‍‍张乐求职意向： 机器学习/NLP

地址：天津市东丽区 |  电话：130-xxxx-0531  | 邮箱：625015879@qq.com

博客：http://www.cnblogs.com/zle1992/ GitHub https://github.com/zle1992

硕士 | 2016.9—2019.6| 航空运输大数据工程 (计算机学院) | 中国民航大学

本科 | 2012.9—2016.6| 通信工程（电子通信物理学院） | 山东科技大学

实习经历

* **2018年11月----2018年12月 CVTE 中央研究院 NLP算法实习**
  1. 利用python GUI 实现了故障检索系统。通过用户输入的故障描述，找到知识库中与之相近的故障现象，并展现相关替换零件的统计情况。
  2. 故障相似度计算使用预训练的bert模型得到句子向量，计算相似度。demo展示效果不错。
* **2018年7月----2018年9月 百度 推荐技术平台部 feed推荐算法实习**
  1. 负责百度三端首页的 Feed 流图文推荐算法研发 （手机百度、 wise、百度浏览器）。
  2. 主要负责召回、 小视频冷启动等算法研发。
* **2017年12月----2018年3月 百度 手百&feed体系 ---广告创意部 NLP策略算法实习**
  1. 参与基于检索的信用卡bot小程序研发
  2. 负责bot中问题解析、知识库构建、答案检索等策略的研发。主要包括：数据爬取及清洗，新词挖掘，关键词提取，文本相似度计算，答案检索等相关工作。
  3. 从0开始，最后程序内部上线（demo），准确率80%左右。
* **2017年7月----2017年9月 海南航空 海航科技---技术研究院 机器学习实习生**

1. 利用航班数据，提取相关特征对航空旅客聚类，构建旅客用户画像。
2. 根据航空公司历史航班数据，构建特征，利用前3年数据对后三个月客流量进行预测，最终RMSLE在0.22左右。

竞赛经历

* 2018 atec蚂蚁开发者大赛----金融智能nlp    排名16/2632 |  2018.5-2018.8

1. 问题相似度计算，即给定智能客服中用户描述的两句话，用算法来判断是否表示了相同的语义。
2. 搭建 CNN、 RNN、 ABCNN 、 ESIM、 注意力等深度学习模型，并挖掘一些特征，加到深度学习模型中，对 40w 句子语意建模，识别语意是否相同**。**。

* 2017ccf大数据竞赛----360 人机大战（机器文章识别）    排名13/589 |  2017.10-2017.12

1. 根据新闻的标题与新闻内容，识别文章是人写的还是机器生成的。
2. 作为团队（3人）主要负责人，初赛时，在20w数据集进行了数据清洗，分词，挖掘规则特征，xgboost分类，取得了top10的效果。复赛时，word2vec生成词向量，并利用CNN 、R-CNN深度模型对文章进行分类 。

* 2018全国高校大数据应用创新大赛---旅游评论打分预测   华北赛区 一等奖 |  2018.5-2018.7

1. 组委会提供10万条正式比赛数据，这些数据仅有评论内容，参赛队伍需要根据评论内容计算出评论等级。预赛将根据评论等级计算准确率计算参赛队伍的得分。
2. 利用Keras搭建CNN、RNN、Attention等深度学习模型，并结合传统tfidf+w2v+svm/nb/gbdt等分类模型，进行stacking模型融合。

* 2016 CCF大数据竞赛----搜狗用户画像构建   排名61/984  |  2016.10-2016.12

1. 根据用户历史的query与用户的人口属性标签（性别、年龄、学历），对新用户的人口属性进行判定。
2. 作为团队（3人）主要负责人，进行了 数据清洗、分词，TFIDF，主题模型，特征选择，SVM等工作。

* 2017中国高校大数据挑战赛---机器鼠标轨迹识别 排名20/1100 |  2017.5-2017.7

1. 根据鼠标轨迹信息（x,y,t），通过分类算法来判断轨迹是人或机器。
2. 作为团队（3人）主要负责人，初赛在10w数据集上分析数据、特征工程 、利用 Xgboost进行分类。 复赛在腾讯云平台实现200w数据的特征工程及5个模型的模型融合工作。

* 2017 JDD京东金融全球数据探索大赛 –登录行为识别 排名17/806 |  2017.11-2017.12

1. 根据用户的登录行为与交易行为，来识别当前用户的交易是否为风险交易。
2. 独立完成在10w数据上进行数据挖掘，提取上百维特征，特征选择，训练Xgboost模型进行stacking融合等工作。

* 2017全国研究生数学建模比赛 全国三等奖 |  2017.9

专业技能

* 熟悉数据预处理、特征工程、特征选择等数据挖掘工作。熟悉SVM 、LR、NB、RF、 Adaboost、GBDT等常用机器学习算法。熟悉TFIDF、Word2vec等NLP算法。熟悉CNN、RNN 、Attention等深度学习算法，熟悉常用数据结构算法。
* Python :最常用，所有的竞赛基于Python完成。 熟悉Numpy、Pandas等数据处理工具， 熟悉Scikit-Learn，Xgboost等机器学习工具， 熟悉Jieba、Gensim等NLP相关工具，会使用Keras、Pytorch 深度学习工具。
* Java： 熟悉使用，利用Java 开发过Android APP。C： 了解，本科期间使用，做过机器人的比赛，获得省一等奖。
* 会用Git ，会使用Liunx、简单shell。