

**NAME: MADHUMITHA.M**

**CLASS: III-CSE-B**

**SUBJECT: NEURAL NETWORKS**

**TOPIC: NAME ENTITY RECOGNITION**

# NAME ENTITY RECOGNITION

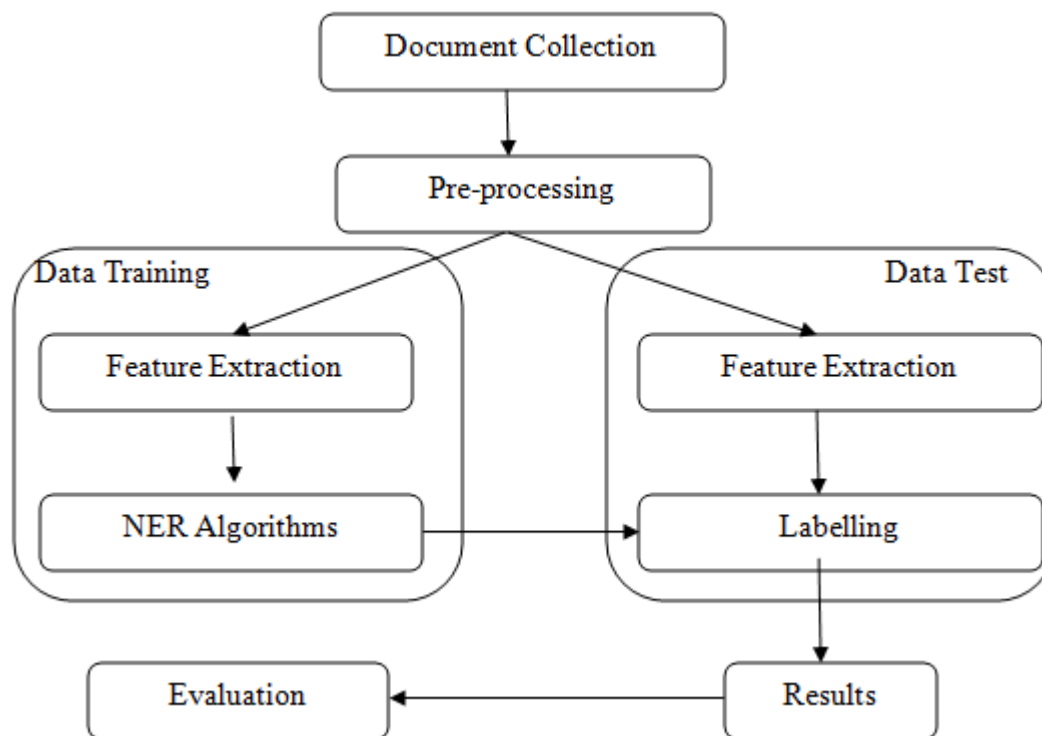
Named Entity Recognition is a popular task in Natural Language Processing (NLP) where an algorithm is used to identify labels at a word level, in a sentence.

NER is an information extraction technique to identify and classify named entities in text. These entities can be pre-defined and generic like location names, organizations, time and etc, or they can be very specific like the example with the resume.

## Approaches to NER:

- **Classical Approaches:** mostly rule-based. here is the link to a short amazing [video](#) by Sentdex that uses NLTK package in python for NER.
- **Machine Learning Approaches:** there are two main methods in this category: **A-** treat the problem as a multi-class classification where named entities are our labels so we can apply different classification algorithms. The problem here is that **identifying and labelling named entities require thorough understanding of the context of a sentence and sequence of the word labels in it**, which this method ignores that. **B-** Another method in this category is **Conditional Random Field (CRF) model. It is a probabilistic graphical model that can be used to model sequential data such as labels of words in a sentence**

- **Deep Learning Approaches:** It is common to use accuracy while training a neural network in different iterations (epochs) as an evaluation metric. However, in case of NER, we might be dealing with important financial, medical, or legal documents and precise identification of named entities in those documents determines the success of the model. In other words, **false positives and false negatives have a business cost in a NER task.**



Other common tasks include classifying of the following:

- date/time
- expression

- Numeral measurement (money, percent, weight, etc)
- E-mail address

## **Ambiguity in NE:**

- For a person, the category definition is intuitively quite clear, but for computers, there is some ambiguity in classification.

## **Methods of NER:**

- One way is to train the model for multi-class classification using different machine learning algorithms, but it requires a lot of labelling. In addition to labelling the model also requires a deep understanding of context to deal with the ambiguity of the sentences. This makes it a challenging task for simple machine learning .
- Another way is that Conditional random field that is implemented by both NLP Speech Tagger and NLTK. It is a probabilistic model that can be used to model sequential data such as words.

## **Named entity recognition used for:**

Named entity recognition (NER) **helps you easily identify the key elements in a text, like names of people, places, brands, monetary values, and more.** Extracting the main entities in a text helps sort unstructured data and detect important information, which is crucial if you have to deal with large datasets.

## **Applications of NER:**

- **Classifying content for news providers:**

A large amount of online content is generated by the news and publishing houses on a daily basis and managing them correctly can be a challenging task for the human workers.

Named Entity Recognition can automatically scan entire articles and help in identifying and retrieving major people, organizations, and places discussed in them. Thus articles are automatically categorized in defined hierarchies and the content is also much easily discovered.

- **Automatically Summarizing Resumes:**

The majority of such tools use the NER software which helps it to retrieve such information. Also one of the challenging tasks faced by the HR Departments across companies is to evaluate a gigantic pile of resumes to shortlist candidates.

A lot of these resumes are excessively populated in detail, of which, most of the information is irrelevant to the evaluator. Using the NER model, the relevant information to the evaluator can be easily retrieved from them thereby simplifying the effort required in shortlisting candidates among a pile of resumes.

- **Powering Recommendation systems:**

NER can be used in developing algorithms for recommender systems that make suggestions based on our search history or on our present activity.

This is achieved by extracting the entities associated with the content in our history or previous activity and comparing them with the label assigned to other unseen content. Thus we frequently see the content of our interest.