

SDAIA ACADEMY

T5 Boot Camp

LINEAR REGRESSION PRESENTATION

Supervised by: Mr. Mohamed El-Deeb

Linear Regression module

27 September 2021

AGENDA

- Goal
- Data Collection
- EDA
- Data Cleaning
- Pre-Processing
 - Feature preparation
 - Feature Engineering
 - Feature Selection
- Modeling
- Models Evaluation

GOAL

- Predicting the Discount Percentage.

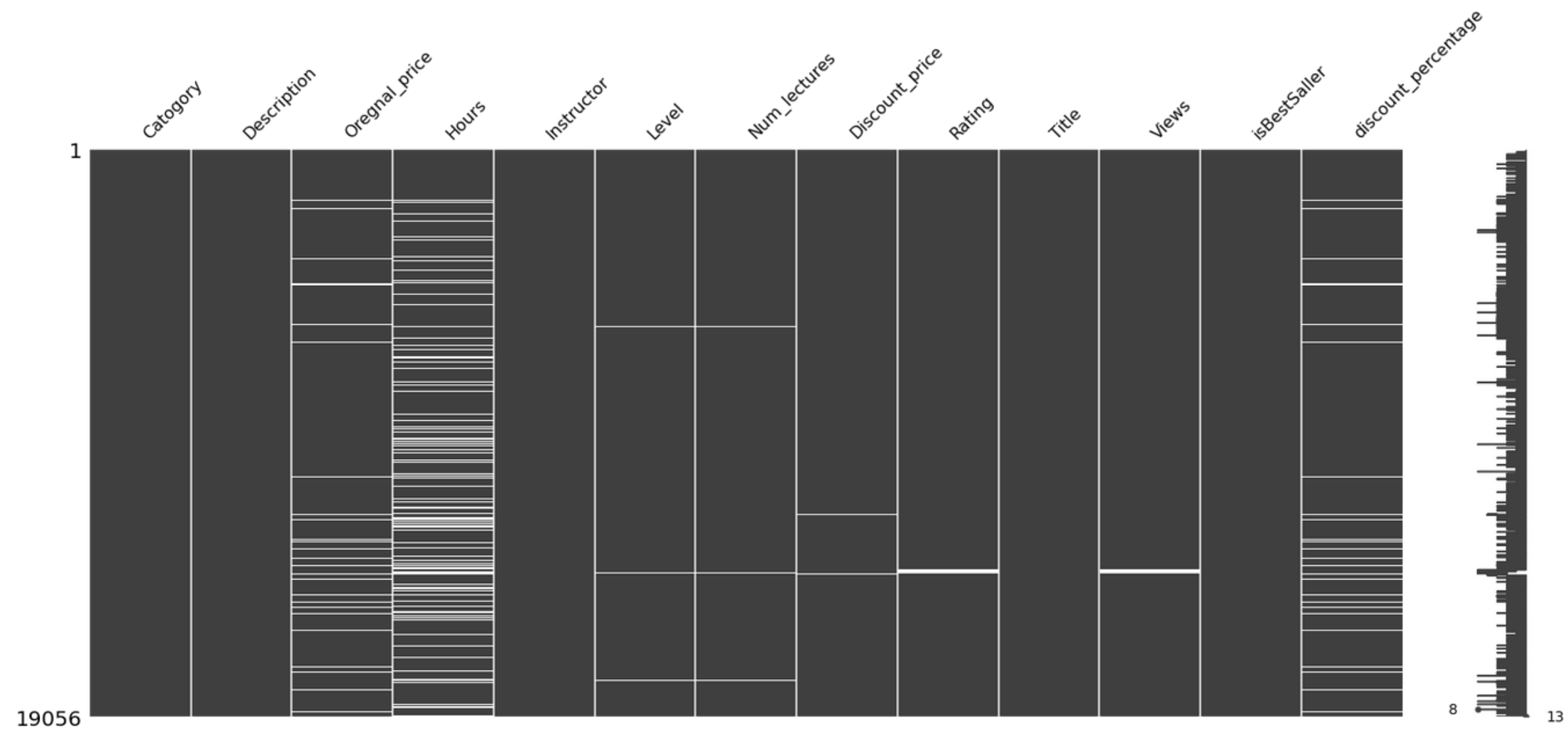
DATA COLLECTION

- Web Scrubbing "Udemy"
- 12 Features.
- Almost 20K Record.

EDA

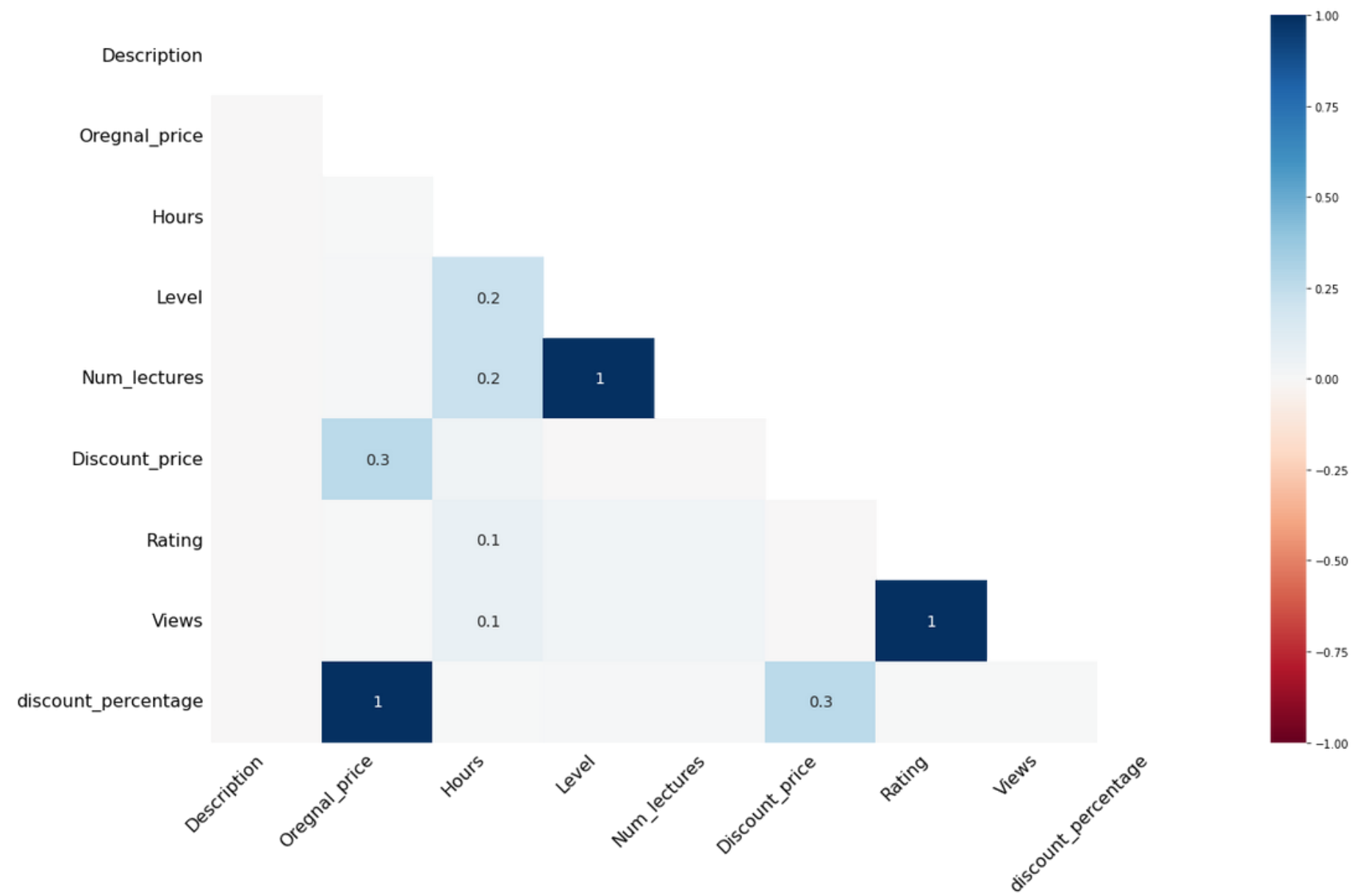
05

- Null Values



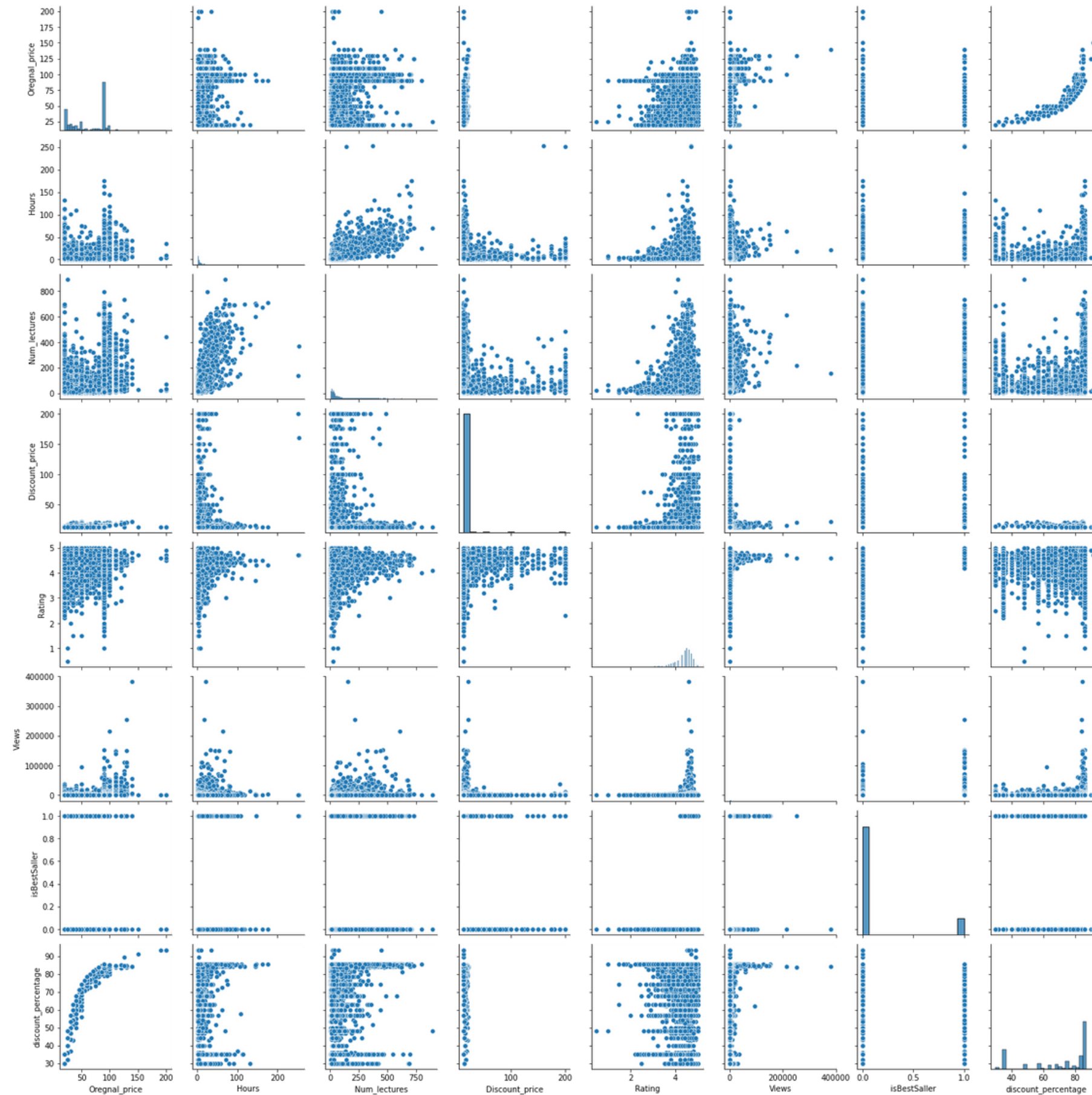
EDA

- Corr



EDA

- Corr



DATA CLEANING

Null
Handling

⋮

Drop
4K

Outlier
Handling

⋮

no Outlier

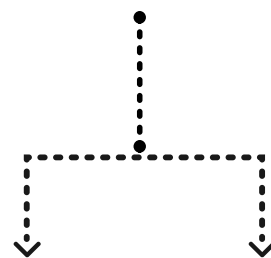
PRE-PROCESSING

Feature
preparation



Categorical
into Dummy

Feature
Engineering



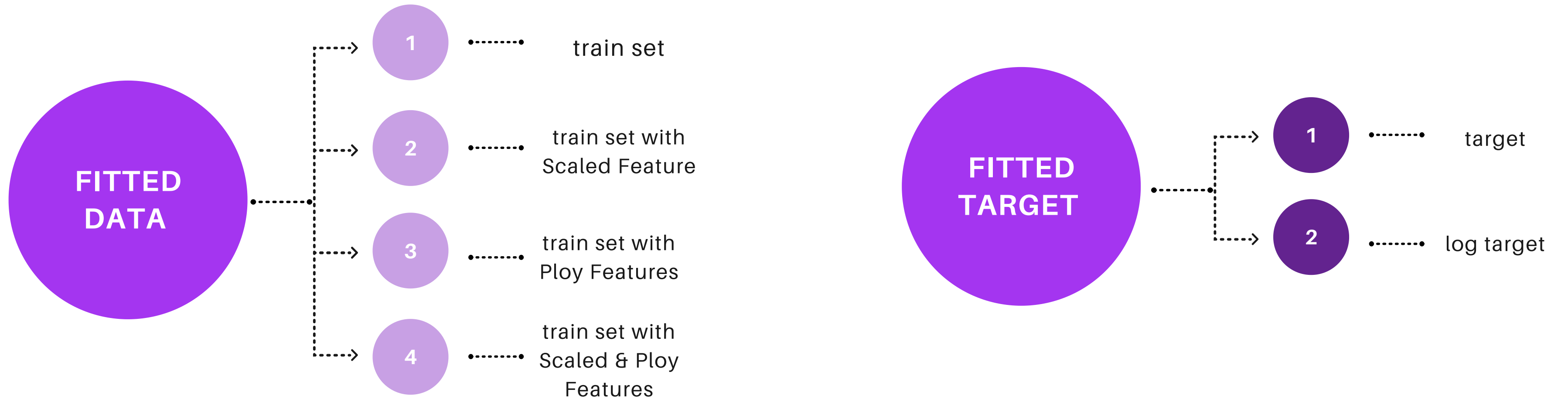
Scale Ploy

Feature
Selection



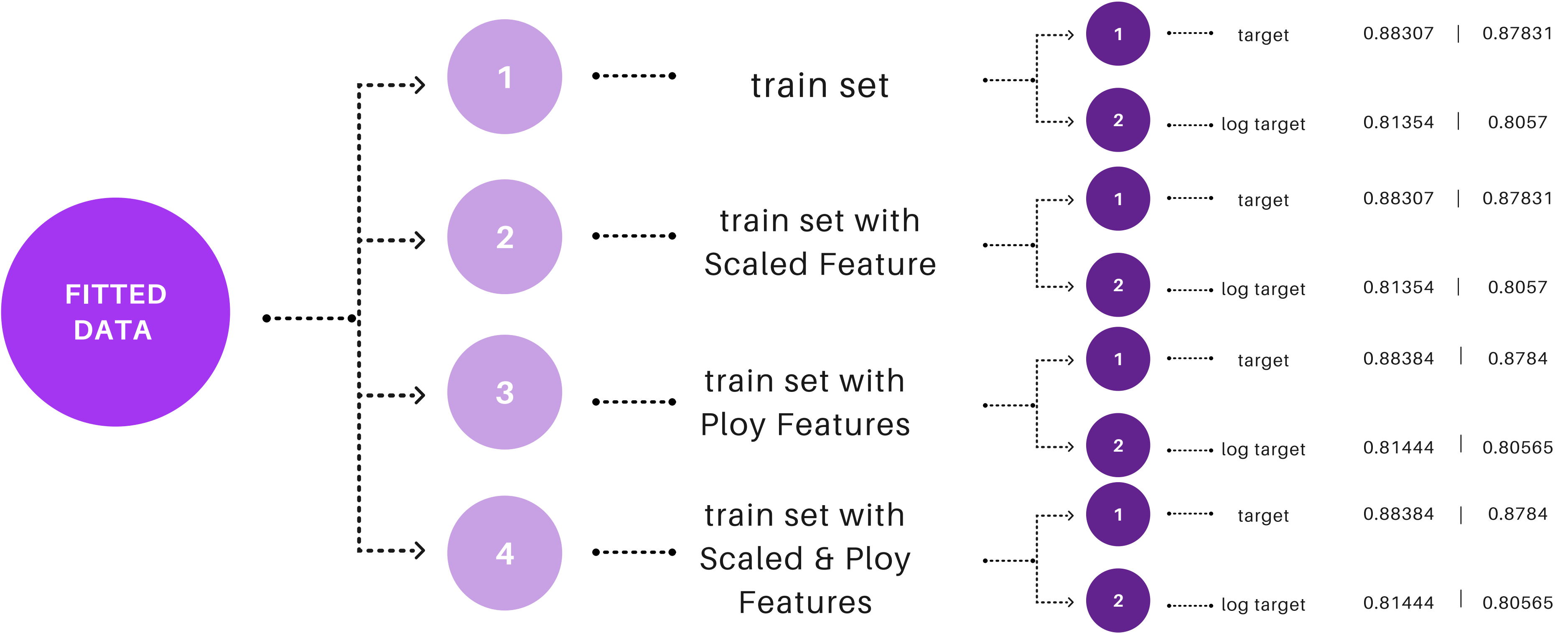
LASSO

MODELING



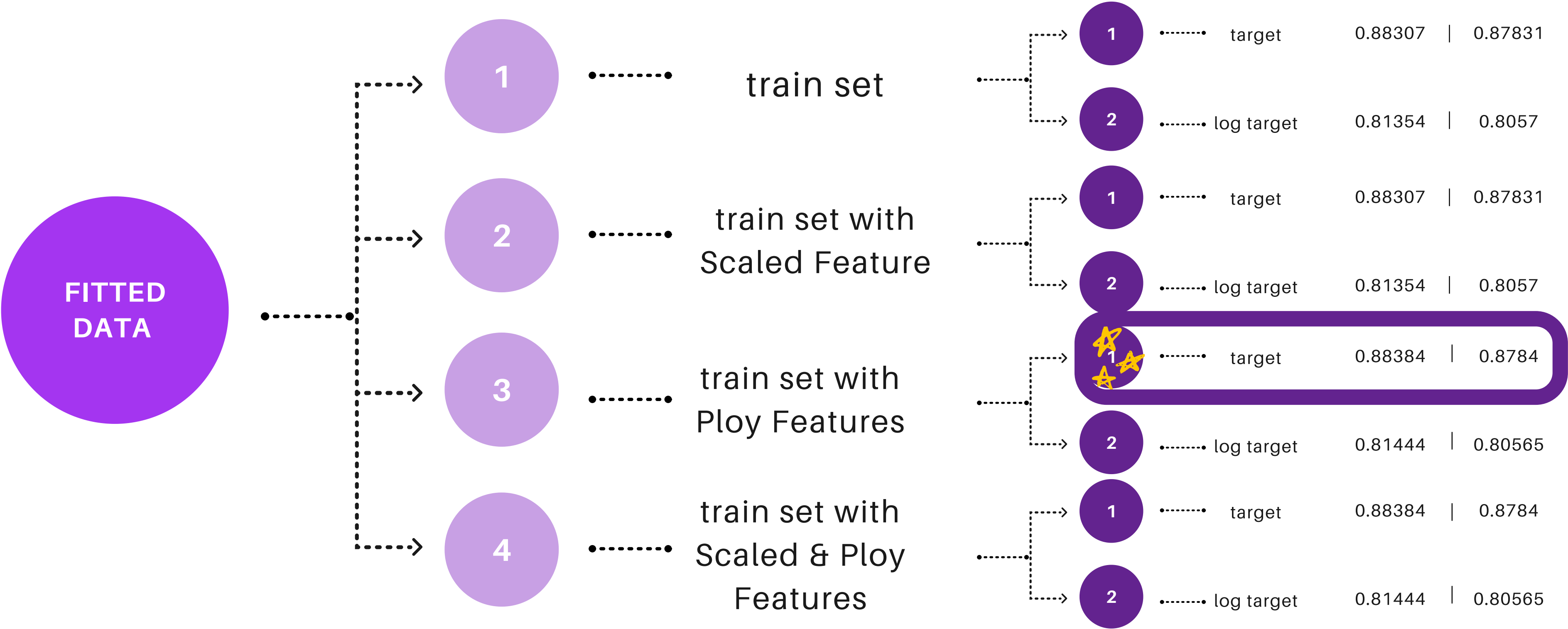
MODELS EVALUATION

Score	
train	validation

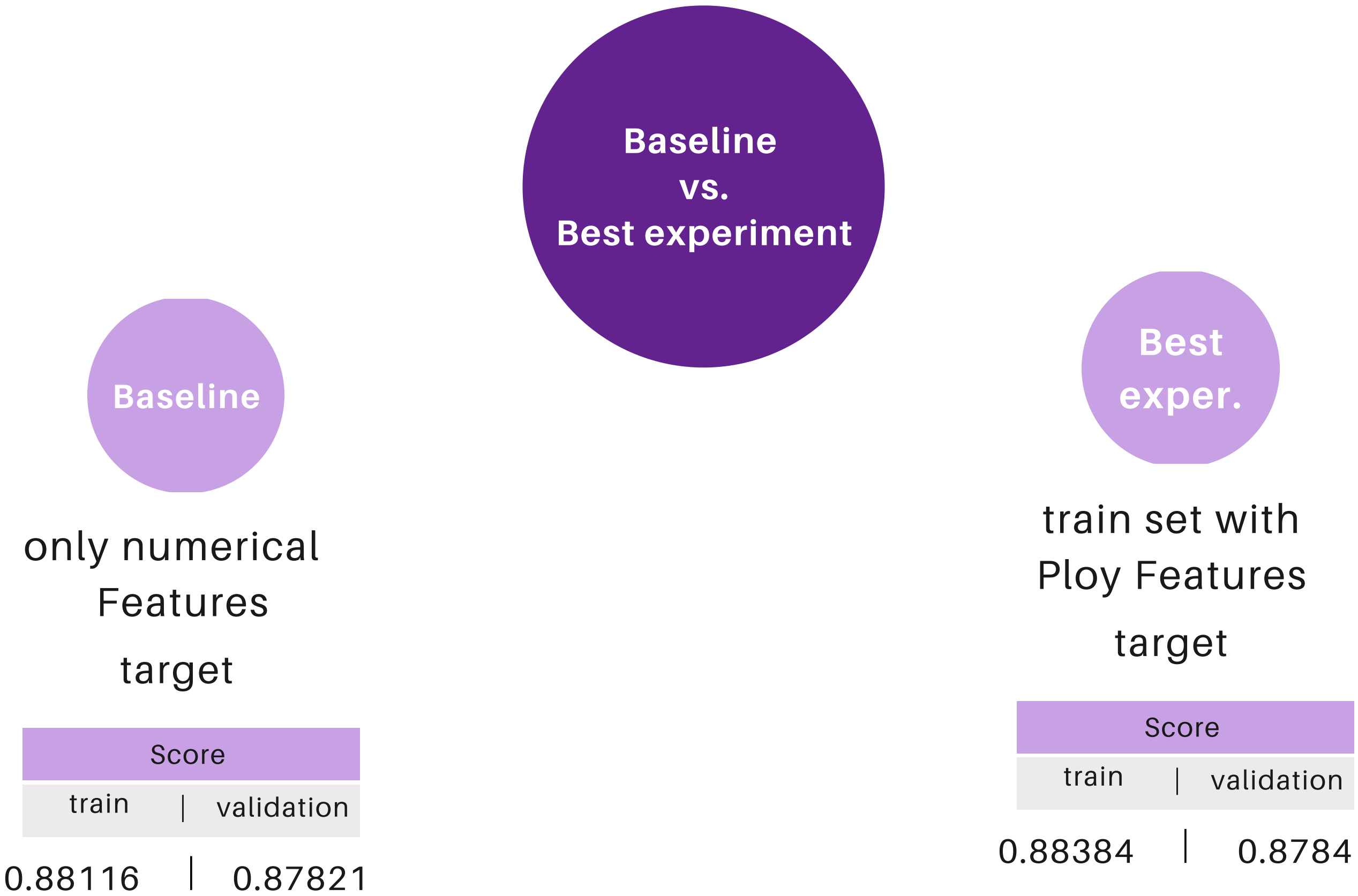


MODELS EVALUATION

Score	
train	validation



MODELS EVALUATION





Udemy

THANK YOU FOR
LISTENING

27 September 2021