

Indian Health Care System is Ready to Fight Against COVID-19 A Machine Learning Tool for Forecast the Number of Beds

- ✓ Multiple Linear regression model is used for machine learning and is applied on the dataset to predict the number of required beds for the patients. An automated model is designed which can give the predictions of beds on the any entered number of patients.
- ✓ The following Objectives have been covered and implemented-
 - Processing of population of Indian states, number of hospitals, number of beds available in hospitals and number of CoV-19 infected patients.
 - Applied machine Learning algorithm for processing the data and making the predictions of required number of beds for the CoV-19 patients
 - Designed a Generalized graphical model for better predictions of number of beds required on the basis of number of patients.
- ✓ The dataset for the availability of number of beds and hospitals available in India was accessed from Kaggle named “Hospitals and beds in India (state wise)” as depicted . The dataset of the state wise populations and number of CoV-19 infected cases is also compiled from the “Ministry of health and family welfare” portal.
- ✓ The objective of the evaluation metrics is to accurately predict the outcome of number of beds required . The accuracy of the results is very important where correlation is positive, standard error is 0.0724, predicted value is acceptable 0.629 which is greater than 0.5 proves that proposed model is acceptable.
- ✓ A machine learning model is applied on the provided dataset, where multiple linear regression model is applied on the various attributes. State-wise population vs hospitals and beds with respect to number of infected cases of CoV-19 are considered for predicting the futuristic requirements. The prediction is acceptable with 95% accuracy.

A Machine Learning Algorithm Predicts Duration of hospitalization in

COVID-19 patients

- ✓ Main aim to develop a machine learning algorithm to predict prolonged LOS, defined as >8 days, for patients admitted with COVID-19. Eight days was selected as the threshold of prolonged length of stay based on magnitude of deviation from mean and median length of stays.
- ✓ Models were validated on three similar tasks: prediction of LOS with
 - ✓ data from only day 1,
 - ✓ data from the first 2 days of hospitalization, and
 - ✓ data from the first 3 days of hospitalization.
- ✓ Models evaluated include variations on Elastic-net, gradient boosted trees, random forest, support vector machines, logistic regression, a Eureka classifier, generalized additive models, a Vowpal Wabbit classifier, K-nearest neighbors classifiers, re-sidual neural network, a Rulefit classifier, and ensemble models, which were a combination of other model.
- ✓ Models were created to predict short LOS on days 1, 2 and 3 of hospitalization. A total of 42 models were trained on data from 80% of patient population, with all models representing variations on the 12 base models. A total of 966 patients were included in this study: 525 of whom had a LOS of ≤ 8 days, while 441 patients had an LOS of >8 days.

Prediction of Patients' Length of Stay at Hospital During COVID-19 Pandemic

- ✓ Machine learning models are proven to be helpful in improving many aspects of the healthcare system[8].It used multi-layer network(MLP), Naive Bayes, K-NN and decision tree to predict patients' LOS at admission.
- ✓ Their results showed that MLP and Naives Bayes had the highest classification accuracy of around 85%, and K-NN performed poorly on the dataset with only 63.6% accuracy.
- ✓ In our project, we employed K-NN, SVM and Random Forest to classify COVID-19 patients' LOS in hospitals. In the dataset of AV: Healthcare Analytics II[13], the stay length is separated into 11 classes. The first 10 classes correspond to 0-10 days, 11-20 days...91-100 days respectively, and the last class is more than 100 days.
- ✓ KNN (K-nearest neighbor) was used to predict the length of stay in hospital of each person according to their different features and SVM was used to analyze the dataset. The data contains two types of data, numerical and categorical.
- ✓ Random Forest Classifier is an ensemble learning method used to classify samples from different classes
- ✓ the highest classification accuracy of 35.41% given by our models is a bit too low to be useful as a diagnosis tool in practice. Since the classification classes are highly imbalanced, with most of the patients spending less than 40 days in hospital, the classification labels could be reformatted into fewer classes concentrating on patients' stays less than 40 days.

Prediction of Length of Stay in the Emergency Department for COVID-19 Patients: A Machine Learning Approach

- ✓ Four machine learning models, namely logistic regression (LR), gradient boosting (GB), decision tree (DT), and random forest (RF), across different data processing stages to predict COVID-19 patients with an ED LOS of less than or greater than 4 hours.
- ✓ The following objectives have been covered and implemented:
 - Prediction framework for COVID-19 patient ED length of stay. Data extracted from the hospital database is pre-processed and cleaned for analysis. Models are developed to predict COVID-19 patients ED LOS.
 - Decision Tree model to build a tree that identifies all possible attribute combinations from the predictive model, and the proportion of COVID-19 patients within the tree experiencing ED LOS less than or greater than 4 hour was calculated.
 - In the present study, machine learning models were applied to all confirmed COVID-19 patients present in the ED and incorporated attributes related to patient complaints, medical history, and initial ED characteristics
- ✓ The performance of each model was evaluated based on accuracy and the F-1 score. The F-1 score is the harmonic mean of precision (positive predictive value) and recall (sensitivity), which complements the accuracy metric.
- ✓ With further validation, the model and results of this study can serve as an effective decision-support tool to improve healthcare delivery/resource planning and help clinicians develop effective interventions to address patient outcomes (e.g., reducing prolonged LOS).

A Machine Learning Algorithm Predicts Duration of hospitalization in COVID-19 patients

- ✓ The models' predictive capabilities improved sequentially over time, reaching an accuracy of 0.765, with an AUC of 0.819 by day 3. These models, developed using readily available data, may help hospital systems prepare for bed capacity needs, and help clinicians counsel patients on their likelihood of prolonged hospitalization.
- ✓ The following objectives have been covered and implemented:
 - LOS prediction models were created using high-dimensional, patient level EHR data. Models were validated on three similar tasks: prediction of LOS with i) data from only day 1, ii) data from the first 2 days of hospitalization, and iii) data from the first 3 days of hospitalization.
 - Patient day to day location was included to classify patients into intensive care unit (ICU) and non-ICU rooms and also to count the number of days a patient had spent in the ICU.
 - Models were evaluated based on the AUC for predicting short LOS, when applied to a set of holdout data, selected based on random stratification of eligible patients.
- ✓ Machine learning models developed prior to the COVID-19 pandemic have demonstrated the ability to predict prolonged LOS, reaching AUCs as high as 0.84.
- ✓ A benefit of such model development, however, includes the ability to adapt as new factors, including treatment modalities, are captured in the EHR. Finally, the testing of multiple ML models raises potential limitations around model tuning and multiple testing.