**《信息论基础》第五次上机作业**

--科学进步是站在巨人的肩膀上的吗？

牛顿有一句经典名言，“如果我想要看得更远，那是因为我站在巨人的肩膀上”。这一名言反映了一个观点：科学的进步主要依赖于少数伟大的科学家（巨人）的工作，后被称为牛顿假说。然而，Ortega等学者持相反的观点，认为在学术界，促进科学进步和发展的主要力量是科研工作者中的大众群体，平庸的科学家为科学进步提供了支撑作用。

**实验目标**：请用分别用如下两组数据集，思考如何识别论文参考文献中的巨人，并分析科学的进步是否是站在了巨人的肩膀上。最终结果提交一份分析报告，根据数据分析的结果为科研管理部门提供政策建议。

1. **诺奖论文数据集**

* 文件np\_physics.csv 包含了物理学领域108篇诺贝尔奖获奖论文的详细信息，id字段表示论文的编号，year表示论文的发表年份，citeN 表示论文的截止到2020年的被引量，refN表示论文参考文献数量，Prize year表示论文的获奖年份。
* 文件np\_physics\_ref.csv包含了108篇诺奖论文参考文献的信息，其中id字段表示诺奖论文的id，ref表示参考文献的id， ref\_year表示参考文献的发表年份，citeN5、citeN10、citeN30分表表示参考文献在5年、10年和30年引文时间窗下的被引次数。

1. **APS论文数据集**

美国物理协会（American Physical Society， APS）旗下有19本期刊，这些期刊全部是物理专业的核心期刊，在物理学界及相关领域具有较高的学术权威。期刊的名称及缩写如表1所示。

表1 APS 期刊信息

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| id | Full name | Abbr. | id | Full name | Abbr. |
| 1 | Physical Review | PR | 11 | Physical Review Letters | PRL |
| 2 | Physical Review A | PRA | 12 | Physical Review Materials | PRMATERIALS |
| 3 | Physical Review Accelerators and Beams | PRAB | 13 | Physical Review Physics Education Research | PRPER |
| 4 | Physical Review Applied | PRAPPLIED | 14 | Physical Review Research | PRRESEARCH |
| 5 | Physical Review B | PRB | 15 | Physical Review Special Topics - Accelerators and Beams | PRSTAB |
| 6 | Physical Review C | PRC | 16 | Physical Review Special Topics - Physics Education Research | PRSTPER |
| 7 | Physical Review D | PRD | 17 | Physical Review X | PRX |
| 8 | Physical Review E | PRE | 18 | PRX Quantum | PRXQUANTUM |
| 9 | Physical Review Fluids | PRFLUIDS | 19 | Reviews of Modern Physics | RMP |
| 10 | Physical Review (Series I) | PRI |  |  |  |

文件APS\_1990.csv是1990年APS期刊发表的所有论文。字段doi表示论文的DOI号，title表示论文的标题，year表示论文的发表年份，journal表示论文发表的期刊， volume和issue分别表示论文的卷、期。

请自行获取所有论文引用的参考文献，分析论文的参考文献的特征，例如引用了多少高质量论文，分析科学的进步是否是站在巨人的肩膀上。

请于当天**18:00**前将**PDF文档**发放到邮箱： informatics\_course@163.com。请将文件命名为:学号\_姓名\_HW5。