摘要

本文针对单车租赁数量的预测任务，采用Transformer模型进行建模与预测。首先，对历史数据进行预处理，包括特征选择与归一化。然后，构建基于Transformer编码器的预测模型，引入位置编码以捕捉时间顺序信息，并通过多头自注意力机制提取序列特征。模型训练过程中，使用均方误差（MSE）和平均绝对误差（MAE）作为评价指标，进行五轮实验以评估模型的稳定性和性能。最后，通过对比预测结果与真实值，验证模型的有效性。

1. 问题介绍

单车租赁数量预测旨在根据历史租赁数据，预测未来特定时间段内的租赁需求。这有助于租赁公司合理调度车辆，提升服务效率。本任务要求基于过去T=96小时的数据，分别预测未来0-96小时（短期）和0-240小时（长期）的租赁数量变化曲线，并分别训练模型。

2. 模型

2.1 数据预处理

特征选择：从数据中选取相关特征，如小时、日期、天气、温度、湿度、风速及租赁数量（cnt）。

归一化：使用MinMaxScaler对特征进行归一化处理，确保数据在相同尺度上。

序列构建：将历史数据转换为固定长度的序列，每个序列长度为96小时，目标为序列末尾的租赁数量。

2.2 Transformer模型

位置编码：引入固定的位置编码，帮助模型理解时间顺序信息。

编码器层：堆叠多个编码器层，每个编码器层包含多头自注意力机制和前馈神经网络。

输出层：最终通过全连接层输出预测的租赁数量。

2.3 模型实现

3. 结果与分析

实验设置：

序列长度：96小时

特征维度：7

编码器层数量：4

前馈维度：64

多头注意力头数：8

Dropout率：0.1

批大小：32

Epochs：50

实验轮数：5

评估指标：

均方误差（MSE）

平均绝对误差（MAE）

实验结果： 通过五轮实验，计算MSE和MAE的平均值及标准差，评估模型的预测性能和稳定性。