Yiting Zheng

dd.zheng001@gmail.com

EDUCATION

Master of Information Technology - Research Pathway Jul. 2024 – Now
Monash University, Australia

WAM: 82/100 GPA: 3.5/4.0

Main Course: IT Research Methods (87/100, HD), Foundations of Data Science (90/100, HD), Cloud Computing and Security (85/100, HD)

Bachelor of Computer Science and Technology Sept. 2017 – Jun. 2021
College of Online Education, Nankai University

WAM: 87.57

RESEARCH EXPERIENCE

Exploring Learning Performance through Knowledge and Interaction Analysis in Medical Training Dialogues (In preparation)
Researcher, supervised by Guanliang Chen at Monash University, Clayton

Research Interests:

(1) Conducted a systematic integration and analysis of the dynamic interplay between cognitive processes and socio-emotional dynamics in medical training dialogues, advancing interdisciplinary research at the intersection of dialogue-based learning, socio-emotional interaction, medical education, and computational analysis.

(2) Dedicated to transforming rich but context-dependent insights from qualitative research—such as "pathophysiological reasoning" —into quantifiable indicators that can be processed by computational methods (e.g., Natural Language Processing, NLP), through the design of detailed coding schemes and multi-level computational analysis.

Core Research Methods:
NLP, ML, Epistemic Network Analysis (ENA), Sequential Pattern Mining (SPM), Process Modelling, Cluster Analysis, Sequence Analysis/Process Mining.

PROJECT EXPERIENCE

Comprehensive Feature Engineering and Predictive Analysis of FLoRA Dialogue Usefulness Scores
Researcher, supervised by Guanliang Chen, at Monash University, Clayton

Engineered and integrated both structural (e.g., turn count, word ratio) and semantic (TF-IDF, weighted keywords, PCA) features from raw chatbot-student dialogue data to model interaction quality.

Developed and optimized an XGBoost model using grid search and cross-validation , achieving a 20.79% improvement in RMSE and a 22% improvement in MAE compared to baseline prediction results.

Employed Shapley Additive exPlanations (SHAP) to provide both local and global interpretability for the model's predictions , gaining deep insights by combining SHAP results with XGBoost's built-in feature importance metrics.

Core Tech: XGBoost, SHAP, Scikit-learn, PCA, TF-IDF, Linear Regression, Random Forest.

Hospital Discharge Trend Prediction and Resource Optimization: A Feasibility Study
Researcher, supervised by Guanliang Chen, at Monash University, Clayton

Integrated and preprocessed large-scale heterogeneous data, combining hospital financial reports (CMS) with public health surveillance data (CDC).

Designed and compared multiple time-series models, including ARIMA and enhanced Linear Regression models incorporating exogenous features (e.g., ILI rates, DSH payments).

Achieved high predictive accuracy, with the final feature-enhanced model reaching a MAPE of 0.98%, significantly outperforming baseline and simpler models.

Core Tech: R, auto.arima, lm, ggplot2, Time-Series Analysis, EDA.

Residential Market Price Prediction and Analysis of Influencing Factors Using K-Nearest Neighbors (KNN) Regression
Researcher, supervised by Guanliang Chen, at Monash University, Clayton

Built and tuned a KNN Regression model to predict property prices, implementing advanced feature engineering using regular expressions to extract features from text descriptions.

Ensured model robustness through rigorous 10-fold cross-validation, systematic K-value grid search, and comprehensive data preprocessing (one-hot encoding, scaling).

Core Tech: R, caret, grepl, KNN, Regex, Cross-Validation.

BirdTag: A Serverless Media Analysis and Management Platform on AWS
Researcher, supervised by Guanliang Chen, at Monash University, Clayton

Designed and deployed a scalable, serverless cloud application. Utilized AWS Lambda and Python to implement AI-driven auto-tagging of media files upon S3 upload and provided a comprehensive media management service via a RESTful API.

Core Tech: Python, AWS, Serverless Architecture, RESTful APIs, Vue3, TailwindCSS.

WORK EXPERIENCE

Full-Stack Development Engineer at Beijing Corefire Technology Co. Nov. 2020 - Jun. 2024

Executed full-stack development for multi-platform applications, utilizing Java for backend services and Vue3, TypeScript for front-end interfaces.

Engineered backend services and data pipelines, building RESTful APIs to power data-driven features such as merchant revenue analysis.

Transformed complex business data into interactive visualization dashboards using Echarts and Antv to support merchant decision-making.

TECHNICAL SKILLS

- **Languages:** Python, R, JavaScript/TypeScript, Java
- **Machine Learning:** XGBoost, Random Forest, Linear Regression, KNN, ARIMA, SHAP
- **Analytics & NLP:** Epistemic Network Analysis (ENA), Sequential Pattern Mining (SPM), Cluster Analysis, TF-IDF, PCA
- **Cloud & Engineering:** AWS, Serverless Architecture, Node.js, RESTful APIs, Vue3, mpx
- **Databases & Tools:** MySQL, Oracle, Postman, Figma, QGIS, Echarts, Antv, Git

ADDITIONAL INFORMATION

Certifications

Project Management Professional (PMP) Certification

Patents

Heng, Deyu, and Zheng, Yiting. (2024). 医用多舱全自动清洗机 (Medical multi-chamber fully automatic cleaning machine). China Patent No. ZL 202430714714.3.