

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pertumbuhan *social media* di Indonesia beberapa tahun terakhir ini sangat tinggi mengingat internet adalah kebutuhan penting dalam mendapatkan dan berbagi informasi. Dengan kata lain, Internet adalah media penyimpan informasi terbesar di dunia yang terbagi menjadi banyak bentuk seperti artikel, postingan, multimedia, komentar, dan *review*. Hampir semua masyarakat lebih aktif di media sosial daripada dunia nyata dalam menyampaikan opini dan buah pikiran mereka tentang suatu masalah tertentu[2]. Dengan berkembangnya internet, masyarakat sudah menemukan cara baru dalam mendapatkan informasi ataupun memberikan informasi. Opini masyarakat tentang suatu masalah yang berlangsung merupakan suatu informasi yang berharga. Sebagai contoh dalam dunia bisnis, masyarakat sudah tidak lagi membutuhkan toko untuk berjualan, mereka dengan mudahnya memanfaatkan media sosial untuk mempromosikan produk mereka. Dalam dunia politik juga, para politisi lebih mudah mengenalkan diri mereka serta tindakan mereka melalui media sosial[2]. *Sentiment Analysis* bertujuan untuk mengidentifikasi dan mengklasifikasikan opini atau sentimen pada *social media* seperti *Facebook*, *Twitter* dan lain-lain yang diekspresikan dalam bentuk *text*. Dengan begitu kita bisa memahami pola pikir dan opini mereka dengan mengklasifikasikannya ke dalam kelas-kelas tertentu.

Learning Vector Quantization (LVQ) adalah sebuah metode klasifikasi dimana setiap unit *output* dari algoritma merepresentasikan sebuah kelas. LVQ digunakan untuk pengelompokan dimana jumlah kelompok sudah ditentukan arsitekturnya (target/kelas sudah ditentukan). Dengan kelas yang sudah ditentukan tersebut akan didapatkan hasil sentimen berdasarkan bobot data yang sudah diberikan sebelumnya.

1.2. Perumusan Masalah

1. Bagaimana menganalisis sentimen positif dan negatif dalam suatu masalah pada media sosial ?
2. Bagaimana mengukur performansi algoritma *Learning Vector Quantization* (LVQ) dalam klasifikasi text di media sosial pada *Machine Learning* ?

1.3. Tujuan

Tujuan dari tugas akhir ini adalah sebagai berikut :

1. Mengimplementasikan algoritma *Learning Vector Quantization* (LVQ) untuk klasifikasi *text* untuk menganalisa sentimen pada media sosial.
2. Menganalisa performansi algoritma *Learning Vector Quantization* (LVQ) pada pada sistem analisis sentimen yang dibangun.

1.4. Batasan Masalah

1. Studi kasus yang diambil adalah pemilihan gubernur DKI Jakarta 2017.
2. Data yang digunakan merupakan *tweet* berbahasa Indonesia.
3. Data yang digunakan bersifat *offline* (tidak terintegrasi secara *realtime* dengan Twitter).
4. Data *tweet* yang diambil hanya komentar teks saja, tanpa atau dengan mengabaikan URL dan juga *hashtag* yang sering muncul dalam *tweet*.
5. Pengambilan data sekitar bulan Agustus – November 2016.
6. Data yang akan dianalisis nantinya akan dibatasi menjadi 100, 200, dan 300 data.
7. Analisis sentimen yang dilakukan dibatasi hanya untuk *tweet* yang tidak mengandung sentimen dalam bentuk negasi.

1.5. Hipotesis

Algoritma *Learning Vector Quantization* (LVQ) berbeda dengan algoritma yang lain karena dapat kita gunakan untuk meminimalisir kesalahan dalam pengklasifikasian dengan menggunakan sifat dari algoritma ini yang mendistribusikan kelas *vector*. Sehingga *machine learning* yang kita gunakan untuk menganalisis sentimen pada *social media* diharapkan dapat menghasilkan performansi yang akurat.

1.6. Metodologi Penyelesaian Masalah

Dalam menyelesaikan Tugas Akhir ini dilakukan pendekatan sistematis untuk memecahkan masalah yang telah dijabarkan di atas. Adapun metode penelitian yang akan dilakukan adalah sebagai berikut :

1. Studi Literatur

Pada tahap ini akan dilakukan pengumpulan informasi yang berhubungan dengan Tugas Akhir ini khususnya pada metode yang diterapkan (*Learning Vector Quantization*) melalui sumber-sumber seperti jurnal penelitian, buku-buku teori, dan sumber-sumber lain termasuk sumber yang didapatkan melalui internet.

2. Pencarian dan Pengumpulan Data

Pada tahap ini dilakukan pencarian dan pengumpulan data Twitter mengenai komentar terhadap para calon Gubernur DKI Jakarta masing-masing berdasarkan nama calonnya yaitu Agus, Ahok, dan Anies. Pengumpulan data ini menggunakan Crawler menggunakan bahasa pemrograman Python yang menggunakan API dari Twitter Developer.

3. Analisis Kebutuhan dan Perancangan Sistem

Pada tahap ini dilakukan analisis kebutuhan dan perancangan sistem untuk pengembangan aplikasi dalam menghasilkan analisis sentimen dengan menggunakan metode *Learning Vector Quantization*.

4. Implementasi Sistem

Pada tahap ini, data-data dari Twitter yang telah dikumpulkan diolah dan dilakukan ekstraksi fitur dalam tahap *preprocessing* dan dilakukan klasifikasi sentimen menggunakan metode *Learning Vector Quantization*.

5. Pengujian dan Analisis Hasil

Pada tahap ini dilakukan pengujian terhadap sistem dengan menggunakan data yang telah didapat dari proses sebelumnya. Kemudian dilanjutkan dengan menganalisa hasil dari pengujian tersebut, dimana analisis mengacu pada rumusan masalah dan tujuan dari penelitian ini.

6. Penyusunan Laporan Tugas Akhir

Pada tahap ini dilakukan penyusunan Tugas Akhir sesuai dengan ketentuan-ketentuan yang telah diterapkan oleh institusi yang berfungsi sebagai dokumentasi dari apa yang selama ini telah dikerjakan.

1.7. Sistematika Penulisan

BAB I **Pendahuluan.** Pada bagian ini akan dibahas mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, batasan masalah dan sistematika penulisan.

BAB II **Dasar Teori.** Pada bagian ini akan dibahas mengenai landasan teori yang berkaitan dengan penelitian tugas akhir ini.

BAB III **Perancangan Sistem.** Pada bagian ini akan dibahas mengenai deskripsi sistem dan proses perancangan sistem secara keseluruhan.

BAB IV

Analisis dan Pengujian. Pada bagian ini akan dibahas mengenai pengujian terhadap sistem. Kebutuhan untuk pengujian, skenario dan analisis hasil dari pengujian sistem.

BAB V

Penutup. Pada bagian ini akan dibahas mengenai kesimpulan dari seluruh penelitian tugas akhir ini dan saran untuk penelitian selanjutnya.