KESHIKA ARUNKUMAR

arunkumar.k@northeastern.edu | (857) 423 5586 | LinkedIn | GitHub | Tableau | Portfolio

EDUCATION

Northeastern University **Boston, MA**
Master of Science in Data Analytics Engineering (GPA – 3.84) Expected May 2025
Key courses: Data Management for Analytics, Foundations of Data Analytics Engineering, Data Mining, Computation and Visualization

Anna University **Chennai, IN**
Bachelor of Engineering in Computer Science Engineering Jun 2023
Key courses: Probability and Random Processes, Database Management Systems, Python Programming, Machine Learning

TECHNICAL SKILLS

Programming Languages: Python (Pandas, NumPy, Matplotlib, Scikit-learn, Seaborn, TensorFlow, PyTorch), SQL, R, Java, C/C++
Analytics Tools: Tableau, Power BI, MS Excel (VLookups, Pivot Tables), Jupyter, Google Colab, Google Sheets
Databases: MySQL, Oracle, SQL Server, MongoDB, AWS (S3, EC2, Lambda), ETL(Snowpipe), Snowflake
Machine Learning: Regression, Classification, Decision Trees, SVM, Hypothesis Testing, Clustering, Neural Networks
Other Tools and OS: Git, MS PowerPoint, MS Word, MS Teams, MS Outlook, Microsoft Office Suite, Windows, Linux

WORK EXPERIENCE

AssetPlus Consultancies Pvt. Ltd. **Chennai, IN**
Data Analyst Mar 2023 – Jun 2023
- Developed a **CNN-based model** for digit recognition from IGL gas meters.
- Managed data collection, **preprocessing**, and **deployment** of the model.
- Achieved 90% accuracy with the digit recognition model.
- Conducted **predictive analysis** using **ARIMA, Random Forest, XGBoost** to reduce gas loss from 3.23% to 0.97%.
- Created an interactive **Tableau** dashboard for **visualizing** gas loss data, facilitating detailed analysis, real-time monitoring, and strategic decision-making.

Suven Consultants and Technology Pvt.Ltd. **Chennai, IN**
Data Analyst Coding Intern Oct 2022 - Dec 2022
- Contributed to advanced exploratory data analysis processes, using **Excel** for documentation, and improving team efficiency.
- Implemented **support vector classification** for recognizing handwritten digits, achieving a remarkable accuracy score of 99%.
- Prepared a PowerPoint deck and video presentation to **communicate** key insights for the client and internal stakeholders..

ACADEMIC PROJECTS

Churn Reduction Framework (Python, Scikit-learn, Pandas) **Boston, MA**
Northeastern University Feb 2024 - Mar 2024
- Applied **machine learning** models (Logistic, Decision Tree, Random Forest) to predict customer churn with 84.2% precision.
- Employed **data preprocessing** techniques (one-hot encoding, log transformation, SMOTE) to achieve 89.4% predictive accuracy.
- Leveraged **scikit-learn** and **pandas** for machine learning tasks, yielding precise customer churn predictions.
- Demonstrated proficiency in optimizing model parameters for improved predictive accuracy using feature engineering for large datasets.

An In-Depth Analysis of Employment Trends in New York City (Application - Link) **Boston, MA**
Northeastern University Oct 2023 - Nov 2023
- Developed a **Streamlit** application to visualize employment trend in NYC, including analysis of salary ranges and work locations.
- Employed **Exploratory Data Analysis** to derive custom metrics such as career level distribution, trends by month and popular work units.
- Implemented slide bars to allow users to adjust **visualization** parameters, enhancing dashboard interactivity dynamically.
- Integrated diverse datasets to provide insights into employment and unemployment trends across different demographics.

Employee Performance and Rewarding System (MySQL, Python, MongoDB) **Boston, MA**
Northeastern University Sep 2023 - Dec 2023
- Designed an Employee Performance and Rewarding System using **MySQL** and **Python**, adhering meticulously to relational database principles.
- Executed **stored procedures, complex joins, CTE, functions,** and **SQL views**, enhancing efficiency and query execution time.
- Employed techniques like ER modeling, UML diagrams, normalization, and Hierarchical Data Modeling to ensure system scalability.

Asthma Prediction Using Convolution Neural Networks **Chennai, IN**
Anna University Mar 2023 - May 2023
- Executed data preprocessing techniques on lung X-ray images to ensure **data integrity** and **accuracy**.
- Designed and implemented a classification module using **CNN**, engineering a **ConvNet** architecture and attained an accuracy of 84.7%
- Developed an interface using **Django** for X-ray image upload and asthma classification and prediction, emphasizing user engagement.