Assignment: Sustainable AI

This assignment is designed to help you critically reflect on the challenges involved in the design of an AI system through a case study. Your objective is to not only treat the functionality of the AI system, but also demonstrate awareness of the broader sustainability implications of your choices through a dedicated sustainability assessment.

Case study: AI for Wildlife Monitoring

A conservation NGO working in the rainforests of Borneo wants to monitor endangered orangutans and sun bears. Poaching and deforestation are rapidly reducing their populations, and researchers need accurate data to guide conservation strategies. They propose an AI system that can automatically identify species in camera trap images. However, resources are limited, and powering equipment deep in the forest is difficult.

Your objective is to design an AI system to support the NGO in its mission, and assess its environmental footprint. You do not need to actually train any models! Your main deliverable is a max. 6 page report (excluding title page, table of contents, appendix, bibliography) that follows the LaTeX template attached separately. You may slightly deviate from the given template as long as you successfully fulfill all tasks in the assignment.

Part 1: System Design

The following tasks need to be performed in order to complete the system design. For each task, we set up a few pointers to help you, but you have a lot of freedom in your design.

1. **Select training data source:** there are few labeled images of these rare animals. Should the team use transfer learning, or crowdsource labeling from volunteers? Should they augment the dataset with simulated/augmented images? Justify your decision.
2. **Design training strategy:** running deep neural networks on cloud GPUs is possible, but the NGO is concerned about costs. Would training on smaller, lightweight models sacrifice too much accuracy? Justify your decision, especially in light of the training data source you selected.
3. **Design deployment strategy:** what hardware will you deploy on? Should inference run directly on edge devices in the forest (low energy use, but computational limits), or should raw data be transmitted to a central server (better compute, but requires high bandwidth and electricity)? How do you handle the balance between real-time (if needed) detection and sustainability? Make sure to account for the operational conditions, and justify your design by trading off the prospective options.

Part 2: Environmental Impact Assessment

After the design of your AI system, your objective is to assess its environmental impact. The following tasks need to be performed. For each task, we set up a few pointers to help you, but you have a lot of freedom in your assessment. Make sure to clearly document your assumptions at each point.

1. **Define system boundary:** which parts of your system are included in or excluded from your assessment? You can freely decide: do you take into account upstream processes like the manufacturing of your infrastructure (cameras, servers)? What time period does your assessment cover? Justify your decision.
2. **Estimate system energy consumption:** what is the size of your model? Where are the computations being performed? What are its energy costs for training and inference? We do not expect exact results, estimates based on literature are sufficient.
3. **Estimate system carbon footprint:** you can adapt one of the methodologies presented in the course.
4. **Estimate system water footprint:** you can adapt one of the methodologies presented in the course. When estimating the water footprint of a data center, you can use average yearly values for PUE/WUE - though more detail is appreciated!
5. **Discuss system material footprint:** we do not expect a complete quantitative assessment. Instead, briefly explain how your design decisions (for example, centralized or decentralized deployment) qualitatively influence the material footprint of your system.
6. **Perform sensitivity analysis:** there is inherent uncertainty in assessing environmental impact, for example in the assumptions you have made until now. Quantify the effect of these assumptions through a simple sensitivity analysis on at least two parameters (for example, the carbon intensity of the electricity consumed by the system). Justify your choice of parameters.

Tip: don't be afraid to make assumptions! This is a developing (and secretive) field, so data isn't exactly abundant. The accuracy of the actual numbers isn't as important as reaching reasonable conclusions.

Part 3: Reflection

With your design and environmental impact assessment complete, it is time to reflect on the process. What limitations did you observe in your system design process? After the environmental impact assessment, do you think that you made the right choices for your design, or would a different option have been more sustainable? Would cost considerations have impacted your decision-making? Did the sensitivity analysis support your conclusions, or refute them? These are a few suggestions for some aspects you could address.

Note on large language models: you may use LLMs for coding or writing support, or to explore a certain topic. However, LLMs are not considered valid references. Always use trustworthy sources for your claims and be critical about the information you obtain.

Grading Scheme: Sustainable AI

Part 1: System Design

Part 2: Environmental Impact Assessment

1. Report clearly defines spatial and temporal system boundary **AND** embedded non-operational emissions are either quantitatively assessed **OR** receive a comprehensive qualitative discussion **OR** their exclusion is well-justified.
2. Report provides a quantitative estimate for the energy consumption of the system **AND** explains the method by which it was obtained sufficiently so as to be independently replicated **AND** documents all assumptions made in the process.
3. Report provides a quantitative estimate for the carbon footprint of the system **AND** explains the method by which it was obtained sufficiently so as to be independently replicated **AND** documents all assumptions made in the process.
4. Report provides a quantitative estimate for the water footprint of the system **AND** explains the method by which it was obtained sufficiently so as to be independently replicated **AND** documents all assumptions made in the process. **BONUS:** estimate has detailed temporal resolution, water consumption is calculated e.g. per season.
5. Report contains a reasoned discussion of the impact of previous design decisions on the material footprint of the system. **BONUS:** quantitative figures are given.
6. Report provides clear sensitivity analysis results for two parameters **AND** explains the method by which they were obtained sufficiently so as to be independently replicated **AND** contains a solid justification for the selection of parameters. **BONUS:** more than two parameters evaluated.

Part 3: Reflection

Reflection demonstrates critical attitude towards design process **AND** includes quantitative figures where possible.