Limei Hou

Cleveland, OH 44109 | limeihou1@gmail.com
https://linkedin.com/in/limei-hou | https://github.com/limeiCode
https://codingpurple.com

Data Analyst

Data analyst with a passion for condensing large data into a simple to understand story. Adept at working independently and collaborating with teams across multiple functions to break down concepts in layman terms to diverse audiences. Bring experience obtained through a cross-functional organization to streamline data analysis, reporting and auditing. Committed to continuous learning. Prepared to excel in solving complex problems. Fluent in English and Chinese.

Technical Skills

Languages: Python(Pandas,NumPy), VBA, SQL, C#
Databases: Microsoft SQL Server, postgreSQL, mongoDB, SQLite
Visualization: Matplotlib, Seaborn, Plotly.js, D3.js, Leaflet.js
Business Intelligence Tools: Tableau
Web: JavaScript, HTML/CSS, Bootstrap, Dashboarding, Web scrapping, Flask, Jinja
Machine Learning: LinearRegression, LogisticRegression, RandomForest, KNN, SVM, DeepLearning, NeuralNetwork, K-means
Big Data: Hadoop(MapReduce), PySpark, NLP, AWS(EC2, S3, RDS), Cloud ETL
Other: APIs, JSON/GeoJSON, Beautiful Soup, Git/GitHub
SAS Certified Specialist: Base Programming Using SAS 9.4 Certificate Verification #2T9G1ZT2CBF11FS4

Experience

Risk Management Data Analyst / Quant Modeling Associate 2012 – 2019
KeyBank Cleveland, OH

- Implemented commercial C&I new LGD and KMF risk rating models using VBA to assist in the underwriting of new loans and assist management in predicting changes in portfolio quality and the subsequent financial impact.
- Designed SQL Server database schemas and developed C# Entity Framework applications harvesting data in excel models into SQL databases for further analyzing.
- Provided monthly model rating exception analyzing reports for different line of businesses helping modelers and credit officers monitoring as well as validating models and making business decisions.

Associate Professor 2002 – 2007
Liaoning University Shenyang, China

- Taught Data Structure, Database Fundamental and Visual Foxpro Programming courses.
- Coach five students in designing and developing software application each semester.
- Won Outstanding Teaching Award given by student organization.

Projects

Home Equity PD Machine Learning Web Model
https://github.com/limeiCode/HomeEquityPDMachineLearningWebModel

- This project is designed to develop a machine learning model to predict default or not of a home equity loan and develop a Flask web application for deploying the machine learning model using Python language.
- Three models Scikit-learn Logistic Regression, SVC and Keras Deep Learning are built through serveral steps including preprocessing the dataset, separating the data into training and testing data, selecting features, buulding model, tuning model hyperparameters using GridSearch, comparing model performance and predicting. Dataset used includes loan characteristics and delinquency information for 5,960 home equity loans.

ETL And Analyze Amazon Cloud Review Data
https://github.com/limeiCode/ETLAndAnalyzeAmazonCloudReviewData

- This project is designed to perform the ETL process on Amazon review data in the cloud and upload the final PySpark DataFrame to an AWS RDS instance.
- First creating tables based on review data's structure in AWS RDS PostgreSQL instance then within ZELP note book environment using PySpark to Extract Amazon review dataset from AWS S3 bucket into PySpark dataframe then Transform the dataset and Load the DataFrames into the AWS RDS PostgreSQL instance. Last investigating whether Vine reviews are free of bias using PySpark.

World Wide Weather Analyzing And Web Dashboard Demonstrating
https://github.com/limeiCode/WorldWideWeatherAnalyzingAndWebDashboardDemonstrating

- This project is designed to analyze how weather changes as get closer to the equator(Latitude=0).
- By making CityPi API call and OpenWeatherMap API call to generate weather data of over 500 world wide cities from 1500 sample geo-coordinates drawn from Numpy Uniform Distribution and these data are plotted using Matplotlib and analyzing results are demonstrated in an interactive and responsive visualization web dashboard.

Life Trend Analyzing And D3 Multi-Axes Plotting
https://github.com/limeiCode/LifeTrendAnalyzingAndD3MultiAxesPlotting

- This project designed to show the nine possible correlations of Life Trend analyzing among the three X Axile variables and three Y Axile varialbes by building only one scatter plot using D3.js.
- By using Scalable Vector Graphics(SVG) to create a responsive web page which scales itself dynamically according to the size of the window and the image is not distorted by resizing browser so it can be viewed in browser and mobile devices.

Web Scraping Mission To Mars Data
https://github.com/limeiCode/WebScrapingMissionToMarsData

- This full stack Flask application does web scraping using Selenium framework from several websites to gather Mars related information like latest Mars news, current featured Mars images and diameter mass facts of Mars.
- When user send scrape request to Flask server these data are scraped and parsed by using Beautiful Soup then are stored in Mongo DB, when user send browsing request to server the Mars data in Mongo is queried and rendered into Jinjia Template webpage returned to user.

Education

Case Western Reserve University**,** Cleveland, OH
Data Analytics Bootcamp Certificate
An 24-week intensive program focused on gaining technical programming skills in Python, JavaScript, SQL, postgreSQL, SQLite, mongoDB, noSQL, SQLAlchemy, Jason, GeoJSON, Pandas, Matplotlib, Seaborn, Plotly.js, Beautiful Soup, D3.js, Leaflet.js, Tableau, HTML, CSS, Bootstrap, Flask, Machine Learning, Big Data, Cloud ETL, AWS, Hadoop, Spark.

Cleveland State University**,** Cleveland, OH
Master degree, Computer Engineering
Ph.D course credits in Computer Engineering
Finished all Ph.D courses and passed Ph.D qualify test.

Liaoning University**,** Shenyang, China
Master degree, Computer Science
Bachelor degree, Computer Software

References Available Upon Request