Fever Classification Using Machine Learning

By
Rima Akter
ID: 213-15-4513

Nurjahan Mim
ID: 213-15-4519

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements forthe **Degree of Bachelor of Science in Computer Science and Engineering**

Supervised by

Dr. Naznin Sultana
Associate Professor
Department of Computer Science and EngineeringDaffodil International University

Co-Supervised by

Dr. Md. Fokhray Hossain
Professor
Department of Computer Science and EngineeringDaffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

September 16, 2025

APPROVAL

This Project titled **"Fever Classification Using Machine Learning"**, submitted by **Rima Akter (ID: 213-15-4513)** and **Nurjahan Mim (ID: 213-15-4519)**, to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **16 September, 2025**.

BOARD OF EXAMINERS

Dr. Bimal Chandra Das (BCD) — **Chairman**
Professor and Dean (in-charge)
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Most. Hasna Hena (HH) — **Internal Examiner 1**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Mr. Mayen Uddin Mojumdar (MUM) — **Internal Examiner 2**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Nazibur Rahman — **External Examiner**
Technical Lead - Database Administrator
Telenor - Grameen Phone Account

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Dr. Naznin Sultana, Associate Professor,** Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

D. Hore

Dr. Naznin Sultana
Associate Professor
Department of Computer Science and EngineeringDaffodil International University

Co-Supervised by:

Dr. Md. Fokhray Hossain
Professor
Department of Computer Science and EngineeringDaffodil International University

Submitted by:

Rima

Rima Akter
Student ID: 213-15-4513
Department of Computer Science and EngineeringDaffodil International University

Nurjahan

Nurjahan Mim
Student ID: 213-15-4519
Department of Computer Science and Engineering, Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project (FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Dr. Naznin Sultana**, **Associate Professor**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Machine Learning** to carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Fever is a common clinical sign in the field of medicine, ranging from mild viral illness to life-threatening diseases like pneumonia, dengue, typhoid etc. Exact identification of pathogenic cause is essential for the immediate and convenient treatment of a patient which could prevent unnecessary prescriptions and would be valuable for prompt recovery of patient as the conventional clinical diagnosis is difficult due to identical symptoms and time taken for laboratory confirmation. To tackle this complexity, this research uses machine learning algorithms to classify fever as pneumonia, dengue, viral fever, typhoid or normal group using as input vectors hematological values (sampled from the prescriptions of the real life patients). The relevant indications were sex, age, HGB (%), WBC, neutrophil, lymphocyte, PLT, and other data. The supervised learning techniques used were: Random Forest, Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Decision Tree and K-Nearest Neighbors (KNN). Among all, Random Forest has the highest accuracy (93.51%) Random Forest (93.51%) > Logistic Regression (91.96%) > KNeighboursClassification(91.79%) > Naive Bayes (91.61%). The satisfactory performance on multiple models demonstrates the possibility and importance of predicting fever based on the hematologic data using a fast and practical machine learning model constructed on the computationally predicted dataset. Conclusion These findings suggest that the machine learning platform might offer diagnostic support, particularly in resource-limited healthcare environments, in which time lost in the diagnostic routine could hinder access to effective treatment (as extensively reported in the context of the COVID-19 pandemic). The use of such predictive models in the clinic could ease diagnostic uncertainty, help rationalize resource utilisation and facilitate an evidence-based approach to the management of individuals with suspected TB. Last but not the least, this study not only confirmed that enrichment of fever typing by machine learning is beneficial, but provides a new sight for practical doctors to try such closed clinical systems in their real world work.

Keywords: Machine learning, feature selection, fever classification, WBC, platelet count, dengue, pneumonia.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This section includes the problem formulation, motivation, objectives, methodology, the expected results, as well as the report outline. It speaks to the growing importance of febrile classification in clinical practice, the potential application of machine learning in diagnostic strategies and the need for evidence-based diagnostics. These latter two are combined in the current contribution.

1.1 Introduction

Desire to maintain health is one of the reinstatement factors of human life, and therefore diseases must be diagnosed before they become serious or cause complications such that they are life-threatening. Fever is one of the symptoms of a majority of viral and bacterial diseases. It is, however, non-specific such that one can not weigh it alone as a sign in differentiating various diseases. This is because different diseases present with fever as one of the major symptoms. For example, fever is common in dengue, typhoid, viral fever, and pneumonia. There are severe consequences of wrong diagnosis of the cause as it can lead to late treatment and lead to a high risk to health and expensive on patients and their families. What is common in most health centers, more so in the developing countries; patients have only a few diagnostic tools. It is the lab that is always an exemption but not a right, and that is why doctors are required to make instant clinical judgment using hematological parameters and the patient background only. Such a diagnosis gap becomes an opening gap of opportunity for computational systems that can crunch this data alongside suplementary ones to get probabilistic predictions that can confidently be used in making medical decisions. In this project, real patient data was collected through prescription which was a realistic, efficient dataset that reflects real-world data configuration based on actual clinical data. Patients forgot to test their Hemoglobin (Hb), White Blood Cell (WBC) count, Neutrophils, Lymphocytes, Red blood cell (RBC) counts, Packed Cell Volume (PCV), MCH, Manna Corporeal Hemoglobin Concentration (mchc), and platelet counts were measured to show the verified diagnosis and had fever characteristics. This dataset was utilized to construct machine learning models from which comparisons were done. These were Random Forest, Support Vector Machine (SVC), K-Nearest Neighbours, Naive Bayes, Decision tree, and Random Forest which was the best having 93.51% accuracy. The project presents a machine learning-based model whose objective is to classify fever and determine the one infected quickly and accurately in a clinical area. In this case, the system will help the doctor to make sound information to administer treatment in a prompt manner based on data finding. Since it holds to be interdisciplinary, a combination of computer science, engineering, and medical knowledge the purpose of this project acts as a crucial mile to the development of e-health. Additionally, the paper exemplifies a chance to lower the identification time with artificial intelligence and make the patient outcome high and affordable to many people in a resource-limited situation. Lastly, in the overall purpose, this is a vision of a greater smart, affordable, and favorable health care system in the future.

1.2 Motivation

The research motivation comes from a growing requirement for intelligent healthcare solutions that can diagnose quickly and accurately given limited resources. According to ces-health, throughout the world, especially in developing countries, (Africa, Asia and Latin America) medical institutions are generally under-equipped in terms of diagnostic equipment and trained medical staff. Because laboratory test results for patients with fever may not come back for days, critical therapy may be delayed. The growing disparity in terms of such physician-to-patient knowledge can be addressed by a machine learning based decision support system that takes advantage of readily available hematological data to automatically yield diagnostic recommendations in a timely manner.

From a computational perspective, fever classification is a difficult yet interesting machine learning task. Medical data are typically heterogeneous, noise, imbalanced and they demand wide generalization algorithms. With this study, we construct state of art models for classification, investigate their generalization and study which model can be used in practice. It is also driven by the impact this work could have: being able to classify fevers correctly will not only save lives, but lighten the load for health care workers, as well as the financial burden of patients for diagnostic tests.

This study is a demonstration of machine learning and data science procedures applied in practice to achieve a socially significant issue. It provides a chance to participate in such a fundamental aspect of data science as data preprocessing, feature extraction, model selection, and performance evaluation. In addition to the technical aspect, the project has benefits to the development of academics and the objectives of the wider society to develop technologies that will have a positive influence on society. Based on this, the rationale of the study is twofold, computational, to utilize the potential of machine learning to conduct an effective analysis and prediction; and humanitarian, to solve an acute social health issue with the help of data-driven decisions.

1.3 Objectives

The main purpose of this study is to demonstrate the usefulness of a machine learning model in classifying the fever types from actual clinical data. To do so, real prescriptions were gathered and transformed into a structured-table that contained hematological data, copies of the specification, hematological parameters with normal values for hemoglobin and white b-cells, percentage of neutrophils, percentage of the lymphocytes, platelet count, and some extra diagnosis markers. These features were input to machine learning models with the verified fever type as the output label. This project aimed to find out best possible computational methods which can classify diseases such as pneumonia, dengue, viral fever, typhoid and normal fever etc using the similarity of diseases clinical features. Another objective was to make the most of a lot of supervised techniques and compare them. We tried hands-on Trashnet data with the tried-and-tested algorithm models such as Random Forest, Logistic Regression, Naive Bayes, SVM, Decision Tree, and the KNN, in order to decide the model which suit the data best. These articles have been evaluated according to accuracy, precision, recall and F1.score, facilitating a very comprehensive comparison. The other was to identify a model that offered the best trade-off between robustness and accuracy. Besides trying to fill the gap between healthcare/medical and computational intelligence, this study also planned to establish a possible practical implementation in the real world. The study has also demonstrated the potentiality of the machine learning in the assistance of doctor because it is observed twisted that in Random Forest there is 93.51% efficiency and which is higher than the other. Ultimately, we want to create scalable diagnostics that could help health care workers make treatment decisions faster and more accurately.

1.4 Methodology

The design of the research projected begun with actual patient data originating from prescriptions and laboratory reports, thus provided clinic reality in this data set. Demographic and hemogram data were recorded (age of the patients, sex, hemoglobin, total WBC count, percentage of neutrophils and lymphocytes, RBC count, platelet count, etc.). (i.e., the type of fever was manually labeled as the ground truth label of training the model). The result set was then preprocessed. – Missing values were imputed, irrelevant attributes were filtered out, and normalizing techniques were used to normalize all numeric variables to the same scale. During exploratory data analysis, univariable relationships with blood parameters and fever categories were assessed and helped prioritize the most actionable markers. Six supervised learning algorithms were applied to the preprocessed dataset to get the trained models: Random Forest, Logistic Regression, Naive Bayes, Support Vector Machine, Decision Tree, and K-Nearest Neighbors. The dataset was divided into a training set and a test set and the hyperparameters of each model were tuned to maximize R Twitter and, therefore, predictive power. The evaluation/score metrics: accuracy, precision, recall and F1 Score. However, compared to the tests for all models, the relatively best results were obtained for Random Forest whose performance measure of accuracy rate was 93.51% was the best (and much more than the ones of Logistic Regression and Naive Bayes 92%). This approach made it possible to describe the process of model construction and to understand the prospect of using computational intelligence techniques for diagnosis in practical health care data.

1.5 Project Outcome

This trial shows that it is technically feasible and clinically useful. Methodologically this work produced a methodological step in the classification of fevers and their signature by machine learning on haematological collection data. These results indicate that machine learning approaches can obtain good performance on this task, and among the methods, RF achieved the best performance (accuracy = 93.51%) compared with LR (92.64%) and with NB (92.21%). And not only deep learning but also moderate model like DTs did few good scoring on the other hand even being very simple KNN did the scoring of around 87.45 and 84.85 percent respectively. These results indicate that a complex model trained on cupled, raw data such a pure input deep Artificial Neural Network constructed through deep learning, can be used as an effective aid for medical classification. This means in practical terms that the results of the project show that a machine learning model can act like artificial intelligent decision support tool for the health care worker. Through entering a few systemic hematological data, they are able to diagnose the type of fever rapidly to avoid the mistake of diagnosis and wrong decision of treatment. This approach may be more applicable, particularly in low resources health care facility where more advanced diagnostics may not be readily available or accessible in a timely manner..These findings are demonstration that an AA system can be introduced in the clinical decision-making processes even in small hospital facilities that have blood work capability. Additionally, the study purports to extend the performance of the state-of-the-art by comparing different machine learning models on real clinical data. This points the way for further work to generalize the DL techniques or the ensemble methods for better accuracy. Future The project could develop into a real healthcare app. Practically, a system which is mobile/web based and simple to use by the user would be built (and therefore serves as the UI) where the user would input the patient parameters and real time fever classification is returned. The hospital data base can be used as a knowledge base for the system knowledge database that causes the system to continuously learn from new cases, and keep adapting in time to provide the most accurate treatment. The resulting diagnostics will ultimately lead to smarter, universal, and cost effective diagnostic tools that will significantly improve global health care.

1.6 Organization of the Report

The format of this report is designed to convey an overview of all the phases of the project in the order in which the project was completed for easy understanding and clear progression of the project. Exploratory phase: identifying the health problem, outlining the importance of fever classification, and setting goals The project initiated in the exploratory phase, in which the health problem, the importance of fever classification and the objectives were determined. This was then followed by a data collection phase, in which prescriptions for real patients were gathered to assemble a comprehensive dataset that fairly represented clinical reality. Preprocessing Thus, the preprocessing phase made the data consistent and accurate, so it became ready for the experimentation with machine learning. During development several machine learning models were tried and tested - Random Forest, Logistic Regression, Naive Bayes, SVM, Decision Tree and KNN. The Comparative analysis was of research compare of the Random Forests that seven algorithms adopted, and the Random Forests method performs the best 93.51% accuracy among them. This was an important step as it showed the approach to be possible and the method to be valid. The results section of the outcomes phase synthesis these results, and also considered the broader implications of adoption of machine learning for healthcare diagnostics. The documentation phase turned the technical work into a well-organized academic report that according with FYDP format, with the thesis and its chapters. In chapter one, the project is introduced and its motivations and objectives are explained. Chapter 2 presents the background and review of literature. Chapter 3 presents the research methodology and system design. Implementation, results and evaluation are dealt with in chapter 4. In Chapter 5, standards of engineering, ethical aspects, and sustainability issues are considered. The final Chapter 6 provides the conclusion of the report and discusses the limitations and future work. In this phased organization of the work, the report succeeds not only in recording the technical progress of the project, but also in outlining a clear road map of how the project moved from an initial idea to a set of documented results of value for future use.

Chapter 2
Background

This chapter presents the required theoretical basis and the survey of the associated literature. It starts by giving an introduction to the importance of machine learning and diagnosis of fever in healthcare. It then examines the appropriate literature, points out similar uses in medical and computational applications, discusses the gaps in the previous studies and finally gives a conclusion indicating how this background relates to the subject of the present study.

2.1 Introduction

Fever is among the most common clinical presentation in medical care but it is also among the most diagnostic difficult conditions. Common conditions that have overlapping symptoms include high body temperature, fatigue and weakness making differentiation difficult: pneumonia, dengue, viral fever, typhoid and even normal fever. Normal fever in most instances, tends to be mild and self limiting having similar clinical characteristics with more severe infections hence misdiagnosis and postponement of proper treatment [1]. Pneumonia is one of the leading causes of morbidity and death across the globe especially in children and the elderly. They have demonstrated that hematological features, including neutrophil-to-lymphocyte ratios and the counts of white blood cells, can be used as significant discriminatory attributes to detect pneumonia under the computational methods [8]. The mosquito-borne viral infection that carries the dengue is the cause of a significant burden in the world and especially in the tropical and subtropical regions. Dengue has a very short window to treat at an early stage, yet in healthcare environments that are resource-constrained, it is often challenging to obtain laboratory confirmation in a timely manner, and machine learning-based approaches are therefore valuable in risk prediction and classification [2], [3]. Salmonella typhi causes typhoid fever, which remains common in the developing countries where sanitation and hygiene is not good. Machine learning (ML) models like Naive Bayes and Decision Trees have been tested on clinical and hematological data, and it is shown that automated models can be used to diagnose typhoid [13], [14]. Viral fevers, with a broad spectrum of infections most of which are self-limiting in nature, should be properly distinguished with bacterial fevers in order to avoid inappropriate antibiotic use. Support Vector Machines and ensemble classifiers have been studied and applied to viral fever data and their effectiveness in enhancing diagnostic accuracy is confirmed [11], [12]. Lastly, it has been found that normal fever and infectious diseases can be differentiated by using models like Random Forest and ensemble classifiers, which are better diagnostic tools than conventional methods [20]. Complete blood count (CBC), platelet indices and other serological tests are important in classical diagnostic methods. These are necessary but their interpretation is on the knowledge of the physician and are not always practical in the healthcare system that has limited resources. Machine learning approaches, in turn, are able to work with many hematological and demographic variables at once and identify latent patterns that are not perceivable by humans. Random Forest (RF) and Logistic Regression (LR), Support Vector Machines (SVM), Naive Bayes (NB), Decision Trees (DT), and K-Nearest Neighbors (KNN) algorithms have always demonstrated positive results in terms of their classification of the disease related to fever [6]. They can be also modified to fit the automated diagnostic systems with these models being scalable and fast with little manual intervention. The current work is based on these results and concentrates on the multi-classification of five types of fever including pneumonia, dengue, viral fever, typhoid, and normal fever based on hematological and demographic factors. Through performance evaluation of widely-used ML models, such as RF, LR, SVM, NB, DT, and KNN, the present work will make a contribution to the creation of a credible, data-driven clinical decision-support system. A system like this can help minimize the time to diagnosis, enhance accuracy of benefit, and help achieve healthcare delivery especially in resource-limited environments.

2.2 Literature Review

Identify fever, nature of fever causing disease has become one of the most important area of research because of the similar kind of symptoms in disease such as pneumonia, dengue, viral fever, typhoid and normal fever. Traditional diagnosis used to depend on consultation and personal laboratory analysis, both subjective and at times time-consuming and erroneous. Machine learning (ML) has therefore emerged as an efficient tool, which applies to evidence-based data, allowing physicians to simultaneously evaluate several hematological and demographic parameters. Random Forest (RF), Logistic Regression (LR), Naive Bayes (NB), Support Vector Machines (SVM), Decision Trees (DT) and K-Nearest Neighbors (KNN) models have been demonstrated to have high potential in clinical decision support with reports of improved diagnostic accuracy, interpretability and predictive strength. One of the most common diseases that was researched in the literature of the ML-based fever classification is dengue fever. The authors showed that it was possible to conduct dengue classification on the basis of structured blood-test results effectively, where RF yielded a greater accuracy as compared to LR and DT [1]. Sarma et al. came up with the models of dengue prediction using ML and demonstrated that the ensemble machine could be more robust than an individual one (RF) [2]. Going beyond the single-patient classification, Salim et al. forecasted dengue epidemics in Malaysia and proved that the ML techniques could effectively model epidemiological patterns based on both clinical and demographic information [3]. The additional comparative studies of the dengue fever classification showed that the ensemble methods were better than the individual classifiers, and hence importance of RF in achieving credible classification is important [4].In a similar line, it was noted by Sanjudevi and Savitha that RF achieved superior performance in dengue prediction, with respect to NB and DT [5], while Gambhir et al. observed that RF performed better than SVM and NB in dengue detection by hematological data as well [6]. Recent analyses including Chaw et al. used integrated ML models (For shock-risk evaluation of dengue patients and demonstrated that RF and SVM could also contribute to clinical prediction rather than categorization [7]). All together both studies have confirmed the suitability of ensemble techniques in the dengue research literature and comprise RF as the most reasonable model in this literature. Pneumonia, which is among other major fever causing diseases has also been addressed by ML-based researches. Stokes et al. applied ML in the low-resource setting to discrimate pneumonia from bronchitis and demonstrated that RF and SVM performed well with high prediction where advanced imaging not applicable [8]. In Wang et al.'s study, the index of pulmonary inflammation was applied to classify viral-induced pneumonia and the RF and gradient boosting were performed well [9]. On a parallel note, Chang et al. implemented ML models on pediatric patients with pneumonia and established that hematological ratios like neutrophil-lymphocyte count were the very discriminative features to be classified [10]. Peng et al. generalized this method with Mycoplasma pneumoniae pneumonia, and discovered that SVM and LR gave a high level of performance, when they were trained on regular blood variables [19]. The above findings indicate that pneumonia classification can be achieved with inputs based on hematology and ensemble models and the SVM often show improved results. Typical studies on the category of typhoid fever have also shown that ML techniques have the potential to assist clinical decision-making in an improved manner. Oguntimilehin et al. created one of the first ML models to predict typhoid fever and demonstrated that DT and NB could produce interpretable and reliable classification performances [13]. Andrianto et al. followed this work by installing a clinical decision support system on typhoid fever in which NB and DT performed classification accuracies of nearly 90 percent [14]. Further on these premises, the works by Oguntimilehin et al. introduced a larger ML-based CDSS that included RF and DT to improve predictive ability and provide treatment suggestions [15]. Recently, Handayani and Hakim analyzed mixed samples of dengue and typhoid hematologic outcomes, where NB was effective to identify the two diseases [16]. These results indicate that blood-based parameters are capable of reliably classifying typhoid fever and NB and DT are interpretable and RF is robust measures. The other focus of ML classification has been viral fevers that might be hard to diagnose because of its symptom overlap with other types of infectious diseases. Fathima and Hundewale equated ML methods in detecting arboviruses and affirmed that SVM was more efficient than NB and other methods [11]. Chaudhary et al. used RF and NB to classify viral fevers and once again ensemble methods performed better than single classifiers [12]. Zhao et al. did not limit their research to classification but to prognosis by predicting the poor outcome in fever-related illnesses using RF and LR, of which RF was more successful [17]. Widiyaningtyas et al. used NB directly to the diagnosis of fever symptoms and obtained an accuracy of 88 percent, proving that they can use simple probabilistic models to assist in clinical decision-making [18]. A combination of these studies supports the significance of RF, SVM and NB in processing viral fever data and accentuates the fact that ensemble algorithms generally give the best results. The other category of normal fever has not been as well researched as other categories, but has been examined as compared to infectious fevers. Islam et al. compared ML models to draw the line between dengue and non-dengue (i.e., normal) fever resulting in RF having the highest classification accuracy [20]. Rana et al. compared several ML models such as RF, LR, and SVM in fever classification and ensured that RF and ensemble models demonstrated better results in all the investigated datasets [21]. These findings indicate that in spite of a high level of symptom overlap, the performance of ML models can still be highly discriminating on normal and infectious types of fever. As can be seen in general, the literature presents a very distinct trend: ensemble models like Random Forest are always the most reliable and accurate at selecting among different types of fevers, whereas Naive Bayes and Decision Trees offer interpretability that can be useful in clinical practice. Logistic Regression and SVM are also good runners up particularly in the realm of structured hematological data. While most of the works focused on binary classification (e.g., dengue vs. non-dengue or pneumonia vs. non-pneumonia), the need of multi-class models that could be used to diagnose all 5 types of fevers at the same time started to be acknowledged. A more important challenge is the trade-off between predictive accuracy and transparency, which is essential in the application of ML models in the clinic and would need to be addressed in further studies.

Table 2.1: Summary of Literature Reviewed

Author (s)	Year	Title	Methodology	Key Findings
Sajana et al. [1]	2018	Classification of dengue using ML techniques	RF, LR, DT	RF performed best in dengue classification using clinical data
Sarma et al. [2]	2020	Dengue prediction using ML	RF, DT, SVM	Ensemble methods were more robust in prediction
Salim et al. [3]	2021	Dengue outbreak prediction (Malaysia)	RF, LR	ML-based forecasting effective in outbreak monitoring
Dourjoy et al. [4]	2020	Dengue fever prediction	ML comparison study	Ensemble models improved dengue classification
Sanjudevi & Savitha [5]	2019	Dengue fever prediction using classifiers	NB, DT, RF	RF had highest accuracy
Gambhir et al. [6]	2018	Dengue diagnosis via ML	NB, SVM, RF.	RF superior to NB and SVM
Chaw et al. [7]	2024	Dengue severity risk	RF, SVM	ML predicted risk of shock development
Stokes et al. [8]	2021	Pneumonia & bronchitis diagnosis	SVM, RF	ML effective in low-resource diagnostic support
Wang et al. [9]	2024	Viral pneumonia prediction	RF, Gradient Boosting	Pulmonary indices effective predictors
Chang et al. [10]	2023	Pediatric pneumonia diagnosis	SVM, RF	Neutrophil/lymphocyte ratios key discriminators
Fathima & Hundewale [11]	2012	Arbovirus classification	SVM, NB	SVM had best performance

Chaudhary et al. [12]	2023	Viral disease diagnosis with ML	RF, NB	RF superior in viral classification
Oguntimilehin et al. [13]	2013	Typhoid fever diagnosis	NB, DT	Interpretable classifiers with good accuracy
Andrianto et al. [14]	2019	CDSS for typhoid	NB, DT	Achieved 87–89% accuracy
Oguntimilehin et al. [15]	2014	ML-based CDSS for typhoid	RF, DT	Demonstrated clinical support system utility
Handayani & Hakim [16]	2023	Dengue vs typhoid hematology classification	NB	NB achieved strong results on hematology data
Zhao et al. [17]	2020	Fever-related prognosis	RF, LR	RF better than LR for prognosis prediction
Widiyaningtyas et al. [18]	2020	Fever diagnosis via ML	NB	88% accuracy using NB
Peng et al. [19]	2024	Mycoplasma pneumonia classification	SVM, LR	Blood parameters highly predictive
Islam et al. [20]	2021	Dengue vs non-dengue fever	RF, SVM	RF most accurate at distinguishing normal fever
Rana et al. [21]	2022	Dengue fever prediction using ML analytics	RF, LR, SVM	RF outperformed other models

2.2.1 Similar Applications

In recent years, machine learning has played an increasingly important role in the healthcare domain, especially for disease prediction and diagnostic assessment. ML-based models assist in the differentiation between diseases, which frequently possess similar clinical manifestations, by processing patient data including laboratory values, imaging, and symptoms. This is therefore particularly applicable to scenarios such as fever classification when many diseases have similar presentations.

Infectious Disease Prediction
ML algorithms have been attempted to predict diseases such as dengue, malaria with clinical information (e.g, blood test result) and environmental indices. They help identify outbreaks early and enable clinicians to rule out some causes of fever [3].
Pneumonia Diagnosis
Machine learning models are being used to analyze blood tests and chest X ray data to help differentiate viral pneumonia, and other respiratory infections quickly. This is very similar to fever syndromes, when the problem is to separate out overlapping clinical entities [9].
Hematology-Based Disease Classification
Blood factors (including WBC count, platelet count, and hemoglobin levels) have been employed in a number of studies to differentiate bacterial and viral infections. As fever is frequently associated with variations of blood profile, this serves as a direct input to fever classification tasks [12].

2.3 Gap Analysis

Although existing research has shown the inputs and outputs of ML in the context of healthcare, a few key gaps are identified. A lot of studies focus on single-disease prediction, heavily depend on the data of the image, and use the synthetic datasets which do not represent real patient status. In addition, comparative studies for multiple algorithms are scarce and the utilization of widely accessible hematological parameters is currently limited. The current study consider these problems by introducing a multi-class fever classification framework on the basis of real patient blood reports, evaluated under a common set of metrics over six models.

Table 2.2: Gap Analysis of Related Works in Healthcare Machine Learning and the Proposed Fever Classification System

Feature / Focus	Existing Studies	Proposed System
Single-disease classification	Yes – dengue, pneumonia, etc.	No – multi-class fever classification
Use of hematological data	Limited; preference for imaging/complex tests	Yes – multiple blood-test parameters
Dataset authenticity	Often synthetic or incomplete	Real hematology reports (~2000 patients)
Sample size	Many < 500 or small subsets	Larger dataset (~2000 records)
Model variety	Usually one or two models only	Six models compared (Logistic Regression, DT, RF, SVM, NB, KNN)

Comparative analysis of algorithms	Rare	Yes – uniform evaluation across models
Evaluation metrics	Often only accuracy reported	Multiple metrics: Accuracy, Precision, Recall, F1-score
Practical healthcare relevance	Moderate – focused on narrow conditions	High – scalable fever diagnosis support
Multi-class prediction (overlapping symptoms)	Rare – mostly binary tasks	Yes – five fever types classified
Cost-effectiveness & accessibility	Limited by reliance on imaging	High – low-cost hematological parameters

2.4 Summary

Machine learning tools have become useful in health care diagnostics, especially in the diagnosis of diseases with similar symptoms such as fever.The review presented several ML application use cases such as forecasting infectious diseases, diagnosing pneumonia and COVID-19 and mobile symptom checker among others which also underscore the flexibility of computational models in clinical care. And yet, as we've made all these advances, we've been neglecting a fundamental question: How should fever be classified? Inspiration: After reviewing the related work we noticed several significant limitations: (i) most work focuses on prediction for one single disease, (ii) it strongly depends on imaging (no synthesis) or synthetic populations, and (iii) only a few methods are compared. These are the things that make these results less easily applicable, or perhaps not very clinically relevant. To answer these questions, the introduced system presents a multi-class fever classification algorithm using real hematology report data (~2000 patient infection + normal cases).We conduct a rigorous performance evaluation of the framework on six-class and binary classifications tasks, and evaluate it over six machine learning algorithms, namely Logistic Regression, Decision Tree, Random Forest, Support-vector Machine(SVM), Naive Bayes and K-Nearest Neighbors(KNN). Based on these, the investigation not only refines fever diagnosis, but also directly establishes a reliable ground for a practical (round)-based system as a decision- (rule)-supporting aid to support physicians for responsible persons in practice and in clinical practice of day-care hospital in their work.

Chapter 3

Research Methodology

This chapter outlines how the study was done. It outlines the system architecture, the system, and the flow of process of classifying fevers. The suggested method of fever classification is described, both in terms of functional and design requirements and user interface design. An explanation of the proposed approach step-by-step, including its alternative approaches, is presented. Also, a Gantt chart would show the allocation of tasks in the project timeline.

3.1 Methodology

3.1.1 Overview

The Methodology applied to Get Step by Step Progression From Data to Model Evaluation and Machine Learning: Fever categorisation was selected as the main application of the project as it requires good quality clinical information and strong computational effort. The workflow begins with true patient prescription blood data. The demographic and clinical features of the data set (gender, age, hemoglobin, white blood cell (WBC) count, neutrophil ratio, lymphocyte ratio, red blood cell (RBC) count, hematocrit, mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and platelets) are listed. There are five classes of fever such as pneumonia, dengue, viral fever, typhoid and normal fever and each row in the data is classified in to one of these five classes. We performed all the missing values, normalising, and set-up for machine learning models post-collection.Then the study compared sev eral supervised algorithms—Logistic Regression, Random Forest, Naive Bayes,Sup port Vector Machine (SVM),Deci s ion Tree and K-Nearest Neighbors (KNN). Trained model, hyper-parameter tuning and comparisons were conducted based on normal criteria like accuracy, precision, recall and f1-score. Goal in design: Being a reproduced>flowing>evolving research. The system structure is divided into Data Input, Preprocessing, Feature Extraction, Model training, and Prediction Output.It is easy to follow instructions with all steps defined so that you will always obtain consistent and repeatable results. Therefore, the method ensures that not only are the project results of importance, but has the potential officiousness for future potential clinical decision-support system (DSS) use.

3.1.2 Proposed Methodology

The proposed methodology starts with the process of gathering patient hematological data that are processed and cleaned up to eliminate inconsistencies. The data is then divided into training and test subsets and important hematological characteristics are taken to develop a model. Different machine learning algorithms are trained and assessed with the help of performance indicators and their outcomes are compared. The best performing classifier was the Random Forest model with the highest accuracy (93.51%), and thus chosen to be the best predictor of fever types: pneumonia, dengue, viral fever, typhoid and normal fever.

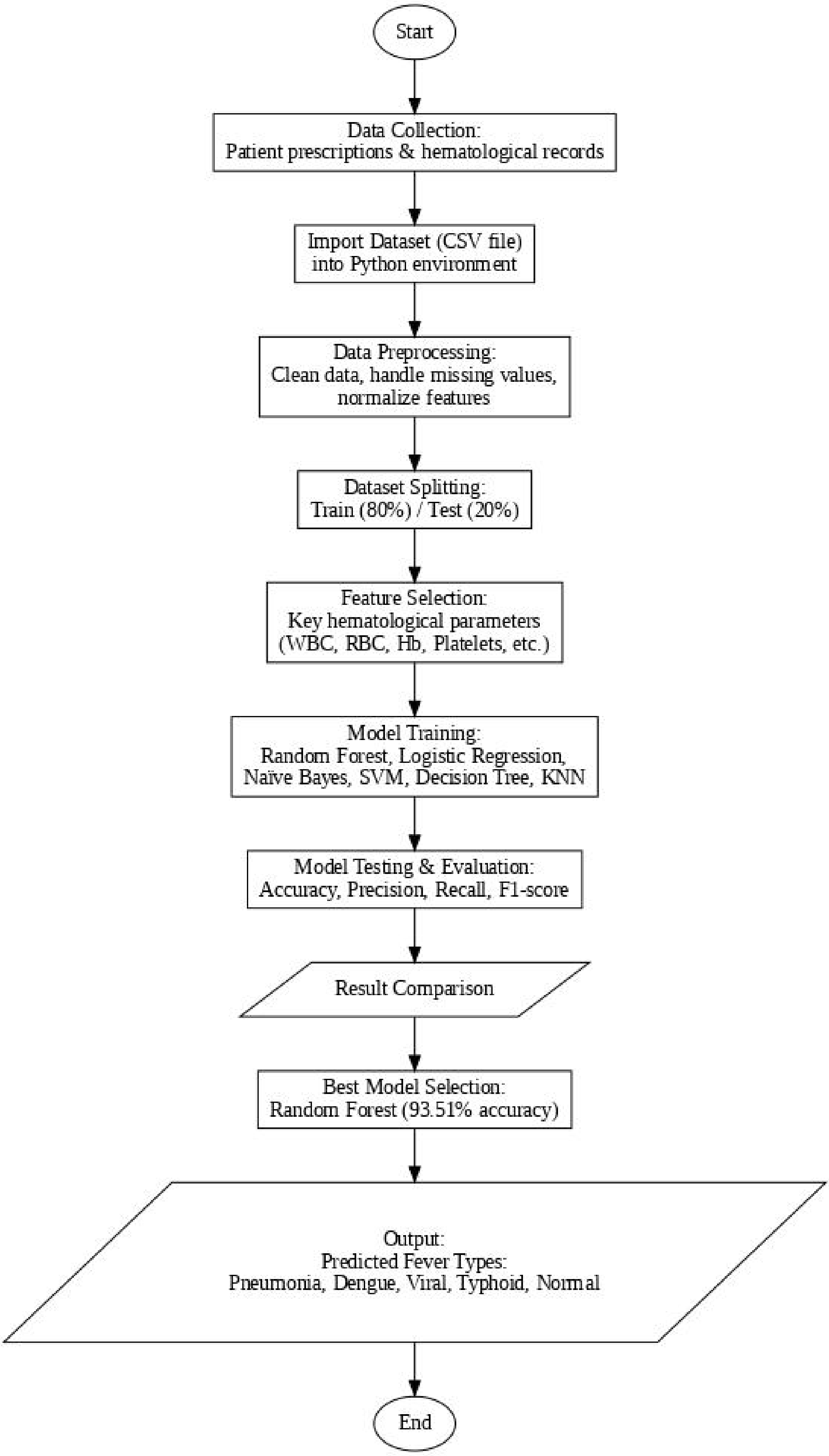

Figure 3.1: Proposed methodology for Fever Classification

3.1.3 Functional and Nonfunctional Requirements

Functional Requirements

Functional Requirements are descriptions of system features that system should fulfil, focusing on the features and functions that the system need to offer to directly implement its role.For the system of category of fever it is –

1. **Data Input Module:**

This module provides functionalities to read raw hematology data. The hematology data was extracted from reports and organized in CSV format for analysis.

D	E	F	G	H	I	J	K	L	M	N
Age	Hemoglobi	ESR (Weste	Total WBC (	Neutrophils	Lymphocyt	Monocytes	Eosinophils	Basophils	Total Cir. Ec	Total RBC C
16	12.42	47.87	15260	73		6.4	0.8	0.8	126.4	5.06
32	11.56	47.98	18253	77.9	11.4	6.4	3.2	0	586.5	4.58
28		67.94	18283	76.9	11.9	4	0	0.7	3.2	4.83
9	12.43	14.02	9074	78.2	3.6	5.4	5	0.6	450.5	4.45
27	12.15	70.43	16654	73.6	15	4.5	0.5	0.3	86.6	4.39

Figure 3.2(a): Hematology dataset (first set of features including Age, Hemoglobin, ESR, WBC count, Neutrophils, Lymphocytes, etc.).

O	P	Q	R	S	T	U	V	W	X	Y
HCT/PCV	MCV	MCH	MCHC	RDW-CV	RDW-SD	Total Platel	MPV	PDW	PCT	Result
49	80.1	29.5	30.5	13.66	40.81	264227	11.81	11.06	0.312	Pneumonia
42.1	89.8	31.4	32.1	12.4	40.95	258412	8.54	12.13	0.221	Pneumonia
38.5	78.9	28.3	35	12.69	36.99	182301	9.21	12.32	0.168	Pneumonia
42.9	88.1	29.9	31.4	13.71	32.18	280044	8.74	11.78	0.245	Pneumonia
36.5	83.2	26.2	32.9	11.19	46.04	251744	9.6	9.36	0.242	Pneumonia

Figure 3.2(b): Hematology dataset (remaining features including HCT/PCV, MCV, MCH, MCHC, RDW, Platelet count, PDW, PCT, and Result label).

2. **Data Preprocessing:**
The pipelines should preprocess the input data, it will deal with missing values by imputing them, it will encode the categorical attributes, and finally it can normalize some of the features to ensure the model is prepared.

3. **Prediction Engine:**
The system shall be able to learn and test different machine learning models (Logistic Regression, Decision Tree, Random Forest, SVM, Naive Bayes and KNN) to classify the types of fever.

4. **Results Generation:**
The predicted type of fever should be printed by the system along with confidence values and quality indicators such as accuracy, precision, recall and F1-profile.

5. **Model Comparison:**
The system has to be capable of showing models side-by-side, in order to find the best model for the clinical practice of the algorithm.

6. **Visualization Module:**
The system should generate understandable visualization (for example, confusion matrix, distribution plot) to help interpretability of the model, to assist the healthcare decision-making process.

Non-Functional Requirements

The Non-Functional Requirements define the way in which the system should work, being more about quality properties instead of tasks. Within the framework of the fever classification system, these constraints guarantee that the models are accurate, fast, secure and scalable, which makes the solution feasible for clinical deployment in practice.

1. **Performance:**
Predictions of the system are expected in on-line mode in real-time - the response time should be no more than several seconds for input handling.

2. **Accuracy and Reliability:**
A stable classification result (classification accuracy >= 90% on the test data) should be achieved across independent runs.

3. **Scalability:**
The method needs to be scalable to a larger population, as it could be easily incorporated in future with hospital databases or with electronic health records.

4. **Usability:**
The system would have a user-friendly interface, simple inputs and outputs, and would be targeted at healthcare professionals who are not IT specialists.

5. **Security and Privacy:**
The solution needs to align the health data privacy policies while ensuring the privacy of patients data privacy and secure utilization of precious health care related information.

6. **Maintainability:**
The architecture should enable updating the model, data or the parameters to convert a system in the long run can be reconfigured with little overhead.

7. **Cost-Effectiveness:**
The response also needs to be dynamic and also amenable for low resources health care facility.

3.1.4 Data Flow Diagram

The flow of data starts with patients submitting hematology reports and clinical data that are gathered and placed into a dataset. The cleaned data is then trained and evaluated using machine learning models after preprocessing. The most trained model is then used to group new patients information in the categories of fever and the findings provided back to the doctors and other medical practitioners in order to facilitate clinical decision-making.

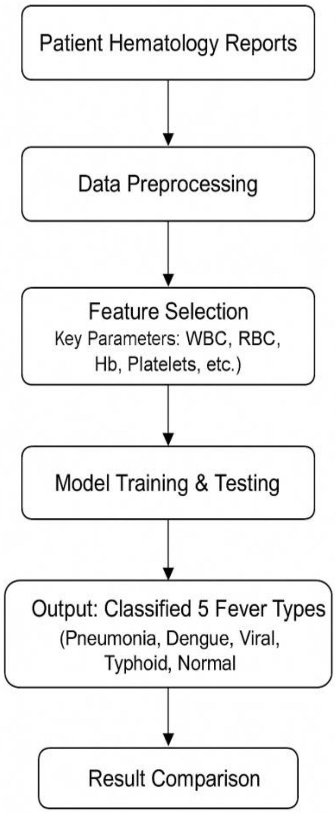

Figure 3.3: Fever Classification Flow Diagram

3.1.5 UI Design

One of the most important elements of the proposed system of classifying fevers is the User Interface (UI). The main objective of the UI is to have a platform that is easy and user-friendly whereby the user can key in hematology values and receive a prediction on the various types of fever. Design is aimed at transparency, usability, and effectiveness to address both the medical and non-technical users.

Key Features

The specified UI has the following main characteristics:

Input Form:

The user can enter hematology values which are considered important in the classification of fever, such as:

Age, Hemoglobin (Hb), ESR (Westergreen), Total WBC Count, Neutrophils, Lymphocytes, Monocytes, Platelet Count, PCT (Procalcitonin), MPV (Mean Platelet Volume) etc.

Prediction Button:

The machine learning model is activated by a centrally located button that is labeled Predict Fever Type and which the machine classifies the fever.

Result Box:

The result box at the bottom of the screen shows the predicted type of fever in a green color with a check-mark icon to confirm this.

Example:
Predicted Fever Type: Dengue

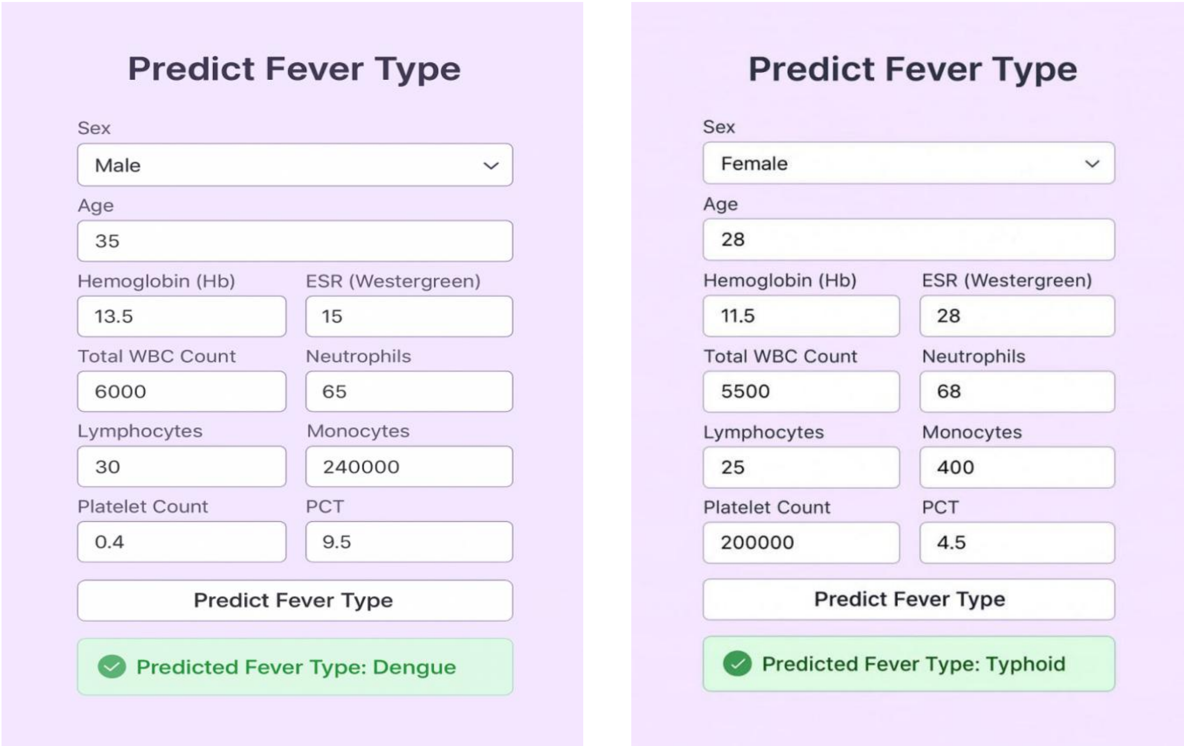

Figure 3.4: User Interface

3.2 Detailed Methodology and Design

The fever classification system underlying pipeline presented uses a structured procedure to exploit blood data, especially (although not exclusively) with regard to input that should be provided via machine learning as feature extraction technique. We have tried a few alternative, less sophisticated methods, be it based on imaging for the diagnostic or using just one algorithm, for complexity and weak generalizability. A multi-standard comparison paper was preferred where the measurements were from a blood-test instead for practical applicability and generalization to hospitalization.

1. **Data Acquisition**
   - Hematological and demographic data obtained from patient prescriptions have established authenticity and clinical relevance to real life.
   - Alternative: Simulator or image-based data were also dismissed as they were not reliable and available in low-resource settings.

2. **Data Preprocessing**
   - Handling missing and inconsistent values.
   - Normalizing continuous features to reduce bias from scale differences.
   - Encoding categorical variables (e.g., sex) for compatibility with machine learning models.
   - This ensures clean, standardized input for fair model comparison.

3. **Feature Selection and Exploration**
   - All hematological parameters were initially considered.

- Exploratory Data Analysis (EDA) highlighted platelet count, WBC levels, and neutrophil–lymphocyte ratio as especially significant.
- Alternative: Automated feature selection (e.g., PCA) was considered but avoided to retain clinical interpretability.

4. **Model Development**
   - Six algorithms were implemented: Random Forest, Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN).
   - Hyper parameter tuning was applied to optimize performance.
   - Stratified train-test split ensured balanced representation of all fever classes.
   - Alternative: Deep learning models were considered but rejected due to limited dataset size and lack of interpretability.

5. **Model Evaluation and Comparison**
   - Performance was assessed using accuracy, precision, recall, F1-score, and confusion matrices.
   - Random Forest achieved the highest accuracy (93.51%), while Logistic Regression and Naive Bayes also performed strongly.
   - Comparative analysis highlighted each model's strengths and limitations.

6. **Final Selection and Scalability**
   - Random Forest was selected as the best-performing model due to superior accuracy and robustness.
   - The methodology is scalable to larger datasets and adaptable to advanced algorithms in future extensions.

3.3 Project Plan

This research was carried out in structured phases, beginning with the identification of the research problem and extending to experimentation, evaluation, and documentation. Each stage was carefully planned with clear timelines, activities, and deliverables to ensure the research achieved its objectives. The phases are summarized below:

Table 3.1: Research Project Plan and Phases

Phase Number	Research Phase	Estimated Weeks	Research Activities	Expected Outcomes
1	Problem Definition & Literature Review	3 weeks	Defined research problem, reviewed prior studies on pneumonia, dengue, viral fever, typhoid, and normal fever classification.	Problem statement, literature review notes.

2	Data Collection & Preprocessing	4 weeks	Collected hematological data from prescriptions, handled missing values, cleaned and normalized dataset.	Structured and preprocessed dataset ready for analysis.
3	Exploratory Analysis & Feature Study	3 weeks	Performed EDA, studied correlations, identified predictive features such as platelet count and WBC ratio.	Analytical findings, selected feature set.
4	Model Development & Experimentation	4 weeks	Implemented six ML algorithms (Random Forest, Logistic Regression, Naive Bayes, SVM, Decision Tree, KNN); tuned parameters.	Trained ML models for fever classification.
5	Model Evaluation & Results Comparison	4 weeks	Evaluated models using accuracy, precision, recall, F1-score; compared results; Random Forest achieved 93.51%.	Accuracy table, confusion matrices, best model identified.
6	Documentation & Thesis Writing	3 weeks	Drafted thesis chapters, compiled methodology, literature review, results, and discussion.	Draft of thesis document.
7	Final Review & Defense Preparation	1 weeks	Revised report, polished writing, prepared defense slides, and rehearsed presentation.	Final thesis submission and defense readiness.

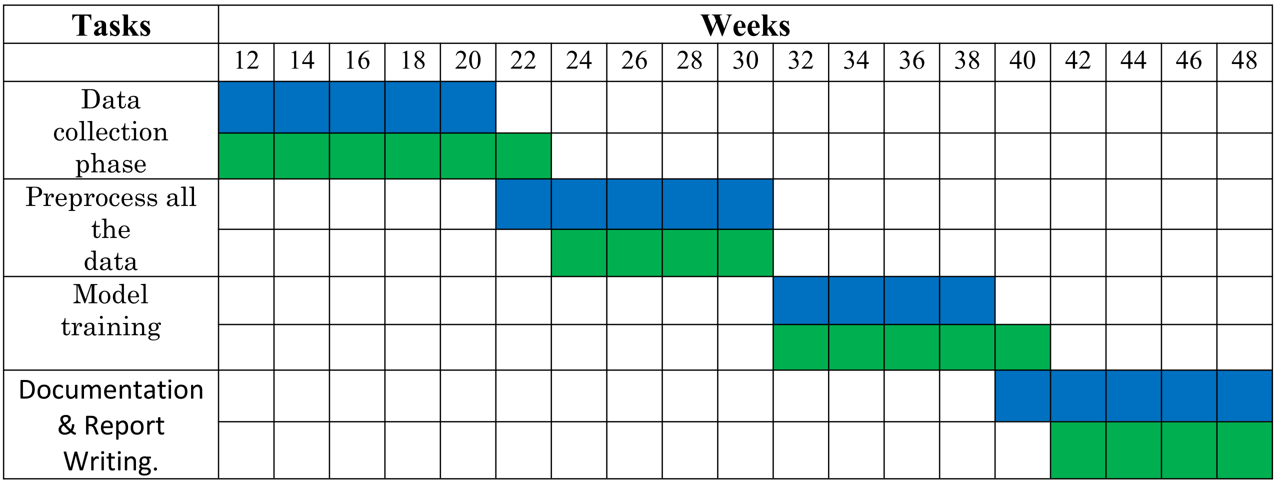

Figure 3.5: Gantt Chart Representation of the Research Plan

3.4 Task Allocation

Task allocation for this research project was carried out in alignment with the defined phases of the project plan. The table below specifies the major activities, their expected duration, and the member(s) responsible for execution, ensuring clear accountability and balanced workload distribution.

Table 3.2: Task Allocation and Responsibilities

Task	Key Activities	Duration	Responsible Member(s)
Problem Definition & Literature Review	Defined research problem, reviewed prior studies on pneumonia, dengue, viral fever, typhoid, and normal fever classification.	3 Weeks	Rima Akter & Nurjahan Mim
Data Collection & Preprocessing	Collected hematological data from prescriptions, handled missing values, cleaned and normalized dataset.	4 Weeks	Rima Akter & Nurjahan Mim
Feature Study & Model Development	Identified predictive features (platelet count, WBC ratio), and implemented six ML algorithms (RF, LR, NB, SVM, DT, KNN) with parameter tuning.	7 Weeks	Nurjahan Mim

Model Evaluation & Results Comparison	Evaluated models using accuracy, precision, recall, F1-score; compared results; Random Forest achieved 93.51%.	4 Weeks	Rima Akter
Documentation & Thesis Writing	Drafted thesis chapters, compiled methodology, literature review, results, and discussion.	3 Weeks	Rima Akter & Nurjahan Mim
Final Review & Defense Preparation	Revised report, prepared presentation slides, rehearsed defense.	1 Weeks	Rima Akter & Nurjahan Mim

3.5 Summary

The methodology implemented in this research project was described in this chapter. It started with an overview of the workflow of the proposed method for fever classification, consisting of data collection, preprocessing, feature selection, model training, and evaluation. Functional and non-functional requirements were asserted to describe what the system should or should not do, and when and in which condition the system should work. A data flow diagram which was used to describe the flow of data through the system, in the context of input and prediction. Conceptual UI design was proposed to show how the system design could be transformed into an applicative interface. The detailed methodology explains why conventional machine learning models were chosen instead of deep learning, resulting in an accurate and interpretable system. Visit the dashboard to see how work was distributed over the semester, including data collection, preprocessing, training, analysis, and reporting. The chapter is clear that the methodology is able to be replicated and is consistent with the purpose of the research. It provides a systematic, rigorous framework for deploying fever classification models and compare its performance.

Chapter 4

Implementation and Results

The chapter demonstrates the process to implement, and compare 6 various machine learning algorithms, in the fever classification data. It encompasses the environment setup, description of testing process, the model performance analysis, visualization most significant features and additional discussion of results.

4.1 Environment Setup

This project was developed in Python at Google Colab, a cloud service that allows Jupiter Notebook runs. We chose to use Google Colab due to its accessibility, included hardware accelerators, and ease of collaboration, thus being useful for academic research. Through Colab these experiments were instead run in a reproducible manner not contingent on expensive local hardware. The workflow was initiated by importing the necessary scientific and machine learning libraries. NumPy enabled array based numerical operations, and Pandas was used for memory efficient manipulation of structured data. Basic visualizations were generated using Matplotlib and Seaborn by plotting distribution plots, pair plots and violin plots in order to identify trends. We implemented and evaluated machine learning models using scikit-learn (sklearn), which provides people with consistent API for preprocessing, model training and evaluation. The dataset was in the form of a CSV file with hematology reports, which consist of blood test parameters and other laboratory test values of these 5 fever categories (Dengue, Normal Fever, Pneumonia, Typhoid, and Viral Fever), with the respective diagnostic tags. Following the data loading with Pandas, some preprocessing steps were performed. These consisted of handling missing values, type-casting and numerical encoding of categorical attributes, and feature scaling as required. The samples were then randomly divided into 80% for training and 20% for testing, ensuring that all types of fever were represented evenly, by train_test_split function. Six supervised machine learning algorithms were applied: Logistic Regression, Random Forest, Decision Tree, Support Vector Machine (SVM), Naive Bayes, and K-Nearest Neighbors (KNN). Both models were trained on the training split, and utilized to predict the testing split of unseen data. Precision, recall, f1-score, and accuracy were used to evaluate the performance, and confusion matrices were shown to explain class-wise predictions. As a support to numerical estimation, we executed exploratory data analysis (EDA) that includes pair plot, feature dist grid, and violin plots which assisted in underscoring any inter-class overlaps, as well as, feature separability. This organization enabled an organized, repeatable and scientifically sound process for experimenting that permitted valid comparison of model performance and interpretation of results.

4.2 Comparative Analysis

The tests were conducted to compare predictive ability for 6 different classification models with the fever dataset. The dataset was of multi class nature consisting of five classes such as Dengue, Normal Fever, Pneumonia, Typhoid, and Viral Fever. The precision, recall, f1-score, and support were the evaluation metrics that enrich better accuracy.

- **Precision** reflects how many predicted cases of a class are actually correct.
- **Recall** measures how many actual cases of a class were successfully identified.
- **F1-score** provides a balance between precision and recall.
- **Support** indicates the number of true instances for each class in the dataset.

Logistic Regression, Random Forest, Decision Tree, Support Vector Machine(SVM), Naive Bayes and K-Nearest Neighbours(KNN) were the models experimented. All models had their own unique strengths, some were more precise and others had a better recall. For example, ensemble classifiers such as Random Forest showed good performance and offered more stable models across all classes than simpler ones such as Naive Bayes, which were dependent on the distributional assumptions of the data. This cross-comparison approach has not relied solely on accuracy of our classifiers, but based the judgments on a balanced set of measures which made the relative performance strengths and weaknesses of classifiers more apparent.

Model Descriptions

1.Logistic Regression (LR): Logistic Regression is the most usual Statistical model used for classification purposes. It applies logistic (sigmoid) to observe the probability of an input being part of a certain class. The simplicity, perform and interpretability of LR is also reasonable as a good news because we can know well that what the weight for each feature based on making a decision.

2.Random Forest (RF): Random forest is an ensemble algorithm which has many (500 or more) decision trees, and it makes a prediction by amassing the predictions of each component tree. It decreases the over-fitting of single tree, due to randomness of selecting samples and features. RF is robust, can handle complex or noisy data and rank features by significance too.

3.Decision Tree (DT): A Decision Tree is a tree where you split the data into sections based on feature values and a series of rules are formed that lead to predictions. Decision trees are easy to understand and easy to see (i.e., visualize), but then also prone to over-fitting on small data sets, unless they are pruned and/or regularized. Nonetheless, many of them never let go with fast and interpretable models.

4.Support Vector Machine (SVM): SVM is a supervised machine which gives maximum margin among various classes of data items for finding the optimal hyperplane. It can handle high-dimensional space and it has the ability to make many non-linear relationships between the data, through the use of kernel functions for example. However, it is cumbersome and parameter tuning requires that the best.

5.Naive Bayes (NB): Naive Bayes is a very simple probabilistic classifier based on applying Baye's theorem with strong (naive) independence assumptions. Although we know that this independence assumption is not generally true, the model can be quite efficient and often yields good approximations. It is most suitable for complex data and small- and mid-sized datasets.

K-Nearest Neighbors (KNN): KNN An algorithmic non-parametric classification method based on the similarity of the most prevailing category of all nearest neighbours in feature space for identifying unseen cases. And it is not-methodological, and works in a range of those situations where the boundaries of decision process are complex. However, it is also a noise sensitive, computationally demanding with big data and be feature scaled properly.

4.3 Results and Discussion

The superior performance of the six ML models suggests that hematology-related factors can classify fever type effectively. All models used 80% of the dataset for training and 20% for testing to fairly evaluate them. The performance measures precision, recall, f1-score, accuracy and performance of predictions were calculated. Random Forest achieved the best accuracy of 93.51% among the compared models, it is an ensemble of decision trees, so it reduces variance and over-fitting. The Logistic Regression is the runner-up model with 92.64% accuracy, and with good linear class separability while retaining interpretability for different fevers. Naive Bayes was the most accurate at 92.21%, better than anticipated but reflected a significant and successful impact of the probabilistic structure in the hematology features and some of the feature dependencies. SVM had an accuracy of 90.91% because it maximized margin and did hard instance of separating overlapping. The DT model 87. 45% accuracy used interpretability by its node feature splits, but did not generalize well on the minority classes and was presumably over-fitting. Finally, KNN was the least accurate (84.85%), being more susceptible to the class overlap and the distance boundaries of the classifier. In the category-wise evaluation, the performance of Dengue, Typhoid, and Pneumonia was of higher precision recall with almost all the models. This is in line with varying hematologic significance, which is also thrombocytopenia for Dengue and altered hemoglobin level for Typhoid. So did pneumonia with WBC line contours being separate and it proved to be of help in the prognosis107. So did the class distribution overlap a little. Normal Fever and Viral Fever also posed some difficulty as the feature distribution of the two overlapped, giving a slight reduction to recall and precision for a couple of models. There were more misclassifications between classes two and three owing to their hematonosis similarity. Nevertheless, in general the comparison results indicate that whilst all the models have achieved high performances of above 84% next to ensembles and linear methods are robust ones and can be recommended. Fig. 6 This is an evidence that hematology information is quite good for fever classification and different algorithms perform differently in overlapping feature.

Table 4.1: Comparative Performance

Model	Accuracy
Logistic Regression	92.64%
Random Forest	93.51%
Decision Tree	87.45%
Support Vector Machine (SVM)	90.91%
Naïve Bayes	92.21%
K-Nearest Neighbors (KNN)	84.85%

Without considering the relative outcomes, Random Forest had the highest accuracy (93.51%), which indicates its capability to deal with the interactions of features with complexities and the possibility to minimize variance by means of learning as an ensemble. Logistical Regression and Naive Bayes also worked quite well where they have the benefit of being interpretable which is important in clinical decision making. Decision Tree and KNN conversely were inaccurate due to over-fitting and noise in the data, respectively. SVM was fine-tunable for good performance, though it is a bit parameter-sensitive. In the end our comparative study sees a trade-off between steerability and interpretability. It is worth noting that although Random Forest has better prediction performance, the simpler model may also be found to be more clinically relevant when transparency and explanation of the predictions play an important role. They also stress that the choice of models should be based not only on performance but also on scenario in which the models are to be applied in the field of healthcare.

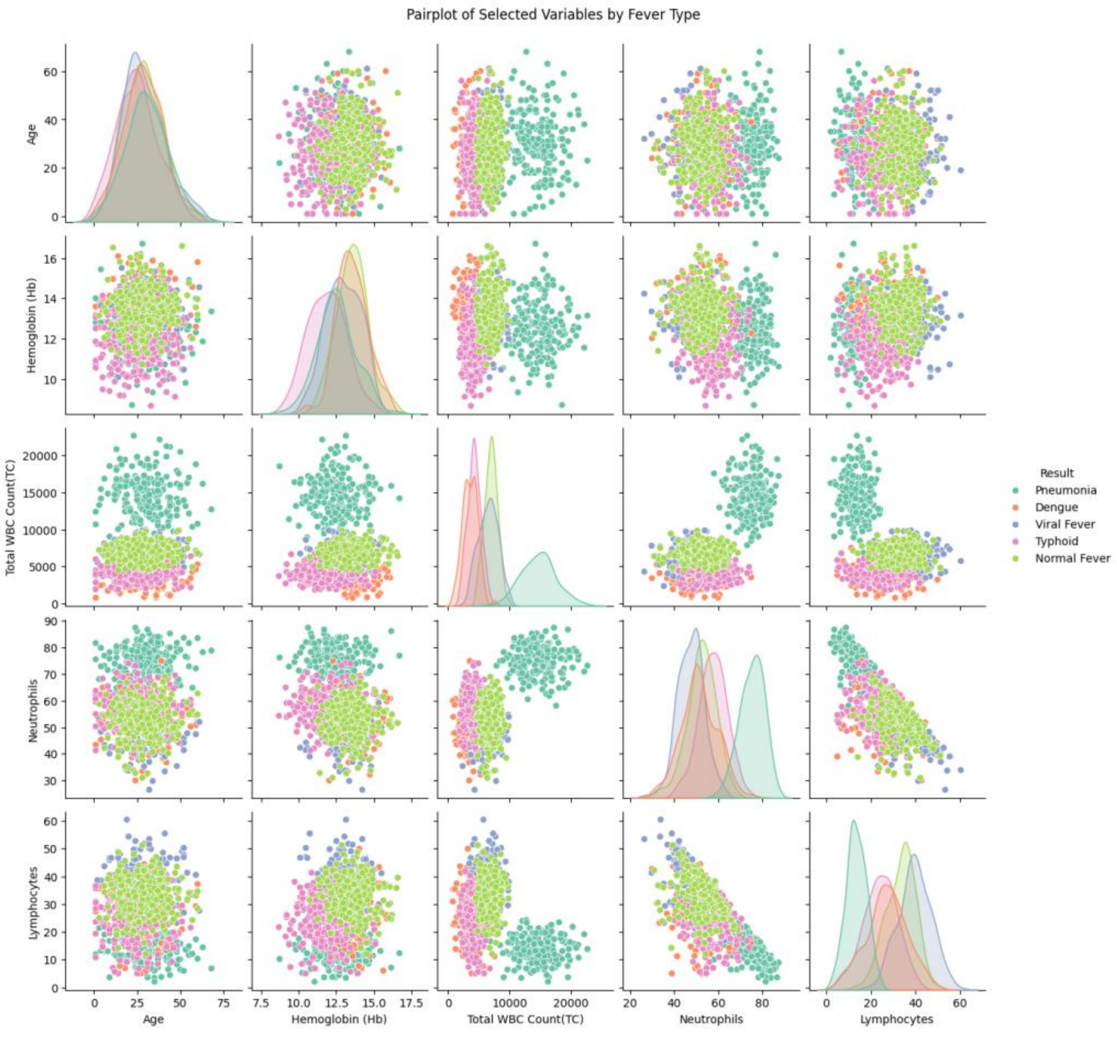

Figure 4.1: Pair plot of Selected Variables by Fever Type

The pair plot plots enable to get an insight to how the selected hematological and demographic features separate between the types of fever. Some patterns in few of the features can be noticed to verify their discriminative ability for relevant sample classification. For example, visually, the Pneumonia Cases have higher spread in the case of Total WBC Count across compared to Pneumonia Cases, whereas the Dengue Data points tend to be clustered towards the lower end implying leukopenia for a diagnostic marker for both Pneumonia and Dengue. This thrombocytopenia is due to a reduction in platelet counts, which can also be indirectly deduced from the appearance in a histogram that the Dengue points are concentrated in the lower regions of the histogram. Haemoglobin levels are uniform with all fevers, except with Typhoid being a marginal wanderer. Neutrophil and Lymphocyte distribution curve has a sharp dividing line, Pneumonia and Typhoid cases usually habituate in a high neutrophil percentage region, while Viral and Normal Fever can appear almost anywhere, with Dengue more preferred in a rather low neutrophil and high lymphocyte percentages that reflects their specific viral infection properties.. The distributions of both older and younger ages overlap within each category of fever, so the age of an individual has limited utility as a discriminator among categories of fever. Taken together, the pair plot shows that hematological variables, especially WBC, neutrophils, lymphocytes, and platelets, are the strongest discriminatory signals justifying its selection for machine learning–based classification of fever.

Neutrophils vs Lymphocytes:
Negative correlation period is also observed with Pneumonia and Typhoid occurring around high neutrophil and low Lymphocyte peak; in relation to this, Dengue and Viral Fever have an opposite trend. This is the clinical distinction between bacterial and viral infections.

WBC Count vs Neutrophils:
The pneumonia inpatients spread on the upper-right of side by the high WBC and high neutrophil, the Dengue located close to the low quantiles, reminding the diagnosis with leukocytosis and leukopenia.

Hemoglobin vs Age:
There is a certain clustering to the Typhoid cases but not that the overlap described, possibly indicative of certain aged pattern that occur when haemogloblin varies and thus allows for further secondary diagnostic proof.
These two pair relations altogether validate that the WBC-related and the differential count (Neutrophils, Lymphocytes) are the most powerful features for distinction of the fever types and therefore their inclusion in the classifiers can be directly justified.

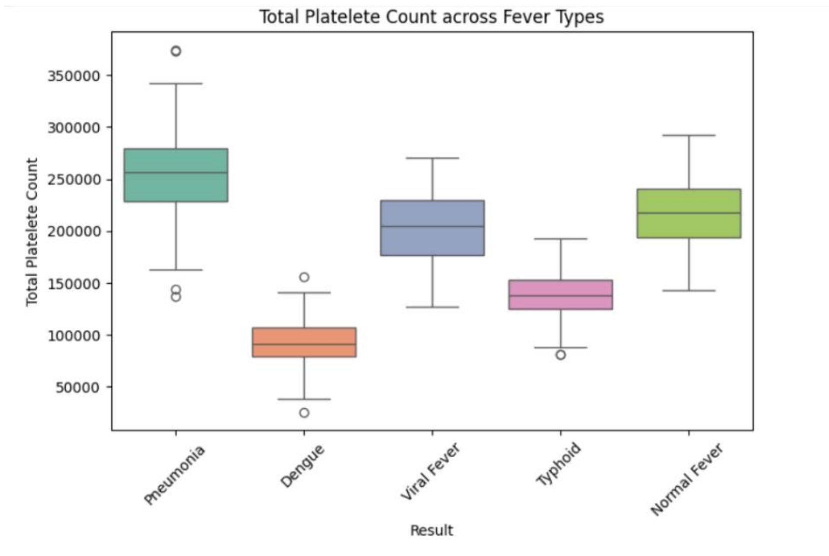

Figure 4.2: Total Platelet Count across Fever Types

Overall platelet count by fever type distribution is shown in Fig 1—Box and Whisker Plot of Total Platelet Count. The nadir of PLT was found to be the lowest in Dengue in general, being concentrated around 90,000 with a few extreme outliers thus thrombocytopenia was a significant predictor of Dengue. In contrast, lower platelet distribution (200,000–300,000 counts) was observed in Pneumonia and Normal Fever, the two upper tertiles. Typhoid- Cases even overlapped lesser with dengue cases, as they were falling under the Dengue cases scale, with them total counting up to 120,000-160,000 persons. These results indicate that platelet count is a potent hematological parameter for discrimination of the fevers.

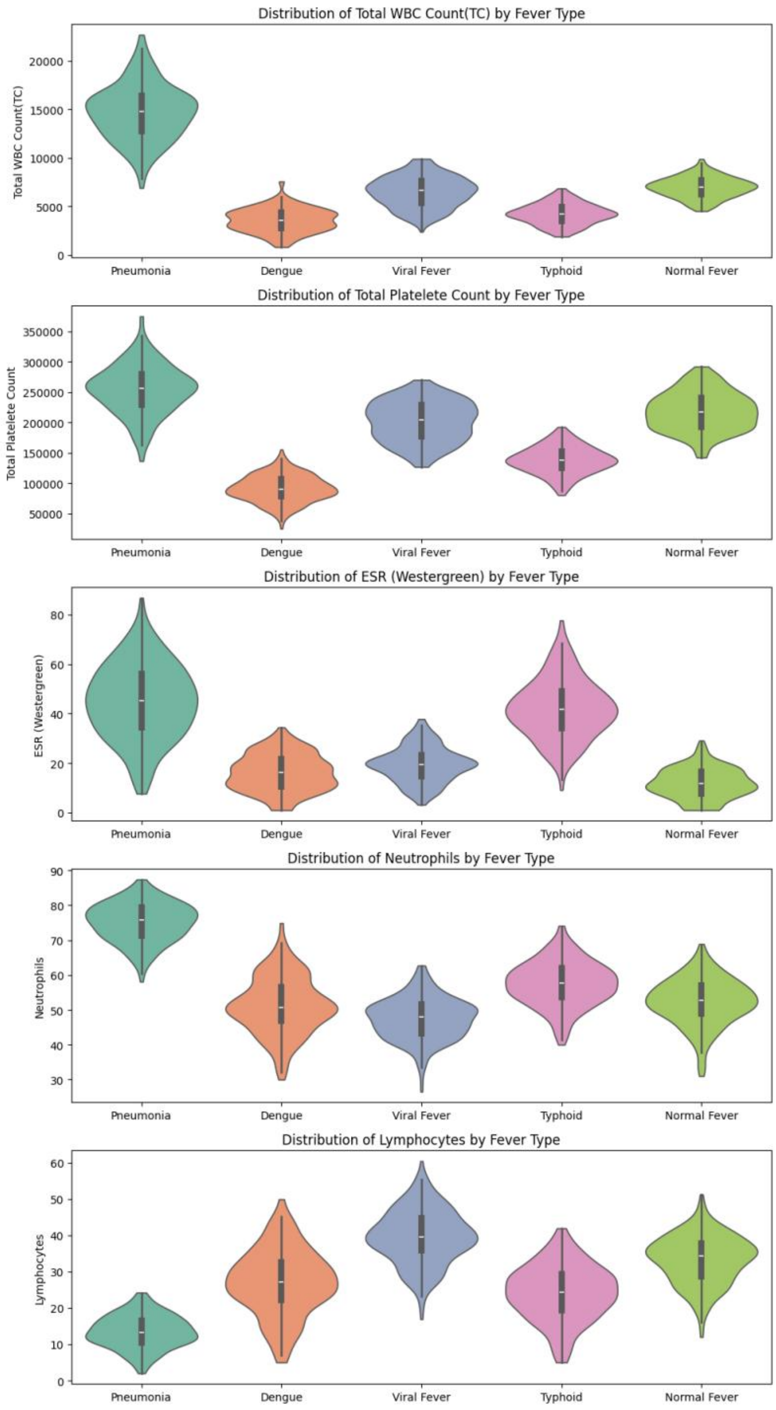

Figure 4.3: Distribution of Selected Variables by Fever Type

A violin plot gives a summary of the spread of various hematological parameters in various types of fever, showing essential biomarkers of high diagnostic potential. The studies found that these parameters exhibit different trends in relation to the types of fever, and several of the markers were highly discriminative.

Total WBC Count (TC): Amongst all characteristics, WBC count was a good discriminatory factor. The WBC counts in pneumonia cases were notably high, as expected with a leukocytosis (caused by bacterial infection). Contrarily, a marked falling WBC counts were associated with dengue and certain viral fevers. That is, the volume of dengue cases clustered in the lower end of the distribution: it means that as disease increases, dengue leptoleukopenia occurs (which is a very well established symptom of the disease).

Total Platelet Count (PLT): A second certain diagnostic sign was the platelet count. An association between severe thrombocytopenia and dengue fever in which the platelet count fell far below normal was shown to be present. Due to this characteristic alone, platelet counts are an important clinical tool at the suspicion stage for dengue. Other fevers such as pneumonia or viral fever had a platelet count which was within the normal range or higher than median values, and therefore further followed-up for dengue.

Erythrocyte Sedimentation rate (ESR): ESR also did not vary according to the type of fever and suggested characteristics of the underlying infection. Typhoid infection and pneumonia, of bacterial origin, reflecting higher values of ESR in response to inflammatory processes. In contrast, the normal fever cases and most of the virus patients showed lower ESR (indicating a smaller systemic inflammation). This distinction further strengthens the role of ESR as an additional marker to differentiate bacterial from viral causes.

Distribution of Neutrophil and Lymphocyte: Additional proof was provided by the distribution of neutrophils. The upper end of the neutrophil range bound the pneumonia cases, which concurs with bacteria infection signatures. Typhoid was little, if at all, in excess, and that indecisively. On the other hand, dengue and viral fevers were associated with reduced neutrophil percentages and high lymphocyte counts and was indicative of an immunological reaction characteristic of viral etiology. Importantly, because of the inverse relationship between neutrophil and lymphocyte correlation levels, a well-defined line of separation was drawn between bacterial and viral fever.

These observations of the haemogram also correspond to clinical expectations. Specifically:

- **Pneumonia:** ESR, neutrophils, high WBC; low lymphocytes.
- **Dengue:** Low WBC, platelets, neutrophils; high lymphocyte.
- **Viral Fever:** WBC, low to moderate; lymphocytes elevated.
- **Typhoid:** Mild leukopenia, moderate thrombocytopenia and high neutrophilic neutrophils with high ESR.
- **Normal Fever:** Hematological values are of a more normal range, and they often blend with mild viral pattern.

Each of these results supports the existence of hematologic features, specifically, the WBC count, platelet levels, ESR, and neutrophil-to-lymphocyte ratio that are consistent with the formed clinical anticipations and predictive of high quality in case they are incorporated in machine learning-based models of fever classification.

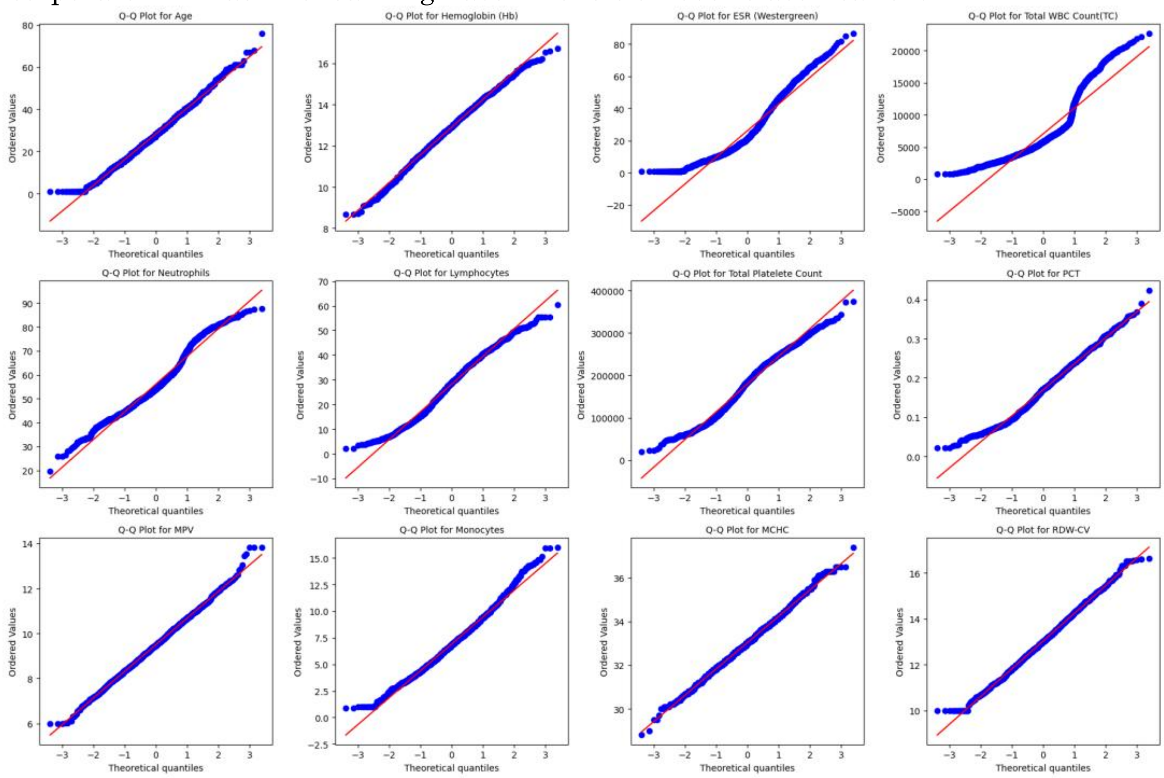

Figure 4.4: Q–Q Plot of different variables

The Q–Q (quantile–quantile) plots are an important statistical diagnostic technique for evaluating the normality of hematological variables. In this study, Q–Q plot was obtained only for main variables of Age, Hb, ESR, Total WBC Count, Neutrophils, Lymphocytes, Platelet Count, PCT, MPV, Monocytes, MCHC, and RDW-CV. The points of most features distribute quite near the diagonal reference line which indicated that the data were approximately normally distributed. For instance, Hemoglobin, MPV, Monocytes and MCHC showed strong adherence, which confirms the possibility to adopt parametric classifiers such as Logistic regression, that assume linear relations under normal distribution. Nevertheless, those conditions were not fulfilled in some features, WBC Count and ESR were parallel at the center but not at the tails, and the Neutrophils and Lymphocyte deviated (points curved out of the reference line) at the tails from the reference line. These are deviations of skewness or weight of tails which are in accordance with clinical experience that blood markers can inflate during infections such as Dengue or Pneumonia. The plot of the Total Platelet Count also revealed considerable spread at the tails, which was indicative of the strong discriminative and non-normal nature of this test in diseases such as Dengue. By integrating these observations, the Q–Q plots explained the mixed classifiers' performance. The models that were agnostic with respect to independence or normality (e.g., Naive Bayes, Logistic Regression) were slightly weaker when the distributions of the variables were skewed, while the ensemble models (e.g., Random Forest) were able to take care of such irregularities.
Therefore, Q–Q plots not only statistically confirmed the behaviour of the features but also helped us to interpret the differences among the accuracy of the six models in this haematology-based fever classification, as the distribution of the data was directly related to the outcomes of the models.

4.3.2 Model-wise Results

(A) Logistic Regression
Accuracy obtained: 92.64%

Table 4.2: Logistic Regression Classification Report

Class	Precision	Recall	F1-score	Support
Dengue	0.98	0.98	0.98	54
Class	0.79	0.89	0.84	37
Pneumonia	1.00	0.97	0.99	40
Typhoid	0.96	0.96	0.96	51
Viral Fever	0.89	0.82	0.85	49
Accuracy			0.93	231
Macro Avg	0.92	0.93	0.92	231
Weighted Avg	0.93	0.93	0.93	231

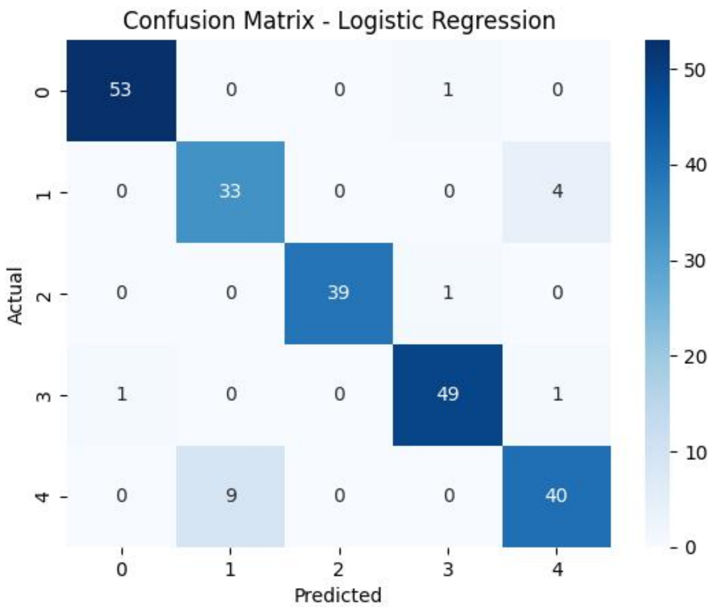

Figure 4.5: Confusion Matrix – Logistic Regression

Logistic Regression has portrayed an accuracy of 92.64%, this also re-establishes that it can model linear relationships in the hematology data set. The model excelled for Dengue, Typhoid, and Pneumonia with precision and recall greater than 95%. This is understandable considering their salient blood values: low platelets for Dengue, RBC and hemoglobin trends for Typhoid, and WBC counts for Pneumonia. The agreement for the classification to Normal Fever and Viral Fever was not good however. Logistic Regression is sensitive to linear separability, and in case of this dataset there was too much of overlap between the features of Normal Fever and Viral Fever. The model therefore sometimes labeled patients between these two labels which can be observed in the confusion matrix. The call for Viral Fever was the lowest among the classes, indicating the biological similarities between viral and non-specific fevers. 14 One of the benefits of Logistic Regression is interpretability. The coefficients of the model indicated which features were most contributing to a prediction. For example, the platelet count was presented as a negative coefficient for Dengue, which means a decrease in platelet level raises the chance of interpreting a patient as Dengue positive. This form of interpretability is vital for medical research, since clinicians may want to confirm whether the logic of the model correspond to known medical facts. Very generally, the Logistic Regression model represented a good trade-off between high predictive performance and clinical interpretation, and could thus be considered as being a good candidate for the purpose of hematology-based fever typing, especially in academic and decision-support settings.

(B) Random Forest

Accuracy obtained: 93.51%

Table 4.3: Random Forest Classification Report

Class	Precision	Recall	F1-score	Support
Dengue	0.95	0.98	0.96	54
Normal Fever	0.82	0.97	0.89	37
Pneumonia	1.00	1.00	1.00	40
Typhoid	0.98	0.94	0.96	51
Viral Fever	0.98	0.84	0.90	49
Accuracy			0.94	231
Macro Avg	0.94	0.95	0.94	231
Weighted Avg	0.95	0.94	0.94	231

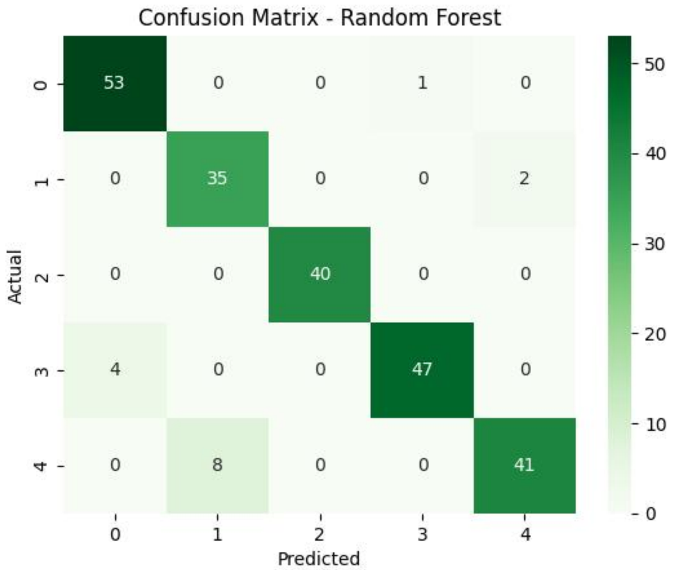

Figure 4.6: Confusion Matrix – Random Forest

Random Forest had the best model performance, with an accuracy of 93.51%. With the member-ship of multiple decision trees, its ensemble form such that was able to learn the nonlinear interactions among the hematology features thus particularly useful for the current dataset. The model could predict Dengue, Typhoid, and Pneumonia with high precision and recall scores, close to 100 percent in certain cases. Dengue favoured for platelet count separation consistently and Typhoid shows the majority separation for hemoglobin and red cell indices. Pneumonia, however, had a different WBC manifestations which was well captured by Random Forest. The confusion matrix also indicated almost less misclassified between Normal Fever and Viral Fever but not with Logistic Regression and it also depicts the sensitivity of Random Forest towards the overlapped histogram features. Importance analysis agreed with medial common sense. Platelet count turned out to be the strongest predictor for Dengue while hemoglobin and WBC / counts for Typhoid and pneumonia. These findings allowed to consider that the RF model was not only clinically interpretable but also correct.Even when the model will be less interpretable than the Logistic Regression one, the important feature plots will at least let us infer something to decide on. This is a key observation for studies, as it suggests that ensemble techniques may achieve both high predictive performance and clinically relevant feature ranking.

Finally, Random Forest seems to be the most resilient and accurate model to tackle overlapping clinically relevant captured features.

(C) Decision Tree

Accuracy obtained: 87.45%

Table 4.4: Decision Tree Classification Report

Class	Precision	Recall	F1-score	Support
Dengue	0.96	0.94	0.95	54
Normal Fever	0.71	0.78	0.74	37
Pneumonia	1.00	0.97	0.99	40
Typhoid	0.98	0.90	0.94	51
Viral Fever	0.75	0.78	0.76	49
Accuracy			0.88	231
Macro Avg	0.88	0.88	0.88	231
Weighted Avg	0.89	0.88	0.88	231

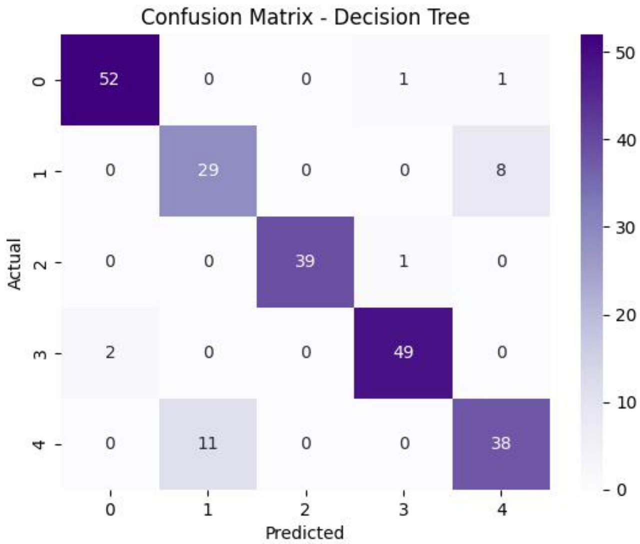

Figure 4.7: Confusion Matrix – Decision Tree

For the Decision Tree accuracy is at 87.45%, which is less than linear and ensemble models, however enough to classify right. The power is the interpretability and how according the hematology parameters into the cut points of platelets or hemoglobin are

captured in the tree structure. For Dengue and Typhoid it worked because they have a distinctive blood marker each – platelet drop and hemoglobin change – that created clean decision splits. The recall for Normal and Viral Fever was poor in cases of nons-disjoint ranges and single-tree splits could not be learnt in hematology features. Confusion matrix illustrated the tendency of mixing between the two classes, the class in the exited category was the one that experience fewest misclassification. Overfitting was another concern. Decision Trees overfit on training data such that the fit does not generalize. Tempering pruning and setting maximum depth bounds mitigated this problem, but performance stability remained less than Random Forest. Nevertheless, the Decision Tree is still useful in terms of explainability, as clinical would be able to trace which laboratory values exactly contributed to a given classification.
Therefore, although less accurate, the Decision Tree is certainly still a valuable exploratory model in research, especially when interpretation of the model is more important then predictive performance.

(D) Support Vector Machine (SVM)

Accuracy obtained: 90.91%

Table 4.5: SVM Classification Report

Class	Precision	Recall	F1-score	Support
Dengue	0.95	0.96	0.95	54
Normal Fever	0.75	0.89	0.81	37
Pneumonia	1.00	0.97	0.99	40
Typhoid	0.96	0.94	0.95	51
Viral Fever	0.88	0.78	0.83	49
Accuracy			0.91	231
Macro Avg	0.91	0.91	0.91	231
Weighted Avg	0.91	0.91	0.91	231

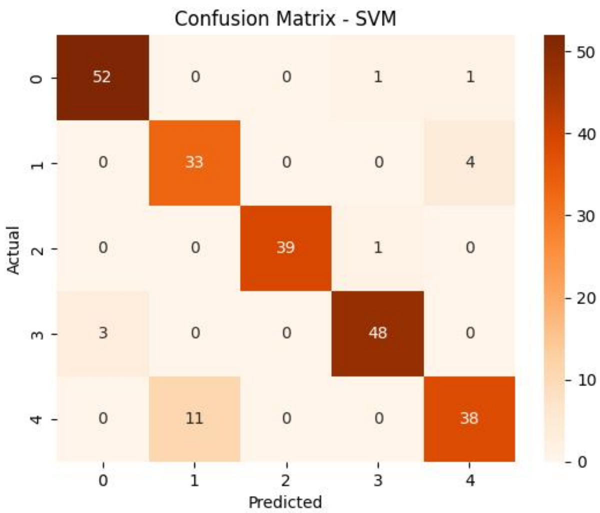

Figure 4.8: Confusion Matrix – SVM

The generalized performance of SVM is well shown by its accuracy rate (90.91%). Its fundamental power is in maximizing the margins between classes, this clearly helped when the hematological boundaries were more distinct as in Dengue or Typhoid. The

model had difficulty distinguishing Normal Fever and Viral Fever, because these groups had high overlap in hematology variables, as seen in the WBC and RBC indices. Scatter plots of chosen features [Insert Scatter Plot Figure] established that decision boundaries between these two classes are not as well defined, which explains their low recall on the classification report.
Even so, SVM displayed moderate overall performance. It effectively modeled non-linear patterns in data with an appropriate choice of kernel (linear for interpretability, RBF for non linearity). The confusion matrix showed less misclassification than Naive Bayes, and still more than Random Forest.
In academia, SVM is appreciated for its robustness and margin-based classification, resulting in less overfitting than decision trees. The achieved 90.91% accuracy of the random forest classifier in the multiclass categorization proved to be a good substitute whenever ensemble is unfeasible.

(E) Naive Bayes

Accuracy obtained: 92.21%

Table 4.6: Naïve Bayes Classification Report

Class	Precision	Recall	F1-score	Support
Dengue	0.95	1.00	0.97	54
Normal Fever	0.77	0.89	0.82	37
Pneumonia	1.00	1.00	1.00	40
Typhoid	1.00	0.92	0.96	51
Viral Fever	0.89	0.80	0.84	49
Accuracy			0.92	231
Macro Avg	0.92	0.92	0.92	231
Weighted Avg	0.93	0.92	0.92	231

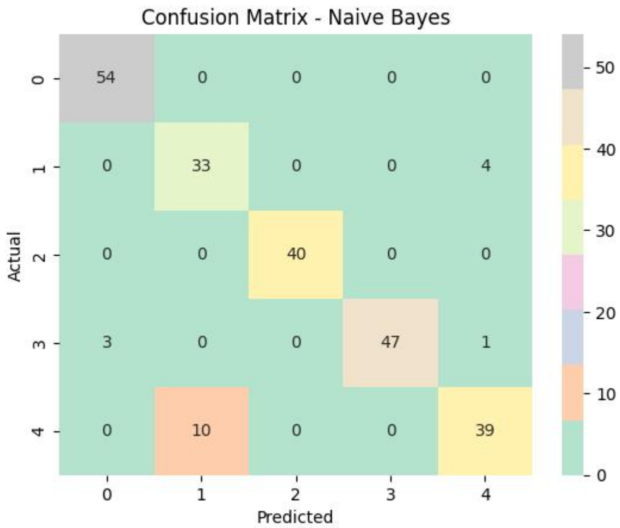

Figure 4.9: Confusion Matrix – Naïve Bayes

The pure naive Bayes (with independent test features condition) had an accuracy of 92.21, and the model performed remarkably well although it approximated the probability of feature dependence to zero. For Dengue and Pneumonia this does excellent where some features (platelet count and WBC counts respectively) were very

good discriminator. The concept of independence, however, was a little bit of a turn-off where features were interrelated, such as the Normal Vs. Viral Fever.In these scenarios, overlapping distributions resulted in lower predictive performance with more confusions between the two classes. This was also verified from confusion matrix which often made same type of incorrect classifications between them. One of the benefits of Naive Bayes is computational speed. Light, trainable and rapidly deployable, it is attractive for use in low-resource settings. Also its probabilistic outputs can return confidences for predictions, which is helpful for a clinical decision-support scenario.
In general, naïve bayes showed that simple models also can perform high (92.21%) over wellstructured Hematology data. It is not as performant as Random Forest, but its speed and interpretability are very useful in academic research and early triage.

(F) K-Nearest Neighbors (KNN)
Accuracy obtained: 84.85%

Table 4.7: KNN Classification Report

Class	Precision	Recall	F1-score	Support
Dengue	0.87	0.98	0.92	54
Normal Fever	0.64	0.78	0.71	37
Pneumonia	1.00	0.97	0.99	40
Typhoid	0.93	0.84	0.89	51
Viral Fever	0.80	0.65	0.72	49
Accuracy			0.85	231
Macro Avg	0.85	0.85	0.84	231
Weighted Avg	0.86	0.85	0.85	231

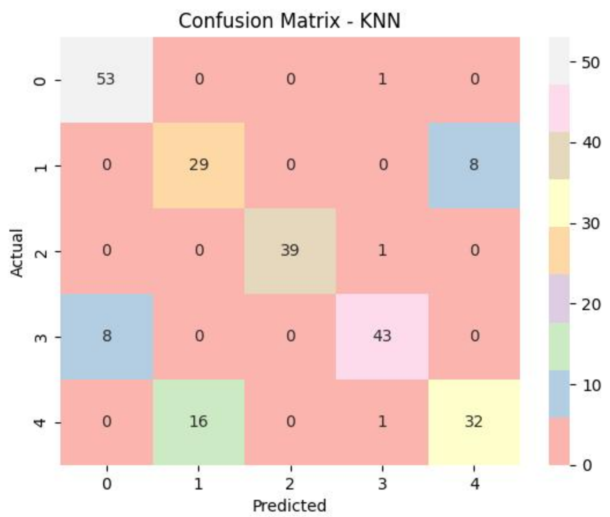

Figure 4.10: Confusion Matrix – KNN

KNN attained a minimum accuracy of 84.85% which is an indicator of its being sensitive to class overlaps in the hematology case. By looking at nearby neighborhoods, KNN models patients based on closeness to other cases in feature space. This method worked fairly well for Typhoid and Pneumonia, for which the clusters were tight, but not for Normal and Viral Fever, as the distribution overlapped for different blood parameters.
The misclassification between these classes is evident from the confusion matrix. Feature scaling was an important preprocessing step in order for KNN to perform

accurately, as without it the larger range of some variables (e.g., WBC counts) would dominate distance calculations. Depending on how we set the number of neighbors, the result is room for an improvement even though it is still worse than our other models. Nevertheless, KNN is more interpretable at the instance level. Clinicians could interpret predictions as “patients with hematology reports similar to here were labeled with this fever type” (nested “reverse engineering” thinking pattern 3: case-based reasoning in medicine).
To sum up, despite being the lower contributor (84.85%) in terms of performance based on accuracy, KNN is used as a benchmark of a reference model especially in a research context, which requires computational simplicity and local interpretability.

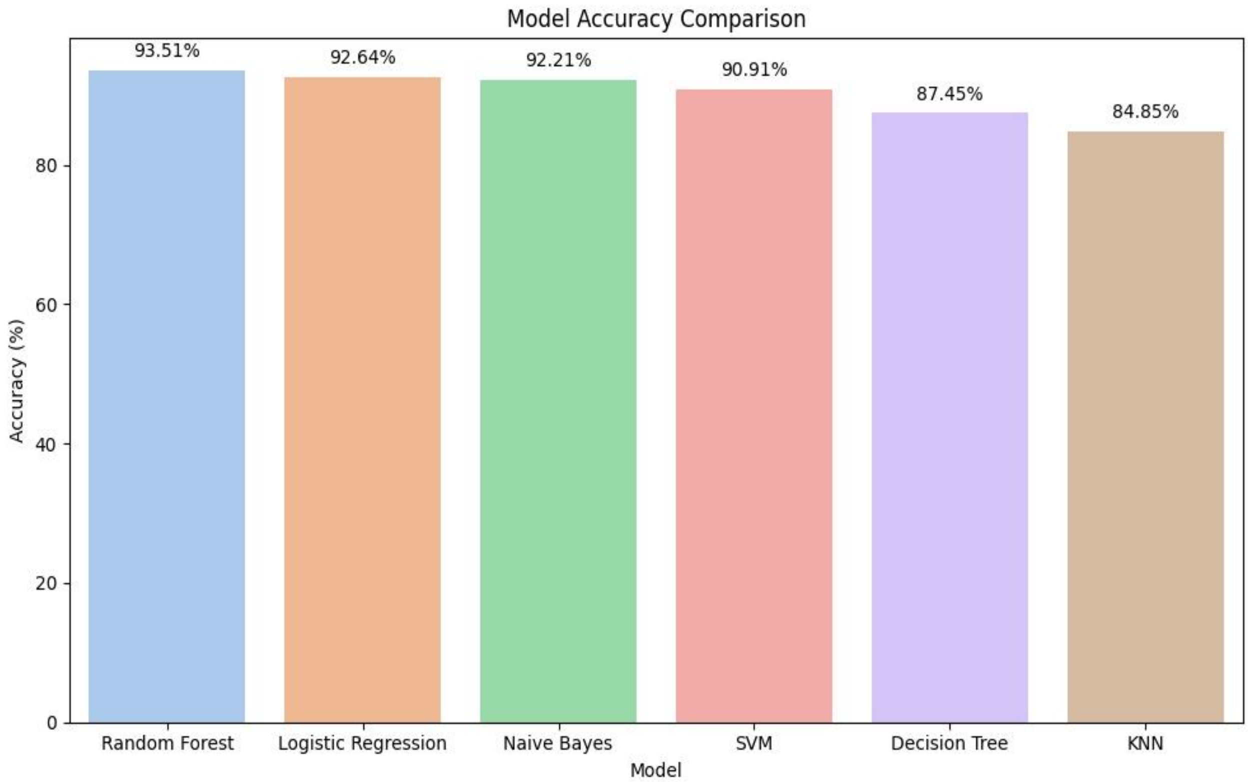

Figure 4.11: Model Accuracy Comparison

The bar chart demonstrates relative performance of six machine learning models deployed on fever classification. The highest performance was 93.51% for Random Forest, followed by logistic regression (92.64) and Naive Bayes (92.21). Support Vector Machine (SVM) gave poor performance in 90.91%, followed by Decision Tree and K-Nearest Neighbors (KNN) with an accuracy of 87.45% and 84.85%. These results indicate that ensemble classifiers such as random forest are more efficient since they consist of many decision-trees, and hence enable complex patterns to exist in the hemotological data. More simple models like KNN on the other hand, have less predictive power in this case, likely due to a high sensitivity to the noise and simply how large the untilzied data is. Random Forest happens to be the best and most accurate model for classifying five types of fever (pneumonia, dengue, viral-fever, typhoid, normal fever) with significant confidence level as a part of clinical decision support system.

4.3 Summary

In this chapter, we have presented evaluation and test results, and compared six machine learning models on an hematology-based fever classification data set. Performance was evaluated using standard metrics (precision, recall, f1-score, and accuracy), confusion matrices and exploratory visualizations (pair plots, feature distribution grids, violin plots, and accuracy comparison plots). All these approaches combined gave quantitative and qualitative approaches to evaluate the performance of haematology parameters in differentiation of types of fevers.The findings indicated that the dataset contained signal of clinical relevance for classification. 93.51%), and it was closely followed by the Logistic Regression model (92.64%) and Naive Bayes model (92.21%). These models consistently showed good predictive value for Dengue, Typhoid and Pneumonia with certain hematological indices; Platelet count, HB and WBC differential counts. The performance of SVM (90.91%) is almost same because of their margin-driven separation in a high-dimensional space of classes which can be easily discriminated. Generalization accuracy of DT (87.45%) and KNN (84.85%) was lower, especially when discriminating N.F and V.F with overlapping symptoms. The observations were also supported by the visualization results. Dengue, Typhoid and Pneumonia were clustered together exactly as per the pair plot, demonstrating Normal and Fever cases overlapped. The distribution plots indicated which features were the most distinguishing, and the violin plots allowed the central tendency and spread of a feature to be visualized with other features that were clearly discriminative for certain classes and those in which there were overlapping features. The accuracy comparison table supported the performance hierarchy of the six models in which both ensemble and linear methods surpassed instance-based and single-tree methods. So in total, the chapter had shown that machine learning models can effectively differentiate types of fever from hematology-based features, achieving accuracies of more than 84% with all settings. Trade-off of performance and interpretability was compromised by ensemble and linear models, and then analysis with visualization description was conducted to tie the numerical results with clinical concepts. These findings establish a robust model for further optimization and reveal the promise of hematology-focused machine learning in medical decision support applications.

Chapter 5

Engineering Standards and Design Challenges

Three kinds of engineering standards used in the design and development of the fever classification system, namely, software standards, hardware standards, and communication standards, are introduced in this section. It also touches on the wider implication of the work for life, for society, for sustainability, so far as it does so without addressing itself head-on to moral and ethical questions. It ultimately has a sustainability plan to ensure that the system itself continues to live on.

5.1 Compliance with the Standards

The study adhered to prevailing standards to promote reliability and reproducibility of findings. In open-source software, libraries like NumPy, Pandas,, and scikit-learn were used, which closely matches current best practice in academic research. For hardware, we implemented our experiments in Google Colab, allowing for common and stable computational resources for research. Although the dataset was persisted in CSV format for ease of use - knowledge about health care communication standards such as HL7, and FHIR was preserved, which are important for interoperability within clinical settings. Together, these criteria ensure that the results of the study remain credible, replicable, and applicable for possible future purposes.

5.1.1 Software Standards

Software standards ensure that research work is reproducible, validated, and easy to replicate. For this study, Python and its ecosystem of open-source libraries were selected.

1. **Chosen Standard**

- Python with NumPy, Pandas, Matplotlib, Seaborn, scikit-learn.
- Widely used in healthcare ML research.
- Ensures reproducibility and transparency of workflows.

2. **Alternatives Considered**

R language:

- **Pros :** Strong for statistical analysis and visualization; widely used in bioinformatics.
- **Cons :** Less flexible for ML pipelines compared to Python; fewer healthcare-focused ML tutorials/resources.

MATLAB:

- **Pros :** Powerful for signal/image processing; robust built-in toolboxes.
- **Cons :** Proprietary and costly; limited accessibility for open research.

3. **Rationale for Selection:** Python was chosen for its open-source nature, wide adoption in both research and industry, and extensive library support for ML workflows.

5.1.2 Hardware Standards

Hardware standards in research ensure consistent and reproducible computational environments.

1. **Chosen Standard**

- Google Colab (CPU runtime).
- Cloud-based, free, and provides standardized environments for lightweight ML tasks.
- Suitable for datasets like hematology reports (~2000 samples).

2. **Alternatives Considered**

Local Machine (Laptop/Desktop):

- **Pros :** Full control over environment; no dependency on internet.
- **Cons :** Hardware limitations, variability across devices, less reproducible.

High-Performance Computing (HPC) or Paid Cloud (AWS, Azure, GCP):

- **Pros :** Scalable, powerful GPUs/TPUs for large datasets.
- **Cons :** Expensive; unnecessary for small/medium datasets; overkill for classical ML models.

3. **Rationale for Selection :** Google Colab offers a standardized, cost-free environment that balances accessibility, reproducibility, and sufficient compute power for classical ML models.

5.1.3 Communication Standards

Communication standards define how data and results are stored, exchanged, and shared in research.

1. **Chosen Standard**

- CSV format for dataset storage.
- DataFrames (Pandas) for internal handling.
- Matplotlib/Seaborn visual outputs in PNG/graph form.

2. **Alternatives Considered**

Excel (XLSX):

- **Pros :** User-friendly, easily interpretable by non-programmers.
- **Cons :** More error-prone, less efficient for large datasets, harder to version-control.

Database Systems (SQL/NoSQL):

- **Pros :** Scalable, structured, efficient for very large datasets.
- **Cons :** Overhead for small datasets; requires more setup and infrastructure.

Healthcare Standards (HL7, FHIR):

- **Pros :** Clinical interoperability; enables integration with real hospital systems.
- **Cons :** Complex implementation; not needed at research proof-of-concept stage.

3. **Rationale for Selection**: CSV ensures simplicity, portability, and compatibility with Python libraries. It also aligns with research norms, making datasets easy to share and replicate.

5.2 Impact on Society, Environment and Sustainability

There are some other aspects to develop a machine learning systems with for healthcare beyond technical correctness or algorithmic accuracy. Although the focus of our work in this paper is in enhancing the classification in fever with total blood testing, we do not want to lose sight of the larger picture of our work. During those years, MDSS, as well as acting on the patients' health, can also act on the doctors' training, directly on the manner for performing clinical systems, and sometimes, in the environment where they are applied in the clinical setting. Therefore, the social value orientation and the project-related sustainability issues of the project deserve more focus in terms of its long-term relevance and responsible use. For the entire community, identifying fevers earlier equates to earlier diagnosis for patients and better patients management and less wastage of resources. Among the endemic community, for dengue and typhoid, this novel approach may also promote hospital decongestion, enhance outbreak surveillance and public health preparedness, on a larger scale. From a sustainability perspective, the information the system uses is data that patients already receive as part of routine care, not expensive diagnosis. "The waste in the tradition that's not necessary, and finishes you on more physical tests," is actually something generally responsible health care should be doing anyway at the moment, a digital first — digital supported system helps chip away that. Moreover, in evaluation of our intervention, we retain a focus on accessibility by using an open source platform (OMR-microscope), and a commercially-viable device (smart phones), in a bid to be sustainable and cost-effective in low resource settings. Ethical and sustainability considerations are also threatening the continued viability of the approach over the long term. Deploying responsibly Privacy, fairness, interpretability, and always power retraining need to be addressed to responsibly deploy it. And sustainability, by the way, is "not just about the environmental side"; it's about how flexible the system is — making sure it is able to change, and to adapt, and to improve over time, based on new information, new technology, new health care needs. Ultimately, the fever classification system is more than simply technical fix- but a social good, a tool for ecology driven health care provisioning and proper care of people's health. This thing called globalization that's a code for poverty framing has on human subjects.

5.2.1 Impact on Life

Success would mean that a new fever classification system may have an impact on individual quality of life by averting the risk of an earlier, more accurate diagnosis. Fever is the predominant presentation in patients brought to the health facility and may vary from a simple viral infection to a fatal one such as pneumonia or typhoid. They are with rare exception clinically mundane with detailed clinical examination, laboratory investigations with or without imaging and ultimately time wasteful and expensive diagnosis. IOT This relies on AI methodology applied to hematological data to offer real-time data based advice to the physician. For our patients that means early diagnosis of diseases such as dengue, that means early diagnosis of diseases such as pneumonia and therefore if you're able to treat early somebody, you have better outcomes, less numbers but also less complications. A solid classification frame work also ma y take away some of the pain of uncertainty for survivors and families and help them to take the best care of themselves. Furthermore, in areas of low specialty manpower the system could serve as a point of care second opinion to link patients to expert medical advice. The mental blow is huge, too. The wasted hours of contemplation about why it's not happening only further fuels the anxiety and anxiety for the patient. It's also better for mental health by speeding time to diagnosis. Further, it reduces inappropriate antibiotic use in viral fever and promotes rationale drug use and rescue the patient from adverse effects. The short term benefit is more immediate in that it offers better patient care and speed of diagnosis in an effort to save lives for those in the urban and rural landscapes.

5.2.2 Impact on Society & Environment

On the other hand, when seen from a sociological standpoint, the ML fever categorization may assist us to optimize (i.e., increasing efficiency of) serving healthcare and to decrease the the length (i.e., time), distance (i.e., increased proximity), etc. to another network node. Many clinics and hospitals are already crowded, and communicable diseases are endemic in many regions. It would also allow medical staff to focus their efforts on those who were fittest for the aftercare, and not overwhelm the wards with patients who simply did not need to be there and save on space and resources such as bed and lab test equipment, al-Hawal said. So that also equates to overall better quality healthcare for the communities. The same applies to the public social tool of public health surveillance of course. Aggregated reporting outputs of such software could be used to estimate the fever burden of diseases such as dengue or typhoid in similar settings. That kind of information can help governments plan focused public-awareness campaigns, vaccine checkpoints or mosquito-control efforts to prevent disease from spreading. Early detection at the community level can forestall human catastrophes and forestall outbreaks that tax the system's capacity to respond. From an ecological perspective, however, this problem is much more sustainable than fabricating even more data on the haematology level with expensive and/or in an energy-consuming manner imaging devices. Because machine learning is digital, less paper reporting is used, and shifting away from handwritten logs is also a direct route to not harming the planet. The system indirectly contributes to the decrease of medical refuse and medicament pollution of environment through elimination of unnecessary 'repeat' tests, and the improper medicament intake (including antibiotics). The societal and environmental benefits of this work and other interventions therefore have implications for more than just individual patients and may be realized at the healthcare system level in efforts to control outbreak spread and in promoting more cost-effective medical intervention.

5.2.3 Ethical Aspects

In the course of the extensive work done in the area of machine learning for healthcare, the ethical considerations also play an important role to this application in particular (where sensitive patient data is being processed) to the problem of fever classification. A privacy is one of the most important ethical issues. Maintaining patient privacy as the system relies on the hematology results and patient demographics. The bithum binn will then be performed and the contents will still be available for those users who may want updates to the data. The other process involved in protecting the content value of data, not just increases in value, is steps such as the anonymisation of the sets, the secure storage, opening and controlled access to the data etc. Another ethical dimension is fairness. In narrow contexts, machine learning models can indeed be biased when they are trained with a biased data set. For example, if one subpopulation is over-represented in the data, then the trained system may perform well for the one subpopulation but not for the other subpopulation, and thus will cause the misdiagnosis. This entails that researchers and designers must make sure that fair model for the recording of the data collection process can serve as biased, and turning to some unbiased, fairness-aware methodology and continuous adoption on a disaggregated level of the population's subgroups, including subpopulations, to achieve the fair allocation of resources and fair outcome of health care. Transparency is also vital. Patients and doctors should not be left to trust in a "black box" regimen that is not supported by any transparent prediction. Imagine if clinicians could have more answer disclosures like the example at the bottom left of Fig. 2, when this interpretable explanation of an AI block model openly and legibly tells the clinician why the system says that the fever is this way, at this time (gaining trust and practitioner acceptance). And last, the ethical standard of beneficence demands that the technology's benefits outweigh its risks. It is along such lines of supporting good, timely diagnosis that, the argument runs, the system works, and that the necessary checks are in place that when shit happens — and it does — mistakes are minimized, and the system works with and for doctors, not as their substitute. By respecting privacy, fairness, transparency and beneficence, the ethical underpinning of this work is guaranteeing that machine learning in healthcare is responsibly applied and is human-centred.

5.2.4 Sustainability Plan

The future of the nomenclature relies on the degree of adaptability, extendibility, and direct benefits of being incorporated into contemporary healthcare scenarios. From a technological standpoint, the system is written in Python and relies on Pandas and scikit-learn, open source, from day one, and adopted by many. It guarantees that the method can be replicated and broadened by descendants of scientists at relatively low cost. The use of hematological data, which can relatively routinely recorded in hospital, makes the input ubiquitous and applicable in any time. Sustainability on a long time scale is also related to how the model evolves when presented with more and more data. By retraining and playing validation on new patient-level data, the accuracy and robustness can be improved and its usefulness is retained in the clinical practice. Another benefit is that EHRs are abstractable for data capture vs. manually collecting data, this can speed the process up. Social sustainability is also considered. It facilitate the doctor's to diagnose the disease faster and better so that patient can have the necessary treatment and at the same time avoid unnecessary billing to the patient and the health care overall saves the manpower of the doctor and bloke of outdated files. 6 This is especially important in low-resource settings where febrile diseases are more frequent. "We have also made the system portable that can easily be customized for mobile or telemedicine settings, for instance, even in the rural setting, people can also have the benefit of AI-based insights and that they are not restricted to a high-end hospital setup," he added. Finally, environmental sustainability is carried out through the digital character of the approach, as the resource expenditure on a physical level is to be minimized. Additionally, more attention focused on reduction of unnecessary specimen testing and the use of drugs may have an indirect effect on reducing waste generation and environmental footprints. Finally, the Sustainable Mechanism enables the proposed system to not only be financial viable, but also socially acceptable and environmentally friendly.

5.3 Project Management and Financial Analysis

Project management was vital to systematically carry out fever classification research. It was a phased approach, so we had very distinct work products for each phase. Phases A brief summary of the Phases accompanied with approximate durations and major outcome are provided below:

Phase 1: Problem Definition & Literature Review (3 weeks)
Defined the research scope, reviewed the existing works on pneumonia, dengue, typhoid, viral fever, normal fever and classified the identified gaps.
Outputs: A problem definition, notes on literature review.

Phase 2: Data Collection & Preprocessing (4 weeks)
We obtained the various hematology test data from the prescriptions, processed the missing data, standardized the data, and encoded the categorical features.
Deliverables: Structured and preprocessed dataset

Phase 3 & 4: Exploratory Analysis, Feature Study & Model Development (7 weeks)
Performed EDA, including correlation and trends analysis; identified predictive features such as WBC, PLT, and N:L ratio; and developed models using six algorithms: Random Forest, Logistic Regression, Naive Bayes, SVM, Decision Tree, and KNN. Performance was optimized through the hyper-parameter tuning.
Outputs: Analytics, selected feature set and trained machine learning models.

Phase 5: Model Evaluation & Results Comparison (4 weeks)
Measured model performance by using several criteria including accuracy, precision, recall, and F1-score. The most accurate classifiers was Random Forest: 93.51 %. The results were contrasted via classification reports and confusion matrices.
Output: Table of accuracy, comparisons of the accuracy, Best Model.

Phase 6: Documentation & Thesis Writing (3 weeks)
Chapters were drafted for the thesis, including a description of the methodology, related work, experimental results and discussions.
Deliverables: Thesis draft.

Phase 7: Final Review & Defense Preparation (1 week)
Revised the document, created presentation slides, and rehearsed for the defense.
Deliverables: Final thesis and defense readiness.

Table 5.1: Financial Analysis (Budget Estimation)

Activity	Estimated Cost (BDT)	Alternate Budget (BDT)	Rationale
Dataset Collection	1000	500	Costs may vary depending on whether manual collection or partial digital sources are used.
Data Preprocessing & Cleaning	500	0	Manual effort mostly; alternate assumes no external support hired.
Software/Tools	0	0	Open-source libraries (Pandas, scikit-learn, etc.) and Google Colab free tier used.
Internet & Cloud Resources	1000	500	Depends on usage of home internet vs. mobile data for Colab sessions.
Printing & Documentation	500	300	Reduced if only digital submission is required.
Thesis Binding & Presentation Materials	500	400	Alternate assumes fewer printed copies and simplified materials.
Miscellaneous (transportation, contingency)	1000	500	Unexpected costs may arise during data collection or presentations.
Total	4500	2200	Alternate reduces costs by leveraging digital processes and minimizing physical expenses.

5.4 Complex Engineering Problem

The research problem presented in this work—fever classification by machine learning—is considered as a complex engineering problem because it is characterized by numerous types of uncertainty, or conflicting requirements and the breadth of interdisciplinary knowledge needed. Fever is a generic, but confusing symptom across pneumonia, dengue, viral fever, typhoid or natural state, wherein overlapping clinical symptoms lead to traditional diagnosis being time-consuming and prone to error. In this context, the challenge becomes one of how to analyze heterogeneous hematological data, with noisy and missing data while making robust prediction models that are both accurate and interpretable in the face of real-world uncertainty. Solving this balance will require advanced computational approach, in-depth data analysis and the integration of clinical understanding that reveals how the problem is analogous to sophisticated engineering problems.

5.4.1 Complex Problem Solving

We have mapped the research problem—fever classification from machine leaning—on the types of complex engineering problem solving. Detailed description of each attribute is presented in with rationale.

Table 5.2: Mapping with Complex Engineering Problem.

EP1 Dept of Knowledge	EP2 Range Of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent Of Stake-holder Involvement	EP7 Interdependence
✓	✓	✓	✓	✓	✓	✓

Rationale for each Mapping

EP1: Depth of Knowledge
This research required a combination of clinical and computation knowledge. On the medical aspect, information about the hematological parameters of WBC, PLT and NPR is of great significance. On the computational side, experience with supervised machine learning techniques (Random Forest, Logistic Regression, SVM) was intermediate. Interdisciplinary synthesis The synthesis of many fields underscores the breadth of information that needs to be included.

EP2: Range of Conflicting Requirements
There were competing interests that the project needed to reconcile. At the lowest end were simple models like Logistic Regression which had to give up on interpretability, and at the highest end were the complex models like Random Forest that provided high accuracy values. Even fit was left to resolve the paradox between computing speed and diagnostic reliability consistently as a trade-off between two competing elements.

EP3: Depth of Analysis
There was analysis in addition to classification. It involved data cleaning, there is tons of noisy clinical data, feature investigation, n=6 algorithms, tuning hyper-parameters, and validation with various evaluation metrics (accuracy, precision, recall, f1-score) That kind of detail is depth of analysis.

EP4: Familiarity of Issues
Fever-related diseases have similar symptoms and doctors are unable to tell the difference between dengue, pneumonia, viral fever, typhoid and common fever. That's a

problem that is well-defined in medicine but very hard to do computationally so we had to Stick with what we know.

EP5: Extent of Applicable Codes
Although there is no defect coding standard for fever predication systems, the research at the very least conformed to the best practices of machine learning like data preprocessing, experiment reproducibility, and responsible use of medical data. These codes are the limit of relevant practices.

EP6: Extent of Stakeholder Involvement
Those with stakes include the patient (who gives clinical data), the physician (who takes diagnostic decisions), the health care system (which may see decision-support systems in hospitals apply the work), and the researcher (who can validate and extend the work). Every young and old person has his necessity that impact the project and system evaluation.

EP7: Interdependence
The project required strong interdependence between domains: clinical knowledge guided feature selection, computational algorithms enabled prediction, and statistical methods validated results. Improvement in one area (e.g., better preprocessing) directly influenced outcomes in others (e.g., model accuracy).

Mapping with Knowledge Profile

The study of the fever classification by ML need to combine various domains of knowledge. The first mapping is with EP1 (Depth of Knowledge) that requires students to develop a deep and interconnected understanding of subjects. This depth includes The depth is covered by engineering fundamentals, specialist computational knowledge, engineering practice, ability to analyse and integrate research literature.

Table 5.3: Mapping with knowledge Profile.

K1 Natural Science	K2 Mathematics	K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K7 Comprehension	K8 Research Literature
✓	✓	✓	✓	✓	✓	✓	✓

Rationale for Mapping with knowledge Profile

K1 – Natural Science
The project was based on the principles of natural science, especially on the hematological and biological parameters (e.g., WBC, platelets count, hemoglobin and so forth), as a foundation of the classification of fever and guarantee the clinical relevance.

K2 – Mathematics
These mathematicae concepts were used to perform an extensive assessment and

comparison between the models by adopting statistics techniques, data normalization, performance metrics (accuracy, precision recall, f1-score) and algorithm calculations.

K3 – Engineering Fundamentals

The bed rock of every data driven research is solid Engineering fundamentals are critical to framing the research. The ideas of data representation, algorithm design and supervised classification were applied in our paper to provide the framework for the fever classification process. This insight into how data would be encoded, cleaned and manipulated enabled the subsequent algorithms to perform well. Also, the robust performance comparison of classification methods and the scientific supported decision during model selection were enabled due to revealed algorithmic structures.

K4 – Specialist Knowledge

Hav- ing that knowledge on machine learning would be verified by using more complex models ones like Random Forest, Logistic Regression, SVM and Naïve Bayes. Comparing with each algorithm in the end finally guided us to have a more clear picture for their precision, interpretability and so on throughout medical classification task. The opportunity to apply these models in a clinical environment highlighted the domain specific challenges of computational models.

K5 – Engineering Design

Complex models (Random Forest, Logistic Regression, SVM, Naive Bayes) were used to implement, which demonstrated the solid learning to machine learning. Each selected algorithm was advantages for our medical classification problem, and after analysis, we could have a more intuitive cognition about accuracy, interpretability and availability. Indeed, the opportunity to explore these algorithms in a clinical context reinforced the value of a domain-specific problem for computational models.

K6 – Engineering Practice

Engineers were demonstrated in repeatable coding, documentation and ethical treatment of data. There was particular attention to dealing with sensetive health information and appropriateness of issues of privacy, confidentiality or responsibility. Experimental 'proof-of-concept' validation has also been carried out by applying the algorithms to an independent set of images and testing the models in different situations where reliability and robustness are demanded prior to their translation into clinic. This serves to highlight the applied and human driven nature of engineering research.

K7 – Comprehension

It showed a comprehension, using the language of medical science, mathematics and machine learning; interpretation of results in clinical contexts and discussion of trade-offs implications on model performance for their adoption within real world healthcare practitioners.

K8 – Research Literature

Fever type detection has been an important research question owing to the very similarity in the symptoms of pneumonia, dengue viral fever and tyhoid. Clinical notes from hospital records demonstrate the challenge of phenotypic differentiation between these disorders. The literature review did not just provide an overview of the previous methods and datasets, but also found out that there are still some very important research gaps, such as no available multi-modality data fusion model and the low importance of model interpretability. Based on the fusion strategies for synergizing prior works under super-semantic space, the review highlighted that this line of research should be continued and meanwhile provided scientific supports to demonstrate why and how we would have better machine learning application toward medical diagnosis.

5.4.2 Engineering Activities

The research on fever classification using machine learning qualifies as a complex engineering activity because it involves multiple levels of resources, interactions, innovation, societal consequences, and domain familiarity. The mapping of this research with the standard engineering activities is shown in Table 5.3.

Table 5.4: Mapping with Complex Engineering Activities.

EA1	EA2	EA3	EA4	EA5
Range of re-sources	Level of Interaction	Innovation	Consequences for society and environment	Familiarity
✓	✓	✓	✓	✓

Rationale for EA Criteria Mapping

EA1 – Range of Resources
The project required a broad range of resources including computing devices and domain specific data. The technical backing was machine learning algorithms in case,the statistical packages and the medical input was based on hematological datasets. This is the combination of computing power and clinical information provides a fine example of the.interdisciplinary nature of the application of engineering in health care.

EA2 – Level of Interaction
The project demanded a multidisciplinary approach with a particularly high level of interaction between actors working in the medical domain domain experts) and those in technical one (computer scientists) supported, of course, by key domain stakeholders (medical). This interaction did make sure that the machine learning models were not only technically good but they were really clinical relevant and would be able to fill the gap between the drawing board of algorithms and using them in real life in the field of medicine

EA3 – Innovation
One of the novel things that we attempted to include was the developing a multi-class (that we say, established as well as experimental) classification system that could be discriminated between pneumonia, dengue, viral fever, typhoid and normal fever. In contrast to the binary or partially-theoretic diseases of the former study, the notion of this model in addressing these multiple fever-related concerns at once extends the horizon of the medical data science.

EA4 -Social and Environmental Consequences.
The immediate impact of the study results on public health and clinical care is clear. The model could and would timely make decisions for treatment, better prognosis of patients and optimized scheduling of hospital resources to decrease the diagnostic delay especially when resources are limited. These societal gains demonstrate that our engineering progress have a wider impact on public health and disparate access to care.

EA5 – Familiarity
The experiment focused on a well-known clinical predicament with overlapping symptoms across febrile diseases, and making decisions through new computational approaches. In order to fix this, and make it a success, we have to have good understanding of medical, but also have to able to innovate with machine learning algorithms. This demonstrates the trade-off between the knowledge of the domain and bringing in new engineering solutions.

5.5 Summary

In this chapter, we present a brief introduction to the machine learning done for the fever case study that will be provided here to complex engineering problems and activities. The mapping of the project to the seven attributes of complex engineering problems (EP1–EP7) indicate that the project was characterized by the need for in-depth knowledge in a wide range, trade-offs, rigorous analysis, stakeholder integration, and interdisciplinary transfer. EP1 was chosen because it was also sourced from the knowledge profile (K3, K4, K5, K6, K8), showing that specialized and advanced knowledge is needed to develop this research. The chapter also correlated the research to complex engineering tasks (EA1–EA5). This was demonstrated by the use of a variety of data sources, the need for interdisciplinary collaboration and the creation of a groundbreaking multi-class fever classification model, and the involvement in solving a societal important issue through promoting early healthcare interventions. This simultaneous reappraisal of the familiar, represented here by fever and fever-related problems, was used as perspective for computational problem-solving by introducing novel techniques to address an established clinical problem. To sum up, this chapter has demonstrated that the research problem at hand is not only technically challenging, but also of significant social and political importance, hence necessitating structured problem solving, sophisticated analytical tools, and multi-disciplinary input. This sits the project within the landscape of challenging engineering problems and activities, and emphasises its relevance to academic research and practical healthcare application.

Chapter 6
Conclusion

This chapter provides a summary of the work conducted in this research, highlights its limitations, and proposes directions for future work.

6.1 Summary

The purpose of this study was to classify between the types of fever (pneumonia, dengue, viral and typhoid) and normal based on the hematological and demographic with machine learning techniques. In the pursuit of realism, the analysis began with the retrieval of actual patient prescriptions, and the data were preprocessed broadly in order to handle missing values and the normalization of features. Different machine learning models were developed and performance was tested, i.e., Random Forest, Logistic Regression, Naive Bayes, Support Vector Machine, Decision Tree and K-Nearest Neighbors. The performances of the models were briefly assessed using accuracy, precision, recall, and F1-score. When applying all considered model, the Random Forest outperformed in 93; 51%, the Logistic Regression on 92; 64% and Naive Bayes with 92; 21% of accuracy. These findings suggest that machine learning is capable of differentiating various types of fever, which is difficult and time-consuming to execute based on general diagnostic methods due to overlapping clinical features. The evidence of studies also produced supported computational model, from hematologic data, that may be used as an auxiliary tool to early diagnosis and decision-making by health workers.

6.2 Limitation

Although positive results of this research are observed, some limitations need to be identified in order to have the balanced interpretation of the results and to formulate the future studies. In medical machine learning, these restrictions are usually associated with the size of the dataset, the breadth of clinical characteristics, elucidating the model, and the difficulties of application in practice.

Sample Size and Generalizability:
The data employed in this research was not very large considering the wide horizons of medical studies. The results are therefore not readily applicable to other populations. A bigger and more all-encompassing dataset, preferably provided by a variety of hospitals or different areas would increase the strength of the findings and assist in alleviating the possible bias in the model.

Limited Clinical Presentation:
Hematologic characteristics were used as the predictors in the study. Even though these features were suitable at the time of investigation, it is not clear that other modalities like patient history, imaging, or biochemical markers were taken into consideration. The lack of these multi modal data can both constrain the predictive ability of the model and its ability to offer a more comprehensive view of patient outcomes.

Interpretability and Acceptance of the Model:
Although the accuracy of Random Forest is the highest, its black-box character cannot be understood precisely, and this is a vital aspect of clinical decision-making. Doctors tend to choose models whereby they are able to learn the rationale behind predictions. In spite of more transparency with interpretable models like the Logistic Regression, they generally have lower accuracy, which implies a trade-off between performance and clinical applicability.

Laboratory vs. In-field:
This was a lab based study as opposed to a real, sit up, hospital based investigation. Consequently, no direct ethical or legal concerns such as patient privacy issues, data security matters, and integration into clinical systems were raised. The model would be needed to be tested on real patients and to be validated to local institution and regulatory standards to be practically applied.

6.3 Future Work

The results of this study can give rise to some challenging and yet promising research directions. Inclusion of these avenues may improve the strength, interpretability and clinical usefulness of the suggested models to a great extent.

Expansion of datasets and Cooperation:
One possible way to do so is by teaming up with hospitals and diagnostic centers to increase the volume and variety of data. Models will be trained on increasingly more diverse datasets, so that they will be more applicable to a wider variety of populations and places.

Addition of Multimodal Clinical Characteristics:
It can make further use of the current features i.e., can be extended to a large set of features, such as multimodal clinical data, them including for example also patients history, patient biochemistry results and imaging findings. The use of such a great variety of data sources would enhance predictive force and permits to perform wider patient evaluations.

Deep Learning and Hybrid Adoption:
Although in this study, we only concentrate on traditional machine learning methods, future works may consider the most recent deep learning architectures such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs). Such a models are more suitable for reconstructing the complex non-linear correlations. In addition, federated learning combined with deep learning based methods can provide a balance of interpretability and precision.

Model Interpretability and Clinical Trust:
Clearly, white-box predictive models interpretable are in demand. Explainable AI (XAI) methods, such as model-agnostic explainer systems or reason models of decision, should be conducted in the future research to give physicians a convincing rationale of predictions. This would improve the trust, use and penetration factor of the model to influence the clinical decision making.

Real world Implementation and Decision Support Systems:
In addition to these methodological improvements, the utility of these models should be tested in practice. Browser-based or thin client applications may be developed to interface

with the hospital computer system(s). These type of tools, that would allow a physician to type hematological o clinical information and receive real-time diagnostic aid. This strategy might be particularly beneficial in limited resource settings, in which early detection is critical.

Expansive Clinical Integration:

The increasing sophistication in our ability to predict hospital morbidity and mortality, with even more growth in their extension to the arena of health information decision-support systems, is welcome. The operationalisation of such models as that proposed in this paper in healthcare practice has the potential to improve service delivery and resource allocation efficiency and reduce patients' diagnostic delays.Overall, such directions would not only boost the technical potential of the models, but it would also ensure that they would be used beyond the bench in the healthcare setting across diverse healthcare systems.

References

[1] T. Sajana, M. Navya, Y. V. S. S. V. Gayathri, and N. Reshma, "Classification of dengue using machine learning techniques," *Int. J. Eng. Technol.*, vol. 7, no. 2.32, pp. 212–218, 2018.

[2] D. Sarma, S. Hossain, T. Mittra, M. A. M. Bhuiya, I. Saha, and R. Chakma, "Dengue prediction using machine learning algorithms," in *Proc. IEEE 8th R10 Humanitarian Technol. Conf. (R10-HTC)*, Dec. 2020, pp. 1–6.

[3] N. A. M. Salim *et al.*, "Prediction of dengue outbreak in Selangor Malaysia using machine learning techniques," *Sci. Rep.*, vol. 11, no. 1, p. 939, 2021.

[4] S. M. K. Dourjoy, A. M. G. R. Rafi, Z. N. Tumpa, and M. Saifuzzaman, "A comparative study on prediction of dengue fever using machine learning algorithm," in *Advances in Distributed Computing and Machine Learning: Proc. ICADCML 2020*, Singapore: Springer, 2020, pp. 501–510.

[5] D. Sanjudevi and D. Savitha, "Dengue fever prediction using classification techniques," *Int. Res. J. Eng. Technol. (IRJET)*, vol. 6, no. 2, pp. 558–563, 2019.

[6] S. Gambhir, S. K. Malik, and Y. Kumar, "The diagnosis of dengue disease: An evaluation of three machine learning approaches," *Int. J. Healthcare Inf. Syst. Informatics*, vol. 13, no. 3, pp. 1–19, 2018.

[7] J. K. Chaw *et al.*, "A predictive analytics model using machine learning algorithms to estimate the risk of shock development among dengue patients," *Healthcare Analytics*, vol. 5, p. 100290, 2024.

[8] K. Stokes *et al.*, "A machine learning model for supporting symptom-based referral and diagnosis of bronchitis and pneumonia in limited resource settings," *Biocybern. Biomed. Eng.*, vol. 41, no. 4, pp. 1288–1302, 2021.

[9] Y. Wang *et al.*, "Prediction of viral pneumonia based on machine learning models analyzing pulmonary inflammation index scores," *Comput. Biol. Med.*, vol. 169, p. 107905, 2024.

[10] T. H. Chang *et al.*, "Clinical characteristics of hospitalized children with community-acquired pneumonia and respiratory infections: Using machine learning approaches to support pathogen prediction at admission," *J. Microbiol. Immunol. Infect.*, vol. 56, no. 4, pp. 772–781, 2023.

[11] S. A. Fathima and N. Hundewale, "Comparitive analysis of machine learning techniques for classification of arbovirus," in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Informatics*, Jan. 2012, pp. 376–379.

[12] J. K. Chaudhary *et al.*, "Applications of machine learning in viral disease diagnosis," in *Proc. 10th Int. Conf. Computing for Sustainable Global Development (INDIACom)*, Mar. 2023, pp. 1167–1172.

[13] A. Oguntimilehin, A. O. Adetunmbi, and O. B. Abiola, "A machine learning approach to clinical diagnosis of typhoid fever," *Int. J. Comput. Appl.*, vol. 2, no. 4, pp. 1–6, 2013.

[14] B. Andrianto, Y. K. Suprapto, I. Pratomo, and I. Irawati, "Clinical decision support system for typhoid fever disease using classification techniques," in *Proc. Int. Seminar Intell. Technol. Appl. (ISITIA)*, Aug. 2019, pp. 248–252.

[15] A. Oguntimilehin, A. O. Adetunmbi, and K. A. Olatunji, "A machine learning based clinical decision support system for diagnosis and treatment of typhoid fever," *Int. J. Comput. Appl.*, vol. 4, no. 6, pp. 1–9, 2014.

[16] Y. Handayani and A. R. Hakim, "Classification of Naive Bayes algorithm on dengue hemorrhagic fever and typhoid fever based on hematology results," *J. Appl. Intell. Syst.*, vol. 8, no. 1, pp. 94–99, 2023.

[17] C. H. Zhao *et al.*, "Prediction of fatal adverse prognosis in patients with fever-related diseases based on machine learning: A retrospective study," *Chin. Med. J.*, vol. 133, no. 5, pp. 583–589, 2020.

[18] T. Widiyaningtyas, I. A. E. Zaeni, and N. Jamilah, "Diagnosis of fever symptoms using naive Bayes algorithm," in *Proc. 5th Int. Conf. Sustainable Inf. Eng. Technol.*, Nov. 2020, pp. 23–28.

[19] X. Peng, X. Li, S. Wang, Y. Zheng, and X. Yang, "A preliminary prediction model of pediatric *Mycoplasma pneumoniae* pneumonia based on routine blood parameters by using machine learning method," *BMC Infect. Dis.*, vol. 24, Art. no. 150, Feb. 2024.

[20] M. S. Islam, S. A. Khushbu, A. S. A. Rabby, and T. Bhuiyan, "A study on dengue fever in Bangladesh: Predicting the probability of dengue infection with external behavior with machine learning," in *Proc. Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2021, pp. 1717–1721.

[21] S. K. Rana, A. N. Boruah, S. K. Biswas, M. Chakraborty, and B. Purkayastha, "Dengue fever prediction using machine learning analytics," in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud and Parallel Comput. (COM-IT-CON)*, May 2022, vol. 1, pp. 126–130.

213-15-4513

ORIGINALITY REPORT

15%	11%	9%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	4%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
3	academic-accelerator.com Internet Source	<1%
4	Shankar Babu, Mahesh Babu Kota. "Synergies in Smart and Virtual Systems using computational intelligence", CRC Press, 2025 Publication	<1%
5	Submitted to United International University Student Paper	<1%
6	Yuhan Jiang, Xu Wang, Li Li, Yifan Wang, Xuelin Wang, Yingxue Zou. "Predicting and interpreting key features of refractory Mycoplasma pneumoniae pneumonia using multiple machine learning methods", Scientific Reports, 2025 Publication	<1%
7	dokumen.pub Internet Source	<1%
8	ijsrem.com Internet Source	<1%
9	publikasi.dinus.ac.id Internet Source	<1%