

**INSTITUCIÓN UNIVERSITARIA POLITÉCNICO GRANCOLOMBIANO  
FACULTAD DE INGENIERÍA, DISEÑO, E INNOVACIÓN  
MAESTRÍA EN INGENIERÍA DE SISTEMAS  
GRUPO DE INVESTIGACIÓN CITIC  
LÍNEA DE PROFUNDIZACIÓN: CIENCIA DE DATOS**

**“ESTILOS DE CONDUCCIÓN DE AUTOMÓVILES. RECONOCIMIENTO  
AUTOMÁTICO USANDO LOS SENSORES DE LOS SMARTPHONES”**

**PRESENTA:**

**JOVAN FARIK FERNÁNDEZ JOYA**

**1910020014**

**ASESOR TEMÁTICO:**

**WILSON JAVIER SARMIENTO, PhD**

**CO ASESOR TEMÁTICO:**

**GABRIEL EDUARDO ÁVILA BUITRAGO, MSc**

**JUNIO 2021**

## TABLA DE CONTENIDO

RESUMEN .....	7
INTRODUCCIÓN .....	8
CONTRIBUCIONES.....	12
OBJETIVOS DE LA PROPUESTA.....	13
ESTRUCTURA DEL DOCUMENTO .....	14
REVISIÓN DEL ESTADO DEL ARTE .....	16
BÚSQUEDA SISTÉMICA.....	20
HALLAZGOS .....	24
ESTRATEGIA METODOLÓGICA .....	28
FASE 1.....	28
FASE 2.....	29
FASE 3.....	34
RESULTADOS.....	36
SELECCIÓN DE CARACTERÍSTICAS.....	36
RECONOCIMIENTO DE EVENTOS DE CONDUCCIÓN.....	37
EVALUACIÓN CON CARACTERÍSTICAS MIXTAS .....	37
EVALUACIÓN CON CARACTERÍSTICAS ESTADÍSTICAS.....	39
RECONOCIMIENTO DE TIPOS DE CONDUCCIÓN.....	40
EVALUACIÓN CON CARACTERÍSTICAS MIXTAS .....	41
HALLAZGOS .....	44
CONCLUSIONES Y RECOMENDACIONES FINALES .....	45
REFERENCIAS.....	50
ANEXOS .....	57



## LISTA DE TABLAS

Tabla 1. Muertes por accidentes de tránsito en el mundo.....	8
Tabla 2. Accidentes de Tránsito (casos). Colombia, 2005 - 2014 (anual). .....	9
Tabla 3. Artículos por año primera ejecución .....	22
Tabla 4. Artículos por año segunda ejecución.....	22
Tabla 5. Criterios de exclusión de artículos.....	23
Tabla 6. Uso de los sensores de los smartphones.....	24
Tabla 7. Artículos científicos por sensor.....	24
Tabla 8. Eventos capturados por los smartphones. ....	25
Tabla 9. Métodos de aprendizaje de máquina implementados. ....	25
Tabla 10. Métodos de aprendizaje por artículo científico. ....	26
Tabla 11. Estilos de conducción identificados.....	27
Tabla 12. Reducción de etiquetado de los eventos capturados. ....	33
Tabla 13. Configuración de clasificadores.....	34
Tabla 14. Mejores resultados de sensibilidad para datasets de características mixtas del subconjunto de eventos.....	38
Tabla 15. Mejores resultados de especificidad y sensibilidad para datasets de características estadísticas subconjunto de eventos.....	40
Tabla 16. Mejores resultados de especificidad y sensibilidad para datasets de características mixtas subconjunto tipo de conducción.....	42
Tabla 17. Mejores resultados de especificidad y sensibilidad para datasets de características estadísticas subconjunto tipo de conducción. ....	43

## LISTA DE FIGURAS

Figura 1. Factores que influyen el estilo de conducción.....	18
Figura 2. Resultados de una búsqueda sobre sistemas en desarrollo que están indagando sobre técnicas, métodos y sensores usados en el reconocimiento de estilos de conducción.....	20
Figura 3. Etapas de la revisión sistemática.....	20
Figura 4. Sensor Data Recorder.....	30
Figura 5. Mapa Gran Granada - Bogotá D. C.....	32
Figura 7. Gráfica de puntos en el espacio ROC mejores resultados datasets de características mixtas subconjunto de eventos.....	39
Figura 8. Gráfica de puntos en el espacio ROC mejores resultados datasets de características estadísticas subconjunto de eventos.....	40
Figura 9. Gráfica de puntos en el espacio ROC mejores resultados datasets de características mixtas subconjunto de tipo de conducción.....	42
Figura 10. Gráfica de puntos en el espacio ROC mejores resultados datasets de características estadísticas subconjunto tipo de conducción.....	44

## **ANEXOS**

Anexo 1. Composición del archivo del dataset estadístico.....	58
Anexo 2. Estructura del dataset de características frecuenciales. ....	60
Anexo 3. Datos tabulados de los procesos de clasificación. ....	62
Anexo 4. Artículo científico.....	73

## **RESUMEN**

Los estilos de conducción son aquellos aspectos únicos que identifican la forma en que cada conductor conduce un vehículo. Cada persona tiene un estilo propio que lo identifica y que ha sido desarrollado a través del tiempo y se ha visto influenciado por factores humanos y medioambientales. El presente proyecto busca identificar los estilos de conducción con el apoyo de un conjunto de sensores presentes en los smartphones, con el fin de clasificar a los conductores de vehículos. Este proyecto busca generar un aporte en el estudio y caracterización, por medios computacionales, de uno de los factores que es tercera causa de accidentes de tránsito, cuyos resultados podrán ser usados a futuro en aplicación como: tarificación diferencial de pólizas de seguros de automóviles, mejorar el consumo de gasolina, conducción autónoma, generación de rutas de tráfico vehicular seguro, entre otras.

## **ABSTRACT**

Driving styles are those unique aspects that allow to identify the way a person drives a car. Each person has a unique style of driving, which has evolved with time, and has been influenced by human and environmental factors. This work proposes a contribution to the study and characterization, by computational methods, of a factor that is the third cause of vehicular traffic accidents. The results of this work. Must be used in differential vehicular insurance policies, improve gas consume, automatic driving, generation of safe vehicular paths, among others.

## **PALABRAS CLAVE**

Estilos de conducción, teléfono inteligente, conductor de vehículo.

## **KEY WORDS**

Driving styles, smartphone, vehicle driver.

## INTRODUCCIÓN

En la actualidad, una gran cantidad de ciudades en el mundo, particularmente en países en vía de desarrollo, se presenta un alto flujo vehicular el cual genera congestión en sus diferentes vías. Debido a esto, las personas que conducen vehículos tienden a presentar estilos de conducción agresivos, que en muchas ocasiones se convierten en accidentes de tránsito. El estilo de conducción agresivo se considera como la tercera causa generadora de accidentes de tránsito [1].

Con el objetivo de realizar un seguimiento a los accidentes de tránsito, en el año 2013, la Organización Mundial de la Salud (OMS) pudo establecer que, en promedio, unas 3000 personas mueren a diario en todo el mundo a causa de accidentes de tránsito, donde: el 22% son peatones, el 23% son motociclistas y el 5% son ciclistas. El estudio también reveló que cada año mueren cerca de 1,24 millones de personas y que entre 20 y 50 millones resultan heridas en las diferentes carreteras del mundo [1]. Este mismo estudio realizado en el 2015, estableció que la tasa de muertes aumentó llegando a 1.35 millones de personas, convirtiéndose en una de las principales causas de muerte en personas de todas las edades y la primera causa de muerte de adultos jóvenes entre los 15 y 29 años, sin embargo, también es posible afirmar que las medidas de seguridad implementadas en diferentes países ha evitado un mayor crecimiento teniendo en cuenta el aumento de la población y la compra de automóviles [2]; en la Tabla 1 se muestra la cantidad de personas fallecidas según estudios de la OMS.

Tabla 1. Muertes por accidentes de tránsito en el mundo.

2010	2013	2015
1.20	1.24	1.35

*Fuente: Elaboración propia a partir de los datos de los informes sobre la situación mundial de la seguridad vial de 2011, 2013 y 2015. Valores dados en millones.*

Un informe entregado por el Instituto de Medicina Legal en el año 2012 resalta que las cifras de muertes por accidentes de tránsito en Colombia aumentaron un 1.5% en el 2011 con respecto al 2010 [3] y, en el Plan Nacional de Seguridad Vial se resalta el incremento de muertes en accidentes de tránsito del 2.1% en el 2014



frente al 2013, convirtiéndose en la segunda causa de muerte violenta y la primera causa de muerte en las personas jóvenes de menos de 30 años en Colombia [4].

Debido a que la accidentalidad vehicular se ha convertido en un tema de salud pública, la Organización de las Naciones Unidas (ONU), decretó en marzo de 2010 el decenio de la seguridad vial, el cual comenzó a regir en el 2011 y terminó en el 2020. Este plan buscaba mejorar el comportamiento de los conductores al volante, lo cual está altamente relacionado con las observaciones de la OMS, quién definió la generación de hábitos más seguros de comportamiento en la conducción de vehículos [5] como un pilar fundamental de prevención de accidentes viales.

La OMS en el año 2015 se planteó como meta la disminución en un 50% de las muertes y las lesiones por causa de accidentes de tránsito antes de que finalizara el 2020, meta que está asociada a unos de sus Objetivos de Desarrollo Sostenible [2].

En Colombia se desarrolló el Plan Nacional de Seguridad Vial 2011-2021, el cual actualizó el Plan Nacional de Seguridad Vial 2011-2016; en este plan se estableció como meta reducir en un 26% las muertes por accidentes de tránsito desde ese año hasta el 2021. En cifras, Colombia ha tenido un total de 1.836.373 accidentes de tránsito entre el 2005 y 2014, dejando lesionadas a 411.956 personas y a 58.121 personas fallecidas [4]. Las cifras de accidentes de tránsito anuales entre el 2005 y el 2014 se pueden observar en la Tabla 2.

Tabla 2. Accidentes de Tránsito (casos). Colombia, 2005 - 2014 (anual).

2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
209.568	166.739	185.568	187.270	180.349	173.393	183.775	197.359	194.659	157.693

*Fuente: Plan Nacional de Seguridad Vial 2011-2021.*

Alcanzar este tipo de metas requiere hacer las cosas de una forma diferente y ahí es donde entra a participar la tecnología, la cual está jugando un papel cada vez más importante en la vida de las personas facilitando las tareas cotidianas [6], y el Internet, el cual ha revolucionado el mundo de las comunicaciones permitiendo la difusión de información a cualquier parte del mundo en solo segundos [7]. Ahora, sumémosle a todo este paradigma un integrante más, la internet de las cosas (IoT),

cuyo concepto base es tener una gran cantidad de dispositivos interactuando entre sí para lograr objetivos comunes [8], como RFID, sensores y smartphones, por nombrar algunos, facilitando el diario vivir de las personas.[6].

Apoyado en lo anterior y en diferentes estudios de investigación científica realizados en algunos países de Europa y en Norte América, se han desarrollado procesos que logran identificar los estilos de conducción de los conductores con la ayuda de la IoT, ya sea instalando sensores como acelerómetros, barómetros y GPS en los vehículos [9], o con los sensores presentes en los smartphones [6], [10], [11]. Todo indica que los smartphones serán algunos de los dispositivos que serán utilizados en el futuro para estas tareas, ya que su bajo costo y facilidad de configuración, uso y disponibilidad, permitirán llegar a más conductores. Entre los datos que pueden ser capturados por los diferentes dispositivos IoT tenemos: la velocidad, la aceleración, la desaceleración, los giros, el cambio de carril y el frenado, entre otros. Hay que tener en cuenta que todos estos datos pueden verse afectados por las condiciones de la vía, lo que permitirá evidenciar cómo reaccionan los conductores a este tipo de circunstancias [12].

Hay que tener claro que los dispositivos IoT por sí solos no cumplen labor alguna y, por consiguiente, se deben determinar las variables necesarias que permitirán realizar el seguimiento, con el fin efectuar una clasificación sólida, ya que los resultados dependerán totalmente de esta selección. Después de tener claramente identificadas las variables, se aplican uno o varios algoritmos de aprendizaje de máquina, con el fin de evaluar e identificar el estilo de conducción del conductor [12]. Son muchos los algoritmos que pueden ser utilizados solos o en combinación. Algunos de ellos pueden ser: Dynamic Time Warping (DTW), redes neuronales artificiales multicapa, árboles de decisión, máquinas de soporte vectorial, modelos ocultos de Markov (HMM) y Naïve Bayes entre otros [6].[12]. Son muchos los algoritmos que pueden ser utilizados solos o en combinación. Algunos de ellos pueden ser: Dynamic time warping (DTW), redes neuronales artificiales multicapa, árboles de decisión, máquinas de soporte vectorial, modelos ocultos de Markov (HMM) y Naïve Bayes entre otros [6].

El reconocimiento de estilos de conducción trae en sí muchas aplicaciones, las cuales pueden ser trabajadas en temas cotidianos que muchas veces son pasados por alto. A continuación, se nombran algunas de las aplicaciones que pueden implementadas:

- Tarifación de pólizas de seguros de automóviles basados en estilos de conducción agresivos o pasivos, que incrementan o decrementan el porcentaje de la tarifa a ser liquidada de forma mensual según como condujo el vehículo el mes inmediatamente anterior [13].
- Mejora del consumo de gasolina evidenciando los estilos de conducción que generan un mayor consumo [12].
- Evidencia de los estilos de conducción agresivos en el transporte público, que pongan en riesgo la vida de los pasajeros [11].
- Identificación de personas en estado de alteración, ya sea por estrés, ebriedad o bajo el efecto de sustancias psicoactivas, ya que su estilo de conducción es errático.
- Vehículos autónomos [12].

El proceso de conducir un vehículo es complejo y dinámico donde intervienen 3 elementos: el vehículo, el entorno y el conductor [1], siendo el conductor el único que toma decisiones y realiza acciones. Esta es la razón por la cual los estudios sobre la accidentalidad vial se enfocan en este elemento con el fin de poder disminuirla apuntando a la identificación de los estilos de conducción.

El estilo de conducción de una persona suele ser afectado por los entornos sociales y dependen de factores internos como externos al vehículo. Por ejemplo, de forma empírica sabemos que un mismo conductor tendría un comportamiento en la ciudad de Bogotá y otro muy diferente en una ciudad pequeña como Zipaquirá. Usando el mismo ejemplo, lo que es percibido como un conductor agresivo en Zipaquirá, puede no serlo en un sector donde el tráfico es denso en Bogotá.

Por esta razón, muchos de los estudios relacionados con los estilos de conducción en otras latitudes no pueden ser interpolados de forma directa a nuestro país, lo cual crea la necesidad de realizar estudios que permitan entender las dinámicas del

tráfico vehicular y su relación con los estilos de conducción en nuestro contexto local. Esto incluye el desarrollo de herramientas, conjuntos de datos y conocimiento, que permitan adelantar tales estudios.

Es en este punto donde se centró el desarrollo de esta tesis de grado de maestría. Puntualmente, este trabajo buscó definir un modelo computacional que permitiera reconocer estilos de conducción, usando conjuntos de datos y sensores comunes en nuestro contexto. Por tal razón, se propuso el uso de los sensores presentes en los smartphones, debido a que la gran mayoría de los vehículos en nuestro país no cuenta con sistemas de monitoreo adecuados, muy contrario al uso de los smartphones el cual si se encuentra masificado.

## **CONTRIBUCIONES**

Las contribuciones de este trabajo de maestría se pueden puntualizar en los siguientes aspectos.

En primer lugar, se realizó una actualización del estado del arte mediante un protocolo de búsqueda sistemática, sobre la base datos de SCOPUS. El objetivo fue identificar en la literatura publicaciones donde se utilizaron los smartphones como un dispositivo que capturara señales de los sensores presentes en dicho dispositivo, para así identificar cuáles eran los sensores más utilizados. En este proceso, se identificaron los métodos de aprendizaje de máquina utilizados para clasificar los estilos de conducción. Los resultados del proceso anterior permitieron sometimiento del artículo titulado “*Smartphones, Suitable Tool for Driver Behavior Recognition. A Systematic Review*” al Iberoamerican Workshop on Human-Computer Interaction, HCI Collab 2020, documento que fue aceptado, presentado y publicado en las memorias del evento [14].[14].

El segundo aporte corresponde a un dataset de conducción formado por tres (3) secuencias de conducción, que corresponden aproximadamente a 45 minutos, debidamente etiquetado con 14 eventos de conducción. El dataset fue segmentado en dos subconjuntos de datos, uno orientado a los eventos de la conducción misma (giros, arranque, frenado, entre otros), y el segundo, a la forma de la conducción,

separando una conducción normal, de una conducción agresiva. Se recalca que en la revisión del estado de arte no se identificó un dataset similar en Latinoamérica.

El tercer aporte fue un proceso de experimentación computacional que permitió 1) la caracterización de las señales objetivas y etiquetadas (dataset), tanto estadística como frecuencialmente, usando un proceso de ventana deslizante, 2) la selección del conjunto adecuado de características, 3) la aplicación de seis (6) algoritmos distintos de aprendizaje de máquina para el reconocimiento de los eventos como de los tipos de conducción, y 4) un análisis basado en métricas de calidad de reconocimiento de máquina (especificidad y sensibilidad) para elegir el mejor clasificador y tamaño de la ventana deslizante.

## **OBJETIVOS DE LA PROPUESTA**

Lo que se planteó en el desarrollo de la presente tesis de grado de maestría fue poder identificar estilos de conducción de las personas que conducen vehículos con la ayuda de los sensores presentes en los smartphones.

Para cumplir con el propósito de esta tesis de maestría se definió el siguiente objetivo general:

**Establecer un modelo computacional que permita reconocer estilos de conducción, utilizando los sensores presentes en los teléfonos inteligentes.**

El cual fue apoyado por los siguientes objetivos específicos:

- Identificar los sensores, métodos de procesamientos y enfoques de aprendizaje de máquina, para la identificación de estilos de conducción mediante smartphones, reportados en la literatura.
- Construir un conjunto de datos, con base en las señales proporcionadas por los sensores presentes en los teléfonos inteligentes, que permita la evaluación y ajuste de los enfoques seleccionados para el reconocimiento de estilos de conducción.

- Determinar cuál enfoque de aprendizaje de máquina presenta un mejor desempeño en la identificación de estilos de conducción, mediante esquemas de validación cruzada.

## **ESTRUCTURA DEL DOCUMENTO**

Este documento se realizó con un total de 5 capítulos y 4 anexos. El capítulo 1 presenta la introducción de lo que se encontrará en el texto junto a las contribuciones.

En el capítulo 2 se encuentra la revisión del estado del arte donde se explica el proceso de búsqueda sistémica realizada y los hallazgos del proceso, donde se identifican los sensores de los smartphones que son utilizados en la captura de señales y los métodos de aprendizaje de máquina que se implementaron en los diferentes artículos.

Para el capítulo 3 contiene la estrategia metodológica desarrollada la cual se dividió en 3 fases, donde se buscó la aplicación de captura de datos, se levantaron las señales de los sensores, se generó el dataset y se aplicaron los diferentes clasificadores encontrados en la revisión del estado del arte.

En el capítulo 4 presenta los resultados obtenidos del proceso realizado, donde se identifican los mejores clasificadores con su respectivo mejor parámetro de proceso, la ventana en la que mejor se desempeñó el clasificador y cuál conjunto de datos generó mejores resultados.

El capítulo 5 expone las conclusiones del desarrollo del proyecto y discusión de trabajo futuro y consideraciones finales.

Adicional al documento se presenta el anexo 1 que muestra la configuración del archivo que genera el programa de Python para las características estadísticas; el anexo 2 muestra la configuración del archivo que genera el programa de Python para las características frecuenciales; el anexo 3 muestra toda la tabulación realizada para determinar la mejor venta, el mejor clasificador y los mejores

parámetros para la clasificación y el anexo 4 contiene el artículo científico sometido y expuesto.

## REVISIÓN DEL ESTADO DEL ARTE

En los últimos años, la IoT ha permitido que procesos que anteriormente eran complicados de realizar, hoy sean muy sencillos. Y, ¿qué es la IoT? Según lo explican en [15], la IoT es solo una colección de objetos físicos interconectados, que generan y transmiten información y que pueden estar conectados a internet; en cambio para [16] la IoT tiene una definición más conceptual: una infraestructura de red global y dinámica con capacidad de autoconfiguración sobre protocolos de comunicación estándar e interoperables, en el que las "cosas" físicas y virtuales poseen identidad, atributos físicos y virtuales, utilizan interfaces inteligentes y se integran entre ellas. A través de la IoT se genera información relevante que será procesada y analizada para cumplir un fin específico.

Para algunas consultoras, en los próximos años existirán más dispositivos conectados a internet que personas, intercambiando información relevante en procesos distribuidos y colaborativos [17][15].[17][15].[17][15].

La IoT está generando un gran impacto en la sociedad, el medioambiente y la economía, agregando valor en una gran cantidad de ámbitos que van desde la logística, el transporte y el seguimiento de activos, hasta los entornos inteligentes (hogares, edificios, infraestructuras), pasando por la energía, la defensa y la agricultura como lo es expresado en [18]. En esencia, la IoT tiene el potencial de influir significativamente en todas las facetas de la sociedad.

La visión de la IoT es ir más allá de los smartphones y portátiles como lo dicen en [19], la cual ha generado una reestructuración de los diferentes dispositivos inteligentes, las arquitecturas de software y las arquitecturas de hardware como en las redes de telecomunicaciones, con solo una finalidad, recolectar el mayor número de datos sin la intervención de personas.

No cabe duda de que la industria automotriz ha desarrollado grandes investigaciones buscando reducir las muertes en accidentes de tránsito que han sido impulsadas por los grandes fabricantes de vehículos como se menciona en [5], dando origen a los sistemas avanzados de asistencia a la conducción (ADAS).



También se hace un breve recuento sobre la historia de las tecnologías implementadas para la seguridad en la conducción, la cual se divide en 2 etapas: La primera etapa comprende entre los años 1950 y 2000, mencionando que en 1958 aparecen los primeros sistemas automáticos como el control de velocidad, la bolsa de aire que aparece en 1973 y el control de estabilización electrónica que aparece en 1987. La segunda etapa se desarrolla desde el año 2000 y es cuando se desarrollan los ADAS, el cual han permitido que las cifras de muertes por accidentes de tránsito hayan disminuido.[5], dando origen a los sistemas avanzados de asistencia a la conducción (ADAS). También se hace un breve recuento sobre la historia de las tecnologías implementadas para la seguridad en la conducción, la cual se divide en 2 etapas: La primera etapa comprende entre los años 1950 y 2000, mencionando que en 1958 aparecen los primeros sistemas automáticos como el control de velocidad, la bolsa de aire que aparece en 1973 y el control de estabilización electrónica que aparece en 1987. La segunda etapa se desarrolla desde el año 2000 y es cuando se desarrollan los ADAS, el cual han permitido que las cifras de muertes por accidentes de tránsito hayan disminuido.

En los últimos años se viene usando la IoT con el fin de identificar estilos de conducción y así ayudar a disminuir la accidentalidad vial, tema que está siendo estudiado ampliamente a nivel mundial y el cual ha permitido identificar las principales razones por las cuales se presentan accidentes de tránsito.

Según [1], al conductor se le imputa el 90% de los accidentes de tránsito que se presentan en el mundo; a la distracción se le atribuye más de la mitad de los accidentes de tránsito siendo el primer factor de accidentalidad vial; la fatiga es el segundo factor de accidentalidad vial y representa entre el 25% y el 35% de estos; el tercer factor generador de accidentes de tránsito es la conducción agresiva.

La distracción de los conductores está asociada a aquellas actividades que generan que el conductor de un vehículo deje de ver el entorno de conducción y por consiguiente pierda la atención de lo que está realizando.

Con relación a la fatiga, en [20] comentan de los estudios que han realizado y que buscan determinar los diferentes factores que permiten identificarla y entre los

cuales tenemos: somnolencia, agotamiento, letargo y apatía. Estos describen un cansancio y debilidad ya sea físico o mental.

En [12] hablan de que los estilos de conducción son influenciados por muchos factores que están divididos en factores humanos y factores medioambientales. Entre los factores humanos tenemos: los sentimientos, la edad, el género, la personalidad, la actitud, el riesgo que se asume, la capacidad de agresión, la procedencia demográfica, la toma de decisiones, la familiaridad con el vehículo, la experiencia y formación de conducción y ciertas condiciones como drogadicción y alcoholismo entre otras; entre los factores medioambientales tenemos: el tráfico, la estación, el clima, el tipo de vía, los otros conductores, la visibilidad, el estado del camino, el momento del día, los peatones, los ciclistas y las condiciones lumínicas, entre otros. Todos estos factores se esquematizan en la Figura 1.



Figura 1. Factores que influyen el estilo de conducción.

En [1] nos cuentan sobre los sistemas de control de conducción y asistencia (DMAS), cuyo propósito es mejorar la experiencia de conducción con la ayuda de múltiples sensores que miden parámetros del conductor, el vehículo y el entorno, buscando una conducción segura y cómoda. El DMAS permite vigilar los ojos, los brazos, las piernas o cualquier movimiento del cuerpo que permita identificar la

fatiga, la distracción o la conducción agresiva. Existen DMAS centrados en el conductor y centrados en el vehículo.

En [12] hablan del papel de los sensores tales como la unidad de medida inercial, el acelerómetro, el sistema de posicionamiento global (GPS), el sistema de información geográfica (SIG), el sensor de radar y los smartphones entre otros, que son utilizados para reconocer los estilos de conducción de las personas. Los vehículos más recientes vienen con algunos de estos sensores ya incluidos en sus especificaciones básicas. Estos sensores permiten capturar eventos específicos tales como: frenado, seguimiento del vehículo, mantenimiento de distancias, rotondas, manejo y el retardo en la reacción a los eventos.

Según [12], los estilos de conducción pueden ser capturados por señales que están estrechamente relacionadas con los criterios de clasificación tales como: clases directas e indexación continua; y la implementación de algoritmos de reconocimiento en los cuales tenemos: implementación a través de reglas, implementación a través de modelos y a través de algoritmos de aprendizaje.

En [12], cuentan que los algoritmos basados en reglas son el enfoque más sencillo para el reconocimiento de estilos de conducción ya que se definen umbrales para las variables que van a hacer monitorizadas, las cuales se definen para eventos específicos. También cuentan que los algoritmos basados en modelos determinan el estilo de conducción definiendo un conjunto de ecuaciones de características predefinidas; hay que tener en cuenta que cada modelo tiene un estilo de conducción asociado. Estos algoritmos pueden ser supervisados, no supervisados o la mezcla de estos dos.

Adicionalmente, resultados parciales de una búsqueda sobre sistemas en desarrollo, que están indagando sobre las técnicas, métodos y sensores usados en el reconocimiento de estilos de conducción, muestra que de 805 trabajos encontrados en SCOPUS, ninguno tiene como origen Latinoamérica (Ver *Figura 2*), lo cual evidencia la necesidad de desarrollar trabajos orientados a las particularidades propias de nuestra región con respecto a parque automotor, características de nuestras ciudades y de los conductores.

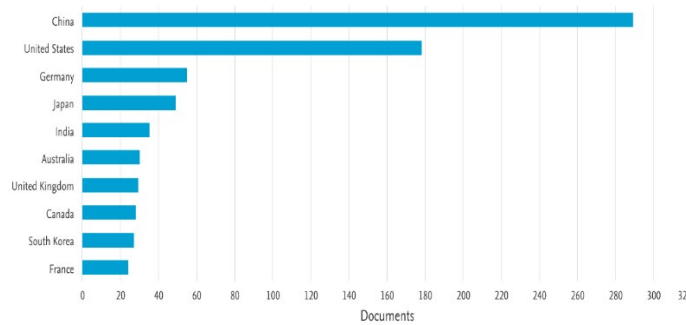


Figura 2. Resultados de una búsqueda sobre sistemas en desarrollo que están indagando sobre técnicas, métodos y sensores usados en el reconocimiento de estilos de conducción.

## BÚSQUEDA SISTÉMICA

Se desarrolló una revisión sistemática que se dividió en una serie de etapas, las cuales se describen en la Figura 3. En la primera etapa se establecieron las preguntas de investigación que permitiera establecer los parámetros del proceso de la búsqueda; la segunda etapa tuvo como objetivo construir la ecuación de búsqueda con la cual se identificarían los artículos a ser trabajados; en la tercera etapa se ejecutó la ecuación de búsqueda en la base de datos de SCOPUS; con el fin de refinar la ecuación de búsqueda y mejorar el resultado de la búsqueda, la etapa 2 y la etapa 3 fueron ejecutadas en varias ocasiones; en la cuarta etapa se aplicaron criterios de exclusión y en la quinta etapa se realizó la extracción de la información.

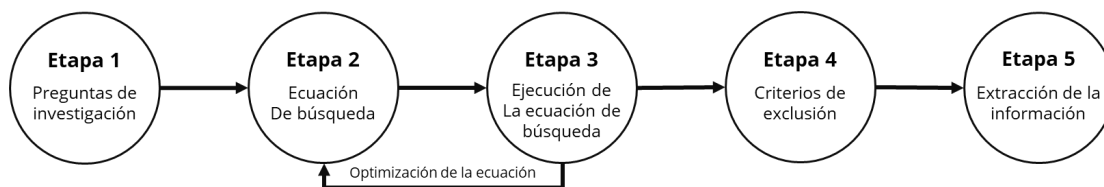


Figura 3. Etapas de la revisión sistemática.

**ETAPA 1.** Como resultado de la ejecución de esta etapa, se plantearon las siguientes preguntas de investigación:

- ¿Cuáles han sido los estudios realizados que identifican los estilos de conducción de vehículos usando los smartphones?

- ¿Existe algún estudio publicado que desarrolle la identificación de estilos de conducción de vehículos usando los smartphones en Colombia y Latinoamérica?
- ¿Cuáles han sido los sensores de los smartphones que han sido utilizados como apoyo para la identificación de estilos de conducción de vehículos?
- ¿Qué métodos de análisis de datos han sido utilizados como apoyo en la identificación de estilos de conducción de vehículos?

**ETAPA 2.** En esta etapa se identificaron las siguientes etiquetas con el fin de obtener la ecuación de búsqueda:

- Etiquetas para los dispositivos móviles
  - Mobile
  - Cellphone
  - Smartphone
- Etiquetas para los estilos de conducción
  - Driving pattern
  - Driving mode
  - Driving style
  - Driving comportment
  - Driving bearing
  - Driving kind
  - Driving type
- Etiquetas de personalización de la búsqueda
  - Recognition
  - Classification
  - Characterization
  - Identification
  - Detection
  - Learning

Por último, solo se tuvieron en cuenta artículos científicos publicados en los años 2020, 2019, 2018 y 2017 y que fueran de los tipos artículo de revista y artículo en conferencia.

**ETAPA 3.** La primera ejecución con la ecuación de búsqueda ya optimizada se realizó el 12 de febrero de 2020, obteniendo como resultado un total de 64 artículos; la distribución de los artículos por año se muestra en la Tabla 3.

Tabla 3. Artículos por año primera ejecución

Año de publicación	Cantidad de artículos
2019	32
2018	15
2017	17

*Fuente: Elaboración propia a partir del resultado de SCOPUS.*

A este resultado se le aplicó la ejecución de la etapa 4, sin embargo, con el fin de mejorar los resultados de la primera búsqueda, se realizó una segunda ejecución de la ecuación de búsqueda el 5 de julio de 2020 donde se obtuvieron un total de 73 artículos; la distribución de los artículos por año se muestra en la Tabla 4.

Tabla 4. Artículos por año segunda ejecución.

Año de publicación	Cantidad de artículos
2020	7
2019	34
2018	15
2017	17

*Fuente: Elaboración propia a partir del resultado de SCOPUS.*

**ETAPA 4.** Los 73 artículos científicos obtenidos se sometieron a un proceso de selección aplicando los siguientes criterios de exclusión:

- Exclusión por artículos repetidos,
- Exclusión por abstract
- Exclusión por introducción.

El primer criterio de exclusión eliminó de la lista 3 artículos científicos que equivalen al 4% del total de artículos obtenidos; este proceso dejó 70 artículos científicos por seguir trabajando.

El segundo criterio de exclusión eliminó de la lista un total de 25 artículos científicos que equivalen al 34% del total de artículos obtenidos; este proceso dejó 45 artículos por seguir trabajando.

El tercer criterio de exclusión eliminó de la lista un total de 20 artículos científicos que equivalen al 27% del total de artículos obtenidos; al final de todo el proceso de exclusión quedaron un total de 25 artículos científicos a ser trabajados. El resumen del proceso de exclusión se puede observar en la Tabla 5.

Tabla 5. Criterios de exclusión de artículos

Motivo de exclusión	Artículos excluidos	Porcentaje de artículos
Artículos repetidos	3	4%
Por abstract	25	34%
Por introducción	20	27%

*Fuente: Elaboración propia.*

En esta etapa también se realizó una evaluación de la calidad de los artículos científicos con el fin de determinar si generaban valor al proceso; esta evaluación condujo a la exclusión de 3 artículos adicionales debido a que se salían del contexto de las preguntas de investigación; al final, para la etapa 5 quedaron 22 artículos científicos a ser trabajados.

**ETAPA 5.** En esta etapa se identificaron los siguientes puntos:

- Sensores de los smartphones que funcionan para la identificación de estilos de conducción de vehículos.
- Eventos que son capturados para determinar estilos de conducción de vehículos.
- Métodos de aprendizaje de máquina que fueron implementados para la identificación de estilos de conducción de automóviles.

- Estilos de conducción que fueron identificados en los diferentes artículos científicos.

## HALLAZGOS

En la Tabla 6 se muestra el resultado de la cantidad de veces que fueron utilizados los sensores en los diferentes artículos científicos, y en la cual se pueden evidenciar que los más utilizados son el acelerómetro, el giroscopio, el GPS y el magnetómetro; los otros sensores tienen un menor uso debido a que cumplen funciones más específicas.

Tabla 6. Uso de los sensores de los smartphones.

Sensor	Cantidad de artículos que lo implementaron	% de artículos
Acelerómetro	19	86%
Giroscopio	18	82%
GPS	13	59%
Magnetómetro	11	50%
Cámara	5	23%
Micrófono	1	5%
Sensor de rotación	1	5%
Sensor de luz	1	5%

*Fuente: Elaboración propia a partir del resultado del análisis de la extracción de la información.*

La distribución de los artículos científicos por cada uno de los sensores trabajados se puede visualizar en la Tabla 7.

Tabla 7. Artículos científicos por sensor.

Sensor	Artículos científicos
Acelerómetro	[21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39]
Giroscopio	[40], [41], [42], [24], [25], [43], [26], [27], [28], [29], [31], [32], [33], [34], [44], [37], [45], [46]
GPS	[41], [42], [25], [47], [27], [28], [29], [30], [33], [34], [37], [45], [46]
Magnetómetro	[40], [41], [24], [25], [27], [28], [30], [33], [44], [45], [46]
Cámara	[42], [48], [32], [44], [46]
Micrófono	[42]
Sensor de rotación	[44]



*Fuente: Elaboración propia a partir del resultado del análisis de la extracción de la información.*

En la Tabla 8 se muestran los diferentes eventos que son capturados en los diferentes artículos científicos.

Tabla 8. Eventos capturados por los smartphones.

<b>Evento</b>	<b>Número de veces que fue capturado en artículos</b>	<b>Porcentaje de artículos</b>
Aceleración	15	68%
Frenado	15	68%
Giros	10	45%
Cambio de carril	6	27%
Exceso de velocidad	5	23%
Velocidad	5	23%
Uso del smartphone	5	23%
Bostezos	3	14%
Cierre de los ojos	3	14%
Movimientos de la cabeza	3	14%
Parpadeo	2	9%
Posición del vehículo en la vía	2	9%
Otros eventos	9	41%

*Fuente: Elaboración propia a partir del resultado del análisis de la extracción de la información.*

En el ítem de otros eventos podemos encontrar todos aquellos que fueron mencionados una sola vez, como, por ejemplo: ruido en la cabina, distancia recorrida, postura, rasgos faciales, subidas y bajadas inclinadas entre otros.

En la Tabla 9 se evidencian los diferentes métodos de aprendizaje de máquina que fueron implementados para la identificación de estilos de conducción de vehículos.

Tabla 9. Métodos de aprendizaje de máquina implementados.

<b>Método de aprendizaje de máquina</b>	<b>Número de veces que fue implementado</b>	<b>Porcentaje de artículos</b>
SVM	7	32%
Cross validation	6	27%
Random forest	5	23%
C4.5	4	18%

Naïve Bayes	4	18%
K-means	4	18%
Red neuronal de perceptrón de múltiples capas	4	18%
BayesNet	3	14%
Inferencia difusa	2	9%
Red neuronal recurrente	2	9%
J48	2	9%
Red neuronal convulsionada	2	9%
Otros métodos de aprendizaje de máquina	14	64%

*Fuente: Elaboración propia a partir del resultado del análisis de la extracción de la información.*

En el ítem de otros métodos de aprendizaje de máquina se encuentran todos aquellos que fueron nombrados una sola vez y en los cuales podemos encontrar: red neuronal profunda, red neuronal artificial, red neuronal de fusión de base radial, gradient boot tree, entre otros.

En la Tabla 10 se puede observar la distribución de los artículos científicos por cada uno de los métodos de aprendizaje de máquina.

Tabla 10. Métodos de aprendizaje por artículo científico.

<b>Método de aprendizaje de máquina</b>	<b>Artículos científicos</b>
Random forest	[41], [43], [26], [31], [45]
C4.5	[41], [30], [31], [37]
Naïve Bayes	[41], [26], [30], [31]
BayesNet	[41], [26], [44]
K-means	[41], [29], [30], [44]
SVM	[41], [30], [31], [32], [35], [44], [45]
Red neuronal de perceptrón de múltiples capas	[24], [26], [30], [37]
Inferencia difusa	[24], [30]
Red neuronal recurrente	[44], [45]
J48	[47], [26]
Red neuronal convulsionada	[43], [44]
Otros eventos de aprendizaje de máquina	[47], [43], [27], [30], [32], [33], [37], [45]

*Fuente: Elaboración propia a partir del resultado del análisis de la extracción de la información.*

Otra de las acciones implementadas en los artículos [41], [26], [30], [31], [35], [37] fue la evaluación mediante validación cruzada.[40], [26], [30], [31], [35], [37] fue la validación cruzada.[40], [26], [30], [31], [35], [37] fue la validación cruzada.[40], [26], [30], [31], [35], [37] fue la validación cruzada.

En la Tabla 11 se muestran los estilos de conducción que fueron identificados en los diferentes artículos científicos.

Tabla 11. Estilos de conducción identificados.

<b>Estilos de conducción</b>	<b>Número de veces que fue identificado</b>	<b>Porcentaje de artículos</b>
Distracción en la conducción	11	50%
Conducción agresiva	9	41%
Conducción anormal	5	23%
Conducción normal	4	18%
Otros estilos de conducción	3	14%

*Fuente: Elaboración propia a partir del resultado del análisis de la extracción de la información.*

En el ítem de otros estilos de conducción, están todos aquellos que fueron nombrados una sola vez por los diferentes autores de los artículos científicos.

Al final de todo el proceso se pudo establecer que son muchas y muy variadas las formas de trabajo que son implementadas para la identificación de estilos de conducción de vehículos, y que estas dependen del objetivo de cada investigación.

## ESTRATEGIA METODOLÓGICA

Para el desarrollo de esta tesis de grado de maestría se realizaron tres (3) fases, orientadas al cumplimiento de cada uno los objetivos específicos.

### FASE 1

La primera fase tenía como finalidad la identificación de los diferentes sensores presentes en los smartphones que son usados para capturar información y se buscó determinar los modelos computacionales y algoritmos de aprendizaje de máquina usados en el estado del arte. Para esto se realizó una búsqueda sistémica en la base de datos SCOPUS. Todos los detalles del proceso están explicados en la sección *BÚSQUEDA SISTÉMICA*, del capítulo *REVISIÓN DEL ESTADO DEL ARTE*.

Es necesario mencionar nuevamente que los sensores más utilizados en el estado del arte son el acelerómetro, el giroscopio, el GPS y el magnetómetro. Se decidió usar únicamente el acelerómetro y el giroscopio, debido a que son los más usados en la literatura y tienen un menor consumo de batería, lo que facilita el desarrollo de ampliación para el usuario final.

Otro resultado de la búsqueda sistémica fue la identificación de los métodos de aprendizaje de máquina más usados, lo que permitió elegir los siguientes clasificadores:

- K-Means
- Máquinas de soporte vectorial (SVM)
- Redes neuronales de perceptrón multicapa (RNPM)
- Random forest
- Árbol de decisión J48
- Naïve Bayes

Un hallazgo importante que se pudo determinar es un método de reconocimiento de estilos de conducción, basado en eventos de conducción. Entre los principales eventos usados como indicadores de estilos de conducción se encontraron los giros

(derecha, izquierda, en U), la aceleración, el frenado y los cambios de carril. En cuanto al reconocimiento de un grupo de estilo de conducción particular no hay un consenso, dado que, cómo se ha comentado anteriormente, se puede interpretar desde muchos aspectos socioculturales. Sin embargo, lo que sí es común en todos los casos es la necesidad de identificar un estilo de conducción agresivo, debido al impacto que tiene en la seguridad vial.

## **FASE 2**

En la segunda fase se construyó un dataset ajustado al enfoque diseñado que permitiera identificar los estilos de conducción.

Para el desarrollo de esta fase, se buscaron aplicaciones en las tiendas de aplicaciones de dispositivos con sistema operativo Android (Google Play) y con sistema operativo IOS (App Store) que permitieran la captura de las señales de los sensores de los teléfonos celulares; las pruebas iniciales para la selección de las dos (2) aplicaciones candidatas para el proceso se realizaron en una mesa, donde se activaba cada aplicación y se generaban varios movimientos al celular y después de desactivar la aplicación se analizaban los datos generados, su facilidad de uso y de ser compartidos.

Teniendo como base el proceso realizado, se escogieron dos (2) aplicaciones para realizar las primeras evaluaciones de captura de señales en el proceso de conducción de un vehículo. La primera aplicación evaluada fue AndroSensor que funciona sobre sistema operativo Android, la aplicación captura las señales de los sensores del acelerómetro, giroscopio, magnetómetro y GPS. Las desventajas de esta aplicación son la imposibilidad de captura voz dentro de la misma aplicación durante el experimento, lo que obliga a usar una aplicación adicional para realizar esta labor lo que dificultaba la sincronización de la captura de las señales con la descripción del evento ejecutado; adicionalmente, el proceso para compartir los resultados del experimento es un poco tedioso. La segunda aplicación evaluada fue Sensor Data Recorder (Ver Figura 4). que funciona sobre sistema operativo IOS. Esta aplicación captura señales de los sensores del acelerómetro y del giroscopio y

tiene una ventaja con respecto a la anterior y es que permite la captura de voz durante la ejecución del experimento, fuera de que permite compartir los resultados de una forma sencilla.

Después de las validaciones realizadas con las dos (2) aplicaciones y de analizar las ventajas y desventajas que tenían cada una de ellas en su uso, se decidió trabajar con la aplicación Sensor Data Recorder como aplicación para captura de señales de los sensores.

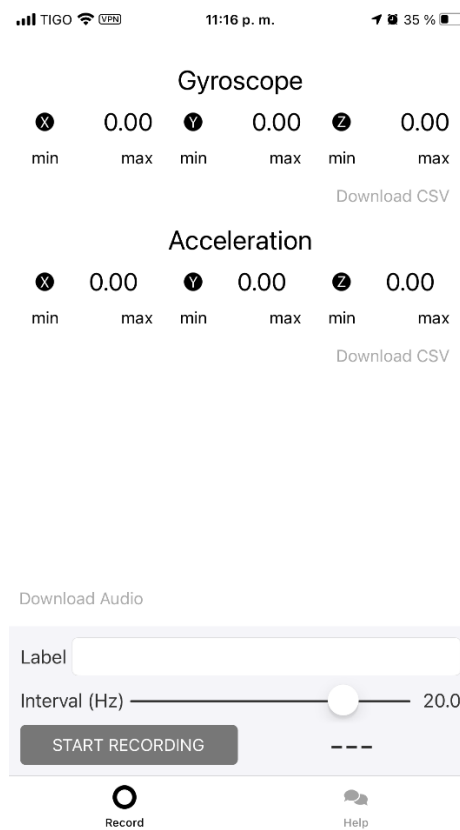


Figura 4. Sensor Data Recorder

El smartphone utilizado para el proceso de captura de señales fue un iPhone 6S Plus, el cual cuenta con múltiples sensores y que cumple con los parámetros necesarios para la ejecución de los experimentos; al smartphone se le instaló la aplicación anteriormente indicada y fue el único utilizado para todo el. El vehículo utilizado para esta labor fue un Renault Logan de transmisión mecánica.

Para el proceso de captura de señales del smartphone fue necesario tipificar los eventos que iban a ser capturados de acuerdo con la búsqueda sistémica realizada, lo que permitiría la identificación de los estilos de conducción; estos eventos fueron:

- Giro a la derecha.
- Giro a la derecha prolongado.
- Giro a la izquierda.
- Giro a la izquierda prolongado.
- Giro en U.
- Cambio de carril a la derecha.
- Cambio de carril a la derecha agresivo.
- Cambio de carril a la izquierda.
- Cambio de carril a la izquierda agresivo.
- Arranque.
- Detención.
- Detención brusca
- Vehículo en movimiento.
- Vehículo detenido.

Se realizaron 3 recorridos entre 10 y 20 minutos cada uno en el sector residencial de Gran Granada de la localidad de Engativá en la ciudad de Bogotá D.C. (Ver mapa en la Figura 5 de las cuadras donde se realizó el recorrido)

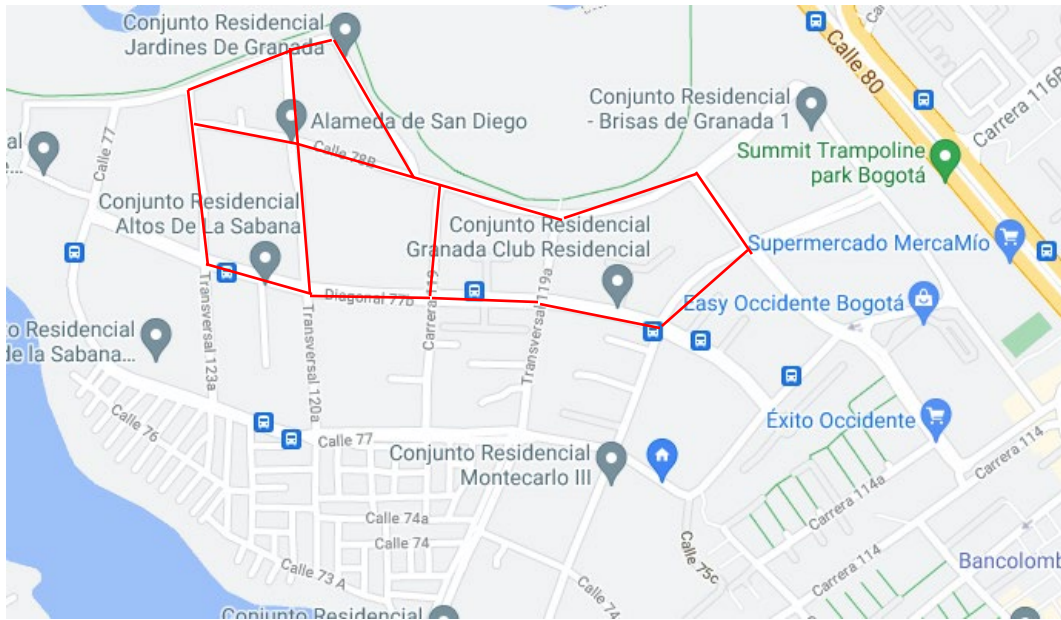


Figura 5. Mapa Gran Granada - Bogotá D. C.

En cada uno de los recorridos realizados se ejecutaron en numerosas ocasiones los eventos anteriormente mencionados, acompañando cada evento con la grabación de la voz de cuando iniciaba y cuando terminaba el evento.

Para trabajar con los datos de las señales capturadas, se unificaron los datos del acelerómetro y del giroscopio en una sola hoja de trabajo ya que la aplicación genera los datos en archivos independientes, para realizar esta labor se utilizó Microsoft Excel, relacionando los datos de las señales por el valor del campo Timestamp.

El proceso anterior permitió identificar registros que no tenían un dato correlacionado, para lo cual se decidió eliminar esos registros y así poder tener registros pares por cada uno de los sensores. A continuación, se etiquetó el evento a cada uno de los registros restantes teniendo como base el audio de la captura. Hay que tener en cuenta que esta asignación se hizo de forma manual y se basó en aproximaciones de tiempo e interpretación del audio.

Para tener una asignación de datos más limpia de los eventos de conducción, se decidió eliminar los primeros cinco (5) registros y los últimos registros, esto se realizó con el fin de eliminar registros del evento anterior o posterior que pudieron quedar mal asignado debido a incertidumbre normal en el proceso.



Para facilitar todo el trabajo, los eventos fueron divididos en diferentes hojas y almacenado en Microsoft Excel y en CSV por compatibilidad.

Además de los datos crudos (raw) de los sensores, se agregó al dataset características estadísticas y con características frecuenciales de la señal. El dataset se dividió señales de estudio usando un esquema de ventana deslizante. El proceso se realizó generó para cuatro (4) tamaños diferentes de ventana los cuales fueron: 10 registros, 15 registros, 20 registros y 25 registro. Teniendo en cuenta que 10 registros equivalen a 1000 milisegundo, en términos de tiempo fueron cada ventana tuvo un tamaño 1000, 1500, 2000 y 2500 milisegundos. En todos los casos se empleó un desplazamiento constante de la ventana de 3 registros (300 milisegundo). El dataset fue generado con una serie de scripts desarrollados en lenguaje Python 3.9 y ejecutados en la herramienta Google Colab.

Posteriormente, se realizó un proceso de agrupación y disminución de etiquetas, dirigido a facilitar el proceso de reconocimiento tanto de los eventos de conducción, como del tipo de conducción. De esa forma se generó dos subconjuntos de datos, uno orientado a los eventos de la conducción misma, y el segundo, a la forma de la conducción, separando una conducción normal, de una conducción agresiva. El proceso de agrupación y disminución de etiquetas se puede observar en la Tabla 12.

Tabla 12. Reducción de etiquetado de los eventos capturados.

Etiqueta Original	Agrupación	
	Eventos	Tipo de conducción
Giro a la derecha	Giro a la derecha	
Giro a la derecha prolongado		
Giro a la izquierda	Giro a la izquierda	
Giro a la izquierda prolongado		
Giro en U	Giro en U	
Cambio de carril a la derecha	Cambio de carril a la derecha	Normal
Cambio de carril a la derecha agresivo		Agresivo
Cambio de carril a la izquierda	Cambio de carril a la izquierda	Normal

Cambio de carril a la izquierda agresivo		Agresivo
Arranque	Arranque	
Detención	Detención	Normal
Detención brusca		Agresivo
Vehículo en movimiento	Vehículo en movimiento	
Vehículo detenido	Vehículo detenido	

*Fuente: Elaboración propia a partir del proceso ejecutado.*

Luego, con el fin de tener registros balanceados en el dataset generado, se eliminaron registros del evento movimiento y del estilo de conducción normal desde Excel, aplicando una fórmula que permitiera eliminar siempre los mismos registros del dataset.

### **FASE 3**

La tercera fase evaluó y ajustó los métodos de aprendizaje de máquina escogidos. se utilizó la aplicación Weka versión 3.8.5, la cual fue resultado de una sucinta revisión basada en pruebas rápidas de viabilidad técnica, facilidad de uso, métricas de evaluación disponibles y la experiencia previa.

En la primera parte fue necesario determinar cuáles atributos generaban la mejor clasificación, para esto se utilizó la aplicación Weka versión 3.8.5 y en la sección de selección de atributos se escogió el evaluador de atributos aplicando el algoritmo RELIEF con el método de búsqueda Ranker, cómo modo de selección de atributos se escogió una validación cruzada de 10 paquetes (10-fold Crossvalidation).

Para el proceso de clasificación también se utilizó la aplicación Weka versión 3.8.5. Este proceso se realizó mediante scripts que se ejecutaron por la línea de comandos y cuyos resultados fueron guardados en archivos de texto. Teniendo claros los clasificadores a trabajar y los atributos del dataset, se seleccionaron los parámetros de ejecución para cada clasificador. En la Tabla 12 se muestran las configuraciones de ejecución por cada uno de ellos.

Tabla 13. Configuración de clasificadores.

<b>Clasificador</b>	<b>Parámetros de ejecución</b>
---------------------	--------------------------------

K Means	Se varió el parámetro de K donde tomó el valor de: 1, 3, 5, 7, 9 y 11
Máquinas de soporte vectorial	Se varió el parámetro del kernel donde tomó el valor de: 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5 y 5.0.
Redes neuronales de perceptrón múltiple	Se varió el parámetro de capas ocultas donde tomó el valor de: 1, 2, 3, 4, 5, 6 y 7.
Random forest	Se varió el parámetro del número de iteraciones donde tomó el valor de: 50, 60, 70, 80, 90 y 100.
Árbol de decisión J48	Se varió el parámetro del factor de confianza donde tomó el valor de: 0.5, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45 y 0.50.
Naïve Bayes	Se ejecutó el modelo estándar.

*Fuente: Elaboración propia a partir del proceso generado.*

Todos los clasificadores se evaluaron con na validación cruzada de 10 paquetes (10-fold Crossvalidation). Una vez generado el dataset a trabajar, se ejecutaron los scripts de los clasificadores y se exploraron los resultados que se almacenaron en archivos de texto, se disidió trabajar con las siguientes columnas de los resultados:

- Sensibilidad (Recall): Predicciones positivas correctas entre el número total de positivos.
- Especificidad (FP Rate): Predicciones negativas correctas entre el número total de negativos.

Teniendo claro los parámetros para tener en cuenta en el análisis de los resultados de los clasificadores, se procedió a tabular y validar los resultados obtenidos del procesamiento del dataset.

## RESULTADOS

### SELECCIÓN DE CARACTERÍSTICAS

Para el proceso de selección de características ejecutado en WEKA se determinó tomar aquellas características con un *average merit* mayor a 0.05 y cuya característica está presente en 2 o más ventanas para el mismo subconjunto de datos.

Para el subconjunto de datos de eventos de las características estadísticas se seleccionaron las siguientes características:

- AX-M1 (Acelerómetro eje X, Momento de orden 1)
- AX-M2 (Acelerómetro eje X, Momento de orden 2)
- AY-M1 (Acelerómetro eje Y, Momento de orden 1)
- AZ-M1 (Acelerómetro eje Z, Momento de orden 1)
- GX-M1 (Giroscopio eje X, Momento de orden 1)
- GX-M2 (Giroscopio eje X, Momento de orden 2)
- GY-M1 (Giroscopio eje Y, Momento de orden 1)
- GZ-M1 (Giroscopio eje Z, Momento de orden 1)
- GZ-M2 (Giroscopio eje Z, Momento de orden 2)

Para el subconjunto de datos de eventos de las características frecuenciales se seleccionaron las siguientes características:

- GZ-D1 (Giroscopio eje Z, Pico 1)
- GZ-D2 (Giroscopio eje Z, Pico 2)
- GZ-D3 (Giroscopio eje Z, Pico 3)
- GZ-D4 (Giroscopio eje Z, Pico 4)
- GZ-D5 (Giroscopio eje Z, Pico 5)

Para el subconjunto de datos de tipos de conducción de las características estadísticas se seleccionaron las siguientes características:

- AX-M1 (Acelerómetro eje X, Momento de orden 1)

- AX-M2 (Acelerómetro eje X, Momento de orden 2)
- AY-M1 (Acelerómetro eje Y, Momento de orden 1)
- AZ-M1 (Acelerómetro eje Z, Momento de orden 1)
- GX-M1 (Giroscopio eje X, Momento de orden 1)

Para el subconjunto de datos de tipos de conducción de las características frecuenciales se seleccionaron las siguientes características:

- AX-D4 (Acelerómetro eje Z, Pico 4)
- AX-D5 (Acelerómetro eje Z, Pico 5)
- AZ-D3 (Acelerómetro eje Z, Pico 3)
- AZ-D4 (Acelerómetro eje Z, Pico 4)
- AZ-D5 (Acelerómetro eje Z, Pico 5)

Mediante Excel se ajustó el dataset eliminando las características excluidas tanto del subconjunto de eventos como del subconjunto de tipos de conducción.

## **RECONOCIMIENTO DE EVENTOS DE CONDUCCIÓN.**

Los resultados de la ejecución de los clasificadores al dataset que tuvo un enfoque en los datos de eventos fueron analizados y tabulados de tal forma que permitieran la identificación de cual ventana fue la de mejor resultado y parámetro del algoritmo de clasificación. La tabulación de los datos se puede observar en el Anexo 1. A continuación, se muestran los resultados por cada uno de los clasificadores. Se aclara que inicialmente se evaluó el rendimiento usando un conjunto de carteristas mixtas, frecuenciales y estadísticas. Debido a que los resultados tuvieron el rendimiento deseado, se probó eliminando las características frecuenciales, obteniendo excelentes resultados. Se hizo lo mismo, eliminado las características estadísticas, pero los resultados fueron muy bajos, razón por la cual no se incluyeron.

## **EVALUACIÓN CON CARACTERÍSTICAS MIXTAS**

Los resultados de la ejecución de K-Means muestran que la mejor ventana de procesamiento es la de 1000 milisegundos para K igual a 11 dando una sensibilidad

de 0.337 y una especificidad de 0.225. Para SVM los resultados muestran que la mejor ventana de procesamiento es la de 1000 milisegundo para un kernel igual a 30 y 35 dando una sensibilidad de 0.383 y una especificidad de 0.222. El mejor resultado que se visualiza en la red neuronal de perceptrón múltiple está sobre la ventana de los 1000 milisegundos para 4 capas con una sensibilidad de 0.376 y una especificidad de 0.227. La ventana que genera el mejor resultado para random forest es la de 1000 milisegundos para 90 y 100 iteraciones, generando una sensibilidad de 0.330 y una especificidad de 0.213. La ventana de 1500 milisegundos en la que genera el mejor resultado para J48 con factor de confianza igual a 0.05, obteniendo una sensibilidad de 0.395 y una especificidad de 0.222. Para Naïve Bayes se presenta como mejor ventana la de 2500 milisegundos generando una sensibilidad de 0.359 y una especificidad de 0.240.

Una vez determinados los mejores procesos de clasificación de cada uno de los clasificadores trabajados, se escogió el mejor resultado para este tipo de datos, el cual se soporta en la Tabla 16 donde se visualizan los valores de especificidad y sensibilidad por clasificador y en la Figura 7 se muestra la gráfica de los puntos de cada clasificador.

Tabla 14. Mejores resultados de sensibilidad para datasets de características mixtas del subconjunto de eventos.

<b>Clasificador</b>	<b>1- Especificidad</b>	<b>Sensibilidad</b>
K-Means	0.225	0.337
SVM	0.222	0.383
RNPM	0.227	0.376
RF	0.213	0.330
J48	0.222	0.395
NB	0.240	0.359

*Fuente: Elaboración propia a partir del proceso generado.*

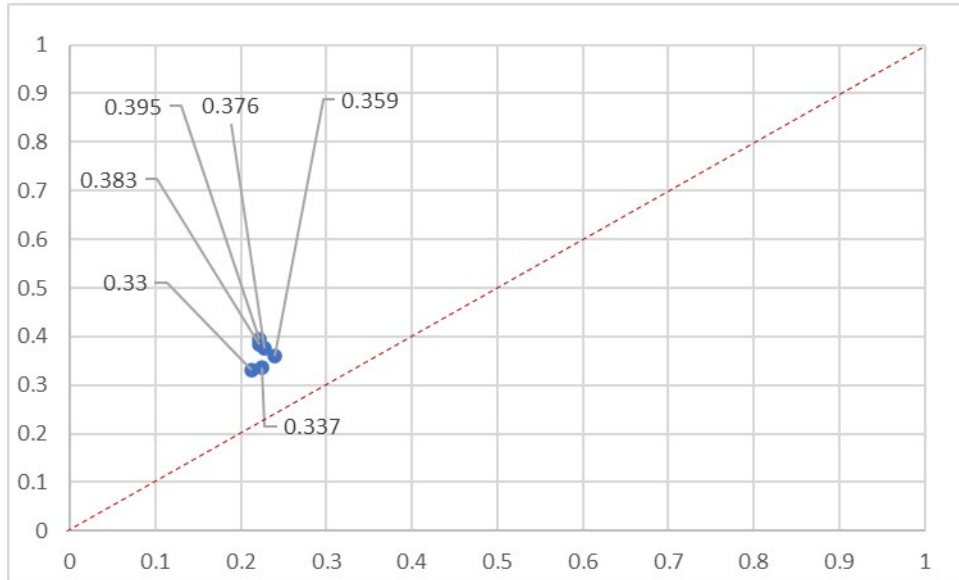


Figura 6. Gráfica de puntos en el espacio ROC mejores resultados datasets de características mixtas subconjunto de eventos.

## EVALUACIÓN CON CARACTERÍSTICAS ESTADÍSTICAS

Los resultados de la ejecución de K-Means muestran que la mejor ventana de procesamiento es la de 2500 milisegundos para K igual a 1 dando una sensibilidad de 0.899 y una especificidad de 0.022. Para SVM los resultados muestran que la mejor ventana de procesamiento es la de 2500 milisegundo para un kernel igual a 50 dando una sensibilidad de 0.749 y una especificidad de 0.085. El mejor resultado que se visualiza en la red neuronal de perceptrón múltiple está sobre la ventana de los 2500 milisegundos para 6 capas con una sensibilidad de 0.758 y una especificidad de 0.064. La ventana que genera el mejor resultado para random forest es la de 2500 milisegundos para 60, 70 y 90 iteraciones, generando una sensibilidad de 0.918 y una especificidad de 0.023. La ventana de 2500 milisegundos en la que genera el mejor resultado para J48 con los factores de confianza de 0.20, 0.30, 0.35, 0.40 y 0.45, obteniendo una sensibilidad de 0.861 y una especificidad de 0.035. Para Naïve Bayes se presenta como mejor ventana la de 2500 milisegundo generando una sensibilidad de 0.409 y una especificidad de 0.094. Una vez determinados los mejores procesos de clasificación de cada uno de los clasificadores trabajados, se escogió el mejor clasificador para este tipo de datasets, el cual se soporta en la Tabla 16 donde se visualizan los valores de

especificidad y sensibilidad por clasificador y en la Figura 7 se muestra la gráfica de los puntos de cada clasificador.

Tabla 15. Mejores resultados de especificidad y sensibilidad para datasets de características estadísticas subconjunto de eventos.

Clasificador	1- Especificidad	Sensibilidad
K-Means	0.022	0.899
SVM	0.085	0.749
RNPM	0.064	0.758
RF	<b>0.023</b>	<b>0.918</b>
J48	0.035	0.861
NB	0.094	0.409

Fuente: Elaboración propia a partir del proceso generado.

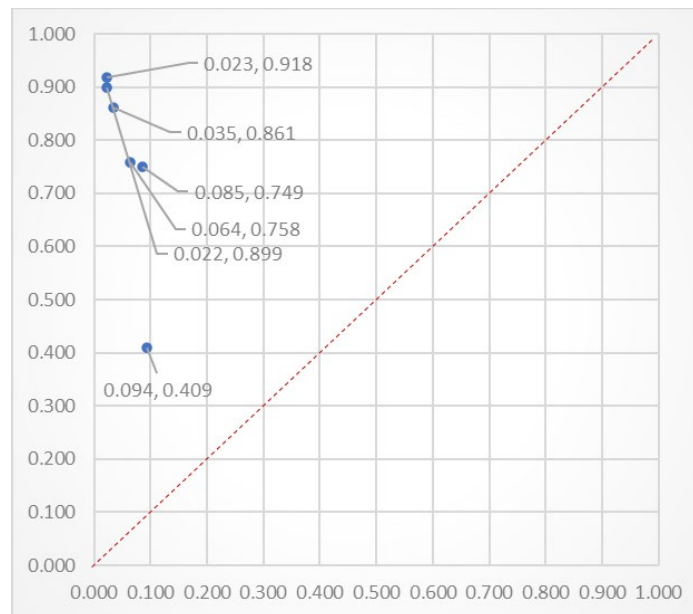


Figura 7. Gráfica de puntos en el espacio ROC mejores resultados datasets de características estadísticas subconjunto de eventos.

## RECONOCIMIENTO DE TIPOS DE CONDUCCIÓN.

De forma similar al procedimiento anterior, los resultados fueron analizados y tabulados de tal forma que permitieran la identificación de cual configuración y ventana fue la de mejor resultado y parámetro del clasificador generaba mejor resultado. La tabulación de los datos se puede observar en el Anexo 3. A continuación, se muestran los resultados por cada uno de los clasificadores. De



forma similar que el proceso anterior, inicialmente se evaluó el rendimiento usando un conjunto de carteristas mixtas, frecuenciales y estadísticas. Debido a que los resultados tuvieron el rendimiento deseado, se probó eliminando las características frecuenciales, obteniendo excelentes resultados. Se hizo lo mismo, eliminado las características estadísticas, pero los resultados fueron muy bajos, razón por la cual no se incluyeron.

## **EVALUACIÓN CON CARACTERÍSTICAS MIXTAS**

La ejecución de los diferentes clasificadores a los datasets de características mixtas del conjunto de tipo de conducción genera resultados que deben ser validados. A continuación, se muestran los resultados por cada uno de los clasificadores.

Los resultados de la ejecución de K-Means muestran que la mejor ventana de procesamiento es la de 2500 milisegundos para K igual a 3 dando una sensibilidad de 0.900 y una especificidad de 0.119. Para SVM los resultados muestran que la mejor ventana de procesamiento es la de 2500 milisegundo para un kernel igual a 20 dando una sensibilidad de 0.840 y una especificidad de 0.213. El mejor resultado que se visualiza en la red neuronal de perceptrón múltiple está sobre la ventana de los 2500 milisegundos para 5 capas con una sensibilidad de 0.840 y una especificidad de 0.195. La ventana que genera el mejor resultado para random forest es la de 2500 milisegundos para 80, 90 y 100 iteraciones, generando una sensibilidad de 0.860 y una especificidad de 0.191. La ventana de 2000 milisegundos en la que genera el mejor resultado para J48 con los factores de confianza de 0.10, obteniendo una sensibilidad de 0.767 y una especificidad de 0.278. Para Naïve Bayes se presenta como mejor ventana la de 2500 milisegundo generando una sensibilidad de 0.810 y una especificidad de 0.242.

Una vez determinados los mejores procesos de clasificación de cada uno de los clasificadores trabajados, se escogió el mejor clasificador para este tipo de datasets, el cual se soporta en la Tabla 16 donde se visualizan los valores de especificidad y sensibilidad por clasificador y en la Figura 7 se muestra la gráfica de los puntos de cada clasificador.

Tabla 16. Mejores resultados de especificidad y sensibilidad para datasets de características mixtas subconjunto tipo de conducción.

Clasificador	1- Especificidad	Sensibilidad
K-Means	<b>0.119</b>	<b>0.900</b>
SVM	0.213	0.840
RNPM	0.195	0.840
RF	0.191	0.860
J48	0.278	0.767
NB	0.242	0.810

Fuente: Elaboración propia a partir del proceso generado.

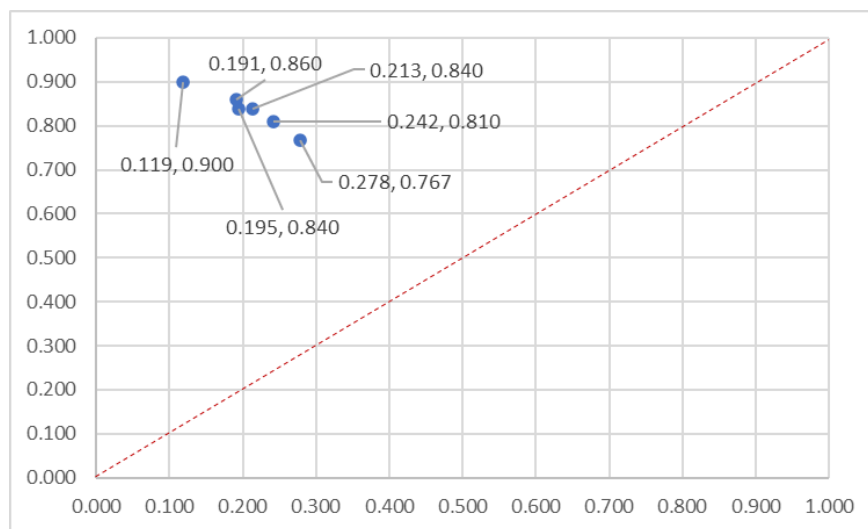


Figura 8. Gráfica de puntos en el espacio ROC mejores resultados datasets de características mixtas subconjunto de tipo de conducción.

### **Evaluación con características estadísticas**

Los resultados de la ejecución de los diferentes clasificadores del dataset de características estadísticas del conjunto de tipo de conducción generaron los resultados que se muestran a continuación.

Los resultados de la ejecución de K-Means muestran que la mejor ventana de procesamiento es la de 2500 milisegundos para K igual a 1 dando una sensibilidad de 0.930 y una especificidad de 0.091. Para SVM los resultados muestran que la mejor ventana de procesamiento es la de 2500 milisegundo para un kernel igual a 30 dando una sensibilidad de 0.850 y una especificidad de 0.198. El mejor resultado

que se visualiza en la red neuronal de perceptrón múltiple está sobre la ventana de los 2500 milisegundos para 7 capas con una sensibilidad de 0.920 y una especificidad de 0.079. La ventana que genera el mejor resultado para random forest es la de 2500 milisegundos para 50 y 60 iteraciones, generando una sensibilidad de 0.892 y una especificidad de 0.144. La ventana de 2500 milisegundos en la que genera el mejor resultado para J48 con los factores de confianza de 0.35, 0.40, 0.45 y 0.50, obteniendo una sensibilidad de 0.820 y una especificidad de 0.226. Para Naïve Bayes se presenta como mejor ventana la de 2500 milisegundo generando una sensibilidad de 0.800 y una especificidad de 0.230.

Una vez determinados los mejores procesos de clasificación de cada uno de los clasificadores trabajados, se escogió el mejor clasificador para este tipo de datasets, el cual se soporta en la Tabla 17 donde se visualizan los valores de especificidad y sensibilidad por clasificador y en la Figura 8 se muestra la gráfica de los puntos de cada clasificador.

Tabla 17. Mejores resultados de especificidad y sensibilidad para datasets de características estadísticas subconjunto tipo de conducción.

Clasificador	1- Especificidad	Sensibilidad
K-Means	0.091	0.930
SVM	0.198	0.850
RNPM	<b>0.079</b>	<b>0.920</b>
RF	0.144	0.890
J48	0.226	0.820
NB	0.230	0.800

*Fuente: Elaboración propia a partir del proceso generado.*

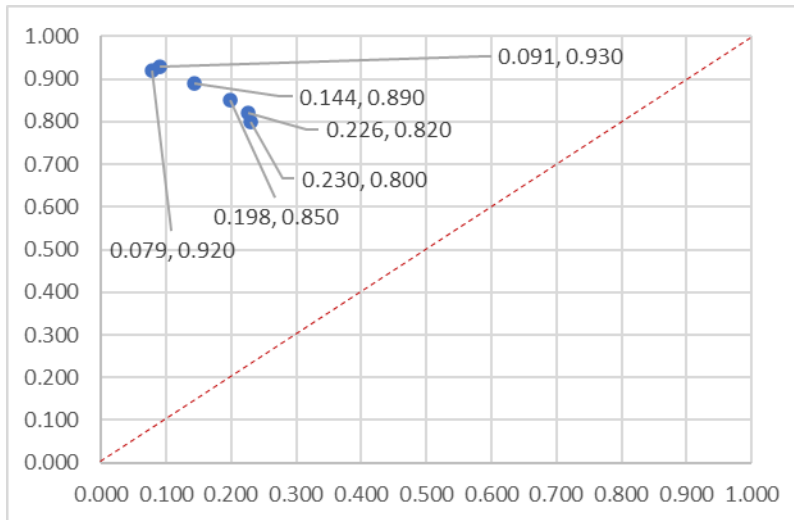


Figura 9. Gráfica de puntos en el espacio ROC mejores resultados datasets de características estadísticas subconjunto tipo de conducción.

## HALLAZGOS

En el análisis de los diferentes resultados arrojados por el proceso se pudo determinar que las características estadísticas generan un mejor resultado que las características frecuenciales para el subconjunto de datos de eventos y el subconjunto de datos de tipos de conducción. Se pudo observar que solo son necesarios los momentos estadísticos de orden 1 y 2 en las características estadísticas para realizar la clasificación del subconjunto de eventos y que para este mismo subconjunto solo es necesario la información del giroscopio para realizar la clasificación por características frecuenciales. También se encontró que para el subconjunto de datos de tipo de conducción es más adecuado el sensor de acelerómetro para clasificar el tipo de conducción. Adicionalmente es posible afirmar que las redes neuronales de perceptrón multicapa de 7 capas presenta el mejor resultado en la clasificación para el subconjunto de tipos de conducción y que random forest con 60, 70 y 90 iteraciones lo hacen para el subconjunto de eventos. Por último, se resalta que la ventana de tiempo en la que los clasificadores generaron el mejor resultado es la de 2500 milisegundos.

## CONCLUSIONES Y RECOMENDACIONES FINALES

Este trabajo de tesis de maestría proponía identificar los estilos de conducción apoyado en los sensores de los smartphones, para lograr esta meta se desarrolló una estrategia metodológica que se dividió en 3 fases. La primera fase tenía como propósito identificar los diferentes sensores de los smartphones que son utilizados para la captura de señales y determinar que modelos de computación y algoritmos de aprendizaje de máquina se han implementado por el estado del arte. La segunda fase buscaba desarrollar todo el trabajo de campo para construir un dataset que permitiera identificar estilos de conducción. La tercera y última fase buscaba evaluar y ajustar los métodos de aprendizaje de máquina escogidos haciendo uso de herramientas tecnológicas para este fin, para esto se debió determinar los atributos que generaban la mejor clasificación y generar procesos de clasificación con diferentes configuraciones de sus parámetros de ejecución.

Se realizó una revisión del estado del arte, usando un protocolo de búsqueda sistémica, con el cual se pudo establecer que de los sensores presentes en un smartphone los más utilizados según las publicaciones consultadas son el acelerómetro, el giroscopio, el GPS, el magnetómetro y la cámara. Con respecto a los algoritmos de aprendizaje de máquina que se implementaron en el estado del arte, se pudo evidenciar que los métodos que más se implementaron fueron las máquinas de soporte vectorial (SVM), random forest, los árboles de decisión C4.5, las redes neuronales de perceptrón múltiple (MLP) y los algoritmos de Naïve Bayes y K-Means; También se logró identificar que en la mayoría de los artículos se utiliza más de un algoritmo con el fin de mejorar la efectividad en la identificación de los estilos de conducción; en muchos de estos casos se incluyen métodos de validación cruzada en la evaluación de resultados. Como la identificación de estilos de conducción puede ser un proceso ambiguo, lo que se encontró en muchos de los artículos fue el reconocimiento de eventos de conducción ya que la forma como se generaba este evento puede clasificar a un conductor dentro de un estilo de conducción.

Dados los resultados de la revisión del estado del arte se tomó la decisión de trabajar con los sensores del acelerómetro y el giroscopio como sensores de trabajo, debido a su amplia documentación y adecuados resultados de su uso, además de no requerir alto consumo de batería. Con respecto a los algoritmos de aprendizaje de máquina se escogieron K-Means, máquinas de soporte vectorial, redes neuronales de perceptrón multicapa, random forest, árbol J48 y Naïve Bayes, por razones similares, amplia documentación y adecuados resultados reportados en la literatura.

Se construyó un dataset de aproximadamente 45 minutos de conducción, el cual era necesario para realizar el análisis de los datos. Para esto se revisaron diferentes aplicaciones para celular que capturaran las señales de los sensores escogidos y así determinar la que mejor se adaptara a nuestro proceso. Se realizó un proceso de levantamiento de señales en 3 recorridos diferentes, que luego fueron etiquetadas por evento (giros a la derecha, izquierda y en U, Cambio de carril a la derecha y a la izquierda, arranque, detención, vehículo en movimiento y detenido) y por tipo de conducción (normal y agresiva). El dataset recibió un proceso de balance de clases, para que fuera adecuado al ser usado por los distintos algoritmos de aprendizaje de máquina escogidos. Todo el proceso se realizó con herramientas de ofimática, lenguajes de programación y con herramientas de aprendizaje de máquina.

Se buscó identificar los mejores clasificadores y sus parámetros de ejecución, así como el mejor tamaño ventana de la señal para el procesamiento de datos analizando, usando las métricas de especificidad y sensibilidad en un procedimiento de validación cruzada de 10 paquetes (10-fold Crossvalidation), comparando los resultados en espacio ROC.

Con respecto a los resultados, se pudo identificar que la mejor ventana de ejecución del proceso de clasificación es la de 2500 milisegundos y la que menos funciona es la de 1000 milisegundos. También es posible concluir que los clasificadores generan mejores resultados con las características estadísticas que con las características frecuenciales. Se aclara que la experimentación computacional permitió establecer

que los momentos estadísticos, usados como características, que mejor se adaptan a los procesos de clasificación son el orden uno (1) y el dos (2) y que los momentos estadísticos de orden superior a tres (3) no son necesarios de procesar. Para las características frecuenciales el mejor sensor que apoya esta clasificación es el giroscopio y para las características estadísticas es necesario hacer uso de los datos de los dos (2) sensores.

Con respecto a las clasificaciones que se realizan por eventos de conducción, el mejor clasificador para datos estadísticos es random forest y el que menos se recomienda es Naïve Bayes. Para la evaluación con características mixtas el mejor clasificador es el árbol J48 y el que menos procede es random forest. Sin embargo, el resultado para este proceso de características mixtas es muy bajo ya que no llega ni al 50% de sensibilidad por lo que se recomienda emplear solo características estadísticas. Lo anterior se puede deber a un nivel alto de ruido de las señales, propio de su naturaleza, que permite una mejor diferenciación con características estadísticas que frecuenciales.

Para la clasificación por tipo de conducción (normal, agresivo) con características mixtas el mejor clasificador es K-Means y el menos recomendado es J48. De igual forma, los mejores resultados se lograron solo con características estadísticas, donde el mejor clasificador es la red neuronal de perceptrón multicapa y el menos recomendado es Naïve Bayes.

La principal limitación en el desarrollo del proyecto fue generar un dataset usando un solo conductor, un solo tipo de vehículo y una sola zona urbana; hizo falta poder levantar información con personas de sexo y edad diferente, que se condujera por avenidas y carreteras y en otro tipo de vehículos como SUV, de transmisión automática entre otros. Sin embargo, actualmente no se encuentra un dataset de conducción en Colombia y Latinoamérica lo que se constituye en una contribución interesante en esta área.

También se propone como trabajo futuro incluir en el dataset otros conductores, otros tipos de vehículos y otras zonas en Bogotá e inclusive en el país. También se

propones explorar otros métodos de aprendizaje de máquina diferentes a los trabajados en este proyecto que porten una forma de clasificar distinta.

Otro tema que puede ser tenido en cuenta para trabajos futuros es el desarrollo de aplicaciones que detectan en línea la forma en que se conduce un vehículo y que aporten un mejor proceso de detección.

Con respecto las aplicaciones prácticas, en el capítulo de introducción se abordó de manera amplia las posibilidades que puede tener el reconocimiento automático de los estilos de conducción en la mejora de la seguridad vial. Pero es necesario hacer tener en cuenta que este tipo de tecnologías puede llegar a capturar mucha información del vehículo y su conductor, que puede considerarse sensible, afectando potencialmente la privacidad de las personas. En este sentido, el uso de un consentimiento informado del uso de una aplicación toma relevancia, pues al igual otras aplicaciones que capturan datos, la aceptación de las condiciones por parte de los usuarios no siempre es clara y transparente. Adicionalmente, al permitirse el almacenamiento de información de los diferentes tipos de sensores en las bases de datos, para favorecer su uso y el reconocimiento adecuado de los estilos de conducción, será posible abrir la puerta a los ataques de seguridad y la explotación de la información con otros fines.

Una de las aplicaciones que incluso ya está viendo de esta tecnología, es su uso como herramienta para la expedición de pólizas de seguro. Al respecto se debe hacer una clara información al usuario de los datos que se van a obtener de su vehículo o teléfono. Por lo tanto, una persona que firme un seguro también debería ser enterado de toda la información que se podrá adquirir y del uso que hagan de esta. Si estos datos pueden ser modificados por terceros, se genera una situación de desconfianza hacia las aplicaciones, sus usos y los algoritmos en los que estén basadas. Otra situación que puede presentarse es la mala calibración o detección errónea por parte de los algoritmos de reconocimiento, que puedan afectar la calificación de un usuario en una central de riesgos, los costos de seguro o en el caso de usar estos parámetros entregados por el sistema en una posible demanda por accidentes de tránsito. En este caso, es necesario hacer lo posible por



disminuir los errores en la detección de estilos riesgosos de conducción y la confianza en los datos almacenados. En resumen, los resultados de este proyecto final de maestría, al igual que trabajo previo publicado en el estado del arte, tiene alta factibilidad de aplicación, pero su uso no debe obseder únicamente a la facilidad técnica, intereses públicos o comerciales, debe hacer un debate ético de las implicaciones de estos desarrollos en la cotidianidad de nuestras vidas.

## REFERENCIAS

- [1] M. Q. Khan and S. Lee, "A Comprehensive Survey of Driving Monitoring and Assistance Systems," *Sensors*, vol. 19, no. 11, p. 2574, Jun. 2019, doi: 10.3390/s19112574.
- [2] V. and I. P. (NVI) Management of Noncommunicable Diseases, Disability, "Informe Sobre la Situación Mundial de la Seguridad Vial 2015," World Health Organization, 2015.
- [3] S. J. Valbuena Cortés, "Muertes y lesiones no fatales por accidentes de transporte, Colombia, 2011," Instituto Nacional de Medicina Legal y Ciencias Forenses, 2012.
- [4] Ministerio de Transporte, "Plan Nacional de Seguridad Vial Colombia 2011 - 2021," Bogotá D.C., 2015.
- [5] J. C. Martínez Rojas, C. E. Villalobos Cuadrado, and E. Benítez, "Elaboración de un modelo de negocio de asistencia a la conducción mediante una propuesta de IoT para la prevención de accidentes de tránsito en la ciudad de Bogotá," Chía, 2018.
- [6] R. I. Lugo Hernández, "Análisis y reconocimiento de patrones de conducción temeraria mediante algoritmos de aprendizaje computacional," 2016.
- [7] B. M. Leiner *et al.*, "A Brief History of the Internet," *Baseline*, 1999. <https://arxiv.org/abs/cs/9901011> (accessed May 16, 2019).
- [8] I. Machorro-Cano, G. Alor-Hernández, N. A. Cruz-Ramos, C. Sánchez-Ramírez, and M. G. Segura-Ozuna, "A Brief Review of IoT Platforms and Applications in Industry," Springer, Cham, 2018, pp. 293–324. doi: 10.1007/978-3-319-56871-3\_15.
- [9] S. Moosavi, B. Omidvar-Tehrani, R. B. Craig, A. Nandi, and R. Ramnath, "Characterizing Driving Context from Driver Behavior," in *G/S*:

*Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, 2017, vol. 2017-Novem, pp. 1–4. doi: 10.1145/3139958.3139992.

- [10] I. Skog *et al.*, “Insurance telematics: opportunities and challenges with the smartphone solution.”
- [11] J. Engelbrecht, M. J. Booyen, F. J. Bruwer, and G.-J. van Rooyen, “Survey of smartphone-based sensing in vehicles for intelligent transportation system applications,” *IET Intelligent Transport Systems*, vol. 9, no. 10, pp. 924–935, Dec. 2015, doi: 10.1049/iet-its.2014.0248.
- [12] C. Marina Martinez, M. Heucke, F. Y. Wang, B. Gao, and D. Cao, “Driving Style Recognition for Intelligent Vehicle Control and Advanced Driver Assistance: A Survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3. Institute of Electrical and Electronics Engineers Inc., pp. 666–676, Mar. 01, 2018. doi: 10.1109/TITS.2017.2706978.
- [13] A. Chowdhury, T. Banerjee, T. Chakravarty, and P. Balamuralidhar, “Smartphone based estimation of relative risk propensity for inducing good driving behavior,” in *UbiComp and ISWC 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the Proceedings of the 2015 ACM International Symposium on Wearable Computers*, Sep. 2015, pp. 743–752. doi: 10.1145/2800835.2804392.
- [14] J. F. Fernández Joya, G. Á. Buitrago, H. Luna-García, and W. J. Samiento, “Smartphones, Suitable Tool for Driver Behavior Recognition. A Systematic Review,” in *Communications in Computer and Information Science*, 2020, vol. 1334. doi: 10.1007/978-3-030-66919-5\_24.
- [15] Lu Tan and Neng Wang, “Future internet: The Internet of Things,” in *2010 3rd International Conference on Advanced Computer Theory and*

*Engineering(ICACTE)*, Aug. 2010, pp. V5-376-V5-380. doi: 10.1109/ICACTE.2010.5579543.

- [16] O. Vermesan *et al.*, "Internet of Things Strategic Research Roadmap," 2009.
- [17] O. Moreno and J. Díaz Fernández, "Diseño y despliegue de una arquitectura IoT para el análisis de datos en tiempo real," 2016.
- [18] L. Coetzee and J. Eksteen, *The Internet of Things – Promise for the Future? An Introduction*. 2011.
- [19] D. L. Pinzón Niño, "Panorama de aplicación de internet de las cosas (IoT)."
- [20] S. Kaplan, M. A. Guvensan, A. G. Yavuz, and Y. Karalurt, "Driver Behavior Analysis for Safe Driving: A Survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3017–3032, Dec. 2015, doi: 10.1109/TITS.2015.2462084.
- [21] D. M. Vlachogiannis, E. I. Vlahogianni, and J. Golias, "A reinforcement learning model for personalized driving policies identification," *International Journal of Transportation Science and Technology*, no. xxxx, 2020, doi: 10.1016/j.ijtst.2020.03.002.
- [22] Z. Ouyang, J. Niu, Y. Liu, and X. Liu, "An Ensemble Learning-Based Vehicle Steering Detector Using Smartphones," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 5, pp. 1964–1975, 2020, doi: 10.1109/TITS.2019.2909107.
- [23] A. Kashevnik, I. Lashkov, and A. Gurtov, "Methodology and Mobile Application for Driver Behavior Analysis and Accident Prevention," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 6, pp. 2427–2436, 2020, doi: 10.1109/TITS.2019.2918328.
- [24] H. R. Eftekhari and M. Ghatee, "A similarity-based neuro-fuzzy modeling for driving behavior recognition applying fusion of smartphone sensors,"

*Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 23, no. 1, pp. 72–83, 2019, doi: 10.1080/15472450.2018.1506338.

- [25] E. Papadimitriou, A. Argyropoulou, D. I. Tselentis, and G. Yannis, “Analysis of driver behaviour through smartphone data: The case of mobile phone use while driving,” *Safety Science*, vol. 119, no. May, pp. 91–97, 2019, doi: 10.1016/j.ssci.2019.05.059.
- [26] A. A. Rahman, W. Saleem, and V. V. Iyer, “Driving Behavior Profiling and Prediction in KSA using Smart Phone Sensors and MLAs,” *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology, JEEIT 2019 - Proceedings*, pp. 34–39, 2019, doi: 10.1109/JEEIT.2019.8717533.
- [27] N. Lourenco, B. Cabral, and J. Granjal, “Driving Profile using Evolutionary Computation,” *2019 IEEE Congress on Evolutionary Computation, CEC 2019 - Proceedings*, no. MI, pp. 2466–2473, 2019, doi: 10.1109/CEC.2019.8790154.
- [28] D. I. Tselentis, E. I. Vlahogianni, and G. Yannis, “Driving safety efficiency benchmarking using smartphone data,” *Transportation Research Part C: Emerging Technologies*, vol. 109, no. November, pp. 343–357, 2019, doi: 10.1016/j.trc.2019.11.006.
- [29] E. G. Mantouka, E. N. Barmponakis, and E. I. Vlahogianni, “Identifying driving safety profiles from smartphone data using unsupervised learning,” *Safety Science*, vol. 119, no. January, pp. 84–90, 2019, doi: 10.1016/j.ssci.2019.01.025.
- [30] M. M. Bejani and M. Ghatee, “A context aware system for driving style evaluation by an ensemble learning on smartphone sensors data,” *Transportation Research Part C: Emerging Technologies*, vol. 89, no. February, pp. 303–320, 2018, doi: 10.1016/j.trc.2018.02.009.

- [31] Y. Guo, B. Guo, Y. Liu, Z. Wang, Y. Ouyang, and Z. Yu, "CrowdSafe: Detecting extreme driving behaviors based on mobile crowdsensing," *2017 IEEE SmartWorld Ubiquitous Intelligence and Computing, Advanced and Trusted Computed, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation, SmartWorld/SCALCOM/UIC/ATC/CBDCCom/IOP/SCI 2017 -* , pp. 1–8, 2018, doi: 10.1109/UIC-ATC.2017.8397522.
- [32] A. U. Nambi *et al.*, "Demo: HAMS: Driver and driving monitoring using a smartphone," *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, pp. 840–842, 2018, doi: 10.1145/3241539.3267723.
- [33] F. Tahmasbi, Y. Wang, Y. Chen, and M. Gruteser, "Poster: Your phone tells us the truth: R identification using smartphone on one turn," *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, no. October, pp. 762–764, 2018, doi: 10.1145/3241539.3267769.
- [34] L. Kang and S. Banerjee, "Practical driving analytics with smartphone sensors," *IEEE Vehicular Networking Conference, VNC*, vol. 2018-Janua, pp. 303–310, 2018, doi: 10.1109/VNC.2017.8275595.
- [35] S. K. Al-luhaibi, A. M. Said, and M. S. Najim Al-Din, "Recognition of driving maneuvers based accelerometer sensor," *International Journal of Engineering and Technology*, vol. 9, no. 11, pp. 1542–1547, 2018.
- [36] C. Streiffer, R. Raghavendra, T. Benson, and M. Srivatsa, "DarNet: A deep learning solution for distracted driving detection," *Middleware 2017 - Proceedings of the 2017 International Middleware Conference (Industrial Track)*, pp. 22–28, 2017, doi: 10.1145/3154448.3154452.

- [37] E. I. Vlahogianni and E. N. Barmounakis, "Driving analytics using smartphones: Algorithms, comparisons and challenges," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 196–206, 2017, doi: 10.1016/j.trc.2017.03.014.
- [38] X. Xu, S. Yin, and P. Ouyang, "Fast and low-power behavior analysis on vehicles using smartphones," *2017 6th International Symposium on Next Generation Electronics, ISNE 2017*, 2017, doi: 10.1109/ISNE.2017.7968748.
- [39] W. Cho and S. H. Kim, "Multimedia sensor dataset for the analysis of vehicle movement," *Proceedings of the 8th ACM Multimedia Systems Conference, MMSys 2017*, pp. 175–180, 2017, doi: 10.1145/3083187.3083217.
- [40] D. M. Vlachogiannis, E. I. Vlahogianni, and J. Golias, "A reinforcement learning model for personalized driving policies identification," *International Journal of Transportation Science and Technology*, no. xxxx, 2020, doi: 10.1016/j.ijst.2020.03.002.
- [41] Z. Ouyang, J. Niu, Y. Liu, and X. Liu, "An Ensemble Learning-Based Vehicle Steering Detector Using Smartphones," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 5, pp. 1964–1975, 2020, doi: 10.1109/TITS.2019.2909107.
- [42] A. Kashevnik, I. Lashkov, and A. Gurtov, "Methodology and Mobile Application for Driver Behavior Analysis and Accident Prevention," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 6, pp. 2427–2436, 2020, doi: 10.1109/TITS.2019.2918328.
- [43] R. Wang, F. Xie, B. Zhang, W. Liu, W. Qian, and W. Xian, "Detecting abnormal driving behaviors by smartphone sensors based on multi-feature convolutional neural network," *Chinese Control Conference, CCC*, vol. 2019-July, no. 2, pp. 6639–6644, 2019, doi: 10.23919/ChiCC.2019.8865705.

- [44] C. Streiffer, R. Raghavendra, T. Benson, and M. Srivatsa, "DarNet: A deep learning solution for distracted driving detection," *Middleware 2017 - Proceedings of the 2017 International Middleware Conference (Industrial Track)*, pp. 22–28, 2017, doi: 10.1145/3154448.3154452.
- [45] X. Xu, S. Yin, and P. Ouyang, "Fast and low-power behavior analysis on vehicles using smartphones," *2017 6th International Symposium on Next Generation Electronics, ISNE 2017*, 2017, doi: 10.1109/ISNE.2017.7968748.
- [46] W. Cho and S. H. Kim, "Multimedia sensor dataset for the analysis of vehicle movement," *Proceedings of the 8th ACM Multimedia Systems Conference, MMSys 2017*, pp. 175–180, 2017, doi: 10.1145/3083187.3083217.
- [47] T. L. L. Mon and T. L. L. Thein, "Design and implementation of smart alert system for reducing road traffic accidents in Myanmar," *AIP Conference Proceedings*, vol. 2129, no. July, 2019, doi: 10.1063/1.5118022.
- [48] I. Lashkov, A. Kashevnik, N. Shilov, V. Parfenov, and A. Shabaev, "Driver dangerous state detection based on OpenCV & dlib libraries using mobile video processing," *Proceedings - 22nd IEEE International Conference on Computational Science and Engineering and 17th IEEE International Conference on Embedded and Ubiquitous Computing, CSE/EUC 2019*, pp. 74–79, 2019, doi: 10.1109/CSE/EUC.2019.00024.



## **ANEXOS**

Anexo 1. Composición del archivo del dataset estadístico.

Nombre de la columna	Descripción de la columna
Evento	Evento ejecutado / Tipo de conducción
AX-M1	Media para el eje X del acelerómetro
AX-M2	Varianza para el eje X del acelerómetro
AX-M3	Asimetría para el eje X del acelerómetro
AX-M4	Curtosis para el eje X del acelerómetro
AX-M5	Momento estadístico de orden 5 para el eje X del acelerómetro
AX-M6	Momento estadístico de orden 6 para el eje X del acelerómetro
AX-M7	Momento estadístico de orden 7 para el eje X del acelerómetro
AY-M1	Media para el eje Y del acelerómetro
AY-M2	Varianza para el eje Y del acelerómetro
AY-M3	Asimetría para el eje Y del acelerómetro
AY-M4	Curtosis para el eje Y del acelerómetro
AY-M5	Momento estadístico de orden 5 para el eje Y del acelerómetro
AY-M6	Momento estadístico de orden 6 para el eje Y del acelerómetro
AY-M7	Momento estadístico de orden 7 para el eje Y del acelerómetro
AZ-M1	Media para el eje Z del acelerómetro
AZ-M2	Varianza para el eje Z del acelerómetro
AZ-M3	Asimetría para el eje Z del acelerómetro
AZ-M4	Curtosis para el eje Z del acelerómetro
AZ-M5	Momento estadístico de orden 5 para el eje Z del acelerómetro
AZ-M6	Momento estadístico de orden 6 para el eje Z del acelerómetro
AZ-M7	Momento estadístico de orden 7 para el eje Z del acelerómetro
GX-M1	Media para el eje X del giroscopio
GX-M2	Varianza para el eje X del giroscopio
GX-M3	Asimetría para el eje X del giroscopio
GX-M4	Curtosis para el eje X del giroscopio
GX-M5	Momento estadístico de orden 5 para el eje X del giroscopio
GX-M6	Momento estadístico de orden 6 para el eje X del giroscopio
GX-M7	Momento estadístico de orden 7 para el eje X del giroscopio
GY-M1	Media para el eje Y del giroscopio
GY-M2	Varianza para el eje Y del giroscopio
GY-M3	Asimetría para el eje Y del giroscopio
GY-M4	Curtosis para el eje Y del giroscopio

GY-M5	Momento estadístico de orden 5 para el eje Y del giroscopio
GY-M6	Momento estadístico de orden 6 para el eje Y del giroscopio
GY-M7	Momento estadístico de orden 7 para el eje Y del giroscopio
GZ-M1	Media para el eje Z del giroscopio
GZ-M2	Varianza para el eje Z del giroscopio
GZ-M3	Asimetría para el eje Z del giroscopio
GZ-M4	Curtosis para el eje Z del giroscopio
GZ-M5	Momento estadístico de orden 5 para el eje Z del giroscopio
GZ-M6	Momento estadístico de orden 6 para el eje Z del giroscopio
GZ-M7	Momento estadístico de orden 7 para el eje Z del giroscopio

## Anexo 2. Estructura del dataset de características frecuenciales.

Nombre de la columna	Descripción de la columna
Evento	Evento ejecutado / Tipo de conducción
AX-D1	Pico 1 de la derivada de la transformada de Fourier para el eje X - Acelerómetro
AX-D2	Pico 2 de la derivada de la transformada de Fourier para el eje X - Acelerómetro
AX-D3	Pico 3 de la derivada de la transformada de Fourier para el eje X - Acelerómetro
AX-D4	Pico 4 de la derivada de la transformada de Fourier para el eje X - Acelerómetro
AX-D5	Pico 5 de la derivada de la transformada de Fourier para el eje X - Acelerómetro
AY-D1	Pico 1 de la derivada de la transformada de Fourier para el eje Y - Acelerómetro
AY-D2	Pico 2 de la derivada de la transformada de Fourier para el eje Y - Acelerómetro
AY-D3	Pico 3 de la derivada de la transformada de Fourier para el eje Y - Acelerómetro
AY-D4	Pico 4 de la derivada de la transformada de Fourier para el eje Y - Acelerómetro
AY-D5	Pico 5 de la derivada de la transformada de Fourier para el eje Y - Acelerómetro
AZ-D1	Pico 1 de la derivada de la transformada de Fourier para el eje Z - Acelerómetro
AZ-D2	Pico 2 de la derivada de la transformada de Fourier para el eje Z - Acelerómetro
AZ-D3	Pico 3 de la derivada de la transformada de Fourier para el eje Z - Acelerómetro
AZ-D4	Pico 4 de la derivada de la transformada de Fourier para el eje Z - Acelerómetro
AZ-D5	Pico 5 de la derivada de la transformada de Fourier para el eje Z - Acelerómetro
GX-D1	Pico 1 de la derivada de la transformada de Fourier para el eje X - Giroscopio
GX-D2	Pico 2 de la derivada de la transformada de Fourier para el eje X - Giroscopio
GX-D3	Pico 3 de la derivada de la transformada de Fourier para el eje X - Giroscopio
GX-D4	Pico 4 de la derivada de la transformada de Fourier para el eje X - Giroscopio
GX-D5	Pico 5 de la derivada de la transformada de Fourier para el eje X - Giroscopio
GY-D1	Pico 1 de la derivada de la transformada de Fourier para el eje Y - Giroscopio
GY-D2	Pico 2 de la derivada de la transformada de Fourier para el eje Y - Giroscopio
GY-D3	Pico 3 de la derivada de la transformada de Fourier para el eje Y - Giroscopio
GY-D4	Pico 4 de la derivada de la transformada de Fourier para el eje Y - Giroscopio
GY-D5	Pico 5 de la derivada de la transformada de Fourier para el eje Y - Giroscopio
GZ-D1	Pico 1 de la derivada de la transformada de Fourier para el eje Z - Giroscopio
GZ-D2	Pico 2 de la derivada de la transformada de Fourier para el eje Z - Giroscopio
GZ-D3	Pico 3 de la derivada de la transformada de Fourier para el eje Z - Giroscopio
GZ-D4	Pico 4 de la derivada de la transformada de Fourier para el eje Z - Giroscopio
GZ-D5	Pico 5 de la derivada de la transformada de Fourier para el eje Z - Giroscopio



Anexo 3. Datos tabulados de los procesos de clasificación.

K-Means Subconjunto de eventos

<b>Ventana (milisegs)</b>	<b>1</b>	<b>3</b>	<b>5</b>	<b>7</b>	<b>9</b>	<b>11</b>
1000	0.503	0.510	0.506	0.511	0.507	0.495
1500	0.466	0.480	0.480	0.492	0.486	0.481
2000	0.457	0.461	0.472	0.477	0.484	0.481
2500	0.482	0.484	0.496	0.501	0.502	0.496

<b>Ventana (milisegs)</b>	<b>1</b>	<b>3</b>	<b>5</b>	<b>7</b>	<b>9</b>	<b>11</b>
1000	0.515	0.383	0.373	0.376	0.390	0.402
1500	0.472	0.420	0.427	0.380	0.395	0.394
2000	0.463	0.471	0.354	0.358	0.364	0.363
2500	0.485	0.496	0.495	0.493	0.497	0.488

<b>Ventana (milisegs)</b>	<b>1</b>	<b>3</b>	<b>5</b>	<b>7</b>	<b>9</b>	<b>11</b>
1000	0.741	0.734	0.720	0.715	0.707	0.701
1500	0.795	0.784	0.759	0.745	0.730	0.728
2000	0.862	0.846	0.828	0.815	0.798	0.775
2500	0.899	0.878	0.860	0.845	0.826	0.812

<b>Ventana (milisegs)</b>	<b>1</b>	<b>3</b>	<b>5</b>	<b>7</b>	<b>9</b>	<b>11</b>
1000	0.057	0.064	0.069	0.072	0.074	0.077
1500	0.045	0.048	0.055	0.062	0.065	0.068
2000	0.030	0.034	0.039	0.043	0.049	0.056
2500	0.022	0.027	0.031	0.035	0.040	0.045

SVM Subconjunto de eventos

<b>Ventana</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
----------------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

<b>(milisegs)</b>										
1000	0.428	0.512	0.582	0.615	0.628	0.634	0.645	0.642	0.640	0.633
1500	0.495	0.540	0.608	0.634	0.642	0.656	0.678	0.684	0.678	0.657
2000	0.522	0.569	0.625	0.660	0.677	0.692	0.702	0.704	0.688	0.679
2500	0.548	0.601	0.643	0.684	0.704	0.722	0.738	0.742	0.724	0.718

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
1000	0.333	0.406	0.439	0.462	0.430	0.487	0.477	0.613	0.623	0.623
1500	0.409	0.409	0.402	0.446	0.436	0.501	0.594	0.678	0.666	0.649
2000	0.416	0.454	0.497	0.491	0.489	0.567	0.591	0.692	0.681	0.675
2500	0.415	0.459	0.531	0.518	0.575	0.706	0.731	0.738	0.722	0.718

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
1000	0.481	0.511	0.524	0.540	0.561	0.579	0.593	0.612	0.622	0.632
1500	0.508	0.534	0.549	0.577	0.600	0.619	0.632	0.649	0.668	0.688
2000	0.534	0.564	0.595	0.615	0.636	0.652	0.670	0.699	0.711	0.722
2500	0.562	0.596	0.627	0.650	0.671	0.685	0.708	0.727	0.734	0.749

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
1000	0.206	0.190	0.183	0.174	0.164	0.155	0.148	0.138	0.134	0.129
1500	0.193	0.179	0.171	0.157	0.146	0.137	0.130	0.122	0.112	0.102
2000	0.182	0.165	0.151	0.140	0.130	0.122	0.114	0.102	0.097	0.091
2500	0.168	0.151	0.136	0.126	0.115	0.109	0.100	0.093	0.090	0.085

### Red Neuronal de Perceptrón Múltiple Subconjunto de eventos

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
1000	0.539	0.609	0.636	0.629	0.632	0.641	0.642
1500	0.550	0.638	0.665	0.679	0.670	0.692	0.679
2000	0.555	0.653	0.689	0.696	0.708	0.724	0.720

2500	0.576	0.679	0.683	0.705	0.731	0.747	0.741
------	-------	-------	-------	-------	-------	-------	-------

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
1000	0.188	0.393	0.422	0.444	0.422	0.425	0.468
1500	0.191	0.413	0.440	0.446	0.479	0.654	0.478
2000	0.192	0.373	0.484	0.513	0.681	0.710	0.696
2500	0.203	0.430	0.493	0.616	0.712	0.735	0.731

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
1000	0.539	0.606	0.636	0.655	0.660	0.665	0.670
1500	0.549	0.640	0.669	0.692	0.693	0.699	0.692
2000	0.560	0.648	0.686	0.695	0.726	0.719	0.734
2500	0.574	0.681	0.703	0.706	0.745	0.758	0.746

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
1000	0.161	0.115	0.113	0.103	0.101	0.096	0.095
1500	0.158	0.104	0.103	0.087	0.084	0.083	0.085
2000	0.154	0.100	0.098	0.090	0.073	0.073	0.070
2500	0.152	0.094	0.089	0.082	0.069	0.064	0.067

### Random Forest Subconjunto de eventos

<b>Ventana (milisegs)</b>	<b>50</b>	<b>60</b>	<b>70</b>	<b>80</b>	<b>90</b>	<b>100</b>
1000	0.787	0.791	0.791	0.792	0.792	0.793
1500	0.847	0.849	0.852	0.850	0.850	0.846
2000	0.881	0.883	0.884	0.885	0.885	0.886
2500	0.915	0.915	0.917	0.917	0.917	0.916

<b>Ventana</b>	<b>50</b>	<b>60</b>	<b>70</b>	<b>80</b>	<b>90</b>	<b>100</b>
----------------	-----------	-----------	-----------	-----------	-----------	------------



<b>(milisegs)</b>						
1000	0.789	0.794	0.795	0.795	0.792	0.793
1500	0.848	0.849	0.853	0.851	0.851	0.846
2000	0.882	0.883	0.884	0.886	0.886	0.886
2500	0.915	0.915	0.917	0.918	0.918	0.917

<b>Ventana (milisegs)</b>	<b>50</b>	<b>60</b>	<b>70</b>	<b>80</b>	<b>90</b>	<b>100</b>
1000	0.790	0.793	0.793	0.792	0.792	0.793
1500	0.856	0.860	0.864	0.864	0.863	0.861
2000	0.890	0.893	0.893	0.892	0.894	0.895
2500	0.916	0.918	0.918	0.916	0.918	0.917

<b>Ventana (milisegs)</b>	<b>50</b>	<b>60</b>	<b>70</b>	<b>80</b>	<b>90</b>	<b>100</b>
1000	0.059	0.058	0.058	0.058	0.058	0.057
1500	0.040	0.039	0.038	0.038	0.038	0.038
2000	0.031	0.030	0.029	0.030	0.029	0.029
2500	0.024	0.023	0.023	0.024	0.023	0.023

#### J48 Subconjunto de eventos

<b>Ventana (milisegs)</b>	<b>0.05</b>	<b>0.10</b>	<b>0.15</b>	<b>0.20</b>	<b>0.25</b>	<b>0.30</b>	<b>0.35</b>	<b>0.40</b>	<b>0.45</b>	<b>0.50</b>
1000	0.720	0.720	0.718	0.710	0.701	0.696	0.697	0.696	0.696	0.695
1500	0.767	0.767	0.763	0.758	0.756	0.756	0.755	0.754	0.753	0.752
2000	0.806	0.808	0.808	0.807	0.806	0.803	0.803	0.802	0.802	0.800
2500	0.841	0.843	0.842	0.843	0.843	0.843	0.843	0.843	0.843	0.841

<b>Ventana (milisegs)</b>	<b>0.05</b>	<b>0.10</b>	<b>0.15</b>	<b>0.20</b>	<b>0.25</b>	<b>0.30</b>	<b>0.35</b>	<b>0.40</b>	<b>0.45</b>	<b>0.50</b>
1000	0.706	0.709	0.710	0.703	0.695	0.693	0.691	0.690	0.690	0.690
1500	0.764	0.765	0.761	0.757	0.755	0.756	0.755	0.754	0.753	0.752
2000	0.805	0.807	0.807	0.807	0.805	0.803	0.802	0.801	0.801	0.799

2500	0.842	0.842	0.841	0.842	0.842	0.842	0.842	0.842	0.842	0.842	0.840
------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
1000	0.728	0.731	0.730	0.725	0.723	0.722	0.723	0.723	0.723	0.723
1500	0.777	0.776	0.778	0.782	0.781	0.778	0.778	0.779	0.779	0.779
2000	0.816	0.820	0.820	0.822	0.821	0.817	0.817	0.817	0.817	0.817
2500	0.855	0.859	0.859	0.861	0.860	0.861	0.861	0.861	0.861	0.860

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
1000	0.069	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068	0.068
1500	0.053	0.053	0.052	0.051	0.051	0.052	0.051	0.051	0.051	0.051
2000	0.046	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0.045
2500	0.036	0.036	0.036	0.035	0.035	0.035	0.035	0.035	0.035	0.035

### Naïve Bayes Subconjunto de eventos

<b>Ventana (milisegs)</b>	<b>Resultado</b>
1000	0.341
1500	0.356
2000	0.384
2500	0.409

<b>Ventana (milisegs)</b>	<b>Resultado</b>
1000	0.485
1500	0.496
2000	0.521
2500	0.552

<b>Ventana (milisegs)</b>	<b>Resultado</b>
1000	0.330
1500	0.357

2000	0.377
2500	0.409

Ventana (milisegs)	Resultado
1000	0.087
1500	0.093
2000	0.095
2500	0.094

### K-Means Subconjunto de tipos de conducción

Ventana (milisegs)	K1	K3	K5	K7	K9	K11
<b>1000</b>	0.673	0.692	0.607	0.598	0.636	0.645
<b>1500</b>	0.692	0.635	0.615	0.635	0.606	0.644
<b>2000</b>	0.738	0.728	0.738	0.718	0.680	0.670
<b>2500</b>	0.730	0.900	0.690	0.700	0.720	0.710

Ventana (milisegs)	K1	K3	K5	K7	K9	K11
<b>1000</b>	0.366	0.373	0.480	0.507	0.474	0.481
<b>1500</b>	0.377	0.463	0.506	0.493	0.536	0.494
<b>2000</b>	0.297	0.279	0.330	0.376	0.443	0.466
<b>2500</b>	0.302	0.119	0.402	0.395	0.382	0.407

Ventana (milisegs)	K1	K3	K5	K7	K9	K11
<b>1000</b>	0.738	0.869	0.729	0.766	0.738	0.748
<b>1500</b>	0.779	0.798	0.808	0.837	0.846	0.798
<b>2000</b>	0.864	0.825	0.845	0.816	0.854	0.786
<b>2500</b>	0.930	0.860	0.820	0.800	0.810	0.770

Ventana	K1	K3	K5	K7	K9	K11
---------	----	----	----	----	----	-----

<b>(milisegs)</b>						
<b>1000</b>	0.286	0.153	0.299	0.293	0.313	0.313
<b>1500</b>	0.242	0.214	0.199	0.164	0.166	0.244
<b>2000</b>	0.148	0.173	0.144	0.180	0.154	0.257
<b>2500</b>	0.091	0.145	0.180	0.220	0.214	0.286

SVM Subconjunto de tipos de conducción

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
<b>1000</b>	0.607	0.626	0.682	0.701	0.682	0.710	0.720	0.664	0.673	0.673
<b>1500</b>	0.673	0.692	0.692	0.673	0.692	0.712	0.731	0.760	0.760	0.750
<b>2000</b>	0.689	0.767	0.748	0.738	0.777	0.757	0.777	0.738	0.718	0.709
<b>2500</b>	0.790	0.790	0.800	0.840	0.800	0.800	0.780	0.760	0.760	0.740

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
<b>1000</b>	0.548	0.487	0.366	0.353	0.338	0.319	0.305	0.358	0.352	0.352
<b>1500</b>	0.467	0.408	0.377	0.383	0.362	0.341	0.305	0.270	0.278	0.285
<b>2000</b>	0.437	0.303	0.307	0.297	0.263	0.276	0.255	0.289	0.310	0.317
<b>2500</b>	0.310	0.310	0.285	0.213	0.239	0.239	0.252	0.283	0.283	0.305

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
<b>1000</b>	0.626	0.673	0.692	0.710	0.710	0.710	0.701	0.682	0.692	0.710
<b>1500</b>	0.702	0.760	0.779	0.769	0.760	0.769	0.750	0.750	0.788	0.817
<b>2000</b>	0.718	0.796	0.796	0.806	0.806	0.825	0.825	0.816	0.825	0.825
<b>2500</b>	0.760	0.830	0.840	0.840	0.840	0.850	0.830	0.830	0.840	0.830

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
<b>1000</b>	0.522	0.441	0.407	0.381	0.374	0.360	0.367	0.380	0.366	0.333
<b>1500</b>	0.425	0.339	0.303	0.302	0.309	0.295	0.315	0.308	0.259	0.224
<b>2000</b>	0.393	0.275	0.267	0.252	0.244	0.215	0.206	0.213	0.190	0.190

<b>2500</b>	0.348	0.238	0.223	0.213	0.223	0.198	0.210	0.201	0.195	0.210
-------------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

### Red Neuronal de Perceptrón Múltiple Subconjunto de tipos de conducción

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
<b>1000</b>	0.664	0.579	0.645	0.579	0.607	0.645	0.682
<b>1500</b>	0.750	0.712	0.750	0.702	0.721	0.692	0.702
<b>2000</b>	0.786	0.748	0.670	0.689	0.709	0.709	0.709
<b>2500</b>	0.800	0.800	0.790	0.800	0.840	0.790	0.780

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
<b>1000</b>	0.372	0.465	0.385	0.458	0.439	0.371	0.352
<b>1500</b>	0.246	0.341	0.277	0.332	0.319	0.331	0.332
<b>2000</b>	0.224	0.266	0.359	0.329	0.300	0.317	0.325
<b>2500</b>	0.239	0.211	0.236	0.230	0.195	0.236	0.252

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
<b>1000</b>	0.692	0.692	0.720	0.710	0.748	0.757	0.738
<b>1500</b>	0.798	0.788	0.779	0.808	0.808	0.827	0.808
<b>2000</b>	0.738	0.806	0.816	0.864	0.825	0.825	0.835
<b>2500</b>	0.810	0.840	0.900	0.890	0.870	0.910	0.920

<b>Ventana (milisegs)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
<b>1000</b>	0.339	0.359	0.353	0.319	0.293	0.293	0.313
<b>1500</b>	0.198	0.220	0.242	0.222	0.222	0.194	0.215
<b>2000</b>	0.281	0.178	0.180	0.131	0.190	0.182	0.167
<b>2500</b>	0.251	0.167	0.110	0.117	0.139	0.095	0.079

### Random Forest Subconjunto de tipos de conducción



<b>2000</b>	0.757	0.767	0.748	0.757	0.738	0.738	0.738	0.738	0.738	0.738
<b>2500</b>	0.740	0.720	0.720	0.720	0.720	0.720	0.730	0.730	0.730	0.730

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
<b>1000</b>	0.373	0.379	0.379	0.366	0.352	0.352	0.352	0.352	0.352	0.352
<b>1500</b>	0.344	0.316	0.302	0.302	0.275	0.275	0.275	0.275	0.275	0.275
<b>2000</b>	0.268	0.278	0.283	0.268	0.272	0.272	0.272	0.272	0.272	0.272
<b>2500</b>	0.305	0.318	0.318	0.318	0.318	0.318	0.302	0.302	0.302	0.302

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
<b>1000</b>	0.570	0.561	0.561	0.570	0.570	0.570	0.570	0.570	0.570	0.570
<b>1500</b>	0.731	0.731	0.731	0.731	0.740	0.740	0.740	0.740	0.740	0.740
<b>2000</b>	0.786	0.786	0.786	0.786	0.786	0.796	0.796	0.796	0.796	0.796
<b>2500</b>	0.780	0.790	0.800	0.810	0.810	0.810	0.820	0.820	0.820	0.820

<b>Ventana (milisegs)</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>25</b>	<b>30</b>	<b>35</b>	<b>40</b>	<b>45</b>	<b>50</b>
<b>1000</b>	0.458	0.458	0.458	0.424	0.424	0.424	0.424	0.424	0.424	0.424
<b>1500</b>	0.298	0.298	0.298	0.298	0.276	0.276	0.276	0.276	0.276	0.276
<b>2000</b>	0.224	0.224	0.224	0.224	0.224	0.209	0.209	0.209	0.209	0.209
<b>2500</b>	0.279	0.264	0.248	0.242	0.242	0.242	0.226	0.226	0.226	0.226

Naïve Bayes Subconjunto de tipos de conducción

<b>Ventana (milisegs)</b>	<b>Resultado</b>
<b>1000</b>	0.673
<b>1500</b>	0.750
<b>2000</b>	0.786
<b>2500</b>	0.810

<b>Ventana (milisegs)</b>	<b>Resultado</b>
---------------------------	------------------

<b>1000</b>	0.400
<b>1500</b>	0.308
<b>2000</b>	0.265
<b>2500</b>	0.242

<b>Ventana (milisegs)</b>	<b>Resultado</b>
<b>1000</b>	0.692
<b>1500</b>	0.769
<b>2000</b>	0.796
<b>2500</b>	0.800

<b>Ventana (milisegs)</b>	<b>Resultado</b>
<b>1000</b>	0.387
<b>1500</b>	0.287
<b>2000</b>	0.250
<b>2500</b>	0.230



Anexo 4. Artículo científico.

Se anexa artículo científico en formato PDF sometido y expuesto en el evento denominado SEGUNDO WORKSHOP EN SISTEMAS INFOTAINMENT Y SISTEMAS INTELIGENTES – 2WOSI2.