

CHAPTER 4

DETECTING HATE SPEECH THROUGH MACHINE LEARNING

F. H. A. SHIBLY,¹ UZZAL SHARMA,² and H. M. M. NALEER³

¹PhD Research Scholar, Assam Don Bosco University and Senior Lecturer, South Eastern University of Sri Lanka, Sri Lanka, E-mail: shiblyfh@seu.ac.lk

²Assistant Professor, Assam Don Bosco University, Assam, India

³Senior Lecturer, South Eastern University of Sri Lanka, Sri Lanka

ABSTRACT

Hatred and abusive speeches are identified as huge crime that have been incrementing very recent years, and this has been not only in the specific interaction done face to face but preferably also in the online sharing of information. A considerable number of factors have been contributing to this. There has been a specific study that has well provided a kind of critical overview on the way of detecting such speeches within post or text has been highly evolved over the last few years. However, it has been observed that there are few studies that have been published in detecting hate speech automatically from the perspective of computer science. There is a great way or process with the support of which hate speech can be detected. It has been observed that both the automatic speech recognition and machine learning (ML) have been together complementing each other in the current past, and this has been because both the paradigms are very much deeply ingrained in each other. Therefore, this chapter aims to find out the relationship between hate speech detection and ML and find out the feasible ML algorithms to control hate speeches in social media. It has been observed that there is the implementation of a huge range of