# NABEEL KHAN

### **EDUCATION**

Master of Science, Data Science, University at Buffalo, SUNY

Sep 2021-Jan 2023

Relevant course work: Numerical Maths for Data Science, Statistical Data Mining, Data Model & Query Language, Probability Theory, Predictive Analytics, Machine Learning,

GPA: 3.92

Reinforcement Learning.

Bachelor of Technology, IIT Roorkee, Industrial Engineering

May 2013-May 2017

### **SKILLS**

• Languages: Python, MATLAB, R, SAS

- **ML Tools :** Classification, Regression, Clustering, Tree Based Algorithms, Bagging & Boosting. Libraries: TensorFlow, Keras, PyTorch, and scikit-learn.
- o Data Management: SQLite, MySQL, pandas, Numpy, Excel
- Tools & Data analysis: Power BI, Google OR Tools, Time Series Forecasting, Version Control (git)

## **WORK EXPERIENCE**

# Associate Data Scientist, Merilytics, India

Apr 2020-Aug 2021

Developed a Demand forecast model with 1800 SKUs for a major European EV supplier

- Leveraged Time Series forecasting packages such as fb-prophet, neural prophet, & seq2seq neural nets.
- Reduced time taken for forecasting from 1 week to 4 hours with the automated pipleline.

Created a heuristic based driver **scheduling algorithm** to incorporate driving, terminal, & regulatory constraints.

- Automated driver scheduling, reducing the time taken from multiple days to 15 minutes.
- Produced a simulation of bills movement across 34 terminals for a long-haul trucking client.

Built an Automated Valuation Model (**AVM**) using TensorFlow-based custom nearest-neighbor architecture to identify comparable properties for 4 major property types.

- Delivered ~\$25 million annual savings by reducing ~2700 man hours weekly by automating property valuation.
- Collaborated with client on biweekly calls to **deploy the AVM** in the client's environment by establishing an end-to-end data pipeline, & refactoring the code using PEP guidelines.

## Senior Data Science Analyst, Merilytics, India

Feb 2019-Apr 2020

- Built Sales Forecast Model using Keras for creating promotion strategy for an American online clothing chain.
- Developed an end-to-end data pipeline on Azure utilizing Azure functions to automate ETL.
- Utilized KNN for determining comparable real estate properties, incorporating weights for different features.

## Research Intern, Vidooly, India

Dec 2018-Jan 2019

 Developed several Keras models for classifying YouTube Thumbnails and established data pipeline for extracting thumbnails using YouTube API.

## **PROJECTS**

• **Time Series Forecasting :** Time Series Analysis, customer segmentation, and interactive dashboard for 100k orders from an e-commerce platform.

Ongoing

 Resume Syncronization: Managed 6 profile versions on a single spreadsheet to keep changes in sync & avoid repetitive editing. Ongoing

- Utilized Python, HTML, & CSS for formatting & rendering, and Excel for handling data to create this version of my resume.
- 8 Ball Pool: Predicted & visualized ball trajectories using OpenCV, & made preemptive optimal decisions.
- Smartphone Price Prediction: Mined data from GSMArena and performed feature engineering by mapping Centurian Mark Score using fuzzy logic
- **Wildfires Analysis :** Visualized the clusters of different wildfires using geopandas, and predicted the Arson wildfire by performing EDA & appropriate feature engineering.
- Poultry Price Forecast: Scraped 2 years of data using Selenium and built a Time Series forecast model.
- · Clustered grocery items using kmeans for optimal positioning & proximity of similar items.
- Sentiment Analysis: Used Glove Word Embedding (NLP) for analyzing sentiments from twitter.

## **ACCOMPLISHMENTS**

**Achieved 99.76 Percentile** in the Joint Entrance Examination (All India exam with 1.4 million candidates). **Achieved 98.98 Percentile** in eLitmus pH Test for Problem Solving.