

Preprocessing:

- Lowercase
- Stemming
- Lemmatization
- Custom stopwords
- Tokenization
- Removed special character,emojis,text other than english,numbers

BERT Model

- Installed pre Trained Model
- Clustering Algo: HDBSCAN
- Dimensionality reduction algo: UMAP
- N_grams:1

Topics

- **Siddham:** bapatla sidham, chandragiri sidham,Raptatu Siddam is a super success
- **Next CM Jagan:** next, confirm, best
- **MLA Confirm :**Buchepalli Sivanna Darshi Mula Confirm,CM Malli confirmed by G Jagan
- **Asking Help:** muppavaram kondapi mandal prakasam district water supply problem.
- Also Middle Class & Higher People Not Getting Any Benefits From YCP
- **Water Bottle:** Intelligence check on water bottles.,Water Bottles Must Print PH value,
- Take Care Water Bottles to Avoid Cancer, Diabetes.
- Maintain Drinking pH value must be 6.8 with 6.9.
- **Vote:** Balireddy Pratap Reddy,M.L.A. Contestant of Chandragiri and Punganur Constituency,Pulicherla,Tirupati,
- **Government:** Good Government,Due To One Of Government Employee Rest Of The Family Members Are Not Getting Government Schemes Rest Members No Jobs
- **People Praising Jagan :** God,hero,superstar,Jesus,Brother,Lord
- **Salary/Pension:** Raise Volunteer Salary,increase pension in your new manifesto

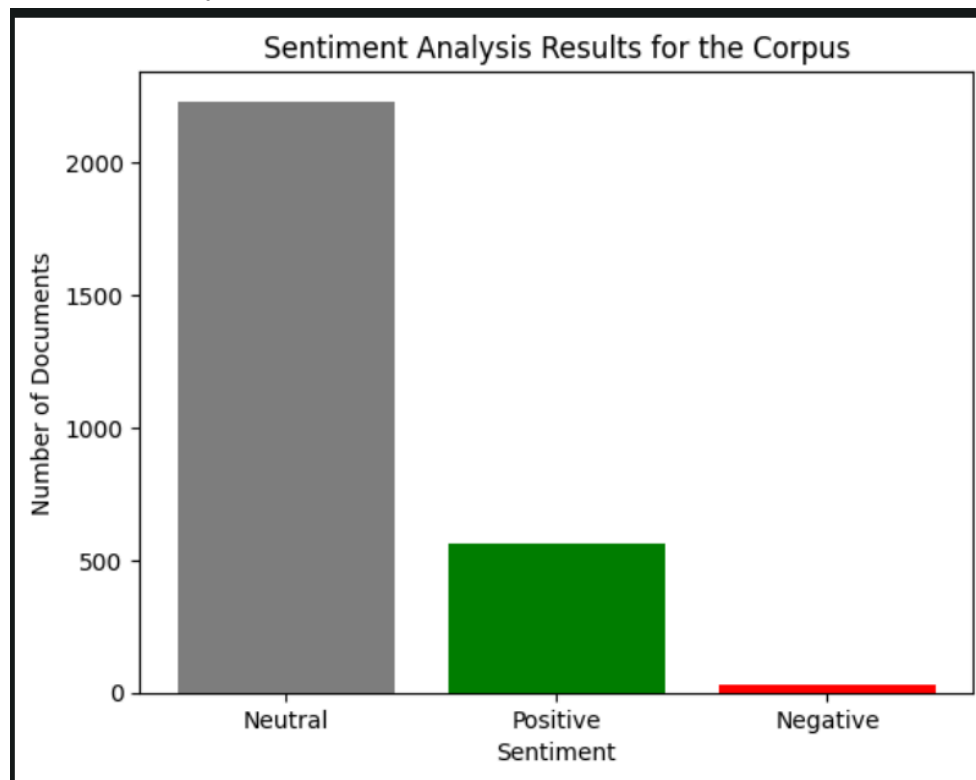
BERT Model:

- KMeans
- PCA
- Word Cloud

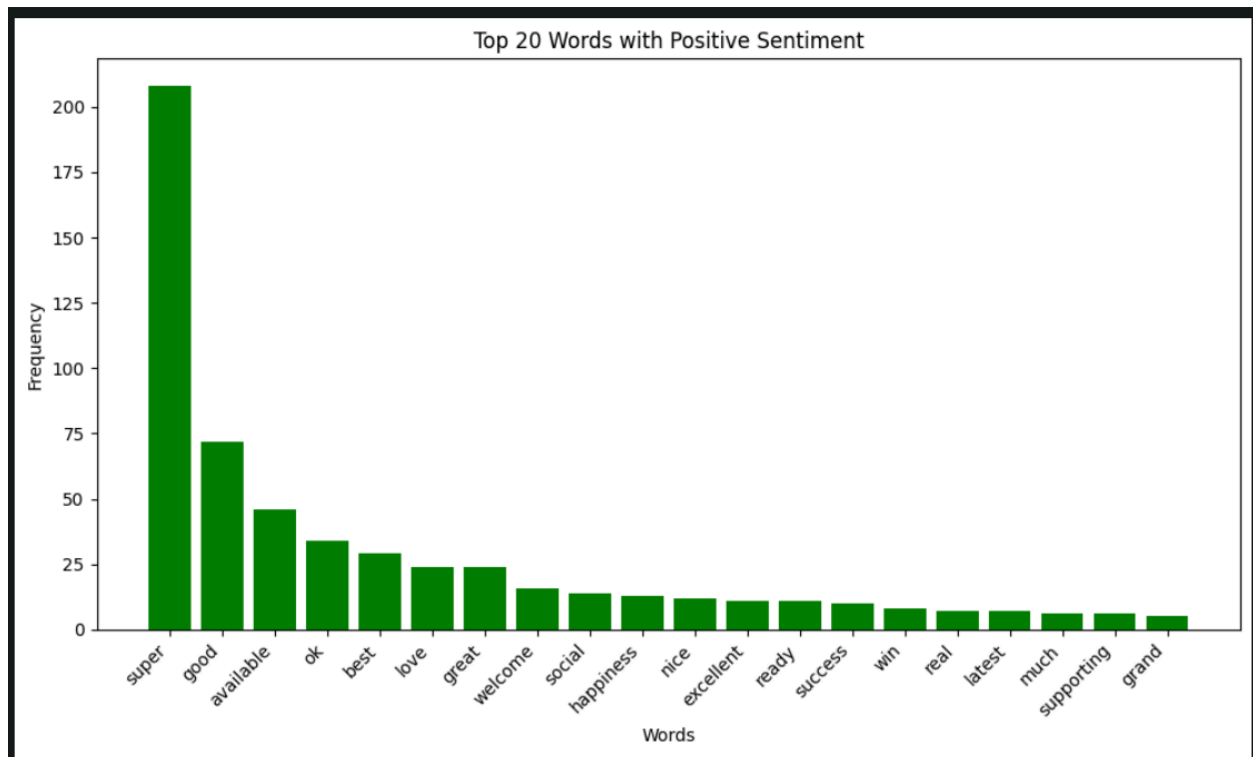
[illegible][illegible]

ramalingareddy
har
army
convenor
menu
p
ga
website
army
rapthadu
convenor
facebook
ap
naa
home

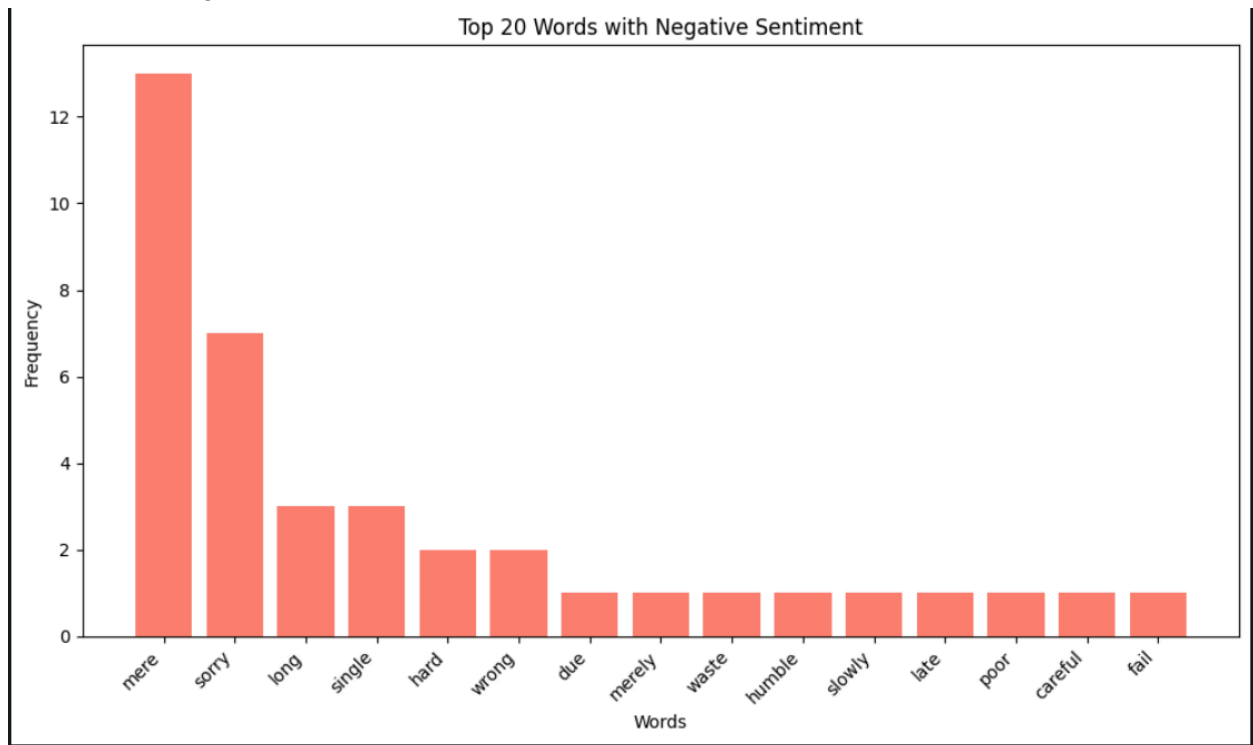
Sentiment Analysis: Used Textblob



Words With positive Sentiment



Words With Negative Sentiment



Sentence using word sorry:

Sorry jagan sahab anantapur somnath nagar 47 division no voltage