IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 9, September 2021

DOI: 10.17148/IARJSET.2021.8973

A Study on the Analysis of Real-Time Medical Diagnosis using Machine Learning

Dillip Narayan Sahu[1]

[1]Lecturer, Department of MCA, Gangadhar Meher University (GMU), India

The biggest problem before the world in the field of Artificial Intelligence is to create a machine that exactly behaves like a human being. A system capable of the autonomous acquisition and integration of knowledge is so called Machine Learning. Learning simply means the changes in a particular system that enable another system to do the same task more effectively and efficiently the next time. In this paper, the medical diagnosis in real time system has been analyzed and shown a clear picture by the help of machine learning tool.

Keywords: Algorithm, Classifier, Decision Tree, Machine Learning, Weka.

1. INTRODUCTION

Learning is an ongoing process. Learning is the constructing or modifying some representations of what is being experienced. –By, Ryszard Michalski
Learning is making useful changes in our minds. –By, Marvin Minsky
Machine learning is mainly concerned with the accuracy and effectiveness of the any system to gives proper result[1][2].

1.1 Types of Machine Learning-

Basically there are three types of Machine Learning

1. Supervised Learning (labeled examples)
2. Unsupervised Learning (unlabeled examples)
3. Reinforcement Learning (reward)[3]

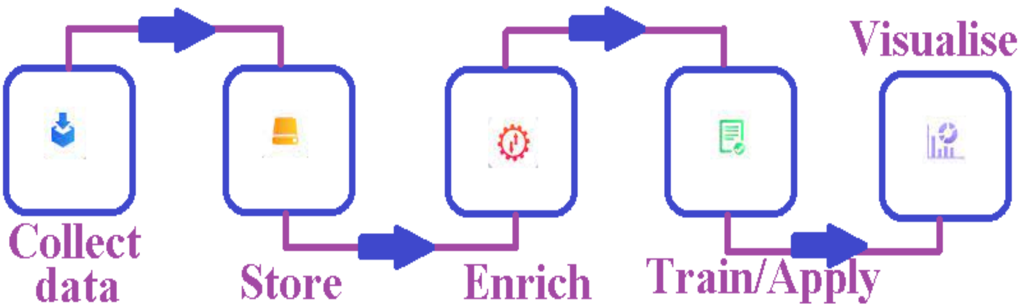

Fig.1. Machine Learning Steps

Supervised Learning-
Data scientists provide input, output and feedback to build model. Supervised learning is based on training.
Examples of some algorithms are- Linear regressions, Support Vector Machines[4].

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 9, September 2021

DOI: 10.17148/IARJSET.2021.8973

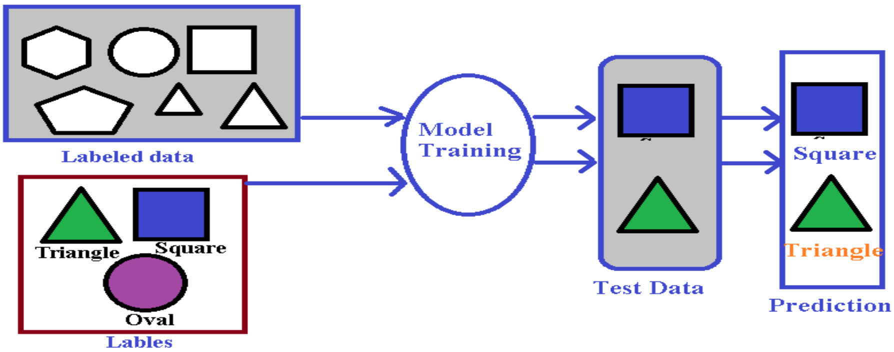

Fig.2. Supervised Learning

Unsupervised Learning-

This use deep learning to arrive at conclusions and patterns through unlabelled training data.
Examples of some algorithms are- Apriori, K means Clustering[4].

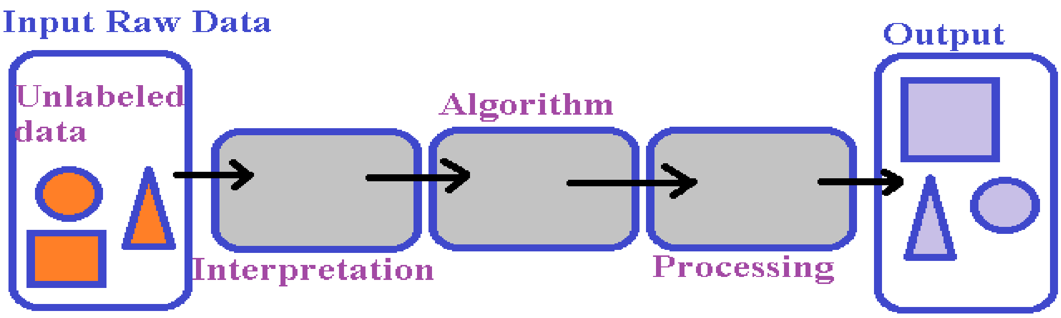

Fig.3. Unsupervised Learning

Reinforcement Learning-

Self interpreting but also based on a system of rewards and punishments learned through trial and error method, seeking maximum reward.
Examples of some algorithms are- Q Learning, Model based value estimation.[5]

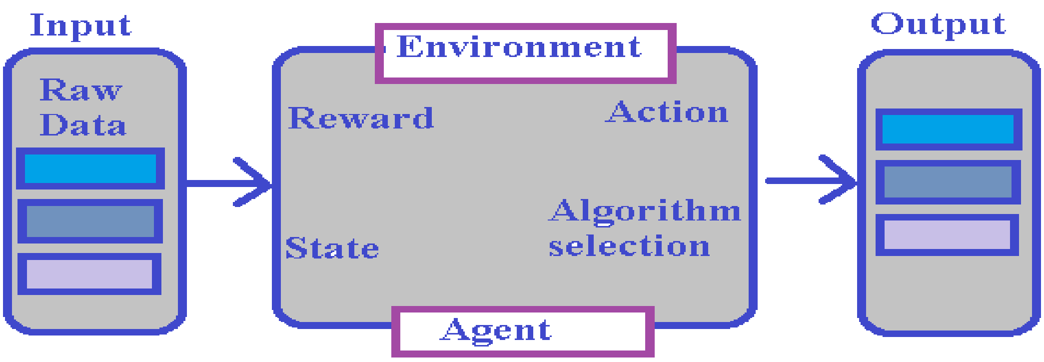

Fig.4. Reinforcement Learning

Decision Trees-

Decision tree is a graphical representation of a tree structure based on the key concept in which the agent will take efficient decision and predict the necessary output.[5]

ISSN (O) 2393-8021, ISSN (P) 2394-1588

IARJSET

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 9, September 2021

DOI: 10.17148/IARJSET.2021.8973

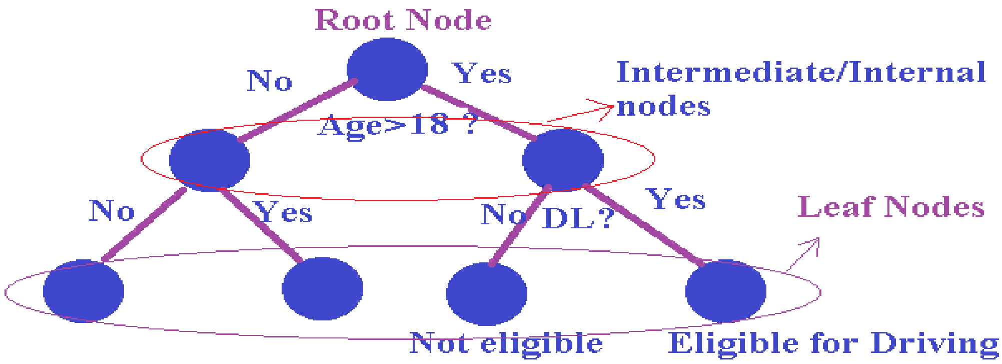

Fig.5. Decision Tree

2. EXPERIMENTS AND OBSERVATIONS-

Machine learning tool that have used in this model is Weka- i.e. Weikato Environment for Knowledge Analysis. Weka is a collection of machine learning algorithms for data mining tasks. It contains tools for data preparation, classification, regression, clustering, association rules mining, and visualization[6].

Sample Dataset Name – Medicine4.arff (Attribute Relation File Format)

ARFF-Viewer - D:\DATA SET\MEDICINE4.arff

File Edit View

MEDICINE4.arff

Relation: MEDICINE3

No.	1: NAME	2: AGE GROUP	3: SEX Nomina	4: SYMPTOM'	5: SYMPTOM2 Nominal	6: SYMPTOM3 Nominal	7: SYMPTOM4 Nominal	8: SYMPTOM5 Nominal	9: MEDICINE Nominal	10: ANTIBIOTIC Nominal	11: PAIN KILLER Nominal	12: COMBINATION OF DRUGS Nominal	13: **DIAGNOSA** String
1	RAM	CHILD	M	FEVER	RUNNING ...	SNEEZING	FATIGUE	COUGH	PARACET...	AZITHAL 500	ACECLOFEN...	PARACETAMOL, AZITHAL 5...	1
2	SH...	ADULT	M	BODY ...	SORE THR...	SNEEZING	NASAL CO...	COUGH	BENADRYL	CEFIXIME 200	DICLOFENAC	BENADRYL, CEFIXIME 200, ...	0
3	HARI	CHILD	M	FEVER	RUNNING ...	BODY PAIN	FATIGUE	WATERY E...	PARACET...	AZITHAL 500	ACECLOFEN...	PARACETAMOL, AZITHAL 5...	1
4	MA...	ADULT	M	COUGH	SORE THR...	SNEEZING	NASAL CO...	MALAISE	BENADRYL	CEFIXIME 200	DICLOFENAC	BENADRYL, CEFIXIME 200, ...	0
5	AMIT	CHILD	F	FEVER	NASAL CO...	BODY PAIN	FATIGUE	COUGH	PARACET...	AZITHAL 500	ACECLOFEN...	PARACETAMOL, AZITHAL 5...	1
6	MA...	ADULT	F	COUGH	MALAISE	WATERY E...	NASAL CO...	COUGH	BENADRYL	CEFIXIME 200	DICLOFENAC	BENADRYL, CEFIXIME 200, ...	0
7	MU...	ADULT	F	COLD	SORE THR...	SNEEZING	FATIGUE	MALAISE	CETRIZINE	CEFPODOXI...	ACECLOFEN...	CETRIZINE, CEFPODOXIM...	1
8	KU...	CHILD	M	FEVER	RUNNING ...	BODY PAIN	FATIGUE	WATERY E...	PARACET...	AZITHAL 500	ACECLOFEN...	PARACETAMOL, AZITHAL 5...	0
9	TU...	CHILD	M	COLD	MALAISE	WATERY E...	NASAL CO...	COUGH	CETRIZINE	CEFPODOXI...	ACECLOFEN...	CETRIZINE, CEFPODOXIM...	1
10	LU...	CHILD	F	FEVER	SORE THR...	SNEEZING	NASAL CO...	MALAISE	PARACET...	AZITHAL 500	ACECLOFEN...	PARACETAMOL, AZITHAL 5...	NO RESULT
11	AK...	ADULT	F	COLD	RUNNING ...	BODY PAIN	FATIGUE	COUGH	CETRIZINE	CEFPODOXI...	ACECLOFEN...	CETRIZINE, CEFPODOXIM...	0

Fig.6. Medical Dataset in ARFF File Format

2.1 Observation-1

Weka-Preprocess-Filter-Randomize:
Weka.filters.unsupervised.instance.randomize-
CAPABILITIES
Class -- Binary class, Date class, Empty nominal class, Missing class values, No class, Nominal class, Numeric class, Relational class, String class, Unary class
Attributes -- Binary attributes, Date attributes, Empty nominal attributes, Missing values, Nominal attributes, Numeric attributes, Relational attributes, String attributes, Unary attributes
Interfaces -- Randomizable, UnsupervisedFilter, WeightedAttributesHandler, WeightedInstancesHandler

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 9, September 2021

DOI: 10.17148/IARJSET.2021.8973

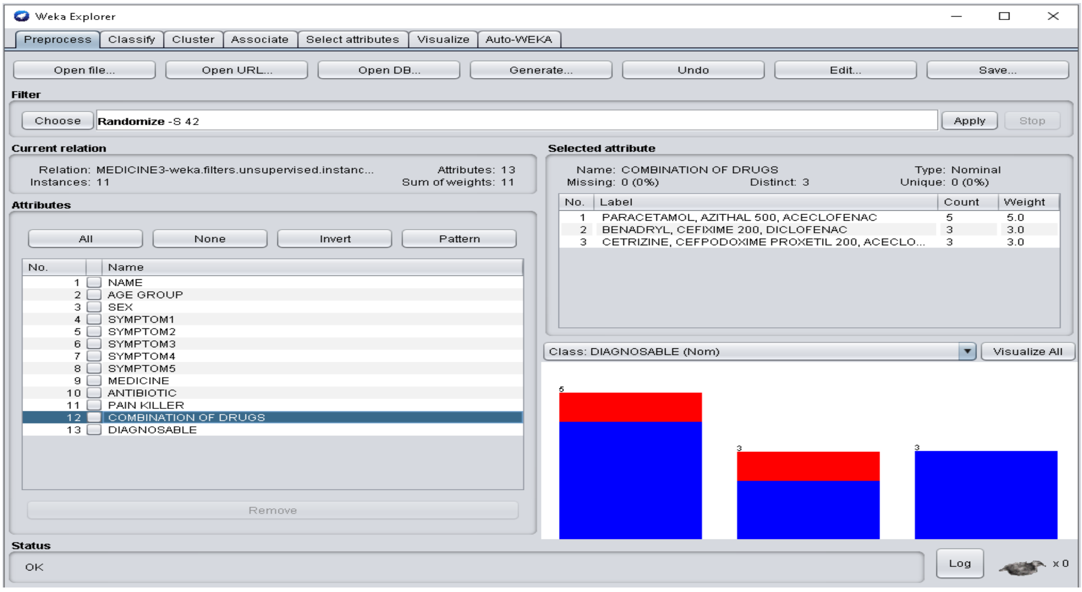

Fig.7. Use of Randomize Filter in the Dataset

2.2 Observation-2

Weka-Classify-ZeroR Classifier-

CAPABILITIES

Class -- Binary class, Date class, Missing class values, Nominal class, Numeric class

Attributes -- Binary attributes, Date attributes, Empty nominal attributes, Missing values, Nominal attributes, Numeric attributes, Relational attributes, String attributes, Unary attributes

Interfaces -- Sourcable, WeightedInstancesHandler

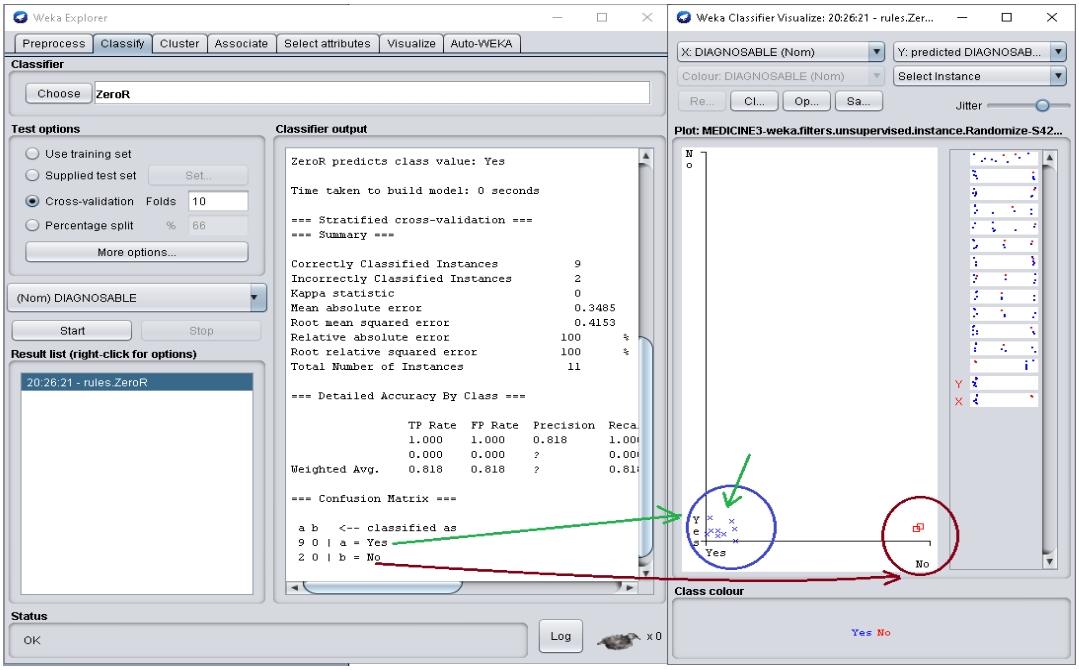

Fig.8. Use of ZeroR Classifier in the Dataset

Classifier Output-

Scheme: weka.classifiers.rules.ZeroR

Relation: MEDICINE3-weka.filters.unsupervised.instance.Randomize-S42

Instances: 11

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 9, September 2021

DOI: 10.17148/IARJSET.2021.8973

Attributes: 13
Test mode: 10-fold cross-validation
=== Classifier model (full training set) ===
ZeroR predicts class value: Yes
Time taken to build model: 0 seconds
=== Stratified cross-validation === == Summary ===

Correctly Classified Instances	9	81.8182 %
Incorrectly Classified Instances	2	18.1818 %
Kappa statistic	0	
Mean absolute error	0.3485	
Root mean squared error	0.4153	
Relative absolute error	100 %	
Root relative squared error	100 %	
Total Number of Instances	11	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	1.000	0.818	1.000	0.900	?	0.028	0.764	Yes
	0.000	0.000	?	0.000	?	?	0.028	0.141	No
Weighted Avg.	0.818	0.818	?	0.818	?	?	0.028	0.651	

=== **Confusion Matrix** ===
a b <-- classified as
9 0 | a = Yes
2 0 | b = No

2.3. Observation 3

Scheme: weka.clusterers.EM -
Relation: MEDICINE3-weka.filters.unsupervised.instance.Randomize-S42
Instances: 11
Attributes: 13
Test mode: evaluate on training data
=== Clustering model (full training set) ===EM==
Number of clusters selected by cross validation: 4
Number of iterations performed: 0
Time taken to build model (full training data) : 0.06 seconds
=== Model and evaluation on training set ===
Clustered Instances
0 1 (9%)
1 2 (18%)
2 5 (45%)
3 3 (27%)
Log likelihood: -9.84923

2.4. Observation 4

Scheme: weka.associations.Apriori-
Relation: MEDICINE3-weka.filters.unsupervised.instance.Randomize-S42
Instances: 11
Attributes: 13
=== Associator model (full training set) = ==Apriori===
Minimum support: 0.5 (5 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 10
Generated sets of large itemsets:
Size of set of large itemsets L(1): 14
Size of set of large itemsets L(2): 22
Size of set of large itemsets L(3): 22

ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 8, Issue 9, September 2021
DOI: 10.17148/IARJSET.2021.8973

Size of set of large itemsets L(4): 15
Size of set of large itemsets L(5): 6
Size of set of large itemsets L(6): 1
Best rules found:
1. AGE GROUP=CHILD 6 ==> PAIN KILLER=ACECLOFENAC 6 <conf:(1)> lift:(1.38) lev:(0.15) [1] conv:(1.64)
2. SYMPTOM4=FATIGUE 6 ==> PAIN KILLER=ACECLOFENAC 6 <conf:(1)> lift:(1.38) lev:(0.15) [1] conv:(1.64)
3. SYMPTOM1=FEVER 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
4. MEDICINE=PARACETAMOL 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
5. ANTIBIOTIC=AZITHAL 500 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
6. COMBINATION OF DRUGS=PARACETAMOL, AZITHAL 500, ACECLOFENAC 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
7. MEDICINE=PARACETAMOL 5 ==> SYMPTOM1=FEVER 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)
8. SYMPTOM1=FEVER 5 ==> MEDICINE=PARACETAMOL 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)
9. ANTIBIOTIC=AZITHAL 500 5 ==> SYMPTOM1=FEVER 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)
10. SYMPTOM1=FEVER 5 ==> ANTIBIOTIC=AZITHAL 500 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)

2.5. Observation 5

weka.associations.FilteredAssociator
weka.filters.MultiFilter \"weka.filters.unsupervised.attribute.ReplaceMissingValues
weka.associations.Apriori –
- Relation: MEDICINE3-weka.filters.unsupervised.instance.Randomize-S42
Instances: 11
Attributes: 13
=== Associator model (full training set) ===
FilteredAssociator using weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1 on data filtered through weka.filters.MultiFilter -F "weka.filters.unsupervised.attribute.ReplaceMissingValues "
Filtered Header
@relation MEDICINE3-weka.filters.unsupervised.instance.Randomize-S42-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.MultiFilter-Fweka.filters.unsupervised.attribute.ReplaceMissingValues
@attribute NAME {RAM,SHYAM,HARI,MAMBO,AMIT,MADHAB,MUNA,KUNA,TUNA,LUNA,AKASH}
@attribute 'AGE GROUP' {CHILD, ADULT}
@attribute SEX {M,F}
@attribute SYMPTOM1 {FEVER,'BODY PAIN', COUGH, COLD}
@attribute SYMPTOM2 {'RUNNING NOSE','SORE THROAT','NASAL CONGESTION', MALAISE}
@attribute SYMPTOM3 {SNEEZING,'BODY PAIN','WATERY EYES'}
@attribute SYMPTOM4 {FATIGUE,'NASAL CONGESTION'}
@attribute SYMPTOM5 {COUGH,'WATERY EYES', MALAISE}
@attribute MEDICINE {PARACETAMOL, BENADRYL, CETRIZINE}
@attribute ANTIBIOTIC {'AZITHAL 500','CEFIXIME 200','CEFPODOXIME PROXETIL 200'}
@attribute 'PAIN KILLER' {ACECLOFENAC, DICLOFENAC}
@attribute 'COMBINATION OF DRUGS' {'PARACETAMOL, AZITHAL 500, ACECLOFENAC','BENADRYL, CEFIXIME 200, DICLOFENAC','CETRIZINE, CEFPODOXIME PROXETIL 200, ACECLOFENAC'}
@attribute DIAGNOSABLE {Yes, No}
@data
Associator Model= =Apriori====
Minimum support: 0.5 (5 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 10
Generated sets of large itemsets:
Size of set of large itemsets L(1): 14
Size of set of large itemsets L(2): 22
Size of set of large itemsets L(3): 22
Size of set of large itemsets L(4): 15
Size of set of large itemsets L(5): 6

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 9, September 2021

DOI: 10.17148/IARJSET.2021.8973

Size of set of large itemsets L(6): 1

Best rules found:

1. AGE GROUP=CHILD 6 ==> PAIN KILLER=ACECLOFENAC 6 <conf:(1)> lift:(1.38) lev:(0.15) [1] conv:(1.64)
2. SYMPTOM4=FATIGUE 6 ==> PAIN KILLER=ACECLOFENAC 6 <conf:(1)> lift:(1.38) lev:(0.15) [1] conv:(1.64)
3. SYMPTOM1=FEVER 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
4. MEDICINE=PARACETAMOL 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
5. ANTIBIOTIC=AZITHAL 500 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
6. COMBINATION OF DRUGS=PARACETAMOL, AZITHAL 500, ACECLOFENAC 5 ==> AGE GROUP=CHILD 5 <conf:(1)> lift:(1.83) lev:(0.21) [2] conv:(2.27)
7. MEDICINE=PARACETAMOL 5 ==> SYMPTOM1=FEVER 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)
8. SYMPTOM1=FEVER 5 ==> MEDICINE=PARACETAMOL 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)
9. ANTIBIOTIC=AZITHAL 500 5 ==> SYMPTOM1=FEVER 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)
10. SYMPTOM1=FEVER 5 ==> ANTIBIOTIC=AZITHAL 500 5 <conf:(1)> lift:(2.2) lev:(0.25) [2] conv:(2.73)

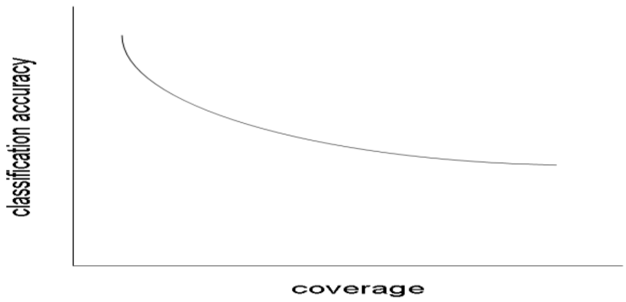

Fig.9. Performance Evaluation

CONCLUSION

I have taken 5 different observations to decide the acceptability of a particular domain in the machine learning model. In the study of the above real time medical dataset implementation and in different observations, it is found that the confidence level using the Wekato Machine Learning tool is very much satisfactory, having good accuracy result and so will be acceptable.

REFERENCES

Journals

1. Ai T., Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, L. Xia, "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases", Radiology, 296 (2) (2020), pp. E32-E40.
2. A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks", Communications of the ACM, 60 (6) (2017), pp. 84-90
3. M. Paknezhad, S. Marchesseau, M. Brown, "Automatic basal slice detection for cardiac analysis", Journal of Medical Imaging, 3 (3) (2016), Article 034004
4. P. Radau, Y. Lu, K. Connelly, G. Paul, A. Dick, G. Wright, "Evaluation framework for algorithms segmenting short axis cardiac MRI", The MIDAS Journal (2009)
5. A. Altan, S. Karasu, "Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique", Chaos, Solitons and Fractals, 140 (2020), Article 110071
6. X. Zhuang, L. Li, C. Payer, D. Štern, M. Urschler, M.P. Heinrich, "Evaluation of algorithms for multi-modality whole heart segmentation: An open-access grand challenge", Medical Image Analysis, 58 (2019), Article 101537
7. S.P. Kasiviswanathan, H.K. Lee, K. Nissim, S. Raskhodnikova, A. Smith, "What can we learn privately?", SIAM J. Comput., 40(3):793–826, 2011.

Book

8. S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach", Pearson 3rd Edition.
9. O. Theobald, "Machine Learning For Absolute Beginners", 2nd Edition
10. A. Ethem, "Introduction to Machine Learning", PHI, 3rd Edition
11. S. S. Shwartz and S. B. David, "Understanding Machine Learning: From Theory To Algorithms".

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 9, September 2021

DOI: 10.17148/IARJSET.2021.8973

12. A. Srinivasaraghavan and V. Joseph, "Machine Learning", Wiley Eastern Ltd.

Conference Proceedings

13. Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, " Gradient-based learning applied to document recognition", Proceedings of the IEEE, 86 (11) (1998), pp. 2278-2324
14. B.E. Boser, I. Guyon, V.N. Vapnik (1992), "A training algorithm for optimal margin classifiers", InProceedings of the Fifth Annual Workshop of Computational Learning Theory, 5, 144–152. Pittsburgh, ACM.
15. L. Bottou, C. Cortes, J.S. Denker, H. Drucker, I. Guyon, L.D. Jackel, Y. LeCun, E. Sackinger, P. Simard, V. Vapnik, U.A. Miller(1994), " Comparison of classifier methods: A case study in handwritten digit recognition", Proceedings of 12th International Conference on Pattern Recognition and Neural Network.

