



Kristu Jayanti College

AUTONOMOUS

Bengaluru

Reaccredited A++ Grade by NAAC | Affiliated to Bengaluru North University

DEPARTMENT OF COMPUTER SCIENCE (PG)

**PROCEEDINGS OF THE
NATIONAL STUDENT
RESEARCH SYMPOSIUM
2023**

Editors

Dr. Aruna Devi K

Dr. Kumar R



ISBN NO: 978-81-964360-5-6

Proceedings of the National Student Research Symposium 2023

Organized by



DEPARTMENT OF COMPUTER SCIENCE (PG)
KRISTU JAYANTI COLLEGE (AUTONOMOUS)
Kothanur P. O, Bengaluru, Karnataka, India ,560077

October 18, 2023

ISBN NO: 978-81-964360-5-6

Message from Principal



Fr. Dr. Augustine George
Principal

In the world of technology, where innovation is a constant and information knows no boundaries, symposium like this one, serve as vibrant hubs for the exchange of knowledge, ideas, and experiences. This annual event is a testament to the remarkable progress achieved in computer science and its pivotal role in shaping the future of our digital age. As we navigate a rapidly changing technological environment, it is essential for us to come together to discuss and deliberate upon the latest breakthroughs, challenges, and opportunities that lie ahead. Student involvement in research offers a wide range of benefits, from academic and intellectual growth to practical skills development and opportunities for future success. It enriches the educational experience and empowers students to make meaningful contributions to their expertise. Kristu Jayanati College always foster and promote student research by creating an effective platform that facilitates interaction and knowledge sharing among the research community. With this goal in mind, we continuously organize the National Student Research Symposium every year to engage in enlightening discussions, witness groundbreaking research, and connect with fellow professionals who share the passion for the possibilities of technology. I congratulate the Department of Computer Science (PG) for the initiatives taken to bring the researchers across the nation in a common platform. With great pleasure and pride, I welcome all the participants and convey my best wishes for SRS 2023. Our distinguished keynote speakers and session presenters, have worked diligently to bring you cutting edge insights, demonstrating the boundless potential of computer science across various domains. Their collective expertise will undoubtedly inspire and spark new ideas that will help chart the course of our field's future. The knowledge sharing discussions that transcend the boundaries of traditional silos and The connections made and knowledge shared during this conference have the potential to shape the trajectory of your careers and the field itself. As we embark on this journey of exploration and discovery, let us remember that innovation flourishes when diverse perspectives and backgrounds converge. Let us embrace this diversity as a strength and a source of inspiration. Best Wishes

Message from Dean



Dr. Calistus Jude AL
Dean, Faculty of Sciences

I am delighted to learn about the National Student Research Symposium being organized by the Post Graduate Department of Computer Science. This event offers a valuable platform for students to showcase and deliberate on their research endeavors. It promises to be a catalyst for enhancing the research acumen of both the department's students and participants from other institutions, empowering them to fortify their professional skills. I look forward to witness the way the symposium stimulate the cultivation of high quality and innovative research, encourage cross-disciplinary collaboration, and reshape our understanding of traditional and contemporary concepts and processes within the field of information technology. My heartfelt appreciation goes out to the organizing team, the dedicated student coordinators, and all participants, and I extend my best wishes to the symposium for its success.

Message from Chief Convener



Dr. R. Kumar
Head

Department of Computer Science (PG)

Cutting-edge computer education operates at the edges of the current sociotechnical systems in our society. Consequently, there is a growing focus on enhancing a wide range of computer applications across various sectors and objectives. Meeting the demands of the latest information systems necessitates optimal training in computation. The National Student Research Symposium (SRS) 2023 aspires to be one of the most comprehensive gatherings dedicated to the diverse realms of Computer Science and Information Technology. It will unquestionably serve as a platform for academic and industry experts to engage in conversations about the ongoing advancements in this rapidly evolving technological era.

I appreciate the faculty coordinator, student coordinators, and the organizing committee for their meticulous planning and execution of SRS 2023. I have confidence that the participants will gain valuable insights from the symposium. I also eagerly anticipate the contributions made during discussions and publications, as they will lay the groundwork for future advancements in all digital sectors. On behalf of the Organizing and Advisory Committee, it is our distinct pleasure to extend a warm welcome to all delegates, where we collectively strive to advance the field through the sharing of our expertise.

Conference Committee

ORGANISING COMMITTEE

Chief Patron

Fr. Dr. Augustine George, Principal

Advisory Committee

Fr. Lijo P Thomas, Vice Principal & Chief Finance Officer

Fr. Som Zacharia, Director, Infrastructure Planning and Development

Fr. Emmanuel P. J., Director, Kristu Jayanti College of Law

Fr. Jais V Thomas, Financial Administrator

Fr. Joshy Mathew, Director, Library and Information Centre

Fr. Deepu Joy, Director, Student Welfare Office

Chair Person

Dr. Calistus Jude A L, Dean, Faculty of Sciences

Chief Convenor

Dr. Kumar R, Professor & Head, Department of Computer Science PG

FACULTY MEMBERS

Dr. Velmurugan R

Dr. Muruganantham A

Dr. Vinothina V

Dr. Ranjitha M

Ms. Divya M O

Dr. Nirmala Devi K

Dr. Bharathi V

Mr. Prakash K

Mr. Tomin Joseph

Mr. Jeo Joy

Mr. Sevuga Pandian

Mr. Bino Joseph

FACULTY CONVENORS

Dr. Aruna Devi K

Ms. Reeja Jose Pavanathil

STUDENT CONVENORS

Mr. Francis Benildus R

Ms. Janani D

Ms. Jessica John

Table of Contents

Attendance Monitoring System using DWT & HOG based Face Recognition Technique.....	4
Tanmay M Malik, Anoop R, Grupreet Singh, Naveen Kumar R	
Will Quantum Computers Make or Break the Future.....	4
P. Maa Pranav, Ms. R. Padma	
Motion Capture Technology in Animation	5
Qudsia Fathima, Rabiya Fathima	
Study on Advancing Cognitive Neuroscience: Brain Fingerprinting for Enhanced Neurological Research	5
Rishita Sri Kotapati	
Security Issues in the IoT Cloud Environment: A Review	6
Yash Tripathi, Shamitha Reddy, Kelam Hema Thejaswini, Ritish Singh, Dr.Krishnachalitha.K C	
Periodic Function In Ferris Wheel	6
Gowri. P, Karthikeyan. C, Dhansree. S, Soundarya P. P	
Cloud Native Monitoring Application.....	7
Rakesh Das	
Analyzing Real World Impacts on Global Gaming Growth: A Steam Data-Driven Study	7
Sabi Joseph	
An Extensive Case Study on Cloud Computing Services	8
Nasreen Anjum, R Lekha and Dr. Bharathi V	
Truth or Deception: Advancements in Fake News Detection Technology	8
Shaon Kundu, V Prem Kumar	
Advancements in Eye Disease Image Classification using Machine Learning	9
Jenifer Berlin A, Vaishnavi S	
Online Food Order Prediction with Machine Learning	9
Divya Ghodke M	
Credit Card Fraud Detection using Supervised Machine Learning Algorithms.....	10
Lohit S, Jayanth Kumar S	
Movie Recommendation System-Using K-Mean Clustering	10
Sukhdev Yadav , Manish Chomal, Ayush Rawat	
AI-Driven Analysis of Smoking and Drinking Behavior Patterns.....	11
Gunjan Kumari, Sharon K John	
Aviation crashes analysis using machine learning	11
Harini Srija S , Solomon Wilson R	

Employee Attrition & Performance Analysis	12
Akhila M, Harshitha M	
A Lens to the Legislation Governing Credit Card Fraud with Machine Learning	12
Ritheka Maha S, Tejasvi Maha S and Rachana Pemmaiah K	
Online news Popularity Prediction using Machine Learning Algorithms	13
Jassim D, Navya P N	
Flight Delay Prediction using Data Science	13
Bettina K Thomas, Shriiyaa Sridhar	
Sleep Health and Lifestyle Analysis Using Machine Learning and Data Science	14
Khusina M Marak, Ileen Xaxa	
Parkinson’s Disease Detection using Supervised Machine Learning Algorithms	14
K Harikarthik , Sneha Grace Thomas	
Enhancing Customer Retention in Telecom: A Churn Analysis Approach	15
Avin Saji and Rithen Koshy	
Electric Vehicle Analysis using Supervised Learning Algorithms	15
Pranav K H , Pothula Manoj Joseph B	
Fake Currency Detection Using CNN	16
Arathi Johnson and Francis Benildus	
Predicting Liver Failure using Supervised Machine Learning Algorithms	16
Maldon Romarieo Pinto and Jintu Brahma	
Unveiling the Synergy: Analyzing the Complex Relationship between Smoking and Drinking Behaviors Through Physiological Signals	17
Ashish K and Mercia Mary Varghese	
Deciphering Digital Feelings: A Comprehensive Tech Firm Sentiment Analysis	17
Aswin P K and Tomjo Shibi	
Mortality Rate Prediction in ICU	18
Johnpaul K Jeju, Jeswin Jose	
Real or Fake Job Posting Prediction Using Catboost Algorithms	18
Saloni Gaidhane, Varun Govind B	
Machine Learning Enabled Analysis of Hair Loss Risk: Feature Based Likelihood Prediction for Enhanced Healthcare and Dermatological Insights	19
Amal Martin and Joyal Saji	
Ransomware Detection Using Machine Learning Algorithm	19
Piu Gayen	

Customer Personality Analysis using Unsupervised Machine Learning Algorithms.....	20
Sarthak M Bobde, Vephito Desai	
Predicting Email Spamming using Supervised Machine Learning Algorithms.....	20
Sandeep S Naik , Sanjeevini	
Student Performance Analysis Using Machine Learning.....	21
Abba Jason Samuel and Edric Alvin	
Screen Time Analysis – WhatsApp and Instagram.	21
Abitha K P, Ankitha S, Ashwin K M	
Machine Learning-Based Classification and Regression Models for Water Quality Assessment and Pollution Prediction.....	22
Christina Hart and Jessica John	
A Comprehensive Analysis of Predictive Models for Employee Attrition	22
Samson Raja S, Vijay S	
The Impact of Insomnia Spectrum Disorders on Lifestyle using Machine Learning Algorithms	23
Maria Theresa A, Mohana Priya	
Exploring Factors Affecting Restaurant success in Bengaluru: A Zomato Dataset Analysis	23
Rizin PP and Muhammed Fardheen Naushad	
Enhancing Anesthesia Safety with Neuromuscular Monitoring Anomaly Detection using Machine Learning Technique.....	24
Likitha S, Seshu Madhavan, Peter Abir	
Steam Game Analysis Using NLP And Linear Regression.....	24
Mohiuddin Jeelani, Shreyas Satish Halbhavi, Noel Joseph	
Skin Cancer Prediction using CNN.....	25
Abin Xavier, Arya PV	
Patient Readmission Analysis Using Machine Learning Algorithm	25
Basil Lal and Mervin Moses	
Music Recommendation Based on Face Emotion Recognition	26
Akshata Chandrappa, Albin PB and Sudurshan	
Breast Cancer Detection using Data Science	26
Dolkar Yangkyi and Phurailatpam Girish Sharma	
Challenges and Opportunities in Multilingual Sentiment Analysis: Beyond English	27
Ephraim Godfrey, Sohan Immanuel, Sophy Jose and Dr. Bharathi. V	

Attendance Monitoring System using DWT & HOG based Face Recognition Technique

Tanmay M Malik, Anoop R, Grupreet Singh, Naveen Kumar R

Department of Electronics, M S Ramaiah College of Arts, Science and Commerce, MSR Nagar, Bangalore, India

Email: tanmaymallik08@gmail.com

Abstract: Marking attendance in classroom is one of the most crucial tasks in institutions for gauging the performance of students. Furthermore, if the student strength is big, using manual methods to maintain attendance in the classroom is one of the most difficult portions. Several facial recognition systems are currently in use in a variety of fields. This method is still in its infancy in Indian educational institutions, but it is being adopted by a small number of wealthy nations. In this research, a simple facial recognition algorithm is utilized to keep the attendance system running on small datasets. The system was trained using a set of preloaded image samples of people. The preloaded dataset of human image samples was used to train the algorithm. The image is decomposed using DWT to extract a low frequency sub band of the data set for recognition. A HOG-based face recognition technique is then utilized to extract the feature of the low frequency facial data. When a captured image is compared to preloaded facial images, HOG features and predicts attendance.

Keywords: Image processing, Face Detection method, DWT.

Will Quantum Computers Make or Break the Future

P. Maa Pranav, Ms. R. Padma

Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai

Email: maapranav2002@gmail.

Abstract: Quantum Theory is one of the most impactful discoveries that have the power to influence the course of the world in the process of scientific progress during the twentieth century. It has presented a new line of sight in scientific thought and possibilities we never dreamt of a few years ago, Predicted entirely inconceivable situations, and influenced several domains of modern technologies and still remains to be out of the physical reaches of many to even fathom the immense potential of these new generation of computers. In this paper, We describe the possible ways Quantum Computing can make and break the world as we know it today. In a world of rapid advancements and evolutions, We need to have a clear view of what we are up against to make a proper evaluation of the ways to tackle the possible problems at hand, As we go through this fast changing modern age of technological advancements. This is a project report on the general impact of Quantum Computing and Information Processing from a layman's point of view.

Keywords: Quantum Computing, Computation, Future Computation, Modern Quantum Computers.

Motion Capture Technology in Animation

Qudsia Fathima, Rabiya Fathima

Department of Computer Science, Bishop Cotton Women's Christian College, Bengaluru

Email: Qudsiafathima03@gmail.com

Abstract: Motion capture is a pivotal technology that has revolutionized various industries, from entertainment and sports to healthcare and robotics. This paper provides a concise overview of motion capture technology, including its methods, applications, and the associated advantages and disadvantages. Motion capture is a versatile tool with applications in fields such as entertainment, healthcare, and robotics. The advantages of precision and efficiency it offers are balanced by challenges related to cost, equipment, and privacy concerns. Understanding these aspects is crucial for harnessing the full potential of motion capture in various industries.

Keywords: Motion capture, Optical mocap, Non-optical mocap, CGI and mocap, animation, rotoscoping, virtual reality, filmmaking, sensors, VFX, optical capture, non-optical capture.

Study on Advancing Cognitive Neuroscience: Brain Fingerprinting for Enhanced Neurological Research

Rishita Sri Kotapati

CSE student, Alliance University, Bangalore

Email: rishitakotapati@gmail.com

Abstract: By detecting brain-wave reactions to phrases or images associated with crimes that are displayed on a computer screen, Brain Fingerprinting is a novel computer-based technique for conclusively and scientifically identifying the perpetrator of a crime. The theory behind brain fingerprinting technology is that when a person experiences a well-known event, their brain produces a distinctive pattern of brain waves. Stimuli that are relevant in the current context cause an EEG event-related potential known as a P300-MERMER. P300-MERMER reactions to words or images associated with a crime scene, terrorist training, bomb-making expertise, etc. are picked up by BF. The cognitive information processing of BF is measured in order to detect information. Lies, tension, or emotion cannot be detected by BF. The data available or if it is lacking for each individual judgment is determined by BF, together with a statistical confidence level. There have been no incorrect positives or negatives in laboratory or field testing conducted by the FBI, CIA, US Navy, and other organizations. All conclusions reached were true to the letter. A mere 3% of findings were deemed "indeterminate." The use of BF has been authorized in legal proceedings involving criminal offenses. The new technique uses brainwaves to determine if the participant of the test can recall specifics of the incident. Even if the subject willfully conceals the required details, the brain wave issuer will catch him. By detecting brain-wave reactions to phrases or images associated with crimes that are displayed on a computer screen Brain Fingerprinting is a novel computer-based technique for conclusively and scientifically identifying the perpetrator of a crime. The theory behind brain fingerprinting technology is that when a person experiences a well-known event, their brain produces a distinctive pattern of brain waves. Stimuli that are relevant in the current context cause an EEG event-related potential known as a P300 - MERMER. P300-MERMER reactions to words or images associated with a crime scene,

terrorist training, bomb-making expertise, etc. are picked up by BF. The cognitive information processing of BF is measured in order to detect information. Lies, tension, or emotion cannot be detected by BF. The data available or if it is lacking for each individual judgment is determined by BF, together with a statistical confidence level. There have been no incorrect positives or negatives in laboratory or field testing conducted by the FBI, CIA, US Navy, and other organizations. All conclusions reached were true to the letter. A mere 3% of findings were deemed "indeterminate." The use of BF has been authorized in legal proceedings involving criminal offenses. The new technique uses brainwaves to determine if the participant of the test can recall specifics of the incident. Even if the subject willfully conceals the required details, the brain wave issuer will catch him.

Security Issues in the IoT Cloud Environment: A Review

Yash Tripathi¹, Shamitha Reddy¹, Kelam Hema Thejaswini¹, Ritish Singh¹, Dr.Krishnachalitha.K C²

¹Department of Computer Science & Engineering, Alliance College of Engineering & Design (ACED), Alliance University, Bangalore

²Assistant Professor, Department of Computer Science & Engineering, Alliance College of Engineering & Design (ACED), Alliance University, Bangalore

Abstract: Significant data storage and security issues have come up across organizations because of the Internet of Things (IoT) widespread implementation and the subsequent boom in data. Existing methods use cloud servers for distant data storage to get around troubles with translation. The Internet of Things (IoT) paradigm is made possible by the scalability, dependability, and on-demand services provided by cloud computing. These systems do, however, include honest but curious characteristics that might make data vulnerable to attackers. A storage system should disperse data among various organizations to prevent illegal modification provide resilience for data recovery and enable quick data retrieval and verification in order to improve IoT data security. The paper aims on Identification of systematic literature review technique, comparison of the chosen sources, and performance analysis of the comparison will be the main goals.

Periodic Function In Ferris Wheel

Gowri. P¹, Karthikeyan. C¹, Dhansree. S², Soundarya P. P²

¹ Assistant.Professor,Department of Mathematics, Sri Krishna Arts and Science College, Coimbatore

² III B.Sc.Mathematics, Department of Mathematics, Sri Krishna Arts and Science College, Coimbatore

Abstract: The periodic function for the Ferris wheel can be expressed using trigonometric functions such as sine and cosine, which have a repeating wave-like pattern. The function takes into account the height of the passenger above the ground and the angle of rotation of the wheel at any given time. By using the periodic function, it is possible to calculate the position of each passenger on the Ferris wheel at any time during the ride. In this paper we will explore how periodic functions can be used to accurately model the movement of ferris wheels and how this can be used to predict and analyze the ride experience for passengers.

Cloud Native Monitoring Application

Rakesh Das

PG Student, Department of Computer Science, Presidency College, Bengaluru

Abstract: Cloud-native monitoring applications have become essential tools in today's dynamic and complex IT environments. As organizations increasingly adopt cloud-native architectures and microservices, traditional monitoring approaches prove inadequate. This abstract provides an overview of a cloud-native monitoring application designed to address these challenges. Cloud-native monitoring applications leverage containerization, orchestration platforms like Kubernetes, and serverless computing to monitor applications and services in a scalable and efficient manner. The Key features of a cloud-native monitoring application include Auto-Discovery, Scalability, Metrics Collection, Alerting and Notification, Distributed Tracing, Log Aggregation: These tools aggregate and analyze logs from various sources, making it easier to troubleshoot issues and gain insights into application behavior, Security Monitoring, Dashboard and Visualization and Integration with DevOps Tools. In summary, cloud-native monitoring applications are vital for ensuring the reliability, performance, and security of modern cloud-native applications.

Analyzing Real World Impacts on Global Gaming Growth: A Steam Data-Driven Study

Sabi Joseph

PG Student – MS CS, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcs21@kristujayanti.com

Abstract: The global gaming community has witnessed a surge in recent years, with approximately 3.09 billion active gamers worldwide. This growth, exceeding 1 billion players in just seven years, representing a remarkable 32% increase, is projected to reach 3.32 billion by 2024. Simultaneously, this paper explores the intricate relationship between real-world geopolitical events, societal pressures, and gaming preferences. Focusing on Valve's Steam platform, it investigates how external factors such as wars, disease outbreaks, and celebrity media influence player choices. By analyzing player behavior patterns, we aim to reveal the underlying psychological mechanisms influencing gaming preferences during times of heightened societal stress. Our study utilizes a comprehensive dataset of player activity on Steam, examining game download and playtime statistics. Preliminary findings indicate significant fluctuations in player base and game adoption rates corresponding to specific geopolitical and sociocultural events. Through statistical analysis and qualitative assessments, we elucidate the factors driving these fluctuations, shedding light on the emotional and psychological processes at play. Understanding how real-world pressures affect gaming choices contributes to a deeper comprehension of the complex interplay between the digital and physical realms. The findings may have implications for the gaming industry and broader discussions on escapism, coping mechanisms, and the enduring appeal of interactive entertainment in an increasingly turbulent world.

Keywords: Statistical Significance, Big Data Analysis, Time Series Analysis Gaming community, Video gamers, Time Series Analysis, Geopolitical events

An Extensive Case Study on Cloud Computing Services

Nasreen Anjum¹, R Lekha¹ and Dr. Bharathi V²

¹PG Student - MCA, Department of Computer Science PG

²Assistant Professor, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka

Email: 23mcaa33@kristujayanti.com

Abstract: A cloud computing architecture is a type of computer design that makes use of the Internet to supply IT resources, such as platforms, infrastructure, and applications, as a service. It offers the necessary computer and data processing infrastructure. A new system in the computer sector has been developed as a result of the advancements in processing power, adaptability, and capacity brought about by contemporary technology. Many big companies have moved their processing and storage to the cloud due to the benefits of cloud computing. Presently, enormous and well known businesses have switched to cloud computing and moved their processing and storage there. In this paper, we present a general overview of cloud computing and call attention to its services.

Keywords: Cloud Computing, Services, Cloud providers.

Truth or Deception: Advancements in Fake News Detection Technology

Shaon Kundu, V Prem Kumar

PG Student – MS CS, Department of Computer Science PG

Kristu Jayanti College, Autonomous, Bengaluru, Karnataka

Email: 22mcs24@kristujayanti.com

Abstract: Classification of fake news on social media has gained a lot of attention in the last decade due to the ease of adding fake content through social media sites. In addition, people prefer to get news on social media instead of on traditional televisions. These trends have led to an increased interest in fake news and its identification by researchers. This study focused on classifying fake news on social media with textual content (text classification). In this classification, four traditional methods were applied to extract features from texts (term frequency-inverse document frequency, count vector, character level vector, and N-Gram level vector), employing 5 different machine learning and deep learning classifiers to categorize the fake news dataset. The acquired findings demonstrated that bogus news containing textual content may be categorized, particularly when utilizing a convolutional neural network. Using several classifiers, this study produced an accuracy range of 81 to 100%.

Keywords: fake news, text classification, TF-IDF, N-Gram, character vector.

Advancements in Eye Disease Image Classification using Machine Learning

Jenifer Berlin A, Vaishnavi S

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka

Email: 22mcaa71@kristujayanti.com

Abstract: Eye diseases are a significant public health concern, affecting millions of people worldwide. Early and accurate diagnosis of these conditions plays a crucial role in preventing irreversible visual impairment. In recent years, advancements in medical imaging and machine learning have opened up new avenues for improving diagnostic accuracy. The study presents a novel approach for classifying eye disease images by combining Principal Component Analysis (PCA) with machine learning techniques. The proposed methodology begins with the application of PCA to reduce the dimensionality of the input image data while retaining the most important features. The reduced feature set obtained from PCA is then used as input to various machine learning algorithms, including but not limited to, Support Vector Machines (SVM), Random Forest. These algorithms are trained on a diverse dataset of annotated eye disease images, encompassing conditions such as diabetic retinopathy, glaucoma, macular degeneration, and more. The results showcase that the combination of PCA and machine learning contributes to enhanced accuracy and robustness in classifying eye disease images. The reduced dimensionality introduced by PCA not only aids in improving computational efficiency but also assists in mitigating the curse of dimensionality. Moreover, the approach demonstrates its potential for real-world applications, paving the way for more efficient and accurate diagnosis of eye diseases.

Keywords: Machine Learning Algorithms, Dimensionality reduction, KNN, Random Forest Classifier, Decision Tree, Visualization

Online Food Order Prediction with Machine Learning

Divya Ghodke M

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka

Email: 22mcaa16@kristujayanti.com

Abstract: The complete online food delivery system, which includes companies like Swiggy and Zomato will always try to improve them. The primary goal of these companies in serving their customers is to ensure timely food delivery. To expedite the delivery process, these companies pinpoint regions with high demand for online food orders and increase their fleet of delivery partners in those areas. This strategy results in quicker food delivery in regions with a higher volume of orders. We also find the prediction that the customer would order the food again. Here in this paper a comparative analysis of four of the supervised Machine learning Algorithms, Decision Tree Classifier, ANN Classifier, Naïve Bayes Classifier, and Random Forest Classifier is performed. From the results it is assured that ANN performs better than the other two classifiers in terms of accuracy, Precision and Recall.

Keywords: Machine learning algorithms, Decision Tree, ANN, Naïve Bayes Classifier, Random Forest, Accuracy, Precision, Recall.

Credit Card Fraud Detection using Supervised Machine Learning Algorithms

Lohit S, Jayanth Kumar S

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
22mcaa36@kristujayanti.com

Abstract: Financial fraud cases, such credit card fraud, have increased as a result of recent developments in e-commerce and e-payment systems. Therefore, it is essential to create systems that can identify credit card fraud. When using machine learning to detect credit card fraud, features of these frauds must be carefully chosen because they play a significant part in the process. The genetic algorithm (GA) is proposed in this research as the feature selection method for a machine learning (ML) based credit card fraud detection engine. The suggested detection engine employs the ML classifiers Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Artificial Neural Network (ANN), and Naive Bayes (NB) after selecting the optimum features. The suggested credit card fraud detection engine's performance is validated using a dataset created from European cardholders. The outcome proved that our suggested strategy performs better than current systems.

Keywords: Credit card fraud detection; Data mining; Machine learning; Fraud prevention; Anomaly detection; Artificial intelligence; Fraud detection algorithms

Movie Recommendation System-Using K-Mean Clustering

Sukhdev Yadav , Manish Chomal, Ayush rawat
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa38@kristujayanti.com

Abstract: The amount of digital material available now and the number of users have both increased significantly. Additionally, there are now significantly more online data transactions. This has potentially created a problem of information overload, which delays prompt access to the online data that is available. Because there are more users, the amount of data is growing far too quickly. The issue comes when a person must spend a lot of time searching to find the video or information they want. The recommendation engine aids in resolving this issue. By anticipating their past behavior and offering them pertinent information, the recommendation system aids users in finding the preferred item. They are, in essence, an integral component of websites that feature movies and music. After analysing different types of machine learning algorithms (K-Mean Clustering, k-nearest neighbor), there is a clear picture of where to apply which algorithm in different areas of industries such as recommender systems, e-commerce, etc. Then there is an illustration of how implementations and working of the proposed system are used for the implementation of the movie recommender system. Various building blocks of the proposed system such as Architecture, Process Flow, Pseudo Code, Implementation and Working of the System are described in detail.

Keywords: Machine learning algorithm, datasets, k-mean clustering, k- nearest neighbor, attributes

AI-Driven Analysis of Smoking and Drinking Behavior Patterns

Gunjan Kumari, Sharon K John
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa64@kristujayanti.com

Abstract: The worldwide public health challenges stemming from smoking and alcohol consumption, associated with chronic diseases and societal repercussions, necessitate a comprehensive understanding of the factors shaping these behaviors. Leveraging algorithms like ANN, Naive Bayes and ID3 is used in this study aims to analyze the contributors to smoking and drinking patterns. Utilizing a diverse dataset that encompasses demographics, socioeconomic conditions, environmental influences, and psychological aspects, a holistic model is developed to predict and understand these habits. AI algorithms facilitate the identification of risk factors, early warning signs, and behavior patterns, contributing to the refinement of predictive models for policymakers, healthcare professionals, and public health advocates. Additionally, this research sheds light on the intricate dynamics of addiction and behavior change.

Keyword: AI algorithms, Risk factor, Behavioral analysis, Environmental influence.

Aviation crashes analysis using machine learning

Harini Srija S , Solomon Wilson R
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa23@kristujayanti.com

Abstract: Airplanes are widely regarded as the quickest and highly favored means of transportation globally. Safety is paramount, given the vast number of travelers crossing borders daily. Nonetheless, air travel carries inherent risks, particularly the possibility of accidents and plane crashes. The aviation industry prioritizes meticulous precautions to prevent crashes, necessitating an understanding of factors contributing to accidents. Extracting valuable insights from extensive databases is challenging, and data mining is the key to deriving knowledge from raw data. This research focuses to provide insights that can contribute to enhancing aviation safety and preventing accidents in the future. In this paper, the aim is to determine which operators and aircraft types exhibit higher rates of accidents, aiding in the development of targeted safety measures and recommendations within the aviation industry using supervised machine learning algorithms like random forest for significantly enhancing accuracy.

Keywords: Data mining, Aviation crashes, Machine-learning, Flight data, analysis, Random forest, Aircraft safety.

Employee Attrition & Performance Analysis

Akhila M, Harshitha M

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa24@kristujayanti.com

Abstract: In this paper, Attrition investigates the gradual reduction of staff numbers within organizations, primarily stemming from retirements and resignations without immediate replacements. It examines how human resources experts use the term to describe intentional downsizing, with a focus on voluntary attrition when employees depart without being replaced. The study explores a range of attrition causes, including low pay, restricted development opportunities, challenging working conditions, and client or customer attrition due to aging or shifting demographics. The paper delves into the consequences of attrition for businesses, especially when it is strategically employed to reduce employee numbers, often using hiring freezes as a cost-control method to avoid layoffs. It delves into both voluntary and involuntary attrition causes, shedding light on their effects on morale and workplace dynamics. It emphasizes that attrition occurs as the workforce gradually decreases over time due to personal or professional decisions. Furthermore, the paper highlights the challenges posed by employee attrition when it outpaces new hires, frequently beyond an employee control. It illustrates this through typical scenarios like the opening of new offices requiring relocation.

Keywords: gradual reduction; intentional downsizing; human resources; shifting demographics; employee attrition.

A Lens to the Legislation Governing Credit Card Fraud with Machine Learning

Ritheka Maha S, Tejasvi Maha S and Rachana Pemmaiah K

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa54@kristujayanti.com

Abstract: This research paper provides a thorough examination of the regulations governing credit card fraud, with a focus on the critical role that machine learning techniques play in its prevention. Credit card fraud is a significant concern for financial institutions, consumers, and law enforcement organizations worldwide. In reaction to the increasing sophistication of credit card fraud schemes, politicians have established a slew of rules and regulations aimed at putting a stop to these illegal acts. This research assesses the effectiveness of existing legal frameworks in deterring and penalizing credit card fraud perpetrators. It looks at how financial institutions and credit card firms use machine learning models and algorithms to detect and reduce fraudulent transactions in real time. The study emphasizes the significance of these technologies in improving the security and integrity of financial systems, protecting both consumers and businesses. While evaluating the legal and technical elements, the study also dives into the ethical dimensions of using machine learning for fraud detection. To guarantee that these preventative actions are reasonable and equitable, privacy concerns, algorithmic bias, and the need for openness and accountability are thoroughly investigated.

Keywords : Section; Credit card fraud; Machine Learning; Prevention; Fraud Detection.

Online news Popularity Prediction using Machine Learning Algorithms

Jassim D, Navya P N

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa26@kristujayanti.com

Abstract: Predicting popularity of Online News has become inevitable and crucial for media professionals, marketers, and content providers to make predictions about the popularity of news stories. A strategy to predict the popularity of online news stories is presented in this paper using machine learning techniques. The proposed approach makes use of text content, metadata, temporal properties, and social media interactions from news stories. Likes, shares, and comments on social media are used to simulate the popularity of news stories. Random Forest, Ada Boosting and Logistic Regression algorithms are used to construct predictive models. The effectiveness of the suggested strategy will be assessed using a large datasets of online news with related attributes and popularity metrics. Key insights regarding the parameters affecting news article popularity are revealed here. This involves identifying key elements, time trends, content categories, themes, and patterns of user interaction. It also clarifies how social media, writers, sources, and geographic differences affect online news popularity. To increase readership and user engagement, these insights can influence editorial decisions, recommendation systems, and content initiatives.

Keywords: Popularity Prediction, Logistic Regression, News Articles, Social Media, Re tweets.

Flight Delay Prediction using Data Science

Bettina K Thomas, Shriiyaa Sridhar

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa6@kristujayanti.com

Abstract: The "Flight Delay Prediction" is a powerful tool designed to improve air travel experiences by forecasting flight delays with high accuracy. This theory utilizes historical flight data, machine learning techniques, and predictive analytics to provide valuable insights to both airlines and passengers. The theory will benefit a variety of stakeholders, including passengers, airlines, and airport operators. Passengers will be able to use the model to make informed decisions about their travel plans, such as whether to book an earlier flight or to pack extra supplies in case of a delay. Airlines will be able to use the model to take proactive measures to reduce delays, such as by rescheduling flights or by pre-positioning spare aircraft. Airport operators will be able to use the model to better manage their resources and to provide more accurate information to passengers about flight delays. The theory will be conducted in several phases. First, the historical flight data will be cleaned and prepared for analysis. Next, a variety of machine learning algorithms will be evaluated to select the best model for predicting flight delays. The selected model will be trained and evaluated on a held-out test set. Finally, the model will be deployed to a production environment so that it can be used to predict flight delays in real time.

Keywords: Data mining; Performance Metric; Machine Learning Algorithms; Flight Fare Detection.

Sleep Health and Lifestyle Analysis Using Machine Learning and Data Science

Khusina M Marak, Ileen Xaxa
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa25@kristujayanti.com

Abstract: Sleep quality is essential for overall health and well-being, but it can be disrupted by a variety of factors. Accurate prediction of sleep quality can enable timely interventions to improve sleep and prevent adverse health consequences. Machine learning is a powerful tool that can be used to predict sleep quality based on a variety of data sources. This study aims to develop personalized machine learning models for predicting sleep quality using health records and lifestyle surveys. The findings of this study have the potential to improve sleep hygiene and improve quality of life for individuals with sleep disturbances.

Keywords: sleep quality, machine learning, health records, lifestyle surveys, sleep hygiene,

Parkinson's Disease Detection using Supervised Machine Learning Algorithms

K Harikarthik , Sneha Grace Thomas
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa25@kristujayanti.com

Abstract: Parkinson's Disease is a long-term condition that has a progressive impact on the nervous system and the areas of the body that are regulated by nerves. The onset of symptoms typically begins with a mild tremor in one hand. Although tremors are the most common symptom, the disorder can also lead to stiffness or a decrease in movement. The field of Parkinson disease prediction is a field of active research in the healthcare sector and machine learning. Despite the fact that Parkinson's disease is not widely known globally, its adverse effects are significant and should be taken into account. Additionally, due to the fact that people are often preoccupied with their day-to-day lives, they often overlook the early signs of the condition, which may become more pronounced as the disease progresses. There are a variety of techniques for predicting Parkinson's disease. While there is no cure for Parkinson's disease, medications may improve symptoms significantly. The main aim of this paper is to accurately predict the occurrence of the disease among the individuals by using supervised machine learning algorithms such as ID3, AdaBoost and also a hybrid approach comprising of the above two with modified parameters so as to gain valuable insights out of the dataset to successfully predict the disease in a better way using various variations related to audio signals emitted by the patients, and categorise them as healthy or not. The models have been trained using the above stated algorithms with the dataset taken from the University of California at Irvine (UCI) machine learning repository. ID3 performed with an overall accuracy of 97.59%, AdaBoost performed with an overall accuracy of 97%, and the hybrid approach yielded an overall accuracy of 98.56%. The experimental results of this research imply that the proposed method can be used to reliably predict PD and can be easily incorporated into healthcare for diagnosis purposes.

Keywords- Parkinson's disease, tremor, machine learning, AdaBoost, ID3

Enhancing Customer Retention in Telecom: A Churn Analysis Approach

Avin Saji and Rithen Koshy

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa10@kristujayanti.com

Abstract: The telecommunications industry faces a critical challenge of subscription churn, causing revenue loss and diminished customer base. This study comprehensively analyzes subscription churn, considering dimensions such as customer demographics, service usage patterns, satisfaction, and competition. Leveraging data analytics and the XGBoost (XGB) classifier, key churn predictors are identified and predictive models are built. The study also evaluates retention strategies like personalized offers and service quality improvements. The findings provide actionable insights for optimizing retention efforts, enhancing customer satisfaction, and mitigating churn, crucial for sustainable growth in this highly competitive market.

Keywords: Subscription churn, telecommunications, customer retention, data analytics, XGBoost classifier, predictive modeling, customer satisfaction, retention strategies.

Electric Vehicle Analysis using Supervised Learning Algorithms

Pranav K H , Pothula Manoj Joseph B

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa52@kristujayanti.com

Abstract: This code presents an exploratory data analysis (EDA) and modeling approach for electric vehicle (EV) data. The dataset contains information about various aspects of EVs, including their make, model, price, and range. The code begins by importing necessary libraries and loading the dataset, followed by data preprocessing steps such as handling missing values and outlier detection. The EDA section includes data visualization using Matplotlib and Seaborn, exploring the distribution of variables, and examining the geographical distribution of EVs across counties. Additionally, market share analysis, growth trends over the years, and clustering using K-means and isolation forest techniques are performed. Two predictive models are built: one for predicting electric vehicle range based on the model year and base MSRP, and another for classifying electric vehicle types using a decision tree classifier. Model accuracy is evaluated for classification. The code serves as a comprehensive analysis of EV data, offering insights into the market, trends, and predictive modeling possibilities.

Keywords: Exploratory Data Analysis; Data Preprocessing; clustering K-means; Isolation Forest Techniques; Comprehensive Analysis.

Fake Currency Detection Using CNN

Arathi Johnson and Francis Benildus

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka

Email: 22mcaa19@kristujayanti.com

Abstract: The detection of fake currency is an important concern for financial institutions, businesses, and society at large. This study presents an approach to fake currency detection using CNN, aiming to develop an automated system capable of identifying genuine currency from fake ones. The process involves collecting a complete dataset that gives both genuine and fake currency samples. Relevant features are extracted from the dataset, which includes various characteristics and patterns specific to genuine and fake notes. Convolutional Neural Networks (CNNs), is trained and evaluated to identify the most effective approach for fake detection. The system is tested strictly using unseen data, measuring its accuracy, precision to ensure robust performance. The results describe the potential of machine learning in effectively identifying fake currency, contributing to develop security measures and a safer financial environment. The proposed approach signifies a proactive step toward reduce the risks associated with fake currency, thereby encourage trust and stability within the financial ecosystem.

Keywords: Social Analysis, Machine Learning Algorithms, Currency Detection

Predicting Liver Failure using Supervised Machine Learning Algorithms

Maldon Romariego Pinto and Jintu Brahma

PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka

Email: 22mcaa37@kristujayanti.com

Abstract: Viral hepatitis stands as the primary cause of liver diseases, with liver cancer emerging as the leading cause of cancer-related fatalities. Regrettably, liver cancer is often diagnosed at advanced stages, which poses substantial challenges to effective treatment. This study harnessed the capabilities of deep learning (DL) models to forecast the onset of liver cancer within a cohort of hepatitis patients. The research initiative began by analyzing a substantial sample of one million individuals from Taiwan's National Health Insurance Research Database (NHIRD). The focus was on patients with viral hepatitis, encompassing the years from 2002 to 2010. Using DL models, the study aimed to predict instances of liver cancer by scrutinizing the medical histories of this hepatitis patient group. The results uncovered a concerning surge in the annual prevalence of hepatitis in Taiwan, steadily increasing from 2002 to 2010, with an average annual percentage change (AAPC) of 5.8% (95% CI: 4.2-7.4). On a more positive note, a declining trend was observed among younger individuals aged 16 to 30 years, with an AAPC of -5.6 (95% CI: -8.1 to -2.9). Remarkably, when it came to predicting cases of liver cancer, the deep learning models, particularly the convolutional neural network (CNN) model, outperformed others, boasting an impressive accuracy of 0.980 (AUC: 0.886). In summary, this study has unearthed an alarming increase in hepatitis prevalence over the years in Taiwan, while also shedding light on a promising decrease among young individuals. The utilization of the CNN model for liver cancer prediction in hepatitis patients demonstrates its remarkable accuracy and potential for early detection.

Keywords: Liver failure, machine learning, causes, prediction.

Unveiling the Synergy: Analyzing the Complex Relationship between Smoking and Drinking Behaviors Through Physiological Signals

Ashish K , And Mercia Mary Varghese
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa41@kristujayanti.com

Abstract: Smoking and drinking are prevalent behaviors with known health implications. This study investigates the interplay between these behaviors using physiological signals as a novel approach, aiming to uncover potential predictive patterns. Physiological signal data was gathered from a diverse sample and machine learning classifiers, including Decision Trees, Random Forests, K-Nearest Neighbors, and Gaussian Naive Bayes, were applied to model the relationship between physiological signals and smoking and drinking behaviors. The analysis uncovered distinct patterns in the physiological data that are associated with both smoking and drinking behaviors. This research not only advances understanding of the intricate relationship between smoking and drinking but also highlights the potential of physiological signals as valuable predictors. Using several classifiers, this study produced an accuracy of range 80 to 100%.

Keywords: Smoking, Drinking, Physiological Signals, Machine Learning, Health Behavior, Predictive Models

Deciphering Digital Feelings: A Comprehensive Tech Firm Sentiment Analysis

Aswin P K , And Tomjo Shibi
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa70@kristujayanti.com

Abstract: In today's digital age, the vast amount of textual data generated on various online platforms, including social media, product reviews, and discussion forums, provides a valuable source of insights into public sentiments. Sentiment analysis, a sub field of Natural Language Processing (NLP), is the process of identifying and classifying views expressed in the form of text, to understand the writer's view or feelings towards a particular subject or product, etc., in a negative, positive or a neutral way. In this paper, we present a comparative study of three prominent sentiment analysis models - Logistic Regression, Random Forest Classifier, and Support Vector Machine (SVM) - within the context of Social Sentiment Analysis, and analyze which model gives the best result. Our analysis specifically focuses on customer sentiments regarding various tech firms engaged in the production and sale of mobiles, computers, laptops, and related technologies. Various challenges have been explored and identified in sentiment analysis in the context of product reviews.

Keywords: Sentiment Analysis, Machine Learning, Social Media, Natural Language Processing, Logistic Regression, Support Vector Machine, Random Forest Classifier.

Mortality Rate Prediction in ICU

Johnpaul K Jeju, Jeswin Jose
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa30@kristujayanti.com

Abstract: Predicting Mortality Rates in the Intensive Care Unit (ICU) presents a critical endeavor within healthcare. The ICU is dedicated to patients facing life-threatening conditions, where the ability to forecast mortality rates accurately becomes paramount. The challenge lies in the myriad of factors contributing to variable mortality rates within the ICU, necessitating precise predictions to guide healthcare decisions. To tackle this issue, this research leverages health records and data science techniques. By emphasizing personalization, the research employs the cutting-edge tool of Artificial Neural Networks (ANN) to enhance prediction accuracy significantly. The ultimate objective is to revolutionize ICU care by optimizing resource allocation and decision-making, ultimately leading to lives saved. This research bridges healthcare and data analytics, contributing to the ever-evolving healthcare landscape, where precision predictions play a pivotal role in shaping critical decisions and patient outcomes.

Keywords: Mortality Rates, Intensive Care Unit (ICU), Prediction Accuracy, Resource Allocation.

Real or Fake Job Posting Prediction Using Catboost Algorithms

Saloni Gaidhane, Varun Govind B
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa72@kristujayanti.com

Abstract: The rapid digitalization of job markets has ushered in a new era of opportunity and convenience for job seekers. However, this transition has also given rise to a concerning challenge - the proliferation of fraudulent job postings. These deceptive listings not only pose financial risks to job seekers but also threaten their personal well-being. To address this pressing issue, we present a comprehensive study that harnesses the power of machine learning and natural language processing (NLP) techniques. Our primary objective is to develop classification models capable of accurately identifying fraudulent job descriptions within a dataset of 2,430 job postings, of which approximately 866 have been flagged as fraudulent. The research journey is multifaceted, commencing with meticulous data pre-processing, including handling missing values, feature selection, and outlier management. The critical phase of splitting the dataset into training and testing sets ensures robust model evaluation. Our approach integrates state-of-the-art machine learning algorithms, with a particular focus on the CatBoost algorithm, renowned for its effectiveness in managing imbalanced datasets and handling categorical features. Experimental results demonstrate the model's remarkable ability to distinguish between legitimate and deceptive job descriptions. With an impressive testing dataset accuracy of 98.0%, the model exhibits its proficiency in this task. Precision and recall metrics further corroborate the model's excellence, with a low false positive rate of 2.0%, indicating a minimal false alarm rate in detecting deceptive postings. In conclusion, our study represents a significant step toward addressing the issue of fraudulent job postings and fostering a secure digital job market. Future research can further enhance the accuracy of detection and explore more sophisticated approaches to safeguard job seekers in the ever-evolving landscape of the digital job.

Keywords: Fraudulent job postings, Machine learning, CatBoost algorithm, Data visualization

Machine Learning Enabled Analysis of Hair Loss Risk: Feature Based Likelihood Prediction for Enhanced Healthcare and Dermatological Insights

Amal Martin and Joyal Saji
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa04@kristujayanti.com

Abstract Hair loss is a widespread concern transcending age and gender, impacting self-esteem and well-being. In recent years, data science and machine learning techniques have emerged as valuable tools for identifying contributing factors and constructing predictive models for assessing the risk of hair loss. Factors such as dandruff severity, stress, pressure levels, hair greasiness, coffee consumption, brain activity duration, sleep patterns, swimming habits, and libido levels have been recognized as potential contributors to this issue. Employing machine learning algorithms, we predict the likelihood of hair loss based on these features, refining models for maximum accuracy. This research has the potential to provide individuals with a reliable risk assessment tool and offers insights into the complex interplay of these factors in the fields of dermatology and healthcare, ultimately improving the lives of those affected by this common condition through the power of data science.

Keywords: hair loss, prediction, models, algorithms, dermatology, healthcare

Ransomware Detection Using Machine Learning Algorithm

Piu Gayen
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa50@kristujayanti.com

Abstract: Ransomware attacks have emerged as a pervasive and persistent cybersecurity threat, causing significant economic and operational disruptions. Detecting ransomware early is crucial to prevent data loss and financial damage. This research explores the application of machine learning algorithms for ransomware detection. Leveraging various data sources, including system logs, network traffic, and user behavior, machine learning models are trained to distinguish between normal system activity and ransomware-related behavior. The study evaluates the effectiveness of different machine learning algorithms and feature engineering techniques in identifying ransomware patterns. The results show that machine learning-based ransomware detection systems can provide accurate and timely alerts, enabling proactive responses to mitigate the impact of attacks. By continuously adapting to evolving ransomware tactics, these models play a vital role in strengthening organizational cybersecurity defenses and safeguarding data integrity. This research contributes to the development of proactive measures against ransomware threats and reinforces the significance of data-driven security approaches in today's digital landscape.

Keywords: Machine learning, Catboost algorithm, Cybersecurity

Customer Personality Analysis using Unsupervised Machine Learning Algorithms

Sarthak M Bobde, Vephito Desai
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa73@kristujayanti.com

Abstract: In today's data-driven business landscape, comprehending customer behavior becomes essential for organizations striving to secure a competitive advantage. This research paper presents an extensive study on the clustering of customers grounded in their purchasing behavior. The principal objective is the categorization of customers into discernible groups, enabling businesses to adapt marketing strategies, optimize product offerings, and enrich customer engagement. The study makes use of a robust dataset comprising customer attributes and their purchase history. Through a methodical approach to data preprocessing, feature engineering, and clustering techniques, the intention is to unearth concealed patterns and insights within the customer base. A variety of clustering algorithms, encompassing k-means, hierarchical clustering, and DBSCAN, are employed to ascertain customer segments grounded in their preferences, purchase frequency, and spending patterns. Additionally, the research delves into the pragmatic applications of customer clusters in real-world scenarios. The implications of these clusters on the effectiveness of marketing campaigns, product recommendation systems, and customer retention strategies are explored. By aligning business objectives with data-driven insights, organizations have the opportunity to elevate customer satisfaction and profitability. The findings of this research make a notable contribution to the burgeoning field of customer analytics, providing valuable guidance to businesses keen to harness the potential of data for optimizing their customer-centric strategies. Ultimately, the capability to cluster customers effectively based on their purchasing behavior empowers organizations to personalize their interactions, cultivate customer loyalty, and foster sustainable growth.

Keywords: Customer Personality analysis, Unsupervised machine learning, clustering, k-means algorithm, Data analytics.

Predicting Email Spamming using Supervised Machine Learning Algorithms

Sandeep S Naik , Sanjeevini
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa60@kristujayanti.com

Abstract: Email attacks are the one emerging now-a-days and are greatest threat to National Security. Many of the attackers are targeting high end audience and critical infrastructures. It is thus essential to identify and eliminate users and machines misusing e-mail service. There is a need to find a technique by which user can differentiate between the genuine and spoofed Email. As there are many ways to spoof an Email but this paper mainly focuses on mails spoofed by using either PHP Script or Anonymous Mailers. This paper represents machine learning techniques by which user can find whether the mail is spoofed or not.

Keywords: Spam filters, machine learning, causes, prediction

Student Performance Analysis Using Machine Learning

Abba Jason Samuel and Edric Alvin
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa01@kristujayanti.com

Abstract: Predicting student performance is a crucial task in the field of education that can aid in identifying students who may require additional support and intervention. In this research, an approach using machine learning is used to predict a student's performance based on various parameters. The framework used for this project encompasses data preprocessing for cleaning the dataset, model building, and custom input for personalized predictions. The data used for this study is sourced from student records, which may contain missing values, empty rows, and duplicate entries. To ensure data quality, a data cleanup process is employed, including the removal of empty rows, interpolation to fill missing values, and the elimination of duplicate entries. Categorical variables are transformed into numerical features through one-hot encoding. The dataset is split into training and testing sets to facilitate model evaluation. The predictive model is built using linear regression, leveraging features such as extracurricular activities, study environment, procrastination level, class participation, social support, note-taking habits, and online collaboration. Evaluation metrics, including Mean Squared Error (MSE) and R-squared, are utilized to assess the model's performance on test data. The trained model demonstrates its ability to predict student performance with reasonable accuracy. This research employs machine learning to predict student performance, addressing data quality issues and using linear regression for effective predictions. Evaluation metrics like MSE and R-squared demonstrate the Model's accuracy. It's a significant step towards data-driven student support systems in education.

Keywords: Student Performance Prediction, Machine learning, Data Preprocessing, Linear Regression, Categorical Variables.

Screen Time Analysis – WhatsApp and Instagram.

Abitha K P, Ankitha S ,Ashwin K M
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcab01@kristujayanti.com

Abstract: The increasing use of smartphones has led to concerns about screen time and its impact on our lives. It involves the process of examining and evaluating the amount of time individuals spend on electronic devices, such as smartphones, tablets, computers, or televisions. It involves tracking and analyzing the duration and frequency of screen-based activities, including the use of specific applications. This study is to analyze screen time data for WhatsApp and Instagram to gain understandings into how these apps are used and how they can be managed more effectively.

Keywords: Screen Time, WhatsApp, Instagram, Apps, Opened, Notifications, Usage.

Machine Learning Based Classification and Regression Models for Water Quality Assessment and Pollution Prediction

Christina Hart and Jessica John
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa29@kristujayanti.com

Abstract: Water quality is a global concern, with millions of lives at risk due to contaminated drinking water. It is also a growing concern about water pollution and its impact on public health and the environment, accurate and timely prediction of water quality is of paramount importance. This paper addresses the critical problem of assessing water portability using machine learning techniques. A comprehensive analysis of four different models, including Decision Tree, Random Forest, K- Neighbours, and SVM in order to predict water potability accurately. The work contributes to the development of reliable and accessible tools for safeguarding public health by ensuring access to safe drinking water. This study advances environmental science and water resource management by showcasing machine learnings potential for accurate pollutant classification and water quality prediction. It emphasizes how crucial data-driven strategies are to protecting our water supplies and guaranteeing a sustainable and safe environment.

Keywords: Water quality, Machine learning, Logistic Regression, Decision Tree, Contaminated drinking water, Water pollution, Public health SVM.

A Comprehensive Analysis of Predictive Models for Employee Attrition

Samson Raja S ,Vijay S
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: samsonrajas03@gmailcom

Abstract: Employee churn has become an increasingly common problem in the modern, dynamic corporate environment, and organizations place a high focus on retaining current employees. In this study, the goal is to apply machine learning models, more especially the Random Forest Classifier and XGBoost Classifier, to effectively forecast employee attrition using a real-world dataset. The data is divided into training (70%) and testing (30%) sets to ensure a thorough model evaluation, and model performance is optimized through hyperparameter tweaking. The effectiveness of the model is assessed using key assessment criteria such as accuracy, F1 score, and mean squared error. For improving staff retention methods in contemporary firms, this study provides insightful information. The strategic imperative of minimizing attrition and fostering a more stable and productive workforce is furthered by this study's insightful recommendations for improving employee retention practices in modern organizations. This study's recommendations will support sustainable growth and success in the cutthroat business environment of today.

Keywords: Machine learning, Extreme Gradient Boosting Classifier, Random Forest Classifier, Employee attrition, Hyperparameter Tuning.

The Impact of Insomnia Spectrum Disorders on Lifestyle using Machine Learning Algorithms

Maria Theresa A , Mohana Priya
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa39@kristujayanti.com

Abstract: Sleep is crucial for overall well-being, impacting physical and mental health significantly. Lifestyle choices strongly influence sleep quality and can lead to sleep disorders. This study aims to create a predictive model for identifying individuals at risk of sleep disorders based on lifestyle factors. Lifestyle changes driven by increased technology usage have contributed to various health issues, with insomnia being a notable concern. Insomnia, or sleeplessness, can become chronic and harm the brain if left untreated. Typically, detecting insomnia requires costly and time-consuming medical tests, which are often inaccessible in developing countries. To address this, we propose an intelligent machine learning model to predict chronic insomnia, offering a more accessible and cost-effective method for identifying at-risk individuals.

Keywords: Sleep health, Machine-learning, Sleep-disorder, Logistic regression, Sleep health.

Exploring Factors Affecting Restaurant success in Bengaluru: A Zomato Dataset Analysis

Rizin PP Muhammed Fardheen Naushad
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa56@kristujayanti.com

Abstract: Bengaluru, India, boasts a flourishing restaurant industry with over 12,000 establishments catering to a diverse clientele. This sector thrives due to the city's status as the IT capital of India, where time constraints often drive the population to dine out. Nevertheless, amidst fierce competition and an array of similar offerings, understanding the factors that determine a restaurant's success becomes paramount. This paper takes a deep dive into the rich source of data provided by Zomato, aiming to extract valuable insights through advanced machine learning techniques. It pursues the following objectives:

This research explores the restaurant scene in Bengaluru through linear regression analysis, random forest regression, XGBoost regression, and the Extra Trees Regressor. These methods, coupled with data visualisation and statistical analysis, provide comprehensive insights for restaurant owners and investors, enabling them to navigate the Bengaluru market more effectively. Furthermore, these results contribute to enhancing the dining experience for residents and visitors by aligning restaurant offerings with customer preferences.

Keywords: Linear Regression, Random Forest Regressor, XGBoost Regressor, Extra Trees Regressor.

Enhancing Anesthesia Safety with Neuromuscular Monitoring Anomaly Detection using Machine Learning Technique

Likitha S, Seshu Madhavan, Peter Abir
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa48@kristujayanti.com

Abstract: This research paper aims to present an enhanced approach to detecting outliers. Anomaly detection also referred to as outlier analysis, revolves around identifying deviations in data patterns from the expected norm. This field holds substantial significance within the realm of machine learning and has seen extensive research and implementation across various research domains and application areas. Neuromuscular monitoring is vital during surgeries where neuromuscular-blocking drugs (NMDs) are administered, affecting approximately 60% of all global surgeries each year. These drugs, often called "curares," paralyze muscles to aid medical procedures. However, the need for precise monitoring is critical to prevent post-operative complications. Neuromuscular monitoring relies on the Train of Four (TOF) electrical stimulation pattern. An international benchmark of TOF 0.9 ensures safe extubation, yet residual neuromuscular blockages still occur at high rates, between 20% to 40%. The challenge lies in artifactual recordings from TOF monitoring devices, which lack a clear explanation, making clinical providers sceptical about their reliability. In this research paper, we adopt a comprehensive machine learning technique to improve the outlier detection method's methodology, techniques, and procedures. The specific domain chosen for the implementation of anomaly detection in this study is neuromuscular monitoring, a subfield within the domain of anaesthesia.

Keywords Anomaly detection; Neuromuscular monitoring; Outlier analysis; Machine learning.

Steam Game Analysis Using NLP And Linear Regression

Mohiuddin Jeelani, Shreyas Satish Halbhavi, Noel Joseph
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa65@kristujayanti.com

Abstract With the rapid growth of the video gaming industry, Steam, a digital distribution platform, has become a hub for players and game developers. This research paper talks about the landscape of Steam, which deals in providing an analysis of player engagements and monetization done with the platform. This research paper provides an insight to the game developers and industry stakeholders. It offers guidance on optimizing the games and making the user have a seamless performance, the engagement of different communities and of course ultimately the success of games on the platform itself. Furthermore, this paper also helps in investigating the role of user generated content and modding which helps in enhancing and making the games more appealing on Steam.

Keywords: Steam Games, Digital Distribution Platform, User generated content, monetization.

Skin Cancer Prediction using CNN

Abin Xavier, Arya PV
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa02@kristujayanti.com

Abstract: Skin cancer is the most common type of cancer, and early detection is essential for successful treatment. Computer-aided diagnosis (CAD) systems can help dermatologists to diagnose skin cancer more accurately and efficiently. This paper compares two skin cancer detection codes: a simple CNN model and a ResNet model, using 'LeakyReLU' as activation function instead of 'relu'. The performance of both models is evaluated on the HAM10000 skin cancer dataset. The ResNet model outperforms the simple CNN model in terms of accuracy and F1 score for the malignant class, achieving an accuracy of 98.2% and an F1 score of 85.6%. This suggests that the ResNet model is better at identifying malignant lesions. Both models are open source codes, which makes them a valuable resource for the development of skin cancer detection systems.

Keywords: Skin Cancer, Convolutional Neural Network

Patient Readmission Analysis Using Machine Learning Algorithm

Basil Lal and Mervin Moses
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa12@kristujayanti.com

Abstract: A hospital readmission is a situation in which a patient who had been released from the hospital is admitted once more within a predetermined time frame. Readmission rates are increasingly being used as an indicator of success in studies of health services and as a standard of excellence for healthcare organizations. Hospital readmissions are viewed as a compounding economic factor for healthcare systems. The ability to predict a patient's readmissions enables timely intervention and improved post-discharge strategies, preventing subsequent life threatening events and lowering medical costs for the patient or the healthcare system. Four machine learning models are employed in this study to predict readmissions. There are a total of 17 independent variables in the dataset, which consists of 25000 actual hospitalization records. According to the experimental findings, the Random Forest classifier, Naive Bayes classifier, the Support Vector Machine (SVM) with an RBF kernel, and the Support Vector Machine with a linear kernel were all assessed. With a sensitivity AUC value, the Random Forest classifier outperformed the SVM models among these models

Keywords: Artificial Intelligence, Machine Learning, Readmission, Random Forest Classification, Naive Bayes classifier, the Support Vector Machine

Music Recommendation Based on Face Emotion Recognition

Akshata Chandrappa, Albin PB and Sudurshan
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcab@kristujayanti.com

Abstract: This paper presents a novel approach for emotion recognition using a Convolutional Neural Network (CNN) trained on the FER 2013 dataset. The system captures real-time facial expressions from a webcam feed and predicts one of seven emotions. Based on the predicted emotion, a music playlist is recommended using Spotifys API through a wrapper, which is displayed on the screen. Our proposed system aims to reduce computational time and overall costs while enhancing the user experience.

Keywords: Machine Learning, CNN.

Breast Cancer Detection using Data Science

Dolkar Yangkyi and Phurailatpam Girish Sharma
PG Student - MCA, Department of Computer Science PG
Kristu Jayanti College, Autonomous, Bengaluru, Karnataka
Email: 22mcaa17@kristujayanti.com

Abstract: Breast cancer is one of the most prevalent and life-threatening diseases among women worldwide. Early detection plays a pivotal role in improving patient outcomes and reducing mortality rates. This research focuses on breast cancer detection, emphasizing the significance of early detection in improving patient outcomes. It employs four machine learning algorithms (Naive Bayes, Decision Tree, Random Forest Classifier, and ID3) to assess their performance in terms of accuracy, sensitivity, specificity, and computational efficiency. A comprehensive dataset, including clinical and demographic attributes, is used for model training and evaluation. The research follows a systematic workflow, involving data preprocessing, feature selection, and model development. Different algorithms leverage various principles to classify breast cancer cases and enhance predictive accuracy. Additionally, ROC curves and AUC values measure the algorithms' ability to distinguish between benign and malignant tumors. The research aims to provide valuable insights into the suitability of these algorithms, aiding healthcare professionals and researchers in making informed decisions for early diagnosis and treatment planning, contributing to ongoing efforts to enhance breast cancer detection. The models have been trained using the above stated algorithms with the dataset taken from Kaggle repository. ID3 performed with an overall accuracy of 95.34%, an overall recall of 90.47%, an overall precision of 100%, and f1-score of 95%. In addition, Decision Tree performed with an overall accuracy of 95.34%, an overall recall of 94.73%, an overall precision of 94.73%, and f1-score of 94.73%. The Naïve Bayes yielded an overall accuracy of 93.02%, an overall recall of 90%, and overall precision of 94.73%, and f1-score of 92.3%. In addition, Random forest Classifier performed the best among the four algorithms with an overall accuracy of 97.67%, an overall recall of 95.83%, an overall precision of 100%, and f1-score of 97.87%. The experimental results of this research imply that the proposed method can be used to reliably predict breast cancer and can be easily incorporated into healthcare for diagnosis purposes.

Keywords: Breast cancer detection, Early detection, Life-threatening disease, Model evaluation, Systematic workflow, Data preprocessing, ROC curves, AUC values.

Challenges and Opportunities in Multilingual Sentiment Analysis: Beyond English

Ephraim Godfrey¹, Sohan Immanuel¹, Sophy Jose¹ and Dr. Bharathi. V²

¹PG Student - MCA, Department of Computer Science PG

²Assistant Professor, Department of Computer Science PG

Kristu Jayanti College, Autonomous, Bengaluru, Karnataka

Email: 22mcaa17@kristujayanti.com

Abstract: With the arrival of the internet, people have actively expressed their thoughts about various items via blogs, social media, and website comments. It becomes clearer and clearer how complicated and advantageous multilingual sentiment analysis of social media is. Sentiment analysis research is expanding incredibly quickly. Sentiment analysis in English alone is insufficient due to the diversity of languages and cultures present on the internet, which has a significant impact on social analytics and social listening. Multilingual sentiment analysis assists businesses in overcoming language barriers and capturing priceless insights in real-time by realizing that sentiment is inextricably tied to language and culture. This paper is a review on multi-language sentiment analysis, which was presented by many studies published over the last decade addressing the difficulties and possibilities.

Keywords: Sentiment Analysis, Multilingual Sentiment Analysis (MSA), Social Media.



Proceedings of the National Student Research Symposium 2023

Editors

Dr. Aruna Devi K, Dr. Kumar R

Published by

Kristu Jayanti College, Autonomous
Kothanur P. O, Bengaluru, Karnataka, India
www.kristujayanti.edu.in
info@kristujayanti.com

Copyright © Kristu Jayanti College, Autonomous (2023)