Ali Majedi

Data Analyst • +60 192877844 • aamajedi@gmail.com

[Portfolio](#) • [Linkedin](#) • [Github](#)

An ambitious and self-motivated person who enjoys finding unique approaches for solving different tasks. I can convey my ideas clearly and concisely in a team environment while being able to deliver quality projects with minimum supervision.

Skills

- Python
- Scikit-learn
- Seaborn
- MySQL

Education

2016/09 – 2020/07

Bachelor of Science: Artificial Intelligence
University of Malaya- Malaysia
3.6/4 CGPA

2015/09 – 2016/06

Pre-University: Mathematics Major
Iranian School in Kuwait - Kuwait
94.05/100 Average Grade

Work History

2020/05 – 2020/08

AI Engineer Internship
Key ASIC Bhd, Malaysia
- Combined different pre-processing techniques and transfer learning on blood cell dataset to achieve over prediction 64% accuracy
- Improved Bilirubin prediction model accuracy to 90% by using multi-linear regression model

2018/08 – 2019/10

AI Engineer Part-time
Fylix, Malaysia
- Added user feedback loop to improve the performance of data extraction tool
- Designed database structure based on information acquired by meeting different field experts and analyzing the data
- Developed custom string-matching algorithm to increase data extraction performance up to 95%

- Expanded supported document formats to Excel, Word and PDF
- Crawled, cleaned and processed twitter data to develop analysis tool

Projects

Hybrid movie recommender
- Collected, cleaned and processed data from IMDB and MovieLens
- Personalized the recommendation by tweaking the 1500 data points after each recommendation loop
- Improved recommendation performance by calculating genre correlation and including it as an attribute
- Improved recommendation speed by applying Bag-Of-Word and KNN and reducing dimensions with KNN in the background to achieve 82% accuracy in focus group

Forecast cryptocurrency price movement using historical data and Reddit submissions
- Scraped Reddit submissions using multiple APIs dating back to 2017
- Scraped cryptocurrency prices from Binance for over 10 coins in multiple timeframes in a custom database
- Preprocessed and encoded the submissions using BERT algorithm
- Achieved over 65% forecast accuracy by training an LSTM model on historical data and Encoded features
- Optimized the database and algorithm to stream data from database for maximum efficiency during training as well as enabling real-time scraping and prediction

Reddit Analysis tool
- Developed pipeline to collect data from multiple APIs, clean and process it on a daily basis
- Designed the database architecture to support millions of data with less than a second retrieval time on any query
- Implemented APIs and UI with interactive charts to view information on any keyword or subreddit from any time
- Applied Mann Kendell to identify trends