|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Serial no.** | **Paper Title** | **Year of publication** | **Datasets** | **Feature extraction techniques** | **Models** | **Optimization techniques** | **Result** | **Evaluation** |
| 1. | Analysis of Classifiers for Fake News Detection | 2020 | LIAR dataset | simple bag-of-words and n-grams and then term frequency  like tf-idf weighting | Naïve Bayes, Logistic Regression, Linear SVM, Stochastic  Gradient Classifier and Random Forest Classifiers |  | The results show that SVM and logistic regression classifier have the best performance on this dataset in the model,  with SVM having a slightly better performance than logistic regression classifier | Naïve Bayes precision = 0.64,  Logistic Regression precision = 0.62,  SVM precision = 0.62,  Stochastic gradient classifier precision = 0.29,  Random Forest classifier precision = 0.6 |
| 2. | A Literature Review of NLP, Approaches to Fake News Detection and Their Applicability to Romanian - Language News Analysis | 2020 | LIAR |  | Multinomial Naïve Bayes, SVM, Random Forest classifier |  | Drawing on the  existing international literature, and in places on our  own related research findings, we also outlined potential  approaches to the automated detection of Romanian  fake news articles and the key points at issue. |  |
| 3. | Identifying Fake News on Social Networks Based on Natural  Language Processing: Trends and Challenges | 2021 | Buzzface,  Fake news net,  Liar, Emergent,  Fake.Br Corpus,  Fever, Credbank,  Buzzfeed news,  Pheme,  BuzzFeed-Webis | The Bag-of-Words (BoW),  TF-IDF | Support Vector Machine (SVM), Random Forest (RF),  k-Nearest Neighbours,  Partitioning-Based Algorithms,  DBSCAN, Hierarchical algorithms |  | The literature  shows that Natural Language Processing (NLP) has been used to detect fake news. We  discussed how NLP could be used to evaluate information from social networks and  compare the different machine learning methods |  |
| 4. | Fake News Classification Bimodal using Convolutional Neural Network and Long  Short-Term Memory | 2020 | Fake news dataset from Kaggle |  | Decision tree, Naïve Bayes, ANN  (Artificial Neural Network), CNN (Convolutional Neural  Network), LSTM (Long Short-Term Memory) | Adam optimizer | It is a good idea to use  CNN+LSTM model to classify text data based on  sentiments. | After  training it for 2 iterations, we were able to get 99.7% and  97.5% accuracy of training and testing. |
| 5. | A Survey on Natural Language Processing for Fake News Detection | 2019 | McIntire’s fake-real-news-dataset |  | Fakebox |  | we evaluate a fake news detector Fake-  box on adversarial attacks, including fact-distortion,  subject-object exchange and cause confounding attack | Fakebox’s accuracy on McIntire’s dataset is  52.77%, false rate is 31.79% and for the other 15.44%  samples |
| 6. | A Survey on Natural Language Processing for Fake News Detection | 2018 | LIAR dataset | Term Frequency-Inverse Document Frequency (TF-IDF) and Linguistic Inquiry and Word | Count (LIWC) are frequently use to Support Vector Machine (SVM Natural Language Processing for Fake News Detection) and Naive Bayes Classifier (NBC),  Recurrent Neural Network (RNN) |  | LSTM based models achieve higher accuracy than CNN based models. | SVM = 0.255,  CNN = 0.274,  MMFD = 0.348,  LSTM = 0.415 |
| 7. | Overview of the GermEval 2021 Shared Task on the  Identification of Toxic, Engaging, and Fact-Claiming Comments | 2021 | 4,000 Facebook user comments, which is drawn  from the Facebook page of a German political talk. |  | Hierarchical Classification |  | A high-level summary of the results by the partic-  ipants in the different subtasks | the best F1-scores  reached in the different subtasks range from 69.98  (subtask 2) to 76.26 (subtask 3) |
| 8. | Sentiment Analysis Using Convolutional Neural Network | 2020 | public  IMDB dataset | Word2vec | CNN |  | Through experiment, we found that the CNN neural network still has good  performance for processing a sequence of data, achieving an accuracy of 88.22%. | After 14500 iterations we achieve  test accuracy of 45.5% |
| 9. | Sentiment Analysis of Health Care Tweets: Review of the Methods Used | 2018 |  |  |  |  | A total of 12 papers studying the quantitative measurement of sentiment in the health care setting were found. More than half of these studies produced tools specifically for their research, 4 used open-source tools available freely, and 2 used commercially available software |  |
| 10. | A Survey on Text Mining and Sentiment Analysis for  Unstructured Web Data | 2015 | Survey |  | Text mining and NLP |  | This exponential increase in the amount of data (termed Big Data) has brought with it a  number of issues related to data management and information extraction. |  |
| 11. | Performance Analysis of Ensemble Methods on  Twitter Sentiment Analysis using NLP Techniques | 2015 | WordNet database, election database, tweets database, movie review dataset |  | Ensemble method, SVM, Baseline, Maxent and Naive Bayes. |  | The precision of ensemble method  outperforms the other solitary methods. | Baseline precision = 83%,  Naïve Bayes precision = 84%,  SVM precision = 87%,  MaxEnt precision = 83%,  Ensemble precision = 90% |
| 12. | Sentiment Analysis with NLP on Twitter Data | 2019 | Twitter database | Bag of words, Tf-Idf | Naive Bayes algorithm, maximum entropy and support vector  Machine. |  | we have seen that  the iPhone phones are more popular than Samsung Phone. |  |
| 13. | Sentiment Analysis in Healthcare: A Brief Review | 2019 | wellbeing remarks (PatientJo),  Twitter information and books surveys (LABR) |  | Deep Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) |  | Applying Convolutional Neural Networks (CNNs) for Sentiment Analysis (SA) has got better results. CNN’s are groundbreaking at separating a various levelled portrayal of the contribution by stacking numerous convolutional and pooling layer. |  |
| 14. | Application of Machine Learning Techniques to  Sentiment Analysis | 2016 | IT Industry (Apple), Bank (ICICI),  Telecom (BSNL) | Term presence and frequency, Parts of speech tagging, Opinion words and phrases, Negation | Naïve Bayes classifier, Support vector machines, Decision trees |  | This work uses  Apache Spark to obtain accurate results fast,  Decision tree takes very less time for predicting  unseen data compared to multinomial Naïve Bayes.  Multinomial Naïve Bayes does not perform as  expected when supplied with small training dataset. | Multinomial naïve baye’s /  Apple accuracy = 84.5%,  ICICI Accuracy = 84.8%,  BSNL Accuracy = 85.6%.  Decision Tree/ Apple accuracy = 100%,  ICICI Accuracy = 100%,  BSNL Accuracy = 100% |
| 15. | A survey on sentiment analysis methods, applications, and challenges | 2022 | Stock market dataset | TF-IDF | Naive Bayes (NB), SVM, Logistic regression, Decision tree, Maximum entropy, KNN, RNN |  |  |  |
| 16. | Product Review Sentiment Analysis by Using NLP and Machine Learning in Bangla Language | 2020 | we collected valid  and understandable Bangla reviews from e-commerce sites | TF-IDF | KNN, Decision tree, SVM, Random Forest, Logistic Regression |  | we observed that  SVM delivered the leading accuracy by using 30% of the  test data. Other algorithms were similarly conducted pretty  well. As SVM produces the most reliable performance, we  determined to employ it to predict the sentiment of the  reviews. | SVM accuracy = 88.81,  Random Forest accuracy = 85.92,  KNN accuracy = 80.14,  Logistic Regression accuracy = 88.09,  Decision tree accuracy = 83.03. |
| 17. | Sentiment analysis on twitter tweets about COVID-19 vaccines  using NLP and supervised KNN classification algorithm | 2021 | Twitter data |  | KNN |  |  |  |
| 18. | Sentiment Analysis of Students’ Feedback with NLP and Deep  Learning: A Systematic Mapping Study | 2021 | ISEAR, SemEval |  | NLP |  |  |  |
| 19. | Analysing the Pulse of Twitter: Sentiment Analysis  using Natural Language Processing Techniques | 2016 | Social Media | CountVectorizer( | NLP, ANN, SVM, Case based reasoning |  | Nevertheless, the result will be classified into 2 categories, that are encoded and un-encoded. According to  security issue for using a data, some of the results will be represented in an ID form like string ID. |  |
| 20. | AIT FHSTP at GermEval 2021:  Automatic Fact Claiming Detection  with Multilingual Transformer Models | 2021 | ClaimBuster dataset |  | mBERT, GottBERT, XLM-R |  | the GottBERT model (run 2)  achieves the highest results in our experiments.  These results indicate that the model that is pre-  trained on German data allows for a better model-  ing of the semantics of the task than a multilingual  model. All other models are also beyond the zero-  rule baseline which is at 66% for the test set | The XLM-R fine-tuned on the Ger-  mEval 2021 data achieved the lowest F1-score  on the test set (69.27%),  The fine-tuned GottBERT on the Ger-  mEval 2021 data achieved an overall F1-score of  74.62% on the test set,  The mBERT seems to generalize well, as  the F1-score on the test set of 72.84% is at a  similar performance level as on the validation se |
| 21. | Gradient-Based Adversarial Training on Transformer Networks  for Detecting Check-Worthy Factual Claims | 2020 |  |  | BERT Transformer. |  | We evaluate our new transformer-based claim-spotting models on  both the Classification and Ranking Tasks (Section 3.1.1). We com-  pare against re-trained and refined versions of past ClaimBuster  models [ 15 , 23 ] and the top-two performing systems from the 2019  CLEF-CheckThat! Challenge. |  |
| 22. | A Machine Learning Approach to Fake News Detection Using Knowledge Verification and Natural Language Processing | 2019 |  |  | Logistic Regression |  |  |  |
| 23. | A transformer-based architecture for fake news classification | 2021 | LIAR dataset,  LIAR PLUS dataset |  | BERT,  support vector machine (SVM), logistic regression (LR) and a bidirectional long short-term memory neural network (Bi-LSTM) | Adam | the test our proposed models with the LIAR and LIAR PLUS datasets under two situations, with binary labels and with six-way labels. In binary labels condition, the ‘true’, ‘half true’ and ‘mostly true’ labels are clubbed into one common one, ‘true’. Similarly, ‘false’ is an amalgamation of the labels ‘mostly false’, ‘false’ and ‘pants-fire’ |  |
| 24. | Fake News Detection Using Machine Learning Approaches | 2021 | The LIAR-PLUS Master | Count Vectorizer and Tiff  Vectorizer | k-Nearest Neighbors (k-NN), Linear Regression,  XGBoost, Naive Bayes, Decision Tree, Random Forests and Support Vector Machine (SVM). |  | As shown the XGBOOST is depicting the highest  accuracy with more than 75%, next is SVM and Random Forest with approximately 73% accuracy | XGboost accuracy = 80%,  Random forest accuracy = 70%,  Naïve Bayes accuracy = 65%,  KNN accuracy = 70%,  Decision tree accuracy = 70%,  SVM accuracy = 80% |
| 25. | Rapid detection of fake news based on machine learning methods | 2021 | l data set containing news, both political and world news | Term Frequency-Inverted Document Frequency (TF-IDF) | Support Vector Machine (SVM), The random forests, CART, Bagging, AdaBoost |  | he best results for all measures used are obtained when the SVM algorithm is used (accuracy 0.9419). However,  it should be noted (Tab. 3) that the running time of the SVM algorithm is the largest. It takes an average of 186.12  seconds, which is more than 78 times worse than the fastest of the algorithms – AdaBoost | CART accuracy = 88.3%,  SVM accuracy = 94%,  Random forest accuracy = 92.6%,  AdaBoost accuracy = 82%,  Bagging accuracy = 89% |
| 26. | Fake News Classification using transformer based enhanced LSTM  and BERT | 2022 | FakeNewsNet dataset |  | BERT, LSTM |  |  |  |
| 27. | Development of Fake News Model Using Machine  Learning through Natural Language Processing | 2022 | Fake news articles were collected from an open  source Kaggle dataset [33] that was published during the 2016  election cycle | Count Vectorizer, TFIDFVectorizer | RapidMiner, Naïve Bayes, Support vector machine, Passive agressive |  | Our proposed combination works  well and obtains performance above the baseline 0.50 | Naïve bayes accuracy = 85%, Passive aggressive accuracy = 93%,  SVM accuracy = 84% |
| 28. | Classifying Fake News Articles Using Natural  Language Processing to Identify In-Article  Attribution as a Supervised Learning Estimator | 2020 | ISOT fake news and LIAR |  | SVM, LSVM, Decision tree,  KNN, Sgd,  Linear Regression |  | The best result of the model is  achieved by using history as metadata. The results show that  this model can perform better than state-of-the-art baselines,  including hybrid CNN and LSTM with attention by 3.1%  on the validation set and 1% on the test set | SVM accuracy = 86%, LSVM accuracy = 92%,  KNN accuracy = 83%,  Decision tree accuracy = 89%,  SGD accuracy = 89%,  Linear regression accuracy = 89% |
| 29. | Natural Language Processing based Online Fake  News Detection Challenges – A Detailed Review | 2020 | Cover news  stories from  major news  outlets | Ti-idf, Bag of words | Naïve Bayes  ,Logistic Regression, Decision Tree,  Support Vector Machine,  EnsembleMethod, Random Forest, XG-Boost |  |  | Total accuracy = 81%, |
| 30. | Effect of Corpora on Classification of Fake News using  Naive Bayes Classifier | 2020 | true and false news from  the train and test data collected from the Kaggle dataset | Count vectorizer, TFIDF vectorizer | Naive Bayes classifier |  | Naive Bayes is one of the prominent Machine Learning Algorithms to solve text classification  problem. This algorithm is suitable when the dimensionality of the inputs is high and still the  method is relatively simple |  |
| 31. | Fake News Detection: A Deep Learning Approach | 2018 | Emergent Dataset created by  Craig Silverman | Bag of words, Tf-Idf, GloVe,  Word2Vec | Convolutional Neural Networks (CNN), Recurrent neural networks (RNN) |  | Tf-Idf word vector representations when propagated through a  dense neural network gives an accuracy of 94.21% |  |
| 32. | Deep neural approach to Fake-News identification | 2020 | George McIntire dataset | GloVe word embeddings | LSTM and FNN |  | When it came to LSTM model using word2vec representation, we could see a significant increase of 8% in accuracy on average for the model  with mined features than the one without. | The LSTM model without mined secondary features performed similar to its  FNN counterpart with accuracy ranging from 82-84%. But the LSTM model with additional mined features showed  significantly improved accuracy in the range of of 91-94% |
| 33. | Attention-based C-BiLSTM for fake news detection | 2021 | Social media datsets |  | Convolutional Bidirectional Long Short-Term Memory (AC-BiLSTM) model |  | we present the performance of the proposed AC-BiLSTM model using the standard evaluation metrics such as confusion matrix, accuracy, precision, recall, F1-score, macro average, micro average, and weighted average [ |  |
| 34. | The COVID-19 fake news detection in Thai social texts | 2021 | iSAI-NLP, PRICAI |  | SVM, MLP, Naïve Bayes |  |  |  |
| 35. | Fake News Detection Using a Deep Neural Network | 2018 | We commonly collect various datasets from  https://www.kaggle.com | Tf-Idf | Neural Network, SVM, Naive Bayes, Random Forest, Decision tree, KNN |  |  | Accuracy: 96% |
| 36. | Pakistani Media Fake News Classification using  Machine Learning Classifiers | 2019 | Fake and True News from online available news websites, Fake News has  collected from well-known News agencies in Pakistan. | Bag of Words (BoW) and Term Frequency  Inverse Document Frequency (TF-IDF) | Logistic regression (LR), Support Vector  Machine (SVM), K-Nearest Neighbour (KNN), Random  Forest Classifier (RFC), Naïve Bayes (NB) and Decision  Tree Classifier (DTC). |  | Experiment results shows linear support vector machine  (LSVM) gives 92% accuracy with term frequency-inverse  document frequency (TF-IDF) as finest text vectorization  technique | Logistic Regression precision = 0.69,  Naïve Bayes precision = 0.68,  KNN precision = 0.64,  SVM (Linear kernel) precision = 0.68, SVM (sigmoid kernel) precision = 0.22,  Random forest precision = 0.65, Decision tree precision = 0.63. |
| 37. | Fake News Detection Using Deep Learning  Techniques | 2019 | Dataset of news is  taken from online. |  | Logistic regression, Random Forest, Deep neural network, SVM, Naïve Bayes |  | In  proposed system LR, RF, SVM NB and DNN classification  techniques are utilized that will help to detect fake news.  Classification techniques like LR, RF, SVM NB and DNN for  feature selection and extraction utilized, DNN will work fine  in execution time and accuracy cases but it needs large  memory than other | LR accuracy = 75%, RF accuracy = 77%, SVM accuracy = 79%,  NB accuracy = 89%,  DNN accuracy = 91%. |
| 38. | Fake News Detection using Machine Learning  Algorithms | 2021 | LIAR | Bag-Of-Words, N-Grams, Tf-Idf | Naïve Bayes classifier, Random Forest, Logistic Regression, Passive Aggressive Classifier |  | Implementation was done using the above algorithms  with Vector features- Count Vectors and Tf-Idf vectors  at Word level and Ngram-level. Accuracy was noted  for all models. We used K-fold cross validation  technique to improve the effectiveness of the models. | Naïve Bayes precision = 0.59, Random Forest precision = 0.62, Logistic Regression precision = 0.69 |
| 39. | A Review of Natural Language Processing  Techniques for Sentiment Analysis using  Pre-trained Models | 2020 |  | Word2Vec  and GloVe | ULMFiT, Transforme, OpenAI’s GPT-2, BiGRU, Google’s BERT, Transformer-XL, XLNet |  | Extracting polarity from dataset in Sentiment Analysis is a  challenging phase. Pre-trained models can achieve this task  easily without the need to build a model from scratch. |  |
| 40. | Sentiment Analysis of Multilingual Twitter Data using Natural Language  Processing | 2018 | Twitter database | Bag of words | Naive Bayes (NB), RNN |  | It has been cleared from the observations  that RNN classifier has significant ahead from the other  classifier in the task of predicting the sentiments with almost  an accuracy of 96% | RNN accuracy = 96.15%,  Naïve Bayes accuracy = 77.16% |
| 41. | The Effects of Natural Language Processing on Big  Data Analysis: Sentiment Analysis Case Study | 2018 | Cornell University movie review, Amazon movie review data set | Tf-Idf | Naïve Bayes Classification |  | Although, the running time has been increased, the  integration of the NLP and linguistic preprocessing in the  classification process achieved improvements in the  classification accuracy. Using PoS tagging, adjectives were  specified and assigned more weight than other terms in the  tweets | Accuracy = 70.2% |
| 42. | Sentiment Analysis Using Support Vector  Machine | 2014 | Comment database | TF-IDF | Support vector machine |  |  |  |
| 43. | Overview of the GermEval 2021 Shared Task on the  Identification of Toxic, Engaging, and Fact-Claiming Comments | 2021 | We manually annotated a dataset of more than  4,000 Facebook user comments, which is drawn  from the Facebook page of a German political talk  show of a national public television broadcaster |  | Support Vector Machines, Logistic Re-regression, Forest, CNN, GRU, or LSTM |  | It provides summary statistics on the macro-  average F1-score, which is the metric that was used  as the official ranking criterion in the shared task |  |
| 44. | A Literature Review of NLP  Approaches to Fake News  Detection and Their  Applicability to Romanian-  Language News Analysis | 2020 | LIAR dataset |  | Claim Buster, Multinomial Naïve Bayes, SVM, Random Forest |  | In this paper, we highlighted the impact of fake news at  a social level and examined the challenges faced when  combating this phenomenon, which can be successfully  addressed by devising automated detection methods |  |
| 45. | Fake News Detection using Deep Learning | 2020 | News dataset can be collected from many different  sources such as Kaggle, UCI Machine Learning and more | N-gram | In the Keras neural network model, the layers 1,  3 and 5 are using rectified linear unit (RELU) as the  activation function with 64 nodes, 16 nodes and 2 nodes  respectively. In between these 3 layers, 2 dropout layers of  20% dropout rate are also added to the network. The  purpose of adding dropout layers is used to avoid from  overfitting by dropping some node and therefore,  generalize better |  |  | Accuracy = 90.3% |
| 46. | Identification of Fake News Using Machine  Learning | 2020 | FakeNewsAMT, | Tf-Idf | Deep neural network, Naïve Bayes, Passive agressive |  | t is observed that the JRU dataset with only  three attributes has performed quite well with almost 100  percent accuracy. While one could chalk up the success of  the dataset to its small size, | Naïve Bayes accuracy = 96%,  Passive aggressive accuracy = 100%,  DNN accuracy = 100% |
| 47. | Fake News Detection in social media using Graph Neural  Networks and NLP Techniques: A COVID-19 Use-case | 2020 |  | Bag of words | Naïve Bayes classifier |  | For TMD, we submitted six different runs mainly relying on two  approaches, namely BERT and BoW, under two late fusion schemes.  Three of the runs are based on binary classification while the three  deal the task as ternary classification problem |  |
| 48. | Thai Fake News Detection Based on Information Retrieval, Natural Language Processing and Machine Learning | 2021 | WWW, Google, YouTube, Google+, Facebook, Twitter, Instagram datasets |  | CNN-LSTM, Naïve Baye, Decision Tree, and Support Vector Machine, logistic regression, KNN |  | Fake news data are very dynamics. It is not an easy task to build a fake news detection system that generalizes all unseen data. Our idea is to exploit the news data on the Internet and social media by using it as inputs fed to the classifier |  |
| 49. | Evaluating Pretrained Transformer-based Models for  COVID-19 Fake News Detection | 2021 | Online dataset |  | BERT, DistilBERT, ALBERT |  | we observes that domain-specific language  models perform better than their counterparts pretrained on the  general domain as they address data tailored to a particular  domain (COVID-19) |  |
| 50. | Fake news detection: A hybrid CNN-RNN based deep learning approach | 2021 | CREDBANK, FA-KES dataset |  | KNN, RNN, LSTM, CNN, RF, MNB, SGD, DT, CNN, CNN + RNN |  | Decision Tree classifier, the accuracy of their methods ranges between 0.85 and 1.00. It also seems that they used only part of the ISOT dataset, since they reported that only 25.2 thousand articles were used. Taking into account the statistical significance of the results (0.99 0.02), the proposed method performs comparatively better than the state-of-the-art method | LR accuracy = 0.49,  RF accuracy = 0.53, MNB accuracy = 0.38, SGD accuracy = 0.47, KNN accuracy = 0.57, DT accuracy = 0.55, AB accuracy = 0.47, CNN accuracy = 0.50, RNN accuracy = 0.50, CNN + RNN accuracy = 0.60 |
| 51. | Fake News classification using Bi-directional LSTM-Recurrent  Neural Network | 2021 | ISOT Fake News Dataset, IEEE Fake News Dataset | Global Vectors (GloVe), TF-IDF | ConvNet (CNN), Recurrent Neural Network (RNN), Bi-LSTM |  | the test set is  directly used to define the accuracy whereas the training and validation set undergo Bi-  LSTM-RNN layer and the model is trained recursively in a loop to reduce loss function  and to improve accuracy. Then the hyper parameter value is used to control the learning  process and finally, the accuracy is calculated |  |
| 52. | A Sensitive Stylistic Approach to Identify  Fake News on Social Networking | 2020 | collection of tweets | Tf-Idf | SVM |  | In the results of the reduction methodology with training,  the proposal demonstrates a more homogeneous performance  among the metrics, standing out mainly for the high accuracy  and more significant percentage of sensitivity among the three  methodologies |  |
| 53. | An Efficient FAKE NEWS DETECTOR | 2020 | Fake news dataset |  | NAIVE BAYES, SVM, Random Forest |  | The experimental study of a dataset carried out on  each classifier independently and then applied on  proposed model and results obtained | SVM accuracy = 57%, Naïve Bayes accuracy = 60%, Random Forest accuracy = 63%. |
| 54. | New explain ability method for BERT-based model in fake news detection | 2021 | Dataset from Kaggle.com |  | BERT-bidirectional encoder representations from transformers |  | Bidirectional LSTM has an accuracy of 92%. The precision for real news is 94%, while for fake news it is equal to 90%. The recall equal to 92% is the same for both categories. The f1-score achieved for the real news classes is 93% and 91% for fake news. | BERT-based fake news detector has accuracy of 98%. Precision for real news is 97%, while for fake news it is equal to 99%. The recall is reversed, with the 99% score for the real news and the 97% for fake news, while f1-score for both classes is 98% |
| 55. | Fake News Detection on social media using  Geometric Deep Learning | 2019 | Snopes, PolitiFact, BuzzFeed datasets |  | We used a four-layer Graph CNN with two convolutional  layers (64-dimensional output features map in each) and two fully connected layers (producing 32-  and 2-dimensional output features, respectively) to predict the fake/true class probabilities |  | We considered two different settings of fake news detection: URL-wise and cascade-wise, using the  same architecture for both settings. In the first setting, we attempted to predict the true/fake label of  a URL containing a news story from all the Twitter cascades it generated. On average, each URL  resulted in approximately 141 cascades |  |
| 56. | A Tool for Fake News Detection | 2018 | Online fake news dataset | Bag-of-Words model, N-gram model, erm Frequency-Inverse Document Frequency mode | Naive Bayes Classifier, Linear Support Vector Machines Classifier. |  | the obtained results by com-  bining each classifier with every text representation model.  To compute the prediction accuracy of a classifier, we used  the metrics class from Scikit-learn library. Then we  represented these results as a ROC Curve. The accuracy is  calculated based on the test data we had when we split the  dataset to train data and test data |  |
| 57. | Social media sentiment analysis based on COVID-19 | 2020 | Tweeter datasets |  | Support Vector Machine, Recurrent Neural Network |  | For the RNN model again, the neutral result is 0% and only 8.60% is weakly negative | we use a Recurrent Neural Network (RNN) to classify emotions on tweets. We developed a model to analyse the emotional nature of various tweets, using the recurrent neural network for emotional prediction, searching for connections between words, andmarking them with positive or negative emotions. |
| 58. | Sentiment Analysis on Interview Transcripts: An  application of NLP for Quantitative Analysis | 2018 | e conducted 18 interviews across five industries  covering nine females and nine males. The dataset consists of  eight senior managers, two HR managers and eight millennial-  als. Table I represents the distribution of interviews over the  industries, gender and position held in the organisation. The  questions and their answers are available in the dataset |  | Linear Model |  | We developed an innovative technique to apply NLP to de-  rive the average sentiment of interview transcripts. We demon-  strated the use of quantitative analysis to derive meaningful  insights from the sentiment analysis leading to hypothesis  validation and building a linear model. We built a linear model  with two simple variables to explain 72% |  |
| 59. | The Impact of Features Extraction on the Sentiment Analysis | 2019 | SS-Tweet dataset | Tf-Idf, N-gram | SVM, Linear Regression, Naïve Bayes, Random Forest, Decision tree, KNN |  | In this paper, we considered two features TF-IDF (word level) and N-Grams (value of n=2) on the Twitter sentiment  analysis dataset (SS-Tweet). Table 1 shows the output(four performance parameters i.e accuracy, precision, recall,  and f-score) of six classification techniques (Random Forest, Decision Tree, Naive Bayes, SVM, Logistic  Regression, and KNN) using TF-IDF feature. | KNN accuracy = 46%, Decision Tree accuracy = 46%, SVM accuracy = 46%, LR accuracy = 57%, Naïve Bayes accuracy = 53%, Random Forest accuracy = 51%. |
| 60. | Sentiment Analysis Based on Dictionary Approach | 2015 | Twitter datasets | Bag of words | Naïve Bayes |  | In order to validate the analysis are performed. It is necessary to employ quantitative measures and  qualitative measures. After such a validation it may be necessary to return to one or more of the  previous step so as to perform modification and try alternatives. |  |
| 61. | Two Stage Transformer Model for COVID-19 Fake News Detection and  Fact Checking | 2020 | fake news on COVID-  19. | Tf-Idf, GloVe | BERT, ALBERT, Mobile BERT |  | We find that a combination of the  best performing Model A (BERT) and best performing Model B (ALBERT) yielded the highest MRR,  Recall@10 and Accuracy on the test set for fact checking | BERT accuracy = 0.810, ALBERT accuracy = 0.825 |
| 62. | Hostility Detection and Covid-19 Fake News  Detection in social media | 2021 | Fake news dataset | TF-IDF | BERT, SVM |  | As an individual model, Hindi BERT Embedding with BERT  model yields the best result for coarse-grain classification and FastText-LSTM  with metadata yields best result on fine-grain classification on test data. | SVM accuracy = 93.59%, BERT accuracy = 93.74 |
| 63. | Detection of Fake News Text Classification on COVID-19 Using  Deep Learning Approaches | 2021 | COVID Fake News Dataset | TF-IDF | SVM, Logistic Regression, Naïve Bayes, AdaBoost, KNN, Decision tree, Random Forest, Multilayer perceptron, CNN, RNN, LSTM |  |  | LR accuracy = 96%, RF accuracy = 97%, DT accuracy = 96%, SVM accuracy= 96%, KNN accuracy = 97%, AdaBoost accuracy = 96%, Naïve Bayes accuracy = 95% |
| 64. | Aspect Based Sentiment Analysis using Support Vector Machine Classifier | 2013 | he data set chosen for the work is a collection  of reviews of digital cameras. A review set consisting of  around 2500 user reviews from Amazon.in, ebay.in  and other popular sites is used for training and feature  identification purposes. The reviews of Canon A2300,  Canon Powershot SX260, Sony DSC-RX100, Nikon D5100  and Canon A70 were used for training purpose. |  | Support vector machines (SVM) |  | The result of the work is a quantitative analysis  conveniently depicted using graphical plotting utilities. The  result of the sentiment analysis involves the polarity orien-  tations of each aspect of the product | Accuracy = 77.98%. |
| 65. | FEATURE SELECTION AND  CLASSIFICATION APPROACH FOR  SENTIMENT ANALYSIS | 2015 | polaritydatasetV2.0 | TF-IDF | Naïve Bayes, SVM |  | The dataset used for the experiments was divided into two classes, positive and negative. For a  given classifier and a document there are four possible outcomes: true positive, false positive, true  negative and false negative. If the document is labelled positive and is classified as positive it is  counted as true positive else if it is classified as negative it is counted false negative |  |
| 66. | Sentiment Analysis and Classification of  Restaurant Reviews using Machine Learning | 2020 | The data used in this research was collected from a  social media community named as “SWOT’s Guide to  KARACHI’s Restaurants Cafes Dhabas HBFE &  Takeouts”. |  | NLP, SVM |  | his paper presents a framework for Pakistani restaurant  customer reviews sentiment analysis and category  classification of reviews. For this purpose, the dataset has  been created by collecting users’ posts from social media  group “SWOT” which contains customer reviews regarding  different Pakistani restaurants |  |
| 67. | Comparing deep learning architectures for sentiment analysis on drug reviews | 2020 | drug reviews taken from the Drugs.com website. |  | CNN, LSTM |  | To evaluate our systems, we have used the standard metrics for text classification tasks: precision, recall and F1. These metrics can be extended to multi-classification problems using the micro-averaged and macro-averaged versions |  |
| 68. | Sentiment Analysis and Product Review  Classification in E-commerce Platform | 2020 | We have collected product reviews from various e-commerce  sites namely Daraz, BDshop, and Evally which are the topmost  online shops in Bangladesh. | Bag of words | DNN |  | . Firstly, we have  generated the learning curves to analyze training and validation  performance which has helped us to determine if there is any  underfitting or overfitting issue. And then, we have performed  a prediction on test dataset to analyze the performance of the  model in reality. | Accuracy = 70% |
| 69. | COVID-Fact: Fact Extraction and Verification of Real-World Claims on  COVID-19 Pandemic | 2021 | FEVER like dataset |  | SBERT |  | Our pipeline provides a strong  baseline with F1 score of ≈ 32. For comparison,  the baseline system in FEVER (Thorne et al., 2019)  achieves the F1 score of 18.26. Note Top 5 evi-  dence retrieval performs worse than gold since we  evaluate how the system performs with automati-  cally negated claims as well, for which we re-run  the Google+SBERT metho |  |
| 70. | Fake news detection using parallel BERT deep neural  networks | 2022 | Online fake news dataset |  | BERT, CNN |  | We then encoded them using two parallel BERT networks and fed them into a  linear classifier. The results showed that the choices made were the right choices.  To get better results in fake news detection using deep neural networks, what we need most  is more data. With a larger dataset, we can also use the larger transformer models, which can  learn better with more parameters. | Accuracy = 0.854 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |