

## 论文相似性检测报告（详细版）

报告编号: 7936d12b-7ebb-4669-937d-a29000b813e0

原文字数: 47,926

检测日期: 2013年12月10日

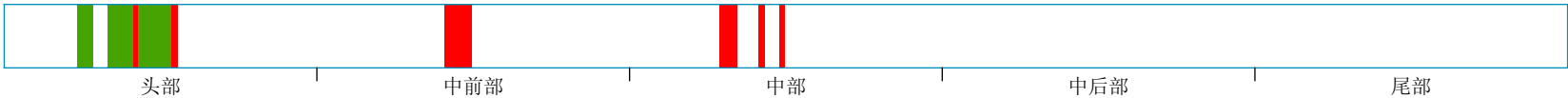
检测范围: 中国学术期刊数据库（CSPD）、中国学位论文全文数据库（CDDb）、中国学术会议论文数据库（CCPD）、中国学术网页数据库（CSWD）

检测结果:

### 一、总体结论

总相似比: 12.75% (参考文献相似比: 7.54%, 排除参考文献相似比: 5.22%)

### 二、相似片段分布



注: 绿色区域为参考文献相似部分, 红色区域为其它论文相似部分。

### 三、相似论文作者（举例10个）

[点击查看全部举例相似论文作者](#)

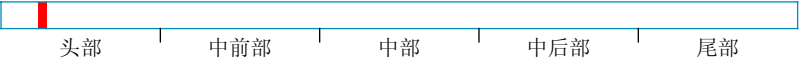
### 四、典型相似论文（举例128篇）

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
1	5.80%	Web文本分类技术研究现状述评	参考文献	期刊论文	高淑琴	图书情报知识	2008
2	1.45%	中文短文本分类的相关技术研究	参考文献	学位论文	崔争艳	河南大学	2011
3	4.35%	Web自动文本分类技术研究综述		期刊论文	蒲筱哥	情报学报	2009
4	4.06%	自动文本分类方法研究述评		期刊论文	蒲筱哥	情报科学	2008

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
5	2.32%	<u>分类方法的新发展:研究综述</u>		期刊论文	张丽娟 等	计算机科学	2006
6	2.03%	<u>中文文本分类关键技术研究</u>		学位论文	何钟莉	西安电子科技大学	2009
7	1.74%	<u>新闻搜索分析平台的相关技术研究</u>		学位论文	李丹东	哈尔滨工程大学	2009
8	1.45%	<u>网络环境下文本自动分类方法研究综述</u>		期刊论文	吴波 等	鸡西大学学报	2009
9	1.45%	<u>基于领域词语本体的短文本分类研究</u>		学位论文	宁亚辉	重庆邮电大学	2009
10	1.45%	<u>文本分类技术进展</u>		期刊论文	胡文静 等	知识经济	2011
11	1.16%	<u>网页与文本自动分类综述</u>		期刊论文	王明文 等	南昌工程学院学报	2005
12	1.16%	<u>中文Web文本自动分类的研究与实现</u>		学位论文	刘磊	长春理工大学	2007
13	1.16%	<u>基于RBF网络的文本自动分类的研究</u>		学位论文	王晓庆	江西师范大学	2003
14	0.87%	<u>基于语义的网络知识获取相关技术研究</u>		学位论文	郭勇	国防科学技术大学	2007
15	0.87%	<u>文本分类综述</u>		期刊论文	靳小波 等	自动化博览	2006

[点击查看全部举例相似论文](#)

## 五、相似论文片段（共14个）

1	送检论文片段	相似论文片段 <span>【1.45%】</span>
	<p>位置：</p> <p>截至2011年, 互联网已经走过了40多年历程, 相对人类上下五千年文明史来讲, 40年是极为短暂的, 但仅仅是这40多年, 互联网为整个世界带来了翻天覆地的变化。正因为互联网的普及, 人们的生活才变得更加便利, 更加丰富多彩, 人与人、与社会、与世界之间的距离也越</p>	<p>来源: <a href="#">[参考文献]</a>中文短文本分类的相关技术研究</p> <p>[学位论文]崔争艳, 2011年 河南大学</p> <p>历程, 相对人类上下五千年文明史来讲, 40年是极为短暂的, 但仅仅是这40多年, 互联网为整个世界带来了翻天覆地的变化。正因为互联网的普及, 人们的生活才变得更加便利, 更加丰富多彩, 人与人、与社会、与世界之间的距离也越拉越近。据2011年1月19日中国互联网络信息中</p>


3


	<p>了样本矢量的数据矩阵，那么随后的文本分类过程和英文文本分类相同，也就是随后的文本分类过程独立于语种。因此，当前的中文文本分类主要集中在如何利用中文本身的一些特征来更好地表示文本样本。国内对于文本自动分类的研究起步较晚，但从简单的查词典的方法，到后来的基于统计语言模型的分词方法，中文分词的技术已趋于成熟</p>	<p>是随后的文本分类过程独立于语种。因此，当前的中文[作者简介]高淑琴，女，1968年生。馆员，从事图书馆数字信息资源建设研究。81万方数据 2008年第3期 闰名. 钰祖. 加诹文本分类主要集中在如何利用中文本身的一些特征来更好地表示文本样本。国内对于文本自动分类的研究起步较晚，但从简单的查词典的方法，到后来的基于统计语言模型的分词方法，中文分词的技术已趋于成熟</p>
3	<p><b>送检论文片段</b></p> <p>位置:</p>  <p>近年来，文本分类已成为众多领域研究者的热门研究课题，研究者们从不同的角度把越来越多的知识引入文本分类领域，推动着文本分类的不断发展，产生了许多新的方法。</p> <p>(1) 基于群的分类方法。这种方法可以看作是进化计算的一个新的分支，它模拟了生物界中蚁群、鱼群和鸟群在觅食或者逃避敌人时的行为。纵观文献中对基于群的分类方法的研究。对ACO或者PSO在数据挖掘中应用的研究</p>	<p><b>相似论文片段</b> <b>【0.58%】</b></p> <p>来源: [参考文献]Web文本分类技术研究现状述评</p> <p>[期刊论文]《图书情报知识》，2008年 高淑琴</p> <p>近年来，文本分类已成为众多领域研究者的热门研究课题，研究者们从不同的角度把越来越多的知识引入文本分类领域，推动着文本分类的不断发展，产生了许多新的方法。(1) 基于群的分类方法 (swarm-based approaches)。这种方法可以看作是进化计算的一个新的分支，它模拟了生物界中蚁群、鱼群和鸟群在觅食或者逃避敌人时的行为。纵观文献中对基于群的分类方法的研究，我们将这种方法分为两类：一类是蚁群算法或者</p>
4	<p><b>送检论文片段</b></p> <p>位置:</p>  <p>方法用到实际的大规模数据挖掘中还需要做大量的研究工作。</p> <p>(2) 基于模糊粗糙集的文本分类模型。文本分类过程中由于同义词、多义词、近义词的存在导致许多类并不能完全划分开来，造成类之间的边界模糊。粗糙集理论有机的结合了模糊集理论与粗糙集理论在处理不确定信息方面的能力。它们处理的是两种不同类别的模糊和不确定性，将两者结合</p>	<p><b>相似论文片段</b> <b>【0.58%】</b></p> <p>来源: 文本分类语料库自动创建系统的研究与实现</p> <p>[学位论文]吴韦，2009年 武汉理工大学</p> <p>方法可以看成是进化计算的一个新的分支，它模拟自然界中的生物种群的行为，目前有两种：蚁群优化算法 (ACO) 和PSO (Particle Swarm optimisers)! 矧。基于模糊. 粗糙集的文本分类模型，文本分类过程中由于同义词、多义词、近义词的存在导致许多类并不能完全划分开来，造成类之间的边界模糊。模糊. 粗糙集理论有机的结合了模糊集理论与粗糙理论处理不正确信息和不完全知识的能力</p>
5	<p><b>送检论文片段</b></p> <p>位置:</p> 	<p><b>相似论文片段</b> <b>【1.16%】</b></p> <p>来源: [参考文献]Web文本分类技术研究现状述评</p> <p>[期刊论文]《图书情报知识》，2008年 高淑琴</p>

<p>完全划分开来,造成类之间的边界模糊。粗糙集理论有机的结合了模糊集理论与粗糙集理论在处理不确定信息方面的能力。它们处理的是两种不同类别的模糊和不确定性,将两者结合起来的模糊粗糙集理论能更好地处理不完全知识。</p> <p>(3) 多分类器融合的方法。实际应用的复杂性和数据的多样性往往使得单一的分类方法不够有效。因此学者们对多种分类方法的融合(fusion)进行了广泛的研究,取得了一系列研究成果。纵观文献中的研究,可以大致将多分类器的融合技术分为以下几类:投票机制(Voting)、行为知识空间方法(Behavior Knowledge Space BKS)、证据理论(Dempster shafer theory)、贝叶斯方法和遗传编程(Genetic programming GP)。</p> <p>(4) 基于 RBF 网络的文本分类模型。把监督方法和非监督方法相结合,通过两层映射关系对文本进行分类,首先利用非监督聚类方法根据文本本身的相似性聚出若干个簇,使得</p>		<p>它们处理的是两种不同类别的模糊和不确定性。将两者结合起来的模糊一粗糙集理论能更好地处理不完全知识。(3) 多分类器融合(fusion)的方法。实际应用的复杂性和数据的多样性往往使得单一的分类方法不够有效。因此学者们对多种分类方法的融合(fu—sion)进行了广泛的研究,取得了一系列研究成果。纵观文献中的研究,可以大致将多分类器的融合技术分为以下几类:投票机制(voting)、行为知识空间方法(Behavior—KnowIedge Space BKS)、证据理论(Dempster—Shafer theory)、贝叶斯方法和遗传编程(genetic programming GP)。采用投票机制的方法主要有装袋(bagging[ ” ])和推进(boostin— ” ] )。Buhlmann P.和Yu B[151对bagging进行了深入的分析;Hothorn T.and Lausen B.在文献[16—17]中将bagging用于决策树,并在文献[18]中对bagging进行了发展,采用boot strap aggregation来融合分类器</p>	
6	<p>送检论文片段</p> <p>位置:</p> <div><div></div><div>头部</div><div>中前部</div><div>中部</div><div>中后部</div><div>尾部</div></div> <p>网络的文本分类模型。把监督方法和非监督方法相结合,通过两层映射关系对文本进行分类,首先利用非监督聚类方法根据文本本身的相似性聚出若干个簇,使得每个簇内部的相似性尽可能高而簇之间的相似性尽可能低,并由此产生第一层映射关系即文本到簇的映射,然后通过监督学习方法构造出第二层映射关系,即簇集到目标类集合的映射,然后为每一个簇定义一个相应的径向基函数(Radial Basis Function, RBF),并确定这些基函数的中心和宽度,利用这些径向基函数的线形组合来拟合训练文本,利用矩阵运算得到线性组合中的权值,在计算权值时,为了避免产生过度拟合的现象,采用了岭回归技术,即在代价函数中加入包含适当正规化参数的权值惩罚项,以保证网络输出函数具有一定平滑度。</p> <p>(5) 潜在语义分类模型。潜在语义索引方法,已经被证明是对传统的向量空间技术的一种改良,可以达到消除词之间的相关性,化简文档向量的目的</p>	相似论文片段	【1.45%】
<p>来源: [参考文献]Web文本分类技术研究现状述评</p> <p>[期刊论文]《图书情报知识》,2008年 高淑琴</p> <p>监督方法和非监督方法相结合,通过两层映射关系对文本进行分类,首先利用非监督聚类方法根据文本本身的相似性聚出若干个簇,使得每个簇内部的相似性尽可能高而簇之间的相似性尽可能低,并由此产生第一层映射关系,即文本到簇的映射,然后通过监督学习方法构造出第二层映射关系,即簇集到目标类集合的映射[25。。然后为每一个簇定义一个相应的径向基函数(Ra—dial Basis Function, RBF),并确定这些基函数的中心和宽度,利用这些径向基函数的线形组合来拟合训练文本,利用矩阵运算得到线性组合中的权值,在计算权值时,为了避免产生过度拟合的现象,采用了岭回归技术,即在代价函数中加入包含适当正规化参数的权值惩罚项,以保证网络输出函数具有一定平滑度。(5) 潜在语义分类模型。潜在语义索引方法,已经被证明是对传统的向量空间技术的一种改良,可以达到消除词之间的相关性,化简文档向量的目的</p>			
7	<p>送检论文片段</p> <p>位置:</p> <div><div></div></div>	相似论文片段	【0.87%】
<p>来源: [参考文献]Web文本分类技术研究现状述评</p>			



	<div> <div> <div>头部</div> <div>中前部</div> <div>中部</div> <div>中后部</div> <div>尾部</div> </div> <p>缺陷：第一，由于要存储所有的训练实例，所以对大规模数据集进行分类是低效的；第二，KNN分类的效果在很大程度上依赖于k值选择的好坏。实验证明，基于KNN模型的方法在分类精确度上与C5.0和标准的KNN相当。另外，针对KNN方法的第一个缺陷，Nong Ye and Xiangyang Li将聚类方法和经典的KNN方法结合起来，提出了一种新颖的分类方法，称为CCA2S。CCA2S能够处理大规模数据集，可伸缩性好，并且支持增量式学习。但CCA2S只能处理连续属性，而且只针对类别为两类的分类问题。如何扩展 CCA2S，以使其能够处理多类别的问题，还有待进一步研究。</p> <p>(7) 支持向量机(SVM)方法</p> </div>	<div> <p>[期刊论文]《图书情报知识》，2008年 高淑琴</p> <p>提高了分类速度和精确度。实验证明，基于KNN模型的方法在分类精确度上与C5.0和标准的K-NN相当。另外，针对K-NN方法的第一个缺陷，Nong Ye and xiangyang Li将聚类方法和经典的K-NN方法结合起来，提出了一种新颖的分类方法，称为CCA—S[3引]。CCA—S能够处理大规模数据集，可伸缩性好，并且支持增量式学习。但CCA—S只能处理连续属性，而且只针对类别为两类的分类问题。如何扩展CCA. S，以使其能够处理多类别的问题</p> </div>
8	<div> <p>送检论文片段</p> <p>位置：</p> <div> <div>头部</div> <div>中前部</div> <div>中部</div> <div>中后部</div> <div>尾部</div> </div> <p>处理连续属性，而且只针对类别为两类的分类问题。如何扩展 CCA2S，以使其能够处理多类别的问题，还有待进一步研究。</p> <p>(7) 支持向量机(SVM)方法的新发展：SVM是进行分类、聚类和时间序列分析的有效数据挖掘工具。但是，由于 SVM 的训练时间会随着数据集的增大而增加，所以在处理大规模数据集时，SVM往往需要较长的训练时间。而实际的数据挖掘应用</p> </div>	<div> <p>相似论文片段 <b>【0.58%】</b></p> <p>来源：Web自动文本分类技术研究综述</p> <p>[期刊论文]《情报学报》，2009年 蒲筱哥</p> <p>支持向量机(SVM)方法的新发展SVM 是进行分类、聚类和时间序列分析的有效数据挖掘工具。但是，由于 SVM 的训练时间会随着数据集的增大而增加，所以在处理大规模数据集时，SVM往往需要较长的训练时间。而实际的数据挖掘应用往往包含了数以百万计的数据，这使得 SVM 很难发挥作用。针对这个问题，文献[46, 47]用选择性采样或者主动学习方法来训练</p> </div>
9	<div> <p>送检论文片段</p> <p>位置：</p> <div> <div>头部</div> <div>中前部</div> <div>中部</div> <div>中后部</div> <div>尾部</div> </div> <p>进行分类、聚类和时间序列分析的有效数据挖掘工具。但是，由于 SVM 的训练时间会随着数据集的增大而增加，所以在处理大规模数据集时，SVM往往需要较长的训练时间。而实际的数据挖掘应用往往包含了数以百万计的数据，这使得 SVM 很难发挥作用[12]。</p> <p>1.2.2中文分词算法</p> </div>	<div> <p>相似论文片段 <b>【0.58%】</b></p> <p>来源：自动文本分类方法研究述评</p> <p>[期刊论文]《情报科学》，2008年 蒲筱哥</p> <p>进行分类、聚类和时间序列分析的有效数据挖掘工具。但是，由于S、M的训练时间会随着数据集的增大而增加，所以在处理大规模数据集时，svM往往需要较长的训练时间。而实际的数据挖掘应用往往包含了数以百万计的数据，这使得SVI很很难发挥作用</p> </div>

	在国外，已经有很多诸如Akimist的厂商在做Spam Detection方面的产品。然而在中国，由于中文的特殊性，需要	
10	<p><b>送检论文片段</b></p> <p>位置:</p>  <p>过程如下所示:</p> <p>1) 请求签名</p> <p>所有的OAuth请求使用同样的算法来生成(signature base string)签名字符基串和签名。</p> <p>base string是把http方法名, 请求URL以及请求参数用&amp;字符连起来后做URL Encode编码。具体来讲, base string由http方法名, 之后是&amp;, 接着是过url编码(url-encoded)之后的url和访问路径及&amp;。接下来, 把所有的请求参数包括POST方法体中的参数, 经过排序(按参数名进行文本排序, 如果参数名有重复则再按参数值进行重复项目排序), 使用%3D替代=号, 并且使用%26作为每个参数之间的分隔符, 拼接成一个字符串。</p> <p>2) 获取Request Token</p> <p>获取request token是进行用户认证的第一步。这一步主要有两个目的:</p> <p>第一, 告诉新浪微博你将要做什么</p> <p>第二, 告诉新浪微博你在callback里要做什么</p> <p>3) 用户认证</p> <p>这一步主要是发送你获取的oauth_token, 并且获得用户的授权。一般来说, WEB应用会简单的重定向到相应的页面, 桌面应用程序会给出URL并要求用户自行验证。</p> <p>新浪微博开放平台的验证URL是http://api.t.sina.com.cn/oauth/authorize</p> <p>要求必须以oauth作为参数, 一般来说请求格式如下:</p> <p>http://api.t.sina.com.cn/oauth/authorize?oauth_token=81XF0ZH5tAwj6vzJYuLQp10WUEYtWc</p> <p>如果用户没有登录新浪微博, 则会要求用户登录。否则将会出现一个页面, 用户可以在此页</p>	<p><b>相似论文片段</b> <b>【1.74%】</b></p>

	<p>面上一键同意或者拒绝对此应用授权。用户授权后，web应用页面将会重定向至你指定的 <code>oauth_callback</code>，如果是桌面应用，将会显示PIN码，用户需要将PIN码输入你的应用中来完成授权过程。</p> <p>4) 获取 <code>access_token</code></p> <p>新浪微博开放平台 <code>access token</code> 请求</p>	
11	<p><b>送检论文片段</b></p> <p>位置:</p>  <p>一个节点上做决策时，会使用到多少个变量。<math>(m &lt; M)</math></p> <p>3. 从N个训练样例中以可重复取样的方式，取样N次，形成一组训练集（本质上是bootstrap取样）。并使用这棵树来对剩余预测其类别，并评估其误差。</p> <p>4. 对于每一个节点，随机选择m个基于此点上的变量。根据这 m 个变量，计算其最佳的分割方式。</p> <p>5. 每棵树都会完整的成长而不会对其进行剪枝（Pruning）（这有可能在建完一棵正常树状分类器后会被采用）。</p> <p>4.2.3 Random Forest优点</p> <ol style="list-style-type: none"> <li>1. 在资料种类非常多的情况下，它可以产生高准确度的分类器。</li> <li>2. Random Forest可以处理具有大量规模输入的变量。</li> <li>3. Random Forest可以在预测类别的同时评估变量在做决策时的重要性。</li> <li>4. 在建造森林时，它可以在内部对于一般化后的误差产生不偏差的估计。</li> <li>5. Random Forest包含一个好方法可以估计弥补丢失的训练样例，而且在资料大部分缺失的情况下，分类其本身仍可以维持一定的准确度。</li> <li>6. Random Forest提供了一个实验方法，可以去侦测变量之间的相互影响。</li> <li>7. 对于不平衡的分类训练集来说，它可以平衡误差。</li> <li>8. Random Forest会计算各个样例数据中的亲近度，对于数据挖掘、侦测偏离者（outlier）</li> </ol>	<p><b>相似论文片段</b> <b>【1.45%】</b></p>



	和处理	
12	<p><b>送检论文片段</b></p> <p>位置:</p>  <p>实现的, 它根据每次训练集之中每个样本的分类是否产生了正确的结果, 以及上层分类器的总体分类的准确率, 来确定每个样本数据的权值。将修改过权值的新训练集送给下层分类器进行训练, 最后将每次训练得到的分类器最后融合起来, 作为最终可以使用的决策分类器。使用Ada boost构建的分类器可以</p>	<p><b>相似论文片段</b> <b>【0.58%】</b></p> <p>来源: <u>一种多视角人脸检测方法</u>              [期刊论文]《计算机工程与科学》, 2011年 马颖哲 等</p> <p>实现的。根据每次训练集中每个样本的分类是否正确。以及上次的总体分类的准确率, 来确定每个样本的权值。将修改过权值的新数据集送给下层分类器进行训练, 最后将每次训练得到的分类器最后融合起来, 作为最后的决策分类器。使用AdaBoost分类器可以排除一些不必要的训练数据, 并将重点放在关]36键的训练数据</p>
13	<p><b>送检论文片段</b></p> <p>位置:</p>  <p>训练集送给下层分类器进行训练, 最后将每次训练得到的分类器最后融合起来, 作为最终可以使用的决策分类器。使用Ada boost构建的分类器可以自动无视掉一些的训练集中的数据特征, 并将核心的分类点放在最关键的训练数据上面。              AdaBoost方法针对具备一定深度学习能力的分类器</p>	<p><b>相似论文片段</b> <b>【0.58%】</b></p> <p>来源: <u>基于差值区域的实时人脸定位</u>              [期刊论文]《辽宁工程技术大学学报(自然科学版)》, 2011年 姜庆玲 等</p> <p>确定每个样本的权值。将修改过权值的新数据集送给下层分类器进行训练, 最后将每次训练得到的分类器最后融合起来, 作为最后的决策分类器。使用AdaBoost分类器可以排除一些不必要的训练数据, 并将重点放在关键的训练数据上面。通过AdaBoost算法生成了由重要特征组成的强分类器, 将强分类器串联在一起形成级联分类器</p>
14	<p><b>送检论文片段</b></p> <p>位置:</p>  <p>得到一个弱分类器。</p> <p>3. 将1和2都错了的样本加上其他的新样本构成另一个新的训练集, 通过对这个样本的学习得到第三个弱分类器。</p> <p>4. 循环重复至新分类其精度提高小于某个阈值为止。</p> <p>5. 组合所有的弱分类器构成一个提升的强分类器, 每个弱分类器都需要某个权值来控制, 即</p>	<p><b>相似论文片段</b> <b>【0.58%】</b></p> <p>来源: <u>嵌入式系统上基于近红外图像的人脸检测的研究</u>              [学位论文]王基帆, 2009年 华东师范大学</p> <p>运算, 3次加减运算, 在取值时计算偏移地址时需要乘加运算各1次, 在理想状况下共15个时钟周期。一个弱分类器的判定结果需要取3个矩形的像素总和, 共45个时钟周期, 还需要乘以偏置方向, 与阈值比对和对结果赋值, 共3个时钟周期。因此一个弱分类器得到结果需要48个时钟周期。强分类器包含5个弱分类器, 每个弱分类器的值需要乘以权重, 共245个时钟周期。再与强分类器的阈值</p>

某个数据被分为哪一类要通过数个弱分类器

## 六、全部举例相似论文作者（共10个）

序号	作者	典型片段总相似比	剩余相似比
1	高淑琴	6.09%	6.67%
2	崔争艳	1.45%	11.30%
3	蒲筱哥	1.16%	11.59%
4	刘爽	0.58%	12.17%
5	陈晓	0.58%	12.17%
6	王基帆	0.58%	12.17%
7	姜庆玲	0.58%	12.17%
8	吴韦	0.58%	12.17%
9	马颖哲	0.58%	12.17%
10	孙劲光	0.58%	12.17%

## 七、相似论文（举例128篇）

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
1	5.80%	Web文本分类技术研究现状述评	参考文献	期刊论文	高淑琴	图书情报知识	2008
2	1.45%	中文短文本分类的相关技术研究	参考文献	学位论文	崔争艳	河南大学	2011
3	4.35%	Web自动文本分类技术研究综述		期刊论文	蒲筱哥	情报学报	2009

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
4	4.06%	<u>自动文本分类方法研究述评</u>		期刊论文	蒲筱哥	情报科学	2008
5	2.32%	<u>分类方法的新发展:研究综述</u>		期刊论文	张丽娟 等	计算机科学	2006
6	2.03%	<u>中文文本分类关键技术研究</u>		学位论文	何钟莉	西安电子科技大学	2009
7	1.74%	<u>新闻搜索分析平台的相关技术研究</u>		学位论文	李丹东	哈尔滨工程大学	2009
8	1.45%	<u>网络环境下文本自动分类方法研究综述</u>		期刊论文	吴波 等	鸡西大学学报	2009
9	1.45%	<u>基于领域词语本体的短文本分类研究</u>		学位论文	宁亚辉	重庆邮电大学	2009
10	1.45%	<u>文本分类技术进展</u>		期刊论文	胡文静 等	知识经济	2011
11	1.16%	<u>网页与文本自动分类综述</u>		期刊论文	王明文 等	南昌工程学院学报	2005
12	1.16%	<u>中文Web文本自动分类的研究与实现</u>		学位论文	刘磊	长春理工大学	2007
13	1.16%	<u>基于RBF网络的文本自动分类的研究</u>		学位论文	王晓庆	江西师范大学	2003
14	0.87%	<u>基于语义的网络知识获取相关技术研究</u>		学位论文	郭勇	国防科学技术大学	2007
15	0.87%	<u>文本分类综述</u>		期刊论文	靳小波 等	自动化博览	2006
16	0.87%	<u>论文投稿系统评审专家自动推荐模型研究</u>		学位论文	刘一星	重庆大学	2009
17	0.87%	<u>因特网文本智能挖掘的模糊聚类算法研究</u>		期刊论文	王晓勇 等	计算机仿真	2009
18	0.87%	<u>Web文本分类研究及应用</u>		学位论文	柯慧燕	武汉理工大学	2006
19	0.87%	<u>基于XML的异构数据处理技术研究</u>		学位论文	周永健	北京工业大学	2007
20	0.87%	<u>汉语主客观文本分类及预处理研究</u>		学位论文	张霄凯	上海交通大学	2009
21	0.87%	<u>基于聚类的主题模型短文本分类方法研究</u>		学位论文	黄鑫	重庆邮电大学	2010

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
22	0.87%	<u>基于SVM的多类文本分类研究</u>		学位论文	梁秀娟	中南财经政法大学	2008
23	0.87%	<u>一种多视角人脸检测方法</u>		期刊论文	马颖哲 等	计算机工程与科学	2011
24	0.87%	<u>基于用户行为的高校BBS热帖预测模型</u>		期刊论文	于兴隆 等	计算机应用与软件	2013
25	0.87%	<u>基于支持向量机的网页自动分类方法研究与实现</u>		学位论文	周兰珍	北京工业大学	2008
26	0.87%	<u>基于随机前沿和随机森林法的沿海开发区发展效率研究</u>		学位论文	倪强	天津大学	2011
27	0.87%	<u>基于差值区域的实时人脸定位</u>		期刊论文	姜庆玲 等	辽宁工程技术大学学报（自然科学版）	2011
28	0.87%	<u>基于人工免疫机制的Web文本分类研究</u>		学位论文	杜英	电子科技大学	2006
29	0.87%	<u>基于嵌入式系统人脸识别方法的研究</u>		学位论文	田洪贞	青岛科技大学	2012
30	0.87%	<u>基于Adaboost的超分辨率重建算法</u>		学位论文	付廷杰	西安电子科技大学	2010
31	0.87%	<u>嵌入式系统上基于近红外图像的人脸识别研究</u>		学位论文	钱艺	华东师范大学	2009
32	0.87%	<u>基于肤色和Adaboost算法的人脸检测方法研究</u>		学位论文	艾学轶	华中师范大学	2009
33	0.87%	<u>SVM与AdaBoost算法的应用研究</u>		学位论文	宋静	大连海事大学	2011
34	0.87%	<u>智能移动机器人的定位与环境建模研究</u>		学位论文	赵黎明	重庆邮电大学	2011
35	0.87%	<u>高速监控视频中的人脸检测研究与实现</u>		学位论文	陈抒	华东师范大学	2009
36	0.87%	<u>基于潜在语义分析的Web文本分类研究</u>		学位论文	王剑锋	华北电力大学(保定)	2006
37	0.87%	<u>统一利用外部语义知识改进文本分类的研究</u>		学位论文	侯俊琦	南开大学	2009
38	0.87%	<u>中文Web文本分类技术研究</u>		学位论文	曹勇	厦门大学	2007

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
39	0.87%	<u>基于Web的中文文档自动分类的研究与实现</u>		学位论文	张楠	哈尔滨理工大学	2007
40	0.87%	<u>Web文本分类关键技术研究</u>		学位论文	尹世群	西南大学	2008
41	0.87%	<u>灾难现场中人体腿脚部检测</u>		学位论文	赵宏	哈尔滨工程大学	2010
42	0.87%	<u>基于AdaBoost和Krawtchouk矩的刀具形状分类识别</u>		期刊论文	赵敏 等	锻压装备与制造技术	2012
43	0.87%	<u>一种潜在语义索引差异模型及其应用</u>		学位论文	米晓芳	烟台大学	2007
44	0.58%	<u>基于机器学习的文本分类方法综述</u>		期刊论文	陈祎荻 等	渤海大学学报（自然科学版）	2010
45	0.58%	<u>基于人脸检测的智能监控追踪系统</u>		学位论文	王飞	华中科技大学	2011
46	0.58%	<u>Deep Web数据源发现与分类研究</u>		学位论文	戚春超	南京信息工程大学	2011
47	0.58%	<u>增强现实人机交互系统的研究</u>		学位论文	高净业	杭州电子科技大学	2011
48	0.58%	<u>面向嵌入式手机平台人脸跟踪技术研究</u>		学位论文	赵季方	长春理工大学	2012
49	0.58%	<u>基于图模型的Web文档分类方法研究</u>		学位论文	张炼	内蒙古科技大学	2010
50	0.58%	<u>含光伏电站配电网无功优化的研究</u>		学位论文	郭康	华北电力大学(保定)	2012
51	0.58%	<u>嵌入式系统上基于近红外图像的人脸检测的研究</u>		学位论文	王基帆	华东师范大学	2009
52	0.58%	<u>基于科技文献的中文文本分类算法研究</u>		学位论文	王俊英	燕山大学	2007
53	0.58%	<u>基于类别结构的文本层次分类方法研究</u>		学位论文	祝翠玲	山东大学	2011
54	0.58%	<u>文本分类语料库自动创建系统的研究与实现</u>		学位论文	吴韦	武汉理工大学	2009
55	0.58%	<u>离线阿拉伯手写体光学文字识别方法的研究</u>		学位论文	程艳芬	武汉理工大学	2009
56	0.58%	<u>基于ASM自动人脸特征点定位及应用的研究</u>		学位论文	孙砚铭	北京交通大学	2010

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
57	0.58%	<u>基于openCV的人脸检测系统的设计</u>		期刊论文	陈志恒 等	电子设计工程	2012
58	0.58%	<u>基于Log-Gabor和AdaBoost的人脸识别算法研究</u>		期刊论文	李迪 等	微型电脑应用	2011
59	0.58%	<u>特征选择方法对英文作文自动评分性能影响的研究</u>		学位论文	崔爱国	苏州大学	2009
60	0.58%	<u>人脸识别在公安布控系统中的应用研究</u>		学位论文	郭凤	上海大学	2010
61	0.58%	<u>文本情感分类及观点摘要关键问题研究</u>		学位论文	张冬梅	山东大学	2012
62	0.58%	<u>基于视频的实时人脸识别的研究与实践</u>		学位论文	徐敏	中南大学	2010
63	0.58%	<u>视频监控系统中基于神经网络的人物行为识别研究</u>		学位论文	王艳丽	华北电力大学(保定)	2012
64	0.58%	<u>随机生存森林在结直肠癌预后分析的应用</u>		学位论文	洪远芳	中山大学	2010
65	0.58%	<u>文本分类及其相关技术的研究</u>		期刊论文	刘冬雪 等	科技创新导报	2010
66	0.58%	<u>人工免疫优化与分类算法及其应用研究</u>		学位论文	王炼红	湖南大学	2009
67	0.58%	<u>基于特征的人脸分类算法研究</u>		期刊论文	李霞 等	微型电脑应用	2010
68	0.58%	<u>基于视频图像的人脸检测与统计</u>		期刊论文	徐麒 等	计算机与现代化	2010
69	0.58%	<u>基于强分类器的神经网络三维目标识别</u>		期刊论文	何灏 等	计算机测量与控制	2012
70	0.58%	<u>基于图像的人脸检测和识别进展研究</u>		期刊论文	张卫 等	大众科技	2010
71	0.58%	<u>驾驶疲劳的视频研究</u>		学位论文	刘鹏	中国科学技术大学	2009
72	0.58%	<u>军用信息自动分类的研究与实现</u>		学位论文	吴娟	南京理工大学	2004
73	0.58%	<u>搜索引擎优化技术的研究与实现</u>		学位论文	黄平珍	四川师范大学	2010
74	0.58%	<u>BBS舆情智能分析系统研究与实现</u>		学位论文	任晓霞	沈阳理工大学	2011



序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
75	0.58%	<u>中国马克思主义大众化历史发展规律研究</u>		学位论文	綦玉帅	苏州大学	2011
76	0.58%	<u>互联网舆情分析关键技术的研究与实现</u>		学位论文	章栋兵	武汉理工大学	2010
77	0.58%	<u>基于网络透视理论的网络链路性能测量研究</u>		学位论文	许鑫	北京工业大学	2010
78	0.58%	<u>基于协议分析和免疫原理的入侵检测技术研究</u>		学位论文	曹传林	江苏科技大学	2011
79	0.58%	<u>马克思主义“人的本质”理论对思想政治教育的指导作用</u>		学位论文	刘瑜	四川师范大学	2010
80	0.58%	<u>基于关键字的搜索引擎搜索结果中文网页分类研究</u>		学位论文	周莹	南开大学	2005
81	0.58%	<u>中文产品评论挖掘关键技术研究</u>		学位论文	黄永文	重庆大学	2009
82	0.58%	<u>基于特定领域的Web文本信息获取系统的设计与实现</u>		学位论文	赵栓柱	太原理工大学	2005
83	0.58%	<u>浅谈基于数据挖掘文档分类方法研究</u>		期刊论文	孟海东 等	现代计算机（专业版）	2010
84	0.58%	<u>面向新闻领域的文本数据获取系统的设计与实现</u>		学位论文	王宝龙	北京邮电大学	2010
85	0.58%	<u>大学生理想信念教育环境探究</u>		学位论文	张学宝	天津工业大学	2009
86	0.58%	<u>略论企业政工网络的建设</u>		期刊论文	解学刚 等	商场现代化	2010
87	0.58%	<u>网络语言的社会语言学考察</u>		期刊论文	康忠德	石家庄学院学报	2011
88	0.58%	<u>当代大学生网络生活现状分析</u>		期刊论文	马磊 等	安徽文学（下半月）	2010
89	0.58%	<u>我国东西部互联网发展的比较分析</u>		期刊论文	石屹 等	学海	2010
90	0.58%	<u>当下阅读的缺失与拯救</u>		期刊论文	姜陵陵 等	铜陵学院学报	2010
91	0.58%	<u>阅读方式多元化与高校图书馆信息资源优化配置</u>		期刊论文	吴利萍 等	现代情报	2010
92	0.58%	<u>从媒介变迁史看网络阅读</u>		期刊论文	李晓源	图书与情报	2010

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
93	0.58%	<u>数字化阅读与图书馆服务创新</u>		期刊论文	陈秀钦	科技情报开发与经济	2010
94	0.58%	<u>C2C电子商务用户信息搜寻行为研究</u>		学位论文	刘鹏	中山大学	2010
95	0.58%	<u>浅谈财会专业学生网上创业能力的培养</u>		期刊论文	杨向苗 等	才智	2012
96	0.58%	<u>中国联通德州分公司营销能力提升策略研究</u>		学位论文	崔亦刚	山东大学	2010
97	0.58%	<u>装傻、戏仿与反讽——网络恶搞图像与公共话语权的关系研究</u>		学位论文	何艺忠	中山大学	2010
98	0.58%	<u>基于云计算的跨越我国城乡数字鸿沟研究</u>		期刊论文	李培维 等	情报探索	2011
99	0.58%	<u>新媒体时代的思考:网络媒介与手机媒介知识信息服务在图书馆的应用</u>		期刊论文	黄幼菲 等	图书馆理论与实践	2010
100	0.58%	<u>推动全民阅读共享书香校园——以驻马店市黄淮学院图书馆为例</u>		期刊论文	白永红 等	金山	2011
101	0.58%	<u>大众化:新媒体环境下档案利用的发展路向</u>		期刊论文	陈菲 等	山西档案	2011
102	0.58%	<u>电子邮件过滤系统的研究与设计</u>		学位论文	杨盛	中南大学	2005
103	0.58%	<u>神经网络作为分类器的算法研究及在信息检索中的应用</u>		学位论文	周瑛	安徽大学	2006
104	0.58%	<u>基于支持向量机的中文文本分类研究</u>		学位论文	魏新	中国民用航空学院	2004
105	0.58%	<u>中文科技论文文本分类研究</u>		学位论文	陈频	厦门大学	2006
106	0.58%	<u>基于支持向量机的文本分类在并行环境中的性能分析及应用</u>		学位论文	殷勇	上海交通大学	2006
107	0.58%	<u>具有Anti-spam功能的安全电子邮件系统的研究与实现</u>		学位论文	王凯	上海交通大学	2003
108	0.58%	<u>基于主题概念树的科技文本分类方法研究和实现</u>		学位论文	张焕哲	华北电力大学(北京)	2006
109	0.58%	<u>基于词语网络的关键词自动提取方法及在中文网页分类中的应用研究</u>		学位论文	温安国	复旦大学	2009

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
110	0.58%	<u>基于用户兴趣模型的新闻自动推荐系统</u>		学位论文	李伟	复旦大学	2009
111	0.58%	<u>快速图像文档分类的研究</u>		学位论文	王谦	北京邮电大学	2005
112	0.58%	<u>基于词跨度的中文文本关键词提取及在文本分类中的应用</u>		学位论文	谢晋	浙江工业大学	2011
113	0.58%	<u>HTML文本自动分类技术的研究与工具的实现</u>		学位论文	刘冬梅	内蒙古大学	2006
114	0.58%	<u>基于支持向量机的文本分类问题研究</u>		学位论文	代亮	大连海事大学	2007
115	0.58%	<u>文本分类特征选取技术研究</u>		学位论文	郑伟	内蒙古大学	2008
116	0.58%	<u>基于领域辞典的航空安全报告自动分类研究</u>		学位论文	杨春宝	中国民航大学	2008
117	0.58%	<u>基于支持向量机的文本分类问题的研究</u>		学位论文	邸锦	北京交通大学	2008
118	0.58%	<u>多类别科技文献自动分类系统</u>		学位论文	陈玉芹	华中科技大学	2008
119	0.58%	<u>基于内容管理的文本自动分类的研究与应用</u>		学位论文	王艳	大连海事大学	2009
120	0.58%	<u>基于遗传算法与模糊聚类的文本分类研究</u>		学位论文	于水英	哈尔滨理工大学	2009
121	0.58%	<u>基于粗糙集和最大熵融合的文本分类方法研究</u>		学位论文	陈光	上海大学	2007
122	0.58%	<u>基于混合并行遗传算法的文本分类及聚类研究</u>		学位论文	戴文华	华中师范大学	2007
123	0.58%	<u>基于关联的中文文本分类系统的设计与实现</u>		学位论文	尚旭祥	中国人民解放军信息工程大学	2006
124	0.58%	<u>基于向量空间模型的中文文本分类的研究</u>		学位论文	吴科	山东大学	2004
125	0.58%	<u>中文文本分类算法的研究与实现</u>		学位论文	张志华	北京工业大学	2004
126	0.58%	<u>利用表格信息的Web文本分类研究与实现</u>		学位论文	廖涛	上海大学	2008
127	0.58%	<u>基于支持向量机的手机垃圾短信过滤器的研究</u>		学位论文	钱升华	河北工业大学	2006

序号	相似比	相似论文标题	参考文献	论文类型	作者	来源	发表时间
128	0.58%	基于RBF网络的中文文本自动分类的研究		学位论文	郁婵嫫	华东师范大学	2007

[查看全文报告请点击](#)

说明：

1. 总相似比 $\approx$ 送检论文与检测范围全部数据相似部分的字数/送检论文总字数
2. 参考文献相似比 $\approx$ 送检论文与其参考文献相似部分的字数/送检论文总字数
3. 排除参考文献相似比=总相似比-参考文献相似比
4. 剩余相似比 $\approx$ 总相似比-典型片段总相似比
5. 本报告为检测系统算法自动生成，仅供参考