

PROYECTO APLICADO I

**PAULO ANDRES MARTINEZ MENDEZ
EDIER GUZMAN MORALES**

ANTEPROYECTO

**PREDICCION DEL TRAFICO DE DATOS DE LAS ZONAS WIFI PUBLICAS DE
SANTIAGO DE CALI**

**DIRECTORA
GLORIA INÉS ALVAREZ VARGAS**

**CODIRECTOR
DIEGO LUIS LINARES OSPINA**

INTRODUCCION

En la ciudad de Cali, donde actualmente se tienen aproximadamente de 2 millones de habitantes, se hace imprescindible para el gobierno local mantener a su población conectada en distintos sitios públicos, entre ellos parques o zonas recreativas. Es por eso que se han instalado en diferentes zonas de la ciudad zonas WiFi públicas que ayudan a tener conectada a la población.

Sin embargo, la alta demanda de este servicio trae consigo problemas de rendimiento en la red, incluyendo congestión en horas pico, desconexiones frecuentes o zonas que, aunque tienen infraestructura WiFi, no se utilizan. Estos problemas no solo afectan la satisfacción del ciudadano, sino que también limitan el potencial de aprovechar el WiFi como una herramienta que pueda conectar a los ciudadanos, brindando soluciones como acercarse al conocimiento, tener oportunidades de empleo, acceder a servicios públicos en la red y así cerrar la brecha digital.

Para abordar estos desafíos, el presente proyecto busca implementar un enfoque de ciencia de datos que permita analizar en profundidad el uso de las redes WiFi en diferentes zonas públicas de la ciudad de Cali. A través de la recopilación de datos de conexión, tráfico de datos y tiempos de uso, se pretende realizar un análisis que identifique horarios y zonas de alta demanda, áreas de congestión o baja conectividad. Estos hallazgos permitirán proponer ajustes estratégicos en la red, como la reubicación de puntos de acceso y la adaptación dinámica del ancho de banda según la demanda.

La solución se basa en el uso de modelos de aprendizaje supervisado para anticipar la carga de la red en diferentes momentos, lo que permitirá tomar decisiones informadas y ajustar los recursos de manera eficiente para identificar picos de demanda inesperados con el fin de responder rápidamente y minimizar los impactos en el servicio.

Con este proyecto, se espera mejorar significativamente la calidad del servicio WiFi, optimizando su distribución y funcionamiento en base a datos reales y patrones de uso. Los resultados esperados incluyen un aumento en la satisfacción del ciudadano, una reducción de los problemas de conectividad, y una herramienta de análisis que permitirá al distrito gestionar de forma proactiva la infraestructura WiFi. Además, el análisis de datos ofrecerá conocimientos valiosos que podrán ser aprovechados para iniciativas comerciales, eventos públicos y el crecimiento urbano entorno a dichas zonas.

1. DEFINICION DEL PROBLEMA

1.1. PLANTEAMIENTO DEL PROBLEMA

Actualmente las zonas WiFi públicas en Cali se ven sujetas a una serie de sobrecostos, ya que hay zonas wifi a la que se les ha asignado más recursos y que no son tan demandadas como otras.

Lo anterior es producido por la falta de un sistema que permita predecir el tráfico de datos en diferentes ubicaciones y horarios, esta ausencia limita la capacidad de anticipar la demanda, lo que resulta en una asignación ineficiente del ancho de banda, sobrecarga en puntos de acceso, y una experiencia deficiente para los ciudadanos.

1.2. FORMULACION DEL PROBLEMA

Pregunta de formulación:

¿Cómo predecir el tráfico de datos para las zonas WiFi públicas de Cali?

Preguntas de sistematización:

¿Cómo preparar los datos?

¿Cómo entrenar los modelos predictivos?

¿Cómo evaluar el desempeño de los modelos entrenados?

¿Cómo poner el modelo en funcionamiento?

2. OBJETIVOS DEL PROYECTO

2.1. OBJETIVO GENERAL

Desarrollar un sistema para predecir el tráfico de datos de las zonas WiFi públicas de Cali aplicando técnicas de aprendizaje automático supervisado.

2.2. OBJETIVOS ESPECIFICOS

1. Preparar los datos con herramientas de limpieza y transformación.
2. Entrenar los modelos predictivos de aprendizaje supervisado.
3. Evaluar el desempeño de los modelos entrenados.
4. Implementar un prototipo para el modelo en funcionamiento.

2.3. RESULTADOS ESPERADOS

- Data Set curado para ser utilizado en el desarrollo de modelos predictivos.
- Un modelo final seleccionado como el más adecuado según las métricas obtenidas y las varias versiones de modelos predictivos entrenados.
- Un prototipo básico en entorno local que muestre gráficos y predicciones en base al tiempo sobre el tráfico de datos en las zonas WiFi Publicas de Cali.
- Documento final detallado sobre el desarrollo, implementación y validación de los modelos predictivos que nos permiten la redistribución del ancho de banda y la predicción de tráfico del mismo.

3. ALCANCE

- El proyecto aplicado específicamente se basa en el análisis y modelado utilizando dataset que son generados por las Zonas WiFi públicas de Cali, no se incluyen datos de redes privadas.
- El proyecto se limita al análisis de tráfico de datos y predicciones de tráfico de las zonas wifi urbanas de Cali, excluyendo zonas rurales o corregimientos.
- No incluye mejoras directas al hardware o cambios físicos en los puntos de acceso WiFi.
- El prototipo de la interfaz será funcional, pero no se integrará con los sistemas existentes de la administración municipal de Cali y operará en entornos locales.
- Las recomendaciones de expansión o redistribución de infraestructura WiFi se plantearán únicamente como sugerencias futuras, y no como un objetivo directo del proyecto aplicado.

4. JUSTIFICACION

El acceso a internet se ha convertido en una necesidad esencial para la participación en actividades sociales, educativas y económicas en el mundo actual. En la ciudad de Cali, las zonas WiFi públicas desempeñan un papel fundamental en la reducción de la brecha digital, proporcionando conectividad a sectores de la población que carecen de acceso a servicios de internet en sus hogares. Sin embargo, la infraestructura WiFi pública enfrenta desafíos significativos, como la congestión en horas pico, desconexiones frecuentes y subutilización en ciertas áreas.

Estos problemas generan no solo insatisfacción en los ciudadanos, quienes dependen de estos servicios para actividades críticas, sino también sobrecostos para la administración municipal, que no cuenta con herramientas para optimizar el uso de la infraestructura actual. La falta de información confiable sobre el tráfico de datos en estas zonas dificulta la planificación, el mantenimiento proactivo y la expansión eficiente de la red.

La viabilidad del proyecto radica en la creciente disponibilidad de datos proporcionados por las entidades municipales sobre el uso de las zonas WiFi, lo que permite el desarrollo de un sistema predictivo basado en técnicas de aprendizaje automático.

El aporte específico de nuestra propuesta se centra en el diseño e implementación de un sistema que utiliza modelos predictivos para anticipar la demanda de tráfico en las zonas WiFi públicas. A diferencia de enfoques tradicionales que solo reaccionan a los problemas de forma parcial, nuestra propuesta introduce un modelo proactivo que facilita una mejor planificación y redistribución de recursos, reduciendo costos operativos y mejorando significativamente la experiencia del ciudadano.

Además, el prototipo funcional desarrollado servirá como una herramienta base que el gobierno local podrá adaptar y expandir en el futuro para integrar más zonas WiFi, apoyar la planificación urbana y fortalecer las estrategias de conectividad digital en la ciudad. Este proyecto también abre la posibilidad de aplicar enfoques similares en otros contextos, posicionando a Cali como un referente en el uso de la ciencia de datos para abordar problemas de la conectividad urbana.

5. MARCO DE REFERENCIA

5.1 Marco Teórico

El desarrollo de este proyecto requiere un fundamento teórico sólido que abarque los aspectos esenciales relacionados con el análisis y predicción del tráfico en las zonas WiFi públicas de Cali. En esta sección se abordarán tres temas principales: primero, se analizará el contexto y los retos actuales de las zonas WiFi públicas en Cali, destacando su importancia en la reducción de la brecha digital.

En segundo lugar, se explorarán los conceptos y técnicas del aprendizaje supervisado, centrándose en los algoritmos de predicción más adecuados para este proyecto, como la regresión lineal, los bosques aleatorios (Random Forest) y las redes neuronales. Finalmente, se discutirán las métricas de evaluación empleadas para medir el desempeño de los modelos desarrollados y se presentarán antecedentes relevantes que aportan conocimiento y precedentes valiosos para la implementación de esta propuesta.

Esta introducción permite estructurar la discusión y sentar las bases teóricas necesarias para garantizar la viabilidad y la efectividad del proyecto.

- **Zonas WiFi Públicas en Cali**

La ciudad de Cali cuenta con una red de zonas WiFi públicas distribuidas estratégicamente en áreas urbanas. Estas zonas tienen como objetivo reducir la brecha digital, brindando acceso gratuito a internet a los ciudadanos. Sin embargo, la creciente demanda ha generado problemas de congestión y deficiencias en la calidad del servicio. Según datos recientes (referencia al dataset disponible), estas zonas presentan picos de uso en horarios específicos, lo que destaca la necesidad de una gestión más eficiente.

- **Aprendizaje Supervisado en Predicción**

El aprendizaje supervisado es una técnica de Machine Learning que utiliza datos etiquetados para entrenar modelos capaces de realizar predicciones. Para este proyecto, se emplearán modelos como:

Regresión lineal: método estadístico y de aprendizaje supervisado que modela la relación entre una variable dependiente "Y" (lo que queremos predecir) y una o más variables independientes "X" (las características o inputs).

Ecuación General:

La relación se expresa como una línea recta:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

- Y : Variable dependiente (resultado o predicción).
- X_1, X_2, \dots, X_n : Variables independientes (predictores o inputs).
- β_0 : Intercepto (valor de Y cuando $X = 0$).
- $\beta_1, \beta_2, \dots, \beta_n$: Coeficientes que indican la influencia de cada X en Y .
- ϵ : Error residual (diferencia entre el valor observado y el predicho).

Random Forest: algoritmo de aprendizaje supervisado que combina múltiples árboles de decisión para hacer predicciones más precisas y reducir el riesgo de overfitting. Se utiliza tanto para clasificación como para regresión.

Ecuación conceptual:

Para Regresión: La predicción final (\hat{y}) es el promedio de las predicciones individuales (\hat{y}_i) de los árboles (T):

$$\hat{y} = \frac{1}{T} \sum_{i=1}^T \hat{y}_i$$

- T : Número total de árboles en el bosque.
- \hat{y}_i : Predicción del árbol i .

Redes neuronales: modelos de aprendizaje supervisado inspirados en la estructura y funcionamiento del cerebro humano.

Estructura básica:

$$z = \sum_{i=1}^n w_i x_i + b$$

$$a = \sigma(z)$$

- x_i : Entrada a la neurona i .
- w_i : Peso asociado a la entrada x_i .
- b : Sesgo (bias), un término adicional para ajustar la salida.
- σ : Función de activación (ReLU, Sigmoide, etc.).
- a : Salida de la neurona.

Estas técnicas son ampliamente utilizadas en problemas similares debido a su capacidad para identificar patrones y realizar predicciones precisas basadas en datos históricos.

1.3 Métricas de Evaluación de Modelos

Para evaluar el desempeño de los modelos predictivos, se emplearán las siguientes métricas:

Error Cuadrático Medio (MSE) El MSE mide la media de los errores al cuadrado entre las predicciones del modelo y los valores reales. Penaliza más los errores grandes debido al término cuadrático, lo que lo hace útil para identificar grandes desviaciones.

Ecuación:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- y_i : Valor real de la variable dependiente.
- \hat{y}_i : Predicción del modelo.
- n : Número total de observaciones.

Error Absoluto Medio (MAE): El MAE mide la media de los valores absolutos de las diferencias entre las predicciones del modelo y los valores reales. Es una métrica menos sensible a los errores extremos que el MSE, lo que la hace adecuada para medir la precisión general.

Ecuación:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Coeficiente de Determinación (R^2): Evalúa la proporción de la varianza total de la variable dependiente que es explicada por el modelo. Un valor de R^2 cercano a 1 indica un buen ajuste del modelo, mientras que valores cercanos a 0 indican que el modelo no explica bien la varianza de los datos.

Ecuación:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

- \bar{y} : Media de los valores reales.

5.2 Antecedentes

Los antecedentes proporcionan una base de trabajos previos relacionados con la temática del proyecto.

Wi-Fi Data Analysis Based on Machine Learning: (Análisis de datos WiFi basado en Machine Learning)

- **¿Qué hace el proyecto?**

Este proyecto basado también en las redes WiFi usan Machine Learning para detectar brechas de seguridad de los distintos dispositivos y redes conformadas por cada uno de ellos.

- **¿Cómo lo hicieron?**

Recolección de Datos: Recopilaron datos de señales Wi-Fi, como la intensidad de la señal (RSSI)

Preprocesamiento de Datos: Limpieza de datos para eliminar valores erróneos o inconsistentes, Normalización de los datos para garantizar uniformidad en las características.

Selección de Modelos: Se utilizaron algoritmos de aprendizaje supervisado, como:

- **Random Forest.**
- **K-Nearest Neighbors (KNN).**
- **Support Vector Machines (SVM).**

- **Resultados Obtenidos:**

El mejor modelo fue Random Forest, ya que proporcionó el equilibrio óptimo entre precisión y eficiencia computacional. Los resultados demostraron la viabilidad del uso de señales WiFi para clasificaciones basadas en aprendizaje automático en interiores.

Link del proyecto: <https://aircconline.com/ijcsit/V15N4/15423ijcsit04.pdf>

Internet Usage Modeling of Large Wireless Networks Using Self-Organizing Maps:

(Modelado del uso de internet de redes WiFi grandes usando mapas auto-organizados)

- **¿Qué hace el proyecto?**

Mediante un concepto llamado mapas auto-organizados los autores de este Proyecto lo usan para descubrir, organizar y visualizar las tendencias de diferentes usuarios en redes inalámbricas extensas, descubriendo patrones de tráfico y detección de anomalías.

- **¿Como lo hicieron?**

El estudio se centra en el análisis de patrones de movilidad y la caracterización de "escenarios sociales" (social scenes) utilizando redes móviles y datos de dispositivos personales. Los pasos principales fueron:

Recolección de Datos:

Datos recopilados de dispositivos personales como teléfonos móviles, que incluyeron:
Conexiones Wi-Fi.
Señales Bluetooth.
Ubicaciones GPS.

Los datos se obtuvieron en diversos entornos sociales, como oficinas, hogares, y espacios públicos.

Preprocesamiento de Datos:

Limpieza de datos para eliminar información incompleta o errónea.
Agrupación de los datos en segmentos temporales para analizar patrones específicos.

Definición de "Escenarios Sociales":

Los escenarios sociales fueron definidos mediante métricas como:

Interacciones entre dispositivos (proximidad Bluetooth).
Patrones de movimiento (trayectorias GPS).
Duración de la conexión a puntos de acceso Wi-Fi.

Clustering y Análisis:

Se utilizaron algoritmos de agrupamiento (clustering) como el Self-Organizing Map (SOM) para identificar patrones comunes en los datos de movilidad.

SOM permitió visualizar las relaciones entre escenarios sociales y categorizar los comportamientos en diferentes contextos.

- **Resultados Obtenidos:**

Identificación de Escenarios Sociales:

Los datos permitieron caracterizar varios escenarios, como:

Espacios de trabajo: Altas interacciones Bluetooth con dispositivos cercanos y conexiones Wi-Fi prolongadas.

Espacios públicos: Movimientos más dispersos y conexiones intermitentes a Wi-Fi.

Hogares: Baja movilidad y patrones de conexión Wi-Fi consistentes.

Patrones de Movilidad:

Los patrones de movimiento mostraron diferencias significativas entre los escenarios:

Mayor estabilidad en hogares y oficinas.

Alta dinámica en espacios públicos.

Eficiencia de SOM:

El algoritmo SOM fue efectivo para agrupar y visualizar los patrones de movilidad.

Los resultados tienen aplicaciones en el diseño de redes móviles, optimización de recursos, y comprensión de comportamientos sociales en entornos urbanos.

Link del Proyecto:

<https://www.cise.ufl.edu/~helmy/papers/scenes-cameraready.pdf>

Kaseya Center installing Enterprise Class guest WiFi to capture data and improve the fan experience:

(Kaseya Center, instalando WiFi público de clase Empresarial para capturar datos y mejorar la experiencia de los aficionados de Miami Heat)

- **¿Qué hace el proyecto?**

El caso del Kaseya Center destaca cómo la implementación de WiFi público empresarial permitió capturar datos de uso para mejorar la experiencia de los usuarios durante eventos deportivos. Este enfoque demuestra el valor de analizar patrones de tráfico y comportamiento para optimizar servicios en tiempo real.

- **¿Como lo hicieron?**

Capturar datos demográficos de los asistentes, que en su mayoría eran desconocidos debido a la compra grupal de boletos o a través de terceros, Se utilizaron páginas de inicio personalizadas (splash pages) para permitir que los usuarios se registraran fácilmente, utilizando Facebook o un formulario breve.

- **Resultados Obtenidos:**

- 1. **Uso del WiFi:**

- Más de **290,000 usuarios únicos** conectados a la red WiFi en los primeros 12 meses.
 - En promedio, **25% de los asistentes** a eventos se conectaron al WiFi durante cada concierto o partido.

- 2. **Recopilación de Datos:**

- Se identificaron **700,000 dispositivos únicos**, lo que permitió obtener información demográfica valiosa.
 - El análisis mostró que **90% de los asistentes** eran visitantes únicos, lo que permitió personalizar campañas para maximizar sus gastos.

- 3. **Retorno de Inversión (ROI):**

- Aumento en las interacciones personalizadas con los asistentes.
 - Incremento en las ventas de alimentos, bebidas y productos de la tienda oficial del Miami HEAT a través de ofertas y descuentos personalizados.

Link del proyecto:

<https://purple.ai/case-studies/miami-heat/>

6. METODOLOGIA

6.1 Preparación de los Datos

6.1.1 Recolección de datos: Obtener datasets de entidades como la administración municipal, en donde se encuentren las distintas zonas WiFi y su consumo de tráfico en el tiempo.

6.1.2 Limpieza y transformación de datos: Identificar y corregir errores, datos faltantes o inconsistencias en el dataset mediante herramientas como Pandas para Python o lenguaje R.

Se limpian los datos eliminando o imputando valores erróneos o nulos. Se transforman y normalizan los datos para garantizar uniformidad y escalabilidad.

6.1.3 Integración de datasets: Combinar información de diferentes fuentes en un único conjunto de datos listo para análisis.

Se Integran diversas fuentes de datos en un único dataset estructurado para el análisis.

6.2 Desarrollo de Modelos Predictivos

6.2.1 Selección de variables: Identificar las variables dependientes e independientes (como hora del día, ubicación, tráfico de datos, etc.) más relevantes.

6.2.2 Construcción de modelos: Entrenar varios modelos de aprendizaje supervisado, como regresión lineal, Random Forest y redes neuronales.

6.2.3 Análisis de errores: Identificar patrones en los errores de predicción para realizar ajustes.

6.2.4 Comparación de métricas: Evaluar los modelos según su precisión, eficiencia y robustez. Registrar el desempeño de los modelos para justificar la selección final.

Para todo lo anterior se realizan análisis exploratorio de datos (EDA) para comprender patrones y relaciones entre variables. Se seleccionan los modelos predictivos iniciales (ej. regresión lineal, Random Forest, redes neuronales). Se dividen los datos en conjuntos de entrenamiento, validación y prueba. Se entrena los modelos iniciales utilizando técnicas básicas de ajuste.

6.3. Evaluación de los modelos entrenados.

6.3.1 Ejecución de índices de confiabilidad: El calcular los errores, varianzas y demás índices permitirán evaluar los modelos.

6.3.2 Escogencia del modelo con los mejores índices: Se elige entonces el mejor modelo dependiendo de los índices ejecutados aplicando las métricas de evaluación como el Error Cuadrático Medio (MSE), Error Absoluto Medio (MAE) y el coeficiente de determinación (R^2). Se identifican patrones de error y se ajustan los hiperparámetros de los modelos.

6.4 Implementación de una Interfaz Funcional

6.4.1 Diseño de la interfaz: Crear una interfaz básica que permita ingresar parámetros (fecha, ubicación) y visualizar predicciones. Implementar la interfaz utilizando herramientas como Python o aplicaciones web básicas para integrar el modelo.

Conectar el modelo predictivo seleccionado con la interfaz para generar predicciones. Visualizar los resultados en gráficos que muestren el tráfico estimado por zonas y tiempo. Realizar pruebas del prototipo para garantizar su funcionalidad en entornos locales.

7. RECURSOS A EMPLEAR

7.1. Recursos Humanos

- **Equipo de trabajo:**
 - Estudiantes responsables del proyecto: Paulo Andrés Martínez Méndez y Edier Guzmán Morales.
 - Director del proyecto: Gloria Inés Álvarez Vargas.
 - Codirector del proyecto: Diego Luis Linares Ospina.
- **Colaboradores externos:** Personal de las entidades municipales (como DATIC Cali) para facilitar el acceso a los datos requeridos.

8. CRONOGRAMA

Tareas	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6	Mes 7	Mes 8	Mes 9	Mes 10
Preparar los datos										
Recolección de datos										
Limpieza y transformación de datos										
Integración de datasets										
Desarrollar modelos Predictivos.										
Selección de variables										
Construcción de modelos										
Análisis de errores										
Comparación de métricas										
Evaluar el desempeño de los modelos										
Ejecución de índices de confiabilidad										
Escogencia del modelo con los mejores índices										
Implementar un prototipo funcional										
Diseño de la interfaz										

9. REFERENCIAS BIBLIOGRAFICAS

- [1] Michelle Pierce (2020, Septiembre 3). WiFi Heat Map: A Complete Guide for Better Coverage [Online]. Available: <https://techgrid.com/blog/wifi-heatmaps-guide>
- [2] Cisco Wireless. (2022, Noviembre 11). Cisco Embedded Wireless Controller on Catalyst Access Points (EWC) White Paper [Online]. Available: <https://www.cisco.com/c/en/us/products/collateral/wireless/embedded-wireless-controller-catalyst-access-points/white-paper-c11-743398.html>
- [3] Lexy Savvides. (2015, Junio 3). Staying safe on public Wi-Fi [Online]. Available: <https://www.cnet.com/tech/services-and-software/tips-to-stay-safe-on-public-wi-fi/>
- [4] Sala de Comunicaciones Telefónica. (2024, Junio 10). Aprendizaje supervisado: definición y aplicaciones [Online]. Available: <https://www.telefonica.com/es/sala-comunicacion/blog/aprendizaje-supervisado-definicion-aplicaciones/>
- [5] Economipedia. (2024, Julio 30) Coeficiente de determinación (R cuadrado): qué es y cómo funciona [Online]. Available: <https://economipedia.com/definiciones/r-cuadrado-coeficiente-determinacion.html>
- [6] Andrés Torres (2021, Octubre 14). Aprendizaje automático: Una introducción al error cuadrático medio y las líneas de regresión. [Online]. Available: <https://www.freecodecamp.org/espanol/news/aprendizaje-automatico-una-introduccion-al-error-cuadratico-medio-y-las-lineas-de-regresion/>

Santiago de Cali, 25 de Noviembre de 2024

**Ingeniero
Diego Luis Linares
Director Maestría en Ciencia de Datos
Facultad de Ingeniería y Ciencias
Pontificia Universidad Javeriana - Cali**

Con el fin de cumplir con los requisitos exigidos por la Universidad para llevar a cabo el Proyecto de Grado y posteriormente optar por el título de Magíster en Ciencia de Datos, me permito presentar como Director a: Gloria Inés Alvarez Vargas identificada con C.C. _____ y como codirector a: Diego Luis Linares Ospina identificado con C.C. _____ del Proyecto de Grado denominado PREDICCION DEL TRAFICO DE DATOS DE LAS ZONAS WIFI PUBLICAS DE SANTIAGO DE CALI, el cual será realizado por los estudiantes, Paulo Andrés Martinez Mendez con código 0183321, Edier Guzmán Morales con código 9015769.

Atentamente,

Paulo Andrés Martinez Mendez
C.C. 94424446 de Cali

Edier Guzmán Morales
C.C. 1130618441 de Cali

Gloria Inés Álvarez Vargas
C.C.

Diego Luis Linares Ospina
C.C.

Santiago de Cali, 25 de noviembre de 2025

**Ingeniero
Diego Luis Linares
Director Maestría en Ciencia de Datos
Facultad de Ingeniería y Ciencias
Pontificia Universidad Javeriana - Cali**

Con el fin de cumplir con los requisitos exigidos por la Universidad para llevar a cabo el Proyecto Aplicado y posteriormente optar por el título de Magíster en Ciencia de Datos, nos permitimos presentar a su consideración el anteproyecto de Trabajo de Grado denominado “Predicción del tráfico de datos de las zonas WiFi públicas de Santiago de Cali”, el cual será realizado por los estudiantes Paulo Andrés Martínez Méndez con código 0183321, Edier Guzmán Morales con código 9015769 pertenecientes a la Maestría en Ciencia de Datos, bajo la dirección de la profesora Gloria Inés Álvarez Vargas y la codirección del profesor Diego Luis Linares Ospina.

El suscrito director del Proyecto Aplicado autoriza para que se proceda a hacer la evaluación de este Anteproyecto ante el Tribunal que para el efecto se designe, toda vez que ha revisado cuidadosamente el documento y avala que ya se encuentra listo para ser presentado oficialmente.

Atentamente,



Paulo Andrés Martínez Méndez
C.C. 94424446 de Cali



Edier Guzmán Morales
C.C. 1130618441 de Cali

Gloria Inés Álvarez Vargas
C.C.

Diego Luis Linares Ospina
C.C.