bayesiana, el enunciado significa que creemos que la moneda es igualmente probable que caiga águila o sello en el lanzamiento siguiente [30].

Una gran ventaja de la interpretación bayesiana es que esta puede ser usada para modelar nuestra incertidumbre acerca de los eventos que tienen frecuencias a largo plazo. Por ejemplo, es posible que desee calcular la probabilidad de que la capa de hielo polar se derrita en el 2020. Este evento ocurrirá cero o una vez, pero no puede suceder repetidamente. Sin embargo, debemos ser capaces de cuantificar nuestra incertidumbre acerca de este evento [30].

2.6.1 Teoría de probabilidad: Variables aleatorias discretas

La expresión $p(A)$ denota la probabilidad que el evento A suceda. Por ejemplo, A indicaría la expresión lógica mañana lloverá. requerimos que $0 \leq p(A) \leq 1$, donde $p(A) = 0$ indicaría que definitivamente no sucederá, y $p(a) = 1$ indicaría que si sucederá. Escribimos $p(\bar{A})$ describimos la probabilidad del evento no A .

Podemos extender la noción de eventos binarios por medio de la definición de variables aleatorias discretas X , la cual puede tomar cualquier valor de un conjunto contable finito o infinito $\mathcal{X}$. Denotamos la probabilidad del evento $X = x$ por $p(X = x)$ o solo probabilidad de x $p(x)$. Aquí $p()$ es llamada la función de probabilidad de masa. Esto satisface las probabilidades $0 \leq p(A) \leq 1$ y $\sum_{x \in \mathcal{X}} p(x) = 1$ [30].

2.6.2 Reglas fundamentales

2.6.2.1 Probabilidad de unión de dos eventos

Dados dos eventos, A y B , definimos la probabilidad de A o B como sigue:

$$\begin{aligned} p(A \vee B) &= p(A) + p(B) - p(A \wedge B) \\ &= p(A) + p(B) \text{ si } A \text{ y } B \text{ son mutuamente exclusivos} \end{aligned} \quad (21)$$

2.6.2.2 Probabilidades conjuntas

Definimos la probabilidad de la unión del evento A y B como sigue:

$$p(A, B) = p(A \wedge B) = p(A|B)p(B) \quad (22)$$

Esto es algunas veces llamada regla producto. Dada una distribución conjunta sobre dos eventos $p(A, B)$, definimos una distribución marginal como sigue:

$$p(A) = \sum_b p(A, B) = \sum_b p(A|B = b)p(B = b) \quad (23)$$

Donde se suman todos los posibles estados de B . aplica igual para a $p(B)$,. Esta es algunas veces llamada regla suma o la regla de probabilidad total.

2.6.2.3 Probabilidad condicional

Definimos la probabilidad condicional de un evento A , dado que el evento B es verdad como sigue:

$$p(A|B) = \frac{p(A, B)}{p(B)} \text{ si } p(B) > 0 \quad (24)$$

2.6.2.4 Regla de Bayes

Combinando la definición de probabilidad condicional, la regla de producto y suma nace la regla de Bayes, también llamada teorema de Bayes:

$$\begin{aligned} p(X = x|Y = y) &= \frac{p(X = x, Y = y)}{p(Y = y)} \\ &= (p(X = x)p(Y = y|X = x))/(\sum_{x'} p(X = x')p(Y = y|X = x')) \end{aligned} \quad (25)$$

Variables aleatorias continuas

Ahora se desarrollará el análisis para la incertidumbre de cantidades continuas.

Supongamos que X es alguna incertidumbre de cantidad continua. La probabilidad que X se encuentre en cualquier intervalo $a \leq X \leq b$ puede ser computado como sigue:

Definimos los eventos $A = (X \leq a)$, $B = (X \leq b)$ y $W = (a < X \leq b)$. Tenemos que $B = A \vee W$, y ya que A y W son mutuamente exclusivos, la regla de sumas da

$$p(B) = p(A) + p(W) \quad (26)$$

Y a su vez

$$p(W) = p(B) - p(A) \quad (27)$$

Definir la función $F(q) \triangleq p(X \leq q)$. Esta es la llamada función de distribución acumulativa de X , esta es obviamente una función de incremento monótonico. Usando la notación tenemos

$$p(a < X \leq b) = F(b) - F(a) \quad (28)$$

Ahora define $f(x) = \frac{d}{dx}F(x)$ asumiendo que la derivada existe; esta es llamada la función de densidad de probabilidad (pdf). Dada una pdf, podemos computar la probabilidad de una variable continua, siendo en un intervalo finito como sigue:

$$P(x < X \leq b) = \int_a^b f(x)dx \quad (29)$$

A medida que el tamaño del intervalo se reduce, podemos escribir

$$P(x \leq X \leq x + dx) \approx p(x)dx$$

Requerimos que $p(x) \geq 0$, pero es posible que $p(x) > 1$ para cualquier x dado, siempre que la densidad se integre a 1.

2.6.2.5 Media y varianza

La más conocida propiedad de una distribución es su media o valor esperado, denotado por μ . La varianza es una medición del esparcimiento de una distribución denotado por σ^2 esta es definida como sigue:

$$Var[X] \triangleq E[(X - \mu)^2] = \int (x - \mu)^2 p(x)dx \quad (30)$$

$$\int x^2 p(x)dx + \mu^2 \int p(x)dx - 2\mu \int xp(x)dx = E[X^2] - \mu^2 \quad (31)$$

De la cual nosotros derivamos el resultado útil

$$E[X^2] = \mu^2 + \sigma^2 \quad (32)$$

La desviación estándar se define como

$$std[X] \triangleq \sqrt{var[X]} \quad (33)$$

2.6.3 Distribuciones discretas

2.6.3.1 Distribución binomial

Supongamos que lanzamos una moneda “n” veces. Dejamos que $X \in \{0, \dots, n\}$ ser el numero de águilas. Si la probabilidad de las águilas es θ , entonces decimos que X tiene una distribución binomial, escribimos como $X \sim Bin(n, \theta)$. La pmf está dada por

$$Bin(k|n, \theta) \triangleq \binom{n}{k} \theta^k (1 - \theta)^{n-k} \quad (34)$$

Donde

$$\binom{n}{k} \triangleq \frac{n!}{((n-k)! k!)} \quad (35)$$

Es el número de veces para seleccionar k ítems de n (esto es conocido como el coeficiente binomial).

Esta distribución tiene la siguiente media y varianza:

Media= θ , var= $n\theta(1 - \theta)$

2.6.3.2 Distribución multinomial

La distribución binomial puede usarse para modelar lanzamientos de monedas. Para modelar los resultados de lanzar un dado, Podemos utilizar la distribución multinomial. Esto se define de la siguiente manera: sea $x = (x_1, \dots, x_K)$ un vector variable, donde x_j es el número de veces que el lado j del dado tiene lugar. Entonces x tiene la siguiente pmf:

$$Mu(x|n, \theta) \triangleq \binom{n}{x_1 \dots x_K} \prod_{j=1}^K \theta_j^{x_j} \quad (36)$$

Donde θ_j es la probabilidad que el lado j aparezca, y

$$\binom{n}{x_1 \dots x_K} \triangleq \frac{n!}{x_1! x_2! \dots x_K!} \quad (37)$$

Es el coeficiente multinomial (el número de formas de dividir el conjunto de tamaño $n = \sum_{k=1}^K x_k$ medidas en subconjuntos con medidas x_1 a x_K).

2.6.4 Distribuciones continuas Distribución normal (gaussiana)

La más utilizada distribución en estadística y aprendizaje automático es la distribución normal o gaussiana. Esta pdf está dada por

$$\mathfrak{N}(x|\mu, \sigma^2) \triangleq \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad (38)$$

Aquí $\mu = E[X]$ es la media, y $\sigma^2 = var[X]$ es la varianza. $\sqrt{2\pi\sigma^2}$ es la constante de normalización necesaria para asegurarse de integrar la densidad a uno. Escribimos

$X \sim \mathcal{N}(\mu, \sigma^2)$ para denotar que $p(X = x) = \mathcal{N}(x|\mu, \sigma^2)$. Si $X \sim \mathcal{N}(0,1)$, decimos que X sigue una distribución estándar normal

La distribución Gaussiana es la más libremente usada distribución en estadística. Hay varias razones para esto. Primero, esta tiene dos parámetros los cuales son fácilmente interpretado, y los cuales capturan algunas de las más básicas propiedades de una distribución, media y varianza. Segundo el teorema del límite central dice que la suma de variables aleatorias independientes tiene una distribución aproximadamente normal, haciéndola una buena elección para modelar errores residuales o ruido. Tercero la distribución gaussiana hace el menor número de suposiciones (tiene máxima entropía) [30].

2.6.5 Distribuciones de probabilidades conjuntas

En esta sección iniciaremos nuestra discusión de problemas más difíciles de construir, distribuciones de probabilidad conjuntas en múltiples variables aleatorias relacionadas.

Una distribución de probabilidad conjunta tiene la forma $p(x_1, \dots, x_D)$ para un conjunto de variables $D > 1$ y modela las relaciones (estocásticas) entre las variables. si todas las variables son discretas, nosotros podemos representar la distribución conjunta como un gran arreglo multidimensional, con una variable por dimensión. sin embargo, el número de parámetros necesarios para definir tal modelo es $O(K^D)$, donde K es el número de estados de cada variable.

2.6.5.1 Covarianza y correlación

La covarianza entre dos variables aleatorias X y Y mide el grado en el que X e Y están relacionados (linealmente). La covarianza está definida como

$$\text{cov}[X, Y] \triangleq E[(X - E[X])(Y - E[Y])] = E(XY) - E(X)E(Y) \quad (39)$$

Si x es un vector aleatorio d-dimensional, esta matriz de covarianzas está definida a tener la siguiente simetría matriz positiva definitiva:

$$\text{cov}[X, Y] \triangleq E[(x - E[x])(x - E[x])^T] \quad (40)$$

$$= \begin{pmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) & \text{cov}(X_1, X_d) \\ \text{cov}(X_2, X_1) & \text{var}(X_2) & \text{cov}(X_2, X_d) \\ \text{cov}(X_d, X_1) & \text{cov}(X_d, X_2) & \text{var}(X_d) \end{pmatrix}$$

Las covarianzas pueden estar entre 0 e infinito. Algunas veces es más inconveniente para trabajar mediciones normalizadas, con un límite superior finito.

El coeficiente de correlación de Pearson entre X y Y está definido como

$$corr[X, Y] = \frac{cov[X, Y]}{\sqrt{var[X]var[Y]}} \quad (41)$$

Y la matriz de correlación tiene la forma

$$R = \begin{pmatrix} corr[X_1, X_1] & \cdots & corr[X_1, X_d] \\ \vdots & \ddots & \vdots \\ corr[X_d, X_1] & \cdots & corr[X_d, X_d] \end{pmatrix}$$

Uno puede notar que $-1 \leq corr[X, Y] \leq 1$. Por lo tanto, en una matriz de correlación, cada entrada en la diagonal es 1, y las otras entradas están entre -1 y 1. La mejor forma para pensar en el coeficiente de correlación es como un grado de linealidad.

2.6.6 La multivariada normal o multivariada gaussiana

La pdf es la más libremente usada función de probabilidad de densidad conjunta para variables continuas la pdf de la multivariada normal en D dimensiones está definida como sigue

$$\mathfrak{N}(x|\mu, \Sigma) \triangleq \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right] \quad (42)$$

Donde $\mu = E[x] \in R^D$ es la media del vector, $\Sigma = cov[x]$ es la $D \times D$ matriz covarianza [30].

2.6.5.2 Distribución t student multivariada

Una alternativa más robusta que la MVN es la distribución t de student multivariada, la cual tiene una pdf como sigue

$$\tau(x|\mu, v) = \frac{\Gamma\left(\frac{v}{2} + \frac{D}{2}\right) |\Sigma|^{-\frac{1}{2}}}{\Gamma\left(\frac{v}{2}\right) v^{\frac{D}{2}} \pi^{\frac{D}{2}}} \exp \left[-\frac{1}{v} (x - \mu)^T \Sigma^{-1} (x - \mu) \right]^{-\frac{(v+D)}{2}}$$

$$= \frac{\Gamma\left(\frac{v}{2} + \frac{D}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} |\pi \mathbf{V}|^{-\frac{1}{2}} x [1 + (x - \mu)^T \mathbf{V}^{-1} (x - \mu)]^{-\frac{(v+D)}{2}} \quad (43)$$

Donde Σ es llamada la matriz escalar (no es exactamente la matriz covarianza) y $\mathbf{V} = v\Sigma$. Esta tiene colas más gordas que una gaussiana. Cuanto menor sea v , más gordas serán las colas, si v tiende al infinito la forma tendera a una gaussiana.

$$\text{Media} = \mu \text{ mediana} = \mu, \text{ Cov} = \frac{v}{v-2} \Sigma$$

2.6.6 Transformaciones de variables aleatorias

Si $x \sim p()$ es alguna variable aleatoria, y $y = f(x)$, ¿cuál es la distribución de y ? esta pregunta da paso a esta sección.

Transformaciones lineales

Supongamos que $f()$ es una función lineal $y = f(x) = Ax + b$. En este caso, podemos fácilmente derivar la media y la covarianza de y como sigue. Primero, para la media, tenemos $E(y) = E[Ax + b] = A\mu + b$. Donde $\mu = E(x)$ la linealidad de la expectativa. Si $f()$ es una función de valor escalar, $f(x) = a^T x + b$, el correspondiente resultado es

$$E[a^T x + b] = a^T \mu + b \quad (44)$$

Para la covarianza, tenemos

$$\text{Cov}[y] = \text{cov}[Ax + b] = A \Sigma A^T \quad (45)$$

Donde $\Sigma = \text{cov}[x]$. Si $f()$ es un valor escalar, el resultado viene siendo

$$\text{var}[y] = \text{var}[a^T x + b] = a^T \Sigma a \quad (46)$$

Nota que la media y la covarianza únicamente definen completamente la distribución de y si x es gaussiana. En general deberíamos usar las técnicas descritas a continuación para derivar la distribución completa de y , en contraposición a sus dos primeros momentos [30].

2.6.6.1 Transformaciones generales

Si X es una variable aleatoria discreta, nosotros podemos derivar la pmf para y simplificando la probabilidad para todos las x 's tal como $f(x) = y$:

$$p_y(y) = \sum_{x:f(x)=y} p_x(x) \quad (47)$$

Por ejemplo, si $f(X) = 1$ si X es par y $f(x) = 0$, y $p_x(x)$ es uniforme sobre los sets $\{1, \dots, 10\}$, luego $p_y(1) = \sum_{x \in \{2,4,6,8,10\}} p_x(x) = 0.5$, y $p_y(0) = 0.5$ similarmente. Nota que en este ejemplo f es una función muchos a uno.

Si X es continua, no podemos usar la ecuación anterior ya que $p_x(x)$ es una densidad, no una pmf, y no podemos sumar las densidades. En lugar de esto, trabajamos con cdf's y escribimos

$$P_y \triangleq P(Y \leq y) = P(f(X) \leq y) = P(X \in \{x|f(x) \leq y\}) \quad (48)$$

Podemos derivar la pdf de y diferenciando la cdf.

En el caso de funciones invertibles y monotonicas, podemos escribir

$$P_y(y) = P(f(X) \leq y) = P(X \leq f^{-1}(y)) = P_x(f^{-1}(y)) \quad (49)$$

Tomando las derivativas obtenemos

$$p_y(y) \triangleq \frac{d}{dy} P_y(y) = \frac{d}{dy} P_x(f^{-1}(y)) = \frac{dx}{dy} \frac{d}{dx} P_x(x) = \frac{dx}{dy} p_x(x) \quad (50)$$

Donde $x = f^{-1}(y)$. Podemos pensar en dx como una medida de volumen en el espacio- x ; similarmente la medida de volumen para dy en el espacio y . así la medida dx/dy del cambio en el volumen. Ya que el signo de este cambio no es importante, nosotros tomamos el valor absoluto para obtener la expresión general:

$$p_y(y) = p_x(x) \left| \frac{dx}{dy} \right| \quad (51)$$

Esto es llamado formula de cambio de variable. Podemos entender este resultado más intuitivamente como sigue. Las observaciones que caen dentro del rango $(y, y+dy)$ se transformarán en $(x, x+dx)$, donde $p_x(x)dx \approx p_y(y)dy$. Por lo tanto $p_y(y) \approx p_x(x) \left| \frac{dx}{dy} \right|$. Por ejemplo, suponiendo

$X \sim U(-1,1)$, y $Y = X^2$. Luego $p_y(y) = \frac{1}{2} y^{-\frac{1}{2}}$.

2.6.7 Teorema del límite central

Ahora consideremos N variables aleatorias con pdf's (no necesariamente gaussianas) $p(x)$, cada una con media y varianza. Asumimos que cada variable es independiente e idénticamente distribuida. Sea $S_N = \sum_{i=1}^N X_i$ ser la suma de las variables aleatorias. Esta transformación de variables aleatorias es simple, pero es libremente usada. Se puede demostrar, que a medida que N aumenta, la distribución de esta suma se aproxima

$$p(S_N = s) = \frac{1}{\sqrt{2\pi N\sigma^2}} \exp\left(-\frac{(s - N\mu)^2}{2N\sigma^2}\right) \quad (52)$$

Ya que la distribución de la cantidad

$$Z_N \triangleq \frac{S_N - N\mu}{\sigma\sqrt{N}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} \quad (53)$$

Converge a la normal estándar, donde $\bar{X}$ es la media de la muestra. Este es el llamado teorema del límite central.

2.6.8 Modelos generativos

En esta sección se discutirá como clasificar un vector características x aplicado la regla de Bayes a un clasificador generativo de la forma

$$p(y = c|x, \theta) \propto p(X|y = c, \theta)p(y = c|\theta) \quad (54)$$

La clave para usar tales modelos es especificar una forma adecuada para la densidad de clase condicional $p(y = c|x, \theta)$, la cual define que tipo de datos esperamos ver en cada clase. Además de evaluar datos discretos se discutirá como inferir parámetros desconocidos θ , de tales modelos [30].

2.6.8.1 Concepto bayesiano

Consideremos ahora un niño como aprende el significado de una palabra, tal como perro. presumiblemente los padres del niño le den ejemplos positivos diciendo: mira qué lindo perro o simplemente mira el perrito, etc. sin embargo es poco probable que proporcione ejemplos negativos. Es muy improbable que proporcionen ejemplos negativos, diciendo “miren a ese no

perro”. Ciertamente ejemplos negativos deberían obtenerse con un proceso activo de aprendizaje. Ejemplos como el niño dice mira un perro y los padres dicen este es un gato querido no un perro; pero la investigación científica ha mostrado que las personas pueden aprender conceptos de solo ejemplos positivos.

Podemos pensar que, al aprender el significado de una palabra como equivalente al concepto de aprendizaje, que a su vez es equivalente a la clasificación binaria. Para ver esto, definimos $f(x) = 1$ si x es un ejemplo del concepto C , y $f(x) = 0$ en otro caso. Luego el objetivo es aprender la función indicador f , la cual solo define cuales elementos están en el set C . Al permitir la incertidumbre sobre la definición de C , podemos emular la teoría de conjuntos difusos, pero utilizando el cálculo de probabilidad estándar [30].

2.6.8.2 Probabilidad vs Verosimilitud (*likelihood*)

Como se ha visto la probabilidad se utiliza antes de que los datos estén disponibles para describir posibles resultados. El *likelihood* se utiliza después de que los datos estén disponibles para describir una función de un parámetro o vector de parámetros para un resultado dado.

2.6.8.3 Clasificador Naive Bayes

Aquí se discutirá como clasificar los vectores de característica de valores discretos, $x \in \{1, \dots, K\}^D$, donde K es el número de valores de cada característica, y D es el número de características. Se usará un enfoque generativo. Esto nos permite escribir la densidad condicional de la clase como un producto de densidades unidimensionales:

$$p(x|y = c, \theta) = \prod_{j=1}^D p(x_j|y = c, \theta_{jc}) \quad (55)$$

El modelo resultante es llamado un clasificador Naive Bayes (NBC). El modelo se llama ‘Bayes’ ingenuo ya que no espera que las características sean independientes, ni siquiera condicionales en la etiqueta de clase. Sin embargo, incluso si la hipótesis ingenua no es cierta, a menudo resulta en clasificadores que funcionan bien [30]. Una de estas razones es que el modelo es demasiado simple (solo tiene parámetros $O(CD)$, para C clases y D características) y por lo tanto es relativamente inmune a sobreajuste.

2.6.8.3.1 Ajuste del modelo

Ahora se discutirá como entrenar un clasificador Naive Bayes.

Estimación del máximo likelihood (MLE) para NBC

La probabilidad de un solo caso de datos está dada por

$$p(x_i, y_i | \theta) = p(y_i | \pi) \prod_j p(x_{ij} | \theta_j) = \prod_c \pi_c^{I(y_i=c)} = \prod_j \prod_c p(x_{ij} | \theta_{jc})^{I(y_i=c)} \quad (56)$$

Ya que el log-likelihood está dado por

$$\log p(D | \theta) = \sum_{c=1}^C N_c \log \pi_c + \sum_{j=1}^D \sum_{c=1}^C \sum_{i: y_i=c} \log p(x_{ij} | \theta_{jc}) \quad (57)$$

Vemos que esta expresión se descompone en una serie de términos, uno relativo a π , y términos DC que contienen los θ_{jc} 's. por lo tanto, podemos optimizar todos estos parámetros por separado.

De la ecuación 3.48, el MLE para la case prior esta dado por

$$\hat{\pi}_c = \frac{N_c}{N} \quad (58)$$

Donde $N_c \triangleq \sum_i I(y_i = c)$ es el número de ejemplos en la clase c.

2.6.9 Modelos gaussianos: Análisis discriminante gaussiano

Una aplicación importante de MVNs es definir las densidades condicionales de la clase en un clasificador generativo, es decir,

$$p(x | y = c, \theta) = \mathfrak{N}(x | \mu_c, \Sigma_c) \quad (59)$$

La técnica resultante es llamada Análisis discriminante gaussiano o GDA. Si Σ_c es diagonal, este es equivalente a Naive Bayes.

Podemos clasificar un vector de características usando la siguiente regla de decisión derivado de la ecuación 2.13

$$\hat{y}(x) = \operatorname{argmax}_c [\log(y = c | \pi) + \log p(x | \theta_c)] \quad (60)$$

Cuando computamos la probabilidad de x bajo cada clase de densidad condicional, estamos midiendo la distancia de x al centro de cada clase, μ_c , usando la distancia de mahalanobis. Esto puede ser pensado como un clasificador de centroides más cercano.

Como un ejemplo, Figura 4.2 muestra dos densidades de clase condicional en 2d, representando el peso y la altura de hombres y mujeres. Podemos ver que las características están correlacionadas como se espera las personas más altas tienden a pesar más. Las elipses para cada clase contienen el 95% de la probabilidad si tenemos una probabilidad a priori uniforme sobre las clases, podemos clasificar un nuevo test como sigue

$$\hat{y}(x) = \underset{c}{\operatorname{argmin}} (x - \mu_c)^T \sum_c^{-1} (x - \mu_c) \quad (61)$$

2.6.9.1 Análisis discriminante cuadrático (QDA)

El posterior sobre las etiquetas de la clase esta dado por la ecuación 2.13. podemos obtener una mayor comprensión de este modelo mediante la conexión de la definición de la densidad gaussiana, como sigue:

$$p(y = c|x, \theta) = \frac{\pi_c |2\pi \Sigma_c|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (x - \mu_c)^T \Sigma_c^{-1} (x - \mu_c) \right]}{(\sum_{c'} \pi_{c'} |2\pi \Sigma_{c'}|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (x - \mu_{c'})^T \Sigma_{c'}^{-1} (x - \mu_{c'}) \right])} \quad (62)$$

Reduciendo esto resulta en una función cuadrática de x. el resultado se conoce como análisis discriminante cuadrático (QDA).

2.6.9.2 Análisis discriminante lineal (LDA)

Ahora consideramos un caso especial en el cual las matrices covarianza están unidas o compartidas entre las clases $\Sigma_c = \Sigma$. En este caso podemos simplificar la ecuación 4.33 como sigue:

$$\begin{aligned} p(y = c|x, \theta) &\propto \pi_c \exp[\mu_c^T \Sigma^{-1} x - \frac{1}{2} x^T \Sigma^{-1} x - \frac{1}{2} \mu_c^T \Sigma^{-1} \mu_c] \\ &= \exp[\mu_c^T \Sigma^{-1} x - \frac{1}{2} \mu_c^T \Sigma^{-1} \mu_c + \log \pi_c] \exp[-\frac{1}{2} x^T \Sigma^{-1} x] \end{aligned} \quad (63)$$

Ya que el termino cuadrático $x^T \Sigma^{-1} x$ es independiente de c , cancelaremos el numerador y en el denominador. Si definimos

$$\gamma_c = -\frac{1}{2} \mu_c^T \Sigma^{-1} \mu_c + \log \pi_c$$

$$\beta_c = \Sigma^{-1} \mu_c$$

Luego podemos escribir

2.6.10 Transformaciones a un nuevo espacio para visualizar

Muchas veces el visualizar los datos es imposible ya que son decenas, cientos hasta miles de variables a analizar, sin embargo, también existen métodos para reducir estas variables y expresarlas en espacios más pequeños, como seres humanos vivimos en un espacio en tres dimensiones y no podemos expresar más allá de estas tres, al expresar decenas, centenas de variables en solo dos o tres dimensiones sin duda existe una deformación de la información pero es meramente un intento para ver cómo es que se comportan estos datos, hoy en día son tantas las características medibles de un sistema que seleccionar un número adecuado que exprese lo más cercano posible la información que requerimos es un problema, muchas de esas veces sabemos que todas las características a evaluar son importantes y sin embargo ninguna nos permite obtener la información que necesitamos. Para aumentar la diferenciación de los datos se utiliza comúnmente transformar a nuevos espacios, crear nuevas características a evaluar para obtener la información tal como la necesitamos.

En problemas de clasificación la tarea es asignar a nuevas entradas una de un numero de clases discretas o categóricas. Sin embargo, hay muchas otras tareas de patrones de reconocimiento a las cuales llamaremos problemas de regresión en los cuales las salidas representan los valores de variables continuas. En efecto, como se vio el termino regresión refiere a una específica forma de función definida en términos de promedio sobre una cantidad variable. Ambos problemas de regresión y clasificación pueden ser vistos como casi particulares funciones de aproximación. En el caso de problemas de regresión se establece una función de regresión con la cual nosotros deseamos aproximar mientras para problemas de clasificación las funciones las cuales nosotros deseamos aproximar son las probabilidades de los miembros de diferentes clases expresadas

como funciones de las variables de entrada. Muchos de los puntos clave los cuales necesitamos dirigirnos a abordar en problemas de patrones de reconocimiento son comunes a clasificación como a regresión [20].

2.6.11 Normalización

Normalizar un conjunto de datos se refiere a ajustar este conjunto a una distribución normal estándar y una distribución normal estándar por definición indica que, si x es una observación de una distribución con media μ y desviación estándar σ , el valor estándar de x lo es $z = \frac{x - \mu}{\sigma}$ a este valor estándar también se le conoce como valor z .

El valor z nos indica cuantas desviaciones estándar esta la observación original de la media y en qué dirección. Las observaciones mayores que su media toman valores positivos cuando se estandarizan mientras que los valores que son menores que su media toman valores negativos.

2.6.12 Selección secuencial de atributos

La selección de atributos se refiere a seleccionar un subconjunto de características existentes sin una transformación. Existen diferentes estrategias de búsqueda para definir que atributos son los que perdurarán, por ejemplo:

- Selección secuencial hacia adelante
- Selección secuencial hacia atrás
- Selección plus-r minus-r
- Búsqueda bidireccional
- Búsqueda flotante

Por definición dado un conjunto de características $X = \{x_i | i=1 \dots N\}$, encontrar un conjunto Y_M , con $M < N$, que maximice una función objetivo $J(Y)$, idealmente $P(\text{correcto})$

$$Y_M = \{x_1, x_2, x_3, \dots, x_{iM}\} = \operatorname{argmax} J\{x_i | i = 1 \dots N\}$$

El seleccionar un subconjunto de características requiere una estrategia de búsqueda para seleccionar el mejor candidato del conjunto total, también requiere una función objetivo para evaluar estos candidatos.

En muchos de los casos una búsqueda exhaustiva (esto quiere decir que evaluar todas las combinaciones posibles para n número de características) es imposible, por ejemplo, el número

de subconjuntos de características aumenta exponencialmente, por lo tanto, es imposible evaluar cada uno de estos subconjuntos. La función objetivo evalúa la bondad de cada subconjunto para compararlos con los otros subconjuntos y elegir el mejor.

Método envolvente

Usa un clasificador para evaluar los subconjuntos por su rendimiento predictivo sobre datos de prueba a través de validación cruzada.

Algoritmos secuenciales

-Selección secuencial hacia adelante

El siguiente pseudocódigo muestra cómo trabaja este algoritmo:

Seudocódigo

SUBCONJUNTO= 0; MEJOR_EVAL = 0;

ESPACIO=1...L

SALIDA= falso

WHILE NOT SALIDA

FOR i=1 **A** |ESPACIO|

TEMP(i)= J (SUBCONJUNTO $\cup$ ESPACIO(i))

END

MEJOR_i = ARGMAX(TEMP)

IF TEMP(MEJOR_i)>MEJOR_EVAL

subconjunto = SUBCONJUNTO +

ESPACIO(MEJOR_i)

MEJOR_EVAL= TEMP(MEJOR_i)

ESPACIO = ESPACIO – ESPACIO(MEJOR_i)

ELSE

SALIDA = verdad;

END

END

RETURN SUBCONJUNTO, MEJOR_EVAL

2.6.13 Trazo de las matrices dispersas y su proyección en el espacio de medias

El trazo de las matrices dispersas el cual mide la separación entre poblaciones agrupadas se muestra estrictamente preservado por el esfereo de los datos, seguida por una proyección en el espacio de las medias de la población. [31]. Las proyecciones lineales de baja dimensionalidad de altas dimensiones son cada vez más importantes para obtener información sobre dominios complejos. El método es genérico y está destinado a la reducción de la dimensionalidad y la visualización de las agrupaciones de la población, sin embargo, se identifican, ya sea dividiéndolos en grupos o con etiqueta de clase predeterminadas. Los marcos genéricos para la visualización de baja dimensión generalmente se dividen en tres categorías, los métodos puramente lineales utilizan frecuentemente valores singulares que abarcan la mayor varianza de los datos, por ejemplo, los biplanos ampliamente utilizados. Sin embargo, la estructura proyectiva lineal del mapa identifica ejes direccionales a lo largo de los cuales las covariables individuales cambian mono tónicamente. Si bien este enfoque es útil para verificar visualmente las correlaciones conocidas entre los ejes de atributos, es generalmente el caso que los primeros dos o tres valores singulares que explican una proporción relativamente pequeña de la varianza de los datos, con la consecuencia de que los grupos (clusters) compactos verdaderos son severamente mezclados cuando se proyectan puntos individuales con este método [31].

Un enfoque alternativo es relajar el concepto de linealidad y definir una proyección no lineal para optimizar la correspondencia entre las distancias vecinas más cercanas en el espacio de entrada original y entre las proyecciones bidimensionales de los puntos de datos individuales. Una metodología eficaz y bien conocida es la escala multidimensional (MDS) o cartografía de Sammon. Sin embargo, estos mapas son computacionalmente costosos de calcular y pueden ser muy sensibles al ruido en los datos alterando radicalmente las proyecciones de los datos, incluso cuando solo un pequeño número de puntos varían o son añadidos o eliminados del conjunto de los datos [31].

Un tercer enfoque relacionado para generar mapas topográficos mediante la proyección de datos en un tejido de superficie curva a través de los datos y cortando a través del ruido. El primer método bien conocido en esta categoría son los mapas auto-organizados (SOM), que fue modelado en las correlaciones de la célula central envolvente típico de las vías sensoriales tempranas en sistemas nervioso-biológicos. Desde entonces este enfoque se ha modelado dentro

de un marco estadístico bayesiano para el modelado de variables latentes en forma de mapas topográficos latentes (GTM) [31].

A medida que los métodos computacionales ganan popularidad para procesar el incremento de datos es de interés definir una visualización de baja dimensión que es rápida para calcular, preferiblemente lineal para mostrar la dirección de los ejes de coordenadas originales. Sino que también representa datos con compacidad dentro de los datos y con la mayor separación entre ellos.

La varianza total de los datos se dividirá en $S_T = S_W + S_B$

$$S_T = \sum_{i=1}^N ((x_i - m)^T (x_i - m)) \quad (64)$$

$$S_W = \sum_{j=1}^{N_c} \sum_{i=1}^{N_j} (x_i - m)^T (x_i - m) \quad (65)$$

$$S_B = \sum_{j=1}^{N_c} N_j (m_j - m)^T (m_j - m) \quad (66)$$

Donde los datos son particionados en N_c grupos (clústeres) cada uno con N_j puntos y medias m_j .

Las medidas escalares para la separación entre los grupos se obtienen fácilmente tomando los trazos de las matrices dispersas, definiendo la suma de cuadrados dentro y entre los grupos [31].

Método para visualización basado en agrupación usando matrices dispersas

Proyección de los datos sobre el espacio de las medias de los grupos

Dada la importancia de los medios de agrupación como prototipos representativos para los propios clústeres, es natural proyectar los datos en el subespacio generado por las medias de los clústeres. Esto se logra fácilmente definiendo un conjunto ortonormal de vectores de base b_i $i = 1 \dots N_c$ por ejemplo mediante la ortogonalización de Gram-Schmidt, generando la primera representación proyectiva compacta de este método [31].

2.6.13.1 Algoritmo ortogonalización de Gram-Schmidt

Sea V un espacio vectorial con un producto interno.

Supongamos $X_1, X_2, \dots, X_n$ es una base para V

$$V_1 = X_1;$$

$$V_2 = X_2 - \frac{[X_2, V_1]}{[V_1, V_1]} * V_1$$

$$V_3 = X_3 - \frac{[X_3, V_1]}{[V_1, V_1]} * V_1 - \frac{[X_3, V_2]}{[V_2, V_2]} * V_2$$

...

$$V_n = \frac{[X_n, V_1]}{[V_1, V_1]} * V_1 - \dots - \frac{[X_n, V_{n-1}]}{[V_{n-1}, V_{n-1}]} * V_{n-1} \quad (67)$$

Donde $V_1, V_2, V_3 \dots V_n$ son ortogonales.

Sin embargo, los límites de separación en la descomposición de Voronoi creada por el agrupamiento no serán generalmente preservados por este mapa proyectivo, causando la mezcla entre los cúmulos. Esto resultará en una reducción en el valor numérico del índice de separación de clúster.

Una forma de preservar tanto como sea posible la separación de clúster después de comprimir los datos proyectándose en el espacio ocupado por los medios de agrupamiento es linealizar la métrica de Mahalanobis en el espacio de datos original por blanqueamiento (*sphering*) de los datos. Esta transformación se aplica únicamente con el propósito de reducción de dimensionalidad y visualización, por lo tanto, no influye de ninguna manera en el proceso de agrupación.

2.6.13.2 Blanqueamiento de los datos

El blanqueamiento (*sphering*) de los datos es un común paso de preprocesamiento para transformar variables aleatorias a ortogonalidad. Una transformación de blanqueamiento es una

transformación lineal que transforma un vector de variables aleatorias con una matriz de covarianza conocida en un conjunto de nuevas variables cuya covarianza es la matriz de identidad, lo que significa que no están correlacionadas y todas tienen una varianza [32].

2.6.13.2.1 Decorrelación: transformación de datos para tener una matriz de covarianzas diagonal

Digamos que tenemos una matriz de datos X compuesta de K dimensiones y n observaciones (X tiene tamaño $[K \times n]$). Supongamos también que las filas de X han sido centradas (la media ha sido sustraída en todas las observaciones). La covarianza Σ de cada una de las dimensiones con respecto a la otra es

$$\Sigma = \text{cov}(x) = E[X X^T] \quad (68)$$

Donde la matriz de covarianza $E[X X^T]$ puede ser estimada de la matriz como sigue

$$E[X X^T] \approx \frac{X X^T}{n} \quad (69)$$

Donde la matriz de covarianzas puede ser definida como un producto de dos simples matrices E y D usando un procedimiento llamado descomposición de valores propios (eigenvalores)

$$\Sigma = E D E^{-1} \quad (70)$$

La matriz E es una matriz de $[k \times k]$ elementos donde cada columna es un vector propio (*eigenvector*) de Σ y D es una matriz diagonal en la cual los elementos de la diagonal D_{ii} son valores propios que corresponden a los vectores propios de la i -ésima columna despejando la ecuación 72 tenemos

$$E \Sigma E^{-1} = D \quad (71)$$

La intención es transformar la matriz original a una nueva matriz Y , $Y = W_D X$. Donde sus dimensiones no están correlacionadas, Y tiene una covarianza diagonal D , como queremos determinar la matriz de transformación W_D

$$D = cov(Y) = E[YY^T] \quad (72)$$

Con las ecuaciones 73 y 74 obtenemos la siguiente

$$D = \frac{W_D X (W_D X)^T}{n}$$

$$D = W_D W_D^T \Sigma \quad (73)$$

$$E \Sigma E^{-1} = W_D W_D^T \Sigma$$

$$\Sigma^{-1} E \Sigma E^{-1} = \Sigma^{-1} W_D W_D^T \Sigma$$

$$E^T E = W_D W_D^T$$

$$E^T = W_D$$

Esto indica que podemos transformar X a un conjunto no correlacionado de variables por medio de una multiplicación de la matriz con la transpuesta de los vectores propios de los datos de la matriz de la covarianza Σ [32].

2.6.13.2.2 Blanqueamiento

Ahora tenemos una forma de transformar nuestra matriz a una forma donde las variables no están correlacionadas, sin embargo, esto únicamente nos da una matriz de covarianza diagonal, no una matriz de covarianza identidad.

En orden a obtener una covarianza identidad necesitamos escalar cada dimensión para que su varianza sea igual a 1. ¿Cómo determinamos esta transformación? Sabemos como transformar nuestros datos para que la covarianza sea igual a D.

Si podemos determinar la transformación que deja $D=I$, luego podemos aplicar esta transformación a nuestra covarianza decorrelacionada para darnos la transformación deseada (blanqueamiento).

$$D^{-1} D = I \quad (74)$$

$$D^{-1} = D^{-\frac{1}{2}} I D^{-\frac{1}{2}}$$

De 73 y 76 tenemos

$$E \Sigma E^{-1} D^{-\frac{1}{2}} D^{-\frac{1}{2}} = I$$

Ahora con $Y = W_W X$ donde W_W es la matriz de transformación para el blanqueamiento de los datos. Desarrollando los mismos pasos para W_D tenemos que

$$\begin{aligned} W_W &= D^{-\frac{1}{2}} E^T \\ &= W^{-\frac{1}{2}} W_D \end{aligned}$$

2.6.13.2.3 Interpretación del blanqueamiento

La primera operación desnaturaliza los datos mediante pre-multiplicación con la matriz de vectores propios E^T , calculada a partir de la covarianza de los datos.

Esta decorrelación puede ser pensada como una rotación que reorienta los datos de manera que los ejes principales de los datos están alineados con los ejes a lo largo de la mayor de sus varianzas (siempre ortogonales) [32].

La segunda operación D^{-1-2} puede ser pensada como una compresión de los datos si la varianza a lo largo de una dimensión es mayor a uno o un alargamiento de los datos si la varianza a lo largo de una dimensión es menor a uno. El estiramiento y la compresión de los datos forman un tipo de esfera. Con el fin de observar los datos en el espacio original, a menudo es costumbre desrotar los datos de nuevo a su espacio original

$$W = E^{-1} D^{-\frac{1}{2}} E^T$$

2.6.13.3 Calculo de las matrices dispersas en el espacio de las medias de los grupos

Reducir la dimensionalidad de los datos que se van a visualizar desde su valor original a N_c (con un número completo de las clases), requiere una caída en el rango de solo una unidad para la matriz de dispersión calculada en el espacio de los medios de agrupamiento

$$S_W^c = \sum_{j=1}^{N_c} \sum_{i=1}^{N_j} (X_i^c - m_j^c)^T (X_i^c - m_j^c) \quad (75)$$

$$S_B^c = \sum_{j=1}^{N_c} (N_j (m_j^c - m^c)^T (m_j^c - m^c)) \quad (76)$$

$$M^c = (S_W^c)^{-1} S_B^c \quad (77)$$

Consecuentemente una diagonalización de la nueva matriz de dispersión M^c muestra, típicamente, que el trazo de la matriz contiene los valores propios más grandes.

2.7 Comunicación Matlab-Opensim

2.7.1 Matlab/Simulink

Millones de ingenieros y científicos en todo el mundo usan Matlab, es considerado como uno de los lenguajes más cómodos para trabajar, aunque actualmente ha sido un poco desplazado por lenguajes de tipo libres. Matlab es utilizado para analizar y diseñar los sistemas y productos que están transformando nuestro mundo.

La plataforma de Matlab es una multi herramienta utilizada para resolver problemas ingenieriles y científicos. Su lenguaje es uno de los más simples, además permite realizar operaciones vectorizadas lo que implica una ventaja con otros lenguajes de programación. Posee una de las más grandes librerías y herramientas de lenguajes de programación para muchas áreas en ingeniería [33]. El ambiente invita a la experimentación, exploración y descubrimiento.

2.7.1.1 Procesos en tiempo real

Como se describió la base de nuestro trabajo es un arreglo de datos los cuales fueron procesados en Matlab y teniendo en cuenta que como objetivo era el proyectar una ruta para una implementación en tiempo real de un clasificador de movimientos. Se planeo que cada operación que se le fuera a realizar a la base de datos fuera de forma que pudiese ser aplicado en tiempo real por lo tanto que fuese una operación rápida, cada una de estas operaciones no fue más que sumas/restas multiplicaciones/divisiones. En Matlab se decidió generar el código para establecer el tiempo de procesamiento.

2.7.2 Programas de simulación evaluados

Hay que hacer un paréntesis al momento de hablar sobre modelos y simulación, aunque bien se ha venido manejando los términos de simulación en un ambiente de programación ahora tendrá otro significado, modelo representará la forma de describir matemáticamente un comportamiento

de determinada circunstancia. Cuando comenzamos a hablar sobre programas de simulación del cuerpo humano se estará hablando de un modelo musculoesquelético evaluado en determinado software, y como modelo la reproducción del sistema anatómico humano representado como un sistema musculo esquelético virtual.

Dentro de los programas de simulación evaluados encontramos a Opensim, SIMM, MoCap (visual 3D profesional cmotion), Anybody (Anybody technology), lifeMOD (lifemodeler)

¿Porque seleccionar Opensim sobre otros programas de simulación?

Las características que buscamos en un software de simulación son que fuese libre, que poseyera modelos de miembro superior, poder hacer análisis musculoesqueléticos y que sea extensible lo que quiere decir que podamos programar sobre este software, modificar modelos y que puedan realizar cálculos en tiempo real, el único que poseía la mayoría de estas características sin contar evaluación en tiempo real aunque se está trabajando en esta aplicabilidad fue Opensim; en seguida mostraremos las características de dicho programa.

2.7.2.1 Opensim

Opensim es un Sistema programable extensible por el usuario validado y libre que deja a los usuarios desarrollar modelos de estructuras músculo esqueléticas y crear simulaciones dinámicas de movimiento. Es una plataforma para modelado del cuerpo humano, animales, robots y ambientes simulando su interacción y movimiento. Opensim tiene una interface gráfica de usuario para visualización de modelos y generación y análisis de simulaciones. El código abierto y extensible también incluye una interface de programación de aplicaciones (API) que los desarrolladores pueden usar para extender el programa o agregar funcionalidades a este.

- Opensim es un software que realiza análisis biomecánicos del cuerpo humano.
- A través de mediciones con cámaras de movimiento y marcadores de posición es posible reproducir movimientos.
- Está documentado como los movimientos que se realizan en el software son muy semejantes a la realidad. Los análisis de movimiento y de fuerzas tienen un error muy bajo con respecto a uno natural.

La plataforma de Opensim tiene un conjunto de capacidades, herramientas y beneficios. Con Opensim se puede: crear y editar modelos, analizar y simular modelos y movimientos, desarrollar flujos de trabajo y extender el programa, usar una plataforma de código abierto para

investigación y con fines didácticos para la enseñanza. Encontrando soporte, comunidades y fuentes de investigación.

2.7.2.1.1 Crear y editar modelos:

Opensim permite construir modelos con un amplio rango de estructuras musculoesqueléticas y muchos otros mecanismos. Los modelos pueden incluir cualquier combinación de cuerpos rígidos, simples y complejas articulaciones, resortes, límites, amortiguadores, controladores, músculos y actuadores. La interface gráfica de usuario te permite cargar y visualizar los modelos y editar cualquiera de estas propiedades.

Opensim utiliza modelos musculares de tipo Hill, captura las fuerzas activas y pasivas generando propiedades de músculos y son basadas en modelos bien probados de dinámicas musculo tendón de la literatura.

2.7.2.1.2 Librería de modelos musculoesqueléticos

Opensim consta con un conjunto extenso de modelos musculoesqueléticos validados de las extremidades superiores e inferiores. En adición con modelos de grado de investigación, hay modelos simplificados para prototipado y enseñanza.

2.7.2.1.3 Analizar y simular modelos y movimientos

Opensim consta de herramientas para importar datos experimentales, puede importar y visualizar datos de marcadores, uniones cinemáticas y fuerzas externas. Los usuarios de Opensim tienen creados y han compartido herramientas compatibles con una gran cantidad de sistemas de captura de movimiento. También posee una herramienta para escalar los modelos musculoesqueléticos, basados en datos medidos experimentalmente.

Otras de sus herramientas son la dinámica inversa, la optimización estática y la dinámica hacia adelante herramientas muy robustas que nos permiten hacer análisis del movimiento evaluado y resolver problemas musculares basados en algoritmos en la literatura. El control muscular computarizado puede generar activaciones musculares en el tiempo para determinado movimiento. El algoritmo está siendo usado para simular movimiento incluyendo el caminar, correr, saltar y evaluar la caminata patológica.

Opensim permite probar modelos y simulaciones para analizar virtualmente cualquier parte de interés. Por ejemplo, se puede usar la graficadora para crear graficas de momentos musculares, uniones cinemáticas, fuerzas musculares y más.

En el Opensim GUI se puede visualizar virtualmente cualquier componente del modelo. Opensim soporta los archivos para importar con extensión '*.stl*', '*.obj*', con esto es posible visualizar geometrías creadas bajo este formato.

Ofrece también una interface con Matlab y el GUI. Las completas librerías de Opensim de herramientas y modelos es completamente validas a través de Matlab y GUI. Esto permite a los programadores de Matlab y usuarios de GUI configurar un lote de rutinas o extender la funcionalidad de Opensim sin utilizar programación en C++.

La API provee acceso programacional a la infraestructura computacional detrás de Opensim y a los algoritmos empleados por la interface gráfica. Esto permite a los usuarios con habilidades de programación escribir sus propios programas y "*plug-ins*". En esta forma, ellos pueden combinar las herramientas existentes, la funcionalidad en nuevas formas y desarrollar nuevos componentes del modelo por ejemplo controladores y músculos. Plug-ins son accesibles a través de la interface gráfica y es una forma conveniente de compartir innovaciones técnicas con otros.

Opensim está construido sobre un rápido y robusto motor dinámico Simbody. Simbody es una fuente libre y abierta e incluye su propio conjunto de documentación y ejemplos.

El solo hecho de utilizar una plataforma de código libre es importante para la investigación y el avance del conocimiento porque barreras como la económica ya no son problema. El desarrollo de modelos y simulaciones en la plataforma de Opensim ¿pueden ser compartidas y reproducidas por otros usuarios. La habilidad para reproducir y verificar o extender los resultados de publicaciones científicas es vital para la investigación.

La plataforma de Opensim está abierta a aplicaciones académicas como a comerciales. Se le da la bienvenida a la academia, laboratorios clínicos, y más. Además de una comunidad de expertos. Construyendo nuestro propio código y modelos aislados está cambiando. Opensim provee una comunidad de expertos, no únicamente de la universidad de Stanford sino de laboratorios a lo largo de todo el mundo.

Aparte de todos estos beneficios posee un sitio web con un largo y gran repositorio de fuentes incluyendo una guía de usuario, tutoriales, ejemplos, foros, videos, webinarios y fuentes para desarrolladores. Además, regularmente se dan seminarios y cursos de entrenamiento alrededor del mundo [34].

2.7.2.1.3 Modelo seleccionado

En esencia Opensim tiene alrededor de 5 modelos de brazo con los cuales cada investigador puede modificarlos y crear sus propios modelos de miembro superior, para nuestros propósitos se decidió utilizar el que mejor se adaptaba a estos. Originalmente creado por Delp y Gonzalo para investigar el movimiento de la muñeca y el gasto que se visualizaba en procedimientos quirúrgicos. Este modelo consiste en todos los huesos del brazo y 10 grados de libertad. Pronación, supinación, flexión-extensión desviación ulnar y radial son todos incluyendo los grados de libertad para el codo, pulgar y dedo índice. Un total de 23 músculos actuadores del control del movimiento [34].

2.7.2.1.4 Programas utilizados para configurar el modelo

2.7.2.1.4.1 Rhinoceros 5

Rhino es un software de modelado 3D, se utiliza mucho para creación de cubiertas de formas libres, componentes repetitivos, estructuras complejas, tiene una interface gráfica de usuario muy cómoda para trabajar aplicable a arquitectura e ingeniería. Uno de los beneficios es que los formatos de conversión que maneja son amplios dentro de los cuales se encuentra el formato ‘.stl’ necesario para modificar las figuras geométricas de nuestro interés. Desafortunadamente no es un software de uso libre sin embargo la versión de prueba te permite trabajar todas sus funcionalidades.

2.7.2.1.4.2 Notepad++

Notepad++ es un editor de código fuente libre, dentro de los que evaluamos es el más amigable con el usuario, te permite identificar rápidamente secciones de tu código, con separación en bloques, por color, su forma de separar el código por bloques, basado en el poderoso editor Scintilla, es escrito en C++ y usa exclusivamente Win32 y STL (standard template library) el

cual provee la más alta velocidad de ejecución con la programación más pequeña. Actualmente se encuentran trabajando en reducir muchas de las rutinas sin perder amigabilidad.

2.7.2.1.4.3 Paraview 5.0.1

Paraview es una aplicación para visualización y análisis de datos multiplataforma de código libre. Los usuarios de *paraview* pueden rápidamente construir visualizaciones para analizar sus datos usando técnicas cuantitativas y cualitativas. La exploración de los datos puede ser hecho interactivamente en 3D o programáticamente usando las capacidades de procesamiento de *paraview*. *Paraview* fue desarrollado para analizar extremadamente bases de datos amplias usando fuentes computacionales de memoria distribuida. Esto puede correr sobre supercomputadoras para analizar bases de datos en una escala de petas tan bien como en laptops para datos pequeños, está siendo una herramienta en muchos laboratorios nacionales, universidades e industrias y está ganando mucho terreno en el rendimiento computacional.

2.7.3 Comunicación entre Opensim y Matlab

Como se ha visto *Opensim* es nuevo, posee gran aceptación en el mundo científico, posee foros donde los investigadores pueden publicar sus dudas con respecto al software o sus ideas para obtener consejos, *Matlab* es uno de los lenguajes de programación más usados por la investigación, por lo tanto, ya era una necesidad crear una interface de comunicación entre estos dos grandes. Actualmente existen librerías para desarrollar interfaces entre *Opensim* Y *Matlab*, sin embargo, esta interface es únicamente para generar cálculos, pasar datos a *Opensim* o viceversa. Es posible generar tablas, graficas, su comunicación se encuentra en la etapa donde un software utiliza las características principales del otro software. *Opensim* probablemente se quede corto al momento de ejecutar algoritmos computacionalmente costosos, con lo que *Matlab* puede ser ayudarlo, en cambio *Matlab* no cuenta con el diseño gráfico de *Opensim* para hacer este tipo de análisis biomecánicos virtuales por lo que se ayuda de *Opensim*, la comunicación que buscamos crear es desarrollar en *Matlab* un modelo de clasificación utilizable en una prótesis multifuncional, validando el movimiento con *Opensim*.

Capítulo 3 Experimentación

3.1 Estructura de la base de datos evaluada

La Tabla 3.1 muestra la subestructura de datos que simboliza a una repetición de alguno de los 15 movimientos de la base de datos. La frecuencia de muestreo a la que se capturaron los datos fue de 4 KHz, como se explicó en la sección 2.3.1, y se capturan 8 canales de señales sEMG. Todos los movimientos en la base de datos tienen una duración de cinco segundos, por lo que cada movimiento se compone de 8 columnas por 20,000 renglones, y cada dato muestra es de un byte.

Tabla 3. 1 Matriz de datos para una repetición de un movimiento.

	no. muestra	sEMG Electrodos (Canal)							
		1	2	3	4	5	6	7	8
5 seg a una frecuencia de muestreo de 4 KHz	1	Dato 8 bits	Dato 8 bits	Dato 8 bits	Dato 8 bits	Dato 8 bits	Dato 8 bits	Dato 8 bits	Dato 8 bits
	2								
	3								
	4								
	...	...	...	...	...	...	...	...	...
	20,000								

La siguiente composición de los datos sería la matriz de datos crudos por movimiento que incluye las 12 repeticiones, como se muestra simbólicamente en la Tabla 3.2. Esta tabla se compone de 8 columnas por 240,000 renglones.

Cada 20,000 renglones de la Tabla 3.2 representa una Tabla 3.1.

Tabla 3. 2 Matriz de datos para un movimiento con 12 repeticiones.

HC	sEMG Electrodos (Canal)							
no. muestra	1	2	3	4	5	6	7	8
1 a 20,000	Repetición 1							
21,000 a 40,000	Repetición 2							
40,001 a 60,000	Repetición 3							
...	Repetición 4							
...	Repetición 5							
...	...							
220,001 a 240,000	Repetición 12							

Un nivel más de conformación de los datos se muestra en la Tabla 3.3, donde se muestra los datos de los 15 movimientos con sus doce repeticiones para un mismo sujeto. Esta matriz llega a tener el tamaño de 8 columnas por 3,600,000 renglones.

Tabla 3. 3 Representación completa de los datos de un sujeto de prueba.

Movimientos	1	2	3	4	5	6	7	8
1.-HC	Repetición 1 -12							
2.-I_I	Repetición 1 -12							
3.-I_M	Repetición 1 -12							
4.-IMR	Repetición 1 -12							
5.-L_L	Repetición 1 -12							
6.-M_M	Repetición 1 -12							
7.-M_R	Repetición 1 -12							
8.-MRL	Repetición 1 -12							
9.-R_L	Repetición 1 -12							
10.-R_R	Repetición 1 -12							
...	Repetición 1 -12							
15.-T_T	Repetición 1 -12							

El mismo procedimiento se realizó para cada uno de los 8 sujetos de prueba. Teniendo como resultado la Tabla 3.3 para cada sujeto.

3.2 Valores atípicos

En esta etapa es común pensar que se pueden tomar varias rutas para eliminar valores atípicos, una de las posibles es eliminar dichos valores de un movimiento en particular a partir de la Tabla 3.2 y realizarlo para cada movimiento, o en lugar de eso eliminarlos de la representación completa de movimientos en general con la Tabla 3.3. Se procedió a obtener estos valores por las dos rutas ya que era necesario corroborar la similitud de los resultados y ver cómo afectan los enfoques en la eliminación de valores atípicos.

Ruta 1: Eliminación de valores atípicos de un movimiento en particular para después repetir el proceso en cada uno de los movimientos (Tabla 3.2).

Se formaron matrices de cada uno de los movimientos y se procedió a eliminar los valores atípicos.

Ruta 2: Eliminación de valores atípicos del conjunto de los 12 movimientos (Tabla 3.3).

Con la matriz completa mostrada en la Tabla 3.3 se procedió a calcular los valores atípicos.

La eliminación de los valores atípicos por cualquiera de las dos rutas se realizó por medio de la distancia Mahalanobis y el valor crítico de una distribución chi cuadrada.

La eliminación de valores atípicos determina que valor en el espacio se aleja de todas las mediciones como se determinó en la sección 2.4.1. Al obtener las distancias mahalanobis de cada punto, este valor se compara con el valor crítico de una distribución chi cuadrada con 7 grados de libertad, por lo tanto, se elimina todo punto muestral donde su valor dM supere el valor crítico.

Después de identificar que datos eran valores atípicos y cuáles no, se procedió a remover los datos atípicos para crear una matriz de datos ilustrada en la Tabla 3.4 representando la ruta 2. En esta tabla ya solo existen los valores que no son atípicos por lo tanto los movimientos ya no poseen el total de las mediciones, poseen n mediciones donde $n < N$, y N es el número total que son 240,000 datos para un movimiento Tabla 3.2 y el subíndice representa el tipo de movimiento.

Tabla 3. 4 Datos mantenidos de un sujeto de prueba después del procesamiento de valores atípicos.

Movimientos	1	2	3	4	5	6	7	8
1.-HC	n_1 valores mantenidos							
2.-I_I	n_2 valores mantenidos							
3.-I_M	n_3 valores mantenidos							

4.-IMR	n_4 valores mantenidos
5.-L_L	n_5 valores mantenidos
6.-M_M	n_6 valores mantenidos
7.-M_R	n_7 valores mantenidos
8.-MRL	n_8 valores mantenidos
9.-R_L	n_9 valores mantenidos
10.-R_R	n_{10} valores mantenidos
...	$n_{...}$ valores mantenidos
15.-T_T	n_{15} valores mantenidos

3.3 Ventaneo

Una vez que se tienen solo valores representativos de los movimientos, sin distorsión se separaron cada uno de los movimientos en matrices como se muestra en la Tabla 3.5.

Tabla 3. 5 Matriz para cada movimiento.

Movimiento i	sEMG Electrodo (Canal)							
No. muestra	1	2	3	4	5	6	7	8
De 1 a n_j	Valores mantenidos							

Los movimientos se separan para procesar las características que seleccionamos en la sección 2.4.2.1, el método para separar es fácil ya que al momento de eliminar los valores atípicos se tienen los índices indicadores, de la matriz original (Tabla 3.3) y de como quedaron (Tabla 3.4). Y cada matriz de movimiento Tabla 3.5 se dividió en ventanas de 250 muestras, como se ilustra en la Tabla 3.6 para obtener m número de ventanas, donde $m < M$, y M es el número total de ventanas (960) si no se hubiesen eliminado valores atípicos.

Tabla 3. 6 Ventana de 250 datos.

HC	sEMG Electrodo (Canal)							
No. muestra	1	2	3	4	5	6	7	8
1								
2								
3								
...								
250								

A cada canal de la ventana (columna) se le aplicaron 12 operaciones, explicadas en la sección 2.4.2.1 entre la que encontramos a MAV, MSV, SSI, NZC, RMS, VAR, TM3, TM4, TM5, WL, SSC, MPR.

Y se obtuvo un nuevo vector con 96 columnas como se muestra a continuación en la Tabla 3.7.

Tabla 3. 7 Nuevas características para un número de N ventanas ($N \times 96$) por cada movimiento.

No. Ventana	MAV								MSV	SSI	NZC	RMS	...	MPR
	E1	E2	E3	E4	E5	E6	E7	E8	...	...	...	...	...	...
1	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...								...	...	...	...	...	...
M	...								...	...	...	...	...	...

La nueva matriz se etiqueta en su última columna (97) con el indicador del tipo de movimiento. por ejemplo:

Matriz para movimiento 1 (HC).

Tabla 3. 8 Matriz / movimiento

No. Ventana	MAV	MSV	SSI	NZC	RMS	...	MPR	Tipo de movimiento
1	...	...	...	...	...	...	...	1
2	...	...	...	...	...	...	...	1
...	...	...	...	...	...	...	...	...
M	...	...	...	...	...	...	...	1

3.4 Clasificación

Hasta este momento poseemos 15 matrices del tipo de la Tabla 3.8, una para cada movimiento. A partir de esto se crea una nueva matriz con los 15 movimientos y el total de ventanas obtenidas de cada movimiento (ver Tabla 3.9).

Tabla 3. 9 Matriz con los 15 movimientos y el número total de ventanas

No. Ventana	MAV	MSV	SSI	NZC	RMS	...	MPR	Tipo de movimiento
1	0.001	...	...	15	...	...	...	1
2	0.002	...	...	19	...	...	...	1
...	0.004	...	...	28	...	...	...	...
m_1	0.002	...	...	22	...	...	...	1
1	...	...	...	...	...	...	...	2
2	...	...	...	...	...	...	...	2
...	...	...	...	...	...	...	...	...
m_2	...	...	...	...	...	...	...	2
...	...	...	...	...	...	...	...	...
1	...	...	...	...	...	...	...	15
2	...	...	...	...	...	...	...	15
...	...	...	...	...	...	...	...	...
m_{15}	...	...	...	...	...	...	...	15

El mismo procedimiento se realiza con los ocho sujetos evaluados. Al final tendremos una matriz con 96 características con su etiqueta de clase (tipo de movimiento) para cada sujeto evaluado.

Se utilizan 3 algoritmos para crear modelos de clasificación Naive Bayes, Análisis discriminante cuadrático y Análisis discriminante lineal.

3.4.1 Naive Bayes gaussiano

Implementación de Naive Bayes Gaussiano.

Se calculan las probabilidades a priori de los valores de entrada para cada clase usando la frecuencia. Obtenemos las probabilidades a priori calculando la frecuencia de cada una de las clases. $P(\text{clase } 1)$, $P(\text{clase } 2)$, ..., $P(\text{clase } 15)$.

Con entradas de valor real, podemos calcular la media y desviación estándar de los valores de entrada x para cada clase para resumir la distribución. Esto significa que además de las probabilidades para cada clase, es decir también almacenamos la media y la desviación estándar para cada variable de entrada para cada clase.

Entonces se calcula una media y una desviación estándar para cada clase de cada variable.

Esto es tan simple como calcular los valores de desviación media y estándar de cada variable de entrada (x) para cada valor de clase.

Las probabilidades de nuevos valores de x se calculan utilizando la función de densidad de probabilidad gaussiana (PDF).

$$probabilidad(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(x-\mu)^2}{2\sigma^2}}$$

Al hacer predicciones, estos parámetros pueden ser conectados al PDF Gaussiano con una nueva entrada para la variable y , a cambio, el PDF Gaussiano proporcionará una estimación de la probabilidad de ese nuevo valor de entrada para esa clase.

Se realiza la operación para la probabilidad a posteriori para cada clase en nuestro caso se realizó para las 15 clases.

En este momento ya se tiene la probabilidad a posteriori de cada clase se multiplica por la probabilidad a priori de cada clase respectivamente.

La clase que obtenga la probabilidad a posteriori más alta es la seleccionada.

3.4.2 LDA

La regla de Bayes dice que deberíamos elegir una clase que tenga la probabilidad máxima posterior dada el vector de características x .

Determinamos el número de clases y la cantidad de cada una de ellas en nuestro set de entrenamiento.

Determinamos los datos de cada una de las clases.

Obtenemos la media y la covarianza de cada una de las clases.

Calculamos la función discriminante para cada una de las clases

$$G^x = \arg \max_k (x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log(\pi_k))$$

La clase que posea la más alta probabilidad para la función discriminante es a la que pertenece nuestro conjunto evaluado

3.4.3 QDA

Como vimos QDA no es realmente muy diferente de LDA, excepto que asumimos que la matriz de covarianza puede ser diferente para cada clase y así, vamos a estimar la matriz de covarianza Σ_k por separado para cada clase k , $k = 1, 2, \dots, K$.

Función discriminante cuadrática:

$$G(x) = \arg \max_k (-\frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) + \log \pi_k)$$

Esta función discriminante cuadrática es muy parecida a la función discriminante lineal, excepto que debido a que Σ_k , la matriz de covarianza no es idéntica, no se pueden descartar los términos cuadráticos. Esta función discriminante es una función cuadrática y contendrá términos de segundo orden. La regla de clasificación es similar también a LDA.

3.5 Normalización

Debido a que las nuevas variables estaban a diferentes escalas (ver Tabla 3.9), por ejemplo, los números de cruces por cero estaban en números enteros y algunos promedios evaluados estaban aún en milivolts, se procedió a normalizar las características dejando a todas las nuevas características a una misma escala.

Se realizó la normalización muestral para nuestro conjunto de datos.

$$\text{Datos Normalizados} = \frac{X - \bar{X}}{S}$$

Por lo que se sacó la media muestral de cada variable $\bar{X}$ y se le resto a cada uno de los datos X y este se dividió sobre su varianza S .

Con esta nueva tabla se procede a generar un modelo de clasificación para los 15 movimientos a partir de las nuevas características con Naive Bayes, análisis discriminante lineal y análisis discriminante cuadrático. La Figura 3.1 muestra el procedimiento general para la ruta número dos.

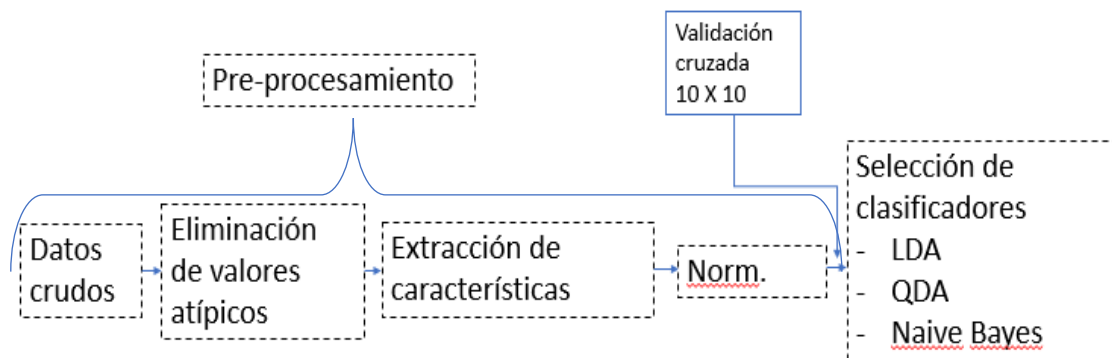


Figura 3. 1 Secuencia de experimentación inicial, involucra procesos desde los datos crudos hasta la normalización para conocer y crear un primer modelo de clasificación referente.

3.6 Selección de características hacia adelante

Uno de los métodos utilizados para reducir la dimensionalidad de los datos es una selección de características hacia adelante tomando como método para seleccionar dichas características un clasificador y su porcentaje de reconocimiento, es decir aquel subconjunto dentro del conjunto total de características evaluadas que de un mejor porcentaje de reconocimiento es el que se mantendrá ver Tabla 3.10 donde se muestra el pseudocódigo. Esta es una técnica de selección de atributos

$$\begin{array}{c} X_1 \\ X_2 \quad X_1 \\ X_3 \rightarrow X_2 \\ \dots \quad X_M \\ X_N \end{array}$$

Por definición dado un conjunto de características $X = \{x_i | i=1 \dots N\}$, encontrar un conjunto Y_M , con $M < N$, que maximice una función objetivo $J(Y)$, idealmente $P(\text{correcto})$

$$Y_M = \{x_1, x_2, x_3, \dots, x_{iM}\} = \operatorname{argmax} J\{x_i | i = 1 \dots N\}$$

El Pseudocódigo mostrado en la Tabla 3.10 es un procedimiento aplicable a cualquier lenguaje de programación.

Tabla 3. 10 pseudocódigo de SFS

```

SUBCONJUNTO= 0; MEJOR_EVAL = 0;
ESPACIO=1...L
SALIDA= falso
WHILE NOT SALIDA
FOR i=1 A |ESPACIO|
TEMP(i)= J (SUBCONJUNTO v ESPACIO(i))
END
MEJOR_i = ARGMAX(TEMP)
IF TEMP(MEJOR_i)>MEJOR_EVAL
subconjunto      =      SUBCONJUNTO      +
ESPACIO(MEJOR_i)
MEJOR_EVAL= TEMP(MEJOR_i)
ESPACIO = ESPACIO – ESPACIO(MEJOR_i)
ELSE
SALIDA = verdad;
END
END
RETURN SUBCONJUNTO, MEJOR_EVAL

```

Habiendo hecho la selección de características se obtendrá un vector de características obtenidas y uno de características eliminadas. Con la Figura 3.2 podemos ver la secuencia de la experimentación incluyendo la selección de características del mismo modo podemos ver que la Figura 3.2 es la misma que la Figura 3.1 con la adición del bloque SFS.

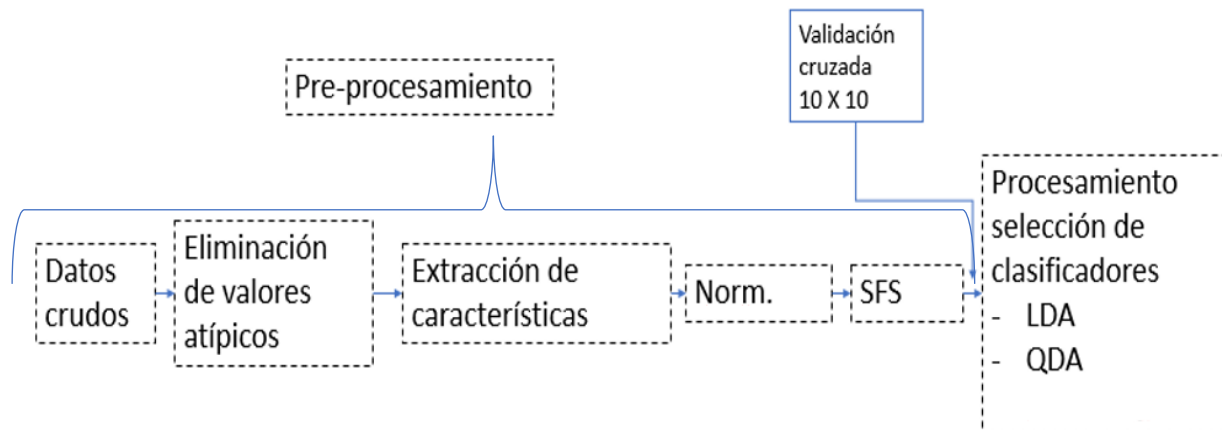


Figura 3. 2 Secuencia de experimentación 2. Se agrega un algoritmo de selección de características (SFS) para mejorar el modelo de clasificación de referencia.

3.7 Clasificación en la secuencia de experimentación 2

En esta nueva clasificación se utilizaron sólo los modelos de clasificación QDA y LDA, ya que por los resultados obtenidos en la secuencia de experimentación inicial el costo computacional del clasificador Naive Bayes gaussiano es muy alto.

3.8 Algoritmo de matrices dispersas

Con el subconjunto de los datos seleccionados obtenidos a partir de la sección 3.6, se procede a implementar una transformación de nuestros datos a un nuevo espacio con el algoritmo de visualización basado en matrices dispersas. El algoritmo usado se utiliza en primera instancia para aumentar las distancias entre clases y reducir las distancias entre los datos de cada una de las clases utilizando la etiqueta de clase; Se muestra la descripción de este método en la Figura 3.3.

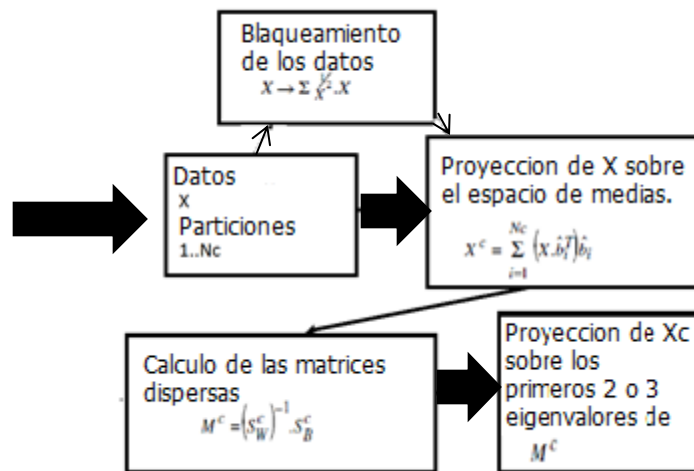


Figura 3. 3 secuencia de procesamiento para el algoritmo de matrices dispersas.

- En este algoritmo los datos se blanquean como se describió en la sección 2.6.13 obteniendo la *media de cada una de las variables de todos los datos*¹ y restándola a los mismos datos (Xn) posteriormente multiplicando los datos Xn por la *matriz de blanqueo* W^2 para obtener Xw, que son nuestros datos ya blanqueados.
- Después los datos blanqueados se ortogonalizan multiplicándolos por la *matriz* $\hat{b}^3$ del algoritmo de Gram-schmidt como se mostró en la sección 2.6.13.1. De este modo se proyectan los datos ya blanqueados a este espacio ortogonalizado para obtener Xc.
- Se calculan las medias por clase y la media total de los datos originales.
- Se calculan las matrices Sw y SB mostradas en la ecuación 77 y 78
- Se calcula MC con la ecuación 79
- Se calcula la covarianza de MC
- Se calculan los valores y vectores característicos de la covarianza de MC.
- Usando datos Xc (datos blanqueados y ortogonalizados) y los valores y vectores característicos de MC se crea una *matriz de proyección* (Mp^4). Donde esta matriz la utilizamos para evaluar valores de entrada originales o en su caso nuevos que se proyectan al nuevo espacio, donde es posible seleccionar las columnas con mayor valor característico, las cuales surgen en orden ascendente de la nueva matriz.

De esta forma tenemos una nueva proyección de los datos donde podemos seleccionar cuantas variables incluir en el nuevo modelo. Esto es con un número menor de características en el nuevo espacio podemos representar la mayoría de la información del total de los datos.

¹ El vector de medias generado en este procedimiento se reutilizará en el proceso tiempo real.

² La matriz de blanqueo se reutilizará en el proceso tiempo real.

³ La matriz de ortogonalización también se reutilizará en el proceso tiempo real.

⁴ La matriz de proyección es la última multiplicación para transformar nuestros datos en este algoritmo.

Así creamos un modelo con un menor número de variables, incluso podemos bajar el número de variables evaluadas por ejemplo de tener 49 variables a solo 3.

En este algoritmo como se observa para poder transformar nuevos datos de entrada es necesario almacenar las matrices de transformación indicadas con superíndices.

3.9 Clasificación en la secuencia de experimentación 3

La siguiente secuencia de experimentación es la que se muestra en la Figura 3.4 donde se le ha agregado la experimentación de la sección anterior a la experimentación 2, el bloque de conversión al espacio de medias y matrices dispersas.

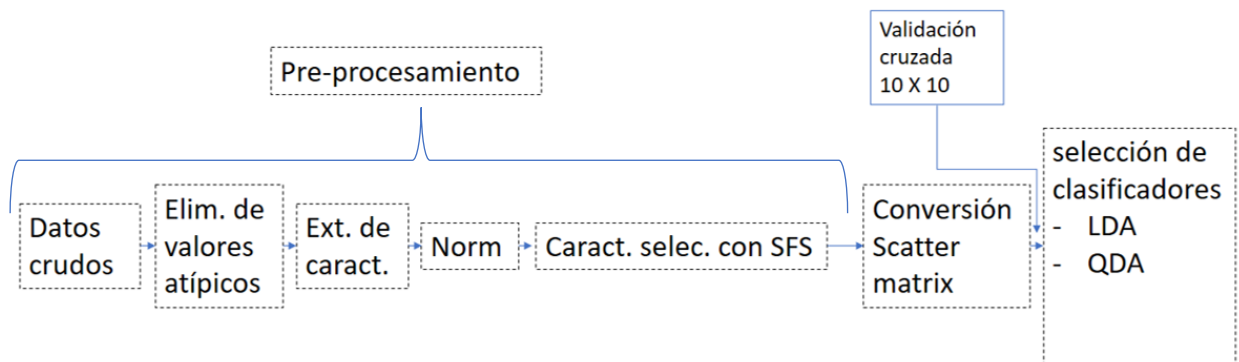


Figura 3. 4 Secuencia de experimentación 3. Adición de una conversión a un nuevo espacio para simplificar el modelo de clasificación de la experimentación 2.

3.10 Secuencia de procesamiento cuando ya se tiene el modelo de clasificación creado

Después de haber creado los modelos de clasificación a través de los algoritmos QDA y LDA, se estableció una ruta para evaluar nuevas clasificaciones en tiempo real con el retraso ya indicado de 62 ms que es el tiempo de duración del ventaneo seleccionado, la ruta se muestra en la Figura 3.5 que describe el diagrama general de este procedimiento.



Figura 3. 5 Diagrama para el procesamiento y evaluación de nuevos datos, en la experimentación 1 2 y 3.

3.11 Técnica para aumentar el porcentaje de reconocimiento en un grupo muestral

Después de implementar y observar los modelos de clasificación en las tres secuencias experimentales descritas anteriormente se encontró que los resultados individuales eran mucho mejores que los del grupo muestral, como se podrá ver en la sección de resultados, por esta razón se pensó en que debía haber alguna relación entre los individuos o mejor aún alguna explicación donde las mediciones individuales de cada uno de ellos les permitiera ser buenos diferenciadores de sus movimientos cosa que no ocurría cuando se evaluaba un grupo de ellos donde había un error mucho mayor. Esto va de acuerdo con la premisa de que fisionómicamente los seres humanos nos dividimos en grupos en función de nuestra semejanza, por lo tanto, es fácil concluir que nuestras mediciones serán un tanto parecidas si nuestra fisionomía es semejante, para aprovechar esta ventaja se decidió extender el modelo de clasificación grupal agregando distintas

clases a un mismo movimiento siempre y cuando fuesen de sujetos distintos, y luego reduciendo al espacio de categorías originales. En la siguiente sección se detalla a profundidad el procedimiento.

3.11.1 Secuencia del procesamiento con las clases modificadas con el modelo de clasificación ya creado en las experimentaciones anteriores

Se creó una técnica para aumentar el porcentaje de reconocimiento como se describe a continuación:

- Empezaremos con la matriz ya definida anteriormente la Tabla 3.9. como mencionamos en la sección 3.7.1 este procedimiento solo se realiza cuando evaluamos el grupo muestral de los sujetos 1 al 7 (S 1-7) por lo tanto a cada sujeto evaluado se le da un nuevo número de clases para cada uno de sus movimientos por ejemplo S1=clases 1 al 15, S2=clases 16 al 30, S3=31 al 45 así sucesivamente hasta el sujeto 7 con el número de categoría 105.
- Se implementa la experimentación ya sea 1, 2 o 3 con la nueva matriz completa (sujetos del 1 al 7).
- Se regresan las clases originales a los datos. Clases del 1 al 15 para todos los sujetos lo que indica que la clase 16, 31, 46, así sucesivamente seria en realidad la clase 1, la clase 17, 32, 47, ..., serían la clase 2 y así hasta la clase 15. Lo que representaría la operación modulo.
- Y se calcula el error.

En esta técnica se utilizó una tabla de extensión de clases, 15 movimientos por 7 sujetos (105 clases), pero la secuencia de aplicación es semejante a la experimentación 1, 2 y 3 como se ilustró en la Figura 3.5. Lo que faltaría en la Figura 3.5 es el proceso final al cambiar a la clase correspondiente entre 1 y 15. Recordando que en la experimentación 3 se seleccionan el número de variables deseadas de la nueva matriz transformada para el modelo de clasificación utilizado.

3.12 Opensim

Opensim como se describió en el capítulo 2 sección 2.7.2.1 es un software de simulación musculoesquelético virtual que se utiliza para realizar las simulaciones de movimiento.

El modelo musculoesquelético que se decidió utilizar como ya se indicó en el capítulo 2 es de Gonzalez, R. V., Buchanan, T. S. y Delp, S. L [34] en la sección 2.7.2.1.3.

3.12.1 Configuración del modelo seleccionado en Opensim

3.12.1.2 Modelo Opensim

El modelo seleccionado se encuentra en un archivo con extensión ‘. osim’ llamado wristmodel ver Figura 3.7.

```
1  <?xml version="1.0" encoding="UTF-8" ?>
2  <OpenSimDocument Version="30000">
3    <Model name="Wrist Model">
4      <credits>Gonzalez, R.V., Buchanan, T.S., Delp, S.L.</cr
5      <publications>How muscle architecture and moment arms a
6      <length_units>meters</length_units>
7      <force_units>N</force_units>
8      <!--Acceleration due to gravity.-->
9      <gravity> 0 -9.80665 0</gravity>
10     <!--Bodies in the model.-->
11     <BodySet>
4607   <!--Constraints in the model.-->
4608   <ConstraintSet>
4649   <!--Forces in the model.-->
4650   <ForceSet>
7381   <!--Markers in the model.-->
7382   <MarkerSet>
7386   <!--ContactGeometries in the model.-->
7387   <ContactGeometrySet>
7391   <!--Controllers in the model.-->
7392   <ControllerSet name="Controllers">
7396   <!--Additional components in the model.-->
7397   <ComponentSet name="MiscComponents">
7401   <!--Probes in the model.-->
7402   <ProbeSet>
7406   </Model>
7407 </OpenSimDocument>
7408
```

Figura 3. 6 Archivo en formato .osim del modelo utilizado visto en notepad++, este es el formato como representa el código este programa se puede observar como el archivo del modelo se encuentra dividido en secciones que definen al modelo.

En la sección de Bodyset encontramos todos los elementos óseos que componen al modelo, la forma en la que se llama a archivos externos nos referimos a cada una de las figuras geométricas que componen al modelo (ver Figura 3.8).

```

<DisplayGeometry>
  <!--Name of geometry file .vtp, .stl, .obj-->
  <geometry_file>radius.vtp</geometry_file>
  <!--Color used to display the geometry when visible-->
  <color> 1 1 1</color>
  <!--Name of texture file .jpg, .bmp-->
  <texture_file />
  <!--in body transform specified as 3 rotations (rad) followed by 3 tra
  <transform> -0 0 -0 0 0 0</transform>
  <!--Three scale factors for display purposes: scaleX scaleY scaleZ-->
  <scale_factors> 1 1 1</scale_factors>
  <!--Display Pref. 0:Hide 1:Wire 3:Flat 4:Shaded-->
  <display_preference>4</display_preference>
  <!--Display opacity between 0.0 and 1.0-->
  <opacity>1</opacity>
</DisplayGeometry>

```

Figura 3. 7 Sección específica BodySet del archivo donde encontramos definida una figura geométrica con formato ‘.vtp’.

La geometría en la Figura 3.8 que se está llamando desde un archivo externo ‘radius.vtp’ desde la sección de BodySet es la del radio.

Ahora nos podemos ir directamente sobre el archivo, todas estas figuras que componen el modelo están almacenadas dentro de un folder en la carpeta de instalación de Opensim buscamos el nombre del archivo por ejemplo ahora ‘radius.vtp’.

Al encontrar dicho archivo se abrirá con el programa *paraview* para modificar el formato a la figura a la cual se le aplicará un filtro y se creará un nuevo archivo con formato ‘. sto’ que puede ser leído con el programa *Rhino* donde se modificará.

A la Figura 3.9 se le aplicará un filtro triangular para poder guardarla con un formato ‘. sto’ y modificar su posición en el espacio en otro programa. Como ya se mencionó Paraview solamente se utiliza para modificar el formato a la figura.

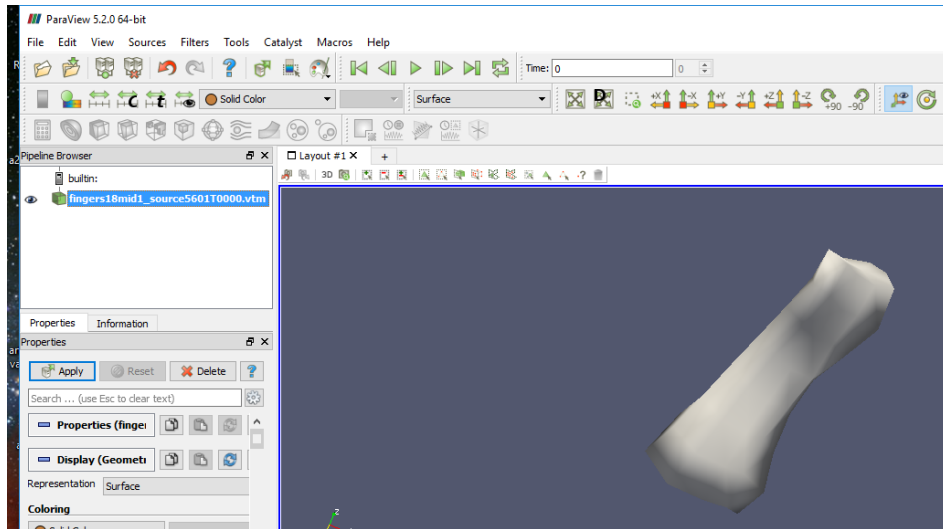


Figura 3. 8 Radius.vtp leído con paraview donde se le aplicara un filtro y se guardara con formato '. sto'

En *Rhino* se hace el cambio en la posición del origen de la figura para adaptarla al nuevo modelo. Posteriormente se vuelve a pasar a *Paraview* y se regresa al formato original. Y se guarda en vez del archivo original en la carpeta de Opensim donde están todas las geometrías.

3.11.1.3 Configuración del modelo musculoesquelético, cambiando el origen de las figuras geométricas

El modelo Opensim como ya se mencionó posee solo movimiento en los dedos pulgar e índice los otros tres están desprovistos de flexión por lo tanto era necesario modificarlos.

Una observación en el modelo original es que los dedos medio, anular y meñique, aunque poseían figuras por separado, había una figura para cada hueso, estaban espacialmente definidas solo para hacer el acople entre las tres figuras la falange, falangeta y la falangina.

Lo que se debe hacer es cambiar el punto de origen de la figura original y acomodar esa figura exactamente en el punto creado para la flexión del movimiento, por lo que era necesario cambiar el punto de referencia para el origen de las nuevas figuras. El modelo completamente modificado puede verse en la Figura 3.10.

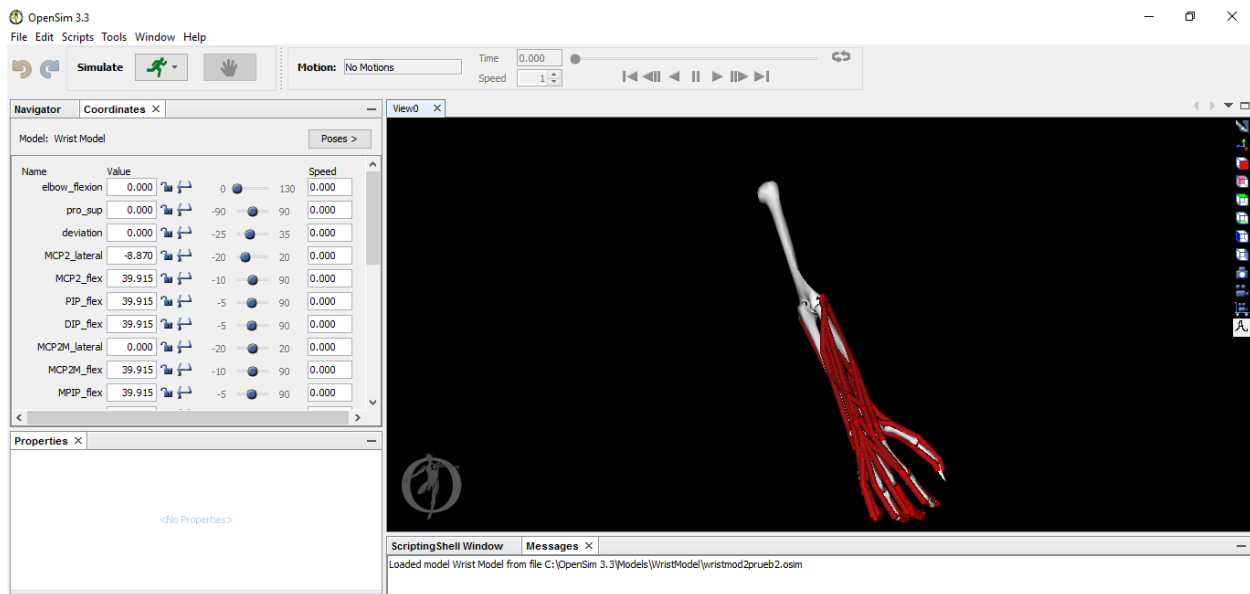


Figura 3. 9 Modelo terminado de mano con todos sus movimientos

3.12.2 Desarrollo de los movimientos

Para reproducir un movimiento en OpenSim es necesario un archivo que describa el movimiento de cada una de las articulaciones del modelo musculoesquelético evaluado, como primer paso creamos cada uno de los 15 archivos de movimiento, teniendo en cuenta que cada 62.5 ms (tiempo de duración del ventaneo) correspondería a un cuadro de avance de la imagen para el movimiento. Donde un vector de avance correspondería a este grado de rotación para cada una de las articulaciones en el desarrollo del movimiento. OpenSim necesita de estos vectores para establecer los cuadros de avance del modelo virtual.

Procedimiento:

Se tenía la posición inicial de cada uno de los movimientos evaluados, por ejemplo mano extendida en semireposo, tomando en cuenta la información inicial de la base de datos. Se tenía la posición final (flexión completa) general de cada uno de los movimientos evaluados ver la Figura 3.10.

Como definimos OpenSim se utilizará solo para la visualización y evaluación de los movimientos es por eso por lo que las posiciones iniciales y finales pueden ser representativas y por esta razón es por eso por lo que decidimos utilizar un movimiento ideal para todos los movimientos evaluados. Como movimiento ideal podemos suponer que la posición inicial empieza con las

articulaciones extendidas y como movimiento final las articulaciones flexionadas como se muestran en la Figura 3.10. El tiempo de duración de cada uno de los movimientos fue de 5 segundos.

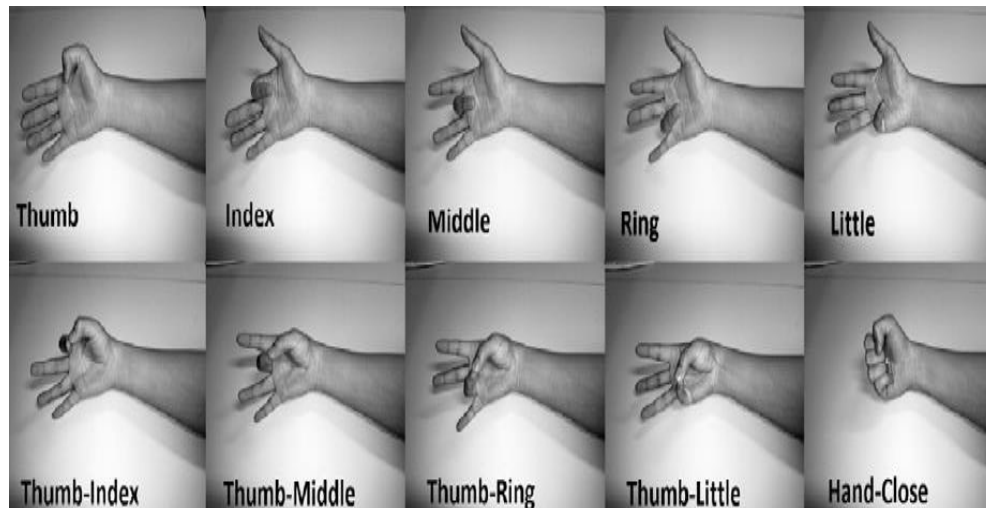


Figura 3. 10 Posición final de 10 de los 15 movimientos evaluados.

Un archivo de movimiento en Opensim consta de un encabezado describiendo la información del archivo a que modelo pertenece y el cuerpo del archivo que consta de n número de columnas tituladas con el nombre de la articulación del modelo como se puede observar en la Tabla 3.11. La primera columna a la izquierda pertenece al tiempo del movimiento realizado. En cada una de estas columnas (articulaciones del modelo) se encuentra la medición en grados sexagesimales de movimiento en el tiempo.

Tabla 3. 11 A continuación se representa una sección del formato de un archivo ‘.mot’ para reproducir el movimiento en un modelo Opensim. Normal_close representa el nombre del modelo al que se le aplicará el movimiento, versión es la versión del archivo, nRows y nColumns son tanto las filas como las columnas del archivo completo, inDegrees indica si las mediciones están en grados, endheader indica el fin de los datos del encabezado para dar inicio a los datos en grados de cada una de las articulaciones, donde la primer columna representa el tiempo y las siguientes cada uno de los grados de libertad del modelo al que se le aplicará este archivo de movimiento (elbow_flexion, MCP2_flex, PIP_flex y DIP_flex).

normal_close				
version=1				
nRows=51				
nColumns=26				
inDegrees=yes				
endheader				

time	elbow_flexion	MCP2_flex	PIP_flex	DIP_flex
0	0	-10	-5	-5
0.02	0	-8	-3.1	-3.1
0.04	0	-6	-1.2	-1.2
0.06	0	-4	0.7	0.7
0.08	0	-2	2.6	2.6
0.1	0	0	4.5	4.5

Al poseer las tablas para cada uno de los 15 movimientos del modelo musculoesquelético ver Tabla 3.11, donde cada cuadro de avance se realizaba en 62.5 ms, lo que indica que en 5 segundos había 80 cuadros de avance para el movimiento musculoesquelético del modelo virtual. Se contó con 16 vectores de avance (uno para cada movimiento dos para el movimiento de mano cerrada) guardados para en cada decisión tomada por el modelo de clasificación se sumará a la matriz de movimiento final.

3.12.3 Estimación del movimiento en Matlab

La reproducción de los movimientos de la mano se realizó en función del modelo de clasificación obtenido a partir de QDA y su precisión para calcular los movimientos en el tiempo. Aunque sabemos de antemano que poseemos modelos de clasificación particulares (para cada sujeto o para un grupo muestral de 7 sujetos) con un porcentaje de reconocimiento que en promedio para cada uno de los sujetos evaluados por separado es de 94.66% y para el grupo muestral de 7 sujetos es de 91.37%, como se puede corroborar en la Tabla 4.20 de la sección de resultados. Lo que indica que el modelo matemático es capaz de reproducir el movimiento virtual en esos porcentajes.

Se crearon los vectores de avance para cada movimiento en Matlab, donde un vector de avance es un vector que posee el grado de avance en grados de cada una de las articulaciones en cada decisión.

3.12.4 Diagrama general de procesamiento

En la sección anterior se explicó el procedimiento de cómo se pensó en la secuencia del procesamiento de la toma de decisiones para reproducir un movimiento a partir de las clases obtenidas a partir del modelo cuadrático se definió la secuencia del procedimiento (ver Figura 3.11 y Tabla 3.12).

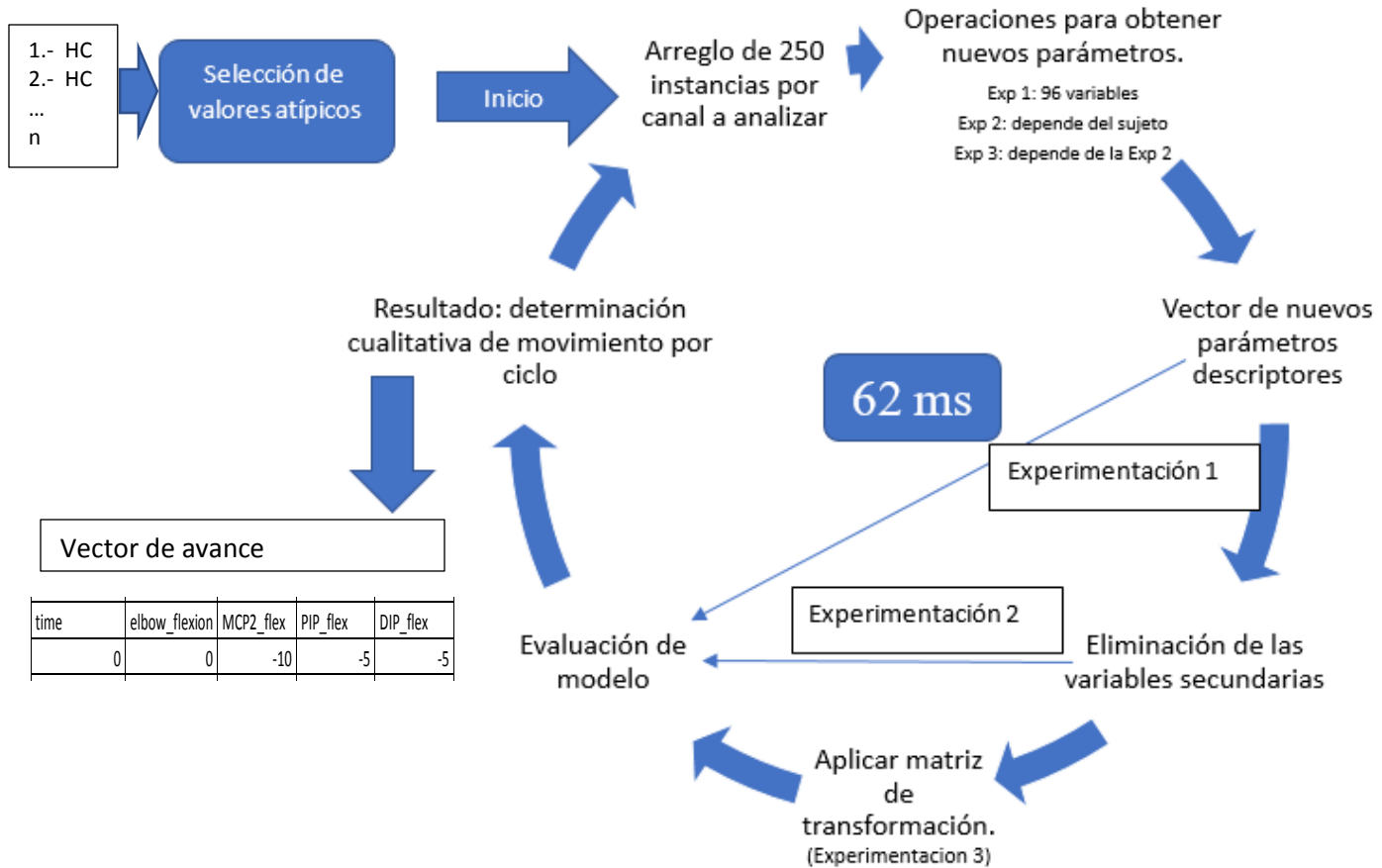


Figura 3. 11 Diagrama para la reproducción del movimiento a partir de las sEMG hasta Opensim.

Tabla 3. 12 Matriz para el movimiento predicho Donde cada fila corresponde a un vector de avance..

n. decisiones	time	elbow_flex	MCP2_flex	PIP_flex	DIP_flex
1	0	0	-10	-5	-5
2	0.02	0	-8	-3.1	-3.1
3	0.04	0	-6	-1.2	-1.2
4	0.06	0	-4	0.7	0.7
5	0.08	0	-2	2.6	2.6
6	0.1	0	0	4.5	4.5

El valor en cada vector de movimiento se va guardando y en caso de que haya un cambio de movimiento se toma el valor del nuevo vector de movimiento. Después de tener el modelo de clasificación representado en la Figura 3.11 para un proceso en tiempo real, el modelo tomará una decisión con la cual el vector del movimiento seleccionado será cargado en una nueva matriz de movimiento final (Tabla 3.12) como se muestra en la Figura 3.12. Esta última matriz (Tabla 3.12) matriz se guardará en un archivo ‘. mot’ y se cargará en Opensim para realizar la simulación y el análisis.

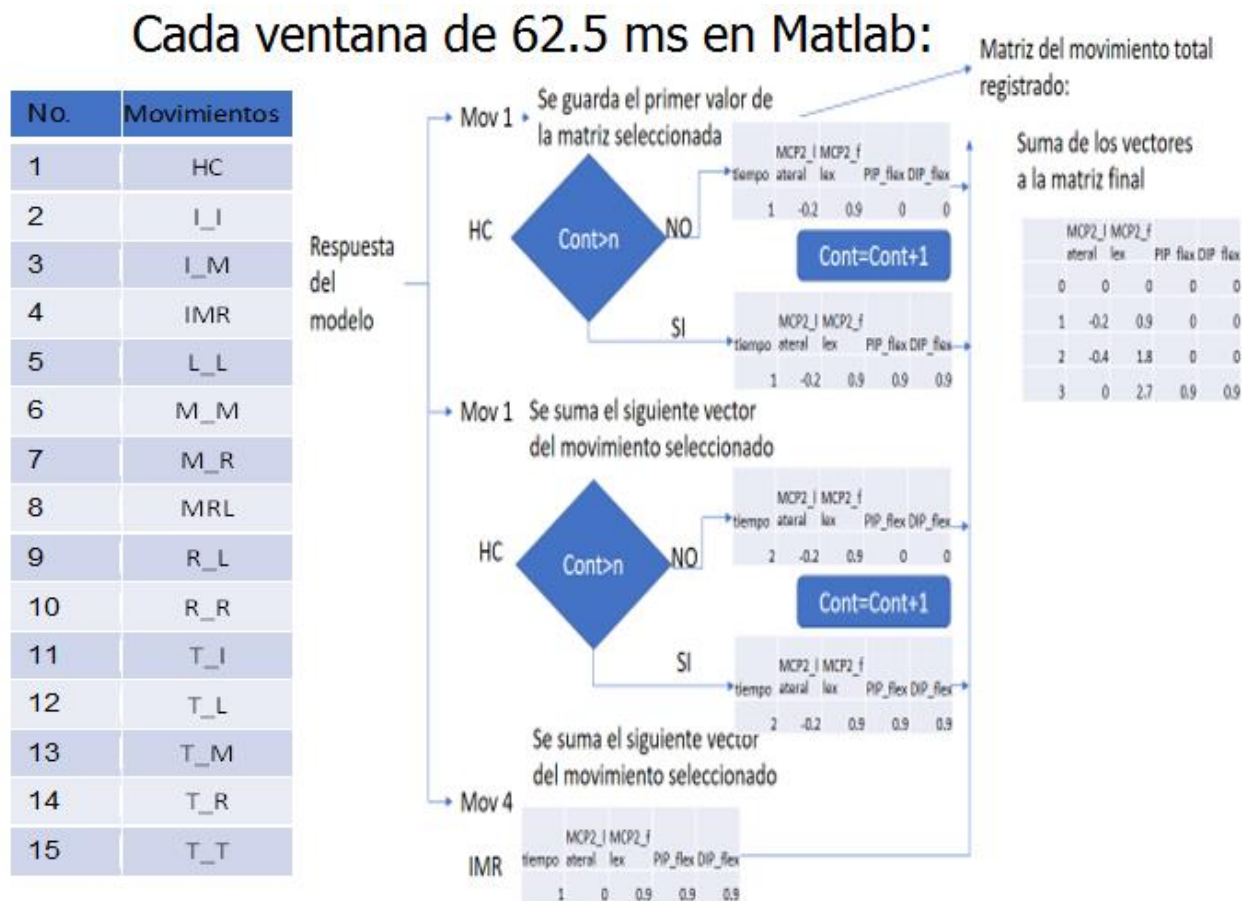


Figura 3. 12 Proceso general para la creación de la matriz final de movimiento en Matlab. El único movimiento que posee dos matrices es mano cerrada para poder controlar el movimiento del dedo pulgar contra los otros dedos. Es por eso que posee un contador, cuando el movimiento mano cerrada ha tenido un número mayor de 18 empieza a activarse el movimiento 2 donde el pulgar empieza su movimiento.

Capítulo 4 Resultados y discusión

Resultados

4.1 Eliminación de valores atípicos

4.1.1 Ruta 1: Eliminación de valores atípicos por movimientos separados

En la Tabla 4.1 se muestra la cantidad de muestras de la sEMG que fueron eliminadas, por ser consideradas atípicas, así como la cantidad de lecturas que se mantuvieron para medir las características de la señal. En las columnas de los sujetos 6 y 7 se muestran unas celdas marcadas porque coinciden exactamente los valores de muestras mantenidas y eliminadas de los sujetos, lo cual es imposible que pase en un experimento. Hay que recordar lo que en el capítulo anterior ya se expuso: esta ruta pertenece a la eliminación de valores atípicos por tipo de movimiento.

Tabla 4. 1 Valores atípicos eliminados y mantenidos por sujetos por separado (Ruta 1)

Mov.		Sujetos procesados outliers movimientos por separado							
		Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Sub 7	Sub 8
1.-HC	M	217272	223983	221437	216805	227971	226615	218416	218400
	E	22728	16017	18563	23195	12029	13385	21584	21600
2.-I_I	M	225721	224783	226875	221674	228287	219215 ⁵	219215	224286
	E	14279	15217	13125	18326	11713	20785	20785	15714
3.-I_M	M	223494	226483	222973	221258	228425	227648	226330	221529
	E	16506	13517	17027	18742	11575	12352	13670	18471
4.-IMR	M	225996	224761	223739	220023	226226	230395	223183	221584
	E	14004	15239	16261	19977	13774	9605	16817	18416
5.-L_L	M	222205	229107	225573	220311	227168	222203	222203	223210
	E	17795	10893	14427	19689	12832	17797	17797	16790
6.-M_M	M	222726	224247	228020	222556	226860	222770	222770	224417
	E	17274	15753	11980	17444	13140	17230	17230	15583
7.-M_R	M	223064	229080	222731	220829	230074	232135	221891	221199
	E	16936	10920	17269	19171	9926	7865	18109	18801
8.-MRL	M	221225	225863	223451	219925	226073	227340	217001	223160
	E	18775	14137	16549	20075	13927	12660	22999	16840
9.-R_L	M	224074	225811	222322	220211	225698	228040	220721	223940
	E	15926	14189	17678	19789	14302	11960	19279	16060
10.-R_R	M	222426	224814	227842	222112	225205	220823	220823	224564
	E	17574	15186	12158	17888	14795	19177	19177	15436

⁵ El color amarillo indica los valores que son idénticos en los sujetos 6 y 7.

11.-T_I	M	224407	222326	226173	220143	232664	220640	220640	226266
	E	15593	17674	13827	19857	7336	19360	19360	13734
12.-T_L	M	221658	226742	224295	221638	227826	228835	216129	224762
	E	18342	13258	15705	18362	12174	11165	23871	15238
13.-T_M	M	223189	223801	225026	219538	227477	222983	222983	223229
	E	16811	16199	14974	20462	12523	17017	17017	16771
14.-T_R	M	221250	223715	225320	222047	226148	223485	223485	222522
	E	18750	16285	14680	17953	13852	16515	16515	17478
15.-T_T	M	223347	224420	231549	219013	225431	218517	218517	222570
	E	16653	15580	8451	20987	14569	21483	21483	17430

4.1.2 Ruta 2: Eliminación de valores atípicos por movimientos en general

En la Tabla 4.2 se muestran de igual forma que la tabla anterior los datos que se consideraron atípicos y los que no, ahora con un enfoque distinto, se seleccionaron valores atípicos con respecto al movimiento en general.

Tabla 4. 2 Valores atípicos eliminados y mantenidos por sujetos por separado (Ruta 2)

Mov.		Sujetos procesados outliers movimientos por separado							
		Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Sub 7	Sub 8
1.-HC	M	161081	204907	123759	167387	205732	196329	159698	105875
	E	78919	35093	116241	72613	34268	43671	80302	134125
2.-I_I	M	239720	226955	238790	232981	236208	211648	227954	237890
	E	280	13045	1210	7019	3792	28352	12046	2110
3.-I_M	M	233065	238946	226278	231157	237744	239420	239851	236490
	E	6935	1054	13722	8843	2256	580	149	3510
4.-IMR	M	238416	220213	230995	215420	223053	239318	238008	236249
	E	1584	19787	9005	24580	16947	682	1992	3751
5.-L_L	M	229439	227970	234418	228828	223933	201330	231358	229822
	E	10561	12030	5582	11172	16067	38670	8642	10178
6.-M_M	M	239114	228727	236841	234033	237508	234847	238673	230973
	E	886	11273	3159	5967	2492	5153	1327	9027
7.-M_R	M	232121	237293	223812	232782	231256	239928	232055	237530
	E	7879	2707	16188	7218	8744	72	7945	2470
8.-MRL	M	204752	220243	219905	206450	228374	233628	213616	236009
	E	35248	19757	20095	33550	11626	6372	26384	3991
9.-R_L	M	211310	229956	219264	222096	193815	235708	196752	237456
	E	28690	10044	20736	17904	46185	4292	43248	2544
10.-R_R	M	238992	171898	237596	234402	234288	217591	236338	235641
	E	1008	68102	2404	5598	5712	22409	3662	4359
11.-T_I	M	236551	217929	236078	215915	237438	223674	227625	237863

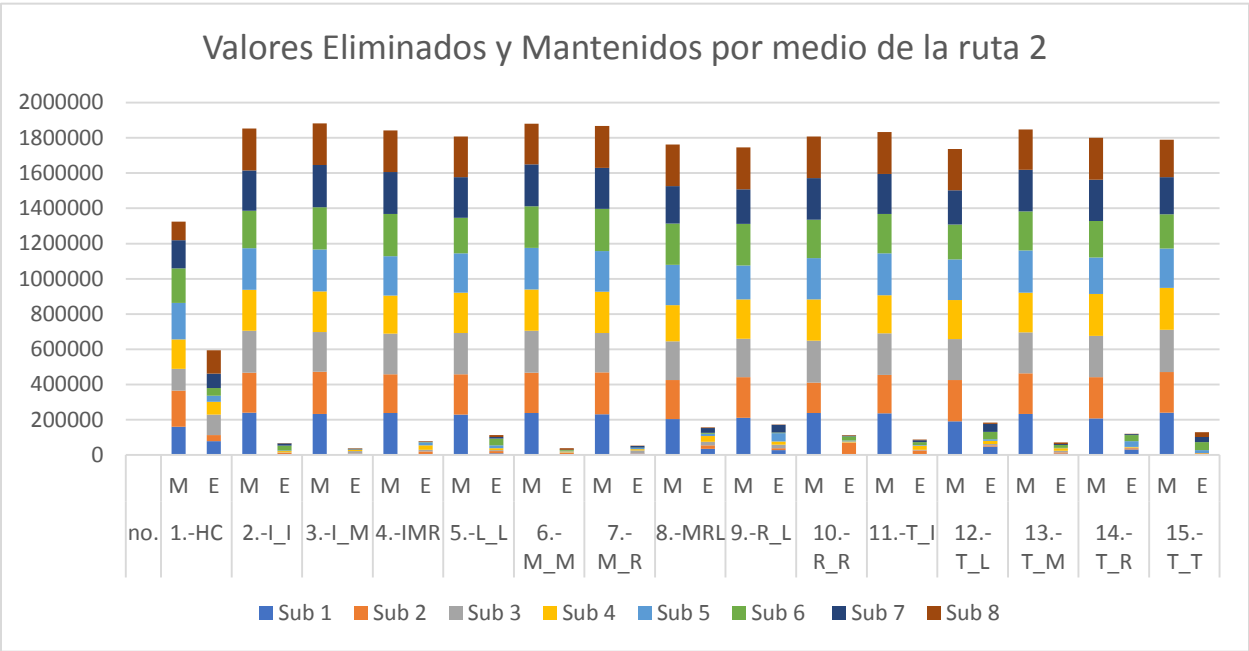


Figura 4. 2 Valores atípicos por medio de la Ruta 2, donde M son los valores mantenidos y E los eliminados.

4.2 Ventaneo

Después de la eliminación de valores atípicos por cualquiera de las dos rutas se procedió a separar en ventanas de datos para procesar la información. Una ventana se forma con la agrupación de 250 muestras (valores mantenidos).

4.2.1 Ventaneo Ruta 1

La Tabla 4.3 muestra los datos mantenidos junto con el número de ventanas que alcanzan con estos datos (número dentro del paréntesis).

Tabla 4. 3 Valores mantenidos después de la eliminación de valores atípicos por la Ruta 1

Mov.	Sujetos procesados outliers movimientos por separado ruta 1							
	1	2	3	4	5	6	7	8
1.-HC	217272 (870)	223983 (896)	221437 (886)	216805 (868)	227971 (912)	226615 (907)	218416 (874)	218400 (874)
2.-I_I	225721 (903)	224783 (900)	226875 (908)	221674 (887)	228287 (914)	219215 (877)	219215 (877)	224286 (898)
3.-I_M	223494 (894)	226483 (906)	222973 (892)	221258 (886)	228425 (914)	227648 (911)	226330 (906)	221529 (887)
4.-IMR	225996 (904)	224761 (900)	223739 (895)	220023 (881)	226226 (905)	230395 (922)	223183 (893)	221584 (887)
5.-L_L	222205 (889)	229107 (917)	225573 (903)	220311 (882)	227168 (909)	222203 (889)	222203 (889)	223210 (893)
6.-M_M	222726 (891)	224247 (897)	228020 (913)	222556 (891)	226860 (908)	222770 (892)	222770 (892)	224417 (898)
7.-M_R	223064 (893)	229080 (917)	222731 (891)	220829 (884)	230074 (921)	232135 (929)	221891 (888)	221199 (885)
8.-MRL	221225 (885)	225863 (904)	223451 (894)	219925 (880)	226073 (905)	227340 (910)	217001 (869)	223160 (893)
9.-R_L	224074 (897)	225811 (904)	222322 (890)	220211 (881)	225698 (903)	228040 (913)	220721 (883)	223940 (896)
10.-R_R	222426 (890)	224814 (900)	227842 (912)	222112 (889)	225205 (901)	220823 (884)	220823 (884)	224564 (899)
11.-T_I	224407 (898)	222326 (890)	226173 (905)	220143 (881)	232664 (931)	220640 (883)	220640 (883)	226266 (906)
12.-T_L	221658 (887)	226742 (907)	224295 (898)	221638 (887)	227826 (912)	228835 (916)	216129 (865)	224762 (900)
13.-T_M	223189 (893)	223801 (896)	225026 (901)	219538 (879)	227477 (910)	222983 (892)	222983 (892)	223229 (893)
14.-T_R	221250 (885)	223715 (895)	225320 (902)	222047 (889)	226148 (905)	223485 (894)	223485 (894)	222522 (891)
15.-T_T	223347 (894)	224420 (898)	231549 (927)	219013 (877)	225431 (902)	218517 (875)	218517 (875)	222570 (891)

La Tabla 4.3 ya se ha procesado lo que indica que se han eliminado los valores atípicos, eliminando una muestra-renglón.

Por lo que la matriz resultante puede ser de un tamaño menor a los 240,000 renglones.

El tamaño máximo de muestras resultante de la matriz, si no hubiera eliminación de valores atípicos para cada movimiento y sujetos, sería de 240,000 (960).

4.2.2 Ventaneo a través de la Ruta 2

La Tabla 4.4 muestra el número de ventanas que se obtuvieron al momento de dividir los datos crudos no atípicos para calcular las características seleccionadas.

Tabla 4. 4 Número de ventanas de cada uno de los 8 sujetos con cada uno de sus 15 movimientos. Los datos que se encuentran en amarillo muestran los datos que hubieran pasado desapercibidos si no se hubiese llevado a cabo la ruta 1.

Mov.	Sujetos procesados outliers movimientos en conjunto ruta 2							
	1	2	3	4	5	6	7	8
1.-HC	645	820	492	670	823	786	639	424
2.-I_I	959	908	956	932	945	847	912	952
3.-I_M	933	956	906	925	951	958	960	946
4.-IMR	954	881	924	862	893	958	953	945
5.-L_L	918	912	938	916	896	806	926	920
6.-M_M	957	915	948	937	951	940	955	924
7.-M_R	929	950	896	932	926	960	929	951
8.-MRL	820	881	880	826	914	935	855	945
9.-R_L	846	920	878	889	776	943	788	950
10.-R_R	956	688	951	938	938	871	946	943
11.-T_I	947	872	945	864	950	895	911	952
12.-T_L	766	934	935	885	923	791	780	935
13.-T_M	931	926	928	903	956	890	942	919
14.-T_R	834	935	941	949	830	825	937	952
15.-T_T	959	926	957	953	896	777	843	851

En la Tabla 4.5 mostramos un movimiento en mano cerrada y sus 12 repeticiones un ejemplo de cómo están divididos los archivos existentes para un movimiento y el cómo se distribuyen los datos atípicos contra los típicos.

Tabla 4. 5 sEMG sujetos movimiento mano cerrada (HC) para los 8 sujetos y sus 12 repeticiones matriz de Mov. completos ruta 2

no. muestra		1	2	3	4	5	6	7	8
1 a 20,000	M	4218	17052	8369	7096	16737	19404	10469	4297
	E	15782	2948	11631	12931	3263	596	9531	15707
21,000 a 40,000	M	10876	18273	10191	10721	16877	19526	12944	4023
	E	9124	1727	9809	9279	3123	474	7056	15977
40,001 a 60,000	M	15474	18267	8700	13296	17819	19425	12971	4728
	E	4526	1733	11300	6704	2181	575	7029	15272
60,001 a 80,000	M	16123	18280	12995	13498	17976	17380	15179	8334
	E	3877	1720	7005	6502	2024	2620	4821	11666
80,001 a 100,000	M	9993	13504	10034	13630	17247	14269	7619	8143
	E	10007	6496	9966	6370	2753	5731	12381	11857
100... a 120,000	M	16088	15311	11261	13570	17772	14718	13406	7478
	E	3912	4689	8739	6430	2228	5282	6594	12522
120... a 140,000	M	16066	16128	9191	15475	17407	14569	15112	8400
	E	3934	3872	10809	4525	2593	5431	4888	11600
140... a 160,000	M	11968	17585	10920	15880	17620	14376	15953	9823
	E	8032	2415	9080	4120	2380	5624	4047	10177
160... a 180,000	M	10828	16201	8245	13518	17086	15975	7924	7424
	E	9172	3799	11755	6482	2914	4025	12076	12576
180... a 20,000	M	15807	17838	10085	16456	16742	15505	14352	11658
	E	4193	2162	9915	3544	3258	4495	5648	8342
200... a 220,000	M	16830	18117	11063	16936	15975	15605	16305	15775
	E	3170	1883	8937	3064	4025	4395	3695	4225
220,001 a 240,000	M	16810	18351	12705	17338	16474	15577	17464	15792
	E	3190	1649	7295	2662	3526	4423	2536	4208

En los movimientos están incluidas las 12 repeticiones. El tamaño máximo de renglones posible sería de 240,000 si no se elimina ningún valor atípico. Además, si se utilizan ventanas de medición de 250 muestras, tendríamos a lo más 960 ventanas para medir características en cada sujeto por canal.

4.3 Clasificación

En esta sección se muestran los resultados de la sección 3.4 del capítulo de experimentación

4.3.2 Resultados de la experimentación inicial a través de la Ruta 2

En la Tabla 4.6 podemos observar los porcentajes de reconocimiento obtenidos de cada uno de los modelos creados a partir de NB, QDA y LDA con respecto a la experimentación de la figura 3.1, con cada uno de los sujetos y con el grupo muestral de los 7 sujetos. Podemos observar como la tabla está dividida en dos secciones la parte de los modelos generados con los datos de entrenamiento calculando su respectivo porcentaje con validación cruzada (**% REC TRN cross-val**) y la parte donde están los mismos modelos evaluados con los datos de prueba (%Rec de prueba). Podemos observar también que con el grupo muestral el porcentaje de reconocimiento se reduce.

Tabla 4. 6 Resultados obtenidos a partir de la experimentación 1 mostrada en la Figura 3.1.

	% REC TRN cross-val			% Rec de prueba		
	LDA	QDA	N. Bayes	LDA	QDA	N. Bayes
Sujeto 1	94.19	91.02	84.99	94.49	90.59	84.33
Sujeto 2	93.71	92.16	89.95	93.53	92.28	89.78
Sujeto 3	88.03	79.28	72.74	87.44	79.31	71.95
Sujeto 4	86.22	80.32	74.98	85.92	79.43	74.44
Sujeto 5	97.23	96.39	93.39	97.26	96.34	92.84
Sujeto 6	94.48	92.22	89.27	94.54	92.63	89.53
Sujeto 7	93.37	90.67	87.44	93.49	90.57	88.10
7 sujetos	60.14	54.93	29.40	59.81	55.04	29.76

La Tabla 4.7 muestra los resultados con las clases ajustadas a 105 y con las clases reajustadas a 15 como se detalló en la sección 3.11. Recordando que se decidió ajustar el número de clases del grupo muestral dándole clases distintas a cada uno de los sujetos evaluados dentro del modelo para el grupo muestral, es por eso por lo que las clases ajustadas están siendo evaluadas directamente sobre el modelo creado, en cambio la fila de clases reajustadas indica que las decisiones arrojadas por el modelo creado con las 105 clases se cambian a las 15 originales y se evalúa el error.

Tabla 4. 7 Resultados obtenidos a partir de la experimentacion 1 en conjunto con el postprocesamiento donde las clases son ajustadas de 15 a 105 (tecnica de extension de clases). El modelo inicial para los 7 sujetos (fila 1) discrimina para las 105 clases.

	TRN			TST		
	LDA	QDA	N. Bayes	LDA	QDA	N. Bayes
7 sujetos Clases ajustadas	82.58	87.13	81.36	82.76	87.32	81.60
Clases reajustadas				83.31	88.09	84.48

4.4 Selección de atributos hacia adelante

En esta sección se muestran los resultados de la experimentación de la sección número 3.6 llamada Selección de atributos hacia adelante. En la Tabla 4.8 podemos observar los resultados de la sección de datos de entrenamiento a través de validación cruzada como se mostró en la Figura 3.2, en la Tabla 4.9 podemos observar cuantas características se mantuvieron para cada sujeto por separado o para el grupo muestral de 7 sujetos en cada uno de los modelos evaluados.

4.4.1 Resultados de la Experimentación 2 a través de la Ruta 2

En la Tabla 4.8 podemos observar que solo existen resultados para los datos de entrenamiento para los modelos QDA y LDA perteneciente a la experimentación de la Figura 3.2 ya que en el experimento anterior se observó que el tiempo de procesamiento para el modelo creado a partir de Naive Bayes era muy grande en comparación con los otros modelos además su porcentaje de clasificación no valía el tiempo requerido para llevarlo a cabo. Además, como el procedimiento de SFS es iterativo lo que quiere decir que se evaluarán cada uno de los subconjuntos creados a partir del pseudocódigo mostrado en la Tabla 3.10 de la sección 3.6 lo que indica que se crearán N números de subconjuntos de variables para crear N números de modelos de Naive Bayes lo que incrementa N veces el tiempo de procesamiento requerido. De igual forma que en el experimento anterior hay una caída en el porcentaje de reconocimiento del grupo muestral de los 7 sujetos, pero en menor medida.

En la Tabla 4.9 podemos observar el número de características del ultimo subconjunto evaluado para cada sujeto y cada algoritmo (QDA y LDA). El total de características evaluadas eran 96 de las cuales quedaron únicamente las indicadas.

Tabla 4. 8 Porcentajes de los modelos creados con cada uno de los algoritmos utilizados.

	% REC TRN cross-val	
	LDA	QDA
Sujeto 1	94.37	96.17
Sujeto 2	93.53	96.46
Sujeto 3	88.33	89.64
Sujeto 4	85.39	88.91
Sujeto 5	97.53	98.44
Sujeto 6	94.11	97.56
Sujeto 7	93.08	95.50
7 sujetos	60.32	80.54

Tabla 4. 9 Número de características mantenidas en cada uno de los algoritmos para cada sujeto.

Clasificador	Características mantenidas								
	Sujeto 1	Sujeto 2	Sujeto 3	Sujeto 4	Sujeto 5	Sujeto 6	Sujeto 7	Sujeto 8	Suj 1-7
LDA	49	37	36	41	43	46	47	40	77
QDA	23	22	22	26	22	14	22	26	49

En la Tabla 4.10 podemos encontrar una comparación entre el modelo con las clases ajustadas a 105 y el modelo con las clases reajustadas a 15 como se mostró en la sección 3.11.

Tabla 4. 10 Comparación entre el modelo normal y el modelo con clases ajustadas.

	TRN			TST		
	LDA	QDA	N. Bayes	LDA	QDA	N. Bayes

7 sujetos		89.71			89.61	
Clases ajustadas		91.41			91.37	

En la Tabla 4.11 y 4.12 podemos encontrar cada una de las características evaluadas en la selección secuencial de atributos para el modelo lineal (Modelo obtenido a partir de LDA) y el modelo cuadrático (Modelo obtenido a partir de QDA) respectivamente, observamos que para el modelo cuadrático las características empiezan a ser eliminadas en un mayor número. Además, en estas tablas encontramos la relación entre las características evaluadas, los sujetos y la posición de los electrodos, se puede observar en que algoritmo son más útiles ciertas características o sea que características ayudan al modelo de clasificación (LDA o QDA) y cuáles no. En estas tablas podemos observar cuales fueron las características que se mantuvieron (cuadro azul con asterisco) cada renglón representa un sujeto y una característica, las columnas representan los electrodos a los que se le aplicaran las operaciones indicadas en la fila (indicadas en la última columna a la derecha), cada operación posee un bloque de sujetos del 1 al 7, podemos ver estos bloques de características en la última columna a la derecha de las tablas. Los bloques amarillos son bloques que podrían ser eliminados ya que poseen un número muy bajo de operaciones que se mantuvieron con respecto a los otros en la Tabla 4.11 solo 6 operaciones se mantuvieron para la característica de VAR, por otro lado tenemos la tabla 4.12 donde el bloque amarillo solo posee 3 operaciones que se mantuvieron en esta tabla dos de esas operaciones pertenecen a un mismo sujeto, también hay bloques rojos donde solo permanecen 2 operaciones o ninguna, estas dos operaciones vemos que pertenecen a sujetos distintos y electrodos distintos, por lo que estas características pueden ser eliminadas.

Tabla 4. 11 LDA Características Eliminadas de cada uno de los sujetos y posición de electrodo a la que pertenecen (Ruta 2).

	1	2	3	4	5	6	7	8		
s1	*	*	*	*		*			MAV	
s2	*		*	*		*	*	*		
s3						*	*	*		
s4	*		*	*	*	*	*	*		
s5	*	*	*		*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1				*	*	*	*	*	MSV	
s2			*		*	*	*			
s3		*			*	*	*			
s4		*					*			
s5	*	*			*	*	*	*		
s6	*	*			*	*	*	*		
s7					*	*	*	*		
s8						*	*	*		
s1	*	*	*	*	*	*	*	*	SSI	
s2	*	*	*	*	*	*	*	*		
s3		*	*	*	*	*	*	*		
s4		*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1					*	*	*	*	RMS	
s2				*	*	*	*	*		
s3				*	*	*	*	*		
s4				*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1					*	*	*	*	VAR	
s2				*	*	*	*	*		
s3				*	*	*	*	*		
s4				*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1	*	*	*	*	*	*	*	*	TM3	
s2	*	*	*	*	*	*	*	*		
s3	*	*	*	*	*	*	*	*		
s4	*	*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1	*	*	*	*	*	*	*	*	TM4	
s2	*	*	*	*	*	*	*	*		
s3	*	*	*	*	*	*	*	*		
s4	*	*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1	*	*	*	*	*	*	*	*	TM5	
s2	*	*	*	*	*	*	*	*		
s3	*	*	*	*	*	*	*	*		
s4	*	*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1	*	*	*	*	*	*	*	*	WL	
s2	*	*	*	*	*	*	*	*		
s3	*	*	*	*	*	*	*	*		
s4	*	*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1	*	*	*	*	*	*	*	*	NZC	
s2	*	*	*	*	*	*	*	*		
s3	*	*	*	*	*	*	*	*		
s4	*	*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1	*	*	*	*	*	*	*	*	MPR1	
s2	*	*	*	*	*	*	*	*		
s3	*	*	*	*	*	*	*	*		
s4	*	*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		
s1	*	*	*	*	*	*	*	*	SSC1	
s2	*	*	*	*	*	*	*	*		
s3	*	*	*	*	*	*	*	*		
s4	*	*	*	*	*	*	*	*		
s5	*	*	*	*	*	*	*	*		
s6	*	*	*	*	*	*	*	*		
s7	*	*	*	*	*	*	*	*		
s8	*	*	*	*	*	*	*	*		

Tabla 4. 12 QDA Características Eliminadas de cada uno de los sujetos y posición de electrodo a la que pertenecen (Ruta 2).

	1	2	3	4	5	6	7	8			
s1			*		*	*	*		MAV		
s2			*				*	*			
s3		*		*	*	*	*	*			
s4			*	*	*	*		*			
s5							*	*			
s6	*				*	*		*			
s7					*	*	*				
s8					*	*		*			
s1									MSV		
s2											
s3											
s4											
s5											
s6											
s7											
s8											
s1									SSI		
s2	*	*		*	*	*	*	*			
s3	*	*		*	*	*					
s4		*	*	*	*		*				
s5	*	*	*	*	*						
s6		*	*	*	*		*				
s7	*			*	*			*			
s8	*		*	*	*	*	*				
s1									RMS		
s2											
s3											
s4											
s5											
s6											
s7											
s8											
s1									VAR		
s2	*										
s3											
s4											
s5											
s6											
s7											
s8											
s1									TM3		
s2											
s3											
s4											
s5											
s6											
s7											
s8											
s1									TM4		
s2											
s3											
s4											
s5											
s6											
s7											
s8											
s1									TM5		
s2	*	*	*	*	*	*		*			
s3	*	*	*	*	*	*	*	*			
s4	*	*	*	*	*	*	*	*			
s5	*	*	*	*	*	*	*	*			
s6	*	*			*	*		*			
s7	*	*	*	*	*	*	*	*			
s8	*	*	*	*	*	*	*	*			
s1				*			*		WL		
s2											
s3				*	*		*				
s4					*		*				
s5							*				
s6						*					
s7						*	*	*			
s8	*				*	*	*	*			
s1	*				*			*	Nzc		
s2		*			*			*			
s3	*						*	*			
s4	*			*	*		*	*			
s5			*		*						
s6	*	*			*	*	*	*			
s7	*	*			*	*	*	*			
s8	*	*		*	*	*	*	*			
s1	*	*		*	*	*	*	*	MPR1		
s2	*				*	*	*	*			
s3			*	*	*	*	*				
s4	*	*			*		*				
s5	*						*				
s6	*				*						
s7	*				*	*		*			
s8	*	*			*	*		*			
s1	*	*		*	*	*	*	*	SSC1		
s2	*				*	*	*	*			
s3			*	*	*	*	*				
s4	*	*			*		*				
s5	*						*				
s6	*				*						
s7	*				*	*		*			
s8	*	*			*	*		*			

4.5 Algoritmo de matrices dispersas

Como se presentó en la sección 3.8 este algoritmo se utilizó para mantener toda la información posible con la menor cantidad de variables, es decir la información más representativa para diferenciar las clases en un mínimo de variables (transformadas). En las Figuras 4.3 y 4.4 mostraremos como con tan solo 3 de las 14 seleccionadas las clases empiezan a concentrarse en regiones lo que las hace fácil la predicción de nuevos valores.

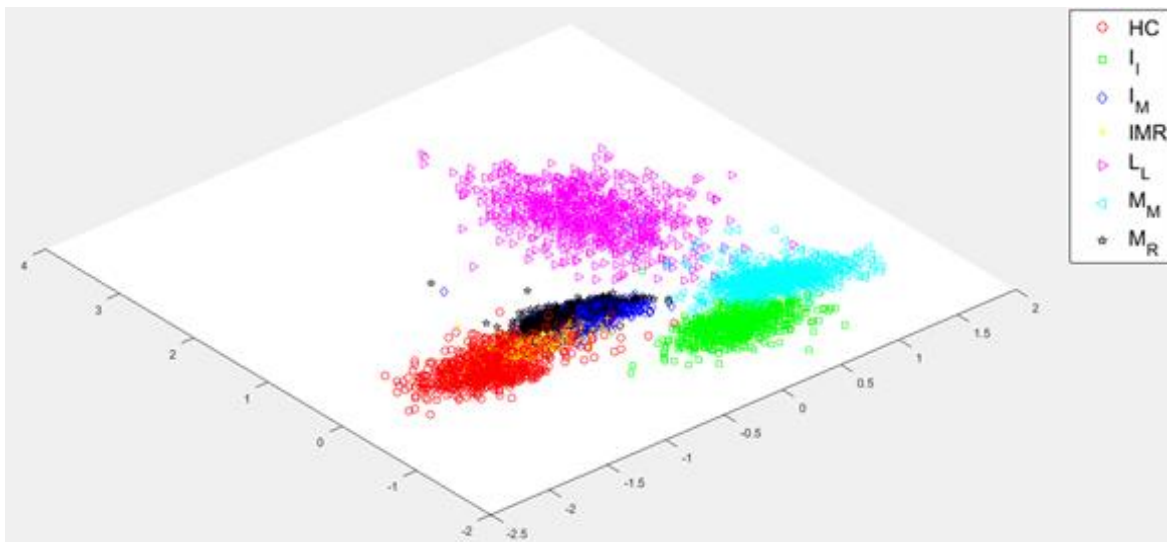


Figura 4. 3 Visualización de los movimientos evaluados con solo 3 características en el nuevo espacio de las 14 originales para el sujeto número 6 después de SFS. Orientación donde es más visible la separación de las clases.

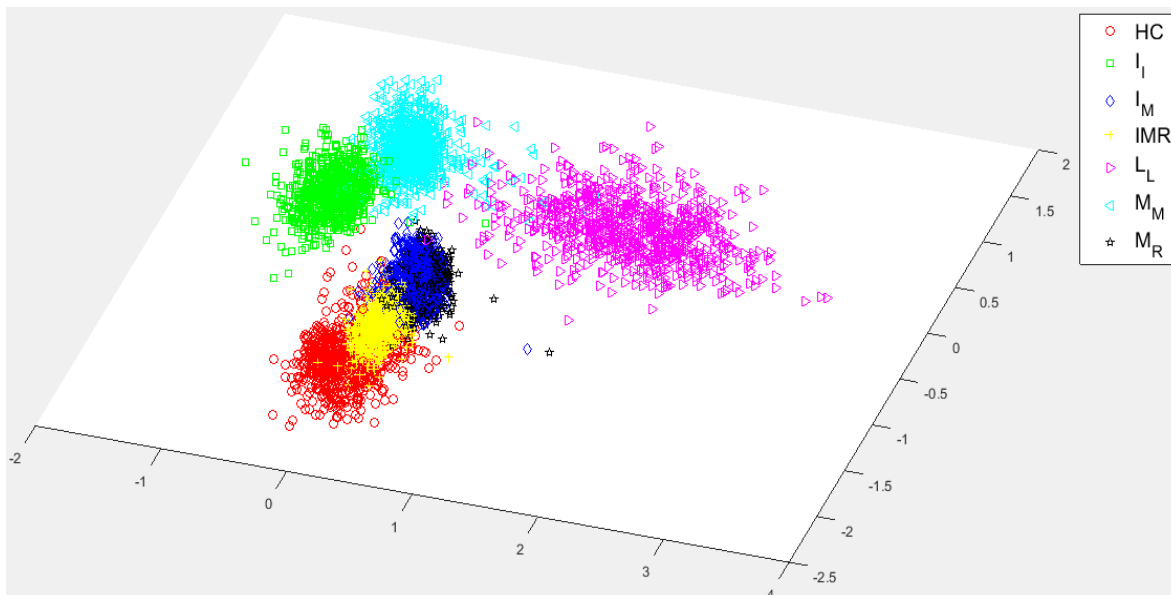


Figura 4. 4 Visualización de 8 movimientos evaluados con solo las primeras 3 características en el nuevo espacio de las 14 originales para el sujeto 6 después de SFS. Orientación donde es más visible la separación de las clases.

La Tabla 4.13 muestra una comparación entre el número de características seleccionadas por SFS y un número de características reducido a solo las 3 primeras del nuevo espacio junto con sus porcentajes de reconocimiento; se puede observar que con tan solo 3 se obtiene un porcentaje cercano al número de características seleccionadas.

Tabla 4. 13 Tabla comparativa del porcentaje de reconocimiento del espacio original contra el nuevo espacio con únicamente 3 características

	Número de características	características (espacio original)	Con los primeros 3 características del nuevo espacio
Sujeto 1	23	96.17	72.26
Sujeto 2	22	96.46	71.09
Sujeto 3	22	89.64	67.21
Sujeto 4	26	88.91	65.87
Sujeto 5	22	98.44	81.93
Sujeto 6	14	97.56	80.08
Sujeto 7	22	95.50	79.16
Sujeto 8	26	91.28	66.30
S 1-7	49	89.61	31.03
S 1-7 Mod	49	91.37	33.39

Como se puede observar en la Tabla 4.13 los porcentajes de reconocimiento de los grupos muestrales es del 31.03 y del 33.39 que son muy bajos lo que indica que 3 características (en el nuevo espacio) es muy poco para 49 de las seleccionadas (después de SFS) por lo tanto se probó aumentando el número de características y se creó la Tabla 4.14 que muestra una comparación entre un número incremental de características y los porcentajes de reconocimiento.

Tabla 4. 14 Tabla comparativa del grupo muestral 1-7 donde se evalúa el porcentaje de reconocimiento frente al número de características en el nuevo espacio, la columna de prueba es el resultado obtenido con la evaluación de las 105 clases, la columna de prueba con clases modificadas es después de haber reducido a 15 clases las 105 y haber calculado el error.

Características	Entrenamiento	Prueba	Prueba clases modificadas
3	31.26	31.03	33.39
6	60.68	60.87	64.17
8	72.76	73.05	75.82
10	80.07	79.97	82.64
12	83.98	84.07	86.58
14	86.57	86.49	88.98

16	87.35	87.31	89.7
18	88.04	87.88	90.2
20	88.61	88.39	90.52

La Figura 4.5 muestra la gráfica de la Tabla 4.14 donde vemos que con tan solo 20 características (en el nuevo espacio) de las 49 obtenidas a partir de SFS sin transformarlas obtenemos la misma información necesaria para obtener un porcentaje semejante.

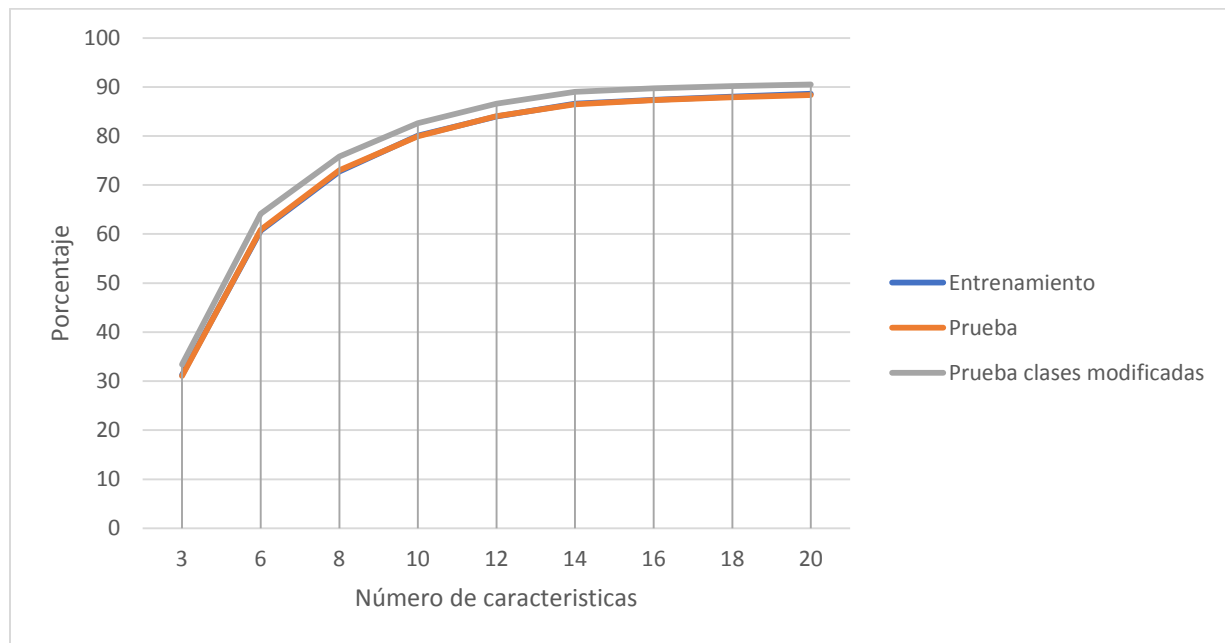


Figura 4. 5 Grafica resultante de la Tabla 4.22.

4.6 Opensim

4.6.1 Articulaciones del modelo creado

Como se mencionó en la sección 2.7.2.1.3 el modelo musculoesquelético original constaba de 8 articulaciones con 10 grados de libertad en los dedos de la mano que le permitían realizar flexión y aducción-abducción del pulgar y flexión del dedo índice. Con la adición de 11 articulaciones más, el modelo musculoesquelético está capacitado para realizar todas aquellas combinaciones de los dedos pulgar, índice, medio, anular y meñique. Y con lo que podemos convertir los 15 movimientos de la base de datos en movimientos bajo la plataforma de Opensim. Los modelos sin la musculatura se pueden observar en las Figuras 4.6 y 4.7.

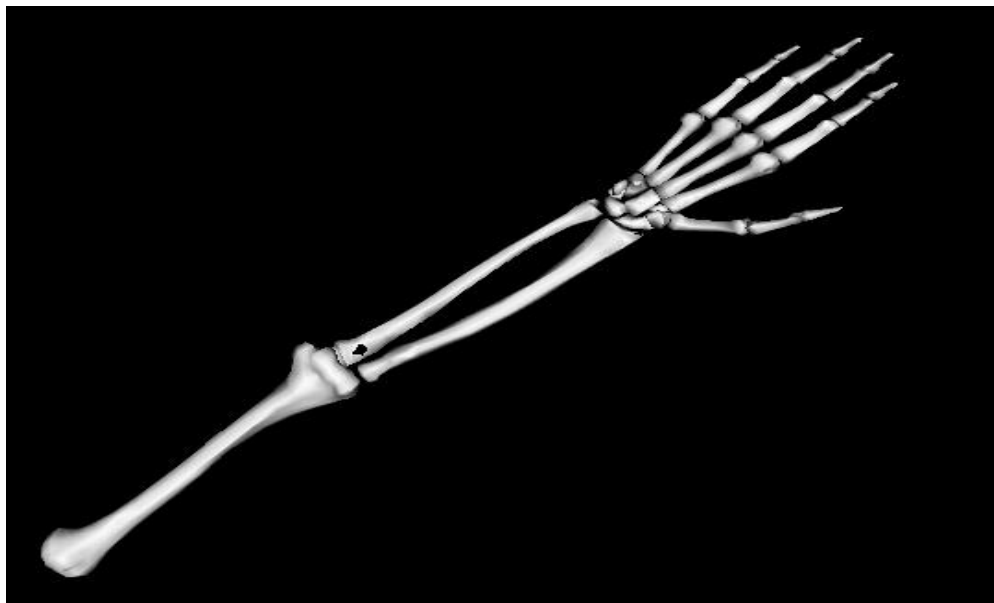


Figura 4. 6 Modelo Original de la muñeca.

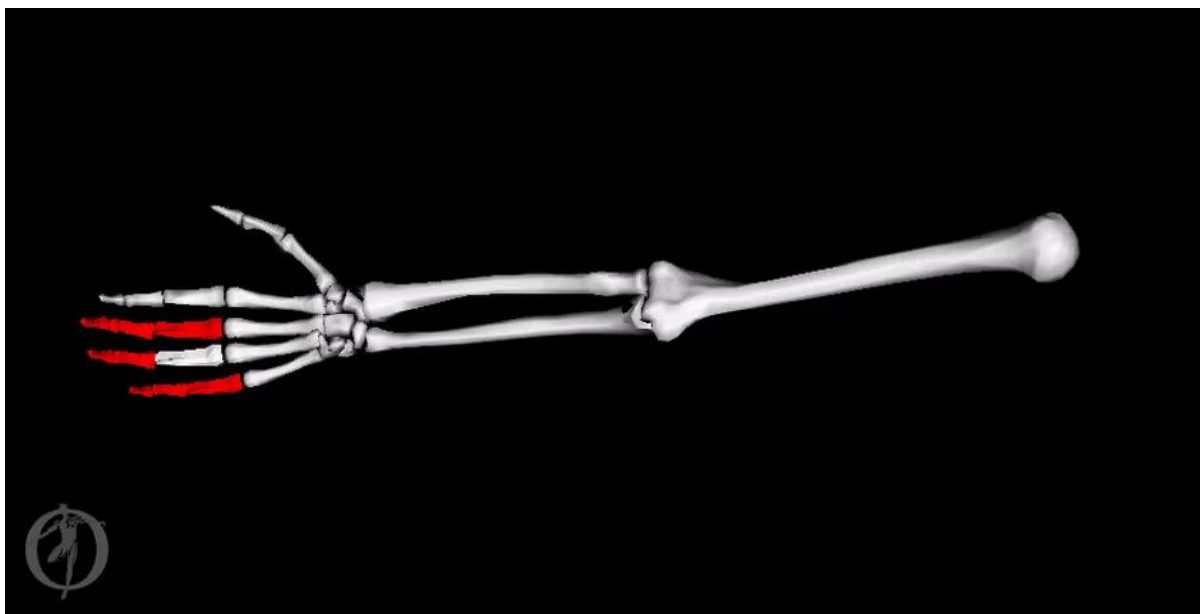


Figura 4. 7 Modelo modificado de la muñeca, los dedos de la mano modificados poseen un color distinto, característica que puede ser modificada al gusto.

4.6.3 Movimientos capaces de realizar

La Tabla 4.15 muestra la comparación entre los movimientos del modelo original y sus grados de libertad y los del modelo modificado.

Tabla 4. 15 Comparación entre movimientos-Grados de libertad (Gdl) entre el modelo original y el modificado

Movimiento modelo original	Movimiento modelo modificado
Flexión dedo índice (4 Gdl)	Flexión dedo índice (4 Gdl)
Flexión dedo pulgar (3 Gdl)	Flexión dedo pulgar (3 Gdl)
Abd/Adu pulgar (1 Gdl)	Abd/Adu pulgar (1 Gdl)
	Flexión dedo medio (4 Gdl)
	Flexión dedo anular (4 Gdl)
	Flexión dedo meñique (4 Gdl)

La Tabla 4.16 muestra los movimientos que evaluaríamos en el modelo de clasificación, aunque este es capaz de realizar una infinidad de estos, solo nos dedicamos a reproducir los que estábamos evaluando.

Tabla 4. 16 Movimientos utilizados.

Movimiento	
Mano cerrada	Flexión pulgar meñique
Flexión dedo índice	Flexión pulgar índice
Flexión dedo medio	Flexión índice medio anular
Flexión dedo pulgar	Flexión medio anular meñique
Flexión dedo anular	Flexión medio anular
Flexión dedo meñique	Flexión índice medio
Flexión pulgar índice	Flexión anular meñique
Flexión pulgar medio	Flexión pulgar anular

4.6.4 Movimientos reproducidos

Acontinuacion en las Figuras 4.8, 4.9, 4.10 y 4.11 se muestran algunos de los movimientos reproducidos con Opensim y su representnacion fisica.



Figura 4. 8 Movimiento Flexión Índice



Figura 4. 9 Movimiento Flexión índice-medio



Figura 4. 10 Movimiento Flexión índice-medio-anular.

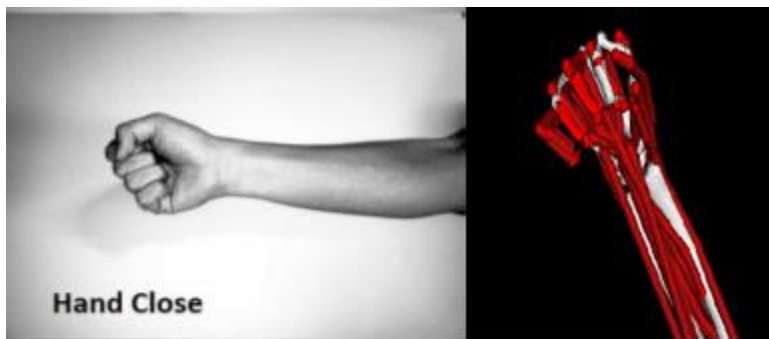


Figura 4. 11 Movimiento Mano cerrada.

4.6.5 Matrices configuradas para cada movimiento en Matlab

Como se describió en la sección 3.12.4, se desarrollaron las matrices (Tablas de la 4.17 a la 4.31) de los movimientos completos, se obtuvieron los siguientes vectores de suma para cada uno de los movimientos. Con estos vectores se realiza la creación de los videos de movimiento del modelo en Opensim, los videos constan de 80 cuadros de movimiento que viene de 5 seg/0.062 seg, como máximo ya que hubo datos eliminados. Cada vector de avance se suma a la matriz final por lo que los ceros de algunas articulaciones no interfieren en la definición del movimiento final.

Tabla 4. 17 vector de suma para mano cerrada cuando cont < 18

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
-0.3225	1.4516	1.4516	1.4516	1.4516	1.4516	1.4516	0.0483	1.4516	1.4516	1.4516	0.3225	1.4516	1.4516	1.4516	0	0	0	0	0

Tabla 4. 18 Vector de suma para mano cerrada cuando cont > 18.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
-0.3225	1.4516	1.4516	1.4516	1.4516	1.4516	1.4516	0.0483	1.4516	1.4516	1.4516	0.3225	1.4516	1.4516	1.4516	0	0.15909	1.43181	-0.45454	0.22727

Tabla 4. 19 Vector de suma para flexión del dedo índice.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
-0.3225	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Tabla 4. 20 Vector de suma para flexión del dedo medio.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0	0	0	0	0

Tabla 4. 21 Vector de suma para flexión del dedo anular.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	0	0	0	0.0483	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0

Tabla 4. 22 Vector de suma para flexión del dedo meñique.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	0	0	0	0	0	0	0	0.3225	1.4516	1.4516	1.4516	0	0	0	0	0

Tabla 4. 23 Vector de suma para flexión del dedo pulgar-índice.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
-0.3225	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0	0	0	0.15909	1.43181	-0.45454	0.22727	0.15909

Tabla 4. 24 Vector de suma para flexión del dedo pulgar-medio.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0.15909	1.43181	-0.45454	0.22727	0.15909

Tabla 4. 25 Vector de suma para flexión del dedo pulgar-anular.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	0	0	0	0.0483	1.4516	1.4516	1.4516	0	0	0	0	0.15909	1.43181	-0.45454	0.22727	0.15909

Tabla 4. 26 Vector de suma para flexión del dedo pulgar-meñique.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	0	0	0	0	0	0	0	0.3225	1.4516	1.4516	1.4516	0.15909	1.43181	-0.45454	0.22727	0.15909

Tabla 4. 27 Vector de suma para flexión del dedo Índice-medio.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
-0.3225	1.4516	1.4516	1.4516	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0	0	0	0	0

Tabla 4. 28 Vector de suma para flexión del dedo medio-anular.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	1.4516	1.4516	1.4516	0.0483	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0

Tabla 4. 29 Vector de suma para flexión del dedo anular-meñique.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	0	0	0	0.0483	1.4516	1.4516	1.4516	0.3225	1.4516	1.4516	1.4516	0	0	0	0	0

Tabla 4. 30 Vector de suma para flexión del dedo índice-medio-anular.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
-0.3225	1.4516	1.4516	1.4516	1.4516	1.4516	1.4516	0.0483	1.4516	1.4516	1.4516	0	0	0	0	0	0	0	0	0

Tabla 4. 31 Vector de suma para flexión del dedo medio-anular-meñique.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	1.4516	1.4516	1.4516	0.0483	1.4516	1.4516	1.4516	0.3225	1.4516	1.4516	1.4516	0	0	0	0	0

Tabla 4. 32 Vector de suma para flexión del dedo pulgar.

MCP2_l	MCP2_f	PIP_flex	DIP_f	MCP2M_f	MPIP_f	MDIP_f	RCP2M_l	RCP2M_f	RPIP_f	RDIP_f	LCP2M_l	LCP2M_f	LPIP_f	LDIP_f	thumb_a	thumb_f	TCP2M_l	TCP2M_f	TCP2M2_f
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.15909	1.43181	-0.45454	0.22727	0.15909

4.6.6 Operaciones y tiempo de procesamiento

Después de poseer todos los elementos nuevo vector de entrada, matrices se procedió a calcular el tiempo de procesamiento de todas las operaciones que deben hacerse a un nuevo vector de entrada en la Tabla 4.32 se muestra cada una de las etapas y el tiempo requerido de Matlab.

Tabla 4. 32 Operaciones contra tiempo de procesamiento, la sección marcada en rojo indica el tiempo requerido para formar una ventana de 250 datos para iniciar el procesamiento, la amarilla es el tiempo de cada operación realizada a la ventana de 250 datos, las franjas negras indican las operaciones suprimidas porque no ayudan al modelo, la naranja el procesamiento de la transformación al nuevo espacio, el azul a la evaluación del modelo creado y por ultimo a la asignación de la matriz según el resultado del modelo de clasificación para el movimiento del modelo virtual de mano. Donde VM=Vector de medias de datos crudos, VC=Valor crítico, SC=subconjunto, VN=Vector nuevo de datos crudos, M=Vector medias de nuevas características, W=Matriz de blanqueo, bhat=Matriz de ortogonalización, Matriz de movimiento final=ceros.

Operaciones	Tiempo de procesamiento (Seg)
Vector de datos crudos nuevo	
$V12 = VN - VM$	0.010
$D1 = s \setminus V12$	0.012
$DM = V12 * D1$	0.014
Comparación DM vs VC	0.014
Ventana completa=250 valores	Hasta conseguir 250 valores para una ventana
MAV (por 8 electrodos)	0.020
MSV (por 8 electrodos)	0.014
SSI (por 8 electrodos)	0.014
RMS (por 8 electrodos)	0.015
VAR (por 8 electrodos)	0.015
TM3 (por 8 electrodos)	0.016
TM4 (por 8 electrodos)	0.017
TM5 (por 8 electrodos)	0.017
WL (por 8 electrodos)	0.025
NZC (por 8 electrodos)	0.021
MPR (por 8 electrodos)	0.024
SSC (por 8 electrodos)	0.024
Vector nuevas caract(96)	
$Xprueba1 = \text{Vector nuevas caract}(SC)$	
$Xn = Xprueba1 - M$	0.012
$Xw = Xn * W$	0.012
$Xc1 = Xw * bhat' * bhat$	0.012
$\text{Proyeccion_}x1 = Xc1 * E_sort$	0.015
$\text{Resultado} = \text{mldQDA}(\text{Proyeccion_}x1)$	0.025
Switch case para seleccionar vector y guardar dato	0.011

Discusión

Como parte de este trabajo de tesis, y como es conocido, se puede observar lo difícil de relacionar una señal superficial del sistema musculo esquelético con un movimiento puntual. Resulta difícil descifrar el código que de la acción de los músculos y de sus miles de fibras aparece en las mediciones de un sEMG. Un individuo puede realizar movimientos iguales, sin embargo, el cómo se activa el movimiento dará como resultado ciertas combinaciones de fibras activándose que se traducirá como una señal semejante pero no igual. Los movimientos pueden ser tan diversos si se varía su grado de flexión, su fuerza y su combinación con otras estructuras, cada persona poseerá señales de distinta magnitud para los movimientos evaluados, es decir la fisionomía de una persona afecta las señales. Los seres humanos no ponemos atención a la cantidad de movimientos que realizamos, todo es automático hacia un objetivo que nos interesa. Consideramos y demostramos que un modelo que clasifique un arreglo pequeño de electrodos puede entender esta codificación y descifrar el tipo de movimiento asociado, con un buen grado de exactitud. La propuesta para determinar los movimientos de la mano utilizando electrodos en el antebrazo puede tener una utilidad en los sistemas de control de una prótesis multifuncional o para aplicaciones de tiempo-real basada en reconocer gestos con la mano. Se utilizó una base de datos ensayada por otros investigadores y que contenía varios movimientos de los dedos de la mano, esto con el fin de que en la tesis nos enfocáramos sólo al modelo de clasificador, aunque la experimentación y resultados nos revelan aspectos sobre la calidad de la base de datos. En este capítulo haremos una revisión exhaustiva de los resultados tratando de seguir la secuencia temática de dicha sección.

La estructura de la base de datos comprendía la señal de 8 electrodos posicionados en el antebrazo, evaluando 15 movimientos, entre flexiones aisladas de los dedos, así como la combinación de flexiones de los dedos. Los patrones de las mediciones se parecerán ya que, por ejemplo, el movimiento flexión de dedo índice, lo comparamos con la combinación flexión de dedo índice y pulgar, las mediciones de flexión del dedo índice estarán inmiscuidas dentro de las mediciones de la combinación índice-pulgar y las hará un tanto parecidas. Ahora si evaluamos la flexión del menique y el anular, son dos movimientos que prácticamente utilizan los mismos músculos para flexionarse, sus mediciones también serán muy parecidas.

Sobre las pruebas iniciales. La tarea de clasificación de movimientos con los datos crudos de la base de datos carece de importancia ya que, si se establece un número pequeño de movimientos tendrás buenos resultados, sin embargo, esto ya es obsoleto, y si estableces un número más alto de movimientos tu porcentaje de clasificación correcta empieza a caer, lo que indicaría que los movimientos cada vez se van diferenciando menos aumentando la complejidad del problema. El tratar de encontrar un modelo que describa el comportamiento de los datos crudos ante más de cinco movimientos es difícil. En nuestro caso, el tratar de utilizar los datos crudos de una base de datos que utilizó un arreglo circular en antebrazo para 15 movimientos fue un desacierto, no es posible aumentar más del 50% de reconocimiento.

Una propuesta para mejorar la clasificación fue la de eliminar los valores atípicos del diseño de los clasificadores. Aunque todo musculo activado en cada movimiento se debiera activar de igual forma, el sistema de adquisición de datos captaría las amplitudes eléctricas de las fibras musculares correspondientes que se sumarán en uno o más electrodos. Pero del universo de datos utilizados en el diseño de los clasificadores, existen valores que se salen de su comportamiento normal y que pueden no representar una acción típica del movimiento. Eliminar estos valores atípicos mejora el desempeño de los clasificadores.

Valores atípicos

Como se estableció existen dos rutas definidas para evaluar los valores atípicos diferenciándose solamente si se tomaban los movimientos de cada sujeto por separado (ruta 1) o si se juntaban todos de manera grupal (ruta 2). Los resultados de la ruta 1 mostrados en la Tabla 4.1 muestran una particularidad, los sujetos 6 y 7 poseen valores mantenidos y eliminados iguales en los movimientos I_I, L_L, M_M, R_R, T_I, T_M, T_R y en T_T, lo que significa que estas señales son el mismo trazo de datos, lo que no debiera ocurrir según el experimento; la causa para nosotros fue desconocida pudiendo deberse a algún error al momento de formar las matrices en la base de datos. En cambio, los resultados de la ruta 2 (Tabla 4.2) no muestran esta particularidad debido al enfoque más englobado que se le dio. La pregunta sería ¿Qué consideramos un valor atípico?, en la ruta 1 se definió un valor atípico como un valor que se alejaba del centro de masa de todos los datos para un movimiento aislado, para la ruta 2 se consideró un valor atípico como un valor que se alejaba del centro de masa de todos los datos

para todos los movimientos en conjunto. Por esta razón es que la ruta 1 y la ruta 2 producen resultados distintos y por esta misma razón la ruta 1 puede detectar los mismos números de datos mantenidos y eliminados de señales idénticas en los canales EMG; lo que nos hace dudar de la calidad de la base de datos; en un experimento bien llevado no puede ser que las señales de dos canales distintos sean idénticas en un movimiento en todas sus repeticiones. Como se observó el cómo seleccionar la forma de eliminar valores atípicos es un trabajo completo:

- 1.- Porque no sabes cómo se comportan tus datos, lo que quiere decir es que no sabes si estos valores atípicos en realidad son valores reales. ¿Se puede decir que existen valores atípicos dentro de las mediciones EMG ya que cada movimiento es distinto?
- 2.- Los movimientos son tan variados y para un conjunto de fibras de un músculo no siempre los conjuntos de fibras se activan de la misma forma.

Podemos ver que los valores atípicos para cada movimiento evaluado de la ruta 1 podrían considerarse más fidedignos en cuanto a la definición de valores atípicos para movimientos por separado, ya que cada registro de datos para cada movimiento está separado y evaluado con el algoritmo de valores atípicos utilizado, sin embargo, los valores atípicos de la ruta 2 que fue la que utilizamos (debido al objetivo general de nuestra investigación: el trazar una ruta para un control del movimiento) son distintos. Se observó como al momento de definir la forma en que consideramos un valor atípico afectó a los movimientos ya que en general hubo una reducción de valores atípicos para los movimientos y el concentrado de valores atípicos fueron para un movimiento en particular, el de mano cerrada en la ruta 2. Esto se puede observar en las Figuras 4.1 y 4.2, probablemente porque es un movimiento que utiliza más fibras musculares y por lo tanto es el que posee mayores variaciones en sus amplitudes.

Ventaneo

Al momento de crear las ventanas no tomamos en cuenta las repeticiones de los movimientos evaluados. Como se mostró en la Tabla 3.4 las repeticiones fueron insertadas en una tabla y después se procedió a separar los datos cada 250 datos lo que indica que una ventana podría contener datos de un final de una repetición de movimiento con un inicio de otra repetición.

Clasificación

Como se observó en la sección de experimentación inicial, así como en la experimentación 2 y 3 de resultados se encontró evidente que los datos de una persona son diferenciables pero los datos

de un conjunto de personas empiezan a ser indiferenciables. Lo podemos ver en la Tabla 4.6 y 4.8, al momento de comparar los modelos de forma individual y el modelo creado a partir del grupo de 7 sujetos.

Selección de características

Dentro de la experimentación inicial las 12 características evaluadas por cada electrodo (96 características) permiten una buena separación entre clases con un modelo de clasificación muy complejo (96 variables) que nos brindaba un porcentaje en promedio para cada sujeto por separado de 92.38% y para el grupo muestral de 59.81% para un modelo lineal (LDA) y para uno cuadrático (QDA) de 88.74% y 55.04% respectivamente. Aunque esta cantidad de operaciones es algo exagerada no sabemos si algunas de ellas son prescindibles, es decir, no sabemos si todas eran realmente útiles por lo tanto al momento de aplicar el algoritmo de selección de características en la experimentación 2 se encontró en la Tabla 4.9 tanto para el modelo lineal como el cuadrático que muchas de ellas perjudicaban al momento de hacer una clasificación quedando en promedio por sujeto por separado en 42.71 características con un promedio en clasificación de 92.33% y con 77 características para el grupo muestral con una clasificación de 59.79% en el modelo LDA y 21.57 características en promedio para sujetos individuales con un promedio en clasificación de 94.66% con 49 características para el grupo muestral con un porcentaje de 80.54% para el modelo QDA, evidentemente hubo un aumento en la clasificación correcta para el modelo cuadrático sin haber aumento en el modelo lineal, ambos sufrieron una disminución de características, como un ejemplo en particular tenemos al sujeto 6 donde solo requirió 14 características de las 96 con un porcentaje de aciertos en las decisiones del 97% con un modelo cuadrático creando así un modelo más simple. Con la selección secuencial de características se vio que muchas de las características con un modelo LDA y QDA son prescindibles, incluso para el modelo QDA se observó que algunas características podrían ser eliminadas para todos los electrodos, para todos los sujetos por separado como es el caso que vemos en la Tabla 4.12 (modelo cuadrático) donde SSI, TM4 y TM5 no contribuyen en absoluto, encontramos también dos características (VAR, TM3) que solo contribuyen en dos operaciones para las 8 para cada sujeto, empezando con VAR para el sujeto 4 y el sujeto 5, electrodo 3 y electrodo 4 respectivamente y continuando con TM3 para el sujeto 3 y el sujeto 4, electrodo 3 y electrodo 4 respectivamente. Estas dos últimas características quedarían como trabajo futuro

calcular las implicaciones al eliminarlas. Con los grupos muestrales es distinto, ya que se utilizan más características (49 variables) para diferenciar los movimientos.

Este algoritmo de selección de atributos sirvió para dos cosas:

Una reducir la cantidad de operaciones realizadas. Dos aumentar el porcentaje de reconocimiento del modelo para el caso del modelo cuadrático. Si seleccionamos el mejor modelo para clasificación nos quedaríamos con el cuadrático.

Algoritmo de matrices dispersas

El reducir aún más la dimensionalidad de los datos es otra tarea que nos empeñamos a realizar, queríamos ver cuanto era el porcentaje de reconocimiento con solo 3 características, ahora bien, el utilizar un algoritmo para transformar a un nuevo espacio fue utilizado.

Algoritmo de matrices dispersas sirvió para transformar al espacio de medias de los grupos y así reducir aún más el número de características evaluadas y obtener la información más representativa de las características originales, aunque hay una reducción en el porcentaje podemos ver que con tan solo las primeras tres características en este nuevo espacio nos brinda un porcentaje promedio de 73.94 % para los 7 sujetos individuales, y para el grupo muestral de 31.03% cifra que es verdaderamente baja, pasándonos a la Tabla 4.13 donde se expone una evaluación de las características iniciando con 3, 6,8,10,12,14,16,18 y 20, donde podemos observar que entre 12 y 14 características evaluadas ya estamos obteniendo un porcentaje entre 84.07% y 86.49%.

Técnica de extensión de clases

La técnica de extensión de clases aplicada en la experimentación 1, 2 y 3 y detallada en la sección 3.11 muestra un evidente aumento en el porcentaje de clasificación como se mostró en las Tablas 4.7 y 4.10 aunque hay implicaciones al momento de aplicar esta técnica, como por ejemplo el modelo tiene una mayor resolución al poseer más clases definidas (105 clases) donde en realidad las únicas operaciones extras son darles clases distintas a los movimientos de cada sujeto, aplicar el modelo y reclasificar las clases después de la evaluación del modelo es decir regresar a 15 clases de las 105 que es capaz de diferenciar el modelo. Tan solo el darle esta resolución al modelo aumenta 9.07% ajustando las clases el aumento es de 10.83%. en la experimentación 2, Tabla 4.7 y Tabla 4.10.

Con la experimentación 3 encontramos en la Tabla 4.14 que el promedio de las clases ajustadas del grupo muestral de 7 sujetos fue de 33.39% con un aumento de 2.36% sobre el modelo que evalúa 105 clases sin rehacer el ajuste a solo 15. Y si evaluamos un espacio entre 12 y 14 características obtendríamos 86.58% y 88.98% en comparación con la experimentación 3 donde no se realiza este cambio de clases.

Modelo Opensim

El tiempo de procesamiento mostrado en la Tabla 4.32 muestra el tiempo aislado de cada una de las operaciones sobre los vectores y matrices calculados sobre la ventana de comandos de Matlab, donde se muestra cada una de las operaciones para un nuevo vector de entrada (vector de datos crudos 8 columnas) y termina con la formación de la matriz para reproducir el movimiento en Opensim del modelo creado y configurado.

Aunque optimizar el tiempo de procesamiento es un proyecto fuera de nuestro alcance por detalles como el sistema de adquisición y de procesamiento, que van más allá de nuestros objetivos, sin embargo, como se mostró en la Tabla 4.32 se creó una secuencia que permitiría procesar en tiempo real a nuestro datos, el flujo en términos de procesamiento, como se observa la mayoría del tiempo requerido o el principal cuello de botella es determinar si un valor es atípico o típico después de ahí la siguiente etapa es la aplicación de las operaciones elegidas únicamente a las características que fueron seleccionadas en el proceso para selección de características, omitiendo como se pueden ver las 5 características de toda operación sombreadas en negro en la misma tabla, después la proyección al espacio de medias que también es rápido requiriendo, posteriormente el evaluar el modelo creado y la asignación del vector de movimiento para crear la matriz que servirá como base para el archivo que utilizará Opensim en formato ‘.mot’.

Al momento de modificar el modelo musculoesquelético mostrado en la Figura 4.7 se completaron los requerimientos para reproducir los movimientos evaluados en Opensim, como parte de los resultados se incluyeron los vectores de avance para cada uno de los movimientos evaluados Tabla 4.17 hasta 4.31 con los que se pueden reproducir cada una de las matrices para cada movimiento completo cada vector puede ser multiplicado por 2,3,4,5..., 80 formar una matriz con 80 filas incluir estos datos en la Tabla 3.11 para formar el archivo ‘.mot’ para recrear el movimiento, también se incluyeron cada uno de los movimientos considerados para el modelo

de clasificación y su representación virtual en este programa de simulación, como fe de ello en este documento se imprimieron imágenes de algunos movimientos reproducidos en Opensim Figura 4.8 hasta 4.11.

Conclusiones

La señal electromiográfica es un conglomerado de señales interpuestas que aun en estos tiempos es difícil de descifrar. Lo que se ha realizado hasta el momento son algoritmos de Aprendizaje automático (*Machine Learning*) para obtener un modelo que pueda relacionar estas señales con los movimientos, con limitantes como un número determinado de personas, y de igual forma con un número limitado de movimientos.

Es evidente que cada persona poseerá señales de distinta magnitud para los movimientos evaluados, es decir la fisionomía de una persona afecta las señales.

La forma en la que se adquiere la señal es el punto más importante de este problema si pudiésemos encontrar puntos aislados donde la información que se registra en cada electrodo es única o que no haya información repetida entre los electrodos evaluados, podríamos entonces obtener un modelo más eficiente.

Las características que más ayudan a la diferenciación entre las clases de las que evaluamos fueron Valor absoluto medio, valor RMS de la señal, la longitud de onda (Wl), cambio de signo de la pendiente y la razón de miopulso.

Las características que pueden ser omitidas porque no ayudan o perjudican a la diferenciación entre las clases son momentos temporales 3, 4 y 5, la varianza e integral cuadrada simple.

La ruta 2 de nuestra experimentación es la que se utilizará para proyectos futuros ya que nos permite clasificar en tiempo real.

El modelo obtenido a través de QDA fue el mejor al momento de evaluar el número de características y el porcentaje de reconocimiento.

La transformación de nuestras variables originales al espacio de las medias (matrices dispersas) aunque aumenta el tiempo de procesamiento, es útil para aumentar el espacio entre las clases y reducir el número de variables involucradas en el modelo para hacer nuestro trabajo aplicable al control de una prótesis.

El utilizar diferentes clases para cada individuo de nuestro grupo muestral fue bueno para aumentar el porcentaje de reconocimiento del modelo evaluado, lo que nos podría indicar que hay similitudes entre grupos de personas y sus movimientos por lo tanto para trabajos futuros con un grupo muestral más grande podríamos tomar en cuenta estas similitudes.

Se encontró evidente que los datos de una persona son diferenciables pero los datos de un conjunto de personas empiezan a ser indiferenciables.

El porcentaje de reconocimiento obtenido es comparable con las otras investigaciones del área con la diferencia de que aquí usamos un análisis entre sujetos por separado, el grupo muestral de 8 sujetos evaluando 15 movimientos específicos el grupo más grande de movimientos en las investigaciones de los últimos años. El gran número de movimientos específicos evaluados en [13] es una ventaja frente a otras investigaciones, la especificidad de los movimientos es el inicio de una gran cantidad de movimientos que podemos evaluar frente a otros no evaluados.

El definir movimientos en un tiempo de 62 ms es el inicio para clasificar los 15 movimientos en tiempo real, además con este modelo de clasificación creado podemos hacer adaptaciones de nuevos movimientos virtuales no evaluados.

El programa Opensim te permite analizar cualquier movimiento del cuerpo humano, esto es parte de los objetivos de nuestra investigación, es importante hacer énfasis en las capacidades de este software y su aplicabilidad al momento de relacionar los movimientos de la mano con las señales electromiográficas, lo que es una forma más detallada de ir entendiendo los movimientos y la representación de las señales en estos términos.

ANEXOS

Anexo 1: Músculos de antebrazo. Imágenes tomadas de Frank H. Netter 2011

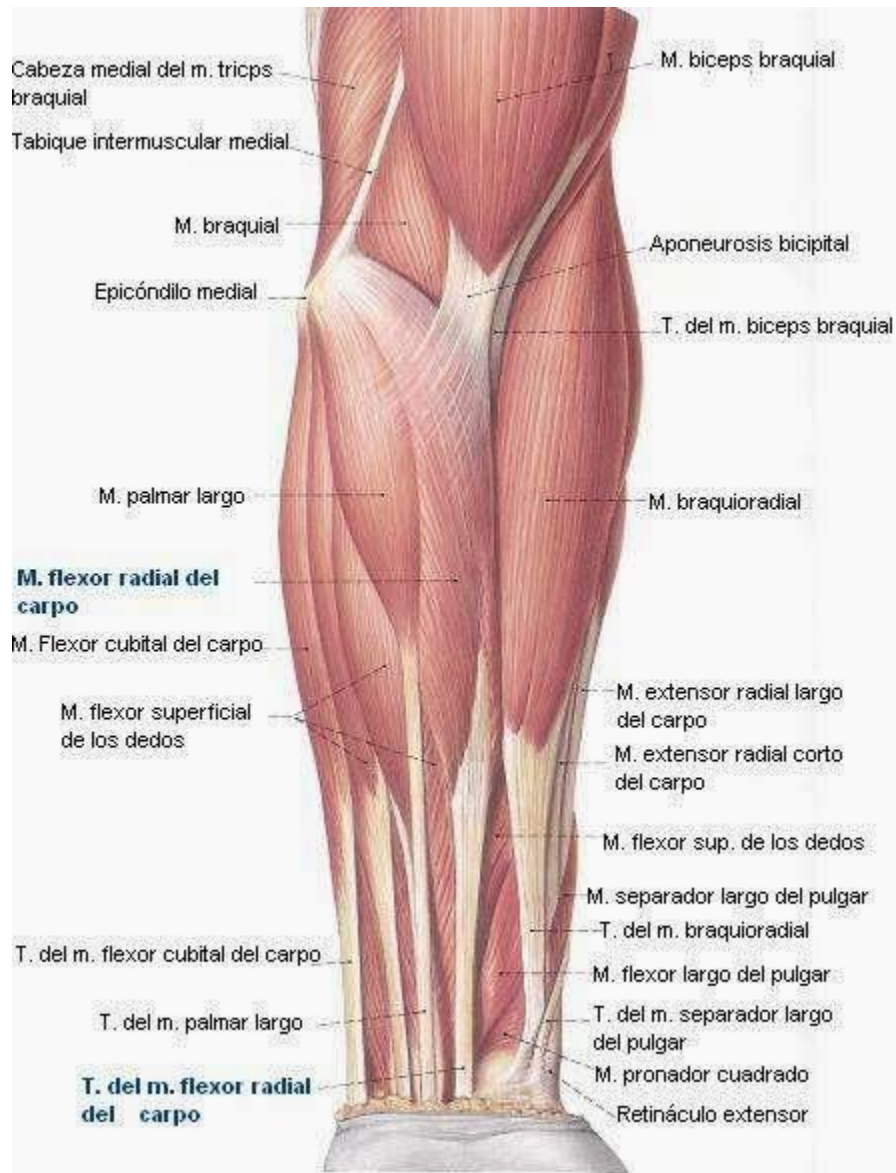
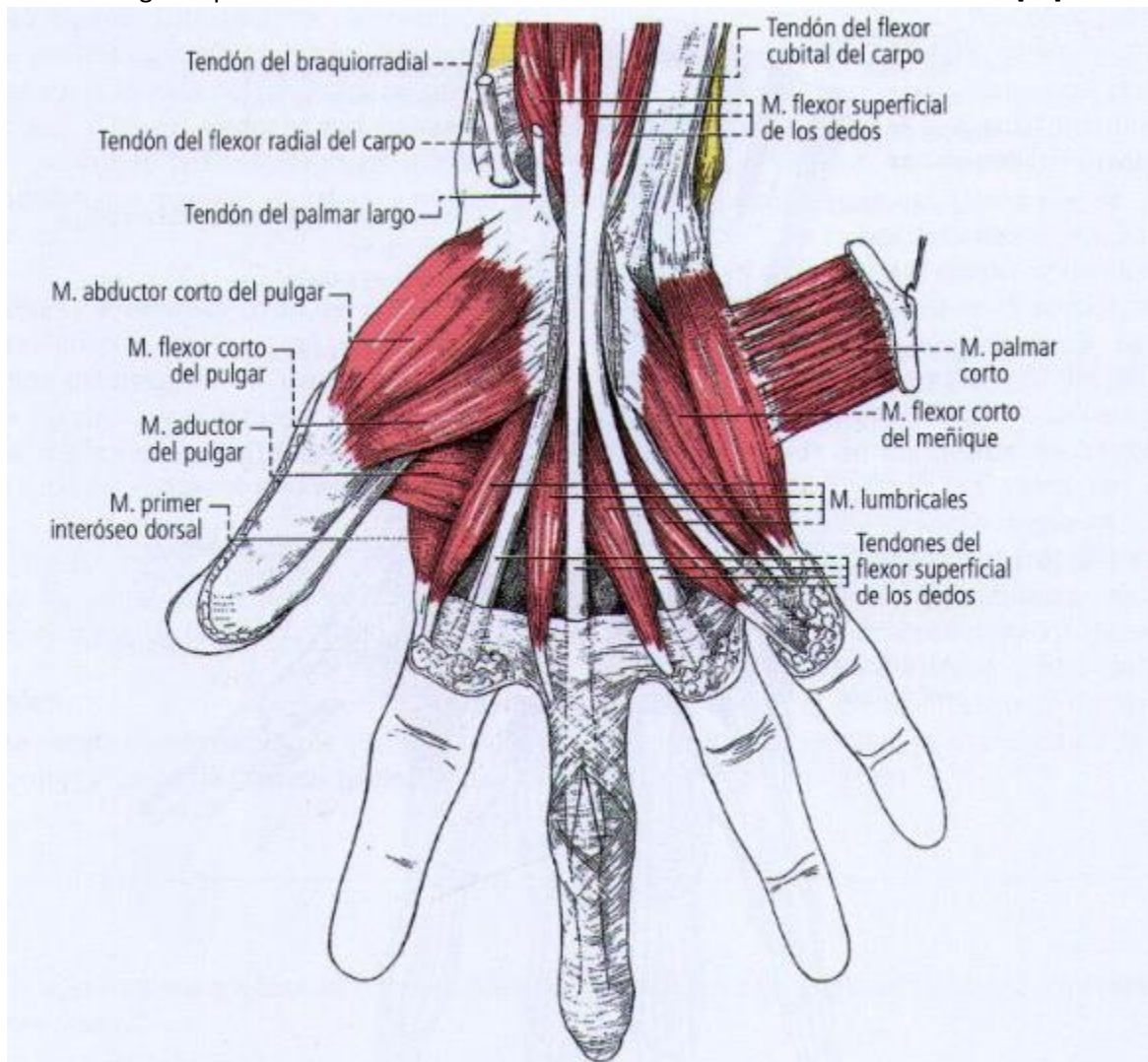


Imagen representativa de los músculos de la mano. Vista anterior. Tomada de [27].



Anexo 2: Código Matlab

Valores atípicos, código para acomodar los datos

%leer todos los movimientos para un sujeto

```
HC_1 = csvread('HC_1.csv');
HC_2 = csvread('HC_2.csv');
HC_3 = csvread('HC_3.csv');
HC=[HC_1;HC_2;HC_3];
clear HC_1 HC_2 HC_3;
I_I1 = csvread('I_I1.csv');
I_I2 = csvread('I_I2.csv');
I_I3 = csvread('I_I3.csv');
I_I = [I_I1;I_I2;I_I3];
clear I_I1 I_I2 I_I3;
I_M1 = csvread('I_M1.csv');
I_M2 = csvread('I_M2.csv');
I_M3 = csvread('I_M3.csv');
I_M=[I_M1;I_M2;I_M3];
clear I_M1 I_M2 I_M3;
IMR1 = csvread('IMR1.csv');
IMR2 = csvread('IMR2.csv');
IMR3 = csvread('IMR3.csv');
IMR=[IMR1;IMR2;IMR3];
clear IMR1 IMR2 IMR3;
L_L1 = csvread('L_L1.csv');
L_L2 = csvread('L_L2.csv');
L_L3 = csvread('L_L3.csv');
L_L=[L_L1;L_L2;L_L3];
clear L_L1 L_L2 L_L3;
M_M1 = csvread('M_M1.csv');
M_M2 = csvread('M_M2.csv');
M_M3 = csvread('M_M3.csv');
M_M=[M_M1;M_M2;M_M3];
clear M_M1 M_M2 M_M3;
M_R1 = csvread('M_R1.csv');
M_R2 = csvread('M_R2.csv');
M_R3 = csvread('M_R3.csv');
M_R=[M_R1;M_R2;M_R3];
clear M_R1 M_R2 M_R3;
MRL1 = csvread('MRL1.csv');
MRL2 = csvread('MRL2.csv');
MRL3 = csvread('MRL3.csv');
MRL=[MRL1;MRL2;MRL3];
clear MRL1 MRL2 MRL3;
R_L1 = csvread('R_L1.csv');
R_L2 = csvread('R_L2.csv');
R_L3 = csvread('R_L3.csv');
R_L=[R_L1;R_L2;R_L3];
```

```

clear R_L1 R_L2 R_L3;
R_R1 = csvread('R_R1.csv');
R_R2 = csvread('R_R2.csv');
R_R3 = csvread('R_R3.csv');
R_R=[R_R1;R_R2;R_R3];
clear R_R1 R_R2 R_R3;
T_I1 = csvread('T_I1.csv');
T_I2 = csvread('T_I2.csv');
T_I3 = csvread('T_I3.csv');
T_I=[T_I1;T_I2;T_I3];
clear T_I1 T_I2 T_I3;
T_L1 = csvread('T_L1.csv');
T_L2 = csvread('T_L2.csv');
T_L3 = csvread('T_L3.csv');
T_L=[T_L1;T_L2;T_L3];
clear T_L1 T_L2 T_L3;
T_M1 = csvread('T_M1.csv');
T_M2 = csvread('T_M2.csv');
T_M3 = csvread('T_M3.csv');
T_M=[T_M1;T_M2;T_M3];
clear T_M1 T_M2 T_M3;
T_R1 = csvread('T_R1.csv');
T_R2 = csvread('T_R2.csv');
T_R3 = csvread('T_R3.csv');
T_R=[T_R1;T_R2;T_R3];
clear T_R1 T_R2 T_R3;
T_T1 = csvread('T_T1.csv');
T_T2 = csvread('T_T2.csv');
T_T3 = csvread('T_T3.csv');
T_T=[T_T1;T_T2;T_T3];
clear T_T1 T_T2 T_T3;
TRN=[HC;I_I;I_M;IMR; L_L;M_M;M_R;MRL;R_L;R_R;T_I;T_L;T_M;T_R;T_T];
%eliminar outliers
[outliers,TRN_N,dm,elim,mant]=ji2(TRN) ;
trn1_mant= mant(mant <= 240000 );
trn1_elim= elim(elim <= 240000 );
trn2_mant= mant(mant <= 480000 & mant >= 240001);
trn2_elim= elim(elim <= 480000 & elim >= 240001);
trn3_mant= mant(mant <= 720000 & mant >= 480001);
trn3_elim= elim(elim <= 720000 & elim >= 480001);
trn4_mant= mant(mant <= 960000 & mant >= 720001);
trn4_elim= elim(elim <= 960000 & elim >= 720001);
trn5_mant= mant(mant <= 1200000 & mant >= 960001);
trn5_elim= elim(elim <= 1200000 & elim >= 960001);
trn6_mant= mant(mant <= 1440000 & mant >= 1200001);
trn6_elim= elim(elim <= 1440000 & elim >= 1200001);
trn7_mant= mant(mant <= 1680000 & mant >= 1440001);
trn7_elim= elim(elim <= 1680000 & elim >= 1440001);

```



```

trn8_mant= mant(mant <= 1920000 & mant >= 1680001);
trn8_elim= elim(elim <= 1920000 & elim >= 1680001);
trn9_mant= mant(mant <= 2160000 & mant >= 1920001);
trn9_elim= elim(elim <= 2160000 & elim >= 1920001);
trn10_mant= mant(mant <= 2400000 & mant >= 2160001);
trn10_elim= elim(elim <= 2400000 & elim >= 2160001);
trn11_mant= mant(mant <= 2640000 & mant >= 2400001);
trn11_elim= elim(elim <= 2640000 & elim >= 2400001);
trn12_mant= mant(mant <= 2880000 & mant >= 2640001);
trn12_elim= elim(elim <= 2880000 & elim >= 2640001);
trn13_mant= mant(mant <= 3120000 & mant >= 2880001);
trn13_elim= elim(elim <= 3120000 & elim >= 2880001);
trn14_mant= mant(mant <= 3360000 & mant >= 3120001);
trn14_elim= elim(elim <= 3360000 & elim >= 3120001);
trn15_mant= mant(mant <= 3600000 & mant >= 3360001);
trn15_elim= elim(elim <= 3600000 & elim >= 3360001);
for n=1:15
switch n
case 1
trn=TRN(trn1_mant,:);
case 2
trn=TRN(trn2_mant,:);
case 3
trn=TRN(trn3_mant,:);
case 4
trn=TRN(trn4_mant,:);
case 5
trn=TRN(trn5_mant,:);
case 6
trn=TRN(trn6_mant,:);
case 7
trn=TRN(trn7_mant,:);
case 8
trn=TRN(trn8_mant,:);
case 9
trn=TRN(trn9_mant,:);
case 10
trn=TRN(trn10_mant,:);
case 11
trn=TRN(trn11_mant,:);
case 12
trn=TRN(trn12_mant,:);
case 13
trn=TRN(trn13_mant,:);
case 14
trn=TRN(trn14_mant,:);
case 15
trn=TRN(trn15_mant,:);

```

```

end
% extracción de características
nvent=length(trn)/250;
nvent=ceil(nvent);
vect(1:nvent,1:96)=zeros;
i=1;
MPR1(1,8)=zeros; %ciclo para crear los nuevos
numvar=8;
for j=1: nvent %vectores de datos
if nvent==j
vent=trn(i:end,1:8);
SSC=[];
SSC(1:length(vent(:,1))-2,1:8)=zeros;
else
vent=trn(i:i+249,1:8);
SSC(1:248,1:8)=zeros;
end
i=250+i;
N=length(vent(:,1));
%Mean absolute value
%AV=vent;
AV=abs(vent);
if N==1
MAV=AV;
else
MAV=mean(AV);
end
%Mean square value
if N==1
MSV=1/N.*((vent.^2));
else
MSV=1/N.*(sum(vent.^2));
end
%Simple squeue Integral
if N==1
SSI=(vent.^2);
else
SSI=sum(vent.^2);
end
%RMS
if N==1
RMS= real(sqrt(1/N.*(vent.^2)));
else
RMS= real(sqrt(1/N.*sum(vent.^2)));
end
%Varianza EMG
if N==1
var=1/(N).*(vent.^2);

```

```

else
var=1/(1-N).*sum(vent.^2);
end
% TM3,TM4,TM5
if N==1
TM3=abs(1/N*(vent.^3));
TM4=abs(1/N*(vent.^4));
TM5=abs(1/N*(vent.^5));
else
TM3=abs(1/N*sum(vent.^3));
TM4=abs(1/N*sum(vent.^4));
TM5=abs(1/N*sum(vent.^5));
end
%Waveform length correido
WL=[0,0,0,0,0,0,0,0];
for k=1:N-1
WL=abs(vent(k+1,:)-vent(k,:))+WL;
end
%Zero crossing
for k=1:numvar
ZC=vent(:,k);
Hzerocross = dsp.ZeroCrossingDetector;
NZC(1,k) = step(Hzerocross,ZC);
end
%Myopulse procentage rate 30mV threshold
for k=1:N
MPR=vent(k,:)>=0.0001;
MPR1=MPR1+MPR;
end
MPR1=1/N.*MPR1;
%Slope sign change
for k=1:numvar
if length(vent(:,1)) == 1
SSC(:,k)=0;
else
SSC(:,k)=diff(sign(diff(vent(:,k))))~=0;
end
end
if length(vent(:,1)) == 1
SSC1(:,k)=0;
else
SSC1=1/(N-1)*sum(SSC);
end

% nuevo vector características
vect(j,:)=horzcat(MAV,MSV,SSI,RMS,var,TM3,TM4,TM5,WL,double(NZC),MPR1,SSC1);
end

```

```

switch n
case 1
nf_HC=vect;
nf_HC(:,97)=ones;
outliers1=outliers;
matrizlimpia1= trn;
nvent1=nvent;
case 2
nf_I_I=vect;
nf_I_I(:,97)=2;
outliers2=outliers;
matrizlimpia2= trn;
nvent2=nvent;
case 3
nf_I_M=vect;
nf_I_M(:,97)=3;
outliers3=outliers;
matrizlimpia3= trn;
nvent3=nvent;
case 4
nf_IMR=vect;
nf_IMR(:,97)=4;
outliers4=outliers;
matrizlimpia4= trn;
nvent4=nvent;
case 5
nf_L_L=vect;
nf_L_L(:,97)=5;
outliers5=outliers;
matrizlimpia5= trn;
nvent5=nvent;
case 6
nf_M_M=vect;
nf_M_M(:,97)=6;
outliers6=outliers;
matrizlimpia6= trn;
nvent6=nvent;
case 7
nf_M_R=vect;
nf_M_R(:,97)=7;
outliers7=outliers;
matrizlimpia7= trn;
nvent7=nvent;
case 8
nf_MRL=vect;
nf_MRL(:,97)=8;
outliers8=outliers;
matrizlimpia8= trn;

```

```

nvent8=nvent;
case 9
nf_R_L=vect;
nf_R_L(:,97)=9;
outliers9=outliers;
matrizlimpia9= trn;
nvent9=nvent;
case 10
nf_R_R=vect;
nf_R_R(:,97)=10;
outliers10=outliers;
matrizlimpia10= trn;
nvent10=nvent;
case 11
nf_T_l=vect;
nf_T_l(:,97)=11;
outliers11=outliers;
matrizlimpia11= trn;
nvent11=nvent;
case 12
nf_T_L=vect;
nf_T_L(:,97)=12;
outliers12=outliers;
matrizlimpia12= trn;
nvent12=nvent;
case 13
nf_T_M=vect;
nf_T_M(:,97)=13;
outliers13=outliers;
matrizlimpia13= trn;
nvent13=nvent;
case 14
nf_T_R=vect;
nf_T_R(:,97)=14;
outliers14=outliers;
matrizlimpia14= trn;
nvent14=nvent;
case 15
nf_T_T=vect;
nf_T_T(:,97)=15;
outliers15=outliers;
matrizlimpia15= trn;
nvent15=nvent;
end
clear vect;
end

```

```
%SFS LDA
data2=[nf_HC;nf_I_L;nf_I_M;nf_IMR;nf_L_L;nf_M_M;nf_M_R;nf_MRL;nf_R_L;nf_R_R;nf_T_L;nf_T_L;nf_T_M;nf_T_R;nf_T_T];
```

Código para calcular valores atípicos a través de la distribución chi cuadrada utilizado en el código anterior

```
function [outliers,matrizlimpia,dm,elim,mant]=ji2(x)
[nmuest,var]= size(x);
media=mean(x); %media de cada variable de los datos
s=cov(x); %sentencia para obtener las covarianzas
V12(nmuest,var)=zeros; %vector para las distancias
for i=1:var %resta de media cada variable a cada dato
for j=1:nmuest
V12(j,i)=abs(x(j,i)-media(i));
end
end
d1=s\V12'; %covinv=inv(s); d1=covinv*V12'

dm(nmuest,1)=zeros;
for i=1:nmuest
%únicamente
dm(i,1)=V12(i,:)*d1(:,i);
end
chi= ncx2inv(0.975,var,0); %valor critico según la distribución ji^2
elim = find(dm > chi); %valores eliminados de mis datos
mant = find(dm <= chi); %valores mantenidos en mis datos
matrizlimpia=x(mant,:); %matriz original limpia de valores atípicos
outliers=x(elim,:); %matriz con val. atípicos para fines gráficos
ord=sort(dm); %se ordenan los datos de mayor a menor
valelim = find(ord > chi);
valmant= find(ord <= chi);
lonvalman=size(valmant);
end
```

Normalización

```
function [Xnewmed0,Xnew2med0var1]=Estddatos(x)
mediadat=mean(x);
[dat,var]=size(x);
Xnewmed0(dat,var)=zeros;
for i=1:var %datos menos la media
for j=1:dat
Xnewmed0(j,i)=x(j,i)-mediadat(i);
end
end
s=std(x); %covarianza de datos originales
%media cero, varianza 1
Xnew2med0var1(dat,var)=zeros;
```

```

for i=1:var
for j=1:dat
Xnew2med0var1(j,i)=Xnewmed0(j,i)/s(1,i);
end
end
Xnew2med0var1=Xnew2med0var1';
Xnewmed0=Xnewmed0';
x=x';
figure; %graficado
plot(x(1,:),x(2,:), 'b.')
grid, axis square
title('Datos Originales');
figure;
plot(Xnewmed0(1,:),Xnewmed0(2,:), 'b.')
grid, axis square
title('Datos Media cero');
figure;
plot(Xnew2med0var1(1,:),Xnew2med0var1(2,:), 'r.',
grid, axis square
title('Datos Media cero Varianza 1');
end

```

Selección de atributos LDA

```

function [lda,mejor_eval,subconjunto]=seleccion_atributos_LDA(data)
class=data(:,97);
data(:,97)=[];
[~,var]=size(data);
mejor_eval=0;
espacio=(1:var);
EXIT=false;
cont=1;
while ~ EXIT % Repetir el ciclo hasta que
temp=[]; % Exit sea true
temp(1:length(espacio))=zeros;
for i=1 :length(espacio) % Ciclo para evaluar el error
if cont==1 % del clasificador
[lda,ldaCVerErr]=LDA(data(:,i),class);
else
[lda,ldaCVerErr]=LDA([data(:,subconjunto),data(:,espacio(:,i))],class);
end
temp(i)=1-ldaCVerErr; % Se guarda el porcentaje de reconocimiento
end % de cada una de las características
[mejor_i,ind]=max(temp); % evaluadas
if mejor_i>=mejor_eval
subconjunto(cont)=espacio(ind);
mejor_eval=mejor_i;

```

```

espacio(ind)=[];
cont=cont+1;
else
EXIT=true;
end
if var==length(subconjunto)
break
end
end
end
end

```

Selección de atributos QDA

```

function [Qdam,mejor_eval,subconjunto]=seleccion_atributos_QDA(data)
class=data(:,97);
data(:,97)=[];
[~,var]=size(data);
mejor_eval=0;
espacio=(1:var);
EXIT=false;
cont=1;
while ~ EXIT                                % Repetir el ciclo hasta que
temp=[];                                    % Exit sea true
temp(1:length(espacio))=zeros;
for i=1 :length(espacio)                    % Ciclo para evaluar el error
if cont==1                                % del clasificador
[Qdam,QdaCVerrm]=QDA(data(:,i),class);
else
[Qdam,QdaCVerrm]=QDA([data(:,subconjunto),data(:,espacio(:,i))],class);
end
temp(i)=1-QdaCVerrm;                        % Se guarda el porcentaje de reconocimiento
end                                          % de cada una de las características
[mejor_i,ind]=max(temp);                  % evaluadas
if mejor_i>=mejor_eval
subconjunto(cont)=espacio(ind);
mejor_eval=mejor_i;
espacio(ind)=[];
cont=cont+1;
else
EXIT=true;
end
if var==length(subconjunto)
break
end
end
end
end

```


Selección de atributos Naive Bayes

```
function [Mdl,mejor_eval,subconjunto]=seleccion_atributos_NB(data)
class=data(:,97);
data(:,97)=[];
[~,var]=size(data);
mejor_eval=0;
espacio=(1:var);
EXIT=false;
cont=1;
while ~ EXIT                                % Repetir el ciclo hasta que
temp=[];                                    % Exit sea true
temp(1:length(espacio))=zeros;
for i=1 :length(espacio)                    % Ciclo para evaluar el error
if cont==1                                % del clasificador
[Mdl,nbCVerErr]=NB(data(:,i),class);
else
[Mdl,nbCVerErr]=NB([data(:,subconjunto),data(:,espacio(:,i))],class);
end
temp(i)=1-nbCVerErr;                        % Se guarda el porcentaje de reconocimiento
end                                          % de cada una de las características
[mejor_i,ind]=max(temp);                    % evaluadas
if mejor_i>=mejor_eval
subconjunto(cont)=espacio(ind);
mejor_eval=mejor_i;
espacio(ind)=[];
cont=cont+1;
else
EXIT=true;
end
if var==length(subconjunto)
break
end
end
end
```

Clasificador Naive Bayes

```
function [Mdl,nbCVerErr]=NB(data,class)
Mdl = fitcnb(data,class,...
'ClassNames',{'1','2','3','4','5','6','7','8','9','10','11','12','13','14','15'});
                                                %crea el clasificador LDA con los datos de entrada
rng(0,'twister');
cp = cvpartition(class,'KFold',10);          %se crean las particiones de los datos con 10 K folds
cvlda = crossval(Mdl,'CVPPartition',cp);     %se evalua el clasificador lda con las particiones creadas
nbCVerErr = kfoldLoss(cvlda);                %se calcula el error de CV
end
```

Clasificador LDA

```
function [lda,ldaCvErr]=LDA(data,class)
lda = fitcdiscr(data,class,'DiscrimType','Linear'); %crea el clasificador LDA con los datos de entrada
rng(0,'twister');
cp = cvpartition(class,'Kfold',10); %se crean las particiones de los datos con 10 K folds
cvlda = crossval(lda,'CVPartition',cp); %se evalua el clasificador lda con las particiones creadas
ldaCvErr = kfoldLoss(cvlda); %se calcula el error de CV
end
```

Clasificador QDA

```
function [Qda,QdaCvErr]=QDA(data,class)
Qda = fitcdiscr(data,class,'DiscrimType','pseudoquadratic'); %crea el clasificador LDA con los datos de entrada
rng(0,'twister');
cp = cvpartition(class,'Kfold',10); %se crean las particiones de los datos con 10 K folds
cvQda = crossval(Qda,'CVPartition',cp); %evalúa el clasificador Qda con las particiones creadas
QdaCvErr = kfoldLoss(cvQda); %se calcula el error de CV
end
```

Algoritmo de matrices dispersas

Blanqueamiento de los datos

```
function [x, XW, W, M] = esfereo(x)
S = cov(x); %Covarianza
M = mean(x); %Media
l = length(x);
M = repmat(M,l,1);
x_minus_mu = x - M;
[E,D] = eig(S); %Valores y vectores característicos de la covarianza
%Potencia a -1/2 de los valores característicos

d = diag(D);
d = real(d.^-.5);
D = diag(d);
% Esfereo de datos
%XW = E*D*x_minus_mu'; %Esfereo normal
W = E*D*E'; %Transformada de esfereo
XW = W*x_minus_mu'; %De rotación para volver al espacio original
XW=XW'; %XW toma el valor de esfereo
end
```

Matrices dispersas

```
function [proyeccion_x] = vis_scatter_mat_clust(x,cl_or_labels)
[n m] = size(x); %Saber cuántas variables y muestras se tienen
dif_labels = unique(cl_or_labels);
```

```

%esfereo, ahora x es XW
x = esfereo(x);
prototipos = zeros(max(dif_labels),m);
for i = 1:length(dif_labels)
    indexes_i = find(dif_labels(i) == cl_or_labels);
    Nj(i) = length(indexes_i);
    prototipos(i,:) = mean(x(indexes_i,:));
end
bhat = zeros(max(dif_labels),m); %Ortogonalizacion usando gram-schmidt algoritmo
proj_current = 0;
count = 0;
for i = 1:max(dif_labels)
    if i == 1
        bhat(i,:) = prototipos(i,:);
        count = count + 1;
    else
        A = prototipos(i,:);
        for j = 1:count
            B = bhat(j,:);
            proj_current = -1*(dot(B,A)/dot(B,B));
            vec_proj_current(j,:) = proj_current*B;
        end
        [R_vec_proj, C_vec_proj] = size(vec_proj_current);
        if R_vec_proj == 1
            bhat(i,:) = A + vec_proj_current;
        else
            bhat(i,:) = A + sum(vec_proj_current);
        end
        count = count + 1;
    end
end
Xc = x*bhat'*bhat;
media_general = mean(Xc);
prototipos_xc = zeros(max(dif_labels),m);
for i = 1:length(dif_labels)
    indexes_i = find(dif_labels(i) == cl_or_labels);
    prototipos_xc(i,:) = mean(Xc(indexes_i,:));
end
SW = 0; %Calcular SW
for i = 1:max(dif_labels)
    media_per_cluster = repmat(prototipos(i,:),n,1);
    SW_temporal = x - media_per_cluster;
    SW_temporal = SW_temporal'*SW_temporal;
    SW = SW + SW_temporal;
    SW_temporal = 0;
end
SB = 0; %Calcular SB
[n1 m2] = size(prototipos);

```

```

for i = 1:max(dif_labels)
    SB_temporal = prototipos(i,:) - media_general;
    SB = SB + Nj(i)*SB_temporal'*SB_temporal;
end
mC = pinv(SW)*SB;
S = cov(mC);
[E,D] = eig(S);
covarianza

d = diag(D);[VM IM] = sort(d,'descend')
E_sort = E(:,IM);
proyeccion_x = x*E_sort;
end

%Calcular mc
%Covarianza
%Valores y vectores característicos de la
%Se acomodan por orden las columnas con

```

Referencias

- [1] Jesus Manuel Dorador González, Patricia Rios Murillo, Itzel Flores Luna & Ana Juarez Mendoza. (2005). Robotica y protesis intelientes . 18 enero de 2005, de UNAM Sitio web: www.revista.unam.mx/vol.6/num1/art01/art01-1.htm
- [2] Kim Norton. (Noviembre / diciembre 2007). Un breve recorrido por la historia de la protésica. inMotion, 17, 1-5. Sitio web: http://www.amputee-coalition.org/spanish/inmotion/nov_dec_07/history_prosthetics.x18949.pdf
- [3] Orthopaedic Prosthetics Market. (2016-11-12). Growth Expected to be Driven by Increasing Ageing Population and Rising Diabetes Rates: Global Industry Analysis and Opportunity Assessment 2016 - 2026. enero 7,2017, de Future Market Insights Sitio web: <http://www.futuremarketinsights.com/reports/orthopedic-prosthetic-devices-market>
- [4] Dr. Grant McGimpsey & Terry C. Bradford, Limb Prosthetics Services and Devices Critical Unmet Need: Market Analysis [white paper] Limb Prosthetics Services and Devices Critical Unmet Need: Market Analysis. 7/5/2017, de <https://pdfs.semanticscholar.org/c3ae/f3563844e2e2835411fcbc2b0fe3091ac30b.pdf>
- [5] touchbionics by osur. (2017). Innovaciones en protésica de miembro superior. 7/5/2017, de touchbionics Sitio web: <http://www.touchbionics.es/>
- [6] Fillauer co. (2010). Body Powered Systems. 7/5/2017, de Filllauer Sitio web: <http://fillauer.com/Upper-Extremity-Prosthetics/body-powered-systems/index.html>
- [7] Oskoei, M. A., & Hu, H. (October 2007). Myoelectric control systems—A survey. Biomedical Signal Processing and Control, 2 (4), 275-294.
- [8] B. Hudgins, P. Parker & R. Scott, A new Strategy for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 40 (1) (1993) 82–94.
- [9] Parker PA, Englehart K & Hudgins B. Myoelectric signal processing for control of powered prostheses. J Electromyogr Kinesiol 2006;16(541):548
- [10] D. Farina & R. Merletti, Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions, J. Electromyogr. Kinesiol. 10 (2000) 337–349.
- [11] Sebastian Amsüss, Peter M. Goebel, Ning Jiang, Bernhard Graimann, Liliana Paredes & Dario Farina. (23 December 2013). Self-Correcting Pattern Recognition System of Surface EMG Signals for Upper Limb Prosthesis Control. IEEE Transactions on Biomedical Engineering, 61 (4), 1167-1176.
- [12] Rami N. Khushaba, Sarath Kodagoda, Maen Takruri & Gamini Dissanayake. (15 September 2012). Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals. Expert Systems with Applications, 39 (12), 10731-10738.
- [13] R. N. Khushaba & Sarath Kodagoda, “Electromyogram (EMG) Feature Reduction Using Mutual Components Analysis for Multifunction Prosthetic Fingers Control”, inProc. Int. Conf. on Control, Automation, Robotics & Vision (ICARCV), Guangzhou, 2012, pp. 1534-1539. (6 pages).
- [14] Boostani, R., & Moradi, M. H. (2003). Evaluation of the forearm EMG signal features for the control of a prosthetic hand. Physiological Measurement, 24(2), 309-319.
- [15] X. Li, S. Chen, H. Zhang, O.W. Samuel, H. Wang, P. Fang, X. Zhang, G. Li. (2016 June). Towards reducing the impacts of unwanted movements on identification of motion intentions. Journal of Electromyography and Kinesiology, 28, 90-98.

- [16] Angkoon Phinyomark, Pornchai Phukpattaranont, Chusak Limsakul. (15 june 2012). Feature reduction and selection for EMG signal classification. *Expert Systems with Applications*, 39 (8), 7420-7431.
- [17] Zecca, M., Micera, S., Carrozza, M. C., & Dario, P. (2002). Control of multifunctional prosthetic hands by processing the electromyographic signal. *Critical Reviews in Biomedical Engineering*, 30(4–6), 459–485.
- [19] Lei, M., Wang, Z., & Feng, Z. (2001). Detecting nonlinearity of action surface EMG signal. *Physics Letters A*, 290(5–6), 297–303.
- [20] Christopher M Bishop. (1995). *Neural Networks for Patter recognition*. Department of Computer science and Aplied matematics Aston University Birmonghman, UK: Claredon Press Oxford.
- [21] Amputee Coalition. (2015). Limb Loss Statistics. 7/5/2017, de Amputee Coalition Sitio web: <http://www.amputee-coalition.org/limb-loss-resource-center/resources-by-topic/limb-loss-statistics/limb-loss-statistics/#.WVzA0-mQzIU>
- [22] Ziegler- Graham K, MacKenzie EJ, Ephraim PL, Travison TG, Brookmeyer R. Estimating the Prevalence of Limb Loss in the United States: 2005 to 2050. *Archives of Physical Medicine and Rehabilitation* 2008;89(3):422- 9.
- [23] U.S. National Library of Medicine. (17 noviembre 2016). Distrofia muscular. 7/5/2017, de Medline Plus Sitio web: <https://medlineplus.gov/spanish/musculardystrophy.html>
- [24] MusculoGraphics, Inc. (2012). MusculoGraphics, Inc. 7/5/2017, de MusculoGraphics, Inc. Sitio web: <http://www.musculographics.com/>
- [25] Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, Guendelman E, Thelen DG. OpenSim: Open-source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering*. (2007) <http://simtk-confluence.stanford.edu:8080/display/OpenSim/About+OpenSim>
- [26] Morgan Jeske. (2014). Planes of Movement. 7/5/2017, de Yinyoga Sitio web: http://www.yinyoga.com/newsletter20_planes%20of%20movement.php
- [27] Latarjet. Michel, Ruiz Liar. (2005). *Anatomia Humana*. Francia: Panamericana.
- [28] Delsys Incorporated. (2014). Bagnoli TM 2-Channel Handheld EMG System User Guide. 7/5/2017, de Delsys Incorporated Sitio web: http://delsys.com/Attachments_pdf/download/users-guides/bagnoli-2.pdf
- [29] Ian H. Witten, Eibe Frank, Mark A. Hall. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier: The Morgan Kaufmann Series in Data Management Systems.
- [30] Kevin P. Murphy. (08/24/2012). *Machine Learning A Probabilistic Perspective*. Cambridge, Massachusetts London, England: MIT PRESS.
- [31] P. J. G. Lisboa, I. O. Ellis, A. R. Green, F. Ambrogi, M. B. Dias. (October, 2008). Cluster-based visualisation with scatter matrices. Elsevier Science Inc *Pattern Recognition Letters*, 29 (13), 1814-1823.
- [32] Agnan Kessy, Alex Lewin, and Korbinian Strimmer. (18 Dec 2016). Optimal Whitening and Decorrelation. 7/5/2017, de arXiv.org Sitio web: <https://arxiv.org/pdf/1512.00809.pdf>
- [33] Mathworks. (1994). Discriminant Analysis. 7/5/2017, de The MathWorks, Inc. Sitio web: <http://www.mathworks.com/help/stats/discriminant-analysis.html>
- [34] Gonzalez, R. V., Buchanan, T. S., Delp, S. L. How muscle architecture and moment arms affect wrist flexion-extension moments, *Journal of Biomechanics* Vol. 30, pp. 705-712, 1997.
- [35] L. Testut, A. Latarjet. (1984). *Tratado de anatomía humana Tomo 1*. Francia: Salvat.

[36] De La Cruz-Oré, Jorge Luis. (agosto, 2013). ¿Qué significan los grados de libertad?. Revista Peruana de Epidemiología, vol. 17, núm. 2, pp. 1-6.

[37] Jeniffer Andrea Muñoz García, Iván Amón Uribe. (enero - Junio de 2013). Técnicas para detección de outliers multivariantes. Revista en Telecomunicaciones e Informática, Vol. 3, No. 5. p. 11-25.