Chenghao Wu

cwu137@syr.edu
3152890209
New York, United States
www.chenghaowu.com
linkedin.com/in/chenghao-wu-65475917a
github.com/ChenghaoWu316?tab=repositories

SKILLS

Python3 | Data warehouse | R | SQL | Tableau | PowerBI | Adobe illustrator | Pyspark | NLTK | AWS | WEKA | ETL | Machine learning | Scrum | Data warehouse | Data pipeline | Accounting | Financial statement analysis | Excel | Kafka | Drill | Azure Data Factory | Docker | Git | Hadoop | Hive | Rest API | Cassandra | Scikit-learn

EDUCATION

M.S. in Business Analytics
Syracuse University, Whitman School of Management
01/2021 - 12/2021 *GPA: 3.9/4 Syracuse, New York*

M.S. in Applied Data Science
Syracuse University, School of information studies
01/2019 - 12/2020 *GPA: 3.7/4 Syracuse, New York*

B.S. in Mechanical Design & Manufacturing
Xi'an University of Technology
08/2014 - 07/2018 *GPA: 3.7/4 Xi'an Shannxi*

WORK EXPERIENCE

Cloud Data Engineer
Hoyos Integrity Corporation
06/2021 - Present *Ft.Lauderdal, FL*

Tasks/Achievements

- Conducted security phone subscriber database design project, created dataverse tables in Power Platform, as well as identified the security level for our clients in power platform.
- Implemented data ETL using Azure data factory, then built dashboard reports for phone inventory, order fulfillment, subscriber's data usage using Power BI and Looker for executive board to make business decisions
- Built mobile app for assigning work orders and monitoring work orders' fulfillment status in Power App. People can use the app to view the working processing stage details.
- Retrieved data from various resources like Postgres database or other severs using API and conducted data pre-processing, reformatting using, pyspark and pandas.
- Designed data flows in automate 365 to automate the data population and data update inside our database, assisted front-end UI developer to build phone admin web portal for our clients.

Graduate Teaching Assistant
Syracuse University School of information
01/2020 - 12/2020 *Syracuse, New York*

Tasks/Achievements

- Assisted the instructor with Data Mining and Natural Language Processing courses, which includes machine learning techniques and algorithms needed for both supervised and unsupervised machine learning using R and Python3.
- Weekly troubleshooting and grading for students. Algorithms taught include: Random Forest, Support Vector Machines and Artificial Neural Networks as well as k-Means clustering and J-48 Association rules. Practical applications include Segmentation, Anomaly Detection, Recommendation Engines and Predictive Modeling, Sentiment Analysis, Speech Recognition, Chatbots and Language Translation.
- Contributed to the development of appropriate teaching materials to ensure content andmethods of delivery meet learning objectives.

WORK EXPERIENCE

Researcher

Syracuse University School of information

01/2020 - 12/2020 *Syracuse, New York*

Tasks/Achievements

- Conducted research on E-sport data by using PySpark. Collected Data by using public API and cleaned the data. Used Random forest to determine the main factors of winning a game and used Association Rule to find out the most frequent ban and pick champion combination for each team by building pipelines. Demonstrated the analyzing result of match data by designing Tableau dashboard.
- Developed chatbot for optimizing faculty teaching effectiveness by collecting textbook data and creating instances and logical relationships using IBM Watson studio. The chatbot content includes 10 articles and more than 200 feasible human-chatbot interactions.

Technical Consultant

Information Technology Experiential Learning Lab

04/2019 - 12/2019 *Syracuse, New York*

Tasks/Achievements

- Supported and maintained a virtual machine environment that contains over 2,000 VMs, the Active Directory, network management utilities, Firewalls, NAS, DHCP, DNS, and other networking applications and services.
- Troubleshoot existing ITELL network infrastructure including Cisco router and switches as well as support Windows, Linux servers, and other workstations.
- Improved system efficiency by consulting with end-users and provided innovative solutions.

SELECTED PERSONAL SIDE PROJECTS

Airbnb pricing strategy analysis. (01/2020 - 05/2020)

- Collected data by using Airbnb Seattle public API and preprocessed the whole dataset using visualization techniques.
- Detected the most important feature for predicting Airbnb price by using spark with random forest, GBT and Lasso regression model.
- Performed sentiment analysis and 5 classes rating score prediction based on guests' reviews by using logistic regression, random forest, SVM, LSTM, and Naïve Bayes model and reached 90% accuracy.
- Developed a recommendation system for airbnb users to find out similar housing resources by using k-means and collaborative filtering based on host description.

Riot games database design. (05/2020 - 12/2020)

- Designed a database of season11 league of legends items and Champions by using Microsoft Azure Data studio and SQL, including conceptual and logical design of database. The database contains 170+ champion and all s11 new items.
- Created transaction-safe procedures that can recommend players to purchase the most suitable item for their champion in different combating phase by setting association tables. Each player is able to record his purchasing history in database as well.
- Deployed the database on Azure cloud and developed a mobile application by using Power App. Players can register on the app to view all recorded match details and purchase history , as well as full champion and item lists.

Yelp's hotel rating prediction (05/2020 - 08/2020)

- Collected review data from more than 100 hotels in the United States, cleaned the data by using Pandas and implement data exploration by using seaborn to find out the possible relationship between the frequency of key words and given rate star.
- Implemented multi-classification tasks by using unigram, bigram tokens converted by TF-IDF vectorizer. Feed the features into Naives Bayes, Linear SVC, random forest as well as LSTM model. Generated models that can predict rating score based on guests' reviews, the highest accuracy achieved 80%.
- Determined the best strategies to optimize the profit for the hotel managers by building Tableau dashboards.

Venture capital analysis. (01/2021 - Present)

- Assisted an engineering start-up to select their business structure and make a growth plan, including product price setting, identify the potential risk and raise funding.
- Made sales forecast for the first 5 years and generated income statement, balance sheet and cashflow statement for each fascial years.
- Evaluated the outside capital investment need and helped the CEO communicate with venture angel investors to raise fund.
- Evaluated the pre-money valuation and exit valuation in 5 years for the company by using burkes method, score card method and VC method.

CERTIFICATES

Tableau Data Analyst (06/2021 - Present)

AWS Certified Cloud Practitioner (06/2021 - Present)

AWS Certified Solution Architect (10/2021 - Present)

Wharton School Entrepreneurship Study (10/2021 - Present)