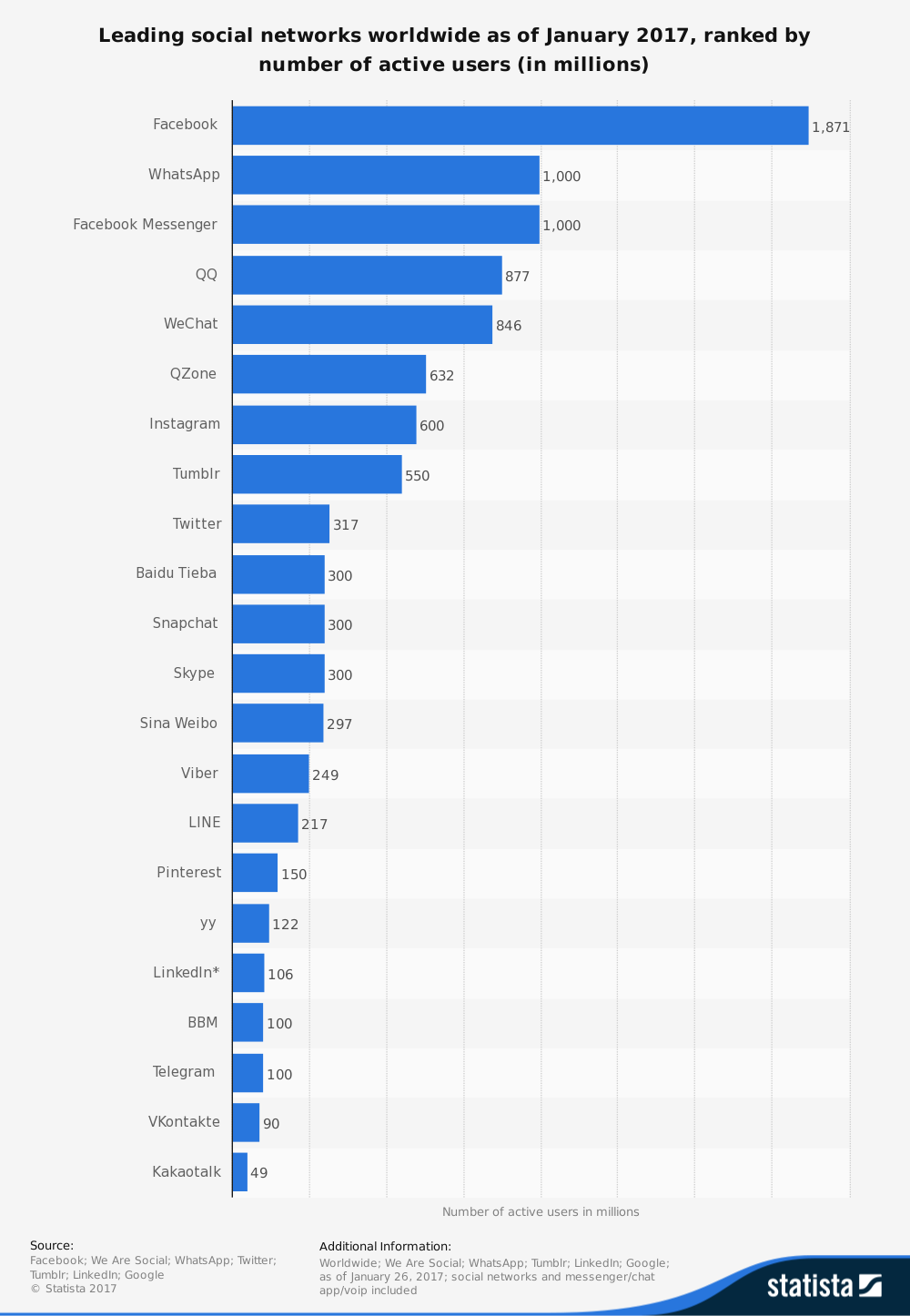
**BAB 1**

**PENDAHULUAN**

**1.1 Latar Belakang**

*Social media* telah menjadi sebuah alat komunikasi dan interaksi antar manusia selama beberapa tahun terakhir. Di zaman dimana hampir seluruh manusia memiliki *smartphone*, interaksi langsung antar manusia sudah jarang terjadi. Sehingga sulit untuk mengenali dan mengetahui kepribadian dari seseorang. Tetapi, hal ini berbeda dengan apa yang terjadi di *social media*.



Gambar 1.1 Survei Pengguna *social media(https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/)*

Pada gambar di atas dapat dilihat bahwa pengguna *social media* di dunia telah mencapai milyaran pengguna dan akan terus berkembang ke depannya. Sehingga *social media* merupakan sarana yang paling tepat untuk memperoleh data dan memprediksikan kepribadian seseorang. Penelitian yang akan kami lakukan akan menggunakan data dari *social media* Facebook. Berdasarkan data pada gambar diatas, Facebook memiliki pengguna terbesar yang mencapai 1,8 Milyar pengguna. Pengguna Facebook umumnya mengungkapkan ekspresi, perasaan maupun opini mereka di *user feed* mereka. Walaupun saat ini Facebook lebih banyak digunakan untuk berbagi foto dan video, penelitian kami kali ini akan berfokus pada *linguistic* pengguna.

Berbagai penelitian psikologi menunjukkan bahwa terdapat korelasi antara kepribadian dengan tingkah laku *linguistic* (Ilmu bahasa) seseorang. Korelasi ini dapat dengan efektif dianalisa dan digambarkan dengan pendekatan NLP (*Natural Language Processing*). Oleh sebab itu penulis membuat sebuah sistem prediksi kepribadian seseorang melalui data dari aktivitas pengguna tersebut di *social media* Facebook. Terdapat banyak kegunaan dari mengetahui kepribadian seseorang maupun kepribadian diri sendiri. Manfaat yang paling utama adalah lebih mengenal diri sendiri. Karena dengan mengetahui pola kepribadian kita sendiri, akan memudahkan dalam berinteraksi dengan orang lain. Kedua, mengenali pola orang lain akan membuat kita memahami lebih baik cara bersosialisasi. Kerap kali kesalahpahaman apalagi terkait faktor emosional terjadi karena ada benturan kepribadian.

Dalam pekerjaan, sangat penting mengetahui tipe kepribadian ini. Karena tentu bekerja adalah agar tujuan bersama dalam perusahaan tercapai. Diperlukan orang-orang yang mampu bekerjasama dalam tim. Maka, dalam tes awal pekerjaan sudah lazim diadakan tes psikologi (psikotes) untuk mendapatkan potret kepribadian dari pelamar. Hasilnya adalah untuk memetakan (*mapping*) orang yang tepat di tempat yang tepat (*the right man on the right place*).

Dengan mengetahui tipe kepribadian kita sendiri ditambah kemampuan menduga kepribadian orang lain, kita akan lebih mudah meraih sukses. Terutama sekali dalam hal membina hubungan dengan orang lain dimana kita sebagai makhluk sosial tentu tak bisa lepas dari interaksi dengan sesama.

Sistem prediksi ini akan dibangun dengan menggunakan model kepribadian “*Big Five Personality Traits*”. Terdapat beberapa model kepribadian lain seperti MBTI (*Myers-Briggs Type Indicator*) atau DISC. Tetapi, Big Five Personality Traits merupakan model yang sedang populer saat ini dan digunakan dalam banyak penelitian yang berhubungan dengan kepribadian seseorang. Sistem prediksi kepribadian ini diharapkan dapat membantu para pekerja menemukan pekerjaan yang paling cocok dengan diri mereka. Ilmu membaca kepribadian seseorang memang bukan hal baru dan sudah dikembangkan beratus-ratus tahun lamanya. Namun, sampai hari ini belum ada teori maupun alat (tes) yang bisa menjelaskan 100% akurat mengenai kepribadian dan perilaku seseorang. Sebab manusia itu unik. Hampir tidak ada manusia yang sama satu sama lain, walaupun mereka kembar identik. Namun demikian setidaknya jika kita menggunakan prinsip hukum 20/80 dari Vilfredo Pareto, yang berarti “kita dapat menggunakan alat ukur yang hanya mengukur 20% saja namun mampu mewakili sebagian besar (80%) aspek yang diukur”, maka kita dapat memahami tentang hasil test ini bahwa paling tidak dapat memberikan gambaran dari Tipe Kepribadian audiensnya.

Berbagai sistem prediksi kepribadian sudah banyak dikembangkan menggunakan berbagai algoritma classifier, seperti SVM (*Support Vector Machine*), *Naïve Bayes*, dan *Logistic Regression*. Penulis akan menggunakan algoritma-algoritma tersebut dan membandingkan hasil dan akurasi prediksinya.

**1.2 Rumusan Masalah**

Berdasarkan uraian latar belakang di atas, maka dapat dirumuskan permasalahan yang ada yaitu :

1. Berapakah selisih akurasi yang dihasilkan dari masing-masing algoritma *classifier* yang akan diaplikasikan (*Naïve Bayes*, *Support Vector Machines*, dan *Logistic Regression*)?
2. Berapakah persentase *acceptance* dari pengguna mengenai hasil kepribadian diri mereka?
3. Apakah dengan sistem prediksi kepribadian dapat membantu pengguna dalam menentukan karir maupun rencana pengguna ke depannya?

**1.3 Ruang Lingkup**

Ruang lingkup yang akan dibahas adalah:

1. Kepribadian yang akan diuji berdasarkan *The Big Five Traits* (*Neuroticism, Extraversion, Openness to Experience, Agreeableness,* dan *Conscientiousness*).
2. Data yang digunakan dalam penelitian didapatkan dari projek myPersonality (http://mypersonality.org/wiki; Celli et al. 2013).
3. Menggunakan feature gabungan untuk mendapatkan hasil prediksi kepribadian dari pengguna yakni meliputi, status, umur, jenis kelamin, *network size*, jumlah *likes*, jumlah foto dan jumlah grup yang dimiliki oleh pengguna.
4. Penelitian akan menggunakan tiga algoritma machine learning dan dibandingkan akurasinya. Tiga algoritma tersebut yaitu, SVM (*Support Vector Machine*, *Naïve Bayes*, *Logistic Regression*).
5. Pengembangan akan menggunakan sistem operasi Windows dan menggunakan bahasa pemograman Phyton.
6. Sistem dari penelitian ini akan diimplementasikan ke dalam sebuah aplikasi *mobile* dengan *platform* *Android.*

**1.4 Tujuan dan manfaat**

Penelitian dan pembangunan sistem prediksi kepribadian ini memiliki tujuan dan manfaat sebagai berikut :

**1.4.1 Tujuan**

1. Menghasilkan sistem prediksi kepribadian berdasarkan The Big Five Traits terhadap pengguna secara akurat
2. Mengetahui algoritma yang paling optimal dan efisien dalam proses penelitian sistem ini.
3. Meningkatkan tingkat akurasi dengan menggunakan *multi feature* sebagai aspek penentu kepribadian.

**1.4.2 Manfaat**

1. Dengan mengetahui kepribadian secara tepat, pengguna dapat mengetahui kelebihan apa yang mereka miliki dan memungkinkan mereka untuk mengembangkannya.
2. Pengguna juga dapat mengetahui kelemahan mereka untuk segera diperbaiki.
3. Pengguna dapat mengetahui potensi-potensi apa yang ada di dalam diri dan memaksimalkannya agar dapat berhasil di dunia kerja dan kehidupannya.
4. Membantu pengguna dalam memilih jenjang karir yang tepat.
5. Membantu pihak HRD perusahaan dalam mengetahui calon karyawan baru secara objektif.

**1.5 Metode Penelitian**

Dalam pengembangan sistem aplikasi ini, digunakan tiga tahap metode penelitian, yaitu metode pengumpulan data, metode perancangan sistem dan metode evaluasi.

**1.5.1 Metode Pengumpulan Data**

1. Studi Literatur

Penelitian dimulai dengan mencari referensi-referensi yang berhubungan dengan sistem yang akan dikembangkan, baik melalui buku, tesis, jurnal atau paper internasional yang dapat memberikan informasi yang berguna dalam pengembangan sistem.

1. Penentuan dataset

Setelah melalui tahap pembelajaran dan melalui analisa serta diskusi, peneliti menetapkan dataset yang akan digunakan dalam penelitian ini.

1. Analisis Penelitian Sejenis

Mempelajari penelitian yang telah dilakukan sebelumnya dan mendapatkan hal penting yang dapat dijadikan acuan sebagai perbandingan dengan penelitian ini.

**1.5.2 Metode Perancangan Sistem**

1. *Collecting Facebook User Data*

Pengambilan data *user Facebook* melalui *myPersonality*. Projek atau aplikasi ini dikembangkan di tahun 2007 dimana aplikasi ini digunakan oleh model untuk mengisi kuesioner yang berhubungan dengan model *The Big Five Traits* dan untuk mendapatkan hasil dari kepribadian mereka berdasarkan kuesioner tersebut. Informasi yang diperoleh berupa status, *network size*, umur dan jenis kelamin

1. *Feature Extraction*

Melakukan pemilihan *feature-feature* yang akan digunakan dan di filter melalui tahap *Feature Extraction*. Penelitian ini tidak hanya menggunakan *linguistic feature* tetapi akan digabungkan dengan *feature* tambahan seperti umur, jenis kelamin, jumlah *likes*, jumlah foto dan jumlah *friends* (*network size*)

1. Classifying Personality

Mengklasifikasi kepribadian berbasis model The Big Five Traits menggunakan teknik *supervised machine learning* dengan *tools machine learning* WEKA untuk mendapatkan hasil prediksi serta akurasi sistem.

**1.5.2 Metode Evaluasi**

Sistem yang dikembangkan dievaluasi dengan menggunakan 2 metode evaluasi, yaitu:

1. Evaluasi Objektif

Hasil evaluasi objektif diperoleh dengan membandingkan hasil akurasi sistem prediksi yang dibangun dengan hasil akurasi pada penelitian sebelumnya. Perbandingan dilakukan dengan menggunakan data training yang sama, namun *feature* yang berbeda untuk dilihat perbedaan akurasinya. Sistem juga akan diujicobakan pada beberapa algoritma classifier lainnya seperti *Naïve Bayes* dan *Logistic Regression*.

1. Evaluasi Subjektif

Hasil evaluasi secara subjektif diperoleh dengan membandingkan hasil prediksi kepribadian menggunakan sistem prediksi yang dibangun dengan hasil prediksi kepribadian oleh pakar psikologi*.*

**1.6 Sistematika Penulisan**

Penelitian ini dibagi menjadi beberapa bagian dengan sistematika penulisan sebagai berikut:

**1. BAB 1 Pendahuluan**

**2. BAB 2 Tinjauan Pustaka**

**3. BAB 3 Metodologi**

**4. BAB 4 Hasil dan Penelitian**

**5. BAB 5 Simpulan dan Saran**