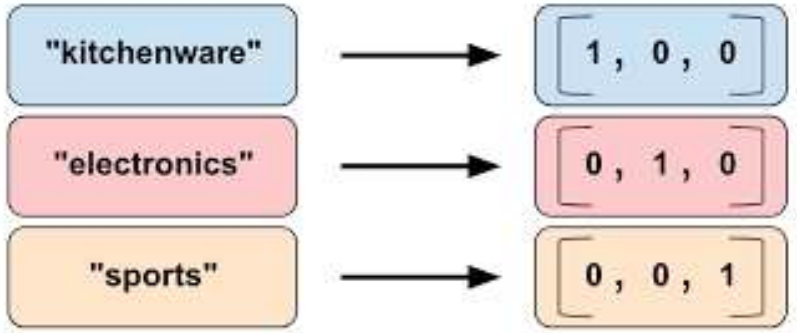
NATURAL LANGUAGE PROCESSING ADVANCEMENTS BY DEEP LEARNING: A SURVEY

INTRODUCTION

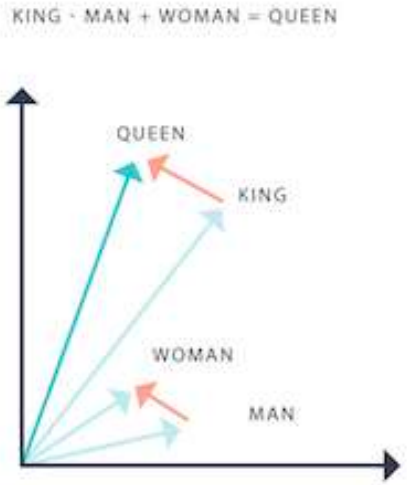
- The process of transferring the knowledge of human-level understanding of human interactions into a machine is the concept of Natural Language Processing.
- Human interactions hold huge information, involving various topics, tones, words, emotions.
- The set of unstructured/unlabelled data are difficult to be processed by machines, requiring extensive efforts, and varied models. It also requires higher computation power, more time, and handcrafting the features through careful analysis.
- The state-of-the-art deep learning models attempt to overcome the above-mentioned complications. It gets an **in-depth representation of the language**, identifies **meaningful information** in a text for further processing, and **constructs features** at different levels.

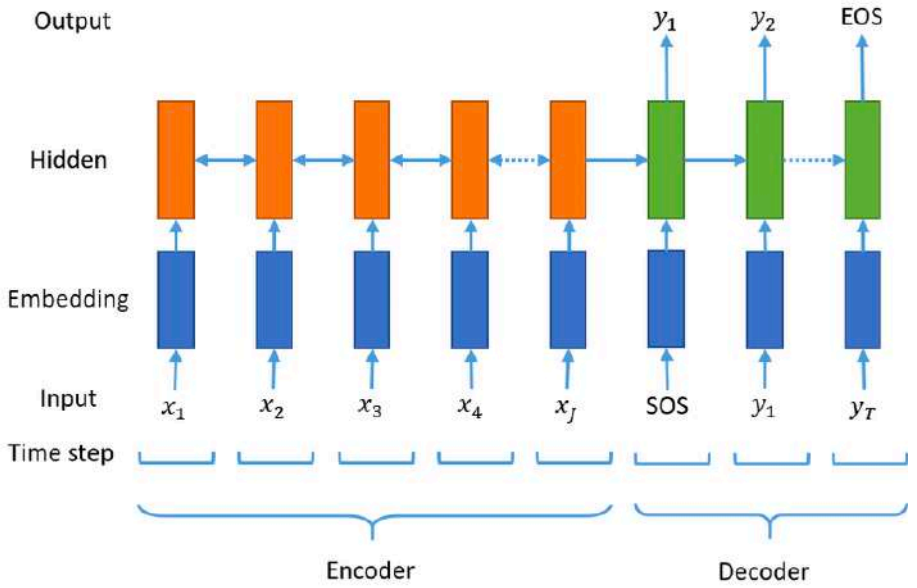
CORE CONCEPTS IN NLP

- **Feature Representation**
  - Converting textual inputs into a machine-understandable format.
  - One-Hot Representation: Representing unique elements in text/categorical variables as binary vectors.
  - Continuous Bag of Words: To predict the missing/target word or word sequence based on the neighbouring words in a context.
  - Word-Level Embedding: Collections of words are converted to vectors of real numbers in representation space.
  - Character-Level Embedding: Each character in a word is one hot encoded.
- **Seq2Seq Framework**
  - The encoder gets the sequence of text inputs, creates a mid-level/hidden output to be ingested to the decoder to produce a sequence of output

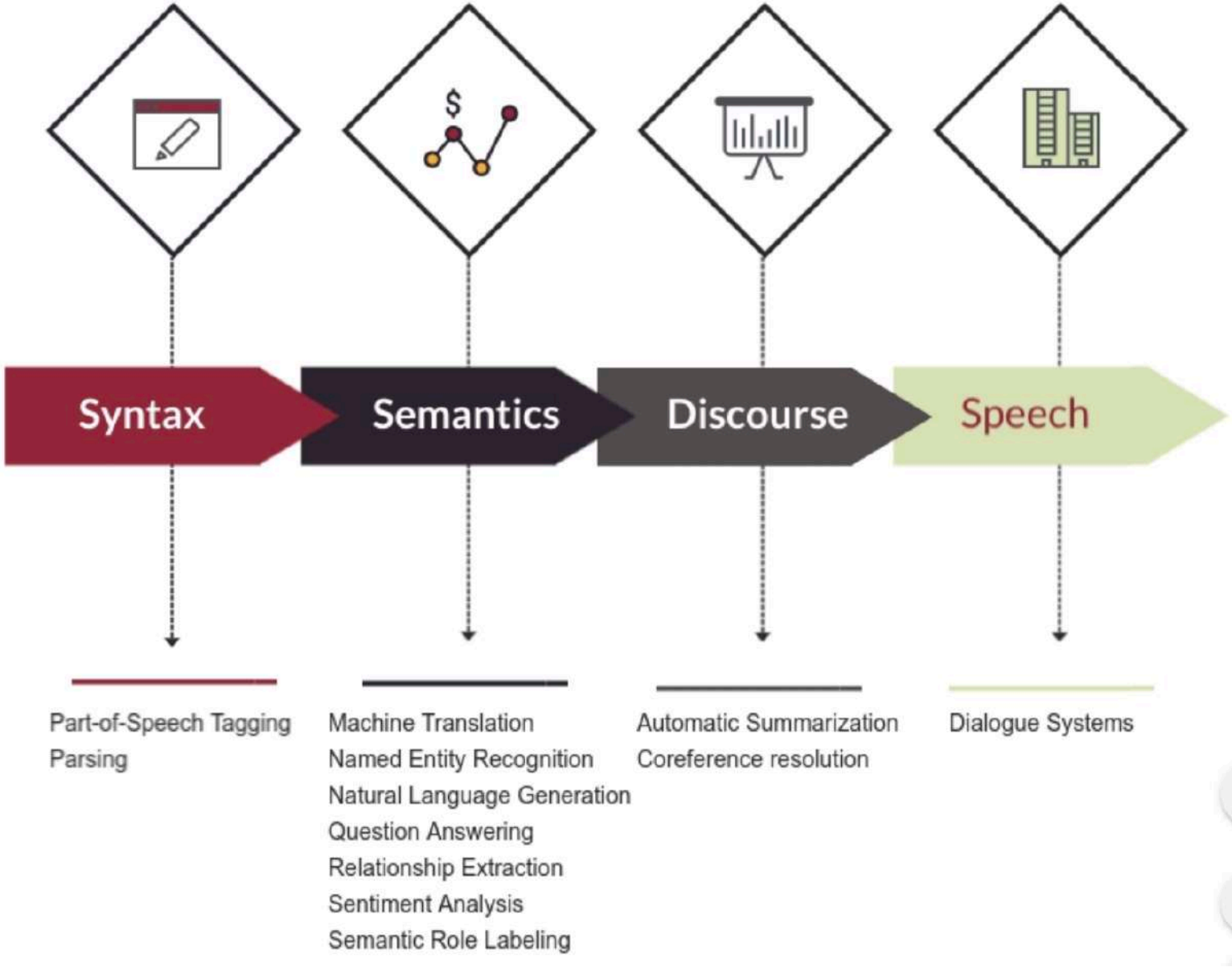

"kitchenware"
"electronics"
"sports"
1 , 0 , 0
0 , 1 , 0
0 , 0 , 1


Context word
Context word
I am
eating
good pizza
now
Target word


KING - MAN + WOMAN = QUEEN
QUEEN
KING
WOMAN
MAN


Output
Hidden
Embedding
Input
Time step
y_1
y_2
EOS
x_1
x_2
x_3
x_4
x_J
SOS
y_1
y_T
Encoder
Decoder

STATE-OF-THE-ART DEEP LEARNING MODELS FOR NLP TASKS

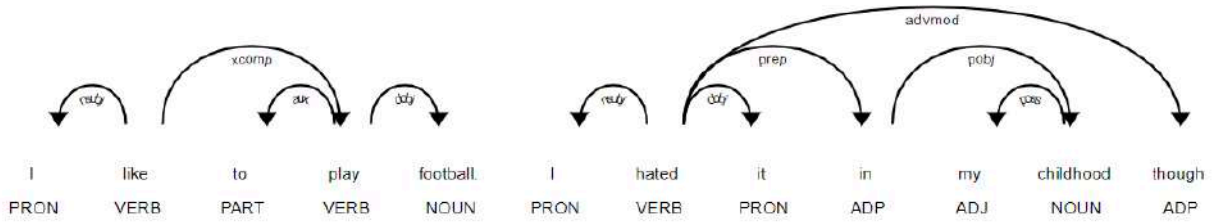

Syntax
Semantics
Discourse
Speech
Part-of-Speech Tagging
Parsing
Machine Translation
Named Entity Recognition
Natural Language Generation
Question Answering
Relationship Extraction
Sentiment Analysis
Semantic Role Labeling
Automatic Summarization
Coreference resolution
Dialogue Systems

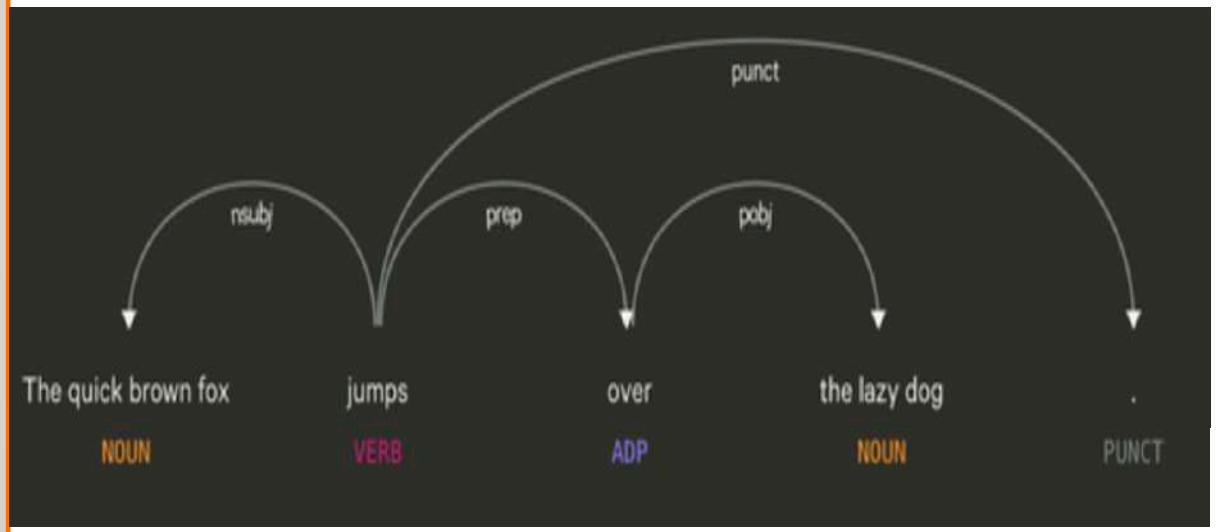
BASIC TASKS IN NLP

- Parts-of-Speech:
  - Labelling word tokens with their part of speech categories — whether POS tag is a noun, verb, adjective, preposition, etc.,
  - Meta-BiLSTM is the current state-of-the-art model for POS tagging with 97.96% accuracy.
- Parsing:
  - Parsing is assigning a structure to a sequence of tokens.
  - Constituency Parsing: Assigns syntactic structure to a sentence. It is also known as deep parsing.
  - Dependency Parsing: Shows structural relationship between tokens in a sentence.
  - Self-Attentive Encoder is the current state-of-the-art model for parsing with 95.1% accuracy.

BASIC TASKS IN NLP

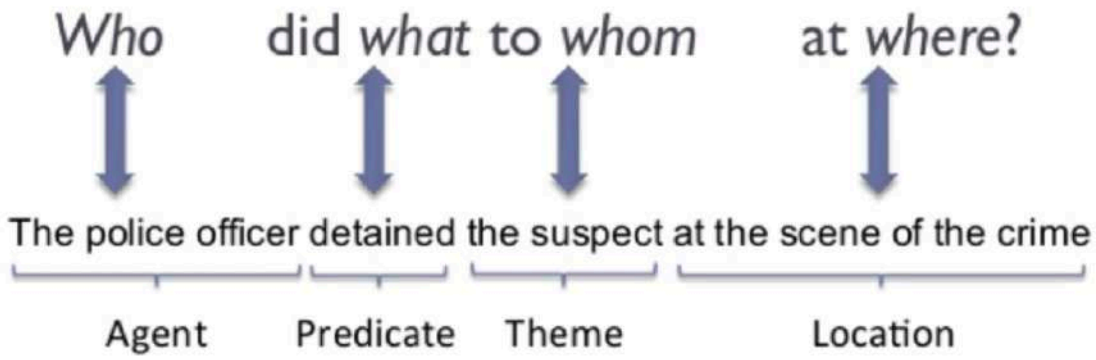
- Semantic Role Labeling :
  - SRL is the process to identify and classify elements in a sentence to determine “who” did “what” to “whom” as well as “how,” “where,” and “when”.
  - Argumented Representation with BiLSTM is the current state-of-the-art model for SRL with 85.3% accuracy.






```
(S
 (NP The/DT quick/JJ brown/NN fox/NN) (VP jumps/VBZ)
 (PP over/IN)
 (NP the/DT lazy/JJ dog/NN)
./.)
```

```
(S
    (NP The/DT quick/JJ brown/JJ fox/NN)
    (VP jumps/VBZ
        (PP over/IN
            (NP the/DT lazy/JJ dog/NN)
        )
    )
    (. .)
)
```



TEXT CLASSIFICATION

- The process of categorizing the text input involving word, sentence or document is text classification.
- A simple CNN model can be used to classify sentences.
- RNN, LSTM-RNN architectures have proved to work well for this particular task.
- Universal Language Model Fine-tuning, also called as ULMFIT, is the current state-of-the-art model for text classification with an accuracy of 94.99%.

INFORMATION EXTRACTION

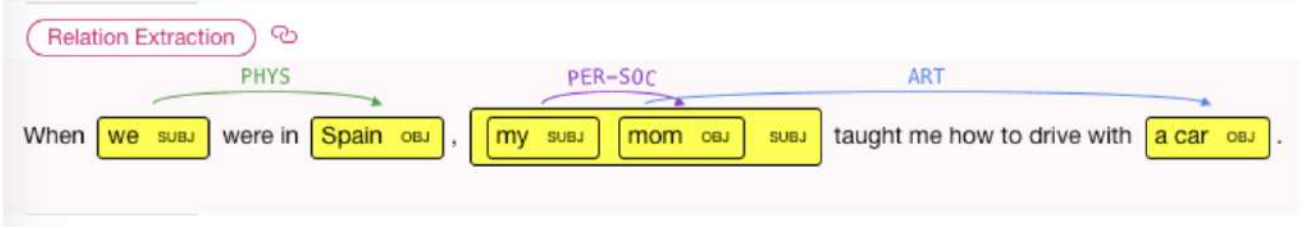
- Distilling the required information, forming structured results as features to a model is called Information Extraction.
- Named Entity Recognition:
  - process of categorizing elements in the text to a predefined category such as Person, Location, Organization, Date, Time.
  - Contextual String Embedding is the current state-of-the-art model with 93.09% accuracy for NER.
- Relation Extraction:
  - getting the semantic relationship between entity pairs.
  - Recurrent Neural Network and Convolutional Neural Network are the proposed models for Relation based classification.

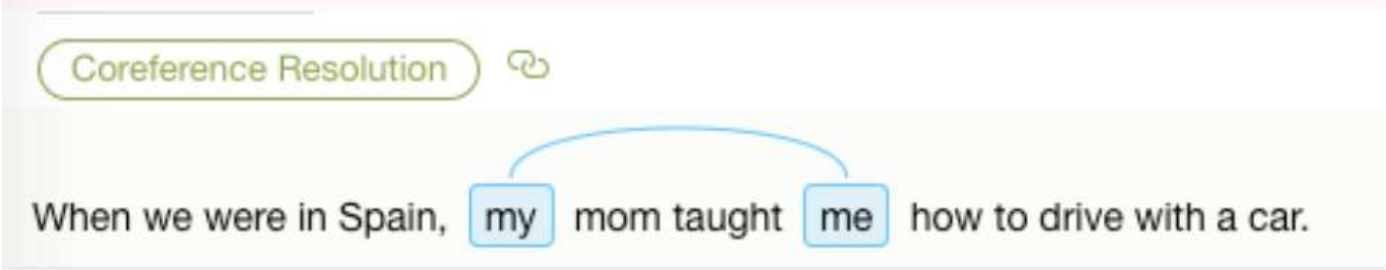
INFORMATION EXTRACTION

- Coreference Resolution:
  - when the same entity is referenced in a sentence/context. For example, the entities "car," "Camry," and "it" could all refer to the same entity.
  - The process of identifying the coreference is called coreference resolution.
  - Reinforcement Learning can be applied. The current state-of-the-art model includes an attention mechanism.
- Event Extraction:
  - A specific type of extracted information from text is an event. This process of recognizing events will also be able to identify trigger words and assign labels to those entities.
  - Convolutional Neural networks are utilized for event extraction/detection.
  - Graph CNN is the state-of-the-art model for event detection with operations that can be applied to syntactically dependent words and also for consecutive words.

In fact, the Chinese NORP market has the three CARDINAL most influential names of the retail and tech space – Alibaba GPE , Baidu ORG , and Tencent PERSON (collectively touted as BAT ORG), and is betting big in the global AI GPE in retail industry space . The three CARDINAL giants which are claimed to have a cut-throat competition with the U.S. GPE (in terms of resources and capital) are positioning themselves to become the 'future AI PERSON platforms'. The trio is also expanding in other Asian NORP countries and investing heavily in the U.S. GPE based AI GPE startups to leverage the power of AI GPE . Backed by such powerful initiatives and presence of these conglomerates, the market in APAC AI is forecast to be the fastest-growing one CARDINAL , with an anticipated CAGR PERSON of 45% PERCENT over 2018 - 2024 DATE .

To further elaborate on the geographical trends, North America LOC has procured more than 50% PERCENT of the global share in 2017 DATE and has been leading the regional landscape of AI GPE in the retail market. The U.S. GPE has a significant credit in the regional trends with over 65% PERCENT of investments (including M&As, private equity, and venture capital) in artificial intelligence technology. Additionally, the region is a huge hub for startups in tandem with the presence of tech titans, such as Google ORG , IBM ORG , and Microsoft ORG .





SENTIMENT ANALYSIS

- Sentiment Analysis is identifying the tone in which the information is presented. It is also called opinion mining
- The main aim is to analyse the human opinion, sentiments, emotions on a product, problem, or any other subject.
- Document-level Sentiment Analysis:
  - Identifying if the document exhibit positive or negative tone on any entity is document-level sentiment analysis.
  - The Gated Recurrent Neural Network can be employed to encode the relations in the semantic structure of a document.
- Sentence-level Sentiment Analysis:
  - Determining if a sentence is expressing positivity, neutrality, or negativity is a Sentence-level Sentiment Analysis.
  - The Sentiment Treebank and Recursive Neural Tensor Networks are the current state-of-the-art model for fine-grained sentiment labelling.

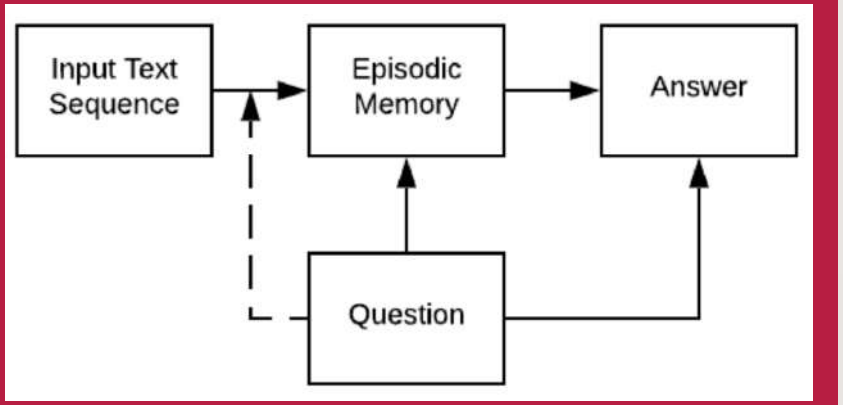
SENTIMENT ANALYSIS

- Aspect-level Sentiment Analysis:
  - Aspect-level works directly on the target of the opinion expressed. Example: A product.
  - Aspect-level Sentiment Analysis is comprised in two stages:
    - Aspect Sentiment Classification — for identifying the tone of the opinion (positive, neutral, or negative)
    - Aspect Extraction — for identifying the target of the opinion
  - Example: For this sentence — "This car is old. It must be repaired and sold!". "car" is the target while the sentiment is negative.
  - Attention-based LSTMs can be used to connect the target and the sentiment.
  - BERT is the state-of-the-art model highly proposed for this specific task.

MACHINE TRANSLATION

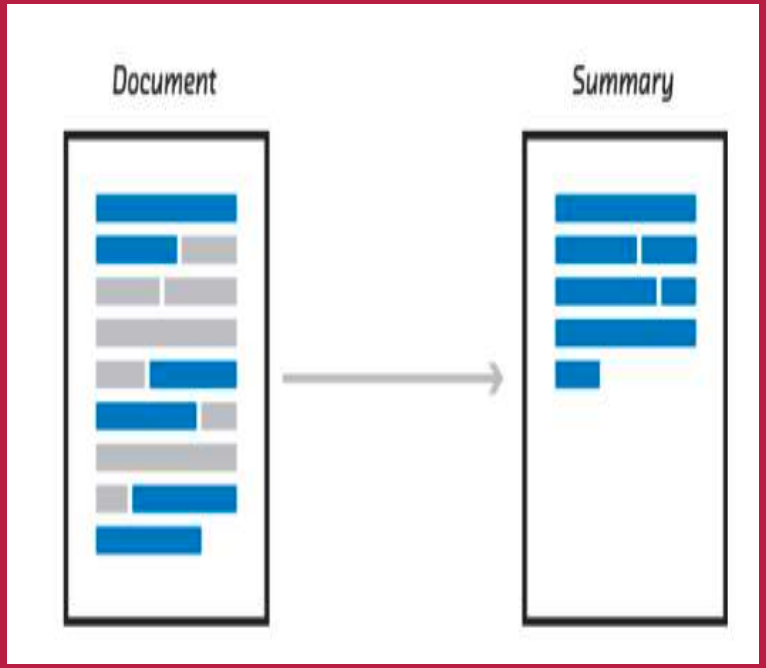
- The process of translating text/speech from one language to other (Speech and Language Processing) by a model is Machine Translation.
- Neural Machine Translation:
  - Traditional Machine Translation was first introduced in 1954, with very limited vocabulary. Human labour was used for evaluation
  - Neural Machine Translation did not require extensive pre-processing and word alignments. The importance was mainly given to network structure.
- Convolutional Seq-to-Seq, Attention Is All You Need, Weighted Transformer, Self Attention, DeepL Translation Machine are some of the models used for machine translation on English-German Dataset.
- Back-translation is the current state-of-the-art model with 35% accuracy for Machine Translation.

QUESTION ANSWERING



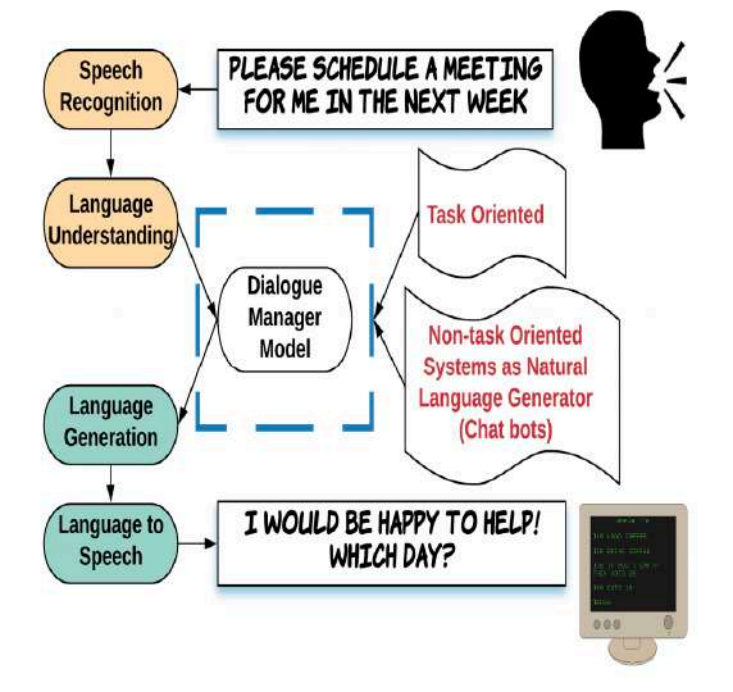
- The detailed form of Information Retrieval is the Question Answering System.
- The expected output from the Question Answering System can be in the form of document, text, image, etc, retrieved from a set of documents.
- Rule-based Question Answering:
  - IBM's QA Systems: 4 components
    - Question/Answer Type Classification
    - Query Expansion/Information Retrieval
    - Name Entity Making,
    - Answer Selection
  - Dynamic Memory Network — DMN can be used for QA system. It also contains 4 components:
    - Input module
    - Question module
    - Episodic memory module
    - Answer module.

DOCUMENT SUMMARIZATION



- Condensing multiple documents into a set of sentences or summary is Document Summarization.
- Extractive Summarization:
  - The model gets the important sentences in the document.
  - Advantage: Long and duplicate sentences as output.
  - Disadvantage: Reveals what the writer is willing to express.
- Abstractive Summarization:
  - The model generates sentences on its own. It tries to remove the unwanted information from the sentence.
  - Advantage: Creates very short summaries.
  - Disadvantage: The model is hard to train.
- A Ranking technique was employed to get the most important sentences(Extractive). The same model was improved by a document encoder to encode sentences and rank them after classification.
- A seq2seq with attention mechanism was used for headline generation which gives a short summary for a passage(Abstractive).

DIALOGUE SYSTEMS



- The well-known application of a Dialogue system is the Automated Customer Service.
- It is also called conversational machines or dialogue machines.
- <u>Task-based Systems:</u>
  - contains — Natural Language Understanding, Dialogue Manager, Natural Language Generation.
  - NLU gets the syntactic and semantic representation from the user's input context.
  - DM gives a closer response, by taking the output of NLU, to get the actual context.
  - NLG generates the actual utterance response from DM.
- <u>Non-task-based Systems:</u>
  - Non-task-based System creates a natural conversation with humans.
- LSTM models along with attention mechanism and sequence matching can be used.
- Single-turn Response Matching selects one suitable response.
- Multi-turn Response Matching considers the current input and previous outputs for creating an accurate response.

THANK YOU.