Kumar Anmol

Noida • k.anmol1437@gmail.com • +91-9582544698 • LinkedIn • GitHub • CodingNinjas

EDUCATION

- **Bennett University (Times of India Group)** — **Greater Noida, Uttar Pradesh**
  Bachelor of Technology in Computer Science | CGPA: 8.58 — September, 2021 - May, 2025
  Specialization in Artificial Intelligence and Machine Learning

SKILLS & INTERESTS

- **Languages** C++ | Python | Java | SQL
- **Frameworks** Streamlit | Flask | OpenCV | BeautifulSoap
- **Utilities** Git | Render | AWS | VS Code | Google Colab | Microsoft Azure
- **Technical Skills** Data Structures & Algorithms | Object Oriented Programming | DBMS | Cloud Infrastructure
- **Data Science & ML** Data Extraction | EDA | Feature Selection | Pandas | Numpy | Matplotlib | NLP | LangChain | LlamaIndex | Prompt Engineering

PROJECTS

Cross-Media Recommender System | *Python, NLTK, Scikit-Learn, Streamlit* GitHub

- Created an end-to-end user-friendly platform using Streamlit for personalized recommendations of 5 similar movies, or books, leveraging **content-based** and **collaborative filtering** algorithms.
- Employed the **TMDB 5000** movie dataset and a comprehensive book recommendation dataset featuring metadata for over **250,000 books** in multiple languages.

Emotion Detection System | *Python, TensorFlow, Keras, ResNet50v2, VGG16, OpenCV* GitHub

- Designed and iterated on **custom CNN models**, including advanced architectures such as VGG16 and ResNet50v2.
- **Addressed class imbalance** in the FER-2013 dataset by employing image augmentation and class weights.
- **Achieved a 66% overall accuracy** on emotion classification, with the final model based on ResNet50v2, across 7 emotion labels.
- **Deployed the model** for real-time emotion detection in live video streams using Gradio and OpenCV.

Credit Risk Assessment System | *Python, Scikit-Learn, XGBoost, Flask* GitHub

- Developed a model to accurately assess the risk factors associated with providing loans to customers.
- Conducted data preprocessing and feature selection on over **80 features**, addressed **multicollinearity** using **sequential VIF**, and effectively reduced the feature set to 39.
- Performed **chi-square test** on categorical data and **ANOVA** on numerical data to calculate **p-values** in relation to the target variable.
- **Achieved an 81% accuracy** rate following hyperparameter tuning, representing an **8% improvement**, with the final fine-tuned model implemented using **XGBoost**.
- Deployed a Flask application, along with an '.exe' file which generates the results in an excel file.

Music Generation | *Python, Music21, TensorFlow, Streamlit, AudioCraft* GitHub

- Developed an advanced model capable of **predicting notes** based on a melody snippet supplied by the user.
- Collaborated with a teammate to incorporate the AudioCraft library, enabling the generation of music from **textual prompts** provided by the user.

Health Insurance Cost Predictor | *Python, Scikit-Learn, Streamlit* GitHub

- Developed a health insurance predictor based on region, age, BMI, and additional factors.
- Employed machine learning algorithms like linear regression, random forest, and gradient boosting.
- Gradient Boosting, with its **accuracy being 88%**, was selected as the most efficient model for prediction.

EXPERIENCE

BloodConnect Foundation | *Volunteer* July 2023 – November 2023

- Engaged with over 120 organizations, residential communities, and educational institutions to coordinate blood donation campaigns
- Attended to over 20 emergency calls on a weekly basis, along with arranging necessary blood/platelet units for recipients.
- Participated in over 15 blood donation campaigns and awareness sessions alongside fellow volunteers.

MISCELLANEOUS

- Achieved a global rank of 390 out of 3,324 participants in CodeChef Starters 132 (Division 3).
- Solved 450+ DSA problem solving questions on coding platforms such as CodingNinjas, GeekForGeeks, and LeetCode.