# PROJECT REPORT

## COMPARISON BETWEEN CNN AND RANDOM FOREST PERFORMANCE IN DETECTING HOAX INDONESIAN NEWS ARTICLES

**FRANCISKA NUGRAHAENI SIWI PRAMUDYA**
**20.K1.0032**

**Faculty of Computer Science**
**Soegijapranata Catholic University**
**2024**

# ABSTRACT

*Hoax news is a serious problem in this era. Many people are easily led by opinions made by certain people without seeing the truth or looking for existing facts. To overcome this, many researchers have conducted hoax news detection using various algorithms. In some studies, it is said that Random Forest has better performance to overcome this hoax news problem. In other studies, it is also said that CNN has the same level of performance as the Random Forest algorithm. In addition, the problem that is often found is the error in prediction due to improper preprocessing methods. Therefore, in this research, the appropriate preprocessing method is searched by using several preprocessing scenarios for the Convolutional Neural Network (CNN) and Random Forest algorithms. Therefore, in addition to finding the right preprocessing method for each algorithm, a performance comparison is also carried out on the CNN and Random Forest algorithms using a dataset of 4000 news facts from Kompas.com and 4000 hoax news from the turnback.hoax site. the results obtained in this study are random forest has an average model accuracy value of 90% and the cnn algorithm has an average model accuracy value of 60% using the same extraction method, namely TFIDF combined with Ngrams worth one or unigram.*

*Keywords : Hoax news detection, Convolutional Neural Network(CNN), Random Forest.*