HAMZA AZIZ

Data Science Associate

+91-8447644188 | hamzaaziz822@gmail.com | linkedin.com/in/hamza2325
Ghaziabad, Uttar Pradesh

CAREER OBJECTIVE

Innovative and scientifically rigorous computer science student with a significant data science learning experience to bring to the table. With a team oriented attitude, I am eager to contribute my abilities in quantitative modelling and experimentation to enhance the user experience around the world.

SKILLS

Programming Languages

Python | NumPy | Pandas | Matplotlib | Seaborn | Scikit Learn | Tensorflow | C/C++ | HTML | CSS | JavaScript | SQL

Mathematical Skills

Linear Algebra | Probability | Statistics | Multivariable Calculus | Regression Techniques | Data Clustering

Data Visualization

MS Excel | Google Sheets | Tableau

VOLUNTEERING

Vice President

RubiX Club @ ABESIT — 2021 - Ongoing

INTERPERSONAL SKILLS

Communication
Easy to have conversation with colleagues and professional talks

Intuitive
Thinks rationally yet a big picture dreamer

Empathetic
Tries to make people comfortable around me by being empathetic to them

FIND ME ONLINE

GitHub
github.com/hamza2325

EDUCATION

Bachelor of Technology

ABES Institute of Technology
2019 - Ongoing
Ghaziabad, Uttar Pradesh
- Computer Science and Engineering

SGPA
8.35 / 10.0

Intermediate - CBSE

Green Field Public School
2018 - 2019
Ghaziabad, Uttar Pradesh
- PCM

%
81.8 / 100

High School - CBSE

Green Field Public School
2016 - 2017
Ghaziabad, Uttar Pradesh

CGPA
10.0 / 10.0

PROJECTS

Fantasy Football Models

- Aggregated and prepped 3 years of fantasy football projection data from 3 independent sources into a MySQL database.
- Created a random forest model in SAS, combining disparate sources into one projection that out-performed the mean absolute error of the next best projection by 15%.

Entertainment Engine

- Aggregated data from IMDB and Rotten Tomatoes, and used k-nearest-neighbors in SAS, constructing an enhanced entertainment selection targeted to reach 15 to 25 year old.
- Improved methodologies to save an average of 12 minutes per movie selection and 3 minutes per song selection.

COVID-19 Data Analysis

- Used AI-assisted document and image extraction techniques to automate the extraction of such data in structured (SQL) form from the state-level daily health bulletins.
- The target is to automate the data extraction and curation for each Indian state, so that once the extraction process of each state is complete, we can be on "autopilot" for that state, requiring little to none continued manual curation (other than to respond to changes in schema).