**2022 ━ 2023学年 交通数据分析与应用 期末大作业**

**请基于提供的主车运行轨迹数据和周边车辆运行轨迹数据，构建主车变道动机识别模型，并对变道行为的风险程度进行分级，并回答以下问题：**

**（1）请简述构建的模型框架，并阐述建模方法的采用原因；**

**（2）请列出模型拟合精度的评价方法，并对所构建模型的缺陷进行分析。**

注1：可以将变道动机识别、变道行为风险程度评估分开建模，也可构建一体化的三分类模型（即输出无变道动机、有变道动机但低风险、有变道动机且高风险）；

注2：本次作业不限定建模软件平台。

附录：数据说明

* 数据源于德国高速公路自然驾驶轨迹HighD数据集，包含3个数据文件，分别为x\_all\_lc.npy、x\_all\_lk.npy、risk\_label.csv，其中x\_all\_lc.npy记录847条车辆的变道轨迹，risk\_label.csv 为847条变道轨迹的风险标签（其中0表示低风险，1表示高风险），x\_all\_lk.npy为对比数据，记录了847条车辆保持车道内行驶的轨迹；
* x\_all\_lc.npy、x\_all\_lk.npy具有相同的数据记录格式，以x\_all\_lc.npy为例展开说明。图1为一个lane change轨迹示例，大小为50×42，其中行代表时间戳，采样频率为25hz；t表示本车变道开始时刻，t-1为变道开始前的上一个采样时刻，依此类推，由上至下分别记录开始变道至开始变道前2秒轨迹；列代表特征变量，为别按照本车、前车、左前车、右前车、后车、左后车、右后车顺序记录了7个车辆的x轴与y轴坐标、速度、加速度，特征变量图表1所示。车辆位置关系以及x轴、y轴方向如图2所示，其中坐标为绝对坐标，左上角为坐标原点（当相对位置上未检测到车辆数据，则以nan值进行填充）；
* risk\_label.csv格式如图3所示，其中样本id与x\_all\_lc.npy的轨迹id相匹配，高风险判断标准为紧急制动下最终车间距离（Potential index for collision with urgent deceleration）（PICUD）<0，表达式如下：

（1）

其中 、 分别表示主车（也称为本车）与后车速度, 为本车与后车的车辆间距, 是驾驶人反应时间, 为最大制动加速度。一共包含高风险样本数为54个，低风险为793个。

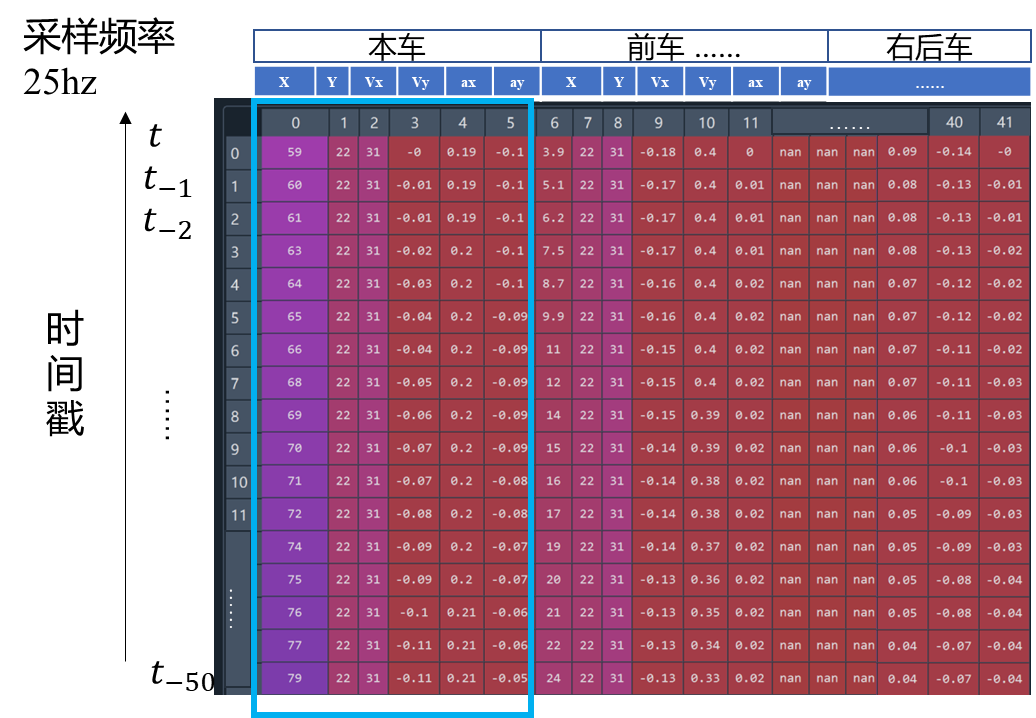


图1变道轨迹样本示例

表1各个特征变量参数含义与单位

|  |  |
| --- | --- |
| Extracted features | Definition |
|  | 车辆前进方向坐标(m) |
|  | 车辆横向坐标(m) |
|  | 车辆前进方向速度 (m/s) |
|  | 车辆水平方向速度 (m/s) |
|  | 车辆前进方向加速度 (m/s) |
|  | 车辆水平方向加速度 (m/s) |

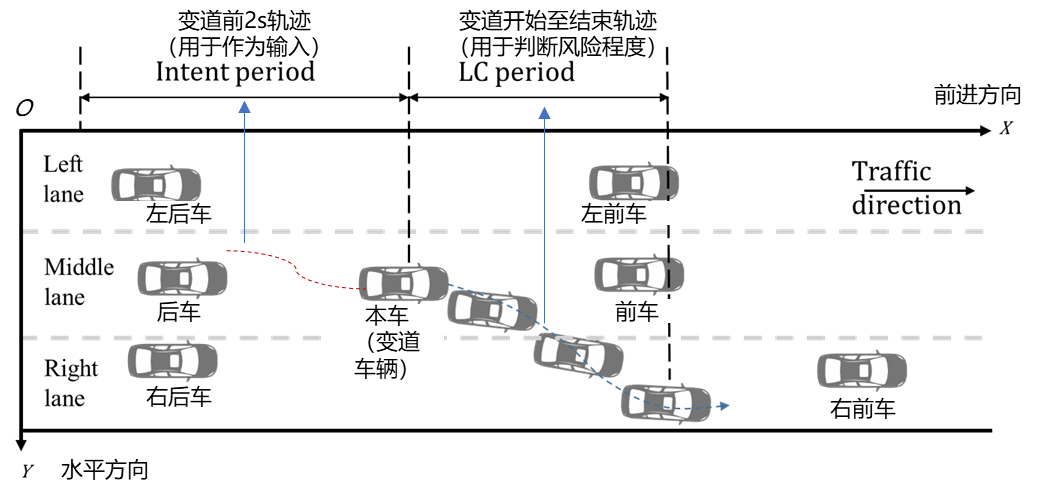


图2 车辆位置关系、运行方向示意图

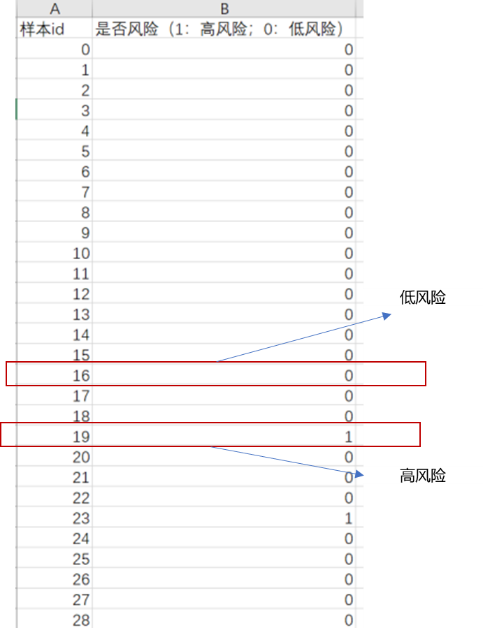


图3 风险标签示意图