

UNDERSTANDING NATURAL LANGUAGES: Samskrita as a case study

Project:

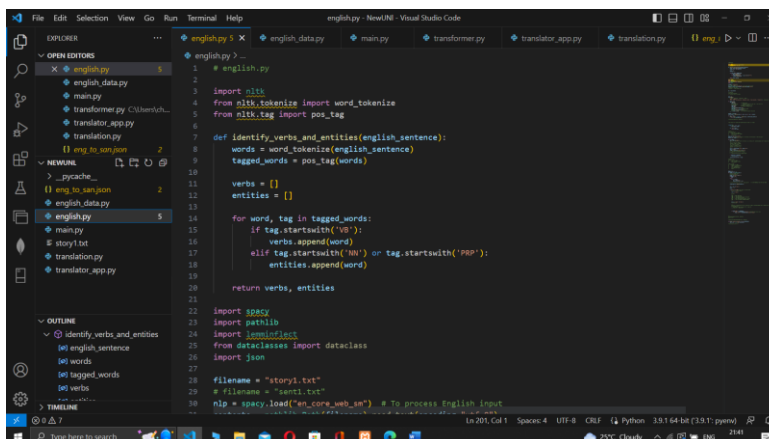
Write code to translate the following paragraph with the Karaka based Interface model

Team Members:

Name	SRN
Chaitanya Madhav R	PES1UG20CS634
SIDDHARTH ARVIND	PES1UG21CS594
SIMRAN MADHUP	PES1UG21CS595

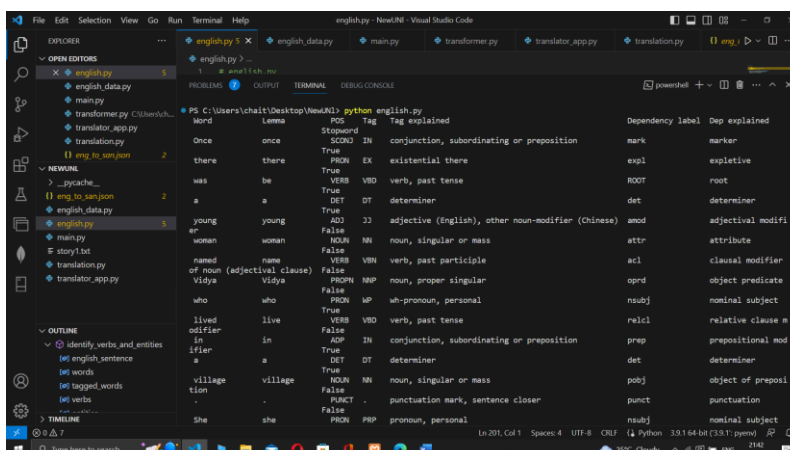
Files submitted:

1. english.py: Contains the implementation of the English Module that reads the English sentence and splits to understand the verb and the entities associated with it.



```
1 # english.py
2
3 import nltk
4 from nltk.tokenize import word_tokenize
5 from nltk.tag import pos_tag
6
7 def identify_verbs_and_entities(english_sentence):
8     words = word_tokenize(english_sentence)
9     tagged_words = pos_tag(words)
10
11     verbs = []
12     entities = []
13
14     for word, tag in tagged_words:
15         if tag.startswith('VB'):
16             verbs.append(word)
17         elif tag.startswith('NN') or tag.startswith('PRP'):
18             entities.append(word)
19
20     return verbs, entities
21
22 import spacy
23 import pathlib
24 import json
25 from dataclasses import dataclass
26 import json
27
28 filename = "story1.txt"
29 # filename = "sent1.txt"
30 nlp = spacy.load("en_core_web_sm") # To process English input
```

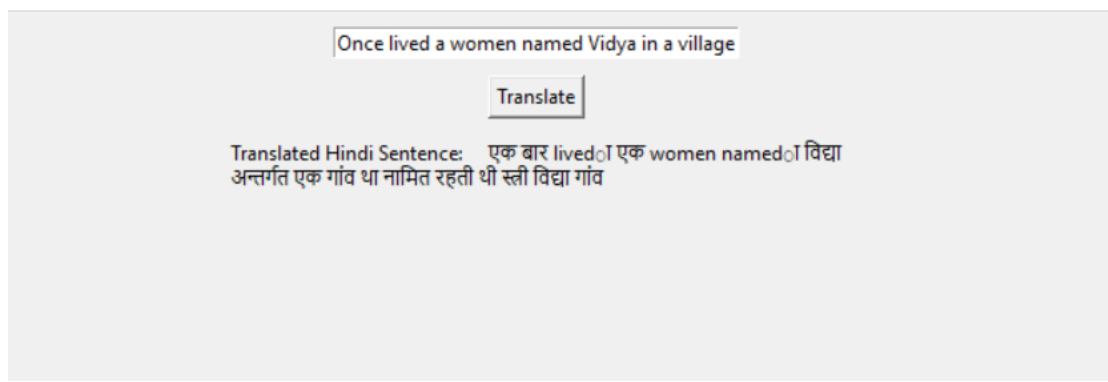
Output of english.py:



word	lemma	POS	Tag	Tag explained	Dependency label	Dep explained
Once	once	ADP	IN	conjunction, subordinating or preposition	mark	marker
there	there	PRON	EX	existential there	expl	expletive
was	be	VERB	VBD	verb, past tense	ROOT	root
a	a	DET	DT	determiner	det	determiner
young	young	ADJ	JJ	adjective (English), other noun-modifier (Chinese)	amod	adjectival modifi
woman	woman	NOUN	NN	noun, singular or mass	attr	attribute
named	name	VERB	VBN	verb, past participle	acl	clausal modifier
of noun (adjectival clause)					opnd	object predicate
Vidya	Vidya	PRON	NP	noun, proper singular	nsubj	nominal subject
who	who	PRON	WP	wh-pronoun, personal	relcl	relative clause m
lived	live	VERB	VBD	verb, past tense	prep	prepositional mod
in	in	ADP	IN	conjunction, subordinating or preposition	det	determiner
if	if	DET	DT	determiner		
a	a	DET	DT	determiner		
village	village	NOUN	NN	noun, singular or mass	poobj	object of preposi
tion		PUNCT	.	punctuation mark, sentence closer	punct	punctuation
She	she	PRON	PRP	pronoun, personal	nsubj	nominal subject

2. English_data.py : class to define the mapping of the entities to the karakas.
3. Main.py: Contains the implementation of the data structure (dictionary) and the rules described for identification of the karakas, vocabulary and grammar rules.
4. Translation.py : Contains the implementation of (Hindi Module) converting the English sentence to Hindi sentence via mapping the karaka framework through the dictionary.
5. Translator_app.py : Contains the interface of Python Tkinter Module to display the result obtained of the translation.

Final Output:



***Thus Our Model is moderately able to translate the given English Sentence into Hindi.**

Note:

1. "Codeboard platform" is unable to support some of the libraries that are used in the project, hence we have displayed the necessary screenshots for reference.