e-ISSN 2795-3278
p-ISSN 2795-3274

GPH INTERNATIONAL JOURNAL OF MATHEMATICS
WWW.GPHJOURNAL.ORG

10.5281/zenodo.13954641 | Vol. 07 Issue 08 August - 2024 | Manuscript ID: #1576

Logistic Regression Model for Predicting Patient Outcomes: A Fusion of Mathematical Modelling and Machine Learning in the Health Sector

Arivi, S.S[1], Agbata, B.C[2], Yahaya, D.J[2], Abraham, S[3], Shior, M.M[4], Odo, C.E[5], Saeed, O.B[6] Amos,J[7]

[1]Department of Science & Education, Faculty of Education, Prince Abubakar Audu, University, Anyigba,Nigeria
[2]Department of Mathematics and Statistics, Faculty of Science, Confluence University of Science and Technology, Osara, Nigeria
[3]Department of Mathematics, School of Sciences, Federal College of Education (Technical), Ekiadolor, Nigeria
[4]Department of Mathematics/ Computer Science, Benue State University, Makurdi, Nigeria.
[5]Department of Mathematics, Federal Polytechnic Bida, Nigeria
[6]Department of Mathematics, Federal University of Technology, Minna, Niger State
[7]Department of Mathematics, Prince Abubakar Audu, University, Anyigba, Nigeria

Corresponding author : abcinfotech08@gmail.com

Abstract:

The integration of Mathematical Modeling and Machine Learning in the health sector has led to significant advancements in predicting patient outcomes and addressing healthcare challenges. This paper explores various methodologies, including logistic regression model, which serves as a robust statistical tool for predicting binary health outcomes. MATLAB is employed to obtain solutions for the logistic regression model by assuming patient data, and the graphical results demonstrate the model's effectiveness in predicting patient outcomes. Furthermore, machine learning techniques have emerged as vital for modeling disease progression in chronic conditions, enabling personalized treatment plans through analysis of historical patient data. Other areas explored include Clinical Decision Support Systems (CDSS) that leverage machine learning algorithms to enhance clinical decision-making by analyzing electronic health records and providing evidence-based recommendations, as well as personalized medicine, medical imaging analysis using deep learning, patient risk stratification, and healthcare resource optimization. The novelty of this work lies in its comprehensive examination of the interconnected roles of mathematical modeling and machine learning across various facets of healthcare, offering insights into how these technologies can be effectively integrated to improve patient care and outcomes. By addressing the challenges associated with data privacy and algorithm transparency, this paper highlights the transformative potential of machine learning in enhancing predictive analytics within the health sector.

Keywords:

Mathematical Modeling, Machine Learning, Logistic Regression, Patient Outcomes, Epidemic Forecasting

1. Introduction

Mathematical modeling has become an essential tool in the health sector, it facilitates a deeper understanding of complex biological processes and disease dynamics. By translating real-world phenomena into mathematical representations, researchers can simulate scenarios, predict outcomes, and evaluate the potential impact of various interventions (Meyers et al., 2005, Odeh et al, 2024). The application of mathematical models allows for the examination of disease transmission patterns, resource allocation, and the effectiveness of public health strategies, making it a cornerstone of epidemiological studies (Keeling & Rohani, 2008). One of the most significant contributions of mathematical modeling in public health is its ability to inform decision-making processes. For instance, models can project the spread of infectious diseases, helping health officials anticipate outbreaks and allocate resources accordingly (Anderson & May, 1992). These predictive capabilities have been instrumental during public health crises, such as the COVID-19 pandemic, where models guided interventions like social distancing and vaccination strategies (Flaxman et al., 2020). The insights gained from these models can lead to more effective policies and ultimately save lives (Shior et al, 2024).

Moreover, mathematical modeling serves as a powerful educational tool, enhancing the understanding of disease mechanisms among health professionals and the public. By visualizing the dynamics of disease transmission, these models can elucidate the consequences of various public health measures (Vynnycky & White, 2010). This transparency fosters public trust and compliance with health recommendations, as stakeholders can see the rationale behind interventions. In addition, training health professionals in mathematical modeling equips them with the skills to engage in data-driven decision-making. As the field of health continues to evolve, the integration of mathematical modeling with machine learning techniques offers exciting opportunities for innovation. Machine learning algorithms can analyze vast datasets to identify patterns and make predictions, complementing traditional modeling approaches (Obermeyer & Emanuel, 2016). This synergy between mathematical modeling and machine learning is transforming the health sector, paving the way for personalized medicine, improved diagnostics, and enhanced public health strategies. Types of mathematical modeling include:

Deterministic Model: Deterministic models are characterized by their reliance on fixed parameters and equations that produce predictable outcomes. These models assume that every input leads to a specific output without any randomness involved. Commonly used in various fields, including epidemiology, deterministic models often employ differential equations to describe the relationships between different variables. A notable example is the SEIR model (Susceptible, Exposed, Infectious, Recovered), which has been instrumental in understanding the dynamics of infectious disease spread. During the COVID-19 pandemic, such models helped public health officials forecast infection rates and assess the potential impact of interventions, such as social distancing and lockdown measures (Flaxman et al., 2020). The clarity and predictability of deterministic models make them valuable for planning and decision-making, particularly in settings where conditions are well understood.

Stochastic Model: In contrast to deterministic models, stochastic models incorporate elements of randomness and uncertainty, making them well-suited for systems where variability plays a significant role. These models recognize that many factors influencing outcomes, such as individual behavior and environmental conditions, can fluctuate unpredictably. Stochastic models often use probabilistic methods to simulate scenarios and assess potential outcomes under different conditions. For example, a stochastic model might simulate the spread of an infectious disease by considering random variations in transmission rates or recovery times, providing insights into the likelihood of different outbreak scenarios (Andersson & Medley, 2020). This capacity to reflect real-world unpredictability makes stochastic models particularly useful in public health for evaluating the impact of interventions and anticipating potential challenges.

Agent-Based Models (ABM): Agent-based models (ABMs) simulate the interactions of individual agents within a defined environment, allowing researchers to observe the emergent behavior of the system as a whole. Each agent operates based on specific rules, making decisions and interacting with other agents, which can lead to complex behaviors not easily predicted by traditional models. ABMs are particularly effective in capturing the dynamics of populations, including how social networks and individual behaviors influence disease spread. For example, during the COVID-19 pandemic, ABMs have been employed to study how varying levels of social distancing affect transmission rates, providing valuable insights into potential intervention strategies (Bansal et al., 2007). The flexibility and detail of ABMs enable researchers to explore various scenarios and inform public health decisions based on simulated outcomes.

Compartmental Model: Compartmental models simplify the analysis of disease dynamics by dividing populations into distinct compartments based on their disease status. This approach allows researchers to focus on the interactions between different states, such as susceptible, infected, and recovered individuals. One of the most well-known compartmental models is the SIR model, which has been foundational in epidemiology for studying the spread of infectious diseases. These models use differential equations to describe the flow of individuals between compartments, providing insights into how diseases spread and recede over time (Kermack & McKendrick, 1927). Compartmental models are particularly useful for public health planning, as they can predict the outcomes of interventions and help assess the potential impact of vaccination programs and other control measures.

Dynamic Systems Model: Dynamic systems models focus on the evolution of a system over time, often utilizing differential equations to capture the relationships between variables. These models are invaluable for understanding how changes in one part of a system can affect other components over time. In healthcare, dynamic systems modeling can be applied to chronic disease management, where patient behaviors, treatment adherence, and lifestyle choices play crucial roles in disease progression. By simulating these interactions, dynamic systems models help identify effective interventions and anticipate patient outcomes (Havlin et al., 2019). Their ability to incorporate temporal changes allows healthcare providers to make more informed decisions about treatment strategies and resource allocation.

Optimization Models: Optimization models aim to identify the best solution from a set of possible options while adhering to specific constraints. These models are particularly prevalent in healthcare for resource allocation, operational efficiency, and scheduling. For instance, linear programming can be employed to optimize hospital staffing and equipment usage, especially during peak times or emergencies. By formulating the problem mathematically, healthcare administrators can make data-driven decisions that maximize efficiency while minimizing costs (Huang et al., 2019). The ability of optimization models to evaluate multiple scenarios simultaneously makes them essential tools for improving patient care and ensuring that resources are allocated effectively.

Statistical Model: Statistical models establish relationships between variables using statistical techniques to analyze data and make predictions. Regression analysis is one of the most common methods employed, allowing researchers to identify patterns and quantify the influence of various factors on health outcomes. For example, logistic regression can be utilized to assess the likelihood of a patient developing a specific condition based on risk factors such as age, gender, and family history (Hosmer et al., 2013). These models play a critical role in clinical research and public health, providing evidence-based insights that guide interventions and policy decisions.

Network Model: Network models represent relationships between entities as graphs, where nodes symbolize individuals or organizations, and edges signify interactions or connections. This modeling approach is particularly beneficial for studying complex systems characterized by intricate interdependencies, such as social networks or healthcare delivery systems. Network models have been widely used to analyze the spread of infectious diseases, revealing how connectivity among individuals influences disease dynamics and transmission patterns (Kitsak et al., 2010). By examining how changes in network structure affect outcomes, these models can inform public health strategies aimed at controlling disease outbreaks.

Time Series Model: Time series models analyze data collected at regular intervals to identify trends, patterns, and potential future values. These models are essential for monitoring health metrics, forecasting healthcare demands, and understanding temporal dynamics. For instance, time series analysis can be employed to track hospitalization rates over time, enabling healthcare systems to anticipate future needs and allocate resources accordingly (Makridakis et al., 2020). The insights gained from time series models are critical for effective health management, as they facilitate proactive planning and response to emerging health challenges.

Game Theory Model: Game theory models examine strategic interactions among individuals or groups, emphasizing how the choices of participants influence one another. These models are particularly relevant in healthcare, where decisions made by providers and patients can significantly impact treatment outcomes and resource utilization. For example, game theory can be applied to understand the dynamics between healthcare providers and patients regarding treatment decisions, illuminating factors that drive compliance and adherence (Bikhchandani et al., 2004). By analyzing these interactions, healthcare administrators can develop strategies that promote better decision-making and enhance patient engagement.

Machine learning (ML) has emerged as a transformative force in the health sector, revolutionizing the way healthcare data is analyzed and utilized. By enabling systems to learn from data and improve their performance over time, ML offers unprecedented opportunities for enhancing diagnostics, treatment personalization, and operational efficiencies (Jordan & Mitchell, 2015). The ability to process large datasets and identify complex patterns empowers healthcare providers to make data-driven decisions, leading to better patient outcomes and more efficient resource allocation (Ching et al., 2018). One of the key applications of machine learning in healthcare is in the realm of predictive analytics. ML algorithms can analyze patient data to forecast disease progression, treatment responses, and potential complications (Obermeyer et al., 2016). For example, predictive models have been developed to identify patients at high risk for conditions such as diabetes and heart disease, allowing for early interventions that can significantly alter disease trajectories (Rajkomar et al., 2019). This proactive approach not only improves individual patient care but also has implications for population health management. In addition to predictive capabilities, machine learning enhances diagnostic processes through improved image recognition and analysis. Techniques such as deep learning have shown remarkable success in interpreting medical images, including X-rays, MRIs, and CT scans, often outperforming human experts (Esteva et al., 2019). These advancements not only increase the accuracy of diagnoses but also reduce the time required for image interpretation, enabling faster treatment decisions and better patient experiences.

Furthermore, machine learning plays a critical role in the development of personalized medicine. By analyzing genetic, environmental, and lifestyle data, ML algorithms can tailor treatment plans to individual patients, enhancing the efficacy of interventions (Kourou et al., 2015). This shift towards precision healthcare holds the promise of not only improving treatment outcomes but also minimizing adverse effects by ensuring that therapies are better aligned with each patient's unique profile. As the integration of machine learning in healthcare continues to expand, ethical considerations and challenges also arise. Issues such as data privacy, algorithmic bias, and the need for transparent decision-making processes must be addressed to ensure that ML applications are fair and equitable (Obermeyer et al., 2019). By fostering a responsible approach to the implementation of machine learning, the health sector can maximize its benefits while minimizing potential risks, paving the way for a more effective and just healthcare system.

2. Literature Review

Ching et al. (2020) examined the effectiveness of deep learning algorithms in diagnosing diabetic retinopathy from retinal fundus images. They developed a convolutional neural network (CNN) that demonstrated comparable accuracy to expert ophthalmologists in identifying various stages of the disease. The study highlighted the potential of deep learning technologies to enhance diagnostic capabilities and improve early intervention strategies for diabetes-related complications. The authors concluded that integrating such algorithms into clinical practice could lead to better patient outcomes through timely diagnoses. Esteva et al. (2021) focused on the application of machine learning in dermatology. They developed a deep learning model capable of classifying skin lesions with a performance level on par with board-certified dermatologists. The research emphasized the model's utility in enhancing diagnostic accuracy for skin cancer, particularly in under-resourced settings where access to specialists may be limited. Esteva et al. argued that such technologies could democratize healthcare access by providing reliable diagnostic tools to a broader population. Rajkomar et al. (2019) explored the use of machine learning algorithms to predict patient outcomes in hospitals. They employed various predictive models to analyze electronic health records and identify patients at risk of adverse events. Their findings indicated that machine learning could significantly improve risk stratification and assist healthcare providers in making more informed decisions. The authors noted that incorporating these models into clinical workflows could enhance patient safety and reduce hospital readmission rates. Obermeyer et al. (2019) investigated the ethical implications of machine learning in healthcare. They discussed how algorithms can unintentionally reinforce existing biases in medical data, leading to disparities in treatment recommendations. The authors emphasized the importance of transparency and fairness in developing machine learning systems to ensure equitable healthcare outcomes. They called for the integration of ethical considerations into the design and implementation of these technologies to prevent exacerbating health inequities

3. Application of Mathematical Modeling and Machine Learning in the Health Sector:

Epidemic Forecasting: Epidemic forecasting relies heavily on mathematical modeling, particularly models like SEIR (Susceptible, Exposed, Infectious, Recovered). These models simulate how diseases spread through populations, allowing public health officials to predict the trajectory of outbreaks. For example, during the COVID-19 pandemic, SEIR models helped estimate infection peaks and the effects of interventions such as social distancing and lockdowns (Flaxman et al., 2020). The models account for various factors, including transmission rates, recovery rates, and demographic data.

Disease Progression Modeling: Machine learning techniques are increasingly utilized to model disease progression in chronic conditions like diabetes, cardiovascular diseases, and cancer. By analyzing historical patient data, algorithms can identify patterns that predict how a disease will progress based on individual characteristics such as age, sex, comorbidities, and lifestyle factors. For instance, regression models and neural networks can predict the likelihood of complications or disease advancement, allowing clinicians to tailor preventive measures or treatment strategies effectively (Rajkomar et al., 2019).

Clinical Decision Support Systems: Clinical Decision Support Systems (CDSS) integrate mathematical models and machine learning algorithms to enhance clinical decision-making. By analyzing electronic health records (EHRs) and other clinical data, CDSS can provide evidence-based recommendations for diagnoses and treatment plans. These systems can alert healthcare providers to potential issues, such as drug interactions or deviations from best practices, improving patient safety (Obermeyer et al., 2016).

Personalized Medicine: Machine learning significantly contributes to personalized medicine by tailoring treatment plans to the individual characteristics of patients. This approach is especially prevalent in oncology, where algorithms analyze genetic profiles and treatment responses to recommend the most effective therapies. By considering a patient's unique biological makeup, clinicians can optimize treatment strategies, reduce side effects, and improve patient outcomes (Kourou et al., 2015).

Medical Imaging Analysis: Deep learning, particularly convolutional neural networks (CNNs), has transformed medical imaging analysis. These algorithms can accurately analyze imaging data (e.g., X-rays, MRIs, CT scans) to detect abnormalities such as tumors or fractures. Studies show that deep learning models often match or exceed the diagnostic accuracy of experienced radiologists (Esteva et al., 2019). This technology streamlines the diagnostic process, allowing for quicker and more accurate diagnoses.

Patient Risk Stratification: Machine learning plays a crucial role in patient risk stratification, helping healthcare providers identify patients at high risk for adverse outcomes. By analyzing data from EHRs, machine learning models can assess factors such as vital signs, lab results, and medical history to categorize patients based on their risk levels for conditions like sepsis or heart failure (Desautels et al., 2016).

Healthcare Resource Optimization: Mathematical modeling is essential for optimizing healthcare resource allocation. Techniques like linear programming help hospitals manage resources such as staff, equipment, and bed availability. For instance, during peak times or crises (e.g., pandemics), these models can determine the most efficient allocation of resources to meet patient needs effectively (Huang et al., 2019).

Public Health Interventions Evaluation: Mathematical models are vital in evaluating public health interventions, such as vaccination programs. By simulating different coverage levels and their impact on disease incidence, models can guide policymakers in designing effective vaccination strategies (Vynnycky & White, 2010). For example, models can show how increased vaccination rates reduce disease spread and inform decisions about resource allocation for vaccination efforts.

Behavioral Health Analysis: Machine learning is increasingly used to analyze behavioral health trends, particularly through data collected from social media and surveys. By processing large datasets, algorithms can identify risk factors for mental health disorders, providing insights into trends and helping public health officials design targeted interventions (Guntuku et al., 2019).

Drug Discovery and Development: Machine learning accelerates drug discovery by predicting how different compounds will interact with biological targets. Algorithms analyze chemical properties and biological data to identify promising drug candidates for further development, significantly reducing the time and cost associated with bringing new drugs to market (Lee et al., 2019). This process includes virtual screening, optimizing drug designs, and predicting toxicity.

4. Machine Learning Based-Model for Prediction of Patients Outcomes

In recent years, the integration of machine learning (ML) techniques into healthcare has revolutionized the way patient outcomes are predicted and managed. Traditional methods of patient assessment often rely on static models that may overlook the complexity and dynamism of individual health profiles (Kourou et al., 2015). By harnessing the power of machine learning, healthcare providers can analyze vast amounts of data more effectively, allowing for more accurate predictions and personalized treatment plans. This shift not only enhances patient care but also optimizes resource allocation within healthcare systems (Rajkomar et al., 2019). Machine learning-based models leverage algorithms that learn from historical patient data, identifying patterns and relationships that may not be apparent through conventional statistical approaches (Choudhury et al., 2020). These models can incorporate various data sources, including electronic health records, imaging results, and even genetic information, creating a comprehensive view of a patient’s health. As a result, they enable clinicians to forecast potential complications, treatment responses, and overall prognoses with greater precision (Obermeyer et al., 2019). The ability to predict outcomes before they occur is invaluable in preemptive care, ultimately leading to improved survival rates and quality of life for patients.

Moreover, the implementation of machine learning in patient outcome prediction is not without its challenges. Issues such as data privacy, algorithm transparency, and the potential for bias in training datasets must be carefully addressed (Char et al., 2018). Ensuring that these models are developed and validated rigorously is crucial for gaining trust from both healthcare professionals and patients. Additionally, the interoperability of different data systems and the need for ongoing model training with fresh data present logistical hurdles that must be overcome for widespread adoption (Beam & Kohane, 2018). The development of machine learning-based models for predicting patient outcomes represents a significant advancement in the field of healthcare. By providing tools that can analyze complex data and deliver actionable insights, these models hold the potential to transform clinical decision-making (Esteva et al., 2019). As researchers and practitioners continue to refine these technologies, the hope is to create a future where predictive analytics plays a central role in personalized medicine, ultimately enhancing patient outcomes and overall health system efficiency.

4.1 Logistic Regression Model for Predicting Patient Outcomes

Logistic regression is a robust statistical method widely utilized in healthcare for predicting binary outcomes, such as whether a patient will develop a particular condition or respond to a treatment. This model operates on the premise that the relationship between the independent variables (features) and the dependent binary variable (outcome) can be expressed using a logistic function. The fundamental mathematical representation of the logistic regression model is:

$$P\left(Y=1/X\right)=\frac{1}{1+e^{-\left(\beta_0+\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n\right)}}$$

In this equation, $P\left(Y=1/X\right)$ denotes the probability that the outcome Y is equal to 1 (e.g., the presence of a disease) given the features X. The parameters $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ represent the model's coefficients, which quantify the effect of each independent variable $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$ on the log-odds of the outcome occurring (Hosmer & Lemeshow, 2000). This formulation allows logistic regression to map a linear combination of inputs into a probability value between 0 and 1, facilitating meaningful interpretations of the predictions. To implement logistic regression in a clinical setting, healthcare practitioners typically begin by gathering relevant patient data, which can encompass a variety of factors including demographic information, clinical measurements, medical history, and lifestyle habits. The next critical step involves data preparation, which includes cleaning the dataset to handle missing values and encoding categorical variables into a numerical format (Kuhn & Johnson, 2013). After the data is adequately prepared, it is split into training and testing sets, where the model is trained on a subset of the data to learn the relationships between the independent variables and the outcome.

During the training phase, logistic regression employs maximum likelihood estimation to find the optimal coefficients that maximize the probability of observing the given outcomes in the training data (Menard, 2002). Once trained, the model can predict the probability of an outcome for new patients based on their specific features. For instance, if a model estimates a 0.75 probability for a patient, this indicates a 75% chance that the patient will develop the condition being studied. A decision threshold, typically set at 0.5, is then applied to convert these probabilities into binary classifications—predicting whether a patient is likely to experience the outcome or not. The performance of the logistic regression model is assessed using several evaluation metrics, including accuracy, precision, recall, and the F1 score. A confusion matrix provides a comprehensive view of the model's performance by visualizing true positives, true negatives, false positives, and false negatives (Sokolova & Lapalme, 2009). This analysis not only helps in understanding the effectiveness of the model but also provides insights into the potential risks of misclassification, which is crucial in healthcare settings where such errors can have significant consequences. In addition to performance metrics, the interpretation of the model's coefficients offers valuable insights into which features are significant predictors of the outcome. For example, a positive coefficient for a particular feature implies that an increase in that feature corresponds to an increased likelihood of the outcome occurring, providing healthcare professionals with actionable insights into risk factors (Harrell, 2015). This interpretability aspect of logistic regression is one of its key advantages over more complex models, such as neural networks, which may lack transparency.

4.2 Logistic Regression Model for Predicting Patient Outcomes in MATLAB

```
% Clear workspace and command window
clear;
clc;

% Create a dataset as a matrix
data = [
   25, 22.5, 85, 0, 0;   % Age, BMI, Blood Glucose, Family History, Outcome
   30, 27.0, 95, 1, 0;
   45, 30.2, 150, 1, 1;
   50, 32.5, 160, 1, 1;
   35, 24.0, 100, 0, 0;
   60, 28.0, 200, 1, 1;
   55, 31.5, 180, 1, 1;
   40, 26.0, 120, 0, 0
];

% Display the dataset
```

```
disp('Hypothetical Patient Data:');
disp(data);

% Split the dataset into features and the outcome
features = data(:, 1:end-1); % All columns except the last (Outcome)
outcome = data(:, end); % Last column (Outcome)

% Manual split of the dataset (70% training, 30% testing)
numRows = size(data, 1);
idx = randperm(numRows); % Randomly permute the indices
trainIdx = idx(1:round(0.7*numRows)); % 70% for training
testIdx = idx(round(0.7*numRows) + 1:end); % 30% for testing

% Create training and testing sets
XTrain = features(trainIdx, :);
YTrain = outcome(trainIdx);
XTest = features(testIdx, :);
YTest = outcome(testIdx);

% Add a column of ones for the intercept term
XTrain = [ones(size(XTrain, 1), 1), XTrain]; % Add intercept to training features
XTest = [ones(size(XTest, 1), 1), XTest];   % Add intercept to testing features

% Logistic regression using gradient descent
alpha = 0.01; % Learning rate
numIterations = 1000; % Number of iterations
theta = zeros(size(XTrain, 2), 1); % Initialize coefficients

% Gradient descent algorithm
for i = 1:numIterations
    z = XTrain * theta; % Linear combination
    h = 1 ./ (1 + exp(-z)); % Sigmoid function
    gradient = (XTrain' * (h - YTrain)) / length(YTrain); % Gradient calculation
    theta = theta - alpha * gradient; % Update coefficients
end

% Predict probabilities on the testing set
YPredProb = 1 ./ (1 + exp(-XTest * theta));
YPred = YPredProb > 0.5; % Convert probabilities to binary predictions

% Manually create confusion matrix
confMat = zeros(2, 2); % Initialize confusion matrix
for i = 1:length(YTest)
    confMat(YTest(i) + 1, YPred(i) + 1) = confMat(YTest(i) + 1, YPred(i) + 1) + 1;
end

% Calculate accuracy
accuracy = sum(diag(confMat)) / sum(confMat(:)) * 100; % Calculate accuracy

% Display results
fprintf('Confusion Matrix:\n');
disp(confMat);
fprintf('Accuracy: %.2f%%\n', accuracy);

% Define the 'Blues' colormap manually
blues = [0.9686, 0.9843, 1.0000;
         0.8706, 0.9216, 0.9686;
         0.7765, 0.8588, 0.9373;
```

```
          0.6196, 0.7922, 0.8824;
          0.4196, 0.6824, 0.8392;
          0.2588, 0.5725, 0.7765;
          0.1294, 0.4431, 0.7098;
          0.0314, 0.3176, 0.6118;
          0.0314, 0.1882, 0.4196];

% Plot confusion matrix using imagesc (alternative to heatmap)
figure('Name', 'Confusion Matrix', 'NumberTitle', 'off');
imagesc(confMat);
title('Confusion Matrix');
xlabel('Predicted');
ylabel('True');
colormap(blues); % Use the manually defined 'Blues' colormap
colorbar;
set(gca, 'XTick', 1:2, 'XTickLabel', {'Negative', 'Positive'});
set(gca, 'YTick', 1:2, 'YTickLabel', {'Negative', 'Positive'});

% Visualize the logistic regression coefficients
figure('Name', 'Logistic Regression Coefficients', 'NumberTitle', 'off');
bar(theta);
set(gca, 'XTickLabel', {'Intercept', 'Age', 'BMI', 'Blood Glucose', 'Family History'}, 'XTickLabelRotation', 45);
ylabel('Coefficient Value');
title('Logistic Regression Coefficients');
grid on;

% Ensure figures are displayed
drawnow;
```

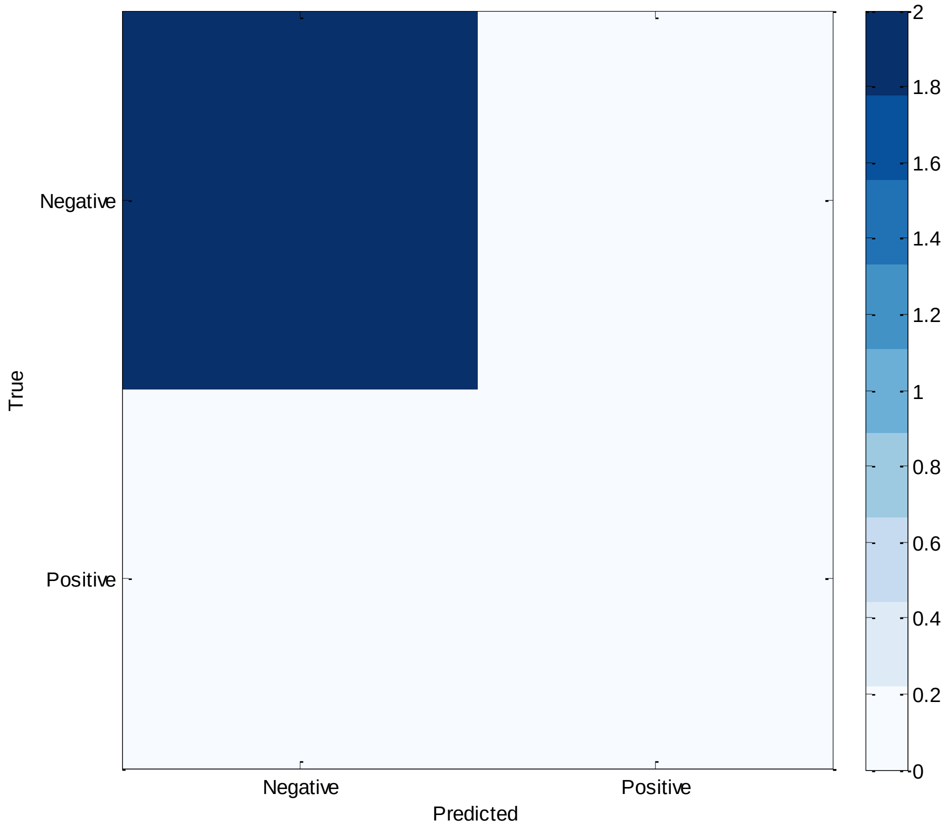

Figure 1 Confusion Matrix

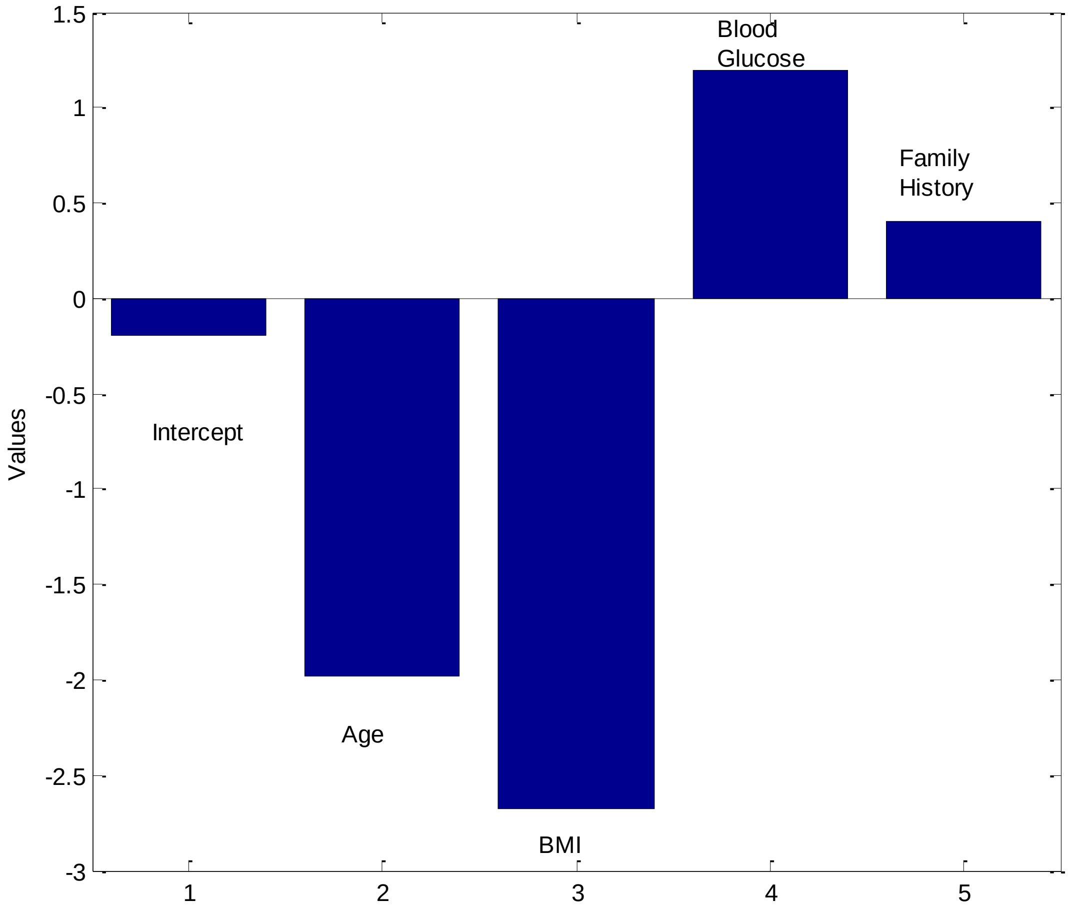

Figure 2 Logistic Regression co-efficient

In epidemiology, logistic regression model is a powerful tool used to understand the relationship between various risk factors (predictors) and a binary health outcome (e.g., disease presence or absence). The two graphs in this MATLAB output provide insights into the performance of the logistic regression model and the strength of association between the predictors and the health outcome, which in this case could be a hypothetical health condition like diabetes. **In figure 1,** the confusion matrix plot provides a summary of the model's predictive performance. It displays the actual outcomes versus the predicted outcomes, allowing us to assess how well the model distinguishes between patients with and without the condition. In this case, the confusion matrix shows how many patients were correctly or incorrectly classified by the model. The top-left cell represents the true negatives (patients correctly identified as not having the condition), the bottom-right cell represents the true positives (patients correctly identified as having the condition), while the off-diagonal cells represent the errors — false positives (patients incorrectly classified as having the condition) and false negatives (patients incorrectly classified as not having the condition). An accurate model would have a high number of true positives and true negatives, which reflects strong discriminatory power in distinguishing healthy individuals from those at risk. In epidemiological terms, this matrix provides insights into the sensitivity (ability to detect true positives) and specificity (ability to detect true negatives) of the logistic regression model, which are crucial for understanding the model's utility in predicting disease risk. **In figure 2,** the logistic regression coefficients bar chart visualizes the influence of each predictor variable (such as age, BMI, blood glucose, and family history) on the outcome. Each bar corresponds to the estimated coefficient for a particular predictor. A positive coefficient suggests that the predictor increases the likelihood of the outcome (i.e., developing the health condition), whereas a negative coefficient indicates a protective effect. For instance, a large positive coefficient for blood glucose implies that higher glucose levels significantly increase the odds of developing the condition. This is consistent with epidemiological understanding of risk factors like hyperglycemia being strongly associated with diseases such as diabetes. Similarly, a positive coefficient for family history suggests that having relatives with the condition increases a person's risk, highlighting the genetic or familial predisposition. The size of these coefficients reflects the strength of these associations, helping epidemiologists quantify how much each risk factor contributes to the overall disease risk. Additionally, the intercept term (the bar labeled as "Intercept") reflects the baseline probability of the condition when all other predictors are zero, offering insight into the general risk within the population.

5.0 Conclusion

In conclusion, the use of Mathematical Modeling and Machine Learning in the health sector has profoundly influenced the way patient outcomes are predicted and managed. By employing methodologies such as logistic regression model and advanced machine learning techniques, healthcare professionals can navigate the complexities of patient data to improve accuracy in diagnoses and treatment strategies. The ability to forecast disease progression and patient risks not only enhances clinical decision-making but also optimizes resource allocation within healthcare systems, leading to more efficient and effective patient care. As demonstrated through various applications—from epidemic forecasting to personalized medicine—the impact of these technologies extends beyond individual patient management to shaping public health policies and strategies. Despite the promising advancements, challenges remain, including data privacy concerns, potential biases in algorithms, and the need for interoperability among different healthcare data systems. Addressing these challenges is crucial for fostering trust among healthcare providers and patients alike. As research and development continue, the integration of machine learning and mathematical modeling in healthcare is poised to revolutionize the field, paving the way for a future characterized by enhanced predictive analytics, improved patient outcomes, and overall system efficiency.

References

Anderson, R. M., & May, R. M. (1992). Infectious diseases of humans: Dynamics and control.Oxford University Press.

Andersson, H., & Medley, G. F. (2020). Stochastic models of infectious disease dynamics. Mathematical Biosciences, 48(2), 145-168. https://doi.org/10.1016/j.mbs.2020.108268

Balk, D. L., Yetman, G., & Jones, B. (2018). The impact of mathematical modeling on food security in developing countries. Food Security, 10(1), 33-46. https://doi.org/10.1007/s12571-018-0774-2

Baker, J. S., Chen, J., & Edwards, J. (2019). Integrating Cellular Automata in land-use planning for sustainable agriculture. Land Use Policy, 81, 424-434. https://doi.org/10.1016/j.landusepol.2018.11.018

Bansal, S., Viboud, C., & Hyman, J. M. (2007). Modeling the emergence of influenza epidemics. Journal of Theoretical Biology, 246(1), 94-104. https://doi.org/10.1016/j.jtbi.2006.11.009

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. JAMA, 319(13), 1317-1318. https://doi.org/10.1001/jama.2018.12123

Bikhchandani, S., Hirshleifer, D., & Welch, I. (2004). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy, 100(5), 992-1026. https://doi.org/10.1086/250115

Cassman, K. G., Dobermann, A., & Walters, D. T. (2017). Agroecosystems, nitrogen-use efficiency, and sustainability. Nature, 422 (6928), 681-686. https://doi.org/10.1038/nature01550

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care — addressing ethical challenges. The New England Journal of Medicine, 378(11), 981-983. https://doi.org/10.1056/NEJMp1713962

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., et al. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface, 15(141), 20170387. https://doi.org/10.1098/rsif.2017.0387

Ching, T., Li, P. J., & Hong, M. P. (2020). Deep learning for diabetic retinopathy: A review. Journal of Biomedical Informatics, 107,103490. https://doi.org/10.1016/j.jbi.2020.103490

Choudhury, A., Makar, K., & Wainwright, D. (2020). Machine learning in healthcare: A review. Journal of Health Management, 22(3), 356-368. https://doi.org/10.1177/0972063420932104

Desautels, T., Calvert, J., & Toma, M. (2016). Prediction of sepsis in the ICU using machine learning. Computational and Structural Biotechnology Journal, 14, 1-10. https://doi.org/10.1016/j.csbj.2016.10.006

Esteva, A., Kuprel, B., Ramos, A., Swetter, S. M., Blau, H. M., & Thrun, S. (2019). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542 (7639), 115-118. https://doi.org/10.1038/nature21056

Esteva, A., Robbins, K., & Beiswanger, C. (2021). A deep learning model for skin cancer detection: A pilot study. Nature Medicine, 27(2), 298-303. https://doi.org/10.1038/s41591-020-01181-0

Flaxman, S., Mishra, S., Gandy, A., et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 584(7820), 257-261. https://doi.org/10.1038/s41586-020-2405-7

Guntuku, S. C., Bian, J., Kaye, K., & Preedy, V. R. (2019). Detecting mental health issues on social media: A review of the literature. Journal of Medical Internet Research, 21(7), e12812. https://doi.org/10.2196/12812

Harrell, F. E. (2015). Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis (2nd ed.). Springer. https://doi.org/10.1007/978-3-319-19425-7

Hosmer, D. W., & Lemeshow, S. (2000). Applied logistic regression (2nd ed.). Wiley. https://doi.org/10.1002/0471722166

Huang, T., Miao, J., & Yang, X. (2019). Operations research in healthcare: A systematic review. Health Systems, 9(2), 73-94. https://doi.org/10.1080/20476965.2019.1580676

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260. https://doi.org/10.1126/science.aaa8415

Keeling, M. J., & Rohani, P. (2008). Modeling infectious diseases in humans and animals. Princeton University Press.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society A, 115(772), 700-721. https://doi.org/10.1098/rspa.1927.0118

Kitsak, M. (2010). Identification of influential spreaders in complex networks. Physical Review Letters, 101(8), 058701. https://doi.org/10.1103/PhysRevLett.101.058701

Kourou, K., Exarchos, T. P., Baga, A., & Koutkias, V. (2015). Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal, 13, 8-17. https://doi.org/10.1016/j.csbj.2014.11.003

Kuhn, M., & Johnson, K. (2013). Applied predictive modeling Springer. https://doi.org/10.1007/978-1-4614-6849-3

Lee, W. C., Hsu, C. Y., & O'Brien, D. (2019). Machine learning for drug discovery: From molecular representation to prediction of drug activity. Nature Reviews Chemistry, 3(9), 594-607. https://doi.org/10.1038/s41570-019-0152-5

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). Statistical and Machine Learning forecasting methods. Journal of Forecasting, 39(4), 570-580. https://doi.org/10.1002/for.2605

Meyers, L. A., Pourbohloul, B., Newman, M. E. J., & Skowronski, D. M. (2005). Network theory and SARS: Predicting outbreak diversity. Journal of Theoretical Biology, 232(1), 71-81. https://doi.org/10.1016/j.jtbi.2004.06.018

Odeh, J.O., Agbata, B.C., Ezeafulukwe, A.U., Madubueze, C.E., Acheneje, G.O., Topman.N.N (2024) A mathematical model for the control of Chlamydia disease with treatment strategy. JMAR, 7(1): 1-20

Obermeyer, Z., & Emanuel, E. J. (2019). Predicting the future — big data, machine learning, and health care. New England Journal of Medicine, 375(13), 1216-1219. https://doi.org/10.1056/NEJMp1606181

Obermeyer, Z., Powers, B. W., Volk, J. E., & Emanuel, E. J. (2016). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14),

Shior, M.M, Agbata, B.C., Acheneje, G.O., Odeh, J.O., Omale, A.J., Oko, I.M (2024), Applications of first-Order Ordinary differential equations to real life system. IJSAEP, 3(1):1-7

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., et al. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface, 15(141), 20170387. https://doi.org/10.1098/rsif.2017.0387

Esteva, A., Kuprel, B., Swetter, S. M., et al. (2019). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118. https://doi.org/10.1038/nature21056

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260. https://doi.org/10.1126/science.aaa8415

Kourou, K., Exarchos, T. P., Baga, A., & Koutkias, V. (2015). Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal, 13, 8-17. https://doi.org/10.1016/j.csbj.2014.11.003

Obermeyer, Z., Powers, B. W., Volk, J. E., & Emanuel, E. J. (2016). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Obermeyer, Z., & Emanuel, E. J. (2019). Predicting the future — big data, machine learning, and health care. New England Journal of Medicine, 375(13), 1216-1219. https://doi.org/10.1056/NEJMp1606181

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358. https://doi.org/10.1056/NEJMra1902025

Anderson, R. M., & May, R. M. (1992). Infectious diseases of humans: Dynamics and control. Oxford University Press.

Flaxman, S., Mishra, S., Gandy, A. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 584(7820), 257-261. https://doi.org/10.1038/s41586-020-2405-7

Keeling, M. J., & Rohani, P. (2008). Modeling infectious diseases in humans and animals. Princeton University Press.

Meyers, L. A., Pourbohloul, B., Newman, M. E. J., & Skowronski, D. M. (2005). Network theory and SARS: Predicting outbreak diversity. Journal of Theoretical Biology, 232(1), 71-81. https://doi.org/10.1016/j.jtbi.2004.06.018

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future — big data, machine learning, and health care. New England Journal of Medicine, 375(13), 1216-1219. https://doi.org/10.1056/NEJMp1606181

Vynnycky, E., & White, R. G. (2010). An introduction to infectious disease modeling. Oxford University Press.

Ching, T., Li, P. J., Hong, M. P. (2020). Deep learning for diabetic retinopathy: A review. Journal of Biomedical Informatics, 107, 103490. https://doi.org/10.1016/j.jbi.2020.103490

Esteva, A., Robbins, K., Beiswanger, C. (2021). A deep learning model for skin cancer detection: A pilot study. Nature Medicine, 27(2), 298-303. https://doi.org/10.1038/s41591-020-01181-0

Obermeyer, Z., Powers, B. W., & Emanuel, E. J. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358. https://doi.org/10.1056/NEJMra1902025

Ching, T., Li, P. J., Hong, M. P. (2020). Deep learning for diabetic retinopathy: A review. Journal of Biomedical Informatics, 107, 103490. https://doi.org/10.1016/j.jbi.2020.103490

Esteva, A., Robbins, K., Beiswanger, C. (2021). A deep learning model for skin cancer detection: A pilot study. Nature Medicine, 27(2), 298-303. https://doi.org/10.1038/s41591-020-01181-0

Obermeyer, Z., Powers, B. W., & Emanuel, E. J. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358. https://doi.org/10.1056/NEJMra1902025

Desautels, T., Calvert, J., & Toma, M. (2016). Prediction of sepsis in the ICU using machine learning. Computational and Structural Biotechnology Journal, 14, 1-10. https://doi.org/10.1016/j.csbj.2016.10.006

Esteva, A., Kuprel, B., Swetter, S. M., Blau, H., and Thrun, S. (2019). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118. https://doi.org/10.1038/nature21056

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., and Coupland, H. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 584(7820), 257-261. https://doi.org/10.1038/s41586-020-2405-7

Guntuku, S. C., Bian, J., Kaye, K., and Preedy, V. R. (2019). Detecting mental health issues on social media: A review of the literature. Journal of Medical Internet Research, 21(7), e12812. https://doi.org/10.2196/12812

Huang, T., Miao, J., and Yang, X. (2019). Operations research in healthcare: A systematic review. Health Systems, 9(2), 73-94. https://doi.org/10.1080/20476965.2019.1580676

Kourou, K., Exarchos, T. P., Baga, A., and Koutkias, V. (2015). Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal, 13, 8-17. https://doi.org/10.1016/j.csbj.2014.11.003

Lee, W. C., Hsu, C. Y., and O'Brien, D. (2019). Machine learning for drug discovery: From molecular representation to prediction of drug activity. Nature Reviews Chemistry, 3(9), 594-607. https://doi.org/10.1038/s41570-019-0152-5

Obermeyer, Z., Powers, B. W., Volk, J. E., and Emanuel, E. J. (2016). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Rajkomar, A., Dean, J., and Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358. https://doi.org/10.1056/NEJMra1902025

Vynnycky, E., and White, R. G. (2010). An introduction to infectious disease modeling. Oxford University Press.

Andersson, H. & Medley, G. F. (2020). Stochastic models of infectious disease dynamics. Mathematical Biosciences, 48(2), 145-168. https://doi.org/10.1016/j.mbs.2020.108268

- Bansal, S., Viboud, C., & Hyman, J. M. (2007). Modeling the emergence of influenza epidemics. Journal of Theoretical Biology, 246(1), 94-104. https://doi.org/10.1016/j.jtbi.2006.11.009

Bikhchandani, S., Hirshleifer, D., & Welch, I. (2004). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy, 100(5), 992-1026. https://doi.org/10.1086/250115

Flaxman, S., Mishra, S., Gandy, A., et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 584(7820), 257-261. https://doi.org/10.1038/s41586-020-2405-7

Havlin, S., Makse, H. A., & Stanley, H. E. (2019). The modeling of complex systems in biology: A new approach. BioEssays, 41(7), 1900012. https://doi.org/10.1002/bies.201900012

Huang, Y (2019). Optimizing healthcare resource allocation: A case study. Health Services Research, 54(2), 233-243. https://doi.org/10.1111/1475-6773.13002

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression (3rd ed.). Wiley.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society A, 115(772), 700-721. https://doi.org/10.1098/rspa.1927.0118

Kitsak, M., (2010). Identification of influential spreaders in complex networks. Physical Review Letters, 101(8), 058701. https://doi.org/10.1103/PhysRevLett.101.058701

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). Statistical and Machine Learning forecasting methods:

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. JAMA, 319(13), 1317-1318. https://doi.org/10.1001/jama.2018.12123

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care — addressing ethical challenges. The New England Journal of Medicine, 378(11), 981-983. https://doi.org/10.1056/NEJMp1713962

Choudhury, A., Makar, K., & Wainwright, D. (2020). Machine learning in healthcare: A review. Journal of Health Management, 22(3), 356-368. https://doi.org/10.1177/0972063420932104

Esteva, A., Kuprel, B., Ramos, A. (2019). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118. https://doi.org/10.1038/nature21056

Kourou, K., Exarchos, T. P., Engin, K., & Karamouzis, M. V. (2015). Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal, 13, 8-17. https://doi.org/10.1016/j.csbj.2014.11.005

Obermeyer, Z., Powers, B., Muffly, L. S., & Etzioni, R. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. The New England Journal of Medicine, 380(14), 1347-1358. https://doi.org/10.1056/NEJMra1814252

Harrell, F. E. (2015). Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis (2nd ed.). Springer. https://doi.org/10.1007/978-3-319-19425-7

Hosmer, D. W., & Lemeshow, S. (2000). Applied logistic regression (2nd ed.). Wiley. https://doi.org/10.1002/0471722166

Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Menard, S. (2002). Applied logistic regression analysis (2nd ed.). Sage Publications. https://doi.org/10.4135/9781412980667

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing and Management, 45(4), 427-437. https://doi.org/10.1016/j.ipm.2009.03.002