

Rate My Tweet: Understanding Comparative Judgement in the Wild

Andy Gray

445348

Submitted to Swansea University in partial fulfilment
of the requirements for the Degree of Master of Science



Swansea University
Prifysgol Abertawe

Department of Computer Science
Swansea University

30th September 2021

Declaration

This work has not been previously accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed (candidate)

Date

Statement 1

This work is the result of my own independent study/investigations, except where otherwise stated. Other sources are clearly acknowledged by giving explicit references. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure of this work and the degree examination as a whole.

Signed (candidate)

Date

Statement 2

I hereby give my consent for my work, if accepted, to be archived and available for reference use, and for the title and summary to be made available to outside organisations.

Signed (candidate)

Date

I would like to dedicate this work to . . .

Abstract

In your abstract you should aim to summarise the core contributions of your work in the context of the problem domain. Start by outlining the domain and the problems posed within it. Discuss how the methods you focus on approach the relevant problems. You should end your abstract by concretely stating the tangible outputs and deliverables you have created in order to complete your work on this document, and whether those outputs represent an improvement or alternative approach to existing methods.

Your abstract should be a couple or so paragraphs long, and roughly approximate the order and flow you then use for structuring the main document. If a viewer has read your abstract then they should already understand at a high level what it is you have created and delivered, and whether it is better than or comparable to existing methods. If your project is driven by a research hypothesis then the reader should know what that is at a high level from this section. Reading on, little should surprise the viewer.

For paper submission of your thesis you should physically sign your name and add the date for each of the above declaration statements (black ink preferred). For digital submissions it is normally enough to simply type your name (see `custard.cls`), though you should sign and date them digitally using a touch or stylus input if at all possible.

Acknowledgements

This is an opportunity to acknowledge and thank those who have supported you throughout your studies. Friends and colleagues who you have studied alongside, your families, and your mentors within the department are the usual suspects.

Contents

1	Introduction	1
1.1	Motivations	1
1.2	Overview	1
1.3	Contributions	1
2	Lit Review	3
2.1	What is Comparative Judgement	3
2.2	Tunnel your internet connection via the university internet	4
2.3	Practice your Google Fu	4
2.4	Organising your citations in BibTeX	4
2.5	Properly using and formatting citations within the text	5
3	Methodology	7
3.1	Tools	7
3.2	Software Development Life Cycle Methodology	9
3.3	Data Set	9
4	Results and Discussion	11
4.1	Contributions	11
4.2	Future Work	11
5	Conclusions and Future Work	13
5.1	Contributions	13
5.2	Future Work	13
	Bibliography	15

Appendices	15
A Implementation of a Relevant Algorithm	17
B Supplementary Data	19

Todo list

1: Add this yourself and submit a pull request?	11
2: Add this yourself and submit a pull request?	13

Chapter 1

Introduction

1.1 Motivations

1.1.1 Objective

1.2 Overview

1.3 Contributions

The main contributions of this work can be seen as follows:

- **A \LaTeX thesis template**

Modify this document by adding additional \TeX files for your top level content chapters.

- **A typesetting guide for useful primitive elements**

Use the building blocks within this template to typeset each part of your document. Aim to use simple and reusable elements to keep your \LaTeX code neat and to make your document consistently styled throughout.

- **A review of how to find and cite external resources**

We review techniques and resources for finding and properly citing resources from the prior academic literature and from online resources.

Chapter 2

Lit Review

2.1 What is Comparative Judgement

Comparative judgement is a mathematical way to determine which observation item is better than the other item also being observed compared to each other. This method first got proposed in 1927 by Louis Leon Thurstone, a psychologist, under the term "the law of comparative judgement" [1, 2]. While comparative judgement is a technique that has been around for almost 100 years, it wasn't until the early nineties that this technique got proposed for use within an educational setting. This first proposal was by Politt and Murry [3], who conducted a study where they tested candidates on their English proficiency within Cambridge's CPE speaking exam. The judges watched 2-minute videos and judged which one out of a pair of videos they deemed better at the requested task in the exam. However, before this, in the ninety seventies and eighties, comparative judgement was presented as a more theoretical basis for educational assessments [4].

With the momentum of his findings, Politt then presented comparative judgement as a tool for exam boards to use to be able to compare the standards of A-Levels from the different exam boards, replacing the direct judgement of a script that was at the time currently being used [5]. In his papers titled, "Let's Stop Marking Exams" [6], he presents a valid argument for using comparative judgement, with the advantages it brings over some traditional types of marking.

How comparative judgement works is to present two options to a marker. The marker then gets asked to pick which one of the two options they think is the better one. The marker will get presented with all possible combinations available, each time picking

which one they think is the better one out of the two. An outputted score is then presented based on the method used. The original method, the Law of Comparative Judgement (LCJ), follows the formula:

$$S_i - S_j = x_{ij} \sqrt{\sigma_i^2 + \sigma_j^2 - 2r_{ij}\sigma_i\sigma_j},$$

Figure 2.1

S_i is the psychological scale value of stimuli i

However, an alternative version derived from Louis Leon Thurstone, referred to as the "Pairwise Comparison" [2], will provide an output based on the difference between the quality values is equal to the log of the odds in respect to object-A will be object-B. This formula gets represented as: $\log \text{odds}(A \text{ beats } B \mid v_a, v_b) = v_a - v_b$.

$$\Pr\{X_{ji} = 1\} = \frac{e^{\delta_j - \delta_i}}{1 + e^{\delta_j - \delta_i}} = \sigma(\delta_j - \delta_i)$$

.

2.2 Tunnel your internet connection via the university internet

2.3 Practice your Google Fu

If the request is simple enough, even a poorly structured query will likely return usable results. For more difficult to find resources you can leverage the language of the search engine to gather relevant papers and resources for your research more efficiently.

<https://www.gwern.net/Search>

"Internet Search Tips" [7] provides an excellent review of methods and tips for scouring the internet for hard to find resources. You will also be less likely to get caught behind journal paywalls when working remotely without a tunnel as your queries can be made to look for raw PDFs that are often released by the authors directly.

2.4 Organising your citations in BibT_EX

BibT_EX is a language for specifying resource citations. Every time you access and read an academic paper, take code from an online repository, or source the media such as

images from existing works, you should create a BibTeX entry in a file that you keep throughout your research. Software such as Mendeley [8] can help automate the process of building your BibTeX library of citations.

```
1 @inproceedings{kaj86,  
2   Address = {New York, NY, USA},  
3   Author = {Kajiya, James T.},  
4   Booktitle = {Proceedings of the 13th Annual Conference on Computer Graphics and  
5     Interactive Techniques},  
6   Isbn = {0-89791-196-2},  
7   Numpages = {8},  
8   Pages = {143--150},  
9   Publisher = {ACM},  
10  Series = {SIGGRAPH '86},  
11  Title = {The Rendering Equation},  
12  Year = {1986}  
}
```

Listing 2.1: An example BibTeX entry for an academic paper published in conference proceedings [9].

The BibTeX code listing above (Listing 2.1) shows an example of how to cite an academic paper; in this case one of the central papers in Computer Graphics research. The key **kaj86** is an arbitrary name chosen as a meaningful identifier for the resource. In the document text we can call on this resource as an inline citation using the L^AT_EX command `\cite{kaj86}`, which produces [9] at the location it is called. As long as a citation has been used at least once somewhere within the document then a formatted full citation will be created in the bibliography at the end of the document with the same citation number that is shown inline.

It is considerably easier to be disciplined in methodically taking note of the resources you access and make use of as you access them than it is to try and hunt them all down again at the time you need to write about them in your document. Invest time in being organised and consistent upfront and it will be easier when you come to write up.

2.5 Properly using and formatting citations within the text

Usually you would not put the URL of the resource you are citing directly in the text like is done previously in Section 2.3. The citation for the resource [7] is sufficient to reference it within the text given that full details of its location are then kept neatly within the bibliography at the end of the document.

In normal usage the purpose of a citation is not to direct the reader away from your thesis, but to justify and back up assertions you are making about the state of the domain. If a reader questions your assertions then they can follow the rabbit hole of papers which will likely also make and justify assertions with even earlier papers from the literature.

In the above case the intention is for the reader of this template to actually go to that resource and read what it has to say directly. The link is therefore shown clearly within the main text to indicate that the reader should visit it. This as opposed to wanting the reader to purely acknowledge that the facts which are within the resource legitimise the points made in this document, in which case a simple inline citation is the best way to back up your assertions. ?? specifically touches on the best practice for how to cite images which you are importing from existing work.

Chapter 3

Methodology

3.1 Tools

To create the web application and insights from the tweets, we required to use several tools. It is a requirement that we develop a full-stack web application with a user UI, an area to input the user's judgements on the tweet, store the results using a database, and extract information from the tweets using NLP techniques. Several factors within the final application needed to be satisfied for the tools to be appropriate for use.

3.1.1 Programming Language

While many programming languages can handle creating a full-stack application and conducting ML, for example, Java, Php and JavaScript. We decided to use the Python language. We decided upon Python due to our familiarity with it over the other main languages and its versatility. We made this decision because Python can make full-stack applications with the use of additional libraries, as well as handle most NLP ML tasks using libraries like NLTK, SpaCy, Sci-Kit Learn and TensorFlow.

3.1.2 Libraries

3.1.2.1 Web Application

For creating the web application, there were two main libraries available. These were Django and Flask.

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source [10].

While Flask is a small framework by most standards—small enough to be called a “micro- framework,” and small enough that once you become familiar with it, you will likely be able to read and understand all of its source code [11].

Flask has three main dependencies. The routing, debugging, and Web Server Gateway Interface (WSGI) subsystems come from Werkzeug; the template support is provided by Jinja2; and the command-line integration comes from Click. These dependencies are all authored by Armin Ronacher, the author of Flask [11].

Flask has no native support for accessing databases, validating web forms, authenticating users, or other high-level tasks. These and many other key services most web applications need are available through extensions that integrate with the core packages. As a developer, you have the power to cherry-pick the extensions that work best for your project, or even write your own if you feel inclined to. This is in contrast with a larger framework, where most choices have been made for you and are hard or sometimes impossible to change [11].

After experimenting with the two frameworks, we decided upon Flask. Flask got decided upon because of the short time frame to put the project together. Additionally, the lightweight nature of the framework also played a fact as we believe that as this will be just an initial prototype, all the other requirements that Django requires would be unessential additional to the project. Therefore, taking focus away from what we believe is the main focus.

3.1.2.2 NLP Tasks

3.1.3 IDE

3.2 Software Development Life Cycle Methodology

3.3 Data Set

3.3.1 Data Capture Method

3.3.2 Pre-Processing

Chapter 4

Results and Discussion

In this document we have demonstrated the use of a \LaTeX thesis template which can produce a professional looking academic document.

4.1 Contributions

The main contributions of this work can be summarised as follows:

- **A \LaTeX thesis template**

Modify this document by adding additional top level content chapters. These descriptions should take a more retrospective tone as you include summary of performance or viability.

- **A typesetting guide for useful primitive elements**

Use the building blocks within this template to typeset each part of your document. Aim to use simple and reusable elements to keep your document neat and consistently styled throughout.

- **A review of how to find and cite external resources**

We review techniques and resources for finding and properly citing resources from the prior academic literature and from online resources.

4.2 Future Work

Future editions of this template may include additional references to Futurama.

1: Add this yourself and submit a pull request?

Chapter 5

Conclusions and Future Work

In this document we have demonstrated the use of a \LaTeX thesis template which can produce a professional looking academic document.

5.1 Contributions

The main contributions of this work can be summarised as follows:

- **A \LaTeX thesis template**

Modify this document by adding additional top level content chapters. These descriptions should take a more retrospective tone as you include summary of performance or viability.

- **A typesetting guide for useful primitive elements**

Use the building blocks within this template to typeset each part of your document. Aim to use simple and reusable elements to keep your document neat and consistently styled throughout.

- **A review of how to find and cite external resources**

We review techniques and resources for finding and properly citing resources from the prior academic literature and from online resources.

5.2 Future Work

Future editions of this template may include additional references to Futurama.

2: Add this yourself and submit a pull request?

Bibliography

- [1] L. L. Thurstone, "Psychophysical analysis," *The American journal of psychology*, vol. 38, no. 3, pp. 368–389, 1927.
- [2] —, "A law of comparative judgment." *Psychological review*, vol. 34, no. 4, p. 273, 1927.
- [3] A. Pollitt and N. L. Murray, "What raters really pay attention to," *Studies in language testing*, vol. 3, pp. 74–91, 1996.
- [4] D. Andrich, "A rating formulation for ordered response categories," *Psychometrika*, vol. 43, no. 4, pp. 561–573, 1978.
- [5] P. Newton, J.-A. Baird, H. P. Harvey Goldstein, and P. Tymms, "Paired comparison methods," 2007.
- [6] A. Pollitt, "Let's stop marking exams," 01 2004.
- [7] G. Branwen. (2020) Internet search tips. [Online]. Available: <https://www.gwern.net/Search>
- [8] RELX Group. (2019) Mendeley. [Online]. Available: <https://www.mendeley.com>
- [9] J. T. Kajiya, "The rendering equation," in *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '86. New York, NY, USA: ACM, 1986, pp. 143–150.
- [10] Django. (2021) Meet django. [Online]. Available: <https://www.djangoproject.com/>
- [11] M. Grinberg, *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.", 2018.

Appendix A

Implementation of a Relevant Algorithm

```
1 #include <stdio.h>
2
3 int main(int argc, char *argv[]) {
4     printf("Hello world.\n");
5     return 0;
6 }
```

Listing A.1: An implementation of an important algorithm from our work.

Appendix B

Supplementary Data

The results of large ablative studies can often take up a lot of space, even with neat visualisation and formatting. Consider putting full results in an appendix chapter and showing excerpts of interesting results in your chapters with detailed analysis. You can use labels and references to refer the reader here for the full data.