Computing at the edge

Deliverable 1: Final year Dissertation

Bsc Computer Science: Artificial Intelligence

Sam Fay-Hunt — sf52@hw.ac.uk

Supervisor: Rob Stewart — R.Stewart@hw.ac.uk

October 7, 2020

DECLARATION

I, Sam Fay-Hunt confirm that this work submitted for assessment is my own and is expressed in my own words. Any uses made within it of the works of other authors in any form (e.g., ideas, equations, figures, text, tables, programs) are properly acknowledged at any point of their use. A list of the references employed is included.

Signed: .....

Date: .....

Abstract: a short description of the project and the main work to be carried out – probably between one and two hundred words

Contents

1 Introduction

Summarising the context of the dissertation project, stating the aim and objectives of the project, identifying the problems to be solved to achieve the objectives, and sketching the organisation of the dissertation.

Edge devices have never been cheaper *citation*, stuff about how IoT devices are ubiquitous

Mention how there is an increasing trend to perform computing at the Edge - real time applications + privacy

These devices are often equipped with some form of AI application: Photo enhancment ect.

Online vs offline learning

Edge-side inference

These models can have a huge number of parameters so inference can sometimes be impractical. [1] - see Table 1

Issues with limited resource computation

outline the document: We start with ..., then we cover x, y, and z ...

This dissertation is an investigation into the effect of pruning on inference latency and accuracy with hardware without optimisations for processing sparse matrices.

1.1 Background

Discussing related work found in the technical literature and its relevance to your project. **Computing at the edge**: Some background on edge computing - maybe a detailed definition **Deep learning**: Types of deep learning & inference **compression types:**: pruning distillation

Quantization

2 Research Methodology/ Requirements Analysis

2.1 Research Methodology

This is required for research projects and should be linked back to the project aim and objectives. It should describe the research methods that will be employed in the project and the research questions that will be investigated.

2.2 Requirements Analysis

This is required for technical projects and should be linked back to the project aim and objectives. It should provide a detailed use case scenario and suitable use

3 Design

4 Evaluation Strategy

5 Project Management

5.1 Timetable

5.2 Risk Analysis

5.3 Professional, Legal & Ethical issues

6 Back matter

6.1 References

References

[1] Y. Chen, B. Zheng, Z. Zhang, Q. Wang, C. Shen, and Q. Zhang, "Deep Learning on Mobile and Embedded Devices: State-of-the-art, Challenges, and Future Directions," *ACM Computing Surveys*, vol. 53, no. 4, pp. 1–37, Sep. 26, 2020, ISSN: 0360-0300, 1557-7341. DOI: 10.1145/3398209. [Online]. Available: https://dl.acm.org/doi/10.1145/3398209 (visited on 10/01/2020).

6.2 Appendices

to include additional material, consult with your supervisor.