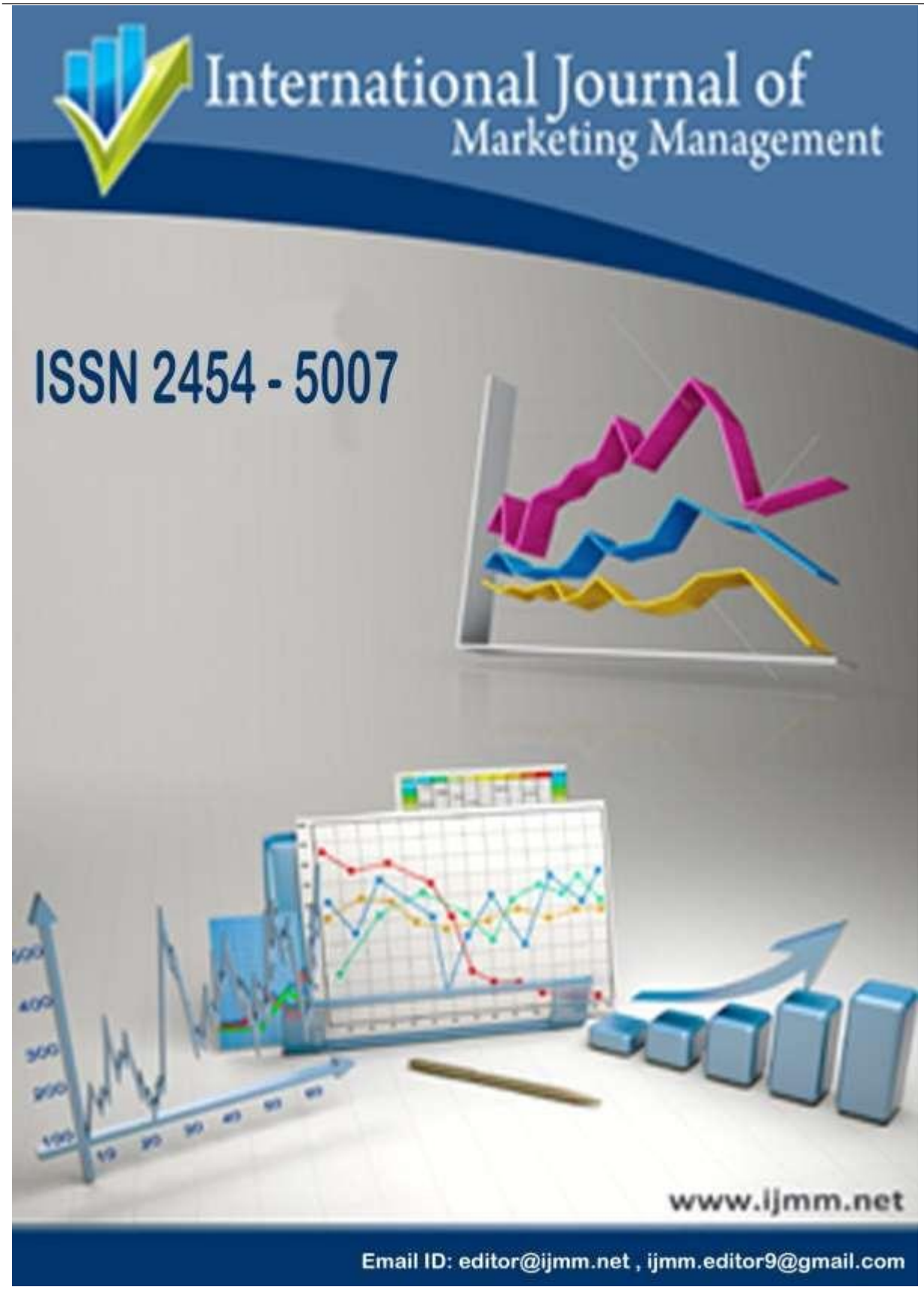
International Journal of
Marketing Management
ISSN 2454 - 5007
www.ijmm.net
Email ID: editor@ijmm.net , ijmm.editor9@gmail.com

International Journal of Marketing Management

ISSN 2454-5007, www.ijmm.net

Vol. 15 ,Issue. 2,April 2023

EFFICIENT THYROID DISEASE PREDICTION USING FEATURES AND MULTI METERS

G. ANUSHA1, K. NAVYASRI 2, K.SAI RISHI 3, H. POOJITHA4

Article Info

Received: 09-01-2023 *Revised: 01 -02-2023* *Accepted: 25-03-2023*

ABSTRACT;

Prediction of Thyroid is a complex axiom in medical research. Machine learning methods are more powerful and compact for the healthcare industry to handle the massive amount of healthcare records. The methods in machine learning provides the facility to use different kind of data values which are used for prediction. Data cleaning techniques are used for enhancing the dataset to provide accurate results. The noisy and missing values are handled using data pre-processing methods. In this work, Adaboost and Bagging techniques are used for thyroid classification. The methods are executed, and the results are compared to show the effective method for thyroid prediction.

Keywords: Machine learning algorithms, Thyroid Disease, Support Vector Machine (SVM), Decision Trees

INTRODUCTION

Nowadays, thyroid disorders destruction the normal functioning of the thyroid gland which causes anomalous production of hormones leading to hyperthyroidism[1][3]. The occurrence of hypothyroidism in the developed world is estimated to be about 4- 5%. Hypothyroidism may cause high cholesterol levels, an increase in blood pressure, cardiovascular complications, decreased fertility, and depression if not properly treated. The technology and information in medical sciences, the computer science professionals are capable of providing expert advisory system (EAS) [4][7] .To diagnose different kinds of diseases with high accuracy. The medical professionals are made to use these systems due to some developed errors during general diagnosis process [5]. Disease diagnosis operations using EAS are performed based on sets of disease symptoms. These systems are based on machine learning technique which helps the physician to minimize the costs and time in effective diagnoses. This work is a hybrid architecture design furnished successfully using machine learning techniques. The purpose of this work is to bring the spotless and smart approach in identifying the thyroid disease in a human[9][10]. There are several mechanisms implemented on thyroid data sets which produced astonished outcomes, but the data considered for the thyroid disease diagnoses (TDD) is inconsistent, redundant and consists of missing attribute values as per my literature survey. The proposed work is to construct an expert advisory system of hybrid architecture, which is to determine the optimistic disease growth because of the thyroid gland. A string matching system (SMS) was at the outset developed, which can predict the actual TDD based on the knowledge available [12] [13].

International Journal of Marketing Management

1ASSISTANT PROFESSOR, DEPARTMENT OF ECE, MALLA REDDY ENGINEERING COLLEGE FOR WOMEN, HYDERABAD.2,3&4UG SCHOLAR, DEPARTMENT OF ECE, MALLA REDDY ENGINEERING COLLEGE FOR WOMEN, HYDERABAD

The proposed a method of invoking machine learning algorithm (MLA) into this proposed work of thyroid disease diagnoses (TDD) [6]. MLA has advantages such as selflearning, high parallelism, speed and error tolerance against noise [7]. The readiness of reuse feature in MLA made us to choose it as an option to develop this proposed work. The thyroid related attributes are converted for suitable representation of rough set analysis. This is done in two levels i.e., discretizing and later a matrix is generated with rows representing similarity score of attributes and columns representing the definition of the attribute for individual rule set

PROPOSED ARCHITECTURE

Applying data mining in the medical field is a very challenging task in medical profession [4]. In medical research the data mining begins with a hypothesis and results are adjusted to fit the hypothesis. This differs from standard data mining practice, which simply starts with datasets without an apparent hypothesis. In the context of data mining, learning technique is generally classified as supervised and unsupervised learning technique both belong to machine learning technique [10]. Classification is supervised learning that focus on the prediction based on known properties [9] [13]. A Classification task begins with a data set in which the class assignments are known. If the target or class label has numerical values then a predictive model uses. Regression algorithm is not a Classification algorithm

existing algorithms are able to extract knowledge from data set that store discrete

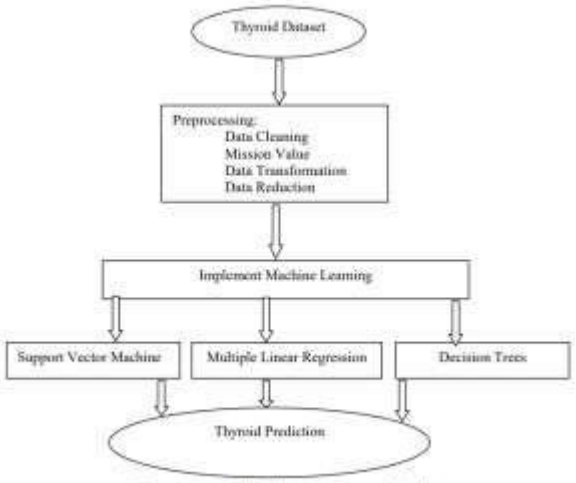

Figure 1 Thyroid Prediction Proposed Architecture

There are many classification algorithms some method are mostly used such as decision tree, Support vector machine, Naive Bayes, KNN etc[1][2]. In the historical verview I emphasize the naive Bayesian classifier, neural networks and decision trees. I present a comparison of some state of the art systems, representatives from each branch of machine learning, when applied to several medical diagnostic tasks. The future trends are illustrated by two case studies. The first describes a recently developed method for dealing with reliability of decisions of classifiers, which seems to be promising for intelligent data analysis in medicine[5][8]. The second describes an approach to using machine learning in order to verify some unexplained phenomena from complementary medicine, which is not (yet) approved by the orthodox medical community but could in the future play an important role in overall medical diagnosis and treatment.

PREPROCESSING The pre-processing is to resolve several types of problems including noisy data, redundant data, missing data values, etc [10]. The high quality data will be lead to high quality results and reduced costs for data mining. Missing data should be pre-processed. It is to allow the whole data set to be processed by a required algorithm and the most of the features. If the features are continuous, the algorithms can be integrated to

creatediscrete attributes [12] [9]. In the proposed work, the data taken from UCI repository has both continuous and discrete data which undergoes pre-processing. The missing value and not a number constraint arechecked using masking method. If the missing value or Not a Number values are present and it is replaced by the mean value of the column. Pre-processing refers to the program that processes its input data and produce output that is used as input to a compiler.

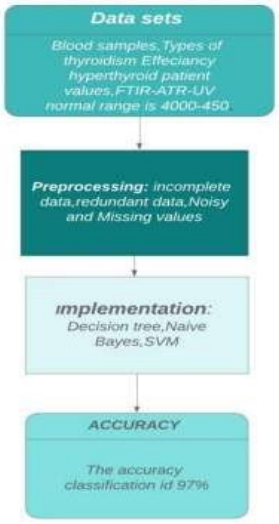

Figure 2.Thyroid Dataset preprocessing

CONCLUSION

The thyroid gland is the primary and biggest gland in the endocrine system. The data mining technique is applied on the hypothyroid dataset to determine the positive and the negative cases from the entire dataset. The objective of this work is aimed to show the classes of thyroid from the available raw medical dataset that helps to the accurate diagnosis. The datas are classified as Hyper, Hypo, Normal and average classes. Various splitting rule for decisiontreeattributeselection had been analyzed and compared. This helps diagnosis thyroid diseases through the extracted rules. It is clear that normalized based splitting rules have high accuracy and sensitivity or true positive rate. It can be extended for any medical datasets. The enhancement can be made by using various optimization algorithms or rule extraction algorithms. It was compared and it can be seen that Decision Trees successfully used to help the diagnosis of thyroid disease. It is observed that the Decision Trees performed in SVM, Multi linear regression, Naïve Bayes with respect to the accuracy of the network to diagnose the thyroid disease. The results are evaluated and based on ten evaluation metrics then the accuracy of classification is 97.97%.

REFERENCES

[1] Sharma, Seema, Jitendra Agrawal, and Sanjeev Sharma. "Classification through machine learning technique: C4. 5 algorithm based on various entropies." International Journal of Computer Applications 82.16 (2013).Geetha, K., and CAPT S. SANTHOSH BABOO. "EFFICIENT THYROID DISEASE CLASSIFICATION USING DIFFERENTIAL EVOLUTION WITHSVM." Journal of Theoretical & Applied Information Technology 88.3 (2016).Margret, J., B. Lakshmipathi, and S. Aswani Kumar. "Diagnosis of Thyroid Disorders using Decision Tree Splitting Rules." International

[2] Journal of Computer Applications 44.8 (2012): 43-46.Prasad, V., T. Srinivasa Rao, and M. Surendra Prasad Babu. "Thyroid disease diagnosis via hybrid architecture composing rough

data sets theory and machine learning algorithms." Soft Computing 20.3 (2016): 1179-1189.