

Estudio del trafico en tiempo real a traves de la continua extraccion y analisis de datos.

Yesid Santiago Carrillo Almeida
Edgar Ronaldo Henao Villarreal
Brayan Andrés Macías Turmequé

Escuela Colombiana de Ingeniería Julio Garavito

Departamento de Ingeniería de Sistemas

September 2022

1 Introducción

Una ciudad inteligente es el concepto de ciudad automatizada por medio de nuevas tecnologías en las cuales se busca la obtención de información completamente digitalizada por medio del uso de máquinas, objetos o servicios que transmiten datos de manera inalámbrica o por la nube, y el uso de tecnologías como Internet de las cosas. La idea primordial de una ciudad inteligente es el intercambio de datos continuo con procesos de análisis y gestión de datos en tiempo real.

2 Justificación

La urbanización es un fenómeno sin fin. Hoy en día, el 54% de las personas de todo el mundo vive en ciudades, una proporción que se espera llegue al 66% para 2050. En combinación con el crecimiento general de la población, la urbanización agregará otros 2,500 millones de personas a las ciudades en las próximas tres décadas. La sostenibilidad ambiental, social y económica es una necesidad fundamental para seguir el ritmo de esta rápida expansión que está desafiando a los recursos de nuestras ciudades.

Afortunadamente, más de 190 países han acordado objetivos para el crecimiento sostenible: la tecnología de ciudad inteligente es primordial para el éxito y el cumplimiento de esos objetivos.

3 Problema

Un problema latente en las ciudades con sobrepoblación como Bogotá, es que a diario se experimentan problemas en la movilidad ya que la gestión y administración de los medios de transporte o la infraestructura de la ciudad no fue planeada a largo plazo si no fue pensada a un corto plazo para solucionar una problemática que en su momento se pensó que iba a ser duradera. Esta situación ha dificultado que la población lleve a cabo sus actividades diarias hasta el punto de que se considera que se pierde más tiempo en el transporte que en el propio día, por las enormes distancias que muchas personas realizan para llegar a sus trabajos o estudios.

Por esta razón nuestra idea se basa en la recopilación de datos de la movilidad en Bogotá por medio de camaras, sensores, reductores de veloci-



Figure 1: Picture of the traffic in Bogota.

dad, aparatos móviles o aplicaciones que nos permitan tomar decisiones en tiempo real sobre las distintas situaciones que pueden ocurrir durante la jornada que dificulten la movilidad, como accidentes, movilizaciones sociales, etc, y de esta manera poder responder ante estos imprevistos de manera dinamica mejorando el flujo vehicular en ese momento.

La movilidad sostenible debe “satisfacer las necesidades actuales de transporte y movilidad sin comprometer la capacidad de las generaciones futuras para satisfacer sus propias necesidades” (Zito Salvo, 2011).

4 Estado del arte

Wi-Fi Direct como Estándar aplicado a Comunicaciones Vehiculares

En este escenario se tiene una investigación sobre el uso de la tecnología Wi-Fi Direct como soporte en el proceso de la comunicación inter-vehicular de forma inalámbrica.

En este caso, se equipa al vehículo con un teléfono móvil con Sistema Operativo Android, en el que está corriendo una aplicación diseñada para este tipo de comunicación. Dicha aplicación, es capaz de inferir el estado actual del tráfico teniendo en cuenta variables como el tiempo en el que está circulando lento o incluso llega a detenerse y el número de mensajes recibidos

de otros vehículos para, de esta forma, poder determinar cuándo se produce una incidencia en carretera. Cada vehículo es capaz de transmitir, a través del dispositivo embarcado, la información recopilada a los demás usuarios de la vía cuando estos están en su rango de cobertura.

Sistemas inteligentes de transporte (ITS) – Caso Zúrich, suiza

Este es un ejemplo de gestión integral sobre la operación completa del transporte público de la ciudad. Parte fundamental de este sistema es combatir los problemas ocasionados por la congestión del tráfico que aumenta el consumo de combustible, la contaminación, y el tiempo de viaje. Este sistema de gestión completa presenta múltiples ventajas, tales como: el incremento de la seguridad, el aumento de la eficiencia del tráfico, la contribución a un transporte sostenible, la mejora de la infraestructura de transportes, entre otros. Así, más puntualmente, el sistema implantado en la ciudad suiza desde hace más de 40 años se encarga de gestionar el transporte público de la misma, formado por tranvías, autobuses urbanos e interurbanos, trolebuses y trenes. Diariamente maneja información de un radio de 90.000 kilómetros y un trasiego de 800.000 pasajeros. Entre las tareas que administra se encuentran:

- Recuento de la recaudación procedente de los pasajes.
- Paneles de información en tiempo real de valor para el pasajero en las estaciones.
- Planificación de conductores y vehículos. Localización de vehículos públicos.
- Señalización del nivel de congestión de las carreteras en tiempo real.

ITS para el monitoreo automatizado del transporte público en tiempo real - México

El objetivo principal es proporcionar información en tiempo real a los usuarios que utilizan los autobuses como medio de transporte, y al mismo tiempo, proporcionar datos históricos de cada unidad para conocer la información relacionada con los horarios y retrasos de las rutas realizadas en los diferentes circuitos; para lograr el objetivo se ha utilizado un conjunto de recursos y medios tecnológicos que se centran en las capacidades de una computadora de placa reducida Raspberry PI. De manera general, el prototipo está formado

por cuatro módulos: el rastreador de GPS, que consta de dos componentes principales: una Raspberry PI y un módulo GPS NEO-6M Ublox. El servicio web y la base de datos se implementan en PostgreSQL, que almacena la información a través de un servicio web RestFul desarrollado en PHP. La aplicación web para la gestión de seguimiento de bus está desarrollada en PHP y utiliza el marco Bootstrap para la interfaz de usuario. Finalmente, la aplicación móvil para Android destinada a usuarios finales, se desarrolló en la plataforma de Visual Studio 2015 con Xamarin y se programó en C.

Smart Cities – España: Santander y La Coruña

En la ciudad de Santander se han desplegado miles de dispositivos que permiten conocer en tiempo real el estado de múltiples servicios tales como agua, residuos, movilidad, alumbrado, etc. Estos datos recogidos por los dispositivos son usados por el ayuntamiento de la ciudad para adecuar la gestión municipal a las necesidades de sus habitantes logrando una mayor eficiencia en el funcionamiento de los servicios públicos. Los ciudadanos pueden acceder a los datos recogidos por los dispositivos mediante aplicaciones móviles o bien mediante paneles informativos desplegados en la ciudad, lo que permite conocer en tiempo real, entre otras cosas, el estado del tráfico.

Respecto a la ciudad de la coruña, se ha desarrollado un proyecto centrado en cuatro áreas diferentes, el agua, la energía, el medio ambiente y la movilidad. Puntualmente en el área de movilidad, que es el que nos interesa, tenemos que cuenta con un sistema de optimización del tráfico en tiempo real mediante la frecuencia del transporte urbano, control de aparcamiento en zonas especiales y una aplicación que permite consultar el estado de las plazas de aparcamiento.

ETA prediction with Graph Neural Networks in Google Maps

En ese paper científico se habla de como como el servicio web de mapeo como lo es google maps es una herramienta sumamente poderosa utilizado en la actualidad sobre todo en las areas metropolitanas, sin embargo su implementacion no es nada faci ya que se debe tener en cuenta la informacion espacial (capturada con la red de la ruta) ademas de una informacion temporal (la evolucion de las condiciones del trafico a o largo del tiempo). En particular se requiere una anticipacion al tiempo demorado en recorrer un segmento de dstancia determinado y por supuesto a condiciones que pueden alterar el trafico como las configuraciones de los semaforo. Por supuesto tambien se debe anticipar a las posibles condiciones que se presenten en una

ruta escogida por un usuario un tiempo mas adelante de haberla recorrido en otras su fluctuacion con el tiempo.

Por eso se hicieron varios estudio como la identificacion de que hay una gran cantidad de datos históricos de viajes disponibles que pueden usarse para respaldar la detección y explotación de tales sutilezas. Como tal, este problema es sustancialmente susceptible a los enfoques de aprendizaje automático

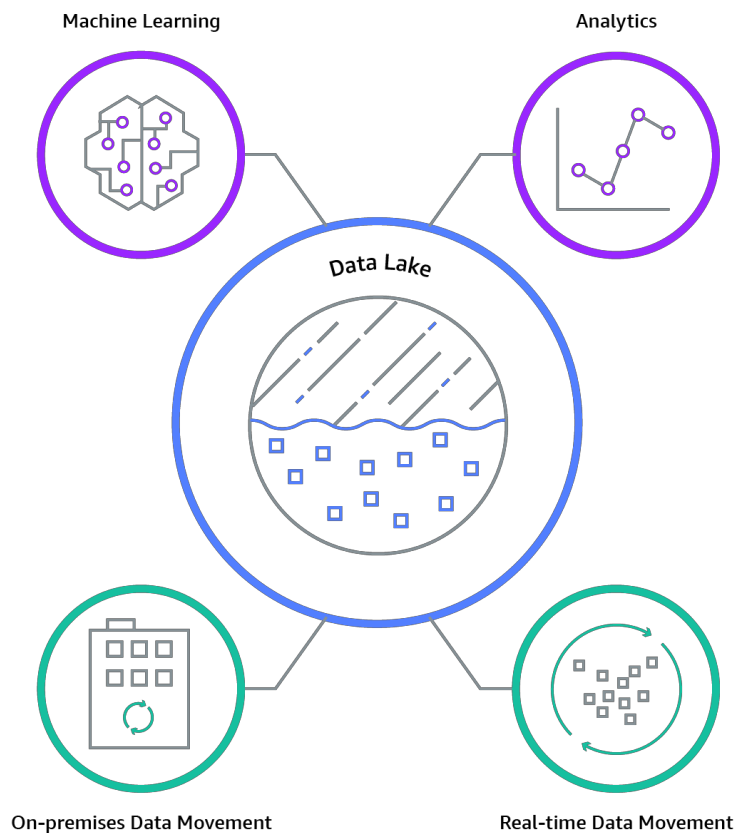
Finalmente se pudo ver cómo se puede diseñar una red neuronal gráfica para predecir con precisión las estimaciones de tiempo de viaje a lo largo de las rutas candidatas. La aplicación de técnicas de estabilización como Meta-Gradients y EMA fue una adición necesaria para que los GNN estén listos para la producción. Implementamos nuestro modelo GNN para la predicción de ETA en Google Maps, donde ahora atiende consultas de usuarios en todo el mundo. Presentamos métricas fuera de línea, evaluaciones en línea y estudios de usuarios: todos muestran mejoras cuantitativas significativas en el uso de GNN para las predicciones de ETA. Más lejos, a través de extensas aplicaciones en varios diseños y características, se brinda asesoramiento prescriptivo sobre el despliegue de este tipo de modelos en la práctica. Creemos que nuestros hallazgos serán beneficiosos para investigadores aplicando GNN a cualquier tipo de red de transporte análisis, que seguramente seguirá siendo una aplicación muy importante área para la representación gráfica del aprendizaje como un todo.

5 Marco Teórico

5.1 Data Lake

¿Qué es un lago de datos?

Un lago de datos es un repositorio centralizado que le permite almacenar todos sus datos estructurados y no estructurados a cualquier escala. Puede almacenar sus datos sin modificarlos y sin tener que estructurarlos primero. También puede ejecutar diferentes tipos de análisis: desde paneles y visualizaciones hasta procesamiento de macrodatos, análisis en tiempo real y machine learning para tomar mejores decisiones.



Data Lake

Fuente: <https://aws.amazon.com/es/big-data/datalakes-and-analytics/what-is-a-data-lake/>

Características

- **Datos:** No relacionales y relacionales provenientes de dispositivos de IoT, sitios web, aplicaciones móviles, redes sociales y aplicaciones corporativas.
- **Esquema:** Escrito al momento del análisis (esquema en lectura).
- **Precio - desempeño:** Resultados de consultas que se tornan más rápidos con almacenamiento de bajo costo.
- **Calidad de los datos:** Cualquier dato seleccionado o no (es decir, datos sin procesar).
- **Usuarios:** Científicos de datos, desarrolladores de datos y analistas de negocios (con datos seleccionados).
- **Análisis:** Machine learning, análisis predictivo, detección de datos y creación de perfiles.

6 Propuesta de Solución

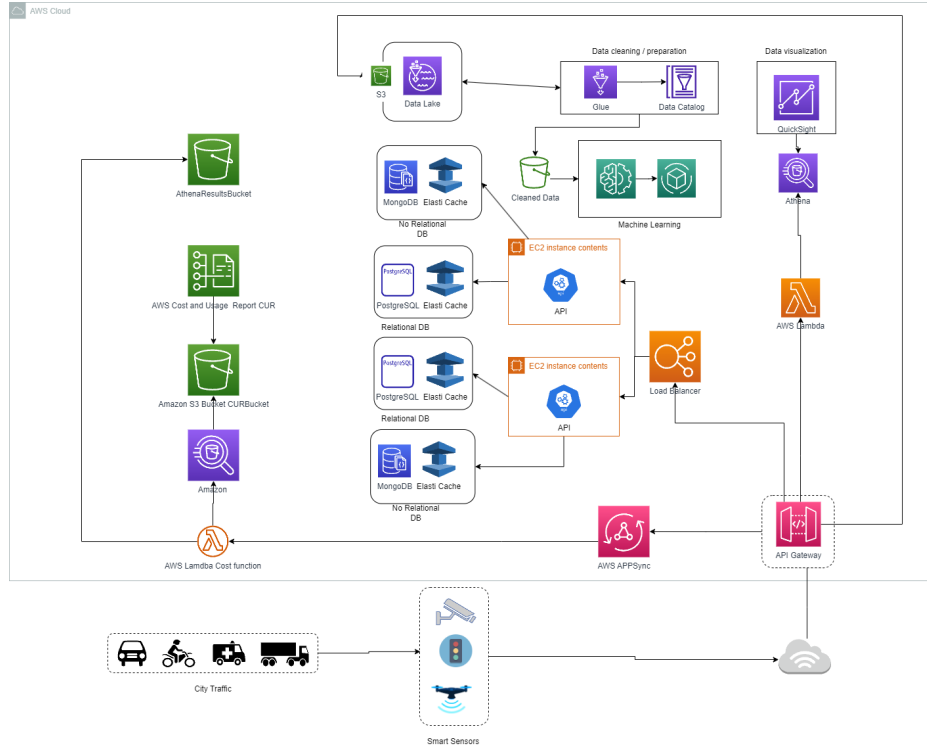


Figure 2: Arquitectura

Una herramienta construida en AWS que a través de una gestión unificada en distintos frentes va a poder dar una cobertura total al tema de la gestión vehicular en Bogotá para poder atacar dicha problemática, moderando su ineficiencia y falta de efectividad. Así que entendiendo que la problemática de la gestión del tráfico es un proceso que debe ser acompañado, vigilado y entendido no como un hecho inmutable y repetitivo a lo largo del tiempo, es que se propone tener una perspectiva dinámica sobre el tráfico vehicular, su ciclo de vida y el de todos los procesos que le rodean. De esta manera, es posible llegar a entender cuál es el nivel de influencia de cada uno de los aspectos relevantes y que son pertinentes a la problemática. Tener toda esta perspectiva es ideal a la hora de comprender la problemática de forma global e integrada, ya que estaremos en línea con nuestro enfoque de mantener un

ritmo dinámico en la gestión de la problemática.

Pero ¿a qué nos referimos realmente cuando decimos mantener un ritmo dinámico? La respuesta es que constantemente se está aprendiendo del sistema, ya que al estar tan cerca de su ciclo de ejecución, nos damos cuenta de cómo este cambia y cómo esto se ve reflejado en el día a día de la problemática, por lo que a diario estamos obteniendo datos sobre su comportamiento y el de los factores influyentes y también de los resultados obtenidos, permitiendo que la problemática pueda ser abordada desde la experiencia continua; así que finalmente todo esto se traduce en que podemos tomar decisiones basadas en información veraz y reciente y que estas decisiones están realmente atacando el problema en un punto específico del tiempo en que debe ser aplicado. Es importante destacar que el objetivo no es tener un proceso a prueba de fallas a la hora de gestionar la toma de decisiones, si no que se pretende fortalecer el proceso mismo de toma de decisiones haciéndolo más preciso y rápido, ya que para cada situación habrá peores o mejores decisiones y la importancia radica en que ante cada falla se pueda tener de manera veloz una nueva propuesta de solución, ya que probablemente una única solución no pueda mitigar el problema.

Ahora, para que todo este proceso pueda ser llevado a cabo, tenemos un gran aliado que es quien nos permite tener acceso a todos estos datos fundamentales: los dispositivos de IoT pertinentes a la infraestructura de Smart City de la ciudad, dispositivos como cámaras especiales y sensores localizados en puntos clave. Una vez todos estos datos son recopilados en tiempo real, son transmitidos directamente a la nube, en donde toda nuestra infraestructura nos permite tener una gestión adecuada de los datos, para que a través de ciertos procesos de análisis estos puedan generar la información necesaria para plantear soluciones reales a través de la toma de decisiones.

Más específicamente, la ruta de planificación a seguir para lograr el objetivo propuesto se basa en la recolección de datos de todo tipo de automóviles, carros, motos, camiones, buses, drones, etc, por medio de sensores inteligentes como lo son las cámaras de seguridad o los semáforos inteligentes para ver el flujo vehicular y como este varía frente a la situación que se presente o la hora del día en la que se este. Luego de eso los datos son capturados a través de la red y son enviados al servicio API Gateway el cual funciona como un intermediario entre lo que está publicado en internet y la lógica de negocio,

en este punto se toman 4 caminos dependiendo lo que se quiere hacer, por el lado de la aplicacion hay un balanceador de cargas que permite la alta disponibilidad de la logica del negocio el cual soporta 2 maquinas virtuales en el servicio de EC2 al igual que su respectivo manejo de base de datos relacional, no relacional y un almacenamiento cache con elastiCache, por otro lado se lleva a cabo una sincronizacion con AWS APPSync mediante el uso de una function Lambda para poder manejar los costos de la operacion, aqui es donde se utilizan servicios como Amazon Athena, Amazon S3 Bucket CUR-Bucket, AWS Cost and Usage Report CUR para lograr el cometido.

En otro punto de conexion del Api Gateway los datos se guardan en buckets de S3 con la ayuda de un Data Lake para su almacenamiento, por otra parte los servicios disponibles pasan por AWS Lambda para permitir su escalabilidad y posterior analisis de datos mediante la herramienta Athena.

La informacion en el Data Lake pasa por una limpieza de datos gracias al servicio Glue y su categorizacion con Data catalog, despues se empaqueta una parte de la informacion recopilada sobre todo la que no se necesita una solucion a corto plazo hacia un bucket centralizado para poder enviarlo de una forma mas portable hacia un entorno de Machine Learning para su tratamiento y de esa forma lograr obtener las soluciones mas optimas a los problemas que se puedan presenta en la movilidad en tiempo real.

Por ultimo la informacion es enviado a Athena al igual que lo que sale de los microservicios con Lambda para su analisis y finalmente se lleva un bosquejo grafico de una gran cantidad de datos con la ayuda de QuickSight para ver los datos mas criticos que requieren de una revision mas rapida y eficaz.

7 Prototipo

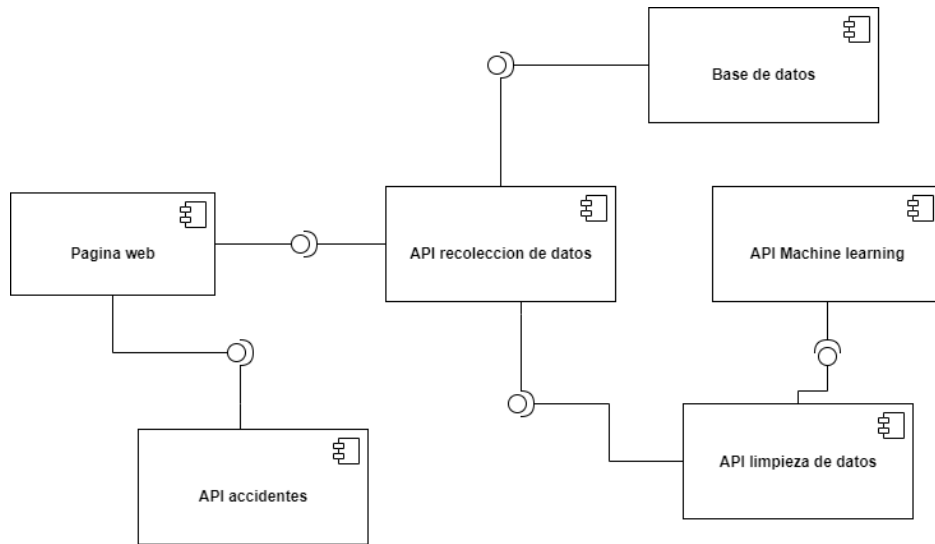


Figure 3: Prototipo

Nuestra propuesta de prototipo es una página web que tenga varios APIs, uno de ellos es el de la recoleccion de datos que por medio de los dispositivos externos tales como camaras, sensores de movimiento, entre otros se encargaran de recolectar datos tales como velocidad, numero de vehiculos registrados, flujo de peatones, accidentes automovilisticos, eventos importantes, posibles manifestaciones, entre otros. El resto de APIs, se encargaran de adecuar la informacion para que el API Machine Learning haga un análisis detallado de posibles soluciones para plantear rutas de planificacion. La informacion sera almacenada en una base de datos PostgreSQL (Ver figura 3)

8 Evaluacion

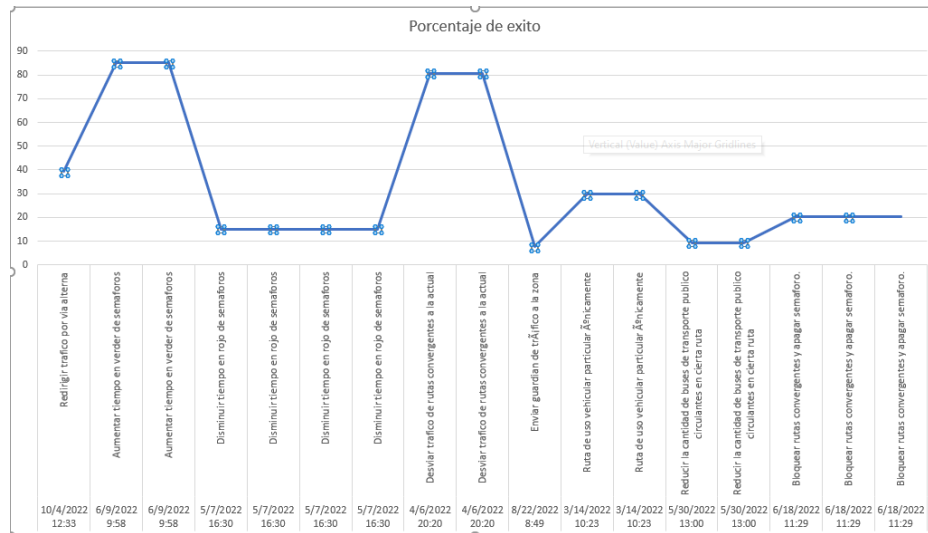


Figure 4: Grafico porcentaje exito

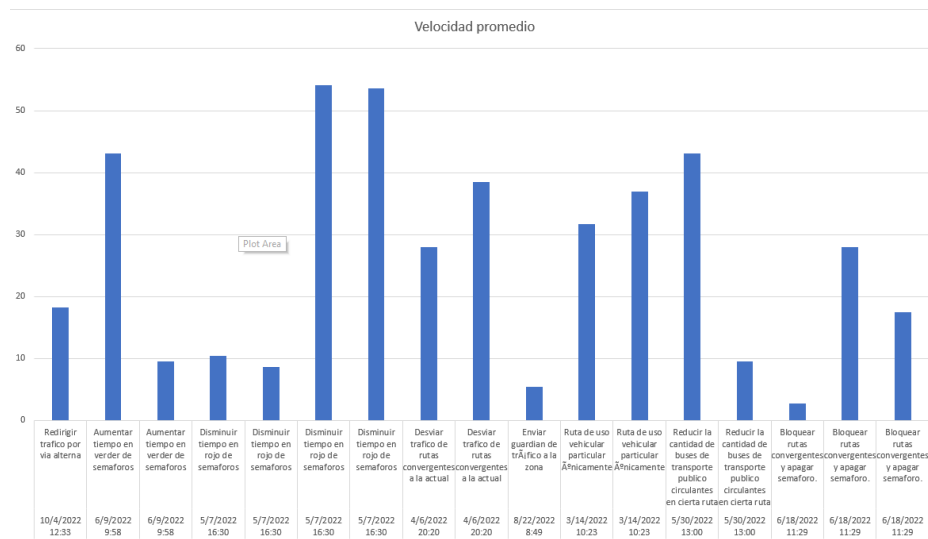


Figure 5: Grafico velocidad promedio

reportDatetime	solutionName	successRate	additionDate	averageSpeed	isMainStreet	
10/4/2022 12:33	Redirigir trafico por via alterna	39.36		11/14/2022	18.21	0
6/9/2022 9:58	Aumentar tiempo en verde de semaforos	85.21		11/22/2022	43.12	1
6/9/2022 9:58	Aumentar tiempo en verde de semaforos	85.21		11/22/2022	9.54	1
5/7/2022 16:30	Disminuir tiempo en rojo de semaforos	15.08		7/22/2022	10.4	0
5/7/2022 16:30	Disminuir tiempo en rojo de semaforos	15.08		7/22/2022	8.63	0
5/7/2022 16:30	Disminuir tiempo en rojo de semaforos	15.08		7/22/2022	54.16	0
5/7/2022 16:30	Disminuir tiempo en rojo de semaforos	15.08		7/22/2022	53.55	0
4/6/2022 20:20	Desviar trafico de rutas convergentes a la actual	80.65		11/1/2022	27.91	0
4/6/2022 20:20	Desviar trafico de rutas convergentes a la actual	80.65		11/1/2022	38.45	0
8/22/2022 8:49	Enviar guardian de tráfico a la zona	7.71		11/22/2022	5.38	1
3/14/2022 10:23	Ruta de uso vehicular particular Ánicamente	29.81		10/22/2022	31.63	1
3/14/2022 10:23	Ruta de uso vehicular particular Ánicamente	29.81		10/22/2022	36.88	1
5/30/2022 13:00	Reducir la cantidad de buses de transporte publico circulantes en cierta ruta	9.29		9/22/2022	43.12	1
5/30/2022 13:00	Reducir la cantidad de buses de transporte publico circulantes en cierta ruta	9.29		9/22/2022	9.54	1
6/18/2022 11:29	Bloquear rutas convergentes y apagar semaforo.	20.19		9/22/2022	2.77	0
6/18/2022 11:29	Bloquear rutas convergentes y apagar semaforo.	20.19		9/22/2022	27.92	0
6/18/2022 11:29	Bloquear rutas convergentes y apagar semaforo.	20.19		9/22/2022	17.5	0

Figure 6: Datos

Durante la realizacion del prototipo planteado del API desplegado en AWS con la ayuda de servicios como EC2 se llevo a cabo una simulacion donde pudimos extraer datos de las inserciones de la base de datos por medio de la herramienta de EXCEL para lograr visualizar la informacion en diferentes graficos para poder llegar a unas conclusiones mas acertadas y aterrizadas a la realidad basandonos en lo recopilado y de esta forma dar respuesta a lo que se tenia planteado en la propuesta de solucion.

9 Conclusiones

Luego de haber planteado la problematica que se va a tratar a cerca de como poder recopilar y analizar la informacion del trafico en tiempo real para poder mejorar la movilidad de las personas en la ciudad Y mejorar la productividad en la poblacion en sus labores diarias, nos disponemos a utilizar nuestro conocimiento aprendido en la asignatura de AREP y basado en proyectos hechos previamente de diferentes lugares del mundo para idear una solucion tecnologia basadandonos en una arquitectura utilizando servicios de la plataforma en la nube de AWS la cual nos ofrece multiples servicios que nos garanticen atributos de calidad muy valiosos como lo son disponibilidad con la ayuda de balanceadores de carga, portabilidad de la aplicacion con servicio de Docker, entre otros.

Teniendo la arquitectura lista nos disponimos a idear un prototipo para poder visualizar las ideas palanetadas de una mejor manera, para ellos creamos una appi con una base de datos en postgresql la cual recibe JSON con los datos de los usuarios al registrar la actividad del trafico en tiempo real, luego de esto montamos la aplicacion en una maquina virtual de EC2 en AWS y miramos que quede desplegada y operativa, esto es lo que se planeo para el prototipo ya que en terminos del machine learning o la inteligencia artificial

se sale un poco del alcance del curso por lo que nos limitamos a su implementación sin embargo es de gran importancia tenerla en el radar ya que en esta reside el valor agregado de nuestra propuesta al manejar las peticiones de las personas y en base a esta tomar buenas decisiones para poder tener una mejor movilidad en la ciudad.

Finalmente se llevo a cabo una insercion de datos de prueba en la base de datos creada y montada sobre la appi desplegada en EC2 para luego su extraccion mediante la herramienta de EXCEL para su analisis y evaluacion, gracias a esto se concluye que las dos formas de mejorar la movilidad en la ciudad son, primero disminuyendo los tiempos de los semaforos en rojo, ya que estos no estaban siendo equitativo con los semaforos en verde por lo que estos tienen mayor duracion que los verdes evitando un mejor flujo vehicular, y segundo evitando las bifurcaciones innecesarias cuando se esta en vias convergentes para evitar que si ya hay congestiones existentes la situacion se agrave aun mas.

References

- [1] Redes vehiculares aplicadas a la movilidad inteligente y sostenibilidad ambiental en entornos de ciudades inteligentes. Jose Antonio Sánchez. 2017. Tesis Doctoral. Universidad de Oviedo.
Disponible en: <https://dialnet.unirioja.es/servlet/tesis?codigo=207357>
- [2] Diseño de una infraestructura y servicio de monitorización de tráfico en Ciudades Inteligentes. Fernando Martín Gil. 2018.
Disponible en: <https://ruidera.uclm.es/xmlui/bitstream/handle/10578/16634/TFM%20Fernando%20Mart%C3%ADn.pdf?sequence=1>
- [3] Sistema inteligente para el monitoreo automatizado del transporte público en tiempo real. Revista Ibérica de Sistemas e Tecnologias de Informação. Quiñonez, Y., Lizarraga, C., Peraza, J., Zatarain, O. 2019.
Disponible en: <https://pdfs.semanticscholar.org/de9d/e1e897699ffc2d6aa257b7049ded61374719.pdf>
- [4] Simulación de tráfico en ciudades inteligentes con SUMO. Francisco Aparicio. 2018.
Disponible en: <https://uvadoc.uva.es/handle/10324/33032>
- [5] ETA Prediction with Graph Neural Networks in Google Maps. 2021.
Disponible en: <https://arxiv.org/pdf/2108.11482.pdf>