



**UNIVERSIDAD PEDAGOGICA
NACIONAL**

Educadora de educadores

SMART ROAD DETECTION: SISTEMA PARTICIPATIVO PARA LA
DETECCIÓN DE IRREGULARIDADES EN LA MALLA VIAL DE LA AVENIDA
CARRERA 68.

Camila Andrea Cubillos Rojas.

Brayan Eduardo Niño Pinzón.

Universidad Pedagógica Nacional de Colombia.

Facultad de ciencia y tecnología.

Departamento de tecnología.

Licenciatura en Electrónica.

Bogotá D.C.

2023

SMART ROAD DETECTION: SISTEMA PARTICIPATIVO PARA LA
DETECCIÓN DE IRREGULARIDADES EN LA MALLA VIAL DE LA AVENIDA
CARRERA 68.

Trabajo de grado para optar por el título de:

Licenciados en electrónica.

Autores:

Camila Andrea Cubillos Rojas

Brayan Eduardo Niño Pinzón

Director:

David Alejandro Martínez Vásquez, Ph.D

Universidad Pedagógica Nacional de Colombia.

Facultad de ciencia y tecnología.

Departamento de tecnología.

Licenciatura en Electrónica.

Bogotá D.C.

2023

Agradecimientos

Dedicamos este trabajo a nuestras familias, padres y hermanos por su apoyo incondicional en el transcurso de nuestras vidas, en especial por el paso por la Universidad Pedagógica Nacional.

Agradecemos al profesor David Alejandro Martínez Vásquez por su acompañamiento y asesoría en la realización del presente proyecto.

Tabla de contenido

1.	Presentación	2
1.1.	Justificación	2
1.2.	Planteamiento del problema.....	4
1.3.	Delimitación y alcance.....	5
1.4.	Objetivos.....	6
1.4.1.	Objetivo general	6
1.4.2.	Objetivos específicos.....	6
2.	Antecedentes.....	7
3.	Marco de referencia	11
3.1.	Marco teórico.....	11
3.1.1.	Machine Learning.....	11
3.1.2.	Clasificadores	11
3.1.2.1.	Asociación.....	11
3.1.2.2.	Clustering.....	12
3.1.2.3.	Modelo de regresión lineal.....	12
3.1.3.	Técnicas de aprendizaje computacional	13
3.1.3.1.	Redes neuronales.....	13
3.1.3.2.	Máquinas de soporte vectorial	14

3.1.3.3.	Árboles de decisión.....	14
3.1.3.3.1.	Características de los árboles de decisión.....	16
3.1.4.	Características usadas en clasificadores.....	16
3.1.4.1.	Amplitud de transformada de Fourier.....	17
3.1.4.2.	Media móvil.....	17
3.1.4.3.	Desviación estándar.....	17
3.1.5.	Internet de las cosas.....	18
3.1.6.	Sistema de Posicionamiento Global GPS.....	18
3.1.7.	Sensores Inerciales.....	18
3.1.7.1.	Acelerómetro.....	19
3.1.7.2.	Giroscopio.....	19
4.	Metodología.....	19
4.1.	Metodología.....	19
5.	Desarrollo del Proyecto.....	20
5.1.	Descripción General del Desarrollo de la Aplicación.....	21
5.1.1.	Smart road app.....	21
5.1.2.	Tratamiento de datos.....	31
5.1.2.1.	Caracterización de anomalías.....	31
5.1.2.1.1.	Reductores de velocidad vial.....	31

5.1.2.1.2.	Huecos o baches.....	32
5.1.2.1.3.	Vías Irregularidades	33
5.1.2.1.4.	Vías planas sin irregularidades.....	33
5.1.3.	Herramientas para análisis de datos.....	34
5.1.4.	Depuración de datos	36
5.1.5.	Análisis de datos.....	38
5.1.6.	Extracción de características	39
5.1.6.1.	Amplitud de transformada de Fourier vs Media móvil.....	39
5.1.6.2.	Amplitud de transformada de Fourier vs Desviación estándar.....	41
5.1.6.3.	Fase de transformada de Fourier vs Media móvil.....	42
5.1.6.4.	Media móvil vs Desviación estándar	43
5.1.7.	Clasificador de árbol de decisión.....	44
6.	Resultados.....	51
6.1.	Resultados.....	51
6.2.	Conclusiones.....	52
6.3.	Trabajos futuros	53
7.	Referencias.....	55

Tabla de ilustraciones

<i>Ilustración 1. Ruta de desarrollo SRD.</i>	6
<i>Ilustración 2. Gráfica del modelo de regresión lineal.</i>	13
<i>Ilustración 3. Ejemplo árbol de decisión.</i>	15
<i>Ilustración 4 Grafica de un árbol de decisión.</i>	16
<i>Ilustración 5. Metodología Top-Down de Smart Road Detection.</i>	20
<i>Ilustración 6. Primera pantalla de visualización Smart App (interfaz de usuario).</i>	21
<i>Ilustración 7 Lógica de bloques Screen 1</i>	22
<i>Ilustración 8. Pantalla 2 registro de usuario Smart App.</i>	23
<i>Ilustración 9 Almacenamiento de cuentas</i>	24
<i>Ilustración 10. Pantalla 4 inicio de sesión Smart App.</i>	25
<i>Ilustración 11 Lógica inicio de sesión</i>	26
<i>Ilustración 12 Capturando datos</i>	27
<i>Ilustración 13 Capturar de datos 1</i>	28
<i>Ilustración 14 Capturar datos 2</i>	29
<i>Ilustración 15 Jerarquía de datos</i>	30
<i>Ilustración 16 Almacenamiento de datos</i>	30
<i>Ilustración 17. Irregularidad policía acostado amarillo(a) y policía acostado blanco(b)</i>	32
<i>Ilustración 18. Irregularidad hueco peligroso.</i>	32
<i>Ilustración 19. Vía con irregularidad constante.</i>	33
<i>Ilustración 20. Vía en buen estado.</i>	33
<i>Ilustración 21. Captura de datos de la avenida carrera 63 de Bogotá.</i>	35
<i>Ilustración 22 Boxplot datos Avenida Carrera 68.</i>	37
<i>Ilustración 23 Grafica Boxplot Datos Filtrados en cada sensor.</i>	38

<i>Ilustración 24. Ejes de sensor acelerómetro de un celular ubicado en posición perpendicular(a) y posición del teléfono en un vehículo(b).</i>	39
<i>Ilustración 25 Grafica de dispersión amplitud de Fourier vs Media Móvil.</i>	40
<i>Ilustración 26 Grafica de dispersión amplitud de transformada de Fourier vs desviación estándar.</i>	41
<i>Ilustración 27 Grafica de dispersión fase de la transformada de Fourier vs Media móvil</i>	42
<i>Ilustración 28. Grafica de dispersión relación entre media móvil vs desviación estándar.</i>	43
<i>Ilustración 29 Árbol de decisión de 4 ramificaciones.</i>	46
<i>Ilustración 301 Nodos finales del árbol de decisión de 4 ramificaciones.</i>	48

Tabla de tablas

Tabla 1 *Características relevantes de los trabajos antecedentes a SRD* 10

Tabla 2 *Comportamiento de los ejes del sensor acelerómetro de acuerdo con la posición del dispositivo móvil.* 34

Resumen

El presente trabajo propone el desarrollo de un sistema participativo para detección de irregularidades de la malla vial en una de las avenidas más importantes de la ciudad de Bogotá, la avenida carrera 68, para este proyecto se implementó una metodología llamada Top-Down que permitió dividir en dos fases el desarrollo del proyecto, una enfocada en la creación del sistema y otra en la extracción de información de datos, en la primera fase se crea una aplicación para dispositivos móviles llamada Smart Road App, esta se encarga de tomar datos desde los sensores inerciales del teléfono, como lo son acelerómetro y giroscopio y determinar la ubicación de las irregularidades mediante GPS.

Estos datos son enviados a una base de datos que los almacenará para su posterior tratamiento. El estudio se enfocará principalmente en las irregularidades y obstáculos más comunes como lo son imperfecciones generadas por el desgaste de la vía y reductores de velocidad; con base en esos datos almacenados, se implanta la segunda fase, donde se creó un algoritmo de árbol de decisión con el cual se realiza la clasificación de las irregularidades, logrando una predicción del 92%. Sus características principales son la media móvil y la desviación estándar con una importancia de 13,4% y 86,5% respectivamente.

Introducción

Generar un ambiente óptimo y de buena calidad es un reto para las ciudades en la actualidad, uno de los factores que generan malestar en la sociedad, es todo lo relacionado con la movilidad y como ésta puede determinar el crecimiento económico y el bienestar de comunidades. En Bogotá, por ejemplo, la inconformidad de los ciudadanos es general, al tener que pasar por vías que están en mal estado, afectando así la vida útil de los automóviles y medios de transporte. Por lo anterior, esta investigación se enfoca en la detección de irregularidades en la malla vial en una de las avenidas más importantes de Bogotá, y los aspectos en los que se dará mayor relevancia son los tipos de irregularidades que más se encuentran en la avenida carrera 68.

Crear iniciativas tecnológicas que contribuyan con la movilidad en una de las ciudades con más niveles de congestión de tráfico y con vías en estado regular a malo, pueden contribuir en una mejor calidad de vida para sus ciudadanos. Smart Road Detection es un sistema participativo que busca generar información del estado de la avenida carrera 68, por medio de dispositivos móviles y los sensores inerciales que estos tienen. Para encontrar los tipos de irregularidades se hace uso de exploración estadística específicamente de algoritmos de clasificación, que permiten hacer análisis predictivos; En este proyecto se implementa el modelo de árboles de decisión, el cual logra clasificar tres tipos de irregularidades con una efectividad del 92% de predicción. Estos resultados pueden generar información que oriente a los entes encargados del estado de las vías y así generar planes de contingencia para las ciudades con ayuda de la participación ciudadana y de las tecnologías.

1. Presentación

1.1. Justificación

Para el año 2023, Colombia registra una población de 51,96 millones de personas y un total de 73,68 millones de conexiones móviles celulares, cifra que equivale al 141,8 por ciento de la población total, según el último informe de (We Are Social, 2023). Esto indica que en el país hay más líneas de telefonía celular que habitantes; El uso de estos dispositivos es indispensable para la cotidianidad de su población y hace viable la implementación de desarrollos tecnológicos que mejoren la gestión urbana y den respuestas a los retos que las ciudades actuales enfrentan.

Por otro lado en Bogotá, los ciudadanos gastan un promedio de 65 minutos para desplazarse de la vivienda al trabajo, mientras que la media mundial de este tiempo de desplazamiento es de 46 minutos, (INDRA, 2014) . Este resultado refleja el reto que tiene la ciudad para mejorar su movilidad, y pensar en iniciativas que ayuden por ejemplo a reducir el tiempo de desplazamiento, diagnosticar el estado o localización de las irregularidades de la malla vial con ayuda de la tecnología.

El estado de las vías en Bogotá según (Instituto de Desarrollo Urbano, 2021), se clasifican en dos, el Índice de Condición del Pavimento (PCI) que determina pavimentos con superficie flexible, rígida o articulada(adoquín) y el Índice de Condición de Vías Sin Superficie (URCI) que determina pavimentos de superficie de afirmado y tierra en la malla rural. Para la clasificación PCI se observa que un 22% de la malla se encuentra en estado pobre y grave, y un 9% en estado fallado y para URCI el estado es mucho más grave, ya que predomina el estado muy pobre y pobre con un 30% y 22% respectivamente. Estos

índices conllevan consecuencias que pueden acarrear daño en vehículos o accidentes viales. Es por lo anterior, que se buscó desarrollar un apoyo o alternativa tecnológica que permita recopilar datos del estado de las carreteras, con el fin de evitar siniestros viales, mitigar lesiones, pérdidas económicas e incluso reducir la tasa de mortalidad de los ciudadanos.

El uso de las tecnologías móviles para localizar dónde se encuentran las irregularidades sobre las vías, se realiza mediante el GPS (Global Position System por su nombre en inglés). Un ejemplo de estas tecnologías se dio inicialmente en la ciudad de Boston, donde se implementó un sistema móvil que detecta e informa las condiciones de la superficie de las carreteras, utilizando una red de taxis que están equipados con sensores inerciales, llamado Patrulla de baches (The Pothole Patrol por su nombre en inglés) (Eriksoon, Girod, Hull, Newton, Madden, Balakrishnan, 2008). Por otro lado, una investigación desarrollada por Allouch, Koubaa, Abbes, & Ammar (2017), demostró que el método de recopilación de datos basado en teléfonos inteligentes es muy útil porque elimina la necesidad de implementar sensores especiales en los vehículos y logra una precisión de 98,6% de detección.

De acuerdo con la experiencia de los trabajos anteriores y viendo los resultados positivos que tuvieron en las ciudades donde se desarrollaron estos proyectos, se considera que es viable implementar un sistema como SRD (Smart Road Detection), ya que le entrega a la ciudad de Bogotá un diagnóstico sobre el estado de la malla vial de la avenida carrera 68, y la clasificación de las irregularidades encontradas.

1.2. Planteamiento del problema

Las ciudades inteligentes (comúnmente conocidas como “Smart Cities” por su nombre en inglés) plantean mejorar los servicios públicos brindados y la conectividad en las ciudades, llevando a cabo procesos de planificación colaborativa y de participación ciudadana; como lo menciona el Banco Interamericano de Desarrollo (2016), una ciudad inteligente es aquella que coloca a las personas en el centro del desarrollo, haciendo uso de las Tecnologías de la Información y Comunicación (TIC). Esta iniciativa es atractiva para los ciudadanos, empresarios y trabajadores, ya que mejora los servicios y crea un ambiente de innovación que garantiza una mayor calidad de vida a largo plazo.

Asimismo, se viene trabajando la Internet de las Cosas (IoT: *Internet of things*, por sus siglas en inglés) que incluye todos los dispositivos y objetos cuyo estado puede ser alterado a través de Internet, con o sin participación de personas, (OCDE, 2018). Por consiguiente, se busca poder unir estos dos conceptos, ya que trabajan en la interconexión de los ciudadanos con todo lo que los rodea.

Para IoT existe la posibilidad del diseño e implementación de un sistema colaborativo y participativo, que haga uso de los dispositivos móviles o “Smartphone”, ya que, para enero del 2023, según el informe dado por *We Are Social* (2023), en Colombia hay más líneas de teléfonos móviles que habitantes. Puntualmente, en la ciudad de Bogotá D.C. en el año 2018 el 84,5% de los ciudadanos contaban con un teléfono inteligente (DANE, 2018). Está marcada masificación del teléfono celular y su continua conexión a internet permiten plantear sistemas de recolección y análisis de datos dentro de un contexto de ciudad inteligente tales como SRD: *Smart Road Detection* que contribuya a

la obtención de datos relacionados con las irregularidades de la malla vial en la ciudad, con el fin de conocer su estado actual y contribuir con la mejora de la calidad de vida de los ciudadanos.

1.3. Delimitación y alcance

El presente proyecto tiene como alcance la realización de un sistema participativo¹ en el que se hace uso de dispositivos móviles y de sus sensores inerciales como acelerómetro, giroscopio y el sistema de posicionamiento global GPS, para obtener datos de la avenida carrera 68 de Bogotá.

Los datos son capturados a partir de una aplicación móvil, desde la que se envían la información capturada a una base de datos, para luego realizar un preprocesamiento manual con el que se seleccionan, limpian y transforman con el fin de obtener la mayor información posible.

Para realizar la clasificación de las irregularidades se hace uso de la inteligencia artificial específicamente del *machine learning*, donde se hace uso de un algoritmo de clasificación para el análisis predictivo, el modelo implementado es el de árboles de decisión, con el que a partir de jerarquías de condicionales crea la clasificación de las irregularidades a partir de las características seleccionadas para la predicción. A

¹ Sistema participativo: conjunto de elementos que relaciona una aplicación móvil, conexión con base de datos y análisis de datos para clasificación de irregularidades, donde los usuarios toman parte activa de la captura de datos.

continuación, se observa una ilustración en la que se muestra la ruta de para el desarrollo de SRD.

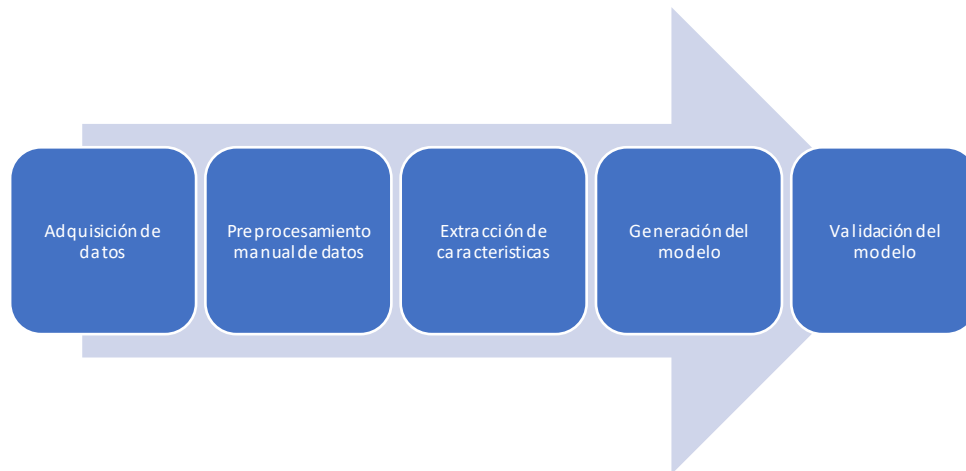


Ilustración 1. Ruta de desarrollo SRD.

1.4. Objetivos

1.4.1. Objetivo general

Desarrollar un sistema participativo para la detección de irregularidades en la malla vial de la avenida carrera 68 en la ciudad de Bogotá, haciendo uso de teléfonos inteligentes e Internet de las cosas.

1.4.2. Objetivos específicos

- Diseñar una aplicación para dispositivos móviles, que permita la recolección de datos de los sensores inerciales y de geolocalización.
- Identificar las diferentes irregularidades por medio de exploración estadística e inteligencia computacional.
- Clasificar los tipos de irregularidades y su respectiva ubicación.

2. Antecedentes

El trabajo desarrollado en la ciudad de Boston (Estados Unidos) (Eriksson , y otros, 2008), muestra la implementación de una aplicación móvil para informar sobre las condiciones de la superficie de las carreteras a través de una selección rigurosa de datos y el entrenamiento de un modelo predictivo a través de estos. Se demostró que detectar e informar las condiciones de la superficie de las carreteras, por medio de vehículos equipados con sensores de vibración y GPS, logra tener una detección de irregularidades acertada mayor al 90%.

Mediante el avance tecnológico que estamos viviendo, los proyectos que implementan tecnologías y participación ciudadana a partir de las IoT toman gran ventaja, ya que ahora contamos con dispositivos móviles conectados a internet que vienen equipados con sensores inerciales que facilitan la captura de los datos requeridos para este tipo de desarrollo, un ejemplo de esto es:

El proyecto desarrollado (Wang, Huo, Li, Wang, & Wang, 2018) toma relevancia ya que hace uso de teléfonos inteligentes y de los sensores que estos integran como el acelerómetro, ellos hacen uso del método MTS(Mahalanobis-Taguchi), con el que se logra clasificar en categorías las condiciones de la carretera, tales como la reducción de velocidad debida a la presencia de baches e irregularidades. Esta implementación tuvo una tasa de error aproximadamente de 5,72%, siendo un buen resultado para la detección. Por otro lado, este proyecto es uno de los primeros en explorar con los sensores inerciales de los dispositivos móviles, donde su experiencia contribuyo con la decisión de la elección del sensor del acelerómetro como sensor principal.

Otra aproximación en la que se usan teléfonos inteligentes es la propuesta de (Edwan, Sarsour, & Alatrash, 2019), en la cual se capturan datos como la posición, la velocidad y la aceleración, que permiten detectar protuberancias y ubicación, con ayuda de un algoritmo llamado Z-Threshold y la geolocalización, los datos son almacenados localmente en la aplicación y cuando este se conecta a internet las ubicaciones de los baches se transfieren a la base de datos.

Por otro lado, en India el departamento civil de ingeniería realizó una investigación donde se compara que detección es más efectiva entre vibración y visión. Ambos experimentos se realizaron en las mismas carreteras mediante levantamientos manuales, y pudieron identificar que la detección de baches por vibración arroja una un 80% de precisión frente a un 84% que muestra los datos de visión, el estudio revela que el análisis basado en vibración es suficiente para el propósito de monitoreo, sin embargo, el método basado en visión es más apropiado para un análisis más detallado. (Lekshmiathy, Samuel, & Velayudhan, 2020)

RoadSense es una aplicación para Android que recoge datos en tiempo real, prediciendo automáticamente la calidad de la carretera, este se basa en triaxiales como acelerómetro y giroscopio. En la investigación, se usa un clasificador de árbol de decisión y lo aplican a los datos de entrenamiento para clasificar el estado de diferentes tramos viales. Ellos crearon un modelo experimental que muestra una precisión constante de 98,6%, con este enfoque visualizan la calidad de las carreteras. Este proyecto supera la precisión de la implementación del modelo de SRD debido a que la toma de datos fueron

de forma experimental en diferentes medios de transporte y a diferentes velocidades. (Allouch, Koubaa, Abbas, & Ammar, 2016)

En (Leizerovych, Sindenko, Kondratenko, & Kondratenko, 2020) , los autores analizan la condición de las carreteras usando una aplicación que recopila automáticamente los datos de acelerómetro, giroscopio y GPS, los datos recolectados pasan a ser analizados por redes neuronales y luego a los datos de entrenamiento para clasificarlos los tramos de la carretera y así construir el modelo, todos los datos se enviaron a la nube IoT de almacenamiento para su posterior procesamiento.

El siguiente estudio (Sattar, Li, & Chapman, 2021), propone un método que usa algoritmos de aprendizaje automático no supervisado para detectar anomalías en las carreteras, usa datos de la vibración del acelerómetro con un enfoque de agrupamiento de k-medias para eliminar los incidentes causados por giros y aceleraciones. El giroscopio reorienta los datos del sensor acelerómetro, por lo tanto, le da libertad de uso a los usuarios para ubicar su teléfono en cualquier orientación. Sin embargo para SRD la posición del dispositivo móvil es fija en el medio de transporte para disminuir ruido y datos erróneos en las señales.

A continuación, en la tabla 1, se presenta un resumen de los trabajos de investigación o desarrollo que sirvieron como base para el desarrollo SRD.

Tabla 1*Características relevantes de los trabajos antecedentes a SRD*

Autor	Año	Datos utilizados para detección y clasificación de anomalías	Algoritmo	Actuación Métrica
Wu et al.	2020	Sensor Acelerómetro	Random Forest	95.7%
Bustamante et al.	2022	Sensor Acelerómetro	K.Nearest Neighbor	95.55%
Ferjani y Ali.	2022	Sensor Acelerómetro	Decision tree	94%
Tiwari et al.	2020	Sensor Acelerómetro y GPS	Convulational Neural Network	98.5%
Dimitriou et.al	2019	Sensor Acelerómetro	Red Neuronal artificial (ANN)	90%
Jagtap y Bhoir.	2020	Sensor Acelerómetro y Giroscopio	Red neuronal	94.78%
Sagar y Annapurna.	2018	Sensor Acelerómetro y Giroscopio	Clasificador de árbol de decisión	97,6%
Kulkarni et al.	2014	Sensor Acelerómetro	Red neuronal aprendizaje (Encog)	95%
Eriksson et al.	2008	Sensores de vibración y GPS	Aprendizaje automático	90%
Wang et al.	2018	Sensor Acelerómetro y Giroscopio	Mahalanobis-Taguchi	95%

3. Marco de referencia

3.1. Marco teórico

En el siguiente apartado se dará a conocer la base teórica, conceptual y tecnológica, seleccionada para dar viabilidad al proyecto.

3.1.1. Machine Learning

El Machine Learning o el aprendizaje de máquina es un campo de investigación entre la estadística, la inteligencia artificial y la informática. Consiste en tratar de extraer conocimientos de los datos y aprender de ellos. Por ejemplo, una aplicación de este aprendizaje son las recomendaciones automáticas sobre películas, libros, música, compra de productos, o diversos problemas científicos de descubrimiento, inmersión o de hallazgo. Esta disciplina tiene la capacidad de identificar patrones en una cantidad de datos y con ellos elaborar predicciones, que permitan dar información relevante de un tema. (C.Müller & Guido, 2016)

3.1.2. Clasificadores

Los clasificadores son técnicas o modelos que permiten hacer tareas de caracterización o predicción. A continuación se nombran algunas técnicas para el modelamiento de datos y/o el análisis predictivo , se describen de la siguiente manera:

3.1.2.1. Asociación

El modelo de asociación se basa en un conjunto de datos que tienen identificadores para casos individuales y para los elementos que contienen los casos. Se componen de una serie de conjuntos de elementos y reglas que describen como estos elementos se agrupan

dentro de los casos. Las reglas que el algoritmo identifica pueden utilizarse para predecir, por ejemplo, las probables compras de un cliente futuro, basándose en elementos existentes en el carro de compra actual del cliente. (Microsoft, Algoritmo de asociación de Microsoft, 2021)

3.1.2.2. Clustering

Clustering es un algoritmo de segmentación que itera en los casos de un conjunto de datos para agruparlos en nubes o conjuntos de datos que contengan características similares. Estas agrupaciones son útiles para la exploración de datos, la identificación de anomalías en los datos y la creación de predicciones. El objetivo es ordenar las observaciones de los grupos tales que el grado de asociación natural es alto entre los miembros del mismo grupo y bajo miembros de diferentes grupos. (Microsoft, Algoritmo de clústeres de Microsoft, 2021)

3.1.2.3. Modelo de regresión lineal

El modelo de regresión lineal analiza la relación existente entre la variable dependiente o de respuesta y un conjunto de variables independientes o predictoras. Esta relación se expresa como una recta que predice la variable de respuesta como una función lineal de los parámetros. Esto con el fin de seleccionar parámetros que minimicen la suma de los errores al cuadrado.

En la imagen a continuación, se ve un ejemplo del uso de regresión lineal simple. Para el ejemplo se toma una variable dependiente y otra independiente, esto con el fin de generar una función lineal que predice los valores de la variable dependiente (Timón, 2017)

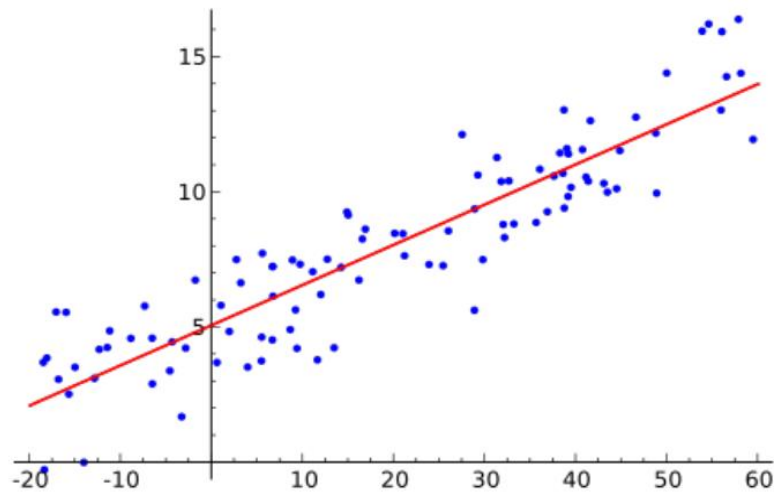


Ilustración 2. Gráfica del modelo de regresión lineal.

3.1.3. Técnicas de aprendizaje computacional

El aprendizaje computacional desarrolla técnicas que permite a las computadoras “aprender” de la cuales se incluyen una serie de métodos estadísticos avanzados para la regresión y la clasificación. Tienen aplicación en una amplia variedad de campos. (Timón, 2017)

3.1.3.1. Redes neuronales

Una red neuronal es un método de la inteligencia artificial que enseña a las computadoras a procesar datos de una forma inspirada en como lo hace el cerebro humano. Este hace un proceso de *machine learning* llamado aprendizaje profundo, que utiliza las neuronas interconectadas en una estructura de capas capaces de aprender de sus errores y mejorar continuamente. (C.Müller & Guido, 2016)

3.1.3.2. Máquinas de soporte vectorial

Las máquinas de soporte vectorial (SVM) son un algoritmo de aprendizaje supervisado que se utiliza en problemas de clasificación y regresión, el objetivo es encontrar un hiperplano que separe dos clases diferentes de un conjunto de datos. El algoritmo solo puede encontrar el hiperplano en problemas que permitan la separación lineal. Los tipos de SVM más comunes son: Lineal, Polinomial, Sigmoides. (Timón, 2017)

3.1.3.3. Árboles de decisión

Los árboles de decisión son modelos altamente utilizados para las tareas de clasificación y regresión, principalmente aprenden de una jerarquía de condiciones (if/else) que lleva a tomar una decisión. Cada vez que el árbol avanza en su toma de decisiones va eliminando opciones basadas en la condición anterior. En la siguiente imagen se puede ver gráficamente, como estaría compuesto el árbol de decisiones (Müller & Guido, 2016).

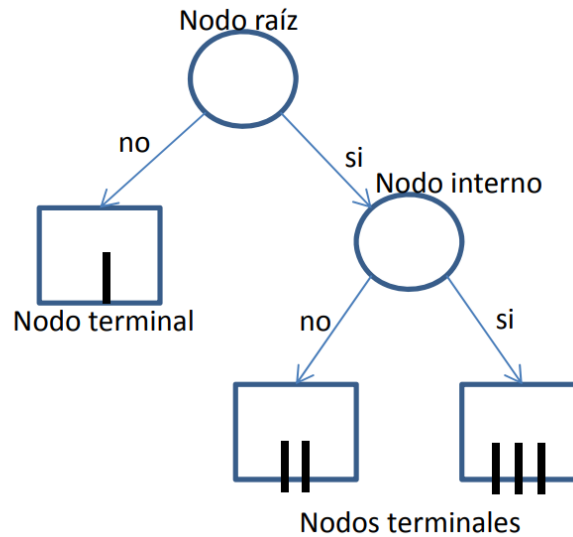


Ilustración 3. Ejemplo árbol de decisión.

A continuación, se toma el ejemplo del autor que plantea lo siguiente:

“Imagina que quieres distinguir entre los siguientes cuatro animales: osos, halcones, pingüinos y delfines. Su objetivo es llegar a la respuesta correcta haciendo la menor cantidad posible de preguntas (if/else). Puede comenzar preguntando si el animal tiene plumas, una pregunta que reduce sus posibles animales a solo dos. Si la respuesta es “sí”, puedes hacer otra pregunta que podría ayudarte a distinguir entre halcones y pingüinos. Por ejemplo, podría preguntar si el animal puede volar. Si el animal no tiene plumas, sus posibles opciones de animales son delfines y osos, y deberá hacer una pregunta para distinguir entre estos dos animales, por ejemplo, preguntar si el animal tiene aletas.”

En la siguiente imagen el autor ilustra de manera gráfica como se representa el árbol de decisión.

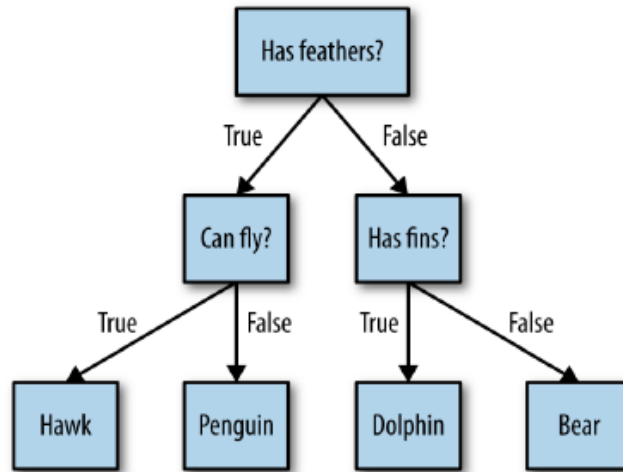


Ilustración 4 Grafica de un árbol de decisión.

3.1.3.3.1. Características de los árboles de decisión

Las características son entradas utilizadas para la predicción o clasificación, estas características son un conjunto de datos que calculan el número de muestras que le llegan y cuanto mayor sea el valor más importante será la característica, y así el algoritmo podrá realizar las tomas de decisiones para la predicción². (Molnar, 2021)

3.1.4. Características usadas en clasificadores

Las características de los datos permiten realizar los procesos de clasificación de nueva información. Por ejemplo, la característica **plumas** permite identificar si un animal es ave o no en el proceso de un árbol de decisión. En nuestro proyecto, hemos identificado las siguientes características como relevantes según los resultados que logramos obtener:

² La predicción es el valor que el modelo de aprendizaje automático pronostica, en función de las características dadas.

3.1.4.1. Amplitud de transformada de Fourier

La transformada de Fourier es una función que transforma una señal muestreada en el dominio tiempo al dominio de la frecuencia, esta función puede dar características importantes de una señal y es por lo que se quiere revisar su comportamiento a partir de la magnitud de la amplitud de la transformada de Fourier como característica para el clasificador, la transformada se define por la siguiente ecuación. (kreyszig, 2003)

$$y_{k+1} = \sum_{j=0}^{n-1} \omega^{jk} x_{j+1}$$

La transformada de Fourier se define para un vector x con puntos n muestreados uniformemente por la ecuación anterior, donde $\omega = e^{-2\pi i/n}$ es una de las raíces complejas de unidad n donde i es la unidad imaginaria. Para x e y , los índices j y k oscilan entre 0 y $n-1$.

3.1.4.2. Media móvil

La media móvil es un dato que calcula la media aritmética de un conjunto de datos ya sea en una serie de tiempo o a partir de un cierto número de periodos anteriores a una serie de tiempo. (Triola, 2006)

3.1.4.3. Desviación estándar

La desviación estándar es una medida que se utiliza para calcular la variación o dispersión de los datos que difieren de la media aritmética, la desviación indica que tan cerca están los datos de la media móvil. (Triola, 2006)

3.1.5. Internet de las cosas

Según la OCDE (Organización para la Cooperación y el Desarrollo Económico, 2018), no hay una definición clara de lo que es IoT, sin embargo, algunas organizaciones internacionales como la Organización Internacional de Normalización (ISO) y el Organismo de Reguladores Europeos de Comunicaciones Electrónicas (ORECE), están trabajando en su definición. Se define el IoT como “sensores y actuadores conectados mediante redes a sistemas informáticos. Estos sistemas pueden supervisar o gestionar el estado y las acciones de los objetos y máquinas conectados. Los sensores conectados también pueden monitorear el mundo natural, las personas y los animales” (McKinsey, 2015, citado por OCDE, 2018, p 8 19).

3.1.6. Sistema de Posicionamiento Global GPS

El sistema de posicionamiento global (GPS) es un sistema basado en el espacio, que proporciona información precisa y continua de la posición y la velocidad tridimensional de sus usuarios, este se compone de satélites que orbitan alrededor de la Tierra, las estaciones terrestres de seguimiento y control y los receptores del GPS. (Kaplan & Hegarty, 2006). De esta manera, se quiere usar este servicio para hallar la ubicación de las irregularidades sobre la carretera.

3.1.7. Sensores Inerciales

Dado el desarrollo tecnológico, existe la posibilidad de medir diferentes magnitudes físicas, o alteraciones en el entorno. Según (Fericean, 2019), los sensores son instrumentos de mediciones físicas que son transformadas a señales eléctricas. En el caso de los sensores inerciales estos miden la aceleración lineal con que se mueve el sensor,

los giroscopios la velocidad angular y el magnetómetro da información acerca del norte magnético, SRD obtendrá datos de los siguientes sensores:

3.1.7.1. Acelerómetro

El acelerómetro es un sensor que permite conocer la aceleración que se produce al realizar un movimiento a lo largo del eje que se esté trabajando (Espin, 2010). El uso de este sensor para SRD es fundamental, ya que aquí es donde se ven reflejado los cambios producidos al impactar con una irregularidad vial.

3.1.7.2. Giroscopio

El giroscopio es un dispositivo que posibilita conocer como varía un ángulo en el tiempo mientras se encuentra rotando (velocidad angular). (Espin, 2010). Este sensor, le brinda a SRD la clasificación de las irregularidades, ya que puede describir el ángulo en el que se encuentra el vehículo, y así determinar qué tipo de irregularidad es.

4. Metodología

4.1. Metodología

Para la elaboración de este trabajo se adoptó una metodología basada en el desarrollo de software, por consiguiente, se elige la metodología Top-Down que permite dividir el proyecto por conjuntos que puedan ser manejados y sencillos de resolver (Restrepo, 2009). Para SRD los problemas y subproblemas son los siguientes:

- Desarrollo del sistema: Aquí se diseñará la aplicación móvil y la implementación de la base de datos en **Firestore**.

- Adquisición de datos: Funcionamiento de la aplicación y posteriormente el análisis de los datos recolectados.
- Clasificador de irregularidades: Una vez los datos son recolectados y procesados, se implementará una técnica de predicción de datos.

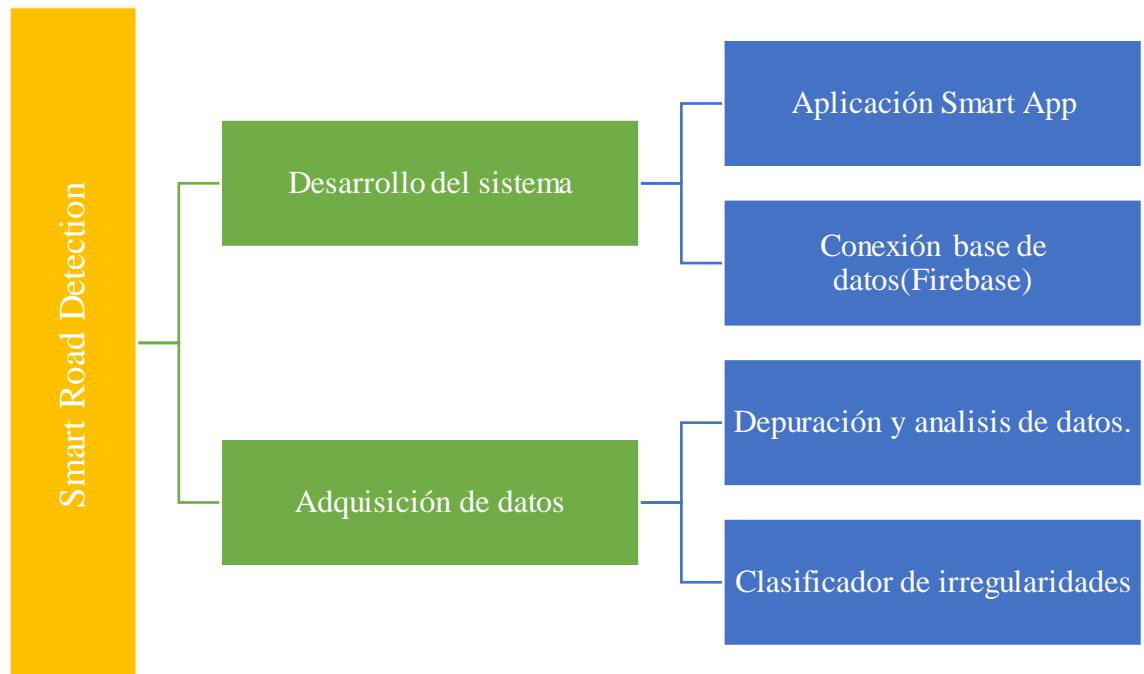


Ilustración 5. Metodología Top-Down de Smart Road Detection.

5. Desarrollo del Proyecto

En el siguiente apartado del documento se muestran los aspectos más relevantes que se abordaron durante la realización del proyecto.

5.1. Descripción General del Desarrollo de la Aplicación

5.1.1. *Smart road app*

Para SRD se requería una aplicación que capturara los datos de los sensores inerciales como: acelerómetro, giroscopio, GPS y la velocidad a la que iba el usuario, para esto se hizo uso de dispositivos móviles o Smart Phone, mediante una APP desarrollada en un software creado por Google Labs, denominada App inventor, la cual es un entorno que permite programar por bloques y que permite crear aplicaciones de gran alcance. El nombre que se le dio a la aplicación creada es Smart App. Esta cuenta con los siguientes 4 escenarios:

Primer escenario: La pantalla inicial con la que se encontrará un usuario será la pantalla de decisión, en la cual tendrá que decidir si inicia sesión o si se desea registrar.

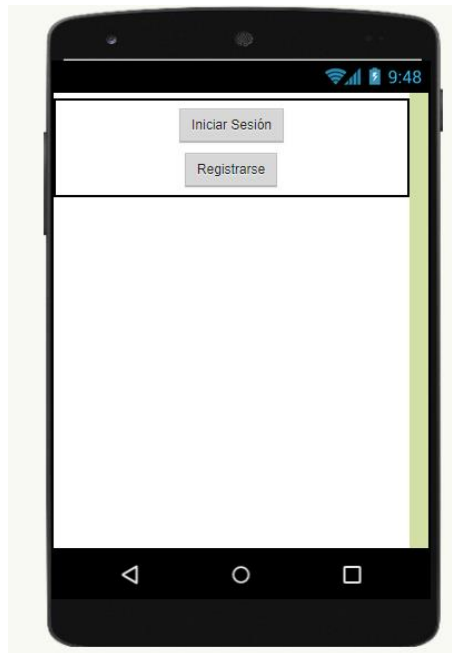


Ilustración 6. Primera pantalla de visualización Smart App (interfaz de usuario).

Screen1: Esta pantalla es una interfaz de usuario, y su programación en bloques es la siguiente:



Ilustración 7 Lógica de bloques Screen 1

De la ilustración podemos aclarar que la app tiene una lógica de desarrollo sencilla, ya que está condicionada por la decisión del usuario, y el código responde a la necesidad. Para iniciar sesión, se redirigirá a la pantalla de acceso. Si el usuario decide crear una cuenta nueva será redireccionado a la creación de la cuenta.

Luego de esto tendremos dos escenarios posibles, veamos primero al registro de un nuevo usuario.

Screen2: Aquí el usuario encontrará los cuadros de texto para ingresar un nombre deseado y una contraseña, la cual le permitirá entrar las veces que desee a la aplicación, ya que se guardará su usuario en la base de datos para que sea recordado.

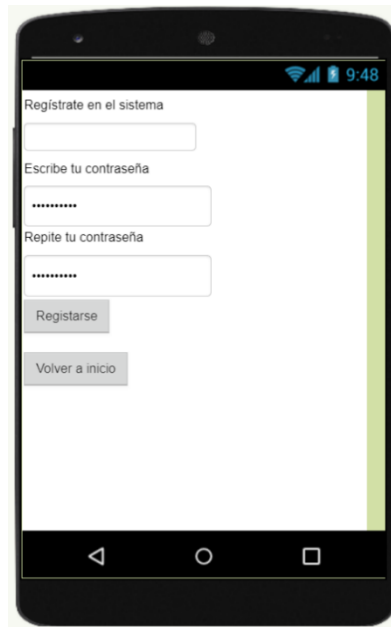
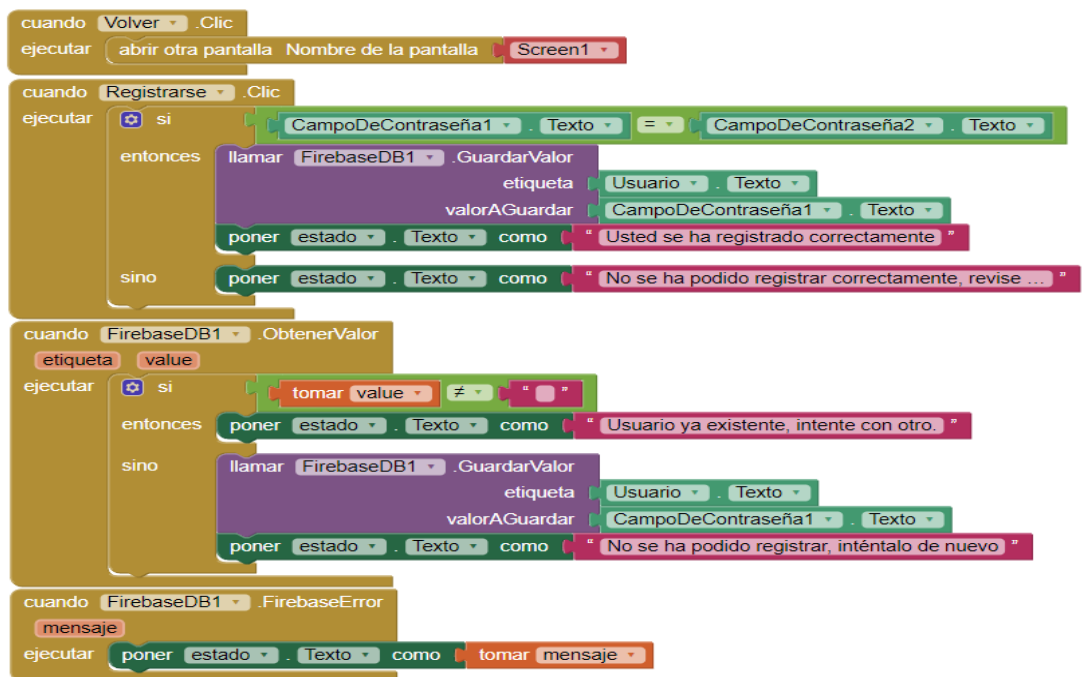


Ilustración 8. Pantalla 2 registro de usuario Smart App.

Su lógica de funcionamiento es la siguiente:



La codificación de la pantalla 2 tiene varios escenarios, el primero, un botón de volver al inicio, se asignó por posibles errores en la toma de decisiones en el Screen 1. En el segundo escenario, la creación del usuario, se habilita un campo de texto para ingresar el nombre deseado y dos cuadros de texto para la contraseña y se respectiva confirmación, esto para que el usuario tenga pleno conocimiento de su contraseña. La lógica diseñada está pensada para retroalimentar al usuario, debido que pueden existir dos posibles mensajes. Primer escenario: el nombre de usuario no existía y la contraseña coincide en los dos campos asignados para ella. Se retorna un mensaje de *“Usted se ha registrado correctamente”*. Segundo escenario: debido a que tenemos la base de datos de los usuarios registrados exitosamente, existe la posibilidad de que dos usuarios deseen el mismo nombre y esto arrojará error, ya que podremos evaluar si el usuario ya existe en la base de datos.

A continuación, se podrá ver el almacenamiento de los usuarios creados exitosamente en la base de datos:

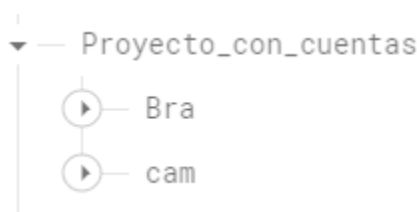


Ilustración 9 Almacenamiento de cuentas

De la ilustración anterior podemos ver dos cuentas de usuario creadas en el nodo “Proyecto_con_cuentas” en la base de datos.

Ya hemos visto los escenarios preliminares para poder lograr un inicio de sesión.

Screen3: Es la pantalla de inicio de sesión. Se verifica si el usuario ya está creado y que la contraseña coincide. Para esto se dejaron dos campos de texto y dos botones (para iniciar sesión y volver a inicio).

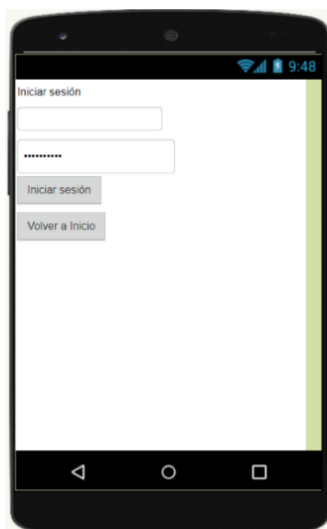


Ilustración 10. Pantalla 4 inicio de sesión Smart App.

El código implementado para esta pantalla es el mostrado en la siguiente ilustración. Se pueden ver las condiciones que se tiene para realizar el inicio y funcionamiento de la aplicación.

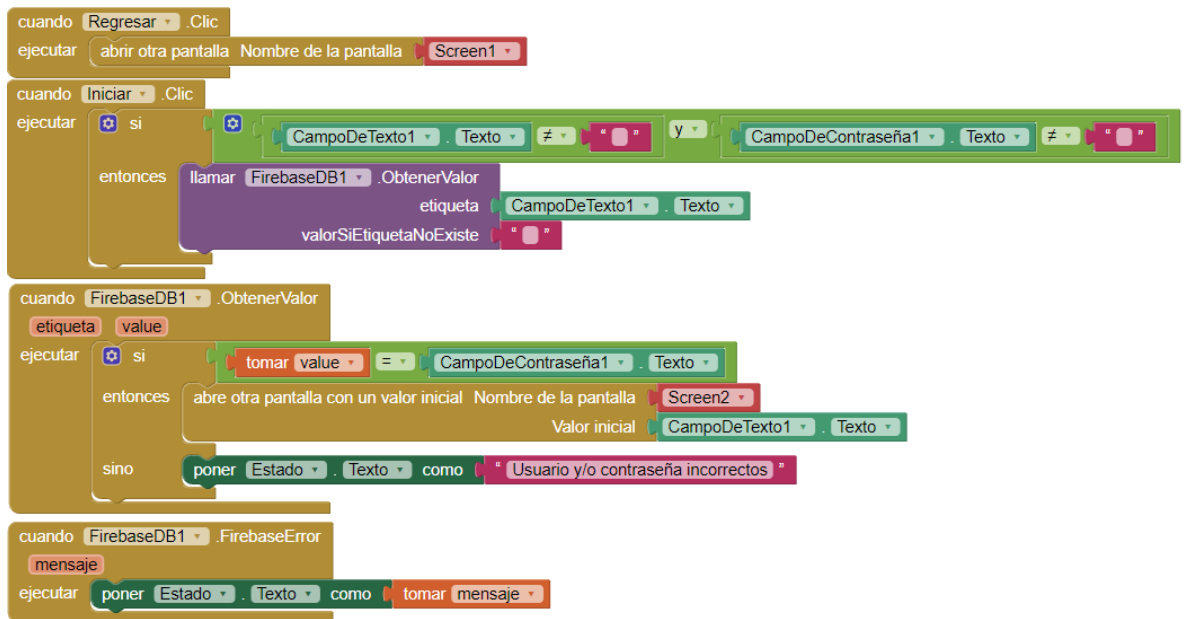


Ilustración 11 Lógica inicio de sesión

Lo primero que validamos es que no se permita el acceso cuando los campos de texto están vacíos. Una vez se valida que tengamos datos en estos dos campos, se realiza el llamado a la base de datos para verificar si el nombre de usuario y la contraseña coinciden. Si estos dos campos coinciden, se dará acceso al funcionamiento de Smart Road App, de lo contrario, se retornará un mensaje de “*Usuario y/o contraseña incorrectos*” y el usuario se mantendrá en la Screen actual para que pueda reintentar su acceso.

Por último, tenemos el funcionamiento de Smart Road App en el cual capturamos los datos mencionados anteriormente para poder realizar el estudio deseado.



Ilustración 12 Capturando datos

Screen 4: El usuario llegará a esta pantalla una vez cree una cuenta y realice el proceso de inicio de sesión. Estas son las únicas acciones que debe realizar el usuario para participar de este proyecto. Esta pantalla final cuenta con un campo de texto que retorna un saludo de bienvenida al usuario que se encuentra activo. Cuenta con otros tres campos de texto, los cuales son asignados para los sensores de GPS, Acelerómetro y Giroscopio. Finalmente, un campo de texto el cual expone la velocidad a la que va el usuario, esta velocidad está dada en m/s.

Su respectiva lógica será plasmada en las ilustraciones 13 y 14.

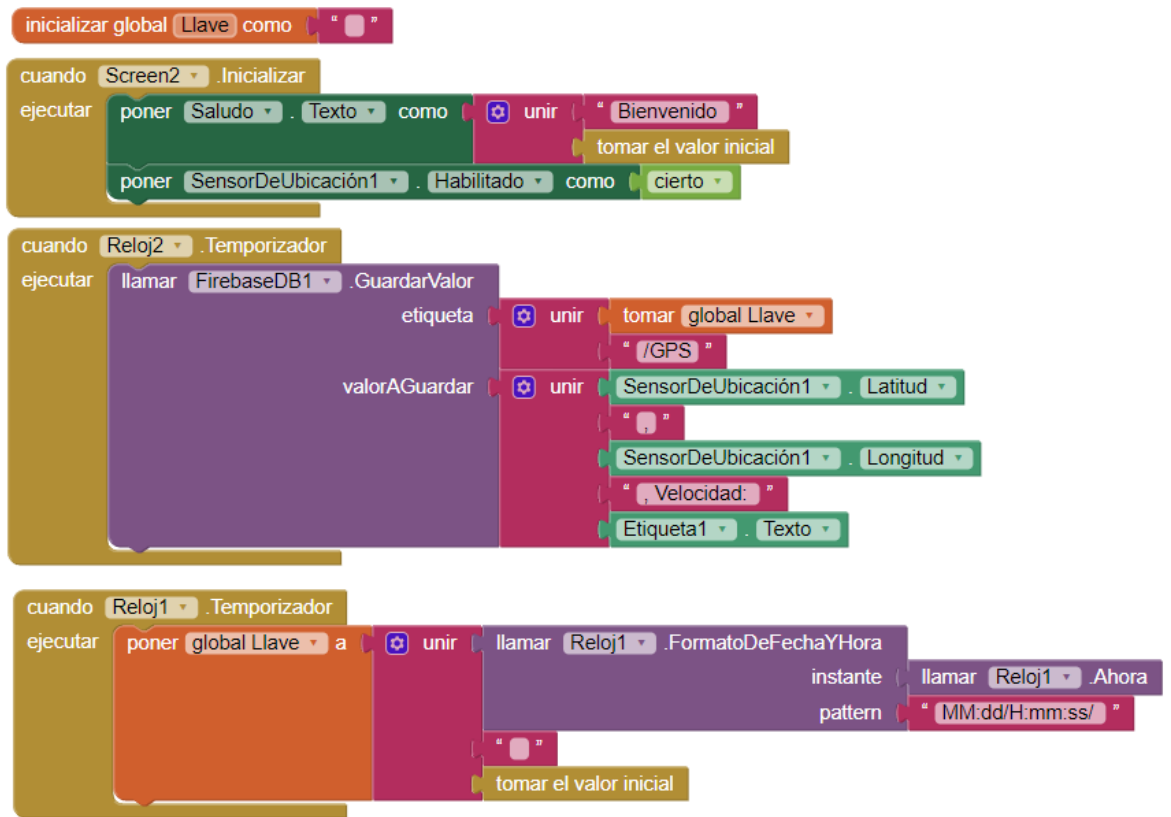


Ilustración 13 Capturar de datos 1

Inicialmente se programa el saludo para el usuario y solicita activar el sensor GPS para capturar los datos de Latitud y Longitud para tener la posición actual y la velocidad. El último bloque está programado para la configuración del reloj para que todos los sensores que requerimos sean sincronizados y se realice un envío exitoso

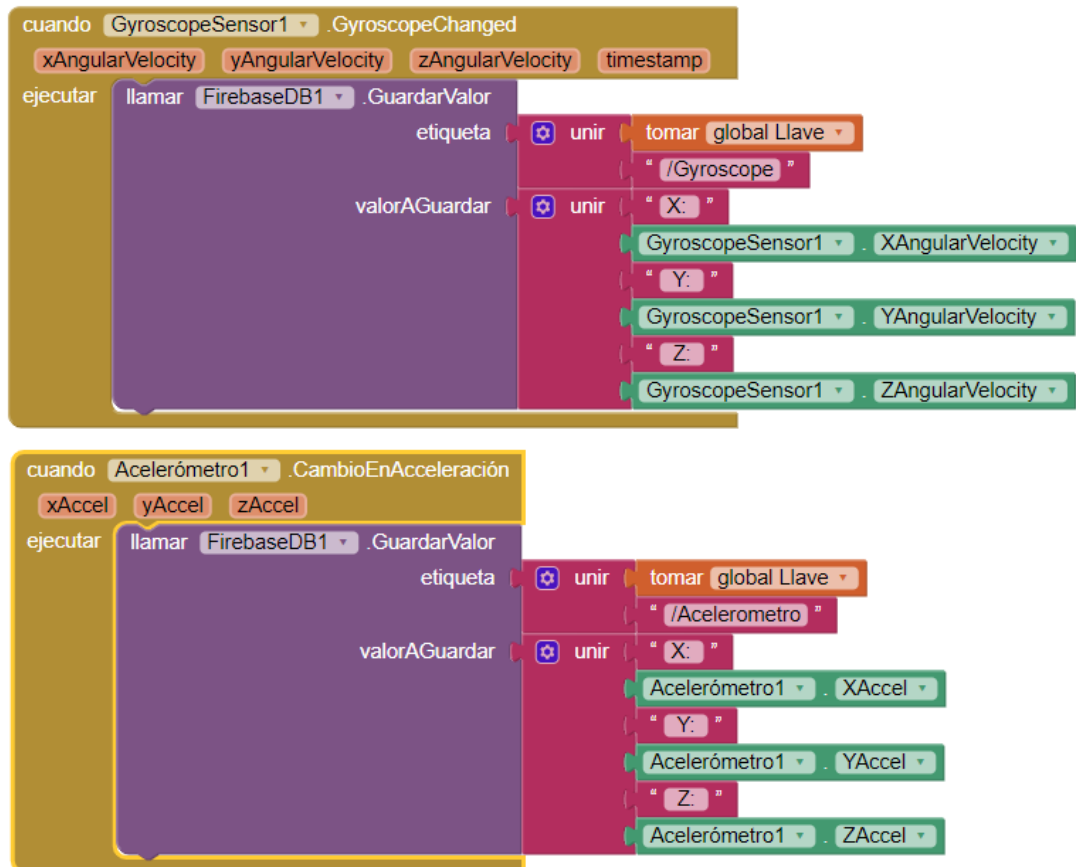


Ilustración 14 Capturar datos 2

En la ilustración 14 que pertenece a la Screen 4, se solicita activar los sensores de giroscopio y acelerómetro y capturamos sus valores en las tres coordenadas X, Y, Z que son necesarias para realizar el análisis. Estos datos capturados son enviados a Firebase con el siguiente orden jerárquico:



Ilustración 15 Jerarquía de datos

Ilustración 16 Almacenamiento de datos

De la ilustración 15 podemos ver el orden jerárquico que toman los datos. El primer nodo es la fecha 03:20 (20 de marzo). El siguiente nodo corresponde a la hora (formato 24H). Los datos fueron tomados entre las 12:48 en adelante. Se evidencia que se

realiza un envío de datos cada segundo, con el fin de tener una precisión más alta en cuando a los datos que recibimos. Tenemos un tercer nodo con el usuario que estaba activo en ese momento capturando datos, para el ejemplo tenemos al usuario “Bra”. Finalmente, tenemos los nodos de los sensores que son necesarios para el estudio, acelerómetro y giroscopio con sus tres coordenadas correspondientes y el sensor de GPS con la posición y la velocidad que mantenía el usuario en ese segundo.

5.1.2. Tratamiento de datos

5.1.2.1. Caracterización de anomalías

Para la elaboración de este proyecto se toman como objeto de estudio las siguientes anomalías viales: los reductores de velocidad, los huecos o baches, las vías con constantes irregularidades y la vía plana.

5.1.2.1.1. Reductores de velocidad vial.

Los reductores de velocidad, son elementos que se encuentran en la superficie de las carreteras y tiene como objetivo hacer que los conductores vayan más lento. En Colombia, existen dos materiales de producción: cemento y plástico. Estos reductores de velocidad, generalmente están ubicadas en las zonas escolares o de alto flujo peatonal.



(a)



(b)

Ilustración 17. Irregularidad policía acostado amarillo(a) y policía acostado blanco(b)

5.1.2.1.2. Huecos o baches.

Los huecos o baches son hoyos o desigualdades que se encuentran en la superficie de la carretera, y que se crean a partir de grietas en el pavimento de las avenidas donde se van expandiendo debido al tráfico vehicular y las condiciones asfálticas, un ejemplo de esta irregularidad se puede observar en la ilustración 18.



Ilustración 18. Irregularidad hueco peligroso.

5.1.2.1.3. Vías Irregularidades

Las vías con irregularidades se presentan debido al deterioro de las carreteras. Se estima que son generadas de manera natural por el alto flujo de vehículos, la falta de mantenimiento vial y demás factores externos. Se catalogaron tres tipos de irregularidades viales: Alta, Media y baja. A continuación, se muestra una imagen de este tipo de irregularidad.



Ilustración 19. Vía con irregularidad constante.

5.1.2.1.4. Vías planas sin irregularidades

Las vías planas y sin irregularidades, son vías en las cuales el estado del asfalto se encuentra en buenas condiciones, sin deterioro en el pavimento y sin desigualdades, así como se puede observar en la siguiente imagen.








Ilustración 20. Vía en buen estado.

5.1.3. Herramientas para análisis de datos

El comportamiento de los sensores inerciales es fundamental para el abordaje del problema planteado. Es por eso que se debe identificar el comportamiento de los sensores dependiendo de las diferentes posiciones que podría tomar el dispositivo móvil dentro de un vehículo, y así poder determinar cuál es la posición óptima para la recolección de datos. En ese sentido, realizamos diferentes pruebas con los dispositivos móviles y obtuvimos la siguiente tabla:

Tabla 2

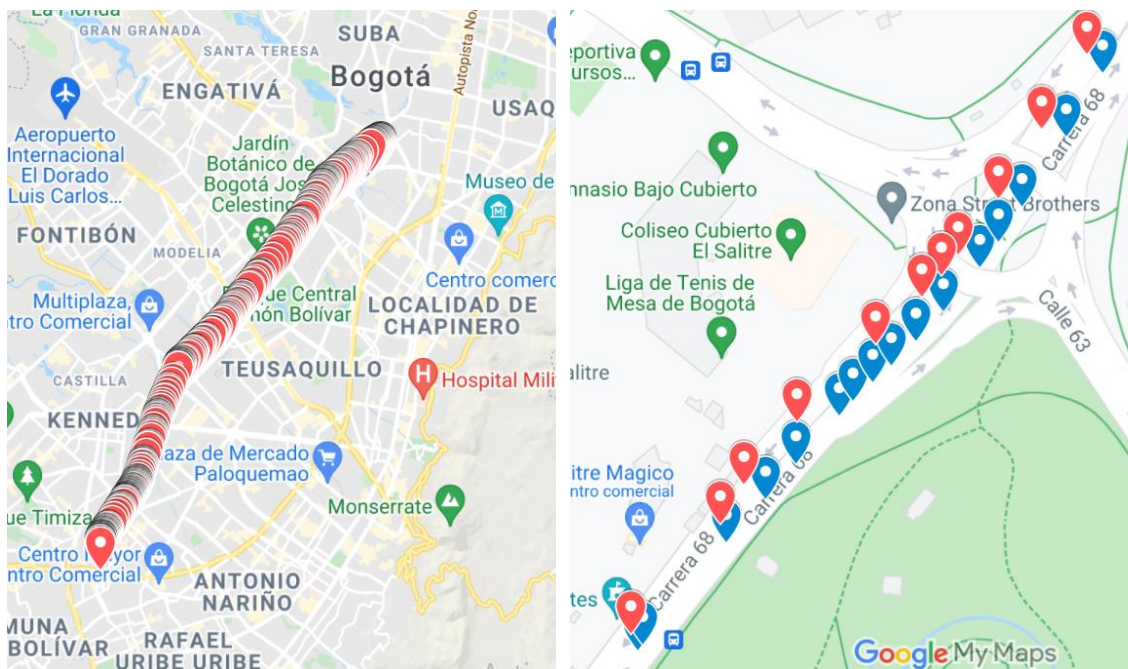
Comportamiento de los ejes del sensor acelerómetro de acuerdo con la posición del dispositivo móvil.

Posición	Explicación	Posición del celular	Ejes Acelerómetro
Posición natural	El celular va en el pantalón de la persona.		
Posición perpendicular	El celular está en posición perpendicular al carro.		
Posición arriba	El celular esta con pantalla hacia arriba.		
Posición abajo	El celular está con pantalla hacia abajo.		

Luego de analizar los datos que nos arrojaba cada uno de los sensores en las diversas posiciones, se decidió trabajar con la posición perpendicular al vehículo, ya que

en esta posición el comportamiento y la afectación del (eje y) del acelerómetro nos brindaba más información en cuanto a los fenómenos que sufriría el dispositivo móvil dentro del vehículo.

Una vez fijamos una posición, vamos a geolocalizar los datos ya obtenidos del sensor GPS para ubicarlos en el mapa de Bogotá, especialmente los datos obtenidos de la avenida carrera 68 que comprende localidades como Kennedy, Puente Aranda, Fontibón, Teusaquillo, Engativá, Suba y Barrios Unidos, partiendo desde el barrio Venecia en la Calle 51 sur con Carrera 53, y concluyendo su recorrido en el barrio la Floresta de la localidad de Suba, donde se convierte en la Avenida España o Avenida Calle 100.



(a)

(b)

Ilustración 21. Captura de datos de la avenida carrera 63 de Bogotá.

En la ilustración anterior, se observan algunos datos tomados de la avenida carrera 68, en cada uno de los sentidos de color azul para sentido Sur-Norte y de color rojo para sentido Norte-Sur.

5.1.4. Depuración de datos

Con el conjunto de datos obtenidos desde Smart Road App, se inicia el tratamiento de estos, primero, se descargan los datos de Fire base en un archivo de formato json, y con la intención de darles una mejor visualización, se cargan a Excel. Aquí se inicia una limpieza de datos, eliminando todos los datos nulos (aquellos que llegan en cero o null), en segunda medida se corrigen los errores estructurales que se genera en la transferencia de archivos, pues traen una nomenclatura extraña, que pueden causar inconsistencias en el análisis, es por eso que son corregidas manualmente. Por último, se realiza una validación y control de calidad de los datos, con el fin de prever datos contaminados que puedan llegar afectar el análisis. A continuación, se muestra una visualización del comportamiento de los datos sin filtrar en el software de Matlab.

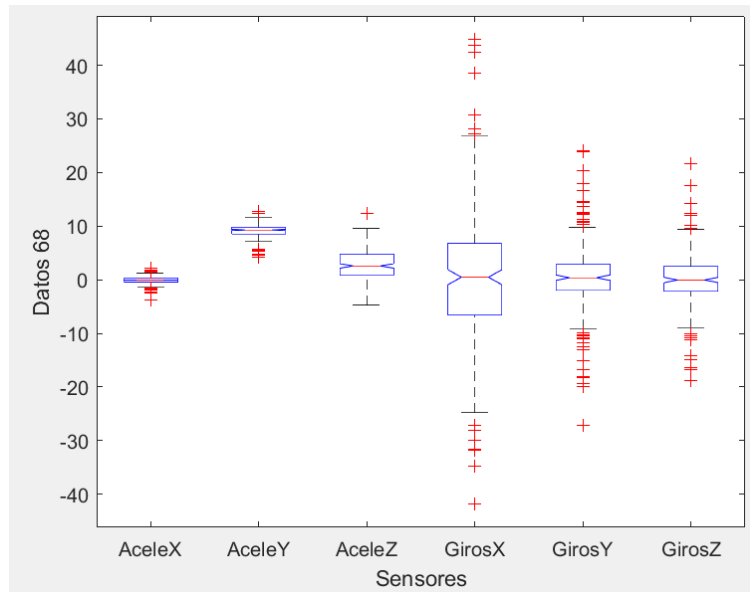


Ilustración 22 Boxplot datos Avenida Carrera 68.

En la ilustración anterior, se evidencia el comportamiento de los datos sin filtrado a partir de una gráfica de Boxplot con el que se quiere visualizar la distribución de los datos, y revisar si tienen buena simetría o si se alejan de la distribución normal. Esto permite identificar la importancia del filtrado de datos, y se logra determinar que los datos marcados en rojo no serán tenidos en cuenta ya que son valores atípicos en cada uno de los ejes. Debido a esto, se realiza una validación y control de calidad de los datos para llevar a cabo un óptimo desarrollo.

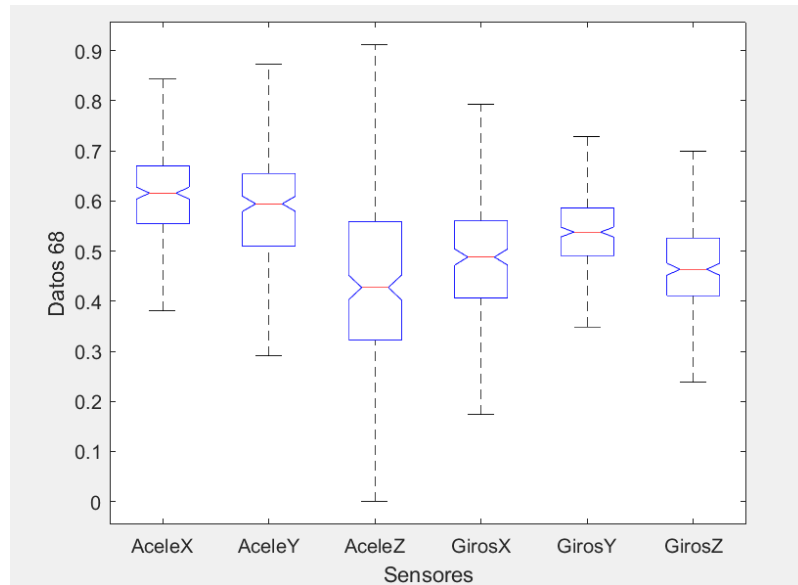
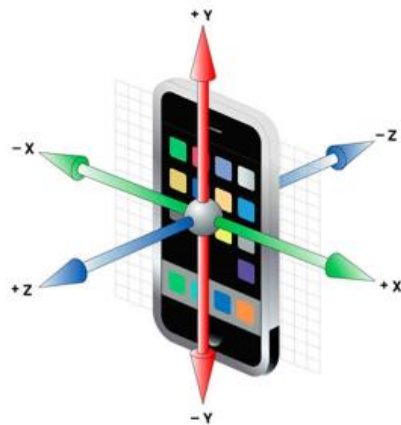


Ilustración 23 Grafica Boxplot Datos Filtrados en cada sensor.

Una vez filtrados y normalizados los datos, se tiene como resultado la gráfica de la ilustración anterior, en la que se evidencia el filtrado de los datos (Ya no se encuentran anomalías) y se puede continuar con el desarrollo y análisis del proyecto.

5.1.5. Análisis de datos

El vehículo en movimiento permite obtener variaciones en el sensor inercial (acelerómetro) que se pueden ver como una serie de tiempo muestreada cada segundo a partir de que el usuario inicia la obtención de datos. Cuando se utiliza el acelerómetro para detectar anomalías viales, se espera que la aceleración varíe en diferentes “picos”.



(a)



(b)

Ilustración 24. Ejes de sensor acelerómetro de un celular ubicado en posición perpendicular(a) y posición del teléfono en un vehículo(b).

5.1.6. Extracción de características

Esta sección describe las características obtenidas a partir de los datos del sensor inercial, las cuales serán utilizadas para clasificar, mediante un árbol de decisión, las anomalías de la malla vial. Para esto se tomaron las siguientes anomalías viales: reductor de velocidad amarillo (policía amarillo), reductor de velocidad blanco (policía blanco) y vía plana. A continuación, se observan las gráficas de dispersión de estas irregularidades y el comportamiento que tienen las características seleccionadas que son: amplitud de la transformada de Fourier, la media móvil y la desviación estándar.

5.1.6.1. Amplitud de transformada de Fourier vs Media móvil

La transformada de Fourier es un método que descompone una señal en sus componentes espectrales individuales y así proporciona información sobre su comportamiento en el dominio de la frecuencia. Para este proyecto se evalúa esta característica de las señales respecto a la media móvil, con el fin de encontrar información

relevante para la clasificación. Los resultados se ven reflejados en la siguiente grafica de dispersión.

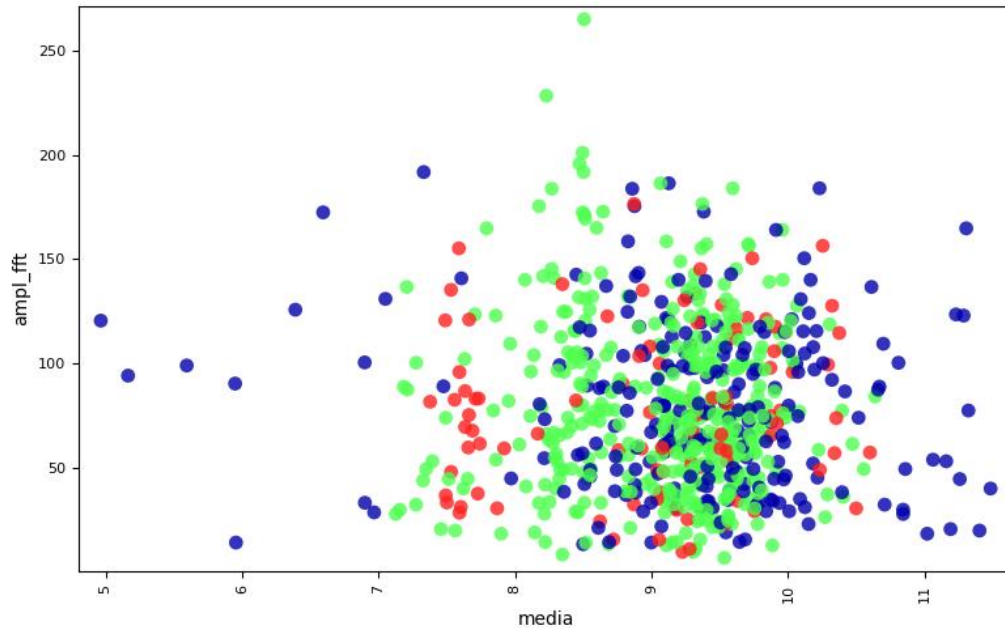


Ilustración 25 Grafica de dispersión amplitud de Fourier vs Media Móvil.

Como se puede observar, en la gráfica de dispersión, no se logra identificar un patrón que describa alguna clasificación de las irregularidades, por el contrario, se agrupan todos los datos mayormente entre los valores de 7 a 11 para la media móvil y de 0 a 200 de amplitud para la transformada de Fourier. Con esto se dificulta la caracterización ya que cualquier nuevo dato que entre será sectorizado entre estos valores y no se conocerá el tipo de irregularidad. Por esta razón, no se tiene una facilidad de clasificación entre estas dos características ya que no logra diferenciar las irregularidades, con lo que se puede concluir que no es viable tomar como características principales para el clasificador la amplitud de la transformada rápida de Fourier respecto a la media móvil.

5.1.6.2. Amplitud de transformada de Fourier vs Desviación estándar

Otra de las características de extracción de información que se evaluó para hacer parte del clasificador fue la amplitud de la transformada de Fourier respecto a la desviación estándar. A continuación, se muestra la gráfica de dispersión con estas características.

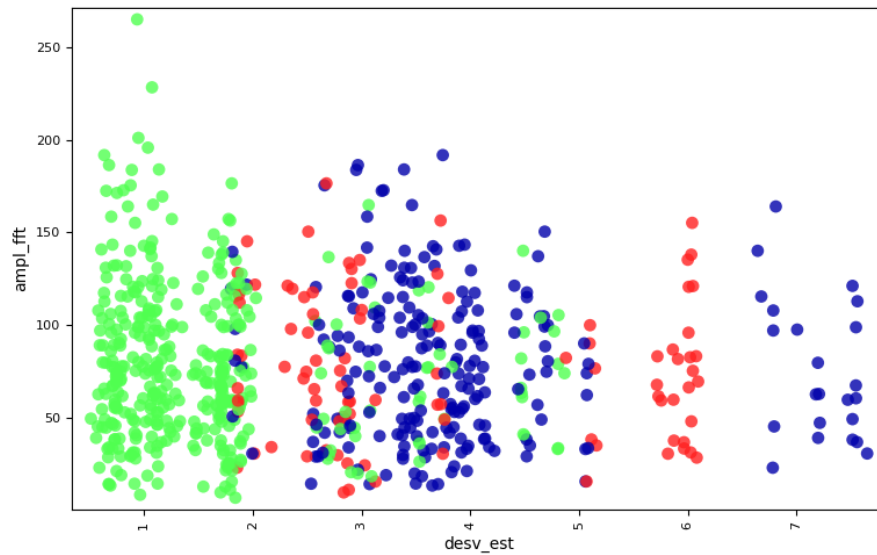


Ilustración 26 Gráfica de dispersión amplitud de transformada de Fourier vs desviación estándar.

En la gráfica anterior se puede observar el comportamiento de los datos agrupados de acuerdo a la clasificación realizada con los datos de entrenamiento, para los datos de vía plana (color verde). Se puede identificar que hay mayor concentración de estos entre los rangos de 0 a 2 de la desviación estándar y amplitud de transformada de Fourier de 0 a 250, teniendo así una clasificación buena para esta característica sobre la vía, ya que no se encuentran datos que hagan parte a otra anomalía. Por otro lado, en cuanto a los datos del policía amarillo (color rojo) y de hueco medio (color azul), se observa que estos tienen

el mismo comportamiento ya que se encuentran mezclados entre los rangos de 2 a 7 en desviación estándar y de 0 a 200 para amplitud de transformada de Fourier indicando así unos bajos niveles de clasificación para estas dos características. Con lo anterior se puede concluir que, para la vía plana las características de la amplitud de la transformada de Fourier respecto a la desviación estándar le permiten tener una buena clasificación, sin embargo, para el policía amarillo y hueco medio no se logra identificar una clasificación entre ellos.

5.1.6.3. Fase de transformada de Fourier vs Media móvil

Ahora se analizará el comportamiento de la fase de la transformada de Fourier respecto a la media móvil, en la siguiente gráfica de dispersión.

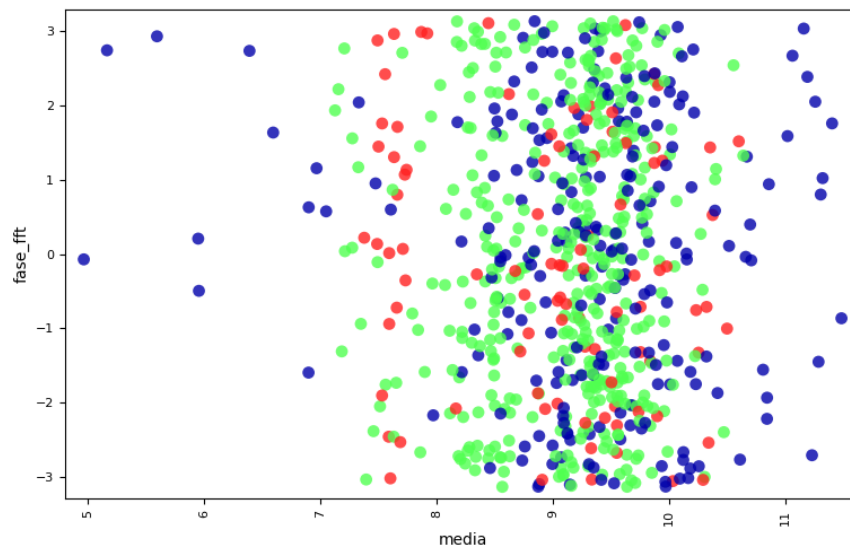


Ilustración 27 Gráfica de dispersión fase de la transformada de Fourier vs Media móvil

En la ilustración 27, se observa un comportamiento muy similar al de las características de la amplitud de transformada de Fourier respecto a la media móvil. Aquí se observa que los datos están concentrados entre los rangos de 7 a 11 para la media y de

-3 a 3 para la fase. Este comportamiento es general para todos los datos, siendo así que si ingresa un dato nuevo no se sabrá con certeza que tipo de irregularidad es, ya que lo agrupará entre este rango de valores. En ese sentido, la fase de la transformada de Fourier respecto a la media móvil no son buenas características para el clasificador de irregularidades.

5.1.6.4. *Media móvil vs Desviación estándar*

Una de las características de extracción de información utilizada fue la media móvil. Esta característica se obtiene promediando un rango de datos denominado ventana. Con estos datos, se pueden identificar tendencias que nos ayuden a clasificar las irregularidades. A continuación, se muestra una gráfica de dispersión con el comportamiento de la media móvil respecto a la desviación estándar.

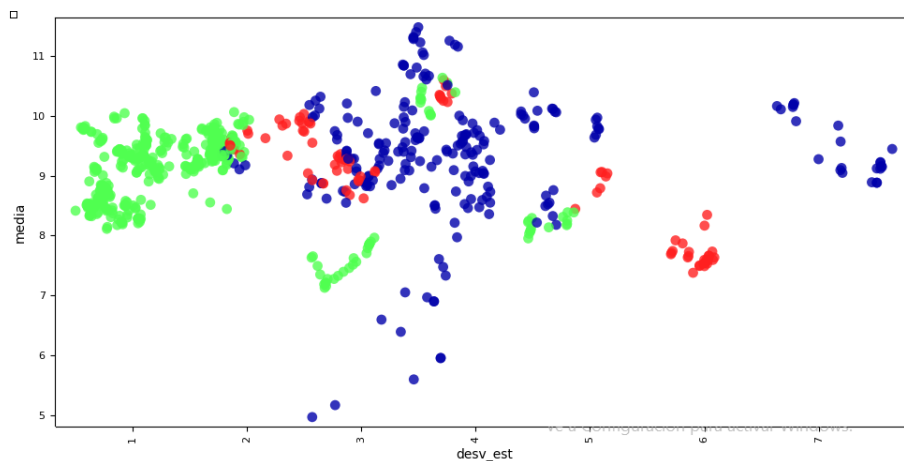


Ilustración 28. Gráfica de dispersión relación entre media móvil vs desviación estándar.

Como se puede observar en la gráfica de dispersión, hay una clasificación de tres irregularidades, hueco medio en color verde, donde se agrupan los datos entre los rangos

de valores de 1 a 3 para la desviación y entre 7 y 11 para la media móvil. Se evidencia que hay datos fuera de la agrupación que se describió, lo cual nos permite conocer que la precisión no será del 100%. En color azul, se encuentran los datos de vía plana, en los que se identifica un rango de valores mayormente en la zona de 2 a 5 para la desviación y de 5 a 11 para la media móvil. También hay datos que se encuentran fuera de estos rangos, lo cual también permite ver que la precisión para esta irregularidad no será del 100%. Los datos rojos representan los reductores de velocidad blancos, donde el conjunto de datos está un poco más disperso y no se logra identificar un patrón ya que tiene datos muy similares a los del grupo anterior. Esto se debe a las características de la irregularidad. Los datos se encuentran de 2 a 6 para la desviación y de 7 a 10 para la media, además de que algunos datos son atípicos. Con esta gráfica se puede concluir que, en el grupo de datos, se pueden identificar las tres clasificaciones esperadas, dando así un buen indicio para usar la desviación estándar y media móvil como características más representativas para el algoritmo de clasificación.

Como resultado de las diferentes exploraciones que se hicieron, se puede concluir que las dos características que más información proporcionan para la clasificación de irregularidades son la media móvil y la desviación estándar, respecto a las otras características que se implementaron.

5.1.7. Clasificador de árbol de decisión

El clasificador de árbol de decisión se seleccionó para este proyecto debido a su facilidad de implementación, ya que se basa en la toma de decisiones a partir de

condiciones, su preparación de datos no es muy estricta y se pueden aprovechar para tareas de clasificación y de regresión.

Para la implementación de este clasificador se hizo uso la librería *sklearn*, útil para el *machine learning* en Python. Esta es una librería de código abierto donde proporciona una gama de algoritmos de aprendizaje supervisados y no supervisados.

La librería de *sklearn* está estandarizado y tiene una serie de pasos para su uso, como primer paso se elige el modelo a trabajar, para este proyecto se eligió el árbol de decisión y librería denominada *DecisionTreeClassifier*, en segunda medida se establecen los parámetros del modelo o las características que permitirán hacer la clasificación, para el tercer paso se entrena el modelo, utilizando el método de ajuste y por último se prueba el modelo con datos nuevos.

El algoritmo de árbol de decisión implementado tuvo 4 ramificaciones con una predicción del 92,6%. Las características usadas en este árbol de decisión fueron la desviación estándar con una importancia de 86,5% y la media móvil con una importancia de 13,5%.

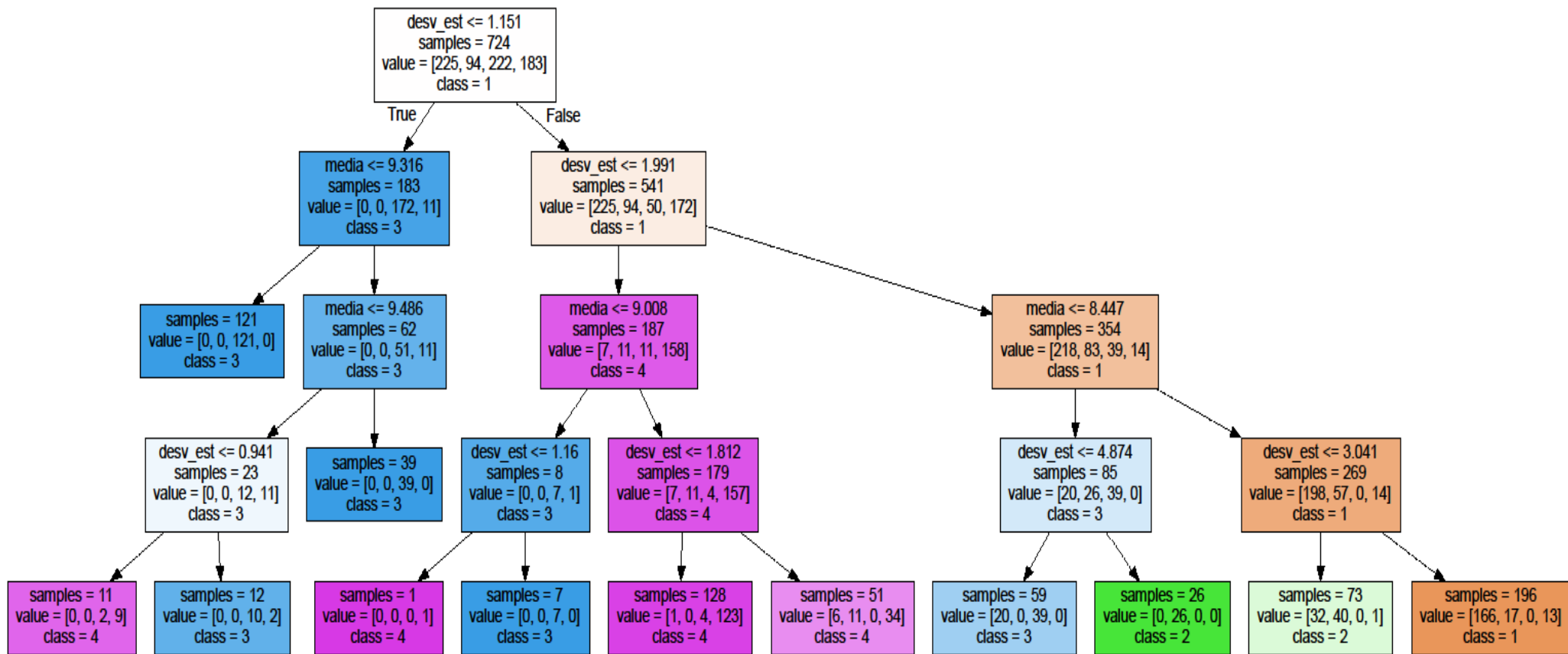


Ilustración 29 Árbol de decisión de 4 ramificaciones.

En la raíz del árbol se tiene una entrada de 724 datos, y se tienen 4 tipos de irregularidades: policía acostado amarillo, policía acostado blanco, hueco medio y vía plana. El árbol inicia haciendo las validaciones de agrupación correspondientes y respecto a ese resultado, filtra teniendo en cuenta las características seleccionadas (media móvil y desviación estándar). Respecto a esto, los datos que cumplen con las condiciones toman una ramificación u otra, y este proceso se repite 4 veces para lograr la agrupación más acertada y precisa.

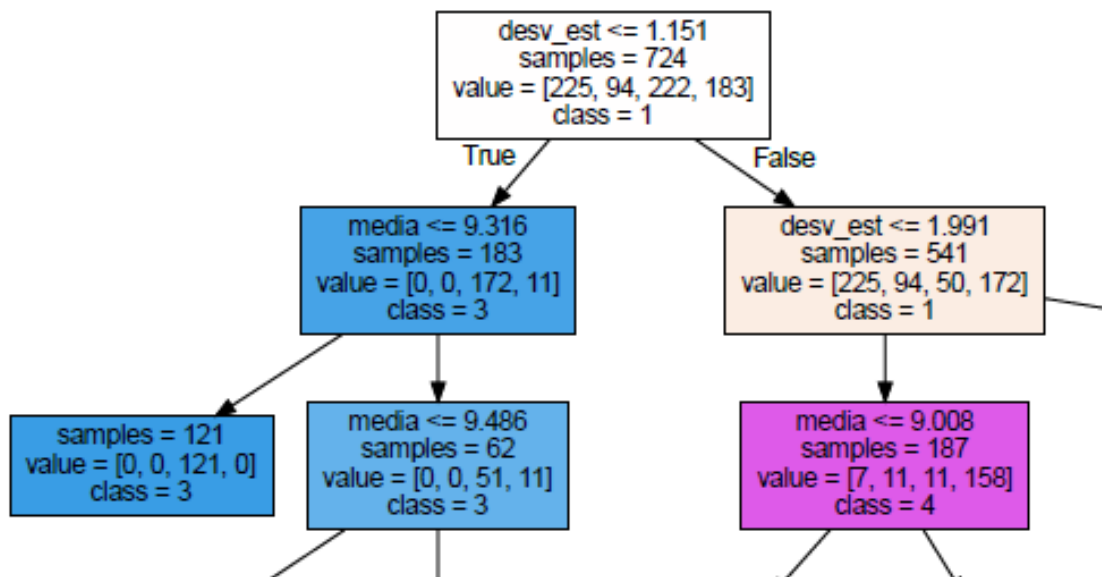
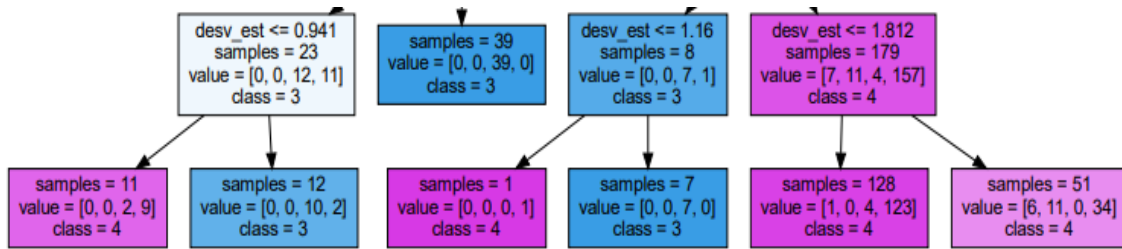


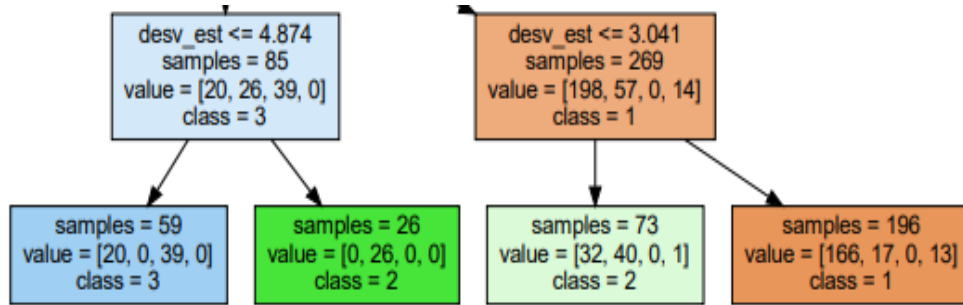
Ilustración 30 Primer nodo final del árbol de decisión.

En la imagen anterior se puede ver que hay un nodo final en la segunda ramificación del árbol, debido a que encontró que, de los 724 datos, 121 corresponden a la clase 3 de tipo de irregularidad vía plana, mientras que para los 603 datos restantes no ha encontrado una clasificación exacta o aproximada. En tal caso, el árbol sigue su clasificación establecida.

Finalmente, el árbol termina su clasificación agrupando los datos en las 4 características descritas anteriormente. En la siguiente imagen se puede observar el final del árbol con su última ramificación.



(a)



(b)

Ilustración 301 Nodos finales del árbol de decisión de 4 ramificaciones.

También se implementó un árbol de decisión de 5 ramificaciones. Este árbol de decisión tuvo una precisión es de un 91,3%, y la importancia de sus características es: para la desviación estándar es de 82,7% y de la media móvil es de 17,2%.

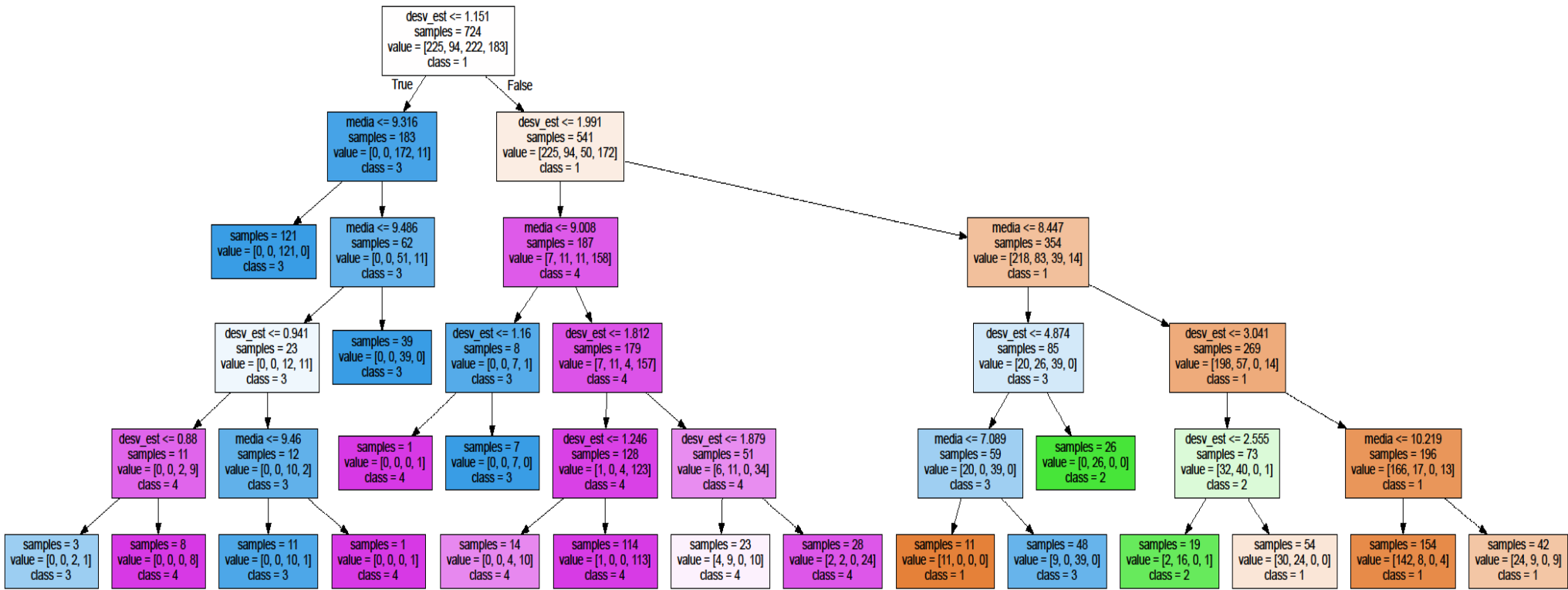


Ilustración 32 árbol de decisión de 5 ramificaciones.

También se implementó un árbol de decisión de 3 ramificaciones, donde se disminuye la precisión a un 87,6%. Respecto a las características, se encontró que la importancia de la desviación estándar es de 89,3% y la media móvil es de 16,1%.

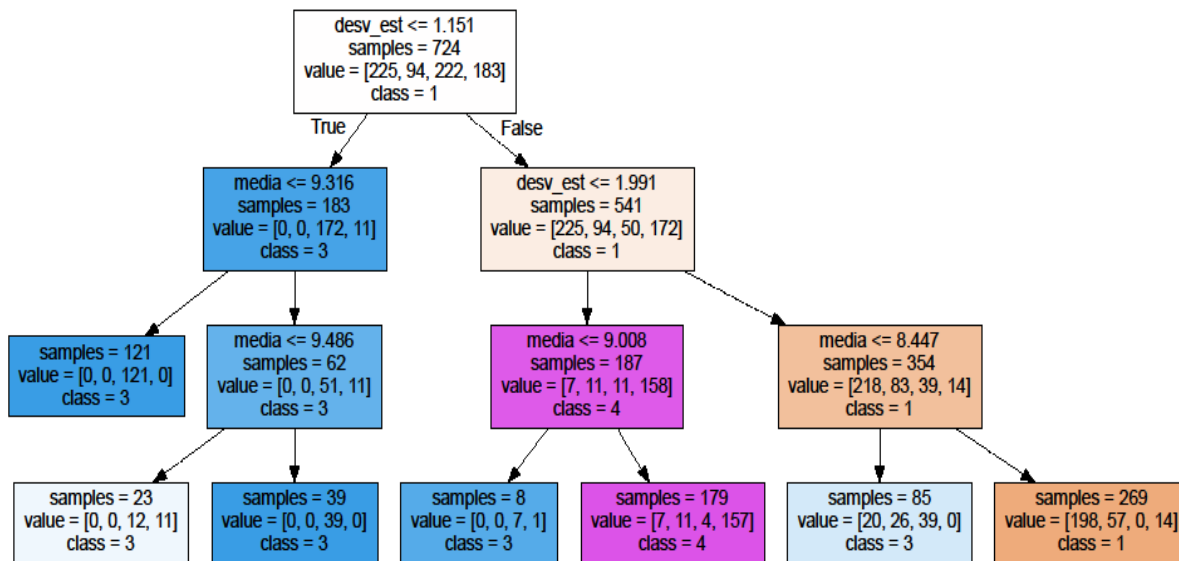


Ilustración 33 Árbol de decisión de 3 ramificaciones.

De acuerdo con los resultados de los árboles de decisión, se puede concluir que para el algoritmo de clasificación de acuerdo con las características de desviación estándar y media móvil, la mejor precisión obtenida es 92,6%, lograda con 4 ramificaciones. Además, se puede afirmar que entre más ramificaciones tenga el árbol de decisión la precisión disminuye, de igual manera sucede con muy pocas ramificaciones. Esto significa que se debe buscar la precisión más cercana a 100%, y este es un proceso experimental.

6. Resultados.

6.1. Resultados

Para el presente trabajo se hizo uso de los sensores inerciales de acelerómetro y giroscopio, con los que se realiza la toma de datos usados para encontrar las irregularidades de la malla vial. Sin embargo, se puede decir que el giroscopio no fue un sensor que nos mostrara resultados relevantes, ya que al tomar alguna irregularidad si se ven alteración de sus ejes, pero el vuelve a tomar su estado natural, es por lo que a la hora de la clasificación los datos tomados de este sensor no fueron tenidos en cuenta.

Para la captura de datos a través de SmartApp, se determina que el celular estará ubicado de forma perpendicular al vehículo, con el fin de reducir ruido y datos no deseados, teniendo el celular en esta posición, se puede ver que el eje de estudio del acelerómetro es el eje Y, debido a que este eje capturara los sobresaltos que presente la vía. En esta posición, el eje Z toma información acerca de la aceleración en la que se encuentra el automóvil, y en el eje X se muestran los movimientos laterales.

Para el desarrollo de la aplicación SmartApp, se hizo uso del programa App Inventor, con el cual se pudo crear una aplicación móvil capaz de capturar los datos de los sensores con un muestreo de un segundo, mostrando así el potencial que tiene este tipo de proyectos de investigación. Por otro lado, para la toma de datos, se presentaron inconvenientes con el envío de datos del celular a la base de datos, ya que la conexión de red presenta caídas por sectores en las que la cobertura no era estable.

La clasificación con la técnica del árbol de decisión es efectiva, ya que nos arroja una exactitud aproximada del 92,6%. Si esta predicción se compara con los antecedentes, por ejemplo, el trabajo desarrollado por (Allouch, Koubaa, Abbas, & Ammar, 2016) donde se obtuvo una precisión del 98,6%, pues es algo baja respecto a este proyecto, y es por eso que se deben tener en cuenta otras variables que perjudican la precisión del proyecto, como lo es la cantidad de datos recolectados, los medios de transporte implementados para la captura de datos o incluso la velocidad en la que se tomaron los datos. Sin embargo, la precisión de este proyecto es superior a las predicciones más bajas, como fue el proyecto realizado por (Lekshmiathy, Samuel, & Velayudhan, 2020) que obtuvieron una efectividad del 80%, lo cual nos sitúa en una posición buena respecto a los demás trabajos realizados con este enfoque.

6.2. Conclusiones

Se puede afirmar que por medio de los dispositivos móviles y de los sensores con los que estos dispositivos están equipados, se puede detectar condiciones del estado de las carreteras, además de una clasificación de las irregularidades dentro de un nivel aceptable del 92,6% de precisión que nos entrega el árbol de decisión, respecto a las características se obtuvo una importancia del 86,5% para la desviación estándar y un 13,4% para la media móvil.

Por otro lado, una de las complicaciones que se obtuvo fue la captura de los datos, ya que al ser una aplicación que necesita una conexión a internet estable, se presentaron problemas cuando se viajaba en zonas de la ciudad donde la conectividad es nula o muy inestable, lo que conlleva a que la aplicación se desconecte de la base de datos y no realice

envío a la misma. Por otra parte, cuando la conexión es inestable se realiza un envío erróneo de datos ya que llena de ceros o null todos sus atributos, los cuales deben ser removidos manualmente en el proceso de manipulación de datos.

Por otro lado, al momento de clasificar las irregularidades se encontró varias similitudes en el comportamiento de los datos respecto a las anomalías, como en el caso de los reductores de velocidad generalmente llamados “policías acostados” (medio y alto) donde el rango de datos tiene características muy similares y no se logra identificar con precisión que tipo de anomalía es. Esto perjudica la clasificación y así mismo el porcentaje de exactitud en la predicción que se quiere alcanzar.

6.3. Trabajos futuros

El presente trabajo se puede complementar implementando un front-end con la tecnología de Angular o React que son frameworks utilizados para este tipo de desarrollos, donde se cree una interfaz gráfica que realice un feed-back a los interesados sobre la ubicación aproximada de las irregularidades viales, pueden ser vía web, vía aplicación móvil o híbrida, con el fin de informar sobre el estado de la vía a la población de la ciudad de Bogotá.

Respecto a la aplicación móvil, se puede realizar una migración a lenguajes de programación híbridos o multiplataformas como lo son Ionic y React Native, frameworks de JavaScript que son más completos y pueden garantizar la mejor optimización de los recursos del dispositivo móvil, para asegurar una buena recolección de datos.

Por otro lado, para un trabajo futuro se pueden agregar más características representativas que ayuden a tener más información sobre las vías, no solo de una ciudad capital, sino para el territorio nacional.

7. Referencias

- Allouch, A., Koubaa, A., Abbes, T., & Ammar, A. (2016). RoadSense: Smartphone Application to Estimate Road Conditions using Accelerometer and Gyroscope. *IEEE*.
- Bustamante, R., Garcia, A., Arce, L., Curiel, L., Izquierdo, J., & Ramirez, R. (2022). *Visualizing Street Pavement Anomalies through Fog*. Ciudad de Mexico.
- C.Müller, A., & Guido, S. (2016). *Introduction of Machine Learning with Python*. United States of America: O'REILLY.
- Date, C. (1975). *Introducción a los sistemas de bases de datos*. Pearson Prentice Hall.
- Edwan, E., Sarsour, N., & Alatrash, M. (2019). Mobile Application for Bumps Detection and Warning Utilizing Smartphone Sensors. *IEEE*.
- Eriksson, J., Girod, L., Newton, R., Madden, S., Balakrishnan, H., & Hull, B. (2008). The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring. *IEEE*.
- Espin, P. (2010). Diseño y construcción de una plataforma didáctica para medir ángulos e inclinación usando sensores inerciales como acelerómetro y giroscopio.
- Fericean, S. (2019). *Inductive Sensors for Industrial Applications*. Artech House.
- Ferjani, I., & Ali, S. (2022). *How to get best predictions for road*. Arabia Saudita.
- Hayes, A. (12 de Octubre de 2021). *Investopedia*. Obtenido de <https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp>

- INDRA. (2014). *Smart Cities encuentra global sobre el estado de smart cities*. Bogotá.
- Instituto de Desarrollo Urbano. (2021). *Boletín Técnico Extensión y Estado de la Malla Vial de Bogotá D.C.* Bogotá.
- Kaplan, E., & Hegarty, C. (2006). *Understanding GPS Principles and Applications*. Artech House.
- kreyszig, E. (2003). *Matemáticas Avanzadas para Ingeniería Vol. II*. Limusa Wiley.
- Kulkarni, A., Mhalgi, N., Gurnani, S., & Giri, N. (2014). *Pothole Detection System using Machine Learning on Android*. Mumbai.
- Kyriakou, C., Christodoulou, S., & Dimitriou, L. (2019). *Smartphone-Based Pothole Detection Utilizing*. Chipre.
- Leizerovych, R., Sindenko, L., Kondratenko, G., & Kondratenko, Y. (2020). IoT-complex for Monitoring and Analysis of motor highway condition using artificial neural networks. *IEEE*.
- Lekshmiathy, J., Samuel, N., & Velayudhan, S. (2020). Vibration vs Vision: best approach for automated pavement distress detection. *Springer*.
- Mandziuk, D. (2007). *Challenges for Computational Intelligence*. Springer .
- Martínez, F., Escalante, H., González, L., Aragón, M., & Ricardo, M. (2018). *Evaluation of Detection Approaches for Road Anomalies Based on Accelerometer Readings—Addressing Who's Who*. Chihuahua.

- Microsoft. (2021). *Algoritmo de asociación de Microsoft*. Obtenido de <https://docs.microsoft.com/es-es/analysis-services/data-mining/microsoft-association-algorithm?view=asallproducts-allversions&viewFallbackFrom=sql-server-ver15>
- Microsoft. (13 de 09 de 2021). *Algoritmo de clústeres de Microsoft*. Obtenido de <https://docs.microsoft.com/es-es/analysis-services/data-mining/microsoft-clustering-algorithm?view=asallproducts-allversions&viewFallbackFrom=sql-server-ver15>
- Mitchell, T. (1997). *Machine Learning*. New York: McGraw-Hill Science/Engineering/Math.
- Molnar, C. (2021). *Interpretable Machine Learning a guide for making black box models explainable*. Obtenido de github: <https://fedefliguer.github.io/AAI/index.html>
- Müller, A., & Guido, S. (2016). *Introduction to Machine Learning with Python*. Boston: O'REILLY.
- Organización para la Cooperación y el Desarrollo Economico . (2018). *Medicion y aplicaciones de IoT*. Mexico.
- Pawar, K., Jagtap, S., & Bhoir, S. (2020). *Efficient pothole detection using smartphone sensors*. India.
- Ramírez, R. (2019). *Métodos para el desarrollo de aplicaciones móviles*. Cataluña: Universidad de cataluña.

- Sagar, V., & Annapurna, V. (2018). *ANDROID APPLICATION FOR POTHOLE*. Ananthapuramu.
- Sattar, S., Li, S., & Chapman, M. (2021). Developing a near real-time road surface anomaly detection approach for road surface monitoring. *Journal of the international Measurement*.
- Timón, C. E. (2017). *Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo-herramientas Open Source que permiten su uso*. Cataluña.
- Tiwari, S., Bhandari, R., & Raman, B. (2020). *RoadCare: A Deep-learning Based Approach to Quantifying*. Mumbai.
- Triola, M. (2006). *Estadística Novena Edición*. Mexico: Pearson Educación.
- Wang, H., Huo, N., Li, J., Wang, K., & Wang, Z. (2018). A Road Quality Detection Method Based on the Mahalanobis-Tagucho System. *IEEE*.
- We Are Social. (2023). *Digital 2021 Colombia*.
- Witten, I., & Frank, E. (2005). *Data Mining Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers.
- Wu, C., Wang, Z., Hu, S., Lepine, J., Na, X., Ainalis, D., & Stettler, M. (2020). *An Automated Machine-Learning Approach for Road*. Hangzhou.