# ANSHIKA BAJPAI

India anshikabajpai23@gmail.com LinkedIn GitHub

## EDUCATION

**Jaypee Institute of Information Technology (JIIT)** **May 2021**
*B.Tech. in Computer Science and Engineering (Honors)* GPA: 8.0

## RESEARCH INTERESTS

Machine Learning, Data Science, Data Analytics, Artificial Intelligence, RNN, Generative AI, NLP, CNN

## INDUSTRY EXPERIENCE

**Optum, UnitedHealth Group** | *Java | ML | Python | Spring boot | Maven | ReactJS| Talend | SQL* **Jun 2021 – Present**
*Senior Software Engineer* *Bangalore, India*

- Led the design and development of critical software components for health insurance systems (UHOne).
- Developed end-to-end architectural designs for various features of the product.
- Led a successful hackathon project focused on cardiovascular disease prediction, achieving ***84.21% accuracy*** with ***Logistic Regression*** and ***82.89% with Random Forest models***.
- Refactored the existing legacy system (VUE coverage feed) with high code quality standards thereby, reducing code execution time from ***500+ minutes to less than 3 minutes***.
- Enhanced the report generation system by optimizing the file sizes from ***7426 Kb to 14 Kb***, hence improving the latency and minimizing the production cost.
- Built a system to send automated Compliance Notices through emails as part of "The Paperless Initiative", which saved the company ***$500,000 annually***.
- Created a real-time Grafana dashboard for the product to analyze metrics such as generated errors and letters, etc thereby, improving team efficiency by saving ***$15,000 annually.***
- 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.

**Taiyo LLC, San Francisco, USA** **Dec 2019 – Apr 2020**
*Machine Learning Scientist Intern* *Remote*

- Developed an algorithm for a stock market prediction interface using Machine Learning and Deep learning.
- Implemented Hyperparameter tuning service using Bayesian Optimization algorithm to fetch best parameters for the model.
- Collaborated with team members in researching and evaluating cutting-edge machine learning techniques and technologies for potential adoption in future projects.
- Used Python and Scikit-learn to build predictive models for time series forecasting of stock financial data.

## TECHNICAL SKILLS

Python, Machine Learning, Deep Learning, C++, Java, Data Structures & Algorithms, SQL, ReactJS, Azure, Design Patterns, Git, Jenkins

## PROJECTS

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

- Developed a sentiment analysis model for Hindi tweets, addressing the research gap in the Hindi language on Twitter(X) tweets.
- Developed and compared approaches: Lexicon-based (Hindi-SentiWordNet) and Machine Learning-based (Naïve Bayes Classifier).
- Identified an enhanced lexicon-based method, utilizing positive and negative word counts, as the most accurate for sentiment analysis in the Hindi language, accuracy improved from 41% to 65%.

**Hindi Question Answering System** | *Levenshtein model | NLP, Python* **Mar 2023 – Jul 2023**

- Developed a Hindi question-answering system using the Levenshtein model, achieving an accuracy ***of 93.4***%.
- Thoroughly analyzed model performance, achieving 100% accuracy for exact matches, 90% for dissimilar queries, and addressing challenges in character-to-character matching for random queries.

**Handwritten Digit Recognition** | *Deep learning | Feedforward networks | CNN | RNN, Python* **Apr 2021 – Jun 2021**

- Developed a convolutional neural network to recognize handwritten digits using Python and TensorFlow with an ***accuracy of 95.1%***.
- Trained the model using stochastic gradient descent (SGD) with backpropagation.

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

- Generated plant recommendations by analyzing a user's location and solar capabilities.
- Utilized a U-Net model for image segmentation to identify the boundary between the sky and ground and then encoded the results in a run-length format with an ***accuracy of 70%***.

**Mental Health prediction of Instagram users** | *Machine Learning and RNN | Python* **Aug 2019 - Nov 2019**

- Leveraged deep learning and machine learning methods to forecast the mental health status of Instagram users.
- Performed multi-class and multi-label classification on Instagram posts.

## PUBLICATIONS

**Bajpai A.,** Garg M., and Mittal H., **Hindi Sentiment Analysis on Tweets**, International Conference on Business Intelligence and Data Analytics (BIDA 2024)**. [SCOPUS] - Under Review**

## VOLUNTEERING EXPERIENCE

**National Service Scheme - Education and Cultural Team Head** **Aug 2018 – Jul 2020**

- Organized and managed blood donation camps, marathons, and plantation drives.
- Led educational initiatives, designed curricula, and participated in community projects, including food distribution and cleanliness drives.

**Saksham: Child Welfare & Educational Trust** **Dec 2019 - Feb 2020**

- Actively participated in initiatives aimed at promoting child welfare and education, contributing to the organization's overarching goals.

**Vegan Outreach - Volunteer** **Feb 2019 - Mar 2019**

- Engaged in vegan outreach, promoting plant-based living and animal welfare. Collaborated on events and campaigns to raise awareness.