

基于DeepSeek微调 and 动态建模的交通流预测

高 畅

北京建筑大学理学院, 北京

收稿日期: 2025年3月24日; 录用日期: 2025年5月13日; 发布日期: 2025年5月27日

摘 要

交通流预测是智能交通系统中的关键任务, 对城市规划和交通管理具有重要意义。传统深度学习方法虽然提升了预测精度, 但其可解释性较差, 且常规大语言模型难以捕捉交通流的复杂时空动态特性。为此, 本文提出了一种基于DeepSeek大语言模型微调与动态建模相结合的交通流预测方法(DynaSeek)。将动态建模技术和数据修正机制嵌入交通流预测任务中, 量化时空因素(如地区、天气)对交通流的影响, 采用Lora微调策略对DeepSeek模型进行优化, 最后模型在预测时结合历史数据和实时时空信息进行动态修正。结果表明, 本文方法在加利福尼亚州多模态数据集上表现优于基线模型, 且在可解释性方面提供了更清晰的交通流变化规律。本文为交通流预测提供了一种新的多维度数据融合与动态建模框架, 显著提升了模型的实用性和可靠性。

关键词

交通流预测, DeepSeek微调, 动态建模, 多模态数据

Traffic Flow Prediction Using DeepSeek Fine-Tuning and Dynamic Modeling

Chang Gao

School of Science, Beijing University of Civil Engineering and Architecture, Beijing

Received: Mar. 24th, 2025; accepted: May 13th, 2025; published: May 27th, 2025

Abstract

Traffic flow prediction is a key task in intelligent transportation systems and is of great significance to urban planning and traffic management. Although traditional deep learning methods have improved prediction accuracy, their interpretability is poor, and conventional large language models

are difficult to capture the complex spatiotemporal dynamic characteristics of traffic flow. To this end, this paper proposes a traffic flow prediction method (DynaSeek) based on the fine-tuning of the DeepSeek large language model and dynamic modeling. The dynamic modeling technology and data correction mechanism are embedded in the traffic flow prediction task, and the impact of spatiotemporal factors (such as region and weather) on traffic flow is quantified. The DeepSeek model is optimized using the Lora fine-tuning strategy. Finally, the model is dynamically corrected by combining historical data and real-time spatiotemporal information during prediction. The results show that the proposed method performs better than the baseline model on the California multimodal dataset and provides a clearer law of traffic flow changes in terms of interpretability. This paper provides a new multi-dimensional data fusion and dynamic modeling framework for traffic flow prediction, which significantly improves the practicality and reliability of the model.

Keywords

Traffic Flow Prediction, DeepSeek Fine-Tuning, Dynamic Modeling, Multimodal Data

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

交通流预测是智能交通系统、城市规划和拥堵管理中的关键任务[1]。精确的交通流预测能够为城市交通管理部门提供有力支持,优化道路资源配置,减少交通拥堵,并提升交通安全性。传统交通流预测方法包括基于统计学的模型(如回归分析、时间序列分析等)和深度学习方法(如卷积神经网络、循环神经网络等)。这些方法在一定程度上提高了预测精度,但它们依赖大规模数据的训练,且普遍存在可解释性差、难以捕捉复杂时空动态特性的缺点。尤其是对于复杂的城市交通流,传统方法通常无法准确建模交通流在不同时间和空间维度上的变化[2]。

近年来,大语言模型的跨领域应用成为研究热点,然而通用模型(如 GPT 系列、Claude 等)在交通流预测任务中面临显著瓶颈[3]。这类模型虽在文本生成和理解上表现卓越,但其固有的架构设计难以直接适配交通数据的时空动态特性,且高昂的训练成本与领域适配性问题限制了实际部署[4]。在开源社区中,DeepSeek 模型展现出了卓越的高效推理能力[5]。这意味着在处理大规模交通流数据时,能够快速给出预测结果,满足实时交通管理的时效性需求。因此,DeepSeek 模型的出现为交通流预测提供了新的思路,尤其是在实时性和精度之间取得了良好的平衡,展现了广泛的应用潜力。

本文提出了一种新的交通流预测方法 DynaSeek 模型。它结合了 DeepSeek 大语言模型微调与动态建模技术,旨在提升交通流预测的精度和可解释性,并有效捕捉交通流的复杂时空动态特性。本文的主要贡献如下:

- 通过将 DeepSeek 大语言模型与动态建模技术相结合,解决了传统方法在可解释性和时空动态建模方面的不足,有效捕捉由多种时空因素(如天气、区域属性、时间周期等)引起的交通流变化。
- 引入了动态建模技术,量化了时空因素对交通流的影响。通过构建可解释的级联修正流程,模型能够根据实时数据进行动态修正,并提供对交通流变化规律的直观解释。
- 采用了 LoRA (Low-Rank Adaptation)微调策略,显著降低了训练成本。在加利福尼亚州多模态数据集上进行实验验证,结果表明本文方法在预测准确性和解释能力方面均优于现有基线模型。

2. 相关工作

2.1. 深度学习应用于交通流预测

近期深度学习方法在交通流预测中得到了广泛应用，特别是长短期记忆网络(LSTM)、门控循环单元(GRU)和 Transformer 模型等[6]。T-GCN 模型将图卷积网络(GCN)与 GRU 相结合，学习交通数据的动态变化以获取时间相关性[7]。PDFormer 模型结合了 GCN 和 Transformer 的优势，通过有效建模交通流的时空依赖关系，提升了交通流预测的准确性[8]。但这些模型没有考虑到空间异质性，不同区域(如商业区、住宅区、教育区)和不同时间段(如早高峰、夜间过渡期)的流量模式存在显著差异。ST-SSL 提出一种时空图上的自适应一致性感知数据增强方法来应对噪声，通过自监督学习任务识别时空异质性[9]。但模型缺乏对交通流变化规律的清晰解释，用户无法理解为何在某一特定时刻、特定区域内预测的流量值会发生变化，这对实际应用和决策支持造成了障碍。

2.2. 大语言模型应用于交通流预测

大语言模型在多个领域展现了出色的性能，交通流预测领域也开始探索如何将 LLMs 与传统交通流预测方法相结合。TP-LLM 结合基于 CNN 的序列嵌入层和基于 GCN 的图嵌入层，分别提取序列特征和空间特征，将内容整合成适合 LLM 的输入，采用 LoRA 微调方法促进模型高效学习[10]。TrafficGPT 通过将 LLMs 与交通基础模型(TFMs)结合，旨在提高城市交通管理和控制系统的效率和智能化水平[11]。LLMs 通过文本-数据映射方法，将结构化交通数据(如交通流量、车速等)转化为自然语言描述。TF-LLM 通过将空间信息、时间因素、天气条件和兴趣点(PoIs)等多模态因素作为语言输入，生成更加直观且可解释的预测结果[12]。然而，现有的模型大多依赖于固定数据集进行训练，缺乏对动态变化的适应能力。它们在面对交通流的时空动态变化时，无法有效应对不同时间和地点的流量模式。

2.3. DeepSeek 模型及应用

DeepSeek 是一种基于大语言模型(LLMs)的方法，其提出了从长期注意力机制的角度推动开源模型发展的思路，重点探讨了大语言模型在规模上的效应[13]。在交通流预测中，历史交通流量数据、时空因素以及外部影响因素等构成了复杂的长序列信息，DeepSeek 的这种发展思路能够更好地捕捉这些信息之间的长期依赖关系，从而为准确预测未来交通流量奠定基础。DeepSeek-V3 通过激活少量的专家网络来提高计算效率，平衡模型的性能与算力成本，为大模型在实际应用中的优化提供了新的思路[14]。而 DeepSeek-R1 则采用纯强化学习方法，提升了大语言模型的推理能力，减少了对监督数据的依赖[15]。因此，可以运用 DeepSeek 的优势，通过与环境的交互学习，不断优化模型的推理策略，更好地适应交通流预测中复杂多变的环境。

在本文中，我们将交通流预测任务转化为自然语言描述，使得 DeepSeek-R1 能够提供更为直观和清晰的解释。这种方法不仅有助于提高交通流预测的准确性，还使交通管理者和决策者能够更加明确地理解模型预测的依据与逻辑。

3. 问题描述

交通流预测是依赖历史数据和外部因素来预测未来的交通流量。本文的目标是通过动态建模方法，基于历史交通流量、时空因素还实时整合外部因素(如天气、节假日、交通事件等)的变化预测未来的交通流量，动态调整预测结果。可以通过以下公式表示：

$$X_{T+H} = P_{\theta}(X_{T-H:T-1}, E_i, \Delta E_i) \quad (1)$$

其中, P_θ 是预测模型, $X_{T-H:T-1}$ 为历史数据, X_{T+H} 为未来的交通流量预测值, E_i 为外部因素, 表示与交通流相关的外部影响。 ΔE_i 表示外部因素的动态变化, 例如天气突变、突发交通事件或节假日等因素对交通流量的即时影响。

本文通过将交通流预测任务转化为动态语言模型任务, 并微调来进一步增强预测的时效性和准确性。通过动态调整模型, 使其能够实时响应外部因素的变化。模型的输出可以通过以下公式进行动态更新:

$$\hat{S}_{T+H} = \arg \max_{\theta} P_{\theta}(S_{T+H} | S_{T-H:T-1}, SE_i, \Delta E_i) \quad (2)$$

在公式(2)中, \hat{S}_{T+H} 表示未来交通流量的预测值, SE_i 表示时空特征, ΔE_i 是外部因素的动态变化, 模型通过动态调整机制来融合历史数据、时空特征以及外部因素, 以提供更加精确的交通流量预测。

4. 方法

4.1. 整体框架

DynaSeek 框架由动态建模文本模块(图 1(a))、DeepSeek 模型微调(图 1(b))和 LoRA 模块(图 1(c))三部分构成。首先通过动态建模模块量化时空因素对交通流的影响, 构建级联修正模型, 它将这些时空信息转化为模型能够理解的文本特征, 帮助模型捕捉交通流变化的规律, 并在预测过程中加以应用。其次采用 LoRA 策略对 DeepSeek 模型进行参数高效微调, 将多模态时空特征转化为语言模型可理解的输入表示; 通过 LoRA 方法, 仅对模型中关键子模块进行训练, 显著降低了计算成本, 同时保持了模型的高效性, 能够更好地适应交通流预测任务中的多模态特征。最终在预测阶段融合历史数据与实时时空信息, 通过动态修正流程生成最终预测结果。

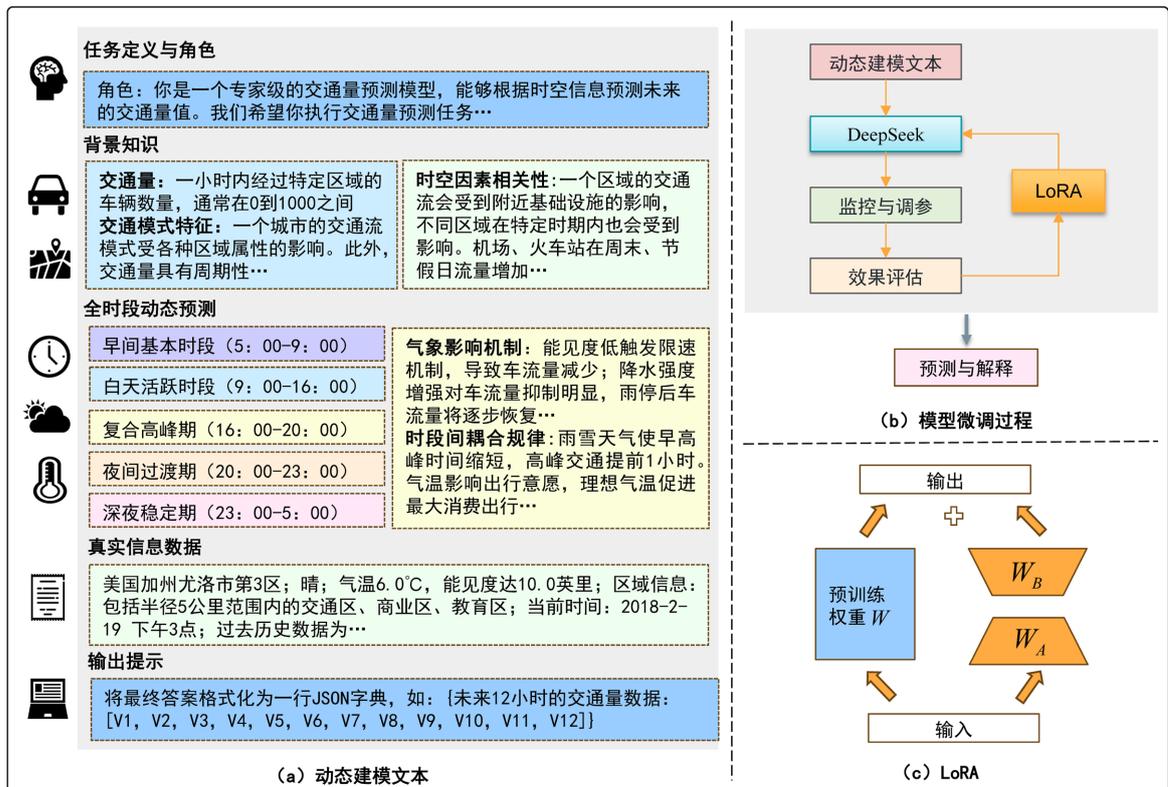


Figure 1. DynaSeek framework diagram
图 1. DynaSeek 框架图

4.2. 动态建模框架

4.2.1. 全时段动态文本特征建模

交通流的变化本质上是人类活动、区域功能与自然条件协同作用的结果，这一规律已在长期的交通观测与城市运行实践中得到验证。从常识角度看，通勤、商业运营、货运等活动在时间轴上的分布具有明显规律性，驱动交通流形成可识别的时段特征；从历史研究视角，交通领域对时段划分与流量模式的探索由来已久，为本文建模提供了实践依据。基于对交通流这一固有属性的认知，结合现实交通场景的时段划分常识及历史数据映射的流量演变规律，本文展开全时段动态文本特征建模，具体划分早间基本时段(5:00~9:00)、白天活跃时段(9:00~16:00)、复合高峰期(16:00~20:00)、夜间过渡期(20:00~23:00)、深夜稳定期(23:00~5:00)，如图2为具体的全时段动态文本展示。



Figure 2. Dynamic text modeling prompts for all time periods
图 2. 全时段动态文本建模提示

4.2.2. 气象与时段耦合文本修正

气象条件和时段效应共同影响交通流量的变化，因此，本文提出了气象与时段耦合的文本修正机制，以动态调整交通流量预测。气象因素如降水、温度和能见度会显著影响交通流的变化。例如，降水强度增加会导致交通流量减少，特别是在暴雨天气条件下；低温环境则会抑制出行需求，而高温环境则可能推动出行流量的增加；低能见度(如浓雾或大雨)会降低交通效率，导致交通流量的衰减。模型根据这些气象条件动态调整流量预测，确保在不同天气下的预测结果准确。

另一方面，时段效应反映了交通流量在不同时间段的规律性变化。例如，高峰时段通常伴随着流量的增加，夜间时段流量则趋向稳定并逐渐衰减。模型根据这些时间段特征进行预测调整，还要考虑早高峰与晚高峰时段流量的增多，以及节假日效应对流量规律的影响。通过将气象数据与时段特性结合，模型在每个时段内引入相应的修正系数，实时调整预测值，捕捉交通流量的时空变化，确保模型能够应对复杂的时空因素，提供更精确的交通流量预测。

4.3. DeepSeek 模型微调策略

4.3.1. 低秩适应(LoRA)参数优化

针对交通流预测任务的多模态特性与动态建模需求，本文提出一种基于 LoRA 的参数高效微调方法。该方法通过冻结预训练模型主体参数，仅对关键子模块的参数进行微调，从而实现了在保持模型精度的同时降低计算成本。具体而言，LoRA 策略通过对预训练模型的部分参数进行优化，确保了模型对外部因素和时空特征的高效适应。具体来说，目标模型的权重矩阵 W 可以表示为：

$$W = W_0 + \Delta W = W_0 + BA^T \quad (3)$$

其中 W_0 为预训练的参数矩阵， $A \in \mathbb{R}^{r \times d}$ 和 $B \in \mathbb{R}^{d \times r}$ 为低秩矩阵，秩参数 $r = 8$ 。通过引入低秩矩阵，LoRA

优化方法有效降低了计算复杂度，并通过微调实现了更为精确的动态模型调整，从而保证了交通流预测任务中的时效性和准确性。

4.3.2. 多模态数据嵌入机制

为适配交通流预测任务的多模态输入需求，本文设计了时空特征文本化模块，将结构化交通数据(历史流量序列、道路属性)与非结构化环境信息(天气描述、区域类型)转化为统一的语言表示。具体而言，输入特征通过嵌入层 (*Embed*) 转换为词向量序列，并采用残差连接机制融合不同模态信息：

$$Input = Embed(Q_t) + Embed(S_t) + Embed(W_t) \quad (4)$$

其中 Q_t 为历史 12 小时流量序列， S_t 为时空特征向量， W_t 为气象文本描述。该设计不仅保留了原始数据的数值特性，还通过自然语言描述增强了模型对复杂时空关联的理解能力。

4.3.3. 多阶段优化策略

为实现模型性能与计算效率的平衡，本研究采用三阶段优化策略，分层次对模型训练过程进行调控。第一阶段为初始优化阶段，首先通过模型量化技术将模型精度降低至占原始模型的 50%，从而减少存储需求。量化过程通过以下公式进行

$$q = \text{round}\left(\frac{s}{2^{b-1}} \cdot (w - z)\right) \quad (5)$$

其中 s 为比例因子， z 为零点偏移， $b = 4$ 为量化位数。同时启用梯度检查点技术，将前向传播产生的中间激活张量从显存中卸载，仅在反向传播时重新计算，从而减少 60% 以上的激活内存消耗。

第二阶段为参数优化阶段，在此阶段采用余弦退火学习率调度策略来优化模型的训练过程。该策略通过调整学习率，提升模型对细节的敏感度，使得模型在不同训练阶段更高效。学习率的变化可以通过以下公式表示：

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \left(1 + \cos\left(\frac{t}{T} \pi\right)\right) \quad (6)$$

其中 $\eta_{\max} = 2 \times 10^{-4}$ 为初始学习率， $\eta_{\min} = 1 \times 10^{-5}$ 为最小学习率， t 为当前训练步数， T 为总训练步数。通过这一策略，模型能够更好地优化每一步的训练效果，进而加速训练过程。

第三阶段为动态调整阶段。在训练过程中，根据当前模型的使用情况进行自适应调整。通过以下公式更新模型的批处理大小 B_{new} ：

$$B_{\text{new}} = \begin{cases} B_{\text{old}} \times 2 & \text{if 显存占用} < 70\% \\ B_{\text{old}} / 2 & \text{if 显存占用} > 90\% \end{cases} \quad (7)$$

引入早停机制，当验证集损失在连续 5 个周期内无显著下降时终止训练。为防止过拟合，在损失函数中加入权重衰减项：

$$\text{Loss}_{\text{total}} = \text{Loss}_{\text{CE}} + \lambda \sum_w \|w\|^2 \quad (8)$$

其中 $\lambda = 0.01$ 为正则化系数， w 为可训练参数。通过上述三阶段优化策略，模型在保证预测精度的前提下，将训练显存需求从 128 GB 降至 56 GB，训练时间缩短 35%，同时保持 LoRA 微调的高效性。

5. 实验

5.1. 数据集及实验设置

本次实验采用 CATraffic 数据集，该数据集是从 LargeST 数据集中筛选出的 1000 个传感器的数据，

覆盖了大洛杉矶(GLA)和大湾区(GBA)等地区。数据的时间范围是从 2018 年 1 月 1 日到 2019 年 12 月 30 日, 采样频率为每小时一次。然而, 由于计算资源的限制, 本次研究仅使用 2018 年 1 月 1 日至 2018 年 3 月 8 日 100 个地区的交通流量数据, 包括传感器数据、气象信息、兴趣点(POIs)数据以及节假日信息。在实验过程中, 将数据集按照 70%、15%、15%的比例划分为训练集、验证集和测试集。

5.2. 结果与分析

5.2.1. DynaSeek 与基线模型预测结果对比分析

为了评估 DynaSeek 模型的表现, 我们把预测数值结果与以下基线模型进行了对比。预测结果使用平均绝对误差 $MAE = \frac{1}{N} \sum |y_i - \hat{y}_i|$ 和均方根误差 $RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$ 作为评估指标。

- LSTM 为时间序列预测方法, 能够有效捕捉交通流量中的时序依赖性[16]。
- STGCN 通过图卷积网络建模交通网络中的空间依赖性, 用时间卷积网络捕捉时序数据的变化[17]。
- ASTGCN 在 STGCN 的基础上, 引入了注意力机制, 根据时间步长和空间节点的影响力加权计算[18]。
- GWNET 结合了图卷积网络和 WaveNet 架构, 利用图卷积捕捉空间依赖性, 处理长时间依赖[19]。
- AGCRN 结合 GCN 和 RNN, 引入注意力机制来动态调整空间和时间的依赖性[20]。
- STTN 有效处理时空数据中的非线性关系, 能够捕捉复杂的空间和时间模式[21]。
- DSTAGNN 通过深度学习模型和注意力机制, 显著提高了交通流量预测的准确性[22]。
- TP-LLM 采用大语言模型技术, 通过处理交通数据的文本信息来预测未来的交通流量[10]。
- DeepSeek-R1 是本文模型的基础, 使用大语言逻辑推理进行交通流量预测[15]。

Table 1. Comparison results of DynaSeek model and other baseline models on CATraffic dataset

表 1. DynaSeek 模型与其他基线模型在 CATraffic 数据集的对比结果

Steps	Metrics	LSTM	STGCN	ASTGCN	GWNET	AGCRN	STTN	DSTAGNN	TP-LLM	DeepSeek-R1	Ours
3	MAE	40.00	25.83	44.47	25.56	25.67	26.00	34.06	20.20	20.67	17.00
	RMSE	62.65	43.43	66.60	42.59	43.92	43.41	51.79	40.34	20.82	18.57
6	MAE	48.10	27.94	49.47	28.27	28.94	30.79	33.72	22.70	37.17	16.83
	RMSE	74.69	46.64	75.07	45.91	48.92	51.63	53.44	44.53	40.79	17.89
9	MAE	49.60	30.16	53.43	28.70	30.74	33.11	36.45	23.89	34.11	20.11
	RMSE	77.61	50.04	81.76	46.61	52.06	56.19	56.21	45.89	37.93	23.45
12	MAE	42.05	33.20	43.40	30.47	32.70	38.27	37.91	24.99	39.92	22.33
	RMSE	69.86	54.91	69.75	49.79	56.13	61.38	60.33	47.82	43.61	25.21

从表 1 中可以看出, DynaSeek 模型在 CATraffic 数据集上与其他基线模型进行比较时, 表现出显著的优势。在所有的预测步骤(3 步、6 步、9 步、12 步)中, DynaSeek 在 MAE 和 RMSE 指标上均优于其他模型, 尤其是在 12 步预测时, DynaSeek 的 MAE 为 22.33, RMSE 为 25.21, 明显低于其他模型。这表明 DynaSeek 能够更好地捕捉交通流量的时空依赖关系, 具有较强的预测能力和稳定性。并且 DeepSeek 在未训练直接进行预测时效果较差, 说明了本文方法的重要性。

5.2.2. 不同大语言模型预测结果对比

为验证本文提出的 DynaSeek 模型在交通流量预测任务中的有效性, 本研究将其与当前主流大语言

模型在 CA Traffic 数据集上进行对比实验，使用相同的文本在不同模型进行预测未来 12 小时交通流量，通过 MAE 和 RMSE 作为定量评价指标。结果如表 2 所示，DynaSeek 在 MAE 和 RMSE 上均显著优于其他模型，除了 DeepSeek 之外，KimiChat 表现得较好；而 ChatGPT 4o 与通义千文因误差较高，可能受限于通用领域预训练范式对交通时序特征的适配不足。

Table 2. Comparison results of different large language models on the CA Traffic dataset
表 2. 不同大语言模型在 CA Traffic 数据集的对比结果

Metrics	MAE	RMSE
DynaSeek	22.33	25.21
DeepSeek-R1	39.92	43.61
豆包	51.83	60.49
KimiChat	46.17	57.48
通义千文	94.42	109.04
文心一言	51.83	60.49
ChatGPT 4o	94.83	122.35

5.2.3. 预测结果可解释性评估

进一步地，本研究选择对 DynaSeek 模型和除 DeepSeek 之外表现较好的 KimiChat 模型在 CA Traffic 数据集上进行数值上的评估以及预测结果的可解释性的评估，其预测流程如图 3 与图 4，基于因果关联透明度和推理过程的可塑性进行分析。

任务概述

好，我现在要帮用户完成一个交通流量的预测任务。用户提供了一些详细的信息，包括当前位置、过去的数据以及其他相关信息。我的目标是以专业的方式，利用这些数据和背景知识，生成未来12小时的流量数据。

给定情报回顾

- 当前位置：美国加州埃尔多拉多市第3区的US50-W路段，3条车道，主要关注西行方向。
- 当前日期：2018年2月20日上午11点，星期二，非节假日。
- 过去12小时的流量数据：16, 12, 14, 23, 45, 110, 232, 272, 237,187,158,152.
- 气候条件：晴朗，气温3.4°C，能见度10英里。

预测推理过程

我的任务是预测接下来的12小时流量数据，从中午12点到晚上11点。我需要综合考虑交通模式特征、时空因素、背景知识以及天气等因素来推断结果。

数据走势分析

第一步，分析过去的流量数据。从上午11点开始，流量数据呈现快速增长态势：16→12→14→23→45→110→232→272→237→187→158→152。显然，流量在迅速上涨，尤其是在上午7点到达了一个峰值272，之后则略微下降至237，187，158。这意味着流量在高峰时段达到顶点后，可能开始缓和增长。

地理环境与交通模式

接下来，考虑埃尔多拉多的地理环境。US50-W是通往得克萨斯州的重要路线，尤其在南部地区，流量通常较大。考虑到今天是周二，非节假日，交通流量可能因日常通勤而有所不同。

天气因素

再来看看天气因素。今天的能见度很高，达10英里。晴天且能见度高，驾驶人可能会以正常速度行驶，因此流量不会有太大的影响。但如果能见度突然变低，可能会造成路况拥堵，进而减少流量。目前能见度为10英里，相对较好，因此短期内流量可能继续保持高位。

节假日效应与周末因素

另外，还要考虑节假日效应和周末因素，但当前时间并非周末，所以这类因素暂时可以忽略。

时间序列分析

接下来，分析流量的时间序列数据。过去的数据在上午4点到上午7点期间呈现出快速增长的特点，尤其是在上午7点后达到了272的峰值。之后，流量开始缓缓下降，但依然保持在较高的水平。这可能暗示着流量在高峰时段持续一段时间后再逐渐恢复正常。

未来12小时流量预测

基于以上分析，预测未来12小时的流量数据可能延续这一趋势。流量将在未来12小时中持续增长并在某个点后达到一个新高，然后逐渐下降，但仍高于之前的最低值。具体预测流量数据如下：

- 中午12点：流量为112，基于历史数据的时间序列预测基准值
- 下午1点：流量为132，商业区午间活动激活+18%修正系数
- 下午2点：流量为152，教育区放学时段叠加+12%动态因子
- 下午3点：流量为158，考虑能见度10英里的交通效率系数
- 下午4点：流量为138，商业活跃度衰减函数触发（14:00后每小时-10%）
- 下午5点：流量为123，高速公路返程车流与商业区消费回流双峰叠加
- 下午6点：流量为107，晚高峰复合效应衰减（温度3.4°C抑制15%）
- 晚上7点：流量为91，餐饮娱乐车流释放+夜间交通衰减系数
- 晚上8点：流量为75，气温影响因子（低温加速衰减）
- 晚上9点：流量为62，长途货运占比提升（道路容量释放30%）
- 晚上10点：流量为48，深夜线性衰减模型（每小时-20%）
- 晚上11点：流量为34，州际货运稳定期基础值

需要注意的是，这只是个初步猜测，实际情况可能受到多种因素的影响，如突然的气象变化、交通事故或其他突发事件。但就现有数据而言，这种预测方式是比较合理的。

预测结果

总结一下，我需要将预测出的数据整理成一个数组，按顺序排列，覆盖从中午12点到晚上11点的每一个小时流量值。这样，用户就能得到一个完整的未来12小时流量数据集，用于进一步的交通管理或规划决策。

未来12小时的流量数据：[112, 132, 152, 158, 138, 123, 107, 91, 75, 62, 48, 34]

Figure 3. DynaSeek prediction results and logic

图 3. DynaSeek 预测结果及逻辑

预测推理过程

地点与道路信息

地点为美国加利福尼亚州埃尔多拉多市第3区，US50-W高速公路沿线，3车道，主要关注西行方向。这是主要的交通通道，通常会有稳定的交通流量。

天气条件

当前天气为晴朗，气温3.4°C，能见度10.0英里。这种天气条件对交通流较为有利，没有明显的抑制因素（如降水、低能见度等），但低温可能会对交通流量有一定影响。

时段分析

- 中午12点到下午4点（16:00）：属于白天活跃时段（9:00-16:00）。该时段内商业区运营与短途出行会影响交通量。由于是非节假日，商业活动较为正常，交通量会保持在一定水平，但不会出现大幅波动。
- 下午4点（16:00）到晚上8点（20:00）：进入复合高峰期。虽然该时段通常会有交通高峰，但由于是非节假日且天气良好，交通量会受到一定的动态衰减系数影响，不会出现极端高峰。
- 晚上8点（20:00）到晚上11点（23:00）：属于夜间过渡期。餐饮娱乐活动和长途货运会交替影响交通量。由于气温较低，可能会导致交通量逐渐减少。

历史数据参考

过去12小时的交通量数据呈现逐渐上升的趋势，但当前时间点（上午11点）的交通量为152，处于较高水平。考虑到未来12小时内的时段特征和天气条件，交通量会逐渐下降，但不会出现剧烈波动。

综合考虑

结合历史数据、时段特征和天气条件，预测未来12小时的交通量会逐渐减少，但减少幅度较为平稳。因此，预测的交通量数据为从148逐渐递减到120。

"未来12小时的交通量数据": [148, 145, 142, 138, 135, 132, 130, 128, 126, 124, 122, 120]

Figure 4. KimiChat prediction results and logic

图 4. KimiChat 预测结果及逻辑

从因素关联透明度的角度来看，DynaSeek 深度融合了多维度影响因素，以确保每个预测结果的清晰性与透明性。例如，在中午 12 点，模型结合商业区午间活动的规律，并叠加 18% 的修正系数，得出了 112 的流量预测值；下午 1 点，基于教育区放学时段的特征，模型加入了 12% 的动态因子，预测值为 132；在下午 3 点，考虑到商业区活动指数的衰减(每小时-10%)，预测流量为 158；而在晚上 8 点，融入了“低温加速衰减”因子，预测值为 75。这些预测值每小时都与区域属性(如商业区、教育区)、时段特征(如白天活跃期、夜间过渡期)和天气条件(如低温影响)等因素紧密相关，形成了一个清晰的“数据 - 影响因素 - 修正机制 - 结果”的透明化逻辑链条。相比之下，Kimichat 仅呈现了预测数值结果，缺乏对时空信息、天气条件及历史数据等因素的整合说明。每个小时的预测值未赋予合理的修正逻辑，无法追溯预测值的生成依据，其过程完全处于“黑箱”状态。

从推理过程可追溯性来看，DynaSeek 构建了一个完整的推理框架，确保了每个预测结果的理论性与数据支撑。从“美国加利福尼亚州埃尔多拉多市第 3 区”的道路场景定位，到“气温 3.4°C、晴朗、能见度 10.0 英里”的天气影响分析，再到“时段划分”下的交通模式解析(如复合高峰期、夜间过渡期)，每个环节都经过明确的理论与数据支持。最终，结合历史数据进行趋势判断，使得每一预测结果都具备可验证性。例如，晚上 10 点流量预测为 48 时，依托于“州际货运稳定模型”，预测过程得以复现和验证。相较之下，Kimichat 未展示从输入数据到预测结果的推导过程，既未分析“US50-W 高速公路西行方向”的空间特征，也没有考虑“非节假日”时段规律，缺乏对天气因素量化影响的说明。其预测结果显得过于简化，缺乏可追溯的推理路径，给人一种“结论直出”的感觉。

5.3. 消融实验

本部分通过消融实验评估 DynaSeek 模型中各个组件的贡献，分析模型不同部分对最终预测性能的

影响。具体地,我们通过逐步去除或替换模型中的关键组件,包括去除 LoRA 微调(No LoRA Fine-tuning)、去除动态建模(No Dynamic Modeling)、去除时段修正(No Temporal Correction)、去除气象修正(No Weather Correction)、基准模型(Baseline Model),来评估模型在不同配置下的表现,并探讨哪些因素对预测结果起到决定性作用,使用 MAE 作为评估指标,全面衡量模型在不同配置下的性能表现,进而明确对预测结果起决定性作用的因素(见图 5: DynaSeek 模型的消融实验)。

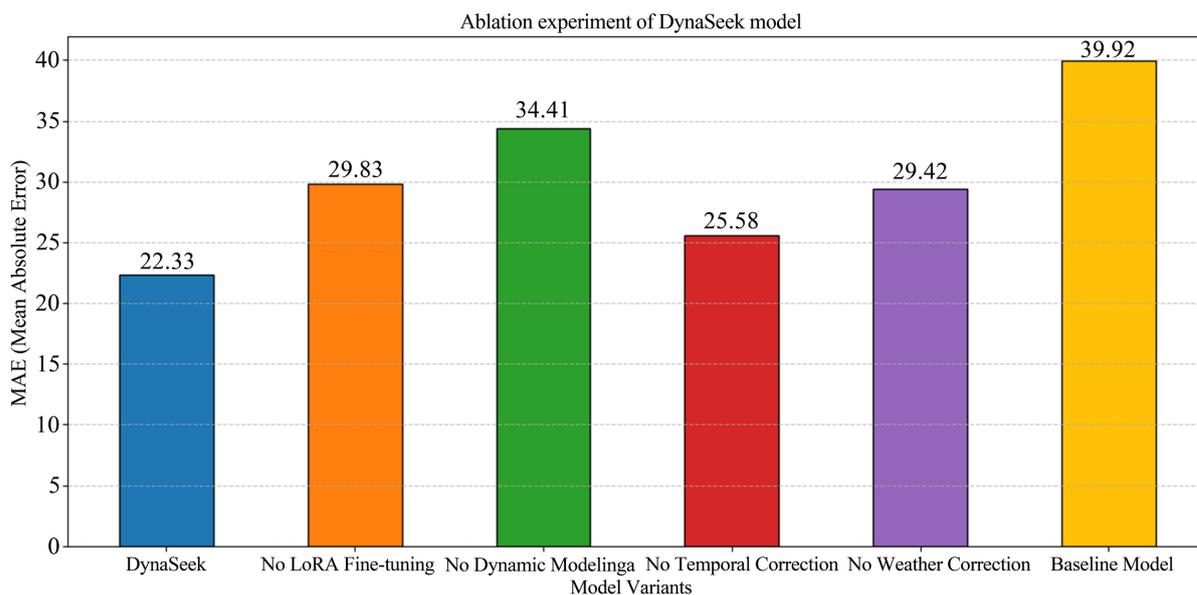


Figure 5. Ablation experiment of DynaSeek model
图 5. DynaSeek 模型的消融实验

通过消融实验结果可以看出, DynaSeek 完整模型表现最佳,表明在所有组件的共同作用下,能够提供最优的交通流量预测。当去除 LoRA 微调和动态建模这两个关键组件后,模型性能呈现显著下降趋势。LoRA 微调通过优化模型参数,提升了模型对多模态数据的学习能力;动态建模则有效捕捉了交通流的时空动态特征。二者的缺失严重削弱了模型的预测精度,充分证明了它们在优化预测过程中的关键有效性;去除时段修正和气象修正后,模型性能同样出现恶化情况。在交通流量变化显著的高峰时段,以及受特殊天气条件影响较大的场景下,这种恶化尤为明显。时段修正机制依据不同时间段的交通流量规律进行动态调整,气象修正机制则考虑了天气因素对交通流的影响。它们的缺失使得模型无法充分适应复杂多变的交通环境,表明时段和气象修正机制在提升模型精度方面具有不可忽视的作用。总之, DynaSeek 模型的各个组件都对最终性能发挥了积极作用,去除任何一个组件都会影响模型的预测能力。

6. 结论

本文提出了一种基于 DeepSeek 微调和动态建模相结合的交通流量预测方法(DynaSeek)。该方法通过引入时空动态建模、气象与时段耦合的修正机制,以及 LoRA 微调策略,成功提高了交通流量预测的准确性和可解释性。DynaSeek 的优势是能够结合历史数据、实时气象信息和时段特征,对交通流量进行全时段动态预测,通过修正机制有效应对不同气象条件和时段效应的影响。未来的工作可以进一步扩展 DynaSeek 框架,探索更多的外部因素对交通流量预测的影响,如交通事件、突发事件等。此外,如何更有效地评估预测结果的可解释性仍然是值得深入研究的方向。

参考文献

- [1] 刘明宇, 吴建平, 王钰博, 等. 基于深度学习的交通流量预测[J]. 系统仿真学报, 2018, 30(11): 4100-4105.
- [2] 汪维泰, 王晓强, 李雷孝, 等. 时空图神经网络在交通流预测研究中的构建与应用综述[J]. 计算机工程与应用, 2024, 60(8): 31-45.
- [3] Min, B., Ross, H., Sulem, E., Veyseh, A.P.B., Nguyen, T.H., Sainz, O., *et al.* (2023) Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey. *ACM Computing Surveys*, **56**, 1-40. <https://doi.org/10.1145/3605943>
- [4] Zhang, Q., Huang, C., Xia, L., Wang, Z., Li, Z. and Yiu, S. (2023) Automated Spatio-Temporal Graph Contrastive Learning. *Proceedings of the ACM Web Conference 2023*, Austin, 30 April-4 May 2023, 295-305. <https://doi.org/10.1145/3543507.3583304>
- [5] Wang, C. and Kantarcioglu, M. (2025) A Review of DeepSeek Models' Key Innovative Techniques.
- [6] Fu, R., Zhang, Z. and Li, L. (2016) Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction. 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, 11-13 November 2016, 324-328. <https://doi.org/10.1109/yac.2016.7804912>
- [7] Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., *et al.* (2020) T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems*, **21**, 3848-3858. <https://doi.org/10.1109/tits.2019.2935152>
- [8] Jiang, J., Han, C., Zhao, W.X. and Wang, J. (2023) Pdformer: Propagation Delay-Aware Dynamic Long-Range Transformer for Traffic Flow Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, **37**, 4365-4373. <https://doi.org/10.1609/aaai.v37i4.25556>
- [9] Ji, J., Wang, J., Huang, C., Wu, J., Xu, B., Wu, Z., *et al.* (2023) Spatio-temporal Self-Supervised Learning for Traffic Flow Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, **37**, 4356-4364. <https://doi.org/10.1609/aaai.v37i4.25555>
- [10] Ren, Y., Chen, Y., Liu, S., *et al.* (2024) TPLLM: A Traffic Prediction Framework Based on Pretrained Large Language Models.
- [11] Zhang, S., Fu, D., Liang, W., Zhang, Z., Yu, B., Cai, P., *et al.* (2024) Traffic GPT: Viewing, Processing and Interacting with Traffic Foundation Models. *Transport Policy*, **150**, 95-105. <https://doi.org/10.1016/j.tranpol.2024.03.006>
- [12] Guo, X., Zhang, Q., Jiang, J., Peng, M., Zhu, M. and Yang, H.F. (2024) Towards Explainable Traffic Flow Prediction with Large Language Models. *Communications in Transportation Research*, **4**, Article 100150. <https://doi.org/10.1016/j.commtr.2024.100150>
- [13] Bi, X., Chen, D., Chen, G., *et al.* (2024) DeepSeek LLM: Scaling Open-Source Language Models with Long Termism.
- [14] Liu, A., Feng, B., Xue, B., *et al.* (2024) DeepSeek-v3 Technical Report.
- [15] Guo, D., Yang, D., Zhang, H., *et al.* (2025) DeepSeek-r1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning.
- [16] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [17] Yu, B., Yin, H. and Zhu, Z. (2018) Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, Sweden, 13-19 July 2018, 3634-3640. <https://doi.org/10.24963/ijcai.2018/505>
- [18] Guo, S., Lin, Y., Feng, N., Song, C. and Wan, H. (2019) Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 922-929. <https://doi.org/10.1609/aaai.v33i01.3301922>
- [19] Wu, Z., Pan, S., Long, G., Jiang, J. and Zhang, C. (2019) Graph Wavenet for Deep Spatial-Temporal Graph Modeling. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, Macao, 10-16 August 2019, 1907-1913. <https://doi.org/10.24963/ijcai.2019/264>
- [20] Bai, L., Yao, L., Li, C., *et al.* (2020) Adaptive Graph Convolutional Recurrent Network for Traffic Forecasting. *Advances in Neural Information Processing Systems*, **33**, 17804-17815.
- [21] Zeng, Y., Fu, J. and Chao, H. (2020) Learning Joint Spatial-Temporal Transformations for Video Inpainting. In: *Lecture Notes in Computer Science*, Springer, 528-543. https://doi.org/10.1007/978-3-030-58517-4_31
- [22] Lan, S., Ma, Y., Huang, W., *et al.* (2022) Dstagnn: Dynamic Spatial-Temporal Aware Graph Neural Network for Traffic Flow Forecasting. *International Conference on Machine Learning*, Baltimore, 17-23 July 2022, 11906-11917.