



MVSR Engineering College, Hyderabad - 501 510



Department of Computer Science and Engineering

Project Title : Question Answering System in Telugu

Batch.No:

2451-18-733-128 G. Sreehith Reddy

2451-18-733-129 A.V.S.Shirish

2451-18-733-136 D.Ganesh

Under the guidance of

Name : Venkata Ramana Bhattula

Designation : Assistant Professor

Problem Statement :



- 60% Indian language internet users stated limited language support and content to be largest barrier for adoption of online services.
- Question answering systems have been extended in recent years to encompass additional domains of knowledge. But, there is no advancement in native languages.
- So, building Question answering system in a native language like telugu would be very much useful

Objective :



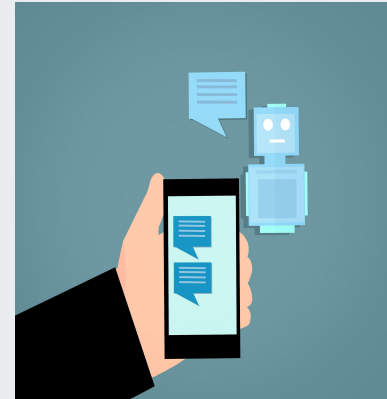
- Construction of a robust program that predicts relevant answer for the given question according to the context.

Project scope :

- Conversion of Telugu language into encoded format using BERT and giving the prediction of the task.
- Presently our project is limited to prediction of answer from given raw text of minimum size but could be modified by supplying documents of large sizes.

Applications :

- Personal Assistants
- Search Engines
- Chat Bots
- E-Learning



Requirements :

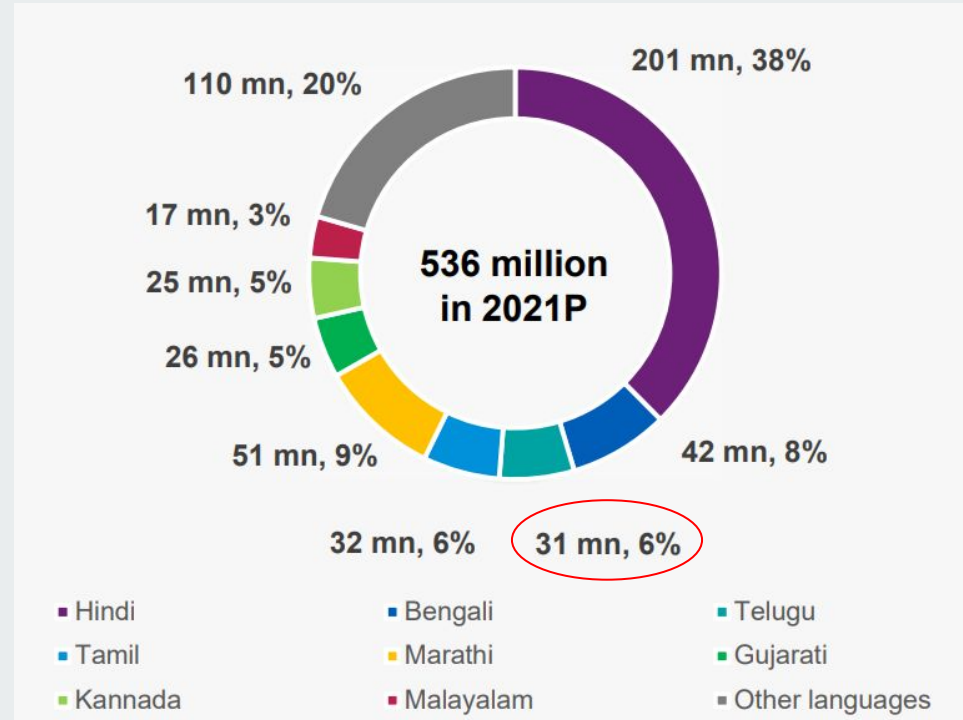
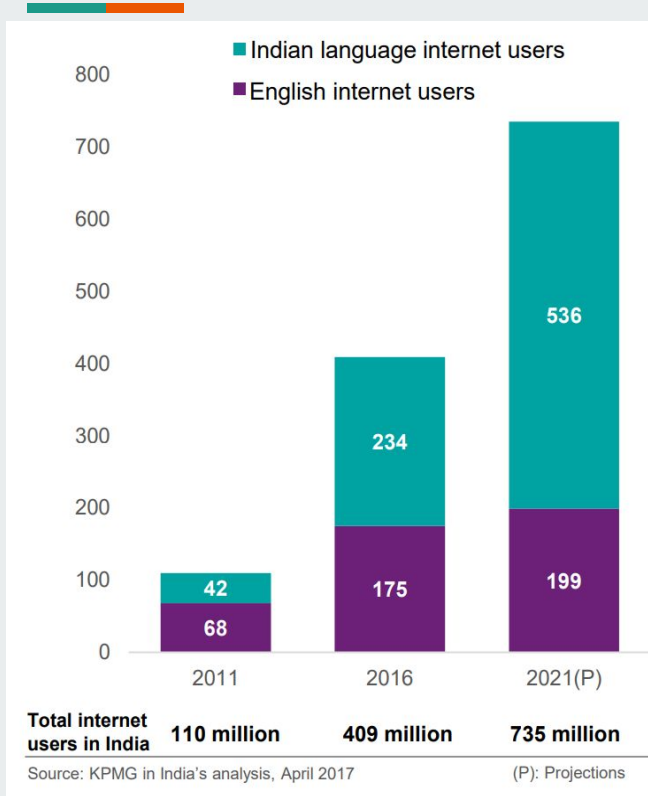
- Related Python libraries
- Pretrained cross lingual NLP model
- Tokenizer
- High-end RAM and GPU enabled machine



Challenges :

- Telugu Language Corpus is limited over the internet.
- Model Selection.
- Ambiguity

Indian Language Sources on Internet :



A Taxonomy of QA Systems:

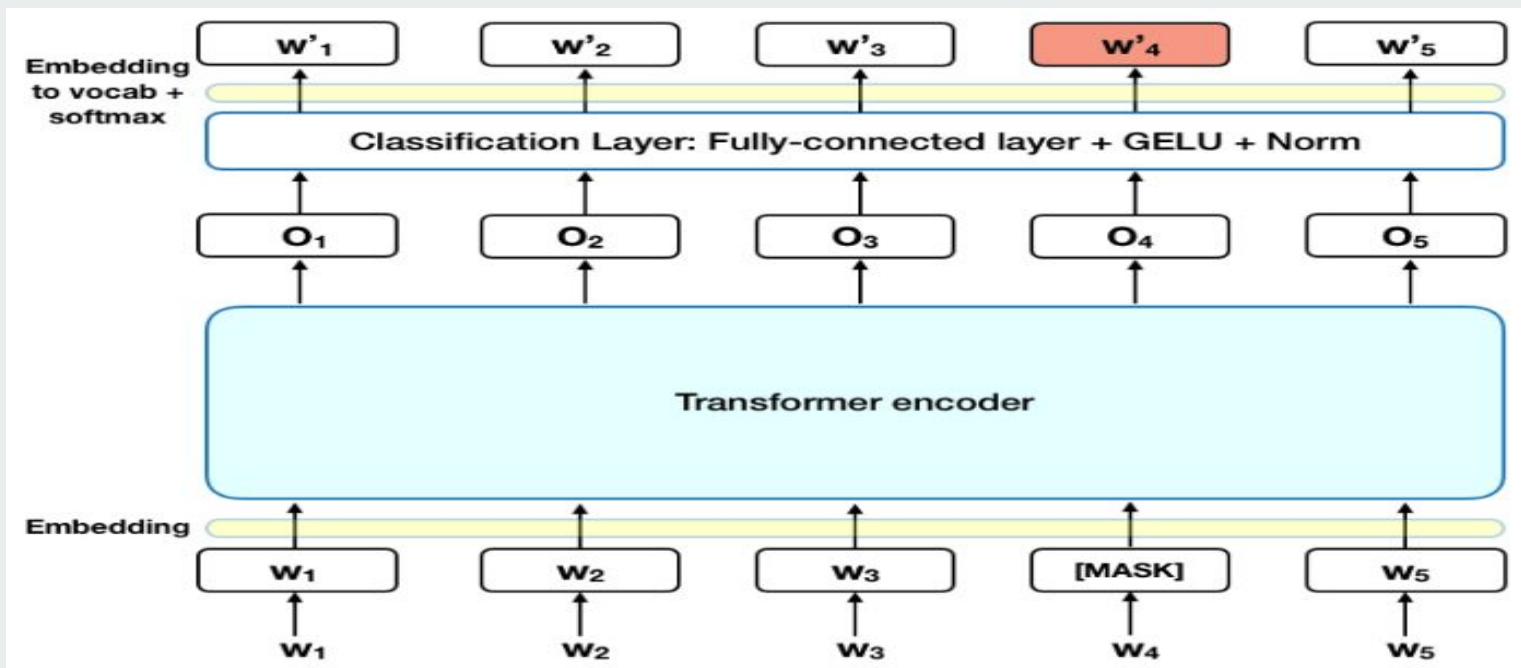
Class	Type	Example
1	Factual	Q : భారతదేశ రాజధాని <u>ఏమిటి</u> ? A : భారతదేశ రాజధాని <u>న్యూ డిల్లీ</u>
2	Simple Reasoning	Q : సోక్రటీస్ <u>ఎలా</u> చనిపోయాడు? A : <u>విషం</u> కారణంగా చనిపోయాడు.
3	Boolean (True / False)	Q : సూర్యుడు పశ్చిమాన <u>ఉదయిస్తాడా</u> ? A : తప్పు
4	Multiple Choice	Q: ఇంద్రధనస్సుకు <u>ఎన్ని రంగులు</u> ఉన్నాయి? 1) 4 2) 5 3) 6 <u>4) 7</u>
5	Opinion Based	Q : 2018 ఎన్నికలు న్యాయంగా జరిగాయో?

BERT (BiDirectional Encoder Representations From Transformers)



- Recently (2018) published paper by researchers at Google AI .
- Makes use of Transformer , a mechanism that learns contextual relation between words in a text.
- Has two mechanisms which includes, encoder that reads input and decoder that produces prediction of the task.
- As the name suggests it is Bidirectional which means reads texts sequentially (left- to-right or right-to-left)
- Many variants of BERT have been reproduced like roBERTa by Facebook, DistilBERT, IndicBERT.

Working of BERT :



Working Example :

Passage :

M.V.S.R 1981లో ఉస్మానియా యూనివర్సిటీకి అనుబంధంగా ఎంపీఎస్ఆర్ ఇంజనీరింగ్ కళాశాల స్థాపించబడింది. ప్రస్తుతం, కళాశాలలో MBAతో పాటు సివిల్, CSE, ECE, EEE, IT, మెకానికల్ మరియు ఆటోమొబైల్ ఇంజనీరింగ్లో B.E ప్రోగ్రామ్లు, CSE, ECE మరియు మెకానికల్లో PG ప్రోగ్రామ్లు ఉన్నాయి. ఈ కళాశాలను ఉస్మానియా విశ్వవిద్యాలయం పిహెచ్డి అవార్డుకు దారితీసే పరిశోధనలను కొనసాగించే కేంద్రంగా గుర్తించింది. CSE, ECE మరియు మెకానికల్ ఇంజనీరింగ్ విభాగాలలో డిగ్రీ. మొత్తం మీద ఏడాదికి 1000 మంది విద్యార్థులు అడ్మిషన్ తీసుకుంటారు. ఈ కళాశాల రాష్ట్రంలోని ప్రైవేట్ ఇంజనీరింగ్ కళాశాలల్లో మొదటిది, మరియు కళాశాల ప్రారంభం నుండి అద్భుతమైన మరియు ఆదర్శప్రాయమైన ప్రమాణాలను నిర్ధారిస్తుంది, ఇది కళాశాలను రాష్ట్రంలోని అత్యున్నత మరియు కోరుకునే కళాశాలలలో ఒకటిగా నిలిపింది.

Question:

M.V.S.Rలో PG కోర్సులు ఏవి అందించబడతాయి?

Answer:

M.V.S.Rలో CSE, ECE మరియు మెకానికల్ PG కోర్సులు అందించబడతాయి.

Q & A

