

Improving Movie Recommendations with Copious Social Media Data

Jonas Myrlund

December 2013

PROJECT THESIS

Department of Computer and Information Science

Norwegian University of Science and Technology

Supervisor 1: Professor Heri Ramampiaro

Supervisor 2: Professor Helge Langseth

Preface

This is a project thesis about movie recommendations, specifically answering the question: “can movie recommendations be significantly improved by copious social media data?” I use the term “copious data” to emphasize the point that I am looking at big-scale trends, public sentiment, and hype; and that I do not delve into the realm of personalized recommendations. @TODO: Extend, refine and rewrite.

Trondheim, October 2, 2013

Jonas Myrlund

Acknowledgment

I would like to thank the following persons for great help in the work with this thesis.

@TODO: Acknowledge.

J.M.

Summary and Conclusions

@TODO

Contents

Preface	i
Acknowledgment	ii
Summary and Conclusions	iii
1 Introduction	2
1.1 Background	2
1.2 Objectives	4
1.3 Limitations	4
1.4 Approach	4
1.5 Structure of the Report	4
2 Equations, etc	6
2.1 Simple Equations	6
2.2 Including Figures	7
2.3 Including Tables	8
2.4 Copying Figures and Tables	8
2.5 References to Figures and Tables	10
2.6 A Word About Font-encoding	10
2.7 Plagiarism	11
3 Summary	12
3.1 Summary and Conclusions	12
3.2 Discussion	12
3.3 Recommendations for Further Work	13

<i>CONTENTS</i>	1
A Acronyms	14
B Additional Information	15
B.1 Introduction	15
B.1.1 More Details	15
Bibliography	16
Curriculum Vitae	17

Chapter 1

Introduction

The first chapter of a well-structured thesis is always an introduction, setting the scene with background, problem description, objectives, limitations, and then looking ahead to summarize what is in the rest of the report. This is the part that readers look at first—*so make sure it hooks them!*

1.1 Background

In this section, you should present the problem that you are going to investigate or analyze; why this problem is of interest; what has, so far, been done to solve the problem, and which parts of the problem that remain.

Problem Formulation

You should define your problem in a clear and unambiguous way and explain why this is a problem, why it is of interest—and to whom. It is also important to delimit the problem area.

Literature Survey

You should here present the main books and articles that treat problems that are similar to what you are studying. If you, later in your thesis, describe the “state of the art” – with a detailed literature survey, you may just give a very brief survey here (approx. a quarter of a page). If

this is the only literature survey, you need to go into more details. An objective of the literature survey is to show the reader that you are familiar with the main literature within your field of research – so that you do not “reinvent the wheel.”

References to literature can be given in two different ways:

- As an *explicit* reference: It is shown by [Lundteigen and Rausand \(2008\)](#) and partly also by [Rausand and Høyland \(2004\)](#) that
- As an *implicit* reference: It is shown (e.g., see [Rausand and Høyland, 2004](#), Chap. 4) that

In the example above, we have used “author-year” references, which is the preferred format.

Remark: Following agreement with your supervisor, you may also refer by numbers, for example, [1]. To do this, open the file `ramsstyle.sty` and comment out (by %) the command `\usepackage{natbib}` and un-comment the corresponding command `\usepackage[numbers]{natbib}`.¹

You may include a link to the Internet in the text or in a footnote by using a command like: <http://www.ntnu.edu/ross>.

When you refer to the scientific literature, you should always write in *present* tense. Example: [Rausand and Høyland \(2004\)](#) show that

Remark: Hyperlinks are included by the command `\usepackage{hyperref}` in `ramsstyle.sty`. If you feel that the hyperlinks are disturbing when you enter the text, or want to avoid the hyperlinks in printed text, you may either comment out or edit this command in `ramsstyle.sty`.

What Remains to be Done?

After you have defined and delimited your problem – and presented the relevant results found in the literature within this field, you should sum up which parts of the problem that remain to be solved.

¹Notice the strange way we have to write the “backslash” in the text. This is because the “backslash” is a command in \LaTeX .

1.2 Objectives

The main objectives of this Master's project are

1. This is the first objective
2. This is the second objective
3. This is the third objective
4. More objectives

All objectives shall be stated such that we, after having read the thesis, can see whether or not you have met the objective. “To become familiar with ...” is therefore not a suitable objective.

1.3 Limitations

In this section you describe the limitations of your study. These may be related to the study object (physical limitations, operational limitations), to the thoroughness of the analysis, and so on.

1.4 Approach

Here you should describe the (scientific) approach that you will use to solve the problem and meet your objectives. You should specify the approach for each objective.

If there are any ethical problems related to your approach, these should be highlighted and discussed.

1.5 Structure of the Report

The rest of the report is structured as follows. Chapter 2 gives an introduction to ...

Remark: Notice that chapter and section headings shall be written in lowercase, but that all main words should start with a capital letter.

The report should be no longer than 60 pages in this format (+ the CV).

Chapter 2

Equations, Figures, and Tables

The content of this chapter will vary with the topic of your thesis.

Remark: If you want a shorter chapter or section title to appear in the Table of Contents and in the headings of the chapter, you just include the short title in square brackets before the title of the chapter/section.

2.1 Simple Equations

This is how a simple equation is included:

$$F(t) = \int_0^t \exp(-\lambda x) dx \tag{2.1}$$

The equations are automatically given equation numbers – here (2.1) since this is the first equation in Chapter 2. Note that you can refer to the equation by referring to the “label” you specified as part of the equation environment.

You can also include equations without numbers:

$$F(t) = \sum_{i=1}^n \binom{n}{i} \sin(i \cdot t)$$



Figure 2.1: This is the logo of NTNU (rotated 15 degrees).

More Advanced Formulas

Please consult the \LaTeX documentation.

If you want to include a definition of a term/concept in the text, I have made the following macro (see in `ramssstyle.sty`):

✎ **Reliability:** The ability of an item to perform a required function under stated environmental and operational conditions and for a stated period of time.

2.2 Including Figures

If you use pdf \LaTeX (as recommended), all the figures must be in pdf, png, or jpg format. We recommend you to use the pdf format. Please place the figure files in the directory **fig**. Figures are included by the command shown for Figure 2.1. Please notice the “path” to the figure file written by a *forward* slash (/). You should not include the format of the figure file (pdg, png, or jpg) – just write the “name” of the figure.

Each figure should include a unique *label* as shown in the command for Figure 2.1. You can then refer to the figure by the *ref* command. Notice that you can scale the size of the figure by the option `scale=k`. You may also define a specific width or height of the figure by replacing the scale options by `width=k` or `height=k`. The factor `k` can here be specified in mm, cm, pc, and many other length measures. You may also give `k` as a fraction of the width of the text or of the height of the text, for example, `width=0.45\textwidth`. If you later change the margins of the text, the figure width will change accordingly. As illustrated in Figure 2.1, you may also rotate the figure – and also do many other things (please check the documentation of the package `graphicx` – it is available on your computer, or you may find it on the Internet).

Table 2.1: The degree of newness of technology.

Experience with the operating condition	Level of technology maturity		
	Proven	Limited field history or not used by company/user	New or unproven
Previous experience	1	2	3
No experience by company/user	2	3	4
No industry experience	3	4	4

In \LaTeX all figures are floating objects and will normally be placed at the top of a page. This is the standard option in all scientific reports. If you insist on placing the figure exactly where you declare the figure, you may include the command `[h]` (here) immediately after `\begin{figure}`. If you will force the figure to be located either at the top or bottom of the page, you may alternatively use `[t]` or `[b]`. For more options, check the documentation.

Large figures may be included as a *sidewaysfigure* as shown in Figure 2.2:¹

2.3 Including Tables

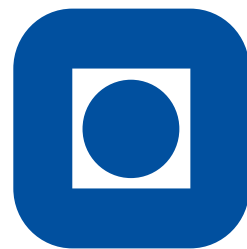
\LaTeX has a lot of different options to include tables. Only one of them is illustrated here.

Remark: Notice that figure captions (Figure text) shall be located *below* the figure – and that the caption of tables shall be *above* the table. This is done by placing the `\caption` command beneath the command `\includegraphics` for figures, and above the command `\begin{tabular*}` for tables.

2.4 Copying Figures and Tables

In some cases, it may be relevant to include figures and tables from other publications in your report. This can be a direct copy or that you retype the table or redraw the figure. In both cases, you should include a reference to the source in the figure or table caption. The caption might then be written as: *Figure/Table xx: The caption text is coming here (Rausand and Høyland, 2004).*

¹You can use a similar command for large tables.



NTNU – Trondheim
Norwegian University of
Science and Technology

Figure 2.2: This is the logo of NTNU.

In other cases, you get the idea from a figure or table in a publication, but modify the figure/table to fit your purpose. If the change is significant, your caption should have the following format: *Figure/Table xx: The caption text is coming here (adapted from [Rausand and Høyland, 2004](#)).*

2.5 References to Figures and Tables

Remember that all figures and tables shall be referred to and explained/discussed in the text. If a figure/table is not referred to in the text, it shall be deleted from the report.

2.6 A Word About Font-encoding

When you press a button (or a combination of buttons) on your keyboard, this is represented in your computer according to the *font-encoding* that has been set up. A wide range of font-encodings are available and it may be difficult to choose the “best” one. In the template, I have set up a font-encoding called UTF-8 which is a modern and very comprehensive encoding and is expected to be the standard encoding in the future. Before you start using this template, you should open the Preferences ->Editor dialogue in TeXworks (or TeXShop if you use a Mac) and check that encoding UTF-8 has been specified.

If you use only numbers and letters used in standard English text, it is not very important which encoding you are using, but if you write the Norwegian letters æ, ø, å and accented letters, such as é and ä, you may run into problems if you use different encodings. Please be careful if you cut and paste text from other word-processors or editors into your \LaTeX file!

Warning

If you (accidentally) open your file in another editor and this editor is set up with another font-encoding, your non-standard letters will likely come out wrong. If you do this, and detect the error, be sure *not* to save your file in this editor!!

This is not a specific \LaTeX problem. You will run into the same problem with all editors and word-processors – and it is of special importance if you use computers with different platforms

(Windows, OSX, Linux).

2.7 Plagiarism

Plagiarism is defined as “use, without giving reasonable and appropriate credit to or acknowledging the author or source, of another person’s original work, whether such work is made up of code, formulas, ideas, language, research, strategies, writing or other form”, and is a very serious issue in all academic work. You should adhere to the following rules:

- Give proper references to all the sources you are using as a basis for your work. The references should be give to the original work and not to newer sources that mention the original sources.
- You may copy paragraphs up to 50 words when you include a proper reference. In doing so, you should place the copied text in inverted commas (i.e., “Copied text follows ...”). Another option is to write the copied text as a quotation, for example:

Birnbaum’s measure of reliability importance of component i at time t is equal to the probability that the system is in such a state at time t that component i is critical for the system.

Rausand and Høyland (2004)

Chapter 3

Summary and Recommendations for Further Work

In this final chapter you should sum up what you have done and which results you have got. You should also discuss your findings, and give recommendations for further work.

3.1 Summary and Conclusions

Here, you present a brief summary of your work and list the main results you have got. You should give comments to each of the objectives in Chapter 1 and state whether or not you have met the objective. If you have not met the objective, you should explain why (e.g., data not available, too difficult).

This section is similar to the Summary and Conclusions in the beginning of your report, but more detailed—referring to the the various sections in the report.

3.2 Discussion

Here, you may discuss your findings, their strengths and limitations.

3.3 Recommendations for Further Work

You should give recommendations to possible extensions to your work. The recommendations should be as specific as possible, preferably with an objective and an indication of a possible approach.

The recommendations may be classified as:

- Short-term
- Medium-term
- Long-term

Appendix A

Acronyms

FTA Fault tree analysis

MTTF Mean time to failure

RAMS Reliability, availability, maintainability, and safety

Appendix B

Additional Information

This is an example of an Appendix. You can write an Appendix in the same way as a chapter, with sections, subsections, and so on.

B.1 Introduction

B.1.1 More Details

Bibliography

Lundteigen, M. A. and Rausand, M. (2008). Spurious activation of safety instrumented systems in the oil and gas industry: Basic concepts and formulas. *Reliability Engineering and System Safety*, 93:1208–1217.

Rausand, M. and Høyland, A. (2004). *System Reliability Theory: Models, Statistical Methods, and Applications*. Wiley, Hoboken, NJ, 2nd edition.