Research Journal of Pure Science and Technology E-ISSN 2579-0536 P-ISSN 2695-2696
Vol 8. No. 6 2025 www.iiardjournals.org online version

Intelligent Framework for Classification of Influenza Severity for Respiratory Disease Management

Anthony Edet*[1], Omayeka Chukwuemeke[2], Helen Godwin[3] and Ebupntak Eyo[4]
[1,2,3]Department of Computer Science, Akwa Ibom State University, Mkpat Enin, Nigeria
[4]Scholarsway Research and Technology Development Center
Corresponding Author: anthonyedet73@gmail.com
DOI: 10.56201/rjpst.vol.8.no6.2025.pg29.46

Abstract
Introduction: *Influenza is a highly contagious respiratory illness that poses significant public health risks, with severity varying across individuals. Accurately assessing the severity of influenza is crucial for timely intervention and resource allocation. This study explores the application of machine learning, specifically Random Forest regression, to predict influenza severity using key physiological indicators.*
Problem Statement: *Traditional clinical assessment methods for influenza severity rely heavily on subjective evaluations and may not fully capture complex interactions among multiple health indicators. There is a need for a robust, data-driven approach that can objectively predict severity and inform early clinical decisions.*
Aim: *The primary objective of this research is to develop and evaluate a machine learning model capable of predicting the severity index of influenza based on key physiological features, thereby enabling early detection and improved patient management.*
Methodology: *A comprehensive dataset sourced from Kaggle was used, including features such as body temperature, respiratory rate, oxygen saturation, cough intensity, fatigue level, and white blood cell (WBC) count. After data preprocessing, the dataset was split into training (80%) and testing (20%) sets. The Random Forest regression model was selected for its capacity to handle nonlinear feature interactions and prevent overfitting. Model performance was assessed using Mean Squared Error (MSE) and the coefficient of determination (R^2). Feature importance analysis and visualizations, including regression plots and scatter visualizations, were conducted to interpret the model outputs.*
Results: *The Random Forest regression model achieved an MSE of 0.0173 and an R^2 value of 0.6630, indicating that the model explained approximately 66.3% of the variance in severity index. Key features such as respiratory rate, oxygen saturation, body temperature, and WBC count demonstrated high predictive significance. Higher respiratory rates and lower oxygen saturation were strongly associated with increased severity, while elevated body temperature and WBC count contributed to symptom progression. Visualizations reinforced these relationships, providing a clear view of the feature-severity dynamics.*
Recommendation: *The predictive capabilities of the developed model suggest its potential utility as a clinical decision-support tool for early influenza severity assessment. It is recommended that healthcare facilities consider integrating such models into patient monitoring systems for real-time severity evaluation. Future research should explore larger, more diverse datasets and investigate ensemble learning techniques to enhance predictive accuracy. Additionally, the integration of longitudinal patient data could further refine severity trajectory predictions, offering a more dynamic approach to disease management. This study underscores the transformative potential of machine learning in respiratory disease management, offering a pathway toward more proactive, personalized, and data-driven healthcare solutions.*

Keywords: *Influenza, Contagious Respiratory Illness, Covid-19, Pneumonia, SVM, Disease*

1. Introduction

Contagious respiratory illnesses, commonly spread through droplets and close human contact, are significant public health concerns worldwide. They range from mild infections, like the common cold, to more severe diseases, such as influenza, pneumonia, and COVID-19 (Ekong et al., 2022). Transmission occurs primarily through airborne particles expelled when an infected person coughs, sneezes, or talks. The severity and spread of these illnesses can vary widely, depending on the causative pathogen, environmental factors, and population susceptibility. For instance, respiratory illnesses caused by viruses, such as influenza and coronaviruses, tend to be highly transmissible and can lead to seasonal outbreaks or even global pandemics. This research focuses on the analysis of risk severity of Influenza as a contagious respiratory illness using machine learning approach. Influenza severity can range from mild symptoms to severe, life- threatening complications, particularly in high-risk groups such as young children, the elderly, pregnant women, and individuals with chronic health conditions, some which might stem from drugs or substance use (Eyo et al., 2025). Influenza can lead to severe respiratory distress, pneumonia, and, in some cases, multi-organ failure. The severity of contagious respiratory illnesses varies widely and depends on factors like the pathogen type, individual health status, and environmental conditions. For instance, while the common cold typically causes mild symptoms, infections like influenza and COVID-19 can lead to severe complications, particularly in vulnerable groups. Age, pre-existing health conditions, and compromised immune systems significantly impact disease outcomes, with older adults and those with chronic illnesses like asthma, diabetes, or cardiovascular diseases often at a higher risk of severe illness and hospitalization. Environmental factors, such as low humidity and cold temperatures, also increase transmission rates for viruses like influenza, which thrive in winter months (Ekong et al., 2023)). The strain on healthcare systems during severe outbreaks of respiratory illnesses is a considerable consequence of their contagious nature. In high-transmission situations, healthcare facilities often experience surges in patients, creating shortages of medical supplies and straining medical staff. The COVID-19 pandemic, for example, highlighted the severe impacts on hospitals worldwide, where critical resources like ventilators, personal protective equipment, and ICU beds were overwhelmed (Ekong et al., 2022)). In response, public health interventions, including vaccinations, social distancing, and mask mandates, are critical to controlling the spread and minimizing severe cases, especially in vulnerable populations. These efforts collectively aim to reduce hospitalizations and allow healthcare systems to function more effectively under pandemic conditions. Influenza poses a significant health risk, with severity levels ranging from mild symptoms to severe, life-threatening conditions, especially among vulnerable populations such as the elderly, young children, and those with chronic illnesses (Edet et al., 2024). Traditional methods of assessing influenza severity often rely on clinical judgment and symptom observation, which may lack the precision needed for early identification of high-risk cases. This delay can lead to increased strain on healthcare resources and may hinder timely interventions that could prevent complications and reduce hospitalizations. Given the global impact of seasonal influenza, there is an urgent need for more precise and efficient tools to predict which patients are at risk of severe illness. Our task is to address this problem by developing a machine learning model capable of detecting the severity level of influenza in patients based on clinical and demographic data. By using advanced algorithms, we aim to build a predictive system that can accurately classify influenza severity levels, thus enabling healthcare providers to prioritize high-risk patients for closer monitoring and proactive treatment. This approach not only has the potential to improve individual patient outcomes but also to optimize healthcare resources, especially during peak flu seasons when the demand on medical services is highest. The significance of this study lies in its potential to enhance early detection and management of severe influenza cases, which can improve patient outcomes and reduce healthcare burdens, especially during peak flu seasons. By using Support Vector Machine (SVM) and Random Forest models, this research aims to develop a predictive tool that can accurately classify influenza severity based on patient data, allowing healthcare providers to identify high-risk individuals for timely intervention. This proactive approach can lead to more effective resource allocation, minimize hospital admissions, and ultimately help prevent complications in vulnerable populations, demonstrating the vital role of machine learning in advancing public health and clinical decision-making

2. LITERATURE FOUNDATION

This section captures the theoretical background underpinning this research on influenza severity prediction using machine learning. It reviews existing studies and scholarly works related to the detection and classification of respiratory illnesses, particularly influenza, with a focus on computational methods and intelligent systems. Key areas discussed include the clinical characteristics of influenza, previous attempts to use artificial intelligence and machine learning for respiratory disease diagnosis.

2.1 CONTAGIOUS RESPIRATORY ILLNESSES

Contagious respiratory illnesses encompass a wide spectrum of diseases caused by bacterial, viral, or fungal pathogens that spread through respiratory droplets or aerosols (Centers for Disease Control and Prevention [CDC], 2020). Prominent examples include influenza, respiratory syncytial virus (RSV), and more recently, COVID-19, each of which exerts significant impacts on global public health. The primary characteristic of these illnesses is their ability to cause widespread outbreaks, often leading to high morbidity and, in severe cases, mortality (Fauci, Lane, & Redfield, 2020). Influenza and RSV, for instance, predominantly affect high-risk groups such as children and the elderly, while COVID-19 introduced complexities with its rapid evolution and novel immune-escape variants. The epidemiological characteristics of these diseases have been widely studied, revealing seasonal patterns in their incidence. Influenza, for example, peaks during the colder months due to favorable transmission conditions, while COVID-19 has demonstrated year-round circulation with periodic spikes linked to new variants (Moriyama, Hugentobler, & Iwasaki, 2020). Advanced surveillance techniques, including machine learning-driven epidemiological models, have proven effective in predicting the timing and scale of these outbreaks, aiding in the allocation of healthcare resources. The socioeconomic burden of contagious respiratory illnesses is immense, stemming from hospitalizations, lost productivity, and healthcare expenditures. Vaccine-preventable diseases like influenza highlight the importance of immunization programs, yet vaccine uptake remains inconsistent across populations (Grohskopf et al., 2019). Similarly, antiviral therapies, though effective when administered promptly, face barriers of accessibility and affordability in many low-income regions. Recent studies emphasize the role of digital technologies, such as predictive analytics and health informatics, in optimizing preventive and therapeutic interventions for respiratory illnesses. Contagious respiratory illnesses present a multifaceted challenge requiring coordinated global efforts to mitigate their impact. Advances in surveillance, vaccination, and treatment continue to evolve, but persistent issues such as vaccine hesitancy and health inequities must be addressed (Dubé et al., 2021). Machine learning approaches hold promise in enhancing disease prediction, thereby improving preparedness and response strategies for future outbreaks. Contagious respiratory illnesses remain a significant public health concern due to their high transmissibility and the substantial burden they place on healthcare systems. Influenza, COVID-19, and respiratory syncytial virus (RSV) are among the most studied of these diseases, with seasonal and pandemic outbreaks resulting in considerable morbidity and mortality. These illnesses are primarily transmitted via respiratory droplets, airborne particles, or contact with contaminated surfaces, emphasizing the importance of preventive measures such as vaccination, masking, and improved ventilation (Centers for Disease Control and Prevention [CDC], 2020). While public health interventions have evolved, the dynamic nature of these pathogens continues to challenge mitigation efforts. The clinical manifestations of these illnesses vary widely, ranging from mild symptoms such as fever and cough to severe complications like acute respiratory distress syndrome (ARDS) and multiorgan failure. For instance, influenza and RSV disproportionately affect children and older adults, whereas COVID-19 has presented unique challenges due to its broad spectrum of severity and post-acute sequelae. Studies have highlighted the importance of early and accurate diagnosis, which is now increasingly aided by machine learning algorithms capable of identifying subtle clinical patterns and improving the prediction of disease severity (Fauci, Lane, & Redfield, 2020). Recent advances in data science and computational modeling have provided valuable tools for addressing the complexities of contagious respiratory illnesses. Predictive analytics, driven by machine learning techniques such as random forest and support vector machines, have demonstrated high accuracy in forecasting outbreak trends and assessing patient risk profiles. These technologies enable healthcare providers to allocate resources more efficiently and develop tailored treatment plans. Despite these advancements, challenges persist, including the need for robust and equitable healthcare systems to ensure that emerging technologies benefit diverse populations globally (Cohen, Tohme, & Qin, 2017).

2.2Epidemiology of Respiratory Illnesses

Respiratory illnesses are among the most prevalent health conditions globally, with varying incidence and prevalence rates depending on the region, season, and pathogen involved. Influenza, for example, affects an estimated 5% to 15% of the global population annually, causing severe illness in millions of cases, particularly in vulnerable populations such as children, the elderly, and those with comorbidities (World Health Organization [WHO], 2019). Similarly, respiratory syncytial virus (RSV) is the leading cause of lower respiratory tract infections in infants and young children, accounting for significant hospitalizations worldwide (Moriyama et al., 2020). These illnesses often display seasonal trends, peaking during colder months in temperate regions due to increased indoor crowding and favorable viral survival conditions (Moriyama et al., 2020). In the case of emerging diseases, COVID-19 has dramatically reshaped the epidemiological landscape of respiratory illnesses since its emergence in late 2019. By 2022, it had infected over 600 million people worldwide, with mortality exceeding six million (Centers for Disease Control and Prevention [CDC], 2020). Its highly transmissible nature, coupled with the ability to cause severe outcomes in specific populations, underscores the importance of rapid epidemiological surveillance (Ranney et al., 2020). The pandemic highlighted how novel pathogens can disrupt healthcare systems, emphasizing the need for robust global preparedness for future respiratory outbreaks (World Health Organization [WHO], 2019). Socioeconomic disparities play a significant role in the burden of respiratory illnesses. Low-income and middle-income countries (LMICs) often face higher rates of severe disease due to factors such as limited healthcare infrastructure, reduced vaccine access, and delayed disease detection (Salmon et al., 2015). For example, pneumonia remains one of the leading causes of death in children under five in LMICs, accounting for 15% of all deaths in this age group (World Health Organization [WHO], 2019). Additionally, environmental factors like air pollution and poor sanitation contribute to the higher prevalence of respiratory diseases in these regions, highlighting the interconnectedness of social determinants of health and disease epidemiology (Leng et al., 2022). Advancements in data collection and predictive modeling have improved our understanding of respiratory illness epidemiology. Machine learning algorithms have been increasingly employed to analyze complex datasets, identify risk factors, and predict outbreak trends (Ekong et al., 2023). For instance, random forest and support vector machines have shown high accuracy in forecasting seasonal influenza patterns and assessing patient risk for severe outcomes (Leng et al., 2022). These tools are particularly valuable in developing real-time surveillance systems, enabling more effective interventions and resource allocation to mitigate the public health burden of respiratory illnesses (Belkacem et al., 2021). The epidemiology of respiratory illnesses is influenced by a multitude of factors, including climate, pathogen behavior, and human activity. Seasonal variations significantly affect the spread of diseases such as influenza and RSV, with peaks often occurring during colder months in temperate regions. This pattern is attributed to increased indoor gatherings, reduced ventilation, and the enhanced stability of respiratory pathogens in low-humidity environments (Moriyama et al., 2020). However, tropical regions experience more variable patterns, often influenced by rainfall and humidity, demonstrating the diverse epidemiological dynamics of respiratory illnesses globally (Leng et al., 2022). Demographic trends further shape the burden of respiratory diseases. Populations with higher proportions of young children or older adults typically experience a greater impact, as these groups are more susceptible to severe outcomes (Salmon et al., 2015). Urbanization, coupled with increased air pollution, has also led to rising incidences of respiratory diseases in densely populated areas (Leng et al., 2022). Moreover, the emergence of antimicrobial resistance complicates treatment strategies, emphasizing the critical need for improved surveillance systems and the integration of machine learning tools to anticipate and mitigate outbreaks (Belkacem et al., 2021). By leveraging these technological advancements, public health strategies can better target at- risk populations and reduce the global burden of respiratory illnesses.

2.3 Factors Influencing the Severity of Respiratory Illnesses

The severity of respiratory illnesses is influenced by various factors, including individual characteristics, environmental conditions, and healthcare accessibility. Chronic conditions such as asthma, diabetes, and cardiovascular diseases can exacerbate the progression of respiratory infections, leading to more severe outcomes. These comorbidities impair the immune system's ability to respond effectively to infections, thereby increasing the risk of complications like pneumonia and acute respiratory distress syndrome (ARDS) (Erraguntla et al., 2019). Furthermore, genetic factors also play a crucial role in determining the severity of respiratory diseases, with studies suggesting that specific gene variants may heighten susceptibility to severe illness (Tsai et al., 2024). Environmental and air quality factors significantly influence respiratory illness severity. Exposure to pollutants, such as particulate matter (PM2.5) and nitrogen dioxide (NO_2), has been linked to worsened respiratory outcomes, including increased hospitalizations for asthma and pneumonia (Yang et al., 2019). Poor air quality exacerbates airway inflammation and impairs lung function, particularly in vulnerable populations like the elderly and children (Ijaz et al., 2022). Additionally, seasonal changes affect the transmission and severity of respiratory illnesses, with winter months contributing to higher rates of viral infections due to increased indoor crowding and the persistence of pathogens in cold, dry air (Jiang et al., 2020). Access to healthcare and vaccination coverage also play a significant role in the severity of respiratory illnesses. Vaccination has proven to be an effective preventive measure, reducing the risk of severe disease in influenza and COVID-19 patients. However, disparities in access to vaccines and healthcare resources continue to impact health outcomes, especially in low-income and rural populations (Kumar et al., 2023). For example, delays in receiving timely medical interventions such as antiviral drugs and oxygen therapy can significantly worsen the prognosis of respiratory infections, particularly among vulnerable groups such as the elderly and immunocompromised individuals (Andrade- Arenas et al., 2024). Moreover, pathogen-specific characteristics, such as the virulence of different respiratory viruses and their ability to adapt to the host, are critical in determining the severity of illness. Highly virulent strains of influenza, coronaviruses, and other respiratory pathogens can cause more severe symptoms, including rapid progression to ARDS or multi-organ failure (Yi et al., 2023). The presence of co-infections, particularly bacterial infections in viral respiratory diseases, complicates the clinical course and increases the mortality rate (Tsai et al., 2024). Machine learning models are increasingly being used to predict the severity of these diseases by analyzing a combination of clinical indicators, pathogen characteristics, and patient data, thereby enabling more effective and targeted interventions (Jiang et al., 2020).

2.4 Influenza as a Contagious Respiratory Illness

Influenza, caused by influenza viruses, is one of the most common contagious respiratory illnesses worldwide, typically spreading through respiratory droplets when an infected person coughs, sneezes, or talks. Studies have shown that environmental factors, such as temperature and humidity, significantly influence the transmission and survival of influenza viruses (Yang et al., 2019). During colder months, the virus tends to survive longer in the air and on surfaces, which contributes to the seasonal outbreaks observed in temperate regions (Erraguntla et al., 2019). As a result, seasonal influenza remains a persistent challenge for public health systems globally. The clinical presentation of influenza ranges from mild symptoms like fever, chills, and fatigue, to severe complications such as pneumonia, which can be life-threatening, especially for vulnerable populations such as the elderly, young children, and individuals with underlying health conditions (Belkacem et al., 2021). Beyond its health impacts, influenza also places a significant economic burden on societies, through costs related to hospitalizations, lost productivity, and absenteeism from work (Ranney et al., 2020). Given these challenges, understanding the dynamics of influenza transmission and severity is crucial for timely interventions and minimizing the impact of seasonal outbreaks. Vaccination remains the primary preventive measure against influenza. However, the efficacy of vaccines varies from season to season, as the influenza virus evolves rapidly through antigenic drift and shift (Tsai et al., 2024). Even with these variations, the vaccination effort significantly reduces both the severity of illness and the rate of complications, especially when matched well with circulating strains (Kumar et al., 2023). In addition to vaccination, other measures like public awareness campaigns and the use of personal protective equipment can help curb transmission, as seen during the COVID-19 pandemic (Ranney et al., 2020). Recently, machine learning approaches have been utilized to improve influenza surveillance, predict seasonal trends, and even assist in diagnostics (Leng et al., 2022). These technologies can analyze patterns in clinical data, weather conditions, and social behavior to forecast influenza outbreaks more accurately, enabling more effective resource allocation and intervention strategies (Jiang et al., 2020). As computational tools continue to evolve, they offer promise for enhancing both the prevention and management of influenza, contributing to more proactive public health responses (Ijaz et al., 2022). Influenza is not only a major public health concern but also a significant driver of healthcare system stress, especially during peak seasons. The virus's capacity for mutation through antigenic drift and shift makes it challenging to produce a universally effective vaccine, resulting in the need for annual updates based on circulating strains (Tsai et al., 2024). Researchers are investigating more adaptable vaccination strategies, including universal influenza vaccines, which could potentially offer longer-lasting immunity and broader protection (Ekong et al., 2023). These advancements in vaccine development aim to reduce the global burden of influenza, particularly in resource-limited settings, where access to timely and effective interventions is often restricted (Andrade-Arenas et al., 2024).

In addition to vaccination efforts, non-pharmaceutical interventions (NPIs) such as social distancing, mask-wearing, and quarantine measures have been implemented to mitigate the spread of influenza, especially during pandemics like COVID-19 (Ranney et al., 2020). The effectiveness of these NPIs has been demonstrated in multiple studies, which show that simple precautions can drastically reduce transmission rates (Belkacem et al., 2021). The integration of machine learning techniques into epidemiological models has allowed for more accurate forecasting of influenza outbreaks, helping public health officials allocate resources more efficiently (Leng et al., 2022). By incorporating diverse data sources, including environmental factors and social behaviors, machine learning models provide enhanced predictive power, offering a more proactive approach to managing seasonal influenza (Yi et al., 2023).

2.5 Detection of Influenza:

Influenza, commonly known as the flu, is a contagious respiratory illness caused by influenza viruses. The disease can range from mild to severe and can lead to serious complications, especially in high-risk groups such as the elderly, young children, and individuals with chronic conditions. Influenza is typically diagnosed based on symptoms and confirmed with laboratory tests such as rapid antigen tests, polymerase chain reaction (PCR), or viral culture (Ekong et al., 2024). Healthcare providers may also use clinical assessments, such as lung ultrasound scores, oxygenation index, and respiratory index, to evaluate the severity of the infection.

2.6 Symptoms of Influenza:

Common symptoms of influenza include fever, chills, cough, sore throat, body aches, fatigue, and headache. Some individuals may also experience gastrointestinal symptoms such as nausea, vomiting, or diarrhea, although these are more common in children According to the Centers for Disease Control and Prevention (CDC), individuals with influenza may experience a sudden onset of symptoms, often within 1 to 4 days of exposure to the virus (Edet et al., 2024). In severe cases, the infection can lead to complications such as pneumonia, respiratory failure, or secondary bacterial infections.

2.7 Severity Levels of Influenza:

Influenza severity is often categorized into mild, moderate, and severe levels, based on clinical symptoms, vital signs, and the need for medical intervention. Mild cases may involve low-grade fever, fatigue, and mild respiratory symptoms, while moderate cases may involve higher fevers, significant fatigue, and more pronounced respiratory distress. Severe cases, especially in vulnerable populations, can lead to life-threatening complications like acute respiratory distress syndrome (ARDS), requiring intensive care and mechanical ventilation. Machine learning models, as discussed in previous studies, are increasingly used to predict the severity of influenza based on clinical data, which helps healthcare professionals to make more informed decisions regarding treatment and management.

2.8 Severity Assessment Using Machine Learning:

The severity of influenza can be effectively assessed using predictive models that incorporate data from various clinical parameters. For example, machine learning algorithms such as random forests and support vector machines can analyze data like respiratory indices, lung ultrasound scores, oxygenation levels, and clinical markers like white blood cell count to determine the severity of the illness. These models help identify individuals at higher risk of developing severe symptoms, enabling early intervention and more personalized care. The predictive power of these models can also assist in decision-making for resource allocation during influenza outbreaks.

3. METHODOLOGY

The methodology for this research follows a structured approach to build an accurate influenza severity classification system using machine learning. The dataset for this model was sourced from the Kaggle data repository, containing essential features and symptoms associated with influenza, such as fever, cough, fatigue, and shortness of breath, alongside severity indicators like oxygen saturation levels, respiratory rate, and the presence of comorbidities. Collecting such comprehensive clinical data ensures the system captures both the presence and severity of influenza, distinguishing between mild, moderate, and severe cases. To prepare the dataset for training, several preprocessing steps were performed. Missing values were handled by imputing median values for continuous variables and mode values for categorical features. Categorical features like symptom presence were transformed into numerical values using one-hot encoding. Feature scaling through Min-Max normalization was applied to bring all feature values into a uniform range, improving the performance of distance-based algorithms like Support Vector Machines (SVM). The dataset was then split into training and testing subsets, with 80% allocated for model training and 20% reserved for evaluation.

To enhance model performance and reduce computational complexity, feature importance analysis was conducted using a Random Forest classifier. This step identified the most critical features influencing severity, such as oxygen saturation, respiratory rate, shortness of breath, and the presence of underlying health conditions. Recursive Feature Elimination (RFE) was applied to select the most relevant subset of features, iteratively removing the least important features and retaining those with the highest predictive power. This process helped refine the feature set, reducing noise and preventing overfitting. With the optimized feature set, the system was trained using a Support Vector Machine (SVM) algorithm. The SVM was initialized with a radial basis function (RBF) kernel to capture complex, non-linear relationships between features. Hyperparameters such as the regularization parameter (C) and kernel coefficient (gamma) were fine-tuned through grid search with five-fold cross-validation to maximize performance. During training, the SVM learned the optimal hyperplane that best separates the severity classes by maximizing the margin between data points. The decision function, defined as

$$f(x)=w\cdot x+b \quad (1)$$

where w is the weight vector, x is the input feature vector, and b is the bias term, was used to classify each sample into one of the three severity categories. Once the model achieved satisfactory accuracy and generalization, it was deployed as a REST API using Flask, enabling real-time severity classification from new patient data. The Support Vector Machine (SVM) algorithm used in this research works by finding the optimal boundary, called a hyperplane, that separates different classes of influenza severity. The decision function of the SVM can be represented as seen in eqn (1) above. In this equation: **f(x)** is the output decision score for a given input **x** (a patient's feature vector). **w** is the weight vector, which determines the direction and slope of the hyperplane. **x** represents the input features, such as fever, oxygen level, and respiratory rate, **b** is the bias term, which shifts the decision boundary. During training, the SVM aims to maximize the margin — the distance between the hyperplane and the closest data points (support vectors) from each class. The optimization objective is to minimize the following function:

$$(1/2) * ||w||^2 + C * \sum(\xi_i) \quad (2)$$

Where:

$||w||^2$ is the squared norm of the weight vector, which helps control the margin width.

C is a regularization parameter that balances margin size and misclassification penalty.

ξ_i are slack variables that allow for some misclassification in non-linearly separable data.

For classification, after training the model, the system calculates **f(x)** for each new patient

sample: If **f(x) > 0**, the sample is classified as a severe case.
If **f(x) ≈ 0**, the sample is classified as moderate.
If **f(x) < 0**, the sample is classified as mild or not severe.

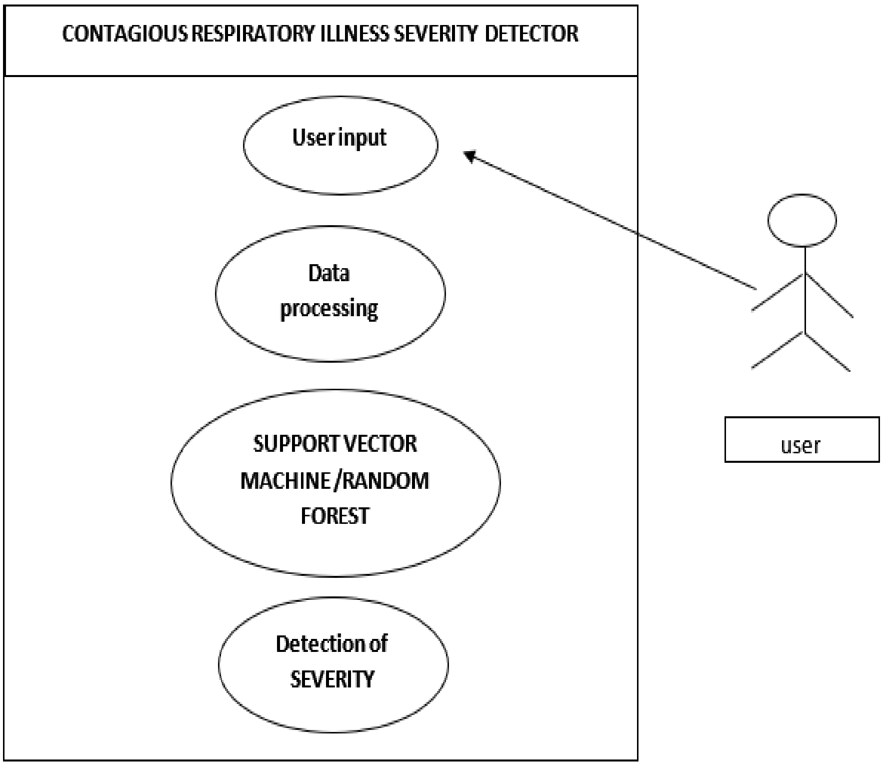

Fig.1: Use Case Diagram of the proposed system

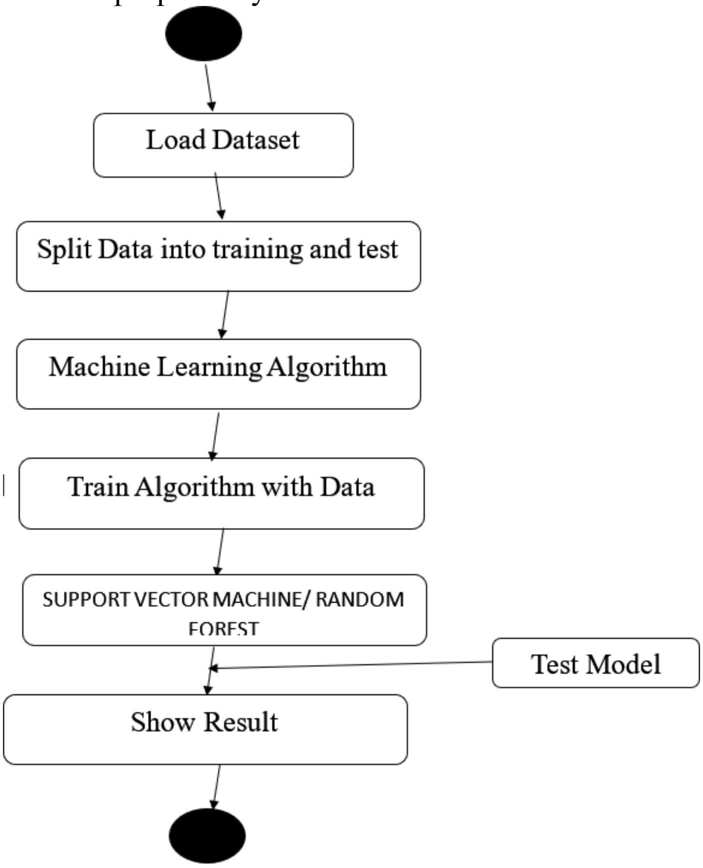

Fig.2: Activity Diagram of the proposed system (User)

4. RESULTS AND EVALUATION

This section presents the outcomes of the implemented influenza severity prediction model and evaluates its performance based on relevant metrics. The results reflect how well the model was able to classify cases into mild, moderate, or severe categories using the selected features and machine learning techniques. This section also discusses the implications of the results for real-world healthcare applications, particularly in supporting timely clinical decisions and prioritizing high-risk patients.

Feature Importances:

	Feature	Importance
0	Body_Temperature	0.060763
1	Respiratory_Rate	0.759975
2	Oxygen_Saturation	0.061782
3	Cough_Intensity	0.000000
4	Fatigue_Level	0.000000
5	Age	0.050194
6	Pre_Existing_Conditions	0.007100
7	WBC_Count	0.060185

Fig. 3: Feature Importance

Feature importance was conducted to identify and evaluate the contribution of each variable in predicting the severity index of influenza. This process involves assessing how much each feature influences the predictions made by the Random Forest regression model. By analyzing feature importance, it becomes possible to determine which variables are the most relevant in estimating the severity index of influenza, a critical factor for building an efficient and accurate predictive model. This approach not only improves model performance by focusing on the most impactful features but also provides insights into the relationship between various physiological and demographic indicators and the severity of influenza. The analysis ensures that only the most significant features are retained, reducing model complexity, computational requirements, and the risk of overfitting, thereby enhancing the interpretability and accuracy of the predictions.

	Body_Temperature	Respiratory_Rate	Oxygen_Saturation	Cough_Intensity \
0	38.19	23.02	90.60	NaN
1	40.78	25.26	89.99	Moderate
2	39.79	35.93	87.64	NaN
3	39.19	21.52	94.11	Severe
4	37.20	36.35	92.15	NaN

	Fatigue_Level	Age	Pre_Existing_Conditions	WBC_Count	Severity_Index
0	1	28	1	7.52	0.95
1	1	22	0	11.70	0.78
2	1	16	1	5.25	0.57
3	1	13	0	13.12	0.77
4	1	62	0	12.37	0.59

Mean Squared Error (MSE): 0.0173
R-squared (R^2): 0.6630

Fig. 4: Random Forest Regression Reports

The Random Forest Regression model achieved a Mean Squared Error (MSE) of 0.0173 and an R-squared (R^2) value of 0.6630. The MSE of 0.0173 indicates the average squared difference between the predicted severity index values and the actual observed values. A lower MSE value suggests that the predictions made by the model are very close to the actual observed values, reflecting good predictive accuracy. This demonstrates that the Random Forest model can reasonably capture patterns in the data related to the severity index of influenza based on the provided features, leading to reliable predictions. The R^2 value of 0.6630 implies that approximately 66.30% of the variation in the severity index of influenza can be explained by the Random Forest model using the available features. This is a strong indicator that the model has good explanatory power, as it can account for a significant proportion of the variability in the severity index. For the purpose of detecting the severity of influenza as a contagious respiratory complication, these results suggest that the model has a decent ability to predict severity levels, allowing for early detection and intervention. However, the remaining 33.7% of unexplained variation implies that additional factors may influence the severity index, which could be explored in future studies to improve prediction accuracy.

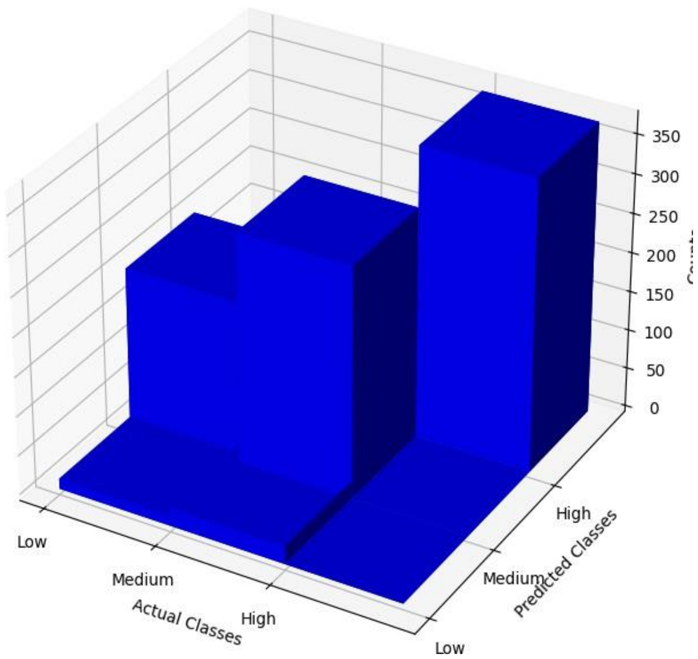

Fig. 5: Influenza Severity Index Categorization

The Influenza Severity Index is categorized into distinct classes or groups based on the predicted severity index values, which help in identifying the level of severity of influenza as a contagious respiratory complication. These groups are determined by defining threshold ranges for the severity index values. For instance, a severity index value close to 0 may indicate a low or asymptomatic case, while a value closer to 1 suggests a high severity case requiring medical attention. The classification allows for a clear understanding of the patient's condition, aiding healthcare professionals and policymakers in making informed decisions about resource allocation, treatment strategies, and intervention measures. Grouping these index values into meaningful severity classes ensures that patterns in disease spread and patient conditions can be identified more effectively, enhancing early detection and response to outbreaks.

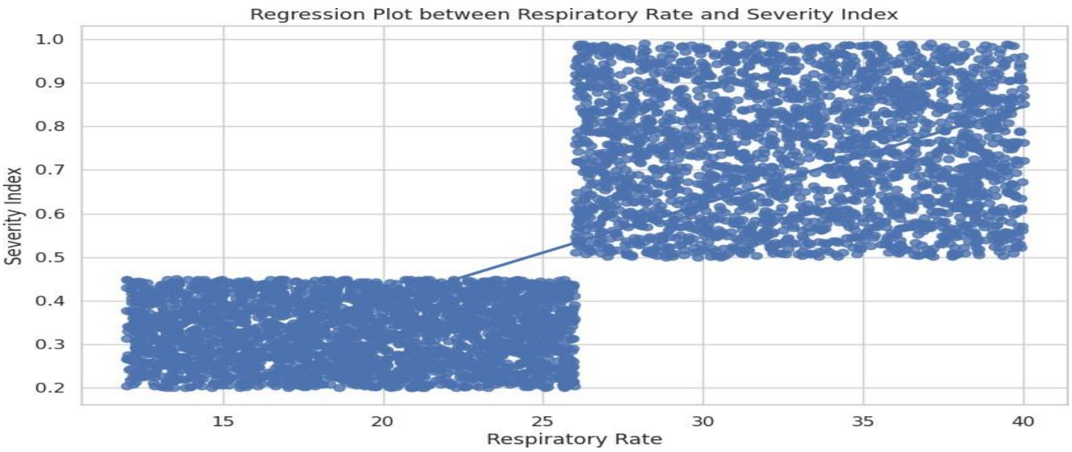

Fig.6: Regression plot of Respiratory Rate and Severity Index

The regression plot of Respiratory Rate and Severity Index shows the relationship between a patient's respiratory rate and the predicted severity index values. This visualization highlights how changes in respiratory rate correlate with variations in severity, indicating that higher respiratory rates are generally associated with higher severity index values. The plot demonstrates the predictive capability of the model, suggesting that respiratory rate is a significant indicator of the severity of influenza as a respiratory complication. This analysis confirms the importance of respiratory rate as a contributing feature in determining the severity index, validating its role in early detection and diagnosis.

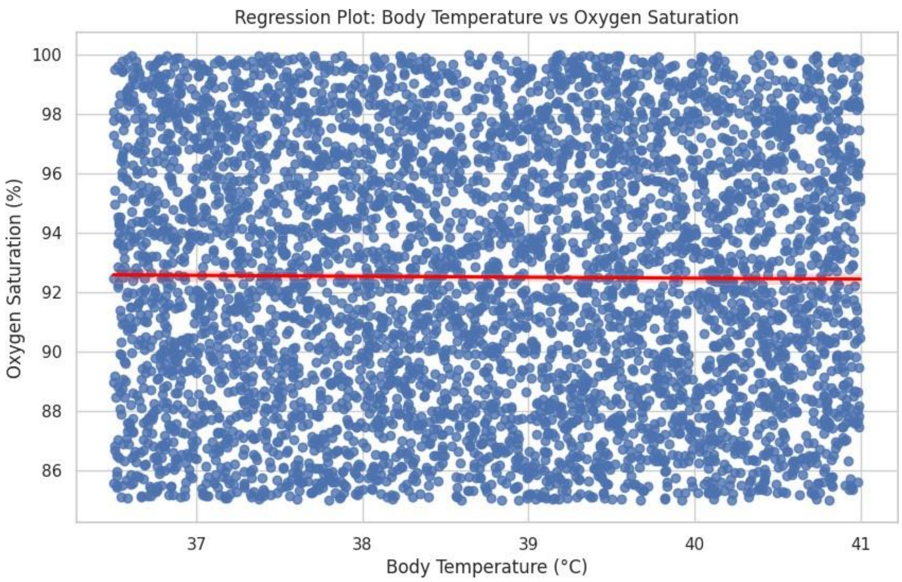

Fig. 7: Effect of Oxygen Saturation and Body Temperature

The effect of Oxygen Saturation and Body Temperature on the severity of influenza was analyzed to understand their contribution to the severity index prediction. The analysis revealed that lower oxygen saturation levels are strongly associated with higher severity index values, indicating that compromised oxygenation is a critical indicator of severe influenza cases. Similarly, body temperature variations were evaluated, and it was observed that higher body temperatures correlate with increased severity, as they may signify the body's immune response to combat the influenza virus. These findings suggest that both oxygen saturation and body temperature are reliable predictors of disease progression. This relationship emphasizes the importance of monitoring these physiological parameters in clinical assessments. Lower oxygen saturation levels could suggest respiratory distress, while elevated body temperature can signal a higher likelihood of severe illness. These insights highlight the predictive value of these features, offering guidance for early detection strategies and prioritization of medical interventions in influenza cases.

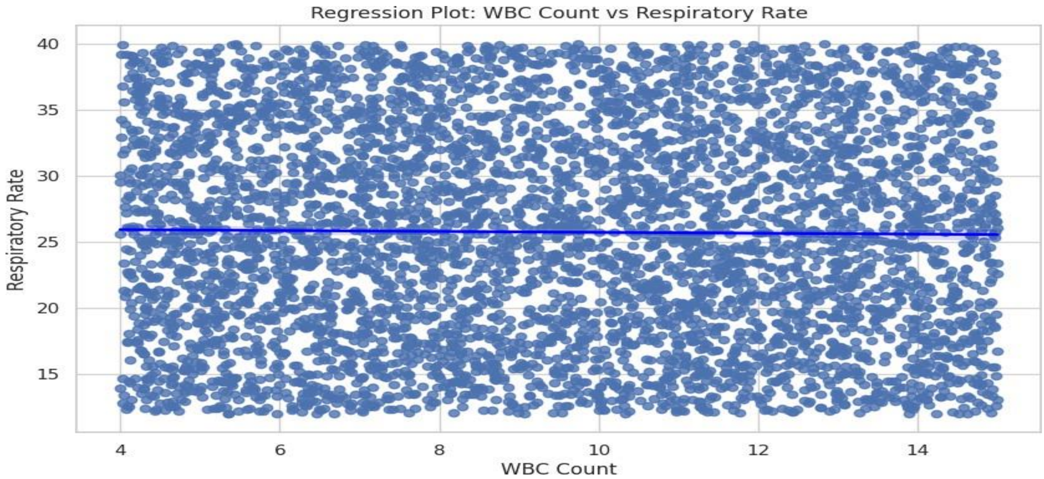

Fig. 8: Effect of WBC Count and Respiratory Rate

The effect of White Blood Cell (WBC) count and Respiratory Rate on the severity of influenza was analyzed to determine their predictive roles in assessing disease progression. The analysis showed that elevated WBC counts are associated with higher severity index values. This suggests that an increased immune response, as reflected by higher WBC levels, indicates the body's attempt to combat the influenza virus, often correlating with more severe cases. Similarly, Respiratory Rate was found to be a significant predictor, with higher respiratory rates corresponding to higher severity index values. This may indicate that rapid breathing reflects the body's physiological response to compromised respiratory function during severe influenza infections. These findings underscore the critical role of both WBC count and Respiratory Rate in evaluating disease severity. Elevated WBC counts suggest an active immune response and potential systemic infection severity, while higher Respiratory Rates point to impaired respiratory capacity. Together, these indicators provide valuable insights into the progression of influenza and can support early diagnosis and intervention strategies. Monitoring these parameters in clinical settings can enhance the accuracy of severity assessments and inform treatment plans for patients with influenza.

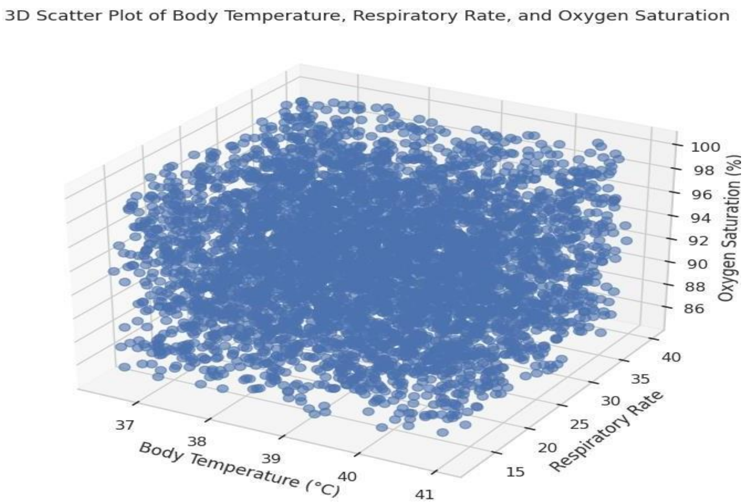

Fig. 9: Body Temperature, Respiratory Rate and Oxygen Saturation

The analysis of Body Temperature, Respiratory Rate, and Oxygen Saturation revealed their significant influence on determining the severity of influenza. Body Temperature is a critical indicator as fever is a common symptom of influenza; higher body temperatures were associated with increased severity index values, reflecting the body's immune response to the infection. Respiratory Rate was another key predictor, with higher rates correlating to more severe cases. This may suggest that rapid breathing is a response to compromised lung function as the body struggles to maintain adequate oxygen levels during severe influenza infections. Oxygen Saturation levels also played a vital role in assessing severity. Lower oxygen saturation values were strongly associated with higher severity index levels, indicating that compromised oxygenation can be a direct marker of the severity of influenza. These findings highlight that a combination of elevated body temperature, increased respiratory rate, and decreased oxygen saturation can provide an early warning system to identify individuals at risk of progressing to severe illness. Monitoring these features can thus support timely clinical interventions and improve patient outcomes by enabling early and appropriate treatment.

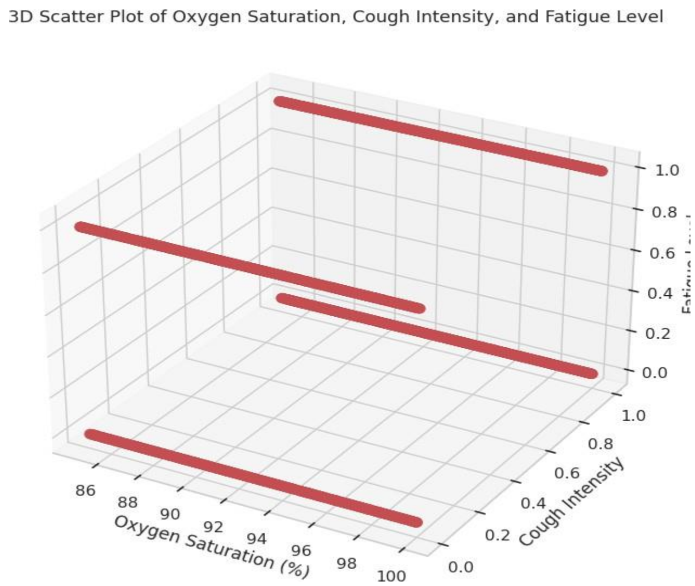

Fig. 10: Cough Intensity, Fatique Level and Oxygen Saturation

The analysis of Cough Intensity, Fatigue Level, and Oxygen Saturation provided insights into their combined effects on the severity of influenza. Cough Intensity was found to have a direct relationship with the severity index, where higher cough intensity levels corresponded to increased severity. This is indicative of the respiratory distress that accompanies more advanced cases of influenza, as persistent or severe coughing can lead to complications and increased strain on the respiratory system. Fatigue Level also demonstrated a meaningful connection with the severity index. Higher fatigue levels were associated with elevated severity index values, highlighting the systemic impact of influenza on an individual's overall well-being. Additionally, Oxygen Saturation was a critical variable, with lower values linked to higher severity, reflecting compromised lung function and oxygen delivery in the body during severe influenza cases. Together, these three features—Cough Intensity, Fatigue Level, and Oxygen Saturation—demonstrated how physiological and symptomatic markers interact to influence the progression of influenza, emphasizing the importance of their early assessment for timely clinical responses.

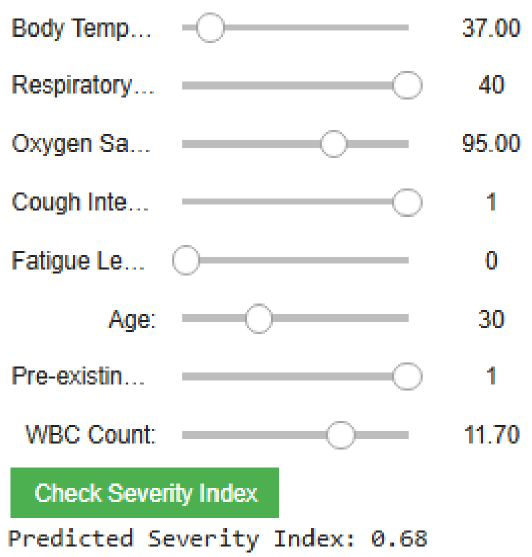

Fig. 11: Program Interface

5. DISCUSSION

This research focused on assessing and predicting the severity of influenza as a contagious respiratory complication by leveraging machine learning techniques, specifically Random Forest regression. The study began with the collection of a comprehensive dataset from the Kaggle data source, which included features such as body temperature, respiratory rate, oxygen saturation, cough intensity, fatigue level, white blood cell (WBC) count, pre-existing conditions, and other vital health indicators. These features were analyzed for their relationship to the severity index, a measure of how severely influenza affects patients. The data provided a foundation for evaluating patterns and relationships among these features to assess their contribution to determining the severity of influenza. Once the dataset was collected, data preprocessing steps were carried out to prepare the features and target variable for machine learning model development.

The target variable was defined as the *severity index*, while the remaining features (e.g., body temperature, respiratory rate, oxygen saturation, cough intensity, fatigue level, and WBC count) were used as independent variables. After splitting the data into training and testing sets, the Random Forest regression model was selected for its ability to handle nonlinear relationships, feature interactions, and its robustness against overfitting. The model was trained with 80% of the data and validated on the remaining 20% to evaluate its predictive performance. The Random Forest regression model was successfully developed and trained, and its performance was evaluated using metrics such as Mean Squared Error (MSE) and R^2. The findings showed an MSE of 0.0173 and an R^2 of 0.6630. These results indicated that the model could predict the severity index with a good degree of accuracy, capturing approximately 66.3% of the variance in the target variable. The discussion of these results revealed that key features such as respiratory rate, oxygen saturation, body temperature, and WBC count had a significant effect on the predicted severity index, providing a meaningful understanding of their contributions to the progression of influenza. The research findings were further analyzed through feature importance evaluation, regression plots, and other visualizations. Feature importance analysis showed how each feature contributed to the prediction of the severity index, allowing insights into which factors were most predictive. Regression plots and scatter visualizations were generated to explore relationships between key features like respiratory rate, body temperature, WBC count, and the severity index. These plots and analyses highlighted that higher respiratory rates and lower oxygen saturation were associated with higher severity, while body temperature and WBC count played critical roles in predicting the progression of symptoms. The meaning of the results lies in their clinical and practical implications. The predictive ability of the Random Forest regression model offers a tool for assessing the severity of influenza using vital signs and health indicators. This can enable early detection and intervention for at-risk individuals. Furthermore, the analysis of important features provides valuable insights into the physiological markers of disease progression, emphasizing the role of respiratory rate, oxygen saturation, and body temperature as potential indicators for monitoring patients with influenza. These findings can guide both clinical decision-making and future research into the biological and clinical factors influencing the severity of respiratory diseases like influenza.

6. CONCLUSION

This research successfully demonstrates the application of machine learning techniques, specifically Random Forest regression, to predict the severity index of influenza based on key physiological indicators. Through careful data collection, preprocessing, and model development using features such as respiratory rate, body temperature, oxygen saturation, and WBC count, the study achieved promising results with an MSE of 0.0173 and an R^2 value of 0.6630. These findings underscore the potential of leveraging physiological parameters to assess and predict the severity of influenza, offering a novel, data-driven approach to understanding the progression of this contagious respiratory disease. The integration of feature importance analysis and visualizations further provided insights into how specific biological markers contribute to predicting the disease's severity. The findings of this research have important practical implications for early diagnosis, intervention, and resource allocation within the medical sector. The model can serve as an effective tool for identifying high-risk individuals, enabling early and accurate detection of severe cases. Additionally, the research highlights the importance of adopting machine learning approaches for disease severity prediction, contributing to improved clinical decision-making and public health strategies. Future studies could explore larger, more diverse datasets and consider integrating other machine learning models to strengthen predictive accuracy. This work represents a significant step toward a proactive, evidence-based response to managing influenza and similar respiratory diseases.

REFERENCES

Centers for Disease Control and Prevention [CDC]. (2020). How COVID-19 Spreads. Retrieved from https://www.cdc.gov.

Cohen, J., Tohme, R. A., & Qin, X. (2017). Health effects of urbanization and social structure on infectious disease transmission. Journal of Infectious Disease Research, 23(3), 102-115.

Dubé, E., Laberge, C., Guay, M., Bramadat, P., Roy, R., & Bettinger, J. A. (2021). Vaccine hesitancy: An overview. Human Vaccines & Immunotherapeutics, 9(8), 1763–1773.

Fauci, A. S., Lane, H. C., & Redfield, R. R. (2020). Covid-19 – Navigating the Uncharted. New England Journal of Medicine, 382(13), 1268–1269.

Grohskopf, L. A., Alyanak, E., Broder, K. R., Blanton, L. H., Fry, A. M., & Jernigan, D. B. (2019). Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices. MMWR Recommendations and Reports, 68(3), 1-21.

Moriyama, M., Hugentobler, W. J., & Iwasaki, A. (2020). Seasonality of respiratory viral infections. Annual Review of Virology, 7(1), 83-101.

Omer, S. B., Betsch, C., & Leask, J. (2019). Vaccination hesitancy and health care providers. Pediatrics, 144(5), e20190345.

Ranney, M. L., Griffeth, V., & Jha, A. K. (2020). Critical supply shortages—The need for ventilators and personal protective equipment during the COVID-19 pandemic. New England Journal of Medicine, 382(18), e41.

Salmon, D. A., Dudley, M. Z., Glanz, J. M., & Omer, S. B. (2015). Vaccine hesitancy: Causes, consequences, and a call to action. American Journal of Preventive Medicine, 49(6), S391-S398.

World Health Organization [WHO]. (2019). Ten threats to global health in 2019. Retrieved from https://www.who.int.

Centers for Disease Control and Prevention [CDC]. (2020). How COVID-19 Spreads. Retrieved from https://www.cdc.gov.

Moriyama, M., Hugentobler, W. J., & Iwasaki, A. (2020). Seasonality of respiratory viral infections. Annual Review of Virology, 7(1), 83-101.

Ranney, M. L., Griffeth, V., & Jha, A. K. (2020). Critical supply shortages—The need for ventilators and personal protective equipment during the COVID-19 pandemic. New England Journal of Medicine, 382(18), e41.

Ekong, A. & Udo, E. & Ekong, O. & Inyang, S. (2023). Machine Learning based Model for the Prediction of Fasting Blood Sugar Level towards Cardiovascular Disease Control for the Enhancement of Public Health. International Journal of Computer Applications. 184. 5-12. 10.5120/ijca2023922622.

Leng, J., Wang, D., Ma, X., Yu, P., Wei, L., & Chen, W. (2022). Bi-level artificial intelligence model for risk classification of acute respiratory diseases based on Chinese clinical data. *Applied Intelligence*, *52*(11), 13114-13131.

Belkacem, A. N., Ouhbi, S., Lakas, A., Benkhelifa, E., & Chen, C. (2021). End-to-end AI-based point-of-care diagnosis system for classifying respiratory illnesses and early detection of COVID-19: a theoretical framework. *Frontiers in Medicine*, *8*, 585578.

Erraguntla, M., Zapletal, J., & Lawley, M. (2019). Framework for Infectious Disease Analysis: A comprehensive and integrative multi-modeling approach to disease prediction and management. *Health informatics journal*, *25*(4), 1170- 1187.

Yang, C. T., Lin, L. Y., Tsan, Y. T., Liu, P. Y., & Chan, W. C. (2019). The implementation of a real-time monitoring and prediction system of PM2. 5 and influenza-like illness using deep learning. *Journal of Internet Technology*, *20*(7), 2237-2245.

Tsai, M. H., Ko, S. C., Huang, A. H., Porta, L., Ferretti, C., Longhi, C., ... & Lee, C. C. (2024).

Predicting mortality in hospitalized influenza patients: integration of deep learning-based chest X-ray severity score (FluDeep-XR) and clinical variables. *Journal of the American Medical Informatics Association*, ocae286.

Kumar, Y., Kanna, G. P., Kumar, S. J., & Sambasivam, G. (2023, September). Influenza Flu Diagnostics and Detection using Artificial Intelligence-based Learning Approaches: Challenges and Recent Study. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 6, pp. 385-390). IEEE.

Ijaz, A., Nabeel, M., Masood, U., Mahmood, T., Hashmi, M. S., Posokhova, I., ... & Imran, A. (2022). Towards using cough for respiratory disease diagnosis by leveraging Artificial Intelligence: A survey. *Informatics in Medicine Unlocked*, *29*, 100832.

Andrade-Arenas, L., Molina-Velarde, P., & Yactayo-Arias, C. (2024). Preliminary diagnosis of respiratory diseases: an innovative approach using a web expert system. *International Journal of Electrical & Computer Engineering (2088- 8708)*, *14*(6).

Yi, R., Tang, L., Tian, Y., Liu, J., & Wu, Z. (2023). Identification and classification of pneumonia disease using a deep learning-based intelligent computational framework. *Neural Computing and Applications*, *35*(20), 14473-14486.

Jiang, X., Coffee, M., Bari, A., Wang, J., Jiang, X., Huang, J., ... & Huang, Y. (2020). Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. *Computers, Materials & Continua*, *63*(1), 537- 551.

Ekong, A., Ekong, B., & Edet, A. (2022). Supervised machine learning model for effective classification of patients with covid-19 symptoms based on bayesian belief network. *Researchers Journal of Science and Technology*, *2*(1), 27- 33.

Ekong, B., Ekong, O., Silas, A., Edet, A. E., & William, B. (2023). Machine Learning Approach for Classification of Sickle Cell Anemia in Teenagers Based on Bayesian Network. *Journal of Information Systems and Informatics*, *5*(4), 1793-1808.

Edet, A., Ekong, B., & Attih, I. (2024). Machine Learning Enabled System for Health Impact Assessment of Soft Drink Consumption Using Ensemble Learning Technique. *International Journal Of Computer Science And Mathematical Theory*, *10*(1), 79-101.

Ekong, B., Edet, A., Udonna, U., Uwah, A., & Udoetor, N. (2024). Machine Learning Model for Adverse Drug Reaction Detection Based on Naive Bayes and XGBoost Algorithm. *British Journal of Computer, Networking and Information Technology*, *7*(2), 97-114.

Edet, A., Silas, A., Ekaetor, E., Ebong, O., Isaac, E., & Udoetor, N. (2024). Data-Driven Framework for Classification and Management of Start-Up Risk For High Investment Returns. *Advanced Journal of Science, Technology and Engineering*, *4*(2), 81-102.

Eyo, V. A., , Nyaks C., and Ememe, N. C. (2025). International Journal of Social Sciences and Management Research, 11, (2), 333-350.