



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

**METODOLOGÍA PARA LA
EVALUACIÓN DE LA CONDUCCIÓN DE
CHOFERES DEL TRANSPORTE
PÚBLICO A PARTIR DEL USO DE
DISPOSITIVOS INERCIALES**

HERNÁN FELIPE CATALÁN SALAS

Tesis para optar al grado de:
Magíster en Ciencias de la Ingeniería

Profesor Supervisor:
JUAN CARLOS HERRERA MALDONADO

Santiago de Chile, Mayo, 2018
© 2018, Hernán Catalán Salas



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

**METODOLOGÍA PARA LA
EVALUACIÓN DE LA CONDUCCIÓN DE
CHOFERES DEL TRANSPORTE
PÚBLICO A PARTIR DEL USO DE
DISPOSITIVOS INERCIALES**

HERNÁN FELIPE CATALÁN SALAS

Tesis presentada a la Comisión integrada por los profesores:

JUAN CARLOS HERRERA MALDONADO

HANS LOBEL DÍAZ

CÉSAR NUÑEZ SEPÚLVEDA

HERNÁN SANTA MARÍA OYANEDEL

Para completar las exigencias del grado de
Magíster en Ciencias de la Ingeniería

Santiago de Chile, Mayo, 2018

*A todos aquellos que han
participado en la serie que es mi vida.*

AGRADECIMIENTOS

Primero que todo, quisiera agradecer a los profesores Juan Carlos Herrera y Hans-Albert Löbel por todo su apoyo y guía a lo largo del desarrollo de esta investigación.

Además, me gustaría agradecer a la profesora Patricia Galilea, quien logró conquistarme con la mención de Transporte, una especialidad que se convertiría en mi vocación. A Ignacio Tiznado por llevarme a ella ante mi crisis vocacional.

No puedo dejar de agradecer a mis mejores amigos: Rodrigo González, Sebastián Muñoz, Lorena Silva, Fernanda Corvillón y Tomás Schmitz. Cada uno de ellos estuvo siempre dispuesto a ayudarme con mis decisiones más difíciles.

También quisiera agradecer a todos los amigos que formé en Transporte: Mariana, Nicolás, Ignacio, Owen, Felipe, Jaime, Elizabeth y Francisco. Los momentos vividos durante el Magíster fueron inolvidables.

Finalmente, me gustaría agradecer a mi familia: mis papás, mi hermano y mi abuela, sin ellos no estaría aquí. Su gran esfuerzo por entregarme un futuro mejor es el mejor regalo que alguien podría haberme dado.

A todos, gracias, un abrazo y para lo que me necesiten...aquí estoy.

TABLA DE CONTENIDO

1	INTRODUCCIÓN	1
1.1	Hipótesis	2
1.2	Objetivos.....	3
1.3	Estructura de la Tesis	3
2	REVISIÓN BIBLIOGRÁFICA.....	5
2.1	Calidad de Servicio	5
2.2	Modelación de la Calidad de Servicio y el rol de la conducción.....	7
2.3	Uso de dispositivos inerciales en el transporte público.....	11
2.4	Índice de Calidad de Atención al Usuario de Transantiago	18
3	METODOLOGÍA.....	23
3.1	Diseño del Experimento	23
3.1.1	Herramientas Utilizadas.....	24
3.1.2	Recolección de datos	25
3.1.3	Procesamiento de datos	28
3.2	Métodos de Clasificación	32
3.2.1	Máquinas de Vectores de Soporte (SVM).....	33
3.2.2	Árboles de Decisión	35
3.2.3	K vecinos más cercanos (KNN).....	37
3.2.4	Caso de la muestra no balanceada.....	39
3.2.5	Elección de Hiperparámetros.....	40
3.3	Modelación	42
3.3.1	Medidas de Rendimiento	42
3.3.2	Calibración de los modelos.....	43
4	RESULTADOS	51
4.1	Rendimiento de los Métodos de Clasificación utilizados.....	51
4.1.1	SVM	51
4.1.2	Árboles de decisión	53

4.1.3	KNN	57
4.2	Comparación de los métodos calibrados	58
4.3	Análisis de la clasificación realizada	60
5	CONCLUSIONES	68
	REFERENCIAS	73

ÍNDICE DE ILUSTRACIONES

Figura 2-1. Círculo vicioso del transporte público.....	6
Figura 2-2. Modelos estructural que explica la relación entre calidad de servicio y el comportamiento de los usuarios hacia el sistema.....	7
Figura 2-3. Atributos de Calidad de Servicio por cuadrantes	9
Figura 2-4. Transformada de Fourier aplicada a los valores de la aceleración según su frecuencia.....	14
Figura 2-5. Relación lineal entre la percepción del panel de evaluación y el indicador propuesto.....	16
Figura 2-6. Mapa temático según los valores de N_{VD} para un recorrido de prueba.....	17
Figura 2-7. Histograma de los valores de N_{VD} para un recorrido de prueba.....	18
Figura 3-1. Ejemplo de la interfaz de la aplicación SensorLog utilizada en el experimento	24
Figura 3-2. Mapa de tramos evaluados durante el experimento	26
Figura 3-3. Ejes de la aceleración capturada por el dispositivo, según la dirección del bus	27
Figura 3-4. Movimientos de rotación del giroscopio capturados por el dispositivo	28
Figura 3-5. Aceleración longitudinal del bus 213E.....	29
Figura 3-6. RMS de la Aceleración longitudinal con las alertas reportadas (puntos sobre la curva)	30
Figura 3-7. Estructura de la nueva de base de datos creada para incorporar la progresión de los atributos en el tiempo.....	31
Figura 3-8. Ejemplo de un árbol de decisión que clasifica en dos grupos.....	36
Figura 3-9. Ejemplo de un algoritmo KNN con dos valores de constante	38
Figura 3-10. Descripción gráfica de una validación cruzada de 3 iteraciones.....	41
Figura 3-11. Variación de las medidas de desempeño según el valor del parámetro C1 (SVM).....	45

Figura 3-12. Variación de las medidas de desempeño según el valor del parámetro C1, -1 (Árboles de Decisión).....	47
Figura 3-13. Variación de las medidas de desempeño según el valor del parámetro C1 (KNN).....	49
Figura 4-1. Árbol de decisión calibrado para la combinación 1	54
Figura 4-2. Box plots de las variables consideradas en el método de árboles de decisión	55
Figura 4-3. Trayectoria del bus 213E	62
Figura 4-4. Trayectoria del bus 213E en coordenadas oblicuas.....	63
Figura 4-5. Curva por la cual transitan los buses 213E, 210 y 210V.....	65
Figura 4-6. Velocidad a lo largo de la curva por la cual transitan los buses 213E y 210V	66

ÍNDICE DE TABLAS

Tabla 2-1. Dimensiones del modelo SERVQUAL.....	8
Tabla 2-2. Definición Índice de Comodidad según la vibración experimentada.....	13
Tabla 2-3. Percepción de la comodidad según valor de N_{VD}	16
Tabla 2-4. Atributos para evaluar calidad de atención al usuario en ruta.....	19
Tabla 2-5. Estructura de Descuentos según ICA obtenido	21
Tabla 2-6. Nivel de cumplimiento según atributo evaluado en el ICA para el año 2015 22	
Tabla 3-1. Detalle de cada tramo utilizado para la recolección de datos.....	25
Tabla 3-2. Intersecciones o puntos de referencia del inicio y término de cada recorrido	27
Tabla 3-3. Matriz de Costos usada en árboles de decisión y KNN.....	39
Tabla 3-4. Matriz de Confusión	42
Tabla 3-5. Exactitud del modelo según tipo de Kernel usado.....	44
Tabla 3-6. Parámetro C escogido según caso y para cada combinación	46
Tabla 3-7. Parámetros escogidos del modelo árboles de decisión para cada combinación	47
Tabla 3-8. Parámetros escogidos del modelo KNN para cada combinación según caso	49
Tabla 4-1. Resultados del modelo SVM para ambos casos	51
Tabla 4-2. Resultados del método árboles de decisión.....	53
Tabla 4-3. Proporción de maniobras clasificadas como inapropiadas según su posible explicación	57
Tabla 4-4. Resultados del modelo KNN para el caso 1	58
Tabla 4-5. Tabla resumen del rendimiento obtenido por cada método de clasificación .	59
Tabla 4-6. Resultado de la clasificación según recorrido	61
Tabla 4-7. Proporción de maniobras bruscas dada una reducción de velocidad significativa	63
Tabla 4-8. Velocidad comercial promedio para los recorridos que comparten trayectoria	64

RESUMEN

Actualmente el sistema de transporte público de Santiago emplea un Índice de Calidad de Atención al Usuario (ICA) que aborda superficialmente la calidad de servicio. Además, la metodología actual no permite realizar un seguimiento ni incentiva la mejora continua del servicio.

El objetivo de esta investigación es utilizar datos de acelerómetros y giroscopios para formular una nueva metodología para la clasificación del atributo de conducción del ICA de Transantiago. Para probar la metodología propuesta se diseñó un experimento donde se recolectó información de 5 recorridos de diferentes operadores y con una duración similar. Luego, se evaluaron 3 métodos de clasificación: máquinas de soporte vectorial (SVM), árboles de decisión y k vecinos más cercanos (KNN).

Los resultados de la modelación muestran que SVM tiene el mejor rendimiento, árboles de decisión entrega un modelo muy interpretable y KNN es el método más simple de aplicar. Respecto a la conducción, los resultados muestran que la mayor cantidad de maniobras bruscas se debe a variaciones de velocidad. Los recorridos con peor conducción fueron 210 y 210V, donde más de la mitad de las variaciones de velocidad fueron clasificadas como bruscas. En cambio, el bus 213E obtuvo resultados bastante mejores incluso en el mismo trayecto que los recorridos anteriormente mencionados.

Por otro lado, el uso del GPS y etiquetar según conductor permite identificar focos a mejorar el servicio mediante capacitación y, en casos puntuales, modificaciones al trazado del recorrido.

Finalmente, la metodología es replicable más allá de Transantiago. Como beneficio, se requiere una pequeña muestra para calibrar un modelo de buen rendimiento. Como futura línea de investigación, se propone evaluar la calibración de modelos diferentes para casos particulares (recorridos en pendiente, en autopista, etc.). Es importante señalar que los modelos calibrados son válidos para el experimento realizado. Para aplicarlo a mayor escala solo se debe seguir la metodología expuesta.

ABSTRACT

Currently, Santiago's public transit system employs an User Service Quality Index (ICA) to measure the quality of service offered in a superficial manner. Additionally, the existing methodology does not allow for a proper follow-up nor it incentives for the continuous improvement of the service provided.

The main objective of this research is to formulate a new methodology that uses the information provided by accelerometers and gyroscopes to classify the driving style attribute considered on ICA from Transantiago. To test the proposed methodology, an experiment was designed to gather data from 5 bus lines belonging to different operators and with a similar travel duration. After this, 3 classification methods were evaluated: support vector machines (SVM), decision trees and k-nearest neighbours (KNN).

The results of the modelling show that the SVM method provides the best performance, the use of decision trees delivers a very interpretable model and that the KNN method is simplest one to apply. Regarding the driving classification, the results show that the greatest number of abrupt manoeuvres is due to speed variations. The bus lines with the worst driving style were 210 and 210V, where more than half of the speed variations were classified as abrupt. On the other hand, the bus line 213E obtained better results even on the same route as aforementioned bus lines.

Furthermore, the use of GPS and labelling according to the driver allows to identify areas to improve the service through focalised training and, in specific cases, modifications to the route.

Finally, the methodology is replicable beyond the case of Transantiago. As an advantage, a small samples is enough to calibrate a good performance model. As a future line of research, the author propose to calibrate different models for particular cases (such as routes with slopes, motorways, etc.). It is important to note that the calibrated models shown on this thesis are valid for the data gathered from the experiment. To apply it on a larger scale, the proposed methodology should be followed.

1 INTRODUCCIÓN

El transporte público en Santiago ha sido tema de conversación y debate durante varias décadas debido a los importantes y complejos problemas que presenta. Durante ese período, el sistema ha sufrido múltiples modificaciones para lograr un servicio de calidad. A pesar de ello, la percepción de los usuarios sigue siendo regular. El 2016, en una escala de 1 a 7, los usuarios evaluaron al sistema con un 4,3. Cuando se toma en cuenta la percepción por los recorridos, en vez del sistema en general, esta nota sube a 4,9 (DTPM, 2016a). Las autoridades consideran esta última como más apropiada, ya que tiene una mayor correlación con los atributos operacionales del servicio, mientras que la evaluación general tiene una componente más social de rechazo (DTPM, 2015).

Con el objetivo de mejorar el servicio entregado a los usuarios, el Gobierno ha incorporado gradualmente indicadores en los contratos de concesión para asegurar un nivel de servicio mínimo. Uno de ellos corresponde al Índice de Calidad de Atención al Usuario (ICA) que entra en funcionamiento el 2012 y consideraba 14 atributos asociados al servicio entregado al usuario, tanto como el que va a bordo del bus como el que espera en el paradero (DTPM, 2014).

Para su medición, se utiliza la metodología de pasajero incógnito, quien debe evaluar cada atributo de manera binaria: cumple o no cumple. El ICA para cada bus se calcula como el promedio simple del valor de cada atributo. En el contrato de licitación se estipula que el nivel de cumplimiento mínimo es 85%. Cuando no se llega a este nivel los operadores se ven expuestos a descuentos monetarios.

Si bien este enfoque es un buen paso hacia formular un indicador que asegure un mejor nivel de servicio, la metodología empleada presenta problemas importantes. Usar una persona como pasajero incógnito implica una subjetividad en la medición. Además, como la toma de datos no es continua y la muestra se escoge de manera aleatoria, se corre el riesgo de que el índice entregue valores alejados a la realidad. Esto también impide que se realice seguimiento a los conductores para detectar focos de buenos o malos resultados con el fin de entregar incentivos o capacitaciones, según corresponda.

Por otro lado, al calcularse el indicador como el promedio simple de los 14 atributos, el cumplimiento global no refleja el real cumplimiento de la forma de conducción. Este atributo tuvo un promedio de cumplimiento de 55% durante el 2015. En cambio, el atributo asociado a la amabilidad del conductor llega a 99% de cumplimiento (DTPM, 2016b).

La forma en que el chofer del bus conduce es de gran relevancia para la percepción que el usuario tiene del servicio. Del total de reclamos del año 2016, el 27% corresponden al manejo y comportamiento del conductor, siendo la segunda razón más frecuente de reclamos (después de la no detención en paraderos con un 40% de los reclamos). Es claro que la forma de conducción de los choferes del Transantiago no es adecuada, pero bajo la metodología actual no se incentiva a los operadores a mejorarla.

En la última década y con el desarrollo de nuevas tecnologías, el uso de dispositivos inerciales (IMU por su sigla en inglés) se ha masificado debido a su utilidad y bajo costo. Un dispositivo inercial es aquel que posee: giroscopio, acelerómetro y GPS. Permite capturar cuantitativamente el movimiento, aceleración y vibraciones experimentadas en los tres ejes por un objeto, en este caso, un bus. Con esta información se puede formular un indicador de calidad de servicio de una forma más objetiva que con un pasajero incógnito esporádico.

1.1 Hipótesis

La hipótesis de esta tesis es que se puede inferir el aspectos del estilo de conducción de un chofer de bus de transporte público a partir de las vibraciones y movimientos experimentados a bordo del vehículo. Las vibraciones y movimientos a considerar son una consecuencia directa de la conducción del chofer y no dependen de factores externos como el estado del bus.

1.2 Objetivos

El objetivo general de esta tesis es formular una nueva metodología para el cálculo del atributo de conducción del ICA de Transantiago, basada en datos recolectados con dispositivos inerciales. Esto se logra al clasificar la forma de conducción tomando en cuenta las vibraciones, aceleraciones y movimientos que experimentan los pasajeros a bordo de los buses.

Adicionalmente, esta investigación tiene los siguientes objetivos específicos:

- Evaluar la factibilidad de usar métodos de aprendizaje supervisado para la clasificación.
- Validar la metodología con datos reales, pero en una muestra controlada.
- Determinar fortalezas y limitaciones de los métodos aplicados.

Es relevante destacar que no es un objetivo de esta tesis calibrar un modelo para ser aplicado a todo el sistema de Transantiago.

1.3 Estructura de la Tesis

Además de este primer capítulo, esta tesis considera otros cinco. En el segundo capítulo se presenta una detallada revisión bibliográfica, orientada a comprender cuáles son los atributos considerados para describir la calidad de servicio, la importancia de la comodidad dentro de esos atributos y los distintos usos para los dispositivos inerciales en el transporte público (con un enfoque en la comodidad). Además, se hace una revisión detallada de la metodología para calcular el ICA de Transantiago.

En el tercer capítulo se explica la teoría detrás de los métodos de clasificación utilizados (SVM, árboles de decisión y KNN). El cuarto capítulo describe la metodología utilizada, incluyendo las herramientas utilizadas, una descripción de la recolección de datos, del procesamiento de datos y la modelación realizada. El capítulo cinco expone los

resultados de la modelación. Finalmente, el capítulo seis menciona las conclusiones y los alcances del trabajo de investigación. Además, se presentan futuras líneas de investigación.

2 REVISIÓN BIBLIOGRÁFICA

En este capítulo se presenta la revisión bibliográfica realizada en el contexto de la investigación. Se comienza con una revisión sobre la calidad de servicio y el impacto que tiene en los usuarios o consumidores. Luego, se exponen los trabajos de distintos autores que han modelado la calidad de servicio de distintos sistemas de transporte público. Además, se realiza un énfasis en el rol que juega la conducción en cada modelo. En tercer lugar, se presentan distintas investigaciones que utilizan dispositivos inerciales en sistemas de transporte público. Finalmente, se expone la metodología utilizada por Transantiago en el cálculo de su indicador de calidad de atención a los usuarios.

2.1 Calidad de Servicio

La definición más aceptada de calidad de servicio fue entregada por Parasuraman, Zeithaml & Berry (1985), quienes postulan que la percepción de la calidad de servicio (QoS) es la diferencia entre la percepción del servicio recibido (PS) y el servicio esperado (ES). Esto se ve reflejado en la ecuación 2.1.

$$QoS = PS - ES \quad (2.1)$$

Asimismo, una buena percepción de la calidad del servicio impacta positivamente en la demanda del transporte público, lo que ha sido ampliamente estudiado en la literatura. El “círculo vicioso del transporte público” (ver Figura 2-1) es una representación inversa de esta situación. Una baja calidad de servicio, representada en este caso por la congestión a la que los buses se ven expuestos, es uno de los factores que reduce el uso de los buses favoreciendo otros medios de transporte como el automóvil (Ortúzar & Willumsen, 2011).

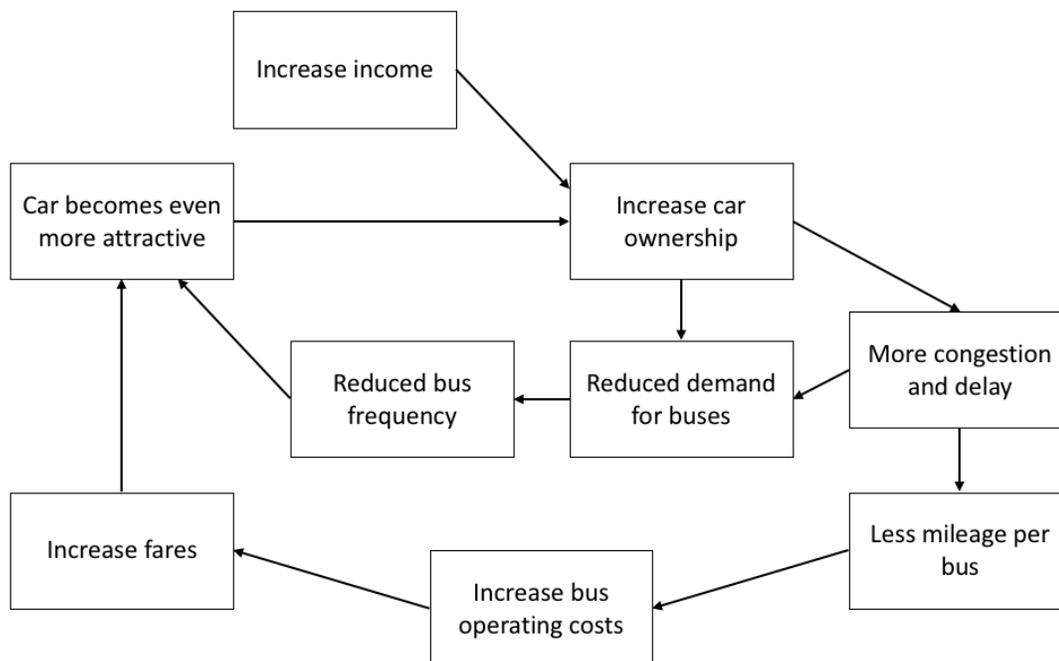


Figura 2-1. Círculo vicioso del transporte público. Fuente: Ortúzar & Willumsen, 2011.

Por otro lado, Cascetta & Carteni (2014) investigan los efectos que tuvieron políticas de transporte de largo plazo aplicadas en la región de Campania, Italia. Luego de realizar una comparación de la demanda a lo largo de la década estudiada, se concluye que las políticas generaron una nueva demanda igual al 10,7% de la actual. Se condujo, además, una encuesta de satisfacción que determinó que las razones para ese incremento de demanda se debían a: disminución de la congestión (23%), aumento de la comodidad (22%), disminución del tiempo de viaje (22%), entre otros.

En Lai & Chen (2011) se encuesta a los usuarios del sistema de metro de Kaohsiung, China, acerca de distintos atributos asociados a la calidad del servicio entregado. Además, se les realiza diferentes preguntas relacionadas a la apreciación personal respecto del transporte público en general. Luego de realizar un análisis factorial de las respuestas, se escogen los atributos significativos y se construye un modelo estructural (ver Figura 2-2), cuyas relaciones resultan significativas. Se aprecia cómo la calidad de servicio (*Service*

Quality) influye directamente en la satisfacción, el valor percibido, la pertenencia y el comportamiento de los usuarios hacia el sistema de transporte público.

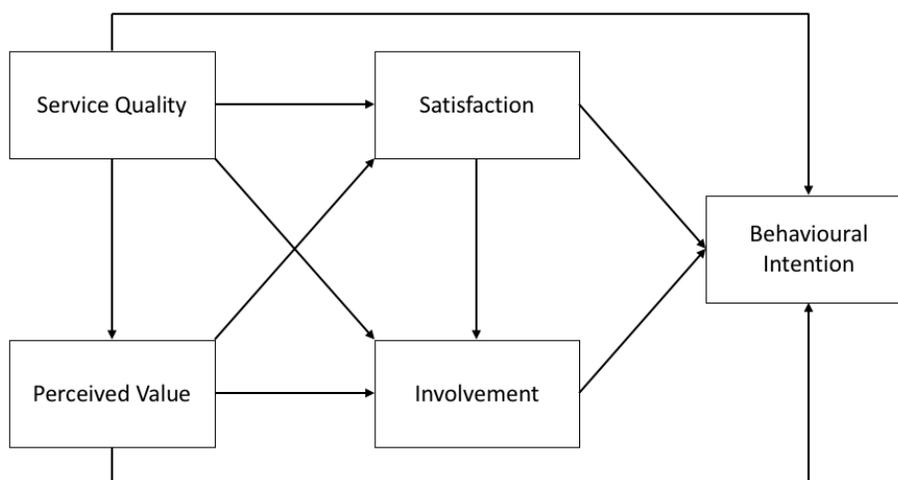


Figura 2-2. Modelos estructural que explica la relación entre calidad de servicio y el comportamiento de los usuarios hacia el sistema. Fuente: Lai & Chen, 2011.

Los autores concluyen que enfocar la calidad de servicio en el usuario es crucial para que las empresas logren cumplir las expectativas de sus usuarios. Con esto, se logra reforzar su lealtad y, mediante recomendación, se puede aumentar la demanda del servicio.

2.2 Modelación de la Calidad de Servicio y el rol de la conducción

Dado que la calidad de un servicio influye en la demanda de este, es relevante comprender aquellos atributos o variables que la explican. Uno de los primeros trabajos en este campo es el desarrollado por Parasuraman, Zeithaml & Berry (1988). Los autores calibran un modelo (llamado SERVQUAL) en base a un cuestionario de 22 preguntas enfocadas a comprender la percepción de usuarios de tiendas comerciales sobre la calidad

del servicio entregada. Estas preguntas se agrupan en cinco dimensiones (ver Tabla 2-1) que resultaron ser estadísticamente significativas.

Tabla 2-1. Dimensiones del modelo SERVQUAL. Fuente: Parasuraman *et al.*, 1988.

Dimensiones	Explicación
Tangibles	La apariencia de las instalaciones físicas, los equipos, el personal y los materiales de comunicación.
Confiabilidad	La habilidad de realizar el servicio prometido de manera fiable y precisa.
Capacidad de respuesta	La disposición a ayudar a los clientes y proporcionar un servicio rápido.
Confianza en la atención	El conocimiento y cortesía de los empleados y su capacidad para transmitir confianza.
Empatía	La prestación de una atención preocupada e individualizada a los clientes.

Dada la definición de percepción de calidad de servicio expuesta (ecuación 2.1), se calcula la diferencia para cada dimensión como el promedio de la diferencia entre percepción y expectativa de cada atributo que compone dicha dimensión. Si esta diferencia es negativa, significa que la empresa debe tomar medidas para mejorar el servicio entregado. Si bien este modelo ha sido utilizado en variados rubros con buenos resultados, cada dimensión es muy abstracta y no es posible cuantificarlas. Esto se vuelve relevante cuando se habla de transporte público, ya que el servicio entregado es a gran escala y replicar dicho modelo se vuelve muy costoso.

En Cronin & Taylor (1992), se cuestiona el enfoque del modelo SERVQUAL, ya que consideran que la definición de calidad de servicio (ecuación 2.1) es inapropiada. Los autores desarrollan un modelo basado en SERVQUAL, pero que no toma en cuenta las expectativas del consumidor al momento de definir la calidad del servicio (el modelo se llama SERVPERF). El estudio se aplica a consumidores de ocho empresas pertenecientes a cuatro industrias distintas en una ciudad del sureste de Estados Unidos. Los autores

concluyen que el enfoque en el rendimiento real y no en las expectativas tiene mejores resultado, ya que explica de mejor manera la variabilidad de la calidad de servicio que el modelo SERVQUAL. Además, determinan que la percepción de la calidad de servicio impacta en la satisfacción del consumidor y, finalmente, es esta dimensión la que mayor relación tiene con la intención de compra del cliente.

Estudios más recientes modifican el enfoque de modelación, ya que el consenso es descifrar la importancia que el usuario le entrega a cada atributo particular que define al servicio entregado.

En Grujičić *et al.* (2014), se realiza un Análisis de Importancia-Rendimiento (IPA, por sus siglas en inglés; Martilla & James, 1977) con los datos de una encuesta aplicada a 449 usuarios de transporte público en Belgrado, Serbia. La metodología considera 24 atributos asociados a la calidad del servicio, donde deben evaluar su importancia y nivel de satisfacción (en una escala de 1 a 5). Con el promedio de las respuestas por atributo se genera un mapa de cuadrantes según nivel de satisfacción e importancia (ver Figura 2-3).

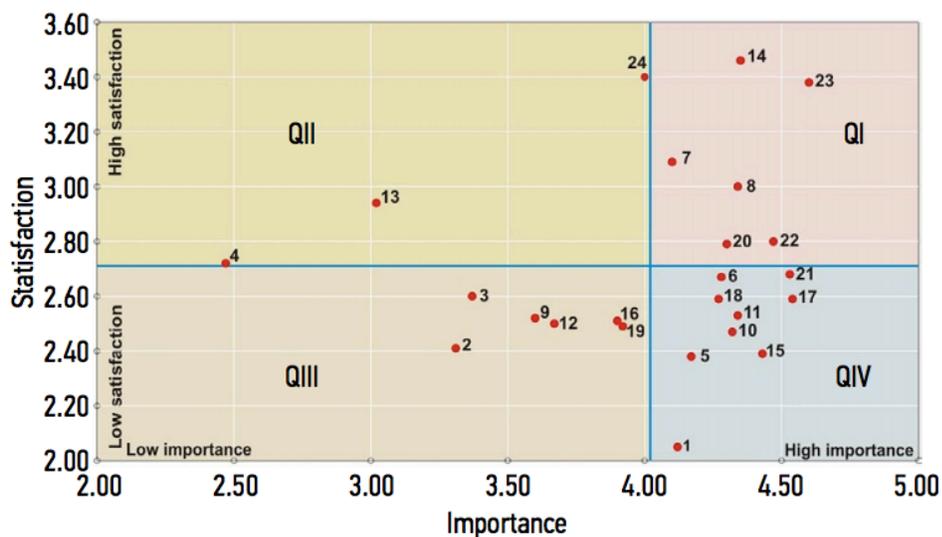


Figura 2-3. Atributos de Calidad de Servicio por cuadrantes. Fuente: Grujičić *et al.*, 2014.

El cuadrante IV de la Figura 2-3 corresponde a aquel que contiene los atributos que deben ser mejorados, al ser muy importantes y con una satisfacción baja. Los tres atributos con menor satisfacción de este cuadrante son: espacio disponible en el vehículo (1), el precio del pasaje (15) y la limpieza de los demás pasajeros (5). Este último atributo tiene una alta importancia declarada debido a que la encuesta se realizó en los meses de verano. Por otro lado, si solo se observan los atributos más importantes (cuadrantes I y IV) tenemos que una conducción segura es el más importante, pero para el caso estudiado tenía un buen nivel de satisfacción.

En de Oña *et al.* (2015), se aplica un cuestionario a 200 usuarios del Metro de Sevilla, España, quienes debían calificar 36 atributos asociados a la calidad del servicio (en la escala del 1 al 10). Además se les preguntaba su nivel de satisfacción global con el servicio y su disposición a volver a utilizarlo. Los autores utilizaban el Análisis de Componentes Principales (PCA, por su sigla en inglés) para agrupar los atributos que estuvieran correlacionados. Con esto se llega a ocho variables explicativas.

Como conclusión se tiene que la variable “Servicio Tangible” es la que más relación tiene con la evaluación global del servicio. Además, dentro de esta variable se encuentra el atributo “conducción apropiada”. Si bien, al observar la importancia de los 36 atributos la conducción no es la más relevante, el estudio si demuestra una relación positiva con la percepción de la calidad de servicio. Esto se puede deber a que en un servicio de Metro la conducción no es tan relevante, ya que no hay adelantamientos, ni virajes ni detenciones inesperadas como en los buses.

Este trabajo también demuestra que la satisfacción con el servicio impacta drásticamente y positivamente la disposición a volver a utilizar el metro para los usuarios no captivos. Esa conclusión refuerza la importancia que tiene mejorar el transporte público para atraer más usuarios.

En el trabajo realizado por dell’Olio *et al.* (2011) se aplicó una encuesta a los usuarios de transporte público de Santander, España, que evaluaba seis variables asociadas a la calidad del servicio. Se escogieron estas seis luego de realizar un *focus group*. Con los resultados se formuló un modelo de elección discreta multinomial y se concluyó que

la variable de comodidad presenta una contribución alta a la función de utilidad dentro de todos los usuarios. A pesar de ello, se nota un incremento en la contribución para los usuarios mayores de 60 años y aquellos usuarios casuales del servicio (no cautivos). Esto tiene sentido debido a que los primeros presentan movilidad reducida por lo que la disponibilidad de asientos y una conducción apropiada se vuelven relevantes. En los segundos, un servicio cómodo se vuelve competitivo respecto de otros modos de transporte (ej: automóvil). El tiempo de espera es la única variable del modelo que supera en contribución a la comodidad.

En Sánchez-Flores y Romero-Torres (2010), realizan un estudio en un corredor de transporte público de la ciudad de Toluca, México, que es servido por más de un recorrido de bus. Los autores realizan una encuesta de preferencias declaradas donde cada usuario debía escoger entre el servicio actual y un teórico “autobús mejorado” que cobra una tarifa más alta. Se consideraron 13 variables asociadas a la calidad del servicio, obtenidas luego de realizar un *focus group*. Luego de aplicar modelos logit binarios, los autores concluyen que la forma de conducción es el segundo atributo más significativo para los usuarios al momento de escoger un recorrido, luego del estado físico del autobús. Este análisis excluye el costo y tiempo de viaje, ya que al ser considerados son estas variables las que más impactan en la función de utilidad.

A lo largo de las investigaciones expuestas en esta sección se refuerza la importancia que tiene la conducción en la percepción de calidad por parte de los usuarios, siendo uno de los atributos más relevantes.

2.3 Uso de dispositivos inerciales en el transporte público

Un dispositivo inercial (IMU por su sigla en inglés) es un dispositivo de alta precisión que posee: giroscopio, acelerómetro y GPS. Permite capturar cuantitativamente el movimiento, aceleración y vibraciones experimentadas en los tres ejes por un objeto. Al ser dispositivos de bajo costo, existe una versatilidad en el uso de los IMU. A continuación se presentará cómo se emplean en algunos sistemas de transporte público.

Han (2013) estudia el caso de los trenes de alta velocidad en Corea que deben reducir su velocidad en las curvas. Debido a que el objetivo de este medio de transporte es la velocidad, algunos incorporan sistemas con dispositivos inerciales para medir y contrarrestar la inclinación del vagón en las curvas. De esta manera los pasajeros no experimentan incomodidad y el tren puede tomar las curvas a mayor velocidad.

El autor propone utilizar, además, el GPS para almacenar la información del dispositivo inercial de manera geolocalizada y así predecir los lugares donde se requerirá contrarrestar la inclinación. De esta manera se comienza este proceso antes de que sea requerido y la sensación para el usuario no es tan brusca, al contrario de la situación actual que es reactiva (se activa cuando se superan umbrales de aceleración en cada eje).

Al poner la metodología a prueba en la línea Gyeongbu (entre Seúl y Busan) se llega a que la velocidad es entre un 20% y un 30% mayor a los trenes convencionales. Además, se consiguen resultados un 30% mejor en los indicadores de comodidad (que miden aceleraciones a bordo) según la norma ENV 12299 (BSI, 1999).

En Zhang *et al.* (2014) utilizan dispositivos capaces de capturar el ruido, la temperatura, la vibración y la aceleración experimentada a bordo de 13 líneas de bus del sistema de transporte público de la Nanjing, China. Se les realizó un cuestionario para evaluar la percepción de los usuarios respecto de la calidad de servicio del sistema y del recorrido específico.

Para cuantificar la comodidad percibida, los autores emplean distintas metodologías. Primero, para abordar la aceleración experimentada, los autores utilizan el concepto de *jerk*, es decir, la tasa a la cual cambia la aceleración. Definen $\pm 0,6 \text{ m/s}^3$ como el umbral del *jerk*, lo que se traduce en que valores mayores (o menores, en caso de un *jerk* negativo) implican cambios en la aceleración incómodos para el usuario.

Por otro lado, se mide la vibración a bordo basándose en la norma ISO 2631 (ISO, 1997) que clasificar el nivel de comodidad experimentado según la vibración medida en RMS (*Root Mean Square*, m/s^2), como se aprecia en la Tabla 2-2. Es importante notar que G corresponde a la aceleración de gravedad ($9,81 \text{ m/s}^2$).

Tabla 2-2. Definición Índice de Comodidad según la vibración experimentada.
Fuente: Castellanos, Susin y Fruett (2011).

Vibración (G)	Percepción
< 0,032	Cómodo
0,032 – 0,057	Un poco incómodo
0,057 – 0,092	Suficientemente incómodo
0,092 – 0,145	Incómodo
0,145 – 0,229	Muy incómodo
>0,229	Extremadamente incómodo

Los autores concluyen que la temperatura, la aceleración y la vibración a bordo eran las razones que más impactaban en la evaluación de los usuarios a bordo del bus. Con los resultados de la encuesta se ajustó una regresión múltiple que probaron en un recorrido específico, donde cada vez que caía la comodidad se le enviaba una alerta al conductor. Los resultados mostraron que la calidad del servicio mejoró un 21% cuando se aplicaba el sistema de alertas en base a la regresión ajustada.

En Figueiredo y Silva (2015) se utiliza la misma norma ISO 2631 para evaluar las vibraciones que experimentan los usuarios y el chofer a bordo de un bus del sistema de transporte público de una ciudad de Minas Gerais, Brasil. Se instalan acelerómetros en el asiento del conductor y de uno de los pasajeros y se miden las vibraciones.

Se concluye que los valores para el conductor exceden los límites permitidos por la norma ISO, tanto en valor bruto como en la duración de la exposición. Además, observan que el asiento del conductor, a pesar de ser acolchado, no reduce las vibraciones respecto de los asientos de los usuarios (siendo que estos son estructuras rígidas). Los autores advierten que esto podría provocar severos problemas de dolor lumbar en los choferes. Por otro lado, se observa que los valores percibidos por el usuario son mayores a los permitidos por la norma, pero que la duración de la exposición no es suficiente para generar problemas de salud en el corto plazo.

En Castellanos, Susin y Fruett (2011) se realizó un estudio en la ciudad de Campinas, Brasil, con IMUs instalados en el piso del bus. Este dispositivo posee acelerómetros en los tres ejes y GPS. Los autores definen tres métodos que pondrán a prueba para ver el que mejor describa los casos de incomodidad asociada a la conducción. El método A establece un umbral de $\pm 0,5G$, donde G corresponde a la aceleración de gravedad. Es importante destacar que este umbral considera que a la aceleración del eje Z (perpendicular al piso del bus) se le reste $1G$ correspondiente a la gravedad.

El método B considera un umbral para el *jerk* longitudinal. Para evitar errores en su cálculo, los autores aplican un *Low Pass Filter* (LPF) con una frecuencia de 2,5 Hz. En la Figura 2-4 se puede apreciar los resultados de aplicar una transformada de Fourier (FFT, por sus siglas en inglés) a los valores de la aceleración. Los valores de baja frecuencia se asocian a movimientos relacionados con la conducción del chofer, mientras que las frecuencias entre 10 Hz y 40 Hz se pueden deber a las vibraciones intrínsecas del bus. Los autores determinan que 2,5 Hz es la frecuencia límite para el análisis. Finalmente, el umbral considerado para el *jerk* longitudinal es $\pm 0,4 \text{ m/s}^3$. Este criterio es más estricto que el utilizado por Zhang *et al.* (2014).

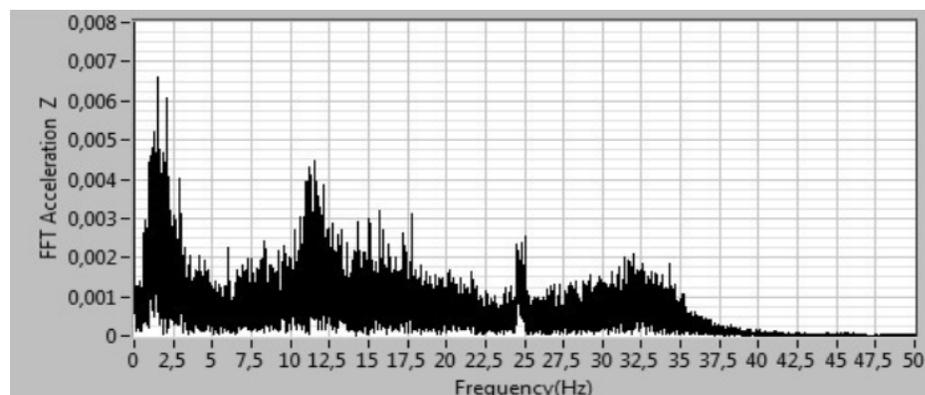


Figura 2-4. Transformada de Fourier aplicada a los valores de la aceleración según su frecuencia.
Fuente: Castellanos, Susin y Fruett, 2011.

Finalmente, el método C consiste en aplicar la norma ISO 2631 metodológicamente de la misma manera que en Zhang *et al.* (2014) y Figueiredo y Silva (2015). A diferencia de estos trabajos, el índice de comodidad se calcula cada 15 segundos y no de manera global. Los autores concluyen, luego de realizar pruebas con las tres metodologías, que el método C era el que mejor se relacionaba con la información reportada por los sujetos participantes del experimento (quienes reportaban los incidentes en una bitácora).

Finalmente, en Bodini *et al.* (2014) los autores adaptan la norma ENV 12299 que define indicadores de comodidad para el sistema ferroviario europeo. Esta norma considera indicadores para personas sentadas, de pie y un indicador global. En este trabajo se adapta, específicamente, el indicador para personas de pie (N_{VD}) según la ecuación 2.3.

$$N_{VD} = 3 \cdot \sqrt{16 \cdot a_x^2 + 4 \cdot a_y^2 + a_z^2} + 5 \cdot a_y \quad (2.3)$$

En esta ecuación tenemos que a_x es la aceleración longitudinal, a_y es la aceleración lateral y a_z es la aceleración vertical. A mayor valor de N_{VD} , mayor es la incomodidad percibida por los usuarios. El N_{VD} es calculado y almacenado en un servidor cada segundo con una posición asociada (gracias al GPS presente en el IMU), para luego analizar espacial y estadísticamente la información recolectada.

Los autores realizan un panel de prueba con 30 pasajeros para capturar su percepción del servicio respecto de distintas infraestructuras de un recorrido establecido (por ejemplo: calles rectas, cruces peatonales, rotondas, etc.). Se les solicitaba evaluar en una escala del 0 (muy bueno) al 4 (muy malo) después de cada ocurrencia (15 segundos de medición). Al mismo tiempo, el IMU instalado calculaba el N_{VD} asociado. Los resultados se pueden apreciar en la Figura 2-5.

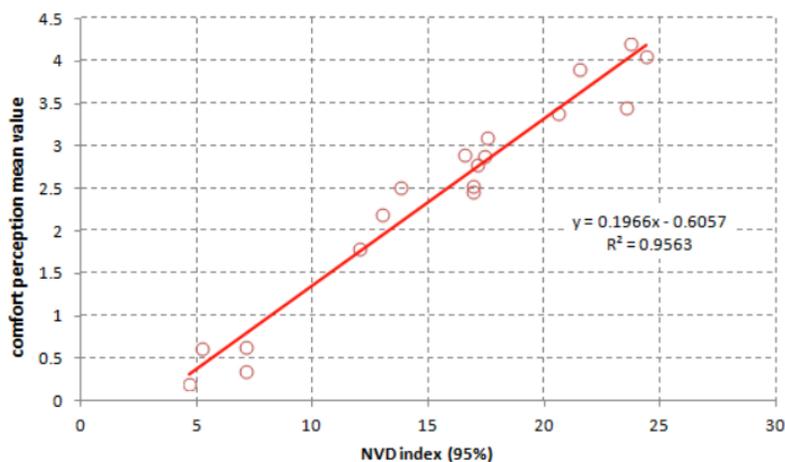


Figura 2-5. Relación lineal entre la percepción del panel de evaluación y el indicador propuesto. Fuente: Bodini *et al.* (2014).

Es claro como existe una gran correlación entre la percepción de las personas y el valor calculado por el indicador propuesto. A partir de esta regresión, los autores establecen una escala que relaciona el valor de N_{VD} con un nivel de comodidad (ver Tabla 2-3).

Tabla 2-3. Percepción de la comodidad según valor de N_{VD} . Fuente: Bodini *et al.* (2014).

N_{VD}	Percepción
< 5	Muy Bueno
5 – 10	Bueno
10 – 20	Regular
20 – 25	Malo
> 25	Muy Malo

En la Figura 2-6 se puede ver un ejemplo de análisis espacial donde cada punto corresponde a una medición de la comodidad asociada a la posición del GPS de la

ocurrencia para el recorrido 10. Se nota como la comodidad tiende a disminuir en las intersecciones y en las curvas es donde alcanza sus peores valores.

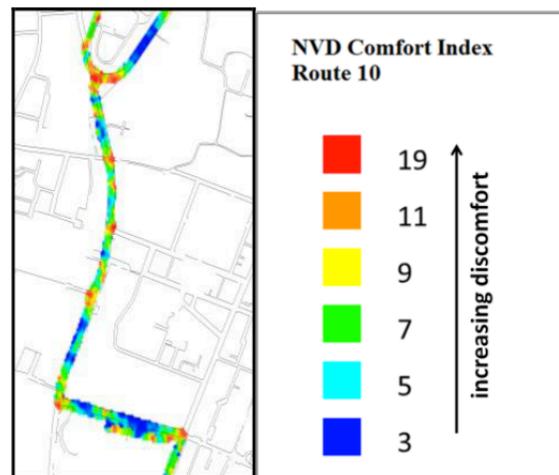


Figura 2-6. Mapa temático según los valores de N_{VD} para un recorrido de prueba.
Fuente: Bodini *et al.* (2014).

En la Figura 2-7 se puede apreciar un análisis estadístico que indica la frecuencia de ocurrencia de un valor de N_{VD} . Es importante destacar como gran parte de las ocurrencias se encuentran en el rango clasificado como “bueno”, seguido por el rango “regular”.

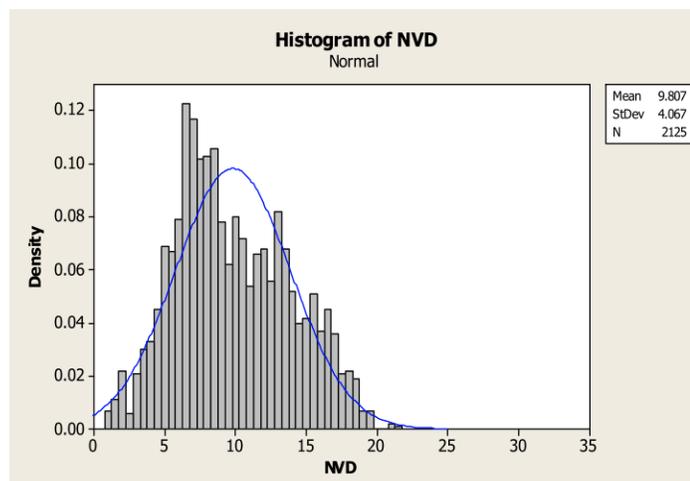


Figura 2-7. Histograma de los valores de N_{VD} para un recorrido de prueba. Fuente: Bodini *et al.* (2014).

Estas herramientas son muy útiles ya que permiten detectar recorridos con problemas y tomar medidas específicas para mejorarlos (por ejemplo: cambiar el trazado, arreglar el asfalto, etc.). Hay que destacar que los autores atribuyen los resultados del indicador a condiciones de la infraestructura o el bus y no a maniobras de conducción del chofer. Se puede apreciar cómo en la literatura no hay una manera específica de medir la comodidad y, específicamente, la conducción. A pesar de ello, se desprende que medir las vibraciones en los distintos ejes puede ser una metodología apropiada para cuantificar la conducción en el transporte público.

2.4 Índice de Calidad de Atención al Usuario de Transantiago

El año 2012 se incorpora a los contratos de licitación de Transantiago el Índice de Calidad de Atención al Usuario (ICA), que consideraba 14 atributos asociados al servicio entregado al usuario tanto como el que va a bordo del bus como el que espera en el paradero (DTPM, 2014). El objetivo de este índice es asegurar un mínimo nivel de servicio

entregado a los pasajeros. En la Tabla 2-4 se muestran los 14 atributos considerados por el ICA.

Tabla 2-4. Atributos para evaluar calidad de atención al usuario en ruta. Fuente: DTPM, 2014.

Nomenclatura	Descripción
A01	El conductor abre y cierra oportunamente las puertas al finalizar e iniciar el movimiento
A02	El conductor conduce sin frenazos ni movimientos bruscos
A03	El conductor es amable con los usuarios
A04	El conductor detiene el bus cuando debe, es decir, cuando algún usuario requiere subir o bajar
A05	El conductor detiene el bus donde debe, es decir, sólo en paradas autorizadas
A06	El conductor aproxima el bus correctamente al paradero, sin detenerse en segunda fila o lejos de la acera
A07	El conductor conduce sin fumar, ni conversa por celular o con un pasajero o un acompañante mientras conduce
A08	El conductor se detiene ante todas las luces rojas de los semáforos y señales Pare, señales Ceda el Paso y Pasos de Cebra
A09	Los letreros de recorrido están en buen estado, bien ubicados y exhiben información correcta respecto del sentido del servicio
A10	El panel superior variable está encendido, en buen estado y exhibe información correcta respecto del sentido del servicio
A11	La señalización interior está correctamente instalada
A12	La señalización interior está en buen estado
A13	El bus cuenta con leyenda "Informaciones y Reclamos" bien ubicada y legible
A14	El bus tiene funcionando el velocímetro

En el contrato se detalla la metodología empleada para el cálculo de este índice. Un pasajero incógnito a bordo del bus es quien evalúa cada atributo de manera binaria: cumple (1) o no cumple (0). El ICA para cada bus j se calcula según la ecuación (2.4), que corresponde al promedio simple del valor de cada atributo k (DTPM, 2014).

$$ICA_j = \frac{1}{14} \cdot \sum_{k=1}^{14} a_{j,k} \in \{0,1\} \quad (2.4)$$

La metodología establece ciertos criterios para la evaluación de cada atributo que buscan reducir la subjetividad que el ser humano puede incorporar. En el caso específico del atributo a02: “El conductor conduce sin frenazos ni movimientos bruscos”, se establecen 5 criterios simples:

- Es considerado frenazo o movimiento brusco cuando un pasajero pierde estabilidad a raíz de la maniobra del conductor.
- Debe disminuir gradualmente la velocidad para pasar por un lomo de toro.
- Debe conducir con la concentración necesaria que le permita detenerse sin frenazos bruscos ante una parada solicitada; luz roja; pare; ceda el paso; paso de cebra y cualquiera tipificada en la ley de tránsito.
- Condiciones de la calzada, virajes estrechos, baches, líneas de tren u otras ajenas al conductor no serían consideradas incumplimiento.
- Si el frenazo o movimiento brusco se realiza con el fin de evitar un accidente, este no será considerado como incumplimiento.

En el contrato se estipula 85% como el mínimo nivel de cumplimiento con el cual los operadores no se ven expuestos a descuentos monetarios por este indicador. Dichos descuentos siguen la estructura mostrada en la Tabla 2-5. Además, en caso de que un atributo se encuentre bajo el 60%, se aplica una multa adicional de 100 UF¹.

¹ Valor de la Unidad de Fomento (UF) al 30 de Abril de 2018 = \$ 27.004,63 (SII, 2018).

Tabla 2-5. Estructura de Descuentos según ICA obtenido. Fuente: DTPM, 2014.

Valor de ICA	Descuento Aplicado
$ICA \geq 85\%$	-
$80\% \leq ICA < 85\%$	Hasta 200 UF
$ICA < 80\%$	Hasta $200 UF \cdot (1 + (0,8 - ICA))^3$

Los contratos estipulan que cada mes se medirá el indicador aleatoriamente en el 25% de los buses de una empresa (DTPM, 2014). Esto impide que se realice seguimiento a los conductores para detectar focos de buenos o malos resultados con el fin de entregar incentivos o capacitaciones, según corresponda. Además, los resultados obtenidos pueden no ser representativos de la realidad de toda la empresa. Esto se debe, principalmente, a la metodología de pasajero incógnito empleada.

Durante el año 2015, el promedio de cumplimiento fue de 77%, con lo cual las unidades de negocio se vieron expuestas a multas promedio de hasta 219 UF mensuales. El atributo asociado a la forma de conducción del conductor (A02 en la Tabla 2-4), tiene un promedio de cumplimiento de 55% en el mismo periodo. En cambio, el atributo asociado a la amabilidad del conductor (A03 en la Tabla 2-4), llega a 99% de cumplimiento (DTPM, 2016b). Pero, al ser el ICA un promedio simple entre todos los atributos (ver ecuación 2.4), ambos valores ponderan del mismo modo por lo que el resultado tiende a ocultar los malos rendimientos en ciertos atributos. En la Tabla 2-6 se puede apreciar el nivel de cumplimiento según atributo durante el año 2015. Aquellos atributos donde el cumplimiento está por debajo del 85% requerido se encuentran destacados en amarillo. Aquellos que están bajo el 60%, en rojo.

Tabla 2-6. Nivel de cumplimiento según atributo evaluado en el ICA para el año 2015.
Fuente: DTPM, 2016b.

Atributo	Nivel de Cumplimiento	de Atributo	Nivel de Cumplimiento
A01	93%	A08	89%
A02	55%	A09	51%
A03	99%	A10	86%
A04	86%	A11	37%
A05	79%	A12	82%
A06	62%	A13	82%
A07	79%	A14	94%

En específico, los descuentos se aplican en base al valor del ICA, entonces la multa aplicada no genera un real incentivo a capacitar a los choferes para lograr una mejor conducción. A modo de ejemplo, si se considerara solo el atributo A02 en la fórmula de descuentos, este llegaría hasta las 491 UF (un 124% más que la multa real aplicada).

Con lo expuesto de la literatura en este capítulo, queda demostrado que el uso de dispositivos inerciales entrega atributos (como la aceleración en los tres ejes) que pueden ser relacionados con la comodidad a bordo de un vehículo. Por otro lado, con el uso de una metodología de clasificación de la conducción basada en la metodología empleada actualmente en el cálculo del ICA, es posible relacionar ese conjunto de atributos con la forma de conducción del chofer.

En última instancia, esto permitiría la formulación por parte de la autoridad de un indicador para la conducción que sea más simple y barato de obtener, además de escalable a todos los buses del sistema.

3 METODOLOGÍA

Como se estableció en el capítulo 1, esta tesis tiene como objetivo formular una nueva metodología para el cálculo del atributo de conducción del ICA de Transantiago. Una maniobra de conducción se entiende, para estos efectos, como una forma de conducir responsable y cómoda para los pasajeros sin aceleraciones y desaceleraciones bruscas, sin exceder la máxima velocidad permitida en ruta, en curvas o al cambiar de pista, etcétera.

De la revisión bibliográfica se pudo establecer que existe una relación entre variables inerciales y la percepción de la conducción. Por lo tanto, utilizando máquinas de aprendizaje se espera calibrar un modelo que sea capaz de clasificar la conducción a partir dichas variables.

Para lo anterior es necesario contar con información real que se obtuvo como pasajero incógnito en algunos recorridos del Transantiago. De esta forma se clasificó la forma de conducción en cada instante y, paralelamente, se registró con un IMU el valor de distintas variables inerciales. Con estos datos se logra construir una base de datos donde cada vector de atributos corresponde a un instante del tiempo y los atributos al valor de las variables inerciales.

En este capítulo se describirá la metodología propuesta en esta investigación, la cual se divide en dos partes: diseño del experimento que permita la recolección de dichos datos y su posterior modelación. Además, se explicarán los métodos de clasificación considerados para dicha modelación.

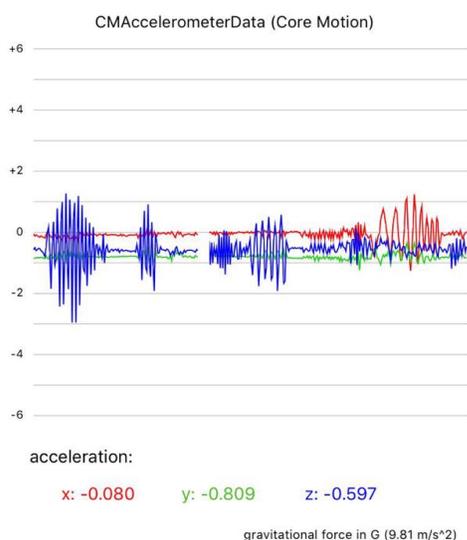
3.1 Diseño del Experimento

En esta sección se describe el experimento realizado en el marco de la investigación. Se comienza con la descripción de las herramientas utilizadas. Luego, se detalla el proceso de recolección de datos. Finalmente, se explica el procesamiento realizado a los datos para generar una base de datos apta para la etapa de modelación.

3.1.1 Herramientas Utilizadas

Una de las ventajas que tienen los dispositivos inerciales (acelerómetros, giroscopios, GPS) es que se encuentran al alcance todas las personas, ya que están presentes en cada *smartphone*. Estos equipos usan IMUs para controlar la posición de la pantalla (vertical, horizontal), como una extensión de realidad aumentada en juegos (manubrio en un juego de carreras), etc. Por lo general, la información de los IMUs está disponible para desarrolladores de aplicaciones de *smartphone*.

Para este trabajo se utilizó un iPhone 6S de 32 GB, modelo A1633, el cual posee un acelerómetro triaxial Bosch Sensortec 367LA, un giroscopio de 6 ejes InvenSense MP67B y un GPS Qualcomm WTR1605L. Para recuperar la información de los dispositivos se utilizó la aplicación “SensorLog 1.9.2” (Thomas, 2017) disponible en AppStore. Dentro de las ventajas que tiene esta aplicación es el uso de una interfaz amigable que muestra en tiempo real gráficos con los datos de cada chip incorporado en el iPhone. Por ejemplo, en la Figura 3-1 se muestra la aceleración en los tres ejes.



**Figura 3-1. Ejemplo de la interfaz de la aplicación SensorLog utilizada en el experimento.
Fuente: Thomas, 2017.**

La aplicación almacena la información en un archivo separado por comas (.csv) con una frecuencia escogida por el usuario que va desde 1Hz a 100 Hz. Para efectos de la recolección de datos realizada, se escogió una frecuencia de 50 Hz. Es importante señalar que el número real de datos recolectados por segundo varía levemente, es decir, a veces en un segundo se pueden registrar 49 o 51 vectores de atributos.

Por otro lado, para registrar las instancias de conducción inapropiada se utilizó un iPad mini de 32 GB, modelo A1454, en conjunto de la aplicación “Counter – Advanced Tally Counter 7.1” (Mikrasya, 2016) también disponible en AppStore. Esta aplicación, como su nombre lo sugiere, es un contador que almacena la información en un archivo .csv con el *timestamp* de cada instancia.

3.1.2 Recolección de datos

Para la recolección de datos se escogió cinco tramos de duración y distancia recorrida variable. En cada tramo se midió un recorrido de Transantiago diferente. En total, el experimento recolectó 232.058 vectores de atributos de recorridos operados por tres de las siete empresas licitadas. Una descripción detallada de cada recorrido se puede encontrar en la Tabla 3-1.

Tabla 3-1. Detalle de cada tramo utilizado para la recolección de datos. Fuente: Elaboración propia.

Tramo	Recorrido	Distancia (km.)	Tiempo (min.)	Tipo de bus	Nº Datos	Operador	Ranking ICA
1	213E	8,0	21	No articulado	62.932	STP Santiago S.A.	1º
2	210	3,0	9	Articulado	26.525	Subus Chile S.A.	7º
3	210V	4,9	14	Articulado	40.660	Subus Chile S.A.	7º
4	102	4,2	10	No articulado	30.311	Alsacia S.A.	2º
5	712	10,4	24	No articulado	71.630	STP Santiago S.A.	1º

En la Figura 3-2 se puede apreciar un mapa con cada tramo descrito en la Tabla 4-1. Además, en la Tabla 3-2 se muestra la intersección o punto referencial de inicio y término de cada recorrido.

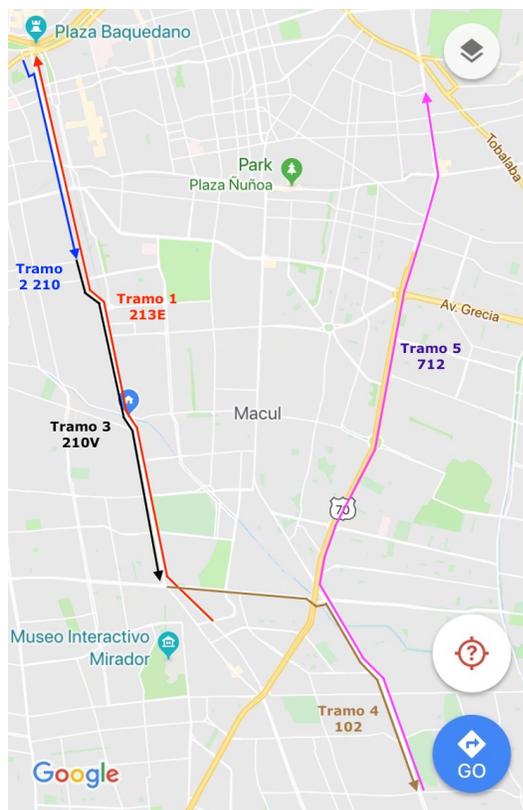


Figura 3-2. Mapa de tramos evaluados durante el experimento.
Fuente: Elaboración propia en base a Google Maps.

Tabla 3-2. Intersecciones o puntos de referencia del inicio y término de cada recorrido.
Fuente Elaboración Propia.

Tramo	Recorrido	Desde – Hasta
1	213E	Metro Mirador – Plaza Baquedano
2	210	Plaza Baquedano – Av. Vicuña Mackenna con Maule
3	210V	Av. Vicuña Mackenna con Maule – Metro Pedrero
4	102	Metro Pedrero – Av. La Florida con Rojas Magallanes
5	712	Av. La Florida con Rojas Magallanes – Av. Américo Vespucio con Emilia Téletz

El día que se realizó la medición fue el miércoles 03 de Mayo de 2017 entre las 13:00 y 17:00. El dispositivo se fijó temporalmente al piso del bus alineado con la dirección del mismo para no distorsionar los valores de la aceleración por cambios en los ejes. Es importante señalar que según la posición del *smartphone*, cambia la nomenclatura de cada eje de la aceleración. En la Figura 3-3 se puede apreciar un diagrama con los ejes del dispositivo en la posición usada para medir. Por ejemplo, comúnmente la aceleración en el eje y es la aceleración lateral, pero debido a la posición del equipo, la aceleración en el eje x corresponde a la lateral.

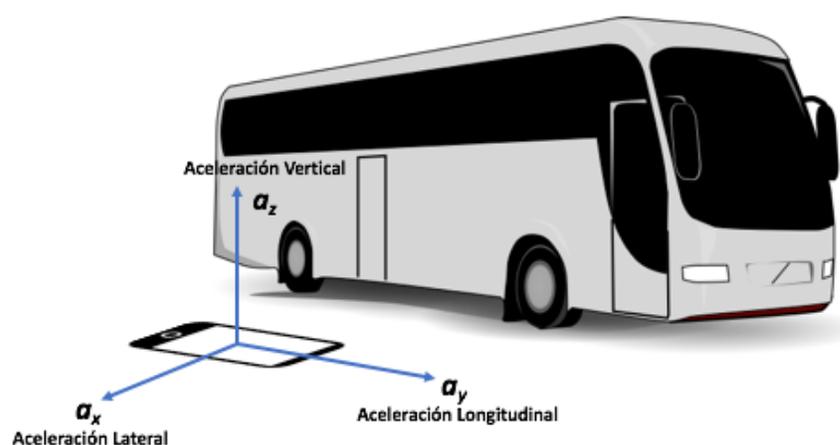


Figura 3-3. Ejes de la aceleración capturada por el dispositivo, según la dirección del bus.
Fuente: Elaboración propia.

En la Figura 3-4 se muestra la nomenclatura usada por los giroscopios para describir la rotación respecto a un eje determinado. Por ejemplo, el movimiento respecto al eje vertical (eje z) se conoce como *yaw rotation* y su nomenclatura es g_z .

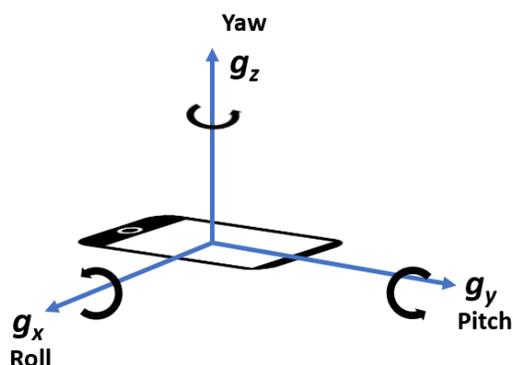


Figura 3-4. Movimientos de rotación del giroscopio capturados por el dispositivo.
Fuente: Elaboración propia.

Para determinar si la conducción fue apropiada o no, se sigue la metodología estipulada en el cálculo del cumplimiento del ICA (DTPM, 2014) y que fue descrita en la sección 2.5 de esta tesis.

Una vez instalado el dispositivo se comienza la grabación de los datos y cada vez que se genera una maniobra brusca se registra el incidente en el contador y se anota en una bitácora la razón. Las instancias de conducción brusca se registran en el contador por segundo. Del total de 4.569 segundos, solo 106 fueron etiquetados como conducción brusca.

3.1.3 Procesamiento de datos

Para la modelación se considerarán las aceleraciones en los tres ejes (a_x , a_y y a_z) capturadas por los acelerómetros del dispositivo, así como las rotaciones respecto a los mismos tres ejes (g_x , g_y y g_z) capturadas por el giroscopio. Debido a que cada instancia

de conducción inapropiada está asociada a un segundo, esto puede llevar fácilmente a errores si es que el pasajero incógnito registró el incidente uno o más segundos después. Para evitar el problema descrito y reducir el número total de observaciones de la base de datos, se decide obtener un promedio de los datos del dispositivo inercial cada 3 segundos. En otras palabras, durante dicho intervalo se obtienen aproximadamente 150 vectores de atributos que son promediados (en cada atributo) para obtener solo un vector de atributos representativo de lo que ocurre en esos 3 segundos.

Es importante destacar que debido a que la aceleración puede ser positiva o negativa (ver Figura 3-5), un promedio simple cancelaría el efecto que se busca captar.

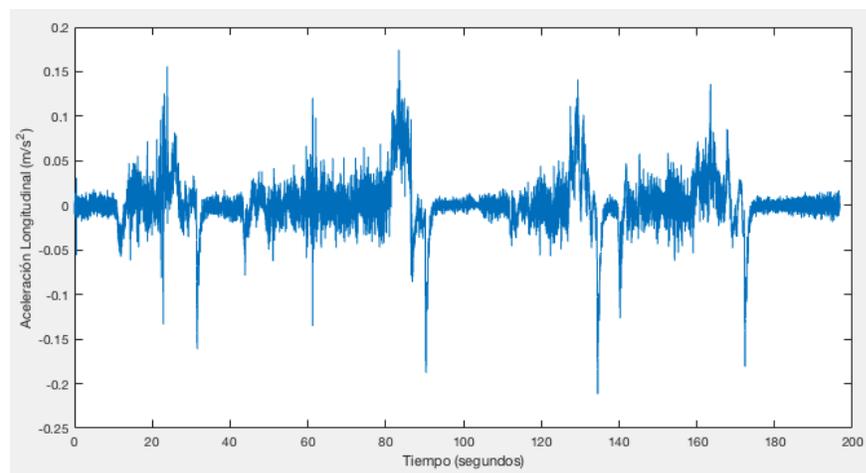


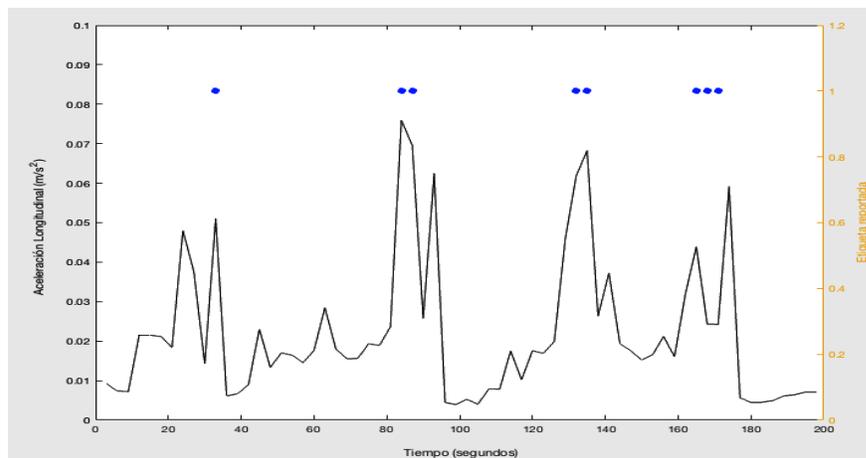
Figura 3-5. Aceleración longitudinal del bus 213E. Fuente: Elaboración propia.

Por consiguiente, se utilizó el valor cuadrático medio (RMS, por su sigla en inglés) que se define según la ecuación 3.1. Se tiene que x_i corresponde al valor del atributo x (aceleraciones y movimientos del giroscopio) en el instante i y N es el número de vectores de atributos que se registraron en el intervalo de 3 segundos.

$$x_{RMS} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N x_i^2} \quad (3.1)$$

El uso de RMS también permite eliminar el ruido blanco presente en los datos entregados por el dispositivo inercial. Este ruido es inherente a la naturaleza de los atributos y no se debe al aparato con que se mide.

Con esto, los datos se ven como en la Figura 3-6. Como es de esperar, todos los valores son positivos y se logra capturar los instantes donde la magnitud de la variable aumenta considerablemente.



**Figura 3-6. RMS de la Aceleración longitudinal con las alertas reportadas (puntos sobre la curva).
Fuente: Elaboración propia.**

Junto a la aceleración longitudinal, la Figura 3-6 incluye los instantes donde se registraron instancias de conducción inapropiada (alertas) registradas por el pasajero incógnito. Es posible notar una gran relación entre los *peaks* de aceleración longitudinal y la percepción de maniobra de conducción brusca.

Como resultado de aplicar RMS a los datos, la cantidad de instancias se reduce a 1.523, de las cuales 106 corresponden a conducción brusca.

Debido a que los datos deben separarse aleatoriamente entre un set de entrenamiento y un set de validación, se pierde la secuencialidad de la información. En otras palabras, un vector de atributos del set (entrenamiento o validación) ya no tiene relación con el vector anterior ni el siguiente. Esta relación es importante de mantener, ya que cuando se realiza una maniobra brusca, esta es progresiva en el tiempo. Si bien la alerta es en un intervalo t^* , el RMS de la aceleración aumentó en el intervalo $t^* - 1$ y luego se reduce en el intervalo $t^* + 1$.

Por lo tanto, se crea una nueva base de datos que incorpora esta progresión y que se puede apreciar gráficamente en la Figura 3-7. Donde cada columna corresponde a los atributos obtenidos por el dispositivo inercial y cada fila asociada a un intervalo de tiempo t^* está compuesta por el vector de atributos del intervalo anterior $f(t^* - 1)$, el actual $f(t^*)$ y el siguiente $f(t^* + 1)$. Además, la fila incluye la etiqueta o clase que describe al vector.

TIEMPO	ATRIBUTOS			CLASE
t^*	$f(t^* - 1)$	$f(t^*)$	$f(t^* + 1)$	$c(t^*)$

Figura 3-7. Estructura de la nueva de base de datos creada para incorporar la progresión de los atributos en el tiempo. Fuente: Elaboración propia.

Con esta modificación, la base de datos queda con 1.513 vectores de atributos asociados a cada intervalo de 3 segundos (en adelante instantes). Es decir, se tiene una muestra desbalanceada donde solo el 7% de las etiquetas corresponde a la clase positiva.

Los modelos de clasificación tienen un mejor rendimiento cuando los valores de las variables se encuentran en el rango $[-1,1]$ (Hsu *et al.*, 2016). Es por esta razón que cada atributo (las columnas de la matriz) se normaliza dividiendo por el módulo de cada columna. Finalizada esta acción, la base de datos está lista para ser dividida en sets de entrenamiento y validación. La proporción de repartición de datos es 80% para el set de entrenamiento y 20% para validación.

Debido a que la muestra es desbalanceada, es necesario que la proporción de datos positivos y negativos se mantenga dentro de los sets creados aleatoriamente. Para efectos de esta base de datos, dicha proporción es 1:7,5.

Finalmente, debido a que la creación de los set de datos es aleatoria, los resultados pueden ser poco representativos en caso de que la muestra de datos sea tal que solo se logre un buen rendimiento para específicas combinaciones de datos de entrenamiento y validación. Para evitar este problema, se generan 5 combinaciones diferentes de sets de entrenamiento/validación y se procede a la modelación descrita en la siguiente sección para cada combinación.

3.2 Métodos de Clasificación

Un método de clasificación es aquel que busca determinar la etiqueta (en muchos casos binaria) que describe a un conjunto de atributos. Una forma de abordar este problema es mediante el aprendizaje supervisado, que forma parte del aprendizaje de máquinas (Mohri, Rostamizadeh y Talwakar, 2012). Así, se puede calibrar un modelo con datos conocidos para luego predecir la etiqueta de clasificación de datos a los cuales no se les conoce su etiqueta.

A los datos conocidos se les conoce como datos de entrenamiento y que, generalmente, están compuestos por vectores de atributos. Para cada vector de atributos se le asocia una etiqueta binaria que describe a esos atributos.

Por otro lado, los algoritmos del aprendizaje de máquinas deben ser calibrados y entrenados con los datos de entrenamiento. La calibración consiste en encontrar los hiperparámetros de cada método de aprendizaje. El entrenamiento consiste en ajustar el algoritmo a un set de datos específico.

En esta sección se explica la teoría detrás de los tres métodos de clasificación utilizados en esta tesis y que están enmarcados en el aprendizaje de máquinas: máquinas de vectores de soporte, árboles de decisión y k vecinos más cercanos. Luego se expone sobre el problema de usar muestras no balanceadas y cómo se adaptan los métodos para superarlo.

A continuación, se describe la metodología para la elección de los parámetros que considera cada método. Finalmente, se describen las medidas de rendimiento con las cuales se evaluarán los modelos calibrados según cada método.

3.2.1 Máquinas de Vectores de Soporte (SVM)

Una Máquina de Vectores de Soporte (SVM, por sus siglas en inglés) se basa en la teoría de aprendizaje estadístico y en el principio de minimización de riesgo estructural. Es útil para regresiones (análisis de series de tiempo) y reconocimiento de patrones (problema de clasificación), entre otros problemas.

SVM permite manejar de manera explícita problemas de clasificación con muestras desbalanceadas. Además, tiene la ventaja de poder enfrentar apropiadamente problemas de sobre-entrenamiento que reducen la generalización del modelo, es decir, aprender solo los datos de entrenamiento (Wang *et al.*, 2015).

La idea tras el funcionamiento de SVM consiste en que los vectores de atributos de los datos de entrenamiento son llevados a un espacio de dimensión superior por una función kernel. SVM encuentra un hiperplano lineal que sea capaz de separar estos vectores en el espacio de dimensión superior con el máximo margen posible, tomando en

cuenta la etiqueta de clasificación buscada. Para evitar el sobre-entrenamiento, se utiliza un parámetro positivo C que es la penalización aplicada al clasificar un vector de atributos erróneamente. La formulación teórica de SVM se puede encontrar en la literatura (Boser *et al.*, 1993; Cortes y Vapnik, 1995).

La función kernel, como se mencionó anteriormente, es aquella que permite llevar los datos a un espacio de dimensión superior. Esto se hace ya que en muchos casos la separación de la información no es factible en el espacio original y uno de mayor dimensión facilita la clasificación.

A continuación se aprecia la formulación genérica de la función kernel:

$$K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j) \quad (3.2)$$

donde la función ϕ varía según el tipo de kernel utilizado. Las más utilizadas en la literatura son:

- Lineal:
$$K(x_i, x_j) \equiv x_i^T x_j \quad (3.3)$$

- Radial Basis Function (RBF):
$$K(x_i, x_j) \equiv \exp(-\gamma \|x_i - x_j\|^2) \quad (3.4)$$

con $\gamma > 0$, parámetro de la función kernel RBF.

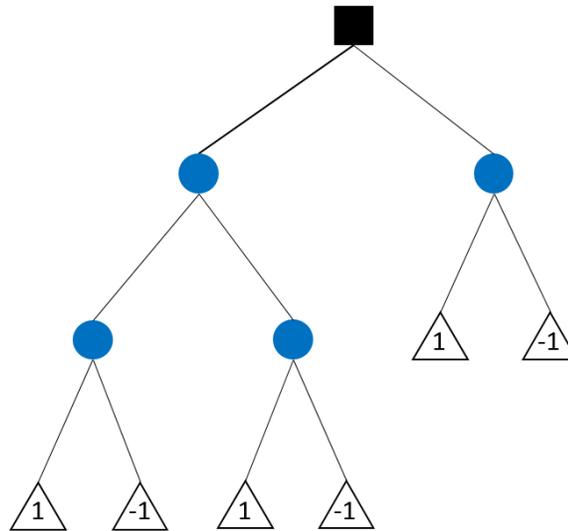
El uso de la función kernel RBF es recomendado debido a que entrega mejores rendimientos cuando el espacio no se puede separar fácilmente usando una función lineal. Por lo general, la mayoría de los problemas de clasificación suelen ser más complejos por lo que una primera aproximación siempre es la función kernel RBF. A pesar de ello, en aquellos casos donde usar una función kernel lineal otorga un mejor rendimiento, se aconseja utilizar esta debido a que se reduce el tiempo de computación considerablemente.

3.2.2 Árboles de Decisión

Los métodos basados en árboles son bastante populares en la minería de datos (Rokach y Maimon, 2008; Hastie, Tibshirani y Friedman, 2001), siendo utilizados tanto en clasificación como en regresión. Estos modelos son apropiados cuando hay un número elevado de vectores de atributos y existe incertidumbre sobre la manera en que las variables explicativas (atributos) deberían introducirse en el modelo (Quinlan, 1986).

Un árbol de decisión consiste en una división jerárquica y secuencial de los datos con el objetivo de crear un modelo capaz de predecir el grupo al que pertenecen los datos de entrada. En el nodo inicial se escoge aquel atributo de los datos de entrenamiento que permite dividirlos en dos conjuntos lo más homogéneos posibles. Por ejemplo, dado una atributo x_1 se determina un punto de corte (U) tal que se obtienen dos subconjuntos: aquellos con $x_1 \leq U$ y los que tienen $x_1 > U$. Cada subconjunto corresponde a otro nodo donde se vuelve a repetir el proceso de seleccionar un atributo y un punto de corte para dividir la muestra en dos partes más homogéneas. El proceso termina cuando se hayan clasificado todas las observaciones correctamente en su etiqueta (Quinlan, 1986; Breiman *et al.*, 1984).

Este método permite obtener finalmente una representación gráfica de todas las decisiones aplicada a los datos, así como comprender qué atributos son considerados y de qué manera.



**Figura 3-8. Ejemplo de un árbol de decisión que clasifica en dos grupos.
Fuente: Elaboración Propia.**

En la Figura 3-8 se aprecia un ejemplo genérico de un árbol de decisión. El nodo cuadrado corresponde al nodo inicial o raíz. Los nodos circulares corresponden a los nodos de decisión. Finalmente, los nodos triangulares u hojas representan la etiqueta final entregada (1 o -1) al clasificar los atributos de los datos de entrenamiento. Esta representación visual es una diferencia clara con los modelos de “caja negra” como SVM.

Para decidir cuál variable se utiliza en cada nodo de decisión se calcula la proporción de observaciones que pasan por el nodo $t = 1, \dots, T$ hacia los subconjuntos $g = 1, \dots, G$. Se define $p(g|t)$ como la probabilidad que una observación que llegue al nodo t pertenezca al grupo g . Luego, se define la impureza del nodo t (ver ecuación 3.5) como una medida de la entropía.

$$I(t) = - \sum_{g=1}^G p(g|t) \cdot \log p(g|t) \quad (3.5)$$

La entropía es máxima cuando $p(g|t) = 1/G$. Finalmente, la variable que se escoge para el nodo t es aquella que minimiza la impureza que resulta al dividir el nodo (Quinlan, 1986; Breiman *et al.*, 1984).

Cuando el número de atributos es muy grande, el árbol puede contener un número excesivo de nodos. En estos casos se debe aplicar procedimientos de simplificación como establecer:

- Número máximo de divisiones: este número establece un límite para la cantidad de nodos de decisión dentro del árbol.
- Tamaño mínimo de un nodo de decisión: este número establece un límite inferior para la cantidad de datos que ameritan una división del nodo. Es decir, si no se llega a la cota entonces el nodo no se divide y se convierte en una hoja.

3.2.3 K vecinos más cercanos (KNN)

El método k vecinos más cercanos (KNN por sus siglas en inglés) es uno de los más simples en el campo de máquinas de aprendizaje. Corresponde a un método no paramétrico utilizado tanto para clasificación como regresión. KNN tiene la particularidad que la función de probabilidad de ser clasificado con una etiqueta se aproxima localmente y el algoritmo se aplica sobre los datos de entrenamiento cada vez que se solicita la clasificación de un nuevo vector de atributos. Esto se conoce como aprendizaje vago.

En comparación con los métodos anteriormente expuestos, SVM y árboles de decisión entrenan *a priori* un modelo en base a los datos de entrenamiento que luego es utilizado en la clasificación de nuevos datos.

La fase de entrenamiento de KNN consiste simplemente en almacenar los vectores de atributos hasta que se solicite una clasificación. Cuando esto ocurre, el algoritmo busca los K vectores más cercanos, a este nuevo vector sin etiquetar, dentro de los datos de entrenamiento. El criterio más utilizado es la distancia euclidiana que se muestra en la ecuación 3.6:

$$d(\vec{x}, \vec{x}_j) = \sqrt{\sum_{r=1}^p (x_r - x_{rj})^2} \quad (3.6)$$

donde p corresponde a la cantidad de atributos del vector, $\vec{x} (x_1, x_2, \dots, x_p)$ es el vector que se desea clasificar y $\vec{x}_j (x_{1j}, x_{2j}, \dots, x_{pj})$ es el vecino j de dicho vector. Finalmente, se etiqueta este nuevo vector considerando la clase mayoritaria dentro de los K vecinos más cercanos.

Es importante destacar que la constante K es definida *a priori* por el modelador. En la Figura 3-9 se puede apreciar un ejemplo donde existen dos clases en los datos de entrenamiento: cuadrados y triángulos. Si se quiere clasificar el círculo rojo considerando un $K = 3$, entonces se llega a que corresponde a un triángulo debido a que solo un vecino es cuadrado. En cambio, al considerar un $K = 6$, la mayoría de los vecinos pasa a ser de la clase cuadrado por lo que la etiqueta cambia.

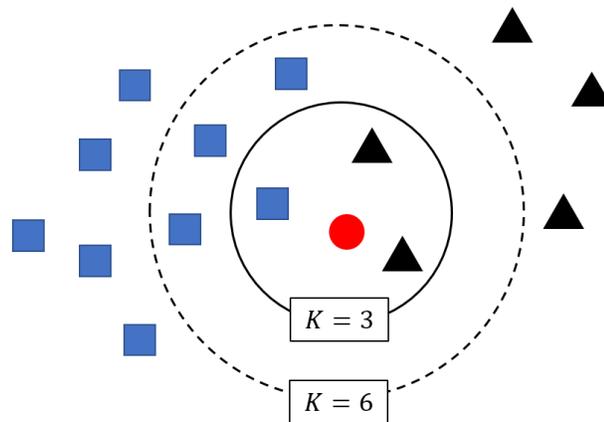


Figura 3-9. Ejemplo de un algoritmo KNN con dos valores de constante.
Fuente: Elaboración Propia.

3.2.4 Caso de la muestra no balanceada

Una muestra es no balanceada cuando la cantidad de datos positivos (con etiqueta 1), es decir, aquellos que se desea identificar es mucho menor a la cantidad de datos negativos (con etiqueta -1). Esto se vuelve un problema en casos extremos como, por ejemplo, cuando los datos positivos son menores al 5% de la muestra.

En estas situaciones, las máquinas de aprendizaje optan por no aprender la clase positiva, ya que equivocarse está dentro del margen de error aceptable para cualquier método de clasificación. Para superar esta dificultad, las máquinas de aprendizaje pueden incorporar un costo a clasificar erróneamente cada clase. Es importante señalar que este costo es diferente para cada clase, por lo que al darle una mayor prioridad a los datos positivos se puede romper el desbalanceo presente en la muestra.

En el caso de SVM, se incorpora la constante C al problema de optimización, que corresponde a la penalización por clasificar erróneamente un vector de atributos. Al abordar el problema de muestras no balanceadas, esta constante se divide en el número de clases de la muestra. En otras palabras, se tiene un parámetro C_i para cada clase. Al asignar parámetros tal que $C_1 > C_{-1}$, se fuerza al SVM a aprender la clase positiva.

Tanto para árboles de decisión como el método de KNN, se puede incorporar una matriz de costos como en la Tabla 3-3.

Tabla 3-3. Matriz de Costos usada en árboles de decisión y KNN. Fuente: Elaboración Propia.

C_{ij}	Predicción Positiva	Predicción Negativa
Condición Positiva	= 0	$\gg 0$
Condición Negativa	> 0	= 0

Se tiene que si la clasificación entregada por el método es correcta de acuerdo a la información real, $C_{ij} = 0$. En caso contrario, $C_{ij} > 0$. Del mismo modo que en SVM, el costo de equivocarse en la clase positiva debe ser mayor al de equivocarse con la clase negativa.

3.2.5 Elección de Hiperparámetros

Para la calibración de cada uno de los modelos expuestos en este capítulo se requiere la elección de ciertos hiperparámetros, cuya cantidad depende de cada método.

En el caso de SVM, el número exacto de hiperparámetros depende de la función kernel escogida. Para el caso de kernel lineal, se requiere calibrar dos parámetros: C_1 y C_{-1} . Para el kernel RBF, C_1 , C_{-1} y γ . Cuando la cantidad de parámetros es superior a uno, se deberán probar múltiples combinaciones de parámetros hasta encontrar la mejor. Esto se determina luego de realizar una validación cruzada dentro del set de entrenamiento que entrega la exactitud del modelo. El tipo de validación cruzada es *n-fold* y consiste en separar la muestra en n grupos. Secuencialmente, se deja un grupo fuera como set de validación y se entrena con los $n - 1$ restantes. En la Figura 3-10 se puede apreciar el procedimiento de manera gráfica para una validación cruzada *3-fold* o de 3 iteraciones.

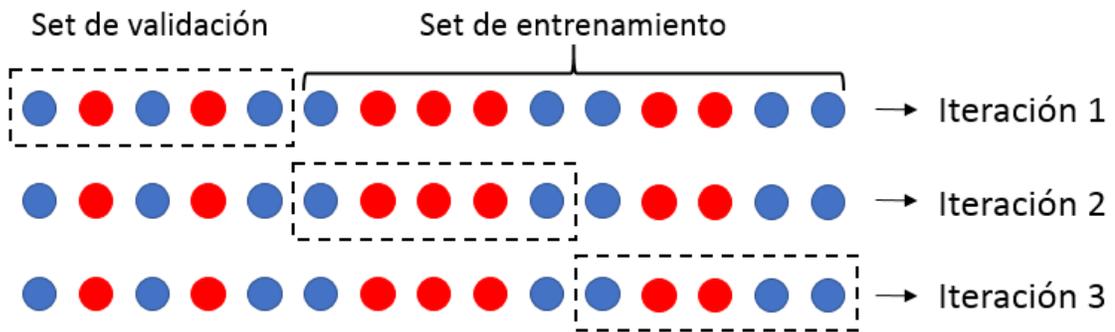


Figura 3-10. Descripción gráfica de una validación cruzada de 3 iteraciones.
Fuente: Elaboración propia.

Para facilitar la búsqueda de la mejor combinación de parámetros, se utiliza el algoritmo de *grid-search*. Es ampliamente aceptado usar incrementos exponenciales hasta encontrar la vecindad con mejores resultados (Hsu *et al.*, 2016). Claramente este procedimiento se vuelve más complejo mientras más parámetros que escoger debido a la gran cantidad de combinaciones que se debe probar.

En el caso de árboles de decisión, se deben escoger en total cuatro parámetros: $C_{1,-1}$, $C_{-1,1}$, el número máximo de divisiones y el tamaño mínimo de un nodo de decisión. Es importante recordar que $C_{1,-1}$ y $C_{-1,1}$ corresponden a los costos de equivocarse en la clasificación. El procedimiento para la elección es similar al de SVM, con validación cruzada y considerando el rendimiento del modelo se obtienen los parámetros finales para la calibración.

Finalmente, para KNN que es uno de los modelos más simples de máquinas de aprendizaje solo se requiere escoger tres parámetros: k , $C_{1,-1}$ y $C_{-1,1}$. El procedimiento es igual que con el método de árboles de decisión.

3.3 Modelación

A continuación se explicarán las medidas de rendimiento que se consideraron para evaluar el rendimiento de los modelos. Luego, se detalla el proceso de calibración de los modelos realizado para cada método de clasificación.

3.3.1 Medidas de Rendimiento

Se consideraron tres indicadores que se evalúan sobre el set de validación. Estos se basan en la matriz de confusión de la Tabla 3-4.

Tabla 3-4. Matriz de Confusión. Fuente: Elaboración Propia

Población		Predicción	
		Positiva	Negativa
Realidad	Positiva	Verdadero Positivo (VP)	Falso Negativo (FN)
	Negativa	Falso Positivo (FP)	Verdadero Negativo (VN)

El primero, la exactitud del modelo, se define según la norma ISO 5725-1 como la cercanía de la predicción a la condición real. En términos de la matriz de confusión se expresa según la ecuación 3.7.

$$Exactitud = \frac{VP + VN}{población\ total} \quad (3.7)$$

Como regla general, se aconseja utilizar la exactitud para evaluar el rendimiento de un método de clasificación, pero dado que se tiene una muestra desbalanceada se deben incluir más indicadores.

El segundo, el error tipo I del modelo, se define como el error que se comete al rechazar la hipótesis nula cuando ésta era verdadera en la población. Para efectos de esta investigación, la hipótesis nula es: “el conductor conduce sin frenazos ni movimientos bruscos”. Es decir, un error tipo I equivale a acusar a un chofer de mala conducción cuando no lo hizo. En términos de la matriz de confusión se expresa según la ecuación 3.8.

$$Error\ Tipo\ I = \frac{FN}{VP + FN} \quad (3.8)$$

El tercer indicador es el error tipo II del modelo y se define como el error cometido al no rechazar la hipótesis nula cuando ésta era falsa en la población. Es decir, para esta investigación, corresponde a catalogar de buena conducción a las maniobras bruscas realizadas por un chofer. En términos de la matriz de confusión se expresa según la ecuación 3.9.

$$Error\ Tipo\ II = \frac{FP}{VN + FP} \quad (3.9)$$

3.3.2 Calibración de los modelos

- SVM

Como se mencionó anteriormente, en la muestra existe una proporción entre datos positivos y negativos de 1:7,5. Como primera aproximación, se consideró esta proporción como los valores de los parámetros de penalización del error de SVM. Es decir:

$$C_1 = 7,5 \quad ; \quad C_{-1} = 1 \quad (3.10)$$

Utilizando la librería LIBSVM (Chang y Lin, 2011) en el *software* MATLAB se entrenó el modelo con un kernel RBF y lineal con el objetivo de determinar aquella función que entrega una mejor exactitud. El kernel RBF tiene una constante extra que debe ser determinada (γ). Luego de aplicar un algoritmo *grid-search* se obtiene que el valor que maximiza el rendimiento es $\gamma = 0,01$.

Finalmente, se realiza una validación cruzada de 10 iteraciones y la exactitud obtenida con cada kernel se encuentra en la Tabla 3-5.

Tabla 3-5. Exactitud del modelo según tipo de Kernel usado. Fuente: Elaboración propia

Tipo de Kernel usado	Exactitud del modelo (Validación cruzada 10-fold)
RBF	86,5%
Lineal	89,3%

Como el kernel lineal entrega una mayor exactitud y es simple de implementar, se escoge para entrenar el modelo definitivo. A continuación se deben encontrar los parámetros C_1 y C_{-1} que entreguen el mejor rendimiento dadas las medidas definidas anteriormente.

A medida que reducimos el error tipo II, el error tipo I y la exactitud del modelo empeoran. Esto lleva a concluir que se debe generar un compromiso entre estos indicadores hasta llegar a los niveles deseados. En la Figura 3-11 se puede notar un ejemplo de esta situación.

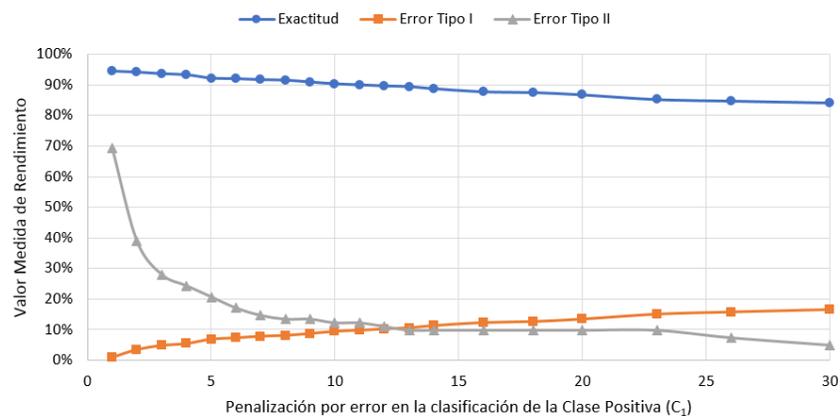


Figura 3-11. Variación de las medidas de desempeño según el valor del parámetro C_1 (SVM). Fuente: Elaboración propia.

De la Figura 3-11 se puede notar que para $C_1 = C_{-1} = 1$ el modelo tiene la mejor exactitud (94,5%), pero el error tipo II es cercano a un 70%. Al aumentar el valor de este parámetro, el error tipo II cae rápidamente a valores menores que 30%, mientras que la exactitud se mantiene cercana al 94% y el error tipo I, al 5%.

La decisión de cuando dejar de aumentar el parámetro C_1 depende completamente de los objetivos del modelador al implementar la metodología. Si se desea estimar la mayor cantidad de malas maniobras, hay que tener en cuenta que el costo es aumentar el error tipo I (acusar a un inocente).

Como esa decisión escapa de los objetivos de este trabajo, se calibrarán modelos según dos casos definidos arbitrariamente:

- Caso 1: se reduce el error tipo II a un valor cercano al 20%, manteniendo el error tipo I bajo el 10%.
- Caso 2: se reduce el error tipo II a un valor similar a aquel de error tipo I.

Debido a que la metodología se aplica a cada una de las cinco combinaciones de datos de entrenamiento y validación, se debe escoger la función kernel y los parámetros asociados para cada combinación. Para todas las combinaciones utilizadas, la función

kernel con mayor exactitud fue la lineal. Es importante señalar que si bien esto ocurrió para la muestra recabada, no necesariamente ocurrirá para una muestra distinta. Es decir, se debe encontrar el mejor kernel para cada muestra de datos a la que se le aplique esta metodología.

Por otro lado, como el valor de C_{-1} se mantiene invariable e igual a 1, en la Tabla 3-6 se muestran solo el valor del parámetro C_1 para ambos casos.

Tabla 3-6. Parámetro C escogido según caso y para cada combinación. Fuente: Elaboración propia.

	Combinación				
	1	2	3	4	5
Caso 1	7	6	6	7	7
Caso 2	12	10	13	14	12

La constante de penalización para la clase positiva es estable en cada combinación para el caso 1. Además, el valor es similar a la proporción inicial de los datos.

- **Árboles de decisión**

Similar al procedimiento usado para SVM, se debe encontrar la matriz de costo que penaliza equivocarse en la clasificación por parte del modelo (ver Tabla 3-1). En este caso el costo $C_{1,-1}$ equivale al costo C_1 utilizado en SVM. Del mismo modo, $C_{-1,1}$ equivale a C_{-1} .

Se fija $C_{-1,1} = 1$ y se varía el valor de $C_{1,-1}$ observando el impacto que tiene en las medidas de rendimiento. En la Figura 3-12 se aprecia un ejemplo para uno de los sets de entrenamiento.



Figura 3-12. Variación de las medidas de desempeño según el valor del parámetro $C_{1,-1}$ (Árboles de Decisión). Fuente: Elaboración propia.

A diferencia de lo ocurrido con SVM, al utilizar árboles de decisión solo se puede reducir el error tipo II hasta un mínimo (en este caso 33%). Tanto la exactitud y el error tipo I se mantienen estables ante variaciones de la matriz de costo.

Luego se procede a buscar el valor que toman los otros dos parámetros considerados por el método árboles de decisión: número máximo de divisiones y tamaño mínimo de un nodo de decisión. Aplicando el algoritmo *grid-search* se llega a los parámetros finales para cada combinación de datos de entrenamiento y validación (ver Tabla 3-7).

Tabla 3-7. Parámetros escogidos del modelo árboles de decisión para cada combinación. Fuente: Elaboración propia.

	Combinación				
	1	2	3	4	5
$C_{1,-1}$	7	6	6	7	7
Número máximo de divisiones	10	8	8	9	8
Tamaño mínimo de un nodo de decisión	10	10	10	10	10

El valor de los parámetros $C_{1,-1}$ y $C_{-1,1}$ toman los mismos valores que sus contrapartes en SVM para el caso 1. Esto lleva a concluir una consistencia en la forma en que los modelos abordan los datos. Por otro lado, el tamaño mínimo de un nodo de decisión se mantiene constante en 10. Finalmente, el número máximo de divisiones se mantiene estable entre 8 y 10. Es importante destacar la importancia de este último parámetro, ya que, sin él, los árboles se ramificarían de gran manera complicando su interpretabilidad y arriesgando que el modelo se enfrente a un problema de sobreentrenamiento.

- **KNN**

Siguiendo la línea del análisis realizado con los anteriores métodos, se debe encontrar la matriz de costo que penaliza equivocarse en la clasificación por parte del modelo. Primero se debe determinar un parámetro k inicial que optimice el rendimiento del modelo. Para esto se utiliza la función incorporada en el *software* MATLAB. Luego, se procede a variar el costo $C_{1,-1}$ observando el comportamiento de las medidas de rendimiento, lo que se puede apreciar en el ejemplo de la Figura 3-13.

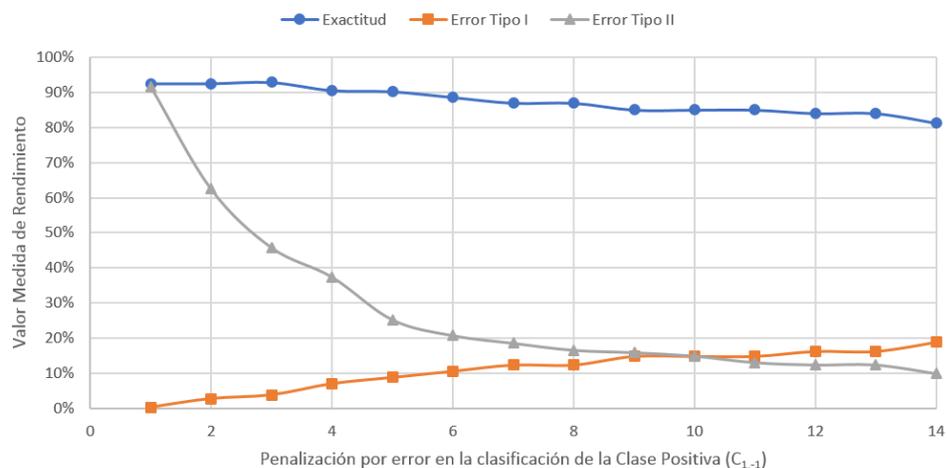


Figura 3-13. Variación de las medidas de desempeño según el valor del parámetro C_1 (KNN).
Fuente: Elaboración propia.

Tal como se observó con el método SVM, a medida que se reduce el error tipo II, la exactitud disminuye y el error tipo I aumenta. Por lo tanto, dependiendo del objetivo que quien implemente esta metodología, el valor de $C_{1,-1}$ que se escoge. Del mismo modo que con SVM, se calibrarán modelos para los mismos dos casos:

Una vez que se escoge el parámetro $C_{1,-1}$ para cada caso, se debe volver buscar el valor de k que optimice el rendimiento. Los parámetros finales se pueden encontrar en la Tabla 3-8 para ambos casos.

Tabla 3-8. Parámetros escogidos del modelo KNN para cada combinación según caso.
Fuente: Elaboración propia.

		Combinación				
		1	2	3	4	5
Caso 1	C_1	7	6	6	7	7
	k	19	3	36	31	57
Caso 2	C_1	12	10	13	14	12
	k	46	25	33	29	80

Se aprecia cómo una vez más, el valor de los parámetros $C_{1,-1}$ y $C_{-1,1}$ toman los mismos valores que sus contrapartes en SVM. Con esto, se confirma la consistencia en la forma en que los modelos abordan los datos.

4 RESULTADOS

En el presente capítulo se comenzará analizando el rendimiento obtenido por cada método considerando las 5 combinaciones de datos de entrenamiento y validación. En segundo lugar, se compararán los métodos y se plantearán las ventajas y desventajas de cada uno. Luego, se analizará los resultados de la clasificación realizada escogiendo uno de los tres métodos como referencia. Finalmente, se expresarán las consideraciones que tiene la aplicación de esta metodología.

4.1 Rendimiento de los Métodos de Clasificación utilizados

A continuación se mostrarán las medidas de rendimiento explicadas en la sección 3.3.1 para los tres métodos de clasificación calibrados: SVM, Árboles de Decisión y KNN.

4.1.1 SVM

Los resultados de la modelación se presentan en la Tabla 4-1 para ambos casos expuestos en la sección 3.3.2. Para las tres medidas de rendimiento se entrega el promedio y la desviación estándar obtenida para cada combinación de datos de entrenamiento y validación. Es relevante recalcar que el rendimiento se mide sobre el set de validación.

Tabla 4-1. Resultados del modelo SVM para ambos casos. Fuente: Elaboración propia.

	Caso 1		Caso 2	
	Promedio	Desviación Estándar	Promedio	Desviación Estándar
Exactitud	91%	1,10%	88%	1,61%
Error Tipo I	8%	1,19%	12%	1,63%
Error Tipo II	17%	4,17%	12%	4,56%

El caso 1 tenía como objetivo reducir el error tipo II hasta llegar a un valor cercano al 20%. Se puede apreciar como este objetivo se logra y el error tipo I se mantiene en niveles aceptables y con cierta estabilidad. Además, al utilizar SVM se logra una alta exactitud, sobre el 90%, con una muy baja desviación estándar. Esto implica que sin importar la combinación de datos de entrenamiento y validación utilizada, la exactitud y el error tipo I se mantendrán estables. Para el caso del error tipo II, si bien se tiene una desviación estándar mayor que en las otras dos medidas de rendimiento, se mantiene en los rangos aceptables según el objetivo buscado.

Debido a que los datos de validación representan un 20% del total, estos contienen apenas 21 datos positivos en promedio. En la combinación con mayor error tipo II, se clasifican correctamente 16. En la combinación con menor error tipo I, 18. Es por esto que la desviación estándar presentada se explica por tener una muestra desbalanceada con un bajo número de datos positivos, donde equivocarse en la predicción de un dato positivo más, aumenta el error tipo II considerablemente.

Por otro lado, el caso 2 tenía como objetivo igualar los errores tipo I y II. Al observar la Tabla 5-2, se concluye que este objetivo se cumple y ambos errores se mantienen en niveles aceptables. Si bien la exactitud en este caso es alta, es levemente inferior a la obtenida en el caso 1. Del mismo modo que con el caso 1, se observan similares niveles de desviación estándar, los que se mantienen relativamente estables en los tres indicadores. Esto permite concluir que las medidas de rendimiento se mantienen estables sin importar la combinación de datos de entrenamiento y validación utilizada.

Es importante señalar que la desviación estándar para todas las medidas es superior a la observada en el caso 1. Esto se puede explicar porque a medida que el modelo se vuelve menos exacto, existe mayor variabilidad en la predicción. Por otro lado, al igual que en el caso 1, la desviación estándar del error tipo II se explica por tener una muestra desbalanceada.

4.1.2 Árboles de decisión

En la Tabla 4-2 se muestran los resultados de la modelación para este método.

Tabla 4-2. Resultados del método árboles de decisión. Fuente: Elaboración propia.

Árboles de decisión		
	Promedio	Desviación Estándar
Exactitud	90%	1,40%
Error Tipo I	9%	1,29%
Error Tipo II	17%	5,10%

Los resultados son similares a los obtenidos con SVM para el caso 1. La exactitud y el error tipo I tienen una baja desviación estándar. En cambio, el error tipo II tiene una mayor desviación explicada por la naturaleza de la muestra. Al utilizar árboles de decisión se logra un modelo de alto rendimiento que es capaz de clasificar con gran exactitud. Este método, a diferencia de los otros dos utilizados en esta investigación, no es de caja negra. Esto implica que el modelo calibrado sí es visible al modelador y, de manera específica, este método entrega una representación gráfica que permite un mayor análisis posterior.

Como se calibra un modelo para cada combinación de datos de entrenamiento y validación, en la Figura 4-1 se puede apreciar el árbol calibrado para la combinación 1. El análisis que se expondrá se puede replicar fácilmente para las cuatro combinaciones restantes y sus respectivos árboles.

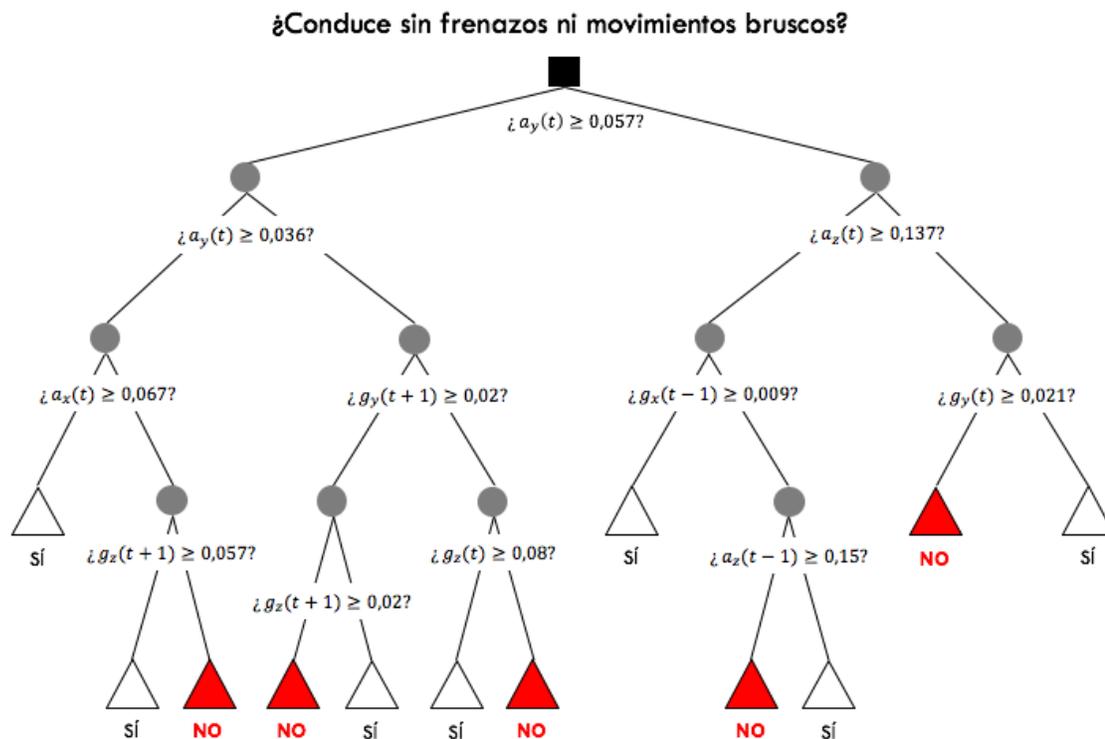
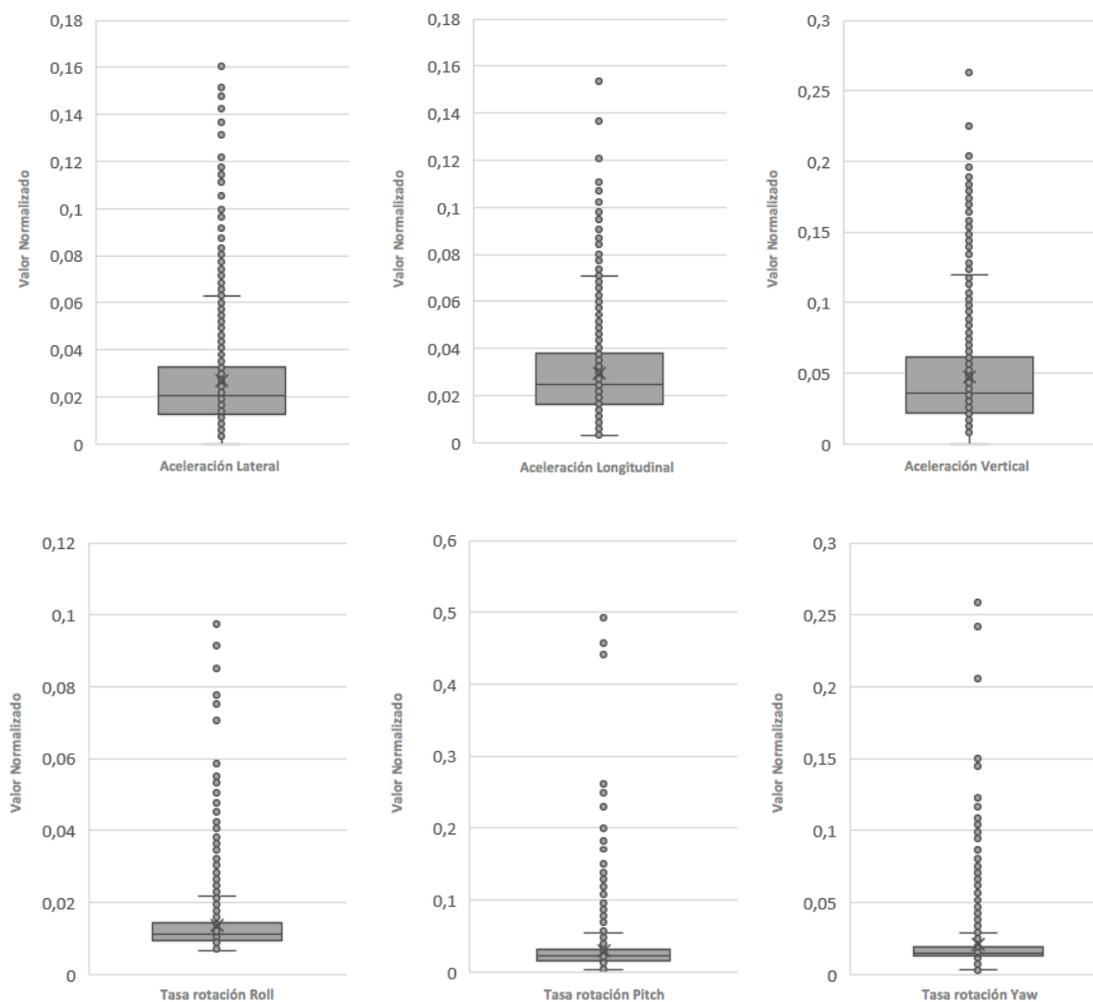


Figura 4-1. Árbol de decisión calibrado para la combinación 1. Fuente: Elaboración propia.

En la Figura 4-1, si la condición de un nodo se cumple se escoge la rama derecha; en caso contrario, la izquierda. Los árboles de decisión tienen la particularidad de ir jerarquizando las variables que considera para la clasificación. En el caso del árbol representado, la variable más relevante y que integra la primera condición es el RMS de la aceleración longitudinal (a_y). Luego, el modelo contempla el RMS de la aceleración vertical (a_z) y lateral (a_x) expresadas en [G], pero en diferentes niveles. Además, se consideran las rotaciones respecto a los tres ejes expresadas en $\left[\frac{rad}{s}\right]$.

Por otro lado, el uso de árboles permite encontrar una explicación a cada rama que lleva a clasificar la conducción como mala. Esto es posible gracias a que se conocen las variables consideradas por el modelo y los rangos numéricos con los que se evalúa cada condición. En la Figura 4-2 se presentan las dispersiones de los valores normalizados para las seis variables consideradas por el método.



**Figura 4-2. Box plots de las variables consideradas en el método de árboles de decisión.
Fuente: Elaboración propia.**

De la misma Figura 4-1, la rama que lleva al primer “NO” está caracterizada por una aceleración longitudinal bajo el percentil 75, una aceleración lateral sobre el valor máximo (sin considerar *outliers*) y una tasa de rotación *yaw* sobre el valor máximo en el instante siguiente. Esto se podría explicar por el conductor realizó un viraje de una pista a otra a alta velocidad lo que finalmente se traduce en una maniobra incómoda para los pasajeros.

La rama que lleva al segundo “NO” está caracterizada por una aceleración longitudinal sobre el percentil 75 y bajo el valor máximo, una tasa de rotación *pitch* bajo el percentil 50 en el instante siguiente y una tasa de rotación *yaw* bajo el percentil 75 en el instante siguiente. Al considerar bajas tasas de rotación, se podría asumir que el conductor no realizó movimientos que llevan a un giro del bus, sino que son vibraciones basales del dispositivo. Por lo tanto, dado el rango de la aceleración longitudinal se podría decir que el movimiento es un frenazo o partida fuerte para detener el bus ante un semáforo o en congestión.

En la misma línea del caso anterior, la rama que lleva al tercer “NO” está caracterizada por una aceleración longitudinal sobre el percentil 75 y bajo el valor máximo, una tasa de rotación *pitch* sobre el percentil 50 en el instante siguiente y una tasa de rotación *yaw* sustancialmente mayor al valor máximo en el instante actual. Al igual que en el caso anterior el rango de aceleración longitudinal explicarían una detención o partida brusca. Pero, por las tasas de rotación se puede suponer que esta detención o partida es en un paradero, ya que el vehículo debe acercarse a la vereda para dejar o recoger pasajeros.

En cambio, la rama que lleva al cuarto “NO” está caracterizada por una aceleración longitudinal sobre el valor máximo, una aceleración vertical bajo el percentil 25 y una tasa de rotación *roll* sobre el percentil 25 en el instante anterior. Esto se podría explicar porque el conductor cambió bruscamente la velocidad ante por ejemplo un lomo de toro, que es lo que ocasionaría el giro respecto al eje *x*. Si bien es esperable una reducción de velocidad considerable ante lomos de toro, esta debiera ser gradual.

Finalmente, la rama que lleva al quinto “NO” está caracterizada por una aceleración longitudinal sobre el valor máximo, una aceleración vertical sobre el percentil 25 y una tasa de rotación *pitch* bajo el percentil 50. Esto se podría explicar por alguna irregularidad leve en el pavimento que produce incomodidad en los pasajeros por la alta velocidad a la que circula el bus. Si la tasa de rotación *pitch* fuera alta, implicaría que existe una irregularidad mayor (por ejemplo: un bache) y que según las consideraciones de la metodología, no se debe atribuir a las maniobras del conductor.

Al conocer la estructura del árbol se puede determinar que proporción de las maniobras clasificadas como inapropiadas se deben a alguna de las razones expuestas anteriormente. Los resultados de este análisis se pueden apreciar en la Tabla 4-3.

Tabla 4-3. Proporción de maniobras clasificadas como inapropiadas según su posible explicación.
Fuente: Elaboración propia.

Rama	Proporción	Posible Explicación
1°	5%	Viraje a alta velocidad
2°	46%	Detención o partida brusca ante un semáforo o en congestión
3°	30%	Detención o partida brusca en un paradero
4°	12%	Lomo de toro a alta velocidad
5°	7%	Irregularidad leve en el pavimento

Como era de esperar, existe una consistencia entre la proporción clasificada y la posible explicación a por qué se consideró maniobra inapropiada. En un recorrido común, como los analizados en el experimento, la mayor cantidad de maniobras bruscas se produce en las detenciones. Cuando el conductor se detiene ante un semáforo o cuando se encuentra en congestión, es el escenario que más produce maniobras incómodas. En segundo lugar se tiene la detención ante paraderos. Dado el trazado de los recorridos, no existen muchos virajes lo que explica la baja proporción de dicha rama. Del mismo modo, maniobras bruscas ante irregularidades en el pavimento que se puedan atribuir al conductor son muy pocas dado las consideraciones de la metodología.

4.1.3 KNN

En la Tabla 4-4 se muestran los resultados de la modelación para los dos casos considerados.

Tabla 4-4. Resultados del modelo KNN para el caso 1. Fuente: Elaboración propia.

	Caso 1		Caso 2	
	Promedio	Desviación Estándar	Promedio	Desviación Estándar
Exactitud	88%	1,45%	85%	2,26%
Error Tipo I	11%	1,43%	15%	2,77%
Error Tipo II	24%	6,18%	14%	4,75%

Tal como se observó con los métodos SVM y árboles de decisión, la exactitud y el error tipo I tienen una baja desviación estándar. En cambio, el error tipo II tiene una mayor desviación que se explica por la muestra no balanceada. El caso 1 tenía como objetivo reducir el error tipo II bajo un 20%. Si bien el método KNN no logra ese objetivo, consigue reducirlo a un cuarto de toda la clase positiva sin incrementar considerablemente el error tipo I. Aun cuando la exactitud es inferior al 90%, esta se puede seguir considerando como alta.

Por otro lado, el caso 2 tenía como objetivo igualar los errores tipo I y II. De la Tabla 4-2 se puede concluir que este objetivo se cumple y ambos errores se mantienen en niveles aceptables. Además, se logra una exactitud superior al 85%, lo que se puede considerar un buen rendimiento. Del mismo modo que con el caso 1, se tiene una baja desviación estándar para la exactitud y el error tipo I. Es decir, estas medidas de rendimiento se mantienen estables sin importar la combinación de datos de entrenamiento y validación utilizada.

4.2 Comparación de los métodos calibrados

De la sección anterior se concluye que el uso de máquinas de aprendizaje entrega resultados con buen rendimiento en la clasificación de la forma de conducción. Además, se pudo ver consistencia en los resultados de los tres métodos a lo largo de las cinco

combinaciones de datos de entrenamiento y validación. Con esto se puede determinar que los modelos no presentan un problema de sobre-entrenamiento de la muestra.

En la Tabla 4-5 se puede apreciar una tabla resumen con el promedio obtenido en cada una de las medidas de rendimiento por los tres métodos para ambos casos.

**Tabla 4-5. Tabla resumen del rendimiento obtenido por cada método de clasificación.
Fuente: Elaboración propia.**

	SVM		Árboles de decisión	KNN	
	Caso 1	Caso 2		Csso 1	Caso 2
Exactitud	91%	88%	90%	88%	85%
Error Tipo I	8%	12%	9%	11%	15%
Error Tipo II	17%	12%	17%	24%	14%

De los tres métodos, SVM es aquel que obtiene mejores resultados en todas las medidas de rendimiento, aunque el método árboles de decisión obtiene un rendimiento bastante similar. Si bien KNN es el método con el peor rendimiento, se mantiene un nivel de predicción aceptable.

Al comparar SVM y KNN para el caso 2, el primero es el método que obtiene mejores resultados en todas las medidas de rendimiento, lo que demuestra el poder de las máquinas de soporte vectorial por sobre k vecinos más cercanos. En general, y como era de esperar, el rendimiento de todos los métodos es inferior respecto al obtenido en el caso 1. Al ser una muestra desbalanceada, al aumentar el costo de penalización se logra aprender algunos elementos más de la clase positiva, pero clasificando incorrectamente muchos más elementos de la clase negativa.

Al considerar la simplicidad de la implementación de cada método, KNN es el más sencillo de ejecutar. En cambio SVM, el método con mejor rendimiento, requiere de una calibración más compleja que el resto de los métodos evaluados. A pesar de lo anterior, una vez calibrado cada método, la aplicación es simple.

Respecto a la interpretabilidad del modelo calibrado, tanto SVM como KNN son métodos de clasificación de caja negra, es decir, entregan resultados pero no un modelo visible del funcionamiento de la máquina para llegar desde los datos de entrada a los de salida. En cambio, los árboles de decisión son capaces de entregar un diagrama que muestra la jerarquía asignada a las variables de entrada, así como mostrar cuáles son las variables que considera el modelo. El nivel de análisis de los resultados y del modelo mismo es mucho más detallado que con los métodos de caja negra.

Cabe señalar que no está dentro de los alcances de esta investigación determinar el mejor método para la clasificación, ya que depende de los objetivos que tenga la autoridad o empresa operadora que implemente la metodología. A pesar de esto, si el objetivo se alinea con el caso 1 expuesto, se recomienda el uso de árboles de decisión por su buen rendimiento y alta interpretabilidad del modelo calibrado. Del mismo modo, si el objetivo se alinea con el caso 2, se recomiendan las máquinas de soporte vectorial por su gran rendimiento por sobre k vecinos más cercanos.

4.3 Análisis de la clasificación realizada

Debido a que la metodología busca clasificar la forma de conducir de los choferes de los recorridos considerados, en esta sección se procede a realizar un análisis de los resultados de dicha clasificación. Para efectos ilustrativos, se presentan los resultados obtenidos al aplicar el método árboles de decisión sobre el set de validación de la combinación 1. Del total de 303 instancias (o vectores de atributos) del set de validación, donde cada uno representa una maniobra de 3 segundos, el modelo indicó que 40 de ellas correspondían a maniobras bruscas de conducción. Esto es, un 13% del total.

En la Tabla 4-6 se presenta la cantidad de maniobras clasificadas como bruscas según recorrido.

Tabla 4-6. Resultado de la clasificación según recorrido. Fuente: Elaboración propia.

Recorrido	Maniobras bruscas	Total de maniobras	Proporción del Total	Ránking ICA	Tipo de Bus
102	4	40	10%	2°	No articulado
210	8	35	23%	7°	Articulado
210V	10	53	19%	7°	Articulado
213E	7	82	8%	1°	No articulado
712	9	93	10%	1°	No articulado

Se puede apreciar cómo los tres recorridos (102, 213E y 712) que pertenecen a empresas operadoras bien evaluadas por el ICA tienen menos maniobras clasificadas como bruscas y, que en este caso, operan con buses no articulados. Del total de maniobras para estos recorridos, aproximadamente un 10% es considerado como conducción brusca. En el caso de los dos recorridos restantes (210 y 210V), se llega a aproximadamente un 20% de las maniobras clasificadas como bruscas. Por lo tanto, se nota una correlación entre las maniobras clasificadas por el modelo calibrado con el tipo de bus y/o el ranking otorgado por la metodología actualmente empleada.

Si bien la metodología considera la conducción y no factores externos como el tipo de bus, la correlación encontrada se puede explicar porque los buses articulados suelen estar en peor estado. Es esperable que si el chofer conduce a una alta velocidad, las vibraciones a bordo y los movimientos de la parte trasera (articulada) del bus serán percibidos de manera más brusca que en un bus no articulado circulando a la misma velocidad. De todos modos, es importante señalar que debido a que el experimento es muy acotado, no se puede concluir que en todos los buses articulados se presentarán peores resultados en la conducción. Sin embargo, es un análisis que se puede realizar una vez implementada la metodología a gran escala.

De la Tabla 4-3 se pudo inferir que casi el 90% de las maniobras clasificadas como bruscas se deben a cambios de velocidad ante una detención o partida. Por lo tanto, un análisis relevante sería conocer la proporción del total de variaciones de velocidad que

fueron incómodas para los pasajeros. Debido a que no es trivial detectar una reducción o aumento de velocidad de la trayectoria (ver Figura 4-3), se aplicará el concepto de curvas oblicuas (Muñoz & Daganzo, 2002) para una detección más sencilla.

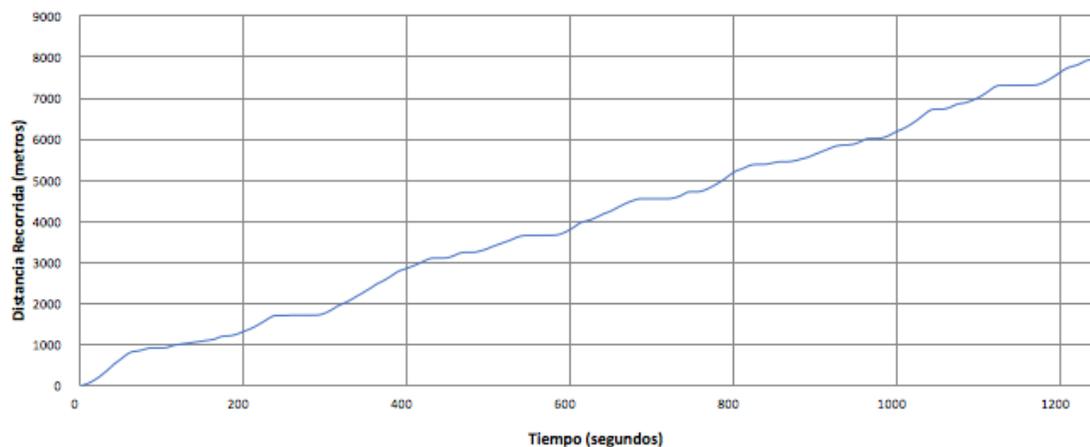


Figura 4-3. Trayectoria del bus 213E. Fuente: Elaboración propia.

En la Figura 4-4 se puede apreciar la trayectoria con curvas oblicuas, donde cada vez que se cambia de signo en la pendiente existe una variación de velocidad significativa. Cuando el cambio es de una pendiente positiva a una negativa, significa que el vehículo redujo su velocidad. En caso contrario, la incrementó. Los puntos de color naranja toman valor 1 cuando el instante corresponde a una variación de velocidad. Este proceso se realiza por inspección visual.

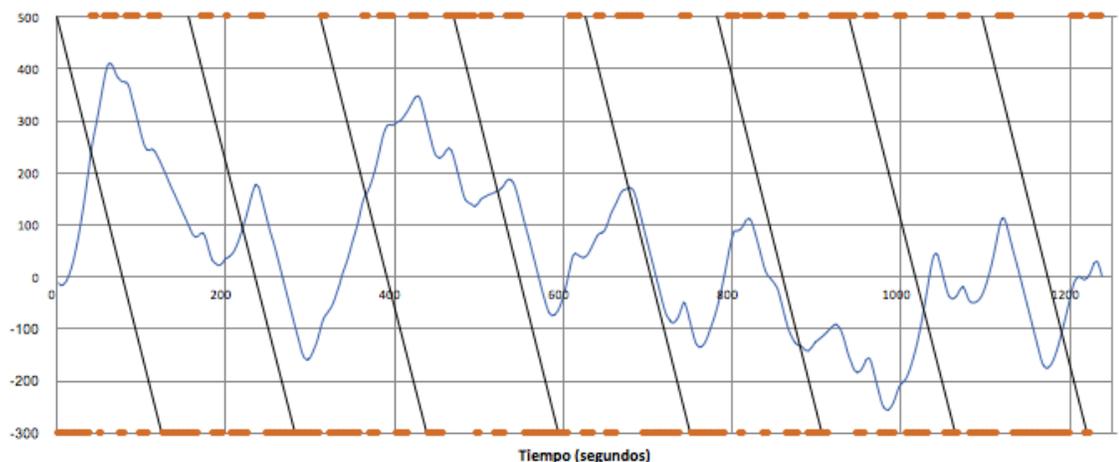


Figura 4-4. Trayectoria del bus 213E en coordenadas oblicuas. Fuente: Elaboración propia.

Una vez identificados los instantes que implican una variación de velocidad significativa, se procede a obtener la proporción de maniobras que fueron clasificadas como bruscas respecto de esa muestra. Los resultados se pueden apreciar en la Tabla 4-7.

Tabla 4-7. Proporción de maniobras bruscas dada una reducción de velocidad significativa. Fuente: Elaboración propia.

Recorrido	Proporción maniobras bruscas
102	25%
210	64%
210V	55%
213E	25%
712	28%

En coherencia con lo apreciado anteriormente, los recorridos 210 y 210V presentan los peores resultados respecto a conducción superando el 50%. Los recorridos 102, 213E y 712 se encuentran en torno al 25% de las variaciones de velocidad clasificadas como bruscas.

Cuando se observa la proporción obtenida en la Tabla 4-6 que es calculada respecto del total de maniobras del conductor, se puede concluir que considerar todos los instantes no es una medida representativa. Si el bus va en línea recta y a velocidad constante es difícil que se genere una maniobra brusca, por tanto, al utilizar la proporción de variaciones de velocidad catalogadas como bruscas, se puede generar un estándar para evaluar la conducción que puede ser incorporado en un nuevo atributo del ICA. Se puede establecer un umbral donde proporciones mayores pueden ser sancionadas monetariamente.

Es relevante mencionar que los recorridos 210, 210V y 213E comparten gran parte del trayecto (a lo largo de Av. Vicuña Mackenna). Como se puede observar en la Figura 3-2 de la sección 3.1.2, el recorrido 213E comparte 3 kilómetros de trayecto con el recorrido 210 y 4,9 kilómetros con el recorrido 210V. En la Tabla 3-8 se presenta la velocidad promedio de estos tres recorridos.

Tabla 4-8. Velocidad comercial promedio para los recorridos que comparten trayectoria.
Fuente: Elaboración propia.

Recorrido	Velocidad Promedio $\left(\frac{km}{h}\right)$
210	20
210V	21
213E	23

Se aprecia como las velocidades son similares, pero la conducción es notoriamente distinta. Incluso, el recorrido 213E presenta la mayor velocidad promedio, pero los mejores resultados de comodidad percibida. Esto lleva a concluir que una conducción apropiada no implica un desmedro en la operación del servicio, es decir, se puede recorrer un trayecto igual de rápido pero sin frenazos bruscos.

En la misma línea que el análisis anterior, gracias al GPS incorporado en el dispositivo inercial, se puede geolocalizar cada instante medido. Como se mencionó

anteriormente, el recorrido 213E comparte parte de su trayecto con el recorrido 210V. A lo largo de la ruta, existe una curva pronunciada (ver Figura 4-5) de 300 metros de largo. Si esta curva es tomada a alta velocidad se produce incomodidad para los pasajeros.

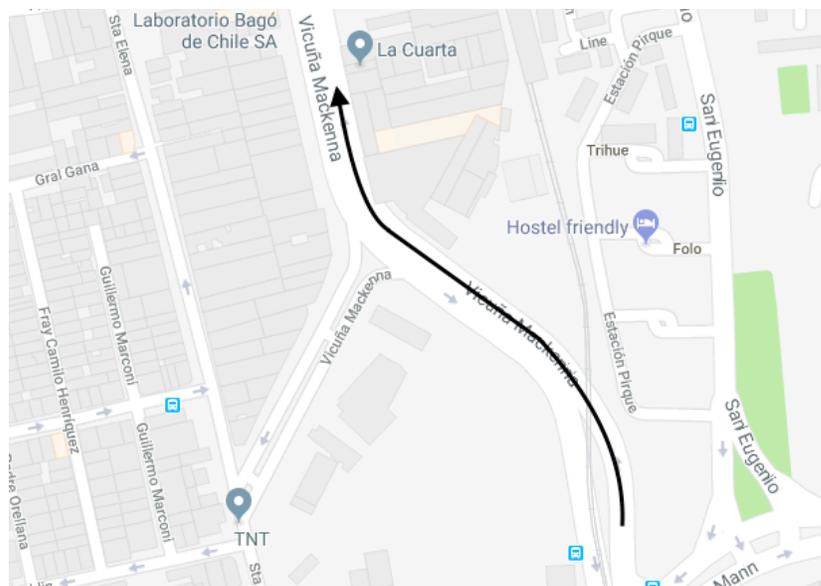


Figura 4-5. Curva por la cual transitan los buses 213E, 210 y 210V. Fuente: Elaboración propia.

En la Figura 4-6 se presenta la velocidad a lo largo de la curva de Figura 5-4 para los buses 213E y 210V. Es importante señalar que el recorrido 213E lo hace en el sentido de la flecha y el recorrido 210V, en el contrario.

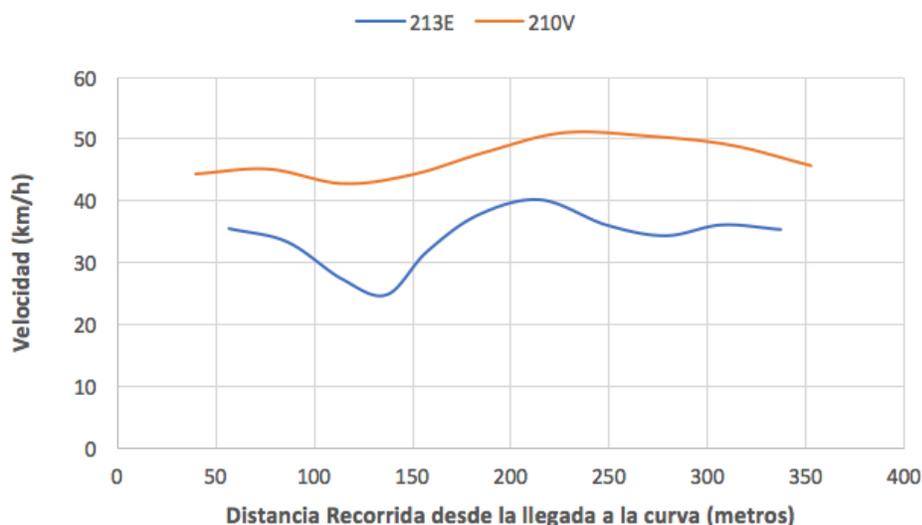


Figura 4-6. Velocidad a lo largo de la curva por la cual transitan los buses 213E y 210V.
Fuente: Elaboración propia.

Del gráfico se puede notar como el bus 210V llega a la curva con una mayor velocidad que el bus 213E. A lo largo de la curva circula a una velocidad un 38% mayor en promedio. Se aprecia como el chofer del bus 213E reduce su velocidad en 11 km/h para tomar la curva y luego acelera, en cambio, el chofer del bus 210V apenas disminuye su velocidad en 3 km/h. Esto se traduce en una peor percepción de la comodidad por parte de los pasajeros.

El efecto que juega el tipo de bus toma importancia en estas circunstancias. El bus 210V al ser articulado vibra más y la parte trasera se mueve mucho más que en el caso de un bus no articulado. Si bien solo se deben considerar las maniobras del conductor, al manejar un vehículo en peor estado él debería manejar con mayor cuidado para evitar producir incomodidad en los usuarios.

Se puede apreciar la importancia que tiene este análisis comparativo de la conducción de distintos conductores y recorridos, ya que al observar ciertas prácticas (como reducir la velocidad ante una curva) con buenos resultados de comodidad se pueden generar instancias de capacitación a los choferes con malos resultados. Por otro lado, si se geolocalizan puntos donde la gran mayoría de los conductores obtienen malos resultados,

se puede observar en terreno qué problemas tiene la infraestructura (calles en mal estado) o el trayecto (curvas muy cerradas). De este modo, y en la medida de lo posible, se pueden tomar medidas correctivas como capacitaciones o cambios en el trayecto del recorrido.

5 CONCLUSIONES

Uno de los objetivos específicos de esta investigación corresponde a evaluar la factibilidad de usar métodos de aprendizaje supervisado para la clasificación de la conducción en el transporte público. Para cumplir con ello, se consideraron tres métodos: máquinas de soporte vectorial (SVM), árboles de decisión y K vecinos más cercanos (KNN).

Además, esta tesis busca validar la metodología con datos reales, pero en una muestra controlada. Para ello se consideraron 5 recorridos de Transantiago: 102, 210, 210V, 213E y 712. Mediante pasajero incógnito se clasificó la conducción en cada instante y se registró con un dispositivo inercial las aceleraciones en los tres ejes y los movimientos del giroscopio en cada eje. Luego se construye una base de datos con la cual se entrenaron y validaron los métodos de clasificación.

De los tres métodos mencionados, SVM es aquel que presenta un mayor rendimiento medido según tres indicadores: exactitud, error tipo I y error tipo II. Si bien se aconseja utilizar solo el primero para métodos de aprendizaje, dado que la muestra es desbalanceada se deben incorporar los errores para controlar el aprendizaje apropiado de ambas clases (positiva y negativa).

A pesar de lo anterior, árboles de decisión es el segundo método con mayor rendimiento, apenas un 1% inferior a SVM. Pero, este método entrega un producto interpretable con el cual se puede analizar mucho mejor la clasificación realizada. Esta es una particularidad que no poseen los métodos de caja negra como SVM y KNN. Respecto de KNN, si bien es el método con menor rendimiento, es el más simple de calibrar.

A partir del árbol calibrado se pueden conocer las variables que el modelo considera y la jerarquía en la cual son evaluadas. En el caso observado se llega a que la aceleración longitudinal es la variable más importante, consistente con lo determinado a partir de la inspección visual. Por otro lado, del árbol se puede buscar una posible explicación a cada rama que lleva a determinar una conducción brusca en base a las variables y umbrales encontrados.

Ilustrativamente, un árbol llegó a que la detención brusca ante un semáforo o en congestión es la principal instancia donde se generan maniobras bruscas (46%). En segundo lugar se tiene la detención en paraderos (30%) y en tercer lugar, tomar un lomo de toro a alta velocidad (12%). Las otras dos razones explican menos del 10% de las maniobras bruscas cada una. Estas posibles explicaciones son consistentes con lo observado en el experimento.

Para analizar la clasificación realizada, se toma en cuenta el método de árboles de decisión como referencia. Se concluye que existe una clara relación entre la proporción de maniobras consideradas bruscas y el ranking entregado por la actual metodología del ICA. De los recorridos 102, 213E y 712 se tiene que aproximadamente un 10% de la duración del recorrido presenta maniobras bruscas. En contraste, los recorridos 210 y 210V, un 20% aproximadamente.

Más aún, se observa que el tipo de bus (articulado o no articulado) puede influir de manera indirecta a la cantidad de maniobras bruscas detectadas. La explicación se debe a que la misma forma de conducir de un chófer produce mayor incomodidad si el bus se encuentra en mal estado (como es el caso de varios buses articulados). Es decir, se requiere que el conductor sea más precavido al conducir estas máquinas, lo que puede indicar ir a menor velocidad e incumplir las metas operacionales impuestas por la Empresa a cargo del recorrido.

Al realizar un análisis mediante curvas oblicuas de la trayectoria de cada bus, se pudo filtrar la base de datos según maniobras que involucran una variación de velocidad considerable (ya sea para detenerse o acelerar). Respecto a este nuevo universo, se llega a que los recorridos 210 y 210V presentan una proporción considerablemente mayor de variaciones de velocidad bruscas, específicamente, un 50% en contraste al 25% obtenido en los recorridos 102, 213E y 712.

Además, se propone que se utilice la proporción de variaciones de velocidad catalogadas como bruscas como un nuevo atributo de conducción del ICA. Un nivel de 25% se puede considerar aceptable, donde proporciones mayores sean sancionadas monetariamente.

Por otro lado, gracias al GPS incorporado en el dispositivo inercial, se puede geolocalizar cada instante medido. Del experimento realizado se tiene que el recorrido 213E comparte parte de su trayecto con el recorrido 210V. Se pudo comprobar que una curva pronunciada presente en el trayecto fue tomada por el chofer del bus 210V a una velocidad un 38% mayor que su contraparte del bus 213E. Además, se apreció como el este último chofer redujo su velocidad antes de comenzar la curva, mientras que el conductor del bus 210V no lo hizo. Sumado al hecho de que el bus 210V es articulado y en mal estado, la percepción de comodidad fue notoriamente inferior.

Este análisis comparativo de la conducción de distintos conductores y recorridos es muy relevante, ya que se pueden detectar buenas prácticas (como reducir la velocidad ante una curva) y capacitar a aquellos conductores que no las realizan. Por otro parte lado, si se geolocalizan puntos donde la gran mayoría de los conductores obtienen malos resultados, se pueden tomar medidas correctivas para mejorar el servicio.

Respecto de la metodología propuesta, esta permite una evaluación constante de todos los conductores y líneas de un operador. Además, se registra la conducción de todo el viaje y no solo una ventana de tiempo como se aplica actualmente en la metodología del ICA. Adicionalmente, la nueva metodología permite realizar un seguimiento de la conducción de un chofer, ya que se registran sus maniobras para una misma ruta a lo largo del tiempo. En conjunto con las multas por incumplimiento estipuladas en los contratos, se genera un incentivo para que exista una mejora continua del servicio entregado.

La nueva metodología permite realizar un seguimiento de la conducción de un chofer, ya que es posible registrar las maniobras para una misma ruta a lo largo del tiempo. En caso de que se detecte un chofer con mal desempeño, se puede capacitar al mismo para que mejore la forma de conducción. Además, se puede premiar a aquellos conductores que se destaquen dentro de una empresa operadora.

La nueva metodología permite calibrar un modelo con buen rendimiento sin importar variables externas como tipo de bus, conductor, horario del día, etc. Además, se destaca por ser simple de implementar. No se necesita una gran cantidad de datos para obtener buenos resultados, por lo que el muestreo utilizado actualmente (20% de los

buses) es suficiente para calibrar el modelo. Los buses utilizados para calibrar el modelo deben ser escogidos de manera estratégica y no aleatoria, velando por obtener una muestra diversa en tipo de material rodante, tipos de calles recorridas (avenidas, corredores segregados, pistas solo bus, autopistas, etc.) y condiciones geográficas (recorridos en cerros o calles con pendiente).

Debido a que pueden existir cambios en las condiciones externas a la conducción (renovación del material rodante, cambios en el pavimento o modificaciones en los recorridos), se aconseja realizar una evaluación con pasajeros incógnitos de manera periódica a fin de validar el modelo. El intervalo exacto entre mediciones debe ser determinado por la autoridad que implemente la metodología y dependerá de la naturaleza de los cambios externos.

Hay que destacar el bajo costo que tienen los dispositivos inerciales, ya que están presentes en todos los *smartphones* comercializados actualmente. Esto permite una implementación a gran escala y a un bajo costo de implementación. Se analizó el caso de Transantiago debido a que ya existen indicadores implementados y el sistema ha sido modernizado considerablemente en los últimos años, creando nuevas oportunidades de investigación con cada nuevo avance. A pesar de ello, la metodología propuesta es válida para implementar en otros sistemas de transporte público (regiones de Chile o, incluso, otros países).

Se desea recalcar que esta investigación no tiene como objetivo calibrar un modelo para que sea aplicado a todo Santiago, en cambio, se presenta una nueva metodología que puede ser replicada a gran escala. En esta línea, el experimento consideró una muestra acotada del total de recorridos disponibles en la capital.

Como futuras líneas de investigación se sugiere modificar la clasificación binaria de la conducción por una que incorpore al menos tres niveles: buena, regular o mala. Esto debido a que las máquinas de aprendizaje pueden cometer errores en la clasificación si es que la línea que separa las actuales categorías es difusa. Adicionalmente, se sugiere evaluar el uso de máquinas de aprendizaje no supervisado, ya que esto permitiría ahorrar

aún más recursos en horas hombre al no ser necesaria la clasificación por parte de pasajeros incógnitos.

Otra arista de esta investigación sería probar la metodología en circunstancias geográficas radicalmente distintas (ejemplo: los cerros de Valparaíso). Esto permitiría comprobar la versatilidad que tiene la metodología propuesta. Finalmente, se sugiere programar un algoritmo que permita realizar la clasificación de la conducción directamente en el dispositivo inercial. De esta manera solo se almacenaría la información de las maniobras de los conductores, logrando un ahorro considerable del almacenaje de información asociado a la implementación de la metodología a gran escala.

REFERENCIAS

- Bodini, I., Lancini, M., Pasinetti, S., y Vetturi, D. (2014). Techniques for on-board vibrational passenger comfort monitoring in public transport. *Acta IMEKO*, 3(4), 32-37.
- Boser, B., Guyon, I., & Vapnik, V. (1992). A Training Algorithm for Optimal Margin Classifiers. Artículo presentado en *5th Annual ACM Workshop on Computational Learning Theory*, Pittsburgh, EE.UU.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.
- British Standards Institution. (1999). *Railway applications: Ride comfort for passengers – Measurement and evaluation* (DD ENV 12299:1999). London: Author.
- Cascetta, E., & Carteni, A. (2014). A quality-based approach to public transportation planning: theory and a case study. *International Journal of Sustainable Transportation*, 8, 84-106.
- Castellanos, J.C., Susin, A., y Fruett, F. (5-7 Octubre 2011). Embedded sensor system and techniques to evaluate the comfort in public transportation. Artículo presentado en *14th International IEEE Conference on Intelligent Transportation Systems*, Washington D.C., EE.UU.
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3), 273-297.
- Cronin, J., & Taylor, S. (1992). Measuring service quality: a reexamination and extensión. *Journal of Marketing*, 56(3), 55-68.
- Chang, C., & Lin, C. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1-27.

de Oña, R., Machado, J.L., & de Oña, J. (2015). Perceived Service Quality, Customer Satisfaction, and Behavioral Intentions. *Transportation Research Record: Journal of the Transportation Research Board*, 2538, 76-85.

dell'Olio, L., Ibeas, A., & Cecin, P. (2011). The quality of service desired by public transport users. *Transport Policy*, 18(1), 217-227.

Directorio de Transporte Público Metropolitano. (2014). *Manual de Medición ICA 2013*.

Dirección de Transporte Público Metropolitano. (2015). *Informe de Gestión Transantiago 2014*.

Directorio de Transporte Público Metropolitano. (2016a). *Estudio Satisfacción de Empresas Operadoras*.

Dirección de Transporte Público Metropolitano. (2016b). *Consolidado ICA a Abril 2016* [Base de Datos].

Figueiredo, M., y Silva, L.F. (12-16 Julio 2015). Public Transport: Whole-Body Vibration and Comfort of Passengers, Drivers and Collectors. Artículo presentado en *22nd International Congress on Sound and Vibration*, Florencia, Italia.

Grujičić, D., Ivanović, I., Jović, J., y Đorić, V. (2014). Customer perception of service quality in public transport. *Transport*, 29(3), 285-295.

Han, S. (20-23 Octubre 2013). A study on the ride comfort control using automatic generating data base on board for Electric railway vehicle. Artículo presentado en *13th International Conference on Control, Automation and Systems*, Kimdaejeung Convention Center, Gwangju, Corea del Sur.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2001). *The elements of statistical learning : Data mining, inference, and prediction*. New York: Springer Verlag.

Hsu, C., Chang, C., & Lin, C. (2016). *A Practical Guide to Support Vector Classification*.

International Organization for Standardization. (1997). *Mechanical vibration and shock: Evaluation of human exposure to whole-body vibration – Part 1: General requirements (ISO 2631-1:1997)*. Geneva: Author.

Lai, W., & Chen, C. (2011). Behavioral intentions of public transit passengers—The roles of service quality, perceived value, satisfaction and involvement. *Transport Policy*, 18(2), 318-325.

Martilla, J., & James, J. (1977). Importance-performance analysis. *Journal of Marketing*, 41(1), 77-79.

Mikrasya (2016). *Counter – Advanced Tally Counter (Versión 7.1)* [Software de Aplicación Móvil].

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). *Foundations of machine learning*. Cambridge, MA: The MIT Press.

Ortúzar, J. de D., & Willumsen, L. (2011). *Modelling Transport*. West Sussex, Reino Unido: John Wiley & Sons Ltd.

Parasuraman, P., Zeithaml, V., & Berry, L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49(4), 41-50.

Parasuraman, P., Zeithaml, V., & Berry, L. (1988). SERVQUAL: A multiple – Item Scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12-40.

Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning*, 1, 81-106.

Rokach, L., & Maimon, O. (2008). *Data mining with decision trees: theory and applications*. World Scientific Pub Co Inc.

Sánchez-Flores, O., y Romero-Torres, J. (2010). Factores de calidad del servicio en el transporte público de pasajeros: estudio de caso de la ciudad de Toluca, México. *Economía, Sociedad y Territorio*, 10(32), 49-80.

Thomas, B. (2017). SensorLog (Versión 1.9.2) [Software de Aplicación Móvil].

Wang, X., An, K., Tang, L., & Chen, X. (2015). Short term prediction of freeway exiting volume based on SVM and KNN. *International Journal of Transportation Science and Technology*, 4(3), 337-352.

Zhang, K., Zhou, K., & Zhang, F. (2014). Evaluating bus transit performance of chinese cities: Developing an overall bus comfort model. *Transportation Research Part A: Practice and Policy*, 69, 105-112.