

INTISARI

PENDEKATAN DATA MINING UNTUK ANALISIS KEPRIBADIAN REMAJA BERDASARKAN PENGUNAAN DATA SELULER

Edo Syahputra
12/340063/PA/15114

Psikoinformatika merupakan salah satu sub disiplin ilmu baru yang terlahir dari kombinasi disiplin ilmu psikologi dan informatika. Berbeda dengan informatika sosial yang membahas mengenai masyarakat di era digital secara umum, psikoinformatika mengupas hubungan manusia sebagai individu dengan perkembangan teknologi khususnya teknologi informasi.

Penelitian psikoinformatika yang dilakukan menggunakan teknik penggalian data untuk menentukan kepribadian seseorang berdasarkan tes *Myers-Briggs Type Indicator* dan menganalisis pola penggunaan data seluler. Algoritma yang digunakan meliputi algoritma klasifikasi Naïve Bayes, pohon keputusan tipe C4.5 dan algoritma asosiasi FP-Growth yang biasa digunakan untuk analisis “keranjang belanja”. Berbagai metode pra-proses digunakan untuk memperoleh hasil terbaik. Normalisasi data juga dilakukan untuk mengecilkan jarak yang diakibatkan dari persebaran data yang terlalu lebar.

Pada pengujian dengan Naïve Bayes, akurasi tertinggi didapatkan ketika dilakukan eksperimen terhadap sub kepribadian *perceiver* dan *judger* hingga 66.15% dengan mempertimbangkan jenis kelamin dan usia seseorang. Pembangunan pohon keputusan yang terbaik terjadi ketika jenis data yang digunakan diubah ke dalam bentuk persentase, artinya nilai menggunakan data dalam *megabyte* diubah ke dalam proporsi tertentu bagi tiap orang. Hasil yang didapatkan menunjukkan akurasi sebesar 74.34%.

Analisis asosiasi dengan algoritma FP-Growth yang digunakan pada penelitian ini menghasilkan beberapa *strong rules* dan menunjukkan *frequent itemset* yang dalam hal ini adalah sekumpulan pilihan aplikasi yang paling sering digunakan oleh remaja di Kota Yogyakarta. Hasilnya, aplikasi seperti Instagram, Line, Whatsapp dan Google Chrome menjadi pilihan remaja perempuan sedangkan remaja laki-laki lebih sering menggunakan Facebook dan Youtube.

Kata kunci: MBTI, data mining, Naïve Bayes, pohon keputusan C4.5, FP-Growth

ABSTRACT

DATA MINING APPROACH FOR ANALYZING PERSONALITY IN TEENAGERS BASED ON CELLULAR DATA USAGE

Edo Syahputra
12/340063/PA/15114

Psycho-informatics is a new emerging sub-discipline from the combination of psychology and informatics. In contrast with social informatics, which discuss about human as a society in digital world generally, psycho-informatics is focused on the relationship between human as an individual with technological development especially in technology of information.

This research is try to do experiment in psycho-informatics area by implementing data mining technique to determine someone's character based on Myers-Briggs Type Indicator test by analyze their celullar data usage pattern. The algorithm which are used in this research including classification algorithm of Naïve Bayes, decision tree type C4.5 and FP-Growth as association algorithm which is very popular in "market basket" analysis. Many pre-processing methods were conducted to ensure the best result. Data normalization also conducted to minimize the gap between datas from one instances to others.

In Naïve Bayes experiment, the highest accuracy is shown when the model classify sub-character of "perceiver" and "judger" from dataset with the value up to 66.15%. This result also consider gender and age attribute as addition in dataset which is designed to only analyze people's character by their cellular data usage.

The best decision tree from this experiment is built when the value of datas in each attribute are converted from megabyte into percentage. This percentage represent the proportion of data usage from each instances. The accuracy of this decision tree is 74.34%.

Association analysis with FP-Growth algorithm which is used in this research, resulted in showing some strong rules and frequent itemset of mobile applications that chosen by teenagers in Yogyakarta City. Instagram, Line Whatsapp and Google Chrome are the most popular applications for female teenagers while the male teenagers prefer Facebook and Youtube.

Keywords: MBTI, data mining, Naïve Bayes, Decision Tree C4.5, FP-Growth