

## ABSTRAK

Perdagangan elektronik merupakan pertukaran yang menggunakan teknologi sebagai perantara antara berbagai pihak (individu atau organisasi) serta kegiatan pertukaran intra atau antar organisasi berbasis elektronik. Antar muka yang paling sering digunakan adalah *web browser*. Perilaku pengunjung di situs web dapat dianalisis menggunakan algoritma klasifikasi terhadap data urutan kunjungan halaman situs atau biasa dikenal sebagai *web clickstream*. Pada penelitian ini digunakan algoritma *weighted k-nearest neighbor* untuk mengklasifikasikan *user* yang berpotensi melakukan pembelian barang online berdasarkan data publik *web clickstream* sebuah *e-commerce*. Metode *weighted k-nearest neighbor* melakukan klasifikasi berdasarkan label yang memiliki jumlah bobot jarak terbesar pada k-tetangga terdekat dari data yang diklasifikasi. Keluaran dari sistem ini adalah klasifikasi *user* yang berpotensi melakukan pembelian barang online dan mana yang tidak. Peneliti melakukan percobaan seleksi atribut dengan variasi k dan variasi jumlah atribut berdasarkan hasil perhitungan *information gain* dan memperoleh 2 atribut yang paling berpengaruh dalam klasifikasi yakni *Page Values* dan *Exit Rates*. Pengujian untuk mencari k yang paling optimal dilakukan pada 7.632 data menggunakan *3-fold cross validation* dan menghasilkan akurasi tertinggi yakni 86.2028% pada k = 65.

**Kata kunci :** *e-commerce*, *web clickstream*, *weighted k-nearest neighbor*, klasifikasi

## **ABSTRACT**

*Electronic commerce is an exchange using technology as an intermediary between various parties (individuals or organizations) as well as electronic-based intra or interorganizational exchange activities. The most commonly used interface is the web browser. Visitor's behavior on a web site can be analyzed using a classification algorithm on the data order of site visits or commonly known as web clickstream. In this study the weighted k-nearest neighbor algorithm is used to classify users who have the potential to make online based purchase on public data on a web clickstream data of an e-commerce. The weighted k-nearest neighbor method makes the classification based on the label that has the highest number of distance weights at the nearest k-neighbor from the classified data. The output of this system is the classification of users who have the potential to purchase goods online and those who do not. Researchers conducted an attribute selection experiment with variations in k and variations in the number of attributes based on the results of rank information gain and obtained the 2 most influential attributes in the classification namely Page Values and Exit Rates. Testing to find the most optimal k was done on 7,632 data using 3-fold cross validation and produced the highest accuracy of 86.2028% at k = 65.*

**Kata kunci** : *e-commerce, web clickstream, weighted k-nearest neighbor, classification*