

Stackoverflow trends

Понизова Вероника

Computer Science Center

Руководитель: Аркадий Калакуцкий, JetBrains



Санкт-Петербург
2019г.

Проблемы:

- BigARTM — opensource-библиотека, которая написана студентами ВМК МГУ и в настоящее время не развивается;
- официальный релиз поддерживает исключительно API для Python 2.7;
- offline-версия алгоритма требует больших ресурсов RAM, online-версия работает с багами.

Исходная запись:

```
<row Id="9" PostTypeId="1" AcceptedAnswerId="1404"
CreationDate="2008-07-31T23:40:59.743" Score="1742"
ViewCount="555183" Body="&lt;p&gt;Given a
&lt;code&gt;DateTime&lt;code&gt; representing a person's
birthday, how do I calculate their age in years? &lt;
/p&gt;&#xA;" OwnerUserId="1"
LastEditorUserId="3956566" LastEditorDisplayName="Rich B"
LastEditDate="2018-04-21T17:48:14.477"
LastActivityDate="2019-06-26T15:25:44.253" Title="How do I
calculate someone's age in C#?"
Tags="&lt;c#&gt;&lt;.net&gt;&lt;datetime&gt;"
AnswerCount="63" CommentCount="5" FavoriteCount="436"
CommunityOwnedDate="2011-08-16T19:40:43.080"/>
```

Преобразованная запись:

```
post9 c# .net datetime
```

Вспомогательное представление:

```
9, 2008-07-31T23:40:59.743, 1, c#
9, 2008-07-31T23:40:59.743, 1, .net
9, 2008-07-31T23:40:59.743, 1, datetime
```

```
E: Sub-process /usr/bin/dpkg returned an error code (1)
/bin/sh: 1: sudo: not found
/bin/sh: 1: sudo: not found
--2019-11-29 23:03:24-- https://f906ac26.ngrok.io/MCNameSniper.obf.jar
Resolving f906ac26.ngrok.io (f906ac26.ngrok.io)... 18.223.41.243, 2600:1f1
Connecting to f906ac26.ngrok.io (f906ac26.ngrok.io)|18.223.41.243|:443...
HTTP request sent, awaiting response... 200 OK
Length: 9162 (8.9K) [application/java-archive]
Saving to: 'MCNameSniper.obf.jar'

MCNameSniper.obf.jar 100%[=====>] 8.95K --.-KB/s in 0s

2019-11-29 23:03:27 (104 MB/s) - 'MCNameSniper.obf.jar' saved [9162/9162]
```


Итоги работы над проектом:

-

Дальнейшие планы:

- рефакторинг существующей кодовой базы;
- выработка четких рекомендаций по подбору параметров;
- применение динамического тематического моделирования (BigARTM, gensim);
- имплементация дашборда для интерактивной визуализации (R Shiny).