WANG CHANGQIN

(+65) 90500157 | changqin.wang@u.nus.edu | github.com/archiewang0716

EDUCATION

National University of Singapore **Singapore**
Master of Science in Data Science and Machine Learning *Aug 2021 – Present (exp. Jan 2023)*

- **CAP:** 4.64 / 5.0
- **Coursework:** Machine Learning, Deep Learning, Computer Vision, Natural Language Processing, Data Mining

Chongqing University **Chongqing, China**
Bachelor of Engineering in Computer Science and Technology *Sep 2017 – Jun 2021*

- **GPA:** 87.3 / 100 (Top 15%), **Major GPA:** 90.4 / 100
- **Honors/Awards:** MCM Honorable Mention; First-class Merit Scholarship; Outstanding Student; Excellent Graduate Student
- **Coursework:** Data Structure, Design and Analysis of Algorithm, Database Systems, Computer Graphics, Intelligent Systems

SKILLS

Databases: MySQL, Redis, MongoDB
Programming Languages: Python, SQL, R, MATLAB, Java, C++, C#, HTML/CSS/JavaScript
Libraries, Frameworks & Tools: PyTorch, TensorFlow, Keras, Scikit-Learn, Matplotlib, Seaborn, NLTK, Git, GCP, AWS
Certifications: Machine Learning (Stanford CS229), Natural Language Processing (Stanford CS224n)

INTERNSHIP EXPERIENCE

NTUC Enterprise Digital **Singapore**
Data Science Intern *Jul 2022 – Present*
Designed a recommender engine that provides explanations for EDM grocery recommendations to 5,000k customers nationwide

- Constructed an aspect-based fine-grained opinion extraction model based on GTS-BERT for brand-level opinion mining
- Implemented a centroid text summarizer based on word embedding that provides customer feedback for each product

Asian Institute of Digital Finance **Singapore**
NLP Research Intern *Nov 2021 – Apr 2022*
Designed a default prediction language model for Indonesian motivated by issue-centric sentiments in social media and business press for over 73,000 global corporations

- Constructed an aspect-based sentiment analysis model to predict credit risk for each company by masking company names based on the sequence-labeling model, achieving an F1-score of 70.58%
- Provided credit risk sentiment of global corporations as alternative data for the default prediction platform, and predicted default events with a stacking classifier using LightGBM and XGBoost

Chongqing University **Chongqing, China**
Research Assistant *Feb 2021 – Jun 2021*

- Project Title: *Stacking Ensemble Convolutional Neural Networks with Split-Attention and Residual for advanced recognition*
- Implemented ResNeSt on three transfer learning tasks – image classification, scene recognition, and face recognition
- Conducted experiments and compared the performance of ResNeSt with CNNs, including ResNet and VGGNet
- Explored Stacking Neural Networks (S-NN) and conducted an ablation study to examine the performance of different S-NNs

The S-NN model with ResNeSt-50 and VGG-16 achieves 74.7% accuracy on the advanced face recognition datasets

Sichuan Hwadee Information Technology Co., Ltd **Chengdu, China**
Data Engineer Intern *Jun 2020 – Aug 2020*

- Performed data cleansing and EDA on the bank marketing dataset, and imputed missing values by a KNN classifier
- Refined the unbalanced dataset leveraging over-sampling, and explored feature engineering on the dataset
- Adopted the RF and SVM algorithms and conducted experiments by 10-fold cross-validation with grid search

Accomplished an analytical data system for bank marketing data, visualizing the predicted campaign strategies (MCC: 0.83)

PROJECT EXPERIENCE

Attention-based Graph Neural Networks for Multi-class Fake News Detection *Dec 2021 – May 2022*

- Augmented Long Short-Term Memory networks (LSTMs) to adopt the attention mechanism
- Explored a model that attaches the attention-based LSTM layers as input to Graph Convolutional Networks (GCNs)
- Improved 4-way classification by 2% on the Labeled Unreliable News dataset (F1-score: 98.7%) and out-of-domain dataset

Movie Recommendation Systems Based on Collaborative and Content Filtering *Aug 2021 – Nov 2021*

- Built recommendation algorithms from scratch, trained and predicted the movie ratings on the MovieLens-1M dataset
- Implemented collaborative filtering algorithms and content-based filtering algorithms, including User-CF, Item-CF, SVD++ (SVD with Implicit Feedback), and TF-IDF (based on item profile correlation)