Middle Black Sea Journal of Health Science /
Mid Blac Sea J Health Sci
April 2021; 7(1):104-114

RESEARCH ARTICLE
DOI: 10.19127/mbsjohs.889492

Prediction of Renal Cell Carcinoma Based on Ensemble Learning Methods

Adem Doganer[1(ID)], Cemil Colak[2(ID)], Faruk Kucukdurmaz[3(ID)], Caner Olmez[4(ID)]

[1]Kahramanmaras Sutcu Imam University, Department of Biostatistics and Medical Informatics, Faculty of Medicine, Kahramanmaras, Turkey
[2]Inonu University, Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey
[3]Sanko University, Department of Urology, Faculty Of Medicine, Gaziantep, Turkey
[4]Diyarbakır Gazi Yaşargil Educational and Research Hospital, Department of Urology, University of Health Science, Diyarbakır, Turkey

Received: 02 March 2021, Accepted: 31 March 2021, Published online: 30 April 2021
© Ordu University Institute of Health Sciences, Turkey, 2021

Abstract

Objective: In recent years, ensemble learning methods have gained widespread use for early diagnosis of cancer diseases. In this study, it is aimed to establish a high-performance ensemble learning model for early diagnosis and classification of renal cell carcinomas.

Methods: In the study, hemogram and laboratory data of 140 patients with renal cell carcinoma and 140 patients without renal cell carcinoma were included in the study. The data set includes 27 predictors and 1 dependent variable. The data were obtained retrospectively. In the study, classification performances of machine learning methods and ensemble learning methods were compared. In the study, classification performances of boosting, bagging, voting and stacking ensemble learning methods as well as IB1, IBk, Kstar, LWL, REPTree, Random Forest and SMO classifiers were compared.

Results: REPTree classifier provided the highest performance among machine learning methods (Accuracy = 0.867). Among the ensemble learning methods, the Stacking ensemble learning method provided the highest performance in Model 6 (Accuracy = 0.906). Stacking ensemble learning methods performed higher than boosting, voting, bagging ensemble methods and machine learning methods.

Conclusion: Stacking ensemble learning methods provide successful results in the early diagnosis of renal cell carcinomas. Stacking ensemble learning methods can be used as an alternative to existing methods for diagnosing renal cell carcinoma. In order to further increase the classification performance of the stacking ensemble learning method, it is recommended to choose a meta classifier suitable for the data set and variable types.

Key words: Ensemble Learning Methods, Meta Classifier, Renal Cell Carcinoma, Early diagnosis

Suggested Citation: Doganer A, Colak C, Kucukdurmaz F, Olmez C. Prediction of Renal Cell Carcinoma Based on Ensemble Learning Methods. Mid Blac Sea Journal of Health Sci, 2021; 7(1):104-114

Address for correspondence/reprints:

Adem Doganer

Telephone number: +90 507 377 23 80

E-mail: adem_doganer@hotmail.com

Introduction

Renal cell carcinoma is a type of malignancy tumor that originates from the renal cortex and grows starting from the inner surface of the renal tubules (1). Renal cell carcinoma accounts for approximately 2-3% of all cancer types. In recent years, there has been an increase of 2% in the incidence of renal cell carcinoma disease. Renal cell carcinomas constitute approximately 85% of kidney malignancies (1). Clear cell renal carcinomas, papillary cell renal carcinomas and chromophobe cell renal carcinomas are the most common types of renal cell carcinoma (2). Mostly, in advanced stages, renal region pain, weight loss, high fever, hematuria and defatigation are the physical findings of renal cell carcinoma (3-6). Physical examination, laboratory tests, ultrasonographic and radiological imaging techniques and renal biopsy are generally used for the diagnosis of renal cell carcinoma. Renal cell carcinomas generally progress asymptomatically even in advanced stages. In the physical examination phase, the possibility of early diagnosis of the disease is low. This situation makes it difficult to diagnose the disease early. Early diagnosis is one of the most important factors for successful treatment in cancer diseases. Important advances have been made for early diagnosis of cancer diseases with technological advances in medical imaging techniques. Another important development for early diagnosis of cancer diseases has been achieved with artificial intelligence and machine learning methods. Machine learning methods have produced solutions for many issues in the field of health in recent years. Machine learning methods are applied in early diagnosis of diseases, risk estimation, genetic sequencing, clinical decision support systems, classification of diseases, and identification of patterns for medical images. Machine learning methods can infer and classify real data by learning the patterns in the training data set and the relationship structures between data. Machine learning methods generally perform the classification of data to diagnose diseases (7-8). Although machine learning methods are generally very good classifiers, they cannot provide the desired classification performance for some data sets. Overfitting and underfitting problems are the main causes of failure in this classification (9-11). There are different reasons due to the data set not being suitable for the model or the poor quality of the data set (12). Different solution suggestions have been developed to overcome these problems. One of these solutions is ensemble learning methods. Ensemble learning methods provide a common classification result from the estimates of each classifier by classifying the data of more than one machine learning algorithm separately, rather than classifying the data set by a machine learning algorithm. Thus, common prediction results obtained from more than one machine learning method offer more accurate, more reliable and higher performance compared to the prediction results of a machine learning method (13). Ensemble learning methods are based on the principle that more than one classifier can perform classification with higher accuracy than a single classifier predicts.

A great deal of research has been carried out to predict renal cell carcinomas using machine learning methods. Liu et al. conducted a study on the data they obtained from gene expressions to predict renal cell carcinomas. They used the K-NN algorithm and genetic algorithms to classify renal cell carcinomas. Algorithms have provided successful results in the classification of renal cell carcinomas. Machine learning methods can be successful in the classification of patterns (14). Won et al. classified renal cell carcinoma and other urological diseases with C4.5 algorithm in their study. In the classification process, the data set was obtained from texture patterns. Machine learning methods have shown that they can be used for the early diagnosis of renal cell carcinomas by using patterns (15). Lee et al. used the K-NN classifier (K-NN), the Support vector Machine (SVM) classifier and the Random Forest classifier (RF) to classify renal cell carcinomas in their studies (16). Fuchs et al. stated that pathology findings could be used to predict the survival time of patients with renal cell carcinoma in their study. The researchers who used the random forest classifier to estimate the survival times determined that the findings were compatible with the findings of the pathologists (17). Lin et al. made use of CT images in their study to classify renal cell carcinomas with gradient boosting method (18). Tabibu et al. carried out the classification of panrenal cell carcinomas using convolutional neural networks. A new SVM model was also proposed in the study (19).

Although the ensemble learning method has emerged recently, many researchers have included ensemble learning methods in their studies. Tan and Gilbert can classify cancer diseases in their study by using ensemble learning methods and gene expression values (20). Luo and Cheng benefited from ensemble learning methods in the diagnosis of breast cancers in their studies. The findings of the model provided higher performance than single classifiers (21). Wang used the Stacking ensemble learning method for the diagnosis of prostate cancer. The ensemble model has produced successful results in the diagnosis of prostate cancer (22). Onan made use of ensemble learning methods for breast cancer classification in his study. They have observed that ensemble models provide better classification performance than single classifiers. In addition to classifying diseases, ensemble learning method can also be applied to predict the success of treatment methods (23). Shayesteh et al. predicted the results of treatment methods applied to rectal cancer patients with the MRI-based ensemble learning model in their study (24).

In this study, it is aimed to create an ensemble learning model that will serve as a basis for artificial intelligence applications for early diagnosis of renal cell carcinoma. It was aimed to predict renal cell carcinoma with high accuracy performance using the ensemble learning model. In addition, it was aimed to determine suitable ensemble models and classifiers to predict the disease with the highest accuracy performance.

Methods

The data of the study consisted of 140 patients diagnosed with renal cell carcinoma and 140 patients diagnosed other than renal cell carcinoma, who applied to the Urology service of Kahramanmaras Sutcu Imam University Health practice and research hospital. Permission was obtained from the Clinical Research Ethics Committee of Kahramanmaras Sutcu Imam University (Ethics committee approval Protocol No: 2018 / 07-21) to conduct the study. The data were obtained retrospectively. Age, gender, hemogram and biochemical laboratory variables of the patients constitute the data set of the study. The variables and their properties in the data set are given in table 1. Power analysis was used to determine the sample size of the study. Considering the values of Mean ± SD: 12.7 ± 7.2 and Mean ± SD: 15.7 ± 7.6 in the reference studies; it was planned to include α: 0.05 first type error level, β: 0.20 type 2 error level, n: 97 for each group with a power of 0.80 test, and at least n: 194 patients in total. The high number of data in data mining, machine learning and multivariate research increases the learning ability of the model. Therefore, a total of n: 280 patient data, n:140 for each group, were included in the study. The data set of the study includes 1 dependent (target) and 27 independent (predictor) variables.

Table 1. Variables in the model

Variables	**Variable Type**	**Definition**	**Role**
Renal Cell Carcinoma (RCC)	Categorical	RCC Present/ RCC Absent	Dependent (Target)
Age	Numerical	Positive Real Number	Independent (Predictor)
Gender	Nominal	Male/Female	Independent (Predictor)
Pathology	Nominal	Positive/Negative	Independent (Predictor)
White Blood Cell (WBC)	Numerical	Positive Real Number	Independent (Predictor)
Hemoglobin (Hb)	Numerical	Positive Real Number	Independent (Predictor)
Hematocrit (Hct)	Numerical	Positive Real Number	Independent (Predictor)
Mean Corpuscular Volume (MCV)	Numerical	Positive Real Number	Independent (Predictor)
Neutrophil	Numerical	Positive Real Number	Independent (Predictor)
Lymphocyte	Numerical	Positive Real Number	Independent (Predictor)
Eosinophil	Numerical	Positive Real Number	Independent (Predictor)
Basophil	Numerical	Positive Real Number	Independent (Predictor)
Red Blood Cell (RBC)	Numerical	Positive Real Number	Independent (Predictor)
Platelet	Numerical	Positive Real Number	Independent (Predictor)
Platelet-Lymphocyte Ratio (PLR)	Numerical	Positive Real Number	Independent (Predictor)
Mean Platelet Volume (MPV)	Numerical	Positive Real Number	Independent (Predictor)
Blood Urea Nitrogen (BUN)	Numerical	Positive Real Number	Independent (Predictor)
Creatinine	Numerical	Positive Real Number	Independent (Predictor)
Sodium	Numerical	Positive Real Number	Independent (Predictor)
Aspartate aminotransferase (AST)	Numerical	Positive Real Number	Independent (Predictor)
Alanine aminotransferase (ALT)	Numerical	Positive Real Number	Independent (Predictor)
Protrombine Time (PT)	Numerical	Positive Real Number	Independent (Predictor)
International Normalized Ratio (INR)	Numerical	Positive Real Number	Independent (Predictor)
Partial Thromboblastin Time (PTT)	Numerical	Positive Real Number	Independent (Predictor)
Potassium	Numerical	Positive Real Number	Independent (Predictor)
Mean Corpuscular Hemoglobin Concentration (MCHC)	Numerical	Positive Real Number	Independent (Predictor)
Neutrophil to Lymphocyte Ratio (NLR)	Numerical	Positive Real Number	Independent (Predictor)
Glucose	Numerical	Positive Real Number	Independent (Predictor)

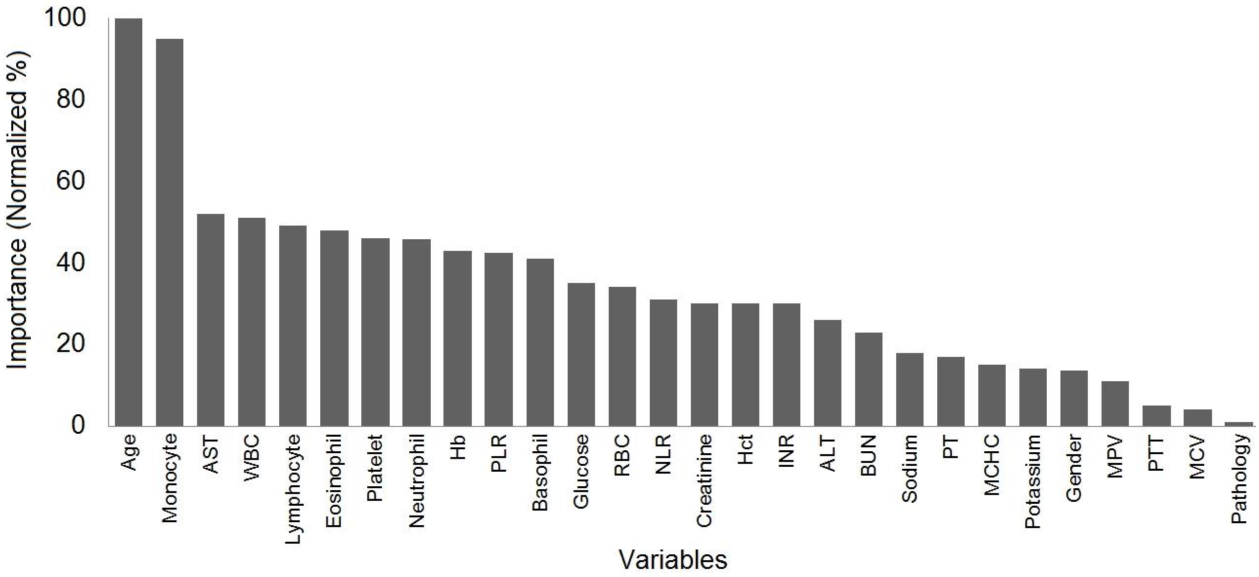

Figure 1. Importance of variables

The Local Outlier Factor (LOF) algorithm has been applied to the data in order to determine the outliers in the data set. LOF algorithm is a powerful method applied in detecting outliers (25-26). As a result of the evaluation, outliers were removed from the data set. Standardization was applied to the quantitative variables in the data set.

Feature selection was carried out to determine the contribution of predictor variables to the classification. Findings regarding feature selection are given in figure 1. According to the findings, the contribution of the "pathology" variable to the classification performance was found to be quite low. The "pathology" variable has been removed from the model.

In the study, in addition to machine learning methods, ensemble learning methods were also applied to predict renal cell carcinoma. IB1, IBk, Kstar, LWL, REPTree, Random Forest and SMO classifiers were applied to classify the data in the model. In the model, besides the individual performances of the classifiers, the performances of voting, boosting, bagging and stacking ensemble learning methods were also evaluated. In the Boosting ensemble learning method, the "Adaboost" algorithm has been applied. In the Stacking ensemble learning method, the "Logistic" classifier is used as a meta classifier. Gridsearch and Multisearch algorithms were used for hyperparameter optimization of classifiers (27-29). Stacking and voting ensemble methods include more than one classifier in the model. For this reason, 7 different models were created to evaluate the stacking and voting ensemble methods. Classifying information for the models is given in Table 2.

Dataset and Model Evaluation

A 10-fold cross validation method was applied to evaluate the performance of classifiers and ensemble learning methods in the model. Weka 3.9 (Waikato Environmental Knowledge Analysis) and R 3.6.0 open source coded softwares were used to evaluate the data with classifiers and ensemble model (30-31). Hyperparameter optimization has been applied to increase the performance of the classifiers. The performances of classifiers and ensemble methods were evaluated with accuracy, sensitivity, specificity, AUC (Area of under curve) and precision metrics. The complexity matrix for the metrics used to evaluate the performance of classifiers and ensemble models is given in table 3.

Table 2. Models and classifiers for voting and stacking enbemble methods

Model	Classifiers
Model 1	IB1
Model 2	IB1, IBk
Model 3	IB1,IBk, Kstar
Model 4	IB1,IBk, Kstar, LWL
Model 5	IB1,IBk, Kstar, LWL, Random Forest
Model 6	IB1,IBk, Kstar, LWL, Random Forest, REPTree
Model 7	IB1,IBk, Kstar, LWL, Random Forest, REPTree, SMO

Table 3. Confusion matrix

		Disease		
		Positive	**Negative**	**Total**
Prediction	**Positive**	True Positive (TP)	False Positive (FP)	TP+FN
	Negative	False Negative (FN)	True Negative (TN)	FP+TN
	Total	TP+FN	FP+TN	TP+FP+FN+TN

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

Ensemble Learning Methods

Machine learning methods are constantly evolving to provide higher accuracy performance in classification and prediction operations. Ensemble learning methods are also one of the methods developed so that machine learning methods can do classification with higher accuracy. Ensemble learning methods are based on the principle that more than one classifier will provide higher classification performance together rather than the classification performance of a single classifier. Common prediction of more than one classifier has higher reliability than the estimation of one classifier (32). The training of the data in the model is performed with more than one machine learning method, and predictions with higher accuracy can be made than only one machine learning method. In ensemble learning models, rather than combining classifiers, a common predictive value is obtained by combining the estimates obtained by the classifiers (33). Different merging and learning techniques can be applied in ensemble learning methods. Choosing the appropriate merging and learning method for the data set and variables provides an increase in performance.

Boosting Ensemble Learning Methods

The Bootstrap Aggregating (Bagging) ensemble learning method is based on the bootstrap sampling method. In the Bagging method developed by Breiman, many different subsets are obtained from the data set with the bootstrap sampling method Subsets obtained from the data set are trained with classifiers (32). All classifiers train different subsets at the same time. Estimates of a large number of classifiers are combined by majority vote. Estimation of the majority among the classifiers is accepted as the estimate of the bagging ensemble learning method (33).

Bagging Ensemble Learning Methods

The Boosting ensemble learning method is based on providing high performance by combining many weak classifiers instead of using only one classifier. In the Boosting ensemble method, the training of data is carried out by iterative operations (35). At every stage of iterative operations, taking into account the mistakes made in the previous stage, the same mistakes are prevented. At the end of the iterative stages, a powerful classifier providing high performance is obtained (34). Adaboost method is a powerful ensemble method suggested by Freud and Shapire (36). In the Adaboost method, weighting is applied at each iterative stage. With the weighting process, the accuracy performance is increased by reducing the error at each stage (33).

Voting Ensemble Learning Methods

Voting is essentially an aggregation technique of classifiers rather than an ensemble learning model. In the voting technique, estimation of majority among different types of classifiers is accepted as the ensemble estimate. The classification of a single data set by different types of classifiers provides different results in estimates. Different estimates increase the accuracy performance of the results. Increasing the variety of classifiers used for training the data set contributes to the reduction of classification errors (33-35).

Stacking Ensemble Learning Methods

The stacking ensemble learning method was developed by Wolpert (37). Stacking ensemble learning method, similar to voting ensemble method, is based on training the data set by different types of classifiers. The main difference in the stacking ensemble learning method is that it has a meta classifier. Estimates obtained from different types of classifiers constitute the input data for the meta classifier. The estimate obtained by the meta classifier from the input data is accepted as the estimate of the stacking ensemble learning method (37).

Results

In the study, 140 patients with Renal Cell Carcinoma and 140 patients with different kidney diseases other than Renal Cell Carcinoma were classified using machine learning methods and ensemble learning methods. For classification, variables related to the descriptive and laboratory findings of the patients were included in the model. In the study, classification performances of boosting ensemble learning method, bagging ensemble learning method, voting ensemble method and stacking ensemble learning methods as well as classifiers applied without using ensemble method were evaluated. The performances of the methods were compared in terms of accuracy, sensitivity, specificity, precision, and AUC metrics. Renal cell carcinoma classification performances of machine learning methods and ensemble learning methods are shown in table 4.

According to the findings in the table, the Stacking ensemble learning method provided the highest value in terms of accuracy, sensitivity, specificity, precision and AUC metrics (Model 6 Accuracy = 0.906; Sensitivity = 0.906; Specificity = 0.906; Precision = 0.910 and AUC = 0.944 respectively). The Stacking ensemble learning method provided the highest performance in the classification of renal cell carcinomas. It has been observed that Boosting and Bagging ensemble learning methods do not contribute to the performance of some classifiers. The stacking ensemble learning method provided the highest performance in Model 6. Model 6 consists of IB1, IBk, Kstar, LWL, REPTree, Random Forest classifier. The "Logistic" classifier, determined as a meta classifier for the Stacking ensemble learning method, contributed to the classification performance of the ensemble model. The performances of the ensemble methods are shown in figure 2.

Discussion

The incidence of renal cell carcinomas has increased significantly in recent years. Early diagnosis of renal cell carcinomas increases the success rate of the disease in treatment. Medical imaging techniques, laboratory tests, biopsy and physiological examination are important for diagnosing the disease. Recently, artificial intelligence technologies have been applied in addition to traditional methods for early diagnosis of cancer diseases. Machine learning methods are important artificial intelligence techniques applied for early diagnosis and classification of diseases. There have been many studies using machine learning methods for early diagnosis of cancer diseases (38-41).

Machine learning methods provide successful results for early diagnosis and classification of cancer diseases with high prevalence. Skin cancer, lung cancer, breast cancer, prostate cancer, stomach cancer, cervical cancer and colorectal cancer are some of the cancer diseases with the highest prevalence in the population (42). Although many studies have been conducted on cancer diseases with a high prevalence, machine learning methods have been applied less for cancer types with low prevalence. It is easier to reach a high number of data in cancer diseases with a high prevalence. The number of data is low in cancer diseases with low prevalence. In order for machine learning methods to be applied successfully, the number of samples in the data set should be high. In cases where the number of samples in the data set is low, the machine learning model cannot learn the pattern between data and variables well. This situation causes underfitting problem. The model shows low classification performance when the number of samples in the data set is small.

Tablo 4. Comparison Performance of Ensemble Learning Methods

		Performance Metrics				
		Accuracy	**Sensitivity**	**Spesificity**	**Precision**	**AUC**
No Ensemble Learning Methods	IB1	0.847	0.851	0.847	0.892	0.849
	IBk	0.847	0.851	0.847	0.892	0.849
	KStar	0.808	0.814	0.808	0.862	0.932
	LWL	0.682	0.684	0.684	0.671	0.726
	REPTree	0.867	0.867	0.867	0.861	0.932
	Random Forest	0.706	0.706	0.706	0.710	0.729
	SMO	0.753	0.753	0.753	0.753	0.752
Boosting Ensemble Learrning Methods	IB1	0.824	0.830	0.824	0.879	0.855
	IBk	0.847	0.851	0.847	0.891	0.849
	KStar	0.796	0.803	0.796	0.852	0.831
	LWL	0.784	0.784	0.784	0.791	0.847
	REPTree	0.863	0.864	0.863	0.845	0.938
	Random Forest	0.788	0.788	0.788	0.801	0.878
	SMO	0.741	0.741	0.741	0.737	0.782
Bagging Ensemble Learning Methods	IB1	0.843	0.847	0.843	0.884	0.894
	IBk	0.847	0.849	0.847	0.879	0.893
	KStar	0.800	0.807	0.800	0.853	0.853
	LWL	0.717	0.720	0.718	0.702	0.786
	REPTree	0.847	0.851	0.847	0.817	0.919
	Random Forest	0.757	0.762	0.757	0.732	0.832
	SMO	0.737	0.737	0.737	0.739	0.809
Voting Ensemble Learning Methods	Model 1	0.847	0.851	0.847	0.891	0.849
	Model 2	0.847	0.851	0.847	0.891	0.849
	Model 3	0.847	0.851	0.847	0.891	0.942
	Model 4	0.847	0.851	0.847	0.891	0.922
	Model 5	0.854	0.861	0.855	0.907	0.934
	Model 6	0.870	0.874	0.871	0.909	0.928
	Model 7	0.878	0.881	0.878	0.911	0.926
Stacking Ensemble Learning Methods	Model 1	0.847	0.851	0.847	0.891	0.835
	Model 2	0.847	0.851	0.847	0.891	0.835
	Model 3	0.858	0.860	0.859	0.881	0.909
	Model 4	0.863	0.864	0.863	0.883	0.905
	Model 5	0.902	0.902	0.902	0.909	**0.944**
	Model 6	**0.906**	**0.906**	**0.906**	**0.910**	**0.944**
	Model 7	0.902	0.902	0.902	0.909	0.942

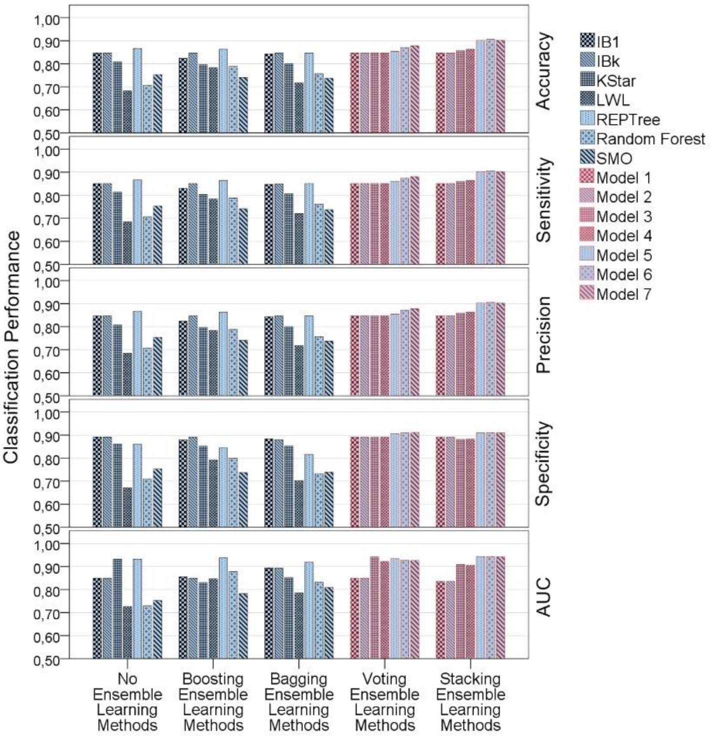

Figure 2. Comparison performance of ensemble learning methods.

In ensemble learning methods, more than one classifier functions in the model. The overall classification error of the ensemble model is reduced by combining the estimates of each classifier. Thus, higher performance and lower errors can be obtained in the ensemble learning method compared to only one classifier (13). Ensemble learning methods can provide higher performance than machine learning methods in the classification and early diagnosis of cancer diseases with low prevalence.

In this study, the classification performance of ensemble learning methods and machine learning methods were evaluated for the prediction of renal cell carcinoma. Classifiers IB1, IBk, KStar, LWL, Random Forest, REPTree and SMO were included in the model. In the model, besides the performance of each classifier, their performances in bagging, boosting, voting and stacking ensemble learning methods are also evaluated. According to the findings of the study, REPTree classifier achieved the highest performance among the classifiers which was not subjected to the ensemble method. The accuracy value of REPTree classifier was obtained as = 0.867 and AUC value = 0.932. In the study, the stacking ensemble learning method achieved the highest performance among the ensemble learning methods. The highest performance in the stacking ensemble learning method was observed in Model 6 (IB1, IBk, KStar, LWL, REPTree and Random Forest). The accuracy value of the Stacking ensemble learning method was obtained as = 0.906 and AUC value = 0.944. The classification performance of the classifiers which was not subjected to the ensemble learning method in the study was found to be compatible with the literature. Sing et al. compared the performance of Naive Bayes, SVM, K-NN algorithm and Random Forest classifiers to predict the progression of late stage papillary renal cell carcinoma. Random Forest classifier provided the highest performance. The accuracy value of the Random Forest classifier was determined as 0.885 (43). Kocak et al. used artificial neural networks to classify renal cell carcinomas in their studies. The accuracy value of the artificial neural network classifier was obtained as 0.692 (44). Jagga and Gupta used Random Forest, SVM, J48 and K-NN classifiers to classify renal cell carcinomas in their studies. Random Forest provided the highest classification performance. The accuracy value of the Random Forest classifier was obtained as 0.77 (45). Bektas et al. classified renal cell carcinomas according to tomography images in their study. They used random forest, K-NN, artificial neural networks and Naive Bayes classifiers in their studies (46). Lin et al. worked on decision tree-based classifiers to predict renal cell carcinomas with tomography images in their studies. They compared the performances of machine learning methods in the model. The highest performance value was achieved as AUC 0.87 (18). The random forest algorithm generally exhibits high performance in the classification of renal cell carcinomas. In this study, REPTree classifier achieved the highest performance among machine learning methods where ensemble learning methods were not used.

In the study, the performances of ensemble learning methods as well as machine learning methods were evaluated to predict renal cell carcinomas. The performance of the boosting, bagging, voting and stacking ensemble methods are evaluated in the model. The REPTree classifier provided the highest performance in the Boosting ensemble learning method. In the Boosting ensemble learning method, the accuracy value of the REPTree classifier was obtained as 0.863. The IBk and REPTree classifiers provided the highest performance in the bagging ensemble learning method. In the bagging ensemble learning method, the accuracy value of both classifiers was obtained as 0.847. Model 7 provided the highest performance in the Voting ensemble learning method. The accuracy value of Model 7 was obtained as 0.878. Model 6 provided the highest performance in the stacking ensemble learning method. The accuracy value of Model 6 was obtained as 0.906. According to these values, Stacking ensemble learning method provided the highest performance among machine learning and ensemble learning methods. Ensemble learning methods generally show higher classification performance than machine learning methods. Classification performances in our study are consistent with the literature. Mohebian et al. worked on the ensemble learning method for predicting breast cancer in their studies. Artificial neural networks, SVM and decision tree classifiers are included in the model. As a result of the classification of the model, the accuracy value was obtained as 0.892 (47). Hsieh et al. used ensemble learning methods to predict breast cancer in their studies. The classification performance of the model was obtained as 0.679 (48). Cai et al. used the ensemble learning method for the classification of lung cancer in their studies. Classification accuracy value of the model was obtained as 0.865 (49). Farahani, Ahmadi and Zarandi in their study, applied ensemble learning methods to detect lung nodules on tomography images. Support vector machine, K-NN and artificial neural networks are applied in the model (50).

In our study, ensemble learning methods showed a higher classification performance compared to machine learning methods in cancer types with low prevalence such as renal cell carcinoma. It has been observed that the ensemble model significantly contributes to the reduction of classification error. In order to increase the classification performance in ensemble learning method, it is very important to select the appropriate ensemble model, merging model and appropriate classifiers for the data set. Choosing the right meta classifier in the stacking ensemble method improves the classification performance. Ensemble learning methods can be successfully applied for the classification and early diagnosis of renal cell carcinomas. There are very few studies in the literature regarding the early diagnosis of renal cell carcinomas by ensemble learning methods. Our study will make a significant contribution to the literature in this field.

Conclusions

In this study, the classification performances of ensemble learning methods and machine learning methods were evaluated for the classification of renal cell carcinomas. High accuracy performance ensemble learning model has been established for early diagnosis of renal cell carcinoma. IB1, IBk, Kstar, LWL, REPTree, Random Forest and SMO classifiers were applied in the study. These classifiers are also included in ensemble learning methods. In the study, boosting, bagging, voting and stacking ensemble learning methods were applied. Stacking ensemble learning method provided the highest classification performance. Model 6 (IB1, IBk, Kstar, LWL, REPTree, Random Forest) provided the highest performance in the stacking ensemble learning method. The accuracy value of Model 6 was 0.906, the sensitivity value was 0.906, the specificity value was 0.906, the precision value was 0.910, and the AUC value was 0.944.

In our study, renal cell carcinomas were successfully classified with high performance using ensemble learning methods. The model created using the ensemble learning method can be successfully applied for early diagnosis of renal cell carcinomas. In our study, stacking ensemble learning methods provided higher classification performance than machine learning methods. Stacking ensemble learning method can diagnose renal cell carcinomas early with high accuracy performance using hemogram and laboratory findings.

Including more variables in the model in ensemble learning methods can increase the classification performance of the model. Increasing the number of samples in the dataset may contribute to the increase in classification performance. It is recommended to choose the appropriate meta classifier for the dataset in the stacking ensemble learning method. In order to increase the classification performance in ensemble learning methods, it is recommended to include different types of classifiers (decision tree-based, function-based and distance-based) into the model.

Ethics Committee Approval: Ethics committee approval was received for this study from Clinical Research Ethics Committee of Kahramanmaras Sutcu Imam University (2018/07-21).

Note: This study was derived from the doctoral thesis prepared by Adem DOĞANER.

Peer-review: Externally peer-reviewed.

Author Contributions:

Concept: A.D.-C.C. *Design:* A.D.-C.C Data Collection and/or Processing: F.K-C.O *Analysis and/or Interpretation:* A.D-C.C. *Writing:* A.D-C.C.-F.K.-C.O.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study hasn't received no financial support.

References

1. Gucer H. The relationship between cox-2 expression, microvessel density and various clinicopatologic parameters in clear cell type renal cell carcinoma. Thesis of Specialization in Medicine. Istanbul: Taksim Educational and Research. 2006.
2. Demirkıran ED. The relationship between tumor volume kidney volume ratio and prognostic factors in renal cell carcinoma. Thesis of Specialization in Medicine. Zonguldak: Bulent Ecevit University, Faculty of Medicine. 2019.
3. Capitanio U, Bensalah K, Bex A, Boorjian SA, Bray F, Coleman J, et al. Epidemiology of renal cell carcinoma. Eur Urol 2019 Jan 1;75(1):74-84.
4. Kim HL, Belldegrun AS, Freitas DG, Bui MH, Han KR, Dorey FJ, et al. Paraneoplastic signs and symptoms of renal cell carcinoma: implications for prognosis. J Urol 2003 Nov;170(5):1742-6.
5. Ficarra V, Prayer-Galetti T, Novella G, Bratti E, Maffei N, Dal Bianco M, et al. Incidental detection beyond pathological factors as prognostic predictor of renal cell carcinoma. Eur Urol 2003 Jun 1;43(6):663-9.
6. Tastekin E. The comparison of angiogenetic and prognostic factors in renal cell carcinomas. Thesis

of Specialization in Medicine. Edirne: Trakya University Faculty of Medicine. 2019.

7. Eble JN, Togashi K, Pisani P. Renal Cell Carcinoma. In : Eble JN, Sauter G, Epstein JI, Sesterhenn IA. editors. World Health Organization Classification of Tumours. Pathology and genetics of Tumours of the urinary system and male genital organs IARC Press, 2004: 9-87.
8. Akman M, Genc Y, Ankarali H. Random Forests Methods and an Application in Health Science. Turkiye Klinikleri Journal Biostat 2011;3: 36-48.
9. Pesch B, Haerting J, Ranft U, Klimpel A, Oelschlägel B, Schill W. Occupational risk factors for renal cell carcinoma: agent-specific results from a case-control study in Germany. Int. J. Epidemiol 2000 Dec 1;29(6):1014-24.
10. Dietterich T. Overfitting and undercomputing in machine learning. ACM Comput Surv 1995 Sep 1;27(3):326-7.
11. Jabbar H, Khan DR. Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). Comp Sci, Comm Instrum Devices 2015:163-72.
12. Lee H, Kim J, Kim S. Gaussian-Based SMOTE Algorithm for Solving Skewed Class Distributions. Int. J. Fuzzy Log. Intell 2017 Dec 25;17(4):229-34.
13. Zhang C, Ma Y. editors. Ensemble machine learning: methods and applications. Springer Science, Business Media, 2012.
14. Liu D, Shi T, DiDonato JA, Carpten JD, Zhu J, Duan ZH. Application of genetic algorithm/k-nearest neighbor method to the classification of renal cell carcinoma. In: Proceedings. 2004 IEEE Computational Systems Bioinformatics Conference, IEEE, 2004: 558-559.
15. Won Y, Song HJ, Kang TW, Kim JJ, Han BD, Lee SW. Pattern analysis of serum proteome distinguishes renal cell carcinoma from other urologic diseases and healthy persons. Proteomics 2003 Dec;3(12):2310-6.
16. Lee HS, Hong H, Jung DC, Park S, Kim J. Differentiation of fat-poor angiomyolipoma from clear cell renal cell carcinoma in contrast-enhanced MDCT images using quantitative feature classification. Med Phys 2017 Jul;44(7):3604-14.
17. Fuchs TJ, Wild PJ, Moch H, Buhmann JM. Computational pathology analysis of tissue microarrays predicts survival of renal clear cell carcinoma patients. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Berlin, Heidelberg, 2008: 1-8.
18. Lin F, Cui EM, Lei Y, Luo L ping. CT-based machine learning model to predict the Fuhrman nuclear grade of clear cell renal cell carcinoma. Abdom Radiol. 2019;44(7):2528-34.
19. Tabibu S, Vinod PK, Jawahar CV. Pan-Renal Cell Carcinoma classification and survival prediction from histopathology images using deep learning. Sci Rep 2019 Jul 19;9(1):1-9.
20. Tan AC, Gilbert D. Ensemble machine learning on gene expression data for cancer classification. Appl Bioinformatics 2003, 2: S75–S83.
21. Luo ST, Cheng BW. Diagnosing breast masses in digital mammography using feature selection and ensemble methods. J Med Syst 2012 Apr;36(2):569-77.
22. Wang Y, Wang D, Geng N, Wang Y, Yin Y, Jin Y. Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection. Appl Soft Comput 2019 Apr 1;77:188-204.
23. Onan A. On the performance of ensemble learning for automated diagnosis of breast cancer. In: Artificial Intelligence Perspectives and Applications, In Artificial Intelligence Perspectives and Applications 2015 (pp. 119-129). Springer, Cham.
24. Shayesteh SP, Alikhassi A, Fard Esfahani A, Miraie M, Geramifar P, Bitarafan-rajabi A, et al. Neo-adjuvant chemoradiotherapy response prediction using MRI based ensemble learning method in rectal cancer patients. Phys Medica. 2019 Jun 1;62:111-9.
25. Breunig MM, Kriegel HP, Ng RT, Sander J. LOF: identifying density-based local outliers. ACM Sigmod Rec 2000 May 16:93-104.
26. Lee J, Kang B, Kang SH. Integrating independent component analysis and local outlier factor for plant-wide process monitoring. J Process Control 2011 Aug 1;21(7):1011-21.
27. Weka Sourceforge. Class GridSearch. Accessed: 24 Nov 2019. Available from: https://weka.sourceforge.io/doc.stable/weka/classifiers/meta/GridSearch.html.
28. Multisearch weka package. Accessed: 24 Nov 2019. Available from: https://github.com/fracpete/multisearch-weka-package.

29. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. ACM SIGKDD explorations newsletter. 2009 Nov 16;11(1):10-8.
30. Team RC. R: A language and environment for statistical computing. Accessed: September 2019. Available from: www.R-project.org.
31. Witten IH, Frank E, Hall MA, Pal CJ. Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2016.
32. Breiman L. Random forests. Mach. Learn 2001 Oct;45(1):5-32.
33. Polikar R. Ensemble learning. In: Zhang C, Ma Y editors. Ensemble machine learning. Springer Science, Business Media, 2012: 1-34.
34. Zhou ZH. Ensemble methods: foundations and algorithms. Chapman and Hall/CRC, 2012.
35. Rokach L. Pattern classification using ensemble methods. World Scientific, 2010.
36. Freund Y, Schapire RE. Schapire R: Experiments with a new boosting algorithm. In: Thirteenth International Conference on ML, 1996 Jul 3 (Vol. 96, pp. 148-156).
37. Wolpert DH. Stacked generalization. Neural Netw 1992 Jan 1;5(2):241-59.
38. Ng EK, Fok SC, Peh YC, Ng FC, Sim LS. Computerized detection of breast cancer with artificial intelligence and thermograms. J. Med. Eng. Technol. 2002 Jan 1;26(4):152-7.
39. Zubi ZS, Saad RA. Using some data mining techniques for early diagnosis of lung cancer. In Proceedings of the 10th WSEAS international conference on Artificial intelligence Knowledge Engineering and Data Bases 2011 Feb 20 (pp. 32-37).
40. Giger ML. Computerized analysis of images in the detection and diagnosis of breast cancer. In Seminars in Ultrasound, CT and MRI 2004 Oct 1 (Vol. 25, No. 5, pp. 411-418). WB Saunders.
41. Abdel-Zaher AM, Eldeib AM. Breast cancer classification using deep belief networks. Expert Syst Appl 2016 Mar 15;46:139-44.
42. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: Cancer J. Clin. 2011 Mar;61(2):69-90.
43. Singh NP, Bapi RS, Vinod PK. Machine learning models to predict the progression from early to late stages of papillary renal cell carcinoma. Comput. Biol. Med. 2018 Sep 1;100:92-9.
44. Kocak B, Yardimci AH, Bektas CT, Turkcanoglu MH, Erdim C, Yucetas U, et al. Textural differences between renal cell carcinoma subtypes: Machine learning-based quantitative computed tomography texture analysis with independent external validation. Eur. J. Radiol. 2018 Oct 1;107:149-57.
45. Jagga Z, Gupta D. Classification models for clear cell renal carcinoma stage progression, based on tumor RNAseq expression trained supervised machine learning algorithms. In BMC proceedings 2014 Oct (Vol. 8, No. 6, pp. 1-7). BioMed Central.
46. Bektas CT, Kocak B, Yardimci AH, Turkcanoglu MH, Yucetas U, Koca SB, Erdim C, Kilickesmez O. Clear cell renal cell carcinoma: machine learning-based quantitative computed tomography texture analysis for prediction of Fuhrman nuclear grade. Eur Radiol. 2019 Mar 2;29(3):1153-63.
47. Mohebian MR, Marateb HR, Mansourian M, Mañanas MA, Mokarian F. A hybrid computer-aided-diagnosis system for prediction of breast cancer recurrence (HPBCR) using optimized ensemble learning. Comput Struct Biotechnol J. 2017 Jan 1;15:75-85.
48. Hsieh SL, Hsieh SH, Cheng PH, Chen CH, Hsu KP, Lee IS, et al. Design ensemble machine learning model for breast cancer diagnosis. J Med Syst 2012 Oct;36(5):2841-7.
49. Cai Z, Xu D, Zhang Q, Zhang J, Ngai SM, Shao J. Classification of lung cancer using ensemble-based feature selection and machine learning methods. Mol Biosyst 2015;11(3):791-800.
50. Farahani FV, Ahmadi A, Zarandi MHF. Lung nodule diagnosis from CT images based on ensemble learning. In: 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, 2015 Aug 12 (pp. 1-7). IEEE.