

Laporan Tugas #1 Regresi

II4035 Sistem Cerdas

13316068/Kriesnawan Handy Sugianto

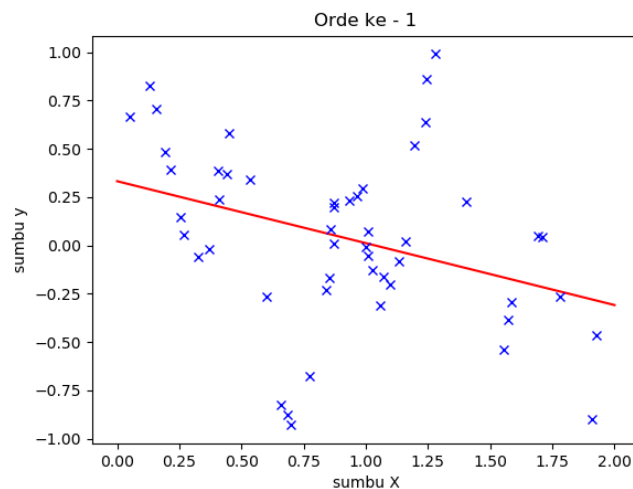
1 Polynomial Regression

1.1 Penjelasan teknik yang dipakai

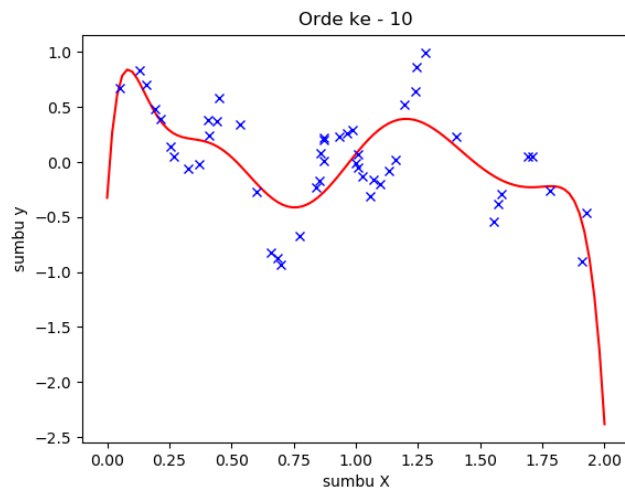
Polynomial Regression merupakan kombinasi dari linear regression dimana terbentuk berdasarkan variable independent x dan variable dependent y yang dapat dimodelkan dengan orde polynomial tertentu. Setiap orde yang diterapkan akan memiliki nilai Root Mean Square Error (RMSE) yang berbeda-beda. Oleh karena itu, dibutuhkan nilai RMSE yang terkecil, namun hal ini harus diimbangi oleh proporsional grafik supaya tidak terjadi *overfit*.

1.2 Grafik hasil prediksi (underfit, overfit, terbaik)

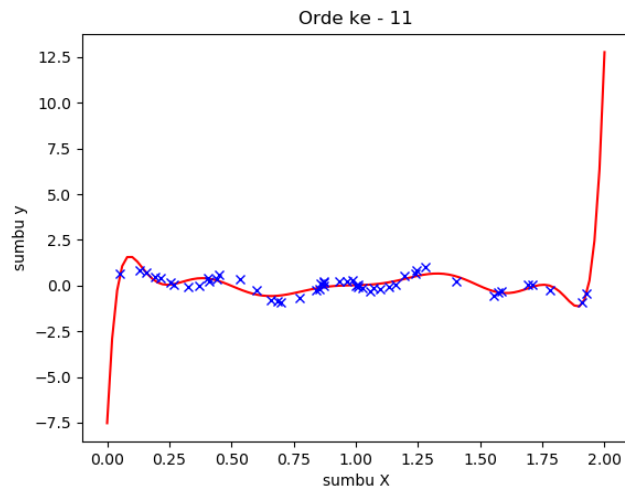
Gambar 1.1 sampai Gambar 1.3 merupakan hasil regresi dimana Gambar 1.1 merupakan hasil regresi under fit, Gambar 1.2 merupakan hasil regresi best fit dan Gambar 1.3 merupakan hasil regresi over fit. Gambar 1.4, 1.5 dan 1.6 merupakan hasil prediksi berdasarkan regresi under fit, best fit dan over fit dengan menggunakan data test pada file [II4035-regresi-test.csv](#).



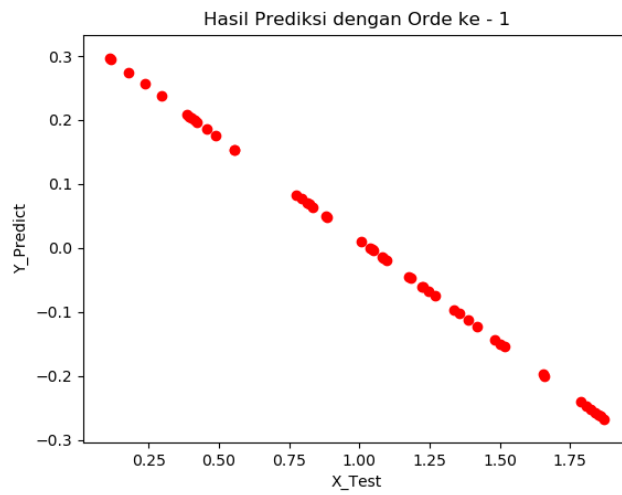
Gambar 1.1. Regresi orde 1 (Under Fit)



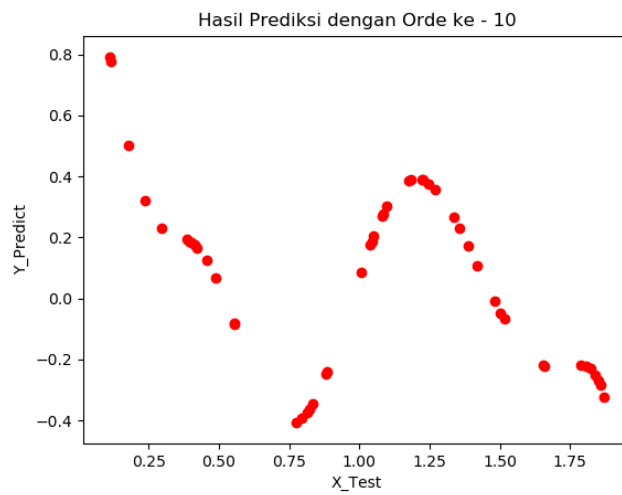
Gambar 1.2. Regressi orde 10 (Best Fit)



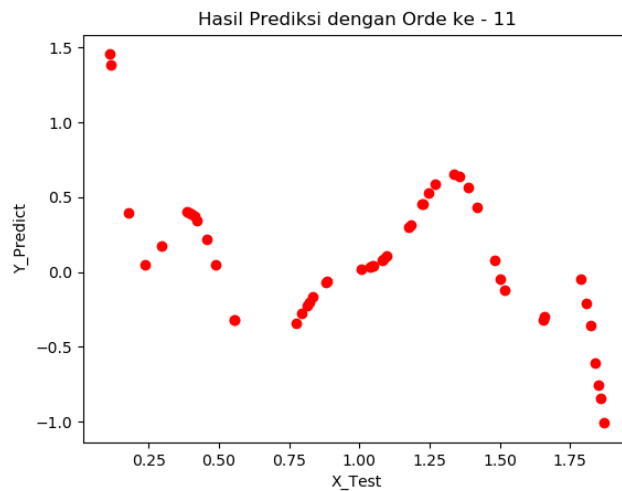
Gambar 1.3. Regressi orde 11 (Over Fit)



Gambar 1.4. Hasil Prediksi Berdasarkan Grafik Under Fit



Gambar 1.5. Hasil Prediksi Berdasarkan Grafik Best Fit

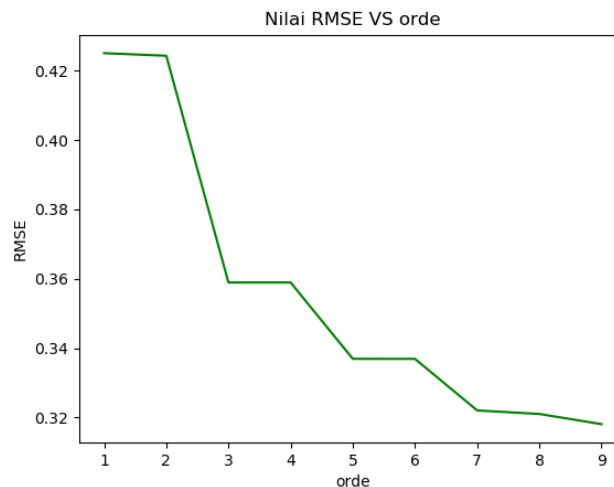


Gambar 1.6. Hasil Prediksi Berdasarkan Grafik Over Fit

1.3 Masalah-masalah yang dijumpai

Ada beberapa masalah yang ditemukan seperti syntax yang belum familiar seperti untuk menampilkan grafik, pengolahan data dan menyimpan data.

1.4 Ulasan



Gambar 1.7. Grafik RMSE terhadap orde

Berdasarkan hasil plot grafik diatas, nilai RMSE akan semakin kecil seiring bertambahnya orde. didapat nilai Root Mean Square Error (RMSE) untuk underfit, bestfit dan overfit berturut-turut sebesar 0,425 , 0,317 dan 1,242. Walaupun nilai RMSE terkecil adalah grafik ke-3(overfit), namun jika dilihat dari grafikya terlalu presisi. Sehingga bestfit memiliki nilai RMSE sebesar 0,317.

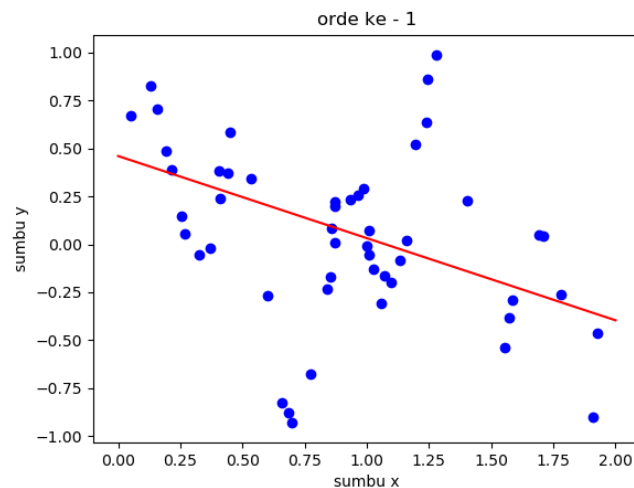
2 SVM Regression

2.1 Penjelasan teknik yang dipakai

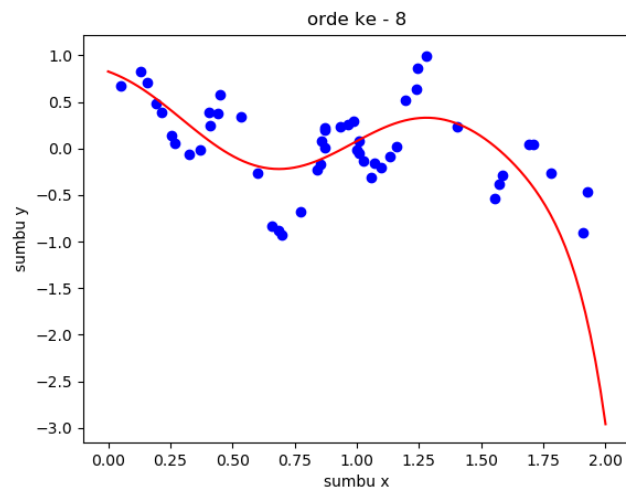
Support Vector Machine (SVM) regression merupakan suatu metoda nonparametric dimana berdasarkan fungsi kernel. SVM regression dapat digunakan untuk regresi linear maupun regresi polynomial. Metoda ini memiliki variable hyperparameter ϵ untuk membatasi kesalahan atau error. Ketika nilai ϵ semakin besar, akan semakin banyak data yang digunakan untuk mengkonstruksi garis regresi namun hal ini akan memberikan nilai error yang semakin besar.

2.2 Grafik hasil prediksi (underfit, overfit, terbaik)

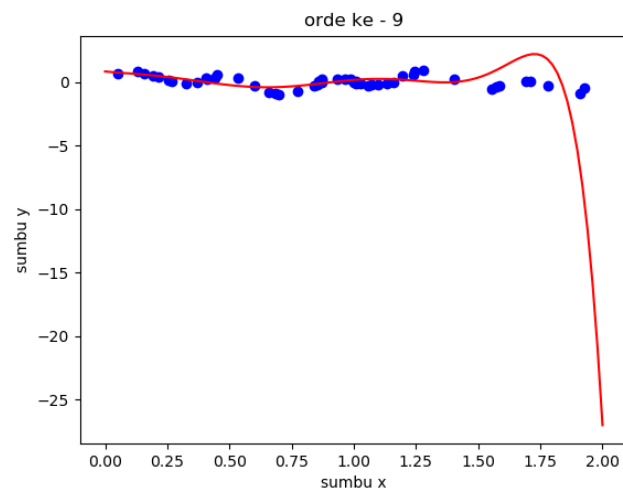
Gambar 2.1 sampai Gambar 2.3 merupakan hasil regresi dimana Gambar 2.1 merupakan hasil regresi under fit, Gambar 2.2 merupakan hasil regresi best fit dan Gambar 2.3 merupakan hasil regresi over fit. Gambar 2.4, 2.5 dan 2.6 merupakan hasil prediksi berdasarkan regresi under fit, best fit dan over fit dengan menggunakan data test pada file [II4035-regresi-test.csv](#).



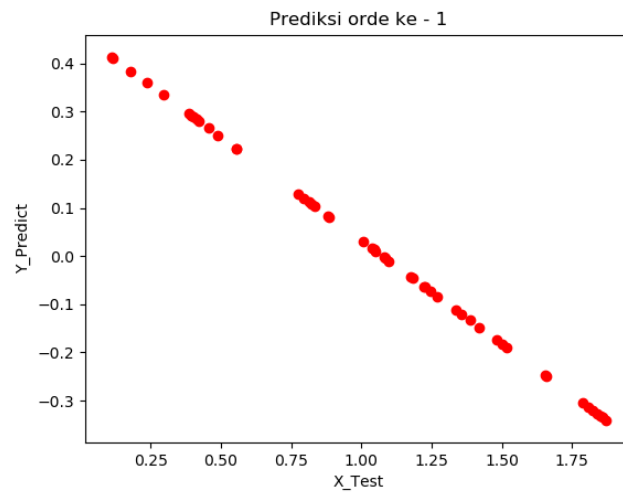
Gambar 2.1. Regresi orde 1 (Under Fit)



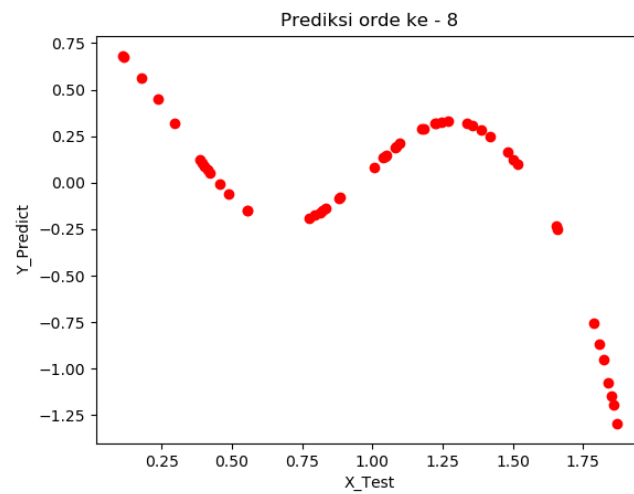
Gambar 2.2. Regressi orde 8 (Best Fit)



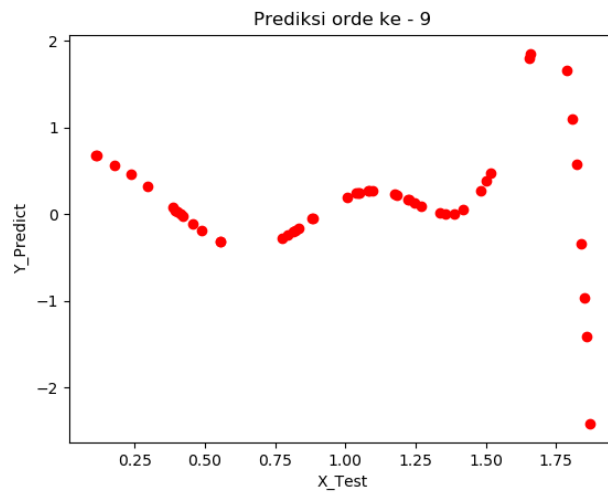
Gambar 2.3. Regressi orde 9 (Over Fit)



Gambar 2.4. Hasil Prediksi Berdasarkan Grafik Under Fit



Gambar 2.5. Hasil Prediksi Berdasarkan Grafik Best Fit

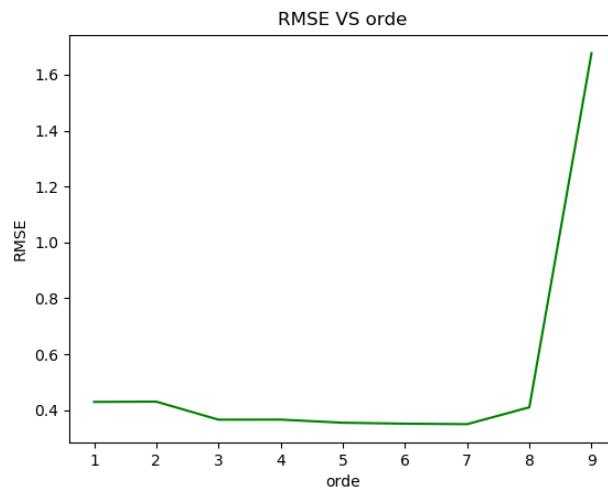


Gambar 2.6. Hasil Prediksi Berdasarkan Grafik Over Fit

2.3 Masalah-masalah yang dijumpai

Ada beberapa masalah yang ditemukan seperti syntax yang belum familiar seperti untuk menampilkan grafik, pengolahan data dan menyimpan data.

2.4 Ulasan



Gambar2.7. Grafik RMSE terhadap orde

Berdasarkan hasil plot grafik diatas, nilai RMSE akan semakin kecil seiring bertambahnya orde pada range orde 1 sampai 7, namun RMSE meningkat drastis pada orde 7 sampai 9. didapat nilai Root

Mean Square Error (RMSE) untuk underfit, bestfit dan overfit berturut-turut sebesar 0,429 , 0,410 dan 1,677. Sehingga bestfit memiliki nilai RMSE yang terkecil.

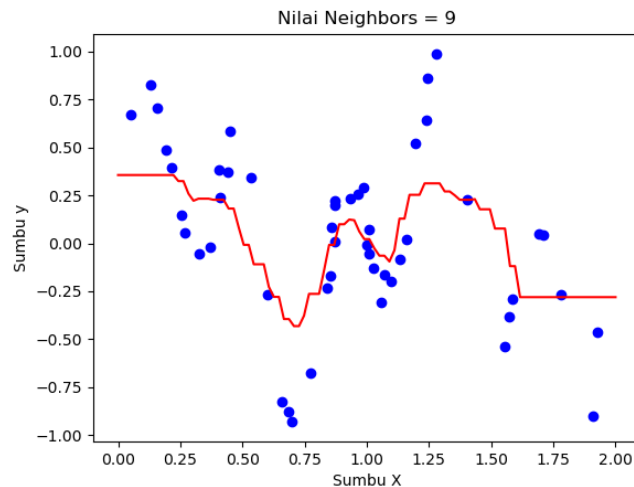
3 KNN Regression

3.1 Penjelasan teknik yang dipakai

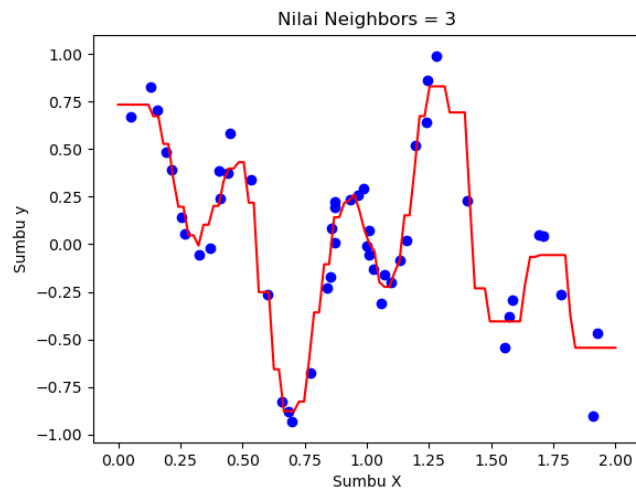
K nearest neighbors (KNN) regression merupakan metoda untuk memprediksi suatu target numerical yang berdasarkan pengukuran data yang sama. Untuk nilai K yang tinggi akan memberikan presisi regresi yang semakin tinggi.

3.2 Grafik hasil prediksi (underfit,overfit, terbaik)

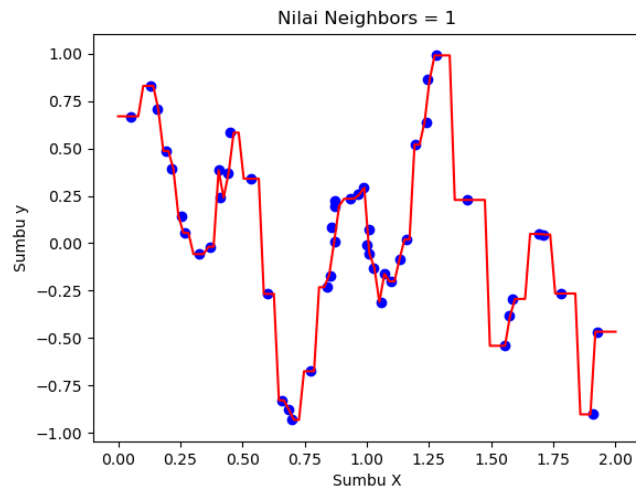
Gambar 3.1 sampai Gambar 3.3 merupakan hasil regresi dimana Gambar 3.1 merupakan hasil regresi under fit, Gambar 3.2 merupakan hasil regresi best fit dan Gambar 3.3 merupakan hasil regresi over fit. Gambar 3.4, 3.5 dan 3.6 merupakan hasil prediksi berdasarkan regresi under fit, best fit dan over fit dengan menggunakan data test pada file [II4035-regresi-test.csv](#).



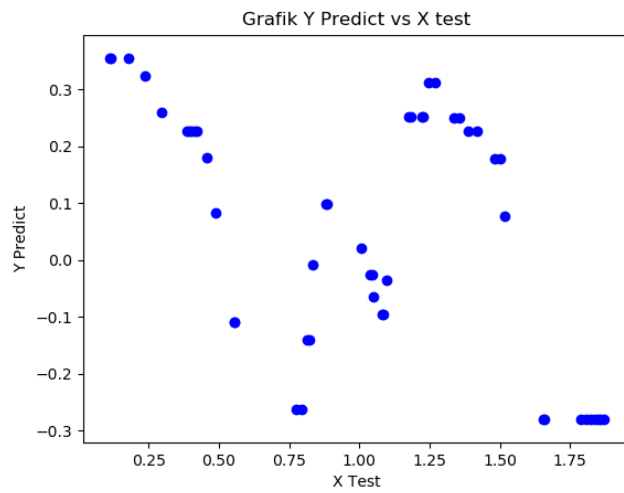
Gambar 3.1. Regresi Neighbors = 9 (Under Fit)



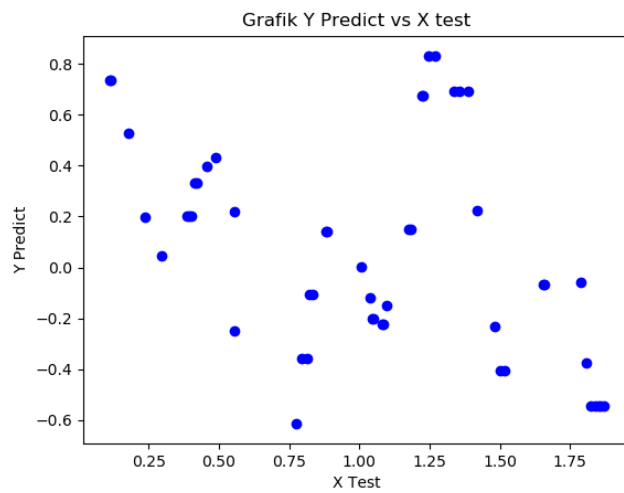
Gambar 3.2. Regressi Neighbors = 3 (Best Fit)



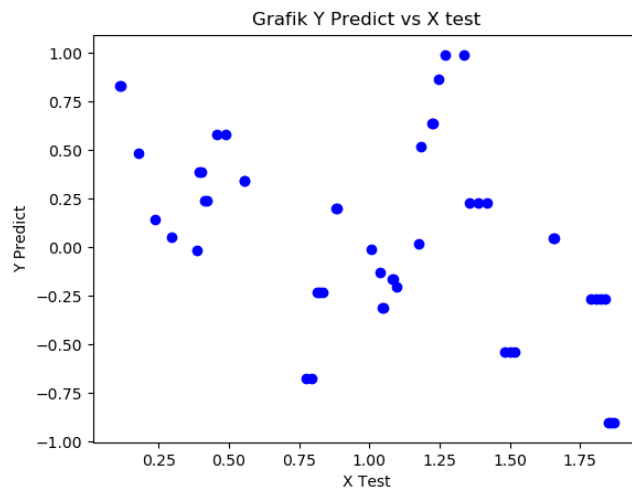
Gambar 3.3. Regressi Neighbors = 1 (Over Fit)



Gambar 3.4. Hasil Prediksi Berdasarkan Grafik Under Fit



Gambar 3.5. Hasil Prediksi Berdasarkan Grafik Best Fit

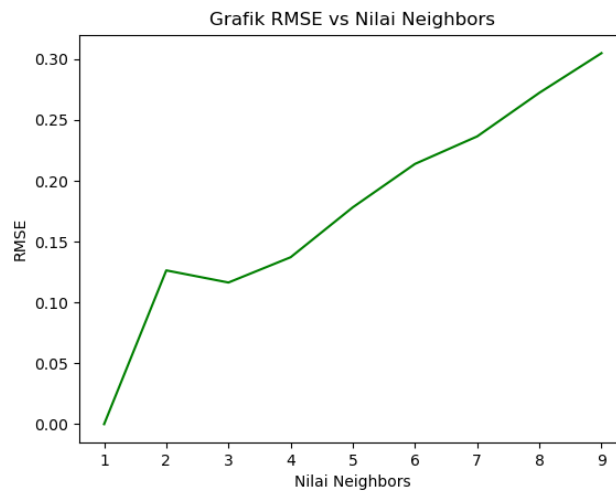


Gambar 3.6. Hasil Prediksi Berdasarkan Grafik Over Fit

3.3 Masalah-masalah yang dijumpai

Ada beberapa masalah yang ditemukan seperti syntax yang belum familiar seperti untuk menampilkan grafik, pengolahan data dan menyimpan data.

3.4 Ulasan



Gambar 3.7. Grafik RMSE terhadap nilai neighbors

Berdasarkan hasil plot grafik diatas, nilai RMSE akan semakin besar seiring bertambahnya orde. didapat nilai Root Mean Square Error (RMSE) untuk underfit, bestfit dan overfit berturut-turut sebesar 0,304 , 0,116 dan 0. Walaupun nilai RMSE terkecil adalah grafik ke-3(overfit), namun jika dilihat dari grafikya terlalu presisi. Sehingga bestfit memiliki nilai RMSE yang 0,116.

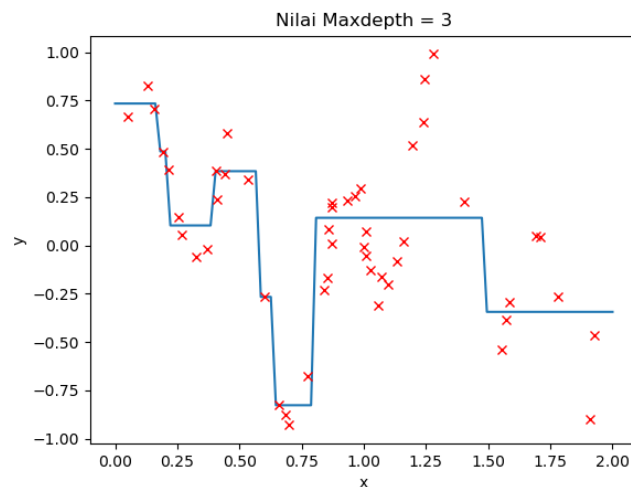
4 Decision Tree Regression

4.1 Penjelasan teknik yang dipakai

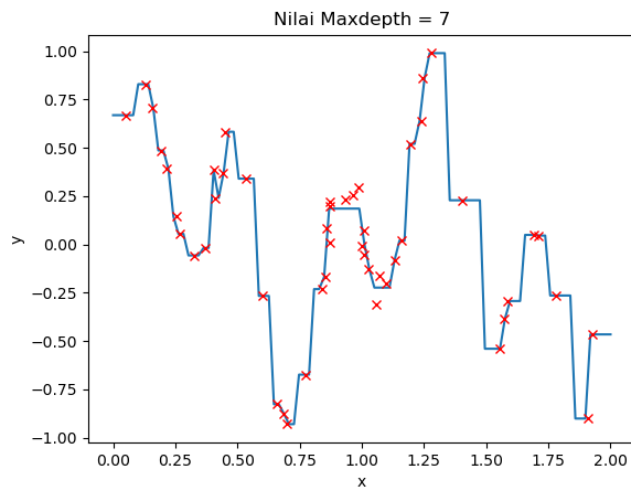
Decision Tree Regression merupakan metoda untuk membagi kumpulan data menjadi beberapa bagian. Setiap kumpulan data akan diwakili dengan nilai rata-ratanya. Banyaknya pembagian data dinyatakan oleh variable `max_depth` pada regresi ini.

4.2 Grafik hasil prediksi (underfit, overfit, terbaik)

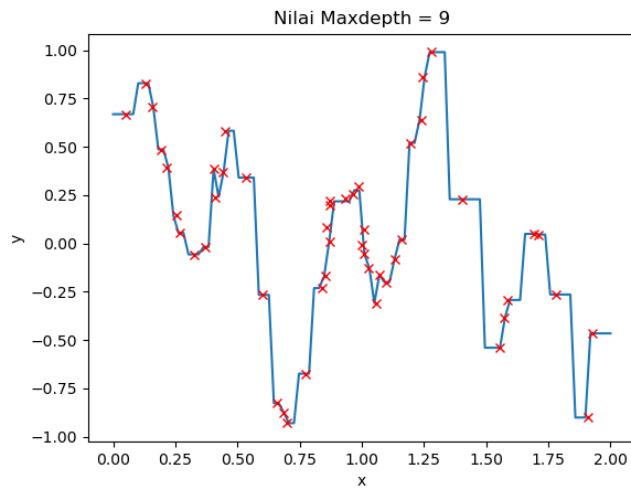
Gambar 4.1 sampai Gambar 4.3 merupakan hasil regresi dimana Gambar 4.1 merupakan hasil regresi under fit, Gambar 4.2 merupakan hasil regresi best fit dan Gambar 4.3 merupakan hasil regresi over fit. Gambar 4.4, 4.5 dan 4.6 merupakan hasil prediksi berdasarkan regresi under fit, best fit dan over fit dengan menggunakan data test pada file [II4035-regresi-test.csv](#).



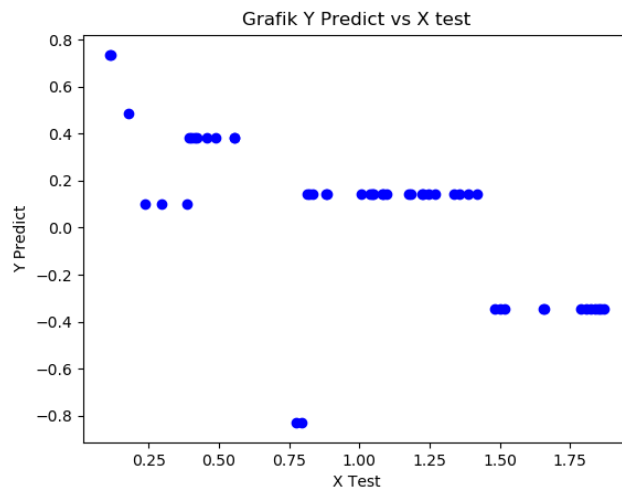
Gambar 4.1. Regresi dengan Nilai Maxdepth = 3 (Under Fit)



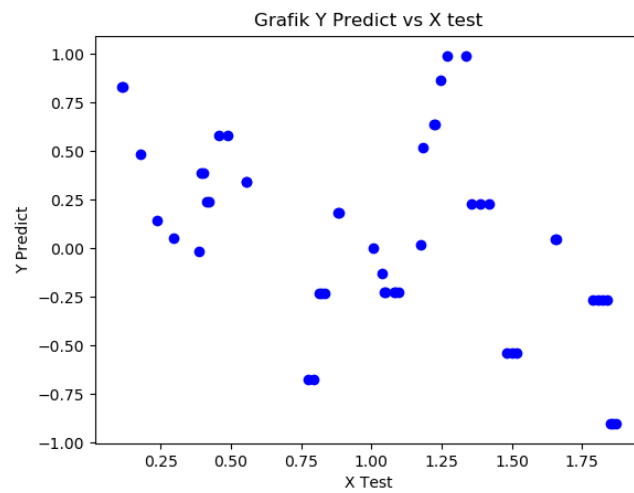
Gambar 4.2. Regresi dengan Nilai Maxdepth = 7 (Best Fit)



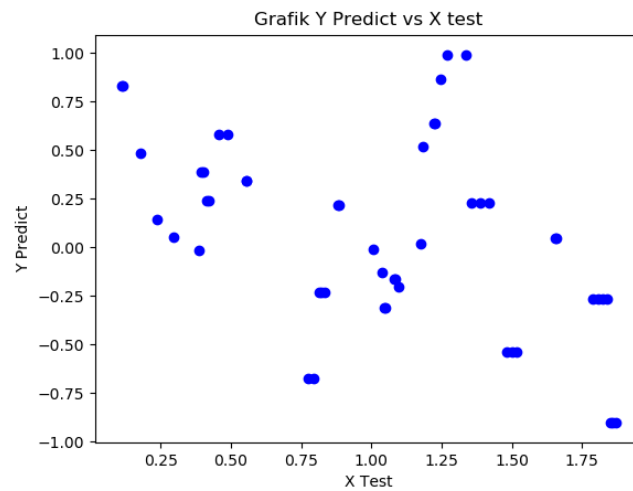
Gambar 4.3. Regresi dengan Nilai Maxdepth = 9 (Over Fit)



Gambar 4.4. Hasil Prediksi Berdasarkan Grafik Under Fit



Gambar 4.5. Hasil Prediksi Berdasarkan Grafik Best Fit

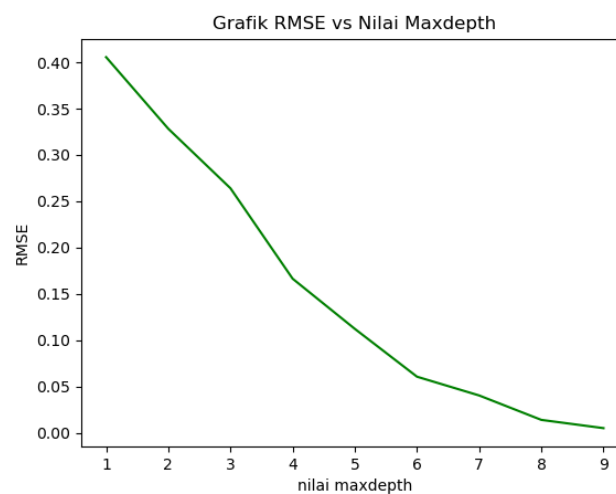


Gambar 4.6. Hasil Prediksi Berdasarkan Grafik Over Fit

4.3 Masalah-masalah yang dijumpai

Ada beberapa masalah yang ditemukan seperti syntax yang belum familiar seperti untuk menampilkan grafik, pengolahan data dan menyimpan data.

4.4 Ulasan



Gambar 4.7. Grafik RMSE terhadap nilai maxdepth

Berdasarkan hasil plot grafik diatas, nilai RMSE akan semakin kecil seiring bertambahnya orde. didapat nilai Root Mean Square Error (RMSE) untuk underfit, bestfit dan overfit berturut-turut sebesar 0,246 , 0,041 dan 0,05 .Walaupun nilai RMSE terkecil adalah grafik ke-3(overfit), namun jika dilihat dari grafikya terlalu presisi. Sehingga bestfit memiliki nilai RMSE yang 0,041.

5 Referensi

Geron, A. (2017). *Hands-On Machine Learning with Scikit-Learn & Tensor Flow*. America: O'REILLY.

MathWorks. (2019). *MathWorks*. Retrieved from MathWorks:
<https://www.mathworks.com/help/stats/understanding-support-vector-machine-regression.html>

saedsayad. (n.d.). *K Nearest Neighbors - Regression*. Retrieved from K Nearest Neighbors - Regression:
https://www.saedsayad.com/k_nearest_neighbors_reg.htm