



**Data Science  
Academy**

[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)

Processamento de Linguagem Natural

Extraindo e Compreendendo as Features

As tags NER são realmente importantes porque ajudam você a entender a estrutura das frases e ajudam as máquinas ou os sistemas de PLN a entender o significado de certas palavras em uma frase. Vamos dar um exemplo. Se você está construindo uma ferramenta de revisão, ferramentas NER podem encontrar o nome de uma pessoa, um nome de organização, símbolos relacionados à moeda, formatos numéricos e assim por diante que ajudarão sua ferramenta de revisão a identificar casos excepcionais presentes no texto. Então, de acordo com a tag NER, o sistema pode sugerir as mudanças necessárias. Considere a sentença:

**“O Bank of America anunciou o seu lucro hoje pela manhã.”**

Neste caso, a ferramenta NER fornece a organização de tags para o Bank of America, o que ajuda nosso sistema a entender melhor o significado da sentença e a estrutura da sentença. As tags NER também são muito importantes se você estiver criando um sistema de perguntas e respostas, pois é fundamental extrair entidades nesse sistema. Depois de gerar as entidades, você pode usar um relacionamento sintático para entender as perguntas. Depois dessa etapa, você pode processar a pergunta e gerar a resposta.

## Mas Existem Desafios

Existem alguns desafios para o sistema NER, que são os seguintes:

- As ferramentas NER treinam em um conjunto de dados de domínio fechado. Portanto, um sistema NER desenvolvido para um domínio normalmente não funciona bem em outro domínio. Isso requer uma ferramenta NER universal que possa funcionar para todos os domínios e, após o treinamento, deve ser capaz de generalizar o suficiente para lidar com situações não vistas.
- Às vezes, você encontrará palavras que são nomes de locais, bem como o nome de uma pessoa. A ferramenta NER não pode manipular um caso em que uma palavra pode ser expressa como o nome do local, o nome da pessoa e o nome da organização. Este é um caso muito desafiador para todas as ferramentas NER. Suponha que você tenha a palavra TATA hospital; a palavra TATA pode ser o nome de uma pessoa, bem como o



nome de uma organização. Nesse caso, a ferramenta NER não pode decidir se TATA é o nome de uma pessoa ou o nome de uma organização.

- Construir uma ferramenta NER especificamente para plataformas web de microblog também é uma tarefa desafiadora.

Vejamos agora uma forma eficiente de compreender e extrair as features de um texto.