

Re: need feedback

From: Ullah, Amjad A.Ullah@napier.ac.uk
To: Faisal, Muhammad 40583155@live.napier.ac.uk
Cc: Thomson, Sarah S.Thomson4@napier.ac.uk
Date: Fri, 17 Nov 2023, 5:08 pm

Hi Muhammad Faisal,

I went through your report. My review notes are as follows:

Introduction: It is clear what you would like to do but should have included more specific examples where you can demonstrate how big the data duplication problem is and for what specific systems it is a big issue

Literature review: This is very weak. The main reasons include:

1. There is a lot of repetition.
2. You described things from a very generic viewpoint. Make things specific to your project. For example, what data duplication techniques you are using? And how it could be beneficial and for what kind of data.
3. How and why hashing is used as a data duplication method? What are the alternative techniques for data duplication?
4. The challenges in general are good. However, which one is important to your project? Which one you tried to solve? What are the challenges related to your project?
5. Most importantly, it is still not clear to me, what is your related work. You discussed various works that worked on data de-duplication, but none of those works are based on hashing techniques. The focus should be on finding and discussing work that is related to you. Even for these works, it isn't clear, how these are related to your project. Or how your work is different from theirs.

Research methodology: This chapter includes many details again but the relevance of most of it to your project isn't clear. For example, what are the specific details about data collection, dataset selection, data processing, data recording, etc? The same goes for the evaluation attributes and ratings. This is, however, you said, didn't yet address.

Analysis and Discussion: This chapter discusses the result first and the implementation details are provided after that. It would have been better to have a separate chapter about the design and implementation, which comes before the Analysis and discussion. My comments for each

one are as follows:

- **Design and Implementation:** Discuss the overall design of your system and explain the different choices you made like the adopted architecture. Then discuss your implementation, the technologies you have used, etc.
- **Analysis and Discussion:** First explain your experiment, the objective of experimentation and how you are going to execute it. Currently, you straight went to the presentation of the results. It is also not clear what are the results and how these are obtained. and how these are relevant to the project. Or how and why these are important? I don't see any improvement in the use of a systematic academic approach for the presentation and discussion of these results.
- The details on the selection of CRC32b include a lot of claims but it isn't clear how these claims are made. E.g. from literature, if yes, where are the references? If by experimentation, where are the results?
- Lastly, no details are found about some of the objectives from Chapter 1, e.g. Comparative analysis and real-world applicability.

References: I couldn't find several references from the list. This raises serious concerns about using AI tools for the generation of content.

In addition to these comments please consider the comments from the last Viva report. Good luck

Br,

Dr Amjad Ullah, FHEA

Lecturer in Software Engineering

School of Computing, Engineering & the Built Environment

Edinburgh Napier University

<https://amjad-ullah.netlify.app/>

On 12 Nov 2023, at 21:19, Faisal, Muhammad <40583155@live.napier.ac.uk> wrote:

Hello, I hope you are doing well. Please have a look at the attached dissertation and provide your feedback. Please note that the evaluation criteria still needs to be improved and I am currently working on that rating and it's screenshots to justify that but other than that is it okay? Could you please confirm?

Thanks!

Muhammad Faisal

40583155

<Dissertation-Final.docx>
