
Predicting Disease Outbreaks Using Healthcare Data Trends

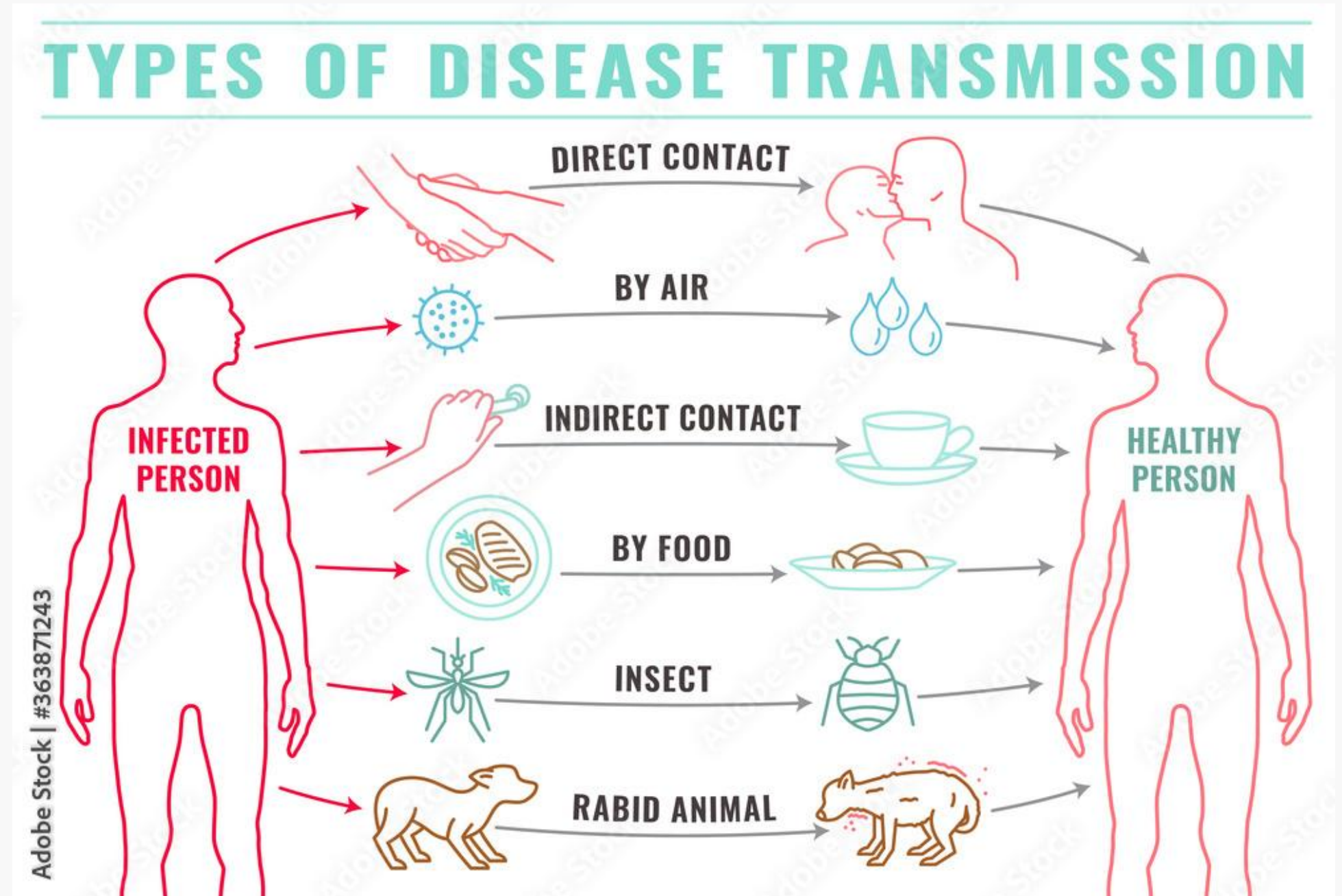
Presented by: [SNEHA BABARIYA]

[STUDENT NUMBER: 2409818]

[Research Methods and Professional Development - ACFA7004]

Definition of Outbreak in NHS

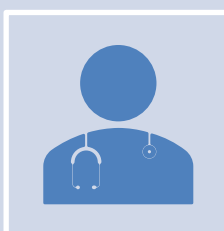
- an incident in which at least 2 or more people affected by the same infectious disease are linked by time, place, or common exposure



Research Context/Background



Global Health Concerns: These rising infectious diseases are global health concerns since they threaten all people's health and need immediate diagnosis systems to contain their harmful potential.



Role of Data Science: Data science has been changing the process of disease prediction by allowing the processing of such a large quantity of healthcare data via the usage of Artificial Intelligence (AI) and Machine Learning (ML) models to search for early warning carriers [1].



Need for Proactive Approaches: ML-driven analytics, when done proactively, helps in disease intervention at a stage before it becomes a nuisance, thereby reducing healthcare responsibility worldwide and helping curb disease transmission rates.

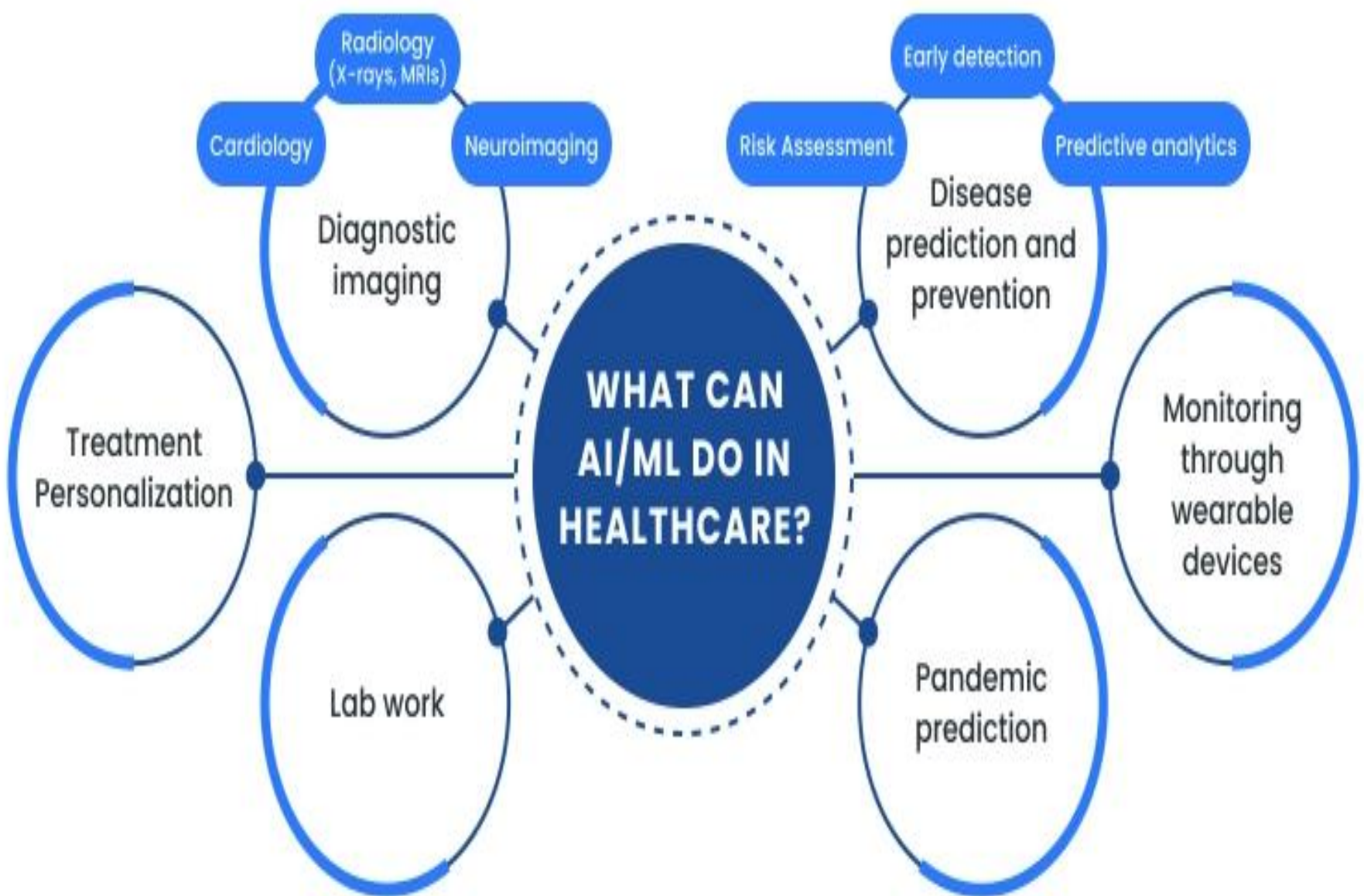


Figure: AI in Health-Care

Research Problem

1

Existing Challenges: Manual reporting generates major delays in traditional disease surveillance, and it is limited in its capacity to respond to outbreaks as they happen.

2

Data Utilization Gap: Existing systems do not utilize these data sources for predictive purposes and in real-time, such as with IoT, electronic health records, and social media [2].

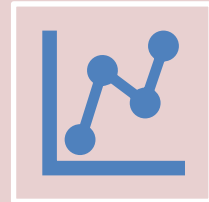
3

Unresolved Issues: Existing predictive models are largely inaccurate due to a lack of data quality consistency, bias, and poor feature selection, which makes it an unresolved issue to reliably detect early outbreaks.

Research Aim and Objectives

Research Aim:

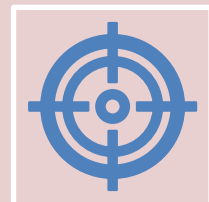
The aim is to develop a machine learning model that can identify the possibility of disease outbreaks and monitor healthcare trends.



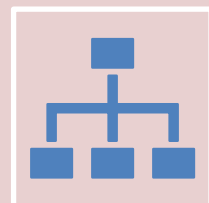
Data-Driven Insights: Collecting and preprocessing healthcare data, including patient records, social media trends, and environmental factors .



Model Development: To perform on implementing and comparing the developed machine learning and deep learning models for outbreak prediction



Performance Evaluation: Assessing model accuracy using key metrics like precision, recall, and F1-score to ensure reliability



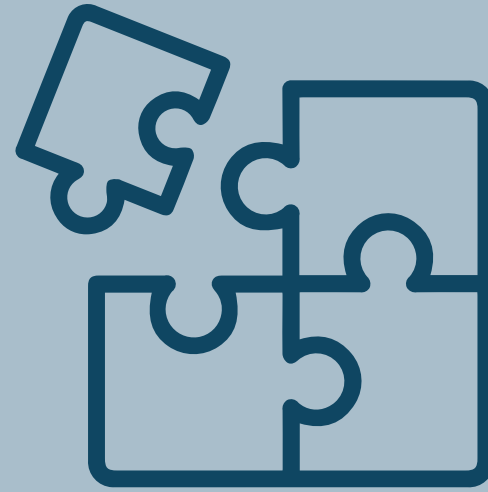
Comprehensive analysis: To discuss and compare the model's performance with traditional processes

Literature Review and Gap Identification



Existing Studies

- A number of current research apply machine learning (ML)-based models to predict disease outbreaks, which show significant promise but struggle to be implemented in real-time with high accuracy and dependability.



Identified Gaps

- Current studies are lacking in the comprehensive integration of multiple healthcare data resources including EHR, IoT, and social media trends [3].



Research Focus

- The goals of this study are to increase the accuracy of prediction by using deep learning techniques to improve outbreak detection.

Research Methodology

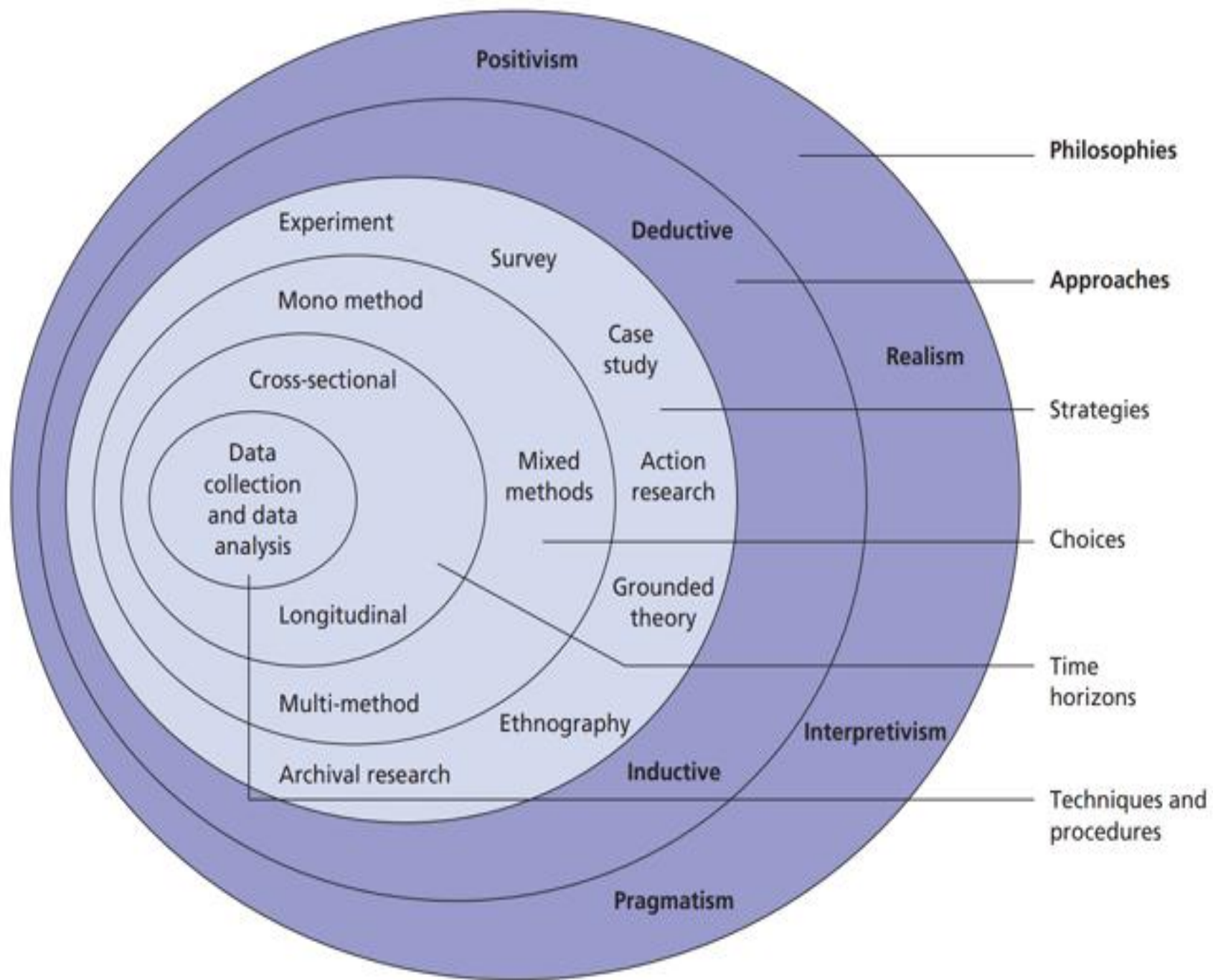


Figure: Research Onion Model [4]

1. Research Philosophy

- **Positivism:** The study will follow a structured approach, using objective data for disease outbreak prediction.

2. Research Approach

- **Deductive:** Hypotheses will be tested using machine learning models on healthcare datasets.

3. Research Strategies

- **Experiment:** Evaluating ML models for accuracy in outbreak prediction.
- **Case Study:** analyzing real-world disease trends for validation.

4. Time Horizon

- **Cross-Sectional:** Data will be analyzed at a given point to identify patterns.

5. Data Collection & Analysis

- **Quantitative Approach:** Structured datasets, ML model evaluation, statistical validation (precision, recall, F1-score).

Methodology (Contd...)

Steps

Data Collection: To Gather disease outbreak datasets from Kaggle, electronic health records.

For now still dataset are not final I am searching on best dataset. Then I will work on model selection.

Model Selection: The predictive model will include ML Tool & techniques such as in python Random Forest and support vector machine (SVM) alongside the deep learning methods Long Short-Term Memory(LSTM) and Convolutional Neural Network (CNN) for pattern detection [5].

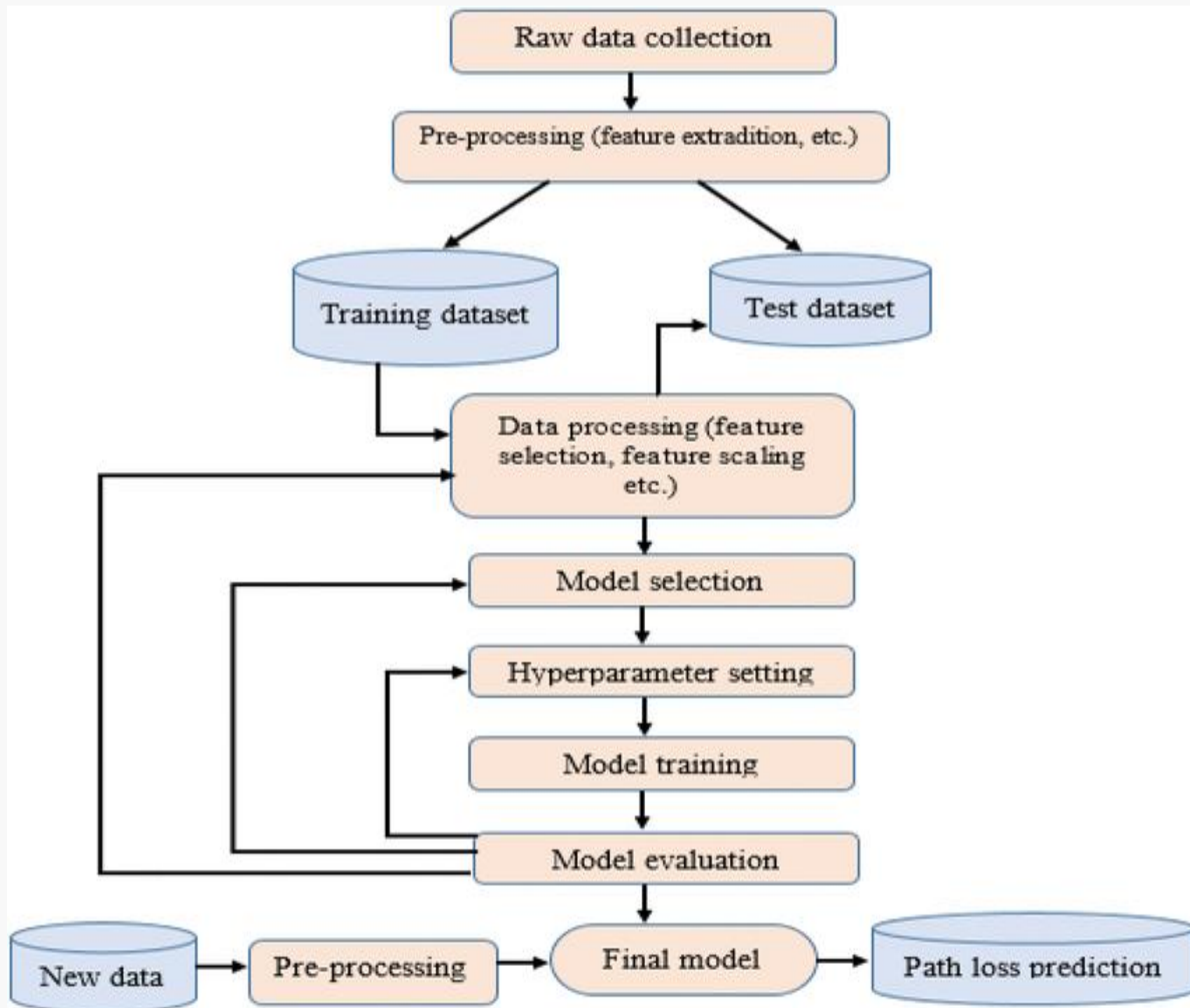


Figure: Figure: Flow chart for Model developemnt

Methodology (Contd...)

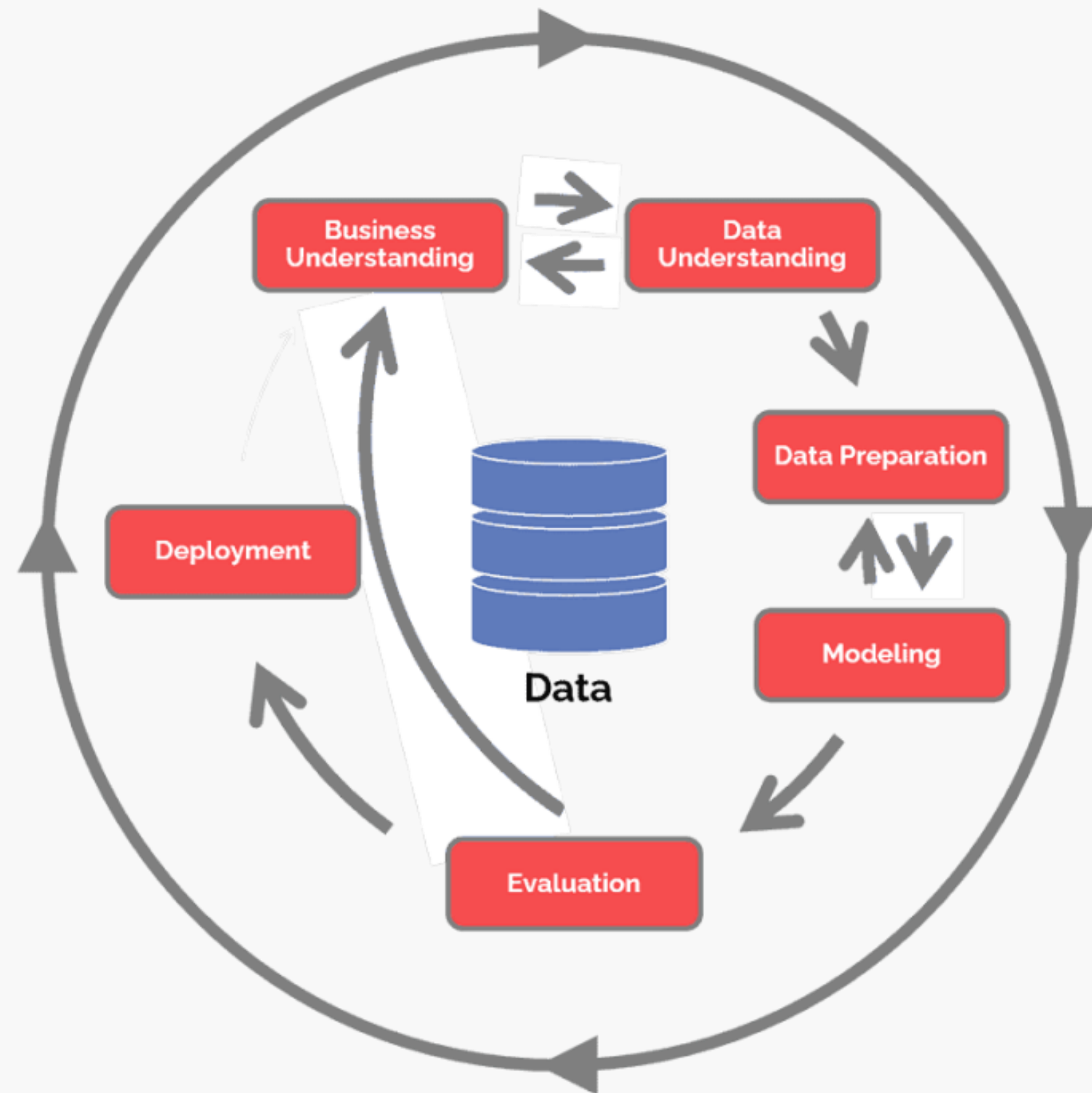


Figure: Experimental Work by following Crisp-DM model

I select CNN and LSTM because this method used in COVID-19 To predict the number of COVID-19 cases in Brazil, India, and Russia, three deep learning models, including CNN, LSTM, and the CNN-LSTM have been developed . The accurate analysis confirmed that the LSTM model had the highest performance.

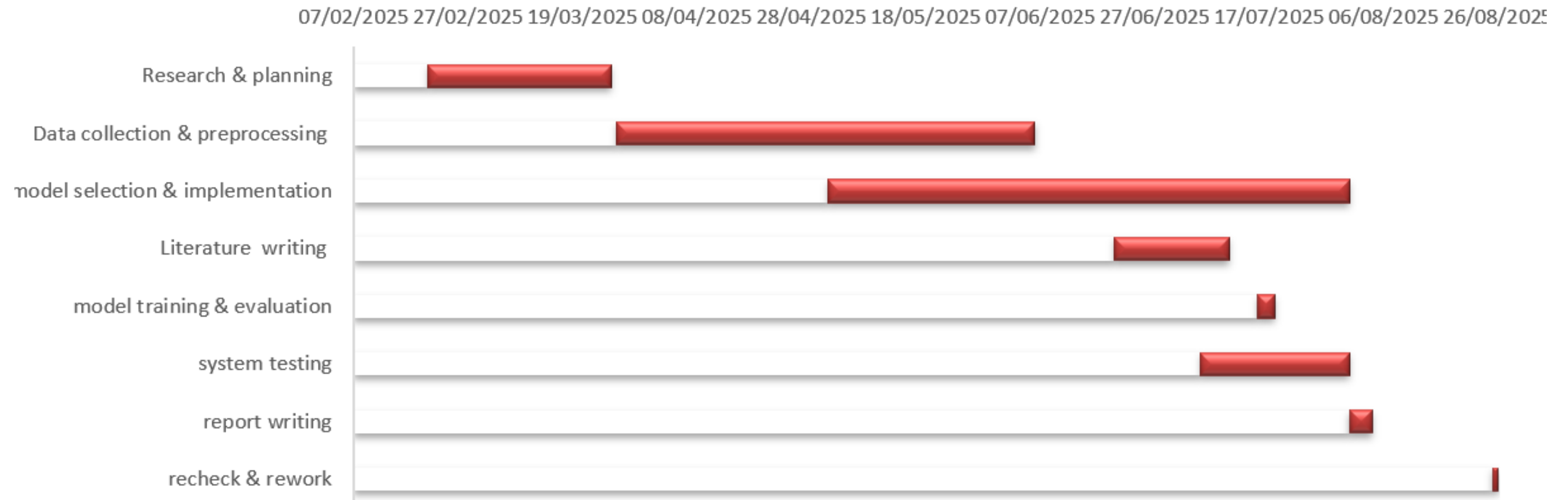
Evaluation Metrics: Need to perform on the evaluation metrics. These include the precision and recall and F1-score and Area Under the Receiver Operating Characteristic curve (AUC-ROC) for validating accurate outbreak predictions and model performance.

Methodological Framework: The famous CRISP-DM framework will be considered for the project work.

Project Planning and Resources

- Timeline: The create project plan will be followed here for completing the work with progressive way.
- Resources Required: This project needs four primary resources which include specialized computing capacity together with various healthcare information sources alongside sophisticated software platforms for machine learning deployments [6].

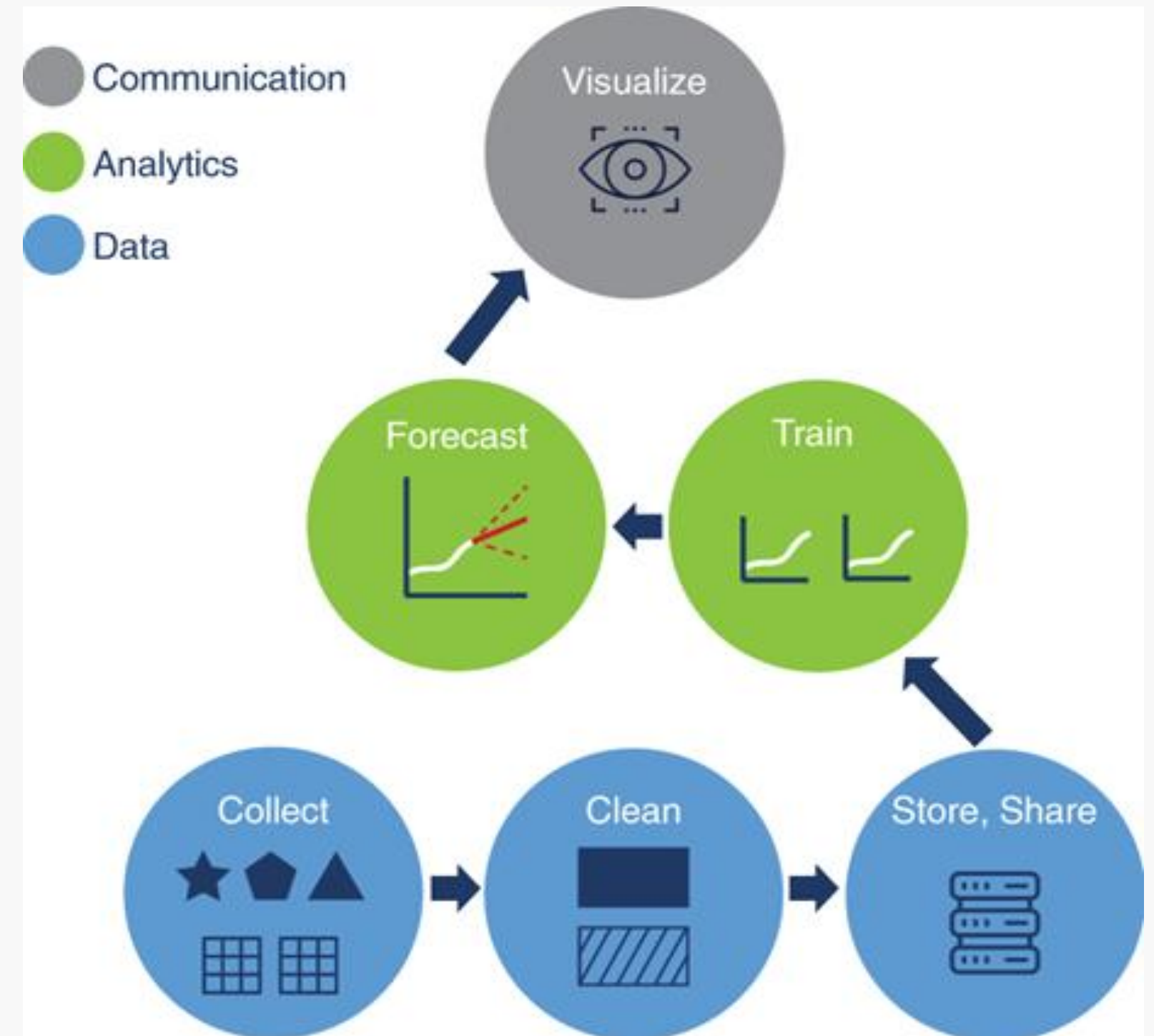




Project Planning and Resources

Project Impact and Limitations

- Potential Benefits: Early detection of disease outbreaks produces benefits where healthcare institutions can implement quick interventions to reduce health system costs as well as enhance public health responses.
 - Limitations:
 - The sensitive information present in healthcare records raises privacy issues that create limitations.
 - For this collecting real-time report data access is the limitation.
-



Conclusion and Next Steps

Summary:

The implementation of ML-based outbreak prediction systems helps global healthcare organizations detect threats in advance and implement preventive measures.

Implementation Strategy:

The developed model will be installed within healthcare platforms to process current data effectively for accurate prediction outputs. Working hand in hand with healthcare groups enables effectiveness.

References

- [1] G. T. Igwama, et al., "Big data analytics for epidemic forecasting: Policy Frameworks and technical approaches," *International Journal of Applied Research in Social Sciences*, vol. 6, no. 7, pp.1449–1460, 2024.
 - [2] O. E. Santangelo, et al., "Machine learning and prediction of infectious diseases: a systematic review," *Machine Learning and Knowledge Extraction*, vol. 5, no. 1, pp.175–198, 2023.
 - [3] X. Y. Leung, et al., "A systematic review of dengue outbreak prediction models: Current scenario and future directions," *PLoS neglected tropical diseases*, vol. 17, no. 2, pp.e0010631, 2023.
 - [4] M. Saunders, P. Lewis and A. Thornhill, *Research methods for business students*, Pearson education, 2009.
 - [5] M. Niaz and U. Nwagwu, "Managing healthcare product demand effectively in the post-covid-19 environment: navigating demand variability and forecasting complexities," *American Journal of Economic and Management Business (AJEMB)*, vol. 2, no. 8, pp.316–330, 2023.
 - [6] N. Absar, et al., "The efficacy of deep learning based LSTM model in forecasting the outbreak of contagious diseases," *Infectious Disease Modelling*, vol. 7, no. 1, pp.170–183, 2022.
-

Thank you
