通过实证验证布拉德福定律，选择区域法，图像法，集中和分散广泛存在（二八定律、马太效应），可以选择某领域/学科/话题的相关论文和对于的期刊进行研究。

研究：

参考文献：[1]王晓芳,王健,袁广林等.基于布拉德福定律的Web被引频次分析[J].计算机科学, 2012, 39(S1): 328-330.

实现：

1.在CNKI中国知网搜索主题“数据挖掘算法”



检索到学术论文728篇，其中有被引频次的论文537篇，按被引频次递减顺序排列。

将数据导出。

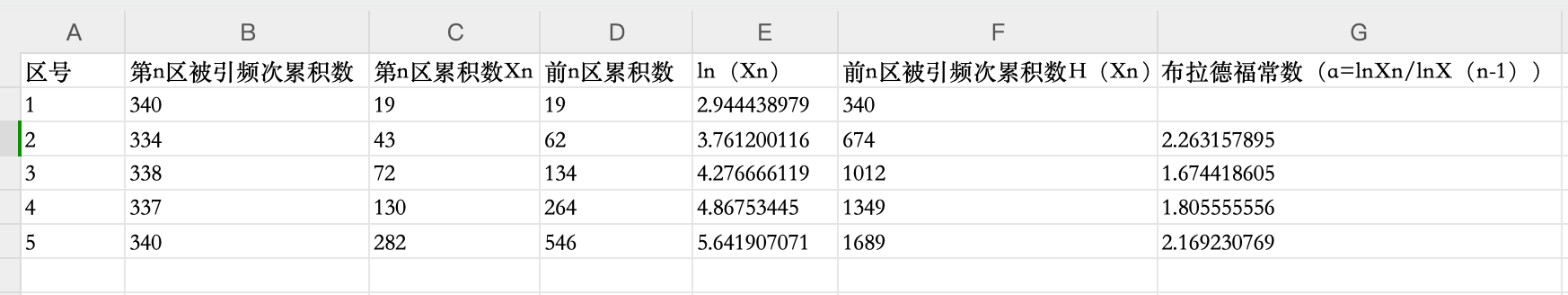
1. 关键词为“数据挖掘算法”的论文被引频次降序表



1. 区域法

设布拉德福分区数为m，当m=5时 ，把上表的数据空间分5个区，令每区被引频次累积数相等，得平均值为A／m= 1689／5=337.8，按此平均值计算各区的论文累积数Xn (n=1，2，3，4，5)，得到分区的结果，如下表所列。其中核心区划分以X(x19)=340为界限，即核心区含6篇论文，该区论文被引频次累积数是m为5时的平均值。

m=5时被引频次区域划分表



布拉德福定律指出，绝大多数的引用集中在少数核心文献上，而大部分文献只有很少的引用量。

根据给出的布拉德福常数值：

布拉德福常数为 2.263157895。这个数值较大，说明引用量较为集中。少数核心文献集中拥有较多的引用量，符合布拉德福定律的特点。

布拉德福常数为 1.674418605。这个数值较小，说明引用量相对较为分散。虽然还是存在较高排名的文献集中拥有一部分引用量，但整体上引用量分散在更多的文献中，与布拉德福定律的集中度要求略有差异。

布拉德福常数为 1.805555556。这个数值较小，说明引用量相对较为分散。引用量分散在更多的文献中，同样与布拉德福定律的集中度要求略有差异。

布拉德福常数为 2.169230769。这个数值较大，说明引用量较为集中。少数核心文献集中拥有较多的引用量，与布拉德福定律的集中度要求相符。

综上所述，根据给出的布拉德福常数值来验证布拉德福定律在数据挖掘算法领域的引用分布情况，可以发现其中有些情况符合布拉德福定律的集中度要求，即少数核心文献拥有较多的引用量，而有些情况则稍有差异，引用量分散在更多的文献中。

1. 图像法

为了验证假设即“论文累积数”与“论文的被引频次累积数”之间具有布拉德福定律描述的函数关系，参照布拉德福的做法，以“论文累积数的自然对数 ln(Xn)”为 x轴，以“论文的被引频次累积数 H(Xn )”为 y轴 ，以表 2中的数据为坐标值作散点图，得到图1。

从布拉德福分散特点上分析，此曲线具备一定的布拉德福分散特征。我们再将曲线图的横坐标变换，以“前区论文累积数 X(Xn )”为 X轴 ，以“学术论文的被引频次累积数 H (Xn )”为y轴，以表 2中的数据为坐标值作散点图，得到图2。

将图1和图2中的曲线作比较，可以发现曲线的走向相似，曲线的直线部分相似，可以认为数据具有布拉德福分布特征。