Final Project

Activities Recognition

Javier Pedrosa Alias
Óscar Palacín Domínguez

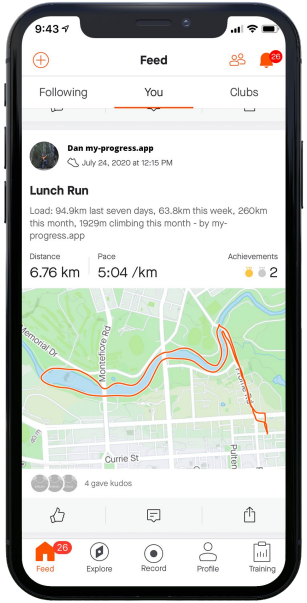
PRML
MUAR - UPC
2020 - 2021

AGENDA

- Introduction
- Strategy
- First Results
- Conclusions and Future Work

INTRODUCTION

- STRAVA



- Activities recognition:
  - Walking, base running, trail running, road cycling, mountain biking, traveling by subway, traveling by bus, traveling by car.
  - Walking, running, cycling, motor vehicle.

STRATEGY

- Transforms TCX files into CSV.
- Data Handling (pandas).
- Dimensionality Reduction (PCA and Correlation).
- Split data into train (70%) and test (30%) to avoid overfitting.
- Supervised Techniques and Unsupervised Techniques.
- Prediction of test data.
- Performance Evaluation.

EXPERIMENTATION

- Handling of NaNs using average values for each activity.
- Dimensionality reduction:
  - PCA: from 32 to 8 features with accumulative explained variance > 95%.
  - Correlation: from 32 to 10 features (maximum altitude, maximum heart rate, maximum default speed, minimum default speed, minimum run cadence, minimum speed computed by the GPS data, minimum acceleration computed by the GPS data, mean acceleration computed by the GPS data, 2nd quantile for the acceleration computed by the GPS data and 3rd quartile for the acceleration computed by the GPS data).

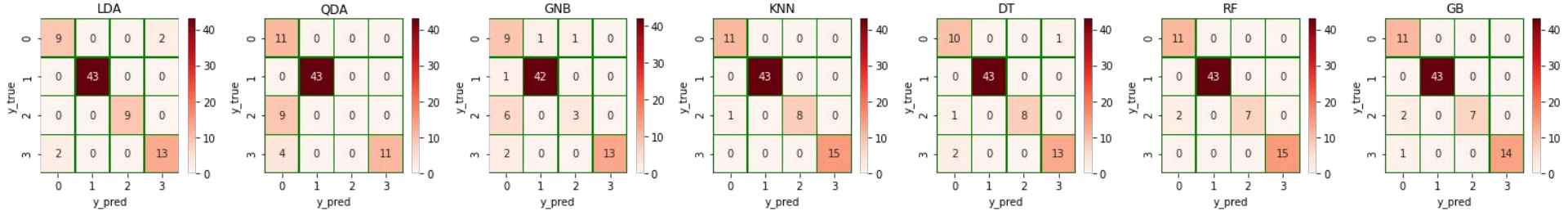
EXPERIMENTATION

- Implementation of supervised techniques:
  - Linear Discriminant Analysis (LDA).
  - Quadratic Discriminant Analysis (QDA).
  - Nayve Bayes (NB).
  - Decision trees (DT).
  - Random Forests (FR).
  - Gradient Boosting (BR)
- Implementation of unsupervised techniques:
  - K-Nearest Neighbours (KNN).

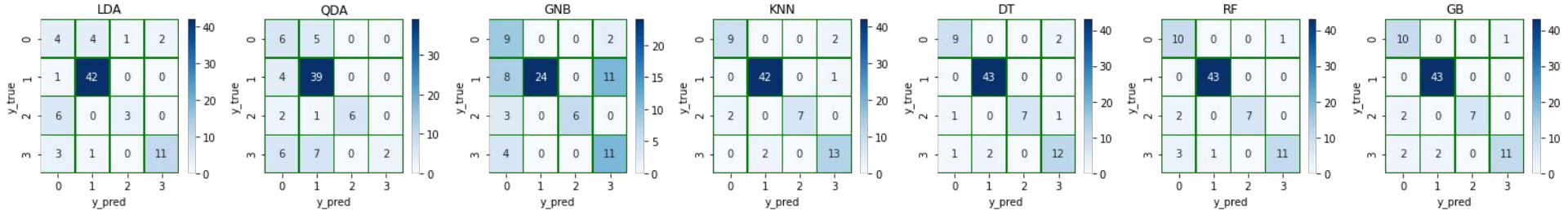
RESULTS

Dim. Reduction	Classifiers						
	LDA	*QDA*	*GNB*	*KNN*	*DT*	*RF*	*GB*
PCA	92%	61%	81%	98%	92%	96%	95%
Correlation	69%	77%	71%	90%	89%	89%	90%

- Performance evaluation:
  - PCA



  - Correlation



CONCLUSIONS

- Best performance dimensionality reduction obtained for PCA.
- Best classification obtained for the following classifiers:
  - KNN (98 % accuracy).
  - RF (96 % accuracy).
  - GB (95 % accuracy).
  - DT (92 % accuracy).

Thank you for your attention!
Any question?