Name: Ashraf Kamal

Supervisor: Dr. Jahiruddin

Co-Supervisor: Prof. Muhammad Abulaish

Title: A Unified Data Mining Approach for Detecting Figurative Language in Twitter

Department: Department of Computer Science, Faculty of Natural Sciences, Jamia Millia Islamia

Due to the emergence of Web 2.0 and easy availability of the Internet world-wide, huge amount of online User-Generated Contents (UGCs) are proliferated on a very large-scale. The UGCs are embedded within unstructured documents and mainly found in online social networking platforms, merchant sites, forums, discussion groups, and blogs. Twitter is a popular micro-blogging social networking platform, which facilitates users to share their viewpoints, opinions, ideas, and emotions. As a result, it generates an enormous amount of textual data through users' posts (aka tweets). Although Twitter has a reputation of being a beneficial source of both structural and textual data for many real-life applications including sentiment analysis, it contains numerous natural language nuances and non-literal expressions. One of the major obstacles that occur due to such nuances is the formation of different Figurative Language (FL) categories, like sarcasm, humor, irony, satire, simile, metaphor, and hyperbole. Due to informal way of communication over the online social media, users are prone to use different FL categories, and it is emerging in everyday conversation threads and individual posts of the users at an unprecedented rate. As a result, it is a highly challenging task for the machines to interpret the semantics of the users' posts in their true sense. Understating FL is very elusive for both machines and humans, because it involves continuous variations, non-literal ambiguous words and phrases, implicit sentiments, domain knowledge, contextual information, incongruity, and opposite polarity. Thus, the development of text information processing applications is a difficult and obscure task due to the presence of FL, because it has immense potential to significantly undermine the efficacy of the text information processing systems.

The research presented in thesis is an attempt to propose a unified data mining approach for detecting different FL categories, mainly in Twitter, which is a rich source of textual data containing FL. The unified approach aims to develop Machine Learning (ML) and Deep Learning (DL)-based classification techniques for detecting four different FL categories viz sarcasm, humor, irony, and satire in Twitter. These FL categories are extensively found on the Web; and over the years, they have become a pervasive phenomenon mainly in short texts like tweets. This thesis also presents a fine-grained analysis of three FL categories -- sarcasm, humor, and irony, and proposes classification techniques for computational detection of self-deprecating sarcasm, self-deprecating humor, and situational irony, respectively.

Date
Sign: 08/10/21

Out of the aforementioned FL categories, sarcasm is the most popular FL which is widely used in online social media to express anger or disregard someone using ridicule tone in a lighter note. As a result, a number of research efforts have been made for sarcasm detection in textual data. However, most of the existing approaches lack in determining the implicit semantics of the sarcastic texts. To this end, we have proposed a new DL-based model, BiCapsulAtion, which consists of an input, embedding, Bidirectional Long Short-Term Memory (BiLSTM), capsule network, and attention layers. The proposed model uses a feature vector consisting of 35 auxiliary feature values to enhance its classification accuracy. The efficacy of BiCapsulAtion is empirically evaluated over three Twitter datasets, and it is also compared with two state-of-the-art and six baselines methods. The experimental results reveal that BiCapsulAtion performs significantly better in comparison to the state-of-the-art and baselines methods.

Self-deprecating sarcasm is a specialization of sarcasm, in which users apply sarcasm over themselves. Interestingly, self-deprecating sarcasm is extensively used in various social media platforms including Twitter. Although, self-deprecating sarcasm detection has been a well-studied problem in interdisciplinary sciences, very rare research efforts have been made for its computational detection. To this end, we have proposed a new ML-based approach for detecting self-deprecating sarcasm in Twitter data. It follows a two-layer approach in which first layer employs a set of rules to identify candidate self-around (self-referential) tweets, and second layer employs feature engineering and classification algorithms over the candidate self-around tweets to determine whether a tweet is self-deprecating sarcasm or not. The efficacy of the proposed approach is empirically evaluated over a Twitter dataset. It is also compared with three state-of-the-art approaches. The experimental results reveal that the proposed approach performs significantly better in comparison to the state-of-the-art approaches. In addition to ML, we have also proposed a DL-based model, CAT-BiGRU, which consists of an input, embedding, convolutional, bi-directional gated recurrent unit, and two attention layers. The efficacy of CAT-BiGRU is empirically evaluated over seven Twitter datasets, and it is compared with two state-of-the-art and five baselines methods. The experimental results reveal that CAT-BiGRU performs significantly better in comparison to the state-of-the-art and baselines methods.

Humor is another FL category, which is mainly used in human communication with an aim to exhibit emotions and sentiments. Like sarcasm, computational detection of humor in UGCs is also a challenging problem due to its complex semantic structure. Self-deprecating humor is a specialization of humor in which users generally criticize and put themselves down, and it is generally found in various online contents, such as celebrities' interviews, politician speeches, and advertisements. Like self-deprecating sarcasm, self-deprecating humor detection is a well-studied problem in interdisciplinary sciences, but there is little research effort for its computational detection. To this end, we have identified commonalities between self-deprecating sarcasm and self-deprecating humor, and proposed a novel ML-based approach for detecting self-deprecating humor in UGCs. It follows a two-layer design approach in which

first layer implements a semi-automated process to filter candidate self-around instances, and second layer focuses on feature engineering and classification tasks for detecting self-deprecating humor. The identified features belong to three different groups viz self-deprecating pattern, exaggeration, and word embedding to capture different aspects of the input texts for classification purpose. The efficacy of the proposed approach is empirically evaluated over three datasets, and it is compared with a state-of-the-art approach. The experimental results reveal that the proposed approach performs significantly better in comparison to a state-of-the-art approach.

Irony is another FL category which is widely used in online social media using polarity reversal phenomenon. Computational detection of irony mainly refers to verbal irony, and it is a well-studied problem in the field of computational science. However, situational irony has been overlooked by the researchers. Situational irony represents a situation which makes strange and unexpected outcome due to the presence of incongruity, and it produces a comical impression on the audience. Like verbal irony, situational irony has a visible presence in social media posts. It is also used in digital marketing-based advertisements, such as memes and viral videos. To this end, we have proposed a new DL-based model, CAL-SNN, detecting situation irony in Twitter. Starting with a data splitting technique, which splits input texts into two fragments, CAL-SNN employs two identical sub-networks to accept input text fragments in parallel to identify incongruous difference between them for detecting situational irony. Each sub-network contains input, embedding, convolutional neural network, long short-term memory, and attention layers, followed by a common subtract layer which measures the incongruous differences between the two sub-networks. The efficacy of CAL-SNN is empirically evaluated over four Twitter datasets, and it is compared with two state-of-the-art and five baselines methods. The experimental results reveal that CAL-SNN performs significantly better in comparison to the state-of-the-art and baselines methods.

Satire is also a FL category that has significant presence in online contents in the form of satirical news, customer reviews, blogs, articles, and tweets. Recently, computational detection of satire has become an emerging research problem in the field of computational science, and most of the researchers highlighted the role of context in satire detection. However, existing literatures mostly consider formal document- or sentence-level textual data, and there is little effort towards satire detection in informal short texts like tweets or social media posts. To this end, we have proposed a new DL-based model, BiSAT, in which a self-attention layer is used between two BiLSTMs to capture contextualized satirical information. The proposed model uses a feature vector consisting of 13 auxiliary feature values to enhance the classification accuracy. The efficacy of BiSAT is empirically evaluated over three datasets, and it is compared with a state-of-the-art and five baselines methods. The experimental results reveal that BiSAT performs significantly better in comparison to a state-of-the-art and baseline methods.

* * *