

封面检测

论文类型“大学硕士毕业设计（论文）”错误，学术型硕士论文应为“硕士学位论文”  
论文类型“大学硕士毕业设计（论文）”每个字之间应包含一个空格  
第 1 行学生信息应为:作者姓名  
第 2 行学生信息应为:学科、专业  
第 3 行学生信息应为:学号  
第 4 行学生信息应为:指导教师  
学生信息有 1 行缺省  
封面中文学校名未居中  
封面英文学校名未居中

学位论文独创性声明检测

论文缺少学位论文独创性声明

摘要检测

中文关键词的内容字体错误，中文应为仿宋\_GB2312，英文应为 Times New Roman  
网络流量分类；哈希编码；大数据近邻检索；  
中文关键词的内容字体错误，中文应为仿宋\_GB2312，英文应为 Times New Roman  
乘积量化  
中文关键词不需加粗  
“Abstract”标题字体错误，应为 Cambria  
英文摘要关键词的内容字体错误，应为 Calibri                      Network Traffic  
Classification; H  
英文摘要关键词的内容字体错误，应为 Calibri                      ash codin  
英文摘要关键词的内容字体错误，应为 Calibri                      g;  
英文摘要关键词的内容字体错误，应为 Calibri                      L  
英文摘要关键词的内容字体错误，应为 Calibri                      arge-  
英文摘要关键词的内容字体错误，应为 Calibri                      scale data Nearest Neighbors r  
英文摘要关键词的内容字体错误，应为 Calibri                      etrieval  
英文摘要关键词的内容字体错误，应为 Calibri                      ;  
英文摘要关键词的内容字体错误，应为 Calibri                      Product Quantization

目录检测

目录缩进错误，应为总体缩进 2 字符    ——摘要 2  
目录缩进错误，应为总体缩进 2 字符    ——Abstract3  
目录缩进错误，应为总体缩进 2 字符    ——引言 1

|                     |                             |
|---------------------|-----------------------------|
| 目录缩进错误, 应为总体缩进 2 字符 | ——1 相关技术概述 4                |
| 目录缩进错误, 应为总体缩进 4 字符 | ——1.1 网络流量分类 4              |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.1.1 网络流 4               |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.1.2 网络流特征 4             |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.1.3 基于端口匹配 4            |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.1.4 基于深度包检测 5           |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.1.5 基于协议解析 5            |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.1.6 基于统计学习 6            |
| 目录缩进错误, 应为总体缩进 4 字符 | ——1.2 基于哈希的大数据近邻检索 6        |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.2.1 近邻检索 6              |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.2.2 基于随机投影的哈希算法 9       |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.2.3 基于学习的哈希算法 10        |
| 目录缩进错误, 应为总体缩进 6 字符 | ——1.2.4 基于乘积量化的哈希算法 10      |
| 目录缩进错误, 应为总体缩进 2 字符 | ——2 基于哈希编码的网络流量分类方法 15      |
| 目录缩进错误, 应为总体缩进 4 字符 | ——2.1 引子 15                 |
| 目录缩进错误, 应为总体缩进 4 字符 | ——2.2 相关算法 15               |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.2.1 K 均值算法 15           |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.2.2 KNN 算法 16           |
| 目录缩进错误, 应为总体缩进 4 字符 | ——2.3 算法描述 16               |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.3.1 线下训练阶段 17           |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.3.2 实时分类过程 18           |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.3.3 算法复杂度分析 19          |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.3.4 算法参数说明 19           |
| 目录缩进错误, 应为总体缩进 4 字符 | ——2.4 与其他工作不同之处 20          |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.4.1 使用汉明距离衡量网络流数据相似度 20 |
| 目录缩进错误, 应为总体缩进 6 字符 | ——2.4.2 对网络流数据进行哈希编码 21     |
| 目录缩进错误, 应为总体缩进 2 字符 | ——3 实验结果与分析 22              |
| 目录缩进错误, 应为总体缩进 4 字符 | ——3.1 实验数据集 22              |
| 目录缩进错误, 应为总体缩进 4 字符 | ——3.2 数据预处理 22              |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.2.1 数据缺失值和空值处理 23       |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.2.2 命名型属性处理 23          |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.2.3 冗余值处理 23            |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.2.4 属性重排列 23            |
| 目录缩进错误, 应为总体缩进 4 字符 | ——3.3 实验环境 23               |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.3.1 硬件环境 23             |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.3.2 软件环境 23             |
| 目录缩进错误, 应为总体缩进 4 字符 | ——3.4 实验结果 24               |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.4.1 分类评价标准 24           |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.4.2 参数分析 25             |

|                     |                          |
|---------------------|--------------------------|
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.4.3 分类性能 25          |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.4.4 和传统 KNN 算法的对比 31 |
| 目录缩进错误, 应为总体缩进 6 字符 | ——3.4.5 实验结果分析 33        |
| 目录缩进错误, 应为总体缩进 2 字符 | ——结论 34                  |
| 目录缩进错误, 应为总体缩进 2 字符 | ——参考文献 36                |
| 目录缩进错误, 应为总体缩进 2 字符 | ——致谢 40                  |

## 正文检测

|   |
|---|
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1 相关技术概述                               |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.1 网络流量分类                             |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.1.1 网络流                              |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.1.2 网络流特征                            |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.1.3 基于端口匹配                           |
| <b>正文标题缩进错误, 应为总体缩进 0 字符: ——1.1.4 基于深度包检测</b>                   |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.1.5 基于协议解析                           |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.1.6 基于统计学习                           |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.2 基于哈希的大数据近邻检索                       |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.2.1 近邻检索                             |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.2.2 基于随机投影的哈希算法                      |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.2.3 基于学习的哈希算法                        |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——1.2.4 基于乘积量化的哈希算法                      |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——2 基于哈希编码的网络流量分类方法                      |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——2.1 引子                                 |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——2.2 相关算法                               |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——2.2.1 K 均值算法                           |
| 正文段落前存在多余空格: 2.2.1 K 均值算法——随机选取 k 个质心                           |
| 正文段落前存在多余空格: 2.2.1 K 均值算法——对 n 个数据                              |
| 正文段落前存在多余空格: 2.2.1 K 均值算法——根据与质心距离划分到...                        |
| 正文段落前存在多余空格: 2.2.1 K 均值算法——对每个聚类                                |
| 正文段落前存在多余空格: 2.2.1 K 均值算法——计算每个聚类中数据均...                        |
| 正文段落前存在多余空格: 2.2.1 K 均值算法——until 质心不再变...                       |
| 正文标题缩进错误, 应为总体缩进 0 字符: ——2.2.2 KNN 算法                           |
| 正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2.2.2 KNN 算法——KNN 算法的伪代码实...  |
| 正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2.2.2 KNN 算法——计算查询数据与训练数...   |
| 正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2.2.2 KNN 算法——排序得到最小的 k 个距... |
| 正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2.2.2 KNN 算法——统计 k 个距         |

离中对应的...

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2.2.2 KNN 算法——输出出现最多次的标签...

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2.2.2 KNN 算法——影响 KNN 算法性能的...

正文标题缩进错误, 应为总体缩进 0 字符: ——2.3 算法描述

正文标题缩进错误, 应为总体缩进 0 字符: ——2.3.1 线下训练阶段

正文此段落后不应有空行 2.3.1 线下训练阶段——End

正文标题缩进错误, 应为总体缩进 0 字符: ——2.3.2 实时分类过程

正文标题缩进错误, 应为总体缩进 0 字符: ——2.3.3 算法复杂度分析

正文标题缩进错误, 应为总体缩进 0 字符: ——1 线下学习阶段

正文标题字体错误, 应为序号 Cambria, 中文黑体——1 线下学习阶段

正文标题字号错误, 应为小三——1 线下学习阶段

正文标题段前距错误, 应为 0 行——1 线下学习阶段

正文标题缩进错误, 应为总体缩进 0 字符: ——2 实时分类阶段

正文标题字体错误, 应为序号 Cambria, 中文黑体——2 实时分类阶段

正文标题字号错误, 应为小三——2 实时分类阶段

正文标题段前距错误, 应为 0 行——2 实时分类阶段

正文标题缩进错误, 应为总体缩进 0 字符: ——2.3.4 算法参数说明

正文标题缩进错误, 应为总体缩进 0 字符: ——2.4 与其他工作不同之处

正文标题缩进错误, 应为总体缩进 0 字符: ——2.4.1 使用汉明距离衡量网络流数据相似度

正文标题缩进错误, 应为总体缩进 0 字符: ——2.4.2 对网络流数据进行哈希编码

正文标题缩进错误, 应为总体缩进 0 字符: ——3 实验结果与分析

正文标题缩进错误, 应为总体缩进 0 字符: ——3.1 实验数据集

正文标题缩进错误, 应为总体缩进 0 字符: ——3.2 数据预处理

正文标题缩进错误, 应为总体缩进 0 字符: ——3.2.1 数据缺失值和空值处理

正文标题缩进错误, 应为总体缩进 0 字符: ——3.2.2 命名型属性处理

正文标题缩进错误, 应为总体缩进 0 字符: ——3.2.3 冗余值处理

正文标题缩进错误, 应为总体缩进 0 字符: ——3.2.4 属性重排列

正文标题缩进错误, 应为总体缩进 0 字符: ——3.3 实验环境

正文标题缩进错误, 应为总体缩进 0 字符: ——3.3.1 硬件环境

正文标题缩进错误, 应为总体缩进 0 字符: ——3.3.2 软件环境

正文标题缩进错误, 应为总体缩进 0 字符: ——3.4 实验结果

正文标题缩进错误, 应为总体缩进 0 字符: ——3.4.1 分类评价标准

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——分类评价术语

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——分类评价术语

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——真正 (True po...

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——真正 (True po...)

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——假正 (False p...)

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——假正 (False p...)

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——真负 (True ne...)

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——真负 (True ne...)

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——假负 (False n...)

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——假负 (False n...)

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——准确率

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——准确率

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——召回率

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——召回率

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——精度

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——精度

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——分类时间

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——分类时间

正文项目编号错误, 应为 (1) (2) 格式, 子编号应为①②格式: 3.4.1 分类评价标准——鲁棒性

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 3.4.1 分类评价标准——鲁棒性

正文标题缩进错误, 应为总体缩进 0 字符: ——3.4.2 参数分析

正文标题缩进错误, 应为总体缩进 0 字符: ——3.4.3 分类性能

正文标题缩进错误, 应为总体缩进 0 字符: ——1 分类准确率

正文标题字体错误, 应为序号 Cambria, 中文黑体——1 分类准确率

正文标题字号错误, 应为小三——1 分类准确率

正文标题段前距错误, 应为 0 行——1 分类准确率

正文此段落后不应有空行 1 分类准确率——下面我们讨论分类模型的参

正文此段落后不应有空行 1 分类准确率——以上分析的是线下训练阶段

正文标题缩进错误, 应为总体缩进 0 字符: ——2 召回率与精度

正文标题字体错误, 应为序号 Cambria, 中文黑体——2 召回率与精度

正文标题字号错误, 应为小三——2 召回率与精度

正文标题段前距错误, 应为 0 行——2 召回率与精度

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2 召回率与精度——Moore 数据集中网...

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2 召回率与精度——如表所示, 对于不同的...

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2 召回率与精度——值得注意的是, 对整体...

正文段落缩进错误, 应为左侧缩进 0 字符, 首行缩进 2 字符: 2 召回率与精度——以上召回率和精度的分...

正文标题缩进错误, 应为总体缩进 0 字符: ——3 空间性能

正文标题字体错误, 应为序号 Cambria, 中文黑体——3 空间性能

正文标题字号错误, 应为小三——3 空间性能

正文标题段前距错误, 应为 0 行——3 空间性能

正文标题缩进错误, 应为总体缩进 0 字符: ——4 分类时间

正文标题字体错误, 应为序号 Cambria, 中文黑体——4 分类时间

正文标题字号错误, 应为小三——4 分类时间

正文标题段前距错误, 应为 0 行——4 分类时间

正文标题缩进错误, 应为总体缩进 0 字符: ——3.4.4 和传统 KNN 算法的对比

正文标题缩进错误, 应为总体缩进 0 字符: ——3.4.5 实验结果分析

---

## 图检测

图与中文图名之间不应有空行 ——图 1.1 近邻检索用于图像处理

图未居中 ——图 1.2 商业识图软件

---

## 表检测

缺少英文表名, ——表 1.1 24 种常用网络流量特征

缺少英文表名, ——表 3.1 数据集流量统计表

表格内文字字号错误, 应全文五号, ——表 3.1 数据集流量统计表

缺少英文表名, ——表 3.1 数据集流量统计表

表格标题之前应空一行, ——表 3.1 数据集流量统计表

表格内文字字号错误, 应全文五号, ——表 3.1 数据集流量统计表

表格的形式应为三线表, ——表 3.1 数据集流量统计表

中文表名序号 M.N 与表名内容之间应空两格, ——表 3.2 5 个数据集交叉验证的准确率

缺少英文表名, ——表 3.2 5 个数据集交叉验证的准确率

表格内文字字号错误, 应全文五号, ——表 3.2 5 个数据集交叉验证的准确率

缺少英文表名, ——表 3.3 V=16、M=32、k=5 时网络流量召回率、精度表

表格内文字字体错误, 应全文宋体, ——表 3.3 V=16、M=32、k=5 时网络流量召回率、精度表

表格内文字应全文居中, ——表 3.3 V=16、M=32、k=5 时网络流量召回率、精度表

表格的形式应为三线表, ——表 3.3 V=16、M=32、k=5 时网络流量召回率、精度表

表格应居中, ——表 3.3 V=16、M=32、k=5 时网络流量召回率、精度表

中文表名字号错误, 应为五号——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

缺少英文表名, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

表格内文字字体错误, 应全文宋体, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

表格应居中, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

中文表名字号错误, 应为五号——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

缺少英文表名, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

表格不能跨页——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

表格标题之前应空一行, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

表格内文字字体错误, 应全文宋体, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

表格的形式应为三线表, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

表格应居中, ——表 3.4 V=16、M=32、k=50 网络流量召回率、精度表

缺少英文表名, ——表 3.5 欧式 KNN 和汉明 KNN 分类准确率对比

表格内文字字号错误, 应全文五号, ——表 3.5 欧式 KNN 和汉明 KNN 分类准确率对比

表格内文字应全文居中, ——表 3.5 欧式 KNN 和汉明 KNN 分类准确率对比

表格的形式应为三线表, ——表 3.5 欧式 KNN 和汉明 KNN 分类准确率对比

缺少英文表名, ——表 3.6 欧式 KNN 和汉明 KNN 分类时间对比

表格内文字字号错误, 应全文五号, ——表 3.6 欧式 KNN 和汉明 KNN 分类时间对比

表格的形式应为三线表, ——表 3.6 欧式 KNN 和汉明 KNN 分类时间对比

缺少英文表名, ——表 3.7 基于汉明的 KNN 分类时间分析表

表格内文字字号错误, 应全文五号, ——表 3.7 基于汉明的 KNN 分类时间分析表

---

## 页眉页脚检测

论文页眉奇数偶数页应不同

第 2 节奇数页页眉内容错误, 应为: 大连理工大学硕士学位论文

第 2 节偶数页页眉内容错误, 应为: 大学硕士毕业设计(论文)基于哈希编码的网络流量分类方法的研究

---

## 结论检测

---

## 致谢检测

---

## 参考文献检测

参考文献总量少于 20 篇

期刊类参考文献数量少于 10 篇

---

**攻读硕士学位期间发表学术论文情况检测**

---

---

**附录检测**

---

---

**学位论文版权使用授权书检测**

---

论文缺少版权使用授权书部分

---

---

**标点符号，量和单位，数字检测**

---

---