

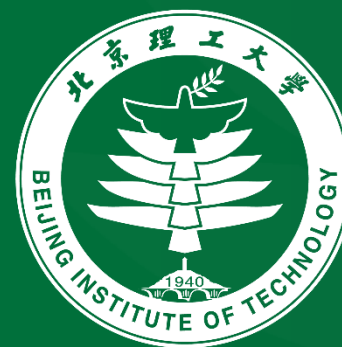
◁ BIT ▷

信息过滤

小组成员：丁兆旭 张文轩 董纪漩 张博文 耿中豪

时间：2022/4/17

德以明理 学以精工



北京理工大学
BEIJING INSTITUTE OF TECHNOLOGY



目录

CONTENTS

1 概述

2 算法

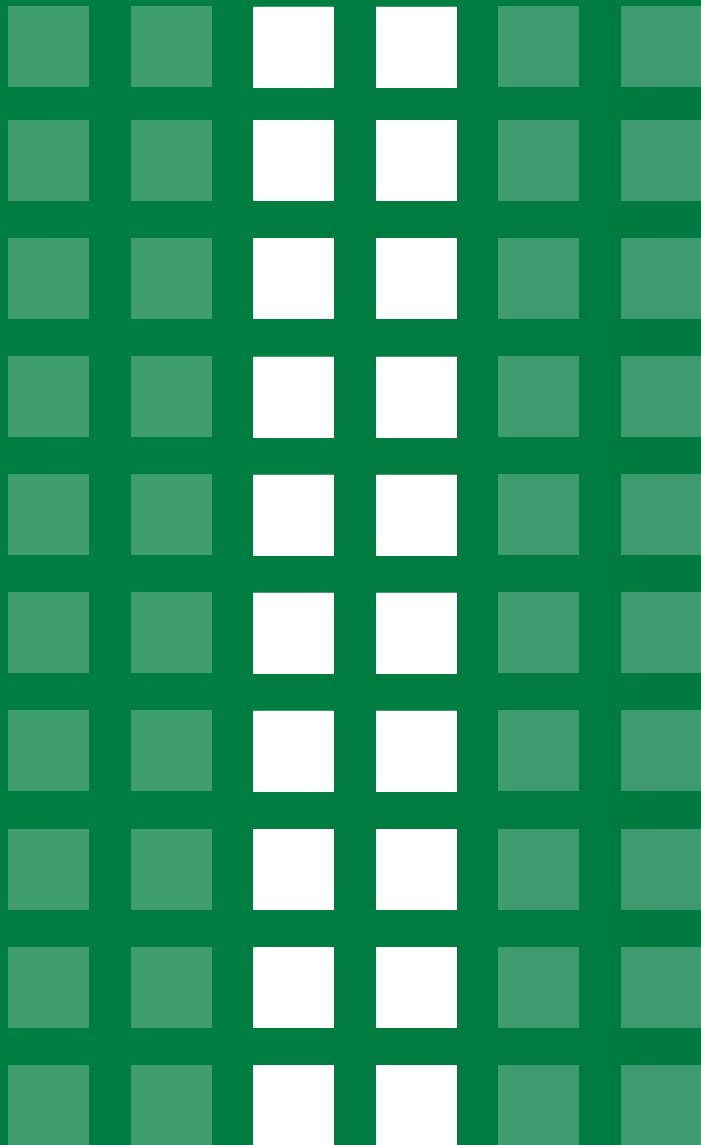
3 前沿

4 demo



PART ONE

信息过滤概述



大数据杀熟



信息茧房





定义

信息过滤是指从动态的信息流中将满足用户兴趣的信息挑选出来，用户的兴趣一般在较长一段时间内不会改变（静态）。信息过滤通常是在输入数据流中移除数据，而不是在输入流中找到数据。

*Communications of the ACM**Dec 1992 v35 n12 p29(10)**Page 1***Information filtering and information retrieval: Two sides of the same coin?***by Nicholas J. Belkin and W. Bruce Croft*

过滤和检索是相辅相成的，可以理解为信息过滤是建立在信息检索的基础上进行的更高级别的操作。信息检索和信息过滤实际上是一枚硬币的两面。他们一起工作，帮助人们获取完成任务所需的信息。

<https://dl.acm.org/doi/10.1145/138859.138861>

1.3

2005-2022相关论文数量



北京理工大学
BEIJING INSTITUTE OF TECHNOLOGY

information filtering

Search

[Advanced search](#)

ACMDL

Informati

Journal

Article type

Subject

Date

All

All

All

All

information filtering

Search

[Advanced search](#)

Journal

Article type

Subject

Date

All

All

All

2005-2022

Sort by: Relevance Date published (new to old) Date published (old to new)

Showing 1-50 of 31434 results

1.4

一些例子



← 骚扰拦截

QQ 邮箱

1元域名

小游戏

游戏直播

微信网页版

人民·新华

7 人读过

阅读

立即抢购



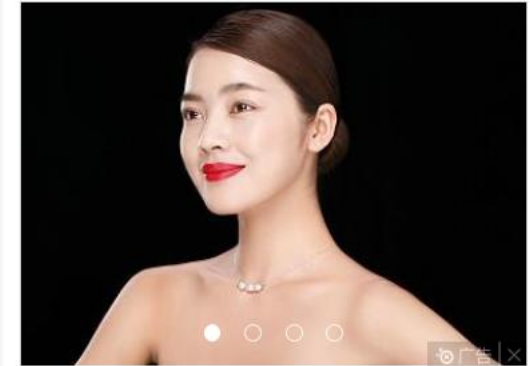
刘看山 · 知乎指南



体育 游戏 房产 财经 旅游

丝, 给女儿取名“紫怡”, 妈妈怒斥: 瞅瞅你

48%的人还浏览了



做下巴整形



为回应符合本地法律要求的通知, 部分搜索结果未予显示。有关详细信息, 请参阅[此处](#)。

一些您可能无法访问的结果已被隐去。

全部删除

号码申诉

拦截规则



1.按展现形式划分

- (1) 文本类
- (2) 图像类
- (3) 数据类

2.按是否过滤的信息预处理

(1) 主动

预先处理网络信息,如预先对网页或网站进行分级、建立允许或禁止访问的地址列表等,当用户访问时根据分级标记或地址列表决定能否访问

(2) 被动

当用户访问时才对地址、文本或图像等信息进行分析以决定是否过滤



脑容量有限

信息过载

海量、繁多的信息对我们感知世界是有帮助的，但带来的更多的是烦恼。因此我们需要信息过滤，来协助我们获得更优质、相关度更高的信息；同时去除有害、不良的信息。



《“十四五”国家信息化规划》专家谈：加强数字化发展治理 推进数字中国建设

2022年02月16日 17:23

来源：人民日报



[【打印】](#) [【纠错】](#)



3月17日下午，国务院新闻办公室举行2022年“清朗”系列专项行动新闻发布会。国家网信办副主任盛荣华出席发布会，介绍2022年“清朗”系列专项行动情况，并回答记者提问。

盛荣华指出，以习近平同志为核心的党中央高度重视网络生态治理工作。习近平总书记多次强调，网络空间是亿万民众的精神家园，要本着对社会负责、对人民负责的态度，依法加强网络空间治理，为广大网民特别是青少年营造一个风清气正的网络空间。2021年，国家网信办部署开展“饭圈”乱象整治、春节网络环境整治、用户账号运营乱象整治等15项“清朗”系列专项行动，累计清理违法和不良信息2200万余条，处置账号13.4亿个，封禁主播7200余名，下架应用程序、小程序2160余款，关闭网站3200余家。网络生态日益向好，网络空间日益清朗，赢得广大网民和社会各界的积极支持。但同时，网络生态乱象极易反弹反复，一些顽瘴痼疾仍然不同程度存在，许多深层次问题还没有根本解决，新情况新问题也在不断出现，这就要求必须持续深入、久久为功。



1 基于规则的过滤方法

设定了一系列规则来定义信息是否为垃圾信息，然后根据这些规则依据一定的算法对信息文本进行匹配，以此判断信息是否是垃圾信息

2 基于机器学习分类的过滤方法

是先获取相关数据，并对数据进行一定的预处理；再对数据进行特征提取，获取信息特征；然后根据一定的算法对信息特征赋予相应的权值；最后在这些基础上采用分类算法对信息进行分类，过滤出垃圾信息。

基于规则	应用场景	优点	缺点
基于黑白名单过滤	手机号、邮箱地址		需维护大量黑白名单
基于URL过滤	恶意、不良的网址	便于操作、实施简单、直观	需不断审查URL地址
基于关键词过滤	搜索引擎、不良、垃圾信息		需不断更新关键词列表

基于机器学习分类

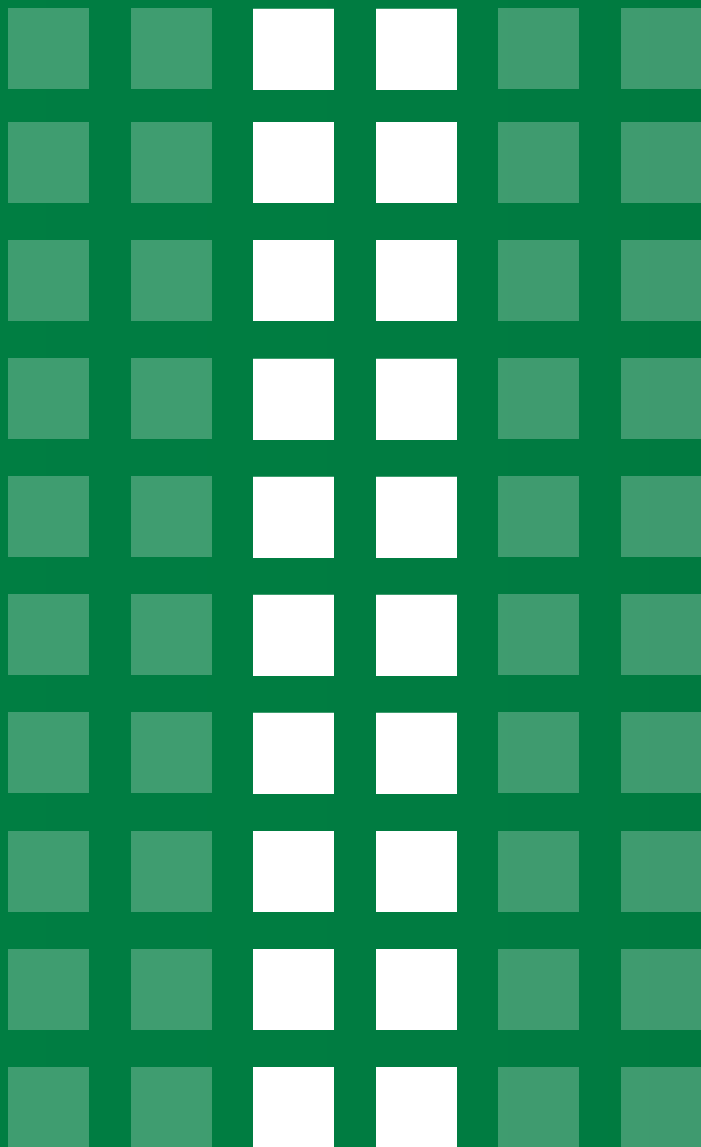
贝叶斯分类 (Bayes Classifier)

决策树分类 (Decision Tree Classifier)

支持向量机算法 (SVM)

PART TWO

信息过滤算法





基于因特网内
容分级平台过
滤(PIcs)



数据库过滤
(IP库、URL
库)



关键词过滤



基于内容理解
的过滤





定制规范

W3C (World Wide Web Consortium)于1995年8月组织当时因特网上的主导力量共同商讨制定一套技术规范。



分级

用户可以通过分级标记了解到分级机构和分级体系的情况。从而在使用时下载合适的过滤系统分级档案。



网页内容分级法

采取人工分级方法过滤不良信息。其错误率相对较低，并可以准确地对图像、视频等多媒体信息进行准确分级。

形成网址库

- 数据库过滤就是通过对网上各种信息进行分类后。精确地匹配 URL和与之对应的页面内容。形成一个预分类的网址库。

网址库的类型

一种为“黑名单”，包括禁止访问的目标网站的URL。另一个是“白名单”。包括允许访问的目标网站的URL。

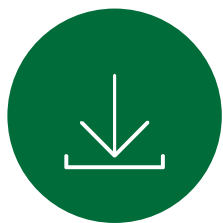


关键字

索引词

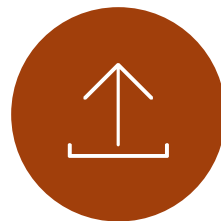


基于内容理解的过滤是指对获取的网络信息内容进行识别、判断、分类，确定其是否为需要过滤的尽标内容，并对已确定曲目标内容 进行过滤等检测控制的技术。它是根据不良信息的一些内容所呈现的 特征来进行判断的一种识别过滤方法。基于内容理解的过滤技术具有 适应性强、精度高等特点，能有效地解决基于数据库和关键词的过滤 技术不能处理好的问题，包括文本内容理解过滤、图像内容分析过滤、 视频内容分析过滤和智能混合过滤等。



第一级过滤

基于查询关键词的过滤和基于数据库的过滤。这两种过滤方式过滤速度快。可以快速剔除一部分不良信息。缓解了系统压力。



第二级过滤

基于内容理解的过滤，是主要的过滤环节。两种过滤方式的有机结合，在一定程度上提高了过滤的效率和准确率。



基于规则的过滤方法

基于规则的过滤方法设定了一系列规则来定义信息是否为垃圾信息，然后根据这些规则依据一定的算法对信息文本进行匹配，以此判断信息是否是垃圾信息。具体的过滤规则包括黑白名单过滤规则、关键词过滤规则等。

黑白名单过滤技术

是通过预先建立黑白名单来过滤信息，多用于垃圾邮件过滤。指当信息发送到接收方时，接收方根据系统设定的黑名单和白名单对信息进行接收或者阻挡。该名单可预先根据一定的信息建立，并且随时更新名单。

关键词过滤技术

基于关键字匹配的过滤方法是创建一个关键字列表，然后按照该列表确定信息是否为垃圾信息。列表是从大量垃圾邮件中总结出来的经常出现的单词，例如“中奖”“特惠”等。该方法需要创建一个庞大的过滤关键字列表并保持更新。



基于机器学习分类的过滤方法

基于机器学习分类的过滤方法就是先获取相关数据，再对数据进行特征提取，获取信息特征；然后根据一定的算法对信息特征赋予相应的权值；最后在这些基础上采用分类算法对信息进行分类，过滤出垃圾信息。

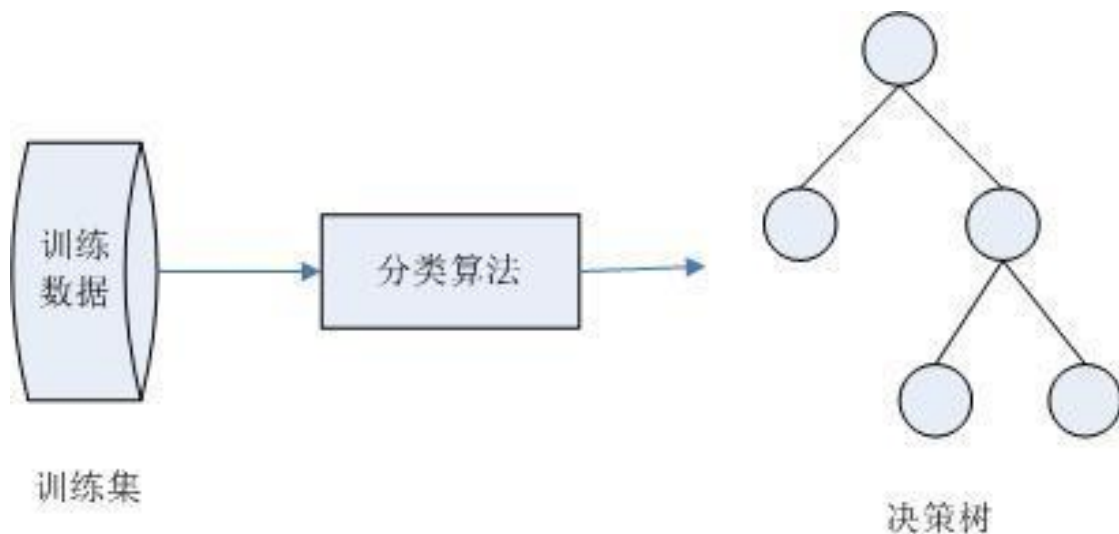


基于贝叶斯分类的垃圾信息过滤方法利用已知的社交网络信息，建立一个垃圾信息和正常信息关键词的贝叶斯模型，根据模型判断信息是否是垃圾信息。

- 简单贝叶斯分类算法原理大致如下，假设文本内容中包含的词汇为W，垃圾信息S，正常信息H。 现有一个文本，内容包含的词汇为W，判断该文本是否是垃圾信息，即计算P (S|W) 这个条件概率。根据贝叶斯定理得：

$$\Pr(S|W) = \frac{\Pr(W|S) \cdot \Pr(S)}{\Pr(W|S) \cdot \Pr(S) + \Pr(W|H) \cdot \Pr(H)}$$

决策树利用树的结构将数据记录进行分类。如图所示





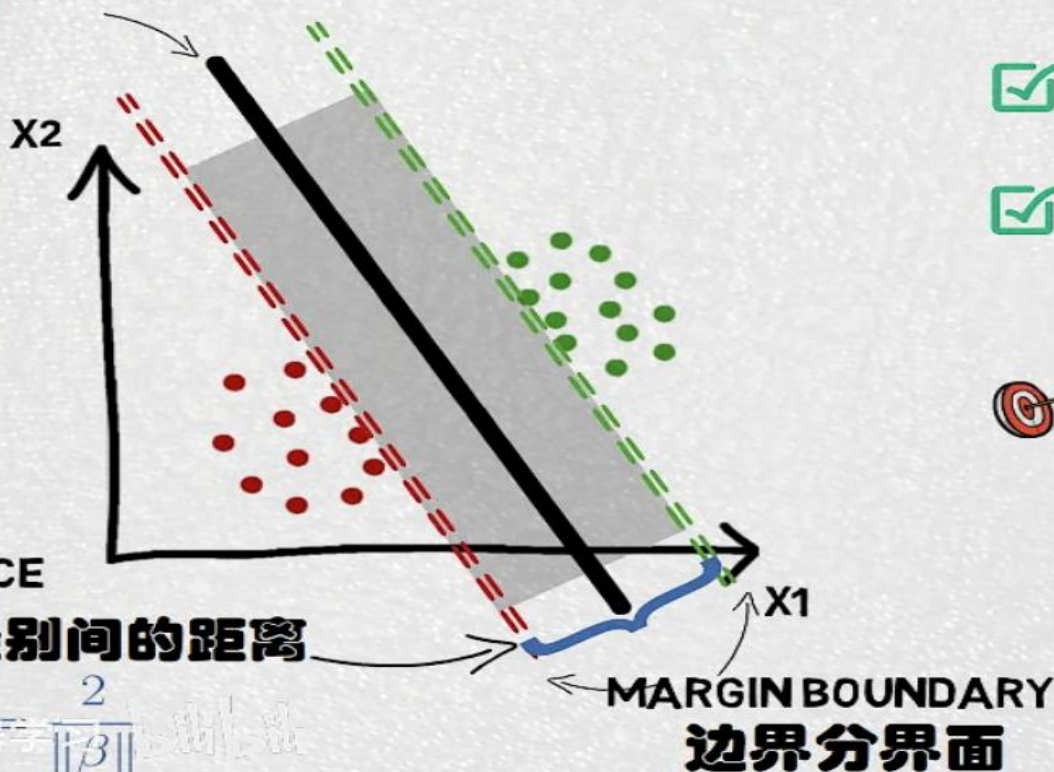
支持向量机（简称SVM），是20世纪90年代提出的一种二分类学习方法。它通过构造最优线性分类面来指导信息分类：引入高维特征空间，利用核函数，将输入空间的非线性决策便捷转化为高维特征空间的线性决策边界。

MATH BEHIND SVM

数学原理

$$\beta_1 x_1 + \beta_2 x_2 + \beta_0 = 0$$

DECISION BOUNDARY
决策分界面



SUPERVISED LEARNING
监督学习



LINEAR CLASSIFIER
线性分类器

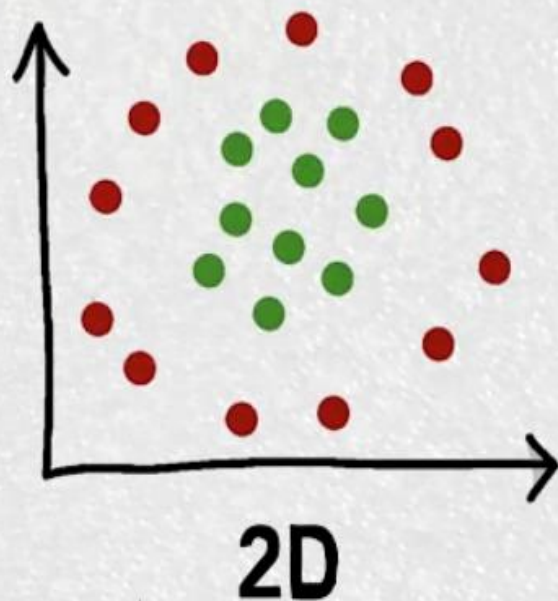


找到合适的 β
来描述决策分界面和边界分界面
(也叫超平面)

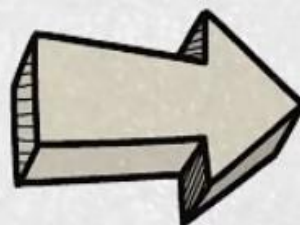
$$\begin{cases} \beta_1 x_1 + \beta_2 x_2 + \beta_0 = +1 \\ \beta_1 x_1 + \beta_2 x_2 + \beta_0 = -1 \end{cases}$$

五分钟机器学习 $dis = \frac{2}{\|\beta\|}$

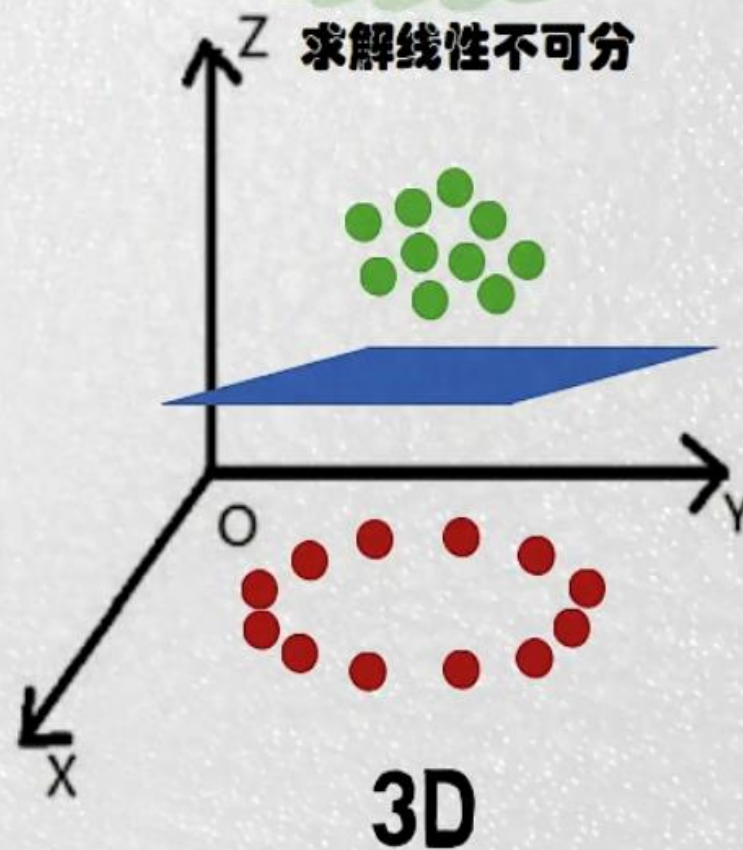
2. 线性分类器,
不可用于求解线性完全不可分的情况



KERNEL



KERNEL SVM



五分钟机器学习

基于规则的过滤方法

关键词过滤使用简单，易于理解，但需要不断维护更新关键词列表。黑白名单技术会因地址过时而效率降低，误判率提高。

基于机器学习的过滤方法

过滤正确率高而且速度快，机器学习方面的研究也较为成熟

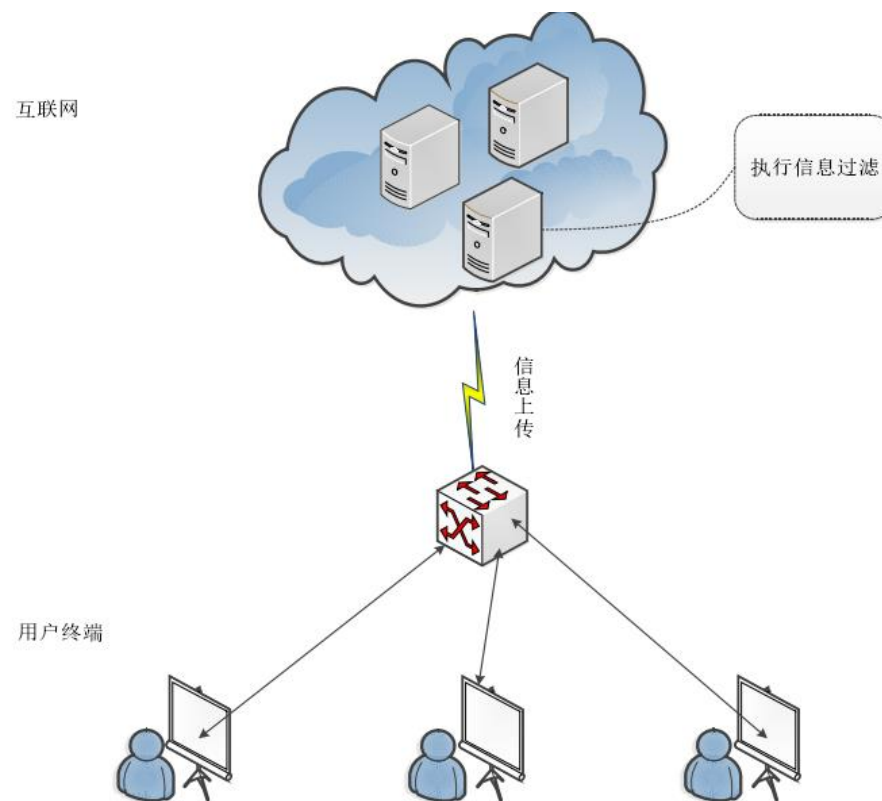
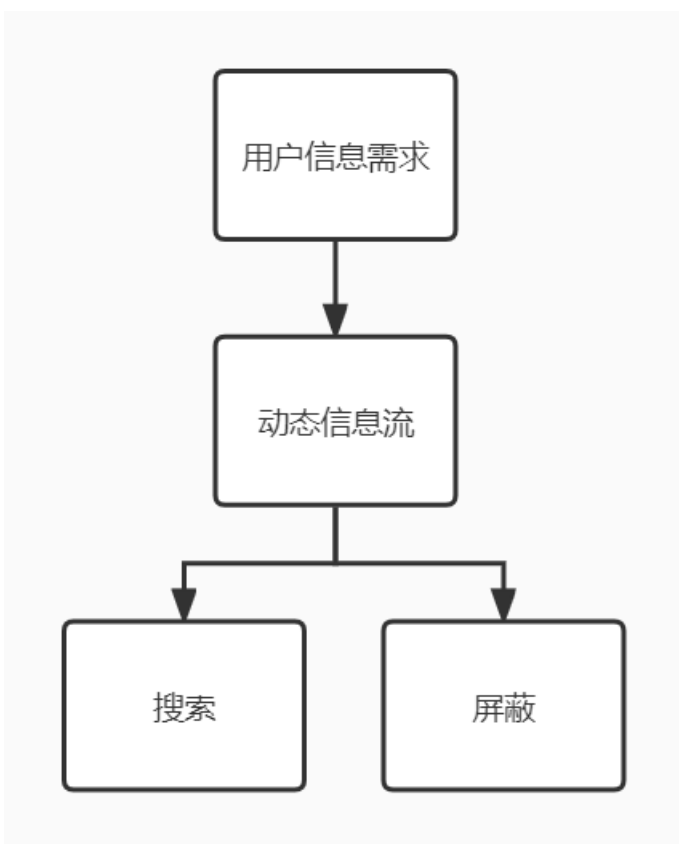
PART THREE

前沿与发展应用



飞速发展

- 网络信息指数级速度增长
- 技术飞速发展
- 世界范围广泛应用





■ 2018年

李泰提出一种大数据环境下海量多媒体信息过滤技术

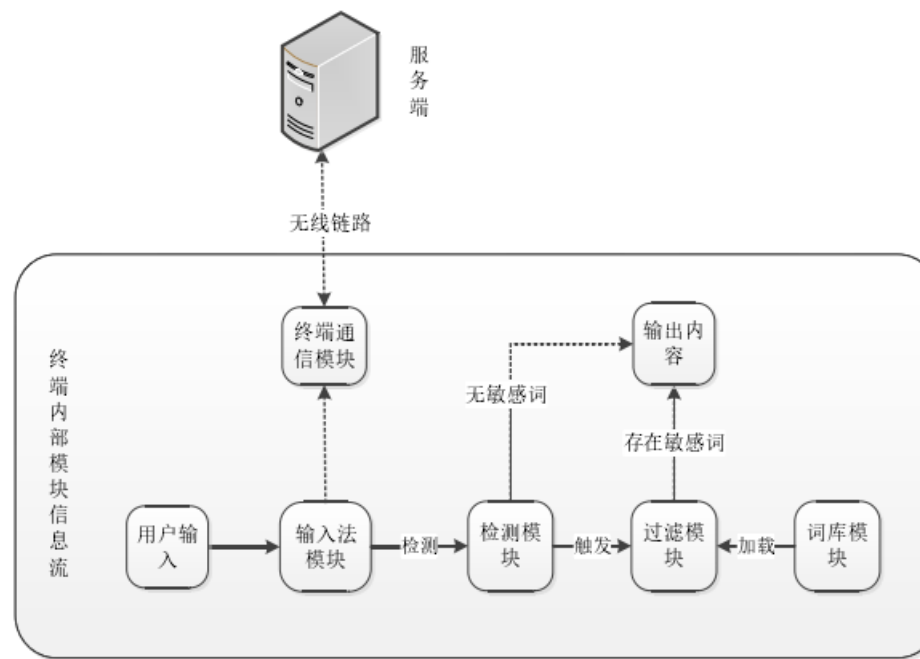
邹苹钧提出个性化手机垃圾信息过滤机制

■ 2019年

赵旭采取图像识别分析来实现敏感信息的过滤

■ 2019年

石宪面向发送端提出了一种基于社团划分和亲密度量化的隐私控制框架,面向接收端设计了一种具有更新能力的垃圾信息过滤系统。

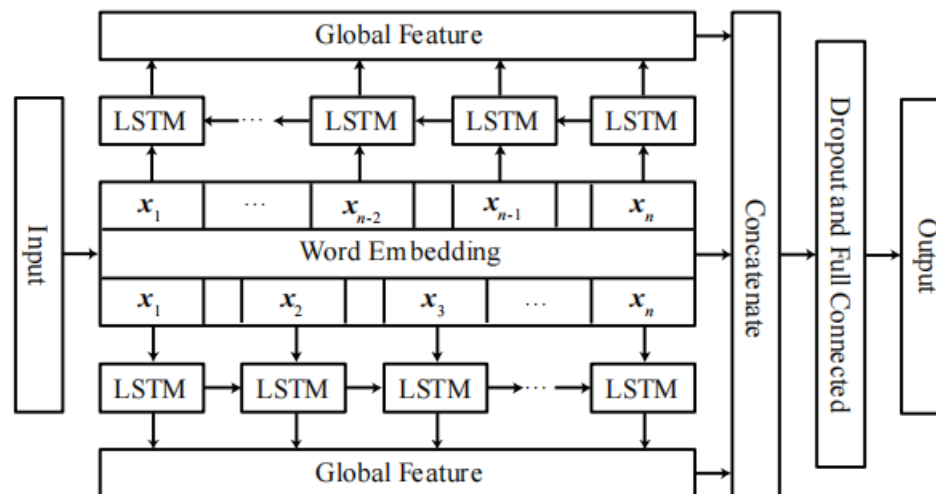


■ 2020年

陈国靖，詹跃明，分析 SVM 算法相比于其他数据挖掘方法的优势

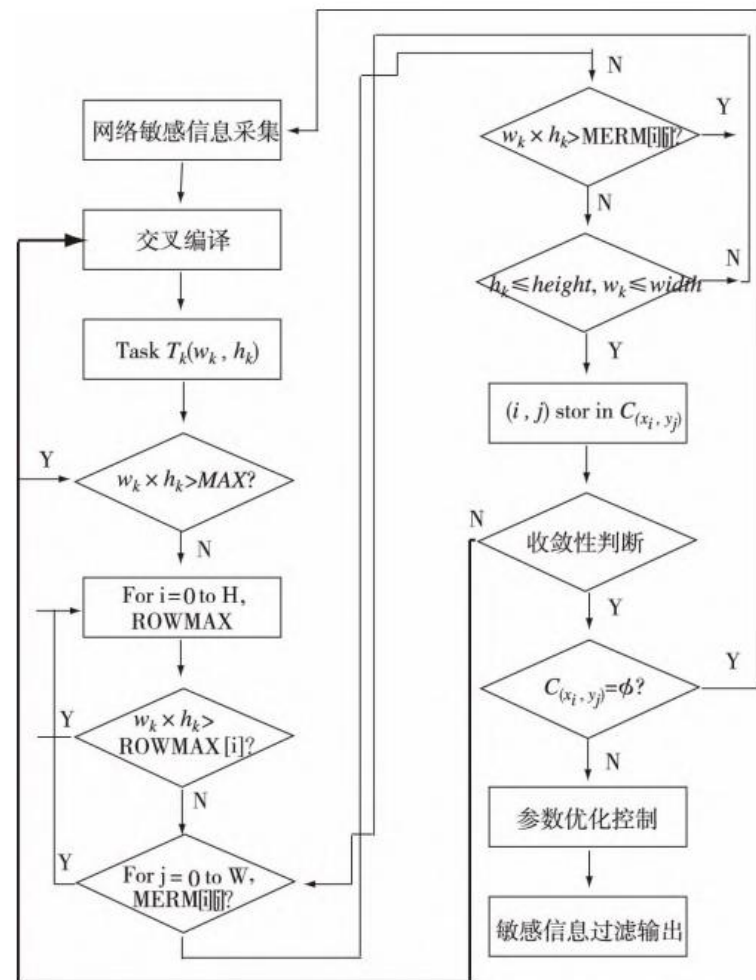
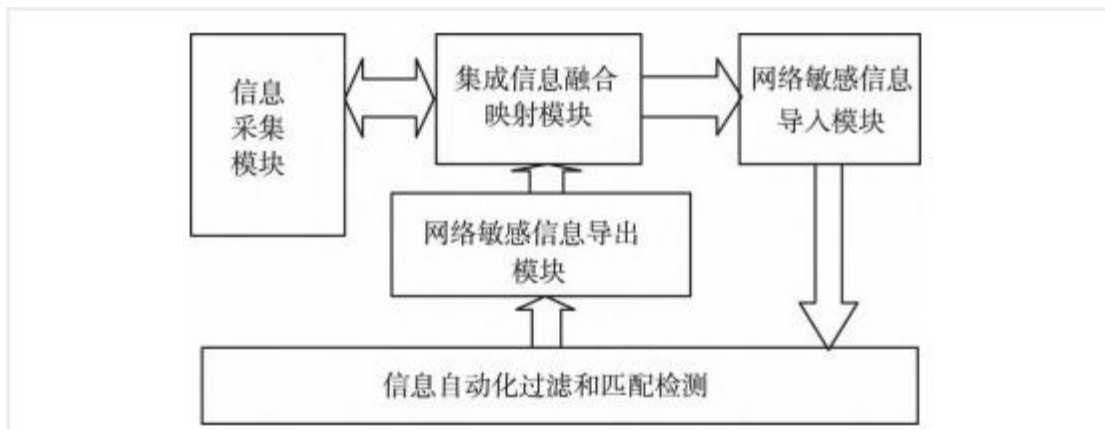
■ 2021年

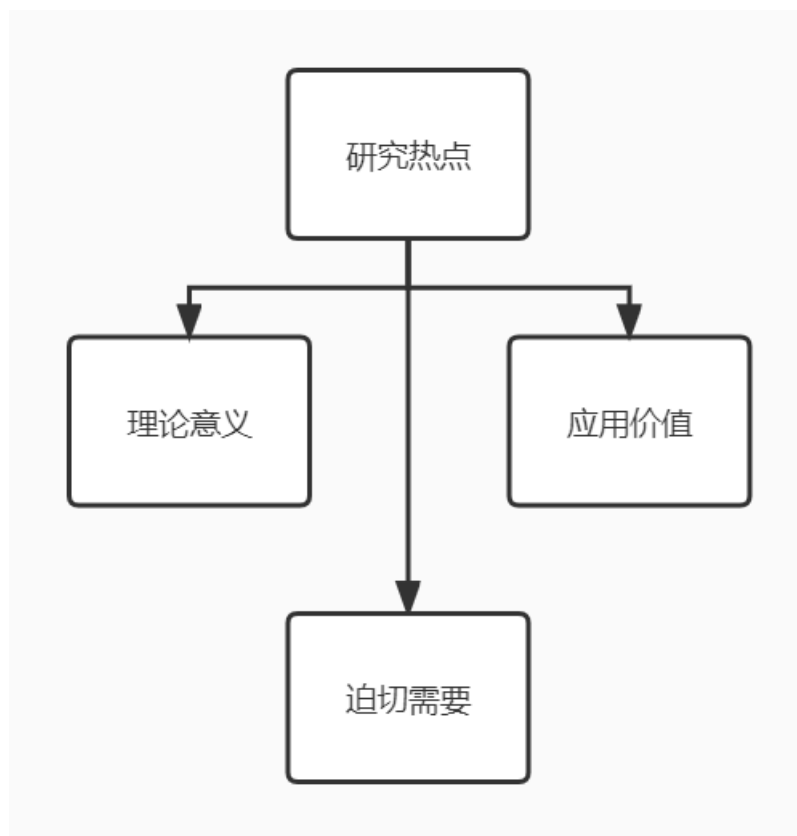
张建，严珂，马祥，以卷积神经网络和循环神经网络为基础，建立了三个模型，并在新数据集上进行识别训练



■ 2021年

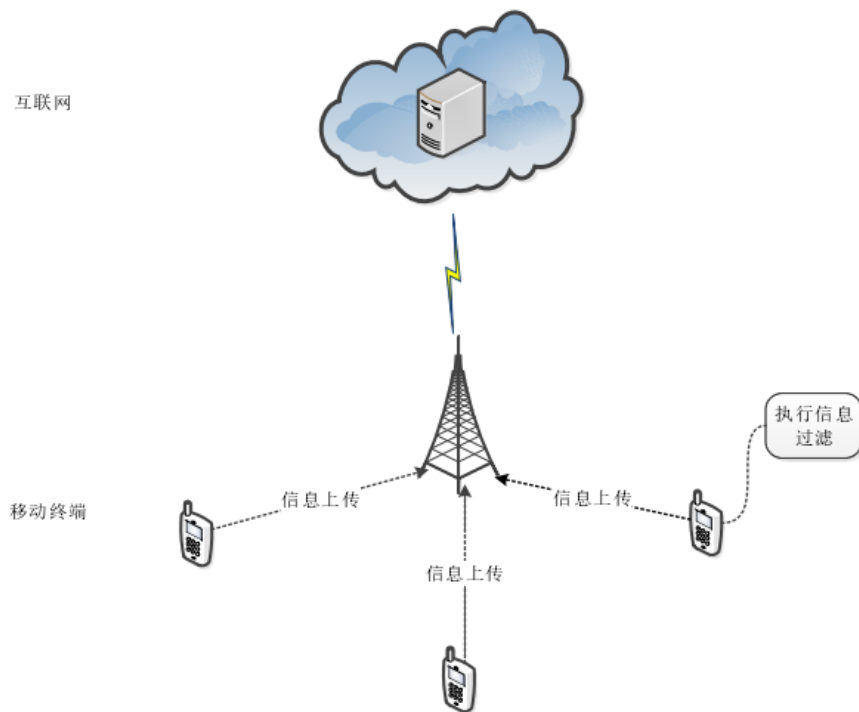
石小兵提出基于 K 近邻分类算法的网络敏感信息自动过滤方法





见解

- 信息过滤领域需进一步研究实践



结合应用

- 互联网交流软件信息过滤
- 健康良好的互联网环境



北京理工大学
BEIJING INSTITUTE OF TECHNOLOGY

谢谢观看
敬请各位老师批评指正

答辩人：北小理
导 师：京小工