WILDCARD WEIGHTING

PLACE NAMES OF GROUP MEMBERS HERE

Abstract. Place abstract here. The abstract summarizes in one paragraph the main question and conclusions draw from your investigation.

1. Research question and overview of the data

The quality of the research question(s) you are asking plays a big role in how good the entire project is. Make a clear case for why your question is interesting, well thought out, precisely formulated, and answerable, at least in principle, with adequate data and the techniques of machine learning.

Briefly review what is already known about your research questions and what techniques others have used to study these questions. The best written reports include references to prior work. Explain the data set, before analysis. Form a thoughtful hypothesis or hypotheses about the data. Answer the following questions and any others that may be relevant to your question and your data set:

- What weaknesses or problems does the data set have?
- Why is this a good choice of data set to answer your research questions (as opposed to other similar data sets)?
- What do you expect your analysis to reveal?
- What other interesting questions will analyzing this data answer?

Also reference articles and sources [**?**] that are relevant or that you used when learning and/or thinking about your project. You should also reference prior work that has considered similar questions.

2. Data Cleaning / Feature Engineering

Tell us what you did when you were cleaning your data and engineering features. Why did you make the choices that you did? What are the consequences of those choices?

3. Data Visualization and Basic Analysis

Analyze the data, draw conclusions, and effectively communicate your main observations and results.

- Calculate appropriate summary statistics.

Date: 4 October 2024.

- Use appropriate plotting techniques, visualizations, and other tools and techniques you have learned, to thoughtfully identify and evaluate what the data are telling you,how well suited the data are to answering your problem,
- Reference figures and plots, like Figure **??**.

4. Learning Algorithms and In-depth Analysis

Analyze the data using the machine learning techniques discussed in class. Explain what research questions you can answer using the machine learning techniques presented this semester, and if applicable, what you think you may be able to answer next semester.

Be able to explain the results of your analysis, whether the results are meaningful, and why you chose the tools that you used.

5. Ethical Implications and Conclusions

Thoughtfully analyze the ethical implications of your research questions, the data you gathered, and the analysis that was performed. Are there privacy or other implications from the collection or use of the data? Could your results and methods be misused or misunderstood? What can and should be done to prevent misuse and misunderstanding? Could your algorithms and methods result in a destructive self-fulfilling feedback loop? How could that be prevented or controlled? What other ethical implications does your work have?

This part should all be done before you get to *page 5*. The bibliography can spill on to page 6, but we won't read text that goes past page 5.