

摘要：随着私家车数量的日益增长，交通堵塞成为了一个亟待解决的问题。本项目实现了利用过往的路段交通流量数据，较精准地预测未来特定时段的交通流量数据，对交通基础设施建设有指导性的作用。

根据 OpenTIS 平台获得安徽宣城 7 月和部分 8 月的交通流量的车道级线圈数据和浮动车 GPS 数据。本组进行了数据预处理，针对线圈数据建立了**考虑时空的 KNN 模型**和基于**路段车辆流量拓扑图的 GCN&LSTM** 深度学习方法，并设置了四个早起实验尝试的对照组以论证前两种方法的创新性和优越性。在浮动车部分，通过 **K-means++**、**整数规划**等算法，并进行数学建模，实现了从浮动车数据到路段整体流量数据的跨越，并将此数据来优化线圈测得的路段数据。

针对线圈数据建模的部分，本组采用 KNN 对问题进行求解，选取欧式距离的倒数衡量交通流的相似度，确定数据间的权重。进而引入时空相似性，建立时空矩阵并采用矩阵 F-范数的倒数作为确定权重的指标。最终训练集 MAPE 结果由 12%降低至 9%左右，效果良好，MAE:4.72，RMSE:6.24。

采用基于路段车辆流量拓扑图的 GCN+LSTM 的模型，本组使用**动态时间规整 DTW** 定义了车辆间流量的相似性，进而建立了路段间车辆流量相关矩阵和拓扑图，再而引入至 GCN+LSTM 进行模型训练，训练集 MAPE:11.55%，MAE:8.43，RMSE:17.55。

引入了四个对照组分别为注意力机制下的 BiLSTM，小波去噪优化后 LSTM，考虑时空相关性后的 LSTM 和结合 CNN 的 LSTM。针对 **BiLSTM** 方法本组进入了 Attention 机制，最终效果 MAPE:14.32%，MAE:8.37，RMSE:11.1。针对引入小波去噪优化后 LSTM，本组得到低频部分结果 MAPE:8.14%，MAE:4.66，RMSE:6.53，高频部分拟合为 MAPE:14.53%，MAE:8.04，RMSE:10.4。针对**考虑时空相关性后的 LSTM**，本组首先证明了空间存在相关性，进而构造了相关系数矩阵。针对 **CNN+注意力 BiLSTM 模型**，本组得到结果 MAPE:11.97%，MAE:6.76，RMSE:8.85。

对于浮动车部分，本组对数据进行预处理和可视化，采用了 K-Means++ 聚类对速度-密度散点图进行处理，绘制了散点肘部图，并最终确定了迭代后的聚类中心。进一步，本组采用**格林希尔兹线性模型**和**格林柏对数模型**对数据进行了拟合。建立模型对交通流密度和车距关系进行分析和预测并运用**整数规划**进行求解，最终预测交通流的流量，是线圈数据的有力补充和佐证。

最终，由于赛题数据限制，本组选取 KNN 和基于路段车辆流量拓扑图的 GCN&LSTM 的两种方式参与了作品数据预提交，并以 GCN&LSTM 的结果作为最终结果。MAPE:24.63%，MAE:8.88，RMSE:11.24。本组进而对结果进行分析，对比了各模型及其优缺点，对本项目未来的在交通运输中应用前景进行了展望。

关键词：交通流预测；时空 KNN；深度学习；图神经网络

