

“A Framework for Fake News Detection Using Various Machine Learning Algorithms”



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Under Guidance of
Prof. Pawan Meena

Presented By
Nikita Katakwar
(0132CS20MT27)

RADHA RAMAN INSTITUTE OF TECHNOLOGY & SCIENCE, BHOPAL (M.P.)

OUTLINE

2

- ☐ Introduction
- ☐ Literature survey
- ☐ Problem Statement
- ☐ Proposed Work
- ☐ Simulation & Result
- ☐ Conclusions
- ☐ Future Work
- ☐ References
- ☐ Publications

INTRODUCTION

3

- ❑ Fake news is false or misleading information presented as news.
- ❑ It often has the aim of damaging the reputation of a person or entity, or making money through advertising revenue.
- ❑ Detection of fake news online is important in today's society as fresh news content is rapidly being produced as a result of the abundance of technology that is present.
- ❑ Fake News appears in different forms and the examples of their features are click bait, propaganda, satire or parody, sloppy journalism, misleading headings and biased or slanted news.

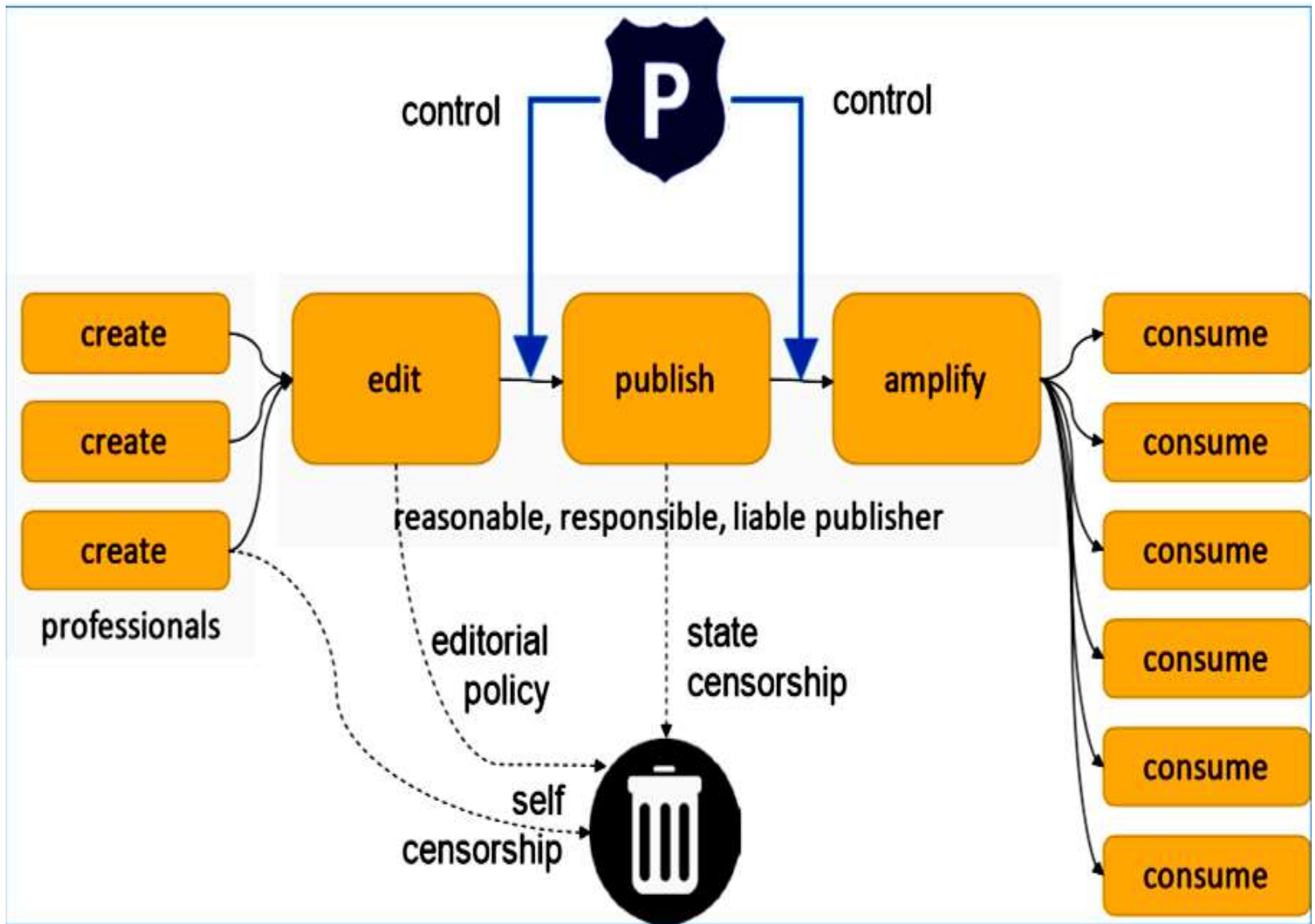


Figure 1: Traditional News Process

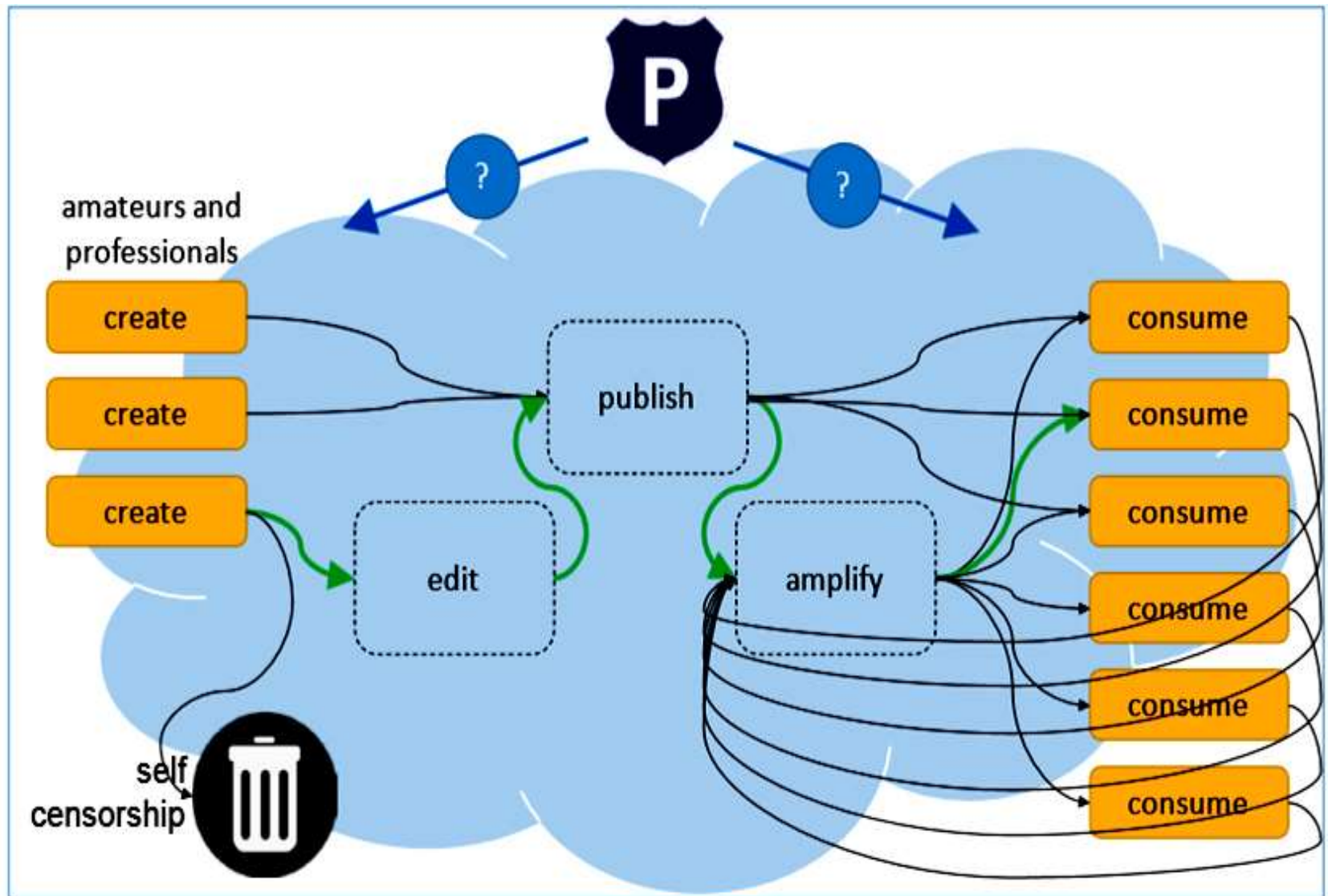


Figure 2: News process in the age of the internet

CONT...

6

- ❑ detection of Fake News becomes more challenging as social media continues to dominate our everyday lives and hence accelerating the speed of which Fake News travel.
- ❑ Fake news detection is still a challenge even to machine learning and deep learning methods because the content of fake news is planned in a way it resembles the truth so as to deceive readers; and without cross referencing and fact checking, it is often difficult to determine veracity by text analysis alone.

LITERATURE SURVEY

7

Sr. No	Author Name	Publication year	Proposed work	Outcome
1	Smitha. N	IEEE, 2020	Performance Comparison of Machine Learning Classifiers for Fake News Detection.	Seven different Machine learning Classification algorithms are trained to classify news as fake or real
2	C. K. Hiramath ,	IEEE, 2019	Proposed deep learning methods	Accuracy is achieved 91% and 9% error rate.

Continued....

8

Sr. No	Author Name	Publication year	Proposed work	Outcome
3	A. Jain	IEEE, 2019	Machine learning and natural language processing based approach.	Correctness of results upto 93.6% of accuracy.
4	R. K. Kaliyar	IEEE, 2019	Experiments have been conducted using a tree-based Ensemble Machine Learning framework	Achieved an accuracy of 86% for multi-class classification of fake news having four classes.

Continued....

9

Sr. No	Author Name	Publication year	Proposed work	Outcome
5	I. Kareem	IEEE, 2019	Supervised Machine Learning (ML) classification algorithms are used	Best performance classifier K Nearest Neighbors (KNN) gives 70% accuracy and logistic regression gives 69% accuracy.
6	C. M. M. Kotteti	IEEE, 2018	TF-IDF vectorization is applied in feature extraction to filter out irrelevant features.	Improves the prediction accuracy by more than 15%.

Continued....

10

Sr. No	Author Name	Publication year	Proposed work	Outcome
7	G. B. Guacho	IEEE, 2018	Feature extraction and supervised classification models leveraging	Proposed method achieves 75.43% of accuracy
8	S. Gilda	IEEE, 2017	Stochastic Gradient Descent model.	Accuracy of 77.2%, with PCFGs having slight effects on recall.

PROBLEM STATEMENT

11

- ❑ The accuracy of machine learning based method is less.
- ❑ More error rate in online fake news detection.
- ❑ Less value of performance parameters like precision, recall f-measure.
- ❑ Existing approach is not so much reliable and optimized.

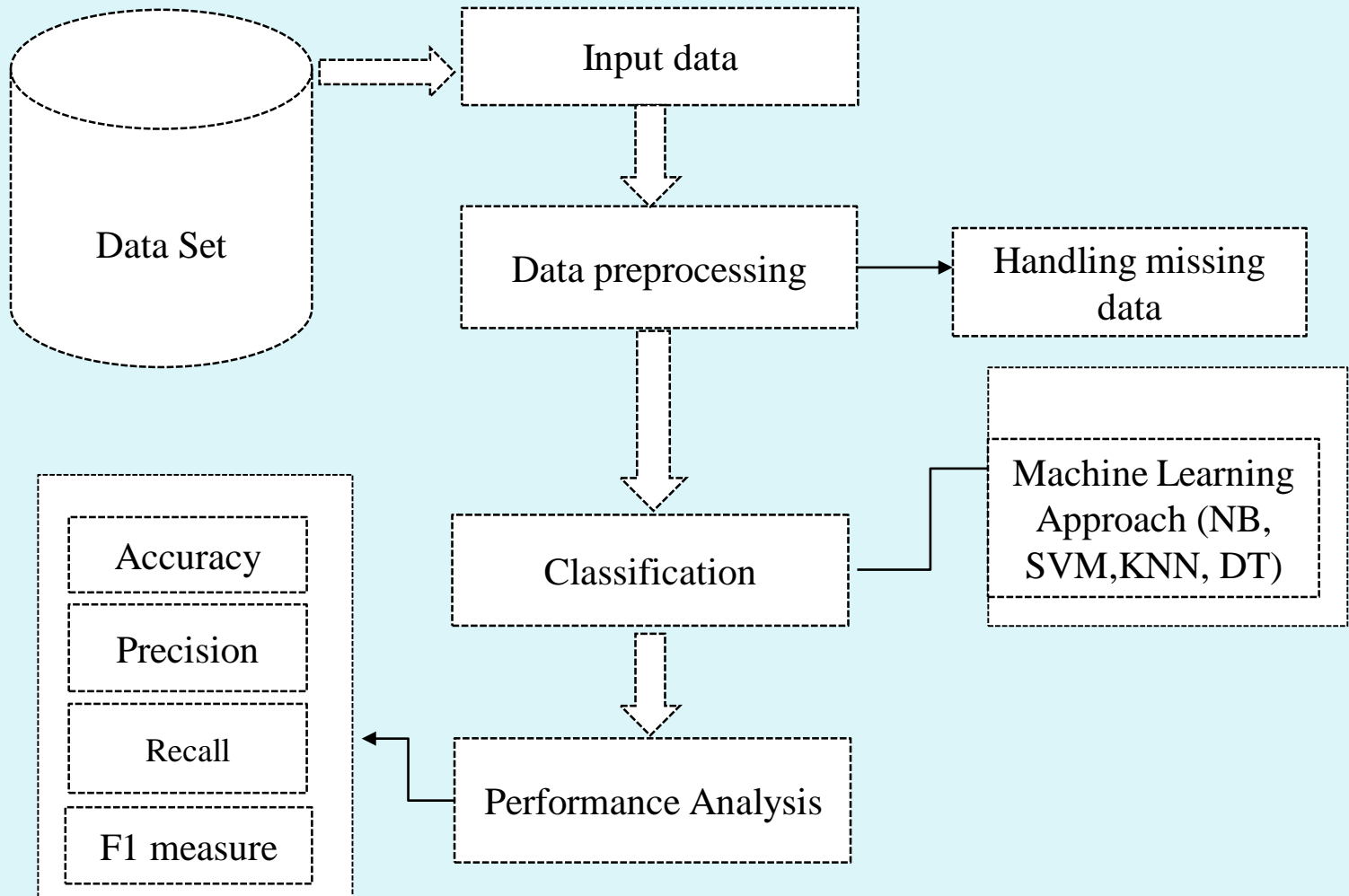
PROPOSED WORK

12

- ❑ now a day's data on web or social media is increasing vastly and it is so hectic to detect news is fake or not by looking all data and it is time consuming so we utilize classification techniques to classify huge data.
- ❑ To propose fake news detection system based on machine learning classification algorithm such as Naïve bayes (NB), Support vector machine (SVM), K Nearest Neighbor (KNN) and decision tree (DT) .
- ❑ To calculate the results parameter and optimize the better approach.

Continued....

13



PROPOSED METHODOLOGY

14

- ❑ Firstly, download the dataset from kaggle website, which is a large dataset provider company for research.
- ❑ Now preprocessing of the data, here handling the missing dataset. Remove the null value or replace from common 1 or 0 value.
- ❑ Now apply the classification method based on the machine learning approach. The Naïve bayes (NB), Support vector machine (SVM), K Nearest Neighbor (KNN) and Decision Tree (DT) machine learning method is applied.
- ❑ Now check and calculate the performance parameters in terms of the precision, recall, F_measure, accuracy and error rate.

Methodology (continued)

15

- ❑ It incorporates expulsion of punctuations, URL's, images, stemming and stop words.
- ❑ At that point classify that information utilizing classifiers, for example, NB, SVM, KNN and DT.
- ❑ News Dataset: Dataset of news is taken from Kaggle website and size of data is 14000 .
- ❑ Processing: Content information needs processing to execute AI on them. There are kinds of methods generally utilized to change over content Information into a structure that is prepared for demonstrating. The information processing steps that are applied.

SIMULATION ENVIRONMENT

16

Software Requirements:

- OS : Windows 10
- Software : Python Spyder 3.6

Hardware Requirements:

- Processor: Intel Core i3 5th generation.
- Ram: 8 GB

SIMULATION PARAMETERS

17

- Precision = True Positive / (True Positive + False Positive)
- Recall = True Positive / (True Positive + False Negative)
- F1-Score = 2x (precision x recall) / (precision + recall)

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

$$\textit{Recall} = \frac{|TP|}{|TP| + |FN|}$$

$$F1 = 2 \cdot \frac{\textit{Precision} \cdot \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

$$\textit{Accuracy} = \frac{|TP| + |TN|}{|TP| + |TN| + |FP| + |FN|}$$

RESULTS (Continued)

Accuracy (%)

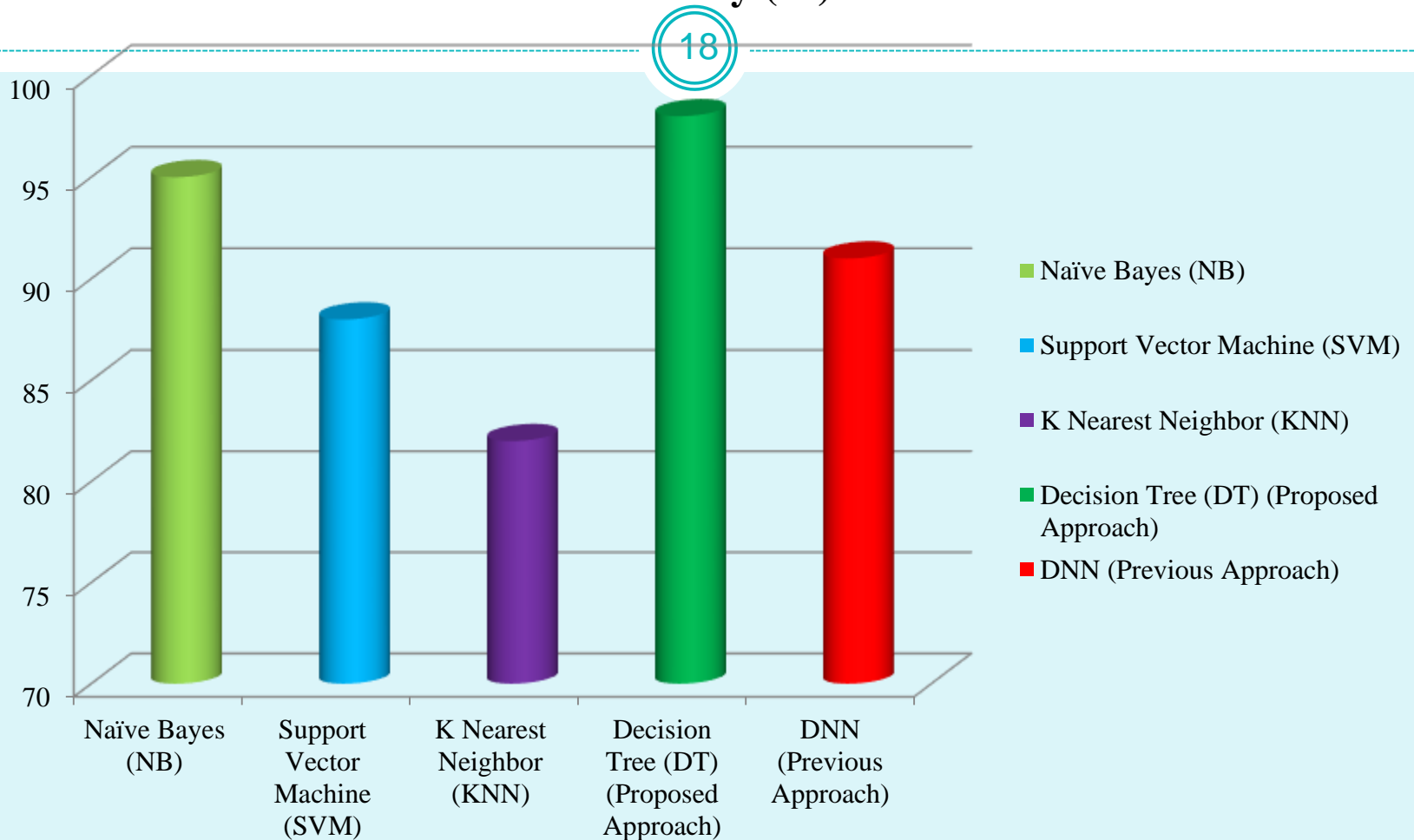


Figure 11: Comparison of accuracy

Figure 10 is showing the comparison of accuracy of various methods. The decision tree method achieves the maximum accuracy that is 98%.

RESULTS (Continued)

19

Figure 2: Comparison of Proposed work with previous work

Sr No	Parameters	Previous Work	Proposed Work
1	Method	Deep Learning	Decision Tree
2	Accuracy	91%	98%
3	Error Rate	9%	2%

RESULTS

20

Table 1: Simulation Results

Sr No.	Methodology	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)	Time (Sec)
1	Naïve Bayes (NB)	96	94	96	95	0.79
2	Support Vector Machine (SVM)	88	89	87	88	0.3
3	K Nearest Neighbor (KNN)	82	83	81	82	2
4	Decision Tree (DT)	98	96	97	98	0.32

CONCLUSION

21

- ❑ News is crucial part of our life. In day to day life current news are helpful to enhance knowledge what happen around the world.
- ❑ This research proposed fake news detection system based on classification such as Naïve bayes (NB), Support vector machine (SVM), K Nearest Neighbor (KNN) and Decision Tree (DT).
- ❑ We compare all machine learning techniques for detecting fake news. Simulation is performed using Python Spyder 3.6 software. Results shows that proposed decision tree method achieves the maximum accuracy that is 98%.

FUTURE WORK

22

- ❑ Future research directions from a data mining perspective and it are outlined in four perspectives: Data-oriented, Feature-oriented, Model-oriented, and Application-oriented.
- ❑ Data-oriented: it focuses on different aspects of fake news data, such as benchmark data collection, psychological validation of fake news, and early fake news detection.
- ❑ Feature-oriented: it aims to explore effective features for detecting fake news from multiple data sources, such as news content and social context.
- ❑ Model-oriented: it opens the door to build more practical and effective models for fake news detection, including supervised, semi-supervised and unsupervised models.
- ❑ Application-oriented: it encompasses research that goes beyond fake news detection, such as fake new diffusion and intervention.

REFERENCES

- A. Uppal, V. Sachdeva and S. Sharma, "Fake news detection using discourse segment structure analysis," 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2020, pp. 751-756, doi: 10.1109/Confluence47617.2020.9058106.
- C. K. Hiramath and G. C. Deshpande, "Fake News Detection Using Deep Learning Techniques," 2019 1st International Conference on Advances in Information Technology (ICAIT), Chikmagalur, India, 2019, pp. 411-415, doi: 10.1109/ICAIT47043.2019.8987258.
- V. M. Krešňáková, M. Sarnovský and P. Butka, "Deep learning methods for Fake News detection," 2019 IEEE 19th International Symposium on Computational Intelligence and Informatics and 7th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics (CINTI-MACRo), Szeged, Hungary, 2019, pp. 000143-000148, doi: 10.1109/CINTI-MACRo49179.2019.9105317.
- A. Jain, A. Shakya, H. Khatter and A. K. Gupta, "A smart System for Fake News Detection Using Machine Learning," 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), GHAZIABAD, India, 2019, pp. 1-4, doi: 10.1109/ICICT46931.2019.8977659.
- R. K. Kaliyar, A. Goswami and P. Narang, "Multiclass Fake News Detection using Ensemble Machine Learning," 2019 IEEE 9th International Conference on Advanced Computing (IACC), Tiruchirappalli, India, 2019, pp. 103-107, doi: 10.1109/IACC48062.2019.8971579.
- I. Kareem and S. M. Awan, "Pakistani Media Fake News Classification using Machine Learning Classifiers," 2019 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, 2019, pp. 1-6, doi: 10.1109/ICIC48496.2019.8966734.

REFERENCES

24

- K. Rajesh, A. Kumar and R. Kadu, "Fraudulent News Detection using Machine Learning Approaches," 2019 Global Conference for Advancement in Technology (GCAT), BANGALURU, India, 2019, pp. 1-5, doi: 10.1109/GCAT47503.2019.8978436.
- B. M. Amine, A. Drif and S. Giordano, "Merging deep learning model for fake news detection," 2019 International Conference on Advanced Electrical Engineering (ICAEE), Algiers, Algeria, 2019, pp. 1-4, doi: 10.1109/ICAEE47123.2019.9015097.
- H. Telang, S. More, Y. Modi and L. Kurup, "Anempirical Analysis of Classification Models for Detection of Fake News Articles," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-7, doi: 10.1109/ICECCT.2019.8869504.
- İ. AYDIN, M. SEVİ and M. U. SALUR, "Detection of Fake Twitter Accounts with Machine Learning Algorithms," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 2018, pp. 1-4, doi: 10.1109/IDAP.2018.8620830.
- R. K. Kaliyar, "Fake News Detection Using A Deep Neural Network," 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-7, doi: 10.1109/CCAA.2018.8777343.
- Y. Seo and C. Jeong, "FaGoN: Fake News Detection model using Grammatic Transformation on Neural Network," 2018 Thirteenth International Conference on Knowledge, Information and Creativity Support Systems (KICSS), Pattaya, Thailand, 2018, pp. 1-5, doi: 10.1109/KICSS45055.2018.8950518.

REFERENCES

25

- C. M. M. Kotteti, X. Dong, N. Li and L. Qian, "Fake News Detection Enhancement with Data Imputation," 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), Athens, 2018, pp. 187-192, doi: 10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00042.
- S. Helmstetter and H. Paulheim, "Weakly Supervised Learning for Fake News Detection on Twitter," 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, 2018, pp. 274-277, doi: 10.1109/ASONAM.2018.8508520.
- S. Khaled, N. El-Tazi and H. M. O. Mokhtar, "Detecting Fake Accounts on Social Media," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 3672-3681, doi: 10.1109/BigData.2018.8621913.
- G. B. Guacho, S. Abdali, N. Shah and E. E. Papalexakis, "Semi-supervised Content-Based Detection of Misinformation via Tensor Embeddings," 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, 2018, pp. 322-325, doi: 10.1109/ASONAM.2018.8508241.
- S. Gilda, "Notice of Violation of IEEE Publication Principles: Evaluating machine learning algorithms for fake news detection," 2017 IEEE 15th Student Conference on Research and Development (SCORED), Putrajaya, 2017, pp. 110-115, doi: 10.1109/SCORED.2017.8305411.

PUBLICATION

26

- [1] Nikita Katakwar, Chetan Agrawal, Pawan Meena “**A Review on Recent Trends of Fake News Detection on Social Media**” International Journal of Innovative Research in Technology and Management, Volume-6, Issue-6, 2022, pp 1-10
- [2] Nikita Katakwar, Chetan Agrawal, Pawan Meena “**A Framework for Fake News Detection Using Various Machine Learning Algorithms**” Engineering Universe for Scientific Research and Management, Volume-14, Issue-4, 2022, pp 1-10

Thank You