



# Text Analysis Using Morphological Operations on a Neutrosophic Text Hypergraph

Dhanya PM<sup>1</sup>, Ramkumar PB<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, RSET ; dhanya\_pm@rajagiritech.edu.in

<sup>2</sup>Department of Mathematics, RSET; ramkumar\_pb@rajagiritech.edu.in

\*Correspondence:Email: dhanya\_pm@rajagiritech.edu.in

**Abstract.** Due to the rise in the operation of platforms on social media, there is more opportunity for users to post content online, out of which some tend to be hate speech. Hate speech is found in almost all domains like sports, politics, religion, government affairs, and personal matters. Its detection and removal from platforms like Twitter, Facebook, etc. are tedious. Over the years, a lot of methods have evolved in this area most of which are more time-consuming machine learning methods. Our objective is to find a better method that considers indeterminacy at the word level and sentence level for the detection and removal of hate speech using fuzzy logic applied to Neutrosophic hypergraphs. A neutrosophic hypergraph is a kind of hypergraph where each node and hyperedge has three associated membership functions namely Indeterminacy, Truth and Falsity. Our work has successfully modeled Text documents into neutrosophic hypergraphs and morphological operators like dilation, erosion etc. are applied to it. Using these operations further operators like thinning, thickening, hit-or-miss, and skeletoning are applied. Finally hate speech is identified and removed. This a novel method in this area. The system is tested with Twitter tweets and the results are promising with an accuracy of 88%.

**Keywords:** neutrosophic; hypergraph; morphology ; hate speech

## 1. Introduction

Since lakhs of contents are posted every day on social media platforms, there is more chance for it to be against the rules of a government, religion, and mostly the society. Filtering the contents and making it suitable for everyone to read it is a tremendous job. In most cases, contents are manually detected after mass protest and are removed or deactivated. Since the readability and reachability of the social media content are higher when compared to the printing media or visual media, there should be good and efficient methods for identifying and removing hate speech. Our system has made efforts in this area by applying the concepts of

---

Dhanya P.M, Ramkumar P.B, Text Analysis Using Morphological operations on a Neutrosophic Text hypergraph

neutrosophic hypergraphs and their morphological operators.

Many real-life problems were solved by modeling hypergraphs with neutrosophic sets and logic. Neutrosophic sets are used to deal with uncertainties in such problems. Neutrosophic sets are to deal with this indeterminacy. Morphological operators like dilation, erosion, thickening, thinning, and skeletoning are useful for various text analysis operations which are discussed in this paper with the main focus on hate speech detection and removal.

### 1.1. *Related works*

The proposed work in this paper has applied neutrosophic hypergraph operations for hate speech detection. There has been a lot of research work on detecting hate speech that is mostly seen in social media. In order to classify hate speech, a novel method namely H-CovBi-Caps [1] was implemented that is a deep learning model based on conventional, BiGRU and Capsule model. Evaluation of this model was done using balanced and unbalanced Twitter data sets. This method gave a recall of 0.80 and f-score of 0.84. Another method used natural language processing strategies and data analysis to make providers of social media responsive to hate speech content [2]. The authors claim that they can surpass the state-of-the-art approach in terms of precision, recall, and F1 scores by approximately 10%. There have been works that focus on the lack of transparency and bias experienced by various hate speech detection and mitigation systems [3]. Using SAS Enterprise Miner's Text Analytics [4], the authors demonstrated how to consider the information in the tweets to classify them as hostile. The tweets were subjected to preprocessing and models were applied and analyzed. The authors claim adequate accuracy. Different approaches to hate speech detection are discussed and compared in a survey [5], where the authors have considered various data sets, features, and machine learning models for comparison.

In all these methods even though the authors claim good accuracy, the works lag a proper mathematical modeling and representation. They have the disadvantage of being costly in terms of time and resources. Since our work is concentrating on finding a solution for this with the help of neutrosophic hypergraphs, let us see some works already done on hypergraphs.

The perspective of a single-valued neutrosophic set, its complement, union, difference, properties of set-theoretic operators etc. was introduced in [6]. The structure of a system can be studied by using hypergraph [7] which is the generalization of a graph. A detailed study of fuzzy graphs and fuzzy hypergraphs [8] and related extensions [9], discusses mathematical models of hypergraphs like intuitionistic, complex, m-polar fuzzy, bi-polar, Pythagorean and

q-rung ortho-pair hypergraphs, and also neutrosophic hypergraphs like single-valued, bi-polar and complex. Graph morphology [10] extracted the structural information from graphs using structuring graphs. Lattice structure on hypergraphs are developed and morphological operators [11], [12] are defined by using vertex-hyperedge correspondence. Also, the classical notion of a dilation/erosion of a subset of vertices is extended to sub-hypergraphs. Several opening, closing and alternate sequential filters are also proposed. Morphology applied on Intuitionistic fuzzy hypergraphs are discussed in [13], [14], [15]. Text summarization using morphological filter [16] is done on intuitionistic fuzzy hypergraphs. Crime Analysis [17] done with the application of graph morphology has successfully tracked the crime rate in various areas. More than 200 neutrosophic graphs [18] are discussed, particularly the bipartite neutrosophic graphs, neutrosophic tree and directed neutrosophic graphs applied in cognitive maps, relational maps and relational equations. The perception of neutrosophic incidence graphs that are single-valued, their cut vertex, blocks and bridges are discussed in [19]. The paper has discussed the neutrosophic incidence graphs and their vertex, edge and pair connectivity. Neutrosophic logic and connectors [20] based on set operations are also defined. The idea of constant single-valued neutrosophic graph (CSVNG) [21], which is the modified form of a single-valued neutrosophic graph has also evolved. The authors applied it to Wi-Fi systems and also discussed the consequences. A methodology of decision-making with multiple criteria [22] applied with a neutrosophic set was developed to handle uncertain data, and the authors have used it in the Logistics Service Sector. A novel adaptable method [23] was used with eleven criteria and ten solar panels in PV which used a neutrosophic set to deal with vague data. Another work in IoT [24] intended to introduce a weight product method based on the neutrosophic framework for the assessment of IoT-based cities that are sustainable and smart. The notion of Fermatean neutrosophic dombi fuzzy graph [25] was initiated which constructed the cartesian, direct, composition of such graphs. A neutrosophic method using type-2 neutrosophic numbers [26] was used in the field of study of risks in power plants. A hybrid approach to decision-making using many criteria under a spherical fuzzy environment was introduced in [27].

Most of the hate speech detection methods developed failed to address the ambiguity aspect of it, our method has included an indeterminacy parameter with every word and every sentence. Even though there are many applications with hypergraphs in the area of image processing, networks, text data etc., the proposed method is the first work that has done hate speech detection and removal of it using a neutrosophic text hypergraph. The preliminaries of the neutrosophic hypergraph and the morphological operations are given in sections 1.2 to 1.4. The section 2 focuses on how a document is converted to a neutrosophic hypergraph and how

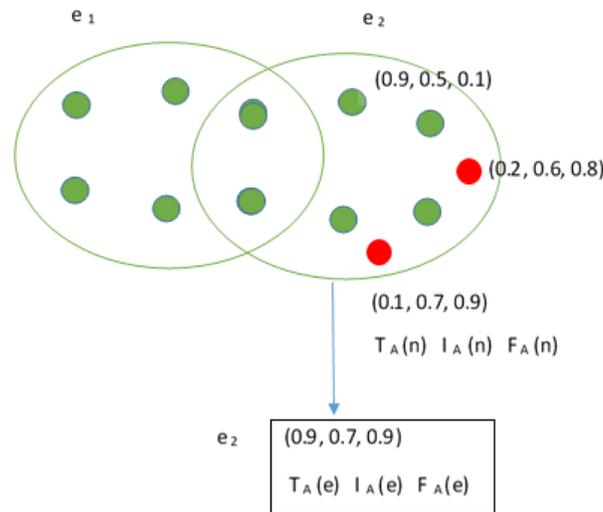


FIGURE 1. Neutrosophic graph with membership degree

operations like hit-or-miss, skeletoning etc. are applied to it. Section 3 deals with operations like thinning and thickening. Section 4 shows how hate speech detection is done using the operations discussed in sections 3 and 4. Finally, section 5 gives a detailed result analysis.

### 1.2. Preliminaries

Let a neutrosophic hypergraph be defined as  $H = (H^n, H^e)$  and is shown in Figure 1, where  $H^n$  is the collection of nodes and  $H^e$  is the collection of hyperedges. For every  $n$  in  $H^n$ ;  $F_A(n) \in [0, 1]$ ,  $I_A(n) \in [0, 1]$ ,  $T_A(n) \in [0, 1]$ , and  $I_A(n) + T_A(n) + F_A(n) \leq 3$ , where  $I_A(n)$ ,  $T_A(n)$  and  $F_A(n)$  are the indeterminacy, truth and falsity value respectively. Set  $A$  which is a neutrosophic set in  $H^n = \{(n, I_A(n), T_A(n), F_A(n)); n \in H^n\}$ . Likewise for every  $e$  in  $H^e$ ,  $T_A(e) \in [0, 1]$ ,  $I_A(e) \in [0, 1]$ ,  $F_A(e) \in [0, 1]$  and  $I_A(e) + T_A(e) + F_A(e) \leq 3$ , where  $I_A(e)$  is the indeterminacy value,  $T_A(e)$  is the truth value, and  $F_A(e)$  is the falsity value. A neutrosophic set  $B$  in  $H^e = \{(e, I_A(e), T_A(e), F_A(e)); e \in H^e\}$ . The edge membership degree,  $(I_A(e), T_A(e), F_A(e))$  is defined as the maximum of respective membership degrees of the nodes and is given by

$$T_A(e) = \vee T_A(n); \forall n \in e \tag{1}$$

$$I_A(e) = \vee I_A(n); \forall n \in e \tag{2}$$

$$F_A(e) = \vee F_A(n); \forall n \in e \tag{3}$$

### 1.3. Special cases of membership values

- case 1:  $[1, 1, 0]$  In weather prediction during the rainy season, the truth value of rain and the indeterminacy is 1. Non-occurrence of rain tends to 0.

- case 2:  $[0, 1, 1]$  In the case of the Nipah virus attack, based on previous experiences in past years, the possibility of a patient being alive is 0. But there is an indeterminacy due to the nature of the virus and the falsity of death is 1. Since indeterminacy is 1, the converse may also happen violating the history and we may get  $[1, 1, 0]$ .
- case 3:  $[1, 0, 1]$  At a particular point of time of hartal or strike, there is a chance of crime or not. Hence at instance  $t$ , the truth value is 1 also falsity can be 1.
- case 4:  $[0, 0, 0]$  In the case of cancer patients, the region not affected by cancer need not be considered for treatment. For this region the indeterminacy is 0, the Truth value is 0 and there is no doubt in the falsity of the disease.
- case 5:  $[1, 1, 1]$  This is a chaotic situation where all the values are 1. In the case of a Tornado, since the system is chaotic, the occurrence of Tornado, the indeterminacy and the falsity is 1.

#### 1.4. Applying morphological operators

Let  $\{H_{NF}, H^n, H^e, (\mu_n, \gamma_n, \kappa_n), (\mu_e, \gamma_e, \kappa_e)\}$  be a neutrosophic hypergraph, where  $\gamma_n$  is the non-membership degree,  $\mu_n$  is the membership degree, and  $\kappa_n$  is the indeterminacy degree defined on a collection of nodes. Let membership degree  $\mu_e$ , non-membership degree  $\gamma_e$  and indeterminacy degree  $\kappa_e$  be defined on a collection of hyperedges of the neutrosophic hypergraph. Here the sum of  $(\mu_n, \gamma_n, \kappa_n) \leq 3$ . Also  $\mu_e$  is the supremum of  $\mu_n$ ,  $\gamma_e$  is the supremum of  $\gamma_n$  and  $\kappa_e$  is the supremum of  $\kappa_n$ .

##### 1.4.1. $(\alpha, \beta, \omega)$ cut of a neutrosophic fuzzy hypergraph

The  $(\alpha, \beta, \omega)$  cut of a neutrosophic hypergraph  $H_{NF}$  is the crisp set of nodes given by  $X_{NF} = H_{\alpha, \beta, \omega} / \alpha \geq m, \beta \geq n, \omega \geq k$  which retrieves a sub hypergraph of  $H_{NF}$ . Once we have  $H_{NF}$ , the parent graph and  $X_{NF}$  as its sub-graph, we can define many morphological operators adjunction, erosion, dilation, closing, and opening filters on it. Figure 2(a) shows a parent neutrosophic hypergraph and Figure 2(b) shows a sub-graph obtained by  $(\alpha, \beta, \omega)$  cut. All the following morphological operations are defined for this parent and sub-hypergraph.

##### 1.4.2. Dilation of $X_{NF}$

The dilation operation can be done to concerning nodes or concerning edges. Dilation concerning nodes can be written as follows:-

$$\delta^n(X_{NF}) = \{n/n \in X_{NF}\} \quad (4)$$

As per e.q(4), it is the collection of nodes present in the sub-hypergraph  $X_{NF}$ . The dilation concerning edges can be written as follows:-

$$\delta^e(X_{NF}^n) = \{e/e \in H_{NF}/n \in X^e\} \quad (5)$$

As per e.q(5), it includes all edges in  $H_{NF}$  such that it contains at least one node in  $X^e$ . Both the dilations are shown in Figure 2(c) and Figure 2(d).

#### 1.4.3. *Erosion of $X_{NF}$*

The erosion operator can be applied in two ways. It can be either concerning nodes or concerning hyperedges. Erosion concerning nodes is written as the following:-

$$\varepsilon^n(X_{NF}^e) = \{n \in X_{NF}/n \notin X_{NF}^{e'}; X_{NF}^{e'} = H_{NF}^e - X_{NF}^e\} \quad (6)$$

According to e.q(6), erosion concerning nodes is defined as the collection of nodes in  $X_{NF}$  which are not present in its complement graph. This is shown in Figure. 2(e). Erosion concerning hyperedges is the collection of edges consisting of nodes of  $X_{NF}$  only. It can be written as the following:-

$$\varepsilon^e(X_{NF}^n) = \{e \in X_{NF}/\forall n \in e n \notin X_{NF}^{e'}\} \quad (7)$$

This is shown in Figure 2(f).

#### 1.4.4. *Adjunction of $X_{NF}$*

We can say that  $(\varepsilon^e, \delta^n)$  are adjunctions iff

$$X_{NF}^e \subseteq \varepsilon^e(Y_{NF}^n) \quad (8)$$

$$\delta^n(X_{NF}^e) \subseteq Y_{NF}^n; X_{NF} \subseteq Y_{NF} \quad (9)$$

#### 1.4.5. *Morphological Opening and Closing*

The morphological opening is of two types:-

- Opening w.r.to edge( $\gamma_e$ )

This Morphological opening

$$\gamma_e = \delta^e(\varepsilon^n(X_{NF}^e)) \quad (10)$$

is a composition of the form  $\delta \circ \varepsilon$  which gives edges in  $X_{NF}$  by applying e.q(6) followed by e.q(5).

- Opening w.r.to node( $\gamma_n$ )

This Morphological opening of  $X_{NF}$  is

$$\gamma_n = \delta^n(\varepsilon^e(X_{NF}^n)) \quad (11)$$

which is a composition of  $\delta \circ \varepsilon$  obtained by applying e.q(7) followed by e.q(4).

- Closing w.r.to edge

This Morphological closing is the set of edges in  $X_{NF}$

$$\phi_e = \varepsilon^e(\delta^n(X_{NF}^e)) \quad (12)$$

which is a composition of  $\varepsilon \circ \delta$  obtained by applying e.q(4) followed by e.q(7).

- Closing w.r.to node

This Morphological closing

$$\phi_n = \varepsilon^n(\delta^e(X_{NF}^n)) \quad (13)$$

is the set of nodes in  $X_{NF}$  which is a composition of  $\varepsilon \circ \delta$  obtained by applying e.q(5) followed by e.q(6)

Repeated application of opening as well as closing operations as mentioned in e.q(10) to e.q(13) results in the same hypergraph. Such operators are called filters. They are shown in Figures 2(g) to 2(j).

## 2. Materials and Methods

### 2.1. Skeleton operation with dilation w.r.to edge

Dilation related to edge is defined as the collection of all edges retrieved from the parent graph  $H$ , which contains all nodes in sub-hypergraph  $X$ . It can be written as  $\delta^e(X^n)$ . The skeleton operation on a graph  $H$ , can be defined as

$$S(H) = H - (\delta^e(X^n))^k \quad (14)$$

Let  $H$  be the hypergraph related to text pertaining to the *sports* domain. Some of the words related to specific sports domains are given in Table 1. Let  $X_1$  be the sub-hypergraph of  $H$ , which is obtained by taking the words related to *cricket*. By applying dilation w.r.to edge,  $\delta^e(X_1^n)$ , we get all the text related to *cricket*. On applying  $S(H_1) = H - (\delta^e(X_1^n))$ , we get a minimal skeleton of *sports* devoid of *cricket*. Now take  $X_2 =$  set of words related to *football*. On applying  $\delta^e(X_2^n)$ , we get all the text related to *football*. Thus  $S(H_2) = S(H_1) - (\delta^e(X_1^n))$  will give us the text devoid of *football*. On repeating this  $K$  times we get the skeleton of sports

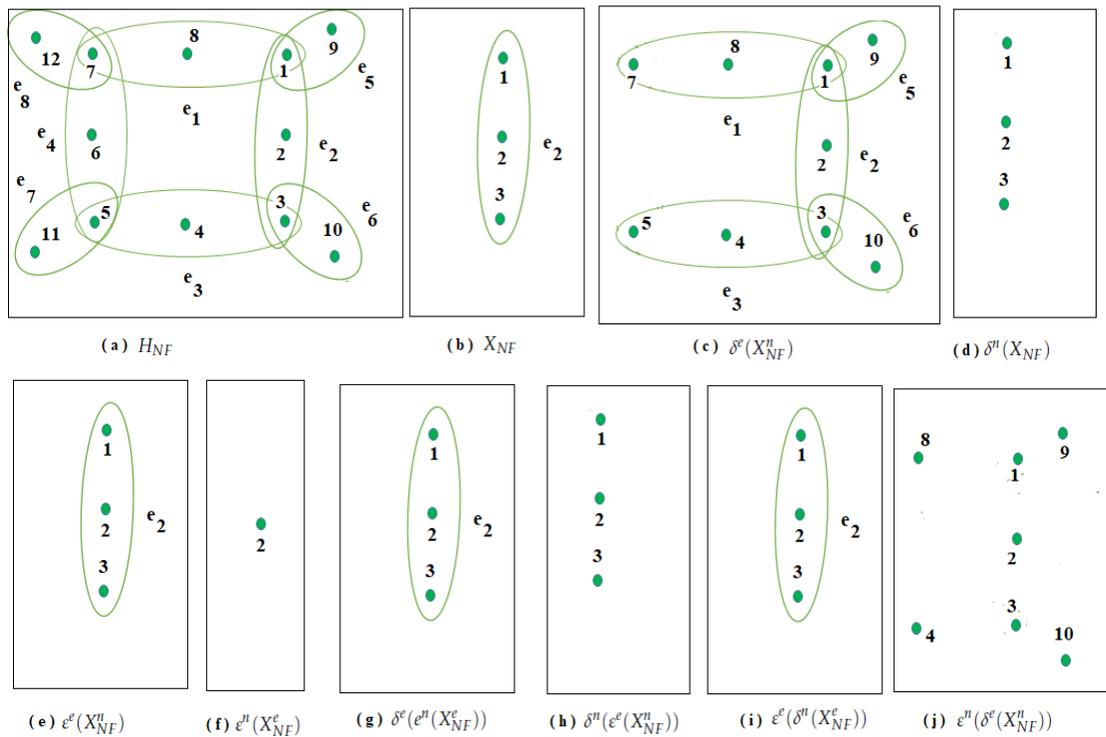


FIGURE 2. Result of morphological operations on a neutrosophic hypergraph

which is devoid of specific sports areas. As a byproduct of this, we get many sub-hypergraphs of  $H$ .

2.1.1.1. *Illustration*

Consider the text given in Figure 3 with words numbered. A hypergraph can be drawn by considering unique words as nodes and sentences as hyperedges. It can be made neutrosophic by giving three degrees to each word based on a criteria. Some of the words in the sports field and the criteria are shown in Table 2. If there are common words across sentences, then edges will overlap as shown in Figure 4. We consider words related to *cricket* first and then apply dilation  $\delta^e(X_1^n)$ . Let  $X_1$  be a sub-hypergraph that consists of words in the *cricket* domain. This dilation is a conditional dilation, which selects the statements in the original text which consists of words in the *cricket* domain. It is subtracted from the hypergraph to get the skeleton  $S(H_1)$ . The first skeleton obtained is shown in Figure 5. Now select sub-hypergraph  $X_2$  which is the set of words in the *football* domain. Apply dilation and select the sentences in the original text related to the *football* domain. On subtracting this we get the next level skeleton which is shown in Figure 6 and Algorithm 1.

TABLE 1. Words related to specific sports

cricket	football
cricket	striker
ICC	Manchester

TABLE 2. Criteria for giving degrees  $T_A(n), I_A(n), F_A(n)$  to the words

Words	$T_A(n)$	$I_A(n)$	$F_A(n)$	Criteria
Cricket	0.9	0.1	0.3	Cricket is a sports game only in a few countries in the world. Even though it is a sports game, it is not seen in the Olympics. So $F_A(n) = 0.3$ and $I_A(n) = 0.1$ . Indeterminacy is less since it is related to sports
ICC				-do-
Tournament				-do-
Football	1.0	0	0	Indeterminacy is 0, falsity is 0. since it is a sports event and seen in Olympics
Manchester				-do-
Olympics				-do-
Badminton				-do-
Game				-do-
Sachin	0.8	0.3	0.3	Depends on person to person and also value varies from person to person. When compared to very popular persons in Football, there is a bit more level of indeterminacy for <i>Sachin</i> for being identified as a sports person.
Stages	0.5	0.5	0.5	This is a word have medium value for all the degrees
Season				-do-
Performance				-do-

<p>അന്താരാഷ്ട്ര ക്രിക്കറ്റ് കൗൺസിൽ ടൂർണമെന്റ് ഇന്ത്യ വിജയിക്കുന്നതിനു സമയത്തിന്റെ പ്രശ്നമാണെന്ന് ഇതിഹാസ ഓസ്ട്രേലിയൻ ക്യാപ്റ്റൻ സ്റ്റീവ് വേ വിശ്വസിക്കുന്നു . 2013ൽ ഇംഗ്ലണ്ടിൽ നടന്ന ചാമ്പ്യൻസ് ട്രോഫി വിജയത്തിനു ശേഷം മെൻ ഇൻ ബ്ലൂ വലിയ ട്രോഫി നേടിയിട്ടില്ല . മത്സരത്തിന്റെ തോൽ കാരണം അത്തരം ടൂർണമെന്റുകൾ വിജയിക്കുക എളുപ്പമല്ലെന്ന് വേ അഭിപ്രായപ്പെട്ടു . ഒരു മാന്സിക ബ്ലോക്ക് കാരണമല്ല ഇന്ത്യ ഐസിസി യുടെ കനോക്ക് നോക്ക് ഔട്ട് ഘട്ടങ്ങളിൽ തകർന്നുകൊണ്ടിരിക്കുന്നത് എന്നും അദ്ദേഹം അഭിപ്രായപ്പെട്ടു .</p> <p>ചൊവ്വാഴ്ചത്തെ ലീഗ് കപ്പ് സെമിഫൈനലിൽ മാഞ്ചസ്റ്റർ യുണൈറ്റഡ് മാനേജർ ഓലെ ഗുനാർ സോൾസ്ജെയർ തന്റെ ആദ്യ പങ്കുതിയിലെ പ്രകടനം, ഈ സീസണിലെ ഏറ്റവും മോശം പ്രകടന മെന് ലേബൽ ചെയ്തു. അംഗീകൃത സ്ട്രൈക്കറില്ലാതെ കളി ആരംഭിച്ചെങ്കിലും ഓൾഡ് ട്രാഫോർഡിൽ നടന്ന മത്സരത്തിൽ 3-1ന് ജയിച്ച സിറ്റി സിറ്റി യുണൈറ്റഡിനെ മറികടന്നു, 21 മിനിറ്റ് ആദ്യ പങ്കുതിയിൽ മൂന്ന് ഗോളുകളും നേടി.</p> <p>ലണ്ടൻ ഒളിമ്പിക്സ് വെങ്കല മെഡൽ ജേതാവ് സൈന മലേഷ്യ മാസ്റ്റേഴ്സ് ബാഡ്മിന്റൺ ടൂർണമെന്റിന്റെ വനിതാ സിംഗിൾസ് ക്വാർട്ടർ ഫൈനലിൽ സൗത്ത് കൊറിയയുടെ ആൻ സെ യൂങ്ങ് തോൽപിച്ചു കടന്നു.</p>	<p style="text-align: center;"><b>Malayalam</b></p> <p style="text-align: center;"><b>Text</b></p> <p style="text-align: center;"><b>and</b></p> <p style="text-align: center;"><b>its</b></p> <p style="text-align: center;"><b>translation</b></p> <p style="text-align: center;"><b>to</b></p> <p style="text-align: center;"><b>English</b></p> <div style="text-align: center; margin-top: 20px;"> </div>																																																																																																																																																																																																																																																			
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center;">1</td><td style="text-align: center;">2</td><td style="text-align: center;">3</td><td style="text-align: center;">4</td><td style="text-align: center;">5</td><td style="text-align: center;">6</td><td style="text-align: center;">7</td><td style="text-align: center;">8</td><td style="text-align: center;">9</td><td style="text-align: center;">10</td><td style="text-align: center;">11</td> </tr> <tr> <td colspan="11"><i>Legendary Australia captain Steve Waugh believes it is only a matter of time before India win a big</i></td> </tr> <tr> <td style="text-align: center;">12</td><td style="text-align: center;">13</td><td style="text-align: center;">14</td><td style="text-align: center;">15</td><td style="text-align: center;">16</td><td style="text-align: center;">17</td><td style="text-align: center;">18</td><td style="text-align: center;">19</td><td colspan="3"></td> </tr> <tr> <td colspan="11"><i>International Cricket Council (ICC) tournament. The 'Men in Blue' haven't won any major trophy since their</i></td> </tr> <tr> <td style="text-align: center;">20</td><td style="text-align: center;">19</td><td style="text-align: center;">21</td><td style="text-align: center;">22</td><td style="text-align: center;">5</td><td style="text-align: center;">23</td><td style="text-align: center;">24</td><td style="text-align: center;">10</td><td style="text-align: center;">16</td><td colspan="2"></td> </tr> <tr> <td colspan="11"><i>Champions Trophy victory in 2013 in England. Waugh stated it is not easy to win such tournaments because of</i></td> </tr> <tr> <td style="text-align: center;">25</td><td style="text-align: center;">26</td><td style="text-align: center;">27</td><td style="text-align: center;">28</td><td style="text-align: center;">28</td><td style="text-align: center;">9</td><td colspan="5"></td> </tr> <tr> <td colspan="11"><i>the level of competition but he also suggested that it is not a mental block because of which India keep on</i></td> </tr> <tr> <td style="text-align: center;">29</td><td style="text-align: center;">30</td><td style="text-align: center;">31</td><td style="text-align: center;">15</td><td style="text-align: center;">16</td><td colspan="6"></td> </tr> <tr> <td colspan="11"><i>crashing in knockout stages of ICC tournaments.</i></td> </tr> <tr> <td style="text-align: center;">32</td><td style="text-align: center;">33</td><td style="text-align: center;">34</td><td style="text-align: center;">35</td><td style="text-align: center;">36</td><td style="text-align: center;">37</td><td style="text-align: center;">38</td><td style="text-align: center;">39</td><td style="text-align: center;">40</td><td style="text-align: center;">41</td><td></td> </tr> <tr> <td colspan="11"><i>Manchester United manager Ole Gunnar Solskjaer has labeled his side's first-half performance in</i></td> </tr> <tr> <td style="text-align: center;">42</td><td style="text-align: center;">43</td><td style="text-align: center;">44</td><td style="text-align: center;">45</td><td style="text-align: center;">46</td><td style="text-align: center;">47</td><td style="text-align: center;">48</td><td style="text-align: center;">49</td><td style="text-align: center;">50</td><td style="text-align: center;">51</td><td style="text-align: center;">52</td><td style="text-align: center;">53</td> </tr> <tr> <td colspan="11"><i>Tuesday's League Cup semi-final, first-leg defeat their worst display of the season. Despite starting the game</i></td> </tr> <tr> <td style="text-align: center;">54</td><td style="text-align: center;">55</td><td style="text-align: center;">56</td><td style="text-align: center;">57</td><td style="text-align: center;">58</td><td style="text-align: center;">33</td><td style="text-align: center;">21</td><td style="text-align: center;">59</td><td style="text-align: center;">60</td><td style="text-align: center;">61</td><td></td> </tr> <tr> <td colspan="11"><i>without a recognized striker, holders City outclassed United in a 3-1 victory at Old Trafford, scoring all three</i></td> </tr> <tr> <td style="text-align: center;">62</td><td style="text-align: center;">63</td><td style="text-align: center;">39</td><td style="text-align: center;">64</td><td colspan="7"></td> </tr> <tr> <td colspan="11"><i>goals in a 21-minute first-half blitz.</i></td> </tr> <tr> <td style="text-align: center;">64</td><td style="text-align: center;">65</td><td style="text-align: center;">66</td><td style="text-align: center;">67</td><td style="text-align: center;">68</td><td style="text-align: center;">69</td><td style="text-align: center;">70</td><td style="text-align: center;">71</td><td style="text-align: center;">72</td><td style="text-align: center;">73</td><td></td> </tr> <tr> <td colspan="11"><i>London Olympics bronze-medallist Saina storms into women's singles quarterfinals of the Malaysia Masters</i></td> </tr> <tr> <td style="text-align: center;">74</td><td style="text-align: center;">16</td><td style="text-align: center;">75</td><td style="text-align: center;">10</td><td style="text-align: center;">76</td><td style="text-align: center;">77</td><td style="text-align: center;">78</td><td colspan="4"></td> </tr> <tr> <td colspan="11"><i>badminton tournament with a straight-game win over South Korea's An Se Young.</i></td> </tr> </table>		1	2	3	4	5	6	7	8	9	10	11	<i>Legendary Australia captain Steve Waugh believes it is only a matter of time before India win a big</i>											12	13	14	15	16	17	18	19				<i>International Cricket Council (ICC) tournament. The 'Men in Blue' haven't won any major trophy since their</i>											20	19	21	22	5	23	24	10	16			<i>Champions Trophy victory in 2013 in England. Waugh stated it is not easy to win such tournaments because of</i>											25	26	27	28	28	9						<i>the level of competition but he also suggested that it is not a mental block because of which India keep on</i>											29	30	31	15	16							<i>crashing in knockout stages of ICC tournaments.</i>											32	33	34	35	36	37	38	39	40	41		<i>Manchester United manager Ole Gunnar Solskjaer has labeled his side's first-half performance in</i>											42	43	44	45	46	47	48	49	50	51	52	53	<i>Tuesday's League Cup semi-final, first-leg defeat their worst display of the season. Despite starting the game</i>											54	55	56	57	58	33	21	59	60	61		<i>without a recognized striker, holders City outclassed United in a 3-1 victory at Old Trafford, scoring all three</i>											62	63	39	64								<i>goals in a 21-minute first-half blitz.</i>											64	65	66	67	68	69	70	71	72	73		<i>London Olympics bronze-medallist Saina storms into women's singles quarterfinals of the Malaysia Masters</i>											74	16	75	10	76	77	78					<i>badminton tournament with a straight-game win over South Korea's An Se Young.</i>										
1	2	3	4	5	6	7	8	9	10	11																																																																																																																																																																																																																																										
<i>Legendary Australia captain Steve Waugh believes it is only a matter of time before India win a big</i>																																																																																																																																																																																																																																																				
12	13	14	15	16	17	18	19																																																																																																																																																																																																																																													
<i>International Cricket Council (ICC) tournament. The 'Men in Blue' haven't won any major trophy since their</i>																																																																																																																																																																																																																																																				
20	19	21	22	5	23	24	10	16																																																																																																																																																																																																																																												
<i>Champions Trophy victory in 2013 in England. Waugh stated it is not easy to win such tournaments because of</i>																																																																																																																																																																																																																																																				
25	26	27	28	28	9																																																																																																																																																																																																																																															
<i>the level of competition but he also suggested that it is not a mental block because of which India keep on</i>																																																																																																																																																																																																																																																				
29	30	31	15	16																																																																																																																																																																																																																																																
<i>crashing in knockout stages of ICC tournaments.</i>																																																																																																																																																																																																																																																				
32	33	34	35	36	37	38	39	40	41																																																																																																																																																																																																																																											
<i>Manchester United manager Ole Gunnar Solskjaer has labeled his side's first-half performance in</i>																																																																																																																																																																																																																																																				
42	43	44	45	46	47	48	49	50	51	52	53																																																																																																																																																																																																																																									
<i>Tuesday's League Cup semi-final, first-leg defeat their worst display of the season. Despite starting the game</i>																																																																																																																																																																																																																																																				
54	55	56	57	58	33	21	59	60	61																																																																																																																																																																																																																																											
<i>without a recognized striker, holders City outclassed United in a 3-1 victory at Old Trafford, scoring all three</i>																																																																																																																																																																																																																																																				
62	63	39	64																																																																																																																																																																																																																																																	
<i>goals in a 21-minute first-half blitz.</i>																																																																																																																																																																																																																																																				
64	65	66	67	68	69	70	71	72	73																																																																																																																																																																																																																																											
<i>London Olympics bronze-medallist Saina storms into women's singles quarterfinals of the Malaysia Masters</i>																																																																																																																																																																																																																																																				
74	16	75	10	76	77	78																																																																																																																																																																																																																																														
<i>badminton tournament with a straight-game win over South Korea's An Se Young.</i>																																																																																																																																																																																																																																																				

FIGURE 3. Sample text for skeletoning

2.1.2. Algorithm: Skeletoning

---

**Algorithm 1:** Skeleton creation of a text hypergraph

---

**Data:** Hypergraph

**Result:** Skeleton

Create a text hypergraph  $H_\tau$ ;

$i = 1$ ;

**repeat**

Create sub-hypergraph  $X_i$  of  $H_\tau$ ;

Apply the dilation  $\delta^e(X_i^n)$ ;

Find the skeleton  $S(H_i) = H_\tau - \delta^e(X_i^n)$ ;

$H_\tau = S(H_i)$ ;

$i = i + 1$ ;

**until**  $X_i = \phi$  or  $i = k$ ;

---

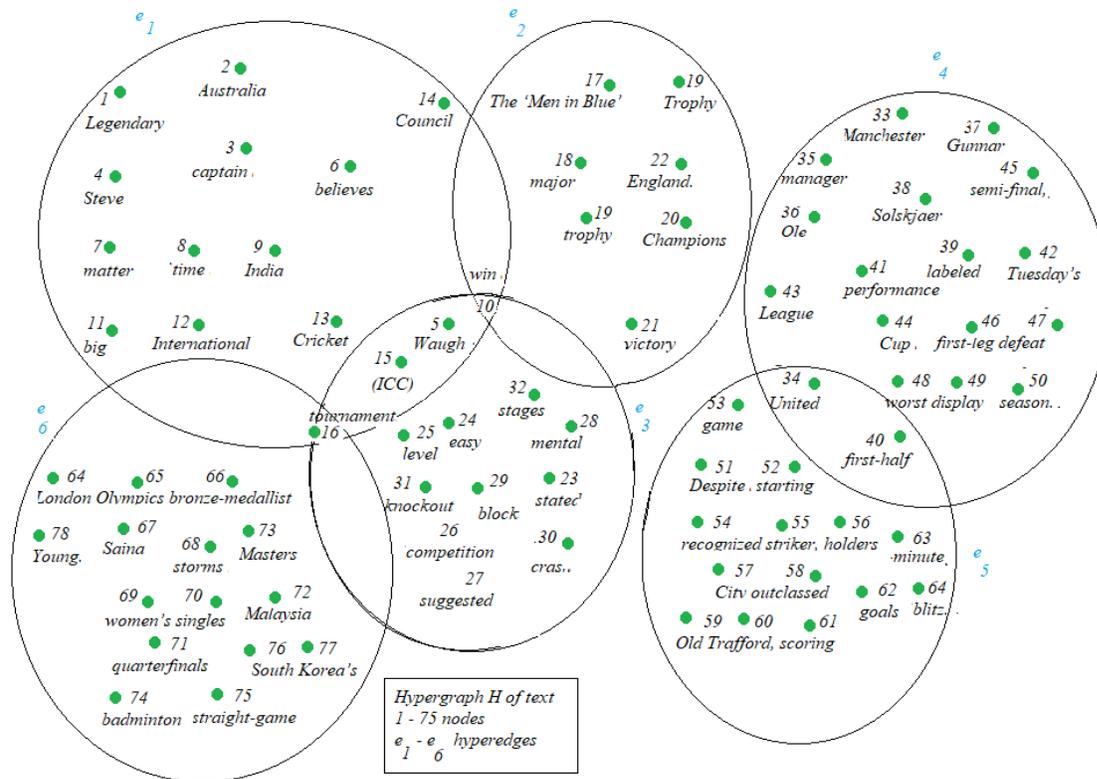


FIGURE 4. Hypergraph formed from text in Figure 3

2.2. Skeleton operation with dilation related to node

Dilation related to node which is written as  $\delta^n(X^e)$  is defined as the set of nodes in  $X^e$  of  $H$ . On applying  $H - (\delta^n(X^e))$  we get the skeleton of  $H$  w.r.to nodes. We can further apply skeleton operation by varying  $X$ .

2.2.1. Illustration

Let us take the same example given in Figure 4. Let  $X_1 = e_1, e_3$  as shown in Figure 4. Now when  $k = 1$ , the skeleton operation  $S(H_1) = H - (\delta^n(X_1)^e)$  results in Figure 5. Now let  $X_2 =$  set of sentences related to *football*. On applying skeleton operation  $S(H_2) = S(H_1) - \delta^n(X_2)^e$ , we get the graph shown in Figure 6. When  $k = 1$ , we get the maximal skeleton. When  $k$  increases the thinning nature of the skeleton increases and we get the minimal skeleton as shown in Figure 7 and Figure 8.

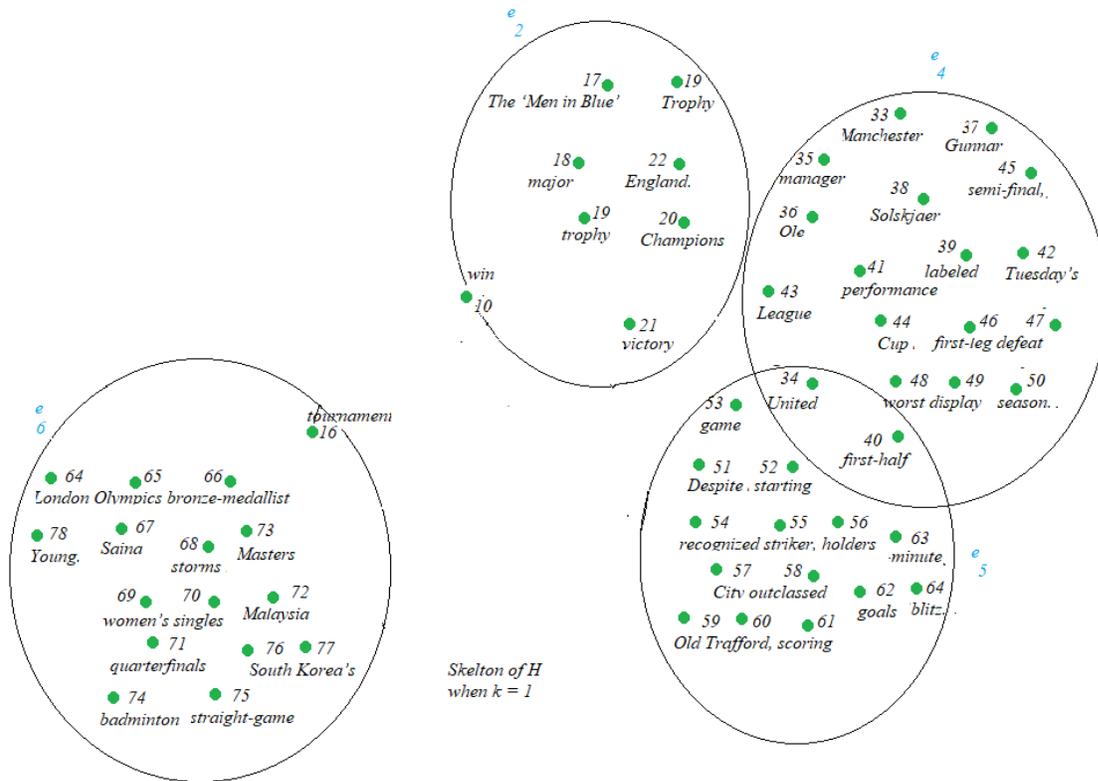


FIGURE 5. Skeleton of H w.r.to edge when k=1

Similarly, skeleton operation can be done with erosion w.r.to edge which can be defined as

$$S(H) = H - (\varepsilon^e(X^n)) \tag{15}$$

where  $\varepsilon^e(X^n)$  is defined as the collection of hyperedges containing only nodes in  $X^n$ .

Skeleton operation using erosion related to node is defined as

$$S(H) = H - (\varepsilon^n(X^e)) \tag{16}$$

where  $\varepsilon^n(X^e)$  is defined as the collection of nodes in  $X^n$ , which are only seen in  $X$  and not in the complement of  $X$ .

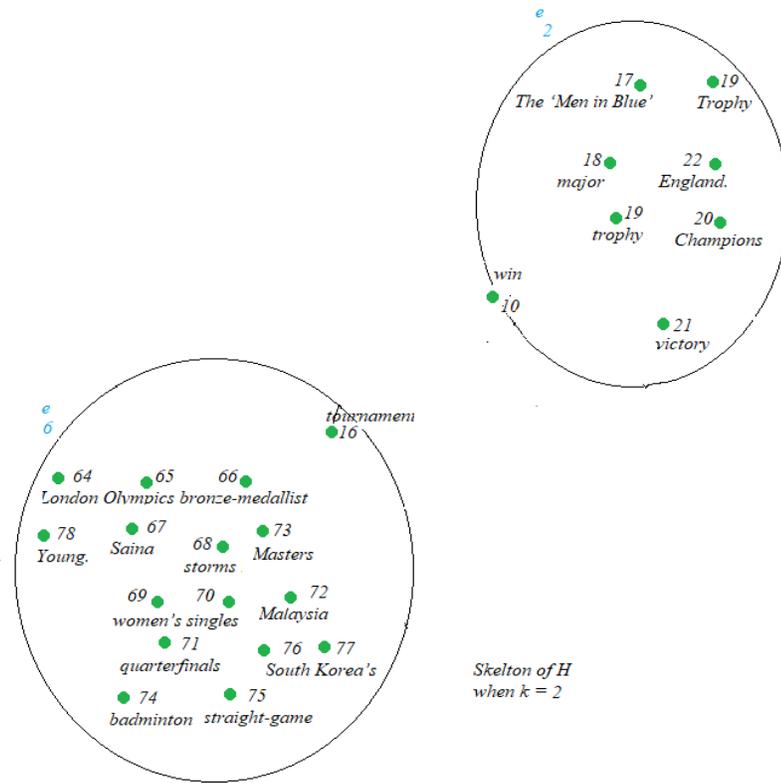


FIGURE 6. Skeleton of H w.r.to edge when  $k=2$

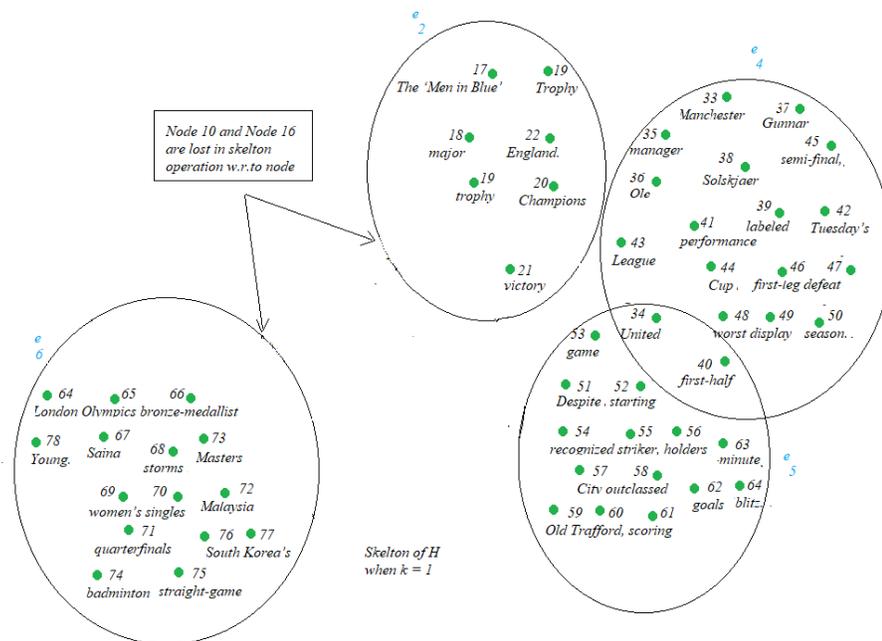


FIGURE 7. Skeleton of H w.r.to node when  $k=1$

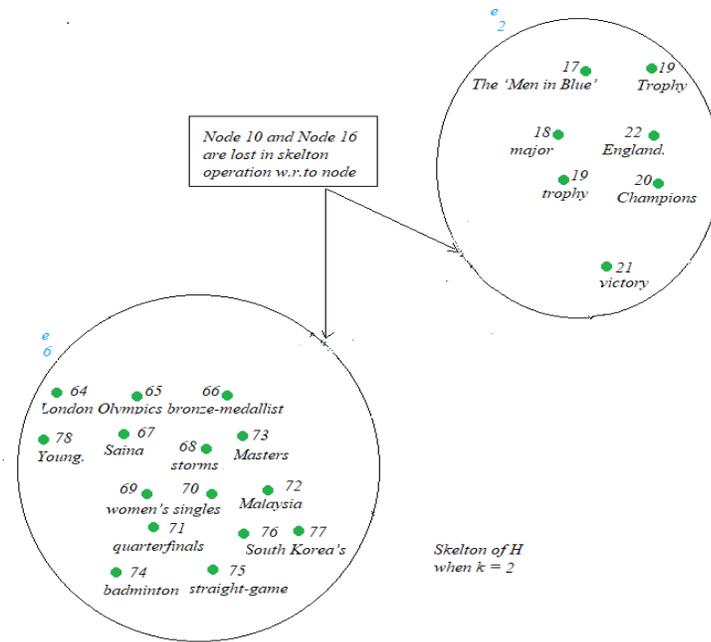


FIGURE 8. Skeleton of H w.r.to node when k=2

2.3. Hit-or-miss algorithm w.r.to dilation

Consider the text related to *Sachin Tendulkar*. He is there in the field of *cricket*, *football* and *politics*. Let us take the set of words related to *sachin* and *cricket* as  $n_{sc}$ , the set of words related to *sachin* and *football* as  $n_{sf}$  and the words related to *sachin* and *politics* as  $n_{sp}$ . Here the word with highest priority is *MP*(Member of Parliament) which comes with in  $n_{sp}$ . Now set  $A = n_{sc} \cup n_{sf} \cup n_{sp}$ . Now let us a take a window  $W$  of  $n_{sp}$  which is the neighbourhood of  $n_{sp}$  obtained as  $\delta^e(n_{sp})$ . This can be defined as the set of social service and charity activities done by *sachin* while he is an *MP*. The hypergraph for the above can be shown in the Figure 9.

Here  $A = n_{sc} \cup n_{sf} \cup n_{sp}$ . which is shown in Figure 9. Let  $X = \text{Text related to sachin while he is an MP}$ . Here *MP* is the node with the highest priority. Let it be named as  $p_{high}$ . Now  $X = n_{sp}$ . Let  $W$ , be the window of  $X$  as shown in Figure 9. The hit-or-miss operation of the hypergraph is defined as

$$HM(H) = (A \varepsilon X) \cap (A' \varepsilon (W - X)) \tag{17}$$

and the same is shown in Figure 10 and the method is shown in Algorithm 2.

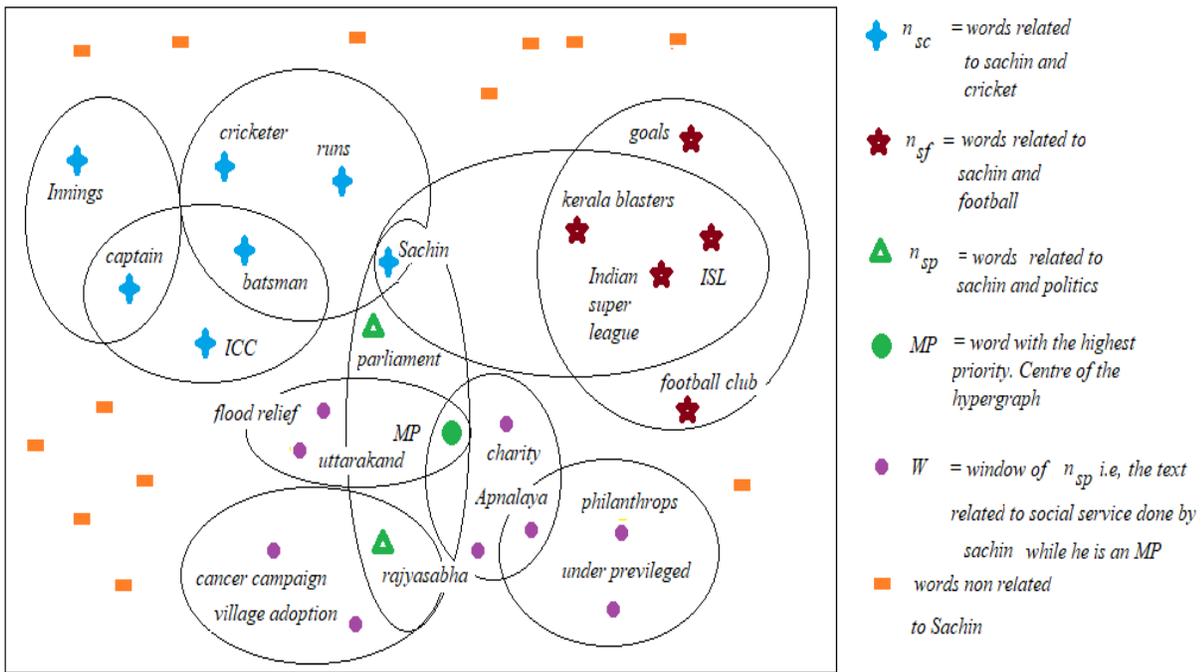


FIGURE 9. Parent Hypergraph  $H$  of text, which contains text related to *sachin*

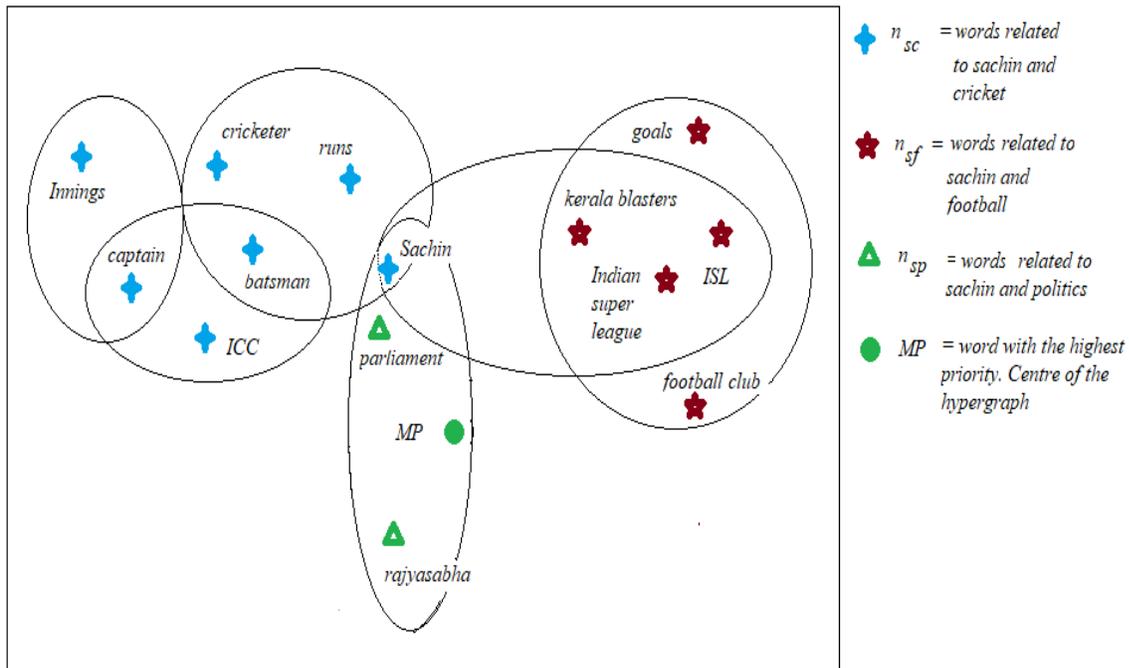


FIGURE 10. Text hypergraph  $A = n_{sc} \cup n_{sf} \cup n_{sp}$  related to *sachin*

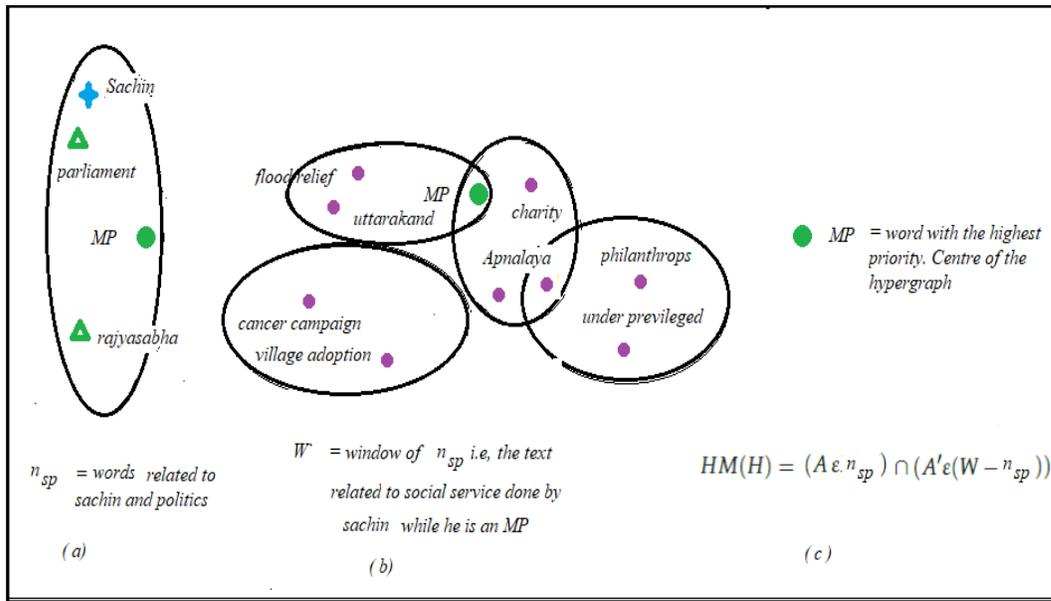


FIGURE 11. Result of hit-or-miss operation on text hypergraph

2.3.1. Hit-miss-algorithm using dilation and erosion

---

**Algorithm 2:** Hit-or-miss algorithm to find the required information node

---

**Data:** Text  $\tau$

**Result:** Hit node  $H_\tau t_n$

Create a text hypergraph  $H_\tau$  as given in Figure 4.;

$i = 1$ ;

Create sub-hypergraphs  $n_i$ , such that node  $p$  is common;

Let  $A = \cup_{i=1}^m n_i$ ;

Let  $p_{high}$  be the node which is the origin of the sub-hypergraph where the node priority  $> 0.9$ ;

**repeat**

    Find  $A \in n_i$ ;

    Calculate the neighbourhood window  $W_i = \delta^e(n_i)$ ;

    Obtain  $W_i - n_i$ ;

    Compute  $A' \in (W_i - n_i)$ ;

    Derive hit node  $H_\tau t_n = (A \in n_i) \cap (A' \in (W_i - n_i))$ ;

**until**  $i = m$  or  $H_\tau t_n = p_{high}$ ;

---

### 3. Thinning and Thickening

#### 3.1. Thinning using hypergraph operations

Thinning operation can be applied to a hypergraph  $H_\tau$  by taking sub-hypergraph  $A$  and taking a hit node  $H_\tau t_n$ . As per Figure 9, hypergraph  $H_\tau$  is the hypergraph related to *sports*, and sub-hypergraph  $A$  is the text related to *Sachin*. Dilation with respect to the edge is done for the hit node as  $\delta^e(H_\tau t_n)$ . All those edges obtained as part of this dilation are removed from the hypergraph  $A$ . The algorithm for the same is shown in Algorithm 3. The result of the thinning operation with respect to hit node  $MP$  is given in Figure 12. Hit nodes can be varied and thinning can be repeatedly done. Thinning with respect to the hit node *Sachin* is given in Figure 13.

---

#### **Algorithm 3:** Thinning algorithm on a text hypergraph

---

**Data:** Text  $\tau$  and hit nodes  $H_\tau t_k$ ; where  $k = 1$  to  $q$

**Result:** Sub-hypergraph  $T^k(H_\tau)$  after thinning

Create a text hypergraph  $H_\tau$  with the text  $\tau$  as given in Figure 4.;

$i = 1$ ;

Create sub-hypergraphs  $n_i$ , such that node  $p$  is common;

Let  $A = \cup_{i=1}^m n_i$ ;

Let  $p_{high}$  be the node which is the origin of the sub-hypergraph where the node priority  $> 0.9$ ;

**repeat**

    Find  $A(n_i) = A \varepsilon n_i$ ;

    Calculate the neighbourhood window  $W_i = \delta^e(n_i)$ ;

    Obtain  $B_i = W_i - n_i$ ;

    Compute  $A(B_i) = A' \varepsilon (W_i - n_i)$ ;

**until**  $i = m$ ;

$k = 1$ ;

$T^1(H_\tau) = H_\tau$ ;

**repeat**

$i = 1$ ;

**repeat**

        Derive hit node  $H_\tau t_k = A(n_i) \cap A(B_i)$ ;

$i = i + 1$ ;

**until**  $i = m$  or  $H_\tau t_k = p_{high}$ ;

    Derive sub-hypergraph  $T^k(H_\tau) = T^{k-1}(H_\tau) - \delta^e(H_\tau t_k)$ ;

$k = k + 1$  ;

**until**  $k = q$  or  $T^k(H_\tau) = T^{k-1}(H_\tau)$ ;

;

---

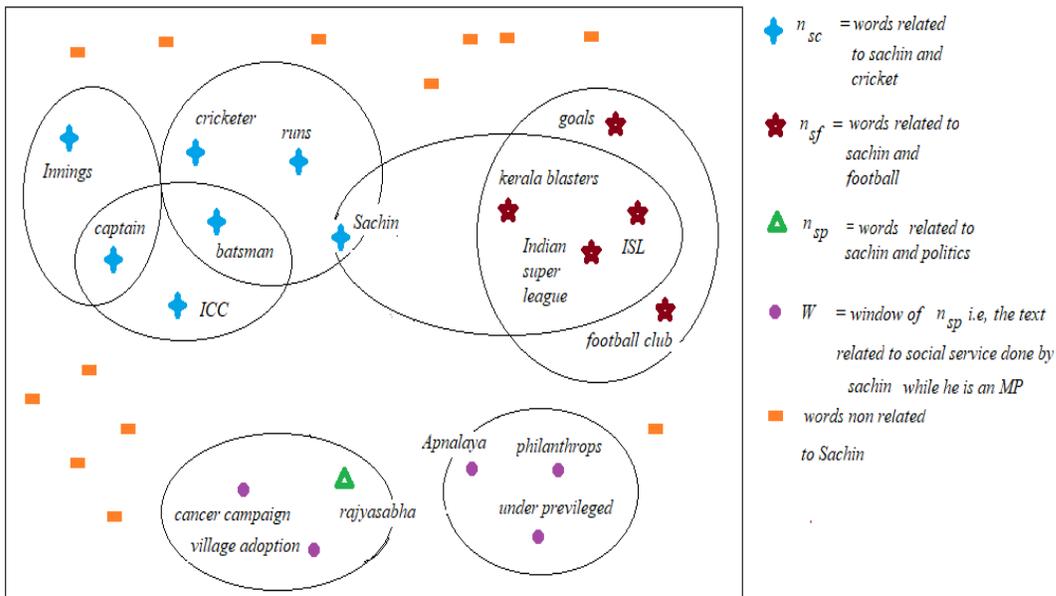


FIGURE 12. Result of Thinning operation on text hypergraph, when hit node = *MP*

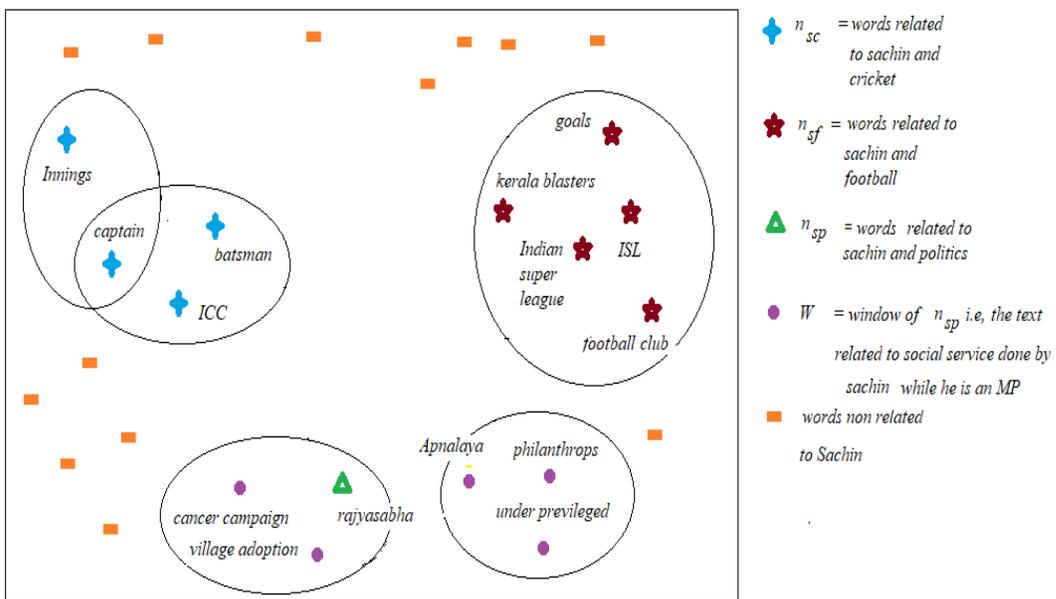


FIGURE 13. Result of Thinning operation on text hypergraph, when hit node = *Sachin*

3.2. Thickening operation of text hypergraph using dilation

Given the parent neutrosophic hypergraph  $H$ , find  $A$  which is the sub-hypergraph of more truth value. Thickening is done by taking the complement of  $A$  and applying its thinning. After thinning  $A'$ , take its complement to get a thickening of  $A$ . The result of thinning of  $A'$ ,

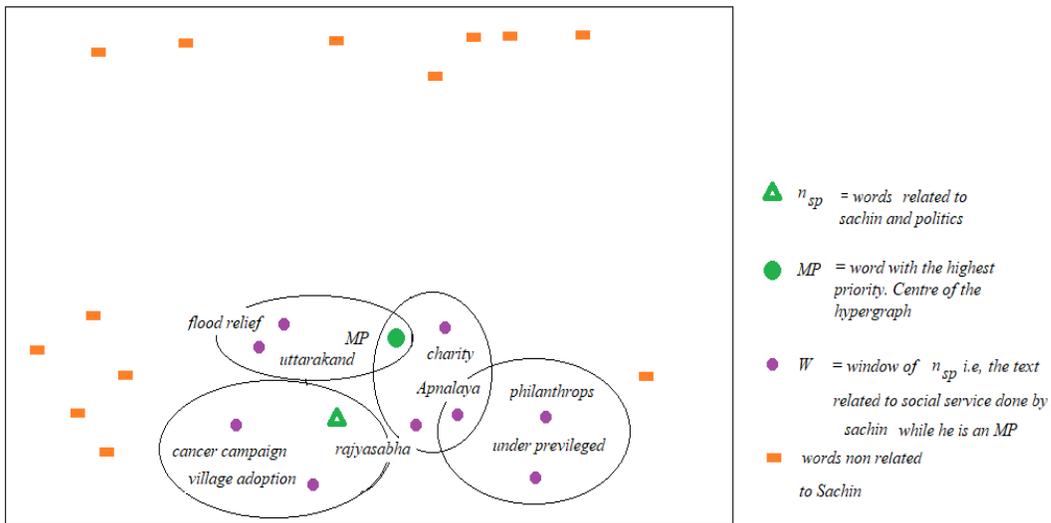


FIGURE 14. Complement of A with respect to Figure. 9.

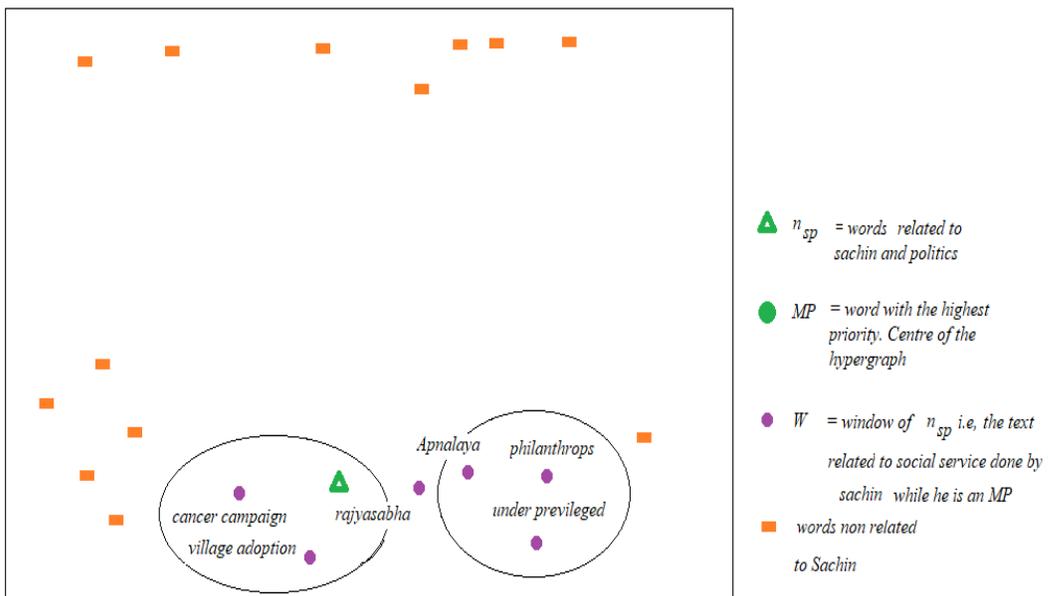


FIGURE 15. Thinning of A'

when hit node =  $MP$  is given in Figure 15. The result of thickening of  $A$ , by thinning  $A'$  and eliminating disconnected components is given in Figure 16. The algorithm for the same is shown in Algorithm 4.

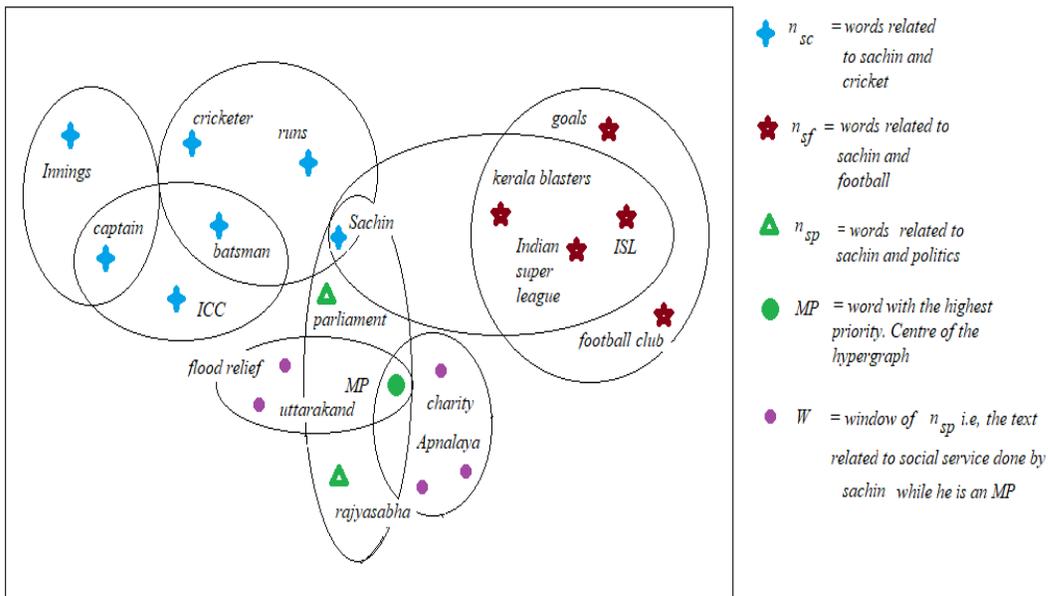


FIGURE 16. Thickening of A

### 3.3. Algebra of morphological operators

**Proposition 1:** Let  $H_1$  and  $H_2$  be the neutrosophic sub-hypergraphs, then

$$S(H_1 \cup H_2) = S(H_1) \cup S(H_2) \tag{18}$$

**Proof:** Let  $e \in S(H_1 \cup H_2)$ , i.e.,  $e \in (H_1 \cup H_2) - (\delta^e(X^n))^k$  i.e.,  $e \in (H_1 \cup H_2)$  and  $e \notin (\delta^e(X^n))^k$  i.e.,  $e \in H_1$  and  $e \notin (\delta^e(X^n))^k$  or  $e \in H_2$  and  $e \notin (\delta^e(X^n))^k$  i.e.,  $e \in (H_1 - (\delta^e(X^n))^k)$  or  $e \in (H_2 - (\delta^e(X^n))^k)$  i.e.,  $e \in S(H_1) \cup S(H_2)$ . Therefore  $S(H_1 \cup H_2) = S(H_1) \cup S(H_2)$ .

**Proposition 2:** Let  $H_1$  and  $H_2$  be the neutrosophic sub hypergraphs, then

$$S(H_1 \cap H_2) = S(H_1) \cap S(H_2) \tag{19}$$

**Proof:** Let  $e \in S(H_1 \cap H_2)$ , i.e.,  $e \in (H_1 \cap H_2) - (\delta^e(X^n))^k$  i.e.,  $e \in (H_1 \cap H_2)$  and  $e \notin (\delta^e(X^n))^k$  i.e.,  $e \in H_1$  and  $e \notin (\delta^e(X^n))^k$  also  $e \in H_2$  and  $e \notin (\delta^e(X^n))^k$  i.e.,  $e \in (H_1 - (\delta^e(X^n))^k)$  also  $e \in (H_2 - (\delta^e(X^n))^k)$  i.e.,  $e \in S(H_1) \cap S(H_2)$ . Therefore  $S(H_1 \cap H_2) = S(H_1) \cap S(H_2)$ .

**Definition:** Let neutrosophic hypergraph be  $H$ , sub hypergraph of  $H$  be  $X$ ,  $S(H)$  be the skeleton of  $H$  obtained as per e.q.(14), then a dilated skeleton  $\delta^e(S(H))$  is defined as

$$\delta^e(S(H)) = \{e/e \in N(S(H)); n/n \in e\} \tag{20}$$

where  $N(S(H))$  is the neighbourhood of  $S(H)$ .

**Proposition 3:** Let  $S(H_1)$  be the skeleton of  $H_1$ ,  $S(H_2)$  be the skeleton of  $H_2$ , where  $H_1$  and  $H_2$  be the sub hypergraphs of  $H$ , then

$$\delta^e(S(H_1) \cup S(H_2)) = \delta^e(S(H_1)) \cup \delta^e(S(H_2)) \tag{21}$$

**Algorithm 4:** Thickening algorithm on a text hypergraph

**Data:** Text  $\tau$  and hit nodes  $H_\tau t_k$ ; where  $k = 1$  to  $q$

**Result:** Sub-hypergraph  $Th^k(H_\tau)$  after thickening

Create a text hypergraph  $H_\tau$  as given in Figure 4.;

$i = 1$ ;

Create sub-hypergraphs  $n_i$ , such that node  $p$  is common;

Let  $A = \cup_{i=1}^m n_i$ ;

Let  $p_{low}$  be the node which is the origin of the  $A'$  where the node priority  $< 0.2$ ;

Create sub-hypergraphs  $x_i$  in  $A'$  where  $p_{low}$  is present.

**repeat**

Find  $A'(x_i) = A' \varepsilon x_i$ ;

Calculate the neighbourhood window  $W_i = \delta^e(x_i)$ ;

Obtain  $B_i = W_i - x_i$ ;

Compute  $A'(B_i) = A \varepsilon (W_i - x_i)$ ;

**until**  $i = m$ ;

$k = 1$ ;

$T^1(H_\tau) = H_\tau$ ;

**repeat**

$i = 1$ ;

**repeat**

Derive hit node  $H_\tau t_k = A'(x_i) \cap A'(B_i)$ ;

$i = i + 1$ ;

**until**  $i = m$  or  $H_\tau t_k = p_{low}$ ;

Derive sub-hypergraph  $T^k(H_\tau) = T^{k-1}(H_\tau) - \delta^e(H_\tau t_k)$ ;

$k = k + 1$  ;

**until**  $k = q$  or  $T^k(H_\tau) = T^{k-1}(H_\tau)$ ;

;

Find  $Th^k(H_\tau) = H_\tau - T^k(H_\tau)$

**Proof:** According to the definition of dilated skeleton,  $\delta^e(S(H_1) \cup S(H_2))$  can be written as  $\{e/e \in N(S(H_1)); n/n \in e\}$  or  $\{e/e \in N(S(H_2)); n/n \in e\}$   
 $= \delta^e(S(H_1)) \cup \delta^e(S(H_2))$

similarly we can write  $\delta^e(S(H_1) \cap S(H_2)) = \delta^e(S(H_1)) \cap \delta^e(S(H_2))$

**Proposition 4:** Let  $S(H_1)$  be the skeleton of  $H_1$  which is the sub hypergraph  $H$ ,  $S(H_2)$  be the skeleton of  $H_2$  which is sub hypergraph of  $H$ , then De morgan's law

$$(S(H_1) \cup S(H_2))^c = (S(H_1))^c \cap (S(H_2))^c \quad (22)$$

holds here

**Proof:** Let  $e \in (S(H_1) \cup S(H_2))^c$ . i.e.,  $e \notin (S(H_1) \cup S(H_2))$ . i.e.,  $e \notin (H_1 - (\delta^e(X^n))^k) \cup (H_2 - (\delta^e(X^n))^k)$ . i.e.,  $e \notin (H_1 - (\delta^e(X^n))^k)$  or  $e \notin (H_2 - (\delta^e(X^n))^k)$  i.e.,  $e \in (H_1 - (\delta^e(X^n))^k)^c$  or  $e \in (H_2 - (\delta^e(X^n))^k)^c$ . i.e.,  $e \in (S(H_1))^c$  and  $e \in (S(H_2))^c$

**Theorem 5:** Let a neutrosophic hypergraph be represented using  $H$ ,  $S(H)$  be the skeleton of  $H$ ,  $H_1, H_2$  be two sub hypergraphs of the neutrosophic hypergraph  $H$  then

$$S(S(H)) = S(H) \quad (23)$$

$$S(H_1) \cup S(H_2) = S(H_2) \cup S(H_1) \quad (24)$$

$$S(H_1) \cup (S(H_2) \cap S(H_3)) = (S(H_1) \cup S(H_2)) \cap (S(H_1) \cup S(H_2)) \quad (25)$$

E.q(18) to E.q(25) give a clear picture of the algebra of skeleton operation.

#### 4. Applications

There are many applications in the field of text analysis using the various operations discussed so far namely thinning, thickening, skeltoning, hit-or-miss operation etc. In this paper, we have applied it in identifying the hate speech in a text and removing it. The system architecture is shown in Figure 17, where the input text is subjected to preprocessing like splitting into sentences and sentences further into words. Stop words are removed from the set of words as they do not contribute to the meaning of the sentence. A neutrosophic hypergraph is constructed out of this by modeling sentences as edges and words as nodes. Lukasiewicz's fuzzy implication is applied as given in Figure 18 and Algorithm 5.

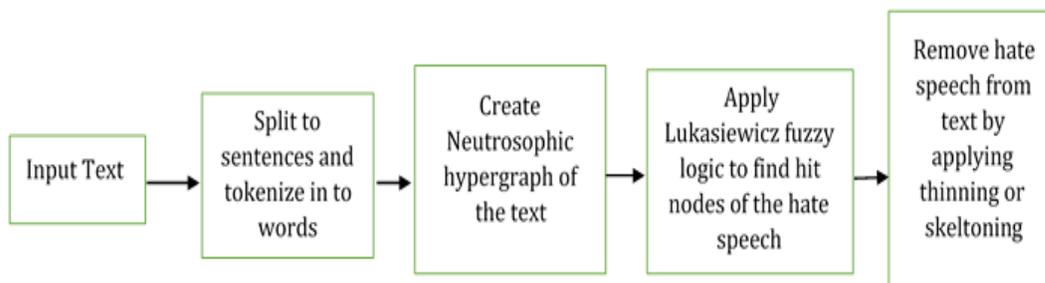


FIGURE 17. System architecture

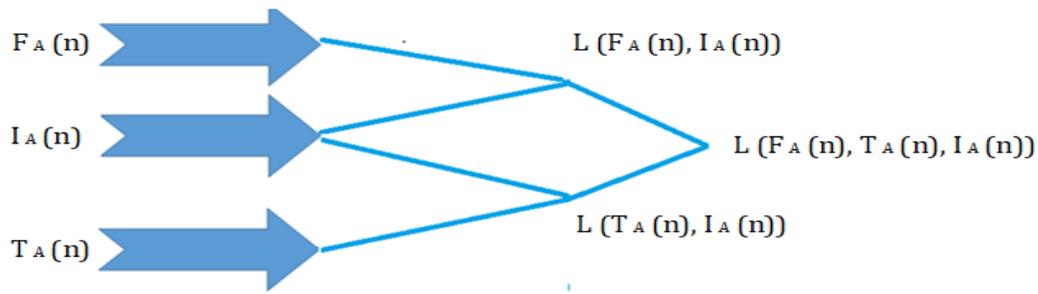


FIGURE 18. Lukasiewicz implication for hatred

**Algorithm 5:** Algorithm 5: Hate speech detection**Data:** Hate Speech Detection Method using fuzzy neutrosophic hypergraph**Result:** Tweets devoid of hate speech

1. Nodes and edges of the hypergraph represent the words and sentences of the document. Weights are assigned as given in section 1.2.  $T_A(n)$  denotes hatred measure of a word,  $I_A(n)$  denotes the uncertainty and  $F_A(n)$  shows the perfectness measure;
2. Now Lukasiewicz's implication is applied to these measures as shown in Figure 17.

- Case 1:  $[0,0,1]$  – Not a hate word. The Lukasiewicz implication would be

$$f \implies (F_A(n), I_A(n)) = \min\{1, 1 - F_A(n), I_A(n)\} = \min\{1, 1 - 1 + 0\} = 0$$

$$f \implies (T_A(n), I_A(n)) = \min\{1, 1 - T_A(n), I_A(n)\} = \min\{1, 1 - 0 + 0\} = 1$$

$$f \implies (F_A(n), T_A(n), I_A(n)) =$$

$$\min\{1, 1 - f \implies (T_A(n), I_A(n)) + f \implies (F_A(n), I_A(n))\} = \min\{1, 1 - 1 + 0\} = 0$$

- Case 2:  $[1,0,0]$  – Definitely, it is a hate word. The Lukasiewicz implication for this case would be  $f \implies (F_A(n), I_A(n)) = \min\{1, 1 - F_A(n), I_A(n)\} =$

$$\min\{1, 1 - 0 + 0\} = 1$$

$$f \implies (T_A(n), I_A(n)) = \min\{1, 1 - T_A(n), I_A(n)\} = \min\{1, 1 - 1 + 0\} = 0$$

$$f \implies (F_A(n), T_A(n), I_A(n)) = \min\{1, 1 - 0 + 1\} = 1$$

- Case 3:  $[1, 0.5, 0]$  – Depends on circumstances even though a Hate word.

$$f \implies (F_A(n), I_A(n)) = \min\{1, 1 - F_A(n), I_A(n)\} = \min\{1, 1 - 0 + 1\} = 1$$

$$f \implies (T_A(n), I_A(n)) = \min\{1, 1 - T_A(n), I_A(n)\} = \min\{1, 1 - 0.5 + 1\} = 1$$

$$f \implies (F_A(n), T_A(n), I_A(n)) = \min\{1, 1 - 1 + 1\} = 1$$

- Case 4:  $[0.5, 1, 0]$  – High indeterminacy, can be a hate word.

$$f \implies (F_A(n), I_A(n)) = \min\{1, 1 - F_A(n), I_A(n)\} = \min\{1, 1 - 0 + 0.5\} = 1$$

$$f \implies (T_A(n), I_A(n)) = \min\{1, 1 - T_A(n), I_A(n)\} = \min\{1, 1 - 1 + 0.5\} = 0.5$$

$$f \implies (F_A(n), T_A(n), I_A(n)) = \min\{1, 1 - 0.5 + 1\} = 1$$

- Case 5:  $[0, 0.5, 1]$  – Depends on circumstances even though a non-hate word.

$$f \implies (F_A(n), I_A(n)) = \min\{1, 1 - F_A(n), I_A(n)\} = \min\{1, 1 - 1 + 0.5\} = 0.5$$

$$f \implies (T_A(n), I_A(n)) = \min\{1, 1 - T_A(n), I_A(n)\} = \min\{1, 1 - 0 + 0.5\} = 1$$

$$f \implies (F_A(n), T_A(n), I_A(n)) = \min\{1, 1 - 1 + 0.5\} = 0.5$$

**Algorithm 6:** Hate speech detection .....continuation of Algorithm 5

3. Assign for each edge  $e$  in  $H^e$ ,  $T_A(e)[0, 1]$ ,  $I_A(e)[0, 1]$ ,  $F_A(e)[0, 1]$  and

$$T_A(e) + I_A(e) + F_A(e) \leq 3;$$

4.  $T_A(e)$  is as per e.q(1) and  $I_A(e), F_A(e)$  is given by

$$I_A(e) = \text{avg}(I_A(n)); \forall n \in e \quad (26)$$

$$F_A(e) = \text{avg}(F_A(n)); \forall n \in e \quad (27)$$

5. Create sub-hypergraph  $X$  by applying  $(\alpha, \beta, \gamma)$  cut such that  $T_A(e) \geq 0.5$ ,

$$I_A(e) \geq 0.3 \text{ and } F_A(e) \geq 0;$$

6. Create a sub-hypergraph  $A$  by applying higher level  $(\alpha, \beta, \gamma)$  cut such that

$$T_A(e) \geq 0.8, I_A(e) \geq 0.3, F_A(e) \geq 0;$$

7. Apply the morphological operations on  $H$  with  $X$

- $\delta^e(X^n)$  – dilation of  $X$ , pertaining edges. This takes all words in  $X$  and fetches all sentences that contain minimum of one such word.
- $\delta^n(X^e)$  – dilation of  $X$  pertaining to nodes. This operation takes all sentences in  $X$  and retrieves all words in those sentences.
- $\varepsilon^e(X^n)$  – is an erosion of  $X$  pertaining to edges. This operation takes all words in  $X$  and retrieves all sentences that contain  $X^n$  only.
- $\varepsilon^n(X^e)$  – is erosion of  $X$  pertaining to nodes. This retrieves all words in  $X$  and not in  $X'$ .

8. Hate speech can be removed in two ways as follows:-

- Apply skeleton operation  $S(H) = H - (\delta^e(X^n))^k$ . Here hate speech is eliminated from the tweets.
- Implement Thinning
  - Obtain  $Hit - or - Miss(H, A) = A \varepsilon X \cap A' \varepsilon (W - X)$  where  $X$  is dilated to get  $W$ . This operation generates intense hate words.
  - Obtain  $H - \delta^e(Hit - or - Miss(H, A))$

9. The sentences obtained after step 8 give tweets without hate speech.

A variation of this method without Lukasiewicz implication is seen in [28]. As per the system architecture shown in Figure 17, Twitter tweets are collected using Twitter APIs, text cleaning is done to remove irrelevant information such as URLs, emojis, hashtags, and punctuation marks. After preprocessing, tokenization is applied to split into words and stop word removal is done. Once the words are separated and stop words are removed, as mentioned in the above algorithm, words are given three membership values namely indeterminacy, truth and falsity. Sentences are also assigned with these three membership values. The truth value of a sentence will be the maximum truth value of the words in it. The indeterminacy value of the sentence will be the average of the indeterminacy values of all the words in it. The

falsity value of the sentence will be the average of the falsity values of all the words in it. Once a neutrosophic hypergraph( $H$ ) is created with these three values for the edges and nodes an alpha, beta, and gamma cut is applied to it to create a sub-hypergraph( $X$ ) which retrieves the sentences which are more likely to have hate speech. Morphological operations namely erosion and dilation are applied with this  $X$  on  $H$  which gives various query results as mentioned in the algorithm. Applying dilation  $k$  times with  $X$  and subtracting it from  $H$  will result in a skeleton of tweets devoid of hate speech. Hit-or-miss operation is also applied which results in retrieval of most hate words. Applying dilation of these words and subtracting it from  $H$  gives thinning.

## 5. Result Analysis

The system is implemented using Python. The data set used in this system is Twitter data(tweets) from which the hate tweets are identified and removed. Results are analyzed using various measures namely:-

- $t_p$  = true positives = Number of tweets which actually consist of hatred words and are classified as hate tweets.
- $t_n$  = true negatives = Number of tweets that do not consist of hatred words and are classified as non-hate tweets.
- $f_p$  = false positives = Number of tweets which are actually non-hate tweets but classified as hate tweets.
- $f_n$  = false negatives = Number of tweets which are actually hate tweets but classified as non-hate tweets.

Further, using the above values we calculate the measures like recall, miss rate, false positive rate, true negative rate, false omission rate, positive predictive value, negative likelihood ratio, negative predictive value, positive likelihood ratio, false discovery rate, accuracy etc. According to our proposed system, recall or sensitivity is the ratio of hate sentences identified by the system to the total hate sentences in the input data set. Our system has shown a better value of 0.87. Precision or specificity is the ratio of non-hate sentences identified by the system to the total number of non-hate sentences in the data set, where our system reported 89% results. The false positive ratio is the ratio between the number of non-hate sentences wrongly identified as hate sentences and the count of non-hate sentences. The system showed a pretty less false positive rate of 0.11. Positive Predictive Value (92%) shows how many are hate out of hate sentences identified by the system. Similarly, other values are also calculated and tabulated in Table 3. Data set 1 is of size 500, Data set 2 is of size 1000 and Data set 3 is of size 5000.

TABLE 3. Result Analysis of the proposed system

Parameter	Data set 1	Data set 2	Data set 3
$t_p$	0.83	0.95	0.87
$t_n$	0.88	0.75	0.92
$f_p$	0.13	0.09	0.03
$f_n$	0.15	0.11	0.42
Recall	0.85	0.89	0.87
Precision	0.87	0.89	0.92
Miss rate(FNR)	0.153	0.103	0.132
False Positive Rate	0.129	0.107	0.083
True Negative Rate	0.871	0.893	0.92
Positive Predictive Value	0.865	0.913	0.97
False Omission Rate	0.145	0.128	0.313
LR+	6.59	8.09	10.8
LR-	0.024	0.124	0.054
Accuracy	0.86	0.898	0.88
False Discovery Rate	0.135	0.098	0.029
Negative Predictive Value	0.854	0.87	0.69

## 6. Conclusions

In this work, we have done a detailed study of various neutrosophic morphological operators like hit-or-miss, thickening, thinning, skeleton etc. This a novel method of representing text as a neutrosophic hypergraph and Illustration of these operators on it. Also, their algorithms are implemented with text as input. As an application of the proposed work, we have applied it to hate speech detection in Twitter tweets and got an accuracy of 88%. It is observed that various compositions of neutrosophic morphological operators may give various results of text analysis. Such a study is very useful for categorizing text with respect to key information provided to the system. This is a novel method for extracting relevant information from text or a document. It is possible to extend the work by analyzing various properties of neutrosophic hypergraphs. Neutrosophic logic has a very important part in the construction of inference systems where connectors like Sheffors and Pierce's connectors may be useful. Since optimality is a major concern in every problem, constructing operators that satisfy various optimality conditions is a future work. Such new operators can be used for the comparison of various data sets and multi-classification of extracted information. This work can be extended to the area of proper fertilizer applications in the area of agriculture, team selection in sports, educational admission systems, and pandemic spread detection and isolation of people.

**Acknowledgements:** We acknowledge RSET, Kakkanad, India for providing us with the computer labs for implementing this work in Python.

**Conflicts of Interest:** There is no conflict in opinion or interest among the authors in publishing the work.

## References

1. S.Khan, Vineeth Kumar S, "HCovBi-Caps: Hate Speech Detection Using Convolutional and Bi-Directional Gated Recurrent Unit With Capsule Network", IEEE Access, Volume 10, pp. 7881-7894, 2022, DOI: 10.1109/ACCESS.2022.3143799
2. Rodriguez A., Yi-Ling Chen, Carlos A, "FADOHS: Framework for Detection and Integration of Unstructured Data of Hate Speech on Facebook Using Sentiment and Emotion Analysis", IEEE Access, Volume 10, pp. 22400-22419, 2022, DOI: 10.1109/ACCESS.2022.3151098
3. Punya Joy Saha, Midhun Das, Binny Mathew, "Hate Speech: Detection, Mitigation and Beyond", WSDM '23: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, pp. 1232-1235, February 2023, DOI: <https://doi.org/10.1145/3539597.3572721>
4. Rajavikram, Mahesh N, "Deep learning based fusion strategies for hate speech detection to combine the classifiers to improve classification performance", AIP Conf. Proc. 2754, 2023, DOI: <https://doi.org/10.1063/5.0161653>
5. Akileng Issac, Raju Kumar, Aruna Bhat, "Hate Speech Detection Using Machine Learning Techniques", Lecture Notes in Electrical Engineering book series (LNEE, volume 840), 31 March 2022, DOI: [https://doi.org/10.1007/978-981-16-9012-9\\_11](https://doi.org/10.1007/978-981-16-9012-9_11)
6. Haibin Wang, "Single Valued Neutrosophic Sets", Technical Sciences and Applied Mathematics Vol.10, 2012, pp. 1-5, 2012, DOI: <https://www.academia.edu/6762666>
7. Allan Breitto, "Hypergraph Theory: An Introduction", Springer, pp. 1-20, 2013, DOI: <https://doi.org/10.1007/978-3-319-00080-0>
8. Mordeson J.N., Nair P.S., "Fuzzy Graphs and Fuzzy Hypergraphs", Studies in Fuzziness and Soft Computing, Vol 46. Physica, Heidelberg, 2000, DOI: <https://doi.org/10.1007/978-3-7908-1854-3>
9. Muhammad Akram, Anam Luqman, "Fuzzy Hypergraphs and Related Extensions", Studies in Fuzziness and Soft Computing, Vol 390, 2020, DOI: <https://doi.org/10.1007/978-981-15-2403-5>
10. H. J. A. M. Heijmanns, P. Nacken, A. Toet, L.Vincent, "Graph Morphology", Journal of Visual Communication and Image Representation, Vol 3(1), pp. 24-38, 1992, DOI: [https://doi.org/10.1016/1047-3203\(92\)90028-R](https://doi.org/10.1016/1047-3203(92)90028-R)
11. Bino Sebastian V, Kannan Balakrishnan, A, Unnikrishnan, Ramkumar P. B., "Morphological filtering on hypergraphs", Discrete Applied Mathematics, Vol. 216, pp. 307-320, 2017, DOI: <https://doi.org/10.1016/j.dam.2015.02.008>
12. A. Unnikrishnan, Kannan Balakrishnan and P. B. Ramkumar, "Mathematical Morphology on Hypergraphs Using Vertex-Hyperedge Correspondence", Hindawi publications, 2014, DOI: 10.1155/2014/436419
13. Dhanya P.M, Sreekumar A., Jathavedan M., Ramkumar P.B, "On Constructing Morphological Erosion of Intuitionistic Fuzzy Hypergraphs", Journal of Analysis, Springer, pp. 583-603, 2018, DOI: <https://doi.org/10.1007/s41478-018-0096-3>
14. Dhanya P.M., Sreekumar A., Jathavedan M., Ramkumar P.B., "Algebra of Morphological Dilation on Intuitionistic Fuzzy Hypergraphs", IJSRSET, Vol. 4(1), pp. 300- 308, 2018.

---

Dhanya P.M, Ramkumar P.B, Text Analysis Using Morphological operations on a Neutrosophic Text hypergraph

15. Dhanya P.M., Sreekumar A., Jathavedan M., Ramkumar P.B., "Metric Induced Morphological Operators on Intuitionistic Fuzzy Hypergraphs", *International Journal of Mathematics and Mathematical Sciences*, Hindawi publications, pp. 1-12, 2018, DOI: <https://doi.org/10.1155/2018/6045358>
16. Dhanya P.M., Sreekumar A., Jathavedan M., Ramkumar P.B., "Text Summarization Using Morphological Filtering of Intuitionistic Fuzzy Hypergraphs", *Journal of Computer Science*, Science publications, Vol 14(6), pp. 837-853, 2018, DOI: <https://doi.org/10.3844/jcssp.2018.837.853>
17. Dhanya P.M., Ramkumar P.B., Nirmaljith Cletus, Paul Joy, "Fuzzy hypergraph modeling, analysis and prediction of crimes", *International Journal of Computing and Digital Systems*, Vol 11(1), pp. 648-661, 2022, DOI:10.12785/ijcds/110152
18. Florentin Smarandache, B. Vasantha Kandasamy, W. B. Vasantha Kandasamy, Ilanthenral Kandasamy, "Neutrosophic Graphs: A New Dimension to Graph Theory", 2015, DOI: 10.6084/M9.FIGSHARE.1574172
19. Muhammad Akram, Sidra Sayed and Florentin Smarandache, "Neutrosophic Incidence Graphs With Application", *Axioms* Vol. 7(3), 2018, DOI:10.3390/axioms7030047
20. Florentin Smarandache, "Neutrosophic Logic - Generalization of the Intuitionistic Fuzzy Logic", *Proceedings of the Third Conference of the European Society for Fuzzy Logic and Technology*, University of Applied Sciences, Zittau/Görlitz, Germany, EUSFLAT, pp. 141-146, 10-12 September 2003, DOI: <https://doi.org/10.48550/arXiv.math/0303009>
21. Naeem Jana, Lemnaouar Zedamb, Tahir Mahmoodc, KifayatUllahd, Said Broumie, Florentin Smarandache, "Constant single valued neutrosophic graphs with applications", *Neutrosophic Sets and Systems*, Vol. 24, pp. 77-89, 2019, DOI:<https://fs.unm.edu/NSS2/index.php/111/article/view/636>
22. Zenat Mohamed, Mahmoud M.Ismail, Amal F.Abd El-Gawad, "Analysis Impact of Intrinsic and Extrinsic Motivation on Job Satisfaction in Logistics Service Sector: An Intelligent Neutrosophic Model", *Neutrosophic Systems with Applications*, vol.4, pp.43-52, 2023, DOI: <https://doi.org/10.5281/zenodo.8202133>
23. Ahmed Sleem, Ibrahim Elhenawy, "Energy Efficiency and Material Cost Savings by Evolution of Solar Panels Used in Photovoltaic Systems under Neutrosophic Model", *Neutrosophic Systems With Applications*, vol.5, pp. 36-43, 2023, DOI: <https://doi.org/10.5281/zenodo.8208871>
24. Samah Ibrahim Abdel Aal, "Neutrosophic Framework for Assessment Challenges in Smart Sustainable Cities based on IoT to Better Manage Energy Resources and Decrease the Urban Environment's Ecological Impact", *Neutrosophic Systems With Applications*, vol.6, pp. 9-16, 2023, DOI: <https://doi.org/10.5281/zenodo.8210126>
25. D. Sasikala, B. Divya, "A Newfangled Interpretation on Fermatean Neutrosophic Dombi Fuzzy Graphs", *Neutrosophic Systems With Applications*, vol.7, pp. 36-53, 2023, DOI: <https://doi.org/10.5281/zenodo.8218317>
26. Mohamed Abdel-Basset ,Abduallah Gamal,Ibrahim M. Hezam ,and Karam M. Sallam, "An Effective Analysis of Risk Assessment and Mitigation Strategies of Photovoltaic Power Plants Based on Real Data: Strategies, Challenges, Perspectives, and Sustainability", *International Journal of Energy Research*, Volume 2023, DOI: <https://doi.org/10.1155/2023/1582795>
27. Mohamed Abdel-Basset ,Abduallah Gamal, Karam M. Sallam, Ibrahim M. Hezam ,and Ahmad M. Alshamrani, "Sustainable Flue Gas Treatment System Assessment for Iron and Steel Sector: Spherical Fuzzy MCDM-Based Innovative Multistage Approach", *International Journal of Energy research*, Volume 2023, DOI: <https://doi.org/10.1155/2023/6645065>
28. Anagha A., Anugraha A., Antony J., Binil Tom, Dhanya P.M, "Hate speech detection in Twitter using different models", *ITM Web Conf.*, Vol. 56, pp. 1-6, 2023, DOI: <https://doi.org/10.1051/itmconf/20235604007>

Received: Aug 13, 2023. Accepted: Dec. 16, 2023