Urvashi Dube

dube.u@northeastern.edu | 236-863-3298 | github.com/ UrvashiDube02| LinkedIn

Technical Skills and Knowledge

- **Framework**s: Anaconda, Jupyter Notebook, Google Colab, PowerBI
- **Languages**: Python, SQL, Microsoft Excel
- **Python Libraries:** NumPy, Pandas, Matplotlib, Scikit-learn, Seaborn, Plotly
- **Technical skills:** Data Science, Machine Learning, AI
- **Algorithms:** Supervised Learning (Regression, Classification), Unsupervised Learning (Clustering), Reinforcement Learning

Professional Experience & Projects

Data Science Intern Jun. 2022 – Aug. 2022

GlobalCert, Singapore, SG

- Endeavored first position against forty other candidates to solve real world problems such as employee screening time reduction, GUI Based Drug recommendation system and e-commerce customer attrition analysis using **MySQL** to manage databases, **PowerBI (EDA)** for visualization, **Random Forest algorithm and Support Vector Machine** for analysis and prediction attaining 92% accuracy
- Attained letter of recommendation from CEO for exceptional work on automated employee selection system built in **Anaconda**

Data Analyst Intern Jan. 2022 – May. 2022

Bonrix Software Systems, Gujarat IN

- Devised a novel IOT-based camera using **OpenCV** for face expression analysis with mask at a specific location, **NumPy, Pandas, Sklearn** to clean, process and visualize data and **Random Forest** to forecast expressions with 91% accuracy
- Diagnosed preferences from customers and formulated products with innovative modifications by incorporating machine learning algorithms such as **Random Forest, Support Vector Machine and Naïve Bayes Algorithm**

Predictive Analyst Intern Aug. 2021 - Sept. 2021

Tevatron Technologies Pvt. Ltd., Uttar Pradesh IN

- Scoured unstructured big data of COVID-19 (age, M/F, vaccinations, deaths) with EDA, **pyplots**, **Excel**, and **PowerBI** for trend analysis
- Refined **Decision Tree** model accuracy incorporating data processing, significant attributes and performed hyperparameter tuning
- Researched about machine learning, deep learning models leveraged in industry and executed **Polynomial Regression** and **Support Vector Regression** in **Spyder** to solve problems pertaining to emotion recognition and face recognition with 88% accuracy

Data Insights Intern Jul. 2021 - Aug. 2021

NITK-STEP, Karnataka IN

- Systemized stock market's operation and stock market developments maneuvering real-time dataset from Bombay Stock Exchange leveraging **PowerBI dashboard** and **functions**
- Administered **Support Vector Machine**, and **Random Forest Algorithm** to predict stock rates post 2 years with 90% accuracy

Projects

Dry Bean Classification Using Machine Learning 2023

- Engineered a high-precision neural network achieving 90.79% accuracy in classifying 14 dry bean varieties, fine-tuning involved **OpenCV, NumPy, Pandas** in **Anaconda, Python**
- Demonstrated effective leadership by guiding a team of five in a rigorous model testing process, resulting in a comprehensive assessment and the identification of the most efficient classifier among three viable choices

Projected Price Prediction for Property in 5 years 2022

- Conducted analysis of Vancouver property data sourced from the City of Vancouver's open data portal leveraging **Python**, specifically **Pandas** and **Sklearn**, emphasizing location, area, and construction year.
- Devised a **Random Forest Regression** model that achieved an impressive 85% accuracy rate, enabling precise forecasts of future property prices in Vancouver.
- Uncovered valuable insights into price trends and predicted the next five years' growth in property prices for entire Vancouver area.

Publications

Cost Effective Railway Track Fault Detection 2020

- Devised a crack detection algorithm in **python** using **OpenCV** to improvise existing system of manual track health checking with 92% accuracy
- Exhibited model before a panel of 5 senior technical engineers at International Conference on **IoT** Based Control Networks & Intelligent Systems - **ICICNIS 2021** and published in SSRN, Elsevier Digital Library

Efficient Pipe Monitoring System and Hazard Detection 2020

- Instituted an efficient pipe monitoring system to monitor pipe health and detect hazards and cracks in a pipe using **ThingSpeak, Raspberry Pi** for **IoT** and Decision Tree Algorithm to predict lifespan of pipe with 95% accuracy
- Presented in front of a panel of 5 senior technical engineers at International Conference on **IoT** Based Control Networks & Intelligent Systems - **ICICNIS 2021** and published in SSRN, Elsevier Digital Library

Education

Master of Science in Data Analytics Engineering Dec. 2023
Northeastern University, Vancouver, BC

- CGPA of 3.84/4.00

Bachelor of Technology in Electronics and Communication Engineering May. 2022
Vellore Institute of Technology, Vellore, Tamil Nadu, IN

- CGPA of 3.58/4.00

Competitions

Responsible AI Symposium at Northeastern University, Vancouver 2023

- Awarded the prestigious recognition at the Responsible Artificial Intelligence Symposium 2023 for pioneering research and application of Responsible AI in healthcare, triumphing over 40 competitors.

Zeal Hackathon 2022

- Acquired the first position by spearheading the creation of an application enabling real-time display of stock availability in nearby stores during COVID 19 pandemic leveraging **Python** and **Power BI** visualizations