

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

Facultad de Ciencias Químicas e Ingeniería

Maestría y Doctorado en Ciencias e Ingeniería



“ANÁLISIS DE PATRONES DE CONDUCTA EJERCIDOS DENTRO DE UN ENTORNO PERSONAL HABITACIONAL, APLICANDO IOT EN EL SENSADO DEL ENTORNO PARA LA DETECCIÓN DE CONDUCTAS ANÓMALAS EN EL INDIVIDUO APLICANDO EDGE COMPUTING”

TESIS PARA OBTENER EL GRADO DE MAESTRO EN CIENCIAS

*Presenta:*

SÁNCHEZ GARCÍA MARÍA DEL ROSARIO

*Director :*

DR. SÁNCHEZ HERRERA MAURICIO A.

*Co-Director*

DR. OCEGUEDA MIRAMONTES THELMA V.

TIJUANA, BAJA CALIFORNIA, MÉXICO

AGOSTO 2021

## 0.1. Agradecimientos

A mi madre que me enseñó a trabajar con calidad en donde me presentara, por su apoyo, atención y por permitirme fallar más no rendirme en lo que fuera que estuviera en mis metas lograr.

A mi novio Nieblas :\* , por su apoyo incondicional en las diferentes etapas y procesos que pase a lo largo de estos años, por brindarme soporte cuando diversas situaciones me hicieron tambalear y por hacer un excelente papel de patito de goma para las decenas de borradores de tesis realizados.

Al individuo de estudio que accedió en plena pandemia mundial a apoyarme para la realización de este trabajo permitiéndome entrar a su vivienda compartiendo su tiempo, vida y espacio.

A los múltiples investigadores alrededor del mundo que hicieron posible el acceso a la información y conocimiento empleado en gran parte del desarrollo de este documento.

A los docentes que apoyaron mi formación a lo largo de la carrera y posgrado, entre los cuales se encuentra mi codirectora la Dra. Violeta Ocegueda, el Dr. Leocundo Aguilar y el Dr. Alejandro Domínguez, a este último le debo un agradecimiento por apoyarme a darme cuenta del potencial que puedo desarrollar y animarme a explotarlo.

Y por último, pero no menos importante a mi director el Dr. Mauricio Sánchez el cual me permitió emprender este camino al depositar su confianza en mí aun sin conocerme, brindándome su constante apoyo a lo largo de esta travesía ya sea a través de materiales, conocimientos o incluso una que otra palabra de aliento cuando hizo falta, gracias, es un excelente docente y persona.

## 0.2. Resumen

La implementación de un entorno IoT permite la obtención de datos del entorno e incluso es posible extender su alcance a registrar la interacción de los individuos con este, brindando información que posteriormente puede ser empleada para fines concretos. Esta tesis propuso hacer uso de los datos que puede generar un entorno IoT para la detección de conductas anómalas mediante el desarrollo de un modelo de aprendizaje de Machine Learning basado en la información obtenida a partir de la aplicación del IoT y Edge Computing en el entorno habitacional de un individuo, donde los datos generados representaban la interacción diaria entre individuo y entorno. Con el fin de realizar un análisis que contemplara las diferentes variables que influían en la vida diaria del sujeto, se propusieron dos tecnologías móviles adicionales (smartwatch Garmin “vivosmart” y una aplicación móvil “Daylio”) que permitieron el seguimiento del individuo aun fuera de su vivienda logrando la obtención de un conjunto de datos significativos para la formación de los patrones de conducta y la detección de anomalías. Los resultados generados por el modelo de final de Machine Learning fueron presentados a expertos en el área de la salud mental con el fin de potenciar su validez y con ello mejorar su calidad para posteriores análisis.

**Universidad Autónoma de Baja California**  
**FACULTAD DE CIENCIAS QUÍMICAS E INGENIERÍA**  
**COORDINACIÓN DE POSGRADO E INVESTIGACIÓN**

Folio No. 312

Tijuana, B.C., a 14 de septiembre del 2021

C. María del Rosario Sánchez García

Pasante de: Maestría en Ingeniería

Presente.-

El tema de trabajo y/o tesis para su examen profesional, en la  
Opción: TESIS

Es propuesto, por los C. Dr. Mauricio Alonso Sánchez Herrera y

Dra. Thelma Violeta Ocegueda Miramontes

Quienes serán los responsables de la calidad del trabajo que usted presente,  
referido al tema: “Análisis de patrones de conducta ejercidos dentro de un  
entorno personal habitacional, aplicando IOT en el sensado del entorno para la  
detección de conductas anómalas en el individuo aplicando Edge Computing”,  
el cual deberá usted desarrollar, de acuerdo con el siguiente orden:

- I. INTRODUCCIÓN
- II. FUNDAMENTOS TEÓRICOS
- III. DESARROLLO EXPERIMENTAL
- IV. CONCLUSIONES Y TRABAJO FUTURO
- V. BIBLIOGRAFÍA



FACULTAD DE CIENCIAS  
QUÍMICA E INGENIERÍA  
CAMPUS TIJUANA

Dr. Mauricio Alonso Sánchez Herrera  
Director de Tesis

Dra. Ana Alejandra Ramírez Rodríguez  
Sub-Directora

Dra. Thelma Violeta Ocegueda Miramontes  
Co-Directora de Tesis

M.C. Roberto Alejandro Reyes Martínez  
Director Provisional

# Índice general

0.1. Agradecimientos . . . . .	2
0.2. Resumen . . . . .	3
<b>1. Introducción</b>	<b>15</b>
1.1. Planteamiento del problema . . . . .	17
1.2. Hipótesis . . . . .	19
1.3. Objetivos . . . . .	20
1.4. Metas . . . . .	21
<b>2. Fundamentos teóricos</b>	<b>22</b>
2.1. Estímulo . . . . .	22
2.2. Hábitos . . . . .	23
2.3. Conducta . . . . .	23
2.3.1. Patrón de conducta . . . . .	23
2.3.2. Observacion y deteccion de patrones de conducta . . . . .	24
2.4. La cuarta revolución industrial . . . . .	25
2.5. El internet de las cosas . . . . .	26
2.5.1. Crecimiento del Internet de las cosas en el 2021 . . . . .	28
2.5.2. Retos del internet de las cosas . . . . .	29
2.5.3. Tecnología en un entorno IoT . . . . .	30
2.6. Edge Computing (EC) . . . . .	38
2.7. El papel de IA en los entornos inteligentes . . . . .	39

2.7.1.	Inteligencia artificial . . . . .	40
2.8.	Inteligencia artificial en la detección de patrones . . . . .	41
2.8.1.	Tipos de aprendizaje . . . . .	42
2.8.2.	Clustering . . . . .	44
2.8.3.	Anomalías . . . . .	51
2.8.4.	Detección de puntos anómalos . . . . .	52
2.9.	Trabajos relacionados . . . . .	55
<b>3.</b>	<b>Desarrollo Experimental</b>	<b>57</b>
3.1.	Análisis general del sujeto de estudio . . . . .	57
3.1.1.	Requisitos para la selección del individuo . . . . .	57
3.1.2.	Entrevista al candidato . . . . .	59
3.1.3.	Evaluación general sobre el estado mental del candidato a estudio . . . . .	59
3.2.	Análisis de conductas relevantes a detectar . . . . .	61
3.2.1.	Espacios estratégicos para la detección de conductas seleccionadas . . . . .	63
3.3.	Sensores a emplear en la vivienda . . . . .	67
3.3.1.	Características por contemplar en sensores . . . . .	68
3.3.2.	Selección de dispositivos en base a la actividad a monitorear . . . . .	68
3.4.	Contexto para emplear cada modo de operación del $\mu C$ ESP8266 dentro del entorno IoT . . . . .	69
3.5.	Funcionamiento lógico general de los dispositivos empleados en la red IoT . . . . .	72
3.5.1.	Contemplar las limitaciones en cada modo operacional implementado . . . . .	72
3.5.2.	Análisis de la relación sensores-modos de operación . . . . .	73
3.6.	Propuesta experimental . . . . .	74
3.6.1.	Arquitectura de red IoT . . . . .	75
3.6.2.	Programación de dispositivos para un bajo consumo . . . . .	77
3.6.3.	Desarrollo físico de dispositivos para un funcionamiento de bajo consumo . . . . .	80

3.6.4. Resultados Finales de la propuesta experimental para el desarrollo de la red IoT . . . . .	81
3.7. Propuesta final para el funcionamiento de la red IoT . . . . .	83
3.7.1. Desarrollo de arquitectura final para la comunicación y almacenamiento de la IoT . . . . .	83
3.7.2. Selección del modo operacional a trabajar en el desarrollo de los dispositivos . . . . .	87
3.7.3. Desarrollo lógico final de los dispositivos empleados en la red IoT . . . . .	88
3.7.4. Desarrollo Físico de los dispositivos . . . . .	95
3.7.5. Instalación final de red con el uso de servidores, AP, tecnologías de comunicación y dispositivos desarrollados . . . . .	97
3.8. Resultado finales de la implementación en red aplicando EC . . . . .	99
3.9. Monitoreo del individuo a través del uso de tecnologías externas al entorno IoT . . . . .	103
3.9.1. Tecnología “Daylio” . . . . .	103
3.9.2. Tecnología Garmin . . . . .	105
3.10. Imprevistos generales de las tecnologías empleadas (Entorno IoT, App “Daylio” y smartwach Garmin) . . . . .	109
3.11. Introducción al análisis de datos . . . . .	111
3.12. Recolección los conjuntos de datos generados por la red IoT . . . . .	112
3.13. Análisis para la construcción del modelo de preprocesamiento que transformara datos en patrones de conducta . . . . .	113
3.14. Diseño experimental . . . . .	114
3.14.1. Detección y estructuración de los patrones de conducta ejercidos por el individuo . . . . .	114
3.14.2. Unión de los conjunto de datos obtenidos de las tecnologías empleadas	124
3.14.3. Reducción de atributos . . . . .	127

3.14.4. Desarrollo del modelo de ML encargado de la detección de anomalías en los patrones de conducta del individuo . . . . .	130
3.15. Diseño Final . . . . .	137
3.15.1. Establecer patrones . . . . .	137
3.15.2. Unión de los conjunto de datos obtenidos de las tecnologías empleadas	145
3.15.3. Reducción de atributos . . . . .	146
3.15.4. Desarrollo Final del modelo de ML encargado del estudio y análisis de los patrones de conducta del individuo . . . . .	147
3.15.5. Desarrollo Final del modelo de ML encargado de la detección de ano- malías en los patrones de conducta del individuo. . . . .	159
3.16. Resultados . . . . .	165
<b>4. Conclusiones y trabajo futuro</b>	<b>167</b>
4.1. Conclusiones . . . . .	167
4.2. Trabajo futuro . . . . .	169

# Índice de figuras

2.1. México: dispositivos conectados a internet por hogar . . . . .	27
3.1. Esquema de la vivienda de estudio . . . . .	64
3.2. Instalación de la paquetería MySQL . . . . .	75
3.3. Información relacionada al servidor conectado a la “red 1” . . . . .	76
3.4. Base de datos general con las tablas . . . . .	77
3.5. Diagrama del funcionamiento del dispositivo para la detección de movimiento en modo de ahorro Deep Sleep . . . . .	79
3.6. Configuración física del dispositivo para el funcionamiento en modo deep sleep	80
3.7. Pruebas de funcionamiento del dispositivo en modo deep sleep . . . . .	81
3.8. Instalación de AP . . . . .	84
3.9. Acceso al servidor 2 conectado a la red 2 . . . . .	84
3.10. Herramienta phpMyAdmin instalada en el servidor 1 para la administración de MySQL . . . . .	85
3.11. Base de datos general desarrollada en el servidor 1 . . . . .	85
3.12. Base de datos general desarrollada en el servidor 2 . . . . .	86
3.13. Monitoreo a distancia de la BD a través de TeamViewer . . . . .	86
3.14. Diagrama del funcionamiento general de los dispositivos bajo el modo opera- cional “Modem sleep” . . . . .	90
3.15. Corrida sin percances del programa diseñado para los dispositivos . . . . .	91
3.16. Corrida del programa diseñado para los dispositivos con servidor desconectado	92

3.17. Alerta enviada a correo destinado con el reporte del problema . . . . .	92
3.18. Diagrama para el funcionamiento del Medidor AC . . . . .	95
3.19. Diagrama para el funcionamiento de PIR–LDR . . . . .	96
3.20. Diagrama para el funcionamiento de Ultrasonico–Interruptor (Switch) . . . . .	96
3.21. Diagrama para el funcionamiento de sensor flexible (Flex sensor) . . . . .	96
3.22. Prueba de dispositivo para el monitoreo de movimiento . . . . .	97
3.23. Diseño y construcción de dispositivos . . . . .	98
3.24. Dispositivo colocado en entrada . . . . .	100
3.25. Dispositivo colocado en refrigerador . . . . .	100
3.26. Dispositivo colocado entre Horno / Gabetta . . . . .	101
3.27. Dispositivo colocado sobre la mesa . . . . .	101
3.28. Dispositivo colocado en baño . . . . .	102
3.29. Servidor ubicado en área de Cocina–Sala . . . . .	102
3.30. Diseño parcial de la app móvil “Daylio” . . . . .	104
3.31. Conjunto final de datos obtenido de la plataforma “Daylio” . . . . .	105
3.32. Análisis de la FC y niveles de Estrés . . . . .	106
3.33. Análisis de sueño . . . . .	106
3.34. Panel de control de la plataforma “Garmin connect” . . . . .	107
3.35. Reporte gráfico del análisis de los niveles de estrés . . . . .	108
3.36. Reporte CSV del análisis de los niveles de estrés . . . . .	108
3.37. Estructura del archivo “csv_A” . . . . .	112
3.38. Estructura del archivo “csv_B” . . . . .	113
3.39. Ejemplo del calculo en la AS de la conducta alimentarse . . . . .	119
3.40. Ejemplo del calculo de la AS sobre la duración del desplazamiento en el entorno	119
3.41. Cronología de las conductas presentadas dentro del área de cocina . . . . .	123
3.42. Cronología de las conductas presentadas dentro del cuarto . . . . .	124
3.43. Ponderacion de las categorias obtenidas del uso diario de la App “Daylio” . . . . .	126

3.44. Ejemplo parcial del archivo csv_Daylio . . . . .	126
3.45. Información parcial de los csv correspondientes a estrés, FC y hrs de sueño .	127
3.46. Ejemplo de matriz de correlación entre 3 variables X, Y, Z . . . . .	128
3.47. Matriz de correlación obtenida a partir de la unión y reducción de atributos entre los datasets (Entorno IoT, smartwatch y App móvil) . . . . .	129
3.48. Ejemplo de tuplas a eliminar del dataset obtenido del entorno IoT . . . . .	131
3.49. Evaluación de los algoritmos empleados . . . . .	134
3.50. Resultados sobresalientes del atributo Hrs_Sueño dentro de la tuplas seleccionada por los modelos . . . . .	136
3.51. Ejemplo final del sistema trabajando en la detección de conductas registradas en el archivo “csv_A” bajo los intervalos de tiempo establecidos . . . . .	140
3.52. Ejemplo final del sistema trabajando en la detección de conductas registrada en el archivo “csv_B” bajo los intervalos de tiempo establecidos . . . . .	141
3.53. Formación de patrones de conducta sobre el archivo “csv_A” . . . . .	143
3.54. Formación de patrones de conducta sobre el archivo “csv_B” . . . . .	144
3.55. Visualización parcial de la unión de conjuntos . . . . .	145
3.56. Atributos seleccionados para el análisis y detección de conductas anómalas por parte del modelo de ML. . . . .	146
3.57. Matriz de correlación sobre los features del archivo “patronesConducta_Final.CSV”	147
3.58. Dataset “conjuntoConductas_Final” de entrada para el entrenamiento del modelo de ML desarrollado. . . . .	148
3.59. Método del codo empleado para la selección de agrupaciones . . . . .	150
3.60. Evaluación de los métodos K-Means y Herarchical clustering . . . . .	151
3.61. Agrupación propuesta por el modelo K-means . . . . .	151
3.62. Agrupación propuesta por el modelo Herarchical Clustering . . . . .	152
3.63. Agrupación propuesta por el modelo Herarchical Clustering . . . . .	152

3.64. Visualización parcial de los elementos de cada cluster con su respectivo horario	154
3.65. Elementos de cada cluster agrupados por sus respectivos horarios . . . . .	155
3.66. Análisis total de los elementos involucrados en el cluster 0 . . . . .	156
3.67. Análisis total de los elementos involucrados en el cluster 1 . . . . .	156
3.68. Análisis estadístico entre las distancias y su respectivo centroide . . . . .	160
3.69. Histograma de distancias por clúster . . . . .	161
3.70. Histograma de distancias por clúster en conjunto . . . . .	162
3.71. Boxplot referente a cada clúster . . . . .	162
3.72. Anomalías detectadas por Kmeans-boxplot . . . . .	163
3.73. Anomalías detectadas por el modelo COPOD() . . . . .	164
3.74. Anomalías detectadas por el modelo de ML desarrollado . . . . .	165
3.75. Análisis de las anomalías detectadas por el modelo de ML desarrollado . . .	166

# Índice de tablas

2.1. Energías y sus formas de medición aplicado en la elaboración de sensores . . .	31
2.2. Características a tener en cuenta al seleccionar los sensores que integrarán una red IoT . . . . .	32
2.3. Tipos de sensores de temperatura . . . . .	34
2.4. Modos de operación del $\mu C$ ESP8266 . . . . .	36
2.5. Aspectos técnicos de la placa Wemos basada en el $\mu C$ ESP8266 . . . . .	38
3.1. Requisitos solicitados al candidato de estudio . . . . .	58
3.2. Encuesta de participación . . . . .	58
3.3. Entrevista de entorno para participación . . . . .	62
3.4. Áreas para la implementación de dispositivos de monitoreo . . . . .	63
3.5. Sensores empleados en la arquitectura IoT . . . . .	67
3.6. Detalles de la implementación de sensores en entorno IoT . . . . .	69
3.7. Características a considerar de los sensores empleados . . . . .	73
3.8. Relación dispositivos - modo de operación . . . . .	74
3.9. Elementos en conexión IoT . . . . .	87
3.10. Funcionalidades generales de los dispositivos . . . . .	89
3.11. Incógnitas para el diseño del modelo de preprocesamiento encargado de la detección y desarrollo de los patrones de conducta ejercidos por el usuario’ .	114
3.12. Ejemplo del procedimiento de detección de conductas en el archivo “csv_A” a partir de los intervalos de tiempo establecidos.’ . . . . .	116

3.13. Abreviaturas de las actividades monitoreadas por los sensores empleados en la “red 1” área de cuarto . . . . .	122
3.14. Abreviaturas de las actividades monitoreadas por los sensores empleados en la “red 2” áreas de cocina sala y baño . . . . .	122
3.15. Abreviaturas de las actividades monitoreadas en el entorno con sus respectivos ID . . . . .	122
3.16. Algoritmos empleados para la evaluación de datos anómalos en el dataset . .	133
3.17. Análisis estadístico de los elementos involucrados en el cluster 0 . . . . .	158
3.18. Análisis estadístico de los elementos involucrados en el cluster 1 . . . . .	159

# Capítulo 1

## Introducción

Las conductas presentan la ventaja de la accesibilidad, esto debido a que no es necesario un dispositivo específico como un microscopio o telescopio para poder llegar a observar alguna, nos encontramos familiarizados con el concepto debido a la constante interacción de un entorno en el cual al menos un organismo presenta un comportamiento, no obstante el contar con una noción de su existencia y relaciones no implicara que las conclusiones que pudiéramos establecer serían apoyadas por los cautelosos métodos de la ciencia. El estudio del comportamiento es un tema difícil, no porque sea inaccesible sino porque es extremadamente complejo dado que no se presenta en un proceso que pueda ser pausado fácilmente para su observación y análisis, es cambiante, fluido y evanescente por lo que supone un gran esfuerzo de técnicas, metodologías y energía por parte del experto para su estudio. [1]

La conducta se define como la acción o respuesta derivada de un estímulo ejercido sobre un individuo, ésta puede ser cognitiva, fisiológica y motora, lo que permite su observación y medición objetiva. [2]. Cada individuo cuenta con un repertorio de conductas que forman parte de su personalidad. En condiciones habituales una persona no manifiesta la totalidad de conductas con las que cuenta, sino que existe el predominio de algunas brindando orden y sentido al comportamiento. [3]

La conducta de un individuo no tiende a ser lo suficientemente flexible para poder ser

modificada o llevada hacia cualquier dirección, incluso el mismo estímulo del cual fue derivada, puede dar diferentes respuestas, dependiendo la condición biológica o psicológica del individuo, sus acciones recientes, el entorno en el que se encuentra, entre otros, esto debido a los sistemas y tendencias con los que cada organismo nace, los cuales establecen límites a la forma en que ocurre el aprendizaje y el impacto que puede tener. [4]

Las acciones (respuestas) no ocurren aisladas unas de otras, por ende, un patrón de conducta consiste en organizar las acciones individuales de una persona en secuencias de conductas funcionalmente eficaces que conforman la personalidad del individuo. [4]. Las conductas se presentan con la finalidad de actuar o adaptarse sobre un entorno, son controladas por las consecuencias que derivan al presentarse, es decir se repiten si las conductas son premiadas y se extinguen si no lo son. [5]

## 1.1. Planteamiento del problema

La calidad de vida y sobrevivencia misma de un individuo depende de la coordinación de conductas que presente frente al entorno en el que se encuentre. Pues la realización de actividades, así como la actitud que el individuo presente ante estas al ejercerlas dependerá en gran medida de los patrones de conducta que esté haya presentado a lo largo del día, por lo tanto, es importante detectar los estímulos que pudieran generar un impacto negativo en estos favoreciendo el desarrollo de conductas de riesgo, las cuales no sólo ponen en peligro la salud, reduciendo la expectativa y calidad de vida de un individuo, sino que con frecuencia afectan a los que comparten el mismo entorno [6]; [7]; [8].

Entre los problemas de salud derivados de las conductas de riesgo se encuentran las enfermedades cardiovasculares, cáncer, sida, enfermedades mentales entre otras. Esto debido a su influencia sobre el desarrollo de hábitos y estilos de vida: alimentación, higiene, sedentarismo, estrés etc. [9]

El cuidado y atención a las conductas se ha vuelto relevante incluso para el sector empresarial, debido a la repercusión de las conductas en la salud de un trabajador y en un entorno laboral, la secretaría del trabajo y prevención social ha emitido la norma oficial mexicana (NOM-035-STPS-2018) , la cual tiene como objetivo establecer los elementos necesarios para identificar, analizar y prevenir los factores de riesgo psicosocial, así como promover en los centros de trabajo entornos organizacionales favorables. [10]; [11]

En la actualidad, la psicología cuenta con la posibilidad de apoyarse en las tecnologías durante la aplicación de estudios que permiten identificar los estímulos causantes de alteraciones en la conducta de un individuo. Tal es el caso de la aplicación del electroencefalograma, estudio que detecta la actividad eléctrica del cerebro lo que permite encontrar y registrar los patrones de las ondas cerebrales de una persona y con ello diagnosticar anomalías dentro de estos [12] ;[13]; No obstante en la actualidad el constante desarrollo tecnológico permite el diseño de nuevos instrumentos para el análisis y detección de conductas anómalas de un

individuo con base a los patrones de comportamiento ejercidos por este a lo largo del día.

Aplicar IoT para sensar un entorno permite obtener grandes cantidades de información sobre la persona que interactúa dentro de éste. Y gracias a la reducción de costos en los sensores y el aumento en su rendimiento su implementación se facilita. [14]

Realizar cómputo en la nube de la información que un entorno IoT puede brindar, requiere una latencia de espera que depende tanto de la información recibida como del ancho de banda y la red proveniente, EC es un paradigma de la computación que permite trabajar los datos generados de manera más cercana a la fuente de origen al permitir realizar un mayor procesamiento de datos dentro de misma red sin necesidad de acceder continuamente a la nube.[15]

Por los motivos expuestos se propone el desarrollo de un sistema que consiste en la recaudación de datos generados a partir de la interacción del individuo con su entorno, así como el registro de actividades que realiza fuera del mismo a través de una aplicación móvil que debe ser actualizada diariamente. Finalmente, un reloj inteligente se encarga de monitorear el estado físico del individuo generando datos estadísticos que complementen la información obtenida con los métodos anteriores.

La recaudación de datos y su procesamiento alimentarán un modelo de aprendizaje que una vez entrenado será capaz de aprender los patrones de conducta que el individuo ejerce en su vida diaria, teniendo la capacidad de detectar cuando una conducta no pertenezca al patrón de comportamiento establecido.

Los beneficios que brinda el implementar *Edge Computing* para el entorno inteligente es menor dependencia de la nube, mayor velocidad en el tratamiento de los datos, se considera la seguridad del consumidor y los requisitos de privacidad, entre otros [9].

## 1.2. Hipótesis

Mediante el uso de tecnologías para la recolección, procesamiento y análisis de datos en el contexto de un entorno sensado y monitoreo del estado físico es posible obtener la información suficiente para el entrenamiento de un modelo de aprendizaje capaz de detectar conductas anómalas dentro de los patrones de conducta ejercidos por una persona.

### 1.3. Objetivos

**Generales:**

Determinar el conjunto de patrones de conducta que permitan la detección de conductas anómalas en un individuo a través de acciones específicas ejercidas dentro su entorno habitacional, fuera de éste y el monitoreo de su estado físico.

**Específicos :**

1. Adquirir información sobre los patrones de conducta y la consistencia de éstos aplicando Edge Computing (EC) en una red IoT.
2. Identificar los patrones relevantes que puedan ser la base para detectar conductas anómalas en un individuo.

## 1.4. Metas

1. Identificar variables de estudio de comportamiento.
2. Implementar una red de monitoreo dentro del entorno habitacional aplicando IoT y Edge Computing.
3. Desarrollar un modelo de ML capaz de detectar patrones de conducta ejercidos por un individuo.
4. Identificar los patrones de conducta relevantes para generar la base de conductas estándar en la persona.
5. Desarrollar un modelo para la detección de conductas anómalas en base a los patrones de conducta establecidos.

# Capítulo 2

## Fundamentos teóricos

La selección natural ha demandado cambios que abarcan desde lo físico (color, forma, tamaño etc.) hasta el surgimiento de conductas necesarias para la supervivencia de los seres vivos, de acuerdo con las exigencias de sus particulares entornos. [3]; [16]; [17]

### 2.1. Estímulo

Se conoce como estímulo a cualquier acto, influencia o agente que pueda crear una respuesta (conducta) en el ser vivo, ejemplos de estímulos son las condiciones climáticas (frío, calor excesivo), condiciones del entorno (polen, ruido, personas), físicos (ardor, comezón), etc. El impacto de un estímulo varía dependiendo de la persona, es decir, el mismo estímulo puede generar diferentes respuestas(conductas).[5]

Nos encontramos expuestos a diversos estímulos en la vida diaria, desde visualizar color, brillantez y textura en una pintura, percibir olores de un entorno, hasta las sensaciones táctiles de la pared donde nos recargamos, el suelo que pisamos o la silla en la que descansamos. Si tuviéramos que responder a todos los estímulos que percibimos nuestra conducta sería desorganizada y caótica. Los efectos de habituación y sensibilización son los que permiten priorizar los estímulos a los que debemos enfocar una respuesta.[4]

## 2.2. Hábitos

Hábito es la sensibilidad o accesibilidad especial a estímulos específicos, de predilecciones establecidas. Su esencia no consiste en la simple repetición de actos específicos, sino en la predisposición adquirida hacia modos o formas de reacción, en otras palabras, voluntad. Si cada hábito se encontrara aislado sin afectar y ser afectado por otros, no existiría el carácter; por lo tanto, la conducta carecería de unidad pues el funcionamiento continuo de los hábitos en cada acto es necesario para brindar unidad y sentido a la conducta. [18]

## 2.3. Conducta

La conducta se define como la acción o respuesta derivada de un estímulo ejercido sobre un individuo, esta puede ser cognitiva, fisiológica y motora, lo que permite su observación y medición objetiva. [2]; [1]

Cada individuo cuenta con un repertorio de conductas que forman parte de su personalidad. En condiciones habituales una persona no manifiesta la totalidad de conductas con las que cuenta, sino que existe el predominio de algunas. Modificando las condiciones todos los individuos pueden manifestar conductas propias con intensidad, frecuencia o durabilidad variables. [3]

Las conductas se presentan con la finalidad de actuar o adaptarse sobre un entorno, son controladas por las consecuencias que derivan al presentarse, es decir, a medida que el entorno cambia las conductas con el fin de adaptarse dan paso a nuevas formas de comportamiento y las conductas antiguas que ya no son adaptativas se desvanecen.[2]

### 2.3.1. Patrón de conducta

Las respuestas no ocurren aisladas unas de otras, la organización secuencial de la conducta implica que las acciones individuales se organizan con el fin de crear una secuencia de conductas funcionalmente eficaces que forman parte del comportamiento de un individuo.

Toda conducta motivada ya sea a alimentarse, buscar una pareja potencial o alimentar crías implica secuencias de acciones organizadas de manera sistemática. [4]

### 2.3.2. Observacion y deteccion de patrones de conducta

La observación y detección de patrones de conducta en los individuos es necesario para el estudio del comportamiento de éstos, el desarrollo de tecnologías para la detección e interpretación de patrones de conducta es de interés en múltiples áreas de las ciencias computacionales, tales como domótica, seguridad, interacción humano computadora, entre otras.[19]

El área de la psicología también se ha visto beneficiada ante el desarrollo de tecnologías que permiten que, en conjunto con la observación y análisis de un experto, la detección de patrones de conducta en un individuo sea factible y eficaz, como ejemplo tenemos el electroencefalograma.

**Electroencefalograma.** En conjunto con la interfaz computacional, el electroencefalograma permite crear el puente entre el cerebro y el mundo exterior, al permitir estudiar el funcionamiento del sistema nervioso a través de la medición y registro de la actividad eléctrica del encéfalo, concretamente de la corteza cerebral. El estudio puede realizarse al individuo durante la vigilia o estando dormido, según el objetivo con el que se realice la prueba. [20]

**Aplicación** Para su aplicación será empleado un instrumento llamado electroencefalógrafo, el cual por medio de electrodos colocados en áreas específicas del cuero cabelludo permite medir la diferencia de voltajes generada entre electrodos colocados en pares, que en su conjunto logran la transmisión de los impulsos eléctricos generados por la comunicación neuronal. Detectar el patrón de funcionamiento estándar del encéfalo es necesario para proseguir con el estudio, para ello se le pide al individuo realizar respiraciones lo más inmóvil posible tanto

con los ojos cerrados como abiertos con la finalidad de poder registrar un ritmo basal. El experto puede solicitar el realizar algunos movimientos con el fin de registrar los cambios generados por éstos en la actividad cerebral. [21]

### **El impacto de los estímulos en el patrón de funcionamiento del encéfalo**

La estimulación externa o interna tiene un impacto en la activación cerebral por lo que la aplicación del electroencefalograma es requerida cuando se requiere conocer la respuesta o respuestas ante un determinado estado, estímulo[13]; [21]. De esta manera se puede explorar el funcionamiento del encéfalo de un paciente en condiciones normales y su reacción ante estímulos o actividades concretas. [22]

## **2.4. La cuarta revolución industrial**

Mencionar la palabra “revolución” implica un abrupto y radical cambio, desde el surgimiento de nuevas tecnologías y procesos que provocan un cambio profundo en los sistemas económicos y las estructuras sociales, las revoluciones han estado presentes.

La cuarta revolución industrial representa formas completamente nuevas donde las tecnologías se integran en la sociedad e industria, a través de su estrecha cooperación con todas las unidades productivas de la economía.

Las tecnologías digitales que cuentan con hardware, software y comunicación a través de redes no son nuevas, pues estas han sido parte de la innovación tecnológica desde la tercera revolución industrial, sin embargo, en la actualidad dichas tecnologías se han vuelto más sofisticadas e integradas dando paso al desarrollo de nuevos avances tecnológicos. [23]

Lo que caracteriza a esta revolución digital es una internet mucho mas ubicua y móvil, sensores mas potentes y pequeños que han logrado disminuir en precios y la aplicación de inteligencia artificial en cada vez más sectores industriales, esto último debido a que

la IA ha logrado un progreso impresionante impulsado por aumentos exponenciales en la computación y la vasta información de datos que pueden ser obtenidos a través de softwares y el monitoreo en espacios ya sean personales o sociales.[24]

## 2.5. El internet de las cosas

El internet de las cosas (*Internet of Things, IoT*) es uno de los pilares fundamentales del desarrollo de la industria 4.0, logrando la comunicación de dispositivos conectados a internet a través de sistemas embebidos, redes de comunicación, mecanismos de computación de respaldo y aplicaciones típicamente en la nube.

Gracias a la comunicación entre los dispositivos se nos permite acceder a información en internet, capturar, almacenar y consultar datos, logrando la interacción entre el usuario, además de otros sistemas y servicios que en su conjunto logran la operación de ambientes cada vez más conectados e inteligentes. [25]

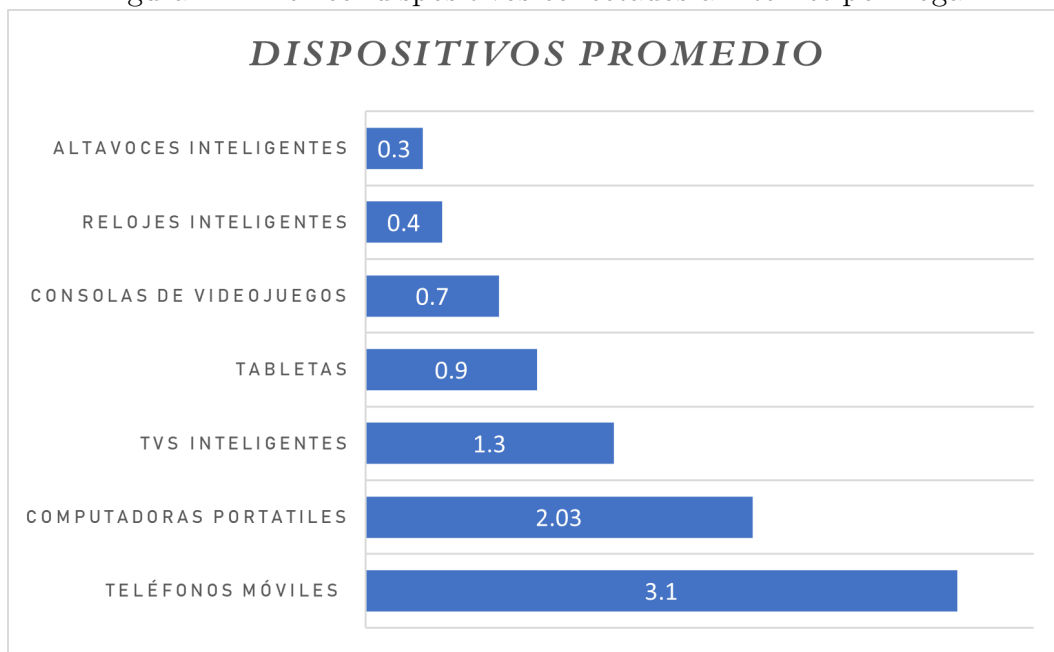
Este nivel de conectividad aumenta la confiabilidad y eficiencia al lograr la accesibilidad a los datos y debido a las grandes aplicaciones del IoT, éste tiene el potencial de remplazar a las personas como el mayor consumidor y productor de información ya que por sí mismo al sensar constantemente un entorno el IoT permite generar grandes cantidades de información sobre la persona que interactúa dentro de éste.[26]

Adicionalmente el Internet de las cosas se ha convertido en un impulsor de nuevos modelos de negocios, debido a su influencia en el mejoramiento de procesos, reducción de costos e integración en áreas de la inteligencia artificial, realidad virtual y aumentada; además de convertirse en un agente recolector de Datos (Big data driver) con relevancia en campos de economía de consumo y procesos industriales.[14]

De acuerdo con la Statista(portal estadístico internacional), se espera que el próximo año el mercado global de IoT alcance los USD \$520 mil millones, esto en base a que la demanda generada por los productos conectados a IoT no muestra desaceleración [27]. En México el consumo de dispositivos conectables a internet dentro de un hogar aumentara

considerablemente en los próximos años. La Fig. 2.1 presenta una encuesta realizada en agosto del 2020 donde los hogares mexicanos reflejaron en promedio 9 dispositivos conectados a internet simultáneamente, los teléfonos móviles tuvieron mayor presencia en los hogares a partir de ese mes, con un promedio de 3.1, mientras que las computadoras o portátiles se llevaron el segundo lugar con 2.03. [27].

Figura 2.1: México: dispositivos conectados a internet por hogar



1

### Impacto de la pandemia global COVID -19 en el IoT

Como resultado de la pandemia mundial COVID-19, el desarrollo de tecnología IoT ha tenido un gran impulso, esto debido a las necesidades de trabajo remoto que se han presentado, una de ellas es el monitoreo de sistemas a distancia que permite un entorno de trabajo seguro sin necesidad de afectar la producción laboral.

Con países aún siendo afectados por esta pandemia y cuarentenas, existe la incertidumbre de un pronto retorno a la normalidad, por lo que el desarrollo de tecnologías IoT continúa

<sup>1</sup>S. R. Department, "México: devices connected internet per household." <https://www.statista.com/statistics/1174769/devices-connected-internet-householdsmexico/>, 2020.

con el fin de garantizar que las personas puedan seguir trabajando en un entorno seguro.[28]

### 2.5.1. Crecimiento del Internet de las cosas en el 2021

El desarrollo de tecnologías basadas en los sistemas embebidos que interactúan en una red IoT será de importancia en diferentes sectores de la economía, como ejemplo tenemos:

#### Seguridad IoT

Muchas industrias han experimentado una baja demanda a causa del COVID-19, no obstante, ese no es el caso del área de seguridad IoT, pues desde el impacto de la pandemia mundial, muchas empresas y sectores han tenido que adaptarse con las tecnologías necesarias, dando lugar a vulnerabilidades que son un blanco a los ciberataques. [29]

#### Cuidado de la salud

El área de la salud también ha sido beneficiada gracias a los avances en las tecnologías, desde el surgimiento de la pandemia mundial Covid-19, entre ellos se encuentran:

1. Diagnósticos digitales: incorporación de dispositivos IoT para realizar el diagnóstico de enfermedades, como ejemplo tenemos el termómetro inteligente Kinsa..
2. Monitoreo remoto: a través de sitio web se ofrece el monitoreo de salud principalmente dirigido a personas de la tercera edad, páginas como Livongo ofrecen servicios de consultas médicas y seguimiento de salud, a través de dispositivos de control que son enviados al paciente.
3. Asistencia robótica: principalmente en China, el uso de robots ha sido implementado con el fin de apoyar en la limpieza y desinfección de las áreas

médicas, al igual que la entrega de medicamentos. [30]

### **Transportación**

Siendo un gran mercado debido a su gran relevancia y demanda, este sector de acuerdo a Allied Market Research reportó que el mercado valuado en \$135 billones de dólares en el 2016 pronostica un crecimiento a \$328 billones de dólares para el 2023. Esto debido a beneficios como:

- Mejorar la experiencia con el cliente.
- Reforzar la seguridad.
- Mejoras ambientales. [31]

#### **2.5.2. Retos del internet de las cosas**

Los dispositivos inalámbricos ofrecen mucha ventaja en términos de costo, flexibilidad, facilidad de instalación y remplazo, no obstante, para su aplicación en IoT se requiere contemplar los desafíos que el entorno va a requerir para un óptimo funcionamiento como:

- Energía
- Escalabilidad
- Tolerancia a fallas
- Consumo de energía
- Problemas de seguridad
- Capacidad computacional
- Aceptación de los usuarios.
- Seguridad en la comunicación

- Autenticación y autorización
- Disponibilidad de Internet en todas partes
- Desarrollo de sistemas de detección inteligente de bajo costo.

[32]

Estos puntos deben ser tomados en cuenta pues trabajar con un entorno IoT implica mayores cantidades de dispositivos inalámbricos conectados dentro de un entorno manteniendo constante comunicación y tareas, las cuales pueden verse afectadas por una gran latencia de espera derivada de la organización y estructura de la red. [33]

### **2.5.3. Tecnología en un entorno IoT**

Un hogar típico en la actualidad cuenta con un gran número de dispositivos embebidos en el entorno, cada uno con funciones específicas y usualmente no situados en un espacio fijo, celulares, tabletas, iPod son ejemplos de ello, no obstante, también en electrodomésticos como arroceras, lavaplatos, refrigeradores. [34]

Los beneficios que brindan los sistemas embebidos van desde el monitoreo de un entorno hasta la transformación de una vivienda en entorno inteligente, pues en complemento con sensores, los sistemas embebidos pueden utilizar los datos generados por los usuarios para así adaptar el entorno a las necesidades de éste y de esta manera mejorar la interacción humano-computadora.[34]

#### **2.5.3.1. Sensores**

Dispositivos que varían en precio, tamaño, consumo y funcionalidad. Conformados por células sensibles que detectan variaciones en una magnitud física, química y biológica del entorno transformándolas en señales eléctricas útiles para medición y posteriormente alimentación de otro sistema.[35]

### Categoría de sensores en base a su forma de energía

Existe una gran variedad de sensores en el mercado, orientados a trabajar con formas específicas de energía que controlan las mediciones resultantes. La Tabla 2.1 presenta ejemplos de estas energías y la medición que pueden derivar.

Tabla 2.1: Energías y sus formas de medición aplicado en la elaboración de sensores

<b>Energías</b>	<b>Formas de medición</b>
Mecánica	Longitud, área, volumen, todas las derivadas del tiempo como velocidad lineal / angular, aceleración lineal / angular, flujo másico, fuerza, par, presión, longitud de onda e intensidad acústicas
Temperatura	Temperatura, calor específico, entropía, flujo de calor, estado de la materia
Eléctrica	Voltaje, corriente, carga, resistencia, inductancia, capacitancia, constante dieléctrica, polarización, campo eléctrico, frecuencia, momento dipolar
Magnética	Intensidad de campo, densidad de flujo, momento magnético, permeabilidad
Radiante	Intensidad, fase, longitud de onda, polarización, reflectancia, transmitancia, índice de refracción
Química	Composición, concentración, velocidad de reacción, pH, potencial de oxidación / reducción

2

### Características a contemplar para la aplicación de un sensor

Es necesario conocer la arquitectura IoT que se desea emplear en el entorno para proceder a la selección de dispositivos de monitoreo, de esta manera se aprovechan al máximo los beneficios que estos pueden brindar al igual que se tienen en cuenta las limitaciones que dichas herramientas presentan, previniendo futuros inconvenientes. La Tabla 2.2 presenta algunas de las características principales a contemplar.

<sup>2</sup>C. Technologies, N. Eror, C. on New Sensor Technologies: Materials, Applications, N. R. Council, D. on Engineering, P. Sciences, N. M. A. Board, C. on Engineering, and T. Systems, Expanding the Vision of Sensor Materials. Online access: National Academy of Sciences National Academies Press, National Academies Press, 1995

Tabla 2.2: Características a tener en cuenta al seleccionar los sensores que integrarán una red IoT

Estático	Dinámico
Precisión	Respuesta de error dinámico
Distorsión	Histéresis
Inestabilidad	Mínima señal detectable
Ruido	No linealidad
Rango de operación	Sensibilidad
Respuesta de paso	Repetibilidad
Umbral	Fuente de alimentación

3

### 2.5.3.2. Consumo de energía

Reducir el consumo de energía en los dispositivos empleados en IoT es un importante factor por considerar, ya que existen eventos que no son propensos a surgir del diario y entornos donde no es posible un acceso constante, tal es el caso de regiones desérticas, laderas empinadas, zonas selváticas entre otros [35].

Para dejar en claro la importancia del consumo de energía de un dispositivo, se facilita el siguiente ejemplo:

*“Se desea medir el promedio de lluvias al año en un oasis”*

Sabiendo que es un área de difícil acceso diario, donde no ocurren lluvias frecuentemente y no hay conexiones eléctricas que puedan suministrar energía permanentemente a nuestros dispositivos, nos damos cuenta de que debemos ser cuidadosos con el consumo y modo de funcionamiento de nuestras herramientas.

En caso de no ser así, se corre el riesgo de un funcionamiento muy limitado por parte de los sensores, ya sea por un bajo desempeño en las tareas o su nulo funcionamiento causado por la falta de alimentación de energía.

<sup>3</sup>C. Technologies, N. Eror, C. on New Sensor Technologies: Materials, Applications, N. R. Council, D. on Engineering, P. Sciences, N. M. A. Board, C. on Engineering, and T. Systems, Expanding the Vision of Sensor Materials. Online access: National Academy of Sciences National Academies Press, National Academies Press, 1995

### 2.5.3.3. Sensores empleados en la arquitectura IoT

En función de la variable a medir o detectar podemos elegir entre sensores de contacto, temperatura, ópticos, magnéticos, infrarrojos, entre otros. A continuación se presenta una breve explicación orientada a los dispositivos empleados en la presente investigación.

#### **Sensores de contacto**

Se emplean para detectar la posición límite de componentes mecánicos, son aplicados en puertas, ventanas o mecanismos que impliquen abrirse y cerrarse. [25]

#### **Sensores optoelectrónicos**

Funcionamiento relacionado directamente con la luz, aplicados en la detección de personas u objetos que interrumpen el haz de luz que recibe el sensor. Se pueden distinguir en 3 grupos:

**Emisores:** emiten luz al ser activados por energía eléctrica.

**Detectores:** generan una pequeña señal eléctrica al ser iluminados

**Fotoconductores:** conducen la radiación luminosa desde un emisor a un receptor.[36]; [37]

#### **Sensores de temperatura**

Resistencias cuyo valor se basa en la temperatura la cual disminuye o aumenta, transformando los cambios de temperatura en cambios de señales eléctricas. [25]

Una de las aplicaciones de estos sensores es la de protección de entornos y equipos hacia las variaciones extremas de temperatura, funcionando como interruptor de corriente en caso de ser necesario. La Tabla 2.3 presenta las categorías correspondientes a este tipo de sensores.

Tabla 2.3: Tipos de sensores de temperatura

<i>Categoría</i>	<i>Termopares</i>	<i>Detectores de temperatura por resistencia</i>	<i>Termistores</i>
<i>Rango de temperatura (grados Celsius)</i>	-180 a 2,320	-200 a 500	-90 a 130
<i>Tiempo de respuesta</i>	Rápido (microseg.)	Lento (segundos)	Lento (segundos)
<i>Tamaño</i>	Largo (~ 1 mm)	Pequeño (5 mm)	Pequeño (5 mm)
<i>Precisión</i>	Baja	Media	Muy Alta

4

### Sensores de movimiento

Su funcionamiento está basado en tecnología infrarroja, en general, este sensor se encuentra constituido por un cristal Fresnel que permite a dos ranuras formar un ampliado exterior, permitiendo crear zonas de detección.

Todo esto gracias a que los materiales que conforman las ranuras responden a los rayos infrarrojos y calor emitidos por el cuerpo, lo que permite captar en tiempo real los movimientos que se generan en un entorno. [25]

El mayor uso que se le da a estos sensores es su incorporación en alarmas de seguridad y electrodomésticos.[35]

### Sensores ultrasónicos

Obtiene la distancia mediante el uso de ondas ultrasónicas, el cabezal emite una onda ultrasónica y recibe la onda reflejada que retorna desde el objeto, lo que mide este sensor es el tiempo entre emitir el pulso y la recepción de éste.[38]

### Sensores de humedad

Estos dispositivos detectan y miden la cantidad de vapor de agua o humedad contenida en el aire, estos sensores tienden a complementarse con los sensores

<sup>4</sup>P. Lea, Internet of Things for architects : architecting IoT solutions by implementing sensors, communication infrastructure, edge computing, analytics, and security. 2018

de temperatura, pues en conjunto son especialmente útiles en los sistemas de ventilación, importantes en la producción o fabricación de productos. [39]

### **Sensores Flex**

Son resistencias analógicas, que producen una resistencia variable en función del grado en el que esta doblada, es decir, convierte el valor de la curvatura en distintos valores de resistencia eléctrica.

Por lo general su presentación es en una delgada tira de 5 cm de largo que varía en resistencia de aproximadamente 10 a 50 kOhms, pero existen presentaciones en forma redondeada de menor tamaño. [40]

#### **2.5.3.4. Sistemas embebidos**

Un sistema embebido o integrado es un dispositivo electrónico controlado por su propio microprocesador o microcontrolador interno y el cual la ROM del microcontrolador es cargada con un conjunto de instrucciones específicas para una determinada tarea, una impresora es un ejemplo de estos sistemas integrados ya que en su procesador interno se encuentra cargado un conjunto de instrucciones para realizar una determinada tarea. [41]

### **Microcontroladores**

Un microcontrolador ( $\mu C$ ) es un procesador con elementos incorporados entre ellos están: RAM, ROM, puertos E/S, reloj externo, algunos con convertidor analógico-digital. Los microcontroladores son ampliamente empleados en productos de sistemas embebidos o integrados, debido a su rapidez, compacto tamaño y coste. [42]; [41]

#### **2.5.3.5. Microcontrolador ESP8266**

Es un microcontrolador producido por Espressif Systems, el cual, debido a su compacto tamaño, bajo costo y adaptabilidad con sistemas embebidos, ha sido aplicado en gran

medida para tecnologías IoT. Pese a que su predecesor ESP32 ha llegado con nuevas funcionalidades y completos, el  $\mu C$  ESP8266 sigue teniendo una gran demanda en la actualidad, para mayor información consultar Tabla 2.5 donde se presentan los aspectos técnicos de la placa.

### Modos de funcionamiento del ESP8266

El  $\mu C$  ESP8266 tiende a ser empleado en dispositivos embebidos aplicados en entornos IoT debido a que este ofrece 3 diferentes modos de operación, lo que permite una mayor flexibilidad en el consumo de energía. La Tabla 2.4 presenta los elementos involucrados en cada modo operacional

Tabla 2.4: Modos de operación del  $\mu C$  ESP8266

<b>Modulo</b>		<b>Modem-sleep</b>	<b>Light-sleep</b>	<b>Deep-sleep</b>
<b>Wi-Fi</b>		OFF	OFF	OFF
<b>Reloj del sistema</b>		ON	OFF	OFF
<b>RTC</b>		ON	ON	ON
<b>CPU</b>		ON	Pendiente	OFF
<b>Corriente empleada</b>		15 mA	0.4 mA	~ 20 $\mu A$
<b>Corriente Promedio</b>	DTIM = 1	16.2 mA	1.8 mA	-----
	DTIM = 3	15.4 mA	0.9 mA	
	DTIM = 10	15.2 mA	0.55 mA	

5

### Modem Sleep

Es el estado por default en el ESP8266, este modo es empleado en caso de que la aplicación requiera funcionando totalmente el CPU, solo puede ser empleado cuando el  $\mu C$  se encuentra configurado en modo estación (WIFI\_STA).

El estado *modem sleep* es interrumpido cuando el dispositivo es conectado al router, mientras tanto, mantiene el circuito del módem WiFi apagado, es decir

<sup>5</sup>E. White, Making Embedded Systems: Design Patterns for Great Software, vol. 2011. 2011.

una conexión WiFi sin datos de transmisión, esto con la finalidad de ahorro de energía.

Con un consumo típico de 15 mA el  $\mu C$  opera bajo tres modos de DTIM que influye directamente en el ahorro de energía como se puede observar en la Tabla 4.

### **Light Sleep**

Este estado permite un consumo de energía típico de 0.4 mA, el cual es menor al consumo *modem sleep*, la causa radica en mantener el circuito del módem Wifi en suspensión al igual que el CPU y en adición el reloj del sistema es apagado.

La suspensión del CPU en este caso no es definitiva, debido a que este es reactivado en caso de ser necesario, por ejemplo, en caso de no existir la transmisión de datos, el circuito del módem Wifi es apagado y el CPU suspendido con la finalidad de disminuir el consumo de energía.

El dispositivo se mantiene activo pero suspendido en algunos módulos, entre ellos el CPU, el cual para ser despertado necesita una interrupción GPIO externa y el proceso no sobrepasa los 3 ms.

### **Deep Sleep**

En caso de requerir un mayor ahorro de energía previo a los dos anteriores, se recomienda el modo deep sleep, con un consumo de aproximadamente 20 el  $\mu A$ , este modo es empleado en proyectos donde se exige el funcionamiento de eventos muy prolongados y haciendo uso de baterías.

A diferencia de los anteriores este mantiene apagado el CPU, además del WiFi y reloj de sistema, siendo el módulo RTC el único responsable de despertar periódicamente el microcontrolador y en caso de ser necesario guardar datos en la memoria de este.

Este modo de funcionamiento permite despertar de dos maneras el  $\mu C$ , a través de un tiempo predefinido y por medio de una interrupción GPIO externa, en ambos casos se requiere una configuración física en el  $\mu C$ .

La conexión necesaria para generar el despertar del  $\mu C$  es puentear D0 a RST, ya que D0 es el único pin que permite realizar el reset necesario.

### Wemos D1 mini

Lolin wemos D1 es una placa basada en el esp8266, de sus mayores beneficios esta su tamaño compacto, lo que permite su uso en proyectos que requieran dispositivos de tamaño reducido con acceso a WIFI. Gracias a su regulador de voltaje, esta placa puede ser conectada directamente a nuestro ordenador a través de un cable USB y programada empleando el entorno de desarrollo integrado (IDE) de Arduino, como fue el caso de uso para el desarrollo de esta investigación.

Tabla 2.5: Aspectos técnicos de la placa Wemos basada en el  $\mu C$  ESP8266

<i>Aspectos técnicos</i>	
Voltaje de operación	3.3 V
Pines E/S digitales	11
Pines de entradas análogas	1 (3.2 v Max)
Velocidad de reloj	80/160MHz
Flash	4M Bytes
Tamaño	34.2 x 25.6mm
Peso	3g

6

## 2.6. Edge Computing (EC)

Realizar cómputo en la nube de la información que un entorno IoT puede brindar, requiere una latencia de espera que depende tanto de la información recibida como del

<sup>6</sup>E. White, Making Embedded Systems: Design Patterns for Great Software, vol. 2011. 2011.

ancho de banda y la red proveniente, variables críticas para aplicaciones (ej: automóviles inteligentes, realidad aumentada) donde el tiempo de respuesta es crucial para un óptimo funcionamiento.[32]

EC es un paradigma de la computación que permite trabajar los datos generados de manera más cercana a la fuente de origen, de manera que se reducen los tiempos de respuesta entre la comunicación de dispositivos conectados en una red IoT. [15]

EC no sólo beneficia disminuir la latencia de espera, también brinda una mayor seguridad en la conectividad al centro de acceso de datos y ante ataques que pongan en peligro la integridad y privacidad de la información compartida [43].

## 2.7. El papel de IA en los entornos inteligentes

Los entornos IoT generan beneficios debido al confort y apoyo que brindan a los usuarios, estos suelen emplear una gran variedad de dispositivos con la finalidad de monitorear características ambientales que posteriormente sirven de base para la detección de cambios en el entorno.

Entre las características ambientales que buscan monitorearse están la luz, temperatura, humedad, presencia en el espacio, la cual es captada a través de sensores de movimiento, entre otros.

Como complemento a estos entornos inteligentes, se puede contar con el apoyo de dispositivos de biodetección, empleados para monitorear el estado de salud de un usuario, un ejemplo es el reloj Garmin vívosmart 3, un reloj inteligente que permite el seguimiento de actividad física, con herramientas que permiten el monitoreo de frecuencia cardiaca, estrés, horas de sueño, estimación de VO2 máximo y demás funciones que lo hacen un dispositivo útil y práctico para el usuario.[44]

Los entornos inteligentes van desde espacios públicos como centros comerciales, aeropuertos, estaciones de transporte, hasta espacios mas privados tal es el caso de una oficina u hogar.[35]

### **Diciplinas involucradas en el desarrollo de un entorno inteligente**

La colaboración de múltiples diciplinas trabajando en conjunto es necesaria para el desarrollo de un entorno inteligente, entre estas podemos encontrar: la inteligencia artificial (IA), visión de computadora, aprendizaje automático (ML), robótica, la ingeniería mecánica y de materiales, así como la arquitectura.[14]

Otros campos involucrados son la interacción humano-computadora, psicología y sociología, estos enfocados más en tratar el aspecto humano relacionado al usuario.[14]

El trabajo en conjunto de estas diciplinas permite que un entorno inteligente apoyado por dispositivos de biodetección sea capaz de reconocer las actividades humanas ejercidas dentro de él y analizar los comportamientos de los usuarios. [35]; [34]

### **El objetivo de desarrollar espacios inteligentes**

Complementar entornos IoT con diciplinas como IA, machine learning (ML) y psicología, apoya a cumplir un objetivo importante en la investigación de espacios inteligentes, el cual es detectar y predecir el comportamiento, así como las preferencias humanas basándose en datos generados por usuarios y detectados por el entorno a través de diversos sensores.

#### **2.7.1. Inteligencia artificial**

El concepto de inteligencia artificial puede ser difícil de definir, debido a preguntas como ¿Qué es la inteligencia?, cada individuo puede tener su propia definición dependiendo del enfoque que éste le quiera dar, algunas personas tendrán como referencia, la rapidez con la que se resuelven cálculos matemáticos, no obstante una calculadora básica puede derrotar muy fácilmente a la mayoría de las personas resolviendo operaciones matemáticas y sin embargo esto no la hace mas inteligente que cualquier persona.

John McCarthy padre fundador de la diciplina, propuso el concepto de inteligencia artificial como:

“La ciencia y la ingeniería de hacer que las computadoras resuelvan problemas y se comporten de formas generalmente consideradas inteligentes.”. [45]

Apoyado en lo anterior complementamos el alcance de la inteligencia artificial de adquirir la experiencia necesaria para resolver problemas sin orientación o instrucciones previas.

## 2.8. Inteligencia artificial en la detección de patrones

Empleada en el área de procesamiento / tratamiento de datos, es la faceta principal después de la recopilación de datos para descubrir observaciones prácticas, localizar patrones y tendencias, que apoyan el rendimiento basado en modelos de análisis de datos o aprendizaje automático, los cuales posteriormente brindan el beneficio de generar recomendaciones personalizadas.

### Relevancia en la detección de patrones

Las máquinas son capaces de realizar una gran cantidad de tareas en menor cantidad de tiempo sin afectar la calidad del resultado obtenido. Con el apoyo de IA este rendimiento puede ser aún mayor, superando ampliamente las capacidades humanas sin perder eficiencia. Entre las aplicaciones de la IA encontramos la detección de patrones o tendencias en un conjunto de datos.[46]

Gracias a la detección de patrones dentro de un conjunto de datos un programa de seguridad sería capaz de anticipar posibles ciber ataques basándose en las anomalías que afectan un patrón de datos establecido en tan solo milisegundos; un sistema de alerta de tsunami detectaría un cambio apenas perceptible en la altura del océano haciendo sonar una alarma; un nuevo fármaco que apoye el tratamiento efectivo de una enfermedad específica podría ser descubierto al encontrar un patrón de arreglos moleculares previamente inadvertido.[46]

Encontrar patrones en conjuntos de datos no está limitado a determinados entornos, su alcance le permite trabajar con información generada en diversos espacios, las viviendas son un ejemplo.[47]

La información que genera el sensar un entorno involucra la interacción del habitante con este, las actividades que realiza en cada sección del hogar, así como su duración y secuencia, generando un patrón de conducta diario. [48]

Una vez establecido el patrón de conducta regular de un habitante, es posible detectar anomalías dentro de éste, al comparar constantemente su desarrollo de actividades actuales, con las previamente establecidas.[48]; [14]

La IA permite realizar esta tarea por medio de diferentes técnicas de ML, el proceso para desarrollar dicha tarea depende de los datos que se manejen, pues en base a estos se trabajará con el aprendizaje supervisado, no supervisado y semi supervisado.

### 2.8.1. Tipos de aprendizaje

#### Aprendizaje supervisado

A partir de un conjunto de variables de entrada se busca encontrar una función que más se aproxime a los valores de salida. Para trabajar el conjunto de valores, se requiere contar con datos etiquetados, es decir, las observaciones de variables de entrada deben contar con su respectiva variable de salida, que constituye nuestra variable dependiente o target.[49]

Debido a que cada entrada cuenta con su respectiva salida, emplear aprendizaje supervisado genera una mayor efectividad a comparación del aprendizaje no supervisado o semi supervisado, sin embargo, el requisito de tener un conjunto de datos etiquetado implica un alto coste, lo que se dificulta al tener grandes cantidades de información.

El aprendizaje supervisado es aplicable en:

- Problemas de clasificación
- Problemas de regresión

[50]

### **Aprendizaje no supervisado**

Este tipo de aprendizaje se caracteriza por no contar con un conjunto de datos etiquetado, es decir, se desconoce la salida a las variables de entrada.

Al no saber qué variables de salida esperar este método no se enfoca en la predicción de valores sino en la detección de patrones o análisis de variables que permita encontrar información relevante en un conjunto de datos.

Aplicar este tipo de aprendizaje resulta más viable en muchos conjuntos de datos dado que la mayoría no cuenta con su variable dependiente (datos etiquetados)[51];[52]

El aprendizaje no supervisado es aplicable en:

- Reducción de dimensionalidad
- Agrupamiento

[50]

### **Aprendizaje Semi supervisado**

El aprendizaje semi supervisado implica un conjunto de técnicas de aprendizaje automático que combina el trabajar con datos etiquetados y no etiquetados con el propósito de facilitar la predicción de salidas.

Este tipo de aprendizaje brinda los beneficios de poder trabajar con datos que cuenten con variable dependiente, lo cual apoya a la la predicción de valores para su posterior clasificación y al mismo tiempo, este método cuenta con datos sin etiquetar que apoyan en la restricción de espacio para posibles reglas de inferencia.

Cabe mencionar que entre los beneficios del aprendizaje semi supervisado se encuentra el mejoramiento predictivo a costa de una pequeña cantidad de datos etiquetados en contraste a una mayor cantidad de datos sin etiquetar.[53]; [52]

### **Aprendizaje por refuerzo**

Los algoritmos de aprendizaje por refuerzo definen modelos y funciones enfocados en maximizar una medida de recompensa, todo en base al entorno donde el agente se desempeña y las acciones.

En el aprendizaje por refuerzo se pretende que el agente se desempeñe basándose en la analogía de vida de los humanos, donde a partir de un entorno nuestras acciones están relacionadas a las recompensas obtenidas de estas, los éxitos o fracasos provenientes de dichas acciones servirán de experiencia para futuras decisiones.[54]

## **2.8.2. Clustering**

El agrupamiento (clustering) en el contexto de la IA, hace referencia a un conjunto de técnicas cuyo propósito es la detección de patrones o agrupamientos dentro de un conjunto de datos, por lo tanto, corresponde a un tipo de aprendizaje no supervisado [55]; [52]. Dentro de la clasificación de agrupamiento destaca:

### **2.8.2.1. Agrupamiento por particiones:**

División del conjunto de datos en subconjuntos no solapados, de manera que cada punto del conjunto pertenece a su propia agrupación.

El agrupamiento puede realizarse a partir de distancia, densidad y modelos, en este trabajo se abordará principalmente el agrupamiento por distancia y densidad.[55]

### **Distancia**

Dado  $k$ , el número de particiones a construir se generan a partir de una partición inicial. Posteriormente aplicando una técnica de reubicación se busca mejorar la calidad de partición, reubicando puntos de un grupo a otro. Ejemplos destacados de este método es K-means, K-medoids, k-nearest neighbor.[52]

### **Densidad**

Se realiza la creación de un clúster, identificando regiones de alta densidad de observaciones rodeadas por regiones de menor densidad. Al contrario de los algoritmos mencionados en el agrupamiento por distancia, el agrupamiento por densidad no requiere una  $k$  inicial, esto debido a que la creación de grupos parte de un mínimo de observaciones vecinas dentro de un radio de proximidad. Ejemplos de esta práctica es el algoritmo DBSCAN y OPTICS.[56];[52]

#### **2.8.2.2. Agrupamiento jerárquico:**

El agrupamiento jerárquico es una alternativa a los métodos de agrupamiento por partición, éste consiste en conjuntos de datos anidados organizados en forma de árbol donde cada nodo (clúster) forma la unión de los subhijos (subclústeres).[56]

Los métodos jerárquicos pueden clasificarse en dos, dependiendo de la estrategia seguida para generar los agrupamientos:

#### **Aglomerativos:**

También llamado algoritmo codicioso, considera cada punto como un clúster que se va fusionando con los pares más cercanos hasta que todos los objetos se encuentren en un solo grupo o hasta que las condiciones

establecidas sean satisfechas.

Esta técnica requiere medir la proximidad entre clústeres, basándose en una medida de similitud se obtiene la distancia entre cada par de unidades, realizando la fusión de los grupos más cercanos en base a la distancia obtenida.[57];[58];[59]

**Enlace Simple:** se calcula la distancia entre todos los posibles pares entre el clúster A y B. La menor distancia obtenida es seleccionada como la distancia entre las dos agrupaciones debido a que se trata de la medida menos conservadora (disimilitud mínima entre grupos)

$$f = \min(d(x, y)) \quad (2.1)$$

**Enlace Completo:** deduce la cercanía entre clústeres tomando la distancia máxima generada al calcular las distancias entre todos los posibles pares entre el clúster A y B (disimilitud máxima entre grupos)

$$f = \max(d(x, y)) \quad (2.2)$$

**Enlace Promedio:** se calcula la distancia entre todos los posibles pares entre el clúster A y B, se realiza el promedio de todas ellas para posteriormente tomarse como distancia final el resultado obtenido. (disimilitud media entre grupos)

$$f = \text{average}(d(x, y)) \quad (2.3)$$

**Enlace centroide:** se calcula el centroide de cada clúster y selecciona la distancia entre ellos como la distancia entre los

dos clústers.

$$d(\text{avg}(X), \text{avg}(Y))d(\text{avg}(X), \text{avg}(Y)) \quad (2.4)$$

Tomamos el promedio de todos los elementos de cada grupo [58];[55]

**Método Ward:** éste es un procedimiento jerárquico aplicable en el agrupamiento aglomerativo donde en cada iteración se identifican aquellas dos agrupaciones cuya fusión conlleva menor incremento de la varianza total intra-clúster. Obtíene las distancias entre dos clústeres usando el incremento de error cuadrático producido al unir dos clústeres. [59]

#### **Divisivo:**

El agrupamiento jerárquico divisivo trabaja de forma descendente, dividiendo el grupo que contiene todos los objetos en grupos más pequeños hasta que cada clúster sólo incluya un elemento, el del objeto mismo o hasta que se satisfaga una condición establecida.

A diferencia del agrupamiento aglomerativo, donde es necesario elegir un tipo de distancia y un método de enlace (linkage), el agrupamiento divisivo sólo requiere la elección de un tipo de distancia(euclidiana, manhattan, máxima,etc).[57]; [59]

### **Ventajas y desventajas del agrupamiento jerárquico**

#### **Ventajas:**

- No es necesario especificar el número de clústers(k) de antemano.
- Es fácil de implementar e interpretar con la ayuda de dendogramas
- El resultado de agrupamiento es siempre el mismo.

### **Desventajas**

- A comparación de otros métodos, éste es más lento, su ejecución depende del conjunto de datos, por lo que si éste es muy grande su tiempo de ejecución incrementará notablemente.[57]

#### **2.8.2.3. Métricas para evaluar el desempeño del modelo empleado**

Cualquier algoritmo de detección de anomalías debe ser evaluado sin importar si fue trabajado a partir del aprendizaje supervisado o no supervisado. La evaluación del algoritmo nos permite saber que tan acertadas fueron sus predicciones, es decir, su desempeño frente a la localización de puntos anómalos dentro de un conjunto de datos. Estos puntos atípicos en comparación con el resto de los datos tienden a encontrarse en cantidades muy reducidas, lo que afecta a las métricas de evaluación que tienden a emplearse, un ejemplo de estas es la precisión (accuracy).

En caso de trabajar con datos supervisados, podemos emplear una matriz de confusión, la cual es un resumen de los resultados de predicción sobre un problema de clasificación que nos permite obtener precisión, valor r-cuadrado, sensibilidad y especificidad. Sin embargo recordemos que en la mayoría de los casos se trabaja bajo datos no supervisados por lo que es necesario aplicar otras métricas que en este caso evalúan aspectos como la separación entre clústeres y la calidad de agrupación de los mismos. Entre algunos de ellos tenemos:

#### **Coefficiente de silueta (Silhouette Coefficient):**

También llamado puntaje de silueta (silhouette score) es una métrica empleada para calcular la calidad de agrupamiento aplicando la siguiente ecuación:

$$S(i) = \frac{(b(i) - a(i))}{\text{máx}\{a(i), b(i)\}} \quad (2.5)$$

Donde  $a$  equivale a la distancia promedio dentro del grupo, es decir, la distancia promedio entre cada punto dentro de un grupo y  $b$  la distancia promedio entre todos los agrupamientos.[60]; [61].

Su rango de valor va de -1 a 1 donde:

- 1: es sinónimo de que los clústeres están asignados de manera incorrecta.
- 0: significa que la distancia entre los grupos no es significativa,
- 1: los grupos están bien separados entre sí y se distinguen fácilmente.

En conjuntos de datos con gran dimensionalidad emplear esta técnica es de gran utilidad, además de que permite guiar en la selección del número óptimo de clústeres a través de sus resultados de evaluación.

### **Calinski-Harabasz**

Siendo un índice de validación interna refleja la similitud entre un grupo y las diferencias entre los otros. La forma de estimar el Calinski-Harabasz se obtiene a partir de la siguiente ecuación:

$$s = \frac{\text{tr}(B_k)}{\text{tr}(W_k)} \times \frac{n_E - k}{k - 1} \quad (2.6)$$

Donde  $k$  equivale al número de clústeres,  $N$  el número de elementos del conjunto de datos,  $\text{Tr}(B_k)$  representa la matriz de dispersión entre los grupos y  $\text{Tr}(W_k)$  la matriz de dispersión dentro del conglomerado.

La métrica es rápida de ejecutar y permite apoyar la elección del número óptimo de agrupaciones.[62] ; [61]

**Davies-Bouldin Index (DB)**

Aplicar el índice Davies- Bouldin (DB) permite obtener la similitud promedio entre las agrupaciones, donde la similitud es una medida que compara la distancia de entre las agrupaciones con el tamaño de los grupos mismos.

Este índice se encuentra definido como:

$$DB = \frac{1}{n} \sum_{i=1}^n \max_{i \neq j} \left\{ \frac{S_n(Q_i) + S_n(Q_j)}{S(Q_i, Q_j)} \right\} \quad (2.7)$$

Donde  $n$  representa el número de clústeres,  $S_n$  la distancia promedio de todos los objetos hacia su centroide y  $S(Q_i, Q_j)$  la distancia entre los centroides del cluster  $i, j$ . [61]

**Dunn (Dunn Index):** Pertenece a los índices de validación interna, debido a que busca el evaluar el desempeño del algoritmo empelado en clustering buscando grupos compactos con una pequeña variación entre los puntos internos y una separación con respecto a las demás agrupaciones.

$$DI_m = \frac{\min_{1 \leq i < j \leq m} \delta(C_i, C_j)}{\max_{1 \leq k \leq m} \Delta_k} \quad (2.8)$$

[63]

Donde el numerador representa la distancia interclúster, es decir la distancia entre el cluster  $i$  y  $j$  y el denominador la distancia intraclúster, la cual es la distancia dentro del cluster  $k$ .

Cuanto mayor sea el índice Dunn mejor será la agrupación por lo que el número de agrupaciones que maximice el valor Dunn se toma con el óptimo número de conglomeraciones.

Uno de los inconvenientes de emplear esta métrica es que el número de agru-

paciones y dimensionalidad está relacionado al costo computacional, por lo que en caso de que exista un gran número de clústeres con alta dimensionalidad el costo computacional tenderá a aumentar significativamente.

Otro punto a tener en cuenta es que el índice Dunn es altamente influenciado por valores atípicos tanto en su numerador como en su denominador.[63]

### 2.8.3. Anomalías

Dentro de un conjunto de datos, una anomalía es sinónimo de elementos, eventos u observaciones que destacan al diferir notablemente del resto del conjunto [64], estas pueden representar problemas o eventos alarmantes que pudieran repercutir o brindar información importante del contexto en el que se trabaja.[64]; [65]

Un problema que se presenta en la detección de anomalías es la disponibilidad de un conjunto de datos etiquetados que permita trabajar el aprendizaje supervisado, por lo que, en la mayoría de los casos, resulta más viable aplicar técnicas de agrupamiento, es decir, trabajar en el aprendizaje no supervisado [66]. Dentro del aprendizaje no supervisado encontramos diferentes métodos de agrupación, la aplicación dependerá del comportamiento de los datos a emplear, en este trabajo nos apoyamos en trabajar principalmente con k-means.

#### **K-means:**

El algoritmo K-means agrupa observaciones en base a  $k$  grupos inicialmente predefinidos, es un algoritmo iterativo que busca el mover los centroides de cada grupo a la posición media de sus puntos constituyentes para posteriormente ajustar en cada iteración los elementos a sus grupos más cercanos. La función objetivo de K-means es determinada por la distancia euclidiana (norma cuadrada) entre los puntos que conforman un grupo a su respectivo centroide, como se muestra en la siguiente ecuación:

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} dist(c_i, x)^2 \quad (2.9)$$

Siendo  $x$  un punto donde  $C_i$  equivale al  $i$ -ésimo cluster y  $c_i$  al centroide del cluster  $C_i$ . [67]; [56]; [59]

#### 2.8.4. Detección de puntos anómalos

Realizar el agrupamiento de un conjunto de datos, no necesariamente implica que estos estén clasificados en datos normales y anómalos, esto debido a que en la mayoría de los casos no trabajamos con conjuntos de datos etiquetados.

Existen técnicas y algoritmos que permiten localizar los puntos aislados de un conjunto de datos, ya sea por medio de distancias, densidad, probabilidad, redes neuronales, etc. [66]. No obstante una gran desventaja que muchos de estos algoritmos presentan es la selección y ajuste de hiperparámetros, números de clústeres para modelos basados en clústeres, número de capas para modelos basados en redes neuronales, etc.

Emplear algoritmos para la detección de puntos anómalos que no requieran parámetros iniciales es un camino viable, pues de esta manera los resultados no se encontrarán afectados por errores humanos en la inicialización del algoritmo. A continuación, se mencionarán los métodos que se emplearon en este trabajo para la detección de puntos anómalos en un conjunto de datos obtenidos a través del monitoreo y sensado de un entorno habitacional.

##### **Diagrama de caja (Boxplot)**

Es una herramienta empleada en el análisis de datos, que permite visualizar la dispersión de los datos en base a los valores centrales, utiliza 5 importantes estadísticas descriptivas:

Valor medio: se requiere ordenar un conjunto de datos con el objetivo de encontrar el punto medio, al cual no referiremos como la mediana. Ésta divide los

datos en dos mitades denominados cuartiles, siendo esta misma el cuartil 2 o Q2.

Cuartil inferior, cuartil superior: también llamados Q1 Y Q3 representan el 25 % y 75 % de los datos. El conjunto de Q1 (25 %), Q2(50 %) y Q3(75 %) recibe el nombre de rango intercuartil o rango intercuartílico (IQR). Valores máximo y mínimo: estos últimos son representados a través de los denominados “bigotes”, los cuales se extienden a una distancia de  $1.5 * IQR$  siendo

$$\text{Min} = Q1 - (1.5 * IQR)$$

$$\text{Max} = Q3 + (1.5 * IQR)$$

En una distribución normal el diagrama de caja representa un 99.3 % de todos los datos, donde los valores fuera de los límites (valores mínimos y máximos) representan el 0.7 % y serán tratados como puntos anómalos.

Algo a tener en cuenta con los diagramas de caja es que a pesar de ser excelentes para mostrar si los datos son simétricos, no brindan información al respecto del tipo de simetría. Si existen variabilidades significativas de frecuencias o estamos tratando con una distribución uniforme de datos, un diagrama de caja no bastaría para brindar dicha información. [68]; [69]

Por lo anterior mencionado se recomienda emplear los diagramas de caja en combinación con otras herramientas de visualización, tal es el caso de los histogramas, esto con la finalidad de lograr mejores resultados de visualización y análisis.[70]

## COPOD

Basado en funciones de distribución acumulativa empírica (cdf), COPOD es un algoritmo de 3 etapas que toma un conjunto de entradas d-dimensionales  $X = (X1, i, X2, i, \dots, Xd, i), i = 1, \dots, n$ , y produce como salida un vector de puntuaciones anómalas de entre  $(0, \text{infinito})$  es decir :  $O(X) = [X1, \dots, Xn]$ .

Las puntuaciones representan una medida relativa de la probabilidad de anomalía de  $X_i$  al compararse con el resto del conjunto de datos, cuanto mayor sea  $O(X_i)$ , es más probable que  $X_i$  se encuentre en un contexto aislado.

El algoritmo copod permite resolver limitaciones como la dimensionalidad, sin importar si el conjunto de datos es reducido o de gran dimensionalidad, copod es capaz de manejarlo sin afectar notablemente el tiempo de desempeño.

Este algoritmo se encuentra en el primer lugar entre los 10 detectores más populares, con un puntaje de 1,5% más en ROC-AUC y 2,7% más en precisión promedio que el segundo mejor detector de rendimiento. [71]

## 2.9. Trabajos relacionados

La detección de eventos anómalos en el contexto de un entorno sensado, brinda información relevante sobre éste, propuestas como [72], abordan la detección de fallas o intrusos como eventos anómalos que pueden ser detectados y tratados incluso si en un inicio no se tiene información sobre el entorno, empleando el aprendizaje no supervisado y algoritmos de machine learning adaptativos trabajan un sistema capaz de aprender el estado de un entorno y en base a esto las anomalías que pueden presentarse. En el caso de la detección de anomalía en el espacio, una alerta de intruso es generada, posteriormente un robot autónomo responde acudiendo al lugar del evento y apoyado de una cámara procede a tratar de detectar los cambios en el entorno.

La información que puede brindar un entorno sensado no está limitada al entorno mismo, un ejemplo es [73] donde se enfocan en monitorear una habitación empleando sensores piroeléctricos, esto con la finalidad de detectar los patrones de comportamiento de un individuo a través de las acciones ejercidas por este dentro del entorno.

Apoyar el mejoramiento de la calidad de vida en adultos mayores ha impulsado propuestas dedicadas a la asistencia y cuidado de este sector a través del desarrollo de herramientas electrónicas capaces de apoyar en las actividades del diario a estas personas [74], al igual que el desarrollo de entornos que permiten el seguimiento de actividades y movimientos [75] que apoyan al cuidado y monitoreo remoto de los adultos.

El complementar la información del estado de un entorno IoT completamente habitacional y las actividades que realiza un individuo en este, genera la posibilidad de detección de patrones de comportamiento basados en las acciones ejercidas constantemente por el habitante [76], duración de estas y el espacio donde se desarrollan.

La detección de anomalías dentro de un patrón de comportamiento desarrollado puede obtenerse a partir de la comparación de un comportamiento previamente establecido como normal y el actualmente obtenido [77]. En estos términos se estaría trabajando bajo un aprendizaje supervisado.

En caso de no contar con un conjunto de datos que representen un comportamiento previamente establecido como normal, la detección de anomalías dentro de un patrón de comportamiento identificado se trabajaría mediante el aprendizaje no supervisado. En este trabajo se propone un modelo capaz de trabajar bajo un aprendizaje no supervisado, bajo el contexto de un entorno IoT aplicando Edge Computing, complementando con sensores de biodetección la información obtenida del entorno y la interacción del individuo con este.

# Capítulo 3

## Desarrollo Experimental

Con el fin de resolver la hipótesis planteada, el desarrollo de este trabajo se encuentra dividido en los aspectos psicológicos, físicos-lógicos y analíticos necesarios para su funcionamiento necesario, en las siguientes secciones se abordarán los temas necesarios para resolver las incógnitas de la creación y operatividad del sistema en conjunto.

### 3.1. Análisis general del sujeto de estudio

Para garantizar el desarrollo constante de este trabajo, se estableció una serie de puntos a evaluar que influyeron tanto en la selección del individuo como en la generación de un conocimiento base sobre la persona de estudio, tal conocimiento fue empleado en análisis posteriores.

#### 3.1.1. Requisitos para la selección del individuo

En la Tabla 3.1 se puede observar una serie de requisitos establecidos para facilitar la selección de candidatos. Estos requisitos se propusieron con la finalidad de mantener la calidad de estudio al evitar acontecimientos que pudieran afectar el desarrollo de este ya sea por el mismo individuo o elementos que influyeran en el entorno.

Tabla 3.1: Requisitos solicitados al candidato de estudio

<b>Requisitos</b>
No contar con discapacidad motora
No consumo de drogas
No consumo de alcohol
Vivienda individual
Mayoría de edad
Disponibilidad
Sin mascotas

En base a estos requisitos se seleccionó al prospecto el cuál cumplió en su mayoría los puntos solicitados, posteriormente una encuesta le fue aplicada vía internet la cuál permitió evaluar su participación en el estudio. La Tabla 3.2 presenta la información referente a la encuesta aplicada

Tabla 3.2: Encuesta de participación

**Sexo:** Masculino

**Edad:** 24 años

**Profesión:** Practicante de ingeniería en electrónica

Marque con una **X** la opción más acertada a su respuesta

<b>Usted cuenta con:</b>	<b>Si</b>	<b>No</b>	<b>Posiblemente</b>
Disponibilidad para formar parte de esta investigación	<b>X</b>		
Vivienda personal (vive solo)	<b>X</b>		
Discapacidad motora		<b>X</b>	
Mascotas	<b>X</b>		
Actividades que deba realizar fuera del hogar	<b>X</b>		
<b>Usted se considera una persona:</b>			
Abierta al cambio	<b>X</b>		
Adaptable a un entorno sensado	<b>X</b>		

### 3.1.2. Entrevista al candidato

Debido al cumplimiento de la mayoría de los puntos clave solicitados al candidato, se procede a realizar una entrevista en vivo con la intención de informar más detalladamente al individuo de los procedimientos necesarios a llevar a cabo para la etapa de experimentación en el estudio.

Una vez que el candidato ha sido informado en su totalidad de los procedimientos a seguir, los dispositivos tecnológicos a implementar en su entorno, las herramientas inteligentes que deberá portar (reloj inteligente Fitbit y celular) para su seguimiento diario y las políticas de privacidad bajo las cuales trabaja este estudio, se prosigue a la siguiente etapa correspondiente a la evaluación psicológica.

### 3.1.3. Evaluación general sobre el estado mental del candidato a estudio

La evaluación realizada en este punto tuvo como propósito el descarte de trastornos mentales graves que pudiera presentar el individuo, influyendo en los resultados finales de este trabajo, por ello una serie de pruebas psicométricas fueron aplicadas a la persona de estudio, siendo estas un gran apoyo a la predicción de tendencias de comportamiento, brindando un panorama general de las conductas que el individuo puede presentar durante el desarrollo del estudio.

Esta evaluación se dividió en 2 etapas:

#### **Aplicación de pruebas psicométricas**

En esta etapa se contó con el apoyo de un experimentado en el área de psicología que por medio de la aplicación de una serie de encuestas al candidato, reunió la información necesaria para una posterior evaluación psicométrica. Las pruebas siguientes abarcan temas relacionados a *estabilidad emocional percepción de riesgos, introspección, hábitos, temores, trastornos de ansiedad, evasión y angustia*

*social, fobias y calidad de sueño*

1. Inventario de Depresión de Beck [78] validado en población mexicana [79].
2. Versión en español del Big Five Inventory con 44 ítems [80].
3. Índice de Calidad del Sueño de Pittsburgh [81] validado en población mexicana [82].
4. Escala de Ideación Suicida de Beck [83] validada en población mexicana [84].
5. Inventario de Ansiedad de Beck [85], validado con población mexicana [86].
6. Escala afectiva positiva y negativa [87] y la versión validada en español [88]
7. Programa de entrevistas sobre trastornos de ansiedad para el DSM-IV [89]. Se aplicó la misma versión similar a [90].
8. índice de Sensibilidad a la Ansiedad - 3 [91], se implementó la versión en español de [92]
9. Cuestionario de claustrofobia [93]. Se implementó la versión en español [94]
10. Inventario de fobia a las inyecciones de sangre [95], elaborado para población española.
11. Escala de sensibilidad y propensión al repugnancia revisada [96], validada para población española por [97]
12. Escala de Evaluación de Miedo a Negativo y Escala de Evasión y Angustia Social [98]. Se implementaron las versiones de [99]
13. Se implementó el Inventario de Ansiedad Rasgo-Estado en español [100]).
14. Inventario de movilidad para la agorafobia [101]. En este estudio se implementó el instrumento en español que utilizaron [102]
15. Programa de encuestas de miedo [103]

16. El Cuestionario de cogniciones agorafobias y el Cuestionario de sensaciones corporales [104]. Se implementó la versión en español de [105] Estas pruebas abarcan los siguientes temas:

### **Evaluación y análisis de las pruebas psicométricas aplicadas**

Para la evaluación de las pruebas realizadas por el candidato, se contó con el apoyo de un psicólogo clínico experimentado, el cual después de un extenso análisis concluyó que al no contar con algún trastorno grave, ideación suicida, depresión o trastorno de ansiedad, el candidato era viable para participar en el estudio necesario para el cumplimiento de los objetivos.

A partir de los resultados concluidos por medio de la evaluación profesional psicométrica, en conjunto de los requisitos cumplidos por el candidato inicialmente y la favorable entrevista aplicada a este, su participación quedó aceptada en el desarrollo de esta investigación, concluyendo la etapa de selección.

## **3.2. Análisis de conductas relevantes a detectar**

Para el estudio de esta propuesta es necesario tener en cuenta que el individuo ejercerá numerosas conductas a lo largo del día, sin embargo, no todas aportarán información relevante para la detección de patrones, por ello es necesario un análisis previo con el fin de establecer las conductas favorables al cumplimiento de los objetivos planteados.

Trabajar sobre un análisis de conductas establecido evita saturar el entorno con dispositivos innecesarios, lo que trae como consecuencia una reducción en costos energéticos, monetarios (luz, materiales de instalación, sensores, microcontrolador entre otros) y mano de obra (programación e instalación de dispositivos).

Parte inicial del análisis de selección consiste en identificar los artículos o dispositivos que tienen una relevancia en la vida diaria de la persona, al formar parte de una constante interacción con ésta. Por ello el llenado de la siguiente encuesta Tabla 3.3 permite

complementar la información obtenida hasta el momento sobre el individuo y su entorno.

Tabla 3.3: Entrevista de entorno para participación

	<i>Cuenta con el producto</i>	<i>Tiempo aproximado que emplea en la utilización de dicho producto al día</i>							
		<i>Sí / No</i>	15 min	30 min	1 Hr	2 Hr	4 Hr	6Hr	8Hr
<i>Televisión</i>	Si								
<i>Laptop</i>	Si					<b>X</b>			
<i>Consola de videojuegos</i>	No								
<i>Escritorio</i>	No								
<i>Silla Escritorio</i>	Si						<b>X</b>		
<i>Cama</i>	SI						<b>X</b>		
		<i>Frecuencia de uso al día</i>							
		1	2	3	4	5	6	7	8
<i>Cafetera</i>		<b>X</b>							
<i>Refrigerador</i>								<b>X</b>	
<i>Alacena</i>					<b>X</b>				

¿Realiza alguna actividad que requiera mucho movimiento dentro de algún espacio de la casa (ejem: ejercicio)? No.

Para generar el conjunto de conductas a monitorear, se trabajó en colaboración con un experto en el área de psicología el cual apoyó en la selección de las siguientes conductas basándose en la información previamente obtenida sobre el individuo y su extenso conocimiento en el estudio del comportamiento humano.

- Recreación (computadora, TV)
- Desplazamiento en el entorno
- Sedentarismo
- Tomar duchas/ Bañarse
- Cocinar

- Comer
- Dormir

A partir de estas conductas se esperaba monitorear, el tiempo, duración, entorno y frecuencia en la que tendían a presentarse y con ello generar un análisis de los posibles patrones en cada una, su interacción con el resto y el contexto en el que se presentaban.

Donde cada una requirió sensores en el entorno que no sólo brindaron información sobre la persona interactuando en éste sino también del entorno mismo. A continuación, la Tabla 3.4 resume las zonas del hogar que fueron seleccionados para la implementación de dispositivos de monitoreo.

Tabla 3.4: Áreas para la implementación de dispositivos de monitoreo

<i>Áreas del hogar</i>	
Cocina	Sala
Cuarto del individuo	Baño

### 3.2.1. Espacios estratégicos para la detección de conductas seleccionadas

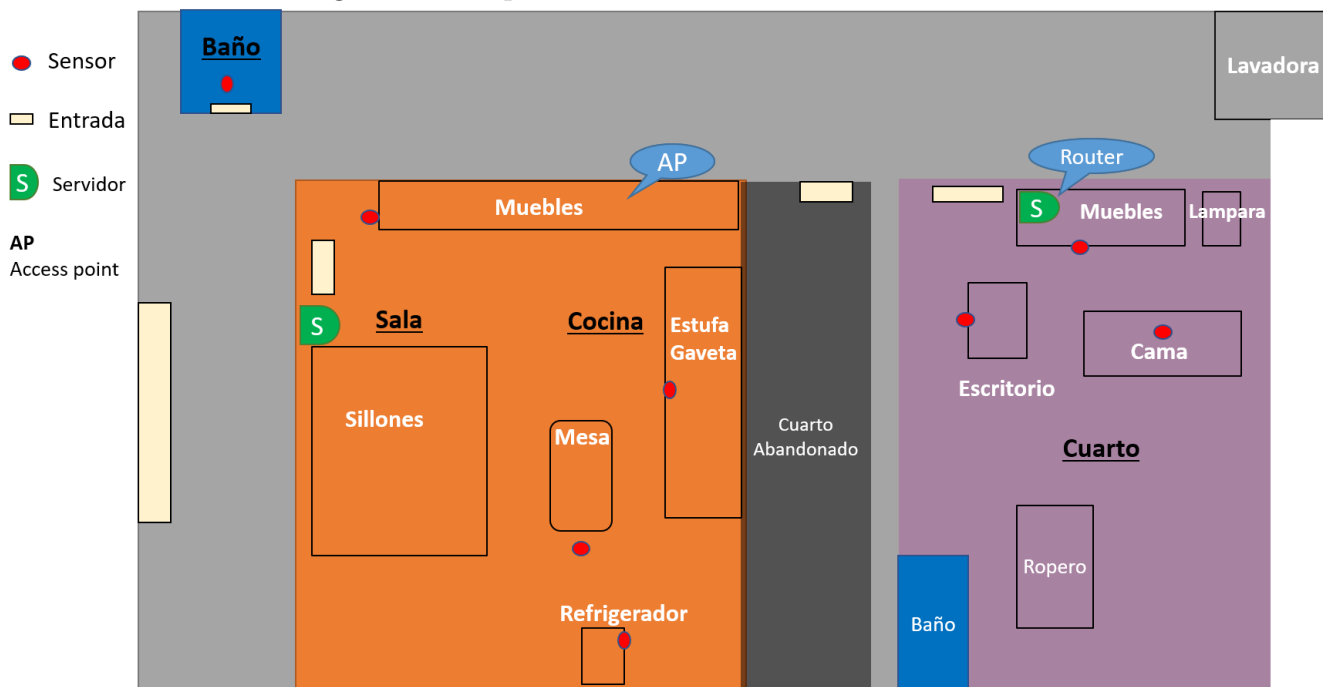
El seleccionar espacios que brinden la información necesaria de las acciones (conductas) que realiza el individuo requiere del conocimiento previo del entorno, sensores a emplear y disposición del sujeto a monitorear, todo esto con el fin de evitar inconvenientes de implementación e inconformidad del sujeto. A continuación, se detalla cada punto.

#### **Conocimiento previo del entorno**

Conocer el entorno donde se planea implementar una red IoT es indispensable para generar un análisis basado en la estructura de la vivienda, en la Fig. 3.1 se presentan las áreas del hogar accesibles para el desarrollo del estudio:

Aplicando el conocimiento del entorno, existen puntos clave respecto a éste que deben ser tomados en cuenta en el análisis para el desarrollo de la red IoT. En

Figura 3.1: Esquema de la vivienda de estudio



este estudio se evaluaron los siguientes:

### Arquitectura del espacio

Donde es necesario observar los espacios de cada área del hogar con la finalidad de visualizar la ubicación que mejor se adapte a las necesidades de cada sensor a emplear y facilite la conectividad tanto física como en red de éste. Algunos puntos para evaluar son:

#### Altura de los cuartos

Los sensores tienen un rango de alcance de detección, al trabajar con sensores piroeléctricos o ultrasónicos, se deben contemplar las características de alcance de cada uno de acuerdo con su funcionamiento, en caso contrario se pueden presentar fallas relacionadas a la detección del individuo, algunas de ellas pueden ser:

- Detección de sujetos no deseados (gatos, perros, pájaros entre

otros.)

- Nula detección del individuo deseado.
- Detección del individuo en áreas no establecidas en un inicio.
- Falsas y acertadas detecciones del individuo sin control alguno.

### **Obstáculos alrededor del entorno**

Existen muchos obstáculos que pueden apoyar o perjudicar el uso óptimo de los dispositivos de monitoreo, un ejemplo claro son los muebles, los cuales juegan un papel beneficioso si se emplean como apoyo para sostener o adaptar los dispositivos, no obstante, debido a su tamaño y ubicación también pueden llegar a generar problemas de conectividad o funcionamiento.

### **Espacios indeseados**

La arquitectura de cada hogar varía, por ende, es necesario conocer si existen espacios que puedan llegar a generar problemas de conectividad entre los dispositivos y el servidor, tal es el caso de las divisiones entre habitaciones, las cuales pueden abarcar una gran separación entre cuartos, influyendo en los tiempos de transmisión de datos por parte de la red IoT.

### **Fuentes de energía**

Una vez localizados los espacios donde queremos situar nuestros sensores, es necesario tener en cuenta que para su funcionamiento estos requieren una fuente constante de alimentación eléctrica, por ello las conexiones de luz deben estar disponibles para su uso y de preferencia lo más próximas posibles a los dispositivos.

Todo esto con la finalidad de evitar el uso de extenso de cableado que tiende a invadir el espacio del individuo afectando su percepción de éste y creando situaciones de riesgo para su traslado e interacción con el entorno.

### **Ubicación de módem**

El módem al ser el puente entre servidor-nube debe permanecer lo más accesible posible, lo que implica cuidar su ubicación y permanente funcionalidad. Esto último hace referencia al servicio de internet el cual dependerá totalmente de la compañía en contrato y por ellos se deben tomar las medidas necesarias para evitar un bajo rendimiento en los servicios.

### **Ubicación de servidor**



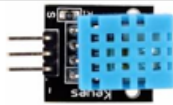




El servidor se encuentra en constante comunicación con más de un dispositivo a la vez y los tiempos en los que puede ser accesado dependen de los eventos generados en el entorno, la importancia de su ubicación radica en mantener una comunicación rápida con los dispositivos, esto con la finalidad de evitar ruido en la información recibida o la pérdida total de esta. La aplicación del paradigma de EC fue indispensable para el funcionamiento en tiempo y forma de la red al permitir la ejecución de tareas dentro del mismo entorno sin necesidad de mantener una constante comunicación a servicios externos que pudieran ocasionar retrasos en conexión y comunicación entre los dispositivos en red.

### 3.3. Sensores a emplear en la vivienda

El repertorio de sensores abarcado en el punto 2.5.3.3 “Sensores empleados en la arquitectura IoT” nos brinda las opciones que fueron implementadas en la vivienda de estudio, no obstante, en este capítulo complementaremos la información brindada con las características que llevaron a la selección y ubicación del conjunto de sensores trabajados.

Además, se espera abordar el tema del acoplamiento empleado entre sensores para su funcionamiento esperado. En la Tabla 3.5 se muestran los modelos específicos de los sensores empleados en la vivienda, así como algunas de sus características contempladas para su uso.

Tabla 3.5: Sensores empleados en la arquitectura IoT

<i>PIR HC-SR501</i>	<i>Ultrasónico HC-SR04</i>	<i>Modulo DHT11</i>
		
<i>Interruptor (switch) magnético</i>	<i>LDR</i>	<i>Medidor de AC Sct-013 Sct013</i>
		
<i>Sensor flex (casero)</i>		
		

Cabe mencionar que los sensores están sujetos a parámetros de funcionamiento establecidos por su fabricante, por ello es necesario realizar la consulta a la hoja de datos del respectivo componente con el fin de evitar problemas en la alimentación energética, manipulación del componente, contraste entre salidas esperadas-obtenidas.

### 3.3.1. Características por contemplar en sensores

- Costo
- Precisión
- Accesible
- Sensibilidad
- Rango de alcance
- Discreción en tamaño
- Facilidad de implementación
- Componentes adicionales para el funcionamiento requerido
- Variables del entorno que influyen en su correcto funcionamiento.

### 3.3.2. Selección de dispositivos en base a la actividad a monitorear

Cada dispositivo estará conformado por un sensor(es) que tendrá una conducta a detectar, por ello su funcionamiento tanto en programación como física variará al resto, por lo que su ubicación debe contemplar que su activación ya sea constante o poco recurrente tendrá un impacto en los datos que generará basados en su algoritmo de funcionamiento, la Tabla 3.6 presenta las variables involucradas para la elaboración e instalación de cada dispositivo en el entorno.

### 3.4. CONTEXTO PARA EMPLEAR CADA MODO DE OPERACIÓN DEL $\mu C$ ESP8266 DENTRO DEL ENTORNO IoT

Tabla 3.6: Detalles de la implementación de sensores en entorno IoT

<i>Área</i>	<i>Sensores x dispositivo</i>	<i>Ubicación</i>	<i>Actividad a detectar</i>	<i>Conducta</i>
<b>Sala</b>	PIR + LDR	Entrada	Movimiento	Desplazamiento en el entorno.
			0/1 Interruptor	Sedentarismo
<b>Cocina</b>	PIR + LDR	Parte superior de mesa	Movimiento	Comer
			0/1 Interruptor	Sedentarismo
			Ultrasónico + Interruptor	Estufa Gaveta
	Interruptor	Refrigerador	Abrir/ Cerrar	Cocinar Comer
<b>Dormitorio</b>	PIR + LDR	Entrada	Movimiento	Desplazamiento en el entorno.
			0/1 Interruptor	Sedentarismo
			Medidor AC	Cargador de Laptop
	Sensor Flex	Cama	Movimiento en cama	Dormir Descanso
<b>Baño</b>	PIR + Modulo DHT11	Entrada	Movimiento Uso de regadera	Tomar ducha

### 3.4. Contexto para emplear cada modo de operación del $\mu C$ ESP8266 dentro del entorno IoT

Hemos mencionado la relevancia de tomar en cuenta el consumo de energía por parte de los dispositivos aplicados en entornos IoT, por ello en este trabajo se ha optado por emplear la placa “Lolin wemos D1” basada en el  $\mu C$  ESP8266. Es posible obtener una mayor flexibilidad en el consumo de energía debido a que el  $\mu C$  permite trabajar bajo 3 modos de operación:

### **Modem Sleep**

Al mantenerse la placa conectada a una fuente de alimentación constante, el funcionamiento de esta no se ve afectado por un consumo energético limitado por lo que este modo de operación permite mantener múltiples funcionalidades del  $\mu C$  habilitadas sin generar problemas en su funcionamiento.

Debido a que la fuente de alimentación de la placa se mantiene constante, la selección de los sensores a implementar no estará condicionada a la alimentación constante requerida o la frecuencia de activación necesaria de estos sino a las fuentes de alimentación disponibles en el  $\mu C$

### **Light Sleep**

El modo de sueño ligero (Light Sleep) funciona si es necesario mantener una conexión con el router y recibir información de éste en tiempo real. El  $\mu C$  se encuentra en un estado suspendido, siendo una interrupción externa vía GPIO la única manera de reactivar el funcionamiento necesario, lo que resulta en una gran ventaja en la implementación de sensores que complementados con un arreglo específico brinden un funcionamiento capaz de realizar dicha acción al existir un cambio en el entorno.

Para el contexto de este trabajo la selección de estos sensores incluye el análisis de la sección “Limitaciones en cada modo de operación implementado”, buscando aquellos que no requieran un consumo de alimentación superior a los 3.3v.

### **Deep Sleep**

Es el modo de operación que permite el mayor ahorro de energía por parte del dispositivo, para esta propuesta experimental se trabajaron sus dos diferentes configuraciones, las cuales consisten en despertar el  $\mu C$  empleando una interrupción externa:

1. Vía GPIO, donde la conexión necesaria es un puente entre los pines D0 y

### 3.4. CONTEXTO PARA EMPLEAR CADA MODO DE OPERACIÓN DEL $\mu C$ ESP8266 DENTRO D

RESET. Aplicando esta configuración se puede despertar al  $\mu C$  repetidamente cada cierto lapso de tiempo establecido.

2. Al poder activar e inicializar el chip mediante un pulso de bajo nivel, la configuración planteada entre D0-RESET puede variar realizando la conexión:

D0- Salida de un sensor, donde a partir del acoplamiento entre el sensor y un arreglo específico el pulso de bajo nivel puede ser generado

Para mayor información consultar Tabla 1.4 “Modos de operación del  $\mu C$  ESP8266”.

Los dispositivos que conforman la red IoT basan su funcionamiento en la interacción entre sensores y la placa wemos, por ello es necesario que trabajen bajo uno de los tres modos de operación. Elegir el modo de operación bajo el cual funcionarían los dispositivos requirió un extenso análisis que contemplo los siguientes puntos:

- Funcionamiento de cada modo de operación
- Limitaciones en software/ hardware de cada modo de operación
- Consumo energético requerido por cada sensor para su funcionamiento óptimo.
- Componentes necesarios para el funcionamiento de cada sensor en la placa wemos bajo los tres diferentes modos de operación.
- Costo de los componentes necesarios para el funcionamiento de cada sensor en la placa wemos bajo los tres diferentes modos de operación.
- Funcionamiento esperado de cada dispositivo en base a la conducta a detectar.
- Arquitectura del entorno de vivienda
- Arquitectura de la red IoT

### **3.5. Funcionamiento lógico general de los dispositivos empleados en la red IoT**

Cada dispositivo varia su funcionamiento dependiendo el modo de operación programado, no obstante, se espera que el procedimiento ideal en general sea:

1. Detectar el evento asignado
2. Establecer conexión a red
3. Establecer conexión a servidor
4. Enviar información al servidor
5. Guardar el evento en la BD correspondiente
6. Repetir procedimiento

#### **3.5.1. Contemplar las limitaciones en cada modo operacional implementado**

Trabajar bajo un modo de operación implica conocer las limitaciones en software/hardware por parte de la placa wemos empleada, entre las que requirieron un mayor enfoque para los fines de este trabajo se presentan a continuación:

- \* Alimentación necesaria para un óptimo funcionamiento en cada modo de operación.
- \* Consumo en cada etapa de funcionamiento (reposo, activación, conexión a red entre otros) por parte de la placa wemos.
- \* Tiempo de respuesta requerida ante la detección de un evento.
- \* Duración de cada conducta (evento) a detectar.
- \* Frecuencia de activación requerida en base a la conducta a monitorear

### 3.5. FUNCIONAMIENTO LÓGICO GENERAL DE LOS DISPOSITIVOS EMPLEADOS EN LA RED

- \* Área de la vivienda en la que el dispositivo estará en funcionamiento
- \* Consumo de energía requerido por parte de cada sensor para su óptimo funcionamiento
- \* Flexibilidad de funcionamiento por parte de cada sensor

#### 3.5.2. Análisis de la relación sensores-modos de operación

Teniendo en cuenta las limitaciones de cada modo de operación disponible para el ahorro de energía, el contexto en el que se recomienda su implementación y la selección de los sensores para la detección de conductas a monitorear; se procede a la asignación de modos de operación bajo los que desempeñará su tarea cada dispositivo. Para ello se requirió un análisis del funcionamiento, características y parámetros de activación de cada sensor que permitió resumir los resultados en la Tabla 3.7 presentada a continuación:

Tabla 3.7: Características a considerar de los sensores empleados

<i>Sensores</i>	<i>Consumo de energía</i>	<i>Facilidad de implementación</i>	<i>Sensibilidad a ruidos</i>	<i>Discreción</i>
<i>HC sr501</i>	Alto	Media	Media	Alta
<i>HC-SR04</i>	Alto	Media	Media	Media
<i>Modulo DHT11</i>	Alto	Media	Media	Alta
<i>Interruptor magnético</i>	Bajo	Baja	Baja	Alta
<i>LDR</i>	Bajo	Baja	Media	Alta
<i>Corriente no *Flex (casero)</i>	Nulo	Alta	Media	Media

Conociendo el consumo de energía aproximado de cada sensor, la selección del método de ahorro de energía a emplear se hace más evidente al conocer las limitaciones de cada uno.

A continuación, en la Tabla 3.8 se presenta una propuesta sobre la elección de sensores y las condiciones de implementación a considerar en caso de elegir un determinado modo de ahorro de energía dentro del contexto establecido este trabajo.

Tabla 3.8: Relación dispositivos - modo de operación

<i>Modo de operación</i>	<i>Condiciones de implementación</i>	<i>Sensores con los que se recomienda trabajar</i>
<b><i>Deep Sleep</i></b> <i>(Sueño Profundo)</i>	<ul style="list-style-type: none"> <li>Solo puede estar compuesto por: <u>Placa wemos + 1 sensor</u> (y los componentes necesarios)</li> <li>El almacenamiento de datos emplea la memoria RTC del <math>\mu C</math></li> </ul>	Fotorresistencia Switch PIR
<b><i>Light Sleep</i></b> <i>(Sueño ligero)</i>	<ul style="list-style-type: none"> <li>Solo puede estar compuesto por: <u>Placa wemos + 1 sensor</u> (y los componentes necesarios)</li> <li>El almacenamiento de datos emplea la memoria EEPROM</li> </ul>	Fotorresistencia Switch PIR
<b><i>Modem Sleep</i></b>	El almacenamiento de datos emplea la memoria EEPROM	Todos los mencionados en la Tabla 3.5 “Sensores empleados en la arquitectura IoT”

Basado en esta propuesta se dio inicio al desarrollo de los programas que permitieran seguir el diseño planteado empezando por el funcionamiento del  $\mu C$  bajo el modo Deep Sleep

### 3.6. Propuesta experimental

Bajo el análisis de estos puntos mencionados y el funcionamiento general determinado se propuso una implementación experimental que tenía como objetivo desarrollar un ambiente IoT con el menor cableado posible dentro del entorno sin afectar el cumplimiento de las tareas asignadas a cada dispositivo de monitoreo, esto gracias a la aplicación de los modos de operación disponibles en el  $\mu C$  para el ahorro de energía, permitiendo de esta manera el funcionamiento de los aparatos a través de baterías recargables.

Con el fin de lograr el cumplimiento del objetivo planteado, el desarrollo de la propuesta se basó en un extenso análisis que trabajó lo siguiente:

### 3.6.1. Arquitectura de red IoT

Una vez elegidas las conductas a detectar, los dispositivos a emplear y los espacios a monitorear se procedió a trabajar la arquitectura en red, donde se buscó que todos los módulos se encontraran en comunicación al servidor y este al router, aplicando el paradigma de EC para así poder mantener una conexión más estable y rápida al no depender de servicios externos para el procesamiento de los datos obtenidos por parte del los dispositivos.

#### 3.6.1.1. Instalación de Servidor

La instalación del servidor a emplear en la red IoT consistió en el uso del dispositivo Raspberry pi 3 modelo b v1.2, al cual se le instalo inicialmente el sistema operativo Raspberry Pi OS (Raspbian) seguido de la paquetería MySQL, el servidor Apache y PHP.

La conexión se realizó a la red principal de la vivienda vía inalámbrica y finalmente se le otorgó un identificador IP correspondiente a 192.168.0.30, consultar la Fig. 3.2 para visualizar parte de la instalación de la paqueteria MySQL y la Fig. 3.3 para visualizar la información entre la conexión del servidor 1 a la red 1

Figura 3.2: Instalación de la paqueteria MySQL

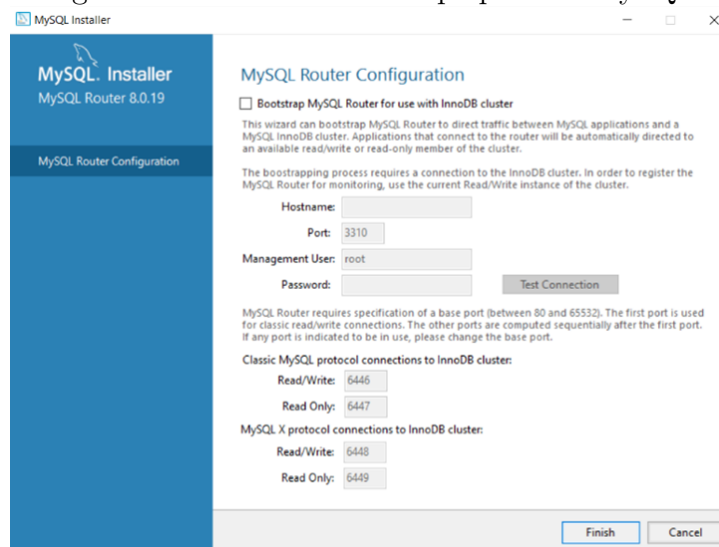
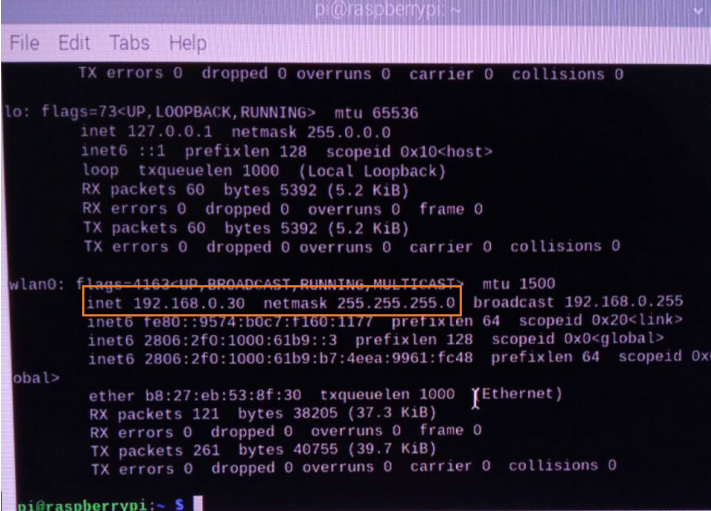


Figura 3.3: Información relacionada al servidor conectado a la “red 1”



```

pi@raspberrypi ~
File Edit Tabs Help

TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
  inet 127.0.0.1 netmask 255.0.0.0
  inet6 ::1 prefixlen 128 scopeid 0x10<host>
  loop txqueuelen 1000 (Local Loopback)
  RX packets 60 bytes 5392 (5.2 KiB)
  RX errors 0 dropped 0 overruns 0 frame 0
  TX packets 60 bytes 5392 (5.2 KiB)
  TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

wlan0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
  inet 192.168.0.30 netmask 255.255.255.0 broadcast 192.168.0.255
  inet6 fe80::9574:b0c7:f160:1177 prefixlen 64 scopeid 0x20<link>
  inet6 2806:2f0:1000:61b9::3 prefixlen 128 scopeid 0x0<global>
  inet6 2806:2f0:1000:61b9:b7:4eea:9961:fc48 prefixlen 64 scopeid 0x0
  global>
  ether b8:27:eb:53:8f:30 txqueuelen 1000 (Ethernet)
  RX packets 121 bytes 38205 (37.3 KiB)
  RX errors 0 dropped 0 overruns 0 frame 0
  TX packets 261 bytes 40755 (39.7 KiB)
  TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

pi@raspberrypi ~ $

```

### 3.6.1.2. Creación de la base de datos

Para el diseño y creación inicial de la base de datos (BD) fue necesario compactar los conjuntos de dispositivos más cercanos entre sí dentro de la red IoT, con la finalidad de seccionar en tablas la información obtenida de cada uno.

Finalmente con el fin de no crear un gran conjunto de bases de datos, se diseñaron dos agrupaciones, donde a la primera se le asignó el nombre “Sala” y contenía las tablas correspondientes a las áreas baño, cocina, sala, la segunda fue nombrada y “Cuarto” almacenaba las tablas relacionadas a los elementos localizados en la recámara del individuo, visualizar la Fig. 3.4. para una visualización parcial a la BD general

Cada una de las dos bases de datos contaba con un conjunto de tablas que se encargaban de almacenar la información registrada por los dispositivos implementados a lo largo del entorno contando con un diseño general que involucran los siguientes atributos:

**Nsensor.** Funciona como contador de registros el cual aumenta conforme un nuevo evento es almacenado.

**Estado.** Esta variable corresponde al evento numérico capturado, si el valor es 0 representa a un evento finalizado o “apagado” y en caso de que un evento inicie o sea activado el valor será igual a 1.

**Time.** Es el registro en fecha y hora del evento capturado con un formato de YY-MM-DD HH:MM:SS este valor será en base al horario configurado del servidor.

**E\_Hrs, E\_min, E\_seg.** En el caso de que un evento sea capturado por algún dispositivo y no pueda ser enviado al servidor o almacenado en la base de datos por cuestiones técnicas, la información será resguardada en la memoria del microcontrolador y una vez que los datos puedan transmitirse se registrará en la base de datos la hora en la que el evento originalmente fue detectado a través de los campos E\_Hrs(EEPROM\_Hora),E\_min(EEPROM\_minutos),E\_seg(EEPROM\_segundos).

**idSensor.** En algunos casos un dispositivo se integró con 2 tipos de sensores diferentes por lo que fue necesario asignar un id a cada uno con la finalidad de identificar los registros de cada elemento en su respectiva tabla.

Figura 3.4: Base de datos general con las tablas

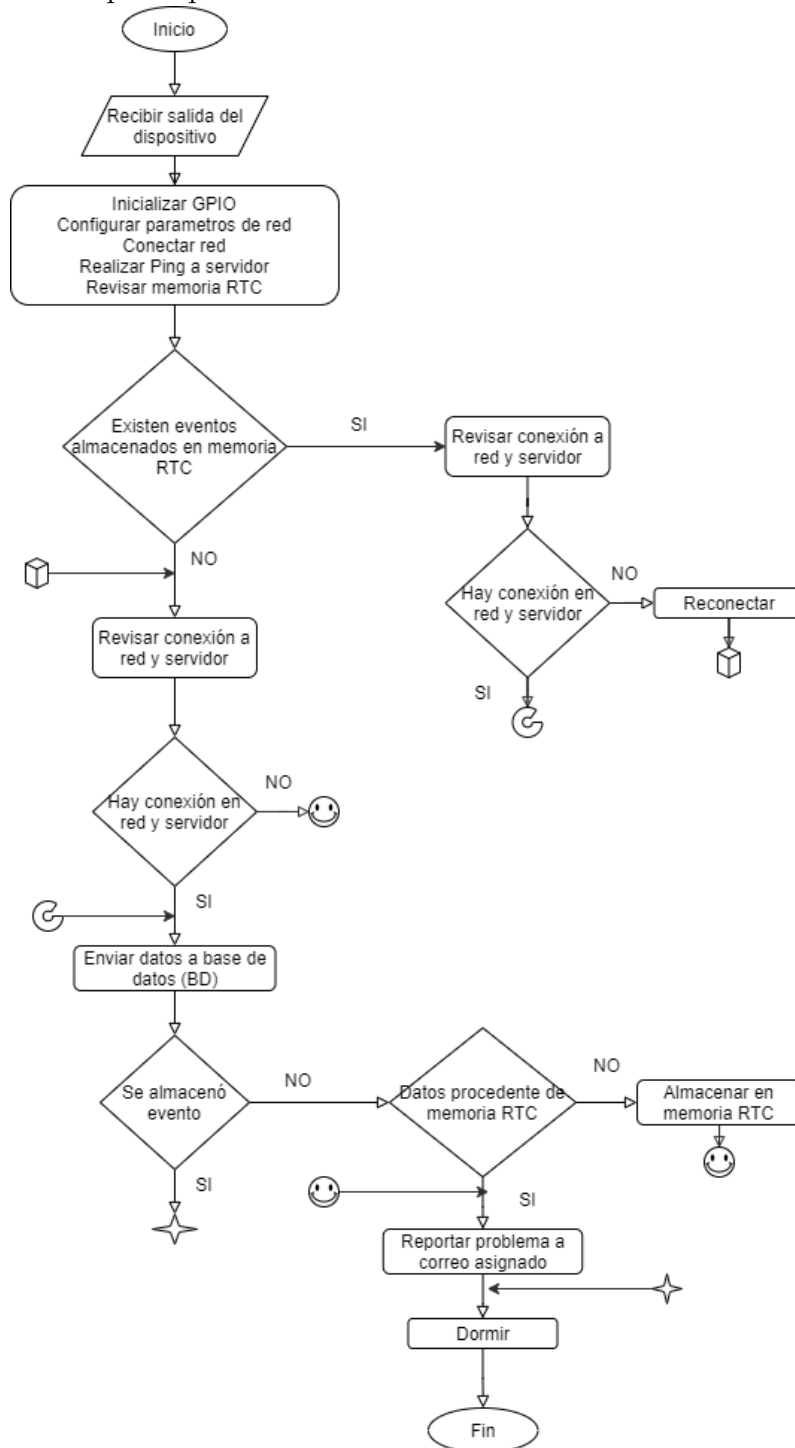
ID	Estado	Time	E_Hrs	E_min	E_seg	idSensor
728	1	2020-10-26 20:05:05	0	0	0	1
729	0	2020-10-26 20:05:40	0	0	0	1
730	1	2020-10-26 20:06:57	0	0	0	1
731	0	2020-10-26 20:07:54	0	0	0	1
732	1	2020-10-26 20:08:24	0	0	0	1
733	1	2020-10-26 20:08:52	0	0	0	0
734	0	2020-10-26 20:09:25	0	0	0	1
735	1	2020-10-27 06:49:04	0	0	0	1
736	0	2020-10-27 06:49:38	0	0	0	0
737	0	2020-10-27 06:50:19	0	0	0	1
738	1	2020-10-27 07:06:23	0	0	0	1

### 3.6.2. Programación de dispositivos para un bajo consumo

A pesar de contar con un esquema de funcionamiento ideal por parte de los dispositivos de monitoreo (subsección 3.5 “Funcionamiento lógico general de los dispositivos empleados

en la red IoT”), al existir limitaciones en cada modo de operación por parte del  $\mu C$ , es necesario adaptar el desarrollo de software a las necesidades que requiere cada uno sin afectar los objetivos esperados. La Fig. 3.5 presenta el diagrama para el funcionamiento de los dispositivos que contaban con la configuración deep sleep y cuyo único sensor incorporado fue el de movimiento (PIR) para la detección de presencia en un área fija. La importancia de su estructura radicó en utilizar un sensor que a través de la circuitería requerida fuera capaz de responder a los cambios del entorno aún dentro del deep sleep sin consumir abruptamente el suministro de energía, también es necesario señalar el uso de la memoria RTC para el almacenamiento de los eventos bajo las condiciones establecidas.

Figura 3.5: Diagrama del funcionamiento del dispositivo para la detección de movimiento en modo de ahorro Deep Sleep



### 3.6.3. Desarrollo físico de dispositivos para un funcionamiento de bajo consumo

El funcionamiento de los programas desarrollados dio paso al análisis del hardware requerido en el completo funcionamiento de los dispositivos programados para un bajo consumo de energía. En estos su fuente de alimentación estaría conformada por:

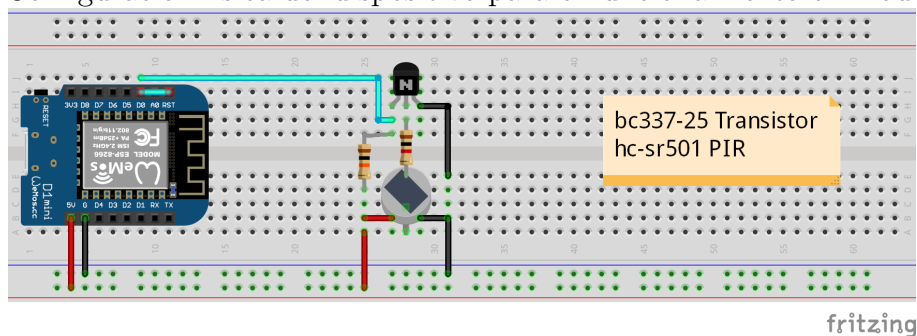
- **Batería Lipo 3.7v — 400mah**
- **Battery shield para wemos d1:** tarjeta de expansión de la placa wemos configurada para recibir el voltaje proveniente de una batería.

Debido a que este modo de operación es el que mas pretende alargar el funcionamiento del dispositivo, se debe buscar el menor consumo por parte de los sensores acoplados a la placa, es decir sólo es recomendada la implementación de un solo sensor por dispositivo.

#### 3.6.3.1. PIR (detección de movimiento, presencia y desplazamiento del individuo)

Se desarrollo la configuración electrónica necesaria para el funcionamiento de este sensor al implementarse en la placa wemos bajo el modo de ahorro Deep sleep. La Fig. 3.6 muestra el esquemático final de este acoplamiento

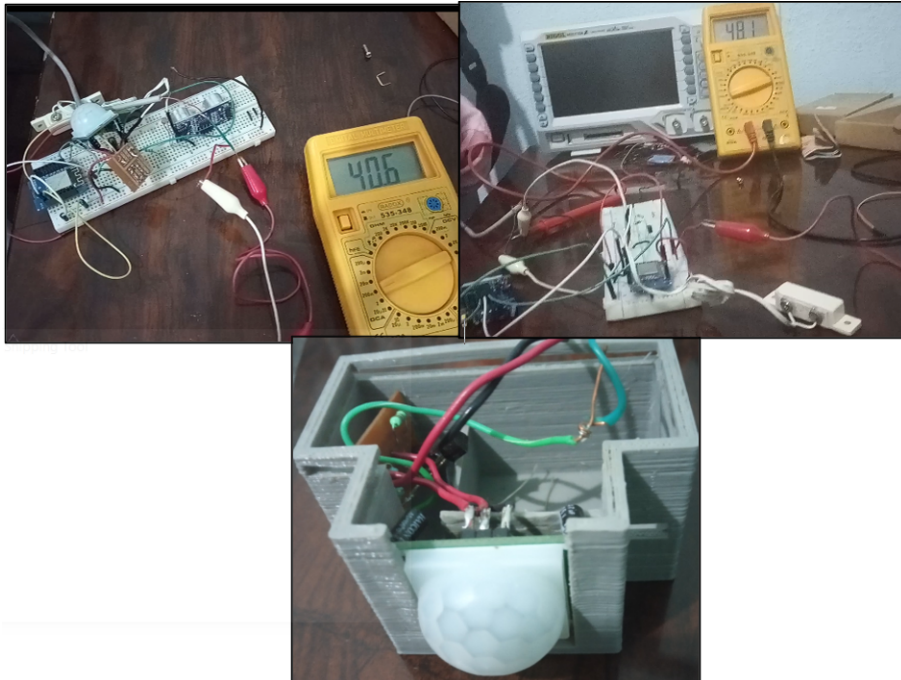
Figura 3.6: Configuración física del dispositivo para el funcionamiento en modo deep sleep



### Pruebas de funcionamiento

A continuación la Fig. 3.7 presenta un conjunto de imágenes relacionadas al funcionamiento del dispositivo conformado por el sensor PIR y  $\mu C$  en modo deep sleep, pruebas de su funcionamiento, medición e implementación con su respectiva protección

Figura 3.7: Pruebas de funcionamiento del dispositivo en modo deep sleep



### Resultado de la implementación PIR en modo Deep Sleep

El rendimiento de la batería logró extenderse hasta un día con un registro aproximado de 15 eventos transcurridos en el entorno.

### 3.6.4. Resultados Finales de la propuesta experimental para el desarrollo de la red IoT

Una vez desarrollado e implementado el diseño experimental se registraron los resultados obtenidos con la finalidad de evaluar el desempeño del sistema en red, su comunica-

ción, funcionalidad lógica y la eficiencia de su instalación física.

#### 3.6.4.1. Propuesta experimental de Red

Durante el periodo de pruebas de la implementación en red empleando

*Dispositivos — Servidor Edge — Router*

**Se concluyó lo siguiente:**

- Los tiempos de latencia entre la comunicación dispositivos-servidor eran altos.
- Los dispositivos ubicados en los extremos de las áreas (Baño, Entrada, Cocina) presentaban mayor dificultad al realizar la conexión hacia el servidor, afectando el registro de los eventos detectados.
- Desconexión constante por parte de los dispositivos a la red.
- Debido a las constantes peticiones por diversos sensores al mismo tiempo, el registro de eventos en la base de datos variaba a los tiempos reales.
- La velocidad en los servicios de internet presentaba tan bajo rendimiento que el sujeto de estudio no podía desempeñar sus actividades recreativas en la misma manera.
- El monitoreo de los dispositivos solamente era posible a través de:
  1. Las alertas generadas por estos al presentarse algún problema en la conexión hacia el servidor.
  2. Revisión presencial
- El manejo de los servidores (consulta, comandos, respaldos) únicamente era posible a través de manipulación física presencial, lo que afectaba constantemente el funcionamiento de la red y además se prestaba para la corrupción de datos obtenidos por los dispositivos.

#### **3.6.4.2. Implementación lógica-física**

La implementación lógica de los dispositivos fue desarrollada y aplicada exitosamente, no obstante, dentro del contexto planteado el rendimiento de la batería provocó la interrupción del funcionamiento de los dispositivos debido a múltiples activaciones en lapsos prolongados, limitando de esta manera la obtención de datos y las limitaciones en tiempo y recursos no permitieron concluir el desarrollo físico de este enfoque experimental.

### **3.7. Propuesta final para el funcionamiento de la red IoT**

Los resultados de la implementación anterior encaminaron el proyecto a una propuesta final enfocada en contemplar los tiempos y condiciones más viables de acuerdo a los recursos disponibles orientando de esta manera el desarrollo de dispositivos adaptables a entornos habitacionales para un alcance del sector socioeconómico del país.

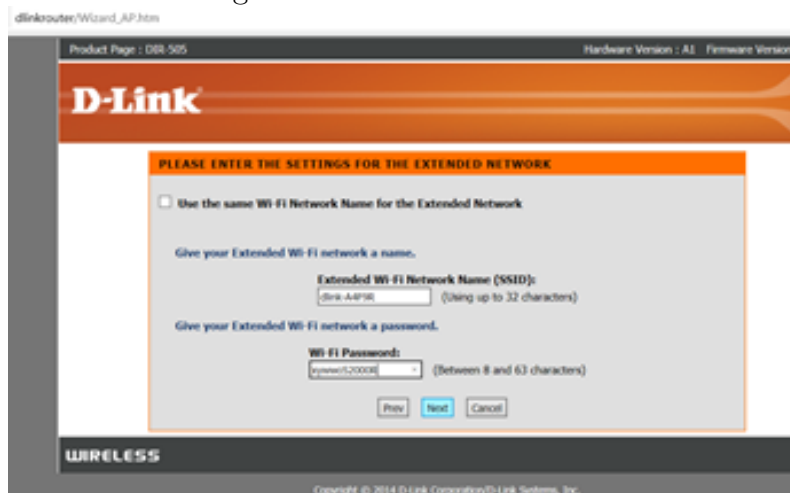
#### **3.7.1. Desarrollo de arquitectura final para la comunicación y almacenamiento de la IoT**

A partir de las conclusiones generadas en el periodo de prueba por parte del funcionamiento de la red se realizó un análisis con el propósito de obtener menos pérdidas en el registro de eventos al mantener una conexión más estable y con un mayor alcance para el beneficio de la comunicación entre los dispositivos que integraban la red IoT, la infraestructura final contó con la ejecución de las siguientes acciones:

1. Extender la cobertura Wi-Fi por medio de un punto de acceso (AP) conectado a la red LAN cableada Fig. 3.8. De esta manera se interconectan los equipos de comunicación inalámbrica construyendo una “red 2” que permanece accesible a los dispositivos

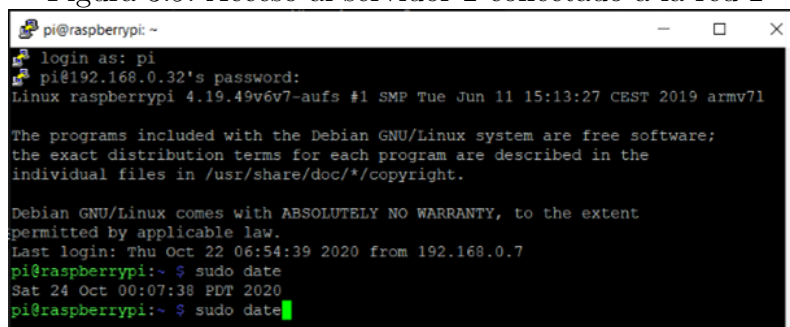
localizados en las áreas de Baño, Cocina y Sala

Figura 3.8: Instalación de AP



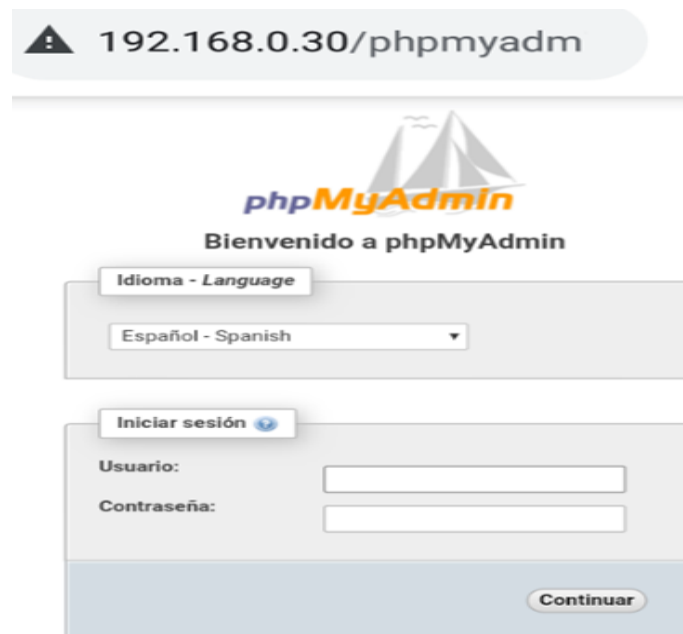
2. Aprovechando la red creada a partir del AP se implementó otro servidor (servidor 2) con identificador IP 192.168.0.32, el cual fue conectado vía inalámbrica a la “red 2” Fig. 3.9 y de esta manera se lograron disminuir los tiempos de latencia en la conexión de dispositivos al servidor y por ende a la red.

Figura 3.9: Acceso al servidor 2 conectado a la red 2



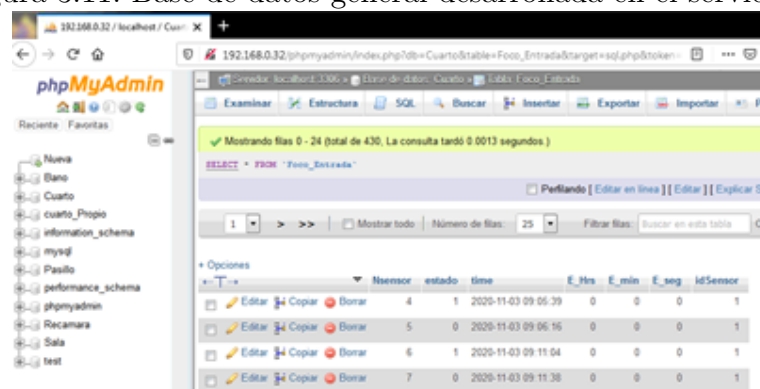
3. Mantener la comunicación de dispositivos en el área “Dormitorio” únicamente hacia la “red 1” Fig.3.10, la cual es creada a partir del router-módem y donde también se encuentra un servidor conectado (Servidor 1) vía inalámbrica para el funcionamiento de esta sección, con un identificador IP de 192.168.0.30

Figura 3.10: Herramienta phpMyAdmin instalada en el servidor 1 para la administración de MySQL



4. Rediseñar las bases de datos correspondientes a cada servidor, creando finalmente en el servidor 1 perteneciente a la “red 1” una base de datos llamada “Sala”, la cual se compone de un conjunto de 5 Tablas relacionadas a los dispositivos localizados en las áreas de *Baño*, *Cocina* y *Sala*, lo anterior mencionado puede observarse en la Fig. 3.11 presentada a continuación.

Figura 3.11: Base de datos general desarrollada en el servidor 1



Y en el caso del servidor 2 conectado a la “red 2” una base de datos llamada “Cuarto”

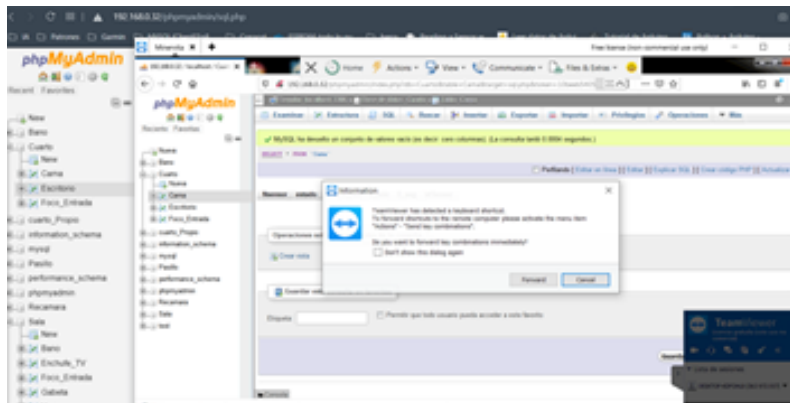
la cual almacena únicamente los valores procedentes de los elementos ubicados en la habitación del individuo Fig 3.12.

Figura 3.12: Base de datos general desarrollada en el servidor 2



5. Con el fin de no afectar el funcionamiento ininterrumpido de la red y a su vez realizar los procesos necesarios de manipulación y monitoreo en los dispositivos y servidores, se instaló un equipo de cómputo equipado con un software capaz de permitir la comunicación vía remota con las redes IoT de la vivienda, consultar Fig.3.13.

Figura 3.13: Monitro a distancia de la BD a través de TeamViewer



6. El funcionamiento de los dispositivos de monitoreo en red se desarrolla en base al paradigma “Edge Computing”, de esta manera se reducirían los tiempos de latencia entre la comunicación de dispositivos a los centros de red (servidores) por ello una gran parte del procesamiento de revisión, validación, monitoreo de entorno, envío de

alertas y almacenamiento es realizado por los dispositivos que forman parte de la red misma.

En la Fig. 3.1 “Esquema de la vivienda de estudio ” se presenta la estructura final de red implementada en el entorno, donde los elementos que la conforman se presentan más detalladamente a continuación en la Tabla 3.9.

Tabla 3.9: Elementos en conexión IoT

<i>Área de ubicación</i>	<i>Elementos</i>	<i>Modelo</i>	<i>Conexión a red:</i>
<i>Consultar tabla 3.6</i>	13 sensores que conforman un total de 8 dispositivos de monitoreo	Consultar tabla 3.5	Revisar puntos 1,2 y 3 de la sección “Implementación final”
<i>Dormitorio</i>	<i>Router-modem</i>	Restringido	Red 1
	<i>Servidor 1</i>	Raspberry pi 3 model b v1.2	
<i>Cocina</i>	<i>Punto de acceso (AP)</i>	D-link DIR 505	Red 2
<i>Sala</i>	<i>Servidor 2</i>	Raspberry pi 3 model b v1.2	

### 3.7.2. Selección del modo operacional a trabajar en el desarrollo de los dispositivos

En la Tabla 3.8 se muestra la relación entre los dispositivos y los modo de operación disponibles donde la comparación beneficia el uso del modo operacional “*Modem Sleep*” en la implementación de todos los sensores propuestos al permitir el funcionamiento constante de estos sin interrupciones derivadas de un agotamiento en la fuente de alimentación.

En base a lo mencionado y los resultados de la propuesta experimental previa se eligió trabajar esta configuración para el desarrollo de todos los dispositivos que conformarían la red IoT, una vez establecido el modo operacional de los dispositivos se procedió a desarrollar

el programa específico para el funcionamiento de cada uno, teniendo en cuenta que los sensores que los conformarían serían de tipo analógico y digital.

### **3.7.3. Desarrollo lógico final de los dispositivos empleados en la red IoT**

El desarrollo lógico se dividió en dos partes *Funcionamiento general* y *Funcionamiento específico*, todo esto basandose en las funciones necesarias para la ejecución en tiempo y forma de todos los dispositivos, así como las funciones específicas para cada uno de acuerdo a su configuración, limitaciones y tarea asignada.

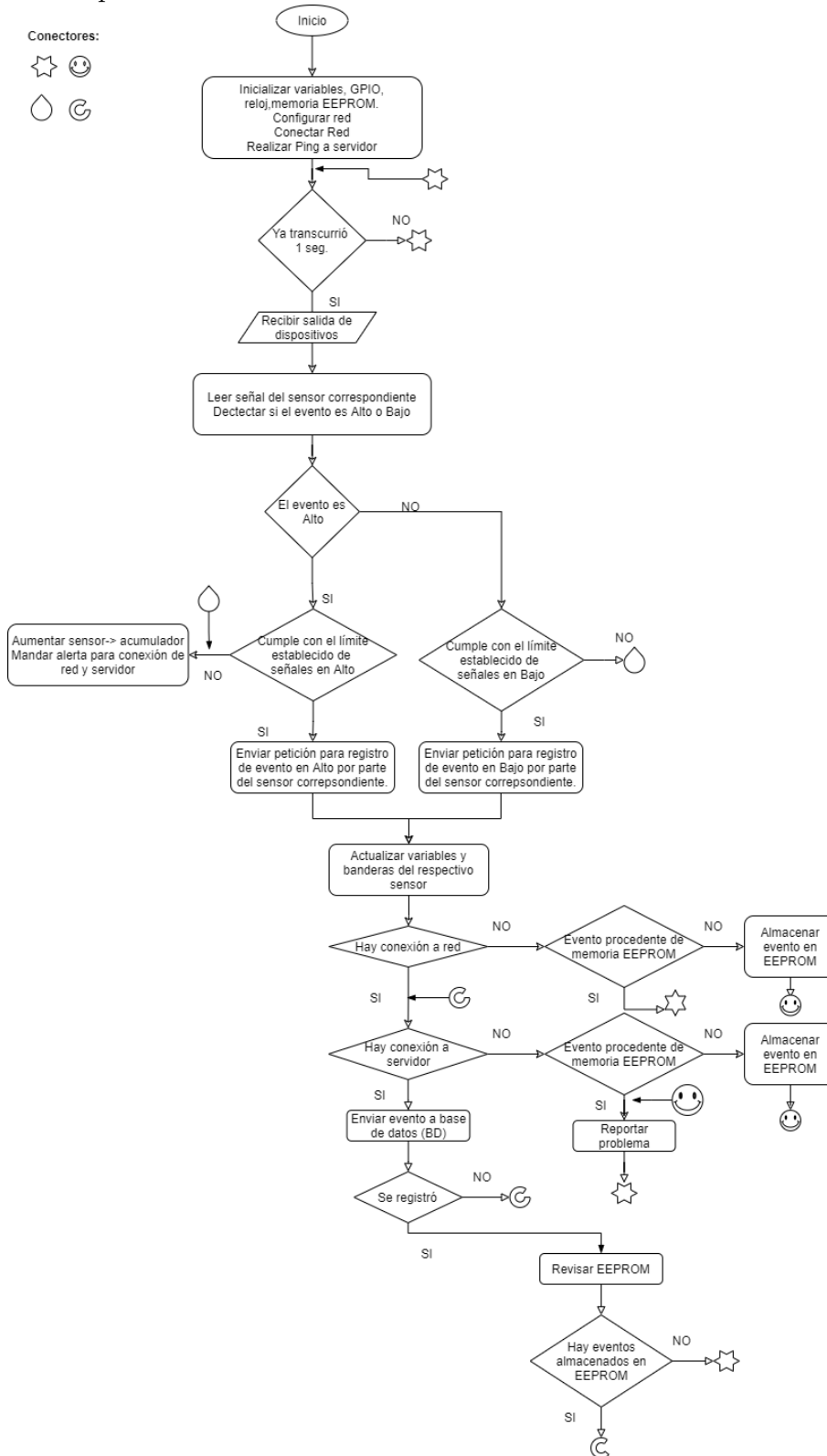
#### **3.7.3.1. Funcionamiento General**

Esta parte del desarrollo lógico se basa en el funcionamiento general que debe cumplir cada dispositivo sin importar los elementos que lo conforman, los procedimientos generales se abordan en la Tabla 3.10 y en la Fig 3.14 se presenta el diagrama de funcionamiento para los dispositivos de monitoreo involucrados en la red IoT.

Tabla 3.10: Funcionalidades generales de los dispositivos

<b>Funcionalidades generales</b>	<b>Elementos implicados</b>	<b>Variables implicadas en el funcionamiento</b>
<b>Establecer conexión</b>	Servidor de la vivienda Red asignada Servidor Blynk Base de datos	- Tiempo - Disponibilidad de red - Disponibilidad de servidores
<b>Detección de evento</b>	Sensores	- Tiempo - Tipo de sensor (analógicos/digitales)
<b>Verificación de conexiones</b>	Servidor de la vivienda Red asignada Servidor Blynk Base de datos	- Tiempo - Disponibilidad de red - Disponibilidad de servidores
<b>Almacenamiento de valores</b>	EEPROM Base de datos	- Tamaño de memoria - Conexión a red - Conexión a servidor - Disponibilidad de base de datos
<b>Revisión de valores almacenados</b>	EEPROM	- Tamaño de memoria
<b>Envío de Alertas</b>	Servidor Blynk Correo electrónico	- Conexión a red - Disponibilidad de servidor - Disponibilidad de base datos - Conexión a servidor Blynk

Figura 3.14: Diagrama del funcionamiento general de los dispositivos bajo el modo operacional “Modem sleep”



### Salida

Una vez programado cada dispositivo de acuerdo a la lógica planteada, la salida de ejecución puede observarse en la Fig 3.15 en la cual no se presenta ningún problema de conexión y todo resulta en el registro en tiempo y forma del evento detectado, la Fig 3.16 presenta un ejemplo de la ejecución del programa en caso de existir la falta de comunicación por parte de los dispositivos hacia el servidor, ya sea por una desconexión de éste o percance presentado, en este procedimiento la conexión a red permite enviar una alerta por medio de la aplicación móvil “Blynk” un correo electrónico al destinatario asignado en la cuál se le informa el problema presentado (Fig 3.17).

Figura 3.15: Corrida sin percances del programa diseñado para los dispositivos

```
0:15:53 0
0:15:54 0
0:15:55 0
0:15:56 1
Me debo preparar
Reconectar
9570650 <- Se conecto a red
time=5ms
4004<-- Me tomo host
Minimum = 5ms, Maximum = 17ms, Average = 8.00ms
ESP conectado
0:16:2 0
0:16:2 0
0:16:3 0
0:16:4 0
0:16:5 0
0:16:6 0
0:16:7 0
0:16:8 0
0:16:9 0
0:16:10 1
Me debo preparar
9710650 <- Se conecto a red
time=2ms
4004<-- Me tomo host
Minimum = 2ms, Maximum = 2ms, Average = 2.00ms
En orden
0:16:16 0
```

Figura 3.16: Corrida del programa diseñado para los dispositivos con servidor desconectado

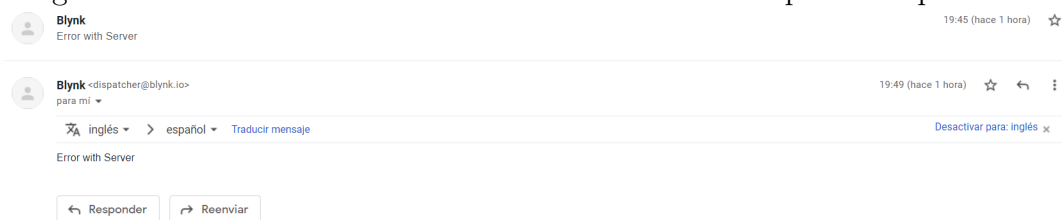
```

EEPROM set data: 1
192.168.100.20
Error with Server
Enviando alerta
cntAlert: 2
0:2:1 0-- 64.06
64.06
0:2:1 1-- 0:2:1 0-- 64.06
64.06
0:2:1 1-- 0:2:2 0-- 64.06
64.06
0:2:2 1-- 0:2:3 1-- 64.06
64.06
0:2:3 1-- 0:2:4 1-- 64.06
64.06
0:2:4 1-- I must check connections
time=0ms
24456Reconnect
1313340 <- Se conecto a red
time=0ms
40030:2:15 1-- 64.06
64.06
0:2:15 1-- Me ire a EEPROM
memory_address Inicial 10

Evento Guardado memory_address Final13
EEPROM set data: 1

```

Figura 3.17: Alerta enviada a correo destinado con el reporte del problema



### 3.7.3.2. Funcionamiento específico

Para la programación de los dispositivos se tomó en cuenta los tipos de sensores que los conformarían, el evento asignado a monitorear y la ubicación establecida de cada uno, dando paso a un análisis enfocado en un funcionamiento específico dividido en:

#### **Entrada de datos**

En su mayoría cada dispositivo se encuentra compuesto por un par de sensores acoplados al  $\mu C$  a través de una circuitería específica, esto implica que tanto el voltaje y pines a utilizar dependerán de la cantidad y tipo de sensor a emplear, lo que influye directamente en la programación inicial para el tratamiento de datos.

#### **Categoría de los sensores a trabajar**

Es necesario contemplar la disponibilidad de pines por parte del  $\mu C$  para el tratamiento de señales:

- Digitales
- Analógicos

#### **Calibración**

Esta sección depende de la conducta a detectar, el entorno y la sensibilidad del detector a emplear, el tratamiento en la entrada de datos varía mucho dependiendo si se trabaja con un sensor tipo analógico o digital.

En base a esta sección se establecen los límites para la toma de decisiones en el funcionamiento de los dispositivos, por lo que es importante realizar un estudio previo en el cual se establezca un panorama general del entorno y a las variables constantes en éste.

En base a un periodo de prueba de alrededor de 3 semanas se detectaron variables que lograban influir en el entorno afectando los resultados esperados por parte de los dispositivos. Estas se mencionan a continuación:

- Variación de temperatura en zonas específicas de vivienda
- Estado climático (días soleados/ nublados)
- Luz interior de la vivienda
- Acciones ejercidas del individuo
- Variación de tiempos en los que el individuo ejercía sus actividades

La variable que requirió una mayor dedicación fue la de tiempo, debido a que parte de los resultados esperados implicaba el conocimiento de la duración de conductas ejercidas por el usuario, por ello el tratamiento de calibración fue una parte prescindible para el funcionamiento del dispositivo basado en las conductas propias de la persona.

### **Condiciones en el funcionamiento de dispositivos**

Como hemos mencionado cada sensor tiene un evento propio a monitorear que en conjunto con el resto brindan un contexto de la conducta que ejerce el individuo en el entorno, no obstante, cada evento transcurre en diferentes tiempos con una duración y frecuencia variable que dependen directamente de la conducta desarrollada en el momento, por ello se deben contemplar los límites de activación establecidos a partir de la calibración de dispositivos.

El almacenar un evento completo indica que el sensor se ha mantenido en los límites de activación por el tiempo establecido (evento en alto) y ha pasado a mantenerse fuera de estos por el tiempo necesario (evento en bajo), indicando que una acción (conducta) tuvo lugar por

un determinado lapso y posteriormente concluyó.

### 3.7.4. Desarrollo Físico de los dispositivos

En comparación con la propuesta experimental, trabajar bajo el modo “*Modem Sleep*” permitió una implementación más viable en tiempos de desarrollo y accesible en recursos disponibles.

Como se ha mencionado la implementación de múltiples sensores requirió un análisis con la finalidad de elegir el mejor acoplamiento entre estos para cumplir la tarea asignada de cada dispositivo. A continuación, las Fig. 3.18-3.21 presentan los diagramas finales de cada dispositivo empleado en el entorno IoT

Figura 3.18: Diagrama para el funcionamiento del Medidor AC

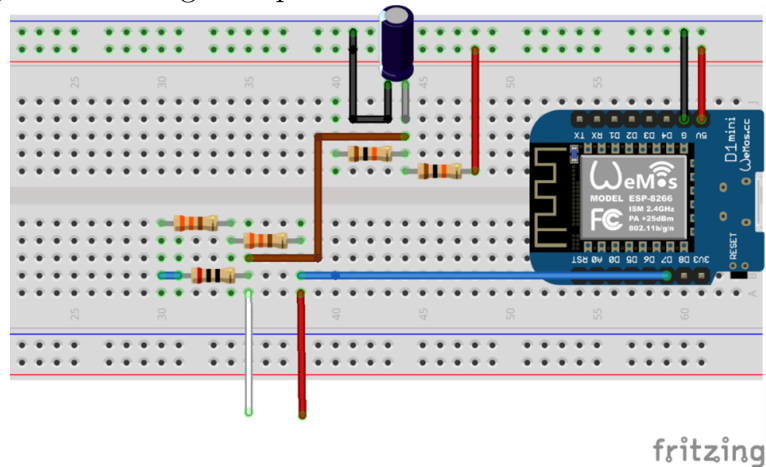


Figura 3.19: Diagrama para el funcionamiento de PIR-LDR

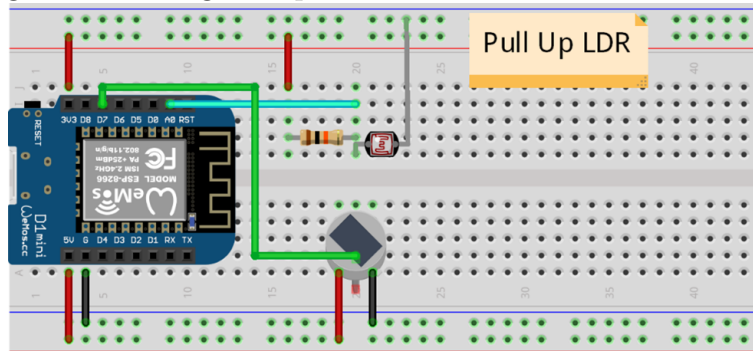


Figura 3.20: Diagrama para el funcionamiento de Ultrasonico-Interruptor (Switch)

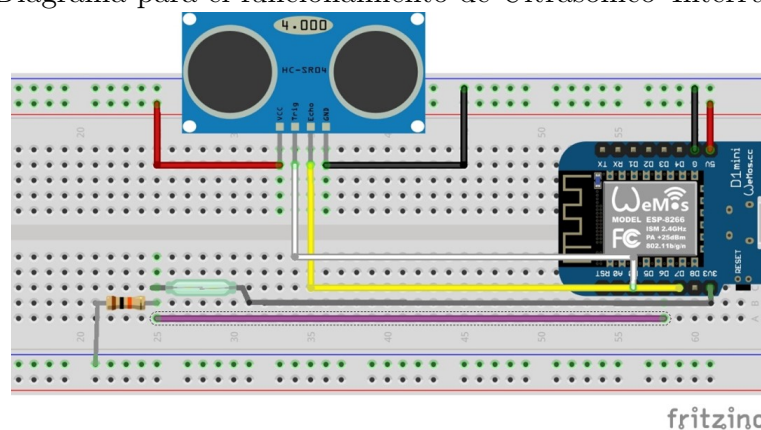
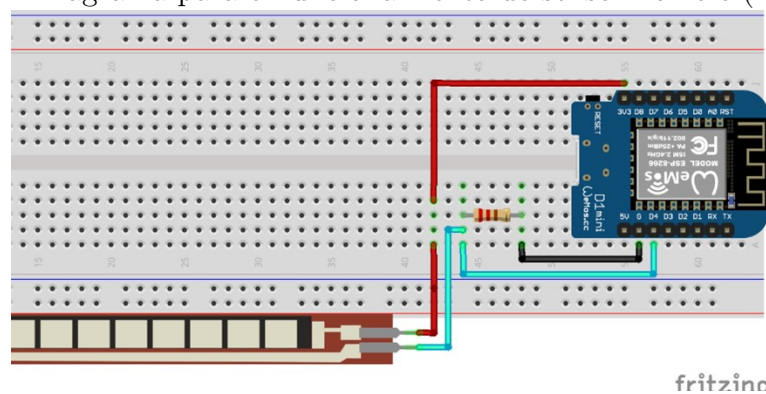


Figura 3.21: Diagrama para el funcionamiento de sensor flexible (Flex sensor)



**Pruebas del funcionamiento Lógico–Físico** La Fig.3.22 presenta las pruebas previas a la implementación final de los dispositivos.

Figura 3.22: Prueba de dispositivo para el monitoreo de movimiento



### 3.7.5. Instalación final de red con el uso de servidores, AP, tecnologías de comunicación y dispositivos desarrollados

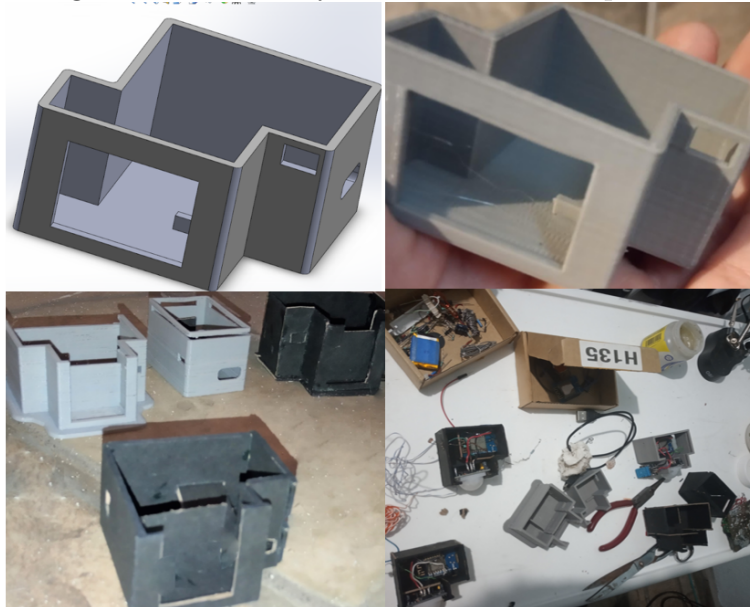
Una vez desarrollado tanto la parte lógica como física de los dispositivos, se procedió a trabajar en su implementación, lo que requirió un análisis para el uso de materiales a emplear en la instalación de red teniendo en cuenta la protección tanto para los dispositivos como de la vivienda.

#### 3.7.5.1. Protección para los dispositivos

Al estar constituido los dispositivos por diversos sensores no existe un diseño de carcasa estándar que cubra las necesidades de protección para cada uno, por ello se requirió elaborar un diseño específico por unidad. En un inicio se hizo uso del software SolidWorks para el diseño de carcasas y se logró la impresión 3D de dos prototipos, no obstante, por cuestiones de aislamiento causadas por la epidemia de coronavirus (COVID 19), el proceso de impresión no se encontraba dentro de las actividades disponibles por lo que se optó por elaborar las car-

casas manualmente basándose en los diseños realizados y empleando materiales arquitectónicos. Los resultados se muestran a continuación en la Fig. 3.23

Figura 3.23: Diseño y construcción de dispositivos



### 3.7.5.2. Protección para la vivienda

La protección del entorno consistió en la selección de materiales que cumplieran con las características de seguridad necesarias para la instalación de red, el cableado, adaptadores múltiples y cargadores debían cumplir con los siguientes requisitos:

- No presentar daños físicos en la estructura
- Estructura completamente sellada y protegida
- Contar con materiales resistentes al calor
- Estar debidamente adaptados para su correcto funcionamiento

### Medidas de seguridad

### 3.8. RESULTADO FINALES DE LA IMPLEMENTACIÓN EN RED APLICANDO EC99

- Dentro de las medidas de seguridad aplicadas en la implementación de red se encuentran:
- No sobrecargar las fuentes de energía eléctrica.
- No realizar instalaciones en entornos no adecuados
- No realizar conexiones que afecten la movilidad del individuo al interactúan con del entorno
- Elegir el cableado adecuado a la instalación a realizar
- Trabajar cada dispositivo bajo la alimentación eléctrica indicada.

### **3.8. Resultado finales de la implementación en red aplicando EC**

Transcurridas las tres semanas de prueba se procedió a iniciar el periodo de estudio final, el cual consistió en monitorear el entorno durante los meses de octubre a diciembre a través de la diseñada e implementada red IoT.

Por parte de la red se obtuvieron resultados prometedores, una parte del procesamiento de datos como la revisión de eventos, validación, monitoreo de entorno, envío de alertas y almacenamiento fue realizado por los dispositivos con la finalidad de evitar la dependencia de conexión a entornos en la nube y respuesta de estos para el correcto funcionamiento de los elementos, por otra parte el almacenamiento de datos funcionó de acuerdo a lo esperado y la información brindada por el trabajo en conjunto de todos los dispositivos fue suficiente para apoyar el análisis, desarrollo e implementación de un modelo de IA alimentado por todo el conjunto de datos obtenido. La imágenes siguientes Fig 3.24-3.29 presentan la mayoría de las instalaciones realizadas dentro del entorno, como se puede observar cada una requirió un cableado que permitiera su incorporación al área asignada, un adherente de acuerdo al tipo de superficie seleccionada y una protección diseñada de acuerdo a las necesidades de propio dispositivo implementado.

Figura 3.24: Dispositivo colocado en entrada



Figura 3.25: Dispositivo colocado en refrigerador



3.8. RESULTADO FINALES DE LA IMPLEMENTACIÓN EN RED APLICANDO EC101

Figura 3.26: Dispositivo colocado entre Horno / Gabeta



Figura 3.27: Dispositivo colocado sobre la mesa



Figura 3.28: Dispositivo colocado en baño



Figura 3.29: Servidor ubicado en área de Cocina-Sala



### **3.9. Monitoreo del individuo a través del uso de tecnologías externas al entorno IoT**

Entre los beneficios del uso de la red IoT se encuentra la accesibilidad a las tecnologías vía remota, lo que permitió monitorear los cambios en el entorno derivados de las actividades realizadas por el individuo dentro de este, no obstante una vez que el usuario abandonaba la vivienda fue necesario el uso de otras tecnologías para el seguimiento de las conductas que ejercía dentro de un entorno laboral o social.

La selección de las tecnologías a emplear implicó una evaluación sobre la información relevante a monitorear y los medios factibles para lograrlo todo esto sin afectar la comodidad, percepción y anonimato del individuo. Finalmente debido a su practicidad, discreción y portabilidad la aplicación móvil “Daylio” y el smartwatch “Garmin” fueron seleccionadas para complementar la información obtenida del entorno y de esta manera registrar un mayor conjunto de datos que brinde la posibilidad de desarrollar un modelo de ML más completo y fiable.

#### **3.9.1. Tecnología “Daylio”**

Daylio es el nombre de una aplicación móvil en formato diario que además de registrar actividades y acontecimientos del diario permite el seguimiento del estado de ánimo por parte del usuario, en la Fig. 3.30 se puede observar parte del diseño en la aplicación.

Figura 3.30: Diseño parcial de la app móvil “Daylio”



A partir de un conjunto de opciones en la plataforma, el usuario fue capaz de adaptar el seguimiento a través de las funcionalidades disponibles, permitiendo de esta manera obtener información sobre las actividades sociales, tareas y objetivos personales que el individuo desarrollaba dentro y fuera del hogar.

### Instrucciones de uso para el individuo

Durante todos los días de la aplicación de estudio se le solicitó al usuario llenar al menos una vez al día las secciones en la aplicación que él mismo había creado de acuerdo a las opciones disponibles por parte del sistema y los gustos y necesidades propias del individuo, para facilitar esta tarea se empleó la funcionalidad de alarma que brindaba la misma aplicación.

### Resultados de implementación

La Fig. 3.31 presenta el formato final obtenido por parte de la plataforma de acuerdo a la información recopilada y disponible del sistema.

### 3.9. MONITOREO DEL INDIVIDUO A TRÁVES DEL USO DE TECNOLOGÍAS EXTERNAS AL EN

Figura 3.31: Conjunto final de datos obtenido de la plataforma “Daylio”

full_date	date	weekday	time	mood	activities	note_title	note
12/10/2020	10 de dicié	jueves	11:58 p. m	meh	amigos   sueño medic	Sigo en el trabajo	son las 12 pm ayuda
12/9/2020	9 de dicien	miércoles	8:01 p. m.	bien	amigos   buen sueño	andar   amabilidad   compras   limpiar	
12/8/2020	8 de dicien	martes	5:26 p. m.	estresado	amigos   buen sueño	Muchas tareas realizadas en el trabajo	
12/7/2020	7 de dicien	lunes	5:26 p. m.	meh	amigos   sueño medic	Muchos pendientes en el trabajo	
12/6/2020	6 de dicien	domingo	8:01 p. m.	increíble	amigos   buen sueño	andar	
12/5/2020	5 de dicien	sábado	11:55 p. m	bien	amigos   buen sueño	comer sano   beber agua   andar	
12/4/2020	4 de dicien	viernes	8:02 p. m.	bien	amigos   buen sueño	andar   compras   limpiar   cocinar	
12/3/2020	3 de dicien	jueves	7:56 p. m.	estresado	amigos   sueño medio	andar	
12/2/2020	2 de dicien	miércoles	4:04 p. m.	nostálgico	buen sueño   andar	Al ayudarle a una señora, y pensar que debería	

#### 3.9.2. Tecnología Garmin

Por parte de la línea Garmin se empleó uno de los dispositivos llamado “Garmin vivo smart 3”, modelo que gracias a sus herramientas de monitorización de la actividad física como el VO2 max, la edad fitness y el entrenamiento de resistencias apoyado por sensores ópticos de frecuencia cardíaca permitía el registro de los niveles de sueño, frecuencia cardíaca y estrés presentados por el individuo.

Las variable finales a evaluar del dispositivos fueorn las siguientes:

- **Frecuencia Cardíaca (FC)**

Este indicador puede brindar informacion del estado físico y salud cardíaca del individuo, generalmente se mide cuando este se encuentra despierto,relajado y en reposo. La frecuencia cardíca media es de 60-100 ppm. Las personas que estan muy en forma pueden observar lecturas de hasta 40ppm [44].

- **Estrés**

El nivel de estrés analizado por el reloj Garmin permite al usuario determinar su nivel en función de la variabilidad de la FC.

El tiempo variable entre cada latido cardíaco está regulado por el sistema nervisos autónomo del cuerpo y la menor variabilidad entre latidos equivale a mayores niveles de estrés, caso contrario ocurre cuando el aumento en la variabilidad indica un menor estrés en el individuo. La Fig. 3.32 presenta la información correspondiente al análisis de la FC y Estrés.

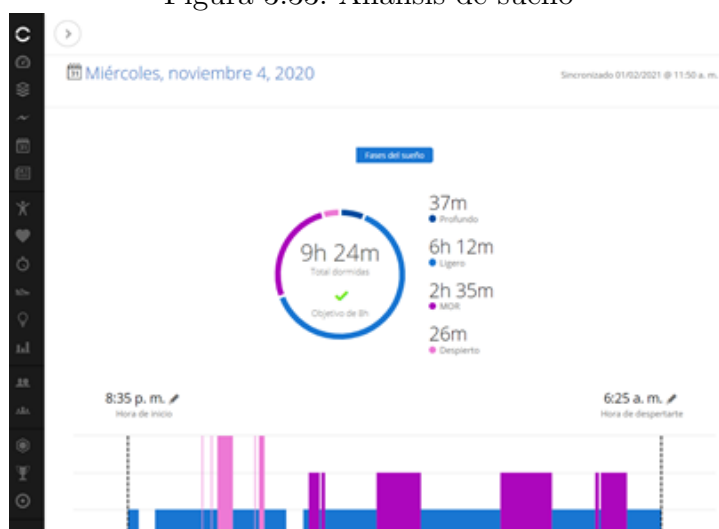
Figura 3.32: Análisis de la FC y niveles de Estrés



### ■ Descanso

El descanso es un periodo de reposos del cuerpo, como el sueño o momentos tranquilos durante el día en los que el cuerpo tiene tiempo para recuperarse del estrés y recuperar la energía perdida en la realización de las actividades diarias para poder mantener la salud, donde un intervalo entre 7 y 8 horas resulta recomendable para lograr un buen descanso. El dispositivo Garmin permite el análisis de sueño a través de la FC del usuario la Fig.3.33 presenta un visualización parcial de la plataforma relacionada al análisis de sueño.

Figura 3.33: Análisis de sueño

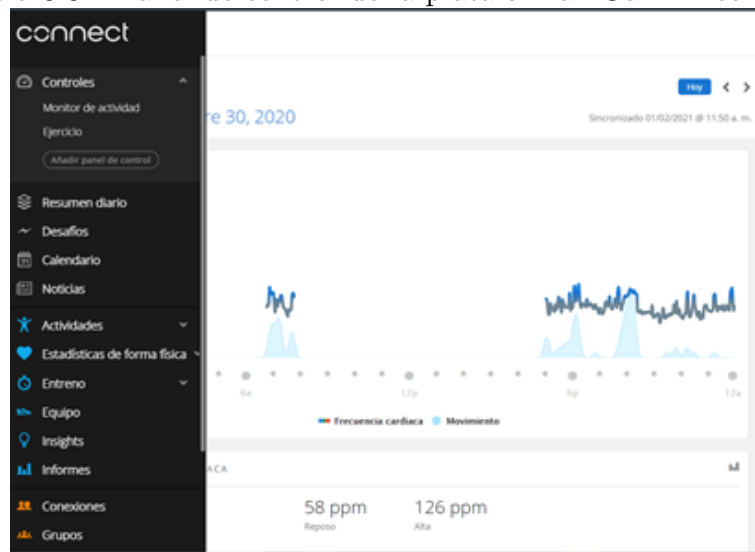


### 3.9. MONITOREO DEL INDIVIDUO A TRÁVES DEL USO DE TECNOLOGÍAS EXTERNAS AL ENTORNO

#### Seguimiento de estado físico del usuario

El análisis y evaluación del estado físico del individuo se visualizaba en lapsos de tiempo a través de la plataforma “Garmin Connect” la cual brindaba la información relacionada a las actividades realizadas a lo largo del día así como su rendimiento físico al ejecutarla. La Fig. 3.34 señala la presentación por parte del sistema de los datos procesados.

Figura 3.34: Panel de control de la plataforma “Garmin connect”



#### Resultados del uso diario del dispositivo Garmin

Por parte de la plataforma fue posible observar el rendimiento y estado físico del individuo en diferentes intervalos de tiempo (días, semanas, meses y años) los cuales se presentan en la Fig 3.35, no obstante la información final brindada por el sistema consistió en un archivo CSV correspondiente a la variable seleccionada y el registro diario del monitoreo hacia el individuo, en la Fig. 3.36 se presenta un ejemplo de lo mencionado.

Figura 3.35: Reporte gráfico del análisis de los niveles de estrés

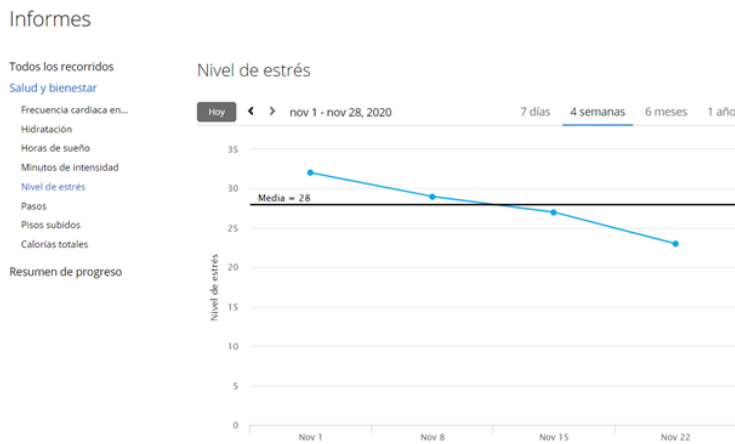


Figura 3.36: Reporte CSV del análisis de los niveles de estrés

	NOV
Dom	31
Lun	30
Mar	40
Mié	19
Jue	27
Vie	38
Sáb.	42
Dom	59
Lun	9
Mar	8
Mié	13
Jue	19
Vie	60

### **3.10. Imprevistos generales de las tecnologías empleadas (Entorno IoT, App “Daylio” y smartwatch Garmin)**

El conjunto de las tecnologías empleadas (Entorno IoT, App “Daylio” y smartwatch “Garmin vivosmart 3”) para el desarrollo de este estudio permitió obtener datos correspondientes a tres diferentes ámbitos del individuo (personal, laboral y social), no obstante el uso de cada una implicó una serie de acontecimientos que intervinieron en el proceso de captura de la información objetivo, a continuación se detallarán estos imprevistos en cada tecnología empleada.

#### **3.10.0.1. Imprevistos del entorno IoT**

Durante el periodo de monitoreo se presentaron en contadas ocasiones eventos que interrumpieron la recolección constante de información, estos fueron:

- Cortes de luz (Apagones)
- Desconexión accidental de dispositivos por parte del individuo
- Visitas inesperadas a la vivienda
- Una corta estancia fuera de la vivienda por parte del individuo
- Calidad y servicio del internet brindado por la compañía telefónica

#### **3.10.0.2. Imprevistos durante el uso de la aplicación móvil “Daylio”**

Durante el periodo de monitoreo se presentaron en contadas ocasiones eventos que no apoyaron la recolección de los datos en tiempo y forma, estos se presentan a continuación:

- El usuario ignoró una de las 2 alarmas establecidas para el llenado de la aplicación.
- El usuario realizó el llenado de la App móvil fuera de los horarios establecidos.
- La información brindada a la aplicación por parte del usuario no contenía la misma cronología de los días previos.

### 3.10.0.3. Imprevistos durante el uso del smartwatch “Garmin”

El uso de esta tecnología resultó muy útil debido a su alta precisión sin embargo es importante tener en cuenta los elementos y circunstancias que pueden afectar su rendimiento, a continuación se mencionan algunos de los ejemplos que se presentaron en el desarrollo de este trabajo.

- El cuidado y mantenimiento del dispositivo brindado por el usuario
- Un nulo funcionamiento debido a su falta de carga
- La colocación del dispositivo por parte del usuario
- Color de piel del individuo
- Falta de condiciones óptimas del dispositivo
- El ajuste a la muñeca correcto, es decir, que éste se encuentre en contacto directo con la piel, sin llegar a apretarla limitando de esta manera el flujo sanguíneo

Una vez contemplados los imprevistos presentados por las diferentes tecnologías fue necesario tomar acción con la finalidad de desarrollar un sistema que tuviera en cuenta los elementos presentados y trabajar bajo un diseño y análisis capaz de tratar la información de manera que no afectara el desempeño y calidad de los datos finales, en las secciones siguientes se abordaran las medidas tomadas para cada caso.

### 3.11. Introducción al análisis de datos

Una vez recopilada la información del entorno a partir de la red IoT desarrollada, se procedió a obtener de cada servidor Edge los conjuntos de datos (datasets) correspondientes a cada dispositivo instalado en las áreas mencionadas en la Tabla 3.6. La información recabada por la aplicación móvil “Daylio” se exportó en formato CSV al igual que la estadística generada a partir del uso del reloj inteligente “Garmin”. El uso en conjunto de estos medios permitió generar una extensa cantidad de información sobre la interacción del individuo dentro y fuera de su entorno habitacional, lo que implicó un aumento en la complejidad de la metodología para el procesamiento de los datos debido a la inexistente relación entre el funcionamiento de los dispositivos entre sí a pesar de su orientación hacia el individuo de estudio, es decir cada elemento trabajaba de manera independiente al resto.

El análisis de la información se centró inicialmente en los conjuntos de datos obtenidos por parte de la red IoT debido a su gran cantidad de extensión, diversidad entre sí y falta de cronología para la formación y detección de los patrones de conducta ejercidos por el individuo dentro del entorno. Posteriormente el resultado fue complementado con la información de la aplicación móvil “Daylio” y el reloj inteligente “Garmin” para su incorporación y procesamiento por medio del modelo de ML desarrollado para la detección de anomalías dentro de los patrones de conducta establecidos. A continuación se detalla el procedimiento desarrollado para el manejo en conjunto de toda la información por medio de un diseño experimental que posteriormente se empleó de base para el diseño y desarrollo del modelo de ML final capaz de detectar las conductas anómalas dentro del patrón de conductas monitoreado.

### 3.12. Recolección los conjuntos de datos generados por la red IoT

A partir de la información almacenada en las bases de datos “Sala” y “Cuarto”, fue posible obtener la frecuencia, fecha, hora y duración de la conducta asignada a monitorear por cada dispositivo empleado en la red IoT.

El análisis inicial consistió en unir en un solo archivo CSV los datos obtenidos de cada base de datos en conjunto, no obstante, al tratarse de información correspondiente a conductas desarrolladas en diferentes áreas de la vivienda se decidió dividir los conjuntos de datos en dos archivos CSV “csv\_A” y “csv\_B” cada uno con la información correspondiente a los servidores implementados, evitando de esta manera la generación de valores “ruido” a causa del desajuste de horarios entre las actividades realizadas.

Recordemos que el archivo “csv\_A” hace referencia a la información correspondiente de las conductas sensadas en las áreas de sala, cocina y baño , consultar Tabla 3.9 y Fig. 3.37 para mayor detalle; y el archivo “csv\_B” corresponde a los datos brindados por la habitación del individuo Fig.3.38.

Figura 3.37: Estructura del archivo “csv\_A”

EVENTO	FECHA	ID
0	10/26/2020 6:49	0
1	10/26/2020 6:50	1
0	10/26/2020 6:51	1
1	10/26/2020 7:09	1
0	10/26/2020 7:10	1
1	10/26/2020 7:11	1
0	10/26/2020 7:13	1
1	10/26/2020 7:13	0

**Nota.** En la figura se presenta el parte del funcionamiento en fecha y hora del sensor de luz (ID,0) y movimiento (ID,1)

### 3.13. ANÁLISIS PARA LA CONSTRUCCIÓN DEL MODELO DE PREPROCESAMIENTO QUE TR

Figura 3.38: Estructura del archivo “csv\_B”

EVEN TO	FECHA	ID
1	10/26/2020 0:17	2
0	10/26/2020 0:17	2
1	10/26/2020 1:25	2
0	10/26/2020 1:49	2
1	10/26/2020 1:50	2
0	10/26/2020 1:51	2
1	10/26/2020 1:51	2

**Nota.** En la figura se presenta parte del funcionamiento en fecha y hora del sensor flexible (ID,2)

### 3.13. Análisis para la construcción del modelo de pre-procesamiento que transformara datos en patrones de conducta

Para el tratamiento de la información presentada en los archivos “csv\_A” y “csv\_B”, fue necesario el diseño y desarrollo de un modelo de preprocesamiento capaz de transformar los datos proporcionados en patrones de conducta ejercidos por el usuario en turno y para para ello primeramente se plantearon las incógnitas iniciales presentadas en la Tabla 3.11.

Tabla 3.11: Incógnitas para el diseño del modelo de preprocesamiento encargado de la detección y desarrollo de los patrones de conducta ejercidos por el usuario'

<i>Incógnitas para el análisis, diseño y desarrollo del sistema de detección de patrones de conducta</i>	<i>Respuesta</i>
“¿Número de datasets a tratar por parte del sistema?”	dos, el primero fue el dataset “csv_A” el cual contenía la información correspondiente de los dispositivos en las áreas de <i>sala, cocina y baño</i> , el segundo al área habitacional del individuo.
“¿Tipos de variables a manipular en los conjuntos de datos?”	Variables tipo entero con valores de 0 y 1 Variables tipo cadena y entero en formato fecha y hora
“¿La información de cuantos dispositivos se debe pre procesar?”.	En el archivo “csv_A” se encuentran los conjuntos de información de un total de 9 sensores.  En el archivo “csv_B” se encuentran los conjuntos de información de un total de 4 sensores
“¿A partir de que datos se espera formar la cronología de los patrones de conducta ejercidos?”	Las variables “Fecha” y “Hora” permiten revisar las conductas presentadas en los intervalos de hora establecidos.
“¿Afecta al conjunto de datos final la gran diferencia de horarios entre el desarrollo de actividades (conductas)?”	Si, debido a que existen actividades como <i>dormir</i> que requieren únicamente la activación de una mínima parte de dispositivos, manteniendo al resto de los elementos “desactivados” al no existir otra conducta a detectar en el momento ( <i>comer, cocinar, trabajar etc.</i> ), lo que significa una gran aglomeración de valores ‘0’ en el conjunto final de datos obtenido lo que puede generar un desbalanceo de valores durante el entrenamiento de ML posterior.
“¿Se espera obtener un conjunto de datos para trabajar el aprendizaje supervisado, no supervisado o semi supervisado?”	No supervisado debido a que el modelo de procesamiento genera un conjunto de datos no etiquetado, es decir no hay una salida esperada a la secuencia de datos de entrada.

### 3.14. Diseño experimental

#### 3.14.1. Detección y estructuración de los patrones de conducta ejercidos por el individuo

En base a las incógnitas evaluadas, se llegó a la conclusión que la cronología final de los patrones de conducta sería desarrollada a partir de los intervalos de fecha y hora, donde

cada renglón del conjunto final de datos representaría el patrón de conductas desarrolladas dentro del lapso establecido. Los lapsos de tiempo se eligieron conforme a los horarios donde el individuo se encontraba en la vivienda realizando las actividades registradas en los archivos “csv\_A” o “csv\_B”, esto con la finalidad de evitar tener renglones o cronologías donde los valores fueran únicamente de ‘0’ debido a la inexistencia de conductas presentadas en horarios “vacíos”.

Para el archivo “csv\_A” se eligieron inicialmente los horarios de 17:00 pm, 20:00 pm, 23:00 pm, 2:00 am, 5:00 am y 8:00 am, todos con 3 horas de diferencia entre sí.

Para el archivo “csv\_B” se eligieron inicialmente los horarios matutinos de 17:00pm, 20:00 pm, 23:00pm, 2:00am, 5:00am y 8:00 am, todos con 3 horas de diferencia entre sí.

En ambos casos a partir de cada inicio de hora establecido se obtenían las actividades realizadas durante la hora siguiente en lapsos de 1 hora, es decir el modelo seleccionaba las conductas que se encontraban dentro de los límites de horarios y formaba la cronología correspondiente al intervalo de tiempo, a continuación, se explica el procedimiento en la Tabla 3.12 con el uso del archivo “csv\_A”

Tabla 3.12: Ejemplo del procedimiento de detección de conductas en el archivo “csv\_A” a partir de los intervalos de tiempo establecidos.’

Horario	Conductas por intervalos de tiempo	Elementos involucrados	Conductas deducidas en base a los sensores activados
Intervalo: 1 hora	Empty DataFrame Columns: [Unnamed: 0, EVENTO, FECHA, ID] Index: []	No datos	N/A
Hora inicial: 17:00pm	<pre>                 Unnamed: 0  EVENTO          FECHA  ID 12          9          0  2020-10-26 18:24:00  0 13          8          1  2020-10-26 18:24:00  1 14         10          0  2020-10-26 18:25:00  1 16         13          1  2020-10-26 18:28:00  1 18         14          0  2020-10-26 18:29:00  1 19         28          1  2020-10-26 18:34:00  2 20         29          0  2020-10-26 18:35:00  2 21         41          1  2020-10-26 18:36:00  3 22         42          0  2020-10-26 18:37:00  3 23         30          1  2020-10-26 18:40:00  2 24         31          0  2020-10-26 18:40:00  2 25         32          1  2020-10-26 18:43:00  2 26         43          1  2020-10-26 18:44:00  3 27         44          0  2020-10-26 18:45:00  3 28         45          1  2020-10-26 18:48:00  3 29         33          1  2020-10-26 18:51:00  2 30         34          0  2020-10-26 18:51:00  2 31         35          1  2020-10-26 18:52:00  2 32         36          0  2020-10-26 18:53:00  2 33         37          1  2020-10-26 18:54:00  2 34         38          0  2020-10-26 18:54:00  2 35         46          1  2020-10-26 18:58:00  3 36         47          0  2020-10-26 18:59:00  3             </pre>	0,1 = Entrada de Sala 2= Refrigerador 3 = Estufa	<ul style="list-style-type: none"> <li>○ Desplazamiento en el entorno de sala</li> <li>○ Cocinar</li> </ul>
Hora final: 20:00	<pre>                 Unnamed: 0  EVENTO          FECHA  ID 38         49          1  2020-10-26 19:07:00  3 39         50          0  2020-10-26 19:07:00  3 40         51          1  2020-10-26 19:08:00  5 41         52          0  2020-10-26 19:09:00  5 42         53          1  2020-10-26 19:10:00  5 43         54          0  2020-10-26 19:11:00  5 44         55          1  2020-10-26 19:12:00  5 45         56          0  2020-10-26 19:12:00  5 47         59          1  2020-10-26 19:14:00  5 49         60          0  2020-10-26 19:15:00  5 50         61          1  2020-10-26 19:21:00  5 51         62          0  2020-10-26 19:21:00  5 52         15          1  2020-10-26 19:39:00  1 53         16          1  2020-10-26 19:40:00  0 54         17          0  2020-10-26 19:41:00  1 55         63          1  2020-10-26 19:47:00  7 56         64          0  2020-10-26 19:48:00  7 57         65          1  2020-10-26 19:49:00  7 58         66          0  2020-10-26 19:50:00  7 59         67          1  2020-10-26 19:51:00  7 60         68          0  2020-10-26 19:52:00  8 61         69          1  2020-10-26 19:54:00  7 62         70          0  2020-10-26 19:54:00  7 63         71          1  2020-10-26 19:56:00  7 64         72          0  2020-10-26 19:56:00  7 66         75          1  2020-10-26 19:59:00  7             </pre>	0,1 = Entrada de Sala 3 = Estufa 5 = Mesa 7,8= Cuarto de baño	<ul style="list-style-type: none"> <li>• Desplazamiento en el entorno de sala</li> <li>• Comer</li> <li>• Ducharse</li> </ul>
17:00-18:00	DateFrame: 2 order is : [0, 1, 2, 3]		
18:00-19:00			
19:00-20:00			

**Nota.** En este archivo se encuentra la información correspondiente a las actividades (conductas) desarrolladas en la sala, cocina y baño dentro el primer lapso de tiempo de 17:00 pm a 20:00 pm.

Una vez obtenido el conjunto de actividades realizadas por los tiempos fijados, se pudo visualizar que las mismas conductas podían presentarse en diferentes horarios, traslapadas entre sí y en constante repetición, ninguna conducta ocurría sin influir en el resto y a pesar de que podía existir un patrón de conducta repetitivo no necesariamente implicaba un mismo orden en las acciones, por ejemplo, cuando el individuo se sentaba a comer, también se levantaba a abrir el refrigerador, se acercaba a la estufa en ocasiones y regresaba a la mesa. Cuando iba a dormir, existía un desplazamiento por la cocina, la sala, un cambio en el estado de luces, desplazamiento en la habitación personal, movimiento en la cama, cambio de luces en caso de prender lámpara de noche.

Los patrones podían ser observados a través de la cronología formada por el modelo, no obstante, al ser muy larga la secuencia por día se requirió sintetizar la información obtenida a través del uso de atributos (Features). Estos atributos permitirían identificar las acciones que habían tenido lugar en cada lapso establecido por día. Con 5 lapsos de tiempo por día (5 renglones del conjunto de datos final) y 47 días de estudio, un total de 235 tuplas conformarían el conjunto de datos final para la detección de anomalías a través del modelo de ML desarrollado.

#### **3.14.1.1. Lógica para el desarrollo cronológico de los patrones de conducta ejercidos diariamente empleando el uso de atributos para la sinterización de información.**

Para el análisis de los patrones de conducta ejercidos por el usuario fue necesario realizar el registro de todas las secuencias en un documento CSV que posteriormente se emplearía en la entrada del modelo de ML desarrollado para la detección de anomalías dentro de los patrones almacenados. Este registro cronológico de conductas requirió el uso de atributos, los cuales al ser aplicados a cada uno de los sensores empleados en la red IoT, debían cumplir en un limitado número de elementos el registro del indicador (sensor), tiempo, duración,

frecuencia y orden en el que las conductas eran desarrolladas por el individuo a lo largo del día, todo lo anterior con el fin de evitar un excesivo número de parámetros en el conjunto final de datos.

Un total de 5 atributos (AS, CF, MIN, MAX, CVT) se emplearon a lo largo de 7 lapsos de tiempo establecido por sensor con la finalidad de registrar los patrones de conducta ejercidos a lo largo del día por el individuo, a continuación, se describe con mayor detalle cada uno de ellos.

El registro final de los sensores implementados a lo largo de la sala, cocina y baño al igual que los del cuarto del individuo implicó el uso 5 atributos diferentes (Escala de anomalía, Cantidad de frames, MIN, MAX, CVT) asignados a cada uno de los 13 sensores empleados, donde la actividad a registrar por cada uno contó con su respectiva abreviación señalada en la Tabla 3.13 y Tabla 3.14. A continuación, se describe con mayor detalle cada uno de los parámetros empleados y para complementar la información presentada se recomienda consultar la Tabla 3.15 “Abreviaturas de las actividades monitoreadas en el entorno con sus respectivos ID”

**Escala de anomalía (AS).** Como se explicó en la Tabla 3.12 “Ejemplos del procedimiento de detección de conductas en el archivo “csv\_A”, a partir de los intervalos de tiempos establecidos” la detección de conductas se realizaba en lapsos de tiempo específicos, los cuales podían ser modificados tanto en la duración de tiempo como en los intervalos (frames) en los que se analizaba el conjunto de actividades desarrolladas en la vivienda.

Basándonos en el ajuste de estos parámetros, fue posible trabajar lapsos específicos, descartando de esta manera los horarios donde existiera nula actividad física dentro del entorno.

La escala de anomalía implica la sumatoria en secuencia de los frames donde el sensor en turno detectó la información de interés a su función, con cada nuevo

intervalo de horario y sensor la escala se reinicia con la finalidad de capturar en cada uno de los 5 lapsos de tiempo establecidos la información referente a la cronología de conductas compuesta por el conjunto de dispositivos empleados.

Un gran valor de sumatoria es sinónimo de que la actividad monitoreada ha sido desarrollada consecutivamente durante los intervalos correspondientes al horario, en caso contrario, un valor mínimo de sumatoria implica una mínima aparición por parte de la actividad en turno. Las Fig. 3.39 y 3.40 explican con mayor detalle el procedimiento para el cálculo de la escala de anomalía

Figura 3.39: Ejemplo del calculo en la AS de la conducta alimentarse

<i>Conducta</i>	<i>Hora inicial</i>	<i>Hora final</i>	<i>Intervalos de tiempo (Frames)</i>	<i>Secuencia de frames total</i>
Comer	5:30	6:30	30 min	1,2
Si la conducta tuvo lugar la primera media hora, la sumatoria de la escala de anomalía sería igual a 1, si ocurrió durante la segunda media el resultado es 2 y en caso de transcurrir durante toda la hora completa sería igual a 3.				

A continuación, se detalla un ejemplo aplicando los parámetros actuales del modelo en la Fig. 3.40

Figura 3.40: Ejemplo del calculo de la AS sobre la duración del desplazamiento en el entorno

<i>Conducta</i>	<i>Hora inicial</i>	<i>Hora final</i>	<i>Intervalo de tiempo (Frames)</i>	<i>Secuencia de frames total</i>
Desplazamiento en el entorno de sala	17:00pm	20:00pm	1 hora	1,2,3
Si la conducta tuvo lugar en la primera hora, la sumatoria en la escala de anomalía sería igual a 1, si ocurrió durante la segunda el resultado es 2, si fue en la última hora sería igual a 3 y en caso de transcurrir durante todas las tres horas completas sería igual 6.				

**Cantidad de frames (CF).** Aplicar la escala de anomalía brinda información de los intervalos de tiempo donde una conducta tuvo lugar, no obstante el resultado puede deberse a una sumatoria en secuencia (1, 2, . . . n) de los intervalos donde existió consecutivamente la actividad o a la operación entre intervalos de tiempo separados (1,3) generando la misma sumatoria, por ello el atributo

“Cantidad de frames” se integró a los parámetros empleados con la finalidad de reducir la probabilidad de errores en la elección de opciones.

Para obtener el resultado de esta variable, el modelo se encargaba de revisar en cada uno de los intervalos en turno la existencia de información relacionada a algún sensor de la red IoT y en caso de existir coincidencia un acumulador referente al sensor se actualizaba para registrar en el conjunto final de datos la cantidad de intervalos en los que la actividad específica había sido detectada en cada uno de los 5 lapsos por día. De esta manera en caso de que el resultado en la escala de anomalías calculara un 4, número que puede obtenerse a través de la sumatoria entre la secuencia 1,3 (1+3) o simplemente abordando el intervalo 4 directamente, si la variable “Cantidad de frames” registraba un 2, automáticamente se descartaría la segunda opción, lo que en otras palabras significaría que una conducta se presentó a lo largo de 2 intervalos de tiempo dentro de los límites de horario en cuestión.

**Valor mínimo de frame (MIN).** Al emplear el atributo CF se obtenía la cantidad de frames en los que una conducta había tenido lugar y en conjunto con el atributo SA era posible determinar el momento aproximado en el que la actividad había transcurrido, no obstante, existían casos en el que el valor de estos atributos concordaba en dos casos diferentes. En el siguiente ejemplo se menciona un caso en el que se respalda la relevancia y utilidad de este atributo.

En un horario de 5:20am-8:00am dividido en 8 frames (intervalos) de 20 minutos la AS registra una sumatoria total de 7 y el valor en CF es igual a 2, por lo que existen dos opciones:

1. La conducta fue detectada dentro de los intervalos 2,5 (2+5) es decir en un horario de 5:40-5:00 am y 6:40-7:00 am
2. La conducta fue detectada dentro de los intervalos 1,6 (1+6) es decir en un

horario de 5:20-5:40 am y 7:00-7:20 am

Con dos opciones a elegir, el atributo MIN permitiría facilitar la elección al brindar el menor valor de frame en el que la conducta fue ejercida por el individuo, es decir el número de intervalo en el que se registró por primera vez la conducta dentro del horario analizado, para el ejemplo anterior en caso de obtener un valor de 2 en el atributo MIN se optaría por registrar la opción 1 y en el caso que el valor sea igual a 1 se seleccionaría la opción 2.

**Valor máximo de frame (MAX).** Este atributo cumple casi el mismo propósito del atributo MIN solo que este registra el valor máximo de frame en el que una conducta tuvo lugar.

**Cantidad de veces en la que la actividad fue detectada (CVT).** Este atributo representa: el número de veces en la que una actividad (conducta) fue detectada dentro de un frame dividido entre 2, se realiza esta operación debido a que, para contabilizar un evento completo, es decir, “ocurrió y dejó de ocurrir” debe haber un inicio y fin del evento. Si por ejemplo la variable CVT del dispositivo localizado en el refrigerador ‘R’ contiene un valor de 4 quiere decir que el evento de abrir y cerrar el refrigerador ocurrió 2 veces ( $4/2 = 2$ ). En caso de existir un número impar de eventos derivados de circunstancias inesperadas como el reset de dispositivos, desconexión de red o cortes de luz, cae en la condición de restar un 1 al número de eventos y dividirlo el resultado entre /2.

Tabla 3.13: Abreviaturas de las actividades monitoreadas por los sensores empleados en la “red 1” área de cuarto

<b>LC</b>	<b>MC</b>	<b>CC</b>	<b>EC</b>
Prender/Apagar luz de entrada de cuarto	Movimiento en el área del cuarto	Presencia en cama	Uso del equipo de computo

**Nota.** La “red 1” involucra la información de los dispositivos almacenada en el “csv\_A”

Tabla 3.14: Abreviaturas de las actividades monitoreadas por los sensores empleados en la “red 2” áreas de cocina sala y baño

<b>LE</b>	<b>M</b>	<b>R</b>	<b>E</b>	<b>G</b>
Prender/Apagar luz de entrada	Movimiento en el área de sala	Abrir/Cerrar refrigerador	Acercamiento a estufa	Abrir/Cerrar gaveta
<b>UT</b>	<b>UL</b>	<b>B</b>	<b>BH</b>	
Movimiento en mesa	Prender/Apagar luz de cocina	Movimiento en el área de baño	Temperatura y humedad en el área de baño	

**Nota.** La “red 2” involucra la información de los dispositivos almacenada en el “csv\_B”

Tabla 3.15: Abreviaturas de las actividades monitoreadas en el entorno con sus respectivos ID

<b>Red 2 “csv_A”</b>	<b>ID</b>	<b>Red 1 “csv_B”</b>	<b>ID</b>
LE	0	LC	0
M	1	MC	1
R	2	CC	2
E	3	EC	3
G	4		
UT	5		
UL	6		
B	7		
BH	8		

### Ejemplo final de sistema trabajando con los datos del archivo “csv\_A”

La cronología siguiente, se obtiene en base a los datos obtenidos del patrón de conducta ejercido por el individuo durante un corto periodo de tiempo, señalado en la Fig. 3.41.

“El día 26 de octubre del 2020 en el periodo de horario de 5:00 am- 8:00 am el individuo no ejerció actividad en el entorno hasta la segunda hora donde a la hora de 6:49 am encendió la luz localizada en la sala (sensor ID,0) y procedió a desplazarse en el entorno hasta la cocina (sensor ID,1) , realizó acercamientos a





### 3.14.2.1. Procesamiento del archivo CSV correspondiente al uso de la app móvil “Daylio”

Como se mencionó en la Fig.3.31 la aplicación móvil “Daylio” permitió obtener información referente al estado de humor, actividades, sueño, metas y notas diarias a través del uso de 8 atributos.

Como puede observarse en la Fig. 3.45 la variable “activities” contiene una serie de datos que parecen no tener relación entre sí, esto es debido al formato en el que la plataforma brinda la información y debido a esto fue necesario realizar el siguiente procedimiento:

1. A través de un análisis entre las variables presentadas se realizó
2. Una serie de categorías que permitían catalogar la información en conjuntos similares entre sí, finalmente las categorías creadas fueron *Emoción*, *Social*, *Metas (Goals)* y *Sueño*.
3. Una vez realizado el seccionamiento de datos a través de las categorías creadas, fue necesario transformar la información textual a valores numéricos con la finalidad de facilitar su manipulación y poder generar una ponderación que apoye la creación de patrones entre las variables. La ponderación consistió en la asignación de valores altos a eventos que implicaran una circunstancia inusual contrario a los elementos comunes que obtenían una ponderación con valores menores. La Fig. 3.43 presenta el total de las categorías con sus respectivas ponderaciones entre elementos.
4. Finalmente, una vez obtenida la ponderación de elementos por categoría, se actualizaron los valores por renglón (día) en el archivo CSV, La Fig. 3.44 muestra parte del contenido final del archivo correspondiente a la información brindada a partir del uso diario por parte del individuo de la aplicación móvil.

Figura 3.43: Ponderación de las categorías obtenidas del uso diario de la App “Daylio”

<i>Emoción</i>	<i>Sueño</i>	<i>Goals</i>	<i>Social</i>
2 Bien	1 Buen sueño	1 Beber agua	2 Familia
4 Enojado	2 Sueño medio	1 Comer sano	1 Amigos
4 Estresado		1 Lectura	1 Escuchar
3 Increíble		1 Andar	1 Cita
1 Meh			1 Donar
4 Nostálgico			1 Amabilidad
			2 Fiesta

Figura 3.44: Ejemplo parcial del archivo csv\_Daylio

1	Date	weekday		Emoción	Sueño	Goals	Social
2	12 de diciembre	sábado	8:30 p. m.	2	2	1	1
3	11 de diciembre	viernes	11:55 p. m.	3	1	1	1
4	10 de diciembre	jueves	11:58 p. m.	1	2	0	1
5	9 de diciembre	miércoles	8:01 p. m.	2	1	1	1

### 3.14.2.2. Procesamiento del archivo CSV correspondiente al uso del smartwatch “Garmin”

Gracias a las herramientas de control incorporadas al reloj inteligente fue posible obtener la información correspondiente a Frecuencia cardíaca (FC), Nivel de estrés (Estrés) y Horas de sueño (HS) correspondientes al estado físico del individuo, el tratamiento de la información consistió en la ejecución de los siguientes pasos:

1. Descargar a través de la plataforma del dispositivo el archivo de la variable correspondiente a trabajar, en la Fig. 3.45 se presenta una muestra de cada documentación final descargada.

Figura 3.45: Información parcial de los csv correspondientes a estrés, FC y hrs de sueño

Lun	13	Lun	50	Lun	8.2	8:10 h
Mar	15	Mar	50	Mar	8.3	8:19 h
Mié	18	Mié	53	Mié	7.3	7:19 h
Jue	7	Jue	54	Jue	8	7:57 h
Vie	65	Vie	58	Sáb.	7.8	7:47 h
Sáb.	39	Sáb.	54			
	NOV		NOV			
Dom	31	Dom	54	Dom	8	8:00 h
Lun	30	Lun	55	Lun	7.1	7:03 h
Mar	40	Mar	53	Mié	9.4	9:24 h
Mié	19	Mié	53	Jue	7.6	7:34 h
Jue	27	Jue	53	Vie	7.3	7:20 h
Vie	38	Vie	56			

**Nota.** La información presentada corresponde a las fechas del 26 Octubre al 6 Noviembre 2020

2. Incorporar los tres documentos en un solo archivo CSV para su posterior incorporación a los otros medios de información.

Una vez obtenidos los archivos correspondientes a la información obtenida del uso diario de la aplicación móvil “Daylio” y el smartwatch “Garmin”, fue necesario realizar la unión entre las 3 documentaciones 1) csv\_patrones de conducta 2) csv\_appMovil 3) csv\_smartwatch; para ello se requirió incorporar los valores de los dos archivos al documento “csv\_patrones de conducta” revisando la concordancia entre los días y valores correspondientes, además fue necesario adaptar los datos de los archivos “csv\_appMovil” y “csv\_smartwatch” a los horarios obtenidos a través del procesamiento de la información de red IoT todo esto con el propósito de evitar el registro de espacios vacíos o valores ‘0’ en el conjunto de datos final evitando de esta manera un desbalance causado por el excedente de ceros.

### 3.14.3. Reducción de atributos

Debido a la gran cantidad de atributos que los tres datasets formaban en conjunto, fue necesario el uso de métodos para la reducción de columnas (features) en el archivo CSV, por ello en conjunto con la técnica Principal Component Analysis (PCA) y el conocimiento experto de los datos hasta ese punto se realizó la reducción de dimensionalidad en el conjunto

de datos apoyándonos siempre de la matriz de correlación como herramienta para en análisis y selección de variables con base en el coeficiente de correlación entre sí. La Fig. 3.46 muestra un ejemplo del funcionamiento de una matriz de correlación.

Figura 3.46: Ejemplo de matriz de correlación entre 3 variables X, Y, Z

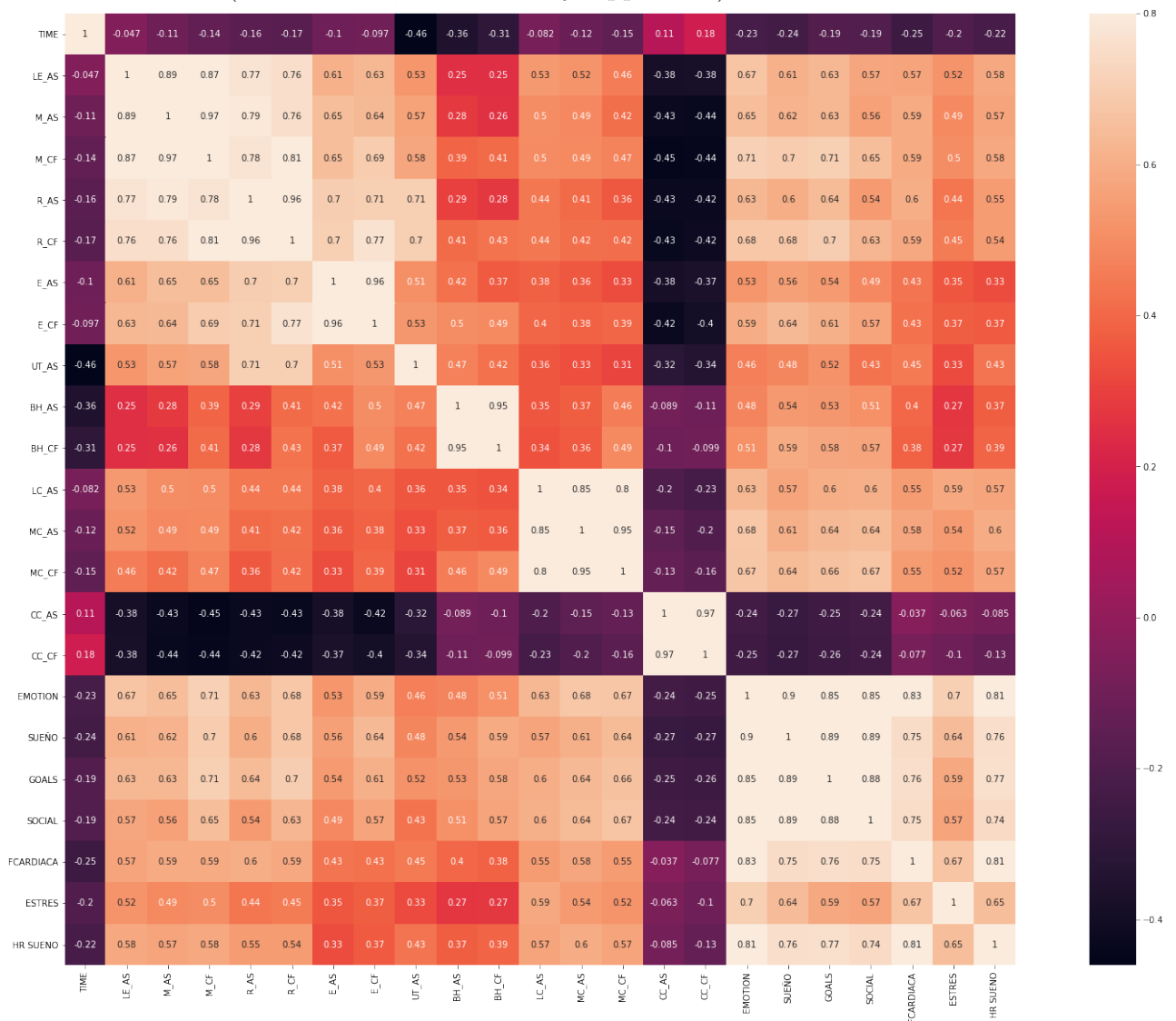
	<b>X</b>	<b>Y</b>	<b>Z</b>
<b>X</b>	$r_{xx}$	$r_{xy}$	$r_{xz}$
<b>Y</b>	$r_{yx}$	$r_{yy}$	$r_{yz}$
<b>Z</b>	$r_{zx}$	$r_{zy}$	$r_{zz}$

### 3.14.3.1. Análisis de la matriz de correlación obtenida.

En base al conjunto de datos finales obtenidos a partir de la reducción de dimensionalidad, se obtuvo la expresión de los conjuntos de coeficientes de correlación entre variables mediante la matriz presentada en la Fig 3.47, la cual expresa una alta correlación entre los dispositivos ubicados en las mismas áreas (Sala, Cocina, Baño y Cuarto ) y menor correlación mas no nula entre los elementos localizados en áreas separadas y una correlación muy negativa respecto a los features “CC\_AS” y “CC\_CF” los cuales corresponden al sensor colocado en la cama para detectar la presencia en esta, por lo que se interpreta como “si está en la cama definitivamente no puede estar en alguna otra área”.

Dentro de la matriz de correlación también se puede observar que las variables correspondientes a las tecnologías de monitoreo (smartwatch y App móvil) presentan un alto coeficiente de correlación entre sí algo que resulta interesante, ya que no existe dependencia de ninguna hacia la otra para su funcionamiento.

Figura 3.47: Matriz de correlación obtenida a partir de la unión y reducción de atributos entre los datasets (Entorno IoT, smartwatch y App móvil)



### 3.14.4. Desarrollo del modelo de ML encargado de la detección de anomalías en los patrones de conducta del individuo

Como hemos visto al largo de este trabajo, una anomalía es un punto dentro de un conjunto de datos que destaca por ser o comportarse diferente al resto, si bien los datos anómalos pueden afectar el resultado del entrenamiento en un modelo de ML también pueden brindar gran información de la información que se trabaja. Para este estudio los patrones de conducta ejercidos por el individuo fueron registrado en un archivo CSV que permitió trabajar toda la información en datos de entrada para el modelo de ML encargado de detectar anomalías dentro de los patrones establecidos, en esta sección se abordará el diseño y desarrollo implicado en su construcción.

#### 3.14.4.1. Normalización del dataset conjunto

Detectar un elemento diferente al resto en un conjunto de datos, implica plantearse la pregunta “¿qué es lo normal?” o “¿cuál es el comportamiento normal de los datos en este conjunto?”, debido a que para localizar un punto diferente al resto se debe conocer un comportamiento “base” de la información. Para el caso del archivo CSV empleado se requirió un preprocesamiento a la unión de los tres datasets procedentes de las diferentes tecnologías empleadas con la finalidad de poder trabajar bajo un formato que permitiera tratar toda la información sin afectar el resultado de entrenamiento debido a la procedencia de los datos, es decir los valores que cada tecnología aportaba al resto debían mantenerse en una escala en común para poder ser ingresados en conjunto al modelo de ML desarrollado.

Es importante considerar que ningún método de normalización beneficia a todos los datasets por igual debido a que cada uno cuenta con un comportamiento diferente de datos, cuestiones como datos faltantes, limitado número de muestras, puntos anómalos o distribución de los datos influye en el desempeño y beneficio de cada método por ello primeramente se decidió trabajar manualmente un proceso de limpieza de datos,

#### **Limpieza de dataset**



datos empleando el método “normalize()” el cual por medio de la selección de parámetros (norm: ‘l1’, ‘l2’, ‘max’) permitió tomar a consideración la presencia de datos anómalos como información relevante para el entrenamiento del modelo de ML.

#### **3.14.4.2. Selección de algoritmos para la detección de anomalías dentro de los patrones de conducta establecidos**

Este diseño y desarrollo experimental solo se enfocó en la selección de algoritmos capaces de detectar tuplas anómalas y apoyar si los datos de entrada podían ser catalogados en agrupaciones normales y no normales, para ello fue necesario el uso de las herramientas de Python Scikit-learn y Pyod [107] dos conjuntos que permiten aplicar algoritmos para el aprendizaje no supervisado donde la diferencia entre estos radica en que Pyod está enfocado directamente al campo de detección de valores atípicos o datos anómalos lo que pueda brindar una metodología más desarrollada para su enfoque.

A partir del uso de cada algoritmo se obtuvo el número de tuplas anómalas presente en el dataset brindado, la agrupación correspondiente a cada punto del conjunto de datos y por último la serie de tuplas que por medio de un puntaje obtenido por los algoritmos se catalogaron como atípicas, estas tuplas representan el patrón de conducta ejercido por el individuo en un determinado intervalo de tiempo u horario.

A continuación, la Tabla 3.16 presenta los algoritmos empleados por las librerías Scikit-learn y Pyod, así como la información mencionada.



### 3.14.4.3. Evaluación del desempeño de los modelos empleados en el diseño experimental

Para la evaluación del desempeño de cada algoritmo se emplearon las métricas silhouette, davies bouldin, calinski\_harabaz y Dunn detalladas en el capítulo 2 “Evaluación de un algoritmo de detección de anomalías”, a continuación la Fig.3.49 “Evaluación de los algoritmos empleados”, presenta los resultados obtenidos de cada modelo empleado. Donde podemos observar que entre los mejores se encuentra COPOD (), una observación importante es que el modelo basado en densidad de datos DBSCAN no resulta conveniente para el entrenamiento de este dataset debido a que las agrupaciones realizadas son afectadas por las aglomeraciones de ceros incluidas en el conjunto.

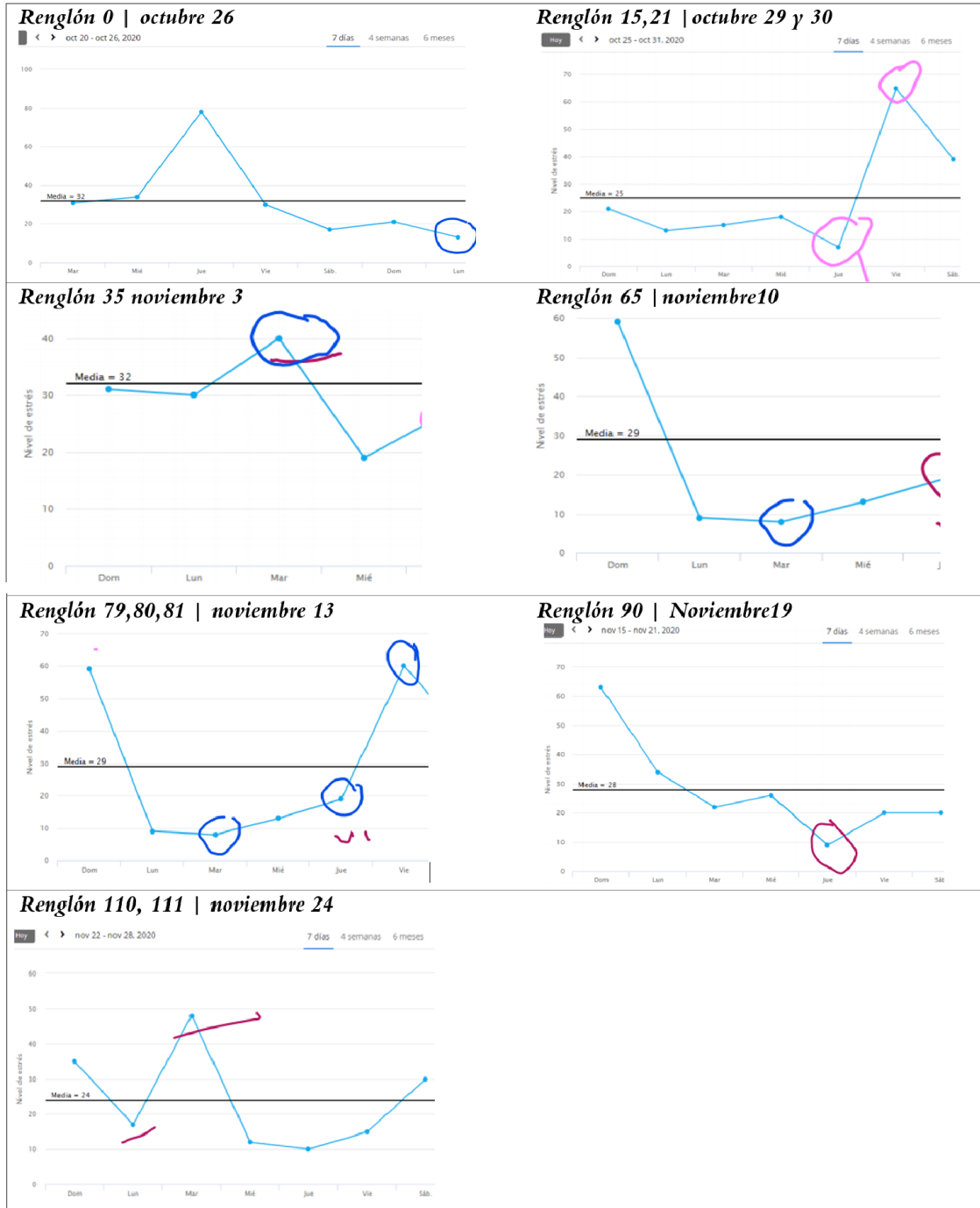
Figura 3.49: Evaluación de los algoritmos empleados

Pyod	Scikit-learn
Model COPOD()	Model K-MEANS()
silhoutte: 0.31689279456842295 davies bouldin: 1.4089143009838219 calinski_harabaz: 34.471837941533565 Dunn: 0.22810325776201795	silhoutte: 0.37881530166294985 davies bouldin: 1.136810731788516 calinski_harabaz: 110.83077599173157 Dunn: 0.26832679193514875
Model KNN PYOD()	Model DB SCAN()
silhoutte: 0.2465364415585439 davies bouldin: 2.0312049399830765 calinski_harabaz: 17.527495971386788 Dunn: 0.22857116172164915	silhoutte: 0.3336299634838019 davies bouldin: 1.5334022211831546 calinski_harabaz: 32.33573329289856 Dunn: 0.31679250109099694
Model ISOLATION FOREST()	Model ISOLATION FOREST()
silhoutte: 0.30284686178525655 davies bouldin: 1.655096585392441 calinski_harabaz: 27.00342012388216 Dunn: 0.2263117645974124	silhoutte: 0.30728049611202163 davies bouldin: 1.5598856897357718 calinski_harabaz: 29.473816808644244 Dunn: 0.23494084335220825

**3.14.4.4. Resultados obtenidos del diseño experimental**

En base en los resultados obtenidos se realizó el análisis de los días y momentos correspondientes a las tuplas señaladas por la mayoría de los modelos para poder detectar los elementos que destacaron del resto, todo esto empleando una bitácora que a lo largo de casi todo el estudio fue llenada por el experto de acuerdo a las actividades sobresalientes realizadas por el individuo entre las que destacaron las realizadas fuera de horarios establecidos y las horas de sueño que ejercía el individuo, los resultados son presentados en la siguiente Fig. 3.50.

Figura 3.50: Resultados sobresalientes del atributo Hrs.Sueño dentro de la tuplas seleccionada por los modelos



## 3.15. Diseño Final

El análisis y desarrollo del diseño final para la construcción del modelo final de ML empleado en la detección de conductas anómalas dentro de los patrones de conducta establecidos se basó en el procedimiento y análisis del diseño experimental empleado. En base a los resultados obtenidos se modificaron algunos parámetros y procedimientos con la finalidad de obtener un mejor desempeño durante el entrenamiento del dataset mejorando la calidad de los datos de entrada y la construcción del modelo final de ML.

A partir de esta sección se presentaron cambios drásticos del diseño experimental respecto al análisis de los datos de entrada al modelo de ML, preprocesamiento y tratamiento para lograr la detección de puntos anómalos en el conjunto total de información recopilada al emplear las tres tecnologías principales en este estudio piloto (Entorno IoT, smartwatch Garmin y la App móvil Daylio).

Como se presentó al inicio en la sección “Recolección los conjuntos de datos generados por la red IoT”, la reunión de toda la información obtenida por la red IoT a través de cada dispositivo fue el primer paso para continuar con el procedimiento en la formación de patrones de conducta ejercidos por el individuo, donde una vez resueltas las incógnitas de diseño y desarrollo del modelo de preprocesamiento capaz de transformar los datos proporcionados en patrones de conducta, consultar Tabla 3.11; se procedió a la selección de los intervalos de tiempo para la creación de la cronología de actividades.

### 3.15.1. Establecer patrones

Al igual que en el diseño experimental, fue necesario establecer intervalos de fecha y hora para obtener el conjunto actividades realizadas dentro del intervalo de tiempo, el procedimiento se repitió en ambos archivos (“csv\_A” y “csv\_B”) obtenidos de la “Unión de los conjuntos de datos generados por la red IoT”.

Pese a que el procedimiento se mantuvo igual al empleado en el diseño experimental la selección de diferentes intervalos de tiempo fue requerida con la finalidad de obtener una

mayor cantidad de tuplas y la reducción de tiempo en los intervalos para un análisis más detallado en las actividades realizadas, a continuación, se presenta el planteamiento de cada uno con el propósito de remarcar los cambios realizados.

### 3.15.1.1. Detección de las conductas ejercidas por el individuo dentro de los intervalos de tiempo seleccionados

- **Diseño experimental.** Para los archivos “csv\_A” y “csv\_B” se eligieron inicialmente los horarios de 17:00pm, 20:00 pm,23:00pm, 2:00am, 5:00am y 8:00 am, todos con 3 horas de diferencia entre sí.

Ambos con un análisis programado cada hora en el que al finalizar el análisis de las 3 horas entre cada uno de los 5 lapsos de tiempo por día (7 renglones del conjunto de datos final) y 47 días de estudio, conformarían el conjunto de datos final de 235 tuplas(renglones) para la detección de anomalías.

- **Diseño actual.** La selección de estos horarios tuvo en cuenta que un intervalo muy reducido para el análisis de los lapsos, al igual que una gran cantidad de intervalos de tiempo por día, generaban una gran cantidad tuplas que en conjunto con las tuplas adyacentes creaban una gran aglomeración de valores ‘0’ por lo que se eligieron los horarios para los archivos “csv\_A” y “csv\_B”

1. 5:30-6:30 am
2. 6:30-7:30 am
3. 17:30-18:30 pm
4. 18:30-19:30 pm
5. 19:30-20:30 pm
6. 20:30-21:30 pm
7. 21:30-22:30 pm

Ambos con un análisis programado cada media hora en el que al finalizar el análisis de la hora entre cada uno de los 7 intervalos de tiempo por día (7 renglones del conjunto de datos final) y 47 días de estudio, conformarían el conjunto de datos final de 329 tuplas(renglones) para la detección de anomalías. A continuación, se explica el procedimiento en la Tabla 3.51 bajo los intervalos de tiempo establecidos ” y la Tabla 3.52 con el uso del archivo “csv\_B”.

Figura 3.51: Ejemplo final del sistema trabajando en la detección de conductas registradas en el archivo “csv\_A” bajo los intervalos de tiempo establecidos

<i>Horario</i>	<i>Conductas por intervalos de tiempo</i>			<i>Elementos involucrados</i>	<i>Conductas deducidas en base a los sensores activados</i>
<u>Intervalo:</u> 30 minutos	EVENTO	FECHA	ID	Entrada de Sala 0,1	Desplazamiento en el entorno de sala
	0	2020-10-26 06:49:00	0		
	1	2020-10-26 06:50:00	1		
	0	2020-10-26 06:51:00	1		
	1	2020-10-26 06:56:00	3		
Hora inicial: 6:30	0	2020-10-26 06:57:00	3		
6:30-7:00 = 30 min	order is : [0, 1, 3]			Refrigerador 2	Cocinar
Y	EVENTO	FECHA	ID		
	1	2020-10-26 07:09:00	1		
	0	2020-10-26 07:10:00	1	Estufa 3	
	1	2020-10-26 07:11:00	2		
7:00-7:30 = 30 min	0	2020-10-26 07:11:00	2		
	1	2020-10-26 07:11:00	1		
	1	2020-10-26 07:13:00	0		
	0	2020-10-26 07:13:00	1		
	order is : [1, 2, 0]				
<u>Intervalo:</u> 30 minutos	EVENTO	FECHA	ID	Refrigerador 2	Cocinar
	1	2020-10-26 18:34:00	2		
	0	2020-10-26 18:35:00	2		
	1	2020-10-26 18:36:00	3		
	0	2020-10-26 18:37:00	3		
	1	2020-10-26 18:40:00	2	Estufa 3	
Hora inicial: 18:30	0	2020-10-26 18:40:00	2		
	1	2020-10-26 18:43:00	2		
18:30-19:00 pm	1	2020-10-26 18:44:00	3		
	0	2020-10-26 18:45:00	3		
	1	2020-10-26 18:48:00	3	Mesa 5	Comer
Y	1	2020-10-26 18:51:00	2		
	0	2020-10-26 18:51:00	2		
	1	2020-10-26 18:52:00	2		
19:00-19:30 pm	0	2020-10-26 18:53:00	2		
	1	2020-10-26 18:54:00	2		
	0	2020-10-26 18:54:00	2		
	1	2020-10-26 18:58:00	3		
	0	2020-10-26 18:59:00	3		
	order is : [2, 3]				
	EVENTO	FECHA	ID		
	1	2020-10-26 19:07:00	3		
	0	2020-10-26 19:07:00	3		
	1	2020-10-26 19:08:00	5		
	0	2020-10-26 19:09:00	5		
	1	2020-10-26 19:10:00	5		
	0	2020-10-26 19:11:00	5		
	1	2020-10-26 19:12:00	5		
	0	2020-10-26 19:12:00	5		
	1	2020-10-26 19:14:00	5		
	0	2020-10-26 19:15:00	5		
	1	2020-10-26 19:21:00	5		
	0	2020-10-26 19:21:00	5		
	order is : [3, 5]				

**Nota.** En este archivo “csv\_A” se encuentra la información correspondiente a las actividades (conductas) desarrolladas en la sala, cocina y baño.

Figura 3.52: Ejemplo final del sistema trabajando en la detección de conductas registrada en el archivo “csv\_B” bajo los intervalos de tiempo establecidos

<b>Horario</b>	<b>Conductas por intervalos de tiempo</b>	<b>Elementos involucrados</b>	<b>Conductas deducidas en base a los sensores activados</b>
<u>Intervalo:</u> 30 minutos	== DataFrame that had activation== SCALE 1 2 3 4 5 6	Entrada de cuarto	Desplazamiento en el cuarto
Hora inicial 6:30 am	<pre>                     Unnamed: 0  EVENTO          FECHA  ID 50          148          1  2020-10-27 06:37:00  2 51          149          0  2020-10-27 06:37:00  2 52          195          0  2020-10-27 06:44:00  0 53          194          1  2020-10-27 06:44:00  1 54          150          1  2020-10-27 06:44:00  2 55          151          0  2020-10-27 06:44:00  2 56          196          0  2020-10-27 06:45:00  1 57          197          1  2020-10-27 06:47:00  1 58          198          0  2020-10-27 06:49:00  1 59          199          1  2020-10-27 06:52:00  1 60          200          0  2020-10-27 06:52:00  1 61          201          1  2020-10-27 06:53:00  1 62          202          0  2020-10-27 06:54:00  1 63          203          0  2020-10-27 06:55:00  0                     </pre>	0,1	
6:30-7:00 am		Cama 2	Acostado
Y			
7:00-7:30 am	<pre> DateFrame: 1 order is : [2, 0, 1] miraaaa 2020-10-27 07:00:00                     Unnamed: 0  EVENTO          FECHA  ID 65          206          1  2020-10-27 07:05:00  1 67          207          0  2020-10-27 07:06:00  1                     </pre> <p>DateFrame: 2 order is : [1]</p>		
<u>Intervalo:</u> 30 minutos	== DataFrame that had activation== SCALE 1 2 3 4 5 6	Entrada de cuarto	Desplazamiento en el cuarto
Hora inicial: 21:30	<pre>                     Unnamed: 0  EVENTO          FECHA  ID 73          214          1  2020-10-27 21:35:00  1 74          215          0  2020-10-27 21:35:00  1 75          216          1  2020-10-27 21:36:00  1 76          217          0  2020-10-27 21:37:00  1 77          218          1  2020-10-27 21:38:00  1 78          219          0  2020-10-27 21:38:00  1 79          152          1  2020-10-27 21:44:00  2 80          209          1  2020-10-27 21:44:00  0 81          154          1  2020-10-27 21:45:00  2 83          155          0  2020-10-27 21:46:00  2 84          156          1  2020-10-27 21:58:00  2 85          157          0  2020-10-27 21:58:00  2 86          158          1  2020-10-27 21:59:00  2 87          159          0  2020-10-27 21:59:00  2                     </pre>	0,1	
21:30-22:00 pm		Cama 2	Acostado
Y			
22:00-22:30 pm	<pre> DateFrame: 1 order is : [1, 2, 0] miraaaa 2020-10-27 22:00:00                     Unnamed: 0  EVENTO          FECHA  ID 88          160          1  2020-10-27 22:02:00  2 89          161          0  2020-10-27 22:02:00  2 90          162          1  2020-10-27 22:03:00  2 91          163          0  2020-10-27 22:03:00  2 92          164          1  2020-10-27 22:05:00  2 93          165          0  2020-10-27 22:06:00  2 94          166          1  2020-10-27 22:19:00  2 95          167          0  2020-10-27 22:19:00  2 96          168          1  2020-10-27 22:20:00  2 97          169          0  2020-10-27 22:20:00  2 98          170          1  2020-10-27 22:23:00  2 99          171          0  2020-10-27 22:24:00  2                     </pre> <p>DateFrame: 2 order is : [2]</p>		

**Nota.** En este archivo “csv\_B” se encuentra la información correspondiente a las actividades (conductas) desarrolladas en la habitación del individuo

### **Desarrollo cronológico final de los patrones de conducta ejercidos diariamente empleando el uso de atributos para la sinterización de información**

El uso de los 5 atributos AS, CF, MIN, MAX, CVT (Escala de anomalía, Cantidad de frames, Valor mínimo de frame, Valor máximo de frame, Frecuencia) fue necesario al igual que en el diseño experimental para el registro de los patrones de conducta ejercidos por el individuo a lo largo del día, proceso que fue aplicado ante los datos de entrada del archivo “cvs\_A” y el archivo “csv\_B”.

A continuación, se presenta el desarrollo cronológico de algunos patrones de conducta obtenidos a través de los intervalos de tiempo seleccionados para este diseño final y el análisis de las acciones detectadas dentro de estos (Tabla 3.53 y 3.54). Recordemos que contamos con una selección de 7 lapsos equivalentes a 7 horarios diferentes con un análisis programado de media hora en cada uno, por lo tanto, el registro de las actividades ejercidas en cada lapso será en base a la revisión y análisis programado cada 30 minutos (2 veces en una hora).

#### **Ejemplo final de sistema trabajando con los datos de entrada del archivo “csv\_A”**

La cronología siguiente, se obtiene en base a los datos obtenidos del patrón de conducta ejercido por el individuo durante un corto periodo de tiempo, señalado en la Fig 3.53 “El día 27 de octubre del 2020 alrededor de las 19:30 horas el individuo se encontraba desplazándose (sensor ID,1) por el área de la cocina, donde a las 19:34 abrió dos veces el refrigerador (sensor ID,2) y durante parte de la primera y segunda media hora estuvo bajo el dispositivo colocado sobre la mesa, por lo tanto, podemos acordar que la secuencia se deriva de un lapso dedicado a la alimentación del individuo”.





solo dataset para incorporar el resto de la información obtenida de las otras dos tecnologías empleadas “smartwatch Garmin” y “App móvil Daylio”.

### 3.15.2. Unión de los conjunto de datos obtenidos de las tecnologías empleadas

El dataset final consistió en la unión de los dos archivos obtenidos a partir de la detección de patrones previamente abordada y la incorporación de la información procedente del uso diario de las tecnologías empleadas “smartwatch Garmin” y “App móvil Daylio” por parte del individuo para el monitoreo de sus actividades fuera del entorno IoT y su rendimiento a lo largo del día Fig 3.55, lo que implicó una recolección y tratamiento de la información obtenida de estas dos tecnologías explicadas previamente en la sección 3.14 del diseño experimental.

Figura 3.55: Visualización parcial de la unión de conjuntos

LE_AS	M_CF	R_AS	R_CVT	E_AS	UT_AS	BH_AS	LC_AS	MC_CF	CC_AS	SOCIAL	GOALS	ESTRES	CARDIACA	HRS_SUEN	TIME	
0	0	0	0	0	0	0	0	0	0	2	1	1	1.8	5.3	8.3	1
3	2	0	0	0	0	0	0	1	2	0	1	1	1.8	5.3	8.3	2
2	1	0	0	0	2	0	0	0	0	0	1	1	1.8	5.3	8.3	3
0	0	3	6	0	3	0	0	0	0	0	1	1	1.8	5.3	8.3	4
2	1	0	0	0	1	2	0	0	0	0	3	2	1.8	5.3	8.3	5
0	0	0	0	0	0	0	1	1	2	0	2	2	1.8	5.3	8.3	6
0	0	0	0	0	0	0	0	1	1	0	1	1	1.8	5.3	8.3	7

A pesar de que la unión de conjuntos siguió el mismo procedimiento y desarrollo del diseño experimental, se tuvo en cuenta que los cambios realizados en los intervalos de tiempo para la detección de las actividades ejercidas por el individuo dentro de estos influirían en el dataset de los patrones de conducta final ya que más tuplas representarían 1 solo día por ello se requirió una minuciosa adaptación de los dataset obtenidos del smartwatch Garmin y la App móvil Daylio para que existiera una concordancia entre fechas y horarios y finalmente obtener un único conjunto de datos que incorporó la información procesada y tratada de las tres tecnologías empleadas en este estudio piloto (Entorno IoT, Garmin y la App móvil Daylio ).

### 3.15.3. Reducción de atributos

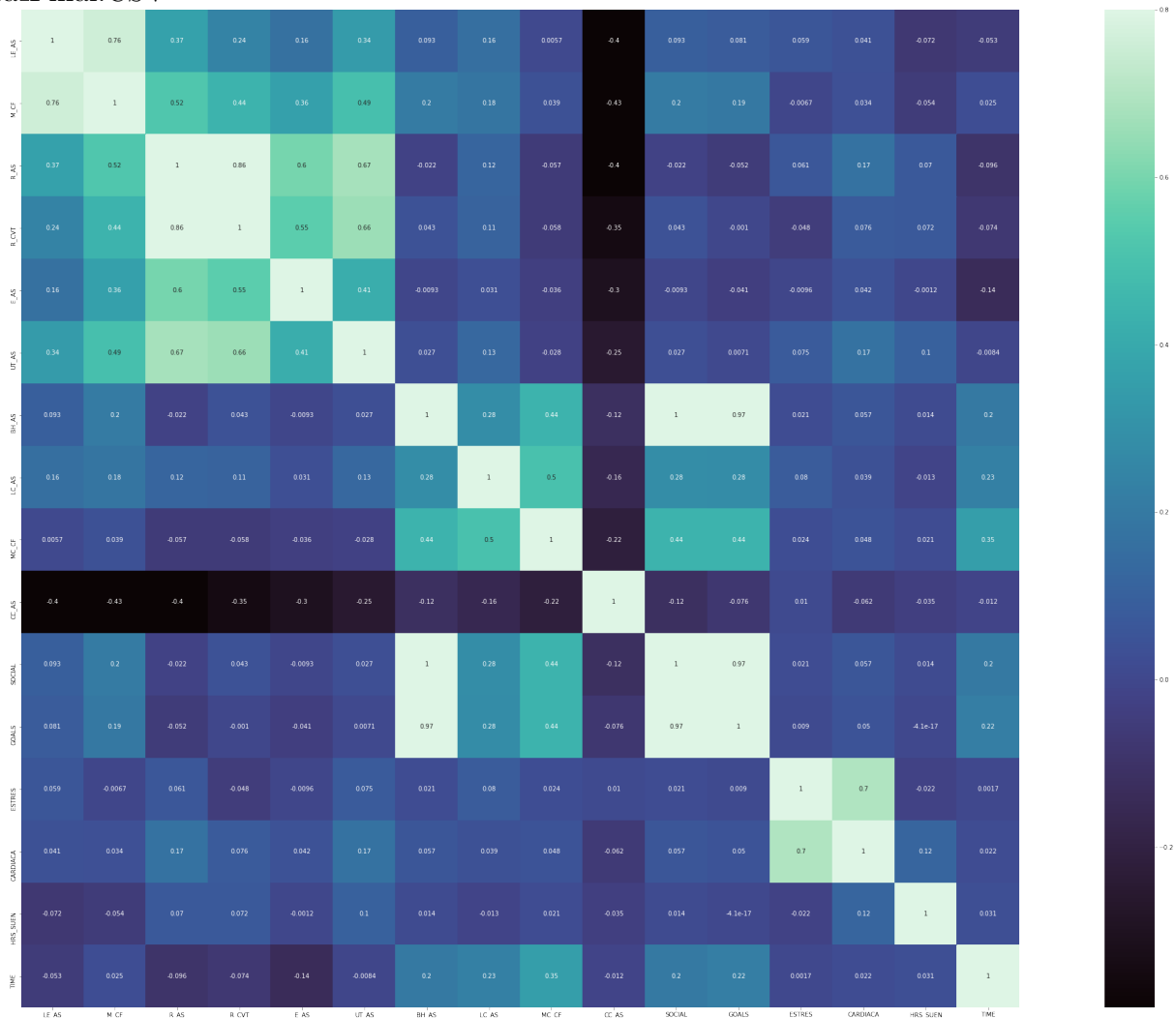
Siguiendo el patrón de análisis del diseño experimental el cual consistió en emplear la técnica PCA a la par del conocimiento experto sobre los dataset del Entorno IoT, Garmin y la App móvil Daylio en conjunto para la reducción de atributos fue posible obtener un conjunto final llamado “patronesConducta.Final.CSV” el cual contenía las variables independientes más relevantes para este estudio piloto las cuales se presentan en la Fig.3.56 con su respectivo orden.

Figura 3.56: Atributos seleccionados para el análisis y detección de conductas anómalas por parte del modelo de ML.

LE_AS	M_CF	R_AS	R_CVT	E_AS	UT_AS	BH_AS	LC_AS	MC_CF	CC_AS	SOCIAL	GOALS	ESTRES	CARDIACA	HRS_SUEN	TIME
-------	------	------	-------	------	-------	-------	-------	-------	-------	--------	-------	--------	----------	----------	------

A partir de estos atributos se obtuvo la matriz de correlación presentada en la Fig. 3.57 la cual muestra una gran correlación positiva entre las actividades realizadas en el área de la cocina y también entre el baño y la habitación personal del individuo, por otro lado, los elementos involucrados en los atributos “Goals” y “Social” también muestran una gran correlación entre si al igual que “Estrés” y “Frecuencia cardiaca”. Por último, resulta interesante la relación de las variables “Social” y “Goals” con el atributo “BH\_AS” correspondiente a la acción de bañarse

Figura 3.57: Matriz de correlación sobre los features del archivo “patronesConducta.Final.CSV”



### 3.15.4. Desarrollo Final del modelo de ML encargado del estudio y análisis de los patrones de conducta del individuo

En base a los resultados obtenidos del modelo desarrollado y aplicado del diseño experimental se llegó a la conclusión de que era necesario conocer más de la relación y comportamiento entre los mismos datos para no limitarse con el uso de un solo modelo de ML para la detección de anomalías sin un contexto que explicara a mayor profundidad el por qué una conducta o variable resultaba diferente al resto de los datos que conformaban el patrón

de conductas del sujeto de estudio, para ello nuevamente se inició por la normalización de datos que involucró la limpieza del dataset nombrado “conjuntoConductas\_Final” el cual contenía la información recopilada de las tecnologías empleadas (Entorno IoT, smartwatch Garmin y App móvil Daylio) .

### 3.15.4.1. Normalización del dataset “conjuntoConductas\_Final”

Previo a la aplicación de la técnica de normalización de datos fue necesario realizar una limpieza al conjunto de datos, para este diseño final se conservó la información procedente a los fines de semana y se optó por eliminar únicamente las tuplas que no contenían suficiente información a analizar, es decir se eliminaron las tuplas que registraban únicamente valores ‘0’ lo que evitó un desbalance radical de datos que influyera en una incorrecta predicción por parte del modelo de ML trabajado, lo que provocó una reducción de muestras para el entrenamiento pasando de 319 a 197 tuplas. Por otra parte, fue necesario la adaptación de la información obtenida de las otras dos tecnologías empleadas (smartwatch Garmin y App móvil Daylio) para que existiera una concordancia en fechas e intervalos de tiempo aun después de la limpieza de tuplas (Fig 3.58). Una vez obtenido un orden y limpieza dentro del conjunto final fue posible aplicar la normalización de datos.

Figura 3.58: Dataset “conjuntoConductas\_Final” de entrada para el entrenamiento del modelo de ML desarrollado.

LE_AS	M_CF	R_AS	R_CVT	E_AS	UT_AS	BH_AS	LC_AS	MC_CF	CC_AS	SOCIAL	GOALS	ESTRES	CARDIACA	HRS_SUEN	FECHA	TIME
3	2	2	1	1	0	0	0	0	0	1	1	1.3	5	7	26	2
2	1	0	0	0	0	0	0	0	0	1	1	1.3	5	7	26	3
0	0	1	5	3	2	0	2	1	2	1	1	1.3	5	7	26	4
3	2	0	0	0	0	3	0	2	2	4	3	1.3	5	7	26	5
0	0	0	0	0	0	0	0	2	3	1	1	1.3	5	7	26	6
0	0	0	0	0	0	0	0	1	1	1	1	1.3	5	7	26	7
0	0	0	0	0	0	0	0	0	3	1	1	1.5	5	8.2	27	1
3	2	0	0	0	0	0	1	2	1	1	1	1.5	5	8.2	27	2
2	1	0	0	0	0	0	0	0	0	1	1	1.5	5	8.2	27	4

La normalización fue la técnica seleccionada para mantener los valores del dataset bajo una escala en común que permitiera su ingreso en conjunto al modelo de ML desarrollado ya que tomaba en cuenta la importancia de los datos anómalos dentro de la aglomeración de valores evitando su descarte.

#### 3.15.4.2. Agrupación de los patrones de conducta registrados

La agrupación (clustering) de datos involucra una relación o semejanza entre los elementos reunidos, por ello con la finalidad de comprender más el comportamiento de los datos entre sí, se desarrolló un proceso en el que a través del algoritmo K-Means fue posible obtener el número óptimo de agrupaciones que reunía las conductas y variables con una similitud entre sí, a continuación se explicará más a detalle cual fue el proceso de selección de este algoritmo, los beneficios obtenidos de su uso, desventajas y resultados.

##### Selección del algoritmo de clustering

Debido al uso de la herramienta Python para el desarrollo de este modelo se cuenta con una serie de opciones para el entrenamiento bajo el aprendizaje no supervisado por medio del uso de modelos de agrupación entre los que se encuentra DBSCAN, K-Means y Hierarchical clustering. Debido a los resultados de desempeño obtenido del diseño experimental se realizó el descarte automático del modelo DBSCAN debido a su falta de fiabilidad para el tratamiento de este conjunto de datos y se mantuvo en uso del modelo K-Means, no obstante, con la finalidad de no basar la credibilidad de este estudio piloto en un solo modelo se propuso el uso a la par del algoritmo Hierarchical clustering, el cual representa sus agrupaciones por medio de un árbol o dendograma creado a partir de agrupaciones fusionadas o divididas sucesivamente.

A partir del uso de ambos algoritmos se llegó a la conclusión de que 2 era el número óptimo de agrupaciones para el dataset presentado basándose en la gráfica “Método del codo” obtenida en la Fig. 3.59 al igual que el uso de la librería “ValidClust” que se puede apreciar en la Fig.3.60.

Figura 3.59: Método del codo empleado para la selección de agrupaciones

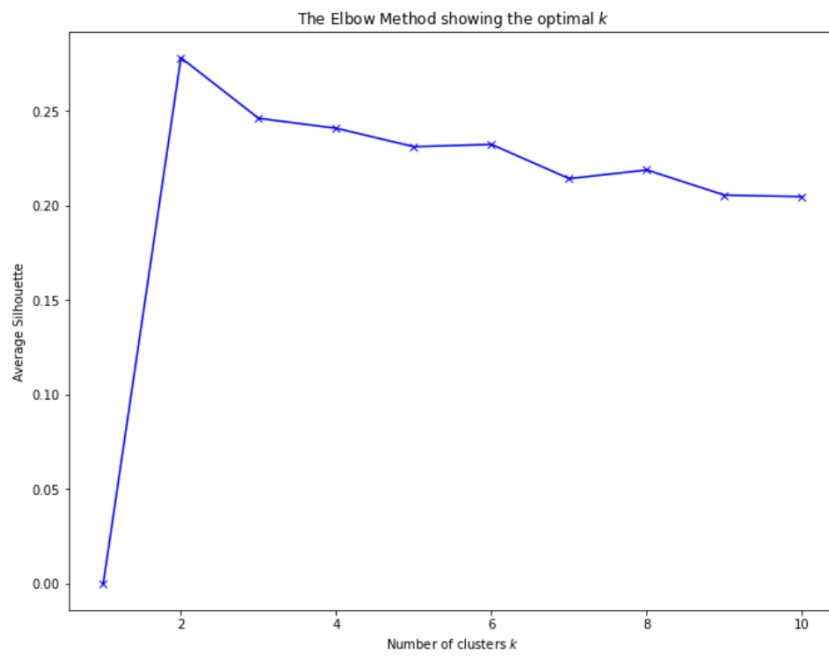
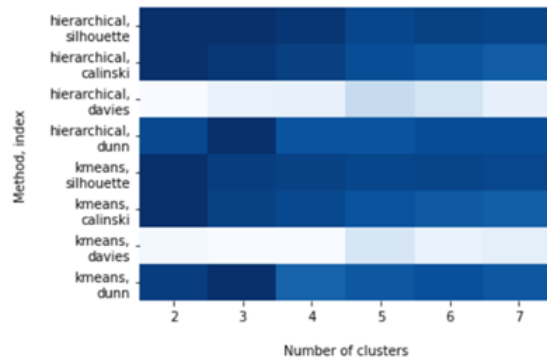


Figura 3.60: Evaluación de los métodos K-Means y Hierarchical clustering

		2	3	4	5 \
method	index				
hierarchical	silhouette	0.242321	0.243696	0.230590	0.202903
	calinski	64.968683	61.205383	57.290118	50.407216
	davies	1.656982	1.562242	1.556809	1.332515
	dunn	0.175382	0.215567	0.157781	0.157781
kmeans	silhouette	0.278130	0.245662	0.240829	0.231073
	calinski	73.135557	63.318323	59.460458	53.719192
	davies	1.539407	1.562347	1.565273	1.381216
	dunn	0.181431	0.203615	0.123841	0.140905

		6	7
method	index		
hierarchical	silhouette	0.212578	0.207164
	calinski	46.438750	42.779016
	davies	1.416075	1.544069
	dunn	0.168809	0.168809
kmeans	silhouette	0.235537	0.229549
	calinski	50.208738	46.467218
	davies	1.488740	1.457854
	dunn	0.154066	0.140905



Una vez seleccionado el número óptimo de agrupaciones se realizó el entrenamiento aplicando ambos modelos con los mismos datos de entrada donde finalmente se obtuvo la asignación propuesta por cada modelo al igual que la evaluación de su desempeño, la Fig 3.61 muestra la división y rendimiento presentada por el algoritmo K-Means y la Figura 3.62 por el algoritmo Hierarchical clustering

Figura 3.61: Agrupación propuesta por el modelo K-means

```
Clustering by row :
[1 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 1 0 0 1 1 1 0 0 0 0 0 1 0 0 1 0 0
1 0 1 0 0 0 0 1 1 1 0 0 0 0 1 0 1 0 0 0 0 1 1 1 1 0 1 0 1 0 0 1 1 0 1 0 1
0 1 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 1 1 0 0 0 0 1
1 0 0 0 0 0 1 0 0 1 0 1 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0
0 1 1 0 0 1 1 1 0 0 0 1 1 1 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 1 0 0
1 1 0 0 0 0 0 1 1 1 0 0 0]

Model DB KMEANS SKLEARN()
silhouette: 0.27812960037852846
davies bouldin: 1.539407443402678
calinski_harabaz: 73.13555695663447
Dunn: 0.1814309700494348
```

Ambos modelos cuentan con sus fortalezas y debilidades, por parte del modelo Hierarchical clustering contar una visualización que presente la división o fusión consecutiva de los datos apoya la comprensión de su procedimiento, visualizar

Figura 3.62: Agrupación propuesta por el modelo Hierarchical Clustering

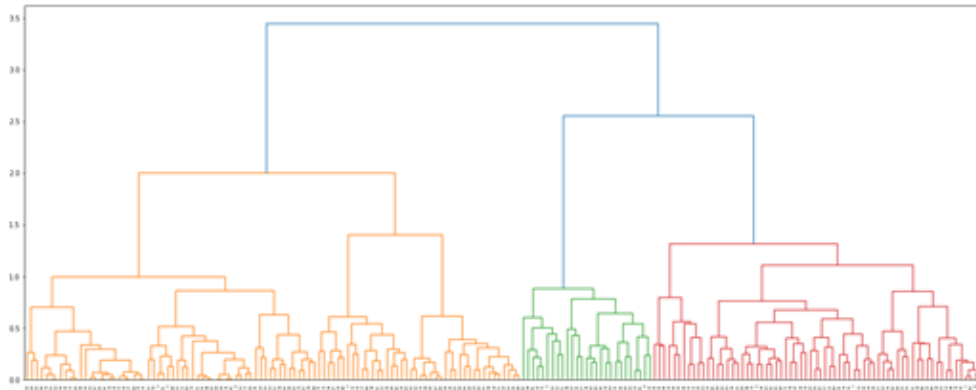
```

No. Leaves (rows): 197
Clustering by row :
[0 0 0 1 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 1 0 1 1 0 1 1
0 0 0 1 1 1 1 0 0 0 1 1 1 1 0 0 0 0 1 1 1 0 0 0 0 1 0 0 0 1 1 0 0 1 0 1 0
0 0 0 1 0 0 0 0 0 1 0 1 1 1 1 1 1 0 0 0 0 1 0 0 0 1 1 1 0 0 0 1 1 1 1 0
0 1 1 1 1 0 1 1 0 0 0 0 1 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1
1 0 0 1 1 0 0 0 1 1 1 0 0 0 1 0 0 0 1 1 1 1 0 0 1 1 1 0 0 0 0 1 1 0 0 0 1 1
0 0 0 1 1 1 1 0 0 0 1 1]
Model Hierarchical Clustering()
silhouette: 0.24232074932210107
davies bouldin: 1.6569815385818119
calinski_harabaz: 64.96868283149301
Dunn: 0.17538220817695743
Clusters : No. elements {0: 95, 1: 102}

```

Fig. 3.63, además de que este algoritmo tiende a mantenerse es su decisión de agrupamiento algo en lo que difiere el modelo K-Means el cual en caso de existir más de dos agrupaciones cambia notablemente la asignación de elementos a los clústeres creados, no obstante basa su análisis en el uso de centroides, un atributo muy característico de esta técnica lo que permitió crear un modelo más robusto por ello y tomando en cuenta los resultados de su rendimiento se optó por emplear este algoritmo para el desarrollo del modelo de ML, cabe mencionar que en caso de que el número de agrupaciones optimas hubiera resultado en un número mayor de 2 el uso del modelo Hierarchical clustering se hubiera considerado como una primera opción.

Figura 3.63: Agrupación propuesta por el modelo Hierarchical Clustering



**3.15.4.3. Análisis de las agrupaciones generadas por el uso del modelo K-Means**

Una vez seleccionado el modelo de agrupación para los elementos registrados en el dataset “conjuntoConductas\_Final”, se procedió a realizar el entrenamiento que brindó las salidas presentadas en la Fig. 3.60 “Agrupación propuesta por el modelo K-means”, no obstante el contar con las 2 agrupaciones no fue sinónimo de tener un conocimiento sobre estas, para ello fue necesario realizar un análisis de los elementos que conformaban cada aglomeración con la intención de comprender la relación de estos entre sí y de esa manera poder obtener un mayor contexto de la información reunida.

Primeramente, se decidió emplear el conocimiento de los intervalos de tiempo que cada tupla representaba al registrar las actividades ejercidas dentro de los horarios establecidos y añadir la variable a cada uno de los elementos de cada clúster, esto con la intención de averiguar si influía la variable tiempo u horas en los elementos reunidos. La Fig. 3.64 presenta los elementos de cada clúster alineado con su respectivo horario.

Figura 3.64: Visualización parcial de los elementos de cada cluster con su respectivo horario

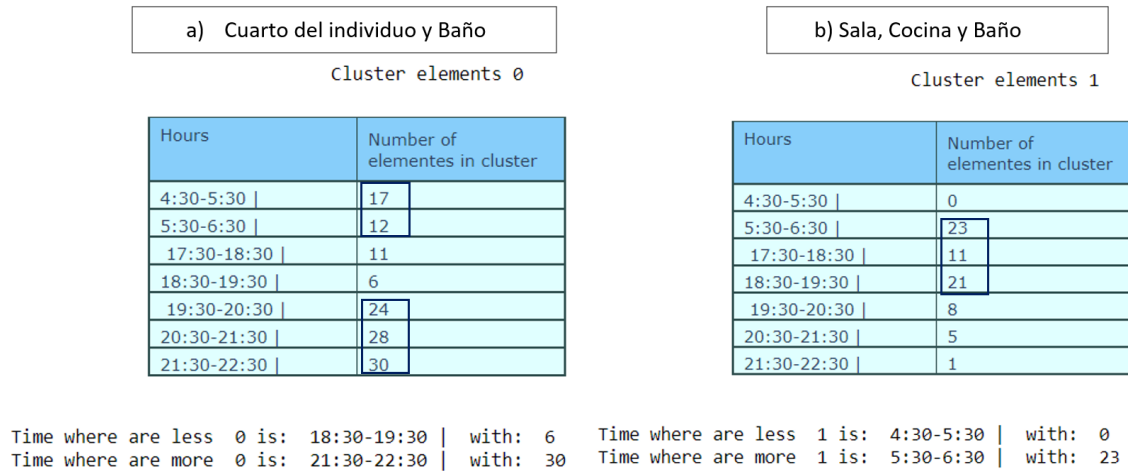
<i>a) Clúster 0</i>	<i>b) Clúster 1</i>
posi: 1 hora 17:30-18:30	posi: 0 hora 5:30-6:30
posi: 3 hora 19:30-20:30	posi: 2 hora 18:30-19:30
posi: 4 hora 20:30-21:30	posi: 9 hora 19:30-20:30
posi: 5 hora 21:30-22:30	posi: 15 hora 18:30-19:30
posi: 6 hora 4:30-5:30	posi: 19 hora 5:30-6:30
posi: 7 hora 5:30-6:30	posi: 20 hora 18:30-19:30
posi: 8 hora 18:30-19:30	posi: 23 hora 5:30-6:30
posi: 10 hora 20:30-21:30	posi: 24 hora 19:30-20:30
posi: 11 hora 21:30-22:30	posi: 25 hora 21:30-22:30
posi: 12 hora 4:30-5:30	posi: 31 hora 19:30-20:30
posi: 13 hora 5:30-6:30	posi: 34 hora 5:30-6:30
posi: 14 hora 17:30-18:30	posi: 37 hora 5:30-6:30
posi: 16 hora 19:30-20:30	posi: 39 hora 18:30-19:30
posi: 17 hora 20:30-21:30	posi: 44 hora 5:30-6:30
posi: 18 hora 21:30-22:30	posi: 45 hora 17:30-18:30
posi: 21 hora 19:30-20:30	
posi: 22 hora 20:30-21:30	

Una vez obtenida la información total, el programa registró en una tabla por clúster la cantidad de elementos por horario, esto con la finalidad de visualizar de una manera simple qué horarios dominaban en cada agrupación y por ende averiguar las actividades relacionadas a estos. En la Fig. 3.65 se pueden visualizar las tablas pertenecientes a cada uno de los dos clústeres, donde se pudo visualizar que en el clúster 0 se encontraba un mayor número de actividades y variable reunidas en horarios mañaneros y nocturnos por lo que se dedujo que esta agrupación estaba dominada por conductas desarrolladas en la habitación y baño del individuo, es decir cuando este se encontraba dormido o alistándose para ir a trabajar y cuando terminaba sus actividades del día dirigiéndose a su cuarto a descansar.

En el clúster 1 se concluyó que la aglomeración de elementos se registraba mayormente en un horario mañanero e intermedio (tarde), lo que concuerda con actividades realizadas en el área de sala, cocina y baño del individuo es decir cuando el individuo desayunaba en

las mañanas, llegaba de trabajar a preparar sus alimentos y finalmente se sentaba a comer.

Figura 3.65: Elementos de cada cluster agrupados por sus respectivos horarios



A partir del conocimiento obtenido del registro de tablas correspondientes a cada clúster se programó el modelo para obtener un análisis más exhaustivo sobre los elementos involucrados en cada agrupación, tipo de actividades, frecuencia de estas, valores mínimos y máximos etc., todo con la finalidad de obtener la suficiente información para detectar y establecer el comportamiento base del individuo con las conductas que este presenta a lo largo del día y de esta manera poder comprender y visualizar el por qué un elemento o conducta se consideraría diferente al resto.

La Fig.3.66 presenta el análisis total de los elementos involucrados en el clúster 0 y la Fig 3.67 del clúster 1 , donde se evaluó la información de las tecnologías aplicadas, es decir los valores incluidos en los atributos que conformaban el dataset “conjuntoConductas\_Final”, por ejemplo el atributo que contara con el nombre de la actividad a detectar seguido de \_AS sólo podía contener los valores disponibles de la escala de anomalía (1,2,3), los que se completaban con \_CF solo podía incluir los valores (1 y 2) correspondientes a los dos lapsos de 30 minutos por hora y para el resto de los atributos correspondiente a las tecnologías Garmin y Daylio los valores planteados equivalían a las agrupaciones que existían dentro de la misma columna.

Figura 3.66: Análisis total de los elementos involucrados en el cluster 0

```

LE_AS {0: 95, 1: 14, 2: 15, 3: 4}
M_CF {0: 88, 1: 32, 2: 8}
R_AS {0: 118, 1: 8, 2: 2}
R_CVT {0: 118, 1: 10}
E_AS {0: 120, 1: 5, 2: 3}
UT_AS {0: 115, 1: 7, 2: 3, 3: 3}
BH_AS {0: 103, 1: 10, 2: 9, 3: 6}
LC_AS {0: 65, 1: 28, 2: 24, 3: 11}
MC_CF {0: 57, 1: 30, 2: 41}
CC_AS {0: 54, 1: 9, 2: 29, 3: 36}
SOCIAL {1: 103, 2: 10, 3: 9, 4: 6}
GOALS {1: 103, 2: 19, 3: 6}

ESTRES {0.6: 4, 0.7: 2, 0.8: 2, 0.9: 8, 1.0: 2, 1.2: 4, 1.3: 6, 1.5: 11, 1.6: 4, 1.7: 3, 1.8: 9, 1.9: 7, 2.0: 9, 2.2: 6, 2.3: 1, 2.5: 2, 2.6: 6, 2.7: 5, 3.0: 6, 3.1: 4, 3.2: 3, 3.5: 7, 3.7: 3, 3.8: 2, 3.9: 3, 4.0: 4, 4.8: 3, 5.9: 2}
CARDIACA {4.6: 3, 4.9: 4, 5.0: 13, 5.1: 24, 5.2: 15, 5.3: 29, 5.4: 24, 5.5: 5, 5.6: 5, 5.7: 1, 5.8: 3, 6.3: 2}
HRS_SUEN {3.8: 4, 4.7: 3, 6.7: 4, 6.9: 5, 7.0: 7, 7.1: 4, 7.2: 12, 7.3: 2, 7.4: 4, 7.6: 13, 7.7: 6, 7.8: 14, 7.9: 5, 8.0: 4, 8.1: 3, 8.2: 12, 8.3: 6, 8.4: 3, 8.5: 3, 8.6: 4, 8.7: 2, 9.0: 1, 9.1: 2, 9.4: 5}

```

Figura 3.67: Análisis total de los elementos involucrados en el cluster 1

```

LE_AS {0: 95, 1: 14, 2: 15, 3: 4}
M_CF {0: 88, 1: 32, 2: 8}
R_AS {0: 118, 1: 8, 2: 2}
R_CVT {0: 118, 1: 10}
E_AS {0: 120, 1: 5, 2: 3}
UT_AS {0: 115, 1: 7, 2: 3, 3: 3}
BH_AS {0: 103, 1: 10, 2: 9, 3: 6}
LC_AS {0: 65, 1: 28, 2: 24, 3: 11}
MC_CF {0: 57, 1: 30, 2: 41}
CC_AS {0: 54, 1: 9, 2: 29, 3: 36}
SOCIAL {1: 103, 2: 10, 3: 9, 4: 6}
GOALS {1: 103, 2: 19, 3: 6}

ESTRES {0.6: 4, 0.7: 2, 0.8: 2, 0.9: 8, 1.0: 2, 1.2: 4, 1.3: 6, 1.5: 11, 1.6: 4, 1.7: 3, 1.8: 9, 1.9: 7, 2.0: 9, 2.2: 6, 2.3: 1, 2.5: 2, 2.6: 6, 2.7: 5, 3.0: 6, 3.1: 4, 3.2: 3, 3.5: 7, 3.7: 3, 3.8: 2, 3.9: 3, 4.0: 4, 4.8: 3, 5.9: 2}
CARDIACA {4.6: 3, 4.9: 4, 5.0: 13, 5.1: 24, 5.2: 15, 5.3: 29, 5.4: 24, 5.5: 5, 5.6: 5, 5.7: 1, 5.8: 3, 6.3: 2}
HRS_SUEN {3.8: 4, 4.7: 3, 6.7: 4, 6.9: 5, 7.0: 7, 7.1: 4, 7.2: 12, 7.3: 2, 7.4: 4, 7.6: 13, 7.7: 6, 7.8: 14, 7.9: 5, 8.0: 4, 8.1: 3, 8.2: 12, 8.3: 6, 8.4: 3, 8.5: 3, 8.6: 4, 8.7: 2, 9.0: 1, 9.1: 2, 9.4: 5}

```

Por otra parte, las Tablas 3.17 y 3.18 presentan un análisis que reunió los valores estadísticos referentes al conjunto de conductas y variables reunidos en cada agrupación con esto fue posible conocer las conductas más repetidas por intervalos de tiempo, las menos repetidas y la media, mínimo y máximo valor correspondiente a cada atributo por clúster.

La Tabla 3.17 corresponde al clúster 0, el cual involucra las áreas del cuarto/baño y por ello se puede observar como los atributos relacionados con las áreas restantes no presentan

valores en las variables “Median” y “Min Value”, esto se puede corroborar al observar que los valores más repetidos corresponden a los atributos relacionados a estas áreas (la interacción con la iluminación, movimiento en el entorno, presencia en cama entre otros), por otra parte algo que resalta en este clúster involucra las variables “SOCIAL” y “GOALS” las cuales parecen tener más presencia en esta zona al registrar un mayor número en la variable “Times”, finalmente analizando el atributo “ESTRES” se puede observar una menor límite máximo de 5.9 a comparación del clúster 1, el cual presenta un valor de 6.5, esto puede ser resultado de las actividades que se realizan en este entorno las cuales pueden reducir el impacto en este atributo.

Tabla 3.17: Análisis estadístico de los elementos involucrados en el cluster 0

Features	Median	Min Value	Max value	Most repeated value	Times	Least repeated value	Times
LE_AS	0	0	3	2	15	3	4
M_CF	0	0	2	1	32	2	8
R_AS	0	0	2	1	8	2	2
R_CVT	0	0	1	1	10	1	10
E_AS	0	0	2	1	5	2	3
UT_AS	0	0	3	1	7	2	3
BH_AS	0	0	3	1	10	3	6
LC_AS	0	0	3	1	28	3	11
MC_CF	1	0	2	2	41	1	30
CC_AS	2	0	3	3	36	1	9
SOCIAL	1	1	4	1	103	4	6
GOALS	1	1	3	1	103	3	6
ESTRES	2	0.6	5.9	1.5	11	2.3	1
CARDIACA	5.3	4.6	6.3	5.3	29	5.7	1
HRS_SUEN	7.75	3.8	9.4	7.8	14	9	1

En la Tabla 3.18 correspondiente al clúster 1 se puede observar un mayor registro en las variables “Median”, “Min Value” y “Max value” correspondientes a los atributos perteneciente a las áreas sala/cocina/baño donde resalta la cantidad de veces máxima en la que la acción abrir/cerrar del refrigerador tiene lugar, no obstante también se puede observar como otras actividades son repetidas constantemente y por un tiempo prologado (1 hora completa = 3) tal es el caso, del manejo de iluminación o presencia en la mesa, con respecto a las variables “Goals” y “SOCIAL” su registro de valores máximo (3 y 4) representan que el individuo logró llegar a realizar esta cantidad de actividades recreativas, no obstante, el predominio se registró en 1 con una repetición de 61 en ambos casos. Finalmente analizando el atributo “Estrés” se pudo observar que en este clúster los niveles alcanzaron un mayor número llegando a un valor máximo de 6.5 y una constante de 2.

Tabla 3.18: Análisis estadístico de los elementos involucrados en el cluster 1

Features	Median	Min Value	Max value	Most repeated value	Times	Least repeated value	Times
LE_AS	2	1	3	3	25	1	5
M_CF	2	1	2	2	35	1	28
R_AS	2	1	3	2	31	1	11
R_CVT	2	1	12	1	25	12	1
E_AS	1	1	3	1	14	3	13
UT_AS	0	1	3	3	23	1	2
BH_AS	0	1	2	2	6	1	1
LC_AS	1	1	3	2	18	1	11
MC_CF	1	1	2	1	28	2	15
CC_AS	0	1	2	1	7	2	4
SOCIAL	1	1	3	1	61	2	2
GOALS	1	1	2	1	61	2	8
ESTRES	2	0.6	6.5	2	7	3	1
CARDIACA	5.3	4.6	6.3	5.3	13	4.6	1
HRS_SUEN	7.6	3.8	9.4	7.6	9	3.8	1

### 3.15.5. Desarrollo Final del modelo de ML encargado de la detección de anomalías en los patrones de conducta del individuo.

El uso del modelo K-Means brindó las dos agrupaciones que reunían las conductas con mayor relación entre sí, permitiendo que el modelo fuera capaz de analizar cada clúster por separado, sus elementos y la información que cada uno brindaba sobre los patrones de conducta del individuo, deduciendo de esta manera los horarios que influían en las actividades que este realizaba, su frecuencia y las áreas de la vivienda involucradas.

Hasta el momento ya fue establecido y analizado el comportamiento base del individuo a través de la serie de actividades y conductas presentadas por este durante el tiempo que este estudio se realizó (1 mes y medio), no obstante, para llevar a cabo la detección de conductas anómalas dentro de los patrones estudiados, se requirió hacer uso del funcionamiento del propio modelo K-Means donde cada clúster con su respectivo centroide y los elementos que lo rodean, permitió basarse en las distancias de los elementos con la finalidad de deducir cuales eran los más arraigados al clúster mismo (distancia próxima al centroide) y cuales tenían una menor relación al resto (distancia lejana al centroide), a continuación se presentará la serie de pasos a seguir para la detección de conductas anómalas por medio del modelo de

ML desarrollado.

1. Por cada clúster se obtuvieron las distancias de cada elemento hacia su respectivo centroide
2. Fue necesario relacionar las distancias obtenidas a sus respectivas tuplas con la finalidad de alinear datos.
3. Se realizó un análisis estadístico que involucró los valores promedio, la mediana, moda, mínimos y máximos de las distancias correspondientes a cada agrupación, los resultados son presentados en la Fig.3.68, donde se puede observar una mayor dispersión de los elementos involucrados en el clúster 1 con respecto al clúster 0, esto en base a las distancias registradas las cuales superan en valor de promedio, mediana, moda y mínima distancia a su contraparte, no obstante la variable “max” la cuál representa la mayor distancia registrada en la agrupación resulta menor a la del cluster 0, lo que indica que muy probablemente dentro de éste exista un elemento anómalo.

Figura 3.68: Análisis estadístico entre las distancias y su respectivo centroide

```

===== Statistics of centroid distances =====

```

Clusters	Average	Median	Mode	Min	Max
Cluster 0	0.34497	0.30745	0.25105	0.1214	0.85545
Cluster 1	0.41464	0.43104	0.51365	0.20785	0.62794

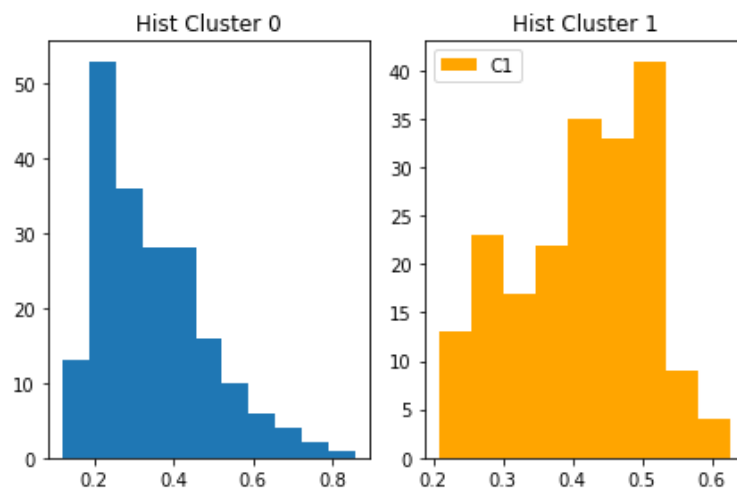
4. Con la finalidad de poder presentar visualmente las distancias que más representaban cada agrupación se realizó su graficación a partir del uso de un histograma para cada clúster presentado en la Fig 3.68 y en conjunto Fig 3.69.

La Figura 3.68 corresponde a la visualización de dos histogramas los cuales representan la distribución de los elementos correspondientes a cada agrupación, donde podemos observar que en el cluster 0 existe un mayor número de elementos dentro del rango de 0.2-0.3 y donde su reducción alcanza a llegar al rango de 0.8-0.9, también

visualizamos una continuidad entre los valores, es decir, no hay intervalos de valores lo suficientemente aislados como para resaltar.

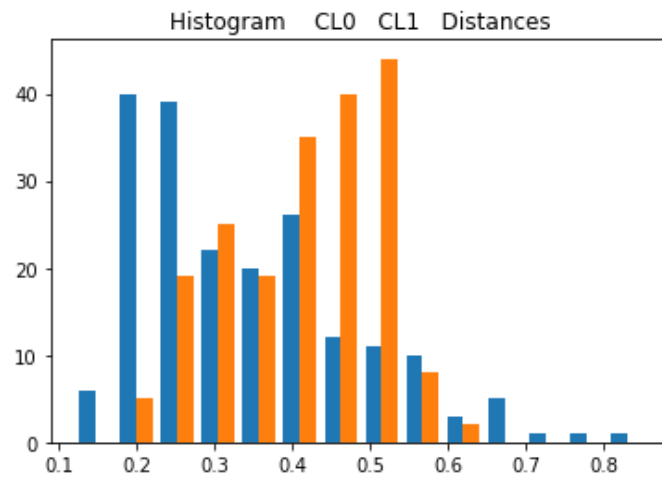
La imagen correspondiente al clúster 1 presenta una distribución más variada, a diferencia de su contraparte no presenta un comportamiento predominantemente descendente sino que existen lapsos donde los valores suben y bajan abruptamente, esto equivale a la concentración de elementos en ciertos intervalos de distancias alcanzando valores de alta densidad entre el rango de 0.40-0.50.

Figura 3.69: Histograma de distancias por clúster



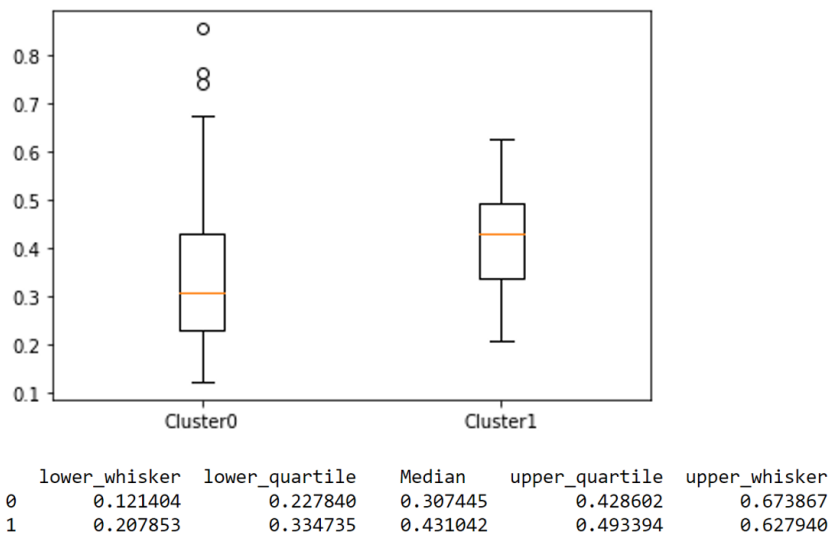
La figura 3.70 presentada a continuación involucra la interacción entre ambos histogramas donde se observa un comportamiento contrario entre ambas agrupaciones, existen dos puntos que resaltan en la visualización, el primero se relaciona con las distancias del clúster 0, las cuales al ser de mayor valor llegan a aislarse en 4 intervalos diferentes, el segundo punto es referente a los intervalos en los que ambas agrupaciones presentan menor diferencia entre sí, estos se encuentran entre los valores de 0.27- 0.37 aproximadamente.

Figura 3.70: Histograma de distancias por clúster en conjunto



5. Se aplicó la técnica de “boxplot” a las distancias obtenidas para poder así detectar los puntos que sobresalían de los valores mínimos y máximos en los bigotes o “whiskers” de cada clúster representado con su respectivo boxplot y presentado en la Fig.3.71, donde se puede observar que de las dos agrupaciones, el clúster 1 muestra una distribución más normal con respecto al clúster 0, éste último además presenta de manera más notoria la presencia de anomalías, las cuales sobresalen de la parte máxima de los límites establecidos.

Figura 3.71: Boxplot referente a cada clúster



6. Una vez localizados los puntos anómalos de cada agrupación en base a su distancia se seleccionaron todo el conjunto de valores alineados a las distancias seleccionadas (Tuplas completas), la Fig 3.72 presenta el conjunto de tuplas que representan una anomalía dentro del dataset “conjuntoConducta\_Final”.

Figura 3.72: Anomalías detectadas por Kmeans-boxplot

```

Upper Anomalies = Cluster : 0
69 0.763803
110 0.673867
122 0.748999
182 0.855448
Name: dist_0, dtype: float64

Under Anomalies = Cluster : 0
114 0.121484
Name: dist_0, dtype: float64

Upper Anomalies = Cluster : 1
91 0.62794
Name: dist_1, dtype: float64

Under Anomalies = Cluster : 1
59 0.207853
Name: dist_1, dtype: float64

Out[13]:

```

	LE_AS	M_CF	R_AS	R_CVT	E_AS	UT_AS	BH_AS	LC_AS	MC_CF	CC_AS	SOCIAL	GOALS	ESTRES	CARDIACA	HRS_SUEN	FECHA	TIME	cluster_
59	2	1	2	2	2	0	0	2	1	0	1	1	3.8	5.6	7.6	6	3	1
69	0	2	3	8	3	3	0	0	0	0	1	1	0.9	5.2	7.9	9	4	1
91	0	0	0	0	0	0	0	0	0	3	1	1	2.6	5.2	3.8	18	1	0
110	0	0	3	8	2	0	0	1	1	0	1	1	2.0	5.4	8.6	21	3	1
114	0	0	0	0	0	0	0	1	1	1	1	1	2.0	5.4	8.6	21	7	0
122	0	2	3	5	3	0	0	0	0	0	1	1	1.7	5.1	4.7	23	4	1
182	0	1	3	12	0	3	2	2	1	0	3	2	0.6	5.1	8.2	7	5	1

7. Con la finalidad de brindar una mayor robustez al modelo desarrollado se incorporó COPOD() para corroborar la selección de conducta anómalas debido al funcionamiento y los resultados de su desempeño obtenidos a partir del diseño experimental realizado por ello se decidió fusionar la selección de ambos métodos con el propósito de obtener y analizar únicamente las tuplas o patrones de conducta que hallan representado una anomalía en ambas metodologías. La Fig.3.73 presenta los resultados obtenidos del uso del modelo COPOD()

Figura 3.73: Anomalías detectadas por el modelo COPOD()

```

=====          COPOD          =====
# Anomalies: 40   Clusters: {0.0: 157, 1.0: 40}

Row index: 0 score: 20.58978340826478
Row index: 2 score: 19.92021476372805
Row index: 3 score: 26.553264160800133
Row index: 19 score: 22.155795849395847
Row index: 20 score: 22.72792195840363
Row index: 21 score: 21.827103561654138
Row index: 23 score: 24.18831691940339
Row index: 24 score: 23.637283259202
Row index: 25 score: 23.79043882777302
Row index: 31 score: 20.219972063053476
Row index: 32 score: 21.281145327517503
Row index: 35 score: 23.802050175542004
Row index: 37 score: 21.780433751416872
Row index: 39 score: 21.405805313656185
Row index: 40 score: 22.77705960575499
Row index: 48 score: 22.51125282155987
Row index: 61 score: 22.45724874961048
Row index: 63 score: 21.071743266526003
Row index: 69 score: 26.002155933030046
Row index: 70 score: 23.548150039701753
Row index: 75 score: 22.88290680588206
Row index: 91 score: 21.025910475173262
Row index: 92 score: 26.473584440668567
Row index: 93 score: 21.975364231730577
Row index: 94 score: 23.13399727206511
Row index: 95 score: 35.95775676530183
Row index: 99 score: 25.594267861220548
Row index: 110 score: 20.642837348159205
Row index: 120 score: 20.30179066097091
Row index: 122 score: 27.849970965950998
Row index: 123 score: 26.426210606794758
Row index: 125 score: 21.431325478667258
Row index: 127 score: 26.652798768624752
Row index: 144 score: 19.90522781769674
Row index: 150 score: 21.234301412434363
Row index: 153 score: 20.93683802550019
Row index: 170 score: 22.350180494999897
Row index: 181 score: 21.514125108576597
Row index: 182 score: 26.47964716437529
Row index: 194 score: 21.959382407483677

```

### 3.15.5.1. Resultados de la selección entre ambos procesos integrados

Una vez obtenidos las conductas anómalas de cada modelo empleado en conjunto (Modelo desarrollado y COPOD ()) se realizó un filtro en el cual se seleccionaron únicamente las tuplas que se encontraran en los resultados de ambos modelos y de esa manera brinda mayor robustez y fiabilidad al modelo final.

La Fig.3.74 muestra el total de las tuplas seleccionadas por ambos algoritmos que contienen en si algún valor o conducta que destaca al resto pero para encontrar una mayor explicación se requiere realizar un último análisis a los resultados para comprar los valores y conductas que conforman los patrones en cada tupla con el resto del conjunto de datos final.

Figura 3.74: Anomalías detectadas por el modelo de ML desarrollado

```

***** Anomalies indexes in common: *****
[ 69  91 110 122 182]

Out[12]:

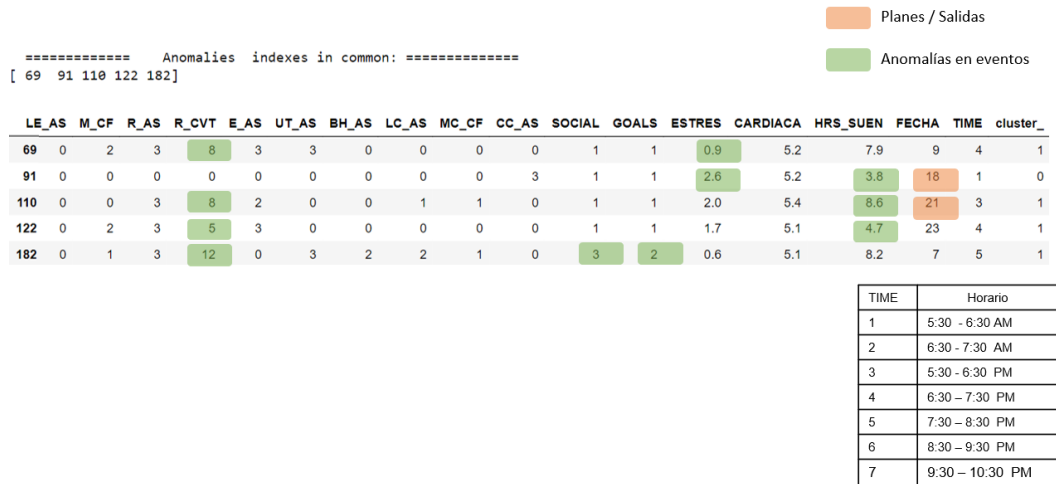
```

	LE_AS	M_CF	R_AS	R_CVT	E_AS	UT_AS	BH_AS	LC_AS	MC_CF	CC_AS	SOCIAL	GOALS	ESTRES	CARDIACA	HRS_SUEN	FECHA	TIME	cluster_
69	0	2	3	8	3	3	0	0	0	0	1	1	0.9	5.2	7.9	9	4	1
91	0	0	0	0	0	0	0	0	0	3	1	1	2.6	5.2	3.8	18	1	0
110	0	0	3	8	2	0	0	1	1	0	1	1	2.0	5.4	8.6	21	3	1
122	0	2	3	5	3	0	0	0	0	0	1	1	1.7	5.1	4.7	23	4	1
182	0	1	3	12	0	3	2	2	1	0	3	2	0.6	5.1	8.2	7	5	1

## 3.16. Resultados

Una vez finalizado el tiempo programado para la duración de este trabajo, se evaluaron los resultados finales obtenidos a través de la aplicación del modelo de ML desarrollado, la Fig.3.75 presenta un análisis realizado a cada uno de los puntos señalados como anómalos con el propósito de detectar las conducta o variables que influyeron en su selección. Para mayor comprensión se recomienda consultar las Tablas 3.17-3.18 y Fig. 3.66-3.67.

Figura 3.75: Análisis de las anomalías detectadas por el modelo de ML desarrollado



Los elementos anómalos resultantes se obtuvieron con base a su relación con el resto de los atributos no solamente comparándose sobre los elementos de sí mismo, es decir, el análisis fue realizado tomando en cuenta la información de cada tecnología reunida, los atributos resultantes, su relación con el resto y los horarios establecidos que permitieron identificar las conductas que se presentaban en conjunto .

Como se puede observar, conductas como abrir un número de veces elevado el refrigerador, desplazarse en áreas del hogar fuera de los horarios normales, presentar niveles de estrés fuera de lo común, realizar salidas u actividades que rompieran con la rutina establecida o presentar alteraciones en el ciclo de sueño formaron parte de las conductas que influyeron en la selección de los puntos por parte del modelo de ML al destacar de los patrones de conducta establecidos en base al entrenamiento de los modelos de ML con los datos reunidos del entorno y del individuo diariamente por el tiempo asignado.

# Capítulo 4

## Conclusiones y trabajo futuro

### 4.1. Conclusiones

La metodología, diseños, desarrollos y análisis propuestos muestra resultados prometedores considerando las tecnologías aplicadas, el tiempo asignado y considerando el sector economico-social al que va dirigido.

Las tecnologías aplicadas permitieron reunir y procesar una gran cantidad de información debido al enfoque y desarrollo con el que fueron seleccionadas tanto para el entorno como para la persona misma, siempre teniendo en cuenta el objetivo a cumplir, las necesidades a cubrir, el impacto en el espacio, la privacidad y el confort por parte del individuo. La información reunida y procesada a través de estas tecnologías se organizó de manera que fue posible la detección de los patrones de conducta que ejercía el individuo a lo largo del día por el tiempo establecido, con base en estos fue posible detectar las conductas y variables que destacaban del resto ya sea por el horario de ejecución, duración, frecuencia, valor, entorno donde se ejercía y circunstancias del espacio.

Una vez obtenidos los patrones de conducta desarrollados por el individuo estos fue necesaria la aplicación de los modelos de ML desarrollados durante el monitoreo del entorno experimental, estos aplicaron una combinación de algoritmos para su funcionamiento, entre los que se encuentran K-means para la formación de agrupaciones y COPOD para la

detección de datos anómalos, por otra parte también fue necesario el uso de técnicas y herramientas estadísticas para el análisis de las agrupaciones creadas (boxplot, media, moda, mediana, histogramas entre otros.) Finalmente los resultados obtenidos cumplieron con los objetivos establecidos al haber sido posible determinar el conjunto de patrones de conducta que permitieran la detección de conductas anómalas en un individuo a través de acciones específicas ejercidas dentro su entorno habitacional, fuera de este y el monitoreo de su estado físico, todo esto sin intervenir-modificar la rutina diaria del individuo o hacer uso de dispositivos que resultaran intrusivos, incómodos o dañinos para el este. Por otra parte existe una concordancia parcial de los resultados finales con respecto a la hipótesis planteada debido a que la información obtenida por parte de las tecnologías empleadas pudo haber sido más explotable si estas hubiesen sido implementadas por un mayor tiempo al disponible permitiendo obtener resultados más concretos y estables a los presentados. Cada uno de los procesos realizados al igual que las tecnologías aplicadas pueden ser mejorados al hacer uso de sensores con mayor potencial, nuevas tecnologías y dispositivos dentro del entorno, no obstante se debe contemplar en cada elección el propósito y sector al que va dirigido, esto con la intención de mantener su implementación lo más accesible a la población mexicana y de esta manera impulsar el desarrollo tecnológico del país permitiendo que cada vez más personas puedan hacer uso de entornos que permitan el seguimiento diario de sus actividades y conductas presentadas apoyando el cuidado de su persona y brindando la posibilidad de mejorar la experiencia en la interacción con su entorno.

Los resultados presentados en este trabajo muestran el potencial de su aplicación en áreas de la salud mental, las cuales a pesar de contar con tecnologías e instrumentos dedicados al seguimiento, desarrollo y detección de conductas de riesgo en los pacientes requieren de nuevas tecnologías capaces de brindar una mayor visión del panorama que vive el paciente en su vida diaria. La información que logra aportar el trabajo desarrollado permite apoyar el seguimiento, evaluaciones y diagnósticos realizados por los especialistas en el área hacia el paciente que lo requiera, brindando la oportunidad de poder dedicar un servicio más

personalizado que beneficie al cuidado mental, emocional y social del individuo.

## 4.2. Trabajo futuro

Dentro de los puntos principales a trabajar se encuentra la implementación de la red IoT con dispositivos que en su mayoría hagan uso del modo de ahorro “deep sleep”, es decir, que estos apliquen la menor cantidad de energía eléctrica para su funcionamiento lo que permitiría prolongar la duración de la fuente de energía y prescindir del uso de cargadores, adaptadores y cierto cableado dentro de la infraestructura disminuyendo la complejidad del diseño en red y los costos de su implementación.

Además se espera expandir el tiempo de la implementación del estudio con la finalidad de obtener una mayor cantidad de información que refuerce los resultados obtenidos y permita tener la oportunidad de cumplir en su totalidad la hipótesis planteada al permitir realizar un mayor entrenamiento a los modelos de ML desarrollados empleando los datos que la red IoT, el smartwatch Garmin y la aplicación móvil Daylio pueden brindar en conjunto.

Se pretende expandir el sistema desarrollado con el diseño, desarrollo e implementación de una plataforma que a través del uso de servicios externos en la nube apoye al especialista y usuario la consulta de la información procesada y obtenida por parte del sistema desarrollado al permitir reunir y registrar toda la información que los modelos de ML y las tecnologías empleadas proporcionan constantemente, de esta manera un especialista en el área de la salud mental podría acceder en el momento necesario a dicha plataforma, realizar un análisis de la información disponible sobre el paciente en cuestión y apoyar sus evaluaciones, diagnósticos y tratamientos con base en el estilo de vida diario de este. No obstante pese a que se propone la aplicación de este sistema para una atención personalizada, el contar con una tecnología que permita reunir la información referente a los patrones de conducta de diversos individuos bajo los términos legales y de confidencialidad deja las puertas abiertas para análisis, detección y obtención de nuevos resultados en torno al cuidado de la salud mental mundial.

Una plataforma que registre los patrones de conducta de diversas personas y la detección de conductas anómalas en ellos tendría el alcance de encontrar patrones dentro de los mismos patrones de conducta sociales favoreciendo la detección de puntos claves para el tratamiento de los pacientes que cuenten con situaciones, entornos o calidad de vida similares.

# Bibliografía

- [1] B. Slife and R. Williams, “Science and Human Behavior,” *What’s Behind the Research? Discovering Hidden Assumptions in the Behavioral Sciences*, pp. 167–204, 2014.
- [2] P. Chance, *Learning and Behavior*. Cengage Learning, 7th ed., 2013.
- [3] B. Jose, *Psicología De La Conducta*. Nueva Edición, Paidós, 2007.
- [4] M. Domjan, *Principios de aprendizaje y conducta*. 2010.
- [5] P. Kleinman, *Psych 101*. 2018.
- [6] N. A. Ruvalcaba-Romero, P. Fernández-Berrocal, J. G. Salazar-Estrada, and J. Gallegos-Guajardo, “Positive emotions, self-esteem, interpersonal relationships and social support as mediators between emotional intelligence and life satisfaction,” *Journal of Behavior, Health Social Issues*, vol. 9, no. 1, pp. 1–6, 2017.
- [7] M. Hakkak, A. Nazarpouri, S. N. Mousavi, and M. Ghodsi, “Investigating the effects of emotional intelligence on social-mental factors of human resource productivity,” *Revista de Psicología del Trabajo y de las Organizaciones*, vol. 31, no. 3, pp. 129–134, 2015.
- [8] B. Mundial, “Las conductas de riesgo constituyen una amenaza creciente para la salud mundial.” <https://www.bancomundial.org/es/news/press-release/2013/11/20/risky-behaviors-growing-threats-global-health>, 2013.

- [9] P. Chance, *Risking Your Health: Causes, Consequences, and Interventions to Prevent Risky Behaviors*. World Bank Publications, 2014.
- [10] S. del Trabajo y Previsión Social, “Norma Oficial Mexicana NOM-035-STPS-2018, Factores de riesgo psicosocial en el trabajo-Identificación, análisis y prevención..” <https://www.gob.mx/stps/articulos/norma-oficial-mexicana-nom-035-stps-2018-factores-de-riesgo-psicosocial-en-el-trabajo-identificacion-analisis-y-prevencion>, 2019.
- [11] S. del Trabajo y Previsión Social, “NOM-035-STPS-2018, Factores de riesgo psicosocial en el trabajo-Identificación, análisis y prevención.” [https://www.dof.gob.mx/nota\\_detalle\\_popup.php?codigo=5541828](https://www.dof.gob.mx/nota_detalle_popup.php?codigo=5541828), 2018.
- [12] C. E. Muñoz-landeros, D. García-silva, R. D. García, E. E. Paypa-jabre, and E. A. García-valadez, “medicina universitaria,” vol. 15, no. 59, pp. 51–58, 2013.
- [13] A. Dominguez-Rodriguez, S. M. Chavez-Valdez, G. C. Avitia, and L. C. Valencia-Espinoza, “Unified protocol for anxiety disorders in two cities of Mexico measuring gamma activity: Study protocol for a randomized controlled trial,” *Contemporary Clinical Trials Communications*, vol. 18, p. 100556, 2020.
- [14] P. Spagnolo, *Human Behavior Understanding in Networked Sensing*. 2014.
- [15] C. Martin Fernandez, M. Diaz Rodriguez, and B. Rubio Munoz, “An edge computing architecture in the internet of things,” *Proceedings - 2018 IEEE 21st International Symposium on Real-Time Computing, ISORC 2018*, pp. 99–102, 2018.
- [16] C. E. Cobo Olivero, “El comportamiento humano,” vol. 19, no. 29, pp. 113–130, 2011.
- [17] D. S. Wilson, “Evolution and contextual behavioral science : a reunification.,” 2018.
- [18] J. Dewey, *Naturaleza humana y conducta. Introducción a la psicología social*. Fondo de Cultura Economica USA, 2014.

- [19] A. Souri, M. Y. Ghafour, A. M. Ahmed, F. Safara, A. Yamini, and M. Hoseyninezhad, “A new machine learning-based healthcare monitoring model for student’s condition diagnosis in Internet of Things environment,” *Soft Computing*, vol. 24, no. 22, pp. 17111–17121, 2020.
- [20] F. L. d. S. M. P. Donald L. Schomer MD, *Niedermeyer’s Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. LWW, 2010.
- [21] D. Acharya, N. Varshney, A. Vedant, Y. Saxena, P. Tomar, S. Goel, and A. Bhardwaj, “An enhanced fitness function to recognize unbalanced human emotions data,” *Expert Systems with Applications*, vol. 166, no. September 2019, p. 114011, 2021.
- [22] D. Devi, S. Sophia, A. Athithya Janani, and M. Karpagam, “Brain wave based cognitive state prediction for monitoring health care conditions,” *Materials Today: Proceedings*, 2020.
- [23] K. Schwab, *The Fourth Industrial Revolution*. World Economic Forum, 2016.
- [24] N. Davis, “The fourth industrial revolution.” <https://www.weforum.org/agenda/2016/01/what-is-the-fourth-industrial-revolution/>, 2016.
- [25] O. Q. onez Muñoz, *Internet de las Cosas (IoT)*. ibukku, 2019.
- [26] H. Sundmaeker, P. Guillemin, P. Friess, and S. Woelfflé, *Vision and Challenges for Realising the Internet of Things The meaning of things lies not in the things themselves, but in our attitude towards them. Antoine de Saint-Exupéry*. 2010.
- [27] S. R. Department, “México: devices connected internet per household.” <https://www.statista.com/statistics/1174769/devices-connected-internet-households-mexico/>, 2020.
- [28] K. L. Lueth, “What ceos talked about during the pandemic.” <https://iot->

- analytics.com/what-ceos-talked-about-during-the-pandemic/?utm\_source = Covid2020.
- [29] M. Rykov, “5 iot security best practices to consider after the covid-19 lockdown.” [https://iot-analytics.com/5-iot-security-best-practices-after-the-covid-19-lockdown/?utm\\_source = Covid2020](https://iot-analytics.com/5-iot-security-best-practices-after-the-covid-19-lockdown/?utm_source=Covid2020).
- [30] K. Lasse, “The impact of covid-19 on the internet of things,” 2020.
- [31] NEC, “Iot in transportation.” <https://www.nec.co.nz/market-leadership/publications-media/iot-in-transportation-5-applications-of-iot-technology-in-transportation/>, 2020.
- [32] K. M. Sadique, R. Rahmani, and P. Johannesson, “Towards security on internet of things: Applications and challenges in technology,” *Procedia Computer Science*, vol. 141, pp. 199–206, 2018.
- [33] M. C. S. and S. N. K, “Internet of Things Challenges and Apportunities,” *Smart Sensors, Measurement and Instrumentation*, vol. 42, no. 3, pp. 287–307, 2015.
- [34] Y. K. e. Dorothy Monekosso, Paolo Remagnino, *Intelligent Environments: Methods, Algorithms and Applications*. Advanced Information and Knowledge Processing, Springer-Verlag London, 1 ed., 2009.
- [35] P. Lea, *Internet of Things for architects : architecting IoT solutions by implementing sensors, communication infrastructure, edge computing, analytics, and security*. 2018.
- [36] Paschotta, Dr. Rüdiger, “Optical sensors,” 2020.
- [37] Universidad de castilla-La mancha, “Optoelectronica,” 2020.
- [38] Keyence, “Fundamentos del sensor,” 2020.
- [39] Thomas, “All about humity sensors,” 2020.

- [40] Rambal, “Sensor flex,” 2020.
- [41] mckinlay and R. danny causey Mazidi, *PIC Microcontroller and Embedded Systems Using ASM C for PIC18*. 2008.
- [42] E. White, *Making Embedded Systems: Design Patterns for Great Software*, vol. 2011. 2011.
- [43] R. Roman, J. Lopez, and M. Mambo, “Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges,” *Future Generation Computer Systems*, vol. 78, pp. 680–698, 2018.
- [44] Garmin, “Garmin vivosmart.” <https://buy.garmin.com/es-MX/MX/p/567813/pn/010-01755-03>, 1996-2021.
- [45] J. McCarthy, “The Little Thoughts of Thinking Machines.” <http://www-formal.stanford.edu/jmc/>, 1983.
- [46] J. Kaplan, *Artificial Intelligence: What Everyone Needs to Know*. Oxford University Press, 2016.
- [47] G. Gebremeskel, C. Yi, C. Wang, and Z. He, “Critical analysis of smart environment sensor data behavior pattern based on sequential data mining techniques,” *Industrial Management Data Systems*, vol. 115, pp. 1151–1178, 07 2015.
- [48] D. Zekri, T. Delot, M. Desertot, S. Lecomte, and M. Thilliez, “Using learning techniques to observe elderly’s behavior changes over time in smart home,” in *The Impact of Digital Technologies on Public Health in Developed and Developing Countries* (M. Jmaiel, M. Mokhtari, B. Abdulrazak, H. Aloulou, and S. Kallel, eds.), (Cham), pp. 129–141, Springer International Publishing, 2020.
- [49] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, “Machine learning: A review of

- classification and combining techniques,” *Artificial Intelligence Review*, vol. 26, no. 3, pp. 159–190, 2006.
- [50] P. Dangeti, *Statistics for Machine Learning Build*, vol. 53. 2013.
- [51] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: A review,” *ACM Comput. Surv.*, vol. 31, p. 264–323, Sept. 1999.
- [52] J. Han, M. Kamber, and J. Pei, “Third Edition : Data Mining Concepts and Techniques,” vol. 53, no. 9, pp. 1689–1699, 2012.
- [53] X. Zhou and M. Belkin, *Academic Press Library in Signal Processing: Volume 1 - Signal Processing Theory and Machine Learning*, vol. 1. Elsevier Masson SAS, 2014.
- [54] M. D. Pra, “Introduction to Reinforcement Learning.” <https://towardsdatascience.com/introduction-to-reinforcement-learning-c99c8c0720ef>, 2020.
- [55] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. P. Patel, A. Tiwari, M. J. Er, W. Ding, and C. T. Lin, “A review of clustering techniques and developments,” *Neurocomputing*, vol. 267, pp. 664–681, 2017.
- [56] A. Alcalde, “Aprendizaje no Supervisado y Detección de Anomalías: ¿Qué es el Clustering? .” <https://elbauldelprogramador.com/aprendizaje-nosupervisado-clustering/>, 2018.
- [57] S. Yildirim, “Hierarchical Clustering-Explained.” <https://towardsdatascience.com/hierarchical-clustering-explained-e58d2f936323>, 2020.
- [58] P. S. University, “Hierarchical Clustering.” <https://online.stat.psu.edu/stat555/node/85/>, 2018.
- [59] J. A. Rodrigo, “Clustering con python .” <https://www.cienciadedatos.net/documentos/py20-clustering-con-python.html>, 2020.

- [60] P. J. Rousseeuw, “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis,” *Journal of Computational and Applied Mathematics*, vol. 20, no. C, pp. 53–65, 1987.
- [61] H. Wei, “How to measure clustering performances when there are no ground truth.” <https://medium.com/@haataa/how-to-measure-clustering-performances-when-there-are-no-ground-truth-db027e9a871c>, 2020.
- [62] T. Calinski and J. Harabasz, “Communications in Statistics A dendrite method for cluster analysis,” *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974.
- [63] J. C. Bezdek and N. R. Pal, “Cluster validation with generalized Dunn’s indices,” *Proceedings - 1995 2nd New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, ANNES 1995*, pp. 190–193, 1995.
- [64] M. Goldstein and S. Uchida, “A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data,” *PLoS ONE*, vol. 11, no. 4, pp. 1–31, 2016.
- [65] V. Bajaj, “Unsupervised Learning For Anomaly Detection.” <https://towardsdatascience.com/unsupervised-learning-for-anomaly-detection-44c55a96b8c1>, 2020.
- [66] S. Agrawal and J. Agrawal, “Survey on anomaly detection using data mining techniques,” *Procedia Computer Science*, vol. 60, pp. 708–713, 2015.
- [67] J. Macqueen, “Some methods for classification and analysis of multivariate observations,” in *In 5-th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.
- [68] M. Galarnyk, “Understanding Boxplots.” <https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51>, 2018.

- [69] M. Alam, “Boxplot for anomaly detection.” <https://towardsdatascience.com/boxplot-for-anomaly-detection-9eac783382fd>, 2020.
- [70] A. Zubova, “Deconstructing the Box and Whisker Plot.” <https://dev.to/annalara/deconstructing-the-box-and-whisker-plot-11f3>, 2019.
- [71] Z. Li, Y. Zhao, N. Botta, C. Ionescu, and X. Hu, “COPOD: Copula-based outlier detection,” *Proceedings - IEEE International Conference on Data Mining, ICDM*, vol. 2020-November, no. 1, pp. 1118–1123, 2020.
- [72] Y. Y. Li and L. E. Parker, “Detecting and monitoring time-related abnormal events using a wireless sensor network and mobile robot,” *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, pp. 3292–3298, 2008.
- [73] T. Mori, A. Fujii, M. Shimosaka, H. Noguchi, and T. Sato, “Typical behavior patterns extraction and anomaly detection algorithm based on accumulated home sensor data,” *Proceedings of Future Generation Communication and Networking, FGCN 2007*, vol. 2, pp. 12–18, 2007.
- [74] D. Zhao, J. Yang, Y. Wang, C. Yu, Y. Jiang, S. Wang, and T. Fukuda, “Multiple Welfare-Robots Architecture of Smart House for Security Assistance of Daily Activities,” *2018 International Conference on Intelligence and Safety for Robotics, ISR 2018*, pp. 595–601, 2018.
- [75] A. A. Chaaoui, J. R. Padilla-López, F. J. Ferrández-Pastor, M. Nieto-Hidalgo, and F. Flórez-Revuelta, “A vision-based system for intelligent monitoring: Human behaviour analysis and privacy by context,” *Sensors (Switzerland)*, vol. 14, no. 5, pp. 8895–8925, 2014.
- [76] J. Reyes-Campos, G. Alor-Hernández, I. Machorro-Cano, J. O. Olmedo-Aguirre, J. L. Sánchez-Cervantes, and L. Rodríguez-Mazahua, “Discovery of resident behavior pat-

- terns using machine learning techniques and IoT paradigm,” *Mathematics*, vol. 9, no. 3, pp. 1–25, 2021.
- [77] D. Zekri, T. Delot, M. Desertot, S. Lecomte, and M. Thilliez, *Using Learning Techniques to Observe Elderly’s Behavior Changes over Time in Smart Home*, vol. 12157 LNCS. Springer International Publishing, 2020.
- [78] A. Beck, R. Steer, R. Ball, and W. Ranieri, “Comparison of Beck Depression 1 in Psychiatric Inventories -1A and - Outpatients,” *Journal of Personality Assessment*, vol. 67, no. 3, pp. 588–597, 1996.
- [79] D. A. González, A. R. Rodríguez, and I. Reyes-Lagunes, “Adaptation of the BDI-II in Mexico,” *Salud Mental*, vol. 38, no. 4, pp. 237–244, 2015.
- [80] V. Benet-Martinez and O. P. John, “Los Cinco Grandes Across Cultures and Ethnic Groups,” *Journal of Personality and Social Psychology Copyright*, vol. 75, no. 3, pp. 729–750, 1998.
- [81] D. J. Buysse, C. F. Reynolds, T. H. Monk, S. R. Berman, and D. J. Kupfer, “Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 1989;28:193–213.” 1989.
- [82] A. Jiménez, E. Monteverde, A. Nenclares, and G. Esquivel, “Confiabilidad y análisis factorial de la versión en español del Índice de Calidad De Sueño de Pittsburgh en pacientes psiquiátricos TT - Reliability and factorial analysis of the Spanish version of the Pittsburg Sleep Quality Index among psychiatric pati,” *Gac Med Mex*, vol. 144, no. 6, pp. 491–496, 2008.
- [83] A. T. Beck, M. Kovacs, and A. Weissman, “Assessment of suicidal intention: The Scale for Suicide Ideation.” *Journal of Consulting and Clinical Psychology*, vol. 47, no. 2, pp. 343–352, 1979.

- [84] S.ález Macip@, A.íaz Martínez@, S.ón@, C.ález Forteza@, and J.ús González Núñez@, “Psychometric characteristics of Beck’s suicidal ideation scale (BSI) in university students of Mexico city [Características psicometricas de la escala de ideacion suicida de Beck (ISB) en estudiantes universitarios de la ciudad de Mexico],” *Salud Mental*, vol. 23, no. 2, pp. 21–30, 2000.
- [85] Beck, Brown, Epstein, and Steer, “An inventory for measuring clinical anxiety: Psychometric properties. Journal of Consulting and,” *Clinical Psychology*, vol. 56, no. 6, pp. 893–897, 1988.
- [86] N. Aguiar, D. Meira, and S. Raquel, “Study on the efficacy of the Portuguese cooperative taxation,” *REVESCO Revista de Estudios Cooperativos*, vol. 121, no. 22, pp. 7–32, 2015.
- [87] D. Watson, L. Clark, and A. Tellegen, “Development and validation of brief measures of positive and negative affect: the panas scales.,” *Journal of personality and social psychology*, vol. 54 6, pp. 1063–70, 1988.
- [88] J. Ortuño-sierra, M. Santarén-rosell, A. P. D. Albéniz, and E. Fonseca-pedrero, “Supplemental Material for Dimensional Structure of the Spanish Version of the Positive and Negative Affect Schedule (PANAS) in Adolescents and Young Adults,” *Psychological Assessment*, vol. 27, no. 3, pp. 1–9, 2015.
- [89] T. Brown, P. D. D. Nardo, C. Lehman, and L. Campbell, “Reliability of dsm-iv anxiety and mood disorders: implications for the classification of emotional disorders.,” *Journal of abnormal psychology*, vol. 110 1, pp. 49–58, 2001.
- [90]
- [91] S. Taylor, M. J. Zvolensky, B. J. Cox, B. Deacon, R. G. Heimberg, D. R. Ledley, J. S. Abramowitz, R. M. Holaway, B. Sandin, S. H. Stewart, M. Coles, W. Eng, E. S. Daly, W. A. Arrindell, M. Bouvard, and S. J. Cardenas, “Robust Dimensions of Anxiety

- Sensitivity: Development and Initial Validation of the Anxiety Sensitivity Index-3,” *Psychological Assessment*, vol. 19, no. 2, pp. 176–188, 2007.
- [92] B. Sandín, R. M. Valiente, P. Chorot, and M. A. Santed Germán, “ASI-3 : Nueva escala para la evaluación de la sensibilidad a la ansiedad,” *Revista de Psicopatología y Psicología Clínica*, vol. 12, no. 2, pp. 91–104, 2007.
- [93] A. S. Radomsky, S. Rachman, D. S. Thordarson, H. K. McIsaac, and B. A. Teachman, “The Claustrophobia Questionnaire,” *Journal of Anxiety Disorders*, vol. 15, no. 4, pp. 287–297, 2001.
- [94] M. Amparo, M. Valls, A. García, C. Botella, and U. J. I. D. Castellón, “Propiedades psicométricas del cuestionario de claustrofobia,” vol. 15, pp. 673–678, 2003.
- [95] M. B. Mas, A. M. L. Jiménez, and M. Á. P. S. Gregorio, “Blood-injection Phobia Inventory ( BIPI ): Development , reliability and validity,” *Anales de psicología*, vol. 26, no. 1, pp. 58–71, 2010.
- [96] W. J. van Overveld, P. J. de Jong, M. L. Peters, K. Cavanagh, and G. C. Davey, “Disgust propensity and disgust sensitivity: Separate constructs that are differentially related to specific fears,” *Personality and Individual Differences*, vol. 41, no. 7, pp. 1241–1252, 2006.
- [97] B. Sandin, P. Chorot, M. Olmedo, and R. Valiente, “Escala de propensión a la sensibilidad a los miedos y a los trastornos compulsivos,” *Análisis y Modificación de Conducta*, vol. 34, pp. 127–168, 062008.
- [98] D. Watson and R. Friend, ““Measurement of Social-Evaluative Anxiety”: Erratum,” *Journal of Consulting and Clinical Psychology*, vol. 35, no. 1 PART 1, p. 69, 1970.

- [99] L. J. García-López, J. Olivares, M. D. Hidalgo, D. C. Beidel, and S. M. Turner, "Psychometric Properties of the Social Phobia and Anxiety Inventory, the Social Anxiety Scale for Adolescents, the Fear of Negative Evaluation Scale, and the Social Avoidance and Distress Scale in an Adolescent Spanish-Speaking Sample," *Journal of Psychopathology and Behavioral Assessment*, vol. 23, no. 1, pp. 51–59, 2001.
- [100] D. G. R. Spielberger Charles D, *Inventario de ansiedad: Rasgo-Estado IDARE*. Manuel moderno, 1975-2002.
- [101] D. L. Chambless, B. A. Sharpless, D. Rodriguez, K. S. McCarthy, B. L. Milrod, S.-R. Khalsa, and J. P. Barber, "Psychometric properties of the mobility inventory for agoraphobia: Convergent, discriminant, and criterion-related validity," *Behavior Therapy*, vol. 42, no. 4, pp. 689–699, 2011.
- [102] H. Villa Martin, C. Botella, A. Garc√ ≠ a – Palacios, and J. Osma, "Virtual reality exposure in the treatment of panic disorder with agoraphobia : A case study," *Cognitive and Behavioral Practice*, vol. 14, no. 1, pp. 58 – –69, 2007.
- [103] J. Wolpe and P. J. Lang, "A fear survey schedule for use in behaviour therapy," *Behaviour Research and Therapy*, vol. 2, no. 1, pp. 27–30, 1964.
- [104] D. Chambless, G. Caputo, P. Bright, and R. Gallagher, "Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire," *Journal of consulting and clinical psychology*, vol. 52 6, pp. 1090–7, 1984.
- [105] M. G. Lorenzo, W. P. Castro, C. T. P. González, J. M. B. Pérez, J. D. L. F. Portero, and R. G. Marco, "Efficacy of virtual reality exposure therapy combined with two pharmacotherapies in the treatment of agoraphobia," *International Journal of Clinical and Health Psychology*, vol. 11, pp. 189–203, 2011.

- [106] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [107] Y. Zhao, Z. Nasrullah, and Z. Li, “Pyod: A python toolbox for scalable outlier detection,” *Journal of Machine Learning Research*, vol. 20, no. 96, pp. 1–7, 2019.