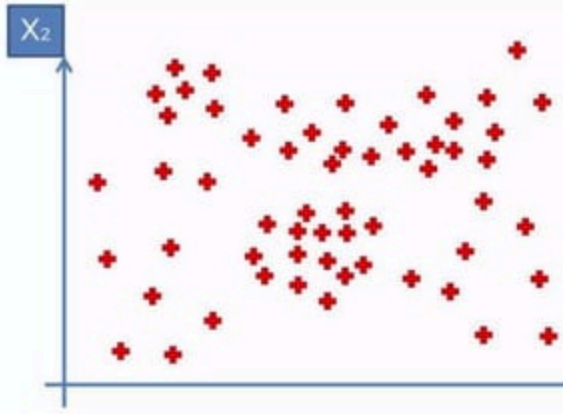
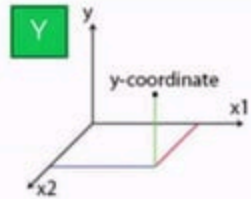
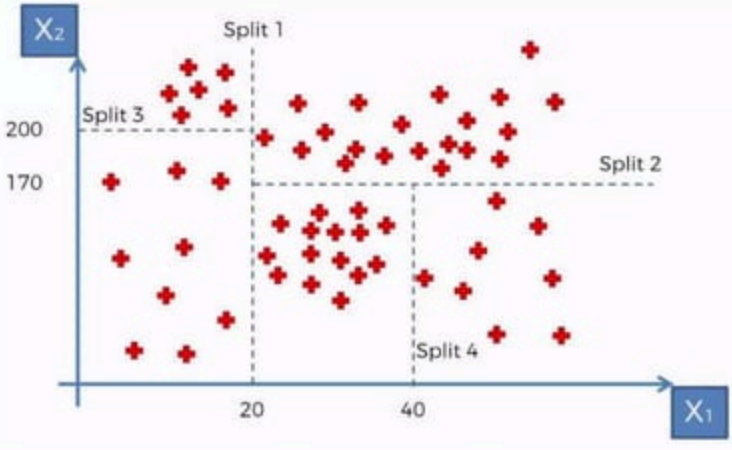
Decision Tree Intuition





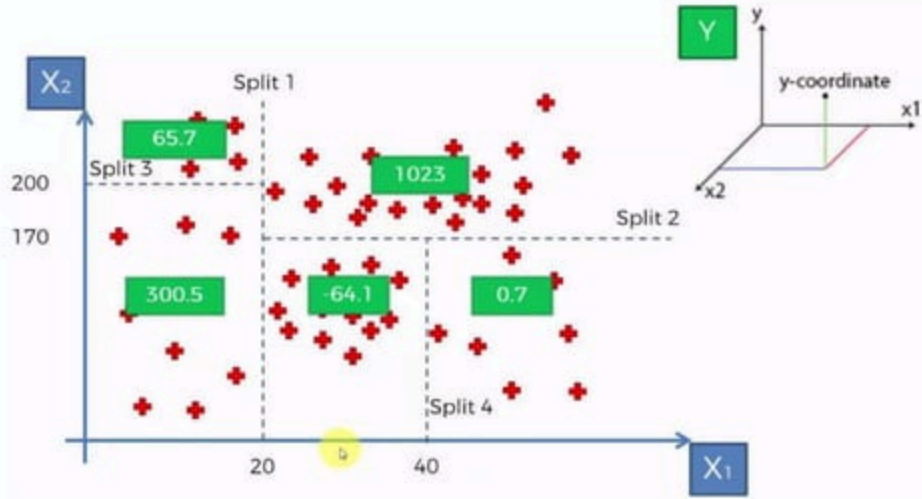
© SuperDataScience

Decision Tree Intuition



© SuperDataScience

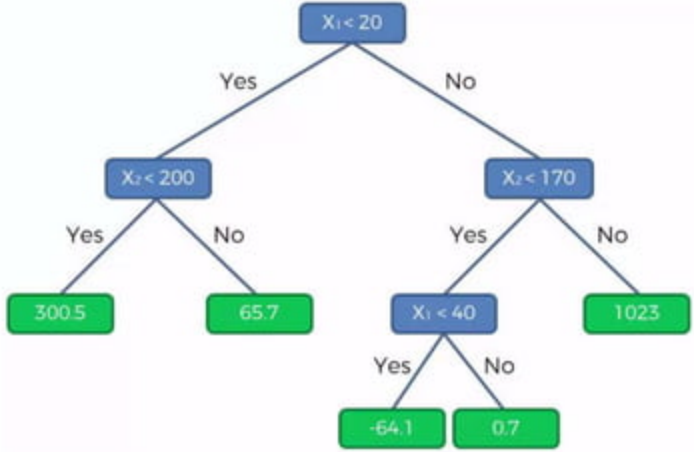
Decision Tree Intuition



The value in green box represents the average of data points in that split

© SuperDataScience

Decision Tree Intuition



© SuperDataScience

Decision Tree Intuition

© SuperDataScience

PYTHON

READING FILE DYNAMICALLY

```
from tkinter import *
from tkinter.filedialog import askopenfilename

root = Tk()
root.withdraw()
root.update()
file_path = askopenfilename()
root.destroy()
```

IMPORTING LIBRARIES

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

IMPORTING DATASET

```
dataset = pd.read_csv(file_path)

X= dataset.iloc[:,1:2].values
y= dataset.iloc[:,2:3].values
```

DECISION TREE REGRESSOR

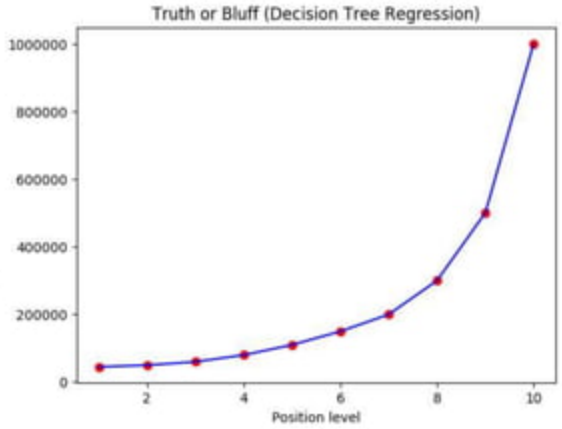
```
from sklearn.tree import DecisionTreeRegressor
regressor = DecisionTreeRegressor(random_state=42)

model = regressor.fit(X,y)
```

PREDICTION

model.predict(6.5)

SIMPLE PLOT

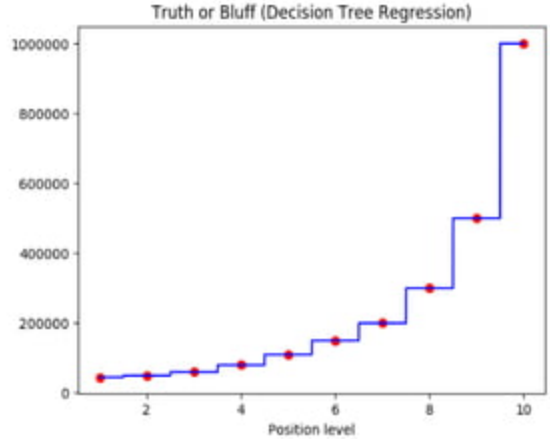
```
plt.scatter(X,y,color="red")
plt.plot(X,model.predict(X),color="blue")
plt.title('Truth or Bluff (Decision Tree Regression)')
plt.xlabel('Position level')
plt.ylabel('Salary')
plt.show()
```



NOTE : Whats wrong here ? Well in the simple plot the Decision Tree Regressor model is treated as a c
But it is not a continuous model. Decision Tree Regressor is a discrete model hence it should be treat
FIX : plotting the same graph with grid with small step size say 0.01 will help us visualize better

UPDATED PLOT

```
X_grid = np.arange(min(X), max(X), 0.001)
X_grid = X_grid.reshape((len(X_grid), 1))
plt.scatter(X, y, color = 'red')
plt.plot(X_grid, regressor.predict(X_grid), color = 'blue')
plt.title('Truth or Bluff (Decision Tree Regression)')
plt.xlabel('Position level')
plt.ylabel('Salary')
plt.show()
```



Note : Here the graph that is plotted gives us the clear discrete structure