

ABSTRAK

Irianto, Dicky. 2022. *Memprediksi Iklan Lowongan Pekerjaan Palsu Menggunakan Algoritma Random Forest*, Skripsi. Tanjungpinang: Jurusan Teknik Informatika, Fakultas Teknik, Univeristas Maritim Raja Ali Haji. Pembimbing I: Nurfalinda, S.T., M.Cs. Pembimbing II: Nola Ritha, S.T.,M.Cs.

Tujuan dari penelitian ini adalah untuk memperoleh hasil prediksi lowongan pekerjaan asli atau palsu. Alasan mengambil judul tersebut karena iklan lowongan pekerjaan palsu sulit dibedakan dengan yang asli. Hal itu disebabkan isi konteks iklan lowongan pekerjaan palsu hampir sama dengan yang asli. Oleh sebab itu, peneliti menggunakan *dataset* yang berisi data iklan lowongan pekerjaan asli dan palsu. Algoritma yang digunakan pada penelitian ini adalah algoritma *Random Forest*. Atribut yang digunakan adalah atribut *title* dan atribut *location* sebagai atribut untuk pencarian data, atribut *has_company_logo*, *has_questions*, *required_experience*, dan *required_education*, dan atribut *fraudulent* sebagai atribut untuk proses prediksi dalam model *Random Forest*. Pengujian menggunakan data sebanyak 17880 dan *tree* sebanyak 2,4, 6, 8, 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, dan 1000. Tingkat akurasi dan *f1-score* terbaik yang didapatkan pada penelitian ini adalah 96,86% dan *F1-score* 0,658, dengan jumlah *tree* yang digunakan sebanyak 2. Peningkatan atau banyaknya jumlah *tree* tidak menghasilkan peningkatan akurasi dan *F1-score*.

Kata kunci: *lowongan pekerjaan, asli, palsu, Random Forest*

ABSTRACT

Irianto, Dicky. 2022. Predicting Fake Job Vacancies Using the Random Forest Algorithm, Thesis. Tanjungpinang: Informatics Engineering Department, Faculty of Engineering, Raja Ali Haji Maritime University. Advisor I: Nurfalinda, S.T., M.Cs. Advisor II: Nola Ritha, S.T.,M.Cs.

The purpose of this study is to obtain predictive results for real or fake job openings. The reason for the title is that fake job ads are indistinguishable from the real thing. Fake job ads are almost identical to real job ads. Researchers therefore use datasets containing real and fake job posting data. The algorithm used in this study is the random forest algorithm. The attributes used are title and location attributes, `has_company_logo` attribute, `has_questions` attribute, `required_experience` attribute, `required_education` attribute as attributes for data retrieval, and incorrect attributes as prediction processing attributes for random forest model. This test used 17,880 data points and 2.4, 6, 8, 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000. The highest precision and f1 values were achieved in the study. This is 96.86. n F1 score 0.658, maximum number of trees used is 2. Increasing and increasing the number of trees does not improve accuracy and F1 score.

Keywords: *job vacancies, real, fake, Random Forest*