

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini menghasilkan *dashboard* visualisasi data dan model *machine learning* yang dihasilkan dengan menggunakan data *training* karyawan PT XYZ. Visualisasi data aktivitas *online training* karyawan PT XYZ berhasil dihasilkan ke dalam 2 buah *dashboard* yang dibangun dengan aplikasi Tableau. *Dashboard* pertama memberikan informasi terkait demografi aktivitas *training online* karyawan secara garis besar sedangkan *dashboard* kedua lebih spesifik terkait waktu aktivitas *training online* karyawan. Informasi yang dihasilkan dalam *dashboard* dapat ditampilkan secara *live* dan telah melalui proses validasi data secara manual dengan data mentah.

Penelitian ini juga menghasilkan dua model *machine learning* yang memiliki tujuan untuk melakukan prediksi atas *group* yang terdiri dari beberapa *skills* dengan kata kunci tertentu untuk diberikan kepada karyawan dengan berbagai kriteria yang sesuai. Model *machine learning* pertama dengan algoritma *decision tree* untuk *personalized learning* berhasil memperoleh nilai akurasi 69,31% dan dengan nilai *precision* 53%, *recall* 69%, dan *F1 score* 58%.

Selain dengan algoritma *decision tree*, dilakukan pembuatan model dengan algoritma *random forest*. Model *machine learning* kedua dengan algoritma *random forest* untuk *personalized learning* berhasil memperoleh nilai akurasi 69,04% dan dengan nilai *precision* 52%, *recall* 62%, dan *F1 score* 57%. Dapat dikatakan bahwa model *random forest* tersebut memiliki hasil evaluasi yang mirip dengan model *decision tree*.

Berdasarkan hasil perbandingan, model *machine learning* yang terbaik adalah menggunakan algoritma *decision tree*. *Decision tree* memiliki hasil evaluasi yang unggul dibandingkan dengan *random forest* dan lebih cepat dalam pembuatan modelnya dengan perbedaan selisih 1 detik untuk jumlah data

yang sama dalam penelitian ini. Untuk tujuan rekomendasi, model *decision tree* telah dapat memenuhi tugasnya dengan baik dan cocok untuk data yang besar.

5.2 Saran

Terdapat beberapa saran yang dapat diaplikasikan dalam penelitian yang memanfaatkan data *training karyawan* dalam pembuatan *dashboard* visualisasi data dan model *machine learning*, antara lain:

- 1) Untuk *dashboard* visualisasi data akan lebih bervariasi dan lebih banyak fitur yang dapat ditambahkan apabila data yang digunakan lebih panjang jangka waktunya seperti seperti yang dapat ditambahkan untuk menampilkan aktivitas *training* per semester, kuartal, dan satuan waktu lainnya.
- 2) Menggunakan algoritma selain *decision tree* dan *random forest* dalam klasifikasi *skills* dari *course* yang perlu dikerjakan oleh karyawan dalam kaitannya dengan *personalized learning* seiring akan terus berkembangnya algoritma *machine learning*.
- 3) Memperoleh data *training* karyawan yang lebih baik dalam komposisi fitur serta target agar dapat menghindari terjadinya *imbalanced class* untuk meningkatkan kualitas model dalam hal akurasi, *precision*, *recall*, dan *F1 score*.

