

```
@tracer(cat_col = ['Gender', 'Education'], numerical_col = [])
def loan_pipeline(f_path = '../pipelines/loan_train.csv'):
    data = pd.read_csv(f_path)

    # Loan_ID is not needed in training or prediction
    data = data.drop('Loan_ID', axis=1)

#     data = data.drop('Loan_Status', axis=1)

    numeric_features = data.select_dtypes(include=['int64', 'float64']).columns
    categorical_features = data.select_dtypes(include=['object']).drop(['Loan_Status'], axis=1).columns
    # do transformer on numeric & categorical data respectively
    numeric_transformer = Pipeline(steps=[
        ('imputer', SimpleImputer(strategy='median')),
        ('scaler', StandardScaler())])

    categorical_transformer = Pipeline(steps=[
        ('imputer', SimpleImputer(strategy='constant', fill_value='missing')),
        ('onehot', OneHotEncoder(handle_unknown='ignore'))])

    preprocessor = ColumnTransformer(
        transformers=[
            ('num', numeric_transformer, numeric_features),
            ('cat', categorical_transformer, categorical_features)])

    # classifier
    pipeline = Pipeline(steps=[('preprocessor', preprocessor),
                                ('classifier', RandomForestClassifier())])

    return pipeline
```

Start Pandas Opeation

```
-----
Inpected data = pd.read_csv(f_path)
-----

-----
Inpected data = data.drop('Loan_ID', axis=1)
-----
```

Start Sklearn Pipeline

```
-----
Operations SimpleImputer on Gender
-----

*****
Changes in categorical features!
```

Gender	
missing_count	-13
num_class	1
class_count	{'Male': 0, 'Female': 0, 'missing': 13}
class_percent	{'Male': -0.0172, 'Female': -0.0039, 'missing': 0.0212}

```
*****
-----
Operations OneHotEncoder on Gender
-----
```

Changes in categorical features!

Gender	
missing_count	0
num_class	-1
class_count	{0.0: 502, 1.0: 112}
class_percent	{0.0: 0.8176, 1.0: 0.1824}

```
*****
-----
Operations SimpleImputer on Education
-----

-----
Operations OneHotEncoder on Education
-----
```

Changes in categorical features!

Education	
missing_count	0
num_class	0
class_count	{1.0: 480, 0.0: 134}
class_percent	{1.0: 0.7818, 0.0: 0.2182}

