

Day-6

Pipeline and column transformer

What is a Column Transformer?

A **Column Transformer** is a tool in machine learning that allows you to apply different preprocessing steps to different columns of your dataset. It is especially useful when dealing with datasets that have both numerical and categorical features, which require different types of preprocessing.

Simple Example:

Imagine you have a dataset with the following features:

- **Age** (numerical)
- **City** (categorical)
- **Income** (numerical)

You want to:

- Standardize the **Age** and **Income** features.
- One-hot encode the **City** feature.

What is a Pipeline?

- **Pipeline** chains together multiple steps and output of each step is input for next step.
- Pipeline make it easy to apply multiple function to columns.