

Random Forest Regression

Importing the libraries

```
In [1]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

Importing the dataset

```
In [2]: dataset = pd.read_csv("Position_Salaries.csv")
x = dataset.iloc[:, 1:-1].values
y = dataset.iloc[:, -1].values
```

```
In [3]: print(x)
```

```
[[ 1]
 [ 2]
 [ 3]
 [ 4]
 [ 5]
 [ 6]
 [ 7]
 [ 8]
 [ 9]
[10]]
```

```
In [4]: print(y)
```

```
[ 45000  50000  60000  80000 110000 150000 200000 300000 500000
1000000]
```

Training the Random Forest Regression model on the whole dataset

```
In [11]: from sklearn.ensemble import RandomForestRegressor
model = RandomForestRegressor(n_estimators = 20, random_state = 0)
model.fit(x, y)
```

```
Out[11]: RandomForestRegressor(n_estimators=20, random_state=0)
```

Predicting a new result

```
In [12]: model.predict([[6.5]])
```

```
Out[12]: array([166000.])
```

Visualising the Random Forest Regression results (higher resolution)

```
In [13]: X_grid = np.arange(min(x), max(x), 0.01)
X_grid = X_grid.reshape((len(X_grid), 1))
```

```
plt.scatter(x, y, color = 'red')
plt.plot(X_grid, model.predict(X_grid), color = 'blue')
plt.title('Truth or Bluff (Random Forest Regression)')
plt.xlabel('Position level')
plt.ylabel('Salary')
plt.show()
```

