ABSTRAK

Eric Stanley Kusuma (03082190025)

PERBANDINGAN ALGORITMA *NAÏVE BAYES DAN K-NEAREST NEIGHBOR* (KNN) DALAM KLASIFIKASI PEMILIHAN PROGRAM STUDI– STUDI KASUS UNIVERISTAS PELITA HARAPAN MEDAN

(xiii + 58 halaman: 18 gambar; 3 tabel; 3 lampiran)

Pemilihan program studi tepat merupakan langkah yang penting untuk membangun karir untuk mencapai kesuksesan dan masa depan yang cerah. Karena itu diperlukan *machine learning* yang dirancang untuk membantu pemilihan yang tepat. Penelitian ini membandingkan menggunakan metode *K-Nearest Neighbor* (KNN) dan *Naïve Bayes* dengan *dataset* yang berasal dari UPH Medan intake 2019 dan 2020 yang mengambil program studi sistem informasi atau informatika. Penelitian ini dikelola dan dilakukan di *Google Colab* dengan bahasa Pemrograman *Python*. Berdasarkan hasil data *training* dan data *testing* didapatkan *accuracy Naïve Besar* sebesar 56,25% pada *class* sistem informasi dengan *precision* 45%, *recall* 83%, *f1-score* 59% dan *class* informatika *precision* 80%, *recall* 40%, *f1-score* 53%. Sedangkan *accuracy K-Nearest Neighbor* (KNN) sebesar 81,25% pada *class* sistem informasi dengan *precision* 67%, *recall* 100%, dan *f1-score* 80% dan pada *class* informatika dengan *precision* 100%, *recall* 70%, dan *f1-score* 82%. Dari hasil performa algoritma *K-Nearest Neighbor* (KNN) lebih cocok digunakan untuk kasus ini

Kata kunci: *machine learning*, algoritma, *K-Nearest Neighbor* (KNN), *Naïve Bayes*, pemilihan program studi, program studi

Referensi:19(2019-2023),

ABSTRACT

Eric Stanley Kusuma (03082190025)

COMPARISON OF NAÏVE BAYES AND K-NEAREST NEIGHBOR (KNN) ALGORITHMS IN THE CLASSIFICATION OF CHOOSING STUDY PROGRAMS - A CASE STUDY OF UNIVERSITAS PELITA HARAPAN MEDAN

(xiii + 58 pages: 18 figures; 3 tables, 3 appendices)

Choosing the right study program is an important step towards building a successful career and a bright future. Therefore, a machine learning approach designed to assist in making the appropriate selection is needed. This research compares the use of the K-Nearest Neighbor (KNN) and Naïve Bayes methods with datasets derived from UPH Medan intake in 2019 and 2020, focusing on the study programs of information systems or informatics. The study was managed and conducted in Google Colab using the Python programming language. Based on the results of the training and testing data, Naïve Bayes achieved an accuracy of 56.25% for the information systems class, with a precision of 45%, recall of 83%, and an f1-score of 59%. For the informatics class, Naïve Bayes achieved a precision of 80%, recall of 40%, and an f1-score of 53%. On the other hand, K-Nearest Neighbor (KNN) achieved an accuracy of 81.25% for the information systems class, with a precision of 67%, recall of 100%, and an f1-score of 80%. For the informatics class, K-Nearest Neighbor (KNN) achieved a precision of 100%, recall of 70%, and an f1-score of 82%. Based on the performance results, the K-Nearest Neighbor (KNN) algorithm is more suitable for this case.

Keywords: machine learning, algorithms, K-Nearest Neighbor (KNN), Naïve Bayes, program selection, study programs.

*Reference:*19(2019-2023)