1. Pendahuluan

1.1 Latar Belakang

Indonesia merupakan negara berkembang yang memiliki jumlah penduduk terbesar ke 4 di dunia dan jumlah penduduknya meningkat setiap tahunnya, dan oleh karena itu banyak juga tenaga kerja terus mengalami peningkatan setiap tahunnya dari berbagai jenis tingkat pendidikan. Disini perusahaan yang sedang mencari pelamar kerja atau pegawai baru pasti akan mendapatkan lebih banyak lamaran kerja dibanding tahun tahun sebelumnya.

Proses pencarian pelamar kerja ini dimulai dari perusahaan yang memberitahukan jika perusahaannya sedang membutuhkan pegawai baru, dan pelamar yang tertarik akan mengirim berkas yang dibutuhkan perusahaan. Disitu perusahaan akan memeriksa berkas berkas yang dikirimkan pelamar untuk memilih siapa yang akan masuk ke tahap selanjutnya. Untuk pelamar yang dipilih akan melakukan *interview* atau wawancara. dan perusahaan akan melakukan seleksi lagi terhadap hasil wawancaranya. Pelamar yang berhasil diterima akan diberitahukan dengan *offering letter* yang menandakan bahwa sang pelamar sudah di terima di perusahaan tersebut.

Banyak perusahaan yang menerima pelamar kerja atau kandidat yang tidak sesuai, hal itu dapat disebabkan oleh adanya ketidaksesuaian, seperti bias dalam penseleksian(*confirmation bias, affinity bias)*, perbedaan standar penilaian pewawancara, waktu dan biaya melakukan wawancara yang lama dan mahal. Dan saat melakukan seleksi juga dibutuhkannya ketelitian dan memakan waktu.

Pada penelitian ini dilakukannya penilaian atau klasifikasi teks (*text classification*) hasil wawancara yang dilakukan oleh pelamar menggunakan pembelajaran mesin. Pembelajaran mesin (*machine learning)* banyak digunakan untuk mengklasifikasikan suatu teks [1]. Dengan menggunakan pembelajaran mesin ini diharapkan dapat membantu proses seleksi, mempercepat proses seleksi pelamar kerja. Oleh karena itu dibutuhkannya pembuatan suatu sistem penilaian atau klasifikasi teks dengan bantuan pembelajaran mesin untuk membuat proses seleksi lebih efisien.

Metode yang di pakai dalam penelitian ini adalah regresi logistik (*logistic regression).* Beberapa keunggulan dari regresi logistik adalah metode ini dapat secara alami memberikan probabilitas dan meluas ke masalah klasifikasi multi-label. [2] [3] regresi logistik menyediakan metode untuk memodelkan respons biner variabel dalam pembelajaran mesinnya, yang mengambil nilai 1 dan 0[4]. Penelitian ini berfokus pada tahap wawancara, dimana data hasil wawancara akan dilakukan pelabelan. Pelabelan dilakukan dengan memberikan tanda kelas atau label terhadap suatu teks [5].

Dataset yang diperoleh dalam penelitian ini didapatkan dari hasil Simulasi wawancara PT. Telkom. Data tersebut berjumlah sebanyak 54 data jawaban wawancara pelamar kerja yang terdiri dari *content* dan *core values*. Dataset ini memiliki 9 (sembilan) *core values* yaitu: *action, enthusiams, focus, imagine, integrity, smart, solid, speed* dan *totality*, dimana masing-masing core values terdiri dengan dua label/kelas yaitu layak atau tidak layak. Data yang didapatkan hanyalah 54 teks hasil wawancara dikarenakan ini adalah penelitian tahap awal, PT Telkom hanya dapat melakukan simulasi terbatas, dan mendanai tim ahli lingustik untuk melakukan labeling tersebut. Terdapat 2 Label didalam dataset, yaitu tidak layak dan layak. Disini hanya terdapat 2 Label saja karena data akan digunakan untuk membuat prototype dan proof of concept apakah memungkinkan adopsi *machine learning* untuk proses pelamaran kerja.

Aplikasi yang akan dibuat akan berupa web yang menggunakan *django framework. Django framework. Django framework ini* digunakan sebagai platform permbuatan website karena django adalah *open-source python web framework* yang sangat berguna untuk mendukung pembuatan aplikasi sistem seleksi pelamar kerja yang akan berupa website. Didalam website ini akan terdapat fitur seleksi pelamar kerja dengan model klasifikasi yang sudah dibuat menggunakan pengklasifikasian regresi logistik dengan dataset yang diperoleh.

1.2 Topik dan pembatasan

Permasalahan yang dibahas dalam tugas akhir ini adalah bagaimana merancang model klasifikasi seleksi pelamar kerja dengan data teks hasil wawancara menggunakan metode regresi logistik (*logistic regression*), bagaimana performa, kinerja dan akurasi dari sistem yang dibangun dan bagaimana membuat aplikasi sistem seleksi pelamar kerja dengan model klasifikasi yang telah dibuat agar nenubunakusur waktu, biaya dan mengurangi keterlibatan pihak ketiga dalam proses seleksi pelamar kerja. Batasan masalah dalam penelitian tugas akhir ini adalah dataset yang digunakan dalam penelitian merupakan data teks wawancara pelamar kerja yang tersedia yaitu 54 *content* dengan 9 (sembilan) *core values*. Dataset disimpan dalam bentuk file *Microsoft Excel Comma Separated Values* (.csv). Dataset yang digunakan pada penelitian ini sangat sedikit dengan jumlah 54 data. Label kelas yang digunakan dalam klasifikasi terbagi menjadi dua kelas, yaitu layak dan tidak layak. Dua label kelas tersebut digunakan untuk membuat *prototype dan proof of concept* apakah memungkinkan adopsi *machine*

learning dalam proses perekrutan karyawan. Hasil dari pengklasifikasian teks hasil wawancara hanya berfokus kepada hasil akhir yang berupa nilai dari 9 core value yang menghasilkan 2 label Tidak layak dan Layak.

1.3 Tujuan

Tujuan dari penulisan tugas akhir ini untuk membuat prototype dan proof of concept untuk mengetahui apakah memungkinkan memakai klasifikasi machine learning untuk seleksi pelamar kerja. klasifikasi data teks hasil wawancara dengan menggunakan metode regresi logistik (*logistic regression*), serta menganalisa hasil kinerja dan akurasi sistem klasifikasi yang dibangun dengan menggunakan metode regresi logistik dan untuk meminimalisir waktu, biaya dan mengurangi keterlibatan pihak ktiga dalam proses seleksi pelamar kerja.