Introduction to NLP fundamentals

Prof. Luca Cagliero
Dipartimento di Automatica e Informatica
Politecnico di Torino

Politecnico
di Torino

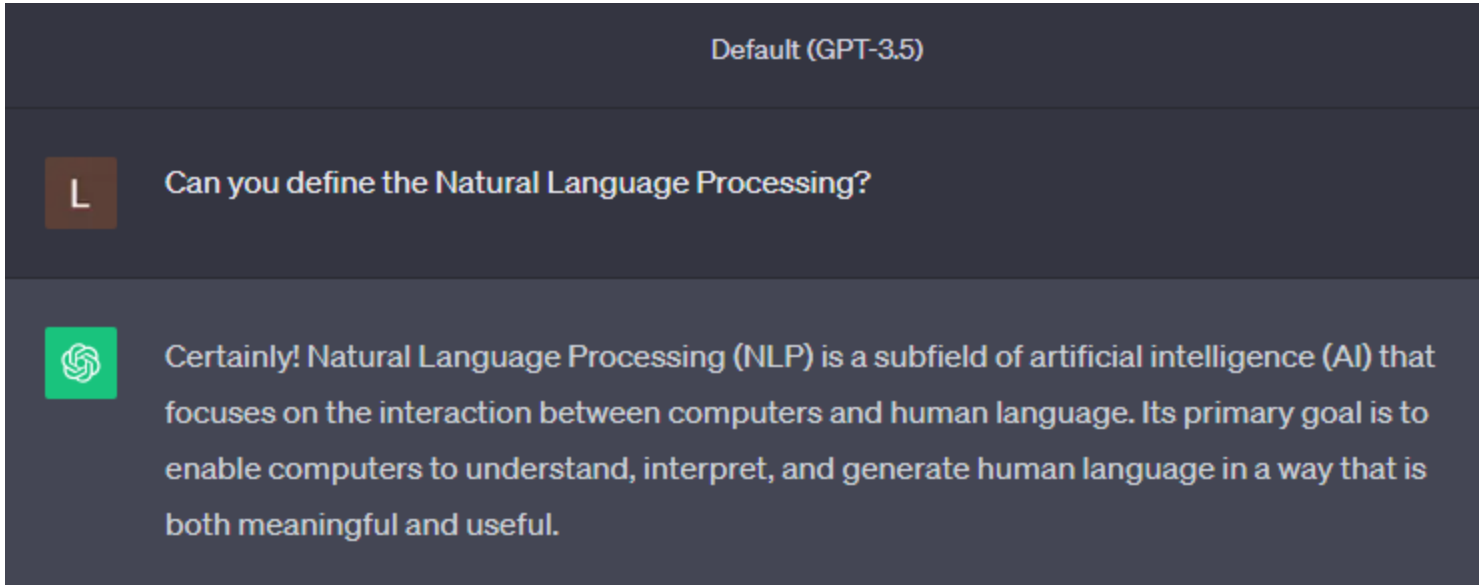
Lecture goal

- Introduction to Natural Language Processing
  - overview of the main objectives, techniques, and application scenarios
- Describe traditional NLP fundamentals
  - The NLP pipeline, text preprocessing

The data deluge

- Newspapers, communication agencies, broadcasters
- Social networks, blogs, and microblogs
- E-learning platforms, digital libraries
- Banks, rating agencies, auditing business


Natural Language Processing



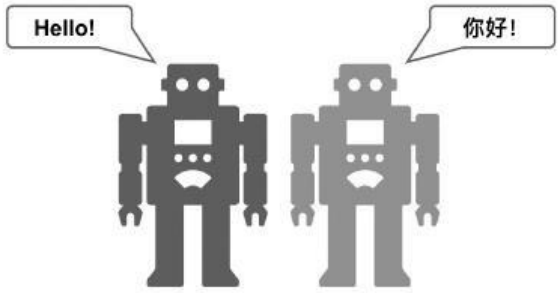
https://chat.openai.com/

Natural Language Processing

- Natural Language Processing (NLP) aims at enabling computers to process human languages intelligently
- Tight connection between several fields related to computer science and linguistics
  - Artificial Intelligence
  - Machine Learning
  - Cognitive science
  - Information processing
  - ...

Natural Language Processing

- NLP studies the connection between machine perception and languages
- Programming computers to fruitfully process large natural language corpora
- Examples of applications
  - Language understanding
  - Machine translation
  - Text generation
  - ...



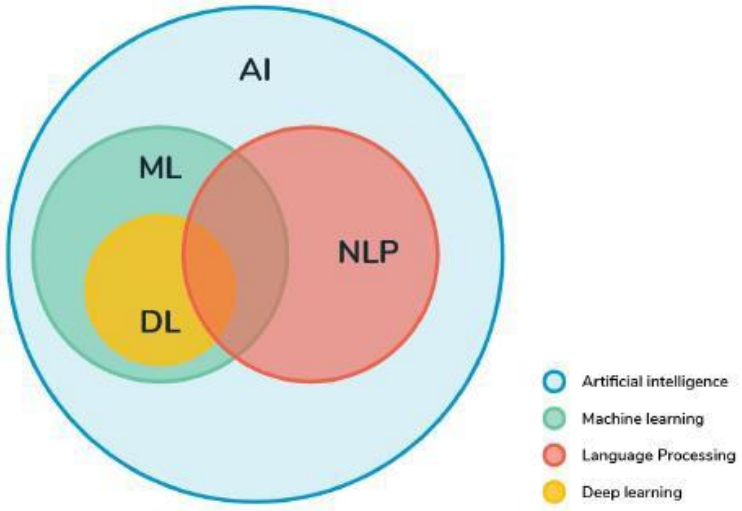
Manning C.D., Schuetze H. Foundations of Statistical Natural Language Processing. MIT Press. 10999. ISBN 9780262133609

Traditional vs. Deep Natural Language Processing

- Traditional NLP techniques heavily relied on ad hoc linguistic models, e.g.,
  - Syntactic rules
  - Dictionaries
  - Text generation
  - ...
- NLP does not necessarily rely on automated learning

Traditional vs. Deep Natural Language Processing

- Deep NLP techniques leverage Machine Learning models to automate the learning process



Related buzzword: text mining

- The process of deriving significant information from text
  - Focus on unstructured and semi-structured text

The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data. Ronen Feldman, James Sanger. 2007. ISBN: 9780521836579

Related buzzword: text mining

- The process of deriving significant information from text
- It encompasses
  - Information Retrieval (IR)
  - Machine Learning (ML)
  - Natural Language Processing (NLP)
  - Knowledge management (KM)

The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data. Ronen Feldman, James Sanger. 2007. ISBN: 9780521836579

Related buzzword: Natural Language Understanding

- Natural Language Understanding or Interpretation (NLU/NLI)
  - NLP subtopic that deals with machine reading comprehension using AI techniques
  - It encompasses
    - Text categorization
    - Entity Recognition and Disambiguation
    - . . .

The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data. Ronen Feldman, James Sanger. 2007. ISBN: 9780521836579

Related buzzword: Natural Language Understanding

- Natural Language Generation (NLG)
  - It aims at generating human-like text or speech based on structured data or specific instructions
  - It encompasses
    - Summarization
    - Machine translation
    - Question Answering
    - . . .

The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data. Ronen Feldman, James Sanger. 2007. ISBN: 9780521836579

Overview of NLP applications

- NLP techniques find application in a wide range of knowledge discovery and decision support systems
- They are commonly part of a **data science pipeline**

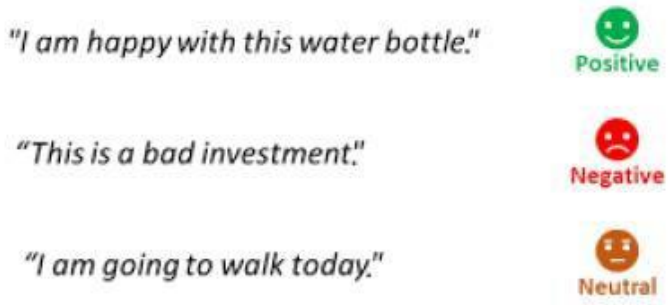
The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data. Ronen Feldman, James Sanger. 2007. ISBN: 9780521836579

Overview of NLP applications

- Hereafter we will focus on
  - Sentiment analysis
  - Text categorization
  - Machine translation
  - Question answering
  - Topic modelling
  - Text summarization

Sentiment analysis

- Extract writer's feeling, opinion, emotions, likes/dislikes
  - Also known as **opinion mining**
- Identify the opinion/human behavior of a person from plain text
- It can rely either on traditional NLP rules or on Machine Learning



Sentiment analysis

- Use cases
  - Hotel review analysis
  - News trading
  - Hate speech detection
  - Advertisement placing
  - . . .

Example of sentiment analyzer: VADER

- Input
  sentiment_analyzer_scores("The phone is super cool.")
- Output
  The phone is super cool --- {'neg': 0.0, 'neu':0.326, 'pos': 0.674, 'compound': 0.7351}

https://github.com/cjhutto/vaderSentiment

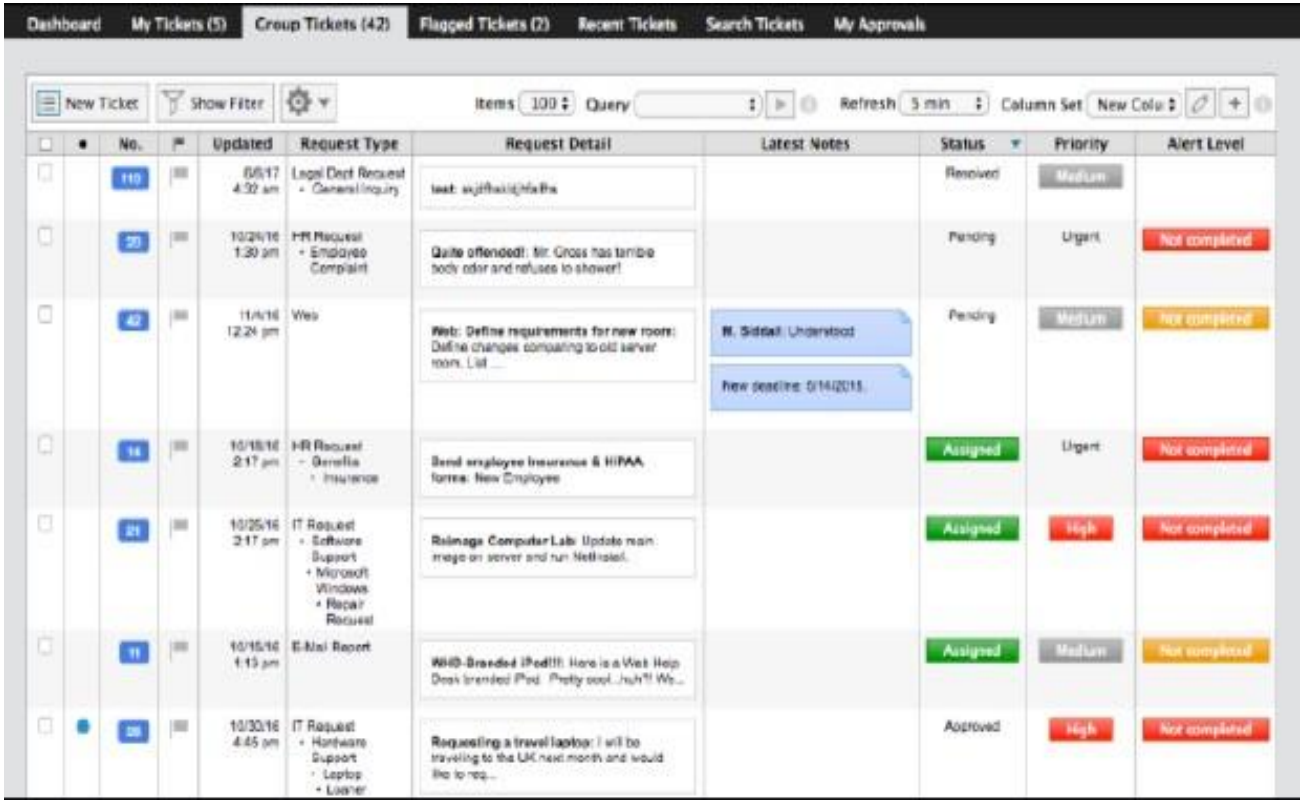
Text categorization

- Assign a predefined label to a document or a text snippet
  - Also known as text classification
- Supervised process
  - given a set of labeled documents/snippets, learn how to automatically label new, unlabeled ones
- If multiple labels are allowed
  - multi-label text classification
- It can rely either on traditional NLP rules or on Machine Learning

Text categorization

- Use cases
  - Spam detection
  - Ticket management
  - Automatic text annotation
  - Document management
  - . . .

Automated ticket labeling in help desk applications



https://www.webhelpdesk.com (latest access: April 2021)

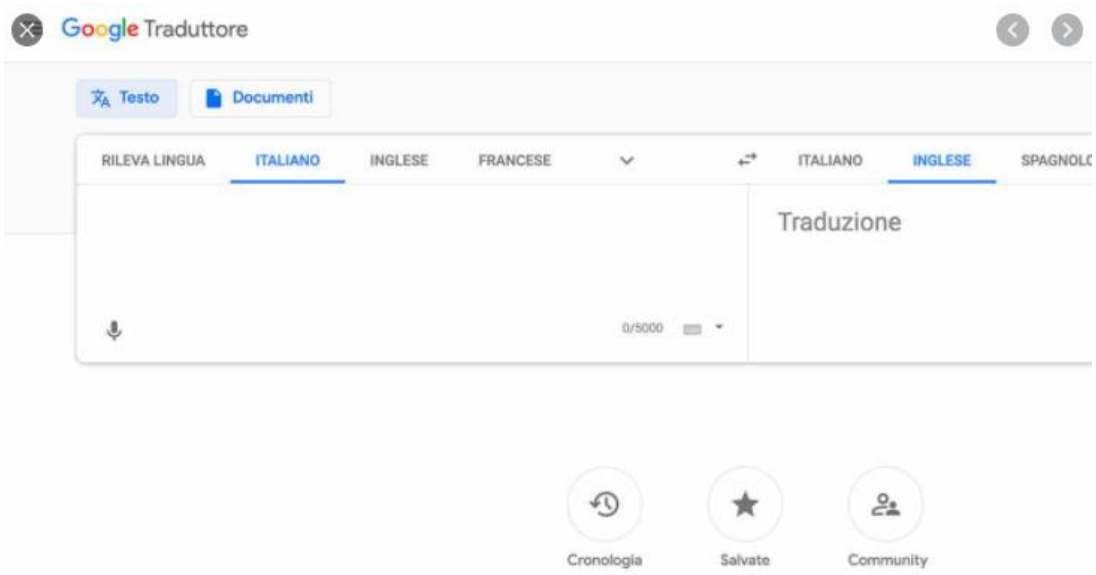
Machine translation

- Automatic translation of text or speech from one language to another
- Approaches
  - Rule- or dictionary-based
  - Statistical methods
  - Deep Learning

memoq (latest access: April 2021)

Rule-based Machine Translation

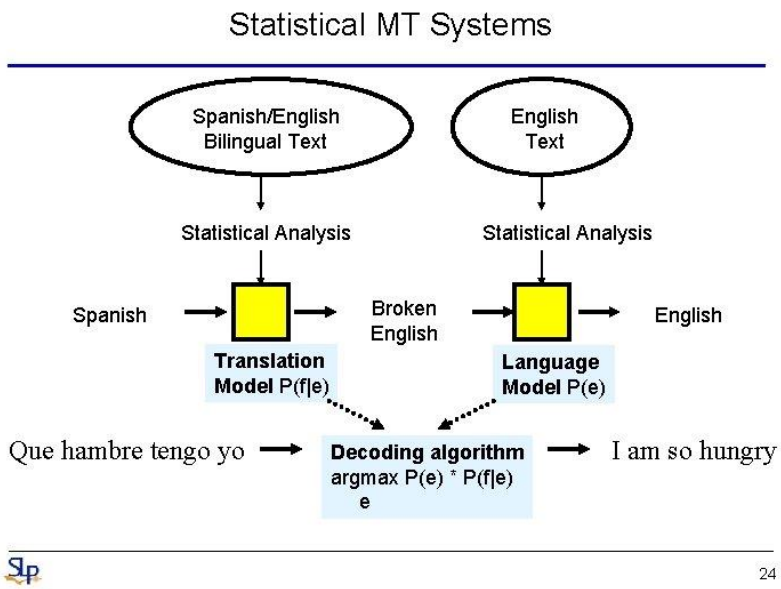
- Machine translation systems based on linguistic information about source and target languages basically retrieved from (bilingual) dictionaries and grammars
- Also known as **Knowledge-Based Machine Translation** or **Classical MT Approach**



https://translate.google.com (latest access: April 2021)

Statistical Machine Translation

- Based on statistical models whose parameters are derived from the analysis of bilingual text corpora.
- Key ideas
  - every sentence in one language is a possible translation of any sentence in the other
  - the most appropriate is the translation that is assigned the highest probability by the system



Introduction to Statistical Machine Learning. Raghav Bashyal.

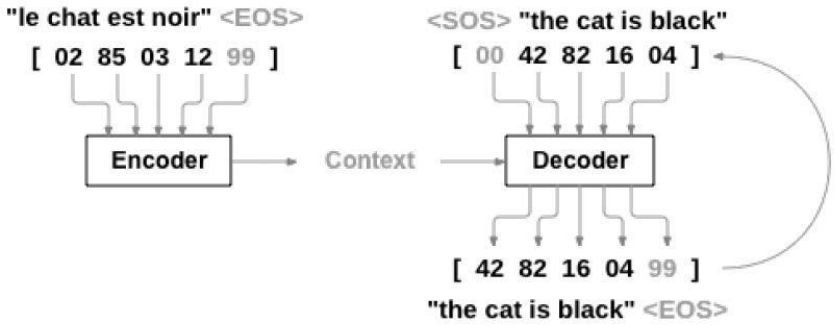
Example-based Machine Translation

- Bilingual corpus with parallel texts
  - set of sentences in the source language (from which one is translating)
  - the corresponding translations of each sentence in the target language with point-to-point mapping
- Translate by analogy

Machine Translation Approaches: Issues and Challenges. M.D. Okpor. IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 5, No 2, September 2014
ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784

Neural Machine translation: the encoder decoder architecture

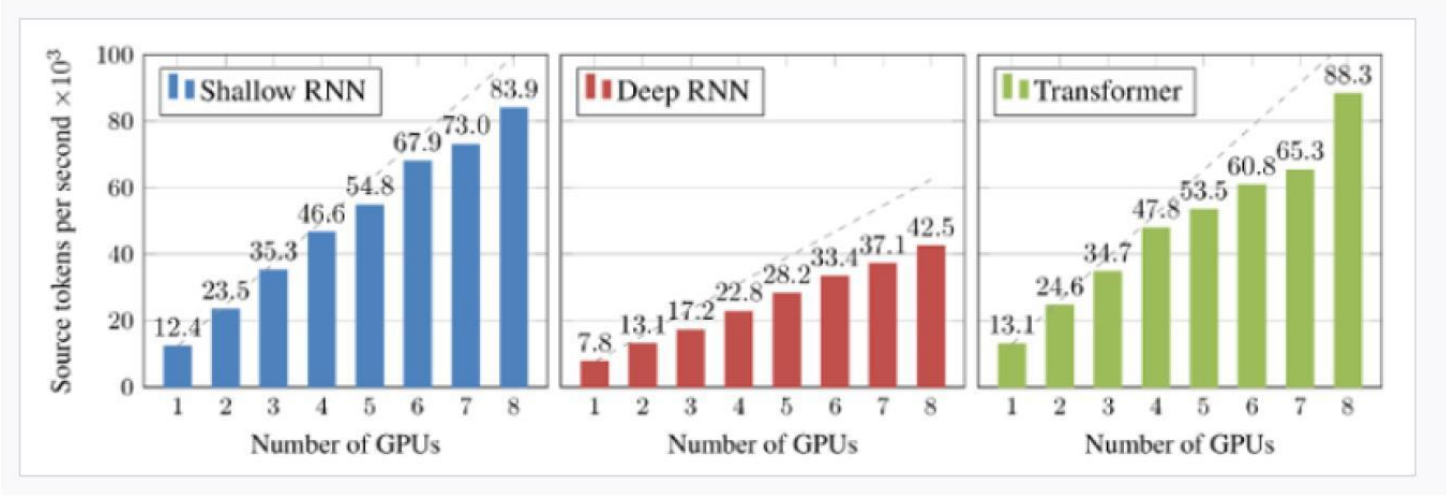
- Input and output are both variable-length sequences
- Encoder: it transforms the input into the encoded vector
  - The vector encapsulates the information for all input elements in order to help the decoder make accurate predictions
- Decoder: maps the encoded state of a fixed shape to a variable-length sequence
  - It codes the state to generate the translated sequence token by token as the output



https://www.google.com (latest access: April 2021)

Neural Machine translation: MarianMT

- Opensource generative Deep Learning model
  - Relies on Seq2Seq Transformers
- Behind the Microsoft Neural Machine Translation service



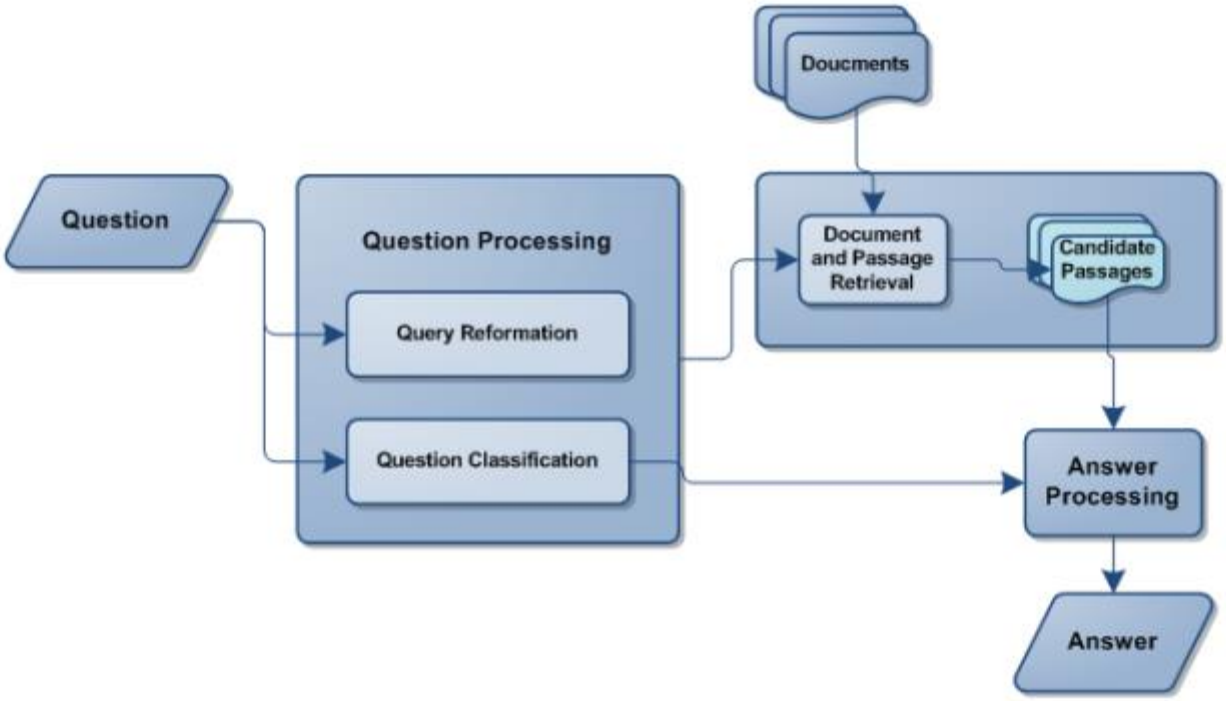
Question answering

- Specialized area in the field of Information Retrieval
- It aims at providing relevant answers in response to questions posed in natural language
- Main steps
  - Question classification
  - Information Retrieval
  - Answer extraction

Question answering

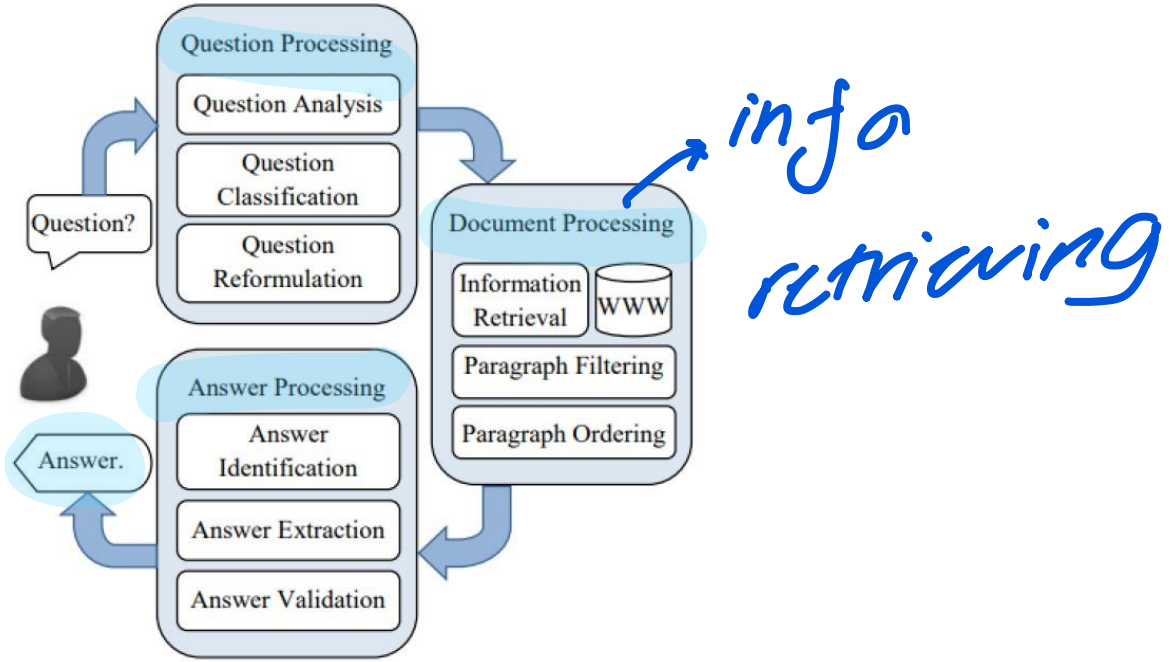
- Key properties
  - Source: Web pages, knowledge bases, social data collections, etc.
- Question formulation
  - domain-specific, factoid, etc.
- Answer type
  - word, sentence, paragraph, image, video, etc.
- Practical use cases
  - Chatbots
  - Remote medical assistance
  - ...

Question answering: traditional approach



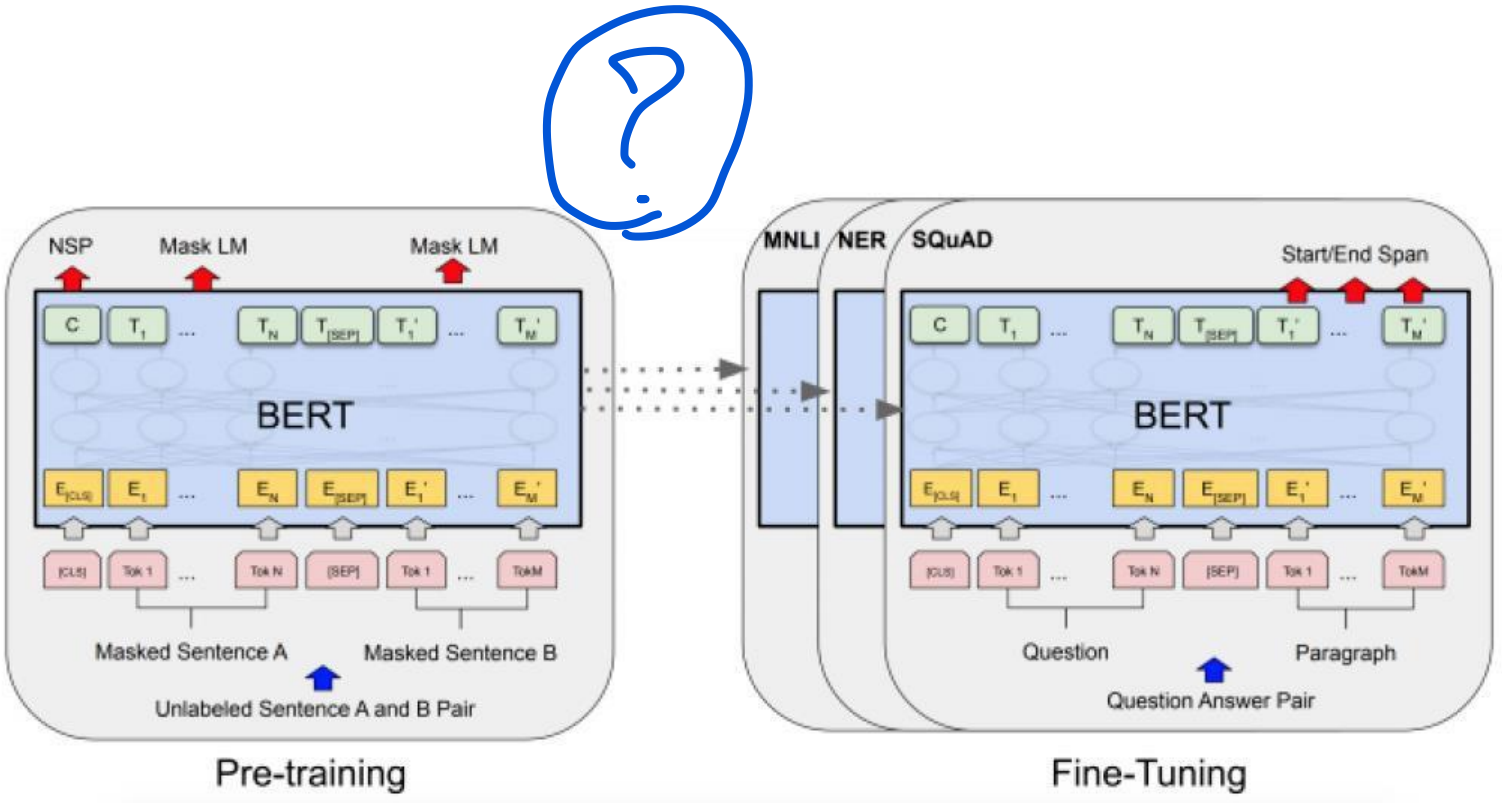
fA Survey of State-of-the-Art Methods on Question Classification Babak Loni

Question answering: traditional approach



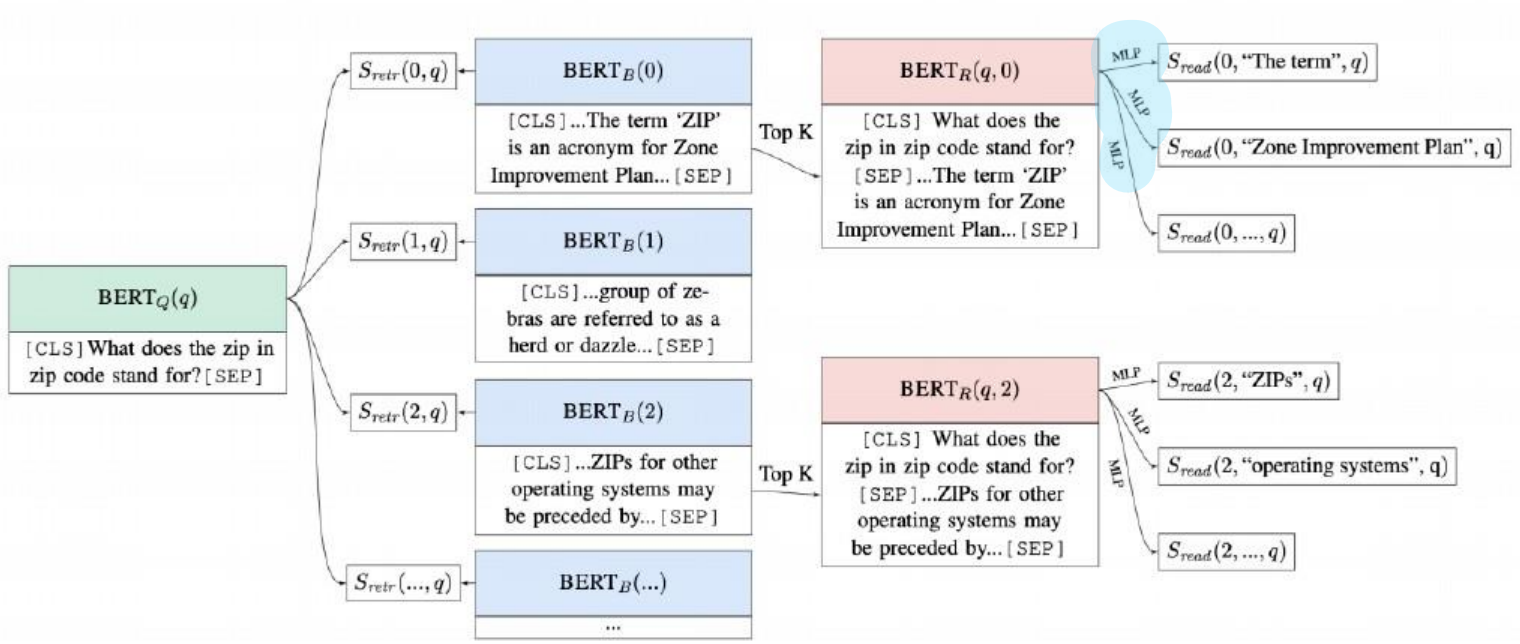
The Question Answering Systems: A Survey. Ali Mohamed Nabil Allam, Mohamed Hassan Haggag. International Journal of Research and Reviews in Information Sciences. 2012

Question Answering: state-of-the-art approach based on Transformers (by Google Inc.)



BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. EMNLP 2019

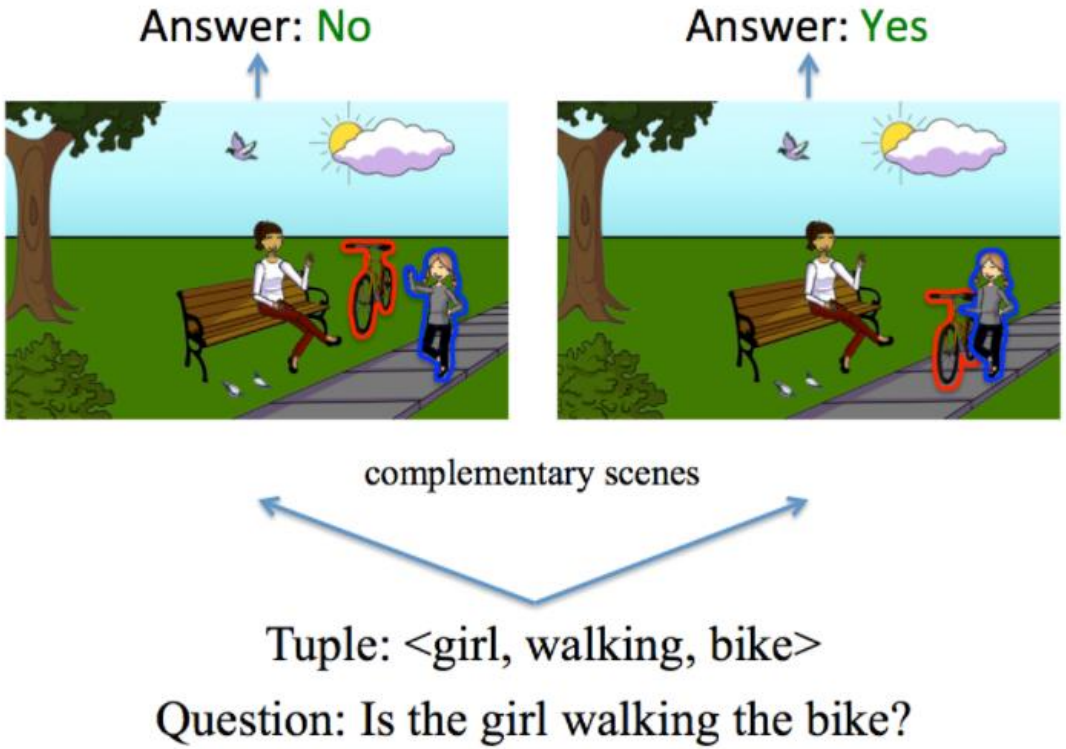
Question answering: state-of-the-art approach based on Transformers (by Google Inc.)



Latent Retrieval for Weakly Supervised Open Domain Question Answering. Kenton Lee Ming-Wei Chang Kristina Toutanova. 2019. https://arxiv.org/pdf/1906.00300.pdf

The Visual QA project

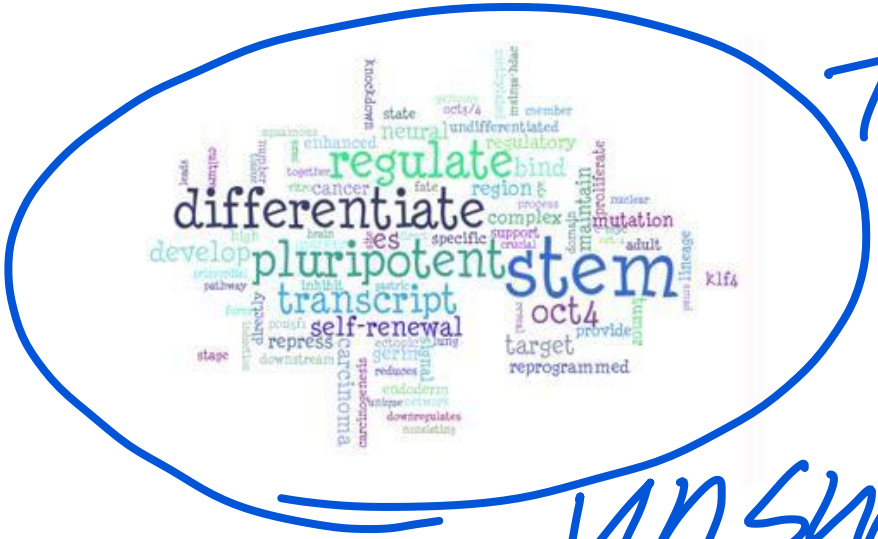
- Combine language understanding with image processing
- understanding of vision, language, and commonsense knowledge to transfer



https://visualqa.org/ (latest access: April 2021)

Topic modelling

- Unsupervised machine learning techniques aimed at detecting the key word and phrase patterns within a document collection
- A topic model is a statistical document representation of the abstract word groups and similar expressions that best characterize a set of documents



Topic modeling

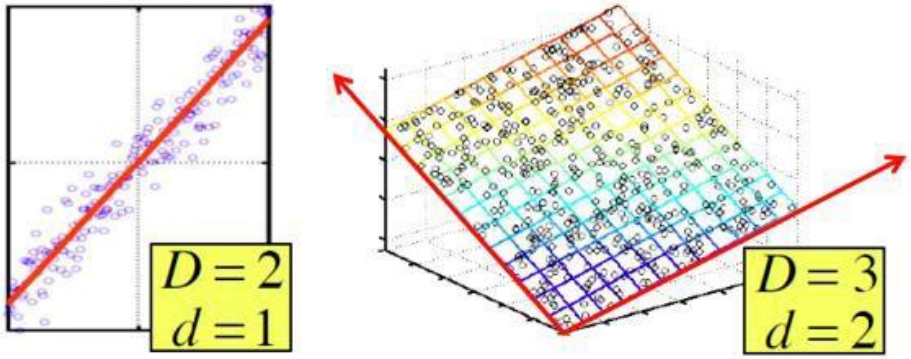
unsupervised

Topic modelling

- Main techniques
  - Latent Semantic Indexing (LSI)
  - Latent Dirichlet Allocation (LDA)
  - Aspect modelling
- Practical use cases
  - Information retrieval
  - Data visualization
  - ...

Topic modelling

- Dimensionality reduction
  - Data lies on or near a low d-dimensional subspace
  - Axes of this subspace are effective representation of the data



Mining of Massive Datasets. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman. 2nd Ed. Cambridge University Press 2014, ISBN 978-1107077232

Text summarization

- Shorten large textual document collections
  - Produce a concise summary that incorporates the most salient content

Transcript

today we are going to show you how to make spanish omelet . i 'm going to dice a little bit of peppers here . i 'm not going to use a lot , i 'm going to use very very little . a little bit more then this maybe . you can use red peppers if you like to get a little bit color in your omelet . some people do and some people do n't .... t is the way they make there spanish omelets that is what she says . i loved it , it actually tasted really good . you are going to take the onion also and dice it really small . you do n't want big chunks of onion in there cause it is just pops out of the omelet . so we are going to dice the up also very very small . so we have small pieces of onions and peppers ready to go .

Summary

how to cut peppers to make a spanish omelette; get expert tips and advice on making cuban breakfast recipes in this free cooking video .

Multimodal Abstractive Summarization for How2 Videos. Shruti Palaskar, Jindrinch Libovicky, Spandana Gella, Florian Metze. https://arxiv.org/abs/1906.07901

Text summarization

- Input
  - Single document
  - Multi-document collection
  - Multimodal data
- Output
  - Extractive summary
  - pick existing content
  - Abstractive summary
  - generate new content
- Approach
  - Supervised
    - e.g., neural summarization
  - Unsupervised
    - e.g., itemset mining, clustering, Latent Semantic Analysis, graph ranking

Text summarization

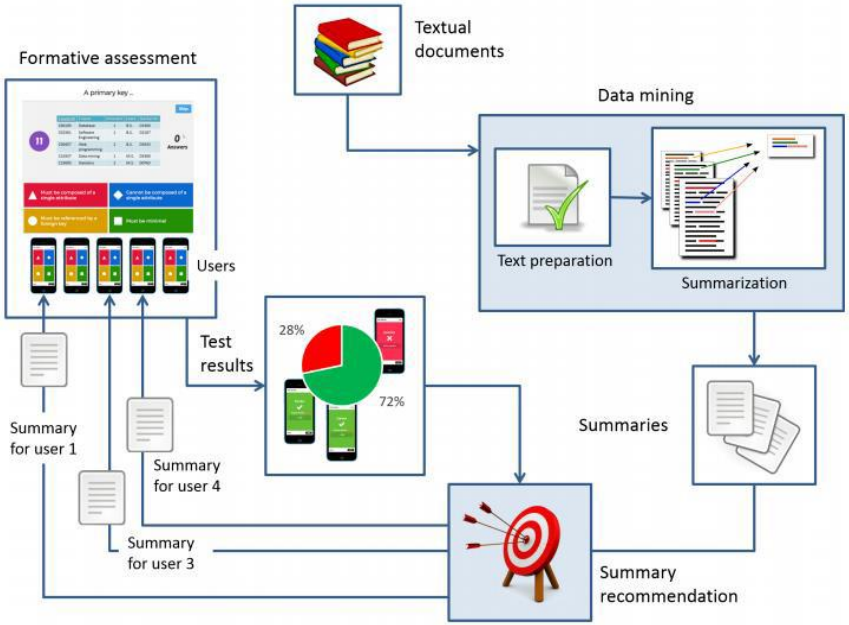
- Language
  - Single-language
  - Multi-lingual
  - Cross-lingual
- Domain specificity
  - General-purpose
  - Context-dependent
  - Query-driven
- Time dependency
  - Temporal summarization
  - Timeline summarization
  - Incremental summarization

Text summarization

- Practical use cases
  - Content curation
  - Accessibility
  - Learning analytics
  - Highlight extraction

Examples of summarization methods

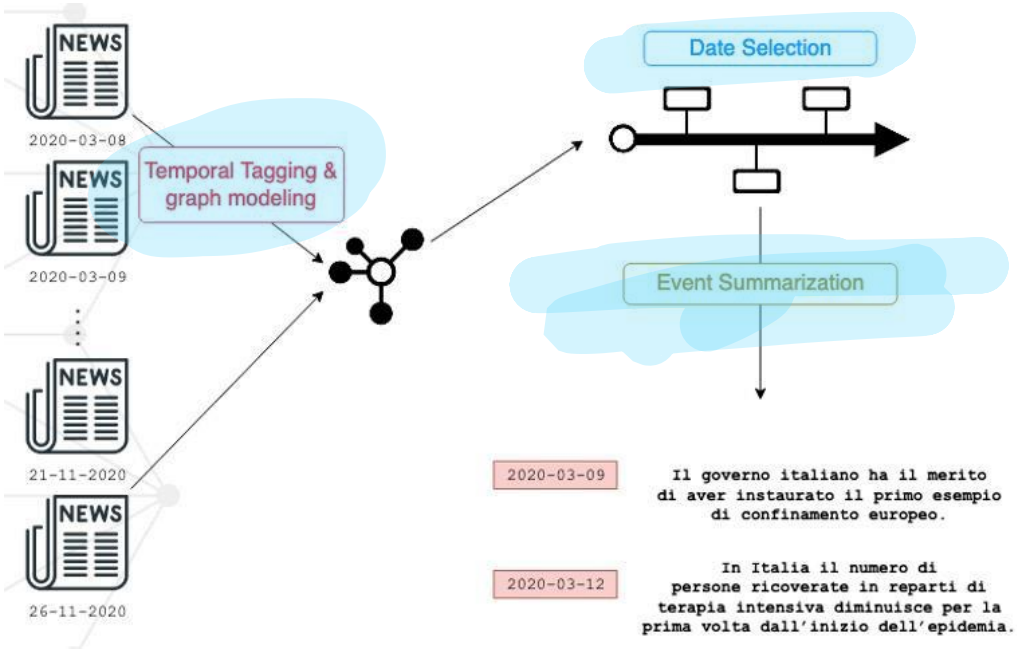
- Summarization of teaching materials for formative assessment



Recommending Personalized Summaries of Teaching Materials. Luca Cagliero, Laura Farinetti, Elena Baralis. IEEE Access. 2019

Examples of summarization methods

- TimeLine Summarization of news articles



Summarize Dates First: a paradigm shift in timeline summarization. Moreno La Quatra, Luca Cagliero, Elena Baralis, Maurizio Montagnulo, Alberto Messina. ACM SIGIR 2021

Text structure

- A text can be divided into units
  - Useful for enabling automated text processing
  - Crucial for extracting relevant knowledge

- Different aggregation levels
  - According to the objective
  - According to the language

Text structure

- Character: smallest text unit
  - Letter, digit, space, special character
  - Low utility unless enriched with positional information
    - Two words are commonly separated by a space
    - A full stop indicates the end of a sentence
    - A question mark indicates an interrogative form

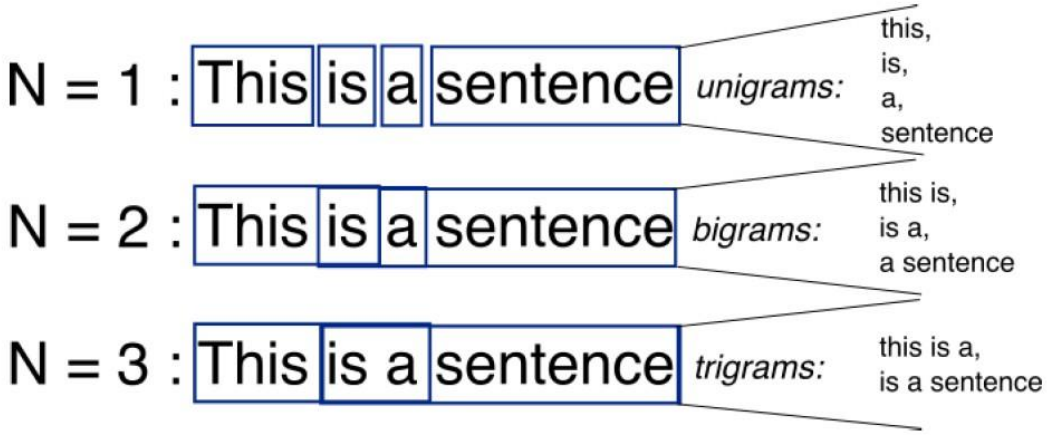
Text structure

- Word
  - series of letters between spaces
  - smallest unit of language that conveys meaning to the majority of people you would like to reach
  - Readers must agree with the language
- N-gram
  - contiguous sequence of N textual units
  - Units can be phonemes, syllables, letters, words
  - Syntactic and semantic relevance are not guaranteed
  - Unigram: 1-gram, bigram: 2-gram, etc.
- Multi-word expressions
  - textual form made up of at least two lexical items

Text structure

- Sentence
  - Text snippet separated by punctuation
    - e.g., full stop, question mark, exclamation mark
- Phrase
  - Part of a sentence consisting of a group of words without subject and verb
    - E.g., the temporal phrase after dinner or the impersonal form waiting for the bus
  - They do not convey a complete thought
- Clause
  - Part of a sentence including a subject and a predicate
  - It can be used by itself as a complete sentence
    - e.g., the independent clause the dog barked at him
  - It provides more insights (e.g., the subordinate clause when the old man came)

Text structure: examples of n-grams

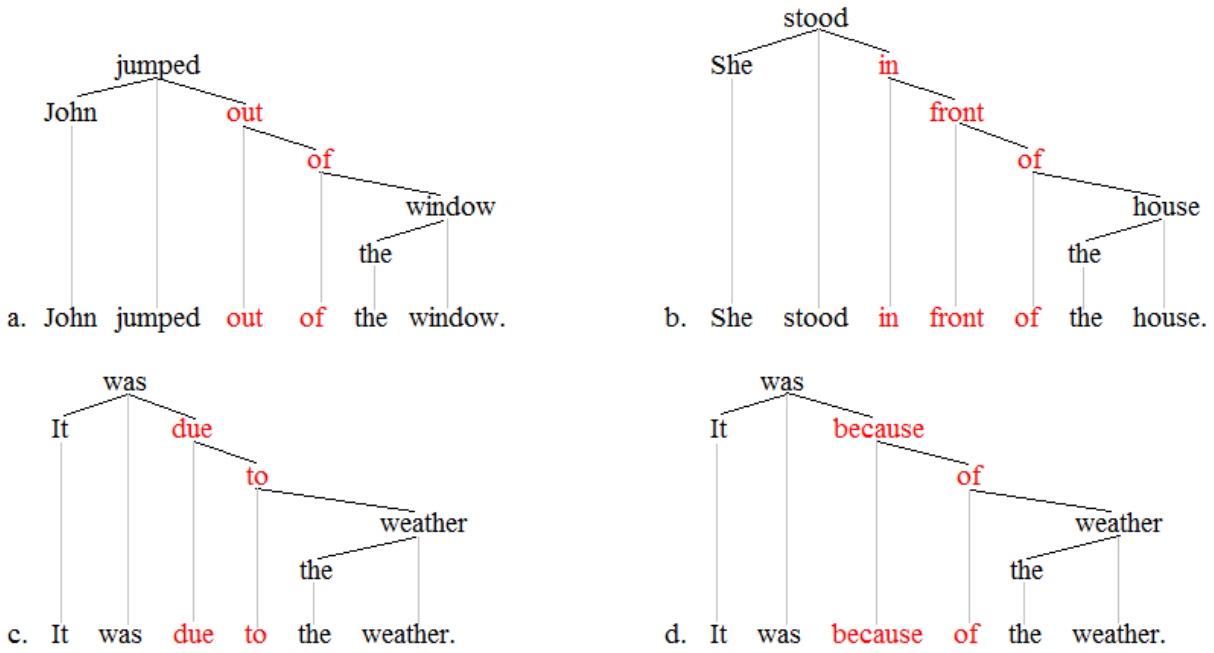


Source: https://deepai.org/ (latest access: April 2021)

Text structure

- Lexical chain
  - sequence of related words in writing
    - spanning short (adjacent words or sentences) or long distances (entire text)

Text structure: examples of lexical chains

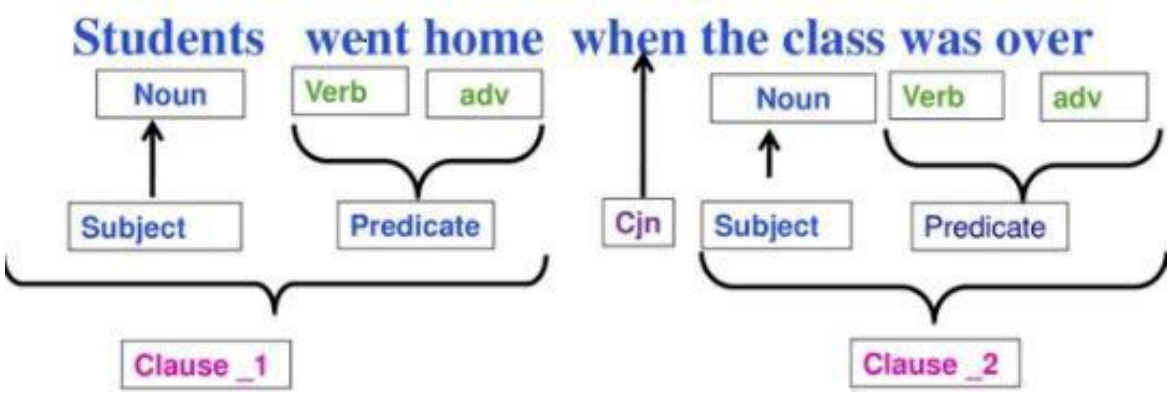


Source: https://Wikipedia.org (latest access: April 2021)

Text structure

- Lemma
  - canonical form of a word or a multi-word expression, chosen from a set of candidate forms in a dictionary
- Lexeme
  - set of all single words or multi-word forms that have the same meaning
  - A lemma is a particular form that is chosen by convention to represent the lexeme
    - E.g., the same word with different inflectional endings
- Stem
  - base form of a word (similar to a lemma) not necessarily derived from a dictionary
  - Typically, static rules to transform inflected forms
- Bag-Of-Word (BOW)
  - unordered set of words
  - Repetitions are removed

Text structure: examples of clauses and phrases

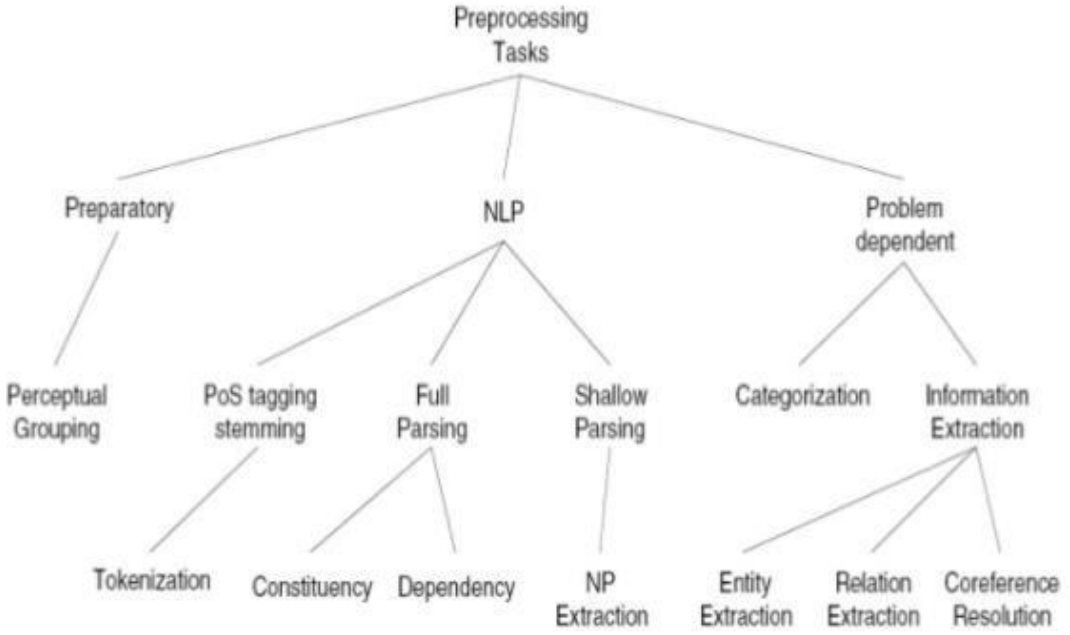


Source: https://knowitinfo.com/ (latest access: April 2021)

Text structure: sectioning

- Paragraph
  - Portion of text consisting of a sequence of sentences
  - Paragraphs can be further partitioned into subparagraphs
- Section
  - Portion of text consisting of consecutive paragraphs
  - Sections can be further partitioned into subsections and subsubsections

Text preprocessing: the taxonomy



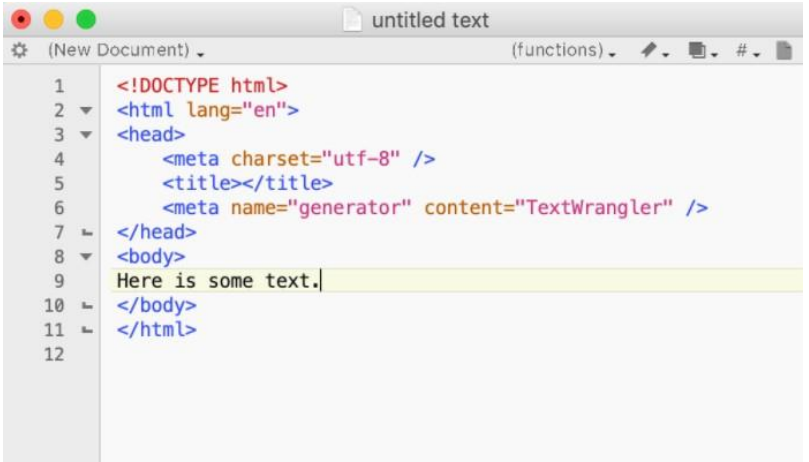
The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data. Ronen Feldman, James Sanger, 2007. ISBN: 9780521836579

Text preprocessing steps

- Cleaning
  - basic filtering steps to remove noise, errors, and redundant content
- Tokenization
  - divide the raw text into units and sub-units
- Stopword elimination
  - remove too frequent words with little semantic meaning
- Part-Of-Speech tagging
  - annotate the text words with the corresponding role in the sentence
- Lemmatization and stemming
  - map word inflections and derivates to their canonical form

Text cleaning

- Remove special characters and extra spaces
  - unrecognized symbols, OCR errors, etc.
- Case normalization
  - handle upper and lower cases
- Data format conversion
  - E.g., HTML, XML, JSON, CSV
  - May require a data schema
- Handle non-textual content



Language dependence

- Cleaning ✓ (partially)
- Tokenization ✓
- Stopword elimination ✓
- Part-Of-Speech tagging ✓
- Lemmatization and stemming ✓

What if you do not know the source language in advance?

Tokenization

- Continuous character stream in documents is broken up into meaningful constituents
- Break text into relevant units and sub-units
- Identify sentence boundaries based on punctuation, capitalization, etc.
- Strongly language-dependent

Source: https://kddnuggets.com (latest access: April 2021)

Example of tokenizer

Text tokenization utility

Source: `R/preprocessing.R`

Vectorize a text corpus, by turning each text into either a sequence of integers (each integer being the index of a token in a dictionary) or into a vector where the coefficient for each token could be binary, based on word count, based on tf-idf...

```
text_tokenizer(
  num_words = NULL,
  filters = "!\"#$%&()*+,-./:;<=>?@[\\]^_`{|}~\t\n",
  lower = TRUE,
  split = " ",
  char_level = FALSE,
  oov_token = NULL
)
```

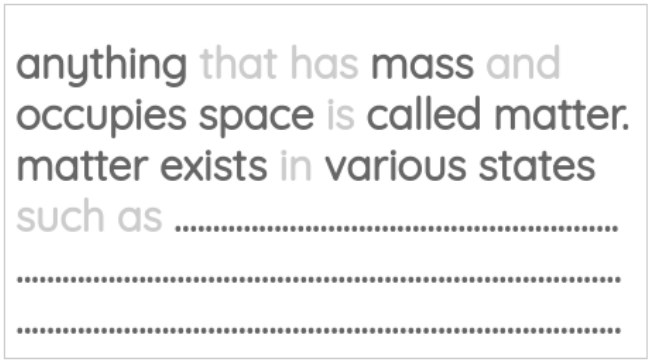
Arguments

num_words	the maximum number of words to keep, based on word frequency. Only the most common `num_words` words will be kept.
filters	a string where each element is a character that will be filtered from the texts. The default is all punctuation, plus tabs and line breaks, minus the ' character.
lower	boolean. Whether to convert the texts to lowercase.
split	character or string to use for token splitting.
char_level	if `TRUE`, every character will be treated as a token
oov_token	`NULL` or string If given, it will be added to `word_index`` and used to replace out-of-vocabulary words during text_to_sequence calls.

Source: https://keras.rstudio.com/reference/text_tokenizer.html (latest access: April 2021)

Stopword elimination

- Stop words refer to the most common words in a language
- They convey a limited amount of information
  - E.g., prepositions, articles, conjunctions
- They are filtered out prior to text processing



Source: https://kddnuggets.com (latest access: April 2021)

Stopword elimination

- Language-dependent step
- Some particular Deep NLP models do not require stopword elimination
- A universal list of stop words used by all natural language processing tools is not available
- Ideally, any group of words can be chosen as the stopwords for a given purpose
  - Different search engines use different stopword lists
  - Some of them remove lexical words, such as want, from a query in order to improve performance

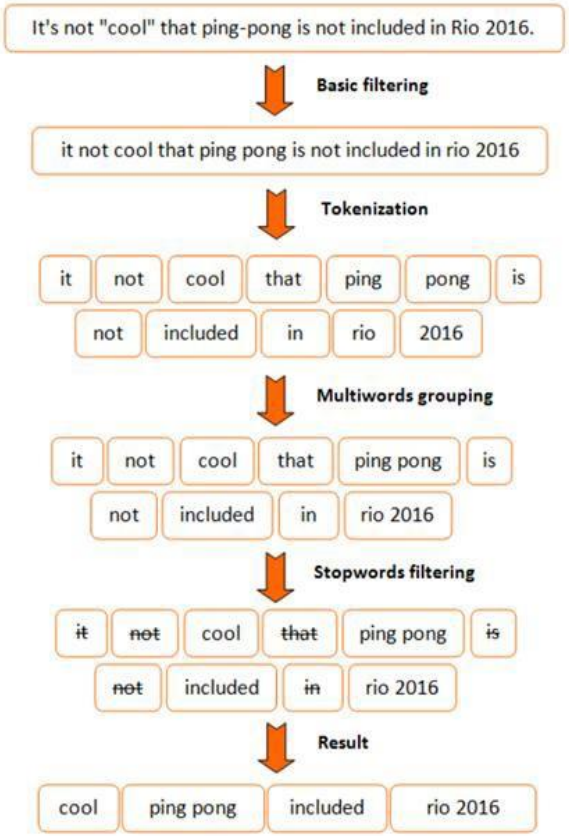
Example of stopword list

NLTK's list of english stopwords

```
i
me
my
myself
we
our
ours
ourselves
you
your
yours
yourself
yourselves
he
him
his
himself
she
her
hers
herself
it
its
itself
they
them
their
theirs
themselves
what
which
who
```

Source: https://gist.github.com/sebleier/554280 (latest access: April 2021)

Example of text preprocessing pipeline



Source: www.meaningcloud.com (latest access: April 2021)

Example of text preprocessing pipeline

- Caveat
  - By removing stopword *Not* the meaning of the text changes!
- To capture the semantics behind the text, stopword elimination is deprecated
  - Deep NLP techniques are recommended

Part-Of-Speech tagging

حرف، نحو، خود، حق

- Part-Of-Speech (POS)
  - word category assigned according to the role it plays in the sentence
- POS tagging
  - label the words in a text with the corresponding POS
    - E.g., article, noun, verb, adjective, preposition, number, and proper noun
- Context and language dependent
- Relies on morphological word analyses ?

Part-Of-Speech tagging example

Source: https://dataingovernment.blog.gov.uk/ (latest access: April 2021)

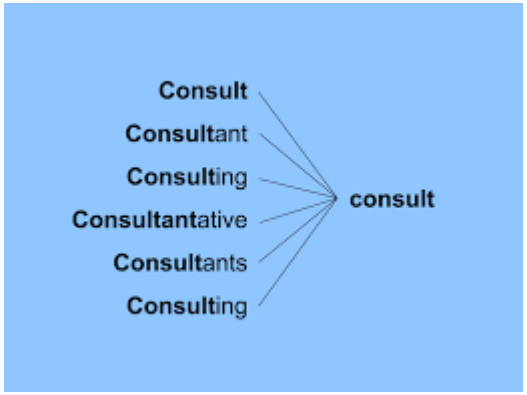
Lemmatization

- Process of determining the lemma of a word by deducing its intended meaning
  - It groups together the inflected forms of a word
  - They can be considered as the same thing and identified by a common lemma
- It can perform together with POS tagging
- It may require analyzing the sentence- or document-level context
  - E.g., article, noun, verb, adjective, preposition, number, and proper noun
- Context and language dependent
- Relies on morphological word analyses

Stemming

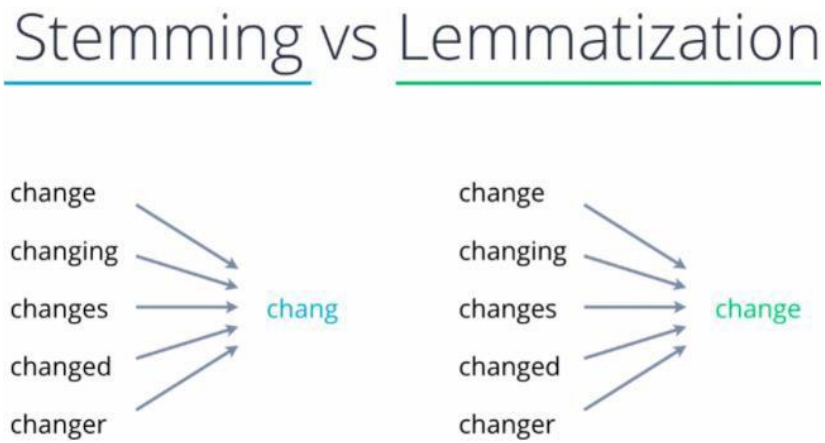
- Simplified version of the lemmatization process
- It determines the stem of the word instead of the lemma
- Lower computational complexity
- Approximated result
- Acceptable in most cases

Form	Suffix	Stem
studies	-es	studi
studying	-ing	study
niñas	-as	niñ
niñez	-ez	niñ



Source: https://devopedia.org/ (latest access: April 2021)

Stemming vs. lemmatization



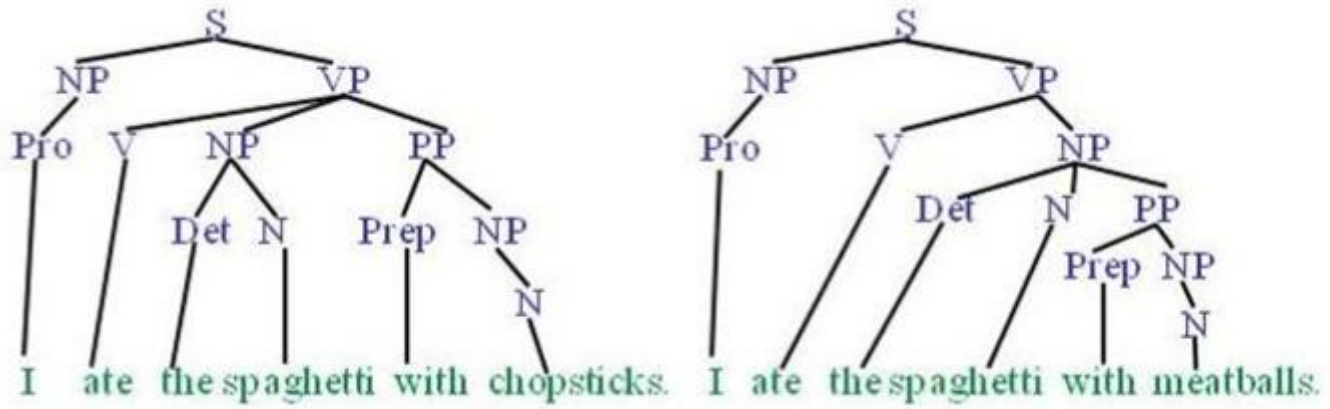
Source: https://dataingovernment.blog.gov.uk/ (latest access: April 2021)

Syntactic parsing

relating to syntaxes.

- Syntactical analysis of the text according to a certain grammar theory
- Constituency grammars describe the syntactical structure of the sentence in terms of recursively built phrases,
  - i.e., sequences of syntactically grouped elements
- Dependency grammars analyze the dependencies between words
  - E.g., an adjective depends on the associated noun
- High computational complexity
- Low quality in many real contexts

Syntactic parsing

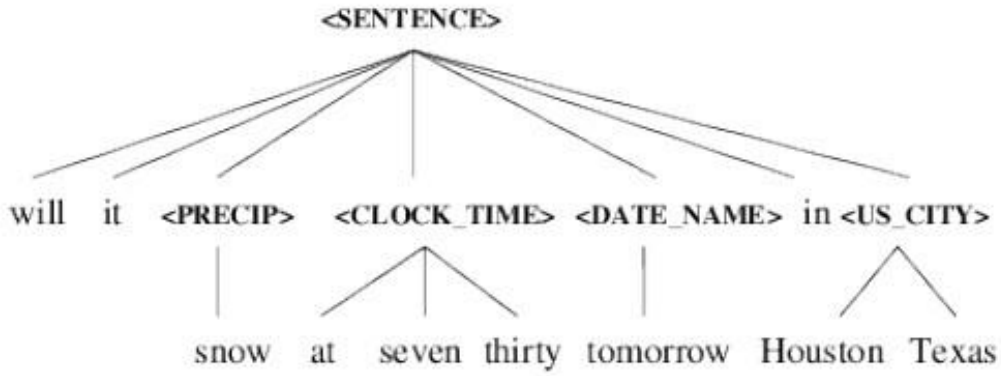


Parsing example: "book that flight"

$S \rightarrow NP\ VP$	$Det \rightarrow that \mid this \mid a$
$S \rightarrow Aux\ NP\ VP$	$Noun \rightarrow book \mid flight \mid meal \mid money$
$S \rightarrow VP$	$Verb \rightarrow book \mid include \mid prefer$
$NP \rightarrow Pronoun$	$Pronoun \rightarrow I \mid she \mid me$
$NP \rightarrow Proper\text{-}Noun$	$Proper\text{-}Noun \rightarrow Houston \mid TWA$
$NP \rightarrow Det\ Nominal$	$Aux \rightarrow does$
$Nominal \rightarrow Noun$	$Preposition \rightarrow from \mid to \mid on \mid near \mid through$
$Nominal \rightarrow Nominal\ Noun$	
$Nominal \rightarrow Nominal\ PP$	
$VP \rightarrow Verb$	
$VP \rightarrow Verb\ NP$	
$VP \rightarrow Verb\ NP\ PP$	
$VP \rightarrow Verb\ PP$	
$VP \rightarrow VP\ PP$	
$PP \rightarrow Preposition\ NP$	

Shallow parsing

- Simplified text parsing
- Syntactical analysis only of the text snippets that are unambiguous
  - E.g., small and simple noun and verb phrases
- Trade-off between complexity and accuracy



Integration Of Supra-Lexical Linguistic Models With Speech Recognition Using Shallow Parsing And Finite State Transducers. Xiaolong Mou, Stephanie Seneff, Victor Zue. MIT press. 2003

Named Entity Recognition

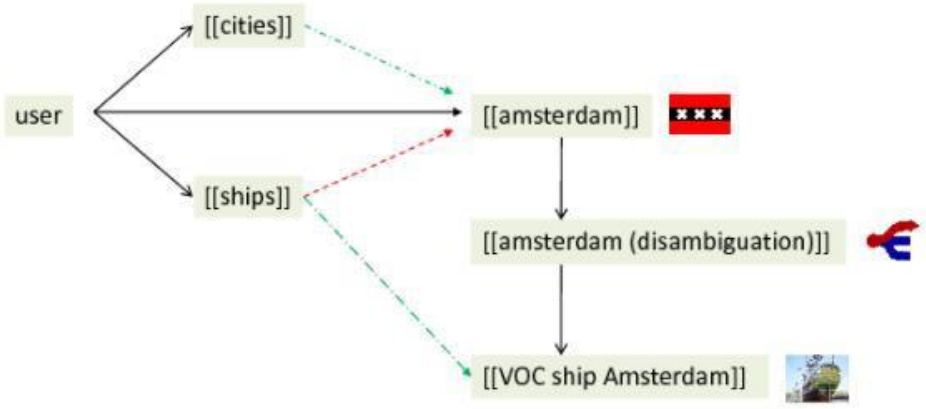
- Locate and classify names in text
- Identify references to information units called entities
- Entities are semantically rich (ontology-based) descriptions of concepts/bodies
  - E.g., persons, organization and location names, numeric
  - expressions including time, date, money, and percent expressions, times, dates, proteins, etc.

Named Entity Recognition

Named Entity Recognition and the Stanford NER Software. Jenny Rose. Stanford University. 2007 27

Word sense disambiguation

- Identifying which sense of a word is used in a sentence, when the word has multiple meanings
- Approaches
  - knowledge base exploration
  - Machine Learning techniques
  - a combination of the above
- Acronym NERD stands for "Named Entity Recognition and Disambiguation"



Named Entity Recognition and the Stanford NER Software. Jenny Rose. Stanford University. 2007 27

Word disambiguation example: “cars”

- Same Google query, different results

Text representation

- Transform the raw text into a data representation suitable for applying Machine Learning and data mining algorithms
  - Most algorithms require a structured data model (relational, key-value form, or equivalent)
  - A fixed-sized model is requested
- Main issue
  - Move from and unstructured or semi-structured text representation to a structured form

Text representation

- Fully unstructured document
  - Raw text without sectioning and paragraphs

Customer XYZ called about Plan A promotion. Explained plan.
Customer thinks roll-over minutes should be included.

Customer ABC called about Plan A promotion. Customer thought
it was ridiculous that roll-over minutes were not in plan.

Potential called about Plan A promotion. Said that plan was expensive.

Potential called about Plan A promotion. Said that 4GB data not enough.

Customer XYT called about Plan A promotion. Said
that data plan was insufficient and stupid.

Text representation

- Weakly structured document
  - Text organized in sections, paragraphs according to a predefined format

1 Section

Hello World!

1.1 Subsection

Structuring a document is easy!

1.1.1 Subsubsection

More text.

Paragraph Some more text.

Subparagraph Even more text.

2 Another section

Text representation

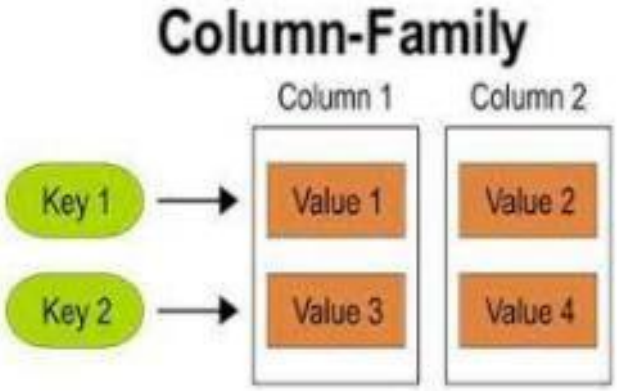
- Semi-structured document
  - Text annotated with tags or with a overlay structure based on markups
    - e.g., XML, HTML

```
<CATALOG>
   <SPRING>
      <TITLE>Garden Sales</TITLE>
      <LINE>Outdoor_Tools</LINE>
      <PAGE>
         <CAPTION>Goodbye, Winter!</CAPTION>
         <ITEM>Gardening Gloves</ITEM>
         <ITEM>Potting Soil</ITEM>
      </PAGE>
   </SPRING>
</CATALOG>
```

Feature-value data model

- Structured way to represent the main information
- A set of features summarizing the key properties of the text
  - E.g., the words that occur in it
- A set of textual units
  - E.g., the sentences

Text representation example: column-based DBs



Feature-value data model

- Structured way to represent the main information
- A set of features summarizing the key properties of the text
  - E.g., the words that occur in it
- A set of textual units
  - E.g., the sentences

Feature-value data model: main steps

- Feature engineering
  - define a set of features and collect the corresponding values
    - considering various textual units and heuristics based on the application domain
- Data transformation
  - build a structured representation of the original text that incorporates the considered features

Feature engineering

- Dimensionality reduction
  - Reduce the number of features to alleviate the issues related to high-dimensional data
  - Main methods
    - Latent Semantic Indexing, Latent Dirichlet Analysis, Principal Component Analysis
- Feature selection
  - Reduce the number of features to improve the performance of Machine Learning models
    - e.g., sentiment analysis models, text classifiers

Feature selection

- Unsupervised methods
  - Do not use the target variable (i.e., remove redundant variables)
    - E.g., Pearson correlation analysis
- Supervised methods
  - Use the target variable (e.g. remove irrelevant variables).
  - Wrapper
    - Search for well-performing subsets of features
  - Filter
    - Select subsets of features based on their relationship with the target
      - Statistical Methods
      - Feature Importance Methods
  - Intrinsic methods
    - Algorithms that perform automatic feature selection during training
      - E.g., Decision Trees

Acknowlegdements and copyright license

- Copyright licence
  - Attribution + Noncommercial + NoDerivatives

- Acknowledgements
  - I would like to thank Dr. Moreno La Quatra, who collaborated to the writing and revision of the teaching content
- Affiliation
  - The author and his staff are currently members of the Database and Data Mining Group at Dipartimento di Automatica e Informatica (Politecnico di Torino) and of the SmartData interdepartmental centre
    - https://dbdmg.polito.it
    - https://smartdata.polito.it

Thank you!