Feature Selection

Dropply constant features :-

```
import pandas as pd
    dab = pd. Dataframe {A: { _____ },
                         B: { _____ }
                         C _____
                         D _____ } }
```

class sklearn.feature_selection.VarianceThreshold (thrushold = 0.0)
   ↳ whichever having variance = 0 it will
      remove that.

```
from sklearn.feature_selection import VarianceThreshold
var_thres = VarianceThreshold (thrushold = 0)
var_thres.sit (data)
- var_thres.get_support ()
```

   ↳ array([[True, True, False, False]))
   data.drop (constant_columns, axis = 1)
      ↳ Variance Threshold.

2) **With Correlation :—**

```
from sklearn.datasets import load_boston
import pandas as pd.
```

        _____   Not correlated
                            or less
If independent ——X—— Dependent
   we remove those one.

— If correlated more than 90% then we ~~remove~~ don't remove them.

③ Feature Selection — Information Gain.
Mutual Information in Classification.

— MI Estimate mutual information for a discreate target variable.

If MI = 0. then lower dependency.

← Entropy Estimation used to find MI

So, $I(x;y) = H(x) - H(x/y)$
↳ Mutual Info.

from sklearn.feature selection import mutualinfo_classif
mutual-info = mutualinfo_classif (X-train, y-train)
↳ Select kBest used to pickup top features.

Info Grain	Mutual Info
←	Synonyms of Info Grain

④ MI in Regression.
Same as above, here import mutual-info-regressi on
⊕ — SelectPercentile import
— Top 20th percentile selected.
⊸ ∟ ∟