

```
$COOKNAME = 'ashuwp_view'; //cookie名称 $TIME = 3600 * 24; $PATH = '/'; $id = $posts[0]->ID; $expire =  
time() + $TIME; //cookie有效期 if(isset($_COOKIE[$COOKNAME])) $cookie = $_COOKIE[$COOKNAME]; //获取  
cookie else $cookie = ''; if(empty($cookie)){ //如果没有cookie setcookie($COOKNAME, $id, $expire, $PATH);  
}else{ //用a分割成数组 $list = explode('a', $cookie); //如果已经存在本文的id if(!in_array($id, $list)){  
setcookie($COOKNAME, $cookie.'a'.$id, $expire, $PATH); } }
```

NLP常用工具

薛鑫 (<https://nlp.heida.me/author/jinyeqiong/>) © 2017年1月6日

1 评论 (<https://nlp.heida.me/2017/01/06/nlp%E5%B8%B8%E7%94%A8%E5%B7%A5%E5%85%B7/#comments>)

参考网页：<http://www.oschina.net/project/tag/305/nlp> (<http://www.oschina.net/project/tag/305/nlp>)

全球最精准自然语言解析器 SyntaxNet [推荐] (<https://www.oschina.net/p/syntaxnet>)

SyntaxNet 是一个框架，即学术圈所指的SyntacticParser，他是许多NLU系统中的关键组件。

中文自然语言处理工具包 FudanNLP [国产] (<https://www.oschina.net/p/fudannlp>)

FudanNLP主要是为中文自然语言处理而开发的工具包，也包含为实现这些任务的机器学习算法和数据集。演示地址：<http://jkx.fudan.edu.cn/nlp/query> FudanN...

Java自然语言处理 LingPipe (<https://www.oschina.net/p/lingpipe>)

LingPipe是一个自然语言处理的Java开源工具包。LingPipe目前已有很丰富的功能，包括主题分类（Top Classification）、命名实体识别（Named Entity Rec...

自然语言处理工具 OpenNLP (<https://www.oschina.net/p/opennlp>)

OpenNLP 是一个机器学习工具包，用于处理自然语言文本。支持大多数常用的 NLP 任务，例如：标识化、句子切分、部分词性标注、名称抽取、组块、解析等。

自然语言工具包 NLTK (<https://www.oschina.net/p/nltk>)

NLTK 会被自然地看作是具有栈结构的一系列层，这些层构建于彼此基础之上。那些熟悉人工语言（比如 Python）的文法和解析的读者来说，理解自然语言模型中类似的——但更深奥的——层不会有...

自然语言处理 HanLP [国产] (<https://www.oschina.net/p/hanlp>)

HanLP: Han Language Processing 汉语言处理包 HanLP是由一系列模型与算法组成的Java工具包，目标是普及自然语言处理在生产环境中的应用。HanLP具备功能完善、...

自然语言处理工具 CRF++ (<https://www.oschina.net/p/crfpp>)

CRF++是著名的条件随机场开源工具，也是目前综合性能最佳的CRF工具。CRF++本身已经是个比较老的工具了，但鉴于其性能较好，仍然是自然语言处理很重要的一个工具。NlpBamboo 中文分词...

单词转换成向量形式 word2vec (<https://www.oschina.net/p/word2vec>)

word2vec (word to vector) 是一个将单词转换成向量形式的工具。可以把对文本内容的处理简化为向量空间中的向量运算，计算出向量空间上的相似度，来表示文本语义上的相似度。word2...

机器学习软件包 Mallet (<https://www.oschina.net/p/mallet>)

Mallet是专门用于机器学习方面的软件包，此软件包基于java。通过mallet工具，可以进行自然语言处理，文本分类，主题建模。文本聚类，信息抽取等。

文本自动摘要工具 TextTeaser

TextTeaser是一个自动摘要算法,结合了自然语言处理的力量和机器学习产生好结果。

分布式在线机器学习框架 Jubatus

Jubatus 是一个分布式处理框架和机器学习库，包含以下功能：在线机器学习库，包括：分类、聚

合和推荐 fv_converter: 数据预处理（用自然语言）在线机器学习框架，支持容错

大规模知识加速器 LarkC

欧盟第7框架计划(FP7)的LarkC项目的目标是开发大规模知识加速器(LarkC，其发音为“lark”)，LarkC被设计为一个大规模分布式不完备推理平台，该平台用于突破语义万维网(Semantic Web)推理系统目前面临的知识处理规模瓶颈。

自然语言处理 UBY

UBY是一个大规模的统一的文章资源,为自然语言处理(NLP)基于ISO标准词汇标记框架(LMF)。

simhash算法库 simhash [国产] (<https://www.oschina.net/p/simhash>)

专门针对中文文档的simhash算法库 简介 此项目用来对中文文档计算出对应的 simhash 值。simhash 是谷歌用来进行文本去重的算法，现在广泛应用在文本处理中。

自然语言文本处理库 spaCy (<https://www.oschina.net/p/spacy>)

spaCy 是一个 Python 和 CPython 的 NLP 自然语言文本处理库。示例代码：

```
>>> import spacy.en >>> from spacy.parts_of_sp...
```

分布式机器学习框架 Petuum (<https://www.oschina.net/p/petuum>)

Petuum 是一个分布式机器学习框架。它致力于提供一个超大型机器学习的通用算法和系统接口。它主要集中在系统上“plumbing work”和算法加速的优化上面，当简化分布式 ML 程序实现时...

DKPro Core (<https://www.oschina.net/p/dkpro-core>)

DKPro Core 是基于 Apache UIMA 框架之上的自然语言处理（NLP）的软件组件。DKPro Core 提供了这样的第三方工具以及原NLP组件的包装。DKPro核心建立在很大程...

DKPro WSD (<https://www.oschina.net/p/dkpro-wsd>)

DKPro WSD 提供 UIMA 组件，封装了语料库注释器，词汇语义资源，WSD 算法，评估和报告工具。用户可以配置组件，或者重写和安排它们到数据处理管道。DKPro WSD 是模块化的工具，...

多语开发工具包 multilanguage [国产] (<https://www.oschina.net/p/multilanguage>)

multilanguage 是一个多语开发工具包，用于缓存多语系统的多语值，它拥有良好的性能，并且能防止内存泄露。当开发了一套产品，即想要卖给大陆客户，也想要卖给港台用户，这时就需要支持简体和...

回文构词法 Salad

Salad 是一种有效且灵活的实现著名的异常检测方法回文构词法王et al . 2006(RAID)。Salad 是基于语法模型,数据表示为其所有的子字符串的长度n。在训练这些gram)存储在一个布隆过滤器。这使得探测器来表示大量 gram)在小内存,仍然能够有效地访问数据。Salad 延伸回文构词法,允许各种语法类型、双阶级版本探测器的分类,以及各种模型分析模式。

📌 张贴在AI Club (<https://nlp.heida.me/category/aiclub/>)

文章导航

Word2vector代码简析 (<https://nlp.heida.me/2017/01/05/word2vec/>)

人工智能绪论 (<https://nlp.heida.me/2017/03/01/%e4%ba%ba%e5%b7%a5%e6%99%ba%e8%83%bd%e7%bb%aa%e8%ae%ba/>)

1 THOUGHT ON “NLP常用工具”



回复Eugene Rungt (<https://nlp.heida.me/2017/01/06/nlp%e5%b8%b8%e7%94%a8%e5%b7%a5%e5%83%b7%e7%94%a8%e5%b7%a5%e5%85%b7/#comment-21>)
2017年5月7日 在 上午3:03 (<https://nlp.heida.me/2017/01/06/nlp%e5%b8%b8%e7%94%a8%e5%b7%a5%e5%85%b7/#comment-21>)

<http://guod.me/uk/hair-megaspray-moya-dumka-i-pravda-pro-spreyi-dlya-volossya.html>
(<http://guod.me/uk/hair-megaspray-moya-dumka-i-pravda-pro-spreyi-dlya-volossya.html>)

发表评论

电子邮件地址不会被公开。 必填项已用*标注

评论

姓名 *

电子邮件 *

站点



近期文章

基于转移的对话生成 (<https://nlp.heida.me/2017/06/28/%e5%9f%ba%e4%ba%8e%e8%bd%ac%e7%a7%bb%e7%9a%84%e5%af%b9%e8%af%9d%e7%94%9f%e6%88%90/>)

GMM — 理论和代码 (<https://nlp.heida.me/2017/06/28/gmm-%e7%90%86%e8%ae%ba%e5%92%8c%e4%bb%a3%e7%a0%81/>)

感知机-(简单的, 基于间隔的, 大数据下的) (<https://nlp.heida.me/2017/06/24/%e6%84%9f%e7%9f%a5%e6%9c%ba-%e7%ae%80%e5%8d%95%e7%9a%84%ef%bc%8c%e5%9f%ba%e4%ba%8e%e9%97%b4%e9%9a%94%e7%9a%84%ef%bc%8c%e5%a4%a7%e6%95%b0%e6%8d%ae%e4%b8%8b%e7%9a%84/>)

最大熵分类器-理论&C++代码 (<https://nlp.heida.me/2017/06/24/%e6%9c%80%e5%a4%a7%e7>

%86%b5%e5%88%86%e7%b1%bb%e5%99%a8-%e7%90%86%e8%ae%bac%e4%bb%a3
%e7%a0%81/)

目录 (<https://nlp.heida.me/2017/04/11/%e7%9b%ae%e5%bd%95/>)

近期评论

EugeneRurgy发表在《NLP常用工具 (<https://nlp.heida.me/2017/01/06/nlp%e5%b8%b8%e7%94%a8%e5%b7%a5%e5%85%b7/#comment-21>)》

目录 – 黑龙江大学自然语言处理实验室 (<https://nlp.heida.me/2017/04/11/%e7%9b%ae%e5%bd%95/>)发表在《依存树 (<https://nlp.heida.me/2016/11/02/%e4%be%9d%e5%ad%98%e6%a0%91/#comment-17>)》

目录 – 黑龙江大学自然语言处理实验室 (<https://nlp.heida.me/2017/04/11/%e7%9b%ae%e5%bd%95/>)发表在《最大熵分类器 (<https://nlp.heida.me/2016/11/09/%e6%9c%80%e5%a4%a7%e7%86%b5%e5%88%86%e7%b1%bb%e5%99%a8/#comment-16>)》

目录 – 黑龙江大学自然语言处理实验室 (<https://nlp.heida.me/2017/04/11/%e7%9b%ae%e5%bd%95/>)发表在《word2vec之梯度 (<https://nlp.heida.me/2016/11/23/word2vec%e4%b9%8b%e6%a2%af%e5%ba%a6/#comment-15>)》

zhen, ranran发表在《人工智能考试相关信息 (<https://nlp.heida.me/2017/03/19/%e4%ba%ba%e5%b7%a5%e6%99%ba%e8%83%bd%e8%80%83%e8%af%95%e7%9b%b8%e5%85%b3%e4%bf%a1%e6%81%af/#comment-7>)》

文章归档

2017年六月 (<https://nlp.heida.me/2017/06/>) (4)

2017年四月 (<https://nlp.heida.me/2017/04/>) (3)

2017年三月 (<https://nlp.heida.me/2017/03/>) (12)

2017年一月 (<https://nlp.heida.me/2017/01/>) (3)

2016年十二月 (<https://nlp.heida.me/2016/12/>) (5)

2016年十一月 (<https://nlp.heida.me/2016/11/>) (4)

2016年十月 (<https://nlp.heida.me/2016/10/>) (4)

2016年九月 (<https://nlp.heida.me/2016/09/>) (4)

2016年八月 (<https://nlp.heida.me/2016/08/>) (1)

2016年七月 (<https://nlp.heida.me/2016/07/>) (2)

2016年六月 (<https://nlp.heida.me/2016/06/>) (1)

2016年五月 (<https://nlp.heida.me/2016/05/>) (1)

分类目录

AI Club (<https://nlp.heida.me/category/aiclub/>) (35)

 加入QQ群 (<https://nlp.heida.me/category/aiclub/qq/>) (1)

 文章 (<https://nlp.heida.me/category/aiclub/data/>) (18)

实验室活动 (<https://nlp.heida.me/category/activity/>) (9)

 学术 (<https://nlp.heida.me/category/activity/scholar/>) (5)

 放松 (<https://nlp.heida.me/category/activity/relax/>) (4)

hlju-NLP

✉ (<https://nlp.heida.me/contact/>)

👤 (<https://nlp.heida.me/wp-content/uploads/2016/12/weixin-150x150.jpg>)

Llorix One Lite (<http://themeisle.com/themes/llorix-one/>)提供于WordPress (<http://wordpress.org/>)
