

Infectious Disease Prediction Algorithms Using Medical Knowledge Base for the Decision Support System Regarding the Risk of Epidemic Threats on Sea-going Vessels – DESSEV

N. Wawrzyniak¹, T. Gregorič², N. Blek³, I. Bodus-Olkowska¹, I. Garczyńska¹, A. Chronopoulos⁴, V. Makar⁵, J. Lahtinen⁶ & G. de Melo Rodriguez⁷

¹ Maritime University of Szczecin, Szczecin, Poland

² Spinaker d.o.o., Portoroz, Slovenia

³ Maria Skłodowska-Curie Medical Academy, Warsaw, Poland

⁴ IDEC SA, Pireas, Greece

⁵ Centre for Factories of the Future Ltd., Alingsås, Sweden

⁶ Satakunta University of Applied Sciences, Pori, Finland

⁷ Technical University of Catalonia, Catalonia, Barcelona, Spain

ABSTRACT: Epidemic outbreaks on sea-going vessels pose a significant health and safety risk, particularly in isolated maritime environments lacking professional medical staff. The COVID-19 pandemic exposed major deficiencies in the preparedness and response protocols of maritime operations, where vessels became de facto quarantine units without access to structured medical decision support. This study presents the development and evaluation of DESSEV, a decision support system designed to assist maritime personnel in identifying and managing infectious disease outbreaks onboard vessels. A comprehensive non-SQL knowledge base was constructed using international epidemiological guidelines (WHO, CDC, ECDC) and medical literature, encompassing 22 infectious diseases and 35 symptoms categorized into 8 clinical domains. To support probabilistic diagnostics, artificial patient profiles were generated reflecting real-world symptom distributions. Three predictive models—Decision Tree, Naive Bayes, and Random Forest were trained. Their performances were assessed through cross-validation and random sampling techniques. Evaluation on a holdout test set of real patient cases showed that the Random Forest model achieved superior performance across all. The Random Forest algorithm was thus selected as the core prediction engine for the DESSEV application. While not a substitute for professional medical care, DESSEV improves situational awareness and supports early risk mitigation actions aboard vessels. Future work will focus on integrating real-time health telemetry and user feedback to further refine diagnostic accuracy and usability.

1 INTRODUCTION

Maritime transport plays a vital role in global logistics, with approximately 90% of goods transported by sea, making vessels both essential assets and potential vectors for disease transmission. The COVID-19 pandemic starkly highlighted the lack of medical infrastructure on a vessel and the absence of real-time, evidence-based guidance for managing infectious diseases at sea. These shortcomings, compounded by international communication barriers, limited access to shore-based healthcare, and diverse crew compositions, necessitated a solution tailored to the maritime environment. Since 2013 the EU SHIPSAN

ACT information system has provided Member-State inspectors and ship operators with a harmonised platform that links pre-boarding risk assessment, voyage monitoring and post-inspection feedback, underpinned by the European Manual for Hygiene Standards and Communicable-Disease Surveillance on Passenger Ships. The manual codifies evidence-based check-lists and response algorithms that are now compulsory during port-state control in many EU countries [1]. Building on SHIPSAN, the EU HEALTHY GATEWAYS Joint Action (2018-2023) expanded the scope from cruise liners to all points of entry, issuing dynamic guidance (e.g., mpox and SARS-CoV-2 advisories) and an electronic catalogue of

best practices that can be imported into local contingency plans [2]. The United States' CDC Vessel Sanitation Program offers a parallel model outside Europe: routine ship inspections, real-time AGE outbreak dashboards and a publicly available operations manual that doubles as a decision-support reference for shipboard environmental health officers [3]. Another approach is to use telemedicine and remote clinical decision support as described in [4] on Telemedical Maritime Assistance Services (TMAS) allowing ashore physicians to run protocol-driven triage and treatment pathways for infectious diseases. And finally, progressing development of early-warning and prediction platforms using various artificial intelligence (AI) methods is present in complementary academic work. Triantafyllou et al. (2024) propose a closed-loop architecture that couples environmental sensors, agent-based epidemic simulation, and optimisation algorithms to recommend ventilation or isolation measures in real time [5]. Similar operational-research studies re-examining the Diamond Princess outbreak distilled an emergency-response mechanism highlighting the value of predictive modelling, structured communication chains and adaptable quarantine layouts [6]. The systems above concentrate on inspection compliance, strategic preparedness, macro-level horizon scanning. None of them supplies ship masters with a low-bandwidth, multilingual, symptom-based diagnostic engine that operates offline and converts probabilistic outputs into plain-language action cards—DESSEV's defining features.

The DESSEV (Decision Support System regarding the risk of Epidemic threats on a sea-going Vessel) project key objective was to develop a Decision-Support System that enhances maritime safety by providing comprehensive risk assessments and mitigation strategies for the spread of infectious diseases on seagoing vessels. DESSEV addresses a longstanding but under-theorized challenge in maritime health security: how to empower shipboard personnel to identify and manage epidemic threats in the absence of medical professionals. Its foundations lie in three integrated components: a data repository, a medical knowledge base, and a decision support platform. It focused on a group of infectious diseases chosen for their significant global health impact, frequency of occurrence, and particular relevance to maritime environments. The selection was done by data from reputable sources including the World Health Organization (WHO), the Centers for Disease Control and Prevention (CDC), and the European Centre for Disease Prevention and Control (ECDC). Additionally, the process was supported by insights from peer-reviewed research, case study analyses, and consultations with medical experts. The materials in the repository are classified by relevance and serve as educational resources for vessel crews and port authorities. More on the development of the repository can be found in [7].

Although the initial concept of the DESSEV project was to build rule-based decision support systems, it quickly became evident that medical cases are not that straightforward, as they rarely follow binary (0–1) logic. Patients may experience the same symptom with varying intensity—or not at all, still suffering for the same disease. For instance, a rule such as "If you have

a sore throat THEN you have pharyngitis" would not always hold true, as some individuals might suffer severely, while others may not feel the symptom so intense at all. On the other hand, the rapid development of machine learning techniques allowed us to pivot toward a more probabilistic and data-driven architecture. This led to the integration of well-established models that can infer complex patterns and non-linear relationships between multiple symptoms and disease classes. Such models are better suited for capturing the inherent uncertainty and variability in clinical presentations. They offer not only higher predictive accuracy but also enhanced generalizability to real-world scenarios where inputs may be incomplete, subjective, or overlapping. Consequently, the DESSEV decision support system evolved into a hybrid platform that maintains the interpretability needed for user trust while leveraging the analytical power of machine learning to support more nuanced and flexible decision-making under uncertainty.

2 MACHINE LEARNING MODELS IN CLINICAL DECISION SUPPORT SYSTEMS (CDSS)

Clinical decision support systems (CDSS) originally relied on manually curated rules (e.g., Arden syntax). The exponential growth of electronic health records and medical imaging has shifted research toward machine-learning pipelines that learn predictive patterns directly from data. Recent systematic reviews document more than 10 000 ML-centred CDSS publications since 2020, with sharp upticks in deep-learning and large-language-model (LLM) approaches [8]. Structured diagnostic codes, laboratory series and vital-sign streams now feed recurrent, Transformer and graph networks that forecast unplanned admissions, sepsis or even cardiovascular events [9]. Convolutional and hybrid composite networks continue to raise the performance ceiling in radiology and pathology [10]. Nevertheless, decision-support systems aimed at front-line users with limited computing resources, such as shipmasters, benefit from classifiers that are auditable, robust to sparse or noisy inputs and computationally lightweight. The classical family of Naïve Bayes (NB), Decision Tree (DT) and Random Forest (RF) models fulfils these requirements, and recent biomedical literature confirms their continued relevance despite the rise of deep learning [11,12].

NB delivers explicit posterior probabilities that map naturally onto the low/medium/high risk language used in maritime health protocols. Instructional materials for health-informatics curricula still present NB as the canonical algorithm for diagnostic testing and symptom triage because it trains in milliseconds and generalises well on small, categorical feature sets. Ensemble bagging of many shallow decision trees boosts discrimination without sacrificing interpretability; variable-importance plots immediately reveal which symptoms drive predictions. Decision trees offer several advantages, making them particularly useful in healthcare settings. They are easy to interpret and visualize, even for individuals without technical expertise, and they can handle both numerical and categorical data effectively. By aggregating the decisions of many independently

trained trees, the Random Forest model provides greater accuracy, resilience to noise, and improved generalization. In medical applications, it is particularly valuable for its ability to handle complex, high-dimensional data and to model intricate interactions between symptoms.

3 MEDICAL KNOWLEDGE BASE CONSTRUCTION: DISEASE SELECTION AND SYMPTOM STRUCTURING

The medical database was structured across two sets of tables. The first set of tables constructed in a nonSQL structure, includes 22 of the most globally impactful infectious diseases, that pose substantial public health risks, including COVID-19, dengue, malaria, and influenza. These diseases were selected due to their potential to spread rapidly among a vessel's crew members, potentially leading to a quarantine and restricted access to ports (Tab. 1).

Table 1. List of selected diseases.

chickenpox	mumps
chickungunya	norovirus
cholera	pertussis
COVID-19	rabies
dengua	rubella
diphtheria	tetanus
ebola	tuberculosis
infectious mononucleosis	typhoid and paratyphoid fever
influenza	hepatitis A
malaria	yellow fever
meningococcal infection	zika

Each disease is mapped up to 35 symptoms organized into eight medical categories (e.g., systemic, respiratory, neurological). Each symptom is described not only medically, but also in user-friendly language to accommodate non-professionals. This structure (Tab. 2) was designed to facilitate accurate symptom recognition and selection by users within the planned application.

Table 2. Categories of infectious disease signs.

General/ Systemic signs:	continuous fever or fever with intervals less than 1 day intermittent fever every 2-4 days lethargy sweating and/or chills head pain lack of appetite and/or weight loss	Hematological symptoms:	bleeding manifestations
Respiratory signs:	chest pain cough phlegm shortness of breath sore throat runny nose	Gastric symptoms:	abdominal pain diarrhea nausea vomiting
Musculoskeletal signs:	back pain joint pain muscle pain lockjaw	Dermatological or associated signs:	neck swelling skin rash yellow skin and/or dark urine
Neurological signs:	blurry vision cognitive difficulties difficulty swallowing dizziness emotional agitation neurological problems with sensation and movement seizures stiff neck and sensitivity to light	Other signs:	fear of water testicular pain eye redness

The final part of first sets of tables outlines recommended actions for managing each disease, offering guidance tailored to both the patient and the vessel's captain. For patients, the instructions include measures such as isolation, maintaining personal hygiene, and using symptomatic treatments like calamine lotion and proper hydration. For captains, the guidance covers the logistics of isolating affected individuals, communicating with the crew, and implementing emergency response protocols. These recommendations are practical and context-specific, reflecting the unique operational challenges of maritime environments.

The second part of the data base serves as training input for predictive algorithms, representing the likelihood of specific symptoms occurring with particular diseases in percentage form. These numbers were derived from a combination of medical literature, documented case studies, and internal expertise based on previously constructed repository. The percentages represent aggregate data collected around the world in many medical studies. To be able to use such knowledge in training process for machine learning models the information they held had to be represented in different way.

For each disease hundreds of artificial patient profiles were generated, each presenting a unique combination of symptoms while collectively preserving the exact statistical distributions from the knowledge base. For example, if 25% of individuals with a specific disease exhibit Symptom 1, 50% display Symptom 2, and 100% show Symptom 3, we were able to construct five synthetic patients whose combined symptom patterns reflect these proportions. This method of randomly simulating artificial patients allows to capture the natural variability in symptom presentation among individuals. Since no two people react identically to the same infection, this approach ensures a realistic and diverse dataset for training the prediction algorithm.

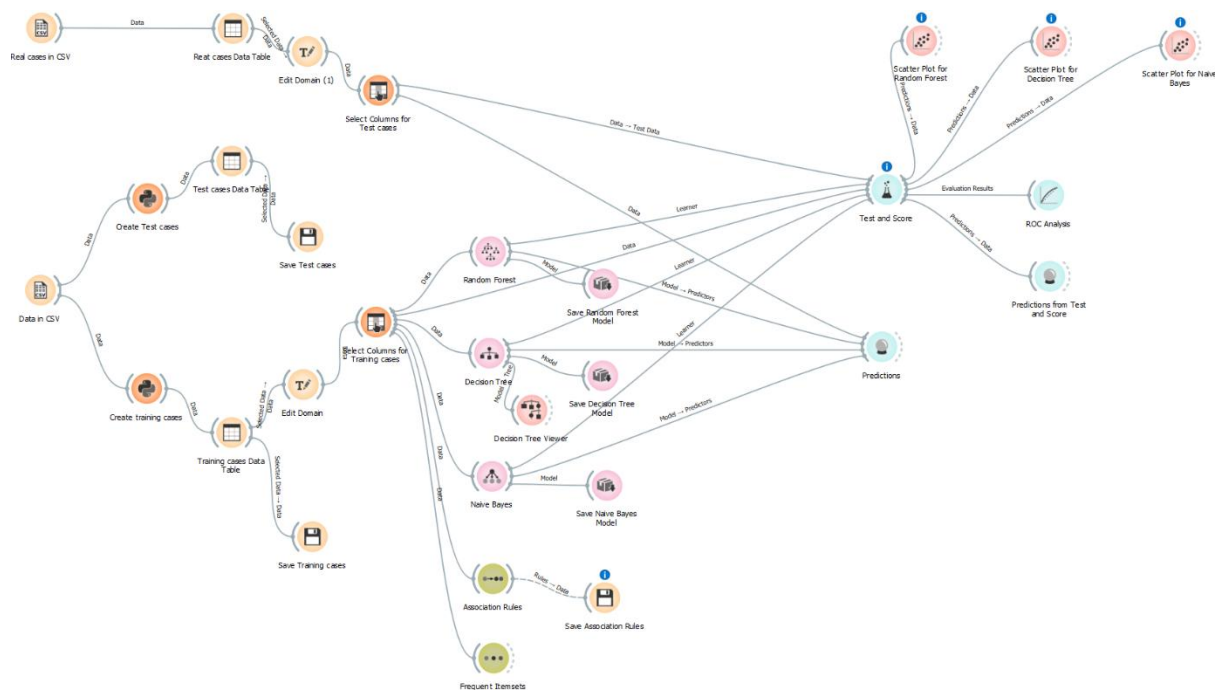


Figure 2. Visual representation of the implemented three models in DESSEV app in Orange software. (Orange screen shot - own elaboration)

4 PREDICTIVE MODEL: DESSEV APPLICATION TESTING RESULTS

All three predictive models were developed using Orange, an open-source data mining and machine learning software. One of its key strengths lies in its flexibility. Models created within the Orange environment can be easily exported and integrated into external applications through the Orange Python library. This capability allows for seamless deployment of predictive models in various systems, including clinical decision-support tools. The solution offers real-time visualization of data processing and results, which helps in understanding model performance and improving transparency in decision-making.

The images below (Fig. 1) present visual representation of the system created using Orange data mining software.

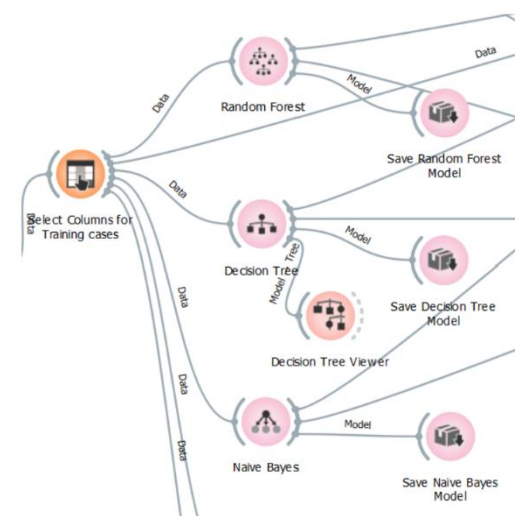


Figure 3. Detailed view of the model, highlighting the implementation of the three algorithms applied in the project.

To test the chosen models, we compiled a small set of real-world cases, with each entry representing a unique patient's combination of symptoms and confirmed disease diagnosis. It is important to note that this test dataset was not included in the training process, ensuring an unbiased evaluation of each model's predictive performance.

To comprehensively assess model accuracy, the project employed a combination of Cross Validation and Random Sampling techniques. This dual approach enabled a robust evaluation framework, offering insights into how well each model generalizes across different subsets of data. Boundary conditions for cross-validation and random sampling are presented in Table 3.

Table 3. Boundary conditions for cross-validation and random sampling.

Cross validation		Random sampling	
number of folds	5	train/ test repeats	10 iterations
stratified sampling	enabled (ensures each fold maintains class distribution for better reliability)	training set size	66 %
		stratified sampling	enabled (preserves class distribution for consistent performance)

Evaluation metrics are quantitative measures used to assess the performance of artificial intelligence (AI) and machine learning models. They provide objective criteria to determine how well a model is making predictions, helping developers and researchers understand its strengths and weaknesses. These metrics are essential in guiding model selection, fine-tuning, and deployment decisions. Common evaluation metrics include accuracy, precision, recall, F1-score, and ROC-AUC, each highlighting different aspects of model performance such as correctness, sensitivity to positive cases, or ability to balance false positives and negatives.

The evaluation metrics used to assess the accuracy of the implemented models are:

- AUC (Area Under Curve): Measures the model’s ability to differentiate between classes.
- CA (Classification Accuracy): The ratio of correctly predicted instances to the total instances.
- F1 Score: The harmonic mean of precision and recall.
- Precision: The ratio of correctly predicted positive observations to the total predicted positives.
- Recall: The ratio of correctly predicted positive observations to all observations in the actual class.
- MCC (Matthews Correlation Coefficient): A measure of the quality of binary classifications.

Table 4 below provides a summary of the evaluation results for the three machine learning models implemented in DESSEV app: Random Forest, Naive Bayes, and Decision Tree.

Table 4. The summary of Decision tree, naive Bayes and Random Forest model performance.

Metric	RF	NB	DT
AUC	1.000	0.998	0.802
CA	0.952	0.857	0.571
F1	0.937	0.825	0.500
Precision	0.929	0.810	0.468
Recall	0.952	0.857	0.571
MCC	0.952	0.854	0.559

The evaluation results clearly demonstrate that the RF model outperforms both NB and DT across all key metrics. With an AUC of 1.000, high classification accuracy (CA) of 0.952, and strong F1-score, precision, recall, and MCC values, Random Forest proves to be the most reliable and robust model. Naive Bayes performs moderately well, while the Decision Tree model shows significantly lower effectiveness in all areas, confirming Random Forest as the optimal choice for disease prediction in the DESSEV project. These results align with broader medical AI trends, where ensemble methods like Random Forest tend to offer superior generalization and resistance to overfitting.

5 SYSTEM-LEVEL EVALUATION

The findings led to the implementation of the RF model in the DESSEV web and mobile applications. A multi-country pilot conducted between January and March 2024 involved 401 maritime professionals from eight nations who ran simulated outbreak scenarios with the DESSEV web/mobile app and answered a 12-item questionnaire. User-interface quality scored highest: > 85 % of respondents agreed that “the app is user-friendly and easy to use”, while 81 % stated that the symptom-input workflow was “quick enough for bridge operations”. Interpretability of the output probabilities mirrored earlier internal tests: 78 % judged the diagnostic report “easy to understand”, and 72 % indicated that they “trust the accuracy of the app’s predictions”. These numbers are congruent with the offline validation in Section 4. With respect to operational value, four out of five users agreed that DESSEV “improves the efficiency and speed of decision-making in response to epidemic alerts”, and a similar proportion reported that the system “enhances overall safety on board and ashore”. The net-promoter item (“How likely are you to recommend the app to a

colleague?”) returned a mean of 8.1 / 10, indicating a favourable adoption climate.

Question 3: The app provides accurate information

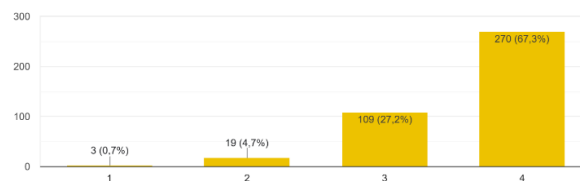


Figure 3 Presentation of user’s evaluation answers on app’s information accuracy.

Free-text feedback echoed the quantitative scores while highlighting two recurrent themes. First, officers requested concise explanations of how individual symptoms influence the model’s vote, confirming the need for transparent machine learning; second, respondents asked for broader language coverage and minor symptom-taxonomy refinements to improve edge-case reliability. Together, these findings validate the choice of an interpretable ensemble as DESSEV’s diagnostic engine and point to the next development steps—embedding lightweight SHAP-style feature-importance visualisations and enabling privacy-preserving federated fine-tuning—so that the system can keep pace with evolving clinical requirements while retaining the usability and trust demonstrated in the field.

6 CONCLUSION AND FUTURE DEVELOPMENT

The DESSEV decision support system is now accessible via a web-based application (<https://dessevproject.eu/app/>) and as a standalone application ready for download from Google Store, available in multiple languages, including Polish, Spanish, Slovenian, Finnish, Greek, and Swedish. The application is optimized for use on mobile devices, given their prevalence on ships, and features a streamlined interface where users can input a minimum of four symptoms from eight categories. The app processes the inputs through the trained model and returns the most likely diagnosis along with practical recommendations for response. Users can also choose to send the results to a designated email address—e.g., ship captain, medical officer, or port authority—and add contextual medical notes such as symptom intensity or pre-existing conditions. The application does not store any personal data, adhering strictly to GDPR standards and ensuring user trust and privacy.

Looking forward, the DESSEV project outlines several possibilities for expansion. These include the integration of real-time health monitoring tools, automated update mechanisms for the knowledge base in response to emerging pathogens and embedding the application within broader maritime safety management systems. The project team also acknowledges the importance of education and training, suggesting that DESSEV could be used in simulation-based drills for epidemic response on vessels. Additionally, feedback loops from users will inform iterative improvements to the tool’s interface and prediction accuracy. DESSEV demonstrates how interdisciplinary collaboration—combining medicine,

data science, maritime policy, and user-centred design—can deliver scalable solutions to complex global challenges. While not a replacement for professional medical advice, the system empowers maritime personnel to take early, informed steps in managing health risks at sea.

ACKNOWLEDGEMENT

This article was created in cooperation with the partners of the DESSEV project implemented by the Maritime University of Szczecin as part of the Erasmus+ program. Financed by the European Union. Contract number: 2022-1-PL01-KA220-VET-000087987. For more information please visit: www.dessevproject.eu

REFERENCES

- [1] European Manual for Hygiene Standards and Communicable Disease Surveillance on Passenger Ships, 2nd ed., 2016. [Online]. Available: <https://www.shipsan.eu>
- [2] EU HEALTHY GATEWAYS Joint Action. [Online]. Available: <https://www.healthygateways.eu>
- [3] Public Health Service Act, 42 U.S.C. § 264, Vessel Sanitation Program, Quarantine and Inspection Regulations to Control Communicable Diseases.
- [4] F. Amenta, M. Di Canio, A. Arcese, F. Bajani, C. Ruocco, and F. Sibilio, "Advanced Telemedicine Solutions for High-Quality Medical Assistance at Sea," *Med. Sci. Forum*, vol. 13, p. 9, 2022. [Online]. Available: <https://doi.org/10.3390/msf2022013009>
- [5] G. Triantafyllou, P. G. Kalozoumis, E. Cholopoulou, and D. K. Iakovidis, "Disease Spread Control in Cruise Ships: Monitoring, Simulation, and Decision Making," in *The Blue Book*, S. T. Rassia, Ed. Cham: Springer, 2024. [Online]. Available: https://doi.org/10.1007/978-3-031-48831-3_8
- [6] X. Liu and Y.-C. Chang, "An emergency responding mechanism for cruise epidemic prevention—taking COVID-19 as an example," *Marine Policy*, vol. 119, 2020. [Online]. Available: <https://doi.org/10.1016/j.marpol.2020.104093>
- [7] I. Bodus-Olkowska, I. Garczyńska, A. Lisaj, M. Mąka, M. Dramski, A. Chronopoulos, T. Gregorič, H. Koivisto, G. de Melo Rodríguez, R. Ziarati, and K. Filipiak, "Repository of Data on Epidemic Situations on Sea Vessels," *TransNav, the Int. J. on Marine Navigation and Safety of Sea Transportation*, vol. 18, no. 1, pp. 133–137, 2024, doi: 10.12716/1001.18.01.12.
- [8] N. A. Aziz, A. Manzoor, M. D. M. Qureshi, M. A. Qureshi, and W. Rashwan, "Explainable AI in Healthcare: Systematic Review of Clinical Decision Support Systems," *medRxiv*, 2024. [Online]. Available: <https://doi.org/10.1101/2024.08.10.24311735>
- [9] N. A. Nasarudin, F. Al Jasmi, R. O. Sinnott, et al., "A review of deep learning models and online healthcare databases for electronic health records and their use for health prediction," *Artif. Intell. Rev.*, vol. 57, p. 249, 2024. [Online]. Available: <https://doi.org/10.1007/s10462-024-10876-2>
- [10] K. A. Abdullah, S. Marziali, M. Nanaa, et al., "Deep learning-based breast cancer diagnosis in breast MRI: systematic review and meta-analysis," *Eur. Radiol.*, 2025. [Online]. Available: <https://doi.org/10.1007/s00330-025-11406-6>
- [11] M. L. Srijia, CH. S. Teja, P. Y. Deep, T. Nandhini, CH. V. S. Narayana, and N. V. M. K. Raja, "Prediction of Disease Based on Symptoms Using Random Forest Classifier," 2023. [Online]. Available: <https://doi.org/10.22214/ijraset.2023.57309>
- [12] V. Jackins, S. Vimal, M. Kaliappan, et al., "AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes," *J. Supercomput.*, vol. 77, pp. 5198–5219, 2021. [Online]. Available: <https://doi.org/10.1007/s11227-020-03481-x>