IM - COMMENTARY

Deep learning algorithms with mixed data for prediction of Length of Stay

Greta Falavigna[1] 🔵

The paper proposed by Bacchi et al. [1] focuses on a very interesting topic for hospital management. Indeed, in a context of scarce financial resources, the estimation and prediction of Length of Stay (LOS) is a relevant factor. Clearly, and above all in the present pandemic situation due to COVID-19, the organization and planning of hospitalization are growing in importance, as the heads of departments try to strike a balance between caring for people's health and keeping the costs of hospitalizations under control [2, 3]. In addition, benefits can also be derived in terms of patients' health, as suggested by the literature (see for example [4–6]). In light of these considerations, research on estimating the LOS is always welcome.

The paper by Bacchi et al. [1] presents a case study of patients admitted to General Medicine at the Royal Adelaide Hospital over a period of 8 months, starting from 01/01/2020. The authors adopt a deep machine learning (or simply deep learning, DL) model and a logistic regression (LR) model to predict discharges in the next two days and seven days and to estimate the correct dates of discharge.

The results of the study suggest that both deep learning and logistic regression are not very effective in predicting discharge within seven days of a given ward round note, but their performance increases when the discharge within two days is examined. The comparison among models analyzes the area under the curve (AUC), accuracy, standard values of correct and incorrect classifications (i.e., true positive, true negative, false positive, false negative), sensitivity and specificity levels. Furthermore, results are also provided considering the cut-off value following Youden's index for discharge in the next two and seven days. For what concerns the area under the curve (AUC), artificial neural networks

(ANNs; i.e., the deep learning model) perform better in predicting discharge in the next two and seven days (0.80 and 0.68) when compared to the logistic regression methodology. As for sensitivity, specificity and discharge in next two days, both ANNs and LR display higher values of true negatives than of true positives, with the exception of the ANN model, which maximizes the sensitivity. On the contrary, the results for discharge in the next seven days are difficult to compare because their sensitivity and specificity values are not significantly different and are both lower than 0.7. However, although both models yield worse results when the LOS is higher than two days, logistic regression is less effective than deep learning, while the correct prediction of the estimated dates of discharge remains an open issue. Similar conclusions have been put forward by [7], which studies the LOS and finds that, in general, artificial neural networks perform better than other models, considering both the AUC and accuracy.

The research by Bacchi et al. [1] brings together two relevant aspects. On the one hand, its authors propose a simple but effective framework, consisting of two networks interconnected by a connection layer before the output one; on the other hand, they work with mixed-input data, in which specific information on patients is scarce, especially if compared to data collected from ward round notes.

The recent literature has applied artificial intelligence (AI) to a large variety of fields with the aim of predicting events, and hybrid models or mixed-input data are used ever more frequently for improving performance. In particular, knowledge of machine learning (ML) and deep learning (DL) allows defining routines and libraries for the majority of software, so that researchers are then able to adapt or combine different frameworks and/or heterogeneous data ([8]). In the medical field, deep learning and artificial neural networks have been widely employed both for predicting/stratifying the risk [2, 3] and for image processing/extraction ([9]). In this field, evidence from the application of convolutional neural networks (CNNs) seems very promising.

✉ Greta Falavigna
  greta.falavigna@ircres.cnr.it

1   Research Institute on Sustainable Economic Growth
    (IRCrES-CNR), National Council of Research of Italy, via
    Real Collegio 30, 10024 Moncalieri, TO, Italy

Indeed, this type of network consists of a multi-layer framework where each layer can be composed of a big number of nodes and is able to identify different features.

Looking at the LOS in the short and long term, researchers have carefully studied its prediction using ANNs and ML. For instance, in [10], an integrated approach featuring fuzzy logic and neural networks outperforms other conventional models, with 79.95% accuracy. Additionally, in [11], the LOS for shoulder arthroplasty is predicted by relying on ANNs, while in [12], good results are achieved in assessing prolonged LOS (i.e., LOS > seven days) for post-acute care in acute coronary syndrome patients undergoing percutaneous coronary intervention.

Another noteworthy point of the paper by Bacchi et al. [1] concerns the application of deep learning for text mining, because these methodologies can be very helpful in improving prediction models. The main idea, as suggested by [13], is to use unsupervised, supervised, self-supervised machine learning and deep learning techniques to be able to capture all available information, text notes included.

Ward round notes and medical records are often written texts and may contain comments that are relevant to predict the LOS or evaluate the health status of a patient. The importance of research in text mining combined with artificial intelligence is increasing because of the very promising results obtained thus far. Text mining would be made easier and more effective when implementing models if hospital documents (i.e., ward round notes and/or medical records) were homogeneous. For this purpose, physicians should agree on a list of standardized terms (i.e., glossary) whose interpretation is unequivocally understood and shared not only by medical practitioners but also across hospital personnel. In this manner, each label in ward round notes and medical records would take on the same meaning, and information would be more suited to being introduced into models as input, thereby reducing their complexity and enhancing their performance.

The design of paper by Bacchi et al. [1] and the results obtained are very stimulating and, together with other recent investigations, suggest that more effort should be devoted to studying the prediction of outcome (i.e., in this specific case, the LOS) through deep learning and hybrid models. Indeed, the model accuracy can impact on the patients' health, the hospital organization (for example, in relation to the LOS) and, consequently, on the management of financial resources.

Declarations

References

1. Bacchi S, Gluck S, Tan Y et al (2021) Mixed-data deep learning in repeated predictions of general medicine length of stay: a derivation study. Intern Emerg Med. https://doi.org/10.1007/s11739-021-02697-w
2. Casagranda I, Costantino G, Falavigna G, Furlan R, Ippoliti R (2016) Artificial Neural Networks and risk stratification models in emergency departments: the policy maker's perspective. Health Policy 120(1):111–119
3. Falavigna G, Costantino G, Furlan R, Quinn JV, Ungar A, Ippoliti R (2019) Artificial neural networks and risk stratification in emergency departments. Intern Emerg Med 14(2):291–299
4. Jeon CY, Neidell M, Jia H, Sinisi M, Larson E (2012) On the role of length of stay in healthcare-associated bloodstream infection. Infect Control Hosp Epidemiol 33(12):1213–1218
5. Heit JA, Silverstein MD, Mohr DN, Petterson TM, O'Fallon WM, Melton LJ (2000) Risk factors for deep vein thrombosis and pulmonary embolism: a population-based case-control study. Arch Intern Med 160(6):809–815
6. Hauck K, Zhao X (2011) How dangerous is a day in hospital? A model of adverse events and length of stay for medical inpatients. Med Care 49(12):1068–1075
7. Bacchi S, Gluck S, Tan Y, Chim I, Cheng J, Gilbert T, Menon DK, Jannes J, Kleinig T, Koblar S (2020) Prediction of general medical admission length of stay with natural language processing and deep learning: a pilot study. Intern Emerg Med 15(6):989–995
8. Patil T, Pandey S, Visrani K (2021) A review on basic deep learning technologies and applications. Data science and intelligent applications. Springer, Singapore, pp 565–573
9. Castiglioni I, Rundo L, Codari M, Di Leo G, Salvatore C, Interlenghi M, Gallivanone F, Cozzi A, D'amico NC, Sardanelli F (2021) AI applications to medical images: from machine learning to deep learning. Phys Med 83:9–24
10. Dogu E, Albayrak YE, Tuncay E (2021) Length of hospital stay prediction with an integrated approach of statistical-based fuzzy cognitive maps and artificial neural networks. Med Biol Eng Comput 59(3):483–496
11. Karnuta JM, Churchill JL, Haeberle HS, Nwachukwu BU, Taylor SA, Ricchetti ET, Ramkumar PN (2020) The value of artificial neural networks for predicting length of stay, discharge disposition, and inpatient costs after anatomic and reverse shoulder arthroplasty. J Shoulder Elbow Surg 29(11):2385–2394
12. Kulkarni H, Thangam M, Amin AP (2021) Artificial neural network-based prediction of prolonged length of stay and need for post-acute care in acute coronary syndrome patients undergoing percutaneous coronary intervention. Eur J Clin Invest 51:e13406. https://doi.org/10.1111/eci.13406
13. Nadif M, Role F (2021) Unsupervised and self-supervised deep learning approaches for biomedical text mining. Brief Bioinform 22(2):1592–1603