# 智能交通背景下出租车载客收益预测案例教学指导手册

**教学适用的课程：**《大数据架构与技术》

**教学适用的专业：**电子信息

**教学目的与用途：**学习实际智能交通场景下的大数据应用；掌握大数据预处理、数据分析相关知识点；掌握大数据处理与分析的基本流程。

**教学内容：**

（1）理论依据：本案例基于大数据技术对智能交通中出租车载客收益预测。案例整合了多种的数据挖掘与深度学习方法，进行数据预处理与基于多种深度学习算法的出租车收益预测，可实现大数据依托的智能交通决策，促进传统交通行业的智能化改进。

（2）涉及知识点：数据预处理、时序数据预测、注意力机制、时间注意力、空间注意力、长短期记忆网络、Seq2Seq模型。

（3）分析路径：首先介绍智能交通与智联网汽车感知的背景，并进行数据集展示，介绍相关字段含义。对数据集进行预处理，关键点在于如何将出租车的轨迹信息与载客信息转化为历史载客收益图。然后基于处理后的数据进行出租车的载客收益预测，主要介绍基于双注意力机制的RNN算法，使学生充分理解时间注意力与空间注意力的含义与作用，另外了解其他经典深度学习算法，包括LSTM、Seq2Seq、以及TAtt。最后，使用MAE指标作为评估标准，对多个算法进行训练与测试，挑选最佳算法，并可进行进一步地调优。

**启发思考题：**

(1) 系统实现：本案例仅仅是停留在数据与算法层面，进一步可以基于该预测算法对预测城市实现原型系统，根据城市出租车的实时GPS数据与载客数据，将预测的载客收益进行实时显示，以便于出租车公司进行调控以提高收益。

(2) 对任务的进一步拓展：考虑对算法进一步拓展，基于算法的载客收益预测结果，实现出租车资源的实时最优分配，以提高交通效率，并可进行高效的智联网感知任务。当然，最优分配问题往往是一种NP-hard 问题，具有指数级时间复杂度，因此需要设计具备可接受时间复杂度的算法寻找较优的分配方案。

**建议课堂计划：**

1. 时间安排：3课时
2. 学生学习准备：了解常见的数据预处理、时序数据预测、深度学习算法的理论知识。
3. 分组及讨论内容：分组讨论启发思考题。
4. 案例开场白：智能交通是大数据技术的一个重要应用场景。随着智能汽车、辅助/无人驾驶汽车、电动汽车的飞速发展和爆炸式普及，当前汽车集成越来越多的传感器，拥有越来越强大的感知、计算和通信能力。近年来，涌现很多基于智联网汽车感知的实际应用。这节课我们将接触一个大数据在智能交通中的实际应用。
5. 结束总结：这节课我们以出租车载客收益预测为例，接触到数据清洗与预处理、时序数据预测、循环神经网络等相关知识，对大数据在智能交通场景中应用有了初步的接触与实践。需要注意的是，课程所学内容是有限的，希望对该领域感兴趣的同学可以自行搜集相关资料，进一步了解大数据在智能交通中的其它应用及相关技术。
6. 案例引导建议：在教师简要分析应用场景后，鼓励学生自主考虑如何使用所学知识解决该场景问题。

**参考文献：**

[1] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT press, 2016.

参考内容：深度学习模型。

**其他教学支持材料：**

1. 本案例提供配套的PPT、视频、数据集与代码等，发布于Github，链接为：https://github.com/Wanghui-Huang/CQU\_bigdata。

2. 本案例涉及到数据预处理以及多种机器学习算法，建议使用python语言进行编写，推荐的工具包有pandas（数据读取与预处理库），Keras与TensorFlow（深度学习算法库），matplotlib（可视化绘图库）。

3. 由于数据版权原因，本案例仅提供重庆出租车轨迹数据集的部分示例数据。出于教学与实践需要，本案例给出以下公开出租车轨迹数据集，以供学生选择使用：

（1）T-Drive，包含北京市2008年02月02日至02月08日共一周时间内10357辆出租车的轨迹数据，由微软提供，下载链接：https://www.microsoft.com/en-us/research/publication/t-drive-trajectory-data-sample/

（2）The epfl/mobility dataset，包含美国旧金山湾区2008年05月17日至2008年06月10日大约500辆出租车的GPS轨迹数据，下载链接：http://crawdad.org/epfl/mobility/20090224/

（3）Smart City Research Group，包括上海市出租车、公交车数据等，来源于香港科技大学，下载链接：https://cse.hkust.edu.hk/scrg/

（4）中华人民共和国交通运输部交通指数，中华人民共和国交通运输部下属交通数据分享平台，具有全国多个城市的出租车GPS轨迹数据，平台网址：https://www.mot.gov.cn/sjkf/