

# Aprendizado de Máquina e Deep Learning

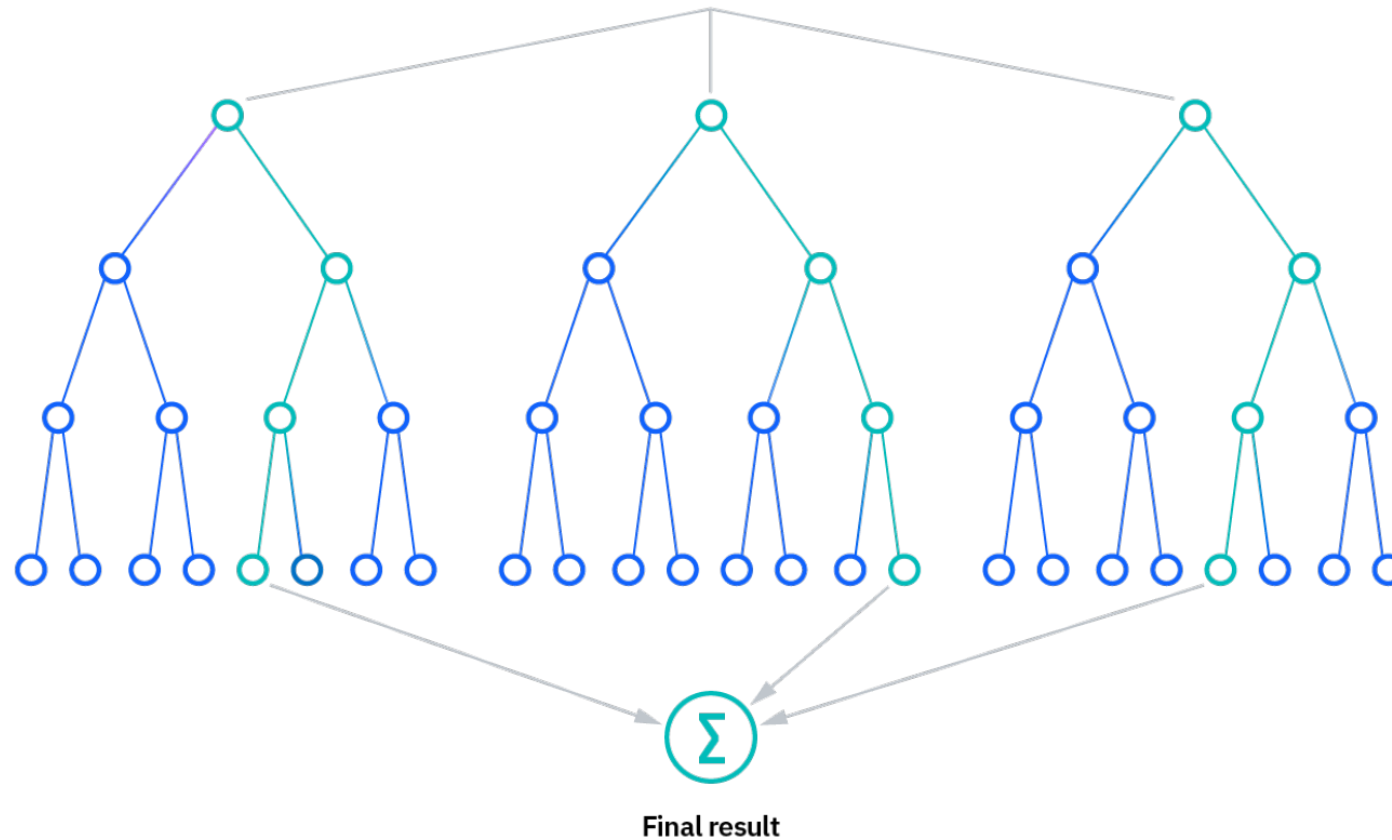
## Random Forest

Prof. Dr. Thiago Meirelles Ventura

# Random forest

- Combina o resultado de diversas árvores de decisão para gerar um único resultado
- Pode ser utilizado tanto para problemas de classificação quanto para regressão

# Random forest



<https://www.ibm.com/topics/random-forest>

# Random forest

- O algoritmo funciona gerando vários subconjunto dos atributos
  - 1 novo subconjunto para cada árvore existente
- Como os subconjuntos são aleatórios, teremos várias árvores diferentes

# Random forest

- Diferente da árvore de decisão simples, que avalia toda a base para determinar os nós, a Random Forest terá árvores que até o nó raiz será diferente entre as árvores existentes
- Isso traz novas perspectivas de estimativas

# Random forest

- Ao prever um dado, pode ser considerado
  - a média dos resultados (em problemas de regressão) ou
  - a classe que mais aparece nos resultados (em problemas de classificação)

# Implementação

- Construção do modelo, treinamento e avaliação

```
from sklearn.ensemble import RandomForestClassifier
```

```
randomforest = RandomForestClassifier(n_estimators=100)
```

```
randomforest.fit(X_train, y_train)
```

```
predicao = randomforest.predict(X_test)
```

# Exercício 7

- Crie um modelo de Random Forest
- Aproveite da classe `RandomForestClassifier`
- Use o conjunto de dados “emprego.csv”



# Implementação

script randomforest.ipynb