

教育

上海交通大学 硕士, 生物医学工程 课题: Retina神经网络编码机制	2010.09 - 2013.03
上海交通大学 本科, 生物医学工程	2006.09 - 2010.06

经历

通用电气 <i>Staff Data Scientist</i> , 资深数据科学家	2016.6 - 现在 GE航空集团数字部门
---	---------------------------

· GE航空集团中国区数据分析技术负责人:

1. 在数据科学家团队中担任技术负责人, 和每个数据分析项目的PL一起制定技术方向以及设计算法, 帮助所有项目优化模型, 并为其他部门的团队提供数据科学方面的咨询。
2. GE全球分析标准委员会的成员, 作为中国区的代表参与GE内部数据分析项目标准化流程和方法的制定, 与全球其他团队共同开发自动化分析模块。
3. 对内负责发动机分析项目, 对外与客户(包括商飞, 海航, 国泰航空等)共同开发产品。

飞利浦心血管院后数据分析系统 <i>Scientist, Project Leader</i>	2015.8 - 2016.6 飞利浦中国研究院
--	-----------------------------

· 用户行为分析与个性化推荐:

1. 利用系统所收集的用户信息(包括年龄、性别、教育程度、家庭情况等所有与用户相关的信息)和系统所记录的用户行为(包括用药依从性、检查录入情况、在系统中的行为等等), 判断哪些用户是该系统的潜在客户, 以及哪些用户在使用该系统后会获得医学收益
2. 使用stacking(RF+GBDT+LR)对用户进行分类
3. 建立知识库, 根据不同的用户分类进行个性化推荐

个性化推荐知识库 <i>Scientist, Project Leader</i>	2015.07 - 2015.12 飞利浦中国研究院
--	-------------------------------

· 为个性化推荐系统建立知识库和规则匹配系统:

1. 提出一个多层知识抽取框架, 将原始的以自然语言描述的“知识”逐层抽象成为计算机可执行的语句
2. 使用wiki作为知识管理平台
3. 编写爬虫将wiki页面上的知识抽取并写入MongoDB
4. 使用Drools管理和执行匹配规则

高风险孕妇预测模型 <i>Scientist, Project Leader</i>	2014.06 - 2015.07 飞利浦中国研究院
---	-------------------------------

· 基于移动平台的孕妇风险监测系统(Mobile Obstetrical Monitoring):

1. 使用Logistic Regression(LR)和Random Forests(RF)算法建立孕妇孕期出现高危高血压以及 preeclampsia疾病的预测模型

2. 对数据进行特征选择, 使用包括在LR中加入正则化惩罚项(Lasso), 和在RF中使用out of bag (OOB)估计变量重要性(variable importance)
3. 试验不同的梯度下降(GD)算法对结果的影响, 包括batch GD, Momentum, Adam, 以及BFGS
4. 在android平台开发一个原型, 用于临床医生对该模型进行验证

个人健康管理 with 个性化推荐系统

Scientist, Project Leader

2014.01 - 2014.12

飞利浦中国研究院

· 个人健康管理系统(Personal Health Management Solution):

1. 使用Logistic Regression(LR)和支持向量机(SVM)算法建立四年期高血压、八年期糖尿病和十年期心血管疾病的风险预测模型
2. 利用不断得到的新数据, 尝试建立了预测模型的在线版本(online LR)

智能检测套件

Scientist

2013.04 - 2013.12

飞利浦中国研究院

· 智能建议模块的算法设计(临床决策支持):

1. 进行文献研究(literature study), 利用文献中已报导的模型(包括Weibull Regression和Cox Regression)对患者进行风险预测
2. 建立知识库, 建立风险预测结果和个性化建议之间的映射关系

技能

- 熟练掌握各种机器学习算法, 包括回归和分类(LR, LAR, DT, SVM等), 聚类(kNN, kMeans), 以及集成学习算法(boosting, bagging和stacking)
- 理解机器学习理论, 包括PAC学习, VC理论, 以及Rademacher复杂度
- 掌握深度学习算法, 包括RBM, DBN等
- 掌握各种优化算法, 包括梯度下降及其变种(BGD, SGD, Momentum, Adam, SAGA, SVRG等), 牛顿法 (newton, BFGS), Proximal, ADMM
- 掌握各种特征选择(feature selection)方法, 包括filter, wrapper, embedded
- 了解Hadoop, MapReduce, Spark
- 熟练掌握Python, Matlab, R
- 掌握Java, C++

发表

期刊

- **Jingyi Bu**, Hao Li, Hai-Qing Gong, Pei-Ji Liang, Pu-Ming Zhang. "Gap junction permeability modulated by dopamine exerts effects on spatial and temporal correlation of retinal ganglion cells' firing activities." in *Journal of Computational Neuroscience*, 2013. (SCI indexed, IF = 2.51)

国际会议

- **Jingyi Bu**, Ning Lan. "An Improved Multi-Channel Cortical Recording And Stimulation System." *International Convention on Rehabilitation Engineering & Assistive Technology*, p. 98-101, 2010. (EI indexed)

专利

- Wang Jin, **Bu Jingyi**. “An Apparatus and Method for Evaluating Multichannel ECG Signals”
WO2015052609A1.16/04/2015