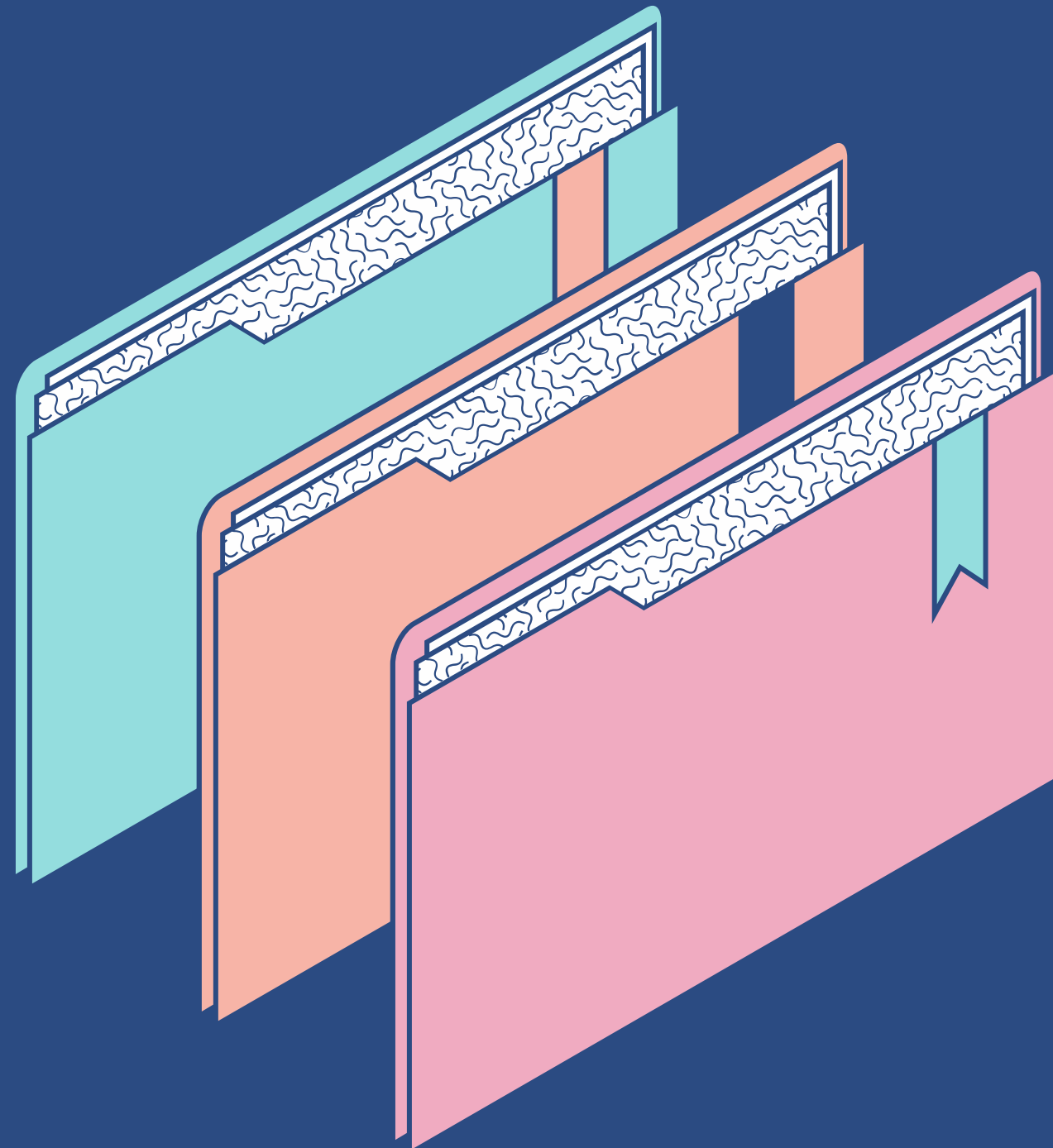




PROPOSAL TUGAS AKHIR

# Analisis Sentimen Terhadap Penurunan Harga Telur pada Data Twitter Menggunakan Metode Support Vector Machine dan K-Nearest Neighbour

Arviandri Naufal Zaki - 064001800035

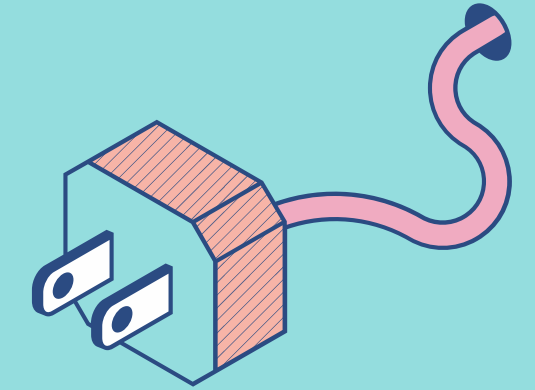


## Latar Belakang

Twitter oleh masyarakat Indonesia dimanfaatkan untuk berbagai hal seperti berkomunikasi dengan orang lain secara publik atau personal, berbagi kabar dan opini pribadi, berjualan, sampai mengkritik suatu hal.

Oleh karena itu, pengguna Twitter dapat beropini yang dipengaruhi oleh emosi yang dapat diklasifikasikan untuk menentukan polarisasinya.

# Rumusan Masalah

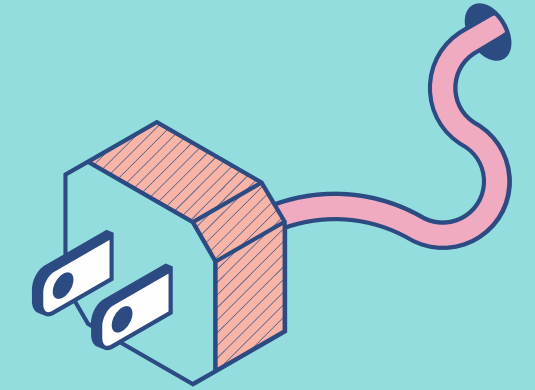


Bagaimana cara mengambil dan mengolah data tweet yang berasal dari Twitter untuk perhitungan Support Vector Machine (SVM) dan K-Nearest Neighbour (KNN)

Bagaimana tingkat keakuratan dari K-Nearest Neighbour (KNN) dan Support Vector Machine (SVM) pada analisis sentimen di Twitter mengenai penurunan harga telur.

Bagaimana hasil klasifikasi dari tweet menggunakan Support Vector Machine (SVM) dan K-Nearest Neighbour (KNN).

# Batasan Masalah



Data yang digunakan adalah tweet berbahasa Indonesia dengan kata kunci “Harga Telur Turun” dan “Harga Telor Turun” dari Twitter.

Metode yang digunakan untuk klasifikasi adalah Support Vector Machine (SVM) dan K-Nearest Neighbour (KNN).



# Tujuan Penelitian

Tujuan dari tugas akhir ini yaitu untuk mengklasifikasi tweet berdasarkan positif dan negatifnya untuk mengetahui keakuratan dari kedua metode ini yaitu Support Vector Machine (SVM) dan K-Nearest Neighbour (KNN) dalam menganalisis sentimen (emosi) pengguna Twitter mengenai penurunan harga telur.

# Manfaat Penelitian

Memperoleh hasil analisis sentimen terhadap penurunan harga telur dengan menggunakan metode SVM dan KNN.

Bagi pemerintah dapat digunakan untuk meningkatkan pengetahuan dan dapat digunakan sebagai rujukan untuk memperbaharui kebijakan yang dikeluarkan.

# Penelitian Sebelumnya



No	Nama Penulis - Tahun	Judul	Hasil Kajian
1	Nabilah Putri Aprilia, Dian Pratiwi, and Anung Barlianto Ariwibowo - 2019	Sentiment Visualization Of Covid-19 Vaccine Based On Naïve Bayes Analysis	Pada penelitian ini peneliti melakukan panggilan API ke twitter untuk mendapatkan data,Kemudiandari hasil dari data yang diambil tersebut dilakukan preprosesing pada data tersebut dan dilanjutkan denganlabelinguntuk menentukan positif dan negatifnya lalu dilakukan ekstraksi data menggunakan TF-IDF untuk pembobotan kata, lalu dilakukan perhitungan menggunakan Naïve Bayes dan mendapatkan nilai akurasi 73,1% ,presisi 73 %, dan recall 83%
2	Ghulam Asrofi Buntoro - 2017	Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter	Dari penelitian ini didapatkan hasil yang cukup tinggi untuk klasifikasi menggunakan Naive Bayes dibandingkan menggunakan SVM yaitu sekitar 95% nilai akurasi, 95% nilai presisi, dan 95% nilai recall.

# Landasan Teori

## Twitter

Twitter adalah platform sosial media yang dapat digunakan untuk mengirimkan suatu postingan (tweet) dalam bentuk foto maupun teks dengan terbatas yaitu 280 karakter

## Python

Python adalah bahasa pemrograman dengan kode sumber yang terbuka (open source) yang dapat digunakan untuk membuat program secara independent (standalone) maupun untuk membuat program scripting.

## Scraping Data

Teknik Scraping menggunakan cara mengambil data dari apa yang ditampilkan oleh website. Pada tahap ini dilakukan penarikan data menggunakan library snsrape.

## Preprocessing

Tujuan dilakukannya preprocessing dokumen adalah untuk menghilangkan suatu hal yang dapat mengganggu jalannya analisis, menyeragamkan bentuk kata dan mengurangi volume kata.

## Lexicon

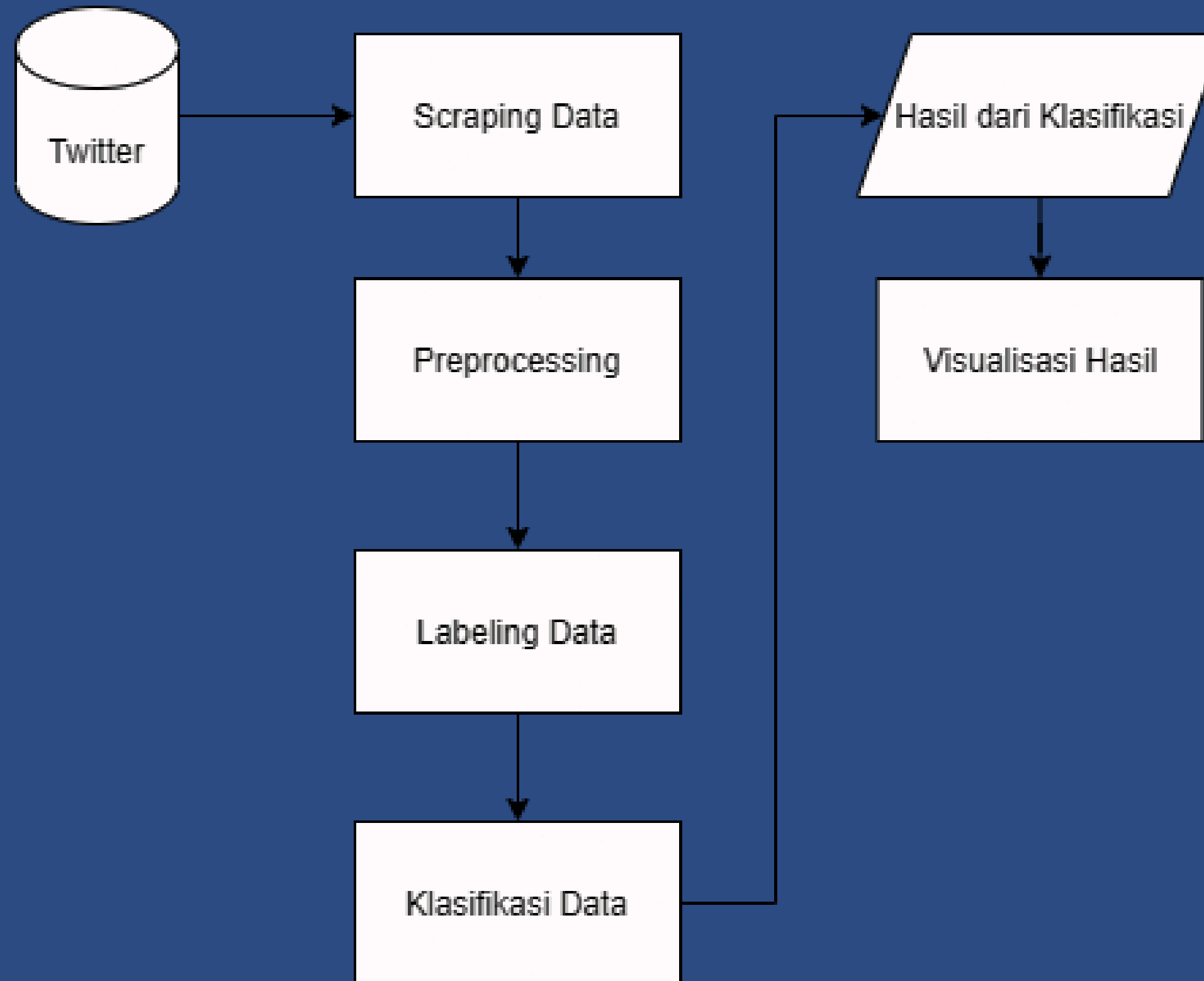
Lexicon merupakan kumpulan kata pada sentimen yang telah diketahui dan dihimpun dalam bentuk dataset

## TF-IDF

TF-IDF merupakan suatu algoritma yang dapat menghasilkan informasi tentang seberapa sering kata tersebut muncul di dalam dataset tersebut dan dimunculkan dalam bentuk berat per kata.



# Metode Penelitian



# Metode Penelitian

## Pengumpulan Data

Pengumpulan data menggunakan teknik Scrapingdata dari Twitter. Data yang dikumpulkan berupa tweet dengan kata kunci "Harga Telur Turun" dan "Harga Telor Turun" dalam rentang waktu 29 Juli 2021 hingga 15 Oktober 2021 dan tidak disertakan posting retweet.

## Pengolahan Data

Proses ini akan mengolah data awal yang masih tidak beraturan untuk dijadikan data teratur yang dapat diterapkan pada proses selanjutnya.

## Labeling Data

Labeling pada data dilakukan secara otomatis menggunakan kamus yang sudah berisi bobot sentimen (lexicon) dan dihitung total dari sentimen berdasarkan jumlah bobot dari seluruh kata pada setiap data.

## Pembobotan Kata

Setelah di berikan label selanjutnya dilakukan pembobotan kata. Pembobotan kata dilakukan dengan menggunakan Term Frequency (TF) dan Inverse Document Frequency (IDF).

## Mengklasifikasikan Data

Proses ini bertujuan untuk mengolah data menjadi opini positif dan opini negatif. Pada penelitian ini, penulis menggunakan metode Support Vector Machine dan K-Nearest Neighbour untuk mengklasifikasikan data.

## Visualisasi

Pada proses ini akan dilakukan visualisasi terhadap data yang dihasilkan dari proses klasifikasi. Tujuan dari proses ini untuk mempermudah membaca maksud dan informasi dari hasil analisis.

# Terima Kasih

