Sanjit Davuluri

Hyderabad, India   +91 9912396296   sanjit.davuluri@gmail.com   LinkedIn   Leetcode

Education

Bachelor of Technology in Computing and Data Science
Sai University, Chennai | 2022 - 2026

Coursework

- Deep Learning, Machine Learning, Artificial Intelligence, Reinforcement Learning, Data Analysis, Advanced Programming, Speech Analysis and Technology, Database Management Systems, Operating Systems
- CGPA : 8.43

Experience

Python Developer Intern | Seneca Global

June 2023-August 2023

- Internship was divided into two phases: training and project work. Acquired valuable skills in web development during the initial weeks.
- Built a RESTful API using Flask to manage orders for the project.
- Gained hands-on experience in designing APIs and working with frameworks like Flask. Explored different approaches and resolved technical challenges encountered during the project.
- Internship Letter

Projects

Image Classification, Deep Learning

- Developed a deep learning model for rock type classification using a dataset of 505 images.
- Implemented Transfer Learning (TL) with pre-trained CNNs (InceptionResNetV2, DenseNet201, ConvNeXtTiny) and Fine-Tuning (FT) to enhance model performance.
- Applied image preprocessing techniques, including resizing, normalization, and label encoding.
- Trained and optimized the model using TensorFlow and Keras, leveraging dropout and adaptive learning rate tuning.
- Evaluated performance using confusion matrix, precision, recall, and accuracy metrics.

Age Group Prediction, Machine Learning

- Developed a classification model to predict age groups using the NHANES dataset.
- Performed data preprocessing, including feature selection, handling missing values, and scaling with StandardScaler. Applied Label Encoding for categorical variables.
- Built and evaluated multiple machine learning models, including: Logistic Regression, Random Forest Classifier, Decision Tree Classifier, Support Vector Classifier (SVC), Gradient Boosting Classifier.
- Optimized model performance using GridSearchCV for hyperparameter tuning.
- Evaluated model accuracy using confusion matrix, precision, recall, and F1-score.

Car Price Prediction, Machine Learning

- Developed a supervised regression model to predict car prices using a dataset with various car features.
- Performed data preprocessing, including handling missing values, feature scaling, and encoding categorical variables.
- Explored feature importance using correlation heatmaps and feature selection techniques.
- Built and evaluated multiple regression models, including: Linear Regression, Decision Tree Regressor, Random Forest Regressor, Gradient Boosting Regressor (XGBoost, LightGBM, CatBoost), Support Vector Regressor (SVR)
- Tuned hyperparameters using GridSearchCV and RandomizedSearchCV to optimize model performance.
- Evaluated model performance using R^2 score, Mean Absolute Error (MAE), Mean Squared Error (MSE), and RMSE.

Spoken Digit Classification, Speech Technology

- Implemented a machine learning pipeline to classify spoken digits (04–09) using the AudioMNIST dataset.
- Extracted MFCC features from audio files using Librosa to represent time-series speech data effectively.
- Applied Label Encoding and StandardScaler for preprocessing and normalization.
- Trained a Support Vector Machine (SVM) with a linear kernel for digit classification.
- Evaluated model performance using accuracy score and classification report from scikit-learn, achieving strong accuracy and class-level insights.
- Ensured robust training with a train-test split and clear error handling during audio file loading.

Fake Voice Detector, Speech Technology and Deep Learning

- Developed and deployed a voice anti-spoofing model to classify .wav audio inputs as bonafide or spoofed, using spectrogram-based deep learning techniques.
- Preprocessed raw audio data using Librosa to generate mel-spectrograms, converting .wav files into 224×224 RGB spectrogram images for CNN training.
- Fine-tuned MobileNetV2 (transfer learning on ImageNet weights) to achieve ~90% accuracy on a real-world dataset of 2,500+ spectrogram images.
- Built a custom end-to-end pipeline including data augmentation, binary classification, and real-time prediction from uploaded audio files.