

Universidad de La Habana
Facultad de Matemática y Computación



Predicción de la siguiente ubicación utilizando contexto basado en sentimientos

Autor:

Yoan René Ramos Corrales

Tutores:

Dra. Aymée de los Ángeles Marrero Severo
Lic. Ernesto Alfonso Hernández

Trabajo de Diploma
presentado en opción al título de
Licenciado en Ciencia de la Computación

Febrero 2025

github.com/YoanRene/My-Thesis

A todos los que me ayudaron
a construir el camino
hasta aquí

Agradecimientos

Agradezco a mi tutora, Aymée, por sus correcciones y sugerencias, y a mi cotutor, Ernesto, por sus ideas y guía durante el desarrollo de este trabajo. También agradezco a mi fiel lectora, la periodista Claudia, por sus comentarios y sugerencias.

A mis profesores, por su dedicación, sus exigencias y enseñanzas, que no solo han sido fundamentales en mi formación académica, sino que también han inspirado mi crecimiento personal y profesional.

Extiendo mi gratitud a mi familia por su apoyo incondicional, en especial a mi tía Liset, a mi mamá y a mi abuela, por su paciencia y comprensión. A mis amigos de la universidad, desde el primer año, en especial a Emmanuel, el Viti y Rubén, hasta el último año, a Roger Moreno, Marquito Puente 100, Kevin C122 y Juana Carla, por los momentos compartidos.

A mi compañera de estudios, Claudia, por transitar juntos este camino de superación y aprendizaje.

Finalmente, agradezco a todos aquellos que no se mencionan, pero que de alguna manera han contribuido a que esta tesis sea posible.

Opinión del tutor

El presente trabajo de diploma aborda un problema relevante en el ámbito de la movilidad humana, combinando técnicas de aprendizaje profundo con el análisis de sentimientos para la predicción de la siguiente ubicación de un individuo. A lo largo del desarrollo de la investigación, el autor ha demostrado una gran capacidad de análisis y síntesis, logrando integrar conceptos complejos de modelado de secuencias, procesamiento de datos espaciales y técnicas de inteligencia artificial.

Uno de los mayores aciertos de la investigación radica en la exploración del contexto basado en sentimientos como un factor que influye en la predicción de movilidad, lo cual constituye un enfoque innovador en el campo. Si bien los resultados experimentales no evidenciaron una mejora estadísticamente significativa con la incorporación de esta variable, el enfoque abre nuevas líneas de investigación en la intersección entre el comportamiento humano y la inteligencia artificial aplicada a la movilidad.

El autor ha mostrado capacidad para el desarrollo experimental, la implementación de modelos avanzados y el análisis crítico de los resultados. Asimismo, la organización del documento y la claridad expositiva hacen que la lectura sea fluida y comprensible. Se reconoce el esfuerzo en la recopilación de datos, la validación de hipótesis y la aplicación de metodologías rigurosas.

En conclusión, se considera un buen trabajo de investigación, tanto en su planteamiento teórico como en su implementación experimental. Representa una contribución al campo de la predicción de movilidad y establece nuevas bases para futuras investigaciones que incorporen variables contextuales en modelos de aprendizaje profundo. Por todo lo mencionado, se recomienda entonces una buena calificación.

Lic. Ernesto Alfonso Hernández
Cotutor

Resumen

La predicción de la siguiente ubicación de un individuo es crucial para analizar el comportamiento de la movilidad humana. Los modelos existentes que indagan en esta problemática a menudo suelen ignorar información de contexto. Por otra parte, el análisis de sentimientos es una rama de estudio que se ha desarrollado en los últimos años, extrayendo información de redes sociales para predecir el estado emocional de sus usuarios. En este trabajo se propone un modelo que busca predecir la siguiente ubicación a partir de la combinación del contexto espacio-temporal con variables de sentimientos, en pos de mejorar la precisión de la predicción. Para ello, se tiene como objetivo diseñar un modelo de aprendizaje profundo que permita extraer las dependencias complejas entre variables y determinar cómo los estados de un individuo afectan la elección de su siguiente ubicación.

Palabras Clave: predicción de la siguiente ubicación, movilidad humana, sentimientos, aprendizaje profundo

Abstract

Predicting an individual's next location is crucial for analyzing human mobility behavior. Existing models that address this issue often tend to ignore contextual information. On the other hand, sentiment analysis is a field of study that has developed in recent years, extracting information from social networks to predict the emotional state of users. In this work, a model is proposed that aims to predict the next location by combining spatio-temporal context with sentiment variables, with the goal of improving prediction accuracy. To achieve this, the proposal is to design a deep learning model capable of extracting complex dependencies between variables and determining how an individual's emotional states influence the choice of their next location.

Keywords: next location prediction, human mobility, sentiment analysis, deep learning

Índice general

Introducción	1
0.1. Motivación	1
0.2. Antecedentes	2
0.3. Problema de Investigación	2
0.4. Pregunta Científica	2
0.5. Objetivos	3
0.5.1. Objetivo General	3
0.5.2. Objetivos Específicos	3
0.6. Estructura	3
1. Estado del Arte	4
1.1. Modelos de Markov	4
1.2. Modelos de Aprendizaje Profundo	5
1.3. Desafíos del Problema	5
2. Propuesta	8
2.1. Definición del Problema	8
2.2. Metodología	10
2.2.1. Representación del contexto como sentimientos	10
2.2.2. Generación de <i>embeddings</i> espacio-temporales	10
2.2.3. Red de autoatención multicabezal	12
3. Detalles de Implementación y Experimentos	14
3.1. Datos y preprocesamiento	14
3.1.1. Conjunto de datos Geolife	14
3.1.2. Generación de los sentimientos	15
3.2. Entrenamiento del modelo	17
3.3. Modelos de predicción de referencia	18
3.4. Métricas de Evaluación	18
3.5. Resultados	20
3.5.1. Resultados de Desempeño	20

3.5.2.	Influencia del contexto de sentimientos	21
3.5.3.	Impacto de las longitudes de entrada históricas	22
Conclusiones		24
Recomendaciones		25
Bibliografía		27

Introducción

El rápido proceso de urbanización ha incrementado los viajes individuales, presentando desafíos para la sostenibilidad de las ciudades. Para alcanzar los objetivos de desarrollo de las Naciones Unidas (Griggs et al. 2013), los cambios en el comportamiento de la movilidad y los nuevos conceptos para promover estos, desempeñarán un papel fundamental (Martin, Reck y Raubal 2021).

En la salud pública, la pandemia de la COVID-19 es otra prueba de la necesidad de conocer esto a fondo, puesto que la movilidad es un factor determinante en la propagación de enfermedades infecciosas (Kraemer et al. 2020).

Definiciones recientes como Movilidad como Servicio (MaaS) (Reck et al. 2022), carga inteligente (Yanyan Xu et al. 2018) y viaje compartido (H. Huang et al. 2019), dependen de la capacidad de ofrecer servicios personalizados adaptados al contexto del viaje y las características individuales (Z. Ma y P. Zhang 2022). La movilidad individual, es decir, predecir cuándo y dónde ocurrirá un viaje, es esencial para el desarrollo de estos conceptos y, por lo tanto, una herramienta clave para el transporte sostenible.

0.1. Motivación

La predicción de la siguiente ubicación de un individuo, a partir de su información histórica de movilidad, es un problema que ha cobrado gran relevancia en la última década. El auge de los modelos de aprendizaje profundo (Luca et al. 2021) ha impulsado el interés de los investigadores en abordar esta cuestión con métodos basados en el aprendizaje. Al formularse como un problema de predicción de secuencias, similar a tareas en procesamiento del lenguaje natural y de audio, se aplican modelos exitosos en estos campos. En particular, el modelo Transformer (Vaswani et al. 2017), con su mecanismo de autoatención multicabezal, ha revolucionado el modelado de secuencias.

0.2. Antecedentes

Los modelos Transformer se están utilizando en la predicción de la movilidad individual por su capacidad para abordar desafíos específicos:

(1) Capturar las múltiples periodicidades en los patrones de visitas a ubicaciones, las cuales varían entre individuos (Feng et al. 2018).

(2) Modelar la dependencia a largo plazo del comportamiento de movilidad, considerando que la movilidad actual se ve influenciada por comportamientos previos (Cherchi et al. 2017; L. Sun et al. 2013).

0.3. Problema de Investigación

La movilidad humana presenta características únicas, como dependencias espacio-temporales complejas (Feng et al. 2018; F. Li et al. 2020) y la estocasticidad de las visitas a ubicaciones (Song et al. 2010), que dificultan la aplicación directa de modelos de aprendizaje de secuencias.

Predecir la siguiente ubicación a partir del historial de visitas es un desafío. Un modelo preciso debe considerar el contexto que influye en la elección de ubicaciones. Estudios de comportamiento de viaje sugieren que la selección de ubicaciones se correlaciona con aspectos como la disponibilidad de modos de transporte (Ye Hong, Martin y Raubal 2022) y el día de la semana (D. B. E. Dharmowijoyo et al. 2016). ¿Pero qué otros factores influirían en esto? ¿Sería posible que los sentimientos o estados de un individuo afecten la elección de su siguiente ubicación? Por ejemplo, si una persona está de mal humor, puede ser que vaya a un lugar de entretenimiento, como un cine o un bar; mientras que si se encuentra en un estado de enfermedad es probable que vaya a un hospital o a una farmacia. De aquí surge la problemática de estudio:

Predicción de la siguiente ubicación de un individuo utilizando contexto basado en los sentimientos o estados de este.

0.4. Pregunta Científica

¿Cómo influyen los sentimientos o estados de un individuo en la predicción de su siguiente ubicación, y cómo se puede integrar esta información en un modelo computacional para mejorar la precisión de la predicción?

0.5. Objetivos

0.5.1. Objetivo General

Desarrollar un modelo de aprendizaje profundo para la predicción de la siguiente ubicación que incorpore la información de sentimiento o estado del individuo, incluyendo el historial de ubicaciones y datos temporales.

0.5.2. Objetivos Específicos

- Representar el sentimiento a partir de la información que se tiene del conjunto de datos.
- Diseñar un modelo de redes neuronales basado en Transformer que utilice información de ubicación, el sentimiento o estado, y tiempo, para predecir la siguiente ubicación.
- Evaluar la mejora en el rendimiento de la predicción al utilizar el contexto del sentimiento.

0.6. Estructura

La tesis está estructurada en tres capítulos. En el Capítulo 1 se realiza un estudio sobre el marco teórico conceptual del problema. El Capítulo 2 aborda tanto la modelación como la metodología utilizada. En el Capítulo 3 se discute sobre la implementación del modelo, así como de su evaluación mediante la experimentación. Finalmente se presentan las Conclusiones, donde se resumen los aspectos fundamentales y se analiza el cumplimiento de los objetivos propuestos; y las Recomendaciones donde se presenta una relación de trabajos futuros vinculados a la investigación.

Capítulo 1

Estado del Arte

El problema de predicción de la siguiente ubicación de una persona ha encontrado aplicación en diversos campos: como sistemas de recomendación (Xue et al. 2021), redes de sensores (Pirozmand et al. 2014) y análisis del comportamiento de movilidad (D. Wang, P. Wang et al. 2021; Yang Xu et al. 2022).

La definición exacta del problema varía entre los estudios dependiendo de los diferentes objetivos y conjuntos de datos empleados. Por ejemplo, las aplicaciones de redes sociales basadas en la ubicación (LBSN¹) se centran en predecir el siguiente punto de interés (POI²) registrado (D. Wang, K. Liu et al. 2022; Xue et al. 2021). En contraste, los estudios del comportamiento de movilidad buscan comprender la siguiente ubicación de un usuario para realizar una actividad (Solomon et al. 2021). Aquí, la investigación está centrada en los métodos propuestos para aplicaciones de movilidad.

1.1. Modelos de Markov

La última década ha presenciado la expansión de estudios centrados en la predicción de la siguiente ubicación. La Cadena de Markov y sus variantes son probablemente los métodos más empleados para esta tarea (Luca et al. 2021). Estos modelos consideran las ubicaciones como estados y construyen una matriz de transición que codifica la probabilidad de transición entre estados para cada individuo.

Ashbrook y Starner 2002 y Gambis et al. 2012 propusieron identificar ubicaciones significativas a partir de datos GPS y construir un modelo de Markov para predecir las transiciones de ubicación.

Posteriormente, las variantes del modelo de Markov que consideran movimientos colectivos (Chen et al. 2014) e incorporan la importancia de la ubicación (Q. Huang

¹LBSN: Location-Based Social Networks

²POI: Point of Interest

2017) aumentaron aún más el rendimiento de la predicción. Sin embargo, los modelos basados en Markov tienen dificultades para representar los complejos patrones secuenciales en la movilidad humana debido a su supuesto inherente de que el estado actual solo depende de los estados de pasos de tiempo previamente limitados (F. Li et al. 2020).

1.2. Modelos de Aprendizaje Profundo

Los avances recientes en el aprendizaje profundo (DL³) también han promovido su aplicación en la predicción de ubicación. Como método de modelado de secuencias ampliamente adoptado, se ha informado que los modelos basados en redes neuronales recurrentes (RNN⁴), como es el caso del modelo de memoria a corto y largo plazo (LSTM⁵) (Solomon et al. 2021) y la RNN espacio-temporal (ST⁶) (Q. Liu et al. 2016), superan a los modelos de Markov por un amplio margen en esta tarea.

Aún así, los modelos RNN convencionales tienden a subestimar las dependencias a largo plazo cuando aumenta la longitud de la secuencia de entrada. Por lo tanto, los estudios emplearon un mecanismo de atención para capturar dinámicamente las dependencias tanto a corto como a largo plazo (Feng et al. 2018; F. Li et al. 2020).

Además, el modelo Transformer, que se basa en el mecanismo de autoatención multicabezal (MHSA⁷) (Vaswani et al. 2017), ha comenzado a ganar interés en el campo. En particular, Xue et al. 2021 propusieron MobTcast para considerar varios contextos con una estructura basada en Transformer y obtuvieron resultados de vanguardia en la predicción de POI para datos LBSN.

A pesar de tener un gran potencial para aprender las complejas dependencias espacio-temporales, pocos estudios han aplicado Transformer al problema de predicción de ubicación debido al acercamiento reciente de estos modelos a la problemática en cuestión.

1.3. Desafíos del Problema

Comprender los factores que afectan la elección de la ubicación de la actividad es beneficioso para predecir la movilidad de las personas, ya que pueden considerarse como conocimiento previo y potencialmente guiar el aprendizaje de los modelos DL. En el entorno de la movilidad, la elección de ubicaciones se considera una parte

³DL: Deep Learning

⁴RNN: Recurrent Neural Networks

⁵LSTM: Large Short-Term Memory

⁶ST: Spatial-Temporal

⁷MHSA: Multi-Head Self-Attentional

integral del comportamiento de actividad-viaje de las personas y se ha estudiado dentro del marco basado en la actividad (Schönfelder y Axhausen 2016).

Los estudios centrados en analizar el comportamiento de viaje a lo largo del tiempo sugieren que se encuentra tanto estabilidad como variabilidad en las elecciones de ubicación de actividad de las personas. Por ejemplo, D. B. Dharmowijoyo et al. 2017 mostraron que la variabilidad de las visitas a ubicaciones es mucho mayor entre los pares fin de semana-día laborable que entre los pares día laborable-día laborable y fin de semana-fin de semana.

Los estudios empíricos también demuestran la correlación de diferentes aspectos del comportamiento de viaje individual. Por ejemplo, Susilo y Axhausen 2014 informaron una alta repetición en las combinaciones de ubicación-modo de viaje, lo que sugiere que las personas usan el mismo modo de viaje para llegar a sus ubicaciones. Ye Hong, Martin, Xin et al. 2022 llegaron a conclusiones similares, donde encontraron que solo un subconjunto de todas las combinaciones de ubicación-modo es esencial para describir el comportamiento de movilidad. Desde esta perspectiva, los aspectos del comportamiento de viaje pueden considerarse restricciones para la elección de ubicaciones de actividad de las personas.

Un problema similar a la predicción de la siguiente ubicación es la formulación del conjunto de opciones de ubicación de un individuo, que es un componente crucial en los modelos de microsimulación de tráfico (Mariante et al. 2018). En lugar de predecir la siguiente ubicación exacta, el problema se centra en generar un conjunto que contenga todas las ubicaciones posibles. Basándose en la teoría de la geografía del tiempo, se ha aplicado el análisis de áreas de rutas potenciales para abordar el problema, lo que sugiere que el conjunto de opciones está limitado por el tiempo de viaje (Scott y He 2012), la hora del día (Yoon et al. 2012) y el modo de viaje disponible (Neutens et al. 2007). Sin embargo, este conocimiento empírico no se utiliza plenamente en los modelos para predicciones de ubicación.

La mayoría de los estudios existentes solo utilizan secuencias de visitas de ubicaciones por los individuos para predecir sus próximas ubicaciones, independientemente del contexto de movilidad actual (Laha y Putatunda 2018). Esto se debe a la complejidad y diversidad de los datos de contexto (Tedjopurnomo et al. 2020).

Con el desarrollo del aprendizaje automático, los estudios han comenzado a explorar métodos para la fusión de datos de contexto y emplearlos para impulsar la investigación de la movilidad humana (Lau et al. 2019; X. Zheng et al. 2018). Sin embargo, los datos de contexto de múltiples fuentes poseen formatos complejos, lo que dificulta su fusión en una representación unificada (Liao et al. 2018).

Los marcos de DL sensibles al contexto proporcionan una solución para combinar información de contexto de múltiples fuentes (G. Sun et al. 2022), pero con mayor frecuencia se centran en el aspecto temporal y pasan por alto las interacciones espacio-temporales complejas.

Mencionar que la evolución de la problemática y el auge de los grandes modelos de lenguaje natural (LLM⁸) han puesto a investigar otras formas de atacar el asunto en cuestión con muy buenos resultados (X. Wang et al. 2024). De aquí y del tema de la utilización del contexto para mejorar la predicción, que se mantenga la necesidad de explorar más a fondo las influencias de la información de contexto del individuo.

En particular, se hará énfasis en el uso de sentimientos o emociones que puedan influir en la elección de la siguiente ubicación. Dichos sentimientos, que en un escenario ideal podrían ser extraídos de las redes sociales, por falta de datos públicos se hace casi inviable. Entonces sería práctico el empleo de los ya mencionados modelos del lenguaje natural para inferir un sentimiento a partir de los datos que se tienen, como lo son la ubicación y el momento actual del individuo. Luego, queda hacer uso de un modelo que tenga en cuenta dichos sentimientos para mejorar la predicción de la siguiente ubicación.

Se propone entonces utilizar el modelo de Aprendizaje Profundo basado en Transformers y MHSA de Ye Hong, Y. Zhang et al. 2023 por sus excelentes resultados en esta problemática y su capacidad de capturar dependencias.

⁸LLM: Large Language Models

Capítulo 2

Propuesta

2.1. Definición del Problema

A continuación aparecen un conjunto de términos y nociones que serán utilizados en el resto del documento y habrá una formulación del problema de predicción de la siguiente ubicación. Los datos de movilidad son recopilados típicamente a través de dispositivos electrónicos y se almacenan como trayectorias espacio-temporales. Cada punto de seguimiento en la trayectoria de un usuario contiene un par de coordenadas espaciales y una marca de tiempo.

Definición 1 (Trayectoria GNSS¹). Sea u_i un usuario del conjunto de usuarios $\mathcal{U} = \{u_1, \dots, u_{|\mathcal{U}|}\}$, una trayectoria $T_i = (q_k)_{k=1}^{n_{u_i}}$ es una secuencia ordenada en el tiempo compuesta por n_{u_i} puntos de seguimiento visitados por el usuario u_i . Un punto de seguimiento se puede representar como una tupla $q = \langle p, t \rangle$, donde $p = \langle x, y \rangle$ representa las coordenadas espaciales en un sistema de referencia, en este caso latitud y longitud, y t es el tiempo de registro.

Los puntos de permanencia se detectan a partir de trayectorias GNSS sin procesar para identificar áreas donde los usuarios permanecen estacionarios durante un período mínimo de tiempo (Q. Li et al. 2008). Luego, las ubicaciones se forman mediante la agregación espacial de puntos de permanencia para caracterizar la semántica del lugar (Hariharan y Toyama 2004; Martin, Ye Hong et al. 2023).

¹GNSS: Global Navigation Satellite System

Definición 2 (Punto de Permanencia). Un punto de permanencia $S = (q_k)_{k=start}^{end}$ es una subsecuencia de la trayectoria T_i donde el usuario u_i estuvo estacionario desde el punto de seguimiento inicial q_{start} hasta el punto de seguimiento final q_{end} . Cada punto de permanencia S se puede representar como una tupla $\langle t, d, g(s) \rangle$, donde t y d representan la marca de tiempo de inicio y la duración de la permanencia, respectivamente, y $g(s)$ denota la geometría, a menudo representada como el centro de sus puntos de seguimiento. S_k es usada para denotar el k -ésimo punto de permanencia en la trayectoria GNSS de un usuario.

Definición 3 (Ubicación). Una ubicación L consiste en un conjunto de puntos de permanencia espacialmente próximos. Se puede representar como una tupla $L = \langle l, g(l) \rangle$, donde l es el identificador de la ubicación, y $g(l)$ denota la geometría de la ubicación, calculada como la envolvente convexa de todos los puntos de permanencia contenidos. Por lo tanto, cada ubicación se define como un área. \mathcal{O}_i es definida como el conjunto que contiene las ubicaciones conocidas para el usuario u_i , y $\mathcal{O} = \{\mathcal{O}_1, \dots, \mathcal{O}_{|\mathcal{U}|}\}$ como el conjunto que contiene todas las ubicaciones.

Mediante la generación de ubicaciones cada punto de permanencia se enriquece, y si es añadida también información de contexto que representa un sentimiento del usuario en dicho punto, llámese a este f , se tiene $S = \langle t, d, g(s), f, l, g(l) \rangle$; y entonces la movilidad de un usuario se puede representar como una secuencia ordenada en el tiempo de N puntos de permanencia visitados $(S_k)_{k=1}^N$. A continuación, el problema de predicción de la siguiente ubicación sería:

Problema 1 (Predicción de la Siguiete Ubicación). Considere una secuencia de puntos de permanencia con información de contexto $(S_k)_{k=m}^n$ visitada por el usuario u_i en una ventana de tiempo desde el paso de tiempo m hasta n . El objetivo es predecir la ubicación que el mismo usuario visitará en el siguiente paso de tiempo, es decir, el identificador de ubicación $l_{n+1} \in \mathcal{O}$.

La longitud de la ventana temporal determina cuánta información histórica se considera en el modelo predictivo. Aquí, se construye la secuencia histórica teniendo como base la movilidad realizada en los últimos $D \in \{0, 1, \dots, 14\}$ días. Por lo tanto, la longitud de la ventana histórica depende del usuario u_i y del paso de tiempo actual n . La predicción de la siguiente ubicación se define como un problema de predicción de secuencia con longitudes de secuencia variables.

2.2. Metodología

Por sus resultados destacados, proponemos hacer uso de la red neuronal que utiliza información de contexto para abordar la predicción de la siguiente ubicación de Ye Hong, Y. Zhang et al. 2023 con algunas modificaciones para adaptarla a la problemática. Primero, representamos el contexto como los sentimientos de los usuarios en los puntos de permanencia. Luego, el modelo utiliza varias capas de *embedding*² para representar los datos heterogéneos de movimiento y contexto. Finalmente, se adapta la red MHSA para aprender las dependencias de la secuencia histórica e inferir la siguiente ubicación visitada. A continuación, se proporciona una descripción detallada de cada módulo.

2.2.1. Representación del contexto como sentimientos

Para capturar el contexto de los puntos de permanencia utilizamos los sentimientos de los usuarios en estos puntos. Sentimientos que son obtenidos a partir de un LLM capaz de inferirlos a partir de la ubicación, hora, día de la semana, así como del punto de interés visitado (hospital, restaurante, etc). Si bien, al ser un LLM no entrenado para este propósito no es capaz de inferir sentimientos con total precisión asumimos que pueda ser suficiente para el propósito de este trabajo.

Los sentimientos utilizados son los siguientes: *miedo*, *hambre*, *enfermedad*, *indiferencia*, *cansancio*. Fueron escogidos estos y no otros por la motivación de establecer puentes futuros entre los resultados de Hernández 2023, y los de este trabajo, en pos de ayudar a la toma de decisiones en el ámbito de la salud pública.

Una vez inferidos por el modelo del lenguaje, los sentimientos pasan a ser representados como vectores de *embedding* y son utilizados para entrenar el modelo de predicción de ubicación; e_{f_k} representa el *embedding* del sentimiento f en el punto de permanencia S_k .

2.2.2. Generación de *embeddings* espacio-temporales

Un modelo preciso de predicción de ubicación requiere una selección y modelado adecuados de la información de la secuencia histórica. Además del identificador de ubicación sin procesar y la hora de visita correspondiente que se incluyen a menudo (F. Li et al. 2020), consideramos la duración de la actividad y las funciones de uso del suelo $g(s), g(l)$ para describir cada punto de permanencia visitado, lo que garantiza una representación completa de su contexto desde una perspectiva espacio-temporal.

²*embedding*: técnica de aprendizaje automático que convierte datos de entrada en representaciones matemáticas permitiendo capturar relaciones semánticas y estructurales entre los datos.

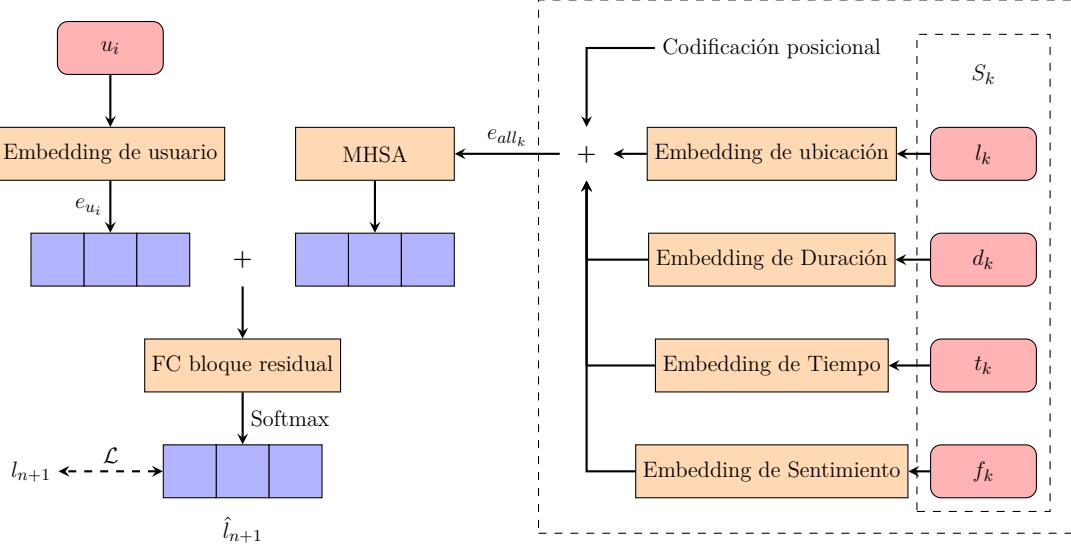


Figura 2.1: Capas de embedding y la red basada en MHA para la predicción de la próxima ubicación.

Además, la información relacionada con el usuario ayuda a descubrir las secuencias recorridas por diferentes de ellos y ayuda a la red a aprender patrones de movimiento específicos del usuario.

Se utilizan capas de *embedding* para representar características del tipo categórico a un vector de valores reales. A diferencia de la representación *one-hot*³ más clásica, los vectores de *embedding* son más compactos y pueden capturar eficazmente la correlación latente entre diferentes tipos de características (Yang Xu et al. 2022). Estas capas son matrices de parámetros que proporcionan mapeos entre la variable original y el vector de *embedding*, optimizadas conjuntamente con toda la red. El proceso de *embedding*, así como la arquitectura del modelo, se muestran en la Figura 2.1.

Operacionalmente, dado un punto de permanencia S_k en la secuencia histórica, su identificador de ubicación l_k , la hora de llegada t_k y la duración de la estancia d_k se introducen en sus respectivas capas de *embedding* para generar representaciones vectoriales:

$$e_{l_k} = h_l(l_k; \mathbf{W}_l), e_{t_k} = h_t(t_k; \mathbf{W}_t), e_{d_k} = h_d(d_k; \mathbf{W}_d) \quad (1)$$

donde e_{l_k} , e_{t_k} y e_{d_k} son los respectivos vectores de *embedding* para l_k , t_k y d_k . En

³*one-hot*: Es una codificación que representa variables categóricas como vectores binarios donde cada elemento corresponde a una categoría. Solo el elemento que representa la categoría activa se establece en 1; todos los demás son 0. Una técnica común para manejar datos categóricos en aprendizaje automático.

el caso de $h(\cdot; \cdot)$ denota la operación de *embedding* y los términos \mathbf{W} son las matrices de parámetros optimizadas durante el entrenamiento. Con *embedding* se convierte por separado los minutos, la hora y el día de la semana a partir de la hora de llegada t_k para capturar diferentes niveles de periodicidad en las visitas históricas.

Finalmente, el vector de *embedding* general e_{all_k} para el punto de permanencia S_k se obtiene añadiendo sus características espacio-temporales, así como su contexto basado en sentimientos e_{f_k} , junto con una codificación posicional PE que codifica la información de secuencia k :

$$e_{all_k} = e_{l_k} + e_{t_k} + e_{d_k} + e_{f_k} + PE \quad (2)$$

El modelo usa la codificación posicional original propuesta por Vaswani et al. 2017 que utiliza funciones seno y coseno. La inclusión de la codificación posicional es esencial para entrenar una red de autoatención, ya que no asume implícitamente el orden secuencial de su entrada (Vaswani et al. 2017). Además, se representa al usuario u_i del cual se registra la secuencia de puntos de permanencia en un vector e_{u_i} con una capa de *embedding* de usuario, es decir, $e_{u_i} = h_u(u_i; W_u)$. La inclusión de la información del usuario asegura que un modelo entrenado con datos de población aún pueda distinguir las trayectorias recorridas por diferentes usuarios. Como resultado, obtenemos el vector de *embedding* general e_{all_k} que codifica las características espacio-temporales y de contexto, y el vector de *embedding* de usuario e_{u_i} para la secuencia.

2.2.3. Red de autoatención multicabezal

Una vez se adquieren los vectores de características espacio-temporales densos en cada paso de tiempo, debemos extraer sus patrones de transición secuencial. Estos patrones históricos se capturan utilizando una red basada en MHSA, un mecanismo propuesto originalmente dentro de la red transformadora para abordar las tareas de traducción de idiomas (Vaswani et al. 2017). Se adopta una arquitectura similar a las redes GPT⁴ que solo incluye la parte del decodificador del modelo transformador (Radford et al. 2018). El decodificador consta de una pila de L bloques idénticos, cada uno con dos componentes. El primero es la red de autoatención multicabezal enmascarada y el segundo es una red de avance con dos capas lineales, separadas por una función de activación ReLU⁵. Son agregadas conexiones residuales, normalización de capa y capas de abandono a cada componente para facilitar el aprendizaje.

La salida del modelo MHSA out_n se añade al *embedding* del usuario e_{u_i} y juntas

⁴GPT: Generative Pre-trained Transformer

⁵ReLU (Rectified Linear Unit): se define como $f(x) = \max(0, x)$, introduciendo no linealidad y siendo computacionalmente eficiente.

se introducen en un bloque residual completamente conectado (FC⁶). Finalmente, la probabilidad predicha de cada ubicación se obtiene mediante una transformación *softmax*⁷:

$$P(\hat{l}_{n+1}) = \text{Softmax}(f_{FC}(\text{out}_n + e_{u_i}; \mathbf{W}_{FC})) \quad (3)$$

donde $f_{FC}(\cdot; \cdot)$ representa la operación del bloque residual FC. Este bloque consta de capas lineales con conexiones residuales, con el objetivo de aprender las dependencias entre la información de la secuencia y el usuario para extraer las preferencias de movilidad personal. $P(\hat{l}_{n+1}) \in \mathbb{R}^{|\mathcal{O}|}$ contiene la probabilidad de que se visiten todas las ubicaciones en el siguiente paso de tiempo.

Durante el entrenamiento, con acceso a la siguiente ubicación real l_{n+1} , la tarea puede considerarse como un problema de clasificación multiclase. Por lo tanto, los parámetros del modelo se pueden optimizar utilizando la pérdida de entropía cruzada multiclase \mathcal{L} :

$$\mathcal{L} = - \sum_{k=1}^{|\mathcal{O}|} P(l_{n+1})(k) \log(P(\hat{l}_{n+1})(k)) \quad (4)$$

donde $P(\hat{l}_{n+1})(k)$ representa la probabilidad predicha de visitar la k -ésima ubicación y $P(l_{n+1})(k)$ es la verdad representada por *one-hot*, es decir, $P(l_{n+1})(k) = 1$ si la siguiente ubicación real es la k -ésima ubicación, y $P(l_{n+1})(k) = 0$ en caso contrario.

⁶FC: Fully Connected

⁷*softmax*: Es una función que convierte un vector de valores reales en una distribución de probabilidad.

Capítulo 3

Detalles de Implementación y Experimentos

3.1. Datos y preprocesamiento

Se demuestra la efectividad del método propuesto mediante la experimentación en un conjunto de datos de seguimiento GNSS.

3.1.1. Conjunto de datos Geolife

El presente es un conjunto de datos público de trayectorias GPS recopilado en el proyecto GeoLife de Microsoft Research Asia por 182 usuarios durante un período de más de tres años (desde abril de 2007 hasta agosto de 2012) (Y. Zheng, Fu et al. 2011, julio). En este conjunto se registra una amplia gama de movimientos al aire libre de los usuarios, que incluyen no solo rutinas diarias como ir a casa o al trabajo, sino también algunas actividades de entretenimiento y deportivas, como compras, turismo, cenas, senderismo y ciclismo.

Una trayectoria GPS de este conjunto de datos está representada por una secuencia de puntos con marca de tiempo, cada uno de los cuales contiene información de latitud, longitud y altitud. Esta serie de datos contiene 17,621 trayectorias con una distancia total de aproximadamente 1.2 millones de kilómetros y una duración total de más de 48,000 horas. Estas trayectorias fueron registradas por diferentes dispositivos GPS y teléfonos con GPS, y tienen una variedad de tasas de muestreo. El 91 por ciento de las trayectorias se registraron en una representación densa, por ejemplo, cada 1-5 segundos o cada 5-10 metros por punto.

Se sigue el marco propuesto por Y. Zheng, Xie, W.-Y. Ma et al. 2010 y se clasifica las trayectorias GPS en puntos de permanencia y etapas de movimiento utilizando la biblioteca Trackintel (Martin, Ye Hong et al. 2023).

Las trazas de movimiento de los estudios de seguimiento GNSS se preprocesan para la predicción de la siguiente ubicación. Pre-filtramos los conjuntos de datos para considerar solo a los usuarios observados por más de 50 días en Geolife, garantizando un tiempo de observación prolongado. Se utiliza la cobertura temporal de seguimiento, que cuantifica la proporción de tiempo en que se registran los desplazamientos de los usuarios, para evaluar así la calidad del seguimiento en la dimensión temporal. Después de este proceso, permanecen 45 usuarios en Geolife.

Las ubicaciones se generan a partir de la secuencia individual de puntos de permanencia visitados. Consideramos un punto de permanencia como una actividad si su duración es superior a 25 minutos. Luego, los puntos de permanencia de actividad se agrupan espacialmente en ubicaciones para considerar visitas al mismo lugar en diferentes momentos. Se utiliza la función proporcionada en Trackintel con los parámetros $\epsilon = 20$ y `num_samples = 2` para generar las ubicaciones del conjunto de datos (Yihong Hong et al. 2021). La Tabla 3.1 muestra estadísticas básicas para el conjunto de datos.

Tabla 3.1: Estadísticas básicas del conjunto de datos de movilidad. Se reportan la media y la desviación estándar entre los usuarios.

	Geolife
Número de usuarios	45
Período de seguimiento (días)	345 ± 413
#Puntos de permanencia por usuario	369 ± 456
#Puntos de permanencia por usuario por día	$2,4 \pm 1,5$
#Ubicaciones por usuario	77 ± 108
Tamaño de las ubicaciones (m^2)	3606 ± 12275
Cobertura de seguimiento (%)	44 ± 24

3.1.2. Generación de los sentimientos

Para la generación de sentimientos, se utilizó el modelo de Google *Gemini-2.0-Flash-Thinking-Exp* como nuestro LLM para inferir los sentimientos a partir de los datos disponibles. Este modelo fue seleccionado debido a su capacidad de procesamiento y a la disponibilidad de una API de acceso gratuito (Google AI for Developers 2024).

Preprocesamiento de Datos

Con el objetivo de que el LLM pudiera inferir los sentimientos a partir de la información de ubicación, hora y día, se realizó un proceso de transformación de estos datos para hacerlos más comprensibles para el modelo. La ubicación, originalmente representada en coordenadas de latitud y longitud, fue convertida a una dirección en lenguaje natural mediante el uso de la técnica de *reverse geocoding*, utilizando la API proporcionada por Geocode Maps 2025. La hora fue transformada al formato *HH:MM AM/PM*, y el día fue representado como el día de la semana correspondiente. Con esta información, se construyó un nuevo conjunto de datos en formato CSV, el cual fue utilizado como entrada para el LLM.

Construcción del Prompt

El prompt utilizado para guiar al LLM en la inferencia de sentimientos fue diseñado a través de un proceso iterativo de prueba y error, con el fin de asegurar que las respuestas generadas fueran coherentes y relevantes. El prompt utilizado fue el siguiente:

```
f""You are a sentiment analysis expert.
You will be given a dataset and you will have to create a
situational context for each row of this dataset, with the provided
information, that it is just the time and location;
from this context your main goal is to identify a sentiment from the
following list of sentiments: {sentiments}.
You will have to return the sentiment that is most prominent in
the situational context.
If you are unable to identify the sentiment,
you will have to return '{default_sentiment}'. You will have to
return only the sentiment and a brief explanation of why you chose
that sentiment.
The Output should be in a structured CSV with columns:
index, sentiment, explanation. Provide only CSV-formatted output.
The index of the output must be the same of the input.
The sentiment must be in lowercase.
The explanation should be between double quotes and
can't have chinese characters.
You must provide output for all rows in the input.
```

Example:

```
Input: 17,968,Friday,Friday,10:35 AM,12:32 PM,
```



```

    "KFC, Chengfu Road, Wudaokou, Dongsheng, Haidian District,
    Beijing, 100190, China"
    Output: 17,hunger,"Being at KFC during late morning/noon
    suggests hunger for lunch."
    """

```

En este prompt, *sentiments* hace referencia a la lista de sentimientos mencionados en la sección 2.2.1, y *default_sentiment* corresponde a uno de estos sentimientos, específicamente la indiferencia.

Validación de Respuestas

Para garantizar la consistencia de las respuestas generadas por el LLM, se implementó un proceso de validación que asegura que:

- El índice de la filas de salida coincida con el índice de las filas de entrada.
- Los sentimientos devueltos pertenezcan al conjunto de sentimientos predefinidos.
- El formato de salida cumpla con la estructura especificada: *index*, *sentiment*, *explanation*.

En caso de que alguna de estas condiciones no se cumpliera, se realizaba una nueva petición al LLM hasta obtener una respuesta válida.

Conversión de Sentimientos a Valores Numéricos

Finalmente, los sentimientos inferidos por el LLM fueron convertidos a valores numéricos en el rango de 0 a 4, inclusive. Estos valores fueron incorporados al conjunto de datos utilizado para el entrenamiento del modelo principal.

3.2. Entrenamiento del modelo

Se divide el conjunto de datos, con la información de los sentimientos añadida, en conjuntos de entrenamiento, validación y prueba sin superposición, con una proporción de 6:2:2 basada en el tiempo. Para cada usuario, las secuencias de puntos de permanencia correspondientes al primer 60% de los días de seguimiento se emplean para el entrenamiento del modelo, mientras que el 20% final se reserva para la fase de prueba. Los parámetros de la red de predicción se ajustan utilizando el conjunto de entrenamiento, y el conjunto de validación se emplea para monitorear la pérdida

del modelo. Se efectúa *grid search*¹ sobre los hiperparámetros en el conjunto de validación. Finalmente, se evalúa y se reporta el desempeño del modelo utilizando el conjunto de prueba.

Durante el entrenamiento, se minimiza la Ecuación (4) con el optimizador Adam sobre lotes de muestras de datos de entrenamiento, con una tasa de aprendizaje inicial de $1e^{-3}$ y una penalización L2 de $1e^{-6}$. Se adopta una estrategia de parada temprana para detener el aprendizaje si la pérdida de validación deja de disminuir durante 3 épocas. Luego, la tasa de aprendizaje se multiplica por 0.1 y el entrenamiento se reanuda desde el modelo con la menor pérdida de validación. Este proceso de parada temprana se repite 3 veces. Además, se implementa un calentamiento de la tasa de aprendizaje durante 2 épocas y una decaída lineal de 0.98 por época posteriormente (Vaswani et al. 2017).

3.3. Modelos de predicción de referencia

Se compara el rendimiento de este modelo con el método clásico de predicción de ubicación basado en Markov y con el modelo de predicción de ubicación sobre el que está construido el modelo propuesto.

- **Markov.** Los modelos clásicos de predicción de ubicación asumen la propiedad de Markov en las visitas a ubicaciones individuales (Ashbrook y Starner 2002). Implementamos la Cadena de Markov de primer orden (1-MMC) (Gambs et al. 2012), ya que aumentar el orden no mejora el rendimiento de la predicción.
- **MHSA Transformer.** El modelo MHSA Transformer es el modelo sobre el que se construye nuestro modelo propuesto (Ye Hong, Y. Zhang et al. 2023). La diferencia entre estos radica en que el primero no utiliza información de contexto adicional para realizar la predicción de ubicación, como es el caso de los sentimientos.

3.4. Métricas de Evaluación

Se utilizan las siguientes métricas para cuantificar el rendimiento de los modelos implementados:

¹*grid search*: Técnica tradicional de optimización de hiperparámetros, que no es más que una búsqueda exhaustiva a través de un subconjunto especificado manualmente del espacio de hiperparámetros de un algoritmo de aprendizaje. Un algoritmo de *grid search* debe guiarse por alguna métrica de rendimiento, normalmente medida por validación cruzada en el conjunto de entrenamiento o evaluación en un conjunto de validación de espera.

- **Exactitud (Accuracy).** Mide la corrección de la ubicación predicha en comparación con la ubicación real visitada a continuación. Prácticamente, se ordena el vector de probabilidades de ubicación $P(\hat{l}_{n+1})$, obtenido de la Ecuación (3), en orden descendente y se verifica si la ubicación real aparece entre las k mejores predicciones, $\text{Acc}@k$ mide la proporción de veces que esto es cierto en el conjunto de prueba. En la literatura sobre predicción de ubicaciones, esta métrica también se conoce como $\text{Recall}@k$ o $\text{Hit Ratio}@k$. Se reportan $\text{Acc}@1$, $\text{Acc}@5$ y $\text{Acc}@10$ para permitir comparaciones con otros trabajos.
- **Puntaje F1 (F1).** Las visitas individuales a ubicaciones son altamente desbalanceadas, con ubicaciones específicas ocurriendo con mayor frecuencia en la rutina diaria que otras. Utilizamos la puntuación F1 ponderada por el número de visitas para enfatizar el rendimiento del modelo en las ubicaciones más importantes.
- **Rango recíproco medio (MRR).** Calcula el promedio del recíproco del rango en el que se recuperó la primera entrada relevante en el vector de predicción:

$$MRR = \frac{1}{N} \sum_{i=1}^N \frac{1}{\text{rank}_i} \quad (5)$$

donde N denota el número de muestras de prueba y rank_i es el rango de la ubicación real en $P(\hat{l}_{n+1})$ para la i -ésima muestra de prueba.

- **Ganancia acumulativa con descuento normalizada (NDCG).** Mide la calidad del vector de predicción por la relación entre la ganancia acumulativa con descuento (DCG) y la ganancia acumulativa con descuento ideal (IDCG):

$$NDCG = \frac{1}{N} \sum_{i=1}^N \frac{DCG_i}{IDCG_i}, \quad \text{donde} \quad DCG_i = \sum_{j=1}^{|\mathcal{O}|} \frac{r_j}{\log_2(j+1)} \quad (6)$$

donde r_j denota el valor de relevancia en la posición j . En el contexto de la predicción de ubicaciones, r_j es binario, es decir, $r_j \in \{0, 1\}$, y $r_j = 1$ si y solo si el j -ésimo elemento en el ordenado $P(\hat{l}_{n+1})$ corresponde a la ubicación real siguiente. $\text{NDCG}@k$ mide la relevancia de los resultados hasta la posición k en el ranking. En la evaluación, se reporta $\text{NDCG}@10$.

3.5. Resultados

3.5.1. Resultados de Desempeño

Tabla 3.2: Resultados de evaluación del rendimiento para la predicción de la próxima ubicación sobre el conjunto de datos Geolife. Se reporta el promedio y la desviación estándar en 12 ejecuciones diferentes.

Método	Acc@1	Acc@5	Acc@10	F1	MRR	NDCG@10
1-MMC	24.1	38.1	39.5	22.7	30.5	32.7
MHSA Transformer	29.4 ± 0.8	53.6 ± 1.4	57.8 ± 1.1	19.7 ± 1.5	40.2 ± 0.8	44.2 ± 0.9
Nuestro modelo	29.6 ± 0.8	53.8 ± 1.5	57.8 ± 1.5	20.2 ± 1.0	40.5 ± 0.7	44.4 ± 0.8

Se presenta primero el desempeño de predicción para todos los métodos considerados en la Tabla 3.2. Para cada modelo basado en aprendizaje, se entrena el modelo doce veces con diferentes inicializaciones aleatorias de los parámetros y reportamos la media y la desviación estándar de los indicadores de desempeño respectivos. Se utiliza la prueba U de Mann-Whitney para verificar si las diferencias de desempeño entre los distintos modelos son significativas. Los modelos de aprendizaje profundo (DL) se entrenan en los datos de toda la población, utilizando identificadores de usuario para distinguir las secuencias registradas de diferentes usuarios. Se introducen secuencias históricas de los últimos $D = 7$ días en todos los modelos DL para garantizar su comparabilidad.

Se reporta que MHSA Transformer supera al método 1-MMC en todos los indicadores excepto en la puntuación F1. La brecha de desempeño es grande en Acc@5, Acc@10, MRR y NDCG@10, lo que implica que MHSA Transformer puede identificar mejor las preferencias del usuario al considerar el conocimiento colectivo de movilidad. La puntuación F1 relativamente alta del método 1-MMC sugiere que este método es práctico si el objetivo es la predicción de ubicaciones esenciales. Sin embargo, sus desempeños siguen siendo significativamente inferiores a los modelos basados en DL en el conjunto de datos considerado. Esta diferencia enfatiza la importancia de considerar dependencias a largo plazo y contextos espaciotemporales en la tarea de predicción.

Tabla 3.3: Resultados de la prueba U de Mann–Whitney para la comparación de MHSA Transformer y el modelo propuesto.

Métrica	Valor p
NDCG@10	0.750832
F1	0.157213
Acc@1	0.772734
Acc@5	0.862312
Acc@10	0.685977
MRR	0.506721

El modelo propuesto, que utiliza como contexto los sentimientos, obtiene los mejores resultados en todos los indicadores, presentando ligeramente mejores promedios en comparación con el MHSA Transformer. Sin embargo, la comparación de los resultados entre ambos enfoques revela diferencias muy pequeñas, con variaciones en el orden de 0.2 a 0.5 puntos en promedio, que se encuentran dentro del rango de las desviaciones estándar reportadas. Además, la aplicación de la prueba de Mann–Whitney U a cada métrica resultó en valores p superiores a 0.05 (Ver Tabla 3.3), lo que indica que no existen diferencias estadísticamente significativas entre los métodos. Estas mínimas discrepancias en el desempeño sugieren que las pequeñas mejoras observadas en nuestro modelo podrían atribuirse a la variabilidad inherente en las 12 ejecuciones experimentales, en lugar de a una ventaja real del uso del contexto de sentimientos. En conjunto, estos resultados apoyan la conclusión de que, aunque el enfoque de incorporar sentimientos como contexto logra obtener promedios ligeramente superiores en los indicadores evaluados, ambos métodos presentan un rendimiento prácticamente similar en la predicción de la próxima ubicación sobre el conjunto de datos Geolife, sin evidencias concluyentes de que uno supere significativamente al otro.

3.5.2. Influencia del contexto de sentimientos

Aunque los resultados experimentales muestran que la incorporación del contexto de sentimientos en nuestro modelo no genera mejoras estadísticamente significativas en comparación con el MHSA Transformer, es importante analizar en detalle cómo este contexto influye en la predicción de la próxima ubicación.

Uno de los factores que pueden haber afectado en el desempeño de nuestro modelo es la manera en que se incorporó el contexto de sentimientos. En este caso, los sentimientos fueron inferidos mediante un LLM utilizando únicamente información como la hora, el lugar y el día de la semana. Dado que el LLM no tenía acceso a información subjetiva o explícita de los usuarios, las etiquetas de sentimientos generadas no

capturan de manera precisa el estado real de los individuos en cada momento.

Además, debido a la aleatoriedad inherente en la generación de texto por parte de los modelos de lenguaje, es posible que la asignación de sentimientos no haya sido consistente a los datos de los que se infiere, lo que introduce una fuente adicional de ruido en los datos. Esta incertidumbre en la inferencia del contexto de estado de los individuos puede haber reducido el impacto del componente de sentimientos en la tarea de predicción, haciendo que las diferencias con el modelo base no sean lo suficientemente pronunciadas para alcanzar significancia estadística.

Otro aspecto a considerar es que, aunque el contexto de sentimientos puede ser útil en ciertos escenarios, su relevancia en la predicción de la próxima ubicación podría depender de la naturaleza de los datos de movilidad. Si los patrones de movimiento de los usuarios están mayormente determinados por factores rutinarios o estructurales (por ejemplo, trabajo, estudios, transporte público), el efecto de los sentimientos en la predicción podría ser marginal. Para evaluar esto, futuras investigaciones podrían explorar la incorporación de fuentes de datos más ricas, como registros de actividad en redes sociales o encuestas directas, para mejorar la calidad de la inferencia emocional y analizar si una mayor precisión en la estimación del estado de los individuos impacta significativamente en la predicción de movilidad.

En conclusión, aunque la integración del contexto de sentimientos en nuestro modelo no mostró mejoras significativas respecto al MHSA Transformer, esto no implica necesariamente que la información de este tipo sea irrelevante. Más bien, su impacto puede depender de la calidad y fiabilidad de los sentimientos inferidos, así como de la influencia real que estos tengan en los patrones de movilidad de los usuarios.

3.5.3. Impacto de las longitudes de entrada históricas

A continuación interesa conocer cuánta información histórica debe considerarse para que la red DL logre el rendimiento deseado. Se identifica el tiempo de visita de cada registro histórico con referencia a la predicción actual y alteramos la longitud de la secuencia de entrada controlando el número de días D a considerar en el pasado. Se puede observar una tendencia general decreciente cuando aumenta el número de días históricos considerados, lo que significa que incluir secuencias más largas, y por lo tanto más información, no necesariamente conduce a un mejor rendimiento del modelo. Además, observamos dos picos en la tendencia de Acc@1 correspondientes a visitas a puntos de estancia en los 7 y 14 días anteriores. Los resultados de las pruebas de Mann–Whitney U muestran que el Acc@1 obtenido de los últimos 7 días no es significativamente diferente de considerar 1 día en el pasado, pero sí es significativamente diferente al obtenido de todas las demás longitudes de entrada. Por lo tanto, concluimos que el modelo propuesto, así como el MHSA Transformer logran el mejor rendimiento al considerar la movilidad realizada en los últimos 7 días. Además,

los picos de rendimiento sugieren que las huellas de movilidad de una o dos semanas atrás llevan información adicional que es beneficiosa para predecir la visita al lugar del día actual.

Conclusiones

En este trabajo, se analiza el desempeño de diferentes modelos para la predicción de la próxima ubicación de usuarios en el conjunto de datos Geolife. Se comparó el modelo basado en MHSA Transformer con nuestro modelo propuesto, el cual incorpora el contexto de sentimientos inferidos mediante un modelo de lenguaje.

Los resultados muestran que ambos modelos basados en aprendizaje profundo superan significativamente al método de primer orden 1-MMC en casi todos los indicadores de desempeño, resaltando la importancia de modelar dependencias temporales a largo plazo en tareas de predicción de movilidad. Sin embargo, la inclusión del contexto de sentimientos en nuestro modelo no resultó en mejoras estadísticamente significativas en comparación con el MHSA Transformer, según la prueba U de Mann-Whitney.

Se identificaron dos posibles razones para esta falta de significancia: (i) la inferencia del contexto de sentimientos basada en información limitada (hora, lugar y día de la semana) pudo no haber capturado con precisión el estado emocional real de los usuarios, introduciendo ruido en los datos, y (ii) la influencia de los sentimientos en la movilidad puede ser marginal cuando los patrones de movimiento están fuertemente determinados por factores estructurales, como el trabajo o la rutina diaria.

Por otro lado, el análisis del impacto de la longitud de la secuencia histórica reveló que la mejor configuración para el modelo propuesto se obtiene al considerar las secuencias de los últimos 7 días. Además, los picos en el desempeño observados en los días 7 y 14 sugieren que los patrones de movilidad pueden tener periodicidades semanales relevantes para la predicción.

En conclusión, si bien el contexto de sentimientos no proporcionó mejoras significativas en la predicción de movilidad en este estudio, su potencial utilidad podría depender de la calidad de la inferencia emocional y del tipo de movilidad analizada.

Recomendaciones

A partir de los resultados, proponemos las siguientes recomendaciones para futuras investigaciones y aplicaciones:

- **Evaluación rigurosa del LLM:** Si se emplea un LLM para la generación de etiquetas de sentimientos, es fundamental realizar una evaluación exhaustiva de su desempeño. Esto incluye verificar la coherencia y estabilidad de las inferencias, comparar los resultados con anotaciones humanas y analizar posibles sesgos o errores sistemáticos que puedan afectar la calidad de los datos.
- **Uso de técnicas clásicas de recuperación de información:** En lugar de utilizar un LLM, una alternativa viable es aplicar enfoques tradicionales de recuperación de información, como la búsqueda de patrones en bases de datos preexistentes de estados de la persona, el uso de diccionarios léxicos de sentimientos o modelos supervisados de clasificación de texto. Estos métodos pueden ofrecer mayor control y transparencia en la inferencia de sentimientos, reduciendo la variabilidad introducida por la generación de texto de los LLM.
- **Mejora en la inferencia del contexto de sentimientos:** Para capturar mejor el estado emocional de los usuarios, se recomienda el uso de fuentes de datos adicionales, como interacciones en redes sociales, encuestas directas o datos biométricos, que puedan proporcionar una inferencia más precisa y menos ruidosa de los sentimientos.
- **Análisis del impacto del contexto emocional en diferentes escenarios:** Se sugiere evaluar el impacto del contexto de sentimientos en distintos conjuntos de datos de movilidad, especialmente en entornos donde las decisiones de movimiento pueden estar más influenciadas por factores emocionales.
- **Validación en contextos reales:** Para evaluar la aplicabilidad práctica de los modelos, se recomienda realizar pruebas en escenarios reales con usuarios que proporcionen retroalimentación sobre la precisión de las predicciones y la relevancia del contexto emocional en su movilidad cotidiana.

En general, estos puntos pueden contribuir a mejorar la precisión de los modelos de predicción de movilidad y a profundizar en la comprensión del papel de los factores de sentimientos en los patrones de movimiento de los usuarios.

Bibliografía

- Ashbrook, D., & Starner, T. (2002). Learning Significant Locations and Predicting User Movement with GPS. *International Symposium on Wearable Computers (ISWC)*, 101-108 (vid. págs. 4, 18).
- Chen, M., Liu, Y., & Yu, X. (2014). NLPMM: A Next Location Predictor with Markov Modeling. *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 186-197 (vid. págs. 4).
- Cherchi, E., Cirillo, C., & de Dios Ortúzar, J. (2017). Modelling correlation patterns in mode choice models estimated on multiday travel data. *Transportation Research Part A: Policy and Practice*, 96, 146-153 (vid. págs. 2).
- Dharmowijoyo, D. B. E., Susilo, Y. O., & Karlström, A. (2016). Day-to-day variability in travellers' activity-travel patterns in the Jakarta metropolitan area. *Transportation*, 43(4), 601-621 (vid. págs. 2).
- Dharmowijoyo, D. B., Susilo, Y. O., & Karlström, A. (2017). Analysing the complexity of day-to-day individual activity-travel patterns using a multidimensional sequence alignment model: A case study in the Bandung Metropolitan Area, Indonesia. *Journal of Transport Geography*, 64, 1-12 (vid. págs. 6).
- Feng, J., Li, Y., Zhang, C., Sun, F., Meng, F., Guo, A., & Jin, D. (2018). DeepMove: Predicting Human Mobility with Attentional Recurrent Networks. *International World Wide Web Conference (WWW)*, 1459-1468 (vid. págs. 2, 5).
- Gambs, S., Killijian, M.-O., & del Prado Cortez, M. N. (2012). Next Place Prediction Using Mobility Markov Chains. *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*, 3:1-3:6 (vid. págs. 4, 18).
- Geocode Maps. (2025). Geocoding API. <https://geocode.maps.co> (vid. págs. 16).
- Google AI for Developers. (2024). Documentación de la API de Gemini. <https://ai.google.dev/gemini-api/docs> (vid. págs. 15).
- Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyam-sundar, P., Steffen, W., Glaser, G., Kanie, N., & Noble, I. (2013). Sustainable development goals for people and planet. *Nature*, 495(7441), 305-307 (vid. págs. 1).

- Hariharan, R., & Toyama, K. (2004). Project Lachesis: Parsing and Modeling Location Histories. *Geographic Information Science*, 106-124. https://doi.org/10.1007/978-3-540-30231-5_8 (vid. pág. 8).
- Hernández, E. A. (2023). *Simulación computacional para la dinámica de enfermedades transmitidas por vectores* [Tesis de Diploma]. Facultad Matemática y Computación, Universidad de La Habana. (Vid. pág. 10).
- Hong, Y. [Ye], Martin, H., & Raubal, M. (2022). How do you go where?: improving next location prediction by learning travel mode information using transformers. *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*. <https://doi.org/10.1145/3557915.3560996> (vid. pág. 2).
- Hong, Y. [Ye], Martin, H., Xin, Y., Bucher, D., Reck, D. J., Axhausen, K. W., & Raubal, M. (2022). Conserved quantities in human mobility. From locations to trips. <https://doi.org/10.3929/ethz-b-000551831> (vid. pág. 6).
- Hong, Y. [Ye], Zhang, Y., Schindler, K., & Raubal, M. (2023). Context-aware multi-head self-attentional neural network model for next location prediction. *Transportation Research Part C: Emerging Technologies*, 156, 104315. <https://doi.org/10.1016/j.trc.2023.104315> (vid. págs. 7, 10, 18).
- Hong, Y. [Yihong], Xin, Y., Martin, H., Bucher, D., & Raubal, M. (2021). A Clustering-Based Framework for Individual Travel Behaviour Change Detection. *11th International Conference on Geographic Information Science - Part II*, 208, 4:1-4:15. <https://doi.org/10.4230/LIPICs.GIScience.2021.II.4> (vid. pág. 15).
- Huang, H., Bucher, D., Kissling, J., Weibel, R., & Raubal, M. (2019). Multimodal Route Planning With Public Transport and Carpooling. *IEEE Transactions on Intelligent Transportation Systems*, 20, 3513-3525 (vid. pág. 1).
- Huang, Q. (2017). Mining online footprints to predict user's next location. *International Journal of Geographical Information Science*, 31(3), 523-541 (vid. pág. 4).
- Kraemer, M., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D., Covid, O., Hanage, W., Brownstein, J., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O., & Scarpino, S. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490), 493-497. <https://doi.org/10.1126/science.abb4218> (vid. pág. 1).
- Laha, A., & Putatunda, S. (2018). Real time location prediction with taxi-GPS data streams. *Transportation Research Part C: Emerging Technologies*, 92, 298-322. <https://doi.org/10.1016/j.trc.2018.05.005> (vid. pág. 6).
- Lau, B. P., Marakkalage, S. H., Zhou, Y., Hassan, N. U., Yuen, C., Zhang, M., & Tan, U.-X. (2019). A survey of data fusion in smart city applications. *Information Fusion*, 52, 357-374. <https://doi.org/10.1016/j.inffus.2019.05.004> (vid. pág. 6).
- Li, F., Gui, Z., Zhang, Z., Peng, D., Tian, S., Yuan, K., Sun, Y., Wu, H., Gong, J., & Lei, Y. (2020). A hierarchical temporal attention-based LSTM encoder-decoder

- model for individual mobility prediction. *Neurocomputing*, 403, 153-166 (vid. págs. 2, 5, 10).
- Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., & Ma, W.-Y. (2008). Mining user similarity based on location history. *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, 1-10. <https://doi.org/10.1145/1463434.1463477> (vid. pág. 8).
- Liao, J., Liu, T., Liu, M., Wang, J., Wang, Y., & Sun, H. (2018). Multi-Context Integrated Deep Neural Network Model for Next Location Prediction. *IEEE Access*, 6, 21980-21990. <https://doi.org/10.1109/ACCESS.2018.2827422> (vid. pág. 6).
- Liu, Q., Wu, S., Wang, L., & Tan, T. (2016). Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts. *AAAI Conference on Artificial Intelligence (AAAI)*, 194-200 (vid. pág. 5).
- Luca, M., Barlacchi, G., Lepri, B., & Pappalardo, L. (2021). A Survey on Deep Learning for Human Mobility. *ACM Computing Surveys (CSUR)*, 55(1), 7:1-7:44 (vid. págs. 1, 4).
- Ma, Z., & Zhang, P. (2022). Individual mobility prediction review: Data, problem, method and application. *Multimodal Transportation*, 1(1), 100002 (vid. pág. 1).
- Mariante, G. L., Ma, T.-Y., & Van Acker, V. (2018). Modeling discretionary activity location choice using detour factors and sampling of alternatives for mixed logit models. *Journal of Transport Geography*, 72, 151-165 (vid. pág. 6).
- Martin, H., Hong, Y. [Ye], Wiedemann, N., Bucher, D., & Raubal, M. (2023). Trackintel: An open-source python library for human mobility analysis. *Computers, Environment and Urban Systems*, 101, 101938. <https://doi.org/10.1016/j.compenvurbsys.2023.101938> (vid. págs. 8, 14).
- Martin, H., Reck, D. J., & Raubal, M. (2021). Using Information and Communication Technologies to facilitate mobility behaviour change and enable Mobility as a Service. *GI_Forum Journal for Geographic Information Science*, 9(1), 187-193 (vid. pág. 1).
- Neutens, T., Witlox, F., Van De Weghe, N., & De Maeyer, P. H. (2007). Space-time opportunities for multiple agents: a constraint-based approach. *International Journal of Geographical Information Science*, 21(10), 1061-1076 (vid. pág. 6).
- Pirozmand, P., Wu, G., Jedari, B., & Xia, F. (2014). Human mobility in opportunistic networks: Characteristics, models and prediction methods. *Journal of Network and Computer Applications*, 42, 45-58 (vid. pág. 4).
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training (vid. pág. 12).
- Reck, D. J., Martin, H., & Axhausen, K. W. (2022). Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility. *Trans-*

- portation Research Part D: Transport and Environment, 102, 103134 (vid. pág. 1).
- Schönfelder, S., & Axhausen, K. W. (2016). *Urban rhythms and travel behaviour: spatial and temporal phenomena of daily travel*. Routledge. (Vid. pág. 6).
- Scott, D. M., & He, S. Y. (2012). Modeling constrained destination choice for shopping: a GIS-based, time-geographic approach. *Journal of Transport Geography*, 23, 60-71 (vid. pág. 6).
- Solomon, A., Livne, A., Katz, G., Shapira, B., & Rokach, L. (2021). Analyzing movement predictability using human attributes and behavioral patterns. *Computers, Environment and Urban Systems*, 87, 101596 (vid. págs. 4, 5).
- Song, C., Qu, Z., Blumm, N., & Barabási, A.-L. (2010). Limits of predictability in human mobility. *Science*, 327(5968), 1018-1021 (vid. pág. 2).
- Sun, G., Qi, H., Shen, Y., & Yin, B. (2022). TCSA-Net: A Temporal-Context-Based Self-Attention Network for Next Location Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 23(11), 20735-20745. <https://doi.org/10.1109/TITS.2022.3181339> (vid. pág. 6).
- Sun, L., Axhausen, K. W., Lee, D.-H., & Huang, X. (2013). Understanding metropolitan patterns of daily encounters. *Proceedings of the National Academy of Sciences*, 110(34), 13774-13779 (vid. pág. 2).
- Susilo, Y. O., & Axhausen, K. W. (2014). Repetitions in individual daily activity–travel–location patterns: a study using the Herfindahl–Hirschman Index. *Transportation*, 41(5), 995-1011 (vid. pág. 6).
- Tedjopurnomo, D. A., Bao, Z., Zheng, B., Choudhury, F., & Qin, A. (2020). A Survey on Modern Deep Neural Network for Traffic Prediction: Trends, Methods and Challenges. *IEEE Transactions on Knowledge and Data Engineering*, 34(4), 1544-1561. <https://doi.org/10.1109/TKDE.2020.3001195> (vid. pág. 6).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, 5998-6008 (vid. págs. 1, 5, 12, 18).
- Wang, D., Liu, K., Xiong, H., & Fu, Y. (2022). Online POI Recommendation: Learning Dynamic Geo-Human Interactions in Streams. *arXiv preprint arXiv:2201.10983* (vid. pág. 4).
- Wang, D., Wang, P., Liu, K., Zhou, Y., Hughes, C. E., & Fu, Y. (2021). Reinforced Imitative Graph Representation Learning for Mobile User Profiling: An Adversarial Training Perspective. *AAAI Conference on Artificial Intelligence (AAAI)*, 4410-4417 (vid. pág. 4).
- Wang, X., Fang, M., Zeng, Z., & Cheng, T. (2024). Where Would I Go Next? Large Language Models as Human Mobility Predictors. <https://arxiv.org/abs/2308.15197> (vid. pág. 7).

- Xu, Y. [Yang], Zou, D., Park, S., Li, Q., Zhou, S., & Li, X. (2022). Understanding the movement predictability of international travelers using a nationwide mobile phone dataset collected in South Korea. *Computers, Environment and Urban Systems*, 92, 101753 (vid. págs. 4, 11).
- Xu, Y. [Yanyan], Çolak, S., Kara, E. C., Moura, S. J., & González, M. C. (2018). Planning for electric vehicle needs by coupling charging profiles with urban mobility. *Nature Energy*, 3(6), 484-493 (vid. pág. 1).
- Xue, H., Salim, F. D., Ren, Y., & Oliver, N. (2021). MobTCCast: Leveraging Auxiliary Trajectory Forecasting for Human Mobility Prediction. *Conference on Neural Information Processing Systems (NeurIPS)*, 30380-30391 (vid. págs. 4, 5).
- Yoon, S. Y., Deutsch, K., Chen, Y., & Goulias, K. G. (2012). Feasibility of using time-space prism to represent available opportunities and choice sets for destination choice models in the context of dynamic urban environments. *Transportation*, 39(4), 807-823 (vid. pág. 6).
- Zheng, X., Han, J., & Sun, A. (2018). A Survey of Location Prediction on Twitter. *IEEE Transactions on Knowledge and Data Engineering*, 30(9), 1652-1671. <https://doi.org/10.1109/TKDE.2018.2807840> (vid. pág. 6).
- Zheng, Y., Fu, H., Xie, X., Ma, W.-Y., & Li, Q. (2011, julio). *Geolife GPS trajectory dataset - User Guide* (Geolife GPS trajectories 1.1). <https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/> (vid. pág. 14).
- Zheng, Y., Xie, X., Ma, W.-Y., et al. (2010). GeoLife: A Collaborative Social Networking Service among User, Location and Trajectory. *IEEE Data(base) Engineering Bulletin*, 33(2), 32-39 (vid. pág. 14).