شبكة المعلومات الجامعية

بسم اللّه الرحمن الرحيم

شبكة المعلومات الجامعية

التوثيق الالكتروني والميكروفيلم

شبكة المعلومات الجامعية

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

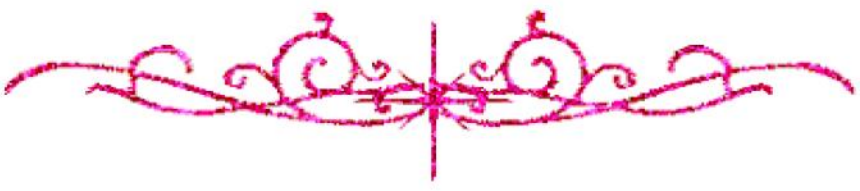
نقسم باللّه العظيم أن المادة التي تم توثيقها وتسجيلها علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات

يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار

شبكة المعلومات الجامعية

بعض الوثائق الأصلية تالفة



بالرسالة صفحات
لم ترد بالأصل

Computer Science Department
Faculty of Computer and Information Sciences
Ain Shams University

Deep Learning for Traits Detection Using Social Networks Interactions

A Thesis submitted to Computer Science Department, Faculty of Computer and Information Sciences, Ain Shams University, in partial fulfillment of the requirements for the degree of Master of Science in Computer and Information Sciences.

By:

Eman Mohammed Hamdi El-Sayed Ibrahim

Supervised by:

Prof. Dr. Mostafa Mahmoud Aref

Professor at Computer Science Department
Faculty of Computer and Information Sciences
Ain Shams University

Dr. Sherine Rady Abdel Ghany

Associate professor at Information Systems Department
Faculty of Computer and Information Sciences
Ain Shams University

Cairo, Egypt
2020

Acknowledgment

First, I am grateful to Almighty ALLAH (S.W.T), the most Generous and the most Merciful. I thank God for providing me with strength and patience to complete this work.

Second, I would like to express my sincere gratitude to my supervisors; Prof. Dr. Mostafa Aref for his support, encouragement and guidance and Dr. Sherine Rady for all the support, patience and special supervision.

Third, I would like to thank my family and my husband for the love and care they give me constantly. Thank you for being with me through the good and the bad times.

My dear friends who have kept on encouraging me to get this work done; Eman Reda, Fatima El-Zahraa, Ghada Hamed, Hadeer Elsaadawy, Hanan Hindy and Marwa Salah. It was a precious journey with your companionship.

At last, I would like to thank all my professors, colleagues and students who believed in me. Thank you for being in my life.

Abstract

Social media networks are one of the main platforms used on a daily basis by millions of people. Feelings, emotions, and opinions are expressed by posting text, images, and videos to express self or to communicate with other people. Using text, the words reflect emotions and indicate behavior towards different topics. Detecting emotions and sentiments helps in many directions such that marketing, political orientation and product reviews. A huge amount of textual data is currently available. There is a need for detecting emotions from social media networks which enhances the machine understanding of humans' perspective.

This thesis contributes to detect positive and negative emotions from short text -tweets- by proposing a deep Convolutional Neural Network (CNN) using different types of word embeddings. CNN is the main building block of the proposed model and is responsible for extracting high-level features from low-level features. Word embeddings are the features fed to the model. The model is constructed by three CNN streams, where each CNN stream contains an embedding layer, a convolutional layer and a max-pooling layer. CNN streams are concatenated and followed by a fully connected layer for classifying text into a positive or a negative emotion class. The used textual features are different types of word embeddings including randomly initialized word embeddings and pre-trained word embeddings. The used pre-trained word embeddings are of different variants such as Word2Vec, Glove and fastText models. The word embeddings in both setups are

trainable and updated through the training phase. After training, the model learns relations between words and generates task-specific word embeddings.

The proposed model has been tested on the Stanford Twitter Sentiment (STS) dataset for classifying emotions. Experiments indicated that the achieved accuracy is 78.5% when using the randomly initialized word embeddings. The accuracy increases 3.6% when using fasttext pretrained word embeddings, 4.5% when using Glove pretrained word embeddings and 6.4% when using Word2Vec pretrained word embeddings. The best tuning for the model is when using Word2Vec pretrained word embeddings which achieves 84.9% accuracy. It is concluded that using CNNs in emotion detection from text is very promising as even when using randomly initialized word embeddings it achieves 78.5% accuracy without any external dataset. Also, not only randomly initialized word embedding can achieve good accuracy in emotion detection from text, it is proven that the power of the pretrained word embeddings helps to achieve a higher competitive accuracy in emotion detection from text.

Keywords: Deep Learning, CNN, Sentiment Analysis, Emotion Detection, Social Media Networks, Word Embeddings.

List of Publications

1- Eman Hamdi, Sherine Rady and Mostafa Aref, “A Survey on Mental Illness Detection using Language via Social Media Networks,” Proceeding of The Seventeenth Conference on Language Engineering (ESCOLE), Cairo, Egypt, 2017.

2- Eman Hamdi, Sherine Rady and Mostafa Aref, “A Convolutional Neural Network Model for Emotion Detection from Tweets,” Proceeding of The Fourth International Conference on Advanced Intelligent Systems and Informatics (AISI), vol. 845, pp. 337-346, Springer, Cairo, Egypt, 2018.

3- Eman Hamdi, Sherine Rady and Mostafa Aref, “A Deep Learning Architecture with Word Embeddings to Classify Sentiment in Twitter,” Proceeding of The Sixth International Conference on Advanced Intelligent Systems and Informatics (AISI), vol. 1261, pp. 115-125, Springer, Cairo, Egypt, 2020.

Table of Contents

List of Figures