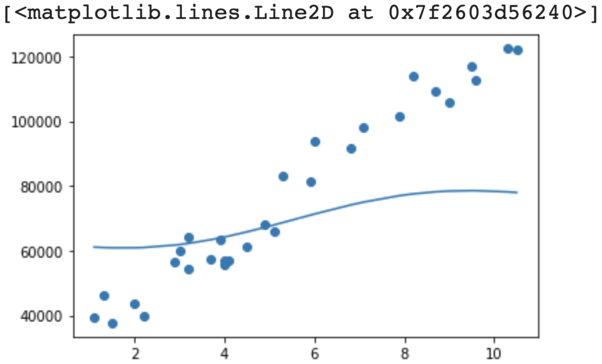
Kode program yang akan diajarkan di sini bisa diunduh di tautan [berikut ini](https://colab.research.google.com/drive/1gJFkAP9QJoBkWxq3yWkDwhjbu3XC6Um_?usp=sharing). Silahkan buat salinan dari kode ini dan jalankan di Google Colab Anda.

**Tujuan**

Untuk latihan kali ini kita akan menggunakan grid search untuk mencari parameter terbaik pada sebuah model SVR. Dataset yang kita gunakan sama seperti di modul 3 yaitu dataset gaji. Pada latihan tersebut model belum mampu menyesuaikan pola pada data dengan baik. Tujuan kita adalah menggunakan grid search untuk mencari parameter yang lebih baik bagi model.



**Tahapan Latihan**

Tahapan latihan kali ini sebagai berikut:

1. Ubah dataset data menjadi Dataframe.
2. Definisikan parameter yang akan diuji.
3. Jalankan fungsi grid search.
4. Latih model dengan parameter terbaik hasil grid search.
5. Buat plot dari model.

**Codelab**

Pertama, unduh berkas "Salary\_Data.csv", dari [tautan](https://www.kaggle.com/vivinbarath/simple-linear-regression-for-salary-data/data) berikut ini. Jika sudah jangan lupa upload ke Google Colab Anda. Berikut caranya.

1. import pandas as pd
3. # membaca dataset dan mengubahnya menjadi dataframe
4. data = pd.read\_csv('Salary\_Data.csv')

Kemudian pisahkan atribut dan label pada dataset. Masih ingat bukan, bahwa jika hanya terdapat 1 atribut pada dataset kita perlu mengubah bentuknya agar bisa dipakai pada pelatihan model.

1. import numpy as np
3. # memisahkan atribut dan label
4. X = data['YearsExperience']
5. y = data['Salary']
7. # mengubah bentuk atribut
8. X = np.array(X)
9. X = X.reshape(-1,1)

Selanjutnya, untuk menggunakan grid search, kita impor library GridSearchCV dari sklearn.model\_selection. Lalu kita buat model yang ingin kita uji dengan grid search dalam hal ini model SVR. Kemudian kita buat sebuah python dictionary yang berisi nama parameter yang akan diuji, serta nilai-nilainya. Selanjutnya kita buat objek grid search dan mengisi parameter-parameternya. Parameter pertama adalah model yang akan kita uji. Parameter kedua adalah dictionary yang berisi kumpulan parameter dari model yang akan diuji. Terakhir kita panggil fungsi fit() pada objek grid search yang telah kita buat.

1. from sklearn.model\_selection import GridSearchCV
2. from sklearn.svm import SVR
4. # membangun model dengan parameter C, gamma, dan kernel
5. model = SVR()
6. parameters = {
7. 'kernel': ['rbf'],
8. 'C':     [1000, 10000, 100000],
9. 'gamma': [0.5, 0.05,0.005]
10. }
11. grid\_search = GridSearchCV(model, parameters)
13. # melatih model dengan fungsi fit
14. grid\_search.fit(X,y)

Setelah grid search mencari parameter terbaik pada model, kita bisa menampilkan parameter terbaik dengan memanggil atribut best\_params\_ dari objek grid search.

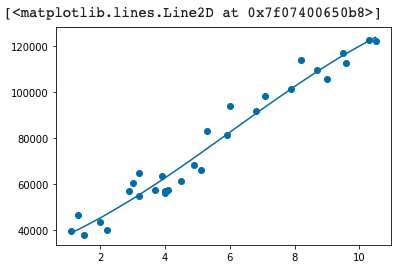
1. # menampilkan parameter terbaik dari objek grid\_search
2. print(grid\_search.best\_params\_)

Selanjutnya Anda bisa mencoba membuat model SVM baru dengan parameter hasil grid search dan melatihnya pada data.

1. # membuat model SVM baru dengan parameter terbaik hasil grid search
2. model\_baru  = SVR(C=100000, gamma=0.005, kernel='rbf')
3. model\_baru.fit(X,y)

Terakhir kita bisa memvisualisasikan SVR dengan parameter hasil grid search. Dapat dilihat dari hasil plot bahwa grid search berhasil mencari parameter yang lebih baik sehingga meningkatkan performa dari model.

1. import matplotlib.pyplot as plt
2. plt.scatter(X, y)
3. plt.plot(X, model\_baru.predict(X))

****

Selamat! Anda telah memahami dan bisa menggunakan grid search untuk mencari parameter terbaik dari suatu model. Dengan grid search Anda akan menghemat banyak waktu dalam mencari parameter terbaik dari model machine learning yang anda kembangkan ke depannya.