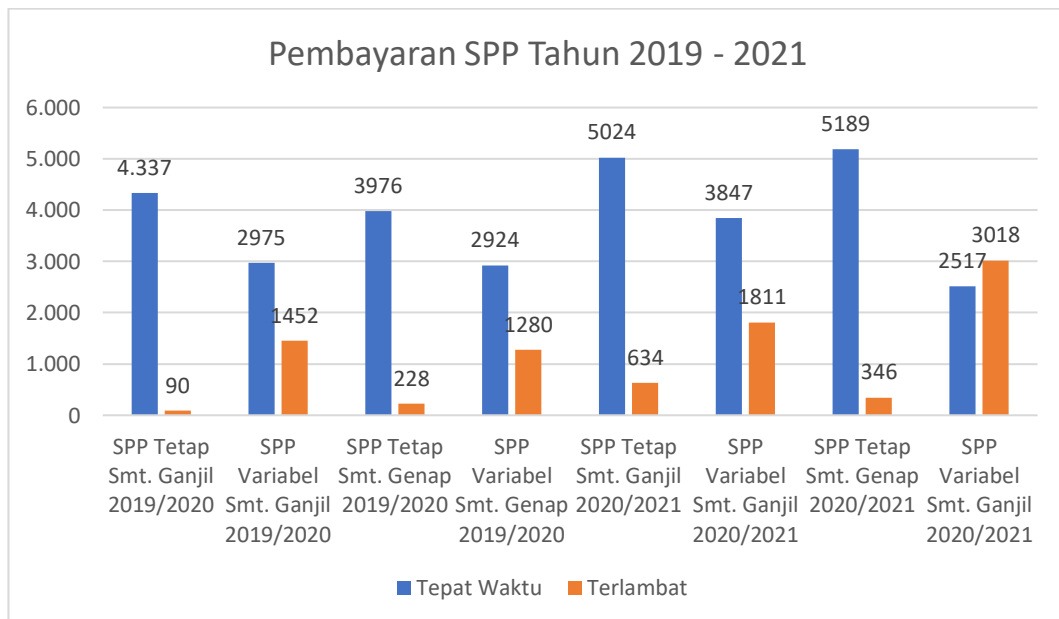


# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Universitas Muhammadiyah Kalimantan Timur (UMKT) merupakan salah satu lembaga pendidikan yang dimiliki oleh persyarikatan muhammadiyah di Provinsi Kalimantan Timur yang pembiayaannya operasionalnya dibebankan kepada mahasiswa, terutama pada pembayaran Sumbangan Pembangunan Pendidikan (SPP). SPP menjadi sumber pendapatan memerlukan perhatian dan pengawasan yang baik. Apalagi pendapatan yang berasal dari SPP menjadi sumber pendapatan utama, terutama untuk perguruan tinggi UMKT.



**Gambar 1.1 Pembayaran SPP Di UMKT Tahun 2019-2021**

Berdasarkan gambar 1.1 dalam kurun waktu 3 tahun terakhir terjadi peningkatan keterlambatan pembayaran SPP yang signifikan di tahun 2020-2021 sebanyak 3.018 mahasiswa dari total keseluruhan 5.535 mahasiswa. Dampak dari keterlambatan pembayaran SPP akan menjadi kendala untuk mengalokasikan biaya pembangunan perguruan tinggi seperti, biaya untuk operasional kegiatan akademik, biaya untuk dosen dan karyawan. Jika tidak melakukan pembayaran SPP sesuai waktu yang ditentukan, maka akibatnya mahasiswa tidak dapat

mengambil mata kuliah (KRS), mencetak kartu ujian dan mengikuti proses perkuliahan semester selanjutnya.

Dari masalah yang sudah dipaparkan diatas maka penting dilakukan sebuah analisis mendalam terkait analisis prediksi agar keterlambatan pembayaran SPP dapat diminimalisir dengan cara menindaklanjuti hasil prediksi yang dinyatakan terlambat. Dengan adanya prediksi dapat memberikan peringatan lebih awal dalam pembayaran SPP sehingga dapat meminimalisir resiko dalam pengelolaan keuangan untuk operasional di UMKT. Salah satu teknik yang dapat digunakan yaitu menggunakan teknik *data mining*, *data mining* merupakan aktivitas pengumpulan informasi atau pengetahuan penting yang berasal dari suatu set data yang besar, informasi yang didapatkan berasal dari *data mining* ini bisa digunakan untuk memperbaiki pengambil keputusan (Santosa & Umam, 2018).

Berikut adalah beberapa penelitian sebelumnya terkait keterlambatan pembayaran SPP seperti dilakukan oleh (Apandi *dkk*, 2019) menggunakan algoritma C4.5. Penelitian (Abdullah, 2019) menggunakan algoritma *k-nearest neighbor*. Penelitian (Ginting *dkk*, 2020) menggunakan algoritma C4.5. Penelitian (Istiana, 2018) menggunakan algoritma C4.5. Penelitian (Muqorobin *dkk*, 2019) menggunakan *algoritma naïve bayes* dengan *feature selection information gain*.

Algoritma *random forest* juga telah banyak digunakan dalam berbagai studi kasus salah satunya di bidang pendidikan seperti yang dilakukan oleh (Rianto & Yunis, 2021) tentang memprediksi jumlah mahasiswa baru menghasilkan tingkat akurasi sebesar 99,8%. Penelitian (Suwardika & Suniantara, 2019) tentang memprediksi ketidaktepatan waktu kelulusan mahasiswa Universitas Terbuka menghasilkan akurasi sebesar 93,23%. Penelitian (triwidiandi *dkk*, 2021) tentang membandingkan metode *decision tree*, *naïve bayes*, dan *random forest* untuk memprediksi prestasi siswa di SMKN 1 Gadingrejo Pringsewu Lampung. Hasil dari pengujian dengan metode *decision tree* menghasilkan akurasi sebesar 94,5%, *naïve bayes* sebesar 96,3%, dan *random forest* sebesar 97,3%. Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa metode terbaik untuk memprediksi prestasi siswa di SMK 1 Gadingrejo Pringsewu Lampung adalah metode *random*

*forest* dan menghasilkan akurasi sebesar 97,3%. Penelitian (Linawati *dkk*, 2020) tentang memprediksi prestasi akademik mahasiswa menggunakan algoritma *random forest* dan C4.5. Hasil dari pengujian dengan metode C4.5 menghasilkan akurasi sebesar 87,1%, sedangkan metode *random forest* menghasilkan akurasi sebesar 92,4%. Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa metode terbaik untuk memprediksi prestasi akademik mahasiswa adalah metode *random forest* dan menghasilkan akurasi sebesar 92,4%. Penelitian (Rosady, 2021) tentang memprediksi kelulusan mahasiswa teknik informatika UIN Suska Riau menghasilkan akurasi sebesar 92,59%. Berdasarkan penelitian yang telah dipaparkan dapat disimpulkan bahwa algoritma *random forest* sangat mumpuni dalam melakukan prediksi di bidang pendidikan dengan memperoleh tingkat akurasi yang baik.

Berdasarkan penelitian sebelumnya tentang prediksi keterlambatan pembayaran SPP sudah banyak dilakukan dan metode yang digunakan berbagai macam seperti C4.5, *k-nearest neighbor*, dan *naïve bayes*. Namun untuk penelitian ini metode yang digunakan ialah *random forest*.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang masalah yang telah dijabarkan diatas maka penulis membuat rumusan masalah berapa persentase akurasi dari algoritma *random forest* dalam prediksi keterlambatan pembayaran SPP di Universitas Muhammadiyah Kalimantan Timur berdasarkan data Biro Administrasi Keuangan dan Biro Administrasi Akademik?

## **1.3. Tujuan Penelitian**

Adapun tujuan penelitian adalah sebagai berikut:

- a) Menentukan indikator yang mempengaruhi keterlambatan mahasiswa dalam membayar SPP.
- b) Mengimplementasikan algoritma *random forest* untuk memprediksi keterlambatan mahasiswa dalam membayar SPP di Universitas Muhammadiyah Kalimantan Timur.

- c) Mengevaluasi hasil kinerja algoritma *random forest* dengan pembagian data 90% data *training* dan 10% *testing* menggunakan *confusion matrix* untuk menghitung nilai akurasi.

#### **1.4. Batasan Masalah**

Agar lingkup permasalahan yang dibahas tidak meluas, maka penulis membatasi permasalahan penelitian sebagai berikut:

- a) Data mahasiswa dan data pembayaran mahasiswa yang dianalisis dan diolah yaitu data pada tahun 2019-2021.
- b) Mengolah dan menganalisa data keterlambatan pembayaran SPP mahasiswa Universitas Muhammadiyah Kalimantan Timur menggunakan aplikasi *RapidMiner*.