

Cooper Sloan

6.UAR: Preparation For Undergraduate Research (Spring 2017)

HW1: Preparation for paper draft

Due date: March 16, 2017, 4:00 PM

1. Features of a field-specific paper: Using a journal article in your field, find the answers to these questions for the paper you need to write. Post this homework on stellar and bring to workshop #2.

My paper is:

Chien, Steven I-Jy, Yuqing Ding, and Chienhung Wei. "Dynamic bus arrival time prediction with artificial neural networks." *Journal of Transportation Engineering* 128.5 (2002): 429-438.

1. Who is the audience of your paper? Field expert, broad technical audience, crossover?

The audience of the paper is a person with some experience in machine learning. They go into specific details and implementations of machine learning algorithms, but understanding these details was not necessary to understanding the paper as a whole. However, someone without some basic knowledge of neural nets would struggle to understand the paper.

2. What type of problem is solved in your field? Implementation, investigation, experimental?

The problem is investigative. The field still hasn't come up with the best solution to the problem, so most research in the field tries different approaches and evaluate their accuracy.

3. How is the real world problem/area of inquiry presented? Motivation, implication, minimal or absent?

The motivation is presented in the introduction. In this case it is very clear, predicting bus arrival times can improve network efficiency and reduce wait times for passengers.

4. What subheaders are typical in your field publications? Generic, project-specific, etc?

Typically for machine learning projects they discuss the model architecture and development. Then they go into model evaluation and finally make conclusions.

5. What information is represented visually? Concepts, data, diagrams, photographs, screenshots?

The model architecture is displayed visually which is well suited for neural nets specifically. They also include graphs of error and performance measures. There are also several data tables.

6. How are figures labeled and captioned?

Figures for the error measure describe what is being plotted across different models. The captions are relatively minimal, but the body of the paper adds context if necessary.

7. How is previous literature used? Where in the paper is it presented and what function does it serve? Is it motivating the problem, locating work in the field, justifying methods?

Previous literature is addressed in the beginning of the paper. It does not motivate the problem, because the problem motivation is relatively clear. Instead it shows some of the previous approaches that have been taken and describes how their approach adds something new.

8. What kind of results are produced? Experimental data, prototypes, benchmarks?

The results produced are accuracy measures. In this case the paper with best accuracy is the goal. They describe the problem and show the utility of their approach.

9. What metrics are used in your field?

Metrics commonly used are root mean square error for continuous valued results, as well as area under the curve. This is typically compared with the size of the test and validation sets, as well as the computation time. The size of the model often also is a factor taken into account.

10. Where are methods presented in the paper, if any? How important is this section?

The methods section of the paper describes the model selection process. It is rather important because model architecture, and feature selection is often the determining factor for accuracy in these kind of applications.

11. To ask your PI or mentor: Which information is assumed/unspoken in your field?

People know that neural nets are basically a brute force approach and may or may not be suitable for every application. Theorists are often skeptical of them because they require little thought.