Packages:

Pandas

Numpy

Matplotlib for visualising tool package

Statsmode: for stat analytics package

For ML:

Keras

# Second class:

Fadn.shape

Look over the length

Clean the data :

1. Extract the data
2. Delete the part of array . fadn.del[]

For cont and discrete variables:

Find what to look for these datsets like: accuracy, accreditation,

Important topic

Model performance evaluation

PSI:

If we don’t know PSI also if we know what variables makes changes, we could make model perform better, by updating model to make new structural index.

Although PSI is not the only technique for detecting model drift, it is a popular metric for measuring a model’s stability for its simplicity, ease of interpretation, and applicability to both categorical and continuous variables. Calculating PSI is quick, requires minimal computational resources, and is easily understood by non-experts. Additionally, PSI works with various data types, making it adaptable across many machine learning applications.

PSI evaluates model stability by comparing category distributions in development and validation populations. Understanding PSI’s mathematical essence empowers analysts to effectively evaluate and maintain their models’ performance over time.

Model evaluation in discrete variables:

Accuracy concept:

Ways to see accuracy:

Kappa : rating kind of data, good , bad type concept