

**LAPORAN PROYEK AKHIR
PRAKTIKUM DATA SCIENCE**

**Analisis Sentimen Review Aplikasi Bibit dengan
Metode Naïve Bayes**



Muhammad Ikhwan Hanif	123200096
Taufik Sahid Fadhil	123200028

**PROGRAM STUDI INFORMATIKA
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK INDUSTRI
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
YOGYAKARTA
2021**

1. PENDAHULUAN

Pada era teknologi yang semakin meningkat dan yang semakin modern berpengaruh terhadap kehidupan individu maupun kelompok. Perkembangan TI dapat membantu dalam mengembangkan tugas-tugas baru pada perusahaan berskala pasar global atau pada instansi pemerintah, TI dimanfaatkan untuk mengatasi meningkatkan kemampuan seseorang dalam meraih keunggulan dan kesuksesan yang handal serta canggih.

Pemanfaatan teknologi informasi juga digunakan dalam aplikasi Bibit yang dapat digunakan disegala aspek dengan menggunakan data yang telah tersedia, data ini akan diolah dengan teknologi yang akan menghasilkan informasi transparan, akurat dan mudah dipahami oleh semua kalangan.

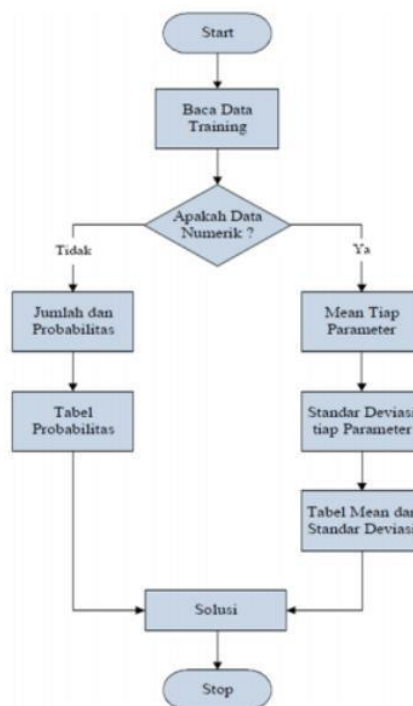
Salah satu kegunaan teknologi informasi dalam museum yaitu untuk mengetahui penilaian dari pengunjung setelah berkunjung ke museum. Analisis tersebut biasa disebut sebagai Analisa Sentimen atau Sentiment Analysis ini adalah proses penggunaan text analisis untuk mendapatkan berbagai sumber data dari internet dan beragam platform media sosial. Tujuannya adalah untuk memperoleh opini dari pengguna yang terdapat pada platform tersebut.

Dengan menggunakan teknologi dan sekumpulan data yang ada, dapat menghasilkan informasi yang akurat juga mudah dipahami tanpa perlu waktu yang lama. Menggunakan metode Naive Bayes Classifier untuk mengolah datanya dan Bahasa R yang digunakan untuk membuat aplikasinya

2. METODE

Data yang kita dapat berasal dari scrapping Aplikasi Bibit di Google Play dengan bahasa python untuk mendapatkan dataset tentang Aplikasi Bibit. Metode yang kami gunakan yaitu Naïve Bayes Classifier merupakan merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas.

Naïve Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu. Keuntungan penggunaan Naïve Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naïve Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan Naïve Bayes Classifier dinilai bekerja sangat baik dibanding dengan model classifier lainnya, yaitu Naïve Bayes Classifier memiliki tingkat akurasi yg lebih baik dibanding model classifier lainnya.

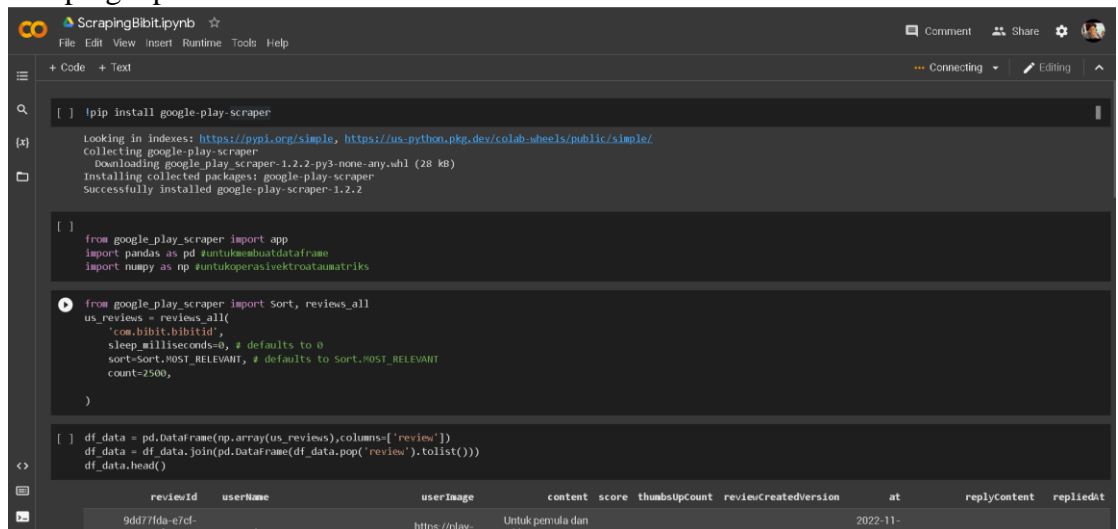


Keterangan:

- Membaca data training.
- Menghitung Jumlah dan Probabilitas dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
- Mendapatkan nilai dalam table mean, Standart Deviasi dan Probabilitas .
- Menghasilkan solusi

3. HASIL DAN PEMBAHASAN

- Scraping Aplikasi Bibit



```
[ ] | pip install google-play-scraper

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting google-play-scraper
  Downloading google_play_scraper-1.2.2-py3-none-any.whl (28 kB)
Installing collected packages: google-play-scraper
Successfully installed google-play-scraper-1.2.2

[ ] | from google_play_scraper import app
import pandas as pd #untukmembuatdataframe
import numpy as np #untukoperasiaktrotaumatriks

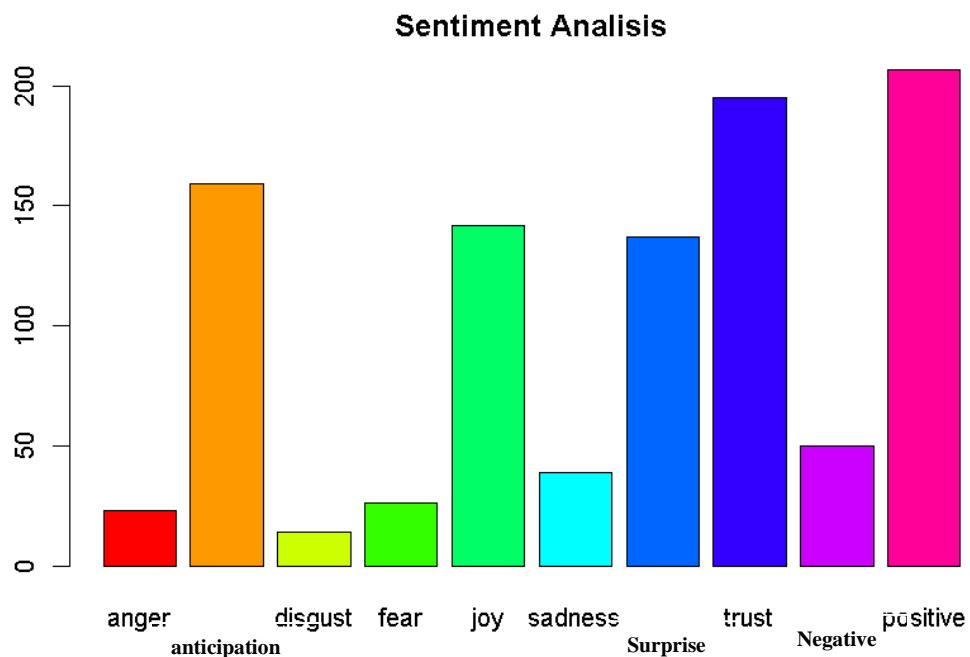
[ ] | from google_play_scraper import Sort, reviews_all
us_reviews = reviews_all(
    'com.bibit.bibitid',
    sleep_allseconds=0, # defaults to 0
    sort=Sort.MOST_RELEVANT, # defaults to Sort.MOST_RELEVANT
    count=2500,
)

[ ] | df_data = pd.DataFrame(np.array(us_reviews), columns=['review'])
df_data = df_data.join(pd.DataFrame(df_data.pop('review').tolist()))
df_data.head()
```

reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt
9dd77fde-e7ef-			https://play-	Untuk pemuda dan			2022-11-		

Pada gambar diatas terdapat proses scraping atau pengambilan data dari review atau ulasan pengguna aplikasi Bibit.

- BarPlot Sentimen Analisis dengan Naïve Bayes



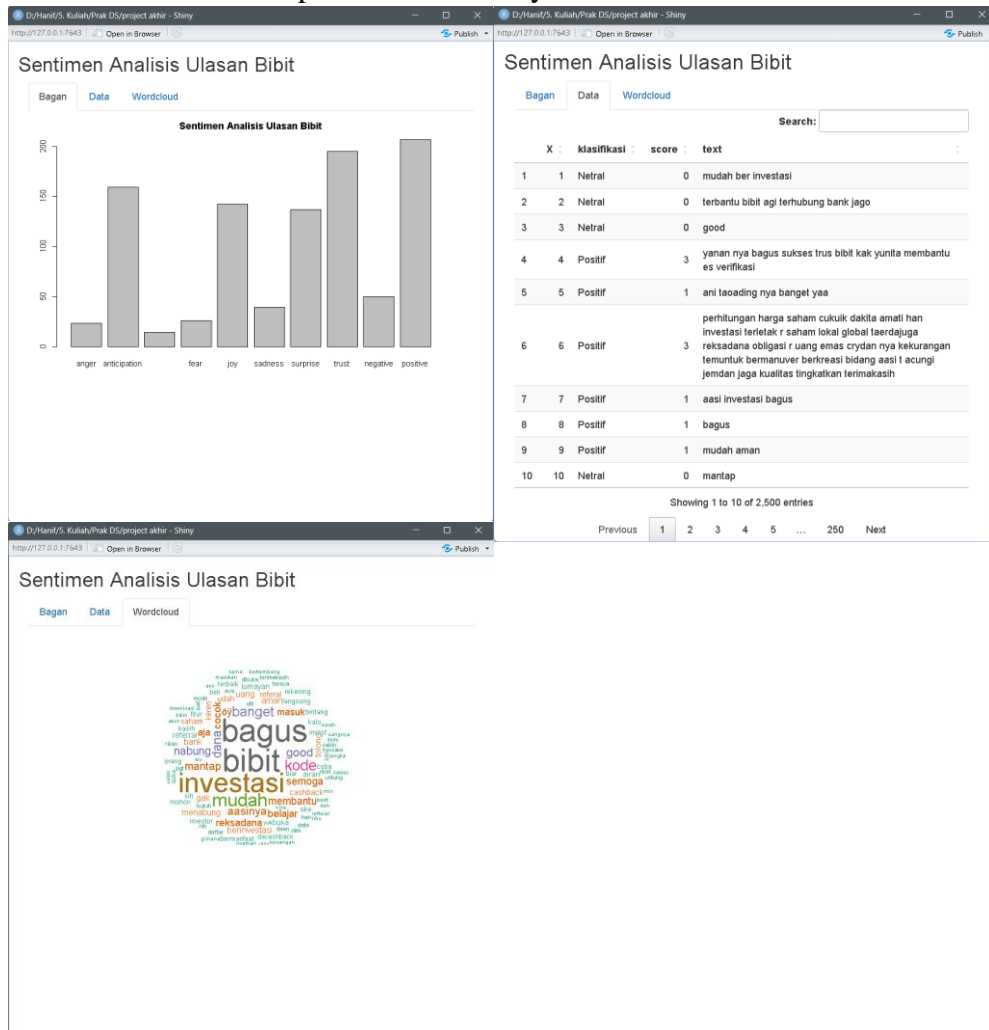
Pada gambar terakhir yaitu terdapat diagram yang berisikan tentang jumlah komentar positif dan negatif yang paling umum.

- WordCloud



Pada gambar ini yaitu terdapat kata-kata yang sering muncul atau yang paling sering disebutkan oleh para pengunjung.

- Visualisasi dengan Shiny
Hasil visualisasi ditampilkan dalam Shiny



4. KESIMPULAN

Dengan adanya Sentiment Analysis ini, pihak developer Bibit dapat mengetahui penilaian dari para pengguna aplikasi Bibit. Dari data yang diperoleh oleh praktikan menyimpulkan bahwa Sebagian besar menilai positif setelah atau saat menggunakan aplikasi Bibit dan ada juga yang merasa kurang puas. Dengan adanya program ini semoga dapat lebih berkembang dan memudahkan developer Bibit.