TÉCNICO LISBOA

DS 2025

Data Science Project

Team nr:	06	**Student 1:**	Maria João Pupo Correia	**IST nr:** 86476
		Student 2:	Ricardo Andrade Silva	**IST nr:** 100071
		Student 3:	Leonardo Pereira	**IST nr:** 103215
		Student 4:	Afonso Carvalho	**IST nr:** 116482

LAB 01

1 Dataset 1 - Traffic Accidents

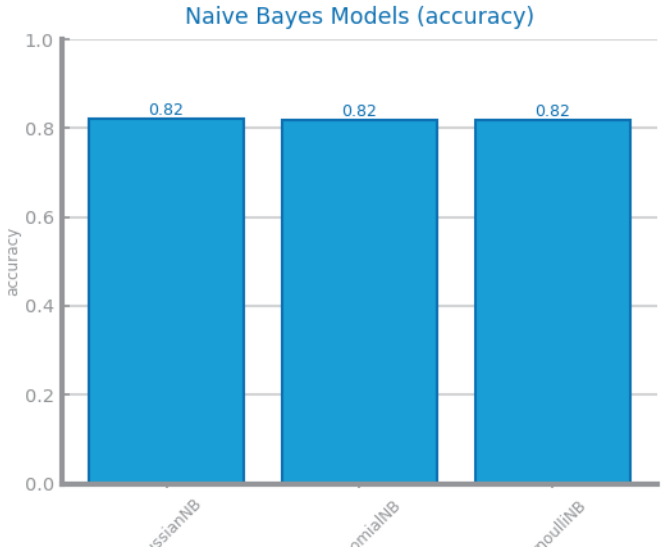

Figure 1: Study of Naive-Bayes alternatives comparing accuracy.

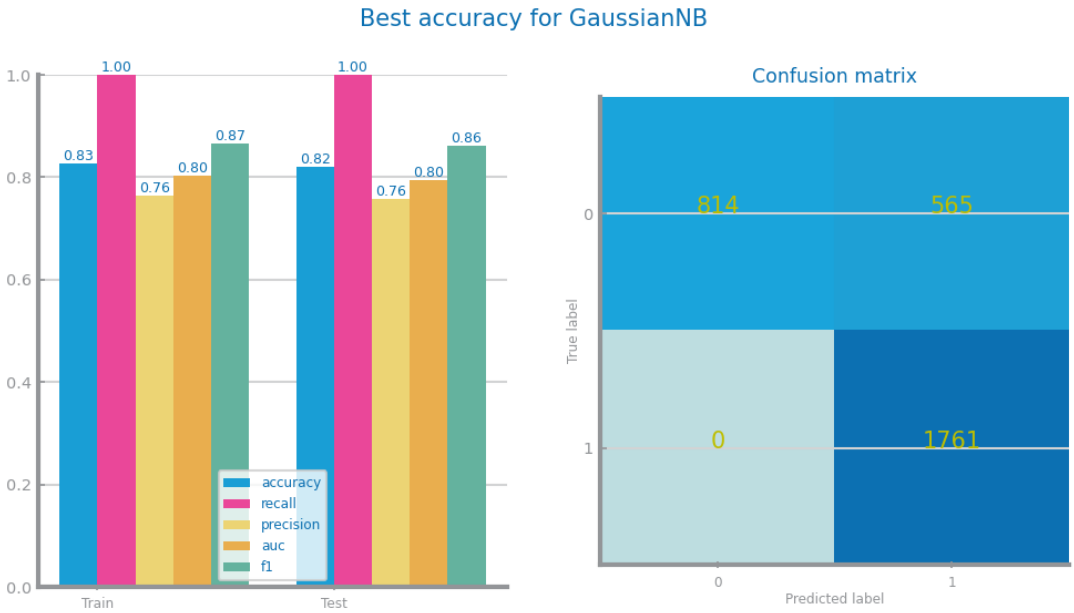

Figure 2: Performance analyses for the Naive-Bayes model with the greatest accuracy.

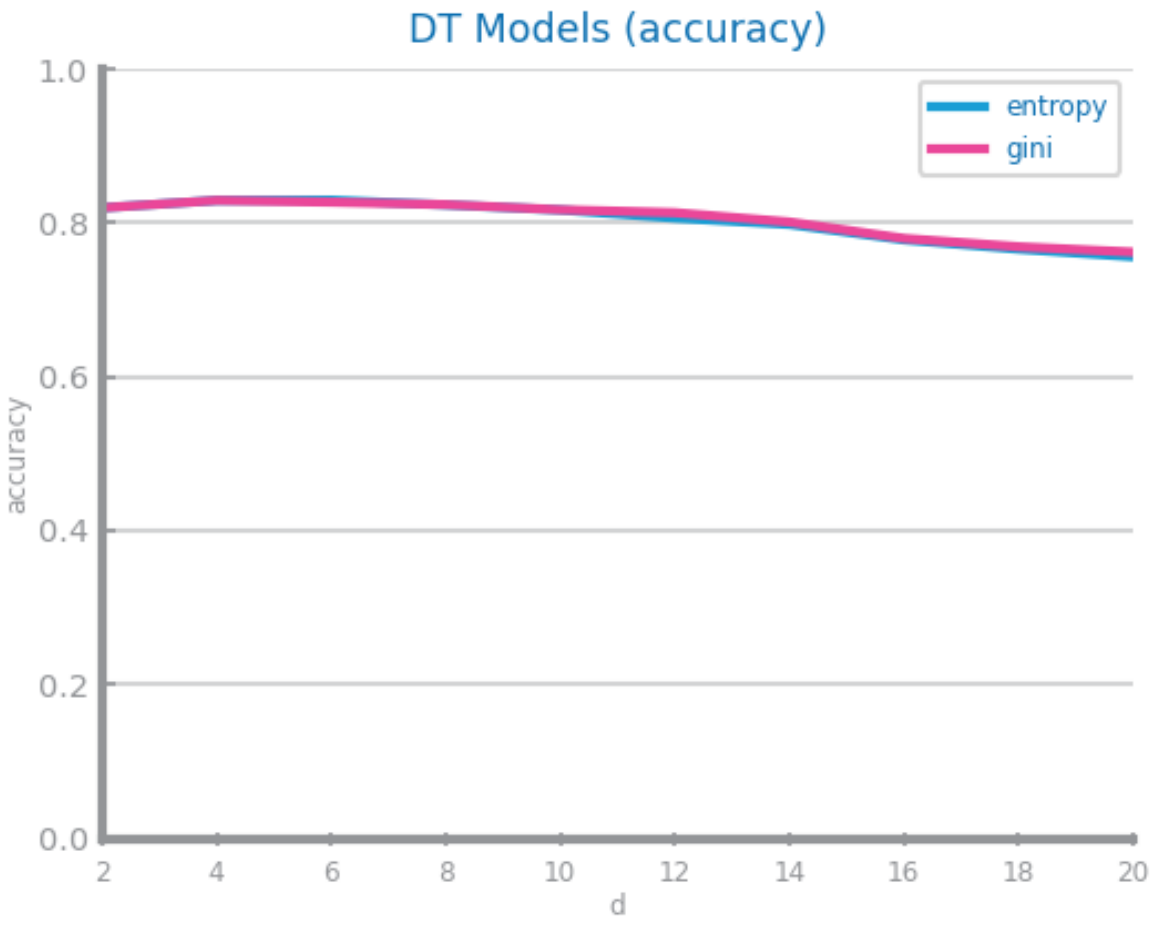

Figure 3: Study of Decision Trees alternatives comparing accuracy.

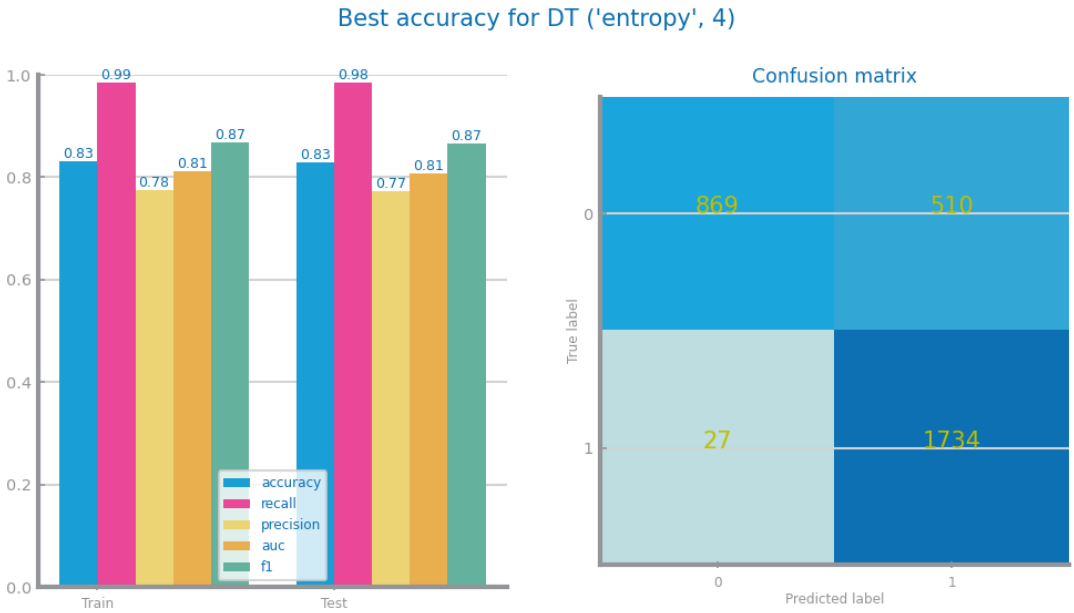

Figure 4: Performance analyses for Decision Trees model with the greatest accuracy.

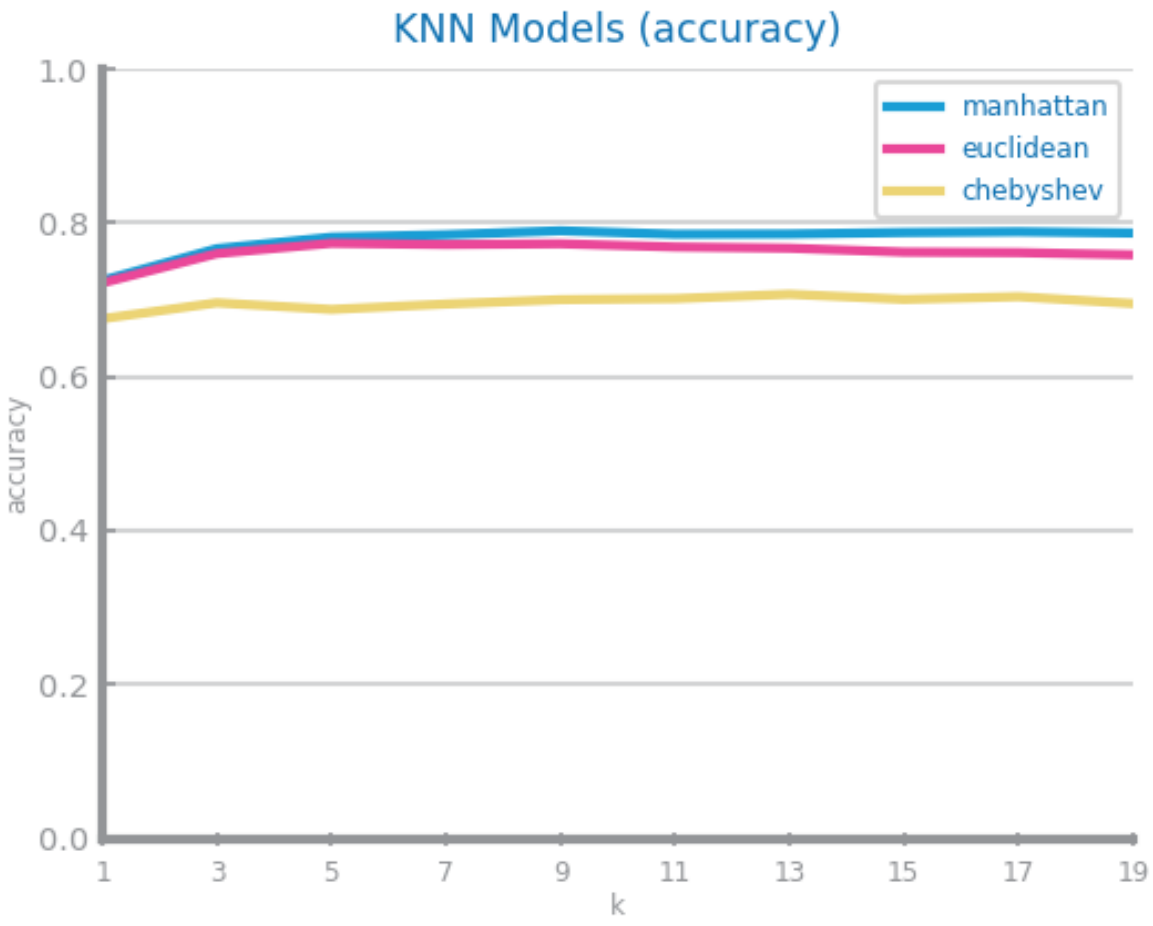

Figure 5: Study of KNN alternatives comparing accuracy.

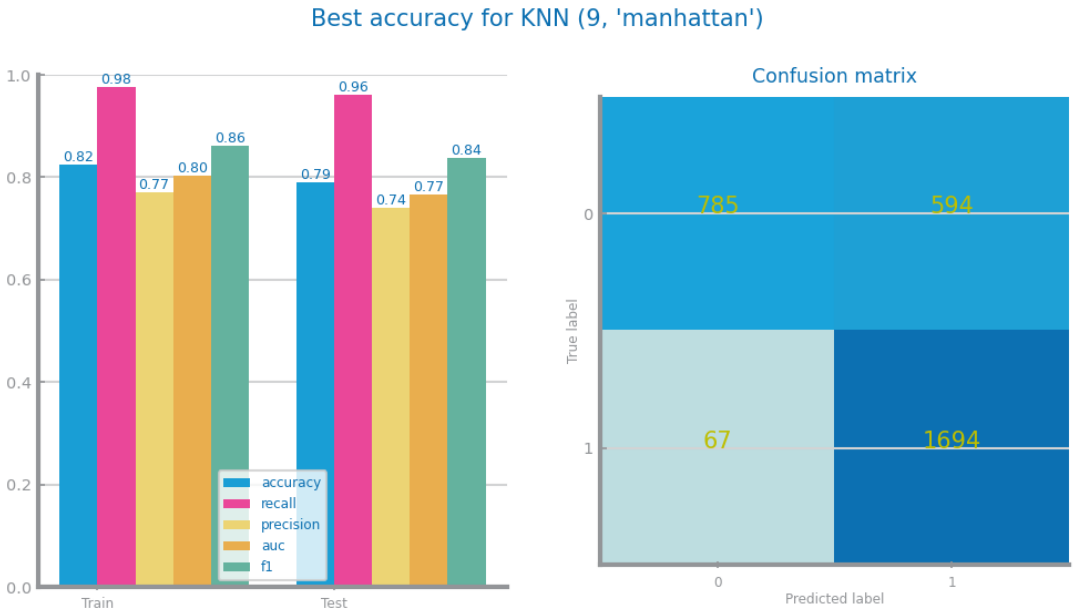

Figure 6: Performance analyses for KNN model with the greatest accuracy.

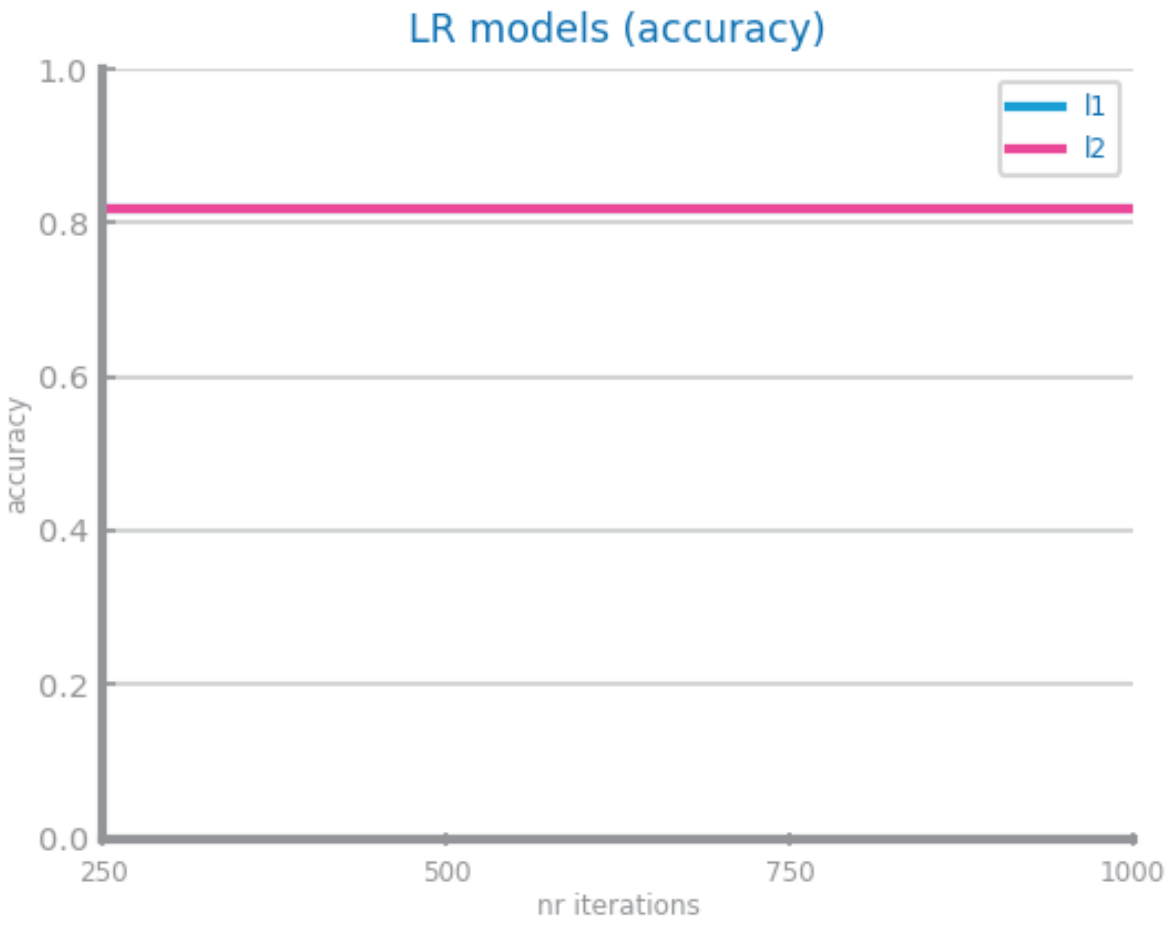

Figure 7: Study of Logistic Regression alternatives comparing accuracy.

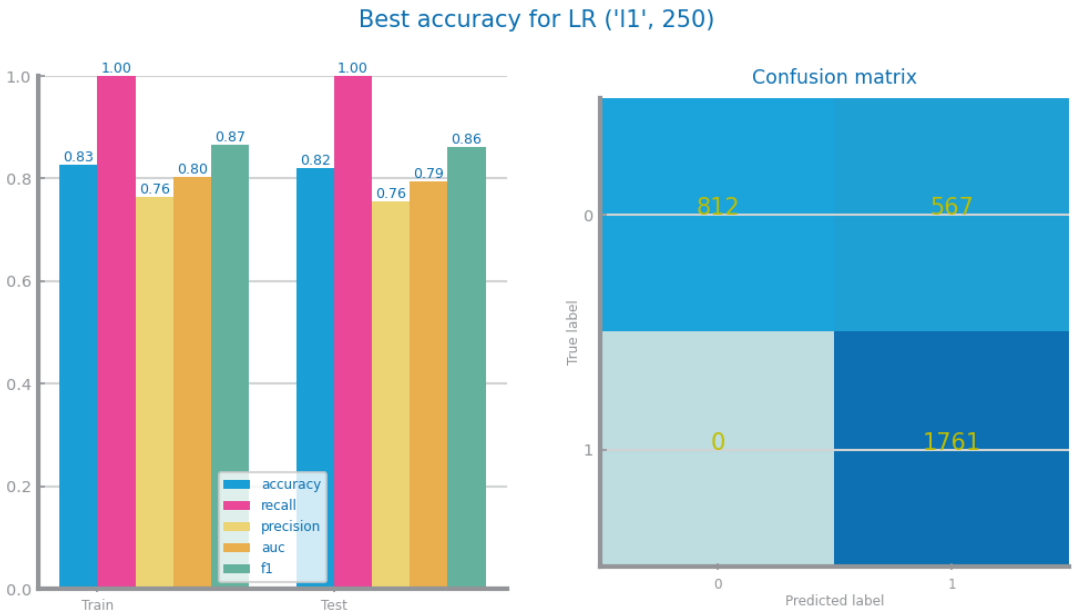

Figure 8: Performance analyses for the Logistic Regression model with the greatest accuracy.

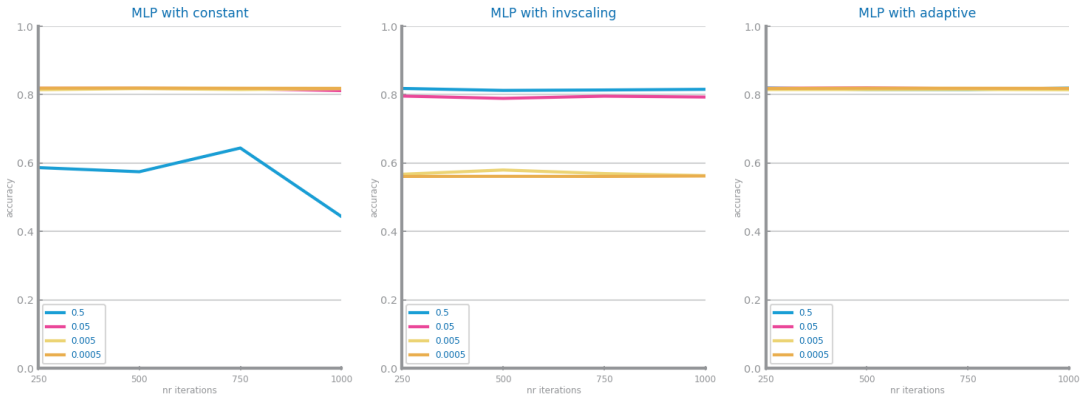

Figure 9: Study of Multi-layered Perceptron alternatives comparing accuracy.

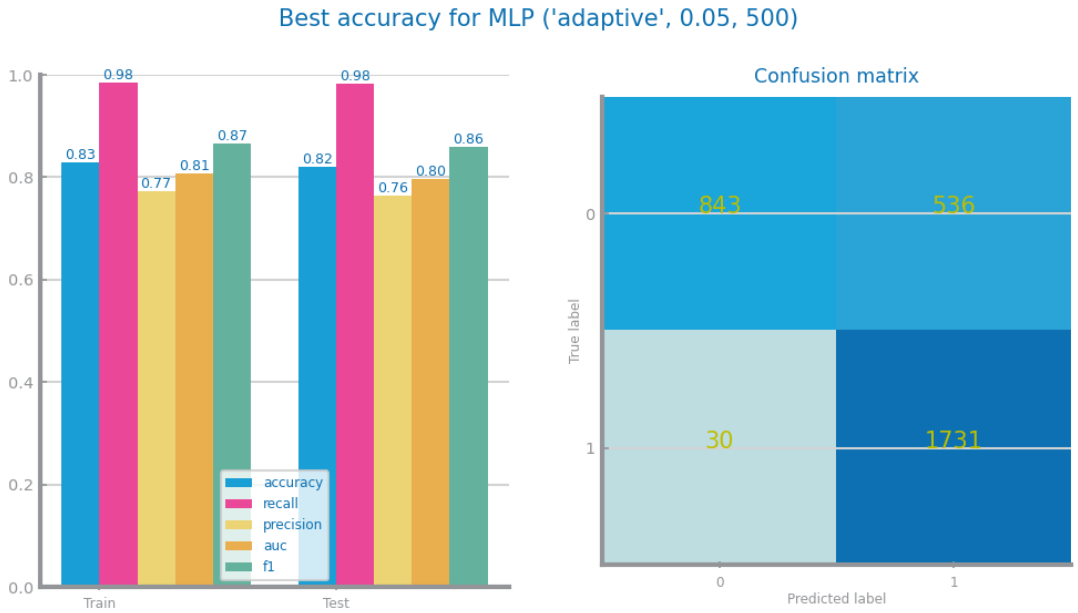

Figure 10: Performance analyses for the Multi-layered Perceptron model with the greatest accuracy.

2 Dataset 2 - Combined Flights

By dropping either the missing values: either the variables or the rows, it is not possible to train models because there are no positive examples. All the canceled flights were dropped since at least of of its variables was empty.