3.2.1 Number of papers published per teacher in the Journals notified on UGC website during theyear 2022-2023 (5)

Title of paper	Name of the author/s	Department of the teacher	Name of journal	Year of publication	ISSN number	Link to the recognition in UGC enlistment
Analysis and Optimization of Surface Roughness in Turning of Wet & Dry Condition of Alloy Steel Using Taguchi Method	yogesh mishra	ME	IJMRT	2022	ISSN 2582-7358	https://ijmrtjournal.com/wp-content/uplo
SIMULATION & DEVELOPMENT OF HIGHER EER FOR OIL FLOODED ROTARY SCREW COMPRESSOR	yogesh mishra	ME	IJPREMS	2023	e-ISSN : 2583-1062	https://ijrpr.com/uploads/V4ISSUE8/IJRPR
EFFECT OF EDM PARAMETERS IN OBTAINING MAXIMUM MRR AND MINIMUM EWR BY USING TAGUCHI METHOD	yogesh mishra	ME	IJPREMS	2023	e-ISSN : 2583-1062	https://www.ijprems.com/uploadedfiles/p
A COMPREHENSIVE REVIEW ON DIFFERENT WELDED JOINTS AND THEIR NUMERICAL ANALYSIS	yogesh mishra	ME	IJRTI	2023	ISSN: 2456-3315	https://www.ijrti.org/papers/IJRTI230420
EVALUATING THE LIFE SPAN OF WELDED JOINTS WITH DIFFERENT EDGE PROFILES	yogesh mishra	ME	IJRTI	2023	ISSN: 2456-3315	https://www.ijrti.org/papers/IJRTI230420
SIMULATION & DEVELOPMENT OF HIGHER EER FOR OIL FLOODED ROTARY SCREW COMPRESSOR USING FEM	yogesh mishra	ME	IJRPR	2023	ISSN: 2582-7421	https://ijrpr.com/uploads/V4ISSUE8/IJRPR
DESIGN ANALYSIS AND OPTIMISATION IN IC ENGINE OF OD PISTON BY ANSYS	yogesh mishra	ME	IJPREM	2023	ISSN: 2582-3930	https://www.ijera.com/papers/Vol4_issue
DESIGN AND ANALYSIS OF BRAKE DISC PLATE USING FINITE ELEMENT ANALYSIS	yogesh mishra	ME	IRJET	2023	ISSN: 2395-0072	https://www.irjet.net/archives/V10/i5/IRJ
ANALYSIS OF EXHAUST MUFFLER IN IC ENGINE	yogesh mishra	ME	IJSRET	2023	ISSN: 2395-	http://www.ierjournal.org/pupload/mitpg

USING FEM METHOD					566X	
A STUDY ON DESIGN AND ANSYS OF CRANE HOOK	yogesh mishra	ME	IJRTSM	2023	ISSN: 2455-9679	https://scholar.google.co.in/scholar?q=A+
DESIGN AND ANALYSIS OF BRAKE DISC PLATE USING FEM	ISHAN KHAN	ME	IRJET	2023	ISSN: 317	https://www.irjet.net/archives/V10/i5/IRJ
COMPUTATIONAL EXPERIMENTAL ANALYSIS OF TURBINE INLET LARGE SIZE BUTTERFLY VALVES BY USUING FEM METHOD	yogesh mishra	ME	IJRTSM	2022	ISSN: 2455-9679	https://www.google.com/search?sca_esv=cGS8I6DAxUmp1YBHflPCHcQkeECKAB6BA
A REVIEW ON FEM ANALYSIS OF RIGID FLANGE COUPLING WITH DIFFERENT MATERIALS TO ENHANCE COUPLING LIFES	Dr. RAJEEV ARYA	ME	IRJMETS	2022	ISSN: 32168	https://www.irjmets.com/uploadedfiles/p
MODELING AND FEM ANALYSIS OF RIGID FLANGE COUPLING WITH TWO DIFFERENT MATERIALS AND DIFFERENT DESIGN OF FLANGES OF COUPLING	Dr. RAJEEV ARYA	ME	IJPREMS	2023	ISSN: 30788	https://www.ijprems.com/paperdetail.phpthpr=Vivek
DESIGN AND OPTIMISATION OF LOWER AND UPPER RAIL FOR CRACK MECHANISAM BY USING FEM METHOD	RAMNARAYAN SAHU	ME	IJMRP	2022	ISSN: 2582-7358	https://scholar.google.co.in/scholar?q=DE
EFFECT OF EDM PARAMETERS IN OBTAINING MAXIMUM MRR AND MINIUM EWR BY USING TOGUCHI METHOD	yogesh mishra	ME	IJMRP	2022	ISSN: 2582-7358	https://www.google.com/search?sca_esv=8I6DAxXVrlYBHc7YBYkQkeECKAB6BAgNEA
A Survey on Chronic Kidney Disease (CDK) Prediction Model Based on Machine Learning Algorithm	Prof. Mahendra Sahare	AIML	IJETT	2022	ISSN – 3699-4213	Volume 3 Issue 1 (January-June 2023) – IJE (ijettjournal.online)
Machine Learning Algorithm Based Prediction Model for Chronic Kidney Disease	Prof. Mahendra Sahare	AIML	IJETT	2022	ISSN – 3699-4213	Volume 3 Issue 1 (January-June 2023) – IJE (ijettjournal.online)
A Review on Machine Learning Based Intrusion Detection System using Sampling Techniques	Prof. Mahendra Sahare	AIML	IJETT	2023	ISSN – 3699-4213	Volume 3 Issue2 (July-December 2023) – IJETT (ijettjournal

Increasing Machine Learning Based Intrusion Detection System using Class Balancing Techniques	Prof. Mahendra Sahare	AIML	IJETT	2023	ISSN – 3699-4213	Volume 3 Issue2 (July-December 2023) – IJETT (ijettjournal.online)
Sentiment Classification of Financial Texts for Stock Markets using LSTM Technique: A Survey	Prof. Mahendra Sahare	AIML	International Journal of Engineering and Techniques	Oct-22	ISSN: 2395-1303	IJET International Journal of Engineering and Techniques (ijetjournal.org)
Sentiment Classification of Financial Texts for Stock Forecasting using LSTM Technique	Prof. Mahendra Sahare	AIML	International Journal of Engineering and Techniques	6-Dec-22	ISSN: 2395-1303	IJET International Journal of Engineering a Techniques (ijetjournal.org)
Remote Aquatic Environmental Monitoring using Internet of things	Prof. Mahendra Sahare	AIML	Patent	20/01/2023	Patent Application Number: 202341001680	https://search.ipindia.gov.in/IPOJournal/J
"Data Mining using face detection with Logic"	Prof. Jaya Malviya	AIML	IJSDR	2022	ISSN:2455-2631	https://www.ijsdr.org/viewpaperforall.php paper=IJSDR2212076
An optimize image classification convolutional neural network over CIFAR-10 Dataset	Prof. Jaya Malviya	AIML	JICR	2022	ISSN:0022-1945	
Machine Learning Algorithm Based Prediction Model for Chronic Kidney Disease	Prof Anurag Shrivastava	CSE	IJETT	2023	ISSN-3699-4213	Volume 3 Issue 1 (January-June 2023) – IJE (ijettjournal.online)
Increasing Machine Learning Based Intrusion Detection System using Class Balancing Techniques	Prof Anurag Shrivastava	CSE	IJETT	2023	ISSN-3699-4213	Volume 3 Issue2 (July-December 2023) – I (ijettjournal.online)
A Review on Sentiment Analysis Using Natural Language Processing and Machine Learning	Prof Vaibhav Patel	CSE	IJETT	2023	ISSN-3699-4213	Volume 3 Issue2 (July-December 2023) – I (ijettjournal.online)
A Survey on Fake News Detection on Social Media Using Natural Language Processing	Prof Vaibhav Patel	CSE	IJETT	2023	ISSN-3699-4213	Volume 3 Issue2 (July-December 2023) – I (ijettjournal.online)
A Survey on Stock Price Prediction Using Machine Learning	Prof Santosh Nagar	CSE	IJETT	2023	ISSN-3699-4213	Volume 3 Issue2 (July-December 2023) – I (ijettjournal.online)
A Survey on Facial Expression Recognition using machine learning Technique	Prof Nitesh Gupta	CSE	IJETT	2023	ISSN-3699-4213	https://ijettjournal.online/?page_id=412

Modified Level Converter Control based Integration of Solar Photovoltaic System with The Grid	Prof Madhu Upadhyay	EX	IJRPR	2023	ISSN-2582-7421	https://ijrpr.com/uploads/V4ISSUE11/IJRP
Multilevel Converter for Solar Photovoltaic System Integration with Modified Level Control	Prof Madhu Upadhyay	EX	IJRPR	2023	ISSN-2582-7421	https://ijrpr.com/uploads/V4ISSUE11/IJRP
MPPT Technique Based Solar Power Extraction for Supplying the Loads Using Boost Converter	Prof Madhu Upadhyay	EX	IJRPR	2023	ISSN-2582-7421	https://ijrpr.com/uploads/V4ISSUE8/IJRPR
100KW Photovoltaic Grid Connected System using MPPT and Converter	Prof Jagdish Koli	EX	IJARSCT	2023	ISSN-2581-9429	https://www.ijarsct.co.in/Paper13627.pdf
Minimizing Power Loss in a Distribution System by Optimal Sizing of DC using PSO	Prof Jagdish Koli	EX	IJARSCT	2023	ISSN-2581-9429	https://www.ijarsct.co.in/Paper13639.pdf
Electric Vehicles with Battery Storage Based dc Charging Station using Solar	Prof Jagdish Koli	EX	IJARSCT	2023	ISSN-2581-9429	https://www.ijarsct.co.in/Paper12304.pdf
A Review on high gain DC-DC Boost converter topologies with recent advances	Prof Sagar tomar	EX	Journal of Computing Technologies (JCT) INTERNATIONAL JOURNAL	2022	2278-3814	http://www.jctjournals.com
High Gain DC-DC Boost Converter for Renewable energy applications	Prof Sagar tomar	EX	Journal of Computing Technologies (JCT)INTERNATIONAL JOURNAL	2022	2278-3814	http://www.jctjournals.com
A literature Survey on transformer less PV system with different filter for current and harmonics minimization	Prof Hemant Patel	EX	Journal of Computing Technologies (JCT) INTERNATIONAL JOURNAL	2022	2278-3814	http://www.jctjournals.com
A modified Transformer Less PV system with double tunned filter for current and harmonics minimization	Prof Hemant Patel	EX	Journal of Computing Technologies (JCT) INTERNATIONAL JOURNAL	2022	2320-0804	http://www.jctjournals.com

Solar with battery storage based DC charging Station for Electric Vehicles	Prof Hemant Patel	EX	Journal of Computing Technologies (JCT) INTERNATIONAL JOURNAL	2022	2278-3814	http://www.jctjournals.com
Renewable energy with battery storage based charging station for electric vehicles	Prof Hemant Patel	EX	Journal of Computing Technologies (JCT) INTERNATIONAL JOURNAL	2022	2278-3814	http://www.jctjournals.com
the Impact of Change and change management in achieving corporate goals and objectives	Dr. Rajesh Sahu	Management	IRJMSH (UGC Care)	7-Jul-23	2277-9809(0)2348-9359 (p)UGC ID-48312	https://srmu.ac.in/pdf2/Research%20Publ
A study on employees' contentment towards performance appraisal system at HDFC Bank	Mr. Vishal Verma	Management	International journal of innovation in engineering research and management	Aug-22	ISSN:2348-4918	https://mrec.ac.in/Research/JournalPublic
A study of Digital inovations in Business and Economy	Dr. Rajesh Sahu	Management	BSSS Jouyrnal of Management(BJOM)	Jun-23	E-ISSN-4643	https://econpapers.repec.org/article/idsijl
Object Detection with YOLO Version 3 for Big Data	Prof. Satyendra Dubey	IT	International Conference on Information Systems and Management Science	Oct. 2022		https://www.mdpi.com/2076-3417/9/18/
A Novel Approach for Service Selection and Ranking in Federated Cloud	Prof. Satyendra Dubey	IT	International Conference on Information Systems and Management Science	Oct. 2022		https://journalofcloudcomputing.springer
Efficient Load Balancing of Resources for Different Cloud Service Providers in Cloud Computing	Prof. Satyendra Dubey	IT	International Conference on Information Systems and	Feb-23	ISSN 2456-3307	https://ijsrcseit.com/archive.php?v=12&i=

			Management Science			
Data Integrity Verification in Cloud Computing	Prof. Satyendra Dubey	IT	International Conference on Information Systems and Management Science	Apr.-2023	ISSN 2456-3307	https://www.researchgate.net/publicatio

INTERNATIONAL JOURNAL OF EMERGING TRENDS & TECHNOLOGIES (IJETT)
ISSN - 3699-4213, DOI - 10.55093/ijett
https://ijettjournal.org/

A Survey of Bankruptcy Prediction Models Based on ML Classification Framework

Rakesh Yadav
MTech Scholar
NIIST, Bhopal
rakeshkumaryadav.1401@gmail.com

Prof. Santosh Nagar
AP, CSE
NIIST, Bhopal
santoshnagar9@gmail.com

Prof. Anurag Shrivastava
AP, CSE
NIIST, Bhopal
anurag.shri08@gmail.com

Abstract - Bankruptcy Prediction is becomes the most demanding topic with the rapid growth of economy. In the finance sector, early prediction of credit defaulters is very helpful in the decision-making process for financial organizations. Many researchers have used methods like statistical modeling and machine learning techniques for bankruptcy prediction. Training Machine learning (ML) based prediction model to predict a person/company goes bankruptcy or not is challenging issue. Hence, Enhanced prediction model is required that increase prediction model performance. This paper contributes towards a systematic understanding of machine learning based classification model to increase performance of bankruptcy prediction and decrease their related shortcomings using proposed classification model which is applied preprocessing, feature selection and ML techniques.

Keywords - Bankruptcy, Bankruptcy Prediction Model, Machine Learning, Classification, Classification framework

1. INTRODUCTION

Bankruptcy prediction has become a most popular research interests over the past decade since it can have significant impact on financial, corporate, government and banking decision-making process. Accuracy is one of crucial performance due to its significant economic impact numerous statistical techniques have been used for improving the performance of bankruptcy prediction models. Bankruptcy prediction has become most important topic for researchers and companies by using various models. The artificial neural network, support vector machine and many machine learning algorithms are used for this purpose. Basically, there are two approaches to predict the bankruptcy: univariate analysis and multivariate analysis. Univariate analysis used to predict financial distress which is the distribution of financial variables for companies that experiencing financial distress is different from companies that don't have financial distress. Deficiency of this model is contradiction between the predicted variables. To solve this problem, multivariate models was developed. The independent variables in this model are the financial ratios that expected to affect bankruptcy, while the dependent variable is the prediction results. But till now, little theoretical discussion only that leads to bankruptcy research, e.g. in the selection of variables that are considered relevant. With at least the theory, bankrupt prediction is more directed to the search for variables that are considered relevant to the trial and error methods [25]. Financial fraud is a growing concern with far reaching consequences in the government, corporate organizations, finance industry, In Today's world high dependency on internet technology has enjoyed increased credit card transactions but credit card fraud had also accelerated as online and offline transaction.[26]

2. LITERATURE REVIEW

An author [1] applies anomaly detection algorithms to the bankruptcy prediction problem in an attempt to suggest a new stable model taking the data distribution into consideration. The efficacy of anomaly detection techniques is tested on bankruptcy prediction datasets Polish banks. Author's empirical evaluation shows that Isolation Forest outperforms multivariate Gaussian distribution, one class SVM, and other classification estimators in terms of the ROC curve.

According to the authors [2] Selection of dataset for training the prediction model, the Machine Learning tool used for prediction and various other factors are essential in building an efficient prediction model. The dataset includes financial ratios as attributes that are derived from the financial statements of various companies. The most Influencing ratios that are required for predicting bankruptcy are selected on the basis of the Genetic Algorithm which filters out the most important ones from different existing bankruptcy models. These ratios of different companies are fed as an input to train the model being implemented in R. The prediction algorithm used is Random Forest, which will enable us to differentiate between bankrupt and non-bankrupt companies.

Authors [4] reduce the imbalance nature of data and then train a deep neural network with the balanced data. The dataset used is that of polish companies which consists of five years of data corresponding to five years of five different tasks. Authors reduce class imbalance using an oversampling method known as SMOTE. Author's model significantly outperforms the previous neural network models and weak learners trained this dataset in terms of AUC.

Authors [26] investigated the design and the application of different ML models to two different tasks related to default events first one is estimating survival probabilities over time and second default prediction using time-series accounting data with different lengths. The entire dataset used for the experiments has been made available to the scientific community for further research and

Copyright 2022 IJETT. ALL RIGHTS RESERVED

INTERNATIONAL JOURNAL OF EMERGING TRENDS & TECHNOLOGIES (IJETT)
ISSN - 2349-4123, DOI - 10.55083/ijett
https://ijettjournal.org/

A Survey on Stock Price Prediction Using Machine Learning

Gopal Kumar
MTech Scholar
CSE, Department
NIIST Bhopal
Pal38916@gmail.com

Prof. Vaibhav Patel
Assistant Professor,
CSE, Department
NIIST Bhopal
vaibhav.bce@gmail.com

Prof. Anurag Shrivastava
HOD
CSE, Department
NIIST Bhopal
Anurag.shri08@gmail.com

Abstract: With theincreasinguse of machine learning techniques for stock price detection has gained significant attention in recent years. This survey aims to provide an overview of the various methodologies and approaches employed in this domain. The study analyzes the application of machine learning algorithms for stock price prediction, focusing on their performance, accuracy, and effectiveness. Different types of features used in modeling, such as technical indicators, sentiment analysis, and financial news, are also explored. Furthermore, this survey presents an evaluation of the challenges and limitations encountered in stock price detection using machine learning, including data availability, model overfitting, and market volatility. By examining the existing literature and research advancements, this survey contributes to a comprehensive understanding of the current state of stock price detection using machine learning and identifies potential areas for future research and improvement.

Keywords—Deep learning, Stock Price Prediction, LSTM, RNN, Neural Network

I. INTRODUCTION

Stock price prediction is a crucial and challenging task in the financial domain, with significant implications for investors, traders, and financial institutions. The ability to forecast future stock prices accurately can potentially yield substantial profits and mitigate risks. In recent years, there has been a surge in interest and research efforts focused on leveraging machine learning techniques to enhance stock price prediction models. This survey aims to provide an in-depth exploration of the various methodologies and approaches used in stock price prediction using machine learning. The study seeks to shed light on the current state of the field by examining the performance, accuracy, and effectiveness of different machine learning algorithms employed for this task.

The predictive power of machine learning algorithms lies in their ability to uncover complex patterns and relationships within historical stock market data. Various types of features, such as historical price trends, trading volumes, technical indicators, sentiment analysis from social media, and financial news, are utilized to capture relevant information for modeling.By analyzing the existing literature and research advancements, this survey aims to identify the strengths and limitations of different machine learning techniques for stock price prediction. Challenges such as data availability, model overfitting, market volatility, and interpretability are also discussed, providing insights into the practical implementation of these models.

Moreover, this survey endeavors to highlight the potential applications and benefits of accurate stock price prediction, including portfolio management, algorithmic trading, risk assessment, and investment decision-making.

Overall, this survey serves as a comprehensive resource for researchers, practitioners, and enthusiasts interested in understanding the current landscape of stock price prediction using machine learning. It provides valuable insights into the state-of-the-art techniques, identifies areas for improvement, and paves the way for future advancements in this exciting and rapidly evolving field. In conclusion, this survey paper serves as a comprehensive guide to the current landscape of news

International Journal of Electronics Communication and Computer Technologies (IJECCT)
ISSN - 3838-4213 DOI - 10.55083/ijeccT

A Survey on Fake News Classification Using Machine Learning and Natural Language Processing

Entezaz Bharti
MTech Scholar
CSE, Department
NIIST Bhopal
Entezaznitrr@gmail.com

Prof. Vaibhav Patel
Assistant Professor,
CSE, Department
NIIST Bhopal
vaibhavbce@gmail.com

Prof. Anurag Shrivastava
HOD
CSE, Department
NIIST Bhopal
Anurag.sri08@gmail.com

Abstract. With the rise of social media and the ease of information dissemination, the problem of fake news has become a significant concern. This survey aims to provide an overview of various methodologies and approaches employed to tackle this issue. It covers the utilization of NLP techniques such as text preprocessing, feature extraction, and sentiment analysis, along with machine learning algorithms including support vector machines, random forests, and deep learning models. The survey also discusses the challenges and future directions in this field to enhance the accuracy and effectiveness of fake news detection systems.

Keywords— NLP, machine learning, fake news classification, social networks

I. Introduction

The proliferation of fake news on social media platforms and online news sources has become a critical issue in recent years. The spread of misinformation can have severe consequences, impacting public opinion, political discourse, and even public safety. As a result, there has been a growing interest in developing effective techniques to automatically detect and classify fake news. Machine learning (ML) and natural language processing (NLP) have emerged as powerful tools in the fight against fake news. ML algorithms can learn patterns and features from large datasets, while NLP techniques enable the understanding and analysis of textual data. This combination offers promising opportunities to develop automated systems capable of distinguishing between genuine and fake news articles.

This survey aims to provide a comprehensive overview of the state-of-the-art techniques used in fake news classification using ML and NLP. It explores the various steps involved in the classification process, starting from data preprocessing, feature extraction, and representation to the selection and evaluation of ML models. Text preprocessing techniques, such as tokenization, stemming, and stop-word removal, play a crucial role in transforming raw text into a suitable format for analysis. Feature extraction methods extract meaningful information from text, including linguistic features, syntactic structures, and semantic representations. The survey also covers different ML algorithms commonly employed for fake news classification, such as support vector machines (SVM), random forests, and deep learning models like recurrent neural networks (RNNs) and convolutional neural networks (CNNs). It examines the strengths and limitations of each approach and highlights their performance in detecting fake news.

Furthermore, this survey discusses the challenges associated with fake news classification, including the presence of evolving news dynamics, the need for large annotated datasets, and the issue of adversarial attacks. It also explores potential directions for future research and suggests ways to enhance the accuracy and robustness of fake news detection systems. Overall, this survey serves as a valuable resource for researchers and practitioners interested in understanding the current landscape of fake news classification using ML and NLP techniques.

A. Machine Learning for Fake news Classification:

Machine learning for fake news classification involves using algorithms and statistical models to automatically detect and classify misinformation. By analyzing

Copyright 2023 IJECT. ALL RIGHTS RESERVED

Sentiment Classification of Financial Texts for Stock Forecasting using LSTM Technique

[1]Kanchan Raipure, [2]Prof. MahendraSahare, [3]Prof. Anurag Shrivastava
M Tech Scholar[1], Associate Professor[2,3]
Department of Computer Science and Engineering
NRI Institute of Information Science and Technology, Bhopal

Abstract: -The stock market is an emerging network that offers an infrastructure for all financial transactions from the world in a dynamic rate called stock value, which is devised using market stability. Prediction of stock values provides huge profit opportunities which are considered as an inspiration for research in stock market prediction. Long short term memory (LSTM) is a model that increases the memory of recurrent neural networks. Recurrent neural networks hold short term memory in that they allow earlier determining information to be employed in the current neural networks. For immediate tasks, the earlier data is used. We may not possess a list of all of the earlier information for the neural node. The long short-term memory (LSTM) and gated recurrent unit (GRU) models are popular deep-learning architectures for stock market forecasting. Various studies have speculated that incorporating financial news sentiment in forecasting could produce a better performance than using stock features alone. This study carried a normalized comparison on the performances of LSTM and GRU for stock market forecasting under the same conditions and objectively assessed the significance of incorporating the financial news sentiments in stock market forecasting. Both the LSTM-News and GRU-News models are able to produce better forecasting in stock price equally.

Keywords: -Stock Market, LSTM, GRU, Neural Network

I. INTRODUCTION

Sentiment analysis can be performed in many classification algorithms considering a single word (unigram) as a feature. In certain cases unigrams can lead to misclassification of sentiments. New methods were proposed to use the N-gram approach for analysis of sentiments. Certain investigations revealed that as the value of N increases the accuracy of classification decreases. N value was tested with only three as trigrams but it needs more training time. Bigrams were used in some works, which is formed by combining the adjacent words results in a slighter increase in accuracy than unigrams. To eliminate the overhead of input the bigrams are checked with the predefined lexicons to extract the essential bigrams. In the sentiment classification machine learning techniques are widely used at various levels. Presently, the machine learning based approaches are performing on par with human level accuracy (Bayes error) in most of the tasks. Human Generated Classification is a way to evaluate an explanation / review is by asking humans to guess the output of a model based on the explanation and the input. This is referred as forward simulation prediction [1, 2]. Analysis of sentiments is done at three levels. Document level is used to classify whether the opinion of the document is positive, negative or neutral [3]. The sentence level is used to determine whether the document's opinion is positive, negative or neutral. The aspect level is used to focuses on all expression of sentiments present within a given document. The advantage of learning-based methodologies is that it is very simple and quick to construct. Online reviews are much more efficient and flexible algorithms are required for sentiment classification than the current approaches, which changes rapidly over time. The sentiment sensitive thesaurus has been used to identify the word's sentiment as a lexicon-based approach [4, 5].Indian stock market has shown extreme volatility in the last decade, with prices surging wildly, without any change in fundamentals. The BSE Sensex, which was more than 5000 points in the year 2000, has dipped to less than 3500 points by the second half of the year 2001. In addition to stock market scams at regular intervals, the fall of information technology shares, which has reached to the highest levels, has brought down the Sensex drastically in the financial year 2000-01. Later in 2008 also SENSEX started at 20,352 in January and drastically fell to 9162 in December. The security market exhibits various patterns and styles which are either seasonal at times and shall not follow any seasonality some more times.

II. SENTIMENT ANANLYSIS

The overview of the sentiment analysis process is shown in Figure 1. The objective of sentiment analysis is to extract the opinions from the reviews, identify the sentiment they exhibit and finally to classify the sentiment polarity.

Sentiment Classification of Financial Texts for Stock Markets using LSTM Technique: A Survey

[1]Kanchan Raipure, [2]Prof. Mahendra Sahare, [3]Prof. Anurag Shrivastava
M. Tech. Scholar[1], Associate Professor[2,3]
Department of Computer Science and Engineering
NRI Institute of Information Science and Technology, Bhopal

Abstract: - Sentiment Analysis (SA) is the current field of research in text mining field. SA is detecting opinions, sentiments, and subjectivity of text. It is the application of natural language processing techniques and text analytics to identify and extract subjective information from the frequently used sources such as web and microblogs. The main objective of sentiment analysis is to analyze reviews of products and services, and determine the scores of such sentiments. The major problem is that the reviews are mostly unstructured and thus, need classification or clustering to provide meaningful information for future use. The study was conducted with different sectors like, IT sector, Automobile sector, Banking sector, Pharmaceutical sector and FMCG sector. All the sectors taken for the study is highly volatile compared to other sectors in NSE/BSE. Hence it is very essential to study on the nexus between the Indian Stock Market and selected companies behavior. The early stage of the share market was very familiar for average investor. Now the markets are wide enough to invest. There are different markets like bond market, forex market, derivative market and other specialty markets. Analysis of the stock price we take the price. By using the artificial neural network, we develop a model. within the neural network, we use a recurrent neural network that remembers each and each information through time.

Keywords: - **National Stock Exchange (NSE), Bombay Stock Exchange (BSE), Neural Network**

I. INTRODUCTION

Sentimental analysis is the study of people's ideas, opinions, feelings, attitudes, thoughts, and their associative features for determining the polarity of the reviews which can be used to improve products and services. Sentimental analysis is also called as opinion mining is a part of data mining activity that analyses text to topics. Sentimental analysis is classified into three types, document level, sentence level and aspect level. The sentence-level does not represent what the reviewers like or not and the Sentimental analysis is done at the document level to predict the opinion. A key problem in this area is sentiment classification, where a document is expressed as a positive or negative evaluation of a target object, the target object may be film, book, movie and business products. Sentiment Analysis supports both supervised and unsupervised learning approaches. Many machine learning approaches are used for effective sentiment classification. Naive Bayes algorithm is proposed for sentimental analysis. The classification result and the accuracy of Naive Bayes, and SVM are analyzed through a comparative experiment [1]. Techniques are also proposed to evaluate the scalability of Naive Bayes Classifier (NBC) in large scale datasets. Naive Bayes classifier is implemented to achieve fine-grain control of the analysis procedure for different implementations. The valuable information is obtained by the typical method to extract the sentiment or opinion from a message which is useful for the business consumer industry was reviewed [2].

The financial markets in India especially stock markets have witnessed a number of changes in terms of trading environment, regulatory framework, increased competition, increase in range of instruments traded and presence of better informed market players. In a financial market investors are exposed to a variety of information on daily basis like updated corporate earnings, revised macroeconomic indices, policy maker's statements and political news. A public company initiates a number of actions, termed as corporate actions, which affect prices of securities issued by company. The corporate actions are - dividend announcements, rights issues, mergers and acquisitions, stock splits, bankruptcy, delisting, de-merger, initial public offer, liquidation, takeover, partial or final redemption, conversion of convertible bonds, warrants issue etc. These corporate actions may have direct or indirect impact on prices, volume, face value etc. of securities when information about the actions is disclosed through any channel of communication. A corporate announcement is information about the corporate action. It is important to understand which news is good or positive and which news is bad or negative for a company [3, 4]. Researchers are interested in understanding effect of different corporate announcements. Studies have been done in past to analyze impact of these corporate announcements on share prices. Stock split is a corporate announcement which has been subject of interest for academicians as well as practitioners and topic of interest in the current study. There is ample scope for academic contributions in this research area, both in terms of appropriately analyzing the impact of stock splits and in terms of extending literature by investigating new aspects of the theme [5].

The present chapter starts by introducing concept of stock splits along with brief discussion of different hypotheses relating to it formulated and tested in different empirical studies done in past covering various aspects of stock splits.

INTERNATIONAL JOURNAL OF EMERGING TRENDS & TECHNOLOGIES (IJETT)
ISSN - 3699-4213, DOI - 10.55093/ijett
https://ijettjournal.online/

Increasing Machine Learning Based Intrusion Detection System using Class Balancing Techniques

Mukesh Goswami
MTech Scholar
CSE Department
NIIST Bhopal
Mukeshgoswami.78.mg@gmail.com

Prof.Mahendra Sahare
Associate Professor
CSE Department
NIIST Bhopal
mahendrasahare1110@gmail.com

Prof.Anurag Shrivastava
HOD, CSE Department
NIIST Bhopal
anurag.shri08@gmail.com

***Abstract*—**With the rapid growth of digital transitions and connecting things with internet huge amount of data collected every day. Classification of this large dataset with high accuracy and low error rate is also an issue. Sampling of data is one the solution of large dataset. An imbalanced class of dataset is another problem for improving classifiers performance. Many researchers have been used machine learning algorithms for improving the classifier performance. This work proposes a sampling technique for obtaining the balanced sampled data. Sampled dataset represent the whole dataset. Imbalanced Classes also balanced by sampling techniques like oversampling and under-sampling. Majority class under-sampled and minority class oversampled for balancing the imbalanced classes. The main purpose of this work is to propose classification model. This model used the sampling technique for class balancing and machine learning algorithms to improve the classifier performance. The Proposed work is tested on basis of Accuracy and Error rate. KDD Cup 99 dataset used in this work.

***Keywords*—Sampling, Classification, Machine learning technique, IDS, Imbalance**

I. INTRODUCTION

We securing information either in private or government sector has become an essential requirement. System vulnerabilities and valuable information magnetize most attackers' attention. Traditional intrusion detection approaches such as firewalls or encryption are not sufficient to prevent system from all attack types. The number of attacks through network and other medium has increased dramatically in recent years. Efficient intrusion detection is needed as a security layer against these malicious or suspicious and abnormal activities. Thus, intrusion detection system (IDS) has been introduced as a security technique to detect various attacks. IDS can be identified by two techniques, namely misuse detection and anomaly detection. Misuse detection techniques can detect known attacks by examining attack patterns, much like virus detection by an antivirus application. However they cannot detect unknown attacks and need to update their attack pattern signature whenever there is new attacks .On the other hand, anomaly detection identifies any unusual activity pattern which deviates from the normal usage as intrusion. Although anomaly detection has the capability to detect unknown attacks which cannot be addressed by misuse detection, it suffers from high false alarm rate. In recent years, and interest was given into machine learning techniques to overcome the constraint of traditional intrusion techniques by increasing accuracy and detection rates. New machine learning based IDS with sampling is used in our detection approach. The advantage of IDS

Copyright 2023 IJETT. ALL RIGHTS RESERVED

INTERNATIONAL JOURNAL OF EMERGING TRENDS & TECHNOLOGIES (IJETT)
ISSN - 3699-4213, DOI - 10.55093/ijett
https://ijettjournal.online/

A Survey on Chronic Kidney Disease (CDK) Prediction Model Based on Machine Learning Algorithm

Shiv Shankar
MTech Scholar
CSE, Department
NIRT Bhopal

Prof. Mahendra Sahare
Assistant Professor,
CSE, Department
NIRT Bhopal

Prof. Anurag Shrivastava
Assistant Professor,
CSE, Department,
NIIST Bhopal

Abstract - Due to irregular and unhealthy life style, India and all over the world chronic kidney disease (CKD) is an increasing and serious disease impacting public health. The symptoms of CKD are often appearing too late and many patients inevitably face pain and expensive medical treatments. The ultimate treatment is frequent dialysis or Kidney transplant. Early detection of disease through symptoms can prevent the disease progression by referral to appropriate health care services. Now days, Machine Learning (ML) techniques have been widely used in healthcare sector. ML techniques can help in identifying the potential risk by discovering knowledge from medical reports of patient. Thus helps in preventing the disease progression. In this paper review the CDK prediction models which is based on machine learning technique. The finding of the survey is that several models for detecting the risk of CKD are proposed by various authors, which are based on machine learning algorithms like decision tree, SVM and neural network. The use of cloud platforms for building the ML model, several literatures uses the platform confirms a considerable performance improvement in running ML tasks.

Keywords - Boosted Decision Tree, Neural Network, Chronic Kidney Disease (CKD), Data Analytics, Health Care, Microsoft Azure, Machine Learning, Prediction Model

I. Introduction

The health care sector of the world is the most promising sector nowadays, due to pandemic. Medical Instruments, Medicines, hospital facility, and other required things are lacking everywhere. In this bad situation, it is required to enhance medical technology and its related research. Chronic kidney disease (CKD) poses a serious burden of disease worldwide with substantially increasing number of patients being diagnosed. According to the study [1] published in 2018 Global Burden of Disease study data and methodologies to describe the change in burden of CKD from 1990 to 2016 involving incidence, prevalence, death, and disability-adjusted-life-years (DALYs). Globally, the incidence of CKD increased by 89% to 21,328,972, prevalence increased by 87% to 275,929,799, death due to CKD increased by 98% to 1,186,561, and DALYs increased by 62% to 35,032,384. Figure 1.1 showing the Google search results for CKD [2]. Also, the cost related to CKD care is too high. So early detection and identification of patients with increased risk of developing CKD on the basis of symptoms can improve care by preventive measures to slow disease progression and timely initiation of nephrology care.

Enormous complex data is being regularly received by healthcare division about diseases, treatment, patients, medical equipments, hospitals and claims etc. The data requires processing for extraction of knowledge. Data Mining is predominantly helpful in healthcare domain when it is difficult to deal a disease with particular treatment option. DM comprise of efficient techniques and tools to apply on healthcare data for making appropriate decisions towards taking preventive measures and predicting risks of disease.

II. Literature Review

An author [1] studies both binary and multi classifications for stage prediction have been carried out. The prediction models used include Random Forest (RF), Support Vector Machine (SVM) and Decision Tree (DT). Analysis of variance and recursive feature elimination using cross validation have been applied for feature selection. Evaluation of the models was done using tenfold cross-validation. The results from the experiments indicated that RF based on recursive feature elimination with cross validation has better performance than SVM and DT.

The research paper [2] authors employ experiential analysis of ML techniques for classifying the kidney patient dataset as CKD or NOTCKD. Seven ML techniques together are utilized and assessed using distinctive evaluation measures such as mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE), root relative squared error (RRSE), recall, precision, F-measure and accuracy.

In paper [3] prediction is performed using Naive Bayes Classifier and K-Nearest Neigbour algorithm. The data

Copyright 2022 IJETT. ALL RIGHTS RESERVED

INTERNATIONAL JOURNAL OF EMERGING TRENDS & TECHNOLOGIES (IJETT)
ISSN - 3699-4213, DOI - 10.55093/ijett
https://ijettjournal.online/

Machine Learning Algorithm Based Prediction Model for Chronic Kidney Disease"

Shiv Shankar
MTech Scholar
CSE, Department
NIRT Bhopal
shiv.om9191@gmail.com

Prof. Mahendra Sahare
Assistant Professor,
CSE, Department
NIRT Bhopal
mahendrasahare1110@gmail.com

Prof. Anurag Shrivastava
Assistant Professor,
CSE, Department,
NIIST Bhopal
anurag.shri08@gmail.com

Abstract—

In India and all over the world chronic kidney disease (CKD) is an increasing and serious disease impacting public health. Now days, Machine Learning (ML) techniques have been widely used in healthcare sector. ML techniques can help in identifying the potential risk by discovering knowledge from medical reports of patient. Thus helps in preventing the disease progression. This research aims at developing a Cloud based Predictive Model to detect the possibilities of CKD and its progression in patients with some health issues like hypertension and diabetes. The proposed model can help physicians to recognize patients at risk and prescribe the treatments and lifestyle changes. The model is trained and tested on the CKD data provided on UCI repository and it is deployed on Microsoft Azure ML platform. The model is built using two algorithms: Two classes Boosted decision tree and two class deep support vector machines. The model built over boosted decision tree algorithm has highest prediction accuracy. The results obtained from proposed model are compared with basic benchmark classifiers.

Keywords— Boosted Decision Tree, Neural Network, Chronic Kidney Disease (CKD), Data Analytics, Health Care, Microsoft Azure, Machine Learning, Prediction Model

I. Introduction

The health care sector of the world is the most promising sector nowadays, due to pandemic. Medical Instruments, Medicines, hospital facility, and other required things are lacking everywhere. In this bad situation, it is required to enhance medical technology and its related research. Chronic kidney disease (CKD) poses a serious burden of disease worldwide with substantially increasing number of patients being diagnosed. According to the study [1] published in 2018 Global Burden of Disease study data and methodologies to describe the change in burden of CKD from 1990 to 2016 involving incidence, prevalence, death, and disability-adjusted-life-years (DALYs). Globally, the incidence of CKD increased by 89% to 21,328,972, prevalence increased by 87% to 275,929,799, death due to CKD increased by 98% to 1,186,561, and DALYs increased by 62% to 35,032,384. Figure 1.1 showing the Google search results for CKD [2]. Also, the cost related to CKD care is too high. So early detection and identification of patients with increased risk of developing CKD on the basis of symptoms can improve care by preventive measures to slow disease progression and timely initiation of nephrology care. Enormous complex data is being regularly received by healthcare division about diseases, treatment, patients, medical equipments, hospitals and claims etc. The data requires processing for extraction of knowledge. Data Mining is predominantly helpful in healthcare domain when it is difficult to deal a disease with particular treatment option. DM comprise of efficient techniques and tools to apply on healthcare data for making appropriate decisions towards taking preventive measures and predicting risks of disease.

II. Literature Review

An author [1] studies both binary and multi classifications for stage prediction have been carried out. The prediction models used include Random Forest (RF), Support Vector Machine (SVM) and Decision Tree (DT). Analysis of variance and recursive feature elimination using cross validation have been applied for feature selection. Evaluation of the models was done using tenfold cross-

Copyright 2022 IJETT. ALL RIGHTS RESERVED

INTERNATIONAL JOURNAL OF EMERGING TRENDS & TECHNOLOGIES (IJETT)
ISSN - 3699-4213, DOI - 10.55093/ijett
https://ijettjournal.online/

A Review on Machine Learning Based Intrusion Detection System using Sampling Techniques

Mukesh Goswami
MTech Scholar
CSE Department
NIIST Bhopal
Mukeshgoswami.78.mg@gmail.com

Prof.Mahendra Sahare
Associate Professor
CSE Department
NIIST Bhopal
mahendrasahare1110@gmail.com

Prof.Anurag Shrivastava
HOD, CSE Department
NIIST Bhopal
anurag.shri08@gmail.com

Abstract—**Rapid development and popularization of information system, network security is a main important issue. Intrusion Detection System (IDS) as the main security defensive technique and is widely used against intrusion. Data Mining and Machine Learning techniques proved useful attention in security research area. Recently, many machine learning methods have also been applied by researchers, to obtain highly detection rate. problem of all those methods is that how to classify attack or intrusion effectively. Looking at such inadequacies, the machine learning technique is used for obtain the high detection rate. Also, use of internet is increasing progressively, so that large amount of data and it security is also an issue. Sampling technique is one the solution of large dataset. This work proposes a sampling technique for obtaining the sampled data. Sampled dataset represent the whole dataset with proper class balancing. Class imbalanced can be balanced by sampling techniques. The purpose of this paper is to review the machine learning based IDS which is used to sampling techniques for improving the performance in terms of accuracy.**

Keywords—**Sampling, Classification, Machine learning technique, IDS, Imbalance**

I. INTRODUCTION

We securing information either in private or government sector has become an essential requirement. System vulnerabilities and valuable information magnetize most attackers' attention. Traditional intrusion detection approaches such as firewalls or encryption are not sufficient to prevent system from all attack types. The number of attacks through network and other medium has increased dramatically in recent years. Efficient intrusion detection is needed as a security layer against these malicious or suspicious and abnormal activities. Thus, intrusion detection system (IDS) has been introduced as a security technique to detect various attacks. IDS can be identified by two techniques, namely misuse detection and anomaly detection. Misuse detection techniques can detect known attacks by examining attack patterns, much like virus detection by an antivirus application. However they cannot detect unknown attacks and need to update their attack pattern signature whenever there is new attacks .On the other hand, anomaly detection identifies any unusual activity pattern which deviates from the normal usage as intrusion. Although anomaly detection has the capability to detect unknown attacks which cannot be addressed by misuse

Copyright 2023 IJETT. ALL RIGHTS RESERVED

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 1, November 2023

Minimizing Power Loss in a Distribution System by Optimal Sizing of DG using PSO

Pravin Kumar[1], Nitish Kumar Nirala[2], Prof. Jagdish Koli[3], Prof. Madhu Upadhyay[4]

Department of Electrical and Electronic Engineering[1,2,3,4]

NRI Institute of Research and Technology, Bhopal, India

Abstract: *The distribution network becomes more unstable as distributed generator penetration levels climb. As a result, employing cutting-edge solutions is required to increase the network's stability and dependability. Finding the best placement and sizing for distributed generators is necessary, keeping in mind the need to optimize the voltage profile and reduce electrical power losses. In this paper, Particle Swarm Optimizer (PSO) methods are presented for network issue solving when several DGs of various sorts are installed. outcomes are compared using PSO methodologies with those with and without DG to demonstrate how the outcomes have improved. When the suggested approach is used with IEEE 33. It is shown by the findings that the voltage profile, active losses, and reactive losses have been significantly improved. For simulations, MATPOWER-DigSILENT software is employed. The performance and efficacy of the suggested strategies are well demonstrated by the simulated outcomes.*

Keywords: Particle Swarm Optimizer (PSO),Radial Distribution System (RDS),Optimal Power Flow(OPS),Distribution Network (DN),Distributed Generator (DG)

I. INTRODUCTION

The presented work focuses on optimizing both DG placement and capacity, addressing the challenge of increasing load demand while injecting both active and reactive power into the distribution system. This research aims to offer a comprehensive solution and compares its results with a reference using analytical methods. [1] Distributed generation is becoming increasingly important due to its potential to enhance the efficiency and reliability of distribution systems. Kansal et al. focus on determining the optimal locations for installing DG units in the distribution network. [2] The reference from the EIA is a valuable source for understanding the modeling of distributed generation within the building sector. Distributed generation, particularly in the context of buildings, has gained significant attention due to its potential for sustainable energy generation.[3] The focus is on improving voltage profiles within the network, which is crucial for maintaining the quality and reliability of the power supply. By considering DG and capacitors in the optimization process, this research contributes to the efficient integration of DG into distribution systems. [4] Network reconfiguration is a crucial aspect of distribution system optimization. This research is likely to delve into the use of mathematical techniques and continuation power flow theorem to optimize the network configuration, allowing for the maximum loading of the system while maintaining stability. [5] The use of the cuckoo search algorithm is a notable feature, highlighting the application of nature-inspired optimization techniques in distribution system management. [6] This technique is an example of nature-inspired optimization methods that have gained popularity in the field of power system optimization., this work contributes to the efficient operation of distribution networks.[7] Network reconfiguration is an important strategy for improving the efficiency of distribution networks, and this paper investigates how it affects loss allocation. [8] This research, although lacking specific publication details, provides insights into the application of AI techniques for distribution system optimization.[9] The application of graph theory in conjunction with these algorithms is a novel approach to solving network reconfiguration problems.[10] They propose an improved binary particle swarm optimization algorithm to tackle this problem, aiming to find the optimal combination of network configuration and capacitor placement to minimize power losses. this work presents an advanced optimization approach for loss reduction in distribution systems.

[11] In this paper present an approach that combines OPF and sensitivity analysis to reduce losses in power distribution systems. OPF is a powerful mathematical tool used in the optimization of power systems, and sensitivity analysis helps

Copyright to IJARSCT
www.ijarsct.co.in

DOI: 10.48175/568

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open Access, Double Blind, Peer Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, July 2023

Electric Vehicles with Battery Storage based DC Charging Station using Solar

Taruendrakumar Tiwari[1] and Prof. Jagdish Koli[2]
Department of Electrical and Electronic Engineering[1,2]
NRI Institute of Information Science and Technology, Bhopal, India,
devshahaeee@gmail.com

Abstract: *In this article, a renewable energy based Electric vehicle charging station (EVCS) is proposed which provides electricity to EV with the support of Battery storage system (BSS). Among all renewable energy sources, the solar PV system is best option because of abundance and easy to operation. However solar PV power fluctuates because of change in irradiance and temperature, and it cannot generate constant power, therefore to compensate the power fluctuation a standby battery storage system is needed to meet up the power demand and maintain the reliability of the EVCS. Thereby, a DC micro grid system has been developed which consist of BSS along with the solar PV system and electric vehicle battery charger. The charging controllers are operated based on the concept of power balance, and constant current/constant voltage charging. Performance of the charging system is validated with simulation and experimental results.*

Keywords: battery storage system, solar PV system, electric vehicle charging station, electric vehicle battery etc..

I. INTRODUCTION

In It is estimated that by 2022, EVs will be over 35 million in the World. The Indian government has set ambitious targets to accelerate the adoption of electric vehicles (EVs) due to potential of Electric Vehicles to reduce pollution along with many other advantages like high torque, easy speed control, and higher efficiency compared to the conventional ICEs. However, if all IC engines were to be replaced by EVs at large, then their high penetration causes heavy electricity demand to the power grid and the electric grid will collapse not being able to withstand the load demand [12, 13].In case of India-like countries, the main sources of electricity are fossil fuels. Currently (as of March 2020),62.8 % of total electricity of the country is being powered fossil fuel-based (coal, lignite, gas diesel) plants. Adding to this, the transmission and distribution losses in the country on an average is around 22%, and in states such as Assam is as high as 38.2%[20, 21].However, with Demand increasing exponentially and the availability of fossil fuels decreasing at two-fold rate, the gap between the resource and the demand is widening at an alarming rate. One efficient approach to relieve the effect is to integrate local power generation such as RESs [18] into the EV charging infrastructure [16, 17]. Renewable energy installations such as solar energy panels generate zero emissions in their generation of electricity. taken advantage of it to its maximum extent, sunlight focused on the earth for 1 hour could meet energy demands of the whole earth for an entire year [22]. The system implemented in this paper incorporates a PV panel fed through boost converter and MPPT algorithm, bidirectional converter, buck converters and a BESS. While the energy generated by solar PV panel, during low solar irradiance conditions, is insufficient to meet the power demanded by the Electric Vehicle battery. Then BESS meets the required power demand. On the other hand, while the solar power generation is greater than the demand, the BESS stores the excess solar energy. Later, the system has been implemented in MATLAB Simulink to verify the system performance. The design, and performance analysis of the proposed system using experimental studies are discussed in further sections of the paper. Figure 1 shows the block diagram of the proposed system. The boost converter uses MPPT algorithm and extract maximum solar power at all instants. The bidirectional converter helps in charging and discharging of the station battery depending on source - load power balance in addition to maintaining a constant 48 V bus at its higher voltage terminals. The buck converters step

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/568
ISSN 2581-9429 IJARSCT

INTERNATIONAL JOURNAL OF EMERGING TRENDS & TECHNOLOGIES (IJETT)
ISSN - 3599-4213, DOI - 10.55803/ijett
https://ijettjournal.org/

A Review on Sentiment Analysis Using Natural Language Processing and Machine Learning

Md Wasim Reza
M.Tech Scholar
CSE, Department
NIRT Bhopal
wasimreza84@gmail.com

Prof. Vaibhav Patel
Assistant Professor
CSE, Department
NIRT Bhopal
vaibhav.bce@gmail.com

Prof. Anurag Srivastava
HOD
CSE, Department
NIRT Bhopal
anurag.shri08@gmail.com

Abstract: Sentiment analysis, also known as opinion mining, is a rapidly evolving field that focuses on extracting subjective information from textual data. It plays a crucial role in understanding public opinion, sentiment trends, and customer feedback across various domains. This review paper provides an overview of sentiment analysis techniques using natural language processing (NLP) and machine learning (ML) approaches. It explores the fundamental steps involved in sentiment analysis, such as data preprocessing, feature extraction, and classification. Additionally, it discusses the challenges associated with sentiment analysis, including linguistic nuances, sarcasm, and the availability of labeled training data. Various NLP techniques, such as tokenization, stemming, and part-of-speech tagging, are examined in the context of sentiment analysis. Furthermore, this review paper presents an extensive survey of machine learning algorithms commonly used for sentiment classification, including Naive Bayes, Support Vector Machines, and Recurrent Neural Networks. It also highlights the recent advancements in deep learning models, such as Convolutional Neural Networks and Transformer-based architectures, for sentiment analysis tasks. Finally, potential applications and future directions in sentiment analysis research are discussed to inspire further advancements in this field.

Keywords—Opinion Mining, Sentiment Analysis, Text Mining, Natural language Processing, Deep Learning

I. INTRODUCTION

In today's digital era, an enormous amount of textual data is generated daily through social media, online reviews, and other online platforms. Extracting valuable insights from this vast amount of data has become a crucial task, particularly in understanding public opinion, customer sentiment, and market trends. Sentiment analysis, also known as opinion mining, has emerged as a powerful technique to automatically analyze and classify sentiments expressed in text. By leveraging the advancements in machine learning (ML) and natural language processing (NLP), sentiment analysis has gained significant attention and witnessed remarkable progress in recent years.

This review paper aims to provide a comprehensive overview and analysis of sentiment analysis techniques utilizing ML and NLP approaches. Sentiment analysis involves determining the sentiment polarity, such as positive, negative, or neutral, expressed in a piece of text. It has broad applications across diverse domains, including market research, brand reputation management, customer feedback analysis, political sentiment tracking, and social media monitoring.The integration of ML and NLP techniques has proven to be effective in handling the inherent challenges of sentiment analysis. ML algorithms provide the ability to learn from data and make predictions or classifications, while NLP techniques enable the understanding and interpretation of human language. Together, they form a powerful combination for sentiment analysis tasks.This review begins by outlining the fundamental steps involved in sentiment analysis. Data preprocessing techniques, such as tokenization, stemming, and removing stopwords, are discussed, as they are essential for transforming raw text into a suitable format for analysis. Feature extraction methods, including bag-of-words, n-grams, and word embeddings, are explored, as

Copyright 2023 IJETT. ALL RIGHTS RESERVED