

---

# *Final project*

---

**In this project there are 4 notebooks, I'm going to write a small description about each notebook :**

**1. Passengers flight satisfaction:** This notebook deals with classifying the passengers with their satisfaction (satisfied/dissatisfied).

In this notebook we were supposed to try improving the accuracy score from last semester, but I have already used models that was learned this semester, so I kept it as it is using KNeighbors, Random Forest, Decision Tree, Logistic Regression classifiers, in addition I also used Grid Search and Voting classifier.

In other words, I stayed with 0.94 accuracy score.

**2. Fashion Mnist :** This notebook contains a dataset of images consisting of a training set of 60,000 examples, each example is a 28x28 grayscale image so in total for each image there are 784 pixels each pixel value from 0 to 255, associated with a label from 10 classes, T-shirt/Top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, Ankle boot.

In this notebook we train our models to get the best results indicating the right piece of clothes.

I used the same classifiers I used in the first notebook "Passengers flight satisfaction", KNeighbors, Decision Tree, Logistic Regression, Random Forest, Voting classifiers and also grid search.

The best results were using Random Forest classifier, giving us 0.88 accuracy score.

But in this notebook I tried applying PCA for dimensionality reduction, but it reduced the accuracy score for 0.86.

**3. Dogs vs Cats :** Here we got a file that contains 25000 different images of cats and dogs, half of them are dog images and the other half is cat images.

We had to train a model to classify the images to cats or dogs, so in order to do that, first we have to compress the images to the same size , so I chose 28x28 and that will give me a 784 pixels for each image, and then I converted the images from RGB to grayscale images.

I trained the model using XGBoost classifier, but I didn't get great results, I got an accuracy score of 0.53.

Then I tried applying PCA for dimensionality reduction but it didn't change anything, I was kept with the same accuracy score 0.53.

**4. Hand movement :** In this notebook we were given a few csv files for each person who participated in article, filled with hand positions.

We were supposed to classify the hand position, for 3 different situations, Alone, Sync, Spontan.

First, I combined all the csv files into one big dataframe, then I dealt with each situation alone, and after that I combined them again into one dataframe.

I trained test split the dataframe, I used Logistic Regression, Random Forest, XGBoost, KNeighbors classifiers, I got great results for each classifier, but surprisingly I got a 1.0 accuracy score using XGBoost classifier !