**ReadMe AI 1 Face Detector**

Tasks completed:

* Skeleton code

Extra tasks:

* Results analysis
* More than 1 algorithm implemented.
* Created confusion matrix heatmaps.
* Cross validations given for each respective algorithm.
* Researched into deep learning algorithms.
* Documentation provided (instructions for setup, algorithm analysis, etc.)

Here is a link to a google drive folder, which contains the datasets which are required to run the assignment: <https://drive.google.com/drive/folders/1P0MyinWETN5SdRqQwTs0AqIcUn7oOULx?usp=sharing>

**Algorithms Used:**

SVM and KNN, each is written about in a paragraph along with results such as accuracy and cross-validation provided.

**KNN (k nearest neighbors)**

KNN is a supervised machine learning classifier (it is non-parametric), as described in the name, it uses proximity to create predictions and clarifications regarding the area of an individual point of data. it is primary used as a classification algorithm, which works off assuming that similar data points can be found near to others. In this case, finding data points close to one another to identity a face within a given image.

The accuracy of the KNN classifier on the scaled training dataset was: 0.9965 (99.65%, to 2f) the accuracy of the KNN classifier on the scaled testing dataset was: 0.9957 (99.57% to 2f)

The cross-validation of the KNN classifier on the scaled dataset was: [0.99629972 0.99306198 0.99537465 0.99167438 0.99352152]

グラフ, 棒グラフ

中程度の精度で自動的に生成された説明

The heatmap shows that 7500 true positives were correctly identified, along with 3200 true negatives. Along with 4 false positives and 43 false negatives.

(reference: <https://www.ibm.com/topics/knn>)

**SVM (support vector machine)**

SVM (Support vector machine) is a supervised machine learning classifier, it maximizes the accuracy of a given model without overfitting the data passed into it. SVM is well suited to analyzing data which contains large numbers (e.g. 1000s). Common uses for the SVM classifier include, facial recognition, bioinformatics, text extraction, voice/speech recognition, etc. SVM works by mapping out data in a high-dimensional feature space, so data points are categorized, even if the data is not linearly separable.

(reference: <https://www.ibm.com/docs/en/spss-modeler/saas?topic=models-about-svm>)

グラフ, 棒グラフ

自動的に生成された説明

The heatmap shows that 7500 true positives were correctly identified, along with 3300 true negatives. Along with 3 false positives and 29 false negatives.

The accuracy of SVM classifier on the scaled training dataset was: 0.9993 (99.93% to 2f)

The accuracy of the SVM classifier on the scaled testing dataset was: 0.9970 (99.70% to 2f)

The cross-validation of the SVM classifier on the scaled dataset was: [0.99768733 0.99491212 0.99491212 0.99398705 0.99490976]

**Conclusion:**

In conclusion, the results from my experiments (accuracy, etc) show that the SVM (Support vector machine) classifier model is the best suited/most accurate out of the two models I have implemented and tested. KNN was found to be somewhat less reliable than SVM through the accuracy results, though this was a small difference. Another measure of accuracy that I used was the confusion matrix, showing that the SVM had a lower rate of false positives and negatives when compared to the KNN classifier.