

AI Women's Health Analyzer Using Machine Learning

Pilla Sasi Prabha, Mrs. T. Siva Kumari

Department of MCA Student, Aditya Degree & PG College, Vizianagaram, Andhra Pradesh, India

Assistant Professor, Department of MCA, Aditya Degree & PG College, Vizianagaram, Andhra Pradesh, India

ABSTRACT: The AI Women's Health Analyzer is an intelligent healthcare application designed to assist women in monitoring and analyzing their health conditions using machine learning techniques. The system uses Natural Language Processing (NLP) and Machine Learning algorithms such as TF-IDF and Naive Bayes to predict possible health conditions based on user symptoms. The application also includes BMI analysis, menstrual cycle tracking, appointment booking, health history management, and PDF report generation. The system is developed using Python, Tkinter, SQLite, and Scikit-learn, providing a secure, offline, and user-friendly healthcare environment. The proposed system improves healthcare accessibility, promotes preventive healthcare awareness, and reduces dependency on manual diagnosis by offering personalized health recommendations and intelligent disease prediction.

KEYWORDS: Machine Learning, Women's Health, Naive Bayes, TF-IDF, Python, Healthcare Analytics, SQLite, Tkinter.

I. INTRODUCTION

Healthcare technology has rapidly evolved with the integration of Artificial Intelligence (AI) and Machine Learning (ML). These technologies help in analyzing large amounts of healthcare data and generating intelligent predictions that support early diagnosis and preventive healthcare. Women's healthcare is one of the most critical areas requiring personalized monitoring and timely medical guidance.

Many women face difficulties in accessing proper healthcare services due to financial, social, and geographical barriers. Traditional healthcare systems mainly depend on manual diagnosis, physical consultations, and delayed medical reports. Existing systems often lack personalized recommendations and intelligent health analysis features.

The AI Women's Health Analyzer is developed to address these challenges by providing an intelligent healthcare platform capable of predicting diseases based on symptoms entered by the user. The application works completely offline, ensuring privacy and accessibility while storing all data securely in a local SQLite database.

The system integrates machine learning algorithms, BMI monitoring, menstrual cycle tracking, appointment scheduling, health report generation, and personalized recommendations within a single desktop application. The proposed system helps improve women's healthcare awareness, preventive healthcare practices, and efficient health record management.

II. LITERATURE SURVEY

Machine Learning and Artificial Intelligence have become increasingly important in the healthcare industry. Several intelligent healthcare systems have been developed for disease prediction, symptom analysis, and patient monitoring.

Traditional healthcare applications mainly focus on specific features such as fitness tracking, appointment scheduling, or BMI monitoring. However, many of these systems lack integrated machine learning support and personalized recommendations. Existing systems also often require internet connectivity and cloud-based services, which may affect user privacy and accessibility.

Research studies show that machine learning algorithms such as Naive Bayes, Decision Trees, Random Forest, and Support Vector Machines are widely used for disease prediction and healthcare analytics. Among these, Naive Bayes

performs efficiently for text classification and symptom-based disease prediction due to its simplicity and high accuracy.

Natural Language Processing (NLP) techniques such as TF-IDF are commonly used to convert text-based symptoms into numerical vectors that machine learning models can process efficiently. Combining TF-IDF with Naive Bayes improves prediction accuracy for healthcare applications.

The proposed AI Women's Health Analyzer combines healthcare management features, machine learning prediction, offline accessibility, and secure local data storage into a single user-friendly platform.

III. METHODOLOGY

The AI Women's Health Analyzer follows a three-tier desktop application architecture consisting of:

1. Presentation Layer
2. Application Logic Layer
3. Data Layer

The Presentation Layer is developed using Python Tkinter and provides a graphical user interface for user interaction. The Application Logic Layer contains all healthcare processing logic, machine learning models, BMI calculations, cycle prediction algorithms, and report generation modules. The Data Layer uses SQLite database for storing user information, health records, appointments, and feedback.

The system workflow begins with user registration and login authentication. After successful login, the user enters health details such as age, symptoms, weight, height, and lifestyle information.

The entered symptoms are processed using the TF-IDF vectorization technique. TF-IDF converts textual symptom descriptions into numerical vectors that can be analyzed by machine learning models.

The Multinomial Naive Bayes classifier is then used to predict possible diseases based on the symptom vectors. The model calculates the probability of each disease and predicts the most relevant health condition.

The system also calculates Body Mass Index (BMI) using the standard BMI formula:

$$BMI = \frac{Weight (kg)}{Height^2 (m^2)}$$

Based on the BMI value, the user is categorized as Underweight, Normal, Overweight, or Obese.

The menstrual cycle tracking module predicts the next period date, ovulation date, and fertile window using date arithmetic calculations.

Finally, the system generates personalized health recommendations, diet plans, exercise suggestions, and downloadable PDF reports.

IV. SYSTEM MODULES

A. User Authentication Module

This module manages user registration and login operations. User passwords are encrypted using SHA-256 hashing before storage in the database. The module validates input fields and prevents unauthorized access.

B. Dashboard Module

The dashboard acts as the main interface of the application. It displays user information, health summaries, appointment counts, BMI values, and navigation options.

C. Health Analysis Module

This module processes user symptoms using machine learning algorithms and predicts possible diseases. The system also provides detailed recommendations including precautions, lifestyle changes, medicines, and specialist doctor suggestions.

D. BMI Calculator Module

The BMI module calculates Body Mass Index using height and weight values. The result is displayed with category classification and personalized health tips.

E. Menstrual Cycle Tracker Module

This module helps users track menstrual cycle information including start date, cycle duration, ovulation prediction, and fertile window estimation.

F. Appointment Booking Module

Users can schedule and manage doctor appointments through this module. Appointment details are stored securely in the database.

G. Health History and Export Module

This module stores past health analysis records and allows users to export reports in PDF and CSV formats for future reference.

V. ALGORITHM USED

A. TF-IDF Vectorization

TF-IDF (Term Frequency – Inverse Document Frequency) is used to convert symptom text into numerical vectors.

$$TF-IDF = TF \times IDF$$

TF measures how frequently a term appears in a document, while IDF measures the importance of the term across all documents.

Advantages:

- Converts text into machine-readable format
- Improves prediction accuracy
- Reduces importance of common words

B. Naive Bayes Algorithm

Naive Bayes is a probabilistic machine learning algorithm based on Bayes' theorem.

$$P(C | X) = \frac{P(X | C)P(C)}{P(X)}$$
$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \approx 0.68, P(B) \approx 0.25$$

$P(B)=0.25P(B|A)P(A)=0.17P(A|B)\sim 0.68$ Posterior = useful evidence / total evidence

Where:

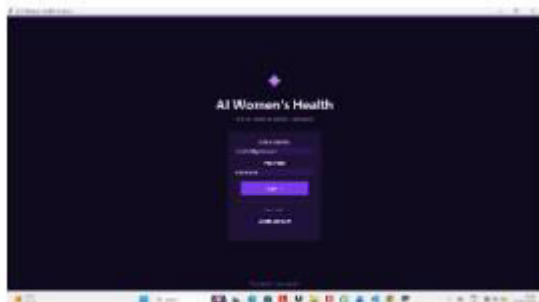
- $P(C | X)$ = Probability of disease given symptoms
- $P(X | C)$ = Probability of symptoms given disease
- $P(C)$ = Prior probability of disease
- $P(X)$ = Probability of symptoms

Advantages:

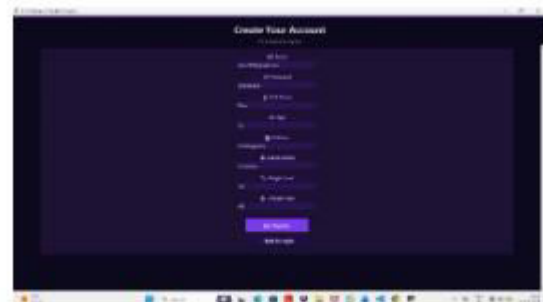
- Fast and efficient
- Requires less training data
- Works well for text classification problems

VI. RESULTS AND DISCUSSION

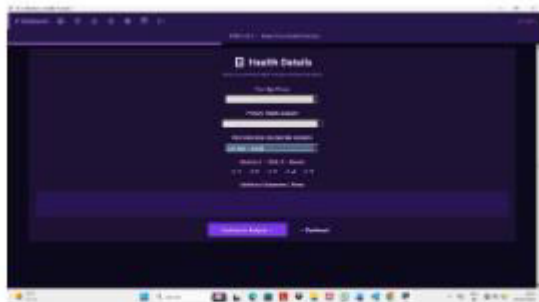
- The AI Women's Health Analyzer successfully predicts women's health conditions based on symptom inputs. The system provides fast and accurate health analysis results along with personalized recommendations.
- The BMI module correctly categorizes users into appropriate BMI ranges and generates health suggestions accordingly. The menstrual cycle tracker predicts ovulation and fertile periods effectively.
- The system securely stores user records and generates downloadable PDF reports. The graphical interface improves usability and user interaction.
- The machine learning model demonstrates efficient classification performance for healthcare prediction tasks and improves preventive healthcare awareness among users.



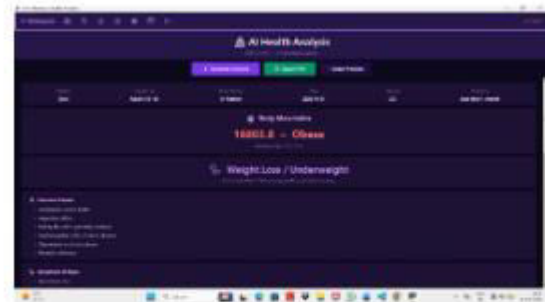
(a)



(b)



(c)



(d)

VII. ADVANTAGES OF THE SYSTEM

- Early disease prediction
- User-friendly graphical interface
- Secure local data storage
- Offline accessibility
- Personalized health recommendations
- PDF health report generation
- Reduced dependency on manual diagnosis
- Efficient healthcare record management

VIII. LIMITATIONS

- Prediction accuracy depends on dataset quality
- Not a replacement for professional medical advice
- Limited to predefined healthcare conditions
- Desktop-only implementation

IX. FUTURE ENHANCEMENTS

- Cloud-based data storage
- Mobile application development
- Integration with wearable devices
- Real-time doctor consultation
- Improved machine learning models
- Online healthcare connectivity

X. CONCLUSION

The AI Women's Health Analyzer demonstrates the successful integration of machine learning techniques into healthcare applications for women. The system provides intelligent disease prediction, BMI analysis, menstrual cycle tracking, appointment management, and personalized healthcare recommendations through a secure offline platform.

The application improves healthcare accessibility, preventive healthcare awareness, and user engagement through a simple graphical interface and intelligent healthcare analysis. Future enhancements such as cloud integration, mobile support, and wearable device connectivity can further improve the efficiency and usability of the system.

REFERENCES

- [1] Tom Mitchell, *Machine Learning*, McGraw Hill, 1997.
- [2] Scikit-learn Documentation for Machine Learning Algorithms.
- [3] Python Software Foundation Documentation.
- [4] Research papers on AI-based healthcare systems.
- [5] Studies on TF-IDF and Naive Bayes classification methods.
- [6] Healthcare analytics and predictive healthcare research articles.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details