iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42959

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

A Survey Paper on Stance Detection of Tweets on Farmers Protests in India

Mr. Prathamesh Badgujar[1], Mr. Aditya Kamble[2], Mr. Anuj Kadam[3], Mr. Dhruv Shah[4], Mr. Anilkumar. J. Kadam[5]
[1, 2, 3, 4, 5]*Dept. of Computer Engineering, AISSMS College of Engineering, Pune, India*

Abstract: ***Protests are a critical part of democracy and a crucial medium for people to deliver their needs and/or dissatisfaction to the authorities. As farmers felt a threat to their rights, there were more and more protests all around the nation. With the development of this technology, additionally there has been a sudden rise in the use of social network sites to trade facts and ideas. In this study, we collected information from the social networking internet site Twitter regarding the Farmers' protest to apprehend the feelings that the twitter users shared on a worldwide platform. In the midst of this protest, social media users had been very lively in voicing their opinion about the matter using the "#FarmersProtest". With lots of people tweeting with the hashtag daily. Through the Stance Prediction of over 850,000 tweets and over 150,000 Users, We intend to decide the Inclination of Common Citizen in addition to Influential People on the now repealed Laws.***

Keywords: Stance Detection, ULMFIT, Language Model, Farmers, Twitter

I. INTRODUCTION

The Farmer protests in India were an ongoing focal point of the Indian and International media, which occurred in the northern part of the country, protesting towards the 3 Farm acts approved by Parliament in September of 2020. The 3 acts are: The Farmers' Produce Trade and Commerce (Promotion and Facilitation) Act, The Farmers' (Empowerment and Protection) Agreement of Price Assurance and Farm Services Act, and The Essential Commodities (Amendment) Act. These Acts, consistent with the authorities, aimed at "transforming Indian agriculture" and "appealing to non-public investment." The Farmers' (Empowerment and Protection) Agreement on Price Assurance and Farm Services Act, 2020, establishes settlement farming, wherein the farmers would produce plants in trade for a price, at the same time remuneration would be negotiated below contracts with company buyers. Protesting farmers were of the thought that effective buyers will have them bound to destructive contracts that were created through important company regulation firms.

More than 41,000 protesters have dedicated themselves to make it certain that their demand to recall these laws is acted upon. The farmers feigned rejection at the authorities' suggestion of postponing the legal guidelines for a period of 18 months and the authorities rather blamed the protests being the result of misinformation caused online. Protests are a critical part of democratic society and are essential in shaping the direction that society follows.

When a community protests, other communities often come forward in assistance of the community or to voice their personal opinions against the community.

It is crucial to understand the feelings behind the online interactions, so as to understand a protest due to the fact that this lets us consider a broader target audience and be inclusive of each direct and indirect participant. Resulting from the extensive protest, there was an influx of opinions and feelings shared on social media on an worldwide stage. These thoughts shared on social media by the general public consisted of a broad range of opinions. The farmers obtained assistance from all around the world, with huge amounts of people expressing their opinions on social media using hashtags like #FarmersProtest, #IAmWithFarmers, #SpeakUpForFarmers, #IStandWithFarmers and had been significantly used on twitter to express the feelings. However, there was also a lot of backlash from people claiming the protests to be a result of anti-national feelings. Additionally people also believed that the farm bills were within the favour of the farmers and that the protests were in-fact just a result of social manipulation through misinformation. In this fast-paced world, spread of information through social media channels is quite fast and it may additionally result in changing people's opinions about certain events. Hence, there is a requirement to be able to correctly understand the sentiments and stances of people online. Sentiment evaluation is a method of processing natural language to examine and understand human emotions. Stance evaluation is an extension of sentiment evaluation where we make use of a certain concept, to detect the stance of people with respect to that concept.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

II. LITERATURE SURVEY

A. Bo Pang, Lillian Lee, Shivakumar Vaithyanathan "Thumbs up? Sentiment Classification Using Machine Learning Techniques"

In this Paper the authors have tackled the problem of classifying files through subject matter, however through basic sentiment, e.g., figuring out whether or not an evaluation is positive or negative. Using film opinions as statistics, they determined that trendy system getting to know strategies definitively outperform human-produced baselines. However, the system getting to know techniques they employed (Naive Bayes, maximum entropy classification, and Support vector machines) do now no longer carry out as properly on sentiment category as on conventional subject matter-primarily based totally categorization

B. Nalini Chintalapudi, Gopi Battineni, Marzio Di Canio, Getu Gamo Sagaro, Francesco Amenta "Text mining with sentiment analysis on seafarers' medical documents"

In this Paper the authors have used 3 years (2018–2020) patient statistics for evaluation. Adoption of each lexicon and Naïve Bayes' algorithms changed into finished to carry out sentimental evaluation and experiments have been carried out over R statistical tools. Visualisation of symptomatic statistics changed into finished via phrase clouds and 96% of the correlation among clinical troubles and analysis final results has been achieved. They tested the sentiment evaluation with greater than 80% accuracy and precision.

C. Ashwin Neogi, Kirti Garg, Ram Krishn Mishra, Yogesh K Dwivedi "Sentiment evaluation and category of Indian farmers' protest the use of twitter statistics"

In this study the authors collected statistics from the microblogging internet site Twitter regarding farmers' protest to apprehend the emotions that the general public shared on a global level.They used fashions to classify and examine the emotions primarily based totally on a group of around 20,000 tweets at the protest.

D. Jonathon Read. 2005. "Using Emoticons to Reduce Dependency in Machine Learning Techniques for Sentiment Classification."

In this paper the authors demonstrates that in shape with appreciate to area and time is likewise essential, and provides initial experiments with schooling statistics labelled with emoticons, which has the capability of being unbiased of area, subject matter and time.They have decided the have an impact on of area, subject matter and time on system getting to know primarily based totally sentiment category.

E. Jeremy Howard, Sebastian Ruder "Universal Language Model Fine-tuning for Text Classification", arXiv:1801.06146, Jan. 18, 2018

III. PROPOSED SYSTEM

Recently many advancements have been made in transfer learning, in which pre-trained models on large datasets are used, instead of having a model trained from scratch, and the pre-trained model is then fine-tuned for specific NLP tasks. Recently methods such as ULMFiT, OpenAI GPT, ELMo and Google AI's BERT have revolutionised the field of transfer learning in Natural Language Processing by making use of language modelling during pre-training, which results in increased performance for a varied number of NLP tasks. It can be argued that the use of all these language modelling methodologies, has resulted in computers showing much improvement in producing semantic meaning of languages.

Here we implement a modern transfer learning approach — using ULMFiT and show how it can perform classification tasks relatively easily by having it fine-tune itself from a different distribution — in this case, stance classification of Tweets as Positive, Negative or Neutral stance towards Farmers' Protest In India.

Our approach aims to tackle the following challenges:

Does our fine-tuned language model classifier interpret the informal and often grammatically incorrect nature of Tweets?

Does our classifier model produce a reasonable enough F1-score without the provision of huge amounts of computational resources and task-specific fine customisation?

©IJRASET: All Rights are Reserved

The ULMFiT and OpenAI GPT are two particular transfer learning strategies have been selected for this project due to the fact they're very comparable in how they make use of language modelling to carry out both unsupervised (training) and supervised (tuning) methods, usage of both resulting in a semi-supervised approach. Uniquely they acquire generalisation by making use of different network architectures. ULMFiT makes use of a 3-layer bi-LSTM architecture, at the same time as the OpenAI method makes use of a transformer network. Any prior information from the word embeddings are only present within the initial layers of the network — a brand new target task may require the entire network to be trained from scratch.

Derivation of meaning from phrases (along with the ones visible in natural language), these models that utilise word embeddings might still want huge amounts of data to be able to find meaning in large sets of words and "learn" from an entirely brand new vocabulary. Benchmarks performed on similar projects have shown that the data stores required for transferring information via phrase embeddings may be quite great, and that can bring about very huge computational costs. In the past few years, the invention of effective pre-trained language models have proven that it is viable to gain better interpretation of language meaning and semantic structure for brand new tasks, in particular for lengthy phrases, the usage of the information gained from unsupervised training on a large text base. A language model attempts to examine the hierarchical structure of natural language semantics, and as a consequence incorporates low-level representational functions and high-level semantic meaning functions. Generative feature of natural language modelling, means that it aims to predict the subsequent elements of preceding series of phrases. Commonly used unsupervised learning methods aimed at larger datasets make it possible for the model to perform such predictions, allowing the model to learn deeper syntactic features of the language than just making use of word embeddings. Unlabelled data is present everywhere throughout the internet so procuring large text corpora is easier, providing a larger dataset to the model, it is now possible to carry out unsupervised pre-training on huge amount of vocabulary while also at the same time creating a context for a in-depth knowledge of language semantics.

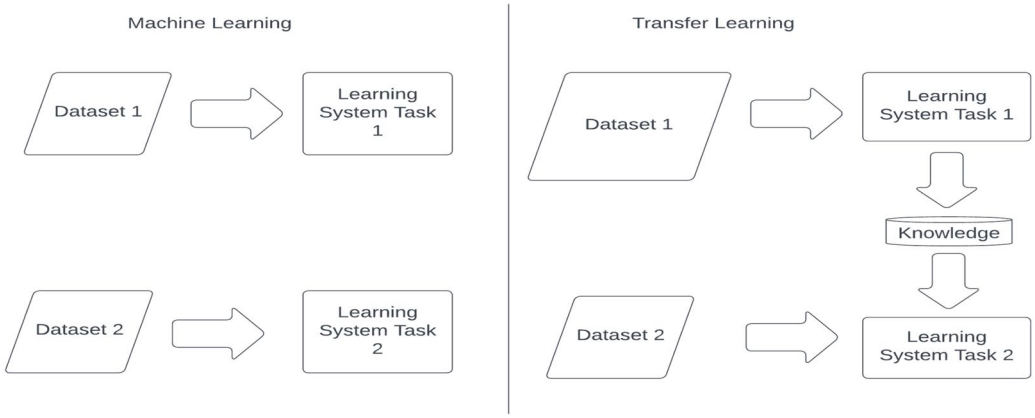


Fig 1. Traditional Machine Learning vs Transfer Learning

IV. FUTURE SCOPE AND CONCLUSION

Social networks have created possibilities to share our opinions, thoughts and critiques for propagating ideas and performing non-public critiques, social networks have increased in usage upto the point of being synonymous with having an "Internet Presence". Being able to analyse the information of social media websites will provide an alternate perspective on human lifestyle and the social environment. This social media presence of the public, provided the Farmers' protest in India a great upward thrust withinside the wide variety of tweets wherein public shared their opinions. The Farmers' Protest resulted in each class of people expressing their thoughts and opinions for/against this issue. Going through various approaches to apprehend the Stance of people and figuring what direction the protests would have gone on if the repeal did not happen. The study also can be elevated to look at how prevalent Covid-19 was all through India due to mass protests and rallies, as Covid-19 was at its peak for the duration of the time when the protests took place. This method can in addition be laboured upon and carried out on a much broader scale through the authorities to gauge the democratic integrity of the country. This study is also valuable withinside the field of public policy, wherein governments can also additionally use NLP techniques to make better policy decisions and analyse public opinions on certain topics. The study can also be useful to examine more than one protest going on the world over with an extra inclusive method to numerous nearby languages.

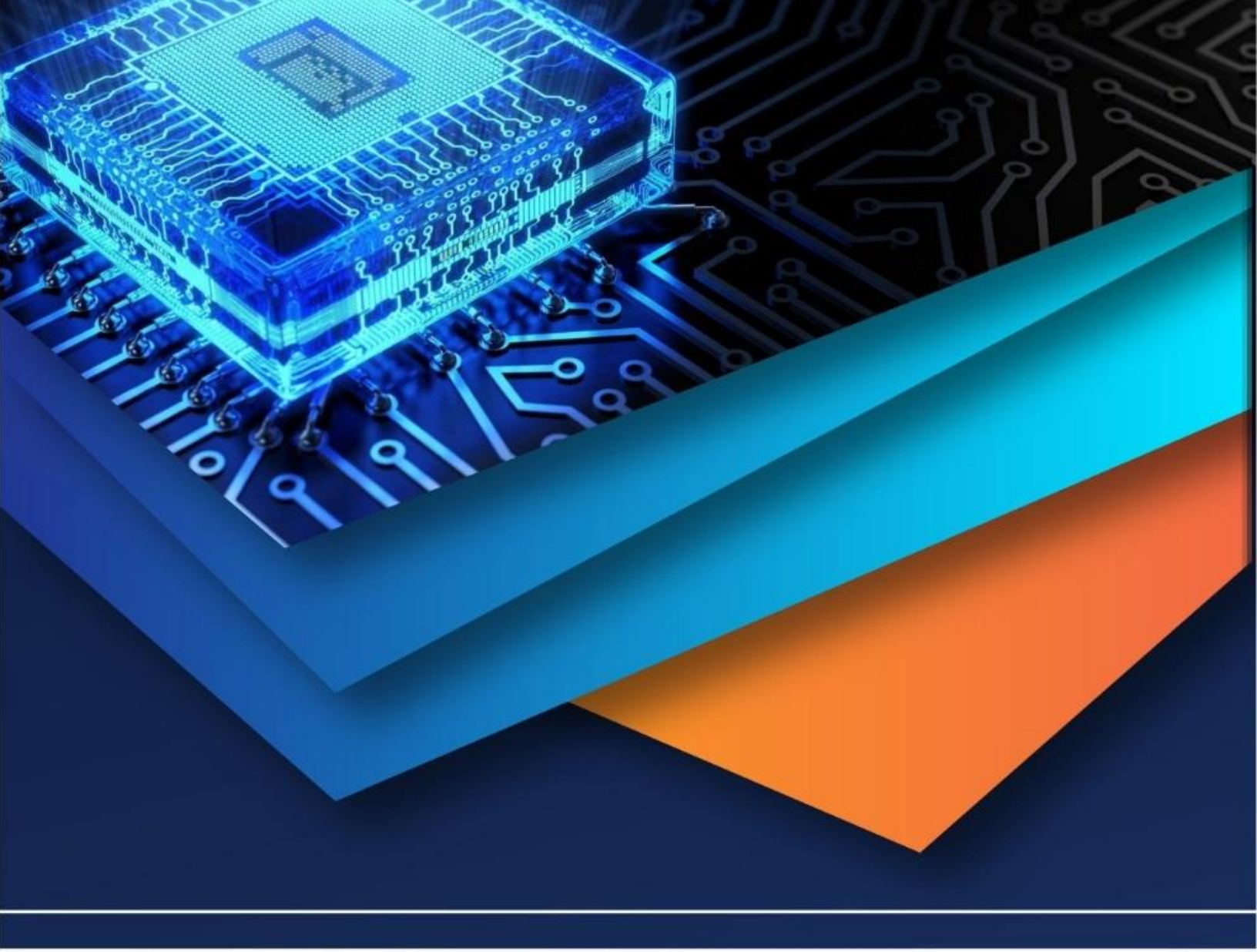
©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. ACKNOWLEDGEMENT

All the faith and honour to our Guide and HOD for his grace and inspiration. I would like to thank all my Friends and Family members who have continually been there to support us. We sincerely thank our Department Head, Project coordinator, our project guide, and all other staff contributors for giving us the guidelines for this paper.

REFERENCES

[1] Bo Pang, Lillian Lee, Shivakumar Vaithyanathan “Thumbs up? Sentiment Classification using Machine Learning Techniques”
[2] Nalini Chintalapudi, Gopi Battineni, Marzio Di Canio, Getu Gamo Sagaro, Francesco Amenta “Text mining with sentiment analysis on seafarers’ medical documents”
[3] Ashwin Neogi, Kirti Garg, Ram Krishn Mishra, Yogesh K Dwivedi “Sentiment analysis and classification of Indian farmers’ protest using twitter data”
[4] Jonathon Read. 2005. “Using Emoticons to Reduce Dependency in Machine Learning Techniques for Sentiment Classification.”
[5] Kareem Darwish, Walid Magdy, and Tahar Zanouda. 2017. Improved Stance Prediction in a User Similarity Feature Space.
[6] Van Engelen, Jesper E., and Holger H. Hoos. "A survey on semi-supervised learning."
[7] Chapelle, Olivier, Bernhard Scholkopf, and Alexander Zien. "Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]." IEEE Transactions on Neural Networks 20.3 (2009): 542-542.
[8] Dutta, Subhabrata, et al. "Semi-supervised Stance Detection of Tweets Via Distant Network Supervision." arXiv preprint arXiv:2201.00614 (2022).
[9] Darwish, Kareem, et al. "Unsupervised user stance detection on twitter." Proceedings of the International AAAI Conference on Web and Social Media. Vol. 14. 2020.
[10] Wei, Penghui, Wenji Mao, and Daniel Zeng. "A target-guided neural memory model for stance detection in Twitter." 2018 International Joint Conference on Neural Networks (IJCNN). IEEE, 2018.
[11] Jeremy Howard, Sebastian Ruder, Universal Language Model Fine-tuning for Text Classification, arXiv:1801.06146, Jan. 18, 2018
[12] https://machinelearningmastery.com/what-is-semi-supervised-learning/
[13] https://towardsdatascience.com/transfer-learning-in-nlp-fecc59f546e4
[14] https://medium.com/technonerds/using-fastais-ulmfit-to-make-multi-label-text-classifier
[15] http://nlpprogress.com/english/stance_detection.html



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)