# Leslie (Liang) Li

Github: github.com/lifinger

linkedin.com/in/leslie-love-data

Email: liangli919@hotmail.com

+1 (424) 527 5577

## Data Scientist

Over 2 years of project and work experience in Data Science and Machine Learning. A proven track record of analyzing emerging trends and changing customer behaviors to create effective business and marketing strategies. Proficient in machine learning, statistics and transferring business problems into analytical models.

**Languages:** Python, R, Scala, SQL | Hadoop, Spark | TensorFlow, Keras | Linux
**Packages:** Scikit-Learn, SciPy, NumPy, Pandas, Matplotlib | data.table, ggplot2, dplyr | MLlib, GraphX
**Specialties:** Machine Learning, Natural Language Processing, Optimization, Regression, GLM, Predictive models

## Professional Experience

**TalkingData - China's largest mobile data firm**                          Beijing, China
*Data Scientist*                                                            *Jan 2017 - Dec 2017*

- **WiFi category prediction**
  - Developed ETL processes to load and transform terabytes of data using Scala and Spark.
  - Established category label by removing outliers, hand engineered features and built predictive models (Logistic regression, SVM and Random Forest) to identify WiFi categories, supporting business case for site selection.
- **Community detection - Graph clustering to segment customers**
  - Devised a community detection model based on WiFi connection and user attributes by using spectral embedding and TF-IDF to construct similarity matrix, clustered similar users using HDBSCAN.
  - Validated and tested the above model by data scraped from website and running A/B test, successfully helped Nestle and New Balance to spot customers, saving over 10% advertisement investment.
- **Products clustering for recommendation**
  - Clustered clothes categories into sub-categories based on purchase history using collaborative filtering and graph clustering algorithm (Louvain) by Python.
  - Presented data driven insights to Bestseller using data visualization drawn by seaborn and ggplot, showing consumer buying behavior and providing modification suggestions for products positioning.

**Intelligent Computing and Machine Learning Lab**                          Beijing, China
*Machine Learning Intern, NLP*                                              *Aug 2016 - Dec 2016*

- **Machine translation in semantic parsing**
  - Devised an encoder-decoder neural network model augmented with attention mechanism in semantic parsing, mapping natural language to their meaning representations by weak supervision using TensorFlow and Python.
  - Utilized dynamic programming to infer latent variables and Bag-of-Words to reduce searching space.
  - Test the model on arithmetic domain which can successfully infer the correct logical forms and learn the word meanings, compositionality and operation orders simultaneously.

**ROJ Company**                                                             Turin, Italy
*Machine Learning Intern, Computer Vision*                                  *Feb 2016 - Jul 2016*

- **Seed detection for precision agriculture seeding machine**
  - Implemented Mean Shift algorithm to segment images and developed Bag-of-Features model to classify objects by employing SURF as feature descriptor and SVM as classifier, achieving classification accuracy 89.51%.
  - Tailored to achieve real-time detection function, and reduced approximately 25% resource waste in seeding process.

## Education

**University of California, Los Angeles**                                    Los Angeles, CA
*M.S. Business Analytics; GPA: 4.0/4.0*                                      *Expected Dec 2018*

Courses: Data Management, Prescriptive Models, Statistical Modeling and Learning, Optimization

**Polytechnic University of Turin**                                         Turin, Italy
*M.S. Mechanical Engineering; GPA: 92/100*                                  *Sep 2015 - Jul 2016*

Courses: Machine Learning, Operational Research, Numerical Modeling and Simulation

**Beihang University**                                                      Beijing, China
*B.E. Materials Science and Engineering; Minor in Statistics; GPA: 3.75/4.0*   *Sep 2011 - Jul 2015*

## Publication

Li L, Li P, Liu Y, et al. Logical Parsing from Natural Language Based on a Neural Translation Model[J]. CCIS781, PACLING 2017