**MOBILE MESSAGE CLASSIFICATION USING NATURAL LANGUAGE PROCESSING AND MACHINE LEARNING ALGORITHMS**

**ABSTRACT**

SPAM: Stupid Pointless Annoying Malware is any kind of unwanted, unsolicited digital communication that gets sent out in bulk. Often spam is sent via email, but it can also be distributed via text messages, phone calls, or social media. Unfortunately, everyone with a mobile must be bothered by spam messages, whether we like it or not. Here this project classifies spam messages. Understanding different spam text classification techniques like extraction, text preprocessing, and NLTK stop words is vital. This project mainly focuses on the spam classification approach using machine learning algorithms such as Random Forest, KNN, Naïve Bayes, Support Vector Machine, decision tree, and NLP algorithms Count Vectorization and TF-IDF

**ALGORITHM:**

Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes (NB), Decision Tree (DT), TF-IDF, Count Vectorization (cv)

**MODULES:**

DATA CLEANING: missing values, null values, class imbalance

DATA PREPROCESSING: Removing punctuations and stop words, lemmatization & stemming

FEATURE EXTRACTION: Bag of words and TF-IDF

MODEL BUILDING: Building the models through algorithms

MODEL TESTING AND COMPARING ALGORITHMS: comparing the accuracies of different algorithms

WEB APP: Deploying model through flask as a web app

**SOFTWARE REQUIREMENTS**

1. Python 2.7 and above.
2. Jupyter notebook and VS code
3. NLTK, NumPy, Pandas, Matplotlib, seaborn
4. Sci-Kit learn

**BATCH NO 10**

GUDIPATI VIKAS      19261A1219

M.V.S KOUSHIK          19261A1230

MIRYALA NITHYA    19261A1237

**Signature Of Students                                                 Signature of Guide**

(MRS.CH. Sudha)