

ABSTRAK

Hasil prestasi akademik siswa pada mata pelajaran Bahasa Inggris sangatlah beragam. Faktor lingkungan yang mempengaruhi hasil prestasi akademik itu pun juga sangat beragam. Berangkat dari kasus tersebut maka dilakukan penelitian untuk mengetahui faktor lingkungan yang mempengaruhi prestasi akademik siswa terhadap mata pelajaran Bahasa Inggris. Adapun faktor lingkungan yang diteliti apakah mempengaruhi nilai Bahasa Inggris atau tidak adalah mengikuti bimbel, dukungan keluarga, lama belajar, ketersediaan internet, kondisi kesehatan, pendidikan ayah, pendidikan ibu, mengikuti ekstrakurikuler, pengasuh, dan jumlah keluarga inti. Data yang digunakan adalah data pribadi siswa kelas X di SMA 1 Bayat Klaten tahun 2020 dan diselesaikan menggunakan metode *Correlated Naïve Bayes* beserta SMOTE. Kemudian data diuji menggunakan beberapa fold yaitu *3-fold cross validation*, *5-fold cross validation*, *7-fold cross validation*, *9-fold cross validation* dan *10-fold cross validation*. Berdasarkan hasil penelitian akurasi terbaik diperoleh ketika menggunakan *7-fold cross validation* yaitu sebesar 65,36%.% ketika tanpa SMOTE dan 64,50% ketika menggunakan SMOTE. Atribut yang paling mempengaruhi prestasi akademik siswa pada mata pelajaran Bahasa Inggris adalah ketersediaan internet.

Kata Kunci :Klasifikasi, Prestasi Akademik Siswa, *Correlated Naïve Bayes*, SMOTE, *K-Fold Cross Validation*.

ABSTRACT

The results of students' academic achievements in English subjects are very diverse. Environmental factors that affect the results of academic achievement are also very diverse. Departing from this case, a research was conducted to determine environmental factors that affect students' academic achievement in English subjects. The environmental factors studied whether influencing the value of English or not are attending tutoring, family support, length of study, internet availability, health conditions, father's education, mother's education, attending extracurricular activities, caregivers, and number of nuclear families. The data used is the personal data of class X students at SMA 1 Bayat Klaten in 2020 and completed using the Correlated Naïve Bayes method along with SMOTE. Then the data is tested using several folds, namely 3-fold cross validation, 5-fold cross validation, 7-fold cross validation, 9-fold cross validation and 10-fold cross validation. Based on the research results, the best accuracy was obtained when using 7-fold cross validation, which was 65.36%.% when without SMOTE and 64.50% when using SMOTE. The attribute that most influences students' academic achievement in English subjects is the availability of the internet.

Keywords : *Classification, Student Academic Achievement, Correlated Naïve Bayes, SMOTE, K-Fold Cross Validation..*