



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH



XNDL

Lab3 - Preprocessing

Preprocessing tasks

- ❖ **Describe the effect on the coefficients of scaling a variable, and show an experiment of it**
- ❖ Find and understand variable correlations, contextualize with coefficients
- ❖ Discuss the best model given the roc curve, considering it's domain
- ❖ Train a DT and compare it with the logistic coefficients

String vars

- ❖ Find out what is the vectorizer doing to the data
- ❖ Change to TF-IDF, understand their differences and test both
- ❖ Discuss the best model given the roc curve, considering it's domain



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH



Closure

The end