

SHARAYU THOSAR

+1 602-849-3746 | thosar.sh@northeastern.edu | [LinkedIn](#) | [GitHub](#) | [Portfolio](#) | Boston, USA

EDUCATION

Northeastern University

Master of Science in Data Analytics Engineering | GPA: 3.85

Boston, USA

Expected May 2024

Relevant Coursework: Data Mining, Computation and Visualization, Data Management

COEP Technological University

Post Graduate Diploma in Data Science and AI | GPA: 3.94

Pune, India

Aug 2022

Vishwakarma Institute of Technology

Bachelor of Technology in Mechanical Engineering | GPA: 3.71

Pune, India

Jun 2021

EXPERIENCE

Northeastern University

Designation: Teaching Assistant (IE:6600 Computation and Visualization for Analytics)

Boston, USA

Feb 2023-Present

- Grade assignments and exams, conduct office hours, solve students' doubts and provide classroom assistance

Designation: Research Aide

Jan 2023-Present

- Investigate academic research paper trends in collaboration with Nano Government officials and the University
- Develop python codes for data cleaning and natural language processing to construct network graphs
- Identify, document, and present insights on emerging and declining research trends in the nanotechnology field

Intelimek Systems

Designation: Business Intelligence Developer

Pune, India

Jan 2022-Jul 2022

- Designed 26 client dashboards using sensor data to enable efficient monitoring and reporting of environmental data, resulting in significant time savings compared to traditional manual methods in Excel
- Guided dashboard users by conducting training sessions on efficient use and management of dashboards

Designation: Data Science Intern

Feb 2021-Aug 2021

- Developed a 98.67% accurate SVM Model to predict types of motor faults, now utilized in company applications
- Implemented Signal Processing, Fast Fourier transforms, and Feature Extraction in python to build ML models
- Interpreted Time Series graphs of IoT data to analyze and gain actionable insights through multiple case studies

TEIM

Designation: Data Analyst Intern

Pune, India

Feb 2021-Aug 2021

- Developed an automated ETL Pipeline to clean, merge and organize the unstructured client data, resulting in an efficient workflow to store data in the SQL server and monitor the KPIs in the Power BI dashboard
- Proposed, initiated, and succeeded in implementing an NLP task to significantly improve pipeline performance
- Curated Project Documentation and conducted code review to aid in smooth knowledge transfer of the work

PROFESSIONAL SKILLS

- **Technical Skills:** Data Management, Machine Learning, Time Series Analysis, Data Exploration, Big Data Analysis
- **Programming Languages:** Python (Pandas, NumPy, Seaborn, Pytorch, SciPy), R (R shiny, tidyverse), SQL, C++
- **BI Tools:** Tableau, Power BI, IBM Cognos Analytics, Data wrapper, Flourish
- **Data Management:** MySQL, Neo4j, Google Big Query, Apache Spark, Google Cloud Platform
- **Others:** GitBash, Git Version Control, Microsoft Suite, AWS, Visual Studio Code

PROJECTS

Database Modeling for Kids Crew App: Creating Conceptual Model (ER and UML), Relational Model, Database Generation in MySQL, Implementation in MySQL and NoSQL (Neo4j), and Applications in Python.

Data Driven Story Telling using US Census Data: Curated a website that features graphs and analysis using US Census data and serves as a valuable resource for exploring significant population trends and patterns in the US.

Human Activity Recognition: Extracting time series features for human activity classification along with visualizing the plots of different body parts for 15 individuals while running, climbing and walking.

Turbine Failure Data Analysis: Data Visualization and Time Series Forecasting in Python to delineate the reasons and predict future breakdowns.

Clustering World Countries: Compare different countries all over the world by implementing K-means and Hierarchical Clustering methods based on GDP, Population, Energy Usage, Mortality Rates, and more.