

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE COMPUTAÇÃO

Exame de Qualificação de Mestrado

08 de maio de 2024

QUANTIFICAÇÃO DE INCERTEZA EM REDES NEURAIIS EM GRAFOS
APLICADA À PREVISÃO DE SÉRIES TEMPORAIS MULTIVARIADAS:
UM ESTUDO DE CASO EM TRÁFEGO URBANO

Candidata: Ioná Santana

Orientador: Prof. Dr. Zanoni Dias

Coorientador: Prof. Dr. Hélio Pedrini

Sumário

1	Introdução	1
1.1	Caracterização do Problema	1
1.2	Objetivos e Contribuições	2
1.3	Hipóteses do Trabalho	3
1.4	Organização do Texto	4
2	Revisão Bibliográfica	5
2.1	Conceitos e Técnicas	5
2.1.1	Séries Temporais	5
2.1.2	Aprendizado de Máquina	6
2.1.3	Aprendizado Profundo	6
2.1.4	Redes Neurais em Grafo (GNNs)	6
2.1.5	Espaço-Temporal em GNNs (STGNNs)	8
2.1.6	Quantificação de Incerteza (QI)	9
2.2	Trabalhos Correlatos	9
2.2.1	GNNs	9
2.2.2	GNNs Aplicadas a Séries Temporais (STGNNs)	10
2.2.3	Previsão de Tráfego	12
2.2.4	QI em Previsão de Séries Temporais	13
3	Material e Métodos	14
3.1	Metodologia	14
3.2	Bases de Dados	16
3.3	Métricas de Avaliação	16
3.4	Recursos Computacionais	18
4	Plano de Trabalho e Cronograma de Execução	19
4.1	Próximos Passos	20
	Bibliografia	20

Resumo

O congestionamento de tráfego afeta a qualidade de vida urbana, impactando as rotinas diárias e a segurança viária. Esta pesquisa aborda a complexidade da previsão do tráfego, considerando fatores como a heterogeneidade de dados e influências externas. Destaca-se o uso das Redes Neurais em Grafos (GNNs) e da Quantificação de Incerteza (QI) como ferramentas nesse contexto. O estudo visa explorar a aplicação conjunta de GNNs e QI na previsão de séries temporais multivariadas, com foco específico em tráfego urbano. Ao investigar métodos avançados de QI adaptados a GNNs, busca-se melhorar a confiabilidade das previsões em ambientes dinâmicos. Utilizando as bases de dados PeMSD4 e PeMSD8, este projeto visa avançar no entendimento das dinâmicas temporais e espaciais do tráfego, contribuindo para uma gestão mais eficiente e segura das vias urbanas.

Capítulo 1

Introdução

O congestionamento de tráfego representa um desafio significativo para a qualidade de vida da comunidade, impactando rotinas diárias e bem-estar em geral [13]. Além disso, acidentes de trânsito sobrecarregam sistemas de saúde, sendo uma das principais causas de morte em todo o mundo, especialmente em países de baixa e média renda [1].

Prever o tráfego é complexo devido a vários fatores inter-relacionados, incluindo a heterogeneidade dos dados, padrões sazonais e a influência de variáveis externas, como eventos especiais e condições climática. Além disso, a escassez de dados históricos de qualidade e a necessidade de lidar com dados ausentes ou inconsistentes representam desafios adicionais para pesquisadores e profissionais da área [15, 20, 27, 40]

Nesse contexto, as Redes Neurais em Grafos (GNNs) emergem como ferramentas para lidar com dados estruturados, como em redes sociais e de transporte, devido à sua capacidade de modelar relacionamentos complexos [39]. Além disso, a Quantificação de Incerteza (QI) em modelos de aprendizado de máquina tem adicionado confiabilidade nas previsões, contribuindo para uma tomada de decisão mais informada [16].

Este estudo visa explorar a aplicação de GNNs e QI na previsão de séries temporais multivariadas, com foco em tráfego urbano. A integração dessas abordagens visam avançar no entendimento das dinâmicas temporais e espaciais em sistemas complexos, impulsionando o campo do aprendizado de máquina e análise de dados estruturados.

Neste capítulo, será descrito o problema abordado neste projeto, bem como apresentado os objetivos, as contribuições esperadas, as hipóteses do trabalho e a organização do texto.

1.1 Caracterização do Problema

Prever o fluxo de tráfego desempenha um papel crucial em várias indústrias, economizando tempo e recursos. Empresas de logística e entrega podem otimizar rotas, reduzir tempos de entrega e aumentar a eficiência operacional, economizando combustível [30]. Além disso, a previsão melhora a segurança viária, identificando áreas propensas a congestionamentos e potenciais pontos de acidentes, permitindo medidas preventivas e aprimorando a gestão do tráfego [23].

Sistemas de navegação e aplicativos de trânsito, como o Waze¹, dependem dessas previsões para fornecer orientações precisas em tempo real, melhorando a experiência dos usuários e promovendo um ambiente de deslocamento mais seguro. Antecipar o tráfego não apenas melhora a mobilidade urbana, mas também contribui para estratégias ambientais, permitindo o controle de emissões e poluição, permite simplificar o deslocamento diário e molda o funcionamento das cidades em um mundo cada vez mais dinâmico, promovendo a segurança e bem-estar dos habitantes [24].

A rápida urbanização e o aumento da mobilidade urbana demandam soluções avançadas para a previsão de tráfego, essenciais para o planejamento eficiente de infraestruturas e serviços. Redes neurais em grafos oferecem uma representação poderosa para modelar a interconexão espacial e temporal entre diferentes pontos em uma rede de tráfego, enquanto a QI traz uma dimensão adicional à confiabilidade das previsões [16].

Apesar dos avanços em técnicas de previsão, a maioria das abordagens negligencia a incerteza associada às previsões, limitando a utilidade prática desses modelos em contextos dinâmicos e imprevisíveis, como o tráfego urbano [16]. Este projeto visa preencher essa lacuna ao investigar métodos avançados de QI adaptados a redes neurais em grafos.

Para conduzir a pesquisa, utilizaremos as bases de dados PeMSD4 e PeMSD8, disponíveis no site da *Caltrans Performance Measurement System (PeMS)*². Essas bases fornecem informações sobre o desempenho do sistema de rodovias da Califórnia, coletadas em tempo real por quase 40 mil detectores individuais. O repositório oferece mais de dez anos de dados históricos para análise e integra informações de outros sistemas de agências locais, como incidentes, fechamentos de faixas, *tags* de pedágio, contagens de tráfego, classificação de veículos, peso em movimento e inventário de rodovias.

A crescente complexidade e interconexão dos sistemas urbanos têm desafiado os métodos tradicionais de previsão, especialmente no contexto de séries temporais multivariadas em ambientes dinâmicos, como o tráfego urbano. A incorporação de incerteza nas previsões torna-se um fator crítico para garantir tomadas de decisão robustas e eficientes. Este projeto de mestrado propõe abordar essa necessidade premente ao explorar métodos de QI em redes neurais em grafos, focando especificamente na previsão de tráfego urbano.

1.2 Objetivos e Contribuições

O principal objetivo deste trabalho é desenvolver uma abordagem capaz de quantificar incertezas na previsão de séries temporais em redes neurais em grafos. Essa abordagem visa aprimorar a interpretabilidade e confiabilidade das previsões, além de promover o desenvolvimento de modelos capazes de aprender com a incerteza inerente aos dados.

Para atingir essa meta, os seguintes objetivos específicos foram delineados:

- 1. Levantamento Bibliográfico:** Realizar um levantamento bibliográfico abrangente, explorando as abordagens utilizadas para a previsão de séries temporais, destacando os avanços mais recentes nesse campo.

¹<https://www.waze.com>

²<https://pems.dot.ca.gov>

2. **Aquisição de Dados:** Coletar e analisar as bases de dados PeMS4 e PeMS8, fundamentais para a validação e teste da metodologia proposta.
3. **Proposição de Métodos para Previsão:** Desenvolver métodos para a previsão de séries temporais multivariadas de tráfego, utilizando redes neurais em grafos como base para o modelo.
4. **Proposição de Métodos para QI:** Propor métodos para a QI em redes neurais em grafos, aplicados especificamente à previsão de séries temporais multivariadas em tráfego urbano.
5. **Realização de Experimentos:** Conduzir experimentos para validar a eficácia e eficiência dos métodos propostos em cenários realistas de previsão de tráfego.
6. **Avaliação e Comparação:** Avaliar e comparar o desempenho do método proposto com outras abordagens disponíveis na literatura, identificando suas vantagens e limitações.
7. **Documentação e Publicação:** Documentar os resultados obtidos, destacando as contribuições específicas para o campo de previsão de tráfego. Publicar esses resultados em veículos científicos relevantes para compartilhar o conhecimento gerado.

1.3 Hipóteses do Trabalho

As seguintes hipóteses serão investigadas neste trabalho:

- O modelo proposto será capaz de generalizar a previsão de fluxo de tráfego para diferentes cidades, demonstrando sua eficácia em ambientes urbanos diversos;
- O modelo apresentará robustez ao ser testado em bases de dados temporais de diferentes períodos, mostrando que sua capacidade de generalização não é limitada por flutuações sazonais ou tendências a curto prazo;
- A incorporação de técnicas de quantificação de incerteza resultará em previsões mais confiáveis, evidenciadas por métricas de avaliação que demonstram menor dispersão nos resultados previstos em comparação com abordagens que não consideram a incerteza;
- A análise das representações aprendidas pelo modelo revelará padrões e correlações entre variáveis que não são facilmente identificáveis por métodos tradicionais de previsão de séries temporais;
- O modelo proposto manterá um desempenho computacional eficiente e escalável mesmo ao lidar com grandes volumes de dados de tráfego urbano.

Este projeto busca contribuir para o avanço da pesquisa em métodos de previsão de tráfego, não apenas oferecendo soluções para desafios específicos do tráfego urbano, mas também proporcionando um impacto mais amplo em domínios que compartilhem complexidades semelhantes.

1.4 Organização do Texto

O Capítulo 2 descreve os conceitos e técnicas relevantes relacionados ao tema sob investigação. O Capítulo 3 descreve a metodologia proposta, a base de dados, as métricas de avaliação e os recursos computacionais que serão empregados no desenvolvimento do projeto. O Capítulo 4 apresenta o plano de trabalho e o cronograma de execução das atividades.

Capítulo 2

Revisão Bibliográfica

Este capítulo descreve conceitos e técnicas relacionados ao tema sob investigação.

2.1 Conceitos e Técnicas

Esta seção apresenta os principais conceitos e técnicas empregados neste trabalho, como previsão de séries temporais, redes neurais em grafos e quantificação de incerteza.

2.1.1 Séries Temporais

Uma série temporal é uma sequência de observações ou medições coletadas em intervalos sucessivos ao longo do tempo. Cada observação está associada a um momento específico, criando uma relação temporal entre os dados [12]. Essas séries temporais podem ser univariadas, envolvendo apenas uma variável ao longo do tempo, ou multivariadas, nas quais várias variáveis são observadas simultaneamente ao longo do tempo.

Séries Temporais Univariada

Uma série temporal univariada é um conjunto de observações ou medidas coletadas ao longo do tempo para uma única variável. Z é uma série temporal e é definida como uma sequência ordenada de n variáveis de valor real:

$$Z = (t_1, \dots, t_n), t_i \in \mathbb{R}$$

A série temporal é um conjunto de observações coletadas de forma sequencial, ao longo do tempo, e conforme uma determinada taxa de amostragem. Nesse sentido, Z pode ser definida como um conjunto de instantes de tempo consecutivos e pode abranger todo o conjunto de dados fornecidos pela observação de um processo [12].

Exemplos comuns de séries temporais univariadas incluem dados diários de temperatura, vendas diárias de um produto específico ao longo do tempo, ou a taxa diária de fechamento de uma ação na bolsa de valores. A análise de séries temporais univariadas geralmente envolve a identificação de padrões temporais, tendências, sazonalidades e a modelagem de comportamentos futuros com base nos dados históricos.

Séries Temporais Multivariada

Uma série temporal multivariada é um conjunto de observações coletadas ao longo do tempo para múltiplas variáveis simultaneamente. Cada ponto na série representa um conjunto de observações correspondentes a diferentes variáveis. Dessa forma, temos k séries temporais, denotadas como Z_1, Z_2, \dots, Z_k , devem ser analisadas. Tais processos multivariados são de interesse em uma variedade de campos, como economia, negócios e ciências sociais [3].

Esse tipo de série temporal oferece a capacidade de analisar as relações entre as diversas variáveis e como essas variáveis evoluem em conjunto ao longo do tempo. A modelagem de séries temporais multivariadas é mais complexa em comparação com a univariada, devido à necessidade de lidar com as interações entre as variáveis. No entanto, pode ser extremamente útil para compreender como mudanças em uma variável afetam outras variáveis ao longo do tempo [3].

2.1.2 Aprendizado de Máquina

A aprendizagem de máquina envolve o reconhecimento de padrões por dispositivos computacionais. Esses dispositivos são dotados de algoritmos capazes de aprender com dados anteriores e realizar previsões para informações futuras que serão analisadas.

O aprendizado de máquina tem diversas aplicações na sociedade. Por exemplo, ele é utilizado para recomendar produtos em lojas online, identificar objetos em imagens e realizar buscas filtradas em navegadores da Internet [18].

2.1.3 Aprendizado Profundo

Aprendizado profundo (*Deep Learning* - DL) é um subcampo da inteligência artificial e aprendizado de máquina que utiliza redes neurais com várias camadas (também conhecidas como redes neurais profundas) para simular o raciocínio humano e a capacidade de aprendizagem. A estrutura dessas redes é inspirada pelo entendimento biológico do cérebro humano, especificamente no que diz respeito à forma como os neurônios processam informações em camadas complexas e interconectadas [31].

Há vários avanços em diferentes frentes usando aprendizado profundo, tais como reconhecimento de imagens e voz, detecção de fraudes, filtragem de e-mails/spam, modelagem de riscos financeiros, carros autônomos, descoberta de medicamentos, reconhecimento de objeto, entre outros [7].

2.1.4 Redes Neurais em Grafo (GNNs)

Uma GNN é um tipo de modelo de aprendizado de máquina que opera em dados representados na forma de grafos. Em um grafo, os dados são estruturados como nós (ou vértices) que estão interconectados por arestas (ou linhas). Mapa de rotas é um exemplo de grafo: as rodovias que conectam cidades são as arestas e as cidades são os vértices. Na Figura 2.1, temos um mapa de rotas entre três cidades (A, B e C).

O funcionamento de uma GNN pode ser dividido nas seguintes etapas [43]:

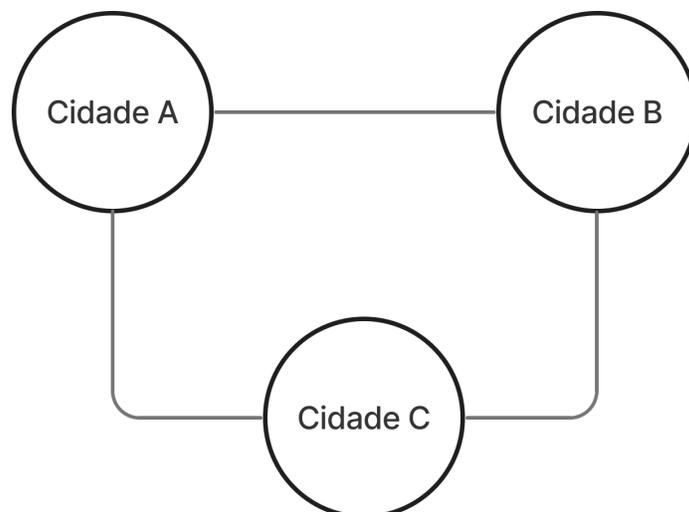


Figura 2.1: Mapa de rotas entre três cidades.

- **Representação dos nós:** Cada nó no grafo é representado por um vetor de características. Essas características podem incluir informações sobre o próprio nó e suas conexões com outros nós.
- **Propagação de mensagens:** A GNN propaga informações através das arestas do grafo. Isso é feito calculando uma mensagem (ou vetor de informação) para cada aresta, com base nas características dos nós conectados por essa aresta.
- **Agregação de informações:** Após a propagação das mensagens, as informações são agregadas em cada nó. Isso geralmente é feito combinando as mensagens recebidas com as características do próprio nó, resultando em uma nova representação do nó que captura informações de seus vizinhos.
- **Atualização dos nós:** Com as informações agregadas, os nós são atualizados com base nas novas representações. Isso pode ser feito usando funções de ativação não lineares, como a função sigmoide ou ReLU (*Rectified Linear Unit*), para transformar as informações agregadas em representações atualizadas dos nós.
- **Iteração:** O processo de propagação de mensagens, agregação de informações e atualização dos nós pode ser repetido por várias iterações, permitindo que as informações se propaguem através do grafo e que cada nó seja atualizado com informações de seus vizinhos.

Ao final do processo, a GNN produzirá uma representação atualizada de cada nó no grafo, que pode ser usada para diferentes tarefas, como classificação, geração ou predição. O treinamento de uma GNN envolve ajustar os parâmetros do modelo para minimizar uma função de perda, comumente usando técnicas de otimização como o gradiente descendente. Na Figura 2.2, temos uma GNN com uma camada de entrada, duas camadas escondidas usando a ReLU como função de ativação e uma camada de saída.

GNNs mostram-se vantajosas para a modelagem de relacionamentos complexos entre entidades interconectadas em dados estruturados, tais como redes sociais, biológicas e de transporte. Elas possuem capacidade de generalização, tornando-as úteis em cenários de previsão e classificação e são capazes de incorporar informações locais e globais em um grafo, permitindo a captura de padrões em diferentes escalas.

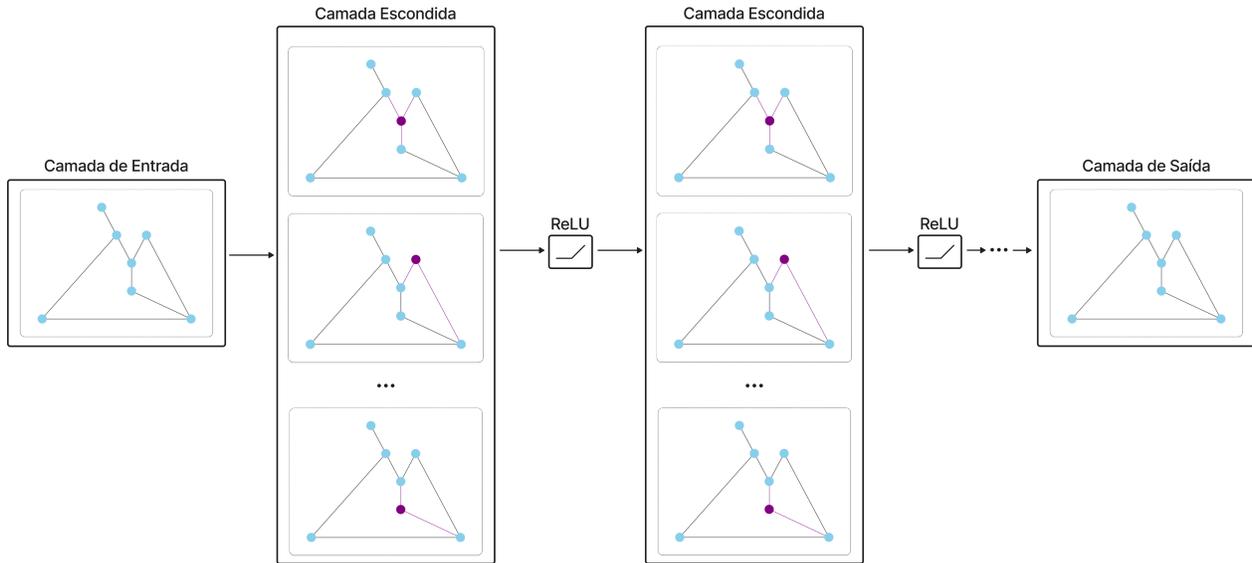


Figura 2.2: Funcionamento de uma GNN.

2.1.5 Espaço-Temporal em GNNs (STGNNs)

Redes neurais em grafos com componente espaço-temporal referem-se à inclusão de informações temporais na estrutura de grafos de uma rede neural. Isso significa que, além das relações espaciais entre entidades (nós) no grafo, também são consideradas as mudanças dessas relações ao longo do tempo.

Cada nó pode representar uma entidade específica, e as arestas indicam as relações entre essas entidades. A dimensão temporal é incorporada através da evolução dessas relações ao longo de diferentes pontos no tempo, como pode ser visto na Figura 2.3. Isso é relevante em cenários onde as interações e influências entre elementos não são estáticas e podem variar ao longo do tempo.

Essa abordagem é comumente aplicada em tarefas como previsão de séries temporais, previsão de velocidade do tráfego, antecipação de manobras de motoristas e reconhecimento de ações humanas, entre outros. A ideia central das STGNNs é considerar dependência espacial e dependência temporal ao mesmo tempo [39]. A inclusão do espaço temporal permite que a rede neural em grafos capture padrões temporais e variações dinâmicas, melhorando assim a capacidade do modelo em entender e prever comportamentos ao longo de diferentes intervalos temporais.

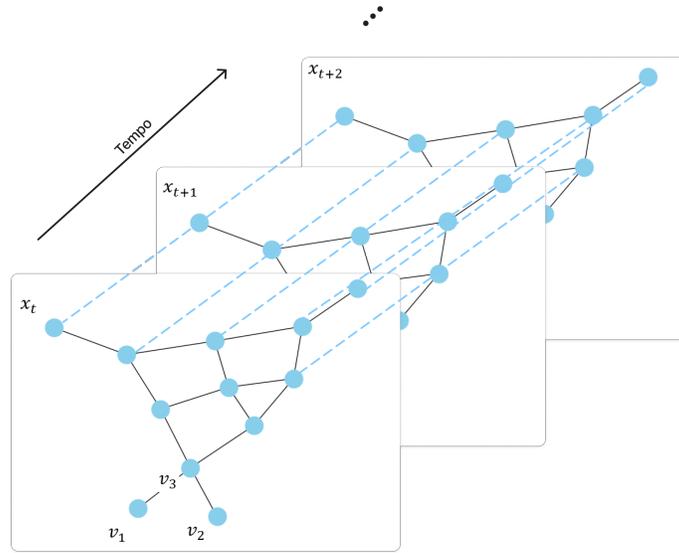


Figura 2.3: Modelagem de dependência espaço-temporal para séries temporais multivariadas, onde as linhas cinzas e linhas tracejadas verdes representam dependência espacial e temporal, respectivamente.

2.1.6 Quantificação de Incerteza (QI)

Vários algoritmos de aprendizado de máquina possuem a característica de serem caixas pretas, resultando em fraca interpretabilidade, falta de robustez, fraca confiabilidade, e vulnerabilidades a ataques adversários e sistemas ruidosos [6]. Assim, a QI em um *framework* para algoritmos de aprendizado de máquina refere-se à capacidade de um modelo não apenas fazer previsões, mas também fornecer uma medida de quão certa ou incerta essa previsão é. Essa incerteza pode surgir de diversos fatores, como a qualidade e quantidade dos dados disponíveis, a complexidade do modelo e as suposições feitas durante o treinamento.

Dois tipos principais de incerteza são identificados: epistêmica (ou incerteza do modelo) e aleatória (ou incerteza dos dados). A incerteza epistêmica é associada ao conhecimento incompleto sobre o modelo ideal e pode ser reduzida com mais dados de treinamento. A incerteza aleatória, por outro lado, é inerente aos dados e não pode ser eliminada. As duas incertezas são ilustradas na Figura 2.4.

2.2 Trabalhos Correlatos

Nesta seção, abordaremos trabalhos correlatos, incluindo GNNs e sua aplicação em séries temporais. Exploraremos também técnicas usadas em previsão de tráfego e QI em séries temporais.

2.2.1 GNNs

GNNs têm aplicações em diferentes domínios, por exemplo, podendo ser usadas em problemas de bioinformática, tais como em tarefas de previsão de fatores associados a doenças e doenças

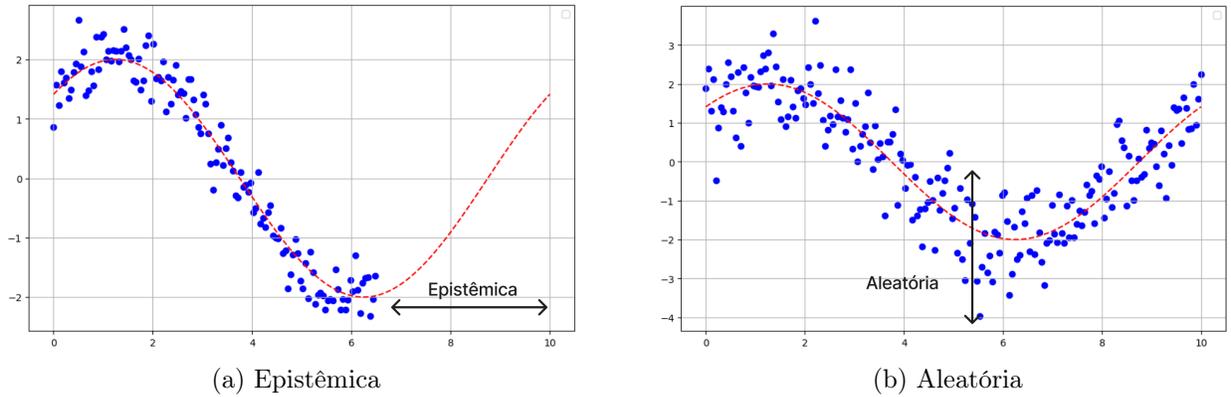


Figura 2.4: Quantificação de incerteza.

associadas a genes [42], na detecção de fraude [21] e em energia sustentável [22], como previsão da velocidade e potência de vento [16]. No campo da visão computacional, elas têm sido aplicadas para resolver problemas de saliência visual, como detecção de objetos salientes e co-saliência [10]. Wu et al. [38] aplicaram GNNs na detecção de anomalias em transporte inteligente, energia inteligente e fábrica inteligente no contexto de Internet Industrial das Coisas (IIoT).

Kipf & Welling [17] propuseram um modelo com solução escalável para aprendizado semi-supervisionado em dados de grafos, empregando redes neurais convolucionais eficientes que operam diretamente nos grafos. Em experimentos realizados em redes de citação e conjuntos de dados de grafos de conhecimento, o modelo superou métodos relacionados, como métodos baseados em *skip-gram*, método *Independent Component Analysis* (ICA) e métodos baseados em regularização do grafo Laplaciano, destacando sua eficácia na classificação de nós em grafos. A escolha da arquitetura convolucional é baseada em uma aproximação de convoluções espectrais de primeira ordem localizadas em grafos.

Apesar do sucesso alcançado, a aplicação de GNNs ainda enfrenta desafios. Por exemplo, a escalabilidade de GNNs para grafos de grande escala e a interpretabilidade dos resultados permanecem áreas de pesquisa ativa [16, 37].

Por fim, Wu et al. [37] destacaram os desafios persistentes em relação à escalabilidade, especialmente ao lidar com grafos em larga escala, contendo milhões de nós e bilhões de arestas. Estratégias de amostragem têm sido propostas para abordar essa questão. Além disso, há uma tendência crescente em tornar os resultados das GNNs mais interpretáveis, com métodos variados, embora muitos sejam computacionalmente intensivos. Paralelamente, a preocupação com a robustez adversarial tem motivado pesquisas sobre como proteger as GNNs contra ataques maliciosos, refletindo uma preocupação mais ampla sobre a confiabilidade da aprendizagem de máquina.

2.2.2 GNNs Aplicadas a Séries Temporais (STGNNs)

Nos últimos anos, as GNNs emergiram como uma ferramenta para aprender representações de dados não euclidianos, abrindo caminho para a modelagem de dados de séries temporais

do mundo real. Integrando GNNs com *frameworks* de modelagem temporal, estudos têm capturado tanto a dinâmica espacial quanto temporal, resultando em avanços em diversas tarefas, incluindo previsão, classificação, detecção de anomalias e imputação de dados em séries temporais [16].

Estudos comparativos têm demonstrado que as STGNNs frequentemente superam modelos tradicionais de séries temporais em termos de precisão e capacidade de capturar padrões complexos [39]. A capacidade das GNNs de incorporar informações de vizinhança e estrutura de rede em suas previsões as torna particularmente eficazes em cenários onde os dados exibem dependências espaciais e temporais [39].

O uso de GNNs na previsão de séries temporais tem sido explorado em diversos contextos. Aplicações anteriores incluem previsões climáticas, onde *Hierarchical Spatio-Temporal Graph Neural Network* (HiSTGNN) [25] foi utilizada para capturar relações espaciais e temporais entre diferentes variáveis meteorológicas e regiões, superando vários modelos bases, tais como Redes Neurais Convolucionais (CNNs) e Redes Neurais Recorrentes (RNNs) em três bases de dados meteorológicas capturando dependências espaciais ocultas e diversas tendências meteorológicas de longo prazo.

Bloemheuvél et al. [2] propuseram a arquitetura TISER-GCN, que processa séries temporais como uma tarefa de regressão multivariada em dois conjuntos de dados sísmicos para prever a intensidade máxima de tremores de terra em cada estação sísmica. Os resultados mostraram uma redução média do MSE de 16,3% em relação as *baselines*, e obtiveram resultados comparáveis usando apenas a metade do tamanho da entrada.

Várias arquiteturas e algoritmos foram propostos para aplicar GNNs na previsão de séries temporais [4]. Um desses modelos é o *Graph Convolutional Recurrent Networks* (GCRNs), que combina a capacidade de modelagem temporal das RNNs com a capacidade de modelagem estrutural das GNNs, sendo que a maioria dos modelos constrói um grafo onde cada nó está associado a uma série temporal, seja univariada ou multivariada, e um exemplo de arquitetura é a *Variational Graph Convolutional Recurrent Network* (VGCRN) [5].

Outra abordagem é o uso de mecanismos de atenção como as Redes Neurais de Atenção de Grafos (GANNs). Esses mecanismos de atenção são utilizados em modelos de séries temporais de grafos, ajudando na modelagem da interdependência entre nós e entre passos de tempo. Uma arquitetura com esse mecanismo é a *Gated Attention Networks* (GaAN) [41]. Esses modelos aproveitam o poder das GNNs para modelar explicitamente as relações intertemporais e interváriáveis. GCRNs e GANNs são exemplos de como as GNNs podem ser aplicadas na previsão de séries temporais para melhorar o desempenho e capturar dependências nos dados.

Há lacunas e possíveis direções em relação às GNNs para séries temporais, como estratégias de pré-treinamento e aprendizado por transferência, visando aprimorar a capacidade preditiva em cenários com dados limitados ou diversos. A interpretabilidade dos modelos é um desafio, sendo necessário desenvolver métodos que proporcionem uma compreensão clara do processo de tomada de decisão das GNNs. Paralelamente, é fundamental fortalecer a robustez desses modelos, com foco em estratégias eficazes para lidar com perturbações nos dados e mudanças na distribuição, especialmente em contextos críticos, como na área da saúde e na previsão financeira [16].

2.2.3 Previsão de Tráfego

O congestionamento de tráfego é um problema com sérias consequências econômicas, sociais e ambientais. Estudos têm se concentrado nas causas do congestionamento de tráfego, suas consequências e métodos para minimizá-lo [9].

Algoritmos de aprendizado de máquina têm mostrado avanços na análise de dados de tráfego, na previsão de congestionamentos e na otimização do gerenciamento. A integração de aprendizado de máquina no gerenciamento de tráfego é vital para atender aos requisitos de sistemas de transporte eficientes e sustentáveis [19]. No geral, esses métodos e tecnologias de previsão visam fornecer informações precisas e oportunas sobre o fluxo de tráfego para ajudar os motoristas e passageiros a tomar decisões e gerenciar o tráfego de forma eficaz [26].

A previsão de tráfego é feita usando diversos algoritmos, como algoritmos de árvore de decisão, *Support Vector Machines* e *Random Florest* [26]. Esses métodos analisam dados de tráfego para prever com precisão o fluxo de tráfego. A precisão da previsão varia dependendo do método utilizado. Por exemplo, o modelo proposto por Slimani et al. [33] compararam 3 métodos, usando abordagem multivariada para o *Multilayer Perceptron* (MLP) e abordagem univariada para os modelos *Seasonal Autoregressive Integrated Moving Average* (SARIMA) e *Support Machine Regression* (SMOreg). Obtendo o menor erro relativo absoluto de 0,57% usando a rede neural MLP, para base de dados da estação S140 “Bounika” localizada entre Rabat e Casablanca.

Trinh et al. [35] mostraram que os modelos de DL, especialmente os modelos multivariados que incorporam dados de fluxo de tráfego, condições climáticas e grafos, onde as conexões são as posições de tráfego, alcançam uma precisão de previsão melhor em comparação com modelos univariados e modelos de aprendizado de máquina. A análise também destaca a importância de considerar as características de diferentes tipos de estradas ao selecionar métodos de previsão para sistemas de transporte inteligentes [35].

A heterogeneidade dos dados, padrões sazonais e dependência de fatores externos como eventos especiais e condições climáticas são alguns dos desafios. Além disso, a escassez de dados históricos de qualidade e a necessidade de lidar com dados ausentes ou inconsistentes representam desafios adicionais para pesquisadores e profissionais da área [15, 20, 27, 40].

Diversos modelos e técnicas foram propostos para enfrentar esses desafios, tais como: decomposição empírica completa de conjunto com ruído adaptativo (CEEMDAN), entropia temporal de estimativa de rede neural (NNetEn), decomposição de modo variacional (VMD) aprimorada pelo algoritmo de otimização Northern Goshawk (NGO), máquina de aprendizado extremo de *kernel* (KELM) aprimorada pela meta-heurística *Artificial Rabbits Optimization* (ARO) e correção de erro (EC) [20]. Outras abordagens incluem correspondência média preditiva (PMM) e vizinho mais próximo (KNN) para imputação de dados, e rede neural de memória de longo e curto prazo bidirecional (BiLSTM) para previsão [40].

Além disso, rede neural espaço-temporal de gráfico dinâmico (DGSTN) obteve melhores resultados nas Bases PeMSD4 e PeMSD8 em relação a modelos clássicos, como ARIMA e métodos avançados de redes neurais usando GCNs e métodos de atenção, fazendo imputação de dados ausentes e com a captura de relacionamentos espaço-temporais [15].

2.2.4 QI em Previsão de Séries Temporais

Nos últimos anos, houve uma ênfase na pesquisa em relação a quantificação da incerteza associada às previsões de séries temporais, a fim de promover estimativas probabilísticas confiáveis e lidar adequadamente com as incertezas inerentes aos dados de séries temporais [16].

Estudos empíricos têm sido conduzidos para comparar a precisão das previsões geradas por diferentes abordagens de QI, levando em consideração fatores como complexidade computacional, robustez a perturbações nos dados e capacidade de capturar diferentes fontes de incerteza [14, 29].

Vários estudos exploraram métodos para estimar e comunicar essa incerteza, visando melhorar a confiança e a utilidade das previsões [8, 11, 34]. Uma abordagem é usar modelos de Processo Gaussiano de Substituição para estimar a incerteza, que podem ser aplicados a vários modelos base [11]. Outro método envolve o uso de RNNs e equações diferenciais estocásticas para a imputação de séries temporais, onde as medições foram realizadas de forma irregular, bem como para a quantificação da incerteza associada a essas imputações [8]. A regressão linear Bayesiana também pode ser usada para prever a incerteza em dados de sensoriamento remoto, com a incerteza do modelo avaliada usando parâmetros de probabilidade a priori e a posteriori [34].

Apesar dos avanços recentes, a aplicação eficaz de técnicas de QI a dados de séries temporais ainda enfrenta desafios significativos. Esses desafios incluem a seleção de modelos e parâmetros apropriados, o tratamento de dados ausentes ou ruidosos e a interpretação e comunicação eficaz dos resultados aos usuários finais [16].

Qian et al. [29] propuseram um método para estimar a incerteza em previsões de tráfego, integrando técnicas avançadas de DL, *Variational Inference* e *Deep Ensembling*. A proposta *Deep Spatio-Temporal Uncertainty Quantification* (DeepSTUQ) destaca-se por sua capacidade de estimar tanto a incerteza aleatória quanto a epistêmica, fornecendo informações para a gestão eficiente de sistemas de transporte urbano. As extensas experimentações realizadas em quatro conjuntos de dados públicos demonstraram que o método proposto supera as abordagens existentes em termos de previsão pontual e QI. No entanto, algumas lacunas em aberto, como a modelagem de incerteza não gaussiana e estratégias de aprendizado de grafo temporal, apontam para possíveis direções futuras de pesquisa para aprimorar a eficácia e precisão do modelo.

A QI para previsão de séries temporais tem algumas desvantagens. Por exemplo, a dependência de suposições específicas de dados, o que pode limitar a aplicabilidade do método e a ineficiência computacional de algumas abordagens, sendo necessário múltiplos modelos adicionais para estimar a incerteza [11]. Além disso, alguns métodos podem não capturar com precisão a complexidade das correlações espaço-temporais nos dados, levando a estimativas de incerteza menos precisas [29, 32]. Essas limitações destacam a necessidade de métodos aprimorados que possam fornecer QI precisa e computacionalmente eficiente para tarefas de previsão de séries temporais.

Capítulo 3

Material e Métodos

Este capítulo descreve a metodologia, a base de dados, as métricas de avaliação e os recursos computacionais que serão utilizados no desenvolvimento do projeto.

3.1 Metodologia

Esta seção apresenta a metodologia proposta que será utilizada na investigação de técnicas de QI em redes neurais em grafos aplicada a previsão de séries temporais multivariadas. A Figura 3.1 representa o fluxo da metodologia proposta.

Aquisição de Dados

A aquisição de dados para este estudo será realizada a partir da base de dados PeMS, conforme detalhado na Seção 3.2. A etapa inicial do processo de análise é o pré-processamento dos dados, discutido na subseção seguinte.

Pré-Processamento

Após a seleção de uma das bases de dados, será conduzida uma análise exploratória da qualidade dos dados, tais como inconsistências e dados faltantes. Com base nessa análise, serão definidas estratégias de pré-processamento dos dados. Essas estratégias podem incluir a imputação de dados faltantes, a normalização dos dados, filtragem, entre outras.

Divisão dos Dados

A base será dividida em duas partes, 90% para treinamento/validação e 10% para teste. Para o treinamento e a validação, utilizaremos o método de validação cruzada em janelas deslizantes (*Time Series Split*). Os atributos da rede serão avaliados para que possamos encontrar os melhores parâmetros. Os resultados obtidos serão avaliados pelas métricas apresentadas na Seção 3.3.

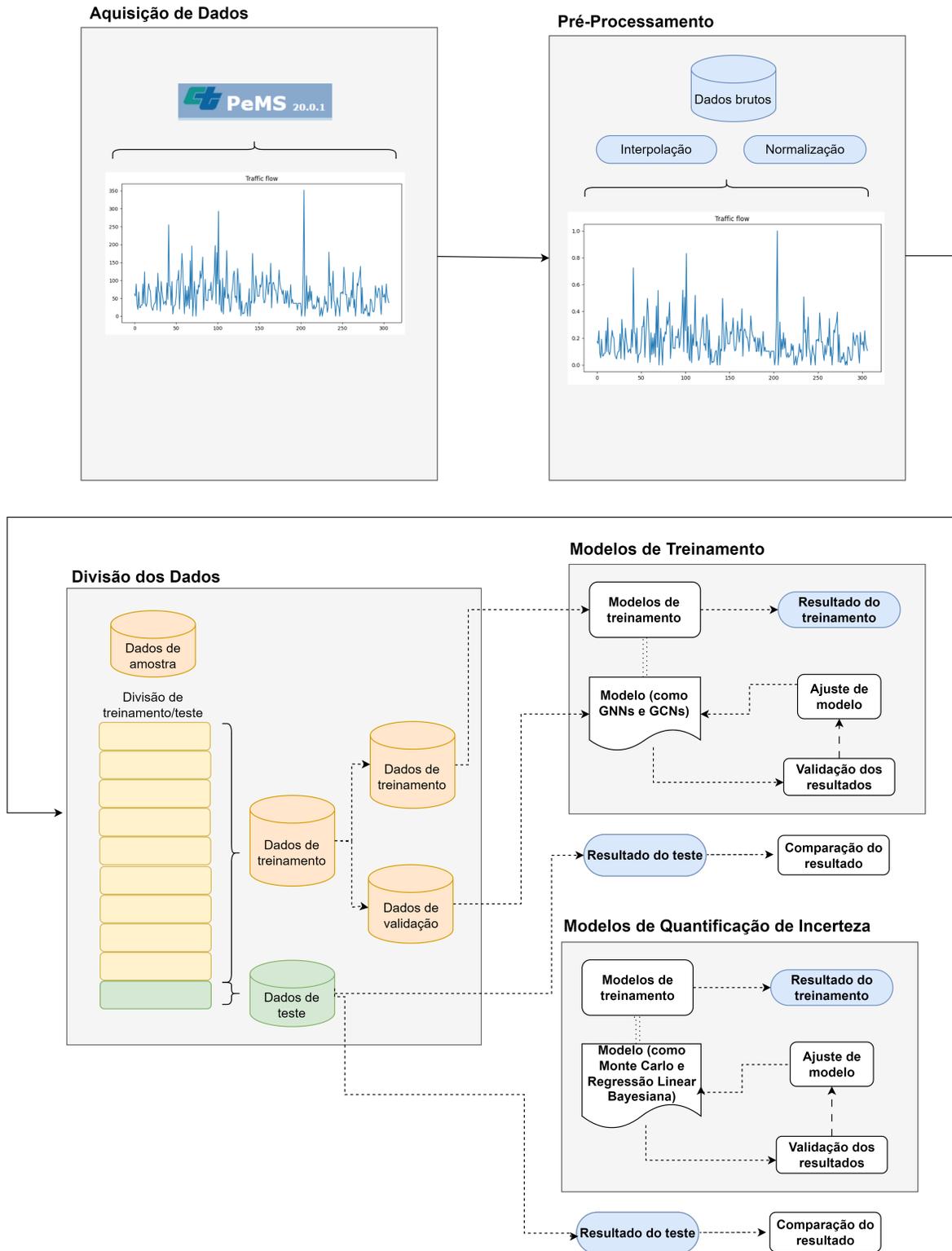


Figura 3.1: Diagrama das etapas da metodologia para previsão de tráfego com quantificação de incerteza.

Modelos para Previsão de Tráfego

Após o pré-processamento dos dados, os algoritmos selecionados serão aplicados aos dados, gerando modelos de GNNs para previsão de fluxo de tráfego. Estes modelos serão avaliados com base nas métricas definidas na Seção 3.3.

Modelos para QI

Após o treinamento dos modelos para previsão de tráfego, será necessária a quantificação da incerteza e para isso seguiremos técnicas levantadas nos trabalhos relacionados, como Monte Carlo e Regressão Linear Bayesiana, entre outras. Os modelos gerados serão avaliados com base nas métricas definidas na Seção 3.3.

3.2 Bases de Dados

A base PeMS¹, foi criada para fornecer dados sobre o desempenho do sistema de rodovias da Califórnia, e pode ser usada para análise, planejamento, operação e manutenção das estradas. A PeMS é um projeto de pesquisa colaborativo entre o Departamento de Transportes da Califórnia (Caltrans), a Universidade da Califórnia em Berkeley e outras instituições acadêmicas e parceiros da indústria.

Os conjuntos de dados são coletados pelo *Caltrans Performance Measurement System* (PeMS) [39] em tempo real, a cada 30 segundos. Os dados de tráfego são agregados em intervalos de 5 minutos a partir dos dados brutos. O sistema possui mais de 39.000 detectores implantados nas rodovias nas principais áreas metropolitanas da Califórnia. Informações geográficas sobre as estações de sensores estão registradas nos conjuntos de dados.

A base PeMSD8 contém dados de tráfego da cidade de San Bernardino, de 01 de julho a 31 de agosto de 2016, com 170 detectores em 8 estradas, com um intervalo de tempo de 5 minutos. A base PeMSD4, contém dados de velocidade do tráfego na região da Baía de São Francisco, contendo 307 sensores em 29 estradas, referente ao período de 01 de janeiro a 28 de fevereiro de 2018.

Para avaliar a eficácia de um modelo de aprendizado de máquina na previsão de tráfego, será adotada uma abordagem de validação cruzada entre conjuntos de dados. A base de dados PeMSD8 será reservada para o treinamento do modelo, enquanto a PeMSD4 será utilizada exclusivamente para testes. Essa estratégia permite uma avaliação mais robusta da capacidade de generalização do modelo, uma vez que o treinamento e o teste serão realizados em conjuntos de dados distintos, evitando vieses de sobreajuste aos dados de treinamento.

3.3 Métricas de Avaliação

O desempenho da metodologia proposta será avaliado utilizando as métricas quantitativas para previsão de fluxo de tráfego e para a QI, descritas a seguir.

¹<https://pems.dot.ca.gov>

Previsão do Fluxo de Tráfego

Para a comparação do desempenho do modelo proposto neste projeto com os encontrados na literatura, serão utilizadas as métricas comumente apresentadas nos trabalhos da área.

Mean Absolute Error (MAE):

MAE é a média das diferenças absolutas entre os valores previstos e os valores reais.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.1)$$

Root Mean Squared Error (RMSE):

RMSE é a raiz quadrada da média dos quadrados das diferenças entre os valores previstos e os valores reais.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.2)$$

Mean Absolute Percentage Error (MAPE):

MAPE é uma medida percentual da média das diferenças percentuais entre os valores previstos e os valores reais.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (3.3)$$

Symmetric Mean Absolute Percentage Error (SMAPE):

SMAPE calcula a diferença percentual média entre os valores previstos e os valores reais, levando em consideração a magnitude dos valores previstos e reais.

$$\text{SMAPE} = \frac{100\%}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{\frac{|A_t| + |F_t|}{2}} \quad (3.4)$$

No período t , temos: F_t é o valor previsto, A_t é o valor real e n é o número total de períodos.

Empirical Correlation Coefficient (CORR):

CORR mede a força e a direção da relação linear entre os valores previstos e os valores reais.

$$\text{CORR} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (3.5)$$

As métricas foram escolhidas por serem complementares. O MAE oferece uma interpretação simples, representando a média das diferenças absolutas, enquanto o RMSE penaliza erros grandes e é comumente usado em modelos lineares. O MAPE expressa os erros como uma porcentagem dos valores reais, sendo independente da escala, mas pode ter problemas com zeros nos dados. O SMAPE, por sua vez, é especialmente útil em séries temporais com valores próximos ou iguais a zero, pois lida adequadamente com essas situações, fornecendo uma medida simétrica da precisão da previsão. Por fim, o CORR mede a correlação linear entre valores reais e previstos, mas é sensível à escala e limitado a capturar relações lineares.

QI

A seguir será apresentada a métrica para QI, baseada no trabalho Qian et al. [29].

Incerteza Combinada:

As equações a seguir estimam a incerteza aleatória e epistêmica:

$$\text{Aleatória} = E_{\theta \sim p(\theta)}[\sigma_{\theta}^2] \quad (3.6)$$

$$\text{Epistêmica} = V_{\theta \sim p(\theta)}[\mu_{\theta}]. \quad (3.7)$$

A função de perda combinada é formulada como:

$$\sigma_{\text{Total}}^2 \approx \text{Aleatória} + \text{Epistêmica}. \quad (3.8)$$

As variáveis utilizadas são as seguintes:

- σ_{Total}^2 : incerteza.
- $E_{\theta \sim p(\theta)}[\sigma_{\theta}^2]$: esperança da variância σ_{θ}^2 sob a distribuição de probabilidade $p(\theta)$, que representa a incerteza aleatória.
- $V_{\theta \sim p(\theta)}[\mu_{\theta}]$: variância da média μ_{θ} sob a distribuição de probabilidade $p(\theta)$, que representa a incerteza epistêmica.

3.4 Recursos Computacionais

A implementação deste projeto será feita em linguagem de programação Python, devido ao grande número de bibliotecas disponíveis e com boa documentação. O projeto utilizará bibliotecas de aprendizado de máquina, aprendizado profundo, funções científicas e numéricas e apresentação de gráficos. Algumas bibliotecas que podem ser destacadas são: NumPy², scikit-learn³, TensorFlow⁴, Keras⁵ e Matplotlib⁶.

²<https://www.numpy.org>

³<https://scikit-learn.org>

⁴<https://www.tensorflow.org>

⁵<https://keras.io>

⁶<https://matplotlib.org>

Capítulo 4

Plano de Trabalho e Cronograma de Execução

O plano de trabalho é composto pelas seguintes atividades:

1. Obtenção dos créditos obrigatórios do mestrado.
2. Estudo e análise das principais técnicas e abordagens disponíveis na literatura.
3. Estudo e seleção de bases de dados.
4. Exame de Qualificação do Mestrado (EQM).
5. Preparação da base de dados.
6. Implementação de um *baseline*.
7. Construção dos modelos de GNNs.
8. Construção dos modelos de QI.
9. Realização de testes.
10. Comparação dos resultados com outros trabalhos disponíveis na literatura.
11. Documentação e publicação dos resultados.
12. Escrita do documento da dissertação.
13. Defesa da dissertação.

O cronograma de execução das atividades propostas, divididas em bimestres, é apresentado na Tabela 4.1.

Atividades	1º ano						2º ano					
	1	2	3	4	5	6	1	2	3	4	5	6
Obtenção dos créditos obrigatórios do mestrado	•	•	•	•	•	•						
Pesquisa bibliográfica	•	•	•	•	•	•	•	•	•	•	•	•
Estudo e seleção de bases de dados		•	•	•	•							
Exame de Qualificação do Mestrado (EQM)				•								
Preparação da base de dados				•	•	•						
Implementação de um <i>baseline</i>					•	•	•					
Construção dos modelos de GNNs							•	•	•	•		
Construção dos modelos de QI								•	•	•	•	
Realização de testes								•	•	•	•	
Comparação dos resultados com outros trabalhos									•	•	•	
Publicação dos resultados										•	•	•
Escrita da dissertação							•	•	•	•	•	•
Defesa da dissertação												•

Tabela 4.1: Cronograma de atividades dividido em bimestres.

4.1 Próximos Passos

Conforme indicado na Tabela 4.1, os próximos passos deste trabalho incluem a preparação da base de dados e a implementação de um *baseline*, para posterior comparação com os métodos propostos nesta dissertação.

Adicionalmente, pretendemos estudar a possibilidade de usar abordagens baseadas em *Transformers* [36] no contexto deste trabalho. Devido à sua capacidade de capturar dependências de longo alcance e modelar padrões complexos ao longo do tempo, os *Transformers* estão sendo usados na previsão multivariada de séries temporais [28].

Bibliografia

- [1] S. K. Ahmed, M. G. Mohammed, S. O. Abdulqadir, R. G. A. El-Kader, N. A. El-Shall, D. Chandran, M. E. U. Rehman, and K. Dhama. Road traffic accidental injuries and deaths: A neglected global health issue. *Health Science Reports*, 6(5):e1240, 2023. 1
- [2] S. Bloemheugel, J. van den Hoogen, D. Jozinović, A. Michelini, and M. Atzmueller. Graph neural networks for multivariate time series regression with application to seismic data. *International Journal of Data Science and Analytics*, 16(3):317–332, 2023. 11
- [3] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Inc., 5th edition, 2016. 6
- [4] H. Chen and H. Eldardiry. Graph time-series modeling in deep learning: A survey. *ACM Transactions on Knowledge Discovery from Data*, 18(5):1–35, 2024. 11
- [5] W. Chen, L. Tian, B. Chen, L. Dai, Z. Duan, and M. Zhou. Deep variational graph convolutional recurrent network for multivariate time series anomaly detection. In *International Conference on Machine Learning (ICML)*, pages 3621–3633, 2022. 11
- [6] S. Cheng, C. Quilodrán-Casas, S. Ouala, A. Farchi, C. Liu, P. Tandeo, R. Fablet, D. Lucor, B. Iooss, J. Brajard, D. Xiao, T. Janjic, W. Ding, Y. Guo, A. Carrassi, M. Bocquet, and R. Arcucci. Machine learning with data assimilation and uncertainty quantification for dynamical systems: A review. *IEEE/CAA Journal of Automatica Sinica*, 10(6):1361–1387, 2023. 9
- [7] K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C. W. Park, A. Choudhary, A. Agrawal, S. J. L. Billinge, E. Holm, S. P. Ong, and C. Wolverton. Recent advances and applications of deep learning methods in materials science. *npj Computational Materials*, 8(1):59, 2022. 6
- [8] S. Dahale, S. Munikoti, and B. Natarajan. A General Framework for Uncertainty Quantification via Neural SDE-RNN. *arXiv preprint arXiv:2306.01189*, 2023. 13
- [9] N. R. Dhange, J. Patil, M. Dhakate, and H. Hingnekar. Literature review for study of characteristics of traffic flow. *International Journal for Research in Applied Science & Engineering Technology*, 11(V):2398, 2023. 12
- [10] G. Dhara and R. K. Kumar. Study and analysis of visual saliency applications using graph neural networks. In *Concepts and Techniques of Graph Neural Networks*, pages 108–131. IGI Global, 2023. 10

- [11] L. Erlygin, V. Zholobov, V. Baklanova, E. Sokolovskiy, and A. Zaytsev. Uncertainty estimation for time series forecasting via Gaussian process regression surrogates. *arXiv preprint arXiv:2302.02834*, 2023. 13
- [12] P. Esling and C. Agon. Time-series data mining. *ACM Computing Surveys*, 45(12), 2012. 5
- [13] W. Ghazali, C. Zulkifli, and Z. Ponrahono. The effect of traffic congestion on quality of community life. In *International Conference on Rebuilding Place (ICRP)*, pages 759–766, 2019. 1
- [14] T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. 13
- [15] M. Jiang and Z. Liu. Traffic flow prediction based on dynamic graph spatial-temporal neural network. *Mathematics*, 11(11):2528, 2023. 1, 12
- [16] M. Jin, H. Y. Koh, Q. Wen, D. Zambon, C. Alippi, G. I. Webb, I. King, and S. Pan. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. *arXiv preprint arXiv:2307.03759*, 2023. 1, 2, 10, 11, 13
- [17] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. 10
- [18] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015. 6
- [19] M. W. E. Leen, N. H. A. Jafry, N. M. Salleh, H. J. Hwang, and N. A. Jalil. Mitigating traffic congestion in smart and sustainable cities using machine learning: A review. In *International Conference on Computational Science and Its Applications (ICCSA)*, pages 321–331. 2023. 12
- [20] G. Li, H. Deng, and H. Yang. Traffic flow prediction model based on improved variational mode decomposition and error correction. *Alexandria Engineering Journal*, 76:361–389, 2023. 1, 12
- [21] Z. Li, P. Hui, P. Zhang, J. Huang, B. Wang, L. Tian, J. Zhang, J. Gao, and X. Tang. What happens behind the scene? Towards fraud community detection in e-commerce from online to offline. In *Web Conference (WWW)*, pages 105–113, 2021. 10
- [22] W. Lin and D. W. 0044. Residential electric load forecasting via attentive transfer of graph neural networks. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2716–2722, 2021. 10
- [23] Y. Lin and R. Li. Real-time traffic accidents post-impact prediction: Based on crowd-sourcing data. *Accident Analysis & Prevention*, 145:105696, 2020. 1
- [24] Z. Lv and W. Shang. Driving the future: Leveraging digital transformation for sustainable transportation. *Green Technologies and Sustainability*, 1:100002, 2022. 2

- [25] M. Ma, P. Xie, F. Teng, B. Wang, S. Ji, J. Zhang, and T. Li. HiSTGNN: Hierarchical spatio-temporal graph neural network for weather forecasting. *Information Sciences*, 648:119580, 2023. 11
- [26] G. Meena, D. Sharma, and M. Mahrishi. Traffic prediction for intelligent transportation system using machine learning. In *3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE)*, pages 145–148. IEEE, 2020. 12
- [27] M. Miao, M. Kang, X. Qian, D. Chen, W. Wu, and W. Yu. Improving traffic time-series predictability by imputing continuous non-random missing data. *IET Intelligent Transport Systems*, 17(10):1925–1934, 2023. 1, 12
- [28] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. *arXiv preprint arXiv:2211.14730*, 2022. 20
- [29] W. Qian, D. Zhang, Y. Zhao, K. Zheng, and J. Q. Yu. Uncertainty quantification for traffic forecasting: A unified approach. In *IEEE 39th International Conference on Data Engineering (ICDE)*, pages 992–1004, 2023. 13, 18
- [30] T. E. Salais-Fierro and J. A. Saucedo Martínez. Demand forecasting for freight transport applying machine learning into the logistic distribution. *Mobile Networks and Applications*, 27:2172–2181, 2022. 1
- [31] K. Sharifani and M. Amini. Machine learning and deep learning: A review of methods and applications. *World Information Technology and Engineering Journal*, 10(07):3897–3904, 2023. 6
- [32] V. L. S. Silva, C. E. Heaney, and C. C. Pain. GAN for time series prediction, data assimilation, and uncertainty quantification. *arXiv preprint arXiv:2105.13859*, 2021. 13
- [33] N. Slimani, I. Slimani, N. Sbiti, and M. Amghar. Machine learning and statistic predictive modeling for road traffic flow. *International Journal of Traffic and Transportation Management*, 3(01):17–24, 2021. 12
- [34] M. Srinivas and P. R. C. Prasad. Modeling uncertainty quantification of ndvi of agricultural fields through bayesian linear regression in time series prediction. *Journal of Environmental Informatics Letters*, 9(1):7–15, 2023. 13
- [35] N.-P. Trinh, A.-K. N. Tran, and T.-H. Do. Traffic flow forecasting using multivariate time-series deep learning and distributed computing. In *RIVF International Conference on Computing and Communication Technologies (RIVF)*, pages 665–670. IEEE, 2022. 12
- [36] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems (NIPS)*, volume 30, 2017. 20

- [37] L. Wu, P. Cui, J. Pei, L. Zhao, and X. Guo. Graph neural networks: foundation, frontiers and applications. In *In Conference on Knowledge Discovery and Data Mining (ACM SIGKDD)*, pages 4840–4841, 2022. 10
- [38] Y. Wu, H.-N. Dai, and H. Tang. Graph neural networks for anomaly detection in industrial internet of things. *IEEE Internet of Things Journal*, 9(12):9214–9231, 2021. 10
- [39] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1):4–24, 2020. 1, 8, 11
- [40] W. Zeng, K. Wang, J. Zhou, and R. Cheng. Traffic flow prediction based on hybrid deep learning models considering missing data and multiple factors. *Sustainability*, 15(14):11092, 2023. 1, 12
- [41] J. Zhang, X. Shi, J. Xie, H. Ma, I. King, and D.-Y. Yeung. GAaN: Gated Attention Networks for Learning on Large and Spatiotemporal Graphs. *arXiv preprint arXiv:1803.07294*, 2018. 11
- [42] X.-M. Zhang, L. Liang, L. Liu, and M.-J. Tang. Graph neural networks and their current applications in bioinformatics. *Frontiers in Genetics*, 12:690049, 2021. 10
- [43] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020. 6