International Journal of Gender, Science and Technology
ISSN: 2040-0748
UGC Care Group I Journal.
Vol-12 Issue-02 Oct 2023

https://ijgst.com.2023.v12.i2.pp136-148

MAKING USE OF UNSTRUCTURED MEDICAL NOTES TO PREDICT PATIENTS' HEALTH OUTCOMES OVER THE LONG RUN

D HAPPY BALASOWRI [1] A.SWATHI [2] K.ANUSHA [3] MANDE PRAVEEN [4]
[1]ASST.PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[2] ASST. PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[3]ASST. PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[4] ASST. PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[1,2,3,4] SRI MITTAPALLI COLLEGE OF ENGINEERING

Abstract—In order to shift focus to accurate risk categorisation, patient data are an essential duty. The digitisation of structured electronic health records is crucial to the most recent iterations. Unstructured clinical notes are the most common way for caretakers to record the progression of their patients' illnesses, but these models fail to take them into account. The availability of such individual patient data presents a once in a lifetime opportunity to create smart systems that shed light on the physiology of patients' diseases. Also, compared to state-of-the-art models, relatively few studies have tried to quantify how well deep neural networks perform on publicly accessible datasets. Important results about the usefulness of deep learning models for the clinical job of ICD-9 group prediction are presented in this paper based on our benchmarking tests. We provide FarSight, an aggregation approach for the long term that can detect the disease's early warning signals. Vector space and topic modelling approaches are used to extract the semantic information from the patient representations. Results from the MIMIC-III database studies showed that proposed unbuilt-data models outperformed a structured EHR-based state-of-the-art model in AUPRC by 19.34% and 5.41%, respectively.

.

Keywords—Clinical decision support systems, disease prediction, healthcare analytics, ICD-9 code group prediction, precision medicine.

I. INTRODUCTION

In order to make important clinical decisions and direct life-saving patient treatment in the early stages, accurate illness forecasts and metrics of early stage diagnosis of patients' health are required [1]. Until recently, patients had limited diagnostic options due to the healthcare sector's conservative treatment policy [2]. In the wake of rapid technology advancement and widespread digitisation, there has been a surge in the use of easily available heterogeneous clinical data for the sake of improving people's quality of life and using this data in evidence-based medicine. The Intensive Care Units (ICUs) in hospitals, for instance, are critical care settings that depend on continuous monitoring of many factors pertaining to the condition of critically ill patients; these units also produce massive amounts of data. Better risk assessment, lower mortality and morbidity rates, and superior treatment recommendations might all result from better use of this data in future iterations of clinical decision support systems (CDSSs). If you want more predictive power, you need the CDSSs. Symptoms, treatments, medications, diagnostic codes, and test results are all part of the structured data that is manually put into Electronic Health

Records. This information is valuable for healthcare purposes. A lot of people are interested in the idea of using machine learning and deep learning to model data from electronic health records (EHRs) for purposes including survival analysis, mortality prediction, causal etetal inference, detecting physiological deterioration, and more [3]. Low adoption of structured EHRs is seen in developing countries, despite the fact that they significantly facilitate processes in precision medicine. Clinical decision-making in these countries' healthcare systems is dependent on human evaluation of unstructured nursing notes. Statistical analysis is a frequent part of building CDSSs used in hospitals using organised EHR data. However, medical images and unstructured clinical writing reveal a great deal about a patient's state. Clinical nursing notes are a primary means of documenting subjective evaluations and worries about patients' medical state (see Fig. 1). The notes include valuable evaluations and insights from the visiting doctors and nurses who constantly keep tabs on the patients. Data mining and modelling on this kind of information may help unearth previously unknown linkages and trends in clinical decision-making. Based on current study in the area of health information [2], the information contained in unstructured nursing notes is based on carer observations and intuitions that don't match the accompanying structured data.

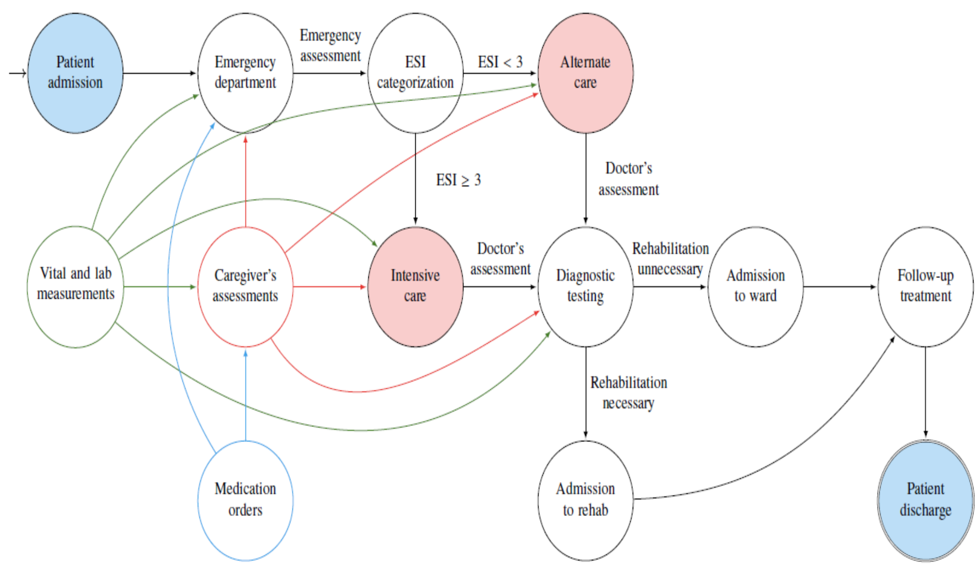

Fig. 1: Changes in a patient's trajectory related to a single hospital stay (across many episodes).

II. RELATEDWORKS

Many initiatives have been launched to tap into the abundance and diversity of health data stored in electronic health records (EHRs). In order to better manage hospital resources, identify patients at high risk, and arrange for tailored treatment, healthcare systems are constantly evolving (e.g., Intelligent Care Delivery Analytics [12] and MatrixFlow [11]). The current equipment and deep learning models used for medical result forecasting will be briefly reviewed in this part. The use of machine learning models, especially feed-to-neural networks, to assess medical risk and mortality prediction yielded promising results in early studies[13], [14]. Additionally, when estimating the probability of mortality among hospitalised patients, neural feed forward networks virtually always outperform LR and other severity of illness measures. This was shown by Celi et al. New neural architectures, made possible by recent advances in deep learning, have shown promising results in a number of tasks, including the prediction of length of stay, hospital mortality, general EHR diagnoses, and diagnostic code groups [16]. Che et al.[17] used pre-existing medical ontology data to construct a scalable, deep neural network for disease diagnosis, which achieved clinically meaningful features. Through a feedforward-neural network, Dabek & Caban[18] have been trying to improve predictability of psychiatric diseases like depression, behavioural disorders, anxiety, and post-traumatic stress disorder (PTSD).

With a rich context and variation dropout, Khin [19] trained a Bi-LSTM model to improve performance and speed convergence in nursing grades. The previous research shows that the deep neural architecture and the leveraging machine are quite effective in health applications. Many studies have attempted to mimic the progression of illnesses by focussing on either acute hospital episodes or chronic diseases. Using a cluster of 45 clinical, physiologic, and intensive care unit therapeutic variables, Cohen et al. [20] have improved patient monitoring and identified complex metabolic states. Zhou et al. [21] combined the lab test, cognitive data, and demographic information to build a model of disease progression based upon fused group lasso formulation. By combining disparate patient datasets with incomplete data, Wang et al. [22] created disease progression models that account for clinical symptoms and co-morbidities. Choi et al. [24] used the context-sensitive Hawkes multivariate technique to model the temporal course and infer disease association networks for patient-specific illness prediction. In order to implicitly include co-morbidities in hidden layers, several research used clinical data from timeseries for feedforward neural networks[24], LSTM networks[25], and ConvNets time-scale prediction [26]. Not long ago, deep neural networks were used to model time series clinical data and sickness [27]. Early patient-specific sickness prediction is essential for the establishment of efficient CDSS, as this article shows. Feldman et al. [29] dismantled the radiology, physicist, clinical care, and electrocardiogram (ECG) stories in order to explore the lexical, structural, and real differences in unstructured medical narratives. All the authors did was provide the groundwork, and our research built upon that to effectively mine clinical data. In their work on patient health, Zalewski et al. [30] laid up a strategy to combine several medical risk assessment techniques. In order to deal with the sparse and high-dimensional nursing billets (found in the MIMIC-II database), their solution employs Hierarchical Dirichlet methods. Nevertheless, the authors have relied on an LR predictor to aid in the computation of mortality without comparing their results to those of deep neural networks and later research. To accurately forecast clinical events and outcomes (such as mortality and diagnostic code group), de-identified healthcare datasets like MIMIC-II and MIMIC-III make it possible to assess machine and deep learning models. The super-student technique, a

collection of several learner models, outperformed typical severity-of-fill methods, according to Pirracchio[7], who used the MIMIC-II database to predict patient-specific fatalities in intensive care units. Though it outperformed traditional prognostic scoring systems, super-learner's algorithm still couldn't compete with state-of-the-art hardware and deep learning models. The challenges of reproducing the findings of 28 recent articles on the publicly accessible MIMICIII database have been investigated by Johnson et al. [8]. They compared the reported performance using the generated characteristics to that of LR and gradient boost models after extracting a basic set of characteristics. Furthermore, the authors emphasised the need for improved performance reporting to address the significant study heterogeneity and allow for fair comparisons of the proposed methodologies. Recent study by Harutyunyan et al. [9] used MIMIC-III and multi-taking recurrent neural networks (RNNs) to experimentally validate four Clinical Prediction Benchmarking challenges. Unfortunately, the author has neglected to evaluate state-of-the-art machine learning models (especially super-learner) and prognostic score systems against their potential clinical prediction results, instead focussing only on LR and LST models. A variety of clinical prediction tasks, such as death prediction, long-term residency, and ICD-9 group prediction using MIMIC-III, were consistently and comprehensively studied by Purushotham et al. [1]. In their investigation, they put a plethora of machine learning algorithms and severity ratings to the test. As a result, clinical nursing notes, which include important information, are often disregarded. To tackle the clinical problem of patient-specific mortality prediction, Krishnan and Kamath have introduced a novel approach to hybrid metaheuristic modelling of functionality for processing large-scale laboratory events [31]. On the other hand, your approach simplifies clinical prediction by using data from large-scale organised laboratory events. Using the methodology developed by Huang et al., the MMIC-III Database was modelled using top-tier deep neural networks to forecast the (top-ten) ICD-9 category. [32]
To ameliorate the diagnostic coding process, Zeng et al.[33] laid up a comprehensive framework for the extraction of domain information from medical area headings. Patients' important information provided in informal nursing notes is ignored in these recent publications that focus on modelling unstructured text for clinical outcome prediction. Furthermore, modelling the clinician notes aids in the establishment of health policies and promotes effective clinical decision making, which alone enable correct billing, in contrast to release summaries. In order to improve CDSS triage accuracy, Stone[34] explored several avenues for assisting medical personnel with a limited history of chronic trauma in high-pressure distraction settings. Expanding the author's e-sorts, we stratify medical risk using full patient-centered information. This helps to lower the underlying risk of clinical deterioration in CDSS, improves triage accuracy, and achieves optimum patient outcomes.

For the purpose of developing intelligent CDSS, our research focusses on unstructured clinical health notes, a resource that is currently underused in the healthcare industry. This approach would be especially helpful for hospitals in low-income rural areas that have not yet adopted formal electronic health records. We want to improve the state-of-the-art research by simulating these reports and their important data, which is often lost when nursing notes are transcribing them into structured EHRs. Furthermore, our study paves the way for extensive comparative studies to examine the efficacy of data modelling approaches for clinical reasons in predicting ICD-9 code groups using various deep neural architectures.

III. PROPOSED SYSTEM

Dataset Specifics and Cohort Selection

An important public critical care database, the MIMIC-III database (v1:4) was created and is maintained by the Massachusetts Institute of Technology Laboratory for Computing Physiology. From 2001 to 2012, over 60,000 patients were admitted to the Beth Israel Deaconess Medical Centre in the United States, and this data set includes a wide variety of de-identified health-related information. Factors such as patient demographics, medical history, test results, treatments, medications, imaging reports, notes from carers, and hospital and non-hospital-related deaths are all part of the vital statistics. In addition, it encourages a wide range of analytical studies, such as those pertaining to epidemiology, electronic instrument development, and clinical decision making.

Data Extraction

Notice events, admissions, patients, and diagnoses icd data from our most recent patient cohort are part of the 26 relational tables that make up the MIMIC-III database. For both inpatient and outpatient care, this table includes unstructured nursing notes, ECG, echo, and radiology data. The admission table is used to get the time of admission to the intensive care unit (ICU) for patients by storing their information. The birth dates of all MIMIC-III participants were extracted from the patient table, which contains charted data. The diagnoses icd table displays the ICD-9 codes for MIMICI-III patients. These pills were chosen for the ICD-9 prediction group work because they include the most relevant data and clinical diagnostic features.

Data Cleansing, Aggregation, and Preprocessing

The data obtained by MIMICIII contains various inaccuracies caused by factors including noise, incorrect or duplicate entries, outliers, clerical mistakes, and missing information. The first step is to identify and remove any nursing notes that have clerical errors or incorrect information by using the iserror feature of the Notes Table. Additionally, duplicate patient data have been removed. Nursing notes from 6,532 patients made up the final patient cohort after processing incorrect inputs; FarSight aggregated the statistics in these notes to establish the onset of the illness based on the earliest symptoms recorded.

.

Clinical Feature Modeling

While rule-based NLP transformations and traditional dictionary transformations show promise, adaptability and automation still need human intervention. For a CDSS's classification models to work, the underlying

https://ijgst.com.2023.v12.i2.pp136-148

group of unstructured nursing notes must be optimally represented vector-wise. In this study, we use topical modelling and vector space to standardise the unstructured medical terminology and guarantee accurate representation of the patient group. Words Bag (BoW) and similar algorithms do a good job of encoding each clinical term once, but models using them have issues with sparsity and high dimensionality. Furthermore, BoW

transformations are unable to grasp the concept of semantically identical clinical notes with similar vectors. Words like "tumour" and "cancer"—which have an enclosed semantine

connection—can be transformed into entries with a big distance in BoW transformation. Take this as an example. Paragraph Vector (PV) networks, such as Doc2Vec, provide a data-driven approach to therapeutic word representations. Numerical variable nursing notes may be transformed into low-dimensional fixed-length document embeddings using Doc2Vec. Connected to the contents, the measurements typically learn the distributions via a shallow hidden layer in a neural network structure. To forecast the next phrase, each corpus makes use of a number of clinical terms. What follows is the main concept. A self-studied embedding matrix is used to convert these clinical words to numerical vectors. Then, a neural network is fed these vectors, and the term output is forecasted. Training the parameters of the neural network and the word embedding matrix is done using the stochastic gradient descent of the mini-batch. Doc2Vec improves upon this basic concept by adding an additional vector that represents the document's semantics. Many natural language processing tasks rely on the ability to detect emotions in texts using Doc2Vec's semantic textual features.

ICD-9 Code **G**roup **P**rediction

Adaptability and automation still need human involvement, despite the promise of rule-based natural language processing transformations and classical dictionary transformations. There has to be an ideal vector-wise representation of the underlying group of unstructured nursing notes for a CDSS's classification models to function. To provide a realistic portrayal of the patient group and standardise the unstructured medical language, we use topical modelling and vector space in this work. While techniques like Words Bag (BoW) efficiently encode each clinical phrase once, models that rely on them struggle with large dimensionality and sparsity. Not only that, but BoW transformations just can't get their heads around the idea of vectors that are semantically equivalent to clinical notes. The embedded semantine link in words like "tumour" and "cancer" allows them to be turned into entries with a significant distance in BoW transformation. Here is an example for you. The paragraph A data-driven approach to therapeutic word representations is provided by vector (PV) networks like Doc2Vec. Use Doc2Vec to convert numerical variable nursing notes to low-dimensional fixed-length document embeddings. In a conventional neural network architecture, the measurements are linked to the contents and learn the distributions via a shallow hidden layer. Each corpus uses a variety of medical phrases to predict the next sentence. The key idea is as

https://ijgst.com.2023.v12.i2.pp136-148

follows. To transform these medical terms into numerical vectors, a self-studied embedding matrix is used. The next step is to feed these vectors into a neural network in order to anticipate the term output. The word embedding matrix and neural network parameters are trained using the stochastic gradient descent of the mini-batch. By including an extra vector that stands in for the document's semantics, Doc2Vec enhances this fundamental idea. The capacity to identify emotional content in texts by using Doc2Vec's semantic textual properties is crucial for several NLP applications.

Deep Neural Architectures

The promise of rule-based NLP transformations and classical dictionary transformations is not enough to make adaptability and automation possible without human intervention. In order for the classification models of a CDSS to work, the underlying group of unstructured nursing notes must be idealised vector-wise. Topical modelling and

vector space are used in this study to

standardise the unstructured medical terminology and provide a realistic representation of the patient group. Models dependent on Words Bag (BoW) and similar approaches have difficulty with sparsity and huge dimensions, even if they effectively encode each clinical phrase once. Beyond that, the concept of vectors that are semantically equal to clinical notes just does not sit well with BoW transformations. Words like "tumour" and "cancer" may be transformed into entries with a considerable distance in BoW transformation because of the inherent semantine connection. Give me an example of what I mean. The text Vector (PV) networks, such as Doc2Vec, provide a data-driven method for therapeutic word representations. Document embeddings with low dimensionality and fixed length may be created from numerical variable nursing notes using Doc2Vec. To learn distributions, a typical neural network design has a shallow hidden layer that is connected to the contents. The following sentence is predicted by each corpus using a different set of medical expressions. What follows is the main point. A self-studied embedding matrix is used to convert these medical phrases into numerical vectors. After that, you may use a neural network to predict the term output by feeding it these vectors. Using the mini-batch's stochastic gradient descent, we train the neural network parameters and word embedding matrix. Doc2Vec improves upon this basic concept by adding an additional vector that represents the document's semantics. An important capability for many natural language processing applications is the ability to detect emotional content in texts by using Doc2Vec' semantic textual features.

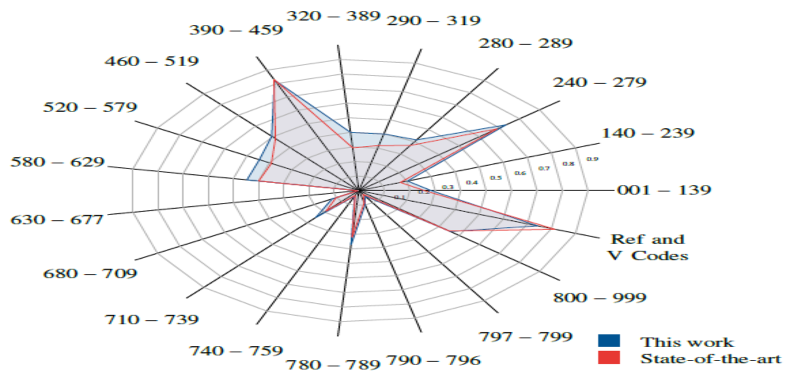

Fig. 2: Comparison with the state-of-the-art model [1] concerning the percentage of patients within an ICD-9 code group.

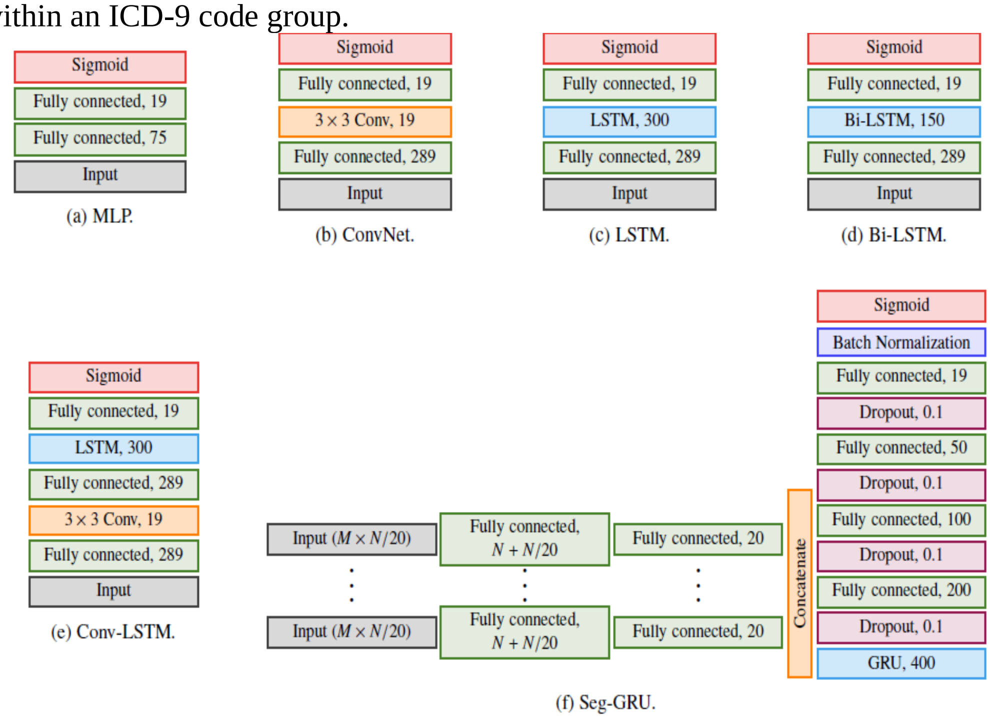

Fig. 3: Schematic overview of the deep neural architectures employed in this study.

IV. RESULTS AND DISCUSSION

In order to validate our method, we thoroughly benchmarked the clinical nursing notes obtained from the MIMIC-III database according to the specified cohort. For a specific clinical nursing note, the Multi-Label Prediction presented a significant challenge in predicting a set of probable ICD-9 coding categories. In order to assess the methods' predictive power, we conduct a two-way comparison of the actual and predicted sets of diagnostic code groups using five-times cross-validation; we also present the standard median and medial errors. Also, to accurately assess the efficacy

https://ijgst.com.2023.v12.i2.pp136-148

of the proposed methods, we used five conventional metrics: Accuracy (ACC), MCC score, F1 score, Area Under the Precise-Recall Curve (AUPRC), and Area Under the ROC curve (AUROC). This research compares the expected and real diagnosis codes. Figure 4 shows that out of 140.792 nursing notes, 520-579, 580-629, and 800-999 had less than 500 mismatches (<0.35%), while the ranks for the diagnostic codes 001-139, 280-289, 320-389, 460-519, 630-677, and 780-789 were below 100 mismatches (<0.007%).

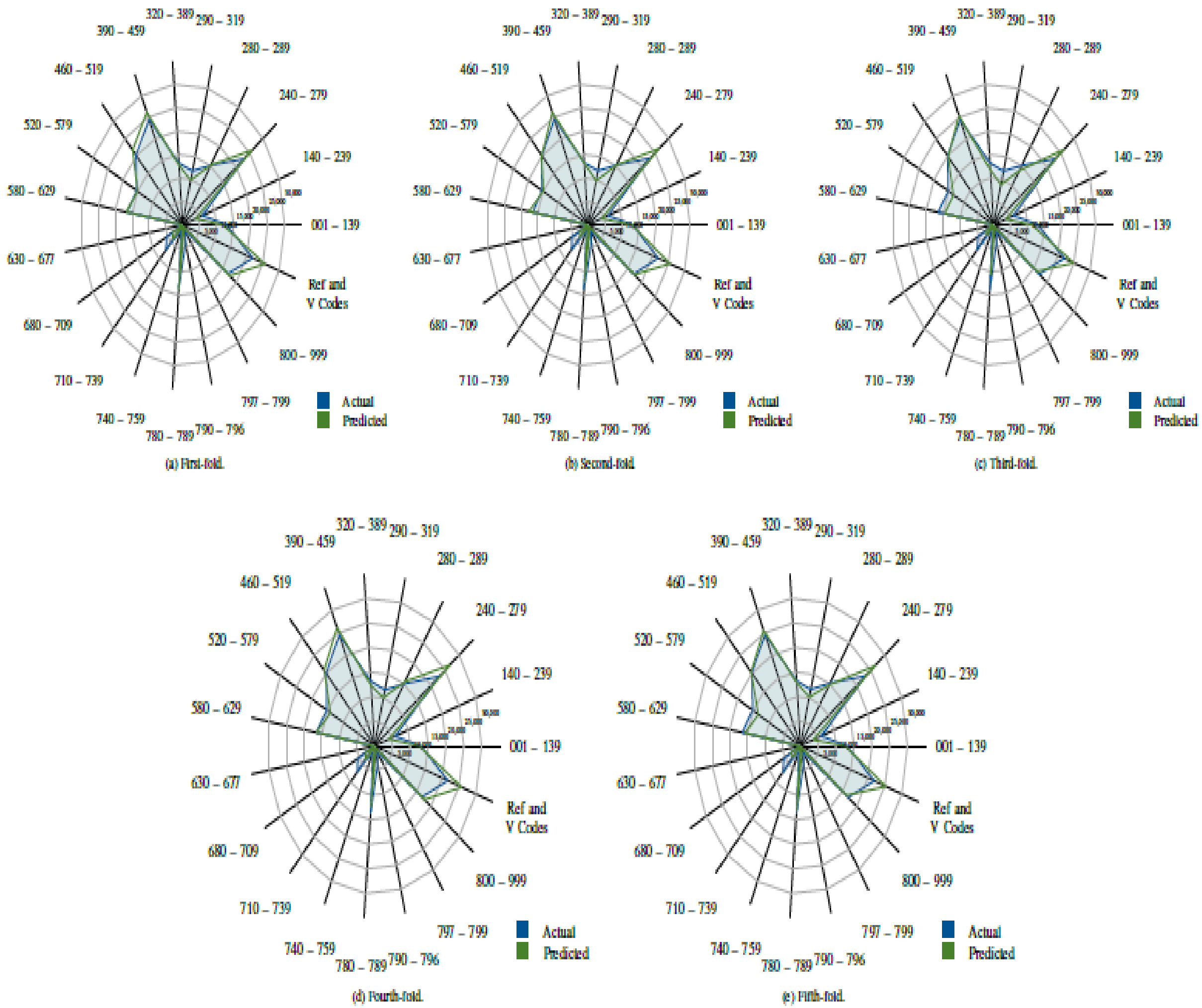

Fig. 4: Comparison of the number of nursing notes concerning actual and predicted (using Conv-LSTM trained on NMF-TW with SC representations) ICD-9 code groups across the cross-validation folds.

V. FUTURE SCOPE AND CONCLUSION

In order to aid in clinical decision-making, this research used FarSight, a method for the long-term accumulation of illnesses with the earliest symptoms recorded, which is essential in the creation of priority care. With a focus on length, heterogeneity, volume, and structural complexity, our approach is built around the valuable patient-provided information found in unstructured nursing notes. To effectively apply deep neural networks for illness forecasting, our technique employs vector space and NMF-theme modelling. The concepts were tested using five conventional evaluation metrics to see how well they performed in various circumstances, including reliability, hit and miss rates, predictability, and the ability to handle imbalanced data. By a margin of 19.34% in AUPRC and 5.41% in AUROC, the suggested FarSight NMF TW with the SC model was able to gather the wealth of information included in informal nursing notes, surpassing the structured EHR model. Also, this particular SC model is brand new. Furthermore, when contrasted with the naïve note aggregation approach, FarSight's unstructured

modelling performance has significantly improved. In addition, our idea lessens the reliance on structured EHRs, which is a key reason why developing nations have low adoption rates of these systems. Although our method is already quite good at doing predictive risk assessments, we are always on the lookout for ways to make it even better. Before eliminating the need for organised EHR data, our modelling approach models the unstructured health care text on the assumption that care notes document all clinical assessments. Additionally, the clinical data is not included into this research in real-time, and MIMIC-III data is exclusive to the current modelling approach. As part of our ongoing effort to replicate structured EHR metrics, we also want to include rich information from nursing notes. In addition, we want to develop more efficient methods for real-time clinical data modelling by evaluating and analysing various instantly obtained clinical indicators. This will help us satisfy the need for accurate and time-consuming CD SS models in practical situations. Clinical data modelling in time series, real-time decision-making, and epidemiology and demography research will all benefit from the CDSS. In addition, the prediction model would be continuously improved by auditing the system's performance and recalculating the accuracy in various clinic scenarios. Thus, the proposed system may be easily connected with an existing hospital information system and is intended as a clinical interface. It may aid carers in conducting real-time assessments that are both patient-specific and based on evidence. Implementing the ICD-10-CM medical taxonomy and creating a more robust CDSS may further improve the study's learning and outcomes.

References

1. Goldstein BA, Navar AM, Pencina MJ, et al. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc 2017;24(1):198–208. CrossRefPubMedGoogle Scholar
2. Khalid S, Yang C, Blacketer C, et al. A standardized analytics pipeline for reliable and rapid development and validation of prediction models using observational health

data. Comput Methods Programs Biomed 2021 doi: 10.1016/j.cmpb.2021.106394.
CrossRefGoogle Scholar

3. Ford E, Carroll JA, Smith HE, et al. Extracting information from the text of electronic medical records to improve case detection: a systematic review. J Am Med Inform Assoc 2016;23(5):1007–15.
CrossRefPubMedGoogle Scholar
4. Hahn U, Oleynik M. Medical Information Extraction in the Age of Deep Learning. Yearb Med Inform 2020;29(1):208.
Google Scholar
5. Spasic I, Nenadic G. Clinical Text Data in Machine Learning: Systematic Review. JMIR Med Inform 2020;8(3):e17984.
Google Scholar
6. Ashutosh N. Aggarwal, Pralay Sarkar, Dheeraj Gupta, Surinder K. Jindal Performance of standard severity scoring systems for outcome prediction in patients admitted to a respiratory intensive care unit in north india Respirology, 11 (2) (2006), pp. 196-204, 10.1111/j.1440-1843.2006.00828.x
7. Peng, Y., Wang, X., Lu, L., Bagheri, M., Summers, R., & Lu, Z. (2018). *NegBio: A high-performance tool for negation and uncertainty detection in radiology reports*. AMIA Summits on Translational Science Proceedings, 2017, 188-196.
8. Perera, S., Mendes, P., Sheth, A., Thirunarayan, K., Alex, A., Heid, C., & Mott, G. (2015). *Implicit entity recognition in clinical documents*. In Proceedings of the Fourth Joint Conference on Lexical and Computational Semantics (pp. 228–238).
9. Perez, J., Perez, A., Casillas, A., & Gojenola, K. (2018). *Cardiology record multi-label classification using latent Dirichlet allocation*. Computer Methods and Programs in Biomedicine, 164, 111–119.
10. PIAAC. (2017). Survey of adult skills: Programme for the international assessment of adult Competencies, Paris: The Organisation for Economic Co-operation and Development.
11. Pylieva, H., Chernodub, A., Grabar, N., & Hamon, T. (2018). *Improving automatic categorization of technical vs. laymen medical words using fasttext word embeddings*. In 1st International Workshop on Informatics and Data-Driven Medicine, IDDM.
12. Pyysalo, S., Ginter, F., Moen, H., Salakoski, T., & Ananiadou, S. (2013). *Distributional semantics resources for biomedical text processing*. In Proceedings of LBM (pp. 39–44).
13. Qiang, J., Li, Y., Zhu, Y., Yuan, Y., & Wu, X. (2020). *Lexical simplification with pretrained encoders*. In Proceedings of the AAAI Conference on Artificial Intelligence (pp. 8649–8656).

14. Rothrock, S. G., Rothrock, A. N., Swetland, S. B., Pagane, M., Isaak, S. A., Romney, J., & Chavez, S. H. (2019). *Quality, trustworthiness, readability, and accuracy of medical information regarding common pediatric emergency medicine-related complaints on the web*. The Journal of Emergency Medicine, 57, 469–477.
15. Rumshisky, A., Ghassemi, M., Naumann, T., Szolovits, P., Castro, V. M., McCoy, T. H., & Perlis, R. H. (2016). *Predicting early psychiatric readmission with natural language processing of narrative discharge summaries*. Translational Psychiatry, 6, e921.
16. Sakakini, T., Lee, J. Y., Duri, A., Azevedo, R. F., Sadauskas, V., Gu, K., … & Walayat, S. (2020). *Context-aware automatic text simplification of health materials in low-resource domains*. In Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis (pp. 115–126).
17. Shardlow, M., & Nawaz, R. (2019). Neural text simplification of clinical letters with a domain specific phrase table. Florence: Association for Computational Linguistics.
18. Tran, T., & Kavuluru, R. (2017). *Predicting mental conditions based on "history of present illness" in psychiatric notes with deep neural networks*. Journal of Biomedical Informatics, 75S, S138–S148.
19. Van den Bercken, L., Sips, R.-J., & Lofi, C. (2019). *Evaluating neural text simplification in the medical domain*. In The World Wide Web Conference (pp. 3286–3292).
20. Van, H., Kauchak, D., & Leroy, G. (2020). *AutoMeTS: The autocomplete for medical text simplification*. Available from https://arxiv.org/abs/2010.10573 (accessed 16 June 2021).
21. Wang, B., Xie, Q., Pei, J., Tiwari, P., & Li, Z. (2021). *Pre-trained language models in biomedical domain: A systematic survey*. Available from https://arxiv.org/abs/2110.05006 (accessed 16 April 2022).
22. Wei, C.-H., Leaman, R., & Lu, Z. (2015). *SimConcept: A hybrid approach for simplifying composite named entities in biomedical text*. IEEE Journal of Biomedical and Health Informatics, 19, 1385–1391.
23. Weissman, G. E., Harhay, M. O., Lugo, R. M., Fuchs, B. D., Halpern, S. D., & Mikkelsen, M. E. (2016). *Natural language processing to assess documentation of features of critical illness in discharge documents of acute respiratory distress syndrome survivors*. Annals of the American Thoracic Society, 13, 1538–1545.

24. Wulff, A., Mast, M., Hassler, M., Montag, S., Marschollek, M., & Jack, T. (2020). *Designing an openEHR-based pipeline for extracting and standardizing unstructured clinical data using natural language processing*. Methods of Information in Medicine, 59, e64–e78.
25. Xu, W., Napoles, C., Pavlick, E., Chen, Q., & Callison-Burch, C. (2016). Optimizing statistical machine translation for text simplification (pp. 401–415). Cambridge, MA: Transactions of the Association for Computational Linguistics.
26. Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). *Deep learning for healthcare: Review, opportunities and challenges*. Briefings in Bioinformatics, 19, 1236–1246.
27. Moradi, M., & Ghadiri, N. (2018). *Different approaches for identifying important concepts in probabilistic biomedical text summarization*. Artificial Intelligence in Medicine, 84, 101–116.
28. Morrison, A. K., Glick, A., & Yin, H. S. (2019). *Health literacy: Implications for child health*. Pediatrics in Review, 40, 263–277.
29. Mukherjee, P., Leroy, G., Kauchak, D., Rajanarayanan, S., Romero Diaz, D. Y. R., Yuan, N. P., Pritchard, T. G. & Colina, S. (2017). *NegAIT: A new parser for medical text simplification using morphological, sentential and double negation*. Journal of Biomedical Informatics, 69, 55–62.
30. NIH (2020). Healthy people 2030. Washington, DC: the U.S. Department of Health and Human Services.
31. Nisioi, S., Štajner, S., Ponzetto, S. P., & Dinu, L. P. (2017). *Exploring neural text simplification models*. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Vol. 2, pp. 85–91), Short papers.