



Digital Discrimination: a research on data-driven social inequalities

some hyped-up tagline

Alice Morano

Supervised by Antonio Vetrò

Co-supervised by Dr Who

Department of Control and Computer Engineering

Faculty of Something

Polytechnic University of Turin

July/Sept, 2020

*A thesis submitted in partial fulfilment of the requirements for the degree
of M.Sc. in Computer Engineering.*



Copyright ©2020 Polytechnic university of Turin

WWW.POLITO.IT

First edition, March 1, 2020

To The Avengers

You know, for saving the world.

Acknowledgements

These are the acknowledgements. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Abstract

This is the abstract. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Contents

- 1 Introduction 1**
 - 1.1 Motivation 1
 - 1.2 Aims and Objectives 2
 - 1.2.1 Showing the Use of Acronyms 2
 - 1.3 Document Structure 2
- 2 Background & Literature Overview 3**
 - 2.1 Possible grounds of discrimination 3
 - 2.2 Some Other Section 5
 - 2.3 A Landscape Table Example 6
 - 2.3.1 Some Sub-technique One 8
 - 2.4 Some Technique Two 9
 - 2.5 Evaluation Criteria 10
 - 2.6 Related Work 10
 - 2.7 Summary 10
- 3 Research goals 13**
 - 3.1 Materials & Methods 14
- 4 Results & Discussion 15**
 - 4.1 An Example of a Table Spanning Multiple Pages 16
 - 4.2 Some Other Section 18
 - 4.3 A Landscape Table Example 19
 - 4.4 Summary 21
- 5 Evaluation 23**

5.1 Summary	24
6 Conclusions	25
6.1 Achieved Aims and Objectives	25
6.2 Critique and Limitations	25
6.3 Future Work	26
6.4 Final Remarks	26
Appendix A Media Content	27
Appendix B Installation Instructions	29
Appendix C User Manual	31
References	33

List of Figures

List of Tables

2.1	Table of possible discrimination grounds. For each ground, are listed known studies regarding cases of discrimination that actually happened and a brief description of them.	3
4.1	Performance of Ligity in HTS mode against the Ligity-compatible DUD-E targets. The mean (and standard deviation in parentheses) values of ROC AUC using Tanimoto is 0.622 (± 0.132), while for Tversky it is 0.671 (± 0.142); the mean $EF_{1\%}$ using Tanimoto is 5.648 (± 8.668), while for $EF_{1\%}$ using Tversky it is 9.047 (± 12.713).	16

List of Abbreviations

CDMA	Code Division Multiple Access	2
GSM	Global System for Mobile communication	2
TDMA	Time Division Multiple Access	2
UA	Used Acronym	2

Introduction

1.1 | Motivation

In the era of big data, it is vitally important to take into consideration the social implications of the scientific research. The idea behind good science is the more discoveries made and more research done, the closer we will be to healing ourselves and the world. Hence, sometimes we are all so smitten by reaching results that we overlook the fact that progress could have consequences on the social sphere. In particular, it happens when science and technology deal with human sensitive data.

It is a fact that we live in a world of pervasive technology and we score a goal when we manage to replace human wearing efforts by automating tasks. However, the results of Artificial Intelligence and Machine Learning can bring us on the one hand to easier solutions, but on the other, to possible unequal conditions.

There are several well known cases - that became real scandals - of the combined use of data and technology, such as the Cambridge Analytica case in early 2018, for political advertising purposes¹. Or the COMPAS software², used as a decision support tool by U.S. courts to determine criminal recidivism. Or again, the AI Amazon recruiting tool, biased against women³.

In some cases, the outcomes were not intentionally discriminating categories of people, but they did, due to either data feedback loops or initial databases that reflected inner social prejudices.

¹See the Wikipedia voice at https://en.wikipedia.org/wiki/Facebook-Cambridge_Analytica_data_scandal.

²See ProPublica article at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

³See an article on Fortune at <https://fortune.com/2018/10/10/amazon-ai-recruitment-bias-women-sexist/>.

Hence, when it comes to human factor, it is important to face the ethics implications of data analysis.

How we should implement algorithms for automated decisions or machine learning? Which data and which types of data should we take into account? Sometimes the legal, regulated aspects give us the answers (e.g. GDPR for data privacy), but in other circumstances, scientists should play a leading role in proper data usage and algorithms implementation, by using common sense.

We might think not to have any responsibility and treat data in the same way we treat any other bare ingredient of our work, as big black boxes, but this should not be the right approach towards sensitive data. What motivated me the most, is in fact the awareness that data are not only the “oil of the 21st century”⁴, but also the identity of billions of human beings, on which we cannot speculate.

1.2 | Aims and Objectives

The aim of this thesis is to research cases of possible social inequalities, arose due to data usage in web platforms.

The objective is to give examples of bad data usage, in order to find some guidelines for avoiding indirect use of sensitive information in algorithms.

1.2.1 | Showing the Use of Acronyms

In the early nineties, GSM was deployed in many European countries. Global System for Mobile communication (GSM) offered for the first time international roaming for mobile subscribers. The GSM’s use of Time Division Multiple Access (TDMA) as its communication standard was debated at length. And every now and then there are big discussion whether Code Division Multiple Access (CDMA) should have been chosen over TDMA.

If you want to know more about Global System for Mobile communication (GSM), Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA) and other acronyms, just read a book about mobile communication. Just to mention it: There is another Used Acronym (UA), for testing.

1.3 | Document Structure

⁴Clive Humby - British mathematician and entrepreneur, 2006.

Background & Literature Overview

2.1 | Possible grounds of discrimination

The following table summarizes cases of possible discrimination related to direct or indirect use of sensitive data.

Table 2.1: Table of possible discrimination grounds. For each ground, are listed known studies regarding cases of discrimination that actually happened and a brief description of them.

Ground	Type of discrimination	Description	Ref.
Search engines	Political orientation	A study on 4,556 undecided voters with different demographic characteristics in United States and India. It demonstrates that biased search rankings can shift the voting preferences of undecided voters by 20% or more.	[5]
Continued on next page			

Table 2.1 – continued from previous page

Ground	Type of discrimination	Description	Ref.
	Gender	Gender unequal distribution in photos retrieved by Bing for the query "person" and for queries based on 68 character traits (e.g., "intelligent person") in four regional markets.	[7]
	Language	A research that shows that cultural stereotyped biases from textual data propagate to artificial intelligence (AI) technologies in widespread use.	[2]
Freelance marketplaces	Gender and Race	Gender and race are significantly correlated with worker evaluations on TaskRabbit and Fiverr. This study is on 13,500 worker profiles: It gathers information about workers' gender, race, customer reviews, ratings, and positions in search rankings.	[6]
Resume search engines	Gender	Indeed, Monster, and CareerBuilder are tools that allow recruiters to proactively search for candidates based on keywords and filters. This study shows gender indirect discrimination that leads to disadvantage rankings for some candidates.	[3]
Continued on next page			

Table 2.1 – continued from previous page

Ground	Type of discrimination	Description	Ref.
Ad delivery	Race	This study concerns the problem of raising questions as to whether Google’s advertising technology exposes racial bias in society and how ad and search technology can develop to assure racial fairness	[?]
Intimate platforms	sexual orientation	Despite the fact that that romantic and sexual choices are understood as intensely personal, this article aims to analyse if intimate platform designers (such as Tinder, OkCupid or Grindr) can help alter troubling patterns of social interaction, without unduly interfering with individual intimate choices.	[?]

2.2 | Some Other Section

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and

an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.3 | A Landscape Table Example

Next is an example of a wide table on a landscape oriented paper.

m	x	y	z	a	A_m	B	C	x	y	z	a	A_m	B	C
1	16.128	+8.872	16.128	1.402	1.373	-146.6	-137.6	16.128	+8.872	16.128	1.402	1.373	-146.6	-137.6
2	3.442	-2.509	3.442	0.299	0.343	133.2	152.4	3.442	-2.509	3.442	0.299	0.343	133.2	152.4
3	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
4	0.993	-0.429	0.993	0.086	0.08	25.6	90	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
5	1.29	+0.099	1.29	0.112	0.097	-175.6	-114.7	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
6	0.483	-0.183	0.483	0.042	0.063	22.3	122.5	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
7	0.766	-0.475	0.766	0.067	0.039	141.6	-122	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
8	0.624	+0.365	0.624	0.054	0.04	-35.7	90	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
9	0.641	-0.466	0.641	0.056	0.045	133.3	-106.3	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
10	0.45	+0.421	0.45	0.039	0.034	-69.4	110.9	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
11	0.598	-0.597	0.598	0.052	0.025	92.3	-109.3	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.3.1 | Some Sub-technique One

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.3.1.1 | Some Sub-sub-technique One

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. Hello, here is some text without a meaning. This text should show what a

printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

2.4 | Some Technique Two with Super Long Title Which Will Overrun In Header

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives

you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Imagine some colourful description on Some Technique Three.

2.5 | Evaluation Criteria

This section should contain information on the metrics and background used to evaluate your work.

2.6 | Related Work

In this section you need to explain (and reference) similar work in literature. Make sure to:

- Give a systematic overview of papers with related/similar work
- Highlight similarities/differences to your work (perhaps in the form of a table)

Note that this section may be sectioned based on the different aspects of your dissertation. Some referenced text, as an example [1, 8, 4].

2.7 | Summary

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This

text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Research goals

This section should include a recipe of what you did (explain what you have done so if someone wants to reproduce the experiment, they can). A flow chart is typically helpful. Also, make sure to define all software that you used including version numbers and OS. Should also include a description of statistical methods used (if any).¹

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information?

¹For more information see: <http://rc.rcjournal.com/content/49/10/1229.short>

Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

3.1 | Materials & Methods

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Results & Discussion

Should include a reiteration of the experiments, and their outcome. Together with a description (discussion). Preamble should include a reminder of the aims and objectives together with a list of experiments to achieve these. Should include many charts and other visualization with appropriate descriptions.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This

text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

4.1 | An Example of a Table Spanning Multiple Pages

The following is an example of a table (Table 4.1) spanning multiple pages.

Table 4.1: Performance of Ligity in HTS mode against the Ligity-compatible DUD-E targets. The mean (and standard deviation in parentheses) values of ROC AUC using Tanimoto is 0.622 (± 0.132), while for Tversky it is 0.671 (± 0.142); the mean EF_{1%} using Tanimoto is 5.648 (± 8.668), while for EF_{1%} using Tversky it is 9.047 (± 12.713).

Target	No. of Ac- tives	No. of De- coys	ROC AUC Tani- moto	ROC AUC Tversky	BEDROC Tani- moto	BEDROC Tversky	EF _{1%} Tani- moto	EF _{1%} Tversky
ABL1	182	10,750	0.563	0.473	0.077	0.077	1.653	2.204
ACE	281	16,877	0.787	0.787	0.336	0.401	12.425	19.525
ACES	453	26,242	0.634	0.645	0.077	0.155	1.766	5.518
ADA	93	5,450	0.724	0.660	0.149	0.147	3.251	3.251
ADA17	532	35,898	0.638	0.728	0.103	0.283	1.317	9.030
ADRB1	247	15,850	0.523	0.647	0.065	0.129	1.619	5.262
ADRB2	231	14,999	0.523	0.589	0.052	0.040	1.735	0.000

(continued...)

Target	No. of Ac- tives	No. of De- coys	ROC AUC Tani- moto	ROC AUC Tversky	BEDROC Tani- moto	BEDROC Tversky	EF ₁ % Tani- moto	EF ₁ % Tversky
AKT1	293	16,450	0.386	0.548	0.039	0.107	2.737	3.080
AKT2	117	6,900	0.511	0.685	0.140	0.194	8.568	8.568
ALDR	159	8,988	0.574	0.610	0.202	0.172	10.747	6.322
AMPC	48	2,845	0.521	0.541	0.049	0.023	0.000	0.000
ANDR	269	14,349	0.722	0.742	0.194	0.354	4.839	24.938
AOFB	121	6,875	0.422	0.464	0.045	0.027	1.652	0.000
BACE1	283	18,100	0.441	0.775	0.017	0.310	0.000	13.062
BRAF	152	9,950	0.612	0.639	0.208	0.165	12.502	5.264
CASP3	199	10,694	0.600	0.734	0.068	0.258	0.502	7.031
CDK2	474	27,838	0.467	0.507	0.021	0.048	0.000	1.055
COMT	41	3,846	0.789	0.889	0.338	0.665	19.447	58.341
CP2C9	120	7,449	0.518	0.634	0.058	0.186	1.660	8.299
CP3A4	170	11,787	0.450	0.493	0.022	0.057	0.000	2.345
CSF1R	166	12,149	0.526	0.542	0.136	0.152	6.031	7.238
CXCR4	40	3,405	0.575	0.722	0.217	0.134	12.665	0.000
DEF	102	5,699	0.732	0.833	0.212	0.379	10.786	15.689
DHI1	330	19,348	0.481	0.595	0.089	0.062	2.422	1.211
DPP4	533	40,941	0.586	0.591	0.154	0.157	4.312	3.937
DRD3	480	34,048	0.484	0.441	0.043	0.046	1.251	0.626
DYR	231	17,196	0.694	0.758	0.210	0.230	6.504	7.371
EGFR	542	35,047	0.593	0.491	0.054	0.037	0.922	0.000
ESR1	383	20,683	0.838	0.861	0.527	0.594	31.281	39.101
ESR2	367	20,199	0.844	0.870	0.563	0.644	20.130	32.644
FA10	537	28,324	0.564	0.674	0.058	0.118	0.930	2.232
FA7	114	6,249	0.762	0.859	0.210	0.332	6.105	8.721
FABP4	47	2,749	0.786	0.744	0.191	0.276	0.000	10.623
FAK1	100	5,350	0.642	0.531	0.111	0.065	2.019	0.000
FGFR1	139	8,698	0.511	0.522	0.036	0.088	0.722	1.445
FKB1A	111	5,799	0.605	0.751	0.162	0.164	8.122	3.610
FNTA	592	51,493	0.411	0.625	0.012	0.132	0.000	4.053
FPPS	85	8,842	0.917	0.985	0.323	0.776	2.360	36.581
GCR	258	14,998	0.805	0.834	0.244	0.324	3.092	8.116
GLCM	54	3,790	0.667	0.685	0.182	0.279	1.873	11.240
GRIA2	158	11,842	0.662	0.684	0.248	0.154	11.392	5.696
GRIK1	101	6,547	0.656	0.668	0.203	0.102	7.978	1.995
HDAC2	185	10,300	0.676	0.734	0.187	0.201	4.318	4.318
HDAC8	170	10,449	0.640	0.819	0.120	0.377	2.946	8.250
HIVINT	100	6,640	0.390	0.554	0.030	0.116	0.000	3.018
HIVPR	535	35,724	0.663	0.872	0.072	0.490	0.187	23.898
HIVRT	338	18,884	0.495	0.475	0.124	0.085	4.443	1.777
HMDH	170	8,750	0.480	0.906	0.068	0.652	2.358	35.963
HS90A	88	4,850	0.635	0.506	0.096	0.083	0.000	3.436
HXK4	92	4,700	0.662	0.803	0.206	0.307	15.192	9.766
IGF1R	148	9,300	0.502	0.575	0.057	0.189	2.037	14.941
INHA	43	2,300	0.493	0.575	0.031	0.045	0.000	0.000
ITAL	138	8,500	0.619	0.465	0.037	0.065	0.000	0.728
JAK2	107	6,500	0.472	0.475	0.073	0.118	2.807	6.549
KIF11	116	6,850	0.755	0.781	0.149	0.219	4.289	2.574
KIT	166	10,449	0.463	0.437	0.045	0.030	0.000	0.000
KITH	57	2,850	0.649	0.838	0.228	0.709	14.069	47.483
KPCB	135	8,699	0.753	0.813	0.220	0.338	8.923	12.641
LCK	419	27,391	0.471	0.437	0.031	0.043	0.000	1.910
LKHA4	171	9,448	0.718	0.694	0.238	0.150	8.203	1.758
MAPK2	101	6,148	0.660	0.670	0.174	0.199	5.988	3.992
MCR	94	5,149	0.816	0.888	0.215	0.454	6.436	19.307
MET	166	11,249	0.566	0.531	0.130	0.065	6.032	0.603
MK01	79	4,550	0.518	0.602	0.121	0.206	5.095	3.821
MK10	104	6,600	0.488	0.489	0.020	0.031	0.962	0.962
MK14	578	35,847	0.511	0.589	0.040	0.064	0.173	0.519
MMP13	572	37,199	0.648	0.753	0.134	0.268	2.446	9.957

(continued...)

Target	No. of Ac- tives	No. of De- coys	ROC AUC Tani- moto	ROC AUC Tversky	BEDROC Tani- moto	BEDROC Tversky	EF ₁ % Tani- moto	EF ₁ % Tversky
MP2K1	121	8,146	0.669	0.569	0.187	0.058	3.293	0.823
NOS1	98	8,028	0.483	0.451	0.109	0.041	3.071	0.000
NRAM	98	6,200	0.853	0.859	0.342	0.290	11.221	3.060
PA2GA	99	5,150	0.793	0.756	0.225	0.153	1.020	3.059
PARP1	508	30,029	0.635	0.692	0.215	0.231	11.234	7.884
PGH1	195	10,798	0.645	0.637	0.077	0.100	0.000	2.050
PGH2	435	23,139	0.716	0.780	0.166	0.291	3.444	9.874
PLK1	107	6,800	0.658	0.531	0.123	0.048	1.871	0.000
PNPH	103	6,946	0.575	0.578	0.161	0.181	4.888	8.799
PPARA	373	19,399	0.783	0.778	0.262	0.280	6.693	7.764
PPARD	240	12,250	0.547	0.544	0.078	0.098	1.665	2.498
PPARG	484	25,299	0.515	0.605	0.055	0.118	0.619	4.955
PRGR	293	15,648	0.740	0.793	0.142	0.318	2.053	14.714
PTN1	130	7,249	0.398	0.538	0.055	0.090	0.000	3.068
PUR2	50	2,700	0.851	0.837	0.281	0.255	7.857	1.964
PYGM	77	3,944	0.403	0.492	0.016	0.137	0.000	3.917
PYRD	111	6,449	0.682	0.710	0.462	0.413	34.027	16.118
RENI	104	6,956	0.720	0.789	0.043	0.138	0.000	0.000
ROCK1	100	6,300	0.347	0.449	0.020	0.084	1.000	4.000
RXRA	131	6,950	0.788	0.900	0.219	0.596	6.091	27.407
SAHH	63	3,450	0.874	0.852	0.598	0.542	35.050	27.084
SRC	524	34,500	0.565	0.477	0.065	0.050	0.382	0.573
TGFR1	133	8,499	0.609	0.639	0.147	0.154	10.565	4.528
THB	103	7,450	0.794	0.762	0.238	0.150	10.614	0.965
THRB	461	27,000	0.605	0.706	0.063	0.166	2.166	5.632
TRY1	449	25,975	0.711	0.815	0.147	0.280	2.898	6.688
TRYB1	148	7,650	0.670	0.670	0.153	0.132	3.378	3.378
TYSY	109	6,745	0.594	0.725	0.071	0.226	0.911	5.468
UROK	162	9,850	0.525	0.650	0.036	0.120	0.000	1.854
VGFR2	409	24,948	0.632	0.578	0.083	0.093	1.465	1.465
WEE1	102	6,150	0.934	0.929	0.789	0.797	59.348	61.294
XIAP	100	5,150	0.752	0.974	0.190	0.897	8.077	51.490

4.2 | Some Other Section

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really?

Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

4.3 | A Landscape Table Example

Next is an example of a wide table on a landscape oriented paper.

m	x	y	z	a	A_m	B	C	x	y	z	a	A_m	B	C
1	16.128	+8.872	16.128	1.402	1.373	-146.6	-137.6	16.128	+8.872	16.128	1.402	1.373	-146.6	-137.6
2	3.442	-2.509	3.442	0.299	0.343	133.2	152.4	3.442	-2.509	3.442	0.299	0.343	133.2	152.4
3	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
4	0.993	-0.429	0.993	0.086	0.08	25.6	90	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
5	1.29	+0.099	1.29	0.112	0.097	-175.6	-114.7	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
6	0.483	-0.183	0.483	0.042	0.063	22.3	122.5	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
7	0.766	-0.475	0.766	0.067	0.039	141.6	-122	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
8	0.624	+0.365	0.624	0.054	0.04	-35.7	90	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
9	0.641	-0.466	0.641	0.056	0.045	133.3	-106.3	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
10	0.45	+0.421	0.45	0.039	0.034	-69.4	110.9	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1
11	0.598	-0.597	0.598	0.052	0.025	92.3	-109.3	1.826	-0.363	1.826	0.159	0.119	168.5	-161.1

4.4 | Summary

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Evaluation

In an ideal world, you should have two kind of evaluations. The first is against some ground truth (perhaps a random model?). The second kind of evaluation is against other people's work (accuracy, speed, etc.). Any dimension which is of interest, should be evaluated. Evaluation should be statistically sound.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the

letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

5.1 | Summary

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Conclusions

This section should have a summary of the whole project. The original aims and objective and whether these have been met should be discussed. It should include a section with a critique and a list of limitations of your proposed solutions. Future work should be described, and this should not be marginal or silly (e.g. add machine learning models). It is always good to end on a positive note (i.e. 'Final Remarks').

6.1 | Achieved Aims and Objectives

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

6.2 | Critique and Limitations

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This

text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

6.3 | Future Work

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

6.4 | Final Remarks

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Media Content

If the dissertation has a DVD or pendrive attached to it, you will need a section which explains what is on the media (structure, files, data, etc.). This could be a table with filename and description.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Installation Instructions

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

User Manual

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed

text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

References

- [1] Pablo Arrighi. Quantum computation explained to my mother. *arXiv preprint quant-ph/0305045*, 2003.
- [2] Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186, 2017.
- [3] Le Chen, Ruijun Ma, Anikó Hannák, and Christo Wilson. Investigating the impact of gender on rank in resume search engines. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery.
- [4] Jean-Paul Ebejer, Michael H. Charlton, and Paul W. Finn. Are the physicochemical properties of antibacterial compounds really different from other drugs? *Journal of Cheminformatics*, 8(1), 2016.
- [5] Robert Epstein and Ronald E. Robertson. The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences*, 112(33):E4512–E4521, 2015.
- [6] Anikó Hannák, Claudia Wagner, David Garcia, Alan Mislove, Markus Strohmaier, and Christo Wilson. Bias in online freelance marketplaces: Evidence from taskrabbit and fiverr. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, page 1914–1933, New York, NY, USA, 2017. Association for Computing Machinery.
- [7] Jahna Otterbacher, Jo Bates, and Paul Clough. Competent men and warm women: Gender stereotypes and backlash in image search results. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 6620–6631, New York, NY, USA, 2017. Association for Computing Machinery.
- [8] Chrislaine Withers-Martinez, Catherine Suarez, Simone Fulle, Samir Kher, Maria Penzo, Jean-Paul Ebejer, Kostas Koussis, Fiona Hackett, Aigars Jirgensons, Paul Finn, and Michael J. Blackman. Plasmodium subtilisin-like protease 1 (SUB1): Insights into the active-site structure, specificity and function of a pan-malaria drug target. *International Journal for Parasitology*, 42(6):597–612, 2012.