Yu-Chieh Chen

+1(626)628-5058 | yuc399@ucsd.edu | LinkedIn | GitHub

EDUCATION

University of California San Diego, *La Jolla, CA*   *Sept. 2017-present*
Bachelor of Science expected in Jun.2021; Major in Data Science, Double Minor in Cognitive Science & Economics
- Overall GPA: 3.79/4.0; Major GPA: 3.85/4.0
- Provost Honors for eight quarters

RESEARCH EXPERIENCE & PROJECTS

Institute of Information Science, Academia Sinica   Research Assistant Intern   *Taipei, Taiwan*
Fake News Identification with CNN models   *Jul-Sept.2020*
- Proposed a framework for the detection of fake news from news headline spread on social media using hybrid of convolutional neural networks and long-short term recurrent neural network models
- Used CNN model to process the dataset (65,639 headlines from content farm and national news agency) with and received 93.99% accuracy on testing data (4,462 headlines)
- Conducted exploratory data analysis and compared content farm headline with national news agency headline, found out three features (wording embedding, sentiment score, parts-of-speech) and applied bigram for more sentiment scores
- Scrapped and cleaned 33,000+ documents as the training data for tfidfVectorizer and Multinomial Naïve Bayes Classifier to predict categories of the news, achieved 0.92 F1-score
- Wrote a paper *"Headline Diagnosis: Manipulation of Content Farm Headlines"* as first author, published on Taiwan Academic Network Conference (TANET) 2020

Prediction of Air Quality Based on K-Nearest Neighbor (KNN)   *Jul-Aug. 2019*
- Used KNN algorithm to select the time and space-related monitoring stations, trained and verified the KNN models to predict PM2.5 values and monitor air quality and implemented data visualization

Convoluted Neural Network (CNN) on CIFAR-10   *June 2020*
- Performed transfer learning by utilizing two layers of ResNet18, receiving 80.75% accuracy on testing data

Machine Learning Analysis on Airline Flights Delay Prediction   *Fall, 2019*
- Proposed the model such as Logistic Regression, K Nearest Neighbors, Decision Tree, etc. to proactively detect flights delay and compared the performance of the models with the 0.62 accuracy and 0.573 F1 score

Data Visualization of Suicide Rate   *Fall, 2019*
- Analyzed the highest suicide rate among different age group between 2005 to 2015
- Performed results by visualizations using HighCharts, JavaScript, HTML/CSS and Python.

WORK EXPERIENCE & ACTIVITIES

Halıcıoğlu Data Science Institute, *La Jolla, CA*   Tutor   *Apr.2019-present*
- Helped in grading and assignments design, and wrote AutoGrader scripts in Python for grading for four courses （DSC20, 30: Data Structure | DSC 80: Practice in Data Science | COGS 185: Advanced Machine Learning Methods）
- Held tutor hours each week, answering students' questions about class materials and debugging code

Chinese Engineering Society, *La Jolla, CA*   Career Development Officer   *Apr. 2019-present*
- Held events for more than 60 students to connect with experienced seniors sharing interview and job techniques
- Invited and communicated with experienced seniors and wrote advertisements for the events

UCSD TripleC (Chinese Computer Community), *La Jolla, CA*   Member   *Sept.2018-Dec.2019*
- Contributed to the group in building a project-based platform of incubating undergraduate start-ups
- Developed a https://github.com/ElvishElvis/Real-time-Face-Detection-Recognition, created local database and pre-store the features for each face for reducing the time consumption, extracted facial landmark with Dlib library and applied supervised learning to predict facial emotion, and achieved highest validation accuracy of 89.43% with logistic regression

Deloitte Touche Tohmatsu Limited, *Taipei, Taiwan*   Audit Department Intern   *July 2018*
- Specialized in audit software, EMS, to update the financial statement
- Audited and analyzed the financial statements of clients and the changes over the quarters

SKILLS

Computer Skills: Java, Python, HTML/CSS, JavaScript, SQL, R, Spark, PostgreSQL, Redis, Cassandra, AWS, Tableau etc.

Language: Chinese Mandarin (Native); English (Fluent)