

INFORMATION VISUALIZATION FOR INTELLIGENT SYSTEMS

Edited By

Premanand Singh Chauhan,
Rajesh Arya, Rajesh Kumar Chakrawarti,
Elammaran Jayamani, Neelam Sharma,
and Romil Rawat

 Scrivener
Publishing

WILEY

Information Visualization for Intelligent Systems

Scrivener Publishing
100 Cummings Center, Suite 541J
Beverly, MA 01915-6106

Publishers at Scrivener
Martin Scrivener (martin@scrivenerpublishing.com)
Phillip Carmical (pcarmical@scrivenerpublishing.com)

Information Visualization for Intelligent Systems

Edited by

Premanand Singh Chauhan

Rajesh Arya

Rajesh Kumar Chakrawarti

Elammaran Jayamani

Neelam Sharma

and

Romil Rawat



WILEY

This edition first published 2025 by John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA and Scrivener Publishing LLC, 100 Cummings Center, Suite 541J, Beverly, MA 01915, USA

© 2025 Scrivener Publishing LLC

For more information about Scrivener publications please visit www.scrivenerpublishing.com.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

Wiley Global Headquarters

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials, or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read.

Library of Congress Cataloging-in-Publication Data

ISBN 978-1-394-30578-0

Front cover images supplied by Adobe Firefly

Cover design by Russell Richardson

Set in size of 11pt and Minion Pro by Manila Typesetting Company, Makati, Philippines

Printed in the USA

10 9 8 7 6 5 4 3 2 1

Contents

Preface	xvii
1 Analysis of Restaurant Reviews Using Novel Hybrid Approach Algorithm Over Convolutional Neural Network Algorithm with Improved Accuracy	1
<i>K. Abhilash Reddy and Uma Priyadarsini P.S.</i>	
Introduction	1
Related Work	3
Existing Methodology	4
Convolutional Neural Network Algorithm	4
Proposed Methodology	5
Novel Hybrid Approach Algorithm	6
Statistical Analysis	7
Results	7
Discussion	12
Conclusion	14
References	14
2 Forecasting of Product Demand Using Hybrid Regression Model in Comparison with Autoregressive Integrated Moving Average Model	17
<i>Adibhatla Ajay Bharadwaj and M. Gunasekaran</i>	
2.1 Introduction	18
2.2 Materials and Methods	19
Novel Hybrid Regression Model	20
Autoregressive Integrated Moving Average	21
Statistical Analysis	21
2.3 Tables and Figures	22
2.4 Results	22
2.5 Discussion	24
Conclusion	26
References	26

3	Identification of Stress in IT Employees by Image Processing Using Novel KNN Algorithm in Comparison of Accuracy with SVM	29
	<i>C. Srinath and S. Parthiban</i>	
	Abbreviations Used	30
3.1	Introduction	30
3.2	Materials and Methods	31
	K-Nearest Neighbor	32
	Support Vector Machine	32
3.3	Statistical Analysis	33
3.4	Results	33
3.5	Discussions	37
3.6	Conclusion	37
	Acknowledgements	38
	References	38
4	Observing the Accuracy of Breast Cancer Using Support Vector Machine with Digital Mammogram Data in Comparison with Naive Bayes	41
	<i>M.A. Aasiya Banu and K. Thinakaran</i>	
	Introduction	41
	Materials and Methods	42
	Support Vector Machine	43
	Naive Bayes Algorithm	43
	Statistical Analysis	44
	Results	45
	Discussion	46
	Conclusion	47
	References	47
5	Analyzing and Improving the Efficiency of Winning Prediction in Chess Game Using AlexNet Classifier in Comparison with Support Vector Machine for Improved Accuracy	49
	<i>Keerthana P. and G. Mary Valantina</i>	
	Introduction	50
	Materials and Methods	51
	AlexNet	51
	Support Vector Machine	52
	Statistical Analysis	53
	Results	53
	Discussion	55

Conclusion	58
References	58
6 Accurate Prediction of Vehicle Number Plate Segmentation and Classification with Inception Compared over Alexnet	61
<i>E. K. Subramanian and V. Sudharshan Reddy</i>	
6.1 Introduction	61
Organization of Chapter	62
6.2 Relevant Works	63
6.3 Proposed Methodology	63
Inception Algorithm	64
Alexnet Algorithm	65
6.4 Resources and Techniques	66
Tables and Figures	66
6.5 Results and Discussion	69
6.6 Conclusion	70
References	70
7 A Novel Method to Analyze a Server Instance's Performance During a Crypto-Jacking Attack Using Novel Random Forest Algorithm Compared with Logistic Regression	73
<i>K. Mahesh Reddy and F. Mary Harin Fernandez</i>	
Abbreviations Used	74
7.1 Introduction	74
7.2 Materials and Methods	75
7.3 Statistical Analysis	77
7.4 Results	77
7.5 Discussion	78
Conclusion	81
Acknowledgements	81
References	81
8 A Comparative Analysis of Twin Segmentation and Classification Over MultiClass SVM and Innovative CNN: An Innovative Approach	85
<i>Prudhvi Venkata Narasimha Varma R. and Senthil Kumar R.</i>	
8.1 Introduction	86
Statistical Analysis	88
Results	89
Discussion	92
Conclusion	93
References	93

9	Prediction of Yields in Semiconductor Using XGBoost Classifier in Comparison with Random Forest Classifier	95
	<i>Soorya K. and Michael G.</i>	
9.1	Introduction	95
	Results	98
	Discussion	100
	Conclusion	102
	References	103
10	A Robust Medical Image Watermarking Scheme with a Better Peak Signal-to-Noise Ratio Based on a Novel Modified Embedding Algorithm and Spatial Domain Algorithm	105
	<i>P. Hemanth and P. Shyamala Bharathi</i>	
10.1	Introduction	106
	10.1.1 Materials and Methods	107
	10.1.2 Statistical Analysis	107
10.2	Result	108
10.3	Discussion	113
10.4	Conclusion	114
	References	114
11	BER Comparison of BPSK-DCO-OFDM and OOK-DCO-OFDM in Visible Light Communication	115
	<i>C. Chandu Ganesh and B. Anitha Vijayalakshmi</i>	
	Abbreviations Used	116
11.1	Introduction	116
11.2	Materials and Methods	117
11.3	Statistical Analysis	119
11.4	Results	121
11.5	Discussions	124
11.6	Conclusion	125
	References	126
12	Improved Accuracy in Blockchain-Based Smart Vehicle Transportation System Using KNN in Comparison with SVM	129
	<i>Mekalathuru Yuvaraj and K.V. Kanimozhi</i>	
	Abbreviations Used	129
12.1	Introduction	130
12.2	Materials and Methods	131
12.3	Tables and Figures	131

	KNN Algorithm	131
	Support Vector Machine	133
	Statistical Analysis	135
12.4	Results	135
12.5	Discussion	136
12.6	Conclusion	136
	References	137
13	Improvement in Accuracy of Red Blood Cells (RBC), White Blood Cells (WBC), and Platelets Detection Using Artificial Neural Network and Comparison with Hybrid Convolution Neural Network	139
	<i>A. Sai Abhishek and T. J. Nagalakshmi</i>	
13.1	Introduction	140
13.2	Materials and Methods	141
	13.2.1 Statistical Analysis	144
13.3	Results	144
13.4	Discussion	147
13.5	Conclusion	149
	References	149
14	Novel Design of Meta Ring Array Antenna Using FR4 for Biomedical Applications	151
	<i>Thota Lakshmi Deekshitha and R. Saravanakumar</i>	
14.1	Introduction	151
14.2	Related Work	153
14.3	Materials and Methods	153
14.4	Results	155
14.5	Discussions	158
14.6	Conclusion	158
	Abbreviations Used	158
	References	159
15	Review: Recommendation System in Tourism and Hospitality Based on Comparison of Different Algorithms	161
	<i>Abhishek Tiwari and Pratosh Bansal</i>	
15.1	Introduction	162
	15.1.1 Recommendation for Tourism Spots	171
15.2	Literature Review	171
	15.2.1 Collaborative Filtering-Based Recommendation Systems for Tourism	172

15.2.2	Content Filtering-Based Recommendation Systems for Tourism	173
15.2.3	Recommendation System from Neural Network	174
15.2.4	CNN in Tourism Recommendation	174
15.2.5	Use of Semantic Analysis in Tourism Recommendation	175
15.2.6	Tourism Recommendation with Artificial Intelligence	176
15.2.7	Genetic Algorithms for Tourism Recommendations	176
15.2.8	Some Other Algorithms that are Used for Tourism Recommendation	177
15.3	Research Gaps	178
15.3.1	Effect of COVID-19 on Tourism	180
15.4	Conclusion	182
15.5	Future Work	183
	Abbreviations Used	185
	References	185
16	Secure and Reliable Routing for Hybrid Network to Support Disaster Recovery and Management	193
	<i>Sanat Jain, Amit Dangi, Garima Jain and Ajay Kumar Phulre</i>	
	Abbreviations	194
16.1	Introduction	194
16.2	Related Work	199
16.3	Proposed Methodology	202
16.4	Experimental Results	205
	16.4.1 Simulation Parameter	206
	16.4.2 Simulation Result	206
16.5	Conclusion	208
	Acknowledgments	208
	References	208
17	Machine Learning Techniques for Sentimental Analysis	213
	<i>Ghanshyam Prasad Dubey, Sahil Upadhyay and Ayush Giri</i>	
	Abbreviations Used	214
17.1	Introduction	214
17.2	Applications of Sentimental Analysis	217
17.3	Related Work	218
17.4	Existing Methodology	220
17.5	Comparison and Discussion	224

17.6	Conclusion	225
	References	226
18	Design of 40-mm Period, 0.8-Tesla Variable-Gap Pure Permanent Magnet Undulator Magnet in RADIA	229
	<i>G. Mishra, Geetanjali Sharma and Vikesh Gupta</i>	
18.1	Introduction	229
18.2	Undulator Modeling in RADIA	231
18.3	Results and Discussion	236
	Acknowledgment	239
	References	239
19	Predicting Academic Performance of Students: An ANN Approach	241
	<i>Priyanka Asthana and Manish Maheshwari</i>	
	Abbreviations Used	241
19.1	Introduction	242
19.2	Literature Survey	244
19.3	Proposed ANN Model	245
	19.3.1 Data Gathering	245
	19.3.2 Data Preprocessing	247
	19.3.3 Splitter	247
	19.3.4 Build Model	247
	19.3.5 Performance Analysis	247
19.4	Experimental Setup	248
	19.4.1 Environmental Setting	249
	19.4.2 Configuration Settings	249
19.5	Result Analysis	250
19.6	Conclusion and Future Scope	254
	Acknowledgements	254
	References	254
20	A Deep Study on Discriminative Supervised Learning Approach	259
	<i>Garima Jain, Sanat Jain, Harshlata Vishwakarma and Shilpa Suman</i>	
20.1	Introduction	259
20.2	Literature Survey	261
20.3	Introductory Information About Deep Learning and Its Features	263
20.4	Methodology of DL Approaches	266
20.5	Deep Learning Network Structures	270

20.6	Conclusion	275
	References	276
21	AI Medical Assistant Machine Learning Techniques	281
	<i>S. Padmakala</i>	
21.1	Introduction	282
21.2	Literature Review	283
21.3	Data and Methodology	286
21.4	Result and Discussion	288
21.5	Conclusion	292
	References	293
22	Early Schizophrenia Prediction Using Wearable Devices and Machine Learning	295
	<i>R. Deepa and A. Packialatha</i>	
22.1	Introduction	296
22.2	Related Works	298
22.3	Proposed Methodology	300
	Methodology	301
22.4	Results and Discussion	304
22.5	Comparison with Existing Methods	309
22.6	Conclusion	311
	References	311
23	Forecasting the Trends in Stock Market Employing Optimally Tuned Higher Order SVM and Swarm Intelligence	315
	<i>Rahul Maheshwari¹ and Vivek Kapoor²</i>	
	Abbreviations Used	316
23.1	Introduction	316
23.2	Related Work	317
23.3	Proposed Methodology	321
23.4	Result	328
	23.4.1 Performance Analysis	329
23.5	Conclusion	330
	Acknowledgements	330
	References	331
24	Social Media Text Classification Analysis and Influence of Feature Selection Methods on Classification Performance	333
	<i>Vedpriya Dongre and Pragya Shukla</i>	
24.1	Introduction	334
24.2	Literature Review	334

24.3	Proposed Work	337
24.4	Results Analysis	343
24.5	Conclusions	345
	References	346
25	4G Versus 5G Communication Using Machine Learning Techniques	349
	<i>S. Padmakala</i>	
25.1	Introduction	350
25.2	Literature Review	351
25.3	Data and Methodology	352
25.4	4G and 5G Methodology	355
25.5	4G and 5G Algorithm	357
25.6	Conclusion	366
	References	368
26	Design and Development of Programmable and UV-Based Automated Disinfection for Sanitization of Package Surfaces	371
	<i>Padmakar Pachorkar, P. S. Chauhan, Akash Pawar, Anil Singh Yadav and Neeraj Agrawal</i>	
26.1	Introduction	372
26.2	Materials and Methodology	373
26.3	Result and Discussion	375
26.4	Conclusion	377
	Funding	377
	Acknowledgements	378
	References	378
27	Fuzzy-Based Segmentations Performance Analysis for Breast Tumor Detection Using Spatial Fuzzy C-Means Filtering with Preconditions (SFCM-P) Over Bilateral Fuzzy K-Mean Clustering Algorithm (BiFKC)	381
	<i>K. Surya Prakash and D. Sungeetha</i>	
27.1	Introduction	382
27.2	Materials and Methods	383
27.3	Results	385
27.4	Discussion	390
27.5	Conclusion	393
	References	393

28 Analysis of Vehicle Accident Prediction Using GoogleNet Classifier Compared with AlexNet Algorithm to Enhance Accuracy	397
<i>Prakash Dilli, Nelson Kennedy Babu C. and A. Akilandeswari</i>	
28.1 Introduction	398
Organization of Chapter	399
28.2 Significance of GoogleNet and AlexNet for Vehicle Accidents	399
28.3 Related Work	400
28.4 Proposed Methodology	401
GoogleNet	401
AlexNet	402
Statistical Analysis	403
28.5 Results Analysis	403
Tables and Figures	403
28.6 Conclusion	407
References	408
29 Maximizing the Accuracy of Fake Indian Currency Prediction Using Particle Swarm Optimization Classifier in Comparison with Lasso Regression	411
<i>Kishore Kumar R., Nelson Kennedy Babu C. and A. Akilandeswari</i>	
29.1 Introduction	412
29.2 Significance of PCO and Lasso Regression	413
Organization of Chapter	414
29.3 Related Work	414
29.4 Proposed Methodology	416
Particle Swarm Optimization (PSO)	417
Lasso Regression (LR)	417
Statistical Analysis	418
29.5 Result Analysis	418
29.6 Conclusion	422
References	423
30 Convolutional Neural Network Algorithm for Proliferative Diabetic Retinopathy Detection and Comparison with GoogleNet Algorithm to Improve Accuracy	427
<i>P. Srinivasan, R. Thandaiah Prabu and A. Ezhil Grace</i>	
Abbreviations Used	428
30.1 Introduction	428
30.2 Materials and Methods	429

30.3	Statistical Analysis	432
30.4	Results	433
30.5	Discussion	437
30.6	Conclusion	438
	Acknowledgements	439
	References	439
31	Conversational AI – Security Aspects for Modern Business Applications	441
	<i>Hitesh Rawat, Anjali Rawat, Jean-François Mascari, Ludovica Mascari and Romil Rawat</i>	
	Abbreviations Used	442
31.1	Introduction	442
31.2	CAI – Security Threats	442
31.3	Literature Review	449
31.4	Mitigation Strategies	452
31.5	CAI Models	452
31.6	Future Research Directions	455
31.7	Conclusion	457
	References	458
32	Literature Review Analysis for Cyberattacks at Management Applications and Industrial Control Systems	461
	<i>Hitesh Rawat, Anjali Rawat, Anand Rajavat and Romil Rawat</i>	
	Abbreviations Used	462
32.1	Introduction	463
	32.1.1 Available Research and Findings	465
	32.1.2 Research Objectives	466
	32.1.3 Contributions	466
32.2	Literature Survey	467
32.3	Research Techniques	472
	32.3.1 Analysis of Observations	472
	32.3.2 Parameters for Manuscript (Inclusion and Exclusion) IE	473
	32.3.3 Outcome Identification	473
	32.3.4 IE-Qualitative	473
	32.3.5 Statistics and Facts Extraction	474
	32.3.6 Statistics and Facts IE	475
	32.3.6.1 Publications	475
32.4	Observational Values	476

32.5	Analysis	477
32.5.1	What are the OSCMN Applications Focused ICSS- RQ1?	478
32.5.2	Analysis of Disparate CCA-CCIE Techniques and Methods-RQ2?	479
32.5.3	Availability of Datasets with CTLI-Related Statistics- RQ3	480
32.6	CICS -CCSC Future Scope	481
32.7	Future Work	482
	Acknowledgements	482
	References	483
33	Fractal Natural Language Semantics and Fractal Machine Learning Engineering: Cultural Heritage Generative Management Systems	489
	<i>Jean-François Mascari, Ludovica Mascari, Hitesh Rawat, Anjali Rawat and Romil Rawat</i>	
33.1	Introduction	490
33.2	Frameworks, Directions, and Domains	490
33.3	CH-GeMS Architecture	493
33.3.1	Material: “Landscapes, Heritage, and Culture” Interaction System	493
33.3.1.1	Components, Tools, and Contexts	494
33.3.1.2	Interaction Networks	496
33.3.1.3	Networks of Networks	498
33.3.1.4	Networks of Networks of Networks N ³	501
33.3.2	Services Dualities and Dynamic Data–Driven Simulations	501
33.3.2.1	Services Dualities	502
33.3.3	Dynamic Data–Driven Applications Systems	504
33.4	Conclusions	508
	References	509
	About the Editors	511
	Index	513

Preface

The book focusses on advanced computing, or machine intelligence, the ability of a technology (a machine, device, or algorithm) to interact with its surroundings intelligently. This means that the technology can make decisions and take actions that will increase the likelihood that its objectives will be met. In contrast to the natural intelligence exhibited by people, artificial intelligence (AI), sometimes referred to as machine intelligence, is intelligence manifested by machines. The modern world is experiencing a period of paradigm shifts. New technologies have contributed to these shifts in part because they offer high-speed computing capabilities that make complicated machine intelligence systems possible. These advancements are paving the way for the creation of new cyber systems, which employ continually generated data to construct machine intelligence models that carry out specific functions inside the system. While the isolated use of cyber systems is becoming more common, the synchronic integration of these systems with other cyber systems to create a compact and intelligent structure that can interact deeply and independently with a physical system is still largely unanswered and has only been briefly discussed from a philosophical perspective in a few works.

Modern civilisation has undergone many paradigm shifts as a result of technological breakthroughs. These developments brought in immense data creation, cloud data storage systems, near-instantaneous worldwide information exchange, very quick computer capabilities, etc. Additionally, they paved the way for the development of cutting-edge cyber systems that employ systematically created data pipelines to carry out certain tasks. For instance, in certain nations, video surveillance imagery is used to detect criminals or possible criminals using machine intelligence (MI) models. Moreover, autonomous MI systems have applications in the medical field, where they enable prompt detection of infections like COVID-19. The chemical industry also uses a variety of applications.

Analysis of Restaurant Reviews Using Novel Hybrid Approach Algorithm Over Convolutional Neural Network Algorithm with Improved Accuracy

K. Abhilash Reddy and Uma Priyadarsini P.S.*

Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India

Abstract

The goal of this endeavor is to assess restaurant evaluations using a novel hybrid approach method in conjunction with the algorithm known as convolutional neural network (CNN). The study presents a novel hybrid approach that uses deep learning to classify restaurant evaluations as either good or negative. A collection of reviews was compiled in order to assess the efficacy of the proposed method. The hybrid approach algorithm (accuracy = 96.1%) analyzes the reviews and increases the measured accuracy over CNN (accuracy = 92.30%) with a statistically significant value of 0.004. The findings from the assessment conducted on the test dataset indicate that, in comparison to alternative methodologies currently in use, the novel hybrid approach technique yields the most precise reviews.

Keywords: Sentiment analysis, novel hybrid approach, convolutional neural network, deep learning, restaurant reviews, polarity, food

Introduction

In recent years, the growth of restaurants through online platforms has been significant, with websites becoming the primary way for customers

*Corresponding author: umapriyadarsini@saveetha.com

to provide their opinions and assess the quality of restaurants and food services. The sentiment of customers can be inferred from these online reviews, which play a crucial role in shaping the reputation of an eatery. The potential to evaluate customers' opinions and make adjustments is provided by the contact between customers and owners via Internet platforms. Training machines with labeled data provides the benefit of more accurate future analysis of customer sentiment (Young *et al.*, 2018). The significance of this study extracted features from reviews and predicted their sentiment using a mixed deep learning technique. The findings of this study will help business owners by offering insightful information for making decisions. The procedure entails removing text from the web, classifying it, and figuring out how it feels. This contribution comprises a dataset comprising one thousand reviews sourced from Bangladesh. Chinsha T.C. proposed a feature based on an analysis of restaurant reviews (Chinsha and Joseph, 2015). The applications of the analysis of restaurant reviews using the novel hybrid approach algorithm include (Sharif, Hoque, and Hossain, 2019) the following: Improving Restaurant Operations: The analysis of customer feedback can be used by restaurants to identify areas for improvement and make informed decisions about menu offerings, food services, and other aspects of their operations. Customer Segmentation: Customers can be divided into groups according to their opinions and preferences using the hybrid approach, allowing restaurants to tailor their customers (Mohammad, Kiritchenko, and Zhu, 2013).

The examination of restaurant review systems has generated a significant amount of scholarly interest in recent years. On Google Scholar, over 191 papers were published, whereas on IEEE, over 97 papers were published. This research aims to analyze individuals' perspectives about restaurants using an innovative method. By utilizing deep learning architectures, this study seeks to achieve higher accuracy in sentiment analysis of restaurant reviews. The innovative combination of a convolutional neural network (CNN) and a novel hybrid strategy is suggested to address the diversity of recent datasets. A sample of 1000 reviews was collected and preprocessed to structure the unstructured and unlabeled data, with labels assigned as positive or negative. The CNN model learns the representation of words, while the novel hybrid approach learns more nuanced representations specific to the classification task. Hyperparameters were optimized before training the model to improve accuracy. Sentiment analysis is a common approach to predict customer reviews, as shown in previous studies by authors like Gan, who assessed restaurants based on factors like the quality of the food, cost, service, or atmosphere, and its context, and Jia, who created a model for categorizing restaurant reviews (Jia, 2020). Using a set of data of 1060

reviews, author Niphat Claypo developed a sentiment analysis framework using a combination of K-means clustering and the MRF feature selection (Claypo and Jaiyen, 2015). The optimal average K-means accuracy was 71.73%. An opinion mining model with 70% accuracy was proposed by author Sun using a dependency parser and Sentiwordnet (Sun, Luo, and Chen, 2017). Author Soujanya Poria developed the CNN methodology for identifying sentiment meaning using aspect uprooting. A multilayer CNN was used for word tagging according to Young *et al.* (2018).

The research gap is to overcome these limitations, a hybrid approach combining deep learning and sentiment analysis is proposed to classify sentiment polarity as either positive or negative. However, it is important to note that the accuracy of the machine learning models is contingent upon the quality of the training data, and any biases or inaccuracies could result in incorrect predictions. Additionally, the interpretability of some machine learning models, such as deep learning networks, can be limited, making it challenging to comprehend their reasoning. This study evaluates the effectiveness of a new hybrid approach method in comparison to a CNN algorithm for improving the accuracy of emotion polarity categorization in restaurant reviews.

Related Work

In the past five years, between 2017 and 2021, there has been an analysis of restaurant review systems that have been the subject of a sizable number of research articles. There are over 191 papers published on Google Scholar and over 97 papers published in IEEE. Hossain, Sharif, and Hoque (2020) conducted an analysis utilizing 4000 Bengali movie reviews, suggested a sentiment analysis model, and achieved a precision of 88.90% for long short-term memory (LSTM) and 82.42% for SVM. Utilizing the multinomial naive Bayes system that has an accuracy rating of 84%, 2000 Bengali critiques of books were utilized to categorize sentiment opposites into positive and negative categories (Hossein, Hoque, and Sarker, 2021). Sarker (2019) offered an LSTM-based sentiment analysis with an accuracy of 55.23% to categorize 1500 tweets into positive, negative, and neutral groups. The study of Wahid, Hasan, and Alom (2019) introduces a method for sentiment analysis utilizing LSTM (long short-term memory) on a dataset of 10,000 comments on Facebook to divide Bengali content into either positive or negative groups with 95% accuracy.

An LSTM-based algorithm was used to classify attitudes from Facebook tweets, achieving an 85% accuracy on a set of 10,000 Bengali messages.

Previous studies on sentiment analysis in Bengali focused on datasets such as Twitter posts, book reviews, and movie ratings, but these datasets were generally small. There is a dearth of research on sentiment analysis of restaurant reviews in the Bengali language other than this study. A total of 6625 restaurant reviews were gathered for the current study from a variety of online sources, including restaurant pages (1763), groups (1940), and Facebook comments (2922). Furthermore, 2000 restaurant reviews were acquired using the Yelp dataset. Data obtained from February 2020 to June 2020 contained inconsistent reviews. To address this, a filter was designed to exclude duplicates, comments with a minimum of three terms, mixed language evaluations, neutral sentiment evaluations, and reviews containing punctuation, numerals, and emojis. The filter produced a refined dataset of 6435 evaluations, which had been manually annotated by three annotators with 12 to 18 months of expertise in natural language processing (NLP). The annotation process entailed preserving the labels of 2000 evaluations from the Yelp dataset. Mohammad, Kiritchenko, and Zhu (2013) employed Cohen's Kappa to assess the inter-rater agreement among annotators for evaluating the annotation quality. The data exhibit good quality, as indicated by the average Kappa value of 0.81 (Kwok & Yu, 2013).

Existing Methodology

Convolutional Neural Network Algorithm

CNNs are a deep learning technique known for their effectiveness in image identification and may also be applied to text classification tasks like sentiment evaluation of restaurant reviews. In the case of restaurant reviews, the text data are transformed into numerical data using techniques such as tokenization, padding, and one-hot encoding. The CNN model is then trained using the numerical data. During training, the model learns the patterns and relationships between the words and phrases in the reviews and the corresponding sentiment (positive, neutral, or negative). After the model has undergone training, it can be utilized to categorize fresh reviews and forecast the sentiment expressed in the review. Evaluating the model's accuracy can be done by utilizing metrics such as precision, recall, and F1-score.

Overall, using a CNN algorithm for sentiment analysis of restaurant reviews can lead to effective and efficient classification results.

Algorithm Steps

Data preprocessing

1. Load the restaurant review dataset
2. Clean and preprocess the text data
3. Tokenize the text data into sequences of words or phrases
4. Pad the sequences to a fixed length
5. One-hot encode the sequences

Model building and training

6. Define the CNN architecture (number of layers, filter size, etc.)
7. Compile the model by defining the loss function, optimizer, and metrics
8. Train the model on the preprocessed data using a suitable batch size and number of epochs

Model evaluation

9. Evaluate the trained model on a test set of reviews
10. Calculate evaluation metrics such as precision, recall, and F1-score

Model deployment

11. Save the trained model for later deployment
12. Load the saved model and use it to classify new restaurant reviews and predict their sentiment.

Proposed Methodology

The dataset used in this research work is based on the concept of restaurant reviews (Govindarajan, 2014). The dataset used in this research is collected from Kaggle. The sample size was determined through GPower software, where two groups of 10 sets were selected. Using the GPower 3.1 tool, a pre-test strength value was determined with $\alpha = 0.05$ and power = 0.80, the necessary parameters for a test of statistical significance comparing two independent means. The study employed a novel hybrid method and CNN algorithm by utilizing Technical Analysis software. Human or animal samples were not used due to the absence of ethical approval requirements. The hardware configuration comprised an Intel i5 core processor paired with 16 GB of RAM. The software utilized included HTML, Python, Java, Tomcat/Glassfish server, Jupyter Notebook, My SQL database, CSS web technologies, and J2SDK1.5 Java version.

Novel Hybrid Approach Algorithm

The novel hybrid approach algorithm for analyzing restaurant reviews combines multiple techniques from NLP to provide a comprehensive analysis of customer opinions and sentiments. The algorithm uses methods such as sentiment analysis, topic modeling, and entity recognition to identify trends and patterns in the data. The objective is to help restaurant owners and managers make informed decisions about their business and food services by understanding the strengths and weaknesses of their establishment through the analysis of customer reviews. The insights obtained from the study can be used to improve the customer experience and enhance overall satisfaction.

Pseudocode for a Hybrid Approach Algorithm for Analyzing Restaurant Reviews:

Step 1: Preprocessing

Input: Raw restaurant reviews

Output: Cleaned and preprocessed reviews

Remove irrelevant information such as punctuation, stop words, and special characters.

Tokenize the reviews into individual words.

Utilize lemmatization or stemming to simplify words to their most basic form.

Step 2: Sentiment Analysis

Input: Cleaned and preprocessed reviews

Output: Sentiment scores for each review

Utilize an analysis of sentiment tool or algorithm to categorize each review as favorable, negative, or neutral.

Calculate the sentiment score for each review based on the classification results.

Step 3: Topic Modeling

Input: Cleaned and preprocessed reviews

Output: Topics and their distributions in the reviews

Apply LDA (latent Dirichlet allocation) for topic modeling to determine the primary subjects covered in the reviews.

Calculate the distribution of topics in the reviews.

Step 4: Entity Recognition

Input: Cleaned and preprocessed reviews

Output: Entities mentioned in the reviews

Use an entity recognition tool or model to identify entities such as food items, services, ambiance, etc. mentioned in the reviews.

Store the entities and their mentions in a data structure.

Step 5: Combination of Results

Input: Sentiment scores, topic distributions, and entity mentions

Output: comprehensive analysis of the restaurant reviews

Combine the results from the sentiment analysis, topic modeling, and entity recognition steps to obtain a comprehensive analysis of the restaurant reviews.

Use the results to identify patterns and trends in the data.

Step 6: Visualization and Reporting

Input: Comprehensive analysis of the restaurant reviews

Output: Visualizations and reports

Use data visualization techniques to present the results of the analysis in an intuitive and easy-to-understand manner.

Statistical Analysis

In order to statistically analyze restaurant evaluations, SPSS software employs a special hybrid approach strategy that outperforms the naive Bayes approach in terms of accuracy. Accuracy is the dependent variable, and efficacy is the independent variable. A sample size of ten has been used for the analysis. The novel hybrid approach algorithm's accuracy is determined using the statistical T-test analysis.

Results

The hybrid approach algorithm and the CNN algorithm were run independently in Jupyter Notebook using an experiment size of 10.

Table 1.1 displays the expected accuracy and loss of the hybrid approach algorithm.

Table 1.2 represents the CNN's expected accuracy and loss. The statistical values that can be utilized for comparison are computed for each of

Table 1.1 An analysis of accuracy and loss was conducted on a novel hybrid approach algorithm using a sample size of 10. The hybrid approach algorithm achieves a precision rate of 96.10% with a corresponding error rate of 3.90%.

Iteration	Accuracy (%)	Loss (%)
1	96.63	3.37
2	97.46	2.54
3	98.25	1.75
4	94.52	5.48
5	95.66	4.34
6	95.68	4.32
7	95.25	4.75
8	96.25	3.75
9	95.20	4.80
10	96.10	3.90

the 10 data samples together with the associated loss values. According to the findings, the CNN had a mean accuracy of 92.30% and the hybrid approach algorithm had a mean accuracy of 96.1%.

Table 1.3 represents the mean accuracy scores for the CNN and the hybrid approach algorithm. When compared to the CNN, the hybrid approach algorithm's mean value is superior, with standard deviations of 2.08555 and 2.18436, respectively.

Table 1.4 shows the hybrid approach algorithm and CNN-independent sample T-test data with a statistically significant value of 0.004 ($p < 0.05$).

Table 1.5 evaluates the convolutional neural network, also known as CNN [20–25], algorithm's accuracy against that of the hybrid approach algorithm. The hybrid approach algorithm attains an efficacy of 96.10%, whereas the CNN (convolutional neural network) algorithm yields a success rate of 92.30%. In terms of accuracy, the hybrid approach algorithm performs better than the CNN model algorithm.

Table 1.2 A CNN with a sample size of 10 is used to analyze its accuracy and loss. 92.30% precision and 7.70% error rates are attained using the CNN.

Iteration	Accuracy (%)	Loss (%)
1	91.15	9.85
2	94.32	5.68
3	93.46	6.54
4	91.07	9.93
5	93.18	6.82
6	94.74	5.26
7	90.56	9.44
8	90.78	9.22
9	91.44	8.56
10	92.30	7.70

Table 1.3 The group statistics for the novel hybrid model show a mean accuracy of 96.10% with a standard deviation of 1.11. In comparison, the LSTM model has a mean accuracy of 92.30% with a standard deviation of 1.52.

	Group	N	Mean	Std. deviation	Std. error of the mean
Accuracy	Hybrid algorithm	10	96.1000	1.11427	0.35236
	CNN	10	92.3000	1.52996	0.48382

Figure 1.1 illustrates the flow chart of a distinctive hybrid methodology that integrates a CNN (convolutional neural network) to analyze the sentiments of online restaurant reviews.

Figure 1.2 shows the average accuracy difference between the hybrid approach algorithm and the CNN. With a mean of 93.3571, a standard

Table 1.4 Performing a T-test on an independent sample to establish significance and compute the standard error. A p-value of less than 0.05 is considered statistically significant, and a confidence interval of 95% was calculated. The hybrid approach algorithm demonstrates superior performance compared to CNNs, with a statistically significant p-value of 0.004, meeting the criteria for significance at a two-tailed level of $p < 0.05$.

Levene's equality of variances test				T-test for means equality					95% Confidence interval for the difference	
		F	Sig	t	df	Sig. (two-tailed)	Mean difference	Std. error difference	Lower	Upper
Accuracy	Assuming equal variance	0.309	0.014	6.349	18	0.004	3.80000	0.59853	2.54254	5.05746
	No assumption of equal variances			6.349	16.451	0.004	3.80000	0.59853	2.53400	5.06600