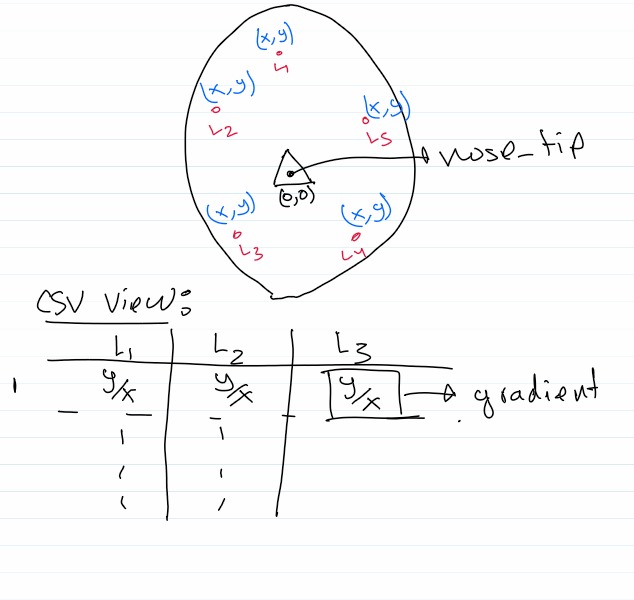
**Facial Landmark Extraction:**

The MediaPipe's face mesh solution is used to detect and extract facial landmarks in an image. For each region (29) (like lips, eyes, eyebrows, nose, and cheeks), multiple landmarks are extracted, each landmark is represented by its x and y coordinates. The landmarks are then normalized to remove any effects of the face's position or size in the image. This is achieved by translating the landmarks so that a reference point (the nose tip in this case) is at the origin, and then scaling the landmarks so that the distance between two other reference points (the outer corners of the eyes) is 1 (Use the same explanation as before) The normalized x and y coordinates of each landmark are then divided (y/x) to get a single value for each landmark (which is the gradient). This is done because when stoke occurs the facial landmarks change positions as the face changed shape which changes the gradient of the line connecting the landmarks and the nose tip.



**CSV File Format**

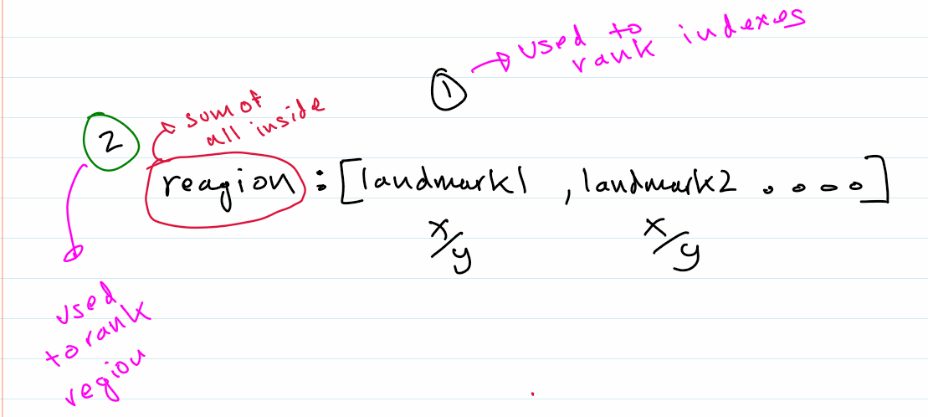
Each normalized landmark value is associated with a group (the region it belongs to) and an index (its position within the group). The group and index are used to create the column names in the CSV file. For example, 'lipsUpperOuter\_61' refers to the normalized value of the landmark at index 61 in the 'lipsUpperOuter' region. So, for each image, a row is added to the CSV file. This row contains the filename , the normalized values of all the landmarks and binary classification of Stroke and Non-stroke Face. The number of columns in the CSV file is 230, which includes one column for the filename, 228 columns for the normalized landmark values and one for the label.

**Training and Testing is Similar to what we did**

**XAI**

The exported CSV file is used to train the data using Random Forest Classifier Using the following Hyperparameters; {'criterion': 'entropy', 'max\_depth': None,'max\_features': 5,'min\_samples\_leaf': 1, 'min\_samples\_split': 2,'n\_estimators': 60}. The model is trained 100 times and the feature (landmark) importances of each feature is calculated using rfc.feature\_importances\_. The feature importances indicate how much each feature contributes to the predictions of the model. For each iteration, the feature importances are calculated and are averaged. The features are grouped into regions (like lips, eyes, eyebrows, nose, and cheeks), and the importances of the features within each region are summed up to get the total importance of that region. A total of 28 (nose tip excluded as its importance is 0) regions are then ranked by average importance.

|  |  |
| --- | --- |
| Region | Importance |
| rightCheek | 14.91738508 |
| leftCheek | 12.39699461 |
| lipsUpperOuter | 6.226201432 |
| lipsUpperInner | 6.087068932 |
| lipsLowerInner | 5.827017569 |
| lipsLowerOuter | 5.154195959 |
| rightEyebrowUpper | 3.365137215 |
| rightEyeLower3 | 3.291075876 |
| rightEyeLower2 | 2.990565034 |
| leftEyeLower3 | 2.97990347 |
| leftEyebrowUpper | 2.96183422 |
| rightEyeLower1 | 2.834417604 |
| leftEyeLower2 | 2.776542398 |
| rightEyeLower0 | 2.761890821 |
| leftEyeLower1 | 2.703630633 |
| leftEyeLower0 | 2.689930877 |
| rightEyeUpper2 | 2.513535948 |
| leftEyeUpper2 | 2.451974963 |
| leftEyeUpper1 | 2.243994292 |
| rightEyeUpper1 | 2.217633024 |
| leftEyebrowLower | 2.216135794 |
| rightEyebrowLower | 2.183288477 |
| rightEyeUpper0 | 2.107114544 |
| leftEyeUpper0 | 2.032525837 |
| noseRightCorner | 0.605737126 |
| noseLeftCorner | 0.559181465 |
| noseBottom | 0.470354611 |
| midwayBetweenEyes | 0.434732195 |
| SUM | 100 |



The model is then trained on the top 5, 10, 15 and 20 regions and evaluated on model accuracy. This step is repeated 100 times to see how many regions do we need to attain the highest accuracy possible. And these are the following results:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Regions | Max | Min | Avg | Accuracy (2sf) | Precision | Recall | F1-Score |
| Top 5 | 0.9325 | 0.905 | 0.918275 | 0.93 | 0.93 | 0.93 | 0.93 |
| Top 10 | 0.935 | 0.9075 | 0.9196 | 0.94 | 0.94 | 0.93 | 0.93 |
| Top 15 | 0.93 | 0.895 | 0.91385 | 0.93 | 0.93 | 0.93 | 0.93 |
| Top 20 | 0.9275 | 0.895 | 0.912625 | 0.93 | 0.93 | 0.92 | 0.93 |

The process is again repeated between the top 1 to 5 regions and these are the results:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Regions | Max | Min | Avg | Accuracy (2sf) | Precision | Recall | F1-Score |
| Top 1 | 0.8925 | 0.87 | 0.881725 | 0.89 | 0.89 | 0.89 | 0.89 |
| Top 2 | 0.9175 | 0.885 | 0.903825 | 0.92 | 0.92 | 0.92 | 0.92 |
| Top 3 | 0.94 | 0.9125 | 0.925175 | 0.94 | 0.94 | 0.94 | 0.94 |
| Top 4 | 0.94 | 0.905 | 0.921025 | 0.94 | 0.94 | 0.94 | 0.94 |
| Top 5 | 0.9325 | 0.905 | 0.918275 | 0.93 | 0.94 | 0.92 | 0.93 |

The Top 3 regions produce the highest possible accuracy and consists of a total of 49 Landmarks.

