Natural Language Processing

I. Extracting the Data

In this module we are going to cover various sources of text data and ways to extract it, which can act as information or insights for businesses.

- Recipe 1. Text data collection using APIs
- Recipe 2. Reading PDF file in Python
- Recipe 3. Reading word document
- Recipe 4. Reading JSON object
- Recipe 5. Reading HTML page and HTML parsing
- Recipe 6. Regular expressions
- Recipe 7. String handling
- Recipe 8. Web scraping

II. Exploring and Processing Text Data

In this module, we are going to cover various methods and techniques to preprocess the text data along with exploratory data analysis.

We are going to discuss the following recipes under text preprocessing and exploratory data analysis.

- Recipe 1. Lowercasing
- Recipe 2. Punctuation removal
- Recipe 3. Stop words removal
- Recipe 4. Text standardization
- Recipe 5. Spelling correction
- Recipe 6. Tokenization
- Recipe 7. Stemming
- Recipe 8. Lemmatization
- Recipe 9. Exploratory data analysis
- Recipe 10. End-to-end processing pipeline

III. Converting Text to Features

In this module, we are going to cover basic to advanced feature engineering (text to features) methods. By the end of this module, you will be comfortable with the following recipes:

- Recipe 1. Lowercasing
- Recipe 2. Punctuation removal
- Recipe 1. One Hot encoding
- Recipe 2. Count vectorizer
- Recipe 3. N-grams
- Recipe 4. Co-occurrence matrix
- Recipe 5. Hash vectorizer
- Recipe 6. Term Frequency-Inverse Document Frequency (TF-IDF)
- Recipe 7. Word embedding
- Recipe 8. Implementing fastText
- Recipe 9. Spacy

IV. Advanced Natural Language Processing

In this module, we are going to cover various advanced NLP techniques and leverage machine learning algorithms to extract information from text data as well as some of the advanced NLP applications with the solution approach and implementation.

- Recipe 1. Lowercasing
- Recipe 2. Punctuation removal
- Recipe 1. Noun Phrase extraction
- Recipe 2. Text similarity
- Recipe 3. Parts of speech tagging
- Recipe 4. Information extraction – NER – Entity recognition
- Recipe 5. Topic modeling
- Recipe 6. Text classification
- Recipe 7. Sentiment analysis
- Recipe 8. Word sense disambiguation
- Recipe 9. Speech recognition and speech to text
- Recipe 10. Text to speech
- Recipe 11. Language detection and translation

V. Implementing Industry Applications

In this module, we are going to implement end-to-end solutions for a few of the Industry applications around NLP.

- Recipe 1. Consumer complaint classification
- Recipe 2. Customer reviews sentiment prediction
- Recipe 3. Text summarization for subject notes
- Recipe 4. Document clustering

VI. Deep Learning for NLP

In this module, we will implement deep learning for NLP:

- Recipe 1. Information retrieval using deep learning
- Recipe 2. Text classification using CNN, RNN, LSTM
- Recipe 3. Predicting the next word/sequence of words using LSTM for Emails