

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perusahaan-perusahaan keuangan semakin banyak menggunakan teknologi untuk meningkatkan efisiensi dan daya tanggap layanan. Salah satu contohnya adalah *Dream Housing Finance*, perusahaan yang berfokus pada pembiayaan perumahan di berbagai wilayah, mulai dari perkotaan hingga pedesaan. Perusahaan *Dream Housing Finance* menangani berbagai permasalahan dengan program Kredit Pemilikan Rumah (KPR) dan alternatif pinjaman berdasarkan data yang diisi konsumen. Nasabah dapat mengajukan pinjaman setelah perusahaan memvalidasi kelayakan nasabah untuk mendapatkan pinjaman. *Dream Housing Finance* memiliki tantangan dalam menangani jumlah pinjaman yang terus meningkat. Jumlah pinjaman yang meningkat membuat perusahaan kesulitan dalam mengolah data konsumen, sehingga perusahaan ingin mengoptimalkan sistem yang ada melalui proses kelayakan pinjaman *realtime* berdasarkan detail data konsumen yang diberikan saat mengisi formulir aplikasi. Perusahaan telah memiliki aturan untuk mengidentifikasi segmen nasabah yang memenuhi syarat untuk pinjaman sehingga perusahaan bisa dengan tepat menargetkan pinjaman sesuai dengan data yang diisi oleh konsumen (Sanjaya dkk, 2020).

Perusahaan ingin mengotomatiskan proses kelayakan pinjaman *realtime* berdasarkan detail nasabah yang diberikan saat mengisi formulir aplikasi. Penerapan metode *machine learning* pada proses seleksi nasabah yang diberikan pinjaman dapat menjadi terobosan yang efektif untuk mengetahui kelayakan nasabah dalam mendapatkan pinjaman.

Machine learning merupakan suatu sistem yang dibangun berdasarkan pembelajaran pada data yang besar. *Machine learning* dilatih untuk mendapatkan hasil yang akurat. Salah satu contoh *machine learning* yaitu *Adaptive Boosting* (Triwahyu, 2023). *Adaptive Boosting* (AdaBoost) adalah metode *ensemble* yang dapat meningkatkan kinerja prediksi dengan menggabungkan hasil beberapa model, dengan cara memberikan bobot pada data yang kemudian akan dimaksimalkan pada iterasi selanjutnya sehingga dapat menghasilkan model yang baik. Penelitian sebelumnya yang dijelaskan oleh Sanjaya dkk. (2020) dengan judul “Prediksi Kelalaian Pinjaman Berbasis *machine learning* Menggunakan Metode *Random Forest* dan *Adaptive Boosting*” memperoleh hasil bahwa metode AdaBoost dapat meningkatkan klasifikasi lemah menjadi klasifikasi kuat. Dalam penelitian ini, penulis akan mengoptimalkan model data dengan pendekatan AdaBoost untuk meningkatkan akurasi dan ketepatan dalam menentukan kelayakan pinjaman.

Metode *Adaptive Boosting* adalah algoritma *supervised* dalam data *mining* yang secara luas diterapkan untuk pembuatan model klasifikasi. Proses AdaBoost melibatkan model *weak learner* seperti pohon keputusan dengan setiap observasi diberi bobot yang sama. Jika prediksi yang pertama dibuat oleh *weak learner* tidak akurat, bobot pada observasi yang salah akan ditingkatkan secara berulang sehingga terbentuk prediksi benar. Metode Adaboost fokus pada data yang sulit diprediksi oleh model sebelumnya sehingga model AdaBoost memastikan model selanjutnya memberikan lebih banyak perhatian pada model yang kurang sesuai (Zulhanif, 2016). Kelebihan dari metode AdaBoost adalah kemampuannya dalam menangani data yang kompleks dengan cara yang efisien. AdaBoost tidak hanya

meningkatkan akurasi model, tetapi juga mengurangi overfitting pada model. AdaBoost dapat dengan mudah diimplementasikan dalam berbagai jenis model *machine learning*, sehingga menjadikannya sangat mudah diterapkan pada klasifikasi dan regresi. Beberapa penelitian sebelumnya yang dilaksanakan oleh Bisri & Wahono (2015) dan Nurzahputra (2017) juga menggunakan metode AdaBoost dalam topik penelitiannya.

AdaBoost memiliki beberapa parameter yang perlu diatur untuk menghasilkan model yang optimal yang dikenal sebagai *hyperparameter* yaitu *learning rate* dan *n estimator*. Metode *tuning hyperparameter* dilakukan dengan memilih model dari berbagai kombinasi nilai *hyperparameter* yang sesuai. Metode *hyperparameter* yang digunakan adalah *GridSearchCV* dan *RandomSearchCV*. *RandomSearchCV* adalah proses pencarian *hyperparameter* yang memilih kombinasi *hyperparameter* secara acak dari parameter yang telah ditentukan sebelumnya. Proses ini mencoba beberapa kombinasi *hyperparameter* secara acak, sehingga memungkinkan eksplorasi yang lebih efisien dari ruang *hyperparameter* yang besar. *GridSearchCV* adalah proses pencarian *hyperparameter* dengan memilih semua kombinasi parameter yang telah ditentukan sebelumnya sehingga dapat lebih optimal dalam menentukan model yang terbaik. Proses ini mencoba semua kombinasi *hyperparameter* untuk memastikan eksplorasi ruang *hyperparameter* yang menyeluruh (Pramudhyta & Rohman, 2024). *GridSearchCV* melakukan pencarian *hyperparameter* dengan mencoba setiap kombinasi dari parameter yang telah ditentukan. *GridSearchCV* dapat menemukan kombinasi parameter terbaik karena mengeksplorasi seluruh ruang parameter yang diuji, namun *GridSearchCV* bisa memakan waktu jika

ruang parameter yang diuji besar dan model yang digunakan memiliki kompleksitas tinggi. *RandomSearchCV* menawarkan solusi yang lebih efisien dengan memilih kombinasi parameter secara acak dari ruang parameter. Meskipun tidak menjamin menemukan kombinasi terbaik, *RandomSearchCV* sering kali cukup efektif dan efisien, terutama ketika kombinasi parameter yang diuji sedikit. *RandomSearchCV* lebih sesuai untuk pengaplikasian yang memerlukan hasil cepat dengan sumber daya komputasi yang terbatas. Kedua metode tersebut digunakan untuk mendapatkan parameter terbaik pada suatu model, sehingga model tersebut mampu memprediksi data secara akurat. Berdasarkan uraian mengenai pentingnya keputusan pemberian pinjaman kepada nasabah, maka penulis menerapkan *Tuning Hyperparameters* pada *Adaptive Boosting* untuk mengoptimalkan prediksi persetujuan pinjaman secara akurat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini terdapat beberapa rumusan masalah yaitu:

1. Bagaimana keakuratan model prediksi menggunakan metode AdaBoost?
2. Bagaimana keakuratan model prediksi setelah dilakukan proses *tuning* menggunakan *RandomSearchCV*?
3. Bagaimana keakuratan model prediksi setelah dilakukan proses *tuning* menggunakan *GridSearchCV*?

1.3 Batasan Masalah

Penelitian ini akan menggunakan metode AdaBoost dengan batasan masalah sebagai berikut:

1. Data *Dream Housing Finance* tahun 2022 diperoleh dari *KAGGLE*.
2. Proses yang digunakan untuk meningkatkan akurasi prediksi adalah *Adaptive Boosting* (Adaboost) dengan *Tuning Hyperparameter RandomSearchCV* dan *GridSearchCV*.

1.4 Tujuan Masalah

Tujuan dari tugas akhir ini adalah:

1. Mendapatkan nilai akurasi terbaik tanpa adanya *overfitting* pada prediksi persetujuan pinjaman menggunakan metode AdaBoost.
2. Mendapatkan nilai akurasi terbaik tanpa adanya *overfitting* menggunakan metode AdaBoost dengan *tuning RandomSearchCV*.
3. Mendapatkan nilai akurasi terbaik tanpa adanya *overfitting* menggunakan metode AdaBoost dengan *tuning GridSearchCV*.