Yueyin Tan

tanyueyin@outlook.com · (+852) 56158538 · LinkedIn · Personal Website

Education

The Hong Kong University of Science and Technology	**2022.9 - 2023.12 (expected)**
Dept. of Computer Science and Engineering & Dept. of Mathematics	Hong Kong SAR
Master of Science in Big Data Technology	**GPA:** 3.8/4

- **Relevant Course:** Machine Learning (A), Mathematical Methods for Data Analytics (A)

Wuhan University	**2018.9 - 2022.6**
School of Mathematics and Statistics & Hongyi Honor College	Wuhan, China
Bachelor of Science in Information and Computational Science	**GPA:** 3.4/4

- **Honors:** Honor Graduate Award (2022), Merit Student (10%, 2021), Outstanding Student (23%, 2020)
- **Scholarships:** Outstanding Undergraduate Scholarship (2019, 2020 & 2021)
- **Relevant Course:** Convex Optimization (92), Numerical Analysis I & II (96 & 90)

Harvard University	**2021.6 - 2021.8**
Summer School	Remote
	GPA: 3.7/4

- **Relevant Course:** Intro to Quantitative Methods (A-)

Work Experience & Projects

Internship

Reghub New Technology (Beijing) Co., Ltd.	**2021.6 - 2021.9**
Algorithm Research Intern	Beijing, China

- Build and tune Random Forest and AdaBoost models to classify sanctioned transactions
- Develop a novel feature engineering method based on data insights, resulting in a 5% increase in Recall
- Build BiLSTM model for Named Entity Recognition, extracting information from SWIFT messages
- Analyze transaction data by SAS and SQL for anti-money laundering regulations

Projects

Graphical Model-Based Algorithms for Clustering on Spatial Transcriptome Data	**2022.3 - 2022.6**
Bachelor's Dissertation	

- Conduct dimension reduction and data denoising by Principal Component Analysis (PCA)
- Implement Louvain algorithm on the similarity network derived from Similarity Network Fusion
- Achieve a 20% raise in Adjusted Rand Index and a 25% raise in Normalized Mutual Information
- Leverage Cloud Computing source from the university and reduce the runtime by 80%

Sentiment Analysis with Deep Learning	**2021.8 - 2021.9**

- Preprocess labelled text data and split datasets by stratified method
- Set up the necessary dependencies, and frameworks to utilize BERT for text classification and analysis
- Set training loops for fine-tuning and predict the sentiment of a sentence

Numerical Schemes for Partial Differential Equations (Link)	**2021.5 - 2021.6**

- Implement Finite Difference Scheme, Crank-Nicolson Scheme, ADI Scheme, PC Scheme and LOD Scheme
- Visualize the distribution of errors for each scheme and estimate the order of errors

Credit Rating Model based on KNN Algorithm	**2020.11**
The First Prize (Provincial) in Chinese Undergraduate Mathematical Modelling Contest	

- Analyzed invoice data and quantified credit risk with Entropy Weighted method on risk indicators
- Implement K-Nearest Neighbors (KNN) algorithm for credit rating
- Calculate credit lines by multi-object programming, maximizing interests and minimizing loss

Skills & Interests

Language: Cantonese (Native), Mandarin (Native, Upper-Level II Certificate), English (Fluent, TOEFL 103)
Programming: Python (Advanced), MySQL (Advanced), C (Intermediate)
Software Tools: Linux CLI, Git, Docker, Spark, Microsoft Office
Interests: Tennis (3.0), Kayaking (3-Stars Award), Photography (Collection), VLOG