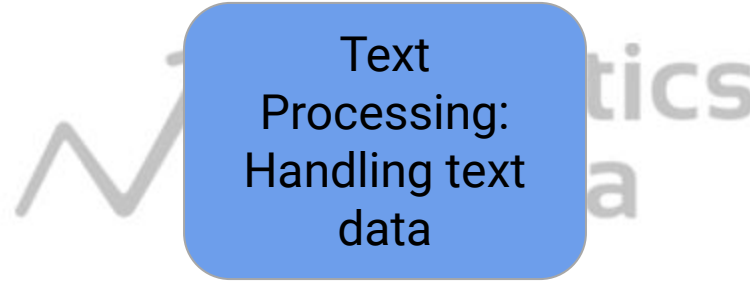


Getting Started with Natural Language Processing (NLP)



Overview of the Course

Overview of the Course



Text
Processing:
Handling text
data

Text Processing: Handling Text Data

Topics to be covered

- Reading Text Data

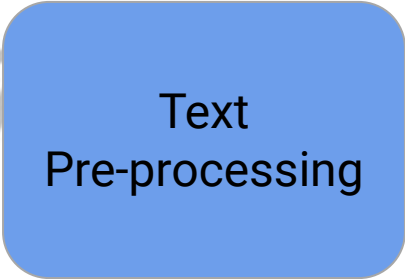


Text Processing: Handling Text Data

Topics to be covered

- Reading Text Data
- Regular Expressions
 - Pattern Matching
 - Implementation in Python

Overview of the Course



Text
Pre-processing

Text Pre-processing

Topics to be covered

- Terminologies



Text Pre-processing

Topics to be covered

- Terminologies
- Word Tokenization, Stopwords removal

Text Pre-processing

Topics to be covered

- Terminologies
- Word Tokenization, Stopwords removal
- Word Normalization (Stemming & Lemmatization)

Text Pre-processing

Topics to be covered


- Terminologies
- Word Tokenization, Stopwords removal
- Word Normalization (Stemming & Lemmatization)
- Text Data Exploration

Text Pre-processing

Project: Exploring the Tweets dataset



Overview of the Course



Information
Extraction

Information Extraction

Topics to be covered

- Applications of Information Extraction in NLP



Information Extraction

Topics to be covered

- Applications of Information Extraction in NLP
- Part-of-Speech (POS) tagging

Information Extraction

Topics to be covered

- Applications of Information Extraction in NLP
- Part-of-Speech (POS) tagging
- Dependency Parsing

Information Extraction

Topics to be covered

- Applications of Information Extraction in NLP
- Part-of-Speech (POS) tagging
- Dependency Parsing
- Named Entity Recognition (NER)

Information Extraction

Topics to be covered

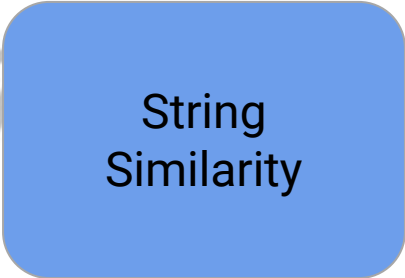
- Applications of Information Extraction in NLP
- Part-of-Speech (POS) tagging
- Dependency Parsing
- Named Entity Recognition (NER)
- Relation Extraction

Information Extraction

Project: Information Extraction from United Nations General Debate dataset



Overview of the Course



String
Similarity

String Similarity

Topics to be covered

- Applications of string similarity in NLP



String Similarity

Topics to be covered

- Applications of string similarity in NLP
- Hamming Distance for string similarity

String Similarity

Topics to be covered

- Applications of string similarity in NLP
- Hamming Distance for string similarity
- Levenshtein Distance for string similarity

Overview of the Course



Information
Retrieval

Information Retrieval

Topics to be covered

- Introduction and applications of information retrieval in NLP



Information Retrieval

Topics to be covered

- Introduction and applications of information retrieval in NLP
- Term - document incidence matrix

Information Retrieval

Topics to be covered

- Introduction and applications of information retrieval in NLP
- Term - document incidence matrix
- Inverted Index

Information Retrieval

Topics to be covered

- Introduction and applications of information retrieval in NLP
- Term - document incidence matrix
- Inverted Index
- Evaluation Metrics for information retrieval (Precision and Recall)

Overview of the Course



Ranked
Retrieval

Ranked Retrieval

Topics to be covered

- Introduction to ranked retrieval



Ranked Retrieval

Topics to be covered

- Introduction to ranked retrieval
- Jaccard Coefficient



Ranked Retrieval

Topics to be covered

- Introduction to ranked retrieval
- Jaccard Coefficient
- Term Frequency (tf) weighting

Ranked Retrieval

Topics to be covered

- Introduction to ranked retrieval
- Jaccard Coefficient
- Term Frequency (tf) weighting
- Inverse Document Frequency (idf) weighting

Ranked Retrieval

Topics to be covered

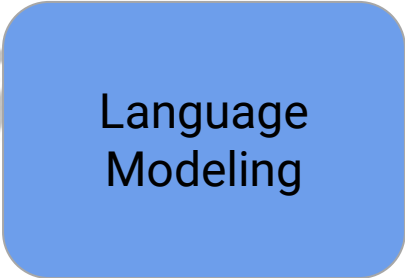
- Introduction to ranked retrieval
- Jaccard Coefficient
- Term Frequency (tf) weighting
- Inverse Document Frequency (idf) weighting
- Term Frequency - Inverse Document Frequency (tf-idf) weighting

Ranked Retrieval

Topics to be covered

- Introduction to ranked retrieval
- Jaccard Coefficient
- Term Frequency (tf) weighting
- Inverse Document Frequency (idf) weighting
- Term Frequency - Inverse Document Frequency (tf-idf) weighting
- Evaluation Metrics for Ranked Retrieval

Overview of the Course



Language
Modeling

Language Modeling

Topics to be covered

- Applications of Language Modeling in NLP



Language Modeling

Topics to be covered

- Applications of Language Modeling in NLP
- Probabilistic Language Models
 - Conditional Probability
 - Markov Assumptions
 - N-gram Language Models

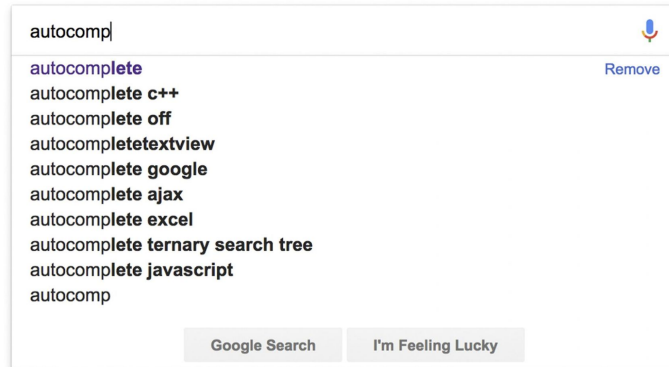
Language Modeling

Topics to be covered


- Applications of Language Modeling in NLP
- Probabilistic Language Models
 - Conditional Probability
 - Markov Assumptions
 - N-gram Language Models
- Evaluation Methods for Language Models (Extrinsic and Intrinsic)

Language Modeling

Project: Next word recommender system



Overview of the Course



Feature
Engineering for
text data

Feature Engineering for text data

Topics to be covered

- Basic Text Features



Feature Engineering for text data

Topics to be covered

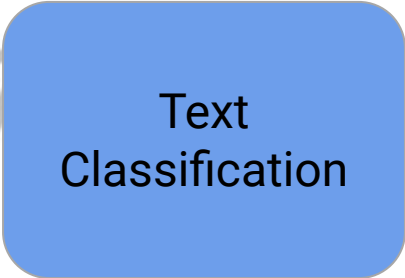
- Basic Text Features
- Text Representation
 - Bag-of-Words (BOW)
 - Tf-idf

Feature Engineering for text data

Topics to be covered

- Basic Text Features
- Text Representation
 - Bag-of-Words (BOW)
 - Tf-idf
- Word Embeddings
 - Word2Vec
 - Glove
 - Fastext

Overview of the Course



Text
Classification

Text Classification

Topics to be covered

- Introduction and applications of text classification



Text Classification

Topics to be covered

- Introduction and applications of text classification
- Supervised approach for text classification

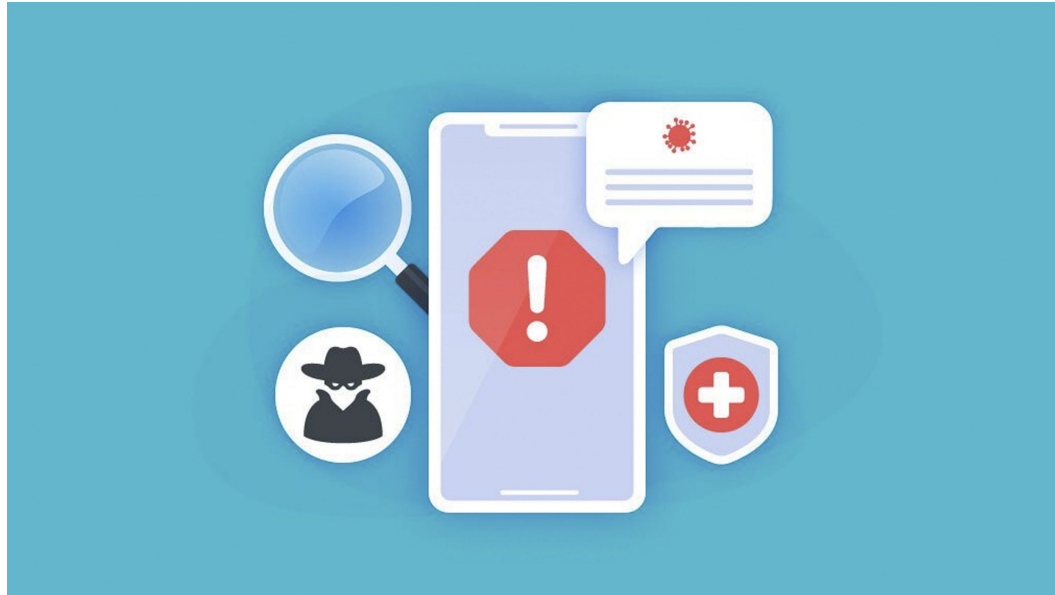
Text Classification

Topics to be covered

- Introduction and applications of text classification
- Supervised approach for text classification
- Evaluation metrics

Text Classification

Project: SMS Spam Classification



Overview of the Course

Unsupervised
NLP -
Topic Modeling

Unsupervised NLP - Topic Modeling

Topics to be covered

- Introduction to topic modeling



Unsupervised NLP - Topic Modeling

Topics to be covered

- Introduction to topic modeling
- LDA



Unsupervised NLP - Topic Modeling

Topics to be covered

- Introduction to topic modeling
- LDA
- LSA

Unsupervised NLP - Topic Modeling

Project: Categorization of sports articles



Overview of the Course

Text
Processing:
Handling text
data

Text
Pre-processing

Information
Extraction

String
Similarity

Information
Retrieval

Ranked
Retrieval

Language
Modeling

Feature
Engineering for
text data

Text
Classification

Unsupervised
NLP -
Topic
Modeling

