INDRAPRASTHA INSTITUTE of
INFORMATION TECHNOLOGY DELHI

MIDAS@IIITD
Multimodal Digital Media Analysis Lab

Identification of Emergency Blood Donation Request on Twitter

## Puneet Mathur, Meghna Ayyar, Sahil Chopra, Simra Shahid, Laiba Mehnaz, Rajiv Ratn Shah

Research Goal

MOTIVATION: Social media sites like Twitter helps in spreading important information for patients in distress to 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.

PROBLEM STATEMENT: Classification of emergency blood donation request (EBDR) and subsequent extraction of the related details regarding the user requirements can help individuals pair up with blood banks, hospitals, and individuals in a critically time-bound manner. We need an automated EBDR assessment tool that can identify such post on social media platforms.

Dataset Constituent

Feature Category	Attributes
Linguistic features (L)	Unigram & Bigram presence and count, TF-IDF vector
User metadata (U)	Retweet count, presence of source of posting, presence of place of posting, user friends count etc
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, etc
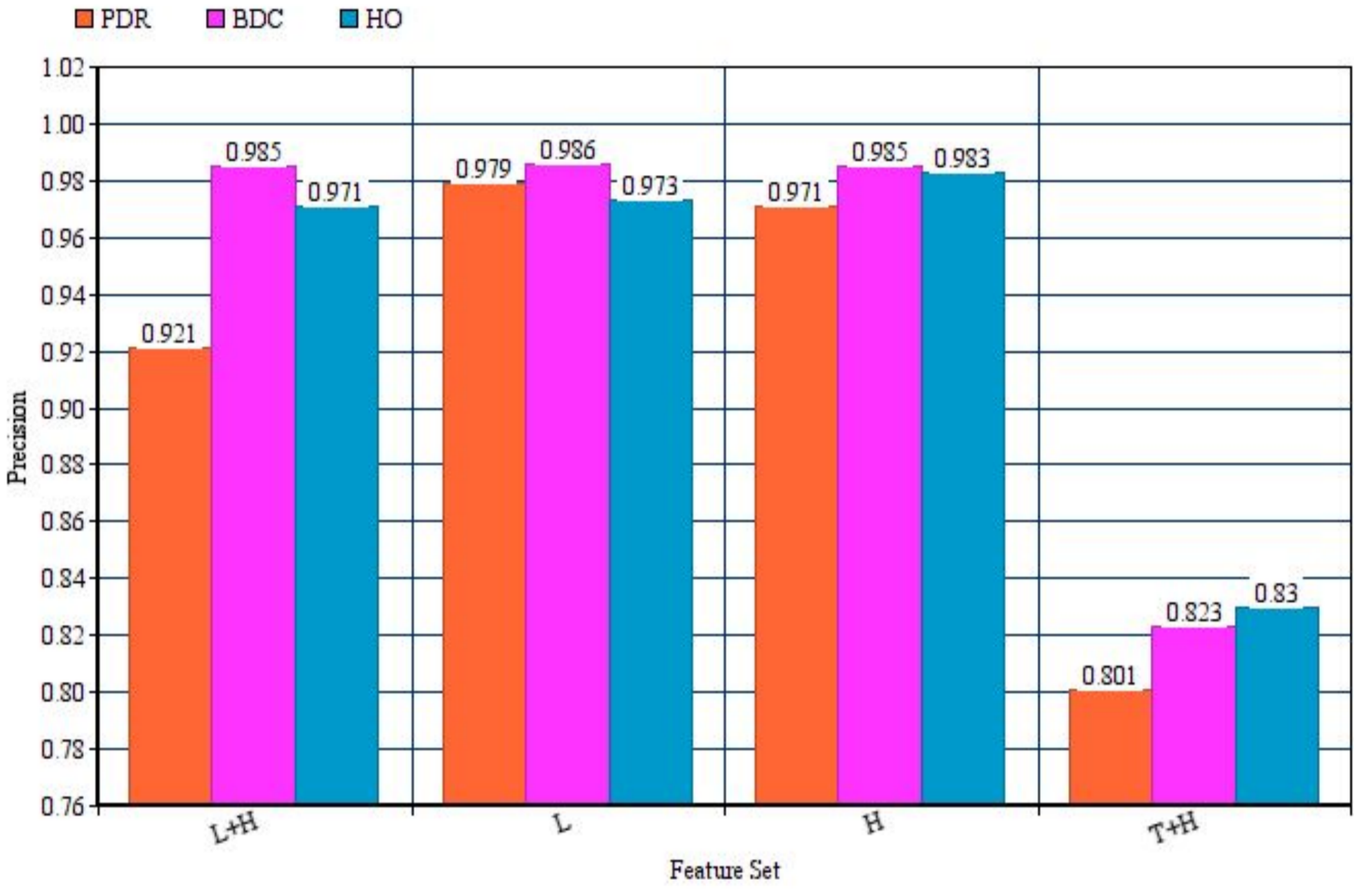
Table: Feature of tweets in EBDR dataset

Figure: Result comparison of feature category

Applications

- Automatic evaluation of EBDR from health posts on social media websites that will help medical agencies and authorities to reach out to patients in time-bound manner.
- Crisis assessment and management through social media monitoring of medical emergency events

- Extraction of patient details, blood group, quantity and other related requirement statistics on the data corpus.

Contributions

- Creation of handcrafted annotated dataset pertaining to specification of emergency blood donation request. This dataset will encourage more research in this area.
- Feature modeling using four independent sets of tweet features: linguistic, handcrafted, user specific metadata and textual metadata.
- Determination of the most relevant set of auxiliary features for SVM based classification.
- Our model performs fairly well with an accuracy of 97.87%.