HUIMING WANG

steve.wang@zju.edu.cn | (+86) 17826823528

EDUCATION

Zhejiang University Hangzhou, China

B.Eng. in Biosystem Engineering (Double Degree: Computer Science) Sep.2016 – Jun. 2020

- GPA 3.69/4.00
- Core CS Courses: C/Python/Java Language Programming, Discrete Mathematics and its Application, Data Structure Basis, Digital Logic Design, Object-oriented Programming, Advanced Data Structure and Algorithm Analysis, Database System, Operating System, Machine Learning,
- Core Mathematics Courses: Calculus, Linear Algebra, Mathematical Statistics, Probability Theory, Ordinary Differential Equations, Complex Variable Functions and Laplace Transform, Partial Differential Equations, Computer Simulation (Monte Carlo method, etc.)
- Awarded the Zhejiang Provincial Scholarship, the Research National Award, MCM/ICM Meritorious Winner

RESEARCH EXPERIENCE

Zhejiang University (Department of Computer Science) Hangzhou, China

Research Assistant to Professor Yueting Zhuang, Dean of the School of Computer Science Mar. 2018 – Present

Speech Text Alignment Based on Neural Machine Translation Model (Multimodal Machine Learning)

- Baseline: Trained an acoustic model with the audio data of Chinese. Used text to train a LM and then decoded the audio to obtain time information of words and sentences. Calculated the distance of the original text with generated text and did the alignment. And this baseline system only got an accuracy of 53%.
- Build a system having an "encoder-decoder" architecture where a sequence of an input audio is projected into a continuous low dimensional space and the output sequence is generated from this representation and alignment is done soon after the output sequence is generated. This end-to-end speech-to-text alignment got an accuracy of 75%. (We used Bi-LSTM to build our system)
- Try to use BERT which Google proposed last year to build another system and get higher accuracy now (use transformer rather than Bi-LSTM as Encoder).

Zhejiang University (Department of Computer Science) Hangzhou, China

Research Assistant to Professor Siliang Tang, Head of Natural Language Processing Group Sep. 2018 – Present

Multilingual Parsing from Raw Text to Universal Dependencies (Natural Language Understanding) and Question Answering in Context on QuAC

- Try to learn syntactic dependency parsers that can work in a real-world setting, starting from raw text, and that can work over many typologically different languages, even low-resource languages for which there is little or no training data, by exploiting a common syntactic annotation standard
- Use BERT to help build intelligent QA system, and try to find something new to modify the model that can help with this system.

WORK EXPERIENCE

Oracle (China) Software Systems Co., Ltd. Beijing, China

Big Data Development Intern & Solutions Specialist Intern Apr 2018 – Jul 2018

- Use the basic knowledge of database design, NoSQL and distributed computing to gain insight and prediction of massive data through data mining;
- Use knowledge of various Oracle databases, PL/SQL usage and the application of Java and Python in database management to help built an online Oracle Big Data Automation Management Platform which can be used in automatic cleaning, inspection and performance monitoring of data through software system, and extract data from it into data sets and provide them for deep neural network training and use;
- Manage the schedule of project using Oracle's project management solution P6 and participate in the development of project management solutions, dedicated to managing project portfolios, controlling costs, and managing cross-project writing and communication;
- Work with the project manager to create a software development lifecycle using Primavera P6, covering the pre-project, documentation, development, commissioning, handover, testing, and product lifecycle phases, and participating in a real-world development. Clarify the entire process of software development.

Eigen Technology Co., Ltd. Hangzhou, China

Back-end web development intern & natural language algorithm intern Jul 2018 – Sep 2018

- Learn the use of flask development frameworks such as flask and sqlalchemy, responsible for the architectural design, functional development and code maintenance of the company's application system background;

- Before transferring to NLP algorithm intern, participate in the development of the company's internal management personnel management OA system and officially launched as a core member of the project;
- After transferring to NLP algorithm intern, participate in the development and optimization of the related model of Eigen Technology Machine Learning, achieve the algorithms and experiments in the papers, and become familiar with the basic algorithms of machine learning and Python programming (including using machine learning library of Python).

COURSE PROJECT EXPERIENCE

Programming in Python Hangzhou, China
Box Office Prediction System May. 2018

- First try from data collection, to data processing, to multi-layer neural network construction, to training algorithm design and implementation of projects that are completed by one person;
- Crawling movie information from multiple domestic and international movie websites through crawler technology, and converting text information into a digital matrix for processing calculations
- Complete the implementation of the neuron class and build a three-layer neural network for calculation. Finally, the box office information can be predicted by several different pieces of information (movie type, director, actor popularity, etc.) and achieve an accuracy of 82% (the rate is obtained by the size of the 10,000 test box office deviations).

Database System Hangzhou, China
Mini Database Design Jun. 2018

- From 0 to 1 using C++ to implement mini-database which can use basic SQL to complete the addition, deletion, and deletion of the data, including index creation and deletion;

Machine Learning Hangzhou, China
Machine Learning/Deep Learning Algorithm Project (including SVM and CNN) Oct. 2018

- Practice SVM and use the lib-SVM library to complete the video-based detection of speech or not (action recognition) procedures;
- Resolve the King's problem and optimize the motion recognition program with artificial neural networks;
- Train the model of CNN using ResNet, and try to judge whether two photos coming from one person

Operating System Hangzhou, China
Operating System Design Jan. 2019

- Complete code execution for three modules including process management, memory management, and file system. The completed operating system is capable of performing basic functions.

ADDITIONAL INFORMATION

Extracurricular Experiences

- Proprietor of Zhejiang University Caring-Heart Campus

Computer and Language Skills

- C/C++, MATLAB, Java, Python, R, SQL, Data Structure and Algorithms, Data Scraping, Machine Learning, LATEX, Monte Carlo method.