Shijie Li

sli41@ncsu.edu | 402.314.9216

EDUCATION

NC STATE UNIVERSITY
PhD in Computer Science
2015 - 2019 Expected GPA: 3.9/4.0

UNIV. OF NEBRASKA-LINCOLN
Master in Physics
2011 - 2013 GPA: 3.8/4.0

SHANDONG UNIVERSITY
Bachelor in Physics
2006-2010 GPA: 87/100

COURSEWORK

SOFTWARE ENGINEER
Software Engineering
Operating System Principles
Database Management Concepts and Systems
Design and Analysis Of Algorithms
Computer Graphics
Computer Vision

DATA SCIENCE
Machine Learning
Automated Learning and Data Analysis
Artificial Intelligence I
Algorithms for Data Guided Business Intelligence
Machine Learning for User-Adaptive Systems

SKILLS

PROGRAMMING
Experienced:
Python • Java • C/C++ • R • SQL • HTML+CSS/JavaScript/jQuery/WebGL
Familiar:
Matlab • PHP • MySQL • Swift

TOOLS/APPLICATIONS
IntelliJ IDEA • Eclipse • PyCharm • Android Studio • WebStorm • RStudio

LINKS

Website: jerry-shijieli.github.io
Github: github.com/jerry-shijieli
LinkedIn: linkedin.com/in/shijie-jerry-li
Facebook: facebook.com/JerryLeeLSJ

PROJECT HIGHLIGHTS

MUSIC RECOMMENDER SYSTEM USING COLLABORATIVE FILTERING IMPLICIT FEEDBACK
- Clean and index the raw data from Last.fm using Resilient Distributed Dataset(RDD) and MapReduce model.
- Implement the collaborative filtering recommender model using Spark MLlib.
- Evaluate the model by top-K recommendation overlapping rate.

REAL-TIME SENTIMENT ANALYSIS OF TWITTER STREAMING
- Set up the data streaming pipeline using Kafka distributed streaming platform, Twitter streaming API and Spark Streaming.
- Code the BOW sentiment analysis method using Spark RDDs and MapReduce API.

TEXTUAL FEATURE EXTRACTION FOR SENTIMENT ANALYSIS
- Implement the text feature extraction models using Python package gensim.
- Implement the sentiment classification and evaluation procedures using Scikit-learn APIs.
- Data visualization using python Matplotlib package.

MATCHING ALGORITHMS FOR THE ADWORDS PROBLEM
- Simulate Google AdWords online advertising marketing mechanism using random query sequence.
- Implement Greedy, Balance and MSVV algorithms to maximize the total revenue from advertisers based on bid prices and budgets.

LIBRARY ADMINISTRATION AND MANAGEMENT SYSTEM
- Organize the library database using Oracle SQL queries.
- Implement Checkout and Return policy in Java using JDBC APIs.
- Design UI, reminder and late penalty messaging components.

RESEARCH EXPERIENCE

PERSONALIZED RECOMMENDER SYSTEMS
My current research focus on the construction of personalized recommender systems with mainly text analytics as well as other machine learning technology. I am developing algorithms and software to combine cutting-edge technology into powerful tools that can provide information accurately and efficiently with excellent user experience.

MODELING OF 3D TOPOLOGICAL INSULATOR
Computational research with professor Evgeny Y. Tsymbal aiming at understanding the fundamental properties of 3D Topological Insulator through first-principle calculation and tight-binging modeling. This research results in a published paper:

- Betancourt, J., **Li, S.**, Dang, X., Burton, J. D., Tsymbal, E. Y., Velev, J. P. (2016). Complex band structure of topological insulator Bi2Se3. Journal of Physics: Condensed Matter, 28(39), 395501.