



UNIVERSIDAD AUTÓNOMA DE ZACATECAS

---

UNIDAD ACADÉMICA DE INGENIERÍA ELÉCTRICA  
DOCTORADO EN CIENCIAS DE LA INGENIERÍA

RECONOCIMIENTO Y CLASIFICACIÓN DE  
ACTIVIDADES INFANTILES UTILIZANDO SONIDO  
AMBIENTAL

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

DOCTOR EN CIENCIAS

PRESENTA:

ANTONIO DE JESÚS GARCÍA DOMÍNGUEZ

ASESORES:

DR. CARLOS ERIC GALVÁN TEJADA

DR. JORGE ISAAC GALVÁN TEJADA



Zacatecas, Zacatecas, 2020

# Dedicatoria

*A mi esposa, mi hija, mi hijo (en camino), mis padres, mis hermanos y a todos los que  
han formado parte de mi vida...*

# Agradecimientos

A mi esposa, Miriam Méndez Saucedo, por su amor, por todo el apoyo que me brindó, su motivación, su confianza y su fe en mí. Por soportar mis cambios de humor, mi estrés, mi cansancio, todo fue más sencillo con ella a mi lado, alentándome siempre.

A mi hija Grettel Valeria García Méndez, por inyectarme esa alegría que siempre está dibujada en su rostro, por hacerme reír cuando más cansado estaba y más lo necesitaba, por su amor y sus abrazos, que aunque ella solo tiene 4 años, transmiten la paz que necesité en momentos difíciles.

A mis padres, Francisco Javier García García y Ma Francisca Domínguez Muñoz, por siempre apoyarme en cada etapa de mi vida, por todas sus enseñanzas, sus consejos y ejemplo. Aunque mi padre ya no está aquí para presenciar esta etapa de mi vida, él es parte de esto.

A mis hermanos, por siempre estar a mi lado cuando los necesito, por alentarme, regañarme, aconsejarme, todos siempre han sido parte de mis logros.

A a mi asesor, el Dr. Carlos Eric Galván Tejada, por toda la orientación brindada en estos años, por el tiempo invertido en mi trabajo y toda la paciencia y confianza hacia mí.

Al Dr. Jorge Isaac Galván Tejada por su disposición, apoyo y tiempo brindado en las diferentes etapas de este proyecto.

Al Dr. Hamurabi Gamboa Rosales, por el apoyo brindado para la publicación de trabajos producto de este proyecto.

Al Dr. Huizilopoztli Luna García, por su tiempo, disposición y comentarios hechos en las distintas revisiones de avances presentadas a lo largo de este tiempo

Al Dr. José María Celaya Padilla, por el tiempo dedicado para aceptar formar parte del comité.

Al Dr. Sodel Vázquez Reyes, responsable del Programa de Ingeniería de Software por siempre brindarme las facilidades y todo su apoyo para poder iniciar y concluir esta etapa.

Al Dr. Ismael de la Rosa, responsable del Doctorado en Ciencias de la Ingeniería, por todo el apoyo y las facilidades brindadas durante mi estancia en este Programa Académico.

A la Universidad Autónoma de Zacatecas, por el apoyo brindado mediante la autorización de la descarga laboral mientras me encontraba realizando estudios de posgrado.

A todas y cada una de las personas que han estado a mi alrededor, no solo en estos años, sino desde siempre.

# Índice general

<b>Dedicatoria</b>	<b>I</b>
<b>Agradecimientos</b>	<b>II</b>
<b>1 Introducción</b>	<b>1</b>
§1.1 Planteamiento del problema . . . . .	1
§1.2 Hipótesis de Investigación . . . . .	5
§1.3 Preguntas de Investigación . . . . .	5
§1.4 Objetivos de la Tesis . . . . .	6
§1.4.1 Objetivo General . . . . .	6
§1.4.2 Objetivos Específicos . . . . .	6
§1.5 Contribuciones de la Tesis . . . . .	7
§1.6 Organización de la Tesis . . . . .	7
<b>2 Estado del Arte</b>	<b>9</b>
§2.1 Introducción . . . . .	9
§2.2 Clasificación de Actividades Humanas . . . . .	10
§2.2.1 Recolección de Datos . . . . .	10
§2.2.2 Extracción de Características . . . . .	13
§2.2.3 Entrenamiento del Modelo . . . . .	14
§2.2.4 Clasificación de la Actividad . . . . .	15
§2.3 Modelos de Clasificación . . . . .	15
§2.4 Clasificación de Actividades Infantiles . . . . .	17
§2.5 Aprendizaje automático . . . . .	20
§2.5.1 Algoritmos Clasificadores . . . . .	22

§2.6	Resumen . . . . .	25
<b>3</b>	<b>Extracción y Selección de Características</b>	<b>27</b>
§3.1	Introducción . . . . .	27
§3.2	Dataset . . . . .	28
§3.2.1	Dispositivos de Grabación . . . . .	29
§3.2.2	Metadatos . . . . .	29
§3.3	Extracción de Características . . . . .	30
§3.3.1	Características de Audio . . . . .	30
§3.3.2	Proceso de Extracción de Características . . . . .	31
§3.4	Selección de Características . . . . .	33
§3.4.1	Métodos de Selección de Características . . . . .	34
§3.4.2	Proceso de Selección de Características . . . . .	36
§3.5	Resumen . . . . .	49
<b>4</b>	<b>Generación de los Modelos de Clasificación</b>	<b>51</b>
§4.1	Introducción . . . . .	51
§4.2	Descripción de los modelos . . . . .	52
§4.2.1	Modelos generados con el Dataset original . . . . .	52
§4.2.2	Modelos generados con el Dataset reducido . . . . .	55
§4.3	Resumen . . . . .	59
<b>5</b>	<b>Discusión</b>	<b>61</b>
§5.1	Clasificación de actividades infantiles mediante sonido ambiental . . . . .	61
§5.2	Modelos de clasificación generados con el conjunto completo de caracaterísticas	61
§5.3	Modelos de clasificación generados con selección de características . . . . .	63
§5.3.1	Modelos generados con características seleccionadas mediante crite- rio de Akaike . . . . .	64
§5.3.2	Modelos generados con características seleccionadas mediante Algo- ritmos Genéticos . . . . .	64
§5.4	Limitaciones Metodológicas . . . . .	65

<b>6 Conclusiones y Trabajo Futuro</b>	<b>66</b>
§6.1 Conclusiones . . . . .	66
§6.2 Publicaciones . . . . .	67
§6.3 Trabajo Futuro . . . . .	68

# Índice de figuras

2.1	Metodología general para la Clasificación de Actividades Humanas. . . . .	10
3.1	Evolución de los modelos generados por Galgo para el método de clasificación k-NN. . . . .	41
3.2	Evolución de los modelos generados por Galgo para el método de clasificación NC. . . . .	42
3.3	Evolución de los modelos generados por Galgo para el método de clasificación ANN. . . . .	42
3.4	Evolución de los modelos generados por Galgo para el método de clasificación RF. . . . .	43
3.5	Evolución de los modelos generados por Galgo para el método de clasificación Rpart. . . . .	43
3.6	Evolución del rango con el clasificador k-NN para las características más frecuentes. . . . .	44
3.7	Las cinco características elegidas para el clasificador k-NN según su frecuencia.	45
3.8	Evolución del rango con el clasificador NC para las características más frecuentes. . . . .	45
3.9	Las cinco características elegidas para el clasificador NC según su frecuencia.	46
3.10	Evolución del rango con el clasificador ANN para las características más frecuentes. . . . .	46
3.11	Las cinco características elegidas para el clasificador ANN según su frecuencia.	47
3.12	Evolución del rango con el clasificador RF para las características más frecuentes. . . . .	47
3.13	Las cinco características elegidas para el clasificador de RF según su frecuencia.	48



3.14 Evolución del rango con el clasificador Rpart para las características más  
frecuentes. . . . . 48

3.15 Las cinco características elegidas para el clasificador Rpart según su frecuencia. 49

4.1 Gráfico de comportamiento de la precisión durante el entrenamiento y vali-  
dación de la Red Neuronal Artificial en la Fase de Experimentación 5. . . . 58

4.2 Gráfica del comportamiento de la función de pérdida durante el entrena-  
miento y validación de la Red Neuronal Artificial en la Fase de Experimen-  
tación 5. . . . . 59

# Índice de tablas

2.1	Características de audio comúnmente utilizadas en modelos de Clasificación de Actividades Humanas. . . . .	13
3.1	Descripción general de las actividades analizadas. . . . .	28
3.2	Clips de audio por actividad. . . . .	29
3.3	Dispositivos de Grabación. . . . .	29
3.4	Metadatos de los clips de audio. . . . .	30
3.5	Características Extraídas. . . . .	32
3.6	Características seleccionadas mediante el AIC . . . . .	38
3.7	Configuración general de parámetros experimentales para el algoritmo genético “GALGO” . . . . .	40
3.8	Subconjunto de cinco características elegido por el algoritmo genético para cada clasificador. . . . .	44
4.1	Precisión obtenida por los modelos en la fase de experimentación 1. . . . .	53
4.2	Verdaderos positivos por actividad para cada modelo de clasificación en la Fase de Experimentación 2. . . . .	54
4.3	Precisión obtenida por actividad para cada modelo de clasificación en la Fase de Experimentación 2. . . . .	54
4.4	Precisión promedio para cada modelo de clasificación en la Fase de Experimentación 2. . . . .	54
4.5	Verdaderos positivos por actividad para cada modelo de clasificación utilizando el conjunto de 27 características extraído mediante AIC en la Fase de Experimentación 3. . . . .	55

4.6	Precisión por actividad para cada modelo de clasificación utilizando el conjunto de 27 características extraído mediante AIC en la Fase de Experimentación 3. . . . .	56
4.7	Precisión promedio para cada modelo de clasificación en la Fase de Experimentación 3 utilizando el conjunto de 27 características extraído mediante AIC . . . . .	56
4.8	Precisión promedio para cada modelo de clasificación en la Fase de Experimentación 4 utilizando el conjunto de 5 características extraído mediante Algoritmos Genéticos. . . . .	57
4.9	Arquitectura propuestas para la Red Neuronal Artificial en la Fase de Experimentación 5. . . . .	57
4.10	Parámetros seleccionados para la implementación de la Red Neuronal Artificial en la Fase de Experimentación 5. . . . .	58

# Capítulo 1

## Introducción

En este capítulo se describe de manera detallada el contexto sobre el cual se desarrolla el presente trabajo a cerca del reconocimiento y clasificación de actividades infantiles utilizando sonido ambiental, como propuesta de tema para la tesis doctoral. A su vez, se presenta el planteamiento específico del problema, analizando los factores que influyen en él y las consideraciones a tomar en cuenta. Se describen además, de manera breve, las soluciones propuestas a través de este trabajo para abordar el problema aquí tratado, mencionando los métodos aplicados para llegar a ellas. También se muestra la hipótesis de investigación, así como el objetivo general y los objetivos específicos. En la parte final del capítulo se presentan las contribuciones hechas con la realización del presente trabajo y la forma en la que está estructurado este documento.

### 1.1. Planteamiento del problema

La Inteligencia Artificial es una área de la tecnología que surge a partir de diversos trabajos desarrollados a mediados del siglo XX [16]. En estos trabajos, principalmente a partir de los presentados por el matemático británico Alan Turing, se comienza a despertar el interés por la idea de lograr que las máquinas puedan igualar la capacidad de pensar del ser humano [102]. A partir de esa época, y con el continuo avance de la tecnología, la Inteligencia Artificial ha sido una área en constante desarrollo, logrando tener influencia en diversos campos de aplicación, mediante la generación de soluciones y propuestas que

tienen como finalidad facilitar la interacción del ser humano con su entorno.

El éxito y gran avance de la Inteligencia Artificial en las últimas décadas ha provocado el surgimiento de distintas subáreas con la finalidad de aplicar soluciones a problemas específicos, atendiendo cada vez un mayor número de necesidades y áreas de oportunidad. En este sentido, cada vez son más los trabajos enfocados a la aplicación de métodos de Inteligencia Artificial en áreas donde previamente esta no había tenido influencia, o había sido muy limitada, logrando de esta manera incidir de forma importante y fundamental en el desarrollo tecnológico de los últimos años.

Una importante subárea de la Inteligencia Artificial en donde ha habido grandes avances y se han desarrollado numerosas aplicaciones en los últimos años es la Inteligencia Ambiental [92, 26]. La Inteligencia Ambiental se caracteriza por dirigir sus trabajos e investigaciones a la creación de espacios donde los usuarios interaccionen de forma natural con su entorno mediante el desarrollo de aplicaciones, interfaces y sistemas inteligentes que resulten invisibles para el usuario, integrándolos con los objetos de uso cotidiano [98]. La Inteligencia ambiental ha sido fundamental en los últimos años para el desarrollo de trabajos enfocados a los conceptos de Internet de las Cosas (IoT) y Ciudades Inteligentes (Smart Cities) [24, 135, 5, 10, 127, 71].

Como se mencionó anteriormente, la Inteligencia Ambiental es una subárea de la Inteligencia Artificial enfocada a brindar soluciones para facilitar la interacción de los usuarios con su entorno. Uno de los temas que ha generado interés en los últimos años es el reconocimiento de actividades humanas (HAR), en el cual el objetivo es determinar la actividad o conjunto de actividades que están siendo realizadas por una persona o grupo de personas en cierto instante, mediante el correcto análisis de datos obtenidos a partir de determinada fuente de datos. Los trabajos desarrollados sobre este tema se enfocan en poder brindar algún tipo de servicio al usuario una vez detectada la actividad que se encuentra realizando.

Para la generación de un modelo de reconocimiento y clasificación de actividades hu-

manas es necesario definir algunos aspectos importantes en los cuales se basa el análisis a realizar:

- Grupo de Interés. Antes de plantear otros aspectos del tipo de análisis a realizar, es necesario identificar el grupo de personas al cual está dirigido el modelo de reconocimiento y clasificación de actividades. Acotar de manera precisa el grupo de interés ayuda a definir de manera más sencilla otros parámetros del modelo, como tipo de datos a analizar, fuente de datos, algoritmos y métodos utilizados, etc.
- Fuente de Datos. Otro parámetro importante a definir para el modelo de clasificación de actividades es la fuente de datos a utilizar. La fuente de datos es el medio que provee datos al modelo y, dependiendo del grupo de interés y tipo de modelo a desarrollar, algunas fuentes de datos son más convenientes que otras. Entre las fuentes de datos más comunes utilizadas en sistemas de reconocimiento y clasificación de actividades humanas se encuentran: audio, video, imágenes y sensores. La estrategia de recolección de datos para el modelo depende del tipo de dato utilizado.

Una vez que se tiene definido el grupo de personas a las cuales está dirigido el estudio y el tipo de datos con los que se va a trabajar, es importante también definir el tratamiento y tipo de análisis que se hará con los datos recolectados para la generación del modelo de clasificación. En los últimos años, se han utilizado diversas técnicas de Aprendizaje Automático para aplicaciones de Inteligencia Ambiental, ya que es posible automatizar, mediante distintos algoritmos, la identificación de patrones o tendencias presentes en los datos recolectados de cierto fenómeno [57, 83, 87].

Los tipos de implementación de Aprendizaje Automático pueden clasificarse en dos categorías:

1. Aprendizaje Supervisado. En el Aprendizaje Supervisado, los algoritmos trabajan con datos “etiquetados”, intentando encontrar una función que, dadas las variables de entrada, les asigne la etiqueta de salida adecuada. El algoritmo se entrena con un “histórico” de datos y así “aprende” a asignar la etiqueta de salida adecuada a un

nuevo valor, es decir, predice el valor de salida [107, 74]. El Aprendizaje Supervisado se suele utilizar en los problemas de clasificación y problemas de regresión, los cuales se diferencian uno del otro por el tipo de variable objetivo. En los casos de clasificación, la variable objetivo es de tipo categórico, mientras que en los casos de regresión es de tipo numérico [108].

2. Aprendizaje no Supervisado. En el Aprendizaje no Supervisado no se dispone de datos “etiquetados” para el entrenamiento, solo se conocen los datos de entrada. Este tipo de aprendizaje tiene un carácter “exploratorio” ya que solo es posible describir la estructura de los datos de entrada para intentar encontrar algún tipo de organización que simplifique el análisis (por ejemplo, agrupamiento basado en similitudes) [60, 74].

En el capítulo 2 se presentan trabajos relacionados a los conceptos anteriormente mencionados involucrados en la generación de modelos de reconocimiento y clasificación de actividades humanas.

Se han desarrollado numerosos trabajos y propuestas en el tema de reconocimiento y clasificación de actividades humanas, tomando en cuenta las posibles variables que se pueden definir. Dependiendo del grupo de personas a las cuales vaya dirigido el análisis, se tienen distintas consideraciones a tomar en cuenta y se atacan distintas problemáticas. Los modelos de clasificación de actividades infantiles regularmente están enfocados a ofrecer soluciones referentes a la seguridad de los niños, ya que es un grupo vulnerable y muy susceptible a accidentes.

La presente tesis se enfoca en la generación de un modelo de clasificación de actividades infantiles mediante el sonido ambiental. Para acotar de una mejor manera el grupo de individuos a los que está dirigido el análisis, se enfoca a niños de entre 12 y 36 meses, debido a que en ese rango de edad realizan actividades similares, además de ser precisamente en ese rango de edad cuando son más propensos a sufrir accidentes, por lo que un sistema de reconocimiento de actividades ayudaría a detectar actividades potencialmente peligrosas para ellos. Los modelos de análisis de actividades en niños es un tema poco explorado, y

las propuestas que se han desarrollado utilizan fuentes de datos que pueden interferir en el comportamiento natural de los niños (por ejemplo, sensores incrustados en las prendas), causando que el proceso de recolección de datos no se lleve a cabo de manera adecuada. Debido a eso, en este trabajo se aplican técnicas de Aprendizaje Automático y análisis del sonido ambiental para la generación de un modelo de clasificación de actividades en niños, puesto que esta fuente de datos no interfiere con las actividades mientras los datos son capturados.

## 1.2. Hipótesis de Investigación

La hipótesis de investigación para el presente trabajo es:

*A través de la utilización de datos del sonido ambiental se puede reconocer y clasificar actividades correctamente en un 70 % en niños de 12 a 36 meses mediante la generación de un modelo clasificador desarrollado con técnicas de aprendizaje automático supervisado.*

## 1.3. Preguntas de Investigación

Una vez revisados los aspectos a tomar en cuenta para la generación de un modelo de reconocimiento y clasificación de actividades infantiles, se puso especial interés en la fuente de datos a utilizar. Se identificó que el sonido ambiental es una fuente de datos que tiene la ventaja de que el proceso de captura de datos es completamente invisible para los niños. En base a eso surgieron las siguientes preguntas de investigación:

- ¿Qué tipo de información provee el sonido ambiental para su análisis en la generación de un modelo de clasificación de actividades infantiles?
- ¿Cuáles son las características posibles a extraer del sonido ambiental para la generación del modelo de clasificación?
- ¿Que subconjunto de características describe mejor las actividades analizadas en el modelo de clasificación?



Además, como se mencionó anteriormente en este documento, se plantea la aplicación de técnicas de Aprendizaje Automático supervisado para la generación del modelo de clasificación. Este sentido, surgen las siguientes preguntas:

- ¿Cuáles algoritmos clasificadores pueden ser aplicados a la generación del modelo de clasificación?
- ¿Qué métodos de selección de características eficientan el modelo de clasificación?
- ¿Cuáles algoritmos clasificadores logran una mayor precisión en la clasificación de actividades infantiles una vez desarrollado el modelo?
- ¿Cuál es el comportamiento del modelo de clasificación al combinar distintas técnicas de selección de características con los distintos algoritmos clasificadores implementados?
- ¿Qué técnicas o métodos complementarios pueden utilizarse para mejorar la precisión del modelo de clasificación?

## 1.4. Objetivos de la Tesis

Se presentan los siguientes objetivos de investigación.

### 1.4.1. Objetivo General

El objetivo general del presente trabajo es: Reconocer y clasificar correctamente actividades realizadas por niños de 12 a 36 meses de edad utilizando el sonido ambiental como fuente de datos.

### 1.4.2. Objetivos Específicos

Para cumplir con el objetivo general, se plantean los siguientes objetivos específicos:

- Definir el conjunto de actividades a tomar en cuenta durante el análisis de datos.

- Recolectar las señales de audio mediante las herramientas o dispositivos seleccionados.
- Extraer las características de las señales de audio.
- Aplicar mecanismos de selección de características para elegir las que mejor describan a las actividades analizadas.
- Optimizar la cantidad de datos requeridos para el análisis.
- Generar e implementar un modelo de análisis y clasificación de los datos obtenidos para las actividades definidas mediante técnicas de Aprendizaje Automático supervisado.
- Evaluar el modelo de clasificación generado con cada algoritmo clasificador implementado y tomando en cuenta el proceso previo de selección de características.

## 1.5. Contribuciones de la Tesis

- Análisis del sonido ambiental como fuente de datos para clasificación de actividades infantiles.
- Selección de características del sonido ambiental para generación de modelo de clasificación.
- Aplicación de métodos de Aprendizaje Automático supervisado en la generación de un modelo de clasificación de actividades infantiles utilizando sonido ambiental.

## 1.6. Organización de la Tesis

Los capítulos siguientes de este trabajo de tesis se organizan de la siguiente manera:

**Capítulo 2.** En el capítulo 2 se presentan los conceptos involucrados en la generación de modelos de clasificación de actividades humanas, especialmente los presentes en trabajos

enfocados a niños. Además se mencionan las tecnologías implementadas y los métodos de Aprendizaje Automático comúnmente utilizados para este tipo de trabajos. De forma paralela también se muestra un recorrido por los trabajos relacionados a la clasificación de actividades humanas y clasificación de actividades infantiles, contrastando las propuestas que se han hecho en base a fuentes de datos utilizadas, tecnologías, métodos, algoritmos, etc.

**Capítulo 3.** En este capítulo se describen las actividades tomadas en cuenta para la generación del modelo de clasificación, además se presentan los métodos de recolección de datos y la forma en la que se compone el dataset utilizado para el análisis. Aquí mismo se presentan los métodos de extracción de características aplicados, y la optimización de datos propuesta mediante el proceso de selección de características, mencionando los algoritmos implementados.

**Capítulo 4.** Este capítulo muestra la forma en la que se generan los modelos de clasificación de actividades infantiles, tomando en cuenta los distintos algoritmos de clasificación implementados y comparando la generación de modelos con subconjuntos de características resultantes del proceso de selección.

**Capítulo 5.** En el capítulo 5 se presentan los resultados obtenidos a partir de la evaluación de los modelos de clasificación generados.

**Capítulo 6.** Aquí se muestra la discusión de los resultados obtenidos para los distintos modelos de clasificación evaluados, comparando sus resultados.

**Capítulo 7.** En el capítulo final de esta tesis se presentan las conclusiones a las que se llega después de la realización de este trabajo de investigación, además de las contribuciones obtenidas y el trabajo futuro.

# Capítulo 2

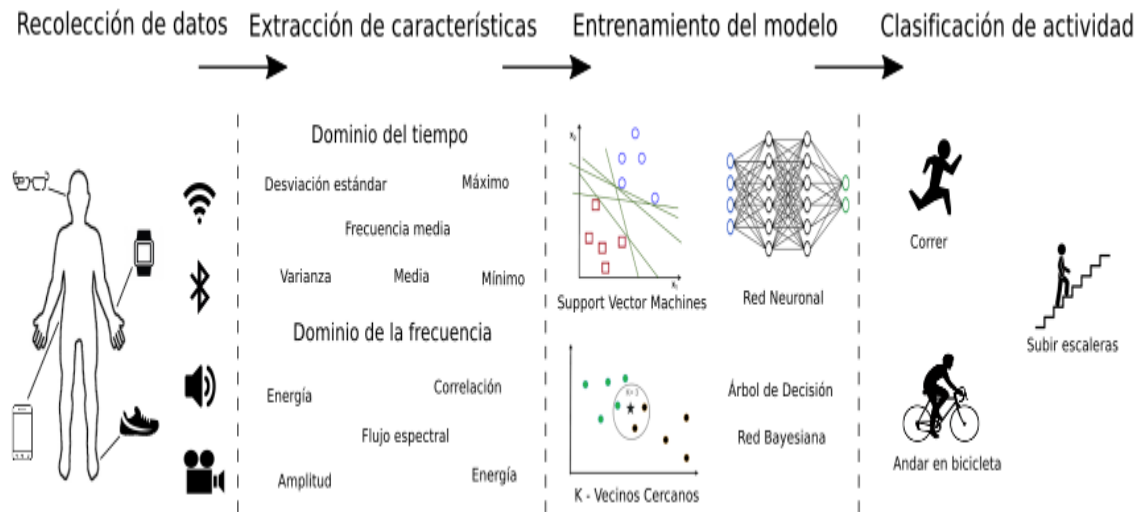
## Estado del Arte

En este capítulo se presentan los conceptos involucrados en el tema de Inteligencia Ambiental y reconocimiento de actividades humanas, sobre los cuales se basa el presente trabajo de investigación, además de aquellos conceptos específicos para el área de reconocimiento y clasificación de actividades infantiles. También se mencionan los métodos de Aprendizaje Automático utilizados en la generación del modelo de clasificación, así como técnicas de selección de características aplicadas y tecnologías comúnmente utilizadas para este tipo de trabajos. De forma paralela a la descripción de los conceptos y métodos, se mencionan trabajos relacionados para contrastar las propuestas y trabajos desarrollados por otros autores con la propuesta mostrada en el presente trabajo.

### 2.1. Introducción

Con el continuo y gran avance de la tecnología en los últimos años, la Inteligencia Ambiental ha sido una área que se ha visto enormemente favorecida. Mediante el desarrollo y surgimiento de nuevos dispositivos y tecnologías, ha sido posible la creación de diversas aplicaciones en esta área, las cuales han impactado de forma positiva en la sociedad y vida de los seres humanos. La Inteligencia Ambiental ha hecho posible el desarrollo de conceptos y disciplinas innovadoras como lo son las Ciudades Inteligentes (Smart Cities) y el Internet de las Cosas (IoT), eficientando las actividades de los individuos mediante aplicaciones que facilitan la interacción con su entorno [35, 47, 104, 13, 48, 93, 30, 100].

Figura 2.1: Metodología general para la Clasificación de Actividades Humanas.



## 2.2. Clasificación de Actividades Humanas

Entre los distintos temas que abarca la Inteligencia Ambiental, uno que ha tenido gran relevancia en los últimos años y sobre el cual se han desarrollado numerosos trabajos es el de clasificación de actividades humanas. El reconocimiento y clasificación de actividades humanas busca, a partir de cierta fuente de datos, detectar la actividad que está siendo realizada por determinada persona o grupo de personas, mediante el análisis de un conjunto de datos recolectados. De manera general, la metodología para la generación de un modelo de reconocimiento y clasificación de actividades humanas se muestra en la Figura 2.1 y se describe a continuación:

### 2.2.1. Recolección de Datos

Cuando se realiza el planteamiento de un modelo de reconocimiento y clasificación de actividades humanas existen varios aspectos a tomar en cuenta. Un elemento fundamental para el análisis es la fuente de datos a utilizar, ya que de ello dependerá el uso de otros elementos para la generación del modelo. Dependiendo de la fuente de datos utilizada se define la forma de recolección de datos para el modelo.

### **Sensores Embebidos**

Una de las fuentes de datos más utilizadas en los sistemas de reconocimiento y clasificación de actividades humanas son los sensores embebidos, los cuales son colocados en las prendas de la persona para medir alguna característica de interés o detectar algún tipo de comportamiento o actividad. Los sensores mayormente utilizados para este fin son los acelerómetros, sensores de presión, sensores de temperatura y dispositivos de radiofrecuencia (RFID) [9, 140, 54, 32, 118, 78, 44, 55, 137].

### **Teléfonos Inteligentes**

El continuo avance en la tecnología de los teléfonos inteligentes ha hecho posible que éstos puedan fungir como fuente de datos en los sistemas de reconocimiento y clasificación de actividades, valiéndose de los diversos sensores con los que cuentan. Los principales sensores presentes en los teléfonos inteligentes utilizados como fuente de datos son:

- GPS (Sistema de Posicionamiento Global)
- Acelerómetro
- Sensor de humedad
- Micrófono
- Sensor de temperatura
- Giroscopio

El tipo de sensor utilizado depende del modelo de clasificación para el que se recolectan los datos, pudiendo ser más conveniente el uso de uno u otro sensor. De manera general, los trabajos de reconocimiento y clasificación de actividades humanas que utilizan algún tipo de sensor presente en los teléfonos inteligentes han venido en aumento en los últimos años [94, 96, 106, 44, 21, 56, 123, 21].

## Audio

El audio, específicamente el sonido ambiental, ha sido utilizado como fuente de datos en distintos trabajos sobre reconocimiento y clasificación de actividades humanas [38, 39, 18, 22], ya que provee información suficiente para llevar a cabo el análisis. A diferencia de otras fuentes de datos, el audio tiene la ventaja de que su captura puede pasar inadvertida para los individuos sujetos al análisis, ya que no es necesario que se porte algún dispositivo como sucede con otras fuentes de datos, si no que simplemente pueden situarse micrófonos en el área donde se estén analizando las actividades.

Esta fuente de datos generalmente se utiliza en modelos de reconocimiento de actividades humanas desarrollados para hogares inteligentes (Smart Homes), ya que resulta sencillo ubicar los sensores de audio de manera estratégica dependiendo de las actividades que se quieran analizar. Comúnmente, este tipo de sistemas busca detectar la actividad realizada por un individuo para brindar algún tipo de servicio.

El uso de los micrófonos presentes en los teléfonos inteligentes se ha vuelto común cuando se trabaja con el audio como fuente de datos para un sistema de reconocimiento y clasificación de actividades humanas, ya que al hacer uso de estos micrófonos no se requiere de dispositivos extra para el proceso de captura de datos, reduciendo así los costos al utilizar dispositivos con los que las personas cuentan con mayor frecuencia cada día.

## Video

Como se ha mencionado anteriormente, el reconocimiento y clasificación de actividades humanas es un tema que puede aplicarse en distintas áreas. Muchos de los trabajos desarrollados enfocados a sistemas de monitoreo y vigilancia se basan en la utilización de video como fuente de datos [52, 88, 51, 111, 53, 69]. Los modelos de este tipo principalmente buscan predecir actividades o comportamientos de los individuos en base a videos capturados de actividades realizadas.

Tabla 2.1: Características de audio comúnmente utilizadas en modelos de Clasificación de Actividades Humanas.

Dominio del Tiempo	Dominio de la Frecuencia
RMS (Root Mean Square) Energy	Fundamental frequency Frequency components
ZCR (Zero Crossing Rate) Maximum amplitude	Spectral centroid Spectral flux
Minimum energy	Spectral density Spectral roll-off

### Dispositivos de Radio Frecuencia

El avance de la tecnología, especialmente en los sensores y la electrónica, ha provocado que cada vez sean más variadas las posibles fuentes de datos para los sistemas de reconocimiento y clasificación de actividades humanas. La tecnología de Identificación por Radio Frecuencia (RFID) ha sido utilizada en este tipo de modelos de manera exitosa, ya que mediante el uso de un dispositivo RFID es posible identificar con gran seguridad y precisión la posición del portador del dispositivo. Generalmente los dispositivos RFID se utilizan en combinación con otro tipo de sensores para aportar mayor información al modelo de reconocimiento y clasificación de actividades humanas [37, 125, 97, 72, 6, 134].

#### 2.2.2. Extracción de Características

Una característica de audio es cualquier aspecto medible cualitativa o cuantitativamente de un sonido, estas medidas son una serie de parámetros que representan de forma compacta la información del sonido y se pueden utilizar para caracterizarla. Mediante un correcto proceso de extracción, es posible obtener distintas características que se pueden utilizar para clasificar las señales de audio. Estas características pueden separarse en dos tipos: dominio del tiempo y dominio de la frecuencia. Algunas de las características comúnmente utilizadas para el análisis de audio en modelos de reconocimiento y clasificación de actividades humanas se muestran en la tabla 2.1.

Aunque es posible obtener un número considerable de características a partir de una



señal de audio, es importante considerar el tipo de análisis que se va a realizar sobre los datos recolectados, ya que ciertas características son más utilizadas que otras para este tipo de trabajos.

### 2.2.3. Entrenamiento del Modelo

Una vez que se han llevado a cabo las etapas de captura de datos y extracción de características, se tienen los elementos necesarios para la generación y entrenamiento del modelo de clasificación. Para la generación del modelo es necesario implementar un algoritmo de clasificación, que será el encargado de analizar los datos capturados y clasificarlos en las clases analizadas. Algunos de los algoritmos de clasificación más utilizados en modelos de reconocimiento y clasificación de actividades humanas son:

- k-Nearest Neighbors (k-NN) [101].
- Nearest Centroid (NC) [28].
- Random Forests (RF) [11].
- Artificial Neural Network (ANN) [70].
- Recursive Partitioning Tree (Rpart) [41].

Cada uno de los algoritmos clasificadores mencionados anteriormente se describen en secciones posteriores en este documento. Entrenar el modelo de clasificación consiste en “alimentarlo” con un conjunto de datos preetiquetados para su análisis, de forma que el algoritmo identifique el “comportamiento” de las características que hacen que una clase se distinga de las demás. Generalmente, de los datos recolectados en el proceso previo de recolección de datos, se destina un 70 % de ellos para la etapa de entrenamiento del modelo, como en los trabajos presentados por Galván et al. [40], Anguita et al. [7] y Reyes et al. [95].

### 2.2.4. Clasificación de la Actividad

Una vez que el modelo ha sido entrenado, es capaz de recibir nuevas muestras y clasificarlas en base al comportamiento que identificó en los datos durante el proceso de entrenamiento. Generalmente, de los datos recolectados en el proceso previo de recolección de datos, se destina un 30 % de ellos para la fase de prueba, a partir de la cual se obtiene la precisión del modelo para clasificar casos nuevos a las clases correctas.

## 2.3. Modelos de Clasificación

Como se ha mencionado anteriormente, al momento de desarrollar un modelo de reconocimiento y clasificación de actividades humanas uno de los aspectos más importantes a considerar es el grupo de personas a las que va dirigido el análisis, ya que de ello dependerá la elección de otros elementos como la fuente de datos, dispositivos de captura, etc. En el área de reconocimiento y clasificación de actividades humanas se han desarrollado trabajos para distintos grupos de interés, cada uno con sus particularidades y consideraciones propias.

Uno de los grupos de personas para los que se han desarrollado modelos de reconocimiento de actividades son los adultos mayores, ya que es un grupo normalmente vulnerable y para el cual un sistema de este tipo puede resultar útil para ofrecer soluciones a distintas problemáticas propias de su edad. Tomando en cuenta lo anterior, muchos de los trabajos desarrollados están enfocados al cuidado de personas mayores en hogares, en donde mediante un modelo de reconocimiento y clasificación de actividades es posible detectar comportamientos o actividades peligrosas para los individuos, ya sea por su edad o por determinada condición en la que se encuentren [89, 17, 8].

Khan et al. [61]. desarrollaron un sistema de reconocimiento y clasificación de actividades humanas basado en video para adultos mayores. En este sistema los autores analizan seis distintas actividades consideradas anormales para el grupo de personas al que está dirigido, las actividades son: caída hacia adelante, caída hacia atrás, dolor en el pecho,

desmayo, vómito y dolor de cabeza.

Ruan et al. [97] proponen un modelo de localización y reconocimiento de actividades, enfocado a personas adultas, haciendo uso de etiquetas RFID pasivas, con lo cual buscan atacar el problema de que tienen muchos otros modelos que utilizan sensores y baterías embebidas en las prendas de las personas o portadas directamente en alguna parte del cuerpo, lo cual resulta poco práctico en situaciones reales.

Otro de los grupos de individuos para el cual comúnmente se han desarrollado trabajos de reconocimiento y clasificación de actividades son las personas que padecen algún tipo de enfermedad. Dependiendo de la condición que padezcan los individuos sujetos al análisis, los sistemas de reconocimiento de actividades ayudan en la detección de actividades que puedan poner en peligro la salud de las personas, sirviendo de esta manera como apoyo en el monitoreo médico [29, 75, 80].

En este sentido, Guo et al. [43] desarrollaron un sistema de reconocimiento de actividades para pacientes con enfermedades como Alzheimer, Parkinson o epilepsia, basado en el acelerómetro y giroscopio presentes en los teléfonos inteligentes, con el cual obtienen una precisión superior al 80 % en el reconocimiento de actividades potencialmente peligrosas para los pacientes. Para lograr detectar la actividad, el paciente debe traer consigo el teléfono inteligente.

Cheng et al. [23] presentan un trabajo enfocado a pacientes con Parkinson para la detección de actividades utilizando el acelerómetro presente en los teléfonos inteligentes. Las actividades de interés para este trabajo son: caminar y cambiar de posición (sentarse y pararse), considerando el tiempo empleado en cada actividad.

El reconocimiento y clasificación de actividades humanas también se ha enfocado en el desarrollo de sistemas para hogares inteligentes, para detectar actividades realizadas por los ocupantes y brindar algún tipo de servicio [62, 112, 138]. Mehr et al. [82] desarrollaron

un sistema de reconocimiento de actividades para hogares inteligentes con el objetivo de detectar actividades simples como bañarse, arreglarse e ir a trabajar, mediante redes neuronales para el análisis y clasificación de los datos.

Chahuara et al. [18] presentaron un sistema de reconocimiento automático de actividades cotidianas para hogares, mediante la combinación de múltiples sensores, actuadores y equipos de automatización, incluyendo sensores de audio. Bianchi et al. [12] desarrollaron un sistema de reconocimiento de actividades humanas, haciendo uso de sensores portátiles y técnicas de aprendizaje profundo para el análisis de datos, enfocado al reconocimiento y la detección de actividades cotidianas dentro de un hogar inteligente. Los autores mencionan que el análisis se realiza sobre 9 distintas actividades y se logra una precisión del 97%.

También se han desarrollado sistemas de reconocimiento y clasificación de actividades humanas enfocados a la seguridad, especialmente en la vigilancia a través de video [121, 3, 20]. Karthikeswaran et al. [59] desarrollaron un modelo para detectar automáticamente comportamientos inusuales en videos de vigilancia, a través de un método de clasificación lineal adaptativa y marcos de Internet de las Cosas (IoT). Los autores reportan una precisión del 97% y mencionan que este sistema puede utilizarse en lugares generales, por ejemplo, centros comerciales, aeropuertos y estaciones de ferrocarril o en locales privados donde la seguridad es la principal preocupación.

Otro de los grupos de individuos para los cuales se han desarrollado trabajos en cuanto a reconocimiento y clasificación de actividades se refiere son los niños. La siguiente sección profundiza en este tema y describe algunos de los trabajos desarrollados en esta área.

## 2.4. Clasificación de Actividades Infantiles

Hoy en día es cada vez más común que tanto el padre como la madre en una familia trabajen fuera de casa, provocando que, en caso de tener hijos pequeños, estos estén al

cuidado de terceras personas, ya sea familiares o escuelas infantiles, y muchas de las veces puedan estar realizando actividades potencialmente peligrosas al no estar monitoreados en todo momento por un adulto. Inclusive al estar en casa, los niños son un grupo muy susceptible a accidentes domésticos, los cuales en algunos casos pueden ser fatales.

Según la Organización Mundial de la Salud (OMS) [1], algunas de las principales causas de muerte infantil son: accidentes de tráfico, ahogamientos, caídas y quemaduras. A excepción de los accidentes de tráfico, las otras causas suceden en su mayoría en el hogar y es por descuido de los padres. Es así como contar con un sistema que ayude a las familias a monitorear el comportamiento de los niños puede beneficiar significativamente, ya que ayuda a detectar comportamientos y actividades potencialmente peligrosas.

Se han desarrollado numerosos y diversos trabajos de reconocimiento y clasificación de actividades enfocados a niños. Boughorbel et al. [14] desarrollaron un sistema de reconocimiento de actividades infantiles mediante la combinación de múltiples sensores (principalmente acelerómetros) situados en un dispositivo portátil. Los autores analizan siete actividades comunes en niños: caminar, acostarse, correr, subir escaleras, caerse, pararse y otra. En este trabajo se reportan precisiones de hasta 99% para algunas de las actividades analizadas, esta precisión varía dependiendo de las características utilizadas para la clasificación.

Adaškevičius et al. [2] presentan un sistema de reconocimiento de actividades infantiles con la finalidad de detectar las actividades y periodos de inactividad que puedan provocar desórdenes en la salud de los niños. Este modelo basa su funcionamiento en un acelerómetro para el proceso de recolección de datos y puede analizar actividades como caminar, trotar, sentarse y saltar (duración y tipo de actividad). Los autores reportan una precisión de 87.56% en la clasificación de las actividades analizadas.

Westeyn [126] presentó un sistema de reconocimiento de actividades para controlar el progreso del desarrollo de los niños con juguetes aumentados, enfocado a detectar niños

con autismo, entre otro tipo de condiciones. En este trabajo, los juguetes se encuentran integrados con sensores como cámara, acelerómetro, micrófono, entre otros, que permitan capturar datos para su análisis.

Kushbu et al. [68] describen un trabajo de reconocimiento de actividades infantiles para niños de 16 a 29 meses de edad mediante el uso de tarjetas de radio frecuencia (RFID) y acelerómetros en la pierna, mano y cintura del cuerpo para prevenir accidentes infantiles como lesiones no intencionales en el hogar. Las actividades analizadas en este trabajo son: rodar, quedarse quieto, sentarse, caminar y gatear. Suzuki et al. [113] desarrollaron un sistema de apoyo para el personal y los padres de los jardines de niños, que ayude a vigilar las actividades que los infantes realicen. Para este trabajo los autores utilizaron un acelerómetro situado en la parte superior del brazo del niño.

Al analizar los trabajos desarrollados en el área de reconocimiento y clasificación de actividades infantiles es posible observar que la gran mayoría de ellos utilizan sensores incrustados en las prendas de los niños, lo cual presenta las siguientes desventajas:

- Interferencia. El hecho de que sea necesario que los sensores sean portados por los niños hace que en determinadas situaciones los propios sensores puedan dificultar la realización de las mismas actividades a analizar.
- Energía. Al utilizar sensores ubicados en las prendas de los niños, se vuelve necesario que las baterías se estén reemplazando o recargando constantemente, para asegurar un funcionamiento adecuado del proceso de captura de datos para el modelo.
- Daños. Debido a la propia naturaleza en el comportamiento de los niños, al utilizar sensores incrustados en las prendas, resulta común que estos puedan sufrir daños ocasionados por la misma realización de las actividades y la ubicación en la que se encuentran los sensores.
- Fiabilidad. Tomando en cuenta todos los factores mencionados en los puntos anteriores, al utilizar sensores incrustados en las prendas de los niños es posible que los

datos capturados presenten variaciones o no sean confiables al 100 %, ya que hay variables que pueden estar afectando la captura.

Como se mencionó anteriormente, los trabajos de reconocimiento y clasificación de actividades infantiles pueden tener la problemática de la fuente de datos que utilizan, ya que el proceso de captura de datos no se realiza de manera inadvertida para los niños. De ahí la importancia de utilizar una fuente de datos que no interfiera con el comportamiento natural de los niños al momento de estar realizando y analizando las actividades.

## 2.5. Aprendizaje automático

Como se ha mencionado anteriormente, es cada vez más común encontrar trabajos que implementan técnicas de Aprendizaje Automático en el área de Inteligencia Ambiental. Especialmente en el área de análisis de datos, el Aprendizaje Automático es utilizado cada vez más en temas de diagnóstico y predicción, debido a la característica de analizar datos de manera eficiente.

El Aprendizaje Automático se centra en el desarrollo de modelos que pueden cambiar o adaptarse cuando se alimentan con datos nuevos. El proceso de Aprendizaje Automático es en parte similar al de la minería de datos, ambos analizan datos para encontrar patrones. Sin embargo, en el Aprendizaje Automático este análisis no se realiza para presentar información para la comprensión humana, como lo es en la minería de datos, sino que se utilizan esos datos para detectar patrones y ajustar las acciones a realizar por el modelo. Como se mencionó en la sección introductoria de este documento, los algoritmos de Aprendizaje Automático se clasifican, de manera general, en supervisados y no supervisados. Los algoritmos supervisados analizan nuevos datos para encontrar similitudes con datos previamente analizados (proceso de entrenamiento), mientras que los algoritmos no supervisados pueden formar inferencias a partir de un conjunto de datos.

Un gran número de problemas cae en el llamado “Aprendizaje Inductivo” [77], la principal diferencia entre ellos radica en el tipo de objetos que se intentan predecir. Algunas clases habituales son las siguientes:

- **Regresión.** En los problemas de regresión, los valores de los conjuntos que se desean reproducir deben tener un valor continuo, siendo este el valor que se desea reproducir. Para ello se seleccionan las características con las que se realizarán las predicciones, a las que se denominan variables independientes. Una vez realizado esto se debe definir una fórmula matemática y calcular los parámetros de modo que al introducir las características se obtenga el valor deseado. Esta se conoce como variable dependiente. Entre los modelos matemáticos más populares que se utilizan en este tipo de problemas son: la regresión lineal [105], los bosques de regresión y las redes neuronales de regresión [109].
- **Clasificación.** En los problemas de clasificación los valores de los conjuntos que se desean reproducir deben ser categorías [66]. Uno de los casos más habituales es tener únicamente dos posibles categorías: verdadero o falso, a lo que se llama clasificación binaria. Cuando se tienen más de dos categorías, los problemas se denominan clasificación multiclase. Al igual que en los problemas de regresión es necesario identificar las variables independientes, la fórmula y los parámetros con los que se predice la variable dependiente. Solo que en este caso la variable dependiente es discreta. Una forma de ver el resultado de los modelos en estos problemas es mediante el uso de la matriz de confusión. Este tipo de matriz es una tabla de resultados en la que se compara los valores reales con los predichos. Algunos términos importantes relacionados con la matriz de confusión son:
  - Verdadero positivo: cuando el resultado real es verdadero y el valor predicho también es verdadero.
  - Verdadero negativo: cuando el resultado real es falso y el valor predicho también es falso.
  - Falso positivo: cuando el resultado real es falso pero el valor predicho es verdadero.



- Falso negativo: cuando el resultado real es verdadero pero el valor predicho es falso.

La calidad y eficiencia de los modelos se mide con diferentes parámetros como la precisión (“Accuracy”), la exactitud (“Precision”), la exhaustividad (“Recall”) y el F1. En primer lugar, la precisión es el porcentaje de acierto. Por otro lado, la exactitud es el porcentaje de acierto en la categoría que se considera positiva. La exhaustividad es el porcentaje de acierto entre los valores que son positivos en la realidad. Finalmente, F1 es un promedio ponderado de la exactitud y la exhaustividad. Los modelos matemáticos más populares que se utilizan en este tipo de problemas son: el K vecino más cercano (k-NN) [90], la regresión logística [64], las máquinas de vectores de soporte (SVM) [27] y las redes neuronales artificiales [141].

### 2.5.1. Algoritmos Clasificadores

Una parte fundamental de las aplicaciones que hacen uso del Aprendizaje Automático es el algoritmo clasificador que se utiliza, ya que existen distintos algoritmos que se pueden implementar, dependiendo del tipo de problema que se quiere solucionar. Algunos de los algoritmos clasificadores más comunes son los siguientes:

#### **k-Nearest Neighbors (k-NN)**

Este método ha sido ampliamente utilizado en diferentes aplicaciones estadísticas. Presenta un enfoque no paramétrico, donde su base consiste en, a partir de un conjunto de datos de entrenamiento, un grupo de muestras  $k$  se identifica por ser el más cercano a las muestras desconocidas. Para desarrollar este proceso, se calcula la distancia euclidiana, definida como  $\|\vec{x} - \vec{y}\|$ , entre un conjunto dado de consultas y las entradas, identificando los puntos de entrada más cercanos para cada consulta. Luego, la salida de las muestras desconocidas se determina calculando el promedio de las características de entrada, en función de las muestras iniciales [116]. Este algoritmo de clasificación ha sido utilizado en diversos trabajos sobre reconocimiento y clasificación de actividades humanas [21, 55].

### Nearest Centroid (NC)

Este método de clasificación ha sido utilizado en diversos trabajos de reconocimiento de actividades humanas [91], presenta un enfoque de agrupamiento particional, donde los objetos de datos se agrupan en  $k$  grupos, dependiendo de su similitud. El parámetro  $k$  debe especificarse inicialmente. La primera selección de centroides se realiza al azar. Para medir la cercanía entre un centroide y un objeto de datos, se calcula la distancia euclidiana. Luego, después de esta primera agrupación, se obtiene un nuevo centroide para cada grupo, así como la distancia entre cada objeto de datos a cada centro, reasignando los puntos de datos de acuerdo con la distancia. El punto en un grupo se trata como el centroide si la suma de todas las distancias entre este punto y todos los objetos del grupo logran la minimización óptima. El objetivo principal de NC es minimizar la suma de distancias entre los objetos de un grupo y su centroide [25].

### Random Forests (RF)

Este clasificador fue desarrollado por Breinman et al. [15] y ha sido utilizado en distintos trabajos [85, 46]. Su desempeño consiste en dos niveles de aleatoriedad para la construcción de los árboles, comenzando con una versión de arranque de un conjunto de datos de entrenamiento, llamado *bagging*, donde un subconjunto de los datos de entrenamiento está destinado a cada árbol, basado en el principio de reemplazo, mientras que los datos restantes se usan para estimar el error, calculando el error fuera de bolsa (OOB). Luego, en el segundo nivel de árboles, se selecciona aleatoriamente un subconjunto de características y se agrega a cada nodo a lo largo del crecimiento de los árboles de decisión. En cada nodo, se selecciona la mejor característica, buscando reducir el error de etiqueta. Este método basa su técnica de clasificación por el principio del voto mayoritario de todos los árboles de decisión. Este proceso se repite de forma recursiva hasta alcanzar una profundidad definida en el bosque o el número de muestras en un nodo no excede un umbral [139, 119].

### Artificial Neural Network (ANN)

En este método, se realiza la búsqueda de una tarea específica basada en la correlación entre características. Su proceso consiste en un aprendizaje o entrenamiento que se asemeja al comportamiento de las redes neuronales biológicas. A través de una serie de etiquetas diferentes contenidas por nodos o neuronas, las ANN intentan encontrar un modelo relacional entre las características de entrada y el resultado. Tres elementos principales están presentes en este método:

1. Un conjunto de sinapsis o conexiones caracterizadas por un “peso”, donde la señal de entrada está conectada a una neurona a través de su producto con el peso en esa conexión.
2. Un sumador, que agrega las contribuciones de una señal golpeada por todos los pesos.
3. Una función de activación, equivalente a una función de transferencia, que está afectando a las neuronas, permitiendo limitar la amplitud del resultado, proporcionando un rango permisible para la señal de resultado en términos de valores finitos [136].

Las redes neuronales artificiales son uno de los métodos de clasificación más utilizados en los últimos años en los trabajos de reconocimiento y clasificación de actividades humanas, ya que han mostrado buenos resultados y desempeño [49, 50, 45, 86].

### Recursive Partitioning Tree (Rpart)

Este es un método estadístico para el análisis multivariado que separa las muestras en diferentes grupos de riesgo homogéneos para determinar los predictores de supervivencia [131, 130]. El algoritmo consiste en seleccionar el predictor que proporciona la división óptima, de modo que cada uno de los subgrupos sea más homogéneo con respecto al resultado. Estos subgrupos se dicotomizan en subgrupos más pequeños y más homogéneos de acuerdo con la característica que presenta las mejores divisiones para cada subgrupo. Se realiza un proceso iterativo continuo hasta presentar pocos valores para divisiones adicionales. El proceso de poda se aplica al árbol de partición original para recortar los árboles hasta el punto

donde se maximiza la precisión predictiva general, evitando el sobreajuste en los datos [73].

## 2.6. Resumen

En este capítulo se presentaron los términos relacionados con el tema de reconocimiento y clasificación de actividades infantiles utilizando sonido ambiental.

Al inicio del capítulo se abordaron los conceptos relacionados con la clasificación de actividades humanas y la metodología general para la construcción de un modelo de este tipo: 1) Recolección de datos; 2) Extracción de características; 3) Entrenamiento del modelo; 4) Clasificación de la actividad. Se detalló cada uno de estos 4 puntos.

Enseguida se presentaron modelos de clasificación diseñados por otros autores, además de explicar las diferencias y particularidades de cada tipo de modelo, en donde se distinguen 2 aspectos importantes a tomar en cuenta al momento de diseñar un modelo de reconocimiento y clasificación de actividades humanas: 1) El grupo de individuos a los que va dirigido el análisis; 2) La fuente de datos a utilizar. En esta parte se revisaron distintos trabajos presentados enfocados a diferentes grupos de personas como lo son adultos mayores, personas enfermas y sistemas de monitoreo.

Avanzando en el capítulo se trató también específicamente el tema de reconocimiento y clasificación de actividades infantiles, mencionando algunos de los trabajos reportados, las fuentes de datos que utilizan, las actividades que toman en cuenta para el análisis y la forma en que lo realizan. Se detectó que los trabajos enfocados a niños generalmente utilizan como fuente de datos sensores incrustados en las prendas de los niños, lo que puede traer diversos problemas al modelo e inclusive interferir con el comportamiento natural de los niños al momento de estar realizando las actividades. En este sentido, se comentó también que el enfoque del presente trabajo es utilizar una fuente de datos que no interfiera con las actividades al momento de la captura de datos, como lo es el sonido ambiental.

En la parte final de este capítulo se presentó el tema de Aprendizaje Automático, mencionando sus principales características y la forma en que sus técnicas se aplican al área de reconocimiento y clasificación de actividades humanas. Además se presentaron los principales algoritmos clasificadores utilizados en la generación de modelos de clasificación. De forma paralela también se fueron mencionando trabajos que hacen referencia al uso de los distintos algoritmos clasificadores en esta área.

# Capítulo 3

## Extracción y Selección de Características

Este capítulo presenta lo referente a los procesos de extracción y selección de características a partir de los audios capturados del sonido ambiental. Se describe cómo está conformado el dataset, cuáles son las características extraídas y analizadas, además de las técnicas utilizadas para la extracción y selección de características.

### 3.1. Introducción

La fuente de datos es un elemento fundamental en la generación de un modelo de reconocimiento y clasificación de actividades humanas, especialmente cuando se enfoca en niños, por las complejidades que esto ocasiona. Las fuentes de datos utilizadas en los trabajos presentados en el capítulo anterior en su mayoría consistían en sensores incrustados en las prendas de los niños. En el presente trabajo se propone la utilización de una fuente de datos que no se base en dispositivos portados por los individuos: el sonido ambiental. Al utilizar el sonido ambiental como fuente de datos se tiene la ventaja de que los niños realizan las actividades sin portar ningún tipo de sensor y los datos recolectados del sonido ambiental proporcionan suficiente información para su análisis y la respectiva clasificación de las actividades.

Tabla 3.1: Descripción general de las actividades analizadas.

Actividad	Descripción
Caminar	Moverse a velocidad media de un lugar a otro
Correr	Moverse rápidamente de un lugar a otro
Jugar	Manipular objetos plásticos
Llorar	Emitir sonido de llanto en reacción a algún evento

## 3.2. Dataset

En la mayoría de los trabajos presentados en el capítulo anterior sobre reconocimiento y clasificación de actividades infantiles, comúnmente se analizan actividades detectables mediante el movimiento, como caminar, correr o saltar, debido a que esos trabajos utilizan como fuente de datos sensores de movimiento, como el acelerómetro. Con el objetivo de analizar diferentes tipos de actividades y, en especial, tener una fuente de datos que no interfiera en las actividades que realizan los niños, en el presente trabajo el conjunto de datos se compone por grabaciones de audio de cuatro actividades comúnmente realizadas por niños de 12 a 36 meses de edad, en un ambiente residencial: caminar, correr, jugar y llorar, dos de las cuales no son detectables a través de sensores de movimiento (jugar y llorar). La Tabla 3.1 muestra la descripción de las actividades analizadas en este trabajo.

Para la conformación del dataset utilizado en el presente trabajo, 10 % de los sonidos fueron generados de manera propia y 90 % fueron tomados de Internet, a través de una búsqueda de audios sobre actividades infantiles, la cual se llevó a cabo el 3 de Octubre de 2018 [36, 132].

A partir del proceso de adquisición de datos, se obtuvieron un total de 146 grabaciones (considerando las grabaciones propias y las tomadas de Internet), las cuales se dividieron en 2716 clips de 10 segundos. Como se muestra en [33, 114], los clips de audio de 10 segundos parecen ser lo suficientemente largos para obtener información potencialmente útil en el proceso de clasificación de actividades mediante el análisis del sonido ambiental. En el conjunto de datos utilizado, las grabaciones de audio con una duración superior a

Tabla 3.2: Clips de audio por actividad.

Actividad	Generadas		Tomadas de Internet		Total	
	Grabaciones	Clips	Grabaciones	Clips	Grabaciones	Clips
Caminar	10	65	30	652	40	717
Correr	9	81	30	611	39	692
Jugar	9	67	17	636	26	703
Llorar	8	72	33	532	41	604

Tabla 3.3: Dispositivos de Grabación.

Dispositivo	Procesador	Sistema Operativo
Lanix Ilium s620	MediaTek MT6582 quad-core	Android 4.2.2
Motorola Moto G4	Snapdragon617	Android 6.0.1
Hisense u965	MediaTek MT6580	Android 8.1.0

10 segundos se dividieron para generar una mayor cantidad de muestras. La Tabla 3.2 muestra el número de grabaciones obtenidas para cada actividad, así como el número de clips de 10 segundos generados.

### 3.2.1. Dispositivos de Grabación

Para realizar las grabaciones de los clips de audio correspondientes a la parte del dataset generado de manera propia se utilizaron 3 dispositivos móviles, de los cuales se muestra su descripción en la Tabla 3.3.

### 3.2.2. Metadatos

En referencia al proceso de grabación de los clips de audio, este se realizó utilizando diferentes dispositivos y diferentes configuraciones, además considerando las grabaciones tomadas de Internet. El conjunto de datos de este trabajo incluye clips de audio con una frecuencia de muestreo entre 44100 Hz y 96000 Hz, en canales estéreo y mono. La Tabla 3.4 muestra los metadatos de los clips de audio presentes en el dataset para cada actividad. Las características presentadas en la Tabla 3.4 aseguran una calidad aceptable para los archivos de audio grabados y definen los parámetros necesarios para futuras grabaciones



Tabla 3.4: Metadatos de los clips de audio.

Actividad	Frecuencia de muestreo	Formato de codificación	Canales
Caminar	44100 Hz - 96000 Hz	mp3,wav,flac,aiff	Mono, Stereo
Correr	44100 Hz - 48000 Hz	mp3,wav,flac	Mono, Stereo
Jugar	44100 Hz - 96000 Hz	mp3,wav	Mono, Stereo
Llorar	44100 Hz - 96000 Hz	mp3,wav	Mono, Stereo

con el fin de expandir el conjunto de datos.

### 3.3. Extracción de Características

Uno de los elementos fundamentales para los modelos de reconocimiento y clasificación de actividades humanas es el proceso de extracción de características. La extracción de características es el proceso mediante el cual se obtiene información de los clips de audio. Esta información se utiliza para diferenciar el tipo de actividad a la que pertenece la grabación, ya que para cada tipo de actividad, los clips de audio contienen diferentes medidas para sus características extraídas.

#### 3.3.1. Características de Audio

Se han propuesto una gran variedad de características de la señal para la clasificación de audio en general. Una gran parte de estas características consiste en características de señal de bajo nivel, que incluyen parámetros tales como la velocidad de cruce por cero, el ancho de banda de la señal, el centroide espectral y la energía de la señal [84, 42, 129, 133]. Por lo general, tanto los promedios como las variaciones de estas propiedades de señal se incluyen en el conjunto de características.

Un segundo conjunto de características importantes consiste en los Coeficientes Ceps-  
trales de Frecuencia de Mel (MFCC) [81, 34]. Esta descripción paramétrica de la envoltura  
espectral tiene la ventaja de ser independiente del nivel y de producir bajas correlacio-  
nes mutuas entre diferentes características [128]. La clasificación basada en un conjunto

de características que no están correlacionadas es típicamente más fácil que la basada en características con correlaciones.

Tanto las características de señal de bajo nivel como los MFCC se han utilizado comúnmente para modelos de clasificación de audio en general. En los siguientes puntos se describe el conjunto de características utilizado para este trabajo de reconocimiento y clasificación de actividades infantiles mediante sonido ambiental.

### 3.3.2. Proceso de Extracción de Características

Este punto describe la forma en la que se llevó a cabo el proceso de extracción de características, así como la forma en la que fueron procesados los clips de audio y el conjunto de características elegidas para llevar a cabo el análisis de las actividades definidas.

#### Preprocesamiento de los audios

Debido a que el conjunto de datos contiene archivos de audio de diferentes longitudes, estos se dividieron en clips de 10 segundos, lo que hace que todas las muestras analizadas tengan la misma longitud.

#### Conjunto de características

Cada clip de 10 segundos se transforma en una matriz, donde cada posición representa la magnitud de la función correspondiente para ese clip de audio. Para evitar problemas con la diferencia en los canales de las grabaciones (Mono y Estéreo), todas las muestras se convirtieron al tipo Mono. La Tabla 3.5 muestra el conjunto de 34 características extraídas para cada clip de audio de 10 segundos.

Se extrajeron un total de 34 características para cada clip de 10 segundos, por lo que la base de datos para la comparación de los algoritmos de clasificación estaba contenida por 2716 registros con 34 características cada uno

Tabla 3.5: Características Extraídas.

Índice	Nombre	Descripción
1	Zero Crossing Rate	Tasa de cambio de signo de la señal durante la duración de un cuadro particular.
2	Energy	Suma de los cuadrados de los valores de señal, normalizados por la longitud respectiva.
3	Entropy of Energy	Entropía de las energías normalizadas de las subtramas. Se puede interpretar como una medida de cambios abruptos.
4	Spectral Centroid	Centro de gravedad del espectro.
5	Spectral Spread	Segundo momento central del espectro.
6	Spectral Entropy	Entropía de las energías espectrales normalizadas para un conjunto de subtramas.
7	Spectral Flux	Diferencia cuadrada entre las magnitudes normalizadas de los espectros de los dos cuadros sucesivos.
8	Spectral Rollof	Frecuencia por debajo de la cual se concentra el 90 % de la distribución de la magnitud del espectro.
9-21	MFCC's	Coefficientes cepstrales en las frecuencias de Mel forman una representación cepstral donde las bandas de frecuencia no son lineales sino que se distribuyen de acuerdo con la escala de Mel.
22-33	Chroma Vector	Representación de 12 elementos de la energía espectral en la que los contenedores representan las 12 clases de tono temperamental de música de tipo occidental (espaciado de semitonos).
34	Chroma Deviation	Desviación estándar de los 12 coeficientes de croma.

Es importante mencionar que este conjunto de características se eligió debido a que estas se han utilizado comúnmente en trabajos relacionados sobre procesamiento y análisis de audio [103, 124, 120], especialmente los Coeficientes Cepstrales en la frecuencia de Mel, siendo una de las características más sólidas en el área de reconocimiento y clasificación de actividades mediante sonido [110, 38, 99, 79].

### 3.4. Selección de Características

La selección de características, también llamada selección de variables o selección de atributos, es uno de los conceptos principales en el Aprendizaje Automático, ya que su implementación impacta directamente en el desempeño del modelo creado. Las características seleccionadas de los datos para entrenar el modelo tienen una gran influencia en el desempeño de este. Las características irrelevantes o parcialmente relevantes pueden afectar negativamente el rendimiento del modelo, y es por esta causa que la selección de características debe ser el primer y más importante paso del diseño de un modelo de clasificación [31].

La selección de características es el proceso en el que selecciona automática o manualmente las características que más contribuyen a su variable de predicción o salida en la que está interesado. Tener características irrelevantes en los datos puede disminuir la precisión de los modelos y hacer que el modelo aprenda basándose en características irrelevantes [63]. La selección de características es diferente a la reducción de dimensionalidad. Ambos métodos buscan reducir el número de atributos en el conjunto de datos, pero un método de reducción de dimensionalidad lo hace creando nuevas combinaciones de atributos, mientras que los métodos de selección de características incluyen y excluyen atributos presentes en los datos sin cambiarlos. En este sentido, el proceso de selección de características trae las siguientes ventajas:

- **Reduce el sobreajuste.** El principal beneficio de la selección de funciones es que reduce el sobreajuste. Al eliminar datos extraños, permite que el modelo se centre solo en las características importantes de los datos y se dejen de lado las características

que no importan [76].

- **Precisión.** Otro beneficio de eliminar información irrelevante es que mejora la precisión de las predicciones del modelo, ya que las características irrelevantes pueden provocar que el modelo se “confunda” durante el proceso de clasificación [58].
- **Tiempo de Cálculo.** El proceso de selección de características también reduce el tiempo de cálculo involucrado para obtener el modelo, ya que resulta más sencillo para el modelo trabajar con un subconjunto reducido de datos en lugar de trabajar con el conjunto de datos original [65].
- **Sencillez.** Tener un número menor de características hace que el modelo sea más interpretable y fácil de entender.

En general, la selección de características es clave para los modelos de clasificación ya que ayuda a aumentar la precisión de los mismos al momento de la predicción de datos.

### 3.4.1. Métodos de Selección de Características

Los métodos de selección de características tienen como finalidad reducir el número de variables de entrada a aquellas que se consideran más útiles para un modelo a fin de predecir la variable objetivo. Algunos problemas de modelado predictivo tienen una gran cantidad de variables que pueden retrasar el desarrollo y la capacitación de los modelos y requieren una gran cantidad de memoria del sistema. Además, el rendimiento de algunos modelos puede verse afectado al incluir variables de entrada que no son relevantes para la variable de destino. Hay tres tipos principales de algoritmos de selección de características [19]:

- Métodos de envoltura
- Métodos de filtro
- Métodos embebidos

### Métodos de envoltura

Los métodos de envoltura calculan modelos con un cierto subconjunto de características y evalúan la importancia de cada característica. Luego iteran y prueban un subconjunto diferente de características hasta que se alcanza el subconjunto óptimo. Dos inconvenientes de este método son el gran tiempo de cálculo para los datos con muchas características, y que tiende a sobreajustar el modelo cuando no hay una gran cantidad de puntos de datos. Los métodos de envoltura más notables de selección de características son la selección hacia adelante (forward selection), la eliminación hacia atrás (backward elimination) y la selección por pasos (stepwise selection) [67, 19].

### Métodos de filtro

Los métodos de filtro utilizan una medida distinta de la tasa de error para determinar si esa característica es útil. En lugar de ajustar un modelo (como en los métodos de envoltura), se selecciona un subconjunto de las características clasificándolas por una medida descriptiva útil. Los beneficios de los métodos de filtro son que tienen un tiempo de cálculo muy bajo y no sobreajustarán los datos. Sin embargo, un inconveniente es que son ciegos a cualquier interacción o correlación entre características. Tres métodos de filtro diferentes son ANOVA, correlación de Pearson y umbral de varianza [67, 19].

Es común utilizar medidas estadísticas de tipo de correlación entre las variables de entrada y salida como base para la selección de características del filtro. Como tal, la elección de medidas estadísticas depende en gran medida de los tipos de datos variables. Los tipos de datos comunes incluyen numéricos (como la altura) y categóricos (como una etiqueta), aunque cada uno puede subdividirse aún más, como números enteros y flotantes para variables numéricas, y booleanos, ordinales o nominales para variables categóricas. Cuanto más se sepa sobre el tipo de datos de una variable, más fácil será elegir una medida estadística adecuada para un método de selección de características basado en filtros.

## Métodos embebidos

Los métodos embebidos realizan la selección de características como parte del proceso de creación del modelo. La regresión Lasso y Ridge son los dos métodos de selección de características más comunes de este tipo, y el árbol de decisión también crea un modelo utilizando diferentes tipos de selección de características [19].

### 3.4.2. Proceso de Selección de Características

Como se mencionó en el punto anterior, existen distintos tipos de métodos de selección de características, en donde para cada método se pueden encontrar algoritmos con diferentes enfoques, pero todos con el objetivo de reducir el número de datos para el modelo de clasificación, con los beneficios ya explicados que esto trae.

A partir del conjunto de datos resultante del proceso de extracción de características, en el presente trabajo se implementaron dos métodos de selección de características:

- Criterio de Akaike
- Algoritmos Genéticos

A continuación se describe la forma de funcionamiento de los métodos de selección de características aplicados así como la forma en la que fueron implementados y los resultados obtenidos.

#### Criterio de Akaike

El criterio de información de Akaike (AIC) [4] es una medida de la bondad de ajuste de un modelo estadístico. Se puede decir que describe la relación entre el sesgo y varianza en la construcción del modelo, o hablando de manera general acerca de la exactitud y complejidad del modelo.

El AIC no es una prueba del modelo en el sentido de prueba de hipótesis. Más bien, proporciona un medio para la comparación entre los modelos de una herramienta para la

selección del modelo [122]. Dado un conjunto de datos, varios modelos candidatos pueden ser clasificados de acuerdo a su AIC, siendo el mejor el modelo que tiene el AIC mínimo. A partir de los valores de la AIC también se puede inferir que, por ejemplo, los dos primeros modelos están más o menos empatados y el resto son mucho peores. En general, el AIC se define como:

$$AIC = 2k - 2\ln(L) \quad (3.1)$$

Donde:

- $k$  es el número de parámetros del modelo.
- $\ln(L)$  es la función de log-verosimilitud para el modelo estadístico.

Para los conjuntos de datos más pequeños, la AICc se aplica la corrección de segundo orden:

$$AICc = AIC + \frac{2k(k+1)}{N-k-1} = \frac{2Nk}{N-k-1} - 2\ln(L) \quad (3.2)$$

Donde:

- $N$  es el tamaño de la muestra de datos.
- $k$  es el número de parámetros del modelo.

El AIC no es una prueba del modelo en el sentido de la prueba de hipótesis, sino que es una prueba entre los modelos, una herramienta para la selección del modelo.

Aplicando lo anterior al presente trabajo sobre reconocimiento y clasificación de actividades infantiles, en esta etapa, se realizó un proceso de selección de características basado en el criterio de Akaike, con el objetivo de reducir el número de características, seleccionando aquellas que presentan la información más significativa para diferenciar las clases a las que pertenecen las muestras de audio.



Tabla 3.6: Características seleccionadas mediante el AIC

Feature Name
Zero Crossing Rate
Energy
Entropy of Energy
Spectral Centroid
Spectral Spread
Spectral Entropy
Spectral Rolloff
MFCCs (1, 3 - 13)
Chroma Vector (2, 4, 6 - 10)
Chroma Deviation

Siguiendo este principio, se generaron modelos que toman en cuenta todas las combinaciones posibles formadas con las 34 características extraídas, calculando posteriormente el AIC para cada uno de estos modelos. Luego, los modelos se clasifican según su AIC, siendo el mejor de ellos el que tiene el AIC más bajo. Las características que forman el mejor modelo, según el AIC, se muestran en la Tabla 3.6.

Se puede observar en la Tabla 3.6 que el modelo seleccionado por este método cuenta con 27 características. Este subconjunto de características extraídas se tomará para la generación de algunos de los modelos de clasificación de actividades infantiles detallados en secciones posteriores.

### Algoritmos genéticos

Para esta etapa, se implementó el algoritmo genético (AG) “Galgo” [117], utilizando cinco métodos de clasificación diferentes, para obtener subconjuntos de cinco características. Estos métodos de clasificación son k-NN, NC, RF, ANN y Rpart.

Galgo es un paquete desarrollado bajo el lenguaje R [115], implementado con un enfoque orientado a objetos y basado en una función de ajuste general para guiar la selección de características. El procedimiento de Galgo está diseñado para comenzar con una población aleatoria de subconjuntos de características, conocidos como cromosomas. Cada

cromosoma se evalúa midiendo su capacidad para predecir una salida o una característica dependiente, en función de la precisión. Para esta evaluación, se incluye un método de clasificación. La idea principal es reemplazar la población inicial por una nueva que incluya características de diferentes cromosomas que presenten una mayor precisión de clasificación. Este proceso se repite suficientes veces para lograr el nivel deseado de precisión. La mejora progresiva de la población cromosómica se realiza mediante un proceso inspirado en la selección natural, basado en tres principios: selección, mutación y cruce.

En resumen, la aplicación de Galgo consta de cuatro etapas principales:

1. **Configuración del Análisis.** En esta etapa se establecen los parámetros del algoritmo, incluidos los datos de entrada, el resultado, el modelo estadístico, la precisión deseada, el esquema de estimación de errores, el método de clasificación, entre otros.
2. **Búsqueda de modelos multivariados relevantes.** Esta etapa consiste en el proceso de selección que comienza con la población aleatoria de cromosomas, basada en un método de clasificación, buscando las mejores soluciones locales.
3. **Refinamiento y análisis de las soluciones locales.** Los cromosomas seleccionados están sujetos a una estrategia de selección hacia atrás, ya que, incluso cuando estos cromosomas presentan la mejor precisión, podría haber características incluidas en el modelo que no contribuyen significativamente al valor de ajuste. El objetivo de esta estrategia es derivar una población cromosómica que contenga solo características que contribuyan efectivamente a la precisión de la clasificación.
4. **Desarrollo de un modelo estadístico final.** Finalmente, se obtiene un único modelo representativo basado en una estrategia de selección directa, donde, de acuerdo con una inclusión gradual, se seleccionan los genes más frecuentes presentados en la población de cromosomas.

Como se mencionó anteriormente, Galgo evalúa la capacidad de los cromosomas para predecir una salida o una característica mediante un método de clasificación. Los métodos de clasificación utilizados para este trabajo en esta etapa son:

Tabla 3.7: Configuración general de parámetros experimentales para el algoritmo genético “GALGO”

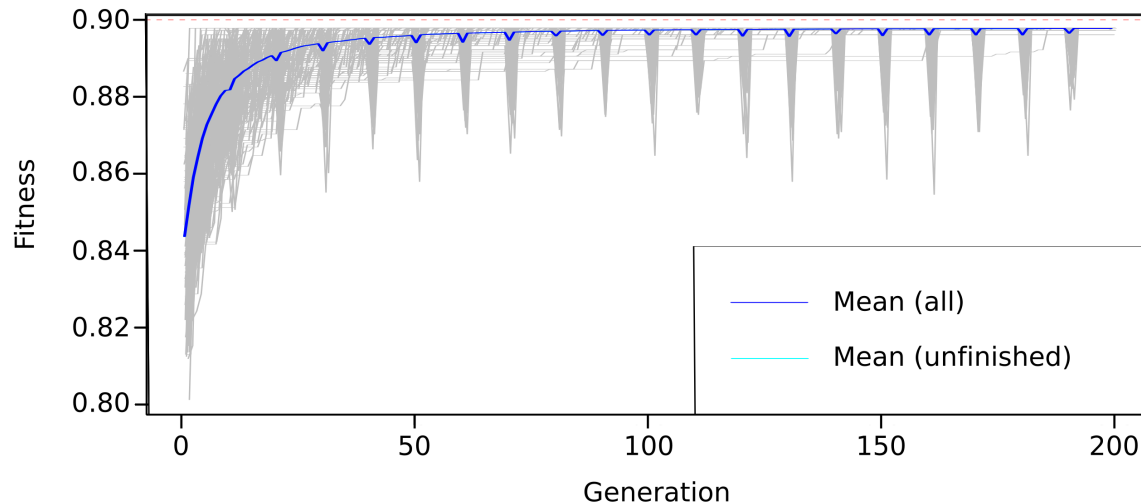
Parámetro	Valor	Descripción
data	Matriz de datos de 2,672 x 34.	Marco de datos que contiene las muestras utilizadas (muestras en columnas y genes en filas).
classification method	“knn”, “nearest”, “nnet”, “randomforest”, “rpart”.	Métodos utilizados para la clasificación.
chromosomeSize	5	Tamaño de los cromosomas (el número de variables / genes incluidos en los modelos).
niches	2	Número de nichos (el nicho representa un conjunto de cromosomas).
goalFitness	0.90	Valor de ajuste deseado (fracción de clasificación correcta).
maxSolutions	300	Cantidad máxima de soluciones recolectadas.

- k-Nearest Neighbors
- Nearest Centroid
- Random Forests
- Artificial Neural Network
- Recursive Partitioning Tree

Para la implementación del Algoritmo Genético en la selección de características, se definió la configuración inicial mostrada en la Tabla 3.7.

Es importante mencionar que para seleccionar el tamaño del cromosoma, la configuración presentada por defecto por Galgo, que es cinco, se respetó debido a diferentes puntos.

Figura 3.1: Evolución de los modelos generados por Galgo para el método de clasificación k-NN.



Primero, debido a que el tamaño del conjunto de datos no es muy grande, por lo que se quería reducir el modelo a un tamaño que no comprometiera el rendimiento de la clasificación, además de querer evitar la aleatoriedad en el desarrollo del modelo final, para probar la robustez de la selección de características. Por otro lado, este es el número que Galgo asigna de forma predeterminada si el usuario no especifica el tamaño del cromosoma porque, según su descripción informada en la literatura, este es el número de cromosomas que ha presentado el mejor rendimiento para la función selección hacia los diferentes enfoques. Este es el mismo caso para la selección del número de nichos [117].

Las Figuras 3.1, 3.2, 3.3, 3.4 y 3.5 muestran el rendimiento del ajuste para cada una de las 200 generaciones desarrolladas para los métodos de clasificación, k-NN, NC, ANN, RF y Rpart respectivamente. A partir de estos gráficos se puede observar que la Figura 3.4, que corresponde a RF, es el ajuste (representada en el eje Y) con mayor precisión desde el comienzo del proceso, siendo estable en términos del proceso evolutivo del algoritmo genético.

Las Figuras 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13, 3.14 y 3.15 muestran la frecuencia de aparición de las características en los 300 modelos desarrollados para los métodos de

Figura 3.2: Evolución de los modelos generados por Galgo para el método de clasificación NC.

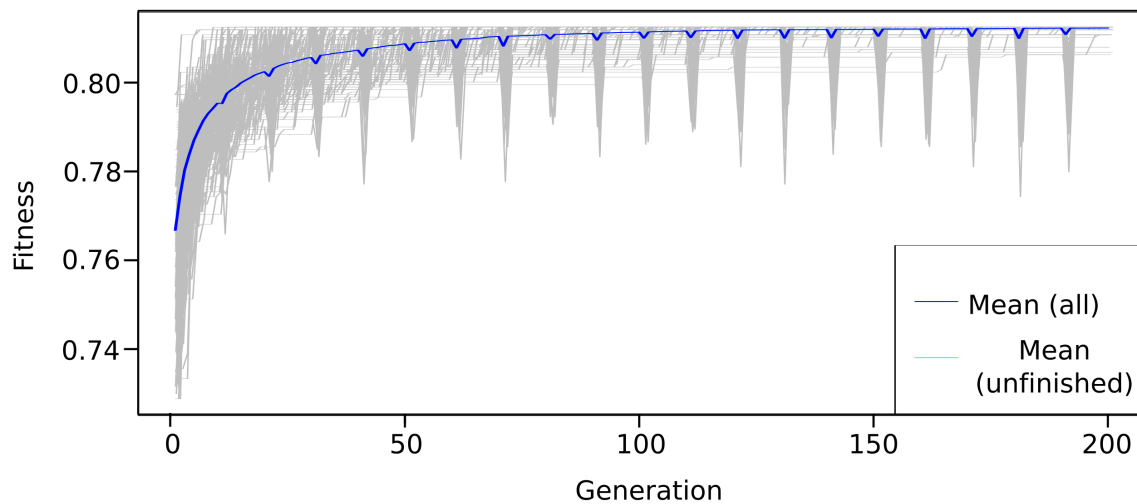


Figura 3.3: Evolución de los modelos generados por Galgo para el método de clasificación ANN.

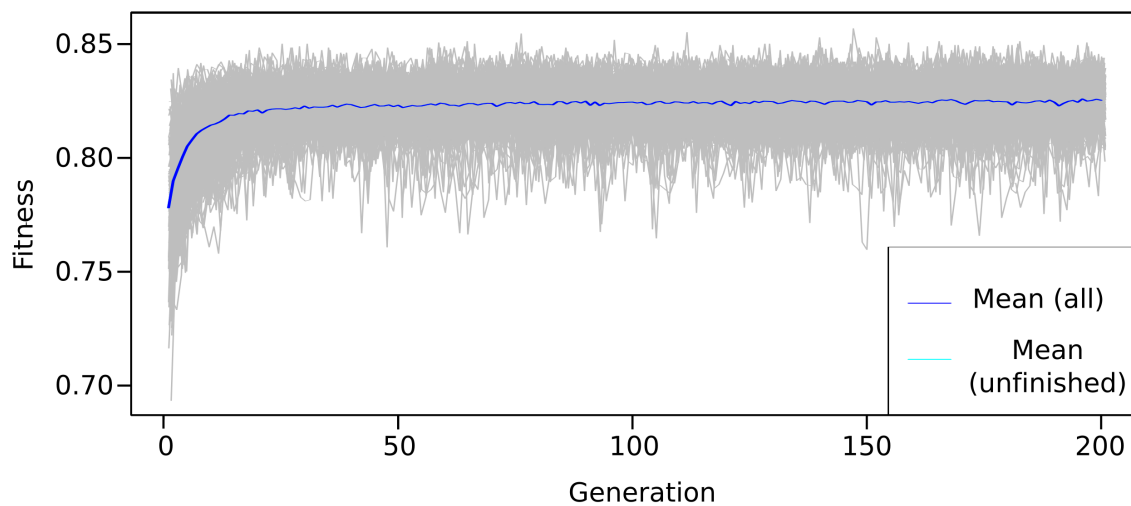


Figura 3.4: Evolución de los modelos generados por Galgo para el método de clasificación RF.

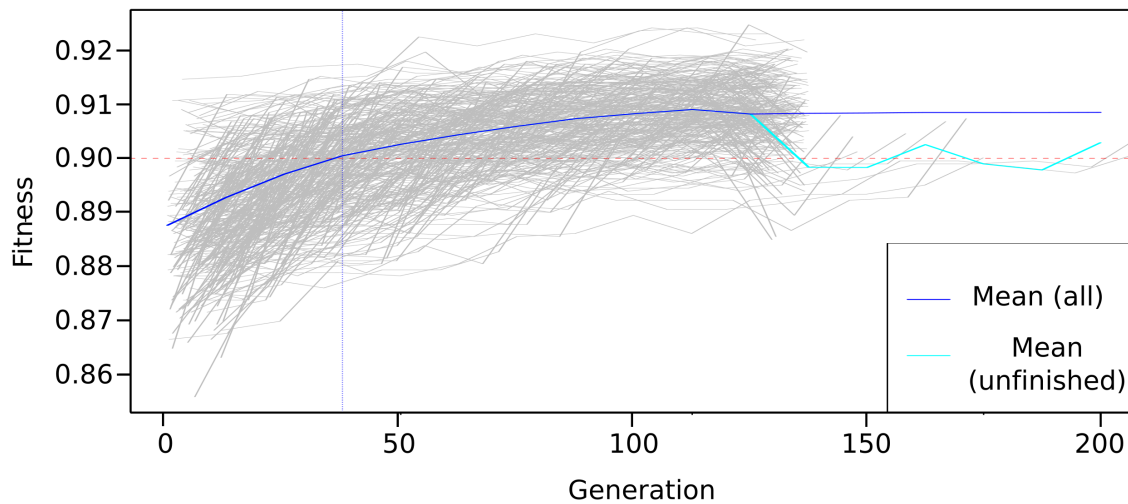


Figura 3.5: Evolución de los modelos generados por Galgo para el método de clasificación Rpart.

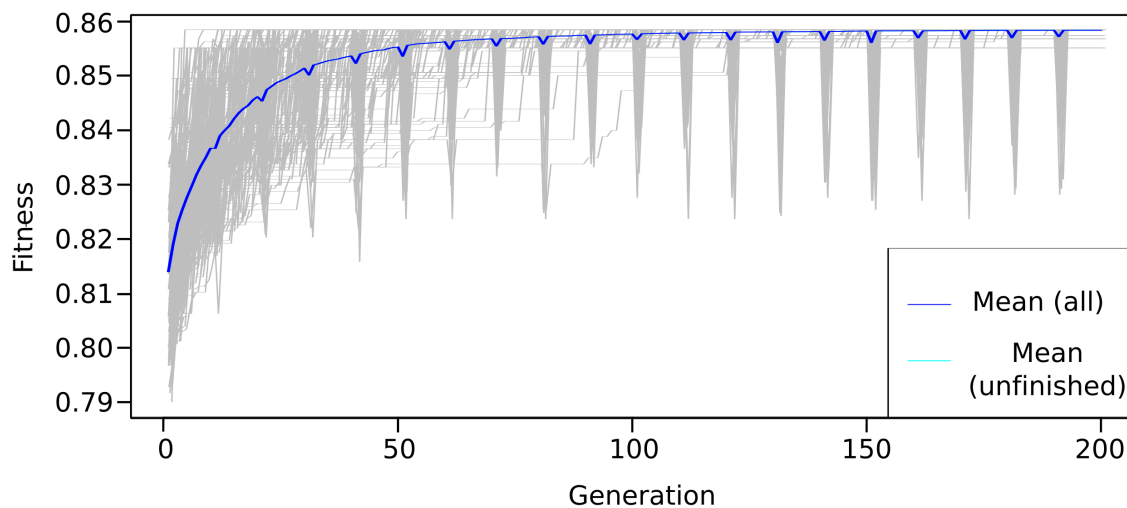


Figura 3.6: Evolución del rango con el clasificador k-NN para las características más frecuentes.

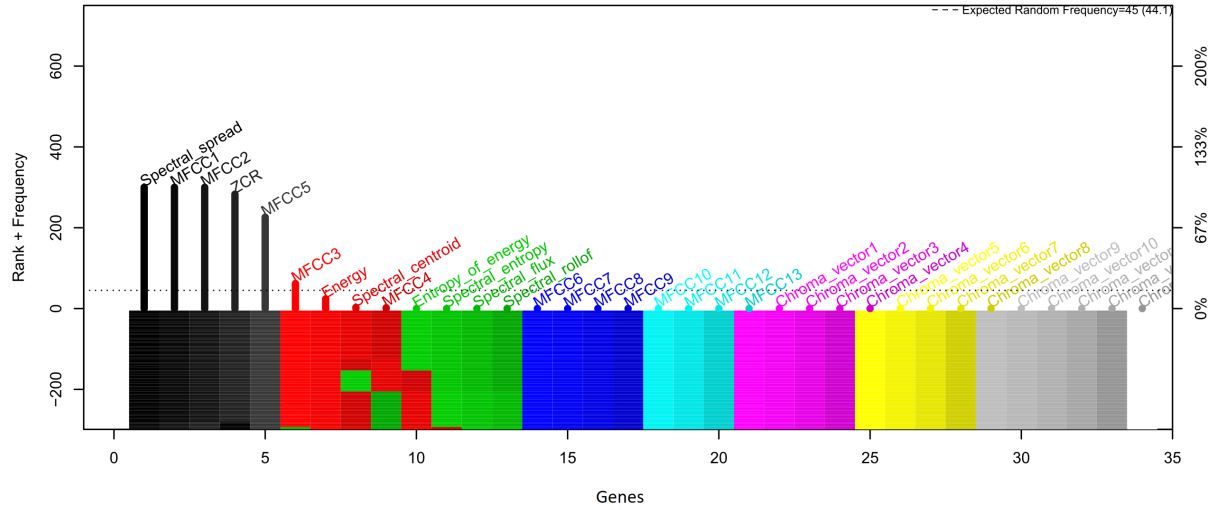


Tabla 3.8: Subconjunto de cinco características elegido por el algoritmo genético para cada clasificador.

Clasificador	Características
k-NN	ZCR, Spectral Spread, MFCC1, MFCC2, MFCC5
NC	Energy, Spectral Spread, MFCC1, MFCC6, MFCC11
ANN	Energy, Spectral Centroid, Spectral Spread, MFCC6, MFCC7
RF	Energy, Spectral Spread, Spectral Rollof, MFCC2, MFCC3
Rpart	Energy, Spectral Centroid, Spectral Rollof, MFCC1, MFCC2

clasificación k-NN, NC, ANN, RF y Rpart, respectivamente. . Esas características presentadas en color negro son las cinco que presentaron la mayor frecuencia de aparición, por lo tanto, seleccionadas como las más representativas. De estos, los algoritmos de bajo costo computacional (como k-NN, Rpart y NC) adquieren una frecuencia estable (mostrada en el eje Y) para las características seleccionadas, mientras que los modelos complejos, como ANN y RF, tienden a clasificar los genes con la capacidad para describir los fenómenos.

A partir de las Figuras 3.6, 3.8, 3.10, 3.12 y 3.14, la Tabla 3.8 muestra el subconjunto de cinco características determinado por Galgo para cada uno de los clasificadores analizados.

Figura 3.7: Las cinco características elegidas para el clasificador k-NN según su frecuencia.

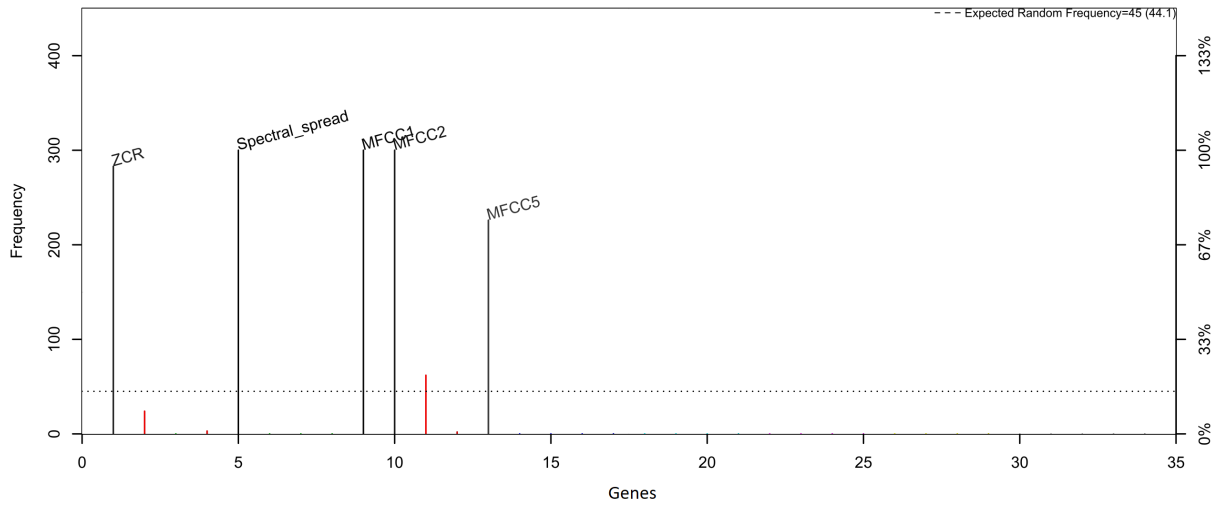


Figura 3.8: Evolución del rango con el clasificador NC para las características más frecuentes.

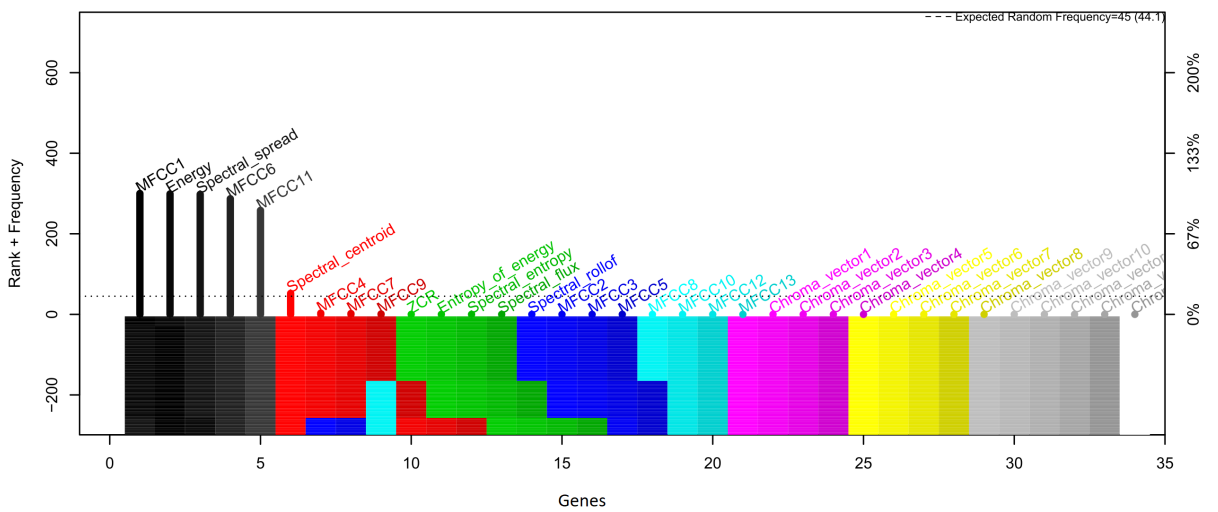




Figura 3.9: Las cinco características elegidas para el clasificador NC según su frecuencia.

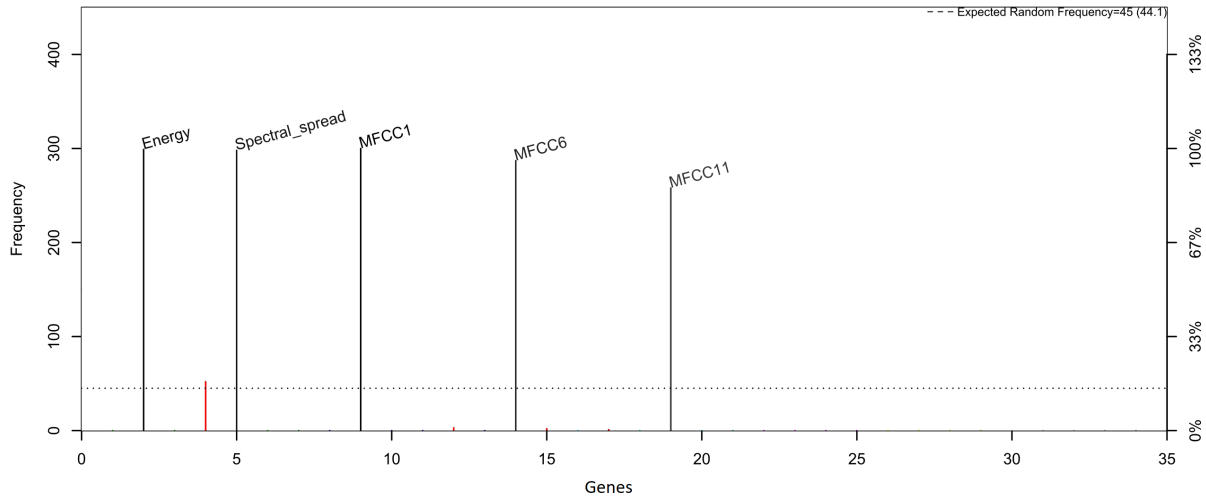


Figura 3.10: Evolución del rango con el clasificador ANN para las características más frecuentes.

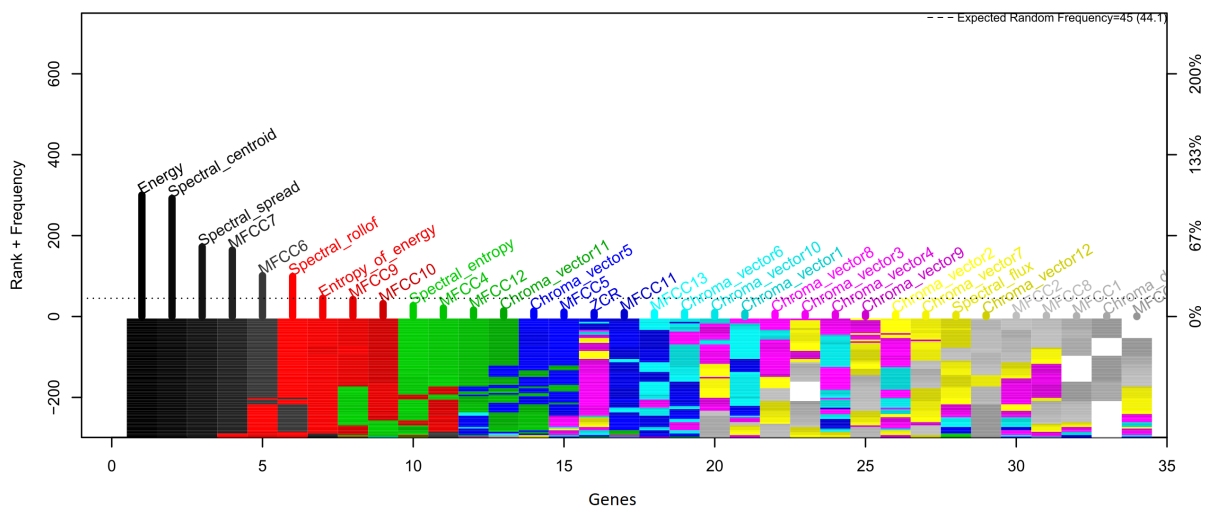


Figura 3.11: Las cinco características elegidas para el clasificador ANN según su frecuencia.

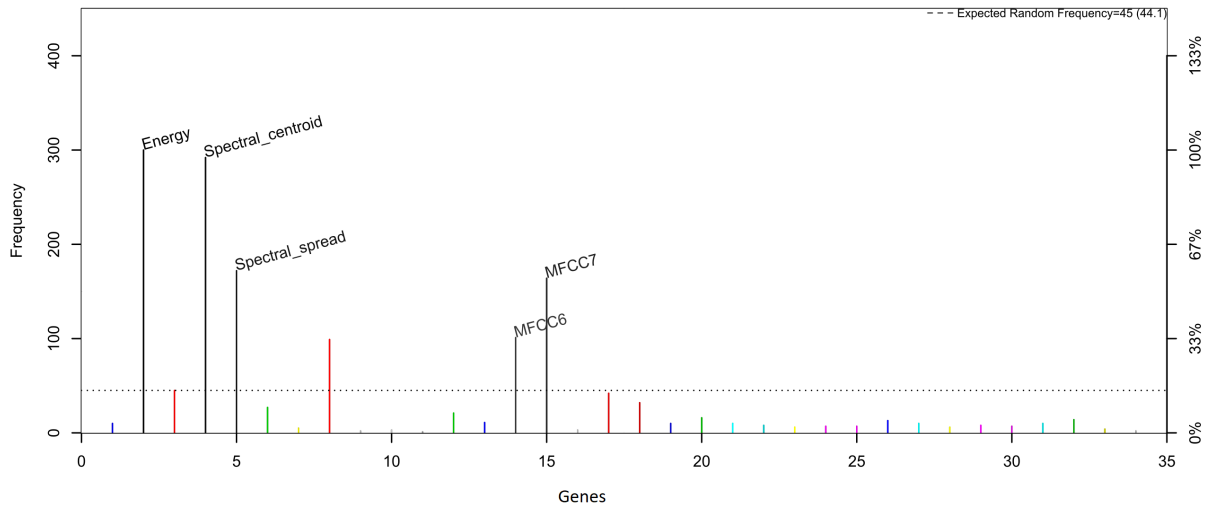


Figura 3.12: Evolución del rango con el clasificador RF para las características más frecuentes.

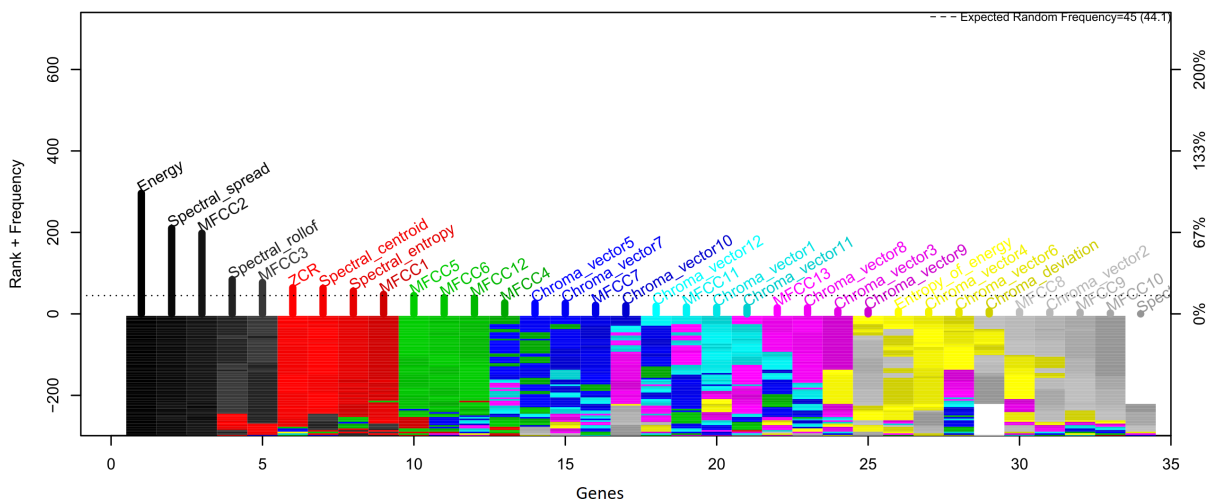


Figura 3.13: Las cinco características elegidas para el clasificador de RF según su frecuencia.

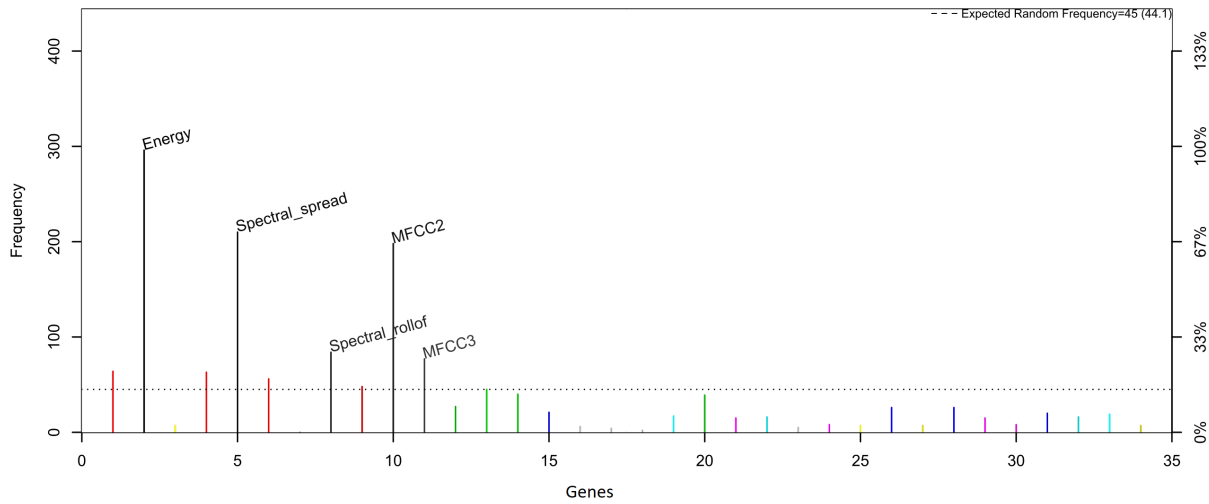


Figura 3.14: Evolución del rango con el clasificador Rpart para las características más frecuentes.

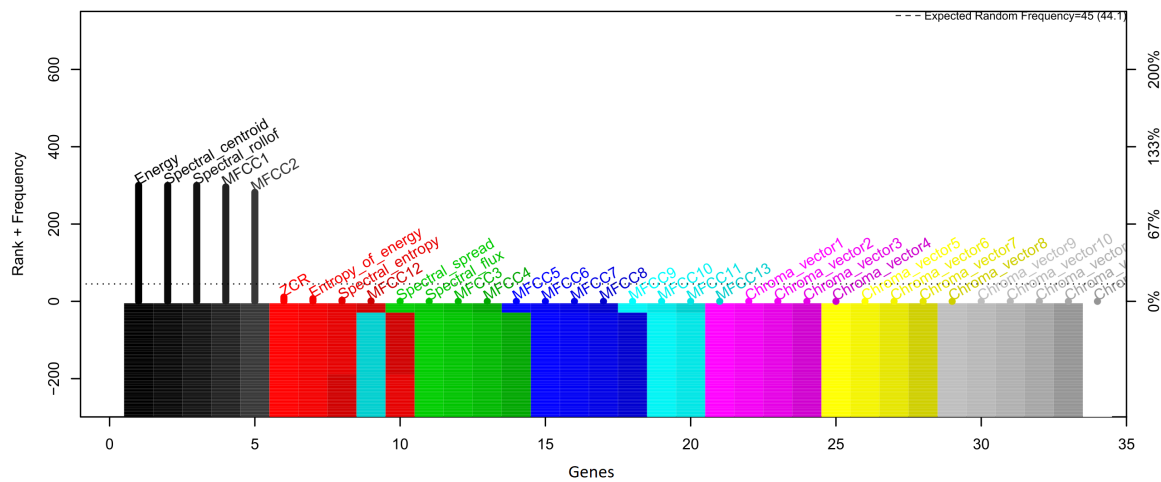
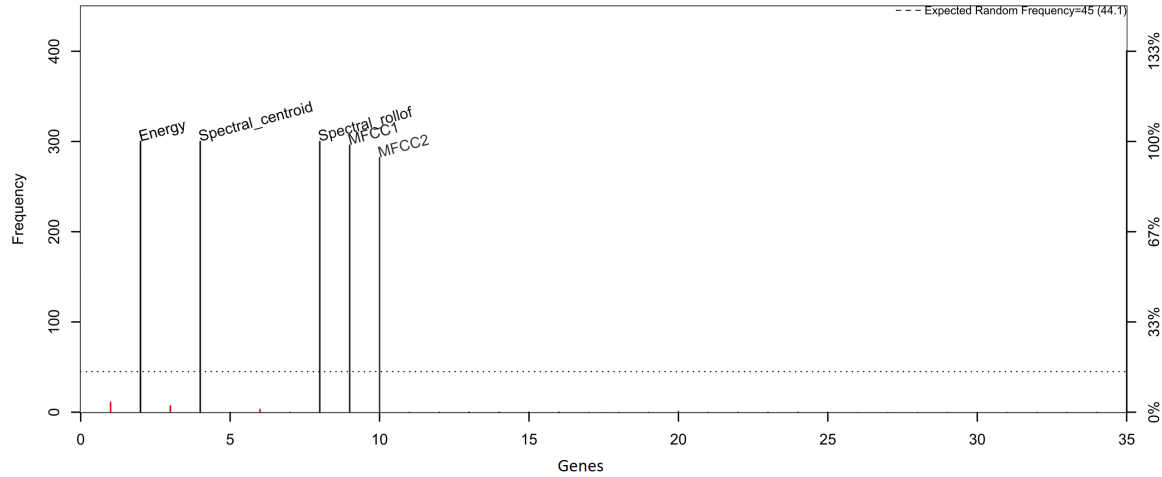


Figura 3.15: Las cinco características elegidas para el clasificador Rpart según su frecuencia.



### 3.5. Resumen

En este capítulo se presentó la forma en la que en este trabajo se llevan a cabo los procesos de extracción y selección de características. En la parte inicial del capítulo se describieron en primer lugar las actividades tomadas en cuenta para su análisis en el presente trabajo sobre reconocimiento y clasificación de actividades infantiles, además se describió el dataset utilizado, así como dispositivos utilizados para su generación.

De manera general, el capítulo describió dos procesos principales e importantes en la generación de modelos de clasificación: extracción de características y selección de características. En la parte de extracción de características se presentaron y describieron las características extraídas de los audios así como la justificación para la elección de estas. En la parte de selección de características se mencionaron los métodos comunes para llevar a cabo este proceso, para después describir cuáles fueron y cómo se aplicaron los métodos implementados en este trabajo.

En la parte final del capítulo se mostraron los resultados obtenidos en el proceso de selección de características con la aplicación de los dos métodos implementados: criterio de Akaike y Algoritmos genéticos. De esta manera, al finalizar este capítulo, se tienen ya los subconjuntos de características a utilizar para la generación de los modelos de clasificación

Una vez teniendo preparado el dataset, así como las características con las que se va a trabajar, en el siguiente capítulo se presentan los modelos de clasificación generados en base a los distintos algoritmos clasificadores que se implementaron, y tomando en cuenta los distintos subconjuntos de características con los que se trabaja. Todo lo anterior para determinar las características que deben de tener los modelos para maximizar el porcentaje de precisión en la clasificación de actividades infantiles.

# Capítulo 4

## Generación de los Modelos de Clasificación

En este capítulo se presentan los modelos de clasificación creados en base a los algoritmos clasificadores mencionados en capítulos anteriores y que se implementan en el presente trabajo, además para la creación de los modelos se consideran tanto las características inicialmente extraídas a los clips de audio como las características resultantes después de la aplicación de los métodos de selección de características implementados.

### 4.1. Introducción

En los capítulos previos de este documento se describen las etapas iniciales de la generación de un modelo de reconocimiento y clasificación de actividades infantiles: extracción y selección de características. Al llegar a este punto, ya se tienen definidas las características extraídas de los clips de audio, así como los distintos subconjuntos de características obtenidos a partir de la selección de características. El siguiente paso es generar los modelos de clasificación mediante la aplicación de algoritmos clasificadores. Finalmente la evaluación de estos modelos se hace tomando en cuenta la precisión alcanzada por cada uno de ellos en la clasificación de las actividades analizadas. Una vez generados y evaluados, los modelos se comparan para obtener y describir aquellos que clasifiquen actividades infantiles con un mayor nivel de precisión.

## 4.2. Descripción de los modelos

Para la generación de un modelo de clasificación es necesario definir el algoritmo clasificador que se utilizará, ya que de ello depende el funcionamiento completo del modelo y el procesamiento que se haga con los datos. Es importante mencionar que otro aspecto importante a considerar en la generación de los modelos de clasificación es la distribución que se hace de los datos para las fases de entrenamiento y prueba del modelo. En este trabajo, para todos los modelos generados, del dataset completo, se utiliza un 70% de los datos para la fase de entrenamiento y un 30% para la fase de prueba.

Durante el tiempo de desarrollo de este trabajo sobre reconocimiento y clasificación de actividades infantiles mediante el sonido ambiental, se llevaron a cabo distintas etapas de experimentación, generando varios modelos de clasificación, estas fases se describen a continuación.

### 4.2.1. Modelos generados con el Dataset original

En esta sección se describen los modelos de clasificación generados tomando en cuenta el conjunto completo de 34 características extraídas a los clips de audio, descritas en secciones anteriores.

#### Fase de Experimentación 1

Una experimentación inicial sobre la generación de modelos de clasificación de actividades infantiles se realizó tomando solo 70 grabaciones de audio, para las cuales se extrajeron el conjunto completo de 34 características mencionado en el capítulo 3. Los algoritmos clasificadores utilizados para la generación de los modelos fueron:

- Support Vector Machine (SVM)
- k-Nearest Neighbor (k-NN)
- Random Forests (RF)

Tabla 4.1: Precisión obtenida por los modelos en la fase de experimentación 1.

Actividad	Clasificadores				
	SVM Precisión	kNN Precisión	RF Precisión	ET Precisión	GB Precisión
Caminar	40.31 %	100 %	76.50 %	100 %	99.93 %
Correr	55.53 %	100 %	71.50 %	100 %	99.95 %
Jugar	74.94 %	100 %	84.00 %	100 %	99.95 %
Llorar	41.93 %	100 %	73.50 %	100 %	99.95 %

- Extra Trees (ET)
- Gradient Boosting (GB)

Se generó y evaluó un modelo de clasificación para cada algoritmo clasificador. Los resultados en base a la precisión obtenida se muestran en la Tabla 4.1.

A partir de los datos presentados en la Tabla 4.1, se puede observar que los modelos generados con los clasificadores k-Nearest Neighbors (k-NN) y Extra trees (ET) fueron los que obtuvieron una precisión mayor (100 %) para la clasificación de actividades infantiles en base a los audios analizados, seguidos por el clasificador Gradient Boosting (GB), el cual obtuvo una precisión del 99.94 % en promedio. Por el contrario, se puede observar que los clasificadores Support Vector Machine (SVM) y Random Forests (RF) clasifican los archivos de audio con una precisión del 53.17 % y 76.37 % en promedio respectivamente.

## Fase de Experimentación 2

En una segunda etapa de experimentación, se generaron modelos de clasificación analizando el dataset completo (146 grabaciones, 2716 clips de 10 segundos) con las mismas 34 características extraídas para los modelos anteriores y utilizando los mismos algoritmos clasificadores. La Tabla 4.2 muestra los verdaderos positivos obtenidos para cada modelo generado y la Tabla 4.3 resume la precisión alcanzada por cada modelo generado con los algoritmos clasificadores implementados, desglosada por actividad.



Tabla 4.2: Verdaderos positivos por actividad para cada modelo de clasificación en la Fase de Experimentación 2.

	Verdaderos positivos				
	SVM	kNN	RF	ET	GB
Caminar	196	199	209	211	211
Correr	180	194	193	195	194
Jugar	177	191	185	191	190
Llorar	167	170	175	177	175

Tabla 4.3: Precisión obtenida por actividad para cada modelo de clasificación en la Fase de Experimentación 2.

	Precisión				
	SVM	kNN	RF	ET	GB
Caminar	0.9245	0.9660	0.9372	0.9548	0.9633
Correr	0.9424	0.9417	0.9847	0.9898	0.9652
Jugar	0.8592	0.9139	0.9257	0.9502	0.9548
Llorar	0.8743	0.9497	0.9665	0.9779	0.9615

La Tabla 4.4 muestra las precisiones promedio de cada modelo considerando todo el conjunto de actividades analizadas.

En esta misma fase de experimentación, se aplicó un proceso de selección de características para obtener subconjuntos de ellas y generar nuevos modelos de clasificación para contrastar los resultados. Esto se describe a continuación.

Tabla 4.4: Precisión promedio para cada modelo de clasificación en la Fase de Experimentación 2.

Precisión promedio	
SVM	0.90
k-NN	0.9425
RF	0.9525
ET	0.9675
GB	0.9612

Tabla 4.5: Verdaderos positivos por actividad para cada modelo de clasificación utilizando el conjunto de 27 características extraído mediante AIC en la Fase de Experimentación 3.

	Verdaderos positivos				
	SVM	k-NN	RF	ET	GB
Caminar	194	200	207	209	209
Correr	178	185	193	198	193
Jugar	177	189	189	193	190
Llorar	169	168	177	177	174

### 4.2.2. Modelos generados con el Dataset reducido

Los modelos generados y evaluados anteriormente trabajan con el conjunto completo de características extraídas. En esta parte del capítulo se presentan los modelos generados que hacen uso de los subconjuntos de características generados en la etapa de selección de características, con el objetivo de comparar los resultados de ambos tipos de modelos.

#### Fase de Experimentación 3

En la fase de experimentación 2 se describieron los modelos creados con el conjunto completo de características y se presentaron los resultados obtenidos. Ligado a esa experimentación, se crearon también, utilizando los mismos algoritmos de clasificación, modelos que utilizan el subconjunto de 27 características resultante de la fase de selección de características mediante el criterio de Akaike (AIC), descrita en el capítulo 3.

A partir del análisis de clasificación basado en las características seleccionadas, los verdaderos positivos obtenidos se muestran en la Tabla 4.5, mientras que la Tabla 4.6 resume la precisión por actividad. La Tabla 4.7 muestra las precisiones promedio para cada técnica de clasificación.

#### Fase de Experimentación 4

Una fase posterior de experimentación en este trabajo sobre reconocimiento y clasificación de actividades infantiles utilizando sonido ambiental consistió en realizar nuevamente

Tabla 4.6: Precisión por actividad para cada modelo de clasificación utilizando el conjunto de 27 características extraído mediante AIC en la Fase de Experimentación 3.

	Precisión				
	SVM	k-NN	RF	ET	GB
Caminar	0.8940	0.9346	0.9539	0.9858	0.9676
Correr	0.9468	0.9343	0.9948	0.9851	0.9747
Jugar	0.8551	0.8915	0.9220	0.9324	0.9179
Llorar	0.8989	0.9545	0.9620	0.9833	0.9721

Tabla 4.7: Precisión promedio para cada modelo de clasificación en la Fase de Experimentación 3 utilizando el conjunto de 27 características extraído mediante AIC .

Precisión promedio	
SVM	0.8975
k-NN	0.9275
RF	0.9575
ET	0.9712
GB	0.9575

un proceso de selección de características mediante otra técnica, Algoritmos Genéticos, como se describió en el capítulo 3. Para la generación de los modelos de clasificación en esta fase se utilizaron las características mostradas en la Tabla 3.8, las cuales fueron las resultantes del proceso de selección de características mediante Algoritmos Genéticos, para cada algoritmo de clasificación implementado. Los algoritmos de clasificación utilizados en esta fase de experimentación fueron:

- k-NN
- NC
- ANN
- RF
- Rpart

En base a los subconjuntos de características utilizados y los algoritmos clasificadores implementados, la Tabla 4.8 muestra la precisión obtenida en cada caso.

Tabla 4.8: Precisión promedio para cada modelo de clasificación en la Fase de Experimentación 4 utilizando el conjunto de 5 características extraído mediante Algoritmos Genéticos.

Algoritmo	Precisión
k-NN	0.890
NC	0.815
ANN	0.850
RF	0.920
Rpart	0.857

Tabla 4.9: Arquitectura propuestas para la Red Neuronal Artificial en la Fase de Experimentación 5.

Inputs	Hidden Layers	Neurons per layer	Outputs	Batch size	Epochs
34	8	20	4	64	200

### Fase de Experimentación 5

Una fase final de experimentación para este trabajo consistió en generar otro modelo de clasificación mediante el diseño de una red neuronal artificial profunda. Para esta fase se utilizó el mismo conjunto de datos que para los modelos generados anteriormente y se utilizó además el conjunto completo de 34 características extraídas de los clips de audio, descritas previamente en este documento. La arquitectura propuesta para la red neuronal creada se presenta en la Tabla 4.9.

En la Tabla 4.10 se presentan los parámetros seleccionados para el desarrollo e implementación del modelo en la interfase Keras con Python.

Una vez que la ANN ha sido diseñada y ejecutada con la arquitectura y la partición de los conjuntos de datos mencionados según lo propuesto, la Figura 4.1 muestra la precisión alcanzada por el modelo a lo largo de las épocas, donde se puede observar que la precisión alcanzada para la clasificación de las actividades de los niños usando el sonido ambiental es 0.9979 para datos de entrenamiento y 0.9451 para datos de prueba.

Tabla 4.10: Parámetros seleccionados para la implementación de la Red Neuronal Artificial en la Fase de Experimentación 5.

Tipo de modelo	Sequential
Función de activación de la capa de entrada	Relu
Función de activación de las capas intermedias	Relu
Función de activación de la capa de salida	Softmax
Función de pérdida	Categorical Crossentropy
Algoritmo de Optimización	Adam

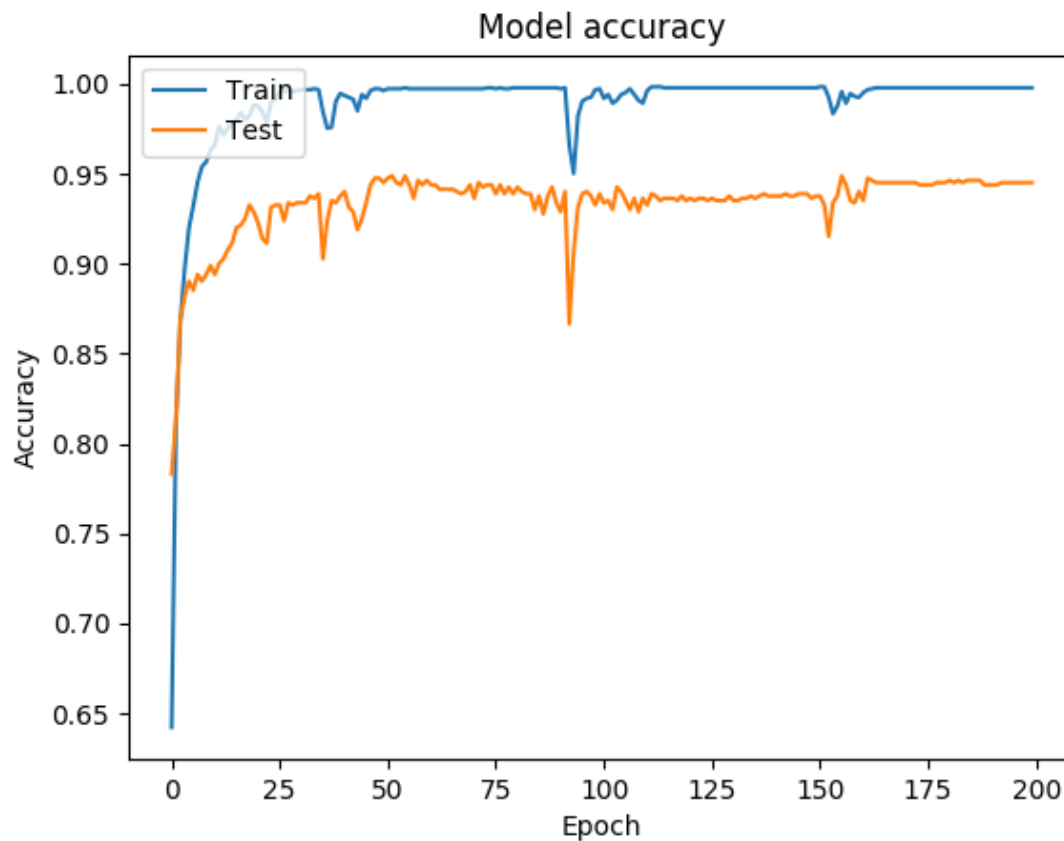


Figura 4.1: Gráfico de comportamiento de la precisión durante el entrenamiento y validación de la Red Neuronal Artificial en la Fase de Experimentación 5.

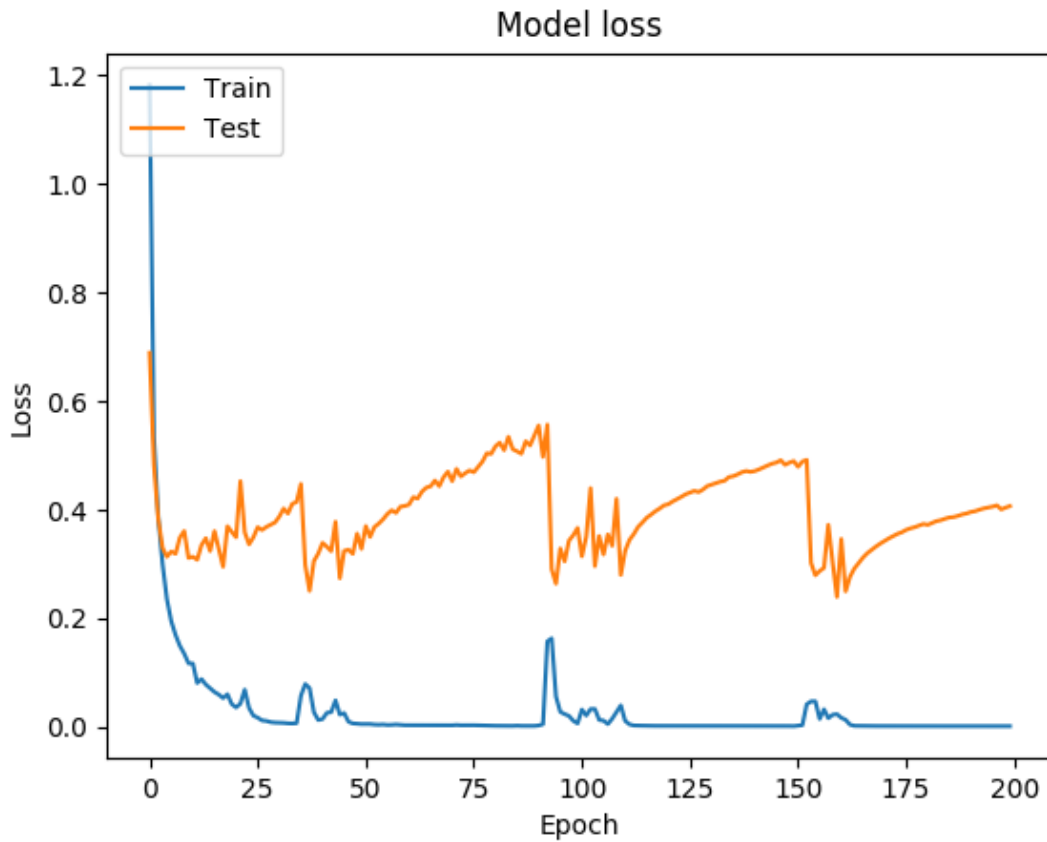


Figura 4.2: Gráfica del comportamiento de la función de pérdida durante el entrenamiento y validación de la Red Neuronal Artificial en la Fase de Experimentación 5.

La Figura 4.2 muestra la pérdida presentada por la ANN, donde se puede observar que la pérdida para los datos de entrenamiento es 0.0018, mientras que para los datos de prueba es 0.4070.

### 4.3. Resumen

En este capítulo se describieron los modelos de clasificación creados a partir de las características extraídas y las características seleccionadas mediante los métodos de selección analizados en el capítulo 3.

Para cada modelo creado se especificó el algoritmo de clasificación utilizado y el conjun-

to de características con el que trabaja. Inicialmente los modelos de clasificación creados utilizaron el conjunto completo de las 34 características extraídas en la etapa de extracción de características, se evaluaron y mostraron sus resultados en términos de precisión en la clasificación de las actividades analizadas.

Finalmente, se describieron también los modelos creados que hacen uso de un subconjunto de características resultante de la etapa de selección de características. En base a eso, se crearon dos tipos de modelos: los que utilizaron el subconjunto de 27 características seleccionadas mediante el criterio de Akaike, y aquellos que utilizaron el subconjunto de 5 características seleccionadas por el método de Algoritmos Genéticos. Para todos los casos se evaluaron los modelos y se presentaron los resultados. En la parte final del capítulo se presentó la última fase de experimentación realizada, que consistió en el diseño e implementación de una red neuronal artificial, con la cual se construyó y evaluó el último de los modelos generados, haciendo uso del conjunto completo de 34 características extraídas.

# Capítulo 5

## Discusión

En este capítulo se presentan las discusiones de los resultados obtenidos en el presente trabajo sobre reconocimiento y clasificación de actividades infantiles utilizando sonido ambiental. Especialmente se discuten los resultados obtenidos de la evaluación de los modelos de clasificación generados, considerando los algoritmos clasificadores implementados y los conjuntos de características analizados.

### **5.1. Clasificación de actividades infantiles mediante sonido ambiental**

En base a lo investigado en el presente trabajo y a las experimentaciones realizadas en la generación de modelos de clasificación, se puede identificar que el sonido ambiental provee suficiente información para el reconocimiento y clasificación de actividades infantiles de manera eficiente.

### **5.2. Modelos de clasificación generados con el conjunto completo de características**

Un elemento fundamental en la generación de modelos de clasificación es el conjunto de características mediante el cual se analizan las actividades. En el presente trabajo se



realizó un proceso de extracción de características sobre los clips de audio teniendo como resultado un conjunto de 34 características, con las cuales se generaron los primeros modelos de clasificación.

Los clasificadores tomados en cuenta en esta etapa fueron: Support Vector Machines (SVM), k Nearest Neighbors (k-NN), Random Forests (RF), Extra trees (ET) y Gradient boosting (GB). Para esta fase inicial de experimentación se trabajó con un dataset de 70 grabaciones de audio. De este análisis se puede identificar lo siguiente:

- Actividades en apariencia similares como caminar y correr pudieron ser clasificadas de forma correcta por los clasificadores ya que contienen suficiente información para diferenciar actividades que generan un sonido ambiental similar.
- k-Nearest Neighbors (k-NN) y Extra trees (ET) resultaron ser los clasificadores más precisos a la hora de evaluar los clips de audio correspondientes a las actividades analizadas.
- En contraste con lo anterior, Support Vector Machines (SVM) y Random Forests (RF) resultaron ser los clasificadores con una menor precisión en la clasificación de las actividades para esta fase de experimentación.

En este sentido, se realizó una segunda fase de experimentación similar a la anterior, solo que ahora tomando en cuenta el conjunto completo de grabaciones de audio (2716 clips), utilizando los mismos algoritmos clasificadores y analizando el mismo conjunto de 34 características extraídas. De este análisis se puede identificar lo siguiente:

- La precisión de los algoritmos clasificadores se mantiene por encima del 85 % para todos los casos.
- La actividad “correr” presenta el mayor porcentaje de precisión en su clasificación, el cual va del 94 % al 98 % dependiendo del algoritmo clasificador utilizado.
- La actividad “jugar” presenta el menor porcentaje de precisión en su clasificación, el cual va del 85 % al 95 % dependiendo del algoritmo clasificador utilizado.

- Tomando en cuenta el conjunto completo de 4 actividades, en promedio, el modelo de clasificación que hace uso del algoritmo Extra Trees tiene el mayor porcentaje de precisión con un 96.75 %, seguido de los algoritmos GB, RF, k-NN y SVM, respectivamente, todos con una precisión promedio igual o superior al 90 %.

Tomando en cuenta también el conjunto completo de grabaciones de audio, de la última fase de experimentación realizada referente a la generación del modelo de clasificación utilizando la red neuronal artificial profunda, se puede identificar lo siguiente:

- Las redes neuronales artificiales profundas son eficientes para generar modelos de clasificación de actividades infantiles a través del sonido ambiental. La red neuronal artificial profunda con la arquitectura propuesta clasifica correctamente las actividades infantiles con una precisión del 94 %, a través del análisis de las características extraídas del sonido ambiental.
- La precisión de la red neuronal artificial profunda es similar a otras técnicas de Aprendizaje Automático reportadas. La red neuronal artificial profunda con la arquitectura propuesta en el presente trabajo logra una precisión similar a la reportada en en otras fases de experimentación, con otras técnicas de Aprendizaje Automático: 100 % para kNN con 34 características y 94.25 % para kNN con 27 características.

### **5.3. Modelos de clasificación generados con selección de características**

Con el objetivo de contrastar los resultados y evaluar el impacto que tiene sobre la precisión de los modelos de clasificación un proceso de selección de características, de manera seguida a la experimentación descrita en el punto anterior, se aplicaron dos métodos de selección de características: criterio de Akaike y Algoritmos Genéticos.

### 5.3.1. Modelos generados con características seleccionadas mediante criterio de Akaike

Como se presentó en el capítulo 3, a partir del proceso de selección de características mediante el criterio de Akaike se obtuvo un nuevo subconjunto de 27 características, con el cual se generaron los modelos descritos en el capítulo 4. Al igual que en la experimentación anterior, los algoritmos clasificadores utilizados fueron Support Vector Machines (SVM), k Nearest Neighbors (k-NN), Random Forests (RF), Extra trees (ET) y Gradient boosting (GB). De los resultados presentados en esta fase de experimentación se puede identificar lo siguiente:

- Mediante la aplicación del proceso de selección de características con el criterio de Akaike, se logró una disminución del 20% en el número de características utilizadas por los modelos para la clasificación de actividades, al pasar de 34 a 27 características.
- Los modelos generados con los algoritmos Random Forests y Extra Trees logran un incremento en la precisión para la clasificación de las actividades.
- De manera general, todos los demás algoritmos mantienen una precisión parecida en estos modelos generados con 27 características en relación a los modelos generados en la fase anterior donde se utilizaba el conjunto completo de 34 características.

### 5.3.2. Modelos generados con características seleccionadas mediante Algoritmos Genéticos

En el capítulo 3 se describió el proceso de selección de características mediante Algoritmos Genéticos, el cual genera subconjuntos de 5 características, uno para cada algoritmo clasificador utilizado. Dichos subconjuntos se reportan en el capítulo 3, y en el capítulo 4 se generan los modelos de clasificación correspondientes que hacen uso de ellos. De los resultados mostrados en esta fase de experimentación se puede identificar lo siguiente:

- El enfoque de generación de modelos de clasificación implementando selección de características mediante Algoritmos Genéticos demuestra una reducción importante

de las características al utilizar solo el 14 % de las características originales, al pasar de 34 a 5 características.

- Las figuras que corresponden al proceso evolutivo de k-NN, NC y Rpart, presentan un comportamiento particular. Estos algoritmos de bajo costo computacional en el proceso de mutación incluyen características que disminuyen la idoneidad del modelo, y finalmente estas características se excluyen casi inmediatamente de esta selección.
- Algoritmos como RF y ANN, tienen un proceso evolutivo suave, dadas las capacidades para encontrar una relación más compleja entre las características.
- Los conjuntos de cinco características determinados por el algoritmo genético utilizado, alcanzan precisiones similares a los modelos generados utilizando el conjunto original de características, pero con la ventaja de que hay una reducción del 80 % en el tamaño del conjunto de datos

## 5.4. Limitaciones Metodológicas

Durante el desarrollo del presente trabajo sobre reconocimiento y clasificación de actividades infantiles utilizando sonido ambiental se llevaron a cabo varios procesos, los cuales fueron acotados a ciertas condiciones que pueden afectar la precisión de los modelos de clasificación generados, en este sentido se identifica lo siguiente:

- Durante el proceso de captura de datos, para la parte del dataset generado de manera propia, se grabaron los audios de las actividades con un solo niño realizándolas de manera separada. Si se requieren reconocer y clasificar actividades de 2 o más niños interactuando o realizando más de 1 actividad a la vez, es necesario realizar otro tipo de análisis que este trabajo no contempla.
- El análisis realizado sobre los clips de audio no contempla métodos de eliminación de ruido, por lo que si se analizan grabaciones que contengan ruidos ajenos a la realización de la actividad, la precisión de los modelos en la clasificación de las actividades podría variar.

# Capítulo 6

## Conclusiones y Trabajo Futuro

### 6.1. Conclusiones

En el presente trabajo se propuso la utilización del sonido ambiental como fuente de datos para la generación de un modelo de clasificación de actividades infantiles. Como primer proceso se describe la recolección de datos, la cual se llevó a cabo en dos etapas, una generándolos y recolectándolos de manera propia y la otra tomándolos de Internet.

Se presentó además el proceso de extracción de características mediante dos métodos: criterio de Akaike y Algoritmos Genéticos. De este proceso se observó que mediante ambos métodos se logra una reducción significativa de características en relación al conjunto de características original, 20 % de reducción con criterio de Akaike y 85 % de reducción con Algoritmos Genéticos. A partir de estos resultados se puede concluir que los datos recolectados del sonido ambiental contiene información que no es relevante para la diferenciación de las actividades a clasificar y pudo ser omitida mediante los procesos de selección de características implementados. La reducción en el número de características es importante porque cuando las técnicas de clasificación están sujetas a grandes cantidades de información, el tiempo de respuesta suele ser largo de manera significativa, aumentando el costo computacional, además del hecho de que el reconocimiento y la clasificación de actividades los modelos generalmente están diseñados para implementarse en aplicaciones móviles, por lo que es importante optimizar la cantidad de datos con los que trabajará el usuario

y reducir el costo de procesamiento.

Analizando la generación y evaluación de los modelos, la mayoría de ellos muestran un comportamiento similar en cuanto a la precisión en la clasificación se refiere. La diferencia radica entonces en el número de características que analiza cada modelo. Por lo tanto, a partir de los resultados obtenidos en el capítulo 4, se puede concluir que los conjuntos de cinco características determinados por el algoritmo genético utilizado, alcanzan precisiones similares a los modelos generados utilizando el conjunto original de características y a los modelos generados utilizando las características seleccionadas mediante el criterio de Akaike, pero con la ventaja de que hay una reducción del 85 % en el tamaño del conjunto de datos, incluidas todas sus implicaciones, como un procesamiento de datos más rápido, que es muy importante cuando se trabaja con aplicaciones para dispositivos móviles ya que sus recursos son limitados.

## 6.2. Publicaciones

Los resultados de este trabajo de investigación se han publicado en 3 trabajos de conferencia (2 nacionales y 1 internacional), además de 1 artículo en revista indizada (Hindawi).

En el primer trabajo de conferencia nacional se presentó un enfoque preliminar para la clasificación de actividades infantiles utilizando sonido ambiental, analizando los archivos de audio mediante su espectro de frecuencia para determinar si las actividades registradas en el sonido ambiental son o no diferenciables.

- Antonio García-Domínguez and Carlos E Galván-Tejada. Reconocimiento de actividades infantiles utilizando sonido ambiental: Un enfoque preliminar. *Research in Computing Science*, 139:71–79, 2017.

En el segundo trabajo de conferencia nacional se comparó el nivel de precisión de los clasificadores Support Vector Machines, k-Nearest Neighbors, Random forests, Extra trees y Gradient boosting en el reconocimiento de actividades infantiles utilizando sonido ambiental.

- Diego M Blanco-Murillo, Antonio Garcia-Dominguez, Carlos E Galván-Tejada, and José M Celaya-Padilla. Comparación del nivel de precisión de los clasificadores support vector machines, k nearest neighbors, random forests, extra trees y gradient boosting en el reconocimiento de actividades infantiles utilizando sonido ambiental. *Research in Computing Science*, 147:281–290, 2018.

En el trabajo de conferencia internacional se propuso el criterio de Akaike como método para selección de características en la generación de los modelos de clasificación.

- Antonio García-Domínguez, Laura A Zanella-Calzada, Carlos E Galván-Tejada, Jorge I Galván-Tejada, and José M Celaya-Padilla. Evaluation of five classifiers for children activity recognition with sound as information source and akaike criterion for feature selection. In *Mexican Conference on Pattern Recognition*, pages 398–407. Springer, 2019

En el trabajo de revista indizada se generaron y evaluaron modelos de clasificación utilizando algoritmos genéticos como método de selección de características.

- Antonio García-Dominguez, Carlos E Galvan-Tejada, Laura A Zanella-Calzada, Hamurabi Gamboa-Rosales, Jorge I Galván-Tejada, José M Celaya-Padilla, Huizilopztli Luna-García, and Rafael Magallanes-Quintanar. Feature selection using genetic algorithms for the generation of a recognition and classification of children activities model using environmental sound. *Mobile Information Systems*, 2020, 2020.

Además, se encuentra en revisión otro trabajo de revista indizada, referente a la generación de un modelo de clasificación mediante una red neuronal artificial profunda.

### 6.3. Trabajo Futuro

Una vez desarrollado este trabajo de tesis, se identificaron los siguientes puntos como trabajo futuro:

- Trabajar con un conjunto diferente de características. Extraer un conjunto de características más grande de los archivos de audio para verificar el comportamiento de los métodos de selección de características analizados con un conjunto diferente de características y contrastar los resultados con los obtenidos en este trabajo para finalmente determinar cuál es el conjunto de características que mejor describe las actividades.
- Combinar métodos de selección de características. Generar modelos de clasificación utilizando subconjuntos de características resultantes de la combinación de los métodos propuesto en el presente trabajo con los métodos clásicos de selección de características como la selección hacia adelante y la eliminación hacia atrás.
- Agregar al análisis más actividades comunes en niños con el rango de edad establecido. Ampliar el alcance del trabajo en cuanto a número de actividades analizadas se refiere, con la finalidad de considerar otro tipo de actividades que también puedan ser potencialmente peligrosas para los niños y que puedan ser detectables mediante el análisis del sonido ambiental.
- Realizar un análisis de validación del conjunto de datos para establecer si el número de características y muestras es óptimo para el tipo de estudio que se está realizando.



# Bibliografía

- [1] Organización mundial de la salud. <https://www.who.int/es>. (Accessed on 02/03/2020).
- [2] Rimantas Adaškevičius. Children activity recognition from accelerometer data. *BIO-MEDICAL ENGINEERING 2016*, 17(1), 2013.
- [3] Jake K Aggarwal and Lu Xia. Human activity recognition from 3d data: A review. *Pattern Recognition Letters*, 48:70–80, 2014.
- [4] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.
- [5] Vito Albino, Umberto Berardi, and Rosa Maria Dangelico. Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of urban technology*, 22(1):3–21, 2015.
- [6] Belal Alsinglawi, Quang Vinh Nguyen, Upul Gunawardana, Anthony Maeder, and Simeon J Simoff. Rfid systems in healthcare settings and activity of daily living in smart homes: A review. *E-Health Telecommunication Systems and Networks*, pages 1–17, 2017.
- [7] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge Luis Reyes-Ortiz. A public domain dataset for human activity recognition using smartphones. In *Esann*, 2013.

- [8] G Anitha and S Baghavathi Priya. Posture based health monitoring and unusual behavior recognition system for elderly using dynamic bayesian network. *Cluster Computing*, 22(6):13583–13590, 2019.
- [9] Ferhat Attal, Samer Mohammed, Mariam Dedabrishvili, Faicel Chamroukhi, Latifa Oukhellou, and Yacine Amirat. Physical human activity recognition using wearable sensors. *Sensors*, 15(12):31314–31338, 2015.
- [10] Michael Batty, Kay W Axhausen, Fosca Giannotti, Alexei Pozdnoukhov, Armando Bazzani, Monica Wachowicz, Georgios Ouzounis, and Yuval Portugali. Smart cities of the future. *The European Physical Journal Special Topics*, 214(1):481–518, 2012.
- [11] Kella BhanuJyothi, K Hima Bindu, and D Suryanarayana. A comparative study of random forest & k-nearest neighbors on har dataset using caret. *IJIRT*, 3:6–9, 2017.
- [12] Valentina Bianchi, Marco Bassoli, Gianfranco Lombardo, Paolo Fornacciari, Monica Mordonini, and Ilaria De Munari. Iot wearable sensor and deep learning: An integrated approach for personalized human activity recognition in a smart home environment. *IEEE Internet of Things Journal*, 6(5):8553–8562, 2019.
- [13] Kamanashis Biswas and Vallipuram Muthukkumarasamy. Securing smart cities using blockchain technology. In *2016 IEEE 18th international conference on high performance computing and communications; IEEE 14th international conference on smart city; IEEE 2nd international conference on data science and systems (HPCC/SmartCity/DSS)*, pages 1392–1393. IEEE, 2016.
- [14] Sabri Boughorbel, Jeroen Breebaart, Fons Bruekers, Ingrid Flinsenbergh, and Warner Ten Kate. Child-activity recognition from multi-sensor data. *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research - MB 10*, 2010.
- [15] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [16] Bruce G Buchanan. A (very) brief history of artificial intelligence. *Ai Magazine*, 26(4):53–53, 2005.

- [17] Nicole A Capela, Edward D Lemaire, and Natalie Baddour. Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients. *PloS one*, 10(4), 2015.
- [18] Pedro Chahuara, Anthony Fleury, François Portet, and Michel Vacher. On-line human activity recognition from audio and home automation sensors: Comparison of sequential and non-sequential models in realistic smart homes 1. *Journal of ambient intelligence and smart environments*, 8(4):399–422, 2016.
- [19] Girish Chandrashekar and Ferat Sahin. A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28, 2014.
- [20] Jose M Chaquet, Enrique J Carmona, and Antonio Fernández-Caballero. A survey of video datasets for human action and activity recognition. *Computer Vision and Image Understanding*, 117(6):633–659, 2013.
- [21] Yufei Chen and Chao Shen. Performance analysis of smartphone-sensor behavior for human activity recognition. *Ieee Access*, 5:3095–3110, 2017.
- [22] Chieh-Feng Cheng, Abbas Rashidi, Mark A Davenport, and David Anderson. Audio signal processing for activity recognition of construction heavy equipment. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, volume 33, page 1. IAARC Publications, 2016.
- [23] Wei-Yi Cheng, Alf Scotland, Florian Lipsmeier, Timothy Kilchenmann, Liping Jin, Jens Schjodt-Eriksen, Detlef Wolf, Yan-Ping Zhang-Schaerer, Ignacio Fernandez Garcia, Juliane Siebourg-Polster, et al. Human activity recognition from sensor-based large-scale continuous monitoring of parkinson’s disease patients. In *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pages 249–250. IEEE, 2017.
- [24] Hafedh Chourabi, Taewoo Nam, Shawn Walker, J Ramon Gil-Garcia, Sehl Mellouli, Karine Nahon, Theresa A Pardo, and Hans Jochen Scholl. Understanding smart

- cities: An integrative framework. In *2012 45th Hawaii international conference on system sciences*, pages 2289–2297. IEEE, 2012.
- [25] Amrita Roy Chowdhury, Tamojit Chatterjee, and Sreeparna Banerjee. A random forest classifier-based approach in the detection of abnormalities in the retina. *Medical & biological engineering & computing*, 57(1):193–203, 2019.
- [26] Diane J Cook, Juan C Augusto, and Vikramaditya R Jakkula. Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4):277–298, 2009.
- [27] Corinna Cortes and Vladimir Vapnik. Support vector machine. *Machine learning*, 20(3):273–297, 1995.
- [28] Federico Cruciani, Ian Cleland, Chris Nugent, Paul McCullagh, Kåre Synnes, and Josef Hallberg. Automatic annotation for human activity recognition in free living using a smartphone. *Sensors*, 18(7):2203, 2018.
- [29] Božidara Cvetković, Vito Janko, Alfonso E Romero, Özgür Kafalı, Kostas Stathis, and Mitja Luštrek. Activity recognition for diabetic patients using a smartphone. *Journal of medical systems*, 40(12):256, 2016.
- [30] Francisco António dos Santos da Silva and Tiago Arruda Ferreira Marques Lopes. Application of smart tourism to nature-based destinations. In *Strategic Business Models to Support Demand, Supply, and Destination Management in the Tourism and Hospitality Industry*, pages 1–20. IGI Global, 2020.
- [31] Manoranjan Dash and Huan Liu. Feature selection for classification. *Intelligent data analysis*, 1(3):131–156, 1997.
- [32] G De Leonardis, Samanta Rosati, Gabriella Balestra, Valentina Agostini, E Pane-ro, Laura Gastaldi, and Marco Knafnitz. Human activity recognition by wearable sensors: Comparison of different classifiers for real-time applications. In *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–6. IEEE, 2018.

- [33] Juan Rubén Delgado-Contreras, Juan Pablo García-Vázquez, Ramon F Brena, Carlos E Galván-Tejada, and Jorge I Galván-Tejada. Feature selection for place classification through environmental sounds. *Procedia Computer Science*, 37:40–47, 2014.
- [34] P Dhanalakshmi, S Palanivel, and Vennila Ramalingam. Classification of audio signals using svm and rbfnn. *Expert systems with applications*, 36(3):6069–6075, 2009.
- [35] Nomusa Dlodlo, Oscar Gcaba, and Andrew Smith. Internet of things technologies in smart cities. In *2016 IST-Africa Week Conference*, pages 1–7. IEEE, 2016.
- [36] Eduardo Fonseca, Jordi Pons Puig, Xavier Favory, Frederic Font Corbera, Dmitry Bogdanov, Andres Ferraro, Sergio Oramas, Alastair Porter, and Xavier Serra. Free-sound datasets: a platform for the creation of open audio datasets. In *Hu X, Cunningham SJ, Turnbull D, Duan Z, editors. Proceedings of the 18th ISMIR Conference; 2017 oct 23-27; Suzhou, China.[Canada]: International Society for Music Information Retrieval; 2017. p. 486-93*. International Society for Music Information Retrieval (ISMIR), 2017.
- [37] Dany Fortin-Simard, Jean-Sébastien Bilodeau, Kevin Bouchard, Sebastien Gaboury, Bruno Bouchard, and Abdenour Bouzouane. Exploiting passive rfid technology for activity recognition in smart homes. *IEEE Intelligent systems*, 30(4):7–15, 2015.
- [38] Carlos E Galván-Tejada, Jorge I Galván-Tejada, José M Celaya-Padilla, J Rubén Delgado-Contreras, Rafael Magallanes-Quintanar, Margarita L Martinez-Fierro, Idalia Garza-Veloz, Yamilé López-Hernández, and Hamurabi Gamboa-Rosales. An analysis of audio features to develop a human activity recognition model using genetic algorithms, random forests, and neural networks. *Mobile Information Systems*, 2016, 2016.
- [39] Carlos E Galván-Tejada, FE López-Monteagudo, O Alonso-González, Jorge I Galván-Tejada, José M Celaya-Padilla, Hamurabi Gamboa-Rosales, Rafael Magallanes-Quintanar, and Laura A Zanella-Calzada. A generalized model for in-

- door location estimation using environmental sound from human activity recognition. *ISPRS International Journal of Geo-Information*, 7(3):81, 2018.
- [40] Carlos Eric Galván-Tejada, Jorge I. Galván-Tejada, José M. Celaya-Padilla, J. Ruben Delgado-Contreras, Rafael Magallanes-Quintanar, Margarita L. Martinez-Fierro, Idalia Garza-Veloz, Yamilé López-Hernández, and Hamurabi Gamboa Rosales. An analysis of audio features to develop a human activity recognition model using genetic algorithms, random forests, and neural networks. *Mobile Information Systems*, 2016:1784101:1–1784101:10, 2016.
- [41] Enrique Garcia-Ceja and Ramon F Brena. An improved three-stage classifier for activity recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 32(01):1860003, 2018.
- [42] Behnaz Ghoraani and Sridhar Krishnan. Time–frequency matrix feature extraction and classification of environmental audio signals. *IEEE transactions on audio, speech, and language processing*, 19(7):2197–2209, 2011.
- [43] Junqi Guo, Xi Zhou, Yunchuan Sun, Gong Ping, Guoxing Zhao, and Zhuorong Li. Smartphone-based patients’ activity recognition by using a self-learning scheme for medical monitoring. *Journal of medical systems*, 40(6):140, 2016.
- [44] Mohammad Mehedi Hassan, Shamsul Huda, Md Zia Uddin, Ahmad Almogren, and Majed Alrubaian. Human activity recognition from body sensor data using deep learning. *Journal of medical systems*, 42(6):99, 2018.
- [45] Mohammed Mehedi Hassan, Md Zia Uddin, Amr Mohamed, and Ahmad Almogren. A robust human activity recognition system using smartphone sensors and deep learning. *Future Generation Computer Systems*, 81:307–313, 2018.
- [46] Chunyu Hu, Yiqiang Chen, Lisha Hu, and Xiaohui Peng. A novel random forests based class incremental learning method for activity recognition. *Pattern Recognition*, 78:277–290, 2018.

- [47] Ling Hu and Qiang Ni. Iot-driven automated object detection algorithm for urban surveillance systems in smart cities. *IEEE Internet of Things Journal*, 5(2):747–754, 2017.
- [48] Barbara Hyman, Zahra Alisha, and Scott Gordon. Secure controls for smart cities; applications in intelligent transportation systems and smart buildings. 2019.
- [49] Andrey Ignatov. Real-time human activity recognition from accelerometer data using convolutional neural networks. *Applied Soft Computing*, 62:915–922, 2018.
- [50] Masaya Inoue, Sozo Inoue, and Takeshi Nishida. Deep recurrent neural network for mobile human activity recognition with high throughput. *Artificial Life and Robotics*, 23(2):173–185, 2018.
- [51] Ahmad Jalal, Shaharyar Kamal, and Daijin Kim. Shape and motion features approach for activity tracking and recognition from kinect video camera. In *2015 IEEE 29th International Conference on Advanced Information Networking and Applications Workshops*, pages 445–450. IEEE, 2015.
- [52] Ahmad Jalal, Shaharyar Kamal, and Daijin Kim. A depth video-based human detection and activity recognition using multi-features and embedded hidden markov models for health care monitoring systems. *International Journal of Interactive Multimedia & Artificial Intelligence*, 4(4), 2017.
- [53] Ahmad Jalal, Yeon-Ho Kim, Yong-Joong Kim, Shaharyar Kamal, and Daijin Kim. Robust human activity recognition from depth video using spatiotemporal multi-fused features. *Pattern recognition*, 61:295–308, 2017.
- [54] Ahmad Jalal, Majid Ali Khan Quaid, and Abdul S Hasan. Wearable sensor-based human behavior understanding and recognition in daily life for smart environments. In *2018 International Conference on Frontiers of Information Technology (FIT)*, pages 105–110. IEEE, 2018.

- [55] Majid Janidarmian, Atena Roshan Fekr, Katarzyna Radecka, and Zeljko Zilic. A comprehensive analysis on wearable acceleration sensors in human activity recognition. *Sensors*, 17(3):529, 2017.
- [56] Wenchao Jiang and Zhaozheng Yin. Human activity recognition using wearable sensors by deep convolutional neural networks. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 1307–1310, 2015.
- [57] Michael I Jordan and Tom M Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.
- [58] Esra Mahsereci Karabulut, Selma Ayşe Özel, and Turgay Ibrikci. A comparative study on the effect of feature selection on classification accuracy. *Procedia Technology*, 1:323–327, 2012.
- [59] D Karthikeswaran, N Sengottaiyan, and S Anbukaruppusamy. Video surveillance system against anti-terrorism by using adaptive linear activity classification (alac) technique. *Journal of medical systems*, 43(8):256, 2019.
- [60] Alboukadel Kassambara. *Practical guide to cluster analysis in R: Unsupervised machine learning*, volume 1. STHDA, 2017.
- [61] Zafar A Khan and Won Sohn. Abnormal human activity recognition system based on r-transform and kernel discriminant technique for elderly home care. *IEEE Transactions on Consumer Electronics*, 57(4):1843–1850, 2011.
- [62] Kibum Kim, Ahmad Jalal, and Maria Mahmood. Vision-based human activity recognition system using depth silhouettes: A smart home system for monitoring the residents. *Journal of Electrical Engineering & Technology*, 14(6):2567–2573, 2019.
- [63] Kenji Kira and Larry A Rendell. A practical approach to feature selection. In *Machine Learning Proceedings 1992*, pages 249–256. Elsevier, 1992.
- [64] David G Kleinbaum, K Dietz, M Gail, Mitchel Klein, and Mitchell Klein. *Logistic regression*. Springer, 2002.



- [65] Daphne Koller and Mehran Sahami. Toward optimal feature selection. Technical report, Stanford InfoLab, 1996.
- [66] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160:3–24, 2007.
- [67] Max Kuhn and Kjell Johnson. *Applied predictive modeling*, volume 26. Springer, 2013.
- [68] Chandrashekar MC Kushbu and MZ Kurian. Design and implementation of child activity recognition using accelerometer and rfid cards. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 3(4):1437–1440, 2014.
- [69] Alok Kumar Singh Kushwaha and Rajeev Srivastava. Multiview human activity recognition system based on spatiotemporal template for video surveillance system. *Journal of Electronic Imaging*, 24(5):051004, 2015.
- [70] Min-Cheol Kwon and Sunwoong Choi. Recognition of daily human activity using an artificial neural network and smartwatch. *Wireless Communications and Mobile Computing*, 2018, 2018.
- [71] Shancang Li, Li Da Xu, and Shanshan Zhao. The internet of things: a survey. *Information Systems Frontiers*, 17(2):243–259, 2015.
- [72] Xinyu Li, Dongyang Yao, Xuechao Pan, Jonathan Johannaman, JaeWon Yang, Rachel Webman, Aleksandra Sarcevic, Ivan Marsic, and Randall S Burd. Activity recognition for medical teamwork based on passive rfid. In *2016 IEEE International Conference on RFID (RFID)*, pages 1–9. IEEE, 2016.
- [73] Jian-Xian Lin, Zu-Kai Wang, Wei Wang, Jian-Wei Xie, Jia-Bin Wang, Jun Lu, Qi-Yue Chen, Long-Long Cao, Mi Lin, Ru-Hong Tu, et al. Development and validation of a new staging system for node-negative gastric cancer based on recursive partitioning

- analysis: An international multi-institutional study. *Cancer medicine*, 8(6):2962–2970, 2019.
- [74] Pierre Lison. An introduction to machine learning. *Language Technology Group: Edinburgh, UK*, 2015.
- [75] Luca Lonini, Aakash Gupta, Konrad Kording, and Arun Jayaraman. Activity recognition in patients with lower limb impairments: do we need training data from each patient? In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3265–3268. IEEE, 2016.
- [76] John Loughrey and Pádraig Cunningham. Overfitting in wrapper-based feature subset selection: The harder you try the worse it gets. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 33–43. Springer, 2004.
- [77] Hema Rao Madala. *Inductive Learning Algorithms for Complex Systems Modeling: 0*. cRc press, 2018.
- [78] Adrien Malaisé, Pauline Maurice, Francis Colas, François Charpillet, and Serena Ivaldi. Activity recognition with multiple wearable sensors for industrial applications. 2018.
- [79] Matteo Mascia, Antonio Canclini, Fabio Antonacci, Marco Tagliasacchi, Augusto Sarti, and Stefano Tubaro. Forensic and anti-forensic analysis of indoor/outdoor classifiers based on acoustic clues. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, pages 2072–2076. IEEE, 2015.
- [80] Fabien Massé, Roman R Gonzenbach, Arash Arami, Anisoara Paraschiv-Ionescu, Andreas R Luft, and Kamiar Aminian. Improving activity recognition using a wearable barometric pressure sensor in mobility-impaired stroke patients. *Journal of neuroengineering and rehabilitation*, 12(1):72, 2015.
- [81] Martin McKinney and Jeroen Breebaart. Features for audio and music classification. 2003.

- [82] Hoday Danaei Mehr, Huseyin Polat, and Aydin Cetin. Resident activity recognition in smart homes by using artificial neural networks. In *2016 4th international istanbul smart grid congress and fair (ICSG)*, pages 1–5. IEEE, 2016.
- [83] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.
- [84] Loris Nanni, Yandre MG Costa, Diego Rafael Lucio, Carlos Nascimento Silla Jr, and Sheryl Brahnham. Combining visual and acoustic features for audio classification tasks. *Pattern Recognition Letters*, 88:49–56, 2017.
- [85] Urbano Miguel Nunes, Diego R Faria, and Paulo Peixoto. A human activity recognition framework using max-min features and key poses with differential evolution random forests classifier. *Pattern Recognition Letters*, 99:21–31, 2017.
- [86] Henry Friday Nweke, Ying Wah Teh, Mohammed Ali Al-Garadi, and Uzoma Rita Alo. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Systems with Applications*, 105:233–261, 2018.
- [87] Ziad Obermeyer and Ezekiel J Emanuel. Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, 375(13):1216, 2016.
- [88] Leonardo Onofri, Paolo Soda, Mykola Pechenizkiy, and Giulio Iannello. A survey on using domain and contextual knowledge for human activity recognition in video streams. *Expert Systems with Applications*, 63:97–111, 2016.
- [89] Venet Osmani, Daqing Zhang, and Sasitharan Balasubramaniam. Human activity recognition supporting context-appropriate reminders for elderly. In *2009 3rd International Conference on Pervasive Computing Technologies for Healthcare*, pages 1–4. IEEE, 2009.
- [90] Leif E Peterson. K-nearest neighbor. *Scholarpedia*, 4(2):1883, 2009.

- [91] V Praveen, V Hemalatha, and P Gomathi. A nearest centroid classifier-based clustering algorithm for solving vehicle routing problem. In *Innovations in Electronics and Communication Engineering*, pages 575–586. Springer, 2018.
- [92] Carlos Ramos, Juan Carlos Augusto, and Daniel Shapiro. Ambient intelligence—the next step for artificial intelligence. *IEEE Intelligent Systems*, 23(2):15–18, 2008.
- [93] Mamata Rath. Big data and iot-allied challenges associated with healthcare applications in smart and automated systems. In *Data Analytics in Medicine: Concepts, Methodologies, Tools, and Applications*, pages 1401–1414. IGI Global, 2020.
- [94] Jorge-L Reyes-Ortiz, Luca Oneto, Albert Samà, Xavier Parra, and Davide Anguita. Transition-aware human activity recognition using smartphones. *Neurocomputing*, 171:754–767, 2016.
- [95] Jorge-Luis Reyes-Ortiz, Luca Oneto, Alessandro Ghio, Albert Samà, Davide Anguita, and Xavier Parra. Human activity recognition on smartphones with awareness of basic activities and postural transitions. In *International conference on artificial neural networks*, pages 177–184. Springer, 2014.
- [96] Charissa Ann Ronao and Sung-Bae Cho. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert systems with applications*, 59:235–244, 2016.
- [97] Wenjie Ruan. Unobtrusive human localization and activity recognition for supporting independent living of the elderly. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pages 1–3. IEEE, 2016.
- [98] Fariba Sadri. Ambient intelligence: A survey. *ACM Computing Surveys (CSUR)*, 43(4):1–66, 2011.
- [99] Justin Salamon and Juan Pablo Bello. Unsupervised feature learning for urban sound classification. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 171–175. IEEE, 2015.

- [100] Jignyasa Sanghavi. Review of smart healthcare systems and applications for smart cities. In *ICCCE 2019*, pages 325–331. Springer, 2020.
- [101] Sadiq Sani, Nirmalie Wiratunga, Stewart Massie, and Kay Cooper. knn sampling for personalised human activity recognition. In *International conference on case-based reasoning*, pages 330–344. Springer, 2017.
- [102] Ayse Pinar Saygin, Ilyas Cicekli, and Varol Akman. Turing test: 50 years later. *Minds and machines*, 10(4):463–518, 2000.
- [103] Eric D Scheirer. Tempo and beat analysis of acoustic musical signals. *The Journal of the Acoustical Society of America*, 103(1):588–601, 1998.
- [104] Veronica Scuotto, Alberto Ferraris, and Stefano Bresciani. Internet of things: applications and challenges in smart cities. a case study of ibm smart city projects. *Business Process Management Journal*, 2016.
- [105] George AF Seber and Alan J Lee. *Linear regression analysis*, volume 329. John Wiley & Sons, 2012.
- [106] Muhammad Shoaib, Stephan Bosch, Ozlem Durmaz Incel, Hans Scholten, and Paul JM Havinga. Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors*, 16(4):426, 2016.
- [107] Osvaldo Simeone. A very brief introduction to machine learning with applications to communication systems. *IEEE Transactions on Cognitive Communications and Networking*, 4(4):648–664, 2018.
- [108] Amanpreet Singh, Narina Thakur, and Aakanksha Sharma. A review of supervised machine learning algorithms. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 1310–1315. Ieee, 2016.
- [109] Donald F Specht et al. A general regression neural network. *IEEE transactions on neural networks*, 2(6):568–576, 1991.

- [110] Johannes A Stork, Luciano Spinello, Jens Silva, and Kai O Arras. Audio-based human activity recognition using non-markovian ensemble voting. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pages 509–514. IEEE, 2012.
- [111] T Subetha and S Chitrakala. A survey on human activity recognition from videos. In *2016 International Conference on Information Communication and Embedded Systems (ICICES)*, pages 1–7. IEEE, 2016.
- [112] Abdul Syafiq Abdull Sukor, Ammar Zakaria, Norasmadi Abdul Rahim, Latifah Munirah Kamarudin, Rossi Setchi, and Hiromitsu Nishizaki. A hybrid approach of knowledge-driven and data-driven reasoning for activity recognition in smart homes. *Journal of Intelligent & Fuzzy Systems*, 36(5):4177–4188, 2019.
- [113] Satoshi Suzuki, Yasue Mitsukura, Hiroshi Igarashi, Harumi Kobayashi, and Fumio Harashima. Activity recognition for children using self-organizing map. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pages 653–658. IEEE, 2012.
- [114] Stephen P Tarzia, Peter A Dinda, Robert P Dick, and Gokhan Memik. Indoor localization without infrastructure using the acoustic background spectrum. In *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pages 155–168, 2011.
- [115] R Core Team et al. R: A language and environment for statistical computing. 2013.
- [116] Phan Thanh Noi and Martin Kappas. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors*, 18(1):18, 2018.
- [117] Victor Trevino and Francesco Falciani. Galgo: an r package for multivariate variable selection using genetic algorithms. *Bioinformatics*, 22(9):1154–1156, 2006.

- [118] Mostafa Uddin, Ahmed Salem, Ilho Nam, and Tamer Nadeem. Wearable sensing framework for human activity monitoring. In *Proceedings of the 2015 workshop on Wearable Systems and Applications*, pages 21–26, 2015.
- [119] Francis Vekeman, Jesus Eric Piña-Garza, Wendy Y Cheng, Edward Tuttle, Philippe Giguère-Duval, Arman Oganisian, Joseph Damron, Mei Sheng Duh, Vivienne Shen, Timothy B Saurer, et al. Development of a classifier to identify patients with probable lennox–gastaut syndrome in health insurance claims databases via random forest methodology. *Current medical research and opinion*, 35(8):1415–1420, 2019.
- [120] Wim Verhaegh, Wim Verhaegh, Emile Aarts, and Jan Korst. *Algorithms in ambient intelligence*, volume 2. Springer Science & Business Media, 2004.
- [121] Sarvesh Vishwakarma and Anupam Agrawal. A survey on activity recognition and behavior understanding in video surveillance. *The Visual Computer*, 29(10):983–1009, 2013.
- [122] Eric-Jan Wagenmakers and Simon Farrell. Aic model selection using akaike weights. *Psychonomic bulletin & review*, 11(1):192–196, 2004.
- [123] Aiguo Wang, Guilin Chen, Jing Yang, Shenghui Zhao, and Chih-Yung Chang. A comparative study on human activity recognition using inertial sensors in a smartphone. *IEEE Sensors Journal*, 16(11):4566–4578, 2016.
- [124] Hualu Wang, Ajay Divakaran, Anthony Vetro, Shih-Fu Chang, and Huifang Sun. Survey of compressed-domain features used in audio-visual indexing and analysis. *Journal of Visual Communication and Image Representation*, 14(2):150–183, 2003.
- [125] Liang Wang, Tao Gu, Xianping Tao, and Jian Lu. Toward a wearable rfid system for real-time activity recognition using radio patterns. *IEEE Transactions on Mobile Computing*, 16(1):228–242, 2016.
- [126] Tracy Lee Westeyn. *Child’s play: activity recognition for monitoring children’s developmental progress with augmented toys*. PhD thesis, Georgia Institute of Technology, 2010.

- [127] Felix Wortmann and Kristina Flüchter. Internet of things. *Business & Information Systems Engineering*, 57(3):221–224, 2015.
- [128] Ziyou Xiong, Regunathan Radhakrishnan, Ajay Divakaran, and Thomas S Huang. Comparing mfcc and mpeg-7 audio features for feature extraction, maximum likelihood hmm and entropic prior hmm for sports audio classification. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03).*, volume 5, pages V–628. IEEE, 2003.
- [129] Yong Xu, Qiuqiang Kong, Wenwu Wang, and Mark D Plumbley. Large-scale weakly supervised audio classification using gated convolutional neural network. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 121–125. IEEE, 2018.
- [130] Zheng Yang, Jonathan Roth, and Rishee K Jain. Due-b: Data-driven urban energy benchmarking of buildings using recursive partitioning and stochastic frontier analysis. *Energy and Buildings*, 163:58–69, 2018.
- [131] Paul R Yarnold. Maximizing classification accuracy of cart® recursive partitioning tree models using optimal pruning. *Optimal Data Analysis*, 8:26–29, 2019.
- [132] LLC YouTube. Youtube. *Retrieved*, 27:2011, 2011.
- [133] Changsong Yu, Karim Said Barsim, Qiuqiang Kong, and Bin Yang. Multi-level attention model for weakly supervised audio classification. *arXiv preprint arXiv:1803.02353*, 2018.
- [134] Aras Yurtman and Billur Barshan. Human activity recognition using tag-based radio frequency localization. *Applied Artificial Intelligence*, 30(2):153–179, 2016.
- [135] Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, and Michele Zorzi. Internet of things for smart cities. *IEEE Internet of Things journal*, 1(1):22–32, 2014.



- [136] Laura A Zanella-Calzada, Carlos E Galván-Tejada, Nubia M Chávez-Lamas, Jesús Rivas-Gutierrez, Rafael Magallanes-Quintanar, Jose M Celaya-Padilla, Jorge I Galván-Tejada, and Hamurabi Gamboa-Rosales. Deep artificial neural networks for the diagnostic of caries using socioeconomic and nutritional features as determinants: Data from nhanes 2013–2014. *Bioengineering*, 5(2):47, 2018.
- [137] Tahmina Zebin, Patricia J Scully, and Krikor B Ozanyan. Human activity recognition with inertial sensors using a deep learning approach. In *2016 IEEE SENSORS*, pages 1–3. IEEE, 2016.
- [138] Ying Zhang, Guohui Tian, Senyan Zhang, and Cici Li. A knowledge-based approach for multiagent collaboration in smart home: From activity recognition to guidance service. *IEEE Transactions on Instrumentation and Measurement*, 2019.
- [139] Xi Zhu, Xiaofei Du, Mike Kerich, Falk W Lohoff, and Reza Momenan. Random forest based classification of alcohol dependence patients and healthy controls using resting state mri. *Neuroscience letters*, 676:27–33, 2018.
- [140] Muhammad Zubair, Kibong Song, and Changwoo Yoon. Human activity recognition using wearable accelerometer sensors. In *2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, pages 1–5. IEEE, 2016.
- [141] Jacek M Zurada. *Introduction to artificial neural systems*, volume 8. West St. Paul, 1992.