1. Data set consists of 700 documents of bomb blast. 101 events dataset is there but not tagged with Person_Accused and so on
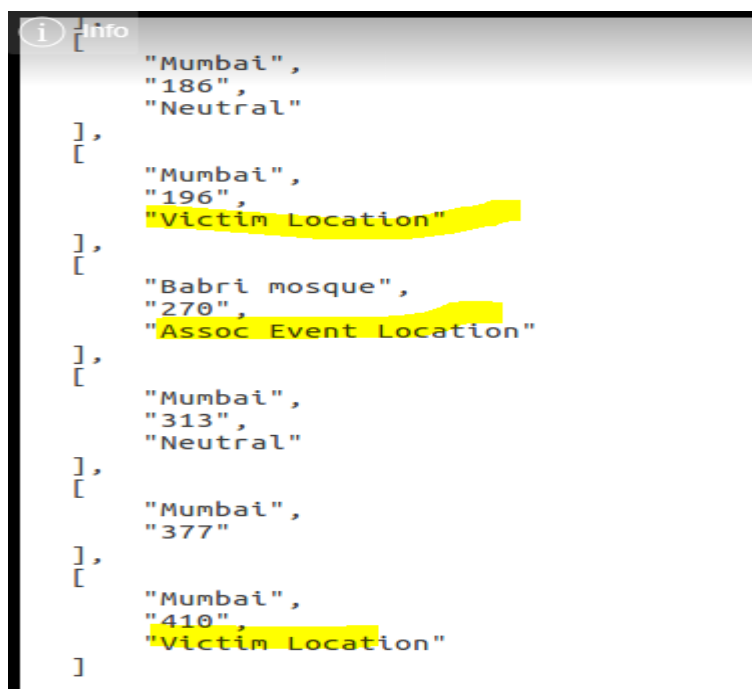
2. Data set has so many issues,

- missing class tags,

- multiple class tags,

- Wrong class assignment tagging

- An entity is assigned two classes. Should I look subjectively the roles. If same entity have different role in the same document, will it be issue in recognizing role at doc level.

- Format of tagging is also different .. causes issue. Need to go through each doc again. I fear due to bad dataset results will be bad.

```
(i) info
    [
        "Mumbai",
        "186",
        "Neutral"
    ],
    [
        "Mumbai",
        "196",
        "Victim Location"
    ],
    [
        "Babri mosque",
        "270",
        "Assoc Event Location"
    ],
    [
        "Mumbai",
        "313",
        "Neutral"
    ],
    [
        "Mumbai",
        "377"
    ],
    [
        "Mumbai",
        "410",
        "Victim Location"
    ]
```

```
            "Neutral"
    ],
    [
        "India-Nepal",
        "45",
        "Accused Location"
    ],
    [
        "Bihars Raxaul town",
        "48",
        "Accused Location"
    ],
    [
        "Punes German Bakery blast",
        "77",
        "Event"
    ],
    [
        "Delhi",
        "84",
        "Event"
    ],
    [
        "Hyderabad",
        "86",
        "Accused Location"
    ],
    [
        "Bangalore",
        "88",
        "Accused Location"
    ],
    [
        "Mumbai",
        "90",
        "Accused Location"
    ],
    [
```

```
    ],
    [
        "Assam",
        "233",
        "Event"
    ],
    [
        "Dhemaji : 349",
        "Event",
        "Event"
    ],
    [
        "Dhemaji",
```

Preprocessing Questions:

1. Should I test the effect of stop word removal, punctuation mark removal.

2. Lemmatization is needed or not. For SRL may be needed. Should I test the effect of these preprocessing also.

3. There are two ways of TAG generation

New Delhi was attacked.

LOC LOC O O // Taking New and Delhi seperate

LOC O O // Taking New and Delhi same

Evaluation Question:

If Others match with Others .. Should we consider that also in evaluation or only classes only. I assume OTHERS should not be considered.

MTP Thesis writing

How??

MTP Phase 1 and Phase 2 Scope??