

ABSTRAK

Eko Putra Wijaya (03082180026)

KLASIFIKASI PESAN *SPAM* PADA *SHORT MESSAGE SERVICE* MENGUNAKAN ALGORITMA *EXTREME LEARNING MACHINE*

(xiv + 54 halaman; 19 gambar; 6 tabel; 1 lampiran)

Layanan pesan singkat (SMS) yang terdapat pada setiap telepon seluler memiliki potensi yang besar untuk menjadi jalur penyebaran pesan *spam*. Kurangnya fitur untuk menyaring pesan singkat berupa *spam* tersebut pada aplikasi pesan bawaan ponsel memberikan peluang bagi pelaku penyebar pesan spam untuk menyebarkan pesan spam ke para calon korban melalui SMS. Penelitian ini dilakukan dengan membuat model pembelajaran mesin menggunakan algoritma *Extreme Learning Machine* untuk mengklasifikasi pesan *spam*. Teks SMS sebelum diklasifikasikan menggunakan model ELM melalui tahap pra-proses dan tahap pembobotan menggunakan algoritma TF-IDF. Setelah 70% dari hasil pembobotan TF-IDF digunakan untuk melatih model pembelajaran mesin yang dikembangkan dan sisa dari data yang ada digunakan untuk evaluasi, model yang dikembangkan dapat melakukan prediksi data pengujian dengan tingkat akurasi tertinggi sebesar 97.37% dengan jumlah titik pada lapisan tersembunyi sebanyak 10,000 titik dan ambang batas klasifikasi di 0.5. Model ini dikatakan dapat diandalkan karena memiliki nilai *precision* sebesar 94.65%, *recall* sebesar 85.09% dan *F1 Score* sebesar 89.62%

Kata Kunci: *Klasifikasi, Machine Learning, Extreme Learning Machine*

Referensi: (15) (2006 – 2021)

ABSTRACT

Eko Putra Wijaya (03082180026)

SHORT MESSAGE SERVICE SPAM MESSAGE CLASSIFICATION USING EXTREME LEARNING MACHINE ALGORITHM

(xiv + 54 pages; 19 figures; 6 tables; 1 appendix)

Short message services (SMS) which provided on any mobile phone has big potential for spreading spam messages. The lack of feature to filter spam messages on mobile phone's native messaging app gives spammers opportunity to spread spam messages via SMS. This research is aimed to develop a machine learning model using Extreme Learning Machine algorithm to classify spam messages. Before the text can be fitted into the model to be classified, it need to be preprocessed and weighted using TF-IDF algoritm. Using 70% of TF-IDF weighted data to train the model and the remaining used to evaluate, the model is able to predict the test data with 97.37% of accuracy on 10,000 hidden nodes count and 0.5 is used for classification treshold. The model is classified as depentable because it has 94.65% of accuracy, 85.09% of recall, and 89.62% of of f1 score.

Keyword: Classification, Machine Learning, Extreme Learning Machine

Reference: (15) (2006 – 2021)