

# Fiona Sijie Feng

fiona.sijie.feng@gmail.com | <https://fsfeng.github.io/home/> | <https://github.com/fsfeng>  
192 Wilson Ave, Brooklyn, NY | +1 (347) 757 1977

## Summary

- 4+ years experience producing novel insights with ML/NLP and statistical data analysis
- Track record of completing self-directed projects using unique synthesis of economic training and ML toolkit
- Extensive public speaking experience in presenting research and teaching undergraduates
- Proven writing proficiency producing research papers and policy reports

## Education

2013–2019	MS & Ph.D. in Economics, NYU Stern School of Business
2011	Bachelor of Commerce (Honours) in Economics, University of Melbourne (Australia)
2007–11	Bachelor of Commerce & Bachelor of Arts in Economics and Philosophy, University of Auckland (New Zealand)

## Work Experience

2013–2019     Ph.D. Research, NYU Stern School of Business

*The Proximity of Ideas: An Analysis of Patent Text Using Machine Learning* ([Paper](#))

- Converted 1.4 mil+ **unstructured** patent texts into meaningful numerical data using **unsupervised NLP** algorithms **Doc2Vec** (similar to Word2Vec) and **topic models** (Latent Dirichlet Allocation) with gensim and SciKit-Learn in **Python**
  - Pre-processed text with **feature selection** in SciKit-Learn, gensim, NLTK
- ETL process: extracted raw data from XML, JSON, and CSV; manipulated data in **Pandas dataframes**; stored data using **Hadoop** and **Spark Parquet** files
- Efficiently performed complex operations (calculating text similarity) for large datasets (1.4 mil+ rows by 100 columns) in 4 minutes using NumPy, SciPy, and SciKit-Learn
- Constructed and implemented **multivariate linear/nonlinear regression** models to measure the effect of innovators' geographic location on flow of ideas with Statsmodels
- **Visualized** estimates of geographic effects with matplotlib, seaborn, ggplot2 (in R)

*Job Satisfaction, Management Sentiment, and Financial Performance* ([Paper](#))

- Trained **supervised** algorithms on job review text to **predict** positive or negative job **sentiment** with 88% accuracy
  - Selected best performing algorithms with **hyperparameter optimisation**; implemented with XGBoost and **deep learning (neural networks)** in Keras/TensorFlow
  - Implemented **sentiment analysis** with **Google Cloud Platform**

- Procured novel data by **scraping** 1.2 mil+ job ratings and review texts of publicly listed companies, then merged with **structured** financial data from Compustat
  - Conducted **multivariate linear** and **panel (dynamic time series) regressions** to find significant relationships between job sentiment and firm's financial performance
- 2018 **Summer Research Associate, Federal Reserve Bank of Minneapolis**
- Designed **data experiments** to find novel evidence of:
    - Strategic bias in patent citation patterns using NLP text similarity
    - Significant influence of patent attorneys (over inventors) on selecting citations
- 2015–2016 **Research Assistant, NYU Center for Urban Studies Program**
- Extracted and managed data held in remote **SQL** server
  - Trained **clustering** algorithm on 100,000+ university job titles with k-means
  - Predicted job categories using **supervised NLP** algorithms (Multinomial Naive Bayes, Random Forest/Boosting models) with over 90% accuracy
- 2012 **Data Consultant, Logicalis Australia (Melbourne, Australia)**
- Analyzed **market research** data using **logit** and **probit** models to find important factors determining winning and losing contract bids for large Australian IT firm

## Technical Skills

Packages	Pandas, NumPy, SciPy, SciKit-Learn, gensim, Keras/Tensorflow, NLTK, BeautifulSoup, Matplotlib, Seaborn, Dask, Jupyter, Statsmodels
Data Stack	Hadoop, Spark Parquet, SQL
Programming	Python, R, Stata, Matlab, Linux scripting

## Other Information

Teaching	Prepared and taught classes as Teaching Assistant for Introductory, Intermediate, and Advanced Macroeconomics (University of Melbourne, 2011–2013); Introductory Microeconomics (NYU Stern, 2016–2017)
Awards	Kauffman Foundation Prize Finalist for Best Student Paper (2018), American Economic Association CSWEF Summer Dissertation Fellowship (2018), Internet Association Policy Research Fellowship (2017), Teaching Commendation at NYU Stern (2016)
Presented at	National Bureau of Economic Research (Boston), Federal Reserve Bank of Minneapolis, Roundtable for Engineering Entrepreneurship Research (Georgia Tech, Atlanta), Transatlantic Doctoral Student Conference (London Business School, UK), National Bureau of Economic Research Digitization Workshop (Stanford, Palo Alto), Department of Economics (NYU Stern), Department of Management and Organizations
Citizenship	New Zealand (Authorized to work for any US employer for 3 years)
Languages	English, Chinese Mandarin