Shiyuan (Eric) Zhou

Tel: (+1) 4254493966 | Github Profile | Email: zhoueric0603@gmail.com
Address: Unit 4805, 25 Telegram Mews, Toronto, Ontario, M5V 3Z1

EDUCATION

University of Toronto（U of T） **September 2019 – May 2023**

Major: BSc (Honours), Data Science Specialist, Department of Computer Science

GPA: 3.87/4.00

Awards: Dean's List 2020 - 2023

EXPERIENCE

BaseBit.ai - Intelligent Medical System **April 2021 – September 2021**

Developer, (Advisor: Dr. Lintao Zhang)

- Improved matrix factorization model accuracy by 7% using Funk SVD algorithm in recommendation system.
- Assisted in constructing intelligent data services including Data Cleaning, Wrangling, and Visualization.
- Used HTML, CSS, and JavaScript to develop a patent application process website for research team.

RESEARCH EXPERIENCE

Peking University – IPP Lab **March 2023 - Present**

*NLP Researcher***,** (Advisor: Prof. Yinyin Zang)

- Sentiment Analysis on Mental Health related social media posts for Mental Illness Prevention.
- Retrieved and compared keywords from scrapped social media posts by TF-IDF, TextRank, KeyBERT, LDA, and Yake.
- Analyzed social platform users' perspectives of mental health services through a Large Language Model (LLM) based on RoBERTa, and concluded that about 30% of users had a negative attitude towards mental health services.

Purdue University - Interactive Intelligent Systems Lab **May 2022 - Present**

CV and ML Researcher, (Advisor: Prof. Tianyi Zhang)

- Rapid Image Labeling via Neuro-Symbolic Learning (RAPID).
- Paper is accepted by KDD 2023 (https://github.com/Neural-Symbolic-Image-Labeling/Rapid).
- RAPID achieved more than 80% accuracy in 4 different image labeling tasks, which includes glaucoma detection, bird species detection, etc.
- Reduced 85% of training data needs during the sensitivity analysis of RAPID via multi-criteria based active learning.
- Implemented a Python Flask backend server containing APIs for data transmission between layers and database manipulation using PyMongo. The whole system is prepared to be submitted to CHI 2024.

University of Toronto - FORCOLAB **September 2022 – June 2023**

NLP Researcher, (Advisor: Prof. Shurui Zhou)

- Aligning Documentation and Stack Overflow through Constrained Decoding with Weak Supervision.
- Paper is prepared to be submitted to ICSME 2023.
- Gathered and augmented training data through web-scrapping Python documentation with Beautiful Soup.
- Achieved 81.8% accuracy using a fine-tuned GPT-2 based model for classifying Stack Overflow posts into Python documentation topics.

University of Toronto **February 2022 – May 2022**

Data Science Researcher, (Advisor: Prof. Meredith Franklin at U of T)

- A Life Expectancy Prediction Project via Machine Learning in R (Full Report).

COURSE PROJECTS

StylEase – A Fashion Mobile App (Github) **January 2023 – May 2023**

- Deployed backend on Heroku and exception tracking on Sentry. Set up CI/CD on Github Actions.
- Integrated fashion item detection algorithm in backend by utilizing Hugging Face API to scan clothing.
- Developed Flask server to respond to client requests and manage databases, achieving 90% backend test coverage.

SKILLS

Languages: Python, R, Java, HTML, CSS, JavaScript

Tools: Pytorch, scikit-learn, Numpy, SQL, Hugging Face API, Pandas, MongoDB, Spacy, OpenCV, Matplotlib, Flask, NLTK, Beautiful Soup, PyMongo, dplyr, Tidyverse, lme4, ggplot, Figma