Inspiratron.org - Natural language processing, machine learning and cybersecurity

Tools

by Nikola Miloševi? - Friday, May 30, 2014

https://inspiratron.org/tools/

Marvin - Semantic text annotator

URL: http://nikolamilosevic86.github.io/Marvin/

Description: Marvin is a semantic text annotation tool that uses various external sources to annotate imputed text. Marvin text annotator can be also used as a java library. Marvin currently supports tagging using Wordnet and DBPedia (linked data version of Wikipedia).

OWASP Seraphimdroid

Web: https://www.owasp.org/index.php/OWASP_SeraphimDroid_Project

GitHub source: https://github.com/nikolamilosevic86/owasp-seraphimdroid

Google Play: https://play.google.com/store/apps/details?id=org.owasp.seraphimdroid

Description: OWASP Seraphimdroid is a privacy and security protection app for android devices. It enables users to protect their devices against malicious software, phishing SMS messages, execution of dangerous USSD codes. OWASP Seraphimdroid also enables users to keep their device safe from theft or loosing with its geofancing feature. OWASP Seraphimdroid has two aims:
- To protect user's privacy and secure the device against malicious features that may cost user money
- To educate user about dangers coming from certain android permissions and setting misconfiguration

Twitter sentiment analyzer

Demo URL: https://inspiratron.org/TwitterSentiment.php

Description: This sentiment analizer analyzes the sentiment of last 100 tweets containing entered search phrase. It will also graphically show how many of the tweets were positive and how many were negative.

Stemmer for Serbian language

Demo URL: https://inspiratron.org/SerbianStemmer.php

Python version: http://nikolamilosevic86.github.io/SerbianStemmer/

Description:In linguistic morphology and information retrieval, **stemming** is the process for reducing inflected (or sometimes derived) words to their stem, base or root form—generally a written word form.

For more information about building this stemmer please read this paper:

Stemmer for Serbian language.pdf

Sentiment analyzer for Serbian language

Demo URL: https://inspiratron.org/SerbianSentiment.php

Description: Sentiment analyzer analyzes sentiment of the sentence on Serbian language. This is example of sentiment analyzer built as my master thesis, classifying sentences into two classes - positive and negative.

Sentiment analysis of sentences on Serbian language (master thesis) - Nikola Miloševi? (on Serbian language) - published 25.1.2013.

Stemmer for English language

Demo URL: https://inspiratron.org/EngStemmer.php

Description: In linguistic morphology and information retrieval, **stemming** is the process for reducing inflected (or sometimes derived) words to their stem, base or root form—generally a written word form. The stem need not be identical to the morphological root of the word; it is usually sufficient that related words map to the same stem, even if this stem is not in itself a valid root. Algorithms for stemming have been studied in computer science since 1968. Many search engines treat words with the same stem as synonyms as a kind of query broadening, a process called **conflation**.

Stemmer algorithm shown in this page is based on Porter Stemmer.

English Sentiment analyzer

Demo URL: https://inspiratron.org/EnglishSentiment.php

Description:Sentiment analyzer for English language is made using machine learning algorithm (Naive Bayes), custom made tokenizer that was doing stemming (using Porter stemming algorithm), negation handling (S. Das and M. Chen negation handling algorithm), punction stripping etc.

Training set of learning algorithm is made out of 1000 positive and 1000 negative movie reviews from IMDB, containing about 700 000 words. Accuracy should be about 85%.

For more information about Sentiment analyzer or API behind it, please contact us.

Eliza - psychotherapist chat-bot

Demo URL: https://inspiratron.org/Eliza/
Description: ELIZA is a computer program and an early example of primitive natural language processing. ELIZA operated by processing users' responses to scripts, the most famous of which was DOCTOR, a simulation of a Rogerian psychotherapist. Using almost no information about human thought or emotion, DOCTOR sometimes provided a startlingly human-like interaction. ELIZA was written at MIT by Joseph Weizenbaum between 1964 and 1966.
