|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Test Description** | **Expected Results** | **Actual Results** | **Pass/Fail** | **Notes** |
| 1 | Install required libraries (praw, SpeechRecognition, pydub, vosk, soundfile) | All libraries installed successfully | All libraries installed successfully | Pass | Dependencies were installed without issues |
| 2 | Convert MP3 to WAV format using FFmpeg | Audio successfully converted | Audio successfully converted | Pass | Conversion worked, but required correct file path |
| 3 | Scrape Reddit posts for player performance analysis | Posts stored in CSV file | Posts stored in CSV file | Pass | Initially tried to scrape data from X however Some posts were missing due to API limitations |
| 4 | Extract player statistics and store in CSV file | Data saved correctly in CSV | Data saved correctly in CSV | Pass | CSV file generated successfully |
| 5 | Transcribe audio to text using Google Speech Recognition | Transcribed text matches the audio content | Transcribed text matches the audio content | Pass | Minor inaccuracies observed in transcription |
| 6 | Perform offline speech-to-text conversion using Vosk | Transcription stored successfully | Transcription stored successfully | Pass | Minor inaccuracies observed in transcription |
| 7 | Extract text from PDFs using pdfminer and OCR | Extracted text saved in CSV | Extracted text saved in CSV | Pass | OCR used when text extraction failed |
| 8 | Clean extracted text (lowercase, remove punctuation, tokenisation) | Cleaned dataset stored successfully | Cleaned dataset stored successfully | Pass | No major issues encountered |
| 9 | Perform sentiment analysis using VADER | Generate sentiment scores | Generate sentiment scores | Pass | VADER provided better accuracy than TextBlob |
| 10 | Perform Exploratory Data Analysis (word cloud, removing NaN, duplicates) | Meet all EDA requirements | Meet all EDA requirements | Pass | Some stopwords needed additional removal |
| 11 | Feature extraction using BOW & TF-IDF | Extracted features successfully | Extracted features successfully | Pass | Accuracy improved with TF-IDF tuning |
| 12 | Train/Test split experimentation (90:10, 80:20, etc.) | Determine the best split | 90:10 gave the best results | Pass | More training data improved accuracy |
| 13 | Compare different ML models (Naïve Bayes, SVC, Logistic Regression) | Generate a comparison plot | Generate a comparison plot | Pass | Random Forest performed best |
| 14 | Run code in Google Colab | Code runs without issues | Code runs without issues | Pass | No major compatibility issues found |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |