FAKE NEWS DETECTION USING MACHINE LEARNING

Thesis Submitted
in Partial Fulfillment of the Requirements for the
Degree of

MASTER OF TECHNOLOGY
in
ARTIFICIAL INTELLIGENCE
By

SHEETAL YADAV
(2K22/AFI/23)

Under the Supervision of

Dr. SANJAY KUMAR

Assistant Professor, Department of Computer Science and
Engineering
Delhi Technological University

To the
Department of Computer Science and Engineering

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042
June, 2024

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

<u>ACKNOWLEDGMENT</u>

I wish to express my sincerest gratitude to **DR. SANJAY KUMAR** for his continuous guidance and mentor ship that he provided during research work. He showed me the path to achieving targets by explaining all the tasks to be done and explain to me the importance of this work as well as industrial relevance. He was always ready to help me and clear our doubts regarding any hurdles in the project. Without his constant support and motivation, this work not have been successful.

Place: Delhi   **SHEETAL YADAV**

Date:

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

<u>CANDIDATE'S DECLARATION</u>

I, **Sheetal Yadav 2K22/AFI/23**, of **M.Tech. (AI),** hereby certify that the work which is being presented in the thesis entitled "**Fake news detection using machine learning**" in partial fulfillment of the requirement for the award of the degree of Master of Technology, submitted in the Department of Computer Science and Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from to under the supervision of Dr. SANJAY KUMAR.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor   **Signature of External Examiner**

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

<u>CERTIFICATE BY THE SUPERVISOR</u>

Certified that **Sheetal Yadav (2K22/AFI/23)** has carried out their research work presented in this thesis entitled "**Fake news detection using machine learning**" for the award of **Master of Technology** from Computer Science and Engineering, Delhi Technological University, Delhi under my supervision. The thesis embodies results of original work, and studies are carried out by the student herself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

(Dr. SANJAY KUMAR)

(Assistant Professor)
(Department of Computer Science and Engineering)
(Delhi Technological University)

Date:

FAKE NEWS DETECTION USING MACHINE LEARNING
SHEETAL YADAV

<u>ABSTRACT</u>

In recent times, with the tremendous use of era, faux news and rumors are spreading too. people and society are greatly impacted with the aid of fake information, which also can be used as phishing tries and a way of stealing their information. in lots of areas of our lives, artificial Intelligence (AI) and machine mastering (ML) have confirmed their effectiveness. furthermore, herbal Language Processing (NLP) has proven promising outcomes in textual content class packages. on this examine, we proposed an experimental look at for detecting fake information using ML fashions. The proposed version analyzes the main textual content of the news the use of NLP strategies and then classifies the news into fake or real information. We used a brand new dataset that mixed a couple of fake information datasets. furthermore, we studied the effect of capabilities extraction strategies at the overall performance of the developed models. eight experiments were achieved the usage of Random wooded area (RF) and support Vector Machines (SVM) fashions, each with a exclusive features extraction method. The SVM version resulted within the first-rate overall performance with an accuracy level of ninety eight%. This result proves the model capability to be deployed and used in actual-global with high reliability, to detect faux information.

TABLE OF CONTENT

List of tables

List of figures

CHAPTER 1

INTRODUCTION

1.1 Definition

Human beings are spending increasingly more time interacting on social media, as the huge adoption of smartphones makes their get right of entry to available almost each time and everywhere, which is not the case with traditional media. further, they facilitate interplay with friends, households, and even strangers thru the remark chains, be it via feedback, discussions, or genuinely like and dislike buttons. This has made social media a primary channel for the dissemination of news. however, new technologies and capabilities can be used through social media platforms to spread fake news on a large scale. Such inaccurate facts would possibly end result either from a deliberate try to deceive or mislead (disinformation) or from an sincere mistake (misinformation). Rumours can fall into both of those classes, relying on the motive of the supply, for the reason that rumours aren't necessarily fake however may additionally become authentic. in contrast to rumours, faux news is, via definition, continually false and, accordingly, may be seen as a form of disinformation. consequently, credible and dependable resources of statistics are wanted in order that the general public does now not fall prey to the intentions of these interested in manipulating fact.

Faux news can characteristic as propaganda or misinformation, however it always appeals to the feelings of the general public and the motive to cover rational responses, analysis, and contrast of information from numerous resources, encouraging inflammation and outrage and can without problems lead to conspiracy theories and partisan biased content that negatively impacts. The most important source of information and records for the public is served by using social media and online news articles because it's miles effortlessly reachable, subsidized and quite simply to be had with one click on. however, simultaneously, it also facilitates to unfold false information that has huge poor effects on society, that is, messages which are intentionally misinformed. It has many similarities with spam messages due to the fact they proportion commonplace capabilities which include grammatical mistakes, false statistics, the usage of a in addition constrained set of words, and that they comprise emotionally coloured facts that influences the reader's opinion. to relieve this trouble, research on the identity of fake news has

received extra consideration these days. in spite of the many computational answers currently to be had to come across faux news, the lack of a comprehensive and community-primarily based false information database has become one of the great limitations. massive-scale news passes over social media makes guide verification not possible, which promotes the design and implementation of automatic systems for detecting fake information.

1.2   Fake news characterization

Fake information definition is made from two elements: authenticity and intent. Authenticity means that fake news content material false facts that may be proven as such, because of this that conspiracy concept isn't protected in fake news as there are difficult to be validated proper or fake in most instances. the second one part, rationale, way that the false records has been written with the aim of deceptive the reader.
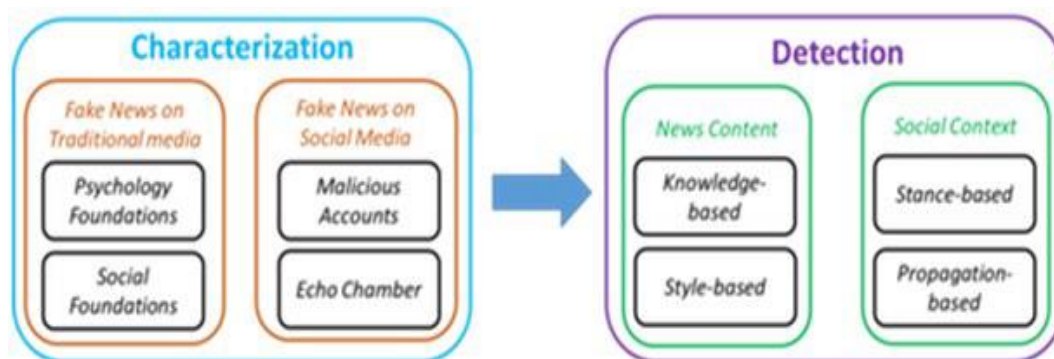
Figure 1.1: Fake news on social media : from characterization to detection.

1.3   Feature extraction

Capabilities can be extracted from those four fundamental additives, with the mains features being linguistic-based totally and visible-based. As defined before, fake information is used to influence the consumer, and so one can do this, they frequently use a particular language so as to draw the readers. alternatively, non-faux information will typically keep on with a distinct language check in, being extra formal. that is linguistic-based capabilities, to which may be added lexical features such as the full range of words, frequency of huge phrases or precise words.

The second features that need to be taken under consideration are visual capabilities. indeed, modified pics are regularly used to add greater

weight to the textual information. For example, the discern 1.2 is supposed to show the progress of deforestation, however the pics are in reality from the identical unique one.

2009   **2019**

Figure 1.2 : Two images provided to show deforestation between two dates are from the same image taken at the same time.

1.4   News content model
1.4.1   Knowledge based model

Now that the exceptional varieties of capabilities available for the information were described, it's far feasible to begin to explain what varieties of fashions can be built the use of these features. The first model that pertains to the news content is based on information: the goal of this model is to check the truthfulness of the information content material and can be completed in 3 one-of-a-kind methods (or a aggregate of them):

• **professional-oriented:** is predicated on experts, along with newshounds or scientists, to assess the information content.

• **Crowd sourcing-orientated:** relies on the wisdom of crowd that announces that if a sufficiently large range of individuals say that something is false or real then it should be.

• **Computational-oriented:** is predicated on automatic reality checking, that might be based on outside resources together with DBpedia.

Those strategies all have professionals and cons, hiring specialists might be costly, and professional are confined in variety and may not be able to treat all of the information this is produced. inside the case of crowd sourcing, it may effortlessly be fooled if sufficient bad annotators smash the system and automatic reality checking may not have the essential accuracy.

1.4.2 Style based model

As defined earlier, fake information generally tries to steer consumer behaviour, and hence typically use a selected fashion which will play at the emotion. those strategies are referred to as deception-orientated stylometric strategies.

The second approach is called objectivity-oriented processes and tries to capture the objectivity of the texts or headlines. these kind of favor is in most cases used by partisan articles or yellow journalism, this is, websites that rely upon headlines without reporting any useful statistics. An instance of these kind of headline will be

you may never believe what he did !!!!!!

This sort of headline plays on the interest of the reader that might click to examine the news.

1.5   Social context model

The ultimate capabilities that have no longer been used yet are social media functions. There are two approaches to use those functions: stance-based and propagation-based totally.

Stance-based strategies use implicit or explicit representation. as an instance, express representation might be nice or poor votes on social media. Implicit illustration needs to be extracted from the publish itself.

Propagation-based totally processes use features associated with sharing along with the number of re tweet on twitter.

CHAPTER 2

2.1 Introduction

In this chapter i can element a piece extra, some associated works which are really worth investigating.

2.2 Supervised learning for fake news detection

Reis et al. use device mastering techniques on buzz-feed article associated with US election. The evaluated set of rules are okay-Nearest Neighbors, Naive-Bayes, Random Forests, SVM with RBF kernel and XGBoost.

So one can feed this network, they used a variety of capabilities including:

• Language features: bag-of-phrases, POS tagging and others for a total of 31 different functions.
• Lexical capabilities: range of precise words and their frequencies, pronouns, etc,
• Psychological capabilities: construct using Linguistic Inquiry and phrase remember which is a selected dictionary construct by way of a textual content mining software program,
• Semantic functions: poisonous rating from Google's API,
• Engagement: number of feedback within numerous time c program language period.

Many other capabilities were extensively utilized, based totally at the source and social meta data.

Their consequences is proven at figure 2.1.
additionally they show that XGBoost is good for deciding on texts that need to be hand-verified, this means that the texts categorized as dependable are indeed dependable, and as a result reducing the quantity of texts the be checked manually. This version is restricted with the aid of the truth they do use meta data that isn't always always to be had. P´erez-Rosas et al.[13] used almost the identical set of features but used linear SVM as a model and worked on a extraordinary data set.

Table 1. Results obtained for different classifiers w.r.t AUC and F1 score.

Classifier	AUC	F1
KNN	0.80±0.009	0.75±0.008
NB	0.72±0.009	0.75±0.001
RF	**0.85±0.007**	**0.81±0.008**
SVM	0.79±0.030	0.76±0.019
XGB	**0.86±0.006**	**0.81±0.011**

RF and XGB performed best.

Figure 2.1: Results by Reis et al.

2.3 CNN for fake news detection

Yang et al. used a CNN with pix contained in article so that it will make the classification. They used kaggle faux information dataset1, further they scrapped real information from depended on source which includes new york times and Washington post.

Their community is made from two branches: one textual content department and one image department. The textual department is then divided of subbranch: textual express: derived records from textual content consisting of duration of the news and the text latent subbranch, that's the embedding of the textual content, restricted to 1000 words.
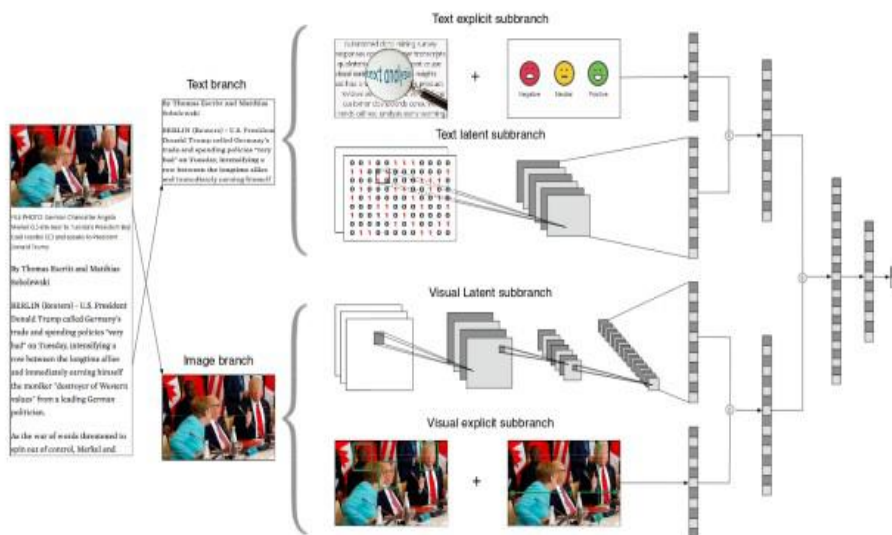

Figure 2.2 TI-CNN

The photograph branch is likewise product of subbranch, one containing statistics including image resolution or the variety of humans present on

the picture, the second one subbranch use a CNN on the picture itself. the full info of the community are at Figure 2.3. And the outcomes are at determine table 2 and display that indeed using pics works better.

2.4 Conclusion

we've got visible inside the preceding sections that most of the associated works attention on enhancing the prediction best by way of including additional functions. The fact is that those functions are now not usually to be had, as an example a few article may not comprise pix. there may be also the truth that the use of social media records is elaborate as it is simple to create a new account on those media and idiot the detection system. That's why I selected to attention on the thing body most effective and spot if it is viable to appropriately come across faux information.

Text Branch		Image Branch	
Textual Explicit	Textual Latent	Visual Latent	Visual Explicit
Input 31×1	Emb 1000×100	Input $50 \times 50 \times 3$	Input 4×1
	Dropout D_α	(2×2) Conv(32)	
		ReLU	
Dense 128	Emb 1000×100	Dropout D_β	Dense 128
	$(3,3)$ Conv1D(10)	(2×2) Maxpo	
	2 MaxPo1D	(2×2) Conv(32)	
	Flatten	ReLU	
BN	Dense 128	Dropout D_β	BN
		(2×2) Maxpo	
	BN	(2×2) Conv(32)	
		ReLU	
	ReLU	Dropout D_β	
		(2×2) Maxpo	
ReLU	Dropout D_β	Flatten	ReLU
		Dense 128	
		BN	
		RelU	
Merge		Merge	
Merge			
ReLU			
Dense 128			
BN			
Sigmoid			

Figure 2.3 TI-CNN

Method	Precision	Recall	F1-measure
CNN-image	0.5387	0.4215	0.4729
LR-text-1000	0.5703	0.4114	0.4780
CNN-text-1000	0.8722	0.9079	0.8897
LSTM-text-400	0.9146	0.8704	0.8920
GRU-text-400	0.8875	0.8643	0.8758
TI-CNN-1000	**0.9220**	**0.9277**	**0.9210**

Table 2: TI-CNN results

CHAPTER 3

3.1 Introduction

In this chapter, we can awareness on the greater conventional methods used in herbal language processing which includes Naive-Bayes, choice bushes, linear SVM and others. these will serve as a baseline for evaluating the performances of the more two advanced fashions a good way to be analyzed later on: LSTM and attention Mechanism. the primary issue to do when operating with text is the do phrases and texts embedding, certainly, with a purpose to use device learning algorithms on texts, a mathematical representation of these texts is required.

3.2 Methodology

All of the methods supplied can be tested in special ways:

• on the liar-liar dataset
• at the faux corpus dataset, aside from the information from beforeitsnews.com and big appleinstances.com

To be extra unique, in the first case, the models can be trained on a schooling set, tuned using validation set and eventually examined the use of test set. inside the 2nd case, the same techniqueology may be used, the dataset has been cut up be choosing 60% of the text from every area for schooling, and 20% for validation and checking out. This manner of splitting has been selected due to the choppy representation of every domain in the dataset as a way to make certain illustration of all of the domains within the tree subsets.

3.2.1 Evaluation matrices

With a view to examine each model, more than one assessment metrics have been used. There are do not forget, precision and f1-rating. it is had to use more than one metrics due to the fact they don't all account for the same values. as an example, it's miles viable to have a model with a recollect of 1 that behave extremely horrific because it truely classifies all of the inputs within the same single class.

Remember that Precision Is defined As

$$Precision = \frac{TP}{TP + FP}$$

(3.1)

which means that that we will have two distinct precision, depending on which training is taken into consideration as being tremendous. this is the share of efficaciously labeled fine elements over the quantity of factors categorised as nice. it's miles equals to 1 whilst there may be no fake tremendous, however it does now not suggest that each one the superb elements are efficiently labeled as it might be a few false negative. The consider allows to remedy this hassle.

It's miles described As

$$recall = \frac{TP}{TP + FN}$$

(3.2)

The f1-rating combines the take into account and the precision.
It's far defined by means of

$$f1 - score = \frac{2 * precision * recall}{precision + recall}$$

(3.3)

it is also viable to have a look at the weighted average of these types of values. for instance, it is possible to compute the worldwide remember via averaging the take into account for each instructions by using the respective class ratio.

In the end, uncooked output may be used by looking at the confusion matrix.

The first parameter to music is the max number of functions utilized by tf-idf. That is the maximum quantity of words with the intention to be stored to create the textual content encoding. The words that are saved are the maximum frequent words.

3.3 Models

3.3.1 Naive bayes

The basic idea of Na¨ıve-Bayes model is that all features are independent of each other.

This is a particularly strong hypothesis in the case of text classification because it supposes that words are not related to each other. But it knows to work well given this hypothesis.

Given an element of class y and vector of features $X = (x1, ..., xn)$. The probability of
the class given that vector is defined as

$$P(y|\mathbf{X}) = \frac{P(y) * P(\mathbf{X}|y)}{P(\mathbf{X})}$$

(3.4)

Thanks to the assumption of conditional independence, we have that
$P(xi |y, x1, ..., xi-1, xi+1, ..., xn) = P(xi |y)$ (3.5)

Using Bayes rules we have that

$$P(y|x_1, ..., x_n) = \frac{P(y) \prod_{i=1}^{n} P(x_i|y)}{P(x_1, ..., x_n)}$$

(3.6)

Because $P(x1, ..., xn)$ is constant, we have the classification rule

$$\hat{y} = \underset{y}{argmax} P(y) \prod_{i=1}^{n} P(x_i|y)$$

(3.7)

3.3.2 Linear SVM

Linear SVM is a method for large linear classification. Given pairs of features-label (xi, yi), $yi \in \{-1, 1\}$, it solves the following unconstrained optimization problem.

$$\underset{w}{min} \frac{1}{2}\mathbf{w}^{\mathbf{T}}\mathbf{w} + \mathbf{C} \sum_{i=1}^{l} \xi(\mathbf{w}; \mathbf{x_i}, y_u)$$

(3.8)

Where ξ is a loss function, in this case L2 loss function has been used, and $C > 0$ a penalty parameter. Class of new examples are assigned by looking at the value of wTw.

The class 1 is assigned if wTw ≥ 0 and the class -1 if wTw < 0.

3.3.3 Decision Tree

Decision tree works by recursively selecting features and splitting the data set on those features. These features can either be nominal or continuous.

In order to find the best split, it uses gini impurity.

$$G = \sum_{i=1}^{C} p(i) * (1 - p(i))$$

wherein p(i) is the probability of class i in the present day branch. The nice split is chosen as the one that decreases the most the impurity. as an instance, starting from the basis, the gini impurity is computed on the whole data set, then the impurity of each department is computed over all capabilities, weighting it by means of the quantity of factors in every department. The chosen feature is the one that has the highest impurity.

3.3.4 Ridge classifiers

Ridge classifier works the same way as ridge regression. It states the problem as a minimization of the sum of square errors with penalization. It can be expressed as in Equation 3.10.

$$\min_{w} ||Xw - y||_2^2 + \alpha ||w||_2^2$$

The predicted class if positive if Xw is positive and negative otherwise.

CHAPTER 4

Result

The net is an critical source for users to achieve records. there are many unique social networking websites, along with fb and Twitter, that help humans connect to others. one-of-a-kind media are also shared on these structures. these days, human beings opt to get their news from those platforms because they're easy to use and get admission to. any other benefit for humans is that those structures permit comments, protests, and so on. It offers options for. These advantages attract users of these platforms. However, no matter how good these platforms are, they are also used by cyber criminals for their own interests. These people can spread false news through these platforms. These platforms also have the feature of sharing articles or news, which has been proven to help spread fake news. People are starting to believe this news and sharing it with others. Researchers say it's miles hard to govern the spread of fake news on social media systems.

Anyone can register on those systems and begin sharing. a person can create a page to make news and spread fake news. those platforms aren't certain whether the person is a reputable advertiser or no longer. in this manner, anyone can submit messages concentrated on people or groups. these fake news can harm a society or a political party. reviews show that it is straightforward to change human beings's minds by means of spreading fake news. consequently, the spread of this fake news must be investigated so that it will save the recognition of the character, political celebration or employer.

1.Why machine learning needed to detect fake news?

Machine learning algorithms can analyze big quantities of facts quickly and efficiently. Because so much news is published every day, it is not possible for people to analyze every article. News organizations and social media platforms can easily detect fake news thanks to the ability of machine learning algorithms to quickly process big data. Machine learning algorithms can detect fake news articles by analyzing words, sources and social media, as well as the news section. The ability to instantly detect fake news based on machine learning algorithms. Building on previous fake news ideas, machine learning algorithms can be trained to detect new trends and new types of fake news. As a result, people have less to do, freeing up time for things like fact-checking and investigative journalism.

Once introduced, these algorithms can be widely used without spending a lot of money.

There are specific gadget getting to know strategies that may be used to come across faux news, so that you can be replied in the next question.

2.What machine learning classes can be used to detect fake news?

Detecting fake information is one in all humanity's maximum difficult duties. faux news can be easily detected using device gaining knowledge of. There are one-of-a-kind gaining knowledge of machines that can help determine whether or not the news is proper or false. in recent times, facts to expose these distributions can be effortlessly collected. different researchers use system getting to know classifiers to fact-test information. Researchers use system mastering classifiers to come across fake information. in line with the researchers' experiments, SVM and Naive Bayes classifiers are the first-class at detecting fake news. these classifiers are better than other classifiers in phrases of the accuracy they provide. A classifier with a higher number is taken into consideration a good classifier. What subjects maximum is what every distributor gives. more correct classifications will help locate extra fake information. Researchers say it is important to analyze fake news because many humans unfold faux information through social media to mislead the public. It should be managed to prevent individuals or corporations from dropping popularity because of fake information. they are saying gadget getting to know can be very useful on this regard. They used many system studying algorithms and additionally observed that logistic regression turned into a better classifier because it provided more accuracy.

Researchers say social media produces more posts. all of us can sign in on these platforms and post some thing. these posts may additionally incorporate false facts approximately a person or commercial enterprise. Detecting these fake information is an important and difficult assignment. To try this paintings, researchers use 3 device mastering methods. those are Naive Bayes, Neural Networks and aid Vector Machines. Naive Bayes gives 96.08% accuracy. on the other hand, techniques, neural community and SVM, deliver ninety.90% accuracy.

Researchers say fake news has a huge impact on social situations. Fake news on social media platforms can change people's minds. People will change their minds based on fake news without evidence. A method that can capture this type of information is needed. Researchers used gadget studying classifiers for this motive. Classifiers utilized by distinctive
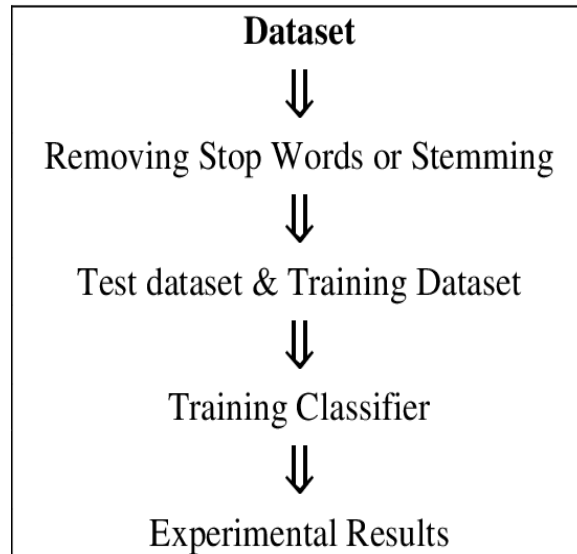
researchers are k nearest neighbor, aid vector device, logistic regression, linear support vector device, pruning tree, stochastic gradient descent. The effects display that the help vector system has accurate accuracy in detecting fake news.

Researchers use gadget getting to know classifiers to discover fake news. They use extraordinary capabilities to educate these training. classroom training is an important venture because trained instructors can deliver better consequences. Researchers say synthetic intelligence can be extra effective at detecting faux information. They used Naive Bayes classifier to discover fake news in fb posts. This class gave them seventy four% accuracy, however they say the accuracy may be progressed. to enhance accuracy, the researchers additionally describe diverse methods within the article. there are various device gaining knowledge of classifiers that can be used to perceive faux news.

3.How to train a machine learning classifier to detect fake news?

Training machine studying classifiers is an critical venture. This plays an vital function in the accuracy of these consequences. The technique need to be nicely educated to apply the right data. unique researchers skilled separate mastering machines to stumble on fake news. the primary problem whilst education these classifications is that the training data is regularly now not equal. Researchers use managed machine mastering to detect faux information. They used three distinctive models for characteristic extraction to teach this classifier. In fact, those features are used to train classifiers. those models are TF-IDF version, N-Gram version, Bag of phrases model. This version extracts capabilities from the training records after which makes use of these functions to teach classes. The researchers did some machine learning training to detect fake news. They use educational materials for education. First of all, they remove unnecessary words and turn the words into a document. Therefore, the educational materials provided for these classifications contain only useful information.

Dataset
⇓
Removing Stop Words or Stemming
⇓
Test dataset & Training Dataset
⇓
Training Classifier
⇓
Experimental Results

CHAPTER 5

5.1 Conclusion

With the increase in internet usage, it is now very easy for fake news to spread.Many people are constantly linked to the net and social media platforms. There are no restrictions on publishing information on these platforms. For this reason, some people use these platforms to spread fake news targeting people or organizations. This could damage one's reputation or impact business. People's opinions about a political party can also be changed through fake news. There must be a way to detect this fake news. Machine learning systems have different applications and can be used to identify false news. The classifier is first trained using a data set called training data. Classifiers can detect fake news.

5.2 Future Scope

One such literature review discusses the maintenance of learning machines that need field data for training. There is no readily available recorded data that could be used to train censors to spot fake news. Future research could include using unsupervised machine learning to detect fake news.

REFERENCES

Abdullah-All-Tanvir, Mahir, E. M., Akhter S., & Huq, M. R. (2019). Detecting Fake News using Machine Learning and Deep Learning Algorithms. 7th International Conference on Smart Computing & Communications (ICSCC), Sarawak, Malaysia, Malaysia, 2019, pp.1-5, https://doi.org/10.1109/ICSCC.2019.8843612

Ahmed, H., Traore, I., & Saad, S. (2017). Detection of online fake news using n-gram analysis and machine learning techniques. Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, 127–138, Springer, Vancouver, Canada, 2017. https://doi.org/10.1007/978-3-319-69155-8_9

Al Asaad, B., & Erascu, M. (2018). A Tool for Fake News Detection. 2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Timisoara, Romania, 2018, pp.379-386. https://doi.org/10.1109/SYNASC.2018.00064

Aphiwongsophon, S., & Chongstitvatana, P. (2018). Detecting Fake News with Machine Learning Method. 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 528-531. https://doi.org/10.1109/ECTICon.2018.8620051

Della Vedova, M. L., Tacchini, E., Moret, S., Ballarin, G., DiPierro, M., & de Alfaro, L. (2018). Automatic online fake news detection combining content and social signals. FRUCT'22: Proceedings of the 22st Conference of Open Innovations Association FRUCT. Pages 272–279. https://dl.acm.org/doi/10.5555/3266365.3266403

Dewey, C. (2016). Facebook has repeatedly trended fake news since firing its human editors. The Washington Post, Oct. 12, 2016.
Donepudi, P. K. (2019). Automation and Machine Learning in Transforming the Financial Industry. Asian Business Review, 9(3), 129-138. https://doi.org/10.18034/abr.v9i3.494

Donepudi, P. K. (2020). Crowdsourced Software Testing: A Timely Opportunity. Engineering International, 8(1), 25-30. https://doi.org/10.18034/ei.v8i1.491

Donepudi, P. K., Ahmed, A. A. A., Saha, S. (2020a). Emerging Market Economy (EME) and Artificial Intelligence (AI): Consequences for the Future of Jobs.

Palarch's Journal of Archaeology of Egypt/Egyptology, 17(6), 5562-5574. https://archives.palarch.nl/index.php/jae/article/view/1829

Donepudi, P. K., Banu, M. H., Khan, W., Neogy, T. K., Asadullah, ABM., Ahmed, A. A. A. (2020b). Artificial Intelligence and Machine Learning in Treasury Management: A Systematic Literature Review. International Journal of Management, 11(11), 13-22.

Granik, M., & Mesyura, V. (2017). Fake news detection using naive Bayes classifier. 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, pp.900-903, https://doi.org/10.1109/UKRCON.2017.8100379

Plagiarism report

Curriculum vitae