

BAB I PENDAHULUAN

1.1 Latar Belakang

Prediksi kebangkrutan perusahaan merupakan salah satu isu paling krusial dalam manajemen risiko keuangan karena memiliki dampak signifikan bagi investor, kreditor, dan perekonomian secara keseluruhan. Kebangkrutan tidak hanya menyebabkan kerugian finansial yang besar, tetapi juga berdampak sosial, seperti pemutusan hubungan kerja dan gangguan pada rantai pasok industri. Oleh karena itu, kemampuan untuk memprediksi kebangkrutan secara dini sangat penting agar perusahaan dan pemangku kepentingan dapat mengambil langkah preventif untuk meminimalkan risiko dan kerugian yang lebih besar. Faktor penyebab kebangkrutan sangat beragam, mulai dari rasio likuiditas yang menurun, profitabilitas yang melemah, tingginya *leverage* (utang), hingga efisiensi operasional yang rendah, yang menjadikan proses klasifikasi menjadi tantangan [1].

Prediksi kebangkrutan penting karena memberikan peringatan dini yang memungkinkan perusahaan melakukan restrukturisasi, efisiensi, atau perubahan strategi bisnis sebelum kondisi memburuk. Hal ini tidak hanya melindungi perusahaan itu sendiri, tetapi juga investor dan kreditor yang berisiko mengalami kerugian. Selain itu, prediksi yang akurat dapat membantu regulator dan pemerintah dalam menjaga stabilitas ekonomi dengan mengantisipasi potensi kegagalan perusahaan yang dapat berdampak sistemik.

Dalam menghadapi tantangan tersebut, metode *ensemble learning* menjadi salah satu pendekatan yang menjanjikan. *Ensemble learning* menggabungkan beberapa model prediksi untuk menghasilkan hasil yang lebih stabil, akurat, dan handal dibandingkan dengan model tunggal [2]. Pendekatan ini efektif mengatasi *overfitting* dan menangani kompleksitas data keuangan yang tinggi, sehingga sangat cocok untuk prediksi kebangkrutan [3]. Berbagai penelitian menunjukkan bahwa model *ensemble* seperti Random Forest, Gradient Boosting, AdaBoost, dan XGBoost unggul dalam menangkap pola non-linear yang kompleks dalam data keuangan. Dalam sektor keuangan, metode ini efektif dalam prediksi kebangkrutan dan penilaian resiko kredit, khususnya dalam menangani kompleksitas data keuangan [4]. Penelitian terbaru menunjukkan bahwa pendekatan *ensemble* dapat meningkatkan tingkat deteksi kebangkrutan hingga 15% dibandingkan dengan pendekatan model tunggal [5]. Selain di sektor keuangan, metode *ensemble* juga diterapkan dalam berbagai bidang lainnya. Dalam sektor kesehatan, metode ini digunakan untuk diagnosis penyakit, prediksi hasil pengobatan, dan analisis citra medis [6]. Di bidang ilmu lingkungan dan klimatologi, metode *ensemble* digunakan untuk prediksi cuaca,

analisis perubahan iklim, dan pemodelan ekosistem untuk menangani ketidakpastian dan variabilitas yang melekat pada sistem kompleks tersebut [7].

Dalam penelitian ini, digunakan delapan model *ensemble learning* yaitu Gradient Boosting, AdaBoost, XGBoost, CatBoost, Voting Classifier, Bagging, Random Forest, dan LightGBM untuk memprediksi kebangkrutan perusahaan berdasarkan dataset Taiwan Economic Journal (1999–2009) yang mencakup 95 atribut keuangan. Dataset ini dipilih karena tidak hanya sering dijadikan *benchmark* dalam penelitian serupa, tetapi juga memiliki representasi yang kuat terhadap indikator-indikator penyebab kebangkrutan. Selain itu, dataset ini mencerminkan kondisi nyata mencerminkan situasi nyata di dunia bisnis di mana kasus kebangkrutan memang jarang terjadi. Hal ini penting untuk menguji ketangguhan model dalam menghadapi tantangan nyata prediksi kebangkrutan. Fokus utama penelitian ini adalah mengevaluasi kemampuan model dan membandingkan kemampuan model *ensemble* dalam mendeteksi perusahaan bangkrut (kelas minoritas) dengan tetap mempertahankan akurasi umum yang tinggi.

Meskipun metode *ensemble learning* mampu memberikan prediksi yang akurat, hasil prediksi tersebut tidak akan optimal jika tidak dapat diakses dan dimanfaatkan dengan mudah oleh pihak-pihak yang membutuhkannya. Saat ini, masih banyak perusahaan, khususnya skala menengah ke bawah, yang belum memiliki sistem pendukung keputusan berbasis data dalam mengevaluasi kondisi finansial mereka secara berkala. Hal ini menyebabkan proses analisis kebangkrutan sering kali dilakukan secara manual, lambat, dan kurang objektif. Bahkan, beberapa perusahaan baru menyadari adanya potensi kebangkrutan ketika kondisi keuangan sudah sangat kritis.

Oleh karena itu sebagai bentuk implementasi praktis dari model yang dikembangkan, pengembangan *dashboard* interaktif menjadi aspek penting dalam penelitian ini. *Dashboard* memungkinkan hasil prediksi dapat diakses oleh pengguna seperti analis keuangan dan *stakeholder* perusahaan tanpa perlu memahami detail teknis algoritma *machine learning*. Dengan adanya fitur seperti unggah dataset, pemilihan model prediksi, serta tampilan hasil dalam bentuk ringkasan yang mudah dimengerti. Keberadaan *dashboard* juga mendukung proses pengambilan keputusan yang lebih cepat dan berbasis data. Tidak hanya untuk mendeteksi risiko kebangkrutan, tetapi juga untuk membantu perusahaan melakukan evaluasi dini terhadap kesehatan keuangannya.

1.2 Rumusan Masalah dan Solusi

Rumusan masalah yang diangkat dalam Tugas Akhir ini adalah: Bagaimana perbandingan performa delapan model *ensemble learning* dalam mendeteksi kasus kebangkrutan perusahaan?

Solusi yang diusulkan adalah melakukan implementasi komparatif terhadap delapan model *ensemble* yang berbeda yaitu Gradient Boosting, AdaBoost, XGBoost, CatBoost, Voting Classifier, Bagging, Random Forest, dan LightGBM, dengan konfigurasi parameter yang telah dioptimalkan. Tujuannya adalah untuk menentukan model yang paling akurat dalam mendeteksi kasus kebangkrutan perusahaan. Evaluasi performa dilakukan langsung pada data asli tanpa menerapkan teknik balancing seperti *oversampling* atau *undersampling*, guna menjaga keaslian distribusi data dan menggambarkan kondisi nyata di mana kasus kebangkrutan merupakan kelas minoritas yang sangat jarang terjadi. Selain itu, untuk memudahkan penggunaan sistem prediksi ini oleh pengguna non-teknis, dikembangkan pula *dashboard* interaktif berbasis Flask yang memungkinkan input data, pemilihan model, klasifikasi, dan pengunduhan hasil.

1.3 Tujuan

Beberapa tujuan dari penelitian ini yaitu:

1. Mengembangkan model prediksi kebangkrutan berbasis *ensemble learning* yang efektif, akurat, dan mampu mendeteksi perusahaan bangkrut.
2. Membangun sistem dashboard interaktif dengan menggunakan Laravel sebagai antarmuka pengguna dan Flask sebagai REST API prediksi, sehingga memudahkan pengguna dalam mengunggah data, memilih model, menjalankan prediksi, serta melihat dan mengunduh hasil klasifikasi.

1.4 Batasan Masalah

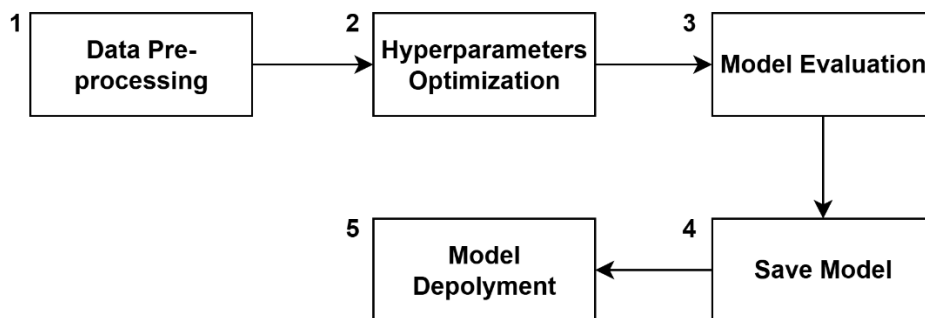
1. Penelitian ini terbatas pada perbandingan delapan metode *ensemble learning*, yaitu Gradient Boosting, AdaBoost, XGBoost, CatBoost, Voting Classifier, Bagging, Random Forest, dan LightGBM. Model-model lain di luar daftar ini tidak dibahas atau dianalisis lebih lanjut.
2. Analisis yang dilakukan hanya berfokus pada klasifikasi biner, yaitu memprediksi status perusahaan sebagai bangkrut atau tidak bangkrut. Penelitian tidak membahas prediksi tingkat risiko kebangkrutan atau klasifikasi multi-kelas.
3. Penelitian ini tidak menerapkan teknik penanganan data *imbalance* seperti SMOTE, *oversampling*, atau *undersampling*. Analisis dilakukan langsung terhadap data asli sebagaimana adanya.
4. Dataset yang digunakan dalam penelitian ini berasal dari Taiwan Economic Journal (TEJ) untuk periode tahun 1999–2009, dan hanya dataset ini yang digunakan dalam proses pelatihan dan pengujian model. Penelitian tidak menggunakan data eksternal lain maupun pembaruan data di luar cakupan tersebut.
5. Pengembangan sistem *dashboard* interaktif difokuskan pada fitur dasar seperti unggah dataset, pemilihan model, dan visualisasi hasil prediksi. Fitur

lanjutan seperti interpretasi model (*explainability*) dan integrasi data real-time belum termasuk dalam cakupan penelitian.

1.5 Metodologi Pengerjaan

Dalam proses pengembangan sistem prediksi kebangkrutan ini, digunakan pendekatan *prototyping* karena metode ini dinilai fleksibel dan responsif terhadap perubahan kebutuhan pengguna. Proses dimulai dengan pembuatan prototipe awal yang mencakup fitur dasar seperti input data, pemilihan model, serta hasil prediksi. Prototipe tersebut kemudian dievaluasi secara berkala dan disesuaikan berdasarkan masukan yang diterima. Tahapan revisi dilakukan secara iteratif hingga sistem berfungsi sesuai dengan kebutuhan dan mudah digunakan. Pendekatan ini memungkinkan pengujian fitur sejak awal, serta mendukung penyempurnaan sistem secara berkelanjutan, baik dari sisi fungsionalitas maupun tampilan antarmuka.

Pekerjaan ini mencakup seluruh tahapan pengembangan model mulai dari *pre-processing* data hingga *deployment* model ke dalam sistem yang siap digunakan. Adapun alur kerja yang dilakukan dalam pengembangan sistem ini adalah sebagai berikut:



Gambar 1.5-1 Alur Pengerjaan

1. Data Preprocessing

Setelah data terkumpul, dilakukan tahap *preprocessing* untuk memastikan kualitas dan kesiapan data sebelum dimasukkan ke dalam model. Tahap ini meliputi *scaling* (normalisasi fitur numerik agar berada dalam skala seragam), serta *splitting* data menjadi dua bagian utama: data *training* dan data *testing*. Pembagian ini bertujuan untuk mengukur kemampuan generalisasi model secara objektif.

2. Hyperparameter Optimization

Melakukan eksperimen terhadap berbagai *algoritma machine learning*, khususnya dari kelompok *ensemble learning*. Beberapa model yang digunakan meliputi: Gradient Boosting, AdaBoost, XGBoost, CatBoost, Voting Classifier, Bagging, Random Forest, dan LightGBM. Tujuannya adalah

untuk membandingkan performa masing-masing model dalam memprediksi kebangkrutan.

Setiap model yang digunakan kemudian melalui proses *hyperparameter tuning* guna memperoleh konfigurasi terbaik yang dapat meningkatkan performa prediksi. Teknik seperti *Grid Search* dan *Random Search* digunakan untuk memilih kombinasi parameter yang paling optimal.

3. Model Evaluation

Setelah proses pelatihan, model diuji menggunakan data uji. Evaluasi dilakukan dengan menggunakan berbagai metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Proses ini bertujuan untuk mengetahui sejauh mana model mampu melakukan prediksi secara andal pada data yang belum pernah dilihat sebelumnya.

4. Save Model

Setelah dilakukan proses evaluasi, model-model yang telah dibuat disimpan menggunakan *library* joblib agar dapat digunakan kembali pada tahap *deployment* di *dashboard*. Berikut merupakan kode untuk save model

5. Model Deployment

Model terbaik yang telah diuji dan dievaluasi kemudian disiapkan untuk tahap *deployment*.

1.6 Penjadwalan Kerja

Jadwal pelaksanaan kegiatan pengembangan sistem prediksi kebangkrutan perusahaan menggunakan *ensemble learning* disusun secara sistematis untuk memastikan tercapainya tujuan penelitian. Berikut adalah rincian jadwal pelaksanaan selama periode Agustus 2024 hingga Mei 2025.

Tabel 1.6-1.1 Tabel Penjadwalan Kerja

No	Deskripsi Kerja	Ags 2024				Sept 2024				Okt 2024				Nov 2024				Des 2024			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Studi Literatur																				
2	Preprocessing Data																				
3	Perancangan Algoritma <i>Ensemble</i>																				
4	Implementasi Model Prediksi																				
5	Evaluasi dan Optimasi Model																				
6	Penulisan Publikasi																				
7	Revisi Paper																				

[illegible]

Tabel 1.6-2. Tabel Penjadwalan Kerja

No	Deskripsi Kerja	Jan 2024				Feb 2025				Mar 2025				Apr 2025				Mei 2025			
		1	2	1	1	2	3	4	2	3	4	3	4	1	2	3	4	1	2	3	4
9	Perancangan Aplikasi Web																				
10	Pengembangan Aplikasi Web																				
11	Integrasi Model dengan Aplikasi																				
12	Penyusunan Dokumen TA																				