

## Detecting scam job advertisement using machine learning techniques

Natha M.S.S.<sup>1\*</sup>, Amath A.A.S.<sup>1</sup>, and Erandi J.D.T.<sup>2</sup>

<sup>1</sup>Department of Computing and Information Systems,  
Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka

<sup>2</sup>Department of Data Science,  
Faculty of Computing, Sabaragamuwa University of Sri Lanka, Sri Lanka

\*mssnatha@std.appsc.sab.ac.lk

Automated detection of scam job advertisements is critical for ensuring safer online recruitment. This study proposes a machine learning–based approach for detecting fraudulent job postings using textual and metadata features. Experiments were conducted using the publicly available Fake Job Posting Prediction Dataset from Kaggle, which contains approximately 17,880 job postings with 18 features, exhibiting a significant class imbalance between legitimate and fraudulent advertisements. Text preprocessing and feature extraction techniques were applied to identify discriminative linguistic patterns associated with scam postings. Multiple machine learning classifiers, including Logistic Regression, Support Vector Machines, Random Forest, and Gradient Boosting, were trained and evaluated. Model performance was assessed using accuracy, precision, recall, F1-score, and ROC–AUC metrics. Experimental results demonstrate that ensemble-based models outperform baseline classifiers, achieving superior detection performance. In addition, feature importance analysis highlights key linguistic indicators such as exaggerated benefits, vague job descriptions, and abnormal salary patterns, contributing to improved interpretability of scam detection. The findings indicate that the proposed approach can serve as an effective automated scam job advertisement detection system to support safer recruitment practices. Future work will focus on validating the proposed framework using real-world job advertisement datasets and deploying it as a practical decision-support tool.

**Keywords:** *Scam Job Detection; Machine Learning; Fake Job Advertisements; Text Classification; Recruitment Security*