LIZHI ZENG

E-mail: lizhizeng@ucla.edu Website: https://lizhiz.github.io Skype: lizhi_zeng@hotmail.com

EDUCATION

- **MSc, Electrical Engineering, University of California, Los Angeles**, United States 03/2017
  Major Field: Machine Learning
  Main Courses: Linear Programming, Pattern Recognition & Machine Learning, Stats Programming.
- **BEng, Electrical Information, Guangdong University of Technology**, China 07/2015
  Major Fields: Signal Processing, Control
  Main Courses: Program Design, Discrete Mathematics, Data Structure, Signals and Linear Systems.
- **Exchange Program, University of California, Berkeley**, United States Winter 2013

TECHNICAL SKILLS

- **Basics:** Python, R, Matlab, SQL, Java, C/C++, Linux **Data:** MySQL, SQLite, HDFS, MapReduce
  Toolkit: Scikit-learn, Numpy, Pandas, NLTK, Matplotlib **Other:** Git, LaTeX, Tableau, Microsoft Office

RESEARCH EXPERIENCES

- **Research Associate, Cedars-Sinai Medical Center**, Los Angeles, United States 07/2016 - 03/2017
  Biomedical Imaging Research Institute (BIRI)
  - Adjusted the homogeneity of the magnetic field with shimming coils to improve MR image quality. Analyzed magnetic data and reconstructed optimization function of calculating values of current. More than 20% improvement in global shimming.
  - Reduced the number of shimming coils for cost reduction. Implemented principal component analysis(PCA) to optimize cardiac MR shimming coil design. Reduced 70% of coils with a small increase(1.5%) in standard deviation.
  - Generated optimal shimming coil designs with mathematical optimization for different needs. Built up a framework in Matlab to obtain optimal solutions with the genetic algorithm.
- **Research Assistant, Guangdong University of Technology**, Guangzhou, China 06/2013 - 06/2014
  - Generated Rayleigh and Nakagami-m distributions for system performance studies. Implemented algorithm for generating bivariate Rayleigh and Nakagami-m distributed random variables(RVs) in Matlab.
  - Optimized the original algorithm with Brent's method and Newton's method. Achieved about 80% improvement in accuracy and 20% in run-time performance.

SELECTED PROJECTS

- **Cross-Device Entity Linking (CIKM Cup 2016)** 02/2017 - 03/2017
  - Identified user's identity at device-level with user-generated data from different devices. Generated a list of user pairs that each pair were predicted to be the same person. Formulated this task as a binary classification problem, and trained a classifier with GBDT and random forest to distinguish true user pairs from possible user pairs.
  - Document representation with TF-IDF metric. Feature engineering with calculating cosine similarity, correlation coefficient and other statistics on URLs, HTML title, Event ID, and time stamps. Gained fine-grained features by evaluating the significance of different URLs.
  - Remedied data imbalance with down-sampling. Reduced the computational cost by performing candidate filtering with k nearest neighbors.

- **PageRank with MapReduce** 02/2017 - 03/2017
  - Realized PageRank algorithm on Wiki dataset.
  - Transferred the relationship between each site into adjacency matrix. Computed the PageRank score based on adjacency matrix in Hadoop environment. Obtained sites' final ranks from converged PageRank scores.
- **Sequence Alignment** 10/2016 - 11/2016
  - Derived the best global DNA/RNA sequence alignment path.
  - Modeled the sequence alignment problem as a three-state HMM with gap opening penalty and gap extension penalty. Implemented Viterbi dynamic programming algorithm in Python to find the best global alignment path.
- **IMDb Database Exploration** 05/2016 - 06/2016
  - Explored the properties of actor/actress network and movie network.
  - Transferred the information from IMDb movie database into weighted directed graphs. Found communities with the Fast Greedy Newman algorithm. Predicted ratings of target movies with ratings of neighbor movies and same community movies. Added famous directors as new feature to improve model performance. The difference between predicted ratings and actual ratings are less than 1.5.
- **Popularity prediction on Twitter** 03/2016 - 04/2016
  - Predicted tweet activity based on hashtags.
  - Analyzed and visualized trends on Twitter by calculating statistics of tweets over different hashtags. Feature engineering. Improved model performance by reselecting features with significant test(T-test). The accuracy varied from 74% to 88% for different hashtags.
- **Statistical classification. Classified textual data from 20 Newsgroups** 02/2016 - 03/2016
  - Text categorization on 20 Newsgroups data set.
  - Excluded stop words, punctuations and different stems of a word.
  - Implemented TFxIDF metric to represent documents. Used Latent Semantic Indexing(LSI) to project data into topic/theme domain. Trained classifier with soft margin SVM with an accuracy of 94%.

COURSE CERTIFICATES

- Using Databases with Python by University of Michigan on Coursera. Certificate earned on January 21, 2017
- Using Python to Access Web Data by University of Michigan on Coursera. Certificate earned on January 11, 2017
- Python Data Structures by University of Michigan on Coursera. Certificate earned on December 26, 2016

AWARDS AND HONORS

- The Second Prize Scholarship (2013 & 2014)
- The Provincial-level Second Prize of the 13th Challenge Cup Contest (2013)
- The Third Prize of SEMIKRON Scholarship (2013)
- Model Student of Academic Records(2012 & 2014)
- National Scholarship (2012)
- The First Prize Scholarship (2012)

REFERENCES

Director of Magnetic Resonance Engineering
Department of Biomedical Imaging Research Institute, Cedars-Sinai Medical Center, Los Angeles, United States
Hui Han, Ph.D. hui.han@cshs.org