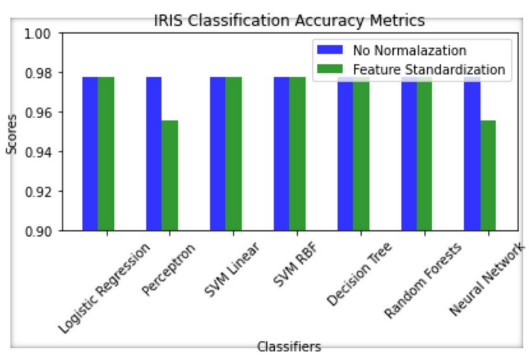
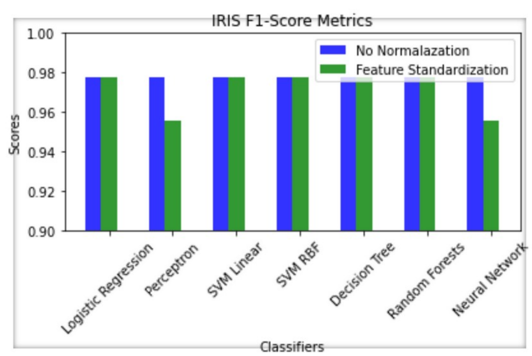
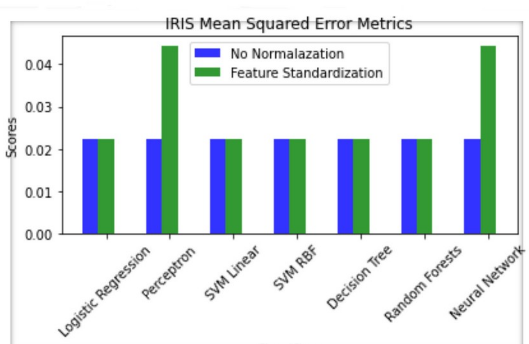
Iris plants dataset

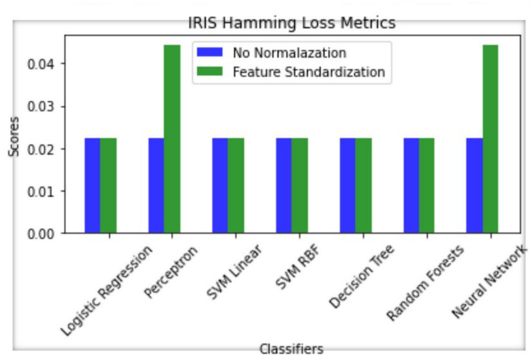
Description of the Iris dataset (samples, features, classes)

- Number of Instances:
  - 150 (50 in each of three classes)
- Number of Attributes:
  - 4 numeric, predictive attributes and the class
- Attribute Information (features):
  - sepal length in cm
  - sepal width in cm
  - petal length in cm
  - petal width in cm
- classes:
  - Iris-Setosa
  - Iris-Versicolour
  - Iris-Virginica









Conclusion about the performance of the classifiers on the Iris plant dataset

As we can observe from comparison plots of the classifiers' performance on the IRIS plants dataset, most of the classifiers score high, around 98%. Accuracy, f1-score are equal in the most of the classifiers. In addition, mean squared error metrics are equal with Hamming

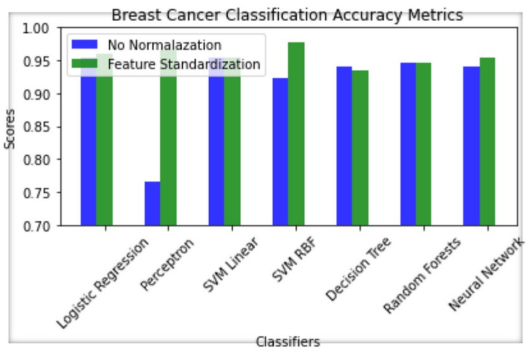
Loss. There is no best classifier for this dataset, but we can tell that, while we implement feature standardization in our dataset. Perceptron and Neural Networks are not as accurate as the other classifiers. However, feature standardization does not affect positively our metrics in all classifiers.

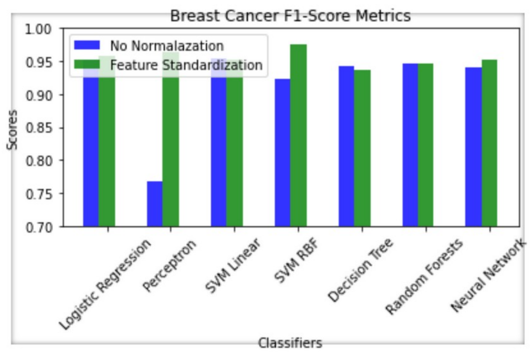
Breast cancer wisconsin (diagnostic) dataset

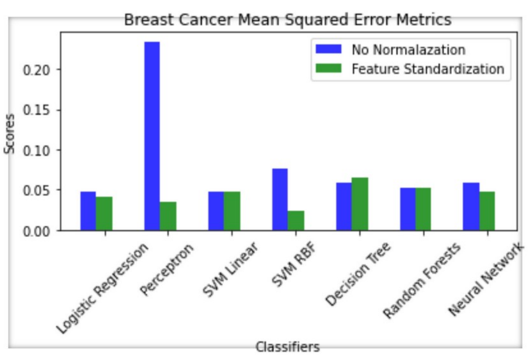
Description of the Breast cancer wisconsin (diagnostic) dataset (samples, features, classes)

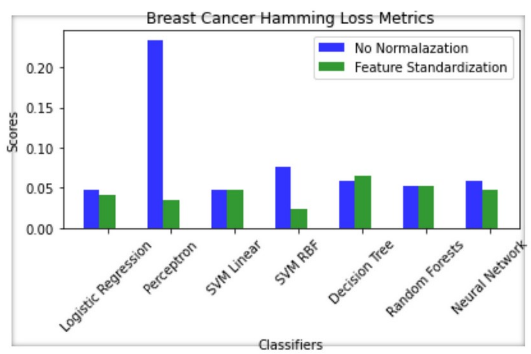
- Number of Instances:
  - 569
- Number of Attributes:
  - 30 numeric, predictive attributes and the class
- Attribute Information:
  - radius (mean of distances from center to points on the perimeter)
  - texture (standard deviation of gray-scale values)
  - perimeter
  - area
  - smoothness (local variation in radius lengths)
  - compactness (perimeter^2 / area - 1.0)
  - concavity (severity of concave portions of the contour)
  - concave points (number of concave portions of the contour)
  - symmetry
  - fractal dimension ("coastline approximation" - 1)
- class:
  - WDBC-Malignant
  - WDBC-Benign

Comparison plots of the classifiers' performance









Conclusion about the performance of the classifiers on the Breast cancer wisconsin (diagnostic) dataset

- In this specific dataset our results are not equal, while we can that sometimes standardization helps, sometimes it does not. We are confident that the best classifier for this dataset is SVM with RBF kernel(with standardization), while has the highest accuracy and f1-score, while has the lowest mean squared error and hamming loss. Also, we are confident that the worst classifier in this specific situation is Perceptron with no normalization (use raw features), because has the lowest accuracy, f1-score and the highest score in mean squared error, hamming loss metrics. In overall, we can say that classification with standardized features is better than using raw features.

Optical recognition of handwritten digits dataset

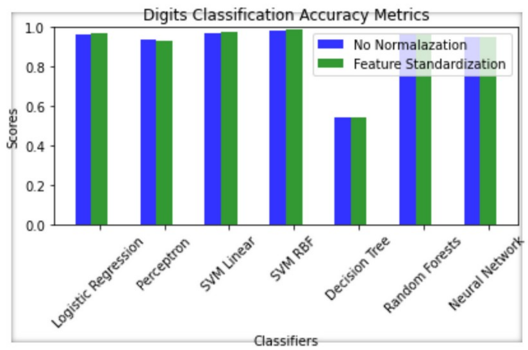
Description of the Optical recognition of handwritten digits dataset (samples, features, classes)

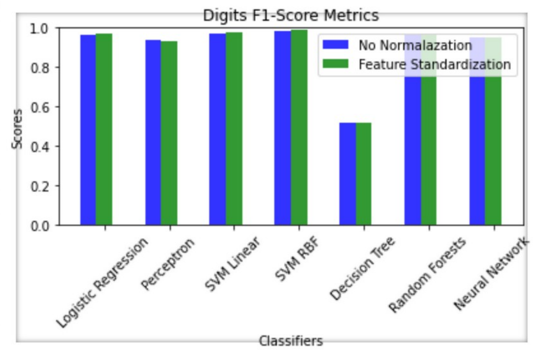
- Number of Instances:
  - 5620
- Number of Attributes:
  - 64
- Attribute Information:

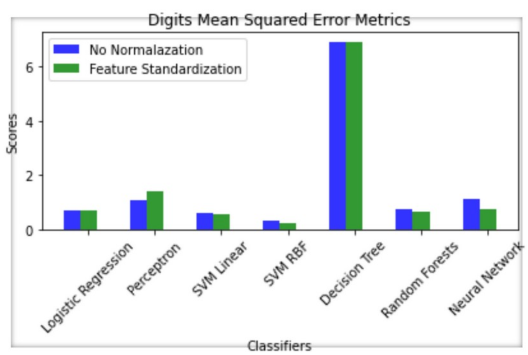
– 8x8 image of integer pixels in the range 0..16.

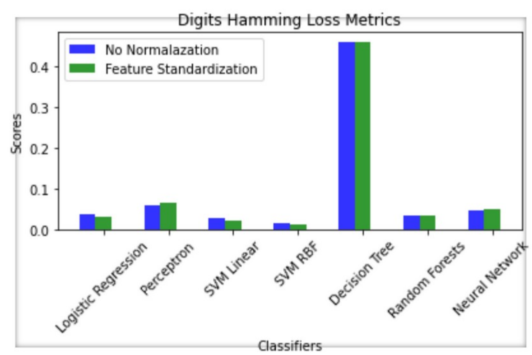
- The last attribute is:
  – the class code 0..9

Comparison plots of the classifiers' performance









Conclusion about the performance of the classifiers on the Optical recognition of handwritten digits dataset

- In optical recognition of handwritten digits dataset, we have many classifiers that provide us very good scores. Meaning, high accuracy, f1-score and lower scores in mean squared error and hamming loss metrics. However, we can say that the best classifier is SVM with RBF kernel, while has the highest scores in accuracy and f1-score (around 0.99), having also the lowest metrics on mean squared error and hamming loss. Both using standardization or not, scores are better than the other metrics, but with standardization of the features, the result is exceptional. However, it's worth mentioning that Decision Tree as a classification method in this dataset has pure performance, and it is not suggested.

Wine recognition dataset

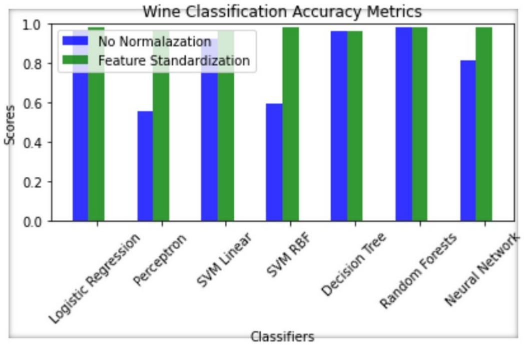
Description of the Wine recognition dataset (samples, features, classes)

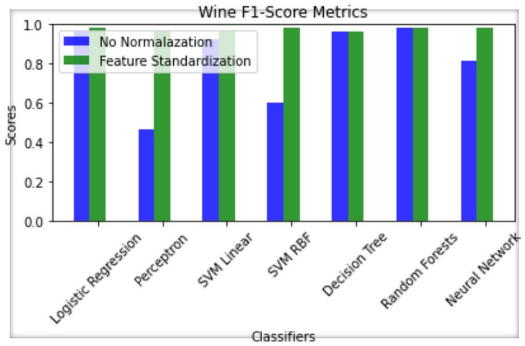
- Number of Instances:
  – 178 (50 in each of three classes)

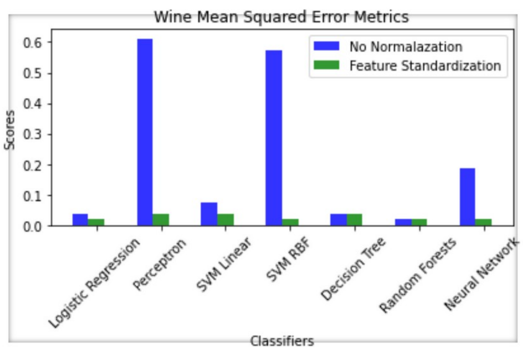
- Number of Attributes:
  - 13 numeric, predictive attributes and the class
- Attribute Information:
  - Alcohol
  - Malic acid
  - Ash
  - Alcalinity of ash

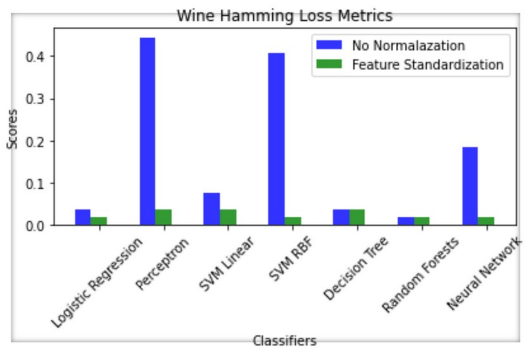
  - Magnesium
  - Total phenols
  - Flavanoids
  - Nonflavanoid phenols
  - Proanthocyanins
  - Color intensity
  - Hue
  - OD280/OD315 of diluted wines
  - Proline
- class:
  - class_0
  - class_1
  - class_2

Comparison plots of the classifiers' performance









Conclusion about the performance of the classifiers on the Wine recognition dataset

- While we examine our results of Wine recognition dataset, it is obvious that the standardization of his features provide us better results, while in the most of the classification methods, standardization of the features provide us great results. Logistic regression, SVM with RBF kernel, Random Forests and Neural Network with features standardized provide us great scores in accuracy and f1-score, and the lowest scores in mean squared error and hamming loss metrics. This example, make us clear that standardization of features many times accelerate the classification. One great example is SVM with RBF kernel, while with no normalization, has low performance, but standardization accelerate the results and make SVM with RBF kernel an exceptional classification method for this dataset (also Perceptron has quite equal results).

20 newsgroups text (vectorized) dataset

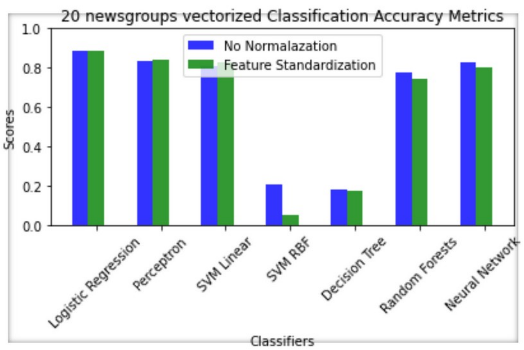
Description of the The 20 newsgroups text (vectorized) dataset (samples, features, classes)

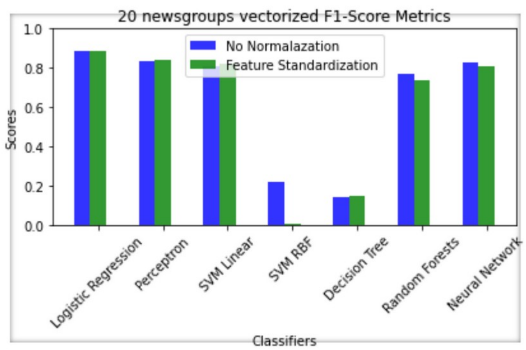
- Number of Samples:
  - 18846
- Features :
  - text
- Dimensionality :

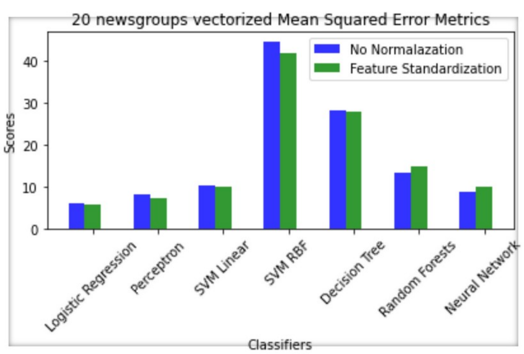
– 1

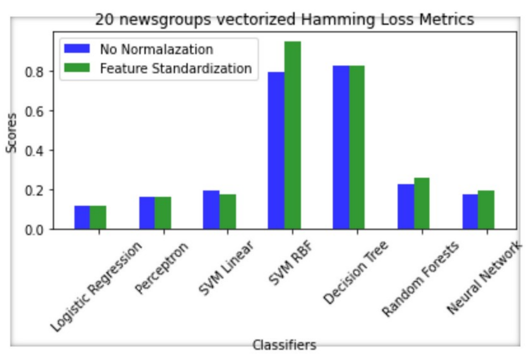
- Number of classes :
  – 20

Comparison plots of the classifiers' performance









Conclusion about the performance of the classifiers on the 20 newsgroups text (vectorized) dataset

- It is very interesting examining results of real '20 newsgroups text (vectorized)' dataset, while using different classifiers, results vary. With the first look we can say that feature standardization does not help every time. Also, it is obvious according to comparison plots that Logistic Regression is the best classifier with the highest accuracy, f1-score metrics and the lowest mean squared error, hamming loss scores. In addition, standardization provide us better metrics, but the difference is not big. Moreover, the other classifier provide us efficient results except from SVM with RBF kernel and Decision tree that their performances are unacceptable.

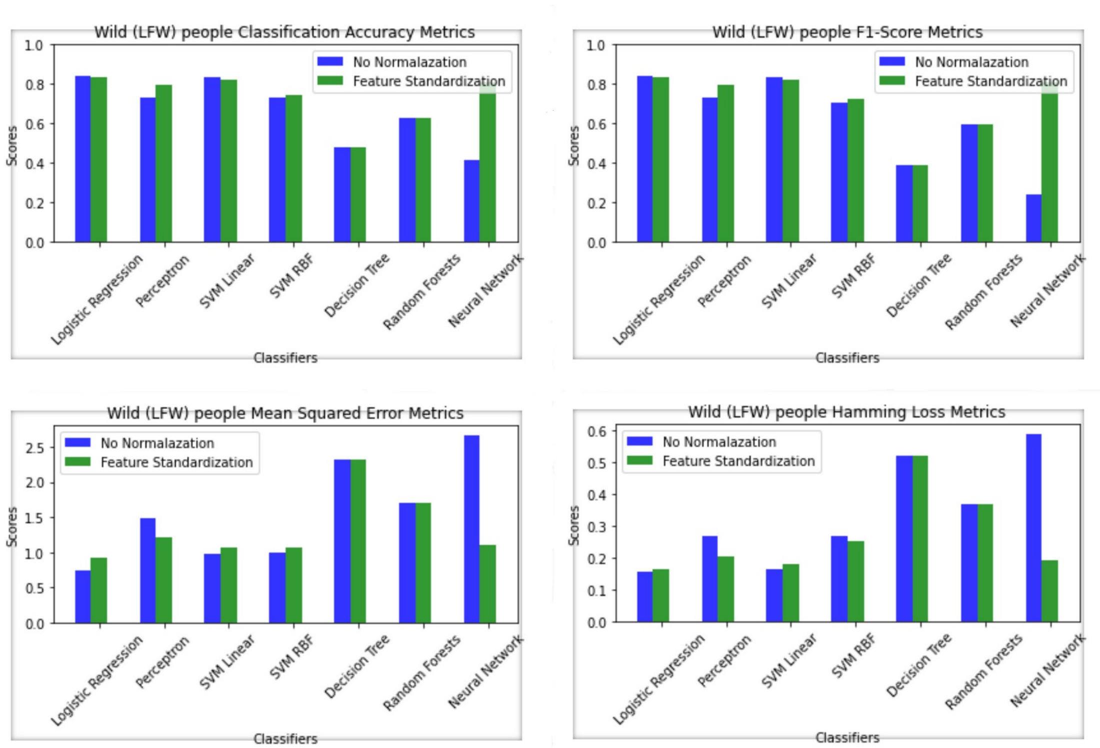
Labeled Faces in the Wild face recognition dataset

Description of the Labeled Faces in the Wild face recognition dataset (samples, features, classes)

- Number of Samples:
  – 13233

- Features :
  - real, between 0 and 255
- Dimensionality :
  - 5828
- Number of classes :
  - 5749

Comparison plots of the classifiers' performance



Conclusion about the performance of the classifiers on the Labeled Faces in the Wild face recognition dataset

- Dealing with pictures in this real dataset example, according to comparisons plots of the classifiers' performance, it is hard to reach an exceptional performance. Many of the classifiers score high but not perfect. Also feature standardization many times does not provide better results. The best performance according to our metrics has been reached by Logistic Regression with raw features. However, other classifiers' performances(i.e. SVM Linear, Neural Network with raw features) are close to logistic regression performance. It is worth mentioning the difference in Neural Network performance with feature standardization and without, while standardizing our features and using neural network has quite low performance contrary using raw features. Although, according to our metrics, decision tree is the worst classifier in overall for this example, performing low either using feature standardization either using raw feature.

Overall Conclusions about the performance of the classifiers on the toy datasets

- Using toy datasets for our experiments on classifiers' performance provide us different results. The most successful classifier for our toy datasets is SVM with RBF kernel while in all datasets performs great (especially with feature standardization). Also, except of specific cases, all classifiers perform acceptable, both using feature standardization and raw features. However, using feature standardization in overall is a good choice because in most cases accelerate the results. We could say that in some cases of our toy datasets, Perceptron and Decision Tree do not perform exceptional, but this does not mean that they are not a good choice because in most of the cases the differences are not huge.

Overall Conclusions about the performance of the classifiers on the text vs face recognition datasets

- The Performance of the classifiers in the text and face recognition datasets are different for some classifiers. Starting with feature standardization, is not always the best option, while many classifiers perform better with raw features. In both datasets, Logistic Regression is the best classifier, because accuracy, f1-score are the highest and mean squared error, hamming loss are the lower metrics comparing to the rest of classifiers metrics. Also, Perceptron and SVM with Linear kernel performs well in both datasets. However, SVM with RBF kernel has disappointing performance in text dataset, something that not happening in face recognition dataset. Identical results for Decision Tree in text dataset, nevertheless in face recognition dataset Decision Tree does not perform well but has not as bad performance as in text dataset. Also, Random forest as a classifier performs better in text dataset. In addition, Neural Network performs well in text dataset, however in face recognition dataset needs feature standardization to perform well.