|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Sr no | Year | Author | Technique | Tool | Result | Conclusion /future scope | Link /DOI |
| 1 | 27 Jun 2022 | Pooja Malhotra  Sanjay Kumar Malik | Spam detection using deep learning | NLP, LSTM  Bi-LSTM | Accuracy-98.5  F1 score -96 | work can be extended by implementing the classifier in real time. | <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4145123> |
| 2 | March 2020 | Vinodhini. M, Prithvi. D, Balaji. S | Spam Detection Framework using ML Algorithm | Machine Learning, Random Forest algorithm, Reviews | Accuracy-  0.9534 | Work can be improved further using  Deep learning | <https://www.ijrte.org/wp-content/uploads/papers/v8i6/F1120038620.pdf> |
| 3 | Nov 2022 | M. Hema Kalyan, 2M. Hari Krishna | Spam E-Mail Detection Using Machine Learning Algorithms | Naïve Bayes, Support vector machines (SVM), Datasets. | Accuracy  0.9623 | Naïve Bayes classifier outperforms all other classifiers | <https://ijrpr.com/uploads/V3ISSUE11/IJRPR7780.pdf> |
| 4 | October 13, 2020 | SIMRAN GIBSON 1 , BIJU ISSAC 1 , LI ZHANG | Detecting Spam Email With Machine Learning Optimized With Bio-Inspired Metaheuristic Algorithms | bio-inspired algorithms, cross-validation, particle swarm optimization, genetic algorithm. | Accuracy  0.9842 | Genetic Algorithm worked better overall for both text-based datasets and numerical-based datasets than PSO | <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9222163> |
| 5 | Apr 2021 | Manoj Sethi1, Sumesha Chandra Vinayak Chaudhary, Yash | Spam Detection using Machine Learning and Neural Networks | Neural Networks, Naive Bayes, Support Vector Classifier, Logistic Regression | Accuracy  .9902 | future work includes testing the model with various standard datasets | <https://www.irjet.net/archives/V8/i4/IRJET-V8I472.pdf> |
| 6 | 30th April 2024 | Rajesh Kumar J, Sudarshan P, Mahalakshmi G | Email Spam Detection using Machine Learning Techniques | Naive Bayes, Support Vector Machine, Natural Language Processing, analysis. | Accuracy  0.97 | Naive Bayes classifiers are much better than that of the SVM for text classification. | <https://iarjset.com/wp-content/uploads/2021/06/IARJSET.2021.8632.pdf> |
| 7 | 6 Jun 2022 | Thaer Sahmoud, Dr. Mohammad Mikki | Spam Detection Using BERT | , BERT, Machine learning, NLP, Transformer, Enron Corpus, SpamAssassin Corpus, | Accuracy  .9928 | high performance of our model is a result of using BERT pre-trained model | <https://arxiv.org/ftp/arxiv/papers/2206/2206.02443.pdf> |
| 8 | 27 Aug  2020 | [Razan Ghanem](https://link.springer.com/article/10.1007/s42452-020-03374-x#auth-Razan-Ghanem-Aff1) &  [Hasan Erbay](https://link.springer.com/article/10.1007/s42452-020-03374-x#auth-Hasan-Erbay-Aff2) | Context-dependent model for spam detection on social networks | BERT,  CNN | Accuracy  .981 | intend to improve the architecture of the proposed model and conduct an empirical study to tune the hyperparameters | <https://link.springer.com/article/10.1007/s42452-020-03374-x#Abs1> |
| 9 | 5 may 2022 | Sonali Kotni1 , Dr. Chandrasekhar Potala2 , Dr. Laxman Sahoo3 | SPAM DETECTION USING DEEP LEARNING MODELS | BERT, DISTILBERT,  ALBERT,  NLP | Accuracy  .9472 | The model performance can be increased by using more data pre-processing techniques. | <https://iaeme.com/MasterAdmin/Journal_uploads/IJARET/VOLUME_13_ISSUE_5/IJARET_13_05_006.pdf> |
| 10 | Jan 2022 | Zeynab Fallah Sokhangoee , Abdoreza Rezapour | A novel approach for spam detection based on association rule mining and genetic algorithm | Genetic algorithm  And rule mining | Accuracy  nil | focused on detecting spam and non-spam on social networks | <https://www.sciencedirect.com/science/article/abs/pii/S0045790621005796?via%3Dihub> |
| 11 | July 2023 | Abishek Sharma, Arjun N | Spam Detection Using Machine Learning Techniques | Support vector Machine Learning, Natural Language Processing | Accuracy  .9558 | use of reputable and verified domain names can be used to identify scam emails | <https://ijrpr.com/uploads/V4ISSUE7/IJRPR15572.pdf> |
| 12 | 2013 | Houshmand Shirani-Mehr | SMS Spam Detection using Machine Learning Approach | Naïve Bayes, SVM, KNN | Accuracy  .9888 | multinomial naive Bayes with laplace smoothing and SVM with linear kernel are among the best classifiers for SMS spam detection | <https://cs229.stanford.edu/proj2013/ShiraniMehr-SMSSpamDetectionUsingMachineLearningApproach.pdf> |
| 13 | 1, July 2018 | E. G. Dada and S. B. Joseph | Random Forests Machine Learning Technique for Email Spam Filtering | Random Forests Machine  Learning | F1 score  0.998  Accuracy  0.999 | RFs is a promising algorithm that can be adopted either at mail server or at mail client side to further decrease the volume of spam messages in email users inbox. | <https://www.researchgate.net/profile/Emmanuel-Dada/publication/327720311_Random_Forests_Machine_Learning_Technique_for_Email_Spam_Filtering/links/5ba0b0c0299bf13e6038ecc1/Random-Forests-Machine-Learning-Technique-for-Email-Spam-Filtering.pdf> |
| 14 | 2021 | p.U. Anitha, dr.C.V. Guru Rao, dr. D. Suresh Babu | Email Spam Filtering Using Machine Learning Based Xgboost Classifier Method | XGBoosting,  Machine learning | Accuracy  .95 | results confirm that the proposed model got better accuracy with 95% compared with the current approaches. | <https://www.proquest.com/openview/9b40fcf6cc3c713db6112dc54ce40e84/1?pq-origsite=gscholar&cbl=2045096> |
| 15 | 2019 | Diksha S. Jawale,  Ashwini G. Mahajan | Hybrid spam detection using machine learning | Naive Bayes, Support vector machine, SVM | Accuracy  .9944 | SVM has great accuracy but slow classification speed and require more dataset and NB has fast classification speed but having low accuracy and requires small dataset | <https://link.springer.com/chapter/10.1007/978-981-13-0761-4_63> |
| 16 | 2010 | V.Christina# , S.Karpagavalli\* , G.Suganya# | Email Spam Filtering using Supervised Machine Learning Techniques | MultiLayer Perceptron,  Decision Tree Induction,  Naïve Bayes Classification | Accuracy  .993 | Email spam filters using this approach can be adopted either at mailserver or at mail client side to reduce the amount of spam messages and to reduce the risk of productivity loss | <https://www.researchgate.net/publication/50235326_Email_Spam_Filtering_using_Supervised_Machine_Learning_Techniques> |
| 17 | 2017 | Priti Sharma,Uma Bhardwaj | Spam E-Mail Detection using Hybrid bagged Approach | Naïve bayes classifier,  J48 classifier, | Accuracy  0.875  F1-score  0.873 | To enhance the system’s performance and results, the concept of boosting approach could be considered for future work | <https://web.archive.org/web/20220320073257id_/http://www.inass.org/2018/2018063001.pdf> |
| 18 | 2017 | Lutfun Nahar Lota,  B M Mainul Hossain | Spam Detection using Bayesian algorithm | Tokenization,  Bayesian classifier | Accuracy  0.99 | scope of further research in this filed | <https://www.researchgate.net/publication/318298908_A_Systematic_Literature_Review_on_SMS_Spam_Detection_Techniques> |
| 19 | 2017 | Nurul Fitriah Rusland, Norfaradilla Wahid, Shahreen Kasim, Hanayanti Hafit | Naıve Bayes Algorithm for Email Spam Filtering across Multiple Datasets | Boosting algorithm, Support Vector Machines (SVM) algorithm [5] and Naıve Bayes algorithm | Accuracy  0.913 | quality of performance Na¨ıve Bayes classifier is also based on datasets that used. | <https://iopscience.iop.org/article/10.1088/1757-899X/226/1/012091/pdf> |
| 20 | 2014 | Sarwat Nizamani, Nasrullah MemonMathies Glasdama , Dong Duong Nguyen | Detection of fraudulent emails by employing advanced feature abundance | J48,SVM,Naïve Bayes,  CCM | Accuracy  0.96 | plan to employ header information of the email for the task of fraudulent email detection task. | <https://www.sciencedirect.com/science/article/pii/S1110866514000280> |
| 21 | 2022 | Jenifer Darling Rosita P a , W. Stalin Jacob | CNN-Based Deep Learning Architectural Scheme for effective spam detection | MOGA–CNN–DLAS,  ML-TSDS | Accuracy  0.95 | MOGA–CNN–DLAS scheme is proposed for superior detection of Twitter spam | <https://www.sciencedirect.com/science/article/pii/S266660302200001X> |
| 22 | 2016 | Hedieh Sajedi1, Golazin Zarghami Parast1 , | SMS Spam Filtering Using Machine Learning Techniques | SVM,KNN,DCA | Accuracy  0.982 | DCA algorithm is the best one as it develops the performance | <https://www.sciencepublishinggroup.com/article/10.11648/j.mlr.20160101.11> |
| 23 | 2018 | Samadhan Nagre | Mobile SMS Spam Detection using Classifiers | Decision Tree,Random Forest,Regression | Accuracy  0.93 | focused on detecting spam and non-spam on various websites | <https://www.jetir.org/papers/JETIR1812A76.pdf> |
| 24 | 2020 | Abdallah Ghourabi , Mahmood A. Mahmood 1,3 and Qusay M. Alzubi | A Hybrid CNN-LSTM Model for SMS Spam Detection | deep learning; CNN; LSTM | Accuracy  0.9837 | As future work, we plan to create a rich framework capable of filtering spam messages in smartphones | <https://www.mdpi.com/1999-5903/12/9/156> |
| 25 | 2018 | G.Vijayasekaran[1], S.Rosi | SPAM AND EMAIL DETECTION IN BIG DATA PLATFORM USING NAIVES BAYESIAN CLASSIFIER | Data mining, Classifier,  Secure Hash Algorithm | Accuracy  0.973 | . In the future work we have a plan to implement other algorithm to our classification method to achieve better performance. | <https://ijcsmc.com/docs/papers/April2018/V7I4201813.pdf> |
| 26 | 2016 | Mohammed Awad and Monir Foqaha2 | EMAIL SPAM CLASSIFICATION USING HYBRID APPROACH OF RBF NEURAL NETWORK AND PARTICLE SWARM OPTIMIZATION | Radial Basis Function Neural Networks, Particles Swarm Optimization | Accuracy  0.843 | proposed a hybrid approach (HC-RBFPSO), that combining RBFNN and PSO in the purpose to classifyEmail spam problems | <https://aircconline.com/ijnsa/V8N4/8416ijnsa02.pdf> |
| 27 | 2013 | Megha Rathi,  Vikas Pareek | Spam Mail Detection through Data Mining – A Comparative Performance Analysis | SVM,  Naïve Bayes,  Decision Tree | Accuracy  0.9971 | we achieve highest accuracy = 99.715%. it is very difficult to achieve 100% accuracy but Random Tree and Random Forest (accuracy>99% | <https://www.researchgate.net/publication/269651895_Spam_Mail_Detection_through_Data_Mining_-_A_Comparative_Performance_Analysis> |
| 28 | 2010 | Tiago A. Almeida Jurandy Almeida Akebo Yamakam | Spam filtering: how the dimensionality reduction affects the accuracy of Naive Bayes classifiers | Naive Bayes spam filters,  Bayesian Classifiers | Accuracy  0.95 | Future works should take into consideration that spam filtering is a co-evolutionary problem, because while the filter tries to evolve its prediction capacity | <https://jisajournal.springeropen.com/articles/10.1007/s13174-010-0014-7> |
| 29 | 2023 | Mrs. Anitha Reddy Kanthaa Harivardhan Reddy , Abhishek | Email Spam Detection Using MachineLearning | NLP,Naïve Bayes,KNN | Accuracy  0.99 | useof ML and NLP for email spam classification can save users valuable time and resources and improve the overall productivity andsecurity of email communication | <https://ieeexplore.ieee.org/document/10170187> |
| 30 | 2022 | Prof. Prachi Nilekar, Tamboli Abdul Salam, Manish Kumar Gupta, Krishna Sharma, Safwan Attar | Email Spam Detection Using Boosting | Random Forest,Decision Tree | Accuracy  0.97 | fixes any false data about the users that they may have. Itis a benefit for the users’ who’s important time | <https://www.irjet.net/archives/V9/i11/IRJET-V9I11154.pdf> |
| 31 | 2021 | Isra’a AbdulNabi , Qussai Yaseen | Spam Mail Detection Using Deeplearning | BERT, Sklearn, BiLSTM, Classifiers | Accuracy  0.973 | Can be used in Enhancing Email Security | <https://www.sciencedirect.com/science/article/pii/S1877050921007493#section-cited-by> |
| 32 | 2029 | Yong Fang  Cheng Zhang; Cheng Huang | Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism | SciKit,NLTKBeautifulSoupJupyter Notebook | Accuracy  0.971 | safeguarding digital communications and thwarting cyber threats. | <https://ieeexplore.ieee.org/abstract/document/8701426> |
| 33 | 9 sep  2013 | Yukti Kesharwani,  Shrikant Lade | Spam Mail Filtering Through Data Mining Approach – | Artificial neural network, Chaid, quest,  SVM,  Pandas. | Accuracy  0.968 | Identify and mitigate threats effectively. | <https://www.ijert.org/research/spam-mail-filtering-through-data-mining-approach-a-comparative-performance-analysis-IJERTV2IS90657.pdf> |
| 34 | 2018 | Sami Smadi, Nauman Aslam, Li Zhang | Detection of online phishing email using dynamic evolving neural network based on reinforcement learning | reinforcement model,NLP,Pandas, numpy | Accuracy  0.992 | This paper shows how a NN with RL can be used to build a powerful model to detect zero-day phishing attacks | <https://www.sciencedirect.com/science/article/abs/pii/S0167923618300010> |
| 35 | Jul 2013 | Sumant Sharma, Amit Arora | Adaptive Approach for Spam Detection | Machine Learning, NumericToBinary Filter, Weka | Accuracy  94.28% | Proposed various ML algorithms for spam detection in emails using SPAMBASE dataset and WEKA. | <https://ijcsi.org/papers/IJCSI-10-4-1-23-26.pdf> |
| 36 | Nov 2010 | V.Christina , S.Karpagavalli, G.Suganya | Email Spam Filtering using Supervised Machine Learning Techniques | Decision Tree, Multilayer Perceptron, Naïve Bayes | Predictive accuracy: 99.3%,  False Positive rate: 1% | Adopt ensemble methods like SVM and Random Forest with deep learning architectures to further improve email spam classification. | <https://www.researchgate.net/publication/50235326_Email_Spam_Filtering_using_Supervised_Machine_Learning_Techniques/fulltext/53ec94460cf2233164944f71/Email-Spam-Filtering-using-Supervised-Machine-Learning-Techniques.pdf> |
| 37 | 3 march 2014 | Sarju S, Riju Thomas, Emilin Shyni C | Spam Email Detection using Structural Features | Apache James Mime4, Machine Learning Algorithms (Naïve Bayes, Random Forest, AdaBoost) | Accuracy : 99.4 %  False positive rate : 0.6% | Further research could expand the methodology to handle various email formats and attachments for comprehensive spam detection. | <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=16f6f6da1ddc7bc3ad7b0bb1bbe1779abc8ab60a> |
| 38 | 2007 | J. Shobana and D. Kanchana | Multinomial Naïve Bayes model Using Passive Aggressive Algorithm | TD-IDF, Multinomial Naïve Bayes, Passive Aggressive Algorithm | Accuracy-89% | Future research may explore the integration of deep learning techniques for improved classification performance. | <https://iopscience.iop.org/article/10.1088/1742-6596/2007/1/012047/pdf> |
| 39 | 2009 | Mehran Sahami, Susan Dumaisy, David Heckermany,  Eric Horvitzy | A Bayesian Approach to Filtering Spam E-Mail | Bayesian Networks, Naive Bayesian Classifier, Vector Space Model, Feature Selection Techniques, | Accuracy :  95.5% | Proposed approach effectively addresses the problem of automatic image captioning, achieving competitive results compared to existing methods. | <http://robotics.stanford.edu/users/sahami/papers-dir/spam.pdf> |
| 40 |  | Ruhul Amin, Md. Moshiur Rahman, and Nahid Hossain | A Bangla Spam Email Detection and Datasets Creation Approach based on Machine Learning Algorithms | Multinomial Naive Bayes (MNB), KNN, Decision Tree, Support Vector Machine (SVM), Random Forest, and AdaBoost | Accuracy :  93.60 % | Used only generic machine learning algorithms. Can use Deep Learning (DL) approaches such as Deep Neural Networks to make the system more robust. | <https://nahid.org/papers/c4.pdf> |
| 41 | 3 June 2016 | Alexy Bhowmick, Shyamanta, M. Hazarika | Machine Learning for E-mail Spam Filtering | Naïve Bayes,  SVM,  Ensemble Classiﬁers,  Clustering Techniques | Best Accuracy:  96.69% | Future research directions include addressing concept drift, reducing false positives, tackling emerging spam threats like snowshoe spam and social network-based spam. | <https://www.researchgate.net/publication/303812063_Machine_Learning_for_E-mail_Spam_Filtering_ReviewTechniques_and_Trends> |
| 42 | 1 Feb  2011 | W.A. Awad and S.M. Elseuofi | Machine Learning Methods for Spam Email Classification | Artificial Neural Network classifier,  Artificial Immune System, Naïve Bayes,  KNN | Accuracy:  96.23% | promising results for less common algorithms like Naïve Bayes and rough sets. Hybrid systems are identified as the most promising approach to enhance the performance of anti-spam filters | <https://airccse.org/journal/jcsit/0211ijcsit12.pdf> |
| 43 | 2020 | N. Sutta,  Z. Liu,  X. Zhang | A Study of Machine Learning Algorithms on Email Spam Classification | Nu Svm, ANN,  KNN,Random Forest, Decision Tree,  Logistic regration | Accuracy:  98.2% | The study evaluates seven machine learning algorithms for spam email classification, highlighting the efficacy of N-Grams and combined datasets in improving accuracy. | <https://easychair.org/publications/download/Jvsw> |
| 44 | 12 Dec  2016 | Prof. S.B. Madankar, Himani Vihare, Aparna Deshmukh , Harshal Walde , Ashok Gaikwad | Content based Classification of Emails for Email Filtering and Spam Detection | Nearest Mean Vector Classification**, Metric Model,** **Naive Learner**,  Pattern recognition. | Accuracy:  91.4% | The Naive Euclidean training method offers rapidity and reasonable accuracy, establishing itself as a CPU-time-efficient baseline for spam detection. | <https://ijarcce.com/upload/2016/december-16/IJARCCE%20116.pdf> |
| 45 | 6 June 2016 | Parhat Parveen , Prof. Gambhir Halse | Spam Mail Detection using Classification | Naïve Bayes, SVM (Support vector machine) ,  J48,  WEKA | Accuracy:  76% | Naïve Bayes Algorithm achieves superior accuracy at 76% and faster processing. | <https://ijarcce.com/upload/2016/june-16/IJARCCE%2074.pdf> |
| 46 | 17 March  2023 | Phani Teja Nallamothu and Mohd Shais Khan | Machine Learning for SPAM Detection | The decision tree, SVM, NB The logistic regression The AdaBoost ,ANN, CNN, The random forest, NB,  MEA. | Accuracy :  98.25% | Research highlight supervised machine learning, particularly SVM and Naive Bayes, as crucial for effective spam detection, despite their dependency on large labeled datasets. | <http://eprint.subtopublish.com/id/eprint/3333/1/dishivam9876mbimph,+Nallamothu612023AJOAIR2465.pdf> |
| 47 | 24 Feb  2012 | Ismaila Idris,  Abdulhamid Shafi’i Muhammad | An Improved AIS Based  E-mail Classification  Technique for Spam Detection | Artificial Immune System (AIS),  Rule Learning,  SVM, Naïve Bayes, Decision Tree. | Accuracy:  86.30% | Artificial Immune System reduce false positives, leveraging both spam and non-spam data, while future research targets precision-recall optimization. | <https://arxiv.org/ftp/arxiv/papers/1402/1402.1242.pdf> |
| 48 | Oct 2014 | Rohit Giyanani, Mukti Desai | Spam Detection using Natural Language Processing | N gram Modeling,  Bayesian Classification,  NLP | Accuracy:  91.63% | Further enhancements can be made in the field of spam detection for online security using Natural Language Processing in future. | <https://www.iosrjournals.org/iosr-jce/papers/Vol16-issue5/Version-4/S01654116119.pdf> |
| 49 | 2014 | Mandeep Choudhary,  V. S. Dhaka | Automatic e-mails Classification Using Genetic Algorithm | Genetic Algorithm (GA),  Numpy,  Matplotlib,  Statistical Analysis | Accuracy :  81% | Achieving over 81% efficiency, dependent on dataset size and algorithm parameters, aiming to reduce false positives/negatives and explore advanced parameter characterization for improved filtering accuracy. | <https://www.ijana.in/Special%20Issue/C9.pdf> |
| 50 | 14 March  2012 | R. Kishore Kumar,  G.  Poonkuzhali,  P. Sudhakar | Comparative Study on Email Spam Classifier using Data Mining Techniques | Principal Component Analysis,  ID3, Log Regression TRIRLS,  Random Forest, SVM. | Accuracy  99% | Fisher and runs filtering feature selection algorithms exhibit superior classification performance. | <https://www.iaeng.org/publication/IMECS2012/IMECS2012_pp539-544.pdf> |
| 51 | 2024 | Happy, Ayush and Vishal | Updated Bert Technique | Google Collab,Pycharm, Jupyter Notebook, TensorFlow,Pandas,Numpy,Sklearn | Accuracy 0.92  F1-score 0.92 | Enhanced Accuracy,  Integration with email platforms,  Integration with Security Systems,  Multi-Modal Spam Detection including images, video | <https://github.com/aaayushsharma/Email_Spam_Predictor> |