

Uma Arquitetura Baseada em Redes Neurais Recorrentes para Predição de Trajetórias Veiculares em Ambientes Urbanos

Breno Krohling¹, Giovanni Comarela¹, Vinícius F. S. Mota¹

¹Programa de Pós-Graduação em Informática (PPGI)
Universidade Federal do Espírito Santo (UFES) – Vitória, ES – Brasil

breno.krohling@edu.ufes.br, {gc,vinicius.mota}@inf.ufes.br

Abstract. *Predicting vehicular trajectories plays a crucial role in optimizing traffic management and enhancing inter-vehicle communication. In this context, this study proposes an architecture based on stacked Long Short-Term Memory (LSTM), where the output of one LSTM serves as the input to the subsequent layer. The multiple LSTM layers enable the proposed architecture to accurately predict vehicle positions in both near and distant future. Two performance metrics have been introduced: the absolute error measured by the distance between the real and predicted positions and the model's accuracy when predicting regions of interest. To evaluate the model, trajectory datasets from taxis in Porto, Portugal, and San Francisco, USA, were employed. The results show the architecture's robust performance in both prediction scenarios, achieving an accuracy of 83% in predicting areas of interest in the near future.*

Resumo. *A predição de trajetórias veiculares permite otimizar o gerenciamento de tráfego e facilitar a comunicação entre veículos. Neste sentido, este trabalho propõe uma arquitetura baseada em Long Short-Term Memory empilhadas (stacked LSTM), na qual a saída de uma LSTM atua como entrada na camada subsequente. As múltiplas camadas LSTM permitem que a arquitetura proposta possa prever a posição de veículos em um futuro próximo e distante. Foram propostas duas métricas de desempenho: o erro absoluto da distância entre a posição real e predita e a acurácia do modelo para prever regiões de interesse. Para avaliar o modelo, foram utilizados datasets de trajetórias de táxis das cidades do Porto, Portugal, e São Francisco, EUA. Os resultados demonstram bom desempenho da arquitetura para ambos os cenários de predição, alcançando uma acurácia de 83% para predição de áreas de interesse no futuro próximo.*

1. Introdução

Com o avanço do uso de tecnologias que permitem coletar e transmitir informações de usuários em tempo real, tais como posição geográfica e sensores, diversos serviços para cidades inteligentes têm sido propostos [Mehmood et al. 2017]. As informações coletadas podem ajudar a lidar com problemas de planejamento urbano, gerenciamento de tráfego [Khansari et al. 2014, Maggi and Vallino 2016] e melhorar a comunicação entre veículos na chamada *Internet of Vehicles* (IoV) [Sharma and Kaushik 2019]. Para isto, é fundamental compreender a dinâmica de movimentação de veículos em grandes cidades.

A dinâmica do tráfego de veículos pode permitir desde a detecção de eventos anômalos em tempo real [Thomé et al. 2020] a predição de trajetórias de veículos para

melhor gerenciamento do tráfego [King et al. 2021]. Por sua vez, a predição de trajetórias permite melhorar a decisão de rotas para motoristas [Zhang et al. 2016], recomendar serviços baseados no destino do veículo [Kong et al. 2017], entre outros.

A predição de trajetória consiste em prever a posição futura de um veículo utilizando informações de suas posições passadas. [King et al. 2021] classificam as propostas de solução para o problema de predição de trajetória em *short-term* e *long-term*. Modelos *short-term* têm por objetivo prever o futuro imediato, ou seja, a posição do objeto na próxima unidade de tempo, enquanto os modelos *long-term* estão interessados na predição do futuro distante, h unidades de tempo a frente da posição atual. Soluções com modelagem *short-term* são mais comuns, por apresentarem um grau de dificuldade inferior a modelos *long-term*. Isso deve-se à complexidade e possibilidades de movimentação encontradas em centros urbanos, tornando as trajetórias menos previsíveis com o avanço temporal.

Dentre as diversas técnicas utilizadas para predição de trajetórias, o uso de redes neurais recorrentes (RNN) tem-se destacado devido às características intrínsecas ao modelo, como a capacidade de armazenamento, que permite a retenção de informações passadas internamente. Essa característica faz com que modelos baseados em RNN's obtenham desempenho superior em relação a outros métodos quando os dados utilizados possuem características temporais ou sequenciais. A RNN *Long Short-Term Memory* (LSTM) foi usada em problemas similares de predição, como predição de pedestres [Zhang et al. 2019] e ciclistas [Saleh et al. 2018].

O problema da predição de trajetória utilizando informações de localização, no formato (latitude-longitude) traz diversos desafios, tais como: i) imprecisão de posição são recorrentes e aceitáveis em GPS comuns; ii) a distância percorrida nas cidades causam apenas pequenas variações nas casas decimais; e iii) a trajetória passada pode ou não indicar as próximas posições. Por exemplo, considere que a última posição (latitude, longitude) conhecida de um veículo seja (-22.905300,-43.133600) em uma região cujo limite de velocidade seja 20km/h (ou 5,55 m/s). Como cada grau de uma latitude ou longitude equivale a aproximadamente 111,11 km, cada minuto da trajetória causaria uma mudança de no máximo a terceira casa decimal da latitude ou longitude, fazendo com que predições com precisões decimais sejam relevantes. Em resumo, predizer uma posição geográfica baseado nas posições passadas ainda é um desafio em aberto.

Neste sentido, este trabalho propõe uma arquitetura baseada em redes neurais recorrentes empilhadas para tratar o problema de predição *short-term* e *long-term* de uma trajetória. A arquitetura proposta utiliza uma sequência de tuplas (*tempo, latitude, longitude*) como entrada em uma LSTM. Contudo, a saída da LSTM é utilizada como entrada em outra LSTM, e assim consecutivamente até um limite de camadas LSTM, definido por hiperparâmetro. Considerando o problema de baixa precisão em leituras de sistemas GPS, é proposto a utilização de duas métricas para avaliar predição de trajetória pela arquitetura proposta: i) distância de Haversine entre a posição real e a posição predita; e ii) região de interesse, que divide área de estudo em uma grade retangular (*grid*) e as posições reais e preditas são mapeadas para uma posição do grid. Deste modo, a métrica visa verificar se o modelo consegue acertar (*hit*) uma região de interesse futura de um veículo.

A arquitetura proposta foi avaliada utilizando datasets com informações de tempo,

latitude e longitude em trajetos percorridos por taxistas na cidade do Porto, em Portugal e em São Francisco, nos Estados Unidos, que representam cidades com topologias e arquiteturas urbanas distintas. A arquitetura foi utilizada para predição *short-term* de trajetórias, representando 10 a 15 segundos a frente, e *long-term*, representando de 100 a 150 segundos a frente. Os resultados obtidos demonstram que a arquitetura proposta pode ser utilizada principalmente em aplicações que se deseja predizer regiões de interesse futura de veículos. Por exemplo, para comunicação entre veículos.

O restante do artigo está organizado da seguinte maneira: A Seção 2 apresenta a revisão de literatura sobre predições de trajetórias. A Seção 3 apresenta a definição do problema e a arquitetura proposta para o problema de predição de posições futuras em trajetórias. Os experimentos e resultados obtidos são apresentados na Seção 4. Por fim, a Seção 5 apresenta as considerações finais.

2. Revisão de Literatura

Uso de dados de localização para predição de trajetórias tem sido abordado através do uso de técnicas de análise de padrões, cadeias de Markov e redes neurais artificiais [Koolwal and Mohbey 2020]. O objetivo principal de tais trabalhos é prever a posição futura de trajetórias de pedestres, ciclistas ou veículos.

Na modelagem por meio de cadeias de Markov são construídas matrizes de transição de estado baseada no histórico de trajetória de cada usuário [Qiao et al. 2017]. Em [Qiao et al. 2017], os dados de trajetórias de coordenadas geográficas são transformados para trajetórias de regiões com uso de diagramas de Voronoy. As trajetórias são então agrupadas por similaridade e posteriormente um modelo de Markov é desenvolvido para cada grupo. Já [Karatzoglou et al. 2018] utilizaram cadeias de Markov multidimensionais que além da localização, utilizam informações como dia e horário do percurso. [Rathore et al. 2019] apresentam um *framework* de duas etapas. Inicialmente todas as trajetórias são agrupadas em *clusters* de acordo com uma função de semelhança proposta pelos autores. Após isso, modelos de cadeia de Markov são construídos para cada *cluster*, que são posteriormente utilizados para realizar a predição de trajetórias.

O uso de redes neurais artificiais vêm sendo relevante para a tarefa de predição de trajetórias, uma vez que tais modelos apresentam bons resultados em extrair padrões de dados complexos. Em [Lv et al. 2018], os autores utilizam redes convolucionais para realizar a tarefa de predição. Os autores propõem representar trajetórias como imagens, visando capturar informações bidimensionais inerentes ao problema/modelagem. Os dados transformados foram utilizados como entradas do modelo convolucional. No entanto, os modelos mais utilizados são baseados em redes neurais recorrentes ([Ip et al. 2021], [Kim et al. 2017], [Feng et al. 2020]), conhecidas por apresentar bons resultados em tarefas com dados de natureza temporal e sequencial.

Especialmente, redes LSTM vêm ganhando popularidade por considerarem as características temporais das trajetórias. [Kim et al. 2017] propuseram uma arquitetura para prever a posição futura de carros em rodovias. O trabalho utiliza dados de localização e velocidade dos veículos e tem como saída do modelo um conjunto de possíveis posições futuras e suas respectivas probabilidades. [Park et al. 2018] utilizaram o mesmo tipo de dado, porém com uma arquitetura do tipo *encoder-decoder* para realizar a tarefa de predição de veículos ao redor de um veículo *target*. Esse tipo de abordagem é utilizada

para veículos autônomos e demonstrou bons resultados para previsões de futuro próximo. [Ip et al. 2021] utilizaram LSTM para previsão de posição futura. A modelagem proposta pelos autores realiza uma transformação nos dados, de modo que as trajetórias representadas originalmente por coordenadas GPS, são mapeadas em células de um *grid*, que são posteriormente utilizadas para o treinamento do modelo proposto. A proposta dos autores visa identificar apenas a célula futura em que um veículo estará. Os autores demonstram que o modelo proposto obtém uma acurácia de até 89%. No entanto, os autores mapearam a cidade de Porto, em Portugal, em um *grid* fixo de 4x4 células, o que representa uma área significativa por célula.

Este trabalho propõe a utilização de redes LSTM empilhadas que usam apenas dados de geolocalizações ordenadas no tempo para treinamento. Desta forma, espera-se que o modelo resultante seja capaz de generalizar padrões de trajetórias de diversos motoristas por toda extensão de uma cidade. Adicionalmente, este trabalho avalia a relação entre a quantidade de camadas LSTM e o desempenho do modelo. Além disso, a análise por *grid* proposta leva em consideração a extensão de toda a cidade e realiza um estudo da influência do tamanho do *grid* nos resultados. Trabalhos semelhantes, como o apresentado por [Ip et al. 2021], utiliza apenas um *grid* de tamanho fixo, que para grandes cidades representam uma área de região de interesse com alta granularidade. Outro ponto a ser destacado é que a arquitetura proposta neste trabalho realiza o mapeamento de coordenadas geográficas para posições no *grid* apenas na etapa de pós-processamento. Isto permite que a arquitetura proposta mantenha as informações originais das trajetórias. Deste modo, o modelo gerado pela arquitetura de LSTM empilhadas permite prever tanto as próximas regiões de interesse quanto as posições futuras da trajetória, estimando o erro pela distância entre a posição real e a posição predita.

3. Arquitetura proposta: LSTM Empilhada

A seguir, são apresentadas as definições do problema e o método utilizado para realizar a previsão de trajetórias e os métodos de avaliação.

3.1. Definição do Problema

Como mencionado anteriormente, problemas de previsão de trajetória podem ser classificados em duas categorias, previsões *short-term* e previsões *long-term*. Neste trabalho, busca-se aplicar previsões *short-term* e *long-term* para analisar a eficiência da previsão ao longo do tempo. Para isso serão utilizadas bases de dados reais de trajetórias de táxis coletadas na cidade do Porto (Portugal) e na São Francisco (Estados Unidos). A seguir são apresentadas definições para formalizar o problema de previsão de trajetórias.

Posição: Seja $x_i = (lat_i, long_i)$ uma posição no sistema de posicionamento global (GPS) no formato Latitude e Longitude (WGS84) de um veículo no tempo i .

Trajeto: Uma trajetória $T = \{x_1, x_2, x_3, \dots, x_t\}$ representa um trajeto completo percorrido por um veículo durante t unidades de tempo, onde x_1 representa a posição inicial e x_t a posição final do veículo.

Sub-trajeto: Dada uma trajetória T de tamanho t , a sub-trajeto, $\tau_{j,w}$, de T é dada pelas w posições consecutivas em T a partir da posição j . Formalmente, tem-se que $\tau_{j,w} = \{x_j, x_{j+1}, \dots, x_{j+w-1}\}$, para $w \leq t$ e $1 \leq j \leq t - w + 1$. Como consequência, T possui um total de $t - w + 1$ sub-trajetos de tamanho w .

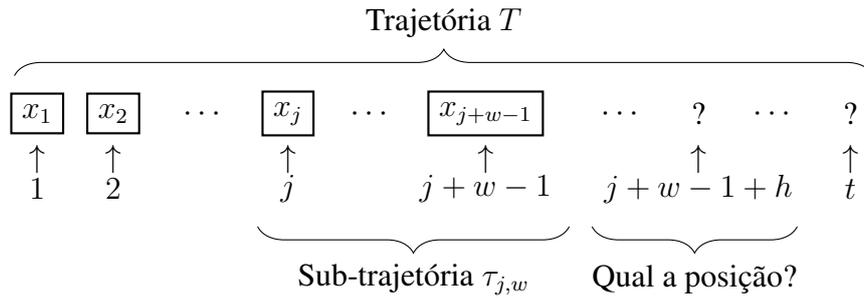


Figura 1. Ilustração do problema de Predição de Trajetória abordado. Dada uma sub-trajetória $\tau_{j,w}$, no tempo $j + w - 1$, é possível prever a posição do veículo h unidades de tempo à frente?

A partir das definições acima, os problemas tratados por este artigo podem ser definidos como a seguir.

Predição *short-term*: dada uma sub-trajetória $\tau_{j,w}$, o objetivo é prever a posição do veículo na próxima unidade de tempo. Na Figura 1, essa versão do problema consiste em tomar $h = 1$.

Predição *long-term*: dada uma sub-trajetória $\tau_{j,w}$, o objetivo é prever a posição do veículo em tempos que reflitam médio e/ou longo prazo. Na Figura 1, essa versão do problema consiste em tomar valores de h significativamente maiores que 1.

3.2. LSTM para predição de trajetórias

RNN's são modelos de aprendizado comumente utilizados para lidar com dados sequenciais e temporais. Já as redes do tipo LSTM são uma variação de RNN's que possuem um mecanismo de memória adaptado em sua arquitetura, possibilitando o armazenamento de informações de estados passados por um longo período [Hochreiter and Schmidhuber 1997]. Devido às características de dados de trajetórias, sequenciais e temporais, o uso de redes LSTM torna-se propício ao problema proposto. Além disso, tais modelos permitem como entrada informações de dimensões variadas [Rumelhart et al. 1986], característica presente nos dados de domínio do problema, uma vez que trajetórias possuem tamanhos variados e cada posição representa uma tupla.

Nesse sentido, é proposto o uso de redes LSTM para a tarefa de predições *short-term* e *long-term*. Dado como entrada uma sub-trajetória $\tau_{j,w}$ percorrida por um veículo qualquer, o modelo deve prever a sua posição futura $\hat{x}_{j+w-1+h}$. Visando extrair informações temporais complexas de padrões de trajetórias, esse trabalho irá utilizar o conceito de LSTM empilhadas, que consiste em diversas células do modelo sequenciais, onde o resultado de saída de uma célula é utilizado como entrada da subsequente. A escolha do uso de LSTM empilhadas em detrimento a um modelo de LSTM padrão, é resultado de experimentos preliminares que indicam a superioridade do uso do modelo empilhado. A arquitetura do modelo é apresentada na Figura 2.

A saída $\hat{x}_{j+w-1+h}$ do modelo é uma localização, *i.e.*, coordenadas GPS no formato latitude e longitude. Devido à natureza desse tipo de dado, avaliações de desempenho por métricas padrão, como o erro quadrado médio (RMSE), por exemplo, tornam-se inviáveis. De fato, alterações na terceira casa decimal das coordenadas causam diferenças

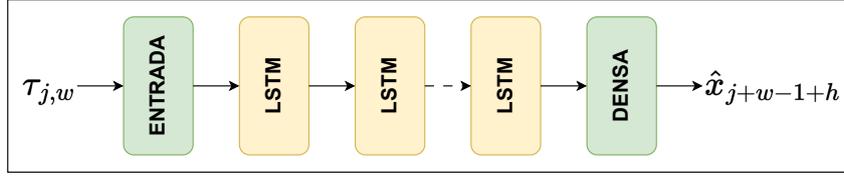


Figura 2. Arquitetura base de uma LSTM com camadas empilhadas. Neste modelo, a saída de uma camada é utilizada como entrada da camada subsequente.

significativas da distância real, em metros. Para contornar esse problema e obter métricas realistas, é realizado um pós-processamento da saída do modelo, mapeando os dados de saída para células de um *grid*. O mapeamento é realizado para um *grid* M de dimensões a por m que cobrem a região da cidade/localidade analisada, onde cada célula do *grid* representa uma região de interesse. Dado uma posição x_i qualquer dentro dos limites do *grid*, sua posição $M_{r,c}(x_i)$ pode ser mapeada conforme as Equações 1, 2 e 3.

$$M_{r,c}(x_i) = a_r(x_i), m_c(x_i) \quad (1)$$

$$a_r(x_i) = \left\lfloor \frac{a(lat_{x_i} - lat_{min})}{lat_{max} - lat_{min}} \right\rfloor, r = 0, 1, \dots, a - 1 \quad (2)$$

$$m_c(x_i) = \left\lfloor \frac{m(long_{x_i} - long_{min})}{long_{max} - long_{min}} \right\rfloor, c = 0, 1, \dots, m - 1 \quad (3)$$

onde lat_{min} , lat_{max} , $long_{min}$, $long_{max}$ são as posições geográficas referentes aos vértices do *grid* M .

Dessa forma é possível analisar o desempenho do modelo por *hits* de áreas de interesse. Nesse caso, as dimensões a e m do *grid* são responsáveis por determinar a granularidade dos resultados. Menor a granularidade, melhor a precisão dos resultados e maior a dificuldade do problema. Esse *trade-off* pode ser analisado segundo os objetivos esperados ao prever as posições futuras do veículo. Na Seção 4, são apresentadas análises referentes à granularidade do *grid*.

Na Figura 3, pode-se observar uma exemplificação do mapeamento de posições para uma célula do *grid* através da função $M_{r,c}$. Quando o mapeamento das posições $x_{j+w-1+h}$ e $\hat{x}_{j+w-1+h}$ é realizado para a mesma célula no *grid*, ocorre um *hit*, como demonstrado para o tempo $t = w + 1$. Do contrário, ocorre um erro, como no tempo $t = w + 2$.

3.3. Métodos de avaliação

Tradicionalmente, modelos de previsão utilizam métricas como RMSE ou Coeficiente de determinação (R^2) para quantificar o erro entre o valor real e um valor predito por um modelo. No entanto, como dito anteriormente, a natureza dos dados do problema acarreta erros de pequena magnitude utilizando tais métricas, mas que podem representar um erro real elevado ao utilizar uma métrica de distância real entre os pontos. Por exemplo, dado os pontos $x_1 = (22.905300, -43.133600)$ e $x_2 = (-22.9067626, -43.1355232)$ obtêm-se as métricas $RMSE = 0.00171$ e $R^2 = 1$, valores que indicam uma ótima previsão. No entanto,

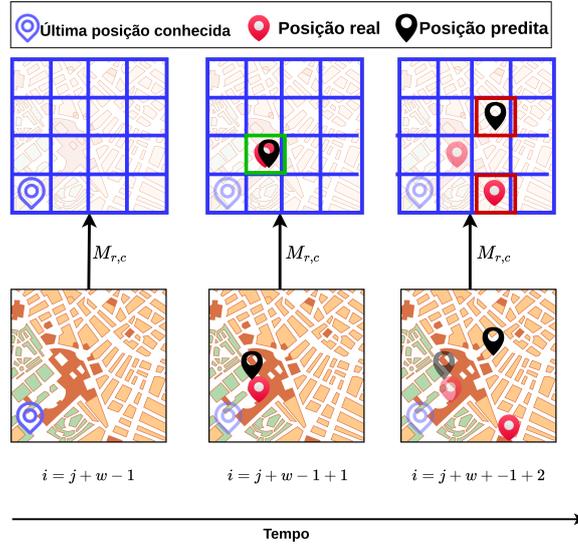


Figura 3. Exemplificação do mapeamento de *grid*. As posições reais x e as posições preditas \hat{x} pelo modelo são mapeadas para uma célula do grid através da função $M_{r,c}$. Se após o mapeamento ambas as posições estiverem na mesma célula é contabilizado um acerto, do contrário é contabilizado um erro.

esses pontos possuem uma distância de aproximadamente 200 metros. Visando representar os resultados de modo realista, serão utilizadas duas métricas principais de avaliação.

Distância: A primeira métrica será calculada baseada na distância real em metros entre o ponto real (x_i) e o ponto predito (\hat{x}_i), calculado através da Distância de Haversine [Van Brummelen 2012], conforme Equação 4.

$$D_{x_i, \hat{x}_i} = 2 \cdot r \cdot \arcsin \sqrt{\sin^2 \frac{\phi_{x_i} - \phi_{\hat{x}_i}}{2} + \cos \phi_{x_i} \cos \phi_{\hat{x}_i} + \sin^2 \frac{\omega_{x_i} - \omega_{\hat{x}_i}}{2}} \quad (4)$$

onde r é o valor do raio da Terra, ϕ representa os valores de latitude e ω os valores de longitude dos pontos x_i e \hat{x}_i .

Região de interesse: Ao contrário da métrica de distância que analisa a posição exata do veículo, a segunda métrica será baseada em uma análise de predição de região de interesse (ROI). Para tal, os pontos x e \hat{x} são mapeados para uma célula de um *grid* M através da Equação 1. Dado um conjunto de teste X e suas respectivas predições \hat{X} , o resultado da métrica de região de interesse é dado pela Equação 5:

$$ROI = \frac{\sum_{x \in X, \hat{x} \in \hat{X}} F(M_{r,c}(x), M_{r,c}(\hat{x}))}{|X|} \cdot 100 \quad (5)$$

onde,

$$F(M_{r,c}(x), M_{r,c}(\hat{x})) = \begin{cases} 1, & M_{r,c}(x) = M_{r,c}(\hat{x}) \\ 0, & M_{r,c}(x) \neq M_{r,c}(\hat{x}) \end{cases} \quad (6)$$

Dessa forma, a avaliação do resultado pela métrica de região de interesse torna-se uma avaliação por acurácia. Destaca-se que maiores as dimensões do *grid* M , maior a precisão da região real onde o veículo se encontra.

4. Avaliação da arquitetura proposta

A seguir, são apresentados os experimentos realizados e os resultados obtidos. Primeiro, serão apresentadas as bases de dados e as configurações experimentais utilizadas nos experimentos. Posteriormente, será apresentada a análise dos resultados experimentais.

4.1. Base de dados

Para a realização dos experimentos foram utilizadas duas bases de dados contendo trajetórias reais de corridas de táxi nas cidades do Porto, em Portugal, e em São Francisco, Estados Unidos.

Porto Táxi: A base de dados *Porto Táxi* [ECML/PKDD 2015] contém dados de corridas realizadas por 442 taxistas na cidade do Porto, Portugal, coletados entre 2013 e 2014. Foram coletadas aproximadamente 1.700.000 trajetórias, a coleta foi realizada obtendo coordenadas de latitude e longitude dos GPS dos táxis em intervalos de 15 segundos durante todos os trajetos.

São Francisco - SFCab: A base de dados São Francisco [Piorowski et al. 2009] contém dados de corridas realizadas por 500 taxistas na baía de São Francisco, Estados Unidos. Os dados foram coletados durante um período de 30 dias e possui aproximadamente 900 mil trajetórias. Os dados de posição de latitude e longitude de cada corrida foram coletados em um intervalo de aproximadamente 10 segundos.

A Figura 4 apresenta os mapas de calor de uma amostragem das trajetórias em São Francisco (esquerda) e Porto (direita). É notável que a cidade de São Francisco possui uma organização urbana com vias mais paralelas e perpendiculares, semelhante a um *grid*. Por outro lado, Porto possui vias mais curvas, sem um padrão regular, o que pode tornar a tarefa de predição de trajetórias mais complexa. Deste modo, a arquitetura proposta será avaliada em cidades com organização urbana distintas.



Figura 4. Mapa de calor para uma amostra da base São Francisco (esquerda) e Porto Taxi (direita). Nota-se que São Francisco concentra o tráfego em suas avenidas mais centrais e a organização de suas ruas se assemelha a um *grid*. Porto possui um tráfego mais disperso e concentrado em sua periferia.

4.2. Configurações Experimentais

Para a realização dos experimentos, optou-se pela utilização de sub-trajetórias de tamanho w como dado de entrada dos modelos de predição. Tal escolha justifica-se pelo fato de que trajetórias completas não possuem tamanho fixo, tornando o modelo mais complexo. Além disso, ao trabalhar com sub-trajetórias é possível extrair informações mais precisas dos dados, uma vez que se está prevendo o comportamento de deslocamento para pequenas divisões de uma trajetória completa.

Nota-se que o tamanho da sub-trajetória (w) é um hiperparâmetro relevante na análise dos resultados. Apesar do problema de predição de trajetórias apresentar uma dependência temporal dos dados, não é explícito qual o tamanho da dependência em relação à janela de tempo. De forma prática, dado a posição atual x_i de um veículo, não se sabe quais as posições anteriores tiveram influência direta em x_i . Ao avaliar diversos modelos com sub-trajetórias obtidas a partir de variados tamanhos de w , torna-se interessante analisar a importância da quantidade de informação passada para prever a posição futura.

De maneira a analisar tal importância, foram gerados conjuntos de dados de sub-trajetórias a partir das bases originais, Porto Taxi e São Francisco, aplicando valores de $w = \{10, 20, 30, 40, 50\}$. Na Tabela 1, é possível observar a quantidade de amostras geradas. Durante a geração de cada conjunto foram ignoradas todas as trajetórias de tamanho inferior a w e trajetórias que ultrapassem as fronteiras delimitadas pelos vértices do *grid*. Para a realização do treinamento, as bases de dados de trajetórias foram divididas em conjuntos de treino e de teste, correspondendo a 80% e 20% dos conjuntos de dados respectivamente. O conjunto de validação foi extraído diretamente do conjunto de treinamento, correspondendo a 25% dos dados. Tais dados foram utilizados para extrair os melhores hiperparâmetros do modelo, de modo que não foram utilizados no treinamento do modelo.

Tabela 1. Quantidade de amostras da base de dados original e a base de dados de sub-trajetória em relação ao parâmetro w para as bases de dados Porto Taxi e São Francisco levando em consideração predições do tipo *short-term* ($h = 1$).

w	Porto Taxi		São Francisco	
	Original	Sub-trajetórias	Original	Sub-trajetórias
10	1059435	3178305	336162	1258728
20	862744	2313720	96123	651918
30	636276	1090208	45108	451080
40	439107	891070	26850	357105
50	290457	775941	18258	296424

Para encontrar o melhor modelo de predição de trajetórias foi realizada uma busca pelos melhores hiperparâmetros da rede LSTM. Os hiperparâmetros analisados no contexto do modelo foram *hidden units* (hu) e *stacked layers* (sl). Os valores utilizados podem ser observados na Tabela 2.

Os hiperparâmetros apresentados na Tabela 1 e Tabela 2 foram utilizados exclusivamente no primeiro experimento, o qual consiste em treinar um modelo de predição *short-term*. A partir dos resultados do conjunto de validação, foram selecionados os melhores hiperparâmetros para serem utilizados no treinamento de modelos de predição *long-term*. Para o segundo experimento, serão utilizados valores $h = \{2, 3, 4, 5, 10\}$. Os resultados

Tabela 2. Conjunto de hiperparâmetros e seus respectivos valores.

Hiperparâmetro	Valores
<i>Hidden Units</i>	{32,64,128,256,512}
<i>Stacked Layers</i>	{1,2,3,4}

obtidos em ambos os experimentos são apresentados e discutidos a seguir.

4.3. Resultados experimentais

Devido à quantidade de combinações possíveis dos hiperparâmetros w , hu e sl , serão apresentados apenas os resultados com os melhores desempenhos. Além disso, a seleção de hiperparâmetros para os experimentos *long-term* foram escolhidos a partir dos melhores resultados obtidos pelo conjunto de validação durante a fase de treinamento dos experimentos *short-term*.

Inicialmente, a Figura 5 apresenta uma análise da influência do tamanho do *grid* para predições *short-term* em relação à região de interesse. Em problemas em que a posição exata do veículo não é um fator determinante e a região a qual ele se encontra é suficiente para tomadas de decisões, essa métrica torna-se relevante. Os *grids* cobrem uma área de 5.3km x 11km e 9km x 11km para as cidades de Porto e São Francisco, respectivamente. Desta forma, aumentar as dimensões a e m do grid implica em diminuir o tamanho da região de interesse e tornar o problema de predição mais complexo.

Para a cidade do porto, Figura 5(a), pode-se observar que a acurácia do modelo é inversamente proporcional ao tamanho do *grid*. Para dimensões 4x4, cujas células são de aproximadamente 1300m x 2700m, o modelo apresentou uma acurácia de aproximadamente 94%. Apesar do resultado, a região de interesse é relativamente grande.

No entanto, para *grids* de dimensões maiores, ainda se pode observar resultados relevantes. A título de comparação, o grid com dimensões 60 x 150 possui células de aproximadamente 90m x 75m.

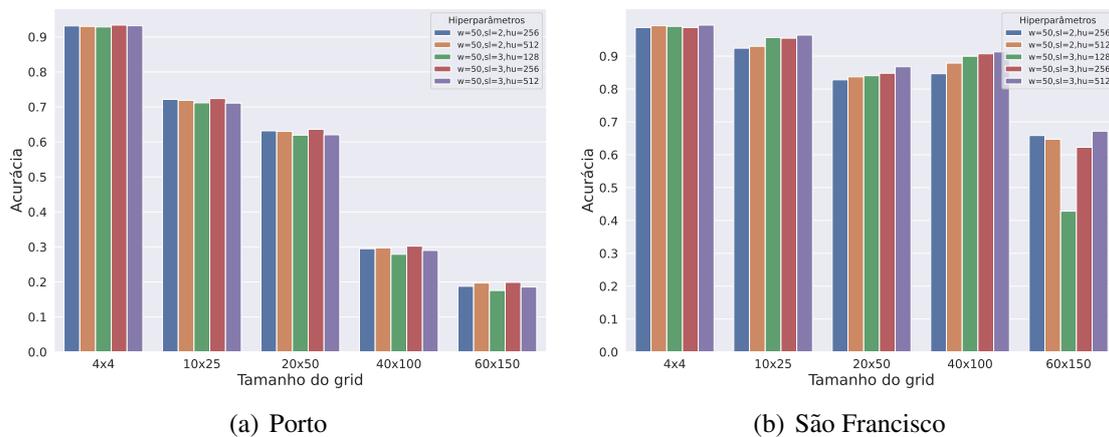
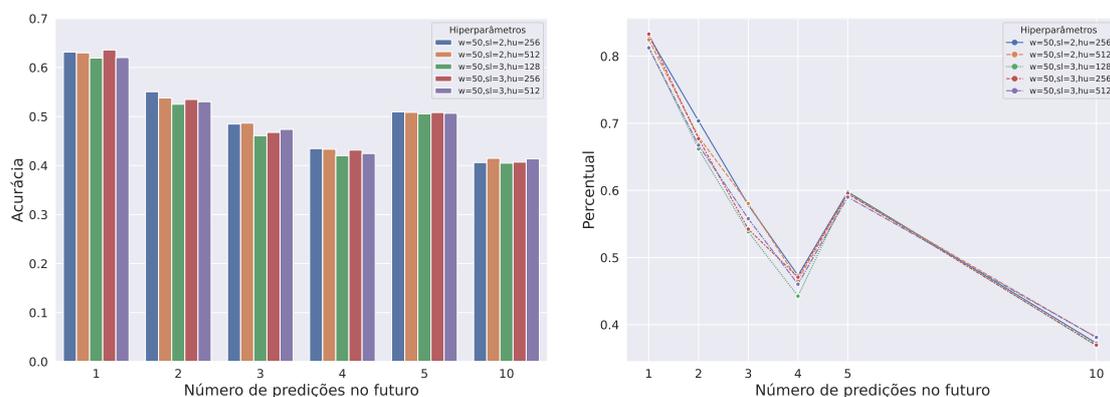


Figura 5. Influência do tamanho do *grid* na acurácia da métrica região de interesse.

A influência do *grid* na cidade de São Francisco pode ser observada na Figura 5(b). Ao contrário dos resultados apresentados para a base do Porto, inicialmente não se observa um decaimento significativo de acurácia ao aumentar o tamanho do *grid*. São Francisco

apresenta naturalmente um formato de *grid* na organização de suas ruas, além de possuir uma área superior em relação à cidade do Porto, o que pode explicar esses resultados.

A Figura 6 apresenta os resultados para predições *long-term* e discute a influência do tamanho do futuro predito pelo modelo. Para a análise da métrica de região de interesse, será adotado um *grid* de dimensões 20 x 50, o que representa células de aproximadamente 240m x 220m para a cidade do Porto. Na Figura 6(a), pode-se observar que a acurácia para predições *short-term* é de aproximadamente 64% para o melhor conjunto de hiperparâmetros. Ressalta-se ainda que no caso específico dessa métrica, um modelo aleatório teria uma probabilidade de acerto aproximada de 11.1%, uma vez que o veículo pode manter-se na posição atual ou mover-se para uma das oito células vizinhas. Em relação a predições do tipo *long-term*, observa-se uma tendência de decaimento da acurácia ao longo do tempo, explicada pela dificuldade do modelo de prever posições distantes em relação ao momento atual. No entanto, para um valor de $h = 10$ o modelo apresenta acurácia aproximada de 42%, valor superior a um modelo aleatório.

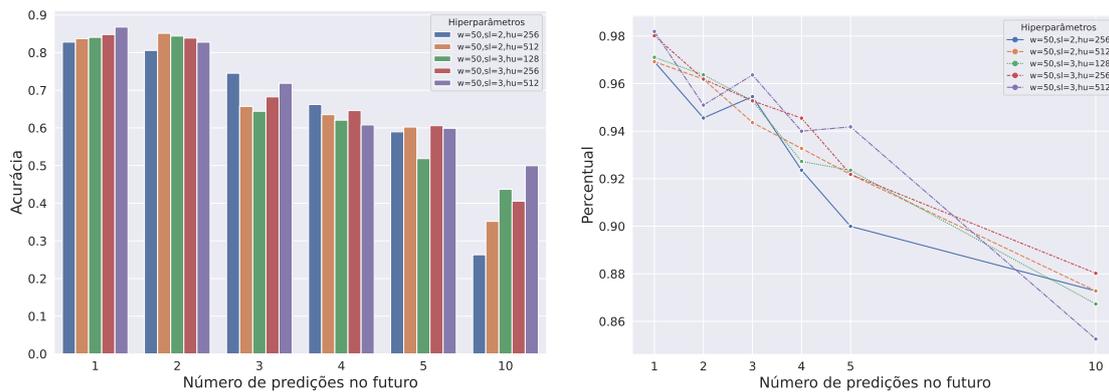


(a) Acurácia da predição de região de interesse (b) Porcentual de predições com erro inferior a 300 metros

Figura 6. Resultados obtidos para a base de dados Porto Taxi para predições *long-term*.

Na Figura 6(b), são apresentados os resultados obtidos através da distância real entre os pontos preditos e reais, ou seja, o erro absoluto em metros. Os resultados mostram que para predições *short-term* o modelo apresenta um erro absoluto inferior a 300 metros para aproximadamente 80% das amostras de teste. Assim como observado para regiões de interesse, esse valor decai para predições *long-term*, apresentando um erro absoluto inferior a 300 metros para aproximadamente 40% das amostras de teste quando $h = 10$.

Os resultados obtidos pela métrica de região de interesse para a cidade de São Francisco, utilizando um *grid* com células de dimensão 450m x 220m, podem ser observados na Figura 7(a). Para predições *short-term* observa-se uma acurácia aproximada de 83% enquanto para predição *long-term* com $h = 10$ foi obtido uma acurácia aproximada de 52%, ao utilizar o melhor conjunto de hiperparâmetros. Nota-se que para os dados de São Francisco, a arquitetura proposta conseguiu uma generalização superior em relação aos dados de Porto. Na Figura 7(b), pode-se observar o erro real em metros. Para $h = 1$ observa-se que 98% das amostras possuem um erro inferior a 300m. Enquanto que para $h = 10$, esse valor é de 88%, seguindo a tendência dos resultados obtidos através da análise de região de interesse.



(a) Acurácia da predição de região de interesse (b) Porcentual de predições com erro inferior a 300 metros

Figura 7. Resultados obtidos para a base de dados São Francisco para predições *long-term*.

Apesar da quantidade de dados inferiores em relação à base de dados Porto Taxi, os resultados obtidos para a base São Francisco demonstra-se superior em relação a ambas as métricas. Essa ocorrência pode ser causada pela organização urbana das cidades onde os dados foram coletados. Enquanto a cidade do Porto possui ruas e rodovias sem um padrão de organização definido, a cidade de São Francisco apresenta uma estrutura de *grid* que pode facilitar o aprendizado do modelo treinado.

De modo geral, os resultados obtidos e mensurados pelas métricas propostas demonstram que é possível aplicar modelos de predição de trajetória *short-term* e *long-term* e obter resultados promissores. Como esperado, quanto maior o futuro a ser predito, menor a acurácia do modelo. Além disso, a possibilidade de mensurar os resultados pelo erro absoluto ou por análise de região de interesse traz flexibilidade para as análises, a depender do problema abordado. Em casos que a posição exata do veículo não é relevante, é possível utilizar micro ou macro regiões da cidade como base de localização, porém sem perder a referência geoespacial. Os resultados obtidos mostram a influência do tamanho do *grid* no desempenho do modelo. Por fim, dado os resultados experimentais, é possível observar que o tamanho da janela temporal utilizada como entrada do modelo, ou seja, o valor do hiperparâmetro w , tem influência direta nos resultados, visto que as melhores métricas foram obtidas exclusivamente com $w = 50$.

5. Conclusão

Neste artigo, foi proposto uma arquitetura de aprendizagem baseada em LSTM's empilhadas para realizar predições *short-term* e *long-term* de trajetórias contendo dados de geoespaciais. Foram propostas duas métricas de avaliação que podem ser utilizadas em cenários onde é necessário um nível de precisão elevado referente a posição atual do veículo e outra que tem por característica analisar regiões de interesse da localidade de estudo. As análises foram realizadas com trajetórias de viagens reais de taxistas nas cidades do Porto e São Francisco.

Apesar dos desafios impostos no tratamento de dados georreferenciados, os resultados obtidos apresentam-se promissores para ambas as métricas, em especial para a métrica que analisa regiões de interesse, demonstrando a capacidade de generalização dos modelos

para suas respectivas cidades alvo. Para a base de dados de São Francisco, utilizando um *grid* com células de 450m x 220m, foi reportado uma acurácia de aproximadamente 83% para predições *short-term* e de 52% para predições *long-term* adotando $h = 10$. Os resultados mostram a viabilidade de utilizar apenas dados de localização em conjunto com um modelo robusto, sem necessidade de dados auxiliares, como proposto em outros trabalhos.

Como trabalhos futuros, pretende-se analisar a capacidade de generalizar um modelo mediante técnicas de aprendizado federado. Apesar da quantidade abundante de dados disponíveis, existem questões referentes à privacidade e ao uso dos dados. Esse problema pode ser superado com modelos locais de usuários únicos que, quando agregados, são capazes de gerar um modelo global de predição de trajetórias para uma determinada localidade, sem a necessidade do compartilhamento de trajetórias.

Agradecimentos

O presente trabalho foi realizado com apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código 001, FAPES Códigos 2022-NGKM5 e 2021-GL60J, CNPq e FAPESP - grant #2020/05121-4 e #2023/00148-0

Referências

- ECML/PKDD (2015). Porto taxi dataset. <https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i>. Acesso: 2024-01-22.
- Feng, J., Rong, C., Sun, F., Guo, D., and Li, Y. (2020). Pmf: A privacy-preserving human mobility prediction framework via federated learning. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1):1–21.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Ip, A., Irio, L., and Oliveira, R. (2021). Vehicle trajectory prediction based on lstm recurrent neural networks. In *IEEE Vehicular Technology Conference*, pages 1–5. IEEE.
- Karatzoglou, A., Köhler, D., and Beigl, M. (2018). Semantic-enhanced multi-dimensional markov chains on semantic trajectories for predicting future locations. *Sensors*, 18(10):3582.
- Khansari, N., Mostashari, A., and Mansouri, M. (2014). Impacting sustainable behavior and planning in smart city. *International journal of sustainable land Use and Urban planning*, 1(2).
- Kim, B., Kang, C. M., Kim, J., Lee, S. H., Chung, C. C., and Choi, J. W. (2017). Probabilistic vehicle trajectory prediction over occupancy grid map via recurrent neural network. In *Conference on intelligent transportation systems*, pages 399–404. IEEE.
- King, S. M., Nawab, F., and Obraczka, K. (2021). A survey of open source user activity traces with applications to user mobility characterization and modeling. *arXiv:2110.06382*.
- Kong, X., Xia, F., Wang, J., Rahim, A., and Das, S. K. (2017). Time-location-relationship combined service recommendation based on taxi trajectory data. *IEEE Transactions on Industrial Informatics*, 13(3):1202–1212.

- Koolwal, V. and Mohbey, K. K. (2020). A comprehensive survey on trajectory-based location prediction. *Iran Journal of Computer Science*, 3:65–91.
- Lv, J., Li, Q., Sun, Q., and Wang, X. (2018). T-conv: A convolutional neural network for multi-scale taxi trajectory prediction. In *2018 IEEE international conference on big data and smart computing (bigcomp)*, pages 82–89. IEEE.
- Maggi, E. and Vallino, E. (2016). Understanding urban mobility and the impact of public policies: The role of the agent-based models. *Research in Transportation Economics*, 55:50–59.
- Mehmood, Y., Ahmad, F., Yaqoob, I., Adnane, A., Imran, M., and Guizani, S. (2017). Internet-of-things-based smart cities: Recent advances and challenges. *IEEE Communications Magazine*, 55(9):16–24.
- Park, S. H., Kim, B., Kang, C. M., Chung, C. C., and Choi, J. W. (2018). Sequence-to-sequence prediction of vehicle trajectory via lstm encoder-decoder architecture. In *2018 IEEE intelligent vehicles symposium (IV)*, pages 1672–1678. IEEE.
- Piorkowski, M., Sarafijanovic-Djukic, N., and Grossglauser, M. (2009). Crawdad data set epfl/mobility (v. 2009-02-24).
- Qiao, J., Li, S., and Lin, S. (2017). Location prediction based on user mobile behavior similarity. In *2017 IEEE 23rd International Conference on Parallel and Distributed Systems (ICPADS)*, pages 783–786. IEEE.
- Rathore, P., Kumar, D., Rajasegarar, S., Palaniswami, M., and Bezdek, J. C. (2019). A scalable framework for trajectory prediction. *IEEE Transactions on Intelligent Transportation Systems*, 20(10):3860–3874.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088):533–536.
- Saleh, K., Hossny, M., and Nahavandi, S. (2018). Cyclist trajectory prediction using bidirectional recurrent neural networks. In *Advances in Artificial Intelligence, Wellington, New Zealand*, pages 284–295. Springer.
- Sharma, S. and Kaushik, B. (2019). A survey on internet of vehicles: Applications, security issues & solutions. *Vehicular Communications*, 20:100182.
- Thomé, M., Prestes, A., Gomes, R., and Mota, V. (2020). Um arcabouço para detecção e alerta de anomalias de mobilidade urbana em tempo real. In *Anais do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 784–797, Porto Alegre, RS, Brasil. SBC.
- Van Brummelen, G. (2012). *Heavenly mathematics: The forgotten art of spherical trigonometry*. Princeton University Press.
- Zhang, J., Meng, W., Liu, Q., Jiang, H., Feng, Y., and Wang, G. (2016). Efficient vehicles path planning algorithm based on taxi gps big data. *Optik*, 127(5):2579–2585.
- Zhang, P., Ouyang, W., Zhang, P., Xue, J., and Zheng, N. (2019). Sr-lstm: State refinement for lstm towards pedestrian trajectory prediction. In *Conference on Computer Vision and Pattern Recognition*, pages 12085–12094.