Chapter 46
Deep Learning Based Sentiment Analysis for Phishing SMS Detection

Aakanksha Sharaff
National Institute of Technology, Raipur, India

Ramya Allenki
UnitedHealth Group, India

Rakhi Seth
National Institute of Technology, Raipur, India

ABSTRACT

Sentiment analysis works on the principle of categorizing and identifying the text-based content and the process of classifying documents into one of the predefined classes commonly known as text classification. Hackers deploy a strategy by sending malicious content as an advertisement link and attack the user system to gain information. For protecting the system from this type of phishing attack, one needs to classify the spam data. This chapter is based on a discussion and comparison of various classification models that are used for phishing SMS detection through sentiment analysis. In this chapter, SMS data is collected from Kaggle, which is classified as ham or spam; while implementing the deep learning techniques like Convolutional Neural Network (CNN), CNN with 7 layers, and CNN with 11 layers, different results are generated. For evaluating these results, different machine learning techniques are used as a baseline algorithm like Naive Bayes, Decision Trees, Support Vector Machine (SVM), and Artificial Neural Network (ANN). After evaluation, CNN showed the highest accuracy of 99.47% as a classification model.

DOI: 10.4018/978-1-6684-6303-1.ch046

INTRODUCTION

Text Classification

Text classification is one of the most important parts of text analysis. It is defined as the process of interpreting and extracting important information from the present textual data this data can be of any type like SMS, Twitter data, emoji, and short messages while talking about classification which is one of the major parts of sentiment analysis; which occurs to be the measuring people's attitude from the piece of text through which they are sharing their views. Views can be of different types based on user intent this can be understood through various examples, we saw over the internet sometimes inappropriate like abusive language and pornographic content; sentiment analysis also deals with classifying those data which helps the policymaker to understand the trend that is running in a market that solely depends on users' reviews, feedbacks, and ratings. From a research point of view, some of the major challenges that could be solved through sentiment analysis like spam filtering, phishing attack, categorization, and summarization well over the decades, spamming and phishing based classification has been some of the most researched topics based on techniques like machine learning and deep learning. A good text classifier is a classifier that efficiently categorizes large sets of text documents in a reasonable amount of time with acceptable accuracy. Many techniques and algorithms for automatic text categorization have been devised.

Applications

There are various applications of text classification:

1. **Document Organization:** Document organization is also known as Document classification. (Rinaldi et al., 2021) discuss the documents that were collected through the different platform in huge amount, but from the information retrieval point of view, not all the data was always relevant, so sometimes the issue of information overloading may generate, for solving this issue the new concept of document classification is introduced while using the text present in the document, we illustrate it through an example given below; here document belongs to a different class (class 1, class 2 and class 3) in the training set as given in Table 1 and associated feature is retrieved from each document class and vector is created from these words of class in Table 2. As we can see "Some" belongs to one class, "Yellow" belong to another based on the specified documents, while testing the new data this labeling helps to understand which word vector belongs to which class.

Table 1. Documents

Document Class 1	Document Class 2	Document Class 3
Some Lion live in the jungle	Yellow is a color	Go to Manhattan city