Homogeneous Ensemble based Support Vector Machine in Breast Cancer Diagnosis

Bouchra El Ouassif[1], Ali Idri[1,2] and Mohamed Hosni[1,3]
[1]*Software Project Management Research Team, ENSIAS, Mohammed V University, Rabat, Morocco*
[2]*MSDA, Mohammed VI Polytechnic University, Ben Guerir, Morocco*
[3]*MOSI, L2M3S, ENSAM-Meknes, Moulay Ismail University, Meknes, Morocco*

Keywords: Breast Cancer, Classification, Support Vector Machine (SVM), SVM Ensemble, Combined Kernel.

Abstract: Breast Cancer (BC) is one of the most common forms of cancer and one of the leading causes of mortality among women. Hence, detecting and accurately diagnosing BC at an early stage remain a major factor for women's long-term survival. To this aim, numerous single techniques have been proposed and evaluated for BC classification. However, none of them proved to be suitable in all situations. Currently, ensemble methods have been widely investigated to help diagnosis BC and consists on generating one classification model by combining more than one single technique by means of a combination rule. This paper evaluates homogeneous ensembles whose members are four variants of the Support Vector Machine (SVM) classifier. The four SVM variants used four different kernels: Linear Kernel, Normalized Polynomial Kernel, Radial Basis Function Kernel, and Pearson VII function based Universal Kernel. A Multilayer Perceptron (MLP) classifier is used for combining the outputs of the base classifiers to produce a final decision. Four well-known available BC datasets are used from online repositories. The findings of this study suggest that: (1) ensembles provided a very promising performance compared to its base, and (2) there is no SVM ensemble with a combination of kernels that have better performance in all datasets.

1 INTRODUCTION

Breast cancer (BC) remains the leading cause of mortality among women in many parts of the world. It is the most common invasive cancer, impacting 2.1 million women each year, and causes the greatest number of cancer-related deaths among women. The causes of BC are not fully understood. However, there are certain factors known to increase the risk of BC such as, age, genetic mutations, family history of BC, overweight, late menopause, late age at first childbirth (Sun et al., 2017). Because of this, it becomes important to diagnose BC as early as possible to provide a better chance for proper medical treatment and to reduce the death rate caused by it (Sun et al., 2017). When breast tumor is spotted, physicians will need to find out whether it is Benign or Malignant. Information technology in form of computer-aided diagnosis (CAD) that was first proposed by Johnston (1994), has made great changes to clinical decision making. In fact, during these last years , data mining models have been well used in clinical (Hosni et al., 2019; Idri et al., 2019, 2020; Kadi et al., 2017), various high-performance models will help physicians detect and predict medical situations, and provide a quick and accurate diagnosis (Topol, 2019). BC is one of the diseases that benefit from CAD, as well as many new data mining techniques (Chlioui et al., 2020; Eltalhi & Kutrani, 2019; Oskouei et al., 2017)

Medical diagnosis is considered as an important and complex task that needs to be carried out accurately and efficiently. Different classification techniques have been proposed and evaluated to classify the breast tumor, using information provided by the mammography to assist the physician to accurately diagnosis BC. According to the systematic map of Idri et al. (Idri et al., 2018), Support Vector Machine (SVM) was the second most frequently classification technique used for BC diagnosis with 25.56% of the 403 selected studies (Artificial Neural Networks were the first with 26.80%). Moreover, it was observed that the use of SVM for BC diagnosis is gradually increasing due to its excellent learning and generalization abilities (Idri et al., 2018). The main advantage of SVM is its ability to model

El Ouassif, B., Idri, A. and Hosni, M.
Homogeneous Ensemble based Support Vector Machine in Breast Cancer Diagnosis.
DOI: 10.5220/0010230403520360
In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2021) - Volume 5: HEALTHINF*, pages 352-360
ISBN: 978-989-758-490-9

complex nonlinear relationships by selecting a suitable kernel function. Indeed, the Kernel function transforms the training data so that a non-linear decision surface is transformed to a linear equation in a higher number of dimensions.

Identifying the most appropriate Kernel function to implement SVM in a given context is a challenging task. In fact, the good choice of the Kernel function can improve the performance of a SVM based classification (Bhavsar & Amit Ganatra, 2016; Trivedi & Dey, 2013). However, few studies in literature of BC classification have dealt with the evaluation of the performance of SVM using various kernel functions (Hussain et al., 2011; Rana et al., 2019), and generally they used one dataset to assess the accuracy of their developed models, which does not allow the findings to be generalized to other contexts. For instance, (Hussain et al., 2011) presented a comparative study of different kernel functions for BC detection; the focus of their study was on classification using SVM with different kernel functions; they employed four kernels (RBF, polynomial, Mahalanobis, and sigmoid) and showed the performance of each of them. They have conducted their experiments on the Wisconsin dataset and they found that SVM with sigmoid kernel showed the best results. Moreover, they suggested the use of a combination of different kernels for better detection results. Rana et al. (Rana et al., 2019) investigated three machine learning algorithms k-nearest neighbor (KNN), MLP and SVM. SVM has been investigated using a linear and quadratic kernel. They used the machine learning algorithms over a Microwave Breast Imaging Clinical Data. The purpose of their study was to develop an intelligent classification system to help clinicians to recognize breasts with lesions. They found that the SVM with the quadratic kernel achieved a higher accuracy when compared to other classification techniques used in the study.

Moreover, it is known that Ensemble Classifiers (EC) have attracted a huge research in the last decade and in general outperformed single classifiers (Dietterich, 2000; Hosni et al., 2019). To address the challenge of searching the most appropriate Kernel function for implementing SVM, the present study aimed to develop and evaluate homogeneous ensembles whose members are four variants of SVM. The four SVM variants used four different kernels: Linear Kernel, Normalized Polynomial Kernel, Radial Basis Function Kernel, and Pearson VII function based Universal Kernel. A MLP is used to combine the outputs of the base classifiers to provide the ensemble decision. The experiments were conducted on four well-known BC datasets available from online repositories.

To this end, we discuss the following research questions (RQs):

RQ1: Does SVM ensembles combining different kernels types perform better than single SVMs?

RQ2: Among the combinations of SVM kernels to construct ensembles, which of them provides a better performance?

The rest of this paper is structured as follows: Section 2 briefly presents the SVM classifier as well as its kernels, and the ensemble concept used. Section 3 presents an overview of related work investigating ensemble techniques in BC classification. Section 4 presents the experimental design followed in this empirical evaluation. Section 5 presents and discusses the empirical results. Conclusions and future works are presented in Section 6.

2 BACKGROUND

2.1 SVM

SVMs are a set of supervised learning methods characterized by the usage of kernels. The first formulation of SVM was proposed in 1992, called maximal margin classifier (Vapnik, 1992). SVMs are based on the search for the optimal margin hyperplane which, when possible, classifies or separates the data correctly while being as far as possible from all observations. The use of a margin-based criterion by SVMs, is attractive for many classification applications like Handwritten digit recognition, Bio-sequence analysis and Speaker Identification, (El Idrissi & Idri, 2020; Luxburg & Schölkopf, 2011; Yu & Kim, 2012). Although SVMs were originally used for linearly separable datasets to find the optimal separating hyper-plane (Bhavsar & Amit Ganatra, 2016) from the large number of separating hyper-plane, that optimally separates the data into two areas SVMs can be generalized to non-linear decision functions by using the so-called kernel trick (Schölkopf & Alexander, 2001).

2.2 Kernel Functions

SVMs are unable to find a linear hyper-plane that can separate the input data into classes in some cases (Kudo & Matsumoto, 2000). This problem can be tackled by transforming the input data that exists in high dimensional space by using some non-linear transformation function. By this process, the input data can be separated out in such a way that linear separable hyper planes can be found in that transformed space (Trivedi & Dey, 2013). However, due to the high dimensionality of the feature space, computation of inner products of two transformed data vectors would be practically unfeasible. This problem is tackled using “Kernel Functions” that can be used in place of the inner product of two transformed data vectors in feature space.

A good choice of Kernel function is very important for effective SVM based classification. An appropriate Kernel function provides learning capability to SVM (Trivedi & Dey, 2013). In the literature, several Kernels have been proposed. In this paper, we attempt to investigate the best choice among SVM kernels namely LK, PUK, NP, and RBF.

2.3 Ensemble Techniques

An ensemble techniques is a machine learning technique that combines several single techniques by means of an aggregation rule in order to produce one optimal predictive model (Idri et al., 2016; Kuncheva, 2014).

Ensemble methods can be categorized into two types Homogeneous or Heterogeneous (Idri et al., 2016; Zhou, 2012): (1) The homogeneous ensemble refers to: (1.a) an ensemble that combines one base method with at least two different configurations or different variants or (1.b) ensemble that combines one base learning model with one meta ensemble, such as Boosting (Schapire & E., 1999), Bagging (Breiman, 1996). (2) The heterogeneous ensemble, meanwhile, refers to an ensemble that combines at least two different base learning models. The current study concerns homogeneous ensembles and for the trained combining rule, a MLP was used to combine the output of the classifiers.

3 EXPERIMENTAL DESIGN

3.1 Performance Measures

To evaluate the accuracy of single and ensembles techniques, we use the performance metrics: Accuracy, Recall and Precision defined by Equations 1 to 3 respectively (Hosni et al., 2019).

$$Accuracy = \frac{TN+TP}{TP+FP+TN+FN} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Precision (Prec)} = \frac{\text{TP}}{\text{TP+FP}} \quad (3)$$

Where FP refers to False positive, TP refers to True positive FN to False Negative and TN to True negative.

3.2 Proposed Ensembles

The purpose of this study was to develop and evaluate SVM homogeneous ensembles with different kernel types: LK, NP, RBF, and PUK. We build eleven SVM ensembles: six ensembles with two members (i.e. two different kernels) each one, four ensembles with three members each one (i.e. three different kernels), and one ensemble with four members (i.e. four kernels). The MLP was used as a combination rule to provide the output of each ensemble. The use of MLP as a combiner rule has been widely investigated in the literature of ensembles (Canuto et al., 2007; Santana et al., 2008; Tsymbal et al., 2005).

To shorten the names of ensembles, the following abbreviation rules were used:

E-Kernel Type1KernelType2
E-Kernel Type1KernelType2 Kernel Type3
E-Kernel Type1 KernelType2 Kernel Type3 Kernel Type4

Where kernel types were abbreviated as follow: L for Linear kernel, N for Normalized Polynomial Kernel, P for PUK and R for RBF

For example, ELNR refers to the ensembles constructed by the three variants of SVM using linear kernel, Normalized polynomial kernel, and RBF kernel respectively.

3.3 Methodology Used for Comparison

The purpose of this empirical study is to build ensembles based on SVMs with different kernels (LK, NP, RBF, PUK) using MLP as a combiner rule, and to compare them with the four single SVM techniques (SVM-LK, SVM-PUK, SVM-NP and SVM-RFB). The comparison is based on the three performance criteria: Accuracy, Recall and Precision. Moreover, we evaluate the statistical significance between ensemble and single SVM techniques by clustering them using the Scott–Knott (SK) test based on error rate (the percent of incorrect classifications; Error rate = 1 - Accuracy). Thereafter, we rank the techniques belonging to the SK best clusters by means of Borda Count based on Accuracy, Recall and Precision. The statistical test was conducted using the R Software and Weka (version 3.9.3) tool was used to conduct the empirical evaluations. Figure 1 presents the experimental process we followed.

3.4 Datasets Descriptions

In this study, three datasets were used to assess the performance of ensemble and single SVM techniques. These datasets were obtained from the online UCI repository and were the most frequently adopted by researchers (Idri et al., 2018).

Table 1 reports the characteristics of each dataset including number of instances, number of features, and the number of missing values. It is worth noting that before we conducted all the experiments, the missing values were removed since their number was very low in each dataset (column "Missing Data" of Table 1). Moreover, the three datasets: BCD, Wisconsin and WPBC represent unbalanced datasets. To address this issue, the SMOTE (Chawla et al., 2002) algorithm was used.

Table 1: Datasets Description.

Datasets	#Features	Missing data?	Instances
BCD	12	Yes(9)	286
WDBC	32	NO	569
Wisconsin	11	Yes(16)	699
WPBC	34	Yes (4)	198

3.5 Single Technique Parameters

The tuning of SVM and MLP parameters is given in Table 2. For SVM, four type of kernels were used (RBF, LK, PUK and NP), while the parameter values of MLP were the default values of Weka.

Table 2: Parameters settings of SVM and MLP.

Technique	Parameters
SVM	Epsilon: 1.0E-12; Complexity : 1.0 Kernel :{RBF, LK, PUK, NP)
MLP	Lerning rate :0.3;Momentum :0.2 Hidden layers:a; Validation threshold: 20

*a = (#attributes + #classes)/2

4 EMPIRICAL RESULTS

4.1 WDBC Dataset

Table 3 shows the performance in terms of three criteria (Accuracy, Recall and Precision) of the four SVM single techniques (SVM-LK, SVM-NP, SVM-RBF and SVM-PUK), as well as the performance of the eleven SVM ensemble. As can be seen from Table 3, the best results were obtained by SVM ensembles combining: LK and NP (ELN); LK and RBF (ELR); LK, NP and RBF (ELNR), since they produce the high performance with accuracy, recall and precision values equal to 98.07%, 98,1% and 98.1% respectively. We observe that the SVM ensembles ELN, ELR and ELNR outperform all the Single SVM techniques and all other SVM ensembles. Moreover, the results from the Table 5 show that the performance of the SVM single with linear kernel is also high 97.89%, 97.9% and 97.9% for accuracy, recall and precision respectively.

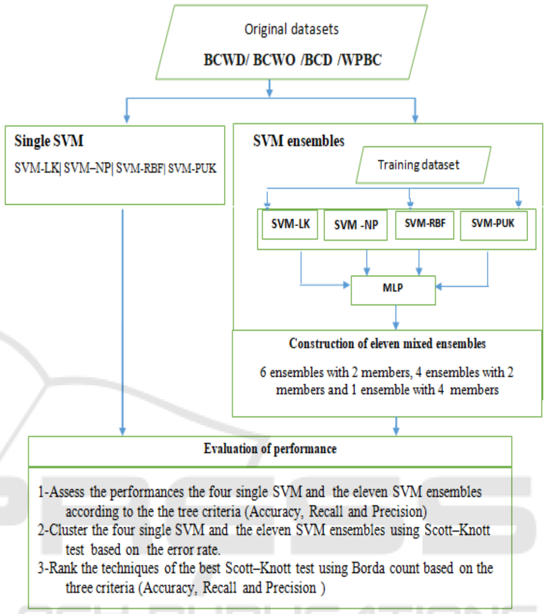

Figure 1: Experimental process.

Moreover, Figure 2 shows the results of the SK test performed based on error rate. We observe that SK test identified three clusters which means that there is a significant difference between ensemble and single SVM classifiers. The best cluster contains all SVM ensembles except the ensemble combining NP and RBF. Moreover, the best cluster included also two single SVM classifiers: SVM-PUK and SVM-LK. Whereas the worst cluster contained only one single SVM technique: SVM-RBF.

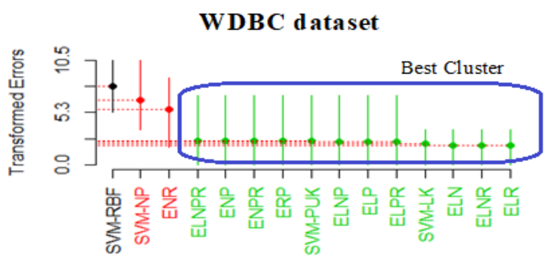

Figure 2: SK test of SVM single and SVM ensemble models on WDBC dataset.

4.2 BCD Dataset

Table 4 presents the performance of single and ensemble SVM techniques over the BCD dataset. As it can be observed from Table 4, the two SVM ensembles ENPR and ELNPR were ranked first with the highest performances: ENPR comes first with 85.43%, 85.4% and 85.7% for accuracy, recall and precision respectively; and ELNPR comes next with 85.18%, 85.2% and 85.4% for Accuracy, Recall and Precision respectively. We note that the single SVM-PUK shows the best performance compared to the other three single SVMs and outperforms six SVM ensembles (ELN, ELR, ELP, ENR, ERP, ELNPR).

Table 3: Performance results: WDBC dataset.

	Tech.	Accuracy	Recall	Prec
Single	SVM-LK	97.89	97.9	97.9
	SVM-NP	93.4	93.5	93.5
	SVM-RBF	92.09	92.1	92.8
	SVM-PUK	97.54	97.5	97.5
Ensemble	ELN	**98.07**	**98.1**	**98.1**
	ELR	**98.07**	**98.1**	**98.1**
	ELP	97.71	97.7	97.7
	ENR	94.38	94.4	94.4
	ENP	97.54	97.5	97.5
	ERP	97.54	97.5	97.5
	ELNR	**98.07**	**98.1**	**98.1**
	ENPR	97.54	97.5	97.5
	ELNP	97.71	97.7	97.7
	ELPR	97.71	97.7	97.7
	ELNPR	97.71	97.7	97.7

Table 4: Performance results: BCD dataset.

	Tech.	Accuracy	Recall	Prec
Single	SVM-LK	78.76	78.8	78.9
	SVM-NP	80	80	79.9
	SVM-RBF	79.26	79.3	79.2
	SVM-PUK	83.21	83.2	84.1
Ensemble	ELN	79.51	79.5	79.4
	ELR	79.75	79.8	79.7
	ELP	81.73	81.7	82.3
	ENR	80	80	79.9
	ENP	83.46	83.5	83.7
	ERP	82.96	83	83.3
	ELNR	79.51	79.5	79.4
	ENPR	**85.43**	**85.4**	**85.7**
	ELNP	84.44	84.4	84.7
	ELPR	84.44	84.4	84.8
	ELNPR	**85.18**	**85.2**	**85.4**

Figure 3 displays the result of the SK test on BCD dataset. Two clusters were identified and the best cluster contains six SVM ensembles (ENPR, ELNPR, ELPR, ELNP, ENP and ERP) and one single SVM (SVM-PUK); the worst cluster contains three single SVM (SVM-LK, SVM-RBFand SVM-NP) and the five SVM ensembles (ELP, ENR, ELR, ELNR and ELN).

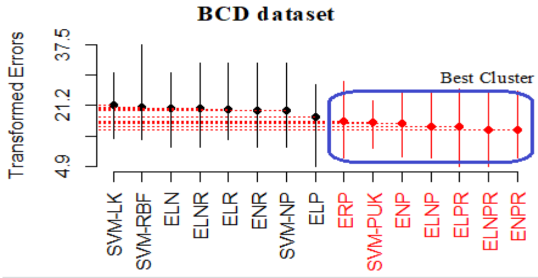

Figure 3: SK test of SVM single and SVM ensemble models on BDC dataset.

4.3 Wisconsin Dataset

Table 5 shows the performance criteria values of the ensemble and single SVM techniques over the Wisconsin dataset. As can be seen from Table 5, we observe that the SVM ensembles ERP and ELNP outperform all other techniques, they produce an accuracy, recall and precision values of 97.07%, 97.1% and 97.1% respectively. The remaining ensembles and single techniques showed almost the same performances.

Table 5: Performance results: Wisconsin dataset.

	Tech.	Accuracy	Recall	Prec
Single	SVM-LK	96.92	96.9	96.9
	SVM-NP	96.92	96.9	97
	SVM-RBF	96.34	96.3	96.3
	SVM-PUK	96.92	96.9	97
Ensemble	ELN	96.49	96.5	96.5
	ELR	96.92	96.9	96.9
	ELP	96.49	96.5	96.5
	ENR	95.46	95.5	95.5
	ENP	96.63	96.6	96.7
	ERP	**97.07**	**97.1**	**97.1**
	ELNR	96.49	96.5	96.5
	ENPR	96.63	96.6	96.7
	ELNP	**97.07**	**97.1**	**97.1**
	ELPR	96.92	96.9	97
	ELNPR	96.92	96.9	97

Figure 4 displays the results when applying the SK test on ensemble and single SVM techniques over Wisconsin dataset. SK identified only one cluster that included all techniques (single and ensemble SVMs), which implies that there is no important difference between them.

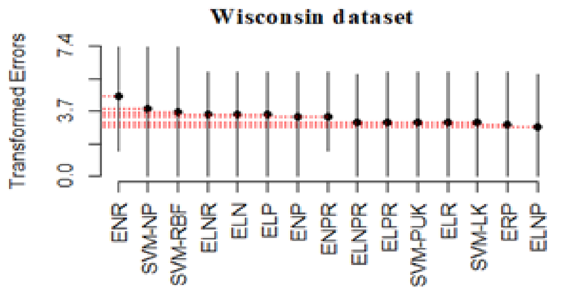

Figure 4: SK test of ensemble and single SVM techniques over Wisconsin dataset

4.4 WPBC Dataset

Table 6 presents the performance results of ensembles and single SVM techniques over the WPBC dataset. We observe that four SVM ensembles (ELP, ENP, ERP, ENRP and ELPR) and the single SVM (SVM-PUK) seem to perform better than all other techniques; they achieved an accuracy, recall and precision of 90.88%, 90.9% and 90.9% respectively. The two SVM ensembles (ELNP and ELNPR) achieved the second best performances with an accuracy, recall and precision of 90.54%, 90.5% and 90.6% of accuracy, recall and precision respectively.

Table 6: Performance results: WPBC dataset.

Tech.		Accuracy	Recall	Prec
SVM single	SVM-LK	74.66	74.8	74.7
	SVM-NP	77.03	77	77.2
	SVM-RBF	65.88	65.9	66.5
	SVM-PUK	**90.88**	**90.9**	**90.9**
SVM ensemble	ELN	77.36	77.4	77.4
	ELR	74.66	74.7	74.8
	ELP	**90.88**	**90.9**	**90.9**
	ENR	76.69	76.7	76.8
	ENP	**90.88**	**90.9**	**90.9**
	ERP	**90.88**	**90.9**	**90.9**
	ELNR	77.70	77.7	77.7
	ENPR	**90.88**	**90.9**	**90.9**
	ELNP	90.54	90.5	90.6
	ELPR	**90.88**	**90.9**	**90.9**
	ELNPR	90.54	90.5	90.6

Figure 5 reports the results of the SK test on ensemble and single SVM techniques over WPBC dataset. SK identified three clusters, the best one contains seven SVM ensembles (ELNP, ELNPR, ELP, ELPR, ENP, ENPR and ERP) and one single SVM (SVM-RBF), while the worst cluster only contains one single SVM (SVM-RBF).

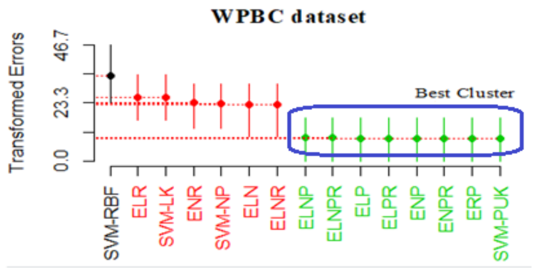

Figure 5: SK test of ensemble and single SVM techniques over WPBC dataset.

4.5 Comparing Ensemble and Single SVM Techniques

In order to investigate the effect of the four kernel techniques L, N, P and R on the performance of ensembles and single SVM techniques, we counted the number of occurrences of each kernel technique in the best SK cluster of each dataset. From Table 7, we note that the P kernel was ranked first in all datasets. However, all the kernels have the same number of occurrences in Wisconsin, and the kernel L was also ranked first in WDBC. Furthermore, the kernels N and R have the same number of occurrences in all datasets.

We can conclude that:

(1) The use of the P kernel instead of L, N and P to build SVM ensembles and single SVM often led to more accurate diagnosis; and

(2) The L, N and R kernels seem have the same impact on the performances of ensemble and single SVM techniques.

Table 7: Number of occurrences of each kernel technique in the best cluster of SK test for each dataset.

Dataset	Kernel technique			
	L	N	P	R
WDBC	8	6	8	6
BCD	3	4	7	4
Wisconsin	8	8	8	8
WPBC	4	4	8	4
Total	23	22	31	22

In order to identify which techniques are the best on all datasets, we ranked the techniques of the best SK cluster of each dataset by using the Borda Count voting system based on Accuracy, Recall and Precision. Table 8 presents the ranking results for each dataset. We observe that:

a. By comparing the number of occurrences of ensembles and single techniques in the best cluster of each dataset, we found that ensembles are the most frequent in the best clusters of all datasets: for WDBC, nine ensembles (ELNR, ELR, ELN, ELP, ELNP, ELPR, ELNPR, ENP, ERP and ENPR) were identified in the best cluster versus two single techniques (SVM-LK and SVM-PUK); for BCD, five ensembles (ENPR, ELNPR, ELPR, ELNP, ENP and ERP) were identified in the best cluster versus one single technique (SVM-PUK); for Wisconsin, eleven ensembles (ERP, ELNP, ELPR, ELNPR , ELR, ENP, ENPR, ELN, ELP, ELNR and ENR) were identified versus four single techniques (SVM-NP, SVM-PUK, SVM-LK and SVM-RBF); and for WPBC, seven ensembles (ERP, ENP, ELP, ENPR, ELPR, ELNP and ELNPR) were identified versus one single technique (SVM-PUK).

b. SVM ensembles outperformed single SVM classifiers in all datasets since the first ranked techniques were ensembles in all datasets: (e.g. ELNR, ELR and ELNR in WDBC).

c. ERP/ENPR ensemble outperformed all the other SVM single and ensemble classifiers in two datasets: Wisconsin and WPBC/BCD.

d. The SVM ensemble ELNPR with four members (i.e. four kernels) was present in the best clusters of all datasets. Moreover, among the 6/4 ensembles with two/three members, 5/4, 2/3, 6/4 and 3/3 were presents in the best clusters of WDBC, BCD, Wisconsin and WPBC respectively.

e. SVM-PUK single technique was present is all the best clusters and was ranked first in WPBC and second in Wisconsin.

To summarize the main findings, we can conclude that:

(1) Ensembles are more accurate than the single classifiers; this confirms the finding of the systematic literature review of (Idri et al., 2016b);

(2) There is no kernels combination (i.e. no SVM ensembles) that outperformed all the others in all datasets. However, the combination P and R seems to perform better.

(3) The use of P kernel instead of L, N and P seems to improve the accuracy of SVM ensembles.

(4) It seems that the performance of SVM ensembles increases with the number of members (i.e. number of kernels).

(5) The best single SVM was SVM-PUK and can be used to overcome the intensive calculation of ensembles.

Table 8: Number of occurrences of each kernel technique in the best cluster of SK test for each dataset.

Rank	WDBC	BCD	Wisconsin	WPBC
1	ELNR[a]	ENPR	ERP	ERP[a]
2	ELR[a]	ELNPR	ELNP	ENP[a]
3	ELN[a]	ELPR	ELPR[a]	ELP[a]
4	SVM-LK	ELNP	ELNPR[a]	ENPR[a]
5	ELPb	ENP	SVM-NP[a]	ELPR[a]
6	ELNP[b]	ERP	SVM-PUK[a]	SVM-PUK[a]
7	ELPR[b]	SVM-PUK	SVM-LK[b]	ELNP[b]
8	ELNPR[b]		ELR[b]	ELNPR[b]
9	ENP[c]		ENP[c]	
10	ERP[c]		ENPR[c]	
11	SVM-PUK[c]		ELN[d]	
12	ENPR[c]		ELP[d]	
13			ELNR[d]	
14			SVM-RBF	

[a,b,c] and [d] mean the same ranks

5 CONCLUSION AND FUTURE WORK

This paper proposed and evaluated SVM ensembles using four kernels (LK, NP, RBF, and PUK) and MLP as a combiner rule over four historical datasets. The SVM ensembles and single SVMs methods were also compared using three criteria: accuracy, precision and recall criteria. The SK test and Borda Count were used to carry out the significance tests and rank the best classifiers respectively. The findings of this study are as follows:

RQ1: SVM ensembles outperformed single SVMs. This means that the use of a combination of kernels with different kernels often lead to better classifiers. Moreover, it seems that the performance of SVM ensembles increases with the number of kernels (i.e. members). Given that ensembles are in general time consuming, the single SVM-PUK can be used.

RQ2: We found that no SVM ensembles (i.e. no combination of kernels) outperformed all the others. However, the use of P and R kernels seems to increase the performance of SVM ensembles. Ongoing work focuses on assessing and comparing homogeneous and heterogeneous ensembles in BC diagnosis. Moreover, the impact of parameters tuning using optimization techniques such as particle swarm and genetic algorithms on the performance of ensemble BC classifiers will be also assessed.

REFERENCES

Bhavsar, H., & Amit Ganatra. (2016). Radial Basis Polynomial Kernel (RBPK): A Generalized Kernel for Support Vector Machine. *International Journal of Computer Science and Information Security*, *14*(April).

Breiman, L. (1996). Bagging Predictors. *Machine Learning*, *24*(2), 123–140.

Canuto, A. M. P., Abreu, M. C. C., de Melo Oliveira, L., Xavier, J. C., & Santos, A. de M. (2007). Investigating the influence of the choice of the ensemble members in accuracy and diversity of selection-based and fusion-based methods for ensembles. *Pattern Recognition Letters*, *28*(4), 472–486.

Chawla, N. V, Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. In *Journal of Artificial Intelligence Research* (Vol. 16).

Chlioui, I., Idri, A., & Abnane, I. (2020). Data preprocessing in knowledge discovery in breast cancer: systematic mapping study. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, *00*(00), 1–15. https://doi.org/10.1080/21681163.2020.1730974

Dietterich, T. G. (2000). Ensemble methods in machine learning. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1857 LNCS*, 1–15.

El Idrissi, T., & Idri, A. (2020). Deep Learning for Blood Glucose Prediction: CNN vs LSTM. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12250 LNCS*, 379–393. https://doi.org/10.1007/978-3-030-58802-1_28

Eltalhi, S., & Kutrani, H. (2019). Breast Cancer Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review. *IOSR Journal of Dental and Medical Sciences (IOSR-JDMS) e-ISSN*, *18*, 85–94.

Hosni, M., Abnane, I., Idri, A., Carrillo de Gea, J. M., & Fernández-Alemán, J. L. (2019). Reviewing ensemble classification methods in breast cancer. *Computer Methods and Programs in Biomedicine*, *177*, 89–112.

Hosni, M., García-Mateos, G., Carrillo-de-Gea, J. M., Idri, A., & Fernández-Alemán, J. L. (2020). A mapping study of ensemble classification methods in lung cancer decision support systems. In *Medical and Biological Engineering and Computing* (Vol. 58, Issue 10, pp. 2177–2193). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s11517-020-02223-8

Hussain, M., Wajid, S. K., Elzaart, A., & Berbar, M. (2011). A comparison of SVM kernel functions for breast cancer detection. *International Conference on Computer Graphics, Imaging and Visualization*, 145–150.

Idri, A., Bouchra, E. O., Hosni, M., & Abnane, I. (2020). Assessing the impact of parameters tuning in ensemble based breast Cancer classification. *Health and Technology*, *10*(5), 1239–1255. https://doi.org/10.1007/s12553-020-00453-2

Idri, A., Hosni, M., & Abnane, I. (2019). *Impact of Parameter Tuning on Machine Learning Based Breast Cancer Classification* (pp. 115–125). Springer, Cham. https://doi.org/10.1007/978-3-030-16187-3_12

Idri, A., Hosni, M., & Abran, A. (2016). Systematic literature review of ensemble effort estimation. *Journal of Systems and Software*, *118*, 151–175.

Idri, Chlioui, I., & EL Ouassif, B. (2018). A systematic map of data analytics in breast cancer. *Proceedings of the Australasian Computer Science Week Multiconference on - ACSW '18*, 1–10.

Kadi, I., Idri, A., & Fernandez-Aleman, J. L. (2017). Knowledge discovery in cardiology: A systematic literature review. *International Journal of Medical Informatics*, *97*, 12–32. https://doi.org/10.1016/j.ijmedinf.2016.09.005

Kudo, T., & Matsumoto, Y. (2000). Japanese dependency structure analysis based on support vector machines. *Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics -*, *13*, 18–25.

Kuncheva, L. I. (2014). *Combining Pattern Classifiers*. John Wiley & Sons, Inc.

Luxburg, U. von, & Schölkopf, B. (2011). Statistical Learning Theory: Models, Concepts, and Results. In *Handbook of the History of Logic* (Vol. 10).

Oskouei, R. J., Kor, N. M., & Maleki, S. A. (2017). Data mining and medical world: breast cancers' diagnosis, treatment, prognosis and challenges. *American Journal of Cancer Research*, *7*(3), 610–627.

Rana, S. P., Dey, M., Tiberi, G., Sani, L., Vispa, A., Raspa, G., Duranti, M., Ghavami, M., & Dudley, S. (2019). Machine Learning Approaches for Automated Lesion Detection in Microwave Breast Imaging Clinical Data. *Scientific Reports*, *9*(1), 1–12.

Santana, L. E. A., Canuto, A. M. P., & Xavier, J. C. (2008). Using feature distribution methods in ensemble systems combined by fusion and selection-based methods. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5163 LNCS*(PART 1), 245–254.

Schapire, & E., R. (1999). A brief introduction to boosting. In *Proceedings of the 16th international joint conference on Artificial intelligence - Volume 2* (pp. 1401–1406).

Schölkopf, B., & Alexander, J. S. (2001). Support Vector Machines, Regularization, Optimization, and Beyond. In *Learning with Kernels* (pp. 1–27).

Sun, Y. S., Zhao, Z., Yang, Z. N., Xu, F., Lu, H. J., Zhu, Z. Y., Shi, W., Jiang, J., Yao, P. P., & Zhu, H. P. (2017). Risk factors and preventions of breast cancer. *International Journal of Biological Sciences*, *13*(11), 1387–1397.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. In *Nature Medicine* (Vol. 25, Issue 1, pp. 44–56).

Trivedi, S., & Dey, S. (2013). Effect of Various Kernels and Feature Selection Methods on SVM Performance for Detecting Email Spams. *International Journal of Computer Applications*, *66*(21), 18–23.

Tsymbal, A., Pechenizkiy, M., & Cunningham, P. (2005). Diversity in search strategies for ensemble feature selection. *Information Fusion*, *6*(1), 83–98.

Vapnik, V. (1992). *Principles of risk minimization for learning theory, in Advances in neural information processing systems,* .

Yu, H., & Kim, S. (2012). SVM tutorial-classification, regression and ranking. In *Handbook of Natural Computing* (Vols. 1–4, pp. 479–506).

Zhou, Z.-H. (Computer scientist). (2012). *Ensemble methods : foundations and algorithms*. Taylor & Francis.