An Early Warning Detection System of Terrorism in Indonesia from Twitter Contents using Naïve Bayes Algorithm

Mediana Aryuni
Information Systems Department,
School of Information Systems
Bina Nusantara University
Jakarta, Indonesia 11480
mediana.aryuni@binus.ac.id

Eka Miranda
Information Systems Department,
School of Information Systems
Bina Nusantara University
Jakarta, Indonesia 11480
ekamiranda@binus.ac.id

Yudi Fernando
Faculty of Industrial Management,
University Malaysia Pahang
Pahang, Malaysia, 26300
yudi@ump.edu.my
Management Department,
BINUS Online Learning,
Bina Nusantara University
Jakarta, Indonesia, 11530
yudi.fernando@binus.ac.id

Tia Mariatul Kibtiah
International Relations Department,
Bina Nusantara University
Jakarta, Indonesia 11480
tia.kibtia001@binus.ac.id

Abstract— **Aware on the benefits of social media as the networking platform, the extremist organization is utilized social media to spread the ideology, recruit new member and guided a suicide bomber alike. There are opportunities to analyze the content of document texts in social media including the terrorism detection and intention by extracting the content evident in their post, comment etc. The objective of this research is to analyze content posted in Twitter and to review whether post and conversation on Twitter will be highly related to terrorism intention or another way around. This study deployed Naïve Bayes classification technique which identified Twitter contents in Indonesian national language. The method has been processed text pre-processing, and dataset divided with hold out technique. Result of F-measure value indicates that 76% and 77% of texts are associated with the accuracy level of terrorism based on macro-averaging and micro-averaging indicators. The finding is contributed to the scanty literature on the early warning detection method in Indonesian language and assist the government to target the extremists' organizations.**

Keywords— terroris, twitter, text categorization, Naïve Bayes

I. INTRODUCTION

The number of text documents accessible online in the Internet has grown exponentially [1, 2]. Social media usage also has enormous increment [3]. Social media like Twitter and Facebook can be used for sharing, expressing and exchanging messages [3]. Like other social media users, there were several ways used by the Islamic State of Iraq and Syria (ISIS) to recruit new members. Since an increasing number of internet users in the world, the online propaganda was chosen as the effective method by ISIS rather than offline approach. Social media is used to spread the ideology and recruit new member by extremist organization [3, 4]. A message is one type of the text document. So, the text documents from social media could be advantageous and important [2, 5, 6].

Research by [1] revealed opportunity using text mining to analyze the content of document texts in social media by extracting the context from them. Text mining is the discovery by computer of the new and previously unknown information from the textual data [1]. Many prior studies which deploy text mining method, namely text categorization [3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]. This process has performed by classify the text which leads to a model building through data training, subsequently, the text would be assigned to predefined categories based on their content [7]. The sources that used for categorization can be from available open datasets [8, 9, 12], e-mails [10], medical documents [15, 20, 21], and social media data [3, 4, 16, 17, 18, 19]. Text categorization can be used for analyzing the content of text documents which was available online in social media sites.

Naïve Bayes (NB) algorithm was used for text categorization in many researches [1, 2, 3, 7, 8, 10, 11, 12, 13, 14, 16, 17, 18, 19]. Besides English text documents, text categorization using NB can be implemented in Chinese [1] and Arabic languages [2, 19]. NB classifier has advantages such as: (1) simple [1, 2, 8, 10, 11, 13, 14, 16], (2) effective when for huge data [11, 13], (3) fast [1, 8, 10], (4) efficient [10, 14, 16], (5) high accuracy [1, 11], (6) capability to predict the parameters essential for classification using a small amount of training data [10], and (7) independence assumptions [1, 14, 16].

There are many previous scholars proposed text categorization to detect terrorism contents [3, 4, 7, 8, 9, 10, 11, 12, 17, 18, 19]. Three common text categorization methods used in those studies were Naïve Bayes [7, 8, 10, 11, 12, 17, 18, 19], Decision Tree [7, 8, 10, 11], and Support Vector Machine [7, 8, 10, 12, 18, 19]. Most of the previous studies used English text documents [3, 7, 8, 10, 11, 12, 17, 18]. This will make the Indonesian government agency hard to identify the terrorism intention.

Hence, the goal of this study is to deploy a classifier model to analyze Twitter content to identify whether it contains terrorism or not using Indonesia text documents as the inputs using Naïve Bayes classifier. This model is considered to support the early warning detection of terrorism for the Indonesian government. Text categorization using NB is deployed to classify Twitter contents due to: (1) It is simple and easy to implement and (2) It doesn't require as much training data [7].

13-14 August 2020