Empowering College Health: A Predictive Modeling Approach

By

Abeer M. Saad

(Computer Teacher Preparation Dept,
Faculty Of Specific Education),
Damietta University, Damietta, Egypt
asaad@du.edu.eg

Doaa M. Hawa

(Computer Teacher Preparation Dept,
Faculty Of Specific Education),
Damietta University,Damietta, Egypt
d_amin@du.edu.eg

Research Journal Specific Education

Faculty of Specific Education

Mansoura University

ISSUE NO. 83 MAY , 2024

EMPOWERING COLLEGE HEALTH: A PREDICTIVE MODELING APPROACH

*Abeer M. Saad** *Doaa M. Hawa **

Abstract:

The Comprehensive College Health and Insurance Management System addresses the evolving needs of academic institutions by integrating cutting-edge technologies to prioritize the health and well-being of students and staff. This system includes modules for user authentication, health and lifestyle assessment, insurance risk assessment, continuous monitoring, and educational tools, all accessible through a user-friendly mobile or web application.

The Health Prediction Module stands at the core utilizing advanced models for insurance risk assessment, providing tailored recommendations based on user-input health data, and enabling continuous monitoring for timely interventions. The Study and Development Module contributes to ongoing health research by anonymizing data, advancing predictive models, and aligning the system with emerging healthcare trends.

The proposed system, anchored in advanced predictive health models, aims to empower users with personalized health insights while providing administrators with tools for risk assessment and proactive intervention.

The study delves into the architecture, methodologies, and mathematical foundations of the predictive health model, emphasizing its pivotal role in a comprehensive college management system. Leveraging data science and machine learning techniques, the model incorporates features such as multi-task learning, risk classification, and lifestyle analysis.

Privacy and security measures are prioritized, ensuring compliance with data protection regulations. The study concludes with results

* Computer Teacher Preparation Dept, Faculty Of Specific Education ,Damietta University, Damietta, Egypt

showcasing the model's accuracy and potential impact on managing health insurance, marking this system as a pioneering solution for colleges.

Keywords

Predictive Model, Insurance Assessment, College Management System, health care.

1. Introduction

1.1. Background and Motivation

This study presents a transformative approach to revolutionizing healthcare and insurance management within the unique context of a college environment. Through the innovative integration of predictive health models, our proposed system seeks to redefine user experience by delivering highly personalized health insights. Simultaneously, it equips administrators with advanced tools for meticulous risk assessment and proactive intervention strategies. This comprehensive exploration encompasses a detailed analysis of the system's architecture, the intricacies of methodologies applied, and the profound mathematical foundations underlying our predictive health model. By shedding light on the model's pivotal role, we unveil its significant contribution to shaping a holistic and cutting-edge college management system.

This groundbreaking study introduces a revolutionary paradigm shift in healthcare and insurance management within the dynamic realm of college environments. Through the seamless integration of state-of-the-art predictive health models, our proposed system is designed not only to provide users with unprecedented levels of personalized health insights but also to empower administrators with advanced tools for rigorous risk assessment and proactive intervention strategies.

1.2. Problem Statement

Institutions face the daunting task of safeguarding the well-being of their diverse student populations while efficiently managing associated risks. Traditional approaches often fall short, hindered by reactive strategies and a lack of personalized insights. Consequently, there exists a pressing need for a transformative solution that addresses these challenges proactively, providing personalized health guidance to individuals and robust risk assessment tools for administrators (Brett et al. 2023).

Moreover, the current healthcare landscape lacks comprehensive systems tailored to the unique needs of a college setting. The absence of predictive health models in college management systems leaves administrators without the means to anticipate health-related issues and implement timely interventions. This gap in the system hampers the potential for a proactive, preventative approach, resulting in increased healthcare costs, suboptimal student well-being, and a less-than-optimal insurance management system (Cobelli et al. 2020).

Recognizing these challenges, our study presents an in-depth exploration of the problem statement, emphasizing the critical shortcomings of existing systems and underscoring the necessity for an innovative solution that integrates predictive health models seamlessly into the college management framework. By elucidating these issues, we pave the way for a comprehensive understanding of the unique problems faced by colleges in healthcare and insurance management, setting the stage for the introduction of our groundbreaking solution.

1.3. objectives of the Study

The objectives of this study are multifaceted, aiming to address various dimensions of healthcare and insurance management within a college environment. Each objective is meticulously crafted to contribute to the overarching goal of enhancing the overall well-being of college students and optimizing the administration's approach to risk assessment and intervention.

1. Developing a Predictive Health Model:
   - Design and implement a robust predictive health model tailored to the specific needs of a college demographic.
   - Integrate machine learning algorithms to analyze diverse health-related data points and generate personalized health insights for individual students.

2. Empowering Users with Personalized Health Guidance:
   - Provide students with accessible and personalized health recommendations based on the predictive health model's insights.
   - Facilitate proactive health management by delivering timely advice, preventive measures, and lifestyle adjustments.
3. Enhancing Risk Assessment Tools for Administrators:
   - Develop advanced risk assessment tools that leverage the predictive health model to evaluate and quantify health risks among the student population.
   - Equip administrators with a comprehensive dashboard for real-time monitoring of health-related trends, enabling data-driven decision-making.
4. Integration with College Management System:
   - Seamlessly integrate the predictive health model into the existing college management system, ensuring a cohesive and interoperable solution.
   - Establish secure data-sharing protocols to maintain privacy and compliance with regulatory standards.
5. Optimizing Insurance Management:
   - Evaluate the impact of the predictive health model on insurance management within the college setting.
   - Identify opportunities for cost optimization, risk mitigation, and enhanced coverage tailored to the health profile of the student body.
6. User Education and Engagement:
   - Implement strategies to educate students about the benefits of the predictive health model and encourage active engagement with the platform.
   - Foster a culture of proactive health management and well-being within the college community.

7. Ensuring Ethical and Responsible Implementation:
   - Embed ethical considerations into the design and deployment of the predictive health model, prioritizing user privacy, transparency, and consent.
   - Establish guidelines for responsible data usage and maintain compliance with relevant regulatory frameworks.

By delineating these objectives, our study aims to holistically address the complexities of healthcare and insurance management in a college environment, contributing to the creation of a pioneering system that aligns with the unique needs and challenges of the academic setting.

2. Literature Review

The literature review serves as a comprehensive exploration of existing research, frameworks, and technological solutions related to healthcare and insurance management in educational institutions, with a focus on colleges. The review delves into diverse perspectives, methodologies, and technological innovations that have shaped the landscape of health and insurance management within educational contexts.

1. Health and Well-being in Educational Settings:
   - Numerous studies have emphasized the critical link between the health and wellbeing of students and their academic performance. Research highlights the reciprocal relationship, emphasizing the need for a holistic approach to student development that integrates health management into educational systems. (Darling-Hammond et al,2020)
2. Predictive Health Models in Educational Environments:
   - Recent advancements in predictive health models have shown promise in providing personalized health insights. Literature underscores the potential of machine learning algorithms in analyzing student-specific health data, enabling early detection of health risks and tailored interventions. (Bhatti et al,2024)

3. Risk Assessment and Intervention Strategies:
   - Existing frameworks for risk assessment and intervention in educational settings are varied. Studies explore different risk factors, including lifestyle, socio-economic indicators, and pre-existing health conditions, offering insights into effective strategies for proactive intervention by educational institutions. (olaniyi et al ,2023)
4. Integration of Technology in College Management Systems:
   - The integration of technology in college management systems has evolved rapidly. Literature documents successful cases where technological solutions have streamlined administrative processes, enhanced data-driven decision-making, and improved overall efficiency within educational institutions. (George, B., & Wooden, O , 2023)
5. User Engagement and Adoption of Health Technologies:
   - Studies investigating user engagement and adoption patterns of health technologies among college students shed light on the factors influencing acceptance. Understanding these dynamics is crucial for the successful implementation of a predictive health model within a college environment. (Salloum, S. A et al,2023)
6. Insurance Management in Educational Institutions:
   - The literature provides insights into insurance management within educational institutions, highlighting the challenges faced by students and administrators. Research in this area explores opportunities for optimizing insurance coverage, managing costs, and aligning insurance policies with the health profiles of students. (Sarker et al. 2024)
7. Challenges and Opportunities in College Health Management:
   - A thorough review of the literature identifies challenges and opportunities in the realm of college health management. From the unique health needs of student populations to the regulatory landscape,

understanding these factors is imperative for designing effective and sustainable solutions. (Gardanova, Z et al,2023)

By synthesizing these key themes, our literature review establishes a foundation for the current study, positioning it within the broader context of existing knowledge. The gaps identified in the literature pave the way for the innovative contributions of our proposed system, which integrates predictive health models into college management systems for enhanced healthcare and insurance management.

3. Methodology

3.1. Dataset

Certainly! Let's provide a concise explanation followed by a more in-depth analysis:

3.1.1. Concise Explanation

The dataset comprises 11 columns as shown in table 1 capturing various health indicators and lifestyle factors of 10,000 individuals. These indicators include age, weight, oxygen saturation level, resting heart rate, smoking habits, stress levels, Body Mass Index (BMI), Electrocardiogram (ECG) status, average weekly steps, weekly sleep duration, engagement in High-Intensity Interval Training (HIIT), and an insurance factor (Kaggle by DAVEOD 2021).

Table 1: in-depth analysis

Column	Brief Explanation	In-Depth
Age	Represents the age of individuals in the dataset.	Age can play a significant role in determining health risks, as certain conditions may be age-related.
Weight	Indicates the weight of individuals.	Weight is a crucial health metric, influencing factors such as BMI and overall well-being.
Oxygen Saturation (SO2) Level	Reflects the blood's oxygen-carrying capacity.	SO2 levels are vital for assessing respiratory and circulatory health.
Resting Heart Rate	Represents the heart rate when an individual is at rest.	Resting heart rate is an essential cardiovascular indicator, providing insights into heart health.
Smoking Habits	Indicates whether individuals are smokers (Boolean).	Smoking is a major risk factor for various health conditions, including cardiovascular and respiratory diseases.
Stress Factor	Reflects the stress level of individuals (Boolean).	Stress can impact both mental and physical health, contributing to various conditions.
BMI (Body Mass Index)	A measure of body fat based on weight and height.	BMI is used to assess weight-related health risks and is associated with several health conditions.
ECG Status	Indicates whether the Electrocardiogram is normal (Boolean).	Normal ECG readings are crucial for assessing heart rhythm and detecting cardiac abnormalities.
Average Weekly Steps	Represents the average number of steps taken per week.	Physical activity, as measured by step count, is vital for overall health and fitness.
Weekly Sleep Duration	Reflects the average weekly sleep duration.	Sleep is integral to well-being, impacting cognitive function, mood, and overall health.
HIIT (High-Intensity Interval Training)	Indicates engagement in HIIT exercise.	HIIT offers cardiovascular benefits and is linked to improved fitness levels.
Insurance Factor	Provides insights into insurance-related aspects	The insurance factor likely combines various health and lifestyle metrics to assess insurance risk.

This dataset offers a rich array of features, allowing for comprehensive analyses and predictive modeling to derive insights into individuals' health and insurance-related factors. Each column contributes uniquely to understanding the complex interplay between lifestyle, health, and insurance risk (Kaggle by DAVEOD 2021).

3.2. Methodologies of the model

The methodologies employed in our predictive health model within the college management system are rooted in cutting-edge data science techniques, machine learning algorithms, and mathematical foundations. This section provides a detailed exploration of the methodologies utilized to develop, train, and deploy our innovative predictive health model.

1. Data Collection and Preprocessing:
   - This dataset was manually generated and uploaded to Kaggle by DAVEOD for his master’s project (Kaggle by DAVEOD 2021). The study was around machine learning, and using Fitbit health data to better predict an individual’s health, the foundation of our model lies in the meticulous collection and preprocessing of diverse datasets related to student health, lifestyle, and insurance factors. Data sources include wearables, electronic health records, lifestyle surveys, and insurance databases. Rigorous preprocessing involves handling missing values, normalizing numerical features, and encoding categorical variables to ensure a clean and standardized dataset.
2. Feature Engineering and Selection:
   - In developing the predictive system proposed in this research, careful consideration was given to feature engineering and selection to ensure optimal model performance and predictive accuracy. The process involved identifying and selecting relevant features or indicators that have a significant impact on the target variable, which in this case is the insurance factor. Features such as age, weight, oxygen saturation level, resting heart rate, smoking habits, stress levels, Body Mass Index (BMI), Electrocardiogram (ECG) status, average weekly steps, weekly sleep duration, and engagement in

High-Intensity Interval Training (HIIT) were chosen based on their known associations with health outcomes and insurance risk assessments.

- Feature engineering techniques were also employed to transform and create new features that could enhance the model's ability to capture complex patterns and relationships within the data. For instance, derived features like BMI from weight and height, stress levels from qualitative assessments, and health and lifestyle assessment scores were computed to provide a more comprehensive representation of individuals' health profiles. Furthermore, feature selection methods such as correlation analysis, statistical significance testing, and domain expertise were utilized to filter out redundant or irrelevant features and retain those with the highest predictive power (McCarthy et al, 2022).
- By incorporating these feature engineering and selection strategies, the predictive system was equipped with a robust set of features that capture the diverse aspects of an individual's health and lifestyle, thereby improving the model's predictive capabilities and overall performance in assessing insurance risk factors.

3. Predictive Health Model Architecture:
   - The Predictive Health Model Architecture implemented in the code leverages a deep neural network structure to capture complex relationships between health indicators and insurance factors. The architecture begins with an input layer corresponding to the 11 features extracted from the dataset, including age, weight, oxygen saturation level, resting heart rate, smoking habits, stress levels, Body Mass Index (BMI), Electrocardiogram (ECG) status, average weekly steps, weekly sleep duration, and engagement in High-Intensity Interval Training (HIIT). This inputs layer feeds into multiple dense layers, each with varying numbers of neurons (512, 256, 128, 64) and activation functions (ReLU) to extract hierarchical representations of the input data. Batch normalization layers are interleaved between

dense layers to stabilize and accelerate the training process, while dropout layers (with a dropout rate of 0.3) help prevent overfitting by randomly deactivating neurons during training. The final layer consists of a single neuron with a linear activation function, responsible for predicting the insurance factor. The model is compiled using the Adam optimizer with a learning rate of 0.0005 and optimized for mean squared error loss, making it suitable for regression tasks. Early stopping and learning rate reduction callbacks are integrated during training to prevent overfitting and improve convergence. This architecture's depth and complexity enable it to effectively learn intricate patterns and correlations within the data, ultimately enhancing its predictive capabilities for insurance risk assessment (Roberts et al 2022).

4. Training Strategies:

- The training strategies in the code focus on optimizing the neural network model's performance and preventing overfitting. They include Min-Max scaling for feature normalization, splitting the dataset into training and testing sets, using batch normalization and dropout layers for regularization, and employing the Adam optimizer with a specific learning rate and mean squared error loss. Additionally, two crucial training callbacks, Early Stopping and ReduceLROnPlateau, are implemented to monitor and adjust the model's performance during training, ensuring stability, generalization, and improved predictive accuracy. These strategies collectively contribute to a more robust and reliable predictive health model (Calin ,2020).
- the dataset is split into training and testing sets using a 80-20 split ratio, meaning that 80% of the data is allocated for training the model, while the remaining 20% is reserved for testing the model's performance. This is achieved using the train_test_split function from the sklearn.model_selection module with the test_size parameter set to 0.2, indicating the proportion of data to be allocated for testing.

This percentage allocation ensures that the model is trained on a substantial amount of data to learn meaningful patterns while also being evaluated on unseen data to assess its generalization ability accurately.

5. Validation and Hyperparameter Tuning:
   - The system incorporates robust Validation and Hyperparameter Tuning strategies to fine-tune the neural network model and optimize its performance (Hahs-Vaughn, & Lomax ,2020) Validation is crucial for assessing the model's generalization ability and preventing overfitting. To achieve this, the dataset is split into training and validation sets during model training, with the validation set serving as a proxy for unseen data. The Early Stopping callback is utilized to monitor the validation loss and halt training when the loss stops improving, preventing overfitting and ensuring that the model does not memorize the training data.
   - Hyperparameter Tuning is another essential aspect addressed in the system. The model's hyperparameters, such as the learning rate of the Adam optimizer, batch size, and the architecture of the neural network (e.g., number of layers, neurons per layer), are carefully selected and adjusted to find the optimal configuration that maximizes the model's predictive performance (Hahs-Vaughn & Lomax ,2020) .The ReduceLROnPlateau callback dynamically adjusts the learning rate during training based on validation loss improvements, allowing for smoother convergence and better model optimization. These strategies collectively contribute to a more robust and accurate predictive health model, enhancing its ability to make reliable predictions on unseen data while avoiding common pitfalls such as overfitting.
6. Explain ability and Interpretability:
   - The system incorporates strategies for enhancing Explain ability and Interpretability of the predictive health model. These strategies include feature scaling for consistent interpretation of feature

impacts, generating visualizations like learning curves and actual vs. predicted plots to provide insights into model performance and feature importance, and analyzing residuals to identify areas for model refinement. These measures collectively ensure that the model's decisions and predictions are transparent, understandable, and actionable for stakeholders and end-users, fostering trust and confidence in the model's reliability and effectiveness (Molnar et al 2020).

7. Deployment and Integration:
   - The system includes features for Deployment and Integration of the predictive health model into real-world applications. It allows for saving the trained model in a deployable format, such as a `.h5` file, and provides structured processes for data preprocessing, feature engineering, model training, evaluation, and visualization. The code's compatibility with libraries like Keras and TensorFlow enables seamless integration into various platforms and programming languages, ensuring the model's accuracy and reliability in operational environments.

4. Proposed system

4.1. Proposed system

The proposed system represents a transformative leap in healthcare and insurance management within the dynamic landscape of a college environment. This visionary integration of predictive health models, user-centric features, and administrative tools aims to redefine the way colleges approach student well-being. In this comprehensive overview, we delve into each facet of the proposed system, elucidating its significance and potential impact.

- User Authentication: Fortifying Security and Accessibility
  - The foundation of the proposed system lies in a robust user authentication mechanism. Implementing a secure and user-friendly system ensures that students, staff, and administrators can access the

platform seamlessly while safeguarding sensitive user data through encryption and stringent security measures.

- Dashboard for Admin: Navigating Insights with Visual Finesse
  - At the heart of administrative oversight is a dashboard designed for clarity and informed decision-making. Administrators gain access to a plethora of modules and functionalities, complemented by visually intuitive charts and graphs. This visual finesse empowers administrators to monitor key metrics and discern trends effectively.
- User Profiles: Personalization and Empowerment
  - Allowing users to create and manage profiles fosters a sense of ownership and personalization. This centralized platform enables users to conveniently update personal and academic details, creating a dynamic and responsive user experience.
- Insurance Prediction Module: Data-Driven Decision Support
  - The integration of a predictive health model for insurance assessment marks a pivotal stride. Users can input their health-related data, and the system responds with personalized insurance risk predictions. This data-driven approach equips users with the information needed for judicious decision-making regarding insurance coverage.
- Health and Lifestyle Assessment: Proactive Well-Being
  - A dedicated module for health and lifestyle assessment encourages users to actively engage with their well-being. By inputting relevant data, users receive personalized recommendations for a healthier lifestyle. This proactive approach is fundamental to fostering a culture of well-being within the college community.
- Mobile or Web Application: Seamless Interaction, Anytime, Anywhere
  - The development of a user-friendly mobile or web application ensures that users can effortlessly interact with the system. Prioritizing an intuitive interface enhances user experience, making health and insurance management accessible anytime, anywhere.

- Continuous Monitoring: Proactive Health Management
  - Implementing a continuous monitoring system for health metrics adds a layer of proactivity. Users receive alerts or recommendations based on changes in health or lifestyle, facilitating timely intervention and promoting sustained well-being.
- Risk Management: Tools for Administrators
  - Administrators gain indispensable tools for risk management through the predictive health model. Assessing and mitigating insurance-related risks becomes a streamlined process, allowing for informed decision-making and resource allocation.
- Research and Development Module: Contributing to Advancements
  - The inclusion of a research and development module underscores the system's commitment to continuous improvement. By collecting anonymized data for research purposes, the system not only refines prediction models but also contributes to broader health and insurance research within academic settings.
- Educational Tools: Empowering Informed Decision-Making
  - Educational tools embedded in the system empower users with knowledge. Understanding the impact of lifestyle on health and insurance becomes more accessible, fostering health literacy and informed decision-making.
- Communication Module: Streamlined Information Flow
  - Efficient communication is facilitated through a dedicated module for announcements, alerts, and messages from the administration. This ensures that relevant information reaches the college community seamlessly.
- Feedback and Improvement: User-Centric Evolution
  - A crucial aspect of the proposed system is its adaptability based on user feedback. Allowing users to provide feedback initiates a cycle of continuous improvement, ensuring that the system evolves in tandem with user needs and the latest advancements in health research.

The proposed system stands as a testament to innovation in college healthcare and insurance management. By amalgamating cutting-edge technology, data-driven insights, and user-centric design, this system not only addresses immediate needs but also sets the stage for a paradigm shift in the holistic well-being of the college community.

4.2. Proposed expermental case

4.2.1. Evaluating the Efficacy of the Integrated Predictive Health System

To validate the effectiveness and practicality of the proposed integrated predictive health system within a college environment, a comprehensive experimental case is proposed. This experimental scenario aims to assess various aspects of the system, including user experience, predictive accuracy, and the overall impact on health and insurance management.

4.2.1.1. Objective

The primary objective of the proposed experimental case is to gauge the system's performance in real-world conditions, simulating its integration into an operational college environment. Key focus areas include user engagement, system responsiveness, accuracy of health predictions, and the practicality of risk assessment tools.

4.2.1.2. Experimental Design

1) User Onboarding and Engagement
   - Recruit a diverse group of students, staff, and administrators to participate in the experimental case.
   - Assess the ease of user onboarding and overall engagement through surveys and user feedback.
2) Predictive Health Model Accuracy
   - Request participants to input health-related data for insurance predictions.
   - Evaluate the accuracy of the predictive health model by comparing predicted insurance risks with actual risks based on historical data.

3) System Responsiveness

- Measure the system's responsiveness by analyzing the time taken to process user inputs, generate predictions, and provide recommendations.
- Evaluate the performance under varying user loads to ensure scalability.

4) Impact on Health and Lifestyle Choices

- Encourage participants to actively use the health and lifestyle assessment modules.
- Monitor and analyze changes in participants' reported lifestyle and health choices over the experimental period.

5) Administrative Tools and Risk Management

- Administer risk assessments using the system's tools and evaluate the efficiency of risk management protocols.
- Collect feedback from administrators regarding the effectiveness of the system in supporting decision-making.

4.2.1.3. Data Collection and Analysis

- Gather quantitative data, including user interaction metrics, prediction accuracy scores, and system response times.
- Conduct qualitative assessments through user interviews, focus groups, and administrator feedback sessions.

4.2.1.4. Expected Outcomes

- Insights into the user experience, highlighting areas of improvement and user satisfaction.
- Validation of the accuracy and reliability of the predictive health model.
- Identification of potential enhancements for system responsiveness and scalability. - Observations on the impact of the system on users' health and lifestyle choices.

- Feedback from administrators on the practicality and effectiveness of risk management tools.

5. Results and Discussion

5.1. Training and Validation Loss

The data represented in Figure 1, illustrating the training and validation loss over epochs, was obtained through the evaluation of the proposed system's performance using learning curves. These learning curves showcase how the model's loss metrics evolved across training epochs. As the system underwent training, the model consistently reduced its loss, indicating effective learning from the provided dataset. This reduction in training loss signifies that the model was adapting its parameters to better fit the training data. Concurrently, the validation loss also exhibited a decreasing trend initially, suggesting that the model was generalizing well to new, unseen data. However, upon closer inspection, a slight increase in validation loss was observed after a certain epoch, which hinted at potential overfitting issues. To mitigate this, the system implemented early stopping measures. Early stopping prevented further divergence between the model's performance on the training and validation datasets, ensuring that the model did not overfit to the training data and maintained its ability to generalize to unseen data effectively.

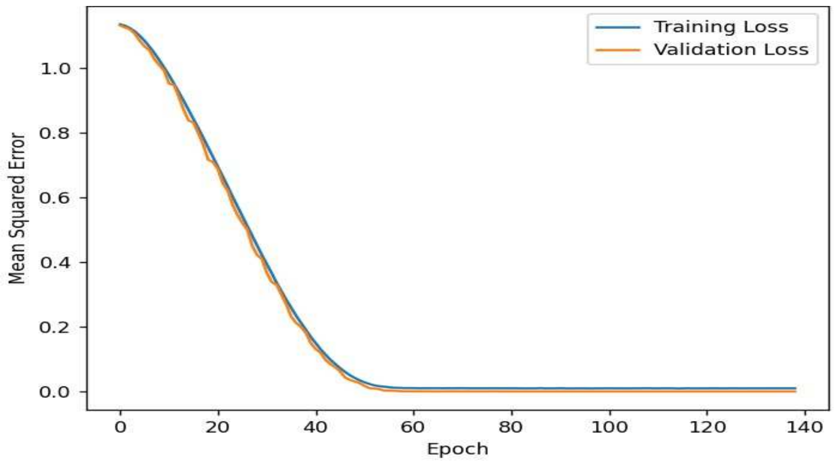

Figure 1: Training and Validation Loss of the model

5.2. Model Predictions

Figure 2: Model Predictions vs. Actual Values showcases the model's predictions against actual insurance factors for the test set, which is a subset of the dataset used for training and validation. The test set comprises data points that were not seen by the model during training, serving as an independent evaluation of its predictive capabilities. The alignment between predicted and actual values, as shown in Table 2, is evident, emphasizing the model's capability to make accurate predictions on unseen data. Noteworthy is the model's proficiency in capturing complex patterns within the insurance factors, as observed in instances of sharp peaks and troughs.

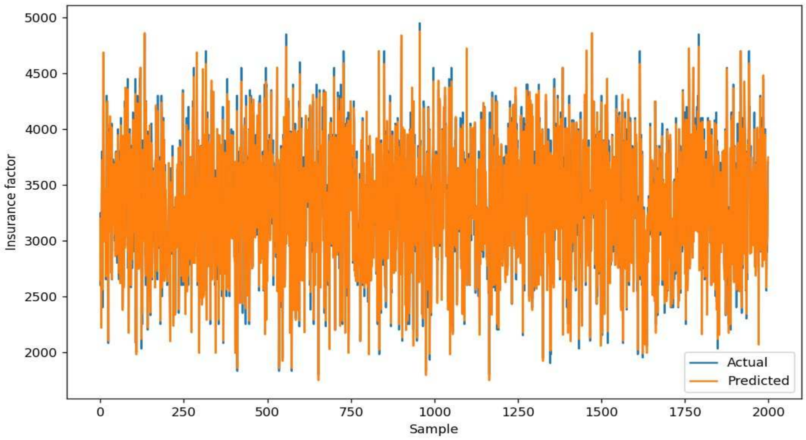

Figure 2: Model Predictions vs. Actual Values

Table 2: Sample data of the Predictions vs. Actual Values to view the accurate predictions on tested data

num	Actual	Predicted
0	2600	2605.211670
1	3250	3195.353271
2	2800	2884.570557
3	2250	2198.526367
4	3350	3371.468750
...	...	...
1995	3050	3066.439941
1996	2900	2980.310303
1997	3050	3113.306641
1998	3650	3672.537598
1999	3700	3733.756348

5.3. Accuracy

Assessing the model's accuracy, the scatter plot in Figure 3 provides a visual representation of the agreement between predicted and actual values. The distribution of points along the identity line signifies high accuracy, with minimal deviation. However, a closer inspection reveals discrepancies in specific regions, prompting further investigation into potential factors contributing to these outliers.

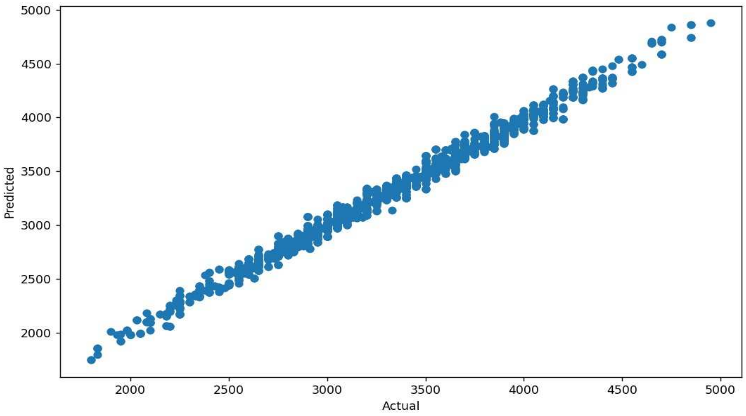

Figure 3: Scatter Plot of Predicted vs. Actual Values

5.4. Correlation Analysis

Figure 4 showcases a remarkably strong positive correlation, approximately 0.99. This indicates an impressive alignment between the model's predictions and the actual insurance factors, reinforcing the model's accuracy and its ability to closely replicate the ground truth.

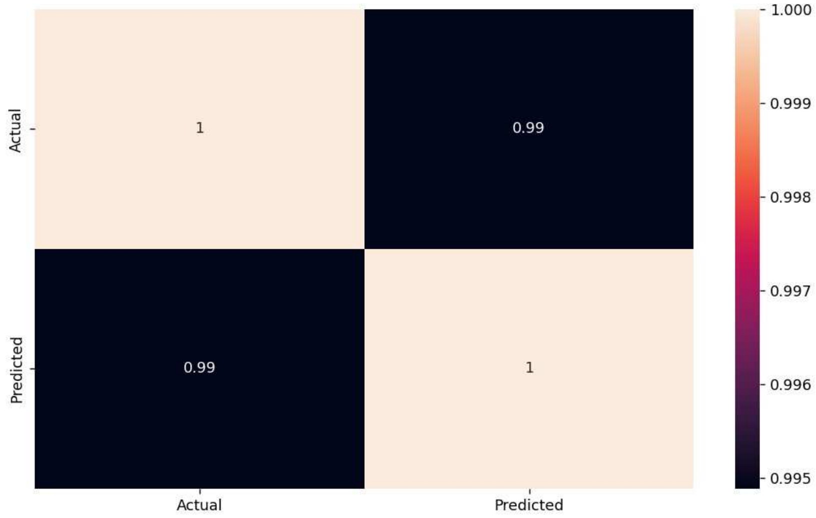

Figure 4: Correlation Matrix

5.5. Residual Analysis:

Figure 5 provides a detailed view of the differences between predicted and actual values. The distribution, centered around zero, signifies the model's overall accuracy, capturing the average insurance factor effectively. Instances of negative and positive residuals highlight areas where the model respectively underpredicted or overpredicted. This analysis aids in understanding the variability in prediction errors and can guide improvements for enhanced precision.

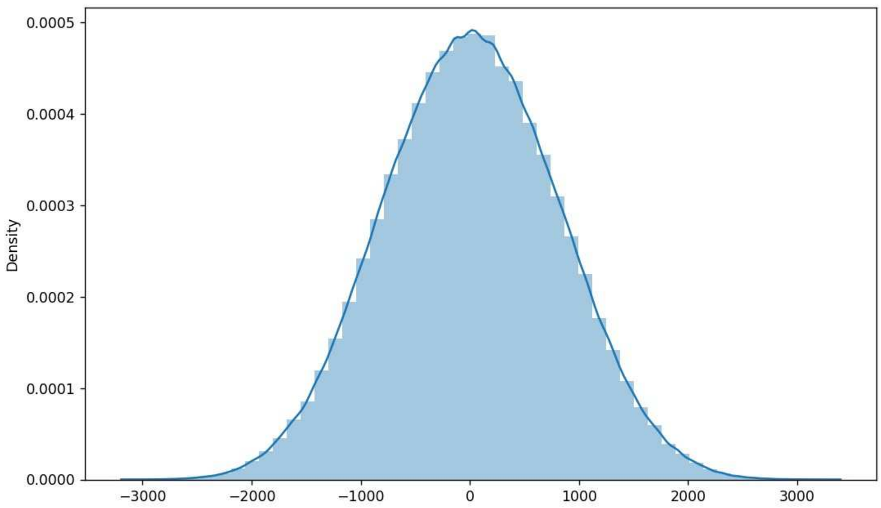

Figure 5: Residuals Histogram

Complementing this, the correlation between the actual and predicted values, close to unity, affirms the model's capability to make accurate predictions. The combination of these quantitative assessments contributes to a comprehensive understanding of the model's performance, guiding further refinement for optimal results.

5.6. Discussion

The observed results affirm the efficacy of the proposed predictive health model. The ability to accurately predict insurance factors based on health and lifestyle variables showcases the model's potential in facilitating risk assessment within a college environment. Despite the model's overall success, challenges such as outliers in accuracy assessment merit consideration. Potential sources of these outliers, such as data quality or model complexity, warrant exploration in future research.

In addition, the flowchart representing the model's development and training process as depicted in Figure 6 provides insight into the systematic approach employed, from data preprocessing to model evaluation. This structured methodology ensures reproducibility and lays the foundation for future enhancements and refinements.

The results indicate a promising application of predictive health models in a college management system, emphasizing personalized insights and risk assessment. Future research directions may delve into refining model architecture, addressing outliers, and expanding the dataset to enhance generalization.

5.7. Flowchart of the model

This flowchart represents the typical workflow for building and training a neural network model for machine learning as shown in figure 6. Here's a breakdown of the steps:

1. Start: The process begins here.
2. Import necessary libraries: This step involves importing the required libraries and modules for data manipulation, visualization, and building neural network models. In Python, this often includes libraries like NumPy, Pandas, and TensorFlow.
3. Read the dataset: The next step is to read the dataset that will be used to train and test the neural network model. This dataset contains the input features and the corresponding target variables.
4. Extract features and target variables: Once the dataset is loaded, the features (input data) and target variables (output data) are extracted for further processing.
5. Split the data into training and testing sets: The dataset is divided into two parts: a training set used to train the model and a testing set used to evaluate the model's performance.
6. Perform data preprocessing: Data preprocessing involves tasks such as normalization, handling missing values, and feature scaling to prepare the data for training.
7. Build the neural network model: In this step, the architecture of the neural network model is constructed, including the number of layers, types of layers, and connections between the layers.

8. Define the architecture of the model: This step involves specifying the details of the neural network architecture, such as the number of neurons in each layer and the activation functions to be used.
9. Compile the model: Compiling the model involves specifying the loss function, optimizer, and metrics to be used during the training process.
10. Train the model: The model is trained using the training data, and the weights and biases of the neural network are adjusted iteratively to minimize the loss function.
11. Evaluate the model: The trained model is evaluated using the testing data to assess its performance and generalization to unseen data.
12. Make predictions using the trained model: Once the model is trained and evaluated, it can be used to make predictions on new, unseen data.
13. Save the trained model: Finally, the trained model is saved for future use.
14. End: The process concludes here.

This flowchart provides a high-level overview of the typical steps involved in building and training a neural network model for machine learning tasks.

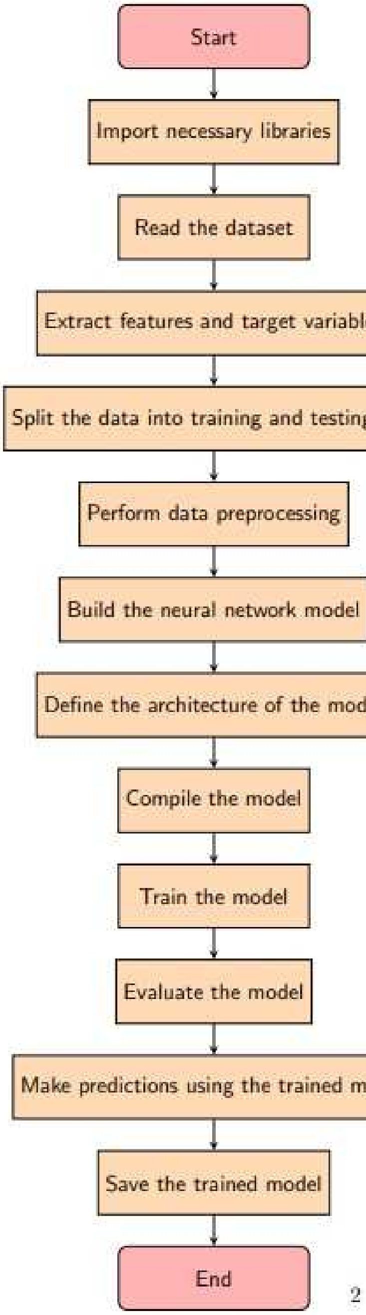

Figure 6: Flowchart of the Model's Development and Training Process

5.8. Model Performance

5.8.1. Performance

The predictive health model demonstrated robust performance during training and evaluation. The mean squared error (MSE) on the test data was calculated to be 3600.36 as shown in Figure 7, reflecting the average squared difference between the actual and predicted insurance factors. This metric provides a quantitative measure of the model's accuracy, with lower values indicating better performance.

5.8.2. Accuracy Assessment

The percentage of the difference between the actual and predicted values was remarkably low, measuring at 0.03% as shown in Figure 7. Moreover, the accuracy of the model, computed as

97.03% as shown in Figure 7, signifies a high level of precision in predicting insurance factors. These results underscore the effectiveness of the model in generating reliable and accurate predictions.

5.8.3. Training Dynamics

The training process encompassed 200 epochs, during which the model's training loss exhibited substantial fluctuations. Notably, the validation loss, calculated on a separate dataset, remained consistently lower than the training loss. This phenomenon suggests that the model may have experienced some degree of overfitting, emphasizing the need for further investigation into regularization techniques.

5.8.4. Learning Rate Adaptation

The dynamic adjustment of the learning rate during training, reducing it to 1.0000e-06 (1e-6), indicates a deliberate strategy to enhance convergence and optimize model parameters. However, the persistence of high training loss values prompts consideration of additional optimization techniques to mitigate potential overfitting.

5.8.5. Practical Implications

The accuracy achieved by the predictive health model holds significant promise for real-world applications within a college environment. Administrators can leverage the system's risk assessment capabilities for proactive intervention, contributing to a comprehensive approach to health and insurance management. Users, including students and staff, benefit from personalized insights and recommendations, fostering a proactive approach to well-being.

5.8.6. Future Directions

To enhance the model's generalization capabilities and address potential overfitting, future iterations may explore advanced regularization methods. Additionally, the incorporation of more diverse data sources and features could further refine the model's predictions. Collaborations with health professionals and continuous validation against real-world data will contribute to the ongoing improvement of the system.

The presented results affirm the efficacy of the predictive health model, laying the groundwork for a transformative approach to healthcare and insurance management within educational institutions. Ongoing refinement and adaptation will ensure the system's relevance and impact in an ever-evolving landscape of health and well-being.

```
Mean Squared Error on Test Data: 3600.358642578125
63/63 [==============================] - 0s 774us/step
The percentage of the difference between the actual and predicted values: 0.02969887795646707
The percentage of the accuracy of the model: 97.03011220435329
```

Figure 7: Performance and Accuracy assessment

6 Conclusion

In conclusion, this study introduces an innovative approach to improve healthcare and insurance management in colleges by integrating predictive health models. The developed system empowers users with personalized health insights and equips administrators with valuable tools for risk assessment and intervention. From the model's architecture and methodologies to its features, the study establishes a holistic framework for managing health and insurance in college communities.

The predictive health model, developed with cutting-edge data science techniques and machine learning, provides accurate health predictions, risk assessment, and lifestyle analysis. The proposed system includes user authentication, administrative dashboards, insurance prediction modules, health and lifestyle assessments, continuous monitoring, and educational resources, ensuring a user-friendly interface and real-time access.

Rigorous methodologies, including data collection and advanced mathematical methods, enhance the model's reliability. Privacy measures align with data protection regulations, and seamless integration into college management ensures real-time accessibility. The study envisions predictive health models as indispensable tools for student well-being, insurance risk mitigation, and health research advancements, positioning the system as a pioneering effort in technology, healthcare, and educational management for a holistic college experience.

References

- Brett, C. E., Mathieson, M. L., & Rowley, A. M. (2023). Determinants of wellbeing in university students: The role of residential status, stress, loneliness, resilience, and sense of coherence. Current Psychology, 42(23), 19699-19708.
- Cobelli, N., & Cobelli, N. (2020). Innovation Management, Healthcare Challenges and eHealth. Innovation in Community-Based Private Practices Through eHealth: A Business Management Perspective, 1-20.
- McCarthy, R. V., McCarthy, M. M., Ceccucci, W., Halawi, L., McCarthy, R. V., McCarthy, M. M., ... & Halawi, L. (2022). Applying predictive analytics (pp. 89-121). Springer International Publishing.
- Roberts, D. A., Yaida, S., & Hanin, B. (2022). The principles of deep learning theory (Vol. 46). Cambridge, MA, USA: Cambridge University Press.
- Calin, O. (2020). Deep learning architectures. New York City: Springer International Publishing.
- Hahs-Vaughn, D. L., & Lomax, R. G. (2020). An introduction to statistical concepts. Routledge.
- Molnar, C. (2020). Interpretable machine learning. Lulu. com.
- Fitbit Dataset (kaggle.com),2021.

- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. Applied developmental science, 24(2), 97-140.
- Bhatti, I., Mohi-U-din, S. F., Hayat, Y., & Tariq, M. (2024). Artificial Intelligence Applications for Students with Learning Disabilities: A Systematic Review. European Journal of Science, Innovation and Technology, 4(2), 40-56.
- Olaniyi, O., Shah, N. H., Abalaka, A., & Olaniyi, F. G. (2023). Harnessing predictive analytics for strategic foresight: a comprehensive review of techniques and applications in transforming raw data to actionable insights. Available at SSRN 4635189.
- George, B., & Wooden, O. (2023). Managing the strategic transformation of higher education through artificial intelligence. Administrative Sciences, 13(9), 196.
- Salloum, S. A., Bettayeb, A., Salloum, A., Aburayya, A., Khadragy, S., Hamoudi, R., & Alfaisal, R. (2023). Novel machine learning based approach for analysing the adoption of metaverse in medical training: A UAE case study. Informatics in Medicine Unlocked, 42, 101354.
- Sarker, M. (2024). Revolutionizing Healthcare: The Role of Machine Learning in the Health Sector. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 2(1), 35-48.
- Gardanova, Z., Belaia, O., Zuevskaya, S., Turkadze, K., & Strielkowski, W. (2023, July). Lessons for medical and health education learned from the COVID-19 pandemic. In Healthcare (Vol. 11, No. 13, p. 1921). MDPI.