

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Banyak industri telah mengalami perubahan dramatis sebagai akibat dari kondisi teknologi digital saat ini. Sistem informasi telah digunakan oleh hampir semua industri untuk merampingkan, mempercepat, dan meningkatkan efisiensi operasional. Teknologi ini tidak hanya merampingkan operasi, tetapi juga menghasilkan data dan informasi yang lebih tepat yang dapat digunakan untuk membuat keputusan strategis [1].

Salah satu sektor penting yang turut merasakan dampak transformasi ini adalah dunia pendidikan. Sebagai fondasi pembentukan kepribadian dan kecerdasan bangsa sebagaimana diamanatkan dalam UUD 1945, institusi pendidikan dituntut untuk terus beradaptasi dan meningkatkan kualitas layanannya. Pemanfaatan sistem informasi terkomputerisasi dalam administrasi pendidikan menjadi sebuah keniscayaan untuk mencapai tujuan tersebut, memungkinkan pengelolaan yang lebih efektif dan transparan [2].

Pesantren yang menganut prinsip-prinsip Islam memiliki kurikulum, tenaga pengajar, dan administrasi sendiri. Sistem administrasi keuangan, khususnya pembayaran Sumbangan Pembinaan Pendidikan (SPP) oleh santri, merupakan komponen penting dalam manajemen pesantren. Permasalahan yang sering dihadapi adalah keterlambatan dalam pembayaran SPP, yang dapat berdampak pada stabilitas keuangan lembaga.

Pada saat ini, proses manual digunakan untuk memulihkan pembayaran yang tertunggak. Hal ini menyulitkan manajemen untuk mengetahui siswa mana yang mungkin menunggak pembayaran dan membuat manajemen tidak memiliki mekanisme pendukung keputusan untuk membantu mereka mengambil tindakan pencegahan. Oleh karena itu, dibutuhkan suatu pendekatan berbasis teknologi untuk membantu mengklasifikasikan keterlambatan pembayaran SPP.

Di tengah kemajuan tersebut, belum semua institusi pendidikan mampu mengadopsi sistem informasi secara optimal. Salah satunya adalah lingkungan pondok pesantren, yang seringkali masih mengandalkan sistem administrasi manual. Pondok Pesantren Al Fajar Tegal, sebagai salah satu lembaga pendidikan Islam, menghadapi tantangan spesifik dalam hal pengelolaan Sumbangan Pembinaan Pendidikan (SPP) santri. Keterlambatan pembayaran SPP menjadi masalah yang kerap terjadi dan dapat berdampak serius terhadap stabilitas keuangan serta kelangsungan operasional pesantren. Tanpa pengelolaan yang baik, kualitas pelayanan pendidikan bagi para santri dapat terganggu [3].

Untuk mengatasi permasalahan tersebut secara efektif, diperlukan sebuah pendekatan yang berbasis data. Pemanfaatan teknologi informasi, khususnya dalam bidang *data mining*, menawarkan solusi untuk mengidentifikasi pola-pola keterlambatan pembayaran berdasarkan data historis. Dengan melakukan klasifikasi terhadap data santri, pihak pesantren dapat mengenali faktor-faktor yang memengaruhi keterlambatan dan mengambil tindakan preventif yang tepat sasaran sebelum masalah meluas.

Kategorisasi data hanyalah salah satu area di mana algoritma lain mungkin berguna; mereka memiliki tujuan yang sama tetapi menggunakan pendekatan yang berbeda. Beberapa contoh di antaranya adalah *SVM*, *Decision Tree*, dan *Naive Bayes*. Agar *Decision Tree* dapat berfungsi, algoritma ini membangun sebuah pohon keputusan dengan menggunakan kualitas yang paling berguna untuk membedakan berbagai jenis data. Proses pembuatan aturan *Decision Tree* untuk kategorisasi lebih transparan dibandingkan dengan K-NN berbasis jarak. Selain itu, ada juga metode seperti *SVM*, yang menentukan *hyperplane* optimal untuk membagi kelas dalam ruang multidimensi, dan *Naive Bayes*, yang menggunakan kemungkinan fitur untuk mengklasifikasikan data dengan mengasumsikan independensi fitur. Semuanya tergantung pada spesifikasi dataset ketika memutuskan metode mana yang akan digunakan. Ketika membangun sebuah model, faktor-faktor seperti kompleksitas data, jumlah fitur, serta akurasi dan kecepatan harus dipertimbangkan ketika memilih sebuah algoritma.

Untuk penelitian ini, *K-Nearest Neighbor (K-NN)* akan menjadi algoritma klasifikasi yang akan digunakan. Kesederhanaan dan keefektifan algoritma ini dalam mengidentifikasi data baru berdasarkan kedekatannya dengan data yang sudah ada menjadi alasan pemilihan algoritma ini [4]. Kemungkinan keterlambatan siswa dapat diprediksi dengan menggunakan K-NN yang diterapkan pada data pembayaran uang sekolah mereka dan variabel terkait lainnya seperti riwayat pembayaran dan tingkat akademis.

Dengan menggunakan teknik K-NN, penelitian ini bermaksud untuk membuat dan menjalankan sistem untuk mengkategorikan pembayaran uang sekolah yang terlambat di Pondok Pesantren Al Fajar Tegal. Proyek ini bertujuan untuk memberikan solusi praktis untuk manajemen keuangan pesantren dan memberikan kontribusi ilmiah pada bidang analisis data dalam pendidikan Islam. Di antara banyak manfaat dari pendekatan K-NN adalah kesederhanaannya, kemampuannya untuk memproses data dengan karakteristik yang sedikit, dan tidak adanya model pelatihan yang rumit. Namun demikian, ada beberapa kelemahan dalam menggunakan K-NN. Salah satunya adalah bahwa K-NN sensitif terhadap ukuran data dan membutuhkan normalisasi sebelum membuat prediksi. Masalah lainnya adalah bahwa K-NN mungkin cukup lamban dengan dataset yang besar karena K-NN harus menghitung jarak antara semua data pelatihan dan data pengujian. Overfitting dan underfitting, yang berdampak pada akurasi model, dapat terjadi akibat penggunaan nilai K yang terlalu kecil atau terlalu besar, oleh karena itu memilih nilai K yang tepat adalah kesulitan lain. Solusi yang dapat diterapkan untuk mengatasi hal ini adalah dengan melakukan evaluasi menggunakan metode validasi silang (*cross-validation*), seperti *K-Fold Cross Validation*, guna menentukan nilai K yang memberikan performa terbaik secara konsisten.

## **1.2. Rumusan Masalah**

Berdasarkan uraian latar belakang di atas maka terdapat permasalahan yang akan di bahas pada penelitian ini dalam bentuk rumusan masalah. Rumusan masalah pada penelitian menggunakan algoritma *K-Nearest Neighbor (KNN)*

untuk mengklasifikasi keterlambatan pembayaran Sumbangan Pembinaan Pendidikan (SPP) sebagai berikut:

1. Bagaimana penerapan algoritma *K-Nearest Neighbor (K-NN)* dalam membangun model klasifikasi untuk mengidentifikasi keterlambatan pembayaran SPP di Pesantren Al Fajar Tegal?
2. Seberapa akurat model klasifikasi K-NN dalam memprediksi status keterlambatan pembayaran SPP, dan bagaimana efektivitasnya dalam mendukung analisis data tersebut?
3. Bagaimana metode *K-Fold Cross Validation* digunakan untuk mengevaluasi performa dan stabilitas akurasi model K-NN saat diuji dengan berbagai variasi pembagian data?

### **1.3. Batasan Masalah**

Agar penelitian ini lebih terarah dan sesuai dengan tujuan yang ingin dicapai, maka diperlukan batasan masalah sebagai berikut:

1. Penelitian ini hanya difokuskan pada klasifikasi keterlambatan pembayaran SPP santri di Pesantren Al Fajar Tegal.
2. Data yang digunakan merupakan data historis pembayaran SPP santri yang telah dikumpulkan dari sistem administrasi pesantren.
3. Metode klasifikasi yang digunakan dalam penelitian ini adalah *K-Nearest Neighbor (K-NN)*.
4. Variabel yang dianalisis mencakup data seperti penghasilan orang tua, pekerjaan orang tua, jenis kelamin, dan riwayat keterlambatan pembayaran sebelumnya.
5. Penelitian ini hanya menggunakan algoritma K-NN tanpa melakukan perbandingan dengan algoritma klasifikasi lainnya.
6. Evaluasi performa model dilakukan dengan menggunakan metode *split data* dan *K-Fold Cross Validation*, serta metrik evaluasi berupa akurasi dan *error rate*.
7. Implementasi dan pengujian dilakukan menggunakan Google Colab dengan bahasa pemrograman *Python*.

#### 1.4. Tujuan Penelitian

Dari rumusan masalah yang ada maka terdapat tujuan dalam penelitian sebagai berikut:

1. Menerapkan algoritma *K-Nearest Neighbor (K-NN)* dalam membangun model klasifikasi yang dapat mengidentifikasi keterlambatan pembayaran SPP santri di Pesantren Al Fajar Tegal.
2. Mengukur tingkat akurasi model K-NN dalam memprediksi status keterlambatan pembayaran (tepat atau terlambat), serta menilai efektivitas model dalam menganalisis pola data keterlambatan.
3. Menerapkan metode *K-Fold Cross Validation* untuk mengevaluasi performa dan kestabilan akurasi model K-NN, sehingga hasil evaluasi dapat menjadi dasar yang andal dalam pengambilan kebijakan oleh pihak pesantren.

#### 1.5. Manfaat Penelitian

Dalam penelitian yang dilakukan, terdapat manfaat penelitian yaitu sebagai berikut:

1. Penelitian ini diharapkan dapat menjadi solusi pendukung pengambilan keputusan dalam pengelolaan administrasi pembayaran SPP santri, khususnya dalam mengidentifikasi potensi keterlambatan pembayaran secara prediktif. Dengan demikian, pihak pesantren dapat melakukan tindakan pencegahan secara lebih efisien dan terarah.
2. Hasil dari penelitian ini dapat membantu meningkatkan efektivitas manajemen keuangan di Pesantren Al Fajar, memperbaiki sistem pencatatan administrasi, serta memperkuat kontrol terhadap ketepatan pembayaran santri secara berkelanjutan.
3. Penelitian ini menjadi sarana pengembangan pengetahuan dan keterampilan dalam penerapan metode klasifikasi berbasis data mining, khususnya algoritma *K-Nearest Neighbor (K-NN)*, serta sebagai pengalaman dalam menerapkan keilmuan teknologi informasi dalam konteks nyata di dunia pendidikan.
4. Penelitian ini dapat menjadi referensi awal dalam mengembangkan penelitian serupa, baik dengan menggunakan algoritma K-NN maupun dalam

membandingkan efektivitas berbagai metode klasifikasi lainnya dalam konteks pengelolaan administrasi pendidikan.