Adoption of a Revised Text Classification Method Using the Laravel Framework

RAJEEV TRIPATHI Computer Science School of Management Sciences Lucknow Uttar Pradesh INDIA

Abstract: - The prevalence of the Internet of Things has increased the frequency of malicious attacks on the many vulnerable devices, such as cellphones. Conventional framework design techniques for web design have a lot of limits, are time-consuming, and have other problems. The industry norm for message exchange and fusion between dissimilar systems has long been web services. We provide a machine learning-based strategy for extracting SPAM messages based on the MVC concept and Laravel. In order to solve these problems, this article offers a web design and implementation technique based on the Laravel framework. By automatically handling connections that lack business logic, Laravel standardises the development process.

Key-Words: - Data Analytics, AI, Text Mining through Laravel, MVC model, Machine Learning

Received: March 9, 2024. Revised: October 12, 2024. Accepted: November 9, 2024. Published: December 11, 2024.

1 Introduction

In this age of fast software development, online applications are gaining inexorable popularity, with many of them serving millions of users each day and reaching towards its peak. Monolithic application architecture is not suited for today's large-scale Internet applications, particularly in dispersed and heterogeneous computer environments, nor is it compatible with multi-team collaborative development.

Text mining is an indexing technique that intends to extract structured text data from unstructured text data. The text mining process entails collecting, preparing, and analyzing documents from a variety of sources. When acquiring structured data from unstructured databases, certain procedures are followed to guarantee that users are satisfied [1]. After that, text mining techniques including information retrieval, classification, clustering, and categorization are employed to guarantee that the data is appropriately processed and created [2]. The processing of structured or semi-structured data has become highly complicated throughout all enterprises as data volumes have raised dramatically. TF, IDF, TF-IDF, and other approaches or algorithms can be used to process data. TF-IDF is a numerical statistic that displays the relevance of keywords to certain documents, or it can be argued that it offers those keywords that may be used to identify or categories specific documents [3]. Term Frequency (TF) is a metric that measures how infrequently a term word appears in the document. Because total document length might range from very short to extremely long, any phrase may appear more frequently in large papers than in small documents [4]. To solve this problem, the word frequency is computed by dividing the occurrence of any term in a document by the total number of terms in the document. IDF (Inverse Document Frequency) provides a lower weight to often occurring words and a higher weight to infrequently occurring terms. The term frequency (TF) and the inverse document frequency (IDF) are easily multiplied [5].

The TF-IDF method has certain drawbacks that must be addressed. The main limitation of TF-IDF is that it cannot detect words with minor tense changes. For instance, the algorithm would regard "go" and "goes" as two different autonomous words, just as it will take "play" and "playing," "mark" and "marking," and "year" and "years" as separate terms. Because of this restriction, when the TF-IDF method is used, it might occasionally provide surprising results. Another drawback of the TF-IDF is that it can't verify the semantics of text in documents, thus it's only helpful up to the lexical level. It also lacks the ability to look for word cooccurrences. Decision Trees, Pattern or rule-based classifiers, SVM classifiers, Neural Network classifiers, and Bayesian classifiers are some of the techniques that may be used to enhance performance and accuracy.

2 Problem Formulation

2.1.1 API Technology

REST is a technological standard for Web application that enables for heterogeneous interoperability. The use of REST can result in an architecture that is simple, scalable, effective, safe, and dependable. REST is a lightweight RPC protocol based on HTTP. Its ease of use and web friendliness make it an unrivalled alternative to SOAP, which has been a popular RPC solution for successfully Many developers have vears. implemented simple and powerful APIs based on Ajax and Restful web services utilising Rest.

2.1.2 Restful API

Representational state transfer (REST) is a method of allowing computer systems on the Internet to communicate with one another. Using a standard and established set of stateless actions, RESTcompliant Web services enable requesting systems to access and alter textual representations of Web resources. Roy Fielding developed and defined the term representational state transfer in his PhD dissertation in 2000. HTTP 1.1 and Uniform Resource Identifiers were designed using REST by Fielding (URI). An application object, a database record, an algorithm, and other types of information that can be accessed are examples of resources. REST represents URI in the form of "/user/name," and actions on HTTP methods GET, PUT, POST, DELETE, HEADER, and OPTIONS result in the next resource being delivered back to the caller. The server side of REST maintains statelessness between multiple encounters, meaning that any server in the cluster can serve the client on any request [16][18].

2.2.1 Model-View-Controller (MVC)

Model-View-Controller (MVC) is a design pattern that segregates an application into three logical components: model, view, and controller. Each of these components is designed to deal with a distinct element of application development. MVC is a popular web development framework for building scalable and adaptable projects [6]. The MVC framework was first used in Small Talk apps to generate user interfaces (Programming Language). The system is separated into three components in this approach: the Model, which shows the Logic area, the View, which displays the user interface, and the Control, which handles changes to the View [7][20]. Fig. 1. Shows the components of MVC model.



Fig. 1. MVC Components

2.2.1 Model: All data-related functionality that the user deals with is represented by the Model component. This might be the data being exchanged between the View and Controller components or any other data pertaining to business logic. A Customer object, for example, will get customer data from a database, alter it, and either update it back to the database or utilise it to render data.

2.2.2View: All of the application's UI functionality is handled by the View component. The Customer

view, for example, will have all of the UI components that the final user interacts with, such as text fields, dropdowns, and so on.

2.2.3 Controller: Controllers serve as a link among the Model and View components, processing all business logic and incoming requests, manipulating data using the Model, and interacting with Views to display the final output. The Customer controller, for example, will handle all interactions and inputs from the Customer View and use the Customer Model to update the database. The Customer data will be viewed using the same controller.

A Model is used to interface with a database, which need not be one. It may be a JSON file or another type of resource. A Controller holds the logic for dealing with a Model, such as verifying form data and saving a resource to the database. The view is the application's User Interface (UI) that includes HTML or presentation markup. Loops and conditionals are examples of logic. Logic is included in Views via template engines. Blade template engine is used by Laravel to provide logic to views [8].

3 Laravel Framework

Laravel is a free, open-source PHP web framework designed by Taylor Otwell for building online applications that follow the model-view-controller (MVC) architectural paradigm. А modular packaging system with a dedicated dependency management is only one of Laravel's features[10]. The laravel framework is simple to grasp and extremely powerful. It includes authentication, routing. session management, caching. IoC container, and a plethora of other widely used components, as well as fantastic database conversion tools and integrated unit testing support.

3.1.1 Text Classification with TNTSearch and Laravel

Online frameworks assist developers in this process by providing a standardised method for developing and deploying web applications. These frameworks are frequently used to decouple the logic and the view, i.e., to implement the Model-View-Controller (MVC) designs [9]. Furthermore, frameworks allow developers to reuse design and implementation by integrating the methods for the many activities that must be completed. These frameworks benefit the project by reducing development time, reducing complexity, increasing productivity, extending extensibility, and ensuring dependability.

Laravel was designed by Taylor Otwell and is now one of the most widely used open-source PHP web frameworks for developing online applications. It is built on Symfony and follows the model-viewcontroller architecture paradigm. Laravel is a PHP MVC framework that was first published in 2011. It is one of the most well-known MVC-based PHP frameworks, and researchers have determined that Laravel surpasses other MVC frameworks, making it the best PHP framework for future web technologies [10][18]. Django is a free and opensource web framework that is based on the Python programming language. It was first launched in 2005 and immediately gained popularity owing to the reusability of components; rapid development enables sophisticated web applications to be developed in less time [11].

It's difficult to deny that machine learning is becoming increasingly common. Everyone is talking about it, and it appears that everyone is using it. True, machine learning activities require a lot of computational power, but language speed isn't everything. Spam Detection, Text Categorization, Sentiment Analysis are the several applications of Text classification [12][22]. Multinomial Naive Bayes is a cutting-edge text categorization method. This is a probabilistic learning approach that determines the likelihood of a document falling into a specific category.

4 Problem Solution

Let's say people running a massages on their mobile or emails and they want to classify their massages. They want to see which massages are SPAMs and which are not. First, I need to install TNTSearch, which is a search engine entirely written in PHP but also has some cool stuff like classification [13], which is part of information retrieval, and this is of course also machine learning. So, installing composer require *teamtnt/tntsearch*

Using above piece of code basically I take the .csv file containing the massages together with their category which can be SPAM and teaches he classifier. Finally the trained model is saved as ./path/to/usefulmassages.cls

In this way we train our model. We only need to do the training one. For more accurate predictions we should take large dataset as possible. Now I have a trained model, predicting if a massage is SPAM or not is easy. In my *predict.php*, I have the following:

<?php

require _____DIR___ . '/vendor/wholedata.php'; use

TeamTNT\TNTSearch\Classifier\TNTClassifier;

\$classifier = new TNTClassifier(); \$classifier->load('./path/to/usefilmassages.cls');

\$predict = \$classfier->predict('This is a not spam
massage');

echo \$predict['label'];

Using above code I load the model from source and ask classifier to tell what the prediction might be. After performing certain tests with large datasets and the performance and accuracy are amazing. The SPAM massages classification test has a score of 98.34753%

5 Conclusion

The notion of using the Laravel framework to construct the fundamental infrastructures of the Restful service is discussed in this article, as well as essential implementation details for our approach. Restful Web Services has become the preferred technology model for micro services applications due to its lightweight, scalability, and HTTP protocol compatibility. This paper shows how to design and construct a web application using the Laravel framework. The development process and automates the processing of non-business logic relationships. It implements a basic Laravel model that allows for automated processing of a portion of the design through php. According to the results of the experiments and simulations, web design based on the Laravel framework has scalability and strong scalability, which improves development productivity.

References:

[1]. Saravanan, D. (2018). Image frame mining using indexing technique. In *Data Engineering and Intelligent Computing* (pp. 127-137). Springer, Singapore.

[2]. Tripathi, R., & Dwivedi, S. K. Identification of QR Code Perspective on enhancement of Text Mining Approaches. ijres.org

[3]. Tripathi, R., & Dwivedi, S.K. "A Quick Review of Data Stream Mining Algorithms", IJIR, Vol-2, Issue-7, 2016.

[4]. Azam, N., & Yao, J. (2012). Comparison of term frequency and document frequency based feature selection metrics in text categorization. *Expert Systems with Applications*, *39*(5), 4760-4768.

[5]. Tripathi R (2021). Interpretive Psychotherapy of Text mining Approaches, Springer Lecture Notes in Networks and System.

[6]. Aniche, M., Bavota, G., Treude, C., Gerosa, M. A., & van Deursen, A. (2018). Code smells for model-view-controller architectures. *Empirical Software Engineering*, 23(4), 2121-2157.

[7]. Jailia, M., Kumar, A., Agarwal, M., & Sinha, I. (2016, November). Behavior of MVC (Model View Controller) based Web Application developed in PHP and. NET framework. In 2016 International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-5). IEEE.

[8]. Sunardi, A. (2019). MVC architecture: A comparative study between laravel framework and slim framework in freelancer project monitoring system web based. *Procedia Computer Science*, *157*, 134-141.

[9]. Deacon, J. (2009). Model-view-controller (mvc) architecture. *Online]*[*Citado em: 10 de março de 2006.] http://www. jdl. co. uk/briefings/MVC. pdf.*

[10]. Olanrewaju, R. F., Islam, T., and Ali, N. (2015). *An Empirical Study of the Evolution of PHP MVC Framework*, pages 399–410. Springer International Publishing, Cham.

[11]. Gupta, P., & Govil, M. C. (2010). Spring Web MVC Framework for rapid open source J2EE application development: a case study. *International Journal of Engineering Science and Technology*, 2(6), 1684-1689.

[12]. Wang, H., Ma, C., & Zhou, L. (2009, December). A brief review of machine learning and its application. In 2009 international conference on information engineering and computer science (pp. 1-4). IEEE.

[13]. Zholudev, V., Kohlhase, M., & Rabe, F. (2010). A [insert XML Format] Database for [insert cool application]. *Proceedings of XML Prague*, 2010(207), 317-339.

[14]. Samet, S., Ishraque, M. T., Ghadamyari, M., Kakadiya, K., Mistry, Y., & Nakkabi, Y. (2019). TouchMetric: a machine learning based continuous authentication feature testing mobile application. *International Journal of Information Technology*, *11*(4), 625-631.

[15]. Stouky, A., Jaoujane, B., Daoudi, R., & Chaoui, H. (2017, November). Improving software automation testing using jenkins, and machine learning under big data. In *International Conference on Big Data Technologies and Applications* (pp. 87-96). Springer, Cham.

[16]. Fielding, Roy Thomas (2000). "Chapter 5: Representational State Transfer (REST)".

[17]. Hatcher, W. G., Maloney, D., & Yu, W. (2016, June). Machine learning-based mobile threat monitoring and detection. In 2016 IEEE 14th International Conference on Software Engineering Research, Management and Applications (SERA) (pp. 67-73). IEEE.

[18]. Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.

[19]. "Fielding discussing the definition of the REST term". groups.yahoo.com. Retrieved 2017-08-08.

[20]. Yu, H. R. (2015). Design and implementation of web based on Laravel framework. *ACSR-Advances in Comptuer Science Research valume*, 6, 302.

[21]. Isele, R., Umbrich, J., Bizer, C., & Harth, A. (2010, November). LDspider: An open-source

crawling framework for the Web of Linked Data. In *Proceedings of the 2010 International Conference on Posters & Demonstrations Track* (Vol. 658, pp. 29-32).

[22]. Tripathi R (2021). Interpretive Psychotherapy of Text mining Approaches, Springer Lecture Notes in Networks and System.

[23]. Tripathi, Rajeev (2023). Inventiveness of Text Extraction with Inspiration of Cloud Computing and ML using Python Logic, 22nd Intelligent Systems Design and Applications (ISDA'22), 2022/12

[24]. Tripathi, Rajeev (2023). Digital Governance augmentation with the assistance of Li-Fi Technologies, International Conference on Digital Governance through IoT Solutions CSI & DigiGov-23, ICDG-103, ISBN: 978-93-5780-241-3.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en US