Identification of Emergency Blood Donation Request on Twitter

Puneet Mathur\textsuperscript{1}, Meghna Ayyar\textsuperscript{2}, Sahil Chopra\textsuperscript{3}, Simra Shahid\textsuperscript{4}, Laiba Mehnaz \textsuperscript{4}, and Rajiv Ratn Shah\textsuperscript{2}

\textsuperscript{1}Netaji Subhas Institute of Technology, NSIT-Delhi
\texttt{pmathur3k6@gmail.com}
\textsuperscript{2}Indraprastha Institute of Information Technology, IIIT-Delhi
\{meghnaa, rajivratn\}@iiitd.ac.in
\textsuperscript{3}Maharaja Surajmal Institute of Technology, MSIT-Delhi
\texttt{sahilc.msit@gmail.com}
\textsuperscript{4}Delhi Technological University, DTU-Delhi
\{simrashahid_bt2k16, laibamehnaz\}@dtu.ac.in

Abstract

Social media-based text mining in healthcare has received special attention in recent times due to the enhanced accessibility of social media sites like Twitter. The increasing trend of spreading important information in distress can help patients reach out to prospective blood donors in a time bound manner. However such manual efforts are mostly inefficient due to the limited network of a user. In a novel step to solve this problem, we present an annotated Emergency Blood Donation Request (EBDR) dataset\textsuperscript{1} to classify tweets referring to the necessity of urgent blood donation requirement. Additionally, we also present an automated feature-based SVM classification technique that can help selective EBDR tweets reach relevant personals as well as medical authorities. Our experiments also present a quantitative evidence that linguistic along with handcrafted heuristics can act as the most representative set of signals this task with an accuracy of 97.89%.

1 Introduction

Sufficiency in the availability of blood in emergency situations can dramatically improve the life expectancy and quality of lives of patients in chronic medical conditions. However, many patients still suffer due to the dual challenges of timely availability and shortfall of required whole blood and blood components. In the case of countries with low rates of blood donation record, blood donation is largely dependent on the families and friends of patients, usually through the word of mouth or peer to peer networking. With the increasing accessibility of social media websites, several instances have emerged where the friends and the family of the patients in need of a blood transfusion have tried to voice their urgent need of blood donation through social media channels. They have reached out to the online community through tweets, Facebook posts, and status updates on popular social media platforms. The effect of such tweets in an emergency situation is largely limited to a user’s first few degrees of connections. Thus, it fails to reach the desired donor within the stipulated critical time.

Problem Statement: Emergency Blood Donation Request (EBDR) detection is the task of identifying tweets that explicitly or implicitly mention a necessity for an urgent blood donor.

1.1 Challenges

The key challenges in preparing the corpus are:

1. More than one topic categorization: The task of EBDR tweets prediction is not limited to segregation into binary classes on the basis of certain keywords (like urgent need, blood required, etc.). It involves identification of multiple textual modalities such as blood group, quantity of blood required, the disease being treated and presence of personal details for authentication.

2. Multiple instances of retweets: It is difficult to obtain a unique set of EBDR tweets as many of the instances of such tweets extracted through the search API are retweets by immediate connections of the original tweet author. In many cases, the tweets describing the same events are spread by rephrasing, morphing or editing the original tweets causing duplication of tweet instances.

1.2 Contributions

The main contributions can be summarized as:

\textsuperscript{1}The dataset and code are available for research purposes at \url{https://github.com/pmathur5k10/EBDR}
• Building a corpus of annotated emergency blood donation request tweets divided into three separate datasets using different tweet extraction methodologies.
• Extraction of ancillary handcrafted features from tweets pertaining to the specifications of the requested blood donation.
• Feature modeling using four independent sets of tweet features: linguistic, handcrafted, user specific metadata and textual metadata for the purpose of tweet classification followed by determination of the most relevant set of auxiliary features for SVM based classification.

2 Related Work

Several successful attempts in health text mining have shown that social media can act as a rich source of information for public health monitoring (Broniatowski et al., 2015). MA and Eldredge developed an annotated dataset from consumer drug review posts on social media. Twitter data has been used previously for identifying mentions of medication intake (Mahata et al., 2018a,b), monitoring prescription drug abuse (Hanson et al., 2013). The domain of health text mining also extends to include mental health. Past work on identifying hateful behaviour (Mathur et al., 2018a,b) and suicidal behavior on Twitter (Sawhney et al., 2018a).

3 Dataset

3.1 Dataset Creation

The tweets were collected between 10 May 2018 to 10 July 2018 using Twitter Streaming API. The major problem of extracting tweets for solicitation of blood donation was the infrequent and sporadic nature of their occurrence. Due to the limitation of time restriction imposed on the search query based retrieval, two parallel strategies were developed to build three separate datasets:

1. Personal Donation Requests (PDR) Dataset

(1311 EBDR, 1511 non-EBDR): A curated list of 53 medical phrases was extracted from selected online blood donation information portals such as American Red Cross\(^2\), Australian Red Cross\(^3\) and NHS Blood and transplant\(^4\) by using TF/IDF (Ramos et al., 2003) to identify the most frequently occurring terms. A few such phrases have been depicted in Fig. 1 which were used to query tweets related to blood donation after removing the stop words. Apart from them, tweets mined by using general medical terms were incorporated to form the complete dataset.

2. Blood Donation Community (BDC) Dataset

(1889 EBDR, 3268 non-EBDR): The list of users present in the tweets in PDR dataset was obtained. Specific Twitter handles of community blood donation groups were identified from these users and historical tweets from their timeline were extracted for the positive class. For the negative class, past tweets from extraneous users were collected.

3. Dataset HO: (741 EBDR, 1072 non-EBDR):

This represents the held-out dataset, with tweets collected using both approaches stated above and no overlap with PDR and BDC.

We utilized the traditional query based tweet mining practice which made the PDR dataset more generic in nature. In addition, we employed a Twitter handle supervision technique in BDC dataset to focus on a larger tweet corpus heavily specific to the positive class. Lastly, the held out dataset was created using both the techniques in conjunction for a fair assessment of real world cases. The collected corpus of tweets was filtered down to remove tweets involving non-English text using Ling-Pipe, non-Unicode characters, duplicate tweets, and tweets containing only URL’s, images, videos or having less than 3 words.

Figure 1: Word cloud of EBDR tweets

3.2 Data Annotation

The datasets were annotated by two independent human annotators. In the case of conflict amongst the annotators, an NLP expert finally assigned the ground truth annotation for ambiguous tweets. A satisfactory agreement between the annotators was inferred from Cohen’s Kappa score of 0.86 and the
Table 1: Features of tweets in the EBDR dataset.

| Feature Category | Attributes |
|------------------|------------|
| **Linguistic features (L)** | Unigram & Bigram presence and count, TF-IDF vector |
| **User metadata (U)** | Retweet count, presence of source of posting, place of posting, user friends count, user followers count, user favorites count, user status count |
| **Textual metadata (T)** | Count of URL’s, hashtags, user mentions and special symbols |
| **Handcrafted features (H)** | Presence of name of reference contact, name of place of requirement, contact number, name of hospital/blood bank, blood group required, quantity of blood required, patient disease information |

Table 2: Distribution of hand crafted features of EBDR tweets across Personal Donation Request, Blood Donation Community and Held-Out datasets

| Handcrafted Features | PDR | BDC | HO |
|----------------------|-----|-----|----|
| Name of Reference contact | 1117 | 1513 | 171 |
| Place of requirement | 1188 | 1844 | 522 |
| Contact number | 1142 | 1783 | 541 |
| Hospital/Blood bank | 1059 | 1832 | 525 |
| Required blood group | 1227 | 1829 | 701 |
| Patient Disease Description | 64 | 267 | 109 |
| Quantity of blood required | 802 | 1842 | 156 |
| **Total EBDR tweets** | 1311 | 1889 | 741 |

5 Evaluation

Table 3 shows the performance of datasets BDC, PDR and HO, where the three datasets have been trained using SVM classifier (Chang and Lin, 2011) by taking a combination of one or more feature sets mentioned in Section 4. The train-test split in each case was fixed to 70:30 and stratified five-fold cross-validation performance is reported.
to account for any imbalance of tweet classes in the datasets that may occur. In each case, linguistic features achieve a marginally better accuracy as compared to the handcrafted features when trained separately, but outperform all other combinations of features when utilized in pair. The extensive under-performance due to inclusion of textual and user metadata prove that these feature sets poorly correlate with the positive class. Dataset BDC consists of a more number of samples having a direct correlation with emergency blood donation requests, as opposed to dataset PDR having a greater abundance of samples relevant to the topic of blood donation. This leads to a higher precision but lower recall in evaluation of Linguistic and Handcrafted feature based PDR dataset. In contrast, the datasets PDR and HO, show a better score, implying the ability to effectively identify posts of EBDR class, thereby reducing the false positive cases. Also, despite the downside of PDR dataset in terms of accuracy, the evaluation metrics follow a similar trend. The best performance in terms of F1-score is shown by using linguistic and handcrafted features in all the three datasets.

The HO dataset performs better in terms of accuracy (97.89%) as compared to both PDR and BDC, implying that training classifiers with tweets covering various other topics and aspects increases its robustness towards noise.

### 5.1 Error Analysis

Some categories of errors that were noticed are:

1. **Rants due to non-availability of blood donors**: Tweets like “Can’t believe we live in a pathetic world, no one came forward to donate a single bottle of B+ve blood ...#HumanityIsDead” are an example of reactionary posts. Such posts do not belong to EBDR. However the supervised classifiers classify such tweets into the same, making it difficult to separate false requests from genuine cases.

2. **Acknowledgment of blood donation**: The tweet “We thank @user for registering as #AB- blood donor ...” was correctly identified by the human annotators but misclassified by the automated classifiers. This can be attributed to the inefficiency of the classifiers to derive contextual meaning from the tweets.

### 6 Conclusion and Future Work

In this paper, we introduced a robust feature based classification system in addition to an annotated corpus to accurately identify Emergency Blood Donation Request (EBDR) tweets and separate them from other unrelated blood donation communication, referred to as non-EBDR tweets. Given the diverse nature of emergency request tweets, we adopted a two-way corpus construction strategy. We mine three datasets to probe various aspects such as robustness and accuracy and manually annotated them to validate the performance of the proposed classification system. In addition we also perform an analysis of the efficiency of four independent feature sets extracted from the tweets. The results point out that the linguistic features like n-grams and TF-IDF statistics along with handcrafted features related to blood donation requirement are best suited for classification.

The EBDR corpus can benefit researchers in various aspects including but not limited to: (i) automatic evaluation of emergency blood donation requests from health posts, (ii) named entity extraction of patient details, blood group and quantity requirement statistics with the help of handcrafted features provided with the tweets, (iii) crisis assessment and management through social media monitoring of medical emergency events and (iv) feature modelling using genetic algorithms as done by Sawhney et al. (2018b,c).

### References

David Andre Broniatowski, Mark Dredze, Michael J Paul, and Andrea Dugas. 2015. Using social media to perform local influenza surveillance in an inner-city hospital: a retrospective observational study, *JMIR public health and surveillance*, 1(1).

Meeyoung Cha, Hamed Haddadi, Fabricio Ben-
Chih-Chung Chang and Chih-Jen Lin. 2011. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):27.

Carl Lee Hanson, Ben Cannon, Scott Burton, and Christophe Giraud-Carrier. 2013. An exploration of social circles and prescription drug abuse through twitter. *Journal of medical Internet research*, 15(9).

Paul Fontelo MA and MD Eldredge. Development of an adverse drug reaction corpus from consumer health posts.

Debanjan Mahata, Jasper Friedrichs, Rajiv Ratn Shah, and Jing Jiang. 2018a. Did you take the pill? detecting personal intake of medicine from twitter. *arXiv preprint arXiv:1808.02082*.

Debanjan Mahata, Jasper Friedrichs, Rajiv Ratn Shah, et al. 2018b. # phramacovigilance-exploring deep learning techniques for identifying mentions of medication intake from twitter. *arXiv preprint arXiv:1805.06375*.

Puneet Mathur, Ramit Sawhney, Meghna Ayyar, and Rajiv Shah. 2018a. Did you offend me? classification of offensive tweets in hinglish language. In *Proceedings of the Second Workshop on Abusive Language Online*.

Puneet Mathur, Rajiv Shah, Ramit Sawhney, and Debanjan Mahata. 2018b. Detecting offensive tweets in hindi-english code-switched language. In *Proceedings of the Sixth International Workshop on Natural Language Processing for Social Media*, pages 18–26.

Juan Ramos et al. 2003. Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning*, volume 242, pages 133–142.

Raquel Recuero, Ricardo Araujo, and Gabriela Zago. 2011. How does social capital affect retweets? In *ICWSM*.

Ramit Sawhney, Prachi Manchanda, Raj Singh, and Swati Aggarwal. 2018a. A computational approach to feature extraction for identification of suicidal ideation in tweets. In *Proceedings of ACL 2018, Student Research Workshop*, pages 91–98.

Ramit Sawhney, Puneet Mathur, and Ravi Shankar. 2018b. A firefly algorithm based wrapper-penalty feature selection method for cancer diagnosis. In *International Conference on Computational Science and Its Applications*, pages 438–449. Springer.

Ramit Sawhney, Ravi Shankar, and Roopal Jain. 2018c. A comparative study of transfer functions in binary evolutionary algorithms for single objective optimization. In *International Symposium on Distributed Computing and Artificial Intelligence*, pages 27–35. Springer.