INTRODUCTION

The goal of the project is the development of a classifier able to verify the correctness of an identity claim based on biometric characteristics of a subject, in details through spoken utterance. The task is to compare an unknown utterance with those belonging to the claimed identity and check if they are similar or not.

The problem can be cast as a binary problem over pairs of utterances. Given a pair of utterances the classifier should decide if the unknown one and the claimed identity one correspond to the same speaker (label 1) or different speaker (label 1).

In this use case the utterances are represented in terms of speaker embeddings. The embeddings are 5-dimensional, continuous-valued vectors, and features have no specific physical interpretation.

INFORMATION ABOUT THE DATASET

Each sample of the dataset is composed of a pair of embeddings for a total dimension of 10.

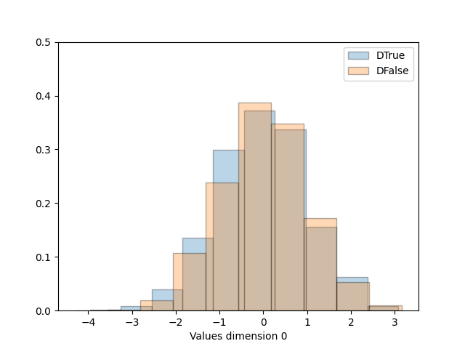
The training dataset is composed of 3500 samples: 43% belong to label 1 and 57% to class 0.

The test dataset instead is composed of 7500 samples in total: 33% belong to label 1 and 67% to class 0.

Both datasets are slightly imbalanced, with the different speaker class (label 0) having slightly more samples. Pairs belong to either the male or female gender, which is not provided by the datasets.

DATASET FEATURES ANALYSIS

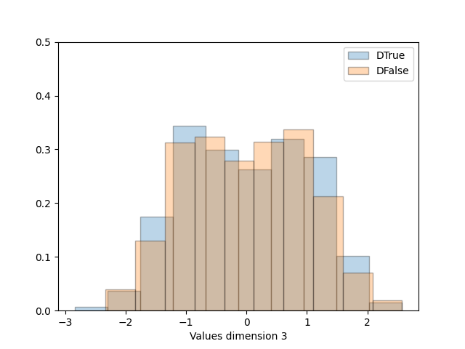
For the analysis of the dataset (which has been normalized first using z-score) is crucial to plot the distribution of the features over histograms. For each feature has been made a distinction between samples belonging to different classes.

A graph of values

Description automatically generatedA graph of values

Description automatically generated

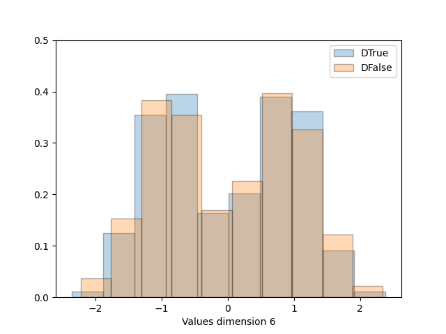
Figure Figure Figure

A graph of values

Description automatically generatedA graph of values

Description automatically generated

Figure Figure Figure

A graph of values

Description automatically generatedA graph of values

Description automatically generated

Figure Figure Figure

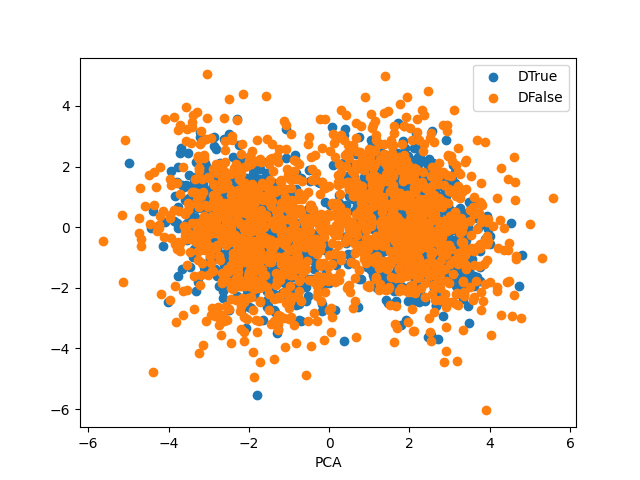
A graph of values

Description automatically generated

Figure

All feature distributions follow a gaussian one with an exception for feature 1, 3, 6, 8. These last reflect a gaussian with 2 components due to the presence of both genders, male and female.

To have a general view over our dataset it has been applied a PCA transformation with 2 dimensions on the left and a LDA transformation on the right. In the PCA graph is evident the presence of two clusters for both the class and in the LDA graph the distribution of the single feature of both classes overlap. These two clusters can be explained by the presence of spoken utterance belonging to the male and female gender. With these results is expected that the data cannot be well separated linearly so a non-linear model should perform better in the classification task.

A graph of a number of objects

Description automatically generated with medium confidence

Figure Figure

DATASET FEATURES CORRELATION ANALYSIS

The graphs below represent the correlation between each feature of the training dataset.

Each Heat Map shows the absolute value of the Pearson Correlation Coefficient. On the left it has been performed on the whole dataset and on the right the graph of the Explained Variance.

A graph of a crossword puzzle

Description automatically generatedA graph with a line

Description automatically generated

Figure Figure

Looking at figure 13 features 2-7 and 4-9 are the highest correlated. For what concern the others one there is a slight correlation such as also no correlation between some of them. Figure 14 tell us that removing two dimensions will keep very high the variance of the dataset so without loosing a lot of information which can be decisive during the training phase of our classifier.

Below there are the heatmap of the Pearson correlation coefficient of the dataset features belonging respectfully on the left to the label 0 (False) and on the right to the label 1 (True). They are pretty similar with exception for features (1-6), (2-7), (3-8), (4-9), (5-10) which are more correlated in the False samples compared to the True samples.

A close-up of a graph

Description automatically generatedA blue squares with white lines

Description automatically generated with medium confidence

QUI VA AGGIUNTA LA PARTE DEL NAÏVE BAYES E DEL TIED COVARIANCE. VA FATTA L’ANALISI DEI DUE MODELLI ANALIZZANDO LE DUE HEATMAP DI CUI SOPRA

TRAINING PROTOCOL

For the training phase it has been used the K-fold cross validation with 5 folds.

The target application working point given by the project is defined by the triple (πT = 0.1, Cf n = 1, Cf p = 1). During this phase two others application working points have been used: (πT = 0.5, Cf n = 1, Cf p = 1) and (πT = 0.9, Cf n = 1, Cf p = 1).

The performance of the different classifiers is measured in terms of Detection Cost Function (DCF). The classifier with the best performance will be the one with the lowest DCF.

GAUSSIAN CLASSIFIER

In this section the classifier has been trained with several gaussian models: MVG (Multi Variate Gaussian), NB (Naïve Bayes), TC (Tied Covariance) and TNB (Tied Naïve Bayes).

RAW Z-Score

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| **Model** | πT=0.1 | πT=0.5 | πT=0.9 | πT=0.1 | πT=0.5 | πT=0.9 |  |
|  |  |  |  |  |  |  |  |
| MVG | 0.186 | 0.070 | 0.130 | 0.209 | 0.070 | 0.130 |  |
| NB | 0.998 | 0.946 | 0.992 | 1.000 | 0.946 | 0.992 |  |
| TC | 1.000 | 0.919 | 0.981 | 1.000 | 0.919 | 0.981 |  |
| TNB | 0.990 | 0.953 | 0.998 | 0.992 | 0.953 | 0.998 |  |
|  |  |  |  |  |  |  |  |

RAW PCA 8 Z-Score PCA 8

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| **Model** | πT=0.1 | πT=0.5 | πT=0.9 | πT=0.1 | πT=0.5 | πT=0.9 |  |
|  |  |  |  |  |  |  |  |
| MVG | 0.186 | 0.070 | 0.130 | 0.209 | 0.070 | 0.130 |  |
| NB | 0.998 | 0.946 | 0.992 | 1.000 | 0.946 | 0.992 |  |
| TC | 1.000 | 0.919 | 0.981 | 1.000 | 0.919 | 0.981 |  |
| TNB | 0.990 | 0.953 | 0.998 | 0.992 | 0.953 | 0.998 |  |
|  |  |  |  |  |  |  |  |