A Comprehensive case study on Applications of Agentic AI in Smart Technology

Dr. Jyoti Jadhav[1], Dr. Tejashri Phalle[2]
[1]*Assistant Professor, School of Information Technology Indira University*
[2]*Associate Professor, School of Information Technology Indira University*

***Abstract*—Agentic AI is a type of artificial intelligence that can act autonomously to achieve complex goals with very little help from people. Unlike older AI that needs strict rules and constant supervision, Agentic AI can make its own advanced decisions, adapt to new situations, and operate independently in a changing world. This paper fully explores the core ideas and methods behind Agentic AI, looking at its unique ability to be self-sufficient and adaptable. We examine its exciting potential uses in areas like medicine, finance, and software that needs to change on its own, stressing how useful these agentic systems are in real-world settings. Finally, the survey addresses the important ethical challenges this technology presents, such as making sure the AI's goals match ours and that it can handle resource limits. It proposes a framework for safely integrating Agentic AI into society to ensure its transformative power is used responsibly**

***Index Terms*—Agentic AI, system, framework, application**

I. INTRODUCTION

The field of artificial intelligence has transitioned from limited, specialized tools to sophisticated AI agents that operate autonomously, perceiving, reasoning, and acting based on environmental feedback. This evolution, driven by the emergence of Large Language Models (LLMs) which serve as robust foundation models augmented with modules for memory, planning, and tool use has enabled individual agents to handle increasingly complex tasks. However, as these single-agent architectures reach their limits when facing real-world challenges requiring diverse skills, parallel execution, and coordinated efforts, there is a growing interest in multi-agent systems. These systems involve collections of specialized agents that collaborate, effectively tackling problems that are too difficult for any one agent to solve alone. Agentic AI marks a qualitative advancement in artificial intelligence, defined by its capability to set and pursue complex goals in uncontrolled, changing environments by autonomously managing its own resources. Unlike most traditional AI systems, which operate as supervised tools excelling only at well-defined tasks within strict boundaries, Agentic AIs remain goal-directed and operational even when facing significant chaos and the necessity to toggle between multiple objectives. This shift is motivated by the critical need for tools capable of operating with flexibility in complex, real-world conditions—such as disaster relief, healthcare, and cybersecurity—where independent and proper decision-making is vital. Agentic AIs promise to fundamentally transform the role of AI from a passive assistant to a strategic planner, allowing them to work collaboratively with humans to reallocate effort, increase productivity, and handle dangerous tasks, thereby reshaping job structures in core industries.

II. LITERATURE REVIEW

This literature review traces the evolution of AI from rigid, rule-based systems and Generative AI to the autonomous framework of Agentic AI, which is defined by its ability to make independent decisions and pursue specific goals in changing environments. By examining architectures like multi-agent systems and hierarchical reinforcement learning, the review distinguishes Agentic AI from its predecessors while identifying critical research gaps in scalability, real-time learning, and ethics. Ultimately, it emphasizes the necessity of integrating safety governance and bias mitigation to ensure these highly autonomous systems are deployed responsibly. The research examines how

AI is used across various fields like healthcare, robotics, and customer service, noting a clear distinction between simple and advanced systems. While standard AI agents are best for straightforward, tool-based tasks like running chatbots or managing schedules, "Agentic AI" is designed for much more complex and autonomous work. This advanced version can manage its own goals and coordinate with other systems, making it ideal for sophisticated applications like medical decision-making, scientific research, or controlling a group of robots that need to work together without constant human intervention.

The conventional AI applications automate routine tasks. In this paper, the authors define the agentic system as capable of interacting across service processes with a degree of self-direction, integrating data inputs, decision making and generating operational outputs. The paper highlights. This paper also highlights the role of agentic agents in operational efficiency, personalization of guest services and strategic decision support. In hospitality organizations while presenting the opportunities of agentic AI. For the hospitalisation, the lid, and the tourism industry the author systematically addresses the practical and ethical challenges associated with deployment. The paper also shows successful adaptation will depend not only on Technological capability, but also on addressing organizational readiness, skill gaps and customer trust factors. Which are underdeveloped in existing research. To guide future research the paper proposes research agenda which should be used within a socio-technical framework. It can also be used for governance, regulations and human-machine interaction in hospitality settings [1].

In this paper, author examines the architectural evaluation of large language models (LLM) based multi agent systems (MAS) and their practical application across diverse application domains. It focuses on how multiple Alarm Power specialized agent can be organized into collaborative networks. These networks. Are capable of handling complex tasks through explicit orchestration and communication mechanisms. This paper. Focuses on. Case studies spanning. Telecommunications Security. Heritage Assess Management. And utilities customer service automation [2].

This study contributes to understand how generative models influence and disciplinary knowledge production, research, design and pedagogical practices in educational settings. The author highlights. Concerns about the validity, reliability, and epistemic trust in AI assisted educational outputs. Within the educational sciences, it expressed caution noting that unchecked integration of generative tools can cause or increase biases, propagate errors, or blur the distinction between original thought and automated suggestion. The fields like healthcare, hospitality and smart systems are increasingly starting to adopt advanced AI paradigms. Whereas. The education research offers a balance between innovation with scholarly rigor and institutional trust [3].

This paper focuses on the emerging paradigm of AI-powered in silico twins (ISTs) high fidelity computational patient models designed to support precision medicine. These IST's integrate with physiological simulation which are adaptive and data-driven components. It consists of the valuation of digital health technologies. Combining the multimodal data sources like electronic health records, wearable sensors, imaging and omics profiles, therapy outcomes, etc. The paper also discussed the technical taxonomy of ISTR architecture and its operational components. It tells us how combining these elements supports adaptive simulation and decision support for patient care continuum. There are few challenges and limitations also, which include issue of model validation, computational cost, data privacy and integration into clinical workflows [4].

The practical integration of interactive conversational agents within case-based research workflows is shown in this paper. This paper consists of how the generative models can be evaluated as interactive resource assistance which supports this case data. The author. Shows how researchers can use this as interface to reformulate research questions, find out anomalies in qualitative data, to find out insights without excessive manual retrieval. [5].

III. SIGNIFICANCE OF AGENTIC BASED ARTIFICIAL INTELLIGENCE TECHNIQUES

Agentic AI systems represent a significant evolution in artificial intelligence, defined by their ability to exhibit autonomous decision-making, engage in goal-oriented behavior, and undergo continuous learning as they interact with dynamic environments. Crucially, Agentic AI adapts its operation based on real-time data and evolving objectives, distinguishing it from traditional AI that typically relies on human oversight or fixed, pre-programmed instructions. These intelligent agents achieve their efficiency and adaptability by leveraging advanced methodologies, including machine learning, reinforcement learning, and sophisticated multi-agent coordination techniques. The Model Context Protocol (MCP) reference architecture establishes a highly adaptable and modular ecosystem designed to orchestrate diverse intelligent agents across varied operational environments. At its foundation, the architecture offers a flexible deployment model that allows organizations to balance the trade-offs between latency, privacy, and computational power; this is achieved by deploying models locally for sensitive tasks, leveraging the cloud for resource-intensive operations, or utilizing hybrid configurations that shift workloads based on real-time requirements. To extend capabilities beyond the limitations of general-purpose Large Language Models (LLMs), the framework incorporates Specialized Function Agents. These purpose-built units utilize dedicated algorithmic approaches—such as classical planning, mathematical optimization, and computer vision—to provide high-precision outputs in niche domains. These specialized agents are seamlessly integrated into the broader ecosystem via standardized adapters that translate internal logic into MCP-compliant interfaces. To ensure this modern framework remains compatible with existing infrastructure, Legacy System Agents serve as sophisticated wrappers, managing protocol translation and state consistency to bridge the gap between established APIs and the new agentic standard.The architecture also formalizes the role of the human expert through Human-in-the-loop Agents, which provide structured interfaces for oversight, delegation, and feedback-driven learning, ensuring that high-stakes decisions remain under human control. These diverse agent types are unified through four primary integration patterns: Hierarchical integration for structured command-and-control, Peer-to-Peer integration for direct collaboration, Service-oriented integration for modular discovery, and Event-driven integration for reactive, state-based workflows. Together, these elements form a cohesive, scalable framework that maintains architectural interoperability while maximizing the specialized strengths of each individual component.

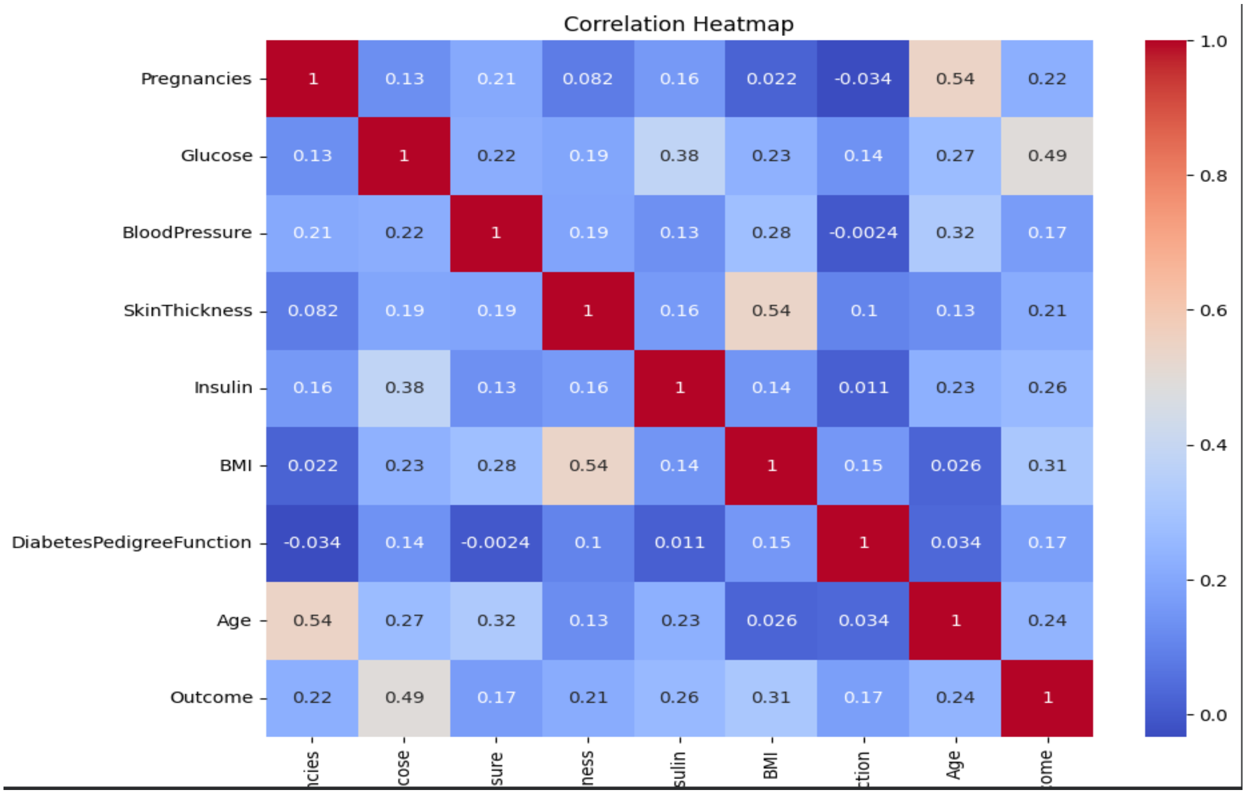
Pattern of Agent	Function	Advantages
Hierarchical	A complex task is decomposed in to simpler e.g., a "Manager" agent delegating to "Coder" and "Tester" agents	Clear accountability, structured reasoning and predictable command flow.
Peer-to-peer	Rapid and direct collaboration between equal agents e.g., two specialized sensors swapping data	Low latency, highly flexible, decentralized with no single point of failure.
Service Oriented	Modular tools and reusable resources. e.g., a translation service used by variety of agents.	High reusability, standard interfaces, clear discovery through registries.
Event-driven	Real-time, reactive systems e.g., a "Security Agent" reacting instantly to a "Log Entry" event	Highly scalable and decoupled

IV. METHODOLOGY

This study proposes a machine learning-based framework for the early prediction of diabetes using the Pima Indians Diabetes Dataset. The methodology consists of data preprocessing, exploratory data analysis, model development, performance evaluation, and hierarchical clustering. Initially, the dataset was imported into the Python environment using the Pandas library, and an initial exploratory analysis was performed to examine the data structure, descriptive statistics, feature distributions, and potential data quality issues. Several clinical attributes, including glucose level, blood pressure, skin thickness, insulin, and body mass index (BMI), contained biologically implausible zero values that were treated as missing observations. These values were replaced with missing values and subsequently imputed using the median of each respective feature to preserve the robustness of the dataset while minimizing the influence of skewed distributions. Furthermore, extreme outliers in the insulin feature were identified using the Interquartile Range (IQR) method and capped at the calculated upper threshold to reduce their impact on model performance.

Following data preprocessing, exploratory data analysis was conducted through count plots, boxplots, and correlation heatmaps to understand feature distributions, identify outliers, evaluate class imbalance, and investigate the relationships between predictor variables and the target variable. The cleaned dataset was then divided into predictor variables and the binary outcome variable, where the outcome represented the presence or absence of diabetes. Subsequently, the data were partitioned into training and testing subsets using a 70:30 train-test split to facilitate unbiased model evaluation.

Four supervised machine learning algorithms, namely Logistic Regression, Gaussian Naïve Bayes, K-Nearest Neighbors (KNN), and Random Forest, were implemented for diabetes prediction. Logistic Regression was selected as a baseline linear classifier for binary classification, while Gaussian Naïve Bayes was employed due to its computational efficiency and probabilistic framework. K-Nearest Neighbors was used to classify patients based on similarity among neighboring instances, whereas the Random Forest algorithm was adopted as an ensemble learning technique capable of handling nonlinear relationships and reducing overfitting through multiple decision trees. Each model was trained using the training dataset and evaluated on the testing dataset.



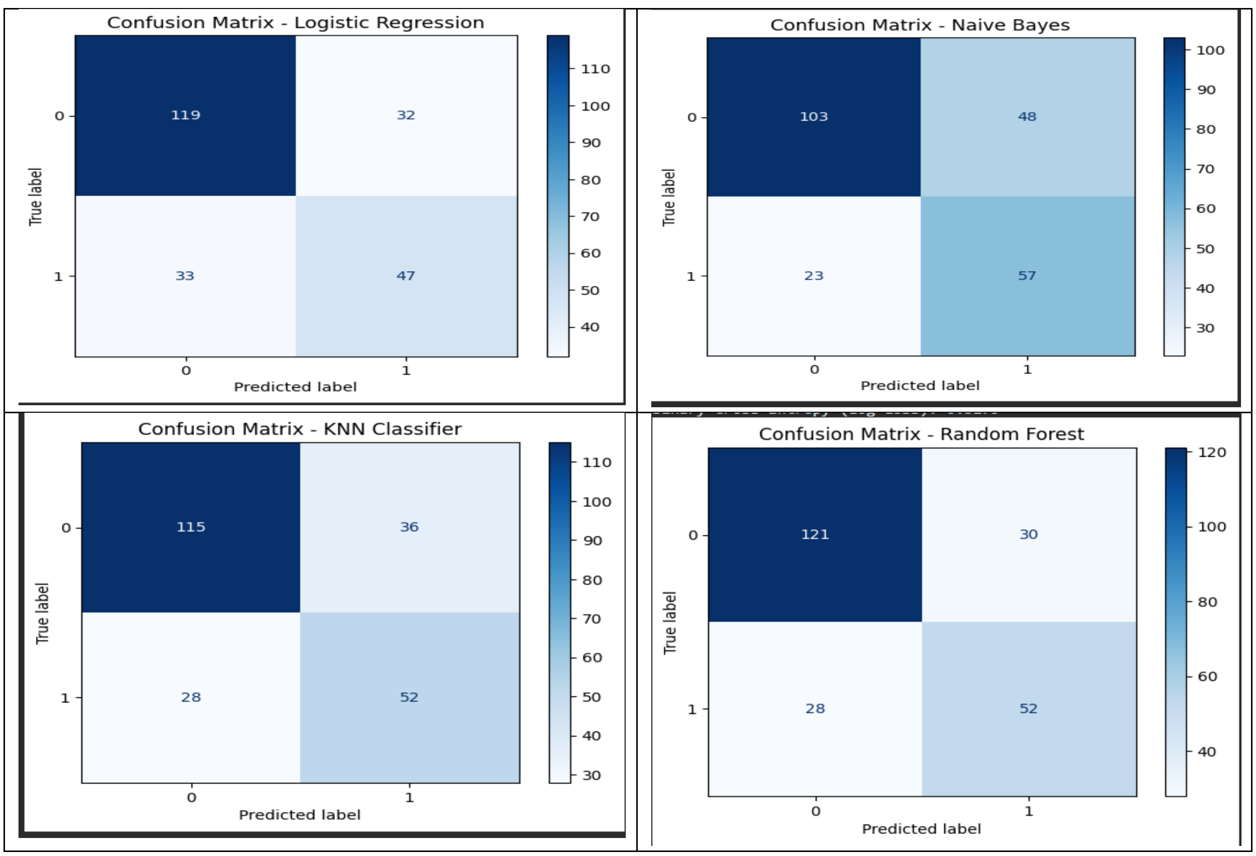
The predictive performance of the developed models was assessed using multiple evaluation metrics, including accuracy, precision, recall, F1-score, and binary cross-entropy (log loss). Confusion matrices were also generated to provide a detailed analysis of true positive, true negative, false positive, and false negative classifications. These evaluation metrics enabled a comprehensive comparison of the classification models to identify the most effective approach for diabetes prediction.

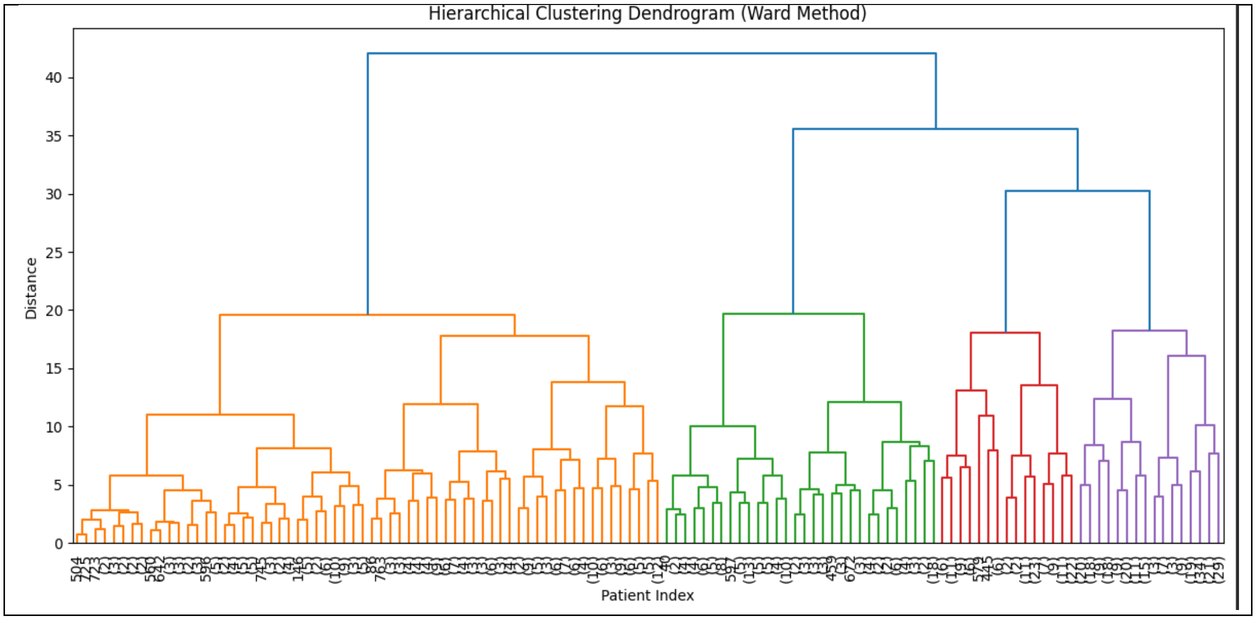
In addition to supervised classification, an unsupervised learning approach was incorporated through hierarchical clustering to explore natural groupings within the patient data. Before clustering, all predictor variables were standardized using the StandardScaler to ensure equal contribution of each feature during distance computation. Ward's linkage criterion was employed to construct a hierarchical clustering dendrogram, which facilitated the identification of patient clusters with similar clinical characteristics. This clustering analysis provides additional insights into potential patient subgroups that may support personalized healthcare interventions and improve clinical decision-making. Finally, the overall performance of all classification models was compared using graphical visualizations and performance metrics to determine the most suitable model for accurate diabetes prediction.

V. RESULT AND DISCUSSION

The performance of the four machine learning models was evaluated using five widely accepted classification metrics: accuracy, precision, recall, F1-score, and log loss. These metrics provide a comprehensive assessment of each model's predictive capability in identifying diabetic and non-diabetic patients. The comparative results indicate that the Random Forest classifier achieved the best overall performance among the evaluated models, while the K-Nearest Neighbors (KNN) classifier demonstrated the weakest predictive capability.

Model	Accuracy	Precision	Recall	F1 Score	Log Loss
Logistic Regression	74.03%	63.51%	58.75%	61.04%	0.5255
Naïve Bayes	74.03%	62.20%	63.75%	62.96%	0.8056
KNN	67.10%	52.27%	57.50%	54.76%	2.0234
Random Forest	74.89%	63.10%	66.25%	64.63%	0.5034

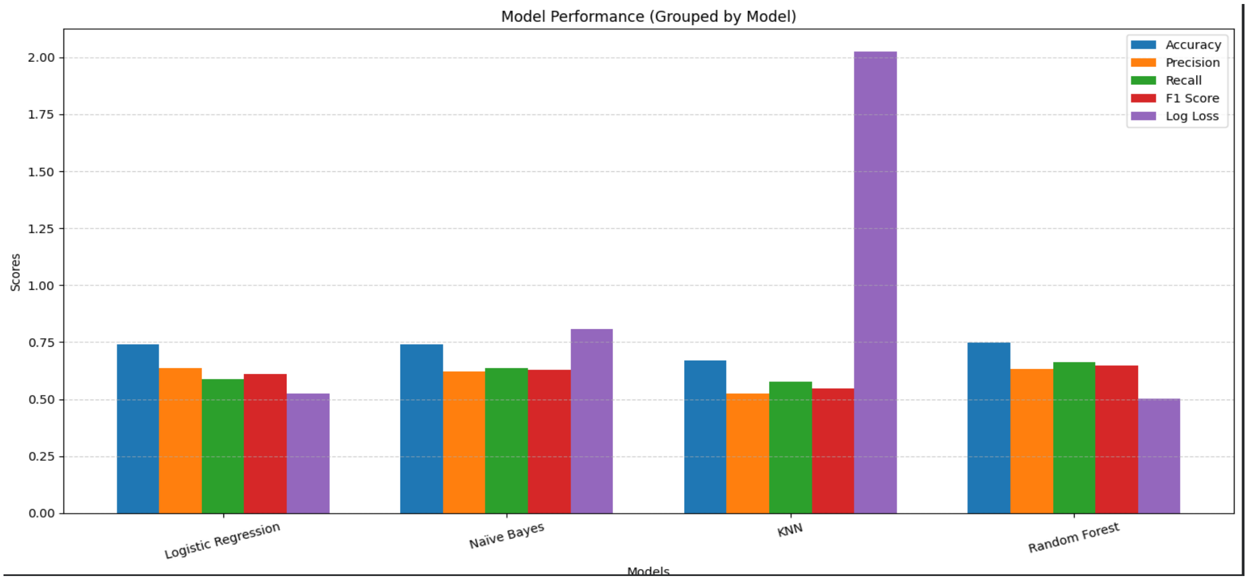




The Logistic Regression model achieved an accuracy of 74.03%, indicating that it correctly classified approximately three-quarters of the test instances. It obtained a precision of 63.51%, demonstrating that nearly two-thirds of the patients predicted as diabetic were correctly identified. The recall value of 58.75% suggests that the model detected slightly more than half of the actual diabetic cases, while the F1-score of 61.04% reflects a moderate balance between precision and recall. Furthermore, the model recorded a log loss of 0.5255, indicating relatively reliable probability estimates with limited prediction uncertainty. These findings suggest that Logistic Regression serves as a strong baseline model for binary diabetes prediction because of its simplicity, interpretability, and stable predictive performance.The Gaussian Naïve Bayes classifier also achieved an overall accuracy of 74.03%, matching the performance of Logistic Regression in terms of correctly classified observations. However, its precision was slightly lower at 62.20%, indicating a marginally higher number of false positive predictions. The recall value of 63.75% was higher than that of Logistic Regression, suggesting that Naïve Bayes was more effective at identifying patients with diabetes. Consequently, the model produced an improved F1-score of 62.96%, reflecting a better balance between precision and recall. Despite these advantages, the log loss increased substantially to **0.8056**, implying that although the model generated acceptable classifications, the confidence associated with its predicted probabilities was comparatively weaker. This behavior is consistent with the strong feature independence assumption inherent in the Naïve Bayes algorithm, which may not fully capture the complex relationships among medical variables.Among all evaluated classifiers, the K-Nearest Neighbors (KNN) algorithm exhibited the lowest predictive performance. The model achieved an accuracy of 67.10%, considerably lower than the other classification techniques. Its precision of 52.27% indicates that only about half of the patients predicted as diabetic were correctly classified. Similarly, the recall value of 57.50% demonstrates limited ability to identify true diabetic patients, resulting in an F1-score of 54.76%, the lowest among all models. Moreover, the KNN classifier recorded the highest log loss of 2.0234, indicating poor probability calibration and greater prediction uncertainty. The comparatively weak performance of KNN may be attributed to the sensitivity of distance-based algorithms to feature scaling, noisy observations, and the relatively small size of the dataset.

The Random Forest classifier outperformed all other models across most evaluation metrics. It achieved the highest classification accuracy of 74.89%, demonstrating the greatest overall predictive capability. The model obtained a precision of 63.10%, indicating a strong ability to correctly identify diabetic patients while minimizing false positive predictions. More importantly, it achieved the highest recall of

66.25%**,** meaning that it successfully detected the largest proportion of actual diabetic cases among all evaluated models. The corresponding F1-score of **64.63%** further confirms the model's superior balance between precision and recall. In addition, Random Forest produced the lowest log loss of 0.5034, indicating highly reliable probability estimates and greater confidence in its predictions. The superior performance of Random Forest can be attributed to its ensemble learning strategy, which combines multiple decision trees to reduce overfitting, improve generalization, and effectively capture complex nonlinear relationships within the medical data.



Overall,
the comparative analysis demonstrates that the Random Forest classifier provides the most effective solution for diabetes prediction in this study. Its superior accuracy, highest recall and F1-score, together with the lowest log loss, indicate that it delivers the best balance between predictive accuracy and probability calibration. Although Logistic Regression and Gaussian Naïve Bayes produced competitive results and may be preferred in situations requiring simpler and more interpretable models, Random Forest offers the greatest potential for supporting clinical decision-making through accurate identification of high-risk diabetic patients. Conversely, the K-Nearest Neighbors algorithm showed comparatively weaker performance and may require additional optimization, such as hyperparameter tuning or advanced feature engineering, before being considered for practical healthcare applications.

VI. CONCLUSION

This study presented a comprehensive machine learning framework for the early prediction of diabetes using the Pima Indians Diabetes Dataset. The proposed methodology incorporated systematic data preprocessing, including the treatment of missing values, outlier handling, exploratory data analysis, supervised classification, and hierarchical clustering to improve data quality and extract meaningful insights. Four widely used machine learning algorithms—Logistic Regression, Gaussian Naïve Bayes, K-Nearest Neighbors (KNN), and Random Forest—were implemented and evaluated using multiple performance metrics, including accuracy, precision, recall, F1-score, and log loss. Among the evaluated models, the Random Forest classifier achieved the best overall performance, with an accuracy of 74.89%**,** the highest recall **(**66.25%**)** and F1-score (64.63%**),** and the lowest log loss (0.5034**),** demonstrating its superior capability in accurately identifying diabetic patients while maintaining reliable probability estimates. Logistic Regression and Gaussian Naïve Bayes also produced competitive results, whereas KNN exhibited comparatively lower predictive performance. Furthermore, hierarchical clustering provided additional insights by identifying natural patient subgroups based on clinical characteristics, which may support personalized healthcare interventions. Overall, the findings demonstrate that machine learning techniques, particularly ensemble learning through Random Forest, can serve as effective decision-support tools for the early detection of diabetes. The proposed framework has the potential to assist healthcare professionals in improving diagnostic accuracy and enabling timely clinical interventions. Future research can further enhance the model by incorporating larger and more diverse datasets, applying advanced ensemble and deep learning techniques, performing hyperparameter optimization, and integrating explainable artificial intelligence methods to improve model transparency and clinical applicability.

REFERENCES

[1] Dwivedi, Y. K., Helal, M. Y. I., Elgendy, I. A., Albashrawi, M. A., Hughes, L., Shawosh, M., ... & Jeon, I. (2026). Artificial intelligence agents and agentic systems in hospitality and tourism: challenges, opportunities and research agenda. International Journal of Contemporary Hospitality Management, 38(1), 27-52.

[2] Banerjee, S. (2026, December). Historical Genealogies of Data Colonialism: From Colonial Censuses to Digital AI Systems. In International Conference on Smart Systems and Social Management (ICSSSM-2 2025) (pp. 991-1004). Atlantis Press.

[3] Brezovec, E., Zelić, M., & Zagode, A. M. (2026). Stabilizing truth in educational sciences: a systematic review of generative AI in education. Kybernetes, 55(13), 1-17.

[4] Alharthi, S. (2026). AI-powered in silico twins: redefining precision medicine through simulation, personalization, and predictive healthcare. Saudi Pharmaceutical Journal, 34(1), 1.

[5] Heinbach, C. (2026). Chat with my case study. Informatik Spektrum, 1-13.