Yao Dou

7545 McCrimmon Pkwy, Cary, NC • (206)-379-5769 • douy@cs.washington.edu

EDUCATION

University of Washington, Seattle, WA

B.S. in Computer Science

- **Expected Graduation:** June 2021
- **GPA:** 3.91/4.0
- **Relevant Coursework:** Natural Language Processing, Algorithm, Data Structures and Parallelism, Software Design and Implementation, Machine Learning, Data Visualization, Artificial Intelligence
- **Extracurricular Activities:** Co-founder of the UW chapter of Triangle Fraternity
- **Honors and Awards: Microsoft Endowed Scholarship** 2020 – 2021; **Dean's List** for Autumn 2017, Winter 2018, Spring 2018, Autumn 2018, Winter 2019, Spring 2019, Spring 2020

EXPERIENCE

Research Intern

Allen Institute for AI, to begin January 2021

Research Assistant

University of Washington, April 2019 – Present

With Prof. Yejin Choi

- Do OOF project and write the paper (to submit to ACL 2021)
- Did MultiTalk project and wrote the paper (accepted to AAAI 2021)
- Presented paper in lab meeting

Teaching Assistant

CSE446 (Machine Learning) at University of Washington, Spring 2020 – Present

With Prof. Kevin Jamieson and Prof. Jamie Morgenstern

- Design assignment question and quiz section worksheet
- Hold weekly office hours and answer online questions
- Attend meetings and grade assignments

Data Science Intern

Noonum, INC, June 2020 – September 2020

- Improved Named Entity Recognition (NER) pipeline 3.5% on overall performance and 12% on PRODUCT entity performance
- Deployed the pipeline to the product
- Wrote a detailed corresponding tutorial for future engineers

Research Assistant

University of Washington Autonomous Flight Systems Laboratory, January 2018 – September 2018

- Flight tested aircrafts and ensured they are tuned appropriately for safe and reliable flight
- Formatted maintenance files of all drones
- Improved interface for the LIDAR sensor

PROJECTS

Ontology of Failure (OOF)

July 2020 – Present

- Develop a new structured, crowd-sourced error annotation scheme called OOF
- Design a tool for annotating text with OOF annotations
- Quantitatively and qualitatively analyze the text generation systems in terms of OOF on different text domains
- To submit to **ACL 2021** conference as first author

MultiTalk

April 2019 – June 2020

- Name of the paper: **MultiTalk: A Highly-Branching Dialog Testbed for Diverse Conversations**
- Collected a large dataset of highly-branching written conversations
- Introduced a scoring algorithm based on the weighted bipartite graph matching
- Did emotion classification on the dataset
- Computed metrics like BLEU on the dataset and generations
- Finetuned the GPT2 model to generate sentence that can lead to a future response with an intended emotion
- Accepted to **AAAI 2021** conference as first author

Exploration of Seattle's housing prices

February 2020 – March 2020

- Built an interactive tool that find desired district of housing by user input preference list
- Analysis the data of Seattle's housing prices, traffic, crime stats, schools, stores, restaurants, companies and transportation
- Did data visualization about the relationship between housing prices and other factors such as traffic, crime stats and etc.
- Collaborated with 3 team members

Campus map route finding application

February 2020 – March 2020

- Built a route-finding tool by implementing Dijkstra's algorithm
- Built a graphical user interface (GUI) that visually draws routes on map

Interactive Shiny Based Stock Analysis Application

February 2018 – March 2018

- Project awarded as the **Best Analysis Project**
- Led a 4-person team to implement an interactive Shiny based application which including functions: visualizing stock data of S&P 500, comparing different companies' stocks and predicting stock in a period
- Collaborated with 3 team members

Building Alexa Skills Kit to voice control a thermostat

December 2017 – January 2018

- Developed a custom Alexa skill to control the thermostat with multi-turn interactions
- Used ESP8266 to record the temperature and insert it into a Cloud database periodically
- Alexa reads or changes data from the database according to the instruction given by people to Alexa