SONALIKA SINGH
8434756189 | Email | LinkedIn | GitHub | Portfolio

Professional Summary

Data Science professional with 1+ year experience at ERNET India and IIT Guwahati, specializing in deep learning, NLP, and MLOps pipeline development. Proven track record of building end-to-end ML solutions achieving 96.8-99.8% accuracy across language classification, and facial recognition projects with expertise in Python, PyTorch, and cloud deployment.

Education

Master of Mathematics, Indian Institute of Technology Madras	7.43 CGPA (2023-2025)
Bachelor of Science (Mathematics), Patliputra University, Patna	62.2% (2018–2021)
Senior Secondary, Bihar School Examination Board	68% (2016–2018)
Secondary, Central Board of Secondary Education	9.4 CGPA (2015–2016)

Experience

Data Analyst Intern, ERNET India, IIT Madras Research Park — July 2024 – June 2025

- Analyzed **FSOC** Communication over a **1.12** km link in Kohima using Python, processing **180+** days of weather and signal data to detect **outage** causes, quantify downtime and improve network resilience.
- Developed **10+** Python visualization tools, **KPI** dashboards and automated reporting scripts to monitor **FSOC** link performance, categorize outage, and enhance **internet** services in remote hilly regions.
- Built groundwater forecasting models (**MLR, RFR, ANN, GRU, LSTM**) on **2.5K+** time-series points, achieving **98%** accuracy and showcasing expertise in neural networks and time series analysis for sustainable resource planning.

Winter Intern – Applications of ML Techniques, E&ICT Academy, IIT Guwahati — Dec 2024 – Jan 2025

- Built and tuned **5+** deep learning models (**FFNNs, CNNs**) on MNIST, CIFAR-10 using TensorFlow/Keras, improving classification accuracy by **15%** through hyperparameters optimization and regularization.
- Fine-tuned VGG and ResNet models, evaluated performance using **6+** metrics (**Accuracy, Precision, Recall, F1-Score, ROC- AUC, Confusion matrix**) and created Python dashboards for result visualization.

Projects

Multilingual Language Classifier: Developed an end-to-end **MLOps** pipeline for classifying **27** Roman-script languages and continent groups on **10,000+** samples using character-level TF-IDF, TruncatedSVD, and Naive Bayes/LDA/QDA models, achieving **96.8%** language and **89.4%** continent accuracy; deployed via **Streamlit** App. [verify]

Music Recommender System: Built and deploy a content-based music recommender by preprocessing **1M+** song lyrics using **NLTK**, leveraging TF-IDF (**5,000** features) and cosine similarity, and delivers real-time **top-5** song recommendations via an interactive **Streamlit** app with Spotify API integration.[verify]

Sequence Modeling for Language Decoding: Built and optimized Seq2Seq transliteration models (**RNN, LSTM, GRU**) for Roman-to-Devanagari on **10,000+** samples, boosting validation accuracy by **8%** through attention mechanisms, beam search decoding and Weights & Biases Bayesian hyperparameter sweeps. [verify]

Facial Recognition Attendance System: Developed a facial recognition web app using **OpenCV**, dlib, and CNNs to automate attendance for **1,000+** users with real-time video processing and reporting; achieved **99.8%** recognition accuracy and deployed using Streamlit and **Docker** containers. [verify]

Skills

Technologies & Tools: Python, SQL, Git, Docker, MLflow, NumPy, Pandas, PyTorch, Streamlit, Power BI
Domains: Machine Learning, Deep Learning, NLP, Transformers, MLOps, Computer Vision.

Relevant Coursework

Core Courses: Foundation of Machine Learning, Introduction to Deep Learning, Data Structures and Algorithm, Data Analytics Laboratory,Linear Algebra, Probability,Applied Statistics, Numerical Optimization, Multivariable Calculus
Online Courses: OCI **Data Science** Professional (Oracle), **Data Science** Job Simulation (Lloyds Banking Group),**Data Analytics** (Deloitte), Intro to **AI** (IITM Summer School), Master in **Data Science** (Udemy), Communication Skills (TCS iON), Advance Diploma in Computer Applications (MICT Wizard-Tech), Power BI Workshop

Leadership, Activities, and Achievements

Scholastic: Secured AIR **417** in IIT-JAM (Mathematics); Qualified **JEE** (2018, 2019) and **GATE** (Data Science, 2025).
Leadership: Volunteer at **Forays 2024** (Math Fest, IIT Madras) managing event operations and participant coordination.
Sports: Represented IIT Madras hostel teams in **Cricket & Football**, achieved 4th position in cricket tournaments and demonstrated teamwork, adaptability, and continuous improvement.