\*\* if (T == (ADJ||V ||ADV ||N N )) then Keyword

\*\* if (Pattern of

(ADV +NN) || (ADV + V )|| (ADJ +V ) || (ADJ +NN) || (NN + ADV ) || (NN +ADJ) ||

(ADV + ADJ + NN) || (V + ADJ + NN) || (V + ADV + ADJ) || (V + NN + NN) || (NN + V + NN) ||(NN + NN + V)

is Matched) then Phrase

1.

भारत दौरे पर तुर्की के राष्ट्रपति NSG पर समर्थन हासिल करने की होगी कोशिश

['NNP', 'NN', 'PSP', 'NN', 'PSP', 'NN', 'PSP', 'PSP', 'NN', 'NN', 'VM', 'PSP', 'VM', 'NN']

keyword/ senti words: 'दौरे', 'तुर्की' , 'राष्ट्रपति', 'समर्थन', 'हासिल', 'करने', 'होगी'

phrase: समर्थन हासिल करने(0.0)

news: भारत-तुर्की व्यापार इवेंट में मोदी ने कहा- आपसी सहयोग बढ़ाने की प्रबल संभावना

['DEM', 'NN', 'PSP', 'PSP', 'NN', 'PSP', 'NN', 'PRP', 'NN', 'NN', 'PSP', 'JJ', 'NN']

Phrase: प्रबल संभावना(0.0)

twitter sentiment=0

twitter phrase=0

news phrase=0

**Not Sarcastic**

\*\* if (Pattern of

(ADV +NN) || (ADV + V )|| (ADJ +V ) || (ADJ +NN) || (NN + ADV ) || (NN +ADJ) ||

(ADV + ADJ + NN) || (V + ADJ + NN) || (V + ADV + ADJ) || (V + NN + NN) || (NN + V + NN) ||(NN + NN + V)

**is Matched) then Phrase**

2.

बिहार में दलितों पर लाठीचार्ज.. मायावती उर्फ देवी तो ये खबर सुनेगी ही नहीं..

['NNP', 'PSP', 'NN', 'PSP', 'DEM', 'NN', 'XC', 'NN', 'CC', 'DEM', 'NN', 'PSP', 'RP', 'NN']

keyword/ senti words: 'दलितों', 'मायावती' , 'देवी', 'खबर', 'नहीं..'

phrase: no phrase(0.0)

news: **मायावती को सता रही बिहार की चिंता, कहा-यहां दलितों की स्थिति खराब**

['NN', 'PSP', 'VM', 'VAUX', 'NNP', 'PSP', 'NN', 'PSP', 'NN', 'PSP', 'NN', 'JJ']

Phrase: **स्थिति खराब**

twitter sentiment=0

twitter phrase=0

news phrase=-0.375

**Sarcastic**

\*\* if (Pattern of

(ADV +NN) || (ADV + V )|| (ADJ +V ) || (ADJ +NN) || (NN + ADV ) || (NN +ADJ) ||

(ADV + ADJ + NN) || (V + ADJ + NN) || (V + ADV + ADJ) || (V + NN + NN) || (NN + V + NN) ||(NN + NN + V)

**is Matched) then Phrase**

3.

अब तो दलितों की पिटाई बिहार में भी हो गयी मायावती को पहले वहा जाना चाहिए वो नजदीक है गुजरात से

['PRP', 'CC', 'NN', 'PSP', 'NN', 'NNP', 'PSP', 'RP', 'VM', 'VAUX', 'NN', 'PSP', 'NST', 'VM', 'VAUX', 'VAUX', 'PRP', 'NN', 'PSP', 'NN', 'PSP']

keyword/ senti words: 'दलितों', 'पिटाई' , 'हो', 'मायावती', 'वहा', 'नजदीक', 'गुजरात'

phrase: no phrase(0.0)

news: **मायावती को सता रही बिहार की चिंता, कहा-यहां दलितों की स्थिति खराब**

['NN', 'PSP', 'VM', 'VAUX', 'NNP', 'PSP', 'NN', 'PSP', 'NN', 'PSP', 'NN', 'JJ']

Phrase: **स्थिति खराब**

twitter sentiment=-0.625

twitter phrase=0

news phrase=-0.375

**Sarcastic**

**short comings:**

1. streaming only on the basis of language is not possible

i.e, languages='hi ' will give 406 OR 420 error

2.sentiwordnet may not have the written unicode hindi word: In that case we are not going to get the sntiment value of the word

3. need to remove user, trending

4. I think we need to remove preposition between two nouns

i.e, NN +PSP+NN -------> NN+NN

and also

NN +PSP+VM -------> NN+VM

it will increase the phrase frequency

**Analysis and proposal of shortcoming 4**

\*\*

if

1.NN +PSP+NN -------> NN+NN

2.NN +PSP+VM -------> NN+VM

and similar

#phrase finding algo

**if (Pattern of**

(ADV +NN) || (ADV + V )|| (ADJ +V ) || (ADJ +NN) || (NN + ADV ) || (NN +ADJ) || (NN +VM)

(ADV + ADJ + NN) || (V + ADJ + NN) || (V + ADV + ADJ) || (V + NN + NN) || (NN + V + NN) ||(NN + NN + V)

**is Matched) then Phrase**

3.

अब तो दलितों की पिटाई बिहार में भी हो गयी मायावती को पहले वहा जाना चाहिए वो नजदीक है गुजरात से

['PRP', 'CC', 'NN', 'PSP', 'NN', 'NNP', 'PSP', 'RP', 'VM', 'VAUX', 'NN', 'PSP', 'NST', 'VM', 'VAUX', 'VAUX', 'PRP', 'NN', 'PSP', 'NN', 'PSP']

keyword/ senti words: 'दलितों', 'पिटाई' , 'हो', 'मायावती', 'वहा', 'नजदीक', 'गुजरात'

phrase: पिटाई बिहार में भी हो गयी , गयी मायावती को पहले, वहा जाना चाहिए वो नजदीक है गुजरात

news: **मायावती को सता रही बिहार की चिंता, कहा-यहां दलितों की स्थिति खराब**

['NN', 'PSP', 'VM', 'VAUX', 'NNP', 'PSP', 'NN', 'PSP', 'NN', 'PSP', 'NN', 'JJ']

Phrase: **स्थिति खराब** , **मायावती को सता, रही बिहार की चिंता**

twitter sentiment=-0.625

twitter phrase=-0.75

news phrase=-0.375

**Not Sarcastic**

**what I propose:**

1. Adding the rules

1.1.NN +(PSP/RP)+NN -------> NN+NN

1.2.NN +(PSP/RP)+VM -------> NN+VM

and similar

2.Adding NN + VM as phrase detector