SAPTHAGIRI COLLEGE OF ENGINEERING

14/5, Chikkasandra, Hesaraghatta Main Road, Bengaluru – 560057.

Department of Computer Science and Engineering

Certificate

Certified that the Project Work entitled **"NETSPAM: A NETWORK-BASED SPAM DETECTION FRAMEWORK FOR REVIEWS IN ONLINE SOCIAL MEDIA"** carried out by **GOVINDARAJU B.G (1SG15CS405), POOJA RAMAPPA MAGADUM (1SG15CS410), PRIYA S (1SG15CS412), RAHUL B.R (1SG15CS416)**, bonafide students of **Sapthagiri College of Engineering**, in partial fulfillment for the award of **Bachelor of Engineering** in **Computer Science and Engineering** of **Visvesvaraya Technological University, Belagavi** during the academic year 2017-2018. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the department library. The project report has been approved as it satisfies the academic requirements in respect of **Project Work (10CS85)** prescribed for the said degree.

13/6/18	13/06/18	
Signature of the Guide	**Signature of the HOD**	**Signature of the Principal**
Latha A	**Dr. Yogish H K**	**Dr. K L Shivabasappa**
Asst. Prof., Dept. of CSE	**Professor & Head**	**Principal**
	Professor & Head of the Department Computer Science Engineering Sapthagiri College of Engineering Bangalore-57	Principal Sapthagiri College of Engineering Chikkasandra, Hesaraghatta Road Bangalore- 560 057

EXTERNAL EXAMINATION

Name of the Examiners	**Signature with Date**
1. MANJESH B V	13/6
2. Chaitra	13/06/18

ABSTRACT

Nowadays, a big part of people rely on available content in social media in their decisions (e.g. reviews and feedback on a topic or product). The possibility that anybody can leave a review provides a golden opportunity for spammers to write spam reviews about products and services for different interests. Identifying these spammers and the spam content is a hot topic of research and although a considerable number of studies have been done recently toward this end, but so far the methodologies put forth still barely detect spam reviews, and none of them show the importance of each extracted feature type. In this study, we propose a novel framework, named NetSpam, which utilizes spam features for modeling review datasets as heterogeneous information networks to map spam detection procedure into a classification problem in such networks. Using the importance of spam features help us to obtain better results in terms of different metrics experimented on real-world review datasets from Amazon websites. The results show that NetSpam outperforms the existing methods and among four categories of features; including review behavioral, user-behavioral, review linguistic, user-linguistic, the first type of features performs better than the other categories.