Sejal Barshikar

Boston, MA | +1 857-397-6984 | barshikar.s@northeastern.edu | LinkedIn | GitHub

EDUCATION

Northeastern University, Boston, MA **Aug 2025 - Dec 2027**
Master of Science, Computer Science Concentration in Artificial Intelligence and Machine Learning GPA: 3.83/4

- **Relevant Courses** - Algorithms, Database Management Systems, Programming Design Paradigms, Deep Learning

Savitribai Phule Pune University, Pune, India **Aug 2021 - May 2025**
B.E in Artificial Intelligence and Data Science GPA: 8.65/10

- **Relevant Courses** - Machine Learning, Pattern Recognition, Computer Vision, Data Structures, Discrete Mathematics, Database Management Systems, Statistics, Probability, Linear Algebra

SKILLS

Programming Languages: Python, C++, JavaScript, SQL
Machine Learning: PyTorch, TensorFlow, Keras, Supervised learning (Regression, Classification), Sequence models (LSTM, GRU, Seq2Seq), Attention mechanisms, Transformers, Model optimization, Feature engineering, Cross-validation
Libraries PyTorch, TensorFlow, Keras, Hugging Face, NLTK, OpenCV, Pandas, Scikit-learn, NumPy, SciPy
Tools: Jupyter, Git, VS Code, MongoDB, REST API

PROFESSIONAL EXPERIENCE

AICTE (All India Council for Technical Education), Remote, India **Feb 2024 – Apr 2024**
Data Science Intern

- Built an end-to-end ML pipeline processing **5M+** transaction records to transform raw data into model-ready features and customer representations
- Engineered behavioral features and trained clustering-based models, evaluating run-to-run stability across **10+** random initializations to ensure consistent model behavior and reduced data preprocessing time by **∼60%** by optimizing Pandas-based ETL workflows
- Deployed a recommendation service via **Flask REST API**, optimizing inference paths to achieve **sub-100ms** latency for real-time usage

PROJECTS

Video Content Analyzer **Dec 2025 – Present**
Developing an end-to-end video analysis pipeline to extract semantic information

- Building an end-to-end video analysis pipeline combining computer vision and NLP models to extract structured metadata from unstructured video content
- Implemented frame sampling and modular inference stages to decouple video decoding from model execution, improving pipeline maintainability
- Evaluating system performance using throughput, latency, and qualitative output consistency rather than raw accuracy metrics

Research Paper Classifier (Fine tuning BERT) | Link **Oct 2025 – Dec 2025**
Classifying research paper based on its title and abstract by Fine Tuning BERT

- Designed and deployed an end-to-end document classification system by fine-tuning BERT-base on **110M** parameters for multi-class classification across **11** academic categories, with a modular, production-oriented codebase and validation pipeline
- Optimized training for a resource-constrained environment by implementing gradient clipping, learning-rate warmup and early stopping, reducing training time to **∼5** minutes per epoch while maintaining **stable convergence**
- Conducted detailed **error analysis** using confusion matrices, identifying conceptual **overlap** between closely related categories and documenting insights to guide future dataset refinement and model improvements

PUBLICATIONS & RESEARCH

Advanced Retrieval-Based Code Summarization using Meta Learning | Link **Apr 2025**
Mukt Shabd Journal

- Developed meta-learning framework using **MAML algorithm** for automated code summarization across **Python** to improve adaptability with minimal fine-tuning