

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

Enero 2025

github.com/YoanRene/My-Thesis

Dedicatoria

Agradecimientos

Agradecimientos

Opinión del tutor

Opiniones de los tutores

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 propone 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
Conclusiones	15
Recomendaciones	16
Bibliografía	17

Índice de figuras

Ejemplos de código

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 *ride-sharing* (H. Huang et al. 2019), dependen de la capacidad de ofrecer servicios personalizados adaptados al contexto del viaje y las características individuales (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 este problema 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 (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 un individuo 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 estructura del documento está dada por 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 discuten sobre la implementación del modelo, así como de su evaluación mediante la experimentación. Terminado este último capítulo se da paso a las Conclusiones, la cual resume los aspectos principales, dando respuesta a los objetivos según los resultados.

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 nos centramos 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 Gambs 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 2017) aumentaron aún más el rendimiento de la predicción. Sin embargo, los

¹LBSN: Location-Based Social Networks

²POI: Point of Interest

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 se 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 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

³DL: Deep Learning

⁴RNN: Recurrent Neural Networks

⁵LSTM: Large Short-Term Memory

⁶ST: Spatial-Temporal

⁷MHSA: Multi-Head Self-Attentional

que se centran 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. 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 (Tadjopurnomo 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; 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 problema

⁸LLM: Large Language Models

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 en que se tengan 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 Hong, Y. Zhang et al. 2023 por sus excelentes resultados en esta problemática y su capacidad de capturar dependencias.

Capítulo 2

Propuesta

2.1. Definición del Problema

Introducimos un conjunto de términos y nociones utilizados en el resto del documento y formulamos el problema de predicción de la siguiente ubicación. Los datos de movilidad se recopilan 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, 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. Usamos S_k 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. Definimos \mathcal{O}_i 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 añadimos también información de contexto que representa un sentimiento del usuario en dicho punto, llamémosle 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, definimos el problema de predicción de la siguiente ubicación.

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, construimos la secuencia histórica basándonos en 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 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 es 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. Además, la información relacionada con el usuario ayuda a descubrir las secuencias

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

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.

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:

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

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 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, $e_{u_i} = h_u(u_i; \mathbf{W}_u)$. 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,

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.

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

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

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} :

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

⁴GPT: Generative Pre-trained Transformer

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

⁶FC: Fully Connected

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

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.

Capítulo 3

Detalles de Implementación y Experimentos

Conclusiones

Conclusiones

Recomendaciones

Recomendaciones

Bibliografía

- Ashbrook, D., & Starner, T. (2002). Learning Significant Locations and Predicting User Movement with GPS. *International Symposium on Wearable Computers (ISWC)*, 101-108 (vid. pág. 4).
- 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. pág. 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. pág. 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. pág. 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. pág. 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. pág. 4).
- Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyamsundar, P., Steffen, W., Glaser, G., Kanie, N., & Noble, I. (2013). Sustainable development goals for people and planet. *Nature*, 495(7441), 305-307 (vid. pág. 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., 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., 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., 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).
- 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., 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. pág. 8).
- 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. *Transportation 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. 5).
- 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. págs. 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. págs. 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. págs. 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. págs. 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. págs. 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, 11, 12).
- 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. págs. 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. págs. 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. págs. 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. págs. 1).

- Xue, H., Salim, F. D., Ren, Y., & Oliver, N. (2021). MobTCast: 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).