Choice_based_on_univariate model

February 17, 2019

In [2]:
```
#https://en.wikipedia.org/wiki/Univariate
import numpy as np
from sklearn.datasets import make_classification
X, y = make_classification(n_samples=800, n_features=100, n_informative=25, n_redundant
```

In [12]:
```
from sklearn.feature_selection import SelectPercentile, chi2, f_classif
from sklearn.preprocessing import Binarizer, scale
Xbin= Binarizer().fit_transform(scale(X))
Selector_chi2 = SelectPercentile(chi2, percentile=25).fit(Xbin, y)
Selector_f_classif = SelectPercentile(f_classif, percentile=25).fit(X, y)
chi_scores = Selector_chi2.get_support()
f_classif_scores = Selector_f_classif.get_support()
selected = chi_scores & f_classif_scores
print(selected)
informative = sum([x for x in selected if x == True])
print(informative)
```

```
[False  True False False False False  True  True False False False False
  True False False False False False False False False False False False
 False False False False False False  True False False False  True  True
 False False  True False False False False False False  True False False
 False False False False False False  True False  True False False  True
 False  True False False False False  True False False False False False
  True False False False  True False False False  True  True False False
  True False False False False  True False False False False False False
 False False False  True]
21
```

In [ ]: