姓 名 ：杨新权

电 话 ：18811171651

邮 箱： starspringcloud@gmail.com

出生年月：1988.05

学 历 ：硕士

住 址：北京市

基本信息

电子科技大学 2010/08-2013/06 硕士 / 模式识别

福建师范大学（一本） 2006/09-2010/07 本科 / 电子信息工程

教育经历

工作经历

**2015.03-至今 阿里集团-高德-搜索 推荐算法专家**

1. 垂类搜索：本地生活（酒店、美食、生活服务等）ctr/cvr建模，用户query为类别词，模型选择偏向推荐系统，侧重用户个性化建模
2. 地图搜索：高德的主搜，用户query有明确的poi意图。负责过搜索词自动补全(suggest)， 和基础NLP以及品牌搜

**2013.07 – 2015.3 奇虎360 算法工程师**

1． NLP任务： 利用基于n-gram的分词算法对短信内容进行分词，使用CRF 进行NER任务。Fasttext进行短信意图分类。

专业技能

1. 熟悉垂类搜索/推荐系统：粗排、精排、重排序； 多行业多任务建模、多样性、冷启动
2. 搜索相关经验：learning to rank，Text Classification， Semantic matching，NER
3. 熟悉tensorflow框架和大数据开发。有良好编程功底Python、c++ 、c、java、hive

项目经验

**2018.05-至今 高德 垂类搜索**

1. 多路召回：geohash+term召回、向量召回、u2Tag2i、冷启动召回
2. 粗排：基于树模型的ctcvr任务，用户画像，商品属性，时空特征之间进行特征工程和特征交叉
3. 精排（rank）：从wide&deep到使用deepfm来学习特征的自动交互、再到使用multi-head attention对用户长期/实时序列学习LastN信息，并使用sim模型的hard-search思路筛选attention候选集。
4. 重排(rerank)：精排只注重item的排序指标，缺乏整个list的视野，做不到整体最优化。PRM模型把精排的输出作为transformer encoder模块的输入，softmax CE作为loss，同时增加辅助loss（多样性）
5. 多行业多任务建模：高德本地生活有多个业务（酒店、美食等），分别维护一套模型成本巨大，且部分场景数据稀少，不利于建模。此外同一个业务也有这多种任务（ctr、cvr）。在综合PLE、ESMM、HiNet等模型优点的基础上，实现了多行业多任务的统一rank模型。上线取得良好的正向收益。
6. 多样性策略（DPP）, 用户冷启动（Airbnb KDD 2018）利用用户历史行为训练user embedding, 基于探索策略的商品冷启动。
   * + 1. **高德 搜索建议**

搜索建议（suggest）：用户在输入框输入query的过程中，推荐用户可能感兴趣的poi/keyword（如美食）

1. 前缀树召回候选结果，利用文本相关性进行粗排，获得候选poi
2. 基于用户历史行为、poi热度特征(点击率)、时空特征，构造2-3阶交叉特征。
3. pair-wise loss + gbrank建模，通过调loss function方法对模型引入先验知识

**2015.03-2017.03 高德 搜索NLP**

1. 使用CRF进行NER任务
2. 使用fastext进行Text Classification任务，进行query的品牌意图识别
3. 基于bert的Semantic matching用于query和poi title的语义相关性任务

项目经验