CLARUSWAY

Lesson Plan

Natural Language Processing

Natural Language Processing (NLP) is a branch of artificial intelligence that helps computers understand, interpret and manipulate human language. In NLP, raw text data is converted into numerical data and put into a form that computers/machines can understand and make inferences. In this course, ML, DL, BERT, pre-trained models and word embedding applications of NLP are shown.

Custodian : Johnson-Instructor (johnson_w@clarusway.com)

In-class Sessions : 6 Sessions / 18 hours

Lab Sessions : 2 Labs / 2 hours

Prerequisites

- Basic Math
- Inferential and descriptive statistics
- Pandas, Numpy, Seaborn, Matplotlib, nltk, Tensorflow, Keras, Gensim
- Machine Learning and Deep Learning Basics
- Python Experience

Course Outline

1. Introduction to Natural Language Processing

    1.1. Introduction
    1.2. NLP Applications Worldwide
    1.3. NLP Theory

2. Data Preparation With Natural Language Processing

    2.1. Tokenize the text
    2.2. Remove punctuation
    2.3. Remove stopwords

2.4. Normalize the text
2.5. Vectorize the text
    2.5.1. Count Vectors (CountVectorizer)
    2.5.2. TF-IDF Vectors (TfidfVectorizer)

3. Practice With Natural Language Processing In ML Algorithms

3.1. Naive Bayes
3.2. Logistic Regression
3.3. K Nearest Neighbors
3.4. Support Vector Machines
3.5. Bagging Models
3.6. Boosting Models

4. Word Embedding

4.1 Embedding Layer
4.2. Word2Vec
    4.2.1. Skip gram
    4.2.2. CBOW
4.3. Global Vectors (Glove)
4.4. Bidirectional Encoder Representations from Transformers (BERT)

5. Recurrent Neural Network (RNN) models With NLP

5.1. Simple RNN
5.2. Long-Short Term Memory (LSTM)
5.3. Gated Recurrent Unit (GRU)
5.4. Practice With Natural Language Processing In Deep Learning Algorithms

6. Bidirectional Encoder Representations from Transformers (BERT)

6.1. Introduction of BERT Architecture
6.2. Transformer and Attention
6.3. Working Logic of Transformer
6.4. Preparing the Data for The BERT Model
6.5. Practice of Fine-Tuning for BERT Model
6.6. Practice for Pre-Trained Models on Hugging Face

7. Project With NLP In ML, DL and BERT

Materials & Resources

- Clarusway Learning Management System (LMS)
- Scikit-Learn Documentation
- Keras Documentation
- Tensorflow Documentation
- Pandas Documentation
- Numpy Documentation
- Seaborn Documentation
- Matplotlib Documentation
- NLTK Documentation
- Gensim Documentation
- Article of " Attention Is All You Need"
- Article of "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding"
- Codebasics
- The A.I. Hacker - Michael Phi

Tools and Software

- Zoom, Slack, Kahoot, Peardeck Applications
- Jupyter Notebook / Google Colab
- Scikit-learn, Numpy, Pandas, Matplotlib, Seaborn,Yellowbrick, Tensorflow, Keras, Nltk, Gensim

Assignments & Projects

Projects

1. Project (sentiment analysis with ML, DL and BERT models)