**tar包说明：**

alge-services.tar 线上服务项目源码包

dwd-alge-services.tar 星堂线上业务逻辑源码包

dwd-xingtang.tar 星堂算法平台源码包

model.tar 离线训练源码包

**【一】转单模型**

**相关对接人：**龙章潮、吴纲维、喻钢、徐然琪、韩硕

**概述**

转单是骑手对派单合理性的重要反馈行为，但不同骑手对同一个订单的偏好程度是不一样的，这和骑手的地域喜好、商家喜好、个人认知度、配送能力、当前状态等因素息息相关。为了提升骑手体验，算法组推出转单模型模型，模型在派单找人时判断骑手是否接受系统派单。当骑手接单概率大时才会派给该骑手，故可以降低骑手转单率能够极大提高骑手的用户体验、接单率和对平台的忠诚度，并能提高对渠道的服务质量。

**离线代码**

位于model项目的model-switch目录下，各文件功能见Readme.txt

**线上代码：**

dwd-alge-services/alge-xingtang-services/src/main/java/com/dianwoda/alge/xingtang/services文件夹。服务代码主文件位于其中的server/ tensorflow/ SwitchPredFeatBuild.java，线上特征工程代码为其中的util/ tensorflow/ SwitchPredFeatBuild.java。

**【二】ETA模型**

**相关对接人：**吴纲维、龙章潮、张其来、张建伟

**概述**

在调度系统超时过滤模块中，需要预估骑手身上各个订单的送达时间，以此作为超时过滤的依据。针对具体的场景，我们用端到端的多任务学习方式同时预估骑手配送顺序和各用户点的送达时间，合理地将骑手的个性化因素嵌入网络结构中，利用self-attention捕捉订单之间的相互关系。

**离线代码**

位于model-eta下, 各文件功能见Readme.txt

**线上代码：**

alge-services包中的alge-services/services-pdst目录下，接口类为:

com.dianwoba.bigdata.services.pdst.provider.EtaServiceImpl

**【三】骑手时效考核模型**

**相关对接人：**吴纲维、张开慧

**概述**

平台对骑手的考核时间点：pdst = fft + 动态调整 + 人工配置修正。

fft是基于骑手实际配送时间预测的时间点，之前的版本是直接取分位数。

饿了么渠道对点我达的时效考核从“参考询T值”改为“饿了么用户T”之后，渠道考核时间 和 骑手考核时间 二者关联性很低，（可能渠道考核时间很紧，但是点我达考核骑手很松），导致点我达的渠道考核超时率飙升，严重影响平台考核收益和超时扣罚。因此，需要对照新的渠道考核规则，设计一套新的骑手考核和奖励规则，以便使得“骑手考核标准”和“渠道考核标准”对应统一。

新的pdst计算逻辑为：pdst=max{最终渠道考核时间，最终兜底时间}

其中，最终兜底时间基于LightGBM分位数回归计算。

**离线代码**

位于model-pdst目录下, 各文件功能见Readme.txt

**线上代码：**

alge-services包中的alge-services/services-pdst目录下，接口类为:

com.dianwoba.bigdata.services.pdst.provider.PdstServiceImpl

**【四】外卖/快递地址命名实体识别**

**相关对接人：**龙章潮、张其来、吴纲维、张建伟

**概述**

本项目为一个基础的地址解析服务项目，从外卖订单地址中，抽取这个地址实体（小区的名字，医院的名字、大学的名字，商场的名字等等）、楼栋、单元，楼层。主要用于这个基于地址实体进行降付费统计，对好配送的地址实体的订单，降低配送费，难配送困难的订单，增加配送费。也用于风控反作弊、商家一键发单、快递订单分组。

**离线代码**

位于model-ner下

**线上代码：**

alge-services包中的alge-services/services-community目录下，接口类为:

com.dianwoba.bigdata.services.community.provider.CommunityServiceImpl

**【五】召回模型**

**相关对接人：**喻钢、白森、韩硕

**概述**

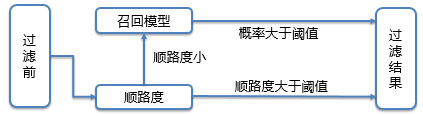
召回模型主要解决订单顺度度问题（通过召回可能顺路的订单来提高系统去化效率）；顺路度订单是分两步决策的：首先过滤掉骑手一定不喜欢的订单，再从骑手可能喜欢的订单中召回顺路度的订单，提高订单去化。

同时在上一版模型上，线下有骑手反馈，我们召回的模型存在一些骑手不喜欢的订单（哪怕是100单中有几单不喜欢的订单，骑手也会有抱怨）。这类问题主要是由于召回订单中存在一小部分顺路度值较低的订单，这类订单对于骑手体验来说极差，所以需要调整顺路度过滤器的判定逻辑，将这部分影响降到最低。

原有逻辑：



调整后逻辑：



**代码**

离线训练代码位于model-recall下的new\_data\_and train\_v3 文件夹。



描述：模型采用骑手app端抢单数据作为训练集，当骑手发生点击抢单行为时，被抢单被标记为1（被抢），订单池中其他订单标记为0（未抢）。模型训练数据、模型与推荐单排序一致，但召回模型训练时将全局顺路度小于0.4的抢单标签改为不抢单。

**线上代码：**

dwd-alge-services/alge-xingtang-services/src/main/java/com/dianwoda/alge/xingtang/services文件夹下，服务代码主文件位于其中的server/ tensorflow/ RecallTensorflowService.java，线上特征工程代码为其中的util/ tensorflow/RecallPredFeatBuildV3.java。

**【五】推荐排序模型**

**相关对接人：**喻钢、白森、张建伟

**概述：**

在正常天气下，调度系统抢单比例在15%左右。当恶劣天气或其它极端情况发生时，抢单比例会飙升至40%~50%。因此，推荐单排序对于改善运力不足情况下的骑手体验，从而加强订单去化有着较大价值。然而，当前基于规则的智能排序策略难以最大程度上解释和表示骑手根据环境变化产生的个性化兴趣与体验，因此，系统提供的智能排序策略需要通过能预测不可知/不确定性较好的模型算法来实现，以提高骑手抢单体验，优化订单去化速度。从业务场景上看，推荐单排序与传统的广告推荐和CTR点击预测是一类问题。

**代码：**

离线训练代码位于model-recall下的new\_data\_and train\_v3 文件夹。

描述：模型采用骑手app端抢单数据作为训练集，当骑手发生点击抢单行为时，被抢单被标记为1（被抢），订单池中其他订单标记为0（未抢）。



**线上代码：**

dwd-alge-services/alge-xingtang-services/src/main/java/com/dianwoda/alge/xingtang/services文件夹下，服务代码主文件位于其中的server/ tensorflow/RankTensorflowService.java，线上特征工程代码为其中的util/ tensorflow/RankPredFeatBuildV1.java。

**【七】骑手招募画像模型**

**相关对接人：**张其来

**概述：**

1. \*\*需求\*\*

1. 对双十一触达骑手进行画像分层，分为高、中、低潜骑手

2. 对于不同层骑手，采取不同的触达策略，达到精细化触达的目的

2. \*\*模型：\*\* Stacking 模型融合【树模型、DNN、LR】

1. 数据从预处理，包括空值填充、归一化、数值变换

2. 基于重要性【熵】对特征排序，筛选特征

3. 基于特征，进行模型的开发

3. \*\*结果\*\*

1. low: 0.9561, mid: 0.0292, high: 0.0147, auc: 0.8988

2. 模型每日预测，到触达前，基于一段时间的预测结果，进行画像分层

3. 评价困难，评价指标定义困难

**代码：**

63服务器：/data1/inst/zql/rider\_p

**【八】商家出餐时间预测模型**

**相关对接人：**吴纲维、龙章潮、张建伟

**概述：**预测每个饿了么订单的出餐时间预测，基于Tensorflow开发深度学习模型，目前模型应用于调度系统路径规划、等餐久过滤、付费、等餐报备等项目中。

**离线代码：**

位于model-pft目录下，各文件功能见Readme.txt

**线上代码：**

alge-services包中的alge-services/services-kael目录下，接口类为:

com.dianwoba.bigdata.services.pdst.provider.ShopServingPredictImpl

**【九】知识库输入联想模型开发**

**相关对接人：**张其来、张建伟

**概述：**

1. \*\*需求：\*\* 便于客服输入、降低使用门槛、优化搜索性能

1. 知识库的搜索输入需要较为精准的输入；提升搜索的准确性、使用的便捷性

2. 使用门槛高，对新手不友好

2. \*\*模型：\*\* 输入联想模型【数据：用户输入的词+标准关键词】

1. 自动补全：当用户输入部分标准关键词情况下，自动补全后续可能的关键词【Trie树查找】

2. 同音纠错：对于可能存在的同音错字的情况下，给出同音正确的关键词【同音映射】

3. 常用词联想：对于用户经常输入的词，给出其可能的候选关键词【词频统计】

4. 同义词联想：对于同义词，给出其可能的候选关键词【词向量的相似性计算】【待优化】

3. \*\*结果\*\*

1. 模型开发完成，Top5 准确率 57%，建伟线上服务开发基本完成

**在线代码**

dwd-alge-services\alge-xingtang-services\src\main\java\com\dianwoda\alge\xingtang\services\server\selfdefine

**【十】骑手社区情感分析**

**相关对接人：**张其来

**概述：**

1. \*\*需求\*\*

1. 骑手社区目前的舆情反馈总量比整体外部平台舆情更多，及时掌握骑手心声，通过骑手舆情来及时诊断问题、调整策略，是保证业务发展、有序迭代、走在行业前列的基础，在业务、策略、产品的发展中形成共识与合力，推动生态长期健康发展。

2. 获取外部常用骑手评价言论，基于竟对骑手言论，实现骑手情感引导和调控

3. 希望算法组能够支持对每一条帖子进行语义分析，判断帖子的“正面”、“负面”、“中立”的性质。基于对骑手言论，实现骑手情感分析

2. \*\*模型\*\* stacking 模型

1. 基于Bert的模型

2. 基于字典模型

3. \*\*结果\*\*

1. 模型开发完成，准确度达到87%左右

**代码：**

位于rider\_sentiment\_analysis文件夹

**【十一】骑手骑行速度预测**

**相关对接人：**吴纲维、张开慧

**概述：**基于骑手历史轨迹数据和贝叶斯推断计算骑手平均骑行速度，利用骑手准实时轨迹数据更新骑手骑行速度，用于路径规划决策、超时过滤、等餐久过滤等项目中。

**离线代码：**位于model-speed目录下, 各文件功能见Readme.txt

**线上代码：**

alge-services包中的alge-services/services-riderspeed目录下，接口类为:

com.dianwoba.bigdata.services.riderspeed.provider.RiderSpeedServiceImpl

**【十二】骑手上楼时间预测**

**相关对接人：**吴纲维、张开慧

**概述：**基于骑手历史轨迹数据预测骑手上楼时间，用于路径规划决策、付费等项目中。

**离线代码：**位于model-upstair目录下, 各文件功能见Readme.txt

**线上代码：**

alge-services包中的alge-services/services-scylladb目录下，接口类为:

com.dianwoba.bigdata.services.syclladb.provider.PathPlanServiceImpl

**【13】在线热力图预测**

**相关对接人：**张建伟、张开慧

**概述：**每5分钟预测一次15分钟后H3网格的订单量情况

**离线代码：**

model包中位于model-dispatch-pressure目录下, 各文件功能见Readme.txt

**线上代码：**

alge-services包中的alge-services/services-order-pressure目录下，接口类为:

com.dianwoba.bigdata.services.orderpressure.provider

**【14】星堂算法平台**

**相关对接人：**张建伟、张开慧

**概述：**已迁移到星堂平台的在线项目都在alge工程下，该工程下的代码在星堂平台上执行产生自动服务。

**线上代码：** 在dwd-xingtang项目下，common是基础模块，core是核心模块，xingtang-server-config是配置模块