

## BAB V

### SIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan data yang diperoleh dari Biro Informasi Akademik UMN, mahasiswa yang terdaftar di UMN sejak 2017 hingga 2023 berjumlah 16,178 data. Dari keseluruhan data tersebut, persentase mahasiswa yang *churn* dapat dikatakan sedikit yaitu hanya sekitar 10% dibandingkan dengan persentase mahasiswa yang tidak *churn* yaitu sekitar 90%. Walaupun demikian, tidak menutup kemungkinan bahwa di masa depan persentase *churn* tersebut dapat meningkat.

Pada penelitian ini, dibuat 3 model klasifikasi dengan algoritma yang berbeda, yaitu Decision Tree, Random Forest, dan XGBoost. Ketiga algoritma memiliki fitur-fitur berbeda yang dianggap penting dalam mengklasifikasikan mahasiswa ke kelas *churn* atau *not churn* atau disebut juga sebagai *feature importance*. Disamping perbedaan *feature importance* untuk tiap model, ditemukan beberapa fitur yang memiliki tingkat *importance* yang mirip untuk ketiga model. Fitur tersebut antara lain adalah *semester\_masuk*, *total\_sks*, *ips\_semester\_3*, dan *ipk\_semester\_4*.

Secara keseluruhan, dapat dikatakan bahwa ketiga model yang dibuat dengan 3 algoritma berbeda memiliki performa yang sangat baik. Model yang dibuat dengan algoritma XGBoost memberikan performa terbaik dengan akurasi sebesar 99%, *precision* sebesar 99%, *recall* sebesar 96%, dan *f1-score* sebesar 97%. Model yang dibuat dengan algoritma Random Forest menghasilkan nilai akurasi sebesar 98%, *precision* sebesar 97%, *recall* sebesar 96%, dan *f1-score* sebesar 96% dan model yang dibuat dengan algoritma Decision Tree menghasilkan nilai akurasi sebesar 97%, *precision* sebesar 93%, *recall* sebesar 94%, dan *f1-score* sebesar 94%. Nilai akurasi, *precision*, *recall*, dan *f1-score* berasal dari performa model terhadap *validation dataset*.

## 5.2 Saran

Output dari penelitian ini dapat dimanfaatkan untuk dijadikan bahan untuk membantu mencegah terjadinya *churn* pada dunia pendidikan, misalnya dengan merancang suatu *student retention strategy*. Pada masa yang akan datang, diharapkan adanya penyempurnaan maupun pengembangan terhadap penelitian ini. Algoritma yang dapat digunakan dalam membuat *classifier* masih sangat banyak diluar 3 algoritma yang telah digunakan. Data yang dapat dimanfaatkan untuk pembuatan model juga dapat diperbanyak lagi, sebab pada penelitian ini hanya dapat mengakses data mahasiswa 2017 sampai 2023 saja.

