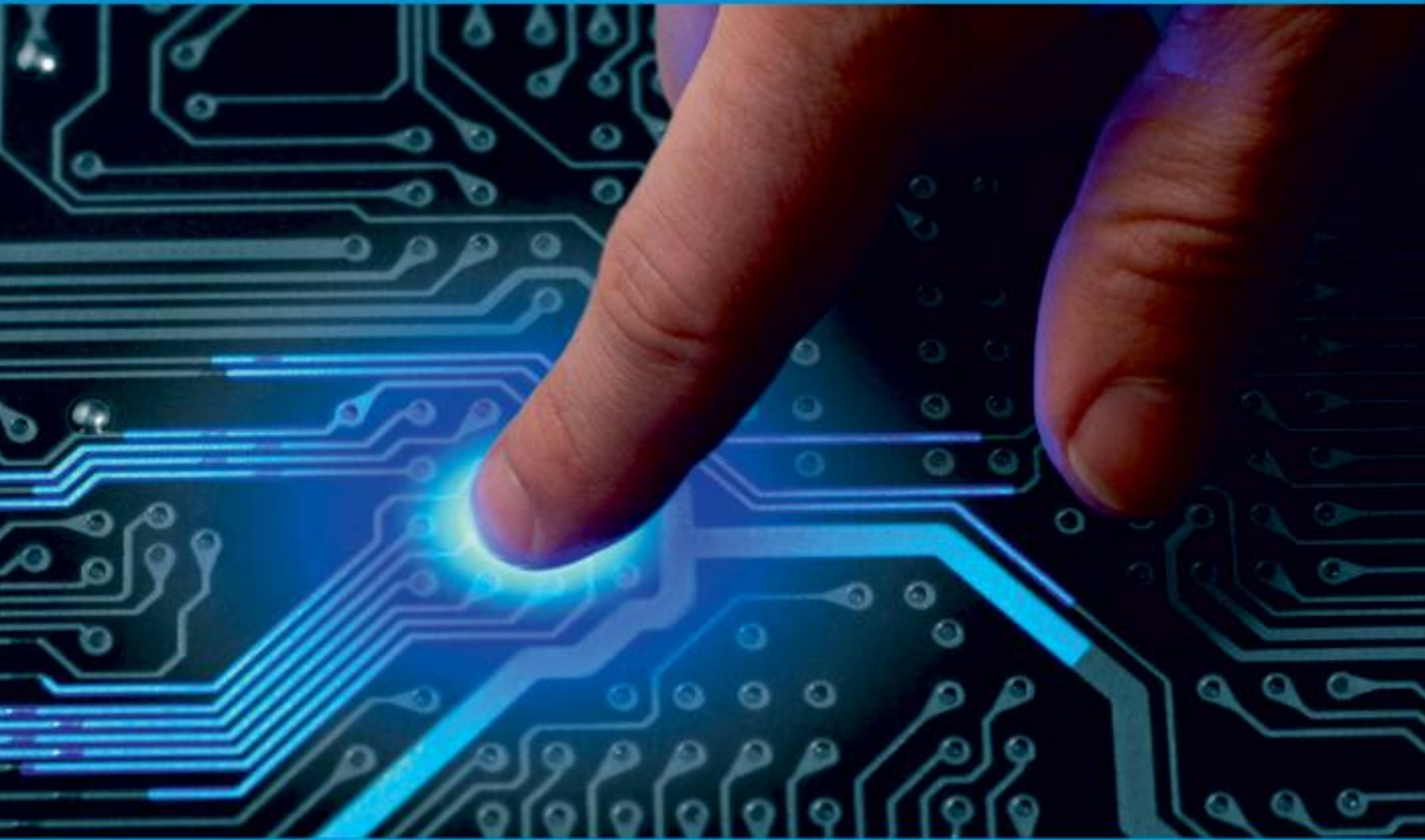
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005074|

Hospital Admission Location Prediction via Deep Interpretable Networks for the Year-round Improvement of Emergency Patient Care

M.Chenna Kesava Rao, B.Mounisha, G.Haritha, D.V.A.Aditya, G.V.Karthik

Associate Professor, Department of CSE, Tirumala Engineering College, NRT, Andhra Pradesh, India

UG Student, Department of CSE, Tirumala Engineering College, NRT, Andhra Pradesh, India

UG Student, Department of CSE, Tirumala Engineering College, NRT, Andhra Pradesh, India

UG Student, Department of CSE, Tirumala Engineering College, NRT, Andhra Pradesh, India

UG Student, Department of CSE, Tirumala Engineering College, NRT, Andhra Pradesh, India

ABSTRACT: This paper presents a deep learning method of predicting where in a hospital emergency patients will be admitted after being triaged in the Emergency Department (ED). Such a prediction will allow for the preparation of bed space in the hospital for timely care and admission of the patient as well as allocation of resource to the relevant departments, including during periods of increased demand arising from seasonal peaks in infections. Methods: The problem is posed as a multi-class classification into seven separate ward types. A novel deep learning training strategy was created that combines learning via curriculum and a multi-armed bandit to exploit this curriculum post-initial training. Results: We successfully predict the initial hospital admission location with area-under-receiveroperating-curve (AUROC) ranging between 0.60 to 0.78 for the individual wards and an overall maximum accuracy of 52% where chance corresponds to 14% for this seven-class setting. Our proposed network was able to interpret which features drove the predictions using a 'network saliency' term added to the network loss function. Conclusion: We have proven that prediction of location of admission in hospital for emergency patients is possible using information from triage in ED. We have also shown that there are certain tell-tale tests which indicate what space of the hospital a patient will use. Significance: It is hoped that this predictor will be of value to healthcare institutions by allowing for the planning of resource and bed space ahead of the need for it. This in turn should speed up the provision of care for the patient and allow flow of patients out of the ED thereby improving patient flow and the quality of care for the remaining patients within the ED.

KEYWORDS: Machine learning algorithms, Multi-layer neural networks, Patient Flow, Hospitals.

I. INTRODUCTION

DEEP neural networks (DNNs) have revolutionised the field of machine learning by providing a way to utilise very large datasets as well as large feature spaces to make meaningful predictions. State of the art performance has been achieved by DNNs in a wide range of tasks proving their efficacy as learning algorithms. Their strength in function approximation has not been overlooked by the medical community, with numerous publications exploiting them to make useful predictions for various healthcare scenarios [17, 30, 21]. One of the challenges of utilising DNNs is that they are nonconvex optimisation problems meaning the best performance that the algorithm is capable of may not be achieved [8]. As a result, much work has been carried out in developing methods R. el-Bouri, T. Zhu and D.A. Clifton are with the Department of Engineering Science, University of Oxford, Oxford, United Kindom. * - authors have equal contribution to the work. of presenting data to the network for training in a structured fashion [10]. This has since been called a curriculum and is widely used when training DNNs today. The aim of this work is to utilise the concept of curriculum training to train a model that will predict where in a hospital a patient will be admitted based on very early information obtained in the ED from the triage nurse. We aim to show that the movement of patients from ED to one of seven different ward types in hospital is predictable. This would allow allocation of a bed and resources for the patient well ahead of admission to ensure that they receive care and treatment in as timely a fashion as possible. We also aim to demonstrate that this prediction can be done given data collected from a patient at point of entry to the ED department, which in turn will improve the flow of patients out of the ED and into the hospital. Difficulties in admitting patients to the optimal hospital ward are often most marked during periods of high demand, such as during peaks in seasonal infections including influenza. We therefore test the performance of our model through out the year. In Section II we discuss the related work and in Section IV we discuss how a curriculum regularises the training of a DNN and how our algorithm is built. Then in Section VI we display the results of our algorithm and discuss these.

II. RELATED WORK

In existing literature, there is currently much work published in the monitoring of patients in hospitals using machine learning techniques [12, 20]. However the application of machine learning to model patient flow is still a relatively new topic with a consequently limited literature. Within this literature, prediction of admission to a particular ward based on measurements within hospital is a well explored area of research [18, 11, 1, 16]. Zhai et al. carried out work in predicting newly-hospitalised children who were likely to need transferral to the paediatric intensive care unit [23]. Logistic regression was used and achieved 89% accuracy. The model however only considered paediatrics, a subset of the total hospital population. While this is useful for the monitoring of the well-being of newly-hospitalised children it is not robust to be used as a general model for patient flow. An investigation into the prediction of ward transition was carried out by Xu et al. in [35]. In this work, “alternating direction method of multipliers” (ADMM) was used in conjunction with discriminative learning of mutually correcting processe to learn and predict the destination of a ward transition. The model produced an overall next location prediction accuracy of 81% when considering all patients for all wards. It would seem that the model is powerful at predicting the transition process within the hospital, however it could also be argued that this is directly due to the data that have been used. In particular, they considered all patients within the hospital and did not discriminate between emergency and non-emergency patients. It is well known that good patient flow is significantly hindered by the ad-hoc introduction of emergency admissions into the hospital [26, 22]. The authors also use the MIMIC-II dataset [14] where the majority of the wards in consideration for transfer are ICU wards. This may not be useful for analysis of patient flow in the hospital as a whole. As a result, we will only consider patients who have been admitted in an emergency, we will consider all the wards within the hospital and we will aim to predict the initial point of entry.

III. NOVELTY

The novelties of this work are as follows: we have developed a novel strategy for the training of neural networks combining a curriculum training phase with a multi-armed bandit phase to maximise prediction performance on noisy biomedical data. This also incorporates a saliency layer before the inputs which allows interpretation of the importance of the input features. To the best of the authors knowledge no other work has proposed the framework of predicting where in the hospital a patient from the ED will be admitted. This is also believed to be the first work to employ deep learning architectures in order to carry out hospital admission prediction.

IV. METHODOLOGY

Curriculum Learning Due to the non-convex nature of optimising artificial neural networks (ANNs), a structured method of presenting data to the network via curriculum learning was introduced with the aim of reducing the likelihood of the weights being optimised into a local minimum [10]. There are similarities between curriculum learning and numerical continuation methods as pointed out in [10], where optimisation of a complex surface is achieved through first optimising over smoother more convex versions of the surface. Consider a family of cost functions $C\lambda(\theta)$ such that C0 is easy to optimise over (and which is likely to be more convex than other functions), $\lambda \in [0, 1]$ is the ranking of “difficulty to optimise" and where C1 is the actual cost function that is to be minimised. By optimising over the network parameters, θ, for C0, as C0 is simply a smoother version of C1 we bring our parameters into the domain of a minimum of C0 as well as C1. We then gradually increase λ while keeping θ at the local minimum. This helps to avoid local minima which may be present in the more complex optimisation space. The aim therefore, is to create batches of data, Q, ranked according to λ (i.e., $Q\lambda$ with $\lambda = 0$ being the “easiest" batch of data to optimise progressing to the “hardest" as λ increases.) These batches

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005074|

are then presented to the network for training in order of increasing λ. Note that the batch Qλ+ will contain all of the data in Qλ for > 0, as an increment in λ represents the addition of more "complex" data to the previous batch.

$$H[Q_\lambda(z)] < H[Q_{\lambda+\epsilon}(z)] \qquad \forall\ \epsilon > 0 \qquad (1)$$

where H is the entropy of data batch Q. The weights of the examples also increase with λ as:

$$W_{\lambda+\epsilon}(z) \geq W_\lambda(z) \qquad \forall\ z,\ \forall\ \epsilon > 0 \qquad (2)$$

Regularisation using a Mahalanobis Curriculum We now postulate how the Mahalanobis curriculum may naturally regularise itself. Let Z be a training data set consisting of datapoints zn where zn ∈ Z and zn consists of input features and a label such that zn = {xn, yn}. We also define the Mahalanobis distance as:

$$d_{m_n} = \left((\boldsymbol{x_n} - \boldsymbol{\mu})^T \boldsymbol{S}^{-1} (\boldsymbol{x_n} - \boldsymbol{\mu}) \right)^{\frac{1}{2}} \qquad (3)$$

where $\boldsymbol{x_n}$ are the (continuous) input features of the datapoint, $\boldsymbol{\mu}$ is the vector of the mean value of each feature, and $\boldsymbol{S}$ is the covariance matrix.

Using this equation we can now create a vector, D_m, of distance of each datapoint from the mean of the assumed p.d.f of the dataset, where $\mathcal{X} \to D_m$, $\forall\ x_n \in \mathcal{X}$, $\mathcal{X} \subset \mathcal{Z}$, $\mathcal{Z} \subset \mathbb{R}$.

We now seek to create N batches of training data of increasing entropy of size k datapoints where $k = \frac{\mathrm{card}(D_m)}{N}$.

We then extract the indices of the lowest entropy features using the following formulation:

$$j_N = index\left(\bigcup_{i=1}^{i=Mk} \min\left((\ldots (d_m \backslash d_{m_1}) \backslash d_{m_2}) \ldots \backslash d_{m_i} \right) \right)$$

Algorithm 1 The multi-armed bandit for training of the network after initially trained with a curriculum

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

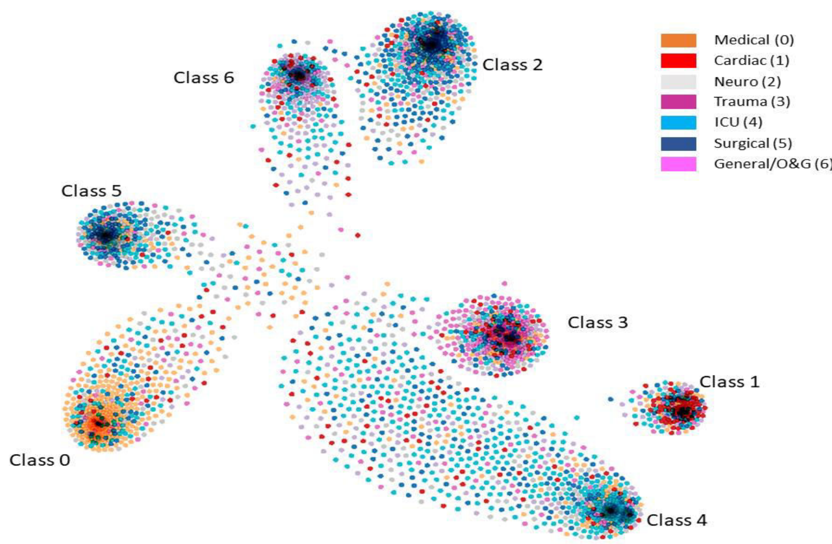
|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005074|

procedure INITIALISATION
  rate of exploration = ϵ
  resource available = N_a
  Prob. curriculum batch gives max reward = P
  Num. training data batches = $N_{choices}$
  Count of number of times batch is chosen = K
  Batches by Mahalanobis distance = $c_{batches}$
  loop:
  for i in N_a **do**:
    if $\epsilon > u \sim U(0,1)$ **then**
      batch $= c_{batches}[int\,(u \sim U(0, N_{choices}))]$
    else
      batch $= c_{batches}[\arg\max(P)]$
    Train on batch and find accuracy on training set
     $i_0 \to A_T^0 = \sum_j^C (\delta_j)$
     $i_{1:N_a} \to A_T^i = \sum_j^C \left((\delta_j^i - \delta_j^{i-1}) / \delta_j^{i-1}\right)$
    Test on validation set
    A_v^i = overall accuracy on validation set
    reward $= A_T^i \times A_v^i$
    $K[\text{batch}] = K[\text{batch}] + 1$
    $\alpha = 1/K[\text{batch}]$
    $P[\text{batch}] = P[\text{batch}] + \alpha \times (\text{reward} - P[\text{batch}])$

VI. RESULTS AND DISCUSSION

The OUH hospital in consideration has a total of 108 unique wards. To create a more meaningful and useful predictor, these were grouped by experienced clinicians working in the hospital into seven ‘ward types’ based on the type of patient that is admitted and the function of the ward. These are medical, cardiac, neurosurgical/neurology (neuro), trauma, ICU, surgical and general / obstetrics & gynaecology (general/O&G) ward types

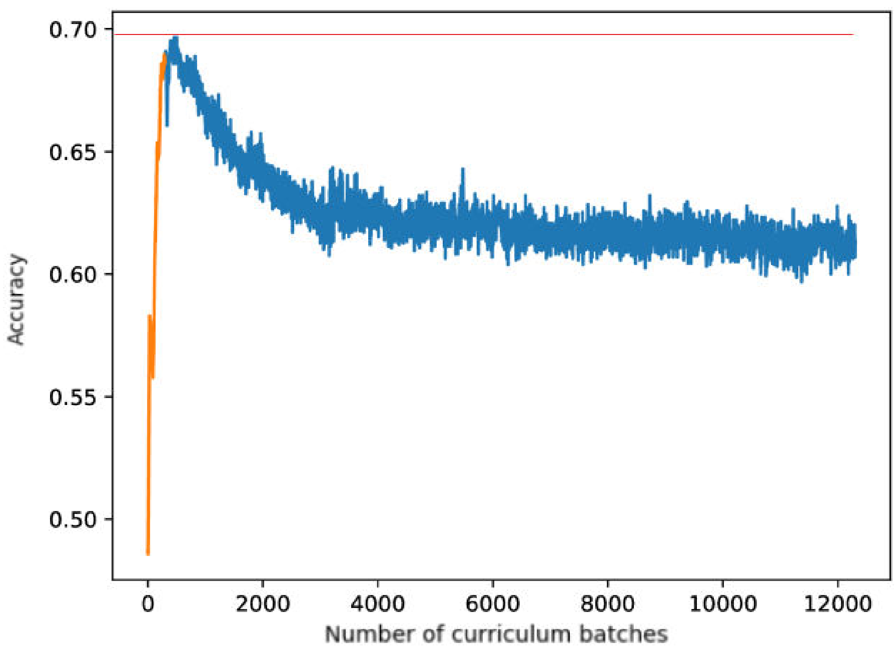


International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005074|



VII. CONCLUSION

In this article we have presented a novel method of training and regularising deep learning model with the aim of predicting where a patient presented to the ED will be admitted in an OUH Trust hospital. This prediction will aid in the provision of timely care and treatment for the patient and those still in the ED. Our model achieves AUC values between 0.60 and 0.78 for the individual ward types. Furthermore, our model also provides an explanation as to the cause of the predictions, allowing the user to incorporate more important features for individual ward types in the future. The authors believe this may be useful for ensuring timely admission to hospital and reducing the time to care. This will in turn improve the quality of care for patients still in the ED due to less crowding. This work may also be useful for resource prediction and optimisation in hospitals more generally.

VIII. FUTURE WORK

The model presented in this work is first trained using a curriculum and then using the curriculum batches a multiarmed bandit is employed to improve the performance. While the algorithm described in Algorithm 1 is non-stationary, it is weakly non-stationary relying on the number of pulls of a certain batch to reduce the probability of choosing said batch. As a result, we will improve this by turning this problem into a full reinforcement learning problem. Treating the weights of the network as the state space, we will train a policy to select the best action to take (batch to train on) given the state space. We believe this will be a much more effective method of training due to the information provided to the trainer about the state of the weights of the network. We would also like to further investigate features that can be obtained from the ED which correlate highly with the individual ward types. In doing so we will be able to reduce the input feature space and advise clinicians in the ED what needs to be measured for this prediction problem. It is hoped that by doing this, we will be able to mitigate the problem of missing features which can commonly happen in models with large input spaces. We will continue investigating methods of identifying when patients were admitted to wards that were not ideal for their treatment. We believe that finding these cases will help to improve the performance of our models due to their reliance on historical data. We will also seek to integrate data on the equipment used during a patient stay to better inform the model of which wards are appropriate for admission.

IX.ACKNOWLEDGMENT

We thank all the people of Oxfordshire who contribute to the Infections in Oxfordshire Research Database. This work uses data provided by patients and collected by the NHS as part of their care and support. Research Database Team (Oxford): R Alstead, C Bunch, DW Crook, J Davies, J Finney, J Gearing (community), H Jones, L O'Connor, TEA Peto (PI), TP Quan, J Robinson (community), B Shine, AS Walker, D Waller, D Wyllie. Patient and Public Panel: G Blower, C Mancey, P McLoughlin, B Nichols. This work was funded by the EPSRC Doctoral Training Award as part of the UK's high-priority "industrial strategy" funding scheme and the National Institute for Health Research Oxford Biomedical Research Centre. Tingting Zhu was supported by the Royal Academy of Engineering Research Fellowship. David Eyre is a Big Data Institute Robertson Foundation Fellow.

REFERENCES

[1] Andrew K Diehl, Max D Morris, and Stephen A Mannis. "Use of calendar and weather data to predict walk-in attendance." In: Southern medical journal 74.6 (1981), pp. 709–712

[2] Bruce D McCarthy, John B Wong, and Harry P Selker. "Detecting acute cardiac ischemia in the emergency department". In: Journal of General Internal Medicine 5.4 (1990), pp. 365–373.

[3]Eugene Somoza and John R Somoza. "A neuralnetwork approach to predicting admission decisions in a psychiatric emergency room". In: Medical Decision Making 13.4 (1993), pp. 273–280.

[4] Roy De Maesschalck, Delphine Jouan-Rimbaud, and Désiré L Massart. "The mahalanobis distance". In: Chemometrics and intelligent laboratory systems 50.1 (2000), pp. 1–18.

[5] Ary L Goldberger et al. "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals". In: Circulation 101.23 (2000), e215–e220.

[6] Jamary Oliveira-Filho et al. "Stroke as the first manifestation of calcific aortic stenosis". In: Cerebrovascular Diseases 10.5 (2000), pp. 413–416. [7] Peter Auer et al. "The nonstochastic multiarmed bandit problem". In: SIAM journal on computing 32.1 (2002), pp. 48–77.

[8] Christopher M Bishop. Pattern recognition and machine learning. springer, 2006.

[9] L.J.P. van der Maaten and G.E. Hinton. "Visualizing High-Dimensional Data Using t-SNE". In: Journal of Machine Learning Research 9 (2008), pp. 2579–2605.

[10] Yoshua Bengio et al. "Curriculum learning". In: Proceedings of the 26th annual international conference on machine learning. ACM. 2009, pp. 41–48

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING