



VIGNAN INSTITUTE OF TECHNOLOGY AND SCIENCE

Near Ramoji Film City, Deshmukhi (V), Yadadri Bhuvanagiri Dist, Telangana - 508 284.

Affiliated to JNTU Hyderabad, Programmes Accredited by NBA, Approved by AICTE, New Delhi



CLASSIFICATION FOR DETECTING INSULTING & ABUSIVE CONTENT

PROJECT GUIDE:

Mr. B. Srinu

Assistant Professor

TEAM MEMBERS:

M.Swetha **16891A0532**

K.Mahesh **16891A0529**

K.Dheeraj **16891A0522**



Contents

- Introduction
- Abstract
- Existing system
- Disadvantages
- Proposed system
- Advantages
- Pre processing & Modeling
- Conclusion
- Future Enhancement
- References



Introduction:

- The sheer simplicity with which abusive and insulting comments can be made on the web
- Utilizing the present computerized correspondence innovations (particularly web based life), is liable for their noteworthy increment and worldwide universality.
- Natural Language Processing advancements can help in tending to the negative impacts of this improvement



Abstract:

- In this project, we assess a lot of classification calculations on two sorts of client created online substance in two English.
- Arrangements of information we deal with aspects like prejudice, sexism, hatespeech etc.
- Main aim is arranging the information as per the commented on qualities utilizing a few Text Classification Algorithms
- Finally detecting which one is more suitable for the problem.



Existing system:

- Classification for Detecting insulting and Abusive Content has been implemented using several algorithms in many different technologies.
- Although, there is no clarity regarding which one to be used in what scenarios. Hence, a comparative analysis needs to be done.
- They've built a model based on traditional machine learning approach (Logistic regression fine tuned model).



Disadvantages:

- Many different algorithms has been implemented with no clarity on which one to use in what type of scenarios.
- Older systems usually need huge amount of training data.



Proposed system:

- In our proposed system, Python based Deep learning algorithms are being implemented for classifying the Abusive and insulting content.
- Further, these algorithms will be compared altogether to understand the most suitable algorithm for the task.
- And also to deploy this Tkinter application.



Advantages:

- Usage of Machine Learning Algorithm makes the system more reliable and accurate.
- Higher accuracy can be used even with less amount of training data

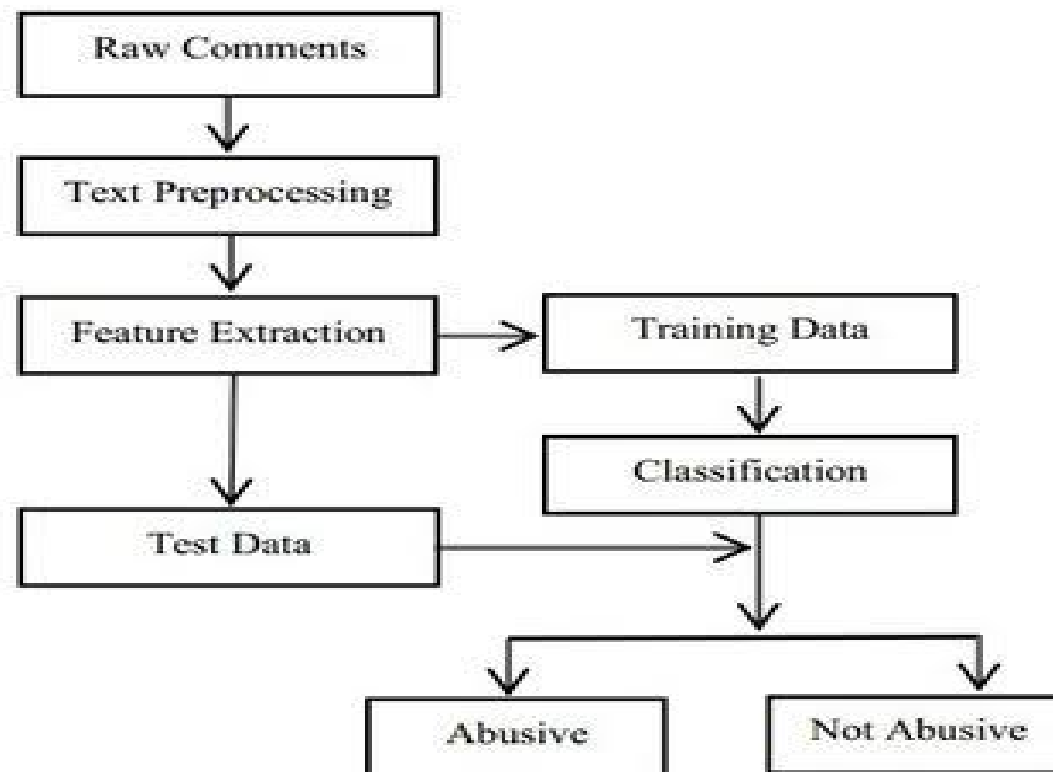


Modules:

- Natural Language Processing
- Comparison of traditional ML algorithms & Deep learning algorithms.
- Deploying through web application.



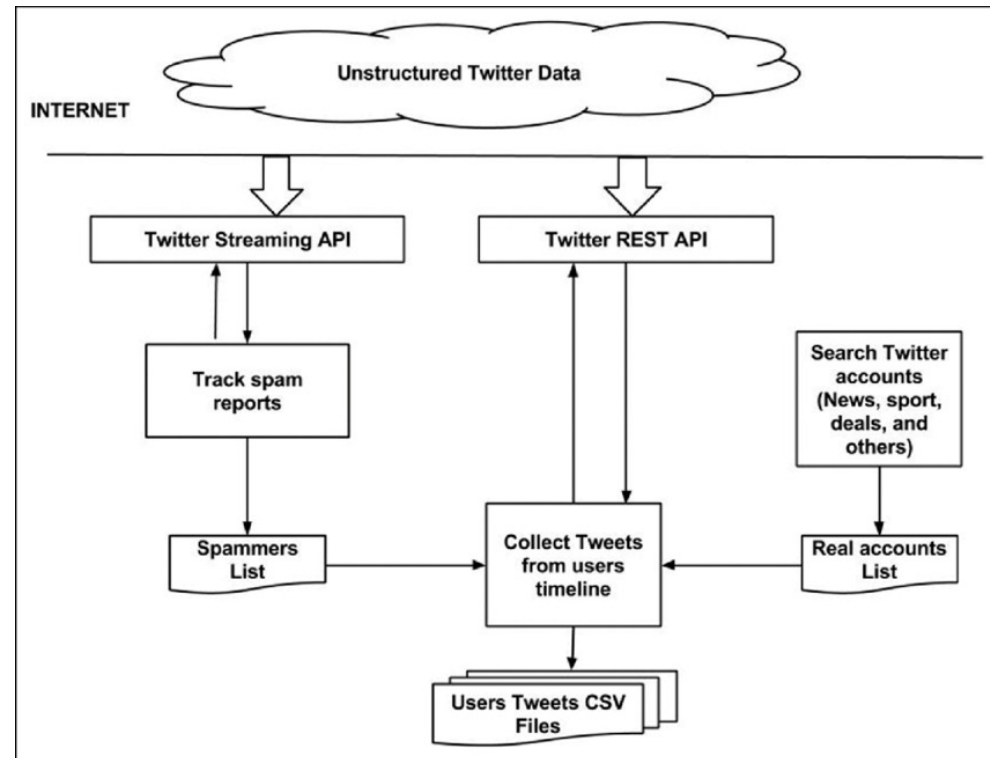
ARCHITECTURE





Data Collection :

- Collected bad words from kaggle
- Collected bad & abusive tweets for these words from twitter , using twitter developer API credentials.





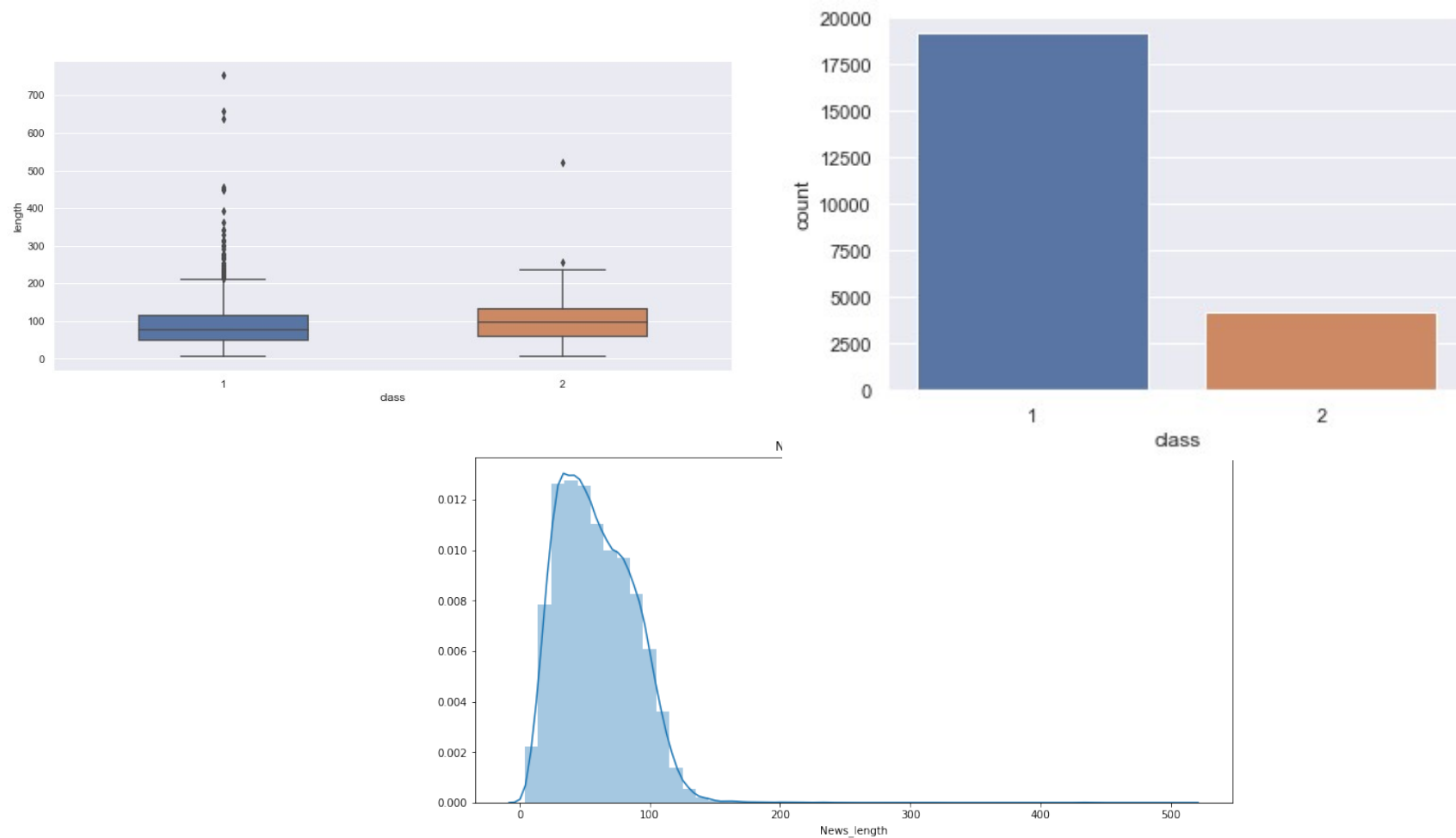
Data Cleaning and Pre processing:

Step is to create features from our data so we can train the machine learning models. The steps followed are:

- cleaning of extra spaces or lines or symbols.
- lower casing
- punctuation signs
- possessive pronouns
- stop words removal and lemmatization.



EXPLORATORY DATA ANALYSIS :





TEXT REPRESENTATION :

- We'll use and form vectors, and divide them according to the given train & test data. that represents the relative importance of a term in the document and the ent stands for Term Frequency, and IDF stands for Inverse Document Frequent.
- We will be forming columns for each word using tf idf vectorizer, so that it would be easier for model to classify.

Train-test split:

- Splitting data into 80% for train 20 % test.
- To test the models on unseen data.

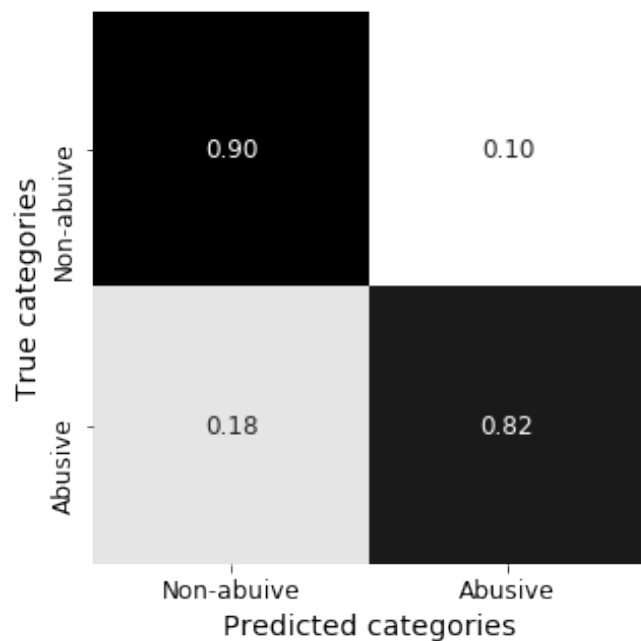


MODELING :

- We have tested several machine learning models to figure out which one may fit better to the data and properly capture the relationships across the points and their labels. For each model, we have given accuracy scores & confusion matrix ,algorithms used are Fine tuned Logistic regression,KNN , Decision Tree and Random Forest classifiers models were built.
- And then we have also implemented neural network i.e.Multilayer perceptron.



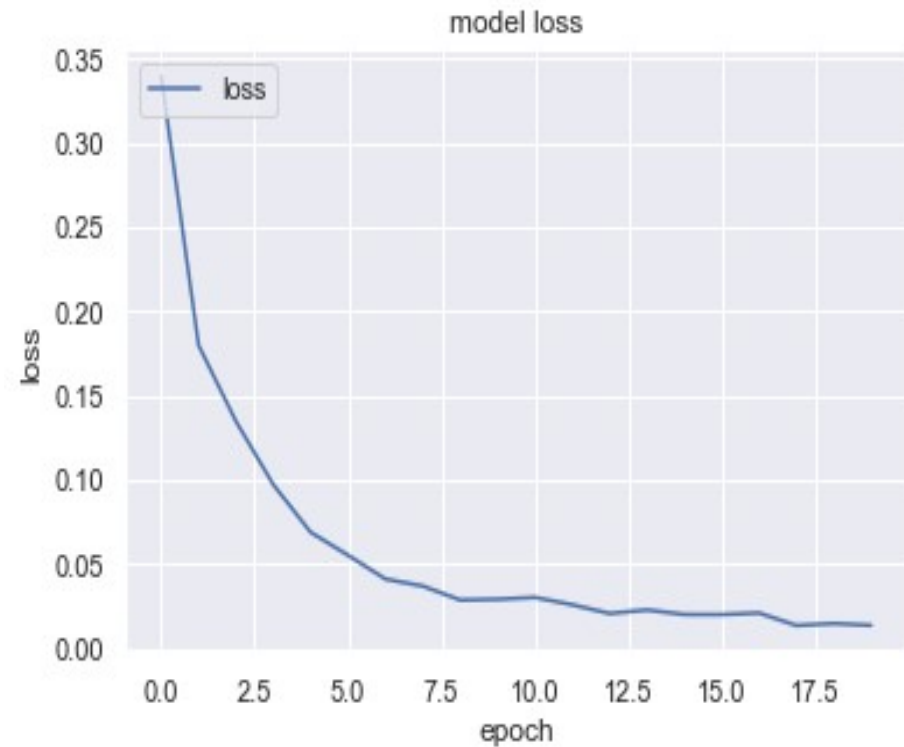
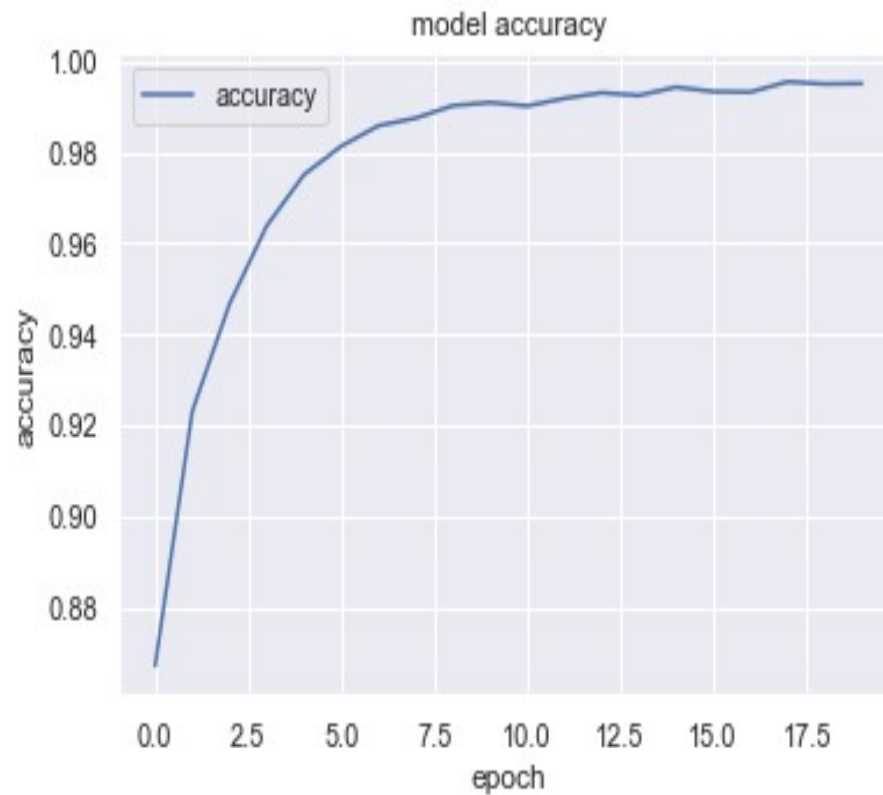
Accuracy Levels & Confuison matrix :



```
-----ACCURACY SCORE-----  
=====  
KNN: 0.827911  
=====  
Decision Tree: 0.854880  
=====  
RandomForestClassifier: 0.880137  
=====  
LR Fine tuned: 0.885702  
=====
```

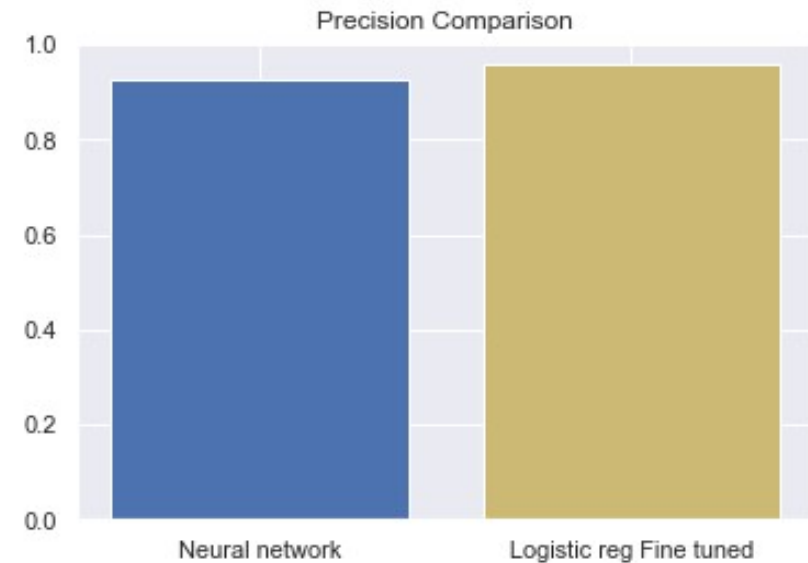
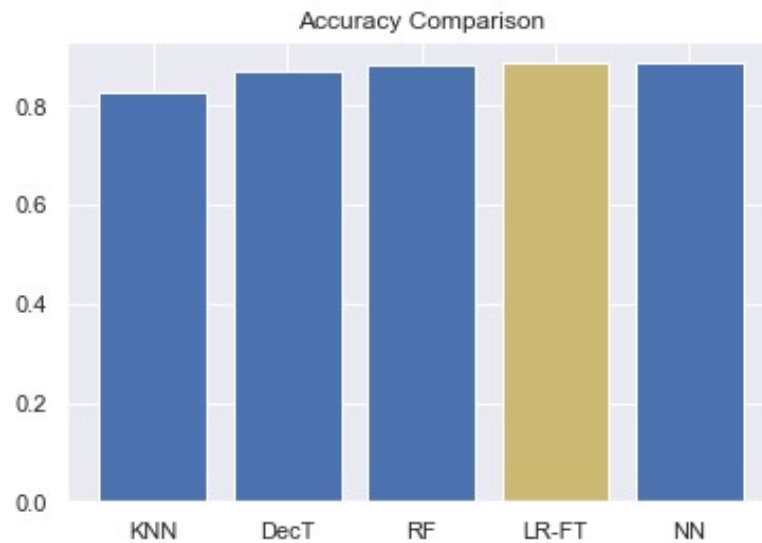


Neural network Results:





Compariosn of traditional ML vs Neural networks:





RESULTS: (Detecting non abusive content)

Detecting abusive content

Detecting abusive content

Enter text

Hello what's up , mate ?

Non Abusive

Detect

HATE SPEECH



RESULTS:(Detecting abusive)

Detecting abusive content

Detecting abusive content



Enter text

I'll kill you , stupid.@@@!!!

Abusive !!

Detect



Conclusion

- .Natural language processing helps us to get trending innovations in improvement of impact of recent innovation.
- In this project, we assess a lot of classification calculations on two sorts of client created online substance in two English.
- The Aim of this project is on arranging the information as per the commented on qualities utilizing a few Text Classification Algorithms finally detecting which one is more suitable for the problem.



FUTURE ENHANCEMENT

- The same technology should be included in video to get audio abusive or in proper content to be removed from audio or video or it may get skip.
- Depending on the abusing content we can retrieve the person information from database to take necessary action.
- Deep learning algorithms can be used for further details and highly accurate analysis.



REFERENCES:

- Agrawal, S., & Awekar, A. (2018). Deep learning for detecting cyberbullying across multiple social media platforms. In ECIR:Advances in Information Retrieval (pp. 141–153).
https://doi.org/10.1007/978-3-319-76941-7_11
- Allen, C. (2011). Islamophobia. Surrey: Ashgate. Amichai-hamburger,Y., & McKenna, K. (2006). The Contact Hypothesis Reconsidered: Interacting via the Internet. Journal of Computer-Mediated Communication, 11(1), 825–843.
<https://doi.org/10.1111/j.1083-6101.2006.00037.>



VIGNAN
Institute of Technology & Science
Deshmukhi (Vill.), Nalgonda

THANK YOU!!

Department of **COMPUTER SCIENCE & ENGINEERING**