SAMIKSHA BARASKAR

Boston, MA | +1(857)-390-5787 | baraskar.s@northeastern.edu | https://www.linkedin.com/in/samikshabaraskar/ | GitHub |

EDUCATION

Master of Science in Data Analytics Engineering, GPA: 4.0/4.0 Expected May 2025
Northeastern University, Boston, USA
Relevant Coursework: Data Mining, Machine Learning, Data Analytics, Data Management for Analytics

Bachelor of Technology in Computer Science and Engineering, GPA: 10/10 July 2021
Dr. Vishwanath Karad's MIT World Peace University, Pune, India
Relevant Coursework: Data Science, Big Data Analytics, Business Intelligence, , **[Innovation Patent]**:Virtual Vision

WORK EXPERIENCE

Junior Software Engineer – Data Platform Aug 2021- Dec 2022
TietoEvry Private Ltd, Pune, India

- Drove **20% increase in deployment efficiency** of engineering applications **Dochotel** and **Share@Poyry**, by **optimizing database architecture** using **Microsoft SQL** and **strategy driven database optimization**.
- Enhanced data retrieval accuracy by **automating data solutions** and **optimizing database schema**, resulting in a **15% reduction in data retrieval time**.
- **Collaborated** with **cross-functional** teams to **streamline customer interactions and communication**, employing strong analytical support and problem solving skills to improve quality assurance and decision making.

Machine Learning – Artificial Intelligence (ML-AI) Intern May 2020 - Jul 2020
Erai Technologies Private Ltd, Pune, India

- Led the development of **Smart Care AI**, harnessing **Python**, **TensorFlow**, and **NLP** to extract real-time skills from resumes. Achieved an outstanding **40% enhancement in skill matching accuracy**, contributing to talent acquisition.
- Engineered and configured the seamless integration of **Apache Kafka** and **Elasticsearch** for real-time data processing. Streamlined **kafka pipelines** utilizing analytical **ETL and ELT** processes to optimize **data collection and distribution**, resulting in a substantial **25% increase in application responsiveness**.
- Applied **text classification techniques** and machine learning to refine candidate skill matching, optimizing data extraction and streamlining the enterprise hiring process employing domain knowledge.

PROJECTS

Spotify Data Analysis [Python, Data Analysis, Machine Learning, Data Visualization Tools]

- Conducted **comprehensive statistical analysis** on a **large-scale** Spotify Music dataset to enhance the user and artist experience, deriving data-driven music recommendations.
- Developed and evaluated **machine learning predictive models**, including multiple linear regression, decision trees, random forests, and neural networks. Achieved an **average accuracy of 85%**, improving model performance.
- Employed **experimental design** techniques for **data collection, preprocessing, cleaning, feature engineering**, and **effective data visualization** to interpret detail oriented meaningful results.
- Implemented a **personalized recommendation system** model, resulting in a **20% increase** in user interaction through collaborative filtering and content-based recommendations.

Foreign Direct Investment in India Analysis [Tableau, Data Analysis, Data Visualization]

- Developed an **interactive Tableau dashboard** for analyzing FDI trends and statistics in India, improving data accessibility and comprehension **by 25%**.
- Implemented **dynamic filters and parameters**, resulting in a **30% improvement in user customization** capabilities. Utilized **advanced Tableau features** to enhance the dashboard's visual appeal and **user experience**.

On-Campus Housing Database Management System [SQL, MYSQL, NOSQL, Python]

- Designed and built housing management system, **cutting assignment process time by 30%** and improving student experience for residents at Northeastern University through **advanced SQL queries** and **analytical scripting**.
- Conceptualized **data mode**l and developed **EER and UML** diagrams to architect **optimal MySQL relational database** design capable of handling over 1 million unstructured and semi-structured housing records.
- Prototyped **NoSQL graph database** in Neo4j to explore feasibility and scalability for housing data analytics.

TECHNICAL SKILLS

Programming Language: Python, R, SQL, C++
Database Systems: Oracle, MS SQL, MySQL, MongoDB, NOSQL
Data Manipulation and Analysis: Pandas, NumPy, Scikit-learn, TensorFlow, Keras, PyTorch, SciPy
Data Science: Model Development and Evaluation, Statistical and Quantitative Analysis, Time Series Analysis, Natural Language Processing, Machine Learning Algorithms, Deep Learning, Recommender Systems, Predictive Modelling
Tools and Software: Git, Excel, Tableau, PowerBI, Amazon Web Services(AWS), PowerPoint, Microsoft office