WEIHAI SHEN

+1 480-793-8982 ◇ wshen24@asu.edu ◇ http://wshen.org

EDUCATION BACKGROUND

Arizona State University **AZ, USA**
Master of Computer Science July 2018 ∼ May 2020
GPA : 3.72/4.0

Nanjing University of Posts and Telecommunications **Nanjing, China**
Bachelor of Geographic Information System July 2010 ∼ June 2014
GPA : 89.3/100, Rank: 1/107

RESEARCH EXPERIENCES

Research in the scientific Visualization ToolKit VADER Lab
Supervised by Dr. Ross Maciejewski Sep. 2019 ∼ Present

- Implemented a WebSocket filter in the Topology-Toolkit to enable real-time data transfer of large datasets between browser and server memory
- Provided an out-of-box JavaScript API to receive and send object data asynchronously
- Used histogram-based visualizations to guide users to select important isovalues based the implemented WebSocket filter
- Focused on Data Visualization and Data Analytics research area

Watermarking on generative adversarial network APG Lab
Supervised by Dr. Yezhou Yang Feb. 2019 ∼ May 2019

- Implemented a sample code to visualize the insights from low dimensions
- Implemented the *Fast Gradient Sign Attack* to impose an adversarial attack on an MNIST classifier
- Attempted a new structure to hide a unique watermarking into generative adversarial networks for intellectual property protection
- Focused on Adversarial Attack research area

WORK EXPERIENCE

ByteDance Technology Co Ltd. Palo Alto, USA
Summer Intern Engineer, Email team May ∼ August 2019

- Implemented the email label system to keep users' inbox neat, tidy and easy-to-manage
- Optimized MySQL slow queries for speed and performance
- Implemented and maintained a Go-based RPC service on a daily basis

ByteDance Technology Co Ltd. Beijing, China
Senior Data Engineer, Data team June 2017 ∼ June 2018

- Led the team to design and accomplish a real-time big data analysis platform that can provide users *Daily Launch User* and *Daily New User* information on a large data using *no-sql, spark, hbase* and *python*
- Designed and implemented a data warehouse to monitor the long-term retention rate of users

- Rewrote distributed real-time processing code and made use of Flink to replace Storm framework

Software Development Engineer, User growth team June 2016 ∼ June 2017

- Built an intelligent ads platform integrating multiple ads engine APIs to control thousands of keywords for user growth marketing
- Implemented different models and strategies to boost user growth

Baidu Inc. Beijing, China

Software Development Engineer, Knowledge search team April 2015 ∼ June 2016

- Designed and built a big-data tool to analyze the performance of the iterative product
- Participated in the application performance optimization project, especially in the aspect of data storage optimization
- Implemented and maintained a php-based API service on a daily basis

58.com & Ganji Inc. Beijing, China

Software Development Engineer, Real estate team July 2014 ∼ April 2015

- Improved the website's search engine ranking greatly by optimizing metadata, publishing relevant contents and improving access speed
- Redesigned the indices and sharding of MySQL to accelerate database access speed
- Implemented and maintained iterative development on a daily basis

RESEARCH PROJECTS

Industrial decision-making platform Aug. 2018 ∼ Dec. 2018

Supervised by Dr. George Runger

- Implemented this project from scratch to online using MVC software design pattern
- Implemented an out-of-box script to import data from CSV to AWS database instances
- Provided configuration center to make the program more scalable

Using BERT for Qualitative Reasoning tasks Jan. 2019 ∼ May. 2019

Implemented in Natural Language processing course

- Implemented *Next Sentence Prediction* to predict the relationship between two sentences
- Transformed raw data into an understandable format that BERT can support
- Fine-tuned a pre-trained BERT model using training dataset

Image Classification on Fashion-MNIST dataset Aug. 2018 ∼ Dec. 2018

Implemented in Fundamentals of Statistical Learning course

- Implemented a 3-layer fully-connected neural network without any deep learning library to classify the fashion-MNIST dataset
- Compared the performance of different SGD algorithms in classification tasks

PATENTS

The big data real-time analysis platform CN108920516A

Weihai Shen; Junxiu Gao; Qi Tan

With the advent of the big data era, one of the most crucial issues the internet companies have to face is how to obtain real-time behavior information of users, which has given birth to real-time big data analysis. The traditional architecture simply analyzes data based on offline data, e.g hive or other warehouse platforms. However, traditional offline analysis cannot guarantee the timeliness of data. In this industrial patent, we creatively propose an architecture to provide users with real-time statistic information, even with huge data.

TECH SKILLS

- **Programming Language**: Python, Golang, PHP, JAVA, C++, Javascript
- **Web**: Nginx, Flask, Linux, Docker, Web framework
- **Data**: Hive, Spark, Hbase, Flink, no-sql, MySQL

COURSES

- **Completed**: Statistical Machine Learning, Natural Language processing, Operating Systems, Fundamentals of Statistical Learning, Perception in Robotics
- **Ongoing**: Distributed Database Systems, Software Security, Data Visualization

AWARDS & ACHIEVEMENTS

- Graduate Service Assistant — Sep. 2019 $\sim$ Present
- Graduate Service Assistant — Aug. 2018 $\sim$ Dec. 2018
- First-Class Scholarship — Oct. 2011
- First-Class Scholarship — Oct. 2011
- Best Student Model — Oct. 2011
- First-Class Scholarship — Nov. 2012
- Best Student Award — Nov. 2012
- Individual Course Award - Advanced Mathematics — Oct. 2013
- Second-Class Scholarship — Oct. 2013
- Excellent Graduation Thesis — July 2014