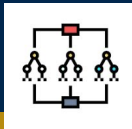


Random Forest



Prof Luiz Barboza
luiz@barboza.me.uk
@profluizbarboza

+

+

.

.

.

.

.



.

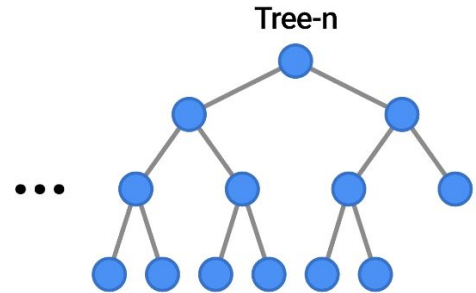
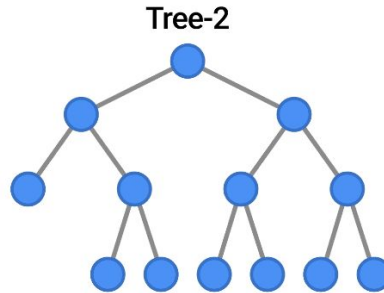
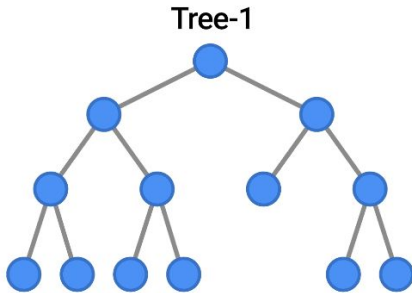
.

.

.

ML: Random Forest

EXAMPLES



...

+

+

.

.

.

.

.

□

.

.

•

•

ML: Random Forest

Uma floresta aleatória usa várias árvores de decisão treinadas aleatoriamente em diferentes partes dos dados e com diferentes características. Cada árvore "vota" em uma previsão, e a mais votada é o resultado final. Isso torna a floresta mais precisa e robusta que uma única árvore, mas também menos interpretável e mais custosa computacionalmente..

```
from sklearn.ensemble import RandomForestClassifier
forest = RandomForestClassifier(n_estimators = 5)
forest.fit(X_train,y_train)
train_accuracy= forest.score(X_train,y_train)*100
train_accuracy
```

93.75

Hyperparameter Tuning

Ajuste de hiperparâmetros é como ajustar as configurações de um modelo de aprendizado de máquina para obter o melhor desempenho. Diferentemente dos parâmetros regulares que o modelo aprende durante o treinamento, os hiperparâmetros são definidos antes do treinamento e controlam como o modelo aprende. Ajustá-los envolve testar diferentes combinações dessas configurações para encontrar aquelas que oferecem os melhores resultados em dados não vistos

```
for n in range(5,50,1):
    forest = RandomForestClassifier(n_estimators = n)
    forest.fit(X_train,y_train)
    train_accuracy= forest.score(X_train,y_train)*100
    preds = forest.predict(X_test)
    test_accuracy = accuracy_score(y_test,preds)*100
    print(n,train_accuracy,test_accuracy)
```

```
5 90.0 65.0
6 91.25 45.0
7 96.25 50.0
8 95.0 50.0
9 98.75 65.0
10 96.25 65.0
11 95.0 70.0
12 96.25 60.0
13 96.25 70.0
```

Exercício: Random Forest - RiscoCredito

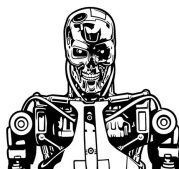
RiscoCredito - okk.csv e RiscoCredito - prever2.csv

- Analise o Histórico de Crédito e verifique quais são as probabilidades de um cliente ter seu crédito aprovado ou não, utilizando o algoritmo de **Random Forest**.
 - Avalie o crédito para um novo conjunto de potenciais clientes.

treino: <https://raw.githubusercontent.com/lcbjrrr/data/main/RiscoCredito%20-%20okk.csv>

test: <https://raw.githubusercontent.com/lcbjrrr/data/main/RiscoCredito%20-%20prever2.csv>

Classificação: Árvore de Decisão



ATIVIDADE: Random Forest

- Utilize a mesma base da atividade de Árvore de Decisão
- Procure realizar a previsão (inferência) de uma variável categórica através de uma Random Forest. Se certifique de medir seus níveis de assertividade. Compare o resultado com sua atividade do kNN. Não esqueça de realizar a interpretação das regras de negócio da árvore.
- Não esqueça de junto com seus códigos realizar suas análises/conclusões (use o botão de +Texto).