A TEMPLATE FOR THE *arxiv* STYLE

A PREPRINT

Shriram Varadarajan
Department of Elecrical Computer engineering
University of Iowa
Iowa city, IA, 52242
shriram-varadarajan@uiowa.edu

Jacob Nishimura
Department of Elecrical Computer engineering
University of Iowa
Iowa city, IA, 52242
jacob-nishimura@uiowa.edu

Joseph Kim
Department of Biomedical Engineering
University of Iowa
Iowa city, IA, 52242
joseph-kim@uiowa.edu

March 16, 2021

ABSTRACT

This is our research paper

Keywords First keyword · Second keyword · More

1 Introduction

- Overview of the problem and the relevant background knowledge must be described in a section titled "Introduction".
- The introduction section should be around 1-1.5 pages, and certainly no longer than 2 pages using the submission template above.
- Your audience should be fellow graduate students/faculty in a different discipline. Hence, you must educate them with an easy-to-understand language, so that the readers are ready to digest the remainder of your project description. You can assume the audience has a basic background knowledge on the topics such as AI/ML/DL.
- Related works should be referenced, so that the readers can have some historical context ("what other people did/do"). However, the introduction section should not be too technical or jargony.

2 Problem Definition

If the intro section was a place for mostly lay-person's description of your project, "Problem Definition" section is where you can use technical terms to define your problem more precisely.
Be very precise and explicit about input-output parameters to your machine learning problem. For example, "I will make a machine learning model that predicts house price" is not a good problem statement. Instead, be more specific about what goes into your model and what will come out of your model. For example, "The model will take a color photograph (RGB image) of a house resized to 224-by-224 alongside other metadata including 'build year,' 'days on market,' 'square footage,' and 'school district,' and predict the dollar amount (normalized in range [0,1]) of the actual market price of the house as an output" is a better way to state your problem. If you have too many parameters to be listed in one sentence, creating a table listing inputs and outputs, as well as their data types (e.g. color image, grayscale image, time-series, scalar, string, ...) would be a great idea. The problem definition section should be around 0.5 - 1 page, but certainly no more than that.

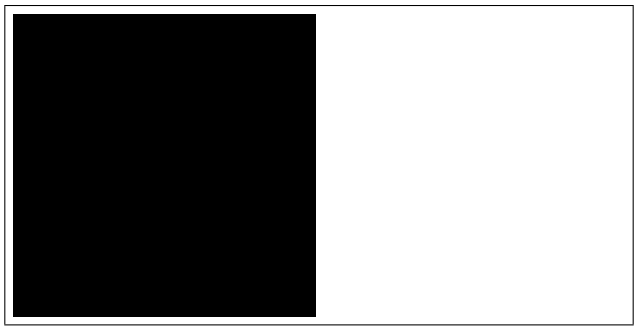

Figure 1: Sample figure caption.

Table 1: Ratings

userId	movieId	Rating	Timestamp
4	5	4	23:12
3	1	5	22:22
1	3	3	11:11

See Section 2.

2.1 Headings: second level

2.1.1 Headings: third level

Paragraph

3 Data

- What is available/not available in the data set (in conjunction with your input-output description in Problem Definition)
  Data that are available in the dataset are the movie ratings and users with the movie tags. The data that is not available are the users' text reviews for further word processing and text similarity. The movies have associated genres and titles which can be used for text preprocessing.
- How do they look like? (insert figures showing some data samples) - See Table 1 and 2
- How are they collected? What device/modality/sensor/etc. was used? The movies were collected from the website `https://www.themoviedb.org/`. These data were created by 162541 users between January 09, 1995 and November 21, 2019. This dataset was generated on November 21, 2019.
- How are they formatted? What do you need to do to parse them? Is there a parser available, or do you need to build your own? They are formatted in the generic csv style format. A csv parser in python would do the trick in extracting the data from the tables.
- (If human subject data) A statement indicating the IRB status and compliance with other human subject research protocols.

3.1 Figures

Figures
See Figure 1. Here is how you add footnotes. [1]

3.2 Tables

See Table 1.

[1] Sample of the first footnote.

Table 2: Tags

userId	movieId	Tag	Timestamp
4	5	extag	23:12
3	1	tag	22:22
1	3	tag	11:11

References

[1] George Kour and Raid Saabne. Real-time segmentation of on-line handwritten arabic script. In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 417–422. IEEE, 2014.

[2] George Kour and Raid Saabne. Fast classification of handwritten on-line arabic characters. In *Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of*, pages 312–318. IEEE, 2014.

[3] Guy Hadash, Einat Kermany, Boaz Carmeli, Ofer Lavi, George Kour, and Alon Jacovi. Estimate and replace: A novel approach to integrating deep neural networks with existing applications. *arXiv preprint arXiv:1804.09028*, 2018.