Name: Nguyen Quang Khoi

SSID: 888881

Part 1: Classification of Facial Expressions

1.2.1.

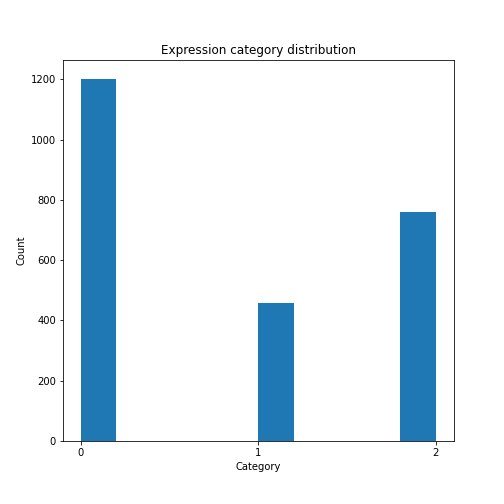


Figure 1: Three images from the dataset

1.2.2

Include your histogram plots

**Answer this question**: Is the dataset balanced?

.

1.3.1

\

List the accuracy of the KNN and other metrics you might use.

Include two examples of the misclassified images.

1.3.2

List the hyperparameters you found.

1.3.3

Create a table of the classifiers you used, there best hyperparameters if you found them and the metrics you found.

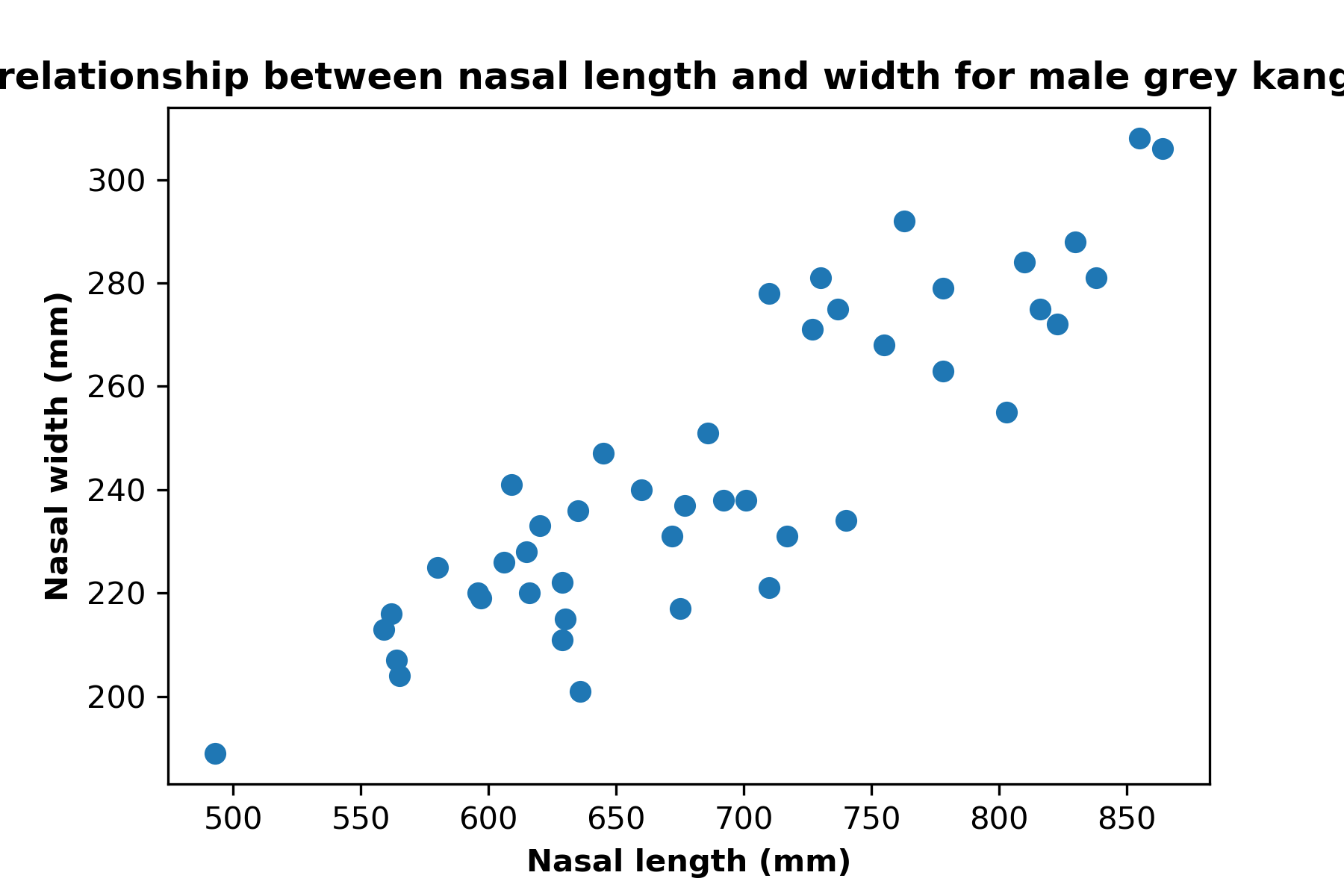
**Answer this question**: Which was your best pipeline or classifier and with what parameters?

**Answer this question**: Did you beat the baseline classifier? If not, why do you think so?

Part 2: Regression to estimate the width of a grey kangaroo’s nose

2.2.1.

Include the plot you created



2.3.2

List the result of your R2 score

R^2 score: 0.7051

2.3.3

List the result of your mean R2 score

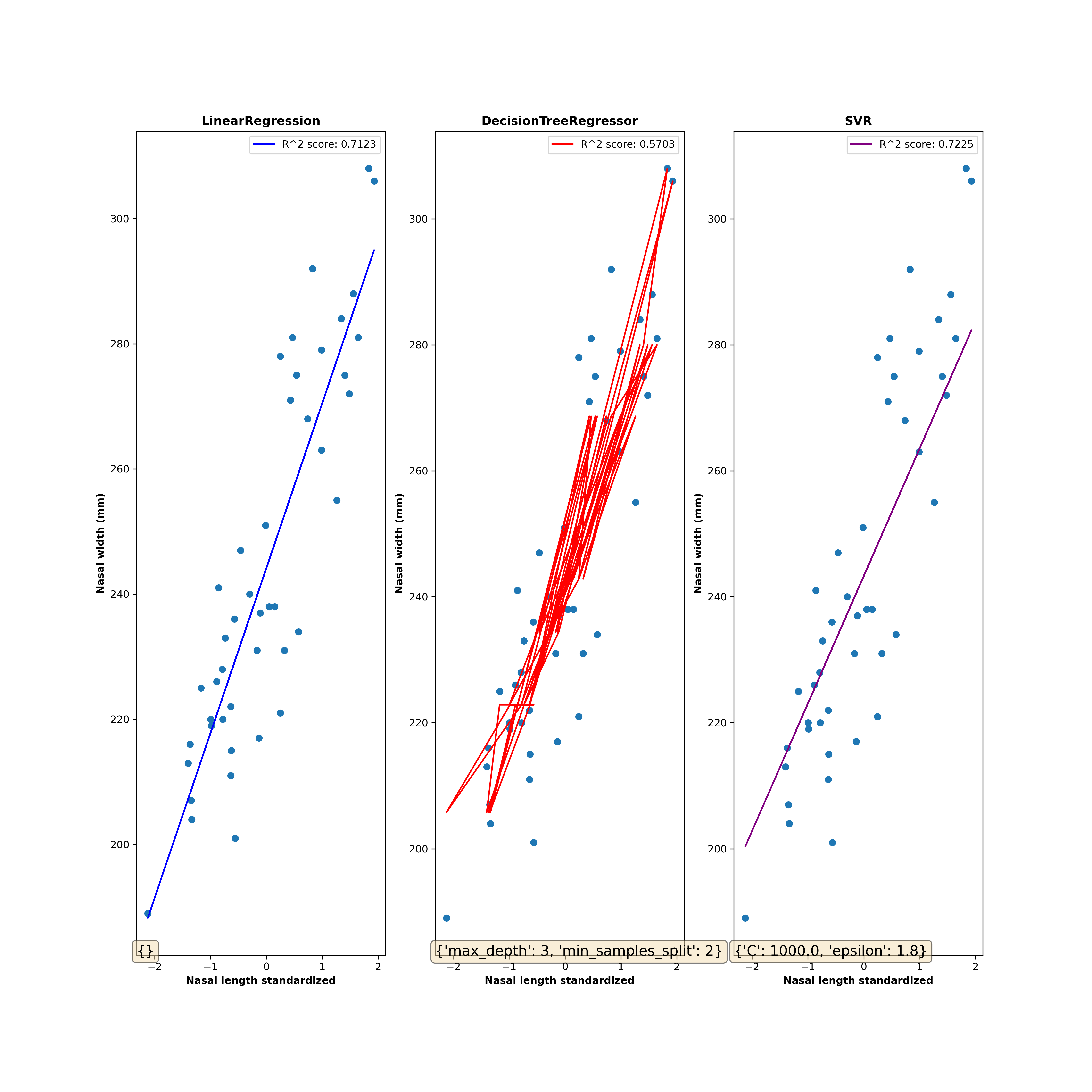
Mean R^2 score: 0.7123

2.3.4

In a table(preferably), display each regressor with it corresponding score. Include the figure displaying your solution in a plot.

|  |  |
| --- | --- |
| Model | Mean R^2 score |
| Linear Regression | 0.7123 |
| Decision Tree Regression | 0.5703 |
| Linear SVM | 0.7225 |

**Answer this question**: Which is the best regressor?



Linear Regression is the best model with the highest R^2 score (0.7311)

2.4.

In a table (preferably), list the data imputation method and evaluation result.

|  |  |
| --- | --- |
| Imputation method | R^2 score on a Linear Regression model |
| Mean imputation | 0.5375 |
| KNN (k = 3) imputation | 0.6289 |

**Answer this question**: Which is the best data imputation method?

The best data imputation method is KNN (k = 3) imputation.