

Prototipo de un sistema para el seguimiento del servicio de transporte público colectivo en tiempo real basado en aprendizaje de máquina y LoRaWAN



Tesis de Trabajo de Grado

Modalidad: Trabajo de Investigación

Andersson Stive Diaz Lucio.

Código: 100616011395

Camilo Andrés Sánchez Diaz.

Código: 100614020986

Director: MSc. Ricardo Salazar Cabrera

Co-Directora: MSc. Mary Cristina Carrascal Reyes

Asesor: PhD. Álvaro Pachón de la Cruz

Universidad del Cauca

Facultad de Ingeniería Electrónica y Telecomunicaciones

Departamento de Telemática

Línea de Investigación: Aplicaciones y Servicios sobre internet

Popayán, marzo 2023



A mi familia y amigos por apoyarme durante mi proceso de formación.

Andersson Stive Díaz Lucio.

Gracias a Dios, mi familia y amigos por brindarme su ayuda en el desarrollo de este trabajo de grado y por el apoyo y motivación durante mi proceso de formación en la Universidad.

Camilo Andrés Sánchez Díaz



Agradecimientos

Expresamos los agradecimientos a nuestro director, el MSc. Ricardo Salazar Cabrera, por guiarnos en el desarrollo del presente trabajo a través de sus correcciones y recomendaciones. Agradecemos su dedicación y paciencia ante los inconvenientes que encontramos a lo largo de todo este tiempo, y por compartir con nosotros sus conocimientos y experiencias.

Al ingeniero Miguel Sotaquirá, creador de la academia *Codificando Bits*, por su contribución en el desarrollo del algoritmo usado en este proyecto.

Queremos expresar nuestra sincera gratitud a nuestra alma mater, la Universidad del Cauca, por proporcionarnos la oportunidad de adquirir una formación de calidad y por apoyar nuestro desarrollo en profesionales.

A la comunidad *Open-Source* por ser una gran fuente de aprendizaje y conocimiento para todos los involucrados, al fomentar la colaboración y el trabajo en equipo a través de la valiosa labor que realiza.

Además, queremos agradecer a nuestros compañeros de estudio, por su amistad y por compartir sus conocimientos y experiencias durante el proceso de formación. Finalmente, queremos agradecer a nuestras familias y a mis amigos por su apoyo incondicional y por estar siempre dispuestos a escuchar nuestras ideas y a ayudarnos en momentos de dificultad.



Abstract

Public transport in a city should be the most efficient, least polluting and most accessible means of transportation for its citizens. However, serious shortcomings have been detected in this means of transport, mainly in intermediate cities in developing countries [1]. These shortcomings are related to the lack of user information (on routes, schedules, stopping points), lack of road safety (due to speeding vehicles, or non-compliance with traffic regulations) [2], low service availability (non-compliance with routes and schedules), continuous stops in inappropriate and/or unauthorized places. Some of these problems contribute to the main mobility problems in cities, which are high accident rates and traffic congestion [1]. The development of tools to improve the characteristics and conditions of public transport service in cities has become an imperative need to improve the quality of life of citizens. The control and monitoring of public transport service is relevant in aspects such as online information of its location to travelers, and the control of transport companies for compliance with speed limits, schedules, routes and stops [1].

This work proposes a public vehicle tracking system by integrating a global positioning system (GPS) and *LoRa* (Long Range) technology using its *LoRaWAN* network protocol, also offering real-time location prediction in the absence of position data. The prediction was implemented using Machine Learning through an algorithm in Python language. The tracking device located in the vehicles integrates a GPS-*LoRa* module card with an IMU inertial measurement unit, to obtain inertial and position data from predefined routes. To predict locations (when necessary due to lack of data), an algorithm was implemented to train and store a prediction model with data previously collected on three routes in the city of Popayán, Colombia. The prediction model consists of two parts, one part to predict future locations, and another part to predict the time of those future locations. With the trained model a set of tests was performed with a prediction algorithm that makes use of the previously saved model. The tests were performed in two scenarios, the first occurs when there is no *LoRa* connection and the second occurs when there is no GPS connection. The results of the system implementation were satisfactory, obtaining an average coverage of 60.4% of the routes in the final field tests, through *LoRa* communication. For the remaining 39.6% of coverage, it was necessary to use location data prediction, with an average accuracy of 177 meters with respect to the actual location.



Resumen

El transporte público en una ciudad debería ser el medio más eficiente, menos contaminante y más accesible para los ciudadanos. Sin embargo, fueron detectados graves falencias en este medio de transporte, principalmente en ciudades intermedias de países en desarrollo [1]. Estas falencias están relacionadas con la falta de información al usuario (sobre rutas, horarios, puntos de parada), problemas de seguridad vial (por exceso de velocidad de los vehículos, o incumplimiento de las normas de tránsito) [2], baja disponibilidad del servicio (incumplimiento de rutas y horarios), paradas continuas en lugares inapropiados y/o no permitidos. Algunos de estos inconvenientes contribuyen a los principales problemas de movilidad en las ciudades que son la alta accidentalidad y congestión vehicular [1]. El desarrollo de herramientas que permitan mejorar las características y condiciones del servicio de transporte público en las ciudades se ha convertido en una necesidad imperativa para mejorar la calidad de vida de los ciudadanos. El control y seguimiento del servicio de transporte público es relevante en aspectos como la información en línea de su ubicación a los viajeros, y el control de las empresas de transporte para el cumplimiento de límites de velocidad, horarios, rutas y paradas [1].

Este trabajo propone un sistema de seguimiento de vehículo público mediante la integración de un sistema de posicionamiento global (GPS) y la tecnología LoRa (Long Range, o largo alcance) haciendo uso de su protocolo de red LoRaWAN, ofreciendo además predicción de ubicación en tiempo real en caso de ausencia de datos de posición. La predicción fue implementada haciendo uso de *Machine Learning* mediante un algoritmo en lenguaje Python. El dispositivo de seguimiento ubicado en los vehículos, integra una tarjeta con módulo GPS-LoRa con una unidad de medida inercial IMU, con el objetivo de obtener datos inerciales y de posición de rutas predefinidas. Para predecir ubicaciones (cuando es necesario por falta de datos) fue implementado un algoritmo para entrenar y guardar un modelo de predicción con datos recolectados previamente sobre tres rutas en la ciudad de Popayán, Colombia. El modelo de predicción consta de dos partes, una parte para predecir las futuras ubicaciones, y otra parte para predecir el tiempo de esas futuras ubicaciones. Con el modelo entrenado fue realizado un conjunto de pruebas con un algoritmo de predicción que hace uso del modelo previamente guardado. Las pruebas fueron realizadas en dos escenarios, el primero ocurre cuando no existe conexión LoRa y el segundo sucede cuando no existe conexión GPS. Los resultados de la implementación del sistema fueron satisfactorios, obteniendo una cobertura promedio del 60.4% de las rutas en las pruebas de campo finales, a través de la comunicación LoRa. Para el 39.6 % restante de cobertura, fue necesario utilizar la predicción de datos de ubicación, con una precisión promedio de 177 metros respecto a la ubicación real.



Tabla de Contenido

1.	Introducción	1
1.1.	Planteamiento del problema	1
1.2.	Objetivos	3
1.2.1.	Objetivo General.....	3
1.2.2.	Objetivos Específicos	4
1.3.	Contribuciones.....	4
1.4.	Metodología.....	4
1.5.	Artículo generado y sometido a revisión	5
1.6.	Estructura del documento	6
2.	Marco teórico.....	7
2.1.	Seguimiento de vehículo público	7
2.2.	Tecnologías LPWAN	7
2.2.1.	<i>Long Range (LoRa)</i>	7
2.2.2.	<i>LoRaWAN</i>	8
2.2.3.	LTE-M	9
2.2.4.	Sigfox	10
2.2.5.	NB-IoT (<i>Narrow-band Internet of Things</i>).....	10
2.3.	<i>Machine Learning</i> para el seguimiento de vehículos	11
2.3.1.	Aprendizaje supervisado	11
2.3.2.	Aprendizaje no supervisado.....	11
2.3.3.	Áreas de aplicación de ML	11
2.4.	Redes neuronales	12
2.4.1.	Redes Neuronales (<i>Neural Networks</i> o NN)	12
2.4.2.	Red de Gran Memoria a Corto Plazo (<i>Long Short Term Memory</i> o LSTM)	13
3.	Estado del arte	14
3.1.	Revisión de la literatura	14
3.1.1.	Fase de alcance de la investigación	14
3.1.2.	Fase de criterios de inclusión/exclusión.....	14
3.1.3.	Fase de búsqueda de la información	15
3.1.4.	Fase de esquema de clasificación.....	15
3.1.5.	Fase de resultados	15
3.2.	Documentos relacionados	17



3.3.	Resumen de evaluación de los documentos relacionados.....	20
3.4.	Revisión del estado del arte.....	21
3.4.1.	Identificación de dispositivos, sensores y variables.	22
3.4.2.	Análisis de dispositivos de recolección de datos.....	29
3.4.3.	Análisis de sensores.....	31
3.4.4.	Análisis de variables.....	32
3.4.5.	Análisis de alternativas de modelos de ML.....	33
3.5.	Selección de dispositivos, variables y sensores.....	38
4.	Selección de modelo de ML para predicción de la ubicación y pruebas iniciales .	40
4.1.	Selección del modelo de ML.....	40
4.2.	Pruebas iniciales de los modelos de ML.....	41
5.	Diseño y Desarrollo del prototipo del sistema	47
5.1.	Diseño del prototipo del sistema	47
5.1.1.	Módulo de recolección de datos	47
5.1.2.	Módulo receptor de datos	48
5.1.3.	Módulo de procesamiento de la información.....	48
5.2.	Diseño de la estructura del conjunto de datos.	50
5.2.1.	Tabla de datos del dispositivo a bordo.....	50
5.2.2.	Tabla general de datos	51
5.3.	Desarrollo e implementación del prototipo del sistema	52
5.3.1.	Desarrollo del módulo de recolección de datos.....	52
5.3.2.	Desarrollo del módulo receptor de datos	56
5.3.3.	Desarrollo del módulo de procesamiento de la información.....	57
5.4.	Pruebas y ajustes del sistema	64
6.	Diseño y ejecución de pruebas de campo y construcción del conjunto de datos .	67
6.1.	Diseño de las pruebas de campo.....	67
6.1.1.	Definición de rutas y elementos a utilizar.....	67
6.1.2.	Diseño de pruebas controladas	75
6.1.3.	Diseño de pruebas no controladas.	77
6.2.	Ejecución de las pruebas.....	78
6.2.1.	Resultados de pruebas controladas.....	78
6.2.2.	Resultados de pruebas controladas representados en mapas.....	79
6.2.3.	Pruebas no controladas.....	88
6.2.4.	Resultados de pruebas no controladas representados en mapas.....	89
6.3.	<i>Datasets</i> generados.....	94
7.	Análisis de resultados obtenidos.....	96
7.1.	Análisis de datos	96
7.1.1.	Análisis de datos pruebas controladas	96



7.1.2.	Análisis de datos de las pruebas no controladas	96
7.1.3.	Análisis general de resultados	97
7.2.	Evaluación de los resultados finales	98
8.	Conclusiones y trabajos futuros	99
8.1.	Conclusiones	99
8.2.	Trabajos futuros.....	100
	Bibliografía.....	101
	Anexo A. Artículo de investigación desarrollado.	108
	Anexo B. Pruebas iniciales de los algoritmos.....	109
	Anexo C. Algoritmos generados de las pruebas iniciales.	110
	Anexo D. Algoritmos generados en las pruebas controladas y no controladas.	111
	Anexo E. Modelos de <i>Machine Learning</i> usados en las pruebas controladas y no controladas.	112
	Anexo F. Aplicación Web desarrollada para el seguimiento de las pruebas realizadas.	113
	Anexo G. Códigos de la aplicación Web desarrollada para el seguimiento de las pruebas realizadas.	114
	Anexo H. <i>Dataset</i> generado en las pruebas controladas y no controladas.....	115



Lista de figuras

Figura 1. Representación gráfica de los paquetes de trabajo.....	5
Figura 2. Elementos de una red neuronal. Fuente: [43].....	12
Figura 3. Mapa de burbujas del mapeo sistemático.	16
Figura 4. Diagrama esquemático del bloque de memoria LSTM [68].	41
Figura 5. Predicción de latitud a partir de datos históricos para una ruta en Rio de Janeiro.....	42
Figura 6. Predicción de longitud a partir de datos históricos para una ruta en Río de Janeiro.....	42
Figura 7. Predicción de latitud en tiempo real, a partir de datos predichos para una ruta en Rio de Janeiro.....	43
Figura 8. Predicción de longitud en tiempo real, a partir de datos predichos para una ruta en Rio de 43	43
Figura 9. Predicción del tiempo a partir de datos históricos para una ruta en Rio de Janeiro.....	43
Figura 10. Predicción de latitud a partir de datos históricos para una ruta en Popayán.	44
Figura 11. Predicción de longitud a partir de datos históricos para una ruta en Popayán.	44
Figura 12. Predicción del tiempo a partir de datos históricos para una ruta en Popayán.	44
Figura 13. Predicción de latitud a partir de datos predichos para una ruta en Popayán (en el escenario 1).	45
Figura 14. Predicción de longitud a partir de datos predichos para una ruta en Popayán (en el escenario 1).	45
Figura 15. Predicción del tiempo a partir de datos predichos para una ruta en Popayán, (en el escenario 1). (Ampliada)	46
Figura 16. Diseño del prototipo propuesto para el sistema.....	47
Figura 17. Diagrama de bloques del dispositivo para recolectar los datos.	48
Figura 18. Escenario 1. Sistema sin comunicación LoRa, ausencia de paquetes en el Gateway (es indiferente si hay o no conexión GPS).	49
Figura 19. Escenario 2. Sistema sin señal GPS y comunicación LoRa activa.	49
Figura 20. Dispositivo recolector de datos (Cube Cell HTCC-AB02S, MPU-6050).	54
Figura 21. Diagrama esquemático del dispositivo recolector de datos (Cube Cell HTCC-AB02S, MPU-6050).	55
Figura 22. Dispositivo recolector de datos (ESP32 LoRa, GPS Ublox 6M, MPU-6050).	55
Figura 23. Diagrama esquemático del dispositivo recolector de datos (ESP32 LoRa, GPS Ublox 6M, MPU-6050).	56
Figura 24. Gateway LoRaWAN LPS8. Fuente:[77].....	56
Figura 25. Arquitectura del algoritmo del sistema general de predicción.	59
Figura 26. Diagrama de flujo del algoritmo de entrenamiento.	60
Figura 27. Diagrama de flujo algoritmo de predicción.	61
Figura 28. Recepción y almacenamiento de los datos provenientes de la TTN.....	62
Figura 29. Interfaz de inicio de la aplicación web.	62
Figura 30. Interfaz de la ruta 3.	63
Figura 31. Interfaz del vehículo 3.	63
Figura 32. Interfaz de las ultimas 100 ubicaciones realizadas por el vehículo 2 en el mapa.	64
Figura 33. Interfaz de los registros obtenidos del vehículo 2.....	64



Figura 34. Adaptador USB para vehículo.....	68
Figura 35. Mapa de la ruta 1 y ubicaciones de los 2 <i>Gateways</i> (mapa tomado de <i>Google Maps</i>).	69
Figura 36. Mapa de la ruta 2 y ubicaciones de los 2 <i>Gateways</i> (mapa tomado de <i>Google Maps</i>).	70
Figura 37. Mapa de la ruta 3 y ubicaciones de los 2 <i>Gateways</i> (mapa tomado de <i>Google Maps</i>).	71
Figura 38. Mapa del recorrido del vehículo 1 y ubicaciones de los 2 <i>Gateways</i> (mapa tomado de <i>Google Maps</i>).	72
Figura 39. Mapa del recorrido del vehículo 2 y ubicaciones de los 2 <i>Gateways</i> (mapa tomado de <i>Google Maps</i>).	73
Figura 40. Mapa del recorrido del vehículo 3 y ubicaciones de los 2 <i>Gateways</i> (mapa tomado de <i>Google Maps</i>).	74
Figura 41. Mapa del recorrido del vehículo 4 y ubicaciones de los 2 <i>Gateways</i> (mapa tomado de <i>Google Maps</i>).	75
Figura 42. Registros en el mapa del escenario 1 para ruta 1 (mapa tomado de <i>Google Maps</i>).	80
Figura 43. Registros en el mapa del escenario 1 para ruta 2 (mapa tomado de <i>Google Maps</i>).	82
Figura 44. Registros en el mapa del escenario 1 para ruta 3 (mapa tomado de <i>Google Maps</i>).	83
Figura 45. Registros en el mapa del escenario 2 para ruta 1 (mapa tomado de <i>Google Maps</i>).	85
Figura 46. Registros en el mapa del escenario 2 para ruta 2 (mapa tomado de <i>Google Maps</i>).	86
Figura 47. Registros en el mapa del escenario 2 para ruta 3 (mapa tomado de <i>Google Maps</i>).	87
Figura 48. Registros en el mapa del recorrido del vehículo 1 (mapa tomado de <i>Google Maps</i>).	90
Figura 49. Registros en el mapa del recorrido del vehículo 2 (mapa tomado de <i>Google Maps</i>).	90
Figura 50. Registros en el mapa del recorrido del vehículo 3 (mapa tomado de <i>Google Maps</i>).	92
Figura 51. Registros en el mapa del recorrido del vehículo 4 (mapa tomado de <i>Google Maps</i>).	93



Lista de tablas

Tabla 1. Resumen de evaluación de los trabajos relacionados.	21
Tabla 2. Hardware utilizado en recolección de datos de vehículos.	22
Tabla 3. Rendimiento de los modelos de ML.	35
Tabla 4. Dispositivos seleccionados.	39
Tabla 5. Datos enviados por el dispositivo a bordo.	50
Tabla 6. Tabla general de datos.	51
Tabla 7. Características de los vehículos.	67
Tabla 8. Ubicación geográfica de los Gateway.	68
Tabla 9. Lista de recorridos planeados para la recolección de datos para el entrenamiento.	76
Tabla 10. Lista de recorridos planeados para las pruebas controladas para el escenario 1.	76
Tabla 11. Lista de recorridos planeados para las pruebas controladas para el escenario 2.	76
Tabla 12. Lista de recorridos planeados para las pruebas no controladas.	78
Tabla 13. Resultados obtenidos de las pruebas controladas ejecutadas para el algoritmo de entrenamiento.	78
Tabla 14. Resultados obtenidos de pruebas controladas para el escenario 1.	79
Tabla 15. Resultados obtenidos de pruebas controladas para el escenario 2.	79
Tabla 16. Datos predichos en las pruebas controladas en la ruta 1 escenario 1.	80
Tabla 17. Datos predichos en las pruebas controladas en la ruta 2 escenario 1.	82
Tabla 18. Datos predichos en las pruebas controladas en la ruta 3 escenario 1.	83
Tabla 19. Datos predichos en las pruebas controladas en la ruta 1 escenario 2.	86
Tabla 20. Datos predichos en las pruebas controladas en la ruta 2 escenario 2.	87
Tabla 21. Datos predichos en las pruebas controladas para la ruta 3 escenario 2.	88
Tabla 22. Resultados obtenidos de las pruebas no controladas.	88
Tabla 23. Datos predichos en las pruebas no controladas ruta 1.	91
Tabla 24. Datos predichos en las pruebas no controladas ruta 2.	92
Tabla 25. Datos predichos en las pruebas no controladas en la ruta 3.	93



Lista de acrónimos

ABP	<i>Activated By Personalisation</i> (Activación personalizada).
ACC	Aceleración.
AES	<i>Advanced Encryption Standard</i> (Estándar de encriptación avanzada).
ANN	<i>Artificial Neural network</i> (Redes neuronales artificiales).
ARIMA	<i>Autoregressive Integrated Moving Average</i> (Modelo autorregresivo integrado de promedio móvil).
ARNN	<i>Attention-based Recurrent Neural Network</i> (Red neuronal recurrente basada en la atención).
B5G	<i>Before 5G</i> (Mas allá del 5G).
BW	<i>BandWidth</i> (Ancho de banda).
CNN	<i>Convolutional Neural Networks</i> (Redes neuronales convolucionales).
CR	<i>Code rate</i> (Tasa de código).
CSS	<i>Chirp Spread Spectrum</i> (chirp de espectro ensanchado).
DBLSTM	<i>Deep Bidirectional Long Short-Term Memory</i> (redes LSTM bidireccionales).
EKF	<i>Extended Kalman Filter</i> (filtro extendido de Kalman).
GAN	<i>Generative Adversarial Network</i> (red adversarial generativa).
GPIO	<i>General Purpose Input/Output</i> (Entrada y salida de propósito general).
GPRS	<i>General Packet Radio Service</i> (Servicio general de paquetes de radio).
GPS	<i>Global Positioning System</i> (sistema de posicionamiento global).
GSM	<i>Global System for Mobile communications</i> (sistema global para las comunicaciones).
GYRO	Velocidad angular.
HTTPS	<i>Hypertext Transfer Protocol Secure</i> (Protocolo de transferencia de hipertexto seguro).
IMU	<i>Inercial measurement unit</i> (Unidad de medida inercial).
INS	<i>Inercial Navigation System</i> (Sistema de navegación inercial).
IoT	<i>Internet of Things</i> (Internet de las cosas).
IP	<i>Internet protocol</i> (Protocolo de Internet).
JSON	<i>JavaScript Object Notation</i> (Notación de objetos javascript).
KF	<i>Kalman Filter</i> (Filtro de Kalman).
LoRa	<i>Long Range</i> .
LoRaWAN	<i>Protocol Wide Area Network of LoRa</i> .
LoS	<i>Line of Sight</i> (Línea de vista).
LPWAN	<i>Low Power Wide Area Networks</i> (redes de área amplia de baja potencia).
LSTM	<i>Long Short-Term Memory</i> (Gran memoria de corto plazo).
EBV	Equipo a bordo del vehículo.
LTE	<i>Long Term Evolution</i> (Evolución a largo plazo).
MAC	<i>Media Access Control</i> (control de acceso al medio).
ML	<i>Machine Learning</i> (Aprendizaje de máquina).
MLP	<i>Multi-Layer Perceptron</i> (Perceptron multi nivel).
MQTT	<i>Message Queuing Telemetry Transport</i> (Transporte de telemetría MQ).
MSE	<i>Mean squared error</i> (Error cuadrático medio).
NN	<i>Neural Networks</i> (Redes Neuronales).
PAN	<i>Person Area Network</i> (Red de área personal).



PMBOK	<i>Project Management Body of Knowledge</i> (Cuerpo de conocimientos de gestión de proyectos).
PMI	<i>Project Management Institute</i> (Instituto de gestión de proyectos).
RNN	<i>Recurrent Neural Network</i> (Red neuronal recurrente).
RSGs	<i>Road Segments</i> (Segmentos de carretera).
SF	<i>Spread Factor</i> (Factor de propagación).
SQL	<i>Structure Query Language</i> (Lenguaje de consulta estructurado).
OTAA	<i>Over-The-Air Activation</i> (método de autenticación LoRaWAN de activación por aire).
TIC	Tecnologías de la información y las comunicaciones.
TTN	<i>The Things Network</i> (Plataforma para el desarrollo de IoT).
UNB	<i>Ultra Narrow Band</i> (Ultra banda estrecha).
UART	<i>Universal Asynchronous Receiver Transmitter</i> (Transmisor receptor universal asíncrono).
URL	<i>Uniform Resource Locator</i> (Localizador de recursos uniforme).
WP	<i>Work Package</i> (Paquete de trabajo).



Capítulo 1

1. Introducción

En este capítulo es presentado el problema planteado, el objetivo general y los específicos, y la metodología usada en el desarrollo del presente trabajo. Además, esta sección presenta la información de un artículo de revisión sometido a una revista relacionada con el área de transporte. Al finalizar, la sección presenta la estructura del resto del documento de la monografía.

1.1. Planteamiento del problema

Los principales problemas de movilidad a nivel mundial son la congestión de tráfico y la alta tasa de accidentalidad [3]. Según los datos suministrados por INRIX respecto a la congestión de tráfico [3], en el 2020 se presentó una reducción considerable del nivel de tráfico, debido principalmente a la pandemia de COVID-19, la cual redujo la movilidad en las vías de todo el mundo. Sin embargo, la gran mayoría de naciones en el mundo están llegando a mediados de 2021 a los niveles de congestión que tenían prepandemia, inclusive, valores superiores [4]. Naciones como el Reino Unido tuvieron una disminución de tráfico del 68% entre 2020 y 2019, pasando de 115 horas en el tráfico en 2019 a 37 horas en el año 2020, a su vez durante este mismo periodo, Estados Unidos presentó una baja del 74%, pasando de 99 a 26 horas. Estos países están llegando a niveles entre el 93% al 107% de los vehículos que transitaban prepandemia [4].

Colombia no escapa de esta problemática, donde los conductores de la ciudad capital Bogotá en el año 2020 perdieron 113 horas al año por la congestión vehicular, frente a las 191 horas del 2019. Pese a la reducción, Bogotá mantuvo el primer lugar como la ciudad donde los conductores pierden más horas en el tráfico en el mundo, adicionalmente los niveles de congestión en 2021 están regresando a los niveles prepandemia [4].

Respecto a los accidentes de tránsito, según el informe anual de medicina legal sobre lesiones de causa externa en Colombia [2], en el 2019 se presentaron 6.892 muertes y 36.832 lesiones por eventos relacionados con transporte. El 18,8% de las lesiones son resultado del exceso de velocidad y el 22,2 % son resultado de desobedecer las señales de tránsito; siendo estas dos las principales causas de accidentes de tránsito en el país. Otro dato relevante respecto a los accidentes de tránsito es que el 3,3% de las muertes y el 8,9% de las lesiones están relacionadas con el transporte público [2]. Los problemas de accidentalidad en el país se dan tanto en las ciudades principales, como en las ciudades intermedias del país, como por ejemplo en la ciudad de Popayán, en 2019, ocurrieron 53 muertes y 376 lesiones vinculadas con accidentes de tránsito, donde el 40,16% y el 41,16% de los datos mencionados respectivamente fueron causados por sobrepasar el límite de velocidad e incumplir las normas de tránsito.

Analizando las principales razones de los accidentes de tránsito en el país (exceso de velocidad y desobediencia de las señales de tránsito), un adecuado seguimiento, control de velocidad y el cumplimiento de las normas de tránsito en los vehículos, permitiría una disminución de la accidentalidad en las ciudades del país. Dicho control es difícil realizarlo en todos los vehículos que transitan en una ciudad. Sin embargo, un grupo representativo como los vehículos de transporte público, que generan el 25% de los



accidentes de tránsito [2], puede ser una opción válida para introducir dicho control y evaluar una posible mejora en la reducción de estos accidentes. Adicionalmente, un adecuado control y seguimiento de los vehículos de transporte público en una ciudad puede incentivar su uso y, de esa forma, contribuir a la disminución del tráfico. Dicho control y seguimiento del transporte público es más significativo, en cuanto a beneficios, para ciudades intermedias de un país en desarrollo como Colombia. En donde se utiliza principalmente el tipo de vehículo colectivo en el transporte público (el cual comparte las vías con los demás tipos de vehículo y no tiene comúnmente un adecuado control y seguimiento). Una ciudad grande de un país en desarrollo comúnmente posee sistemas de transporte masivo (por ejemplo: MIO en Cali, Transmilenio en Bogotá o Metro en Medellín), en los cuales los vehículos tienen vías exclusivas y un adecuado control y seguimiento.

Para realizar dicho seguimiento y control (de velocidad y cumplimiento de normas), con el fin de reducir el número de accidentes de tránsito y reducir los niveles de congestión, es viable usar las tecnologías de la información y las comunicaciones (TIC) como un agregado a las estrategias de mejoramiento de la seguridad vial, como lo son los planes de concientización de buenas conductas en las vías, y la gestión de señales de tránsito [5], [6].

Las tecnologías denominadas redes de área amplia de baja potencia (*Low Power Wide Area Networks* o LPWAN) son caracterizadas por un bajo consumo de energía, largo alcance y transmisión de pequeños datos a larga distancia [7]. Entre estas tecnologías está la denominada "Largo Alcance" (*Long Range* o *LoRa*). *LoRa*, es una técnica de modulación en radio frecuencia derivada de chirp de espectro ensanchado (*chirp spread spectrum* o CSS), el cual le permite alcanzar grandes distancias y servir como soporte para el envío y recepción de la información en escenarios de tipo IoT. Esta tecnología (*LoRa*) es comúnmente utilizada con su protocolo de red (*LoRa* en redes de área amplia o *LoRaWAN*). El uso de este protocolo permite mejorar las características de transmisión y recepción cuando existen múltiples dispositivos operando al mismo tiempo en un área reducida.

LoRaWAN es un protocolo de capa de control de acceso al medio (*Media Access Control* o MAC). Su función es servir como protocolo de red que administra la comunicación entre las puertas de enlace (*Gateways*) y dispositivos finales [8], [9]. Algunos ejemplos del uso de estas tecnologías (*LoRa* y *LoRaWAN*) en sistemas de seguimiento vehicular, puede evidenciarse en [10], [11], [12], [13].

En la actualidad existen diferentes formas de efectuar el seguimiento del servicio de transporte público, basados en un sistema de posicionamiento global (*Global Positioning System* o GPS), *Bluetooth* y/o tecnologías de comunicaciones utilizadas en redes celulares como el sistema global para las comunicaciones (*Global System for Mobile communications* o GSM) y evolución a largo plazo (*Long Term Evolution* o LTE) pero se conocen algunos inconvenientes respecto a costos y algunos aspectos técnicos como alcance y cobertura [12], los cuales pueden ser solucionados con tecnologías LPWAN como *LoRa*.

Las principales ventajas que presenta *LoRa* son: un largo alcance de hasta 48.27 km, baja potencia de alimentación extendiendo hasta 10 años la vida útil de la batería de los dispositivos que operan como nodos, ofrece estandarización LPWAN con la especificación del protocolo *LoRaWAN* con un programa de certificación para garantizar interoperabilidad en múltiples redes, permite el desarrollo de aplicaciones de



localización sin GPS, bajo consumo de energía entre dispositivos en movimiento, permite una gran cantidad de mensajes por *Gateway* y un bajo costo de operación e infraestructura [8]. *LoRaWAN* ofrece las siguientes características: topología en estrella, velocidades de 0.3 - 50 kbps, alta seguridad al usar cifrado de extremo a extremo AES128, comunicación bidireccional, diferentes bandas de frecuencia de acuerdo con la zona geográfica y soporta tres clases de nodos (Clase A, B y C, que se clasifican de esta forma considerando los procesos de conexión, envío y recepción de información, desde y hacia los *Gateways*) o dispositivos según la aplicación y requisitos [14], [15]. Las desventajas que presenta *LoRa* y su protocolo de comunicación, además de la necesidad de una línea de vista (*Line of Sight* o LoS), entre los equipos *Gateway* y los dispositivos terminales, son: las bajas velocidades y el reducido tamaño de paquetes (255 bytes) [16].

La necesidad de una LoS constante entre *Gateways* y dispositivos terminales al utilizar esta tecnología de comunicación en sistemas de seguimiento de vehículos, genera inconvenientes de cobertura. Por lo anterior, es necesario implementar algún tipo de solución que permita mejorar los inconvenientes sobre localización del vehículo, cuando no existe LoS.

La predicción de localización mediante el aprendizaje máquina (*Machine Learning* o ML) es una opción aplicada en algunos trabajos de seguimiento de vehículos ([17], [18], [19], [20], [21], [22], [23], [24]). En dichos trabajos, se pretende predecir la ubicación futura de los vehículos a través de la recolección de datos como la latitud, longitud y paradas realizadas por estos. El conjunto de datos recolectado en estos trabajos es procesado por diferentes modelos de ML (como *Long Short-Term Memory Neural Networks*, *Convolutional Neural Networks*, *Recurrent Neural Network*, *Artificial Neural Network* y *Generative Adversarial Network*) para realizar la predicción de ubicación.

Teniendo en cuenta lo anterior, para mejorar las características de los sistemas de seguimiento del servicio de transporte público e intentar aportar en la reducción de accidentalidad y congestión vehicular, es propuesta la siguiente pregunta de investigación: ¿Cómo mejorar las características técnicas y operativas de los actuales sistemas de seguimiento del servicio de transporte público en el país?

La hipótesis propuesta para dar solución a la pregunta de investigación es la siguiente: Es posible mejorar las características de los actuales sistemas de seguimiento del servicio de transporte público colectivo, mediante el desarrollo de un prototipo de un sistema para el seguimiento del servicio en tiempo real, basado en aprendizaje de máquina y *LoRaWAN*.

1.2. Objetivos

1.2.1. Objetivo General

Proponer un sistema inteligente de seguimiento del servicio de transporte público colectivo en tiempo real que utilice aprendizaje de máquina para la predicción de la ubicación en caso de ausencia de datos por parte del sistema de comunicación principal *LoRa* (y su protocolo de red *LoRaWAN*).



1.2.2. Objetivos Específicos

- Identificar los parámetros y variables relevantes de la tecnología *LoRa* y su protocolo de red *LoRaWAN*, para permitir un servicio escalable de seguimiento de servicio de transporte público colectivo, considerando aspectos como la seguridad y la eficiencia.
- Seleccionar el algoritmo de aprendizaje de máquina que permita realizar la predicción de ubicación de los vehículos de servicio de transporte público colectivo, cuando no hay datos del sistema de comunicación.
- Desarrollar el prototipo del sistema de seguimiento del servicio de transporte público colectivo (*hardware, software* y comunicaciones).
- Evaluar la eficiencia y eficacia del prototipo desarrollado en un caso de estudio, siguiendo el comportamiento de rutas de empresas de servicio de transporte público.

1.3. Contribuciones

El trabajo propuesto fortalecerá la línea de investigación del Grupo de Ingeniería Telemática relacionada con las aplicaciones y servicios sobre Internet. Concretamente los aportes más relevantes que tendrá el trabajo son:

- El diseño e implementación del prototipo del sistema de seguimiento del servicio de transporte público colectivo.
- La construcción de un módulo de predicción, para la ubicación de un servicio de transporte público colectivo utilizando algoritmos (previamente desarrollados y adecuadamente entrenados) de aprendizaje de máquina. Esta predicción es requerida, en los puntos en que no opere adecuadamente la comunicación mediante *LoRaWAN*, debido a inconvenientes de la LoS (por demasiados obstáculos) o a una distancia demasiado grande entre dispositivo y la puerta de enlace (*Gateway*).
- La evaluación de la tecnología *LoRa* y su protocolo de red *LoRaWAN*, en el prototipo de seguimiento implementado. Dicha evaluación debe realizarse mediante una cantidad de pruebas suficientes y adecuadamente diseñadas.
- Un conjunto de datos construido con la información recolectada en las pruebas de campo que se diseñen para el prototipo del sistema.

1.4. Metodología

El desarrollo de los objetivos propuestos fue realizado a través de una serie de actividades según las directrices descritas en el cuerpo de conocimientos de gestión de proyectos (*Project Management Body of Knowledge* o PMBOK), instrumento desarrollado por el instituto de gestión de proyectos (*Project Management Institute - PMI*) [25]. El proyecto consta de los siguientes paquetes de trabajo (*Work Package* o *WP*): Identificación de parámetros relevantes y dispositivos a considerar; desarrollo del prototipo del sistema; pruebas de campo y construcción del conjunto de datos; análisis y evaluación de resultados; preparación de documentación y otros entregables.

La **Figura 1** representa el diagrama de paquetes de trabajo desarrollados en el proyecto.

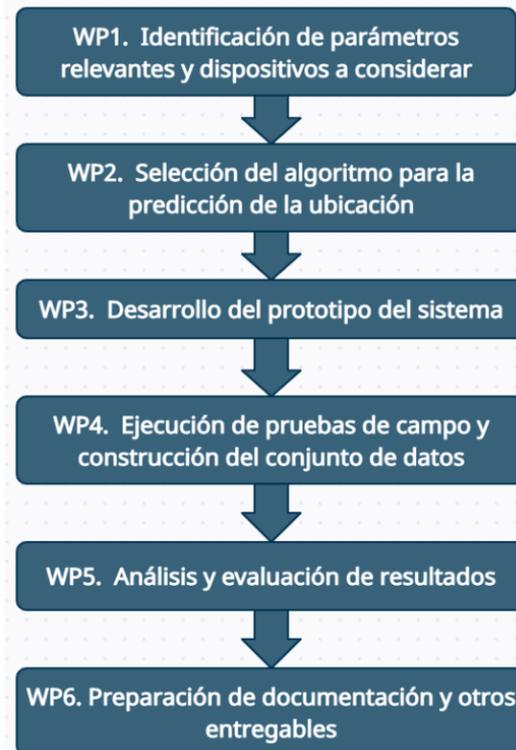


Figura 1. Representación gráfica de los paquetes de trabajo.

1.5. Artículo generado y sometido a revisión

El trabajo desarrollado en este documento fue presentado a la comunidad científica mediante el envío de un artículo a la revista: *Sustainability* (<https://www.mdpi.com/journal/sustainability>). *Sustainability* es una revista internacional, interdisciplinaria, académica, revisada por pares y de acceso abierto sobre la sostenibilidad ambiental, cultural, económica y social de los seres humanos.

ISSN de la revista: 2071-1050

Mejor cuartil en el SJR; Q1

Clasificación en Colciencias: A1 para 2023.

Special Issue de la revista al que fue enviado: *Smart Urban and IoT: Advances, Opportunities and Challenges*

(https://www.mdpi.com/journal/sustainability/special_issues/AC244X9X72)

Autores del artículo: Andersson Stive Diaz Lucio, Camilo Andrés Sánchez Diaz, Ricardo Salazar-Cabrera, Mary Cristina Carrascal, Álvaro Pachón de la Cruz, Juan Manuel Madrid Molina.

Título: Prototype of a smart system for monitoring transit service in real time based on *Machine Learning* and *LoRaWAN*

Estado: Enviado a evaluación.

Texto del artículo: Disponible en el **Anexo A. Artículo de investigación desarrollado.**



1.6. Estructura del documento

A continuación, es descrita la estructura de las demás secciones del presente documento:

- Capítulo 1. Introducción, contiene el planteamiento del problema, los objetivos, las contribuciones, la metodología y la publicación.
- Capítulo 2. Marco teórico, describe los conceptos y tecnologías utilizados en el desarrollo del trabajo.
- Capítulo 3. Estado del arte, presenta la revisión de la literatura mediante el mapeo sistemático, los trabajos relacionados con el presente trabajo, la identificación de variables y dispositivos relevantes. Describe las alternativas y selección de los dispositivos para la recolección de datos, las variables a medir y la identificación de modelos de ML.
- Capítulo 4. Selección de algoritmo para predicción de la ubicación. Este capítulo describe el proceso de selección del algoritmo a utilizar y las pruebas iniciales con el algoritmo.
- Capítulo 5. Desarrollo del prototipo del sistema, especifica cada uno de los componentes funcionales que conforman el sistema desde el dispositivo de recolección de datos hasta el algoritmo de predicción.
- Capítulo 6. Ejecución de pruebas de campo y construcción del conjunto de datos, presenta el diseño y ejecución de las pruebas de campo realizadas y los resultados obtenidos al ejecutarlas.
- Capítulo 7. Análisis de resultados obtenidos. El capítulo presenta el análisis de los resultados obtenidos en la ejecución de las pruebas.
- Capítulo 8. Conclusiones, presenta las conclusiones y aportes del presente documento, además de los trabajos futuros.



Capítulo 2

2. Marco teórico

En este capítulo son presentados los fundamentos teóricos adoptados para el desarrollo del sistema propuesto. Primero, es descrita la forma en la que es posible realizar el seguimiento de vehículo público, luego, son presentadas las distintas tecnologías LPWAN que pueden ser usadas para transmitir la información del receptor GPS, después es abordado el concepto de ML en el contexto del seguimiento de vehículo público como una herramienta para mejorar la eficiencia en los sistemas de rastreo cuando ocurre una interrupción de la señal GPS. Por último, es descrito el concepto de redes neuronales como parte del ML.

2.1. Seguimiento de vehículo público

El seguimiento vehicular nace con la aparición de la tecnología del GPS y su posterior adopción por parte del público, donde un dispositivo a bordo del vehículo recibe información por medio de satélites para después transmitir esa información hacia un centro de monitoreo mediante comunicaciones que pueden ser satelitales, celulares o de radio. Una vez la posición del vehículo es recibida, es desplegada en una plataforma de visualización donde es posible entre otras cosas, gestionar flotas de vehículos públicos, rastrear mercancías, realizar espionaje, brindar servicios de rastreo en caso de robo o recolectar datos para analizarlos y mejorar la eficiencia de las operaciones realizadas. Aunque este seguimiento generalmente es realizado de forma activa, en línea, también es posible hacerlo pasivamente, mediante un dispositivo con memoria.

Como fue mencionado, el seguimiento de un vehículo es realizado en general de manera pasiva a través de dispositivos con memoria o de manera activa donde el conocimiento de la posición es realizado en tiempo real al hacer uso de la telefonía celular, más precisamente al usar la red GSM / Servicio General de Paquetes vía Radio (*General Packet Radio Service* o GPRS), la cual permite brindar un servicio integral y rentable para los usuarios como lo mencionan en [22]. Aunque en los últimos años la telefonía celular (2G/3G/4G) ha sido el puente de comunicación entre vehículos y los centros de gestión, al haber un crecimiento en el IoT (Internet of Things), surgen nuevas posibilidades de nuevas tecnologías de comunicaciones [10].

2.2. Tecnologías LPWAN

Las redes LPWAN (*Low Power Wide Area Network*) son un conjunto de tecnologías de comunicación inalámbrica que permiten transmitir datos entre un dispositivo y una estación base separados por kilómetros de distancia con un consumo energético bajo, generalmente con una topología de red en estrella. La estación base o *Gateway* sirve como pasarela al mundo IP y su arquitectura permite una cobertura de área amplia.

2.2.1. Long Range (LoRa)

Es un tipo de comunicación inalámbrica que trabaja en el espectro no licenciado de la banda de radio industrial, científica y médica (ISM) y basa su funcionamiento en la modulación de CSS. Es una tecnología usada en el campo de las comunicaciones militares y en plataformas del Internet de las Cosas (*Internet of Things* o IoT) debido a



las grandes distancias alcanzadas en kilómetros y la robustez frente a las interferencias [8], [15].

Las principales características de *LoRa* son: Una alta inmunidad a la interferencia; opera en las frecuencias de 868 MHz para Europa, 915 MHz en Estados Unidos y 433 MHz para Asia; bajo consumo energético prolongando la vida útil de las baterías entre 8 a 10 años; un largo alcance en el orden de 10 a 20 Km o hasta 48.27 Km, distancia que depende de la línea de vista entre dispositivos y ubicación geográfica; alta sensibilidad en recepción en el orden de -168 dB; tamaño máximo que puede transmitir es de 255 bytes; bajo costo de los dispositivos; número ilimitado de mensajes diarios; la red puede ser privada o pública [26].

Los parámetros de la comunicación *LoRa* son:

- Factor de ensanchamiento o propagación. Representa el número de símbolos enviados por cada bit de información. Utiliza 6 posibles valores (7 - 12), un incremento de este factor permite un mayor alcance de comunicación, mayor duración de paquete en el aire, una mayor sensibilidad del receptor, pero reduciendo la velocidad de transmisión [27].
- Ancho de banda. Representa el rango de frecuencias usadas en la banda de transmisión. En *LoRa* a mayor ancho de banda, mayor será la velocidad de transmisión, pero con una sensibilidad más baja debido a la integración del ruido. Los anchos de banda utilizados son: 125, 250 y 500 KHz [28], [29].
- Tasa de código. Emplea una corrección de errores hacia adelante, la cual permite la corrección y detección de errores y la protección contra interferencias al agregar una pequeña sobrecarga o redundancia en los datos a transmitir. Las tasas de codificación permitidas son: 4/5, 4/6, 4/7 y 4/8. A mayor tasa de codificación, mayor protección y tiempo de duración en el aire, pero origina una disminución en la velocidad de datos [28], [30].

2.2.2. *LoRaWAN*

LoRaWAN es una especificación que define un protocolo de red de área amplia y baja potencia permitiendo conectar dispositivos que tienen un bajo consumo de alimentación. La especificación permite operar en el IoT ofreciendo servicios de comunicación bidireccional, seguridad de extremo a extremo, movilidad y localización [14].

Arquitectura de la red

La topología de red utilizada por *LoRaWAN* es en estrella en donde los *Gateway* están conectados a un servidor de red a través de una conexión IP y son el puente de comunicación bidireccional entre dispositivos finales y el servidor permitiendo la convertir paquetes de radio frecuencia en paquetes IP y viceversa [14].

Clases de nodos

Se manejan tres tipos de nodos finales clasificados en A, B y C de acuerdo con el requisito que sea necesario.



- Clase A. Este tipo de dispositivo funciona entre tres ventanas, la primera cuando es necesario enviar información hacia el servidor. Una vez confirmada la recepción de los datos, son activadas las dos ventanas restantes que permiten recibir información proveniente del servidor de red. Este tipo de nodo es el que ahorra la mayor cantidad de energía porque solo es activado cuando hay que enviar paquetes de datos el resto de tiempo permanece dormido, ideal donde la fuente de alimentación es una batería [14], [15].
- Clase B. Tiene la misma funcionalidad del nodo clase A con la diferencia que tiene ventanas de recepción adicionales las cuales son activadas en horas programadas aumentando un poco el consumo de energía, pero siendo válido para usar una batería como fuente de alimentación [14], [15].
- Clase C. Este tipo de dispositivo mantiene abierta de forma constante la ventana de recepción incrementando el consumo energía y solo la cierra cuando está transmitiendo información. Esta clase permite que el servidor de red pueda transmitir hacia el nodo en cualquier instante de tiempo asumiendo que el dispositivo está escuchando, reduciendo la latencia. Este tipo de nodo necesita de una fuente de alimentación constante [14], [15].

Tasa de transferencia de datos

Al utilizar la tecnología de espectro ensanchado como modulación, las señales a transmitir prácticamente son ortogonales entre si cuando son usados diferentes valores de factor de propagación. Al aumentar la velocidad de transmisión la duración del paquete es más corta abriendo espacio para que otros nodos transmitan y de esta forma no utilizar el espectro por más tiempo del necesario. La velocidad de datos adaptiva optimiza la batería de cada uno de los nodos. La velocidad datos en LoRaWAN varia de 0.3 a 50 Kbps [14], [15].

Seguridad

El protocolo utiliza dos capas de seguridad, una para red y otra para aplicación haciendo uso de algoritmos AES. La seguridad de la red garantiza la autenticidad del nodo en la red a través de una clave de sesión de red única de 128 bits compartida entre el dispositivo final y el servidor de red. La clave de aplicación no permite que el operador de la red tenga acceso a la información de aplicación del usuario final por medio de la clave de sesión de aplicación única de 128 bits compartida de extremo a extremo en este nivel [14], [15].

Duración de la batería

Los nodos finales pertenecientes a la red funcionan de forma asíncrona, lo que significa que solo establece la comunicación cuando tiene información por enviar, ya sea por un evento que suceda o por si han sido programados [14], [15].

2.2.3. LTE-M

Long Term Evolution for Machines o LTE-M es una tecnología de acceso celular de ancho de banda estrecha definida por el 3GPP y en realese 13, 14 y 15, permitiendo coexistir con las generaciones de telefonía celular 2G, 3G y 4G [31].



Esta tecnología maneja un ancho de banda igual a 1 Mbps, convirtiéndola en su principal ventaja al tener el mayor ancho de banda de todas las LPWAN; permite el envío de voz a través de su servicio VoLTE; transmisión de pequeñas cantidades de datos en periodos largos de tiempo; permite la movilidad, al soportar dispositivos en movimiento; duración de la batería hasta 10 años [32].

2.2.4. Sigfox

Es una tecnología LPWAN desarrollada para ser usada en IoT con la finalidad de ser independiente de la red de telefonía celular y trabaja en frecuencias libres de licencia entre 868 MHz para Europa y 915 MHz para Estados Unidos. Su funcionamiento está basado en el uso de Ultra Banda Estrecha (*Ultra Narrow Band, UNB*). Es una tecnología que transmite por un canal muy estrecho menor a 1 KHz permitiéndole alcanzar distancias de 10 Km en zonas urbanas y 40 Km en zonas rurales consumiendo baja potencia para transmitir. UNB fue diseñada con el objetivo de trabajar a bajas velocidades para la transferencia de datos (10 a 1000 bits por segundo) [33], [34].

Las principales características de la tecnología son las siguientes: el tamaño máximo que pueden transmitir los dispositivos finales es de 12 bytes a una tasa de 100 bps; permite el envío de 140 mensajes en el día, siendo una limitante para aplicaciones como el seguimiento GPS donde tan solo se transmitiría información cada 10 minutos; los nodos finales pueden recibir 4 mensajes diarios con un tamaño máximo de 8 bytes; bajo consumo de energía permitiendo extender la vida útil de las baterías; resistente a la interferencia debido a la utilización de UNB y la diversidad espacial de las estaciones base, debido a que cada mensaje emitido es recibido por múltiples antenas; topología en estrella; permite obtener un alcance de varios kilómetros dependiendo de la topografía del terreno; la infraestructura es propietaria por lo que debe pagarse una suscripción anual por dispositivo, a pesar de operar en frecuencias no licenciadas [32], [34], [35], [36].

2.2.5. NB-IoT (*Narrow-band Internet of Things*)

La banda estrecha para el internet es otro tipo de tecnología LPWAN la cual basa su funcionamiento en telefonía celular para IoT y esta especificada en los *release 13* y *14* del 3GPP permitiéndole coexistir con las redes móviles 2G, 3G y 4G. Esta tecnología también es conocida como *CaT-M2*. A pesar de que NB-IoT y LTE-M están definidas por el 3GPP su diferencia radica en que LTE-M está desarrollado para comunicaciones donde sea transmitida la voz y tiene un mayor ancho de banda. Por el contrario, NB-IoT opera en un ancho de banda menor, pero tiene una mejor penetración en interiores y bajo tierra [26], [37], [38].

Las principales características de NB-IoT son: ocupa un ancho de banda de 200 KHz; permite la conexión de 52.000 dispositivos por canal; el tamaño máximo del paquete es de 1.600 Bytes; velocidad de transferencia de datos de 250 Kbps; no hay límite de envío de mensajes diarios; la duración de la batería de los dispositivos es 10 años; opera en las frecuencias licenciadas para LTE (700 a 800 MHz); al trabajar en frecuencias licenciadas no tienen interferencia con otras tecnologías como Sigfox y LoRa; mayor penetración en interiores y bajo tierra; cobertura de 1 Km en zonas urbanas y 10 Km en



zonas rurales; la cobertura está limitada a la infraestructura desplegada de LTE (número de estaciones base en una región); equipos más costosos; y pérdidas de acuse de recibo [10], [26], [34], [37], [38].

2.3. *Machine Learning* para el seguimiento de vehículos

El ML es un campo perteneciente a la inteligencia artificial (IA) y la estadística que consiste en extraer información y patrones de los datos en bruto a través de un algoritmo o métodos con el objetivo que los sistemas aprendan de la experiencia sin la intervención de los humanos. ML permite tomar decisiones o realizar predicciones a partir de datos estadísticos, recopilados a través del tiempo, permitiendo que una máquina o proceso automatizado realice una tarea en particular [39]. Dentro de ML son utilizados diferentes métodos de aprendizaje, donde los más utilizados son el aprendizaje supervisado y no supervisado. Los cuales son descritos de forma general a continuación:

2.3.1. Aprendizaje supervisado

Este tipo de algoritmos de aprendizaje consiste en utilizar muestras de datos etiquetados que sirven de entrenamiento para el algoritmo y de esta manera lograr aprender sus características, logrando clasificar o predecir datos similares los cuales no están etiquetados. En el aprendizaje supervisado las respuestas o etiquetas se conocen previamente en la fase de entrenamiento. Este tipo de aprendizaje está dividido en dos grupos que son: clasificación y regresión [40], [41].

2.3.2. Aprendizaje no supervisado

A diferencia del aprendizaje supervisado donde ya existen datos de entrenamiento previamente etiquetados para generar una determinada salida. En el aprendizaje supervisado ocurre lo contrario, donde los algoritmos aprenden los patrones, características y relaciones de los datos en bruto sin ningún tipo de ayuda o supervisión. El objetivo principal de este método es la extracción de la información importante de los datos. Este tipo de aprendizaje está dividido en agrupación, reducción de la dimensionalidad, detección de anomalías, y minería de reglas de asociación [40], [41].

2.3.3. Áreas de aplicación de ML

Los diferentes campos de aplicación donde ML puede aplicarse incluyen: la medicina, para la identificación de tumores benignos a través de imágenes; entidades bancarias, detectando transacciones fraudulentas con tarjetas de crédito; oficina de correos, identificación de código postal a partir de dígitos escritos a mano; web, detección de entradas anómalas en un sitio web; sitio de compras, dividir a los cliente en grupos de acuerdo a sus preferencias similares; transporte, predicción del tiempo de salida y llegada de un usuario a su destino, determinar el momento de acelerar, frenar y girar en

vehículos autónomos equipados con diferentes tipos de sensores como por ejemplo GPS [39], [41].

Con el fin de mejorar el rendimiento de los sistemas GPS algunos estudios realizan una estimación de la posición del vehículo rastreado cuando existen cortes en la señal y no hay información ya sea por una interrupción desde los satélites al dispositivo GPS o una interrupción en el sistema terrestre de envío de datos, pudiendo ser este sistema un GPRS/LTE o LPWAN/LoRa (como es el caso de este trabajo). Para ayudar en la precisión de la estimación de la ubicación de un dispositivo GPS, es establecida una relación entre las características de un sistema inercial (la velocidad, la tasa angular y la fuerza específica) y la posición GPS. Cuando los datos del sistema inercial son obtenidos de forma continua, es útil la utilización de un filtro de Kalman (*Kalman Filter* o KF) con el objetivo de eliminar el ruido de las variables medidas.

Los modelos de ML más usados en la predicción son: redes neuronales artificiales (ANN) [22]; el perceptrón multi nivel (*Multi-Layer perceptrón* o MLP) [17] que es una red neuronal de más de una capa; redes neuronales de gran memoria a corto plazo (*Long Short Term Memory* o LSTM) [18] usadas para manejar series de tiempo; redes neuronales LSTM bidireccionales [20] capaces de manejar series de tiempo en dos sentidos; y el Modelo de Promedio móvil integrado autorregresivo (*Autoregressive Integrated Moving Average* o ARIMA) [19], un modelo estadístico para encontrar patrones para realizar una predicción.

2.4. Redes neuronales

En esta sección son explicados los modelos de algoritmos encontrados en la literatura [17], [18], [19], [20], [21], [22], [23] y [24] revisada e implementados para predecir la ubicación de vehículos mediante redes neuronales y otros complementos que mejoran la precisión.

2.4.1. Redes Neuronales (*Neural Networks* o NN)

Las Redes Neuronales (NN) son una forma de programación que emula la forma en la que el cerebro humano procesa la información interconectando unidades de procesamiento que parecen versiones lógicas de una neurona [42]. La **Figura 2** presenta los elementos de una red neuronal.

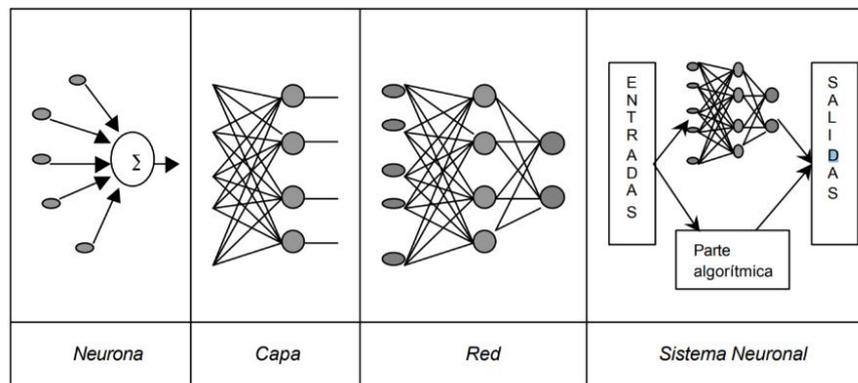


Figura 2. Elementos de una red neuronal. Fuente: [43].



La neurona contiene un conjunto de entradas que representan las variables que determinan un comportamiento de otra variable a evaluar que puede obtenerse a través de las salidas, la neurona procesa la información como una suma ponderada de variables de entrada. Una NN está organizada en capas, normalmente se compone de una capa de entrada encargada de evaluar un resultado o componente del sistema, una o varias capas ocultas y una capa de salida con unidades que representan los datos de salida. La NN aprende a través del entrenamiento generando una predicción y generando ajustes a las ponderaciones cuando realiza predicciones incorrectas hasta alcanzar ciertos criterios predefinidos [42].

2.4.2. Red de Gran Memoria a Corto Plazo (*Long Short Term Memory* o LSTM)

Son un tipo de NN recurrente capaz de aprender la dependencia del orden en problemas de predicción de secuencias. Tiene conexiones de retroalimentación, puede procesar imágenes, voz y video. Puede usarse en reconocimiento de manuscritos, reconocimiento de imágenes, reconocimiento de voz, detección de anomalías en redes de tráfico o reconocimiento de intrusos.

Comúnmente una unidad LSTM está compuesta por una célula, una compuerta de entrada, una compuerta de salida y una compuerta de olvido. La célula recuerda valores arbitrarios de intervalos de tiempo y las tres compuertas regulan el flujo de información que sale de la célula [44].

Una variante de las redes LSTM son las redes de Memoria Bidireccional Profunda a Largo Plazo (*Deep Bidirectional Long Short-Term Memory* o DBLSTM) que son el resultado de combinar los conceptos usados en las redes neuronales convolucionales bidireccionales y la célula de una red neuronal LSTM, Las redes DBLSTM hacen uso de salidas futuras procesando datos en ambas direcciones con dos capas ocultas que son enviadas a una misma capa de salida [45].



Capítulo 3

3. Estado del arte

3.1. Revisión de la literatura

En la revisión de la literatura de este trabajo fue utilizada la metodología del mapeo sistemático [46], que permite la selección, lectura y clasificación de la información disponible mediante cinco fases: alcance de la investigación, criterios de inclusión/exclusión, búsqueda de la información, esquema de clasificación y resultados. El desarrollo que fue llevado a cabo en cada fase es mostrado a continuación.

3.1.1. Fase de alcance de la investigación

En esta fase fue realizada la identificación del objetivo del mapeo sistemático, el cual fue determinar qué tipos de propuestas de sistemas de seguimiento de servicio de transporte público existían, que tuvieran cierto tipo de predicción de ubicación mediante aprendizaje automático. También fueron identificadas las razones del por qué realizar el mapeo sistemático, las cuales fueron: identificar propuestas similares a este trabajo; conocer sobre los estudios existentes con el uso de la tecnología *LoRa*, *LoRaWAN*, y/o aprendizaje de máquina para predecir la ubicación en un sistema de seguimiento; e identificar vacíos, necesidades y posibles soluciones referentes al trabajo de investigación. Por último, fueron establecidas las preguntas de investigación del mapeo sistemático y por cada pregunta fueron identificadas palabras claves y sus sinónimos. Las preguntas identificadas fueron: ¿Qué tipos de algoritmos de predicción son utilizados en el seguimiento de vehículos? (palabras clave: *artificial intelligence*, *Machine Learning*, *prediction*, *location*, *tracking*, *vehicle*), ¿Qué elementos de una red *LoRa* y protocolo *LoRaWAN* son tenidos en cuenta para el seguimiento de vehículos? (palabras clave: *LoRa*, *LoRaWAN*, *vehicle*, *location*, *tracking*).

3.1.2. Fase de criterios de inclusión/exclusión

En esta fase fue establecida e identificada una lista de criterios de inclusión o exclusión clasificándola por búsqueda o análisis. Por medio de esta fase fue posible diferenciar entre lo que se quería encontrar y lo que no era necesario. Los criterios clasificados fueron:

- **Inclusión y búsqueda:** El documento fue publicado en el año 2017 o posteriormente. El documento tiene los términos: "*LoRa*, *LoRaWAN*, *vehicle*, *Machine Learning*, *artificial intelligence*, *prediction*, *tracking*" en el título.
- **Inclusión y análisis:** El documento describe modelos de prototipo de rastreo y seguimiento. El documento determina los algoritmos de predicción de ubicación utilizados.
- **Exclusión y análisis:** El artículo presenta un modelo de cualquier otro tema. El contenido del documento no está relacionado con el seguimiento de vehículos mediante *LoRa* o *LoRaWAN* o un modelo de predicción.



3.1.3. Fase de búsqueda de la información

En esta fase fueron construidas dos cadenas de búsqueda a partir de las palabras claves obtenidas de las preguntas de investigación del mapeo sistemático de la fase uno. La primera cadena de búsqueda fue: ("*Machine Learning*" OR "*artificial intelligence*") AND ("*prediction*") AND ("*location*" OR "*tracking*") AND ("*vehicle*"). La segunda cadena de búsqueda fue: ("*LoRa*" OR "*LoRaWAN*") AND ("*vehicle*") AND ("*location*" OR "*tracking*"). A partir de ellas fue realizada una búsqueda en las bases de datos de Scopus y ScienceDirect de los últimos cinco años. Filtrando los resultados fueron obtenidos un total de 1978 documentos.

3.1.4. Fase de esquema de clasificación

Los tipos de investigación y el contexto de investigación fueron determinados con la revisión inicial de los documentos (revisando *abstract* y título), no fueron tomados de ninguna fuente, fueron establecidos al revisar los documentos, identificando características similares y agrupando (tal como recomienda el autor de [46] respecto a los mapeos sistemáticos). En esta fase fueron establecidos los términos de dos tipos de clasificación. La primera clasificación fue por tipo de investigación. Los tipos de investigación determinados fueron:

- *Evaluation Research* (Este tipo de trabajos realiza una investigación de evaluación, no realiza una propuesta de una solución).
- *Solution proposal*. (Este tipo de trabajos propone una solución, mas no realiza mediciones).
- *Validation research* (Este tipo de trabajos realiza pruebas o experimentos para validar la solución propuesta).
- *Others*.

La segunda clasificación fue por el contexto de la investigación. Los contextos considerados fueron:

- *Vehicle tracking*
- *Location prediction*
- *Public vehicle tracking*
- *Public vehicle tracking with LoRa or LoRaWAN*
- *Vehicle tracking and prediction*
- *Vehicle tracking and prediction with Machine Learning*.
- *Discarded*.

Revisando el título y resumen (*abstract*) de los 1978 documentos y siguiendo los criterios de la fase tres fueron obtenidos 119 documentos clasificados por tipo de investigación y contexto de investigación.

3.1.5. Fase de resultados

En esta última fase fue obtenida una intersección de la cantidad de documentos entre el tipo de investigación y el contexto de la investigación. Esta información está representada en un mapa de burbujas en la **Figura 3**.

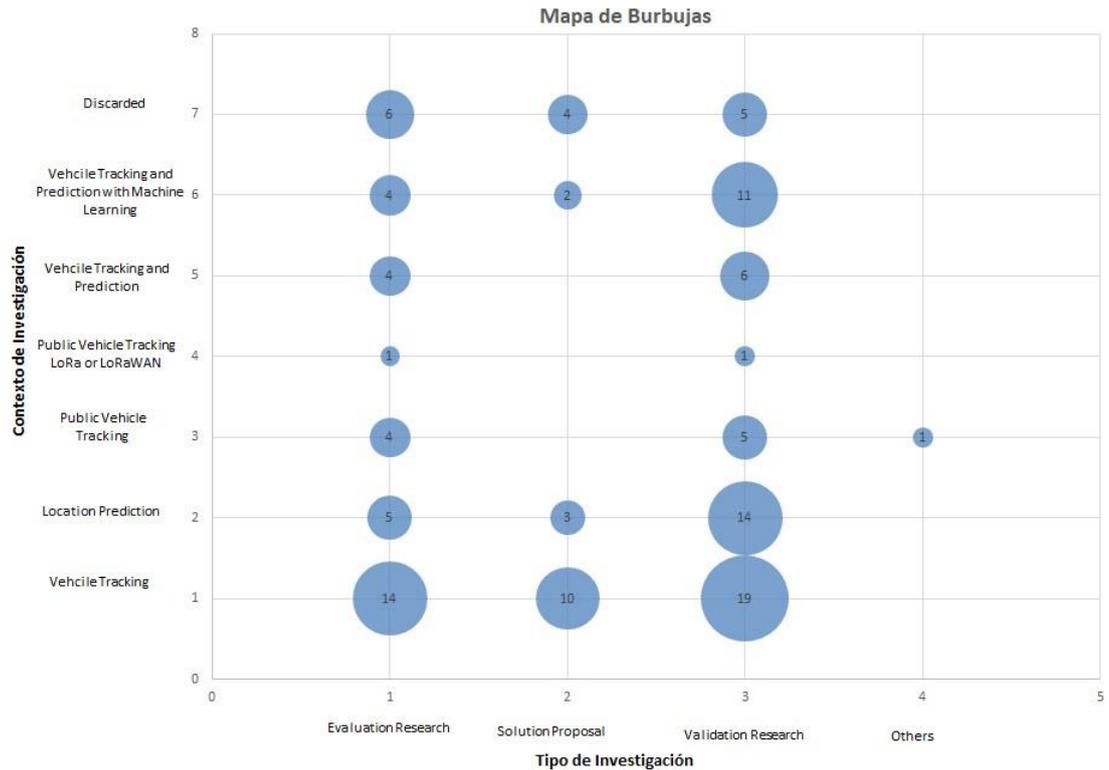


Figura 3. Mapa de burbujas del mapeo sistemático.

De los 119 documentos encontrados y representados en el mapa de burbujas, fueron descartados 100 documentos. Fueron seleccionadas las burbujas de las intersecciones que tuvieran una relación directa con el trabajo de investigación y en las cuales no hubiese demasiados trabajos previos. Fueron seleccionados los 19 trabajos identificados en la fila del contexto de investigación denominado “*Vehicle tracking and prediction with Machine Learning*” (3 con el tipo de investigación denominado “*evaluation research*”, 2 con el tipo de investigación “*solution proposal*”, y 11 con el tipo de investigación “*validation research*”). Adicionalmente fueron seleccionados los 2 trabajos identificados en la fila de contexto de investigación denominado “*Public vehicle tracking with LoRa or LoRaWAN*” (uno de los trabajos con el tipo de investigación “*evaluation research*” y otro con tipo de investigación “*validation research*”).

Después de obtener los 19 documentos a través de las fases del mapeo sistemático, fue realizada una revisión de otras fuentes con el objetivo de encontrar nuevos artículos relacionados, antes de pasar a la etapa de documentos relacionados. Fueron encontrados 2 nuevos documentos relacionados con el uso de *LoRa* y/o *LoRaWAN* en el transporte público. Por lo cual fueron obtenidos un total de 21 artículos a revisar en la siguiente etapa de revisión de literatura.

Con base en lo anterior, no fue encontrada una investigación que trabaje de manera conjunta el seguimiento vehículos de transporte público, mediante el uso de la tecnología *LoRa/LoRaWAN* y que también realice la predicción de la ubicación del vehículo por medio de modelos de ML.



3.2. Documentos relacionados

A partir de los 21 documentos encontrados (con el proceso de mapeo sistemático y la revisión posterior), fueron descartados posteriormente 9 artículos a través de una revisión detallada de los mismos. Los 9 documentos descartados realizaban el seguimiento de automotores por medio de vehículos aéreos no tripulados, cámaras, automóviles como nubes vehiculares e interconexión entre estos, y predicción del número de pasajeros por autobús. Estos temas no están relacionados directamente con el objeto del trabajo por lo cual no fueron considerados.

Los 12 documentos finalmente seleccionados para el estado del arte del trabajo fueron los siguientes.

- El artículo ***Unifying Multi-Radio Communication Technologies to Enable mMTC Applications in B5G Networks*** [10] investiga la utilidad de usar múltiples tecnologías de acceso radioeléctrico en un mismo sistema. Los vehículos de servicio público son rastreados mediante dos tecnologías: NBloT (3gpp) y LoRa usando redes 5G y posteriores (B5G), mostrando que esta combinación aporta una mayor flexibilidad, fiabilidad y mejora de la calidad de servicio global a la conectividad inalámbrica.

Este trabajo utiliza la tecnología LoRa en el caso específico de seguimiento de vehículos, aunque cabe resaltar que se hizo uso de otra tecnología LPWAN como lo es NBloT.

- El artículo ***Experimental Evaluation of LoRa in Transit Vehicle Tracking Service Based on Intelligent Transportation Systems and IoT*** [11] propone un prototipo de seguimiento de vehículos basado en una arquitectura ITS, haciendo uso de la tecnología LoRa en seis diferentes escenarios fijos y móviles, realizando diferentes configuraciones del factor de propagación (*Spread Factor* o SF), ancho de banda (*BandWidth* o BW) y tasa de codificación (*Coding Rate* - CR).

Este artículo ofrece información relevante sobre los experimentos realizados y las configuraciones establecidas. Aspectos de este trabajo como la velocidad, LoS, número de dispositivos, y limitaciones de los experimentos (corto alcance, inconvenientes con la LoS) son relevantes para el diseño de las pruebas del trabajo planeado en este documento. La diferencia entre este trabajo y el planteado (en este documento) es que la propuesta utilizaría algoritmos de predicción de ubicación del vehículo, usando ML cuando no exista LoS, lo cual no es propuesto en este trabajo.

- El artículo ***IoT-based Bus Location System Using LoRaWAN*** [12] presenta un sistema de localización de buses basado en IoT con el fin de mejorar la calidad del servicio y la eficiencia de la gestión de operación. Esta propuesta utiliza dispositivos LoRa a bordo de los buses para recoger información de la posición, procesa datos en un servidor implementado en la nube y luego muestra esa información en las paradas de autobús, teniendo con todo esto un costo de operación menor que usar la tecnología 3G/LTE.

En este trabajo es utilizado LoRa y el protocolo de red LoRaWAN. Los experimentos realizados en este trabajo también nos sirven de base para el diseño de las pruebas en la propuesta descrita en este documento. Al igual que el trabajo anterior fueron evidenciados inconvenientes en la operación de la tecnología LoRa en algunas zonas (de las rutas propuestas), por lo cual se



identifica la importancia de una herramienta alternativa para obtener la ubicación del vehículo en estos casos.

- El artículo ***Proof of Concept of an IoT-Based Public Vehicle Tracking System, Using LoRa (Long Range) and Intelligent Transportation System (ITS) Services*** [13] propone un sistema de seguimiento para vehículo público utilizando la tecnología *LoRa* e implementando la arquitectura ITS ARC-IT, donde por medio de GPS se obtiene la ubicación del vehículo, la velocidad y altitud. Los datos fueron obtenidos por medio de dos tarjetas micro controladas que hacen uso de *LoRa*.

El uso de ciertos parámetros de propagación de la tecnología *LoRa* que utiliza este trabajo son útiles para utilizarlos en la propuesta descrita en este documento. También fueron identificados problemas en el alcance y el porcentaje de paquetes recibidos, ocasionados principalmente por obstáculos en la LoS.

- ***A hybrid fusion algorithm for GPS/INS integration during GPS outages*** [17] propone un algoritmo que relaciona datos del INS como la velocidad, tasa angular y fuerza específica con el incremento de la posición en el sistema GPS con lo que es posible predecir la posición de GPS cuando la señal de este no está disponible.

Este artículo nos permite reconocer la viabilidad de usar modelos de ML para ofrecer datos de ubicación de un vehículo cuando la señal no es recibida.

- ***Joint prediction of next location and travel time from urban vehicle trajectories using long short-term memory neural networks*** [18] explora dos modelos de predicción diferentes usando redes neuronales de gran memoria de corto plazo (*Long Short-Term Memory* o LSTM). Lo anterior lo realiza capturando patrones de la trayectoria en espacio y tiempo para predecir la ubicación siguiente y el tiempo de viaje simultáneamente.

La diferencia principal de este trabajo con la propuesta (descrita en este documento) está en la forma de recolectar datos. Mientras que lo propuesto es la recolección mediante *LoRa*, en el artículo se usa la tecnología *Bluetooth*, con detectores en distintas partes del recorrido separados por una distancia mínima de 100 metros. En este artículo, los autores afirman que considerar el tiempo de viaje juntamente con la predicción de ubicación mejoró notablemente la precisión del modelo pues el tiempo de viaje incluye mediciones sobre tráfico que pueden afectar las tomas de decisiones de ubicación de los conductores.

- ***A Hybrid Machine Learning Based Low Cost Approach for Real Time Vehicle Position Estimation in a Smart City*** [19] trata sobre el entrenamiento de un modelo de NN de la posición del vehículo mientras la señal GPS está disponible. De esta forma, durante las interrupciones de la señal ya sea por pasar por un túnel o una avería en el sistema se pueden proporcionar buenas predicciones de la ubicación. El enfoque propone combinar la anterior NN con un modelo ARIMA para capturar la estructura lineal.

Este estudio nos aporta la división de dos modelos para tener en cuenta, el lineal y el no lineal, además de un modelo de estimación complementario a las redes neuronales para estimar datos. Este estudio es aplicable a cualquier tipo de



vehículo por lo que, para el servicio público, con rutas preestablecidas, la variabilidad de los datos es menor y se tiene menor incertidumbre.

- **Vehicle Trajectory Prediction Method based on Deep Learning under the background of Internet of Vehicles** [20] utiliza CNN y RNN para el desarrollo de un algoritmo de predicción. El algoritmo utiliza datos históricos de trayectoria del vehículo como: detección de matrícula, número de carril y secuencia de GPS, obteniendo las características de trayectoria del vehículo en movimiento para después predecir la siguiente ubicación.

Los autores de este artículo afirman que para el procesamiento de los datos recolectados y lograr predecir la ubicación es más conveniente el uso de RNN, debido a que la posición siguiente y la anterior, no son independientes. Lo anterior es útil para el análisis del modelo de ML ideal, para la propuesta descrita en este documento.

- **An Algorithm of Curved Path Tracking with Prediction Model for Autonomous Vehicle** [21] propone un método predictivo de seguimiento cuando un vehículo autónomo debe girar en las esquinas de una vía (debido a que el método Stanley sufre retrasos en estos puntos). La recopilación de los datos en este trabajo fue realizada con un vehículo sonda equipado con GPS y INS que permite conocer la geometría de la carretera. Además, por medio de curvas de línea polinómica suave básica (*Basis spline* o B-spline) cúbicas se realizó el modelado del camino para finalmente realizar el seguimiento predictivo en la trayectoria curva.

El algoritmo utilizado en este trabajo para la predicción de la ubicación en el cruce podría ser útil para la predicción que pretende realizar la propuesta descrita en este documento.

- El trabajo **Customizable vehicle tracking with intelligent prediction system** [22] establece el seguimiento de vehículos públicos por medio de GPS, GSM y GPRS. Los datos de seguimiento son almacenados en un servidor y en base a estos informa al usuario a través de una aplicación móvil sobre qué tan cerca está un autobús a una de las paradas. Si no hay acceso a internet el sistema predice el tiempo de llegada al destino con base en las paradas que realiza una persona para llegar a su destino. El procesamiento de esta información genera una salida por medio de una ANN.

El artículo se diferencia de la propuesta descrita en las tecnologías de comunicaciones utilizadas (*LoRa/LoRaWAN* para la propuesta presentada) para realizar el seguimiento vehicular y enviar los datos al servidor. En el documento, los autores afirman que el uso de ANN es más preciso para predecir el tiempo de llegada frente a un modelo histórico o modelos de regresión. Este aporte será considerado en el desarrollo del trabajo de la propuesta.

- En **Attention-based Recurrent Neural Network for Urban Vehicle Trajectory Prediction** [23], los autores proponen una mejora a un anterior trabajo desarrollado por ellos, donde utilizaron RNN para la predicción de la ubicación basado en la trayectoria. En este artículo, ellos proponen un modelo ARNN donde no solo fueron consideradas las trayectorias anteriores para pronosticar la localización, sino que también el estado del tráfico que es una entrada no secuencial a diferencia de la secuencia de trayectorias. Los datos de movimiento de los vehículos son obtenidos desde dispositivos *Bluetooth* instalados en las vías e intersecciones. El modelo propuesto en este trabajo puede ser útil en la



definición del modelo a proponer.

- El artículo ***A Novel Generation Adversarial Network-based Vehicle Trajectory Prediction Method for Intelligent Vehicular Networks*** [24], presenta el método de predicción de trayectorias de vehículos basado en una Red Generativa de Adversarios (*Generative Adversarial Network (GAN)-Based Vehicle Trajectory Prediction Method* o GAN-VEEP). Este método utiliza coordenadas espaciales del vehículo para después ser procesadas, garantizando una mejor precisión en la predicción, a la vez realizan una división de la vía en segmentos de carretera (*Road Segments* o RSGs) donde el GAN sólo predice cuando el automóvil se encuentre dentro de esta. Para solucionar esto aplican un modelo de giro basado en la psicología del conductor. Las pruebas son llevadas a cabo en simulador de movilidad urbana (*Simulation of Urban Mobility* o SUMO) obteniendo buenos resultados de GAN-VEEP en escenarios urbanos.

La principal diferencia de este trabajo, respecto a la propuesta presentada, es que las pruebas son realizadas a través de simulación y no en un entorno real para verificar qué tan precisa es la predicción de la ubicación del vehículo.

Dos aspectos que no fueron analizados en cada uno de los artículos individualmente son: el uso de sensores y dispositivos asequibles, y una ciudad intermedia o pequeña como contexto de la investigación. Respecto al primer aspecto, los trabajos que incluyeron sensores o dispositivos de medición fueron [10], [11], [12], [13], [18], [19] y [22], los demás trabajos están más relacionados con el uso de algoritmos u otras herramientas de ML para realizar predicción de la ubicación del vehículo. Todos estos artículos, a excepción de [12] utilizaron elementos que es viable adquirir en el mercado nacional actualmente (*Gateways*, Tarjetas ESP32, dispositivos *Bluetooth*, sensores GPS, módulos GSM, o microcontroladores). En [12] fueron utilizados dispositivos utilizados comúnmente en Asia. Debido a que el rango de frecuencias de operación de *LoRa* es diferente en cada uno de los continentes, dichos elementos no podrían ponerse a operar en el país (continente americano). Respecto al segundo aspecto, los trabajos que tuvieron como contexto una ciudad intermedia o pequeña (población inferior a 1.000.000 de habitantes) a nivel mundial fueron [10], [11], [12] y [13], los demás trabajos tuvieron otro tipo de ciudades como contexto, o no fueron trabajos enfocados a un contexto de aplicación específico (como una ciudad).

3.3. Resumen de evaluación de los documentos relacionados

La **Tabla 1** presenta un resumen de los aspectos considerados en las propuestas de los documentos relacionados que conforman el estado del arte. La propuesta de investigación de este trabajo tiene en cuenta todos los puntos identificados, mientras los demás trabajos solo tienen en cuenta algunos de dichos aspectos.



Tabla 1. Resumen de evaluación de los trabajos relacionados.

Criterios/Propuesta	[10]	[11]	[12]	[13]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]
Emplea modelos de ML para predecir ubicación.					✓	✓	✓	✓	✓	✓	✓	✓
Seguimiento vehicular con LoRa/LoRaWAN.	✓	✓	✓	✓								
Uso de sensores y dispositivos asequibles.	✓	✓		✓		✓	✓			✓		
Ciudad intermedia o pequeña como contexto.	✓	✓	✓	✓								

3.4. Revisión del estado del arte

Dado el sistema inteligente propuesto en este trabajo el cual realiza el seguimiento del servicio de transporte público y la recolección de datos de vehículos, es necesario determinar las características del módulo *Hardware* que permita la comunicación y recolección de ciertas variables y su almacenamiento. Igualmente, es necesario identificar las características de un módulo de *software* encargado de realizar el seguimiento vehicular en tiempo real (a través de la tecnología *LoRa* y aprendizaje de máquina).

Para adaptar los módulos *Hardware* y *Software* de la mejor manera a las necesidades y contexto del presente trabajo, fue realizada una revisión detallada del estado del arte. La revisión de cada artículo fue enfocada en resolver las siguientes preguntas: ¿Qué variables o datos consideró y para qué los utilizó?, ¿Qué *hardware* utilizó para la medición o recolección de los datos?, ¿Cómo fue el experimento o prueba de campo empleado en la recolección? y ¿Qué algoritmos o *software* implementó para el análisis o procesamiento de esos datos?

Las respuestas a las anteriores preguntas permitieron obtener una visión general de diferentes variables, dispositivos, modelos y/o técnicas de procesamiento de datos empleados en mejorar los sistemas de seguimiento vehicular en tiempo real basados en aprendizaje de máquina (similares a la hipótesis de este trabajo), y cómo estos están correlacionados.

En las siguientes **subsecciones 3.4.1 a 3.4.4**, es realizada la revisión y análisis de las diferentes opciones para dispositivos, sensores y variables.



En la **sección 3.4.5** es realizada la revisión y análisis de los posibles modelos de ML encontrados en el estado del arte.

3.4.1. Identificación de dispositivos, sensores y variables.

La **Tabla 2** muestra los dispositivos o sistemas identificados en la literatura, que son usados en la recolección de los datos vehiculares.

Tabla 2. Hardware utilizado en recolección de datos de vehículos.

Dispositivo	Sensores	Variables	Dispositivos adicionales y plataformas (software)	Tipo de experimento	Referencia
Microcontrolador + módulos <i>LoRa</i> y NB-IoT	GNSS	Longitud, latitud, altitud, velocidad, nivel de RSSI.	Infraestructura de comunicación de Ceske´ Radiokomunikace <i>LoRaWAN</i> (<i>Gateway LoRaWAN</i>), Redes NB-IoT de Vodafone (LTE), plataforma de gestión de datos de IoT.	Seguimiento del transporte público por medio de las tecnologías radio <i>LoRa</i> y NB-IoT conmutando entre sí, las mediciones fueron obtenidas de 74 lugares de la ciudad de Brn, Republica Checa.	[10]
Esp32 <i>LoRa</i> Heltec	GPS	Latitud, longitud, distancia Tx-Rx, número de dispositivos Tx, spreading factor, RSSI, potencia de transmisión, frecuencia de transmisión, ancho de banda, tasa de código.	<i>Gateway LoRa</i> , software para hojas de cálculo y gráficas, google maps.	Fue utilizada la tecnología <i>LoRa</i> para realizar seguimiento vehicular. Fueron variados los parámetros SF, BW y CR de esta tecnología, en 6 experimentos. Los 3 experimentos iniciales fueron a distancias fijas, mientras que los 3 finales fueron realizados con vehículos en movimiento.	[11]



<p>Arduino UNO y dispositivo LoRa</p>	<p>GPS</p>	<p>Latitud, longitud, tiempo.</p>	<p>Gateway LoRa, servidor de gestión de autobuses en la nube, paradas de autobús inteligentes (Raspberry PI 3 y e-paper).</p>	<p>Se utilizó la tecnología inalámbrica LoRa para realizar el seguimiento de la ubicación de autobuses recorriendo ciertas rutas. Luego procesar los datos e informar a los pasajeros desplegando información en paradas de autobús inteligente, (distancia del próximo bus y tiempo de retraso). La máxima distancia alcanzada entre el Gateway y el OBE fue de 2.4km. La máxima velocidad en la que se recolectaron datos adecuadamente fue 50 km/h.</p>	<p>[12]</p>
<p>ESP32 LoRa Heltec</p>	<p>GPS</p>	<p>Posición (latitud, longitud), velocidad, altitud, distancia al Gateway. tasa de paquetes recibidos, RSSI, tiempo (fecha y hora de registro).</p>	<p>Gateway LoRa. plataforma TTN, Arduino IDE, plataforma XAMP, mapas de Google.</p>	<p>Consistió en el seguimiento vehicular usando la comunicación inalámbrica LoRa, para lo cual fueron realizados 5 experimentos. En los 4 experimentos iniciales fue verificada la operación adecuada de los dispositivos con distancias fijas, operando inicialmente con comunicación entre dispositivos, luego el uso de un Gateway, luego adicionando el uso de un GPS.</p>	<p>[13]</p>



				Finalmente, en el 5º experimento se utilizaron todos los elementos anteriores en 2 vehículos en movimiento, siguiendo 2 rutas. Se realizaron mediciones de recepción de mensajes.	
No especifica.	<ul style="list-style-type: none"> GPS. giroscopio de fibra óptica, Acelerómetro de quartz. 	Velocidad angular, velocidad, fuerza específica, latitud y longitud.	PHINS (Obtiene los datos de navegación de referencia). IMU	En el experimento un vehículo siguió una trayectoria, provocando una interrupción de 300 segundos en el GPS. La duración del experimento fueron 4700 segundos, la tasa de actualización fue de 200Hz. Durante estos cortes se pusieron a prueba 3 modelos, 2 basados en entradas del GPS y el INS y 1 con el INS solamente. El INS estaba conformado por los sensores de giroscopio de fibra óptica y un acelerómetro de quartz.	[17]



No especifica	<i>Bluetooth</i>	Tiempo de viaje entre sensores, seguimiento del ID de los dispositivos individuales <i>Bluetooth</i> .	No especifica.	El experimento consistió en ubicar sensores <i>Bluetooth</i> en las vías, con el objetivo de registrar el ID, la dirección MAC individual del <i>Bluetooth</i> en el vehículo, sistemas de navegación y dispositivos móviles. A partir del registro de los dispositivos fueron construidos los trayectos que siguieron los vehículos. La velocidad máxima de los vehículos en la red fue de 80km/h. La recolección de los datos fue realizada entre julio de 2015 a junio de 2016.	[18]
Arduino nano y módulo Xbee. Raspberry y módulo Xbee.	Odómetro Giroscopio GPS	Distancia recorrida, velocidad, Velocidad angular, latitud y longitud.	Plataforma VIPALAB.	En el experimento se realizó el seguimiento de un vehículo a través de una trayectoria de 4.2 km durante 28 minutos. Fueron realizados 5 cortes de la señal del GPS con una duración de 100 segundos. Antes de cada interrupción fueron obtenidos nuevos datos del GPS por un periodo de al menos 96 segundos para lograr entrenar y generar las respectivas salidas de los algoritmos de predicción. Los datos del GPS fueron	[19]



				recolectados a una tasa de frecuencia de 1Hz y los del odómetro y giroscopio a 50Hz.	
Dispositivos móviles	GPS. Cámara.	Latitud, longitud, número de carril.	No especifica.	El experimento consistió en realizar un registro de los vehículos cuando pasan por cámaras de registro de matrícula donde las variables a medir son el número de carril y la secuencia GPS proporcionada por dispositivos móviles. La duración del experimento fue de una semana.	[20]
No especifica	GPS INS	Latitud y longitud.	No especifica.	Consistió en enviar un vehículo autónomo equipado con GPS y un INS de alta precisión a través de una carretera. Con el objetivo de capturar los datos de los sensores y obtener información de la geometría de la carretera.	[21]



<p>Arduino Uno y módulo GSM/GPRS</p>	<p>GPS</p>	<p>Intervalos de tiempo, latitud y longitud.</p>	<p>Modem GSM, estación base (Antena), teléfono inteligente, mapa de Google, aplicación móvil, servidor Web de Amazon.</p>	<p>El experimento consistió en realizar el seguimiento de un autobús a través de GPS, donde un usuario a través de una aplicación móvil puede visualizar la ubicación del autobús. El sistema también sugiere al usuario buses alternativos cuando la persona pierde su transporte. Adicionalmente, el sistema predice el tiempo de llegada al destino.</p>	<p>[22]</p>
<p>No especifica.</p>	<p><i>Bluetooth</i></p>	<p>Detección del tiempo de paso de dispositivos <i>Bluetooth</i>, cantidad de vehículos por celda.</p>	<p>No especifica</p>	<p>El experimento consistió en utilizar el conjunto de datos recolectados en marzo de 2016, utilizando sensores <i>Bluetooth</i> ubicados en las carreteras e intersecciones, para realizar una predicción de la ubicación del vehículo. Fue considerado que, si el vehículo no se mueve por más de una hora, el viaje finalizó. El algoritmo de predicción utilizó otra entrada que consistió en medir la cantidad de vehículos por celda (las celdas fueron construidas a partir de las trayectorias identificadas de los vehículos) que identificaría</p>	<p>[23]</p>



				la cantidad de tráfico. Este trabajo está relacionado con [18].	
Datos simulados por software	Datos generados por software	Trayectoria de vehículos, densidad vehicular, posición del flujo de tráfico.	No especifica.	Consistió en simular múltiples trayectorias de vehículos, información de trayectoria de los vehículos (coordenadas, velocidad y la identificación del vehículo) y las características de la carretera (longitud de la carretera y estado de los semáforos) por medio del software SUMO en un mapa de 1.000 metros por 1.000 metros y con una densidad de vehículos de 250/km ² . Donde a las coordenadas de los vehículos se les aplicó un modelo de transformación de coordenadas con el objetivo de normalizar estos datos. Para después utilizarlos por el algoritmo de predicción de la trayectoria.	[24]
Esp32 LoRa Heltec. Heltec Cube Cell	GPS Temperatura y Humedad	Latitud, longitud, velocidad, tiempo de transmisión, paquetes, temperatura, humedad, RSSI.	Software QGIS, Gateway LoRa, Gateway LoRaWAN, Plataforma TTN. Servidor en la nube.	En el experimento se realizó seguimiento vehicular utilizando comunicación LoRa y su protocolo de red LoRaWAN. En la primera fase se realizaron pruebas de operación de los dispositivos LoRa, el Gateway y la plataforma TTN	[47]



				<p>a distancias fijas. En las otras dos fases se utilizaron rutas en las vías de la ciudad de Popayán y vehículos que las recorrieron utilizando cada uno dispositivos LoRa. Los paquetes recibidos en los Gateways fueron almacenados en un dataset. Se recolectaron datos sobre la posición desde donde fueron enviados los paquetes, por lo cual fue posible reconstruir el recorrido del vehículo e identificar las zonas con inconvenientes de comunicación.</p>	
--	--	--	--	---	--

BW: Ancho de banda, CR: Code Rate, Gateway: Pasarela, GNSS: Sistema Global de Navegación por Satélite, GPS: Sistema de Posicionamiento Global, INS: Sistema de Navegación Inercial, IMU: Unidad de Medida Inercial, LoRa: Long Range, LoRaWAN: Long Range Wide Area Network., LTE: Long Term Evolution, NBloT: Narrowband-Internet of Things, PHINS: Sistema de Navegación Inercial de Alto Rendimiento, RSSI: Indicador de fuerza de la señal recibida, SF: Factor de dispersión, SUMO: Simulation of Urban Mobility, TTN: The Things Network.

3.4.2. Análisis de dispositivos de recolección de datos

Las tarjetas micro controladas fueron los dispositivos más utilizados en los trabajos en los que realizaron una recolección de datos de recorridos vehiculares. En los trabajos [10], [11], [12], [13], [19], [22] y [47] fue observado el uso de las tarjetas: Raspberry, Arduino UNO, Arduino Nano, ESP32 LoRa Heltec y Heltec Cube Cell HTCC – AB02S.

Entre la categoría de tarjetas micro controladas están la **Arduino UNO y Arduino Nano**, las cuales basan su funcionamiento en un microcontrolador Atmega. Las ventajas que presentan estas tarjetas son: permiten la conexión de diferentes sensores a través de las entradas y salidas analógicas y digitales del microcontrolador; el *hardware* permite la conexión con otras tarjetas o *shields* (escudos); utilizan interfaz USB tipo B; bajo costo de adquisición; y voltaje de alimentación entre 7 – 12 Voltios [48], [49]. Las desventajas son: no incorporan un módulo de comunicación inalámbrico o algún sensor útil para la medición requerida [48], [49].



Otra tarjeta de esta categoría es la **ESP32 LoRa** del fabricante Heltec. El microprocesador que utiliza esta tarjeta es el Tensilica LX6 doble núcleo a 240 MHz. Las ventajas de esta tarjeta son: permite la programación mediante el IDE de Arduino; incorpora comunicación inalámbrica *Bluetooth*, *LoRa SX1276* y *Wi-Fi*; soporta el protocolo *LoRaWAN*; posee una pantalla OLED de 1 pulgada aproximadamente; posee una interfaz Micro USB tipo B; tamaño reducido; dispone de entradas y salidas de propósito general – GPIO y analógicas, voltaje de alimentación de 5 voltios y bajo costo de adquisición [50]. Entre sus desventajas están la falta de integración de otro tipo de sensores útiles para la medición requerida y el bajo número de pines disponibles (considerando su tamaño) [50], [51].

Por último, la tarjeta micro controlada **HTCC-AB02S** fabricada por Heltec que funciona con un “master chip” ASR6052. Las ventajas que presenta son: compatibilidad con el IDE de Arduino; incorpora un módulo *LoRa SX1262*; admite el protocolo *LoRaWAN*; incorpora un módulo *GPS (AIR530Z)*; tiene una pantalla OLED; tamaño reducido; entradas y salidas de propósito general – GPIO y analógicas; voltaje de alimentación de 5 voltios; y bajo costo de adquisición [52]. Entre sus desventajas están el no tener *Bluetooth* y *Wi-Fi* [52].

Aparte de las anteriores tarjetas micro controladas fue considerado el uso de una “**Raspberry**”, considerada como una minicomputadora con la capacidad de realizar diferentes tareas como un computador de escritorio. El rendimiento de una Raspberry depende de las características técnicas asociadas al modelo. La **Raspberry** presenta ventajas como: sistema operativo basado en Linux; incorpora un procesador y memoria RAM, lo cual permite realizar diferentes tareas de alto procesamiento y al mismo tiempo, trae una unidad de procesamiento gráfico (GPU); incorpora comunicación *Ethernet* permitiendo ser utilizado como un servidor web dedicado; de acuerdo al modelo algunos dispositivos traen *Bluetooth* y *Wi-Fi*; posee puertos USB y salida HDMI para conectar diferentes periféricos; almacenamiento a través de un puerto SD o MicroSD; entradas y salidas de propósito general – GPIO donde es posible conectar diferentes tipos de sensores, voltaje de alimentación de 5 voltios a 3 Amperios; su tamaño es reducido (aunque es mayor a las tarjetas ESP32 y HTCC-AB02S) y posee salida de audio [49], [53], [54]. Entre sus desventajas están: no tiene integrado un convertor análogo–digital; no es posible incrementar la memoria RAM; no incorpora un reloj de tiempo real; no tiene integrado un GPS u otro sensor útil para la medición requerida [55].

Respecto a las diferentes tecnologías de comunicación inalámbrica utilizadas para transmitir los datos obtenidos por los dispositivos, se identificaron las siguientes tecnologías: LoRa, NBloT, *Bluetooth*, Xbee y GSM/GPRS. Las ventajas y desventajas que estas tecnologías presentan se detallan a continuación.

LoRa. Tecnología clasificada como LPWAN, que presenta ventajas como: basa su funcionamiento en el CSS, lo cual le permite alcanzar distancias máximas de 48.27 km; baja potencia de alimentación prolongando la vida útil de los dispositivos que funcionan como nodos hasta 10 años; inmunidad al ruido y baja interferencia [8], [16], [55]; trabaja en las bandas de frecuencia no licenciadas de 433 MHz en Asia, 868MHz en Europa y 915 MHz en América; hasta 20.000 dispositivos conectados por *Gateway*; y presentan un bajo costo de operación e infraestructura [8], [37], [38], [55]. Las desventajas que presenta son: requiere de línea de vista continua entre los dispositivos terminales y los equipos *Gateway*; velocidades hasta de 50 Kbps; y reducido tamaño de paquetes (255 Bytes) . Estas dos últimas pueden considerarse como desventajas dependiendo del proyecto a implementar.



NB-IoT. También hace parte de las LPWAN. Las ventajas presentadas por NB-IoT son: no requiere la utilización de *Gateway* porque funciona bajo la infraestructura de GSM y LTE donde su comunicación es a través de estaciones base; alcanza grandes distancias de cobertura; baja potencia de alimentación prolongando la vida útil de los dispositivos a 10 años; ideal para entornos urbanos densos debido a que depende de la red LTE; tamaño de los paquetes de 1600 Bytes y velocidad de transferencia hasta 250 Kbps [10], [26], [34], [37], [38]. Las desventajas que presenta son: trabaja con frecuencias licenciadas (rango entre 700 y 800 MHz); pérdidas de acuse de recibo; equipos más costosos [26], [34], [37], [38].

Xbee. Es otra de las tecnologías inalámbricas de corto alcance. Sus ventajas son: larga duración de la batería; funcionan en la frecuencia de 2.4 GHz; velocidad de datos de 250 Kbps y ofrece un alcance de 100 metros [56], [57]. La principal desventaja es la necesidad de línea de vista entre dispositivos [56], [57].

GSM y GPRS. Son tecnologías utilizadas en las telecomunicaciones móviles. Las ventajas que presenta GSM son: cobertura a nivel mundial; inicialmente fue creado para voz después fueron agregados servicios de datos; hace uso de un módulo de identidad del suscriptor - tarjeta SIM, el cual posee la información del usuario permitiendo el cambio dispositivo móvil al presentar este una falla; hace un uso eficiente del espectro [58], [59] y [60]. Las desventajas de GSM son: al no soportar IP (*Internet Protocol*) no es posible conectarse a internet; orientado a la conmutación de circuitos, por las limitaciones de costos y el ancho de banda tan solo es posible alcanzar una velocidad de datos de 9.6 Kbps y el cobro en GSM es por tiempo de conexión. GPRS es considerado la evolución de GSM al modificar la forma de cómo son transmitidos los datos. Presenta ventajas como: acceso a internet, permitiendo el envío y recepción de correo electrónico, imágenes, gráficos y etc; en teoría puede llegar a una velocidad de 171.2 Kbps, orientado a la conmutación de paquetes, el costo de tarificación está basado en los datos enviados y no por el tiempo de conexión; asigna la capacidad de la red cuando es necesario enviar paquetes para después ser liberada. Las desventajas que presenta GPRS son: velocidad limitada para algunas aplicaciones en particular, capacidad de celda limitada y modulación sub-óptima debido a que GPRS utiliza una modulación GSMK (*Gaussian Minimum Shift-Keying*) [58], [59], [60].

En los artículos [10] y [47] fue observado, que adicionalmente al uso de *LoRa*, fue utilizado el protocolo de red *LoRaWAN*. *LoRaWAN* presenta ventajas como: topología de red en estrella; los *Gateway* transmiten los datos entre los nodos y un servidor de red central; los *Gateway* se conectan al servidor de red por medio de protocolo *IP* además convierten los paquetes de radio frecuencia en paquetes IP y viceversa; el protocolo permite que los dispositivos finales operen bajo tres clases de nodos (Clase A, B y C) operando de acuerdo a los requisitos de una aplicación; ofrece una alta seguridad al usar un cifrado de extremo a extremo AES128 [14], [15]. Las desventajas que tiene *LoRaWAN* son: la necesidad de línea de vista, y tamaño del paquete [11]. Es necesario resaltar que en [47] utilizan la plataforma TTN (*The Things Network*) para la recepción de los datos cuando es utilizado *LoRaWAN*.

3.4.3. Análisis de sensores

Considerando la literatura, es importante que el dispositivo a ubicar en los vehículos tenga un sensor de *GPS*, un odómetro, y un giroscopio. El dispositivo GPS puede ser el **Ublox 6M**, componente electrónico que permite obtener las coordenadas geográficas (latitud y longitud), velocidad y altura (a nivel del mar) de un objeto. Utiliza el protocolo de comunicación *Universal Asynchronous Receiver Transmitter* (UART) para enviar y



recibir datos a través del puerto serial [61]. El cual fue utilizado en algunas investigaciones anteriores revisadas en la literatura. También es considerado el **Ublox 7M y Ublox 8M** que ofrecen un mejor desempeño al tener una alta sensibilidad en seguimiento o navegación, mayor frecuencia de actualización y compatibilidad con los sistemas GLONASS, BeiDou y Galileo [62], [63], pero aun costo mayor de compra. Mientras que, para realizar las medidas requeridas del giroscopio, puede utilizarse el dispositivo **Melexis MLX90609-N2** utilizado en el artículo de la literatura revisada [19].

En [17] fue identificado el uso de dispositivos adicionales IMU que es una unidad de medida inercial. Este sensor electrónico permite medir la aceleración, velocidad angular y fuerzas gravitatorias de un objeto. Existen diferentes tipos de IMU a un bajo costo de adquisición encontrando la **MPU-6050**, este sensor es un sistema micro-electromecánico (*Micro Electro-Mechanical Systems - MEMS*) de 6 grados de libertad, el cual a través de sus ejes “X”, “Y” y “Z” permite medir la aceleración y velocidad angular de un sistema u objeto en movimiento. Emplea el puerto y protocolo de comunicación I²C (*Inter Integrated Circuit*) para comunicarse por medio de los pines SDA (*Serial Data*) y SCL (*Serial Clock*) [64]. Otra unidad de medición inercial es la **MPU-9250**, la cual utiliza el mismo tipo de comunicación de la MPU-6050 con la diferencia que aparte de medir aceleración y velocidad angular permite conocer el campo magnético en los tres ejes, y un costo mayor de adquisición [65].

3.4.4. Análisis de variables

En la literatura revisada fueron encontradas 27 variables relacionadas con el proyecto de investigación (ver **Tabla 2**), donde las más comunes fueron: posición (latitud y longitud), velocidad y tiempo (hora y fecha). Para seleccionar las variables para tener en cuenta fueron priorizadas aquellas que tuvieran una estrecha relación con la obtención de la localización en línea y la predicción de ubicación mediante modelos de aprendizaje automático. A continuación, lo encontrado al respecto en los artículos revisados.

Las coordenadas geográficas juegan un papel importante, además de algunos factores adicionales como la velocidad o rotación del vehículo que son usadas para disminuir el error en los cálculos mediante los modelos. En dichos trabajos los datos son tratados para prevenir la incorporación de valores por fuera del rango estimado y para establecer el formato de los valores nulos cuando se presenta ausencia de la señal que brinda la ubicación.

La variable de tiempo es tomada como consecuencia de los datos de salida de las herramientas de predicción y también de los sensores utilizados en diferentes puntos a la hora de hacer seguimiento alternativo o que complementarse con el realizado mediante el sensor de GPS. Debido a que los sensores están localizados dentro de los vehículos y no en las paradas de autobús, el tiempo de llegada puede ser sustituido por hora de llegada del paquete a la plataforma de recolección de datos.

En [19], además de las variables de posición, son tenidas en cuenta la distancia recorrida por el vehículo y la rotación de este, así cuando la señal es interrumpida (cuando no se recibe la información del satélite por parte del *GPS*), el algoritmo trabaja con información adicional relacionada directamente con la ubicación y la trayectoria tomada mejorando la precisión de la predicción.

Es importante tener en cuenta en este punto, que el trabajo de investigación busca realizar un seguimiento continuo del vehículo de transporte público en la ruta que tenga



asignada, **por lo cual es importante garantizar un seguimiento adecuado tanto en los casos en los que no se reciba señal por parte del GPS (proveniente de los satélites), como en los casos en los que el paquete de datos no sea recibido en el equipo de red (Gateway)**. El paquete de datos no es recibido en el *Gateway* debido a que la tecnología de comunicaciones utilizada (*LoRa* o *LoRaWAN* en el caso de la investigación) no permite la llegada del paquete (debido a que la distancia del dispositivo a *Gateway* es demasiado larga, o cuando la línea de vista tiene demasiados obstáculos u obstáculos significativos).

Considerando lo anterior, es importante tener en cuenta las siguientes variables a medir por parte de los dispositivos y sensores:

- Latitud.
- Longitud.
- Velocidad.
- Velocidad angular.
- Distancia recorrida.

3.4.5. Análisis de alternativas de modelos de ML.

Los artículos [17], [18] y [19], usaron datos del sistema inercial y datos de la posición GPS para entrenar un módulo de IA con el objetivo de predecir la ubicación de un vehículo cuando la señal GPS no esté disponible. La pérdida de señal del GPS es un escenario que puede presentarse en el sistema que es objeto de la investigación (Sistema para el seguimiento del servicio de transporte público colectivo en tiempo real). Por esta razón, es necesario que el seguimiento considere dos posibles escenarios, que requerirían de una predicción de la ubicación del vehículo: cuando no haya comunicación a través de la tecnología *LoRa* (entre vehículo y *Gateway*) y cuando no haya comunicación GPS.

Los artículos [18] y [23] usaron el mismo *dataset*, que corresponde a una secuencia de diferentes puntos de detección (con separación media de 100 metros) en las vías de una ciudad intermedia. En el artículo [18] estos datos fueron usados para predecir el siguiente punto de ubicación, teniendo en cuenta el punto de ubicación actual y anteriores, también es tenido en cuenta el tiempo transcurrido entre puntos para ofrecer una predicción del tiempo de llegada a un punto específico. Este modelo podría ser viable, cambiando los puntos de detección (con uso de tecnología *Bluetooth*) por los puntos de ubicación que vaya generando los datos enviados por el GPS; además, usando el tiempo transcurrido desde la ausencia de señal *LoRa* como un dato de entrada del módulo de predicción para conocer el incremento de posición del vehículo. En el artículo [23] el mapa de la ciudad es separado en celdas correspondientes a un área alrededor de los puntos de detección del *dataset* usado. El modelo de ML empleado es una RNN donde las dos entradas de la red son: las celdas en las que se ubica el vehículo y el tráfico en cada modelo es la próxima celda en ser alcanzada. Esta solución tiene como inconveniente la necesidad de datos de tráfico, los cuales no es posible conseguir con los sensores y dispositivos propuestos para el sistema. Además, la predicción por celdas hace menos precisa la ubicación ya que la salida es un área y no un punto específico.

En la **Tabla 3** son presentados los resultados de rendimiento de los modelos de ML utilizados por cada uno de los artículos.





Tabla 3. Rendimiento de los modelos de ML.

Modelos de ML	Artículo	Rendimiento														Comentario
		BLEU-1	BLEU-2	BLEU-3	BLEU-4	METEOR	Acc-1	AA	RMSE	MAE	MAPE	MAD	F1-test	Recall	Accuracy	
MLP Network KF	[17]	----	----	----	----	----	----	----	----	----	----	----	----	----	----	<p>Predicción más precisa del incremento de posición que una ANN normal.</p> <p>Puede construir mejor la relación entre las entradas y salidas.</p>
Hybrid LSTM	[18]	46.2%	28.7%	16.6%	9.2%	----	47.8%	----	1.7 min	----	52.6%	----	----	----	----	Valores para una distancia entre detectores de 100 metros
		53.7%	29.9%	19%	9.4%	----	51.4%	----	----	----	57.2%	----	----	----	----	Valores para una distancia entre detectores de 450 metros
52.1%		36.5%	23.9%	15%	----	54.5%	----	1.3 min	----	41.3%	----	----	----	----	Valores para una distancia entre detectores de 100 metros	
65.3%		50.2%	35.7%	22%	----	70.4%	----	----	----	26.9%	----	----	----	----	Valores para una distancia entre detectores de 450 metros	
EKF	[19]	----	----	----	----	----	----	----	864.47 m	----	----	750.64	----	----	----	Quinto corte del GPS al Norte del modelo cinemático



		---	---	---	---	---	---	---	690.97 m	---	---	609.17	---	---	---	Primer corte del GPS al Este del modelo cinemático
NN/ARIMA		---	---	---	---	---	---	---	76.62 m	---	---	61.79	---	---	---	Quinto corte del GPS al Norte del modelo cinemático. Se presenta una mejora del 91% para el RMSE y MAD respecto al KF extendido en el quinto corte del GPS.
		---	---	---	---	---	---	---	36.70 m	---	---	28.53	---	---	---	Primer corte del GPS al Este del modelo cinemático. Se presenta una mejora del 94% para el RMSE y 95% para el MAD respecto al KF extendido en el primer corte del GPS.
DBLSTM CNN	[20]	---	---	---	---	---	---	---	---	---	---	---	75%	64%	76%	Valores para una cantidad de 256 vehículos. Al superar este valor se reduce la capacidad generalizada del modelo.
Predictive Method of Curved Path Tracking	[21]	---	---	---	---	---	---	---	---	---	---	---	---	---	---	El algoritmo reduce los errores laterales en el seguimiento de trayectorias curvas.



ANN	[22]	----	----	----	----	----	----	----	----	----	----	----	----	----	----	>90%	----
<i>Attention-based Recurrent Neural Network</i>	[23]	85%	82%	81%	80%	60%	----	----	----	----	----	----	----	----	----	----	El modelo propuesto presenta una puntuación mejorada en METEOR.
<i>Generative Adversarial Network-based Vehicle trajectory</i>	[24]	----	----	----	----	----	----	95.1%	1.812	1.447	----	----	----	----	----	Valores correspondientes a la velocidad de tráfico.	
		----	----	----	----	----	----	94.9%	1.474	1.043	----	----	----	----	----	Valores correspondientes al flujo de tráfico.	



Error absoluto medio (Mean Average Error o MAE).

Precisión promedio (Average Accuracy o AA).

Raíz de error cuadrático medio (Root Mean Square Error o RMSE).

Porcentaje de error absoluto medio (Mean Absolute Percent Error o MAPE).

Desviación absoluta media (Mean Absolute Deviation o MAD).

Acc-1: Precisión de la predicción en la primera localización consecutiva.

Accuracy: Precisión.

BLEU (Bilingual Evaluation Understudy) y METEOR (Metric for Evaluation of Translation with Explicit ORdering). Son métricas de evaluación de secuencias utilizadas especialmente en la traducción máquina.

BLEU-n "Es una de las puntuaciones de evaluación más utilizadas en el modelado de secuencias, especialmente en el procesamiento del lenguaje natural. BLEU utiliza una forma modificada de precisión para comparar una secuencia de referencia y una secuencia candidata. N son el número de celdas en la que es dividida la secuencia".

METEOR "Crea en primer lugar un alineamiento entre la secuencia de células candidatas y la secuencia de células de referencia. La alineación es un conjunto de asignaciones entre las celdas más similares".

3.5. Selección de dispositivos, variables y sensores

De acuerdo con la literatura consultada para lograr un adecuado seguimiento y predicción de la ubicación futura de un determinado vehículo es muy importante conocer las coordenadas geográficas del vehículo y su movimiento. Su ubicación y su movimiento serán determinados a través de: la latitud, la longitud, la altitud, la velocidad, la velocidad angular, la distancia recorrida y el tiempo (registrando fecha y hora con minutos y segundos). En todo el trayecto de la ruta del vehículo se plantea capturar las variables mencionadas, para lo cual es requerido sensores de GPS, y una IMU para poder capturar los giros del vehículo y su desplazamiento.

La **Tabla 4** muestra los dispositivos *Hardware* considerados convenientes para utilizarlos en el desarrollo de este proyecto de investigación. Los dos tipos de tarjetas son adecuadas para capturar los datos requeridos (ya que las dos tienen integrado un módulo de comunicación *LoRa*). La ventaja de la tarjeta Cube Cell HTCC AB02S es que ya tiene integrado un GPS. Sin embargo, debido al número limitado de dichas tarjetas (HTCC AB02S), es necesario utilizar las tarjetas ESP32 *LoRa*, con un módulo GPS adicional, ya que de dichas tarjetas solo están disponibles algunas unidades.

Fueron investigados diferentes parámetros de configuración de la tecnología *LoRa* y su protocolo de comunicación *LoRaWAN*, para configurar aspectos de acceso a la red el cual es llevado a cabo a través de dos tipos de registros llamados activación, denominados: activación por aire (*Over-The-Air Activation - OTAA*), los dispositivos finales siempre van a realizar una solicitud de unión a la red. Donde serán asignadas las respectivas claves de sesión (AES128), que serán renovadas cada vez que se apague, reinicie o pierda la conexión. Evitando que sea robada la sesión, clonación del dispositivo y permitiéndole a cada dispositivo moverse en redes diferentes. La activación personalizada (*Activated By Personalisation - ABP*), no es necesario hacer una solicitud de acceso a la red. Debido a que las claves de sesión ya están asignadas de forma manual en el dispositivo. Convirtiéndolo en la forma más insegura del sistema porque las claves pueden ser extraídas y clonar el dispositivo. En caso de que el dispositivo final quiera cambiar de red, debe cambiar las credenciales de forma manual [66], [67]. Los dos tipos de activación mencionados mejoran las características de eficiencia y seguridad en la red.



Adicionalmente para capturar las variables se propone utilizar el sensor GPS (Ublox 6M) el cual será conectado a la ESP32 LoRa en conjunto con una Unidad de Medida Inercial IMU (MPU 6050) conformada por un giroscopio y acelerómetro triaxial. Es posible que con dicho dispositivo también pueda medirse el desplazamiento del vehículo (lo cual realizaría el odómetro). A la vez esta IMU será conectada a la HTTC AB02S. Al tener conectados el GPS y la IMU se espera obtener las coordenadas geográficas y variables de movimiento de los vehículos, Mejorando la precisión al predecir la ubicación vehicular. Son considerados el GPS Ublox 6M y la IMU MPU-6050 por tener un costo de adquisición menor en comparación con otras referencias mencionadas en la **subsección 3.4.3.**

De acuerdo con lo anterior las variables que fueron seleccionadas para el trabajo a realizar son las siguientes:

- Latitud.
- Longitud.
- Velocidad.
- Velocidad angular
- Desplazamiento
- Fecha y hora del registro.

Dichas variables serán capturadas por los siguientes sensores:

- GPS
- IMU

En la **Tabla 4** son presentados los dispositivos que permiten leer las variables requeridas en el proyecto, la complejidad de uso de los dispositivos usados, el tipo de comunicación y el costo.

Tabla 4. Dispositivos seleccionados.

Dispositivo y sensores	Características (Features)						Complejidad	Tipo de comunicación	Costo
	Vel	Lat.	Long	Tiempo	Rot	Dist.			
ESP32 LoRa + GPS + IMU	✓	✓	✓	✓	✓	✓	Baja	LoRa	Medio
Cube Cell HTCC – AB02S (LoRa + GPS) + IMU	✓	✓	✓	✓	✓	✓	Media-alta	LoRa	Medio-alto

Vel: velocidad, Rot: tasa de rotación, Tiempo: tiempo de llegada, Lat: latitud, Lon: longitud, Dist: distancia.



Capítulo 4

4. Selección de modelo de ML para predicción de la ubicación y pruebas iniciales

4.1. Selección del modelo de ML.

De acuerdo con la **Tabla 3**, algunos modelos de ML no fueron tenidos en cuenta al no proporcionar los resultados las métricas de desempeño de las pruebas realizadas. Por lo tanto, no fue posible comparar todos los modelos de la tabla. Con base en lo anterior fueron considerados: la red neuronal artificial con un valor de la métrica *Accuracy* (precisión) mayor al 90%. La red neuronal convolucional y la red DBLSTM obtuvieron un valor de *Accuracy* de 76%. Mientras que, las redes las redes neuronales LSTM secuenciales presentaron una *ACC-1* (precisión de la predicción en la primera localización consecutiva) del 70.4%, y el LSTM híbrido alcanzó un 51.4%.

Un aspecto tenido en cuenta fue la viabilidad en la implementación de los modelos, para ello el lenguaje de programación Python cuenta con la librería Keras que permite implementar una gran variedad de modelos de redes neuronales como lo son: CNN, NN y RNN. Dado que las redes neuronales LSTM son una variación de las redes RNN es posible implementar una red LSTM en Python con un grado reducido de complejidad.

Otro aspecto importante considerado fue que el modelo hubiera sido implementado en un experimento lo más cercano al contexto de aplicación de la investigación. Los experimentos realizados con LSTM secuencial y DBLSTM eran los más cercanos al contexto de aplicación requerido.

Además, fue considerado que las redes LSTM pueden extraer y transmitir información relevante automáticamente a lo largo de una larga cadena de secuencias para realizar predicciones [18], resulta adecuada para el aprendizaje de patrones de movimiento secuencial en datos de trayectorias urbanas. La estructura de un modelo LSTM es presentada en la **Figura 4**, contiene una capa de entrada, una capa oculta y una capa de salida. La capa de entrada inicializa los datos de entrada para las capas posteriores, mientras que la capa de salida realiza la tarea de predicción (clasificación o regresión).

La capa oculta es capaz de aprender características de la secuencia de entrada a través de una unidad recurrente denominada bloque de memoria. A diferencia de una red neuronal artificial las cuales no tienen ningún tipo de memoria. El bloque de memoria de la capa oculta contiene una celda, que transfiere información relevante a lo largo de la secuencia, y tres puertas: la puerta del olvido, la puerta de entrada y la puerta de salida, que regulan el flujo de información. El primer paso en un bloque de memoria LSTM es decidir qué información conservar u olvidar a partir de la salida del bloque de memoria anterior, $ht-1$, y la entrada del bloque de memoria actual, Xt . Después, la puerta de entrada captura la información de los datos de entrada Xt y actualiza el estado de la celda. Por último, la puerta de salida decide qué parte del estado de la celda se utiliza para calcular la salida final, ht .

A diferencia de las redes LSTM, en las redes DBLSTM la entrada fluye en ambas direcciones de la red, y es capaz de utilizar la información de ambos lados.

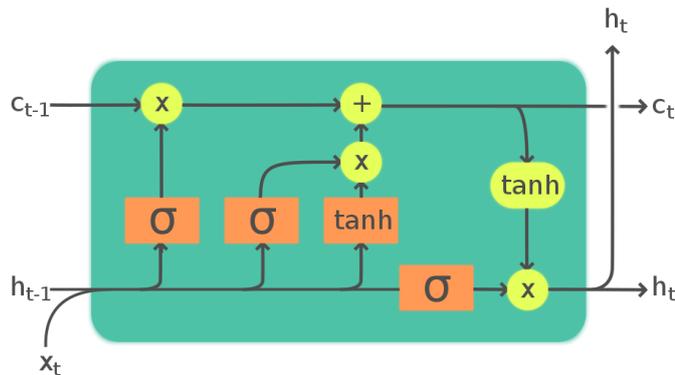


Figura 4. Diagrama esquemático del bloque de memoria LSTM [68].

Con base en los valores de desempeño del 70.4% y 76% de los anteriores modelos y la viabilidad de implementación, se consideró conveniente utilizar LSTM secuencial y DBLSTM en el desarrollo de la predicción de la ubicación de los vehículos, en este proyecto.

4.2. Pruebas iniciales de los modelos de ML

Con estas pruebas iniciales permitió evaluar la precisión del modelo seleccionado haciendo uso de la librería Keras y pandas del lenguaje de programación Python. La comprobación fue realizada con una fracción del *dataset* (186 de 2.500 registros) seleccionado para esta prueba inicial, prediciendo cada uno de los puntos de ubicación de la fracción del *dataset* y obteniendo el error mediante la distancia de separación y el tiempo entre puntos (reales y predichos).

Las condiciones de precisión del modelo fueron realizadas en dos partes. La primera parte consistió en predecir puntos de ubicación a partir de un *dataset* seleccionado. La segunda parte consistió en verificar si el modelo funcionaba en tiempo real, al predecir puntos geográficos tomando como entradas valores predichos.

Las pruebas iniciales fueron ejecutadas en dos partes, primero fue implementada una red LSTM secuencial con el objetivo de verificar su adecuado funcionamiento de predicción de la ubicación (latitud y longitud) y del tiempo a partir de ciertos *dataset*. Aquí fue utilizado un *dataset* generado en un trabajo previo (consultado en una plataforma en línea de *datasets* denominado Kaggle [69]), para realizar predicción de localización (latitud, longitud) y tiempo, fueron utilizados datos históricos (solamente registros de GPS y tiempo) como entrada. Luego para la segunda parte de las pruebas iniciales fue utilizado un *dataset* generado mediante el recorrido de un vehículo y el envío de información utilizando comunicación LoRa y su protocolo de red LoRaWAN. Esto con el fin de realizar predicción de localización (latitud, longitud) y tiempo, pero utilizando como entrada de información los datos de GPS, registro de tiempo, y además los datos tomados de un dispositivo IMU.

En el **Anexo B. Pruebas iniciales de los modelos de ML** son presentados los resultados sobre la precisión del modelo de predicción aplicado sobre dos *datasets* diferentes. El primer *dataset* consta de datos extraídos de un base de datos en la plataforma Kaggle, que contiene datos de posición obtenidos mediante GPS en la

ciudad de Rio de Janeiro, Brasil. El segundo *dataset* fue construido por los autores del presente documento con datos de posición y datos inerciales recolectados en una pequeña ruta en la ciudad de Popayán, Colombia.

Las **Figuras 5 y 6** representan los resultados de predecir la ubicación (latitud y longitud) del vehículo a partir de datos históricos sobre el dataset con datos de Río de Janeiro, validando así el adecuado comportamiento del modelo de ML empleado.

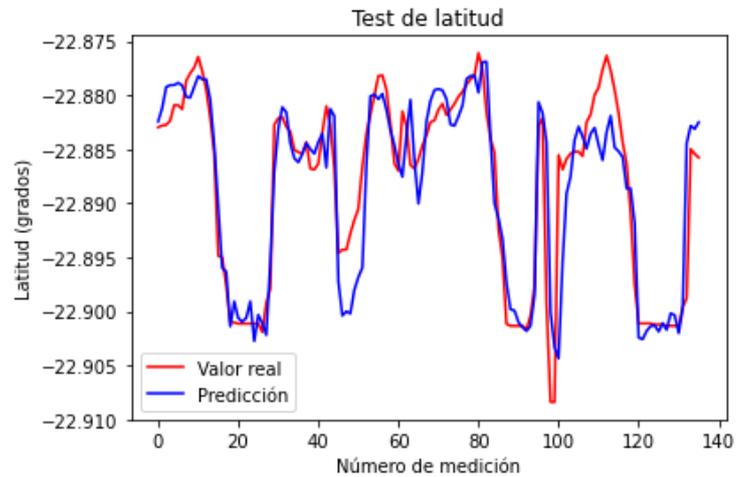


Figura 5. Predicción de latitud a partir de datos históricos para una ruta en Rio de Janeiro.

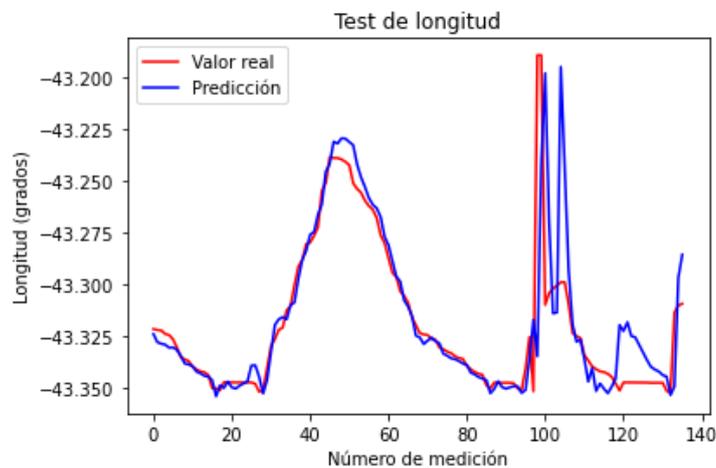


Figura 6. Predicción de longitud a partir de datos históricos para una ruta en Río de Janeiro.

Las **Figuras 7 y 8** representan los resultados de la predicción de la ubicación (latitud y longitud) a partir de valores de ubicación predichos por el algoritmo a partir de datos de posición de un bus en Río de Janeiro, debido a que son necesarios 50 valores de posición para empezar a predecir la siguiente posición, esos valores iniciales son tomados con el mismo valor. Para este caso la **Figura 9** es el resultado de predecir el tiempo de llegada a la futura ubicación.

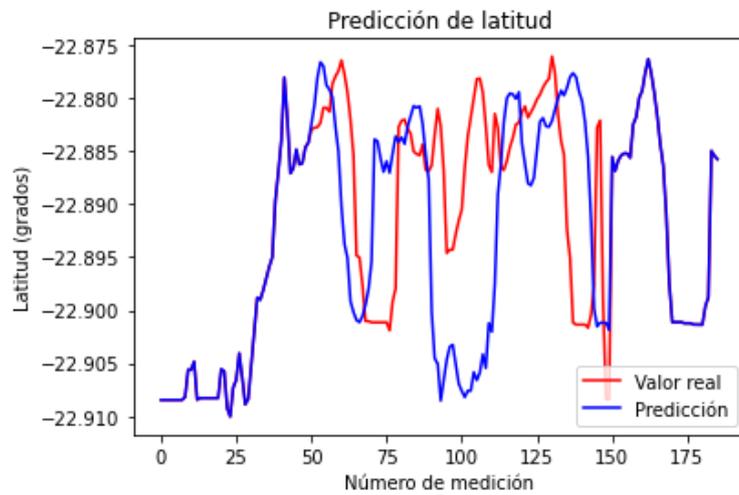


Figura 7. Predicción de latitud en tiempo real, a partir de datos predichos para una ruta en Rio de Janeiro.

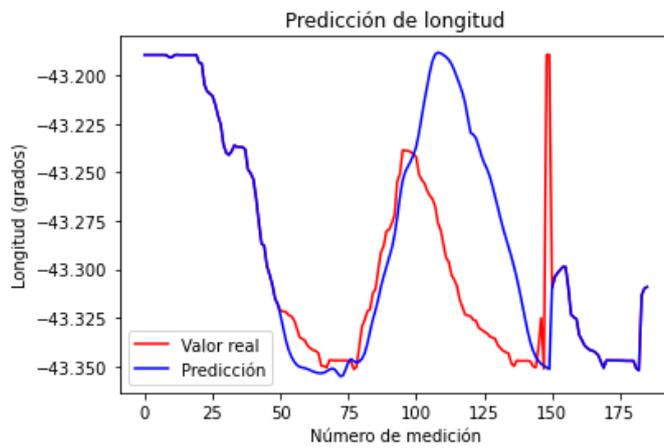


Figura 8. Predicción de longitud en tiempo real, a partir de datos predichos para una ruta en Rio de Janeiro.

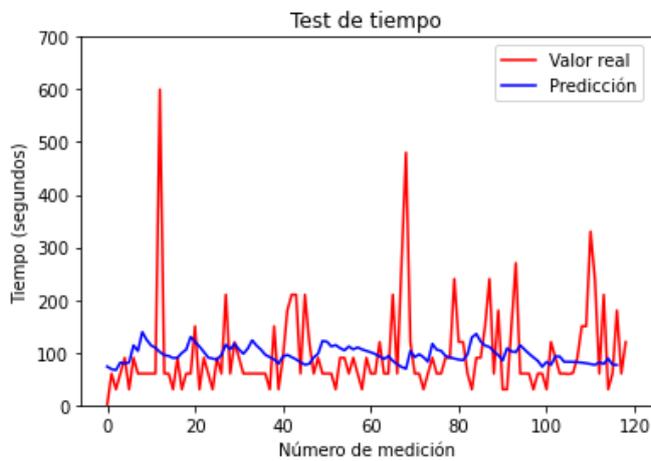


Figura 9. Predicción del tiempo a partir de datos históricos para una ruta en Rio de Janeiro.



Las **Figuras 10 y 11** representan los resultados de predecir la ubicación (latitud y longitud respectivamente) del vehículo a partir de datos históricos sobre los datos recolectados en la ciudad de Popayán. Para este caso la **Figura 12** es el resultado de predecir el tiempo de llegada a la futura ubicación

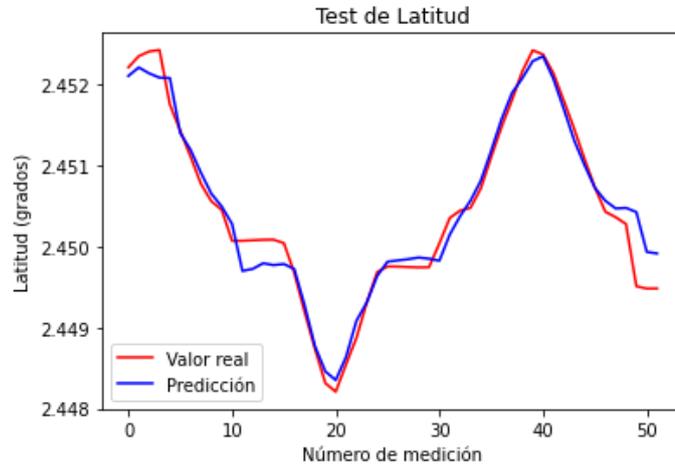


Figura 10. Predicción de latitud a partir de datos históricos para una ruta en Popayán.

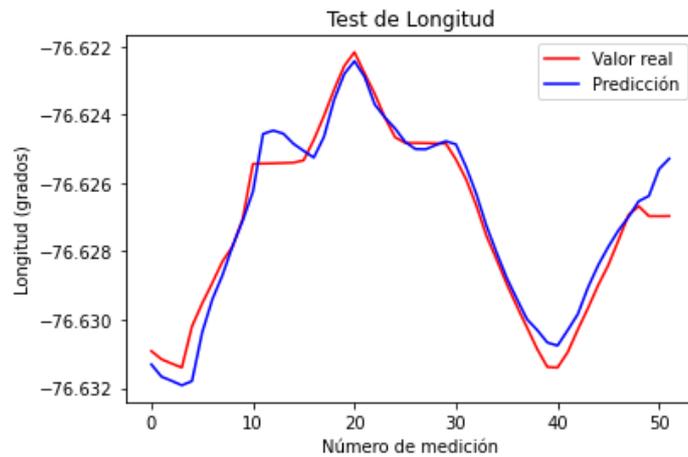


Figura 11. Predicción de longitud a partir de datos históricos para una ruta en Popayán.

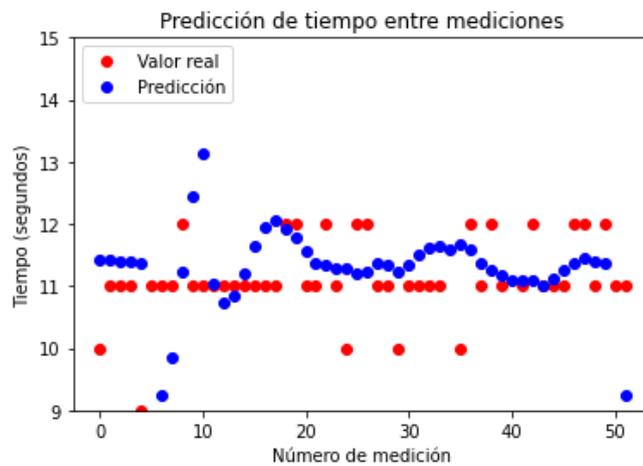


Figura 12. Predicción del tiempo a partir de datos históricos para una ruta en Popayán.



Las **Figuras 13 y 14** representan los resultados de la predicción de la ubicación (latitud y longitud) a partir de valores de ubicación predichos por el algoritmo a partir de datos de posición recolectados en la ciudad de Popayán, debido a que son necesarios 30 valores de posición para empezar a predecir la siguiente posición, esos valores iniciales son tomados con el mismo valor. Para este caso la **Figura 15** es el resultado de predecir el tiempo de llegada a la futura ubicación.

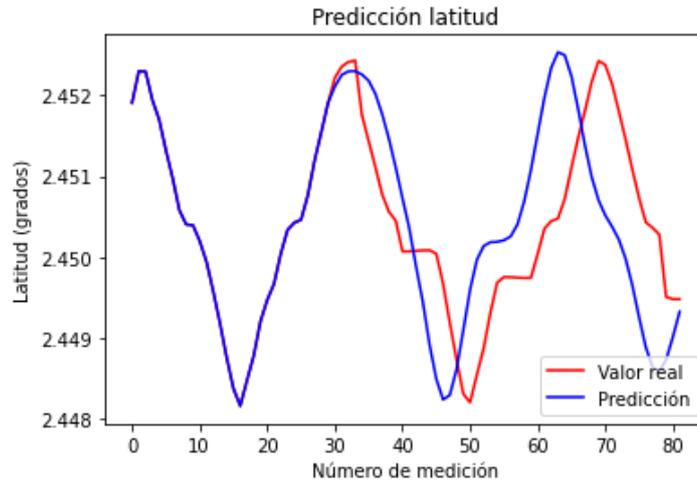


Figura 13. Predicción de latitud a partir de datos predichos para una ruta en Popayán (en el escenario 1).

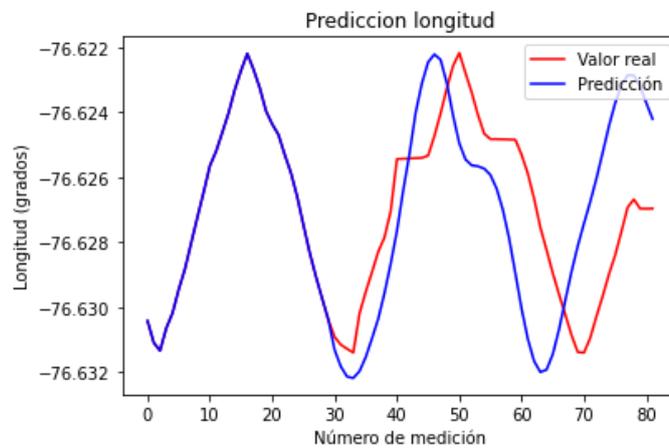


Figura 14. Predicción de longitud a partir de datos predichos para una ruta en Popayán (en el escenario 1).

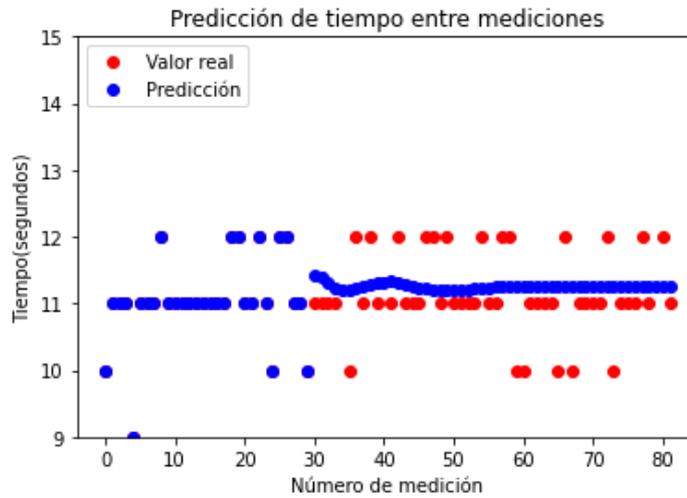


Figura 15. Predicción del tiempo a partir de datos predichos para una ruta en Popayán, (en el escenario 1). (Ampliada)

Los algoritmos generados en las pruebas iniciales son presentados en el **Anexo C. Algoritmos generados de las pruebas iniciales** del presente documento. Estos algoritmos fueron ajustados para las pruebas finales con el fin de que el sistema general funcionara en tiempo real, la versión final es presentada en el **Anexo D. Algoritmos generados en las pruebas controladas y no controladas**.

Capítulo 5

5. Diseño y Desarrollo del prototipo del sistema

5.1. Diseño del prototipo del sistema

El diseño de prototipo del sistema fue realizado con base en la identificación y selección de las variables, parámetros y dispositivos a utilizar en este trabajo, permitiendo definir los componentes *hardware*, *software* y de comunicaciones que debía manejar el sistema. El diseño del prototipo se compone de tres módulos, que son: recolección de datos, recolector de los datos y procesamiento de la información. La **Figura 16** presenta cada uno de los módulos que conforman el prototipo del sistema y el flujo de la información.

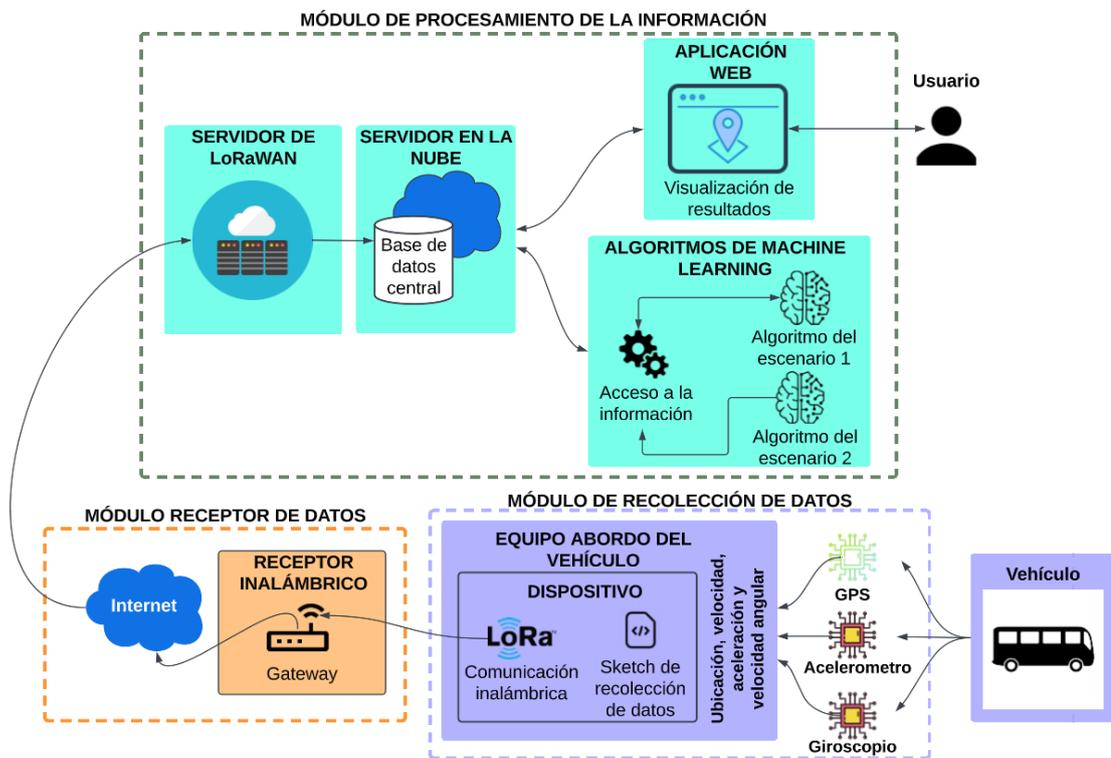


Figura 16. Diseño del prototipo propuesto para el sistema.

La **Figura 16** describe el prototipo del sistema, este diagrama está basado en una arquitectura de Sistemas Inteligentes de Transporte (*Intelligent Transportation Systems, ITS*), específicamente en la arquitectura ARC-IT o Arquitectura Americana [70], que es gestionada por el Departamento de Transporte de Estados Unidos. Fue considerado conveniente utilizar este tipo de notación, para facilitar la comprensión del prototipo propuesto.

A continuación, cada uno de los 3 módulos presentados en la **Figura 16**.

5.1.1. Módulo de recolección de datos

Para realizar el seguimiento vehicular en tiempo real es necesario tener un equipo a bordo del vehículo (EBV) conformado por un dispositivo electrónico con elementos

hardware y *software* encargados de capturar los datos del vehículo. Este dispositivo debe incluir un sensor *GPS* encargado de obtener los datos de ubicación y velocidad; un acelerómetro para medir la aceleración; un giroscopio para detectar la velocidad angular o giro; y un dispositivo que permita la comunicación *LoRa*, para realizar el envío de los datos medidos en el vehículo hacia un receptor ubicado a varios metros de distancia a través de radiofrecuencia. La gestión de captura y envío de los valores se debe realizar a través de una tarjeta micro controlada, encargada de controlar cada una de estas acciones por medio de un código de programación. La **Figura 17** representa el dispositivo al interior del vehículo.

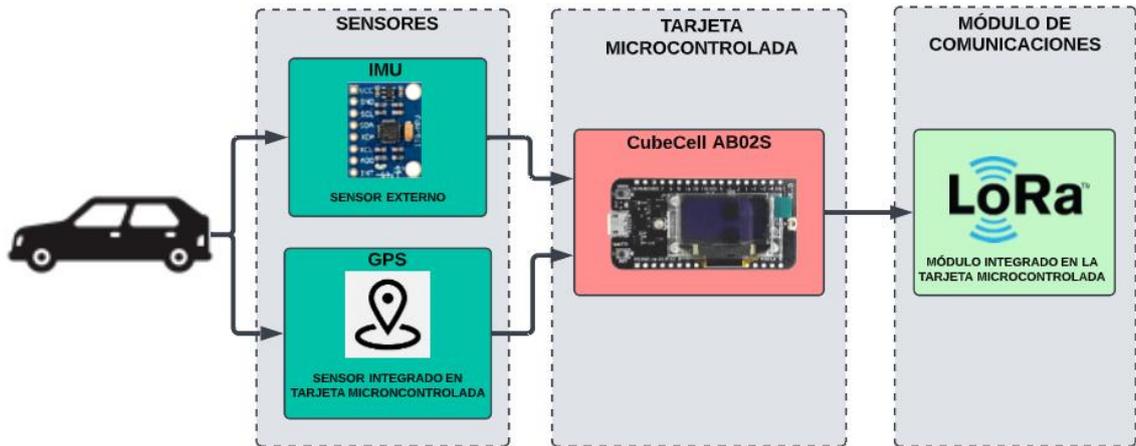


Figura 17. Diagrama de bloques del dispositivo para recolectar los datos.

5.1.2. Módulo receptor de datos

Este módulo debe incluir un componente hardware que utilice comunicación *LoRa* y el protocolo de red *LoRaWAN*, para recibir los datos provenientes del módulo de recolección de datos. Este módulo además de recibir la información recolectada debe de retransmitir los datos recibidos (de los datos del vehículo) hacia un servidor de la nube para su posterior almacenamiento, para realizar esta tarea el *hardware* de este módulo debe de estar en la capacidad de establecer una conexión IP.

Como se menciona en la **Sección 3**, hay dos formas de activación para que los dispositivos finales puedan acceder a la red *LoRaWAN* los cuales son OTAA y ABP. El tipo de activación elegido para el sistema fue OTAA debido a que este brinda mayor seguridad frente a ABP. Con el propósito de cumplir con uno de los objetivos específicos definidos en el trabajo de investigación referente a la seguridad y eficiencia.

5.1.3. Módulo de procesamiento de la información.

El módulo está conformado por un servidor de red *LoRaWAN*, un servidor en la nube, los modelos de ML y una aplicación web.

El servidor de red *LoRaWAN* permite recibir los paquetes de datos provenientes del módulo receptor de datos y reenviarlos a un servidor en la nube.

El servidor en la nube tiene la función de almacenarlos en una base de datos. Dicha base de datos permite almacenar un registro histórico de datos para ser accedido en cualquier instante.

El componente de modelos de ML debe tener la capacidad de acceder a la información alojada en la base de datos, utilizando dicha fuente como entrada para realizar predicciones sobre los datos de ubicación y tiempo de un vehículo en una ruta. Estos modelos tienen la tarea de predecir la ubicación del vehículo bajo dos escenarios establecidos. El escenario 1 se presenta cuando no existe comunicación mediante la tecnología LoRa (entre el módulo de recolección y el módulo receptor). En este caso si hay o no información del GPS o de la IMU es indiferente debido a que dicha información no puede transmitirse, ni almacenar en una base de datos. El escenario 2 se presenta cuando hay conexión LoRa, pero no existe comunicación del GPS. En este caso no es posible recibir las mediciones de ubicación, pero sí de la IMU. Las **Figuras 18 y 19** detallan estos dos escenarios (la **Figura 18** presenta el caso en que no llega el paquete en el Gateway y la **Figura 19** presenta el caso en que no hay de señal GPS).

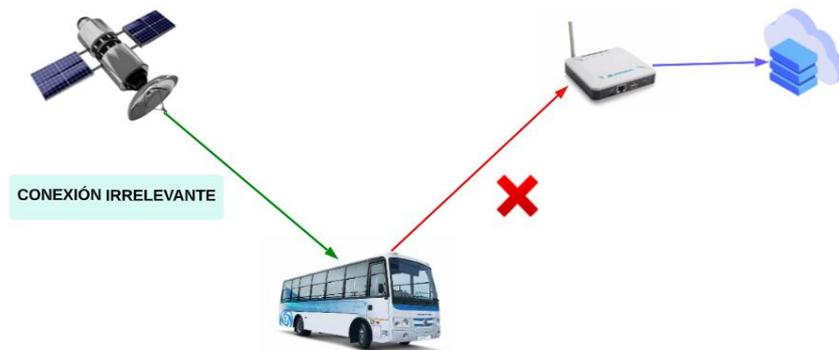


Figura 18. Escenario 1. Sistema sin comunicación LoRa, ausencia de paquetes en el Gateway (es indiferente si hay o no conexión GPS).

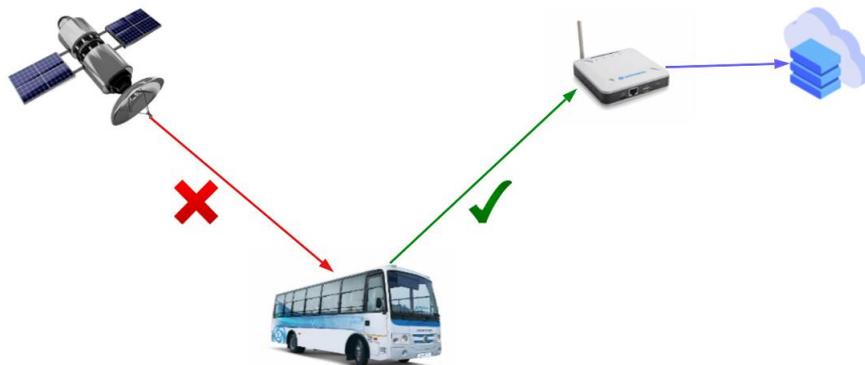


Figura 19. Escenario 2. Sistema sin señal GPS y comunicación LoRa activa.



La aplicación web de este módulo permite que un usuario visualice los resultados obtenidos del seguimiento del vehículo (incluidos los segmentos donde sea requerida la predicción por parte de los modelos).

5.2. Diseño de la estructura del conjunto de datos.

A partir de la identificación de variables, análisis de alternativas y selección de estas, fueron establecidas las variables que conforman la estructura del conjunto de datos para llevar a cabo el desarrollo del prototipo del sistema.

El módulo de recolección de datos debe tener asociado un *dataset* en el cual son almacenados los diferentes datos transmitidos por el dispositivo a bordo del vehículo. A la vez se tendrá un *dataset* general, en la cual son almacenados algunos de los datos recolectados por el dispositivo a bordo y los datos obtenidos de la salida de los modelos de ML. Esto con el objetivo de tener una tabla donde sea posible revisar la ubicación en cualquier instante del vehículo (ya sea una ubicación real enviado por el dispositivo a bordo o la ubicación predicha por el modelo).

Las estructuras de cada uno los *dataset* son presentadas en las **Tablas 5 y 6**. En las tablas son especificados el nombre, la descripción y el tipo de dato de cada una de las columnas.

5.2.1. Tabla de datos del dispositivo a bordo.

En la **Tabla 5** son presentadas las columnas que contienen las variables de la tabla de datos del dispositivo a bordo. Estos datos son detectados a través de los sensores GPS, acelerómetro y giroscopio, conectados a la tarjeta micro controlada. Además de los datos capturados por los sensores también deben ser almacenados los parámetros de comunicación de la tecnología (*LoRa*) que permite la transmisión de la información de forma inalámbrica. El intervalo de recepción de los datos es de 9 a 13 segundos, que corresponde al tiempo de propagación de la señal desde el módulo de recolección de datos hasta el módulo receptor, más el tiempo de procesamiento, porque los datos pasan por varios servidores (servidor de *LoRaWAN* y servidor en la nube), antes de ser almacenados en la base de datos. Por lo cual el registro de tiempo debe tener hora, minuto y segundo.

Tabla 5. Datos enviados por el dispositivo a bordo.

Nombre de la columna (variable)	Descripción	Tipo
ID dato (<i>id</i>)	Indica el orden en que los datos son guardados (auto incremental).	Numérico
ID dispositivo (<i>dev_id</i>)	Nombre del dispositivo que ha enviado el dato.	Texto
Fecha (<i>date</i>)	La fecha en que el dato es enviado (aaaa-mm-dd).	Fecha
Tiempo (<i>hour</i>)	La hora en la que el dato es enviado (hh:mm:ss).	<i>Timedelta64</i>
Latitud (<i>latitude</i>)	Coordenada geográfica del <i>GPS</i> (6 decimales mínimo).	Numérico
Longitud (<i>longitude</i>)	Coordenada geográfica del <i>GPS</i> (6 decimales mínimo).	Numérico



Velocidad (<i>speed</i>)	La velocidad del vehículo en un instante de tiempo (Km/h).	Numérico
Aceleración X (<i>accx</i>)	Aceleración del vehículo en el eje X (m/s^2).	Numérico
Aceleración Y (<i>accy</i>)	Aceleración del vehículo en el eje Y (m/s^2).	Numérico
Aceleración Z (<i>accz</i>)	Aceleración del vehículo en el eje Z (m/s^2).	Numérico
Velocidad angular X (<i>gyrox</i>)	Velocidad de giro del vehículo sobre su eje X ($^{\circ}/s$).	Numérico
Velocidad angular Y (<i>gyroy</i>)	Velocidad de giro del vehículo sobre su eje Y ($^{\circ}/s$).	Numérico
Velocidad angular Z (<i>gyroz</i>)	Velocidad de giro del vehículo sobre su eje Z ($^{\circ}/s$).	Numérico
Spread factor (<i>sf</i>)	Factor de ensanchamiento o esparcimiento. (Rango de 7 a 12)	Numérico
Ancho de banda (<i>bw</i>)	Rango de frecuencia utilizada para transportar la señal (Hz).	Numérico
Tasa de codificación de datos (<i>cr</i>)	Relación entre número de bits y número de bits totales.	Numérico

5.2.2. Tabla general de datos

A continuación, la **Tabla 6** presenta las variables que componen la tabla general. Los campos por almacenar corresponden a los mencionados en la **Tabla 6** y adicionalmente, los campos relacionados con las predicciones realizadas por los modelos.

Tabla 6. Tabla general de datos.

Nombre de la columna (variable)	Descripción	Tipo
ID dato (<i>id</i>)	Indica el orden en que los datos son guardados (auto incremental).	Numérico
ID dispositivo (<i>dev_id</i>)	Nombre del dispositivo que ha enviado el dato.	Texto
Fecha (<i>date</i>)	La fecha en que el dato es enviado (aaaa-mm-dd).	Fecha
Tiempo (<i>hour</i>)	La hora en la que el dato es enviado (hh:mm:ss).	<i>Timedelta64</i>
Tiempo predicho (<i>predicted_hour</i>)	Tiempo predicho por el modelo(hh:mm:ss).	<i>Timedelta64</i>
Tipo de registro (<i>type_record</i>)	Define el tipo de registro realizado a través de un valor numérico: 0: Indica que es un registro proveniente del dispositivo a bordo.	Numérico



	<p>1: Indica que es un registro predicho porque no hubo dato de <i>GPS</i>.</p> <p>2: indica que es un registro predicho porque no hubo conexión <i>LoRa</i>.</p>	
Latitud (<i>latitude</i>)	Coordenada geográfica del <i>GPS</i> (6 decimales mínimo).	Numérico
Longitud (<i>longitude</i>)	Coordenada geográfica del <i>GPS</i> (6 decimales mínimo).	Numérico
Latitud predicha (<i>predicted_latitude</i>)	Latitud predicha por el modelo (° decimales).	Numérico
Longitud predicha (<i>predicted_longitude</i>)	Longitud predicha por el modelo (° decimales).	Numérico
Velocidad (<i>speed</i>)	La velocidad del vehículo en un instante de tiempo (Km/h).	Numérico
Aceleración X (<i>accx</i>)	Aceleración del vehículo en el eje X (m/s^2).	Numérico
Aceleración Y (<i>accy</i>)	Aceleración del vehículo en el eje Y (m/s^2).	Numérico
Aceleración Z (<i>accz</i>)	Aceleración del vehículo en el eje Z (m/s^2).	Numérico
Velocidad angular X (<i>gyrox</i>)	Velocidad de giro del vehículo sobre su eje X ($^{\circ}/s$).	Numérico
Velocidad angular Y (<i>gyroy</i>)	Velocidad de giro del vehículo sobre su eje Y ($^{\circ}/s$).	Numérico
Velocidad angular Z (<i>gyroz</i>)	Velocidad de giro del vehículo sobre su eje Z ($^{\circ}/s$).	Numérico
Spread factor (<i>sf</i>)	Factor de ensanchamiento o esparcimiento. (Rango de 7 a 12)	Numérico
Ancho de banda (<i>bw</i>)	Rango de frecuencia utilizada para transportar la señal (Hz).	Numérico
Tasa de codificación de datos (<i>cr</i>)	Relación entre número de bits y número de bits totales.	Numérico

5.3. Desarrollo e implementación del prototipo del sistema

En la **sección 5.1** es presentado el diseño del prototipo del sistema, el cual se compone de diferentes módulos con una función determinada (ver Figura 7). A su vez, cada uno de estos módulos tiene diferentes componentes que permiten que el sistema opere de una manera adecuada. Esta sección especifica cada uno de los dispositivos electrónicos, tecnologías y herramientas seleccionados para implementar el prototipo a partir del diseño realizado.

5.3.1. Desarrollo del módulo de recolección de datos

Para implementar este módulo se determinó desarrollar un dispositivo, a partir de una tarjeta micro controlada, con ciertos módulos integrados y algunos componentes adicionales. A continuación, es descrito en detalle este dispositivo.



5.3.1.1. Dispositivo

Previamente, fueron analizadas y seleccionadas las diferentes alternativas de dispositivos, sensores, variables y la tecnología de comunicación inalámbrica para realizar el seguimiento vehicular. Con base en lo anterior el dispositivo está conformado por una tarjeta micro controlada Cube Cell HTCC-AB02S del fabricante Heltec.

Debido a la baja disponibilidad de la tarjeta Cube Cell HTCC-AB02S (dos tarjetas adquiridas por los estudiantes y una por el director del trabajo), al momento de desarrollar este trabajo. Fue requerida una segunda opción de tarjeta similar, de la que existiera una mayor disponibilidad. La otra tarjeta seleccionada fue la ESP32 LoRa, también del fabricante Heltec. Esta tarjeta tiene integrado el módulo LoRa, un módulo de comunicación Wi-Fi y Bluetooth, pero no posee el dispositivo GPS que tiene la tarjeta Cube Cell.

El número de dispositivos utilizados fue 4, debido a la limitación de recursos tanto de tarjetas, como de vehículos, y personal para conducirlos.

Los módulos integrados LoRa de las tarjetas seleccionadas debían trabajar bajo la banda de frecuencias no licenciada de 915 MHz. Fue elegida esta frecuencia de operación porque es la asignada para el continente americano por parte de los entes de estandarización del protocolo.

Sensores adicionales para el dispositivo

Además de obtener la ubicación (latitud y longitud) del vehículo a través del sensor GPS, también es requerido su aceleración y velocidad angular, el sensor seleccionado para esta tarea es una IMU. La **subsección 3.4.3** del estado del arte describe el GPS Ublox-6 M y la IMU MPU-6050 que fueron conectados a las tarjetas.

Ambiente de desarrollo

Las herramientas utilizadas durante el desarrollo del dispositivo fueron las siguientes:

- Arduino IDE. Entorno de desarrollo de integrado de Arduino, el cual permite editar, cargar, y compilar el código (*sketch*) [71].
- CubeCell Development Framework. Proporciona la compatibilidad entre las placas Heltec Cube Cell y el IDE de Arduino. Posee las librerías para controlar la transmisión y recepción del módulo LoRa y el protocolo LoRaWAN, y extraer las mediciones del GPS de la tarjeta HTCC-AB02S [72].
- ESP32_LoRaWAN. Posibilita el control de la recepción y transmisión del módulo integrado LoRa y el protocolo de comunicación LoRaWAN de la tarjeta ESP32 LoRa [73].
- Heltec ESP32 Series Dev-boards: Proporciona el adecuado funcionamiento de las placas Heltec ESP 32 LoRa con el IDE de Arduino [74].
- MPU6050_light. Permite obtener de la IMU MPU-6050 la aceleración y la velocidad angular de un objeto en movimiento y calibración del sensor [75].
- TinyGPSPlus. Librería con la función de extraer del módulo GPS Ublox-6M los valores ubicación (latitud y longitud), altitud y velocidad [76].

Implementación del dispositivo

A través del IDE de Arduino fue realizada la integración de las librerías, para que los elementos electrónicos funcionaran como una sola unidad, definiendo la frecuencia de operación de 915 MHz, el intervalo de transmisión de los datos fue de 3000 milisegundos (debido a que fue el tiempo mínimo en que se logró operar) y el método de autenticación *LoRaWAN* de activación por OTAA. A parte de esta integración fueron realizadas las respectivas conexiones entre sensores y tarjetas micro controladas, descritas a continuación:

- La MPU-6050 fue conectada con la tarjeta Cube Cell HTCC-AB02S por medio del puerto de comunicación I²C, los pines utilizados para realizar esta conexión fueron: SCL (pin físico 39) y SDA (pin físico 40). La fuente de alimentación para el sensor fue obtenida del pin VIN y GND que proporcionan una salida fija de 5 Voltios. A la vez fueron conectadas las respectivas antenas para la comunicación *LoRa* y *GPS*. La **Figura 20** presenta el dispositivo (incluyendo el módulo IMU) y la **Figura 21** presenta el diagrama esquemático.

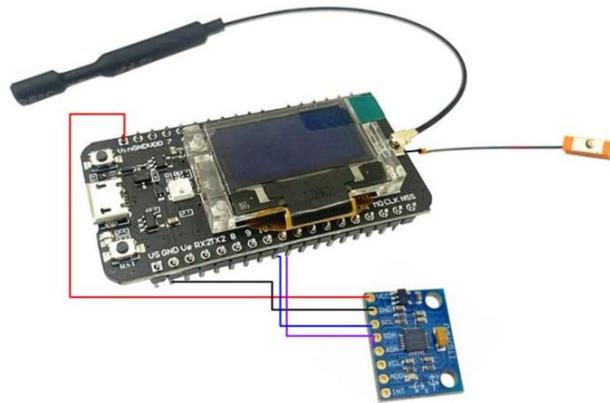


Figura 20. Dispositivo recolector de datos (Cube Cell HTCC-AB02S, MPU-6050).

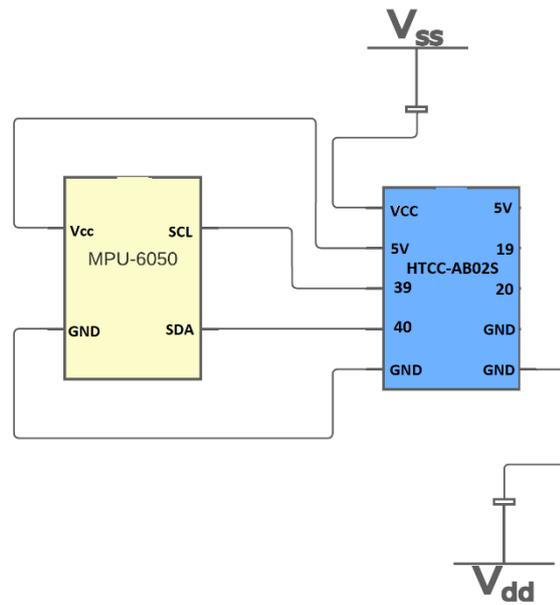


Figura 21. Diagrama esquemático del dispositivo recolector de datos (Cube Cell HTCC-AB02S, MPU-6050).

- Las conexiones entre la ESP32 LoRa y los sensores GPS e IMU son descritas a continuación. El módulo GPS Ublox-6M fue conectado a través de los pines U2_TXD (GPIO17) y U2_RXD (GPIO16) para establecer la comunicación UART. La alimentación de 3.3 Voltios para modulo fue tomada de los pines 3V3 y GND. La conexión de la MPU-6050 se realizó en los pines SCL (GPIO22) y SDA (GPIO21), los pines 5V y GND proporcionaron los 5 Voltios para el sensor. También fue realizada la conexión de antenas para la comunicación LoRa y GPS. El dispositivo (incluyendo el módulo GPS y el módulo IMU) es presentado en **Figura 22** y el diagrama esquemático en la **Figura 23**.

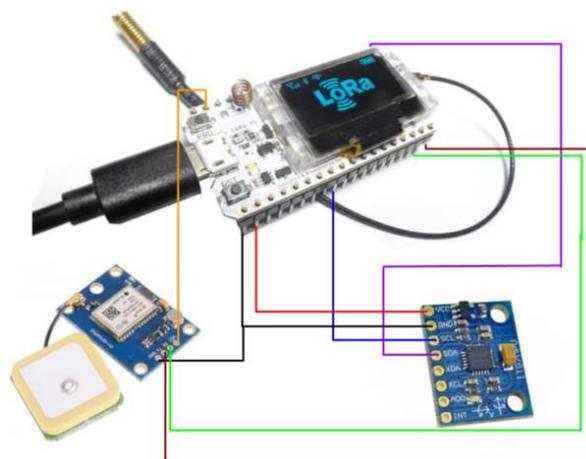


Figura 22. Dispositivo recolector de datos (ESP32 LoRa, GPS Ublox 6M, MPU-6050).

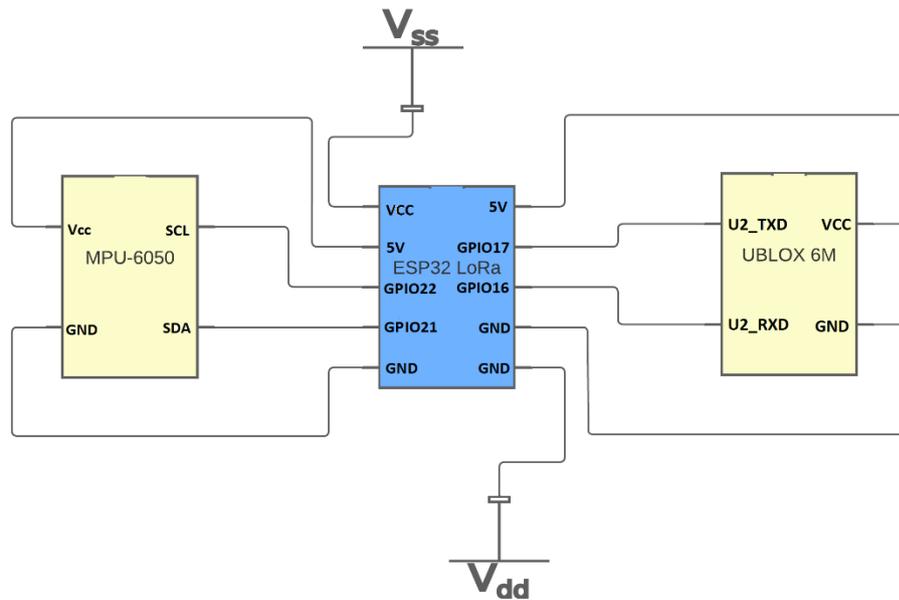


Figura 23. Diagrama esquemático del dispositivo recolector de datos (ESP32 *LoRa*, GPS Ublox 6M, MPU-6050).

5.3.2. Desarrollo del módulo receptor de datos

En la identificación de variables y dispositivos relevantes realizada previamente, fue determinado que el dispositivo encargado de recibir la información proveniente del módulo de recolección de datos era un *Gateway LoRaWAN*. Con base en lo anterior, la disponibilidad de dispositivos, y los equipos disponibles en el mercado, fue elegido el *Gateway LPS8* del fabricante Dragino.

El LPS8 (**Figura 24**) es un *Gateway* que sirve como puerta de enlace para unir una red inalámbrica *LoRa* (utilizando el protocolo *LoRaWAN*) y una red IP a través de Wi-Fi o *Ethernet*, con la funcionalidad de reenviar paquetes. En su interior tiene un chip *LoRa* SX1308, que permite programar 10 formas de demodulación paralelas y opera en las frecuencias de 915, 920 y 923 MHz [77].



Figura 24. *Gateway LoRaWAN* LPS8. Fuente:[77]



La configuración del manual del dispositivo [78], permitió poner en funcionamiento el *Gateway*, donde fue habilitado el plan la frecuencia de 915 MHz y la sub-banda numero 2 conformada por el rango de frecuencias de 903.5 a 905.3 MHz que utiliza los canales del 8 al 15. Esta configuración fue uno de los aspectos considerados, para lograr una mayor eficiencia en el desempeño del sistema, respecto a la comunicación a través de *LoRaWAN*.

5.3.3. Desarrollo del módulo de procesamiento de la información

Previamente, fueron identificados diferentes alternativas de componentes, para desarrollar las respectivas funcionalidades del módulo. En las siguientes subsecciones son descritos los componentes seleccionados.

5.3.3.1. Servidor de red *LoRaWAN*

Este servidor es el encargado de reenviar los datos provenientes del *Gateway LoRaWAN*. La plataforma elegida para realizar esta tarea fue TTN [79]. Esta herramienta fue seleccionada después de revisar las funcionalidades de plataformas similares en el mercado y otros trabajos previos [47].

TTN, es una plataforma que funciona como servidor de red para *LoRaWAN*. La cual permite construir una red global abierta para IoT de bajo costo, máxima seguridad y escalable. A la vez permite la integración de HTTPS (*Hypertext Transfer Protocol Secure*) y MQTT (*Message Queuing Telemetry Transport*) con el objetivo de conectar dispositivos a internet para ser controlados y supervisados [79]. La implementación fue realizada siguiendo los pasos de configuración explicados en [78]. Fueron creados en la herramienta los respectivos dispositivos a usar en este trabajo (tarjetas micro controladas y *Gateway*). Considerando que la información recibida en la TTN se elimina cada vez que es actualizada la página o finaliza sesión, fue configurado el servicio de *Webhooks* [80]. Este servicio permitió integrar HTTPS para reenviar por *Uniform Resource Locator* (URL) en formato *JavaScript Object Notation* (JSON) los datos recibidos hacia un servidor en la nube, para ser almacenados de forma permanente.

5.3.3.2. Servidor en la nube

Este servidor permite almacenar en una base de datos la información proveniente del servidor de red *LoRaWAN*. Para realizar esta función fue elegido el servicio de Hostinger, considerando los costos, y beneficios técnicos y operativos de varias plataformas en el mercado [81]. Este servidor brinda los servicios de alojamiento web, servicio de almacenamiento en bases de datos, y gestión mediante herramientas como *phpMyAdmin*.

- *MYSQL*. Sistema de gestión de base de datos relacional de código abierto. La implementación, gestión y control de este tipo de bases de datos es realizada por medio del *Structured Query Language* (SQL) [82].
- *Structure Query Language* (SQL). Lenguaje de consulta estructurado, con el objetivo de administrar (crear, insertar, consultar y eliminar) la información de un sistema de gestión de datos relacionales [83].



- *PhpMyAdmin*. Aplicación que ofrece una interfaz gráfica para configurar el servicio de *MYSQL*, gestionar las tablas y su información, exportar e importar bases de datos y realizar consultas *SQL* [83].

La implementación de este componente consistió en crear las bases de datos y tablas mencionadas previamente (**sección 5.2**), para almacenar los respectivos datos. Además, fue creado un acceso remoto para consultar y enviar datos desde un equipo de cómputo local. Realizado lo anterior, a través del servicio de alojamiento web que ofrece el servidor en la nube, fue almacenado un programa (explicado en el **Anexo F. Aplicación Web desarrollada para el seguimiento de las pruebas realizadas**) con la finalidad de obtener los datos del mensaje JSON proveniente del servidor de red *LoRaWAN* y guardar la información en el conjunto de datos creado.

La **subsección 5.3.3.4** presenta los detalles de la interfaz de la página web desarrollada donde es posible visualizar los datos de localización en tiempo real de un vehículo en particular.

5.3.3.3. Modelos de *Machine Learning*

El modelo de ML fue seleccionado mediante la revisión de los documentos relacionados en el estado del arte, este modelo es usado en el bloque funcional del módulo de procesamiento de información descrito en la **sección 5.1.3**.

Para el desarrollo de cada uno de los modelos ML que utilizan los algoritmos desarrollados fue utilizada la metodología *Cross Industry Standard Process for Data Mining (CRISP-DM)*. *CRISP-DM* es una metodología para el desarrollo de procesos de minería de datos, que proporciona un modelo de ciclo de vida para proyectos de análisis de datos. Este modelo contiene seis fases (comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue) que están relacionadas entre sí [84].

A partir de la identificación de los objetivos de los modelos de ML en el **Anexo E. Modelos de *Machine Learning* usados en las pruebas controladas y no controladas**, para cada ruta fue desarrollada una arquitectura general del sistema de predicción presentada en la **Figura 25**. El flujo de ejecución de la arquitectura es el siguiente:

- Primero, el bloque de monitoreo de datos es el encargado de acceder a la información alojada en la base de datos del servidor en la nube. A partir del monitoreo de los datos los bloques de ejecución de entrenamiento y ejecución de escenarios entran en funcionamiento.
- Segundo, cuando existen al menos 200 registros en la base de datos, es activado el componente de entrenamiento del modelo, para después generar un modelo entrenado que servirá de entrada para los escenarios de predicción.
- Tercero, cuando el bloque de monitoreo detecta alguno de los eventos que hacen necesario ejecutar los modelos de predicción del escenario 1 o 2, estos son ejecutados, según sea el caso.
- Cuarto, con el modelo ya entrenado cada algoritmo genera las respectivas salidas de cada escenario, para después almacenar estos resultados en la base de datos.

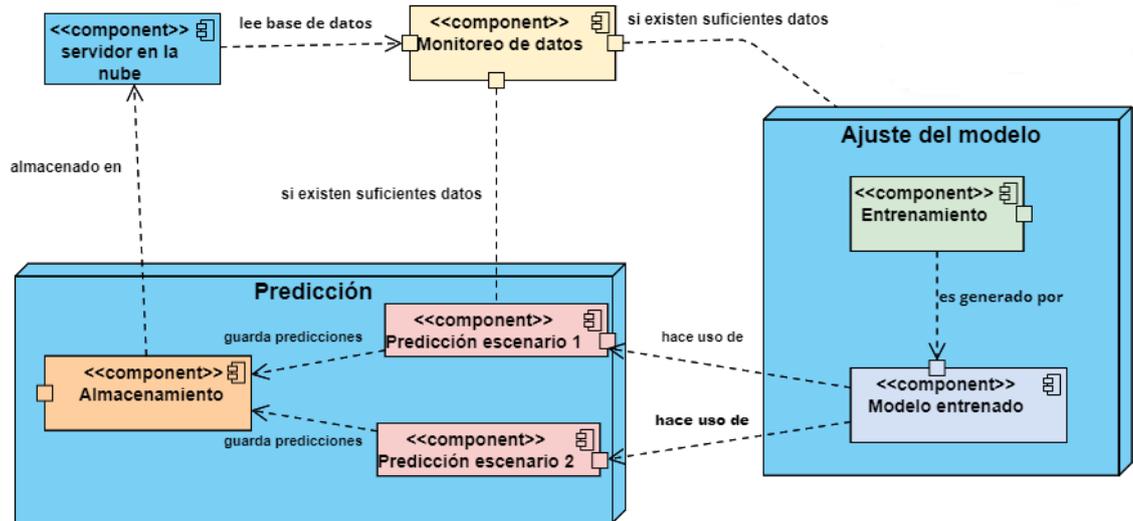


Figura 25. Arquitectura del algoritmo del sistema general de predicción.

Cabe aclarar que en la arquitectura (Figura 25) la ejecución de escenarios y ejecución de entrenamiento entran en funcionamiento de forma independiente, con el objetivo que si es necesario predecir no sea necesario esperar a tener un modelo entrenado. Para poder realizar las respectivas predicciones de los escenarios, es necesario haber ejecutado por lo menos una vez el entrenamiento para obtener el modelo entrenado y así lograr predecir. El Anexo E. Modelos de Machine Learning usados en las pruebas controladas y no controladas presenta una descripción del modelo de Machine Learning utilizado en este proyecto y el Anexo C. Algoritmos generados de las pruebas iniciales contiene el código fuente en GitHub.

El funcionamiento del algoritmo de entrenamiento es explicado con en un diagrama de flujo en la Figura 26.

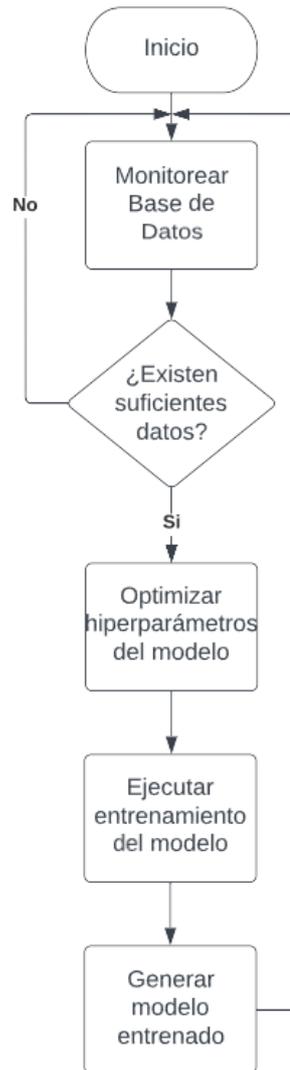


Figura 26. Diagrama de flujo del algoritmo de entrenamiento.

La **Figura 27** muestra el diagrama de flujo del algoritmo de predicción.

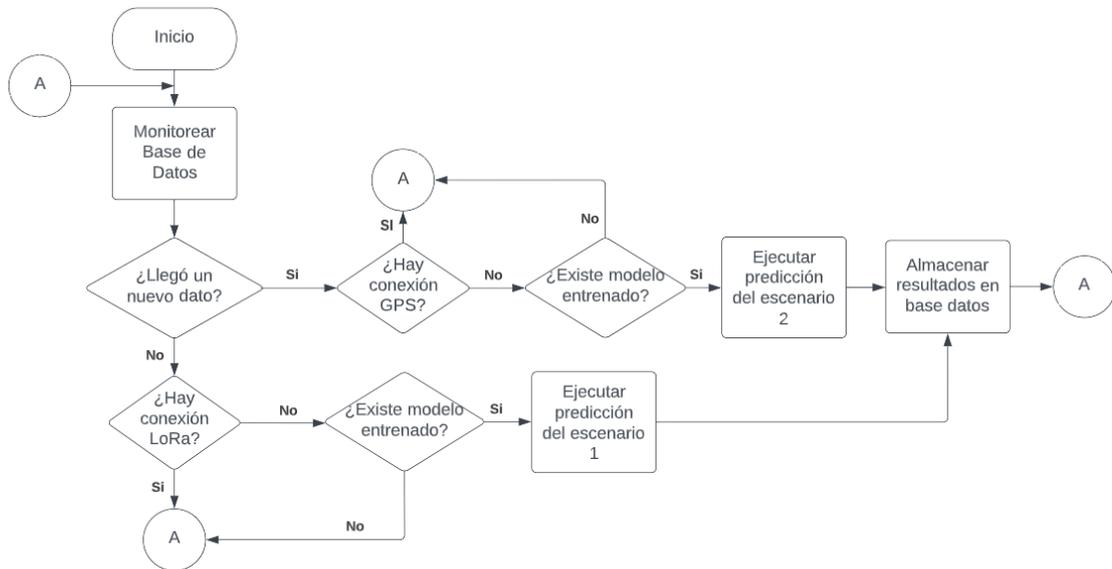


Figura 27. Diagrama de flujo algoritmo de predicción.

5.3.3.4. Aplicación web

El prototipo del sistema de seguimiento del vehículo público propuesto está centrado en el desarrollo de los modelos de ML, más que en la funcionalidad prestada a cierto usuario a través de la aplicación web propuesta. El único usuario considerado para el prototipo del sistema fue un usuario para la visualización de los resultados de las pruebas del prototipo. El usuario puede visualizar las rutas utilizadas en las pruebas, la información de los vehículos utilizados, y revisar a través de mapas y tablas de datos, los recorridos realizados por parte de dichos vehículos.

En el desarrollo de la aplicación web fue utilizado el marco de trabajo SCRUM, que puede ser considerado como una metodología ágil de desarrollo. Al utilizar SCRUM la documentación formal es la mínima, además, la descripción de la funcionalidad se realiza mediante HU, facilitando la descripción e implementación.

La aplicación web fue desarrollada haciendo uso del lenguaje PHP y una API de Google Maps y alojada en un servidor del proveedor Hostinger. Fueron considerados dos componentes en esta aplicación web. El primer (**Figura 28**) componente de la aplicación está encargado de recibir y almacenar los datos provenientes de la plataforma TTN, mientras que el segundo (**Figura 29**) componente es el encargado de visualizar y filtrar los datos.



```
<?php
include "connection.php"; // Connection to Data base information

$content = trim(file_get_contents("php://input")); //Receive the RAW post data.
$decoded = json_decode($content, true); //Attempt to decode the incoming RAW post data from JSON.

if(!is_array($decoded)){ //If json_decode failed, the JSON is invalid.
    throw new Exception('Received content contained invalid JSON!');
}

$mysqli = new mysqli($host, $user, $pw, $db); // The connection to the database is made.
date_default_timezone_set('America/Bogota'); // This line is important when the server is REMOTE and is local
$date = date("Y-m-d");
$hour = date("H:i:s");
//Data decoding
$dev_id=$decoded['end_device_ids']['device_id'];
$latitude= $decoded['uplink_message']['decoded_payload']['latitud'];
$longitude= $decoded['uplink_message']['decoded_payload']['longitud'];
$accx= $decoded['uplink_message']['decoded_payload']['aceleracion_x'];
$accy= $decoded['uplink_message']['decoded_payload']['aceleracion_y'];
$accz= $decoded['uplink_message']['decoded_payload']['aceleracion_z'];
$speed= $decoded['uplink_message']['decoded_payload']['velocidad'];
$gyrox= $decoded['uplink_message']['decoded_payload']['giro_x'];
$gyroy= $decoded['uplink_message']['decoded_payload']['giro_y'];
$gyroz= $decoded['uplink_message']['decoded_payload']['giro_z'];
$spreading_factor= $decoded['uplink_message']['settings']['data_rate']['lora']['spreading_factor'];
$bandwidth= $decoded['uplink_message']['settings']['data_rate']['lora']['bandwidth'];
$coding_rate= $decoded['uplink_message']['settings']['data_rate']['lora']['coding_rate'];
$rssi= $decoded['uplink_message']['rx_metadata'][0]['rssi'];
//Type of record
if($latitude == '0'){
    $type_record = 2;
}else{
    $type_record = 0;
}
//Data is inserted into the tables
if(!is_null($latitude)){
    $sql1 = "INSERT into data_set (dev_id, date, hour, latitude, longitude, speed, accx, accy, accz, gyrox,
    $result1 = $mysqli->query($sql1);
    $sql2 = "INSERT into Tabla_General (dev_id, date, hour, type_record, latitude, longitude, speed, accx,
    $result12 = $mysqli->query($sql2);
}
?>
```

Figura 28. Recepción y almacenamiento de los datos provenientes de la TTN.



Figura 29. Interfaz de inicio de la aplicación web.



En la interfaz inicial de la aplicación web son presentadas tres opciones, la primera opción “*rutas*” permite acceder a los recorridos a realizar en la ejecución de las pruebas de funcionamiento del sistema. La **Figura 30** muestra en un mapa una de las rutas definida para la realización de las pruebas.

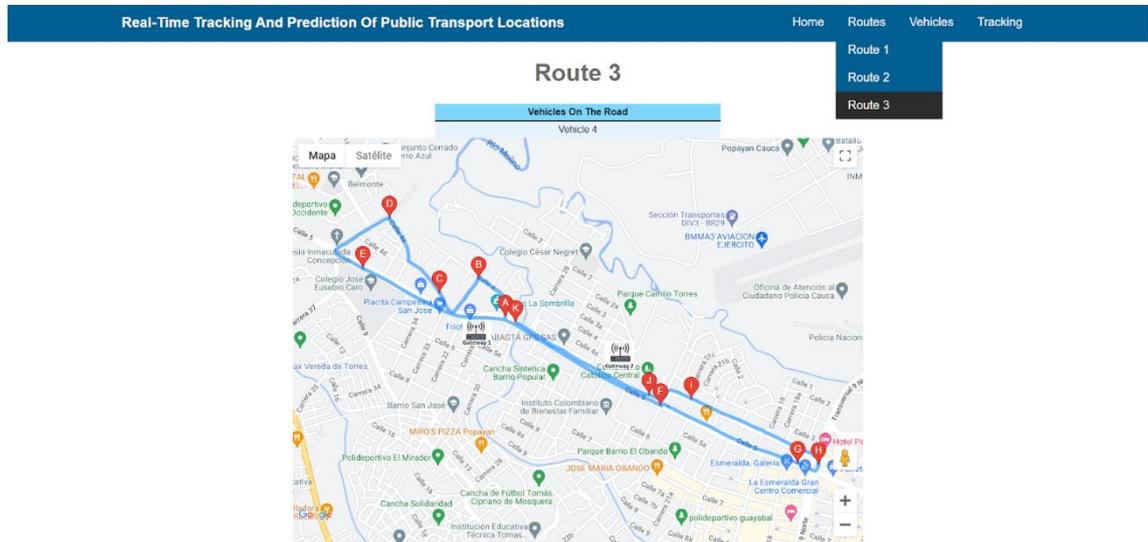


Figura 30. Interfaz de la ruta 3.

La segunda opción “*vehículos*” permite seleccionar los vehículos registrados en el sistema y ver sus características: la ruta asignada, la ubicación de los Gateways y el nombre del dispositivo transmisor a bordo. La **Figura 31** muestra en un mapa el recorrido para uno de los vehículos.

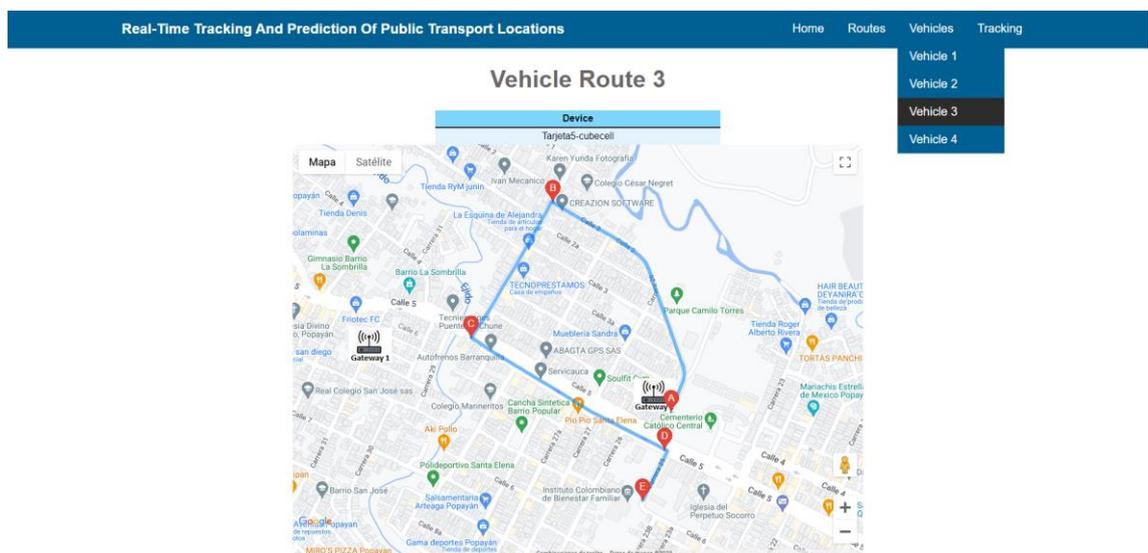


Figura 31. Interfaz del vehículo 3.

Por último, se presenta la tercera opción denominada “*tracking*”, en la cual es posible filtrar los datos a visualizar de cierto vehículo, en un intervalo de tiempo. El **Anexo F. Aplicación Web desarrollada para el seguimiento de las pruebas realizadas**



presenta una descripción detallada de la aplicación web desarrollada. La **Figura 32** muestra en un mapa los últimos 100 registros obtenidos sobre una de las rutas predefinidas. La **Figura 33** muestra una tabla con los valores detallados de los últimos 100 registros obtenidos sobre una de las rutas definidas anteriormente.

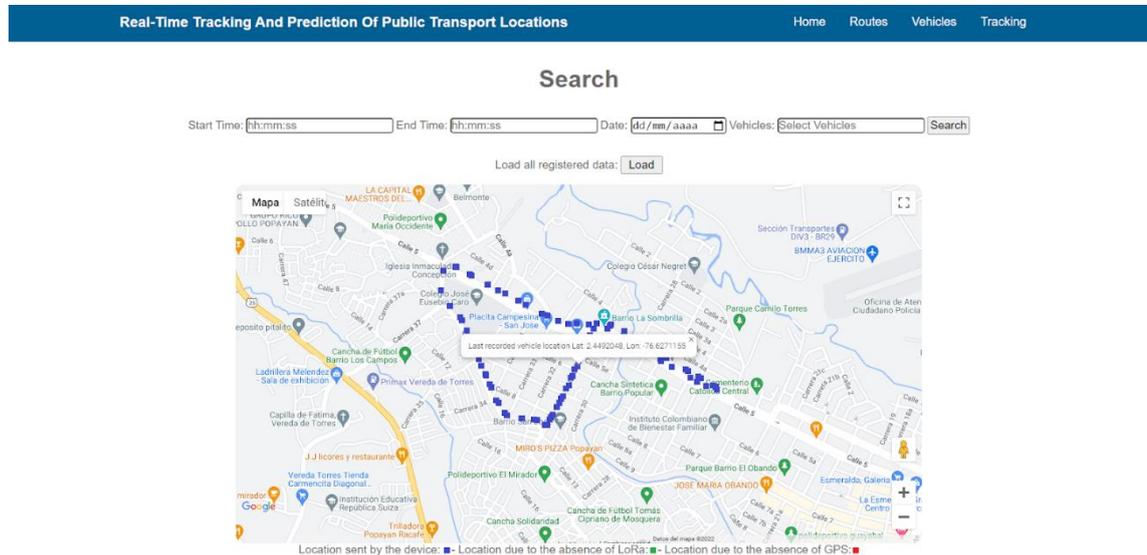


Figura 32. Interfaz de las ultimas 100 ubicaciones realizadas por el vehículo 2 en el mapa.

Registry Number	Device	Date	Hour	Predicted Hour	Latitude	Longitude	Predicted Latitude	Predicted Longitude	Speed	Acceleration	Turn	Spread Factor	Bandwidth	Code Rate	RSST
100	tarjeta2-cubecell	2022-12-04	12:04:29	00:00:00	2.4492048	-76.6271155			11.44536	-0.0358055	-1.6814355	7	125000	4/5	-71
99	tarjeta2-cubecell	2022-12-04	12:04:18	00:00:00	2.4489141	-76.6272539			14.27892	-0.3495531	0.0895568	7	125000	4/5	-77
98	tarjeta2-cubecell	2022-12-04	12:04:07	00:00:00	2.4485985	-76.6273866			15.07528	-0.2624925	0.1658927	7	125000	4/5	-82
97	tarjeta2-cubecell	2022-12-04	12:03:56	00:00:00	2.4482125	-76.6275643			14.6308	-0.0453856	3.234595	7	125000	4/5	-89
96	tarjeta2-cubecell	2022-12-04	12:03:45	00:00:00	2.447857	-76.6277248			8.85256	-0.5267845	0.8529156	7	125000	4/5	-88
95	tarjeta2-cubecell	2022-12-04	12:03:34	00:00:00	2.447562	-76.6278776			16.74208	-0.344763	-5.3913593	7	125000	4/5	-96
94	tarjeta2-cubecell	2022-12-04	12:03:22	00:00:00	2.4471828	-76.6280941			8.57476	-0.5603148	-2.8675424	7	125000	4/5	-96
93	tarjeta2-cubecell	2022-12-04	12:03:11	00:00:00	2.4469173	-76.628228			12.48248	-0.6656957	5.6010074	7	125000	4/5	-94
92	tarjeta2-cubecell	2022-12-04	12:03:03	00:00:00	2.4471131	-76.628688			16.26056	0.776106	-2.8264737	7	125000	4/5	-113
91	tarjeta2-cubecell	2022-12-04	12:02:43	00:00:00	2.4471425	-76.6291339			20.87204	1.3748609	0.8223812	7	125000	4/5	-108

Figura 33. Interfaz de los registros obtenidos del vehículo 2.

5.4. Pruebas y ajustes del sistema

Las pruebas por realizar siguen el siguiente protocolo: primero verificación de dispositivos hardware conectados (GPS, IMU, LoRa), segundo, definición de las rutas a seguir y por último pruebas integración.



Primero: Para comprobar la funcionalidad del sistema, el día 13 de noviembre del 2022 fueron conectados los 4 dispositivos en puntos estáticos al interior de una vivienda (en la ciudad de Popayán) y cerca de los *Gateway* (pocos metros) con la finalidad de observar la conexión a través de *LoRa* y el protocolo *LoRaWAN*. En esta primera prueba fue comprobado que los dispositivos conectaban con el sistema enviando paquetes de los sensores GPS e IMU. Al realizar la prueba, el módulo GPS Ublox-6M conectado a la ESP32 *LoRa* realizaba la comunicación satelital en un tiempo inferior a 7 minutos a diferencia de la tarjeta Cube Cell HTCC-AB02S que a pesar de tener un incorporado un sensor GPS y su respectiva antena, el tiempo en conectarse fue de 15 a 30 minutos. Adicionalmente, mientras se establece el enlace satelital las mediciones realizadas por el sistema de posicionamiento global son iguales a cero. Durante esta prueba inicial, la tarjeta ESP32 *LoRa* a pesar de estar cerca de los *Gateway* realizaba el ajuste de los parámetros *LoRa* (factor de ensanchamiento, ancho de banda y tasa de código). En la base de datos eran insertados estos parámetros *LoRa*, pero no los medidos por los sensores (GPS e IMU). Lo anterior, requirió de ajustes para registrar todos los datos necesarios

Segundo: El día 29 de noviembre del 2022 fueron realizadas diferentes rutas de occidente a oriente de la ciudad de Popayán, en los barrios denominados María Occidente (2.452999, -76.631835), Camilo Torres (2.448446, -76.621751) y Santa Helena (2.447085, -76.623708), recorridos que iniciaron a las 11:17 a.m. hasta las 11:54 a.m. con el objetivo de verificar el alcance de cobertura de los *Gateway*. A partir de estos recorridos fueron seleccionadas las rutas y la ubicación ideal de los *Gateways*, que serán explicadas en detalle en otro documento del trabajo de investigación. Las rutas fueron recorridas en un vehículo (Renault Sandero) en el cual fueron ubicados dos dispositivos basados en la tarjeta Cube Cell. A una de las tarjetas se le asignó un valor de intento de envío de trama igual a 1 y a la otra igual a 2. Para el dispositivo con el valor igual a 1 fueron obtenidos 85 datos y para el valor igual a 2 157 datos. Esta prueba fue realizada con el objetivo de obtener la configuración con la mayor cantidad de datos y que las tarjetas estuvieran transmitiendo información de forma continua.

Los datos recolectados en las pruebas descritas, con dispositivos dentro de un vehículo, siguiendo ciertas rutas, sirvieron para presentar el funcionamiento básico de cada una de las opciones de la aplicación web.

Tercero: Las pruebas de integración consistieron en verificar el funcionamiento de todo el sistema, donde los 4 dispositivos fueron ubicados nuevamente al interior de una vivienda (en la ciudad de Popayán) y a una distancia cercana de los *Gateways*. Con el objetivo que los dispositivos estuvieran transmitiendo de forma simultánea paquetes y recolectar una suficiente cantidad de datos. Una vez recolectados suficientes datos para ejecutar el entrenamiento del algoritmo fueron realizadas algunas pruebas de funcionamiento previas donde al ejecutar al mismo tiempo 8 algoritmos (1 algoritmo de entrenamiento y 1 algoritmo de predicción por cada uno de los 4 vehículos) existían inconvenientes con el servidor en la nube. Esto debido a que era excedido el límite de 500 conexiones simultáneas por hora permitidas por el servidor. Este límite era excedido debido a que el algoritmo para cada vehículo debe de estar consultado de forma constante la base de datos para determinar cuál escenario ejecutar para cada ruta y a la vez estar verificando si hay la suficiente cantidad de datos para ejecutar el algoritmo de entrenamiento. Por lo anterior, es descartada la ejecución de los algoritmos de entrenamiento y predicción al mismo tiempo, la predicción debe realizarse para 3 tres vehículos una vez los modelos estén ya entrenados.



Al realizar las pruebas individuales de la funcionalidad de cada módulo, y las pruebas de integración, fueron detectados algunos detalles por corregir. Los ajustes realizados en los respectivos módulos son descritos a continuación:

- **Módulo de recolección de datos.** El primer ajuste realizado en este módulo fue en la tarjeta Cube Cell HTCC-AB02S la cual trae incorporado un módulo GPS y su respectiva antena. Al probar el funcionamiento de la tarjeta, el tiempo en establecer la conexión satelital tarda entre 15 y 30 minutos. Debido a lo anterior, fueron adquiridas antenas cerámicas para el sensor GPS de estas tarjetas, con esto el tiempo de conexión no era superior a 5 minutos. En los *sketchs* de recolección también fueron realizados algunos cambios. El primero fue establecer un valor igual a 2 en la variable `confirmedNbTrials`, la cual establece el número de intentos realizados para transmitir el mismo paquete, en caso de no recibir un acuse de recibo. Este ajuste fue realizado para que el dispositivo a bordo del vehículo envíe de forma constante y actualizada cada una de las lecturas realizadas por los sensores GPS e IMU y no quede retransmitiendo lecturas demasiado antiguas. El segundo cambio en este fichero fue realizado para que solo sean enviados paquetes cuando la lectura del GPS sea diferente de cero, ya que dicho valor es el obtenido, mientras la comunicación satelital es establecida. El tercer cambio consistió en establecer un valor de 3000 milisegundos para la variable `appTxDutyCycle`, que corresponde al ciclo de trabajo de transmisión de los paquetes. Significa que cada 3 segundos son capturadas las mediciones los sensores y transmitidos. Fue elegido este valor porque al establecer valores inferiores, el módulo dejaba de funcionar (bloqueo del dispositivo) ocasionando la no transmisión de las lecturas realizadas por los sensores. Es de aclarar que a pesar de estar transmitiendo cada 3000 milisegundos los datos eran recibidos por el sistema en un intervalo de 9 a 13 segundos debido a la propagación de la señal en el espacio, más el procesamiento de los datos en cada uno de los servidores antes de ser almacenados. Por último, el módulo fue configurado como un nodo clase. Porque solo era necesario transmitir los datos desde el vehículo hacia el receptor independiente de que este tipo de nodo tenga dos ventanas de recepción de datos.
- **Módulo de procesamiento de la información.** En este módulo fue realizado un ajuste en el fichero PHP encargado de recibir la información proveniente del servidor de red LoRaWAN. Durante el almacenamiento de los datos fue observado que en algunas ocasiones solo eran obtenidos los valores del factor de ensanchamiento, ancho de banda, tasa de código y el indicador de fuerza de la señal recibida, mientras que las variables referentes a las mediciones realizadas por los sensores llegaban con valores nulos. Este comportamiento fue originado, porque el protocolo LoRaWAN es el encargado de configurar los parámetros de comunicación LoRa a medida que aumenta la distancia entre el dispositivo a bordo del vehículo y el Gateway. Estos parámetros influyen en la velocidad de datos efectiva, la resistencia al ruido y la fácil decodificación. Para evitar este inconveniente, en el fichero fue incluida una condición en donde solo sean insertados registros en las tablas cuando las variables correspondientes a las lecturas de los sensores (latitud, longitud, aceleración y giro) sean diferentes de nulos.



Capítulo 6

6. Diseño y ejecución de pruebas de campo y construcción del conjunto de datos

6.1. Diseño de las pruebas de campo.

Las pruebas de campo fueron diseñadas con el objetivo de evaluar el seguimiento vehicular a través de la tecnología de comunicación LoRa y el protocolo de red LoRaWAN. Además, evaluar el desempeño del modelo de ML al predecir futuras ubicaciones cuando se pierde la comunicación LoRa o la conexión del GPS. Para cumplir los anteriores objetivos fueron definidas dos tipos de pruebas: controladas y no controladas. Las pruebas controladas fueron diseñadas para recolectar datos de ubicación de las rutas, y de este modo, entrenar el modelo de ML; además, observar el comportamiento de la predicción en el escenario 1 y 2. Las pruebas no controladas fueron diseñadas para comprobar el funcionamiento del algoritmo de predicción en un entorno real, donde es posible tener los dos escenarios anteriores.

6.1.1. Definición de rutas y elementos a utilizar

Fue planeado ubicar cada uno de los dispositivos al interior de un vehículo, para un total de 4, la ubicación del dispositivo debe realizarse en la parte intermedia del piso de los pasajeros debido a que esta zona es una superficie horizontal que permite una adecuada calibración y lectura de las mediciones a realizar por la IMU.

A pesar de que el objetivo del trabajo es realizar seguimiento del servicio de transporte público colectivo, para el desarrollo de las pruebas fue determinado utilizar vehículos de servicio particular de tipo automóvil, debido a que no fue posible conseguir vehículos de transporte público disponibles para las pruebas. Una condición considerada relevante, en la selección de los vehículos para realizar las pruebas, fueron: disponer en su interior una salida con puerto USB que ofrezca una tensión de 5 voltios, la cual puede servir como fuente de voltaje fija para los dispositivos. El puerto puede ser proveído directamente en el tablero frontal del vehículo, o a través de un adaptador, como el que presenta la **Figura 34**.

La **Tabla 7** presenta las características de los vehículos que fueron seleccionados para las pruebas, especificando la marca, el modelo, y el tipo de automóvil.

Tabla 7. Características de los vehículos.

Vehículo	Marca	Modelo	Tipo
1	Hyundai Accent	2018	Automóvil
2	Renault Sandero	2016	Automóvil
3	Chevrolet Spark	2016	Automóvil
4	Chevrolet N300	2012	Van



Figura 34. Adaptador USB para vehículo.

Los dispositivos *Gateway* fueron ubicados en dos edificaciones con la suficiente altura. Las edificaciones seleccionadas debían cumplir con los siguientes requerimientos: un punto de toma de corriente eléctrica, tener una altura considerable (mayor a 10 metros de altura) para tener una buena línea de vista y lograr evadir la mayor cantidad de obstáculos que puedan interferir en la comunicación. la **Tabla 8** presenta la ubicación geográfica de cada uno de los *Gateways*.

Tabla 8. Ubicación geográfica de los *Gateway*.

Nombre (<i>Gateway</i>)	Sector	Latitud [grados]	Longitud [grados]	Altura Edificación [metros]
LPS8	Barrio San José	2.449317	-76.626888	15
LPS8-2	Barrio Camilo Torres	2.448446	-76.621751	15

A partir de la ubicación de los *Gateway* fueron evaluadas diversas posibilidades de rutas, cercanas a estos puntos. Fueron seleccionadas finalmente tres rutas, para ser recorridas por los vehículos, considerando los resultados obtenidos con pruebas preliminares de recepción de datos. A continuación, la descripción de cada uno de los puntos que conforman las rutas y su respectiva visualización en un mapa en las **Figuras 35, 36 y 37** correspondientes a las rutas 1, 2 y 3. En cada mapa también es ubicado cada uno de los 2 *Gateways*.

- **Ruta 1:**

- *Puntos que conforman la ruta:*

- **A:** Carrera 31° con Calle 5°, barrio San José (2.450315, -76.626607).
- **B:** Calle 5° con Carrera 26°, barrio Santa Helena (2.448086, -76.622215).
- **C:** Calle 5° con Carrera 25°, barrio Camilo Torres (2.447976, -76.621707).
- **D:** Calle 5° con Carrera 30°, barrio La Sombrilla (2.450378, -76.625922).
- **E:** Calle 5° con Carrera 36°, barrio Carlos Primero (2.451983, -76.630474).



- E: Carrera 36° con Calle 5°, barrio Las Palmas (2.452267, -76.631980).
- G: Calle 9° con Carrera 34°, barrio El Mirador (2.447926, -76.630211).
- H: Carrera 31° con Calle 5°, barrio El Mirador (2.446828, -76.628339).
- I: Carrera 31° con Calle 5°, barrio San José (2.450092, -76.626713).

○ *Distancia*: 3.12 Km.

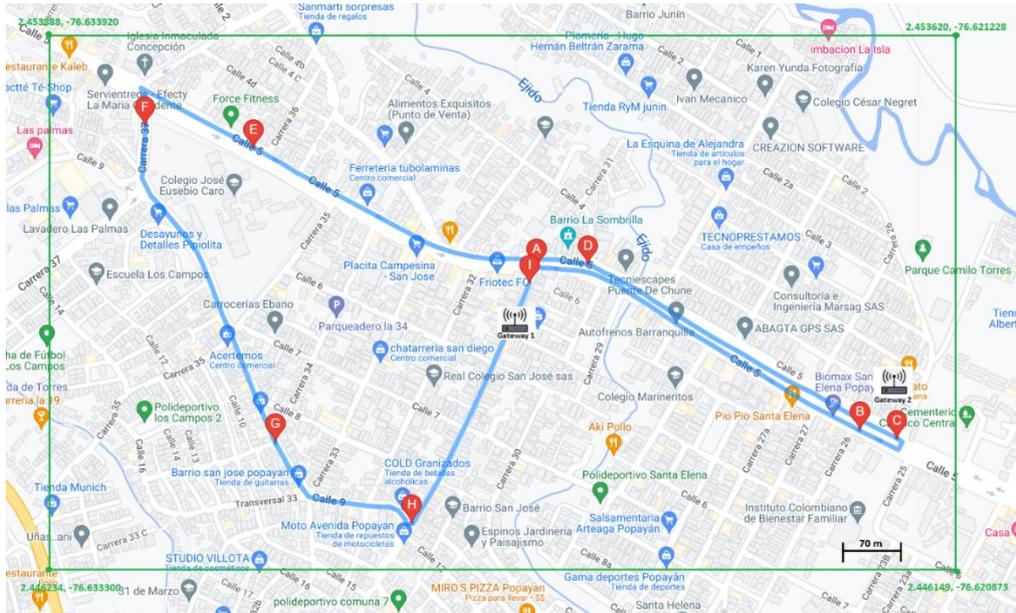


Figura 35. Mapa de la ruta 1 y ubicaciones de los 2 Gateways (mapa tomado de Google Maps).

• **Ruta 2:**

- *Puntos que conforman la ruta:*
 - A: Carrera 26° con Calle 6°, barrio José María Obando (2.446080, -76.620817).
 - B: Calle 5°, Cementerio Central (2.447624, -76.620976).
 - C: Carrera 28° con Calle 2°, barrio Junín (2.452255, -76.623589).
 - D: Calle 5° con Carrera 29°, barrio Santa Helena (2.449769, -76.625077).
 - E: Calle 5°, barrio Santa Helena (2.447738, -76.621584).
 - F: Calle 6° con Carrera 26°, barrio José María Obando (2.446145, -76.621040).

○ *Distancia*: 1.96 Km.

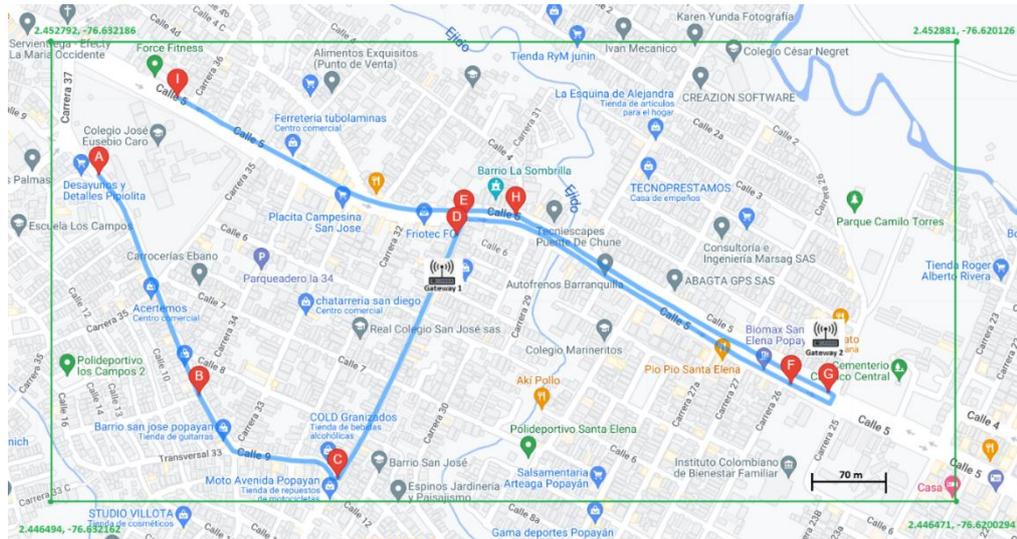


Figura 38. Mapa del recorrido del vehículo 1 y ubicaciones de los 2 Gateways (mapa tomado de Google Maps).

- **Recorrido del vehículo 2 – Renault Logan:**

- Ruta asignada: Ruta 1.
- Punto de inicio:
 - A: Carrera 31° con Calle 5e°, barrio San José (2.449403, -76.627020).
- Puntos intermedios:
 - B: Carrera 31° con Calle 5°, barrio San José (2.450315, -76.626607).
 - C: Calle 5° con Carrera 26°, barrio Santa Helena (2.448086, -76.622215).
 - D: Calle 5° con Carrera 25°, barrio Camilo Torres (2.447976, -76.621707).
 - E: Calle 5° con Carrera 30°, barrio La Sombrilla (2.450378, -76.625922).
 - F: Calle 5° con Carrera 36°, barrio Carlos Primero (2.451983, -76.630474).
 - G: Carrera 37° con Calle 5°, barrio Las Palmas (2.452267, -76.631980).
 - H: Calle 9° con Carrera 34°, barrio El Mirador (2.447926, -76.630211).
- Punto de finalización:
 - I: Calle 9° con Carrera 34°, barrio El Mirador (2.446828, -76.62833).

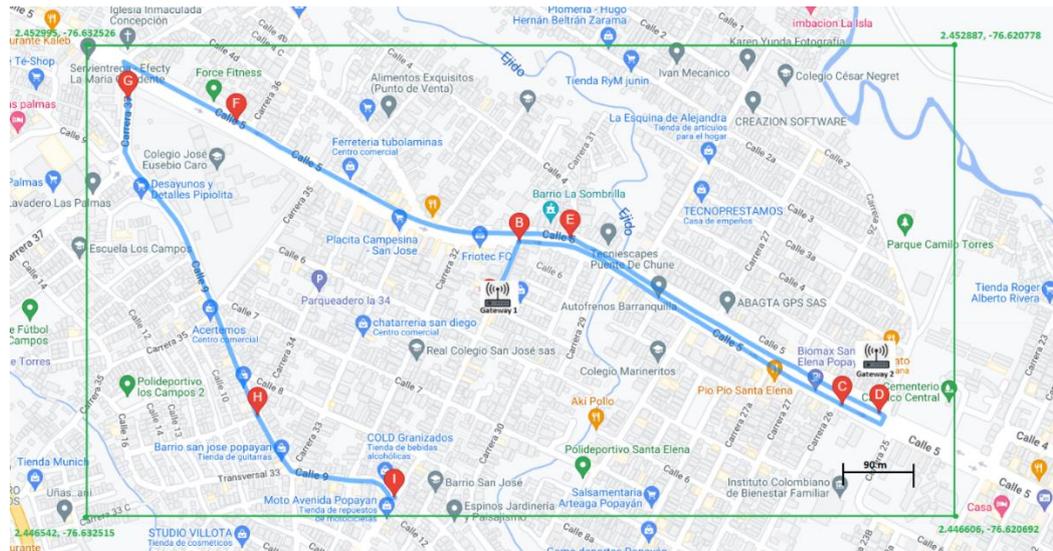


Figura 39. Mapa del recorrido del vehículo 2 y ubicaciones de los 2 Gateways (mapa tomado de Google Maps).

- **Recorrido del vehículo 3 – Chevrolet Spark:**

- *Ruta asignada: Ruta 2.*
- *Punto de inicio:*
 - A: Carrera 25° con 4a°, barrio Camilo Torres (2.448399, -76.621391).
- *Puntos intermedios:*
 - B: Carrera 28° con Calle 2°, barrio Junín (2.452255, -76.623589).
 - C: Calle 5° con Carrera 29°, barrio Santa Helena (2.449769, -76.625077).
 - D: Calle 5°, barrio Santa Helena (2.447738, -76.621584).
- *Punto de finalización:*
 - E: Carrera 25° con Calle 6°, Bienestar Familiar (2.446834, -76.621997).

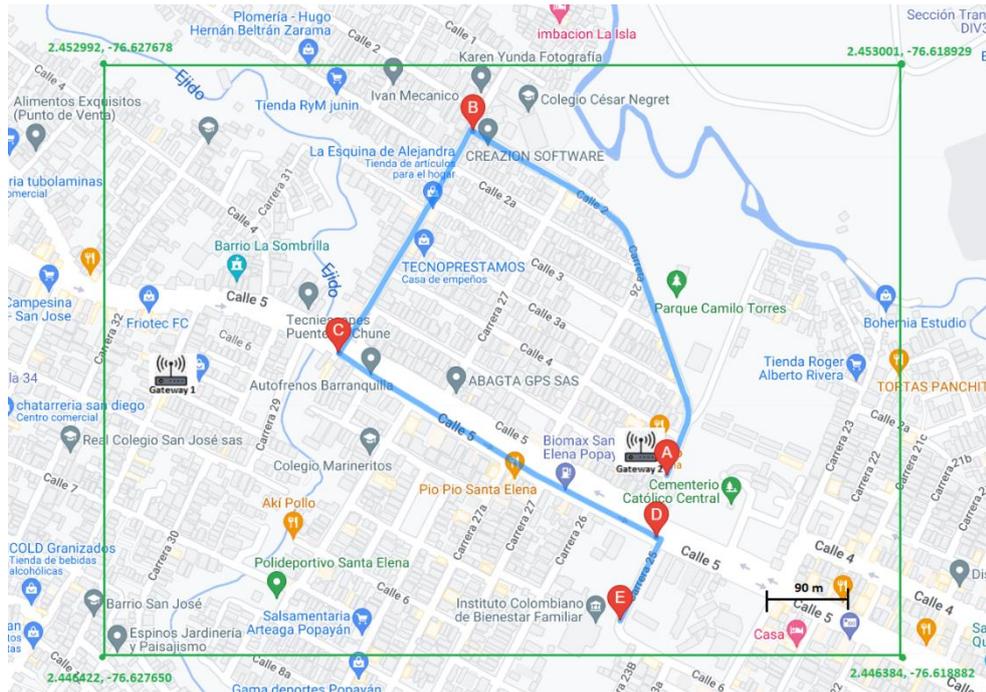


Figura 40. Mapa del recorrido del vehículo 3 y ubicaciones de los 2 Gateways (mapa tomado de Google Maps).

- **Recorrido del vehículo 4 - Chevrolet N300:**
 - *Ruta asignada: Ruta 3.*
 - *Punto de inicio:*
 - A: Calle 5° con Carrera 17°, La Esmeralda (2.445051, -76.615319).
 - *Puntos intermedios:*
 - B: Carrera 17°, Puente Deprimido (2.445000, -76.614593).
 - C: Calle 4° con Carrera 21°, barrio Pandiguando (2.447395, -76.619153).
 - D: Carrera 23°, Cementerio Central (2.447543, -76.620691).
 - E: Calle 5° con Carrera 29°, Puente Chune (2.450176, -76.625507).
 - F: Calle 4°, barrio La Sombrilla (2.450371, -76.625871).
 - G: Carrera 32° con Calle 4°, barrio La Sombrilla (2.451753, -76.626850).
 - H: Calle 4ª, barrio La Sombrilla (2.451253, -76.628241).
 - I: Carrera 37° con Calle 4ª, barrio La María Occidente (2.453975, -76.630095).
 - J: Calle 5°, Colegio José Eusebio Caro (2.452122, -76.631032).
 - *Punto de finalización:*
 - K: Calle 5°, Iglesia del Perpetuo Socorro (2.447135, -76.620283).

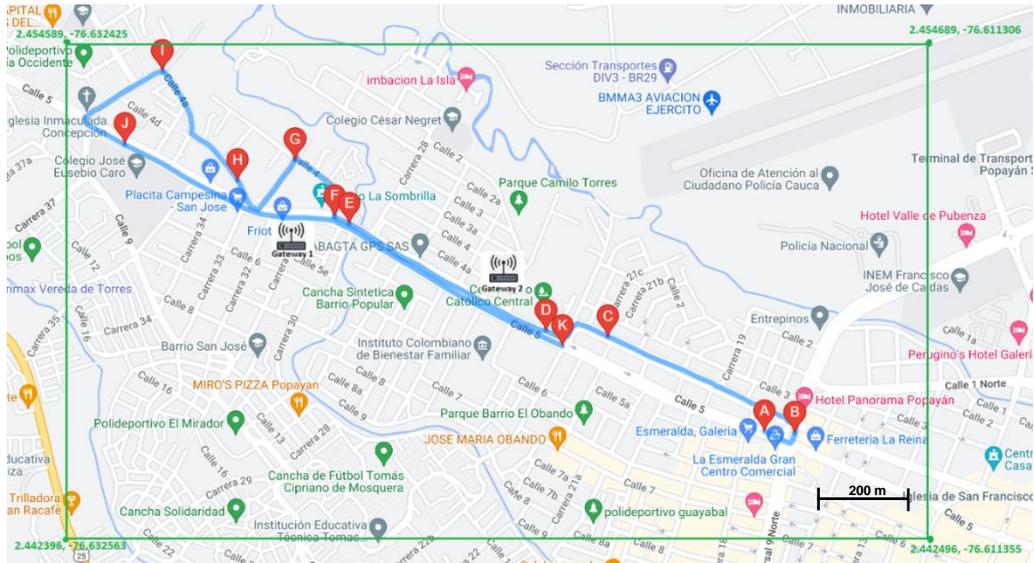


Figura 41. Mapa del recorrido del vehículo 4 y ubicaciones de los 2 Gateways (mapa tomado de Google Maps).

6.1.2. Diseño de pruebas controladas

El diseño de las pruebas controladas estuvo dirigido a recolectar una cantidad suficiente de datos para el algoritmo de entrenamiento. Además, simular la ausencia de la señal GPS y la ausencia de conexión LoRa.

Con base en lo anterior fue planeada la recolección de 300 datos para la ruta 1 y 3, y 200 datos para la ruta 2.

Se planeó recolectar más datos en la ruta 1 y 3 debido a que estas tienen una mayor distancia respecto a la ruta 2. El objetivo de la recolección fue generar un modelo a partir del algoritmo de entrenamiento, para poder ejecutar las pruebas de campo no controladas para las tres rutas descritas en la sección anterior.

Para la recolección de los datos fue planeado utilizar un solo vehículo, con un único conductor para recolectar un total de 300 datos (rutas 1 y 3) por un tiempo aproximado de 65 minutos por cada ruta y 200 datos (ruta 2) en tiempo de 43 minutos. En condiciones ideales, fue planeado recibir mediciones cada 13 segundos.

Como se menciona en la **subsección 5.13**, fueron manejados dos tipos de escenarios que permitieron evaluar el rendimiento de los modelos de ML. En el escenario 1 el dispositivo no es capaz de enviar paquetes LoRa (mediciones del GPS e IMU) al Gateway. Mientras que en el escenario 2 el sistema GPS no es capaz de capturar los datos de ubicación y solo son enviadas las mediciones de la IMU.

Aunque el escenario 1 puede presentarse frecuentemente durante el desarrollo de los recorridos planeados (ya que la comunicación LoRa entre dispositivo y Gateway no es posible establecerla cuando hay una distancia muy grande o cuando hay obstáculos considerables en la línea de vista), se estimó conveniente realizar un procedimiento para simular su ocurrencia.

Para realizar la simulación del escenario 2, debe conectarse un switch en uno de los pines GPIO de los dispositivos, y deben definirse las instrucciones en el código de programación de cada dispositivo por medio del entorno de desarrollo de Arduino. Con lo anterior, cuando sea accionado el switch, el dispositivo enviará en cero los valores de



la latitud, longitud y velocidad y una vez vuelva el switch a su estado inicial serán enviados nuevamente los valores que detecta el sensor GPS.

Para estas pruebas, el vehículo debe realizar los recorridos presentados en las **Figuras 38, 40 y 41** teniendo como partida el punto “A” de cada uno de los respectivos recorridos. Para los recorridos de las **Figuras 38 y 41** el vehículo debe acercarse al *Gateway* y para la **Figura 40** el vehículo estaría alejándose del *Gateway*. La simulación del escenario 1 fue planeada recolectando 30 datos, para después retirar la alimentación de cada dispositivo de recolección de datos y así el algoritmo realice la predicción de la ubicación para cada uno de los recorridos. El anterior procedimiento también fue planeado para el escenario 2 con la salvedad que no se desconectaría la fuente alimentación.

Las **Tabla 9, 10 y 11** describen de forma resumida, el cronograma del plan de pruebas controladas.

Tabla 9. Lista de recorridos planeados para la recolección de datos para el entrenamiento.

Vehículo	Ruta	Fecha programada	Hora de Inicio programada	Hora final programada	N°. de Registros Esperados (Enviados por dispositivo)
2	1	02-dic-2022	10:00 a.m.	11:05 a.m.	300
2	2	02-dic-2022	11:15 a.m.	11:58 a.m.	200
2	3	02-dic-2022	2:00 p.m.	3:05 a.m.	300

Tabla 10. Lista de recorridos planeados para las pruebas controladas para el escenario 1.

Vehículo	Ruta	Fecha programada	Hora de Inicio programada	Hora final programada	N°. de Registros Esperados (Enviados por dispositivo)
2	1	08-dic-2022	10:00 a.m.	10:07 a.m.	30
2	2	08-dic-2022	10:10 a.m.	10:17 a.m.	30
2	3	08-dic-2022	10:20 a.m.	10:27 a.m.	30

Tabla 11. Lista de recorridos planeados para las pruebas controladas para el escenario 2.

Vehículo	Ruta	Fecha programada	Hora de Inicio programada	Hora final programada	N°. de Registros Esperados (Enviados por dispositivo)
2	1	08-dic-2022	10:30 a.m.	10:37 a.m.	30
2	2	08-dic-2022	10:40 a.m.	10:47 a.m.	30
2	3	08-dic-2022	10:50 a.m.	10:57 a.m.	30



Con el objetivo de obtener la ubicación real de los vehículos durante la ejecución de los escenarios de predicción fue desarrollada una aplicación para *Smartphone* que almacena los datos obtenidos mediante el sensor GPS del teléfono en una base de datos local, luego estos datos almacenados son usados para compararlos con datos predichos por el sistema y evaluar así la precisión de la predicción.

6.1.3. Diseño de pruebas no controladas.

Para las pruebas no controladas fue planeado observar el comportamiento de la conexión *LoRa* en el seguimiento de vehículos y predicción de la ubicación (latitud y longitud) del automotor en ausencia de la comunicación *LoRa* o GPS. Para llevar a cabo este procedimiento debe utilizarse el mismo modelo ya entrenado, descrito en la **sección 6.1.2**, donde los ficheros encargados de realizar la predicción estarían ejecutándose de forma simultánea y estarían accediendo de forma constante a la base de datos para determinar si ejecutar o no la predicción ante la ausencia de comunicación *LoRa* o GPS. Para el caso de desconexión *LoRa*, en cada recorrido el algoritmo calcularía la diferencia de tiempo entre la hora actual y la hora del último dato registrado, cuando la diferencia sea mayor a 13 segundos (promedio de tiempo entre datos recibidos durante el entrenamiento) entrarían en ejecución los algoritmos y generarían las posibles futuras ubicaciones del vehículo hasta volver a recibir un nuevo paquete con información de ubicación real desde el servidor. En caso de falta de conexión GPS el valor de las variables registradas latitud, longitud y velocidad son iguales a cero, debe iniciarse la ejecución de los algoritmos y producirse las posibles futuras ubicaciones del vehículo hasta recibir estas variables en un valor diferente de cero.

Los vehículos 1 y 2 deben recorrer la ruta 1, el vehículo 1 acercándose al *Gateway* 1 desde el punto "A" hasta el punto "I" del recorrido mostrado en la **Figura 38**, el vehículo 2 alejándose del mismo *Gateway*, siguiendo el recorrido que va desde el punto "A" al punto "I" mostrado en la **Figura 39**.

Al mismo tiempo, los vehículos 3 y 4 deben recorrer la ruta 2 y la ruta 3 respectivamente. El vehículo 3 comenzando a alejarse del *Gateway* 2, desde el punto "A" hasta el punto "E" del recorrido mostrado en la **Figura 40**. El vehículo 4 comenzando a acercarse al *Gateway* 2 desde el punto "A" hasta el punto "K" del recorrido mostrado en la **Figura 41**. Todos los vehículos empezarán y terminarán su recorrido a la misma hora. Una vez que la ruta más larga, la ruta 3, sea completada 3 veces, debe ejecutarse el algoritmo de predicción para los vehículos 2, 3 y 4, no se predecirá sobre todos los vehículos debido a que en pruebas de funcionamiento anteriores se alcanzaba el máximo de conexiones por hora (500 conexiones) a la base de datos a los 5 minutos de empezar a ejecutar los 4 algoritmos.

Para obtener la ubicación real del vehículo en las pruebas no controladas, se siguió el mismo procedimiento que las pruebas controladas.

La **Tabla 12** describe de forma resumida, el cronograma del plan de pruebas no controladas. El número de registros esperados de la **Tabla 12** es obtenido a partir de la duración estimada de la prueba (1 hora 42 minutos) y la duración media de recepción entre datos que es de 13 segundos en condiciones ideales (línea de vista entre el transmisor y el receptor).



Tabla 12. Lista de recorridos planeados para las pruebas no controladas.

Vehículo	Ruta	Fecha programada	Hora de Inicio programada	Hora final programada	N°. de Registros Esperados (Enviados por dispositivo)
1	1	11-dic-2022	11:55 a.m.	1:37 p.m.	475
2	1	11-dic-2022	11:55 a.m.	1:37 p.m.	475
3	2	11-dic-2022	11:55 a.m.	1:37 p.m.	475
4	3	11-dic-2022	11:55 a.m.	1:37 p.m.	475

6.2. Ejecución de las pruebas

Antes de iniciar la ejecución de las respectivas pruebas fue realizada la sugerencia a cada conductor que mantuvieran una velocidad entre 20 y 25 Km/h. Este rango de velocidades es debido a que en los experimentos realizados en [12] obtuvieron una recepción del 100% de los paquetes al utilizar velocidades de 0 a 20 Km/h y de 21 a 30 Km/h. En [11] de acuerdo con la investigación realizada encuentran que al no exceder una velocidad de 25 Km/h permite no sufrir degradaciones por el efecto Doppler logrando una comunicación confiable para aplicaciones de seguimiento y rastreo de personas. Adicionalmente, en [11], en las pruebas realizadas concluyeron que aumentar la velocidad de 20 a 60 Km/h produce un aumento en la pérdida de paquetes del 6%. No fue establecida una velocidad inferior a 20 Km/h porque en pruebas iniciales se llegó a originar aumento del tráfico vehicular e inconvenientes con otros conductores.

6.2.1. Resultados de pruebas controladas

La **Tabla 13** presenta los resultados obtenidos en la recolección de datos para el entrenamiento del modelo de ML.

Tabla 13. Resultados obtenidos de las pruebas controladas ejecutadas para el algoritmo de entrenamiento.

Vehículo	Ruta	Fecha	Hora Inicio	Hora Final	N° de Vueltas por recorrido	N°. de Registros Obtenidos (Enviados por dispositivo)
2	1	2-dic-2022	10:26 a.m.	12:20 p.m.	7	302
2	2	2-dic-2022	2:03 p.m.	2:53 p.m.	6	215
2	3	2-dic-2022	3:38 p.m.	5:10 p.m.	4	304

La **Tabla 14** presenta los resultados obtenidos en las pruebas controladas forzando la ejecución del escenario 1. Las pruebas no fueron ejecutadas en la fecha en la que fueron planeadas debido a fallas presentadas en el punto de acceso a internet en uno de los



Gateways. Inicialmente las pruebas fueron llevadas a cabo el 23 de diciembre, pero debido a factores climáticos en la ciudad, fueron completadas el día 27 de diciembre.

Tabla 14. Resultados obtenidos de pruebas controladas para el escenario 1.

Vehículo	Ruta	Fecha	Hora Inicio	Hora Final	Nº. de Registros Obtenidos (Enviados por dispositivo)	Nº de datos predichos
2	1	23-dic-2022	2:51 p.m.	3:03 p.m.	31	25
2	2	23-dic-2022	3:14 p.m.	3:29 p.m.	33	27
2	3	27-dic-2022	9:41 p.m.	9:58 p.m.	33	44

La **Tabla 15** presenta los resultados obtenidos de las pruebas controladas forzando la ejecución del escenario 2.

Tabla 15. Resultados obtenidos de pruebas controladas para el escenario 2.

Vehículo	Ruta	Fecha	Hora Inicio	Hora Final	Nº. de Registros Obtenidos (Enviados por dispositivo)	Nº de datos predichos
2	1	27-dic-2022	10:09 p.m.	10:18 a.m.	32	13
2	2	27-dic-2022	10:27 p.m.	10:37 a.m.	33	12
2	3	27-dic-2022	10:46 p.m.	10:58 a.m.	32	11

6.2.2. Resultados de pruebas controladas representados en mapas

En las **secciones 6.2.2.1** y **6.2.2.2** son presentados los resultados de las pruebas controladas en un mapa para el escenario 1 (ausencia de comunicación *LoRa*) y escenario 2 (ausencia de conexión GPS) respectivamente.

La aplicación del *Smartphone* usada para obtener la posición real funcionó adecuadamente durante las pruebas controladas. A partir de los datos reales obtenidos se extraen las coordenadas correspondientes a cada coordenada predicha, según el tiempo en el que fue predicha esa coordenada.

6.2.2.1. Resultados en un mapa para el escenario 1 (Ausencia de LoRa)

Las Figuras 42, 43, y 44 presentan la información obtenida al ejecutar las pruebas controladas. Los datos registrados fueron graficados en un mapa, donde para diferenciar los registros enviados por el dispositivo a bordo del vehículo fue asignado un icono de color azul y para las predicciones de ubicación en ausencia de la comunicación LoRa un icono de color verde. Logrando de esta manera diferenciar el comportamiento del vehículo al realizar cada ruta. A continuación, la información recolectada y predicha para los respectivos vehículos.

- **Vehículo 2 (Recorrido de la Figura 38)**
 - Registros enviados por el dispositivo: 31
 - Registros predichos: 25
 - Error medio: 168 metros

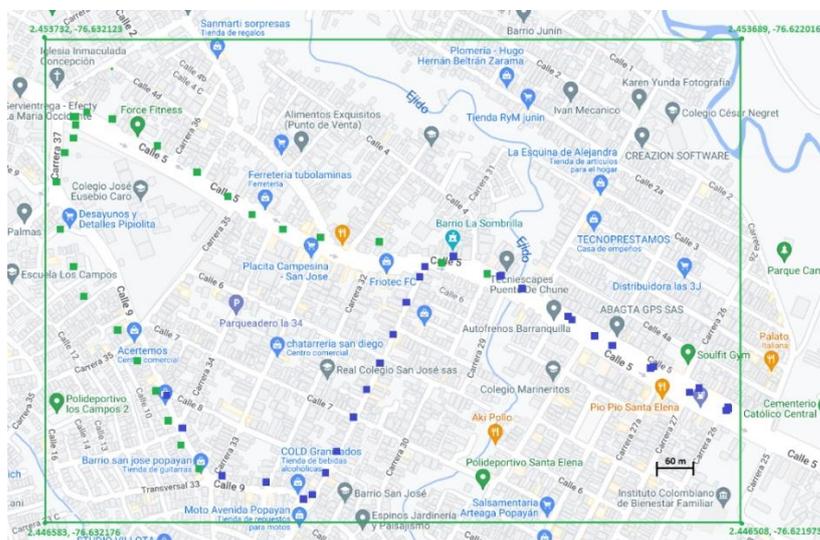


Figura 42. Registros en el mapa del escenario 1 para ruta 1 (mapa tomado de Google Maps).

La Tabla 16 presenta los datos predichos y datos reales obtenidos en las pruebas controladas sobre la ruta 1 en el escenario 1.

Tabla 16. Datos predichos en las pruebas controladas en la ruta 1 escenario 1.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
23-dic-22	14:58:10	2,450337	-76,625764	2,450176035	-76,62568539	19,92094516
23-dic-22	14:58:26	2,450452	-76,626531	2,450342062	-76,62635645	22,92926724
23-dic-22	14:58:34	2,450444	-76,627006	2,450651111	-76,6272706	37,35301151
23-dic-22	14:58:47	2,450571	-76,6279	2,450713968	-76,6281306	30,15870539
23-dic-22	14:59:00	2,450924	-76,62859	2,450829054	-76,62867827	14,41339571
23-dic-22	14:59:15	2,451369	-76,629403	2,451045354	-76,62912087	47,73646031
23-dic-22	14:59:27	2,451723	-76,630034	2,451312697	-76,62947346	77,21809231



23-dic-22	14:59:37	2,452084	-76,630672	2,451671104	-76,62992312	95,05009421
23-dic-22	14:59:48	2,452385	-76,631187	2,452052437	-76,63050896	83,93704766
23-dic-22	14:59:58	2,452407	-76,631223	2,452456116	-76,63116152	8,747633951
23-dic-22	15:00:09	2,452406	-76,631222	2,452551995	-76,63153373	38,25778715
23-dic-22	15:00:22	2,452418	-76,631256	2,452489007	-76,63170285	50,28056783
23-dic-22	15:00:31	2,452578	-76,631569	2,452482636	-76,63172781	20,58981508
23-dic-22	15:00:43	2,452651	-76,631707	2,452369072	-76,63170379	31,35983509
23-dic-22	15:01:04	2,45265	-76,631706	2,452171853	-76,63172815	53,23939312
23-dic-22	15:01:12	2,45265	-76,631706	2,451959798	-76,63185987	78,64983933
23-dic-22	15:01:25	2,45265	-76,631706	2,45152663	-76,6319781	128,5547993
23-dic-22	15:01:39	2,45265	-76,631706	2,450831912	-76,6319192	203,6022303
23-dic-22	15:01:50	2,45265	-76,631706	2,450283487	-76,63173475	263,2377304
23-dic-22	15:02:05	2,452547	-76,632009	2,449846954	-76,63157762	304,1178049
23-dic-22	15:02:22	2,451686	-76,631998	2,449354795	-76,63109223	278,1418933
23-dic-22	15:02:31	2,451343	-76,631922	2,448908025	-76,63080054	298,1294634
23-dic-22	15:02:43	2,450901	-76,631484	2,448470422	-76,63052509	290,5866549
23-dic-22	15:02:54	2,450511	-76,631208	2,448029668	-76,63034997	291,9956586
23-dic-22	15:03:06	2,449903	-76,630965	2,44767503	-76,63016764	263,1735816
23-dic-22	15:03:19	2,449467	-76,630774	2,447323244	-76,62989961	257,4799386
23-dic-22	10:15:57	2,449943	-76,625275	2,449604056	-76,62584657	73,86159484
23-dic-22	10:16:09	2,44994	-76,625263	2,449387076	-76,62538584	62,99627925
23-dic-22	10:16:19	2,449939	-76,625259	2,449183385	-76,62496048	90,35421481
23-dic-22	10:16:31	2,449829	-76,625052	2,449037404	-76,62460512	101,0851379
23-dic-22	10:16:43	2,449408	-76,624437	2,448973871	-76,62444375	48,29234471
23-dic-22	10:16:54	2,44907	-76,623946	2,448730608	-76,62426826	52,03296071
23-dic-22	10:17:05	2,448682	-76,623304	2,448683321	-76,62425543	105,7278548
23-dic-22	10:17:16	2,448301	-76,622671	2,448655797	-76,62442043	198,3702351
23-dic-22	10:17:27	2,448157	-76,622302	2,448655797	-76,62442043	241,8594761
23-dic-22	10:17:38	2,448184	-76,622306	2,448671218	-76,62458788	259,2987515
23-dic-22	10:17:50	2,448188	-76,622288	2,448872031	-76,62482348	291,8459738
23-dic-22	10:18:01	2,448413	-76,622617	2,449300114	-76,62509551	292,5641053
23-dic-22	10:18:13	2,448745	-76,623218	2,449733513	-76,62532478	258,6480984

- **Vehículo 2 (Recorrido de la Figura 40):**
 - Registros enviados por el dispositivo: 33
 - Registros predichos: 27
 - Error medio: 206 metros

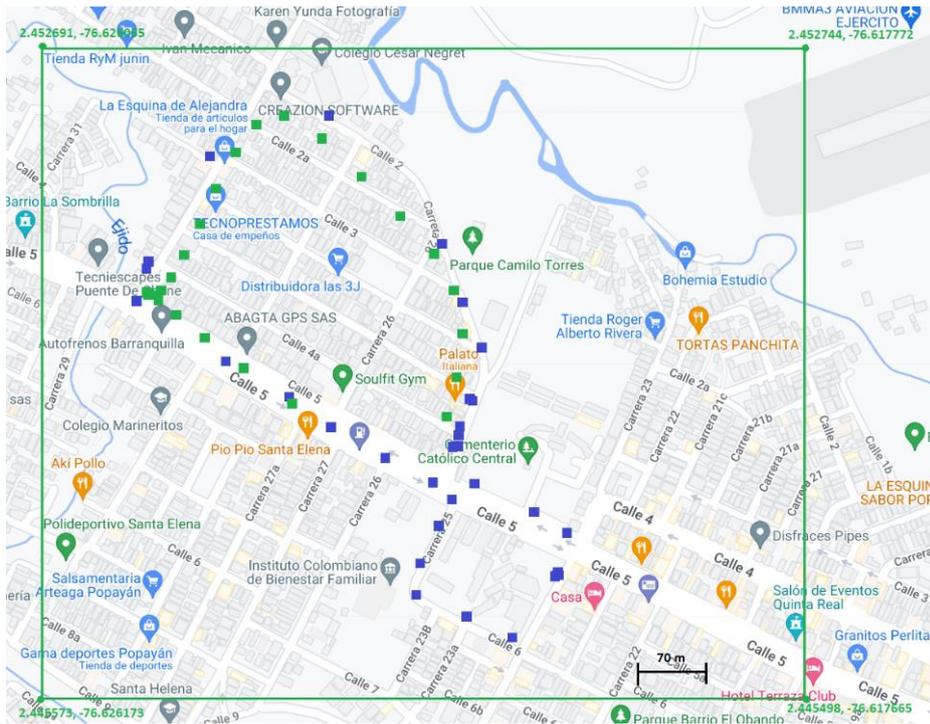


Figura 43. Registros en el mapa del escenario 1 para ruta 2 (mapa tomado de Google Maps).

La **Tabla 17** presenta las coordenadas predichas y coordenadas reales obtenidas en las pruebas controladas sobre la ruta 2 en el escenario 1.

Tabla 17. Datos predichos en las pruebas controladas en la ruta 2 escenario 1.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
23-dic-22	15:23:38	2,450053	-76,621476	2,44858524	-76,62163138	164,1640489
23-dic-22	15:23:51	2,450706	-76,621716	2,449012742	-76,62152273	189,5553977
23-dic-22	15:24:05	2,451302	-76,622026	2,449481544	-76,62146447	211,8793456
23-dic-22	15:24:18	2,451747	-76,622598	2,449951067	-76,62155584	230,8980461
23-dic-22	15:24:31	2,452118	-76,623234	2,450353576	-76,6217789	254,283437
23-dic-22	15:24:44	2,45231	-76,623571	2,450761271	-76,62213493	234,8190364
23-dic-22	15:24:58	2,451893	-76,623856	2,451191022	-76,62255767	164,0484251
23-dic-22	15:25:11	2,451452	-76,624116	2,451602221	-76,62299212	126,003025
23-dic-22	15:25:24	2,450944	-76,624409	2,451851461	-76,62339765	151,0567455
23-dic-22	15:25:37	2,450575	-76,624645	2,451761204	-76,62369661	168,8617306
23-dic-22	15:25:51	2,450253	-76,624833	2,451453135	-76,62391944	167,7043759
23-dic-22	15:26:04	2,450122	-76,624929	2,451057125	-76,62413212	136,6011901
23-dic-22	15:26:17	2,450126	-76,624926	2,450670608	-76,62431046	91,3676813
23-dic-22	15:26:30	2,450124	-76,624929	2,450342049	-76,62447723	55,75432162
23-dic-22	15:26:44	2,450125	-76,624929	2,450098298	-76,62461386	35,1456491
23-dic-22	15:26:57	2,450118	-76,624921	2,449959397	-76,62472353	28,15553987
23-dic-22	15:27:10	2,449817	-76,624994	2,449906486	-76,62480816	22,92519236
23-dic-22	15:27:24	2,449405	-76,62431	2,449905499	-76,62486152	82,79548233



23-dic-22	15:27:37	2,44905	-76,623757	2,449932435	-76,62488475	159,1815698
23-dic-22	15:27:50	2,448653	-76,623105	2,449923147	-76,62486135	240,9374792
23-dic-22	15:28:03	2,448276	-76,622449	2,449847673	-76,62475846	310,5191643
23-dic-22	15:28:17	2,447891	-76,621748	2,449685503	-76,6245607	370,8533422
23-dic-22	15:28:30	2,447626	-76,621651	2,449439407	-76,62424838	352,1237399
23-dic-22	15:28:43	2,447217	-76,621847	2,449109244	-76,62383649	305,2430131
23-dic-22	15:28:56	2,446817	-76,622047	2,448725235	-76,62330543	254,1738188
23-dic-22	15:29:10	2,44654	-76,621565	2,448353673	-76,62274783	240,7721011
23-dic-22	15:29:23	2,446285	-76,621095	2,447945442	-76,62206505	213,8421424

- **Vehículo 2 (Recorrido de la Figura 41):**
 - Registros enviados por el dispositivo: 33
 - Registros predichos: 44
 - Error medio: 474 metros



Figura 44. Registros en el mapa del escenario 1 para ruta 3 (mapa tomado de Google Maps).

La Tabla 18 presenta las coordenadas predichas y coordenadas reales obtenidas en las pruebas controladas sobre la ruta 3 en el escenario 1.

Tabla 18. Datos predichos en las pruebas controladas en la ruta 3 escenario 1.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
27-dic-22	09:50:16	2,450782	-76,627605	2,451218313	-76,62845494	106,1875332
27-dic-22	09:50:30	2,450682	-76,627653	2,451613457	-76,62901244	183,1799261
27-dic-22	09:50:40	2,450601	-76,627869	2,45206765	-76,62946812	241,2241825
27-dic-22	09:50:50	2,450916	-76,628105	2,452416918	-76,62987153	257,6920777



27-dic-22	09:51:01	2,451404	-76,628366	2,452678994	-76,63036041	263,1156135
27-dic-22	09:51:13	2,451746	-76,62864	2,45278595	-76,63092813	279,3406704
27-dic-22	09:51:21	2,452017	-76,628906	2,452763408	-76,63138657	287,8827215
27-dic-22	09:51:32	2,452344	-76,629331	2,452636744	-76,63155653	249,4444112
27-dic-22	09:51:45	2,4529296	-76,629443	2,452465393	-76,63137844	221,1850473
27-dic-22	09:51:53	2,45329	-76,629643	2,452271638	-76,63103402	191,6338213
27-dic-22	09:52:01	2,453658	-76,629883	2,451991899	-76,63067838	205,3131993
27-dic-22	09:52:11	2,453922	-76,630056	2,451660489	-76,63018752	251,9636661
27-dic-22	09:52:20	2,453916	-76,630294	2,451372586	-76,62974918	289,3003073
27-dic-22	09:52:30	2,453799	-76,630545	2,451119865	-76,62925594	330,6321962
27-dic-22	09:52:41	2,453704	-76,630732	2,45086176	-76,628753	385,1000206
27-dic-22	09:52:53	2,453437	-76,631128	2,450689211	-76,62823612	443,4850022
27-dic-22	09:53:03	2,453162	-76,631536	2,450567916	-76,62773174	511,8247391
27-dic-22	09:53:12	2,452866	-76,631963	2,450467026	-76,6272992	582,917619
27-dic-22	09:53:22	2,452553	-76,631832	2,450404688	-76,6269087	597,0037424
27-dic-22	09:53:31	2,452284	-76,631302	2,450337658	-76,62656586	569,0851122
27-dic-22	09:53:42	2,451959	-76,63064	2,450263531	-76,62625217	522,7927993
27-dic-22	09:53:53	2,451667	-76,630104	2,450158664	-76,62590554	495,7984816
27-dic-22	09:54:04	2,451321	-76,62946	2,450030249	-76,62560413	451,8930619
27-dic-22	09:54:16	2,45103	-76,628936	2,449898696	-76,62535068	417,8153452
27-dic-22	09:54:27	2,450726	-76,628363	2,449735385	-76,62505298	383,9729945
27-dic-22	09:54:38	2,450467	-76,62775	2,449534497	-76,62474572	349,5894342
27-dic-22	09:54:46	2,450415	-76,62724	2,449276404	-76,62437085	343,0640332
27-dic-22	09:55:09	2,450246	-76,62573	2,448452832	-76,62308965	354,7785516
27-dic-22	09:55:21	2,449975	-76,625278	2,448254046	-76,62274327	340,5562
27-dic-22	09:55:32	2,449946	-76,625262	2,448085547	-76,62240921	378,5750574
27-dic-22	09:55:44	2,449951	-76,625262	2,447909197	-76,62202938	424,9906335
27-dic-22	09:55:53	2,449952	-76,625264	2,447734747	-76,62163488	472,714423
27-dic-22	09:56:02	2,449952	-76,625263	2,44754108	-76,62120582	524,5728299
27-dic-22	09:56:13	2,449952	-76,625263	2,447251727	-76,6205998	598,9429224
27-dic-22	09:56:23	2,449953	-76,625264	2,446866081	-76,61981603	695,9889462
27-dic-22	09:56:36	2,449681	-76,624808	2,44649748	-76,61905083	731,2173915
27-dic-22	09:56:48	2,44926	-76,624184	2,44620218	-76,61836808	730,3210029
27-dic-22	09:57:00	2,448874	-76,623553	2,445927896	-76,61771911	726,4000403
27-dic-22	09:57:14	2,448391	-76,622806	2,445703816	-76,61715503	695,4640132
27-dic-22	09:57:26	2,447988	-76,622089	2,445528992	-76,61671625	656,7109736
27-dic-22	09:57:37	2,447732	-76,621582	2,445392091	-76,61633229	638,794843
27-dic-22	09:57:46	2,44764	-76,621253	2,445296458	-76,61599098	640,2092199
27-dic-22	09:57:53	2,447633	-76,621196	2,445243725	-76,61571955	664,0639879
27-dic-22	09:57:59	2,447582	-76,621036	2,445226263	-76,6155288	665,719398
27-dic-22	10:56:02	2,451494	-76,626998	2,450720501	-76,62728735	91,84566261
27-dic-22	10:56:13	2,451203	-76,627216	2,450665468	-76,62764878	76,72966443
27-dic-22	10:56:36	2,45062	-76,627674	2,450935161	-76,62772479	35,50558232
27-dic-22	10:56:48	2,450575	-76,627737	2,451258363	-76,62816968	89,93873963



27-dic-22	10:56:59	2,450584	-76,627763	2,451591263	-76,62835734	130,0524814
27-dic-22	10:57:11	2,450579	-76,627762	2,451965533	-76,62879477	192,235872
27-dic-22	10:57:21	2,450653	-76,62805	2,452604653	-76,62963543	279,5731601
27-dic-22	10:57:33	2,451161	-76,628284	2,452744445	-76,6303852	292,4689654
27-dic-22	10:57:54	2,451786	-76,628718	2,452578667	-76,63075676	243,1068612
27-dic-22	10:58:06	2,452224	-76,629122	2,45259458	-76,63111273	225,0262695
27-dic-22	10:58:17	2,452628	-76,629409	2,452508392	-76,63124838	204,8320169

6.2.2.2. Resultados en un mapa para el escenario 2 (Ausencia de GPS)

Las **Figuras 45, 46, y 47** presentan la información obtenida al ejecutar las pruebas controladas. Los datos registrados fueron graficados en un mapa, donde para diferenciar los registros enviados por el dispositivo a bordo del vehículo les fue asignado un icono de color azul y las predicciones de ubicación en ausencia de la comunicación GPS un icono de color rojo. Logrando de esta manera diferenciar el comportamiento del vehículo al realizar cada ruta. A continuación, la información recolectada y predicha para los respectivos vehículos.

- **Vehículo 2 (Recorrido de la Figura 38):**
 - Registros enviados por el dispositivo: 32
 - Registros predichos: 13
 - Error medio: 185 metros.

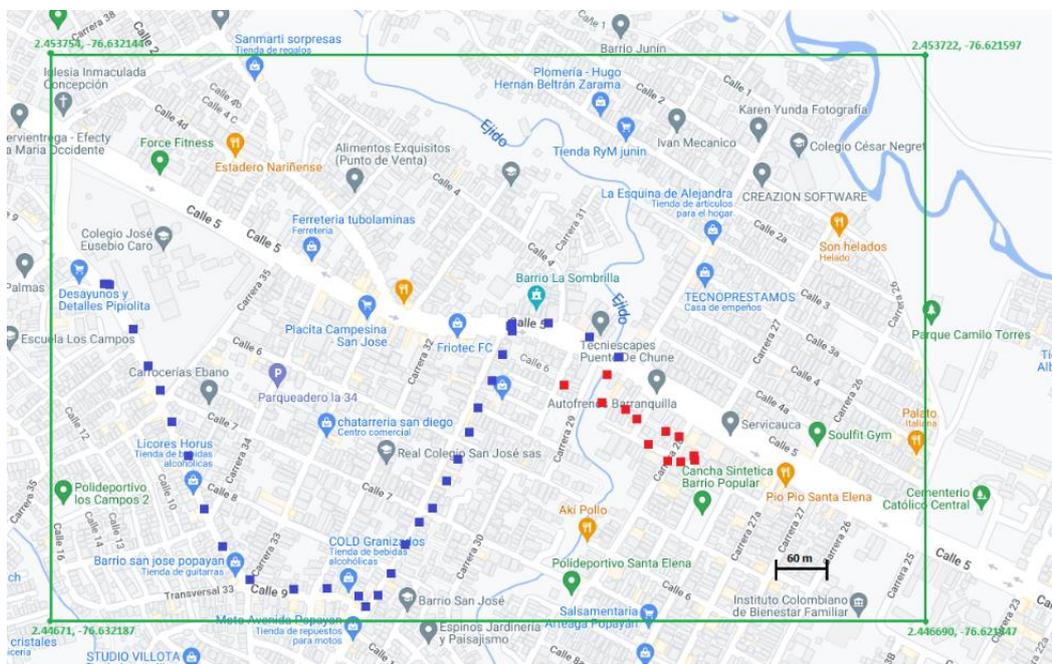


Figura 45. Registros en el mapa del escenario 2 para ruta 1 (mapa tomado de Google Maps).

La **Tabla 19** presenta las coordenadas predichas y coordenadas reales obtenidas en las pruebas controladas sobre la ruta 1 en el escenario 2.



Tabla 19. Datos predichos en las pruebas controladas en la ruta 1 escenario 2.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
27-dic-22	10:15:57	2,449943	-76,625275	2,449604056	-76,62584657	73,86159484
27-dic-22	10:16:09	2,44994	-76,625263	2,449387076	-76,62538584	62,99627925
27-dic-22	10:16:19	2,449939	-76,625259	2,449183385	-76,62496048	90,35421481
27-dic-22	10:16:31	2,449829	-76,625052	2,449037404	-76,62460512	101,0851379
27-dic-22	10:16:43	2,449408	-76,624437	2,448973871	-76,62444375	48,29234471
27-dic-22	10:16:54	2,44907	-76,623946	2,448730608	-76,62426826	52,03296071
27-dic-22	10:17:05	2,448682	-76,623304	2,448683321	-76,62425543	105,7278548
27-dic-22	10:17:16	2,448301	-76,622671	2,448655797	-76,62442043	198,3702351
27-dic-22	10:17:27	2,448157	-76,622302	2,448655797	-76,62442043	241,8594761
27-dic-22	10:17:38	2,448184	-76,622306	2,448671218	-76,62458788	259,2987515
27-dic-22	10:17:50	2,448188	-76,622288	2,448872031	-76,62482348	291,8459738
27-dic-22	10:18:01	2,448413	-76,622617	2,449300114	-76,62509551	292,5641053
27-dic-22	10:18:13	2,448745	-76,623218	2,449733513	-76,62532478	258,6480984

- **Vehículo 2 (Recorrido de la Figura 40):**
 - Registros enviados por el dispositivo: 33
 - Registros predichos: 12
 - Error medio: 350 metros.

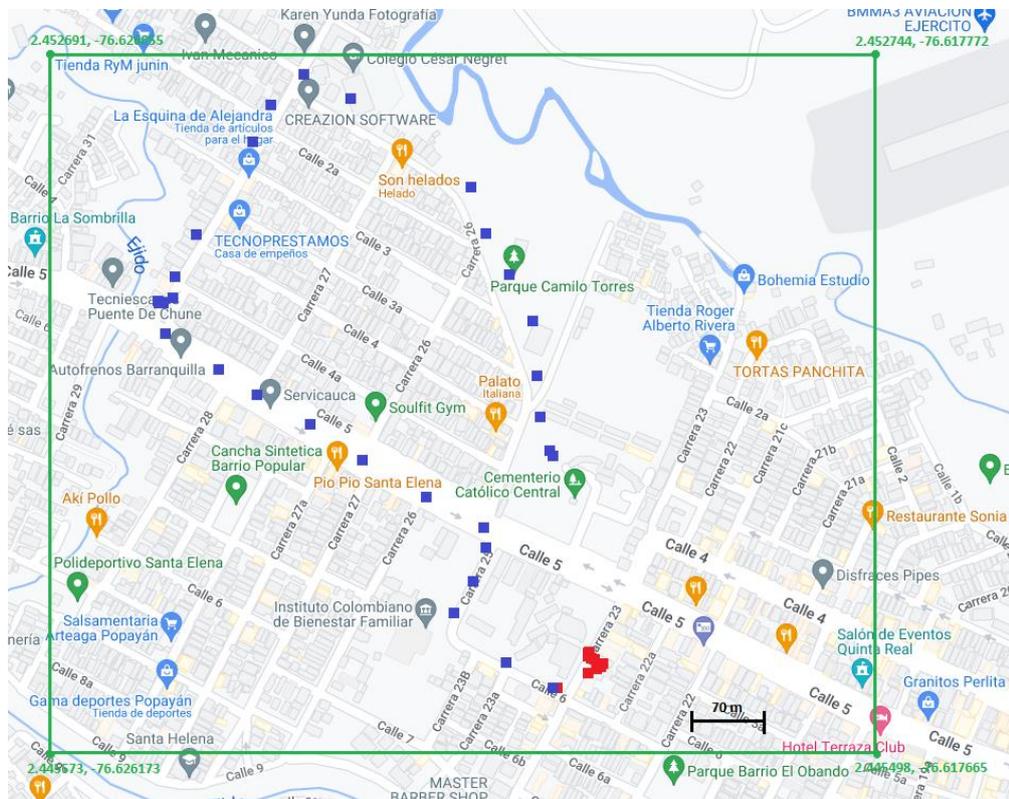


Figura 46. Registros en el mapa del escenario 2 para ruta 2 (mapa tomado de Google Maps).



La **Tabla 20** presenta las coordenadas predichas y coordenadas reales obtenidas en las pruebas controladas sobre la ruta 2 en el escenario 2.

Tabla 20. Datos predichos en las pruebas controladas en la ruta 2 escenario 2.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
27-dic-22	10:35:00	2,446997	-76,620398	2,446139347	-76,62092845	112,1366976
27-dic-22	10:35:11	2,447013	-76,620386	2,446286133	-76,62061644	84,80531639
27-dic-22	10:35:22	2,447111	-76,620341	2,446286133	-76,62061644	96,71791921
27-dic-22	10:35:34	2,447512	-76,62078	2,446328418	-76,62052647	134,626431
27-dic-22	10:35:44	2,447784	-76,621356	2,446354789	-76,62049068	185,786664
27-dic-22	10:36:06	2,448495	-76,621499	2,446390558	-76,62046676	260,664755
27-dic-22	10:36:18	2,44897	-76,621296	2,446428261	-76,6205529	294,5213655
27-dic-22	10:36:29	2,44941	-76,621335	2,446428261	-76,6205529	342,8464271
27-dic-22	10:36:41	2,449954	-76,621499	2,446442755	-76,62058229	403,6098239
27-dic-22	10:36:52	2,450472	-76,621687	2,446465453	-76,62061168	461,3762627
27-dic-22	10:37:09	2,451273	-76,622069	2,446490794	-76,62061104	556,0341047
27-dic-22	10:37:32	2,451898	-76,622984	2,446507416	-76,62060805	655,1331552

- **Vehículo 2 (Recorrido de la Figura 41):**
 - Registros enviados por el dispositivo: 32
 - Registros predichos: 11
 - Error medio: 189 metros

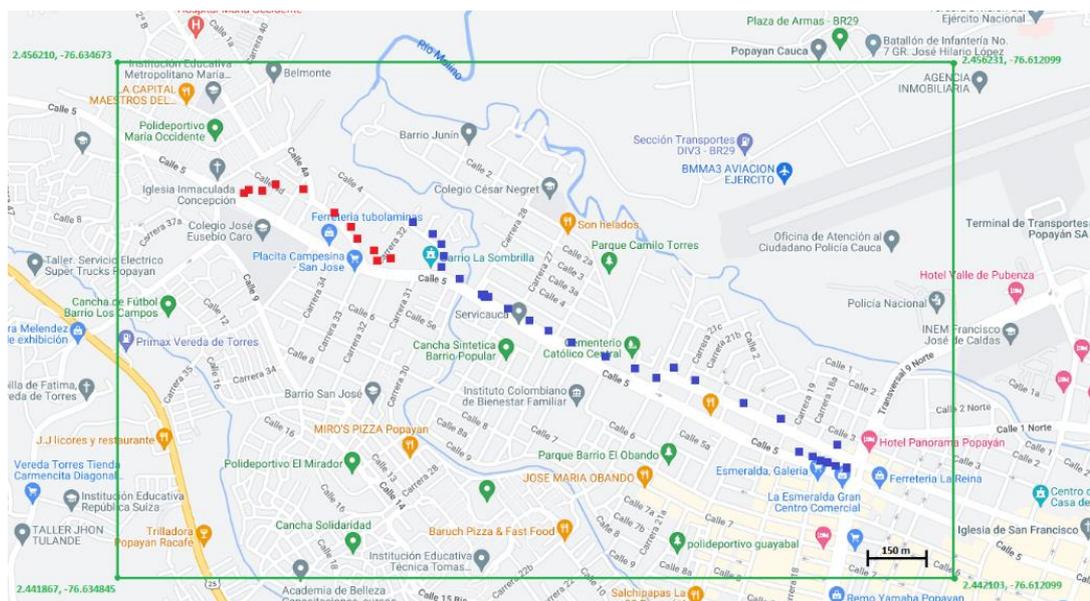


Figura 47. Registros en el mapa del escenario 2 para ruta 3 (mapa tomado de Google Maps).



La **Tabla 21** presenta las coordenadas predichas y coordenadas reales obtenidas en las pruebas controladas sobre la ruta 3 en el escenario 2.

Tabla 21. Datos predichos en las pruebas controladas para la ruta 3 escenario 2.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
27-dic-22	10:56:02	2,451494	-76,626998	2,450720501	-76,62728735	91,84566261
27-dic-22	10:56:13	2,451203	-76,627216	2,450665468	-76,62764878	76,72966443
27-dic-22	10:56:36	2,45062	-76,627674	2,450935161	-76,62772479	35,50558232
27-dic-22	10:56:48	2,450575	-76,627737	2,451258363	-76,62816968	89,93873963
27-dic-22	10:56:59	2,450584	-76,627763	2,451591263	-76,62835734	130,0524814
27-dic-22	10:57:11	2,450579	-76,627762	2,451965533	-76,62879477	192,235872
27-dic-22	10:57:21	2,450653	-76,62805	2,452604653	-76,62963543	279,5731601
27-dic-22	10:57:33	2,451161	-76,628284	2,452744445	-76,6303852	292,4689654
27-dic-22	10:57:54	2,451786	-76,628718	2,452578667	-76,63075676	243,1068612
27-dic-22	10:58:06	2,452224	-76,629122	2,45259458	-76,63111273	225,0262695
27-dic-22	10:58:17	2,452628	-76,629409	2,452508392	-76,63124838	204,8320169

6.2.3. Pruebas no controladas

La predicción fue ejecutada para los vehículos 2, 3 y 4 en el momento en el que la ruta más larga, la ruta 3, fuera completada 3 veces por parte del vehículo 4. Para todos los recorridos la predicción estuvo ejecutándose por 25 minutos, durante este tiempo la conexión con la base de datos estuvo estable, a diferencia de algunas pruebas anteriores donde el número de conexiones por hora del servidor era excedido (500 conexiones por hora).

6.2.3.1. Resultados de las pruebas no controladas

Para estas pruebas en total la ruta 1 fue recorrida 6 veces por los vehículos 1 y 2. La ruta 3 fue recorrida en el vehículo 4 un total de 4 veces y la ruta 2 fue recorrida en el vehículo 3 un total de 10 veces. Durante la ejecución de las pruebas solo fue presentada la ausencia de comunicación *LoRa* debido a que no hubo desconexión del GPS, lo que significa que solo entro en ejecución el escenario 1 de predicción. La **Tabla 22** muestra el número de registros obtenidos en las pruebas no controladas.

Tabla 22. Resultados obtenidos de las pruebas no controladas.

Vehículo	Ruta	Fecha	Hora Inicio	Hora Final	N° de Vueltas por recorrido	N° de Registros obtenidos (Enviados por dispositivo)	N° esperado de registros (Enviados por el dispositivo)	Cobertura (%)	N° de registros predichos
1	1	11-dic-2022	11:55 a.m.	1:37 p.m.	6	353	475	74.3	0



2	1	11-dic-2022	11:55 a.m.	1:37 p.m.	6	374	475	78.7	18
3	2	11-dic-2022	11:55 a.m.	1:37 p.m.	10	248	475	52.2	30
4	3	11-dic-2022	11:55 a.m.	1:37 p.m.	4	173	475	36.4	42

A partir de la **Tabla 22** fue obtenido un valor igual a 60.4% que corresponde al promedio de cobertura de todas las rutas cubiertas con solo dos *Gateway* y cuatro dispositivos en movimiento enviando paquetes al mismo tiempo.

6.2.4. Resultados de pruebas no controladas representados en mapas

En las **Figuras 48, 49, 50 y 51** es presentada la información obtenida al ejecutar las pruebas no controladas. Los datos registrados fueron graficados en un mapa, donde los registros enviados por el dispositivo a bordo del vehículo son mostrados con un icono de color azul y las predicciones de ubicación en ausencia de la comunicación *LoRa* con un icono de color verde. Logrando de esta manera diferenciar el comportamiento de cada vehículo al realizar su respectiva ruta. A continuación, la información recolectada y predicha para los respectivos vehículos.

La aplicación del *Smartphone* usada para obtener la posición real funcionó adecuadamente también durante las pruebas no controladas, al igual que para las pruebas controladas.

- **Vehículo 1:**
 - Registros enviados por el dispositivo: 353
 - Registros predichos: 0
 - La ruta fue recorrida en un total de 6 veces.

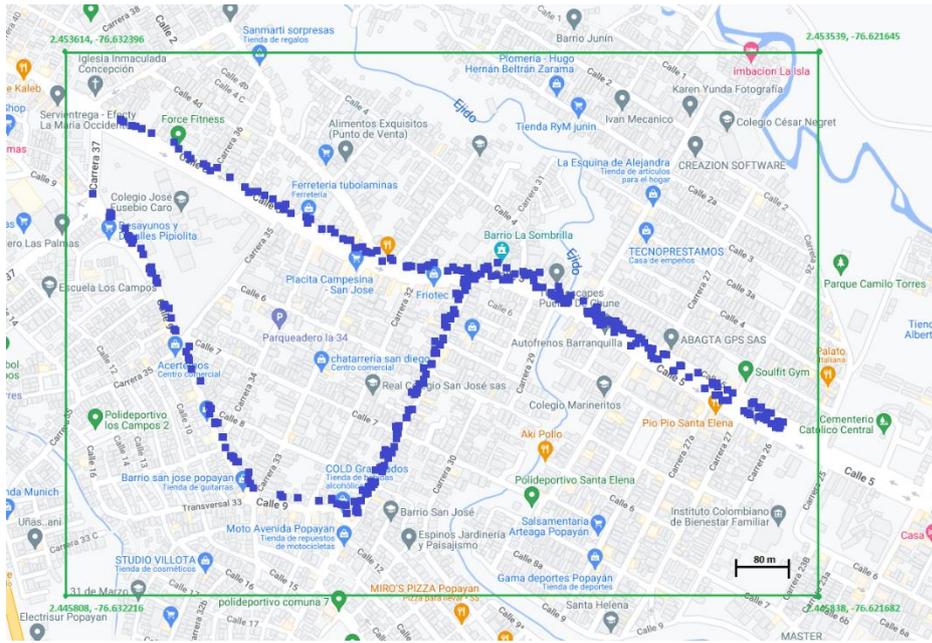


Figura 48. Registros en el mapa del recorrido del vehículo 1 (mapa tomado de Google Maps).

- **Vehículo 2:**
 - Registros enviados por el dispositivo: 374
 - Registros predichos: 18
 - La ruta fue recorrida en un total de 6 veces.
 - Error medio: 87 metros.

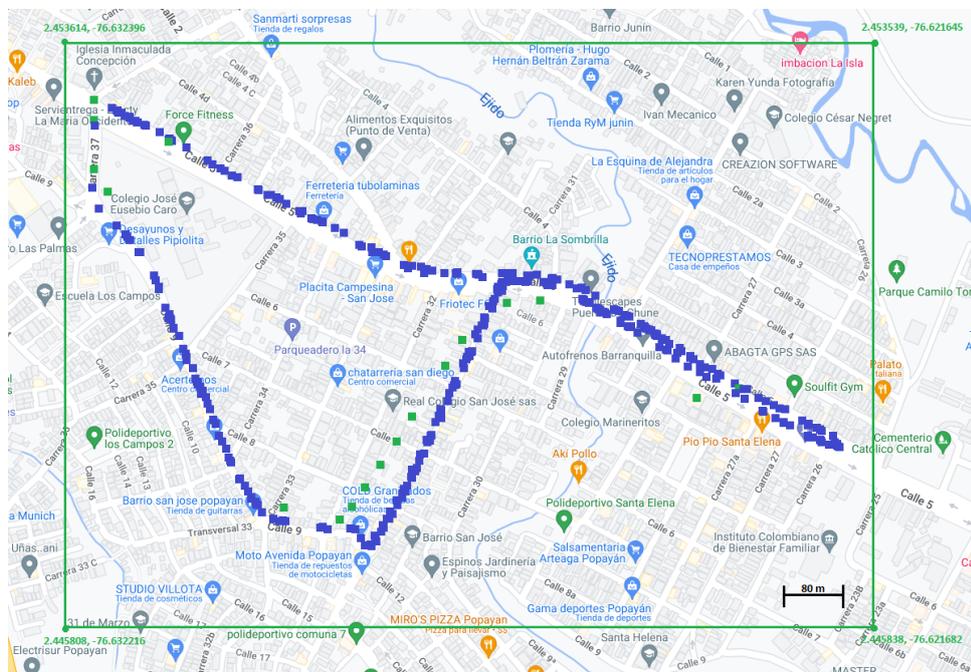


Figura 49. Registros en el mapa del recorrido del vehículo 2 (mapa tomado de Google Maps).



La **Tabla 23** presenta las coordenadas predichas y coordenadas reales en la ruta 1, esta ruta fue recorrida por dos vehículos haciendo dos distintos recorridos, pero solo fue realizada la predicción sobre uno de los vehículos.

Tabla 23. Datos predichos en las pruebas no controladas ruta 1.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
11-dic-22	13:09:18	2,448626	-76,623126	2,448809184	-76,62400609	99,89945125
11-dic-22	13:13:12	2,452503	-76,631433	2,452213993	-76,6309855	59,21317738
11-dic-22	13:14:07	2,451644	-76,632031	2,452503368	-76,6319781	95,76489901
11-dic-22	13:14:07	2,451644	-76,632031	2,452503368	-76,6319781	95,76489901
11-dic-22	13:14:15	2,451325	-76,631903	2,451857769	-76,6319781	59,84270995
11-dic-22	13:14:20	2,451158	-76,631731	2,451549898	-76,6317974	44,20947187
11-dic-22	13:16:30	2,446822	-76,628364	2,447348201	-76,62947037	136,1656306
11-dic-22	13:20:34	2,448593	-76,623075	2,44894016	-76,62344711	56,57634749
11-dic-22	13:26:11	2,451388	-76,632003	2,452775059	-76,6319781	154,3021948
11-dic-22	13:29:09	2,447055	-76,628912	2,447187768	-76,62873015	25,02901275
11-dic-22	13:29:26	2,446886	-76,628404	2,447619508	-76,62839515	81,59122986
11-dic-22	13:29:31	2,446801	-76,628317	2,447357794	-76,62856811	67,9262524
11-dic-22	13:29:56	2,44748	-76,627906	2,448233711	-76,62797872	84,22099424
11-dic-22	13:29:58	2,447581	-76,62786	2,447920826	-76,62819192	52,81229431
11-dic-22	13:30:46	2,449359	-76,627052	2,448562038	-76,62777368	119,5371627
11-dic-22	13:30:46	2,449359	-76,6270522	2,448562038	-76,62777368	119,5222533
11-dic-22	13:32:30	2,449415	-76,626979	2,449581432	-76,6271041	23,14997854
11-dic-22	13:32:35	2,449415	-76,62698	2,449237815	-76,62733382	43,98115696

En la **Tabla 24** son presentadas las coordenadas predichas y coordenadas reales obtenidas durante la ejecución de las pruebas no controladas en la ruta 1, junto con el error en metros para cada coordenada.

- **Vehículo 3:**
 - Registros enviados por el dispositivo: 248.
 - Registros predichos: 30.
 - La ruta fue recorrida en un total de 10 veces.
 - Error medio: 246 metros

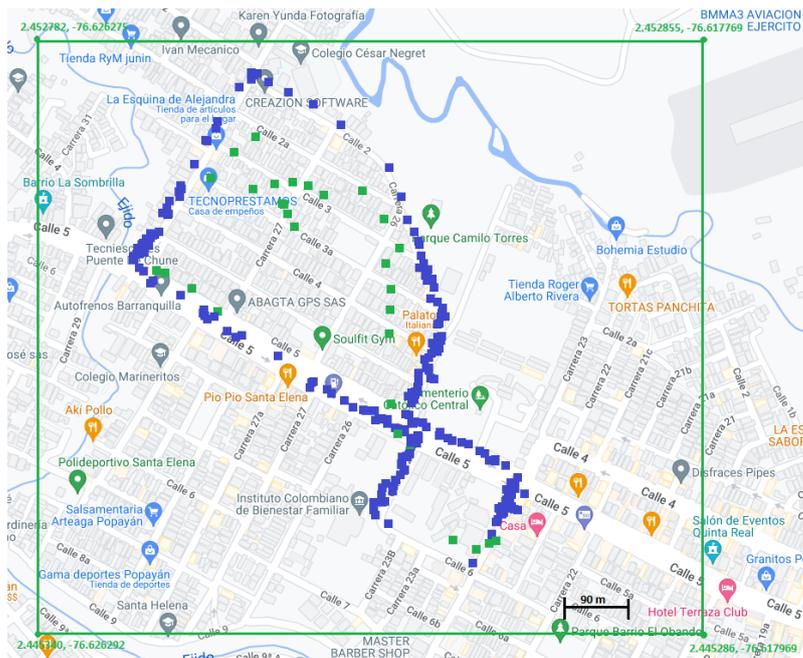


Figura 50. Registros en el mapa del recorrido del vehículo 3 (mapa tomado de Google Maps).

Tabla 24. Datos predichos en las pruebas no controladas ruta 2.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
11-dic-22	13:09:08	2,44654	-76,621634	2,4503589	-76,62306307	453,4774328
11-dic-22	13:09:21	2,446287	-76,621159	2,450521795	-76,62314522	520,1698988
11-dic-22	13:09:35	2,446257	-76,620746	2,450638643	-76,62319178	558,0160624
11-dic-22	13:09:48	2,446667	-76,620522	2,450654835	-76,6232086	534,6674414
11-dic-22	13:10:53	2,447435	-76,620541	2,448147391	-76,62185849	166,4726776
11-dic-22	13:12:35	2,451281	-76,621986	2,449025262	-76,62188045	251,1713429
11-dic-22	13:12:48	2,451599	-76,622386	2,449319448	-76,62187318	259,8714241
11-dic-22	13:13:02	2,451918	-76,622919	2,449581736	-76,6218583	285,3374291
11-dic-22	13:13:15	2,452232	-76,623413	2,44989639	-76,62192343	308,0350701
11-dic-22	13:13:51	2,451448	-76,624076	2,450755308	-76,62270116	171,1056285
11-dic-22	13:14:04	2,451157	-76,624272	2,450880792	-76,62288633	157,0164308
11-dic-22	13:14:18	2,451006	-76,62435	2,450916102	-76,62308377	141,0637151
11-dic-22	13:14:31	2,450729	-76,624505	2,450900589	-76,62331296	133,8324068
11-dic-22	13:14:44	2,450329	-76,624741	2,450837547	-76,62357844	141,0291567
11-dic-22	13:15:51	2,449347	-76,624256	2,449591838	-76,62434099	28,82378176
11-dic-22	13:17:56	2,446151	-76,620868	2,447765992	-76,62177148	205,7833194
11-dic-22	13:18:09	2,446353	-76,620698	2,44754951	-76,62162588	168,3537252
11-dic-22	13:20:53	2,451283	-76,621996	2,450090167	-76,62175734	135,2990783
11-dic-22	13:21:06	2,451611	-76,622421	2,450462046	-76,62194979	138,1056093
11-dic-22	13:21:20	2,451857	-76,6228	2,450813232	-76,62221262	133,1853431
11-dic-22	13:22:13	2,451573	-76,624023	2,451485427	-76,62354341	54,17708879
11-dic-22	13:22:26	2,45124	-76,624233	2,451298226	-76,62381038	47,4081822



11-dic-22	13:22:40	2,450786	-76,624496	2,450970911	-76,62409956	48,61893186
11-dic-22	13:24:06	2,449446	-76,624382	2,449820169	-76,62477467	60,3000823
11-dic-22	13:24:19	2,449042	-76,623793	2,449786071	-76,62466916	127,7842592
11-dic-22	13:24:57	2,447956	-76,621954	2,449303788	-76,62401678	273,893202
11-dic-22	13:26:55	2,44693	-76,622084	2,446440652	-76,62108849	123,2899822
11-dic-22	13:27:08	2,446924	-76,622054	2,446320619	-76,62080084	154,5852975
11-dic-22	13:27:22	2,446696	-76,621889	2,446393361	-76,62063976	142,8449847
11-dic-22	13:27:35	2,446511	-76,621547	2,446423849	-76,62055615	110,5335557

En la **Tabla 25** son presentadas las coordenadas predichas y coordenadas reales obtenidas durante la ejecución de las pruebas no controladas en la ruta 1, junto con el error en metros para cada coordenada.

Vehículo 4:

- Registros enviados por el dispositivo: 173
- Registros predichos: 42
- La ruta fue recorrida en un total de 4 veces.
- Error medio: 198 metros.



Figura 51. Registros en el mapa del recorrido del vehículo 4 (mapa tomado de Google Maps).

Tabla 25. Datos predichos en las pruebas no controladas en la ruta 3.

Fecha	Hora	Latitud real	Longitud real	Latitud predicha	Longitud predicha	Error[metros]
11-dic-22	13:09:02	2,445606	-76,616731	2,446636553	-76,61955376	333,9660445
11-dic-22	13:09:16	2,445585	-76,61668	2,446419721	-76,61911038	285,588432
11-dic-22	13:09:32	2,445458	-76,616271	2,446248252	-76,61871013	284,9439545
11-dic-22	13:09:38	2,445394	-76,616096	2,446112832	-76,61835185	263,1231178
11-dic-22	13:09:51	2,44529	-76,615759	2,445978481	-76,61800265	260,8205339
11-dic-22	13:10:01	2,445218	-76,615551	2,445867738	-76,61768633	248,0497078
11-dic-22	13:10:08	2,445172	-76,615394	2,445771079	-76,61740001	232,6640103



11-dic-22	13:10:19	2,445083	-76,615171	2,445685847	-76,61714491	229,3700333
11-dic-22	13:10:29	2,445078	-76,615148	2,445620007	-76,6169335	207,3698617
11-dic-22	13:10:40	2,445089	-76,615104	2,445568002	-76,61674124	189,5789961
11-dic-22	13:10:54	2,444899	-76,614717	2,445541089	-76,61657755	218,7406287
11-dic-22	13:11:03	2,445142	-76,61455	2,445535312	-76,61645856	216,5533742
11-dic-22	13:11:15	2,445408	-76,61482	2,445552495	-76,61640155	176,4830681
11-dic-22	13:11:23	2,445596	-76,615226	2,445615296	-76,61646203	137,3700988
11-dic-22	13:13:21	2,447857	-76,621435	2,446730294	-76,61906328	291,8342292
11-dic-22	13:15:15	2,45058	-76,625856	2,448922591	-76,62512743	201,3426608
11-dic-22	13:15:21	2,450724	-76,625851	2,449169089	-76,62541633	179,5655907
11-dic-22	13:15:35	2,451112	-76,625892	2,449467888	-76,62575184	183,5305013
11-dic-22	13:15:45	2,451312	-76,626072	2,449840969	-76,62613835	163,7833521
11-dic-22	13:16:00	2,45156	-76,626425	2,45027803	-76,62654803	143,2426803
11-dic-22	13:16:13	2,451775	-76,626719	2,450681874	-76,62691209	123,4631341
11-dic-22	13:16:25	2,451615	-76,62698	2,450944419	-76,62714157	76,71677655
11-dic-22	13:17:12	2,450619	-76,627733	2,451602449	-76,62820383	121,2543719
11-dic-22	13:18:26	2,453107	-76,629481	2,453285521	-76,62999612	60,58838286
11-dic-22	13:18:38	2,453649	-76,629841	2,453379202	-76,63076225	106,6813378
11-dic-22	13:18:49	2,453932	-76,630087	2,453320925	-76,63164909	186,4186249
11-dic-22	13:19:00	2,45385	-76,630399	2,453069739	-76,6319076	188,7739751
11-dic-22	13:19:12	2,453658	-76,630762	2,452763882	-76,6319076	161,5441325
11-dic-22	13:19:23	2,453457	-76,631087	2,452456216	-76,6319076	143,8959786
11-dic-22	13:20:13	2,453122	-76,631648	2,452272691	-76,63138042	99,03456413
11-dic-22	13:20:27	2,45307	-76,631789	2,45177884	-76,63061643	193,9138216
11-dic-22	13:21:10	2,453043	-76,631865	2,451253191	-76,62946917	332,4326481
11-dic-22	13:22:23	2,453043	-76,631865	2,451282135	-76,62923117	352,1681726
11-dic-22	13:22:53	2,452584	-76,631771	2,451474522	-76,62941381	289,5546456
11-dic-22	13:23:33	2,451726	-76,630142	2,451356147	-76,62911685	121,1188083
11-dic-22	13:23:42	2,451511	-76,629742	2,450782615	-76,62813461	196,1349041
11-dic-22	13:25:07	2,45013	-76,625455	2,449805531	-76,62598168	68,75948861
11-dic-22	13:25:35	2,449293	-76,624137	2,449329979	-76,62482307	76,350455
11-dic-22	13:26:01	2,44857	-76,623038	2,448775636	-76,6238127	89,07464782
11-dic-22	13:27:13	2,447412	-76,620553	2,447153882	-76,62055665	28,71235828
11-dic-22	13:28:26	2,446492	-76,618643	2,446077076	-76,61848976	49,19186834
11-dic-22	13:28:37	2,446146	-76,617988	2,445830653	-76,61782958	39,24504699

6.3. Datasets generados

La información recopilada en las pruebas realizadas fue almacenada en un *dataset* compuesto de tres partes en un formato csv. La primera parte contiene los datos de coordenadas recogidas en las tres rutas recorridas para entrenar los modelos encargados de predecir la ubicación. La segunda contiene los datos obtenidos durante la ejecución de las pruebas controladas. La tercera parte contiene la información obtenida durante la ejecución de las pruebas no controladas. Estos *datasets* pueden encontrarse en el **Anexo H. Dataset generado en las pruebas controladas y no**



controladas. El Anexo H contiene el enlace a la plataforma Kaggle (plataforma internacional de *datasets* de diversos temas), en donde se encuentra todo el detalle necesario del *dataset* (descripción, campos, datos, etc.).



Capítulo 7

7. Análisis de resultados obtenidos

7.1. Análisis de datos

7.1.1. Análisis de datos pruebas controladas

Al observar los resultados obtenidos en las pruebas controladas ha quedado en evidencia que, a pesar de tener distancias de separación considerables entre cada una de las ubicaciones reales y predichas, las predicciones son calculadas sobre las rutas previamente definidas o muy cercanas a estas (**Figuras 38, 40 y 41**). Esto quedó evidenciado en el escenario 1 donde al haber ausencia de comunicación *LoRa* las predicciones siguieron la ruta predefinida.

Adicionalmente, en las predicciones para el escenario 2 por ausencia de conexión GPS los datos predichos fueron localizados en un mismo punto o se devolvían (**Figuras 45 y 46**). Este tipo de predicción ocurrió porque al recolectar los datos para el entrenamiento del modelo el vehículo debía hacer paradas por un periodo de tiempo debido a los semáforos que estaban en cada una de las rutas. En dichos casos, cuando el modelo predice, asume que el vehículo está en una de estas paradas. Caso contrario sucedió en la **Figura 47** donde los datos predichos fueron localizados muy cerca de la ruta que debía seguir el vehículo, debido a que en ese tramo de predicción los semáforos estaban a una distancia considerable del automotor.

Al analizar los errores obtenidos entre las ubicaciones reales y predichas se llega a la conclusión que estas distancias significativas entre los puntos geográficos pueden deberse a la falta de sincronización del sistema, presentada cuando las condiciones de tráfico y por tanto velocidad media del vehículo son diferentes al momento de capturar los datos de entrenamiento y al momento de capturar los datos previos a la predicción.

7.1.2. Análisis de datos de las pruebas no controladas

Al observar los resultados obtenidos de esta prueba ha quedado en evidencia que el error cuadrático medio obtenido para el escenario 1 en comparación con las pruebas controladas fue mejor. Esto debido a que, en estas pruebas las interrupciones por ausencia de comunicación *LoRa* duraron poco tiempo y los datos predichos tenían en cuenta datos enviados por el dispositivo. Por el contrario, en las pruebas controladas fueron predichas ubicaciones con una mayor cantidad de valores debido a que la comunicación *LoRa* no estuvo presente un mayor tiempo.

Al igual que en las pruebas controladas cada uno de los datos predichos fue localizado dentro de la ruta establecida o muy cerca de ella, como muestran las **Figuras 49, 50 y 51**.

El sistema funciona de forma autónoma, al predecir coordenadas geográficas que hacían parte de cada uno de los recorridos, pero con la limitación que solo fue posible predecir para tres vehículos al mismo tiempo, debido a que en pruebas iniciales al ejecutar los cuatro algoritmos de predicción al tiempo era superado el límite de conexiones por ahora hacia el servidor porque cada uno de los algoritmos debían estar



consultando la base de datos para saber si debía ejecutar la predicción del escenario 1 o escenario 2.

7.1.3. Análisis general de resultados

Uno de los objetivos específicos del trabajo era analizar la eficiencia y eficacia del sistema de seguimiento de vehículo público. La eficiencia se refiere a cumplir con el objetivo de seguimiento de los vehículos, utilizando la menor cantidad de recursos posibles. En cuanto a la eficacia, se refiere a la capacidad de conseguir lo propuesto en el tiempo indicado. Es posible afirmar que el sistema logra la eficacia buscada, con la tecnología de comunicaciones planteada (*LoRa* y su protocolo de red *LoRaWAN*) entre los dispositivos dentro del vehículo y el *Gateway*; y, además, con el sistema de predicción de ubicación propuesto, ya que fue logrado el objetivo de mantener un seguimiento continuo del vehículo acercándose lo suficiente a la ubicación real (entre 87 y 246 metros). Respecto a la eficiencia, puede afirmarse que fue logrado igualmente debido a que dicha eficacia fue alcanzada buscando también el consumo de la menor cantidad de recursos posibles, en cuanto a inversión económica, recursos computacionales, y elementos HW y SW.

El porcentaje de paquetes que fueron recibidos en el *Gateway*, respecto a los enviados por parte de los dispositivos fueron de 60.4%, el cual es relativamente alto, además, para el 39.6% restante es usada la opción de predicción, logrando entregar para el 100% de la ruta una ubicación lo suficientemente cercana a la real.

La precisión promedio obtenida en las pruebas no controladas fue de 177 metros, dicho valor que es la diferencia entre la ubicación real (medida por la aplicación del Smartphone dentro del vehículo) y la ubicación predicha, es relativamente baja, para un sistema de transporte público, ya que en cuanto a tiempo de desplazamiento representaría algunos pocos segundos (aproximadamente de 10 a 15 segundos).

Adicionalmente, con los resultados de las pruebas no controladas, se logró demostrar que el sistema opera adecuadamente cuando existe más de un vehículo operando, lo cual representa un sistema de seguimiento de vehículo público, en el cual existen varios vehículos operando en una misma zona de la ciudad al mismo tiempo.

Otro aspecto importante con el que es demostrada la eficacia del sistema utilizando la tecnología *LoRaWAN*, es que los vehículos fueron enlazados adecuadamente a la red, independientemente de la distancia que los separaba del *Gateway* (siempre y cuando estuviesen dentro de la zona de cobertura).

Respecto a la limitación del número de algoritmos que pueden ejecutarse al mismo tiempo, es importante tener en cuenta que esto puede solucionarse mejorando las características operativas del servidor en la nube utilizado para esta tarea.

La optimización de recursos con la arquitectura propuesta para el sistema se logró mediante la evaluación de las posibles variables a medir, los sensores que logran medirlas, los dispositivos que están disponibles en el mercado para poder capturar los datos censados y una tecnología de comunicación que permitiera ser eficientes en cuanto a costos de operación y mantenimiento. Esta evaluación y posterior selección de los elementos del sistema fue realizada considerando aspectos de limitaciones en los recursos económicos disponibles. Adicionalmente, las demás herramientas y plataformas de desarrollo utilizadas fueron seleccionadas considerando aspectos como costos de licenciamiento, operación y mantenimiento.



7.2. Evaluación de los resultados finales

Fue complicado comparar el sistema propuesto con otros sistemas que utilizan tecnología de comunicación celular [22] o sistemas de comunicación LPWAN [10] debido a que estos artículos no presentan datos específicos en cuanto a la cantidad de paquetes recibidos y la exactitud de la predicción realizada, en [10] utilizan un híbrido entre *LoRaWAN* y NB-IoT para la comunicación y en [22] usan la red GSM/ GPRS.



Capítulo 8

8. Conclusiones y trabajos futuros

8.1. Conclusiones

Analizando los resultados obtenidos en las pruebas, es posible afirmar que las características de los actuales sistemas de seguimiento del servicio de transporte público colectivo pueden mejorar, mediante el desarrollo de un prototipo de un sistema para el seguimiento del servicio en tiempo real, basado en aprendizaje de máquina y *LoRaWAN*. Teniendo en cuenta las etapas de desarrollo del presente trabajo, es posible concluir lo siguiente:

Las redes neuronales LSTM y LSTM bidireccionales mostraron tener un rendimiento similar en el momento de ejecutar predicciones de ubicación y predicciones de tiempo, por lo que fue establecida una red LSTM para la predicción de ubicaciones y una red LSTM bidireccional para predecir el tiempo de esas ubicaciones.

El servidor en el cual se alojó la aplicación de seguimiento presentó limitaciones en cuanto a la cantidad de conexiones por hora, esto influyó en el periodo de tiempo durante el cual fue posible ejecutar los algoritmos de predicción para varias rutas al mismo tiempo, además influyó en la cantidad de rutas sobre las que fue aplicada la predicción.

Debido a la baja precisión de la velocidad en el dispositivo GPS, fue necesario incluir la aceleración ofrecida por el sensor IMU como entrada al algoritmo de predicción. Esto podría evitarse haciendo uso de un sensor GPS más preciso, pero esto elevaría los costos del dispositivo y limitaría la viabilidad económica.

Recolectando previamente datos inerciales y de ubicación GPS sobre una ruta completa es posible predecir la ubicación de un vehículo de transporte público colectivo durante suspensiones cortas de comunicación (de cinco minutos aproximadamente), a medida que la predicción supera esa barrera, la precisión disminuye de una manera más rápida.

Observando los puntos predichos en los diferentes recorridos, es posible visualizar que estos puntos están sobre la ruta y no dispersos aleatoriamente sobre el punto real, se considera que parte del error obtenido es debido a un problema de sincronización del tiempo predicho con las ubicaciones predichas.

Es conveniente el uso de la tecnología *LoRa* con su protocolo de red *LoRaWAN* para lograr una mayor seguridad, mejores distancias de alcance y poder operar varios dispositivos en los vehículos por cada *Gateway* existente.

Los valores de cobertura (60,4%) y precisión (177 metros de error promedio) obtenidos en las pruebas de campo finales del sistema fueron satisfactorios, considerando que el número de *Gateways LoRaWAN* utilizados en la prueba fue relativamente bajo (solo 2). En una implementación del sistema a gran escala, en una ciudad intermedia de un país en desarrollo, el número *Gateways* será mucho mayor y en sitios con mayor altura, lo cual mejoraría la cobertura de las rutas diseñadas, requiriendo mucho menos el proceso de predicción de ubicación de los vehículos.



En un sistema de seguimiento de vehículo público la adopción de la tecnología LoRaWAN reemplazando la comunicación celular (GSM, LTE) generaría ahorros considerables en la operación. Esto debido a que al utilizar LoRa no se requiere de un pago mensual a un proveedor de servicios de comunicaciones, por usar un rango de frecuencias de uso libre.

8.2. Trabajos futuros

En base al desarrollo del prototipo y los resultados presentados en el trabajo de grado, son propuestos los siguientes trabajos futuros relacionados con la investigación realizada:

- Modificar las características del algoritmo del sistema de seguimiento, que mejoren la precisión de las predicciones. Estas características pueden ser los propios hiperparámetros del modelo, el tamaño del conjunto de datos en la entrada del modelo de ML, o el número de capas de la red LSTM.
- Modificar la forma del entrenamiento del modelo de predicción haciendo que este modifique sus valores a medida que la información de posición vaya agregándose a la base de datos. Además, comprobar los efectos de este tipo de entrenamiento en la precisión de los resultados obtenidos.
- Mejorar las características operativas del servidor en la nube donde son alojados los algoritmos de predicción, con el objetivo de verificar si es posible solucionar las limitaciones de cantidad de vehículos a los que se les puede hacer seguimiento al mismo tiempo.
- Ampliar la cantidad de recursos utilizados en las pruebas de operación del sistema, incrementando el número de *Gateways*, el número de dispositivos de seguimiento, y el número de vehículos. Adicionalmente, incrementar las rutas y los tiempos de recorrido. Esto acercaría mucho más el ambiente de pruebas, al ambiente de producción de un sistema de seguimiento de vehículo público.
- Abordar el uso de *Edge Computing* en el módulo de recolección de datos, para lograr predecir la ubicación del vehículo cuando hay ausencia de la comunicación entre el GPS y el satélite.



Bibliografía

- [1] “Plan maestro movilidad Popayán.” https://drive.google.com/drive/folders/1C1khTb_mHZw7zA8m3B9txhmvCMIO8n5D (accessed Mar. 13, 2023).
- [2] “Forensis - Instituto Nacional de Medicina Legal y Ciencias Forenses.” <https://www.medicinalegal.gov.co/cifras-estadisticas/forensis> (accessed Jan. 08, 2023).
- [3] “Leading Transportation Analytics Solutions | INRIX.” <https://inrix.com/> (accessed Jan. 08, 2023).
- [4] “Global Traffic Scorecard | INRIX Global Traffic Rankings.” <https://inrix.com/scorecard/> (accessed Jan. 08, 2023).
- [5] “Así descubrimos ‘El ojo de Dios.’” <https://www.mintic.gov.co/portal/inicio/Sala-de-prensa/Noticias/103895:Asi-descubrimos-El-ojo-de-Dios> (accessed Jan. 08, 2023).
- [6] G. A. C. Caceres, “El impacto de la implementación de la tecnología en la seguridad vial en Colombia,” *Tecnología Investigación y Academia*, vol. 7, no. 2, pp. 28–44, Aug. 2020, Accessed: Jan. 08, 2023. [Online]. Available: <https://revistas.udistrital.edu.co/index.php/tia/article/view/14916>
- [7] W. Alexis, P. Calderón, O. Yazit, and S. García, “Implementación de una Red LPWAN de Sensores para el Monitoreo de Condiciones Ambientales.”
- [8] “LoRa PHY | Semtech.” <https://www.semtech.com/lora/what-is-lora> (accessed Jan. 08, 2023).
- [9] “LoRaWAN y LoRa.” <http://lorawan.es/> (accessed Jan. 08, 2023).
- [10] R. Mozny, M. Stusek, P. Masek, K. Mikhaylov, and J. Hosek, “Unifying multi-radio communication technologies to enable mMTC applications in B5G networks,” *2nd 6G Wireless Summit 2020: Gain Edge for the 6G Era, 6G SUMMIT 2020*, Mar. 2020, doi: 10.1109/6GSUMMIT49458.2020.9083791.
- [11] F. J. Murillo, J. S. Q. Yoshioka, A. D. V. López, R. Salazar-Cabrera, Á. P. de la Cruz, and J. M. M. Molina, “Experimental Evaluation of LoRa in Transit Vehicle Tracking Service Based on Intelligent Transportation Systems and IoT,” *Electronics 2020, Vol. 9, Page 1950*, vol. 9, no. 11, p. 1950, Nov. 2020, doi: 10.3390/ELECTRONICS9111950.
- [12] T. Boshita, H. Suzuki, and Y. Matsumoto, “IoT-based Bus Location System Using LoRaWAN,” *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, vol. 2018-November, pp. 933–938, Dec. 2018, doi: 10.1109/ITSC.2018.8569920.
- [13] R. Salazar-Cabrera, Á. Pachón De La Cruz, and J. M. Madrid Molina, “Proof of Concept of an IoT-Based Public Vehicle Tracking System, Using LoRa (Long Range) and Intelligent Transportation System (ITS) Services,” *Journal of Computer Networks and Communications*, vol. 2019, 2019, doi: 10.1155/2019/9198157.
- [14] “What is LoRaWAN® Specification - LoRa Alliance®.” <https://loralliance.org/about-lorawan/> (accessed Jan. 08, 2023).



- [15] “What is LoRaWAN® - LoRa Alliance®.” https://loralliance.org/resource_hub/what-is-lorawan/ (accessed Jan. 08, 2023).
- [16] A. Planes Martínez, “Estudio experimental de tecnologías LPWAN en entornos industriales 4.0,” Universitat Politècnica de València, Oct. 2020. Accessed: Jan. 08, 2023. [Online]. Available: <https://riunet.upv.es/handle/10251/152370>
- [17] Y. Yao, X. Xu, C. Zhu, and C. Y. Chan, “A hybrid fusion algorithm for GPS/INS integration during GPS outages,” *Measurement*, vol. 103, pp. 42–51, Jun. 2017, doi: 10.1016/J.MEASUREMENT.2017.01.053.
- [18] J. Sun and J. Kim, “Joint prediction of next location and travel time from urban vehicle trajectories using long short-term memory neural networks,” *Transp Res Part C Emerg Technol*, vol. 128, p. 103114, Jul. 2021, doi: 10.1016/J.TRC.2021.103114.
- [19] I. Belhajem, Y. Ben Maissa, and A. Tamtaoui, “A Hybrid Machine Learning Based Low Cost Approach for Real Time Vehicle Position Estimation in a Smart City,” in *Lecture Notes in Electrical Engineering*, vol. 397, Springer Verlag, 2017, pp. 559–572. doi: 10.1007/978-981-10-1627-1_44.
- [20] W. Ma, H. Xie, Y. Hu, C. Gao, J. Li, and R. Wang, “Vehicle Trajectory Prediction Method based on Deep Learning under the background of Internet of Vehicles,” *J Phys Conf Ser*, vol. 1881, no. 2, p. 022055, Apr. 2021, doi: 10.1088/1742-6596/1881/2/022055.
- [21] J. Yang, H. Bao, N. Ma, and Z. Xuan, “An Algorithm of Curved Path Tracking with Prediction Model for Autonomous Vehicle,” *Proceedings - 13th International Conference on Computational Intelligence and Security, CIS 2017*, vol. 2018-January, pp. 405–408, Feb. 2018, doi: 10.1109/CIS.2017.00094.
- [22] D. Sundararaman, G. Ravichandran, R. Jagadeesh, S. Sasirekha, I. Joe Louis Paul, and S. Swamynathan, “Customizable vehicle tracking with intelligent prediction system,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10595 LNCS, pp. 298–310, 2017, doi: 10.1007/978-3-319-68557-1_27/TABLES/4.
- [23] S. Choi, J. Kim, and H. Yeo, “Attention-based Recurrent Neural Network for Urban Vehicle Trajectory Prediction,” *Procedia Comput Sci*, vol. 151, pp. 327–334, Jan. 2019, doi: 10.1016/J.PROCS.2019.04.046.
- [24] L. Zhao, Y. Liu, A. Y. Al-Dubai, A. Y. Zomaya, G. Min, and A. Hawbani, “A Novel Generation-Adversarial-Network-Based Vehicle Trajectory Prediction Method for Intelligent Vehicular Networks,” *IEEE Internet Things J*, vol. 8, no. 3, pp. 2066–2077, Feb. 2021, doi: 10.1109/JIOT.2020.3021141.
- [25] “A Guide to the PROJECT MANAGEMENT BODY OF KNOWLEDGE (PMBOK® GUIDE) Sixth Edition”, Accessed: Jan. 08, 2023. [Online]. Available: www.PMI.org
- [26] M. N. Silva, “Uso de IoT y loNT en animales: Estudio de herramientas de recolección de datos para el análisis y comparación de datos vinculados a la sanidad animal”, Accessed: Jan. 08, 2023. [Online]. Available: http://sedici.unlp.edu.ar/bitstream/handle/10915/88952/Documento_completo.pdf-f-PDFA.pdf?sequence=1&isAllowed=y



- [27] Dña. M. M. Valenciano López, “Implementación de un radioenlace LPWAN con tecnología LoRa”, Accessed: Jan. 08, 2023. [Online]. Available: <https://uvadoc.uva.es/bitstream/handle/10324/57458/TFG-G5892.pdf?sequence=1&isAllowed=y>
- [28] A. Andreu Dólera Tutor and Á. Perles Ivars, “Modelado de redes LoRaWAN aplicadas a la conservación preventiva de patrimonio cultural”, Accessed: Jan. 08, 2023. [Online]. Available: www.etsit.upv.es
- [29] D. A. Jiménez Ballart, “Estudio de la cobertura de la modulación LoRa en la banda de 915[Mhz] en un ambiente urbano,” 2018, Accessed: Jan. 08, 2023. [Online]. Available: <https://repositorio.uchile.cl/handle/2250/168163>
- [30] “WIRELESS & SENSING Revision 1 July 2013 © 2013 Semtech Corporation 1 SX1272/3/6/7/8 LoRa Modem Design Guide,” 2013, Accessed: Jan. 08, 2023. [Online]. Available: https://www.openhacks.com/uploadsproductos/loradesignguide_std.pdf
- [31] “LTE-M deployment guide to basic feature set requirements,” 2019, Accessed: Jan. 09, 2023. [Online]. Available: <https://www.gsma.com/iot/wp-content/uploads/2019/08/201906-GSMA-LTE-M-Deployment-Guide-v3.pdf>
- [32] “Redes LPWAN: Una guía completa de inicio a fin”, Accessed: Jan. 09, 2023. [Online]. Available: <https://www.redgps.com/hubfs/Ebook%20-%20Redes%20LPWAN.pdf>
- [33] “Sigfox: ventajas y funcionalidades de la tecnología IoT - Productos IoT.” <http://productos-iot.com/sigfox-3/> (accessed Jan. 09, 2023).
- [34] B. L. Cevallos García and S. W. Rubio Echeverría, “Desarrollo de una red IoT con tecnología lora para gestión de invernaderos.” Accessed: Jan. 09, 2023. [Online]. Available: <https://dspace.ups.edu.ec/bitstream/123456789/20297/1/UPS%20-%20TTS365.pdf>
- [35] “Sigfox Technical Overview,” 2018, Accessed: Jan. 09, 2023. [Online]. Available: <https://www.avnet.com/wps/wcm/connect/onesite/03aebfe2-98f7-4c28-be5f-90638c898009/sigfox-technical-overview.pdf?MOD=AJPERES&CVID=magVa.N&CVID=magVa.N&CVID=magVa.N>
- [36] “¿Qué es sigfox? – SIGFOX.” <https://sigfox.com/py/que-es-sigfox/> (accessed Jan. 09, 2023).
- [37] A. M. Pinto Erazo, “Desarrollo de una metodología de análisis de tecnologías de radio IoT, para escenarios urbano y rural, por medio de modelos matemáticos,” 2020. Accessed: Jan. 09, 2023. [Online]. Available: <http://repositorio.utn.edu.ec/bitstream/123456789/10809/2/PG%20798%20TRA-BAJO%20GRADO.pdf>
- [38] J. Saiz Miranda, J. López Vicario, and X. Vilajosana Guillén, “Estudio en detalle de NB-IoT. Comparación con otras tecnologías LPWAN considerando diferentes patrones de tráfico,” 2019. <https://openaccess.uoc.edu/bitstream/10609/94946/6/jsaizmiTFM0619memoria.pdf> (accessed Jan. 09, 2023).
- [39] A. C. Müller and S. Guido, *Introduction to Machine Learning with Python : A Guide For Data Scientists*, First Edition. O’Reilly, 2016. Accessed: Jan. 20, 2023.



- [Online]. Available: <https://www.oreilly.com/library/view/introduction-to-machine/9781449369880/>
- [40] D. Sarkar, R. Bali, and T. Sharma, *Practical Machine Learning with Python*. Apress, 2018. doi: 10.1007/978-1-4842-3207-1.
- [41] I. Vasilev, D. Slater, G. Spacagna, P. Roelants, and V. Zocca, *Python Deep Learning 2nd*, Second edition., no. 2. 2019. Accessed: Jan. 20, 2023. [Online]. Available: <https://www.packtpub.com/product/python-deep-learning-second-edition/9781789348460>
- [42] P. Larranaga and A. Moujahid, “Tema 8. Redes Neuronales”, Accessed: Jan. 09, 2023. [Online]. Available: <https://www.researchgate.net/publication/268291232>
- [43] “Redes Neuronales Artificiales [Parte 2] – Desde mi Lab.” <https://desdemilab.com/redes-neuronales-artificiales-parte-2/> (accessed Mar. 13, 2023).
- [44] R. DiPietro and G. D. Hager, “Deep learning: RNNs and LSTM,” *Handbook of Medical Image Computing and Computer Assisted Intervention*, pp. 503–519, Jan. 2020, doi: 10.1016/B978-0-12-816176-0.00026-0.
- [45] A. Graves, N. Jaitly, and A. R. Mohamed, “Hybrid speech recognition with Deep Bidirectional LSTM,” *2013 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2013 - Proceedings*, pp. 273–278, 2013, doi: 10.1109/ASRU.2013.6707742.
- [46] “Guidelines for performing Systematic Literature Reviews in Software Engineering,” 2007. https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf (accessed Jan. 08, 2023).
- [47] B. J. Rojas Becerra and C. S. Bolaños Portilla, “Development validation of a fleet management and control system, based on Intelligent Transportation Systems (ITS) services and long-range technology (LoRa),” Universidad del Cauca, 2021. Accessed: Jan. 09, 2023. [Online]. Available: <http://repositorio.unicauca.edu.co:8080/xmlui/handle/123456789/5846>
- [48] “Arduino Hardware | Arduino.” <https://www.arduino.cc/en/hardware#boards-1> (accessed Jan. 09, 2023).
- [49] S. M. Casco, “Raspberry Pi, Arduino y Beaglebone Black Comparación y Aplicaciones,” 2014, Accessed: Jan. 09, 2023. [Online]. Available: <http://jeuazarru.com/wp-content/uploads/2014/10/MiniPCs.pdf>
- [50] “WIFI LoRa 32 (V2) (Phaseout) – Heltec Automation.” <https://heltec.org/project/wifi-lora-32/> (accessed Jan. 09, 2023).
- [51] A. E. Heredia Rivadeneria and P. F. Lucero Andrade, “Diseño e implementación de una red inalámbrica de sensores con tecnología LoRa para monitoreo industrial orientado a OPC de arquitectura unificada,” 2021. Accessed: Jan. 09, 2023. [Online]. Available: <https://dspace.ucuenca.edu.ec/bitstream/123456789/35875/1/Trabajo%20de%20Titulaci%C3%B3n.pdf>
- [52] “CubeCell – GPS-6502 – Heltec Automation.” <https://heltec.org/project/htcc-ab02s/> (accessed Jan. 09, 2023).



- [53] D. S. Escalante Cardona and D. M. Vargas Ospina, "Raspberry pi: la tecnología reducida en placa," 2019, Accessed: Jan. 09, 2023. [Online]. Available: <https://studylib.es/doc/9114041/raspberry-pi>
- [54] L. Pino Acosta, "Propuesta de Sistema SCADA código abierto para Sistemas Embebidos." Accessed: Jan. 09, 2023. [Online]. Available: <https://dSPACE.uclv.edu.cu/bitstream/handle/123456789/8220/Lisvet%20Pino%20Acosta.pdf?sequence=1&isAllowed=n>
- [55] M. A. Manrique Latorre, L. M. Buitrago Márquez, and J. Hernández-Gutiérrez, "Redes LoRaWAN. Revisión de componentes funcionales en aplicaciones IoT," 2019, Accessed: Jan. 09, 2023. [Online]. Available: <https://repository.udistrital.edu.co/bitstream/handle/11349/22411/LeidyMarcelaBuitragoMarquez2019.pdf?sequence=1&isAllowed=y>
- [56] "XBee™/XBee-PRO™ OEM RF Modules." <https://www.sparkfun.com/datasheets/Wireless/Zigbee/XBee-Manual.pdf> (accessed Jan. 09, 2023).
- [57] J. Cruz Núñez Pérez, A. Bonilla Rodríguez, and A. Calvillo Téllez, "Estimación del alcance de radiotransmisores Xbee," 2017, Accessed: Jan. 09, 2023. [Online]. Available: https://rcs.cic.ipn.mx/2017_142/Estimacion%20del%20alcance%20de%20radiotransmisores%20Xbee.pdf
- [58] J. A. Sánchez Wevar, "Análisis y Estudio de Redes GPRS." Accessed: Jan. 09, 2023. [Online]. Available: <http://cybertesis.uach.cl/tesis/uach/2005/bmficis211a/doc/bmficis211a.pdf>
- [59] R. M. Castilla Sierra and V. M. Meza Jiménez, "Descripción y evolución de tecnologías para redes de datos en ambiente GSM.," 2005. Accessed: Jan. 09, 2023. [Online]. Available: <https://biblioteca.utb.edu.co/notas/tesis/0032077.pdf>
- [60] J. C. Morales Davila, "Mejoramiento de comunicaciones móviles mediante la tecnología de GPRS." Accessed: Jan. 09, 2023. [Online]. Available: http://biblioteca.usac.edu.gt/tesis/08/08_0193_CS.pdf
- [61] u-blox, "NEO-6 u-blox 6 GPS Modules Data Sheet ." Accessed: Jan. 09, 2023. [Online]. Available: <https://manualzz.com/doc/23733551/neo-6-u-blox-6-gps-modules-data-sheet>
- [62] "NEO-7 u-blox 7 GNSS modules Data Sheet," 2014, Accessed: Mar. 13, 2023. [Online]. Available: www.u-blox.com
- [63] "u-blox NEO-M8 Data Sheet | Manualzz." <https://manualzz.com/doc/68134045/u-blox-neo-m8-data-sheet> (accessed Mar. 13, 2023).
- [64] InvenSense Inc, "MPU-6000 and MPU-6050 Product Specification Revision 3.4." 2013. Accessed: Jan. 09, 2023. [Online]. Available: <https://invensense.tdk.com/wp-content/uploads/2015/02/MPU-6000-Datasheet1.pdf>
- [65] "MPU-9250 Product Specification Revision 1.1 MPU-9250 Product Specification", Accessed: Mar. 13, 2023. [Online]. Available: <https://invensense.tdk.com/wp-content/uploads/2015/02/PS-MPU-9250A-01-v1.1.pdf>



- [66] “End Device Activation | The Things Network.” <https://www.thethingsnetwork.org/docs/lorawan/end-device-activation/> (accessed Jan. 22, 2023).
- [67] “ABP vs OTAA | The Things Stack for LoRaWAN.” <https://www.thethingsindustries.com/docs/devices/abp-vs-otaa/> (accessed Jan. 22, 2023).
- [68] “Long Short-Term Memory - QuantConnect.com.” <https://www.quantconnect.com/docs/v2/research-environment/applying-research/long-short-term-memory> (accessed Mar. 12, 2023).
- [69] “Kaggle: Your Machine Learning and Data Science Community.” <https://www.kaggle.com/> (accessed Jan. 09, 2023).
- [70] “Architecture Reference for Cooperative and Intelligent Transportation.” <https://www.arc-it.net/html/viewpoints/physical.html> (accessed Feb. 05, 2023).
- [71] Arduino, “What is Arduino? .” <https://www.arduino.cc/en/Guide/Introduction> (accessed Jan. 09, 2023).
- [72] Quency-D, “CubeCell-Arduino.” <https://github.com/HeITecAutomation/CubeCell-Arduino> (accessed Jan. 09, 2023).
- [73] Heltec-Aaron-Lee, “ESP32_LoRaWAN.” https://github.com/HeITecAutomation/ESP32_LoRaWAN (accessed Jan. 09, 2023).
- [74] Heltec-Aaron-Lee, “WiFi_Kit_series.” https://github.com/Heltec-Aaron-Lee/WiFi_Kit_series (accessed Jan. 09, 2023).
- [75] rfetick, “MPU6050_light.” https://github.com/rfetick/MPU6050_light (accessed Jan. 09, 2023).
- [76] mikalhart, “TinyGPSPlus.” <https://github.com/mikalhart/TinyGPSPlus> (accessed Jan. 09, 2023).
- [77] Dragino, “LPS8 LoRaWAN Gateway User Manual.” https://www.dragino.com/downloads/downloads/LoRa_Gateway/LPS8/LPS8_LoRaWAN_Gateway_User_Manual_v1.3.2.pdf (accessed Jan. 09, 2023).
- [78] R. Salazar Cabrera, “Manual Para Lograr Una Adecuada Operación, Mediante El Protocolo LoRaWAN y La Plataforma TTN, De Un Gateway Lps8 y Dispositivos Esp32 LoRa Heltec.” 2020.
- [79] “The Things Network.” <https://www.thethingsnetwork.org/> (accessed Jan. 09, 2023).
- [80] The Things Stack, “Webhooks.” <https://www.thethingsindustries.com/docs/integrations/webhooks/> (accessed Jan. 09, 2023).
- [81] Hostinger, “Alojamiento Web .” <https://www.hostinger.co/> (accessed Jan. 09, 2023).
- [82] J. Pavón Mestras, “PHP y MySQL.” Accessed: Jan. 09, 2023. [Online]. Available: <http://www.dominio.com/phpMyAdmin/setup>



- [83] A. E. García, "Manual Práctico de SQL", Accessed: Jan. 09, 2023. [Online]. Available: <https://www.lawebdelprogramador.com/cursos/archivos/ManualPracticoSQL.pdf>
- [84] "What is CRISP DM? - Data Science Process Alliance." <https://www.datascience-pm.com/crisp-dm-2/> (accessed Feb. 02, 2023).



Anexo A. Artículo de investigación desarrollado.

El Anexo A presenta el artículo científico desarrollado como parte de la elaboración de este trabajo de grado, el artículo fue presentado en la revista *Sustainability*.

Disponible en el siguiente enlace:

https://docs.google.com/document/d/1KEBsRHU_YB2Yh60-G1hafivvwD4BN-pJ/edit?usp=sharing&oid=114117981624372996966&rtpof=true&sd=true



Anexo B. Pruebas iniciales de los algoritmos.

El Anexo B presenta las pruebas iniciales realizadas para comprobar el funcionamiento del algoritmo de predicción (sin incluir la característica de tiempo real).

Disponible en el siguiente enlace:

<https://docs.google.com/document/d/1FKiQBuD1DWqziNdV4HHpNiLi3C2BToSf/edit?usp=sharing&oid=114117981624372996966&rtpof=true&sd=true>



Anexo C. Algoritmos generados de las pruebas iniciales.

El Anexo C presenta los archivos generados al desarrollar el algoritmo usado en las pruebas iniciales desarrollado en Python sobre el entorno Jupyter Notebook.

Disponible en el siguiente enlace:

<https://drive.google.com/drive/folders/1-VxdqSjeumuHg4ARVU8vR3PXh-kiTAax?usp=sharing>



Anexo D. Algoritmos generados en las pruebas controladas y no controladas.

El Anexo D presenta los algoritmos y archivos generados y usados en la ejecución de las pruebas finales de este proyecto.

Disponible en el siguiente enlace:

https://github.com/Andersson-Diaz/location_predict_LSTM_LoRa.git



Anexo E. Modelos de *Machine Learning* usados en las pruebas controladas y no controladas.

El Anexo E presenta la descripción de los modelos usados en las pruebas en tiempo real realizadas en este trabajo.

Disponible en el siguiente enlace:

<https://docs.google.com/document/d/11SxyLQd1yNyLtLxxEJTIU221rPSfkzQ7/edit?usp=sharing&oid=114117981624372996966&rtpof=true&sd=true>



Anexo F. Aplicación Web desarrollada para el seguimiento de las pruebas realizadas.

El Anexo F presenta una descripción del funcionamiento de la aplicación web desarrollada para el seguimiento de las pruebas realizadas.

Disponible en el siguiente enlace:

https://docs.google.com/document/d/1dPioxWo_6uOcdllm_eCvp4bl7tYodoir/edit?usp=sharing&oid=114117981624372996966&rtpof=true&sd=true



Anexo G. Códigos de la aplicación Web desarrollada para el seguimiento de las pruebas realizadas.

El Anexo G presenta la carpeta y subcarpetas de los diferentes archivos desarrollados para el funcionamiento de la aplicación web.

Disponible en el siguiente enlace:

<https://drive.google.com/drive/folders/1B7Vsl3FWFI55wDi63qvoCiiLCBHN3PVk?usp=sharing>



Anexo H. *Dataset* generado en las pruebas controladas y no controladas.

El Anexo H presenta el *dataset* generado durante la ejecución de las pruebas controladas y no controladas, estos datos son presentados en un formato csv y alojados en la plataforma Kaggle.

Disponible en el siguiente enlace:

<https://www.kaggle.com/datasets/anderssonstivediaz/vehicle-tracking-using-LoRaWAN>