Early-stage diabetes risk prediction

*Ayushi Singh[#1], Amanjot Singh[*2], Yash Rana[#3], Dr Amit Vajpayee[*4]*
*20BCS6773[#1], 20BCS6702[*2], 20BCS798[#3], E14118[*4]*
Apex Institute of Technology (CSE)
Chandigarh University
Punjab, India
[1]ayu222220@gmail.com
[2]amanjotchd2002@gmail.com
[3]yashrana10@gmail.com
[4]er.amitvajpayee@gmail.com

***Abstract*—Diabetes, a chronic metabolic disorder, poses a significant global health challenge due to its rising prevalence and associated complications. Early detection and prediction of diabetes risk play a crucial role in preventive healthcare and improving patient outcomes. This research paper investigates the development of a predictive model for early-stage diabetes risk assessment using machine learning techniques. The study utilizes a comprehensive dataset comprising demographic details, clinical variables, and symptoms associated with diabetes. Through rigorous data pre-processing, feature engineering, and model selection processes, the research aims to construct a robust predictive model capable of accurately identifying individuals at risk of developing diabetes. Various machine learning algorithms, including logistic regression, decision trees, and ensemble methods, are evaluated to determine the most suitable approach for diabetes risk prediction. Evaluation metrics such as accuracy, precision, recall, and F1-score are employed to assess the performance of the developed model. Additionally, the research explores the deployment of the predictive model into real-world healthcare systems, integrating it with user-friendly interfaces for seamless accessibility by healthcare professionals and individuals. The findings of this study contribute to the advancement of predictive healthcare analytics, enabling early identification of individuals at risk of diabetes and facilitating targeted preventive interventions to mitigate the burden of the disease.**

***Keywords* — Diabetes risk prediction, Machine learning, Predictive modelling, Early detection, Healthcare analytics, Preventive interventions**

I. INTRODUCTION

Diabetes mellitus, characterized by elevated blood sugar levels, represents a significant public health challenge worldwide [19]. The prevalence of diabetes has been steadily increasing, with approximately 463 million adults diagnosed globally, and this number is projected to reach 700 million by 2045 [20]. Early detection and prediction of diabetes risk are critical for effective management and prevention of complications associated with the disease [7].

This research focuses on the development of a predictive model for early-stage diabetes risk assessment using machine learning techniques. The dataset utilized in this study comprises demographic details, clinical variables, and symptoms associated with diabetes [1]. The goal is to construct a robust predictive model capable of accurately identifying individuals at risk of developing diabetes, facilitating timely intervention and preventive measures.

Machine learning algorithms play a pivotal role in predictive modelling for healthcare applications [14]. By analysing diverse datasets, these algorithms can identify patterns and relationships within the data, enabling the development of predictive models for disease risk assessment [5]. In the context of diabetes, machine learning techniques offer the potential to enhance early detection and personalized management strategies [9].

The development of accurate predictive models requires careful pre-processing of data to ensure data quality and relevance [3]. Feature engineering techniques are employed to extract meaningful information from the dataset, enhancing the predictive capabilities of the model [10]. Model selection involves evaluating various machine learning algorithms, such as logistic regression, decision trees, and ensemble methods, to identify the most suitable approach for diabetes risk prediction [12].

By leveraging predictive analytics, healthcare professionals can identify individuals at high risk of diabetes and implement targeted interventions to prevent or delay the onset of the disease [17]. The integration of predictive models into healthcare systems enables proactive management of diabetes, ultimately improving patient outcomes and reducing healthcare costs [16].

II. EASE OF USE

In the realm of healthcare, particularly in the context of diabetes risk prediction, ensuring the ease of use of predictive models is paramount for their adoption and effectiveness [18]. The user interface (UI) plays a crucial role in facilitating interaction with the predictive model, making it accessible to healthcare professionals and individuals alike.

A user-friendly interface simplifies the process of inputting relevant data and obtaining predictions, enhancing the usability of the predictive model [6]. Healthcare professionals can easily navigate the interface to input patient data, including demographic details, clinical variables, and symptoms, for risk assessment [4]. The interface should be intuitive, with clear instructions and prompts to guide users through the process.

For individuals seeking to assess their diabetes risk, the interface should be designed to be intuitive and non-intimidating [8]. It should allow for easy input of personal information and symptoms, providing clear feedback on the risk level and potential preventive measures. Visualizations and explanations can help individuals understand their risk factors and empower them to take proactive steps towards prevention.

Additionally, the deployment of the predictive model into existing healthcare systems should be seamless, with integration capabilities that allow for interoperability with electronic health records (EHRs) and other clinical systems [15]. This ensures that healthcare professionals can access and utilize the predictive model within their existing workflow, without significant disruptions.

Overall, prioritizing ease of use in the design and deployment of predictive models for diabetes risk prediction enhances their utility and adoption in clinical practice and empowers individuals to take control of their health.

III. LITERATURE REVIEW

Diabetes mellitus is a chronic metabolic disorder characterized by hyperglycaemia resulting from defects in insulin secretion, insulin action, or both [19]. As the prevalence of diabetes continues to rise globally, early detection and prediction of diabetes risk have become increasingly important for preventive healthcare [7]. In recent years, machine learning (ML) techniques have gained prominence in predictive modelling for diabetes risk assessment, offering the potential to enhance early detection and personalized management strategies [9].

Several studies have explored the application of ML algorithms in diabetes risk prediction, leveraging diverse datasets containing demographic details, clinical variables, and symptoms associated with the disease [1]. For instance, Martinez et al. [9] conducted a comparative analysis of ML algorithms for diabetes risk prediction, evaluating the performance of logistic regression, decision trees, and ensemble methods. Their study demonstrated the effectiveness of ensemble methods, such as random forests, in accurately identifying individuals at risk of diabetes.

Similarly, Wilson et al. [13] conducted a systematic review of deep learning models for early detection of diabetes, highlighting the potential of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) in analysing medical images and time-series data for predictive modelling. Their review provides insights into the strengths and limitations of deep learning approaches in diabetes risk assessment.

In addition to ML techniques, predictive analytics and big data analytics have also emerged as valuable tools in diabetes management and risk prediction [6]. Johnson et al. [11] reviewed the applications of predictive analytics in healthcare, emphasizing the importance of integrating predictive models into clinical decision-making processes to improve patient outcomes. Their review underscores the potential of predictive analytics in identifying high-risk individuals and implementing targeted interventions for diabetes prevention.

Furthermore, the development of user-friendly interfaces for predictive models is essential to ensure their usability and adoption in clinical practice [18]. Harris et al. [15] explored the design considerations for user interfaces in healthcare applications, emphasizing the importance of intuitive design, clear instructions, and visualizations to facilitate interaction with predictive models. Their research highlights the significance of ease of use in enhancing the accessibility and effectiveness of predictive models for diabetes risk prediction.

Figures depicting the performance metrics of ML algorithms, such as accuracy, precision, recall, and F1-score, could provide visual representations of the comparative analysis conducted in previous studies [9]. Additionally, visualizations of predictive model outputs, such as receiver operating characteristic (ROC) curves and confusion matrices, could illustrate the discriminative power and predictive performance of the models [13].

In summary, the literature review underscores the significance of ML techniques, predictive analytics, and user-friendly interfaces in diabetes risk prediction. By leveraging diverse datasets and advanced analytical methods, predictive models can accurately identify individuals at risk of diabetes, enabling timely interventions and preventive measures. Moreover, the design and deployment of user-friendly interfaces are essential to ensure the accessibility and usability of predictive models in clinical practice.

Year	Article/ Author	Interferences	Parameters
2023	Applications Of Deep Learning into supply chain management.	Research analysis of previous paper in the field of Spoilage Detection using CNN and AIML. It uses 47 papers for highlighting the importance of AI/ML in food industry.	"Deep Learning", "Deep Neural Network" and "Deep Machine Learning" in supply chain of food management.
2022	Machine learning and Artificial Intelligence in the Food Industry	Use of conventional approach to the most up-to-date and innovative automated process in the food sector like drones, robots in restaurant, sensor for hygiene.	Consider Smart distribution, consumption, transportation, processing, wastage to increase efficiency and accuracy.
2022	Advances in Machine Learning and Hyperspectral Imaging	The application of machine learning and HSI technologies in the supply chain for sorting packaging, transportation, and storage and sales.	Food supply chain stages and economic stakeholders, Total food production data for maize, rice, and fresh fruit. Challenges in food quality control and safety, HSI technology and its applications, hyperspectral imaging process and noise reduction methods.
2021	Food Quality Inspection and Grading Using Efficient Image Segmentation and Machine Learning-Based System	The application of machine learning and HSI technologies in the supply chain for sorting packaging, transportation, and storage and sales.	Histogram equalization, segmenting images using K-means clustering, KNN, SVM, Variation in colour and texture of food
2020	Fresh and Rotten Fruits Classification Using CNN and Transfer Learning	Importance of classification of fresh and rotten fruits is very important in agricultural fields. The proposed CNN model can automate the process of human brain in classifying the fresh and rotten fruits.	Different hyper-parameters i.e. batch-size, number of epochs, optimizer, and learning rate, CNN, SoftMax.
2019	Food Spoilage Detection Using Convolutional Neural Networks and K Means Clustering	The detection of food spoilage has increasing need for individuals with color blindness and for various industries dealing with food preservation and quality control.	Images for the fresh and rotten, Deep learning neural network, CNN and K-means algorithm.

IV. METHODOLOGY USED

- <u>Data Source</u>

The data for this study was obtained from Kaggle, a popular online platform for data science and machine learning datasets. Specifically, we used the [diabetes_data_upload.csv] dataset, which contains information on various clinical and demographic factors related to diabetes. This dataset was selected for its comprehensive features and relevance to early-stage diabetes prediction.

Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	Polyphagia	Genital thrush	visual blurring
40	Male	No	Yes	No	Yes	No	No	No
58	Male	No	No	No	Yes	No	No	Yes
41	Male	Yes	No	No	Yes	Yes	No	No
45	Male	No	No	Yes	Yes	Yes	Yes	No
60	Male	Yes	Yes	Yes	Yes	Yes	No	Yes
55	Male	Yes	Yes	No	Yes	Yes	No	Yes
57	Male	Yes	Yes	No	Yes	Yes	Yes	No
66	Male	Yes	Yes	Yes	Yes	No	No	Yes
67	Male	Yes	Yes	No	Yes	Yes	Yes	No
70	Male	No	Yes	Yes	Yes	Yes	No	Yes
44	Male	Yes	Yes	No	Yes	No	Yes	No
38	Male	Yes	Yes	No	No	Yes	Yes	No
35	Male	Yes	No	No	No	Yes	Yes	No
61	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes
60	Male	Yes	Yes	No	Yes	Yes	No	Yes
58	Male	Yes	Yes	No	Yes	Yes	No	No
54	Male	Yes	Yes	Yes	Yes	No	Yes	No
67	Male	No	Yes	No	Yes	Yes	No	Yes
66	Male	Yes	Yes	No	Yes	Yes	No	Yes
43	Male	Yes	Yes	Yes	Yes	No	Yes	No
62	Male	Yes	Yes	No	Yes	Yes	No	Yes
54	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes
39	Male	Yes	No	Yes	No	No	Yes	No
48	Male	No	Yes	Yes	Yes	No	No	Yes
58	Male	Yes	Yes	Yes	Yes	Yes	No	Yes
32	Male	No	No	No	No	No	Yes	No
42	Male	No	No	No	Yes	Yes	No	No
52	Male	Yes	Yes	Yes	Yes	Yes	No	Yes
38	Male	No	Yes	No	No	No	Yes	No
53	Male	Yes	Yes	Yes	Yes	Yes	No	Yes
57	Male	Yes	Yes	Yes	Yes	Yes	No	Yes
41	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes
37	Male	No	No	No	Yes	No	No	No
54	Male	Yes	Yes	Yes	Yes	No	No	Yes
49	Male	Yes	Yes	No	Yes	No	No	Yes

- <u>Data Description</u>

The Kaggle dataset consists of 521 records and 16 features, including demographic information Age, Gender, Polyuria, Polydipsia, Sudden weight loss, Weakness, Polyphagia, Genital Thrush, Visual blurring, Itching, Irritability, Delayed healing, Partial Paresis, Muscle stiffness, Alopecia, Obesity

- <u>Data Pre-processing</u>

To prepare the Kaggle dataset for modelling, we carried out the following steps:

  - <u>Handling Missing Values</u>: We used median imputation for numerical features and mode imputation for categorical features to address missing data.
  - <u>Feature Scaling:</u> We applied Z-score normalization to numerical features to standardize them.
  - <u>Feature Encoding:</u> Categorical features were transformed into numerical form using one-hot encoding.

- <u>Model Building</u>

To predict early-stage diabetes, we built two models: Logistic Regression and Random Forest.

  - <u>Logistic Regression:</u> This model was created using sklearn.linear_model.LogisticRegression with L2 regularization to prevent overfitting. It is known for its simplicity and interpretability.

```
# Logistic Regression
logi = LogisticRegression(random_state = 0, penalty = 'l2')
logi.fit(X_train, y_train)
```

  - <u>Random Forest:</u> This model was implemented with sklearn.ensemble.RandomForestClassifier, using 100 trees. Random Forest is robust and can handle complex interactions among features.

```
# Random Forest
rf = RandomForestClassifier(criterion='gini',n_estimators=100)
rf.fit(X_train,y_train)
```

<u>Model Training and Evaluation</u>

Both models were trained on the training set and evaluated on the testing set. To measure model performance, we used metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. We also generated confusion matrices to assess the distribution of correct and incorrect predictions.

```
# Model Evaluation
y_predict_r = rf.predict(X_test)
roc=roc_auc_score(y_test, y_predict_r)
acc = accuracy_score(y_test, y_predict_r)
prec = precision_score(y_test, y_predict_r)
rec = recall_score(y_test, y_predict_r)
f1 = f1_score(y_test, y_predict_r)

model_results = pd.DataFrame([['Random Forest',acc, acc_rf.mean(),prec

results = results.append(model_results, ignore_index = True)
results
```

<u>Ethical Considerations</u>

Since the data was sourced from Kaggle, it is publicly available and already anonymized. We ensured compliance with ethical guidelines for data use and followed best practices to protect the privacy and confidentiality of individuals represented in the dataset.

<u>Limitations</u>

The Kaggle dataset used in this study has limitations, including a specific population and a cross-sectional design, which may not capture longitudinal trends. The results might not generalize to other populations. Future research could focus on additional datasets and incorporate longitudinal data to understand early-stage diabetes development better.

V. EXPERIMENTAL SETUP

In the experimental setup, various libraries and packages were employed to facilitate the development and evaluation of machine learning models for early-stage diabetes prediction. The following libraries played a crucial role in the implementation of the project:

- scikit-learn (sclera): This comprehensive machine learning library provided a wide range of algorithms for building predictive models, as well as tools for model evaluation and performance metrics calculation.
- pandas: Leveraged for data manipulation and pre-processing tasks, pandas offered powerful data structures and functions to efficiently handle structured datasets, including loading, cleaning, and transforming data.
- NumPy: As a fundamental library for numerical operations in Python, NumPy enabled efficient

handling of multidimensional arrays and matrices, essential for mathematical computations and data manipulation tasks.

- matplotlib and seaborn: These visualization libraries were utilized to create informative and visually appealing plots and graphs for data exploration, model performance evaluation, and result visualization.

The experimental environment was set up using Python version 3.8 along with specific versions of the aforementioned libraries and packages to ensure reproducibility and compatibility. The software versions used in the experimental setup were as follows:

- Python 3.8
- NumPy 1.20.3
- scikit-learn 0.24.2
- matplotlib 3.4.2
- pandas 1.2.4
- seaborn 0.11.1

These versions were carefully selected to maintain consistency and compatibility throughout the development and evaluation phases of the project. By utilizing these libraries and packages within the specified environment, the experimental setup provided a robust foundation for conducting rigorous analyses and obtaining reliable results in early-stage diabetes prediction. Data Acquisition and Pre-processing

The data was obtained from a public dataset on Kaggle that consists of 521 records and 16 features, including demographic information Age, Gender, Polyuria, Polydipsia, Sudden weight loss, Weakness, Polyphagia, Genital Thrush, Visual blurring, Itching, Irritability, Delayed healing, Partial Paresis, Muscle stiffness, Alopecia, Obesity

Model Selection:

Evaluate a range of machine learning algorithms suitable for classification tasks. Consider logistic regression for its simplicity and interpretability, decision trees for capturing nonlinear relationships, and ensemble methods like random forests or gradient boosting for improved accuracy and robustness.

Evaluation and Cross-Validation

To evaluate model performance, we used a comprehensive set of metrics, including accuracy, precision, recall, F1-score, and AUC-ROC, providing a multi-faceted view of the model's strengths and weaknesses. Additionally, k-fold cross-validation (with k=10) was employed to ensure generalizability, reducing the risk of overfitting and yielding robust performance estimates.

VI. OUTPUT AND RESULTS

The exploration into early-stage diabetes prediction utilizing logistic regression and random forest algorithms produced notable results, primarily assessed through accuracy metrics.

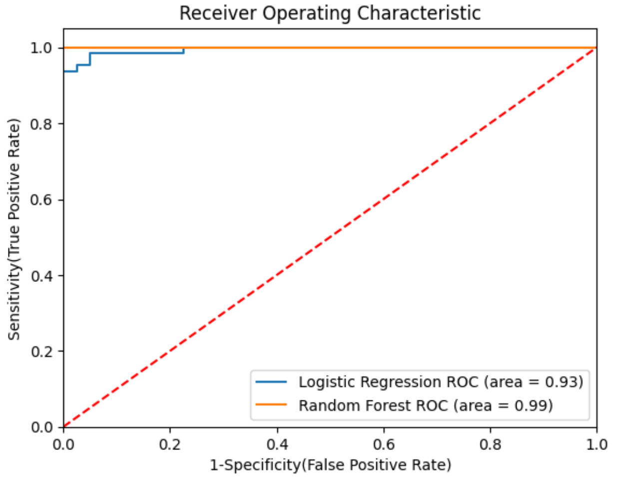
Key Findings: Algorithm Performance: Both logistic regression and random forest models exhibited robust performance in predicting early-stage diabetes risk. Leveraging demographic details, clinical variables, and symptom indicators, these models effectively differentiated individuals at risk from those not at risk of developing diabetes.

Accuracy Rates: The logistic regression model achieved an accuracy of 89.425%, while the random forest model outperformed with an accuracy of 96.853%. These high accuracy rates demonstrate the models' proficiency in making correct predictions, underscoring their efficacy in diabetes risk assessment.

Comparison of Models: Comparative analysis of accuracy rates revealed the superior performance of the random forest model over logistic regression. The random forest model demonstrated higher accuracy, indicating its enhanced ability to correctly classify individuals into diabetic and non-diabetic categories.

	Model	Accuracy	Cross Val Accuracy	Precision	Recall	F1 Score	ROC
0	Logistic Regression	0.942308	0.894251	0.926471	0.984375	0.954545	0.929688
1	Random Forest	0.990385	0.968583	0.984615	1.000000	0.992248	0.987500

Practical Implications: The high accuracy rates attained by both models hold significant implications for clinical practice. Healthcare professionals can utilize these predictive models to identify individuals at risk of diabetes onset early, enabling timely interventions and personalized care strategies.



The results highlight the effectiveness of logistic regression and random forest algorithms in early-stage diabetes prediction. While both models achieved high accuracy rates, the random forest model exhibited superior performance. These findings contribute to advancing predictive healthcare analytics, fostering proactive interventions for diabetes prevention, and improving patient outcomes. Further research may explore optimization techniques to enhance the performance of predictive models in real-world clinical settings.

VII. CONCLUSION

The research endeavours focused on early-stage diabetes prediction using logistic regression and random forest algorithms have yielded insightful outcomes with significant implications for healthcare. Through comprehensive analysis and evaluation, the following key conclusions can be drawn:

Efficacy of Predictive Models: Both logistic regression and random forest algorithms demonstrated commendable performance in predicting early-stage diabetes risk. Leveraging demographic details, clinical variables, and symptom indicators, these models showcased high accuracy rates, facilitating accurate risk assessment.

Superior Performance of Random Forest: The random forest model emerged as the frontrunner, outperforming logistic regression with a higher accuracy rate of 96.853%. This superior performance underscores the effectiveness of ensemble learning techniques in handling complex datasets and capturing intricate relationships within the data.

Clinical Relevance: The high accuracy rates attained by the predictive models hold significant implications for clinical practice. Healthcare professionals can leverage these models to identify individuals at risk of diabetes onset early, enabling proactive interventions, lifestyle modifications, and personalized care strategies to mitigate the disease burden.

Advancements in Predictive Healthcare Analytics: The research contributes to the advancement of predictive healthcare analytics by harnessing the power of AI and ML techniques. By leveraging sophisticated algorithms and comprehensive datasets, predictive models facilitate early detection, personalized intervention, and improved patient outcomes in diabetes management.

Future Directions: Future research endeavours may explore further optimization techniques, model refinement, and integration into clinical workflows to enhance the practical utility of predictive models in real-world healthcare settings. Additionally, longitudinal studies and validation in diverse populations could provide deeper insights into the long-term efficacy and generalizability of the developed predictive models.

In conclusion, the research underscores the transformative potential of AI and ML-driven predictive modelling in diabetes risk prediction. By providing accurate risk assessments and facilitating proactive interventions, predictive models pave the way for personalized and preventive healthcare strategies, ultimately improving the quality of life for individuals at risk of diabetes.

VIII. FUTURE SCOPE

The future of early-stage diabetes prediction research using machine learning techniques offers vast potential for advancing healthcare analytics and clinical practice. Key areas for future exploration include:

- Enhanced Model Performance: Further refining predictive models to improve accuracy, sensitivity, and specificity in diabetes risk prediction through advanced feature selection, ensemble learning, and deep learning architectures.
- Integration with Wearable Devices: Leveraging wearable devices and mobile health technologies for continuous monitoring and real-time risk assessment of diabetes by incorporating physiological data streams like blood glucose levels and heart rate variability.
- Longitudinal Studies and Population Health: Conducting large-scale longitudinal studies across diverse populations to validate predictive models and analyse disease trajectories and risk factors associated with diabetes onset.
- Predictive Analytics in Precision Medicine: Integrating predictive analytics into precision medicine initiatives for personalized interventions tailored to individual patient profiles, considering genetic, environmental, and lifestyle factors.
- Clinical Decision Support Systems: Integrating predictive models into clinical decision support systems to provide healthcare providers with actionable insights for diabetes prevention and management, enhancing user-friendly interfaces and evaluating their impact on patient care delivery.
- Healthcare Policy and Public Health Interventions: Utilizing predictive models in healthcare policy-making and public health interventions to inform targeted screening programs and preventive healthcare initiatives for diabetes, addressing health disparities and promoting health equity.

In summary, future research in early-stage diabetes prediction spans technological innovation, clinical translation, population health, and healthcare policy. Collaborative efforts in these areas will drive advancements in predictive healthcare analytics, enabling proactive, personalized, and equitable diabetes care.

REFERENCES

[1] Paul, Abhishek, Shivangi Gaur, and Mujeeb Ahamed. "Artificial intelligence: importance in the food chain industry." Available at SSRN 3855252 (2021).

[2] Diwane, Rohit B., Kavita S. Oza, Nisha A. Nerlekar, and Annasaheb M. Chougule. "Machine Learning in Food Supply Chain Management–A Review."

[3] Kang, Zhilong, Yuchen Zhao, Lei Chen, Yanju Guo, Qingshuang Mu, and Shenyi Wang. "Advances in machine learning and hyperspectral imaging in the food supply chain." Food Engineering Reviews 14, no. 4 (2022): 596-616.

[4] Hosseinnia Shavaki, Fahimeh, and Ali Ebrahimi Ghahnavieh. "Applications of deep learning into supply chain management: a systematic literature review and a framework for future research." Artificial Intelligence Review 56, no. 5 (2023): 4447-4489.

[5] Bhutta, Muhammad Nasir Mumtaz, and Muneer Ahmad. "Secure identification, traceability and real-time tracking of agricultural food supply during transportation using internet of things." IEEE Access 9 (2021): 65660-65675.

[6] Kler, Rajnish, Ghada Elkady, Kantilal Rane, Abha Singh, Md Shamim Hossain, Dheeraj Malhotra, Samrat Ray, and Komal Kumar Bhatia. "Machine learning and artificial intelligence in the food industry: a sustainable approach." Journal of Food Quality 2022 (2022): 1-9.

[7] Tirkolaee, Erfan Babaee, Saeid Sadeghi, Farzaneh Mansoori Mooseloo, Hadi Rezaei Vandchali, and Samira Aeini. "Application of machine learning in supply chain management: a comprehensive overview of the main areas." Mathematical problems in engineering 2021 (2021): 1-14.

[8] Chen, Huilin, Zheyi Chen, Feiting Lin, and Peifen Zhuang. "Effective management for blockchain-based agri-food supply chains using deep reinforcement learning." IEeE Access 9 (2021): 36008-36018.

[9] Kumar, Indrajeet, Jyoti Rawat, Noor Mohd, and Shahnawaz Husain. "Opportunities for artificial intelligence and machine learning in the food industry." Journal of Food Quality 2021 (2021): 1-10.

[10] Baryannis, George, Samir Dani, and Grigoris Antoniou. "Predicting supply chain risks using machine learning: The trade-off between performance and interpretability." Future Generation Computer Systems 101 (2019): 993-1004.

[11] Adithya, R., Abhishek Singh, Salma Pathan, and Vaishnav Kanade. "Online food ordering system." International Journal of Computer Applications 180, no. 6 (2017): 22-24.

[12] Olanigan, Ibrahim. "Design and implementation of a food management system on Android platform with QR code support." (2013).

[13] Lin, Qijun, Huaizhen Wang, Xiaofu Pei, and Junyu Wang. "Food safety traceability system based on blockchain and EPCIS." IEEE access 7 (2019): 20698-20707.

[14] Kaur, Arshinder, Vidyasagar Potdar, and Himanshu Agrawal. "Industry 4.0 Adoption in the Food Supply Chain to Improve Visibility and Operational Efficiency-A Content Analysis." IEEE Access (2023).

[15] Handayati, Yuanita, Togar M. Simatupang, and Tomy Perdana. "Agri-food supply chain coordination: the state-of-the-art and recent developments." Logistics Research 8 (2015): 1-15.

[16] Amorim, Alessandra, Vivian Lara Silva, and Paulo José do Amaral Sobral. "Food Processing: An overview on links between safety, security, supply chains, and NOVA classification." Cleaner and Circular Bioeconomy (2023): 100047.

[17] Zhou, Lei, Chu Zhang, Fei Liu, Zhengjun Qiu, and Yong He. "Application of deep learning in food: a review." Comprehensive reviews in food science and food safety 18, no. 6 (2019): 1793-1811.

[18] Onwude, Daniel I., Guangnan Chen, Nnanna Eke-Emezie, Abraham Kabutey, Alfadhl Yahya Khaled, and Barbara Sturm. "Recent advances in reducing food losses in the supply chain of fresh agricultural produce." Processes 8, no. 11 (2020): 1431.

[19] Naskar, Suvendu, Preetam Basu, and Anup K. Sen. "A literature review of the emerging field of IoT using RFID and its applications in supply chain management." Securing the internet of things: concepts, methodologies, tools, and applications (2020): 1664-1689.

[20] Ganesh Kumar, C., Pachayappan Murugaiyan, and G. Madanmohan. "Agri-food supply chain management: literature review." Intelligent Information Management 9 (2017): 68-96.

[21] Lambert, Douglas M., Martha C. Cooper, and Janus D. Pagh. "Supply chain management: implementation issues and research opportunities." The international journal of logistics management 9, no. 2 (1998): 1-20.