**DealDirtyDaty**(X, dirty\_col, Detect = None, Method = None) -> List

Find the dirty data and deal with that.

**Parameters**

**X**: DataFrame to be detected dirty data

**dirty\_col**: columns have dirty data

**Detect**: method that detect dirty values

None:[unique()]

Detecting dirty data and replace it with NaN

**method**: list of method that handle dirty values

None[ drop(), fillna()]

Handling Nan values.

**Return**

DataFrame

**How to operate**

**For in Detect** [unique()]

* Detect dirty data and replace with Nan

**For in method** [drop(), fillna()]

* Drop or fillna data
* Append to list

**FindEncodeScaler**(X, m, encode\_col, scaled\_col, scalers=None, encoders=None) -> List

Fine the best combination of scaler and encoder

Return best combination

**Parameters**

**X**: DataFrame to be scaled and encoded

**m**:data that replace NaN values.

**encode\_col**: columns to encode

**scaled\_col**: columns to scaled

**Encoders**: list of encoders:

None: [ OrdinalEncoder(), OneHotEncoder()]

Put columns to encoders.

**scalers**: list of scalers:

None: [ StandardScaler(), MinMaxScaler(), MaxAbsScaler(), RobustScaler()]

Put columns to scalers

Compare with Encoded set and Encoded+scaled set and return higher score.

**Return**

List of dataframes scaled and encoded

**How to operate**

**For in Scalers [**StandardScaler(), MinMaxScaler(),MaxAbsScaler(), RobustScaler()]

* Scaling scaled\_col\_col with each scaler

**For in Encoders** [OrdinalEncoder(),OneHotEncoder()]

* Scaling encode\_col with each encoder
* Compare it with only encoded columns
* Append it to list

**Return** list

**FindBestCombination**(X,y, preprocessing\_list, hyperParameters, models, parameters, to\_fit = None)

Find the best combination of scaler, encoder and fitting algorithm

Return the best combination

**X**: DataFrame except target column

**y**: Target column

**preprocessing\_list**: list that returned at FindEncodeScaler

**parameters**: parameters for models

**hyperParameters**: hyperParameters for svm and k-fold

**models**: list of models

None: [DecisionTreeEntropy(),DecisionTreeGini(),LogisticRegression(),SupportVectorMachine())

If you want to fit others ways, then put the model in the list

**How to operate**

**For in models [**DecisionTreeEntropy(), DecisiontreeGini(), LogisticRegression(),

SupportVectorMachine()]

* Fit X and y to models with variety parameter and hyperparemeters
* Compare accuracy of models