The effect of gamma value on support vector machine performance with different kernels

Intisar Shadeed Al-Mejibli[1], Jwan K. Alwan[2], Dhafar Hamed Abd[3]
[1,2]University of Information Technology and Communications, Iraq
[3]Al-Maaref University College, Iraq

Article Info	ABSTRACT
	Currently, the support vector machine (SVM) regarded as one of supervised machine learning algorithm that provides analysis of data for classification and regression. This technique is implemented in many fields such as bioinformatics, face recognition, text and hypertext categorization, generalized predictive control and many other different areas. The performance of SVM is affected by some parameters, which are used in the training phase, and the settings of parameters can have a profound impact on the resulting engine's implementation. This paper investigated the SVM performance based on value of gamma parameter with used kernels. It studied the impact of gamma value on (SVM) efficiency classifier using different kernels on various datasets descriptions. SVM classifier has been implemented by using Python. The kernel functions that have been investigated are polynomials, radial based function (RBF) and sigmoid. UC irvine machine learning repository is the source of all the used datasets. Generally, the results show uneven effect on the classification accuracy of three kernels on used datasets. The changing of the gamma value taking on consideration the used dataset influences polynomial and sigmoid kernels. While the performance of RBF kernel function is more stable with different values of gamma as its accuracy is slightly changed.

Corresponding Author:

Intisar Shadeed Al-Mejibli,
Smart Cities,
University of Information Technology and Communications,
Al-Nidhal Street, Baghdad, Iraq.
Email: dr.intisar.almejibli@gmail.com; dr.intisar.almejibli@uoitc.edu.iq

1. INTRODUCTION

Support vector machine (SVM) is supervised learning models have the ability of analyzing data for classification and regression purposes. SVM is supported by the theory of statistical learning, which has been developed by Vladimir [1]. Although, there are many algorithms were developed for pattern classification, SVM is the best one among them. Newly, it was adapted for new implementations for instance discovering regression and distribution estimation. Currently, SVM is implemented in different fields for instance bioinformatics, which is regarded as so promised research area for researchers.

In general, the process of analyzing dataset items is based on their labels, where labels represent the class information of related data included in that datasets. Classification methods is implemented in many research fields such as bioscience, economy and forecasting because it support the decisions and summarize the collected data to be analyzed efficiently. Thus, classification methods are used by economists and doctors and not limited to computer scientists [2-4]. The SVM principle is built on binary classification where it uses straight line to separates data points to classify the class label. Nevertheless, straight line cannot be used to separate the data points in some datasets, so the Kernel functions were revealed to tackle this issue [5, 6].

The SVM performance is based on both its parameters and used kernel function, which is assisting SVM to achieve the best separating hyperplane in the feature space. One of the significant SVM parameters is Gamma. The unintuitive parameter gamma defines the width or slope of kernel function. The 'curve' of the decision boundary becomes very low when gamma value is low making the decision region very broad. The 'curve' of the decision boundary becomes high when gamma is high, which creates islands of decision-boundaries around data points. Gamma parameter is involved when the polynomials, RBF and Sigmoid are set [7]. This paper aims to investigate the impact of gamma value with kernel on performance of SVM using different datasets, in order to specify the best performance of SVM. The methodology of this study based on using different values of gamma parameter with different kernel functions to evaluate the performance of SVM method. Various datasets descriptions is used in this study, which give the evaluation task another axis.

In general, this research studies the impact of gamma value, kernel functions and datasets descriptions on SVM performance. Three kernel functions were studied with five different datasets which having various classes instances. Polynomials, radial based function (RBF) and Sigmoid are the investigated kernel functions. The main objective of this study is reveal the performance of SVM using different values of Gamma parameter with the used kernel functions and datasets.

The rest part of this research is organized as following: Section 2 presents related work. Section 3 explaines support vector machines (SVM). Datasets is described in section 4. Section 5 discusses the results. Conclusion is illustrates in section 6.

2.   RELATED WORK

The performance of SVM related to gamma parameter values in kernel functions has been studied by few researchers. The following abstracted some of the published studies:

The authors in [8] presented a review in the efficacy of the choice of kernels for support vector machines and two different dimensionality techniques, which are fixed slope regression and principal component analysis. A leave–one-out classifier used to determine both of dimensionality reduction combination and accuracy. A high-dimensional bio-imaging dataset is used and the results showed that RBF kernel achieved the best accuracy.

In [9] the authors compare the performance of SVM classification algorithm for all different kernels (linear, polynomial, RBF and sigmoid) with different parameters (C-SVC, NU-SVC) to filter junk or spam E-mails. A spam-base dataset from public domain website (UCI respiratory) used to implement the comparison. The SVM obtained best accuracy when linear kernel function and SVC are implemented as the results showed.

Paper in [10] implemented SVM by using eleven mathematical functions called kernels. This research was implemented on images of Arabic alphabetic character. The dataset used in this paper was a database created from character images. The salt and pepper noise from 0.1 to 0.9 are used to corrupt characters. The results showed the relation between noise level and efficiency. When the noise level is increased the efficiency are regularly decreased for both Polynomial and linear kernels.

The paper [11] presented a comparative study for SVM implementation with four different kernels (linear, radial basis, polynomial, and sigmoid). This comparative study used many different datasets. The linear kernel obtained best result as it obtained 88.20 % in accuracy and also showed that the prediction time was faster about 4.078 sec.

In this paper [12] another comparative study have been done to study the implementation of linear and RBF SVM for fMRI classification with voxel selection schemes. This study compared the classification accuracy and time-consuming to decipher brain patterns from functional MRI (fMRI) data. Six different methods of voxel selection have been employed to specify, which voxels of fMRI data would be included in SVM classifiers with linear and RBF kernels in classifying 4-category objects. Results revealed that the RBF SVM was outperform linear SVM when implementing a feature space of a relative low dimensional. Whereas, linear SVM outperformed RBF SVM when implementing a feature space of a relative high dimensional. In [13] the authors investigated the SVM implementation with linear, polynomial and Radial Basic Function kernels using three datasets for the purpose of text independent speaker. The SVM with polynomial kernel achieved the best performance as the results showed.

The authors in [14] presented the implementation of SVM with four kernels: Linear, Radial Basis, Polynomial and Quadratic to solve the speaker identification task using TIMIT corpus. Four datasets were used to implement the experiment. The Quadratic kernel outperform the other three kernels as the results showed.

In [15] the authors analyzed the implementation of different kernel functions in SVM such as Radial Basis, Linear, Quadratic and Polynomial in recognizing the emotions of face. A database is created using

digital camera and based on 51 persons, which includes 714 images of face emotion with seven facial expressions. The Quadratic kernel obtained best accuracy.

The authors in [16] discussed the implementation of kernel functions with the SVM classifier in field of automatic text classification. A few models for text representation including VSM, GVSM, and n-grams, strings are discussed; in most of them text is represented by a set of vectors in a high dimensional space.

The authors in [17] compared the performance of SVM when setting it with linear, polynomial and Radial Basis Function (RBF) kernels. Three datasets are applied to test the three kernels, which are DLBCL, brain tumor, and Tumors. The testes were conducting with feature selection in SVM classification kernel function and without it for comparison purpose. The results showed that linear kernel function obtained high accuracy for the whole dataset without feature selection. Whereas, the dataset that includes informative genes gets high accuracy of 90 percent and above.

In [18] the authors applied detailed assessment of resampling procedures for SVM hyper parameter selection as they aimed to enrich literature in machine learning. The conducted empirical evaluation based on testing 15 procedures of various resampling on 121 binary classification datasets to choose the best SVM hyper parameters. Analyzing the obtained results required using three different statistical procedures, includes the standard multi-classifier/multidata set procedure, the confidence intervals on the excess loss of each procedure, and the Bayes factor analysis. The results showed that the 2-fold procedure is convenient to select the hyper parameters of an SVM for datasets with 1000 or more of data points, whereas a 3-fold procedure is convenient for datasets with smaller data points.

The authors in [19] proposed a kernel function for (OAO) and (OAA) multiclass SVM's. Many metrics have been measured including accuracy, support vector percentage, speed and classification error, for OAO and OAA performance evaluation. The results revealed the ability of generalizing more kernels and it revealed that the polynomial is outperform the other kernel functions in SVM.

In [20] the authors presented an assessment study includes five kernel functions with SVM in term of sensitivity classification based on POS sequences. The evaluated kernel were Linear, Spectrum, Gaussian, Smith-Waterman and Mismatch. This paper explained that the combination of POS sequence classification and text classification was enhance the effectiveness of sensitivity classification. The authors in [21] presented a study of SVM performance with different kernels. The results showed the dataset description and the implemented kernel impact the accuracy of the SVM classifier.

3.   SUPPORT VECTOR MACHINES (SVM)

SVM is machine learning and it is regarded as a part of neural networks that used in data mining and pattern recognition. The main concept of SVM is illustrated in Figure 1. SVM techniques can be explained as a process of separating two different classes in feature space such as positive and negative.

Discovering the hyperplane that separates those classes is the main issue in classification process where this hyperplane is depending on maximal margin. Several optimization problems can be tackled by SVM such as regression issues and data classification issues. Distinguish the positive and negative data points requires identify a hyper-plane to separate the data points as two classes. Figure 1 shows the hyperplane that it expected as decision boundary for linear SVM.
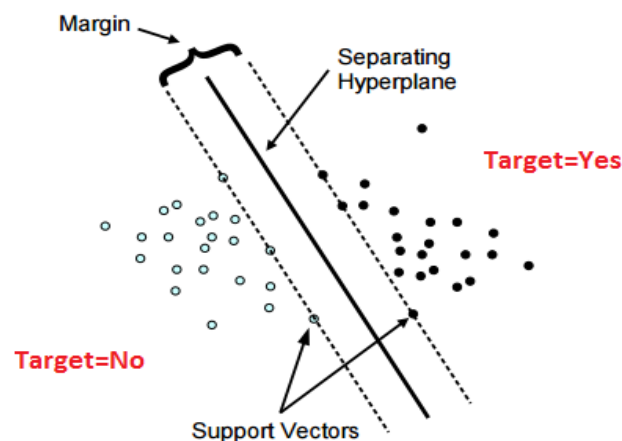

Figure 1. Data classification

SVM has three types, which are linear SVMs, nonlinear SVMs and multiclass SVMs. Linear support vector machines tackles the issue of binary classification of targeted data points by dividing them into two classes. Figure 1 showed points of data are linearly separable (two-dimensional case). When the two classes are non-linearly separable, it called non-separable. Thus, determining the hyperplane specify the ability of classify the data points linearly or nonlinearly (caused by, for example noise) [21].

When linear SVM is not proper for the used dataset, the Nonlinear SVMs provides solution. The concept of linear SVM has been extended to tackle the nonlinear case. The main concept of nonlinear SVM is to locate an optimal separating hyperplane in high-dimensional feature space similar to the linear SVM in input space. This implemented same as in linear SVM but instead of computing the inner products in feature space, which would be computationally expensive because of its high dimensionality. This computations are performed by using a nonlinear kernel function, in input space, which helps speed up the computations [22]. Kernel functions are implemented with SVM to solve many issues such as pattern analysis. Generally, pattern analysis can help in determining and studying various types of relations in dataset such as clusters, classifications, principal components, rankings and correlations.

Multiclass support vector machines are series of binary problems and it has two types. First named as one-versus-rest which divides the K-class problem into K binary classification sub-problems of the type "kth class" vs. "not kth class," k=1, 2, . . .,K. Second type is called one-versus-one, which divides the K-class problem into comparisons of all pairs of classes. Constructing a true multiclass SVM classifier, require considering all K classes simultaneously. In addition, the classifier must reduce the binary SVM classifier if K=2 [22]. This paper focusing on investigating the nonlinear support vector machines with kernel functions.

3.1. General problem formulation

Generating a non-linear separator is the main concept of SVM. The task of mapping data to high dimensional space is performed to facilitate the classification of data in linear decision surfaces. Hence, reformulation data is performed to implicitly mapped data to this space. Reformulation task is showed in Figure 2.
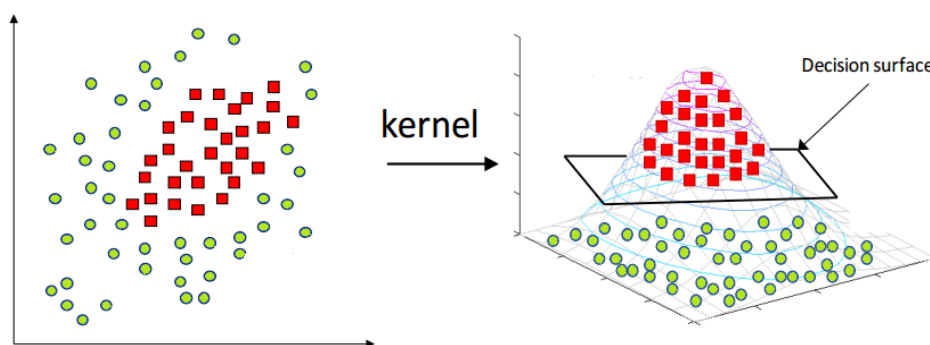

Figure 2. Data classification using kernel

This transformation is performed by kernel functions, so the data is transformed into a space of K(X, Y) which is higher dimensional feature by kernel function to facilitate the process of linear separation. SVMs can be implemented as both linear classifier and non-linear classifier. Non-linear classifier must implement specified kernel function to transform the points of dataset into higher dimensional of feature space [23].

3.2. Kernel functions of SVM

SVM is a technique of discovering best boundary to specify the classes of some features and outlier detection. SVM is named as ideal SVM when the analysis results in a hyper-plane that able to completely separate the data points into two different classes. Kernel functions offer the ability of computing dot products in higher dimensional spaces without the need to explicitly mapping into these spaces. Various mathematical properties must be satisfied by a kernel function. One of the important issues is determining whether a given function k is a kernel function or not ?. Given x1, x2, … xm ∈ X, the (m × m) matrix K with elements $K_{ij} = k(x_i, x_j)$ is named as the Gram matrix (or kernel matrix). Based on Mercer's Theorem, k is a kernel if the Gram matrix is positive definite [24]. SVM implements several kernels to solve different issues.

The gamma parameter is used in some of these kernels. The most common kernels that used this parameter are polynomial, radial basis function (RBF) and sigmoid as explained below [25, 26].
Kernel Functions:

$$K(X,Y) = \begin{cases} (\gamma\, X \bullet Y + c)^d, \gamma > 0 & \text{Polynomial} \\ \exp(-\gamma\, |X \bullet Y|^2), \gamma > 0 & \text{RBF} \\ \tanh(\gamma\, X \bullet Y + c) & \text{Sigmoid} \end{cases} \tag{1}$$

where X and Y act as the vectors in the input space, γ represents the gamma parameter in the kernel function and c is the bias parameter in the kernel function. γ, d, and c, are user-controlled parameters. Gamma value is calculated by dividing 1 to number of featured in the employed dataset. Each kernel function has specific parameters that need optimization to achieve the best results [27]. Here these parameters are named as optimized parameters. The three investigated kernels are as follows [28-30]:

3.2.1. Polynomial kernel

Sometimes noise in data or bad feature representation result in not possibility of linear separable for data. To tackle this issue the data points is mapped into different space. Polynomial kernel function for both hard margin and soft margin are to deal with nonlinear separable patterns. Polynomial kernel function applied by using mathematical function as expressed in (2),

$$K(X,Y) = (X \bullet Y + c)^d \;\; \gamma > 0 \tag{2}$$

where $c \geq 0$ acts as a free parameter in the polynomial that trading-off the impact of higher-order versus lower-order terms [16, 23, 30]. d, γ, c and d are the optimization parameters [31, 32].

The polynomial is regarded as non-stationary kernel. It is appropriate to be used in the issues, which all of its training samples are normalized. The slope gamma parameter should be settled with this kernel. c is the constant term and d is the polynomial degree (where d=3, c=0).

3.2.2. RBF kernel

The Radial basis function also called the RBF kernel, or Gaussian kernel. It is applied by using mathematical function that expressed in (3):

$$K(X, Y) = \exp(-\gamma\, |X - Y|2) \;\; \gamma > 0 \tag{3}$$

To set the "spread" of the kernel, the γ parameter is used [16, 23, 33]. The adjustable parameter γ must carefully tuned as it significant impact on this kernel performance. The exponential behavior will almost change to linearly if this parameter is overestimated. This also will make the higher dimensional projection losing its non-linear power. RBF kernel suffers from lack regularization and the decision boundaries tend to become highly sensitive to noise in training data, if underestimated. Thus, the value of γ parameter must be carefully chosen as it influence the performance of SVM [31].

3.2.3. Sigmoid kernel

As mentioned previously, each kernel must be positive definite to satisfy Mercer's theorem. Although, the Sigmoid kernel function is commonly implemented, it is not positive semi-definite for specific values of this kernel parameters. Consequently, the parameters γ and c must be carefully selected to avoid any mistake in the obtained results.

Mathematical function of sigmoid kernel presented in (4):

$$K(X, Y) = \tanh(\gamma\, X \bullet Y + c) \tag{4}$$

where γ acts as a scaling parameter of the input data and c acts as a shifting parameter, which is supervise the mapping threshold (hence c =0) [15, 16]. γ and c are the optimization parameters [30].

4. DATASETS

UC Irvine Machine Learning Repository [34] is the source of the used datasets in this study. Where five different datasets have been transformed to numeric and implemented in this investigation. Table 1 shows the used datasets.

The downloaded datasets are: Mushroom, Chronic Kidney Disease, Breast cancer, Lung cancer and Heart statlog. Three kernels with SVM were achieved for all the aforementioned datasets. Table 1 describes some features of the datasets used such as attributes, instances, classes in addition to missing values (if exists). Where the described features in Table 1 will used in assessment task later.

Table 1. The datasets description

	Name	No. of Attributes	No. of Instances	No. of Class	Missing in vale Attributes
1	Mushroom	22	8124	2	Yes
2	Chronic Kidney Disease	25	400	2	Yes
3	Breast cancer	9	286	2	Yes
4	Lung cancer	6	59	2	Yes
5	Heart statlog	13	270	2	No

5.   RESULT

To start the training and testing phases of SVM, each of the used datasets have been divided into two sets: training set and testing set based on their number of instances. Table 2 showed training sets and testing sets of used datasets.

Table 2. The size of training set and testing set

Dataset	Training Size	Testing Size
Mushroom	5686	2438
Chronic Kidney	280	120
Breast cancer	200	86
Lung cancer	41	18
Heart statlog	189	81

As aforementioned the gamma value is calculated by dividing 1 to number of featured in the employed dataset.

$$\gamma \text{ (gamma)} = 1/\text{No. of Features} \tag{5}$$

Figure 3 shows the flowchart of used algorithm to assess the implemented kernels with different values of gamma parameter. The dataset represents the input to proposed algorithm, where it will be classified by applying the SVM with three kernels. Then, the resulted validity and accuracy are checked. Where different values of gamma parameter are implemented with kernels to specify the hyperplane.

Applying the SVM successfully require select the proper parameters that SVM and kernel functions need. For example the kernel width gamma (γ), the constant parameter c which is constant in the kernel function and kernel degree (d).
The following parameters' values ere applied in three kernel functions:
γ : Its value is (0.3, 0.6, 0.9).
c : (c= 0)
d : kernel (d= 3).
To validate all implemented experiments, we used 10-fold cross validation.

Table 3 shows the achieved accuracy of Sigmoid, Polynomial and Rbf when the gamma value is calculated based on (5). Three different values of gamma (0.3, 0.6, 0.9) are used in the three kernel functions that are implemented with SVM applied to the five datasets. These values have been proposed, to be close to the gamma value based on equation 5 and to measure the achieved accuracy when the gamma value are changed. The resulted classification accuracy of implemented kernels on datasets are affected by gamma value changing unevenly. The Figures 4-6 present the SVM acuuracy with polynomial, radial basis function (RBF) and sigmoid respectively depending on the used gamma value. In general, the results can be described depending on three factors, which are the changing in the value of gamma parameter, employed dataset and used kernel function. The Polynomial kernel achieved best results when the gamma value is calculated as in (5) with all the employed datasets except the Chronic Kidney dataset, which shows no change in obtained accuracy. From dataset point of view, the Polynomial Kernel achieved best results in accuracy when the Mushroom dataset is employed as the achieved accuracy was no less than 99.9%.

The accuracy of RBF kernel function is affected slightly when the gamma value is changed as shown in Figure 5. The accuracy of RBF kernel when Chronic Kidney and Lung cancer datasets are used show no impact with the changing of gamma values. Same as Polynomial Kernel, the RBF achieved highest accuracy when the Mushroom dataset is employed by no less than 99.876% with all the used gamma values.
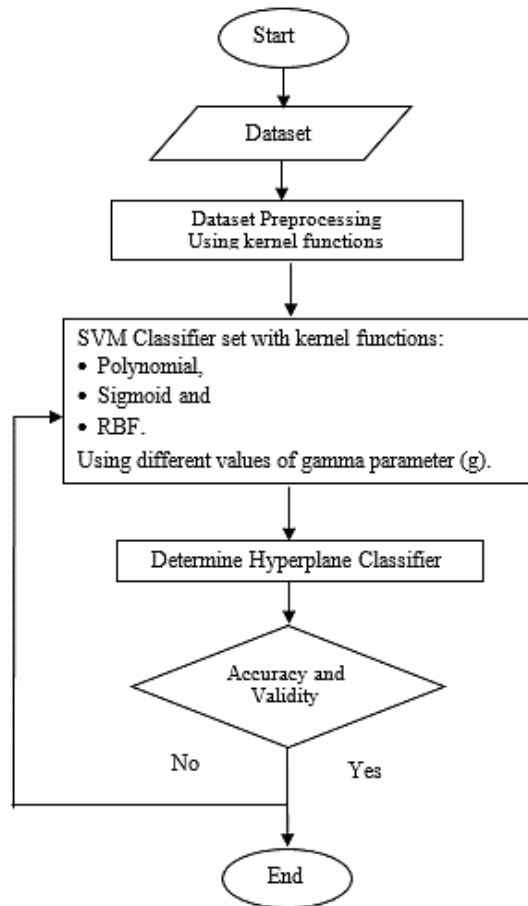
Figure 3. Flowchart of the implemented algorithm

Table 3. The accuracy of polynomial, rbf and sigmoid

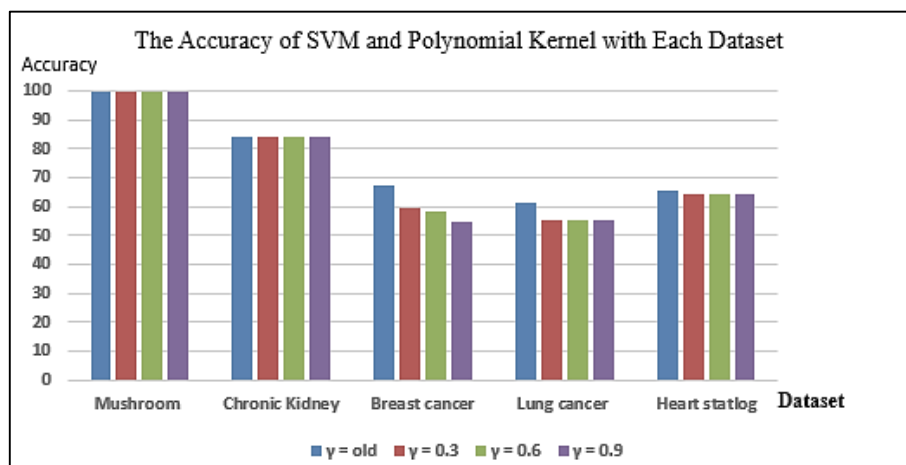| Dataset	Number of features	Gamma value	Polynomial	Rbf	Sigmoid
Mushroom	22	0.0454	100%	100%	24.405%
Chronic Kidney	24	0.0416	84.166%	61.666%	61.666%
Breast cancer	9	0.1111	67.441%	74.418%	62.790%
Lung cancer	6	0.1666	61.111%	33.333%	33.333%
Heart statlog	13	0.0769	65.432%	49.382%	50.617%

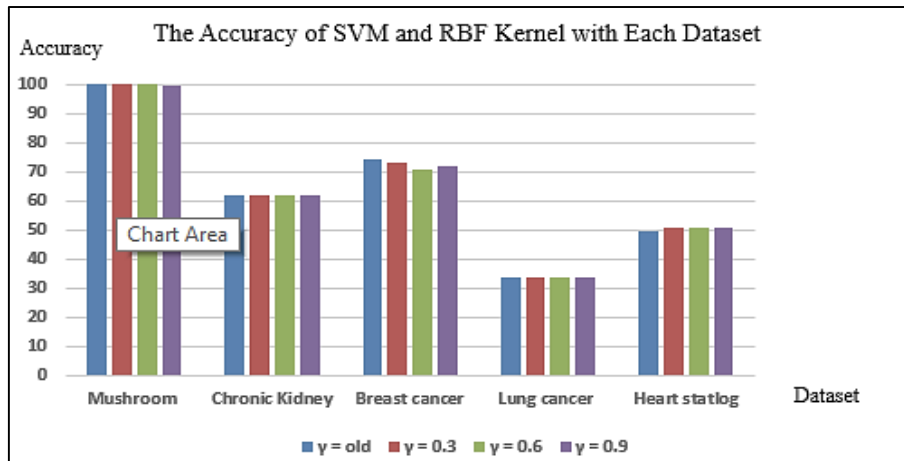Figure 4. SVM accuracy of polynomial kernel according to each dataset

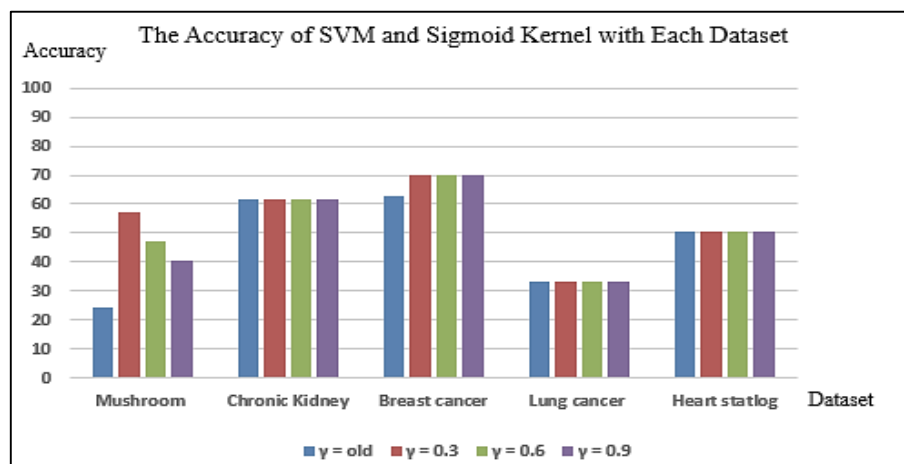Figure 5. SVM accuracy of RBF kernel according to each dataset

Figure 6. SVM accuracy of sigmoid kernel according to each dataset

The accuracy of Sigmoid is influenced by the changing of the gamma value based on the used dataset as Figure 5 shows. Using the new proposed gamma values impact the accuracy of Sigmoid positively. The accuracy of SVM with Sigmoid is less than 70% when employing all the datasets, which is less than the previous two kernels. The Polynomial kernel achieved highest accuracy than RBF and Sigmoid with all the employed datasets except the breast cancer where RBF obtained the highest accuracy rate. Polynomial and RBF kernels achieved highest accuracy than Sigmoid with Mushroom dataset. The value of gamma value affect the accuracy of the SVM with used kernel and not always using (5) in calculating the value of gamma resulting in best accuracy.

6.   CONCLUSION

This study evaluates the performance of three kernels with SVM, which are used for classification based on different values of gamma parameter (γ) with various datasets descriptions. SVMs uses popular kernels for classification purposes where these kernel functions (polynomial, RBF, and sigmoid) used γ parameter. The assessment of applied experimental revealed that the gamma value, implemented kernel function and employed dataset descriptions affect the SVM accuracy. The changing in gamma value unevenly influenced the classification accuracy of three kernels on datasets as shown by the results.

The Polynomial kernel obtained high accuracy when the gamma value is calculated as (5) in all the employed datasets except the Chronic Kidney dataset, which shows no change in obtained accuracy. The RBF kernel function efficiency is stable using various values of gamma as its accuracy is slightly changed. The changing of the gamma value, taking on consideration the used dataset, influences the accuracy

of Sigmoid. Thus, the used dataset affects the performance of Sigmoid in addition to gamma value where some datasets show change in accuracy when the gamma value is changed such as in Mushroom and Breast cancer whereas the accuracy of other dataset were not changed.

The result shows that the performance of the kernel is affected by the number of attributes, instances, and missing values. In addition to, gamma value influences the kernel performance. The accuracy of SVM results increased by increasing of attributes number, instances number for dataset as in Mushroom and Chronic Kidney. From other hand, increasing the number of missing values affects the accuracy negatively. Various adaptations, examinations, and experiments based other SVM optimization parameters and features of datasets can be investigated in future. Future work may presents extensive analysis of SVM with other optimization parameters and datasets.

REFERENCES

[1] Vapnik, V. and A. Lerner, "Pattern recognition using generalized portrait method," *Automation and Remote Control,* vol. 24, pp. 774-780, 1963.
[2] Zhang, D., "Machine Learning in Value-Based Software Test Data Generation," *18th IEEE International Conference on Tools with Artificial Intelligence,* pp. 732-736, 2006.
[3] Wernick, M. N., Yongyi. Y. Brankov, J. G. Yourganov, G. and Strother, S. C., "Machine Learning in Medical Imaging," *IEEE Signal Processing Magazine,* vol. 27, no. 4, pp. 25-38, 2010.
[4] Yoo, P. D., Kim, M. H., and Jan, T., "Machine Learning Techniques and Use of Event Information for Stock Market Prediction: A Survey and Evaluation," *International Conference on Intelligent Agents, Web Technologies and Internet Commerce,* vol. 2, pp. 835-841, 2005.
[5] S Maji, A. C Berg., and J Malik, "Efficient classification for additive kernel SVMs," *IEEE transactions on pattern analysis and machine intelligence,*" vol. 35, no. 1, pp. 66-77, 2013.
[6] F. Gullo, "From Patterns in Data to Knowledge Discovery: What Data Mining Can Do," *Physics Procedia,* vol. 62, pp. 18-22, 2015.
[7] Rimah A., Dorra B. A., and Noureddine E., "Practical Selection of SVM Supervised Parameters with Different Feature Representations for Vowel Recognition," *International Journal of Digital Content Technology and its Applications*, vol. 7, no. 9, pp. 418-424, 2013.
[8] Michelle H., Shridharan C., Sumedha S., and Yang S, "Kernel Selection and Dimensionality Reduction in SVM Classification of Autism Spectrum Disorders," 2013.
[9] Deepak K. A., and Rahul K., "Spam Filtering using SVM with different Kernel Functions," *International Journal of Computer Applications,* vol. 136, no. 5, 2016.
[10] Sayed F., Said G., Mohamed A., Hussein A. S., Amira A., and Asif A., "Investigating the Effect of Different Kernel Functions on the Performance of SVM for Recognizing Arabic Characters," *International Journal of Advanced Computer Science and Applications,* vol. 7, no. 1, pp. 446-450, 2016.
[11] Supriya P., and Deepak S., "Comparison of Various Kernels Of Support Vector Machine," *International Journal for Research in Applied Science and Engineering Technology,* vol. 3, no. VII, pp. 532-536, 2015.
[12] Sutao S., Zhichao Z., Zhiying L., Jiacai Z., and Li Y., "Comparative Study of SVM Methods Combined with Voxel Selection for Object Category Classification on fMRI Data," *PLoS ONE,* vol. 6, no. 2, pp. 1-11, 2011.
[13] IIT Kanpur, "Support Vector Machines," *Programming Club*, Mar. 11, 2016. [online], Available: http://home.iitk.ac.in/~arunothi/pclub/ml/3.pdf.
[14] D. B. A. Mezghani, S. Z. Boujelbene, and N. Ellouze, "Evaluation of SVM Kernels and Conventional Machine Learning Algorithms for Speaker Identification," *International Journal of Hybrid Information Technology,* vol. 3, no. 3, pp. 23-34, 2010.
[15] I. A. Adeyanju, E. O. Omidiora, and O. F. Oyedokun, "Performance Evaluation of Different Support Vector Machine Kernels for Face Emotion Recognition," *SAI Intelligent Systems Conference,* pp. 804-806, 2015.
[16] Celso A. A. Kaestner, "Support Vector Machines and Kernel Functions for Text Processing," *Journal of RITA,* vol. 20, pp. 130-154, 2013.
[17] N. M. Hatta and Z. A.. Shah, "Evaluate the Performance of SVM Kernel Functions for Multiclass Cancer Classification," *UTM Computing Proceedings Innovations in Computing Technology and Applications,* vol. 1, pp. 1-7, 2016.
[18] J. Wainer and G. Cawley, "Empirical evaluation of resampling procedures for optimising SVM hyperparameters," *Journal of Machine Learning Research,* vol. 18, pp. 1-35, 2017.
[19] R. Sangeetha and B. Kalpana, "Performance Evaluation of Kernels in Multiclass Support Vector Machines," *International Journal of Soft Computing and Engineering,* vol. 1, no. 5, pp. 138-145, 2011.
[20] Graham McDonald, Nicol´as Garc´ıa-Pedrajas, Craig Macdonald and Iadh Ounis, "A Study of SVM Kernel Functions for Sensitivity Classification Ensembles with POS Sequences," *The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval,* pp. 1097-1100, 2016.
[21] I. S. Al-Mejibli, D. H. Abd, J. K Alwan, and A. J. Rabash, "Performance evaluation of kernels in support vector machine," *1st Annual International Conference on Information and Sciences,* 2018.
[22] Ho Tu Bao, "Kernel Methods and Support Vector Machines," *Japan Advance Institute of Science and Technology,*
[23] A. N. Khobragade, M. M. Raghuwanshi, and L. Malik, "Evaluating Kernel Effect on Performance of SVM Classification using Satellite Images," *International Journal of Scientific and Engineering Research,* vol. 7, no. 3, pp. 742-748, 2016.

[24]   B. Scholkopf and A. J. Smola, "Learning with Kernels Support Vector Machines, Regularization," *Optimization, and Beyond*," *MIT Press*, 2002.
[25]   Andrew Ng, "Support Vector Machines," Part V, CS229 Lecture notes., [online], Available: http://cs229.stanford.edu/notes/cs229-notes3.pdf.
[26]   D. H. Abd, A. T. Sadiq, and A. R. Abbas, "Classifying Political Arabic Articles Using Support Vector Machine with Different Feature Extraction," in *International Conference on Applied Computing to Support Industry: Innovation and Technology*, Springer, pp. 79-94, 2019.
[27]   Zhang, Y., and Wu, L., "Classification of fruits using computer vision and a multiclass support vector machine," *Sensors,* vol. 12, pp. 12489-12505, 2012.
[28]   D. Abd, J. K. Alwan, M. Ibrahim and M. B. Naeem, "The utilisation of machine learning approaches for medical data classification and personal care system mangementfor sickle cell disease," *Annual Conference on New Trends in Information and Communications Technology Applications,* pp. 213-218, 2017.
[29]   M. Khalaf et al., "Recurrent Neural Network Architectures for Analysing Biomedical Data Sets," *10th International Conference on Developments in eSystems Engineering,* pp. 232-237, 2017.
[30]   Gualtieri. J. A, Chettri. S. R, Cromp. R. F and Johnson.L. F, "Support vector machine classifiers as applied to AVIRIS data," in *Summaries 8th JPL Airborne Earth Science Workshop, JPL*, pp. 217–227, 1999.
[31]   M. A. Nanda, K. B. Seminar, D. Nandika, and A. Maddu, "A Comparison Study of Kernel Functions in the Support Vector Machine and Its Application for Termite Detection," *Information,* vol. 9, no. 5, pp. 1-14, 2018.
[32]   M. Khalaf et al., "An Application of Using Support Vector Machine Based on Classification Technique for Predicting Medical Data Sets," in *International Conference on Intelligent Computing*, Springer, pp. 580-591, 2019.
[33]   Steve R. Gunn, "Support Vector Machines for Classification and Regression," *University of Southampton, Technical Report Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science,* 1998.
[34]   "UC Irvine Machine Learning Repository," *UCI*, [Online], Available: http://archive.ics.uci.edu/ml/index.php.

BIOGRAPHIES OF AUTHORS

Intisar Al-Mejibli, has PhD in Computing and Electronic Systems/ School of Computer Science and Electronic Engineering/ University of Essex/UK. Currently she is working as Assistant Professor in Smart Cities Section/ University of Information Technology and Communications/ Baghdad / Iraq. The interests of Intisar includes Wireless Sensor Networks, Security, (IoT) and healthcare systems. Further, Intisar interests include using Artificial Intelligent techniques to develop solutions for different.

Jwan K. Alwan has M.Sc. of information system 2014 from Osmania University in India and currently she is working as lecturer in University of Information Technology and Communications/Baghdad/Iraq. She is interested in pattern classification, medical computing, medical information systems, patient monitoring, support vector machines, learning (artificial intelligence), optimization.

Dhafar Hamed Abd is working as lecturer in Al-Maaref University College, Iraq. He is interested in diseases, pattern classification, health care, medical computing, medical information systems, patient monitoring, support vector machines, Bayes methods, data analysis, decision trees, expert systems, learning (artificial intelligence), optimization, patient treatment, polynomials, radial basis function networks, recurrent neural nets, smart phones, and software performance evaluation.