Journal of Nonlinear Analysis and Optimization Vol. 14, Issue. 2, No. 2: 2023 ISSN : **1906-9685**



THE PREDICTION OF DISEASE USING MACHINE LEARNING WITH ALGORITHMS

Ms.G.Aswinipriya, M.sc., Assistant Professor, Department of Computer Science, Marudhar Kesari Jain College for Women, Tamilnadu. aswiniindira98@gmail.com

Abstract

Disease Prediction using Machine Learning is the system that is used to predict the diseases from the symptoms which are given by the patients or any user. The system processes the symptoms provided by the user as input and gives the output as the probability of the disease. Naïve Bayes classifier is used in the prediction of the disease which is a supervised machine learning algorithm. The probability of the disease is calculated by the Naïve Bayes algorithm. With an increase in biomedical and healthcare data, accurate analysis of medical data benefits early disease detection and patient care. By using linear regression and decision tree we are predicting diseases like Diabetes, Malaria, Jaundice, Dengue, and Tuberculosis.

Keywords: Disease Prediction, Machine learning, Decision Tree, Random Forest, Naive bayes algorithm.

1.Introduction

Machine Learning is the domain that uses past data for predicting. Machine Learning is the understanding of computer system under which the Machine Learning model learn from data and experience. The machine learning algorithm has two phases: 1) Training & 2) Testing. To predict the disease from a patient's symptoms and from the history of the patient, machine learning technology is struggling from past decades. Healthcare issues can be solved efficiently by using Machine Learning Technology. We are applying complete machine learning concepts to keep the track of patient's health.

ML model allows us to build models to get quickly cleaned and processed data and deliver results faster. By using this system doctors will make good decisions related to patient diagnoses and according to that, good treatment will be given to the patient, which increases improvement in patient healthcare services. To introduce machine learning in the medical field, healthcare is the prime example. To improve the accuracy of large data, the existing work will be done on unstructured or textual data. For the prediction of diseases, the existing will be done on linear, Decision Tree algorithm.

2. ALGORITHM TECHNIQUES

Decision Tree:

A decision tree is a structure that can be used to divide up a large collection of records into successfully smaller sets of records by applying a sequence of simple decision tree. With each successive division, the members of the resulting sets become more and more similar to each other. A decision tree model consists of a set of rules for dividing a large heterogeneous population into smaller, more homogeneous (mutually exclusive) groups with respect to a particular target.

Random Forest

It will find user given symptoms to predict the diease using Random forest algorithms The purpose of the activity is to find the user given symptoms to predict the disease using random forest. The dataset

84

will be stored in the excel sheet for analysis purpose and loaded to the random forest program for training and testing the model.

Naive Bayes:

Naive Bayes is an easy however amazingly powerful rule for prognosticative modeling. The independence assumption that allows decomposing joint likelihood into a product of marginal likelihoods is called as 'naive'. This simplified Bayesian classifier is called as naive Bayes. It will find user given symptoms to predict diease using naïvebayes algorithm's give accurate prediction. The purpose of the activity is to find the user given symptoms to predict diease using Naïve bayes Algorithm and predict to give accurate results. The datasets will be stored in the excel sheet for analysis purpose and loaded to the Naïve Bayes program.

Train Data:

THE.										read-read											
	Marine .		Page Loop	e no		H 14		en landt	es Ter										Series 3	a there	
- 1	tie -	633		144	200	-		-	1	1-1-1-		16	U HEY	100	22	2017	XAN	n A	- Ö		
-Cin	Crev -	100		- 11 -	1.1	- 36		1. 11 B. St.		Contrast.		19.5		1	ET (Tite-	- 3	T PC		
10.0	Round Pale	. *	3.8 - 1	199	4	6 X X	10.10	11 Mage & C	800 C	1.15.1	25.72	Calutta	nat Formation	Gell.	nor b	ate ferrer	# Char	14	Th Faild		
1.0		5	10		-		10400			Test to a			97400	17.55	1.0	10		1004	2003.000		
				10. j. j. a. a. a.																	
a.			v . A	uting .																	-
100	110	100	: 10	0.91118		#10110G	00	H 1	011 0	1 K	11		M. 11		0	P. 13	0 1	0.13	 115 	Ť.	10.3
(EA)	akin ra	en node	Likk contra	web. shive	ring shill	s jekt	t pairstee	mach_jacidili	y storm	, on muscle	C WYORK	ing burn	ing_reports	ng_ tatip	sae wee	pt_gans	iety mid	hancmo	od_sorwei	gre kans	stel.
	-	1	£		0	-8-	0	0		0	8-	0		8-	0	- 5-	0		D	0	
-	- C	2-	-		0.	-2-	- 2-		-2	- 00	-			÷.	- 0				0		
		÷		-	0		n		-2		-	0		÷2	0			100			
	1	1	L		0		0	0	0	0.		.0.	0		0		0	1	0	0	
		1	1		0		0	0	0	0		D	0	0	0		0		Ð	0	
	1	0	1		0		0		0	α		D	0	Ð	0		0	1	0	0	
	1	1	đ		0		0	0	0	α		0	0	0	a		0		0	0	
-	-5	1-	1		0	-8-	0	0	0	0	-	D	0	8-	0		0		0	0	
-	-			-	1	-11-	0	0	2	0	2	0		10	0	1	0	1	0	0	
		6	a		1	1	0	0	2	a	-	B		.0	0		0		D	0	
			à	1	0	1	0.	0		0		0			.0		0		D.	0	
		0	ġ	1	4		0	0	8	0.		0		1	0		0		0	. 0	
		0	0	1	1	1	0	0		0.		0	. 0	.0	0		0	. 8	0	.0	
-		0	0		1	1	0	0		0.		0			0		0		0	0	
-		0	0	-	0	-	0	0	-	0	- C	8		÷.	0		0		0		
		6	0		1	-10-	0	0	1	0	÷.	10		2	0	-	0	1	0		
		0	0	1	i.	1	0	a		a	-		6	2	0	- 6	0	1	0	0	
	_					- 21									.0				0	0	
		0	0		0.		.0,										M				
ally and	Tre	e eining	a d (†)	1	ů ů	-	0	î.	2	î	1	î	a 4		a	1		₩900 11 1 10 1	. 1	.0	10.10
'es	t Da		a d (†)	:	0	0	0	e a	2	0	:	1	े 	1 (2	0	1			i o ei t	0 	1 1 1 1
'es	t Da	ta			0	0	0	: e @	2			1	- -	1	ă	1		14 juli 18 +	001		2 Z 1
'es	t Da	ita			0	0	0	: • •				1			a 23 1 2	:				0	I I I I
`es	t Da	ita			0 0			1 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	1 18	L L L L L L L L L L L L L L L L L L L		1									IC IN IC
`es	t Da					0			1								R A MEY				
es	t Da		a a a a a a a a a a a a a a a a a a a		ο α	0					-				a a a a a a a a a a a		N I I I		1 10 (11) 27 (2)	2	
'es	t Da	ita		- 11					1 88 1 1 1 1 1 1 1	Tanaga I Sanga						Table For					
`es	t Da	ita						1 2 2 2 2 2 2 2 2 2 2 2 2 2	1 	Tanaga I Januar Tanaga I Sanaga Sanaga Sanagan						Table For					
es	t Da	ita		8 8 - [1] - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2				1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2						E - Coli							
es	t Da	ta																			
ſes	t Da																Mit M				
es	t Da																				
es	t Da	ita															MEN ALL				
'es	t Da	ta		E E E E E E E E E E E E E E E E E E E													Mere and a second secon				
'es	t Da	o c c ta															Wer a				
es	t Da															Training for the second					
es		ta																			
es		ta																			
es		ta																			
`es		ta																		0	
`es		e ining																			
es																				0	
es		ta								Nerve 1											

RESULTS

- Comparison of accuracy of algorithm.
- Decision Tree 84.5%
- Random Forest 98.95%
- ➢ Naïve Bayes 89.4%



We found that the Support Vector Machine (SVM) algorithm is widely used followed by the Naïve Bayes algorithm (in 24 studies). However, the Random Forest algorithm showed relatively high accuracy. in which it was used, RF showed the highest accuracy of 98.95%. This was followed by SVM which included 96% of the accuracy considered.

CONCLUSION

The main aim of this disease prediction system is to predict the disease on the basis of the symptoms. This system takes the symptoms of the user from which he or she suffers as input and generates final output as a prediction of disease. Average prediction accuracy probability of 100% is obtained. Disease Predictor was successfully implemented using the grails framework. This system gives a user-friendly environment and easyto use.

As the system is based on the web application, the user can use this system from anywhere and at any time. In conclusion, for disease risk modeling, the accuracy of risk prediction depends on the diversity feature of the hospital data.

This systematic review aims to determine the performance, limitations, and future use of Software in health care. Findings may help inform future developers of Disease Predictability Software and promote personalized patient care. The program predicts Patient Diseases. Disease Prediction is done through User Symbols. In this System Decision tree, Unplanned Forest, the Naïve Bayes Algorithm is used to predict diseases. Forthe data format, the system uses the Machine Learning algorithm Process Data on Database Data namely, Random Forest, Decision Tree, Naive Bayes. System accuracy reaches 98.3%. machine learning skills are designed to successfully predict outbreaks.

References

1. D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, no G. Escobar, "Big data for health care: using analytics to identify and treat high-risk and high-risk patients, Health, vol. 33, no. 7, pages 1123–1131, 2014.

2. K.R. Lakshmi, Y. Nagesh and Mr. Veera Krishna, "Comparison of performance of the three data mines ways to predict survival of kidney disease", International Journal of Engineering Development & Technology, March 2014.

3. Dr.C K Gomathy, Article: A Semantic Quality of Web Service Information Retrieval Techniques Using Bin Rank A Cloud Monitoring Framework Perform in Web Services, International Journal of Scientific Research in Computer Science Engineering and Information Technology IJSRCSEIT | Volume 3 | Issue 5 |ISSN : 2456-3307,May-2018

4. Dr.C K Gomathy, Article: Supply chain-Impact of importance and Technology in Software Release Management, International Journal of Scientific Research in Computer Science Engineering and Information Technology (IJSRCSEIT) Volume 3 | Issue 6 | ISSN : 2456-3307, P.No:1-4, July-2018

5. Dr.C K Gomathy, Article: A Scheme of ADHOC Communication using Mobile Device Networks, International Journal of Emerging technologies and Innovative Research (JETIR) Volume 5 | Issue 11 | ISSN : 2349-5162, P.No:320-326, Nov-2018