Yue Teng

ytteng0212@gmail.com | +61 466618591 | www.linkedin.com/in/yueteng-359b26244

EDUCATION

Master of Data Science **Feb 2022 - Present**
The University of Melbourne
WAM H1

Bachelor of Data Science **Feb 2019 - Dec 2021**
University of Melbourne
WAM H1

Foundation Studies **Aug 2017 - Dec 2018**
Trinity college - The University of Melbourne
Foundation

AREAS OF SPECIALIZATION

- Data Processing and analysis
- Data cleaning and visualization
- Statistics
- Natural Language Processing
- Cloud computing platform builder
- Artifical intelligence
- Machine Learning
- Pytorch toolsets, libraries such as sklearn, numpy, pandas
- SQL based database systems, knowledge about non-sql based databased systems, recovery and concurrency control, SQL programming
- Skilled use of libraries such as ggplot2, MASS in R to implement models within GLM, random forest, subset model selection and so on
- Report authoring and presentation skills
- Classification with different neural networks (BERT, CNN, LSTM...)

RELEVANT EXPERIENCE

Software Engineering developer **Apr 2023 - Jun 2023**
The University of Melbourne, Melbourne, VIC
- Accomplished automation system for resources allocation on Melbourne Research Cloud platform by Ansible.
- Crawl toots from Mastodon with Mastodon API and fault tolerance ensured.
- Accomplish CouchDB cluster and retrieve data with MapReduce functions.

Consulting report on patients undergoing coronary artery **Oct 2022 - Nov 2022**
The University of Melbourne, Melbourne, VIC

- Analysed features related to arteries and attributes of patients age, gender to predict disease a patient belongs to
- Led R to do classification and error analysis done by confusion matrix
- Used ggplot2 library to do data visualization for relationship between features and labels
- Wrote a consulting report to present experiment preprocessing and results

Author Detection Model for given documents   Sep 2022 - Oct 2022
The University of Melbourne, Melbourne, VIC
- Used google colab to design and develop author classification system by LSTM, CNN, MLP models on Pytorch. Achieve top 1 on both public Kaggle leader board and private one
- Managed in data cleaning, debugging, data visualizations to do feature selection and feature engineering

Rumour Classification Model for COVID-19   May 2022 - Jun 2022
The University of Melbourne, Melbourne, VIC
- Developed a rumour detection system by BERT, SVM models on pytorch. Extract useful features and apply logistic regression model on to analyse nature of rumours spread on Twitter
- Managed in debugging, testing, feature extraction and rumour analysis

Quantitative Analysis of Taxi Trip Record Data   Sep 2021 - Nov 2021
The University of Melbourne, Melbourne, VIC
- Explored a 4.5GB CSV format dataset consisting of trip records of taxis in New York City to find attributes are highly correlated with fare amounts
- Do data cleaning and feature engineering are by Python. Use libraries folium and geopandas to visualize data on map
- Implemented different regression models on R to check feature importance
- Wrote a report by latex to present experiment results

Business Analysis of CSL Data   Jul 2021 - Aug 2021
The University of Melbourne, Melbourne, VIC
- Explored a dataset consisting of experiment results in a hospital and apply data cleaning and visualization on it
- Implemented data cleaning by checking its consistency based on specific data metrics (missing ratio, BMF Score), then feature engineering is applied based on requirements
- Made a 30-minute presentation and arrange meetings with team-mate on discord. Write logs after each meeting

Data Analyst & Engineer   May 2021 - Jul 2021
Shanghai J&J certified public accountants firm, Shanghai, China
- Analyse and compare price of public cloud service providers based on computing resources offered.
- Design basic structure of cloud application for client.

KEY STRENGTHS
- Proficiency in Python, R

- Advanced data cleaning, analysis techniques
- Multiple languages: Chinese (Native), English (Advanced)
- Leadership with strong responsibility
- Advanced written and communication skills
- Research and development
- Project management