

AIMAN CHUGHTAI

✉ chughtaia1014@gmail.com

☎ 347-968-0051

📍 Staten Island, New York,
10301

in

<https://www.linkedin.com/in/aiman-chughtai>

📄 AimanChughtai

Skills

PROGRAMMING

Python

SQL

MACHINE LEARNING

Linear Regression

Logistic Regression

Classification

Natural Language
Processing

Random Forest

Gradient Boosting

Naive Bayes

Decision Trees

LIBRARIES AND FRAMEWORKS

Pandas

NumPy

BeautifulSoup

Selenium

NLTK

Gensim

Spacy

TextBlob

StatsModel

SciKit Learn

VISUALIZATION

Matplotlib

Seaborn

Plotly

Tableau

MS Excel

DATABASES

PostgreSQL

Education

Macaulay Honors College:

College of Staten Island

B.S Biology

Experience

Metis

Data Scientist

Completed an immersive data science bootcamp with a focus on Python programming, machine learning, statistical modeling, data visualization, project design, and communication

Designed and built end-to-end projects to utilize data science in consumer market scenarios

Center for Developmental Neuroscience- College of Staten Island

Researcher

Compiled, analyzed, and interpreted data in MS Excel on the impact of intensive rehabilitation experiments to recommend treatment options for spinal cord injuries

Prepared posters in MS Powerpoint to be presented during Society for Neuroscience National Conference

ResCare - Continuing Education

Research Analyst

Compiled data about educational certificate programs that may fit client needs

Staten Island University Hospital

Patient Experience Coordinator

Compiled and analyzed data in MS Excel on patient wait times to recommend workflow optimizations in key tasks to decrease overall patient visit duration and improve patient experience

CityMD

Medical Scribe

Collected and analyzed trends in patient charts in order to classify symptoms and translate findings in electronic health records with accuracy, consistency, and completeness, to be reviewed by an onboard healthcare provider

Projects

Need Car Insurance? - Using Classification Algorithms to Predict Cross-Buying

Leveraged SciKit Learn in Python to compare the performance of multiple classification models (Logistic Regression, Decision Tree, Random Forest, XGBoost) to predict whether health insurance policy-owners would buy car insurance from their insurance provider, based on their customer statistics. The best model was Logistic regression, which increased F2 score (recall) by 12%, effectively capturing 12% more interested customers than the baseline model.

A Coronaspiracy: Using NLP to Debunk Covid-19 Vaccine Myths on Reddit

Scrapped over 2000 comments and posts on Reddit's Anti-Vaccine threads using Reddit's Praw API to discover common myths surrounding the COVID-19 vaccine in order to aid NYC Health with their newest ad campaign. Facilitated NLTK and Spacy for topic modeling using Non-Negative Matrix Factorization, and Latent Semantic Analysis, vectorizing both with CountVectorizer and TF-IDFVectorizer, to categorize vaccine myths based on popularity. Analyzed sentiment surrounding common myths using Textblob. Visualized results in Matplotlib and Seaborn.

Predicting IMB Employee Attrition

Utilized Scikit Learn's logistic regression classification algorithm in order to predict whether or not an employee would attrit from their position at IBM. Explored relationships between features in MS Excel using pivot charts and vlookups. Created Tableau dashboard for visualization purposes in order to recommend solution paths to reduce employee attrition.

Sold! - Zillow House Price Predictor

Scrapped nearly 2000 house listings on Zillow using BeautifulSoup in order to predict home values based on house features. Utilized Statsmodel's Ordinary Least Squares model and Sklearn's linear regression model to make predictions on listing price. Achieved R2 score of .60 and mean square error of about 120K. Visualized results in Matplotlib and Seaborn.