

BAB I

PENDAHULUAN

1.1 Latar Belakang

Twitter merupakan salah satu media sosial yang berfungsi untuk menjalin komunikasi antar pengguna serta mendapatkan informasi atau berita terbaru. Hanya dengan membuat *Tweet* atau ciutan, pengguna *Twitter* dapat secara langsung membagikan informasi. *Twitter* sudah sangat familiar bagi pengguna internet di Indonesia. Menurut catatan [1], total populasi di Indonesia pada bulan Januari 2019 adalah 268,2 Juta jiwa dan diantaranya terdapat 150 Juta pengguna aktif internet. Dari jumlah pengguna aktif internet tersebut, jasa layanan iklan mencatat terdapat 6,43 juta pengguna aktif *Twitter*. Hal ini berarti jumlah pengguna *Twitter* berada di angka tersebut atau bahkan bisa lebih. *Twitter* menjadi incaran pengguna aktif internet karena banyak publik figur, pejabat pemerintah, serta orang-orang penting ikut menjadi salah satu penggunanya. Media sosial *Twitter* juga bermanfaat dalam kebutuhan mencari informasi pekerjaan karena biasanya sebuah perusahaan juga menyebarkan informasi perekrutan pegawai baru menggunakan akun *Twitter Official* masing-masing perusahaan.

Mendapatkan pekerjaan adalah hal yang pastinya diinginkan oleh semua orang. Dengan mendapatkan pekerjaan yang layak maka kualitas hidup seseorang juga akan naik. Hal yang sama juga dirasakan oleh suatu perusahaan. Dengan mendapatkan tenaga kerja yang berkualitas maka daya saing perusahaan itu juga akan meningkat. Suatu perusahaan pastinya akan menyeleksi terlebih

dahulu kriteria apa yang sedang dibutuhkan. Kini banyak perusahaan yang juga menyeleksi tenaga kerja dilihat dari sifat karakteristiknya. Karena *attitude* atau sikap menjadi hal utama dalam bekerja. Dalam ilmu psikologis, sifat dan karakteristik individu adalah hasil dari proses perkembangan psikologis seseorang yang mampu menjelaskan berbagai perbedaan individu yang satu dan individu lainnya, serta menjabarkan sifat manusia dalam berperilaku [2]. Suatu perusahaan tidak ingin mendapatkan kesalahan atau kerugian hanya disebabkan oleh hal kecil yang berasal dari perilaku pegawainya dalam bekerja. Maka dari itu, suatu perusahaan akan menyeleksi para tenaga kerja dengan sistem rekrutmen yang ketat. Sehingga didapatlah tenaga kerja yang memiliki kemampuan intelektual serta sifat yang baik.

Dalam menyeleksi terutama menyeleksi sifat karakteristik seseorang tidaklah mudah. Seorang *HR* atau *Human Resources* pastilah memiliki kriteria dan metode tersendiri dalam menentukan karakter pegawai yang akan direkrutnya sesuai dengan kebutuhan perusahaan. Menyeleksi pegawai baru dengan cara manual biasanya akan memakan waktu lebih lama, apalagi jika jumlah pendaftar banyak. Maka dari itu, dengan bantuan teknologi yang mumpuni dalam bidang *Data Mining* dan algoritma yang mampu mengklasifikasi, pekerjaan seorang *Human Resources* akan jauh lebih dimudahkan. Beberapa algoritma klasifikasi diantaranya adalah *logistic regression*, *decision tree*, *naive bayes*, dan *neural network*. Menurut [3], dalam pengujiannya menganalisis komparasi beberapa algoritma tersebut, didapatkan algoritma *Naive Bayes* memiliki tingkat akurasi yang tinggi yaitu 93,47%.

Dari permasalahan seorang *Human Resources* tersebut, penulis berencana untuk membuat suatu sistem aplikasi berbasis *Web* dimana sistem tersebut dapat secara cepat mengklasifikasikan karakter seseorang sehingga memudahkan dalam perekrutan pegawai baru bagi suatu perusahaan. Sistem ini akan mengambil data calon pegawai baru menggunakan konsep *Data Mining* dari akun *Twitter*nya terlebih dahulu. Data yang diambil adalah semua *Tweet* atau ciutan yang pernah ditulis dari awal membuat akun *Twitter* sampai yang terbaru. Dengan menggunakan Algoritma *Naive Bayes Classifier* yang dianggap memiliki kinerja yang baik sehingga sistem ini diharapkan mampu mengklasifikasi karakter seseorang dengan bantuan metode DISC, yaitu pengukuran kepribadian yang bertujuan untuk melihat dan mengukur perilaku kerja yang berdasarkan pada 4 tipe kepribadian seseorang yaitu *Dominance*, *Influence*, *Steadiness*, *Compliance* sehingga dapat menjadi salah satu pertimbangan apakah seorang calon pegawai kerja baru tersebut layak direkrut atau tidak.

Sistem ini hanya sebatas penunjang keputusan atau memberikan rekomendasi mengenai perekrutan calon pegawai baru. Dengan adanya sistem ini, diharapkan dapat membantu kinerja seorang *Human Resources* sehingga kualitas mutu penyeleksian calon tenaga kerja baru akan menjadi semakin baik dan semakin efisien serta meminimalisir kesalahan dalam perekrutan calon tenaga kerja baru dalam suatu perusahaan.

1.2 Perumusan Masalah

Berdasarkan latar belakang masalah yang telah dijelaskan diatas, maka rumusan masalah dalam penelitian ini adalah :

1. Bagaimana hasil dari analisis profil akun twitter calon tenaga kerja baru dalam pengklasifikasian karakter DISC dengan Algoritma Naive Bayes Classifier?
2. Bagaimanakah akurasi Algoritma Naive Bayes Classifier dalam mengklasifikasi karakter DISC pada akun twitter calon tenaga kerja baru?

1.3 Batasan Masalah

Beberapa batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Luas lingkup hanya meliputi analisis untuk mengklasifikasi karakter seseorang.
2. Hanya menggunakan sistem berbasis *Web*.
3. Objek yang digunakan hanya meliputi data dari akun personal Twitter.
4. Metode pembobotan yang digunakan hanya *W-IDF (Weighted-Inverse Document Frequency)*.
5. Algoritma yang digunakan hanya Naive Bayes Classifier.
6. Penentuan karakter hanya menggunakan konsep DISC (*Dominance, Influence, Steadiness, Compliance*).

1.4 Maksud dan Tujuan Penelitian

Tujuan umum yang ingin dicapai dalam penelitian ini adalah untuk mengetahui pemanfaatan media sosial Twitter dalam mengklasifikasi sifat dan karakteristik seseorang. Tujuan khusus penelitian ini adalah sebagai berikut :

1. Penelitian ini ditujukan untuk menguji Algoritma Naive Bayes Classifier (NBC) dalam mengklasifikasi sifat dan karakteristik seseorang dengan konsep DISC.
2. Penelitian ini juga ditujukan untuk mengukur akurasi Algoritma Naive Bayes Classifier (NBC) dalam mengklasifikasi sifat dan karakteristik seseorang dengan konsep DISC.

1.5 Manfaat Penelitian

Melalui penelitian ini diharapkan dapat diperoleh manfaat sebagai berikut :

1. Bagi peneliti akan mendapatkan informasi mengenai hasil pengujian Algoritma Naive Bayes Classifier (NBC) dalam mengklasifikasi sifat dan karakteristik seseorang dengan konsep DISC beserta tingkat akurasi.
2. Bagi perusahaan akan memudahkan dalam penyeleksian tenaga kerja baru ditinjau dari sifat dan karakteristik yang didapat dari akun twitternya.

1.6 Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah metode eksperimental dan dilaksanakan dalam tahapan-tahapan sebagai berikut :

1.6.1 Metode Pengumpulan Data

Pada tahap ini dilakukan pengambilan data latih dari *Twitter* dengan menggunakan cara *scraping* dengan bantuan library Scrapy untuk mendapatkan seluruh *Tweet* dari linimasa suatu akun.

1.6.2 Metode Analisis

Pada tahap ini dilakukan analisis dan pembersihan terhadap data *Tweet* yang telah berhasil didapat, diantaranya Preprocessing, tokenisasi, dan melakukan pembobotan sehingga pada hasil akhir seluruh *Tweet* yang diolah tersebut dapat diklasifikasikan.

1.6.3 Metode Perancangan

Pada tahap ini dilakukan perancangan dalam pembuatan sistem, database, serta alur dalam mengimplementasikan algoritma.

1.6.4 Implementasi

Tahap ini berisi penulisan kode program, perhitungan algoritma secara manual, dan memasukkan algoritma yang digunakan ke dalam sistem yang telah dibuat.

1.6.5 Metode Pembahasan

Pada tahap ini dilakukan pengamatan mengenai hasil yang didapat setelah melakukan implementasi antara algoritma yang digunakan dengan sistem yang telah dibuat. Selain itu, dilakukan juga perhitungan akurasi dari performa

algoritma Naive Bayes Classifier dalam mengklasifikasi karakter calon tenaga kerja baru. Metode yang digunakan adalah Cross Validation dan diimplementasikan kedalam bentuk penyajian *Confusion Matrix*.

1.7 Sistematika Penulisan

Sistematika penulisan dibuat untuk mempermudah dalam penyusunan skripsi ini maka perlu ditentukan sistematika penulisan yang baik. Sistematika penulisannya adalah sebagai berikut :

BAB I PENDAHULUAN

Bab pendahuluan mendeskripsikan mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan, manfaat dan sistematika penulisan.

BAB II LANDASAN TEORI

Berisi tentang teori-teori yang digunakan dalam penelitian, perancangan dan pembuatan sistem.

BAB III ANALISIS DAN PERANCANGAN

Dalam bab ini penulis mengemukakan analisis tentang kebutuhan sistem serta metode penelitian yang dilakukan dalam perancangan.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Memaparkan dari hasil-hasil tahapan penelitian, mulai dari analisis, desain, hasil testing dan implementasinya.

BAB V PENUTUP

Berisi kesimpulan dan saran dari seluruh penelitian yang telah dilakukan.