

## ABSTRAK

Berkembangnya jaman sekarang ini peran teknologi informasi sangatlah penting, dimana kita sebagai pengguna dapat memanfaatkan teknologi informasi dengan hanya sekali klik. Informasi yang diolah tidak hanya text, gambar, video namun bisa berbentuk multimedia, kita dapat memanfaatkan komputer dengan berbagai cara, mulai menulis, menggambar, edit foto, memutar video/lagu sampau dengan menganalisis data penelitian maupun mengatasi masalah-masalah lainnya. Knowledge Discovery in Database (KDD) adalah keseluruhan proses untuk mencari dan mengidentifikasi pola dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti. Algoritma Naïve Bayes menghitung probabilitas kejadian masa dating dari kejadian sebelumnya di mana masing- masing variable dianggap tidak saling tergantung. Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Penelitian ini dilakukan di Politeknik LP3I Kampus Padang pada program studi Diploma Tiga (DIII) Manajemen Informatika angkatan 2020/2021. Berdasarkan kerangka kerja penelitian yang terdapat pada metodologi penelitian di Bab III, bahwa tahapan kerja terdiri dari mengidentifikasi masalah, menganalisa permasalahan, menentukan tujuan dari penelitian, mempelajari literatur yang digunakan, menganalisa data metode algoritma Naive Bayes, perancangan sistem prediksi keterlambatan pembayaran uang kuliah, implementasi sistem prediksi keterlambatan pembayaran uang kuliah serta hasil dan pembahasan. Data dan Algoritma Naive Bayes yang sudah dianalisis akan diimplementasikan menggunakan bahasa pemrograman PHP dan database MySQL berbasis website. Berdasarkan hasil perhitungan dengan metode Naïve Bayes maka untuk pengujian algoritma dapat dilakuan dengan confusion matrix. Berdasarkan pengujian yang telah dilakukan, maka didapatkan hasil akurasi 80%, presisi 67%, dan recall 33%. Penentuan data testing yang digunakan dapat mempengaruhi hasil pengujian, karena probabilitas yang dimiliki oleh model akan digunakan untuk menentukan kelas pada data testing, sehingga besar kecilnya nilai akurasi, precision dan recall juga dipengaruhi oleh penentuan data testing.

**Kata Kunci : *Teknologi, KDD, Algoritma Naive Bayes, Prediksi, Keterlambatan***

## ABSTRACT

In today's development, the role of information technology is very important, where we as users can utilize information technology with just one click. The information processed is not only text, images, videos but can be in the form of multimedia, we can use computers in various ways, from writing, drawing, editing photos, playing videos/songs to analyzing research data and solving other problems. Knowledge Discovery in Database (KDD) is the entire process of searching for and identifying patterns in data, where the patterns found are valid, new, useful and understandable. The Naïve Bayes algorithm calculates the probability of future events from previous events where each variable is considered independent of each other. Naive Bayes is based on the simplifying assumption that attribute values are conditionally independent of each other given the output values. This research was conducted at the LP3I Polytechnic Padang Campus in the 2020/2021 Diploma Three (DIII) Informatics Management study program. Based on the research framework contained in the research methodology in Chapter III, the work stages consist of identifying problems, analyzing problems, determining the objectives of the research, studying the literature used, analyzing data using the Naive Bayes algorithm method, designing a prediction system for late payment of tuition fees, implementing prediction system for late payment of tuition fees as well as results and discussion. The data and Naive Bayes algorithm that have been analyzed will be implemented using the PHP programming language and a website-based MySQL database. Based on the results of calculations using the Naïve Bayes method, algorithm testing can be carried out using a confusion matrix. Based on the tests that have been carried out, the results obtained are 80% accuracy, 67% precision and 33% recall. Determining the testing data used can influence the test results, because the probability possessed by the model will be used to determine the class of the testing data, so that the size of the accuracy, precision and recall values are also influenced by the determination of the testing data.

Keywords: *Technology, KDD, Naive Bayes Algorithm, Prediction, Delay*