XIN (LEXIE) WEN

86-55 Broadway
New York
NY

814.321.5918
xw2443@columbia.edu
https://www.linkedin.com/in/xin-wen

EDUCATION

COLUMBIA UNIVERSITY   **New York, NY**
Master of Science in Operations Research (MSOR)   *GPA: 3.6/4.0*   Dec 2017

THE PENNSYLVANIA STATE UNIVERSITY   **State College, PA**
Bachelor of Science in Economics, Minor in Mathematics and Statistics   *GPA:3.9/4.0*   May 2016

COURSEWORK/SKILLS

Coursework: Machine Learning in Data Science, Data Analytics, Business Analytics, Cloud Computing, Project Management
Programming: Python, SQL, R, Excel(VBA)
Tools: Tableau, AWS, MySQL, PostgreSQL, Google Analytics, Salesforce, Trello

EXPERIENCE

HEALTHCASTS   **New York, NY**
Data Scientist, Member Services Team   Feb 2018 – Present
- Operationalize data process for a peer-to-peer education platform for physicians (resembles the medical version of Quora), lead weekly, monthly and quarterly business reports, and socialize analytical results into actionable insights for sales and production departments to make business decisions
- Maintain 2 databases (3M records with physicians and programs) on MySQL and PostgreSQL(AWS RDS) including privilege management, pulling data, writing script, being familiar with ETL process
- Initiate a Natural Language Processing(NLP) project in python and implement the final models into production
- Build web scrapers tool in python to automatically fetch physicians data in different online sources
- Design and evaluate A/B tests to reduce user engagement funnel drop-off. Worked with stakeholders to implement product and features changes that improve user engagement by increasing answer/view rate and shortening response time

Business Intelligence Analyst Intern   Oct 2017 – Jan 2018
- Conducted proactive analyses for Core Experience to understand user behavior, validate use cases, and identify growth opportunities. Provided recommendations and business insights to facilitate decision making
- Developed reporting tools on Tableau and Cyfe, made recommendations for improving corporate health metrics

COLUMBIA BUSINESS SCHOOL   **New York, NY**
Data Analytics Research Assistant; Customer Behavior Analysis on Executive Education Program.   Dec 2016 – May 2017
- Applied Logistic Regression, Gradient Boosting Machine to build individual enrollment predictive model, helping team figure out different patterns in applications and predicted enrollments for different program
- Initiated B2B model analysis to find importance of HR representatives behavior in different industries and companies

PROJECTS

Healthcasts Medical Q&A Platform in NLP. (Python)
- Train a high-quality embeddings containing medical vocabulary using Word2Vec(CBoW), Doc2Vec and inferSent. Leverage cosine similarity and spectral clustering results to evaluate
- Perform Supervised learning and semi-supervised learning on the platform data to improve the embedding and reach to find most popular answers through labeling
- Build a user-based Collaborative Filtering Model to predict a physician's answer to a new question, given a physician's response to questions in the past. Achive 66-68% accuracy

Research on the Results of a Speed-Dating Experiment. (Python)
- Applied K-means and Neural Network algorithem to identify the features of successful dating and improved the development of speed dating
- Built a predictive model for different types of participants, matched them and visualized model output by Python (Matplotlib)
- Implemented and constructed a web application in Python Flask to present text and characteristics analysis