**Title of the Project:** Real & Fake Classification using NLTK Technique

**Name of the students:** Kavushick P, Manigandan L, Nareshkumar C.

**Register Number(s):** 211419104133, 211419104159, 211419104177.

**Name of the Guide:** Dr. S. Hariharan, M.E., M.BA., Ph.D.

**ABSTRACT**

To avoid fraudulent Job postings on the internet, we target to minimize the number of such frauds through the Machine Learning approach to predict the chances of a job being fake so that the candidate can stay alert and make informed decisions if required. The model will use NLP to analyze the sentiments and pattern in the job posting and TF-IDF vectorizer for feature extraction. In this model, we are going to use Synthetic Minority Oversampling Technique (SMOTE) to balance the data and for classification, we used Random Forest to predict output with high accuracy, even for the large dataset it runs efficiently, and it enhances the accuracy of the model and prevents the overfitting issue. The final model will take in any relevant job posting data and produce a result determining whether the job is real or fake. The rise of online job postings has led to an increase in fraudulent job postings, which can deceive job seekers and cause significant financial and emotional harm. In this paper, we propose a machine learning approach to classify job postings as real or fake using natural language processing (NLP) techniques. We then applied the K-Nearest Neighbors (KNN) algorithm to classify the job postings as real or fake. Our results show that our model achieved an accuracy of 92% in classifying job postings as real or fake, outperforming the baseline model. Furthermore, we conducted error analysis and feature importance analysis to identify the strengths and limitations of our approach. Our study contributes to the growing body of research on detecting fraudulent job postings using NLP techniques and provides valuable insights for job seekers, job portals, and law enforcement agencies