JIMS
Une série de conférences scientifiques internationales
http://jims-sciconf.com/

8th Health Marketing International Day, EDHEC Paris - 4th June 2024
Paper presentation under the Conference Workshop 1 : Révolution IA et Pharma | **Presided by:** Pr. A. Hamard
Time: 15h - 16h45 - **Hall No.** Amphi A & B

Personal Digital Twins Machine Learning Models for Personalized Healthcare and Ethical Implications with Generative-AI

Dr. Kiran Gulia[1], Phool Kumar[2], Dr. Hiran Patel[1], Dr. Sandeep Mitten[3]

[1]Faculty of Science and Engineering, University of Wolverhampton - UK
[2]IAE - University School of Management, University of Lille - France
[3]Zydus Pharmaceutical Division of Clinical Research and Development, New Jersey - USA

Research made in the research labs of

LUMEN Lille University Management - ULR 4999 | zydus Pharmaceuticals USA | Université de Lille | University of Wolverhampton Knowledge • Innovation • Enterprise

https://lumen.univ-lille.fr | https://zydususa.com/ | www.univ-lille.fr | https://www.wlv.ac.uk

Research Index

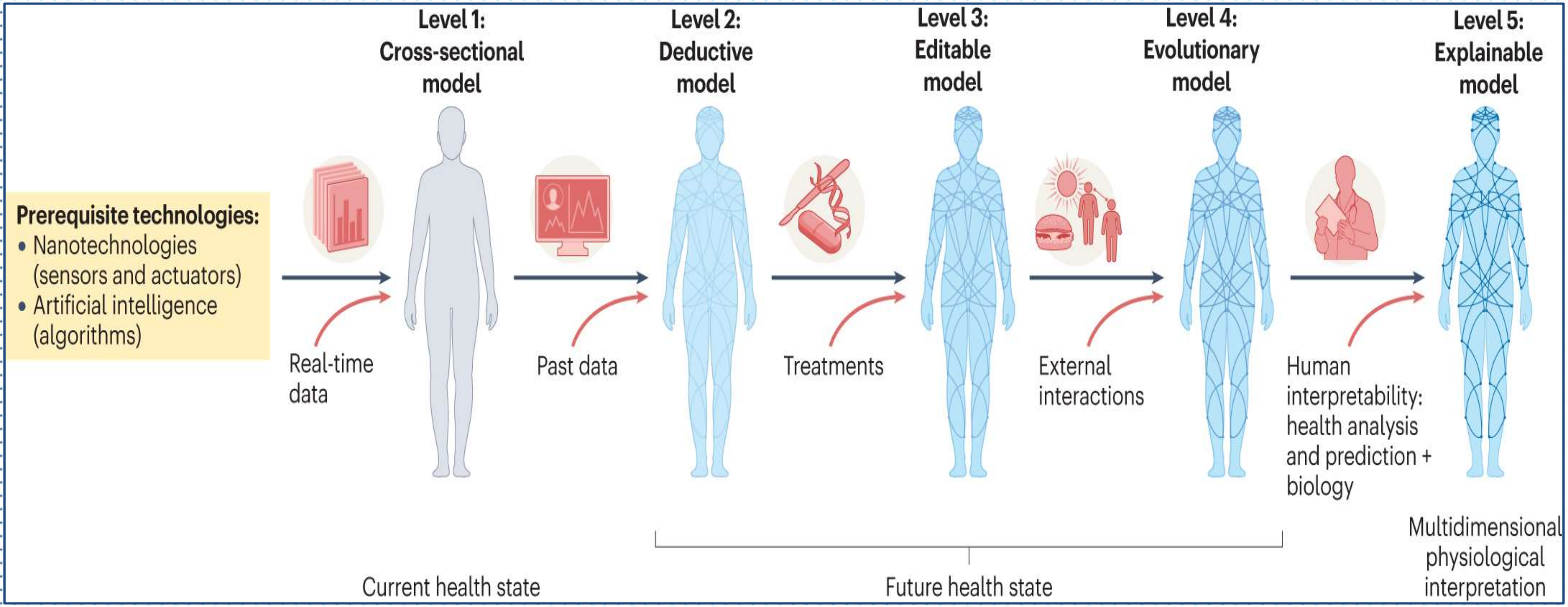
Research Keywords

Personal Digital Twins (PDT), Digital Twin (DT), Machine Learning Models of PDTs, G-AI Base Healthcare, Personalised Medical Interventions, Diabetes Management using PDTs, Algorithmic biasing in PDTs, Ethical implications of PDT with Generative-AI, Personal Digital Twins ethical challenges, Use of Generative-AI in Personal Digital Twins Healthcare, Use of Large Language Models in Personal Digital Twins Healthcare.

Research Objectives

Introduces a Machine Learning (ML) Model:

- **Key benefits:** Early disease prevention, optimised treatment planning, enhanced patient-physician communication, reduced healthcare costs, and improved resource allocation.
- **Ethical implications**



PDTs: Concept, Importance and Ongoing R&Ds

Concept and Importance:

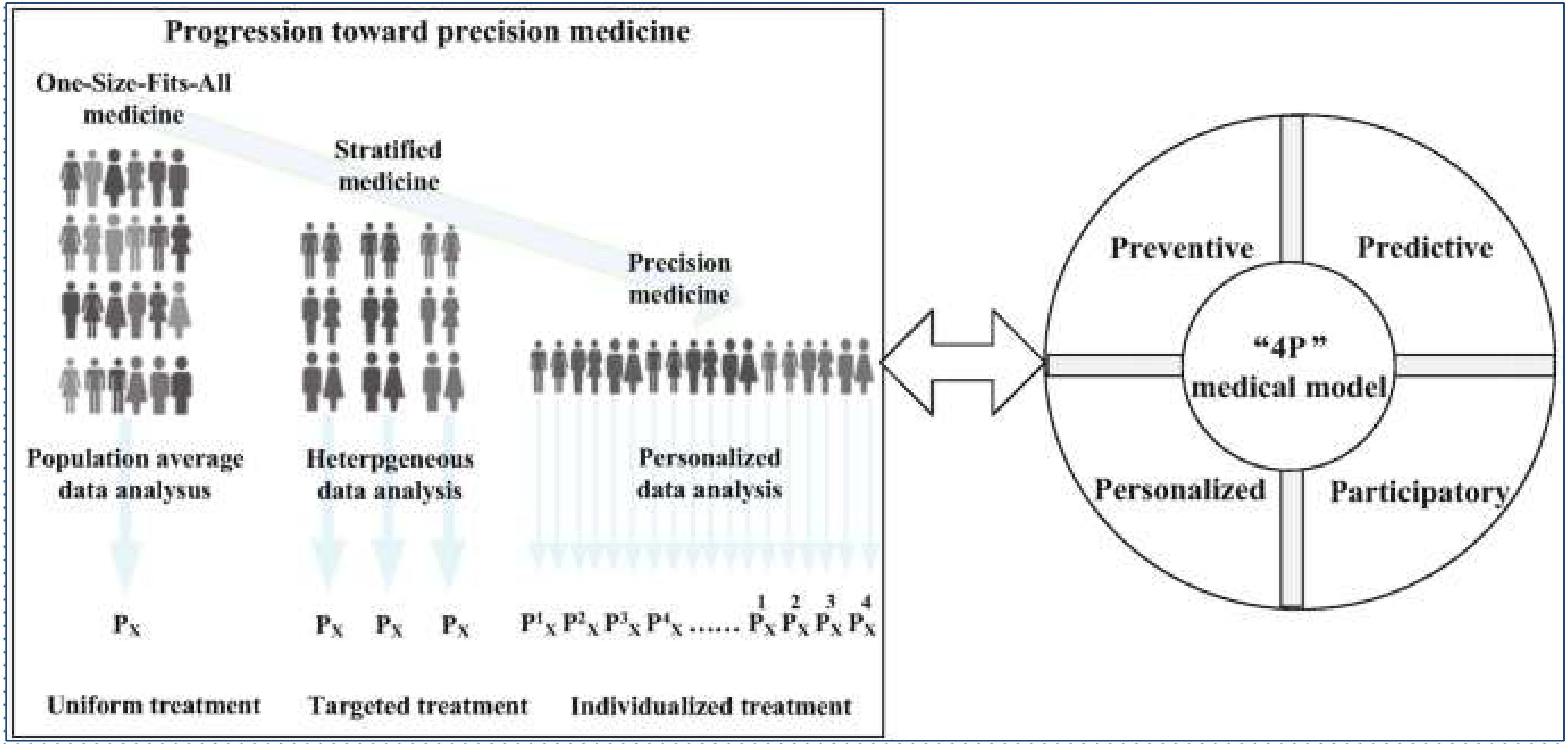
1. Creating Digital Twin Clone of patients to enhance diagnosis and illness treatment procedures, by simulating the human body using that clone for medical treatment as well as the medicine research and developments, *Watts* [2018].
2. PDTs transform healthcare by creating virtual replicas of patients, *Habibzadeh et al.* [2020] & *Internet Things* [2022].
3. DTs derive predictions about diagnosis, prognosis, efficacy, and optimization of therapeutic interventions, *Indiana University* [2021].
4. PDT replicas leverage individual data, including genetics, *Gawel et al.* [2019].
5. PDT produce 3D heart models, *Philips-USA* [2018] & *Copley* [2018].
6. PDT Predictive Algorithms analyse real-time data to detect health risks to prevent diseases before symptoms arise, *Smith-B* & *Miller-EH*.
7. PDTs are equipped with genetic and medical history information helpful in personalizing treatment plans, *Wang-J* & *Popa et al.* [2021].

Ongoing R&Ds:

1. EDITH and CSA are exploring virtual human twin in integrated multidiscipline approach, *EDITH*.
2. MeDigit in Sweden's Linköping University individualized Digital Twins are tested for diagnosis and treatment evaluation in heart disease, *Tilda & Tino*.
3. Philips' Heart Model application offers cardiologists interactive 3D models of patients' hearts for surgical planning, *Philips-USA* [2018], with similar products under development by Siemens Healthineers and GE Healthcare, *Copley* [2018].

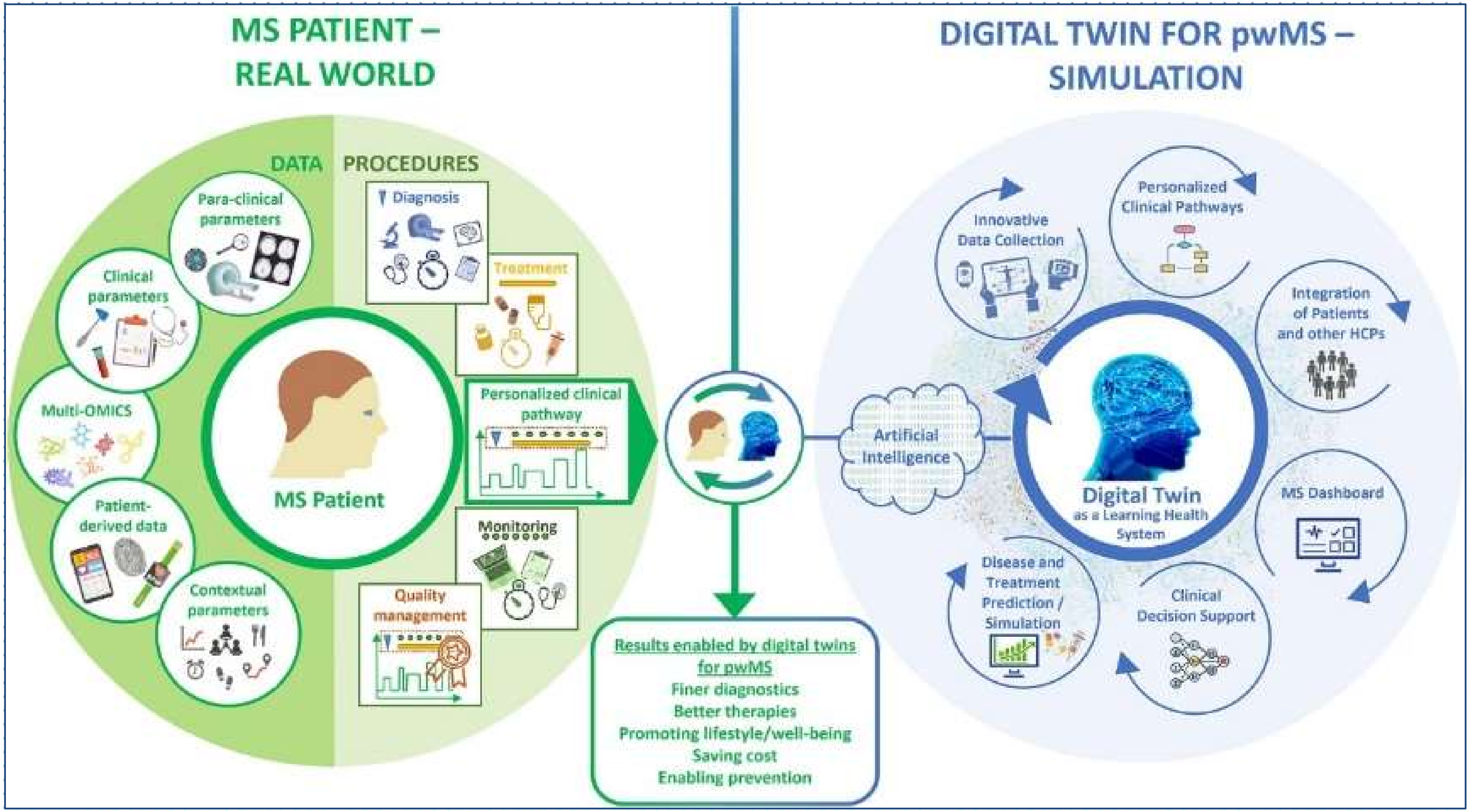
Existing Healthcare Challenges

- **Generic:** Current Healthcare Systems
- **Reactive:** Traditional Healthcare
- AI based Healthcare with **Significant Issues**



Proposed AI Concept

- ML models on **Diabetes management**.
- PDTs provide highly **personalized insights**.
- PDTs **predictive maintenance**



Dataset-1: PIMA Indian Diabetes Dataset

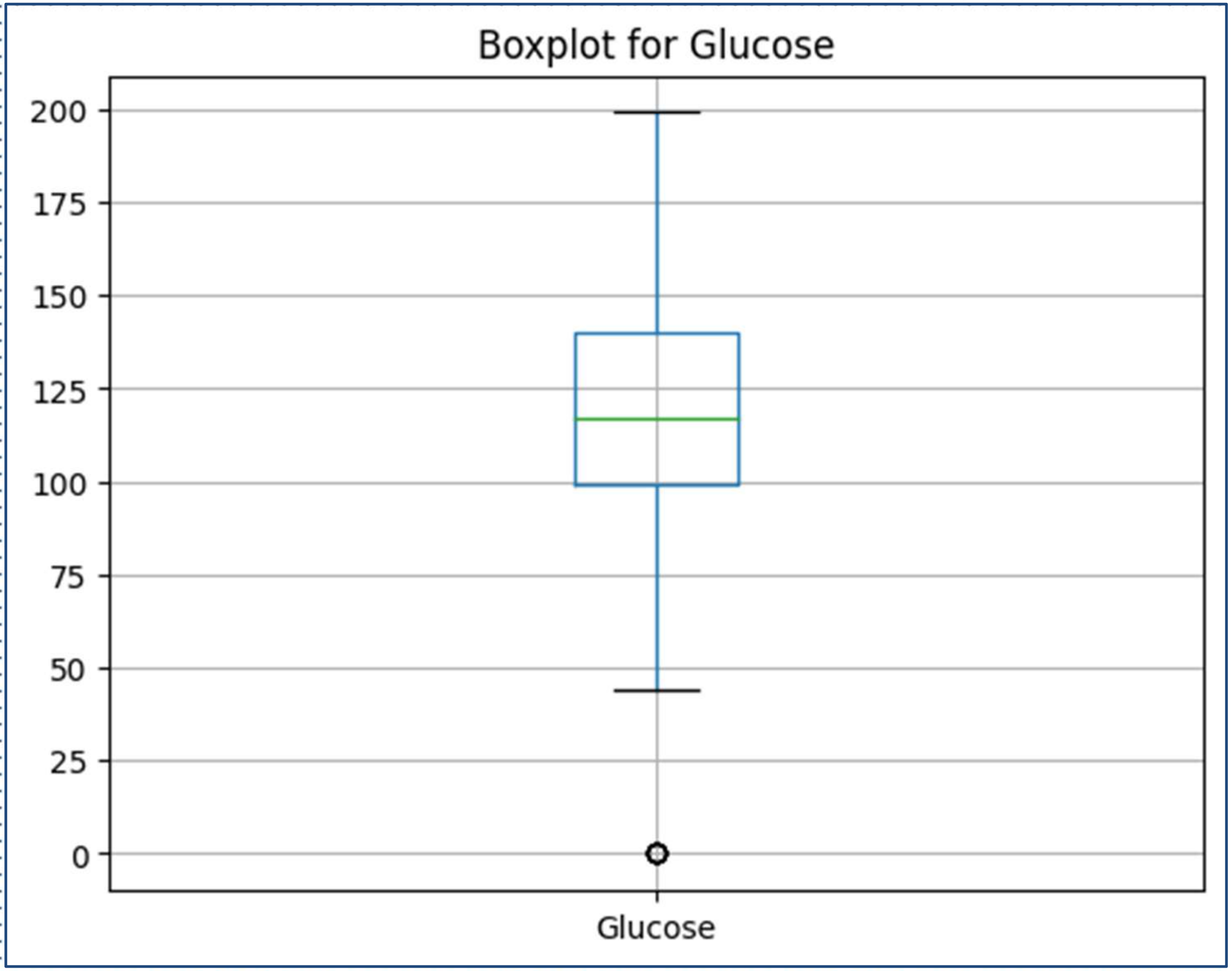
- Collected by the National Institute of Diabetes and Digestive and Kidney Diseases.
- Medical Data Women aged 21 and older.
- Study prevalence of diabetes in this population and developing predictive models.

Key Features Include:

Pregnancies, Glucose, Concentration, Blood Pressure, Skin Thickness, Insulin Level, Body Mass Index (BMI), Diabetes, Age

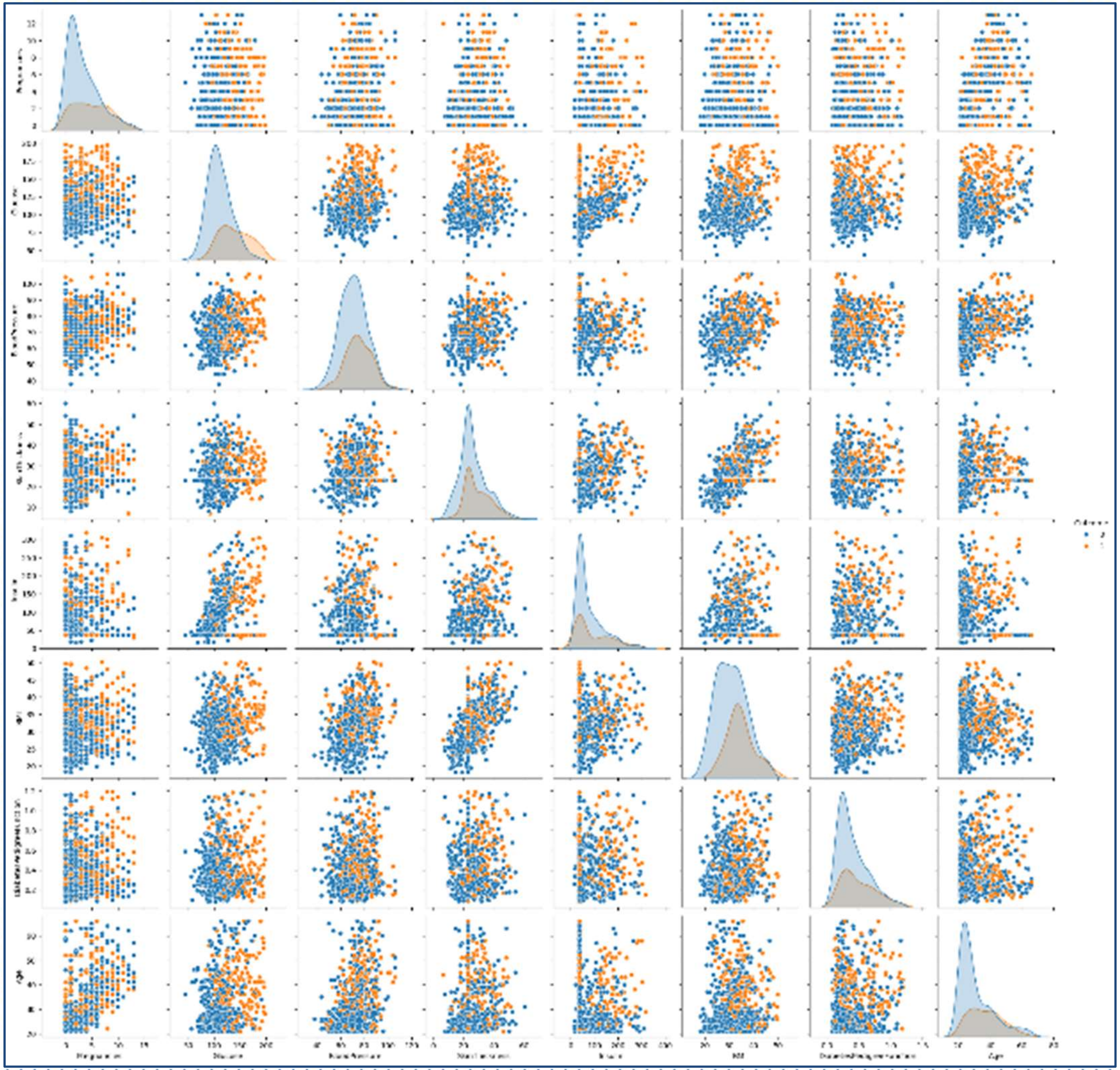
Initial Preprocessing:

- Cleaning the Dataset
- Identified Outliers
- Interquartile Range (IQR) method to remove outlier data



Dataset-1: Exploratory Data Analysis

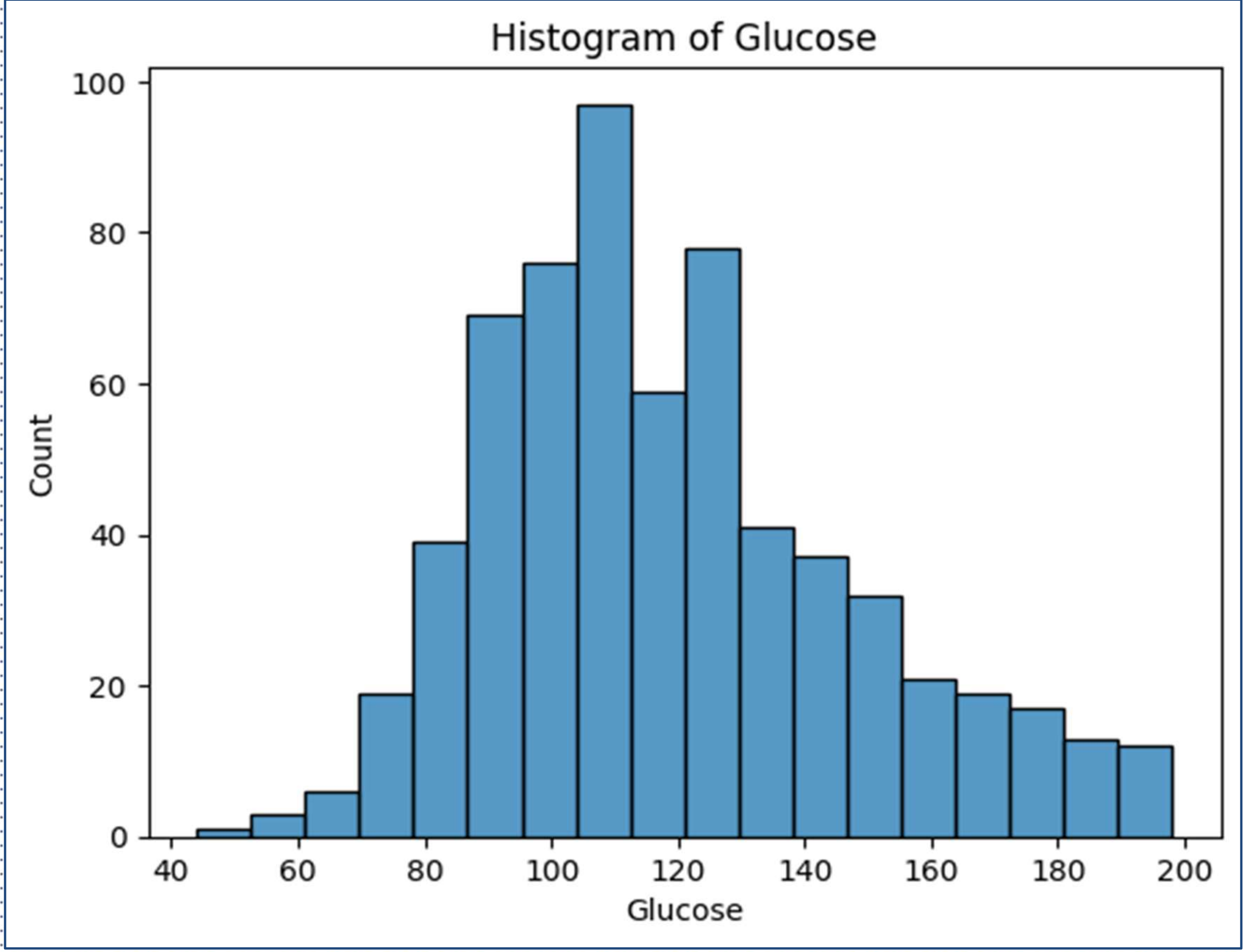
- **Box Plots:** Visualize data distribution and identify outliers
- **Pair Plots:** Illustrate relationships between features.
- **Identifying Distributions and Anomalies**



Dataset-1: Exploratory Data Analysis

Histograms:

- Show the frequency distribution of individual features.
- Assess data normality and skewness.
- Detect unusual data points that may affect model performance.



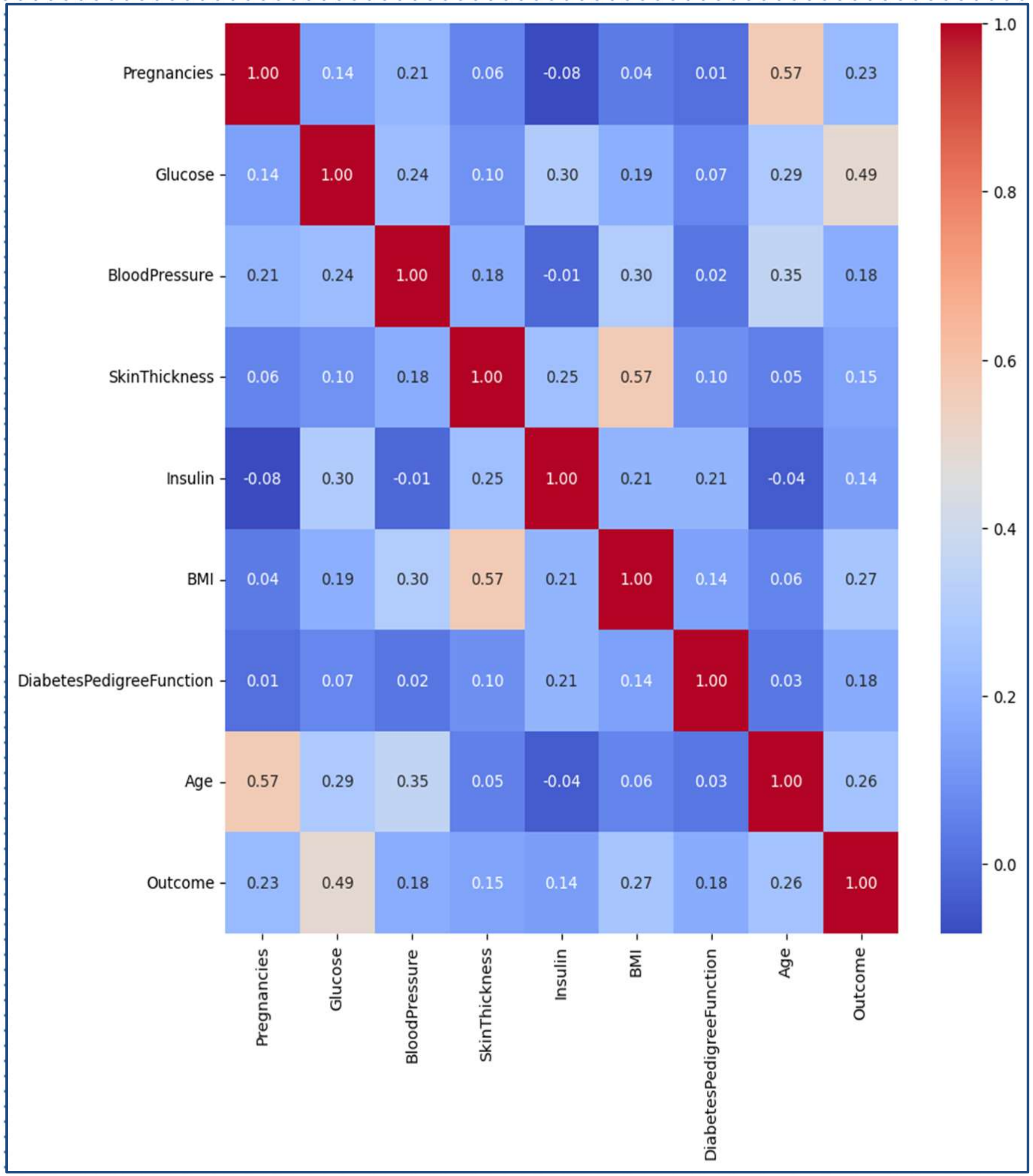
Dataset-1: Correlation and Standardization

Correlation Analysis:

- Heatmap: Visual representation of the correlation matrix.
- Purpose: Understand the relationships and dependencies between features.
- Insight: Identify highly correlated features that may impact model performance.

Data Standardization:

- MinMaxScaler: Applied to normalise the dataset.
- Purpose: Scale features to a uniform range.
- Benefit: Enhances model performance and convergence during training.



Dataset-2: UCI Diabetes Dataset

Dataset Description:

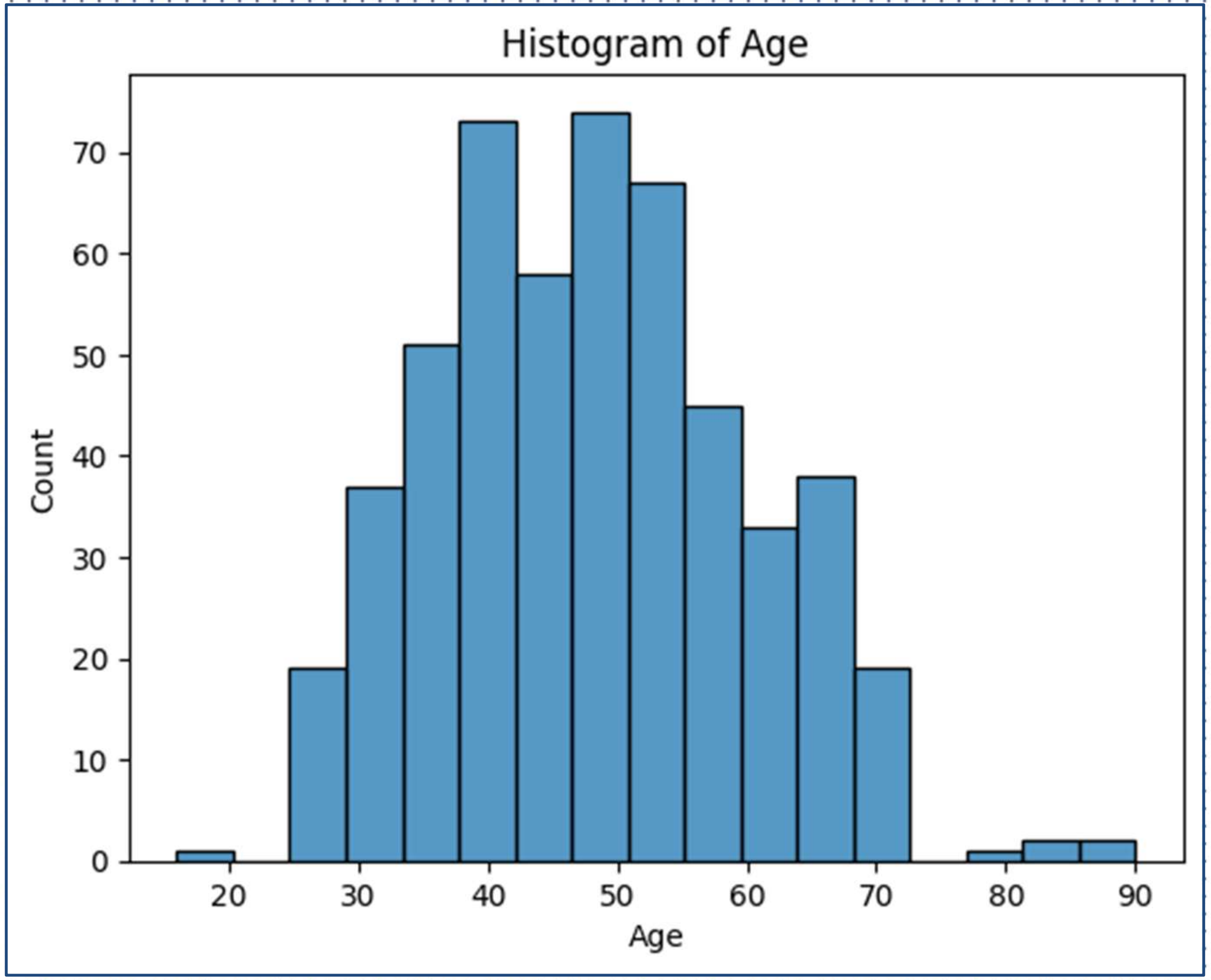
- UCI Diabetes Dataset used for studying diabetes prediction.
- Contains patient data with various features and symptoms.

Key Features:

- Age: Patient's age.
- Sex: Gender of the patient.
- Symptoms: Includes polyuria, polydipsia, sudden weight loss, etc.

Preprocessing Steps:

- Shuffling: Ensures random distribution of data to avoid bias.
- Label Encoding: Converts categorical variables into numerical format for ML algorithms.



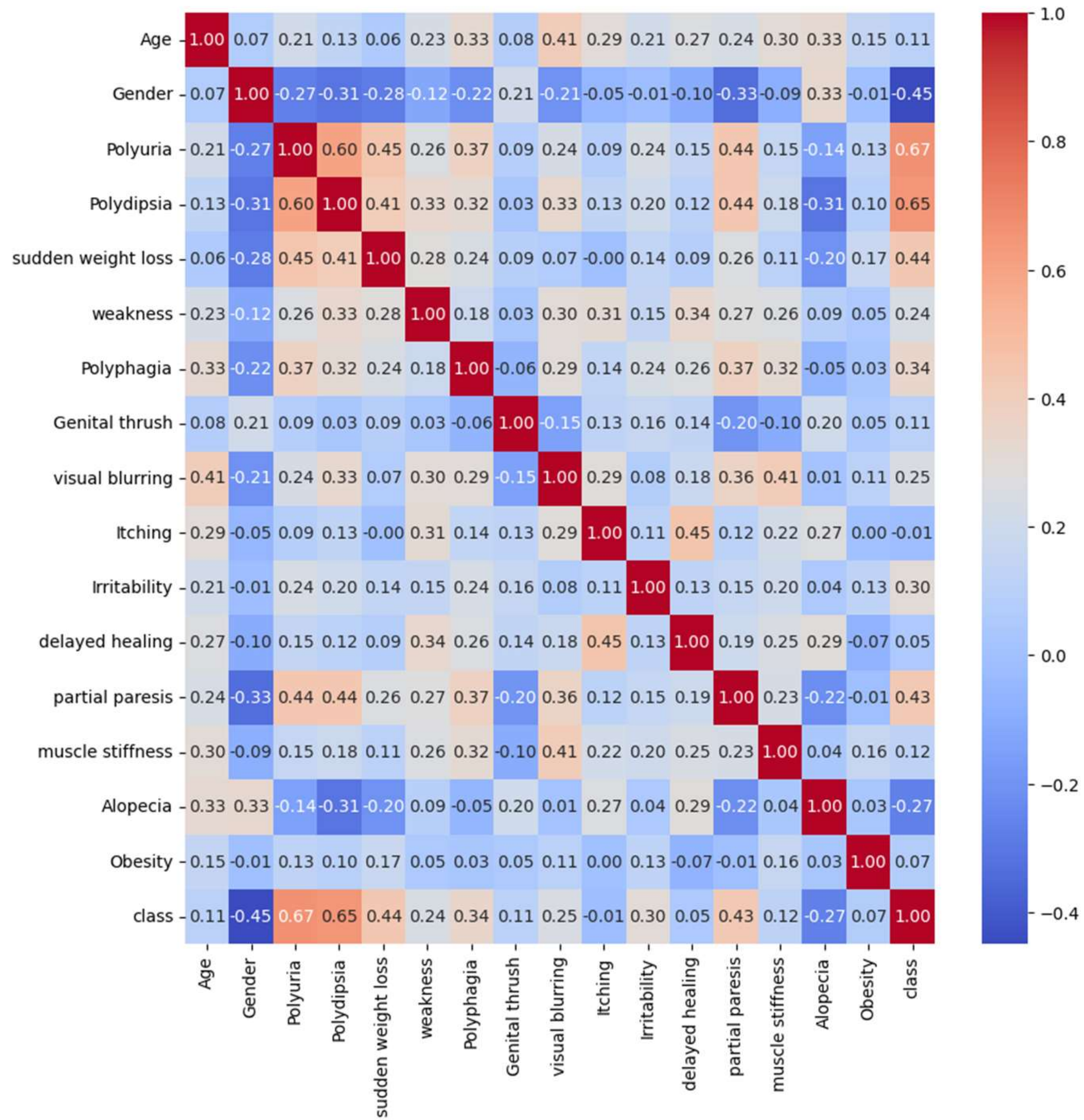
Dataset-2: Correlation and Feature Selection

Correlation Analysis:

- Heatmap: Visualizes the correlation matrix.
- Purpose: Examine relationships and dependencies between features.
- Insight: Identify features that are highly correlated.

Feature Selection:

- Basis: Selection based on correlation analysis and its impact on model performance.
- Removed Features: 'Delayed healing,' 'Itching,' and 'Obesity' were excluded due to low impact or high correlation with other features.



Dataset-2: Model Training and Evaluation

Models Used:

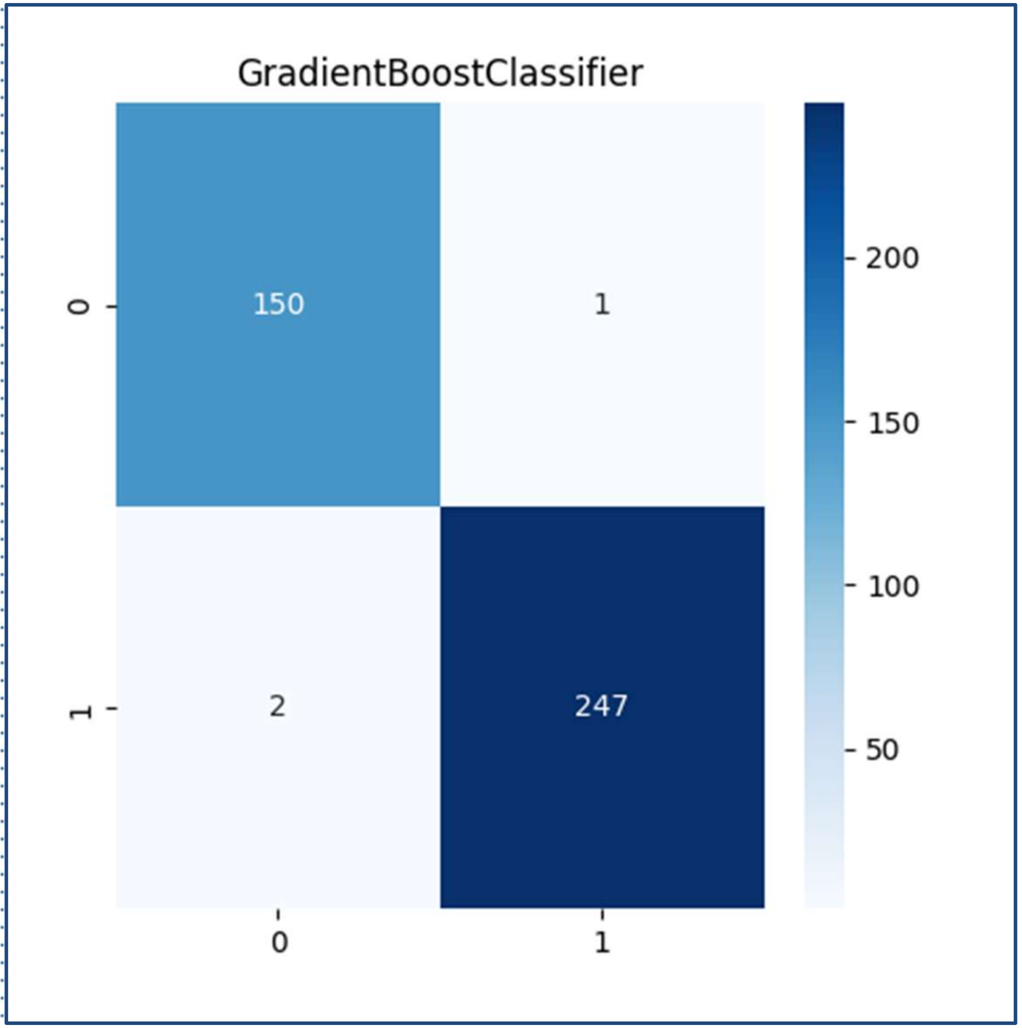
- Logistic Regression: Basic yet effective linear model for binary classification.
- Random Forest Classifier: Ensemble model using multiple decision trees.
- Gradient Boosting Classifier: Advanced ensemble model that builds trees sequentially to improve performance.

Evaluation Metrics:

- Accuracy: Overall correctness of the model.
- Confusion Matrix: Visual tool to understand true positives, true negatives, false positives, and false negatives.
- Classification Report: Provides precision, recall, and F1-score.

Performance:

Gradient Boosting Classifier: Outperformed other models, delivering the highest accuracy and best evaluation metrics.



Comparative Analysis of Models

Gradient Boosting Classifier:

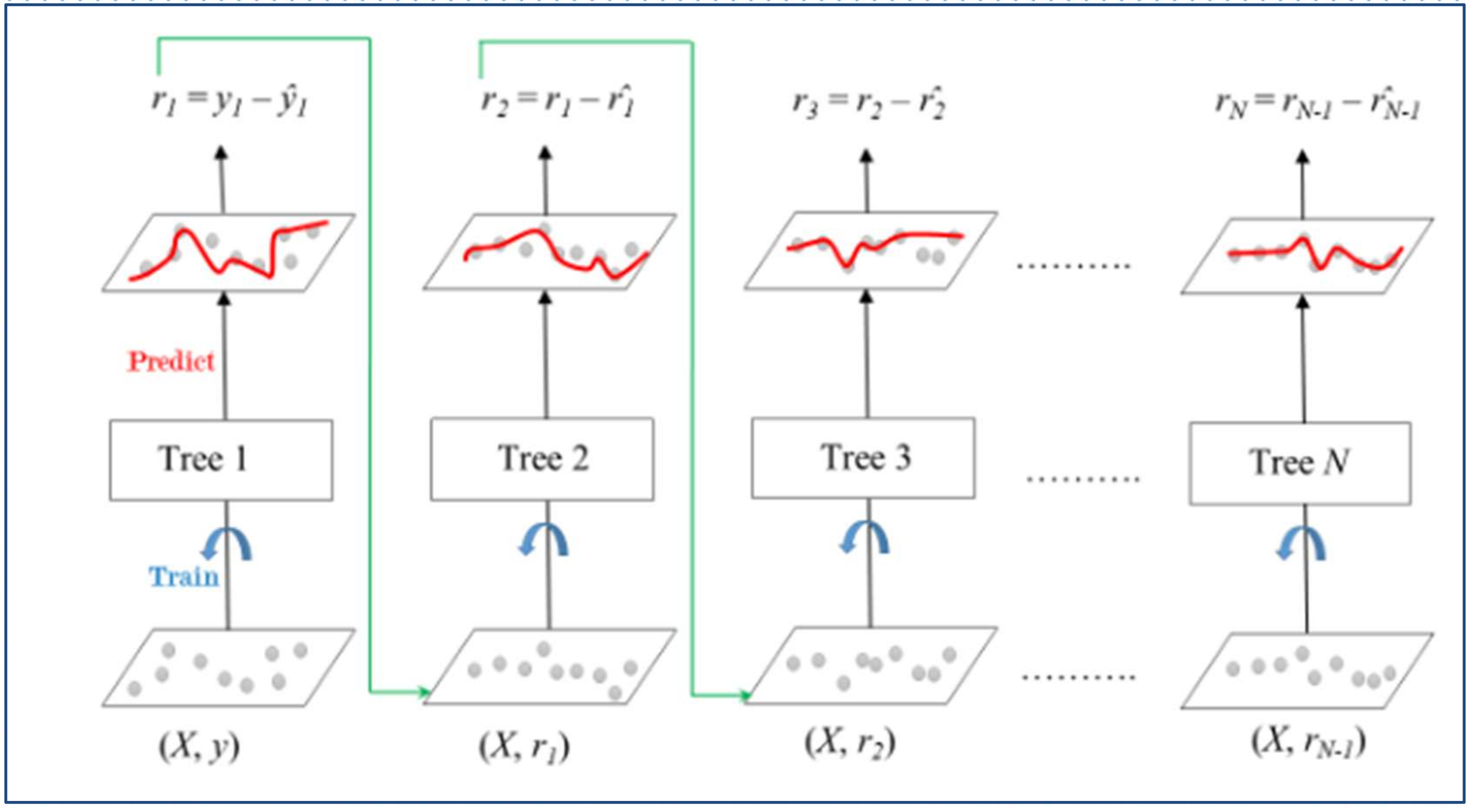
- Consistently achieved the highest accuracy across both datasets.

Logistic Regression:

- Better performance on Dataset 1 compared to Dataset 2.

Ensemble Methods:

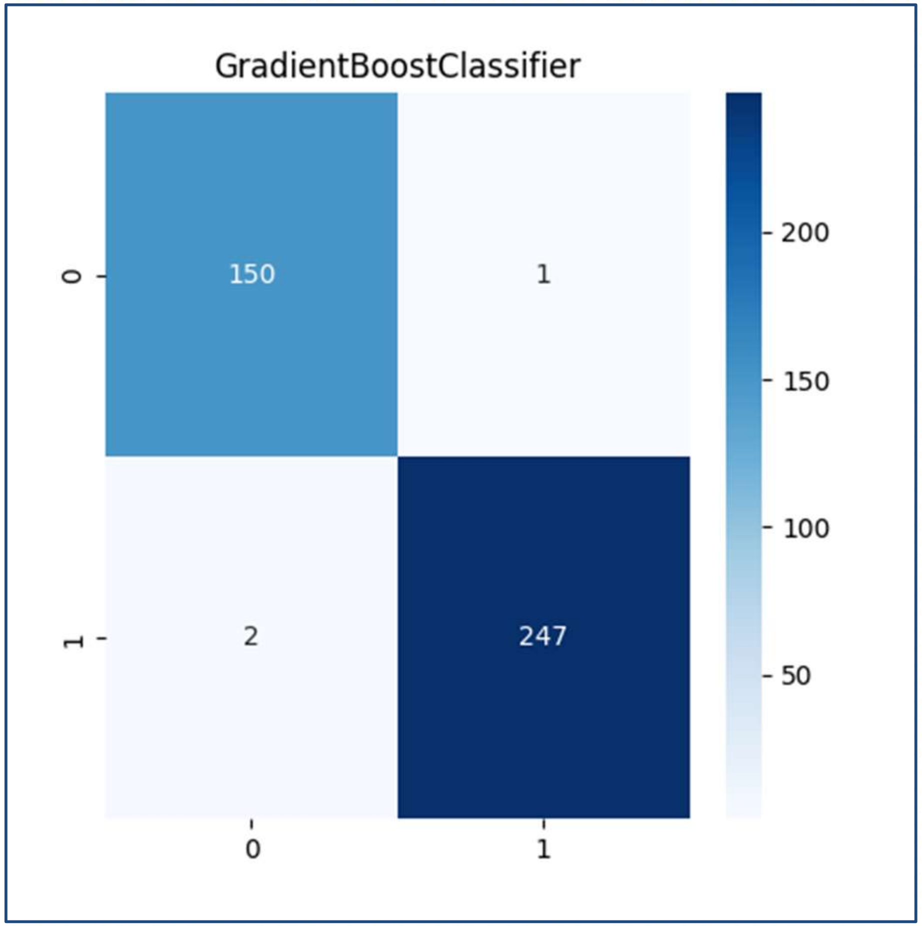
- Reduced false positives and false negatives, enhancing overall model reliability.

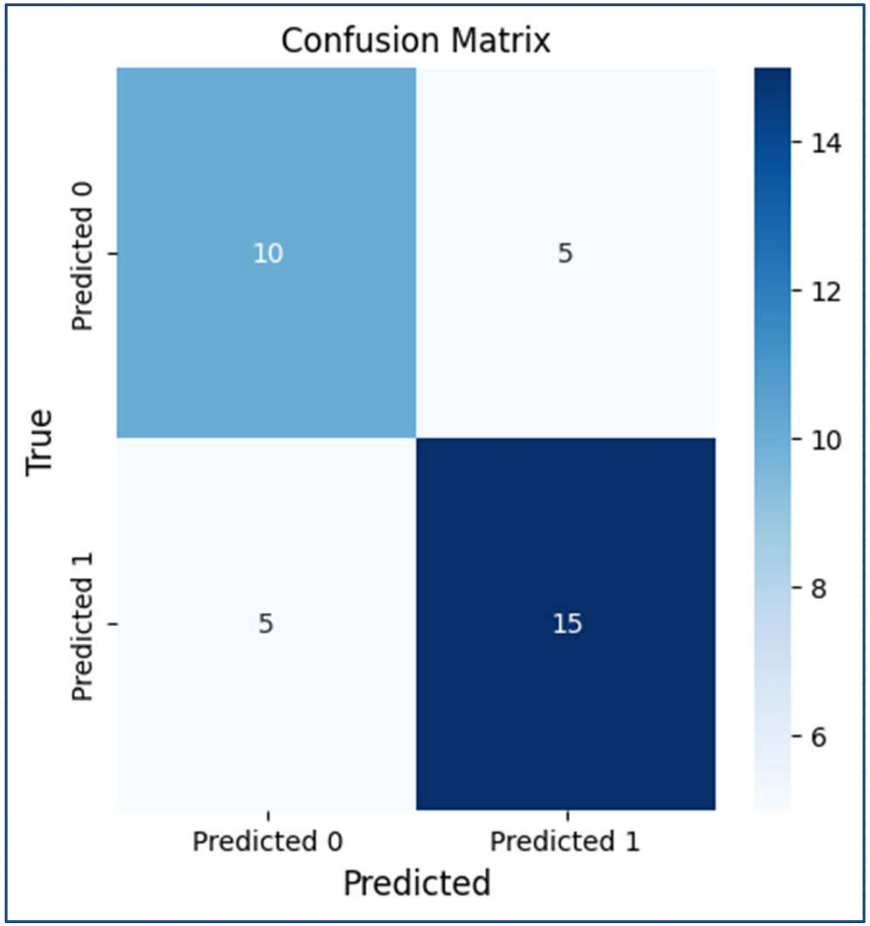


Key Insights:

- Gradient Boosting's sequential learning improved accuracy and robustness.
- Logistic Regression's simplicity proved effective for certain datasets.
- Ensemble methods, like Random Forest & Gradient Boosting, offered better error correction.

Results and Findings





Gradient Boosting Classifier: Emerged as the most robust and accurate model across both datasets.

Feature Engineering: Enhances model performance by selecting and transforming relevant features.

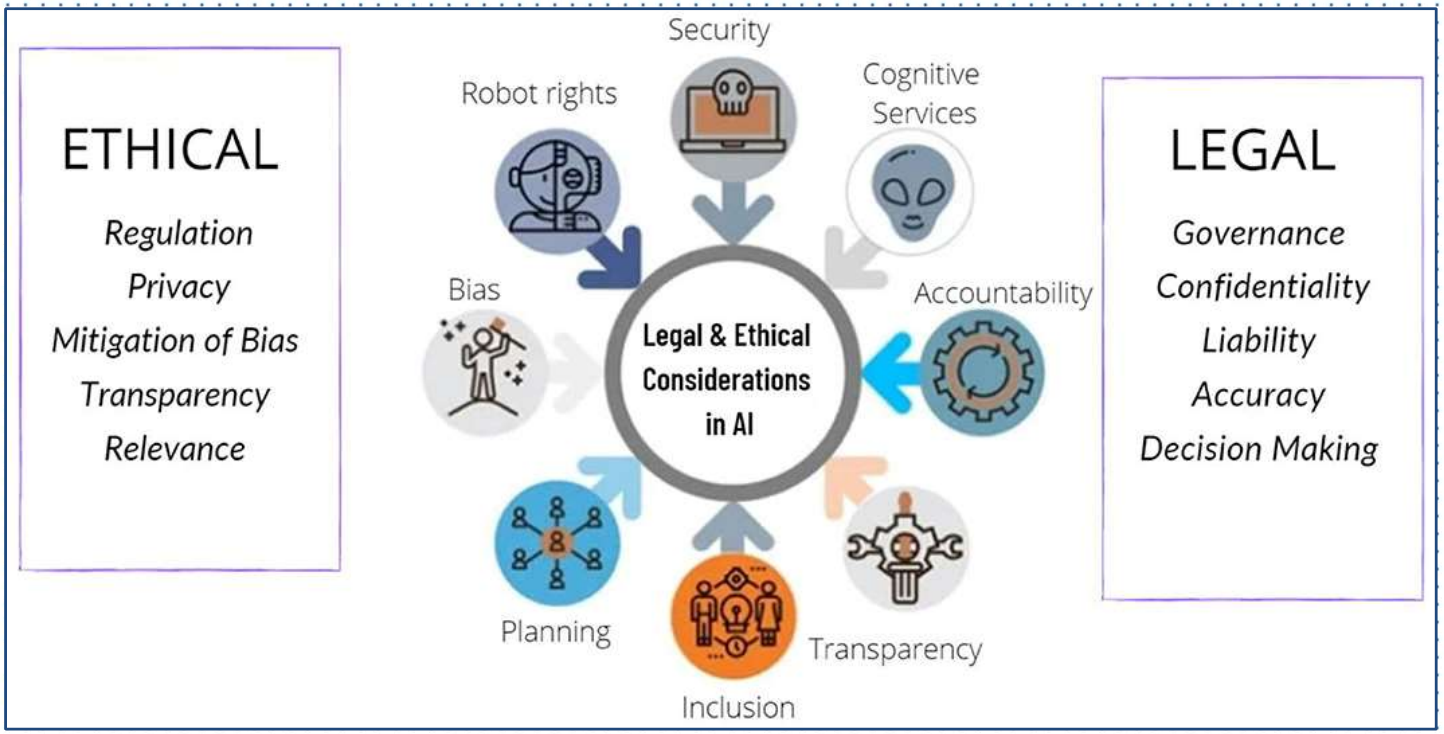
Comparison: Evaluated how different preprocessing methods.

Insights: Significant improvement in the accuracy and reliability of the models.

PDTs Ethical Implications: Issues and Solutions - I

Ethical Implications of G-AI in Healthcare:

- Data Privacy and Security: Ensuring patient data is protected against breaches and unauthorized access.
- Algorithmic Bias: Addressing biases in AI algorithms to prevent unfair treatment and disparities in healthcare outcomes.
- Diagnostic Accuracy: Ensuring the reliability and accuracy of AI-driven diagnostics to avoid misdiagnoses and harm to patients.
- Guidelines and Safeguards: Establishing comprehensive guidelines and safeguards to govern the ethical use of AI in healthcare, including transparency and accountability.



PDTs Ethical Implications: Issues and Solutions - II

Data Privacy Concerns: Data privacy concerns are of big concern, *Ahmadi-Assalemi et al.* [2020]; *Jahankhani et al.* [2020];

Algorithmic Biasing Issue: Algorithmic biases within predictive models may pose severe risks, *Prainsack* [2019]

Treatment Inequalities Issue:

1. DTs can challenge equality, *Bruynseels et al.* [2018]
2. Use of digital replicas of doctors in healthcare raises concerns about the social implications, potential negative impacts on doctor-patient communication, and liability issues, Zalake [2023].
3. There is necessity of addressing ethical considerations, privacy concerns, and potential negative impacts on doctor-patient communication to ensure responsible and equitable use of these technologies, Shah & Bhatt [2023], Cluitmans [2023], Zalake [2023] & Fontanari [2023].

End of Life Considerations:

1. Challenge of preserving child's place in care when using Digital Twin Systems in paediatric care, Maeckelberghe et al. [2023].
2. Ensuring equitable access to DTS and guaranteeing safety of systems for children is sensitive concern, David & Adrien [2022].

Individual Control Issues:

1. Interaction between the represented persons and their simulations raises questions regarding privacy, autonomy, and the control individuals have over their DTs, Iqbal et al. [2022].
2. PDTs presents a range of ethical challenges to data privacy, accuracy, and reliability, Mittelstadt [2021]

1. PDTs present both socio-ethical benefits and risks including privacy and data ownership, individuality, data-driven approach to healthcare, potential impact on a person's identity, disruption of societal structures, and potential for inequality, Popa et al. [2021].
2. PDTs ethical issues include behavioural and management attitudes, privacy, individuality, and the potential impact on equality, Kerckhove [2021].
3. Issues such as patient safety, privacy, and the rigorous ethical review of AI in multiple aspects including legal, humanistic, algorithmic, and informational aspects need to be addressed, Rubeis [2023].
4. Nanoparticle-based measurement used in PDTs may enable body hacking and the stealing of highly sensitive personal health data, Dirk & Javier [2022].

Solutions:

1. A conceptual process map can help in mitigating risks. Identify and address risks in the development of PDTs, Huang et al. [2022].
2. Exacerbate existing social injustices and marginalization must be carefully considered, Braun [2021].
3. By integrating **LLMs** with PDTs, healthcare professionals can leverage the predictive capabilities of LLMs to create more accurate and tailored digital representations of individuals, leading to more effective and personalized healthcare interventions, Shah & Bhatt [2023] and Cluitmans [2023].
4. Combination of **LLMs** and PDTs can democratize predictive access to healthcare, providing daily support to individuals and enabling precision wellness, prevention, and diagnostics, Fontanari [2023].

PDTs Ethical Implications: Issues and Benefits - III

Tabular classification of Issues based on bibliographic references from previous slides

Ethical Issue Level and Type		Ethical Issue
Data collection	Hyper-collection, Data quality and unorthodox use	Autonomy, Informed consent, Right to privacy, Surveillance health care, Distortion of the understanding of health
Data management	Data ownership, data accessibility? data brokerage, Hacking	Autonomy, Health equity, Informed consent, right to privacy, Transparency
Data analysis	Biased algorithms, Biased training data set	Discrimination or injustice, Distortion of the understanding of health
Information use	Decontextualization of disease formation, Epistemic injustice, Overdiagnosis, Quality compromise and secret sharing	Autonomy, Distortion of the understanding of health, Victim blaming, Damage physician–patient relationship, right to bodily integrity, Sensitive, personal content sharing, Third party data use against patient interests, Poor quality of data
Societal disruption	Responsibility and accountability, Patient-practitioner relationship, Disrupted self-care	Annoyance, Improper assessment of treatment risks and cautions, Over-reliance on doctor replica by PDTs, Unmonitored PDTs
Inequality and injustice	Modelling based on the white, healthy, middle-aged male, North-south, rich-poor injustice by access, excluding those who do not use the technology	Racism, Regionalism, Favouritism, Impartiality, Attitudinal approach, Communication gaps between doctor and patient
Other issues	Environmental risks, Technology push and lock-in, Crisis liability	Improper Training Modules, Annoyance towards awakening environment stakeholders, unethical handling of crisis situations

Benefits of PDTs Treatment and Technology	
Patient Health	Better diagnostics, less invasive treatments, Fewer side-effects, less error
Cost Reduction	Faster medicine discovery, shorter treatment periods, faster clinical trials, better logistics and maintenance
Patient autonomy	Patients have more power over treatment parameters, Patients take ownership of their body data
Other benefits	Fair and equal treatment of patients, Less animal suffering
Improved Health	Increased autonomy in life with new activities allowed, improved mental well-being, and decreased cognitive load, Improved social development, Increased participation in sports
Improved access to healthcare	Remote disease management, timely treatment adaptation, standardization of level of care

Conclusion and Further Scope

Summary of Research:

- Potential of PDTs: Personal Digital Twins (PDTs) can transform healthcare by providing personalized, data-driven insights.
- Superior Model: Gradient Boosting Classifier consistently performed best in predictive tasks.
- Importance of Preprocessing: Proper data preprocessing and feature engineering significantly enhance model accuracy and reliability.
- Ethical Considerations: Addressing data privacy, algorithmic bias, and establishing robust guidelines are crucial for ethical AI use in healthcare.

Potential of Personal Digital Twins (PDTs):

- Revolutionise Healthcare: Emphasize how PDTs can transform healthcare by providing personalized, data-driven insights for better diagnosis, treatment, and prevention.
- Improved Outcomes: Highlight the potential for enhanced patient outcomes through tailored healthcare solutions.

Importance of Continued Research:

- Overcoming Challenges: Stress the need for ongoing research to address existing challenges, such as data integration, privacy concerns, and algorithmic biases.
- Future Innovations: Encourage further innovations and advancements in AI and healthcare technologies.

Acknowledgement

- Pr. Amita Yadav, Assistant Professor, Maharaja Surajmal Institute of Technology, New Delhi - India; Contact: amita.yadav@msit.in
- Izhar Ahmad Hamdan, Bachelor's student at the Maharaja Surajmal Institute of Technology, New Delhi - India; Contact: izharhamdan@gmail.com
- Nilay Datta, Bachelor's student at the Maharaja Surajmal Institute of Technology, New Delhi - India; Contact: nilaydatta123@gmail.com
- Yuvraj Singh, Bachelor's student at the Maharaja Surajmal Institute of Technology, New Delhi - India; Contact: yuvrajam12@gmail.com

References: Technical and Bibliographic - I

Datasets:

- **UCI** Diabetes Dataset

- **PIMA** Indian Diabetes Dataset

Software and Tools:

- Key tools and libraries used for data analysis and model training: **Scikit-learn** and **Pandas**.

References: Technical and Bibliographic - II

Bibliography: Conference papers, white papers, and relevant online resources that contributed to the research:

1. *Tigard* [2021] - Daniel W. Tigard; 07 May 2021; Digital twins running amok? Open questions for the ethics of an emerging medical technology; https://jme.bmj.com/content/47/6/407.long
2. *Iqbal et al.* [2022] - Jeffrey David Iqbal, Michael Krauthammer, Nikola Biller-Andorno; 01 Jan 2022; The Use and Ethics of Digital Twins in Medicine; https://www.researchgate.net/publication/365496482_The_Use_and_Ethics_of_Digital_Twins_in_Medicine
3. *Mittelstadt* [2021] - B. Mittelstadt; 2021; Near-term ethical challenges of digital twins; https://philpapers.org/rec/MITNEC-2
4. *Braun* [2021] - Matthias Braun; 2021; Ethics of digital twins: four challenges; https://jme.bmj.com/content/48/9/579
5. *Shah & Bhatt* [2023] - Neel Shah, Nirav Bhatt, Nikita Bhatt; Dec 2023; Future of Large Language Models and Digital Twins in Precision Healthcare: A Symmetric Literature Review; https://www.researchgate.net/publication/376377075_Future_of_Large_Language_Models_and_Digital_Twins_in_Precision_Healthcare_A_Symmetric_Literature_Review
6. *Cluitmans* [2023] - Matthijs Cluitmans; 14 Jul 2023; How Happy is Your Virtual Brain? The Exciting Intersection of Digital Twins and Large Language Models; https://www.transformingmed.tech/p/how-happy-is-your-virtual-brain-the
7. *Zalake* [2023] - Mohan Zalake; 07 Dec 2023; Doctors' perceptions of using their digital twins in patient care; https://www.nature.com/articles/s41598-023-48747-5
8. *Fontanari* [2023] - Adriano Fontanari; 30 Aug 2023; Crossing the Chasm of Quantified Health: Towards LLM-enabled Digital Twins; https://www.linkedin.com/pulse/crossing-chasm-quantified-health-towards-llm-enabled-twins-fontanari
9. *Almazrouie* [2023] - Omar Almazrouie; 31 Mar 2023; Ethical implications of artificial intelligence in healthcare; https://www.journalijar.com/uploads/6427084fa472c_IJAR-42612.pdf
10. *McMillan et al.* [2023] - Hazem Zohny, John McMillan, Mike King; 24 Jan 2023; Ethics of generative AI; https://www.x-mol.net/paper/article/1617931149615644672
11. *Watts* [2018] - Ben Watts; 16 Oct 2018; Digital Twins Providing Personalised Medicine; https://www.challenge.org/knowledgeitems/digital-twin-technology-in-medicine/
12. *Habibzadeh et al.* [2020] - H. Habibzadeh, K. Dinesh, O. R. Shishvan. Boggio-Dandry A, Sharma G and Soyata T; A survey of healthcare internet of things (HIoT): A clinical perspective; https://ieeexplore-ieee-org.ressources-electroniques.univ-lille.fr/document/8863483
13. *Internet Things* [2022] - Internet Things J., vol. 7, no. 1, pp. 53–71, 2020. Y. Yang, H. Wang, R. Jiang et al., "A review of IoT-enabled mobile healthcare: technologies, challenges, and future trends; https://ieeexplore-ieee-org.ressources-electroniques.univ-lille.fr/document/9686065
14. *Gawel et al.* [2019] - Gawel DR, Serra-Musach J, Lilja S, Aagesen J, Arenas A; A validated single-cell-based strategy to identify diagnostic and therapeutic targets in complex diseases. Genome Med. 2019; https://pubmed.ncbi.nlm.nih.gov/31358043/
15. *Philips-USA* [2018] - Dynamic HeartModel A.I (Anatomical Intelligence) for additional insight. Philips. 2018. [2021-08-20]. https://www.usa.philips.com/healthcare/resources/feature-detail/ultrasound-heartmodel
16. *Copley* [2018] - Copley C. Medtech firms get personal with digital twins. Reuters. 2018. [2021-08-20]. https://www.reuters.com/article/us-healthcare-medical-technology-ai-insi-idUSKCN1LG0S0

References: Technical and Bibliographic - III

Continued from previous slide:

1. *Miller-EH* - E. H. Miller, "A note on reflector arrays (Periodical style - Accepted for publication)," *IEEE Trans. Antennas Propagat.*, to be published.
2. *Wang-J* - J. Wang, "Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style - Submitted for publication)," *IEEE J. Quantum Electron.*, submitted for publication.
3. *EDITH* - https://www.edith-csa.eu/edith/
4. *Tilda & Tino* - Tilda Herrgårdh & Tino Ebbers; Medical Digital Twin; https://liu.se/en/research/medigit
5. *Ahmadi-Assalemi et al.* [2020] - Gabriela Ahmadi-Assalemi; Haider Al-Khateeb, Carsten Maple, Gregory Epiphaniou, Zhraa A. Alhaboby, Sultan Alkaabi, Doaa Alhaboby; 07 Apr 2020; Digital Twins for Precision Healthcare; https://link-springer-com.ressources-electroniques.univ-lille.fr/chapter/10.1007/978-3-030-35746-7_8
6. *Ahmadi-Assalemi et al.* [2020] - Gabriela Ahmadi-Assalemi; Haider Al-Khateeb, Carsten Maple, Gregory Epiphaniou, Zhraa A. Alhaboby, Sultan Alkaabi, Doaa Alhaboby; 07 Apr 2020; Digital Twins for Precision Healthcare; https://link-springer-com.ressources-electroniques.univ-lille.fr/chapter/10.1007/978-3-030-35746-7_8
7. *Jahankhani et al.* [2020] - Hamid Jahankhani, Stefan Kendzierskyj, Nishan Chelvachandran, Jaime Ibarra; 2020; Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity; https://link-springer-com.ressources-electroniques.univ-lille.fr/book/10.1007/978-3-030-35746-7
8. *Prainsack* [2019] - Barbara Prainsack; 14 May 2019; Personalized Medicine: Empowered Patients in the 21st Century; New York University Press, Volume 98; Issue 2, Pages 1-3; https://academic.oup.com/sf/article-abstract/98/2/1/5489213?redirectedFrom=fulltext
9. *Braun* [2021] - Matthias Braun; 15 Mar 2021; Represent me: please! Towards an ethics of digital twins in medicine; https://jme.bmj.com/content/47/6/394
10. *Popa et al.* [2021] - Eugen Octav Popa, Mireille van Hilten, Elsje Oosterkamp, Marc-Jeroen Boga; 2021; The use of digital twins in healthcare: socio-ethical benefits and socio-ethical risks; https://lsspjournal.biomedcentral.com/articles/10.1186/s40504-021-00113-x
11. *Huang et al.* [2022] - Pei-hua Huang, Ki-Hun Kim, M. Schermer; 2022; Ethical Issues of Digital Twins for Personalized Health Care Service: Preliminary Mapping Study; https://www.jmir.org/2022/1/e33081/PDF
12. *Bruynseels et al.* [2018] - K. Bruynseels, Filippo Santoni de Sio, J. van den Hoven; 2018; Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm; https://www.academia.edu/58981319/Digital_Twins_in_Health_Care_Ethical_Implications_of_an_Emerging_Engineering_Paradigm
13. *Rubeis* [2023] - Giovanni Rubeis; Jan 2023; Hyperreal Patients - Digital Twins as Simulacra and their impact on clinical heuristics; https://www.researchgate.net/publication/366775988_Hyperreal_Patients_Digital_Twins_as_Simulacra_and_their_impact_on_clinical_heuristics
14. Kerckhove [2021] - Derrick de Kerckhove; 23 Mar 2021; The personal digital twin, ethical considerations; https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2020.0367
15. *Dirk & Javier* [2022] - Dirk Helbing, Javier Argota Sánchez-Vaquerizo; 2022; Digital Twins: Potentials, Ethical Issues, and Limitations; https://arxiv.org/ftp/arxiv/papers/2208/2208.04289.pdf
16. *Maeckelberghe et al.* [2023] - Els Maeckelberghe, Kinga Zdunek, Sara Marceglia, Bobbie Farsides, Michael L. Rigby; 19 Jun 2023; The ethical challenges of personalized digital health; https://www.researchgate.net/publication/371709696_The_ethical_challenges_of_personalized_digital_health
17. *David & Adrien* [2022] - David Drummond; Adrien Coulet; 01 Oct 2022; Technical, Ethical, Legal, and Societal Challenges With Digital Twin Systems for the Management of Chronic Diseases in Children and Young People; https://pubmed.ncbi.nlm.nih.gov/36315239/
18. *Smith-B* - B. Smith, "An approach to graphs of linear forms (Unpublished work style)," unpublished

Thank you for your attention!

https://lumen.univ-lille.fr https://zydususa.com/ www.univ-lille.fr https://www.wlv.ac.uk