

Date	21 JULY 2024
Team ID	team-740025
Project Title	Unlocking Silent Signals :Decoding Body Language With Mediapipe
Maximum Marks	6 Marks

Data Exploration and Preprocessing Report

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

Section	Description
Data Overview	<p><u>Dimension:</u> 4595 rows \times 2004 columns</p> <p><u>Descriptive statistics:</u></p> <pre> 34]: x1 y1 z1 v1 x2 y2 z2 v2 x3 y3 ... z499 v499 x500 y500 z500 v500 0 0.460887 0.595566 -1.121984 0.999785 0.484884 0.513489 -1.069172 0.999606 0.504262 0.512900 ... -0.001009 0.0 0.532710 0.511281 0.033799 0.0 0.5 1 0.461394 0.595729 -1.238655 0.999784 0.483257 0.513369 -1.173208 0.999604 0.501266 0.512291 ... -0.000845 0.0 0.529603 0.512033 0.034787 0.0 0.5 2 0.461530 0.595861 -1.264079 0.999784 0.481376 0.513183 -1.198528 0.999603 0.498597 0.511612 ... -0.000800 0.0 0.528513 0.512119 0.034527 0.0 0.5 3 0.463032 0.599466 -1.253284 0.999765 0.481420 0.515575 -1.181970 0.999566 0.498603 0.513372 ... -0.001560 0.0 0.529767 0.519925 0.032815 0.0 0.5 4 0.465295 0.607626 -1.228310 0.999726 0.482713 0.522248 -1.166696 0.999484 0.499419 0.519148 ... -0.005636 0.0 0.538700 0.526081 0.030072 0.0 0.5 4590 0.684548 0.511105 -1.037075 0.999985 0.706993 0.433930 -0.976545 0.999968 0.723656 0.435579 ... -0.003546 0.0 0.748149 0.430420 0.029645 0.0 0.7 4591 0.684455 0.511044 -0.985705 0.999985 0.706319 0.433798 -0.928487 0.999967 0.722609 0.435238 ... -0.003345 0.0 0.747559 0.434463 0.029979 0.0 0.7 4592 0.684569 0.509095 -1.019472 0.999984 0.706292 0.430161 -0.948244 0.999966 0.722583 0.429946 ... -0.003653 0.0 0.751030 0.432241 0.029848 0.0 0.7 4593 0.687874 0.509134 -1.102308 0.999953 0.708338 0.430230 -1.033438 0.999922 0.724405 0.429943 ... -0.004866 0.0 0.760940 0.428284 0.030739 0.0 0.7 4594 0.691391 0.509307 -1.146814 0.999941 0.710918 0.430336 -1.065239 0.999904 0.726889 0.430056 ... -0.003975 0.0 0.764780 0.431389 0.030545 0.0 0.7 </pre> <p>4595 rows \times 2004 columns</p>

Loading Data	<p>▼ Training model using Scikit Learn</p> <p>Read in collected data and process</p> <pre> 9]: import pandas as pd from sklearn.model_selection import train_test_split 0]: df = pd.read_csv('coords.csv') 1]: df.head() 1]: class x1 y1 z1 v1 x2 y2 z2 v2 x3 ... z499 v499 x500 y500 z500 v500 x501 0 Happy 0.460887 0.595566 -1.121984 0.999785 0.484884 0.513489 -1.069172 0.999606 0.504262 ... -0.001009 0.0 0.532710 0.511281 0.033799 0.0 0.537714 1 Happy 0.461394 0.595729 -1.238655 0.999784 0.483257 0.513369 -1.173208 0.999604 0.501266 ... -0.000845 0.0 0.529603 0.512033 0.034787 0.0 0.534479 2 Happy 0.461530 0.595861 -1.264079 0.999784 0.481376 0.513183 -1.198528 0.999603 0.498597 ... -0.000800 0.0 0.528513 0.512119 0.034527 0.0 0.533444 3 Happy 0.463032 0.599466 -1.253284 0.999765 0.481420 0.515575 -1.181970 0.999566 0.498603 ... -0.001560 0.0 0.529767 0.519925 0.032815 0.0 0.534700 4 Happy 0.465295 0.607626 -1.228310 0.999726 0.482713 0.522248 -1.166696 0.999484 0.499419 ... -0.005636 0.0 0.538700 0.526081 0.030072 0.0 0.543335 5 rows × 2005 columns 2]: df.tail() 2]: class x1 y1 z1 v1 x2 y2 z2 v2 x3 ... z499 v499 x500 y500 z500 v500 x501 4590 Fight 0.684548 0.511105 -1.037075 0.999985 0.706993 0.433930 -0.976545 0.999968 0.723656 ... -0.003546 0.0 0.748149 0.430420 0.029645 0.0 0.753331 4591 Fight 0.684455 0.511044 -0.985705 0.999985 0.706319 0.433798 -0.928487 0.999967 0.722609 ... -0.003345 0.0 0.747559 0.434463 0.029979 0.0 0.752515 </pre>
Data Transformation	<pre> [35]: y [35]: 0 Happy 1 Happy 2 Happy 3 Happy ... 4590 Fight 4591 Fight 4592 Fight 4593 Fight 4594 Fight Name: class, Length: 4595, dtype: object [36]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=1234) [37]: y_train [37]: 1761 Sad 1128 Happy 1416 Sad 1571 Sad 3695 Victorious ... 664 Happy 3276 Victorious 1318 Sad 723 Happy 2863 Sad Name: class, Length: 3676, dtype: object </pre>
Feature Engineering	Attached the codes in final submission.
Save Processed Data	-