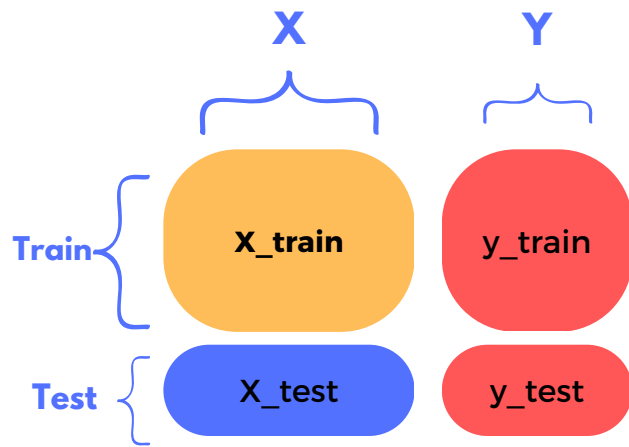


SCIKIT-LEARN

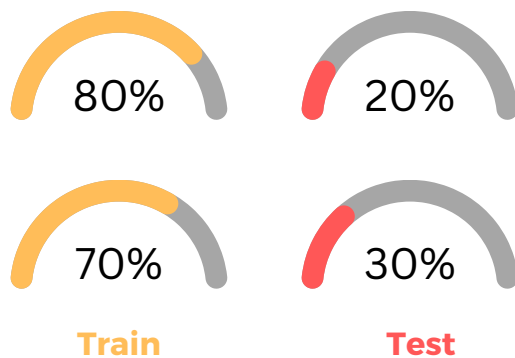
Cheatsheet for your Next Interview

Train-Test Split



- Method of dividing the data into two part: Training and Testing
- Training Data : Model Training
- Testing Data : Model Evaluation

Standard Ratio



```
from sklearn.model_selection import  
train_test_split  
X_train,X_test,y_train,y_test =  
train_test_split(X,y,test_size=0.3)
```

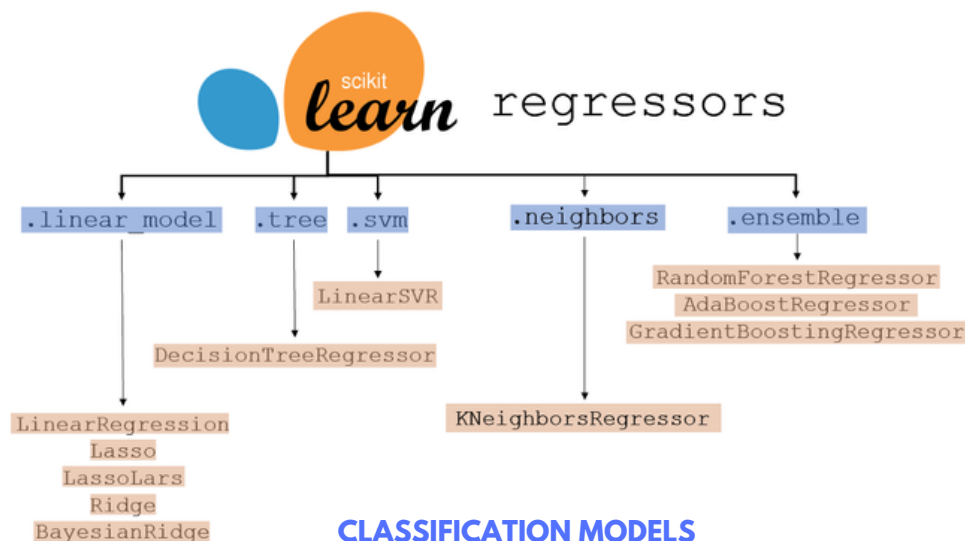
```
# Import  
from sklearn.ensemble import  
model_name
```

```
# Create a Instance  
model = model_name()
```

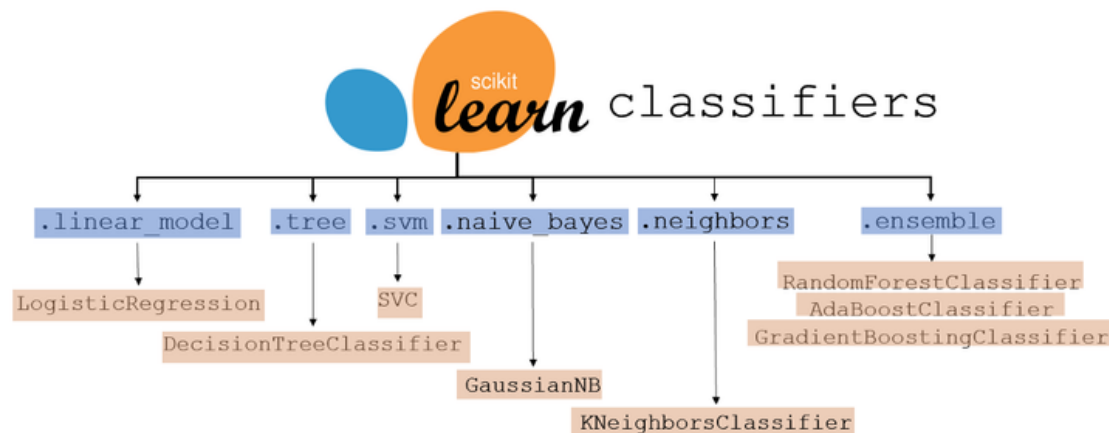
```
# fit model  
model.fit(X_train, y_train)
```

Supervised Learning Models

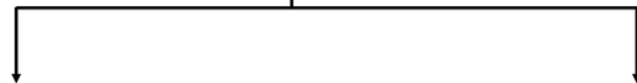
REGRESSION MODELS



CLASSIFICATION MODELS



METRICS



Classification

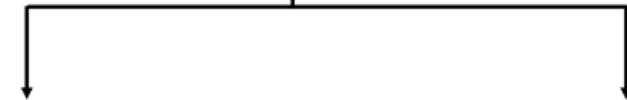
accuracy_score()
balanced_accuracy_score()
classification_report()
cohen_kappa_score()
confusion_matrix()
f1_score()
precision_score()
recall_score()
roc_auc_score()
Jaccard_score()

Regression

explained_variance_score()
mean_absolute_error()
mean_squared_error()
mean_squared_log_error()
median_absolute_error()
mean_absolute_percentage_error()
r2_score()

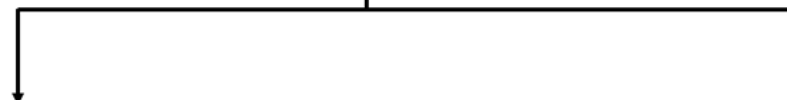
DIMENSIONALITY REDUCTION

CLUSTERING



AffinityPropagation()
AgglomerativeClustering()
Birch()
DBSCAN()
FeatureAgglomeration()
KMeans()
BisectingKmeans()

MiniBatchKmeans()
MeanShift()
OPTICS()
SpectralClustering()
SpectralBiclustering()
SpectralCoclustering



PCA()
KernelPCA()
IncrementalPCA()
SparsePCA()
MiniBatchSparsePCA()
LatentDirichletAllocation()
FastICA()

DictionaryLearning()
FactorAnalysis()
MiniBatchDictionaryLearning()
NMF()
MiniBatchNMF()
SparseCoder()
TruncatedSVD()