IDS 705 Final Report

Project Title

Subtitle [as applicable]

**Team Members**:

Eleanor Jiang

Chao Peter Yang

Kaisen Yao

Uzoma Uwazurike

[in alphabetical order by last name]

**Team Identifier**: ####

# Abstract

[1 paragraph maximum] This should be the one paragraph that captures the significance of what you did and why you did it - this should be a summary of the work and your outcomes in brief.

# Introduction

Provide a description of the problem and the value in finding a solution, motivate your reader as to why they should care about your problem or question.

# Background

This section should cite work that has been previously addressed that relates to your project, and the key takeaways of those studies/projects.

# Data

Describe and visualize your data in the context of the problem you are working on.

# Experiments

Present your machine learning experiments (for supervised learning, a description of any preprocessing, feature extraction, classification/regression techniques, experimental designs and evaluation criteria) and why you made each of the choices you did to achieve your goal. Cite relevant literature to support your claims and any work that is not your own. Also include a flow chart of your methodology so the reader can easily conceptualize your solution. The flow chart of the overall experimental design should clearly articulate your process ([example](https://www.nature.com/articles/sdata2016106/figures/1)). Additionally, for multiple experimental conditions or applications, they should each be represented in your flowchart. Describe your approach to measuring generalization performance, what metric(s) you used and why.

# Results

Include a complete performance assessment that includes your validation approach (cross validation, train/validate/test split, etc.) and the key metrics of performance for the problem (ROC curves, PR curves, confusion matrices if applicable, etc.). You should also compare your outcomes to at least one baseline model to act as a point of reference for interpreting the results of your work as well as chance performance (i.e. random guessing for classification, guessing the mean/median for regression). This section should be supported with visualizations including examples where your method worked well and where it failed, when possible, and hypotheses supported by evidence as to why in each case.

Ethical Considerations

Discuss any ethical aspects or implications of your project. What potential sources of bias have you encountered? What (if any) potential privacy concerns are associated with your project? Is there any potential for negative societal impact from the work (who benefits, who could be harmed, and what steps could you take to minimize negative consequences). Are there any issues around transparency and explainability for your work (e.g. is that important for your application) and if so, how are you planning to address them?

# Conclusions

It is critical to have a strong ending and not just let the energy fizzle out of the report. Many readers, if pressed for time, will simply read your abstract and your conclusions. In fact, you may want to start by writing your conclusions. Very succinctly recap the problem you were studying and what was your approach to the solution. Focus on explaining the key takeaways from your work - these should not be merely a set of bullet points, but fleshed out conclusions. As you're writing your conclusions, think about if the reader took nothing else away from reading your report, what would you want them to know most? Did you identify one particular approach that worked well? Was there a challenge that you faced that opens the door to working on solving a new problem? What avenues of research would you pursue next?

# Roles

Since this is a team project, we want to know what your specific contribution was to this project. Provide detail on your individual role. Each team member should clearly articulate an individual role.

# References

An alphabetical list of references cited in this work. A minimum of 15 are required (a minimum of 10 must be technical papers/reports or conference papers, rather than blogs or websites). Consider using the Zotero citation manager for collecting and compiling your references. These should primarily be research papers and technical reports.

Citations should follow a known model (MLA, IEEE, Chicago, APA, etc.). Citations SHOULD NOT JUST BE HYPERLINKS. When in doubt, use MLA.

# Appendix [optional]

You may have unlimited appendices of supporting materials, although they will not be directly evaluated as part of the report itself.

# Other information

*Note: this section is for information purposes and should* ***not*** *be included in your report (-2 points if included)*

## How to use this template

1. Click File > Make a copy and delete the instructions replacing text with your report content
2. Follow the formatting instructions contained in this document
3. Share your completed version of the file giving general access to “anyone with the link” and at least “commenter” privileges so that we can add comments and suggestions directly to the text

## Length requirements

Your report should be no longer than 2,500 words, not including references and roles. You may have an unlimited number of figures and tables, and those do not count towards the word count.

## Headers

The main sections of the report should have the “Heading 1” style applied (as “Formatting” is shown here). Subsections, like this one, should have the “Heading 2” style applied and so on. This will allow the report to be

## Tables

Tables should have captions before the start of the table, and, like Table 1 below, should be referenced in the text.

Table 1. Experimental conditions investigated in this study. In this case, each experimental condition is listed in each row with the columns discussing the factors that differ across experiments

| **Experimental Conditions** | **Factor 1 (e.g. hyperparameters, model types, training datasets, etc.)** | **Factor 2** | **Factor 3** |
| --- | --- | --- | --- |
| Condition 1 |  |  |  |
| Condition 2 |  |  |  |
| Condition 3 |  |  |  |

## Figures

Figures are highly encouraged, and should each be referenced in the text (such that every figure has a clear point to the story that you tell). Every figure should have a caption, figure number, axis labels (with units if applicable), and legend or direct labeling of data (as shown in Figure 1), if applicable. If you use any figures that are not your own, they should be cited as well. No figure should be superfluous - every figure should be referenced in the text. Figure 1, below, shows an example of a clear figure with a descriptive caption.

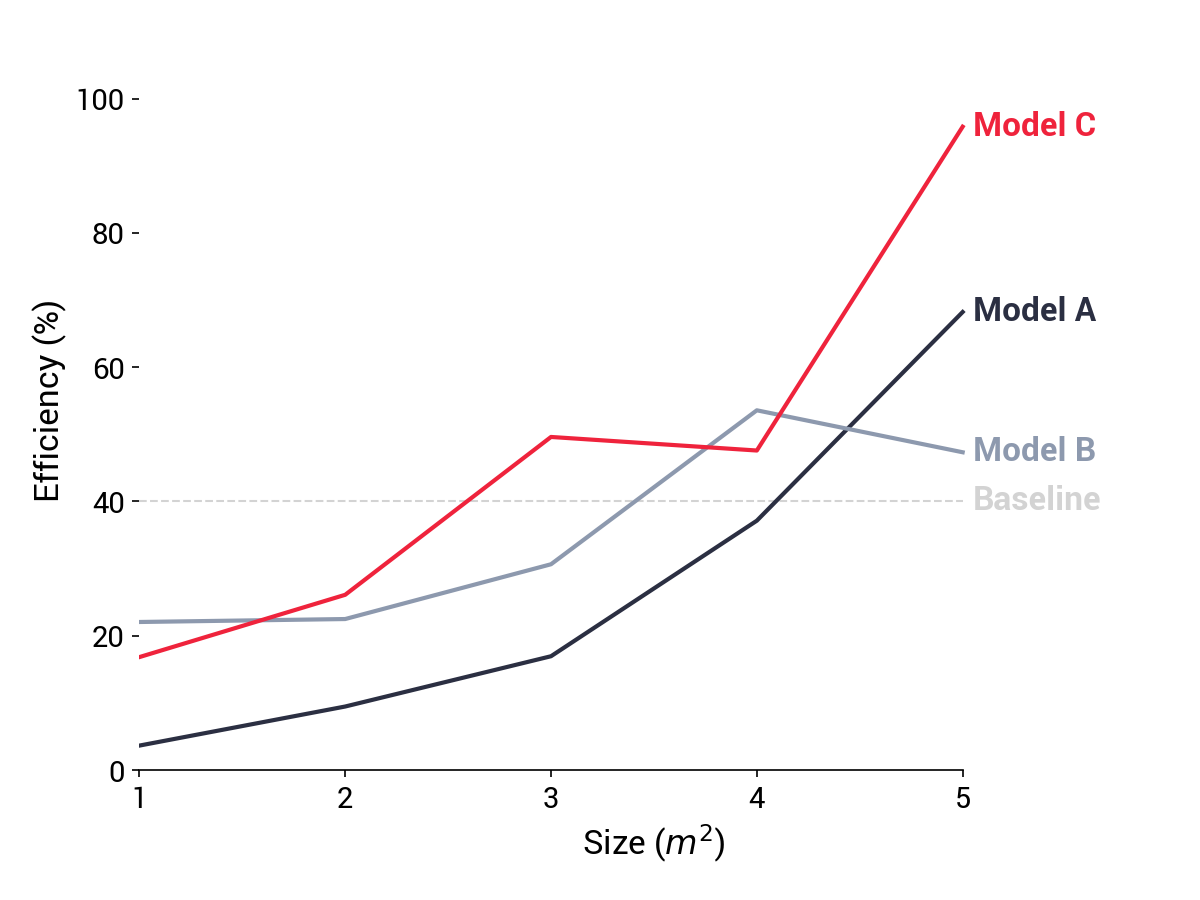


Figure 1. Experimental results from showing that Model C tends to outperform other models compared for the same training dataset.

## Flowcharts

Every report should have a flowchart as described above. Some examples of this include:

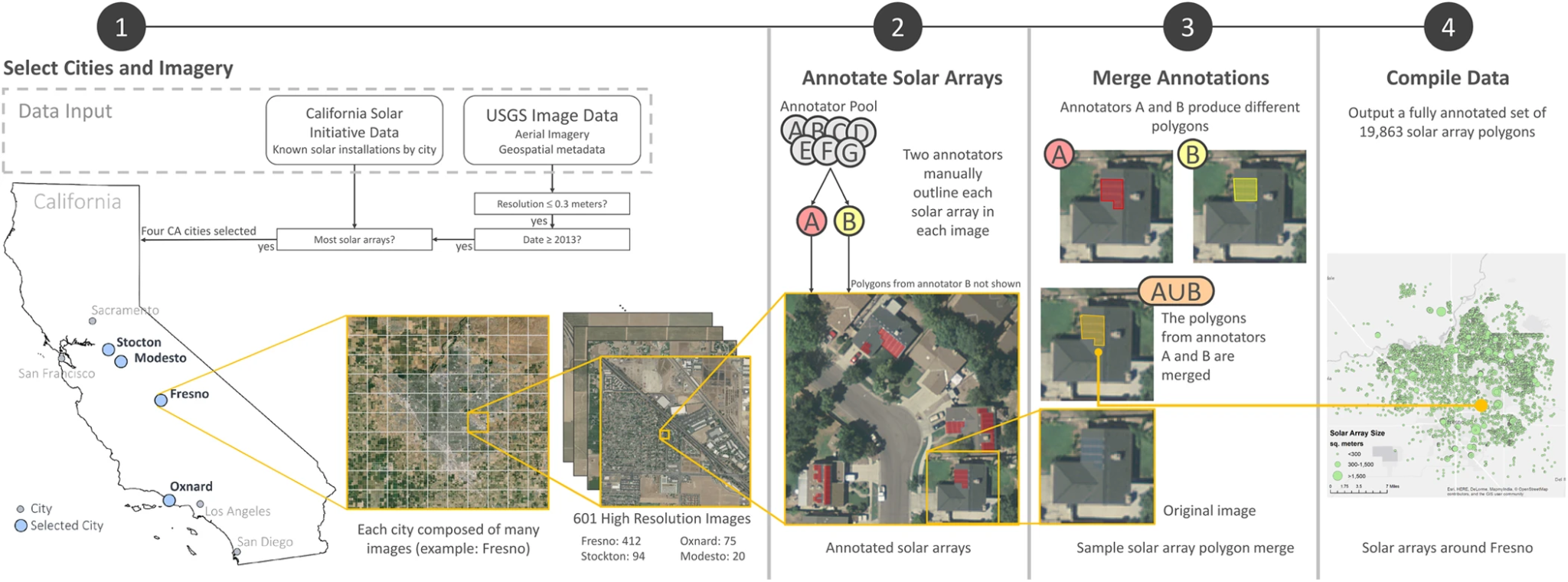


Figure 2. Example flowchart from Bradbury et al. 2016.