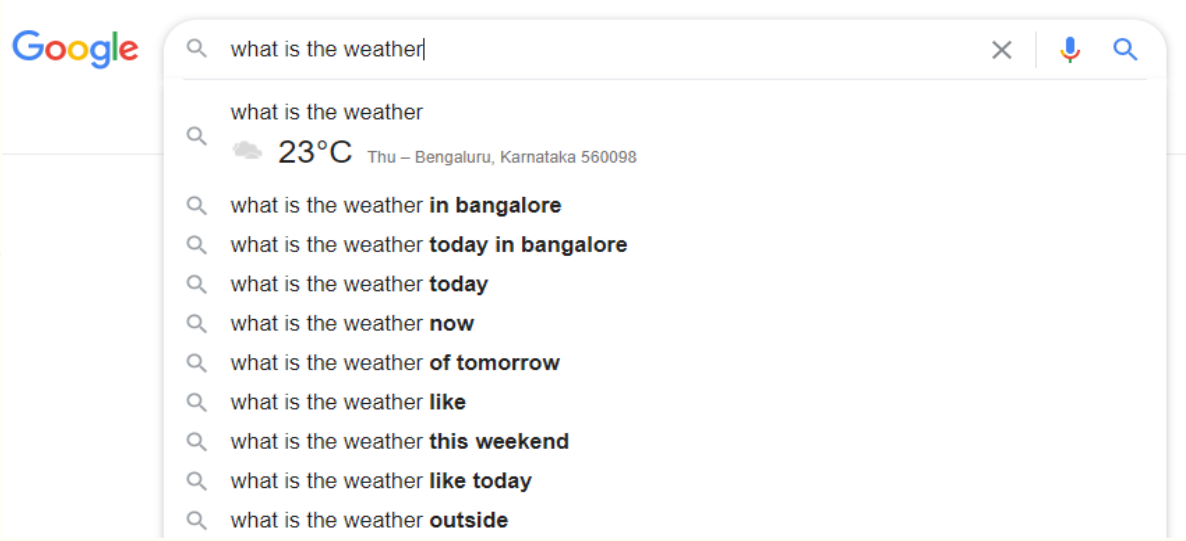
NEXT WORD PREDICTION USING NLP TECHNIQUES



Presented by
Rajashri Chitti

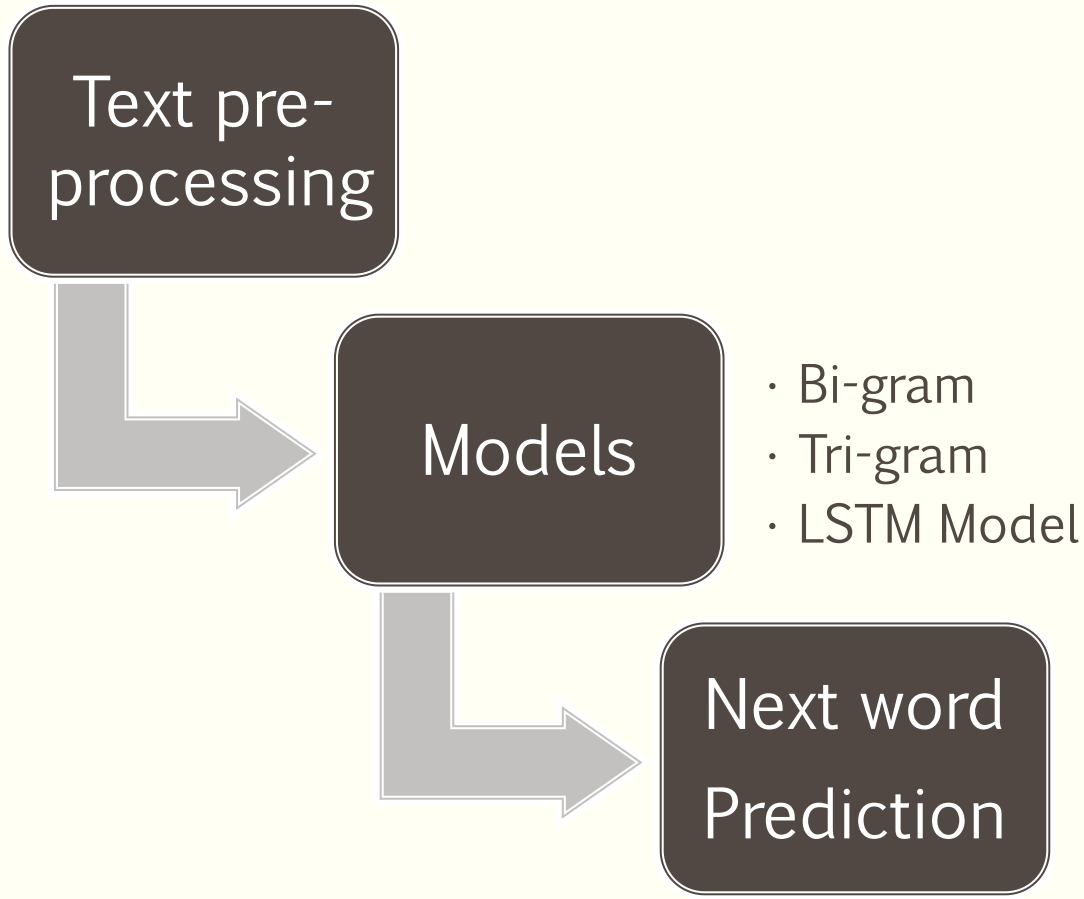
Content

- ✓ Introduction
- ✓ Methodology
- ✓ Flow of the project
- ✓ NLP Techniques/libraries used in the project
- ✓ Screenshot of Output
- ✓ Conclusion

Introduction

- Next Word Prediction is also called Language Modeling that is **the task of predicting what word comes next**.
- It is one of the fundamental tasks of NLP and has many applications.
- Word prediction is an assistive technology tool that **suggests words while a child types**. Kids with writing issues can benefit from this technology. You may have used word prediction on a smartphone, but more advanced tools also exist.
- This can be done by using RNN, HMM, or N-grams model.

Methodology



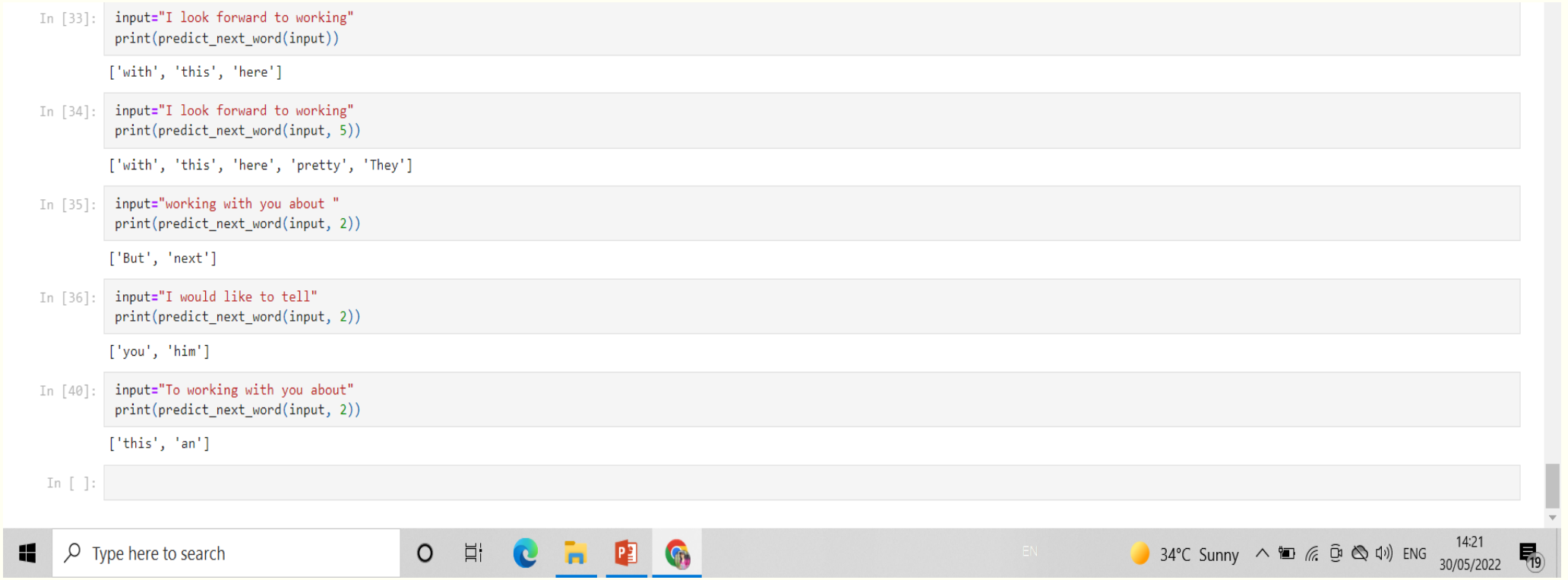
Flow of the project

- Reading the text file(alice.txt and train.json)
- Tokenization
- Use the pre-trained smoothed models that were formed without removal of stop words.
- Using n-grams model (Bi-gram and Tri-gram) of the strings to get the next word predictions based on the highest probabilities.
- Return the most probable word
- Build LSTM(Long Short Tem Memory Model), train and apply.
- Output : Next word will be predicted for Bi-gram model, Tri-gram model as well as LSTM(Long Short Memory Model).

NLP techniques/libraries used in the project

- NLTK library
- Tokenization
- Pre-trained smoothed models for stop of words
- Bi-gram model
- Tri-gram model
- LSTM Model

Screenshots of output



Conclusion

- In this project, "**Next Word Prediction**" is done using Natural Language Techniques and Machine learning models.
- Initially, **N-gram models** is used to predict the next word, also Deep learning model i.e. **LSTM(Long Short Memory Model)** is used for "**Next Word Prediction**".
- Models are tested on two dataset which are publically available on kaggle(alice.txt and train.json).
- The manual prediction can also be tested to predict next word by user.

THANK YOU!