International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Inflammatory Bowel Disease Classification Using Neural Network and Support Vector Machine

Asmaa Mohamed Hassan[1], Yasser M. K. Omar[2]
[1]Department of Computer Science, El Shorouk Academy, Cairo, Egypt
[2]Department of Computer Science, Arab Academy for Science, Technology and Maritime Transport, Cairo, Egypt
Email address: asmaamohamed.am75 @ gmail.com

Abstract— *Due to the importance and difficulty of diagnosis of Inflammatory Bowel Disease (IBD) computer-aided diagnosis plays an essential role in the early detection of the IBD. In this paper we explore the feasibility of Machine Learning (ML) algorithms for classification of the activation of IBD and its subtypes Chon's Disease (CD) or Ulcerative Disease (UC). Two commonly used ML algorithms were applied: The Artificial Neural Network (ANN) and the Support Vector Machine (SVM). A detailed experiment is conducted to evaluate the classification capability of each algorithm. Moreover, using different datasets containing normal and IBD patients, the experiment result reveals that ANN and SVM with linear kernel have comparable performance in classification with accuracy of 80% and 79.9% respectively; however, SVM with Radial Basis Function (RBF) kernel outperforms ANN with accuracy of 70.9% and 67% respectively. ANN performance may become much worse by increasing the training samples and the ANN number of layers.*

Keywords— *Data mining, classification, Inflammatory bowel disease, Artificial neural networks, Support Vector Machine.*

I. INTRODUCTION

IBD is a group of inflammatory feature of the colon and small intestine. Crohn's disease and ulcerative colitis are the principal types of inflammatory bowel disease[1]. An accurate diagnosis of IBD is necessary for a rapid and effective treatment. the incidence of which is increasing [2][3]. The major feature of inflammatory bowel disease is chronic inflammation of the gastrointestinal (GI) tract. Although both Crohn's disease and ulcerative colitis are considered fall into the same disease group, there are differences in disease location within the bowel, that can be observed through endoscopic and histological assessment. Although the Endoscopic investigation of the IBD is macroscopic and can determine initial treatment and provisional diagnosis, this investigation of the gastrointestinal system is not always adequate diagnosis. Therefore, the histological (microscopic) examination of biopsies from the upper and lower GI tracts is vital to determine the disease extent and to confirm that this is the right diagnosis. Typically, Crohn's disease is marked by a non-continuous inflammation of the entire gastrointestinal system, while the inflammation pattern of ulcerative colitis is continuous and limited to the colon and rectum[4]. Diagnosis of IBD is challenging as the cause of the disease is not fully understood and any attempt to manage or predict it would be complex. There are some Tools that assist clinicians to diagnose more accurately. These tools may assist in categorizing the disease better by classifying the disease into several specific phenotypes with implications for how best to treat.

Therefore, Machine learning [5] is a contemporary branch of statistics which is particularly used for the analysis of complex data. The algorithms of Machine learning focus on finding patterns within data in order to use them to predict, classify or infer new knowledge about the disease. These methods are mainly divided into two categories: unsupervised machine learning algorithms and supervised algorithms. supervised algorithms are mainly used to solve classification problems. for instance, when the class of each sample/patient is known a priori, these samples are used to train a model to classify the follwing samples of unknown class. In the proposed model, Several techniqes are used to classify IBD patients. These techniqes are SVM with different kernels (linear,polynomial and rbf) and ANN. This study depend on machine learning (ML) to classify the types of the IBD using endoscopic and histological data of 644 patients diagnosed with IBD. These data were used in order to enhance, train, test and validate a ML model to classify the subtype of disease. This study employs mathematical modeling of endoscopic and histological data which assist in diagnostic accuracy. Section 2 tackles old studied that focus on IBD classification. Section 3 discusses our material and methods used. In section 4, the results will be discussed.

II. RELATED WORK

Many studies have attempted to enhance the accuracy of IBD diagnosis over the last years. In general, there are three techniques for diagnosis: Blood or serum analysis, Radiographic, Endoscopy and histology. Despite the wide availability of a multitude of diagnostic modalities, the correct diagnosis of inflammatory bowel disease depends on the accurate interpretation and correlation of endoscopic and histologic findings[6].

IN [7], the authors concentrate on endoscopic data and histological data. principal component analysis (PCA) for Pre-Processing data and applied recursive feature elimination algorithm combined with a 5-fold cross validation scheme (RFE-CV) to feature extraction. The SVM is used to classify the Paediatric of Inflammatory Bowel Disease (PIBD). However, the problem with this paper is its limitation; it uses a small dataset. In addition, it doesn't measure the sensivity and specificity.

IN [8], the authors focus on results from colonoscopy, the current reference level for CD diagnosis. Intensity statistics,

Asmaa Mohamed Hassan and Yasser M. K. Omar, "Inflammatory Bowel Disease Classification Using Neural Network and Support Vector Machine," *International Research Journal of Advanced Engineering and Science*, Volume 5, Issue 1, pp. 216-221, 2020.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

texture anisotropy and shape asymmetry of the 3D regions are used as charachteristics to differentiate between diseased and normal regions.Random forests (RF), Support Vector Machines (SVM), and a Bayesian Classifier (BC) are employed to detect and localize CD afflicted regions by input abdominal magnetic resonance (MR) volumes without an explicit segmentation of the bowel wall.

IN[9], the main focus of the authors is on the analysis of Ulcerative Colitis Proteomic Data. Random Forest, Logistic Regression, SVM and Parenclitic networks analysis are used to examine proteomic data by children patients, to solve classification task for proteomic data from healthy and diseased. The limitation of this paper is that although it uses a huge number of data, the results are bad.

IN[10], the authors focus on endoscopic colorectal biopsies by IBD patients. four Bayesian network classifiers are used as potential instruments for the histopathological diagnosis of chronic idiopathic inflammatory bowel disease (CIIBD). This paper not used any techniques for preprocessing the data before classification .

It has been observed from old studies that classify subtypes of IBD that the use of SVM and ANN algorithms and preprocessing of data were very important in accurate diagnosis.

III. Material and Methods

SVM with different kernels and ANN systems applied to diagnose subtypes of inflammatory bowel disease. The experiments were carried out by optimizing the parameter of NN model and SVM by trial and error to reduce false negative rate. The overall workflow described in fig. 1 is used to propose crucial steps. At first, chronic idiopathic inflammatory bowel disease (CIIBD) dataset is loaded, then Apply one hot encoding to represent the categorical variables as binary vectors. In this paper, two approaches are Applied by using two algorithms SVM and ANN. In first approach (IBD-H), diagnosis problem was divided into two subproblems. the first subproblem classified using SVM classifiers and ANN classifier, the second makes classification between CD and UC on bases of being activation. In the second approach(IBD-1), we differentiate between normal, active CD, inactive CD, active UC and inactive UC.

A. Dataset

The real-world database of this study is originally from the field of histopathological diagnosis of endoscopic colorectal biopsies which is received at the Department of Histopathology, Royal Hallamshire Hospital in Sheffield, United Kingdom, in 1990 and 1995[11]. This database is divided into three subsets. The first subset is the complete dataset (called “All IBD & Normal”), which contains 809 cases of which 165 are normal, 473 as ulcerative colitis (UC) and 171 as Crohn’s disease. The second case (“All IBD”) contains 644 cases of which 473 as UC and 171 Crohn’s disease. The third (“Active IBD”) contain 370 cases of which 283 are active UC and 87 active Crohn’s disease. The word (“active” here means that is an active inflammation as indicated by polymorphs in the lamina propria). Moreover, there are 23 independent variables and one dependent variable (the outcome) that olso form part of these datasets. The diagnosis was confirmed by endoscopy, radiology and microbiological items results.

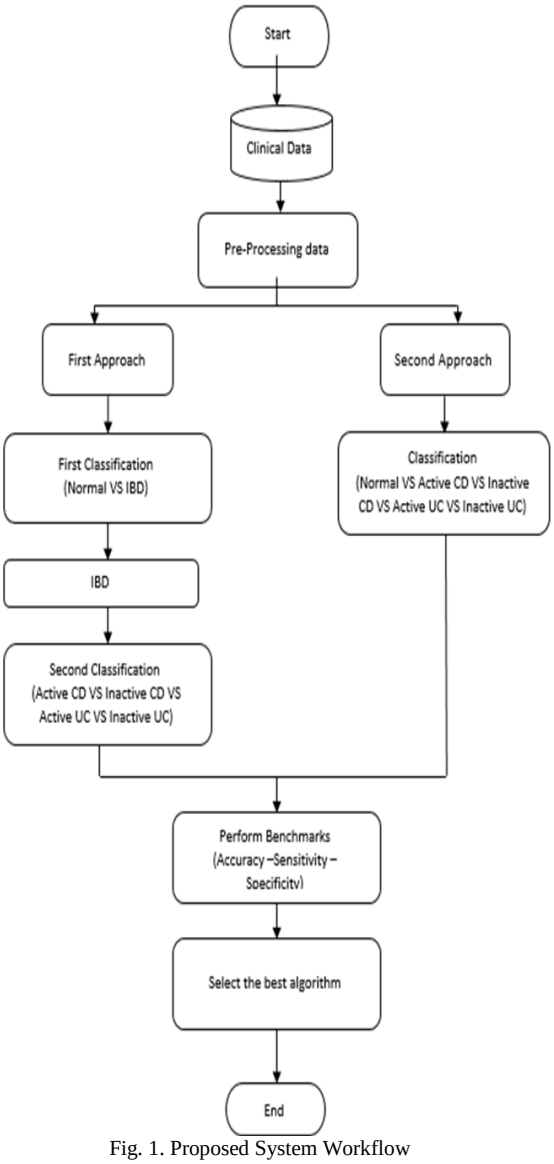

Fig. 1. Proposed System Workflow

The observed features: The biopsies were examined (blind to all clinical details) by a single experienced observer (SSC) using of computer interface which implement the BSG Guidelines in the Initial Biopsy Diagnosis of Suspected Chronic Idiopathic Inflammatory Bowel Disease (Jenkins et al. 1997) with declared images represent examples of each histopathological feature (Cross et al. 1997). Some of the features are dichotomous variables, e.g. the presence or absence of mucosal granulomas, whilst others are ordinal categories, e.g. mucin depletion classified into none, mild, moderate or severe. The observed features and their coding are given in table I. Observation was studied in a period 9 months with no more than 30 biopsies observed at every single day.

Asmaa Mohamed Hassan and Yasser M. K. Omar, “Inflammatory Bowel Disease Classification Using Neural Network and Support Vector Machine,” *International Research Journal of Advanced Engineering and Science*, Volume 5, Issue 1, pp. 216-221, 2020.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

TABLE I. The observed features of IBD.

Feature	Type
Age	Real integer
Sex	Binary
Active inflammation (subset classifier, not observed feature)	Binary
Mucosal surface	Ordinal categorical
Crypt architecture	Ordinal categorical
Crypt profiles	Real integer
Increased lamina propria cellularity	Binary
Mild & superficial increase in lamina propria cellularity	Binary
Increased lymphoid aggregates in lamina propria?	Binary
Patchy lamina propria cellularity?	Binary
Marked & transmucosal increase in lamina propria cellularity	Binary
Cryptitis extent	Ordinal categorical
Cryptitis polymorphs	Ordinal categorical
Crypt abscesses extent	Ordinal categorical
Crypt abscesses polymorphs	Ordinal categorical
Lamina propria polymorphs	Ordinal categorical
Epithelial changes	Ordinal categorical
Mucin depletion	Ordinal categorical
Intraepithelial lymphocytes	Binary
Subepithelial collagen	Binary
Lamina propria granulomas	Binary
Submucosal granulomas	Binary
Basal histiocytic cells	Binary

B. Data Pre-processing

1. Classifying the dependent and Independent Variables. The dependent and independent values are saved in different arrays. X: independent variable set (Takes all rows of all columns except the last column), Y: dependent variable set(takes all rows of the last column).
2. Dealing with Categorical Data SVM and ANN algorithms cannot operate on label data directly. They require all input variables and output variables to be numeric. This includes two steps:
   a. Integer Encoding: As a first step, each unique category value is allocate an integer value.
   b. One-Hot Encoding [12][13]: can be employed to the integer representation. Where the integer encoded variable is extracted and a new binary variable is added for each singular integer value.
3. Splitting the Dataset into Training and Testing sets, database split into 2 subsets: the training and the test sets [13]. The size of the training set is 80% of the data and the remaining 20% of data acts as the test set in each fold (using 5-fold cross validation).

C. Classification

Machine learning techniques has crucial role in the classification purposes. This section describes SVM [14] and ANN [15]. Statistical analyses were performed applying Python3 [16] and the Scikit-Learn (0.17.1) package [17].

SVM is a type of supervised machine learning classification algorithm. Given a set of training examples, each manifest as belonging to one or the other of two categories, an SVM training algorithm constructs a model that allocates new examples to one category or the other using different Kernel, Kernel functions are used for finding data relations in dataset. Kernel methods allow data to activate in higher dimensional space by discovering the kernel function that is appropriate for higher classification accuracy. There are three commonly used kernel functions: linear, polynomial, RBF [18].

1. Linear:

$$K(x_i, x_j) = (x_i \cdot x_j) \quad (1)$$

2. Polynomial:

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^p \quad (2)$$

Where p is the degree of the polynomial.

3. RBF

$$K(x_i, x_j) = \exp\left[-\gamma \left\| x_i - x_j \right\|^2\right] \quad (3)$$

Where γ is specified by keyword gamma, must be greater than 0

ANN[15] uses the processing of the brain as a basis to develop algorithms that can be used to model complex patterns and prediction problems. A network is constructed by including layers of neurons. The first layer on the left is the input layer, and it encloses the neurons that receive input from the outside. The last layer on the right is the output layer, and it contains the neurons that carry the output of the network. One or more hidden layers are positioned between the input and output layers. Hidden-layer neurons are used for execution most of the calculations during the approximation of the function.

D. Cross Validation

The performance of the SVM classifier and ANN are trained and tested by using 5- fold cross validation. SVM classifier and ANN are designed based on 80% of datasets in order to train data, the classifier is partitioned into 5 approximately equal sets. Each set had a similar proportion of individuals. This process is then repeated 5 times for each test set in order to to select the best SVM kernel and the best model of ANN. The remaining of data 20% is considered as a test set used to evaluate the classifier performance.

IV. Results and Discussions

The proposed model is developed and implemented by dividing it to two approaches. The first approach includes two classifiers; the first classifier is to differentiate between normal and IBD, then the second classifier diagnose the disease among its 4 different types. On the other hand, the second approach is proforming just one classifier to diagnose among the the normal and the four types of IBD in one step.
The proposed model was experiemented by using the dataset (Dataset of Observed Features on Endoscopic Colorectal Biopsies from Normal Subjects and Patients With Chronic Inflammatory Bowel Disease (Crohn's Disease and Ulcerative Colitis) by sample as 644 patient and when appling the svm and ANN its has given the following results. The applied Performance evaluation by using accuracy, recall and specificity. Accuracy is used to calculate the proportion of the total number of correct predictions as follows [19]:

Asmaa Mohamed Hassan and Yasser M. K. Omar, "Inflammatory Bowel Disease Classification Using Neural Network and Support Vector Machine," *International Research Journal of Advanced Engineering and Science*, Volume 5, Issue 1, pp. 216-221, 2020.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Where, TP = True Positive Rate, FP = False Positive Rate.

Precision is used to calculate the proportion of the predicted positive cases using True Positive Rate (TP) and FP False Positive Rate (FP), it is computed as:

$$\text{Re}call = \frac{TP}{TP + FN} \quad (5)$$

Where, FN =False Negative Rate

Specificity (also called the true negative rate) measures the proportion of actual negatives that are correctly identified, it is computed as:

$$\text{Specificit y} = \frac{TN}{TN + FP} \quad (6)$$

➢ First Approach SVM

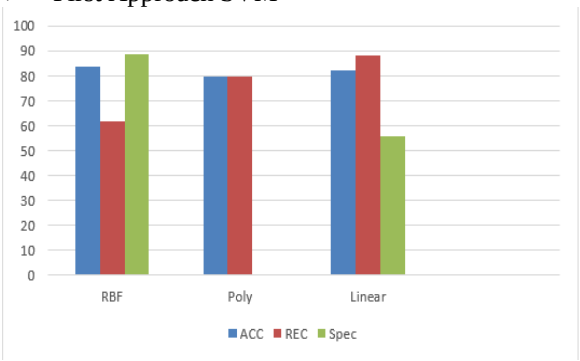

Fig. 2. SVM Classification results: normal and IBD

In terms of accuracy [Fig.2], we see that SVM with RBF kernel has achieved highest score with 4.3% higher accuracy than SVM polynomial kernel and 1.9% higher than SVM linear kernel. In terms of recall, Linear kernel achieved highest accuracy of 88%, 26% higher than RBF kernel and 8.37% higher than polynomial kernel. In terms of specificity SVM RBF kernel achieved 83.7%, 27.7% higher than SVM linear kernel.

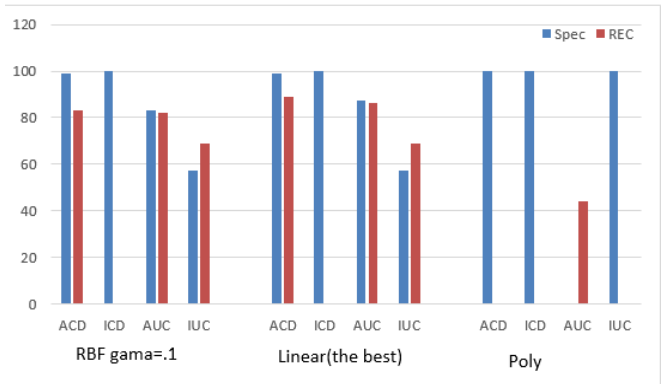

Fig. 3. Classification results: IBD (ACD, ICD, AUC, IAC)

In the next classifier, we classify cases between Active CD(ACD), Inactive CD(ICD), Active UC(AUC) and Inactive UC (IUC). In terms of accuracy [Fig. 4], SVM with Linear kernel achieved 79.07% accuracy, 2.37% higher than SVM with RBF kernel, while SVM with polynomial kernel with the lowest with 44.1%. We see [Fig. 3] that Linear SVM achieved highest score with 88.8% ACD recall (5.5% higher than RBF), 86.1% on AUC (3.8% higher than RBF), while scores on ICD and IUC remains the same. We see that SVM RBF and SVM Linear kernels produced same results in ACD, ICD and IUC, where SVM Linear was better at AUC with specificity of 87.5(4.2% higher than of SVM RBF kernel).

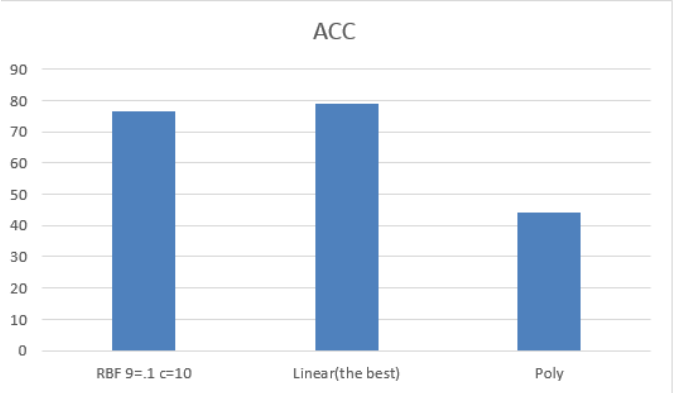

Fig. 4. Accuracy: IBD (ACD, ICD, AUC, IAC)

Comparing according to kernels: In first classifier, RBF produced highest scores in accuracy and specificity, while Linear kernel scored highest score in sensitivity.

In second classifier, using Linear kernel produced the highest scores among all three kernels.

➢ Second Approach SVM

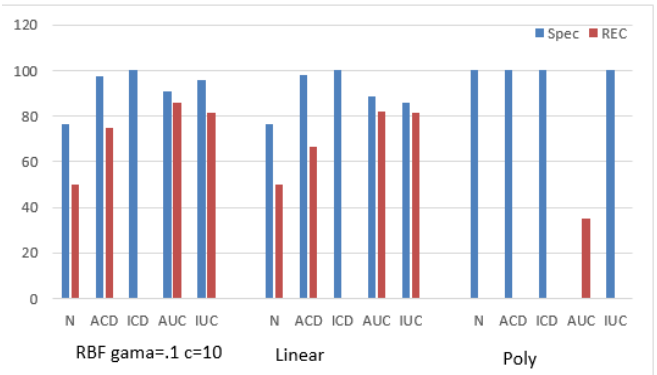

Fig. 5. Classification results: IBD (N, ACD, ICD, AUC, IAC)

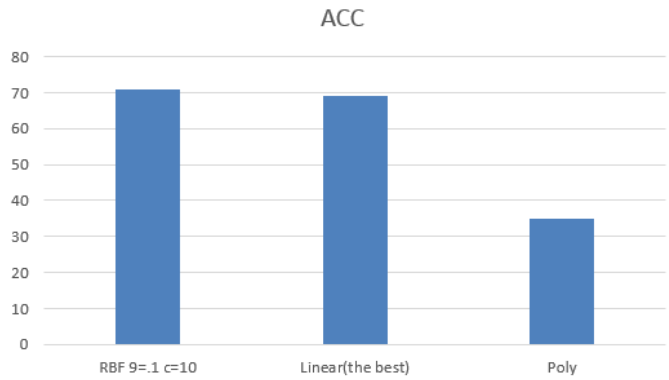

Fig. 6. Accuracy IBD (N, ACD, ICD, AUC, IAC)

Asmaa Mohamed Hassan and Yasser M. K. Omar, "Inflammatory Bowel Disease Classification Using Neural Network and Support Vector Machine," *International Research Journal of Advanced Engineering and Science*, Volume 5, Issue 1, pp. 216-221, 2020.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

In the second approach, we classify between normal, active and inactive CD and active and inactive UC.

Comparing according to kernels

In terms of recall [Fig. 5]:

- SVM RBF scored: 50% of normal, 75% of ACD, 0% of ICD, 85.7% of AUC and 81.4% of IUC
- SVM Linear scored 50% of normal, 66.6% of ACD, 0% of ICD, 81.8% of AUC and 81.4% of IUC
- SVM polynomial scored 0% of normal, 0% of ACD, 0% of ICD, 35.1% of AUC and 0% of IUC

We see that RBF and Linear kernels have same score of normal, ICD and IUC. However, RBF scored higher than Linear in ACD (75% by RBF to 66.6% by Linear) and also of AUC (85.7% by RBF to 81.8% by Linear)

In terms of Specificity [Fig. 5],

- SVM RBF scored: 76.7% of normal, 97.3% of ACD, 100% of ICD, 91% of AUC and 95.9% of IUC
- SVM Linear scored: 76.7% of normal, 97.9% of ACD, 100% of ICD, 88.5% of AUC and 85.9% of IUC
- SVM polynomial scored: 100% of normal, 100% oof ACD, 100% of ICD, 0% of AUC and 100% of IUC

We see that RBF and linear have same results of Normal and ICD, while RBF achieved higher score than Linear of AUC (91% by RBF to 88.5% of Linear) and of IUC (95.9% by RBF to 85.9% of Linear)

In terms of Accuracy [Fig. 6], SVM RBF scored 70.9%, SVM Linear scored 69.1% and SVM Polynomial scored 35.1%, The SVM RBF scored 70.9%, which is a slightly higher score (1.8%) of SVM Linear kernel (69.1%).

Comparing according to approaches using SVM:

We see higher scores in accuracy, recall and specificity in first approach than in the second approach

First Approach Accuracy: RBF: 76.7%, Linear: 79.07%, Polynomial: 44.1%.

Second Approach Accuracy: RBF: 70.9%, Linear: 69.1%, Polynomial: 35.1%.

➢ First Approach ANN

TABLE II. ANN Classification results: normal and IBD

Measures	NN
ACC %	82.7
REC	90
Spec	54.5

In the first classifier [Table II], classifying between normal and IBD cases, Accuracy is 82.7%, Recall is 90% and specificity was 54.5.

In the second classifier [Fig.7], classifying between Active CD(ACD), Inactive CD(ICD), Active UC(AUC) and Inactive (UC). Recall of ACD is 87.5%, of ICD is 60%, of AUC is 85%, of IUC is 72%.

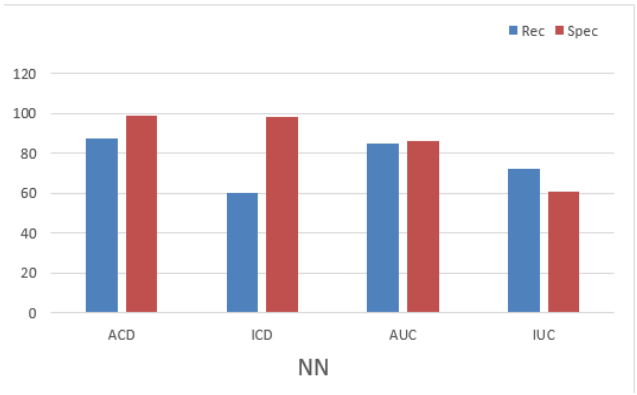

Fig. 7. IBD (ACD, ICD, AUC, IAC)

Specificity of ACD is 99.1%, of ICD is 98.2%, of AUC is 86.1%, of IUC is 60.2%.

Accuracy of the second classifier is 80%.

➢ Second Approach ANN

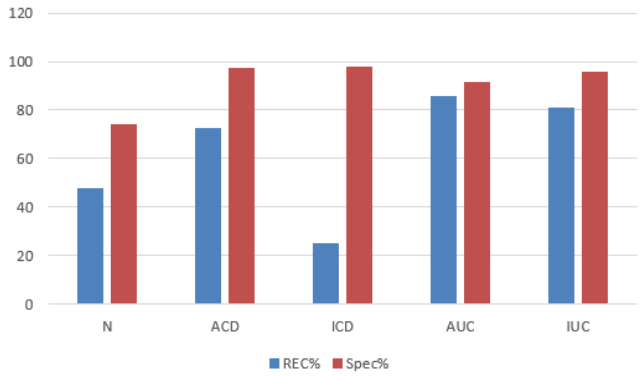

Fig. 8. ANN: Classification results: IBD (N, ACD, ICD, AUC, IAC)

In [Fig. 8] show that Recall of normal is: 47.6%, of ACD is 72.7%, of ICD is 25%, of AUC is 85.7% and of IUC is 80.9% and Specificity of normal is: 74%, of ACD is: 97.4%, of ICD is: 97.9%, of AUC is: 91.5%, of IUC is: 96%.

Accuracy of second approach model is: 67.9%

Comparing between approaches using NN: Accuracy of first approach is higher than(first classifier: 82.7%, second classifier: 80%) of the second approach(67.9%).

TABLE III. Summary Results

Measures	SVM						ANN	
	First Approach			Second Approach			First Approach	Second Approach
	Linear	RBF	Poly	Linear	RBF	Poly		
ACC	79.07	76.7	44.1	70.9	69.1	35.1	80 %	67.9%

V. Conclusions

In regard to the importance and difficulty of diagnosis of IBD disease. there are two approaches model were proposed to classify the activation of disease's sub types. The first approach was tested on dataset" Dataset of Observed Features on Endoscopic Colorectal Biopsies from Normal Subjects and Patients With Chronic Inflammatory Bowel Disease (Crohn's

Asmaa Mohamed Hassan and Yasser M. K. Omar, "Inflammatory Bowel Disease Classification Using Neural Network and Support Vector Machine," *International Research Journal of Advanced Engineering and Science*, Volume 5, Issue 1, pp. 216-221, 2020.

Disease and Ulcerative Colitis" by sample size of 644 patients leads us to a better experimental results using NN with accuracy 80%, with enhancement 0.93 over using SVM with linear kernel with accuracy 79.9%.

In the second approach, by using the same sample size dataset, the SVM is applied again in other approach showed at using of RPF kernel produced accuracy 70.9% which is better than applying NN with accuracy 67%. Accurate diagnosis depends on using varieties of algorithms and choosing its parameters. Using a small dataset and a limited number of layers in neural network justified this accuracy so, increasing the size of dataset with increasing the numbers in layers by neural network may lead to the increase of accuracy.

REFERENCES

[1] Talley, N. J., and S. O'Connor, "Clinical examination: a systematic guide to physical diagnosis," Elsevier Health Sciences, p. 227, 2013.
[2] H. Y. Su, V. Gupta, A. S. Day and R. B. Gearry, "Rising incidence of inflammatory bowel disease in Canterbury, New Zealand," Inflammatory bowel diseases 22.9, 2238-2244, 2016.
[3] S. C. Ng, H. Y. Shi, N. Hamidi, F. E. Underwood, W. Tang, E. I.Benchimol, and J. J. Sung, "Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies," The Lancet 390.10114, 2769-2778. 2017.
[4] Gophna, Uri, et al. "Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis." Journal of clinical microbiology 44.11 (2006): 4136-4141.
[5] Trappenberg, Thomas P. Fundamentals of Machine Learning. Oxford University Press, 2019.
[6] M. Fakhoury, R. Negrulj, A. Mooranian, and H. Al-Salami, "Inflammatory bowel disease: clinical aspects and treatments," Journal of inflammation research , 7, 113, 2014.
[7] E. Mossotto, J. J. Ashton, T. Coelho, R. M. Beattie, B. D. MacArthur and S. Ennis, "Classification of Paediatric Inflammatory Bowel Disease using Machine Learning," Scientific reports 7.1:1-10, 2017.
[8] D. Mahapatra, P. Schueffler, J. A. Tielbeek, J. M. Buhmann, and F. M.Vos, "A supervised learning based approach to detect crohn's disease in abdominal mr volumes," International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging. Springer, Berlin, Heidelberg, p. 97-106, 2012.
[9] A. Ryblov, S. Kolesov, E. Fedulova, N. Karyakin, M. Ivanchenko and A. Zaikin, "Comparison of Machine Learning Methods forA nalysis of Ulcerative Colitis Proteomic Data," Opera Medica et Physiologica 1 ,2017.
[10] N. Cruz-Ramírez, H. G. Acosta-Mesa, R. E. Barrientos-Martínez and L. A. Nava-Fernández, "Diagnosis of chronic idiopathic inflammatory bowel disease using bayesian networks," Iberoamerican Congress on Pattern Recognition. Springer, Berlin, Heidelberg, p. 706-715, 2006.
[11] Cross, Simon S. "Dataset of Observed Features on Endoscopic Colorectal Biopsies from Normal Subjects and Patients with Chronic Inflammatory Bowel Disease (Crohn's disease and Ulcerative Colitis)." *Department of Pathology, University of Sheffield Medical School* (1999): 1-15.
[12] S. Okada, M. Ohzeki, and S. Taguchi, "Efficient partition of integer optimization problems with one-hot encoding," Scientific reports 9.1, 1-12, 2019.
[13] Amal Nair, January 2019. Hands-On Tutorial on Data Pre-Processing in Python [E-reader version]. Retrieved from https://www.analyticsindiamag.com/data-pre-processing-in-python/ , Hands-on Tutorial on Data pre-processing InPytho.
[14] J. Berbić, E. Ocvirk, D. Carević, and G. Lončar, "Application of neural networks and support vector machine for significant wave height prediction," Oceanologia 59.3, 331-349, 2017.
[15] I. Livshin, "Internal Mechanics of Neural Network Processing," Artificial Neural Networks with Java. Apress, Berkeley, CA, p. 7-19, 2019.
[16] Rebala, Gopinath, Ajay Ravi, and Sanjay Churiwala. An Introduction to Machine Learning. Springer, 2019.
[17] Trappenberg, Thomas P. Fundamentals of Machine Learning. Oxford University Press, 2019.
[18] Dataflair Team, November 2018. Kernel Functions-Introduction to SVM Kernel & Examples [E-reader version]. Retrieved from. https://data-flair.training/blogs/svm-kernel-functions/
[19] Alaa Tharwat," Classification assessment methods" , 17 August 2018.