MedicBot: A New Virtual Assistance for the Children with Auditory Processing Disorder

Do Dung Vu
Supervisor: Prof. Sylvie Ratté

Département de génie logiciel et des TI

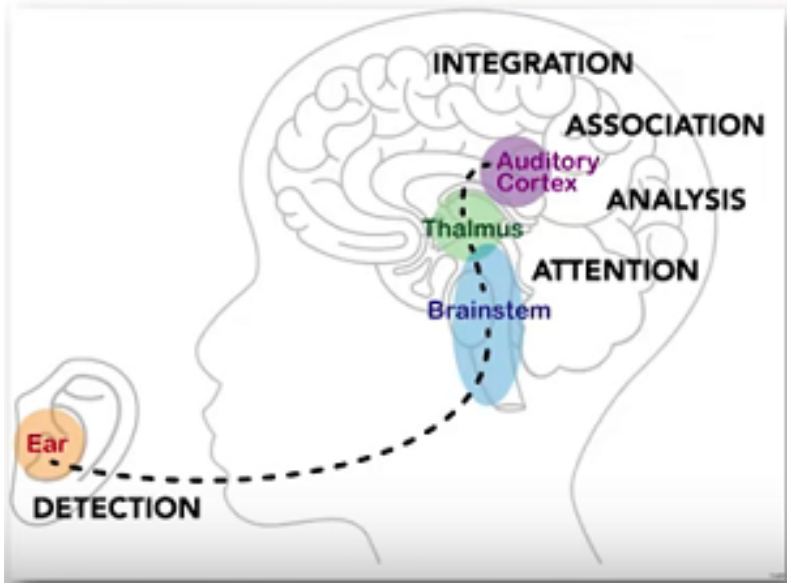
do-dung.vu.1@ens.etsmtl.ca

February 19, 2019

Overview

1. Introduction
2. Methodology
3. Techniques
4. Conclusion

Introduction



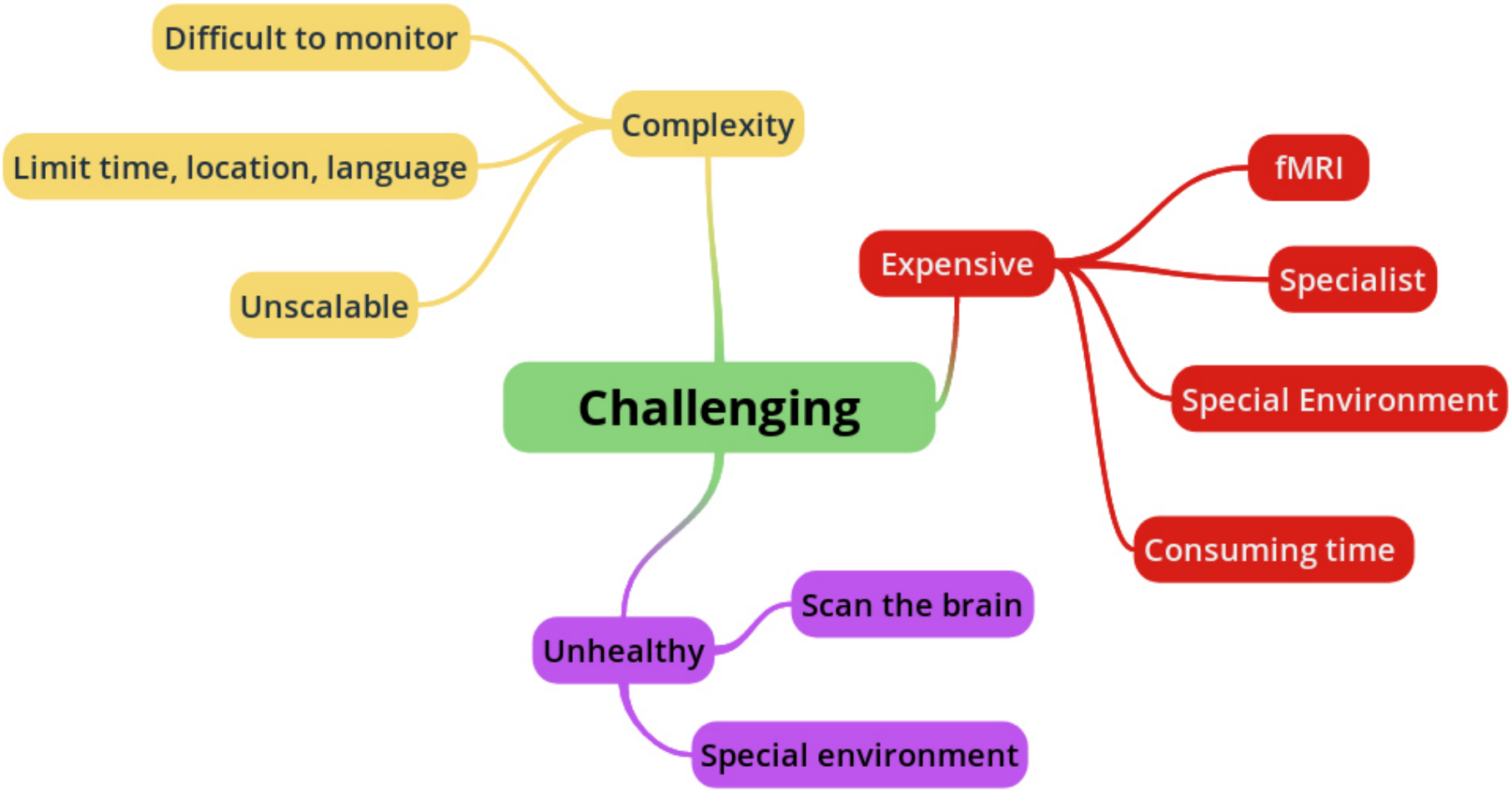
https://www.autismspeaks.org

- Auditory processing is defined as what we do with what we hear [a]
- Auditory Processing Disorder (APD) is a condition where someone has normal hearing, but the auditory system does not faithfully bring information to the brain [b]
- **Approximate 2-4% of school age children have APD** [c]
- Autism and auditory processing disorders often overlap[d]

---

[a] Katz & Tillery, 2004
[b] https://www.sac-oac.ca
[c] http://www.ementalhealth.ca/
[d] https://www.autismspeaks.org

Challenging



Objectives

Propose an AI model (virtual assitance) named **MedicBot** to assit in diagnosing, monitoring, and training of the children with APD problem with low price, healthy, and convinent

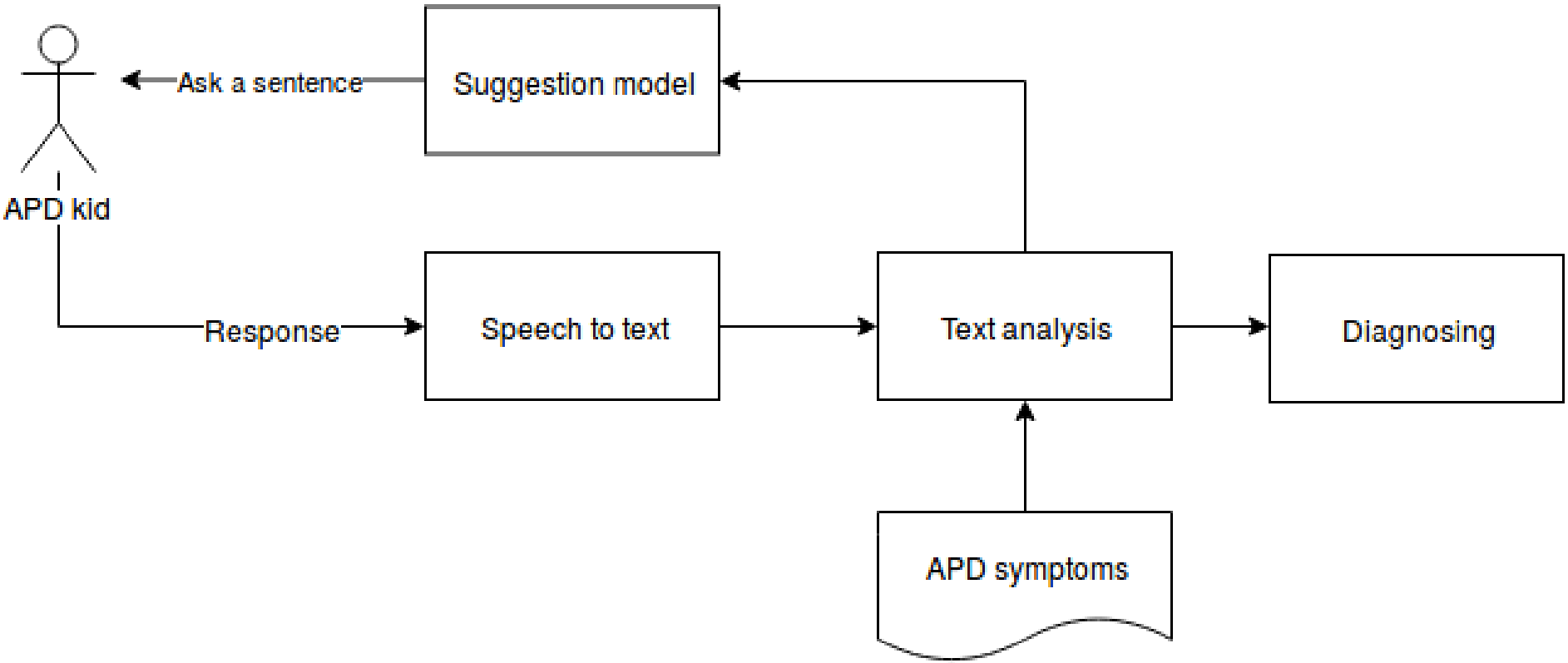
MedicBot

- Diagnose APD symptoms based on conversation with the considered children
- Create a Training Therapy Model Assitance (adaptable)
- Build the Reinforcement Learining (RL) Model to monitor the progress of APD treatment

Methodology

- Analysis the given APD symptoms by speech recognition based on Deep learning
- Analysis the given APD therapy and recommend the treatment to the APD children. Apply a natural language processing (NLP) to generate sentences and exploit Deep learning to understand the context of the speech
- Monitoring the process of APD treatment by using speech analysis based on Deep
  learning

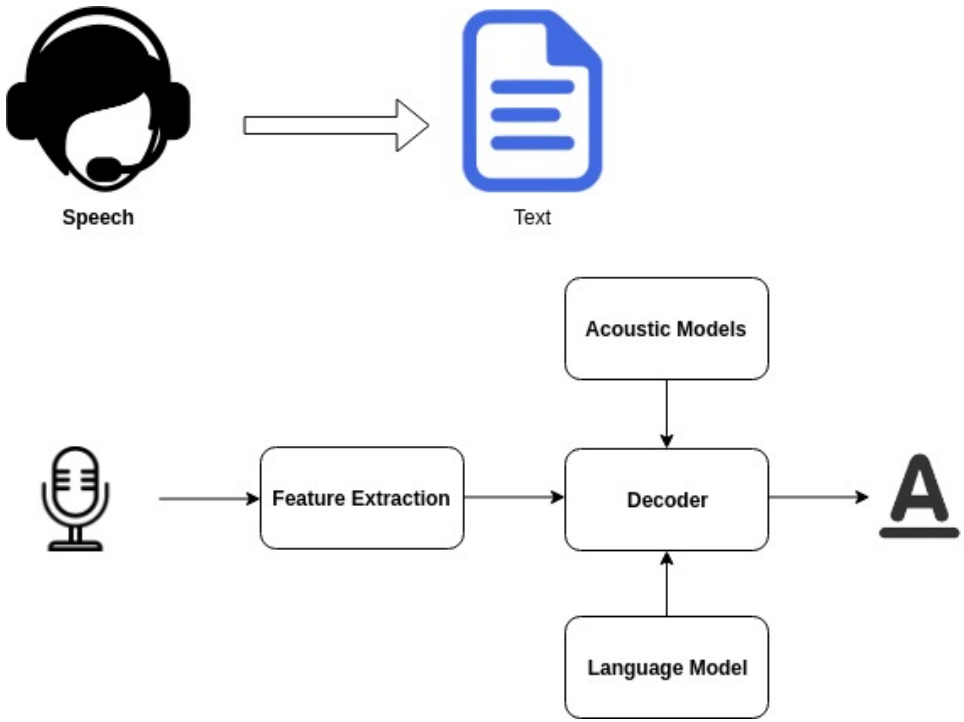
Implementation



Techniques

Convert speech to text

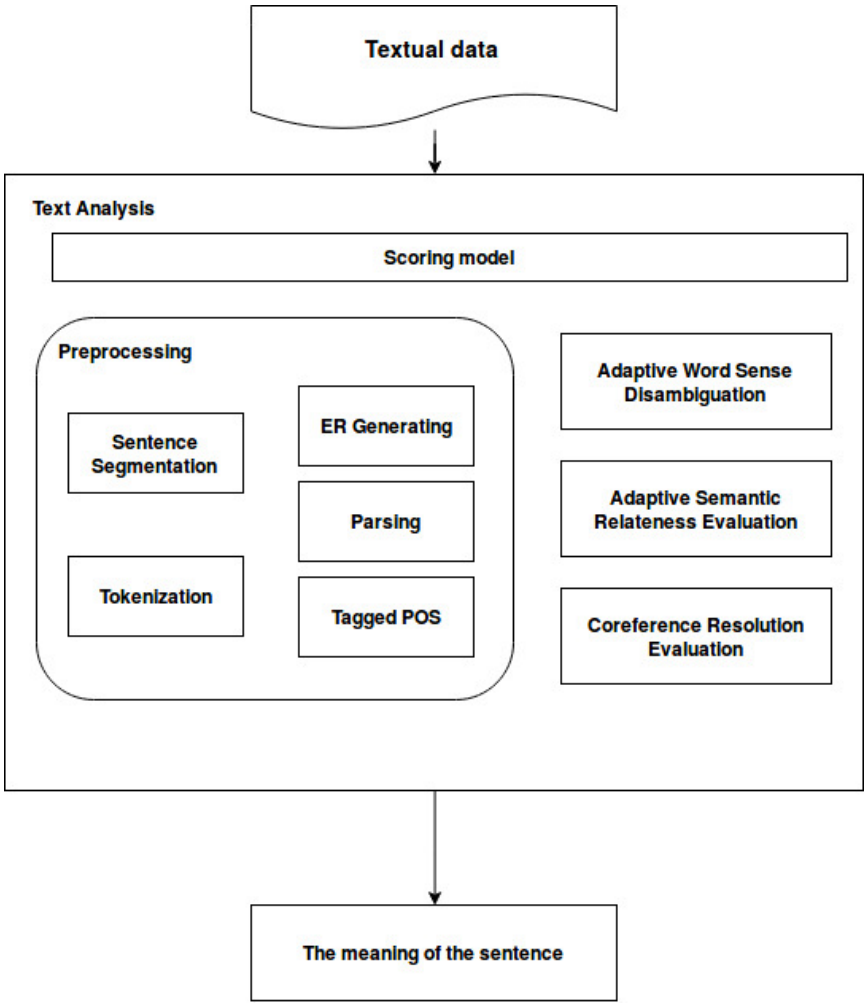
- **Acoustic modeling** represents the relationship between linguistic units of speech and audio signals.
- **Language modeling** matches sounds with word sequences to help distinguish between words that sound similar.



We evaluated the quality of output by using two factors:[1]
Accuracy and **Speed**

[1] https://pypi.org/project/SpeechRecognition/

Proposal for the objective 1 solution



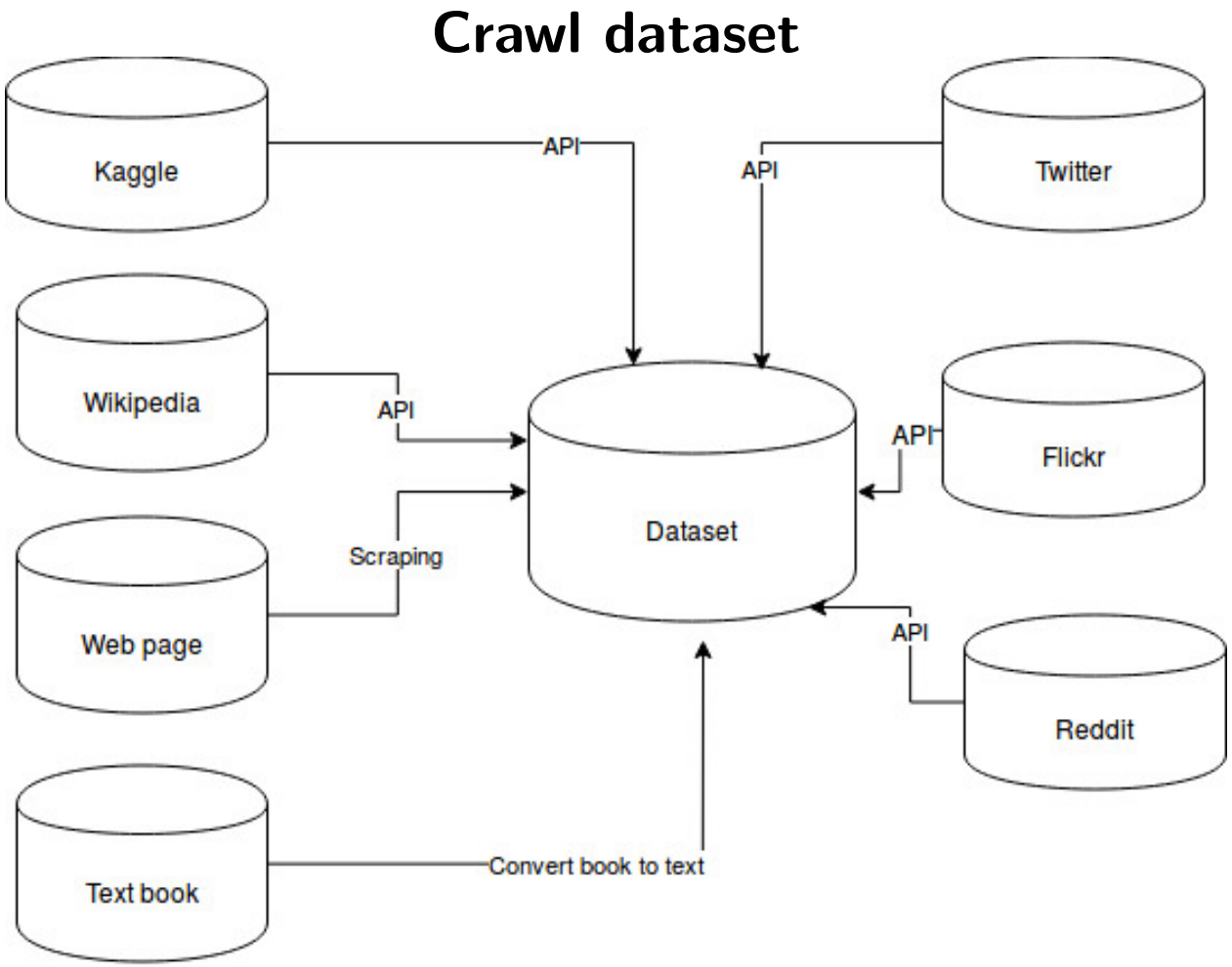
- Step 1: Get the raw text data from the user conversation
- Step 2: Process text go extract and compute the score of features
- Step 3: Adapt word sense disambiguation
- Step 4: Evalue the semantic relateness and coreference resolution
- Step 5: Get the meaning of the sentence

Dataset

Open Dataset

- (NLVR) A Corpus of Natural Language for Visual Reasoning, 2017
- (MS MARCO) MS MARCO: A Human Generated MAchine Reading COmprehension Dataset, 2016
- (NewsQA) NewsQA: A Machine Comprehension Dataset, 2016
- (SQuAD) SQuAD: 100,000+ Questions for Machine Comprehension of Text, 2016
- (GraphQuestions) On Generating Characteristic-rich Question Sets for QA Evaluation, 2016
- (Story Cloze) A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories, 2016
- (Children's Book Test) The Goldilocks Principle: Reading Children's Books with Explicit Memory Representations, 2015
- (SimpleQuestions) Large-scale Simple Question Answering with Memory Networks, 2015
- (WikiQA) WikiQA: A Challenge Dataset for Open-Domain Question Answering, 2015
- (CNN-DailyMail) Teaching Machines to Read and Comprehend, 2015
- (QuizBowl) A Neural Network for Factoid Question Answering over Paragraphs, 2014
- (MCTest) MCTest: A Challenge Dataset for the
- Open-Domain Machine Comprehension of Text, 2013
- (QASent) What is the Jeopardy model? A quasisynchronous grammar for QA, 2007

Dataset


Crawl dataset
Kaggle
API
API
Twitter
Wikipedia
API
Dataset
API
Flickr
Scraping
Web page
API
Reddit
Text book
Convert book to text

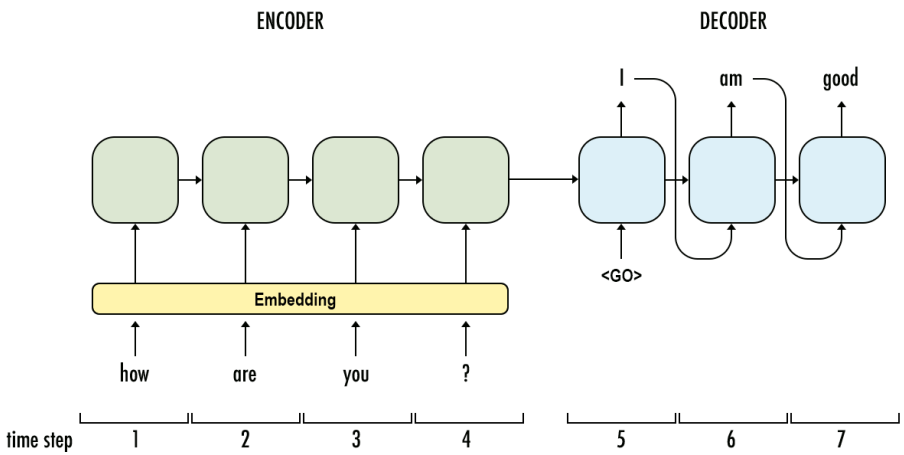
Dataset

Paraphase database (PPDB)

- **PPDB**[2] is an automatically extracted database containing millions paraphrases in 16 different languages.
- The goal of PPBD is to improve language processing by making systems more robust to language variability and unseen words.
- The entire PPDB resource is freely available under the Creative Commons Attribution 3.0 United States License.
- PPDB contains over 150 million paraphrase rules covering three paraphrase types lexical (single word), phrasal (multiword), and syntactic restructuring rules

[2] http://paraphrase.org

Chatbot



- $< PAD >$: During training, inputs in these batches all need to be the same width for the network to do its calculation.
- $< EOS >$: It allows us to tell the decoder where a sentence ends, and it allows the decoder to indicate the same thing in its outputs as well.
- $< UNK >$: If you're training your model on real data, you'll find you can vastly improve the resource efficiency of your model by ignoring words that don't show up often enough in your vocabulary to warrant consideration. We replace those with $< UNK >$.
- $< GO >$: This is the input to the first time step of the decoder to let the decoder know when to start generating output.

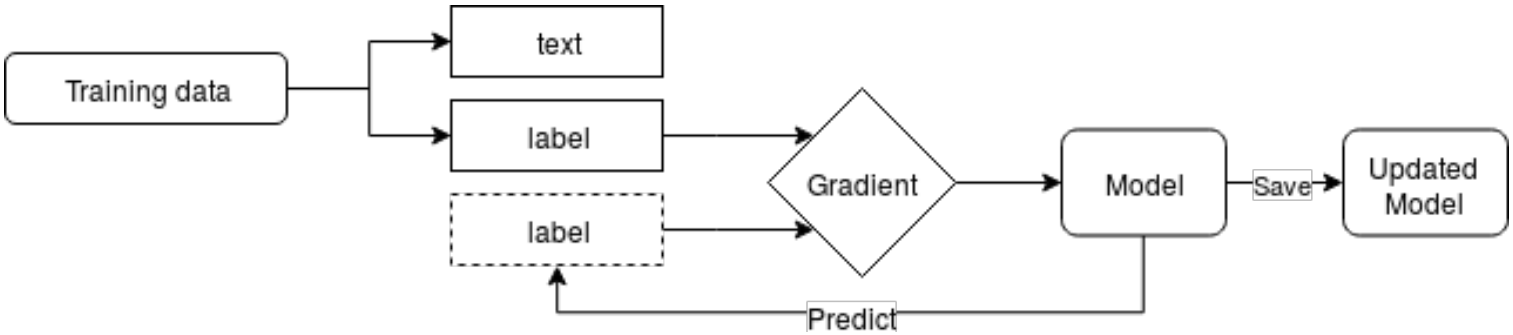
APD Symptoms

Listening
Has difficulty locating a sound source.
Has difficulty hearing in noisy background.
Often asks for repetition or clarification
Speaking
Has difficulty answering open–ended questions
May speak in oversimplified short sentences with difficulties in syntax
Mispronounced words, especially long words.
Phonological Awareness
Has difficulty focusing during conversations
Forgets information that is easily heard

Table: A part of checklist for assessing whether the child with APD[3]

[3] https://kidshear.com.au

Process of creating model

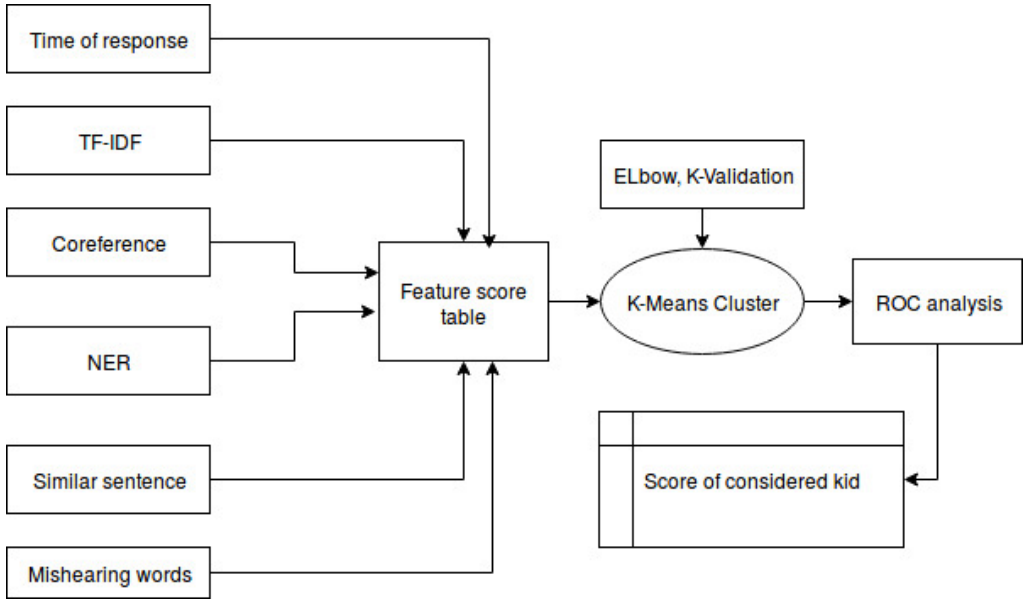


- MedicBot's models are **statistical** and every "decision" she makes is a **prediction**
- This precision is based on the examples the model has seen during training
- We give the model feedback on its prediction in the form of an **error gradient** of the **loss function** that calculates the difference between the training example and the expected output

Techniques

Scoring model

- Create the features for the scoring model
- Compute the score for these ones
- Using the K-Mean Clustering algorithm to cluter the kid
- Apply the Elbow and K-validation algorithm to optimize the K-value of K-Mean Clustering algorithm
- Make the score table of considered kids

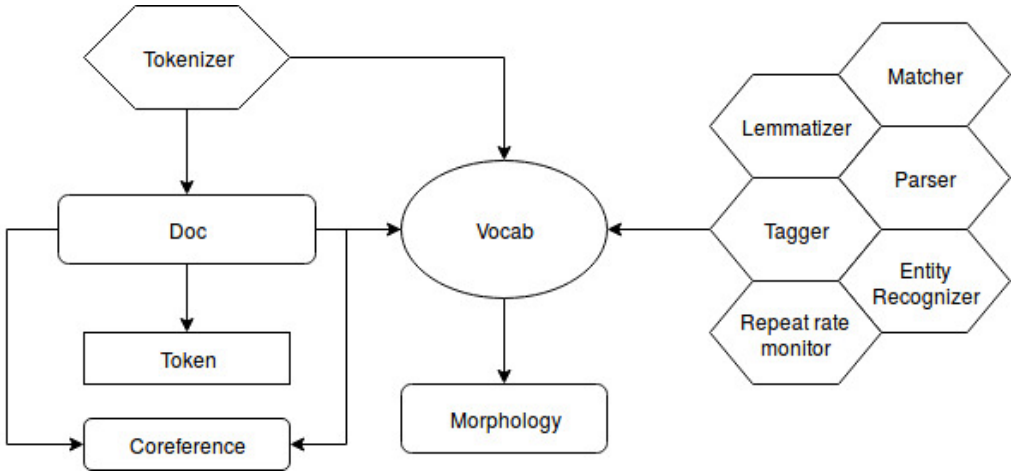


Techniques

Preprocessing

- Recognize the subject, verb, object in a given sentence
- Recoginze noun, adj, adv, preposition of a sentence
- Recognize the entities in a sentence
- Recognize the coreference of a sentence

$\rightarrow$ *create the table of features*



time of response	$\sum_{t=0}^{T}(t_i)$, t_i the duration of one sentence
tf-idf(k,d,D)	$tf(k,d) \times idf(k,D)$, k: term k d: document d; and $d \in D$
coreference	coreference resolution evaluation
ner	name entity recognization
similar sentence	similarity evaluation
mishearing word	spelling and grammar checking evaluation
elbow algorithm	choose a small value of k that still has a low SSE
ROC analysis	Receiver Operating Characteristic analysis

Implementation

Requirements	Content
Python	2.6, 2.7, or 3.3+
PocketSphinx	large vocabulary speaker independent recognition system
PyAudio	0.2.11+ (required only if you need to use microphone input, Microphone)
SpeechRecognition	a process in which a computer or device record the speech of humans and convert it into text
Google API Client	use the Google Cloud Speech API
FLAC encoder	if the system is not x86-based Windows/Linux/OS X

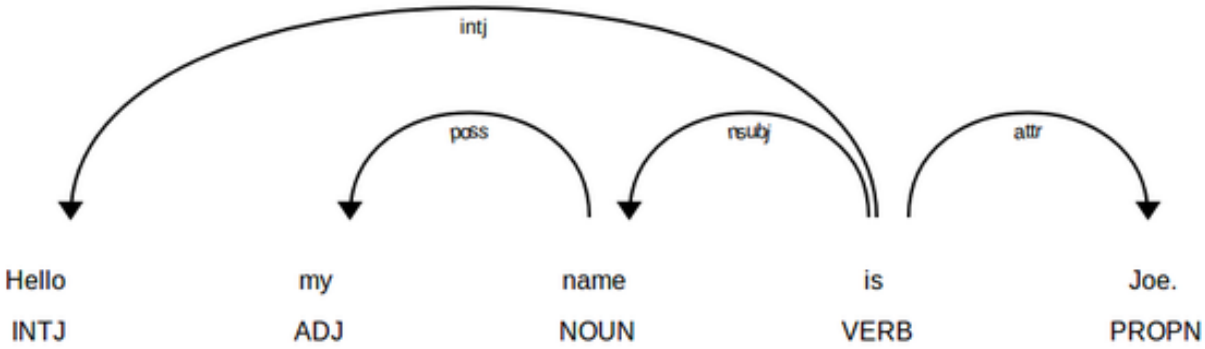
Table: Requirements of environment

Implementation

Requirements	Content
SpaCy	Industrial-Strength NLP
NeuralCoref	Coreference Resolution in spaCy with Neural Networks
NLTK	Natural Language toolkit
Scikit-learn **Tensorflow** **Pytorch**	Machine learning library
ELMo	Embeddings from Language Models

Table: Requirements of environment

Unit tests of the system



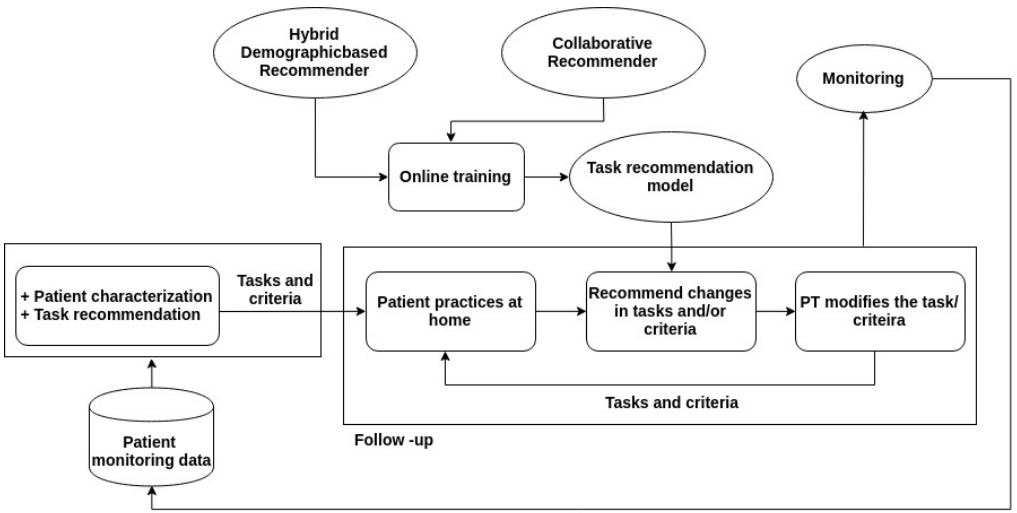
But Google ORG is starting from behind. The company made a late push GPE into hardware, and Apple ORG 's Siri, available on iPhones PRODUCT , and Amazon ORG 's Alexa ORG software, which runs on its Echo GPE and Dot ORG devices, have clear leads in GPE consumer adoption.

```
text 1: My sister has a dog.
text 2:  She loves him
0.016491954893255212  a dog loves him
0.9719385606143631  my sister loves him
0.011569484492381571  she loves him
```

Methodology

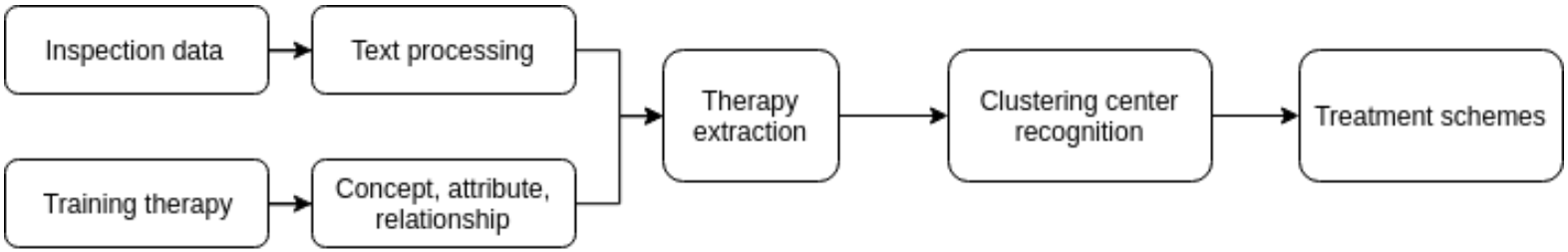
- Analysis the given APD symptoms by speech recognition based on Deep learning
- Analysis the given APD therapy and recommend the treatment to the APD children. Apply a natural language processing (NLP) to generate sentences and exploit Deep learning to understand the context of the speech
- Monitoring the process of APD treatment by using speech analysis based on Deep learning

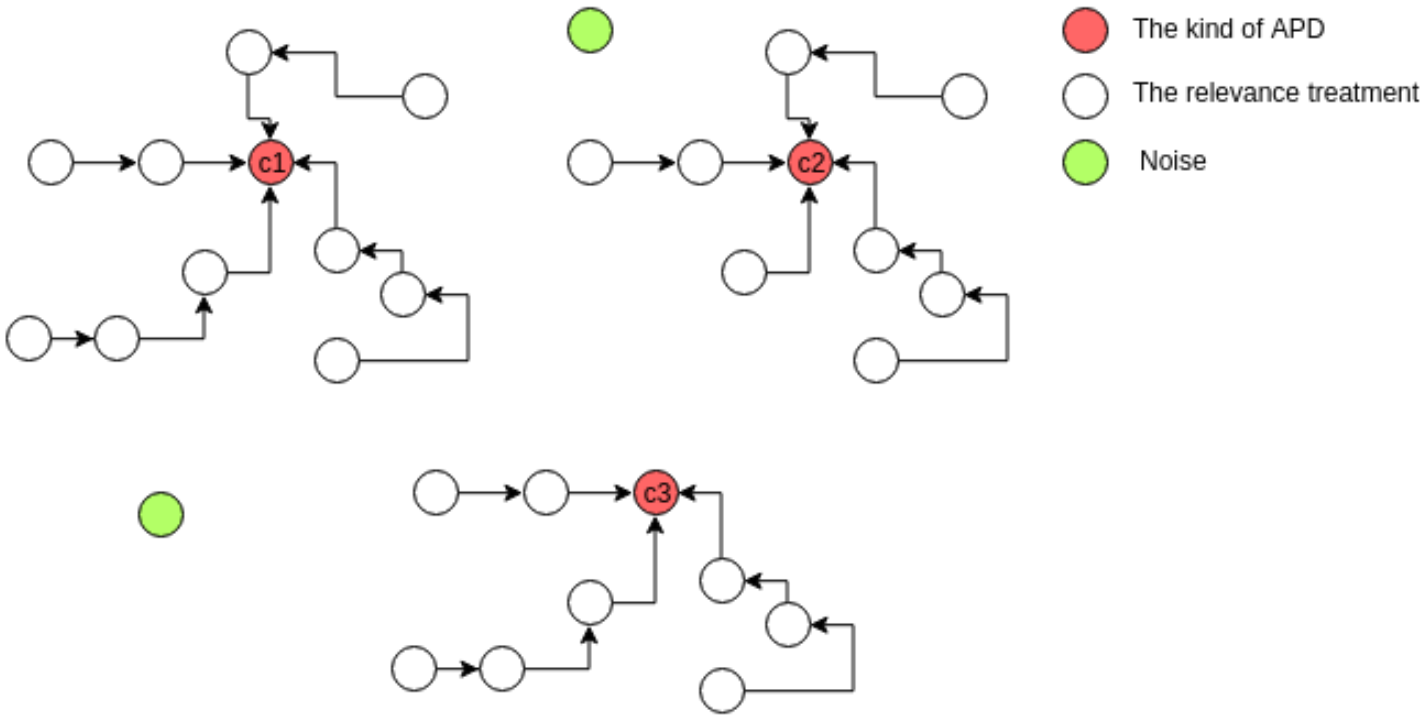
Proposal for the objective 2 solution



- Propose a training therapy to the APD kid based on the diagnosing report and Task recommendation model
- Monnitor the progress of therapy
- Update the monitoring data
- Suggest the fit training therapy by using reinforcement learning based on recommendation system

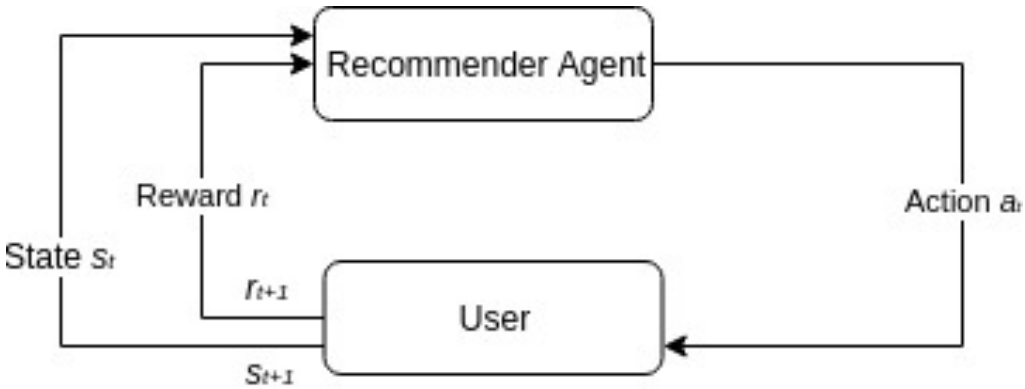
Task recommendation model [4]





[4] https://www.thebsa.org.uk

Tasks and criteria



- State space $\mathcal{S}$: A state $s_t = \{s_t^i\}, i = (1, N)$
- Action space $\mathcal{A}$: An action $a_t = \{a_t^j\}, j = (1, K)$
- Reward $\mathcal{R}$: $r(s_t, a_t)$ according to the user's feedback
- Transition probability $\mathcal{P}$: $p(s_{t+1}|s_t, a_t)$ defines the probebility of state transition from s_t to s_{t+1} when Recommender Agent takes action a_t

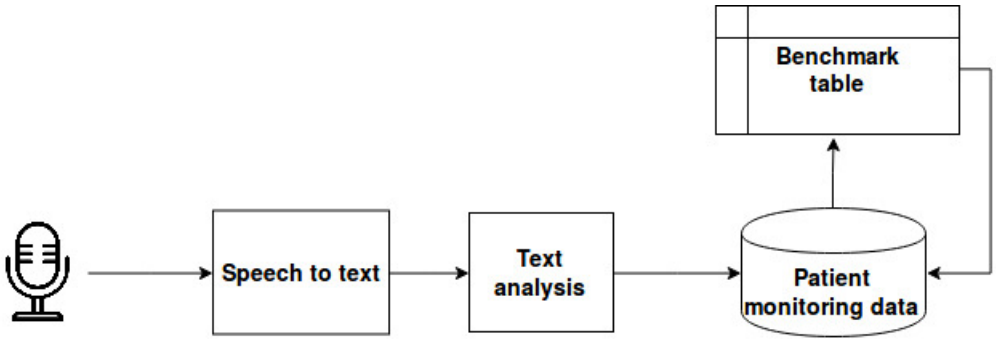
Methodology

- Analysis the given APD symptoms by speech recognition based on Deep learning
- Analysis the given APD therapy and recommend the treatment to the APD children. Apply a natural language processing (NLP) to generate sentences and exploit Deep learning to understand the context of the speech
- Monitoring the process of APD treatment by using speech analysis based on Deep learning

Techniques

Proposal for the objective 3 solution

- Convert the APD speech to text
- Analysis the meaning of the text
- Make the score of the benchmark table
- Make the client-server model to monitor and evaluate the progress of each patient

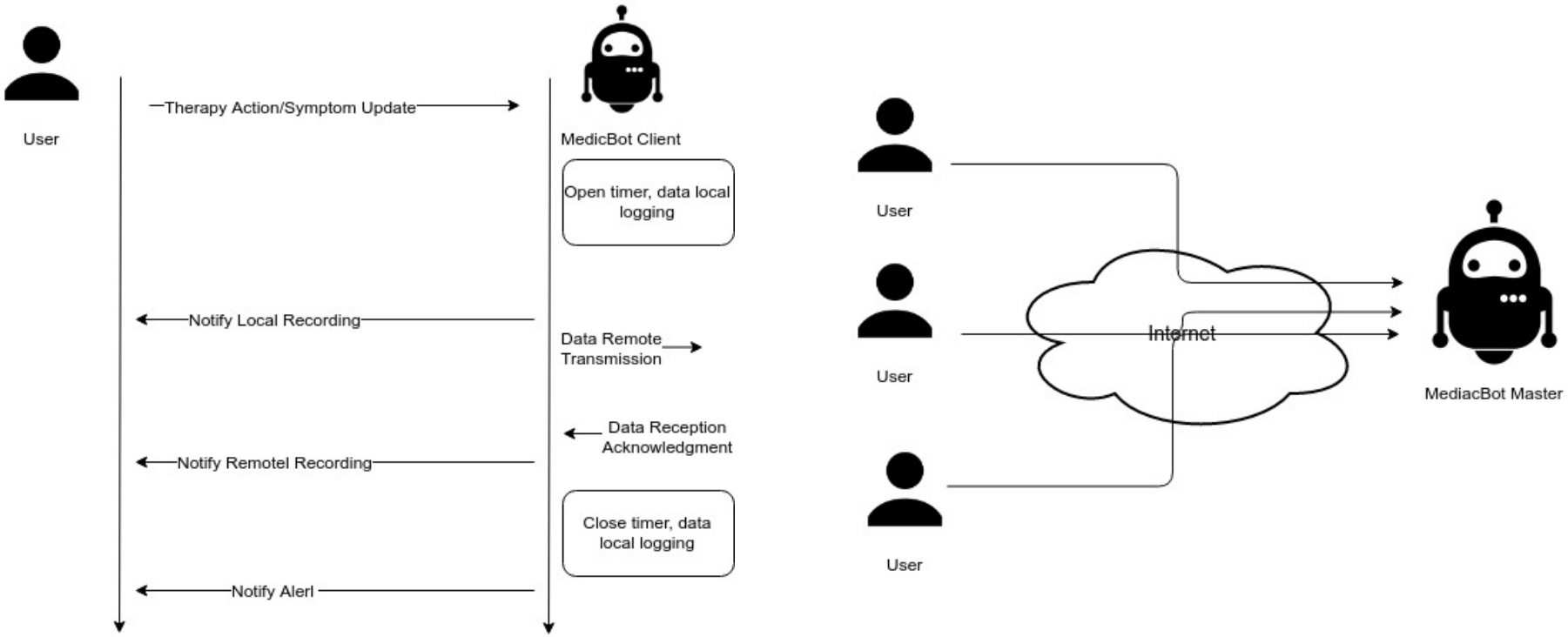


Describe the score of the training for the respective user at home

	Observation	Assessment	Therapy lesson	Other
Listening				
Speaking				
Phonological Awareness				

Table: Benchmark table

Client and Master MedicBot



MedicBot Client collects data from User and then send them to the MedicBot Master to store and analysis more (if the problem needs specialists)

Conclusion

- MedicBot is a new approach for building a Medical assistance for Auditory Processing Disorder (APD) children.
- In this framework, the text analysis is of paramount importance and its strategy should be learned from data and adapted online, to address the possible evolution of the APD problem.
- Many challenges remain to build a complete system based on this proposition such as: out-of-domain handling applications overestimating their scores, adaptation to a specific user , non stationary usages of sets of application and more generally co-adaptation belong to our concerns.

Thank You