**Named Entity Recognition**

**About**

Named Entity Recognition (NER) is a powerful tool that can provide significant benefits across various industries. By automatically identifying and categorizing important information within unstructured text, NER serves as a bridge between raw data and actionable insights.

**Application of NER**

1. Improved Customer Support

NER can help companies better organize and analyse customer feedback and complaints, making it easier to respond promptly and improve support services. By identifying relevant entities like product names, departments, and locations, NER enables faster routing of issues to the appropriate teams.

1. Enhanced Customer Experience

By leveraging NER to extract insights from customer interactions, businesses can gain a deeper understanding of their customers' needs and pain points. This knowledge can inform product development, marketing strategies, and overall customer experience improvements.

1. Competitive Intelligence

NER can help businesses stay on top of industry trends and monitor their competitors by analysing news articles, social media posts, and online discussions. By identifying key entities like competitor names, products, and industry events, NER provides valuable competitive intelligence to inform strategic decision-making.

1. Sentiment Analysis

Combining NER with sentiment analysis techniques allows businesses to gauge brand perception and customer sentiment towards their products or services. By identifying entities and associating them with positive or negative sentiment, businesses can quickly identify areas for improvement or opportunities for growth.

1. Regulatory Compliance

In industries like finance and healthcare, NER can help businesses ensure compliance with regulations by identifying sensitive information like personal identifiers, financial data, and medical terms within large volumes of text. This automation reduces the risk of human error and streamlines compliance processes.

**Relevance of the Project**

Accurate information extraction is crucial for businesses to gain actionable insights from vast amounts of text data. By utilizing advanced NLP techniques, this project aims to streamline the extraction of key entities, such as person names, organizations, and locations, from unstructured text, thus enhancing data-driven decision-making and operational efficiency.

The identified entities can be used for a variety of applications:

Resume Comparison: Enhancing recruitment processes by automatically extracting and comparing key information from candidate resumes

Customer Feedback Analysis: Analyzing customer feedback to identify mentioned products, services, and sentiment, aiding in product improvement and customer satisfaction strategies.

Market Research: Extracting relevant data from market reports and news articles to inform business strategies and competitive analysis.

**Objective**

The objective of this project is to create a Named Entity Recognition (NER) model that can effectively identify and categorize crucial entities such as persons and organizations within unstructured text data. By developing a robust NER model, the project aims to enhance information extraction processes, improve data organization, and facilitate more accurate and efficient analysis of textual data.

**Understanding the dataset**

Data set collected from <https://www.kaggle.com/datasets/abhinavwalia95/entity-annotated-corpus>

**Solution**

Named Entity Recognition (NER) is a task of Natural Language Processing (NLP) that involves identifying and classifying named entities in a text into predefined categories such as person names, organizations, locations, and others. The goal of NER is to extract structured information from unstructured text data and represent it in a machine-readable format. Approaches typically use BIO notation, which differentiates the beginning (B) and the inside (I) of entities. O is used for non-entity tokens.

NER extracts the meaningfull information from the unstructured data, basically detecting the entities and using the similary measurement we can take decisions.

**Solution approach**

Fine tuned the Hugginface bert-based-model to develop a named entity recognition

**Solution Proposed**

* Import necessary libraries and load the dataset.
* Prepare the data for training and tokenization.
* Load a pre-trained BERT model from Hugging Face.
* Define a custom model, train, evaluate, and fine-tune for NER.
* CI/CD deployment

**Deployment:**

Production Grade CI CD deployment using CircleCI