

About the Tutorial

Language is a method of communication with the help of which we can speak, read and write. Natural Language Processing (NLP) is the sub field of computer science especially Artificial Intelligence (AI) that is concerned about enabling computers to understand and process human language. We have various open-source NLP tools but NLTK (Natural Language Toolkit) scores very high when it comes to the ease of use and explanation of the concept. The learning curve of Python is very fast and NLTK is written in Python so NLTK is also having very good learning kit. NLTK has incorporated most of the tasks like tokenization, stemming, Lemmatization, Punctuation, Character Count, and Word count. It is very elegant and easy to work with.

Audience

This tutorial will be useful for graduates, post-graduates, and research students who either have an interest in NLP or have this subject as a part of their curriculum. The reader can be a beginner or an advanced learner.

Prerequisites

The reader must have basic knowledge about artificial intelligence. He/she should also be aware of basic terminologies used in English grammar and Python programming concepts.

Copyright & Disclaimer

© Copyright 2019 by Tutorials Point (I) Pvt. Ltd.

All the content and graphics published in this e-book are the property of Tutorials Point (I) Pvt. Ltd. The user of this e-book is prohibited to reuse, retain, copy, distribute or republish any contents or a part of contents of this e-book in any manner without written consent of the publisher.

We strive to update the contents of our website and tutorials as timely and as precisely as possible, however, the contents may contain inaccuracies or errors. Tutorials Point (I) Pvt. Ltd. provides no guarantee regarding the accuracy, timeliness or completeness of our website or its contents including this tutorial. If you discover any errors on our website or in this tutorial, please notify us at contact@tutorialspoint.com

Table of Contents

About the Tutorial	ii
Audience.....	ii
Prerequisites.....	ii
Copyright & Disclaimer	ii
Table of Contents	iii
1. NLTK — Introduction.....	1
What is Natural Language Processing (NLP)?	1
How does it work?	1
Components of NLP	2
Examples of NLP Applications	3
Implementing NLP	4
Natural Language Tool Kit (NLTK)	5
2. NLTK — Getting Started	6
Installing NLTK	6
Downloading NLTK's Dataset and Packages	8
How to run NLTK script?	9
3. NLTK — Tokenizing Text.....	11
What is Tokenizing?.....	11
NLTK package	11
Tokenizing text into sentences	13
Sentence tokenization using regular expressions	14
4. NLTK — Training Tokenizer & Filtering Stopwords	16
Why to train own sentence tokenizer?	16
What are stopwords?	18
5. NLTK — Looking up words in Wordnet.....	20
What is Wordnet?	20