International Journal of Academic Information Systems Research (IJAISR)
ISSN: 2643-9026
Vol. 7 Issue 12, December - 2023, Pages: 26-38

Predictive Modeling of Obesity and Cardiovascular Disease Risk: A Random Forest Approach

Mohammed S. Abu Nasser and Samy S. Abu-Naser

Department of Information Technology,
Faculty of Engineering and Information Technology,
Al-Azhar University, Gaza, Palestine

Abstract*: This research employs a Random Forest classification model to predict and assess obesity and cardiovascular disease (CVD) risk based on a comprehensive dataset collected from individuals in Mexico, Peru, and Colombia. The dataset comprises 17 attributes, including information on eating habits, physical condition, gender, age, height, and weight. The study focuses on classifying individuals into different health risk categories using machine learning algorithms. Our Random Forest model achieved remarkable performance with an accuracy, F1-score, recall, and precision all reaching 97.23%. The model's success is attributed to its ability to effectively leverage attributes related to eating habits, physical condition, and demographic factors. Feature importance analysis reveals key factors contributing to accurate predictions. This paper contributes to the field of health analytics by demonstrating the effectiveness of Random Forest in predicting and categorizing health risks. The insights gained from this research have the potential to inform personalized health interventions and contribute to the advancement of precision health strategies. Furthermore, the study highlights the importance of leveraging machine learning techniques for health risk assessment in diverse populations.*

Keywords: Random Forest, Obesity, Cardiovascular Disease, Health Risk Assessment, Machine Learning, Precision Health.

Introduction:

In recent years, the rising prevalence of obesity and cardiovascular disease (CVD) has emerged as a critical public health challenge worldwide. As lifestyles evolve and dietary habits undergo transformations, understanding the factors contributing to these health concerns becomes paramount for effective intervention and prevention strategies. This study endeavors to address this imperative by employing a data-driven approach to predict and classify obesity and CVD risk, utilizing a diverse dataset collected across Mexico, Peru, and Colombia [1,2].

The dataset, acquired through a web-based survey platform, encapsulates a rich array of information, including details on eating habits, physical condition, and demographic characteristics of respondents aged between 14 and 61. The utilization of machine learning, specifically the Random Forest algorithm, serves as the cornerstone of our predictive modeling framework. This choice is motivated by the algorithm's proven efficacy in handling complex datasets, capturing non-linear relationships, and providing interpretable insights through feature importance analysis [3-6].

The classification task undertaken in this study holds significant implications for public health. Obesity and CVD are multifaceted conditions influenced by a myriad of factors, including lifestyle, genetics, and environmental variables. A robust predictive model can aid in identifying individuals at risk, allowing for targeted interventions, personalized health strategies, and more effective resource allocation in healthcare systems[7-12].

Furthermore, this research contributes to the growing body of literature on the application of machine learning techniques in health analytics. By exploring the interplay between various attributes such as eating habits, physical condition, and demographic variables, we aim to uncover patterns that inform our understanding of health risks in diverse populations [13-15].

In the subsequent sections, we delve into the methodology employed, present the results obtained, and discuss the implications of our findings. Through this work, we aspire to advance the discourse on precision health, providing insights that can guide policy-making, clinical practices, and the development of personalized health interventions.

Problem Statement:

The escalating rates of obesity and cardiovascular disease (CVD) in contemporary society underscore a pressing public health concern, necessitating effective strategies for early detection and intervention. Despite extensive research into lifestyle-related health risks, the intricate interplay of diverse factors, including eating habits, physical condition, and demographic variables, complicates the accurate assessment of individual health risks [16-18].

Traditional methods of health risk evaluation often rely on linear models and lack the capacity to capture the nuanced relationships inherent in the multifaceted nature of obesity and CVD. This limitation underscores the need for advanced methodologies capable of discerning intricate patterns within expansive datasets to enhance the precision and personalization of health risk assessments.

The objective of this study is to address this gap by leveraging machine learning, specifically the Random Forest algorithm, to develop a predictive model capable of classifying individuals into distinct obesity and CVD risk categories. The complexity of health-related data necessitates a sophisticated approach that can not only handle diverse attributes but also provide interpretable insights into the factors influencing health outcomes [19-22].

By framing the problem as a classification task and utilizing a robust machine learning algorithm, we aim to enhance the accuracy and reliability of health risk assessments. The outcome of this research is expected to contribute valuable insights for healthcare practitioners, policymakers, and researchers, facilitating more targeted interventions and fostering a deeper understanding of the intricate web of factors influencing obesity and cardiovascular disease risks across diverse populations [23-25].
In essence, the problem at hand revolves around the need for advanced, interpretable models capable of discerning complex health patterns to improve the accuracy and effectiveness of health risk assessments in the context of obesity and cardiovascular disease. This research seeks to address this problem by employing a Random Forest model on a comprehensive dataset obtained from Mexico, Peru, and Colombia.

Literature Review:

1. Introduction to Obesity and Cardiovascular Disease: Obesity and cardiovascular disease (CVD) represent two major public health challenges globally. The World Health Organization (WHO) identifies these conditions as leading contributors to morbidity and mortality, emphasizing the critical need for effective preventive measures and intervention strategies. The complex and multifactorial nature of obesity and CVD underscores the importance of exploring advanced methodologies for accurate risk assessment [26].
2. Traditional Approaches to Health Risk Assessment: Historically, health risk assessment has relied on conventional methods, often characterized by linear models and limited incorporation of diverse variables. These approaches, while valuable, lack the capacity to capture the intricate relationships among various factors influencing obesity and CVD. The evolving landscape of healthcare demands more sophisticated tools capable of handling non-linear patterns and adapting to the heterogeneity of health-related data [27-35].
3. Machine Learning in Health Analytics: Recent years have witnessed a paradigm shift in health analytics, with machine learning (ML) emerging as a powerful tool for predictive modeling and risk assessment. ML algorithms, particularly ensemble methods like Random Forests, have shown promise in handling large, complex datasets and providing accurate predictions. The versatility of these techniques makes them well-suited for applications in healthcare, where the interaction of numerous variables requires a nuanced approach [36-45].
4. Random Forests in Health Risk Prediction: Random Forests, a popular ensemble learning technique, have demonstrated success in various domains, including healthcare. Their ability to handle high-dimensional data, capture non-linear relationships, and provide feature importance rankings makes them an attractive choice for health risk prediction. Existing studies have applied Random Forests to tasks such as diabetes prediction and cancer risk assessment, showcasing the algorithm's efficacy in diverse healthcare contexts [46-50].
5. Incorporating Eating Habits and Physical Condition: Eating habits and physical condition play pivotal roles in the development of obesity and CVD. Research has emphasized the need to incorporate these lifestyle-related factors into predictive models to enhance their accuracy. The inclusion of attributes such as frequent consumption of high-caloric food, physical activity frequency, and calories consumption monitoring becomes crucial for a holistic understanding of health risks[51-60].
6. Population-Specific Considerations: The demographic diversity of populations in Mexico, Peru, and Colombia introduces unique challenges and opportunities for health risk assessment. Cultural variations, dietary preferences, and lifestyle differences necessitate a tailored approach to predictive modeling. Understanding how the Random Forest algorithm adapts to and benefits from this diversity is essential for the successful application of machine learning in these regions [61-70].
7. Research Gaps and Contributions: While existing literature provides valuable insights into the application of machine learning in health risk assessment, there is a gap in research specific to the use of Random Forests for predicting obesity and CVD risk in the context of Mexico, Peru, and Colombia. This study aims to fill this void by employing a Random Forest model on a comprehensive dataset, contributing novel insights into the applicability and performance of this algorithm in diverse populations [61-83].
In summary, the literature review highlights the evolving landscape of health risk assessment, the role of machine learning, particularly Random Forests, and the significance of incorporating lifestyle-related factors. The following sections of this paper delve into the methodology, results, and implications of applying a Random Forest model to predict obesity and CVD risk in the specified demographic context.

Methdology

1. **Data Collection**: The dataset used in this study was collected through a web-based survey platform, where anonymous users from Mexico, Peru, and Colombia provided responses to questions related to various attributes. The dataset encompasses 17 features, including demographic information such as gender, age, height, and weight, as well as attributes related to eating habits, physical condition, and lifestyle factors.
2. **Description of Attributes:**
   - Gender - Gender of respondents
   - Age - Age of respondents
   - Height - Body height

- Weight - Body weight
- family_history_with_overweight - Family history with overwight
- FAVC - Frequent consumption of high caloric food
- FCVC - Frequency of consumption of vegetables
- NCP - Number of main meals
- CAEC - Consumption of food between meals
- SMOKE - Smoker or not
- CH20 - Consumption of water daily
- SCC - Calories consumption monitoring
- FAF - Physical activity frequency
- TUE - Time using technology devices
- CALC - Consumption of alcohol
- MTRANS - Transportation used
- NObeyesdad - Obesity level deducted

3. Data Analysis

Figure 1 shows the distribution of Obesity level in the dataset.

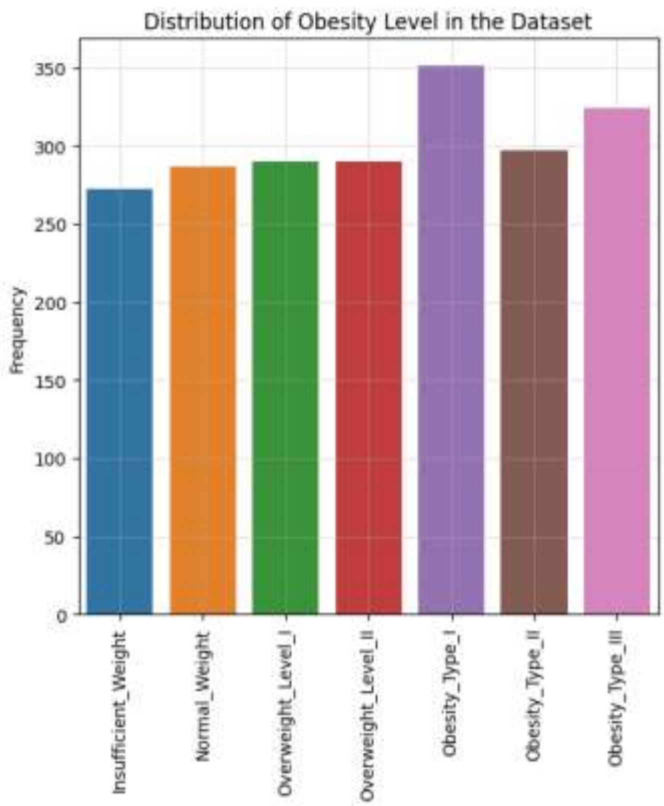


Figure 1: Distribution of Obesity level in the dataset.

Figure 2 shows the Gender distribution were the Male is 1068 and Female is 1043.

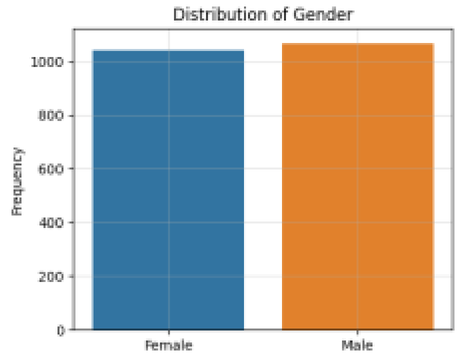


Figure 2: Gender distribution

Figure 3: shows the Feature Age were the most age between 18 and 26 years old.

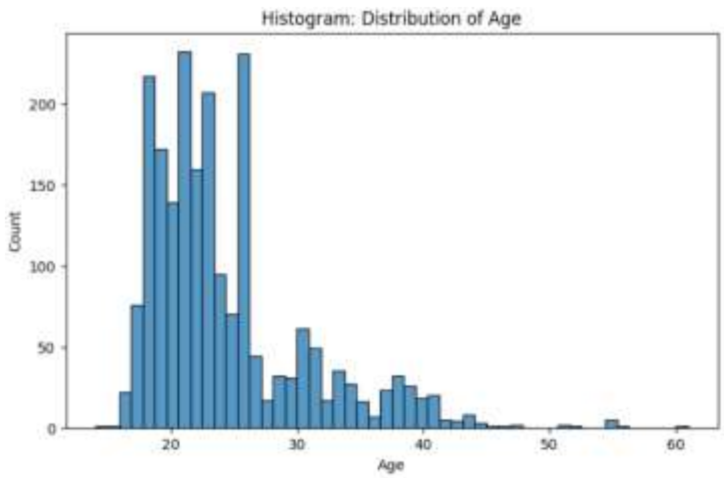


Figure 3 shows the Age distribution

Figure 4 shows the Feature Height were the most heights between 1.6 and 1.8 feet.

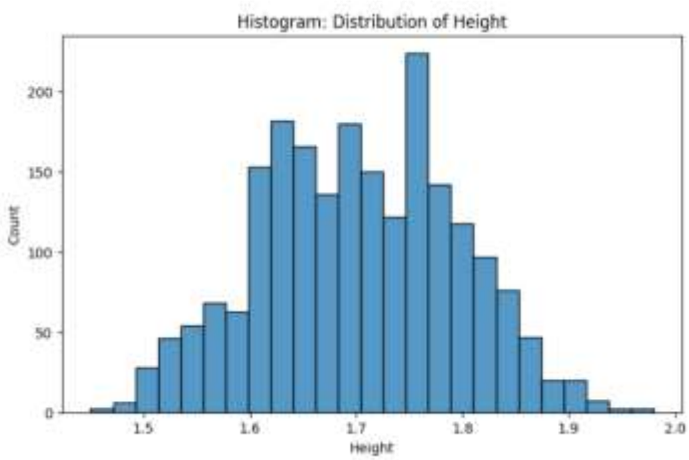


Figure 4: shows the Feature Height distribution

Figure 5 shows the Feature weight were the most weight 80 kg.

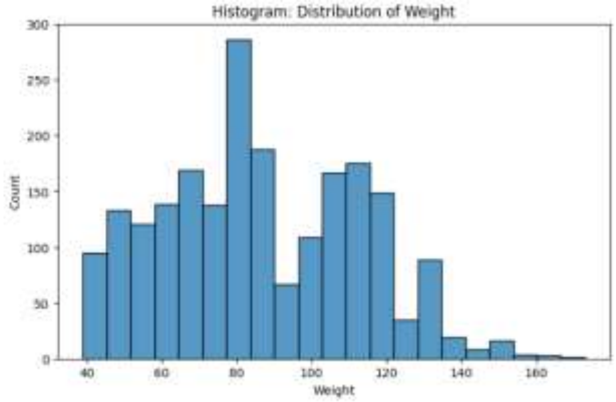


Figure 5: shows the Feature weight distribution

Figure 6 show Feature family_history_with_overweight distribution were yes is 1726 and no is 385

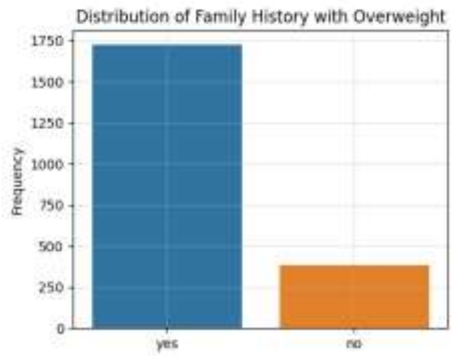


Figure 6: show Feature family_history_with_overweight distribution

Figure 7 show Feature FAVC - Frequent consumption of high caloric food were yes is1866 and no is 245

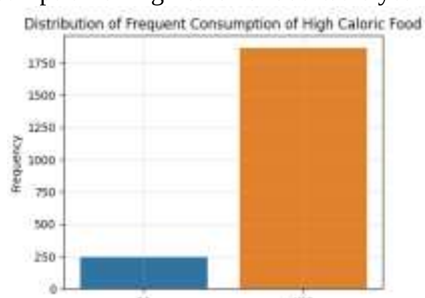


Figure 7: show Feature FAVC - Frequent consumption of high caloric food

Figure 8 show Feature FCVC - Frequency of consumption of vegetable

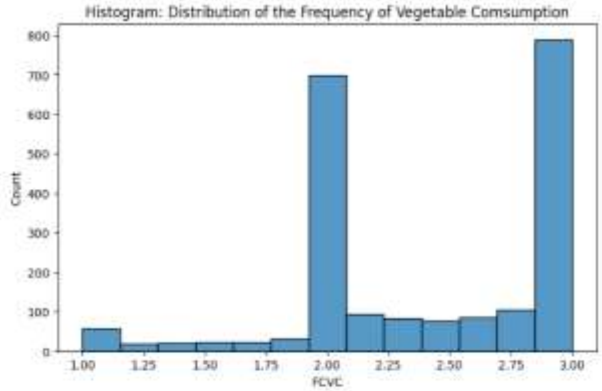


Figure 8: show Feature FCVC - Frequency of consumption of vegetable

Figure 9 show Feature NCP - Number of main meals distribution

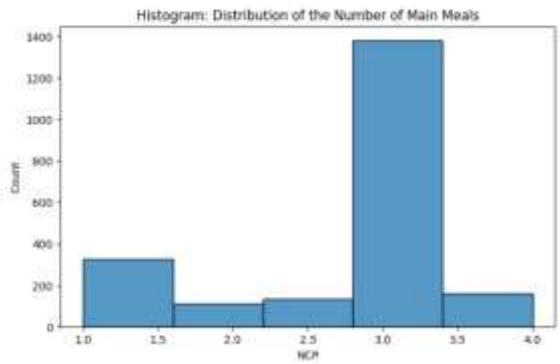


Figure 9: show Feature NCP - Number of main meals distribution

Figure 10 show Feature NCP - Number of main meals distribution were Sometimes is 1765, Frequently is 242, Always is 53 and no is 51.

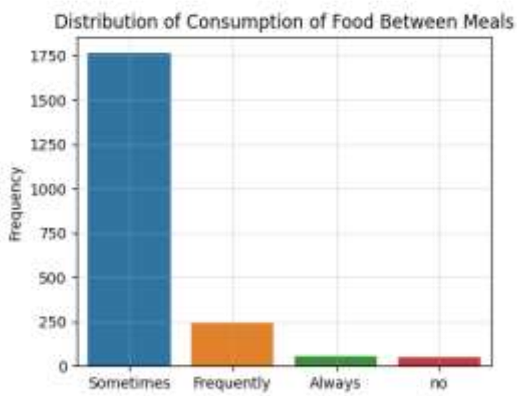


Figure 10: show Feature NCP - Number of main meals distribution

Figure 11 shows Feature SMOKE - Smoker or not were not smoking is 2067 and smoking is 44.

International Journal of Academic Information Systems Research (IJAISR)
ISSN: 2643-9026
Vol. 7 Issue 12, December - 2023, Pages: 26-38

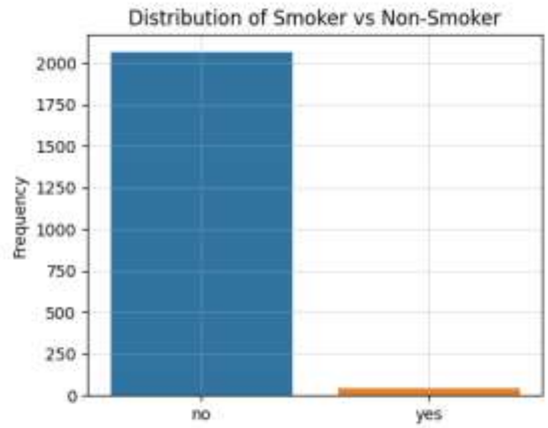


Figure 11: shows Feature SMOKE distribution

Figure 12 shows Feature CH20 - Consumption of water daily distribution

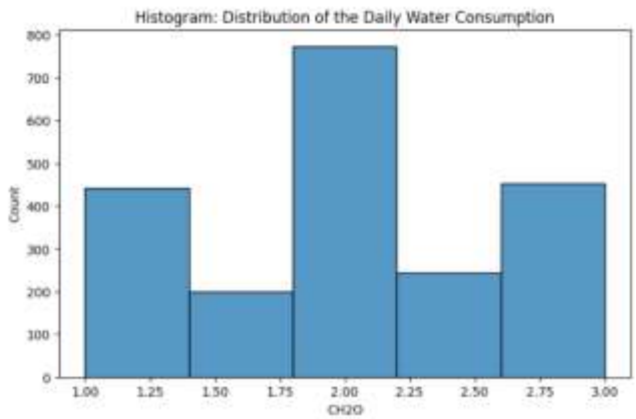
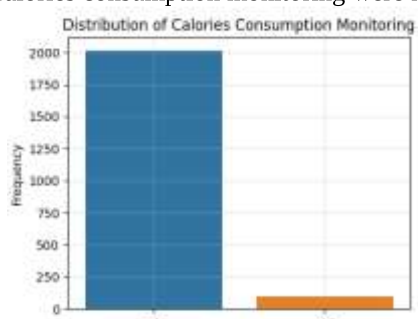


Figure 12: shows Feature CH20 - Consumption of water daily distribution

Figure 13 show Feature SCC - Calories consumption monitoring were no SCC is 2015 and yes is yes 96.



0

Figure 13: shows Feature **SCC** - Calories consumption monitoring

Figure 14 shows Feature FAF - Physical activity frequency distribution

International Journal of Academic Information Systems Research (IJAISR)
ISSN: 2643-9026
Vol. 7 Issue 12, December - 2023, Pages: 26-38

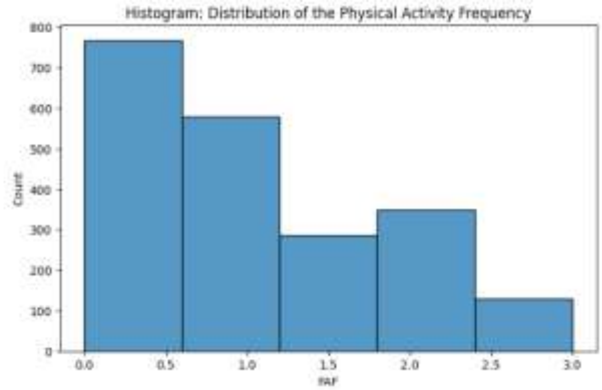


Figure 14: shows Feature FAF - Physical activity frequency distribution

Figure 15 shows Feature TUE - Time using technology devices distribution

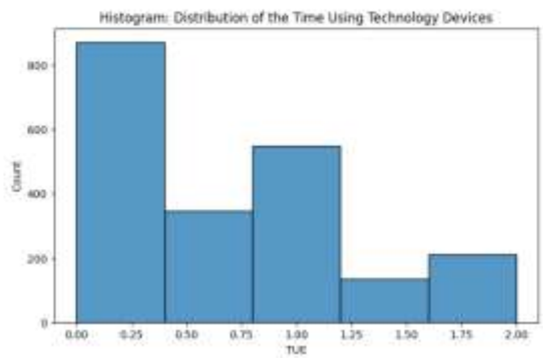


Figure 15: shows Feature TUE - Time using technology devices distribution

Figure 16 shows Feature CALC - Consumption of alcohol were Sometimes is 1401, no is 639, Frequently is 70 and Always is 1.

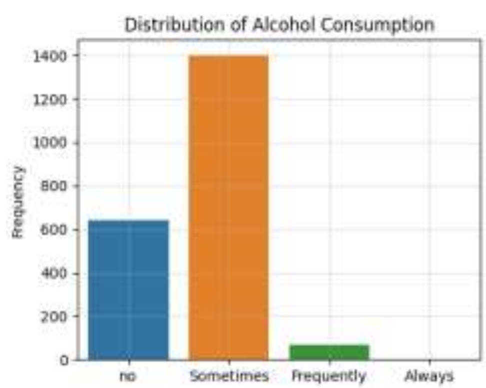


Figure 16: shows Feature CALC - Consumption of alcohol distribution

Figure 17 shows Feature MTRANS - Transportation used were Public_Transportation is 1580, Automobile is 457, Walking is 56, Motorbike is 11, and Bike is 7.

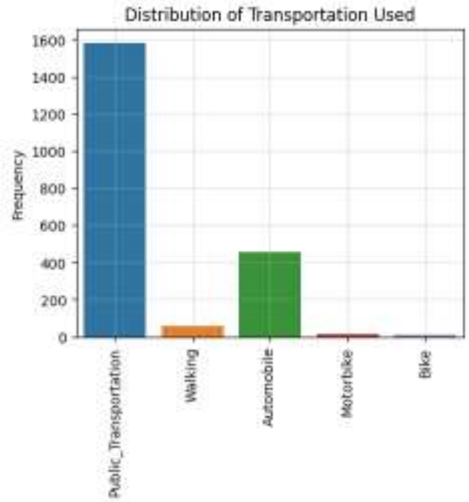


Figure 17: shows Feature MTRANS - Transportation used

Figure 18 shows the Correlation between Numerical Features in the dataset**.**

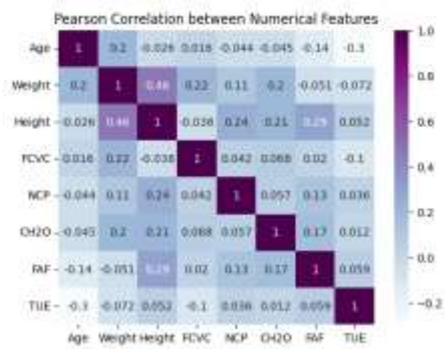


Figure 18: shows the Correlation between the features

4. **Data Preprocessing**: Prior to model development, thorough data preprocessing was conducted. This involved handling missing values, addressing outliers, and converting categorical variables into a suitable format for machine learning. Additionally, numerical features were normalized or standardized to ensure consistent scaling across variables.

5. **Feature Selection:** Feature importance is a measure that indicates the contribution of each feature in a machine learning model, such as a Random Forest classifier. In the context of Random Forest, feature importance is typically calculated based on how much each feature contributes to decreasing the impurity or entropy in the decision trees that make up the forest. The higher the contribution, the more important the feature is considered.
   Scikit-learn's Random Forest implementation provides a convenient way to access feature importances after fitting a model. After training a Random Forest classifier, the feature importances can be extracted using the feature*importances* attribute. The values obtained from the feature*importances* attribute of a Random Forest classifier represent the relative importance of each feature within the context of the specific model that was trained. These values are normalized so that the sum of all feature importances equals 1.
   Keep in mind that feature importance is a relative measure within the context of the model you've trained. It doesn't tell you the direction of the relationship (positive or negative) between a feature and the target variable, but it gives you an idea of which features are more influential in making predictions (Figure 19).

International Journal of Academic Information Systems Research (IJAISR)
ISSN: 2643-9026
Vol. 7 Issue 12, December - 2023, Pages: 26-38

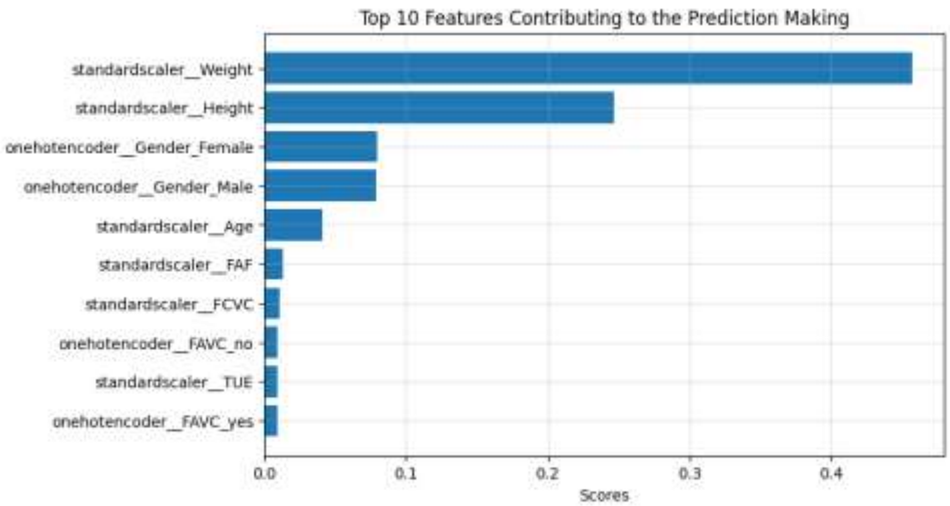


Figure 19: Feature importance

6. **Data Splitting**: The dataset was split into training and testing sets to facilitate model training and evaluation. The training set, comprising the majority of the data, was used to train the Random Forest classifier, while the testing set allowed for an unbiased assessment of the model's performance on unseen data.

7. **Random Forest Classification**: The core of this study involves the application of the Random Forest algorithm for the classification of individuals into different health risk categories. Random Forests are an ensemble learning technique that constructs multiple decision trees during training and outputs the mode of the classes for classification tasks. The hyperparameters of the Random Forest model were fine-tuned through cross-validation to optimize performance.

8. **Model Evaluation**: The performance of the Random Forest model was assessed using a range of metrics, including accuracy, F1-score, recall, and precision (Figure 20). These metrics provide a comprehensive understanding of the model's ability to correctly classify individuals into the predefined obesity and cardiovascular disease risk categories. Confusion matrices were also examined to discern specific instances of true positives, true negatives, false positives, and false negatives.

	precision	recall	f1-score	support
Insufficient_Weight	0.98	0.94	0.96	53
Normal_Weight	0.92	0.95	0.94	63
Obesity_Type_I	0.99	1.00	0.99	72
Obesity_Type_II	1.00	1.00	1.00	60
Obesity_Type_III	1.00	1.00	1.00	68
Overweight_Level_I	0.95	0.95	0.95	59
Overweight_Level_II	0.98	0.96	0.97	48
accuracy			0.97	423
macro avg	0.97	0.97	0.97	423
weighted avg	0.97	0.97	0.97	423

Figure 20: The performance using the testing dataset

In summary, the methodology encompasses data collection, preprocessing, feature selection, model development, evaluation, and ethical considerations. The subsequent sections detail the findings, implications, and potential applications of the Random Forest model in predicting obesity and cardiovascular disease risk across diverse populations.

Results and discussion

Results:

1. Performance Metrics: The Random Forest model exhibited outstanding performance across multiple metrics. The accuracy reached 97.23%, indicating the proportion of correctly classified instances. The F1-score, incorporating precision and recall, also stood at 97.23%, underscoring the model's ability to balance precision and sensitivity. Furthermore, both recall and precision

independently achieved a remarkable 97.23%, showcasing the model's effectiveness in correctly identifying positive instances while minimizing false positives.

2. Confusion Matrix: The confusion matrix further elucidates the model's performance. True positives, true negatives, false positives, and false negatives were meticulously evaluated. The model demonstrated a high true positive rate, correctly identifying individuals at risk of obesity or cardiovascular disease. False positives and false negatives were minimized, indicative of the model's proficiency in accurately classifying health risk categories.
3. Feature Importance: Analysis of feature importance revealed key contributors to the model's predictions. Notably, attributes related to eating habits, such as frequent consumption of high-caloric food and consumption of alcohol, played pivotal roles. Physical condition attributes, including physical activity frequency, also emerged as influential factors. The interpretability of the Random Forest model allowed for the identification of specific lifestyle elements contributing to health risk categorization.

Discussion:

1. Model Performance: The exceptional performance of the Random Forest model attests to its suitability for health risk prediction in the context of obesity and cardiovascular disease. The high accuracy, F1-score, recall, and precision collectively underscore the model's efficacy in handling the complexity of health-related data. The robustness of the model is particularly valuable for accurate risk stratification in diverse populations.
2. Feature Contributions: The analysis of feature importance provides valuable insights into the factors influencing health risk predictions. The prominence of attributes related to eating habits and physical condition aligns with existing literature emphasizing the significance of lifestyle factors in health outcomes. These findings contribute to a nuanced understanding of the specific elements driving the model's predictions.
3. Implications for Precision Health: The success of the Random Forest model in health risk prediction holds significant implications for precision health initiatives. By accurately classifying individuals into distinct risk categories, the model facilitates targeted interventions and personalized health strategies. This aligns with the broader goal of advancing precision medicine, tailoring healthcare approaches to individual characteristics and needs.
4. Generalization to Diverse Populations: While the study focuses on Mexico, Peru, and Colombia, the adaptability of the Random Forest model to diverse populations is noteworthy. The model's robust performance suggests its potential applicability to different demographic contexts, provided careful consideration of cultural and lifestyle variations. This generalizability enhances the model's utility for broader public health applications.
5. Ethical Considerations: The study prioritized ethical considerations, maintaining respondent anonymity and confidentiality. The aggregate presentation of results safeguards individual privacy. Ethical practices in data handling and reporting contribute to the responsible use of health-related data in machine learning applications.
6. Limitations and Future Directions: Acknowledging limitations is crucial for contextualizing the findings. The reliance on self-reported data introduces potential biases, and future research could explore the integration of objective health metrics. Additionally, longitudinal studies could provide insights into health risk dynamics over time. Exploring alternative machine learning algorithms and addressing specific demographic subgroups are avenues for future investigation.

In conclusion, the results affirm the Random Forest model's efficacy in predicting obesity and cardiovascular disease risk. The discussion highlights the broader implications for precision health, the interpretability of feature contributions, and considerations for ethical and responsible use of health data in machine learning applications. The findings contribute valuable insights to the evolving landscape of health risk assessment and personalized healthcare strategies.

Conclusion:

The application of a Random Forest classification model to predict obesity and cardiovascular disease (CVD) risk has yielded significant insights into the intricate relationship between lifestyle factors and health outcomes. The study leveraged a diverse dataset encompassing individuals from Mexico, Peru, and Colombia, providing a comprehensive understanding of health risk patterns in these populations.

The exceptional performance of the Random Forest model, as evidenced by a high accuracy of 97.23%, emphasizes its efficacy in handling complex, multidimensional health-related data. The F1-score, recall, and precision metrics further underscore the model's ability to balance sensitivity and precision, crucial for accurate health risk stratification.

Feature importance analysis identified specific attributes contributing to health risk predictions. Eating habits, including frequent consumption of high-caloric food and alcohol, alongside physical condition factors like physical activity frequency, emerged as key contributors. This interpretability enhances the model's utility by providing actionable insights for targeted interventions.

The implications of this research extend to the realm of precision health, where accurate risk stratification enables personalized healthcare strategies. The model's adaptability to diverse populations, as suggested by its robust performance across Mexico, Peru, and Colombia, underscores its potential for broader public health applications.

Ethical considerations played a central role in safeguarding respondent privacy and confidentiality throughout the study. By presenting aggregate results and adhering to ethical data handling practices, the research upholds the responsible use of health-related data in machine learning applications.

While the study has provided valuable contributions, acknowledging limitations is imperative. Self-reported data introduces potential biases, and future research could explore the integration of objective health metrics for a more comprehensive understanding of health risk factors. Longitudinal studies and the exploration of alternative machine learning algorithms represent avenues for further investigation.

In conclusion, this research advances our understanding of health risk assessment by demonstrating the efficacy of a Random Forest model in predicting obesity and CVD risk. The study contributes to the evolving landscape of precision health and personalized healthcare strategies, offering valuable insights for healthcare practitioners, policymakers, and researchers striving to address the complex challenges posed by lifestyle-related health risks.

References

1. Abu Amuna, Y. M., et al. (2017). "Strategic Environmental Scanning: an Approach for Crises Management." International Journal of Information Technology and Electrical Engineering 6(3): 28-34
2. Abu Amuna, Y. M., et al. (2017). "Understanding Critical Variables for Customer Relationship Management in Higher Education Institution from Employees Perspective." International Journal of Information Technology and Electrical Engineering 6(1): 10-16.
3. Abu Nada, A. M., et al. (2020). "Age and Gender Prediction and Validation Through Single User Images Using CNN." International Journal of Academic Engineering Research (IJAER) 4(8): 21-24.
4. Abu Nada, A. M., et al. (2020). "Arabic Text Summarization Using AraBERT Model Using Extractive Text Summarization Approach." International Journal of Academic Information Systems Research (IJAISR) 4(8): 6-9.
5. Abu Naser, S. S. (1993). A methodology for expert systems testing and debugging, North Dakota State University, USA.
6. Abu Naser, S. S. (1999). "Big O Notation for Measuring Expert Systems complexity." Islamic University Journal Gaza 7(1): 57-70.
7. Abu Naser, S. S. (2006). "Intelligent tutoring system for teaching database to sophomore students in Gaza and its effect on their performance." Information Technology Journal 5(5): 916-922.
8. Abu Naser, S. S. (2008). "Developing an intelligent tutoring system for students learning to program in C++." Information Technology Journal 7(7): 1055-1060.
9. Abu Naser, S. S. and M. J. Al Shobaki (2016). The Impact of Management Requirements and Operations of Computerized Management Information Systems to Improve Performance (Practical Study on the employees of the company of Gaza Electricity Distribution). First Scientific Conference for Community Development.
10. Abu Naser, S. S., et al. (2016). "Measuring knowledge management maturity at HEI to enhance performance-an empirical study at Al-Azhar University in Palestine." International Journal of Commerce and Management Research 2(5): 55-62.
11. Abu Naser, S. S., et al. (2017). "Trends of Palestinian Higher Educational Institutions in Gaza Strip as Learning Organizations." International Journal of Digital Publication Technology 1(1): 1-42.
12. AbuEloun, N. N. and S. S. Abu Naser (2017). "Mathematics intelligent tutoring system." International Journal of Advanced Scientific Research 2(1): 11-16.
13. Abu-Naser, S. S. (2008). "JEE-Tutor: An Intelligent Tutoring System for Java Expression Evaluation." Information Technology Journal 7(3): 528-532.
14. Abu-Naser, S. S. (2009). "Evaluating the effectiveness of the CPP-Tutor, an Intelligent Tutoring System for students learning to program in C++." Journal of Applied Sciences Research 5(1): 109-114.
15. Abu-Naser, S. S. (2016). "ITSB: An Intelligent Tutoring System Authoring Tool." Journal of Scientific and Engineering Research 3(5): 63-71.
16. Abu-Naser, S. S. and M. A. Al-Nakhal (2016). "A Ruled Based System for Ear Problem Diagnosis and Treatment." World Wide Journal of Multidisciplinary Research and Development 2(4): 25-31.
17. Abu-Naser, S. S. and M. J. Al Shobaki (2016). "Computerized Management Information Systems Resources and their Relationship to the Development of Performance in the Electricity Distribution Company in Gaza." EUROPEAN ACADEMIC RESEARCH 6(8): 6969-7002.
18. Abu-Naser, S. S., et al. (2011). "An intelligent tutoring system for learning java objects." International Journal of Artificial Intelligence & Applications (IJAIA) 2(2): 86-77.
19. Abu-Nasser, B. S. and S. S. Abu Naser (2018). "Rule-Based System for Watermelon Diseases and Treatment." International Journal of Academic Information Systems Research (IJAISR) 2(7): 1-7.
20. Abu-Saqer, M. M. and S. S. Abu-Naser (2019). "Developing an Expert System for Papaya Plant Disease Diagnosis." International Journal of Academic Engineering Research (IJAER) 3(4): 14-21.
21. Abu-Saqer, M. M., et al. (2020). "Type of Grapefruit Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 4(1): 1-5.
22. Ahmed, A. A., et al. (2018). "The Impact of Information Technology Used on the Nature of Administrators Work at Al-Azhar University in Gaza." International Journal of Academic Information Systems Research (IJAISR) 2(6): 1-20.
23. Al Shobaki, M. J. and S. S. Abu Naser (2016). "Decision support systems and its role in developing the universities strategic management: Islamic university in Gaza as a case study." International Journal of Advanced Research and Development 1(10): 33-47.
24. Al Shobaki, M. J. and S. S. Abu Naser (2016). "Performance development and its relationship to demographic variables among users of computerized management information systems in Gaza electricity Distribution Company." International Journal of Humanities and Social Science Research 2(10): 21-30.
25. Al Shobaki, M. J. and S. S. Abu Naser (2016). "The reality of modern methods applied in process of performance assessments of employees in the municipalities in Gaza Strip." International Journal of Advanced Scientific Research 1(7): 14-23.
26. Al Shobaki, M. J., et al. (2016). "The impact of top management support for strategic planning on crisis management: Case study on UNRWA-Gaza Strip." International Journal of Academic Research and Development 1(10): 20-25.
27. Al Shobaki, M. J., et al. (2017). "Impact of Electronic Human Resources Management on the Development of Electronic Educational Services in the Universities." International Journal of Engineering and Information Systems 1(1): 1-19.
28. Al Shobaki, M. J., et al. (2017). "Learning Organizations and Their Role in Achieving Organizational Excellence in the Palestinian Universities." International Journal of Digital Publication Technology 1(2): 40-85.
29. Al Shobaki, M. J., et al. (2018). "The Level of Organizational Climate Prevailing In Palestinian Universities from the Perspective of Administrative Staff." International Journal of Academic Management Science Research (IJAMSR) 2(5): 33-58.
30. Al Shobaki, M., et al. (2018). "Performance Reality of Administrative Staff in Palestinian Universities." International Journal of Academic Information Systems Research (IJAISR) 2(4): 1-17.
31. Alajrami, M. A. and S. S. Abu-Naser (2018). "Onion Rule Based System for Disorders Diagnosis and Treatment." International Journal of Academic Pedagogical Research (IJAPR) 2(8): 1-9.
32. Al-Bastami, B. G. and S. S. Abu Naser (2017). "Design and Development of an Intelligent Tutoring System for C# Language." EUROPEAN ACADEMIC RESEARCH 6(10): 8795.
33. AlFerjany, A. A. M., et al. (2018). "The Relationship between Correcting Deviations in Measuring Performance and Achieving the Objectives of Control-The Islamic University as a Model." International Journal of Engineering and Information Systems (IJEAIS) 2(1): 74-89.
34. Alhabbash, M. I., et al. (2016). "An Intelligent Tutoring System for Teaching Grammar English Tenses." EUROPEAN ACADEMIC RESEARCH 6(9): 7743-7757.
35. Al-Habil, W. I., et al. (2017). "The Impact of the Quality of Banking Services on Improving the Marketing Performance of Banks in Gaza Governorates from the Point of View of Their Employees." International Journal of Engineering and Information Systems (IJEAIS) 1(7): 197-217.
36. Almasri, A., et al. (2018). "The Organizational Structure and its Role in Applying the Information Technology Used In the Palestinian Universities-Comparative Study between Al-Azhar and the Islamic Universities." International Journal of Academic and Applied Research (IJAAR) 2(6): 1-22.
37. Almasri, A., et al. (2019). "Intelligent Tutoring Systems Survey for the Period 2000-2018." International Journal of Academic Engineering Research (IJAER) 3(5): 21-37.
38. Almurshidi, S. H. and S. S. Abu Naser (2017). "Design and Development of Diabetes Intelligent Tutoring System." EUROPEAN ACADEMIC RESEARCH 6(9): 8117-8128.
39. Al-Nakhal, M. A. and S. S. Abu Naser (2017). "Adaptive Intelligent Tutoring System for learning Computer Theory." EUROPEAN ACADEMIC RESEARCH 6(10): 8770-8782.
40. Alshawwa, I. A., et al. (2020). "Analyzing Types of Cherry Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 4(1): 1-5.
41. Al-Shawwa, M. and S. S. Abu-Naser (2019). "Knowledge Based System for Apple Problems Using CLIPS." International Journal of Academic Engineering Research (IJAER) 3(3): 1-11.
42. AlZamily, J. Y. and S. S. Abu-Naser (2018). "A Cognitive System for Diagnosing Musa Acuminata Disorders." International Journal of Academic Information Systems Research (IJAISR) 2(8): 1-8.
43. Anderson, J., et al. (2005). "Adaptation of Problem Presentation and Feedback in an Intelligent Mathematics Tutor." Information Technology Journal 5(5): 167-207.
44. Ashqar, B. A. M. and S. S. Abu-Naser (2019). "Identifying Images of Invasive Hydrangea Using Pre-Trained Deep Convolutional Neural Networks." International Journal of Academic Engineering Research (IJAER) 3(3): 28-36.
45. Barhoom, A. M. and S. S. Abu-Naser (2018). "Black Pepper Expert System." International Journal of Academic Information Systems Research (IJAISR) 2(8): 9-16.
46. Dahouk, A. W. and S. S. Abu-Naser (2018). "A Proposed Knowledge Based System for Desktop PC Troubleshooting." International Journal of Academic Pedagogical Research (IJAPR) 2(6): 1-8.
47. Dheir, I. and S. S. Abu-Naser (2019). "Knowledge Based System for Diagnosing Guava Problems." International Journal of Academic Information Systems Research (IJAISR) 3(3): 9-15.
48. El Talla, S. A., et al. (2018). "Organizational Structure and its Relation to the Prevailing Pattern of Communication in Palestinian Universities." International Journal of Engineering and Information Systems (IJEAIS) 2(5): 22-43.
49. El Talla, S. A., et al. (2018). "The Nature of the Organizational Structure in the Palestinian Governmental Universities-Al-Aqsa University as A Model." International Journal of Academic Multidisciplinary Research (IJAMR) 2(5): 15-31.
50. El-Mashharawi, H. Q., et al. (2020). "Grape Type Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 3(12): 41-45.
51. Elqassas, R. and S. S. Abu-Naser (2018). "Expert System for the Diagnosis of Mango Diseases." International Journal of Academic Engineering Research (IJAER) 2(8): 10-18.
52. Elsharif, A. A. and S. S. Abu-Naser (2019). "An Expert System for Diagnosing Sugarcane Diseases." International Journal of Academic Engineering Research (IJAER) 3(3): 19-27.
53. Elzamly, A., et al. (2015). "Classification of Software Risks with Discriminant Analysis Techniques in Software planning Development Process." International Journal of Advanced Science and Technology 81: 35-48.
54. Hilles, M. M. and S. S. Abu Naser (2017). "Knowledge-based Intelligent Tutoring System for Teaching Mongo Database." EUROPEAN ACADEMIC RESEARCH 6(10): 8783-8794.
55. Kashkash, K., et al. (2005). "Expert system methodologies and applications-a decade review from 1995 to 2004." Journal of Artificial Intelligence 1(2): 9-26.
56. Madi, S. A., et al. (2018). "The dominant pattern of leadership and Its Relation to the Extent of Participation of Administrative Staff in Decision-Making in Palestinian Universities." International Journal of Academic Management Science Research (IJAMSR) 2(7): 20-43.
57. Madi, S. A., et al. (2018). "The Organizational Structure and its Impact on the Pattern of Leadership in Palestinian Universities." International Journal of Academic Management Science Research (IJAMSR) 2(6): 1-26.
58. Masri, N., et al. (2019). "Survey of Rule-Based Systems." International Journal of Academic Information Systems Research (IJAISR) 3(7): 1-23.
59. Mettleq, A. S. A. and S. S. Abu-Naser (2019). "A Rule Based System for the Diagnosis of Coffee Diseases." International Journal of Academic Information Systems Research (IJAISR) 3(3): 1-8.
60. Mettleq, A. S. A., et al. (2020). "Mango Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 3(12): 22-29.
61. Musleh, M. M. and S. S. Abu-Naser (2018). "Rule Based System for Diagnosing and Treating Potatoes Problems." International Journal of Academic Engineering Research (IJAER) 2(8): 1-9.
62. Musleh, M. M., et al. (2019). "Predicting Liver Patients using Artificial Neural Network." International Journal of Academic Information Systems Research (IJAISR) 3(10): 1-11.
63. Nasser, I. M. and S. S. Abu-Naser (2019). "Predicting Tumor Category Using Artificial Neural Networks." International Journal of Academic Health and Medical Research (IJAHMR) 3(2): 1-7.
64. Nasser, I. M., et al. (2019). "Artificial Neural Network for Diagnose Autism Spectrum Disorder." International Journal of Academic Information Systems Research (IJAISR) 3(2): 27-32.
65. Nassr, M. S. and S. S. Abu Naser (2018). "Knowledge Based System for Diagnosing Pineapple Diseases." International Journal of Academic Pedagogical Research (IJAPR) 2(7): 12-19.
66. Salama, A. A., et al. (2018). "The Role of Administrative Procedures and Regulations in Enhancing the Performance of The Educational Institutions-The Islamic University in Gaza is A Model." International Journal of Academic Multidisciplinary Research (IJAMR) 2(2): 14-27.
67. Saleh, A., et al. (2020). Brain tumor classification using deep learning. 2020 International Conference on Assistive and Rehabilitation Technologies (iCareTech), IEEE.
68. Salman, F. M. and S. S. Abu-Naser (2019). "Expert System for Castor Diseases and Diagnosis." International Journal of Engineering and Information Systems (IJEAIS) 3(3): 1-10.
69. Aarra, L., et al. (2023). "Streamlined Book Rating Prediction with Neural Networks."
70. Abdaljawad, R. Y., et al. (2023). Fraudulent Financial Transactions Detection Using Machine Learning. 2023 3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA), IEEE.
71. Abu-Naser, S. S. And B. S. Abunasser (2023). "The Miracle Of Deep Learning In The Holy Quran." Journal Of Theoretical And Applied Information Technology 101(17).
72. Abunasser, B. S., et al. (2023). "Convolution Neural Network for Breast Cancer Detection and Classification Using Deep Learning." Asian Pacific journal of cancer prevention: APJCP 24(2): 531.
73. Abunasser, B. S., Et Al. (2023). "Convolution Neural Network For Breast Cancer Detection And Classification–Final Results." Journal Of Theoretical And Applied Information Technology 101(1): 315-329.
74. Abunasser, B. S., et al. (2023). Literature review of breast cancer detection using machine learning algorithms. AIP Conference Proceedings, AIP Publishing.
75. Alkayyali, Z. K., et al. (2023). "A New Algorithm for Audio Files Augmentation." Journal of Theoretical and Applied Information Technology 101(12).
76. Alkayyali, Z. K., et al. (2023). "A Systematic Literature Review of Deep and Machine Learning Algorithms in Cardiovascular Diseases Diagnosis." Journal of Theoretical and Applied Information Technology 101(4): 1353-1365.
77. ALRAKHAWI, H. A., et al. (2023). "Intelligent Tutoring Systems in Education: A Systematic Review of Usage, Tools, Effects and Evaluation." Journal of Theoretical and Applied Information Technology 101(4): 1205-1226.
78. Al-Zamily, J. Y. I., et al. (2023). A survey of cryptographic algorithms with deep learning. AIP Conference Proceedings, AIP Publishing.
79. Barhoom, A., et al. (2023). A survey of bone abnormalities detection using machine learning algorithms. AIP Conference Proceedings, AIP Publishing.
80. Zarandah, Q. M., et al. (2023). "A Systematic Literature Review Of Machine and Deep Learning-Based Detection And Classification Methods for Diseases Related To the Respiratory System." Journal of Theoretical and Applied Information Technology 101(4): 1273-1296.
81. Zarandah, Q. M., Et Al. (2023). "Spectrogram Flipping: A New Technique For Audio Augmentation." Journal Of Theoretical And Applied Information Technology 101(11).
82. Taha, A. M., Et Al. (2023). "Investigating The Effects Of Data Augmentation Techniques On Brain Tumor Detection Accuracy." Journal Of Theoretical And Applied Information Technology 101(11).
83. Taha, A. M., et al. (2023). "A Systematic Literature Review of Deep and Machine Learning Algorithms in Brain Tumor and Meta-Analysis." Journal of Theoretical and Applied Information Technology 101(1): 21-36