

Sandeep Kumar

Delhi, India | +91 60053 62383

22becse37.cse@cujammu.ac.in | github.com/Sandeep37-s | linkedin.com/in/sandeep-kumar

Profile

B.Tech Computer Science and Engineering student with hands-on experience in Machine Learning, Data Science, and Artificial Intelligence. Skilled in developing end-to-end ML applications, from data collection and preprocessing to model deployment. Experienced through academic projects, internships, and certifications in areas like NLP, computer vision, and data analysis. Currently researching Intrusion Detection Systems, with a strong interest in applying AI to solve real-world security and analytical challenges.

Education

Central University of Jammu	2022 – 2026
Bachelor of Technology (B.Tech) in Computer Science and Engineering	
<ul style="list-style-type: none">CGPA: 7.2/10Coursework: Data Structures Algorithms, Artificial Intelligence, Machine Learning, Deep Learning, LLM, NLP, Database Systems	
St Johns Academy, Anjanpeer, Hajipur , Senior Secondary (12th), CBSE Board	2020 – 2022
Mount Litera Zee School, Sonpur , Secondary (10th), CBSE Board	2020

Experience

Advanced Data Analysis , Certification Program	Jan 2025 – Feb 2025
Completed training in advanced data analysis techniques, covering preprocessing, feature engineering, and statistical modeling. Certificate ID: C261545 , Reg. No: IN12510521	
Machine Learning with Python , Training Program	May 2025 – June 2025
Gained hands-on experience in supervised/unsupervised learning, model evaluation, and ML algorithms using Python. Certificate ID: VTP-241963	
Data Science Intern , Pinnacle Labs – Remote	June 2025 – July 2025
Completed a 4-week internship focusing on data preprocessing, model building, and real-world ML applications. Certificate ID: PL/2025/JUNP1/029 (Verify)	

Certificates

Machine Learning Basics - Regression Analysis , Udemy	
<ul style="list-style-type: none">Completed 5.5 hours of training on regression analysis and machine learning fundamentals.Gained practical knowledge in predictive modeling using real-world datasets.	
Advanced Data Analysis , Certification ID: C261545	
<ul style="list-style-type: none">Acquired advanced skills in data preprocessing, feature engineering, and statistical modeling.Strengthened expertise in applying data analysis to business and research problems.	
Build Your Generative AI Productivity Skills , Microsoft & LinkedIn Learning	
<ul style="list-style-type: none">Learned to apply Generative AI tools for business productivity and design applications.Strengthened AI literacy for leveraging modern workplace AI solutions.	

Research

Intrusion Detection Systems: A Systematic Review	2025 – Ongoing
<ul style="list-style-type: none">Conducting a systematic review of various algorithms and approaches for Intrusion Detection Systems (IDS).Analyzing performance metrics such as accuracy, precision, recall, and F1-score across multiple benchmark	

datasets.

- Research paper in progress as part of B.Tech CSE project.

Projects

Mini Movie Search Engine - <https://github.com/Sandeep37-s/movie-searchEngine-word2vec>

- Scraped movie data from online sources to build a custom dataset.
- Developed a semantic search engine using a self-trained **Word2Vec** model combined with pre-trained **Google News vectors (GoogleNews-vectors-negative300.bin)**.
- Enabled movie search and recommendation functionality based on semantic similarity.
- Tools Used: Python, Gensim, scikit-learn, BeautifulSoup

Sentiment Analysis of Prompts - https://github.com/Sandeep37-s/sentiment_analysis

- Collected movie reviews by scraping websites and downloading datasets from Kaggle.
- Performed preprocessing (cleaning, tokenization, stopwords removal) to prepare the dataset.
- Trained **Logistic Regression**, **SVM**, and **Multinomial Naive Bayes** models, achieving **79% accuracy**.
- Tools Used: Python, scikit-learn, NLTK

CodingThunder Blog Website - https://github.com/Sandeep37-s/FlaskBlogApp-coding_thunder

- Built a blogging platform named **CodingThunder** with post creation and management features.
- Backend implemented in **Flask**, frontend in **HTML/CSS/JavaScript**.
- Used **XAMPP (MySQL)** with **SQLAlchemy ORM** for database management.
- Tools Used: Flask, HTML, CSS, MySQL (managed with SQLAlchemy ORM)

Potato Disease Detection - <https://github.com/Sandeep37-s/Potato-disease-detection>

- Developed a CNN model for potato leaf disease classification using the PlantVillage dataset.
- Integrated the model with a **Flask web app** for real-time disease detection and user interaction.
- Tools Used: TensorFlow, Flask, NumPy, Pandas

Skills

Languages: Python, C, C++, Java, SQL

Technologies/Tools: Linux, Git/GitHub, MySQL, SQLite, XAMPP, Computer Vision, Docker

Libraries/Frameworks: TensorFlow, PyTorch, Scikit-learn, NLTK, OpenCV, Flask, FastAPI, Langchain

Achievements

- Captain, B.Tech Kabaddi Team
- Secured **1st Place** in Running (University Level)
- Winner in **Sketching & Story Writing** (University Level)
- Actively participated in university cultural and sports events