SMS Text Classification

Ayleah K. Hill

Professor Ananth Jillepalli

CPTS 315 Data Mining

December 11, 2020

Introduction

Nowadays we are bombarded with plenty of spam messages either in emails, phone calls or texts. With the development of Covid-19 this year, there has been a significant increase in volume of spam messages and phone calls. Between January and October there was an increase in spam calls alone by 159 percent since last year.[1] With people being stuck at home due to the pandemic, some people may have more free time to pay attention to messages that they receive, including spam messages. As spam messages continue to plague us with no end in sight, it would help to be able to build machine learning algorithms and utilize data mining skills to be able to identify whether texts are spam messages based on their content. Spam messages not only have the ability to waste our time but also have the ability to obtain sensitive information from us that can put us in dangerous situations.

The challenge I had with my approach was finding the best way to analyze the data. After having found numerous ways that text classifications could be handled, it took some research to find a model that would be both complex and efficient with a decent amount of accuracy.

My results showed that with the use of a Long-Short Term Memory model, the SMS text message data could be classified as either "spam" or "ham" with around a 98% accuracy.

Data Mining Task

The purposes of data mining is to be able to process data, finding useful information, detecting trends and being able to make predictions[2].The task of this project was to use machine learning algorithms to be able to classify the data with as high of accuracy as possible, followed by the development of visualizations on the findings. I asked a variety of data mining questions: what methods have been used for text classification? How accurate is the use of deep learning models on text classification? Some of the challenges that came about solving this task was determining what ways could be used for text classifications that could provide some of the most accurate results.

Technical Approach

In class, one of the approaches that was taken was using the perceptron algorithm. The perceptron algorithm is designed to find a hyperplane to divide linearly separable data. My initial plan was to replicate this same idea on the SMS text messages data, but I decided to take it a step further. Looking into other methods that have been used for text classifications, I found what is called a Long-Short Term Memory Model (LSTM). An LSTM is a form of Recurrent Neural

[1] Ballard, B. (2020, December 10). Spam calls and emails saw a huge rise this year. TechRadar. https://www.techradar.com/in/news/spam-calls-and-emails-saw-a-huge-rise-this-year

[2] Bushkovskyi, O. (2018, August 14). Data Mining: The Definitive Guide To Techniques, Examples, And Challenges. https://www.Theappsolutions.com/Blog/Development/Data-Mining-Guide.

Network (RNN) that is found in machine learning methods[3]. RNN's also fall within the category of deep learning, which is within machine learning and focuses on neural networks[4]. The purpose of neural networks in general is to understand and comprehend relationships in data sets, mimicking the way the human brain solves problems[5]. More specifically, RNN's are designed to be able to comprehend data when it is provided in a particular sequence and is used in Natural Language Processing (NLP) such as text and speech recognition[6]. RNNs do this by keeping track, or "remembering" old information, taking into consideration in future computations[78].

Evaluation Methodology

The dataset I used was provided by the UCI Machine Learning Group which posted the dataset on Kaggle. That dataset contains 5,574 text messages total, consisting of a mix between legitimate texts and spam texts (identified as either ham or spam accordingly) where 87% of the messages are "ham" and the other 17% are "spam". The two columns in the data set are for labelling the texts as "ham" or "spam" and for holding the content of the texts.
One of the challenges with this dataset was determining how to preprocess the data, ensuring as much of the information in the dataset was well formatted and consistent as possible to ensure the clarity and accuracy of the results. Because this dataset consisted of text messages, the format of the text messages themselves are far more casual and colloquial than formal written texts such as articles or books. I contemplated eliminating the "noise" in messages such as acronyms and misspelled words, but it did not seem to influence the results of the data that much when they were kept in. The output of the data was displayed in an accuracy score and a loss score (on a scale from 0-1), showing how accurate the LSTM is able to classify the data as either "spam" or "ham" based on the SMS texts, and the rate at which it guesses incorrectly.

R**esults and Discussion**

The results of my program show an accuracy rate of 0.983 (or a 98.3% accuracy) and a loss rate of 0.051 (or a 0.051% loss). Judging by the accuracy rates, an LSTM is able to effectively identify the spam messages among all the SMS texts. Furthermore, this shows that there is a

[3] Grandic, I. (2019, May 26). How LSTM's work. Medium. https://www.medium.com/@izzygrandic/how-lstms-work-263ac4e412ba
[4] Brownlee, J. (2016, September 22). What is Deep Learning? Machine Learning Mastery. https://machinelearningmastery.com/what-is-deep-learning/
[5] Simpson, M. (2019, November 19). Machine Learning Algorithms: What is a Neural Network? Www.Verypossible.com. https://www.verypossible.com/insights/machine-learning-algorithms-what-is-a-neural-network
[6] Simpson, M. (2019, November 19). Machine Learning Algorithms: What is a Neural Network? Www.Verypossible.com. https://www.verypossible.com/insights/machine-learning-algorithms-what-is-a-neural-network
[7] Simpson, M. (2019, November 19). Machine Learning Algorithms: What is a Neural Network? Www.Verypossible.com. https://www.verypossible.com/insights/machine-learning-algorithms-what-is-a-neural-network
[8] Kowsari, K. (2019, May 21). Text Classification Algorithms: A Survey. Medium. https://www.medium.com/text-classification-algorithms/text-classification-algorithms-a-survey-a215b7ab7e2d

clear distinction between spam messages and normal text messages. This opens up the possibility of building a program designed to classify SMS texts to flag incoming spam messages. Flagging spam text messages can further protect texter's privacy as well as ensure their security and personal information.

The program also created two different word clouds:

Figure 1 Spam Word Cloud

Figure 2 Ham Word Cloud

The first word cloud shows the most frequently used words found in spam messages. The second word cloud shows the most frequently used words found in messages that were not labelled as spam. These word clouds show what words the LSTM learned to find in the messages to then flag the messages as “spam” or “ham.”

What is interesting when analyzing the word clouds is that while there is some overlap in a few of the most common words found in both the “spam” and “ham” messages, with the use of the LSTM and its ability to do Natural Language Processing, it is still able to differentiate between what is considered a spam message and a standard message based on the contents and their sequence in the message.

Lessons Learned

After having done a form of text classification with the perceptron algorithm in class, it is interesting to see the same work being done using a different method. I learned that text classification has been handled in a variety of ways with multiple algorithms, all with a variety of advantages and disadvantages, and some being more complex than others.

What I could have done differently with the project is ran a variety of machine learning algorithms for text classifications to see whether certain algorithms are more efficient and/or effective at classifying the SMS text messages, such as having ran the perceptron algorithm (making two versions: one with stemming, and one without stemming) on this data set and comparing the results to the LSTM model, as well as K-nearest neighbors, Naïve Bayes, and other forms of deep learning algorithms like the LSTM model, including using an attention model.

Acknowledgements

Dataset

https://www.kaggle.com/uciml/sms-spam-collection-dataset

Code Sources

https://www.datacamp.com/community/tutorials/wordcloud-python

https://towardsdatascience.com/multi-class-text-classification-with-lstm-1590bee1bd17

https://www.kaggle.com/kredy10/simple-lstm-for-text-classification

Bibliography

Ballard, B. (2020, December 10). Spam calls and emails saw a huge rise this year. TechRadar. https://www.techradar.com/in/news/spam-calls-and-emails-saw-a-huge-rise-this-year

Brownlee, J. (2016, September 22). *What is Deep Learning?* Machine Learning Mastery. https://machinelearningmastery.com/what-is-deep-learning/

Bushkovskyi, O. (2018, August 14). *Data Mining: The Definitive Guide To Techniques, Examples, And Challenges*. https://www.Theappsolutions.com/Blog/Development/Data-Mining-Guide.

Grandic, I. (2019, May 26). *How LSTM's work.* Medium. https://www.medium.com/@izzygrandic/how-lstms-work-263ac4e412ba

Kowsari, K. (2019, May 21). *Text Classification Algorithms: A Survey*. Medium. https://www.medium.com/text-classification-algorithms/text-classification-algorithms-a-survey-a215b7ab7e2d

Simpson, M. (2019, November 19). *Machine Learning Algorithms: What is a Neural Network?* Www.Verypossible.com. https://www.verypossible.com/insights/machine-learning-algorithms-what-is-a-neural-network