Survey on Classification of Insincere Questions on Quora

Ms.Tanvi Patel, M.Tech Computer Engineering , BVM Engineering College, V.V.Nagar, India,
pateltanvi49@gmail.com

Dr.Udesang K Jaliya, Asst. Prof Computer Department, BVM Engineering College, V.V.Nagar,
India, udayjaliya@gmail.com

Mrs.Mahasweta Joshi, Asst. Prof Computer Department, BVM Engineering College, V.V.Nagar,
India, mjjoshi@bvmengineering.ac.in

Abstract : With the expansion Usage of internet in general and Social media play substantial role in our life where data is massively exchanged. Online Social media, community Forums and discussion platforms are scant examples of frequent playground of online discussion where people are freely connected. Due to freedom of speech gain the abusing or insincere questions. Quora is one of the platform for post online Q & A and share knowledge. Meanwhile among these users, there are bad actors misuse this freedom to post insincere question(e.g. non - neutral, false information etc.). In this paper, we use a real-world dataset posted by Quora to evaluate different mechanism and algorithms to classify insincere question.

Keywords — *Classification , Insincere, LSTM(Long Term Short Memory) , Machine Learning, Quora, SVM(Support Vector Machine)*

I. INTRODUCTION

Internet play important role in the evolution of world refinement. According to Internet word stats, Almost 4.92 billion people were active internet users as of October27, 2020, encompassing 63.2 percent of the global population[9].

There are many social media platform like yahoo, stack overflow, tweeter. People express their feelings, via social media. It is way for people to gain knowledge and share the knowledge. People are free to post and speak on social media platform. They also post information that insincere or harm full to any community and religion.

Insincere questions are major problem in quora. Quora is like any other online social network that "encourage people to learn from each other". On quora, people can ask question on any topic and get answer form expert or qualified people. Some people ask question in good faith ,and some percentage of bad people post questions that are insincere or problematic. problem that arises is when a user ask question whose purpose is not to look for answer; the question is categorized as an insincere question.

Nowadays there are no restrictive rules in set a question can be categorized as insincere question. As one the platform Quora has given some characteristics that can signify that a question is insincere such as: 1) Non- Neutral Tone 2)

Disparaging or Inflammatory 3) Is-not- Grounded in Reality 4) Sexual Content. "Non-Neutral Tone" defined as question asked by users that have a special an exaggerated tone to underscore a point about a group of people and rhetorical and meant to imply a statement about a group of people. "Disparaging or inflammatory" defined as user ask question that promote discriminatory action against group of people. "Isn't grounded" in reality defined as question that are submit- ted by users based on wrong information or contains absurd assumptions. "Sexual Content" defined as question contain sexual content such as pedophilia, incest, and bestiality the questions are asked not to find the right answers.

At the end, we used real-world dataset. Our goal and objectives are:

1. explore the part of pre-trained word embedding and text preprocessing , feature representation in detecting insincere questions from quora.

2. Investigate the performance of the different machine learning(e.g. SVM, Logistic regression, random forest, decision tree etc.)and deep learning algorithm(CNN, RNN, BiLSTM etc) in detecting insincere questions from quora.

The rest of the paper divided as following: "section 2" summery of relevant research or literature review. "Section 3" contain Quora data analysis and frameworks. "Section 4" presents Result ,and Section 5 contain Conclusion of paper.

II. LITERATURE REVIEW

Previous research related to insincere question classification was carried out by Hendri Priyambowo et al.[8] that use Decision tree, random forest and many other machine learning algorithm and compare their performance using basic feature extraction methods like Unigram, n-gram, Syntax feature like Part of speech(POS) and Word Embedding like Word2vec,Doc2vec.Still this research in early stage and also improve using deep learning and work in what kind of insincere question(miss-information, discriminative, non-neutral tone).

Deepak Kumar Jain et al.[5]propose a pipeline is called deep refinement for text classification that able to verify offensive language. Deep refinement uses capsule networks with attention mechanism and provides a system for the insincere classification in order to enhance monitoring and information quality. Deep refinement uses some state-of-the-art methods, Comparison stat that Deep refinement classification method outperformed previously used text classification methods.

Do Yeon Kim et al.[6]resolve that deep learning framework enhanced pre-trained word embedding with topical information for insincere question classification. Topical Word Embedding model and Word Embeddings utilized for learning and pre-trained large corpora respectively. Latent Dirichlet allocation (LDA) and BiLSTM used for learn the topics of all questions. Another notable observation is that the models which utilized Paragram, GloVe and Word2Vec outperformed the models which utilized FastText. experiment results showed that topic-enhanced word embeddings are able to improve insincere question classification.

Bishal Gaire et al.[3]experimented with various machine learning algorithm (Multinomial Naive Bayes, K-nearest, Logistic Regression)and this results are compare with deep learning framework (RNN,LSTM,GRU).using Word embedding methods with deep learning frame work experiments show that RNN model outperformed with 0.6915 score.

Akanksha Mishra et al.[11]worked with BiLSTM and pre-trained Glove embedding. Data characterize in 6 categories and each question for 900 training in- stance. Further ,analyzed the results with different other embedding like word2vec,Paragram.experiment obtained accuracy of 64.35 per. on test set.

Hao Mao et al.[7]worked with various machine learning models and that can be used with pre-processing methods

that remove the punctuation, tokenization, stemming, removing stop words and perform various models to identify that question sincere or insincere. Naive Bayes, Logistic Regression, Averaged Perceptron and Recurrent Neural Networks and BERT model were used for best result. The paper suggest that best models were Averaged Perceptron, bidirectional RNN with Adam optimization and modified decision boundary.

Vandan Mujadia et al.[16]construct different machine learning and neural network based models. gradient boosting, random forest and 3-nearest neighbor classifiers with majority voting. Data characterize in six categories and each question for 900 training instance.Experiment Comparison stat that Adaboost classifier with word unigram and bigram TF-IDF features performed the best among all the classifiers.

Sreyan Ghosh et al.[15]propose a deep learning model that can be classify insincere questions.CNN and BiLSTM used with capsule network.Before train model data are prepossessed ,using some extra feature like Number of unique words in a sentence ratio of uppercase in total number of letters etc .Cyclic Learning Rate and Word Embedding are used for in this model. Experiments show that F1 score on the train with 0.6820.

Chandra Prakash et al.[4]experiments with Contrast Sentiment in word using machine and deep learning method. This method experiments on tweeter dataset and experiment shows deep learning model outperforms the ML model in term of accuracy and f1 score.

Ashwin Geet D'Sa et al[2]discussed the automatic detect the toxic speech using approaches based on deep learning classifier and word embedding. They worked with binary classification and multi-class classification, for that they used feature-based approaches, FastText and word embedding and also com- pare this configuration with pre-train BERT model.

In Another paper, Mohammad A. Al Ramahi et al.[12]study the fake review and insincere question form quora and discussed the related work. they experiment with different combination of prepossessing methods. Experiments shows that logistic regression is an appropriate predictive model to identify insincere questions.

Here one article Michael [10]conclude that pre- learned embedding perform good with text element feature. They experiment BERT and word2vec embedding with neural network and Logistic regression, study shows that Neural Networks slightly better than LR.Glove2vec slightly looked better though the performance among other embedding are close.

III. DATASET

Datasets used to perform various technique of insincere question and toxic content detection in research community. Here we explore quora and tweeter datasets.

The dataset of Quora Insincere Classification competition available on Kaggle[1].The training set has 13,06,122 samples and the test set has 56,370 samples. it is high imbalanced in nature. There are binary classification [8],[5], [6],[3],[7],[15],[12],[10]which contain two label, Label 0 indicates sincere question or normal question. Label 1 indicates insincere question.

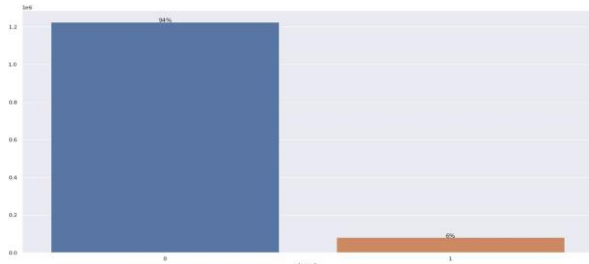Figure 1 shows the dataset representation.

Figure 1: Dataset snap

Multi-class classification [11],[16]characterized in six categories.

IV. TECHNIQUES

Machine Learning Approaches

In recent year, many machine learning methods frame as supervised classification problem. This includes manual feature engineering and then applied classic machine learning algorithms such as Support vector machine(SVM)[8],[16],[10],,Decision Tree[8],[12],Multinomial naive bayes[8],[3],[16], [7]Random Forest[8],[16],[12],Logistic regression[3],[7], [12],[4],[10],Nearest neighbor[8],[3],Multilayer Perceptron[8], Adaboost and Gradient Boost[16].

A.SVM

Support Vector Machine(SVM)[14]is supervised machine learning algorithm that can be used for both classification and regression purposes. SVMs are mostly used in classification problem. In, 1960 they first introduced but they got refined in 1990.

Figure 2: Support Vector Machine [14]

In figure 2, basically representation of different classes in hyperplane .The hyperplane will be generated iterative manner by SVM so that will minimize error. The goal of SVM to divide the datasets into the classes to find maximum marginal hyperplane.

Support Vectors are data points are the closest to the hyperplane. This points helps to separate the line.

Hyperplane is a decision plane which is divided between set of objects having different classes.

Margin defined as the gap between two data lines in the closet data points of the different classes.

First, SVM generate hyperplane iteratively that segregate class in best way. then, it will choose hyperplane that separates class correctly.

Deep Learning Approaches

Deep-learning approaches have shown to be very powerful in classifying insincere question detection[5].In the context of insincere question classification ,CNN extract the character and word from the combination[2] and BiLSTM (Bidirectional Long short term memory)learn the character dependencies in insincere question[5],[15],[2].The authors have investigated three neural network architectures including CNN and LSTM networks with Glove and FastText embeddings[6] for the detection of insincere question detection. BiLSTM and capsule network[5],[15]for Accuracy improvement.

B. Logistic Regression

Logistic Regression is one of the basic and popular algorithm to solve a classification problem. its predicts something is true or false

- Binomial Logistic Regression Used for Binary Classification: The target variable can have only 2 possible outcomes like 0 or 1 which represent Spam/Not Spam ,Dead/Alive, etc.

- Multinomial Logistic Regression Used for Multi-class Classification: The target variable can have 3 or more possible outcomes e.g. Disease A/Disease B/Disease C or Digit Classification.

For more information on Logistic regression refer to article [21].

A. LSTM/RNN Approach

Long Short Term Memory[LSTM][13] networks were introduced by Hochreiter Schmidhuber in 1997.it is special type of RNNs which is capable for learn long term dependencies.

All RNNs have chain form repeating modules. In standard RNNs, the repeating module has simple single layer

structure. However, the repeating module in LSTM more complicated. Instead of having single layer LSTM has Four layers interact with special way. it has two states :hidden state and cell state. LSTM designed for to solve long term dependency problem. To store the information for long way.

Forget Gate : The forget gate decides what information should be memorized or dump. Information from the previous hidden state and information from the current input. it is passed through the sigmoid function. The Output bounded between 0 and 1.The 0 means to forget and 1 means to memorize.

In figure 3 shows that at time t step current input x_t and the output from the previous step h_{t-1}.The weighted and input summation are taken, added a bias b_f , and this value passes through sigmoid activation function. The resulting equation (1)

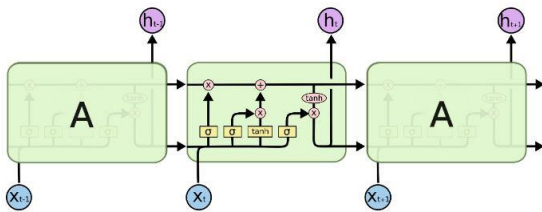$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \qquad (1)$$

Figure 3: LSTM contains four interacting layers [13]

Input Gate: LSTM decide what information to store in cell state. it regulate what information store to cell state. this has two steps. First the sigmoid layer called "input gate" as equation (2) decide which value update. second step tanh layer create new candidate value $\tilde{C}_t$, this two state added to cell state and create updated cell state.

$$i_t = \sigma (W^i x_t + U_f h_{t-1} + b_i)$$
$$\tilde{C}_t = \tanh(W_f x_t + U_f h_{t-1}) \qquad (2)$$

The old state in updated and now new state Ct .

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \qquad (3)$$

Output gate: The output gate decide amount of information inside the cell state that is exposed to external network. First information passes through sigmoid layer then tanh layer and multiply with by the output of the input layer.

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$
$$h_t = o_t \tanh C_t \qquad (4)$$

B .GRU

Gated Recurrent Units (GRU) is a mutant version of the LSTM model and was introduced by Cho et al.[17].GRU has similar architecture as the LSTM, but few parameter lacks an output. Gated recurrent unit solve vanishing gradient problem which comes with a standard recurrent neural network. GRU considered as a variation on the LSTM   because both designed similarly and some cases produce equally result.

GRU has two gates, reset gate and update gate. These are two gates decide what information passed to the output. The special thing about this network they can be keep information long ago, without washing it through time.

Update Gate: Update gate helps decide how much of the past information needs to be passed along to the future.

at time t start calculating the update gate _t using this formula:

$$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1}) \qquad (5)$$

figure 4 shows the update gate , When x_t is plugged into the network unit, it is multiplied by its own weight W(z). $h\_(t-1)$ which holds the information for the previous *t-1* units and is multiplied by its own weight U(z). Both results are added together and a sigmoid activation function is applied and result between 0 and 1.
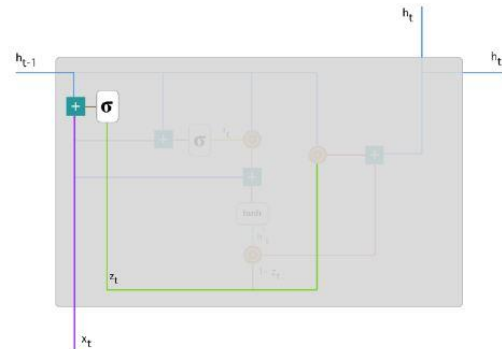

Figure 4: Update gate[18]

Reset gate: Reset gate decide the how much of the past information to forget.

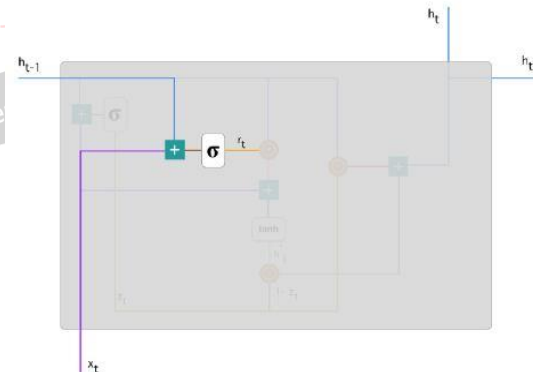$$r_t = \sigma(W^{(r)}x_t + U^{(r)}h_{t-1}) \qquad (6)$$

Figure 5 : Reset Gate[18]

equation 6 is same as Update gate. The difference comes in weight and the gate usage, which shows in figure 5 $h\_(t-1)$ as blue line and x_t as purple line, multiply them with their corresponding weights, sum the results and apply the sigmoid function.

Current memory content: Here introduction of the   new memory content which use reset gate to store the appropriate information from the past. It is calculated as follows

$$h'_t = \tanh(Wx_t + r_t \odot Uh_{t-1}) \qquad (7)$$

1. Multiply the input x_t with a weight *W* and h_(t-1) with a weight U.
2. Calculate the Hadamard (element-wise) product between the reset gate r_t and Uh_(t-1).
3. Sum up the result of step 1 and step 2.
4. Apply nonlinear activation function tanh.
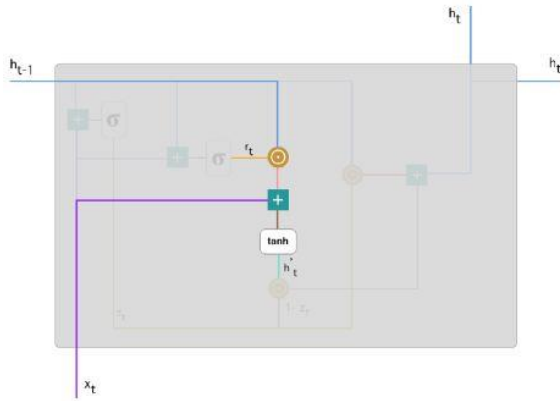
Here steps shows in figure 6,

Figure 6 : Current memory content[18]

Final memory ate current step: At the last, the network needs to value of vector h_t which holds information of current unit and passes it down network. In order to do that need update gate.

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h'_t \qquad (8)$$

1. Apply element-wise multiplication to the update gate z_t and h_(t-1).
2. Apply element-wise multiplication to (1-z_t) and h'_t.
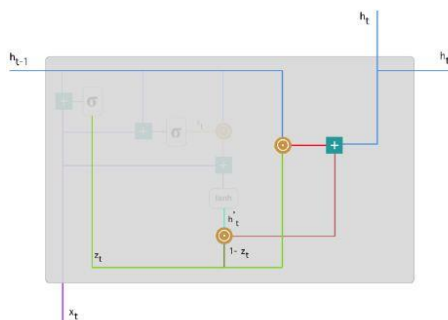3. Sum the results from step 1 and 2.

Figure 7 : Final memory at current time step[18]

Here, in figure 7 z_t used for calculate 1-z_t which, combined with h'_t as bright green line, generate a result in the dark red line. z_t used h_(t-1) as blue line in an element-wise multiplication.

Finally, h_t as blue line is a outcome of the addition of the outputs corresponding to the bright and dark red lines. for more expiation refer to blog[18].

C .Bi-LSTM

Bidirectional LSTMs are extension version of typical LSTMs that can raise performance of model on sequence classification problem. Bi-LSTMs train two LSTMs: one taking the input forward direction and other one backward direction. As shown in figure 8, by this additional layers is added to network that makes model results are faster.
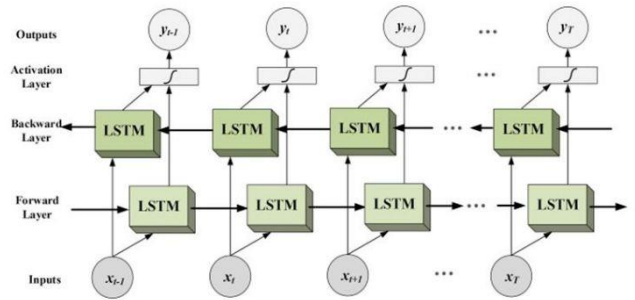

Figure 8 : Bidirectional LSTM[19]

Bidirectional Recurrent Neural Networks (RNNs) is very straightforward. which include two independent RNN together. Using bidirectional run input in tow way, one from past to future and second future to past.

Bidirectional recurrent neural network (BRNN) trained all available input info in the past and future of a particular time-step. for more detail refer to Mike Schuster et al.[20].

Feature set

In several studies in classification feature selection play important role in determining the performance.[8]combine lexical, syntactic and semantic feature selection techniques in the insincere question classification task. Typically unigram[8],[16],[4], bai- gram[8],[5],[16],[4] used, although many alternative feature are used. Authors are also experimented with pre-train word embedding[6],[3],[11],[15],[10].In [6],topic enhancement embedding are able to improve insincere question classification. New BERT pre -trained model[7], [2], [10] introduce for insincere classification.

A. Pre-train Word embedding

Pre-trained Word Embeddings are embeddings that learned one task that are used for solving another similar task. These types of embeddings are trained on large datasets, saved, and then used for solving other tasks. That's why pre-trained word embeddings called Transfer Learning**.**

Pre-trained word embeddings capture the syntactic and semantic meaning of a word as they are trained on large datasets. Learn from the past information and They are capable of boosting the performance of the NLP model. Google's Word2Vec and Stanford's GloVe popular pre-train word embedding.

Word2Vec is based on continuous Bag-of-Words (CBOW) and Skip-gram architectures which can provide high quality word embedding vectors.

GloVe word embedding is a global log-bilinear regression model and is based on co-occurrence and factorization of matrix in order to get vectors. In this study authors used

300-dimension word2vec embeddings trained on Google News and 300-dimension GloVe word embeddings.

For more information on pre-train word embedding refer to the article [22].

V. PERFORMANCE

Table 1 represent the past collected work .This values not directly compare they calculated using varied approaches ,methods, experiment setup, datasets, or metrics. However, the table does present us with rough idea of current performance.

Table 1. Performance Values Form Last Works

Reference no.	Reported performance
[8]	F:87.81%
[5]	Acc:0.96031,P:0.96962,
[6]	F:0.667
[3]	F:0.6913
[11]	Acc:64.35%
[7]	F:0.630
[16]	Acc:66.33%
[15]	F:0.6820
[4]	F:0.698,Acc:0.58449
[2]	F:97.0%(B),F:84.0%(M)
[12]	F:0.8675,Acc:0.8691
[23]	F: 0.677
[24]	F: 0.6732(M)

F:F1-Score, Acc: Accuracy, P:Precision ,B:Binary class

,M:Multi class

VI. CONCLUSION

This paper presents a thorough study of important techniques that classify insincere questions on quora. An effort has been done to cover maximum essential techniques out of all the existing techniques used in this direction. As discussed in paper, there are two common technique for classifying insincere questions detection. We discovered that primitive machine learning approaches working with feature selection technique give effective performance. Deep learning approaches have been effectively used for semantic feature selection, pre-train word embedding methods. After this research, it found that using deep learning with semantic feature give better result. For future work, there are some research in early stage With the advances of deep learning research and applications, we believe that there will be more exciting research of deep learning for sentiment analysis in the near future.

REFERENCES

[1] Quora insincere questions classification. https://www.kaggle.com/c/quora-insincere-questions-classification

[2] D. F. Ashwin Geet D'Sa, Irina Illina, "BERT and fastText Embeddings for Automatic Detection of Toxic Speech", SIIE 2020- Information Systems and Economic Intelligence.

[3] D. G. N. L. S. S. Bishal Gaire, Bishal Rijal, "Insincere Question Classification Using Deep Learning", International Journal of Scientific and Engineering Research vol. 10 (2019) pp. 2001– 2004

[4] S. J. N. Chandra Prakash, Singh Sengar, "Sarcasm Detection in Tweets as Contrast Sentiment in Words Using Machine Learning and Deep Learning Approaches", Machine Learning, Image Processing, Network Security and Data Sciences (2020) pp. 73–84.

[5] Y. U. A. K. X. L. Deepak Kumar Jain, Rachna Jain, "Deep Refinement: capsule network with attention mechanism-based system for text classification", Neural Computing and Applications.

[6] S. W. Y. Z. R. K.-W. L. Do Yeon Kim, Xiao- hang Li, "Topic Enhanced Word Embedding for Toxic Content Detection in Q and A Sites", 2019 IEEE International Conference on Advances in Social Networks Analysis and Mining (2019) 1064–1071. Available: https:http://dx.doi.org/10.1145/3341161.3345332

[7] J. C. s. Hao Mao, Rekha Kumar, Classification of insincere questions on quora (2019).URL cs229.stanford.edu/proj2019aut/

[8] M. A. Hendri Priyambowo, "Insincere Question Classification on Question Answering Forum", 2019 International Conference on Electrical Engineering and Informatics.

[9] I.W.stats, Internet world stats usage and population statistics (2020). https://www.internetworldstats. com/stats.htm

[10] N. R. C. Michael Lanier, Sibi Shanmugaraj, Insincere questions classification on quora using pre-trained word embeddings. http://cs229.stanford.edu/proj2019aut/data/assignment_308875_raw / /26426561.pdf

[11] A. Mishra, S. Pal, Iitbhu at CIQ 2019 : Classification of insincere questions (2019).URLhttps://cse-iitbhu.github.io/ irlab/index.html

[12] I. A. Mohammad A. Al-Ramahi, "Using Data Analytics to Filter Insincere Posts from Online Social Networks. A case study: Quora Insincere Questions", Proceedings of the 53rd Hawaii International Conference on System Sciences. Available: https://scholarspace.manoa.hawaii.edu/handle/10125/64046

[13] C. Olah, Understanding lstm networks. (2015).Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[14] T. Point, Ml - support vector machine(svm). Available: https://www.tutorialspoint.com/machine_learning_with_python/mac hine_learning_with_python_classification_algorithms_support_vect or_machine.htm.

[15] S. L. Sreyan Ghosh, Sonal Kumar, S. S. Jain,Toxic Text Classification (2020) page. 251–260.

[16] D. M. S. Vandan Mujadia, Pruthwik Mishra, Classification of insincere questions with ml and neural approaches (Hyderabad 2019).

[17] Cho,k., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. pages 1724-1734.

[18] https://towardsdatascience.com/understanding-gru-networks 2ef37df6c9be.

[19] https://www.i2tutorials.com/deep-dive-into-bidirectional-lstm/

[20] Mike Schuster and Kuldip K. Paliwal, Member, IEEE, Bidirectional Recurrent Neural Networks, 2673- 2681.

[21] https://medium.com/analytics-vidhya/understanding-logistic-regression-in-depth-intuition-99ad14724464

[22] https://www.analyticsvidhya.com/blog/2020/03/pretrained-word-embeddings-nlp/

[23] Madhu Hegde and Rohit Aggarwal,Qoura Insincere Question Classification.

[24] Sourya Dipta Das, Ayan Basak, and Soumil Mandal, Fine Grained Insincere Questions Classi_cation using Ensembles of Bidirectional LSTM-GRU Model.