ZHUOQUAN CHEN

DATA SCIENTIST

New York, NY | (+1) 718-300-0078 | zhuoquan1223@gmail.com | https://zhuoquan-chen.github.io/portfolio

EDUCATION

- General Assembly | Data Science Immersive Course, Manhattan, NY — Jan 2021
- Brooklyn College | BS. Computer Science, Brooklyn, NY — Aug 2018 – May 2020
- Borough of Manhattan Community College | Computer Science, Manhattan, NY — Aug 2016 – May 2018

SKILLS

- Programming: Python (Numpy, Pandas, Scikit-learn, Scipy, Keras) / SQL / Java / HTML / CSS / JavaScript / Node.js.
- Modeling:
  - Supervised Learning: Linear and Logistic Regression / K-Nearest Neighbors (KNN) / Decision Tree / Random Forest / Super Vector Machines (SVM) / Naive Bayes / Neural Networks / Convolutional Neural Networks (CNNs)
  - Unsupervised Learning: K-Means Clustering / Principal Component Analysis (PCA)
- Data Visualization: Python (Matplotlib, Seaborn, Plotly)
- Natural Language Processing (NLP) / Time Series Analysis

PROJECTS

SENTIMENT ANALYSIS & CLASSIFICATION

Skills: NLP / CountVectorizer / Sampling / Naive Bayes / Logistic Regression / GridSearchCV / Confusing Matrix

The purpose of this project is to determine a customer whether satisfied or not with a product of Amazon by analyzing their feedback.

- Naive Bayes Model with 91.7% accuracy for new data, with 89% accuracy for sensitivity (recall).
- Logistic Regression Model with 94.7% accuracy for new data, with 89% accuracy for sensitivity (recall).

MARKET SEGMENTATION

Skills: Unsupervised Learning / K-Means / PCA / EDA / Data Visualization / Data Cleaning

The purpose of this project is to help the Marketing Department of banks to segment the consumption market by clustering customers based on their credit card consumption data, then launch ads to target it precisely to grow or extend business.

- The customer dataset was divided into 5 clusters by K-Mean algorithms.
- The dimensionality of the dataset was reduced to 2 dimensions for visualization.

CHEST DISEASE CLASSIFICATION

Skills: Residual Neural Networks / Convolutional Neural Networks / Transfer Learning / Drop out / Data Visualization

The purpose of this project is to achieved automatically detecting and classifying different types of chest diseases such as healthy, covid-19, bacterial pneumonia, and viral pneumonia based on X-ray images..

- Model testing validation accuracy: 80% with 50 epochs.
- The model has weak performance at precision for covid-19 & bacterial pneumonia symptoms with 68% & 55% accuracy.

EXPERIENCE

BMCC | College Assistant, Financial Aid Department, Manhattan, NY — Oct 2017 - Present

- Worked for data entry and management of more than 100,000 financial aid applications data each academic year.
- Used Python to clean and filter duplicate data and ensure students' names will not repeat to appear on the calling list.

GENERAL ASSEMBLY | Apprentice , New York, NY — Sep 2020 – Jan 2021

- 500+ hours full-time immersive educational program strengthening Data Science skills including Python, SQL, data cleaning, data visualization, regression models, classification models, web-scraping, APIs, NLP, advanced supervised learning, unsupervised learning, time series analysis, and statistics.