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

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 EXPERIENCE

The WebSocket filter in the scientific visualization toolkit Sep. 2019 ∼ -
Instructor: Dr. Ross Maciejewski, VADER Lab
- Implemented a WebSocket filter in the visualization toolkit to enable real-time interaction between a client and a server in C++
- Optimized data transmission on large scale between a client and server
- Provided an out-of-box javascript API to receive and send object data asynchronously
- focused on Data Visualization research area

Watermarking on generative adversarial network Feb. 2019 ∼ May 2019
Instructor: Yezhou Yang, APG Lab
- Implemented a sample code to visualize the insights from low dimensions
- Attempted a new structure to embed a private key into GAN architecture as a watermarking for intellectual property protection
- focused on Adversarial Attack research area

WORK EXPERIENCE

ByteDance Technology Co Ltd. Palo Alto, USA
Summer Intern Engineer May ∼ August 2019
Email team
- Implemented and maintained a Go-based RPC service on a daily basis
- Collaborated with QA and code owners to improve code quality
- Optimized MySQL slow queries for speed and performance

ByteDance Technology Co Ltd. Beijing, China
Senior Data Engineer June 2017 ∼ June 2018
Data team
- Led the team to design and accomplish a real-time big data analysis platform that can provide users DLU(Daily Launch User) information, even with huge data through 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 used Flink to replace Storm framework

Software Development Engineer June 2016 ∼ June 2017
User growth team
- Built an intelligent ads platform integrating multiple ads engine APIs to control thousands of keywords for user growth
- Implemented different models and strategies to boost growth

Baidu Inc. Beijing, China
Software Development Engineer April 2015 ∼ June 2016
Knowledge search team
- Designed and constructed 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 through no-sql and MySQL redesign
- Implemented and maintained a php-based API service on a daily basis

58.com & Ganji Inc. Beijing, China
Software Development Engineer July 2014 ∼ April 2015
Real estate team
- 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 PROJECT

Face Recognition for Privacy Protection Jan. 2019 ∼ May. 2019
Course: Perception in Robotics
- Used the pre-trained model to recognize the face object from the input image
- Implemented the face replacement algorithm
- Tracked the motion of a person using OpenCV

Using BERT for Qualitative Reasoning Jan. 2019 ∼ May. 2019
Course: Natural Language processing
- Implemented BERT Next Sentence Prediction(NSP) 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
Course: Fundamentals of Statistical Learning
- 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
- Achieved a better understanding of deep learning on neural network

PATENTS

The architecture for real-time big data analysis CN108920516A
Weihai Shen; Junxiu Gao; Qi Tan; Slides

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(like 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, Go, PHP, JAVA, C++, Javascript
- **Web**: Nginx, Flask, Linux, MySQL, Docker, Web framework
- **Data**: Hive, Spark, Hbase, Flink, no-sql

AWARDS & ACHIEVEMENTS (Undergraduate's Period)

- 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