



王伟好

北京邮电大学 | 信息与通信工程 | 学术硕士

籍贯：山西太原

出生年月：1995.08

手机：15501102797

邮箱：wangweiyu5107@163.com

教育经历

- | | | | |
|-------------------|--------|---------|---------|
| 2017.09 ~ 2020.07 | 北京邮电大学 | 信息与通信工程 | 硕士研究生 |
| 2013.09 ~ 2017.07 | 北京邮电大学 | 通信工程 | 本科/工学学士 |
- 研究方向为多媒体与网络大数据。熟练运用 Python 在 Jupyter Notebook 上进行数据分析，熟练掌握 SQL、MongoDB、Hadoop 等工具进行数据挖掘，熟悉常用的机器学习及数据挖掘算法。
 - 在研究生入学考试中以实验室第 1 名的成绩（430 分,学院前 1%）考入本校杨洁教授导师组。

实习经历

- | | | |
|-------------------|------------------------|-------|
| 2019.07 ~ 2019.09 | 中国农业银行总行 信用管理部 资产组合监控处 | 数据分析岗 |
|-------------------|------------------------|-------|
- 代表处室参与“雅典娜杯”数据挖掘竞赛，针对智慧风控场景“线上贷款违约预测”，对用户基本信息、资产信息、信用卡及信贷信息进行数据挖掘，构建的风控预测数据模型 KS=57，进入竞赛排行榜 Top10%。
 - 参与处室其他工作，对贷后管理信息进行文本数据分析和统计，协助整理并撰写部门年中审计的访谈稿等。

项目经历

- | | |
|-------------------|--------------------------------|
| 2018.09 ~ 2019.06 | 研究生毕设：文本分类与聚类算法在网络安全运营系统中的应用研究 |
|-------------------|--------------------------------|
- 项目描述：**研究选用机器学习算法实现网页文本的自动分类以及从海量网络文本中对安全情报话题的自动识别，以帮助系统统计和管控上网内容、进行网络舆情监测等。
 - 网页自动分类：**数据源为数千万条用户上网 URL 访问记录，经去重并获取网页内容后得到的网页文本数据，存储在 MongoDB。经过多次实验，最终对预处理后的文本用 Word2vec 作为特征表示+深度随机森林作为分类器的方式进行网页多分类，对真实网页的分类结果 F1-Score 达到 0.9。
 - 安全话题聚类：**对网络文本进行二分类，判断是否为网络安全相关，并对相关文本提取 TF-IDF 特征，自编码器降维后选用 K-Means 算法进行话题聚类，实验数据集上的聚类召回率 87%以上。
- | | |
|-------------------|------------------------------|
| 2017.03 ~ 2017.06 | 本科毕设：用大数据技术对移动互联网用户流量数据的分析研究 |
|-------------------|------------------------------|
- 实现了用于海量数据的高性能多模匹配算法，提升了数据流量管理和分析的效率。
 - 从网络流量报文中总结出各类移动端业务的模式特征，在 Hadoop 大数据分析平台上改进并实现了 AC 多模式匹配算法，用 MapReduce 分布式计算模型对流量数据进行高效的特征识别和统计分析。

获奖经历

- 2018.09 “华为杯”中国研究生数学建模竞赛 国家级三等奖
 - 独立完成第二问的数学建模与程序编写，为我国城市群制定网络价值最大化的光传送网规划。
 - 与小组成员有效沟通，合作完成选题调研、任务分配、论文撰写等任务。
- 本科及研究生阶段多次获得北京邮电大学一等、二等奖学金等奖励，优秀团员等荣誉。
- 作为校田径队成员，多次在校级运动会的跳远、200 米、4x100 米接力等项目中获得前三名。

其他

- 校园活动：**
 - ①曾担任班级团支部宣传委员、北邮阳光志协多类活动志愿者及国家图书馆长期志愿者；
 - ②热爱并擅长田径运动，作为校田径队主要成员多次代表北京邮电大学参加首都高校大学生运动会。
- 个人评价：**踏实耐心，待人谦逊友好、真诚和善，做事认真负责、积极主动，具有团队合作精神；思维逻辑清晰，对新技术充满学习热情，有较强的理解、学习、分析并解决问题的能力。