

# UNIVERSITAS GUNADARMA



Komputer

SK No. 92 / DIKTI / Kep /1996

Fakultas Ilmu Komputer, Teknologi Industri, Ekonomi, Teknik Sipil & Perencanaan, Psikologi, Sastra

**Program Diploma (D3)** Manajemen Informatika, Teknik Komputer, Akuntansi Komputer, Manajemen Keuangan dan Pemasaran **Terakreditasi**

**Program Sarjana (S1)** Sistem Informasi, Sistem Komputer, Teknik Informatika, Teknik Elektro, Teknik Mesin, Teknik Industri, Akuntansi, Manajemen, Arsitektur, Teknik Sipil, Psikologi, Sastra Inggris **Terakreditasi**

**Program Magister (S2)** Manajemen Sistem Informasi, Manajemen, Teknik

Elektro **Program Doktor (S3)** Ilmu Ekonomi, Teknologi Informasi / Ilmu

## SOAL UJIAN TENGAH SEMESTER

Mata Kuliah	: Pengantar Sains Data	Tanggal	: 1 Juni 2022
Fakultas	: FIKTI	Waktu	: <b>90 Menit</b>
Jenjang / Jurusan	: S1 / Sistem Informasi	Dosen	: Mufni Alida
Tingkat / Kelas	: III / 3KA17, 3KA18	Sifat Ujian	: <b>Buka Buku</b>
Semester / Tahun	: ATA-2021/2022	Juml. Soal	: -

### Selesaikan Tugas di bawah ini!

1. Buka link berikut <https://bit.ly/3PsVQyJ>. Kemudian, baca dengan seksama dan ikuti semua langkah yang ada pada link tersebut.
2. Ubah dataset dengan rentang waktu dari 1 September 2021 – 30 September 2021.
3. Tentukan hasil prediksi untuk tanggal 29 September 2021!
4. Sebutkan model mana yang paling baik untuk menentukan prediksi pada tanggal 29 September 2021! Linear, Polynomial, atau RBF?
5. Jelaskan setiap langkah run yang sudah anda buat!

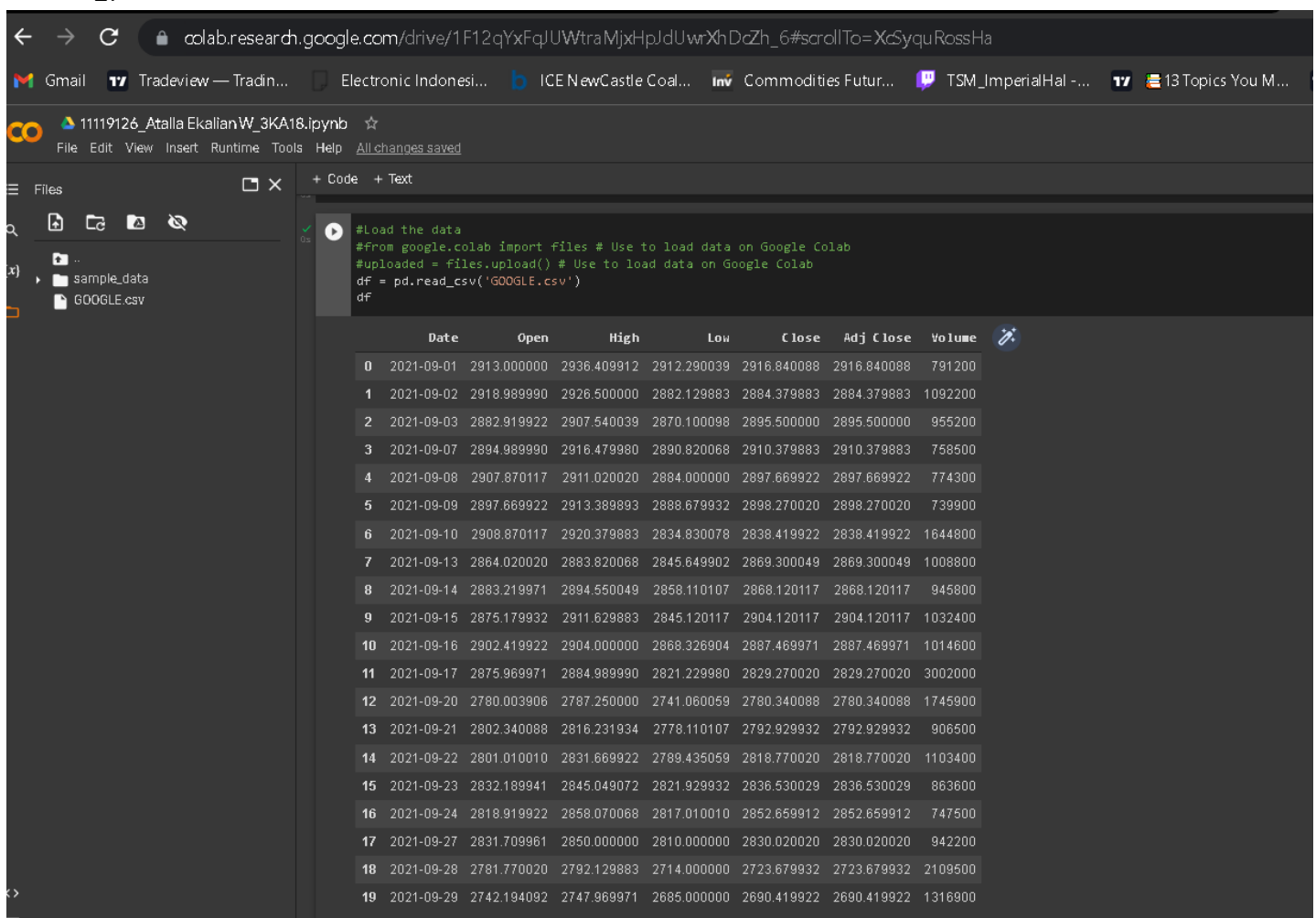
## Jawaban

Nama : Atalla Ekalian W

Kelas : 3KA18

NPM : 11119126

2.



The screenshot shows a Google Colab notebook interface. The left sidebar displays a file explorer with a folder named 'sample\_data' containing a file named 'GOOGLE.csv'. The main area shows a code cell with the following Python code:

```
#Load the data
#from google.colab import files # Use to load data on Google Colab
#uploaded = files.upload() # Use to load data on Google Colab
df = pd.read_csv('GOOGLE.csv')
df
```

Below the code cell, a table of data is displayed. The table has 8 columns: Date, Open, High, Low, Close, Adj Close, and Volume. The data rows are indexed from 0 to 19, representing dates from 2021-09-01 to 2021-09-29.

	Date	Open	High	Low	Close	Adj Close	Volume
0	2021-09-01	2913.000000	2936.409912	2912.290039	2916.840088	2916.840088	791200
1	2021-09-02	2918.989990	2926.500000	2882.129883	2884.379883	2884.379883	1092200
2	2021-09-03	2882.919922	2907.540039	2870.100098	2895.500000	2895.500000	955200
3	2021-09-07	2894.989990	2916.479980	2890.820068	2910.379883	2910.379883	758500
4	2021-09-08	2907.870117	2911.020020	2884.000000	2897.669922	2897.669922	774300
5	2021-09-09	2897.669922	2913.389893	2888.679932	2898.270020	2898.270020	739900
6	2021-09-10	2908.870117	2920.379883	2834.830078	2838.419922	2838.419922	1644800
7	2021-09-13	2864.020020	2883.820068	2845.649902	2869.300049	2869.300049	1008800
8	2021-09-14	2883.219971	2894.550049	2858.110107	2868.120117	2868.120117	945800
9	2021-09-15	2875.179932	2911.629883	2845.120117	2904.120117	2904.120117	1032400
10	2021-09-16	2902.419922	2904.000000	2868.326904	2887.469971	2887.469971	1014600
11	2021-09-17	2875.969971	2884.989990	2821.229980	2829.270020	2829.270020	3002000
12	2021-09-20	2780.003906	2787.250000	2741.060059	2780.340088	2780.340088	1745900
13	2021-09-21	2802.340088	2816.231934	2778.110107	2792.929932	2792.929932	906500
14	2021-09-22	2801.010010	2831.669922	2789.435059	2818.770020	2818.770020	1103400
15	2021-09-23	2832.189941	2845.049072	2821.929932	2836.530029	2836.530029	863600
16	2021-09-24	2818.919922	2858.070068	2817.010010	2852.659912	2852.659912	747500
17	2021-09-27	2831.709961	2850.000000	2810.000000	2830.020020	2830.020020	942200
18	2021-09-28	2781.770020	2792.129883	2714.000000	2723.679932	2723.679932	2109500
19	2021-09-29	2742.194092	2747.969971	2685.000000	2690.419922	2690.419922	1316900

Pada nomor 2 kita masukkan dulu file yang diminta dengan rentang waktu 1 september – 30 september 2021 dengan cara upload file GOOGLE.csv (tergantung nama yang dipakai). Setelah memasukkan file run pada codingan seperti diatas.

```
[14] actual_price = df.tail(1)
      actual_price
```

	Date	Open	High	Low	Close	Adj Close	Volume
19	2021-09-29	2742.194092	2747.969971	2685.0	2690.419922	2690.419922	1316900

3.

Pada bagian ini kita dapat melihat hasil prediksi saham google pada tanggal 29 september 2021

4. Model terbaik dari grafik ini adalah RBF yang merupakan model Regresi Vektor Pendukung yang menggunakan kernel yang disebut fungsi basis radial.

5.

- 1) impor paket /libraries untuk memudahkan penulisan program.
- 2) Selanjutnya memuat data saham Google (GOOG) yang saya dapatkan dari finance.yahoo.com ke dalam sebuah variabel bernama 'df' kependekan dari data frame.
- 3) Dapatkan dan cetak baris data terakhir. Harga Tutup Penyesuaian adalah 2690.419922
- 4) Siapkan data untuk pelatihan. Buat ulang bingkai data dengan mendapatkan semua data kecuali baris terakhir yang akan digunakan untuk menguji model nanti, dan simpan data baru dengan baris terakhir hilang kembali ke 'df'.
- 5) Buat variabel yang akan digunakan sebagai kumpulan data independen dan dependen dengan menyetelnya sama dengan daftar kosong.
- 6) Dapatkan semua baris dari kolom Tanggal simpan ke dalam variabel yang disebut 'df\_days' dan dapatkan semua baris dari kolom Harga Tutup Penyesuaian dan simpan datanya ke dalam variabel.
- 7) Buat kumpulan data independen 'X' dan simpan data dalam variabel 'hari'.
- 8) Buat kumpulan data dependen 'y' dan simpan data dalam variabel 'adj\_close\_prices'. Keduanya dapat dilakukan dengan menambahkan data ke masing-masing daftar.
- 9) Lihat hari apa yang dicatat dalam kumpulan data.
- 10) Selanjutnya, membuat dan melatih 3 model Support Vector Regression (SVR) yang berbeda dengan tiga kernel berbeda untuk melihat mana yang berkinerja terbaik.
- 11) memplot model pada grafik untuk melihat mana yang paling cocok dan mengembalikan prediksi hari itu.
- 12) Model terbaik dari grafik ini tampaknya adalah RBF yang merupakan model Regresi Vektor Pendukung yang menggunakan kernel yang disebut fungsi basis radial. Namun grafik ini bisa menyesatkan.
- 13) Sekarang saya bisa mulai membuat prediksi harga saham saya. Mengingat baris terakhir data yang tertinggal dari kumpulan data asli, tanggalnya adalah 30-09-2021, jadi harinya adalah 30.
- 14) Jadi sekarang akan memprediksi harga yang memberi model nilai atau hari ke-30.
- 15) Dari pengujian kecil ini, model yang tampaknya memiliki performa terbaik tampaknya adalah model RBF SVR. Model ini memperkirakan nilai \$2692.62407997 ketika harga sebenarnya adalah \$2690.419922, jadi hanya naik sekitar \$2 !

