

# JENNIE LE

## DATA SCIENCE & ANALYTICS

### CONTACT

✉ [t1e35@fordham.edu](mailto:t1e35@fordham.edu)  
☎ (347) 449-2616  
📍 New York, NY  
🌐 [linkedin.com/in/jenniele](https://www.linkedin.com/in/jenniele)  
🌐 <https://github.com/bigforehead>  
🌐 [bigforehead.github.io](https://bigforehead.github.io)

### PROFILE

Background in consulting with focus on building solutions using data analytics to improve efficiencies and revenue.

### EDUCATION

2020  
FORDHAM UNIVERSITY [NEW YORK, NY]

#### Master of Science in Analytics

GPA: 3.81

2017  
ELIZABETHTOWN COLLEGE [ELIZABETHTOWN, PA]

#### Bachelor of Science in Finance and Accounting

### TECHNICAL SKILLS

- Data Analysis
- Statistics and Probability
- Machine Learning
- Natural Language Processing
- Data Visualization
- Database Management
- Python, R, SQL, Tableau
- Excel
- SPSS Modeler
- Pandas, Scikit-learn, NLTK, TensorFlow
- Google Cloud Platform (GCP), AWS, Data Bricks, Azure
- Alteryx, QlikView, Spotfire

### AWARD

- Dean's List
- Delta Mu Delta
- The National Society of Leadership & Success

### EXPERIENCE

MAY 2020 – AUG 2020

NEW YORK, NY

#### Data Analytics Intern, Design Lab | Alvarez & Marsal

- Led a team of 5 to conduct ad-hoc data analysis via machine learning, text analytics with **Natural Language Processing**, and **web scraping** to collect Glass Door's 60,384 job descriptions from S&P500 and examine the relationship between financial performance and organizational culture
- Performed deep **exploratory data analyses** using **Tableau** to provide analytical insights and drafted 5 academic research papers (Crowdfunding, Litigation, University Technology, ORCID, Kiva)

JAN 2020 – PRESENT

NEW YORK, NY

#### Graduate Assistant | Gabelli School of Business

- Self-learned theories and created lecture materials (Audit Data Analytics, Consumer Behavior, Big Data)
- Developed workshop materials and gave over 10 lectures to 120 graduate students in **Tableau**, **Databricks**, and **Azure**
- Scraped Instagram metadata of +200 accounts and +7 million medias using virtual machine on **Google Cloud Platform**
- Built statistical models to predict fashion trend and generated model assessment report using Tableau and Python

MAR 2018 – FEB 2019

NEW YORK, NY

#### Economic Research Analyst | Haver Analytics

- Maintained and updated extensive time-series database utilizing Data Link Express (DLX)
- Led a team of two by providing technical guidance and engagement opportunities
- Provided regular reporting on project performance to leadership team and high-level executives
- Streamlined updating process by **50%** by writing **DOS batch** to automate consolidation tasks (data collection, database updates, etc.)

### PROJECTS

#### One Million Song Music Recommendation System

- Built a personalized recommendation system with **Spark** and **Mlib** using a 3GB dataset with Ubuntu Virtual Machine on Google Cloud Platform
- **Queried** metadata from SQLite database and performed data consolidation and **exploratory data analysis**
- Visualized and model 20 topics from the lyrics using **NLTK** and **Latent Dirichlet Allocation**
- Engineered new features of **TF-IDF**, **Word2vec**, and **LDA** on lyrics to build a content-based song recommender
- Applied **Alternating Least Squares** recommender algorithms with **Grid Search** and cross-validated on the prediction results with **9.61 RMSE**

#### Google Question & Answer Topic Classification

- Applied Deep Learning using **Keras** and Gradient Boosting classifiers with **Cross Validation**, **tackling data imbalance using SMOTE** to classify topic for each question & answer, reaching **95% accuracy rate**

#### Twitter Malicious Bots Classification

- Created a bot classification model with a **91.7% accuracy rate**
- U **pandas** and **Beautiful Soup** to scrape data from Tweeter