

1. Pendahuluan

Latar Belakang

Kata karyawan menurut Kamus Besar Bahasa Indonesia (KBBI) adalah orang yang bekerja pada suatu lembaga (kantor, perusahaan, dan sebagainya) dengan mendapat gaji (upah). Karyawan direkrut untuk memenuhi sumber daya yang dibutuhkan oleh perusahaan dalam mencapai tujuannya. Karyawan dapat disebut sebagai salah satu aset perusahaan dikarenakan sebuah perusahaan tidak dapat mencapai tujuannya tanpa adanya bantuan dari karyawan.

Employee attrition atau keluarnya karyawan dari perusahaan adalah sebuah tantangan untuk perusahaan mengingat karyawan merupakan salah satu aset penting bagi perusahaan. Perekrutan karyawan baru tentunya merupakan salah satu solusi untuk menangani dampak dari *employee attrition* ini. Namun, perekrutan karyawan tentunya tidak mudah dan memerlukan waktu, biaya, dan sumber daya lain [1]. Berdasarkan penelitian yang dilakukan oleh Marsden [2], untuk setiap karyawan yang keluar dari perusahaan, perusahaan terbebani sebanyak 1 sampai 1.2 gaji tahunan karyawan tersebut. Oleh karena itu, perekrutan karyawan baru mungkin bukanlah solusi yang “murah”.

Keluarnya karyawan berdampak pada kualitas perusahaan apabila karyawan tersebut merupakan karyawan yang berkualitas. Keluarnya karyawan juga dapat berdampak pada karyawan yang lain, terutama karyawan-karyawan lain yang sebelumnya tergabung dalam tim yang sama. Hal tersebut dapat mengganggu dinamika tim yang sudah ada. Oleh karena itu, dibutuhkannya solusi lain untuk mencegah *employee attrition* ini.

Perlunya untuk menganalisis dan memprediksi *employee attrition* agar dapat melakukan tindakan preventif dan persuasif sehingga dapat mencegah kerugian-kerugian yang dapat disebabkan oleh keluarnya karyawan. Oleh karena itu, dibutuhkannya sebuah *tools* atau alat bantu untuk memprediksi apakah seorang karyawan akan keluar dari perusahaan.

Pada era digital 4.0 ini penggunaan *machine learning* sudah sangat luas. *Machine learning* biasa digunakan untuk melakukan klasifikasi dan klusterisasi data. *Machine learning* melakukan klasifikasi dengan mencari rumus atau pola data yang serupa dengan data yang lainnya. Hasil dari klasifikasi *machine learning* dapat digunakan sebagai prediksi. *Machine learning* dapat membuat prediksi data dengan menggunakan algoritma *support vector machine* (SVM), *k-nearest neighbors* (kNN), dan algoritma klasifikasi lainnya [3]. Prediksi menggunakan *machine learning* ini bersifat otomatis. Hal ini dimanfaatkan untuk memprediksi *employee attrition*.

Hasil klasifikasi dari *machine learning* bergantung pada data set. Data set yang distribusi kelasnya seimbang lebih baik daripada data set yang distribusi kelasnya tidak seimbang, dan data set yang jumlah datanya banyak lebih baik daripada data set yang jumlah datanya sedikit [4]. Di *real-world*, permasalahan klasifikasi biasanya memiliki data set yang distribusi kelasnya tidak seimbang dan jumlah datanya sedikit. Data set yang seperti itu biasanya disebut sebagai *imbalanced dataset* dan memiliki dampak buruk pada klasifikasi *machine learning*. Oleh karena itu, pada penelitian ini dilakukan beberapa cara untuk menangani *imbalanced dataset* agar hasil klasifikasi menjadi lebih baik. Pada penelitian ini juga dibandingkan performansi algoritma SVM dan kNN dikarenakan keduanya sama-sama merupakan algoritma yang berbasis jarak, tetapi memiliki pendekatan yang berbeda.

Topik dan Batasannya

Berdasarkan latar belakang yang telah diuraikan sebelumnya, didapat rumusan masalah sebagai berikut:

1. Bagaimana cara implementasi *machine learning* untuk memprediksi *employee attrition*?
2. Bagaimana performansi algoritma SVM jika dibandingkan dengan algoritma kNN pada permasalahan prediksi *employee attrition*?
3. Bagaimana pengaruh penanganan *imbalanced dataset* pada algoritma SVM?

Kemudian, penelitian ini memiliki batasan sebagai berikut:

1. Dataset yang digunakan adalah dataset IBM HR Analytics Employee Attrition & Performance;
2. Algoritma yang dibahas pada penelitian ini hanya algoritma kNN dan algoritma SVM.

Tujuan

Berdasarkan rumusan masalah yang telah diuraikan sebelumnya, didapat tujuan penelitian sebagai berikut:

1. Implementasi model *machine learning* untuk memprediksi *employee attrition*;
2. Membandingkan performansi algoritma SVM dengan algoritma kNN pada permasalahan prediksi *employee attrition*;
3. Mengetahui pengaruh penanganan *imbalanced dataset* pada algoritma SVM.

Organisasi Tulisan

Terdapat 5 bagian pada jurnal Tugas Akhir ini. Bagian pertama adalah mendeskripsikan latar belakang, perumusan masalah, batasan, dan tujuan Penelitian ini. Bagian kedua adalah membahas mengenai literatur atau

studi yang mengacu dan mendukung tentang penelitian tugas akhir ini yang telah dilakukan sebelumnya. Bagian ketiga adalah pembangunan sistem klasifikasi yang terdiri dari pengumpulan data set, pemrosesan data set, dan membagi data set menjadi 2 bagian (data set latih dan data set uji). Bagian keempat adalah klasifikasi dan analisis hasil evaluasi pengujian dari sistem klasifikasi. Bagian terakhir adalah kesimpulan hasil pengujian beserta saran untuk penelitian selanjutnya.