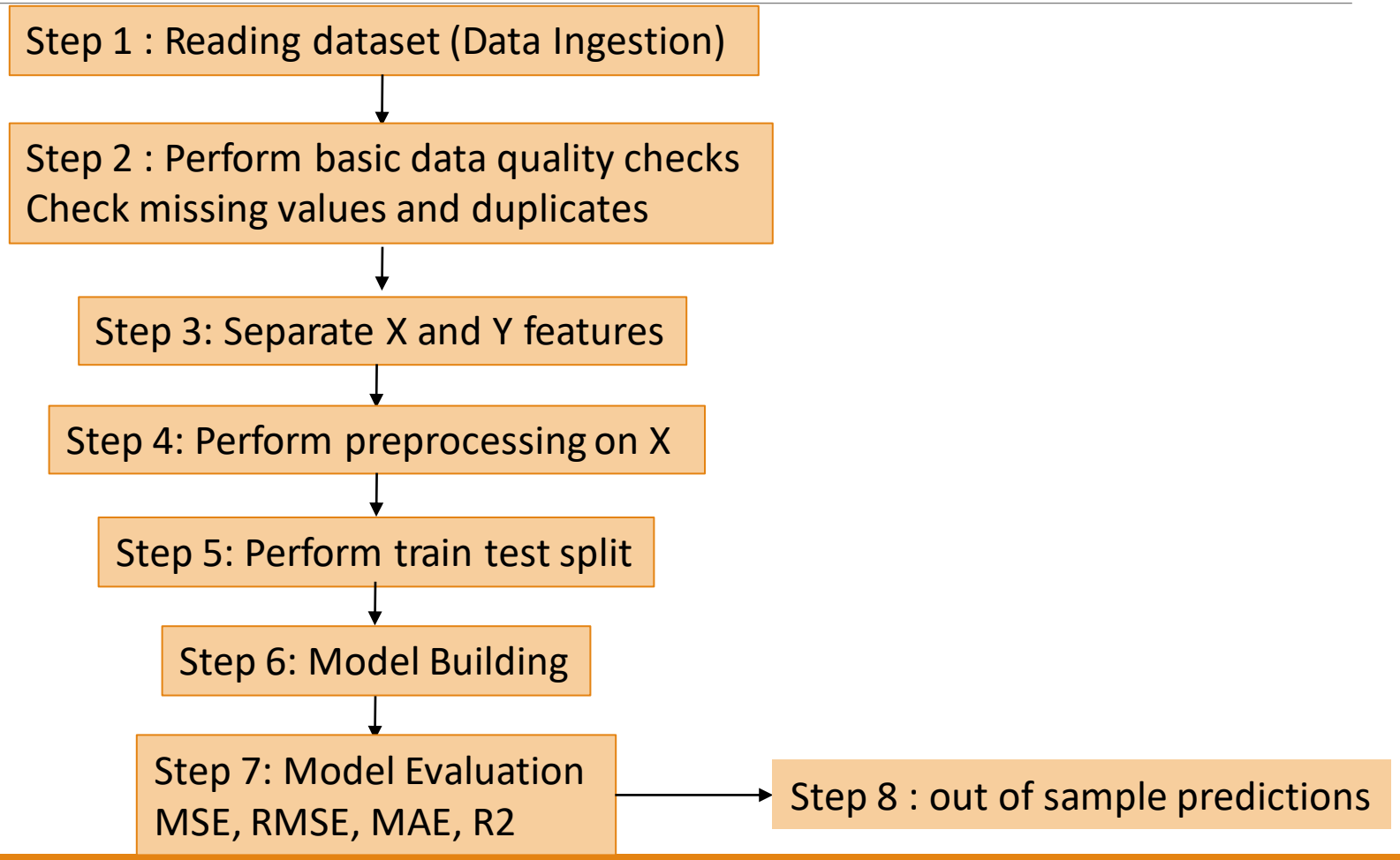
Preprocessing with Sklearn pipeline

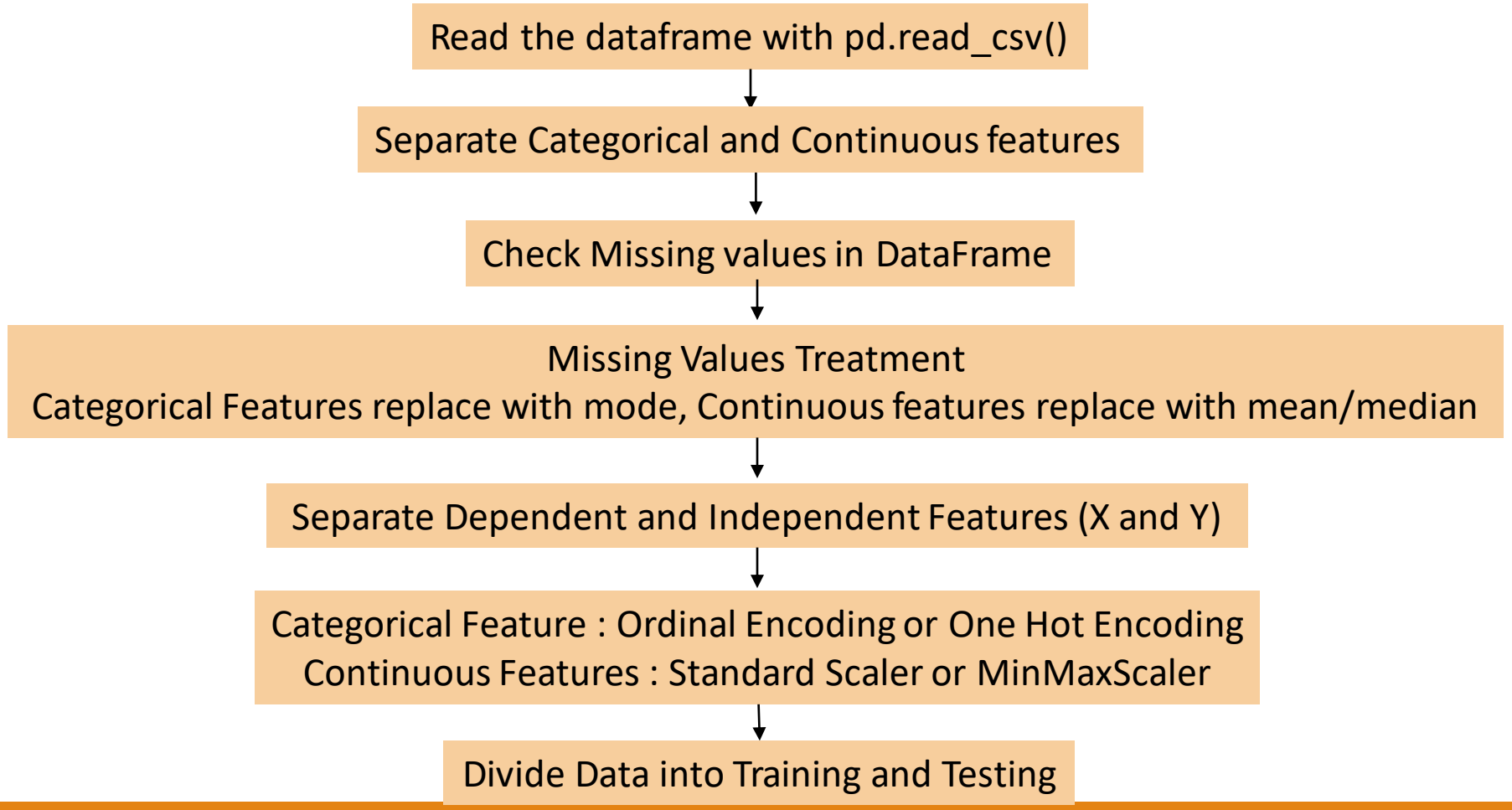
UTKARSH GAIKWAD

CLASS STARTING SHARP AT 3:07 PM

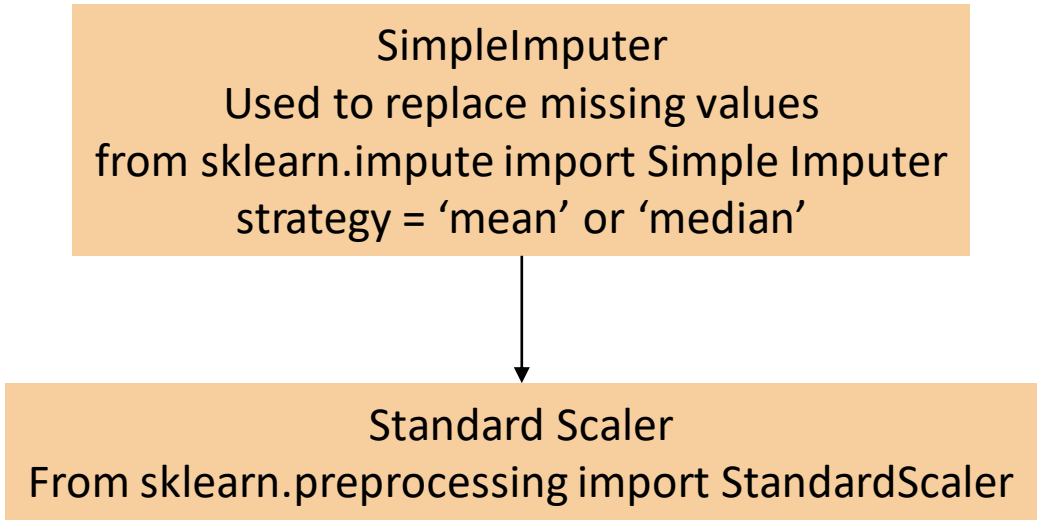
Basic steps in machine learning



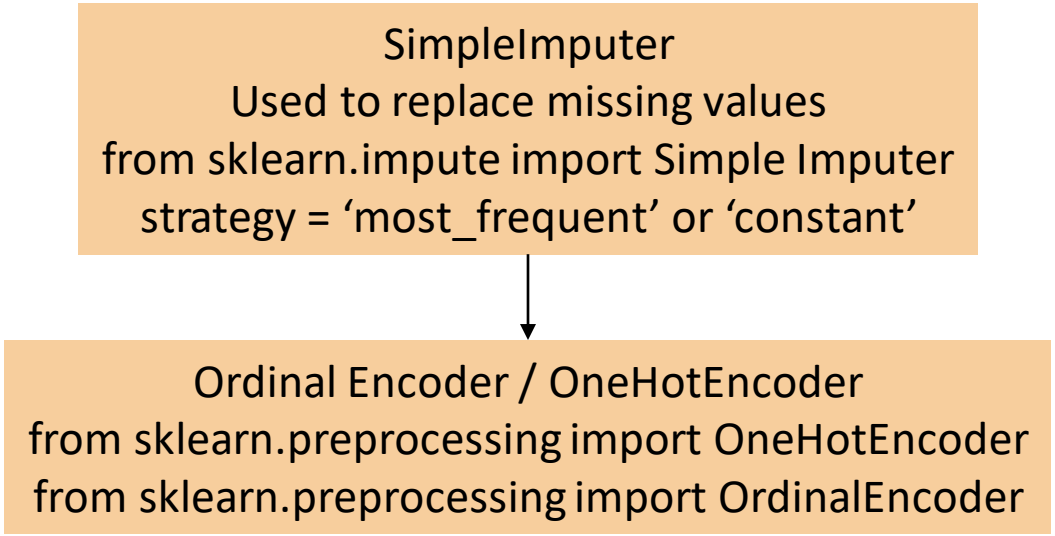
Basic Steps in creating a Data Preprocessing



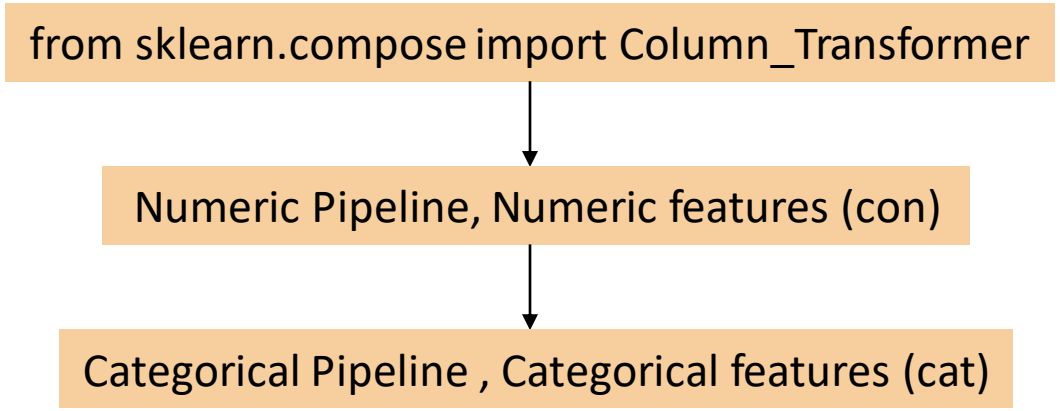
Sklearn pipeline numeric features



Sklearn pipeline for categorical features



Column Transformer



Example Code with all pipeline

```
from sklearn.pipeline import Pipeline
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import OrdinalEncoder, StandardScaler
from sklearn.compose import ColumnTransformer
```

Import the Pipeline functions

```
# Numeric pipeline
num_pipeline = Pipeline(steps=[('imputer',SimpleImputer(strategy='mean')),
                               ('scaler',StandardScaler())])
```

Numeric Pipeline

```
# Categorical Pipeline
cat_pipeline = Pipeline(steps=[('imputer',SimpleImputer(strategy='most_frequent')),
                               ('ordinal_encoder',OrdinalEncoder())])
```

Categorical Pipeline

Combine Cat and num pipes

```
# Column Transformer
preprocessor = ColumnTransformer([('num_pipeline',num_pipeline,con),
                                  ('cat_pipeline',cat_pipeline,cat)])
```

Thank you

PING ME ON SKYPE FOR ANY QUERIES

COMPLETE THE PRACTICAL THAT WAS ALL FOR TODAY