JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 03 Issue 02
2022

Research Article
https://doi.org/10.56979/302/2022/60

A Data Driven Approach for Automated Risk Assessment and Survival Prediction of Coronavirus Using Artificial Neural Networks

Sundas Israr[1], Mehreen Tariq[2], Sajid Iqbal[2*], Qaisar Rasool[2] and Nabeel Asghar[2]

[1]Department of Computer Science, Air University, Multan, Pakistan.
[2]Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan.
[*]Corresponding Author: Sajid Iqbal. Email: sajid.iqbal@bzu.edu.pk

Received: May 22, 2022 **Accepted:** September 20, 2022 **Published:** September 27, 2022

Abstract: Coronavirus is proved to be a severe epidemic diseasethroughout the world. Despite of endeavoring lot of medicalfacilities for mitigating with this pandemic, still the number ofinfected cases increased rapidly, which leads to lack of healthcareresources (i.e. hospitals, doctors and other healthcare amenities). Early stage risk prediction by analyzing several clinical andbehavioral risk factors is considered to be a promising solutionfor prescribing appropriate triage to patients and to reduce themortality rate due to this fetal disease. To cope up with thisproblem, in our study we have proposed a deep learning basedapproach for the early stage prediction of risk of infection andrisk of mortality in individuals possessing certain risk factors. Wehave utilized a publically available covid-19 dataset incorporatingseveral risk factors that may cause this infection. For the selectionof most significant risk factors i.e. with respect to their level ofimportance in risk prediction, we have employed three featuresselection techniques (i.e. f_classif, PCA and Tree). The set ofextracted features are the utilized for the training of proposedANN for the prediction of infection risk and mortality risk due tocovid-19. For the performance analysis of proposed method, fourdifferent evaluation metrics are being employed including: MSE,MAE, ME and EV. The proposed model has achieved a minimum loss (MSE) of 0.00137 for infection risk prediction and MSE of0.000012 for mortality risk prediction.

Keywords: Coronavirus Infection Risk Prediction; CoronavirusMortality Risk Prediction; Covid-19 Risk Factors; Deep Learning; Artificial Intelligence; Artificial Neural Networks.

1. Introduction

Coronavirus (COVID-19) is a major infectious pandemic disease that first emerged in the city of China namely Wuhan and spread up to a serious stage in all over the world [1], [2].The cause of this epidemic disease is a stern virus known as severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) [3]. Sever acute respiratory syndrome (SARS-CoV) and Middle East Respiratory Syndrome (MERS-CoV) are the two major diseases which are originated due to this wide family of viruses. The patients infected from this infection could be classified into three major categories on the basis of their clinical factors i.e. 1) patients having asymptotic corona (i.e.no infectious symptoms are evident), 2) patients enduring mild symptoms and 3) patients suffering from severe infection that may cause malfunctioning of body organs, or may even leads to death [4]. In most patients (i.e. in near about 99% of cases),the symptoms of Covid infection are asymptotic, while severe indications are found in rest of the cases [22]. According to the web statistics, near about 101 million people has been suffered from this infection till 29th of January, 2021, from which 56.1 million people have been recovered, while 2.19 million patients got died. Due to its rapid inflation, the whole world is struggling with this epidemic such as in the form of lack of healthcare facilities for dealing

with millions of infected people (i.e. to provide them with appropriate treatment facilities and in deciding suitable prescription for them) [5],[6]. The scarcity of health amenities and non-accessibility of pertinent triage came to light the critical need of national reinforcement actions in various aspects i.e. to highlight the necessity of appropriate tool for gauging the risk of infection for feasible healthcare services [7], [8].

The proper identification and categorization of patients with respect to the level of infection risk they possess (i.e. patients having low risk of infection and patients those are prone to have critical infection) is crucial to prescribe efficient and appropriate triage to patients. Some of the major clinical factors that are critically associated with the covid-19 infection risk are: age (i.e. patient with older age possess high risk of infection) and presence of a serious disease in patient, hypoxia and liver disorders [9]. In addition to these, there are numerous other clinical and behavioral factors associated with patients that could be utilized for the early stage risk prediction of Covid infection risk and mortality risk due to this infection. Researchers have employed numerous Artificial Intelligence(AI) based approaches for dealing with this pandemic situation in several ways i.e. in automated diagnosis of COVID-19infection using medical images (e.g. Chest X-Rays and CT Scans) [10], forecasting of future cases infected from this infection [11], to automate the task of food and medicines delivery in disinfect areas through drones and robots [12], to automate the task of prescribing appropriate medicines and drugs to infected people [13], in automated diagnosis of infection through patient's cough sound [14] and in automated infection and mortality risk prediction using patient's clinical and behavioral factors [15] etc. Different types of datasets (i.e. medical image-based datasets, textual reports, clinical factors-based datasets or speech based like cough or breathing sounds etc) are being employed for the training and evaluation of these approaches [16]. The intervention of AI greatly assists in automation and performance enhancement of automated risk prediction of coronavirus infection and mortality prediction due to this disease. To perform this task, several AI based approaches (i.e. machine learning and deep learning) are being employed by different researchers, which could assists medical practitioners to a great extent, in prescribing suitable triage for the patients. The traditional machine learning based statistical approaches possess the ability of efficiently mapping input-output relationships in multifaceted problems. Several researchers have employed different machine learning based approaches for the automation of covid-19 risk prediction task. However, only few authors have utilized deep learning approaches to perform the same task (i.e. on the basis of patient's risk factors). Motivating from this fact, in our study we have employed ANN based approach for the automated risk prediction of Covid infection and to evaluate the risk of mortality due to this infection, in a patient possessing certain risk factors. To accomplish this task, we have utilized a publically availablecovid-19 dataset, which incorporates several patient's clinical risk factors. For dataset cleaning, several preprocessing techniques have been employed, while for the selection of most crucial features several statistical features selection techniques have also been employed. Several experiments have been conducted to evaluate the performance of different features selection techniques. Moreover, the impact of different major risk factors (i.e. patient's blood group, gender, alcohol etc.) over infection and mortality risk, have also been gauged on the basis of several experiments. The outline of rest of the paper is organized as follows: section2 incorporates related literature, in section 3 the proposed methodology have been discussed in detail, in section 4 results of conducted experiments have been mentioned, while the last section incorporates conclusion and future work of our study.

2. Literature Review

As we have discussed above, numerous authors have employed different machine learning based approaches in their studies to automate the task of coronavirus risk prediction using patient's clinical factors. A few of these recent studies have been discussed below: A statistical machine learning based approach has been proposed by the author in [17], in which three different statistical models (i.e. two different implementations of XG Boost and logistic regressor) have been utilized to automatically verdict individuals who are at high risk of respiratory infection. The selected list of risk factors for the training of proposed models is chosen on the basis of risk assessment criteria defined by US Centers for Disease Control and Prevention, while the major factors that have been utilized in this study are: lungs disease, age, diabetes and heart disease. The ROC metric has calculated for the performance evaluation of these models, while logistic regression-based approach has achieved ROC of 0.73 and XG Boost has achieved a ROC of0.81.

In another study [18], an AI based approach has been proposed for the automated early stage forecasting of patients who could suffer from severe infection, on the basis of symptoms visible at the early

stage. For the training and evaluation of proposed methodology, a dataset collected from 53 patients incorporating several patients presentation based analyzed factors and hospital findings has been utilized. From the list of available risk factors, the most crucial ones are identified by employing a features selection technique. The final set of setoff selected features are fed to several machine learning-based classifiers including: Logistic regression, KNN, Decision Tree, Random Forest and support vector machine (SVM). The best accuracy of 70%, 70% and 80% has been achieved by decision tree, random forest and SVM respectively. In another recent study [19], automated machine learning (auto ML) has been employed for the development and comparison of different machine learning based models for the early survival chances prediction in patients infected fromCOVID-19. The dataset that has been utilized to perform this task incorporates of 47 biomarkers.

The biomarker that plays the most crucial role in survival prediction in infectious patient has also been investigated and highlighted in this study. To perform this task auto ML has generated twenty different machine learning (ML) based models with different area under the precision-recall curve (AUCPR), while author has selected the model with best AUCPR. The best model has achieved an AUCPR of 0.836, while the final set of most influential biomarkers include: age, blood pressure, pulse rate, respiratory rate, glucose, lactate dehydrogenase, troponin, blood urea nitrogen and d-dimer.

For the survival prediction in patients infected from epidemic coronavirus disease based on certain medical conditions, a ML based approach has been proposed in [20]. The proposed methodology is evaluated over a dataset collected from 10,237patients of Korea, who are tested positive by this virus. The dataset incorporates nine different categories of predictive factors i.e. sex, age, residence, income level, disability, household type, infection routine, symptoms and medical condition. Uni-variable and multi-variable regression has been employed for prioritizing the risk factors on the basis of their level of importance. Moreover, five different machine learning based models (i.e. including LASSO, RBF SVM, linear SVM, KNN and random forest) have been evaluated for the survival prediction of underlying infectious disease, while SVM and LASSO have depicted high sensitivities (i.e. 92.0% and 90.7%respectively). In [21], author has compared five different machine learning based models for the automated mortality risk prediction in a patient suffering from COVID-19.

The five different ML models that have been compared in this study, includes: KNN, SVM, logistic regression, random forest and gradient boost. Four different types of predictors have been utilized for risk assessment, which includes sex, age group, exposure and province. A dataset of 4004 Covid positive patients have been utilized for the evaluation of proposed methodology. From the selected set of ML methods, logistic regression has performed best and achieves and AUC of 0.83.

3. Proposed Methodology

3.1. Dataset

The dataset [23] that has been exploited in our study (i.e. for the training and evaluation of proposed approach) is acquired by data research team of a software company named as “Nexoid”, situated in London, UK. It incorporates a total of 60dependent and independent variables (i.e. related to patient’s geographical location, behavioral, health, medication factors and risk values). However, in our study several attributes (i.e. depicting patient’s geographical location) has been excluded, while a total of 37 factors (i.e. mentioned in Appendix A) are considered as independent variables and two are considered as dependent variables (i.e. opinion infection and opinion mortality).

3.2.Proposed Methodology

In several recent studies, deep learning has depicted promising performance over several types of medical datasets (i.e. in automated diagnosis and risk prediction of various diseases).Inspiring from this fact, in our study we have employed an ANN based approach for the automated risk prediction ofCovid-19 infection and automated mortality risk prediction due to this disease on the basis of certain clinical risk factors.

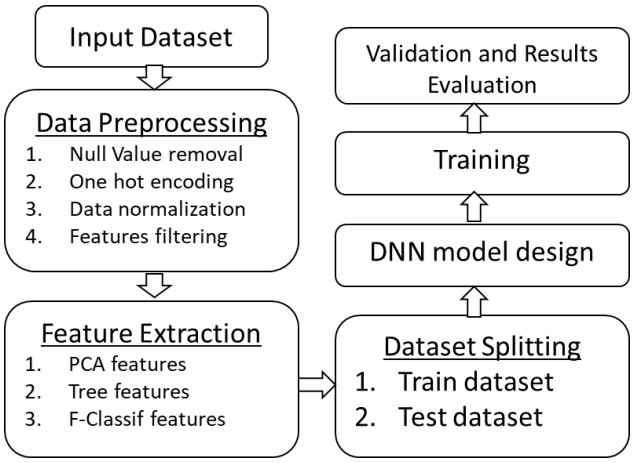

Figure 1. Proposed methodology

3.3. Preprocessing

For the refinement of input dataset, four different preprocessing techniques have been employed over it, which are described in Table 1 below:

Table 1. Preprocessing techniques applied on the dataset

Preprocessing Technique	Description
Null Values Removal	As the dataset that we have used in our study is collected from an online survey, therefore several attributes incorporate null values. To tackle with this issue, in case of numeric features null values are replaced by the mean of that attribute, while in case of categorical attribute null values are replaced by the Mod of that feature.
One-Hot Encoding	For the conversion of categorical data to integer, one hot encoding has been employed.
Data Normalization	For scaling the range of features in range of (0-1), min-max normalization technique has been employed. $$X_{new} = \frac{x - \min(X)}{\max(X) - \min(X)}$$
Features Filtering	From the resultant set of preprocessed data, unneeded features (i.e. geographical location based features) are filtered out, while the resultant data is kept for further processing.

3.4. Features Extraction

A significant phase of machine learning pipeline is features extraction. In most cases, this phase is considered as optional. However, with sparse input data dimension, this phase proves useful to enhance the performance of machine learning model. Basically, in this phase the related feature importance of every input feature with respect to the class to be predicted is calculated. On the basis of calculated features importance, top elements with highest importance or with most impact over the output class are selected and fed to the machine learning or deep learning model for training or evaluation. In our study we have employed three different features selection techniques i.e. classify, tree based and PCA features extraction techniques.

3.5. Model's Architecture

Numerous deep learning-based architectures (i.e. multi-layer perceptron (MLP), convolutional neural networks (CNN), Restricted Boltzman's machine, generative adversarial networks (GAN), long short-term

memory (LSTM) etc) have been utilized by different researchers in literature for the early stage diagnosis and risk prediction of different disease.

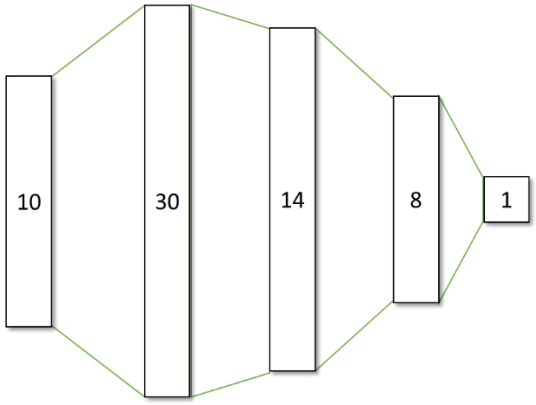

Figure 2. ANN Model used for classification

However, in our study, we have utilized multi-layer perceptron also known as back propagation neural network to perform automated risk prediction task. We have proposed a 5-layered deep neural network (i.e. as depicted in Figure-2) for the automated risk prediction.

3.6. Training Parameters Configuration

The proposed model is trained over 100 epochs with a batch size of 500 samples, while the dataset has been split first before model's training (i.e. 70% for training and 30% for testing). Mean squared error (MSE) is calculated for loss calculation during training. For learning rate optimization ADAM optimizer has been utilized. For the final output layer of proposed model, "linear" activation function has been employed (i.e. to find the risk probability).

4. Experiments and Results

In our study we have performed primarily two types of experiments i.e. 1) for the risk prediction of infection and 2) for the risk prediction of mortality using patient's clinical and behavioral factors. For the performance analysis of proposed approaches, we have employed four evaluation metrics i.e. including 1) mean square error (MSE), 2) mean absolute error (MAE), 3) Max Error (ME) and 4) Explained Variance (EV). The detail of the utilized evaluation metrics and conducted experiments is mentioned below.

4.1. Evaluation Metrics

Table 2. Evaluation metrics used in this study

Metric	Formulae	Description
MSE	$MSE(a,p) = \frac{1}{x}\sum_{i=1}^{x}(a_i - p_i)^2$	Mean square error is used to compute the error or difference between actual value and the predicted value by the model. For x as the total number of samples, a as actual value and p as predicted value the formulae for its calculation is mentioned here.
MAE	$MAE(a,p) = \frac{1}{x}\sum_{i=1}^{x}\lvert a_i - p_i\rvert$	Like MSE, mean absolute error is also employed for the calculation of difference between actual and predicted value, while the formulae for its calculation is also mentioned here.
ME	$ME(a,p) = \max(\lvert a_i - p_i\rvert)$	For the calculation of maximum residual error, max error is employed. This metric is used to find the worst error calculated in between any actual and predicted value. The ideal ME score is 0, however practically this score is impossible to achieve.

EV	$EV(a,p) = 1 - \frac{\text{Var}\{a - p\}}{\text{Var}\{a\}}$	Explained Variance is specifically used for evaluating performance of regression problems. The best EV score that could be achieved by any problem is 1. The formulae for its calculation has been mentioned here, in which a represents actual value, p depicts predicted value and Var is variance.

4.2. Risk Infection Prediction

As already mentioned, that three different features selection techniques have been employed (i.e. f_classif, pca and tree) in our study, for the automatic early stage prediction of covid-19 infection on the basis of several risk factors. We have performed six different types of experiments to perform this task, the detail of which is described below:

4.2.1. f_classif features based Risk Predication

In first experiment, we have extracted a list of 10 features on the basis of their level of importance by employing f_classif features extraction technique, which are further fed to the proposed deep model.

4.2.2. PCA features based Risk Predication

In second experiment, we have extracted again a list of 10 features in accordance to their level of importance by employing PCA features extraction technique, while the final set of extracted features are fed to the proposed deep model.

4.2.3. Tree features based Risk Predication

For the third experiment, we have employed Tree features extraction technique and selected a list of 10 most critical risk factors for automated infection risk prediction, which are subsequently fed to the presented model for the training and evaluation of proposed methodology.

4.2.4. Combined features-based Risk Predication

After gauging the performance of above-mentioned features extraction techniques individually, we have combined the three set of extracted features, while subsequently removed the set of redundant features from them. The final set of 24 concatenated features is then fed to the proposed model for further training and evaluation.

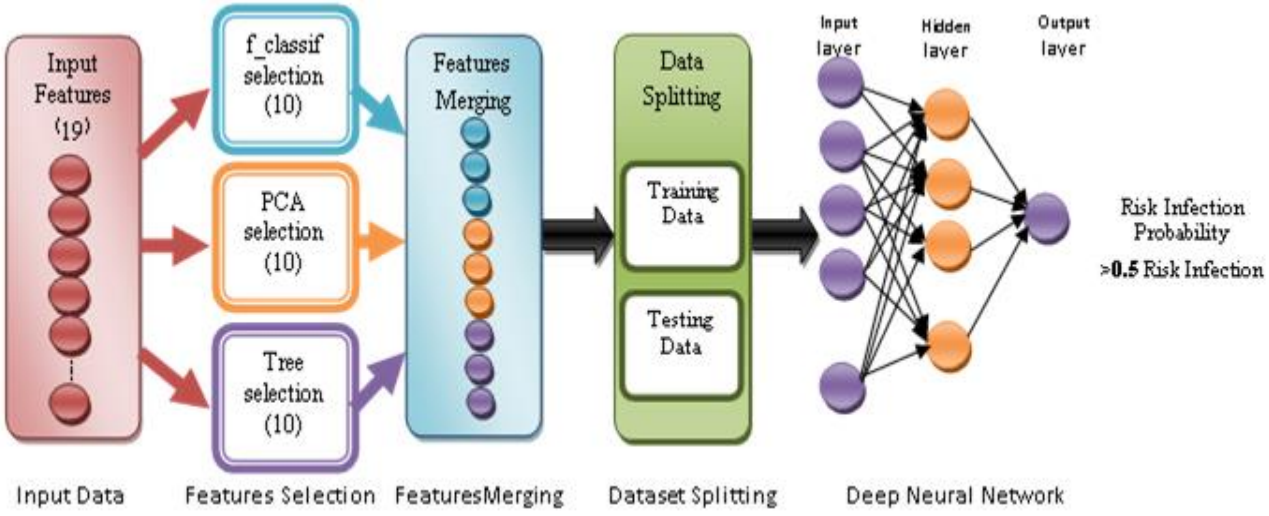

Figure 3. Proposed Architecture with Combined Features Selection Strategy

4.2.5. Ensemble based Risk Predication

In this experiment we have utilized an ensemble-based approach to perform automated risk prediction task. To do so, the proposed model is trained over four different types of features individually (i.e. over f_classif, PCA, tree and combined features individually), while the final prediction is obtained by averaging the predictions of all four models.

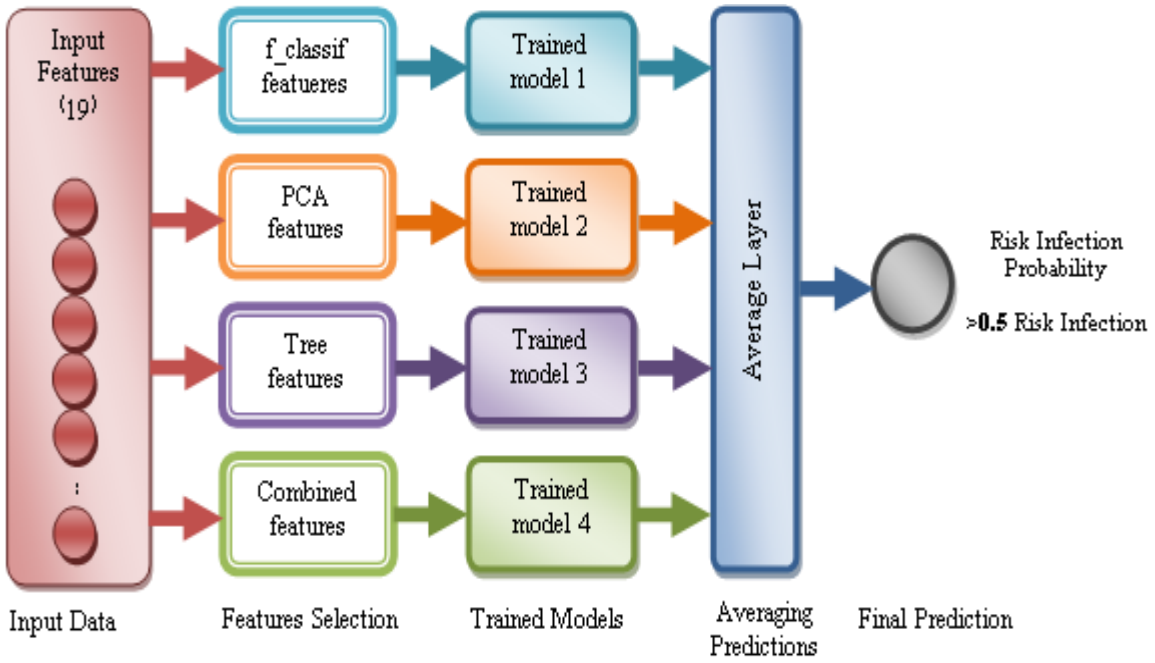

Figure 4. Proposed Ensemble Approach

From above mentioned results it could be concluded, that the combined performance of all three features selection techniques has depicted best performance in terms of both validation loss (i.e. MSE and MAE) and explained variance, while the best (i.e. minimum) ME is achieved by Tree features based regressor. In addition to this it could also be observed from the results that despite of promising performance of ensembles in literature, here in our case the combined features approach has outperformed the ensembles.

Table 3. Performance of proposed model for Infection Risk Prediction

Features Extractor	MAE	MSE	ME	EV
f_classif	0.00685	0.00195	0.94890	0.90361
PCA	0.03355	0.00942	0.94943	0.53468
Tree	0.01366	0.00212	**0.94783**	0.89525
Combined	**0.00416**	**0.00136**	0.94932	**0.93250**
Ensemble(average)	0.01367	0.00208	0.94887	0.85955
Ensemble (weighted average)	0.00590	0.00139	0.94908	0.92313

4.2.6. Gauging the Impact of different Factors over Prediction Results

In addition to the features selected by features selection techniques, we have also utilized some other risk factors that are commonly employed in literature by different researchers. The main motive behind this experiment is to gauge the impact of these selected factors over risk prediction performance. A total of five risk factors have been chosen to perform this task, which includes: 1) sex, 2) age, 3) blood type, 4) alcohol and 5) disease.

The selected risk factors are individually concatenated with the set of features extracted through different features selection approaches and fed to the proposed model for performance analysis.

Table 4. Impact of Gender over Risk of Infection

FeaturesExtractor	MAE	MSE	MaxError	ExplainedVariance
f_classif	0.00937	0.00195	**0.94692**	0.90374
PCA	0.03242	0.00862	0.94937	0.57431
Tree	0.01359	0.00212	0.94763	0.89530
Combined	**0.00472**	**0.00138**	0.94882	**0.93173**

Table 5. Impact of Age over Risk of Infection

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00827	0.00194	**0.94794**	0.90408
PCA	0.04008	0.00958	0.95017	0.52838
Tree	0.01235	0.00210	0.94826	0.89617
Combined	**0.00531**	**0.00137**	0.94886	**0.93229**

Table 6. Impact of Blood Group over Risk of Infection

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00760	0.00195	0.94796	0.90358
PCA	0.03798	0.01036	**0.94238**	0.48879
Tree	0.01263	0.00212	0.94781	0.89521
Combined	**0.00441**	**0.00137**	0.94967	**0.93196**

Table 7: Impact of Alcohol over Risk of Infection

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00895	0.00195	0.94674	0.90377
PCA	0.03064	0.00835	**0.94498**	0.58753
Tree	0.01248	0.00212	0.94694	0.89520
Combined	**0.00490**	**0.00137**	0.95082	**0.93205**

Table 8. Impact of Disease over Risk of Infection

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00696	0.00195	0.94811	0.90365
PCA	0.03356	0.00920	0.95218	0.54528
Tree	0.01186	0.00211	0.94928	0.89543
Combined	**0.00634**	**0.00137**	**0.94716**	**0.93242**

From above mentioned results it could be concluded that, the features or risk factors selected by the features selection techniques have depicted the best results, while the addition of other factors (i.e. age, alcohol, disease etc.) have not improved the results. However, in these experiments also the model trained over combined features has depicted best performance.

4.3. Risk Mortality Prediction

The same set of six experiments has also been conducted for the automated mortality risk prediction due to Covid-19 using same risk factors dataset. All three features selection techniques have been employed in similar manner as in case of infection risk prediction. The results of conducted experiments have been mentioned below.

4.3.1. Results of first five experiments

Table 9. Performance of Proposed Model for Mortality Risk Prediction

	MAE	MSE	Max Error	Explained Variance
f_classif	0.00087	0.00001	0.61977	0.88791
PCA	0.00338	0.00009	0.71641	0.38832
Tree	0.00140	0.00002	0.62771	0.82037
Combined	**0.00063**	**0.00001**	0.62139	**0.92454**
Ensemble (average)	0.00119	0.00001	**0.61084**	0.84373
Ensemble (weighted average)	0.00067	0.00001	0.62086	0.90992

Same as in case of risk infection prediction, in the prediction of risk of mortality, the model trained over all three types of extracted features (i.e. f_classif, PCA and Tree features) has performed best, and has achieved a loss of 0.00001 over validation data.

4.3.2. Impact of different Factors over Mortality Risk Prediction

Table 10. Impact of Gender over Risk of Mortality

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00106	0.00001	0.70866	**0.88345**
PCA	0.00347	0.00008	0.90763	0.42801
Tree	**0.00102**	**0.00001**	**0.62339**	0.87275
Combined	0.00187	0.00002	0.62802	0.85901

Table 11. Impact of Age over Risk of Mortality

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00079	0.00001	0.61880	0.89272
PCA	0.00141	0.00002	0.73756	0.80613
Tree	0.00154	0.00002	0.62458	0.81556
Combined	0.00125	0.00001	0.62556	0.89740

Table 12. Impact of Blood Type over Risk of Mortality

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00085	0.00001	0.63372	0.89502
PCA	0.00321	0.00008	0.77915	0.45250
Tree	0.00128	0.00002	0.62497	0.81621
Combined	0.00096	0.00001	0.62077	0.91292

Table 13. Impact of Alcohol over Risk of Mortality

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00085	0.00001	0.62041	0.89908
PCA	0.00354	0.00008	0.79131	0.44381
Tree	0.00140	0.00002	0.62582	0.82035

Combined	0.00077	0.00001	0.61876	0.91160

Table 14. Impact of Disease over Risk of Mortality

Features Extractor	MAE	MSE	Max Error	Explained Variance
f_classif	0.00086	0.00001	0.74869	0.88946
PCA	0.00311	0.00008	0.79416	0.44021
Tree	0.00176	0.00002	0.62665	0.81430
Combined	0.00094	0.00001	0.62647	0.91706

The same pattern of performance has been observed in case of mortality risk prediction, when impact of different risk factors over model's performance has been gauged i.e. proposed model has depicted best results over the features extracted by the features selector without the addition of any new risk factor.

5. Conclusion and Future Work

In our study, we have designed an automated deep learning-based approach for the early stage risk prediction of risk of infection and risk of mortality due to the recent epidemic disease i.e. known as covid-19. The proposed approach is based upon certain clinical, behavioral and social risk factors found in an individual, which enhance the risk of having the infection or mortality chances due to this disease. A publically available risk factors-based dataset has been employed to perform this task. For the cleaning of data several preprocessing techniques have been employed over it. For the selection of most significant risk factors (i.e. that have the highest role in infection occurrence), three different features selection techniques have been employed. The resultant features extracted by employing these techniques are further fed to the proposed ANN for risk prediction. An ensemble of proposed network has also been tested during experiments. However, the results depicted that the model trained over combined features (i.e. features extracted through all three features selectors) have depicted the best performance.

Funding: No funding was received from any organization for this work.

Acknowledgments: This work is conducted in Center for Artificial Intelligence Research at Bahauddin Zakariya University Multan, Pakistan.

Conflicts of Interest: There is no conflict of interest among authors.

References

1. Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., ... & Chowell, G. (2020). Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. Infectious Disease Modelling, 5, 256-263.
2. Yan, L., Zhang, H. T., Xiao, Y., Wang, M., Sun, C., Liang, J., ... & Xu, H. (2020). Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan. MedRxiv, 27, 2020.
3. Stoecklin, S. B., Rolland, P., Silue, Y., Mailles, A., Campese, C., Simondon, A., ... & Levy-Bruhl, D. (2020). First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020. Eurosurveillance, 25(6), 2000094.
4. Yang, X., Yu, Y., Xu, J., Shu, H., Liu, H., Wu, Y., ... & Shang, Y. (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. The Lancet Respiratory Medicine, 8(5), 475-481.
5. Armocida, B., Formenti, B., Ussai, S., Palestra, F., & Missoni, E. (2020). The Italian health system and the COVID-19 challenge. The Lancet Public Health, 5(5), e253.
6. Ji, Y., Ma, Z., Peppelenbosch, M. P., & Pan, Q. (2020). Potential association between COVID-19 mortality and health-care resource availability. The Lancet Global Health, 8(4), e480.
7. Erika, P., Andrea, V., Cillis, M. G., Ioannilli, E., Iannicelli, T., & Andrea, M. (2020). Triage decision-making at the time of COVID-19 infection: the Piacenza strategy. Internal and Emergency Medicine, 15(5), 879-882.
8. Truog, R. D., Mitchell, C., & Daley, G. Q. (2020). The toughest triage—allocating ventilators in a pandemic. New England Journal of Medicine, 382(21), 1973-1975.
9. Robert, V., Okell Lucy, C., Ilaria, D., Peter, W., Charles, W., Natsuko, I., ... & Ferguson Neil, M. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet infectious diseases, 20(6), 669-677.
10. Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Acharya, U. R. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. Computers in biology and medicine, 121, 103792.
11. Rustam, F., Ahmad, A., Mehmood, A., & Ullah, S. (2020). Byung-won on, Waqar Aslam, and G. Sang Choi," COVID-19 Future Forecasting Using Supervised Machine Learning Models" in IEEE Access, 8, 101489-101499.
12. Chamola, V., Hassija, V., Gupta, V., & Guizani, M. (2020). A Comprehensive Review of the COVID-19 Pandemic and the Role of IoT, Drones, AI, Blockchain, and 5G in Managing its Impact, IEEE Access, vol. 8, no. April, pp. 90225â, 90265.
13. Zhavoronkov, A., Aladinskiy, V., Zhebrak, A., Zagribelnyy, B., Terentiev, V., Bezrukov, D. S., ... & Ivanenkov, Y. (2020). Potential 2019-nCoV 3C-like protease inhibitors designed using generative deep learning approaches.
14. Brown, C., Chauhan, J., Grammenos, A., Han, J., Hasthanasombat, A., Spathis, D., ... & Mascolo, C. (2020). Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. arXiv preprint arXiv:2006.05919.
15. Gao, Y., Cai, G. Y., Fang, W., Li, H. Y., Wang, S. Y., Chen, L., ... & Gao, Q. L. (2020). Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. Nat. Commun. 11, 5033.
16. Junaid, S. (2020). Alanazi Eisa, Alasmary Waleed, Alashaikh Abdulaziz. COVID-19 open source data sets: a comprehensive survey. Applied Intelligence.
17. DeCaprio, D., Gartner, J., Burgess, T., Garcia, K., Kothari, S., Sayed, S., & McCall, C. J. (2003). Building a COVID-19 vulnerability index. arXiv 2020. arXiv preprint arXiv:2003.07347.
18. Jiang, X., Coffee, M., Bari, A., Wang, J., Jiang, X., Huang, J., ... & Huang, Y. (2020). Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. Computers, Materials & Continua, 63(1), 537-551.
19. Ikemura, K., Goldstein, D. Y., Szymanski, J., Bellin, E., Stahl, L., Yagi, Y., ... & Reyes, M. G. (2020). Using automated-machine learning to predict COVID-19 patient survival: identify influential biomarkers. medRxiv.
20. An, C., Lim, H., Kim, D. W., Chang, J. H., Choi, Y. J., & Kim, S. W. (2020). Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. Scientific reports, 10(1), 1-11.
21. Das, A. K., Mishra, S., & Gopalan, S. S. (2020). Predicting CoVID-19 community mortality risk using machine learning and development of an online prognostic tool. PeerJ, 8, e10083.
22. McKeever A. Here's what coronavirus does to the body. National Geographic: https://www.nationalgeographic.com/science/2020/02/here-is-what-coronavirus-doesto-thebody/, 2020
23. COVID-19 risk factor calculator dataset. Nexoid. https://www.covid19survivalcalculator.com/download, 2022