iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42122

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Heterogeneous Multi-Model Ensemble based Length of Stay Prediction on MIMIC III

D. Geethamani[1], Dr. R. Rangaraj[2]
[1]Assistant Professor, Dr. N.G.P. Arts and Science College, Department of Computer Science
[2]Professor & Head, PG & Research Department of Computer Science, Hindusthan College of Arts and Science, Coimbatore – 641 028.

Abstract: Accurately predicting the length of stay of patients in hospital can have positive impacts on both financial and psychological conditions of patients. This can also help hospital administrations to perform effective and correct allocation of critical resources to patients. This work presents an ensemble modeling technique that can be used for accurate determination of length of stay of patients based on several factors that are obtained from hospital records. The multi model nature of the architecture, and the heterogeneity associated with the production model provides a high performing system that can handle the complex hospital data. Experiments were performed on the standard MIMIC III data. Comparisons were performed with the existing state of the art models from literature. Comparisons indicate that the proposed HMME model demonstrations accuracy levels of 95%, indicating that the model can be effectively deployed in hospitals for decision making purposes.

Keywords: Length of state prediction; ensemble modeling; multi model architecture; MIMIC III; correlation based analysis

I. INTRODUCTION

The current days have witnessed a huge and increased emphasis on high value care in hospitals. Hospitals tend to improve the efficiency, quality off the services provided and also aims to reduce the complications for the patients [1]. Delay in discharges tend to be the most troublesome off issues faced by hospitals [2]. From the patient’s perspective, delay in discharges leads to an automatic increase in cost. Further, the increased Length of Stay (LoS) also leads to a risk of several hospital acquired complications. All these factors collectively lead to decreased patient satisfaction [3]. Precise identification of length of stay can aid in better management of bed and other critical resources, which are part of a hospital system. This can also result in effective control of inpatient flow. Controlled in patient flow can effectively reduce waiting time for patients, leading to better patient satisfaction and also results in high level of cost savings for the hospitals [4]. According to research it was identified that a one hour delay in discharge can lead to a 3% increase in the odds of patient mortality [5]. Inpatient admission levels and determining the length of stay can highly influence bed assignment decisions. These are the major challenges faced by hospital strategic and operational management team. Hospital admissions generally require preplanned bed assignments. This is due to the fact that surgery based admissions are mostly inpatient in nature. Current bed allocations and assignments are based on human judgment. These judgments may not be proved to be highly operative. Smooth allocation of beds with less waiting time for patients can be effectively performed with length of stay based analysis.

Length of stay is an effective indicator of hospital activity and can be a clear and defining proxy that can indicate the resources consumed in hospitals [6]. Accurate length of state prediction can have positive impact on health service delivery [7]. Clinical treatment procedures tend to be clear indicators that can be used to determine the length of stay of patients [8]. The clinical treatment processes presents crucial factors that can aid in evaluating health care services of hospitals. The clinical treatment procedures tend to vary with patients [9]. Hence, the treatment process should be continuously monitored to correctly determine the length of stay. Further, treatment variations should be monitored to identify the change in length of stay.

This work presents an effective patient monitoring system that can aid in accurate determination of length of stay of patients. The proposed model is constructed as a heterogeneous multi model ensemble, HMME, to handle the complex nature of hospital data. Hospital data is composed of large number of features. This issue of huge dimensionality is handled by using correlation based feature selection techniques. The preprocessed data is passed to multiple heterogeneous models for training. Multiple predictions obtained from the models are aggregated to obtain the final prediction.

The remainder of this paper is structured as follows; Section 2 presents the related works, Section 3 presents a detailed view of the proposed HMME architecture, Section 4 provides the results and the discussions, and Section 5 concludes the work.

II. RELATED WORKS

This section will cover some of the recent and most popular literature works that are available for length of stay prediction.

A statistical model based method for predicting length of stay for stroke patients was proposed by Lin et al. [10]. This technique is based on identifying statistical properties between features like age, gender, comorbidity levels, and functional independence measure. In addition to this chase square test and PH regression analysis and log transformed linear regression where applied to obtain the length of stay levels. A length of stay prediction model that predicts prolonged length of stay for patients with community acquired pneumonia was proposed by Uematsu et al. [11]. This work uses logistic regression determine the length of stay. The model also includes patient characteristics as prediction indicators. The model has been validated using bootstrapping techniques. Other similar studies based on identifying Length of stay of patients include a pneumonia based study by Cortoos et al. [12], a Swiss multicenter study by Widmer et al. [13], and a healthcare study in Japan by Hamada et al. [14].

A neural network based model for length of state prediction was proposed by He et al. [15]. This is a multi-task learning model that operates based on two factors, patient flow and length of stay. Further, recovery status, the types of surgery performed on patients and the treatments performed for the patients also plays a vital role in determining the length of stay. This work links all the above mentioned factors and creates an artificial neural network based model that performs multitask learning. Multiple model analysis do identify length of stay in newborns was proposed by Thompson et al. [16]. This is a standard analysis architecture that uses existing models from weka to perform analysis. An ensemble based model do identify length of state prediction in brain tumor patients was proposed by Muhlestein et al. [17]. This is a regression based model that aims to improve the predictive performance of the ensemble. A machine learning model developed for predicting length of stay for ICU patients was proposed by Jauk et al [18]. This work uses several machine learning models like neural networks and logistic regression to perform the classification process. Early length of state prediction model was proposed by Solis et al. [19].

A length of state prediction model that uses temporal similarity measures to determine the number of days of state was predicted by Huang et al. [20]. The model performs feature creation at varied stages of clinical treatment. A Case based reasoning methodology is followed to determine similar medical behaviors. Stage based length of state prediction model was proposed by Yang et al [21]. This technique has been performed on records that deal with patients sustained with burns. Prediction is performed after every clinical stage to determine the current length of stay factor. A network based prediction model for cardiac patients was proposed by Tu et al. [22]. This technique is used to determine the length of stay of patients in ICU. Prediction calculator for length of state prediction after spine surgery what's proposed by Lubelski et al. [23]. This technique is based on using demographic variables, comorbidities, treatment processes, and insurance status. The model performs both univariable and multivariable analysis. A technique using individualized single classification models for length of stay prediction in ICU patients was proposed by Ma et al. [24].

Most of the works discussed in this section are specific and disease based analysis. The proposed model is a generic architecture and can be used on all types of patients using their basic treatment records

III. LOS USING HETEROGENEOUS MULTI MODEL ENSEMBLE (HMME)

Predicting length of stay is one of the most significant aspects of a hospital management system. Effectively predicting the length of stay can ensure that the hospital can perform effective resource allocation. Hospital resources are usually critical components like ICU beds, oxygen beds, surgery rooms etc. Effective allocation of these resources can ensure that patients who require these resources are provided with the resources within significantly lower time buffer. This work presents a length of stay prediction model, Heterogeneous Multi Model Ensemble(HMME),to perform faster and more accurate predictions. The proposed HMME model is composed of two major sections; the first section performs statistical feature reduction by using correlation based techniques, the second section performs length of state prediction using multiple heterogeneous machine learning models.

A. Data Preprocessing

Data preprocessing forms the initial and the most significant process in the length of state prediction process. Medical data is used as the training data for this process. Medical data tends to be huge and voluminous in nature. This work uses the MIMIC III data as the model data for the training process. Data is distributed in several files based on the department from where the data has been collected. Hence, it becomes important to analyze the files and identify the most appropriate and the most significant data features for use in the prediction model. This work selects the features that correspond to admission details, discharge details, treatments and test details corresponding to patients.

Followed by the process of manual selection of features, the data is transformed to ensure that it is suitable for the machine learning process. The initial transformation is performed on the date of birth feature. The data contains a feature that shows the date of birth of a patient. This feature is a datetime object, hence, it cannot be directly used by the machine learning model. This feature is transformed to represent age, and is converted to a numerical feature. Diagnosis has been identified as a feature containing too many distinct string values. This is because of performing several unique treatments on patients, as treatments vary from patient to patient. Further, the large number of distinct entries are also attributed to spelling errors. In order to reduce the number of distinct elements, a grouping process is applied to this feature. The treatments that are performed on less than 10 patients are identified and are grouped together as miscellaneous treatments. Further, the length of stay feature is created as a continuous valued feature. Analysis from literature shows that the length of stay of a patient that is less than 10 days does not exhibit any significant impact on the patient's health, as well as the cost. However, the length of stay of more than 10 days effects both the health of the patient and also makes it significantly difficult to perform prior determination of their discharge date. Hence, the threshold of 10 days is a very important threshold factor, and this work converts the continuous valued data do a binary classification problem, based on this threshold.

Statistical analysis of the data also indicates that, the data is also prone with several null values, and inconsistent values. This issue is tackled in two stages. Features that exhibit more than 90% of null values are eliminated. Similarly, instances that exhibit more than 90% of null values are also eliminated. Other null values are imputed using averaging factor for numerical features, and refilling with the mode value for string features. The base data operated upon is medical data, hence the data contains several string features. These features are converted to numerical instances by using encoding techniques. This work uses one hot encoding as the preferred encoding technique. The one hot encoding technique creates a feature for each existing unique value contained in the selected feature. This process results in creating large number of features in the data.

B. Correlation Based Feature Reduction

The presence of large number of stream features string features, and, the presence of several unique values in the features research in results in creating increasing the attribute size of the data to a large extent. This eventually results in the curse of dimensionality. Hence, to ensure effective predictions, feature reduction is included as a mandatory component in this work. The process of correlation based feature selection is performed to reduce features in the data. Correlation refers to the affinity a feature depicts towards another feature. The values of correlation varies between -1 and 1. This work uses the Pearson's correlation [25] method to identify the correlation levels.

$$c = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Where $x_i$and y_i are the values in sample, and $\bar{x}$ and $\bar{y}$ are the mean values if the x variable and the y variable.

Pair based correlation is identified for all the existing features in the data set. Pairs that exhibit greater than 50% correlation in both magnitudes (positive and negative) are selected as the features to be retained. This process is performed for all the features, and the selected feature set is combined, and the duplicates containing the feature set are eliminated to obtain the final feature set. Identifying values like admission ID, hospital ID, and patient ID are retained.

C. Heterogeneous Multi-Model Ensemble (HMME) Creation

The prepared data is passed to the multi model ensemble architecture for prediction. The data is huge and is complex in nature. This mandates the use of multiple models that can handle the complexity of the data to provide effective results. The advantage of using a multi model approach is that issues contained in certain models can be compensated by the positives exhibited by other models. Hence, predictions from a multi model architecture are considered to exhibit better performance compared to using stand-alone machine learning models. This work creates multiple models using decision tree, Gaussian naïve Bayes, and random forest.

Decision tree is the type of machine learning model that creates a tree like model for decision making. The root of the created tree contains the basic split, representing the feature exhibiting high entropy. This is followed by creating features that exhibit next level in entropy levels. The decision tree ends in leaf nodes that provide the final classification result. Decision trees can also be linearized into decision rules. Hence, decision trees are considered to be highly interpretable models. The major advantage of decision trees is that they can be combined with other decision techniques effectively. However, they are very unstable. Even a small change in the base data can lead to a drastic change in the tree structure.

Naïve Bayes is the type of probabilistic classifier that is based on Bayes' theorem. Naïve Bayes is based on the assumption of attribute independence. The model is highly scalable and is highly interpretable. Gaussian naïve Bayes is used when operating on continuous valued features. The major downside of this model is that it considers all the features to be independent.

Random forest is a tree based ensemble modeling technique. They are created by aggregating multiple decision trees. They are effective in reducing the overfitting issue cost in decision trees. They are generally used as black box models. Hence, they lack interpretability provided by decision trees.

The three models, decision trees, Naïve Bayes and random forest are used as the three base models for creating the classifier. The identifier values contained in the data are temporarily removed to create the training data for the multiple models. The entire data is passed to all the three models for the training process. Analysis of the data shows that the data exhibits an imbalance level of eight. Although the imbalance level is not very high, handling the existing imbalance level is mandatory to create a model that can effectively and accurately predict the length of stay. Since, the imbalance level is low, model based imbalance handling approach has been chosen. This is performed by the Naïve Bayes model. The high level of complexity contained in the data is handled by the decision tree model and the random forest model. Further, the overfitting issue exhibited by the decision tree model is effectively handled by the random forest model. This ensures that prediction from each model exhibits its own significance. Hence, all three models are considered with equal importance.

D. Prediction Averaging

Patient records for prediction are passed to the preprocessing module. This is followed by null value elimination and normalization of features. Features that are filtered using the correlation based feature selection method are retained for the prediction process. This preprocessed data is past all the train models. Every model uses the rules that have been generated to provide the prediction. A single prediction is required for each instance of the test data. However, as every model provides its predictions, multiple predictions are obtained for each instance. This work proposes a prediction averaging technique that can be used to combine the generated predictions into a single prediction. The process of the averaging approach is shown in the equation below.

$$P = \frac{\sum_{i=1}^{n} p_i}{n}$$

Where p_i is the prediction obtained from each of the individual models, and n is the total number of models used for prediction.

The final averaged value, represents if the length of stay of the patient is less than or greater than 10 days. Although, this is a binary prediction, it is performed in the initial stages of the admission. Hence, it can be used as a valid indicator for determining the accurate length of state.

IV. RESULTS AND DISCUSSION

The heterogeneous multi model ensemble has been created using the MIMIC III dataset [26]. The ensemble model has been coded in the Python environment using the scikit package. MIMIC III is a single center database that contains clinical information that has been collected from patients admitted in critical care units in Beth Israel Deaconess Medical Center in Boston. The dataset contains details measured from demographics, billing, clinical measurements, medical history, laboratory tests and pharmacotherapy. The records are for a seven year period ranging from 2008 to 2014. The data contains details about 38,597 distinct patients representing 49,785 admissions. Performance of the length of stream prediction model has been analyzed based on performance metrics like, sensitivity, specificity, precision, accuracy, G-Mean and lift. The performances were calculated by creating the confusion matrix and by using the corresponding values in the below equations.

$$TPR\ (or)\ Sensitivity = \frac{TP}{TP + FN}$$

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

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

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

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

$$G - Mean = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}}$$

$$Lift = \frac{TP}{TP + FP} / \frac{TP + FN}{TN + FP + TP + FN}$$

ROC curve for the proposed HMME model has been created and is shown in Figure 1. The ROC plot shows true positive rate in the Y axis and false positive rate in the X axis. An effective classifier is considered to be the one that exhibits its highest point to the top left corner of the plot. The plot created using TPR and FPR levels from the HMME model exhibits its highest point at less than 0.2 FPR and greater than 0.95 TPR. The point could be observed in the top left coordinate of the plot. This shows that the model is capable of predicting with high true positive rate and low error levels.

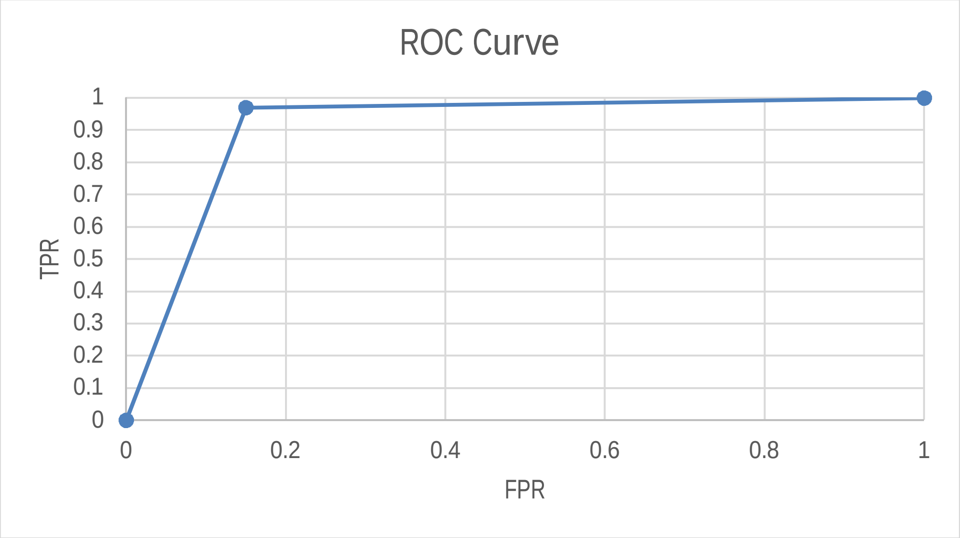

Figure 1: ROC Plot of HMME

The precision recall (PR) curve of the HMME model is shown in figure 2. The precision recall curve is created based on the precision and the recall values which depict the performance of the classifier model in correctly identifying the positive classes. This work considers patients with length of stay greater than 10 days as the positive class. High values for both precision and recall are expected in a good classifier model. An effective indicator, is the alignment of the highest point in a PR plot towards the top right coordinate of the chart. This indicates high precision and high recall. HMME model exhibits its highest point towards the top right. This indicates that the model high efficiency in correctly identifying the positive classes which correspond to patients with a length of state greater than 10 days.

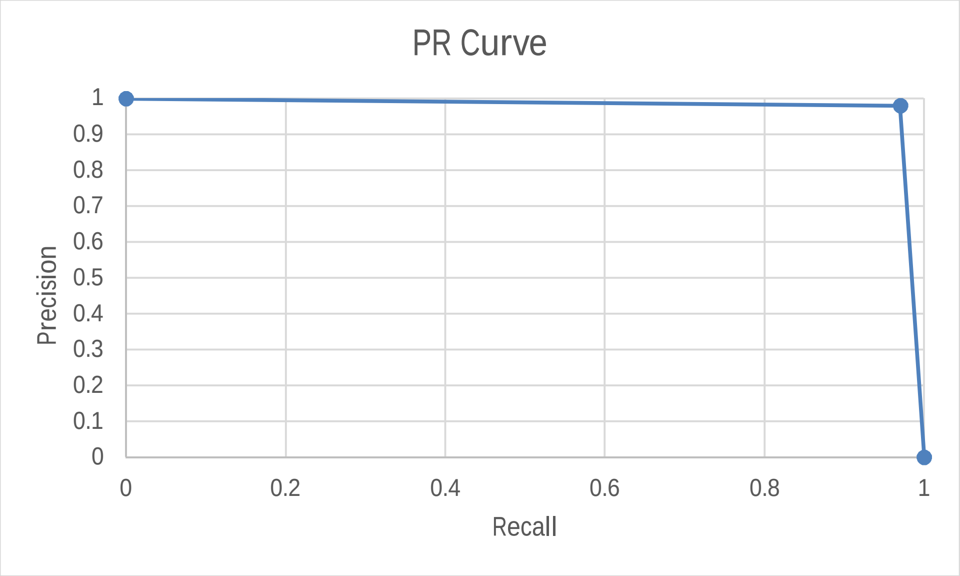

Figure 2: PR Plot of HMME

The classification report of HMME model is shown in table 1. The report shows a view of precision, recall and F1-score. The report also provides a class based performance view. Further, overall metrics like accuracy is also provided in the report. Class based analysis shows that the performance of HMME in identifying patients with less than 10 days of length of stay exhibits an overall F1-Score of 80%. Performance on patients with greater than 10 days, which is a more significant prediction, exhibits an overall F1-Score of 98%.

Table 1: Classification Report of HMME

	Precision	Recall	F1-Score	Support
0	0.76	0.85	0.8	1221
1	0.98	0.97	0.98	11085
Accuracy			0.96	12306
Macro	0.87	0.91	0.89	12306
Avg	0.96	0.96	0.96	12306

A comparison of the aggregatemetrics, accuracy, sensitivity, specificity and precision are shown in figure 3. The comparison is performed with length of stay prediction model by Ma et al [24]. Comparison shows that HMME model exhibits better accuracy and sensitivity levels, while the model proposed by Ma et al. exhibits better specificity and precision levels.

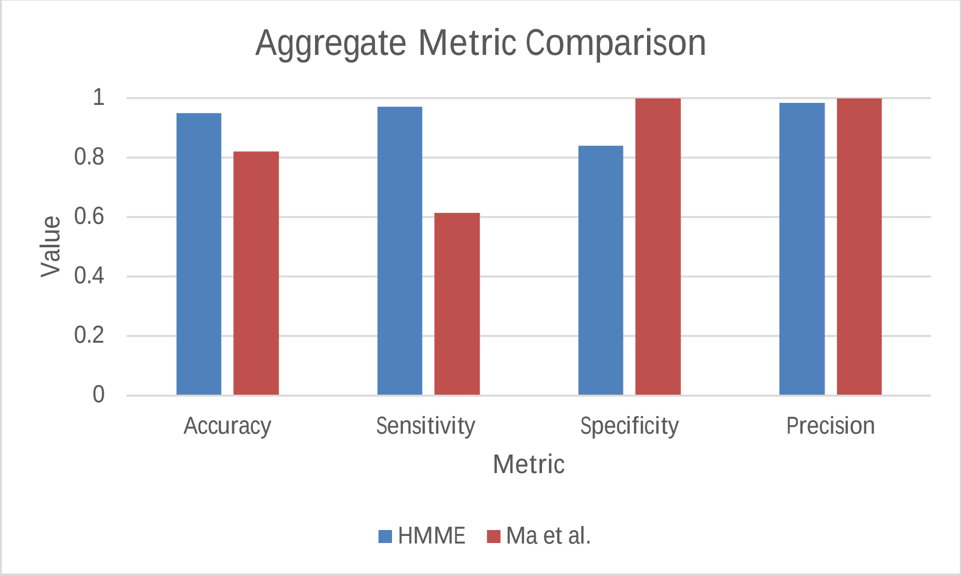

Figure 3: Aggregate Metric Comparison with HMME

An overall comparison of performance is shown in table 2. The best performances and highlighted in bold. It could be observed that the HMME model exhibits 16% reduced specificity levels and 2% reduced precision levels when compared to Ma et al. However, the overall accuracy levels of HMME model is 95%, which is 13% more than the model proposed by Ma et al [24]. Further, the specificity levels were also found to be 97%, which is 36% more improved, and the G-Mean was found to be 90%, which is 12% improved compared to Ma et al. However, the lift levels were found to be slightly low compared to the model proposed by Ma et al. The overall performance indicates that the HMME model outperforms Ma et al. in several metrics, indicating that it is a better performer in predicting the length of state.

Table 2: Overall Performance Comparison with HMME

	HMME	Ma et al.
Accuracy	0.95	0.82
Sensitivity	0.97	0.614
Specificity	0.84	1
Precision	0.983	1
G-Mean	0.9	0.784
Lift	1.09	2.13

V. CONCLUSION

Accurately identifying the length of stay of patients is of paramount importance due to the high cost involved for the patient, and the height resource requirement requests in the hospitals. Correct identification can ensure appropriate allocation of resources to patients who are in need. This work presents a heterogeneous multimodel and ensemble, HMME, which is used for effective discrimination of patients with length of stay of less than 10 days and greater than 10 days in the hospital. Patients with high length of state stay are often prone to several financial issues and also health issues. This model has been designed to determine the length of stay levels during the initial stages of admission. Hence, the model can be used as an effective initial indicator to determine a rough estimate of the patients stay levels. The prediction is performed with an accuracy of 95%. Hence, the hospital and the patients can have a rough estimate of their stay period. Comparison with existing state of the art models shows that the HMME model exhibits good performance which is indicative of accurate length of stay prediction. Limitations of this model are they slightly reduced specificity levels and the precision levels. The model can be further extended into a multiclass prediction model, or a regression model by accurately predicting the number of days of stay.

REFERENCES

[1] Costa AP, Poss JW, Peirce T, Hirdes JP. Acute care inpatients with long-term delayed-discharge: evidence from a Canadian health region. BMC Health Serv Res 2012;12:172https://doi.org/10.1186/ 1472-6963-12-172.
[2] Hwabejire JO, Kaafarani HMA, Imam AM, Solis CV, Verge J, Sullivan NM, et al. Excessively long hospital stays after trauma are not related to the severity of illness: let's aim to the right target!. JAMA Surg 2013;148:956–61 https://doi.org/10.1001/jamasurg.2013.2148.
[3] New PW, Andrianopoulos N, Cameron PA, Olver JH, Stoelwinder JU. Reducing the length of stay for acute hospital patients needing admission into inpatient rehabilitation: a multicentre study of process barriers. Intern Med J 2013;43:1005–11 https://doi.org/10.1111/ imj.12227.
[4] Michael D. Howell, Managing ICU throughput and understanding ICU census, Curr. Opin. Crit. Care 17 (6) (2011) 626–633.
[5] Matthew M Churpek, Blair Wendlandt, Frank J Zadravecz, Richa Adhikari, Christopher Winslow, Dana P Edelson, Association between intensive care unit transfer delay and hospital mortality: a multicenter investigation, J. Hospital Med. 11 (11) (2016) 757–762.
[6] Huang, Z., Lu, X., Duan, H., & Fan, W. (2013). Summarizing clinical pathways from event logs. Journal of Biomedical Informatics, 46(1), 111–127.
[7] Ramos, M. (1999). The successful utilization of financial data in the support of care management. Family and Community Health, 22, 49–63.
[8] Adlassnig, K., Combi, C., Das, A., Keravnou, E., & Pozzi, G. (2006). Temporal representation and reasoning in medicine: Research directions and challenges. Artificial Intelligence in Medicine, 38(2), 101–113.
[9] Doering, L., Esmailian, F., Imperial-Perez, F., & Monsein, S. (2001). Determinants of intensive care unit length of stay after coronary artery bypass graft surgery. Heart Lung, 30(1), 9–17.
[10] Model-based Prediction of Length of Stay for Rehabilitating Stroke Patients
[11] Prediction model for prolonged length of stay in patients with community-acquired pneumonia based on Japanese administrative data
[12] Cortoos P, Gilissen C, Laekeman G, Peetermans WE, Leenaers H, Vandorpe L, et al. Length of stay after reaching clinical stability drives hospital costs associated with adult community-acquired pneumonia. Scand J Infect Dis 2013;45:219e26.
[13] Suter-Widmer I, Christ-Crain M, Zimmerli W, Albrich W, Mueller B, Schuetz P, ProHOSP Study Group. Predictors for length of hospital stay in patients with community-acquired pneumonia: results from a Swiss multicenter study. BMC Pulm Med 2012;12(1):21.
[14] Hamada H, Sekimoto M, Imanaka Y. Effects of the per diem prospective payment system with DRG-like grouping system (DPC/PDPS) on resource usage and healthcare quality in Japan. Health Pol 2012;107:194e201.
[15] Neural network-based multi-task learning for inpatient flow classification and length of stay prediction
[16] Brandon Thompson, Karim Elish, Robert Steele, Machine learning-based prediction of prolonged length of stay in newborns, in: 2018 17th IEEE International Conference on Machine Learning and Applications, ICMLA, IEEE, 2018, pp. 1454–1459.
[17] Whitney E Muhlestein, Dallin S Akagi, Jason M Davies, Lola B Chambless, Predicting inpatient length of stay after brain tumor surgery: Developing machine learning ensembles to improve predictive performance, Neurosurgery 85 (3) (2019) 384–393.
[18] Stefanie Jauk, Diether Kramer, Günther Stark, Karl Hasiba, Werner Leodolter, Stefan Schulz, Johann Kainz, Development of a machine learning model predicting an ICU admission for patients with elective surgery and its prospective validation in clinical practice., Stud. Health Technol. Inf. 264 (2019) 173–177.
[19] Early Prediction of Length of Stay After Pancreaticoduodenectomy

[20] Length of stay prediction for clinical treatment process using temporal similarity
[21] Yang, C., Wei, C., Yuan, C., & Schoung, J. (2010). Predicting the length of hospital stay of burn patients: Comparisons of prediction accuracy among different clinical stages. Decision Support Systems, 50(1), 325–335.
[22] Tu, J., & Guerriere, M. (1993). Use of a neural network as a predictive instrument for length of stay in the intensive care unit following cardiac surgery. Computers and Biomedical Research, 26(3), 220–229
[23] Prediction calculator for nonroutine discharge and length of stay after spine surgery
[24] Length of stay prediction for ICU patients using individualized single classification algorithm
[25] "Correlation Coefficient: Simple Definition, Formula, Easy Steps". Statistics How To.
[26] Data Descriptor: MIMIC-III, a freely accessible critical care database