

李煜



硕士研究生，中共党员， Email: liyu52577@163.com, Tel: (+86)18516255516

工作经历

- 上海大智慧财汇数据科技有限公司** **高级算法工程师** 2017.03- 至今 上海
- 基于统计语言模型的股票涨跌趋势预测，将预测模型应用到 Range Break 区间突破交易系统中，降低了收益回撤和收益波动率。
 - 量化因子的挖掘以及多因子模型的构建，为机器学习模型寻找更多的特征因子。
 - 开发股票时间序列量化波动模型，模型分别使用了 LR, SVM, XGBoost, lightGBM 等传统的机器学习模型，和 Bi-LSTM 的深度学习模型，并且有做过集成学习。并且对参数较多的 XGBoost 和 lightGBM 使用贝叶斯优化方法实现了超参数的自动调优。
 - 基于财经新闻、论坛评论、券商研报语义分析提取相关舆情因子，根据提取的舆情因子进行行业涨跌趋势预测，同时将该模型与基于量价数据的趋势涨跌模型进行融合。
 - 基于 ELK 的实时文本搜索系统搭建和维护，用于关键词的实时搜索推荐，搜索响应时间大大缩短。根据用户的搜索浏览记录向用户推送最新的相关信息资讯。
 - 根据股票的主营产品和主营业务的文字描述计算该股票的标的，构建上市公司竞争关系图谱。

- 上海坤雁数据服务有限公司** **算法工程师** 2016.01-2017.03 上海
- 负责开发基于 KNN 的验证码识别系统 (Python)。
 - 金融行业文本分类，基于 TextRank, Word2Vec 建立一套金融行业文本分类系统，提高了行业分类的准确度 (Scala, Python)。
 - 负责热词系统的改进，建立一套基于贝叶斯平均和牛顿冷却算法的热词候选词的热度评价标准和热度计算标准 (Scala)。
 - 开发基于 TextRank 的新闻文本关键词和文本摘要提取系统 (Scala)。
 - 基于 Spark 集群，进行金融财经类新闻文本主题聚类分析，研究开发相关主题事件抽取模型。
 - 负责市场行情趋势、金融风险等相关的数据挖掘、特征值提取和建模，研发市场情绪指数模型。

- 亿臻能源 (上海) 科技有限公司** **算法工程师** 2014.11-2015.12 上海
- ECT 正问题分析，实现二维/三维管道有限元分析与计算，构造 SVM 训练样本；
 - 使用支持向量机 (SVM) 分类算法进行二维/三维图像多类分类重建模型的建立和算法优化；
 - 负责项目团队管理，项目硬件电路设计部分的研发人员的管理和协调，研发进度的把控。

其他经历

- 参加 THE MTS/IEEE OCEANS'15, Washington, DC** 2015.10
- 参加在美国华盛顿举办的国际海洋年会 THE MTS/IEEE OCEANS'15 并在大会上做论文汇报，论文的主要研究内容是将非负矩阵分解(NMF)算法应用到潮汐信号的盲分离上，研究论文被 EI 检索。

研究生期间参与的科研项目

- 电容层析成像技术应用与发展研究** 2014.11-2016.05
- 编写完成基于 Matlab 的 ECT 系统仿真软件包，实现二维/三维管道有限元分析与计算；
 - 利用二维电磁场有限元仿真软件包，对传感器结构的优化设计和灵敏度场的分布的研究，以期优化传感器的结构参数,从而改善敏感场检测灵敏度和空间分辨率；
 - 使用支持向量分类机(SVM)算法进行二维/三维图像重建模型的建立和算法优化研究。
- 基于 GIS 的海洋环境多参数快速检测与处理关键技术研究(上海市科委项目)** 2013.09-2016.05
- 负责实验室 7 台服务器的安装调试和并进行高性能计算测试；
 - 该项目将物联网技术应用到海洋环境监测中，并结合 GIS 可视化、空间分析等技术，从海洋环境数据获取、数据传输、数据应用等方面进行了研究；
 - 使用非负矩阵(NMF)算法对获取的潮汐数据进行分析处理。

	<p>长江口附近海岸生态修复及生物资源利用技术及示范 2013.09-2016.05</p> <ul style="list-style-type: none"> 搜集长江口附近海岸带生态环境监测资料、近海海域使用情况等数据，完成长江口附近海岸带典型生物环境退化诊断与评价。 根据浒苔遥感监测数据研究浒苔动态变化情况，建立基于支持向量回归机(SVR)的浒苔漂移路径以及面积预测模型。 <p>海洋信息服务产业科技应用与产业发展研究 2013.11-2013.12</p> <p>担任项目的主要执行人，负责该项目前期资料收集，数据整理，以及最终报告的撰写完成。</p>
获奖情况	<p>第十二届“中关村青联杯”全国研究生数学建模竞赛全国二等奖 1次 2015.10</p> <ul style="list-style-type: none"> 担任队长，负责程序的编写和论文撰写，文章涉及子空间聚类、多流形聚类、图像分割、运动分割、人脸识别等具体问题，我们先利用现有的聚类方法（谱聚类等）进行建模编程得出结果，然后给出一种创新性求解方法，将支持向量分类机算法应用于混合流型数据聚类问题，并对实际的一个机器工件外部边缘轮廓图像的多流行聚类问题将轮廓线中不同的直线和圆弧分类。 <p>第十一届“华为杯”全国研究生数学建模竞赛全国三等奖 1次 2014.10</p> <ul style="list-style-type: none"> 担任队长，负责模型的建立和程序编写，文章主要研究的是小鼠视觉感受区电位信号(LFP)与视觉刺激之间的关系，文章通过对小鼠的呼吸原理和受刺激的脑电波变化的生理机制的分析，建立小波支持向量神经网络模型，揭示了呼吸曲线的频率和脑电波信号的某种非线性相关关系，最后根据这种对应关系建模得出了视觉感受区的 LFP 的规律性变化，分析了视觉刺激下的脑电波曲线和视觉刺激曲线的关系。 <p>第一届全国高校物联网应用创新大赛华东赛区三等奖 1次 2014.07</p> <ul style="list-style-type: none"> 物联网传感器节点的性能优化，在 Linux 下使用 TinyOS 操作系统编程实现传感器节点的相互通信，主要负责对节点通讯进行优化，降低传感器节点的能耗。 <p>淮安市创业计划大赛三等奖、国家励志奖学金、校职业生涯规划大赛三等奖校二等奖学金、优秀学生干部、优秀党员、科技文化艺术节先进个人</p>
个人能力	<ul style="list-style-type: none"> ● 熟练掌握 Python、Scala、Matlab、C++开发。 ● 理解分类聚类回归问题，熟悉 SVM, XGBoost 等机器学习算法，熟悉自然语言理解。 ● 理解深度学习,熟悉 Tensorflow、Pytorch 等深度学习框架。 ● 了解知识图谱相关理论（实体识别、语义搜索等）； ● 了解 Spark, Hadoop, kafka 等分布式大数据框架，了解日志监控系统 ElasticSearch; ● 熟悉 Linux 原理和 Docker 技术,了解 Docker 相关技术生态; ● 熟悉 HBase、redis 等主流的 no sql 数据库。
论文发表	<p>1 《Unmixing :A New Direction From Classical Tidal Harmonic Analysis》[J],2015 年国际学术年会,MTS/IEEE OCEANS'15,Washington,DC,第一作者,论文被 EI 检索,2015(9);</p> <ul style="list-style-type: none"> ● 将非负矩阵分解(NMF)算法应用于潮汐分潮的非监督盲分离中，解决了传统分离遇到的数据样本搜集困难，计算复杂等问题。 <p>2 《基于投影梯度的 NMF 盲信号分离算法》[J],第一作者,计算机工程,2016(02);</p> <ul style="list-style-type: none"> ● 使用投影梯度的迭代思想进行非负矩阵分解的迭代寻优,对于稀疏信号的盲分离提高了分解精度、降低了算法的时间复杂度。 <p>3 《基于小波支持向量机的小鼠脑电波与呼吸的关系》[J].中国医学物理学杂志. 2015(03);</p> <p>4 《基于支持向量机的绿潮灾害影响因素的权重分析》[J].中国环境科学. 2015(11)。</p>
自我评价	<p>两年的算法相关工作，熟悉分类，聚类机器学习算法，在自然语言处理领域有一定了解，对 CNN、RNN、LSTM 等深度学习算法有一定的了解。使用过 Spark/Hadoop/HBase 等大数据平台。良好的数据敏感能力，与数据打交道而乐此不疲。工作刻苦踏实，认真负责。为人乐观开朗，乐于助人。组织协调能力强，有很强的沟通能力，擅于发现和解决问题，对新事物的接受和学习能力很强，对细节有足够的洞察力，对数据有足够的敏感性。</p>