

```
1 from nltk.tag import tnt
2 from google.colab import drive
3 import json
```

```
1 drive.mount('/content/drive')
```

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force\_remount=True)

```
1 tagged_sentences = []
2 with open('/content/drive/MyDrive/Cleaned_Sentences_Task/cleaned_sentences.txt', 'r') as file:
3     tagged_sentences = json.load(file)
```

```
1 actual_tags = []
2 with open ('/content/drive/MyDrive/Cleaned_Sentences_Task/tags_original.txt', 'r') as file:
3     actual_tags = json.load(file)
```

```
1 removed = []
2 with open ('/content/drive/MyDrive/Cleaned_Sentences_Task/sentences_without_tags.txt', 'r') as file:
3     removed = json.load(file)
```

```
1 tagged_sents = []
2 for i in tagged_sentences:
3     tagged_sents.append([tuple(i[0]), tuple(i[1])])
```

```
1 print(tagged_sents[0])
```

```
[('आ', 'nn'), ('की', 'psp')]
```

```
1 pos_tagger = tnt.TnT()
2 pos_tagger.train(tagged_sents[0: -4000])
3 print("POS TAGGER ACCURACY = {}".format(pos_tagger.evaluate(tagged_sents[-4000: ])))
```

POS TAGGER ACCURACY = 0.929125

```
1 nltk_tags = []
2 for i in removed:
3     tag_1 = pos_tagger.tag(i)[0][1]
4     tag_2 = pos_tagger.tag(i)[1][1]
5     nltk_tags.extend([tag_1, tag_2])
```

```
1 print(nltk_tags[0: 2])
```

```
☐➞ ['nn', 'psp']
```

```
1 with open ('/content/drive/MyDrive/Cleaned_Sentences_Task/nltk_predictions.txt', 'w+') as file:
2     json.dump(nltk_tags, file)
```