Natural Language Processing – Spring 2021

Momina Atif Dar
P18-0030

Assignment 01

1. Report on Psychological Linguistics in NLP

Psycholinguistics is the study which involves relationship between language and psychological aspects.

Different languages are being spoken in this world. There's so much diversity and it's somewhat possible to work for the betterment of wider audience even if they speak a different language than us. This is what NLP is for. No matter if the text is in English, Urdu, Pashto, Latin, Hebrew or whatever language. Everyone has the freedom to express themselves on social media. When people express themselves, they are reflecting their minds, their thoguht process, their mindset, and most importantly, their mental health. All of the emotions and thoughts are psychological aspects. For example, users pen down their thoughts on social media in different languages. Psycholinguistic would be to interpret what they have written to get a hint of their mental state. This can be used to put up useful ads about mental health or forums where people can connect with people to help them out of misery. For this purpose, Natural Language Processing is essential, as without processing the data found online, without making meaning out of it, one cannot help the people with psychological issues or disorders.

Apart from disorders, thoughts can be monitored to get to study about human brain. For example, in a study conducted in 2011, 509 million Tweets were taken into consideration to get to know more about 'moods'. The study revealed that people are more positive in the morning, and this positivity tends to decrease throughout the day. This study with such large dataset couldn't have been possible without NLP and models. Tweets or such other way of expressions can be considered to look for some unhealthy symptoms in humans i.e. mention of some sort of medicine, some disorder or any depressive symptoms. But again, they have to be distinguished from normal stuff for which NLP comes into play.

Another usage of NLP in psycholinguistics is of suicidal notes. They can help to understand what was the mental state of the person writing the note. Similarities between suicidal notes and hidden-meaning Tweets can be found to work more in this field to raise awareness among people about mental health and how to help such people or directing them to online forums who provide help via advertisements.

Research is being done on such topics on a daily basis. Some even implement their ideas of online help. There are applications for people who have mental health issues which have automation. User communicates with the agent using the specified language i.e. mostly English, and the agent replies accordingly based on its training using. The data is then processed using NLP and the appropriate reply is generated. Agents connect the user with psychologist which takes the conversation further and the solution is sought. But if the agent is not trained properly or NLP is not used properly then the application will be of no use, the user's message won't be interpreted correctly so no appropriate reply will be generated.

Apart from mental health issues, NLP can be implemented in businesses as well. For example, a company sells products, after every purchase from their application or website, they ask user for feedback. Due to diversity, people will give feedback in different languages. The company can collect the feedback dataset on a monthly or even yearly basis and analyze the data using NLP. From the feedback, one can tell what the customers actually look for in a product, what's their mindset and what are their priorities, on what factors they are humble to compromise and more of such sort. This will help company to focus on products which are popular and give the customers what they look for.

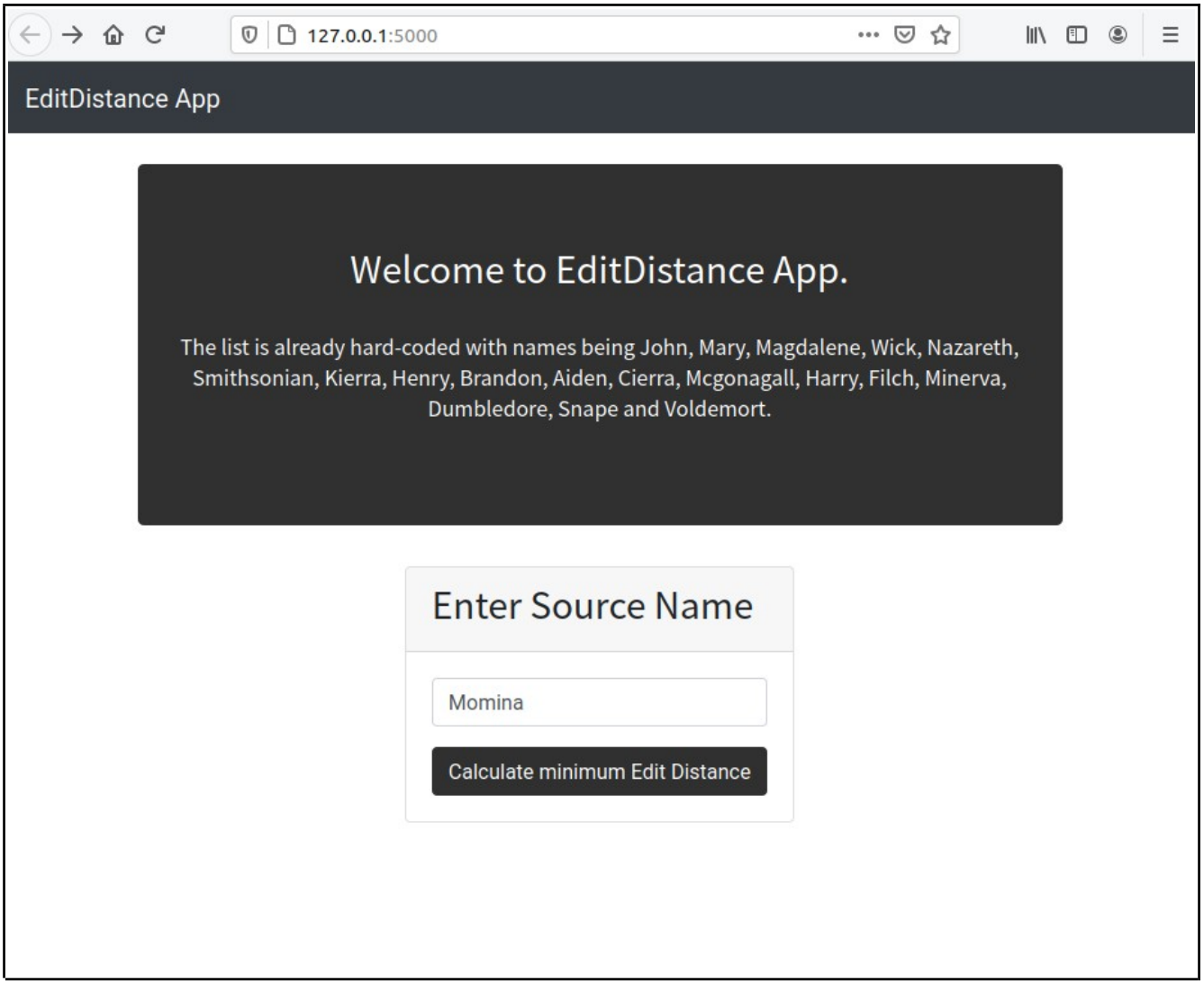
When it comes to text, it can tell a lot about the person that has written it. Text can be processed not just for mental health issues or disorders of any kind but to get a hold of the personality as well, what type of personality traits one possess. Or it can be observed on gender level like data can be collected from female and male both. The responses of female candidates can be compared with responses of male candidates to get to know more about human mind and how it differs according to gender. How women think, what is their thought process and how men think.

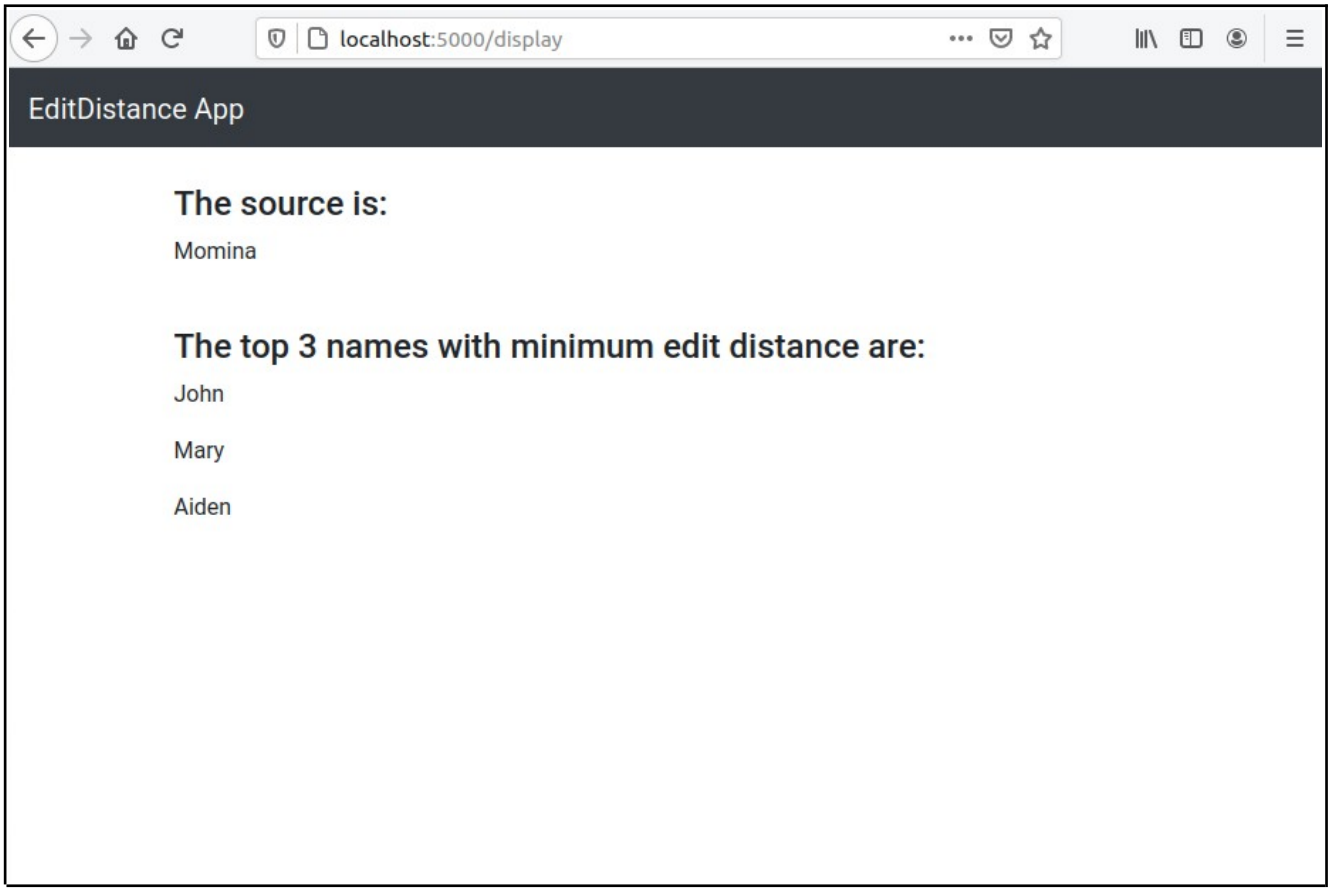
Natural Language plays an important part in processing and understanding humans, how they think, why they do something what might be their mental condition, what sort of personal issues they might be having. The world is still growing, problems are still increasing, we just need to be more empathetic and help people in any way that deems possible even if it requires using NLP to process the thoughts people pen down.

2. Calculate minimum Edit Distance

NOTE: I have used Flask framework to create a basic Web App.

Working:




localhost:5000/display
EditDistance App
The source is:
Momina
The top 3 names with minimum edit distance are:
John
Mary
Aiden

Code in app.py file:

```
lst = ['John','Mary','Magdalene','Wick','Nazareth','Smithsonian','Kierra','Henry',
'Brandon','Aiden','Cierra','Mcgonagall','Harry','Filch','Minerva','Dumbledore','Snape','Voldemort']

#######################################################


class Dummy:
    dic = {}

d = Dummy()

def min_editDistance(lst, dic, src):

    for word in lst:

        add = 0
        s_len = len(src)
        w_len = len(word)

        for i in range(min(s_len, w_len)):

            if src[i] == word[i]:
                continue  #same/equal so pass
            else:
                add = add + 2  #substitution

        if len(src) != len(word):

            if s_len < w_len:

                length = w_len - s_len  #insertion if target is longer

                add = add + length
                dic[word] = add

            if s_len > w_len:
                length = s_len - w_len  #deletion if source is longer

                add = add + length
                dic[word] = add

        dic[word] = add


        #first sorting dictionary and then treating is as list and returning first 3 elements

    return list(dict(sorted(dic.items(), key=lambda item: item[1])))[:3]
```

How to run project: (I am assuming flask and other needed resources are already installed)

Open the ***editDistanceApp*** folder containing ***app.py*** file in terminal and run the command ***python app.py***.
The project will start running and a localhost address will be shown on terminal, open it in browser and you'll be able to see and use project.
