

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Penelitian Terdahulu

Terdapat penelitian serupa yang telah dilakukan pada Analisis sentimen sosial media pada khususnya twitter, di antaranya adalah penelitian tentang Analisis sentimen pada hashtag tertentu. Sehingga dalam melakukan penelitian, penulis akan mengidentifikasi metode yang pernah dilakukan, meneruskan penelitian sebelumnya, serta mengetahui orang lain yang spesialisasi dan area penelitiannya sama dibidang ini.

Penulis memperoleh beberapa studi pustaka diantaranya dari jurnal, tesis, buku dan juga literatur yang lainnya. Beberapa tinjauan literatur tersebut adalah sebagai berikut :

2.1.1 Hashtag #2019GantiPresiden: Sentimen Anti-petahana dan Orientasi Politik Pemilih Pemula menghadapi Pilpres 2019 di Pematangsiantar

Penelitian dilakukan oleh Eronid Litno Damanik pada “Jurnal Ilmu Pemerintahan dan sosial Politik UMA” yang telah dipublikasikan pada bulan Desember 2018. Penelitian dilakukan untuk memahami munculnya hashtag #2019GantiPresiden dan orientasi politik pemilih pemula menghadapi Pemilihan Presiden 2019.

Hashtag #2019GantiPresiden disebarakan melalui media sosial, kaos, stiker, maupun deklarasi hashtag #2019GantiPresiden di berbagai daerah. Hashtag #2019GantiPresiden merupakan penggiringan opini, sentimen Anti-petahana dan dekontruksi *success story* presiden petahana pada Pemilihan Presiden 2019 [1].

2.1.2 Sentiment based Analysis of Tweets during the US Presidential Elections

Penelitian ini dilakukan oleh Usama Yaqub, Soon Ae Chun, Vijayalakshmi Atluri, dan Jaideep Vaidya pada “*18th Annual International Conference on Digital Government Research*” yang dilakukan pada 07-09 Juni 2017 di New York, Amerika Serikat.

Penelitian ini bertujuan untuk melakukan analisis berdasarkan sentimen dari data *Twitter* yang dikumpulkan sebelum dan sesudah hari pemilihan dan mengetahui topik diskusi terpopuler selama pemilihan presiden Amerika Serikat berlangsung [2].

2.1.3 Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis

Penelitian ini dilakukan oleh Widodo Budiarto dan Meiliana pada “*Journal of Big Data*”

yang dipublikasi pada 19 Desember 2018. Penelitian dilakukan ini menggunakan tweet dari Calon Presiden Indonesia (Jokowi dan Prabowo), dan *tweet* dari tagar yang relevan untuk analisis sentimen yang dikumpulkan dari Maret hingga Juli 2018 untuk memprediksi hasil pemilihan Presiden Indonesia.

Para penulis membuat algoritma dan metode untuk menghitung data penting, kata-kata teratas dan melatih model dan memprediksi polaritas sentimen. Hasil eksperimen dihasilkan dengan menggunakan bahasa *R* [2].

2.2 Teori Dasar Yang Digunakan

A. Pemilihan Umum (PEMILU)

Indonesia salah satu negara yang menganut sistem demokrasi. Hal ini ditandai dengan diadakannya suatu pemilihan umum terhadap presiden dan wakil presiden. Pemilihan umum pada suatu negara yang menganut demokrasi bisaanya

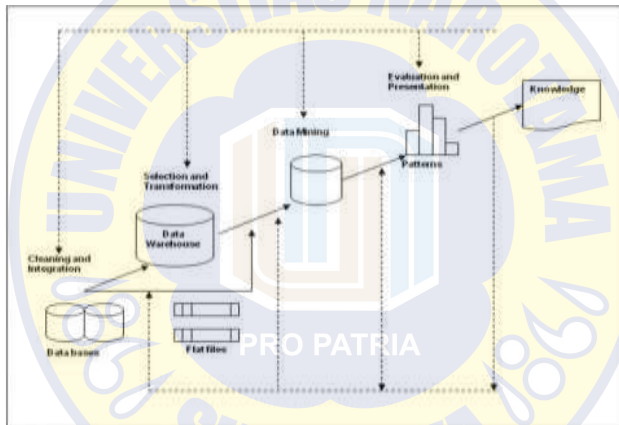
diselenggarakan secara periodik. Pada tahun 2019 akan dilaksanakan pemilihan umum presiden dan wakil presiden [3].

Seorang tokoh politik yang ingin maju sebagai calon presiden tentu akan melihat atau mempertimbangkan popularitas mereka berdasarkan opini dari masyarakat. Dahulu masyarakat mengungkapkan opini, kritik, dan sarannya melalui media cetak yang tidak semua orang mempunyai kemampuan menulis dan kesempatan menerbitkan tulisannya.

Namun, perkembangan teknologi komunikasi saat ini telah merubah kecenderungan kebiasaan masyarakat dalam mengekspresikan opininya pada jejaring sosial. Salah satu jejaring sosial yang populer di kalangan pengguna internet saat ini adalah Twitter.

B. Data Mining

Data mining merupakan kegiatan mengekstraksi atau menambang pengetahuan dari data yang berukuran/berjumlah besar, informasi inilah yang nantinya sangat berguna untuk pengembangan [4].



Gambar 1. Langkah-langkah untuk melakukan *data mining*

1. *Data cleaning* (untuk menghilangkan data yang tidak konsisten) dan *Data integration* (di mana sumber data yang terpecah dapat disatukan).

2. *Data selection* (di mana data yang relevan dengan tugas analisis dikembalikan ke dalam database).
3. *Data transformation* (di mana data berubah atau bersatu menjadi bentuk yang tepat untuk menambang dengan ringkasan performa atau operasi agresi).
4. *Data mining* (proses esensial di mana metode yang intelejen digunakan untuk mengekstrak pola data).
5. *Pattern evolution* (untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan berdasarkan atas beberapa tindakan yang menarik).
6. *Knowledge presentation* (di mana gambaran teknik visualisasi dan pengetahuan digunakan untuk memberikan pengetahuan yang telah ditambang kepada pengguna).

Jenis data dalam *data mining*

1. *Relation Database* : Sebuah sistem *database*, atau disebut juga *database management system* (DBMSnS), mengandung sekumpulan data yang saling berhubungan, dikenal sebagai sebuah *database*, dan satu set program perangkat lunak untuk mengatur dan mengakses data tersebut.
2. *Data Warehouse* : sebuah ruang penyimpanan informasi yang terkumpul dari beraneka macam sumber, disimpan dalam skema yang menyatu, dan biasanya terletak pada sebuah situs.

C. *Twitter API*

Untuk dapat mengakses dan mengembangkan data di *Twitter* harus mempunyai *Twitter API* (*Application Program Interface*) yang merupakan aplikasi yang diciptakan oleh Pihak pengembang *Twitter* untuk lebih mudah mengakses, mengambil dan menerima data tweet

yang ada pada database pusat *Twitter* yang telah terstruktur sehingga dapat dimanfaatkan sebagai bahan pengembangan dan analisis data . Untuk bisa mengakses *Twitter API*, Pengguna mengakses websitenya terlebih dahulu di <https://dev.twitter.com/apps> sebelum membuat permintaan Autentikasi API di *Twitter* [5].

D. *Trending Topic*

Trending Topic atau tren terkini di *Twitter* merupakan fitur yang menampilkan beberapa hashtag berisi topik yang sedang trend saat ini. *Twitter* menggunakan algoritma TF-IDF (Term frequency–inverse document frequency) yaitu menjumlahkan kata atau hashtag yang diposting di hari dan waktu yang sama secara bersamaan [6].

E. Bahasa Pemrograman *R*

R merupakan bahasa pemrograman untuk komputasi statistikal, manipulasi, dan analisis [7]. Untuk menjalankan *R* pengguna menginstal

software *R* terlebih dahulu dengan mengunduh di website <https://www.r-project.org/> sesuai dengan Sistem Operasi Komputer yang digunakan.

F. Analisis Sentimen

Analisis Sentimen merupakan riset komputasional dari opini sentimen dan emosi yang diekspresikan dalam bentuk tekstual [8]. Analisis Sentimen telah banyak dilakukan seperti mengetahui respons opini konsumen pada suatu produk, preferensi politik, pendapat tentang film yang telah dirilis beserta prediksi pendapatan film dan lainnya [9].