



**Modelo de Machine Learning para la normalización y
geocodificación de direcciones en Imbocar S.A.S.**

Laura Herrera Giraldo

Ingeniera Industrial

Modalidad de Práctica

Semestre de Industria o Práctica Empresarial

Asesor

Juan Sebastián Jaén Posada, Doctor (PhD) en Ingeniería

Universidad de Antioquia

Facultad de Ingeniería

Ingeniería Industrial

Medellín, Antioquia, Colombia

2024

Cita	(Herrera Giraldo, 2024)
Referencia	Herrera Giraldo, L. (2024). Modelo de Machine Learning para la Normalización y Geocodificación de Direcciones en Imbocar S.A.S. [Trabajo de grado profesional]. Universidad de Antioquia, Medellín, Colombia.
Estilo APA 7 (2020)	



Centro de Documentación Ingeniería (CENDOI)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Dedicatoria

A quien ha sido mi luz en momentos de oscuridad. Gracias por darme ánimo cuando sentía que no podía más, por aconsejarme con paciencia y recordarme que soy más fuerte de lo que creo. Por celebrar mis victorias como si fueran tuyas y alegrarte sinceramente con cada paso que doy. Tu apoyo incondicional y tu fe en mí han sido mi mayor impulso para no rendirme.

Agradecimientos

A mis padres, por creer en mí y brindarme las herramientas para alcanzar cada meta propuesta. Gracias a su amor, esfuerzo y apoyo incondicional, he llegado a ser la persona que soy.

A mis amigos, por estar siempre en los buenos y malos momentos. Gracias por las experiencias maravillosas que compartimos en la universidad, por cada risa, cada consejo y cada instante que hicieron de esta etapa algo inolvidable.

A mis profesores, por su dedicación, paciencia y pasión por enseñar. A mi Alma Máter, por ser mi segunda casa y brindarme las herramientas para crecer como persona y profesional.

A la empresa Imbocar S.A.S., por abrirme las puertas y permitirme desarrollar este proyecto. Gracias por el apoyo brindado durante este proceso y por ofrecerme un espacio valioso para aprender.

Tabla de contenido

Resumen	8
Abstract	9
1. Introducción	10
2. Objetivos	12
2.1 Objetivo general	12
2.2 Objetivos específicos.....	12
3. Marco teórico	13
3.1 Automatización de Procesos	13
3.2 Calidad de Datos	13
3.3 Normalización o estandarización de direcciones	13
3.4 Geocodificación.....	14
3.5 Procesamiento del lenguaje natural (NLP).....	14
4. Metodología	15
4.1 Entendimiento del Problema	15
4.2 Selección de Herramientas Adecuadas.....	15
4.3 Desarrollo del Código	15
4.3.1 Normalización de Direcciones.....	15
4.3.2 Geocodificación	16
4.4 Pruebas de Funcionamiento.....	17
4.5 Comparación del Proceso Actual y el Propuesto	19
5. Análisis de resultados.....	20
5.1 Comparación de Alternativas	20

5.1.1 En términos de tiempo	20
5.1.2 En términos de costo.....	20
5.1.3 Ventajas y desventajas	21
6. Conclusiones y recomendaciones.....	22
Referencias	23

Lista de tablas

Tabla 1. Resultados direcciones estandarizadas.....	17
Tabla 2. Resultados direcciones geocodificadas	18
Tabla 3. Comparación alternativas en tiempo	20
Tabla 4. Comparación alternativas en costo.....	20

Siglas, acrónimos y abreviaturas

NLP	Natural Language Processing
NER	Named Entity Recognition
API	Application Programming Interface

Resumen

Este proyecto se enfocó en optimizar los procesos de normalización y geocodificación de direcciones en la empresa Imbocar S.A.S., desarrollando una solución automatizada que reduce la carga de trabajo manual y mejora la eficiencia operativa. Mediante el uso de técnicas de Procesamiento de Lenguaje Natural (NLP), incluyendo el Reconocimiento de Entidades Nombradas (NER) para identificar y estructurar los componentes clave de las direcciones, y la integración con la API de Google Maps, se logró diseñar un modelo que disminuye el tiempo de normalización y reduce los costos por dirección en un 209%. La solución también eliminó la dependencia de proveedores externos, otorgando mayor control interno sobre los procesos logísticos y aumentando la confiabilidad de los datos procesados. Este trabajo demuestra el impacto positivo del desarrollo de soluciones automatizadas, reduciendo costos, optimizando tiempos y fortaleciendo la capacidad operativa de Imbocar, mientras sienta las bases para posibles mejoras a futuro.

Palabras clave: Normalización de direcciones, automatización de procesos, Reconocimiento de Entidades Nombradas (NER), geocodificación.

Abstract

This project aimed to optimize the address normalization and geocoding processes at Imbocar S.A.S., developing an automated solution that reduces the manual workload and improves operational efficiency. A model was designed using natural language processing (NLP) techniques, including named entity recognition (NER) to identify and structure the key components of addresses, and integration with the Google Maps API, which reduces normalization time and reduces costs per address by 209%. The solution also avoided dependence on external suppliers, giving greater internal control over logistical processes and increasing the reliability of the data processed. This work demonstrates the positive impact of developing automated solutions, reducing costs, optimizing time and increasing Imbocar's operational capacity, while laying the groundwork for possible future improvements.

Keywords: Address normalization, process automation, Named Entity Recognition (NER), geocoding.

1. Introducción

En el sector logístico y de transporte, la entrega de pedidos de manera eficiente es un factor crítico para garantizar la satisfacción del cliente y mantener la competitividad de las empresas (Mora García, 2023). Imbocar S.A.S., una compañía colombiana especializada en servicios de última milla, se enfrenta a un desafío significativo relacionado con la calidad y estandarización de los datos de direcciones proporcionados por sus clientes.

En Imbocar, los clientes envían bases de datos con direcciones que frecuentemente presentan problemas como la falta de estandarización y la mezcla de información adicional (nombres de edificios, apartamentos o conjuntos residenciales) en el mismo campo de la dirección. Esta situación genera dificultades para realizar una geolocalización precisa, lo que complica la planificación de rutas y aumenta el tiempo necesario para localizar las direcciones. Actualmente, el proceso de estandarización, corrección y verificación de direcciones es manual, lo que consume una gran cantidad de tiempo y recursos humanos, ya que los empleados deben invertir horas, e incluso días, para organizar la información, lo que afecta a productividad del equipo administrativo, aumenta el riesgo de errores y limita el tiempo de los empleados para enfocarse en actividades estratégicas de mayor valor para la empresa.

Para abordar esta problemática es importante revisar trabajos previos que hayan enfrentado y solucionado desafíos similares. Un antecedente relevante en este ámbito es el trabajo de (Vélez, 2022), quien desarrolló una solución automatizada para el sistema de última milla en el Grupo Familia. Este proyecto implementó herramientas como expresiones regulares y una macro en Excel para estructurar y separar direcciones, logrando una reducción del 82,59% en los tiempos de procesamiento y una mejora significativa en la eficiencia operativa. Este estudio demuestra el impacto positivo que tienen las soluciones automatizadas en la gestión de datos logísticos y resalta la importancia de contar con datos estandarizados para para optimizar procesos logísticos.

Considerando la efectividad demostrada en el trabajo de Maria Paulina Vélez, este proyecto busca abordar esta problemática mediante el desarrollo de una solución automatizada que permita normalizar y geocodificar las direcciones de forma eficiente. Al identificar y separar los

componentes principales de una dirección (tipo de vía, número, letra, etc.) y sus complementos adicionales (edificio, apartamento, barrio), se optimizará el proceso de corrección y estandarización de las direcciones y se reducirá la carga de trabajo manual del equipo administrativo. Además, esta solución eliminará la dependencia de proveedores externos para la geocodificación, generando ahorros económicos significativos y proporcionando una base sólida para integrar la información en otros sistemas de la empresa, como herramientas de geolocalización y planificación de rutas. Esto mejorará la calidad de los datos, reducirá los errores operativos y fortalecerá la capacidad de Imbocar para ofrecer un servicio más eficiente, confiable y competitivo.

La metodología empleada en este proyecto incluyó el desarrollo de un modelo automatizado para la normalización y geocodificación de direcciones, basado en técnicas de procesamiento de lenguaje natural y el uso de una API de Google Maps. Los resultados demostraron una mejora significativa en términos de tiempo y costos, reduciendo el tiempo de normalización de 8-12 horas a 19.88 segundos y logrando un ahorro del 209% en costos por dirección.

2. Objetivos

2.1 Objetivo general

Desarrollar una solución automatizada para estandarizar y geocodificar las direcciones proporcionadas por los clientes de la empresa Imbocar S.A.S., con el fin de hacer más eficiente el proceso de normalización y geocodificación, reduciendo la carga de trabajo manual.

2.2 Objetivos específicos

- Seleccionar una herramienta de procesamiento de lenguaje natural (NLP) adecuada para implementar la normalización y separación automatizada de direcciones.
- Establecer las reglas y diccionarios para la separación y estandarización de las partes de una dirección, garantizando un formato homogéneo.
- Desarrollar los códigos para la normalización y geocodificación de direcciones, utilizando técnicas de procesamiento de lenguaje natural (NLP) e integrando una API de Google Maps para obtener las direcciones estandarizadas y sus coordenadas geográficas.
- Realizar pruebas con diferentes conjuntos de datos para evaluar la precisión y el funcionamiento de la solución propuesta en la normalización y geocodificación de direcciones.
- Comparar el proceso actual con el automatizado, midiendo mejoras en términos de tiempo, costos y eficiencia operativa.

3. Marco teórico

3.1 Automatización de Procesos

La automatización de procesos consiste en el uso de tecnologías avanzadas para ejecutar tareas repetitivas y estructuradas, optimizando el tiempo y los recursos disponibles. Según Ghosh, esta práctica ha evolucionado desde su aplicación inicial en la manufactura hacia usos en el sector servicios y procesos administrativos. Entre sus principales beneficios se encuentran una mayor precisión, reducción de costos y la liberación del personal para enfocarse en actividades estratégicas. Además, la automatización inteligente, que integra herramientas basadas en datos y análisis predictivos, no solo mejora la eficiencia operativa, sino que también genera valor estratégico y aumenta la adaptabilidad empresarial (Ghosh et al., 2021).

3.2 Calidad de Datos

La calidad de los datos se refiere a qué tan aptos son estos para ser utilizados en procesos empresariales, siendo un factor esencial para garantizar la precisión y confiabilidad de los sistemas operativos y analíticos. Según el documento “A Review on Data Quality Dimensions for Big Data”, las dimensiones clave de la calidad de datos incluyen la precisión, consistencia, completitud y puntualidad, que son fundamentales para la toma de decisiones informadas. La falta de calidad en los datos puede conducir a decisiones equivocadas, afectando negativamente los resultados organizacionales. Una gestión efectiva de la calidad de los datos permite no solo minimizar errores, sino también garantizar que los modelos predictivos y las herramientas analíticas basadas en estos datos sean precisos y útiles (Ridzuan et al., 2024).

3.3 Normalización o estandarización de direcciones

La estandarización de direcciones es el proceso mediante el cual se transforman direcciones incorrectas, no estándar o incompletas en un formato uniforme que cumpla con los estándares globales. Este proceso incluye la corrección de errores tipográficos, ortográficos o de formato, así como la incorporación de los elementos esenciales de una dirección, como número de calle, nombre

de calle, ciudad, estado y código postal. Aunque la estandarización no garantiza la validez de la dirección, asegura que su formato sea correcto, facilitando su interpretación y uso en aplicaciones como geocodificación, logística y envíos postales. La falta de estandarización puede dificultar la localización de destinos debido a información incompleta, no válida o imprecisa, lo que resalta la importancia de este proceso en entornos donde la precisión y consistencia de los datos son esenciales (ArcGIS, 2022).

3.4 Geocodificación

La geocodificación es el proceso de transformar descripciones de ubicaciones, como direcciones, coordenadas geográficas o nombres de lugares, en ubicaciones precisas sobre la superficie de la Tierra. Este proceso permite convertir información textual o tabular en entidades geográficas que pueden ser representadas en mapas y utilizadas para análisis espacial. La geocodificación tiene aplicaciones en diversos campos, como la administración de datos de clientes, el análisis de patrones espaciales y la optimización de rutas (ArcGIS, 2021).

3.5 Procesamiento del lenguaje natural (NLP)

El NLP es un área de la inteligencia artificial y la lingüística computacional que tiene como objetivo principal permitir que las máquinas comprendan, interpreten y generen lenguaje humano de manera eficaz. Se basa en la creación de modelos computacionales capaces de analizar y procesar grandes volúmenes de datos textuales, facilitando tareas como la extracción de información, el análisis de sentimientos, la traducción automática y la búsqueda de datos relevantes (Moreira et al., 2021).

4. Metodología

La metodología adoptada en este proyecto siguió un enfoque estructurado, dividido en etapas que abordan desde el entendimiento inicial del problema hasta la evaluación de la solución propuesta. A continuación, se describen cada una de las fases:

4.1 Entendimiento del Problema

El análisis inicial se centró en la falta de estandarización de las direcciones y en el hecho de que esta tarea se realiza manualmente en la empresa, lo que consume una cantidad significativa de tiempo del equipo administrativo. Los empleados deben dedicar horas, e incluso días, a organizar y estandarizar la información proporcionada por los clientes. Este proceso manual genera retrasos que impactan negativamente la eficiencia operativa y aumentan la carga de trabajo, limitando la capacidad de los empleados para enfocarse en actividades estratégicas. La información sobre esta problemática fue recolectada mediante conversaciones con el equipo encargado de esta función, quienes detallaron las dificultades enfrentadas, los recursos empleados y el impacto de las inconsistencias en las bases de datos de direcciones.

4.2 Selección de Herramientas Adecuadas

Se evaluaron diferentes herramientas tecnológicas para abordar los problemas detectados. Para la normalización de direcciones, se seleccionó un modelo de NLP basado en la biblioteca SpaCy, dada su capacidad para entrenar modelos NER personalizados. Para la geocodificación, se optó por la integración con la API de Google Maps, que proporciona resultados precisos, incluyendo coordenadas geográficas y direcciones verificadas.

4.3 Desarrollo del Código

El desarrollo se dividió en dos componentes principales:

4.3.1 Normalización de Direcciones

- Se implementó un modelo NER personalizado utilizando la librería spaCy, que permitió identificar y separar los componentes principales de una dirección (tipo de vía, número,

letra, complementos adicionales, etc.). Este modelo fue entrenado utilizando un conjunto de datos etiquetados específicamente para este proyecto, el cual incluía ejemplos representativos de las diferentes estructuras y patrones encontrados en las direcciones. El entrenamiento incluyó varias iteraciones para optimizar la precisión del modelo en la clasificación de los componentes.

- Preprocesamiento de datos: Todas las direcciones fueron convertidas a minúsculas para evitar inconsistencias debidas a diferencias en mayúsculas y minúsculas. Se eliminaron caracteres especiales, como signos de puntuación y símbolos innecesarios, para garantizar que el texto fuera uniforme y libre de elementos que pudieran interferir en el análisis automático, además, se eliminaron espacios adicionales. Estas reglas de limpieza fueron definidas basándose en el conjunto de datos proporcionado y ajustadas tras realizar pruebas iniciales para maximizar la calidad del procesamiento.
- Se crearon diccionarios específicos para estandarizar las direcciones y sus indicaciones adicionales. Estos diccionarios incluyeron una variedad de términos comúnmente utilizados y sus diversas formas de escritura, tales como "Calle", "Carrera", "Avenida", "Transversal", "Diagonal", y sus abreviaciones ("Cl", "Cra", "Av"). A cada variación se le asignó un término estandarizado, asegurando así la uniformidad en el formato de salida. Además, se incorporaron reglas específicas para corregir errores tipográficos comunes y normalizar indicaciones adicionales como "Apartamento", "Torre", "Bloque", entre otras. Esto permitió identificar y facilitar la separación de los elementos complementarios de la dirección principal, facilitando un formato homogéneo para el procesamiento posterior.
- Se desarrolló una interfaz gráfica que facilita la carga de archivos con direcciones y devuelve los resultados en un archivo de excel.
- Finalmente, el código fue compilado en un ejecutable (.exe) para su uso autónomo por parte del equipo administrativo.

4.3.2 Geocodificación

- Se configuró la conexión con la API de Google Maps, definiendo los parámetros necesarios para procesar las direcciones normalizadas y obtener como resultados: latitud, longitud y la dirección verificada por Google Maps.

- Se desarrolló una interfaz gráfica para cargar direcciones normalizadas, procesar las solicitudes y exportar los resultados en un archivo Excel.
- El código fue empaquetado como un ejecutable independiente para facilitar su uso.

4.4 Pruebas de Funcionamiento

Se realizaron pruebas para garantizar la precisión y funcionalidad de ambos componentes:

- Se utilizaron conjuntos de datos con direcciones reales para evaluar el funcionamiento de los procesos de normalización y geocodificación.

La Tabla 1 muestra ejemplos de direcciones antes y después de su estandarización, en la columna “Dirección Original” se presenta la entrada inicial, mientras que en “Dirección Normalizada” se observa el resultado después de la estandarización. Además, se incluyen tres columnas adicionales: una con la dirección separada de su complemento, en caso de que contara con información complementaria, otra con el complemento y la última con la dirección concatenada a la columna “CIUDAD” (si esta estaba presente en el archivo original). Esta última tiene como propósito mejorar la precisión en la búsqueda de la dirección.

Tabla 1. Resultados direcciones estandarizadas

Dirección Original	Dirección Normalizada	Dirección Separada	Complemento	Dirección_ciudad
Calle 23 #45 - 08	cl 23 45 08	cl 23 45 08		cl 23 45 08 GALAPA
CLL 55a #179-33	cl 55 a 179 33	cl 55 a 179 33		cl 55 a 179 33 SOLEDAD
KR 50b#40-89	cra 50 b 40 89	cra 50 b 40 89		cra 50 b 40 89 BARRANQUILLA
CALL 70C#23C-57 SAN FELIPE	cl 70 c 23 c 57 san felipe	cl 70 c 23 c 57	san felipe	cl 70 c 23 c 57 BARRANQUILLA
Call 7 #56-120	cl 7 56 120	cl 7 56 120		cl 7 56 120 GALAPA
KRA 4 sur # 50b- 45	cra 4 sur 50 b 45	cra 4 sur 50 b 45		cra 4 sur 50 b 45 BARRANQUILLA

CARRERRA 17 ## 10 - 143	cra 17 10 143	cra 17 10 143		cra 17 10 143 PUERTO COLOMBIA
TV 44 #75b-29 loc 3	transversal 44 75 b 29 loc 3	transversal 44 75 b 29	loc 3	transversal 44 75 b 29 BARRANQUILLA

Por otro lado, la Tabla 2 presenta ejemplos del resultado del proceso de geocodificación. Utilizando la columna “Dirección_ciudad” generada en el paso anterior, el modelo procesa las direcciones a través de la API de Google Maps, devolviendo las columnas “Latitud”, “Longitud” y “Dirección Verificada”. Esto permite validar la ubicación de las direcciones y garantizar la calidad de los datos geocodificados.

Tabla 2. Resultados direcciones geocodificadas

Direccion Original	Latitud	Longitud	Direccion Verificada
cl 23 45 08 GALAPA	10,899347	-74,8912657	Cl. 23, Galapa, Atlántico, Colombia
cl 55 a 179 33 SOLEDAD	10,9217546	-74,7930037	Cl. 55a, Soledad, Atlántico, Colombia
cra 50 b 40 89 BARRANQUILLA	10,9895598	-74,7810801	Cra. 50b #40-89, Nte. Centro Historico, Barranquilla, Atlántico, Colombia
cl 70 c 23 c 57 BARRANQUILLA	10,9725948	-74,8107773	Cl. 70c #23c-57, San Felipe, Barranquilla, Atlántico, Colombia
cl 7 56 120 GALAPA	10,9206887	-74,8713625	Cl. 7 #56 - 19, Galapa, Atlántico, Colombia
cl 98 b 9 m 66 BARRANQUILLA	10,9606167	-74,8299821	Cl. 98b #9m-66, Palquema, Barranquilla, Atlántico, Colombia

- Se midieron métricas de rendimiento como el tiempo de procesamiento, la precisión en la clasificación de componentes y la coherencia de los resultados generados por la API de Google Maps.

4.5 Comparación del Proceso Actual y el Propuesto

Finalmente, se realizó una comparación entre el proceso actual y la solución automatizada, considerando criterios como:

- **Tiempo de procesamiento:** Se midieron las diferencias en el tiempo requerido para normalizar, geocodificar y verificar un conjunto de direcciones.
- **Costos:** Se compararon los costos por dirección en el proceso actual y el automatizado.
- **Eficiencia operativa:** Se evaluó la reducción de la carga de trabajo manual, calculada principalmente a partir de la disminución significativa en los tiempos de procesamiento tanto en la normalización, geocodificación y verificación de direcciones.

5. Análisis de resultados

5.1 Comparación de Alternativas

5.1.1 En términos de tiempo

La siguiente tabla (Tabla 3) muestra una comparación entre el proceso manual actual y la solución automatizada propuesta en términos de tiempo. Se puede observar que el modelo automatizado logra reducir de manera significativa los tiempos de normalización, geocodificación y validación, ofreciendo una mejora considerable en la eficiencia operativa.

Tabla 3. Comparación alternativas en tiempo

Alternativas	Cantidad de direcciones	Normalización	Geocodificación	Validación
Modelo actual	1.000	8-12 horas (manualmente)	5-10 minutos	20-60 minutos
Modelo propuesto	1.000	19.88 segundos	5.44 minutos	30 minutos

El modelo actual requiere entre 8 y 12 horas para normalizar 1.000 direcciones manualmente, mientras que el modelo propuesto reduce este tiempo a solo 19.88 segundos, representando una mejora del 1809.9%. La geocodificación también se optimiza, pasando de un rango de 5-10 minutos a solo 5.44 minutos, reduciendo el tiempo en un 37.9%. Además, el proceso de validación es más homogéneo en el modelo automatizado, con un tiempo constante de 30 minutos.

5.1.2 En términos de costo

La (Tabla 4) compara los costos asociados al uso de Driv.in y la API de Google Maps para la geocodificación de direcciones. Se observa que la solución basada en Google Maps es más económica por dirección, ofreciendo una alternativa más sostenible a largo plazo.

Tabla 4. Comparación alternativas en costo

Opciones	Costo total	Cantidad de direcciones	Costo por dirección
Driv.In	1.700.000 COP	25.000	68 COP
API Maps	5 USD	1.000	22 COP

El costo por dirección al utilizar Driv.in es de \$1.700.000 pesos colombianos por un paquete de 25.000 direcciones, lo que equivale a un costo unitario de \$68 pesos colombianos. Por otro lado la API de Google Maps ofrece hasta 40.000 consultas gratuitas, y una vez superado este límite, el costo es de \$5 dólares por cada 1.000 direcciones. Considerando una tasa de cambio de 4.400 COP por dólar, esto resulta en un costo de \$22 pesos colombianos por dirección. Esto representa un ahorro del 209% en comparación con Driv.In. Además, la solución basada en la API de Google Maps elimina la dependencia de proveedores externos y proporciona un mayor control interno sobre el proceso.

5.1.3 Ventajas y desventajas

El modelo actual tiene como ventaja principal la familiaridad y experiencia del equipo administrativo con el proceso manual, lo que facilita la adaptación a imprevistos o correcciones inmediatas. Sin embargo, presenta desventajas significativas, como el alto consumo de tiempo (hasta 12 horas para normalizar 1.000 direcciones) y una mayor posibilidad de errores humanos. Además, depende de proveedores externos como Driv.in para la geocodificación, lo que incrementa los costos y reduce el control operativo.

Por otro lado, el modelo automatizado destaca por su rapidez, reduciendo el tiempo de normalización a tan solo 19.88 segundos y los costos a 22 pesos colombianos por dirección gracias al uso de la API de Google Maps. También elimina la dependencia de proveedores externos, otorgando mayor control interno. Sin embargo, una de sus desventajas es la necesidad inicial de capacitación del personal para el uso de la herramienta y la ausencia de funcionalidades avanzadas como la visualización de geolocalización.

6. Conclusiones y recomendaciones

La implementación del modelo automatizado para la normalización y geocodificación de direcciones demostró ser significativamente más eficiente en términos de tiempo y costos. La reducción del tiempo de normalización de 8-12 horas a solo 19.88 segundos resalta el impacto positivo de la automatización en procesos administrativos repetitivos.

El uso de la API de Google Maps permitió disminuir el costo por dirección a 22 COP, en comparación con los 68 COP de Driv.in, logrando un ahorro del 209%. Este resultado refuerza la viabilidad económica del modelo automatizado para volúmenes altos de datos.

La solución automatizada otorga a Imbocar un mayor control sobre sus procesos operativos al eliminar la necesidad de depender de proveedores externos para la geocodificación, mejorando la seguridad y disponibilidad de los datos.

Aunque la solución automatizada ofrece claras ventajas, se identificó la necesidad de capacitar al equipo administrativo en el uso de las nuevas herramientas. Esto asegura una adopción exitosa y maximiza los beneficios de la automatización.

Para futuros trabajos se recomienda explorar la integración de funcionalidades avanzadas, como la visualización directa de la geolocalización en mapas, para ofrecer mayor valor al equipo operativo.

Referencias

- ArcGIS. (2021). *¿Qué es la geocodificación?* <https://desktop.arcgis.com/es/arcmap/latest/manage-data/geocoding/what-is-geocoding.htm#:~:text=La geocodificación es el proceso,la superficie de la Tierra>
- ArcGIS. (2022). *Modelos preentrenados de ArcGIS*. <https://doc.arcgis.com/es/pretrained-models/latest/text/introduction-to-address-standardization.htm#:~:text=La estandarización de direcciones es,acuerdo con los estándares globales>
- Ghosh, B., Prasad, R., & Pallail, G. (2021). *The automation advantage, Embrace the future of productivity and improve speed, quality, and customer experience through AI*.
- Mora García, L. A. (2023). *Gestión Logística Integral* (Tercera ed). Ecoe Ediciones S.A.S.
- Moreira, D., Cruz, I., Gonzalez, K., & Quirumbay, A. (2021). *Análisis del Estado Actual de Procesamiento de Lenguaje Natural*. 126–137.
- Ridzuan, F., Mohd, W., & Wan, N. (2024). A Review on Data Quality Dimensions for Big Data. *Procedia Computer Science*, 234, 341–348. <https://doi.org/10.1016/j.procs.2024.03.008>
- Vélez, M. P. (2022). *Desarrollo de una alternativa de automatización para el procesamiento de direcciones en el sistema última milla del Grupo Familia*.