International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-3, February 2019

A Predictive framework for the assessment of Asthma control level

Pooja M R, Pushpalatha M P

Abstract: ***Asthma is a chronic respiratory disease that is reversible in nature and hence identification of the level of control on the disease can be an important intervention to reduce the morbidity and mortality of the disease. We propose a predictive framework that efficiently predicts the asthma control levels in patients by identifying cells and cytokines in bronchoalveolar lavage (BAL) that contribute significantly to the differences in the controls. We apply various regularized regression techniques to infer the best performing technique on the dataset under consideration. Further, a two class classification problem to distinguish controlled and uncontrolled asthma subjects was handled by deploying binary classifiers and the best performing classifier was adopted. The framework involved the application of feature scoring techniques to identify the risk factors. The work is validated on the data containing subjects including healthy, controlled and uncontrolled subjects, acquired from the Department of Asthma, Allergy and Lung Biology, King's College London School of Medicine, U.K. which was available on the Dryad repository.***

Index Terms: ***Regularization, Dryad, Feature scoring, Binary classifiers, Performance metrics***

I. INTRODUCTION

Asthma is a chronic respiratory disorder of the airways which has a high degree of prevalence. Assessment of level of severity in asthma at an earlier stage is at most important as the disease can be treated with appropriate medications. Cells and cytokines in the bronchoalveolar lavage fluids provide an important means to distinguish asthmatics from healthy control subjects. Also the differences in cell counts and measured level of cytokines in BAL fluids are strong sources to distinguish controlled asthma from uncontrolled asthma. The different cell types and the level of individual cytokines play an important role in identifying varying control levels. Cell counts performed and the absolute numbers and percentage of eosinophils, neutrophils, lymphocytes and monocytes/macrophages are quantified in clinical studies involving BAL fluids. In our work, we make an attempt to predict the level of asthma control using the various fluids and the cell types present in the BAL fluid. Further usage of medications including ICS, SABA, LABA and their dosage also contribute to the possible variations in the control levels. Most of the studies that have been carried out on the dataset under consideration aimed at identifying cytokines and chemokines that could be associated with specific endotypes of asthma such as eosinophil rich and neutrophil rich asthma endotypes [4, 8]. Relationship among the various parameters was inferred by performing linear regression analysis. For the comparison of asthmatic and healthy subjects, logistic regression with elastic net regression was used while for the logistic regression comparing controlled to uncontrolled asthma, a least squares regression was used. In our work, we try to predict the three possible outcomes namely healthy, controlled and uncontrolled asthma using various regression techniques that rely on regularization to effectively perform the prediction. The following variants of regression were tried on the data: (i) Lasso regularized regression (ii) Ridge regularized regression (iii) Elastic net regression (iv) Regression without regularization. It was found that the elastic regression outperformed the other techniques. Further, we deploy two class classifiers to classify the asthmatic subjects into two groups namely controlled and uncontrolled asthma. Some of the standard binary classifiers namely logistic regression, Naïve Bayes, Stochastic Gradient descent, Random forest learner and Support Vector Machine Classifiers which were analyzed for their performance in [18] were deployed for the classification purposes in our work. The Stochastic Gradient descent classifier performed considerably well when compared to the other classifiers. The rest of the paper is organized as follows: We present a review of the related work in Section 2, followed by a discussion of the various regression techniques and binary classifiers deployed on the dataset under consideration in Section 3 under the heading "Materials and Methods". Further, Section 4 provides a discussion on the results obtained by applying the prediction and classification techniques under different scenarios and the last section provides concluding remarks on the work.

Manuscript published on 28 February 2019.
***** Correspondence Author (s)
Pooja M R*, Department of Computer Science & Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, INDIA
Research Scholar at Department of Computer Science & Engineering, Sri Jayachamarajenda College of Engineering, Mysuru, Karnataka, INDIA
Pushpalatha M P, Department of Computer Science & Engineering, Sri Jayachamarajenda College of Engineering, Mysuru Karnataka, INDIA

II. REVIEW OF RELATED WORK

Asthma is a complex heterogeneous disease and the severity levels vary with the difference in the intensity of the risk factors, hence it is at most important to study the factors that contribute significantly to the disease outcome. Good risk assessment models and instruments play a vital role in systems involving population based disease management systems [4]. Commonly used methods for prediction of outcomes related to the asthma disease include Multiple Regression analysis methods as well as the application of classifiers which are binary or multi-class in nature [9].

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Asthma severity levels are based on varying factors, common among which are FEV1, the level of cytokines, count of different cell types and treatment levels as per the guidelines set by ERS/ATS.A prediction model that predicts asthma for symptoms involving wheeze and cough was developed which showed significant validations in primary care when used with symptoms such as male sex, age, eczema, and parental respiratory history which act as important predictors. However the model incorporated only binary outcomes for the features. Missing values for the attributes were interpreted as absence of risk factor [11].

Recent studies show that bronchoalveolar lavage (BAL) contain information on dynamic lung response. Four largely separable quantitative phenotypes were identified in one of the study concerning the physiological data from 1,048 subjects enrolled in the US Severe Asthma Research Program (SARP) [13, 14]. Four different statistical / machine learning methods for predicting each intermediate phenotype using a 76 subject subset of BAL-cytokine measurements were deployed. Comparison of these models using area under the ROC curve and the overall accuracy of classification showed that logistic regression and multivariate regression methods produced results that were more accurate and reliable [15,17].

Patient telemonitoring results in an aggregation of huge amounts of information related to patient disease trajectory. However, the potential utilization of this information for early prediction of exacerbations in adult asthma patients has not been systematically evaluated. Predictive modeling includes preparation of stratified training datasets, selection of predictive features, and evaluation of resulting classifiers [8, 9]. A cloud based predictive model for the prediction of asthma readmission was adopted to a categorical task of prognostication for pediatric asthma readmission utilizing a cohort of case patients with asthma readmission and matching control patients. The predictive modeling was considerably successful in the prediction task through a 5-fold cross validation scheme. The system predicted patients at risk for 12-month asthma readmission with an AUC of 0.69 [14].

A new method of shrinkage and selection called elastic net is adopted which produces a sparse model with good prediction accuracy [1]. The empirical results and simulations show the good performance and the superiority of the elastic net over the lasso. An extension of selection methods are regularized methods including common ridge regression and LASSO. Instead of selecting the optimal subset of variables, an outcome is regressed to all the predictors. In order to deal with many predictor variables, these methods reduce the regression coefficient to 0. This reduction is achieved by applying a penalty to the total estimated coefficients [5, 10]. Regularized methods differ in the way this shrinkage is performed: ridge regression results in a relatively equal shrinkage of all regression coefficients. LASSO regression results in a complete shrinkage of a sub-set. This complete shrinkage works effectively as a means of variable selection [1,2].A notable boosting feature is that the technique can be modified to include an integrated mechanism for shrinking coefficient estimates and selection of variable. This regularization process mechanism makes boosting a suitable method for characterized data analysis involving both small and large sample sizes. In [16,17], a detailed analysis was carried out to identify the important cell types and cytokines that contribute to the identification of different asthma control levels, however there was no work identified on the same which could possibly predict the control levels after identifying the factors that correlated to the different disease levels.

III. MATERIAL AND METHODS

A. Dataset description

The dataset contains information of 36 subjects including 11 healthy, 15 controlled and 10 uncontrolled subjects acquired from the study data of Department of Asthma, Allergy and Lung Biology, King's College London School of Medicine, U.K. which was available on the Dryad repository. A panel of 48 cytokines and chemokines in BAL fluids from healthy control subjects and subjects with controlled and uncontrolled asthma was used for the study. The asthma severity was defined based on FEV1 while on treatment, according to international ERS/ATS guidelines. The asthma attribute was coded as 0 for healthy subjects, 1 for controlled subjects and 2 for uncontrolled subjects. IL-2, IL-4, IL-17, FGF, Eotaxin, GM-CSF, MIP-1α were eliminated as there were no recordings for the same. The healthy subjects were characterized as those who had no history of allergic disease, had PC20 of more than 32 mg per milliliter and normal FEV1 whereas asthmatic subjects were included on the basis of history and a demonstrated airflow limitation which is reversible and increased airway responsiveness to meth choline or both. However no history of other respiratory diseases was reported [17].

B. Regression techniques

A model based on the least square regression does not appear to generalize well on unseen data such as the test data when trained on the training data. Regularization is a way of reducing this variance without increasing its bias. The parameter for regularization also called the tunable parameter controls the bias and the variance. When the parameter increases, the values of regression coefficients are reduced and the overall variance is reduced. After a certain level, however, the bias begins to increase and leads to the adaptation of the model. The selection of the tunable parameter thus plays a key role. Regularization is an important solution to improve the accuracy of regression models. To create a less complex (parsimonious) model when you have a large number of features in your dataset, some of the techniques that can be used to overcome the problem of over-fitting and select the features are: L1 Regularization and L2 regularization [6]. A regression model that adopts L1 regularization technique is called Lasso Regression and the one which adopts L2 is called Ridge Regression.

Lasso and Ridge Regularization techniques

Lasso regression is preferred if we want to use some automatic variable selection or if we handle highly correlated predictors where regression coefficients are too high when standard regression is used. Selection and regularization of features / variables happen almost simultaneously. In this case, L1 standardization is used in which the shrinkage amount is used as the tuning parameter input.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-3, February 2019

The LASSO is a regression method involving the penalization of the absolute size of the regression coefficients. Lasso Regression (Least Absolute Shrinkage and Selection Operator) adds absolute magnitude value of the coefficient to the loss function as a penalty term.

Ridge regression adds square magnitude of the coefficient to the loss function as a penalty term. If lambda is zero then you can imagine that we get OLS back [3, 5]. If Lambda is very large, however, it adds too much weight leading to underfitting. Again, if lambda is zero, we will get OLS (Ordinary Least Square), back whereas very high value will make coefficients zero, which means that it will be inadequate. The key difference between these techniques is that Lasso reduces the coefficient of the less important feature to zero, thereby eliminating some features altogether. This works well for the selection of features if we have a large number of features. Traditional methods such as cross-validation, step-by-step regression to handle overfitting and function selection work well with a small set of functions, but these techniques are a great alternative when dealing with a wide range of functions.

Elastic net

As lasso & ridge regression have their own limitations, elastic net is used to overcome the same, treating each as a special case. The elastic net method connects the Lasso and ridge regression methods. It balances having a parsimonic model with borrowing strength from correlated regressors by solving the problem of regression with restrictions on both the sum of the absolute coefficients and the minimum squares.

More specifically, the elastic net coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_m)$ are the solution to the constrained optimization problem which can be interpreted as.

$$\min ||\mathbf{y} - \mathbf{X}\boldsymbol{\beta}||^2 \quad \text{subject to} \quad \sum_{j=1}^{m} |\beta_j| \le t_1, \sum_{j=1}^{m} \beta_j^2 \le t_2$$

The method can be written as the equivalent Lagrangian form

$$\min ||\mathbf{y} - \mathbf{X}\boldsymbol{\beta}||^2 + \lambda_1 \sum_{j=1}^{m} |\beta_j| + \lambda_2 \sum_{j=1}^{m} \beta_j^2$$

If t_1 is set to a very large value or, equivalently, if λ_1 is set to 0, then the elastic net method reduces to ridge regression. If t_2 is set to a very large value or, equivalently, if λ_2 is set to 0, then the elastic net method reduces to LASSO. If t_1 and t_2 are both large or, equivalently, if λ_1 and λ_2 are both set to 0, then the elastic net method reduces to ordinary least squares regression. Because of the limits observed in the regression of OLS (Ordinary Least Square), the need for a new method was felt. The main drawbacks of OLS are poor predictability of new data and complex model interpretation, if the number of features is high. Hence there was a need for a different model which could infer the relations between the features easily. And the easiest way to do that was to reduce the number of features by taking only the important ones into account. A new subset selection method known as LASSO (Least Absolute Shrinkage and Selection Operator) which addresses this was proposed. Lasso works well with the two cases :(i) When the number of samples is more than the number of features.(ii) If the predictors are independent having no correlation between them. Elastic Net is a lasso-like method that simultaneously shrinks continuously and automatically selects the variable subset. The elastic net penalty is a convex combination of the penalty for lasso and ridge. It cannot consider correlations between variables. If there is a exceedingly correlated group of variables, lasso tends to select only one variable from the group instead of everything and pays no attention to which one is selected.

Elastic network overcomes all these constraints and exceeds predictability and precision. Suppose there are n samples, S={(X1,y1),(X2,y2),... (Xn, yn)} where, Xi=(xi1,xi2,xip) is the input instance of ith and comprises p functions, xij is the value of the jth function and yi is the output of I,for any fixed non-negative criteria, the function constituting elastic net would then be used as:

L(λ1,λ2,β)=|y−Xβ|2+λ2|β|2+λ1

C. Binary classifiers

Binary classifiers are deployed in our work to distinguish between controlled and uncontrolled asthma. The asthma feature is encoded as 1 for controlled and 2 for uncontrolled asthma subjects.

Naïve Bayes (NB)

The Naïve Bayes algorithm works by applying the theorem which is based on the principle of remaining naïve indicating that the features are entirely independent having no preconditions or any prior information. Since the problem under study is a binary classification problem, we deployed the NB classifier which adopts the Bernoulli process for its working.

Random Forest (RF)

This classifier technique when used for classification deploys the mechanism of building an assembly of decision trees and results in a class that is the mode of all classes. Most of the B trees yield features that predict the target attribute strongly. After sufficient repetitions, it was perceived that it was better to choose the total number of trees to be 10 and the number of attributes at each split to be 2 which yielded comparatively better results. The maximum size of the subset was set to 5 and the depth to 3.

Logistic Regression (LR)

Logistic regression is adopted for classification problems when the dependent variable is binary in nature. Here we applied the same to perform binary classification to classify subjects as 1 or 2 representing controlled and uncontrolled asthmatic subjects respectively. It makes use of one or more predictor variables which may be continuous or categorical in nature. But one of its limitations is that it can be used when the underlying data is free from outliers.

Support Vector Machine (SVM)

Support Vector Machines are most suitably adopted for problems involving binary classification, wherein the data available at any point of time can fit into exactly one of the two classes available. The classifier attempts to draw a hyperplane that clearly separates the data instances into different classes. Some of the data however are not classified visibly clearly in the existing feature space though they might show a clear classification in the higher dimensional feature space. Radial basis functions are hence adopted for these learners to achieve better results in terms of classification.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Stochastic Gradient descent (SGD)

In order to avoid redundancy when handling larger data sets, Stochastic Gradient Descent uses the concept of batching data over iterations and takes the idea of batching to the extreme level where only one sample per iteration could be selected at any point of time. The selected single sample is drawn randomly thus characterizing the stochastic approach.

However, the samples must be mixed after each iteration in order to achieve better performance. In addition, since the SGD approach tries to optimize a differential function iteratively, smoother convergences can be achieved.

D. Feature scoring

About 12 features constituting 25% of the total features are chosen by feature ranking by applying feature scoring techniques including Gini, Gain ratio, FCBF and ANOVA.A weighted average of all the features chosen by feature scoring technique is used to form the final feature subset. The following feature scoring techniques have been used in our approach. Information gain: the amount of information expected resulting in reduction of entropy

Gain ratio: the proportion of gained information and the information present intrinsically within the attribute

Gini: the index that discriminates the frequency distance values

ANOVA (Analysis of one way variance): the difference between the average values of the feature in different classes

Chi2: the dependence as measured using chi-square statistics between the class and the feature

ReliefF: the capacity of an attribute to distinguish between classes with respect to the data instances that are similar

FCBF (Fast Correlation Based Filter): an entropy-based measurement which identifies redundancy that arises because of pairwise correlations between features

IV. RESULTS AND DISCUSSION

Table 1. Prediction Results using Test on Test data

Method	MSE	RMSE	MAE
Linear Regression -Lasso	0.350	0.592	0.487
Linear Regression-Elastic Net	0.119	0.344	0.263
Linear Regression-Ridge	3.354	1.831	1.185
Linear Regression-No regularization	3.355	1.832	1.185

Table 2. Prediction Results using Random sampling

Method	MSE	RMSE	MAE
Linear Regression -Lasso	0.273	0.523	0.465
Linear Regression-Elastic Net	0.158	0.398	0.358
Linear Regression-Ridge	2.665	1.633	1.444
Linear Regression-No regularization	2.675	1.636	1.447

Table 3. Prediction Results using Leave One Out technique

Method	MSE	RMSE	MAE
Linear Regression -Lasso	0.245	0.495	0.415
Linear Regression-Elastic Net	0.169	0.411	0.294
Linear Regression-Ridge	1.880	1.371	1.131
Linear Regression-No regularization	1.883	1.372	1.132

Table 4. Prediction Results using 10-fold cross validation

Method	MSE	RMSE	MAE
Linear Regression -Lasso	0.261	0.511	0.420
Linear Regression-Elastic Net	0.207	0.455	0.354
Linear Regression- Ridge	1.838	1.356	1.046
Linear Regression-No regularization	1.841	1.357	1.047

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 5. Actual vs. Predicted asthma outcome using Elastic Net and Lasso

	Asthma Control Level	Predicted Continuous Value	After Round Off
Elastic net	0	0.0316756988463000	0
	1	0.945590652270000	1
	1	1.0437287391000	1
	2	1.91104908762000	2
	2	2.03628012332000	2
	0	0.0316756988463000	0

	Asthma Control Level	Predicted Continuous Value	After Round Off
Lasso	0	2.75770962910000e-06	0
	1	1.00000161481000	1
	1	1.00000019040000	1
	2	2.00000073535000	2
	2	1.99999194402000	2
	0	2.75770962910000e-06	0

Table 6. Classification results using random sampling

Method	AUC	CA	F1	Precision	Recall
SGD	1.000	1.000	1.000	1.000	1.000
LR	0.583	0.700	0.571	0.695	0.700
NB	1.000	0.900	0.889	0.920	0.900
Random Forest Learner	0.833	0.900	0.889	0.920	0.900
SVM Learner	0.875	0.900	0.857	0.914	0.900

Table 7. Classification results using test on test data

Method	AUC	CA	F1	Precision	Recall
SGD	1.000	1.000	1.000	1.000	1.000
LR	1.000	0.400	0.571	0.160	0.400
NB	1.000	0.800	0.667	0.850	0.800
Random Forest Learner	1.000	0.800	0.667	0.850	0.800
SVM Learner	1.000	1.000	1.000	1.000	1.000

Table 8. Feature Subset obtained by combined feature scoring

Features	Information gain	Gain Ratio	Gini	ReliefF	FCBF
ICS (doseμg/day)	0.9709505944546686	0.551374456431497	0.48	0.46239999999999976	0.7079617922707337
%predict FEV1	0.7282129458410014	0.3641064729205007	0.36	0.2262222222222222	5.096685823739945e-05
LABA	0.7166422780956524	0.7672270494333152	0.3876923076923078	0.72	0.720990999143192
ICS (use)	0.6099865470109874	0.6099865470109874	0.32000000000000006	0.564	6.61516116790119e-05
Age	0.590468570732828	0.2973940655707992	0.30000000000000004	0.3140645161290323	4.049612470038811e-05
IL-1RA	0.36724889839732033	0.18362444919866017	0.20000000000000007	-0.0344209685642315	2.6278648793953205e-05
IL-8	0.36724889839732033	0.18362444919866017	0.20000000000000007	0.06956730655679626	2.6278648793953164e-05
%neutrophil	0.27224217190281397	0.13711687689239607	0.14666666666666672	0.044809523809523806	1.9932954297931416e-05
IL-16	0.2427376486136672	0.1213688243068336	0.12	0.006743685848520208	1.7889580085630817e-05

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

M-CSF	0.2427376486136672	0.1213688243068336	0.12	-0.012573200099648438	1.7889580085630817e-05
total cells (x106)	0.2427376486136672	0.1213688243068336	0.12	-0.017178571428571435	1.7889580085630817e-05
% mo0/Mac	0.19046857073282808	0.09593096977777188	0.1216666666666667	-0.03182692307692309	1.3173101925205703e-05

Tables 1, 2, 3 and 4 present the results for prediction in terms of metrics namely MAE, RMSE and MSE through the testing schemes namely, test on test data, random sampling, leave one out technique and 10-fold cross validation. Table 5 presents both actual and predicted asthma control levels via Lasso and Elastic net regularization techniques. Tables 6 and 7 present the two-class classification results in terms of metrics AUC, CA, Precision, Recall and F1 score for the various classifiers using random sampling and test on test data, testing schemes respectively. Random sampling is used with 75% of the data in the training set and 25% in the test set. Table 8 depicts the features constituting the feature subset along with their feature scores obtained using different feature scoring techniques namely Information gain, Gain ratio, Gini, ReliefF and FCBF. About 12 features constituting approximately 25% of the total features were selected by combined feature scoring method which performed average weighting of features chosen by the individual feature ranking techniques. The prediction results obtained by applying multivariate regression with various regularization techniques are presented in the tables above. It was seen that the prediction results were very good in the case of Elastic net regularization. It was better with Lasso regularization when compared to Ridge regularization. The prediction results obtained via regression without any regularization yielded poor results. The training data set had 29 instances, while 6 instances were exclusively used for testing. The predicted values obtained through Elastic and Lasso regression are presented. The predicted values when rounded off yield the outcomes as desired. The classification process was performed by considering only two classes namely uncontrolled and controlled asthma .80% of the total data constituting 20 instances were used in the train set, while the test data consisting of 5 instances, covering 3 instances with controlled asthma and 2 instances with uncontrolled asthma formed the test data set. Various binary classifiers including SGD, Logistic Regression, naive Bayes, random forest learner and SVM were deployed to perform the classification and the classification accuracy in terms of precision, recall ,F1 Measure, AUC and CA were used to assess the classifiers' performance. It was observed that the SGD classifier outperformed the other classifiers when tested on the testing set as well as when random sampling was applied with 75% of the data in the training set.

V. CONCLUSION

The predictive framework presented in the paper can be taken as an important intervention in the early assessment of the degree of asthma control which can lead to targeted therapy for the respective subjects. Prediction results are projected by deploying random sampling, leave one out and 10-fold cross validation techniques and experimental results show that the Elastic net regression outperforms the other regression techniques as can be observed from the values obtained for MAE, MSE and RMSE, when predicting the level of asthma control. Stochastic Gradient Descent classifier performs well compared to the other binary classifiers in all the cases, as can be inferred empirically through metrics such as Classification Accuracy, AUC, Precision, Recall, and F1 measure presented here, which characterize the performance efficiency of classifiers when deployed to distinguish between controlled and uncontrolled asthmatic subjects. The technique yielded optimal results even when tested on the test data. The combined feature scoring technique deployed to identify the reduced features that exclusively predict the asthma outcome has proved to be fruitful as the same was used to perform the binary classification with a higher degree of accuracy.

REFERENCES

1. Zou, Hui, and Trevor Hastie. "Regularization and variable selection via the elastic net." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67.2 (2005): 301-320.
2. Yuan, Ming, and Yi Lin. "Model selection and estimation in regression with grouped variables." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68.1 (2006): 49-67.
3. Kim, Seyoung, Kyung-Ah Sohn, and Eric P. Xing. "A multivariate regression approach to association analysis of a quantitative trait network." Bioinformatics 25.12 (2009): i204-i212.
4. Brasier, Allan R., et al. "Predicting intermediate phenotypes in asthma using bronchoalveolar lavage-derived cytokines." Clinical and translational science 3.4 (2010): 147-157.
5. Schmid, Matthias, et al. "Estimation and regularization techniques for regression models with multidimensional prediction functions." Statistics and Computing 20.2 (2010): 139-150.
6. Xu, Zongben, et al. "L 1/2 regularization." Science China Information Sciences 53.6 (2010): 1159-1169.
7. MartíNez-MartíNez, José M., et al. "Regularized extreme learning machine for regression problems." Neurocomputing 74.17 (2011): 3716-3721.
8. Savenije, Olga EM, et al. "Predicting who will have asthma at school age among preschool children." Journal of Allergy and Clinical Immunology 130.2 (2012): 325-331.
9. Chatzimichail, Eleni, et al. "An intelligent system approach for asthma prediction in symptomatic preschool children." Computational and mathematical methods in medicine 2013 (2013).
10. Yu, Qi, et al. "Regularized extreme learning machine for regression with missing data." Neurocomputing 102 (2013): 45-51.
11. Pescatore, Anina M., et al. "A simple asthma prediction tool for preschool children with wheeze or cough." Journal of allergy and clinical immunology 133.1 (2014): 111-118.
12. Mozaffari, Ahmad, and Nasser L. Azad. "Optimally pruned extreme learning machine with ensemble of regularization techniques and negative correlation penalty applied to automotive engine coldstart hydrocarbon emission identification." Neurocomputing 131 (2014): 143-156.
13. Hosoki, Koa, et al. "Analysis of a panel of 48 cytokines in BAL fluids specifically identifies IL-8 levels as the only cytokine that distinguishes controlled asthma from uncontrolled asthma, and correlates inversely with FEV1." PloS one 10.5 (2015): e0126035.
14. Chen, Robert, et al. "Cloud-based predictive modeling system and its application to asthma readmission prediction." AMIA Annual Symposium Proceedings. Vol. 2015. American Medical Informatics Association, 2015.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

15. Goldstein, Benjamin A., Ann Marie Navar, and Rickey E. Carter. "Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges." European heart journal 38.23 (2016): 1805-1814.
16. Finkelstein, Joseph. "Machine learning approaches to personalize early prediction of asthma exacerbations." Annals of the New York Academy of Sciences 1387.1 (2017): 153-165
17. Hosoki, Koa, et al. "Analysis of a panel of 48 cytokines in BAL fluids specifically identifies IL-8 levels as the only cytokine that distinguishes controlled asthma from uncontrolled asthma, and correlates inversely with FEV1." PloS one 10.5 (2015): e0126035.
18. Pooja, M. R., and M. P. Pushpalatha. "An Empirical Analysis of Machine Learning Classifiers for Clinical Decision Making in Asthma." International Conference on Cognitive Computing and Information Processing. Springer, Singapore, 2017.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)