

Shijie Li

sli41@ncsu.edu | (402)-314-9216 | <http://jerry-shijie.github.io> | Raleigh, NC

EXPERIENCE

Software Research Development | LexisNexis Inc., Raleigh, NC

Customer Profile Verification By Text Information Mining On Emails

Jan. 2016 - May 2016

- Created Python back-end to preprocess labeled text datasets obtained by Amazon Mechanical Turk.
- Applied rule-based algorithm with Talon library as base-line to extract customer email signatures through regular expressions, named-entity recognition(NER) and knowledge-based rules.
- Manipulated text data by word vector and train various classifiers to improve the accuracy of email signature recognition with A/B testing. Applied SMOTE method on unbalanced data and increased precision by 10%.

Full-stack Open-source Development | UNC System, Raleigh, NC

Expertiza Web Application: A Peer-review Learning System

Oct. 2017 - Dec. 2017

- Designed a UI/UX monitor architecture and implemented with MVC framework in Ruby on Rails in TDD.
- Built front-end UI with jQuery and Ajax with RESTful web services, and applied D3.js library to visualize statistics of user behavior data for further research.
- Set up MySQL database to save interaction event data from front end to study user behaviors.

SKILLS

- Algorithms in Machine Learning, Deep Learning, Reinforcement Learning with mainstream frameworks.
- Programming languages: *Python, Java, R, SQL, C/C++* with major IDEs and tools.

PROJECTS

Feature Selection For Reinforcement Learning In Educational Policy Development

Mar.2017 - Apr.2017

- Modified *Python* Markov Decision Process Toolbox to compute Expected Cumulative Reward(ECR) of a Reinforcement Learning system measuring user learning performance by actions in educational contexts.
- Designed a greedy search algorithm to select optimal subset of features using Wrapper method.
- Analyzed relationships among optimal features and generate policy map for instruction decisions.

GMM-based Change Detection On Very High-resolution Satellite Imagery

Feb.2017 - Apr.2017

- Applied symmetric Kullback-Leibler(KL) divergence to temporal changes between Gaussian distributions.
- Built framework of Gaussian Mixture Model(GMM) class for clustering detected changes.
- Collaboratively performed *Python* code testing and parameter optimization for image processing and detected over 85% of significant changes labeled by human experts on temporal satellite images.

Library Database Administration And Management System Development

Aug.2015 - Nov.2015

- Designed E-R diagram to define entities and relationships among components of the library system.
- Created *Java* back-end with Oracle *SQL* to implement administration policy of resource reservations for various user priorities, and monitor usage and penalty statistics in the library database.
- Built command UI and collaborated on software testing and refactoring by Agile development.

EDUCATION

North Carolina State University (NCSSU), GPA: 4.0/4.0

Aug. 2015 - Dec. 2017

Master of Computer Science

University of Nebraska-Lincoln (UNL), GPA: 3.8/4.0

Aug. 2011 - May 2013

Master of Science, Physics

Shandong University, GPA: 87/100

Sep. 2006 - Jun. 2010

Bachelor of Science, Physics

CERTIFICATE

Microsoft Professional Program in Data Science

Dec. 2016 - Feb. 2017