

ANALYTICS FOR HOSPITAL'S HEALTH CARE DATA

LITERATURE SURVEY

TEAMID : PNT2022TMID04321

S.NO	PAPER	AUTHOR	YEAR	METHOD & ALGORITHM	RESULT	OBSERVATIONS
1	Indian Health Care System is Ready to Fight Against COVID-19 A Machine Learning Tool forForecast the Number of Beds	Shakti Nagpal; Vijay Anant Athavale ; Ashwini Kumar Saini; Ravindra Sharma	2020	Multiple Linear Regression model	The prediction is acceptable with 95% accuracy.	The proposed model is designed with limited dataset available on different portals but the information is inconsistent. Further this can be considered in future work.
2	A Machine Learning Algorithm Predicts Duration of hospitalization in COVID-19 patients	Joseph Ebinger, Matthew Wells , David Ouyang , Tod Davis , Noy Kaufman , Susan Cheng , Sumeet Chugh	2021	Models evaluated include variations on Elastic-net, gradient boosted trees, randomforest, support vector machines, logistic regression, a Eureqa 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.	<ul style="list-style-type: none">For all 3 prediction tasks, ensemble-based models performed best (ENET Blender for days 1 and 2 models and Advanced AVG Blender for day 3 model).Model performance improved with increasing data, with the models trained on culminative day 3 data demonstrated the highest sensitivity (0.93), accuracy (0.765) and AUC (0.819).	The proposed model demonstrates that these algorithms are accurate and can be developed for novel disease states for which clinical knowledge is yet unavailable, enhancing clinicians' ability to make early determinations.

3	Using Data Analytics to Improve Hospital Quality Performance	Christine Pitocco, Thomas R. Sexton and Kelly Stickle	2020	<ul style="list-style-type: none"> Extracted data for 2,233,214 discharges in 2014 from 183 hospitals in the state. Found that 20.8% of the facilities were on the quality performance frontier—20.6% of the not-for-profit facilities and 21.4% of the other facilities. 	79.2% of hospitals could improve their quality of care	As an upper bound, if all hospitals increased each quality factor performance to 100%, there would have been 11,722 (24.8%) fewer deaths, 17,840 (15.8%) fewer readmissions, and the statewide average length of stay would have been 0.71 days (13.5%) less
4	Prediction of Length of Stay in the Emergency Department for COVID-19 Patients: A Machine Learning Approach	Egbe-Etu Etu , Leslie Monplaisir, Suzan Arslanturk , Sara Masoud, Celestine Aguwa;	2022	Four machine learning models, namely Logistic Regression (LR), Gradient Boosting (GB), Decision Tree (DT), and Random Forest (RF)	The GB model outperformed the baseline classifier (LR) and tree-based classifiers (DT and RF) with an accuracy of 85% and F1-score of 0.88 for predicting ED LOS in the testing data.	With further validation, the model and results of this study can serve as an effective decision-support tool to improve health care delivery/resource planning and help clinicians develop effective interventions to address patient outcomes (e.g., reducing prolonged LOS).
5	Predicting length of stay and mortality among hospitalized patients with type 2 diabetes mellitus and hypertension	Diana Barsasella, Srishti Gupta , Shwetambara Malwade,Aminin, Yanti Susanti , Budi Tirmadi , Agus Mutamakin f, Jitendra Jonnagaddala , Shabir Syed	2021	Linear Regression , Random Forest and Multilayer Perceptron	<ul style="list-style-type: none"> A Random forest best predicted inpatient LoS (R², 0.70; root mean square error [RMSE], 1.96; mean absolute error [MAE], 0.935), and the gradient boosting regression model also performed similarly (R², 0.69; RMSE, 1.96; MAE, 0.935) 	Linear regression model best predicted LoS and mortality was best predicted using MLP. Patients with primary diseases such as T2DM or HTN may have comorbidities that can prolong inpatient LoS.