**Final Examination of Big Data Processing**

**you will work on your own Big Data project. You will work on a large data set and define a Big Data Analytics application.**

**所提交的论文或者报告均会要上传到知网进行学术不端检测**

**1. 任选一种主题（可以参考后面的建议选题），基于自己的研究与实现，写一篇综述性的论文/报告。有原理，有算法应用，有比较分析以及结论。**

**2. 任选一种主题（可以参考后面的建议选题），基于自己的研究与实现，写一篇应用研究性的论文/报告。**

**3. 也可以选择几个主题的组合，基于自己的研究与实现，写论文/报告。**

**评分标准：**

**(1)摘要与关键词叙述:5分**

**(2)贴题程度:20分**

**(3)论文整体层次与结构:10分**

**(4)算法描述与应用程度:10分**

**(5)图表符号的规范性:5分**

**(6)结果分析方法:10分**

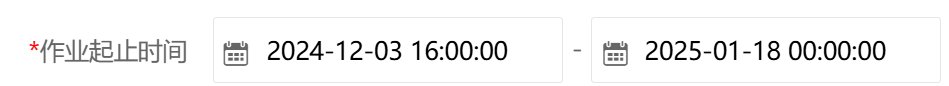
**(7)语言文字的表达能力:10分**

**(8)逻辑推理的合理性:10分**

**(9)实验或分析数据的可靠性与正确性:10分**

**(10)参考文献的新颖性:10分**

**At least 3000 word!**

****

**Deadline : before weekend of the 20th week**

[Collected Advice on Research and Writing (cmu.edu)](https://www.cs.cmu.edu/afs/cs.cmu.edu/user/mleone/web/how-to.html)

[如何写出A+Report？超实用教程+高分模板 - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/555040888)

[How to Write a Report: A Guide With Examples | Grammarly Blog](https://www.grammarly.com/blog/how-to-write-a-report/)

[How to Write a Research Paper | A Beginner's Guide (scribbr.com)](https://www.scribbr.com/category/research-paper/)

**1. Choose any topic (you can refer to the suggested topic below), based on your own research and implementation, write a comprehensive paper/report. There are principles, algorithm applications, comparative analysis, and conclusions.**

**2. Choose any topic (you can refer to the suggested topic below) and write an applied research paper/report based on your own research and implementation.**

**3. You can also choose a combination of several topics and write papers/reports based on your own research and implementation.**

**Grading:**

**(1) Abstract and keyword description: 5 points**

**(2) Posting level: 20 points**

**(3) The overall level and structure of the paper: 10 points**

**(4) Algorithm description and application level: 10 points**

**(5) Standardization of chart symbols: 5 points**

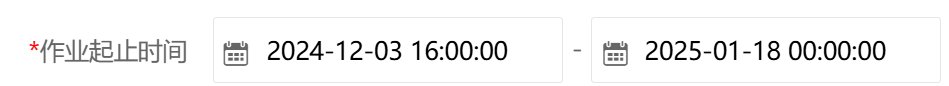
**(6) Result analysis method: 10 points**

**(7) Language expression ability: 10 points**

**(8) Reasonability of logical reasoning: 10 points**

**(9) Reliability and correctness of experimental or analytical data: 10 points**

**(10) Novelty of references: 10 points**

****

**At least 3000 word!**

[登录 (cnki.net)](https://tscheck.cnki.net/cm/)

**Deadline : before weekend of the 20th week**

[Collected Advice on Research and Writing (cmu.edu)](https://www.cs.cmu.edu/afs/cs.cmu.edu/user/mleone/web/how-to.html)

[如何写出A+Report？超实用教程+高分模板 - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/555040888)

[How to Write a Report: A Guide With Examples | Grammarly Blog](https://www.grammarly.com/blog/how-to-write-a-report/)

[How to Write a Research Paper | A Beginner's Guide (scribbr.com)](https://www.scribbr.com/category/research-paper/)

Some suggested topics:

1. Traffic control using Big Data
2. Web Application for Controlling Robotic Car Using IOT cloud
3. Building an Activity Planning Tool on the Cloud
4. Enhanced Cloud Based Smart Security System
5. Cloud Based Smart Storage
6. A Scalable Graph Based Ranking Model for Content Based Image Retrieval
7. Autonomous Vehicle Cloud Data Storage & DNN Training
8. Topics-Based Sentiment Analysis using GCP
9. Medical Insurance Fraud Detection
10. Crime Detection
11. Social Community Event Application
12. Recommendation System
13. Social Media Sentiment Analysis Tool
14. Custom Detection System
15. Disease Prediction Based on Symptoms
16. Anomaly Detection in Cloud Servers
17. Smart Cities Using Big Data
18. Tourist Behavior Analysis

## Reference Books

* [Hadoop: The definitive guide](https://item.jd.com/11789217.html) by Tom White.
* [HBase: The definitive guide](https://item.jd.com/11321037.html) by Las George.
* [Learning Spark: Lightning-fast Data Analysis](https://item.jd.com/11782888.html) by Holden Karau et al.
* [Advanced Analysis with Spark](https://item.jd.com/12371964.html) by Sandy Ryza et al.
* [MongoDB: The definitive guide](https://item.jd.com/11384782.html) by Kristina Chodorow.
* [Cassandra: The definitive guide](https://item.jd.com/12263427.html) by Jeff Carpenter et al.
* [ZooKeeper: Distributed Process Coordination](https://item.jd.com/11384782.html) by Flavio Junqueira et al.
* [Programming Hive](https://item.jd.com/11363739.html) by Edward Capriolo et al.
* [Data Algrithms: Recipes for Scaling Up with Hadoop and Spark](https://item.jd.com/11993447.html) by Mahmoud Parsian.