Predicting clinical outcomes using machine learning algorithm through patient health records

Preethi T
Assistant professor /CSE, KPR Institute of Engineering and Technology, Coimbatore – India, preethiithangavel@gmail.com

Naveen E
Department of CSE, KPR Institute of Engineering and Technology, Coimbatore – India, naveen7904074513@gmail.com

Karan A K
Department of CSE, KPR Institute of Engineering and Technology, Coimbatore – India, karandurai1212@gmail.com

Karthik Raja G
Department of CSE, KPR Institute of Engineering and Technology, Coimbatore – India, kr278504@gmail.com

Abstract

Clinical independent direction is currently molded by data driven machines, and by their assumptions or proposition [1]. Different AI a recent clinical investigation has found application composition, especially for result assumption models, with results going from mortality and cardiovascular breakdown to serious. Mimicked insight procedures are proper to expect clinical outcomes [2]. Before long, AI methods can be viewed as limits that gain capability with the outcomes going with standardized input data to make exact outcome gauges when tried with it can be integrated within existing clinical cycles. For new data. A promising assessment and displaying the utility of patient-uncovered outcome estimates data for developmental assessment, tweaked treatment and precision drug with the help of AI based decision genuinely steady organizations. In this assignment, we summarize the top tier in related works covering data taking care of, derivation, and model appraisal, with respect to result assumption models made using data isolated from electronic prosperity records [3]. We also talk about limitations of unquestionable showing doubts and element important entryways for future assessment.

Keywords: *Machine learning, Clinical outcomes.*

I. Introduction

Late man-made thinking (AI) enhancements look to positively influence drug and clinical practice. Simulated intelligence (ML), another application of AI, utilizes large amounts of clinical data in order to create future measures, resulting in varied NLP, CV, and customized speech recognition have all had successful applications. Employments of ML have been powerful across a couple of clinical regions, for as well as clinical outcome assumption instance, disease estimate using different data modalities, including talk signs and clinical imaging to distinguish rot, similar to cardiovascular breakdown, mortality, or crisis unit affirmation. The aim of the paper is to provide a development investigation of late arrangements with clinical outcome gauge models that address the different areas where they are envisioned.

When in doubt, ML structures can be setup through a multidisciplinary effort that begins with data planning and reaches as far as developing and evaluating a judicious model [4]. Care pathways inside clinical centers shift for the most part as a result of the assortment of yielded patients. Hence, a clinical understanding is essential to develop AI models that treatment option., whereas the last option usually stipulates a standard procedure.

Care pathways inside crisis facilities change for the most part on account of the assortment of yielded patients. As a result, knowledge of the clinical setting is crucial to creating AI models that can be integrated into existing clinical trials. A patient may be admitted to the hospital as an emergency or as an elective procedure; the latter option includes a standard system [5].The purpose of hospitalization is to gather different types of data from the patient with the ultimate objective of noticing. The use of data types such as imaging, talk, and claims can help develop result assumption models. Here, we base our analysis on data obtained from electronic prosperity records (EHR), which are active and sent in crisis facilities all over the world. In clinical centers, EHRs are used to store patients' longitudinal data. Patient economics, significant physical processes, solutions, laboratory data, and any outcomes that might have happened during or soon after a hospitalisation are all included in this data. Machine learning models can be developed and tested using her informative index data. It is regularly split into an arrangement set and a test set1, either by an inconsistent or a nonrandom split taking into testing sets, as it avoids sporadic distributions between them. In model learning, the readiness set is used to overhaul the limits of the model. A pre- arranged model is then evaluated using different performance metrics on the held-out test.

II. CLINICAL OUTCOME PREDICTION

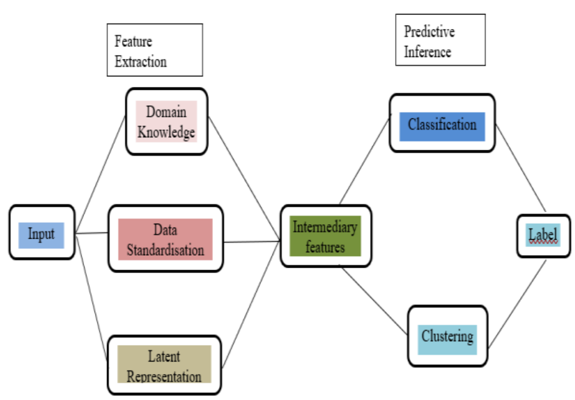
Clinical decision- production in medical care is as of now being impacted by expectations or suggestions made by information driven machines .In the most recent clinical writing, various AI applications have emerged, particularly for outcome forecast models, with results ranging from mortality and coronary failure to severe renal injury and arrhythmia. This exam article summarizes the best- in-class related work dealing with information processing, derivation, and model evaluation in relation to result-expected models built using information removed from electronic medical records [6]. We additionally talk about restrictions of unmistakable demonstrating presumptions and feature open doors for future examination. Patient observing devices, like early admonition frameworks , are inescapable across various medical clinic wards to persistently survey for patient decay. The meaning of what precisely comprises clinical crumbling has developed over the long run in view of the information assortment and handling procedures. Early endeavors to characterize clinical disintegration zeroed in on clinical disregard and its outcome of clinical complexities. Resulting concentrates on zeroed in on more discrete clinical occasions, like serious sepsis, unforeseen heart failure, ICU confirmation or mortality , and will generally choose at least one end-point proportions of clinical disintegration. Such situations have significant financial consequences associated with extended clinic visits, legal fees, staff time, influence upon patients and staff, and additional expenses. Since it enables analysts to divide patients into separate categories, such as crumbling and non-crumbling, and infers the y names implicitly, the final definition is the most well-known.

The structure of result expectation models additionally fluctuates across the writing. A few investigations anticipate the danger of a result just once utilizing the patient's first N long periods of information after confirmation, like 24 or 48 hours . Others decide to anticipate the danger of a result, for example, ICU readmission, utilizing the patient's keep going N long stretches of information before release [7]. Another normal strategy is to foster an ongoing alarming score, which processes the danger of weakening whenever a bunch of clinical perceptions is gathered as in clinical early advance notice frameworks .

A. FEATURE EXTRACTION

In AI, plan affirmation, and picture taking care of, component extraction begins by taking a look at a basic course of action of assessed data and estimating inferred values (features) to be informative and non-abundant, using the resulting learning and theory and influencing better human comprehensions. Dimensionality reduction occurs when features are extracted. When the data a computation is overly expansive to possibly be handled and is anticipated to be a lot, then it can be reduced to a smaller number of components (also called a part vector). Incorporating a subset of features is called incorporating. The picked features are expected to contain the vital information from the data, so it is ideal to utilize this reduced representation rather than the absolute basic data. Extracting features reduces the amount of resources needed to display a massive set of data [8]. An issue that arises when examining staggering data is the number of variables involved. It often takes large amounts of memory and estimation power to examine Incalculable variables, in addition to making a gathering computation over planning tests and summarizing insufficiently to new models. The term feature extraction refers to methods of blending variables together in order to deal with these issues and still depict the data with adequate accuracy. Numerous AI consultants acknowledge that befittingly improved half extraction is that the thanks to possible model flip of events [9]. varied knowledge assessment programming packs oblige feature extraction and viewpoint decline. Typical numerical programming conditions, for example, MATLAB, and also the R language provides a piece of the a lot of clear half extraction techniques (for instance head half assessment) through bound orders. a lot of categorical estimations routinely obtainable as squarely open substance or outcast additional things. There are programming teams zeroing in on categorical programming AI applications that address spectacular skilled in incorporate extraction.

Fig 1. Feature Extraction and predictive interface



B. PREDICTIVE INTERFACE

The differentiation between the settings of revelation and support is because of the intelligent rationale. Popper broadly announced "The underlying stage, the demonstration of imagining or concocting a hypothesis, appears to me to neither call for coherent investigation nor to be helpless of it". That this assertion happens in a book called The Logic of Scientific Discovery has flabbergasted numerous persons. The peculiarity can be to some degree eased. 'Find' is a triumph word. One can't find that the moon is made of green cheddar, since it isn't. To find that p one should think of the speculation that p, or surmise that p, and afterward show that p is valid. It is steady to keep up with that the underlying 'speculating' stage isn't defenseless of legitimate investigation, and that rationale just assumes a part in the subsequent stage, where we show that p is valid, demonstrate it or legitimize it [10]. This just to some extent frees the peculiarity from Popper's case, since he broadly asserts that there is no demonstrating or supporting our speculations by the same token. All he gives us in The Logic of Scientific Discovery is a legitimate examination of the course of experimental testing. Since 'find' is a triumph word, it is odd to discuss finding a bogus speculation. It would be smarter to talk, not of the setting of disclosure, but rather of the setting of development. Then, at that point, we can isolate the topic of creating a theory from the subject of legitimizing it. Be that as it may, 'setting of defense' is certifiably not a blissful expression either, essentially for Popper's situation. He feels that while researchers can normally assess or evaluate theories, they can never legitimize or demonstrate them. So as not to make one wonder against that view, it would be smarter to discuss the setting of examination[11]. These phrased ideas are because of Robert McLaughlin .Were the positivists and Popper right that there is no rationale of creation, no consistent investigation of the underlying phase of developing a speculation? No. Individuals don't normally create theories indiscriminately or through blazes of otherworldly instinct or in their fantasies. Individuals normally create new theories by reason or contention. However, the unavoidable idea is, these thoughts or contentions can't be insightful, for the decision of a substantial derivation contains the same old thing. Subsequently we really want an inductive or ampliative rationale of creation (disclosure). They viewed the 'setting of disclosure' as having a place with the region of brain research than positivists and Popper. They were doubtful with regards to there being any rationale of disclosure.

III. MODULAR DESCRIPTION

A. VM SETUP AND EVALUATION MODULE

In this Module Treatment Evaluation is the assessment of virtual therapy times in the planning process. The assessment depends on the actual utilization of the patient in clinic. Each virtual the cloud is figured and sent to the client for further processing. The Hospital Queuing- Recommendation (HQR) framework is used to determine a productive and helpful treatment plan for patients[12]. The PTTP calculation and HQR framework are highly proficient and low idleness due to their enormous scope, practical dataset and requirement for constant reaction.

B. CALCULATE NEW FEATURE VARIABLES OF THE DATA

Different critical features of the data are not predefined when setting up the PTTP model, such as the patient time used for each treatment, day of week for the treatment, and the length of the treatment. The dimensionality decline process, which isolates and condenses unrefined data to more logical social events, includes feature extraction. Therefore, it will be more direct when you truly want to control [13]. The fact that these amazing illuminating records have countless variables is their main quality. To address these variables, a lot of figuring resources are needed.

C. REGULARIZED WORKFLOW STATES

By and by, the distinguished work process states ought to be spatially coterminous. For instance, we need to distinguish areas of semantically significant areas, for example, 'second floor upper east quiet rooms' and 'storm cellar focal extra spaces'. In this part, to urge every work process state to be an adjoining locale in the structure, we utilize the vicinity between rooms to characterize an earlier on the work process state dissemination.

D. WORK FLOW SCHEDULING

In the forecast of the work process booking the idea for the result can be clinical result expectation through the PTTP calculation gives the better exactness and better result[14].workflow task planning implies that the work process errands put together by clients are apportioned to fitting figuring assets for execution, and the relating charges are paid progressively as indicated by the utilization of assets. For most normal clients, they are primarily worried about the two help quality signs of work process task fulfillment time and execution cost.

E. EQUATION

Similarly, in Random Forest, we train a number of decision trees, and the class with the most votes becomes the final result if the problem is a classification.

$$Gini\ Index = 1 - \sum_{i=1}^{n}\left(P_i\right)^2$$
$$= 1 - [(P_+)^2 + (P_-)^2]$$

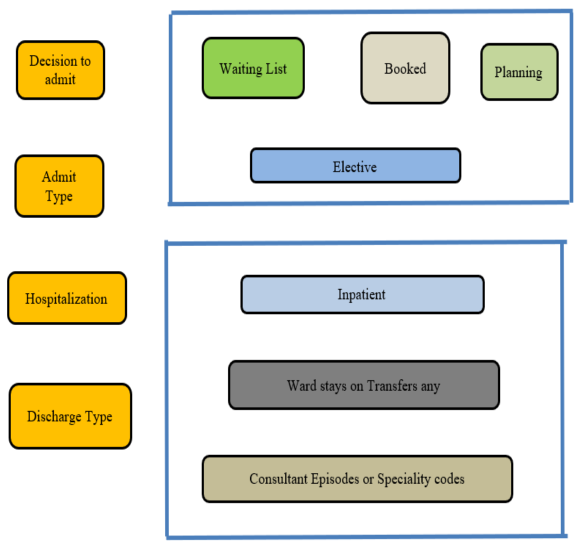
Decision trees display the predictions that result from a series of feature-based splits using a flowchart similar to a tree structure. It begins with a root node and the gini index ends with a leaf decision.

IV. MACHINE LEARNING

Man-made intelligence (ML) is an investigation of artificially created computations that deal with issues through experience and by utilizing data. It is a piece of fake knowledge. An AI model produces conjectures or decisions based on planning data and test data without explicitly changing the model. For instance, machine learning computations are used for drug identification, email isolating, talk verification, and PC vision, when it is irksome or illogical to encourage standard computations to perform the required functions. AI is unquestionably a subset of computational learning, which involves making assumptions with PCs; however, not all AI is quantifiable. A mathematical improvement examination conveys methodology, speculation, and application areas to the field of artificial intelligence[15]. In data mining, performance learning is used to explore data through exploratory data examination. In a couple of cases, artificial intelligence uses data and neural connections in a way that resembles the working of a particular psyche. AI is also called perception analysis in its application to business issues.

Artificial Intelligence emerged out of the quest for human awareness. Early on in the development of AI as an insightful discipline, a few execs were passionate about the idea of making machines learn from data. In the process, the authors overcame various obstacles, such as "neural associations"; perceptron's and other models were used that were seen, in this manner, to be reconsiderations of the summarized straight models of estimations. Probabilistic reasoning was also used, including in electronic clinical diagnosis. However, an extension of the lucid, data-based procedure caused a break between AI and AI. Data acquisition and depiction issues plagued probabilistic systems. By 1980, ace systems had surpassed AI, and estimations were no longer needed. Significant data set learning went inside artificial intelligence, provoking inductive reasoning programming, anyway the real line of study was as of now outside the field of genuine artificial intelligence, in plan affirmation and information recovery. Approximately around the same time, AI and programming had abandoned neural association research. This was labelled, as "connectionism", by experts from various fields, such as Hopfield, Rumelhart, and Hinton. In the 1980s, the back expansion was re-evaluated. A new field, artificial intelligence (ML), developed during the 1990s. The field went from being geared towards mechanized thinking to dealing with doable issues of a helpful type. As the focus shifted away from the important procedures it had gained from AI, it developed techniques and models based on estimations and probability speculations.

Fig 2. Visualization of patient Flow Hospital



Most of the time, the distinction between ML and AI is misunderstood. In ML, a prediction is made based on idle discernments, but AI proposes a teamwork with the environment to learn and take action to increase the chances of success. According to many sources, ML remains a subfield of AI as of 2020. Others think that not all ML can be considered AI, but a 'savvy subset' of it can.

V. RELATED WORK

The current system which uses the isn't exact in distinctive the so the outcome will be going on like the picture of clinical based the emergency of the things. Records of patients could include discrete obvious codes, like finding, medication, or treatment codes.

Several analyses advise using embedding methodology to derive from such components. This methodology is derived from distributional hypotheses. According to the distributional hypothesis, words that appear in similar situations throughout sizable samples of linguistic data have semantically connected presumptions with a satisfactory outcome.

A new paper by and colleagues explains how peripheral uttermost point disease (PAD) affects countless people. Substantial limb ischemia (CLI) is present, and this condition is associated with significant cardiovascular events and a high mortality risk. One year after being diagnosed with CLI, 30% of patients experience the evacuation of an extremity, while 25% die. The majority of CLI patients go through removal without vascular assessment in the preceding year, regardless of the openness of the state- of-the-art revascularization philosophy. Diabetics patients and its inevitable occurrence Diabetes is heavily pushed as a tool for quick robotic identification of patients and as a strategy to undermine the concept of patient cognition. The computerised ascertainment of CLI has undergone testing, even if not all CLI codes in EHRs appear to be authentic. Although not all CLI codes in EHRs appear to be legitimate, the electronic ascertainment of CLI has exhibited testing. Consequently, prior examinations have outlined and endorsed combinations of ICD-9 codes for ascertaining CLI cases. The focus has shifted from charging code computations to setting. The clinical investigation of CLI relies on the presence of signs and secondary effects that are documented in the clinical record while the billing codes serve primarily as administrative codes.

In this paper With tremendous data improvement in biomedical and clinical consideration organizations, exact between examination of clinical data helps early disease acknowledgment, patient thought, and neighborhood. Regardless, the assessment precision is diminished when the idea of clinical data is inadequate. Likewise, different regions show phenomenal traits of explicit regional ailments, which could weaken the assumption for infection eruptions. In this paper, we smooth out AI estimations for strong conjecture of progressing contamination episode in disorder consistent organizations. The insufficient data is overcome with the use of an inert component model. Our study examines a localized, steady disease of cerebral corruption. A new convolutional neural association (CNN)-based method for estimating multimodal disease risk assumptions is presented based on coordinated and unstructured data collected from clinical facilities. The current work in the field of clinical enormous data examination does not focus on these two data types. With the differentiated and normal figure estimations, the assumption accuracy of our proposed computation is 94.8%, which is faster than the CNN-based uni- modal contamination danger assumption computation. An analysis by McKinsey found that half of Americans suffer fromat least one industrial disease, and 80% of American clinical thought charge goes to steady contamination therapy. As the assumptions for regular solaces improve, the event of consistent illness is growing. United States has consistently spent 2.7 trillion dollars on disease treatment. Approximately 18% of the U.S. yearly GDP is contained in this aggregate. Furthermore, clinical consideration of industrial sicknesses is essential in various countries.

As shown by a Chinese report on food and tireless ailments in 2015, 86.6% of death is caused by impending contamination risk factor for CLI, so the number of CLI patients is expected to rise in both developing and developed countries. Countries CLI is also associated with basic clinical benefits and resource utilization. Averaging $4.2 In terms of absolute costs, billions of dollars are spent each year, CLI hospitalizations cost the US public $4.2 billion in 2013- 2014, but the 30-day readmission rates cost $ 624 million in clinical benefits. EHRs have Similarly, performing danger assessments for continuous diseases is crucial. As clinical data improves, gathering electronic prosperity records (EHR) becomes increasingly valuable.

The Alberta Stroke Program Early CT Score (ASPECTS)is an acronym for Alberta Stroke Program Early CT Score. most generally utilized neuro imaging biomarker with intracranial gigantic vessel obstruction, which monitors ischemic changes at 10 express areas inside the district of the middle cerebral course. From the get go, the scoring was applied to figured tomography (CT) pictures and later to diffusion weighted pictures (DWI). It has been represented that ASPECTS scores relate with clinical outcome, and ASPECTS is by and large used in clinical practice for decision making.1,2 ASPECTS scores can be evaluated actually and quickly by the visual examination of psyche pictures, yet they can be reliant upon alterability between investigators. Profoundly. A trademark development was to survey the particular volume itself. Movements in computation and programming have allowed the ischemic focus volume to be resolved normally as shown by a specific breaking point, for instance, an unmistakable spread coefficient worth of 600 to 620 mm2/s or a CT.

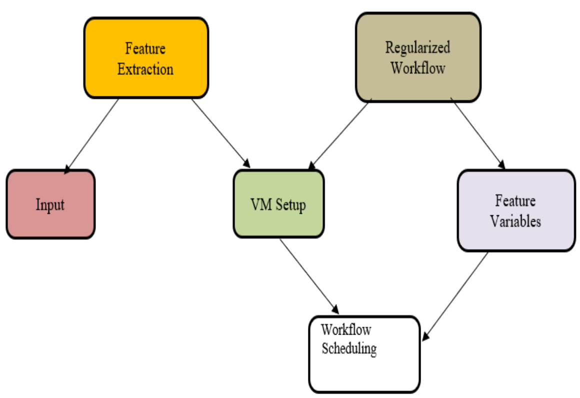
VI. PROPOSED METHOD

This task in issue definition will distinguish the patient circumstances in view of the states of the patient prerequisite. The patient needs to get in to part a first then the proposal will get the space a then the opening

B. Dealing with the patient through the treatment flow. Patient space and time in the line can be distributed to make the treatment time all the more successfully. Keeping away from the undesirable time can be killed all the more successfully through the Patient Treatment Time Scheduling[16]. The distribution for the schedule opening allocation can be made a lot of efficient. In forecasting patient holding up time, the PTTP (Patient Treatment Time Prediction) model is the most effective. perfusion regard. Significantly. These directbiomarkers, which are routinely used in clinical practice, give regards that have been represented to beconnected with clinical outcomes. Significantly, its homogeneity, and sign strength, may add to more correct figures of clinical outcomes.

Regardless, making an assumption model taking into account topological and morphological imaging features would require complex showing considering high- layered data, which is tolerably difficult to achieve with the standard real approach.

Fig 3. Proposed Workflow



VII. RESULT AND DISCUSSION

Features are extracted as part of the dimensionality decline process, and a movement of experiments is conducted to determine whether the proposed computation is accurate and efficient We use CT to evaluate patient treatment time and patient characteristics, which is part of the dimensionality decline process[17]. This part of the dimensionality decline process involves assessing the patient treatment time use of the CT really examine task. Based on the activities and different circumstances, the patient treatment time use of the CT can be calculated [18]. Feature extraction is a step in the dimension decline process, however the improvement Long-term multidisciplinary interactions between ML scientists and physicians are essential to the field's progress. While the two players must talk the same language in order for the process to work, we believe the review will offer both a broad overview and a conclusion as to what the overall ML pipeline entails. Dimensionality decline is the process by which Organizational-factors fuse crisis facility, time, staff ratio, or staff planning.

Fig 4. Dataset

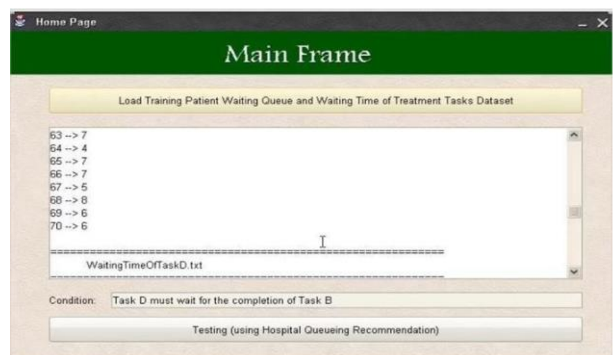


Fig 5. Main Frame

Fig 6. Loading Test Dataset

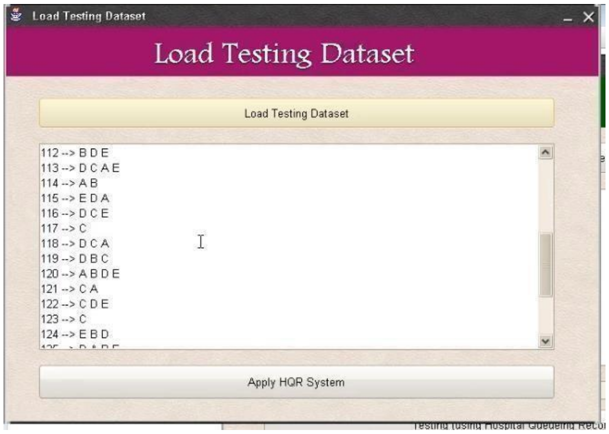
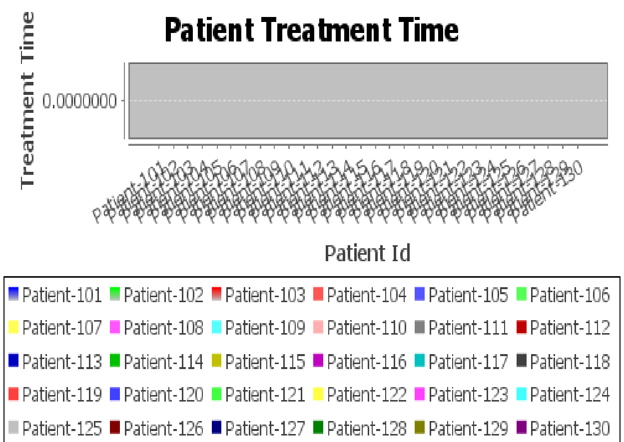


Fig 7. Treatment Time



VIII. CONCLUSION

The dimensionality decline process is a piece of feature extraction that involves neural association taking care of [19]. A piece of dimensionality decline is the feature extraction process, in which models try to discover a way to anticipate a specific outcome or task, which can generally be inferred as 'limit AI'. The dimensionality decrease process, which incorporates learning for downstream applications driving learning for general patient representations, includes this approach as one of its steps. To construct summary models, which might integrate various HER data to execute scheduled activities concurrently, more work needs to be done [20]. Comparable to illness assurance or patient presumption.

References

1. NaveedAfzal, Vishnu Priya, SunghwanSohn, Hongfang Liu, Rajeev Chaudhry, Christopher G Scott, Iftikhar J Kullo, and Adelaide M Arruda-olson. Normal language handling of clinical notes for distinguishing proof of basic appendage ischemia. Worldwide Journal of Medical Informatics, 111:83-89, 2018

2. Min Chen, YixueHao, Kai Hwang, Lin Wang, and Lu Wang.Disease Prediction by Machine Learning over Big Data from Healthcare Communities. IEEE Access, 2017

3. Hidehisa Nishi, Naoya Oishi, Akira Ishii, Isao Ono, Takenori Ogura, Tadashi Sunohara, Hideo Chihara, RyuFukumitsu, Masakazu Okawa, Norikazu Yamana, Hirotoshi Imamura, Nobutake Sadamasa, Taketo Hatano, Ichiro Nakahara, Nobuyuki Sakai, and Susumu Miyamoto. Profound Learning Derived High- Level Neuro imaging Features Predict Clinical Outcomes for Large Vessel Occlusion. Stroke, 2020

4. Nenad Tomaev, Xavier Glorot, Jack W. Rae, Michal Zielinski, Harry Askham, Andre Saraiva, Anne Mottram, Clemens Meyer, SumanRavuri, Ivan Protsyuk, Alistair Connell, Can O. Hughes, Alan Karthikesalingam, Julien Cornebise, Hugh Montgomery, Geraint Rees, Chris Laing, A clinically relevant approach to dealing with the constant threat of severe kidney injury in the future.Nature, 572(7767), pp.116-119 (2019).

5. Sanjay Purushotham, ChuizhengMeng, ZhengpingChe, andYan Liu. Benchmark of Deep Learning Models on Large Healthcare MIMIC Datasets. 2017.

6. Tingting Zhu, Pulkit Sharma, Peter J.

Watkinson, and David A. Clifton. Farah E. Shamout, Tingting Zhu, Peter J. Watkinson, and David A. Clifton. Deep Interpretable Early Warning System for Clinical Deterioration Detection 2019. IEEE Journal of Biomedical and Health Informatics.

7. Jonathan Rubin, CristhianPotes, Minnan Xu- Wilson, JunziDong, AsifRahman, Hiep Nguyen, and David Moromisato. An troupe supporting model for anticipating move to the pediatric emergency unit. Global Journal of Medical Informatics, 112:15- 20, 2018.

8. Alvin Rajkumar, and others. Scalable and precise deep learning for electronic health records 1(1):1- 10, Nature Digital Medicine, 2018.

9. Joonmyoung Kwon, Youngnam Lee, Yeha Lee, Seungwoo Lee, Hyunho Park, and Jinsik Park. Approval of deep learning-based emergency and keenness score utilizing an enormous public dataset. PLOS ONE, 2018.

10. Jimeng Sun, Edward Choi, Andy Schuetz, Walter, F. Stewart Using intermittent neural organization models to detect the onset of cardiovascular breakdown. American Medical Informatics Association Journal, 24(2):361-370, 2017.

11. Jonathan Rubin, Cristhian Potes, Minnan Xu- Wilson, Junzi Dong, Asif Rahman, Hiep Nguyen, and David Moromisato. A group supporting model for anticipating move to the pediatric emergency unit. Global Journal of Medical Informatics, 112:15-20, 2018.

12. Huan Song, Deepta Rajan, Jayaraman J. Thiagarajan, and Andreas Spanias. Join in and Diagnose: Clinical Time Series Investigation utilizing Attention Models. arXiv, 2017.

13. Joseph Futoma, Sanjay Hariharan, and Katherine Heller. Figuring out how to Detect Sepsis with a Multitask Gaussian Process RNN Classifier. In Proceedings of the 34th International Gathering on Machine Learning, 2017.

14. Edward Choi, Andy Schuetz, Walter F. Stewart, and Jimeng Sun. Utilizing intermittent brain network models for early discovery of cardiovascular breakdown beginning. Diary of the American Medical Informatics Association, 24(2):361-370, 2017.

15. Alvin Rajkumar and Others. Versatile and precise profound learning for electronic wellbeing records. Nature Digital Medication, 1(1):1-10, 2018.

16. Joon myoung Kwon, Youngnam Lee, Yeha Lee, Seungwoo Lee, Hyunho Park, and Jinsik Park. Approval of deep learning- based emergency and sharpness score utilizing a huge public dataset. PLoS ONE, 2018.

17. Yoshua Bengio, Aaron Courville, and Pascal Vincent. Portrayal learning: A survey and new points of view. IEEE Exchanges on Pattern Analysis and Machine Intelligence, 35(8):1798-1828, 2013.

18. Jimeng Sun, Fei Wang, Jianying Hu, and Shahram Edabollahi. Grouping Overly-Specific Features in Electronic Clinical Records. ACM SIGKDD Explorations Newsletter, 14(1):16, 2012.

19. Marco A.F. Pimentel, David A. Clifton, and Lionel Tarassenko. Gaussian interaction bunching for the practical characterization of imperative sign directions. In IEEE International

Workshop on Machine Learning for Signal Processing, MLSP, 2013.

20. Glen Wright Colopy, Stephen J. Roberts, and David A. Clifton. Gaussian Processes for Personalized Interpretable Unpredictability Metrics in the Step-Down Ward. IEEE Journal of Biomedical and Health Informatics, 2019