

# DTR - Decision Tree Regression Algorithm

In [2]: *#importing libraries*

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

In [3]: *#importing the dataset*

```
dataset = pd.read_csv(r"C:\Users\Jan Saida\Downloads\emp_sal.csv")
dataset
```

Out[3]:

	Position	Level	Salary
0	Jr Software Engineer	1	45000
1	Sr Software Engineer	2	50000
2	Team Lead	3	60000
3	Manager	4	80000
4	Sr manager	5	110000
5	Region Manager	6	150000
6	AVP	7	200000
7	VP	8	300000
8	CTO	9	500000
9	CEO	10	1000000

In [4]: `x=dataset.iloc[:, 1:2].values` *#independent variable*  
`y=dataset.iloc[:,2].values` *#dependent variable*

In [5]: x

```
Out[5]: array([[ 1],
               [ 2],
               [ 3],
               [ 4],
               [ 5],
               [ 6],
               [ 7],
               [ 8],
               [ 9],
               [10]], dtype=int64)
```

In [6]: y

```
Out[6]: array([ 45000,  50000,  60000,  80000, 110000, 150000, 200000,
                300000, 500000, 1000000], dtype=int64)
```

In [7]: *#decision tree algorithm*

```
from sklearn.tree import DecisionTreeRegressor
```

```
In [8]: tree_reg = DecisionTreeRegressor( splitter="best", random_state=0)
tree_reg
```

```
Out[8]: ▼      DecisionTreeRegressor ⓘ ?
DecisionTreeRegressor(random_state=0)
```

In [9]: tree\_reg.fit(x,y)

```
Out[9]: ▼      DecisionTreeRegressor ⓘ ?
DecisionTreeRegressor(random_state=0)
```

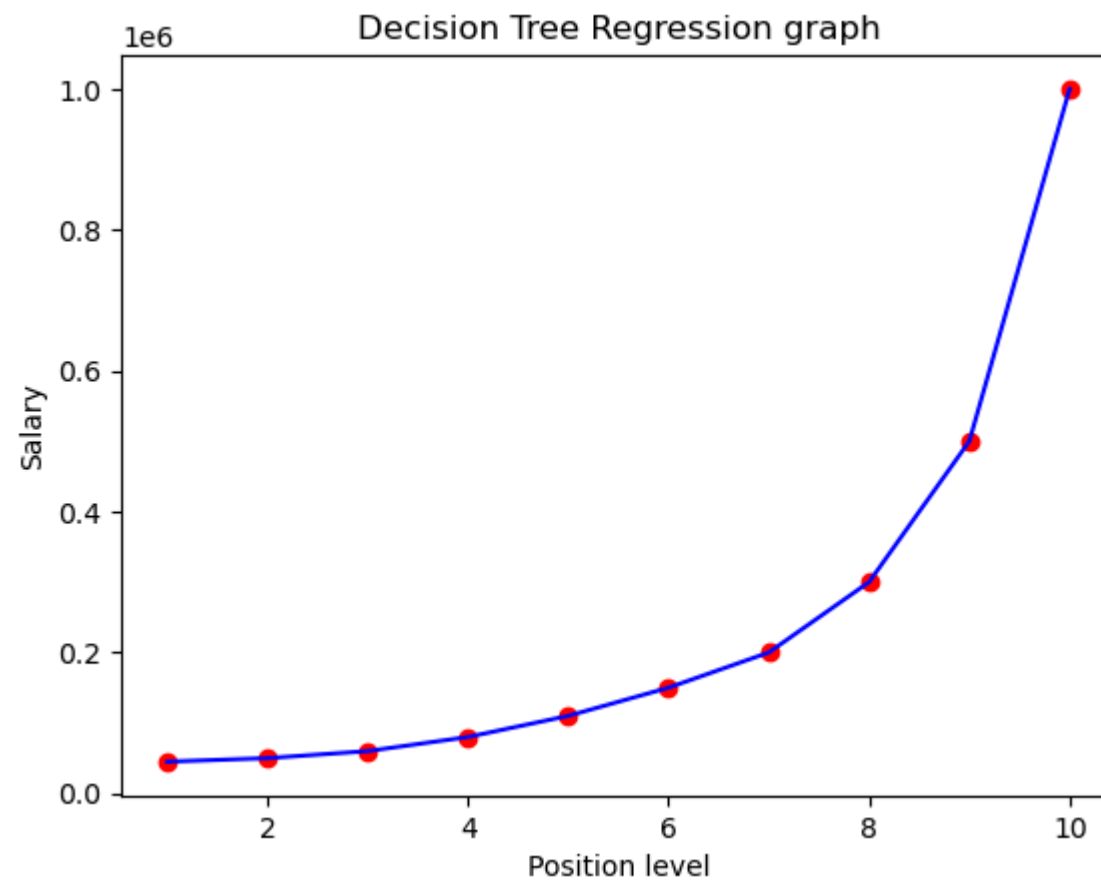
In [10]: *#decision tree Predictions*

```
tree_reg_pred=tree_reg.predict([[6.5]])  
tree_reg_pred
```

Out[10]: array([150000.])

In [11]: *# Desicion Tree regressor Visualizations*

```
plt.scatter(x,y,color='red')  
plt.plot(x,tree_reg.predict(x),color='blue')  
plt.title('Decision Tree Regression graph')  
plt.xlabel('Position level')  
plt.ylabel('Salary')  
plt.show()
```



In [ ]: