Machine Learning - Tag Acquirer / Target in an M&A Deal

Vaibhav Singh
Sushant Mehta
Sarthak Agarwal
Virat Kumar

Problem Statement

- Three dataset files present are:

Corpus.xlsx - 900 rows * 5 columns

system_id	title	org_link	source	date	raw_article

MnA_Training.xlsx - 500 rows * 3 columns
MnA_Test.xlsx - Test data 250 rows * 3 columns

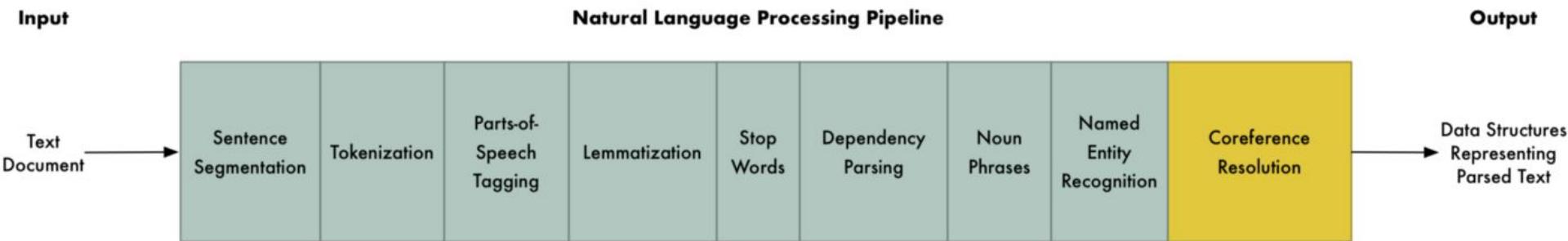| system_id	Acq_Final	Tar_Final

Standard Approach

- Match unique id from corpus set and training set

- Tokenizer splits data into small figments

- Use tagger library which informs about nouns, verbs, adjectives etc.

- Use parser library defines the dependency of words to each other

- Match unique id from corpus set and test set and predict Acq_Final & Tar_Final according to result of above algorithm

What does this mean to you?

- **Tagger**
- **Parser**
- **Tokenizer**

Input Natural Language Processing Pipeline **Output**

Text Document → Sentence Segmentation | Tokenization | Parts-of-Speech Tagging | Lemmatization | Stop Words | Dependency Parsing | Noun Phrases | Named Entity Recognition | Coreference Resolution → Data Structures Representing Parsed Text

Our Approach! Let's Discuss.
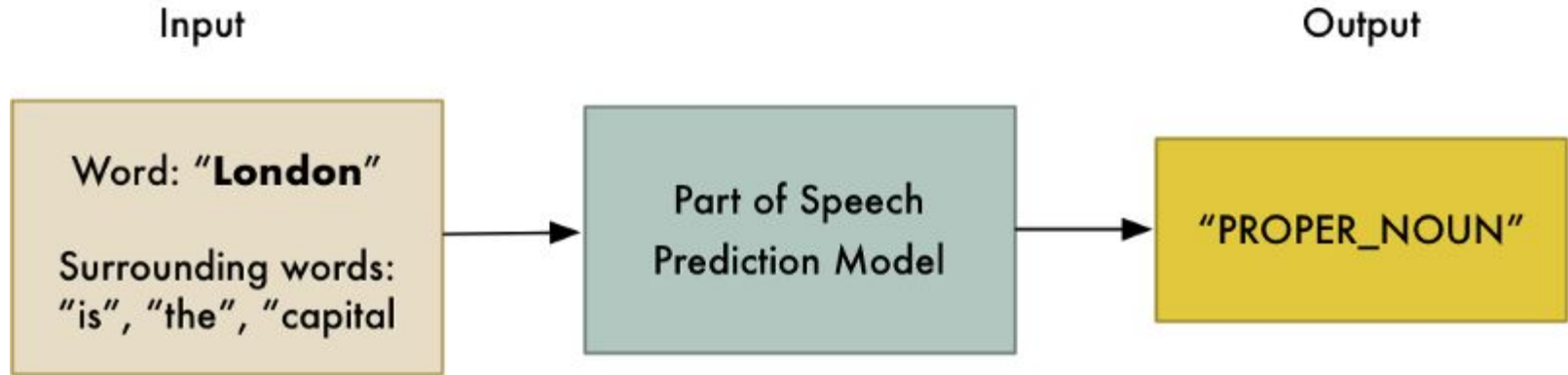
Tokenizer

Stemming

Tagger

Rule Based Approach

Rule Based Approach

Buy: Token after buy is target
Sell: Token after sell is target
Sells: Token after sells is target
Sells to: Token after sells to is acquiring

Thank You