MACHINE TRANSLATION

DR. DINKAR SITARAM

GREESHMA KARANTH, SHIVANGI GUPTA, SUHAAS K, VISHAL KANTEPPA

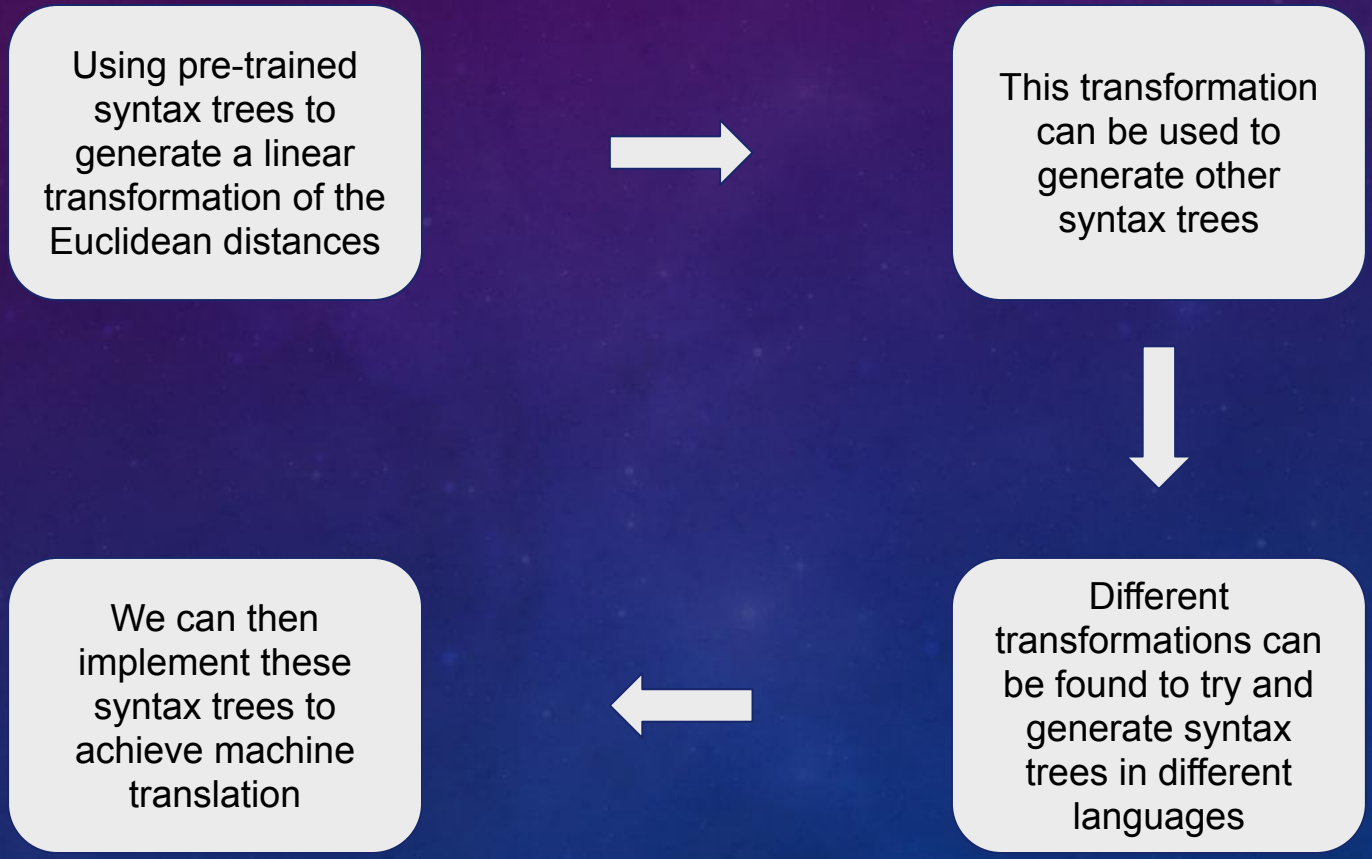
OBJECTIVE OF THE PROJECT

- This project aims to translate Kannada text to Telugu text by using structural probe for machine translation
- The objective for this semester is to train the linear transformation matrix for transforming a sentence and being able to generate a syntax tree

CONFIGURATION/INFRASTRUCTURE

- python 3.7
- BERT (Bidirectional Encoder Representations and Transformers) language model of embeddings

SOLUTION ARCHITECTURE



CURRENT PROGRESS

We have tried to generate the transformation matrix for a random sentence by using gradient descent

- First, the transformation matrix (B) and the syntax_tree is randomly populated with values
- Then, we use BERT Embeddings, which is a module that takes a word and generates its vector representation in the n-dimensional space, say h
- This helps us calculate the transformed distance between each word of a corpus, here, a sentence
- This transformed distance is calculated by taking the inner product of the transformed vector embeddings, i.e., $(Bh)^T(Bh)$
- A randomly initialized syntax tree helps us calculate the tree distance for each pair of words
- We then take a difference of the transformed distance and tree distance and sum it over the entire parsed corpus
- This difference is then minimized using gradient descent to obtain the transformation matrix

INITIAL RESULTS AND FUTURE SCOPE

We start with randomly initialized values of parameters. On running the code, we can observe that gradient descent is working as the error moves over different values to attempt to arrive at the optimum.

After implementing the code, we can see that this gradient descent reaches a minima but doesn't converge with the given configuration. The gradient descent is observed to overshoot the minima.

We intend to fix this and try gradient descent using numpy arrays entirely. We also want to implement stochastic gradient descent to get better results as the code takes too long to run currently.

With the structural probe gained, we will then be able to get syntax trees in different languages and therefore attempt a translation from one language to another.

LITERATURE SURVEY

Paper	Summary	How does this impact my project?
A Structural Probe for Finding Syntax in Word Representations	To find parse tree syntax in contextual embeddings using a linear transformation.	We can use these pre trained syntax trees to find the linear transformations for sentences and use these for translation.

FEEDBACK FROM EVALUATORS