

NABEEL KHAN

📞 716-614-2691 ✉ nkhan6@buffalo.edu 🔗 linkedin.com/in/nabilkhaan 🐙 github.com/lanbeee 📍 Buffalo, NY, 14214

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, Machine Learning, Reinforcement Learning.

GPA: 3.92

Bachelor of Technology, IIT Roorkee, Indian Institute of Technology, Roorkee

May 2013-May 2017

SKILLS

- **Languages** : Python, JavaScript, R
- **ML Tools** : Classification, Regression, Clustering, Tree Based Algorithms, Bagging & Boosting. Libraries: TensorFlow, Keras, PyTorch, and scikit-learn.
- **Deep Learning & NLP** : LSTM, Transformers, CNN, YOLO, Word2Vec, TF-IDF, Topic Modelling, spaCy, NER
- **Data Management** : SQLite, MySQL, pandas, Numpy, Excel
- **Tools & Data analysis** : Version Control (git), Time Series Forecasting, Data Visualization, Machine Learning Interpretability

WORK EXPERIENCE

Associate Data Scientist, Merilytics, India

Apr 2020-Aug 2021

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.

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 pipeline.

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.

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.

Co-Founder & Content Creator, Synergy Learn, India

Jan 2018-Sep 2018

- Animated & Produced lectures for YouTube & garnered 10,000 hours of watch time.
- Managed company website, and social media to promote business.

Python Developer Internship, Modestreet, India

Sep 2017 Dec 2017

- Created a web app using Django & Three.js to make 3D human models with specific dimensions, to be used as a virtual mannequin for an online clothing store.

PROJECTS

- **Reinforcement Learning Model** : Designed environment using numpy to mimic AlphaGo for Ludo Ongoing
- **Time Series Forecasting** : Time Series Analysis, customer segmentation, and interactive dashboard for 100k orders from an e-commerce platform. Ongoing
- **Resume Synchronization** : Managed 6 profile versions on a single spreadsheet to keep changes in sync & avoid repetitive editing. Utilized Python, HTML, & CSS for formatting & rendering, and Excel for handling data to create this version of my resume. Ongoing
- **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