



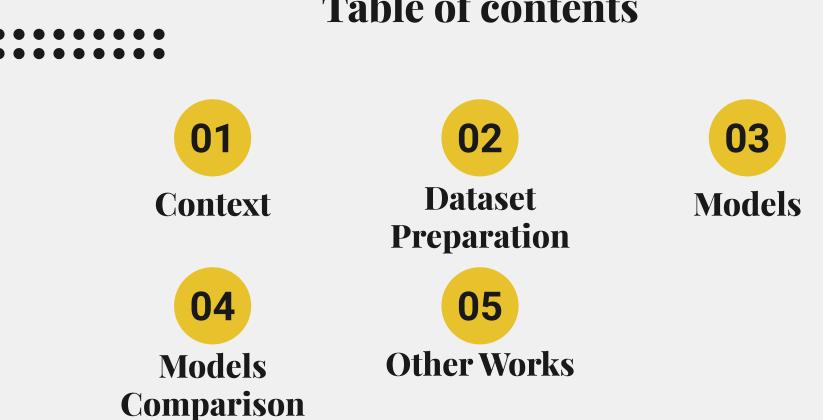


Fundamentos de Aprendizagem Automática

Course Instructor: Pétia Georgieva 2022/2023

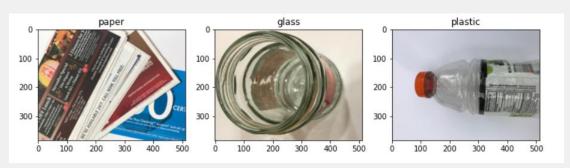
Pedro Monteiro 97484 Eduardo Fernandes 98512

Table of contents



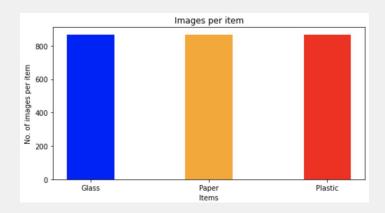
Context

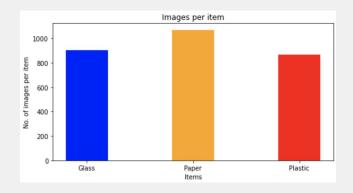
- Why?
- Dataset available on Kaggle
- Different machine learning models implemented:
 - Logistic Regression
 - o SVM
 - Decision Tree
 - Random Forest
 - Neural Networks



Dataset Preparation

- Unbalanced Dataset
 - O Glass 900 images
 - o Paper 1070 images
 - Plastic 868 images

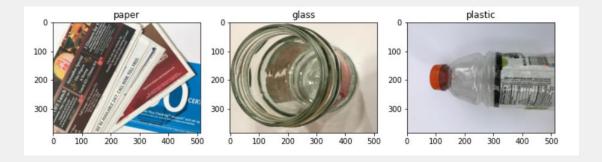




- Balanced Dataset
 - Glass 868 images
 - Paper 868 images
 - O Plastic 868 images

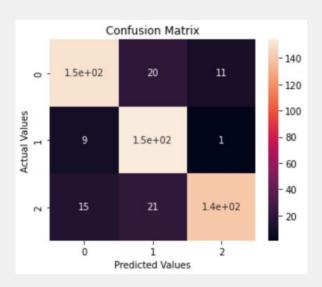
Dataset Preparation

- Image size 512*384
- Large number of features
- Final: 32*32
- Pixel normalization



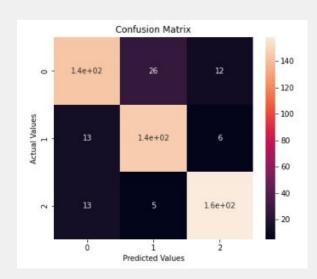
Logistic Regression

	Base	K-fold CV	HyperTuned
Accuracy	84.64%	84.64%	85.60%
F1 Score	84.67%	84.67%	85.62%



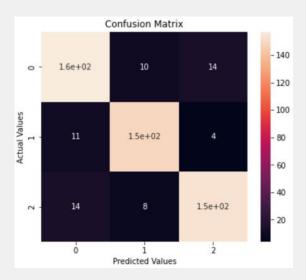
SVM

	Base	K-fold CV	Hypertuned
Accuracy	85.22%	85.22%	90.02%
F1 Score	85.20%	85.20%	90.07%



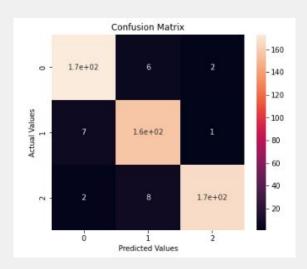
Decision Tree

	Base	K-fold CV	Hypertuned
Accuracy	88.68%	89.25%	86.96%
F1 Score	88.69%	89.25%	86.95%



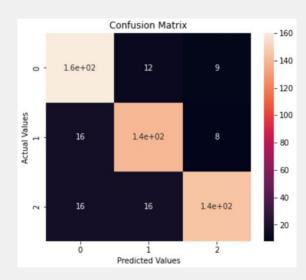
Random Forest

	Base	K-fold CV	Hypertuned
Accuracy	95.01%	93.47%	96.16%
F1 Score	95.03%	93.48%	96.16%



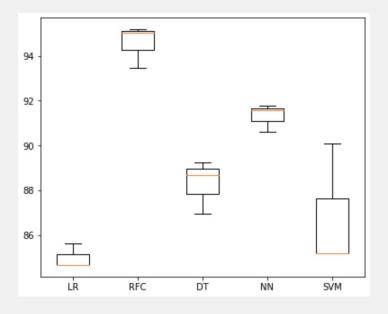
Neural Networks

	Base	K-fold CV	Hypertuned
Accuracy	91.75%	90.59%	91.55%
F1 Score	91.77%	90.63%	91.57%



Models Comparison







Other Works

- Random Forests: An algorithm for image classification, Ned Horning
- Image Classification using Random Forests and Ferns, Anna Bosch, Andrew Zisserman and Xavier Munoz
- Image Classification and Recognition Based on Deep Learning and Random Forest Algorithm, Erhui Xi
- Image recognition using Machine learning, Abhinav N Patil
- Fine-Tuning Models Comparisons on Garbage Classification for Recyclability, Umut Özkaya and Levent Seyfi

Thanks!

