

10/17/2025

# **Named Entity Recognition (NER)**

**Ayush Singh**

# Index

1. Introduction.....	2
2. What is Named Entity Recognition (NER)? .....	2
3. Common Entity Types .....	2
4. Process of NER.....	3
5. Practical Applications of NER .....	4
6. How To Improve NER Results.....	4
7. Advantages of NER .....	4
8. Disadvantages of NER.....	5
9. Tools and Libraries for NER .....	5
10. Conclusion .....	5

## 1. Introduction

In today's digital era, an enormous amount of textual data is generated every second through social media, news articles, research papers, and other online platforms. To extract meaningful insights from this unstructured text, Natural Language Processing (NLP) techniques are essential.

One of the most significant tasks within NLP is Named Entity Recognition (NER). NER focuses on identifying and classifying named entities

## 2. What is Named Entity Recognition (NER)?

NER is an advanced technique in Natural Language Processing (NLP) designed to automatically find and classify words or phrases that represent important entities. These entities include individuals, institutions, places, and dates, making complex data more accessible and useful.

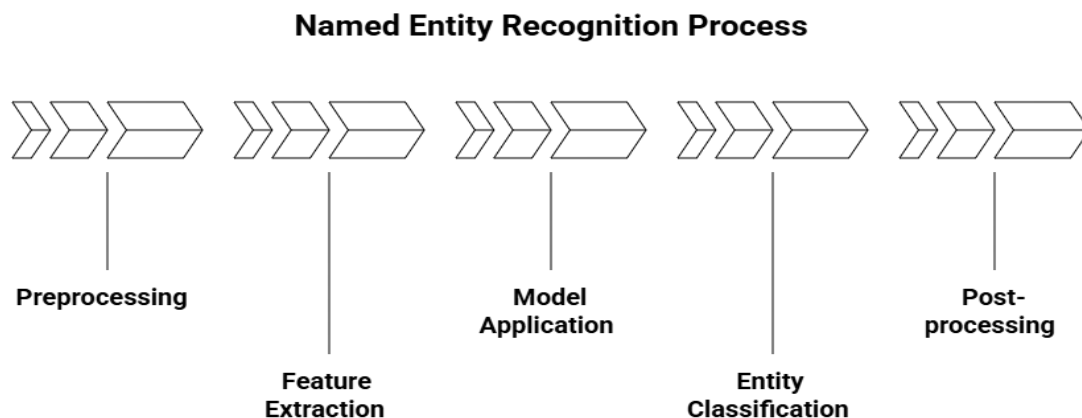
**Example:** "Amazon will ship your order to Delhi on July 12th."

- "Amazon" → Organization
- "Delhi" → Location
- "July 12th" → Date

## 3. Common Entity Types

- **Person:** Names like Ayush or Sachin.
- **Organization:** Entities such as NASA or RBI.
- **Location:** Cities, countries, landmarks (example – India, USA).
- **Date and Time:** Identifies precise dates, months, or centuries.
- **Quantity and Money:** Detects numeric values and currencies.

## 4. Process of NER



### I. Text Input:

- The process begins with raw text data (e.g., sentences, documents, or paragraphs).

### II. Pre-processing:

- Clean and prepare the text for analysis.
- Steps include removing punctuation, converting text to lowercase, and splitting sentences into words (*tokenization*).

### III. Feature Extraction

- Identify useful information about each word, such as:
  - Part of speech (noun, verb, etc.)
  - Capitalization or position in a sentence
  - Surrounding words (context)

### IV. Model Application

- NER systems can be built in two main ways:
  - Rule-Based Systems (Traditional)
    - Use predefined grammar rules and word lists.
  - Machine Learning / Deep Learning Models (Modern)
    - These models learn patterns from labelled data.
    - Common algorithms used:
      - Recurrent Neural Networks (RNN), LSTM)
      - Transformer-based models (e.g., **BERT**)

## **V. Entity Classification:**

- Detected entities are categorized into types like **Person, Organization, Location, Date, Product**, etc

## **VI. Post-Processing**

- Final cleaning and formatting of the results.
- Duplicate or overlapping entities are removed, and output is presented in structured form (e.g., tables or tags).

## **5. Practical Applications of NER**

- **Healthcare:** Finds diseases, drugs, and patient names in medical records, speeding up research and improving care.
- **Finance:** Identifies company names and financial events in reports and news to guide investments and spot risks.
- **Publishing & News:** Tags and organizes articles by featured people, places, and organizations, improving search and recommendations.

## **6. How To Improve NER Results**

- Train with More Data
- User Feedback
- Domain-Specific Models

## **7. Advantages of NER**

- Finds important info quickly.
- Improves search results.
- Helps in automation (like chatbots).
- Works in many fields.
- Makes text more organized.

## **8. Disadvantages of NER**

- Needs lots of training data.
- May not work well in all topics.
- Can confuse similar words.
- Not always accurate.
- Needs regular updates.

## **9. Tools and Libraries for NER**

Several NLP libraries and frameworks provide easy-to-use NER functionalities:

- spaCy
- NLTK (Natural Language Toolkit)
- Stanford NLP
- Hugging Face Transformers

## **10. Conclusion**

Named Entity Recognition (NER) helps computers understand and organize text by finding important names, places, and other details. It makes information easier to use and supports many real-world applications like chatbots, search engines, and data analysis. Although NER has some limits, such as needing training data and updates, it remains a very useful and important tool in today's data-driven world.