

Universidad Autónoma de Baja California

Instituto de Ingeniería

Maestría y Doctorado en Ciencias e Ingeniería



**Modelación Matemática de la Respuesta de un Biosensor
de Glucosa Oxidasa de Segunda Generación mediante
Técnicas de Reconocimiento Estadístico de Patrones**

Tesis que para obtener el grado de:

MAESTRA EN CIENCIAS

Presenta

Livier Rentería Gutiérrez

Director de Tesis:

Dr. Félix Fernando González Navarro

Codirector de Tesis:

Dra. Margarita Stoytcheva

Dedicatorias

A mi padre por haber siempre creído en mi y haberme enseñado a ser quien soy con sus consejos y motivación.

Reconocimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo y patrocinio de la beca para realizar mi proyecto de Tesis para obtener el grado de Maestría.

A la Universidad Autónoma de Baja California (UABC) por permitirme ingresar al proyector de Maestría y doctora en ciencias e Ingeniería y por el apoyo del proyecto y darme la oportunidad de ser miembro de CONACYT.

Al Instituto de Ingeniería primeramente por brindarme la oportunidad de ser alumna del instituto y proporcionarme un lugar y un laboratorio donde trabajé a lo largo de la tesis.

Al director de mi Tesis el Doctor Félix Fernando González Navarro por la paciencia y el apoyo brindado durante el periodo de estudios de maestría , a mi co-directora Dra. Margaria Stoytcheva, así como a cada uno de mis compañeros de los cuales aprendí cada vez más durante el desarrollo de este trabajo.

Resumen

Un biosensor es un instrumento o dispositivo analítico compacto que contiene un elemento de detección biológico o bioreceptor y esta acoplado a un transductor físico-químico que convierte la señal biológica en una señal electrónica. Esta señal es proporcional al compuesto o grupo de compuestos que se quiere determinar y puede recoger información cualitativa o cuantitativa. Las aplicaciones de los biosensores abarcan diversas áreas de investigación y desarrollo como el sector médico y el cuidado del medio ambiente, entre otros. El impacto potencial de los biosensores en diversos campos de la ciencia los ha hecho elementos muy prometedores dadas sus características de simplicidad, tiempo de respuesta rápido y de bajo costo. Sin embargo, esta tecnología está todavía en desarrollo. De esta manera en esta tesis se ofrecen contribuciones al avance del conocimiento en este campo mediante la aplicación de métodos computacionales para modelar su funcionamiento. En particular el objetivo es modelar la respuesta de un biosensor de glucosa oxidasa mediante las técnicas de inteligencia artificial y reconocimiento estadístico de patrones.

Contenido

1. Introducción	1
1.1. Planteamiento del problema	3
1.2. Objetivo General	4
1.2.1. Objetivos Específicos	4
2. Marco Teórico	5
2.1. Diabetes mellitus	5
2.1.1. Definición	5
2.1.2. Prediabetes	7
2.1.3. Diabetes tipo 1	7
2.1.4. Diabetes tipo 2	8
2.1.5. Diabetes gestacional	9
2.2. Glucosa Oxidasa	9
2.2.1. Aplicaciones de la glucosa oxidasa.	10
2.2.2. Modelación de biosensores de glucosa oxidasa	11
2.3. Sensores y biosensores	11
2.4. Tipos de biosensores.	13
2.4.1. Electroquímicos	13
2.4.2. Biosensores de primera generación	14
2.4.3. Biosensores de segunda generación	15
2.4.4. Biosensores de Glucosa Oxidasa	16
2.4.5. Modelación Matemática de Biosensores	17
2.5. Aprendizaje de máquina	20
2.6. Máquinas de vectores de soporte	20
2.7. Redes Neuronales para Regresión	24

2.8. Árboles de clasificación y regresión (CART)	27
2.9. Modelo Generalizado lineal GLM	28
2.10. Modelo Lineal Generalizado GLM	30
2.11. Mínimos cuadrados parciales PLS	31
2.12. Regresión Ridge	33
2.13. Regresión LASSO	34
3. Materiales y Métodos	36
3.1. Metodología	36
3.2. Hardware y Software utilizados	37
3.3. Descripción de los datos	37
3.4. Métodos y técnicas utilizadas	38
3.4.1. Métodos de regresión	39
3.4.2. Modelos lineales	39
3.4.3. Máquinas de vectores de soporte para regresión	39
3.4.4. Parámetros de la red Neuronal de regresión	40
3.4.5. Parámetros de árbol de regressión CART	40
3.4.6. Validación Cruzada o Cross Validation	40
4. Resultados y discusión	42
4.1. Análisis estadístico básico	42
4.1.1. Desempeño de los algoritmos de regresión	43
4.1.2. Modelo de la red neuronal de una capa con 4 neuronas	46
4.1.3. Modelo de CART	49
4.1.4. Modelo de la SVM RBF	49
5. Conclusiones	52
5.1. Conclusiones	52
5.2. Trabajo Futuro	53
A. Análisis de regresión estadística clásica	55
B. Superficie generada por la respuesta del biosensor	58
Referencias	67

Lista de Figuras

2.1. Esquema de un biosensor.	12
2.2. Esquema funcionamiento de una red neuronal. Las neuronas de las capas de entrada representan cada atributo de los datos, las neuronas de la capa de salida representa la predicción de la red, mientras que las neuronas de las capas ocultas permiten que la red aprenda	25
4.1. Visualización gráfica de los coeficientes de correlación R antes y después del LOG de los mejores modelos.	45
4.2. Valores observados y valores predichos renderizados por los modelos de NN, CART y SVM RBF en los datos log.	46
4.3. Valores reales y valores de predicción por medio de los modelos de SVMR en los datos log.	47
4.4. Valores reales y valores de predicción por medio de los modelos de SVMR en los datos log.	48
4.5. Modelo de Red Neuronal final	48
4.6. Modelo gráfico de un árbol de regresión CART	50
4.7. Modelo gráfico de un árbol de regresión podado CART	50
4.8. Modelo gráfico de un árbol de regresión óptimo CART	51
A.1. Visualización gráfica de los coeficientes de correlación R antes y después del LOG en los peores modelos.	56
A.2. Visualización gráfica de los coeficientes de correlación R antes y después del LOG en los peores modelos (continuación).	57

LISTA DE FIGURAS

B.1. Valores reales y valores de predicción por medio de los modelos de NN en los datos log.	59
B.2. Valores reales y valores de predicción por medio de los modelos de NN en los datos log.	60
B.3. Valores reales y valores de predicción por medio de los modelos de RTREE en los datos log.	61
B.4. Valores reales y valores de predicción por medio de los modelos de RTREE en los datos log.	62

Lista de Tablas

2.1. Elementos de los diferentes tipos de biosensores	14
3.1. Parámetros de las condiciones de un biosensor amperométrico. . .	38
4.1. Tabla estadística descriptiva después del pre-procesamiento de los datos.	42
4.2. Variables de correlaciones de Pearson.	43
4.3. Variables de correlaciones de Spearman.	43
4.4. NRMSE y coeficiente de regresión de los modelos matemáticos sin el LOG.	44
4.5. NRMSE y coeficiente de regresión de los modelos matemáticos después de LOG.	44

Capítulo 1

Introducción

La diabetes mellitus es un padecimiento metabólico crónico que se caracteriza por los altos niveles de glucosa en la sangre, causado por la insuficiente producción de insulina en el páncreas que se encarga de convertir la glucosa en energía necesaria (ADM, 2008).

Uno de los temas que han sido de preocupación durante muchos años ha sido el control de los niveles glucosa que existen en la sangre que también se le conoce como glucemia, por la razón de que es considerada la fuente principal de energía de la células del cuerpo humano y es muy importante que los niveles de glucosa se mantengan normales (100 miligramos por decilitro (mg/dL) en ayunas).

Si los niveles de glucemia son superiores se le conoce como *hiperglucemia* lo que puede ser un síntoma principalmente de diabetes o entre otros padecimientos como hipertiroidismo, pancreatitis o cáncer pancreático. Si éstos son inferiores a los niveles normales se considera *hipoglucemia* Las principales síntomas que se presentan al controlar los niveles de azúcar en la sangre se encuentran los temblores, sudoración, hambre, mareos, dolor de cabeza, irritabilidad, convulsiones, pérdida de atención, entre otros.

Según el Diabetes Research Institute (American Diabetes Association, 2013) a nivel mundial este padecimiento afecta a 180 millones de persona y la Organización Mundial de la salud estima que para el 2030 habrá un incremento del 50 % de esta cifra.

México es uno de los países líderes ocupando el séptimo lugar a nivel mundial según especialistas de la Federación Mexicana de Diabetes (FMD). El INEGI

estima que la tasa de mortalidad de cada 100 mil habitantes es de 69.92 en hombres y 74.5 en mujeres (INEGI, 2011).

Actualmente existen diversas técnicas para medir los niveles de glucosa en la sangre. Entre las más comunes están los glucómetros, que es una de las técnicas que actualmente se encuentran en desarrollo mediante el uso de biosensores. Una enzima que se es útil para medir y determinar los niveles de glucemia en los fluidos biológicos se conoce como *glucosa oxidasa*. Los biosensores enzimáticos de segunda generación que datan desde 1984 fueron desarrollados por Cass y Davis (AEG Cass, 1984) donde su mediador era el ferroceno. Éstos convierten la enzima inmovilizándola con un mediador convirtiéndola en un producto medible. Este proceso en particular será explicado más a detalle en secciones posteriores.

De esta manera surge el interés de conocer y plantear modelos matemáticos y herramientas que puedan favorecer el diseño de estos elementos tecnológicos. Esto permitirá tener dispositivos accesibles para el control los niveles de glucemia en las personas y contribuir a la prevención y control de las diferentes enfermedades que conlleva. En particular en este trabajo de investigación se abordó la modelación de un biosensor de glucosa oxidasa de segunda generación por medio de técnicas de reconocimiento estadístico de patrones.

Así, este documento está organizado de la siguiente manera: En el primer capítulo se encuentra la introducción de la presente tesis, en la que se ofrece el planteamiento del problema y los objetivos perseguidos. En el capítulo dos se ofrece el marco teórico el cual consta de las principales definiciones y conceptos en torno a la temática estudiada como la diabetes mellitus, el campo de sensores y biosensores y tipos de biosensores; se explican además los algoritmos de los métodos de aprendizaje de estadístico utilizados durante el desarrollo de la tesis. En el capítulo tres se exponen los materiales y métodos que incluye el conjunto de datos experimentales, los métodos y técnicas utilizadas. En el capítulo cuatro se exponen los resultados obtenidos durante el trabajo experimental. Y finalmente en la última sección se reportan las conclusiones finales así como el trabajo futuro.

1.1. Planteamiento del problema

La Diabetes Mellitus es una enfermedad grave, crónico degenerativa donde los pacientes afectados presentan niveles altos de glucosa en la sangre. Esto es debido a dos principales causas: 1- Una producción insuficiente de insulina o 2- Mal funcionamiento en un tipo de células llamadas islotes. Estas últimas no responden adecuadamente a la insulina en el proceso de regulación de la glucosa en la sangre. En el proceso de la digestión de los seres humanos, el cuerpo convierte los alimentos ingeridos en glucosa. En el páncreas se produce la insulina para convertir esta glucosa en energía. Cuando un paciente es diagnosticado con la enfermedad Diabetes Mellitus todo el proceso presenta una dinámica errática (Ronald *et al.*, 2005).

Debido al incremento de las cifras de pacientes que presentan dicha enfermedad en los últimos años, el interés por el constante monitoreo de los niveles de glucosa en la sangre (glucemia) para el control de la DM ha ido en aumento (D.G.Buerk, 2005). Una de las opciones para tal efecto es el uso de biosensores, debido a su bajo costo de fabricación, alta sensibilidad y confiabilidad. Sin embargo, este campo aun se encuentra bajo desarrollo, sobre todo los biosensores de glucosa oxidasa. Aunque las pruebas con biosensores son sencillas y no se necesitan gran cantidad de instrumentación, aun persisten los mecanismos de prueba y error, el cual puede ser simplificado mediante el uso de modelos matemáticos que contribuyan al apoyo su diseño y desarrollo.

El uso actual de los biosensores se ha extendido a varios campos de la ciencia y la ingeniería. La detección de contaminantes de los recursos hídricos, detección de agentes patógenos en medicamentos y en la industria alimentaria, por mencionar algunos ejemplos. En un escenario normal, un paciente requiere un examen físico y pruebas de laboratorio que tardan algunos días para obtener los resultados necesarios para un correcto tratamiento. Tal demora a veces puede derivar en complicaciones debido a la falta del tratamiento médico adecuado (Ronald *et al.*, 2005). Es por ello que los biosensores se convierten en un dispositivo tecnológico muy prometedor en aplicaciones de salud pública. Por ello esta tesis busca contribuir en el desarrollo de este nuevo campo aportando conocimientos en el

modelado matemático de biosensores que permita expresar su comportamiento y facilitar las tareas de diseño y desarrollo.

1.2. Objetivo General

El objetivo general es obtener modelos matemáticos para la respuesta amperométrica de un biosensor de glucosa-oxidasa de segunda generación a partir de técnicas de aprendizaje estadístico.

1.2.1. Objetivos Específicos

1. Realizar un estudio comparativo de diversos modelos matemáticos de regresión matemática estadística.
2. Caracterizar los principales parámetros de forma de los modelos matemáticos obtenidos.

Capítulo 2

Marco Teórico

En este capítulo, se expondrán las principales definiciones sobre el tema tratado, entre ellos se dará un panorama de lo que es la *diabetes mellitus*, sus principales síntomas y las clases; se expondrá el concepto de biosensores, su tecnología, los diversos tipos y aplicaciones. Por otra parte se ofrecerá una breve introducción al tema de aprendizaje de máquina y una descripción de los métodos de regresión empleados en este trabajo de investigación.

2.1. Diabetes mellitus

2.1.1. Definición

Diabetes mellitus (DB) es un grupo de padecimientos metabólicos caracterizado por la condición llamada hiperglucemia que resulta de la producción de insulina, la acción de la insulina o ambas (ADM, 2008). Existen diversos procesos patológicos que están involucrados en el desarrollo de la diabetes como por ejemplo la destrucción autoinmune de las células beta del páncreas.

Cuando se consumen alimentos se lleva a cabo un proceso químico dentro del intestino delgado donde son convertidos en glucosa. Las células del intestino delgado absorben las moléculas de azúcar la cual se conduce al torrente sanguíneo. Una vez que la sangre llega al páncreas, éste juega un papel importante en el proceso de detección de niveles de glucosa mediante las células-beta. Dichas células son encargadas de la liberación de insulina que permite abrir las células para

convertir el azúcar en energía. En personas que padecen diabetes este importante sistema no funciona correctamente (Ronald *et al.*, 2005).

Según la Asociación Americana de Diabetes, después de realizar diferentes tipos de pruebas en el entorno médico para el diagnóstico como la prueba AC1 que mide el promedio de la glucemia durante 2 o 3 meses anteriores, se diagnostica la diabetes a pacientes cuando dicha prueba es mayor o igual a 6.5 %. Adicionalmente la de tolerancia a la glucosa oral, donde se diagnostica diabetes cuando la glucosa en la sangre es mayor o igual a 200mg/dl y la prueba aleatoria de glucosa plasmática cuando es mayor o igual 200 mg/dl (American Diabetes Association, 2013).

Los síntomas más comunes (American Diabetes Association, 2013) de la DM son:

- Orinar con frecuencia
- Sed excesiva
- Sentir demasiada hambre incluso cuando se está comiendo
- Fatiga extrema
- Visión borrosa
- Las heridas o moretones son lentos en sanar
- Pérdida de peso inexplicable aun cuando se ha consumido mas alimentos (diabetes tipo 1)
- Dolor, entumecimiento en manos y pies (diabetes tipo 2)

México es uno de los países líderes, ocupando el primer lugar a nivel mundial de muertes por este padecimiento según especialistas de la Federación Mexicana de Diabetes (Federación Mexicana de Diabetes, 2010), y según la Encuesta Nacional de Salud y Nutrición la prevalencia en adultos fue de 7 %, y el 7.3 % en mujeres mientras que los hombres fue de 6.5 %. Se registraron además 8 millones de muertes a consecuencia de la diabetes (Palma *et al.*, 2006).

En la literatura médica se clasifican diferentes tipos de diabetes sobre la base de diferentes factores. En las siguientes subsecciones se explicarán los detalles de estos.

2.1.2. Prediabetes

Se caracteriza cuando los niveles de glucosa en sangre están por encima de lo que se considera normal pero no se puede aún considerar como diabetes. Los rangos para las personas con intolerancia a la glucosa presentan 75-g glucosa; dos horas de prueba de tolerancia oral a la glucosa de 140- 199 mg/dL y personas con problemas de alteración de la glucosa en ayunas con valor de 100 a 125mg/dL [Boucher & O'Connell \(2004- 2005\)](#).

2.1.3. Diabetes tipo 1

También conocida como insulino dependiente o Diabetes Juvenil *Diabetes Mellitus* resulta de la destrucción progresiva de las células pancreáticas por el sistema inmunológico ([LeRoith *et al.*, 2004](#)). En esta condición las células-Beta pierden la capacidad de producir insulina lo que tiene como consecuencia altos niveles de glucosa en la sangre entre otras. Actualmente los científicos aun no tienen seguro las causas de este ataque autoinmune. Si no es tratada adecuadamente da lugar a diferentes consecuencias como daño en los ojos, riñones, corazón u ocasionar coma y la muerte ([Diabetes Research Institute, 2010](#)).

El tratamiento de una persona con diabetes tipo 1 consiste en la aplicación de inyecciones de insulina como fuente externa para la adaptación de glucosa en el cuerpo que no solamente depende de la insulina que se aplica sino también de otros factores como la alimentación, el ejercicio, situaciones emocionales y el estrés así como el estado de salud en general ([American Diabetes Association, 2013](#)). Debe existir un balance en la aplicación de insulina ya que si se aplica una cantidad elevada podría llegar a ocasionar hipoglucemia la cual puede llegar a ser mortal de acuerdo con el Institute diabetes research ([Diabetes Research Institute, 2010](#)).

2.1.4. Diabetes tipo 2

Diabetes mellitus tipo 2 (no insulino dependiente). Es una afección en la cual el nivel de la glucosa es muy elevado, el páncreas no produce suficiente insulina o las células resisten los efectos de la insulina o pueden presentarse ambas situaciones. En la diabetes tipo 2 la insulina no es capaz de desbloquear las células, lo que causa que la insulina queda fuera del proceso esto lleva a la hiperglucemia. Cuando los niveles de glucemia son elevados las células-Beta del páncreas producen una sobrecarga de insulina por consecuencia una deficiencia en la capacidad de la producción de insulina para completar correctamente las funciones del sistema(Ronald *et al.*, 2005).

El tratamiento consiste en una dieta y ejercicio y, si los niveles de azúcar en la sangre aún siguen altos, se recetan medicamentos orales para ayudar al cuerpo a utilizar de manera eficiente la insulina.

Los pacientes con diabetes tipo 2 empiezan la mayor parte de las veces con una prediabetes donde los niveles de glucosa son más elevados de lo normal, pero aun no tan elevados como para llamarlo diabetes y presentan resistencia a la insulina. Los factores de riesgo son en mayores de 45 años (ADM, 2008):

- El sobrepeso
- Escasa actividad física
- La medida de la cintura en mujer es superior a las 35 pulgadas y en hombre superior a 40 pulgadas.
- Antecedentes familiares con diabetes tipo 2.
- Padecer del síndrome de ovario poliquístico.
- Presión de sangre mayor a 140/90 mmHg.
- Niveles de colesterol HDL inferiores a 35 mg/dl.
- Triglicéridos elevados 250/dl.

Los síntomas comunes de la diabetes tipo 2 son (ADM, 2008):

- Presentar los síntomas de la diabetes tipo 1
- Infecciones frecuentes
- Vista borrosa
- Las heridas tardan en sanar
- Entumecimiento de manos y pies
- Infecciones de piel, vejiga y encías.

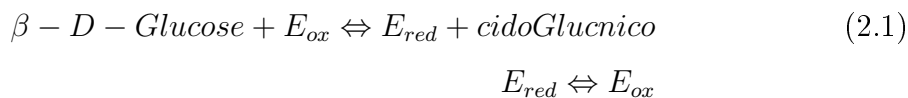
2.1.5. Diabetes gestacional

La diabetes gestacional es la intolerancia a los hidratos de carbono de severidad variable que se diagnostica por primera vez durante la gestación y se caracteriza por la hiperglicemia. Según la organización mundial de la salud estima que en el mundo existen 347 millones de personas con diabetes, Se estima que las muertes de pacientes con este padecimiento aumenten considerablemente entre los años 2005 y 2030. Convirtiéndose en una de las principales causas de muerte a nivel mundial, donde el 80 porciento de estas muertes se registran en países de ingresos bajos a medios [World Health Organization \(2013\)](#). Las personas con diabetes mellitus necesitan llevar un constante monitoreo de sus niveles de glucemia.

2.2. Glucosa Oxidasa

Una enzima es una proteína producida por las células de un organismo que se une a un componente de un sustrato la cual es encargada de regular y facilitar las reacciones químicas en los seres vivos. Ésta es comúnmente utilizada para medir los niveles de glucosa en la sangre en fluidos biológicos ([Bankar *et al.*, 1999](#)).

La glucosa oxidasa (GOD) es una enzima de la clase oxirreductasa que cataliza la conversión de beta-D-glucosa y oxígeno a D-glucono-1, 5-lactona y peróxido ([Bankar *et al.*, 1999](#)). Es una flavoproteína, altamente específica para la Beta-D-glucosa, ver Ecuación 2.1:



La principal fuente de producción de GOD es *Aspergillus*, *Penicillium notatum* y otros hongos como *Saccharomyces* y tiene la actividad antibacteriana en la presencia de glucosa y oxígeno que tienen la capacidad de fermentar carbohidratos. La GOD se utiliza para estimar la concentración de glucosa en muestras de sangre y orina a través de la formación de tintes coloreados por el peróxido de hidrógeno producido en la reacción ([Enzyme Nomenclature, 1992](#)).

La literatura muestra que existen diferentes métodos analíticos utilizados para la producción de esta enzima. Microbiólogos de la Facultad de Biología de la Universidad de La Habana, lograron aislar esta enzima a partir del *Penicillium Funiculosum* (PF), un hongo que sintetiza la glucosa oxidasa (GOD) en un medio adecuado; además, como resultado de la fermentación se obtiene también ácido glucónico (AG) ([B. Valiño & Fernández., 2003](#)). Para la descripción de este fenómeno se construyeron sistemas de ecuaciones diferenciales ordinarias descritas de manera general y suposiciones de algunas ecuaciones muy aproximadas a su experimento considerando los de carácter biológico.

2.2.1. Aplicaciones de la glucosa oxidasa.

Las aplicaciones que se le atribuyen a la glucosa oxidasa durante los últimos años principalmente en el ámbito de la medicina es el monitoreo de niveles de glucosa en pacientes con diabetes mellitus para el control de esta enfermedad mediante glucómetros. La GOD es la enzima que utiliza el biosensor estudiado.

Existen diferentes biosensores que utilizan GOD ([Bankar *et al.*, 1999](#)):

- Tiras de biosensor para sangre (GOD-HPR-dye).
- Biosensor de fibra óptica para analizar las concentraciones de glucosa en bebidas suaves
- Biosensor en tiras desechables para sangre y monitoreo de suero

- Sensor de glucosa para la sangre
- Biosensor de glucosa para el suero de la sangre humana.

2.2.2. Modelación de biosensores de glucosa oxidasa

Actualmente existen diversos métodos para medir los niveles de glucosa en la sangre. Los más populares son mediante el uso de dispositivos pequeños conocidos como glucómetros que tienen el fin de medir el número de electrones que pasan a través de la enzima en el electrodo y a su vez mide la carga resultante. También existe una manera alternativa de medir la glucosa, y que es con el control de cetonas en la orina, esto en el caso de que se vuelva imposible la detección mediante la sangre, aunque éstos son menos precisos. Las cetonas son un ácido que descompone la grasa para convertirla en energía (Bankar *et al.*, 1999).

2.3. Sensores y biosensores

Los sensores son dispositivos transductores que transforman la información física o química en una señal eléctrica útil que pueda ser procesada y, por tanto, que facilite información de interés de una manera rápida y sin necesidad de análisis muy complejos. Los sensores pueden encontrarse en el cuerpo humano naturalmente. Entre los campos más importantes podemos encontrar a los sensores químicos teniendo una mayor implantación en la biomedicina.

Entre este tipo de sensores destacan los biosensores, ya que son los dispositivos con más selectividad (Luna, Noviembre 2008). Un biosensor es un instrumento o dispositivo analítico compacto que contiene un elemento de detección biológico o bioreceptor, que usualmente es una enzima, microorganismos, tejidos y organelas, inmunoreceptores o quimiorreceptores, y esta acoplado a un transductor físico-químico que convierte la señal biológica en una señal electrónica. Esta señal es proporcional al compuesto o grupo de compuestos que se quiere determinar y puede recoger información cualitativa o cuantitativa. Las enzimas son usualmente utilizadas como bioreceptores, debido a la capacidad que poseen para reconocer una molécula específica. En el ámbito de los sensores enzimáticos los biosensores amperométricos son los más utilizados y estudiados sobre todo el sensor de glucosa

2.3 Sensores y biosensores

es el de mayor aplicación (Gatti, 2004). Una de las técnicas que actualmente se encuentran en desarrollo es la medición de la glucosa mediante biosensores basados en sensores electroquímicos enzimáticos.

El desarrollo de biosensores comenzó en el año de 1962 con electrodos enzimáticos por Leland C. Clark. La característica que distingue a los biosensores con otros métodos analíticos es que el analito puede ser medido instantáneamente y directamente sin necesidad de procesos lentos y de equipo de laboratorio, además de ser dispositivos de alta sensibilidad y confiabilidad.

El funcionamiento de un biosensor consiste en el analito o muestra a analizar el cual es un compuesto biológico de reconocimiento, que se pone en contacto con el transductor. Este contacto produce una serie de cambios físico- químicos lo cual da como resultado la conversión a una señal eléctrica cuantificable o semi-cuantificable que puede ser interpretada clínicamente (Antonio & Pérez, 2004-2005).

A continuación se presenta un esquema del funcionamiento de un biosensor –ver figura 2.3.

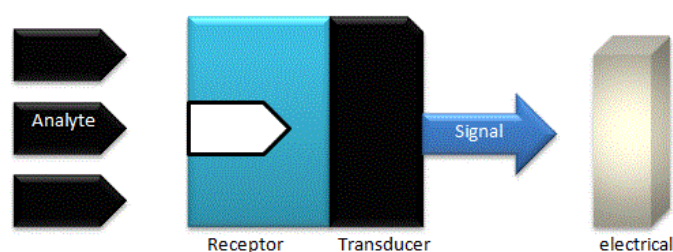


Figura 2.1: Esquema de un biosensor.

El analito es reconocido por el receptor biológico donde se lleva a cabo un proceso de conversión a señal eléctrica mediante el transductor. Los compuestos biológicos pueden ser químicos: hormonas, electrolitos, biomarcadores; o Eléctricos: neurológicos, enzimas.

El correcto funcionamiento de un biosensor depende de la inmovilización del sistema de reconocimiento biológico en el transductor. La inmovilización se refiere a la confinación física o localizada en una región definida de espacio con retención

de sus actividades catalíticas, y cada una de ellas puede ser usada repetidamente y constanmente (Guisan, 2006).

2.4. Tipos de biosensores.

En esta sección se listará la evolución que han tenido los biosensores.

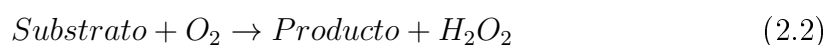
2.4.1. Electroquímicos

El biosensor electroquímico combina las técnicas electroquímicas con la selectividad de los procesos de reconocimiento biológico. La selectividad del biosensor depende del elemento biológico y la sensibilidad se le atribuye al transductor (Reviejo & Pingarrón, 2000).

Son conocidos principalmente por la aplicación de detección de glucosa y la hibridación de tiras de ADN. Los biosensores electroquímicos miden la corriente de un electrodo contenido en el compuesto biológico en base a uno de referencia. Dentro de los biosensores electroquímicos se agrupan tres tipos de sensores (Alegret *et al.*, 2004):

- Voltimétricos
- Potenciométricos
- Amperométricos

Un biosensor amperométrico es un tipo de biosensor electroquímico el cual se caracteriza por la alta sensibilidad y confiabilidad. Un biosensor enzimático es aquel en donde su elemento de reconocimiento es una enzima, dicho biosensor trabaja con el consumo de oxígeno, producción de peróxido de hidrogeno u obtención de b-nicotinamida adenina durante el proceso de la conversión catalítica del sustrato –ver ecuaciones 2.2 y 2.3 .



2.4 Tipos de biosensores.



Dentro de este campo los más estudiados son los biosensores amperométricos enzimáticos. En la literatura podemos encontrar diferentes tipos de elementos que son utilizados en un biosensor los cuales se muestran en la siguiente tabla- 2.1:

Receptor	<i>Seafsico – qumica</i>	Transductores
Tinte	Proteína	Detector Optoelectrónico
Lectina	Sacáridos	transistores de efecto campo
Enzima	Glicoproteína	Electrodo semiconductor
Apoenzima	Substrato	Electrodo Potentiométrico
Anticuerpo	Inhibidor	Electrodo amperométrico
Organelas	Group prostético	Termistores
Microbio	Antígeno	
Tejidos	Hormonas	

Tabla 2.1: Elementos de los diferentes tipos de biosensores

En la tecnología de biosensores podemos encontrar diversos tipos, los cuales con clasificados de la siguiente manera:

2.4.2. Biosensores de primera generación

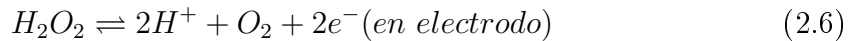
Están basados en la medida de alguno de los productos de la reacción enzimática de una sustancia con propiedades electroquímicas. En 1973, Guilbault y Lubrano desarrollaron un nuevo tipo de biosensor de primera generación basado en la oxidación del agua oxigenada, generada por la reacción enzimática con la utilización de mediadores, lo que dió paso a la segunda generación de biosensores amperométricos. Estos tenían desventajas como por ejemplo, la oscilación en la concentración de oxígeno influía en la señal de salida. La oxidación electroquímica de H_2O_2 no era selectiva en presencia de algunos componentes orgánicos acompañados de glucosa, ya que esta oxidación inhibía a la glucosa, entonces para resolver estos problemas se lograron construir sensores más capaces empleando membranas, circuitos diferenciales, etc., en los cuales se logró encontrarles aplicación analítica (Stoytcheva *et al.*, 2004). El electrodo de Clark pertenece a este grupo, consiste en una membrana compuesta por glucosa oxidasa inmovilizada

sobre un electrodo que determina variaciones de oxígeno. Ésta concentración de glucosa es cuantificada mediante reducción de oxígeno:



Esta reacción enzimática consume oxígeno y produce peróxido de hidrógeno. Aquí el peróxido producido sirve como base para la medida de la glucosa. El peróxido determina, a través de la corriente generada, cuando un potencial de aproximadamente 900 mV se aplica a través del compósito grafito-epoxi ([Alegret et al., 2004](#)).

Las ecuaciones que describen la dinámica de este tipo de biosensores son:



2.4.3. Biosensores de segunda generación

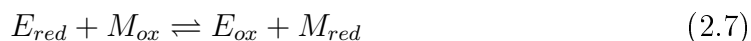
Los biosensores de segunda generación datan desde 1984 desarrollado por Cass y Davis que tenían como mediador al ferroceno. En los biosensores de segunda generación se incorpora un mediador, que se encarga de la transferencia electrónica entre el centro activo de la enzima y la superficie del electrodo. El mediador reacciona rápidamente junto con el centro activo. Además de las enzimas los mediadores más utilizados son, entre otros, el par ferri/ferrocianuro, 1,4-benzoquinona, derivados del ferroceno, TTF, TCNQ, Meldola's blue y azul de metileno ([AEG Cass, 1984](#)).

También se pueden usar polímeros que son conductores, que contienen macromoléculas orgánicas que pueden modificarse para satisfacer varias aplicaciones. Estos polímeros conductores se han utilizado en sensores químicos que se utilizan para reconocimiento. La glucosa oxidasa puede quedar atrapada ¹ en una

¹Consiste en la retención física del enzima en las cavidades interiores de una matriz sólida porosa

de las membranas del poliperrol durante la electrodeposición del polímero, en un electrodo inerte como el platino. Esto se lleva a cabo aplicando un potencial constante a través de una solución tamponada que contiene el monómero del pirrol y la enzima (Reviejo & Pingarrón, 2000). La membrana resultante formada en el electrodo puede usarse como base reactiva de un biosensor amperométrico de glucosa que constituye el electrodo del biosensor, y que se utilizará en el instrumento amperométrico(AEG Cass, 1984).

Un biosensor de segunda generación consta de dos partes, un elemento de reconocimiento que interactúa con el analito y concede selectividad al sensor, y el transductor que permite convertir esa interacción en una señal analítica. Si el transductor es un electrodo entonces el sensor es un electroquímico, lo que hace que sea un biosensor. Debido a los problemas presentados en los biosensores de primera generación se realizaron investigaciones en donde se integraron métodos de inmovilización de mediadores lo que le daba transferencia de electrones entre el centro activo de la enzima y el electrodo. La ecuación de reacción es la siguiente (Reviejo & Pingarrón, 2000):



2.4.4. Biosensores de Glucosa Oxidasa

En la literatura científica se pueden encontrar diversos trabajos e investigaciones de biosensores utilizando glucosa oxidasa que han dado lugar al desarrollo de biosensores electroquímicos. Los científicos de la universidad de Guangongdng en China estudiaron la preparación de un nanocompuesto de electrodo híbrido modificado llamado grafeno-quitosano, el cual permite por medio de cuya superficie de carbón, la mejor absorción de la glucosa oxidasa. Después de un escaneo microscópico se determinó la capacidad de este compuesto para formar un ambiente adaptable de GOD. Dadas las características de adaptabilidad del compuesto, es

comúnmente empleado para la detección de glucosa en biosensores amperométricos. Por otra parte, [nez *et al.* \(2004\)](#) desarrollaron de un biosensor amperométrico de glucosa basado en un biocomposito disperso de grafito-epoxy-platino-glucosa oxidasa con el que se aprovechan las propiedades electrocatalíticas para contruir un biosensor amperométrico de glucosa de primera generación.

[nez *et al.* \(2011\)](#) desarrollaron un biosensor de Glucosa basado en un Biocompósito disperso de Grafito-Epoxi-Platino-Glucosa Oxidasa el cual consistió en desarrollar un biosensor de glucosa basado en las propiedades electroquímicas de un compuesto grafito-epoxi-platino-glucosa oxidasa el cual se fundamenta en la caracterización de la respuesta del biosensor por voltamperometría de barrido lineal y amperometría.

2.4.5. Modelación Matemática de Biosensores

Actualmente los biosensores debido a sus propiedades ya descritas se han convertido en elementos indispensables en la ciencia, la química y la biotecnología, es por ello que su moderación por medio de diversos métodos es de suma importante en las etapas de diseño y desarrollo [Baronas *et al.* \(2009\)](#).

Los biosensores amperométricos se modelan suponiendo que no existen límites de difusión externa y la conversión del compuesto en la membrana biocatalítica sigue las reacciones cinéticas de primer orden. La conversión del sustrato es consecutiva, cíclica y paralela la cual se calcula de la siguiente manera:



Donde S es el sustrato, Z el intermediario, el producto P es electroquímicamente activo, y $K1$ y $K2$ son las constantes de velocidad de primer orden $\frac{V_{max}}{KM}$ de la conversión enzimática de los compuestos. El transporte de la masa y la concentración consecutiva de sustratos en la membrana biocatalítica se describe de acuerdo al siguiente sistema de ecuaciones ([Alegret *et al.*, 2004](#)):

$$\frac{\delta S}{\delta t} = D_e \frac{\delta^2 S}{\delta x^2} - k_1 S, \quad (2.10)$$

$$\frac{\delta Z}{\delta t} = D_e \frac{\delta^2 Z}{\delta x^2} - k_2 Z + k_1 S. \quad (2.11)$$

$$\frac{\delta P}{\delta t} = D_e \frac{\delta^2 P}{\delta x^2} - k_2 Z, \quad (2.12)$$

Donde S, Z, P corresponde a las concentraciones del sustrato, el intermediario del producto y el producto, respectivamente. Las ecuaciones iniciales ($S = 0, Z = 0, P = 0$ cuando $0 < -x < d, t = 0$) y el límite ($S = S_0, Z = 0, P = 0$ cuando $0 < x < d, t = 0$; $S/Dx = 0, \delta Z/\delta x = 0, P = 0$ cuando $x = 0, t = 0$). Estas condiciones conforman el problemas de límites de valores inicial.

Los biosensores amperométricos, los cuales son el objeto de estudio de esta investigación, miden la corriente Faradaica que surge en un electrodo indicador de oxidación electroquímica o la reducción de productos de una reacción bioquímica. Estos biosensores son conocidos por tener fiabilidad, ser baratos y su alto grado de sensibilidad en usos clínicos e industriales (Reviejo & Pingarrón, 2000).

Debido a que los biosensores son utilizados en procedimientos de síntesis combinatoria, el producto puede llegar a ser en cantidades muy escasas y limitadas que se llegan a medir en microgramos a miligramos, además los sustratos dentro de las enzimas no pueden ser medidos con dispositivos analíticos. Para esto, desde los años setentas se han desarrollado y utilizado modelos matemáticos de biosensores amperométricos, como una herramienta importante en el estudio de las características de los biosensores actuales con mejor precisión y sencillez (Reviejo & Pingarrón, 2000).

2.4 Tipos de biosensores.

Una tendencia en la biotecnología, debido al micro volumen en que se desenvuelve el campo de los biosensores, es la miniaturización de éstos. Se han realizado investigaciones de modelos basados en ecuaciones de difusión que contienen en un término no lineal relacionado con la cinética de Michaelis-Menten, que muestra la velocidad de la reacción enzimática con el objetivo de tener sistemas de biosensores basados en arreglos de enzimas microreactoras inmovilizadas por un solo electrodo. El modelo matemático $S \rightarrow E \rightarrow P$ considerando que (S) es el sustrato que es enzimáticamente convertido en el producto (P) (Baronas *et al.*, 2006).

Este modelo matemático consiste en que el producto catalizado por la enzima la tasa de reacción disminuye con el incremento de la concentración del sustrato el cual actúa con un efecto negativo. La ecuación cinética de la inhibición del sustrato es la siguiente : (Baronas *et al.*, 2006)

$$\frac{\delta S}{\delta t} = D_e \frac{\delta^2 S}{\delta x^2} - V_{max} S, \quad (2.13)$$

$$\frac{\delta P}{\delta t} = D_e \frac{\delta^2 P}{\delta x^2} + \frac{V_{max} S}{K_M + S(1 + s/k_s)}, x \in (, d), t > 0 \quad (2.14)$$

Donde x y t están fijas por un espacio de tiempo respectivamente, $S(x, t)$ es la concentración del sustrato, $P(x, t)$ es la concentración de la reacción del producto, d es el espesor de la capa de la enzima, D_s , D_p son los coeficientes de difusión, V_{max} es la tasa máxima enzimática, K_m y K_s son las constantes de ecuación de Michaelis y la inhibición del sustrato, respectivamente (Baronas *et al.*, 2006).



donde S es el sustrato combinado con la enzima E el cual se transforma en un complejo ES que se disocia con el Producto P y genera la enzima.

La modelación matemática representa una herramienta importante en los biosensores electroquímicos por que permiten una reducción considerable del costo y tiempo de diseño y desarrollo (Wang, 2011). Una opción alterna a las diversas formas de modelado de estos dispositivos es aquella basada en técnicas de aprendizaje de máquina, ya que su esquema de caja negra permite modelar la respuesta del biosensor evitando el uso de complicadas ecuaciones matemáticas. Entre los algoritmos de aprendizaje más utilizados se encuentran las Maquinas de Vectores de Soporte (SVM) y las Redes neuronales (RN). Ambas fueron utilizadas en el presente trabajo. Otros algoritmos que pueden ofrecer una aceptable respuesta esta Mínimos Cuadrados Parciales (PLS por sus siglas en ingles), Arboles de Regresión, Regresión de ridge y LASSO (Gatti, 2004).

2.5. Aprendizaje de máquina

El aprendizaje de máquina o aprendizaje automático (ML) es una rama multidisciplinaria de la inteligencia artificial que a través de técnicas, algoritmos y mediante ciertos sucesos, acciones o información se adapta para permitirle aprender a la computadora (Marsland, 2011). En el ML el aprendizaje puede ser de dos tipos, supervisado que se basa en un conocimiento previo el cual es independiente de la experiencia; mientras que el otro es no supervisado, que trata las variables de entrada como un conjunto aleatorio. Entre los principales algoritmos de ML se encuentran las redes neuronales, máquinas de vectores de soporte, mínimos cuadrados parciales, modelos generalizados lineales, árboles de regresión (Zheng & Wang, 2006). En las siguientes subsecciones se explicarán los principales modelos de regresión utilizados en este trabajo de investigación.

2.6. Máquinas de vectores de soporte

Las máquinas de vectores de soporte (SVM por sus siglas en inglés) fueron introducidas a principios de los años 90s y desarrolladas por Vladimir Vapnik y un grupo de científicos de los laboratorios de ATT. Originalmente fueron diseñadas para resolver problemas de clasificación, pero recientemente se han extendido al ámbito de los problemas de regresión (Vapnik & Smola, 1997). Las aplicaciones

2.6 Máquinas de vectores de soporte

más comunes en las que se destaca el uso de las SVM se encuentran el reconocimiento de dígitos, procesamiento de lenguaje natural, biología computacional, visión por computadora, minería de datos, diseño de medicamentos, reconocimiento de discursos y traducción, entre muchas otras.

Dado un conjunto de datos de entrenamiento (x_i, y_i) , $i = 1, \dots, l$ donde $x_i \in \mathbb{R}^n$ y $y \in \{1, -1\}^l$ se requiere resolver la optimización del problema (Chih-Wei Hsu & Li, 2010):

$$\min_{w, b, \xi} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^l \xi_i \quad (2.16)$$

sujeto a

$$y_i (\mathbf{w}^T \phi(x_i) + b) \geq 1 - \xi_i,$$

$$\xi_i \geq 0$$

De esta manera los vectores x_i son mapeados en un espacio de gran dimensión dimensional por la función ϕ . En dicho espacio la SVM busca un hiperplano de separación lineal con el margen máximo. $C > 0$ es el parametro de penalización del término del error. La función kernel es definida como $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ (Hastie *et al.*, 2001). Existen cuatro muy bien kernels propuestos en la literatura:

- Lineal: $K(x_i, x_j) = x_i^T x_j$.
- Polinomial: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$.
- Función de base radial RBF: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$
- Sigmoides: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$

Por otro lado para satisfacer las restricciones de minimización se utiliza la técnica de multiplicadores de Lagrange α , donde $\alpha_i \geq 0 \forall_i$ (Fletcher, 2009):

$$L_P \equiv \frac{1}{2} \|\mathbf{w}\|^2 - \alpha [y_i (x_i \cdot \mathbf{w} + b) - 1 \forall_i] \quad (2.17)$$

$$\equiv \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^L \alpha_i [y_i (x_i \cdot \mathbf{w} + b) - 1] \quad (2.18)$$

$$\equiv \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^L \alpha_i y_i (x_i \cdot w + b) \sum_{i=1}^L \alpha_i \quad (2.19)$$

Para encontrar la w y b que minimiza, y α que maximiza la ecuación 2.18, se necesita derivar L_P con respecto a w y b e igualando a cero:

$$\frac{\delta L_P}{\delta w} = 0 \Rightarrow w = \sum_{i=1}^L \alpha_i y_i x_i \quad (2.20)$$

$$\frac{\delta L_P}{\delta b} = 0 \Rightarrow \sum_{i=1}^L \alpha_i y_i = 0 \quad (2.21)$$

Sustituyendo 2.20 y 2.21 en 2.19 da una nueva formulación para que sea dependiente de α se necesita maximizar:

$$L_D \equiv \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i \cdot x_j \quad (2.22)$$

sujeto a $\alpha_i \geq 0 \forall_i, \sum_{i=1}^L \alpha_i y_i = 0$

$$\equiv \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i H_{ij} \alpha_j \quad (2.23)$$

Donde $H_{ij} \equiv y_i y_j x_i \cdot x_j$

$$\equiv \sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T H \alpha \quad (2.24)$$

Sujeto a $\alpha_i \geq 0 \forall_i, \sum_{i=1}^L \alpha_i y_i = 0$

Esta nueva formulación L_D es conocida como la formal Dual del L_P primario. Habiendo pasado de la minimización L_P a la maximización L_D se necesita encontrar (Fletcher, 2009):

$$\max_a \left[\sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T H \alpha \right] \quad (2.25)$$

sujeto a $\alpha_i \geq 0$ y $\sum_{i=1}^L \alpha_i y_i = 0$

Esta es un problema de optimización cuadrática convexa, el cual se resuelve con cualquier programa para resolver problemas de tipo QP el cual regresará α y 2.20 nos devolverá w . Entonces solo queda calcular b

Cualquier punto de los datos satisface 2.21 el cual es un vector de soporte x_s y tendrá la forma:

$$y_s (x_s \cdot w + b) = 1 \quad (2.26)$$

Substituyendo en 2.20:

$$y_s \left(\sum_{m \in S} \alpha_m y_m x_m \cdot x_s + b \right) = 1 \quad (2.27)$$

Donde S denota el conjunto de índices de los vectores de soporte. S es determinado por encontrar los índices i donde $\alpha_i > 0$. Multiplicando por y_s y luego usando $y_s^2 = 1$ de $x_i \cdot w + b \geq +1$ por $y_i = +1$ $y_s^2 = 1$ de $x_i \cdot w + b \leq -1$ por $y_i = -1$ (Fletcher, 2009)

$$y_s \left(\sum_{m \in S} \alpha_m y_m x_m \cdot x_s + b \right) = y_s \quad (2.28)$$

$$b = y_s - \sum_{m \in S} \alpha_m y_m x_m \cdot x_s \quad (2.29)$$

En lugar de utilizar un vector de soporte x_s arbitrario, es mejor tomar un promedio de todos los vectores de soporte en S (Fletcher, 2009):

$$b = \frac{1}{N_s} \sum_{s \in S} \left(y_s - \sum_{m \in S} \alpha_m y_m x_m \cdot x_s \right) \quad (2.30)$$

Como resultado tenemos las variables w y b que definen la orientación óptima del hiperplano de separación y nuestra SVM. Para el caso de una SVM de regresión, que es lo que ocupa esta tesis, el procedimiento es similar. La variación radica en la penalización de la diferencia entre el valor predicho y el valor verdadero dado por $\|t_i - y_i\| \leq \xi_i$. De esta manera, la función de error estará dada por:

$$C \sum_{i=1}^l (\xi^+ + \xi^-) + \frac{1}{2} \|w\|^2 \quad (2.31)$$

2.7. Redes Neuronales para Regresión

Las redes neuronales (RNA) son un método de inteligencia artificial el cual se inspira en la analogía del funcionamiento del cerebro humano basado en la capacidad de aprendizaje a partir de la experiencia obtenida de casos anteriores para generar conocimiento. Las redes neuronales artificiales se expresan habitualmente mediante un grafo definido con diferentes tipos de nodos y conexiones –ver Figura 2.7. Consiste en el procesamiento de un conjunto de unidades llamados nodos los cuales están interconectados entre sí a través de su topología. Los nodos que no tienen una entrada se les conoce como los nodos de entrada y los que no tienen conexiones hijas se les llama nodos de salida.

Debido a sus ventajas, las RNA ofrecen numerosas aplicaciones en diferentes campos, principalmente en la biología, medio ambiente, medicina, finanzas, manufactura, y aplicaciones militares. El objetivo del aprendizaje consiste en la determinación de los valores precisos de los pesos para todas las conexiones que permita resolver de la forma más eficiente un problema. En términos generales consiste en introducir de forma gradual e ir modificando los pesos de las conexiones basado en un determinado esquema de aprendizaje (nuela & Galván, 1995).

El entrenamiento de una red neuronal depende de el número de capas ocultas. Del número de neuronas que tiene cada capa oculta, del algoritmo de aprendizaje que se utiliza, de la topología empleada y del tipo de función de activación que se escoge para cada neurona. La construcción de una red implica determinar la

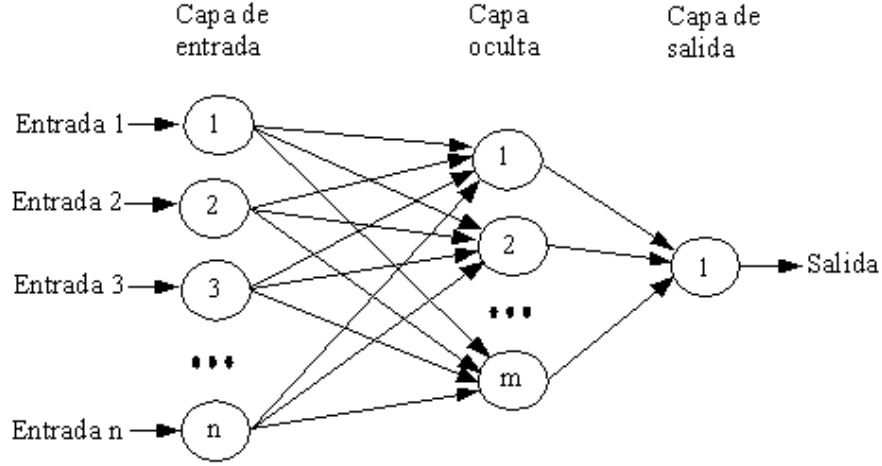


Figura 2.2: Esquema funcionamiento de una red neuronal. Las neuronas de las capas de entrada representan cada atributo de los datos, las neuronas de la capa de salida representa la predicción de la red, mientras que las neuronas de las capas ocultas permiten que la red aprenda

estructura de una red con sus respectivos pesos de las interconexiones que darán como resultado una aproximación de funciones (nuela & Galván, 1995).

La regresión de una variable dependiente Y en una variable independiente X , es el cálculo de un valor más probable de Y para cada valor de X el cual está basado en un número finito de mediciones de X y los valores asociados de Y , donde X y Y son vectores de salida y de entrada, respectivamente. A medida que se implementa, la red puede estimar valores de Y para cualquier nuevo valor de X .

La RNA se inicia con pesos iniciales aleatorios. La topología mayormente utilizada es la de tipo feed-forward. La no lineal están basadas en combinaciones lineales de funciones base no lineales compuestas $\phi_j(x)$ de la forma 2.32:

$$y(x, w) = f\left(\sum_{j=1}^M w_j \phi_j(x)\right) \quad (2.32)$$

donde $f(\cdot)$ es una función de activación no lineal, ésta es la función identidad en la regresión.

2.7 Redes Neuronales para Regresión

Las RNAs pueden dividirse en diferentes tipos en base a la estructura de sus conexiones, dentro de las cuales las redes de conexión hacia adelante (feed-forward) son las mayormente utilizadas. En dichas redes existe una forma de numerar los nodos de la red de tal manera que no existe ninguna conexión de un nodo con mayor número a un nodo con menor número. Las redes feedforward permiten computar funciones escalares con una sola capa oculta. A cada nodo j se le da una variable x_j caracterizado por un estado de activación y_i , mientras que a cada conexión (i, j) se le asocia un peso $w_{ij} \in \mathfrak{R}$, también existe una función de salida f_i la cual está encargada de convertir el estado actual de la activación en una señal de salida $f_i(y_i)$, tomando un valor numérico que va generalmente de $[0, 1]$ o bien $[-1, 1]$. Cuando el algoritmo de aprendizaje tiene una función de base radial, casi cualquier problema puede ser representado con tan solo una o dos capas ocultas con sus respectivas neuronas (nuela & Galván, 1995).

Los modelos de regresión de una red feed-forward son basados en combinaciones lineales de funciones base combinadas no lineales $\phi_j(x)$ (Bishop, 2009).

$$y(x, w) = f\left(\sum_{j=1}^M w_j \phi_j(x)_i\right) \quad (2.33)$$

Donde $f(\cdot)$ es una función identidad. Se produce una serie de combinaciones lineales y no lineales M de las variables de entrada de la siguiente manera:

$$a_j = \sum_{i=1}^D w_{ji}^{(1)} x_i + w_{j0}^{(1)} \quad (2.34)$$

Donde $j = 1, \dots, M$ y el superíndice (1) son los que indican los parámetros correspondientes a la primera capa, los parámetros $w_{ji}^{(1)}$ son los pesos y $w_{j0}^{(1)}$ son los bias.

Dentro de los algoritmos supervisados también encontramos las RNAs de retro-propagación. Para una red entrenada se requiere un patrón de entrada el cual está encargado de transmitir a éste a través de las capas de neuronas hasta que se obtiene una salida, seguido de una etapa de entrenamiento en el que se modifican los pesos de la red (Bishop, 2009).

2.8. Árboles de clasificación y regresión (CART)

Los árboles de clasificación y regresión (CART) son métodos de particionamiento recursivo dentro del aprendizaje automático para la construcción de modelos de predicción a partir de los datos. El resultado es mostrado gráficamente como un árbol de decisión con las asociaciones mas importantes. En el árbol de regresión la variable Y toma los valores ordenados y el modelo de regresión es adecuado a cada nodo para obtener una predicción de los valores de Y . CART tiene la capacidad de demostrar estadísticamente cuales son los factores más importantes en el modelo. La ventaja que tiene utilizar un árbol de regresión es que no considera ninguna hipótesis, lo que lo hace computacionalmente rápido, además de una sencilla representación gráfica. El objetivo de un árbol de regresión es obtener respuestas numéricas que son fáciles de predecir para los nuevos datos.

El proceso estadístico de la clasificación y la regresión en el análisis del árbol es muy similar. Mientras que en los árboles de clasificación los datos se agrupan en clases regularmente binarias por la variable de respuesta, en los árboles de regresión la variable de respuesta es numérica o continua que se utilizan para predecir el resultado. El algoritmo clásico de CART fué popularizado por Breiman en el contexto de las facilidades del análisis de los árboles de clasificación.

El Árbol de regresión o clasificación se construye primeramente particionando el nodo raíz de la muestra en nodos binarios basado en la forma:

$$X \leq d \tag{2.35}$$

donde X variable del conjunto de datos d es un número real

Cada una de las variables independientes debe reducirse al mínimo por lo que emplea el criterio de la división de la suma de cuadrados. El nodo heterogéneo contiene todas las observaciones de las localidades y el objetivo fundamental es convertir esta localidades en un nodos homogéneos binarios. Por lo que CART selecciona las mejores divisiones para crear nodos secundarios. El proceso se repite recurrentemente hasta conseguir el mayor árbol de máximo nodos posibles. Del árbol máximo se pueden obtener el árbol óptimo utilizando un algoritmo conocido

como poda del árbol el cual aplica validación cruzada (Breiman & Friedman, 1984).

La construcción de un árbol de regresión y clasificación CART consiste en 3 simples pasos:

1. Construcción de un árbol máximo el cual es un modelo informativo o una visualización final de las asociaciones y relaciones mas importantes. Se parte de la matriz de datos donde n es el número de observaciones. Utilizando un procedimiento de partición binario, comenzando en la raíz del árbol hasta que hasta llegar al nodo final.
2. Poda del árbol. El árbol obtenido T_{max} es generalmente sobre-ajustando mientras está siendo podado para adecuar el tamaño del árbol.
3. Selección del árbol óptimo mediante un procedimiento de validación cruzada.

2.9. Modelo Generalizado lineal GLM

Modelo generalizado lineal formulado por John Nelder y Robert Wedderburn se refiere a una generalización de una regresión lineal basada en la distribución normal multivariada. Un modelo lineal se compone básicamente de un componente aleatorio que identifica la variable de respuesta junto con su distribución de probabilidad . una variable aleatoria Y con observaciones independientes $(y_1, y_2, \dots, y_i)$ donde cada Y_i indica el número de éxitos de entre un número fijo de ensayos modelada como una distribución binomial. Los modelos se encuentran en el grupo de exponencial de distribución:

$$f(y_i|\theta_i) = \alpha(\theta_i) \cdot b(y_i) \cdot \exp[y_i Q(\theta_i)] \quad (2.36)$$

Donde $Q(\theta)$ es el parámetro natural Seguido de un componente sistemático que es encargado de de especificar variables de predicción que se utilizan en una función predictora lineal. Las variables x_j se relacionan de la forma:

$$\alpha + \beta_i x_i + \dots + \beta_k x_k \quad (2.37)$$

2.9 Modelo Generalizado lineal GLM

El cual es denominado como el predictor lineal que también puede ser expresado en un vector $(\eta_1, \dots, \eta_N)$:

$$\eta_i = \sum_j \beta_j x_{ij} \quad (2.38)$$

Donde x_{ij} es el j -ésimo predictor en el i -ésimo individuo, e $i = 1, \dots, N$. Por último consta de una función link la cual se refiere a la función del valor esperado de Y , $E(Y)$ la cual es una combinación lineal de las variables de predicción. La función link se encarga de especificar $g(\cdot)$ para vincular μ con el predictor lineal denotada de la siguiente manera:

$$g(\mu) = \alpha + \beta_1 x_1 + \dots + \beta_k x_k \quad (2.39)$$

La función link relaciona los componentes de aleatoria y sistemática

$$\mu_i = E(Y_i) \quad (2.40)$$

y

$$\eta_i = g(\mu_i) = \sum_j \beta_j x_{ij} \quad (2.41)$$

$\forall i = 1, \dots, N$ donde $g(\mu)$ es la identidad para dar el modelo de regresión lineal

$$\mu = E(Y) = \alpha + \beta_1 x_1 + \dots + \beta_k x_k \quad (2.42)$$

la media μ de la distribución depende de las variables x por medio de la fórmula:

$$E(Y) = \mu = g^{-1}(X\beta) \quad (2.43)$$

Donde $E(Y)$ es el valor esperado de $\mathbf{Y}$, $\mathbf{X}$ y β es el predictor lineal de la combinación de parámetros desconocidos β y g es la función de enlace. La varianza es típicamente una función $\mathbf{V}$ de la media

$$Var(Y) = V(\mu) = V(g^{-1}(X\beta)) \quad (2.44)$$

2.10. Modelo Lineal Generalizado GLM

El Modelo Lineal Generalizado fue formulado por John Nelder y Robert Wedderburn, y se refiere a una generalización de una regresión lineal basada en la distribución normal multivariada. Un modelo lineal se compone de un elemento aleatorio que identifica la variable de respuesta junto con su distribución de probabilidad; y una variable aleatoria Y con observaciones independientes $(y_1, y_2, \dots, y_i)$ donde cada Y_i indica el número de éxitos de entre un número fijo de ensayos modelada como una distribución binomial. Los modelos se encuentran en el grupo de distribuciones exponenciales:

$$f(y_i|\theta_i) = \alpha(\theta_i) \cdot b(y_i) \cdot \exp[y_i Q(\theta_i)] \quad (2.45)$$

Donde $Q(\theta)$ es el parámetro natural deguido de un componente sistemaático que es encargado de especificar variables de predicción que se utilizan en una función predictora lineal. Las variables x_j se relacionan de la forma:

$$\alpha + \beta_i x_i + \dots + \beta_k x_k \quad (2.46)$$

El cual es denominado como el predictor lineal que también puede ser expresado en un vector $(\eta_i \dots, \eta_N)$:

$$\eta_i = \sum_j \beta_j x_{ij} \quad (2.47)$$

Donde x_{ij} es el j -ésimo predictor en el i -ésimo individuo , e $i = 1, \dots, N$. Por último consta de una función link la cual se refiere a la función del valor esperado de Y , $E(Y)$ la cual es una combinación lineal de las variables de predicción. La función link se encarga de especificar $g(\cdot)$ para vincular μ con el predictor lineal denotada de la siguiente manera:

$$g(\mu) = \alpha + \beta_i x_i + \dots + \beta_k x_k \quad (2.48)$$

2.11 Mínimos cuadrados parciales PLS

La función link relaciona los componentes aleatorio y sistemático:

$$\mu_i = E(Y_i) \quad (2.49)$$

y

$$\eta_i = g(\mu_i) = \sum_j \beta_j x_{ij} \quad (2.50)$$

$\forall i = 1, \dots, N$ donde $g(\mu)$ es la identidad para dar el modelo de regresión lineal:

$$\mu = E(Y) = \alpha + \beta_1 x_1 + \dots + \beta_k x_k \quad (2.51)$$

y la media μ de la distribución depende de las variables x por medio de la fórmula:

$$E(Y) = \mu = g^{-1}(X\beta) \quad (2.52)$$

Donde $E(Y)$ es el valor esperado de $\mathbf{Y}$, $\mathbf{X}$ y β es el predictor lineal de la combinación de parametros desconocidos β y g es la función de enlace. La varianza es típicamente una función $\mathbf{V}$ de la media:

$$Var(Y) = V(\mu) = V(g^{-1}(X\beta)) \quad (2.53)$$

2.11. Mínimos cuadrados parciales PLS

Los mínimos cuadrados parciales son diferentes a los métodos previamente analizados, estos tienen relación con la regresión de los componente principales. Es un método estadístico que tiene como objetivo encontrar una regresión lineal mediante la proyección de las variables de predicción y las variables observables a un nuevo espacio n . El modelo de PLS trata de encontrar el sentido multidimensional en el espacio de X que explica la dirección de la máxima varianza multidimensional en el espacio Y .

2.11 Mínimos cuadrados parciales PLS

La idea básica de los mínimos cuadrados parciales PLS es intentar extraer factores latentes que representen lo más posible el factor de variación manifiesto mientras que se modelan las respuestas. En particular, PLS no produce estimaciones consistentes de los llamados "variables latentes." en modelos de ecuaciones estructurales formales. El objetivo general consiste en utilizar los factores para predecir las respuestas en una población, donde T y U son los factores para predecir las respuestas en la población y las respuestas de los factores incluidos en la muestra, respectivamente. Los factores extraídos T o $X - scores$ se utilizan para predecir el $Y - scores$ U , y luego la Y son los resultados predichos que se utilizan para construir las predicciones de las respuestas.

Los mínimos cuadrados parciales están compuestos de los siguientes elementos:

1. Regresión de Componentes Principales (PCR): Los scores de X se eligen para explicar parte importante de la variación del factor como sea posible. Este enfoque produce direcciones informativas en el espacio factor, pero que no puede estar asociado con la forma de la superficie prevista.
2. Análisis de redundancia máxima: Los puntajes Y se eligen para explicar la mayor cantidad posible de variación de las Y predichas. Este enfoque busca direcciones en el espacio de los factores que están asociados con la mayoría de la variación en las respuestas, pero las predicciones pueden no llegar a ser exactas.

En modelo a solucionar de PLS esta dado de la siguiente forma:

$$X = TP^T + E \quad (2.54)$$

$$Y = UQ^T + F \quad (2.55)$$

donde X es una matriz de $n \times m$ predictores. Y es la matriz de predicción, T y U son matrices $n \times l$ que son proyecciones de X o el $X - score$ y proyecciones de Y , $Y - score$, respectivamente; P y Q son las matrices de carga ortogonales mientras que E y F se refiere a los términos del error.

2.12. Regresión Ridge

La regresión Ridge es un método de regularización conocida también como la regresión de arista. Uno de los objetivos de la regresión de ridge es producir una ecuación de regresión con coeficientes estables. El modelo de la regresión de ridge es el siguiente:

$$Y = Z\theta + \epsilon \quad (2.56)$$

donde Y es el vector de variables de respuesta, $Z = (Z_1, \dots, z_p)$ es una matriz de n observaciones en p variables predictores, θ es el vector de coeficientes, y ϵ es un vector de errores aleatorios. Esto asume que $E(\epsilon) = 0$, $E(\epsilon\epsilon^T) = \sigma^2 I$, donde I es una matriz identidad. Z_j se obtiene de la transformación de la variable original predictora X_j .

Los estimadores mínimos cuadrados por θ son $\bar{\theta} = (Z^T Z)^{-1} Z^T Y$ como se muestra:

$$E[(\hat{\theta} - \theta)^T (\hat{\theta} - \theta)] = \sigma \sum_{j=1}^p \lambda_j^{-1} \quad (2.57)$$

donde $\lambda_1 \geq \lambda \geq \dots \geq \lambda_p$, estos son los eigenvalores de $Z^T Z$ lo que lleva a error cuadrático medio total el cual sirve como medida de la distancia cuadrada de los coeficientes de regresión estimados desde sus valores reales ([Chatterjee & A.Hadi, 2006](#)). La regresión de ridge intenta construir un estimador alternativo que tenga un valor de error cuadrático medio mucho más pequeño. Los estimadores se definen mediante un parámetro k donde $k > 0$

$$\hat{\theta}(k) = (Z^T Z + kI)^{-1} Z^T Y = (Z^T Z + kI)^{-1} Z^T Z \hat{\theta} \quad (2.58)$$

El valor esperado de $\hat{\theta}(k)$ es:

$$E[\hat{\theta}(k)] = (Z^T Z + kI)^{-1} Z^T Z \theta \quad (2.59)$$

En un modelo lineal se representa Y en términos de q variables:

$$y_i = \beta_0 + \sum_{j=1}^q \beta_j x_{ij} + \epsilon_i \quad (2.60)$$

donde β_j son los parámetros y ϵ_i son las alteraciones aleatorias.

El objetivo es predecir el valor de una futura observación o estimar la media de respuesta correspondiente a una observación dada. Cuando se habla de esto, se requiere minimizar el error MSE de la predicción. Las estimaciones de los coeficientes de los modelos descritos en la Regresión lineal se basan en la independencia de los términos del modelo. Cuando los términos están correlacionados y las columnas de la matriz de diseño X tienen una dependencia lineal aproximada, la matriz $(X^T X)^{-1}$ se vuelve singular.

2.13. Regresión LASSO

La regresión LASSO (Least absolute shrinkage and selector operator por sus siglas en inglés) es un método de la inteligencia artificial, alternativo de la versión de los mínimos cuadrados que usa una restricción $\|\beta\|_1$, la L^1 - *norm* del vector de parámetros. La L^1 es utilizado en muchos contextos donde la tendencia es preferir aquellas soluciones donde sus parámetros sean menor que cero, reduciendo el número de variables en la que la solución dada es independiente. LASSO es el encargado de minimizar la suma de los residuos de los cuadrados sujeto a la suma de los valores absolutos de los coeficientes.

Supongamos que tenemos un conjunto de datos (x^l, y_i) donde $i = 1, 2, \dots, N$ $(x_{i1}, \dots, x_{ip})^T$ son las variables predictoras y y_i son las respuestas. Asumiendo que x_{ij} están estandarizadas tal que $\sum_i x_{ij}/N = 0$, $\sum_i x_{ij}^2/N = 1$. Sea $\beta = (\hat{\beta}_1, \dots, \hat{\beta}_p)^T$ el estimador LASSO $(\hat{\alpha}, \hat{\beta})$ definido por:

$$(\hat{\alpha}, \hat{\beta}) = \arg \min \left\{ \sum_{i=1}^N \left\{ (y_i - \alpha - \sum_{j=1}^p \beta_j x_{ij})^2 \right\} \right\} \quad (2.61)$$

sujeto a

$$\sum_j |\beta_j| \leq t \quad (2.62)$$

donde $t \geq 0$ es un parámetro de sintonización, $\forall t$ la solución de α es $\hat{\alpha} = \bar{y}$ entonces se omite α .

El parámetro de $t \geq 0$ se encarga de controlar el aumento del shrinkage o encogimiento aplicado a los estimados. La motivación para el LASSO viene de la propuesta de minimización de los garrotes no negativos. Dado los valores de λ el cual es un parámetro no negativo, la regresión de LASSO resuelve el problema:

$$\min_{\beta_0, \beta} \left(\frac{1}{2N} \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \sum_{j=1}^p (|\beta_j|) \right) \quad (2.63)$$

donde N es el número de observaciones, y_i es la respuesta de la observación i , x_i es el dato, un vector de p valores de la observación i , λ es un parámetro de regularización no negativo correspondiente a uno de los valores de Lambda y los parámetros β_0 y β que son escalares y el p vector respectivamente.

A medida que λ incrementa, el número de los componentes no-cero de β decremantan. Este es el punto donde el problema de LASSO involucra el problema de $L^1 - norm$ de β en contraste con el algoritmo de red elástica .

Para un α estrictamente entre 0 y 1, y no negativa λ , se resuelve el problema de la siguiente manera:

$$\min_{\beta_0, \beta} \left(\frac{1}{2N} \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 + \lambda P_\alpha \right) \quad (2.64)$$

donde

$$P_\alpha(\beta) = \frac{1-\alpha}{2} |\beta|_2^2 + \alpha |\beta|_1 = \sum_{j=1}^p \left(\frac{1-\alpha}{2} \beta_j^2 + \alpha |\beta_j| \right) \quad (2.65)$$

La red elástica es lo mismo que LASSO cuando $\alpha = 1$. Cuando α se contrae a 0, la red elástica se acerca a la regresión de RIDGE.

Para los otros valores de α , el término de penalización $P_\alpha(\beta)$ interpolan entre $L^1 - norm$ de β y el cuadrado de $L^2 - norm$ de β .

Capítulo 3

Materiales y Métodos

En este capítulo se explicará la metodología seguida para desarrollar esta investigación. Se detallarán los datos utilizados en los experimentos, además del hardware y software utilizados como herramientas para la implementación de algoritmos y pruebas.

3.1. Metodología

La Metodología utilizada es la propuesta por Duda and Hart. En la cual se muestran el diseño: (Duda & E.Hart, 1973)

1. Recolección de datos. Recopilación del conjunto de datos que posteriormente será analizado. En la investigación desarrollada, se utilizaron los datos un biosensor el cual contiene las variables de Glucosa, Benzoquinona, temperatura en grados centígrados y la acidez con notación PH como entradas. La variable a predecir es la corriente faradaica del biosensor. En la sección de descripción de datos se ampliará la información sobre éstos.
2. Selección de atributos: Una de las etapas de mayor importancia en un estudio de aprendizaje de máquina es la selección de atributos. Esta consiste en la reducción o eliminación de variables irrelevantes o redundantes hacia la variable a predecir. Debido al tamaño del conjunto de datos, i.e. cuatro variables variables de entrada, no fue necesario aplicar esta etapa de la metodología.

3. Elección de modelo. Es importante seleccionar aquellos algoritmos que mejor se adapten al problema a modelar. De esta manera, se seleccionaron algoritmos que posean el mínimo número de parámetros de ajuste, y en el mejor de los casos, sin parámetros. Los algoritmos seleccionados fueron explicados en la sección 2.5.
4. Entrenamiento de algoritmos de regresión. Esta etapa consta del entrenamiento de los algoritmos de regresión.. Para este propósito, se utilizó el método de validación cruzada en su modalidad de 30x10 –ver sección 3.4.6.
5. Evaluación de algoritmos. Para verificar el desempeño de los algoritmos de regresión, i.e. su capacidad de predecir con el mínimo error la corriente faradaica del biosensor, se utilizó la raíz cuadrada media del error normalizado o NRMSE –ver sección 3.4.6.

3.2. Hardware y Software utilizados

- El software utilizado fué con el entorno de desarrollo matemático matlab diseñado para trabajar con matrices y vectores.
- El Hardware empleado fue una computadora toshiba satellite P855-s5200, con un procesador Core i7 3610QM de 2.30Ghz, con capacidad de memoria RAM 8GB y sistema operativo 64 bit windows 7.

3.3. Descripción de los datos

Los datos que se utilizaron para realizar este trabajo consisten en un conjunto de 320 observaciones proporcionados por un biosensor amperométrico de segunda generación el cual consta de un sensor de grafeno unido covalentemente con glucosa oxidasa. Estas observaciones están conformadas por los parámetros de entrada en diferentes condiciones de $\text{PH} = \in \{4, 5, 6, 7\}$, Glucosa= $\{4, 8, 16, 20\}$ medida en mM, temperatura = $\in \{20, 37, 47, 57\}$ grados Celsius, concentración de Benzoquinona= $\in \{1, 0.8, 0.4, 0.2\}$. Bajo estos parámetros el biosensor devuelve como resultado la medición de la corriente faradaica I medida en microamperes mA. En la Tabla 3.1 se muestran los parámetros utilizados en el biosensor:

3.4 Métodos y técnicas utilizadas

Glucosa G (mM)	PH	Temperatura T (celcius)	p-Benzoquinona (mM)
4	4	20	1
8	5	37	0.8
12	6	47	0.4
20			

Tabla 3.1: Parámetros de las condiciones de un biosensor amperométrico.

Para adquirir los datos se utilizó un polarógrafo Radelkis OH-105. El electrodo de trabajo fué un electrodo de disco giratorio con un diámetro de 6mm, preparado de un grafito espectralmente puro con glucosa oxidasa inmovilizada sobre su superficie y se utilizó un electrodo de calomelanos saturados como el electrodo de referencia, mientras que el electrodo auxiliar era un electrodo de carbón vitrificado.

Una vez que se obtuvieron los datos, se plantearon los experimentos desde dos perspectivas:

- En el primer grupo de experimentos los datos son utilizados tal como fueron obtenidos, es decir, sin ningún pre-procesamiento adicional.
- Para el segundo grupo de experimentos se utilizó el conjunto de los datos mediante una transformación basada en un logaritmo natural lo cual permitió *simetrizar* los datos de entrada.

3.4. Métodos y técnicas utilizadas

La base de datos fué particionada de manera aleatoria para definir los conjuntos independientes de entrenamiento y prueba, por lo que se definió un 70 % de para el conjunto de entrenamiento mientras que el 30 % restante fué utilizado como el conjunto de prueba. De esta manera, el primer conjunto es utilizado para *ajustar* o entrenar el modelo analizado, a partir de esto se calcula los valores de la predicción solo con los datos de prueba sin que estos hayan sido analizados antes. De esta manera se aseguró que los modelos fueran correctamente evaluados, es decir, si sesgo alguno por parte del experimentador.

Algunos de los modelos de regresión son entrenados modificando sus respectivos parámetros de ajuste. Éstos fueron seleccionados utilizando la técnica de búsqueda de grilla la cual tiene como objetivo encontrar para cada modelo sus

parámetros óptimos correspondientes. A continuación se describe su proceso de selección:

3.4.1. Métodos de regresión

Dado un conjunto de datos de entrenamiento de la forma $\{(x_n, t_n)\}_{n=1}^N \in X \times \mathbb{R}$ donde $X = \mathbb{R}^d$ los métodos de regresión consisten en escoger un estimado específico $y(x)$ del valor de t por cada entrada x , donde el promedio de t se condiciona en x lo cual es conocido como la función de regresión.

La configuración de los diferentes métodos de regresión utilizados durante los experimentos se detallan a continuación:

3.4.2. Modelos lineales

- El parámetro λ para los modelos lineales de regresión se varió entre 0.1 y 0.9, a intervalos de 0.1. Éste tiene la función de ser un parámetro de regularización.
- GLM. La ecuación que genera es de la forma $\mu = X\beta$, donde se devuelve un vector β de estimaciones de los coeficientes para una regresión lineal generalizado de las respuestas en y sobre los predictores en X .
- PLS. Calcula una regresión de Y sobre X usando un determinado número de componentes. En este caso el número de componentes analizados fue de $\{1, 2, 3, 4\}$. El algoritmo interno de PLS estandariza el conjunto de datos a media es 0 y varianza 1.
- Ridge. El parámetro se ajustó a $\alpha = 0.0001$, λ se varió de $\{0.1 - 0.9\}$ y el parámetro de estimación se analizó con valores $k = \{1, 2, 3, 4\}$

3.4.3. Máquinas de vectores de soporte para regresión

Las SVM fueron implementadas utilizando las librerías en C++ desarrolladas por Chih-Chung Chang and Chih-Jen Lin ([Chang et al., 2011](#)) que es un software integrado para tareas de clasificación C-SVC y nu-SVC; y regresión épsilon-SVR y nu-SVR). Fueron tres tipos de kernels utilizados: Kernel lineal, Polinomial de

grado 2 y kernel de base radial. Los parámetros C que es el que determina la complejidad del modelo, el parámetro ϵ que puede afectar en el número de vectores de soportes usados para la construcción de la función de regresión fueron utilizados de la siguiente manera:

- El parámetro C fue variado logarítmicamente entre $10^{-1.5}$ y $10^{1.5}$ a intervalos de 0.5.
- El parámetro ϵ fué variado logarítmicamente entre 10^{-3} y 10^0 a intervalos de 0.5.

Estos son valores comunes a todos los tipos de kernel. El kernel de base radial utiliza un parámetro adicional que es γ , el cual especifica el ancho del kernel. Éste fue varió logarítmicamente entre $10^{-1.5}$ y $10^{1.5}$ a intervalos de 0.5.

3.4.4. Parámetros de la red Neuronal de regresión

La arquitectura de la red consistió en una sola capa oculta con cuatro neuronas internas. Las funciones de activación sigmoide y la estrategia de aprendizaje fue el algoritmo de Levenberg-Marquart. Este tiene la ventaja de ser rápido y de eliminar el uso de la tasa de aprendizaje.

3.4.5. Parámetros de árbol de regresión CART

Una de las ventajas que presenta el algoritmo CART es su ausencia de parámetros libres (A.Chipman *et al.*, 1998). Esto permite su utilización en diversos escenarios, independientemente de la naturaleza de los datos.

3.4.6. Validación Cruzada o Cross Validation

El proceso de entrenamiento antes mencionado (sobre el 70 % de los datos), se empotró en un proceso de validación cruzada (Russel & Norving, 2004) de 30 repeticiones por 10 particiones; en otras palabras el conjunto de entrenamiento fue dividido en 10 subconjuntos y a su vez se reprodujo el proceso durante 30 repeticiones. Durante éste se ejecutaron a la misma vez los algoritmos de modelos

generalizados lineales, mínimos cuadrados parciales, regresión de RIDGE, LAS-SO, SVM en sus diferentes tipos de kernel, lineal, polinomial y de base radial, así también como la red neuronal y el árbol de regresión.

El desempeño o la exactitud de la predicción de la corriente farmacia del biosensor de cada ejecución, es decir de cada partición de los resultados, por los distintos modelos de regresión se evaluó por medio de *Normalized Root Mean Square Error* NRMSE. Y esta se refiere al RMSE dividido por el rango de valores observados de una variable que se predice:

$$NRMSE = \sqrt{\frac{\sum_{i=1}^N \left(\frac{e_i}{1_{hii}} \right)^2}{\sum_{i=1}^N (\hat{y} - \bar{y})^2}} \quad (3.1)$$

Capítulo 4

Resultados y discusión

En este capítulo se presentarán los resultados experimentales que fueron logrados mediante la aplicación de los métodos y algoritmos descritos en el capítulo anterior. Los algoritmos fueron analizados en dos modalidades experimentales: por una parte los datos fueron empleados tal cual sin aplicar ningún pre-procesamiento, mientras que por otro lado se aplicó un log natural a los datos como pre-procesamiento.

4.1. Análisis estadístico básico

Después de preprocesar los datos, las principales medidas estadísticas de cada variable se muestran en la tabla 4.1

Glucosa	Benzoquinona	T	PH	Target
Min. :-1.412	Min. :-1.2629	Min. :-1.4797	Min. :-1.3395	Min. : 0.2848
1st Qu.: -0.706	1st Qu.: -0.7893	1st Qu.: -0.5481	1st Qu.: -0.6698	1st Qu.: 1.9575
Median : 0.000	Median : 0.0000	Median : 0.1279	Median : 0.0000	Median : 4.2178
Mean : 0.000	Mean : 0.0000	Mean : 0.0000	Mean : 0.0000	Mean : 12.3994
3rd Qu.: 0.706	3rd Qu.: 0.7893	3rd Qu.: 0.6759	3rd Qu.: 0.6698	3rd Qu.: 14.9122
Max. : 1.412	Max. : 1.2629	Max. : 1.2240	Max. : 1.3395	Max. : 75.5506

Tabla 4.1: Tabla estadística descriptiva después del pre-procesamiento de los datos.

De aquí se puede observar que las variables predictoras son perfectamente simétricas –i.e. media y mediana son iguales, con excepción de la T , cuya distribución esta sesgada hacia la izquierda (sesgo negativo). Un cómputo rápido de la

correlación producto-momento de pearson indica que la variable *PH* ofrece una relación un tanto lineal con respecto a la variable de salida ¹.

Primera variable	Segunda variable	Correlación
PH	Target	-0.332
Glucosa	Target	0.267
T	Target	0.167
Benzoquinona	Target	0.063

Tabla 4.2: Variables de correlaciones de Pearson.

Con el propósito de refinar un poco este análisis, se computó también el coeficiente de correlación Spearman ρ el cual es una medida de asociación entre dos variables aleatorias continuas. Este coeficiente sirve para detectar no linealidades cuadráticas generales pero es eficiente para detectar posibles valores atípicos y no linealidades monótonas – Ver Tabla 4.3.

Primera variable	Segunda variable	Correlación
PH	Target	-0.569
Glucosa	Target	0.405
T	Target	0.195
Benzoquinona	Target	0.082

Tabla 4.3: Variables de correlaciones de Spearman.

De este análisis se puede notar una relación de naturaleza no lineal de las variables de entrada con la salida del biosensor, donde la variable predictora mas importante es *PH*, seguido por la *Glucosa* y la *Temperatura* mientras que la *Benzoquinona* es la variable predictora menos importante.

4.1.1. Desempeño de los algoritmos de regresión

En las tablas 4.4 y 4.5 se muestra el desempeño que tuvieron cada uno de los modelos de regresión en los dos escenarios o variantes, los datos sin el operador LOG y con el operador LOG –ver Capítulo 3:

¹Aunque las verdaderas correlaciones no son iguales a 0 (PH-Target: $t = -6.2685$, $df = 318$, $p\text{-value} = 1.187e-09$, Glucosa-Target: $t = 4.9337$, $df = 318$, $p\text{-value} = 1.303e-06$, T-Target: $t = 3.0135$, $df = 318$, $p\text{-value} = 0.00279$).

4.1 Análisis estadístico básico

Método de regresión	NRMSE-no-log	R
GLM	0.50	0.50865
PLS	0.50	0.50885
LASSO	0.52	0.50987
RIDGE	0.55	0.50859
SVM Lineal	1.44	0.51917
SVM polinomial	0.88	0.59903
SVM RBF	0.03	0.99997
CART	0.04	0.98994
NN	0.11	0.98448

Tabla 4.4: NRMSE y coeficiente de regresión de los modelos matemáticos sin el LOG.

Método de regresión	NRMSE-log	R
GLM	0.26	0.73256
PLS	0.26	0.7626
LASSO	0.33	0.76809
RIDGE	0.29	0.73.25
SVM Lineal	0.28	0.7179
SVM polinomial	0.18	0.82089
SVM RBF	0.01	0.99997
CART	0.03	0.9944
NN	0.05	0.98 362

Tabla 4.5: NRMSE y coeficiente de regresión de los modelos matemáticos después de LOG.

Se puede observar claramente que los mejores valores de NRMSE son obtenidos con el procesamiento LOG de los datos, y son tres los modelos que sobresalen al respecto, la SVM RBF, en árbol CART y la red neuronal NN –ver Figura 4.1. Estos tres alcanzan coeficientes de regresión cercanos a uno. Por otra parte, los modelos que peor se desempeñan son los modelos GLM, PLS, LASSO, RIDGE y la SVM lineal –ver Anexo A. Esto viene un tanto a reforzar los resultados del análisis de correlación de Spearman, es decir la naturaleza no lineal de los datos.

En la figura 4.2 se muestra un gráfico de la respuesta del biosensor vs. la predicción del modelo en particular. Se puede apreciar como los algoritmos ganadores *siguen* de una manera bastante razonable el valor real de la corriente faradaica. Sin embargo, valores altos de la señal presenta dificultades para ser modelada.

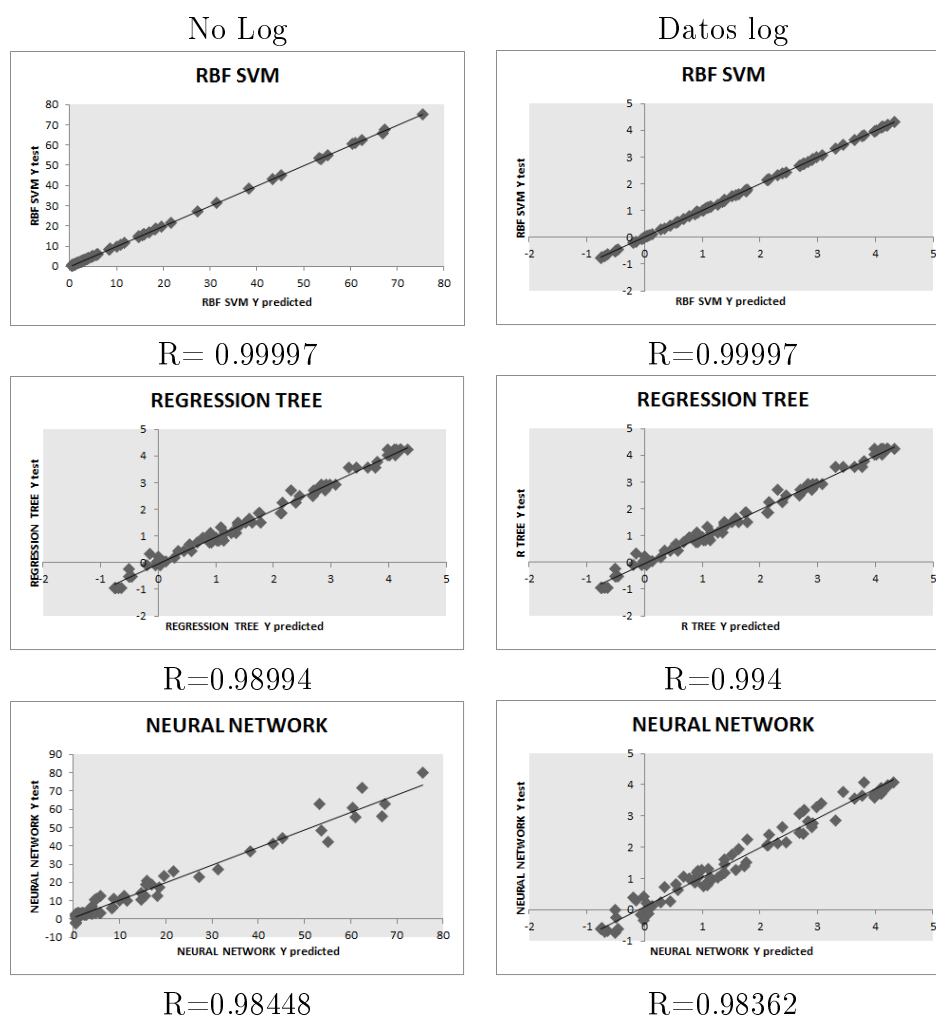


Figura 4.1: Visualización gráfica de los coeficientes de correlación R antes y después del LOG de los mejores modelos.

Con el objetivo de visualizar la superficie de la respuesta del biosensor en función de sus entradas, se generaron gráficas con los mejores modelos. La superficie generada por la SVM RBF es indicada en las figuras 4.3 y 4.4. Las superficies de los algoritmos CART y la red neuronal son plasmadas en el Apéndice B. Este arreglo responde a que se busca resaltar en detalle que las SVM RBF es el algoritmo que mejor desempeño ha mostrado.

Estas figuras fueron ensambladas primero, fijando la Benzoquinona a 0.2 y variando la glucosa. Los ejes X-Y son la Temperatura y el PH; y posteriormente

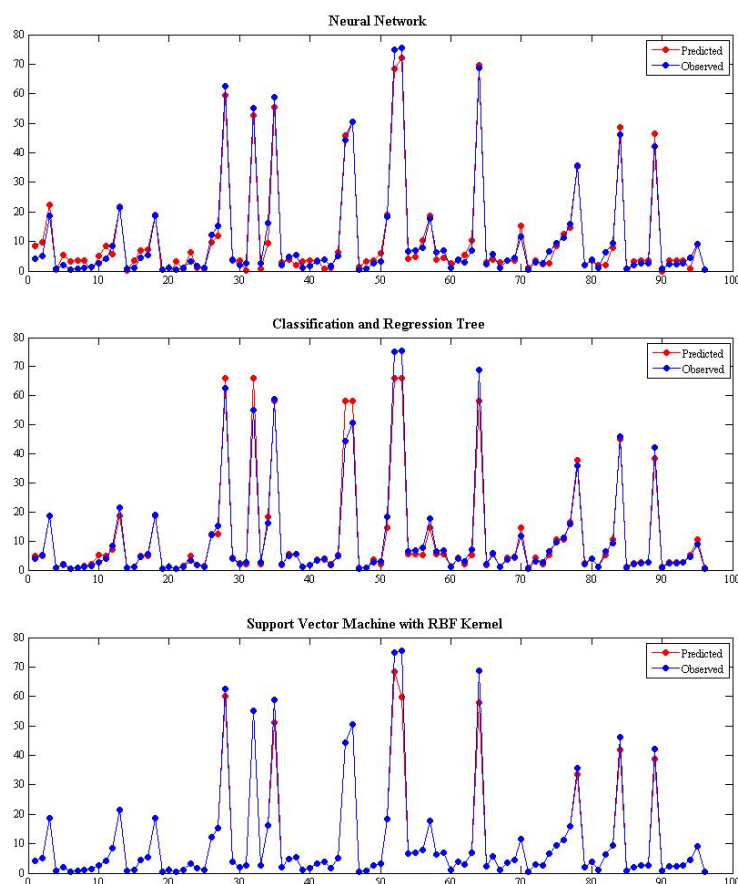


Figura 4.2: Valores observados y valores predichos renderizados por los modelos de NN, CART y SVM RBF en los datos log.

fijando la Glucosa a 4 con diversos valores de Benzoquinona.

En las siguientes secciones se detallarán los modelos matemáticos que mejor describen la respuesta de biosensor en función de sus entradas.

4.1.2. Modelo de la red neuronal de una capa con 4 neuronas

El modelo de Red Neuronal es una estructura básica y sencilla de aproximación, la cual es una ecuación lineal. Dado un conjunto de variables tales: $X_1 = \text{glucosa}$, $X_2 = \text{Benzoquina}$, $X_3 = \text{Temperatura}$ y $X_4 = \text{PH}$ que se relaciona con sus pesos mediante una capa oculta la cual consta de 4 neuronas

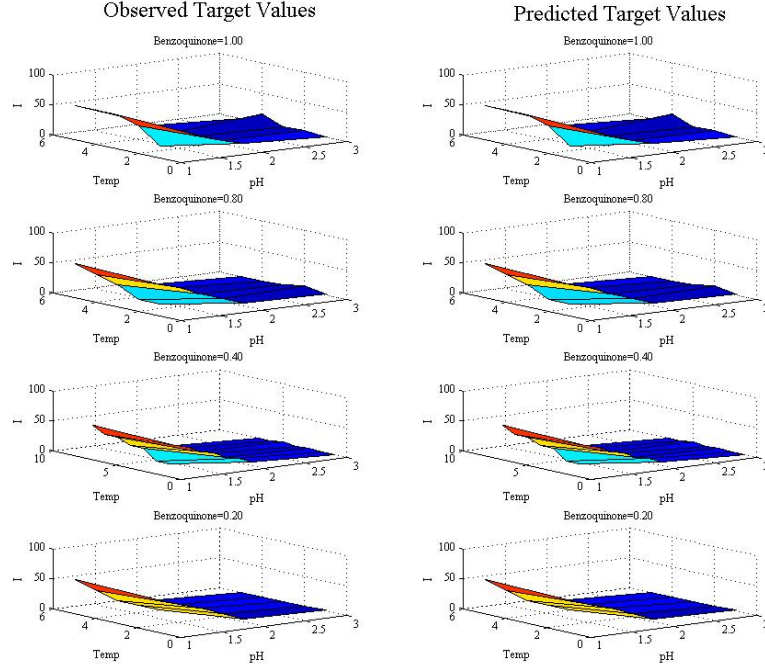


Figura 4.3: Valores reales y valores de predicción por medio de los modelos de SVMR en los datos log.

internas para producir la salida (Intesidad de la corriente faradaica) –ver Figura 4.5.

El modelo matemático final se muestra en la ecuación 4.1:

$$Y = b^* + \sum_{s=1}^4 w_s * n_s \quad (4.1)$$

donde el subíndice s indica la pertenencia a la capa de salida. Por otra parte cada n_s esta definida de la siguiente forma:

$$n_s = \text{tansig} \left[\sum_{i=1}^4 (b_i + w_{is} * X_i) \right]$$

Los pesos w_s y el bias b^* tienes los siguientes valores:

4.1 Análisis estadístico básico

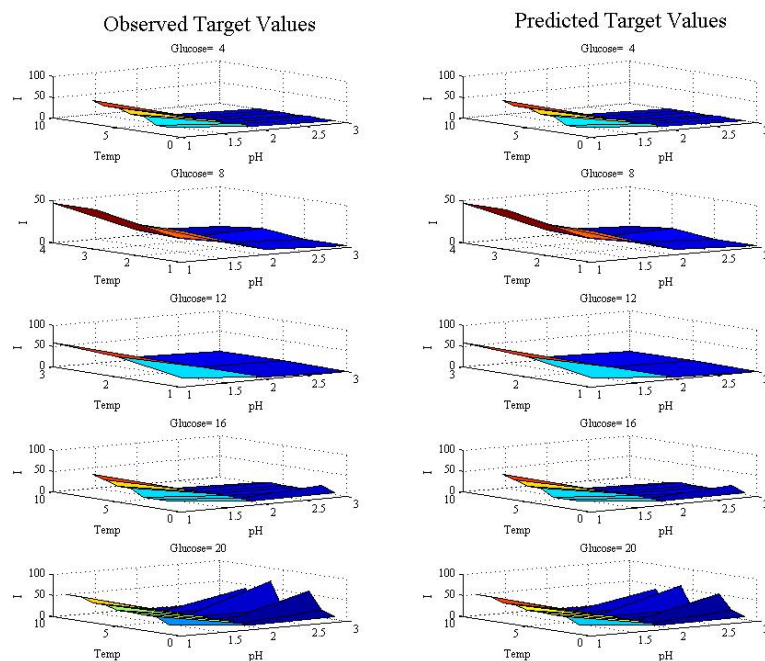


Figura 4.4: Valores reales y valores de predicción por medio de los modelos de SVMR en los datos log.

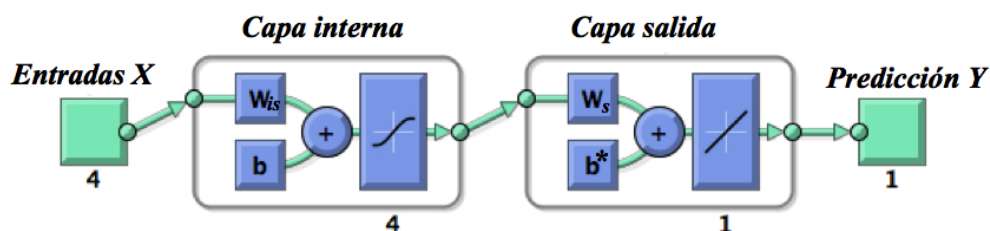


Figura 4.5: Modelo de Red Neuronal final

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ b^* \end{pmatrix} = \begin{pmatrix} -0.345672 \\ -0.628801 \\ 0.362671 \\ -4.781217 \\ -5.092845 \end{pmatrix}$$

Los pesos w_{is} y los bias b para cada nodo de la capa interna son:

$$w(i, s) = \begin{pmatrix} -13.292032 & -0.733296 & 0.643120 & -0.039301 \\ -0.087084 & -0.215409 & 0.043895 & -0.007911 \\ 133.065337 & -0.576762 & 4.903561 & 0.125645 \\ 139.782513 & 5.381696 & 0.824023 & -4.086904 \end{pmatrix}$$

$$b(s) = \begin{pmatrix} -21.431381 \\ 1.841679 \\ 1.449347 \\ -4.925324 \end{pmatrix}$$

4.1.3. Modelo de CART

El modelo gráfico de CART se representa en un árbol binario de decisión que se le conoce como el árbol máximo en el cual se muestran las variables en forma jerárquica a partir del nodo raíz.

Los otros parámetros utilizados fueron: control del tamaño del árbol, control del podado por reducción de error, selección del criterio de partición inicial y asignación de probabilidad de las categorías según la original de la base de datos. Los resultados se muestran en las figuras 4.6, 4.7 y 4.8:

El árbol utilizó las cuatro variables que son: Glucosa, p-benzoquinona, temperatura y PH. El modelo se resume en 97 reglas de decisión para el árbol máximo –ver Figura 4.6, iniciando a partir de la variable más importante x_4 que es PH.

Puede ser necesario reducir el tamaño del árbol, por lo que un proceso de podado fue aplicado. Así se obtuvo un árbol de decisión el cual es el árbol completo sin podar pero con la información óptima del árbol podado, es decir, se selecciona la sección mas importante del CART, tal como se muestra en la figura 4.7:

De acuerdo con CART la variable mas importante se le atribuye a la variable X_4 que en los biosensores representa la variable del PH.

4.1.4. Modelo de la SVM RBF

Los parámetros para generar el modelo de la SVM con kernel RBF, que ofrece el mejor desempeño, son los siguientes:

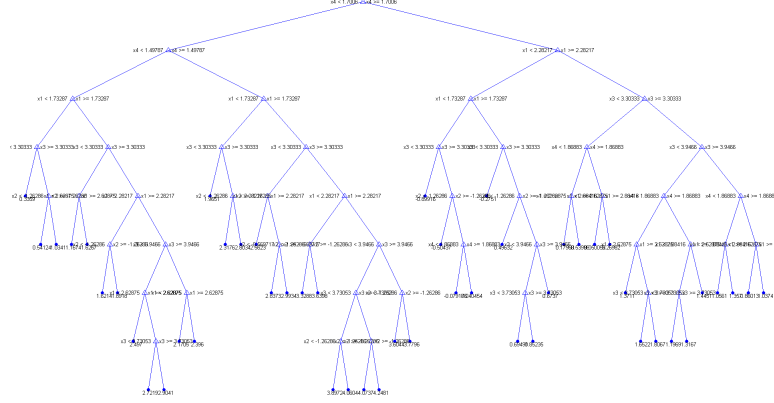


Figura 4.6: Modelo gráfico de un árbol de regresión CART

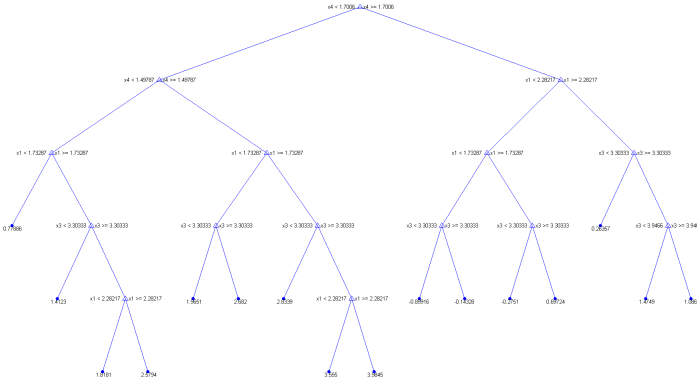


Figura 4.7: Modelo gráfico de un árbol de regresión podado CART

$$\begin{pmatrix} C \\ \gamma \\ \rho \end{pmatrix} = \begin{pmatrix} 16 \\ 5.6569 \\ 0.95 \end{pmatrix}$$

Estos parámetros deberán ser usados para generar un modelo matemático de la siguiente forma:

$$Y = b + \sum_{i=1}^L w_i \times k(x_t, x_i)$$

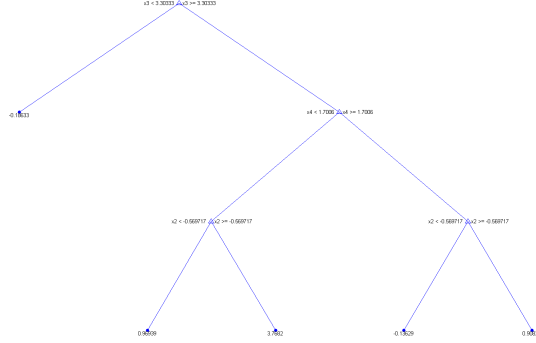


Figura 4.8: Modelo gráfico de un árbol de regresión óptimo CART

donde b es una constante y w_i son los factores de ponderación (pesos) de la función kernel $k(\cdot, \cdot)$. Para el caso de este trabajo, el kernel RBF se define como:

$$K(X_i, X_j) = \exp\left(-\frac{1}{\gamma^2} \|x_i - x_j\|^2\right)$$

Las constantes b y w están dadas por:

$$w = \sum_{i=1}^L \alpha_i y_i x_i$$

y:

$$b = \frac{1}{N_s} \sum_{s \in S} \left(y_s - \sum_{m \in S} \alpha_m y_m x_m \cdot x_s \right)$$

Ambas deberán ser encontradas resolviendo el problema de optimización cuadrática descrito en la sección 2.6.

Capítulo 5

Conclusiones

5.1. Conclusiones

A la luz de los resultados obtenidos durante el desarrollo de esta investigación, podemos concluir lo siguiente:

1. De acuerdo con los resultados obtenidos en las tablas de error NMRSE y el coeficiente de correlación –Ver tablas 4.4 y 4.5, respectivamente, se puede concluir que la relación entre las variables predictivas y la salida del biosensor es de naturaleza no lineal. Por lo tanto, es recomendable en biosensores con las características particulares estudiadas a lo largo de esta investigación, que las técnicas de regresión no lineales para su moderación y análisis.
2. Los modelos de regresión que mejor se desempeñan en la predicción de la corriente faradaica son las SVM con kernel de base radial, los árboles de regresión CART y la red neuronal. Coeficientes de regresión cercanos a uno, dan cuenta del potencial de estos algoritmos de reconocimiento estadístico en etapas de diseño y desarrollo de biosensores. Estos algoritmos de regresión representan un potencial elemento en la construcción de modelos de complejidad superior, donde se necesite la integración de respuestas de redes de biosensores. Su independencia de la naturaleza de los datos, los hace elementos muy atractivos como modelos de caja negra.

3. Del modelo CART –ver figura 4.7, se puede observar en que las variables más relevantes son PH, seguido de la glucosa, temperatura y la p-benzoquinona. Es importante resaltar que la importancia de éstas no depende en sí del método de aprendizaje estadístico, si no de la relación y el comportamiento enzimático del biosensor.
4. Se puede apreciar en la figura ?? que la mayor parte de la señal se predijo satisfactoriamente, sin embargo se pueden notar algunas divergencias entre los puntos, es decir, los modelos como SVM RBF y NN tienen dificultad para predecir las salidas de alto valor.

El monitoreo de glucosa por medio de biosensores de glucosa oxidada puede constituir un importante aliado para pacientes diabéticos. Como una medida para evitar complicaciones crónicas colaterales, un monitoreo rápido, confiable y barato sería deseable para ellos, donde los biosensores son un sólido candidato para tal tarea. Si embargo, su entendimiento aun esta en desarrollo por los especialistas del campo, con mirar a incrementar considerablemente su exactitud y estabilidad. En la tecnología de biosensores, la moderación matemática es una herramienta frecuentemente solicitada, dado que facilita la simulación computacional, evitando pruebas destructivas, además de tiempo y problemas de recursos.

Finalmente, se puede afirmar que la metodología plasmada a lo largo de este trabajo puede ser aplicada en diversos escenarios del desarrollo de tecnología diversa, donde se requiera simular la respuesta física de cualquier fenómeno, en función de una serie de entradas.

5.2. Trabajo Futuro

Como trabajo futuro se plantean las siguientes líneas de investigación:

1. Una posible dirección puede ser el modelar la corriente faradaica de salida en función de su forma. En este sentido es natural pensar en los modelos *Splines* para representarla. Es un procedimiento no-paramétrico de regresión que no hace suposiciones *a priori* acerca de la relación funcional subyacente

entre las variables dependiente e independientes. Dicha relación es construida a partir de un conjunto de coeficientes y un conjunto de funciones base (Hastie *et al.*, 2001):

$$y = f(X) = \beta_0 + \sum_{m=1}^M \beta_m h_m(X)$$

donde la suma es sobre un total de M términos no constantes en el modelo. Durante el proceso de entrenamiento, se va incrementando gradualmente el número de funciones base intentando de maximizar el criterio de bondad de ajuste de mínimos cuadrados.

2. Otra posible línea de investigación es el uso de *Wavelets* para aproximar la salida del biosensor. Considerando un modelo de regresión no paramétrico estándar univariado:

$$y_i = g(t_i) + \sigma \epsilon_i, i = 1, \dots, n,$$

El objetivo es encontrar la función subyacente g de los datos $\mathbf{y} = (y_1, \dots, y_n)'$, sin asumir ninguna estructura en particular de esta función (Antoniadis *et al.*, 2001).

Anexos A

Análisis de regresión estadística clásica

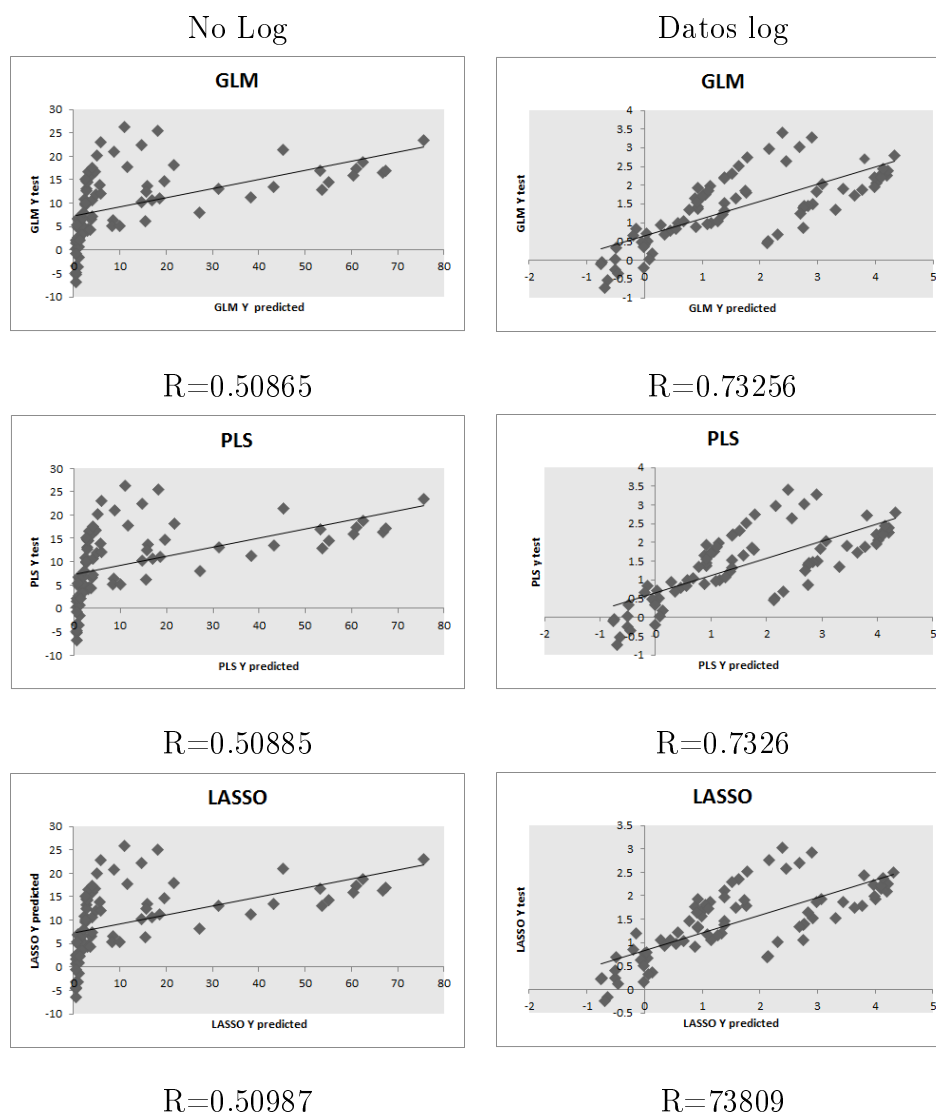


Figura A.1: Visualización gráfica de los coeficientes de correlación R antes y después del LOG en los peores modelos.

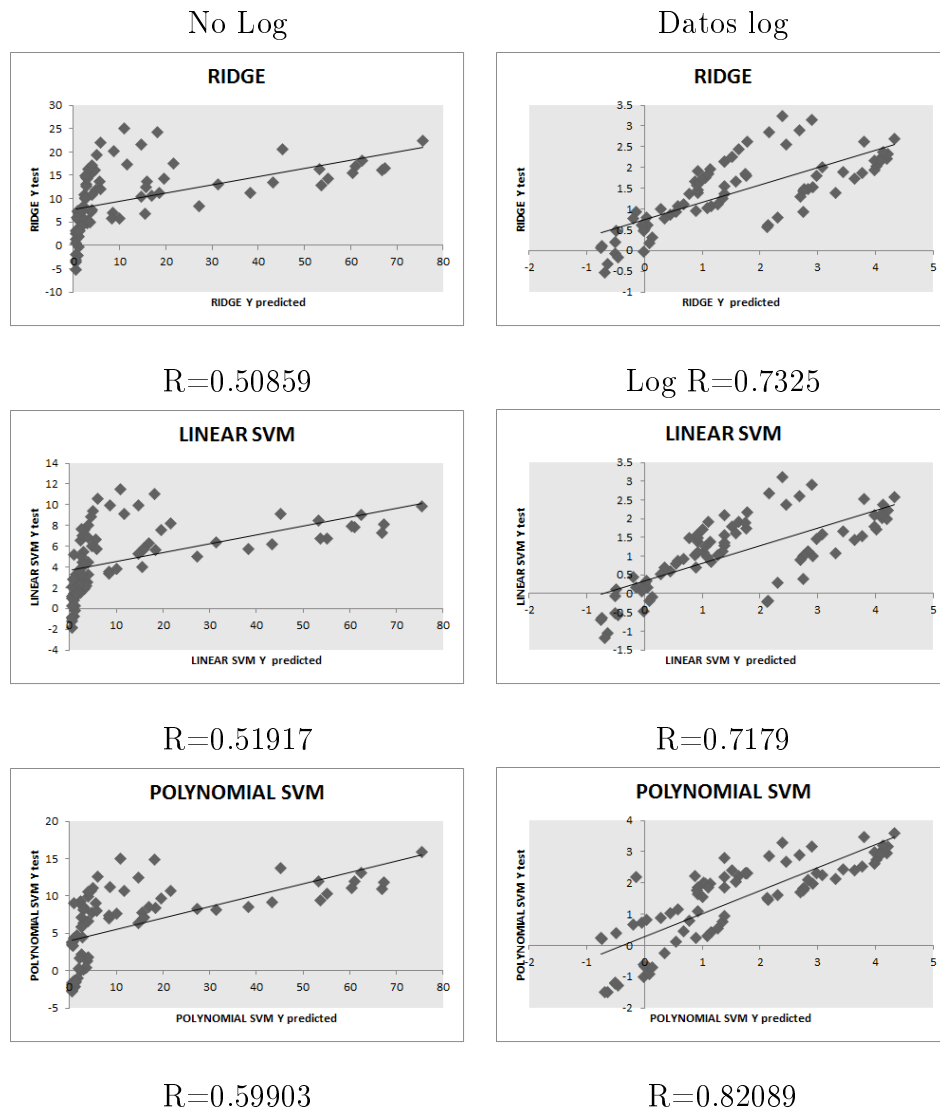


Figura A.2: Visualización gráfica de los coeficientes de correlación R antes y después del LOG en los peores modelos (continuación).

Anexos B

Superficie generada por la respuesta
del biosensor

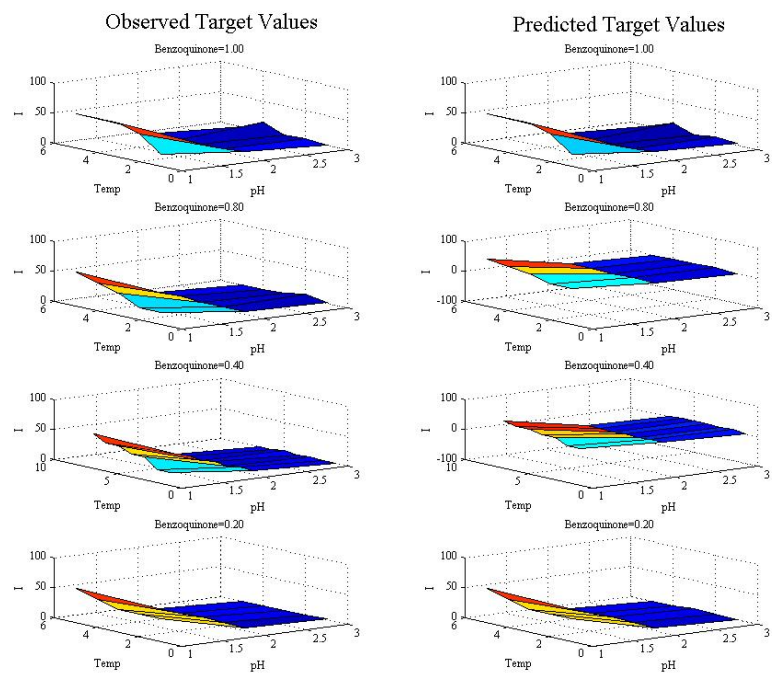


Figura B.1: Valores reales y valores de predicción por medio de los modelos de NN en los datos log.

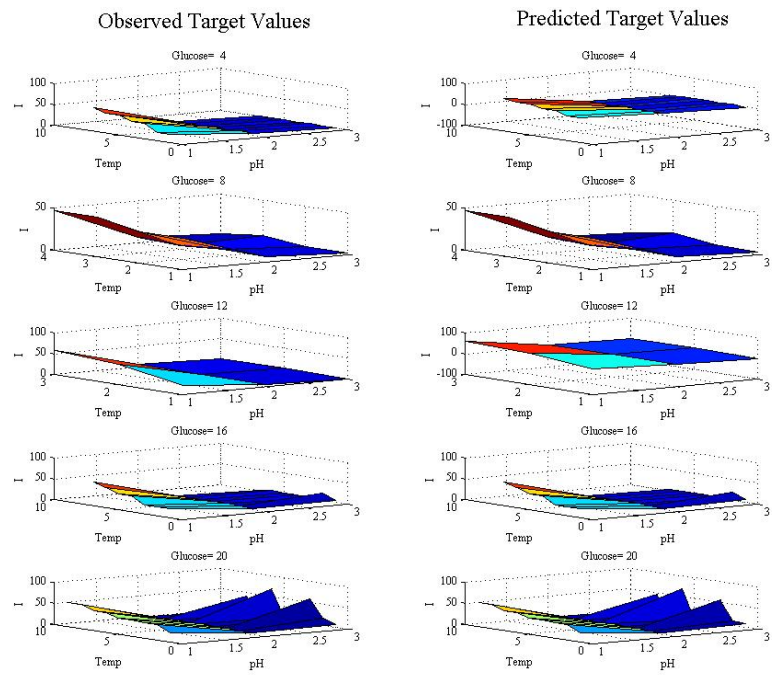


Figura B.2: Valores reales y valores de predicción por medio de los modelos de NN en los datos log.

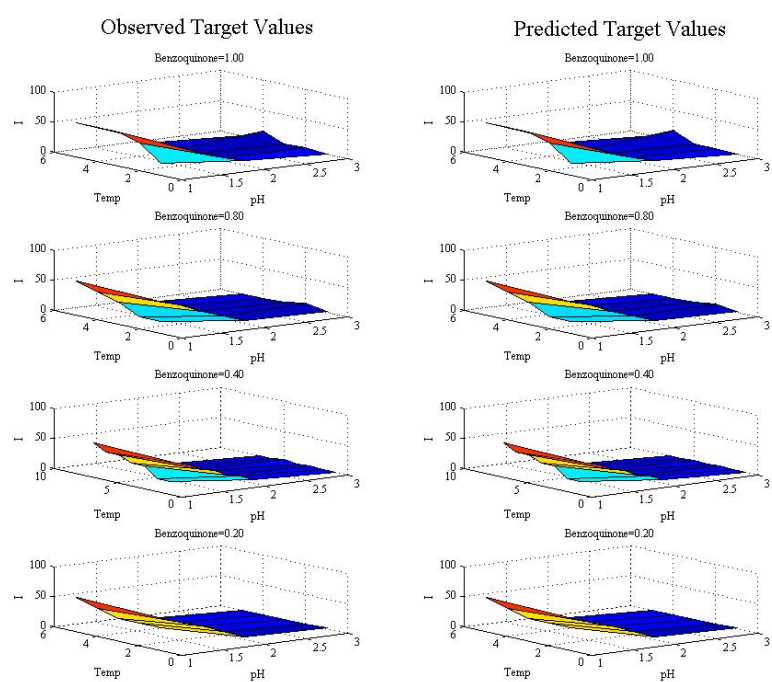


Figura B.3: Valores reales y valores de predicción por medio de los modelos de RTREE en los datos log.

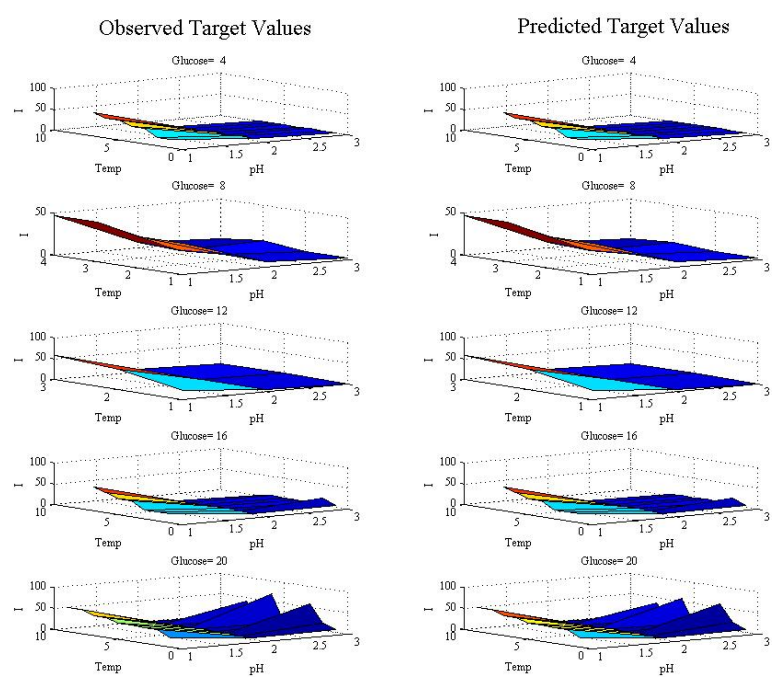


Figura B.4: Valores reales y valores de predicción por medio de los modelos de RTREE en los datos log.

Referencias

- A.CHIPMAN, H., GEORGE, E.I. & MCCULLOCH, R.E. (1998). Bayesian cart model search. In *Journal of the American Statistical Association*, 935–948, Journal of the American Statistical Association. [40](#)
- ADM, A.D.A. (2008). Diagnosis and classification of diabetes mellitus. In *Diabetes Care, VOLUME 31, Supplement 1*, 56–59, Springer-Verlag. [1](#), [5](#), [8](#)
- AEG CASS, G.F., G. DAVIS (1984). Analytical chemistry 56. In *Analytical Chemistry*, 667–670, American Chemical Society. [2](#), [15](#), [16](#)
- ALEGRET, S., DEL VALLE, M. & MERKOCI, A. (2004). Sensores electroquímicos. In *Introducción a los quimiosensores y biosensores*, 14–20, Servei de publicacions. [13](#), [15](#), [17](#)
- AMERICAN DIABETES ASSOCIATION (2013). <http://www.diabetes.org/>. [1](#), [6](#), [7](#)
- ANTONIADIS, A., BIGOT, J. & SAPATINAS, T. (2001). Wavelet estimators in nonparametric regression: A comparative simulation study. *Journal of Statistical Software*, **6**, 1–83. [54](#)
- ANTONIO, P. & PÉREZ, C. (2004- 2005). Aplicaciones de la ingeniería electrónica e informática en medicina. In *Aplicaciones de la ingeniería electrónica e informática en medicina*, 42, American Dietetic Association. [12](#)
- B. VALIÑO, A.M.S. & FERNÁNDEZ., J.G. (2003). Matemática de la producción de glucosa oxidasa. *Instituto de Minas de San Petersburgo*, **1**, 2. [10](#)

REFERENCIAS

- BANKAR, S.B., BULE, M.V., SINGHAL, R.S. & ANANTHANARAYAN, L. (1999). Glucose oxidase. *Elsevier*, **24**, 397–401. 9, 10, 11
- BARONAS, R., IVANAUSKAS, F. & SAPAGOVAS, J.K. (2006). Mathematical modeling of biosensors. In *An Introduction for Chemists and Mathematicians*, 104–230, a. 19
- BARONAS, R., IVANAUSKAS, F. & SAPAGOVAS, J.K. (2009). Mathematical modeling of biosensors based on an array of enzyme microreactors. In *J. Math. Chem Lithuania*, 849, a. 17
- BISHOP, C. (2009). Redes neuronales. In *Pattern Recognition and Machine Learning*, 225–232, Springer. 26
- BOUCHER, J.L. & O’CONNELL, B.S. (2004- 2005). Prediabetes. In *American Dietetic Association Guide to Diabetes, Medical Nutrition Therapy and Education*, 42, American Dietetic Association. 7
- BREIMAN, L. & FRIEDMAN, L. (1984). Classification and regresssion trees. In *Classification and regresssion trees*, Wandsworth, Belmont. 28
- CHANG, CHIH-CHUNG & LIN, C.J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, **2**, 27:1–27:27, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. 39
- CHATTERJEE, S. & A.HADI (2006). Regression analysis by example. In *Regression Analysis by example*, 272–282, wiley. 33
- CHIH-WEI HSU, C.C.C. & LI, C.J. (2010). A practical guide to support vector classification. *Department of Computer Science National Taiwan University, Taipei 106, Taiwan*, **1**, 1–2. 21
- D.G.BUERK (2005). Blood glucose monitoring. In *Biosensors: Theory and Applications*, 73, CRC Press. 3
- DIABETES RESEARCH INSTITUTE (2010). <http://www.diabetesresearch.org>. 7

REFERENCIAS

- DUDA, R. & E.HART, P. (1973). Pattern classification and scene analysis. In *Pattern Classification and Scene Analysis*, wiley. 36
- ENZYME NOMENCLATURE (1992). EC. 1.1.3.4. 10
- FEDERACIÓN MEXICANA DE DIABETES (2010). <http://www.fmdiabetes.org/fmd/pag/index.php>. 6
- FLETCHER, T. (2009). Support vector machines explained. *University College London UCL*, 1, 3–5. 21, 22, 23
- GATTI, P.S. (2004). Electrodo para medir glucosa. *XIII SEMINARIO DE ING. BIOMÉDICA*, 13, 1–5. 12, 20
- GUISAN, J. (2006). Immobilization of enzymes and cells. *Springer*, 22, 1–3. 13
- HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. (2001). *The Elements of Statistical Learning*. Springer Series in Statistics, Springer New York Inc., New York, NY, USA. 21, 54
- INEGI (2011). <http://www.inegi.org.mx/>. 2
- LEROITH, D., TAYLOR, S.I. & OLEFSKY, J.M. (2004). Diabetes type 1. In *Diabetes Mellitus, A fundamental and clinical text*, 591, Lippincott Williams and Wilkins. 7
- LUNA, M.C. (Noviembre 2008). Diseño y aplicación de sensores electroquímicos basados en moléculas orgánicas conductoras. In *Centro de Investigaciones en Nutrición y Salud, Instituto Nacional de Salud Pública. Cuernavaca, Morelos, México*, 19–100, publicaciones uco. 11
- MARSLAND, S. (2011). Machine learning. In *Machine Learning: An Algorithmic Perspective*, 1–16, CRC Press. 20
- NEZ, J.L.M., RAMOS, E.G., ALEGRET, S. & DELGADO, R.J. (2011). Biosensor de glucosa basado en un biocompuesto disperso de grafito-epoxi-platino-glucosa oxidasa. *Información tecnológica*, 22, 29–40. 17

REFERENCIAS

- NEZ, J.M., ARROYO, M. & BUELNA, H. (2004). Desarrollo de un biosensor amperométrico de glucosa basado en un biocomposito disperso de grafito-epoxy-platino-glucosa oxidasa. In *Biosensor amperométrico*, 1–14, Facultad de Ciencias Biológicas de Monterrey. [17](#)
- NUELA, P. & GALVÁN, I. (1995). Redes neuronales. In *Redes Neuronales artificiales, fundamentos, modelos y aplicaciones*, 9–26, RA-MA Editorial. [24](#), [25](#), [26](#)
- PALMA, O., LEVY, T.S., FRANCO, A., OLAIZ, G. & MÉNDEZ, I. (2006). Encuesta nacional de salud y nutrición. In *Centro de Investigaciones en Nutrición y Salud, Instituto Nacional de Salud Pública. Cuernavaca, Morelos, México*, 2, a. [6](#)
- REVIEJO, A.J. & PINGARRÓN, J. (2000). Biosensores electroquímicos una herramienta útil para el análisis medio ambiental y clínico. *Departamento de química analítica Universidad computense de madrid*, **2**, 5–12. [13](#), [16](#), [18](#)
- RONALD, C., W., G.C., KING, G.L., JACOBON, A., MOSES, A. & SMITH, R.J. (2005). Diabetes mellitus. In *Josli's Diabetes Mellitus*, 41–48, Josli's Diabetes Center. [3](#), [6](#), [8](#)
- RUSSEL, S. & NORVING, P. (2004). Crossvalidation, validación cruzada. In *Inteligencia artificial un enfoque moderno*, 755, Pearson Educación. [40](#)
- STOYTCHIEVA, M., NANKOV, N. & SHARKOVA, V. (2004). Amperometric graphite sensor. In *Analytical characterisation and application of a p-benzoquinone mediated amperometric graphite sensor with covalently linked Glucoseoxidase.*, 100 –102, Elsevier. [14](#)
- VAPNIK, S. & SMOLA, A. (1997). Support vector method for function approximation, regression estimation, and signal processing. In *Advances in neural information processing system*, 281–287, petshe. [20](#)
- WANG, Q. (2011). Mathematical methods for biosensor models. *Dublin Institute of Technology*, **11**, 15–20. [20](#)

REFERENCIAS

WORLD HEALTH ORGANIZATION (2013). <http://www.who.int/en/>. 9

ZHENG, Z. & WANG, J. (2006). Advances in neural network research and applications. In *Advances in Neural Network Research and applications*, 849, Department of control Science and Engineering Huazhong University of Science and Technology. 20