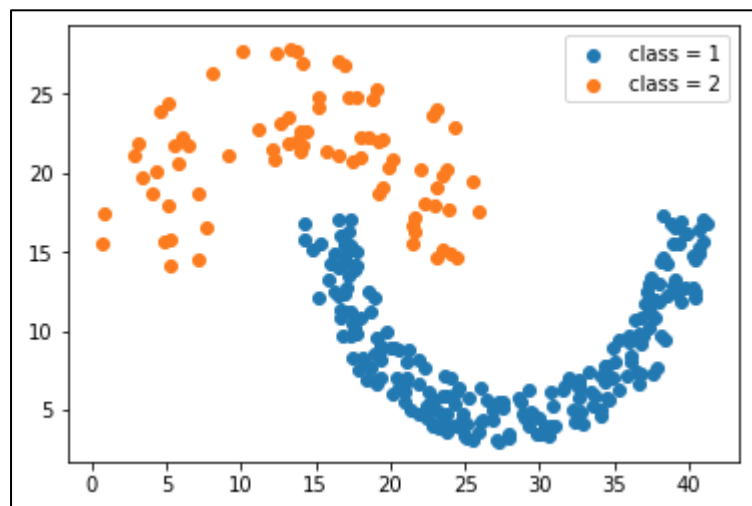


**Soal!**

Diberikan sebuah Trainset berupa himpunan data berisi 298 objek data yang memiliki 2 atribut input (X1 dan X2) dan 1 output (Class) yang memiliki dua label (bernilai 1 atau 2). Bangunlah sebuah sistem klasifikasi menggunakan metode metode Bagging, salah satu teknik Ensemble Learning, berbasis Naïve Bayes untuk menentukan kelas/label data uji dalam Testset. Sistem membaca masukan file TrainsetTugas4ML.csv dan TestsetTugas4ML.csv dan mengeluarkan output berupa file TebakanTugas4ML.csv berupa satu kolom berisi 75 baris yang menyatakan kelas/label baris yang bersesuaian pada file TestsetTugas4ML.csv.

**Analisa :**

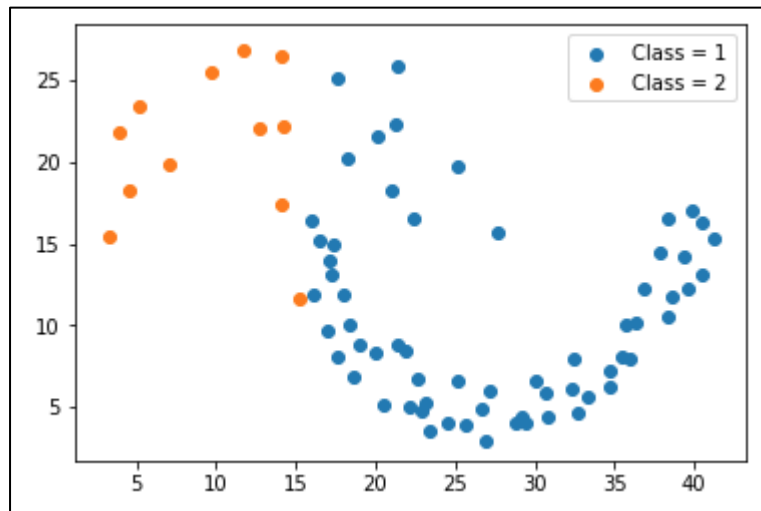
- Terdapat 298 data Trainset, 220 diantaranya memiliki kelas 1, dan 78 lainnya adalah kelas 2
- Nilai terkecil dari X1 pada Trainset adalah 0.75, dan Nilai terbesarnya adalah 41.25
- Nilai terkecil dari X2 pada Trainset adalah 3, dan Nilai terbesarnya adalah 27.85
- Nilai X1 dan X2, dapat divisualisasikan sebagai berikut



- Didalam Testset terdapat 75 data, tidak memiliki kelas
- Nilai terkecil dari X1 pada Testset adalah 3.3, dan Nilai terbesarnya adalah 41.3
- Nilai terkecil dari X2 pada Testset adalah 2.95, dan Nilai terbesarnya adalah 26.85
- Naïve Bayes yang digunakan dalam Ensemble Learning ini adalah Gaussian Naïve Bayes
- Banyak bagging yang dilakukan adalah 25
- Data random yang diambil dari Trainset dalam setiap iterasi bagging untuk inisialisasi Model Gaussian Naïve Bayes adalah 50 data
- Outputkan hasil berupa 1 baris data berisi kelas kedalam .csv

**Penyelesaian Masalah :**

1. Ambil data yang akan digunakan dalam proses pembelajaran dan pengujian
2. Inisialisasi jumlah bagging yang akan dilakukan
3. Selama proses bagging dilakukan akan diinisialisasi 1 model yang kemudian akan dihitung akurasi terhadap datanya sendiri
4. Kemudian setelah semua model diinisialisasi (proses bagging selesai), lakukan prediksi kelas kepada seluruh Trainset untuk mendapatkan akurasi dari pembelajaran mesin terhadap Trainset
5. Kemudian lakukan prediksi kelas ke data Testset
6. Hasil Prediksi kelas disimpan kedalam .csv



*Figure 1 Visualisasi hasil prediksi kelas ke Testset*