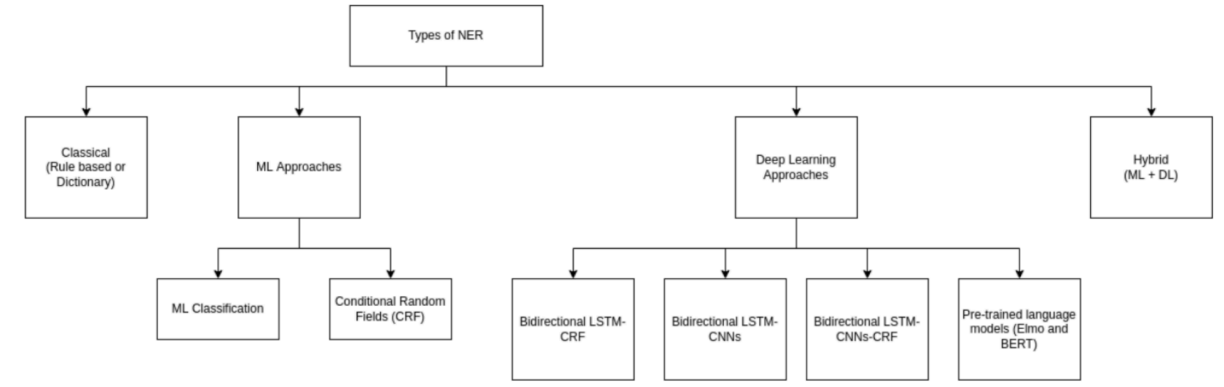
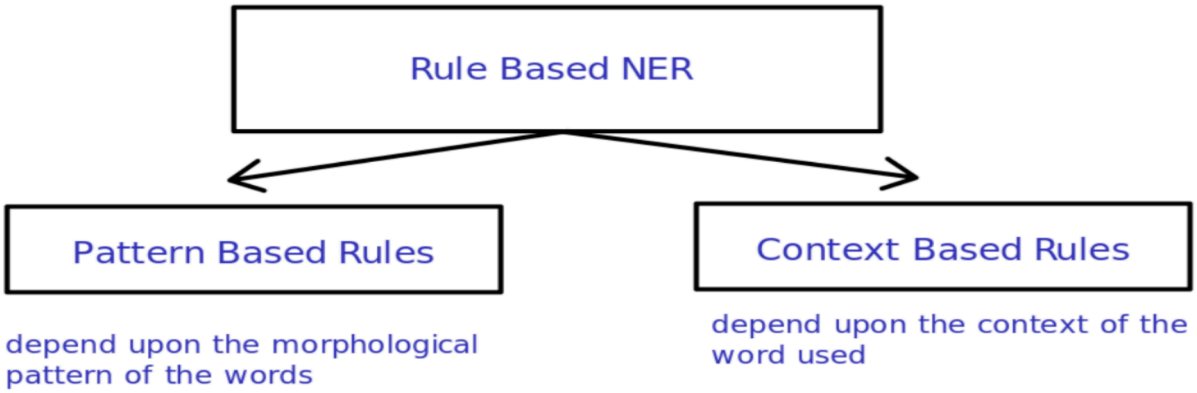
# 8: Named Entity Recognition

**Named entity recognition (NER)** is a subtask of information retrieval concerned with the automatic extraction of named mentions of entities, where the set of possible entity types originally consisted of people, organizations, and locations

| Type | Tag | Sample Categories | Example sentences |
|---|---|---|---|
| People | PER | people, characters | **Turing** is a giant of computer science. |
| Organization | ORG | companies, sports teams | The **IPCC** warned about the cyclone. |
| Location | LOC | regions, mountains, seas | The **Mt. Sanitas** loop is in **Sunshine Canyon**. |
| Geo-Political Entity | GPE | countries, states, provinces | **Palo Alto** is raising the fees for parking. |
| Facility | FAC | bridges, buildings, airports | Consider the **Golden Gate Bridge**. |
| Vehicles | VEH | planes, trains, automobiles | It was a classic **Ford Falcon**. |

## Ways to do NER



## Sequence Labelling

Assignment of labels/tags to each element of a sequence being passed as an input using an algorithm or machine learning model.

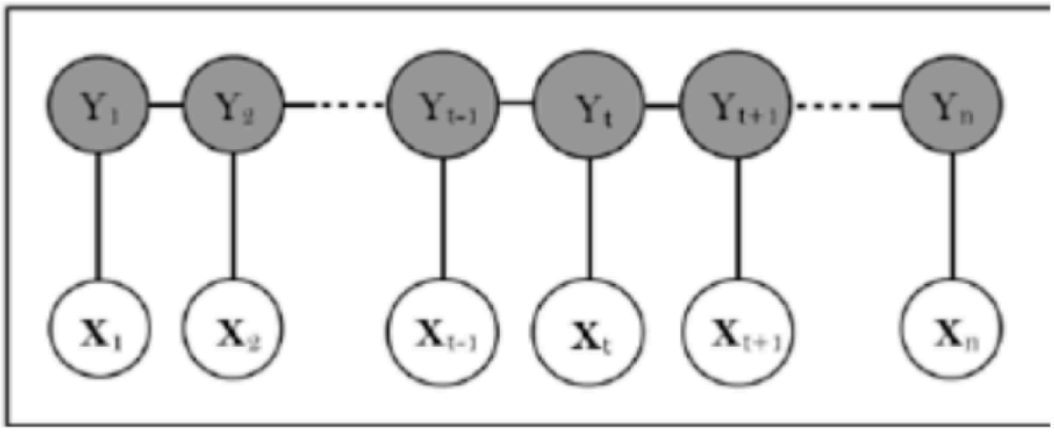
Used for handling: - Segmentation(Detection) ambiguity - Tag assignment(Recognition) ambiguity

## Conditional Random Fields

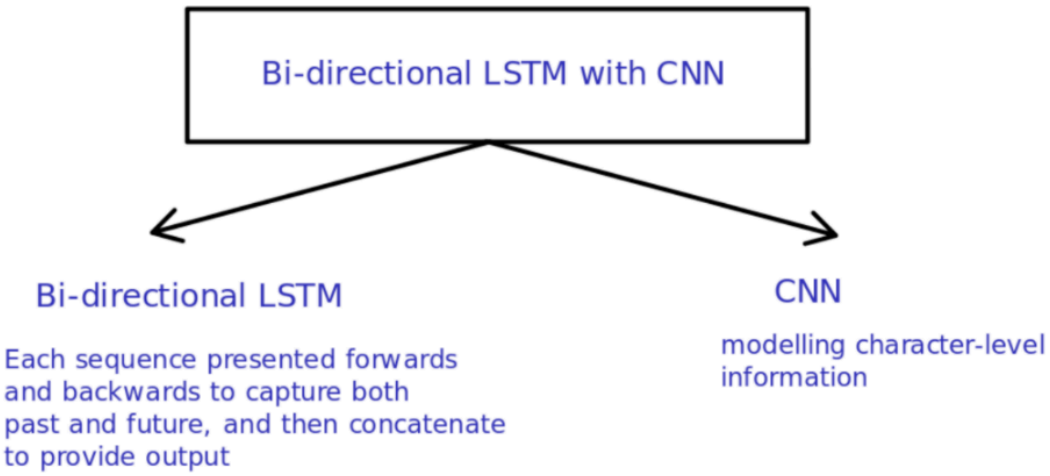
- They are a type of probabilistic graph model that take neighboring sample context into account for tasks like classification.
- Prediction is modeled as a graphical model, which implements dependencies between the predictions.

## Linear CRF

A linear chain CRF confers to a labeler in which tag assignment(for present word, denoted as $y_i$) depends only on the tag of just one previous word(denoted by $y_{i-1}$)



## Deep Learning Approaches

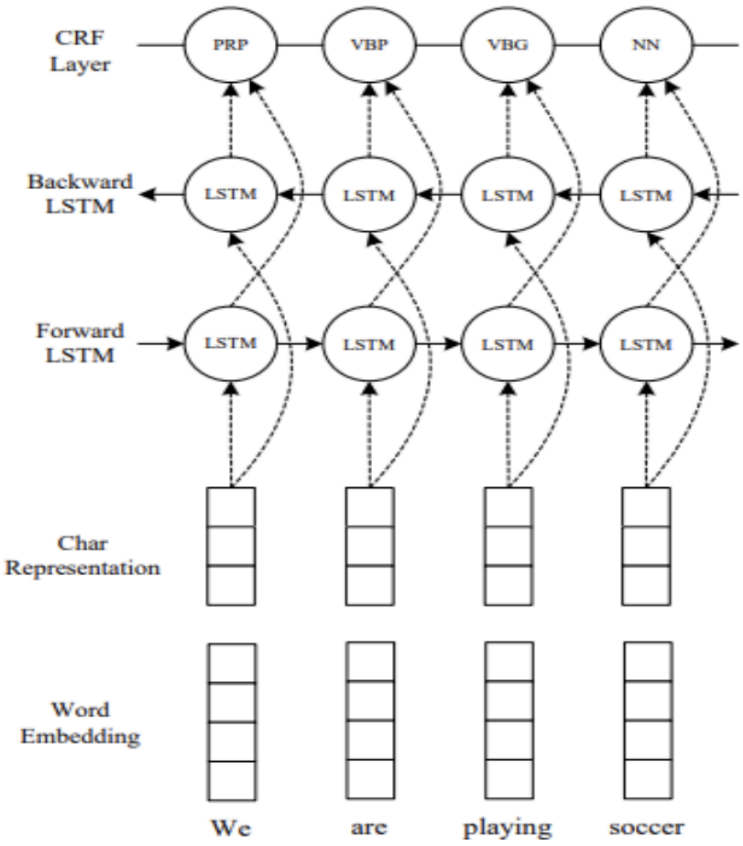



## Bi-directional LSTM with CRF

CRF helps in Sequence Labeling and considering correlations between labels in neighborhoods.

We have how a Bi-directional LSTM model works with another layer of CNN, but for sequence labelling we can consider **CRF**, as

- it is beneficial to consider the **correlations between labels in neighbourhoods** and jointly decode the best chain of labels for a given input sentence.
    - For example, in POS tagging an adjective is more likely to be followed by a noun than a verb,
    - In NER with standard BIO2 annotation I-ORG cannot follow I-PER.
    - Therefore, we model label sequence jointly using a conditional random field (CRF) instead of decoding each label independently.


CRF Layer
PRP
VBP
VBG
NN
Backward LSTM
LSTM
LSTM
LSTM
LSTM
Forward LSTM
LSTM
LSTM
LSTM
LSTM
Char Representation
Word Embedding
We
are
playing
soccer