

ZHUOQUAN CHEN

Data Scientist | Machine Learning

718-300-0078 | Brooklyn, NY 11229 | zhuoquan1223@gmail.com | LinkedIn: <https://www.linkedin.com/in/zhuoquanchen>

Portfolio: <https://zhuoquan-chen.github.io/portfolio> | GitHub: <https://github.com/ZhuoquanChen/Data-Science-Projects>

With an education in both computer science and data science, I believe that every problem has a solution and many more than one. With a graphic design background, I believe more in the power of imagination and creativity, because everything starts with imagination before it becomes reality. As a data scientist with a diverse background, I like to think from different perspectives and use data science methods and models to help people make more informed decisions.

TECHNICAL SKILLS

- Data Processing: data cleaning and data visualization, PCA, dimensionality reduction, feature engineering
- Machine Learning: classification model, regression model, clustering, NLP, Time Series Analysis, Neural Networks
- Methods: Statistical Distributions, Bayesian Analysis, p-Values, Hypothesis Testing
- Programming Languages / Environment: Python (Scikit-learn, Numpy, Pandas, Matplotlib, Seaborn, Plotly), SQL, Java, Jupyter Lab, Google Colab, PySpark

PERSONAL PROJECTS

Sentiment Analysis & Classification

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

This project used NLP technology that can immediately judge a customer whether satisfy a product or not with their sentiment analysis. It can greatly reduce manual judgment, thus saving lots of time and money for enterprises.

- Visualized the phrases base on key judgment results
- Naive Bayes Model with 91.7% accuracy for new data
- Logistic Regression Model with 94.7% accuracy for new data

Chest Disease Classification

Skills: Residual Neural Network / CNNs / Transfer Learning / Drop out / Data Visualization

This project used transfer learning on the pre-trained ResNet (Residual Neural Network) model to automatically diagnose different types of chest diseases such as healthy, covid-19, bacterial pneumonia, and viral pneumonia through chest X-ray images.

- Model training validation accuracy: 93.27% (50 epochs)
- Model testing validation accuracy: 80% (50 epochs)
- Model had weak performance at metric of precision for type 1 (covid-19) and recall for type 2 (bacterial pneumonia) respectively with 64% and 50% accuracy

EXPERIENCE

BMCC (Borough of Manhattan Community College), New York, NY

Oct 2017-Present

College Assistant (Financial Aid Department)

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

EDUCATION

- General Assembly | Data Science Immersive Course *Sep-Dec 2020*
- Brooklyn College | B.S. Bachelor of Science in Computer Science *Aug 2018-May 2020*
- Borough of Manhattan Community College | A.S. Associate of Science in Computer Science *Aug 2016-May 2018*

EXTRACURRICULAR

- **Deep Learning Team | BMCC** *Sep 2019-Jan 2020*
Joined Professor Tang's Deep Learning team researching in CNNs including Neural Network, Softmax, SVM, Regularization, and Fully-connected Neural Network.
- **CUNY Hackathon 2019 | Baruch College** *Dec 2020*
My team's idea for this competition is to use CNNs to identify the distance and direction of nearby objects, so that blind people can know what is near them.