

2017 年 9 月更新

# 杨帆

• 手机: (086) 131-1615-8612

• 电子邮箱: [fanyang0801@yahoo.com](mailto:fanyang0801@yahoo.com)

• 籍贯: 天津

## 专业概述

- 熟悉机器学习和深度学习等算法模型、数据可视化、网页建设、数据库管理 ( 计算机科学硕士 )
- 构建统计模型和假设检验的相关经历 ( 统计硕士+两年统计师 )
- 经济、金融报表、股票管理、固定收益及衍生工具相关知识 ( CFA 一级 )

## 经历

### 研究助理

09/2016 - 09/2017

计算生物医学成像与建模中心 ( CBIM ), 罗格斯大学

新泽西州新布朗斯维克市, 美国

- 在 Konstantinos Michmizos 教授实验室进行医学信号建模研究

### 统计师

05/2013 - 05/2015

Target Health Inc.

纽约州纽约市, 美国

- 审查临床方案并辅助制定统计分析计划书 ( SAP )
- 根据客户需求和美国食品药品监督管理局 ( FDA ) 规定编写 SAS 程序并进行统计分析
- 有数据清洗经验 ( 异常检测, 多重插补法 )
- 工作涉及与临床站点、中心实验室和大型医药公司 ( 辉瑞、礼来等 ) 沟通交流

### 研究助理

12/2007 - 06/2009

天津医院骨科研究所

天津, 中国

- 发表文章: 甲氨蝶呤诱导色素沉着绒毛结节性滑膜炎滑膜细胞凋亡的研究。范新军, 杨帆, 邢国胜, 赵文君, 陆芸, 白人骁《天津医药》2008 年 12 期

## 学历

### 计算机科学, 理学硕士

09/2015 - 10/2017

罗格斯大学-新布朗斯维克分校 GPA 3.7/4.0

新泽西州新布朗斯维克市, 美国

- 毕业论文: 使用改进 k-平均算法方法自动定位下丘脑核

### 统计学, 理学硕士

09/2011 - 05/2013

罗格斯大学-新布朗斯维克分校 GPA 3.8/4.0

新泽西州新布朗斯维克市, 美国

- 免除硕士资格考试

### 生物医学博士项目, 退出

09/2009 - 07/2011

俄勒冈健康科学大学

俄勒冈州波特兰市, 美国

- 活动: 中国学生学者联合会 ( CSSA ) 副主席

### 生物技术与生物信息, 工学学士

09/2004 - 07/2008

天津医科大学 ( 211 工程 ), 生物医学工程系

天津, 中国

## 证书

06/2015

特许金融分析师 ( CFA ) 一级 通过

12/2014

SAS 认证统计商务分析师: 回归分析与建模 ( 证书号: SBARM001399v9 )

07/2011

SAS 认证高级证书 ( 证书号: AP000926v9 )

06/2011

SAS 认证初级证书 ( 证书号: BP003352v9 )

04/2006

国家计算机等级考试 三级 通过

2017 年 9 月更新

## 技能

- |       |  |  |
|-------|--|--|
| 计算机技能 | <ul style="list-style-type: none"><li>• 编程语言：Java, Python</li><li>• 数学统计软件：SAS, R, Matlab</li><li>• 操作系统：Windows, OS, Linux</li><li>• 数据库：MySQL, SQLite</li></ul>                  | <ul style="list-style-type: none"><li>• 机器学习工具包：Scikit-Learn, Tensorflow, Mahout</li><li>• 大数据并行计算：Hadoop, MapReduce, Spark, AWS</li><li>• 网页建设：HTML5/CSS3/Javascript, jQuery/AJAX, Bootstrap, PHP</li></ul> |
| 语言    | <ul style="list-style-type: none"><li>• 普通话（母语），英语（熟练）</li><li>• 英语六级：605/710</li><li>• 托福：106/120</li><li>• GRE: 语文 640/800 (91%) + 数学 800/800 (94%) + 写作 4.5/6.0 (80%)</li></ul> |  |

## 部分项目

### 硕士论文

- **使用改进 k-平均算法方法自动定位下丘脑核**  
09/2016 - 10/2017  
导师：Professor Konstantinos Michmizos  
语言与工具：Matlab, Python

深度大脑刺激（DBS）对于治疗帕金森病具有显著效果。我们的研究旨在找到一种机器学习算法来帮助医生自动定位下丘脑核（STN），从而提高 DBS 手术成功率。我们对深度大脑刺激的手术中收集的大脑微电极信号（MERs）进行波段过滤（0.5kHz – 8kHz）并提取出 18 种特征。其中 4 种是我们自己设计的，而其他 14 种来自过往文献。对这些特征进行平方根转换、标准化和正规化后，将修正过的特征值输入 k-平均算法得出聚类输出。利用 curve length 的特征活动图谱，我们从各个聚类中找到下丘脑核对应的聚类，进而得到下丘脑核的上界和下界。然而传统 k-平均算法的输出结果含有大量噪音。通过将深度信息引入特征值矩阵，改进的算法大大降低了聚类输出中的噪音，这实现了自动化界定下丘脑核的目的。通过与专家划定的真实结果对比，我们的方法对下丘脑核背侧和腹侧两个的边界的预测在 0.5mm 偏差标准下分别达到了 93.5%和 80.7%的成功率，而在 1mm 偏差标准下则达到了 100%和 93.5%。

### 机器学习

- **构建随机梯度下降优化的逻辑回归分类器**  
10/2015 - 12/2015  
合作者：Bharath Joginapally, Sanjay Bharadhwaj Vijayaraghava  
语言与工具：Java, Apache Mahout

通过使用 Java 和 Apache Mahout 我们建立了一个逻辑回归分类器用于对二分变量或者分类变量进行预测。测试该分类器使用了 3 个数据集，样本量分别为 150, 4000 和 4 百万。

- **使用机器学习算法检测网络入侵**  
03/2016 - 05/2016  
合作者：郑洋，程鑫  
语言与工具：Python, NumPy, Pandas, Scikit-Learn, Spark

我们运用不同的机器学习算法建立了一个网络入侵检测系统并对不同方法进行比较。涉及到的算法包括朴素贝叶斯、决策树、逻辑回归分类器、支持向量机、K-平均算法和人工神经网络。数据来自 UCI 机器学习库，全数据的样本量为 5 百万，41 个自变量（特征），要预测的因变量的数值为“正常”（的网络链接）或者不同类型的网络攻击。这个项目分为三个阶段。第一阶段为“二分阶段”：任务是预测出一个网络链接是“正常”还是“网络攻击”。第二阶段为“多分阶段”：目标是给出一个网络链接是“正常”还是某一特定种类的“网络攻击”。介于前两个阶段是单机计算，我们只用了 10%的训练集，大约 40 万数据进行模型训练。第三阶段，我们利用 Amazon AWS Spark 技术结合 K-平均算法对全体 5 百万数据进行训练，验证了第二阶段的结果。

### 大数据分析

- **基于 n-gram 词频统计的谷歌自动完成模型**  
09/2016 - 10/2016  
语言与工具：Java, Hadoop, MapReduce, jQuery, PHP, MySQL

这个算法使用了连续两个 MapReduce 过程。通过 wiki 数据库我们建立了一个 n-gram 短语库。在该短语库基础上，我们建立一个语言模型计算出每个短语出现的概率。运用 PHP 和 MySQL 管理后端数据，jQuery 和 AJAX 做前端演示。

- **电影推荐系统**  
11/2016 - 12/2016  
语言与工具：Python, Hadoop, Amazon EMR

根据各用户过往观看电影打分记录的相似度建立电影推荐系统。使用了 MovieLens 上 2 千万条电影打分数据。

### 网站构建

- **运用 Spring-Hibernate 技术构建具备 CRUD 功能的学生注册系统**  
12/2016 - 01/2017  
语言与工具：Spring MVC Framework, Hibernate Framework, HQL, Java EE, JSP, HTML, CSS, XML, Tomcat 8.5