Anshika Bajpai

United States | anshikabajpai23@gmail.com | LinkedIn | Portfolio | +1 (930) 204 3030

EDUCATION

Indiana University Bloomington, Masters of Science in Data Science, USA **May 2026**

Coursework: *Machine Learning, Algorithms, Statistics, Advance Database Concepts, Exploratory Data Analysis*

Bachelor of Technology in Computer Science and Engineering, Jaypee Institute of Information Technology, India **May 2021**

SKILLS & CERTIFICATION

Programming Language: Python, R, SQL, Java, C++, C

Frameworks/Tools: Talend, Jenkins, Git, AWS, DevOps, Docker, Kubernetes, PySpark, Data Visualisation, ETL

AI/ML: TensorFlow, PyTorch, Keras, NLP, Computer Vision, Keras, Neural Networks (including CNN, RNN, LSTM, GRU, FeedForward Networks), Generative AI, Recommendation System, MLOps, Computer Vision, Anomaly Detection, LLM

Libraries/Software: Numpy, Matplotlib, Pandas, NLTK, SciKit-Learn, OpenCV, Grafana

Machine Learning Certification, Stanford University

WORK EXPERIENCE

Senior Software Engineer, Optum (UnitedHealth Group) **Jun 2021 – Jun 2024**

Coverage Feed

- Reduced time from ***500+ minutes to less than 3 minutes*** *by* refactoring existing legacy systems with high code quality standards.
- Optimized error report generation system, reducing errors from ***25,000 to 20***, improving latency and minimizing production costs.

Insurance Letter Generation

- Led the design and development of critical software components for health insurance systems (UHOne).
- Took initiatives to improve code quality & coding standards by introducing design patterns, writing unit tests with *at least 90% code coverage*, creating code review guidelines, and mentoring juniors across teams.
- Built a system to send automated Compliance Notices as part of ***"The Paperless Initiative"***, saving the company ***$500,000 annually***.
- Created a real-time Grafana dashboard for the product to analyze metrics, improving team efficiency by saving ***$15,000 annually***.

Idea-thon

- Led a successful hackathon project focused on cardiovascular disease prediction, achieving *84.21% accuracy* with ***Logistic Regression*** and *82.89% with* ***Random Forest models***.

Machine Learning Scientist Intern, Taiyo LLC **Dec 2019 – Apr 2020**

- Built a **machine learning pipeline** for stock market prediction, improving **accuracy by 15% over baseline models.**
- Engineered a **hyperparameter tuning system** for 10+ models, boosting predictive performance and user satisfaction.
- Developed time series forecasting models in **Python and scikit learn**, analyzing 50,000 stock price records to uncover trends.

PUBLICATIONS

- **Bajpai A.**, Garg M., Hindi Sentiment Analysis on Tweets, 2nd International Conference on Artificial Intelligence: Theory and Applications (AITA 2024). [SCOPUS][DBLP] - **Accepted & Presented**
- Garg M., **Bajpai A.**, Kumari S., Rathi S., Dahiya T., Ghosh A., AI-Driven Radiomics for Early Detection of Gynaecological Cancers: A Multimodal Approach, International Conference on Smart Cyber Physical Systems (ICSCPS-2024). [SCOPUS] - **Accepted & Presented**

PROJECTS

Rock Type Classification | *Machine learning | Ensemble Learning | Python* **Sept 2024 – Oct 2024**

- Implemented and optimized **Softmax Regression, SVM (82.2%), and Random Forest (76.7%)** models for classifying rock types, leveraging features like texture and stripes.
- Developed and compared ensemble models **(Hard Voting: 81.1%, Soft Voting: 77.8%, Stacked: 78.9%)** against human classification accuracy to evaluate reliability and performance.

Hindi Sentiment Analysis on Tweets | *Naive Bayes | Lexicon-Based Model | NLP, Python* **Oct 2023 – Jan 2024**

- Increased accuracy of Hindi sentiment analysis from **41% to 65%** by enhancing the **lexicon-based method**, addressing a key research gap, and improving sentiment analysis for Hindi tweets.

PLANTIFY - Mobile Application | *Image Segmentation | CNN, Python* **Jan 2020 – May 2020**

- Achieved 70% accuracy in plant recommendation by leveraging a **U-Net model for image segmentation**, effectively identifying sky-ground boundaries, and providing tailored plant recommendations based on location and solar capabilities.

LEADERSHIP & COMMUNITY INVOLVEMENT

Attendee, SWE (Society of Women Engineers) Conference (2024) **Google Developer Student Club Member, Indiana University (2024)**