**Nama : Dedih Kurnia**

**Nim : 1301164077**

**Kelas : IF 40-11**

1. **Algoritma Naive Bayes**

Algoritma Naive Bayes adalah salah satu algoritma yang digunakan untuk klasifikasi atau pengelompokan data. Contoh yang dibahas kali ini adalah mengelompokan sebuah data yang memiliki besar angka yang berbeda. Algoritma ini dapat melakukan proses perhitungan baik data numerik, teks ataupun data yang sudah terkategori.

1. **Analisis Masalah**

Dalam laporan ini *Algoritma Naive Bayes* digunakan untuk menganalisis 40 data dan 160 data yang di mana 40 data adalah data test dan 160 data adalah data train. Data Train adalah sebuah data yang valid atau bisa di sebut dengan data yang sudah di kategorikan sehingga data ini menjadi sebuah data acuan untuk menentukan sebuah kategori data Test untuk nantinya di kelompokan sesuai dengan kategori data Train. Logika dalam kasus ini seperti ini : menentukan apakah golongan tersebut nantinya akan masuk kedalam golongan data Besar atau data Kecil tergantung dari data Train(data yang sudah di kategorikan/data valid). Bahasa pemrograman Python digunakan pada perancangan ini untuk mencari nilai akurasi data yang dimana nantinya kita tahu data tersebut dominan ke data yang mana.

1. **Strategi Pengujian**

Menentukan data yang akan di ujidan kali ini saya memberikan data :

* **Data Test** : 40 data baru (40 data yang di simpan melalui file excel dimana nanti program akan mengload data dari data excel tersebut).
* **Data Train** : 160 data asli (160 data yang di simpan melalui file excel dimana nanti program akan mengload data dari data excel tersebut).
* **Algoritma Function**  : pada algoritma tersebut saya menggunakan 4 fungsi/function :

1. *Function BacaFile :* fungsi ini berfungsi untuk membaca file “TrainsetTugasML.csv” yang dimana nanti data ini sebagai acuan data Test untuk menentukan kategori sebuah data apakah lebih dari 50 Ribu atau kurang.
2. *Function BacaTest :* fungsi ini berfungsi untuk membaca file “TestsetTugasML.csv” yang dimana nanti di gunakan untuk data yang akan di bandingkan dengan data Train.
3. *Function BandingData :* fungsi ini berfungsi untuk membandingkan sebuah data 1 dengan data lainnya yang dimana data tersebut akan di simpan dan di pisahkan sesuai dengan perbandingan.
4. *Function Peluang :* fungsi ini berfungsi untuk mencari nilai yang lebih 50 dan banyaknya. Dan mencari nilai yang kurang 50 dan banyaknya.