**NAMA : Wandi Yusuf Kurniawan**

**NIM : 1301218601**

**KELAS : IFX-45-GAB**

**Perhatikan pemrograman SVM berikut ini:**

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Pertanyaan 1 (point 20):  
Apa yang akan dikeluarkan setelah baris ini di run?

Chart, scatter chart

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application, Word

Description automatically generated

Pertanyaan-2 (point 20):  
Apa output setelah baris ini di run?

Chart, scatter chart

Description automatically generated

Graphical user interface, application

Description automatically generated

Pertanyaan-3 (point 20):  
Apa output setelah baris ini di run? Jelaskan!

Chart, scatter chart

Description automatically generated

Penjelasan: Data pada X\_test di-plot sesuai dengan koordinat dan diberi warna (label) berdasarkan data pada y\_test. Garis pemisah merupakan model yang dibentuk dari hasil data latih X\_train.

Graphical user interface, application

Description automatically generated

Pertanyaan-4 (point 20):  
Jelaskan bagaimana akurasi training atau testing ini 100%!

Data pada X\_train dilatih untuk membentuk model berupa garis pemisah dengan margin terjauh antar-label yang terdekat (label -1 atau merah dengan koordinat (-0.3, -0.1) dan label 1 atau biru dengan koordinat (0.0, 0.3) pada output soal no.2), kemudian X\_train diberikan warna (label) sesuai dengan data pada y\_train. Kemudian model tersebut diuji menggunakan data X\_test dan label pada y\_test. Ternyata, output dari soal 3 menunjukkan pemisahan label yang benar-benar sesuai dengan garis pemisah. Hal tersebut didukung dengan hasil y\_pred\_train dan y\_pred\_test memiliki akurasi sempurna karena kedua array tersebut benar-benar sesuai dengan label yang diberikan oleh y\_train dan y\_test.

Pertanyaan-5 (point 20):  
Sebutkan contoh penggunaan algoritma SVM yang anda ketahui! Jelaskan dengan cara menjabarkan cara kerja algoritma SVM.

Contohnya pada klasifikasi artikel pada berita

1. Sebelum diproses menggunakan SVM, data mentah berupa artikel atau corpus di-prepocessing terlebih dahulu (record yang bolong dihapus, ubah semua kata menjadi huruf kecil kemudian ditokenisasi/dipisah-pisah, hapus stop words seperti kata sambung atau kata ganti, hapus kata yang mengandung angka atau symbol, dan lemmatisasi kata, seperti mengambil kata beri dari memberi).
2. Siapkan variable X\_train (data latih prediktor), X\_test (data uji prediktor), y\_train (data latih target), dan y\_test (data latih target) untuk memisahkan data yang sudah di-preprocessing dengan train\_test\_split. Untuk kasus ini data X (predictor) adalah kata-kata yang sudah disaring dari artikel, dan data y (target) adalah label\_1 atau label\_2 dengan ukuran data uji sebesar 20%.
3. Label pada variable y\_train dan y\_test di hot encode terlebih dahulu untuk memudahkan klasifikasi. Misalnya label\_1 diberi encode 0 dan label\_2 diberi encode 1 dengan .fit\_transform().
4. Gunakan algortima TF-IDF (Term Frequency – Inverse Document Frequency) untuk mengonversi data predictor (kata-kata) menjadi vector dengan .fit() kata-kata saringan (gabungan X\_train dan X\_test) kemudian .transform() setiap X\_train dan X\_test untuk mendapatkan nilai yang bisa digunakan untuk klasifikasi SVM. Nilai tersebut masing-masing dimasukkan ke dalam variable X\_train\_tfidf dan X\_test\_tfidf.
5. Inisialisasi SVM dengan kernel ‘linear’ untuk membuat garis pemisah, kemudian .fit() X\_train\_tfidf dan y\_train. Algoritma tersebut akan mencari jarak terjauh antar label yang terdekat dengan garis pemisah untuk menentukan nilai kemiringan/bobot dan bias (y=wx+b).
6. Setelah model dibuat, .predict() data X\_test\_tfidf, kemudian hitung akurasi dengan y\_test. Semakin tinggi nilai akurasinya, semakin baik modelnya dan direkomendasikan untuk digunakan dalam klasifikasi artikel pada berita (misalnya ‘hiburan’ atau ‘teknologi’).