

## BAB V

### PENUTUP

#### 5.2 Kesimpulan

Berdasarkan hasil dari analisis dan visualisasi yang telah dilakukan maka didapatkan beberapa kesimpulan sebagai berikut:

1. Beberapa faktor yang memberikan kontribusi dalam mempengaruhi pengambilan keputusan dalam model prediksi adalah atribut *duration*, *contact\_cellular*, *poutcome\_success*, *month\_jun*, *day*, *housing*, *pdays*, *age*, *balance*, *campaign*, *education\_tertiary*. Faktor-faktor tersebut dapat dipertimbangkan untuk optimalisasi strategi pemasaran kampanye bank berikutnya.
2. Dalam membangun model prediktif terdapat beberapa hal penting yang perlu untuk dilakukan. Pertama jika isi dari data merupakan kategorikal contohnya seperti dalam atribut *contact* yang berisikan jenis perangkat yang digunakan seperti *unknown*, *telephone*, dan *cellular* yang tidak dapat diolah oleh algoritma pembelajaran mesin perlu dilakukan pembuatan label baru menjadi *contact\_unknown*, *contact\_telephone*, dan *contact\_cellular* agar datanya berubah menjadi bilangan bulat sehingga dapat dilakukan pengolahan dengan algoritma pembelajaran mesin. Kemudian yang kedua jika menggunakan algoritma LightGBM dapat melakukan *hyper-parameter tuning* untuk melakukan pencarian parameter dengan skor terbaik. Selanjutnya yang ketiga melihat *confusion matrix* untuk melihat apakah algoritma memiliki kinerja yang baik dalam mengidentifikasi nasabah yang tidak tertarik pada kampanye (True negatif yang tinggi) dan nasabah yang tertarik pada kampanye (True positif yang tinggi) sehingga dapat meminimalkan upaya dan meningkatkan fokus pada pasar yang tepat. Kemudian

langkah keempat melakukan pengecekan *overfitting* untuk memastikan bahwa model memiliki kemampuan yang baik dalam menggeneralisasi pada data baru dan tidak hanya terlalu fokus pada data pelatihan, selain itu pengecekan ini juga mencegah terjadinya kesalahan dalam interpretasi. Jika terjadi *overfitting* pada model, hasil evaluasi pada data pelatihan akan menunjukkan kinerja yang sangat baik. Namun, ketika model tersebut diterapkan pada data pengujian atau situasi dunia nyata, performanya akan menurun secara signifikan. *Overfitting* menyebabkan model terlalu khusus dan terlalu cocok dengan data pelatihan, sehingga tidak dapat menggeneralisasi dengan baik pada data baru yang belum pernah dilihat sebelumnya.

3. Untuk melakukan interpretasi model prediktif yang dihasilkan oleh LightGBM dengan menggunakan SHAP dapat menggunakan plot-plot yang disediakan oleh SHAP sebagai visualisasi, dengan menggunakan plot yang dibutuhkan. Dalam analisis dan visualisasi yang telah dilakukan, plot yang digunakan antara lain adalah plot *bar*, *beeswarm*, *waterfall*, *decision* yang memiliki fungsi berbeda-beda. Plot bar digunakan untuk melihat atribut yang paling berpengaruh dalam pengambilan keputusan pada model prediksi yang dihasilkan LightGBM, untuk mendapatkan sebuah kesimpulan faktor-faktor apa saja yang penting dan relevan maka perlu dilakukan kombinasi dengan plot *beeswarm*, karena plot tersebut akan memberikan informasi mengenai faktor-faktor yang mempengaruhi keberhasilan kampanye pemasaran. Kemudian yang terakhir SHAP *decision plot* atau plot keputusan dapat digunakan untuk memudahkan eksplorasi antara masing-masing atribut pada model dengan penjelasan visual dari nilai masing-masing atribut mempengaruhi hasil prediksi serta identifikasi bagaimana

cara setiap atribut saling berinteraksi untuk mempengaruhi hasil prediksi, selain itu plot ini juga dapat melihat interaksi setiap atribut dan pengaruhnya pada subset tertentu yang dapat ditentukan sendiri, sehingga memudahkan untuk melakukan pengamatan terhadap subset tertentu. Pada penelitian ini atribut yang berkontribusi tinggi dan relevan dengan tujuan penelitian antara lain adalah *duration*, *contact\_cellular* yang merupakan hasil *encode* dari atribut *contact*, *poutcome\_success* yang merupakan hasil *encode* dari atribut *poutcome*, *month\_jun* yang merupakan hasil *encode* dari atribut *month*, *day*, *housing*, *pdays*, *age*, *balance*, *campaign*, *education\_tertiary* yang merupakan hasil *encode* dari atribut *education*.

## 5.2 Saran

Beberapa saran yang dapat diberikan untuk pengembangan dalam penelitian selanjutnya antara lain sebagai berikut:

1. Membuat model prediktif dengan kinerja yang lebih bagus lagi dalam mengurangi jumlah false positif dan false negatif untuk meningkatkan efektivitas kampanye pemasaran bank dan target yang lebih akurat.
2. Menggunakan data yang lebih baru dari data yang digunakan pada penelitian ini.
3. Melakukan penelitian lebih lanjut untuk melihat lebih dalam lagi interaksi antar atribut pada subset yang lebih banyak.
4. Mempertimbangkan nilai *random state* untuk melihat variasi hasil dari nilai *random state* yang berbeda.
5. Memperbesar range nilai dalam *hyper-parameter tuning* untuk mendapatkan parameter terbaik dan skor evaluasi terbaik.

6. Mencoba interpretasi lain selain SHAP (Shapley Additive exPlanations) untuk melihat hasil interpretasi yang berbeda.