

BAB I

PENDAHULUAN

1.1 Latar Belakang

Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki peranan strategis dalam perekonomian Indonesia dengan kontribusi lebih dari 61% terhadap Produk Domestik Bruto (PDB) dan penyerapan 97% tenaga kerja nasional [1]. Namun, tingginya kontribusi tersebut tidak sejalan dengan tingkat keberlanjutan usaha, ditunjukkan oleh laporan bahwa lebih dari 50–60% UMKM berhenti beroperasi dalam tiga tahun pertama [2]. Kegagalan UMKM tidak hanya berdampak pada pelaku usaha, tetapi juga stabilitas ekonomi, tingkat pengangguran, serta efektivitas kebijakan pemerintah dalam pemberdayaan sektor UMKM. Salah satu kendala yang dihadapi UMKM adalah rendahnya pemanfaatan data dalam pengambilan keputusan bisnis [3]. Pelaku UMKM umumnya memiliki berbagai data yang berkaitan dengan karakteristik usaha dan pemilik, seperti usia pemilik, tingkat pendidikan, kecukupan modal awal, pencatatan keuangan, penggunaan teknologi, serta pengalaman usaha. Data-data tersebut berpotensi digunakan untuk membangun model prediksi yang dapat membantu mengidentifikasi karakteristik UMKM berdasarkan data yang tersedia. Namun, dalam praktiknya data tersebut belum banyak dimanfaatkan sebagai dasar analisis dalam mendukung pengambilan keputusan berbasis data [5].

Sejumlah penelitian telah dilakukan terkait prediksi keberhasilan dan kinerja UMKM menggunakan berbagai pendekatan *machine learning*. Penelitian yang dilakukan dengan mengembangkan kerangka kerja dua tahap berupa *clustering* dan klasifikasi untuk memprediksi tingkat keberhasilan UMKM menunjukkan bahwa model terbaik yang diperoleh adalah *K-Nearest Neighbor* [6]. Penelitian lain melakukan prediksi kinerja UMKM manufaktur menggunakan pendekatan dua tahap berupa *Data Envelopment Analysis* (DEA) dan *machine learning*, dengan model terbaik yang diperoleh adalah LASSO [7]. Selain itu, prediksi keberhasilan UMKM juga telah dikembangkan menggunakan algoritma *Random Forest* berdasarkan data historis usaha, serta mampu mengidentifikasi faktor-faktor penting, seperti akses keuangan, karakteristik bisnis, dan dukungan pemerintah, yang berkontribusi terhadap keberhasilan UMKM [8].

Meskipun penelitian sebelumnya [6][7][8] menunjukkan performa prediksi yang tinggi, masing-masing penelitian menggunakan pendekatan yang berbeda, seperti *framework hybrid clustering–classification*, *hybrid DEA–Machine Learning*, serta

algoritma *Random Forest*. Selain itu, setiap penelitian menggunakan karakteristik *dataset* dan tujuan penelitian yang berbeda sehingga hasil yang diperoleh tidak dapat langsung digeneralisasi pada *dataset* UMKM yang digunakan dalam penelitian ini. Oleh karena itu, penelitian ini mengeksplorasi penggunaan ANN untuk mengevaluasi kemampuannya dalam memprediksi keberhasilan UMKM berdasarkan karakteristik *dataset* yang digunakan. Meskipun ANN umumnya menunjukkan performa yang lebih optimal pada *dataset* berukuran besar, penelitian ini bertujuan mengevaluasi kemampuan ANN pada *dataset* UMKM yang tersedia yang terdiri atas 250 data.

ANN merupakan salah satu metode *machine learning* yang dirancang untuk meniru cara kerja sistem saraf manusia dalam mempelajari pola dari data [9]. ANN mampu memodelkan hubungan kompleks dan nonlinier melalui proses pembelajaran berlapis, sehingga banyak digunakan dalam berbagai permasalahan prediksi berbasis data [10]. Penelitian yang dilakukan dengan menerapkan ANN untuk memprediksi kondisi kesulitan finansial perusahaan dan memperoleh akurasi sebesar 88,88%, yang menunjukkan kinerja lebih baik dibandingkan metode *Support Vector Machine (SVM)* dengan akurasi 80,47% serta PSO-SVM dengan akurasi 83,16% [11]. Penelitian lain membandingkan algoritma ANN dan *naive bayes* dalam prediksi kelahiran prematur, dengan hasil akurasi model ANN sebesar 90,67%, lebih tinggi dibandingkan *naive bayes* yang memperoleh akurasi 84,53% [12]. ANN juga digunakan untuk memodelkan dan memprediksi performa akademik mahasiswa dengan memanfaatkan berbagai variabel latar belakang mahasiswa dan model tersebut menghasilkan akurasi prediksi sebesar 84,8% serta menunjukkan kemampuan ANN dalam mengidentifikasi pola hubungan yang kompleks pada data [13]. Hasil-hasil penelitian tersebut [11][12][13] menunjukkan bahwa ANN mampu menghasilkan performa prediksi yang baik dengan tingkat akurasi yang tinggi dan memiliki kemampuan yang baik dalam melakukan proses pembelajaran pola data.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk membangun dan mengevaluasi model prediksi keberhasilan UMKM menggunakan ANN dengan memanfaatkan karakteristik usaha dan pemilik yang terdapat pada *dataset*. Selain menghasilkan model prediksi, penelitian ini juga bertujuan mengidentifikasi variabel-variabel yang paling berpengaruh terhadap hasil prediksi menggunakan metode *Permutation Feature Importance*. Oleh karena itu, penelitian ini diangkat menjadi judul tugas akhir “Prediksi Keberhasilan UMKM Menggunakan *Artificial Neural Network*”.

1.2 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini yaitu tingginya tingkat kegagalan UMKM dalam menjaga keberlanjutan usahanya serta belum optimalnya pemanfaatan data numerik yang dimiliki UMKM sebagai pendukung pengambilan keputusan bisnis.

1.3 Tujuan

Tujuan dari penelitian ini adalah:

1. Membangun dan mengevaluasi model prediksi keberhasilan UMKM menggunakan algoritma ANN.
2. Mengidentifikasi variabel yang paling berpengaruh terhadap prediksi keberhasilan UMKM.

1.4 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat berupa:

1. Membantu prediksi keberhasilan UMKM berdasarkan karakteristik usaha.
2. Menjadi referensi akademik dan praktis dalam pengambilan keputusan berbasis data bagi pelaku dan pendamping UMKM.

1.5 Ruang Lingkup

Adapun yang menjadi ruang lingkup dari penelitian ini adalah:

1. *Dataset* yang digunakan pada penelitian ini berasal dari *platform* Kaggle dan dapat diakses melalui tautan berikut: [tautan dataset](#). *Dataset* terdiri dari 250 baris data dan 13 kolom (atribut). Variabel yang terdapat di dalamnya meliputi data numerik, biner, dan ordinal yang berkaitan dengan karakteristik usaha dan pemiliknya.
2. Variabel yang digunakan meliputi usia pemilik, tingkat pendidikan, kecukupan modal awal, pencatatan keuangan, penggunaan internet, perencanaan bisnis, usaha pemasaran, kemitraan, pengalaman bisnis orang tua, pengalaman industri, jenis kelamin pemilik, dan frekuensi konsultasi profesional.
3. Variabel target yang menjadi luaran model adalah *status_keberhasilan* (0 = tidak berhasil dan 1 = berhasil).
4. Metrik evaluasi menggunakan *confusion matrix* dan *classification report* dengan metrik *accuracy*, *precision*, *recall* dan *F1-score*, serta membandingkan kinerja model pada data *training* dan data *testing*. Pemilihan metrik ini didasarkan pada relevansi

dalam kasus klasifikasi biner serta kemampuannya dalam menggambarkan keseimbangan antara prediksi benar dan kesalahan model.

5. Analisis fitur yang paling berpengaruh terhadap hasil prediksi menggunakan *Permutation Features Importance*.

