The performance comparison of Machine learning classifiers involved different stages in the required system such as Text collection, Text preprocessing, Feature extraction and the News to be classified. Further,the feature extraction included several constraints like Bag of words, Countvectorizer, TF-IDF and Word embedding using Spacy. The Classifiers further included the constraints like Support vector machine, Logistic regression, Decision trees, Random forest, Gradient boosting and XG-Boost.In Identification of Fake news using Machine learning, we have used ML to develop our devices. Naive bayes model assumes that the features are independent of one another statistically. The Naive Bayes model is famous for its multi-class prediction, it was selected for its case and dynamic nature of predicting the nature of the text. We opted to use tokenizers for feeding into deep neural network models. For datasets with just one attribute, we used sequential DNN and complex networks were developed with the help of functional DNN.

Support vector machine (SVM) is a good device to derive the binary class based on the information given to the model. Its function is to characterise the article into two categories (true or false). SVM is an advanced machine learning algorithm that can be used for different purposes like regression and clarification. It is based on the principle of searching the hyperplanes that best classify the dataset into two classes. SVM has the potential to deal with high dimensional spaces and will be memory proficient.

The performance evaluation of trained models are computed based on three feature sets i.e., a)Normalized frequency of parsed syntactic dependencies, b)Bi-gram term frequency - Inverse document frequency and union of a) and b). Spacy is executed in Cython , which is another set of Python language that allows C code to be used in Python by using Python/C APL. Spacy gives better yield rate for performance especially on tasks like entity recognition, parsing and so on. A neural network based model is proposed which is trained under a very huge number of training samples. We call this a Fake News Article Detector (F-NAD) model, which acts as a binary classifier between fake and real news segments. A total of approximately 45k training samples is generated based on approximately 7.5k news articles downloaded from the Internet. Once the model is trained correctly, the test news article is fed to the device to determine whether it is fake or real. Content information requires processing to implement AI on them. Stemming technique is used to remove suffixes or prefixes from a word. There is a huge amount of data stored in file but it can't be accessed by computer assisted analysis. NLP allows the analysts to find some crucial data. The Data should be in the form of preprocessing. It operates with expulsion of punctuations, URL's images, stemming and stop words. We then classify that information utilizing classifiers (for eg. LR, SVM, NB, RF and DNN).