

CSE704 Seminar

Training-Bot

Team Members:

Lu Dong(ludong@buffalo.edu)

Aparna Ananthakrishnan (aparnaan@buffalo.edu)

Adhesh Garg (adheshga@buffalo.edu)



Background

Video Volunteers

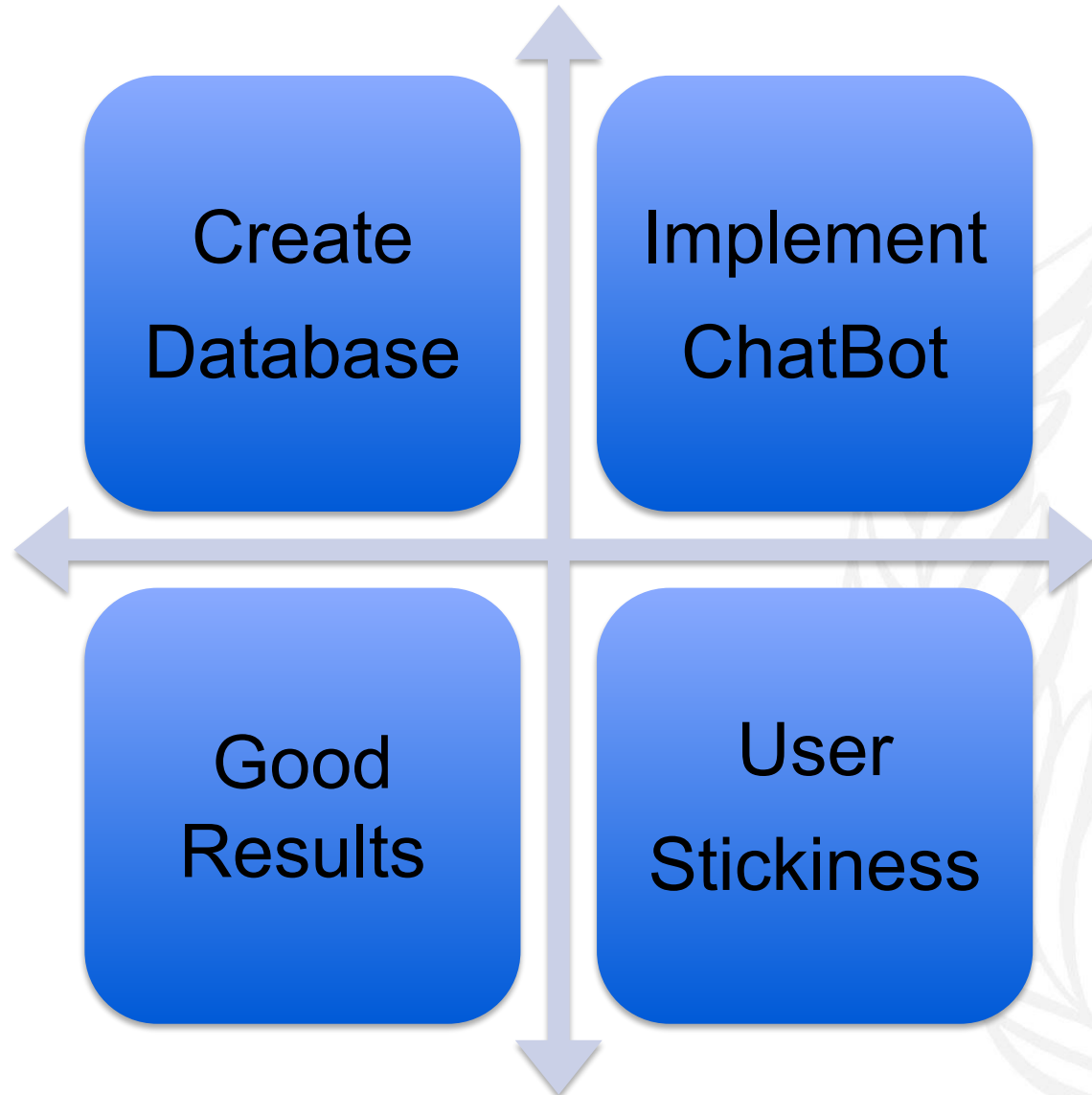
empowers marginalised people to tell their stories and create change campaigns, so that their issues come out from under the rug and become important threads in India's development narrative.



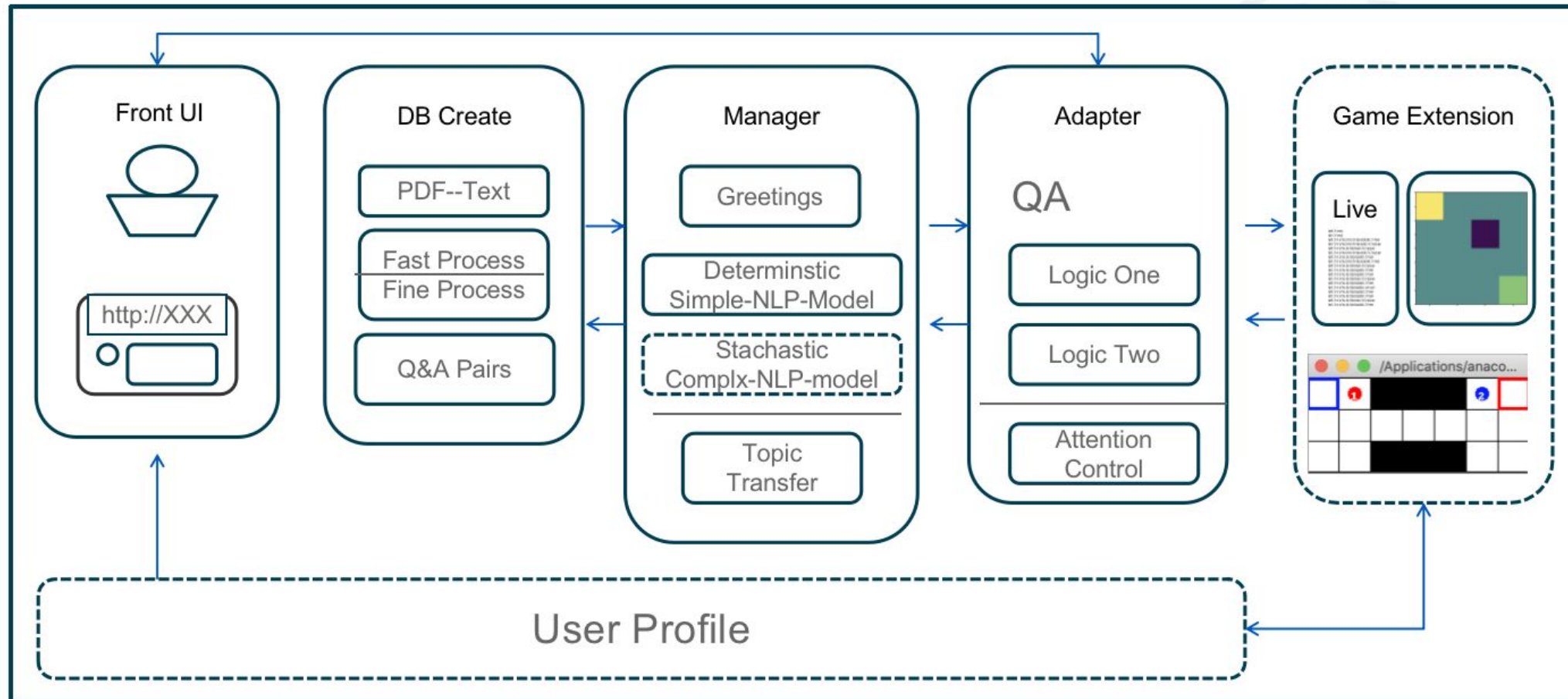
Purpose: Retrieval based bot, that can converse with new recruits as well as existing Community Correspondents & answer FAQ.

Data Sources: Training manuals, Video manuals, Related technology manuals from internet, video4change.org, Online resource, etc.

Challenges

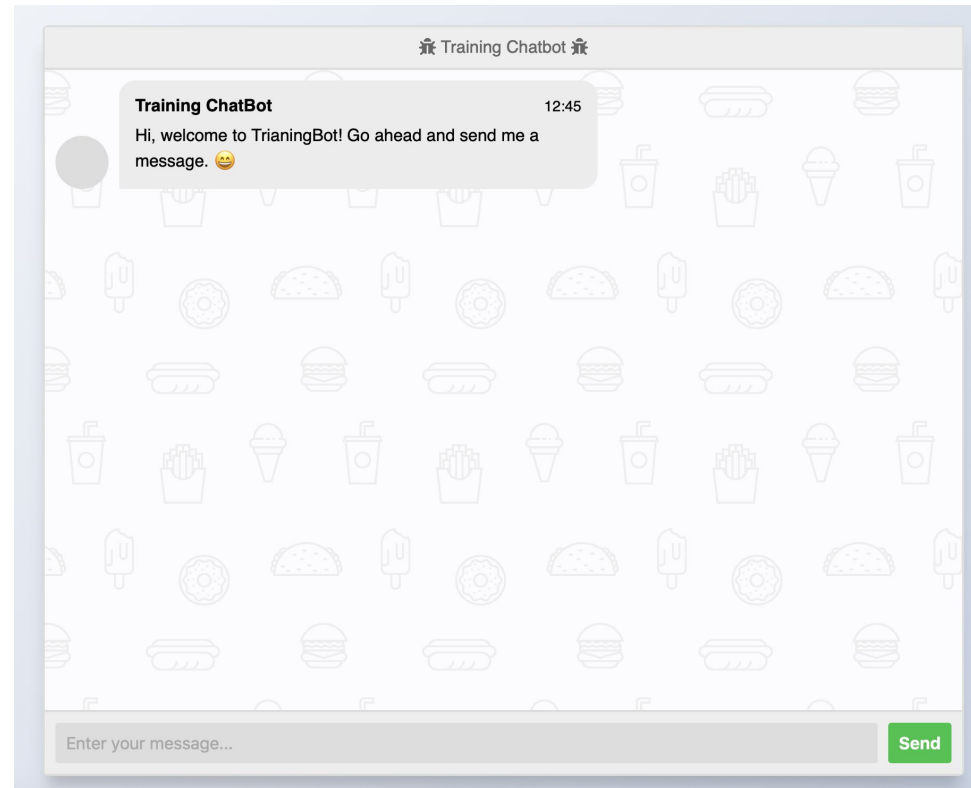
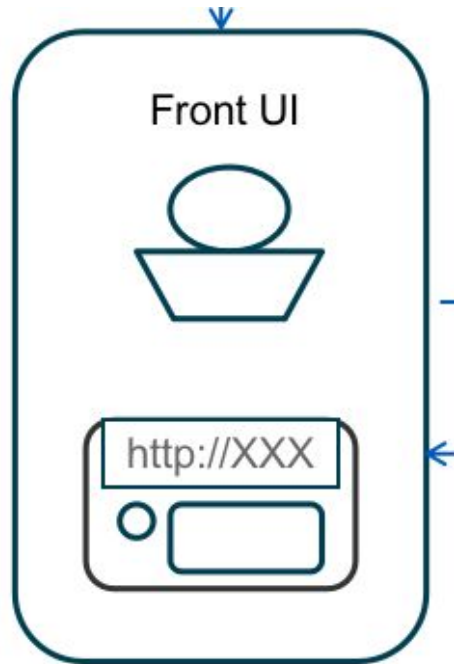


Overview Design for Training-Bot



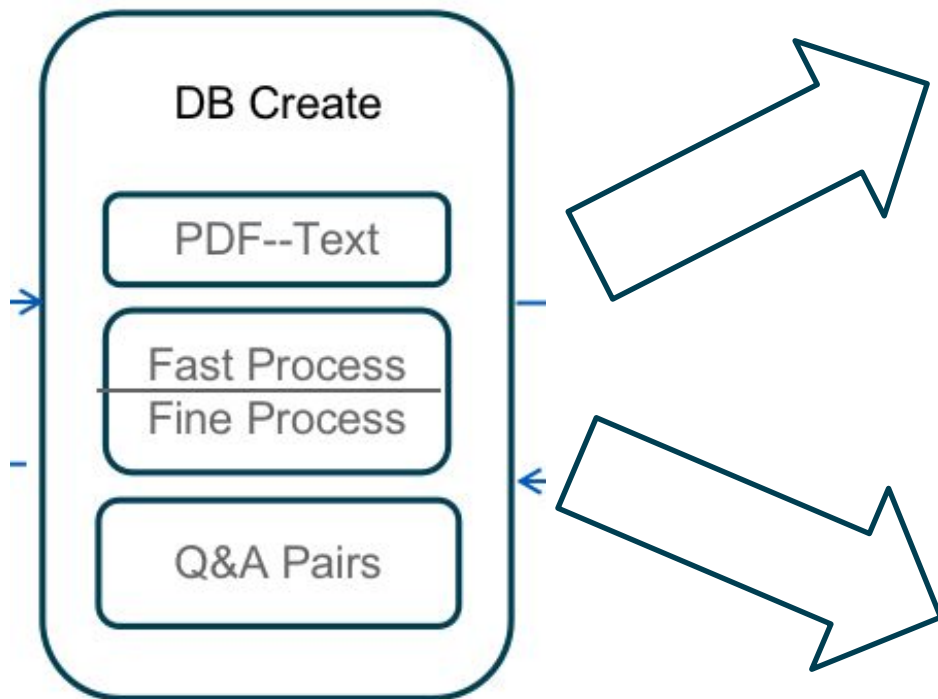
Front UI

`http://127.0.0.1:5000`



Implement with Flask
which is a micro web
framework written in
Python.

DataBase



Manual Processing

Currently Test Good

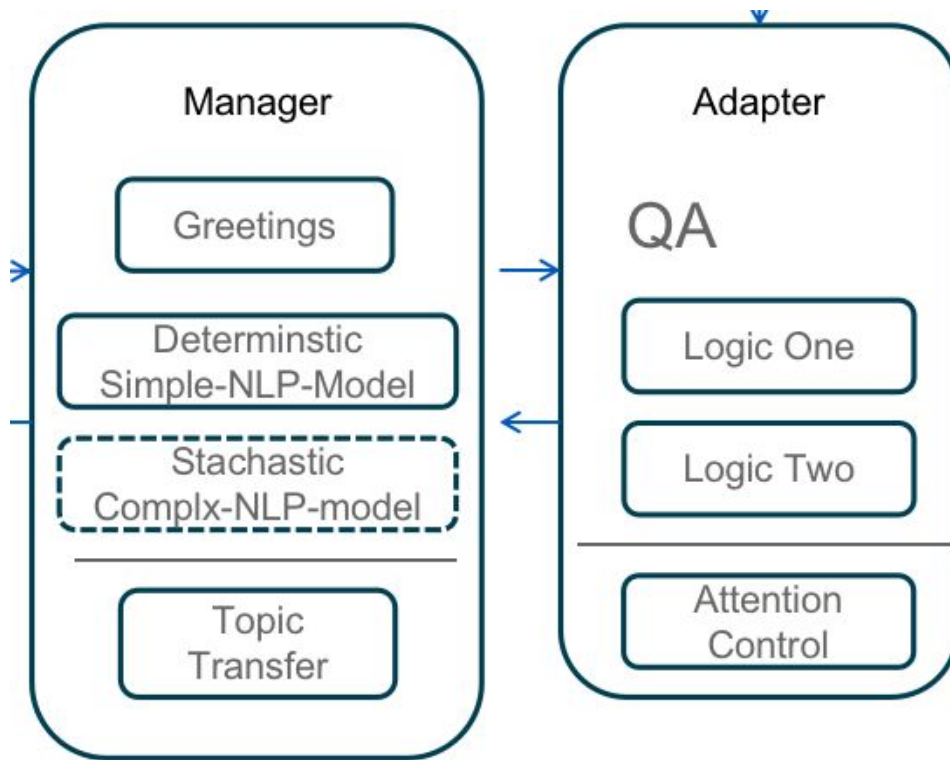
Automatic Processing

Text: XXX	Sign: 1
Text: XXX	Sign: 0
Text: XXX	Sign: 0
Text: XXX	Sign: 1
Text: XXX	Sign: 1
Text: XXX	Sign: 0

QA_Pairs

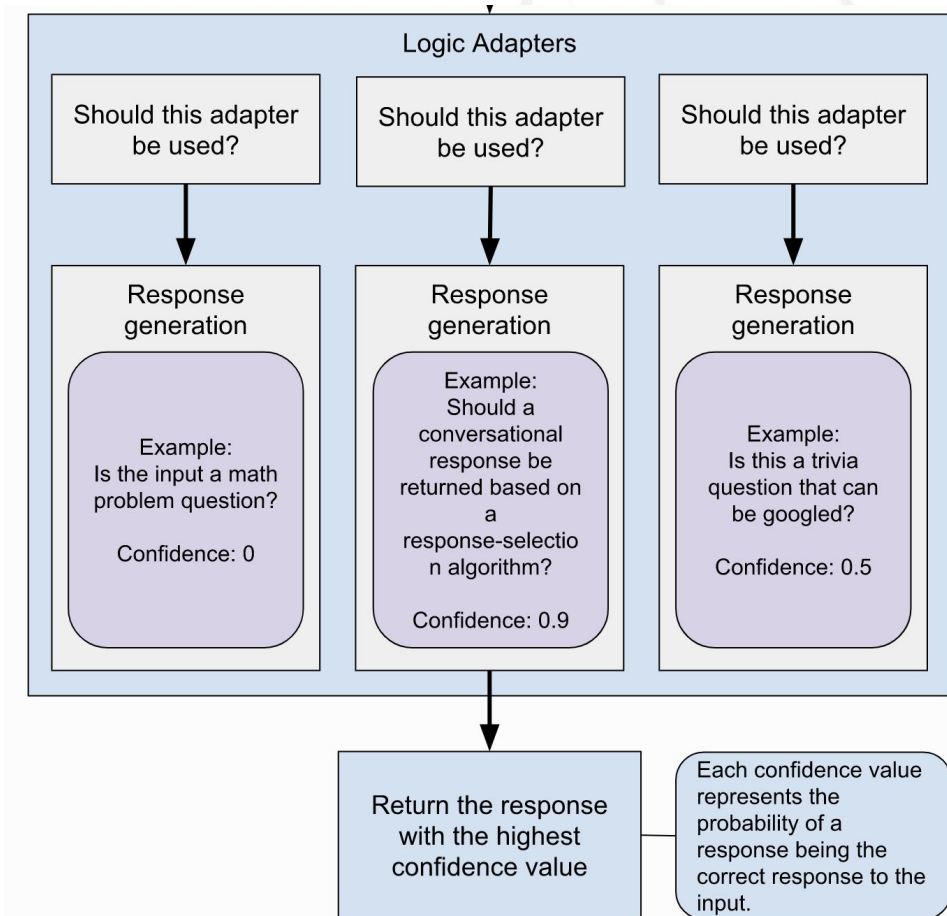
Results are not that Good.

Manager-Adapter

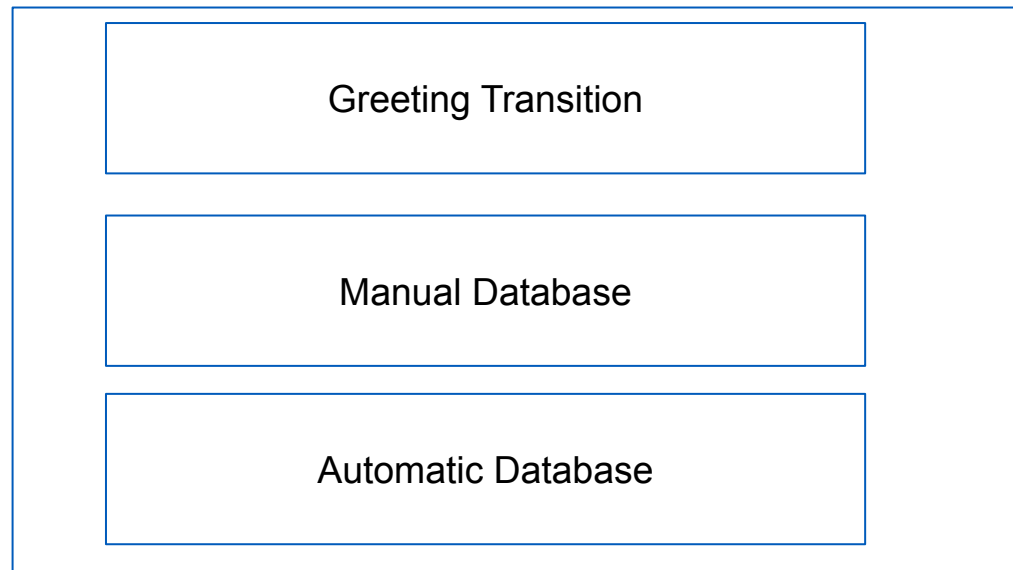


<https://chatterbot.readthedocs.io/en/stable/quickstart.html>

ChatterBot library to generate automated responses to a user's input.



Our Adapter Model



Generate Answer
based on highest scores

Part One Instructions and Demo Presentation

Our database follows format at right side;

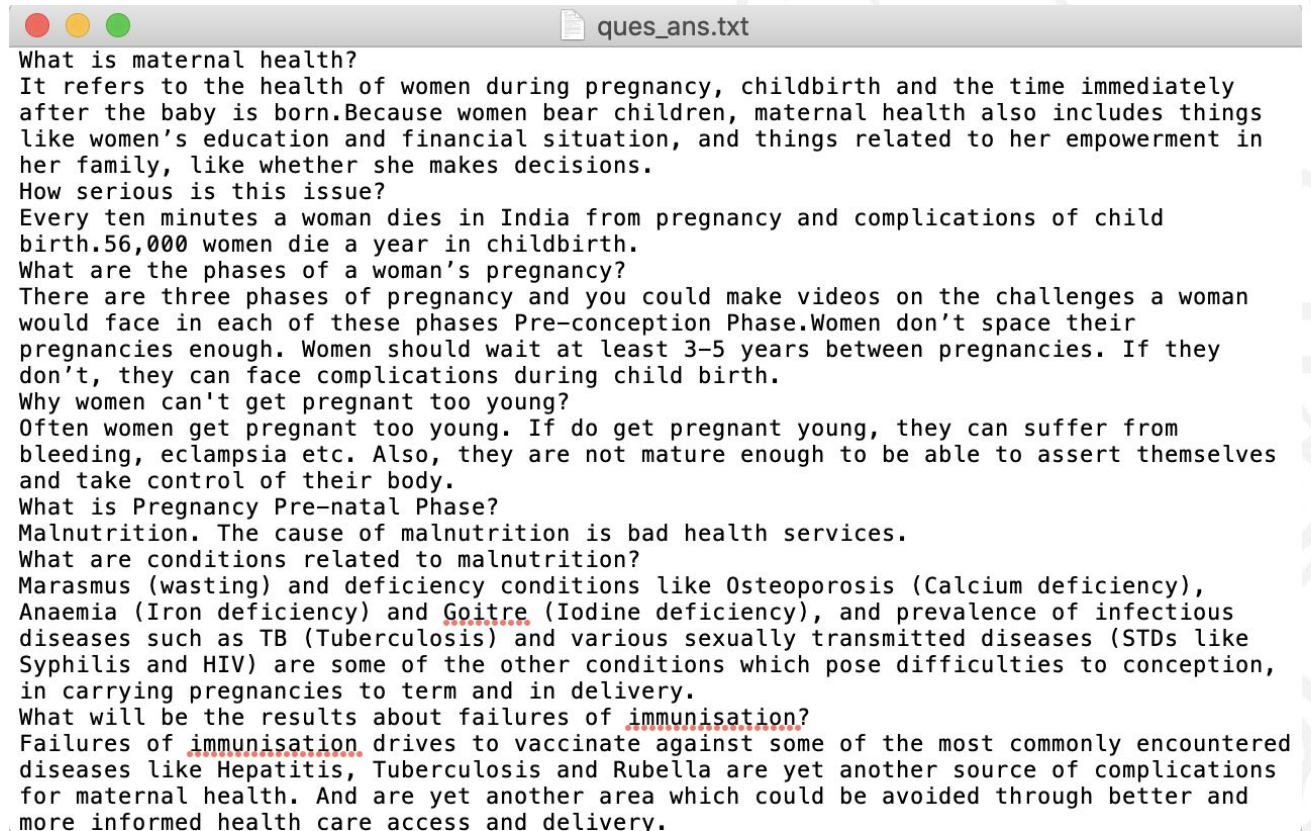
One line question;

One line answer;

If the database can as good as this, it will work very well.

Here we manually deal with one pdf;

http://www.videovolunteers.org/wp-content/uploads/2017/03/14_Community-Monitoring-of-Maternal-Health-English-2015.pdf



```
ques_ans.txt
What is maternal health?
It refers to the health of women during pregnancy, childbirth and the time immediately after the baby is born. Because women bear children, maternal health also includes things like women's education and financial situation, and things related to her empowerment in her family, like whether she makes decisions.
How serious is this issue?
Every ten minutes a woman dies in India from pregnancy and complications of child birth. 56,000 women die a year in childbirth.
What are the phases of a woman's pregnancy?
There are three phases of pregnancy and you could make videos on the challenges a woman would face in each of these phases Pre-conception Phase. Women don't space their pregnancies enough. Women should wait at least 3-5 years between pregnancies. If they don't, they can face complications during child birth.
Why women can't get pregnant too young?
Often women get pregnant too young. If do get pregnant young, they can suffer from bleeding, eclampsia etc. Also, they are not mature enough to be able to assert themselves and take control of their body.
What is Pregnancy Pre-natal Phase?
Malnutrition. The cause of malnutrition is bad health services.
What are conditions related to malnutrition?
Marasmus (wasting) and deficiency conditions like Osteoporosis (Calcium deficiency), Anaemia (Iron deficiency) and Goitre (Iodine deficiency), and prevalence of infectious diseases such as TB (Tuberculosis) and various sexually transmitted diseases (STDs like Syphilis and HIV) are some of the other conditions which pose difficulties to conception, in carrying pregnancies to term and in delivery.
What will be the results about failures of immunisation?
Failures of immunisation drives to vaccinate against some of the most commonly encountered diseases like Hepatitis, Tuberculosis and Rubella are yet another source of complications for maternal health. And are yet another area which could be avoided through better and more informed health care access and delivery.
```

Part One Demo and Instructions

Download vscode;

Open the file with vscode;

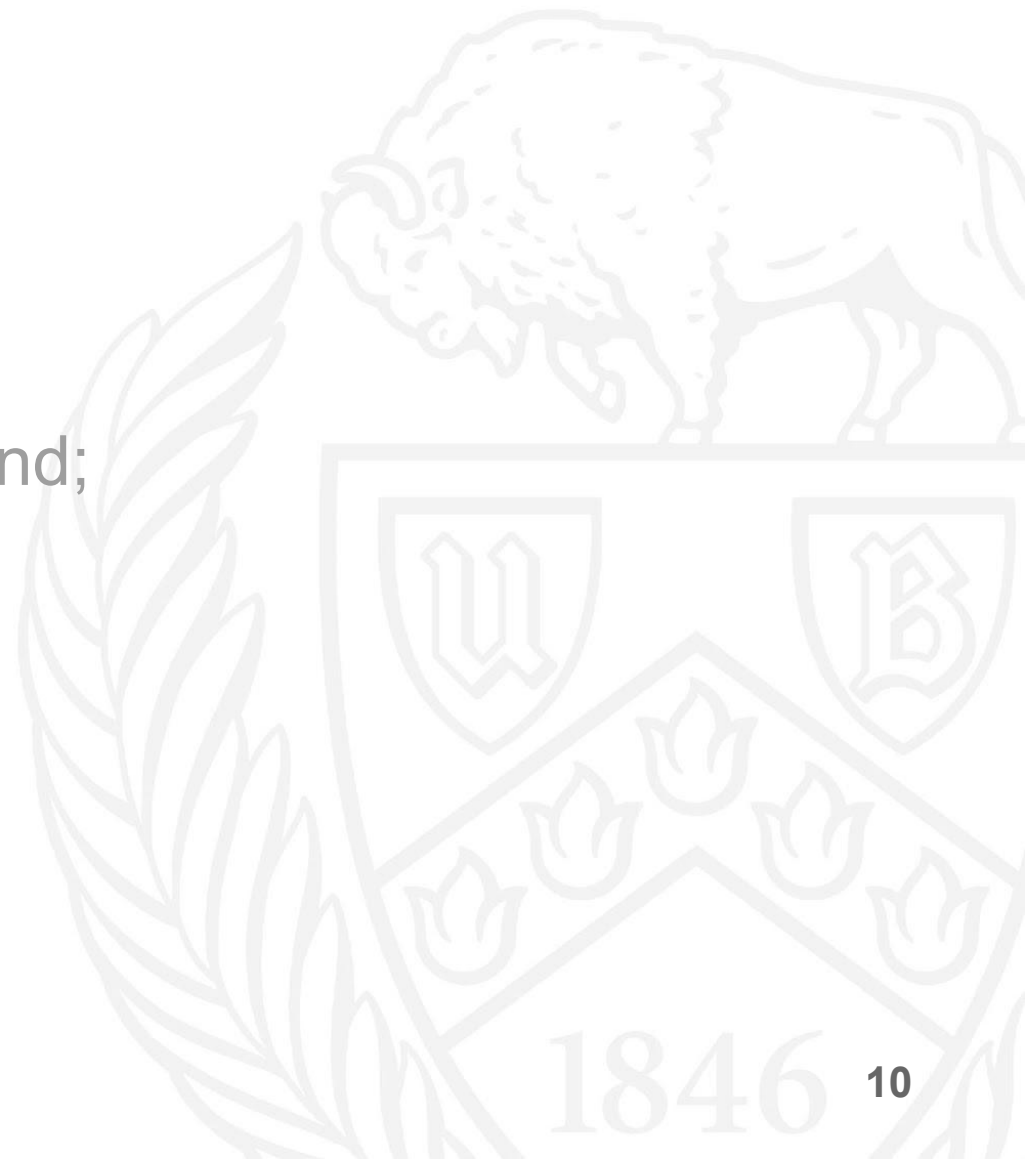
open terminal and run the following command;

source evn/bin/activate

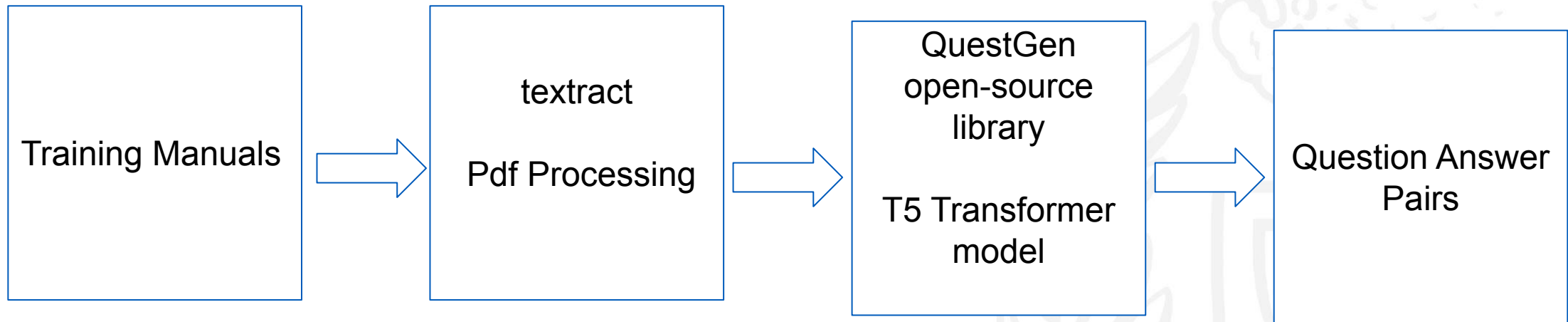
python3 app.py

open <http://127.0.0.1:5000>

start chat



Part Two: Data Annotation



Abstractive/Extractive
Summarization T5,BART,Pegasus
Models



Paraphrasing of sentences
T5,GPT2

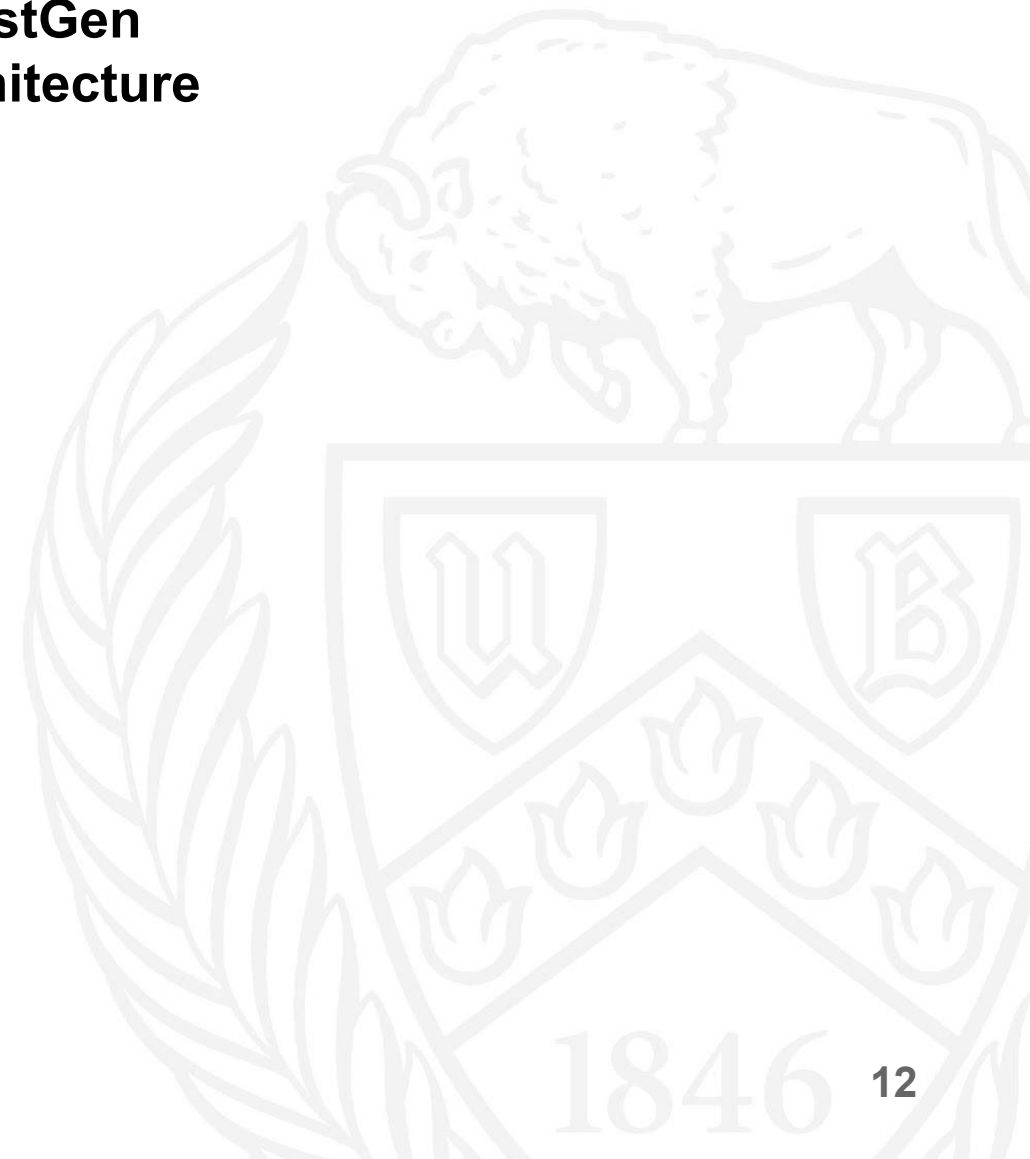


Keyword/Keyphrase extraction
YAKE,TopicRank,KeyBERT



Question Generation
T5,GPT2,GPT3

QuestGen Architecture



Running model for generation

```

{'questions': [{'Question': "What should you make to show that it's not just you but you & community?", 'Answer': 'videos'}],
{'questions': [{'Answer': 'videos',
  'Question': "What should you make to show that it's not just "
    'you but you & community?',
  'context': 'Put bites into your videos like, 'if we are '
    'informed we can do anything!' then your viewers '
    'will see others supporting you and they will be '
    'inspiring • If the community are asking for money, '
    'stopping you, refusing to take your calls - it is '
    'a sign you are doing a bad job in building '
    'community support in your area & you need to find '
    'out what you are doing wrong. • Make videos that '
    'convey a 'collective call to action', so your '
    'videos show that it's not just you but you & '
    'community. • Make videos that convey a 'collective '
    'call to action', so your videos show that it's not '
    'just you but you & community.',
  'id': 1}],
{'Answer': 'media',
  'Question': 'What is the main reason IU is a national network?',
  'context': 'Unlike mainstream media we are interested in the '
    'issues of the most marginalized communities and. • '
    'Explain IU -we're a national network set up '
    'precisely because media is so bad for us. • Tell '
    'them what's wrong with mainstream media and how we '
  
```



Parl AI

A unified platform for training and evaluating dialogue models across many tasks

Data Format

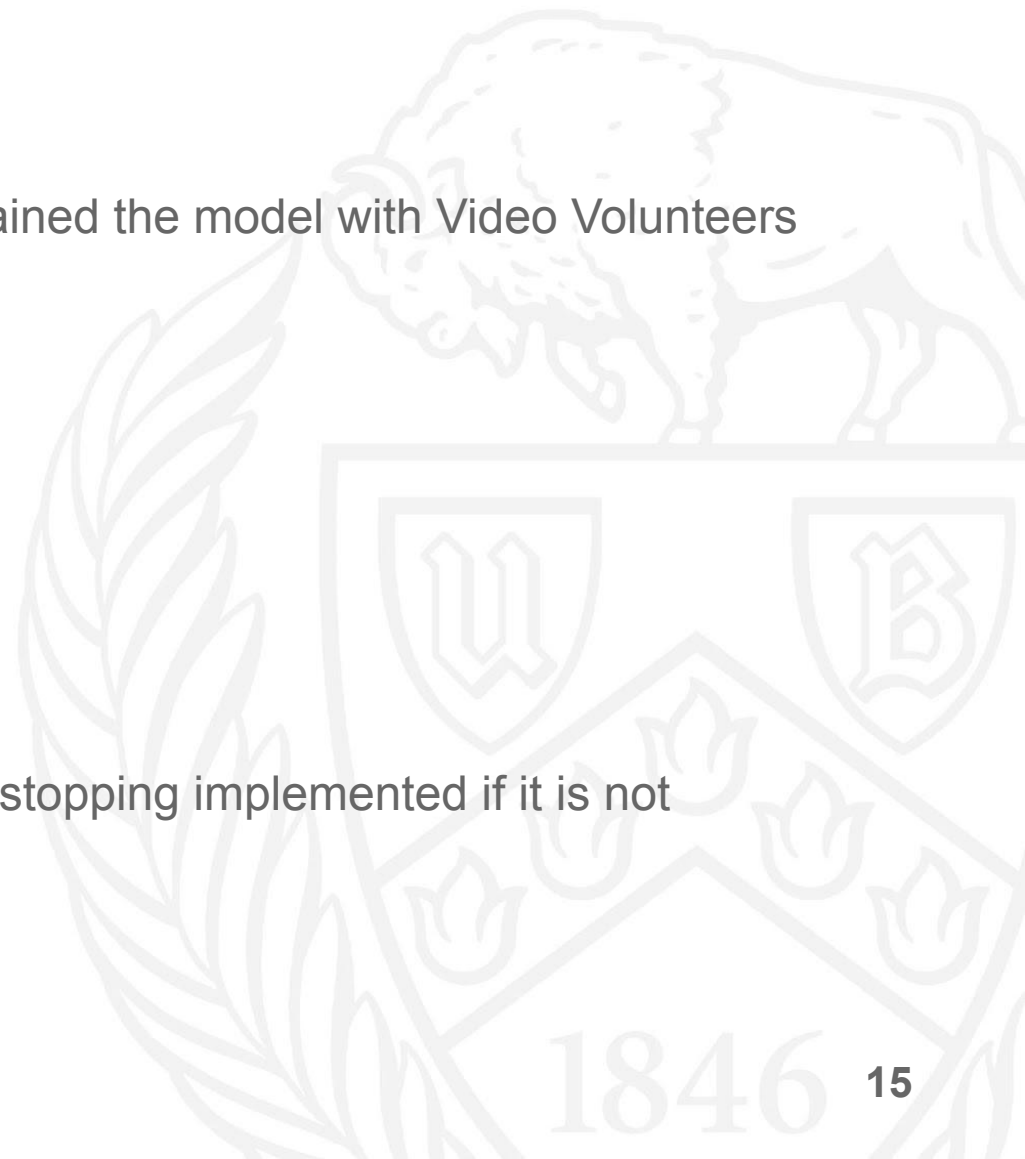
```
text:hello how are you today?  labels:i'm great thanks! what are you doing?  
text:i've just been biking. labels:oh nice, i haven't got on a bike in years!  episode_done:True
```

Training

We have used base model as Pretrained Blender bot 90M and Trained the model with Video Volunteers data using ParlAI by Facebook Research.

Hyperparameters:

1. We have used Adam optimizer for the same
2. We used Gaussian Error Linear unit activation
3. Validation metric used is Perplexity where there is also early stopping implemented if it is not reducing.
4. PPL started with 352.8 and after training it was



Predictions from Model

Who is the star of your story?

labels: • Your video SHOULD HAVE ONE MAIN CHARACTER. • You can have more than one character ONLY if the story absolutely requires it. Other characters will be 'supporting characters' and they should all be connected to the main character. • Spend some time getting to know your character and the environment before filming. What do they do? How do they do it? Where do they go? What is happening? What is interesting about the character?

model: the star of the story is the character who is in the process of getting into a relationship with the main character .

- - - NEW EPISODE: content/data.txt- - -

why reporting on these cases in an evidence based manner with accuracy is important?

labels: gender project design with select group of ccs where ccs are oriented to conduct discussion clubs using videos produced on instances of everyday patriarchy as well gender based violence these discussion clubs aim to create a discourse around the need to dismantle patriarchy within their communities

model: the reporting on these cases is usually done in a reporting format that is not necessarily based on the facts .

- - - NEW EPISODE: content/data.txt- - -

q2 why are we challenging patriarchy?

labels: male gender based violence physical verbal or emotional is a manifestation of patriarchythis phenomenon is rooted in the idea that men are endowed with the power and right tocontrol and if required discipline women hence to prevent gender based discrimination andviolence one needs take on its root cause patriarchy

model: patriarchy is a social construct that is rooted in the belief that men are the root cause of all of this

- - - NEW EPISODE: content/data.txt- - -

q 3 how do i produce a video or mentor a cc to produce a video on gender?

labels: read this production and field guide from end to end this guide is aimed at producing videosthat are different from the iu format hence ccs and mentors need to approach these storiesdifferently this guide will help you through identifying story ideas give you tips on filmingreporting checklists and several short notes that will help you produce videos that chronicle

model: 1) ask your cc to mentor you 2) ask them to do so 3) ask them to do so

- - - NEW EPISODE: content/data.txt- - -

how patriarchy plays out in different areas of life please read the entire story idea right down?

labels: to filming tips before starting out do not skip directly to filming tipswile reading this guide you will find many story ideas that they would like to film

model: read the entire story idea right down

- - - NEW EPISODE: content/data.txt- - -

q 4 how do i plan and approach videos on gender based violence?

labels: khel badal dismantle patriarchy 4 www videovolunteers orgthis guide also aims at ensuring that you report on gender based violence with sensitivityaccuracy stories on gender based violence should go beyond r eporting the act of violenceand examine the effectiveness of recourse mechanisms for survivors remember to read allsections on protocols for reporting and intervention rights of a survivor protocols for policeand doctors and several reporting checklists to make good videos on gender based violence

model: ask the community members for tips on how to approach videos on gender based violence .

- - - NEW EPISODE: content/data.txt- - -

q 5 i have identified a story idea who do i pitch it story to or discuss it with?

labels: gender project story ideas should not be discussed with your field mentors cc mentors butwith your state coordinator or stalin or nupur or manish or rajkumarnote for state coordinators please ask stalin or nupur for guidance if you find it difficult tomentor a story do this only after you have the one liner which is who where when andwhat

model: ask the questions you need to ask the audience and ask them what they think the story should be about .

- - - NEW EPISODE: content/data.txt- - -

q 6 do the gender videos follow the iu format?

labels: yes and nothis guide will equip you to produce gender videos under two categories1 everyday patriarchy2 gender based violencevideos that chronicle everyday patriarchy will n ot follow the iu formatthey don t need to have an opening p2c or ctathey will have vos and a closing p2c the closing p2c will not be an appeal but

model: yes they do and they are always interviewable in the same format

Webhooks chat Interface from ParlAI

We have also used ParlAI to host a chatbot using webhooks that is also based on Blender Bot 90M. We have currently hosted this model with the help of AWS on a public IP. ParlAI also has provision to increase the number of workers for the same which makes it easy for users or Video Volunteers in this case of use this library for their chatbot needs.

ParlAI[1] paper was release by Facebook research in 2018 and this since then this is an active repository on Github with over 160 active contributors. They are constantly improving this library.

Conclusions

- We have better understanding about chatbot.
- We attempt multiple methods to deal with PDF to Database.
- We explore Bayesian neural network model achieving good results and propose to use game strategy to improve user stickiness.
- We also explore the-state-of-art NLP model and also achieve good performance.

References

[1] ParlAI A. H. Miller, W. Feng, A. Fisch, J. Lu, D. Batra, A. Bordes, D. Parikh, and J. Weston. Parlai: A dialog research software platform. *arXiv preprint arXiv:1705.06476*, 2017.

<https://arxiv.org/abs/1705.06476>

