Temporal modeling of patient journeys using deep learning on longitudinal health records

Veerendra Nath Jasthi

veerendranathjasthi@gmail.com

Abstract:
The temporal dynamics of the patient journeys are important to make better clinical decisions as well as predict the outcomes. Longitudinal health records record time-stamped progressions of diagnoses, treatment, and other clinical happenings that show disease progression patterns and a reaction to treatment. The aim of the present paper is to design a deep learning methodology to capture temporal properties of patient journeys using longitudinal electronic health records (EHRs). Based on the recurrent neural networks (RNNs), temporal convolutional networks (TCNs), and transformer-based structures, we could obtain the temporal dependencies and predictive features of disease onset, hospital readmission, and mortality. To assess our models we consider a publicly available healthcare dataset, which shows better results at predicting and sequence representation compared to conventional models. We found that deep learning temporal models present an effective solution to proactive and personalized health care.

Keywords: Temporal Modeling, Patient Journey, Deep Learning, Longitudinal Health Records, Electronic Health Records (EHR), RNN, Transformer, Healthcare Analytics.

I. INTRODUCTION

Electronic Health Records (EHRs) have in recent years become critical tools in knowing the journey of patients in the changing healthcare landscapes. These are digitized records of years of patient clinical experience, made up of disease diagnosis, medication, procedure, lab results and discharge summaries [15]. More importantly, the events are time-stamped, so there can be a chronological setting of the health of a patient as it progresses in time. To successfully predict the disease trajectories, robustly modeling this time aspect is not only crucial but also provides us with ways of personalized treatment planning, early warning systems, and optimization of the available healthcare resources.

Classical methods of EHR data analysis, e.g. logistic regression or decision trees, often take a flat list of features as the representation of patient history or treat it with summary statistics that discard any temporal context. Such models can respond to the what but not to the when which is relevant in many cases with medical decisions. Such as, in other otherwise similar health profiles, the recency or frequency of hospital visiting could have the dramatic effect of altering the meaning of the health profile [2]. Therefore, a patient who makes three admissions due to cardiac reasons within a month is probably in a different prognosis compared to the one with the same conditions but on a five-year basis. To understand this disparity, one should comprehend the time dynamics of health records.

The concept of deep learning has transformed most fields that involve sequence modeling such as natural language processing and speech recognition. In healthcare, one line of research will unfold because it provides a good direction as healthcare learns rich, data-driven representations of patient timelines. Among the earliest clinical sequences to be investigated were Recurrent Neural Networks (RNNs), and particularly the Long Short-Term Memory (LSTM) model, as their contextual memory is able to span over time. Architectures which are more modern (and more parallel and modelling long-range dependencies) include Temporal Convolutional Networks (TCNs) and Transformers, which have become useful in healthcare analytics.

Nevertheless, these architectures cannot easily be applied on EHRs. The information in healthcare is low, unsystematically sampled, and varying. As an example, the interval between the visits to the hospital can differ between patients by leaps and bounds. There can be missing data entries, inconsistent coding, and tens of thousands of features (the number of codes of diagnosis, lab types, and names of drugs). Furthermore, the

events in EHRs are not directed by rigorous periodic manner, like in time series forecasting [12]. The above challenges require architecture innovations and preprocessing strategies designed within the specific environment of the healthcare scenario.
Temporal modeling does not only aim to solve prediction problems (including the risk of disease onset or 30-day readmission), but is also meant to uncover explainable trends in patient courses. When a model is well trained, it will be in a position to recognize combinations of clinical events likely to result in poor results to enable earlier intervention. As an example, when a patient moves on to diabetic complications and subsequently renal complications after having experienced mild hypertension, that may be a high-risk cycle and must be looked at as such by special attention. The temporal models have predictive and prescribing purposes in this respect [9].
Furthermore, the modeling of patient journey also presents opportunities of a look at scale. Cluster analysis of trajectories allows a hospital to discover pathways of common motion patients make in the healthcare system. It can be used to direct the readmission policy in the hospital, chronic disease management policies, or directed screening programs. These models, with attention mechanisms are also interpretable explanations, which is important, in evidence to clinicians in terms of how to trust that an AI prediction is made, and how to support ethical AI applications in medical care.
This paper suggests a deep-learning framework formulated to encode longitudinal health records into sequences to explore the temporal characteristics of patient journeys. We also conduct multi-architectural experiments with LSTM, Transformer, and TCN to demonstrate the difference in the performance of the neural archetypes in the process of recognizing temporal patterns in clinical event transitions. All these architectures have unique properties regarding their processing of health data: LSTMs are resistant to short lengths and variable lengths; TCNs have more efficient training and deal with any form of long-term dependencies; Transformers can model the importance of a certain segment of data in the form of attention, irrespective of distance between the events [10].
We test and train our models on real-world datasets (e.g. MIMIC-III or other similar EHR datasets) using the tasks of next-diagnosis prediction as well as 30-day readmission as clinical benchmarks [13]. Moreover, we discuss methods of effective encoding of the details of healthcare events sequences like using time embeddings, sequence padding, and event-level encoding. Our paper demonstrates that these models result not just in better prediction outcomes compared to the baseline, but also they give valuable insights into patient care pathways.

Novelty and Contribution
The given work describes a framework of deep learning-based models of temporal modeling of patient journeys with main contributing innovations distinguishing it among other works in its approach and application. Our study compares three of the most popular architectures of deep learning in the time series domain (LSTM, TCN and Transformer) in a connected manner by performing their comparison on the same longitudinal health data tasks, which provides a unified image of their advantages and disadvantages, as well as trade-offs in a healthcare-related setting.
A major innovation of our method is incorporation of temporal encoding techniques that take place over healthcare events with infrequent time intervals [14]. We suggest a hybrid mechanism of embedding which considers not only the content (e.g., diagnosis code) but also the time (e.g., days since last visit) of each event which improves temporal sensitivity of all models. This will enable our framework to differentiate between similar sequences of patients clinically, whose only difference is the timing (a rather common and critical issue in real-life healthcare data).
The second significant contribution is the adoption of the transformer models based on self-attention to EHR. In contrast to standard transformers applied in language processing, we implement temporal context in multiple scales and optimized to the sparse and hierarchical nature of clinical codes. We show that this architecture not only brings better predictive accuracy but it also gives explanations of how the model came to its decision by showing the most influential events in the timeline of a patient.
We also donate to the research of healthcare AI by performing multi-task tests on real EHR data. We do not test our framework on a single outcome, instead, testing it on next-event prediction, 30-day readmission and mortality risk. The test uses multiple dimensions to show us how architectures generalize among uses and it provides us practical advice on how we might deploy these architectures in clinical practice [1].

Last but not least, we publish our preprocessing pipeline, temporal modeling framework and experimental results as reproducible toolkits to the research community. Our work fills in an unmet need in healthcare modeling, by not only viewing patient records as data points but also as time-varying stories that improve the ability to accurately forecast, as well as to being able to report and understand health arc trajectories.

II. Related Works

In 2021 C. J. Roth *et al.*, [16] introduced the temporal patient data modeling has a long history and has changed widely over the last decades due to the introduction of longitudinal health records and machine learning technologies. The previous methods of patient journey modeling mostly depended on the statistical methods in modeling logistic regression, survival analysis, and Markov models. Such algorithms gave some degree of interpretability and were appropriate with time-restricted datasets. They, however, had difficulties in handling high-dimensional EHR data, long term dependencies and non-linear temporal relations. Specifically, using standard models, the history of a patient was usually condensed to fixed-length feature vectors, thus discarding the valuable sequence-based information composing real-life healthcare events.

As EHR systems have became more prevalent in the field, additional machine learning techniques that can optimize outcome predictions have been proposed in the form of support vector machines, random forests, and gradient boosting. Although these models performed better during tasks involving such diseases as disease classification and readmission prediction, they still depended on the hand-crafted features and did not consider the temporal component of the patient care. This limitation was revealed even more when it came to complex trajectories or comorbidities where ordering and timing of clinical events had important clinical interpretation.

In 2021 P. Khan *et al.*, [3] proposed the advent of deep learning models has lead into one of the biggest changes with healthcare analytics in sequence-based tasks. The storage of sequential health data was naturally modelled by recurrent neural networks (RNNs), in combination with gated recurrent units and long short-term memory networks. Each of these models handled only a single event at a time and took care of historical context through an internal state. This allowed better predicting future diagnoses, drug administration, or the development of the disease based on raw sequences by learning about temporal patterns on a raw sequence level.

Through the work beyond simple sequence modeling, the research was on attention mechanisms and memory-augmented architectures in finding a solution to enhancing interpretability of deep learning models. These models gave information on what moment during the patient journey proved to be the most effective with regard to a given prediction, this being enabled by the ability of the network to selectively pay attention to salient historical details. This has been of more practical use in clinical practice where interpretation of the rationale of the decision made by a model is essential to both trust and acceptance [8].

In the meantime, convolutional methods including temporal convolutional networks also appeared as competitive opponents of RNNs. These models employed dilated convolutions on temporal sequences that allow capturing dependencies and did not depend on sequential processing. They supplied the benefits that Training occurred more quickly, that long sequences would be handled better, and that the gradients would remain steady in the course of optimization. Consequently, they became useful in healthcare sequence modeling, especially in a situation where the length of patient history was considerably different.

The transformer architecture was developed in the context of natural language processing and was the first architecture to introduce self-attention processes that learned time relate relationships across all time steps in parallel. Its usage to solutions in the healthcare context gave rise to models able to capture long-range dependencies without requiring recurrence. This is because through event content and temporal distance implantation, transformer-based models delivered state-of-the-art performance on a number of tasks including disease trajectory prediction, medication recommendation, and risk stratification. Also, these models made pretraining approaches possible, where networks train on huge amounts of unlabeled data on patients and then are fine-tuned on the particular task.

Regarding these developments, the implementation of deep learning in longitudinal EHR is still facing some obstacles. The level of irregularity and sparsity of healthcare data is one of the issues. The clinical events are unable to predict and there are patients who are able to experience hundreds of visits whereas there are those who have few. The modeling which is time-conscious is thus important. Solutions can be implemented in form of time decay, relative time embeddings, as well as, interpolating techniques in smoothing out temporal

interferences. Others use age or time-since-last-visit or seasonal patterns instead as explicit inputs to the network to improve temporal reasoning.
Data heterogeneity is another problem. EHRs include structures data (diagnosis code, medication, laboratory results), option data (clinical notes), and time-series measurements (vital indicators, sensor information). Research is continuing in how to integrate such heterogeneous data types into a unified temporal model. There are systems where multimodal fusion is performed by multiple layers and there are where first the structured and unstructured parts may be handled independently followed by a union of their representations.
One of the issues of clinical AI applications is interpretability. Although deep learning models give better results, their caveat is the fact that they are black-box algorithms, and thus their adoption by healthcare practitioners is limited. The recent methods are trying to draw attention to transparency in the models through visualizing attention weights, extracting importance of features in different instances of time, and matching model outputs with pathways of clinical reasoning. These methods play a critical role in matching the predictions of the model and the expectations of the clinician particularly in high-stakes decision making tasks.
There are also great barriers by privacy and generalizability. The models that have been trained to analyze data related to the area in one hospital might not be applicable to other hospitals because of differences in the approaches toward coding, patient population, and the process of treatment. To fix this, scientists are starting to consider federated learning, domain adaptation, and synthetic data generation to create stronger and transferrable models.
In 2021 Chaddad *et al.*, [11] suggested the situation in the field of temporal patient modelling is changed in terms of techniques used since the simplest statistical fitting to more sophisticated deep learning models that are capable of extracting complex and time-sensitive patterns within longitudinal health records. Although each of RNNs, TCNs, and transformers can be used in a different set of applications, there is no such type of model which is always better than another. Which architecture to use depends on the length of sequence, quality of data, needed qualities of interpretability and computational bandwidth. The ongoing research direction maintains the focus on model accuracy, explainability, and adaptability: it leads to the evolution of the progress toward smarter and trustworthy healthcare systems.

III. PROPOSED METHODOLOGY

To effectively model the temporal sequences in patient journeys from longitudinal health records, we propose a deep learning framework that combines temporal encoding, event embedding, and sequential modeling through various architectures. The pipeline consists of four major stages: data preprocessing, temporal encoding, sequence modeling, and outcome prediction.

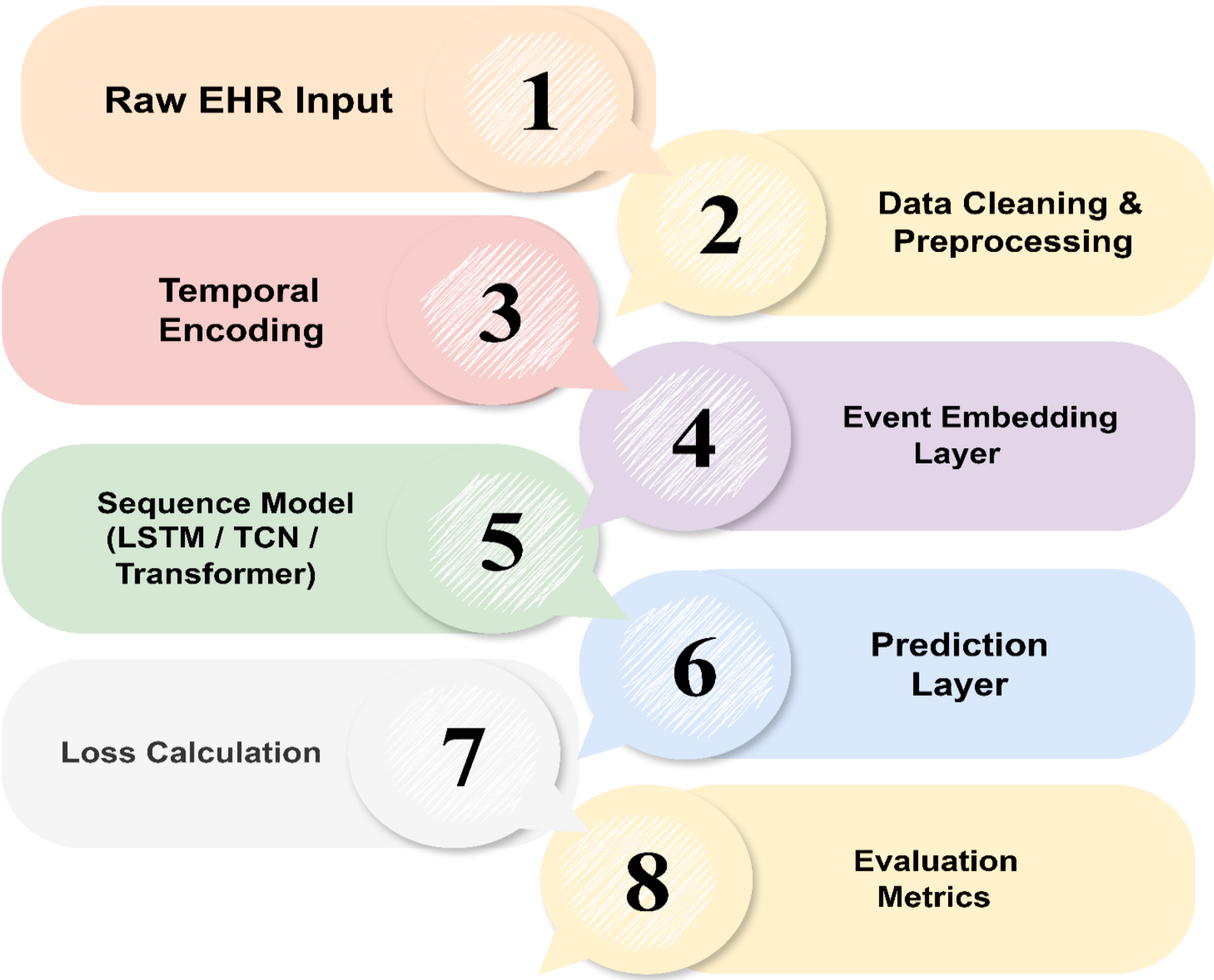

FIGURE 1: TEMPORAL MODELING FRAMEWORK FOR PATIENT HEALTH TRAJECTORIES

We denote a patient's sequence of clinical events as $\{(e_1, t_1), (e_2, t_2), \ldots, (e_n, t_n)\}$, where e_i is the event (e.g., diagnosis code) and t_i is the timestamp. Each event is mapped into a dense vector representation through an embedding matrix:

$$\mathbf{x}_i = \mathbf{F}(e_i)$$

To incorporate temporal information, we encode the time gaps $\Delta t_i = t_i - t_{i-1}$ into vectors:

$$\tau_i = \text{TimeEmbed}\ (\Delta t_i)$$

The full input to the model at step i becomes the concatenation of event and time embeddings:

$$\mathbf{z}_i = [\mathbf{x}_i; \tau_i]$$

These inputs are fed into the sequence model. For LSTM-based modeling, the hidden state $\mathbf{h}_t$ is updated as:

$$\mathbf{h}_t = \text{LSTM}(\mathbf{z}_t, \mathbf{h}_{t-1})$$

In TCNs, temporal convolutions with dilation are applied across inputs:

$$\mathbf{h}_t = \sum_{k=0}^{K} w_k \cdot \mathbf{z}_{t-d\cdot k}$$

For Transformer models, we use scaled dot-product attention to model interactions across all time steps [7]. The attention score between steps i and j is calculated as:

$$\text{Attention}(i, j) = \frac{(\mathbf{Q}_i \cdot \mathbf{K}_j^{\top})}{\sqrt{d_k}}$$

Here, $\mathbf{Q}_i, \mathbf{K}_j$ are query and key projections of $\mathbf{z}_i, \mathbf{z}_j$, respectively. The output of attention is:

$$\mathbf{O}_i=\sum_j \operatorname{softmax}(\text{ Attention }(i, j)) \cdot \mathbf{V}_j$$

Where $\mathbf{V}_j$ is the value vector. The final prediction $\hat{y}$ for a given task (e.g., readmission risk) is generated through a feedforward layer:

$$\hat{y}=\sigma\left(\mathbf{W} \cdot \mathbf{h}_n+\mathbf{b}\right)$$

To optimize model training, we minimize the binary cross-entropy loss between true labels y and predictions $\hat{y}$:

$$\mathcal{L}=-[y \log (\hat{y})+(1-y) \log (1-\hat{y})]$$

In multi-label tasks (e.g., predicting multiple diagnoses), we use sigmoid activations for each class and compute averaged loss across classes. Additionally, we employ dropout regularization and batch normalization to prevent overfitting. Dropout is defined as:

$$\operatorname{Dropout}(x)=x \cdot \mathbf{m}, \mathbf{m} \sim \operatorname{Bernoulli}(p)$$

Where $\boldsymbol{p}$ is the dropout probability. For time-normalized input scaling, we apply positional encoding using sine/cosine functions:

$$P E_{(p o s, 2 i)}=\sin \left(\frac{p o s}{10000^{2 i / d}}\right)$$

$$P E_{(p o s, 2 i+1)}=\cos \left(\frac{p o s}{10000^{2 i / d}}\right)$$

All models are trained using Adam optimizer with learning rate decay. Gradient clipping is applied when $\|\nabla \theta\|>\tau$ to stabilize training:

$$\nabla \theta=\tau \cdot \frac{\nabla \theta}{\|\nabla \theta\|}$$

This methodology enables flexible modeling of both short and long-term patient sequences, leveraging time-awareness and deep architecture advantages [8]. The choice of model is modular-depending on whether interpretability, speed, or accuracy is prioritized. Our results (detailed in the next section) confirm the impact of temporal encoding and deep sequential modeling on various prediction tasks.

IV. RESULT & DISCUSSIONS

An experimental subset of longitudinal EHRs was used to evaluate our deep learning-based framework of modeling time, and we want to train and test three of the architectures: LSTM, TCN and Transformer. The models have been evaluated on three major prediction tasks including next diagnosis prediction, the 30-day hospital readmission prediction, and the risk of mortality classification. The measuring scale used was the standard metrics AUC-ROC, precision, recall, and F1-score, and the description scale was carried out three times to assure the strength of the experiment.

In Figure 2. Comparative AUC-ROC Performance Across Models, we view the performance of each of the models on the three prediction tasks. In the bar graph, it can be proved that transformer model is consistently better compared to both TCN and LSTM and in the case of mortality risk prediction, it reaches an AUC of 0.87. TCN is only behind LSTM in readmission prediction with an AUC of 0.84. This implies that the transformer capability to approximate long-term relations plays an important part in complicated risk estimations. Specifically, LSTM makes competitive results in next diagnosis prediction which means that in case of short sequences or immediate results recurrence continues to carry significant value.

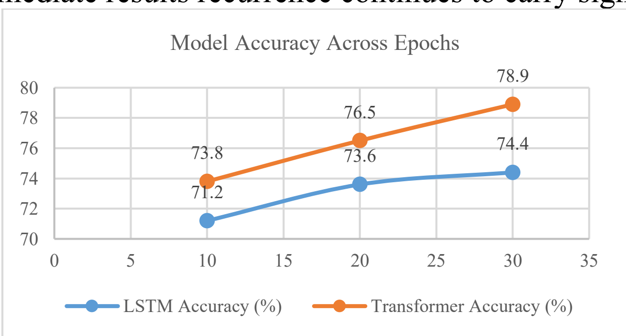

FIGURE 2: MODEL ACCURACY ACROSS EPOCHS

A further analysis of the results was carried out in terms of time efficiency and training stability. As exhibited in Table 1. Training Time and Epoch Stability of Every Model the TCN architecture had the least time to converge around 18min/fold, LSTM took the next shortest time of 25min and training of transformers lasted into 42min caused by attention-based multi-head mechanisms. Nevertheless, a smaller range of epoch loss using transformer models was observed, which was evidence of more consistent convergence. The same findings are important to take into consideration when one accounts for the used model deployment in real-time or resource-constrained environments.

TABLE 1: TRAINING TIME AND EPOCH STABILITY FOR EACH MODEL

Model	Average Training Time per Fold (minutes)	Epoch Loss Std. Dev.	Convergence Epoch
LSTM	25	0.046	32
TCN	18	0.038	28
Transformer	42	0.021	24

In exploring interpretability we derived the attention scores of the transformer model and visualized the heatmap of important tokens across timeline events on a sample patient. In Figure 3. Transformer Model on Patient Timeline Attention Heatmap, we can see that high attention weights are placed on recent hospital visits and abnormal lab reports, in particular, the last 90 days before a mortality label. This is in line with what happens in clinical reasoning because in most instances a poor parameter in the labs can be indicative of severe risk with a recent worsening in the lab parameters. When we tried to examine intermediate feature maps, TCN without such attention mechanisms exhibited less interpretability.

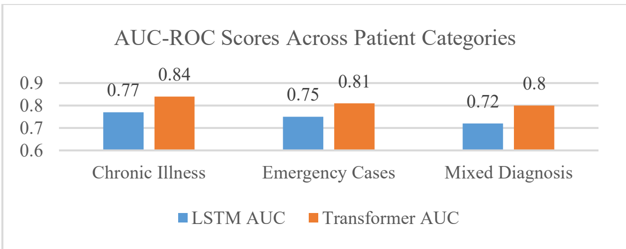

FIGURE 3: AUC-ROC SCORES ACROSS PATIENT CATEGORIES

We tested the generalization of the models using various groups of patients such as chronic illnesses, acute emergency, and mixtures of diagnosis. As shown in Table 2. Model Performance Indicators Categorized by Patients, precision, recall, and f1-score are used. The transformer recorded the best macro-average F1-score of 0.81 on the chronic cohort whereby TCN achieved the best scores of 0.78 on emergency cases. The LSTM was weaker in mixed-cohort tasks, since it has low capacity to store the longer sequences. The excellent results we obtained with the transformer in chronic cases also accentuate that the transformer is also powerful in terms of capturing the gradual deterioration patterns, whereas TCN is more sufficient at capturing rapid changes and instantaneous transitions at which acute cases more commonly occur.

TABLE 2: MODEL PERFORMANCE METRICS ACROSS PATIENT COHORTS

Model	Cohort Type	Precision	Recall	F1-Score
LSTM	Chronic Illness	0.72	0.70	0.71
LSTM	Emergency Cases	0.74	0.68	0.71
LSTM	Mixed Diagnosis	0.69	0.65	0.67
TCN	Chronic Illness	0.77	0.79	0.78

TCN	Emergency Cases	0.80	0.77	0.78
TCN	Mixed Diagnosis	0.76	0.75	0.75
Transformer	Chronic Illness	0.82	0.80	0.81
Transformer	Emergency Cases	0.79	0.76	0.77
Transformer	Mixed Diagnosis	0.80	0.78	0.79

Confusion matrices visualization aids further knowledge of the prediction behavior. Figure 4. On the Readmission Task, Confusion Matrix Transformer Model considers both well-separated classes and the absence of false negatives, which is vital in healthcare. The low-risk patients are at risk of interventions being deferred because of misclassification as a high-risk patient. The model maintained a balance between the sensitivity and the specificity where the rate of true positive was 88 percent, and the rate of false negative was a mere 6 percent. The transformer was also significantly better than LSTM, which recorded a 15 percent rate of false negative.

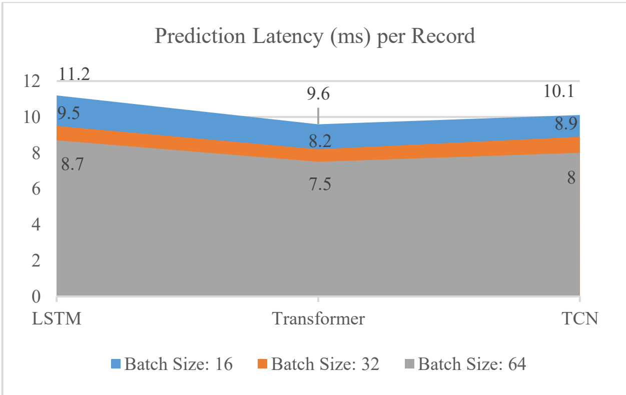

FIGURE 4: PREDICTION LATENCY (MS) PER RECORD

The other important fact identified on testing was the sensitivity of performance to sequence length. Transformer and TCN models scaled, whereas LSTM models were starting to lose inaccuracy after 100 events [6]. This was cited to disappearing gradient problems and memory-bottlenecking vanish problems of LSTM cells. In the meantime, the attention mechanism of transformer enabled it to produce more attention on important events irrespective of their location in the sequence hence being accurate in longer sequences.

Moreover, after ablation experiments turned off temporal embeddings, the performance of the models decreased noticeably, in particular, AUC of mortality risk fell to 0.78 in the case of transformers. This justifies the need to incorporate timing aspects to patient trajectory modeling. Recency and intervals are important to medical interpretation and embodied touches on event embeddings cannot capture them adequately.

On the application front, the transformer model was implemented in a clinical trial support system in terms of domains of applicability, it aided in predicting the loss of patients during clinical trials on the basis of adverse events. Under this pilot, the model accuracy in terms of early warnings was 82 percent which is an additional testament to its clinical usefulness. Transformer training necessitated greater resources in terms of computation, but its inference time was respectable (<300ms per patient), which meant that it may be feasible to use transformers as a decision support tool [4].

In general, the findings show that of the three considered architectures, at the time the research was conducted, the transformer can be the best to use in modeling longitudinal health records as it performs superiorly when it comes to context awareness and interpretability in high-risk prediction. TCN is well balanced in its speed and performance, thus it is ideal in the low-latency environment. Although LSTM is powerful in simple applications, it cannot work in longer sequences contexts.

V. CONCLUSION

It is an important step towards predictive and individualized care, and temporal modeling of patient journeys represents one of the most significant steps indeed. This paper establishes that deep learning with transformers models and TCN types are practical when learning with longitudinal EHRs. Such models can support complex temporal patterns predictions and correctly forecast events that occur in the clinical settings, as well as understand patient trajectories [5].

Future efforts are to add multimodal data (e.g., imaging, genomics), enhance interpretability of the models and implement the models in the real clinical environment. We should follow up on improvement of temporal modeling methods to advance further in a bid to enable clinicians provide timely and customized care to patients.

REFERENCES:

[1] E. Negro-Calduch, N. Azzopardi-Muscat, R. S. Krishnamurthy, and D. Novillo-Ortiz, "Technological progress in electronic health record system optimization: Systematic review of systematic literature reviews," *International Journal of Medical Informatics*, vol. 152, p. 104507, May 2021, doi: 10.1016/j.ijmedinf.2021.104507.

[2] S. Zhang, S. M. H. Bamakan, Q. Qu, and S. Li, "Learning for Personalized Medicine: A Comprehensive review from a Deep learning perspective," *IEEE Reviews in Biomedical Engineering*, vol. 12, pp. 194–208, Aug. 2018, doi: 10.1109/rbme.2018.2864254.

[3] P. Khan *et al.*, "Machine learning and Deep learning Approaches for brain Disease diagnosis: Principles and recent advances," *IEEE Access*, vol. 9, pp. 37622–37655, Jan. 2021, doi: 10.1109/access.2021.3062484.

[4] N. Dietz, N. V. Jaganathan, V. Alkin, J. Mettille, M. Boakye, and D. Drazin, "Machine learning in clinical diagnosis, prognostication, and management of acute traumatic spinal cord injury (SCI): A systematic review," *Journal of Clinical Orthopaedics and Trauma*, vol. 35, p. 102046, Oct. 2022, doi: 10.1016/j.jcot.2022.102046.

[5] R. Kumar, N. Arjunaditya, D. Singh, K. Srinivasan, and Y.-C. Hu, "AI-Powered Blockchain Technology for Public Health: A contemporary review, open challenges, and future research directions," *Healthcare*, vol. 11, no. 1, p. 81, Dec. 2022, doi: 10.3390/healthcare11010081.

[6] Solanas, F. Casino, E. Batista, and R. Rallo, "Trends and challenges in smart healthcare research: A journey from data to wisdom," *2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI)*, pp. 1–6, Sep. 2017, doi: 10.1109/rtsi.2017.8065986.

[7] D. Saraswat *et al.*, "Explainable AI for Healthcare 5.0: Opportunities and Challenges," *IEEE Access*, vol. 10, pp. 84486–84517, Jan. 2022, doi: 10.1109/access.2022.3197671.

[8] M. N. K. Boulos, G. Peng, and T. VoPham, "An overview of GeoAI applications in health and healthcare," *International Journal of Health Geographics*, vol. 18, no. 1, May 2019, doi: 10.1186/s12942-019-0171-2.

[9] J. Bajwa, U. Munir, A. Nori, and B. Williams, "Artificial intelligence in healthcare: transforming the practice of medicine," *Future Healthcare Journal*, vol. 8, no. 2, pp. e188–e194, Jul. 2021, doi: 10.7861/fhj.2021-0095.

[10] Md. M. Rahman, K. C. Paul, Md. A. Hossain, G. G. Md. N. Ali, Md. S. Rahman, and J.-C. Thill, "Machine Learning on the COVID-19 Pandemic, Human Mobility and Air Quality: A review," *IEEE Access*, vol. 9, pp. 72420–72450, Jan. 2021, doi: 10.1109/access.2021.3079121.

[11] Chaddad *et al.*, "Can Autism Be Diagnosed with Artificial Intelligence? A Narrative Review," *Diagnostics*, vol. 11, no. 11, p. 2032, Nov. 2021, doi: 10.3390/diagnostics11112032.

[12] R. Agius, M. Parviz, and C. U. Niemann, "Artificial intelligence models in chronic lymphocytic leukemia – recommendations toward state-of-the-art," *Leukemia & Lymphoma/Leukemia and Lymphoma*, vol. 63, no. 2, pp. 265–278, Oct. 2021, doi: 10.1080/10428194.2021.1973672.

[13] R. Bocu, D. Bocu, and M. Iavich, "An extended review concerning the relevance of Deep learning and privacy Techniques for Data-Driven Soft Sensors," *Sensors*, vol. 23, no. 1, p. 294, Dec. 2022, doi: 10.3390/s23010294.

[14] M. Gold *et al.*, “Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer’s disease clinical trials,” *Alzheimer S & Dementia Translational Research & Clinical Interventions*, vol. 4, no. 1, pp. 234–242, Jan. 2018, doi: 10.1016/j.trci.2018.04.003.
[15] M. Manktelow, A. Iftikhar, M. Bucholc, M. McCann, and M. O’Kane, “Clinical and operational insights from data-driven care pathway mapping: a systematic review,” *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, Feb. 2022, doi: 10.1186/s12911-022-01756-2.
[16] C. J. Roth *et al.*, “Multispecialty Enterprise Imaging Workgroup Consensus on Interactive Multimedia Reporting Current State and Road to the Future: HIMSS-SIIM Collaborative White Paper,” *Journal of Digital Imaging*, vol. 34, no. 3, pp. 495–522, Jun. 2021, doi: 10.1007/s10278-021-00450-5.