

Exercise Sheet 7 (programming part)

Exercise 3: Implementing the Fisher Discriminant (25 + 25 P)

We would like to implement and test the Fisher Linear Discriminant on the Breast Cancer Dataset available in sklearn. The following code loads the dataset, standardizes the input features, and stores points of the two classes in the variables `XM` and `XB` (for malignant and benign).

```
In [1]: %matplotlib inline

import numpy, sklearn, sklearn.datasets
X, T = sklearn.datasets.load_breast_cancer(return_X_y=True)
X = (X - X.mean(axis=0)) / X.std(axis=0)
XM, XB = X[T==0], X[T==1]
```

(a) Create a function `w = fisher(X1, X2)` that takes as input the data for the two classes (two numpy arrays of size $N_1 \times d$ and $N_2 \times d$ respectively) and returns the Fisher linear discriminant.

```
In [2]: def fisher(X1, X2):

    ## -----
    ## TODO: replace by your code
    ## -----
    import solutions
    w = solutions.fisher(X1, X2)
    ## -----

    return w
```

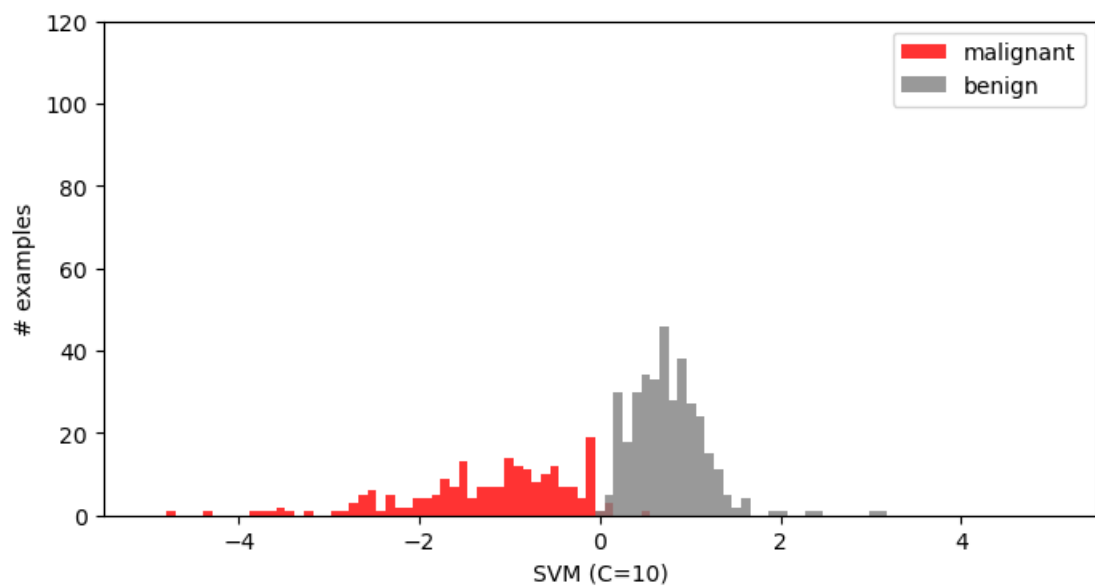
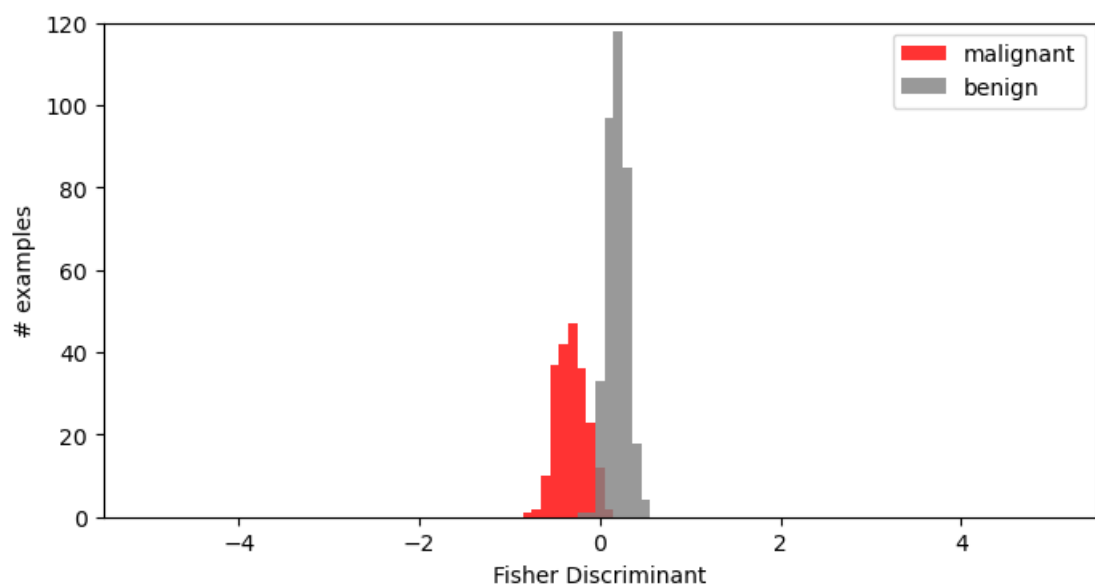
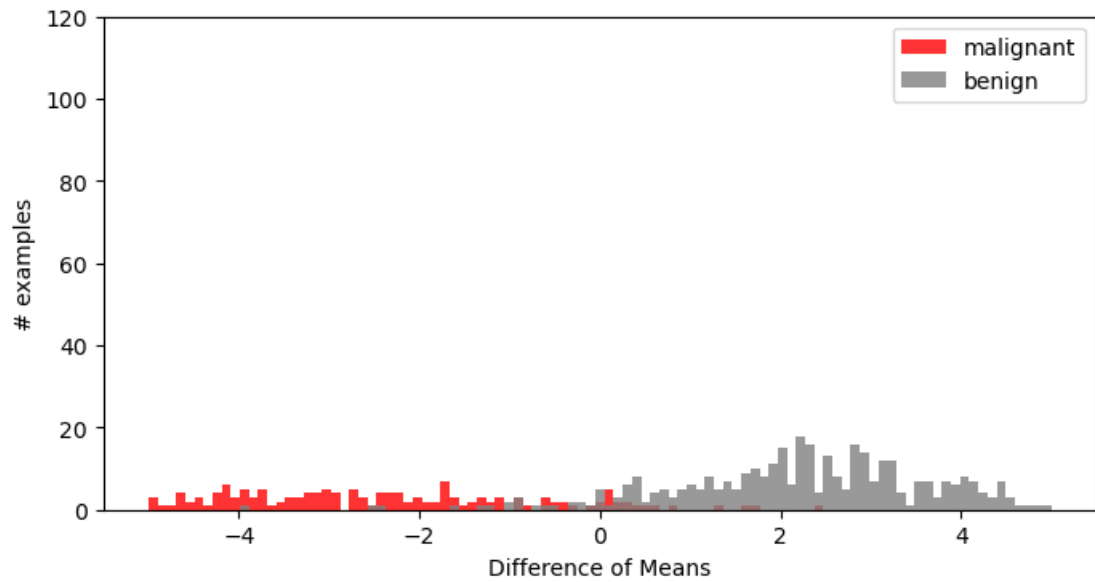
(b) Plot the histogram of projected data for several projection vectors `w`, in particular

1. `w` is the difference between the mean vectors of the two classes
2. `w` is the Fisher discriminant, and
3. `w` is the weight vector of a linear SVM (with hyperparameter $C=10.0$).

For the SVM, you can make use of the sklearn implementation (`sklearn.svm.SVC`). Before applying the projection and plotting, ensure that `w` is rescaled to have norm 1 in each case. When plotting, use fixed axis ranges and bin sizes in order to ease comparison.

```
In [3]: %matplotlib inline

## -----
## TODO: replace by your code
## -----
import solutions
solutions.analysis(X, XM, XB, T)
## -----
```



We observe that different discriminants extract different aspects of the data. For example, the difference of means algorithm produces the highest average distance between classes. The Fisher vector produces rather small distance between classes but also very low variance within classes. Lastly, the SVM focuses on achieving a low overlap between the two classes through learning a separating hyperplane.