

VAISHNAVI RAJPUT

Brooklyn, NY - (332) 207-9599 - vrjpt11@gmail.com - [linkedin.com/in/vaishnavirjpt10](https://www.linkedin.com/in/vaishnavirjpt10) - github.com/vrjpt10 - vrjpt10.github.io/

EDUCATION

New York University - New York, USA

Master of Science in Computer Engineering

Sep 2021 - May 2023

Dr. A. P. J. Abdul Kalam Technical University - Noida, India

Bachelor of Technology in Electrical and Electronics Engineering

Aug 2016 - Sep 2020

PROFESSIONAL EXPERIENCE

Development Data Analyst Intern - Block Convey, New York

Jun 2023 - Present

- Analysed data, fine-tuned model, and collaborated with the team to build an advanced machine learning system integrated with blockchain technology for the early detection of brain tumors

Data Science Intern (Remote) - British Airways, New York

Feb 2023 - Mar 2023

- Developed a pipeline for data cleaning, analysis, and a prediction model that predicts customer buying behavior with an 85% accuracy for personalized marketing strategies
- Employed web scraping to extract 7GB of customer reviews from the Skytrax review website and performed sentiment analysis leveraging NLP techniques to understand customer sentiments comprehensively and identify key areas of concern

Graduate Assistant - NYU, New York

Oct 2022 - May 2023

- Administered over 300 students during doubt-clearing sessions and exams for undergrad courses - Data Structures and Algorithms and Introduction to Programming

Software Engineer - YM Grad, India

Jun 2020 - Jul 2021

- Developed an administration system for service plan selection, including CRUD operations, and improved SEO ranking of the website by 200%

ACADEMIC PROJECTS

Recommender System - NYU Big Data

Mar 2023 - May 2023

- Built a music recommendation system on the ListenBrainz dataset (30GB) and fabricated a popularity baseline model and implemented collaborative filtering to improve song suggestions per user
- Incorporated light FM and fast search to compare single-machine execution with cluster-based execution achieving a 20.6% reduction in model fitting time and a 30.1% reduction in accuracy calculation time
- Deployed and fine-tuned alternating least square method in Spark to learn latent factor representation of users to increase recommendation precision by 5.7%

Churn Prediction - NYU Big Data

Apr 2023 - May 2023

- Performed EDA on a large-scale telecommunication customer data containing over 70,000 data points to identify the factors that influence customer churn
- Leveraged decision tree, random forest classifier, and PCA to construct a prediction with up to 94% accuracy
- Designed a Flask application to predict customer churn confidence level and enable real-time access to the churn prediction model

Mu-GAN: Generating Adversarial Modified Music - NYU Machine Learning for Cybersecurity

Aug 2022 - Dec 2022

- Formulated a DC-GAN extension architecture to generate novel music of a particular genre or modify existing music to a new genre using 7000 songs from GTZAN, FMA, and GACMIS datasets
- Produced adversarial modified genre transferred clips of duration more than 30 seconds that can bypass copyright claims and fine-tuned synthesized music for 7 distinct genres employing Librosa library

ETL System Development for Data Warehousing (Capstone Project) - NYU Advanced Project

Aug 2022 - Dec 2022

- Constructed an e-commerce website using HTML, CSS, PHP, and Bootstrap, enabling efficient user data collection
- Implemented ETL system using MySQL and PostgreSQL relational databases and NodeJS and created a data warehouse. Applied necessary data cleansing, validation, and normalization processes for smooth integration of user data into the ETL pipeline
- Generated data visualizations and analyses with Tableau, providing insights on consumer demographics, buying tendencies, and sales trends for informed decision-making and improved business operations

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, HTML, CSS, C++

Libraries and Frameworks: Numpy, Pandas, Matplotlib, Seaborn, Pytorch, Sklearn (Scikit-learn), Scipy, OpenCV, Pyspark

Big Data Tools: Apache Spark, MapReduce, Hadoop

Other Tools: Tableau, Microsoft Excel, Git, DataProc, AWS SageMaker, MATLAB

Databases: MySQL, PostgreSQL, PL/SQL

Data Science & Analysis: Data Warehousing, Data Analytics, Data Modeling, ETL Processing, Data Visualization, Database Querying