Named Entity Recognition (NER)

Entities and Named Entities

- Entities: common things, belongs to noun family

The president will meet the chairman of the company in the capital city

- Named Entities: entities having proper name

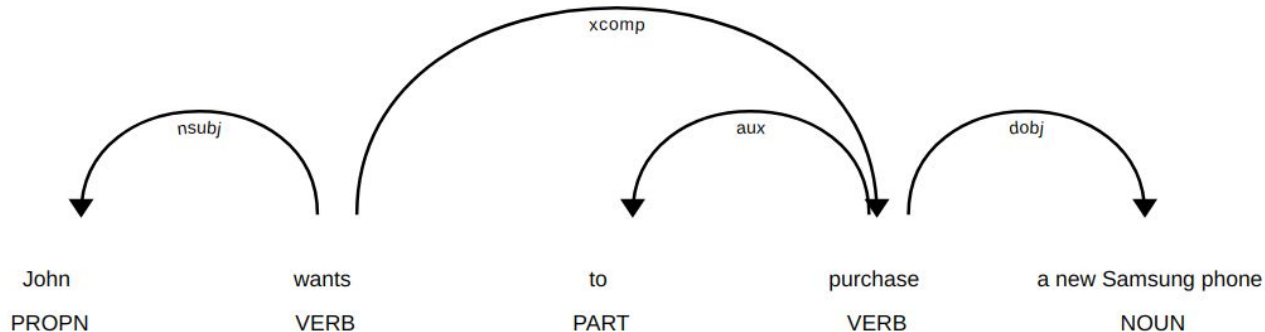People Places Companies Products

Donald Trump will meet the chairman of Google in New York City

Named Entity Recognition (NER)

- Process of finding named entities from a given text

- Two steps:

  - Named Entity Extraction

  - Named Entity Classification

Named Entity Extraction

- Identification of noun phrases

- Noun phrases: connected by direct subject or object relationships

- Sentence: John wants to purchase a new Samsung phone

Named Entity Extraction

NNP	VBZ	TO	VB	DT	JJ	NNP	NN
John	wants	to	purchase	a	new	Samsung	phone

Named Entities: John, Samsung

NNP	NNP	VBZ	NN	IN	NNP
Satya	Nadella	is	CEO	of	Microsoft

Named Entities: Satya Nadella, Microsoft

NNP	VBZ	VBN	NN	IN	NNP	NNP	.
Facebook	has	bought	stake	in	Reliance	Jio	.

Named Entities: Facebook, Reliance Jio

Analytics Vidhya

Named Entity Classification / Linking

- Assigning the class / category of the named entity

- Classes: Person, Organization, Location, Drug etc.

- Entity Linking with a Knowledge Base

- Linking Approaches:

    - Domain Dictionary Lookup (Ontologies)

    - Knowledge Base: Wikipedia, Google

    - APIs - Google Maps, News API
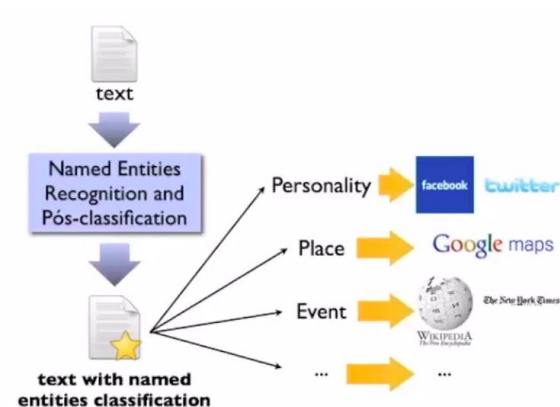
Challenges

- Updated Ontologies

- Name Variation

  - Singh Yuvraj, Yuvraj Kr Singh

- Multiple Categories

  - London: Name of a person, Name of a city

  - Donald Trump Park: Name of a park(location), Name of a person

- Non Capitalization

Thank You