Voice Based Prescription Generation using Artificial Intelligence

Kajal Ghadage[1], Laxman Reddy[2], Deepak Borate[3], Omkar Dalavi[4], Prof. S.P.Aundhakar[5]

[1-4]U.G Student, Dept. of computer Engineering, Tssm's PVPIT, Pune, Maharashtra, India
[5]Professor, Dept. of computer Engineering, Tssm's PVPIT, Pune, Maharashtra, India

---***---

Abstract – *Since prescriptions in hospitals are still written by hand, illegible handwriting may lead to major problems like taking the wrong drug or taking the wrong quantity of the drug by patient. This may cause the serious issue of health or death of the patient. To solve this issue voice-based prescription generation system came into the picture where the prescription is taken as input in voice format with help of Google speech recognition API to convert speech to text. This text transcript is obtained to conduct name entity recognition (NER) task to extract medical entities from text. In this way, the digital prescription will be generated. For performing the NER task Bi-LSTM and CRF networks are used. This system is built using MERN (MongoDB*, Express JS, React JS, and Node JS) *stack where react JS technology is used for the frontend along with this node JS and express JS is used with MongoDB database as backend. This independent platform web-based system will improve the procedure of generating prescriptions.*

Key Words: Speech Recognition, Speech to text, Annotations, Natural language processing (NLP), Named entity recognition (NER), Electronic Health Record (EHR), *MERN (MongoDB*, Express JS, React JS, and Node JS).

1. INTRODUCTION

A prime problem today in India and abroad is most prescriptions are still written by hand. If a doctor has given some medicine, for example, "Vyvanse tablet" such medicine is only readable by pharmaceutical people like chemist due to illegible handwriting of doctor causes the non-medical background people to interpret the prescribed medicines erroneously. This cause problem, a patient will not able to read prescription correctly and also not able to verify medicine given by chemist is as per prescription or not. If medicine given by chemist is wrong by misunderstanding then these cause lots of damages or adverse drug reactions (ADRs) to patient [1]. Furthermore, the problem with all hospitals is not having any method like an electronic health record system to know the history of the patient and what types of tablets consumed in the past of a particular patient. Writing prescriptions by hand takes time which leads to doctors attending only fewer patients in the scheduled time.

Now, It is the time of the computer era where, everything is computerized, boosting the pace of human life. To make the prescription generation system computerized, the scenario of voice based prescription generation system using AI comes into picture. A solution to the above-mentioned problems is to create an application that can be used to reduce the work of doctors. By using this system, the Doctor will be able to dictate his prescription to the patient and at the same time, this dictation is gets recorded by the system. This recorded prescription is converted to text and extracts medical entities from text like a drug, drug frequency, drug dosage, etc. Medication entities matching the above examples are demonstrated in Figure-1. In the end, a PDF prescription was generated with help of classified tokens. Additionally using this app, the Doctor should be able to edit the prescription, sign the prescription and also send it to the patient directly on his email ID.

Drug	Duration	Strength	Route	form	Dosage	Frequency
Crosine	For 2 days	500mg	by mouth	tablet	1	Every 6 hours

Fig-1: Medication entities with example

Due to this world-changing solution, doctors will be able to handle more patients in a small amount of time. Now the chemist will able to read the prescription rightly, also the patient will be able to verify prescription given by the chemist is the same as written in the prescription. In Addition, a prescription is sent on a patient's email id the patient can show all previous prescriptions to the doctor in a secure way even though the patient goes to a distinct hospital all prescriptions can be viewable by a doctor. This will also prevent the illegal use of patient prescription and provide security.

The objective of this project is to design a system which will generate voice-based prescription where Patient will get prescription in PDF format through the mail. So, there is no chance of wrapping, burning, and loss of prescription. If this solution is implemented in a real-time hospital system then it will lead to saving lots of time for doctors to write a prescription as well as patients to search previous prescriptions.

2. LITERATURE REVIEW

The Electronic prescription system [3] will generate prescriptions using speech recognition API and natural language processing techniques. They proposed an approach of using Speech Recognition technology to speed up the process of prescription generation. As we were known, speaking a sentence will consume less amount of time than writing it. And For speech-to-text conversion, Google's speech recognition API has been used. This is the best available API as it supports Indian English. NLP is used to extract prescription information from the transcript.

MedLEE [4] is one of the earliest clinical information extraction (CIE) systems, built to extract, structure, and encoding clinical information within the raw text of patient reports. MedLEE application extracts medication information using hand-written rules.

The healthcare NER model solution [5] is used for extracting structured information from unstructured text such as EHR and medical records. This system required us to extract key entities such as prescribed drugs, dosage, etc mentioned in the EHRs. The extracted entities would be processed further downstream to find out relationships between entities like drug and dosage.

3. Data

We have collected sample medical notes data from various publicly available websites. Using these datasets defined 8 types of entities: Drug, Duration of a drug, strength of a drug, route to take a drug, form of a drug, dosage of a drug, and frequency of a drug. We have collected sample medical transcripts from MT samples [7] and MT example [8]. These both provide open access to large big collected transcript medical reports for reference purposes. To perform annotation on data collected from these websites Doccano annotation tool [9] is used with its 8 predefined entities.

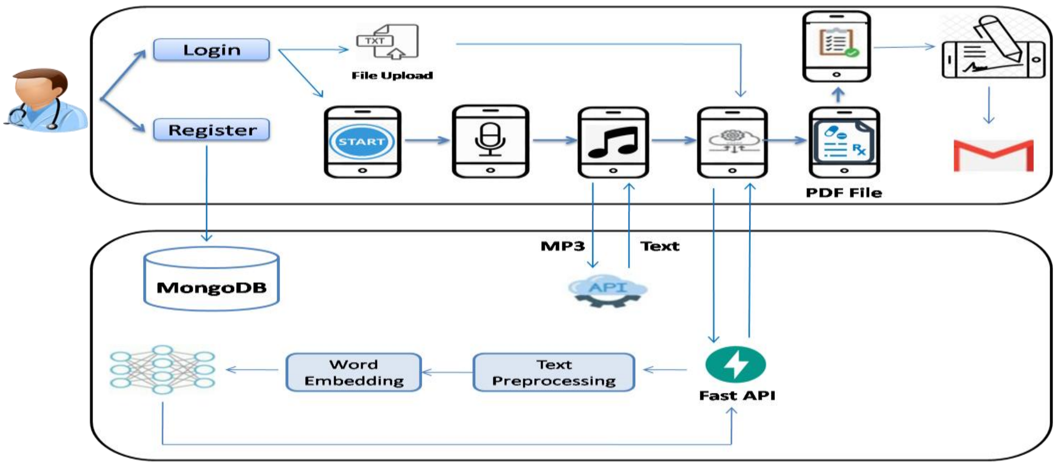

Fig-2: Voice Prescription System Architecture

Along with these we have also used the N2C2 2018 shared task gold standard corpus [10] on adverse drug events and medication entities extraction in electronic health records. Table-1 shows N2C2 dataset details used in a project to train and test the model. This Dataset belongs to Harvard Medical School and consists of 404 Electronic Health Records and their annotations files.

Table -1: N2C2 2018 dataset details

No of tokens in dataset	Train set	Test set
# as a Drug	20,684	5797
# as a Strength	8507	2387
# as a Duration	741	231
# as a Route	6992	1978
#as a Form	8602	2399
# as a Dosage	5325	1530
# as a Reason	8034	2254

4. ARCHITECTURE OVERVIEW

The voice Prescription system architecture is shown in Figure-2. First of all, a doctor has to perform a login for authentication purposes. After successful login in the system, the doctor can detect prescription to patient and prescription can be recorded by clicking the "START" button on a web page. This recorded voice is sent over Google speech recognition API. Google speech recognition API provides medical transcripts in text format. Along with Speech recognition, an alternative option is to upload a prescription written raw text File.

This text is in raw format to make this useful and to extract entities from this raw text, backend request is sent FastAPI consisting of Bi-LSTM and CRF model.

FastAPI [11] is one of the fastest Python-based web frameworks. Before extracting predefined entities from text, text preprocessing is performed. Furthermore, Word embedding is performed. Word embedding is phrases or words in vocabulary that are mapped to vectors of real numbers. Word embedding is a term used for representing text in vector format. As we know computers cannot understand text format but they can understand numbers.

This processed text is given to the trained model to perform the NER task. The Bi-LSTM and CRF model gives output in {token: entity} format. Here, a token is peace of word from given input text and entity represents in which of predefined entity this word belongs. For example {"Crosine": "drug"}. Furthermore, this model result is sent back to the app and by using the model result, a prescription is generated which is editable by doctors. In prescription, a doctor is authorized to add a new medicine, delete existing medicine, and edit existing medicine.

Now, the most important step where a doctor has to verify the prescription (to correct if there is any transcription error) and perform a digital signature. After the verification process, the prescription in PDF format is sent to the patient's mail securely.

4.1 Text Preprocessing

EHR records are usually lengthy and it is not possible to perform NER tasks on such lengthy records. To solve this problem, a function is defined to divide a large chunk of data into small pieces this process is known as tokenization. a function would split the EHR records based on a maximum sequence length parameter. The function tries to include maximum number of tokens, maintaining as much context as possible for every token.

The data is tokenized using a modified ScispaCy tokenizer for BiLSTM + CRF model which just removes the tokens with whitespace characters after ScispaCy tokenizes them. Each sequence of labels or tokens in the data was represented using the IOB (Inside, Outside, Beginning) tagging scheme for BiLSTM models.

4.2 Algorithm for NER

To perform name entity recognition task on EHR Bidirectional LSTM (Bi-LSTM) coupled with a conditional random field (CRF) classifier algorithm. This model is very efficient for a variety of sequence tagging tasks.

Bi-LSTM : Just a Bi-LSTM network is enough to classify each token into various entities along with its class B and I (i.e. B: beginning of entity and I: inside of entity) for example as shown in Figure-4 "for three days" represent duration. But in result, word "for" will have entity B-Duration means beginning of duration and all further words will have entity I-Duration. And if token is not a part of any of the pre-defined entities we are looking for then it is classified as O (i.e. O: outside) but we witnessed some common errors of misclassification. Because the outputs of Bi-LSTM of each word are the label scores, we can select the label which has the highest score for each word.

CRF : with this Bi-LSTM scheme, we may end up with invalid outputs, for e.g.: I-Drug followed by I-Frequency or B-Drug followed by I-Frequency. Hence we are using CRF (Conditional Random Field) algorithm to calculate the loss of our Bi-LSTM network as it could add some constraints to the final predicted labels to ensure they are valid. These constraints are learned by CRF automatically from the training dataset during the training process. CRFs considers the context as well rather than just predicting label for a single token without considering neighboring samples [12].Figure-4 shows output of trained Bi-LSTM and CRF network where each word has assigned token.

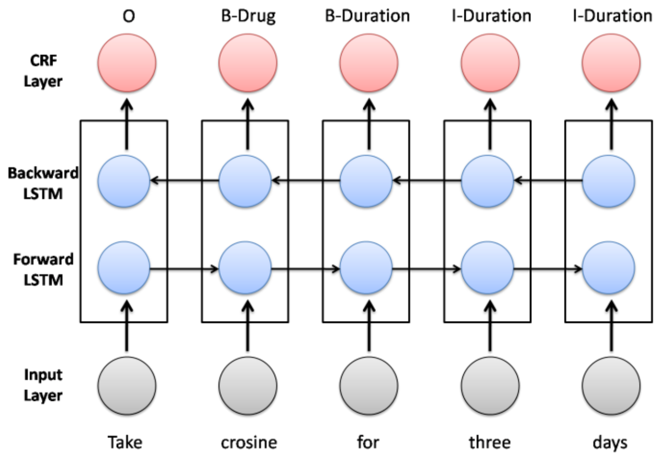

Fig-4: Output of Bi-LSTM model

Figure-3 represents steps to train Bi-LSTM and CRF for NER task with N2C2 dataset.

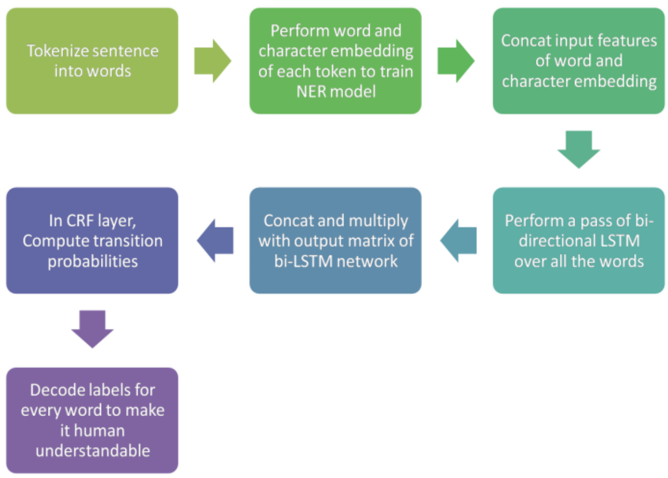

Fig-3: Training procedure of NER task using bidirectional LSTM-CRF

5. React-JS WEB APPLICATION

The proposed system is a web-based application developed using MERN stack this makes system platform-independent. For web application frontend development React JS technology is used with Google cloud speech API for transcription. And for backend concerns, Node JS is used and for a database storing purpose, Mongo DB is used.

5.1 Authentication

Authentication is the primary step that a doctor has to perform in the system. Security of patient is the first concern for that purpose authentication is a way to give authority to doctor to prescribe a prescription. Figure-5 shows the Login page of a system which consists of 2 fields Email-ID and password. The doctor has to login into the system using Email-ID and password provided at the time of registration.

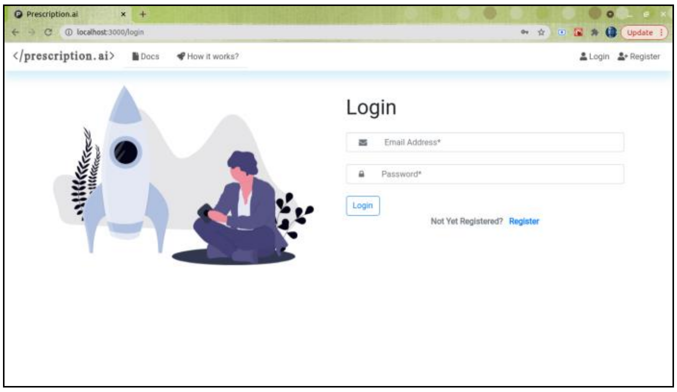

Fig -5: Login Page for authentication

5.2 Give voice input for prescription

Once login is done doctor is ready to prescribe a prescription. There are two options available to provide input, first is Voice input. Figure-6 shows the prescription input page. Here, to record voice prescription start button should click. After clicking the start button, it starts recording voice prescriptions. And the other available option is text file input.

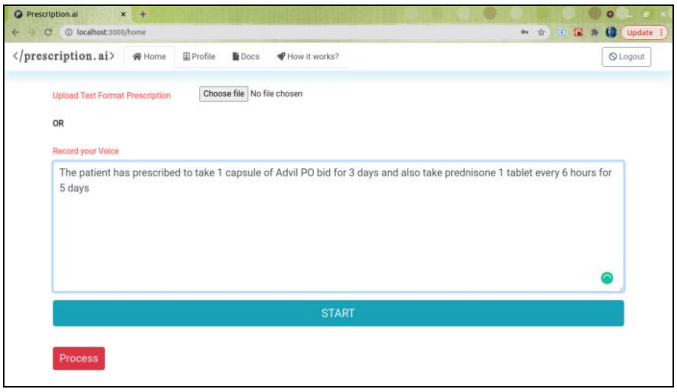

Fig -6: Prescription voice input page

When Proceed button is clicked, a request is sent to the fastAPI framework to obtain entities for each input word. When the response is obtained from fastAPI, a page will show all tokens with their entities.

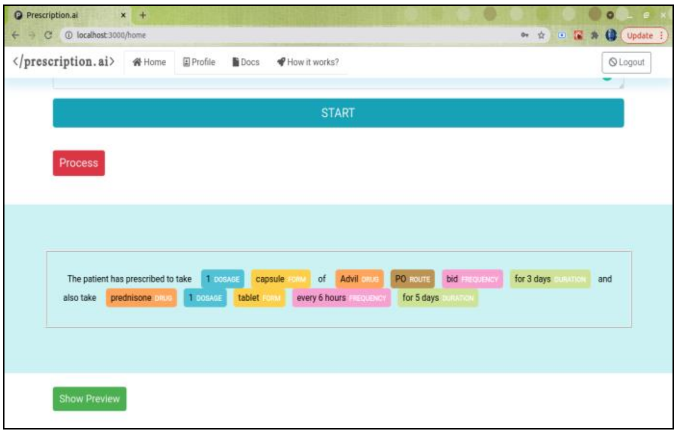

Fig -7: Visual representation of NER output

As shown in Figure-7 Advil is classified as a drug, for 3 days is classified as duration, etc. Visualization of data is one of the best ways to make it easy to understand so doctors can make quick reviews of classified entities.

5.3 prescriptions Generation and Send Mail to patient

Figure-8 shows generated prescription format. Prescription has three input fields Patient name, Patient email-ID, and patient ID. This information is used to send prescription to patients Email-ID. Also, it has seven columns to show prescription in readable form drug name, Duration, Strength, Route, form, dosage, and frequency. A digital signature is also part of a prescription. The need for a digital signature is to shows the authenticity of a document.

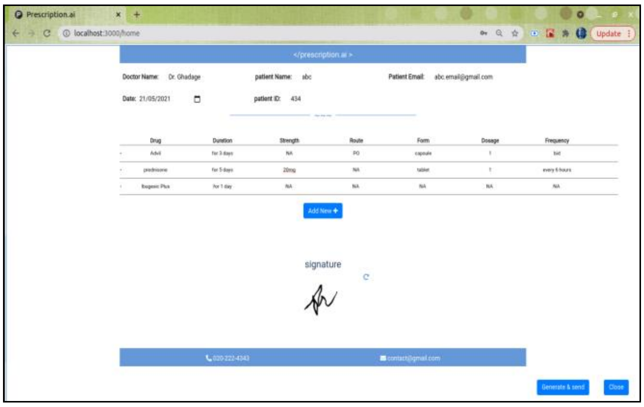

Fig -8: Generated prescription

Figure-9 shows a screenshot of mail sends on patient Email-ID successfully with prescription attachment.

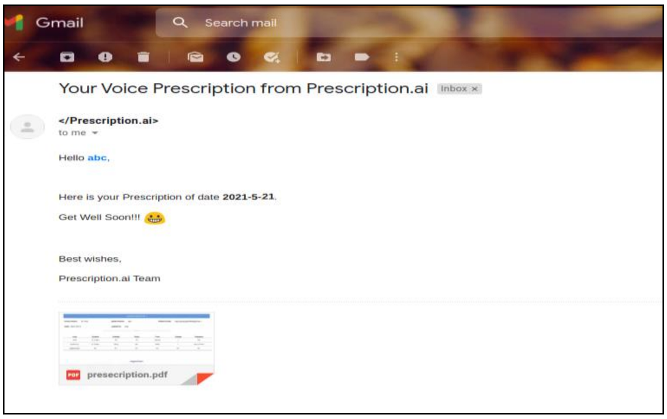

Fig -9: Mail sent successfully

6. RESULT

The Bi-LSTM and CRF model trained on 15 epochs with an F1 score of training is 89.82% also an F1 score of testing is 87.02%. Chart-1 shows a plot of the training accuracy VS validation accuracy over the number of epochs.

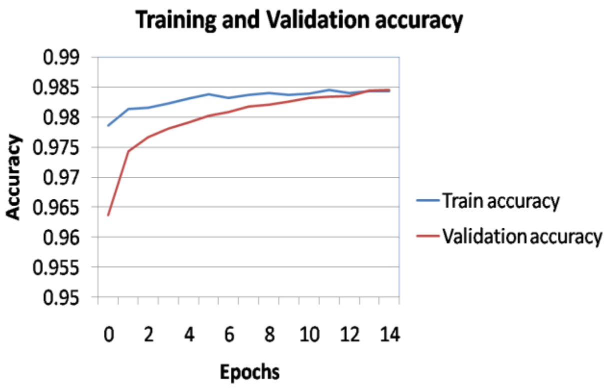

Chart -1: train and validation accuracy chart

The proposed system is a web based application that does not need any infrastructure to use it. The implemented system is associated with two major advantages over the existing healthcare system.

1) This System saves doctors time in writing prescriptions as well as paper

2) Make it easy to comprehend doctor's notes

The one major advantage associated with this system is that it perform NER task on both normal English medical phrase as well as doctors abbreviations.

Table-2 contains doctor abbreviations and its respective normal medical phrases for how often to take medicine and how to use or intake medicine. For example, a bid is an abbreviation used by a doctor to say that take medicine twice a day, PO is abbreviation used by doctor to say that take medicine by mouth, etc.

Table -2: Medical Abbreviations

How often to take medicines	
bid	Twice a day
q3h	Every 3 hours
Q4h	Every 4 hours
qd	Every day
qid	Four times a day
How to use your medicine	
PO	By mouth
OU	Both eyes
OD	Right eye
OS	Left eye
AD	Right ear
AL	Left ear

7. FUTURE SCOPE

In this work, we have used Google speech recognition API for speech recognition. Google speech recognition spells every word it recognizes correctly. Typically, it recognizes 5–10% of words incorrectly to overcome this we will build our speech recognition model.

Now our proposed system is only supported the Indian English language but in near future, it can also support multiple languages.

In the future, we intend to improve the performance of the system by including Unified models for both NER and Relation Extraction. This would also allow the doctors to easily find out relationships between drugs and ADE so that such drugs can be monitored carefully.

8. CONCLUSIONS

The voice prescription system will need a minimal change in the workflow of doctor's and it will also create a huge impact in developing a digital EHR system for patients and doctors. A Voice prescription system helps in managing electronic health records in real-time while maintaining the patient's privacy. This digital system will reduce the patient's record access time and maintain high security and privacy of patient data.

ACKNOWLEDGEMENT

The authors would like to thank publishers and researchers for making their resources available and teachers for their guidance. We thank to the college authority for providing the required infrastructure and technical support.

REFERENCES

[1] Trithankar Dasgupta, Aakash Anuj, Manjira Sinha, Ritwika Ghose, Anupam Basu "Voice Mail architecture in desktop and mobile devices for the Blind people", 2012 4th International Conference on Intelligent Human Computer Interaction, www.ieee.org.

[2] Shriram KV "AI based Voice Prescription System for Doctors | Intel DevMesh ", Intel software developer zone.

[3] Jitendra Mahatpure, Dr Piyush Kumar Shukla, Dr Mahesh Motwani, "An Electronic Prescription System powered by Speech Recognition, Natural Language Processing and Blockchain Technology", International Journal Of Scientific & Technology Research, VOLUME 8, ISSUE 08, AUGUST 2019 ISSN 2277-8616 1454 IJSTR©2019 www.ijstr.org.

[4] Carson Tao Department of Information Science, State University of New York at Albany, NY, USA, Michele Filannino Department of Computer Science, State University of New York at Albany, NY, USA and Özlem Uzuner, "Prescription Extraction Using CRFs and Word Embedding", published online 2017 Jul 4.

[5] Amogh Kamat Tarcar, Aashis Tiwari, Vineet Naique Dhaimodker, Penjo Rebelo, Rahul Desai, Dattaraj Rao,"Healthcare NER Models using language Model Pretraning", www.arxiv.org

[6] Xi Yang, Jiang Bain, William R Hogan, Yonghui Wu, "Clinical concept extraction using transformers "| Journal of the American Medical Informatics Association | Oxford Academic (oup.com)

[7] Collection of transcribed "medical transcription sample reports and examples"

[8] Medical transcript examples "MT example"

[9] Daccano annotation tool- "doccano · PyPI"

[10] N2c2 2018, "Adverse drug event and medication extraction in EHR's"

[11] FastAPI web framework " https://fastapi.tiangolo.com/ " python

[12] Anuj Ananad, Viral Pandey, Smit Kiri "EHR Relation extraction Github".