

Final Project – semester B

Renana Rimon

In this project, the main methods I have used is:

1. **Scale**: standardize the dataset by standard-scaler.
2. **PCA**: dimensionality reduction.
my goal was to reach the best accuracy with lowest number of components, so I have tried to get the best balance between both.
3. **GridSearch**: find best parameters in each model (using **cross-validation**).
4. **pipeline**: scaling, pca, and fit model.
5. **Ensamble learning models**:
 - a. Logistic regression
 - b. KNN
 - c. Random forest
 - d. XGBoost
 - e. Gradient Boost
 - f. Voting
 - g. Stacking

Models I used but didn't perform well, so isn't in the notebooks:

- h. Ada Boost
- i. K-means