COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

姚 华

+1-(646) 318-8174 | hy2632@columbia.edu
LinkedIn: linkedin.com/in/yaohuacn | 个人主页: https://hy2632.github.io/about/

教育经历

哥伦比亚大学 **2019 年 8 月 – 2020 年 12 月**
管理科学与工程 硕士 **美国 纽约**

- GPA: 4.06 / 4.33
- 荣誉: 研究生奖学金 Frydman Graduate Fellowship (10%)
- 相关课程: 运营咨询，高级公司金融，资本市场与投资，商业分析，商业分析策略，数据分析工具，概率论和数理统计建模，优化方法，随机过程与建模

清华大学 **2019 年 7 月**
土木工程 学士 **北京**

- GPA: 3.56 / 4.0

项目经历

马克雅各布 (Marc Jacobs International)，LVMH 集团 **2019 年 8 月 – 2020 年 5 月**
数据分析顾问 (兼职)，全球零售计划部 **美国 纽约**

采购工具设计: 基于随机森林模型的产品销售量预测

- 收集产品分地区及分季节的历史销售数据，使用 Pandas 进行数据清洗和预处理
- 使用 matplotlib 和 seaborn 工具包将数据可视化，定性分析市场规模变化/区域效应/色彩流行趋势/款式等因素，并基于采购部门主题专家的意见和模型调试验证，定义并确定了关键属性
- 建立随机森林 (Random Forest) 回归模型， 根据历史数据和关键属性预测新产品的销售量，实现优化采购规模、提高售罄率 (Sell-through rate)

诺和诺德(Novo Nordisk Inc., U.S.) **2019 年 8 月 – 2020 年 5 月**
数据分析顾问 (兼职)，商业洞察和分析部 **美国 新泽西**

基于 IQVIA 数据库建立聚类模型，确定抗肥胖药物 Saxenda 的受众群体的优先级

- 基于 IQVIA 医疗数据库，使用 SQL 创建一个以病人 ID 为主键，包含诊断、医疗过程、处方、社会经济数据的数据库，并提供代码供公司周期性同步 IQVIA 数据库的最新数据
- 使用 K-means、Gaussian Mixture 等聚类算法开发基于患者对于该药物接受程度的优先级排名系统，解释每个患者群体的典型特征，使用 Tableau 生成高优先级患者的地理热力图，并提供关于优先发展区域的业务见解
- 向诺和诺德商业策略和发展团队的主题专家作两次专题报告，得到商业洞察和分析部总监高度评价

学术研究

CS224N (深度学习与自然语言处理) 课程项目: 基于 SQuAD2.0 数据集的问题回答系统 **2020 年 7 月**

- 阅读论文，用 PyTorch 框架实现 BiDAF(with Char-Embedding) 和 QANet 两种无预训练上下文词嵌入的模型

DROM8110 (商业分析策略) 研究报告: Uber 盈利前景和经营策略研究 **2020 年 5 月**

- 运用贡献价值、可持续性、支付意愿等商业分析框架，研究 Uber 主营业务——叫车 (Ride-hailing) 的盈利前景和可行的经营策略 (https://hy2632.github.io/2020/08/22/ResearchOnUber/#more)

其他

数据分析/计算机语言: Python (pandas, NumPy, Matplotlib, scikit-learn), Java, SQL, Tableau, Excel

深度学习: PyTorch, Git, Microsoft Azure, CS224N (自然语言处理), CS231N (卷积神经网络与图像识别)

证书: CFA II 级考生

语言: 英语 (TOEFL 108, GRE 328)，法语 (基础)

兴趣爱好: 摄影 (https://500px.com/yaohua)，艺术，足球

H U A Y A O

(646) 318-8174 | hy2632@columbia.edu
LinkedIn: https://www.linkedin.com/in/yaohuacn |Homepage: https://hy2632.github.io/about/

EDUCATION

Columbia University, New York, NY
M.S. in Management Science & Engineering **Aug 2019 – Dec 2020**

- Cumulative GPA: 4.06/4.33
- Award: Frydman Graduate Fellowship (10%)
- Coursework: Operations Consulting, Advanced Corporate Finance, Capital Markets and Investment, Business Analytics, Business Analytics Strategy, Tools for Data Analytics, Probability Statistics and Simulation, Optimization Methods, Stochastic Modeling

Tsinghua University, Beijing, CN
B.S. in Civil Engineering **Jul 2019**

- Cumulative GPA: 3.56/4.0

PROJECT EXPERIENCE

Marc Jacobs International, New York, NY
Data Analytics Consultant (Part-time), Global Retail Planning **Aug 2019 – May 2020**
Buy tool design: product sales forecast based on Random Forest model

- Collected historical sales data of products by region and season and used Pandas for data cleaning and preprocessing
- Drew qualitative conclusions on regional effects, general trend, etc. through data exploration and visualization, using matplotlib and seaborn, identified and defined key attributes based on the opinions of procurement SMEs and model validation
- Built a Random Forest regression model to predict sales volume of new products based on historical data and key attributes, optimizing the buying process and increasing the sell-through rate

Novo Nordisk, Plainsboro, NJ
Data Analytics Consultant (Part-time), Insights and Analytics **Aug 2019 – May 2020**
Established a clustering model to identify patients' receptiveness to their anti-obesity medicine Saxenda

- Used SQL and IQVIA database to create a database containing patients' DX, PX, RX and socioeconomic data with a primary key of patient ID and provided SQL source code for database auto-updating
- Used K-means, Gaussian Mixture, and other clustering algorithms to develop a ranking system based on patients' receptiveness to the medicine
- Explained the patterns of each patient group and used Tableau to generate a choropleth map, visualizing the geographical distribution of high priority patients, and provided business insights on determining prioritized regions
- Gave two informative presentations to SMEs from Novo Nordisk US's Business Strategy & Development team and was highly rated by Director of Commercial Insights & Analytics

RESEARCH

CS224N (NLP with Deep Learning) Final Project
Question Answering on SQuAD 2.0 Dataset **Jul 2020**

- Implemented two Non-PCE models: BiDAF (with char-embedding) and QANet using PyTorch

DROM8110 (Business Analytics Strategy) Research Paper
Research on Uber's Prospect of Profitability and Strategy to achieve Sustainability of Business **May 2020**

- Conducted analysis on Uber's core business, Ride-hailing, from Contributed Value, Sustainability, WTP, etc.
- Link: https://hy2632.github.io/2020/08/22/ResearchOnUber/#more

SKILLS

Data Analytics & Programming: Python (pandas, NumPy, Matplotlib, scikit-learn), Java, SQL, Tableau, Excel
Deep Learning: PyTorch, Git(github.com/hy2632/), Microsoft Azure, CS224N (NLP), CS231N (Visual Recognition)
Certificate: CFA Level I Passed
Interests: Photography (https://500px.com/yaohua), Fine Art, Soccer