**商洛学院本科生毕业论文（设计）答辩情况登记表**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 学 院 | 电子信息与电气工程学院 | 学号 | 15033114 | 姓名 | 罗鹏娟 | 专业  班级 | 电子信息工程1501班 |
| 题 目 | 基于朴素贝叶斯算法的垃圾邮件过滤 | | | | | | |
| 指导  教师 | 王博 | 答辩时间 | | 2019.6.12 | | 答辩  地点 | 2#206 |
| 答  辩  小  组  成  员 | 姓 名 | 职 称 | | 研究方向 | | 答辩  成绩 | 成员签名 |
| 李华 | 副教授 | | 电路与系统 | | 73 |  |
| 刘萌 | 讲师 | | 嵌入式系统、 图像处理 | | 74 |  |
| 唐卫斌 | 讲师 | | 微电子 | | 76 |  |
| 答辩总评成绩（分数） | | | | | 75 | |
| 答辩记录（记录答辩中提出的主要问题、回答的要点等）：   1. 你的分类的结果没有达到100%正确，如果将正常邮件分类为垃圾邮件的话，可能会给人们带来损失，你是如何解决这个问题的？   答：运用词云。在分类结束后，为了防止将正常邮件分类为垃圾邮件，在分类结果为垃圾邮件时，通过词云突出展现邮件的关键词，词云可以过滤掉大量的文本信息。   1. 如果邮件中的词不在训练的样本中，即它的概率为0，这样的话你的计算结果也会为0，那你是如何计算这封邮件的概率的？   答：分类算法主要是计算邮件中每个词语在样本中出现的概率，在进行分类时，多个概率乘积得到类别，但是如果有一个概率为0,则最后的结果为0，因此采用了的“拉普拉斯修正”算法，在计算时，在分子加1，分母加2，再来计算这个词语的概率。   1. 利用朴素贝叶斯文本分类算法来过滤垃圾邮件的缺点是什么？   答：（1）无法适应变化的垃圾邮件特征。（2）训练成本高。由于新的垃圾邮件不断出现，垃圾邮件的特征也是经常变化的，为了能反映出这些变化，必须学习这些新的特征，使系统能够识别具有新特征的垃圾邮件。  答辩秘书（签名）：  年 月 日 | | | | | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 答辩小组评语：  该生的毕业设计，能基本正确地运用电子信息科学与技术专业所学理论和专业知识，按期完成毕业论文，有一定的独立工作能力。论文立论正确，分析合理，条理分明，文字通顺，图表、公式、符号规范，论文格式符合要求，但论文水平一般。该生在答辩过程中，表达清晰，能基本正确地回答答辩小组提出的主要问题。  答辩小组组长签字：  年 月 日 | | | |
| 指导教师  评定成绩（30%） | 评阅教师  评定成绩（20%） | 答辩总评成绩  （50%） | 毕业论文（设计）  综合成绩 |
| 78 | 78 | 75 | 77 |
| 经综合指导教师评定成绩、评阅教师评定成绩及答辩成绩，该同学的毕业论文（设计）综合等级确定为 中等 。    院答辩委员会主任（签章）：  年 月 日 | | | |