

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Dalam sebuah perusahaan, karyawan merupakan salah satu bagian utama yang menjalankan perusahaan. Tanpa karyawan, perusahaan tidak akan dapat beroperasi sama sekali. Karyawan yang keluar akan diganti dengan karyawan baru, atau dapat disebut juga dengan pergantian karyawan atau *Employee Churn*. *Employee Churn* seringkali merugikan perusahaan, dan telah banyak peneliti yang meneliti permasalahan ini [1]. Kerugian yang disebabkan oleh pergantian karyawan antara lain biaya rekrutmen baru, pelatihan karyawan baru, dan potensi hilangnya karyawan berpengaruh [2]. Ada beberapa hal yang menyebabkan seorang karyawan untuk keluar dari perusahaan, seperti lingkungan kerja, lokasi kerja, kesetaraan *gender*, kesetaraan gaji, dan lain-lain [3]. Strategi yang diterapkan berbagai organisasi untuk mengurangi *Employee Churn* antara lain kenaikan gaji, kesempatan untuk jalan-jalan, rotasi pekerjaan yang sering, dan lain-lain [4].

Peneliti-peneliti sebelumnya telah meneliti data karyawan dengan berbagai model. Model-model tersebut antara lain *Resume Prediction* [5], [6], *Churn Prediction* [7], [8], dan *Uplift Modeling* [9], [10]. Ketiga model tersebut menggunakan algoritma *Machine Learning* yang berbeda-beda, seperti *Random Forest*, *Neural Network*, dan *Extreme Gradient Boosting*. Selain dari ketiga algoritma tersebut, *Logistic Regression* adalah salah satu algoritma *Machine Learning* yang juga dapat digunakan untuk memprediksi. Algoritma ini menggunakan teknik regresi. *Logistic Regression* juga telah digunakan pada penelitian sebelumnya untuk meneliti prakiraan hujan es [11], *customer churn* [12], [13], dan performa akademik [14].

Dalam penelitian ini, penulis menerapkan model *Uplift Modeling* yang membandingkan variabel *Treatment* dalam memprediksi. Model ini juga telah digunakan untuk meneliti pencegahan *dropout* [15], pemasaran [16], dan periklanan [17]. Penelitian ini menggunakan dataset *IBM HR Analytics* [18] yang juga telah diteliti oleh peneliti-peneliti sebelumnya [19]–[22]. Dataset ini berisi data pribadi dan pekerjaan karyawan. Penelitian ini menggunakan algoritma *Logistic Regression* dan

model *Uplift Modeling*. Diharapkan penelitian ini dapat menambah wawasan tentang model *Uplift* dalam memprediksi dan mempreskripsi *Employee Churn*.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah bagaimana cara mengatasi *Employee Churn* menggunakan *Uplift Modeling* dan algoritma *Logistic Regression*.

1.3. Tujuan

Tujuan dari penelitian ini adalah untuk menganalisa penerapan model *Uplift Modeling* menggunakan algoritma *Logistic Regression* dalam permasalahan *Employee Churn*.

1.4. Manfaat

Manfaat yang didapat dari penelitian ini antara lain:

- a. **Bagi Universitas Prima Indonesia:** memberikan referensi bagi mahasiswa/i Universitas Prima Indonesia tentang penerapan model *Uplift Modeling* dengan algoritma *Logistic Regression* dalam mengatasi masalah *Employee Churn*,
- b. **Bagi Instansi Terkait:** mengurangi masalah *Employee Churn* dengan lebih efektif.

1.5. Batasan Masalah

Batasan masalah pada penelitian ini antara lain:

- a. *Dataset* yang digunakan berupa *dataset IBM HR Analytics Employee Attrition & Performance* dari Pavansubhash [18],
- b. Model yang digunakan adalah *Uplift Modeling*,
- c. Algoritma yang digunakan untuk memprediksi dan mempreskripsi adalah *Logistic Regression*,
- d. Bahasa pemrograman yang digunakan adalah *Python 3.0* dengan *software Jupyter Notebook*.

1.6. Keterbaruan

Penerapan *Uplift Modeling* dalam permasalahan *Employee Churn* telah diteliti dengan berbagai algoritma seperti *Random Forest* dan *Extreme Gradient Boosting*. Sedangkan, metode *Uplift Modeling* yang telah diteliti antara lain *Pessimistic Uplift Modeling* [23], *Two-Model Uplift Modeling* [24], dan *Lai's Generalized Weighted*

Uplift Method (LWUM) [25]. Penelitian ini menggunakan metode LWUM dengan algoritma *Logistic Regression*.