Sajid Hussain

(551)334-8760 sh739@njit.edu www.linkedin.com/in/sajidcodes sajidcodes.github.io

Experience

Applied AI Program **June 2022 – December 2022**

Student

- Built an end-to-end project on Shark Tank data using complex SQL queries on MS SQL Server.
- Performed hypothesis testing, One Sample, Two Sample t-test, Confidence Interval in Python.

Machine Learning Intern **April 2022 – June 2022**

iNeuron Pvt. Ltd

- Built a model that predicts the person's premium for health insurance, thus saving costs.
- Selected Gradient Boosting with an RMSE score of 0.388 and R-squared score of 0.818 vs Random Forest with an RMSE score of 0.396 and an R-squared score of 0.810.

Freelance Projects

Kaggle Competition **April-May 2023**

- Successfully achieved a score of 70.34 in the *AMP®-Parkinson's Disease Progression Prediction* hosted by Kaggle, using protein and peptide data measurements of patients.

HR Dashboard *[Image] [Github]* **December 2022**

- Designed and developed an HR Dashboard using Tableau to identify attrition rate for women.

Recommendation system **December 2022-present**

- Performed Exploratory data analysis on **100 million** ratings from 480,000 randomly chosen, anonymous Netflix customers over 17000 movie titles and computed similarity matrices.
- Presently building a recommendation system for a fast-growing startup with more than 200,000-page visits.

Credit Card customer Attrition rate **November 2022**

- Experimented with ensemble models, Logistic regression, Decision tree, Random Forest, and Light GBM to forecast and reduce customer attrition by analyzing more than 10,000 data points and 21 features.
- Achieved an F1-score of 0.97 as compared to other algorithms.

Research Experience

New Jersey Institute of Technology **Sep 2023- Dec 2023**

Debiasing Pre-trained Language Models

- Selected for a research project under the supervision of a distinguished professor Dr. Guiling (Grace) Wang for debiasing Pre-trained language models.

BioSMART Center **June 2023-present**

Research laboratory at NJIT

- Pain Perception from Brain fMRI • Applying a data-driven technique—Constrained Principal Component Analysis for fMRI (fMRI-CPCA) to perform whole-brain analyses of fMRI BOLD signal data.• Whole-brain functional images were collected on a 3 T Philips Achieva TX scanner at Columbia University's Program for Imaging in Cognitive Science (PICS). • This work may have far-reaching consequences in demonstrating the potential of data-driven, whole-brain functional network techniques for the analysis of pain imaging data.

Driverless Car: Autonomous Driving Using Deep Reinforcement Learning **Mar 2023**

- Reviewed Driverless Car: Autonomous Driving Using Deep Reinforcement Learning.

Education

New Jersey Institute of Technology **Sep. 2022 – May 2024**

MS Data Science *Machine learning, Applied Statistics, Data Mining*

Visvesvaraya Technological University

BE Computer Science

Skills Summary

Languages	SQL, Python, and R.
Frameworks	Natural language processing (NLP), Numpy, Pandas, Seaborn, MatplotLib.
Tools	AWS, Sagemaker, Hadoop, MS Office.