

数据仓库与数据挖掘

大作业评分标准（100 分）

一、 报告主题

针对提供的关于总书记对科技创新内容的论述，结合数据挖掘中关联规则和聚类算法，实现对于不同年份科技创新数据的知识挖掘，并基于所挖掘的知识，对科技创新方向的趋势给出深入分析。报告题目自拟。

二、 报告内容

报告内容主题明确、思路清晰、要点阐述清楚、方案可行性好，需包含各个步骤运行结果截图，针对关联规则任务使用两种及以上挖掘方法，分析不同参数对最终结果的定量和定性影响，给出结果的分析和总书记关于科技创新发展方向和趋势论述的总结性文字和可视化分析（如发展脉络图），内容充实，论述充分。

报告包含但不限于以下内容：

题目、

摘要、（5 分）

介绍（课题背景和总体研究方案等）、（5 分）

关联规则任务的数据挖掘结果及分析（共 40 分，其中使用两种及以上算法进行任务实现 20 分，分析不同参数对结果的影响 5 分，挖掘结果的可视化 10 分（可使用不同的数据图形，也可以根据自己的理解画图，要求不少于两类可视化方法），针对结果的分析 5 分）

聚类任务结果及分析、（共 20 分，根据字典构建向量 5 分，根据聚类算法得出结果 5 分，结果可视化 5 分，结果及参数分析（定量和定性）5 分）

总体任务结果及分析（共 15 分，总结两类任务所挖掘知识对科技创新方向的脉络梳理图形 10 分、结果分析 5 分）、

开放性提出及思考（10 分）（如讨论数据、挖掘、可视化等中存在的问题及未来的解决对策）

结论、（5 分）

注：排版不符合要求扣 5 分，报告中新增以上未提到的创新性思考内容加分（5-20 分不等）。**不局限于课内介绍的关联规则和聚类方法，可以根据对象的特性通过分析选择或构造创新性的方法或规则评估指标。**

三、 具体要求

1. 报告结构至少包括以下部分
 - a) 题目

题目应主题明确，体现数据挖掘技术与具体领域的结合与应用，应是整个报告总体内容的体现，要引人注目，力求简短，严格控制在 25 字以内。

b) 摘要

摘要含内容摘要、关键词两部分。摘要应说明研究目的、方法、结果和结论等，字数为 300-600 字。中文关键词紧随其后，另起一段。关键词的数量为 3-8 个。中文摘要原则上不能超过一页。（不需要英文摘要）

c) 目录

目录应列至三级标题，包含正文及其后的各部分，并附有相应页码。目录的文字应与相应标题文字完全一致。

d) 正文（按照要求格式排版后，正文不少于 15 页）

正文的内容一般包括：课题介绍（背景、意义、动机等）、相关工作、设计的方案、预期结果或实验结果（不具备实验条件应做可行性分析）、性能分析、结果讨论、结论等。请同学们根据报告内容要求部分自行设计。

正文中的图、表应具有“自名性”。图要精选、简明，切忌与表及文字表述重复。图中的术语、符号、单位等应同文字表述一致。表中参数应标明量和单位的符号。表的编排建议采用国际通行的三线表。公式要标准、通用。

e) 参考文献

参考文献一律放在论文参考文献部分中，不应放在章节末尾或页面脚注中。

2. 报告内容排版格式：

a) 排版格式整体应整洁美观，内容充实，图表清晰，编排有序，正文各部分占比均衡。

b) 封面

封面采用统一封面。

c) 纸张和页面

采用 A4 纸张（大小为 210×297mm）。页边距为：上：27.5mm；下 25.4mm；左：35.7mm；右：27.7mm。页眉距边界 15.0mm；页脚距边界 17.5mm。如果纸质版报告，要求双面打印，每一章从另页右页（奇数页）开始。

d) 报告内容排版格式以《天津大学关于本科生学位论文统一格式的规定》中的要求为准，包括摘要、目录、字体、段落、行距、多级标题、图标公式、参考文献等。

e) 页眉和页脚

页眉从正文开始到论文结尾，页眉内容为“天津大学数据挖掘基础课程报告”。页眉的字号为五号，居中。其余样式与正文相同。

页脚的字号为小五号，居中。从中文摘要开始，页脚从 I 开始顺序编页码，为大写罗马数字格式。正文第一页开始，用阿拉伯数字重新从 1 开始顺序编

页码，至报告结尾。

天津大学

《数据仓库与数据挖掘》课程报告



题目名称

学 号_____

姓 名_____

邮 箱_____

学 院_____

专 业_____

年 级_____

任课教师_____

年 月 日