

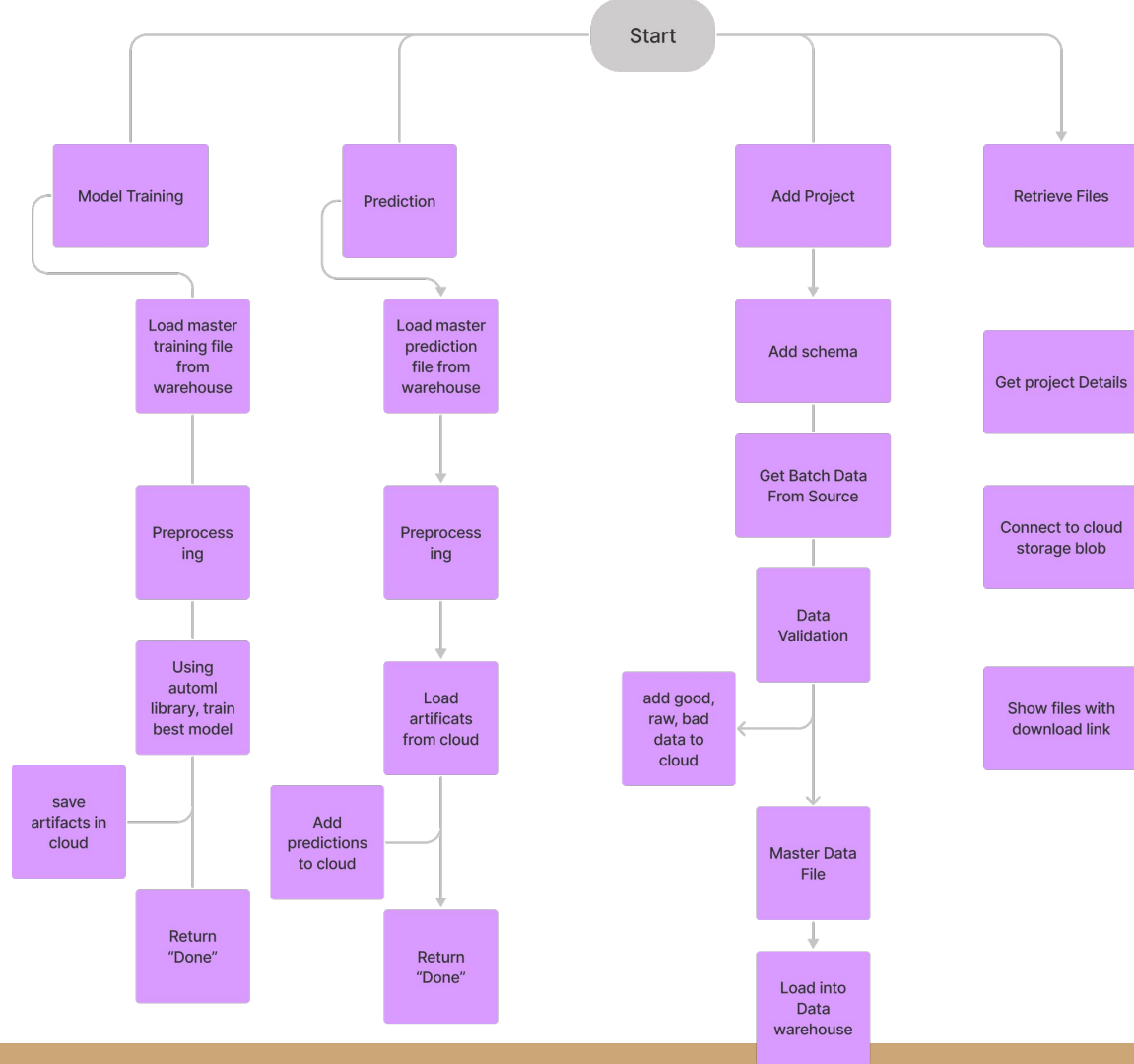


Automated Machine Learning

By Atufa Shireen



Web Architecture



Validation Process

Convert To HDFS

File name validation

Column name, Number of columns
validation

validate missing columns based on data

creating master data

Loading in warehouse

storing good,
bad and
raw data in
folders

Preprocessing Steps

Remove Duplicates

Missing Imputations

Remove Zero variance cols

Remove perfect colinearity

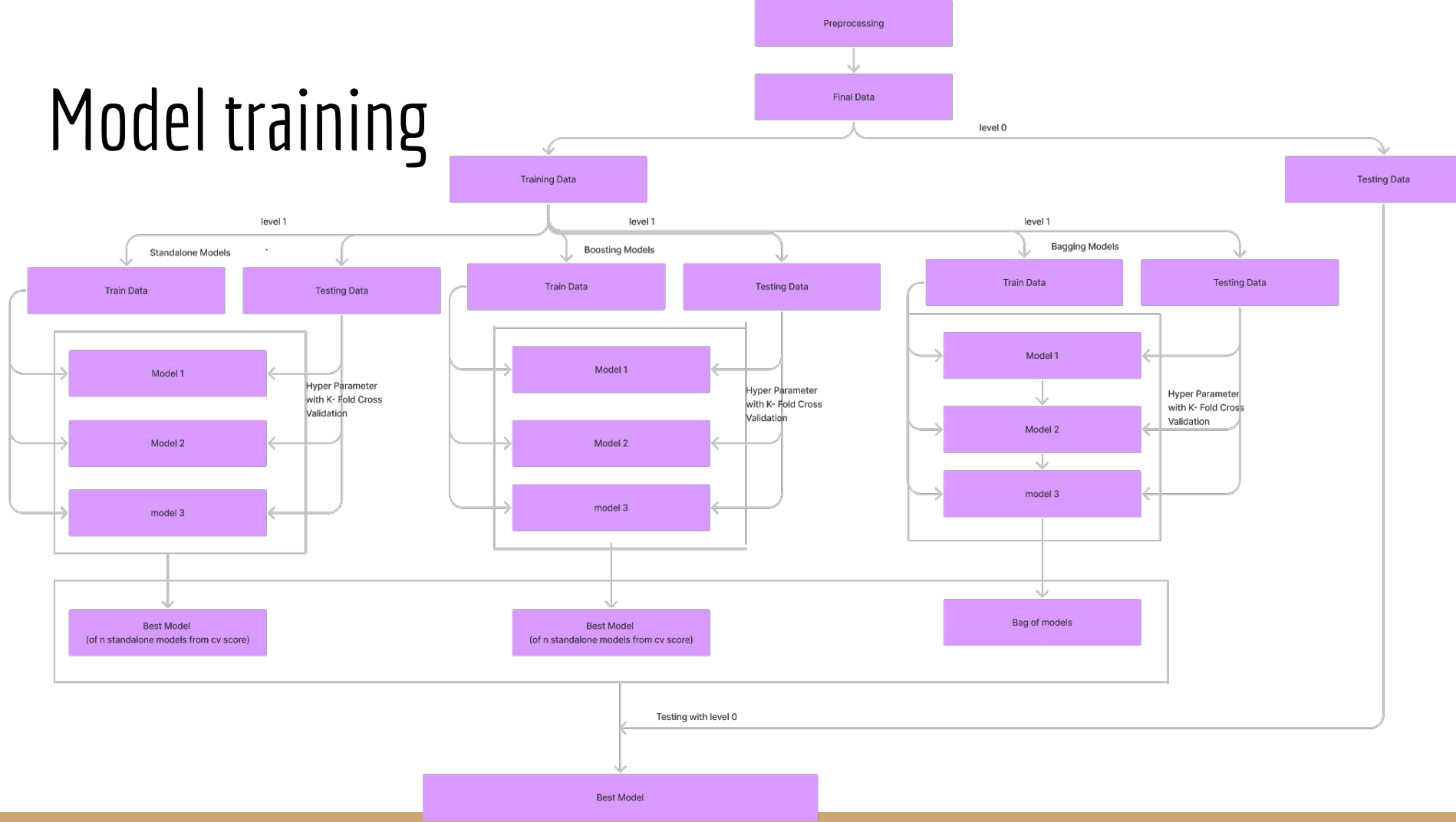
OHE

Data Transformation

Fitting and Adding cluster label

Label Encode Target
(classification)

Model training



Thank You