



数据说明 Data descri...



乘用车细分市场销量预测

主办方：中国计算机学会 & 深瞳云涂

预测回归

数据挖掘

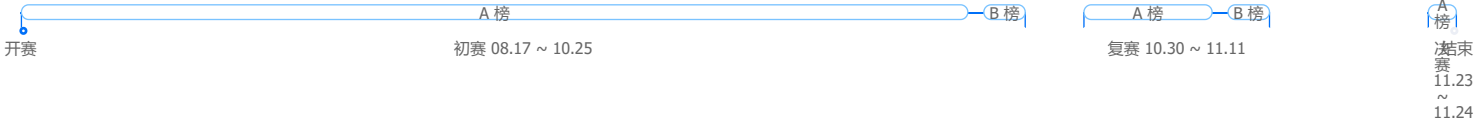
队伍 / 人数

奖励

【2019/09/24 15:00:12】DF1569308376081参加了乘用车细...

1943 / 2264

¥ 100,000



2019/08/23 16:30:50 赛题数据

待预测测试集

(报名后可下载数据) 下载

2019/08/23 16:31:10 赛题数据

提交样例

(报名后可下载数据) 下载

2019/08/23 16:31:29 赛题数据

训练集

(报名后可下载数据) 下载

数据说明 Data description

历史销量数据包含60个车型在22个省份，从2016年1月至2017年12月的销量。参赛队伍需要预测接下来4个月（2018年1月至2018年4月），这60个车型在22个省份的销量；参赛队伍需自行划分训练集数据进行建模。

历史销量数据包含60个车型在22个省份，从2016年1月至2017年12月的销量。参赛队伍需要预测接下来4个月（2018年1月至2018年4月），这60个车型在22个省份的销量；参赛队伍需自行划分训练集数据进行建模。

1.[训练集]历史销量数据：train_sales_data_v1.csv

[Training Set] Historical sales data: train_sales_data_v1.csv

字段名称 Field name	字段类型 Field type	字段说明 Field description
province	String	省份
adcode	int	省份编码
model	String	车型编码
bodyType	String	车身类型
regYear	int	年
regMonth	int	月
salesVolume	int	销量

2.[训练集]车型搜索数据：train_search_data_v1.csv

[Training Set] Vehicle Type Search Data: train_search_data_v1.csv

字段名称	字段类型	字段说明
province	String	省份
adcode	int	省份编码
model	String	车型编码
regYear	int	年

数据说明 Data descri...

提交要求

提交示例 Submissio...

评分标准 Scoring crit...

popularity	int	搜索量
------------	-----	-----

3.[训练集]汽车垂直媒体新闻评论数据和车型评论数据：train_user_reply_data_v1.csv

[Training Set] Vehicle Vertical Media News Review Data and Vehicle Model Review Data: train_user_reply_data_v1.csv

该数据集包含了垂直媒体中，各车型的每月（不分地域）每月新闻评论数据、车型下的评论数据两部分，这两个数据没有任何包含关系。

The data set contains two parts: the monthly news review data of each vehicle type (regardless of region) and the review data under the vehicle type in the vertical media. The two data sets have no relationship with each other.

字段名称	字段类型	字段说明
model	String	车型编码
regYear	int	年
regMonth	int	月
newsReplyVolum	int	对车型相关新闻文章的评论数量
carCommentVolum	int	对车型的评价数量

4.[评测集]2018年1月至4月的各车型各省份销量预测：evaluation_public.csv

[Evaluation Set] Sales Forecast for each vehicle type in each province from January to April 2018: evaluation_public.csv

字段名称	字段类型	字段说明
id	int	数据的唯一标识，不可更改
province	String	省份
adcode	int	省份编码改
model	String	车型编码
regYear	int	年
regMonth	int	月
forecastVolum	int	预测销量，参赛队伍使用建立的模型得出的销量预测结果 Forecast sales volume—the results of sales forecasting obtained by the team using the established model

(注：以提交样例为准)

提交要求

[初赛]：以csv格式进行提交包含预测结果的“evaluation_public.csv”文件，每支队伍每天最多可提交三次，在线评分。其中A榜实时评测，B榜定时评测。

[Preliminary]: The “evaluation_public.csv” file containing the forecast results is submitted in CSV format. Each team can submit up to three times a day; Scoring will be done online. Among them, A-list evaluation will be performed in real time and B-list evaluation will be performed at a given timing.

[复赛]：以csv格式进行提交包含预测结果的“evaluation_public.csv”文件，每支队伍每天最多可提交三次，在线评分。其中A榜实时评测，B榜定时评测。

[Semi-final]: The “evaluation_public.csv” file containing the forecast results is submitted in CSV format. Each team can submit up to three times a day and Scoring will be done online. Among them, A-list evaluation will be performed in real time and and B-list evaluation will be performed at a given time.

[决赛]：各队需提交方案文档、源码、程序部署运行文档，并进行现场答辩和专家评审。为保证赛题作答真实有效，方便企业及选手之间进行测评，所提供源码必须保证按照选手提供的文档，可部署运行并通过接口方式进行结果输出。赛题给出多个车型销售数据，选手在答题过程中需要考虑是否可以使用同一套算法模型进行结果输出。各参赛选手完成赛题的准确度、完整度，以及所提交的模型的适应性、可扩展性、代码的工程性将影响最终答辩结果。

[Final]: Each team is required to submit project documents, source code, program deployment and operation documents, and conduct on-site debate and receive expert review. In order to ensure the authenticity and validity of questions answering and facilitate the