RESEARCH ARTICLE OPEN ACCESS

Disease Prediction using Symptoms with Remedies

Raj Mulik, Henil Ahire, Sahil Pimpalshende, Riya Polade,
Department of Computer Engineering ,Smt. Kashibai Navale College of Engineering Pune, India
Prof. Sneha Patil ,
Department of Computer Engineering, Smt. Kashibai Navale College of Engineering , Pune,India

Abstract:

This project focuses on leveraging machine learning techniques to develop accurate, efficient, and accessible disease prediction models, thereby revolutionizing healthcare delivery and improving patient outcomes. The project begins with a comprehensive data collection process, gathering a diverse range of patient data, including medical history, genetic information, lifestyle factors, and environmental variables. A dataset of this magnitude is meticulously pre-processed, cleaned, and standardized to ensure data quality. Features are extracted, and dimensionality reduction techniques are applied to enhance the computational efficiency of the predictive models. Various machine learning algorithms, including but not limited to decision trees, random forests, support vector machines, and deep learning models, are employed to analyse the data and build predictive models. These models are trained and validated using advanced techniques such as cross-validation and hyperparameter tuning to ensure robustness and reliability.By employing this cutting-edge approach to healthcare, the project aims to significantly reduce the burden of disease, enhance early detection, and promote preventative measures, ultimately leading to improved patient care and a more efficient healthcare system. Disease Prediction Using Machine Learning is poised to revolutionize the medical field, providing a proactive, data-driven approach to health management and disease prevention.

INTRODUCTION

The "Disease Prediction Using Machine Learning" project is an innovative initiative designed to leverage the capabilities of machine learning and data analytics in healthcare. Its primary objective is to develop accurate and accessible disease prediction models, enabling early detection, prevention, and improved patient care. "Disease Prediction Using Machine Learning" represents a transformative approach to healthcare, offering a data-driven strategy to enhance patient care, prevent disease, and ultimately improve the efficiency of the healthcare system. By combining the power of machine learning with comprehensive healthcare data, the project aims to make a lasting impact on the field of medicine and patient well-being.

LITERATURE SURVEY

In the study "Disease Prediction using Machine Learning" published in the International Research Journal of Engineering and Technology (IRJET), the authors, Raj H. Chauhan, Daksh N. Naik, Rinal A. Halpati, Sagarkumar J. Patel, along with Mr. A.D. Prajapati, have delved into the critical realm of disease prediction through machine learning. Their system leverages

predictive modeling, utilizing a Decision Tree Classifier to calculate the probability of diseases based on user-provided symptoms. The significance of this work lies in its application within the healthcare sector, where early disease detection and patient care are paramount. The study highlights how machine learning, with its training and testing phases, offers a powerful platform for improving healthcare services and enhancing patient diagnoses, positioning it as a transformative technology in the medical field.

The study's objective is to develop a user-friendly system capable of predicting diseases without requiring a physical visit to a physician. This system, referred to as 'AI THERAPIST,' integrates machine learning to boost prediction accuracy and analyze both structured and unstructured data. The authors, affiliated with the Department of Computer Engineering at R.N.G. Patel Institute of Technology in Gujarat, India, underscore the limitations of traditional medical approaches and advocate for the adoption of medical decision support systems. By leveraging machine learning for rapid data analysis, this technology empowers doctors to make informed decisions, leading to improved patient healthcare services. The article provides a comprehensive overview of its data sources and the algorithms employed, particularly focusing on the gain ratio decision tree, while also mentioning other decision tree variants. The results indicate that the Random Forest algorithm outperforms others and concludes by emphasizing the system's potential in supporting individuals concerned about their health and contributing to accurate disease prediction based on symptoms, even extending its impact to mental health issues like depression and anxiety.

DESIGN AND ANALYSIS MODEL

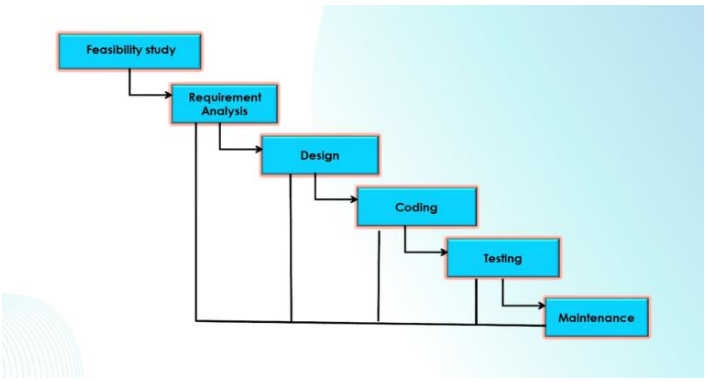


Figure 1 : Iterative Waterfall Model

The Iterative Waterfall Model refines the traditional Waterfall approach by introducing iterative cycles between its well-defined phases. While maintaining the structured progression of the Waterfall Model, these iterations allow for ongoing refinement and adaptation to evolving software requirements. This model strikes a balance between the rigidity of a linear process and the flexibility required to address changing project needs, making it a valuable methodology for projects with dynamic requirements. The iterative feedback loops provide a mechanism for continuous improvement, enhancing the adaptability of the development process throughout the software development life cycle.

PROPOSED ARHITECTURE

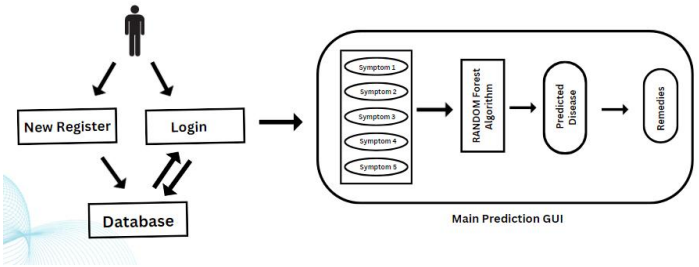


Figure 2 : Entire Working Model

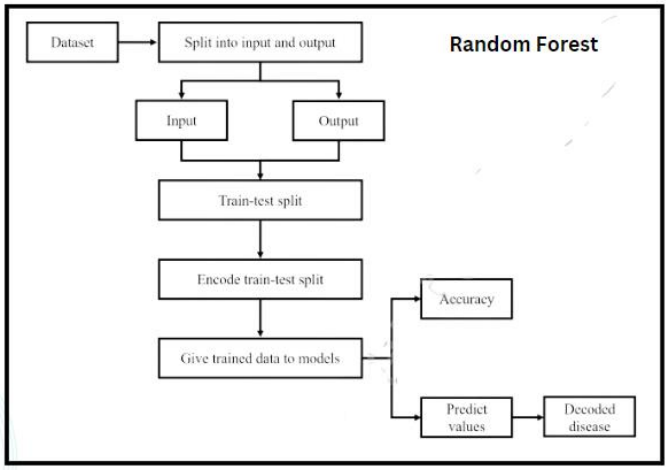


Figure 3 : Random Forest Algorithm

CODE IMPLEMENTATION

Disease Prediction Using machine learning involves leveraging various algorithms and techniques to analyze medical data and predict the likelihood of a person having a particular disease or condition.

TOTAL OF 132 SYMPTOMS USED FOR PREDICTIONS

```
l1=['itching','skin_rash','nodal_skin_eruptions','continuous_sneezing','shivering','chills','joint_pain',
'stomach_pain','acidity','ulcers_on_tongue','muscle_wasting','vomiting','burning_micturition','spotting_ urination','fatigue',
'weight_gain','anxiety','cold_hands_and_feets','mood_swings','weight_loss','restlessness','lethargy','patches_in_throat',
'irregular_sugar_level','cough','high_fever','sunken_eyes','breathlessness','sweating','dehydration','indigestion',
'headache','yellowish_skin','dark_urine','nausea','loss_of_appetite','pain_behind_the_eyes','back_pain','constipation',
'abdominal_pain','diarrhoea','mild_fever','yellow_urine','yellowing_of_eyes','acute_liver_failure','fluid_overload',
'swelling_of_stomach','swelled_lymph_nodes','malaise','blurred_and_distorted_vision','phlegm','throat_irritation',
'redness_of_eyes','sinus_pressure','runny_nose','congestion','chest_pain','weakness_in_limbs','fast_heart_rate',
'pain_during_bowel_movements','pain_in_anal_region','bloody_stool','irritation_in_anus','neck_pain','dizziness','cramps',
'bruising','obesity','swollen_legs','swollen_blood_vessels','puffy_face_and_eyes','enlarged_thyroid','brittle_nails',
'swollen_extremities','excessive_hunger','extra_marital_contacts','drying_and_tingling_lips','slurred_speech','knee_pain','hip_joint_pain',
'muscle_weakness','stiff_neck','swelling_joints','movement_stiffness','spinning_movements','loss_of_balance','unsteadiness','weakness_of_one_body_side',
'loss_of_smell','bladder_discomfort','foul_smell_of urine','continuous_feel_of_urine','passage_of_gases','internal_itching','toxic_look_(typhos)',
'depression','irritability','muscle_pain','altered_sensorium','red_spots_over_body','belly_pain','abnormal_menstruation','dischromic _patches',
'watering_from_eyes','increased_appetite','polyuria','family_history','mucoid_sputum','rusty_sputum','lack_of_concentration','visual_disturbances',
'receiving_blood_transfusion','receiving_unsterile_injections','coma','stomach_bleeding','distention_of_abdomen','history_of_alcohol_consumption',
'fluid_overload','blood_in_sputum','prominent_veins_on_calf','palpitations','painful_walking','pus_filled_pimples','blackheads','scurring','skin_peeling',
'silver_like_dusting','small_dents_in_nails','inflammatory_nails','blister','red_sore_around_nose','yellow_crust_ooze']
```

TOTAL OF 41 DISEASE USED FOR PREDICTION BASED ON SYMPTOMS

```
disease=['Fungal infection','Allergy','GERD','Chronic cholestasis','Drug Reaction',
         'Peptic ulcer diseae','AIDS','Diabetes','Gastroenteritis','Bronchial Asthma','Hypertension',
         ' Migraine','Cervical spondylosis',
         'Paralysis (brain hemorrhage)','Jaundice','Malaria','Chicken pox','Dengue','Typhoid','hepatitis A',
'Hepatitis B','Hepatitis C','Hepatitis D','Hepatitis E','Alcoholic hepatitis','Tuberculosis',
'Common Cold','Pneumonia','Dimorphic hemmorhoids(piles)',
'Heartattack','Varicoseveins','Hypothyroidism','Hyperthyroidism','Hypoglycemia','Osteoarthristis',
'Arthritis','(vertigo) Paroymsal  Positional Vertigo','Acne','Urinary tract infection','Psoriasis',
'Impetigo']
```

RANDOM FOREST ALGORITHM USED FOR PREDICTION

The Random Forest (RF) classifier is a powerful machine learning algorithm employed in the field of predictive modeling and classification. In the provided code snippet, a Random Forest Classifier is implemented using the RandomForestClassifier class from the popular scikit-learn library.

The classifier is initialized with 250 decision trees (n_estimators=250) and a fixed random state for reproducibility (random_state=42). The fit method is then utilized to train the model on the input data (X) and corresponding labels (y), where X represents the features and y is the target variable.

Following the training phase, the classifier's predictive performance is evaluated on a test set (X_test) using the predict method. The accuracy of the model is then calculated by comparing the predicted labels (y_pred) with the actual labels of the test set (y_test). The accuracy score is printed, indicating the proportion of correctly classified instances.

Furthermore, the code includes a mechanism for obtaining input symptoms (psymptoms) from the user interface. These symptoms are then matched with a predefined list of symptoms (l1), and a corresponding binary feature vector (l2) is created, where a value of 1 indicates the presence of a symptom.

This Random Forest classifier demonstrates robust classification capabilities, making it suitable for a variety of applications, including medical diagnosis based on symptom analysis. Its ensemble of decision trees enables efficient handling of complex relationships within the data, contributing to its effectiveness in predictive modeling scenarios. The code snippet, therefore, encapsulates the essential steps of training, evaluation, and symptom-based prediction within the context of the Random Forest algorithm.

RANDOM FOREST CODE SNIPPET:

```
def RF_classifier():
    from sklearn.ensemble import RandomForestClassifier
    clf = RandomForestClassifier(n_estimators=250, random_state=42)
    clf.fit(X, np.ravel(y))
    from sklearn.metrics import accuracy_score
    y_pred = clf.predict(X_test)
    print(accuracy_score(y_test, y_pred))
    #print(accuracy_score(y_test, y_pred, normalize=False))
    print("Accuracy : ",accuracy_score(y_test, y_pred)*100)
    #print("Accuracy Score : ",accuracy_score(y_test, y_pred, normalize=False))

    psymptoms = [Symptom1.get(),Symptom2.get(),Symptom3.get(),Symptom4.get(),Symptom5.get()]

    for k in range(0,len(l1)):
        for z in psymptoms:
            if(z==l1[k]):
                l2[k]=1

    inputtest = [l2]
    predict = clf.predict(inputtest)
    predicted=predict[0]

    h='no'
    for a in range(0,len(disease)):
        if(disease[predicted] == disease[a]):
            h='yes'
            break

    if (h=='yes'):
        t3.delete("1.0", END)
        t3.insert(END, disease[a])
        global prectictedDisease
        prectictedDisease=disease[a]
        search_remedy()
    else:
        t3.delete("1.0", END)
        t3.insert(END, "No Disease")
```

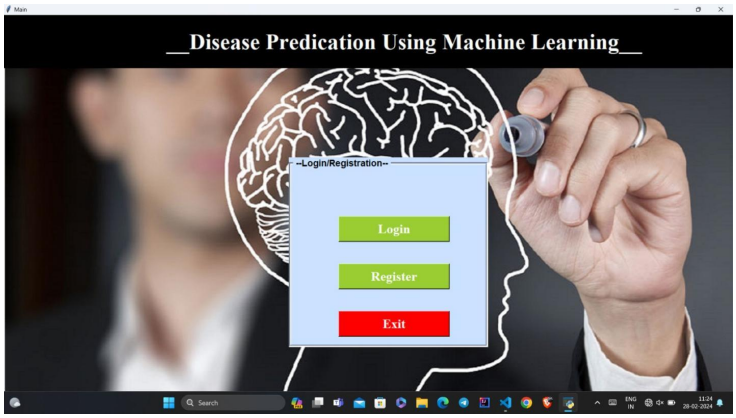
CREATION OF GUI USING PYTHON TKINTER

```
root = Tk()
root.geometry('2000x1000')

# Set background image
background_image = PhotoImage(file="newcropedimg.png")  # Replace with your image path
background_label = Label(root, image=background_image)
background_label.place(relwidth=1, relheight=1)

root.title(" Disease Prediction and Remedies From Symptoms")
root.configure()

# Change color schemes
root.config(bg="lightblue")  # Set the background color of the root window

w2 = Label(root, justify=CENTER, text=" Disease Prediction and Remedies From Symptoms ", bg="lightblue")  # Set background color
# w2 = Label(root, justify=CENTER, text=" Disease Prediction and Remedies From Symptoms ", bg="lightblue")  # Set background color
w2.config(font=("Monospace", 30))
w2.grid(row=1, column=0, columnspan=2, padx=50,pady=50)  # Adjusted columnspan and pady
# w2.pack(expand=True, fill='both', side='top', anchor='center')

Namelb1 = Label(root, text="Enter Your Symptoms: ")
Namelb1.config(font=("Elephant", 20))
Namelb1.grid(row=5, column=1, pady=10, sticky=W)

Namelb1 = Label(root, text="(Atleast Enter 3 Symptoms)")
Namelb1.config(font=("Elephant", 9))
Namelb1.grid(row=6, column=1, pady=10, sticky=W)

Symptom1 = StringVar()
Symptom1.set(None)
Symptom2 = StringVar()
Symptom2.set(None)
Symptom3 = StringVar()
Symptom3.set(None)
Symptom4 = StringVar()
Symptom4.set(None)
Symptom5 = StringVar()
```

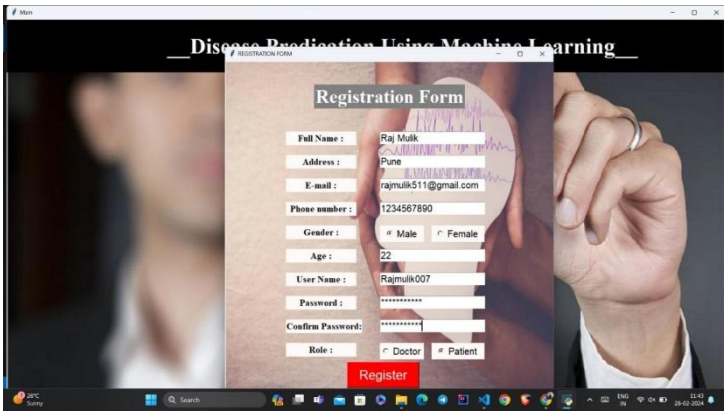
WORKING OF REMEDIES SUGGESTION BASED ON PREDEFINED REMEDIES FOR ALL 41 DISEASES

```
def search_remedy():
    disease = prectictedDisease.strip().title()

    # Use str.contains for a case-insensitive search
    mask = gf['Disease'].str.contains(f'\\b{disease}\\b', case=False, regex=True)

    if mask.any():
        remedy = gf.loc[mask, 'Remedies'].values[0]
        result_label.config(text=f"For {disease} : {remedy}",font=("Elephant", 15))
        #result_label.config(text=f"Remedies for {disease} are : {remedy}",font=("Elephant", 12))
```

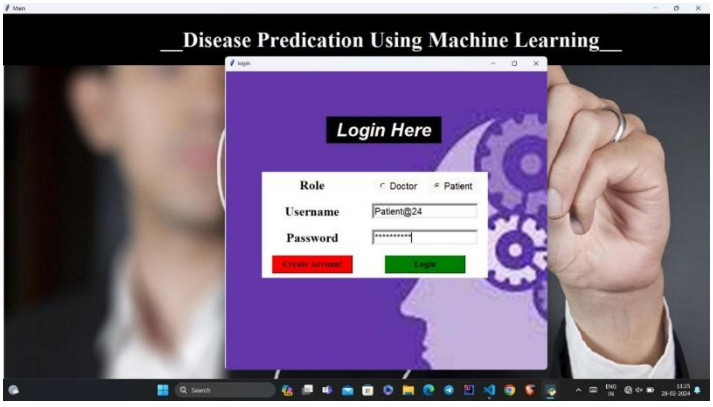
GUI RESULTS :



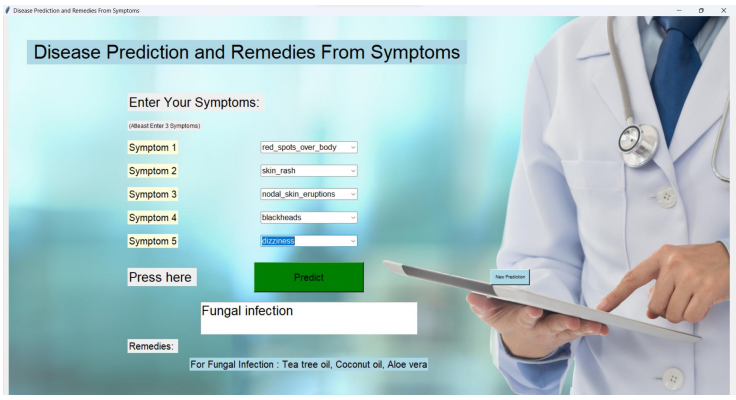
REGISTRATION FORM



LOGIN GUI :



PREDICTION PAGE:



APPLICATIONS

Machine learning applications in disease prediction are revolutionizing the healthcare industry. By harnessing the power of algorithms and data analysis, machine learning is transforming early disease detection, risk assessment, and personalized medicine. It plays a pivotal role in analyzing medical records, diagnostic images, and genetic data, enabling healthcare professionals to detect diseases at their earliest stages, stratify patient risks, and tailor treatments to individual needs. This technology also enhances the accuracy and efficiency of medical imaging analysis, assists in drug discovery, and contributes to proactive public health measures by predicting disease outbreaks. Furthermore, machine learning facilitates remote patient monitoring and fraud detection while empowering epidemiological studies to identify health trends and track the spread of diseases. Overall, the applications of machine learning in disease prediction are shaping a future of more precise, efficient, and personalized healthcare, improving patient outcomes and public health management.

CONCLUSION

Our project has the potential to improve health treatments by predicting proper disease with the help of symptoms.It will also help to reduce the burden of healthcare system, and promote a healthy approach to life with the use of Machine Learning.India's population puts more burden on healthcare system so implementation of our project model will definitely reduces the excessive load of our healthcare infrastructure.With the right expertise, resources,and strong commitment to transparency and ethics, disease prediction through machine learning has the potential to

revolutionize healthcare, making it more proactive and efficient.

REFERENCES

1. Disease Prediction using Machine Learning Algorithms, 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020
2. Disease Prediction using Machine Learning ,International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
3. Disease prediction from various symptoms using machine IEEE papers IEEE PapersSSRN-id3661426
4. THE PREDICTION OF DISEASE USING MACHINE LEARNING,International Journal of Scientific Research in Engineering and Management (IJSREM)
5. Disease Prediction using Machine Learning ,IJCRT2105229 International Journal of Creative Research Thoughts (IJCRT)
6. Common human diseases prediction using machine learning based on survey data,Bulletin of Electrical Engineering and Informatics pp. 1~1x ISSN: 2302-9285, DOI: 10.11591/eei.v11i6.3405
7. Disease prediction based on symptoms. International Journal of Creative Research Thoughts (IJCRT) arXiv:2307.08290v
8. Identification and Prediction of Chronic Diseases Using Machine Learning Approach,Hindawi Journal of Healthcare Engineering Volume 2022,Article ID 2826127