Lecture Notes in Artificial Intelligence 12117

Subseries of Lecture Notes in Computer Science

More information about this series at http://www.springer.com/series/1244

Denis Helic · Gerhard Leitner ·
Martin Stettinger · Alexander Felfernig ·
Zbigniew W. Raś (Eds.)

Foundations of Intelligent Systems

25th International Symposium, ISMIS 2020
Graz, Austria, September 23–25, 2020
Proceedings

Springer

Editors
Denis Helic
Graz University of Technology
Graz, Austria

Gerhard Leitner
University of Klagenfurt
Klagenfurt, Austria

Martin Stettinger
Graz University of Technology
Graz, Austria

Alexander Felfernig
Graz University of Technology
Graz, Austria

Zbigniew W. Raś
University of North Carolina at Charlotte
Charlotte, NC, USA

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-030-59490-9 ISBN 978-3-030-59491-6 (eBook)
https://doi.org/10.1007/978-3-030-59491-6

LNCS Sublibrary: SL7 – Artificial Intelligence

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains the papers selected for presentation at the 25th International Symposium on Methodologies for Intelligent Systems (ISMIS 2020), which was held in Graz, Austria, 2020. The symposium was organized by the Department of Software Technology at the Technical University of Graz, Austria. ISMIS is a conference series that started in 1986. Held twice every three years, it provides an international forum for exchanging scientific, research, and technological achievements in building intelligent systems. In particular, major areas selected for ISMIS 2020 included: Explainable AI (XAI), machine learning, deep learning, data mining, recommender systems, constraint based systems, autonomous systems, applications (configuration, Internet of Things, financial services, e-health, etc.), intelligent user interfaces, user modeling, human computation, socially-aware systems, digital libraries, intelligent agents, information retrieval, natural language processing, knowledge integration and visualization, knowledge representation, soft computing, and web and text mining. This year, the following sessions were organized: special sessions for invited talks, for best paper nominees and best student paper nominees, thematic sessions on natural language processing, deep learning and embeddings, digital signal processing, modeling and reasoning, machine learning applications, and finally a section containing short paper presentations on diverse related topics.

We received 79 submissions that were carefully reviewed by three or more Program Committee members or external reviewers. After a rigorous reviewing process, 35 regular papers, 8 short papers, and 3 invited papers were accepted for presentation at the conference and publication in the ISMIS 2020 proceedings volume. An additional selection of papers was assigned to a specific industrial track. It is truly a pleasure to thank all the people who helped this volume come into being and made ISMIS 2020 a successful and exciting event. In particular, we would like to express our appreciation for the work of the ISMIS 2020 Program Committee members and external reviewers who helped assure the high standard of accepted papers. We would like to thank all authors of ISMIS 2020, without whose high-quality contributions it would not have been possible to organize the conference. We are grateful to all the organizers of and contributors to a successful preparation and implementation of ISMIS 2020.

The invited talks for ISMIS 2020 were: "Complementing Behavioural Modeling with Cognitive Modeling for Better Recommendations," given by Marko Tkalcic, University of Primorska in Koper, Slovenia; "Fairness is not a number: Methodological implications of the politics of fairness-aware systems," given by Robin Burke, University of Colorado Boulder, USA; and "Brevity*," given by Robert West, EPFL Lausanne, Switzerland. We wish to express our thanks to all the invited speakers for accepting our invitation to give plenary talks. We are thankful to the people at Springer (Alfred Hofmann, Anna Kramer, and Aliaksandr Birukou) for supporting the ISMIS

2020. We believe that the proceedings of ISMIS 2020 will become a valuable source of reference for your ongoing and future research activities.

July 2020

Denis Helic
Gerhard Leitner
Martin Stettinger
Alexander Felfernig

Organization

General Chair

Alexander Felfernig	Graz University of Technology, Austria

Program Co-chairs

Denis Helic	Graz University of Technology, Austria
Gerhard Leitner	University of Klagenfurt, Austria
Martin Stettinger	Graz University of Technology, Austria

Steering Committee Chair

Zbigniew Ras	UNC Charlotte, USA

Steering Committee Members

Troels Andreasen	Roskilde University, Denmark
Annalisa Appice	Università degli Studi di Bari, Italy
Jaime Carbonell	CMU, USA
Michelangelo Ceci	Università degli Studi di Bari, Italy
Henning Christiansen	Roskilde University, Denmark
Juan Carlos Cubero	University of Granada, Spain
Floriana Esposito	Università degli Studi di Bari, Italy
Alexander Felfernig	Graz University of Technology, Austria
Mohand-Said Hacid	Université Claude Bernard Lyon 1, France
Nathalie Japkowicz	American University, USA
Marzena Kryszkiewicz	Warsaw University of Technology, Poland
Jiming Liu	Hong Kong Baptist University, Hong Kong
Jerzy Pawel Nowacki	Polish-Japanese Academy of IT, Poland
George A. Papadopoulos	University of Cyprus, Cyprus
Olivier Pivert	Université de Rennes 1, France
Zbigniew Ras	UNC Charlotte, USA, and Polish-Japanese Academy of IT, Poland
Henryk Rybinski	Warsaw University of Technology, Poland
Andrzej Skowron	Polish Academy of Sciences, Poland
Dominik Slezak	University of Warsaw, Poland

Programm Committee

Esra Akbas	Oklahoma State University, USA
Marharyta Aleksandrova	University of Luxembourg, Luxembourg

Aijun An	University of York, UK
Troels Andreasen	Roskilde University, Denmark
Annalisa Appice	Università degli Studi di Bari, Italy
Martin Atzmueller	Tilburg University, The Netherlands
Arunkumar Bagavathi	Oklahoma State University, USA
Ladjel Bellatreche	University of Poitiers, France
Robert Bembenik	Warsaw University of Technology, Poland
Petr Berka	University of Economics, Prague, Czech Republic
Maria Bielikova	Slovak University of Technology in Bratislava, Slovakia
Gloria Bordogna	CNR, Italy
Jose Borges	University of Porto, Portugal
François Bry	Ludwig Maximilian University of Munich, Germany
Jerzy Błaszczyński	Poznań University of Technology, Poland
Michelangelo Ceci	Università degli Studi di Bari, Italy
Jianhua Chen	Louisiana State University, USA
Silvia Chiusano	Politecnico di Torino, Italy
Roberto Corizzo	Università degli Studi di Bari, Italy
Alfredo Cuzzocrea	ICAR-CNR, University of Calabria, Italy
Marcilio De Souto	LIFO, University of Orleans, France
Luigi Di Caro	University of Torino, Italy
Stephan Doerfel	Micromata, Germany
Peter Dolog	Aalborg University, Denmark
Dejing Dou	University of Oregon, USA
Saso Dzeroski	Jožef Stefan Institute, Slovenia
Christoph F. Eick	University of Houston, USA
Tapio Elomaa	Tampere University of Technology, Finland
Andreas Falkner	Siemens AG, Austria
Nicola Fanizzi	Università degli studi di Bari “Aldo Moro”, Italy
Stefano Ferilli	Università degli Studi di Bari, Italy
Gerhard Friedrich	University of Klagenfurt, Austria
Naoki Fukuta	Shizuoka University, Japan
Maria Ganzha	Warsaw University of Technology, Poland
Paolo Garza	Politecnico di Torino, Italy
Martin Gebser	University of Klagenfurt, Austria
Bernhard Geiger	Know-Center GmbH, Austria
Michael Granitzer	University of Passau, Germany
Jacek Grekow	Bialystok University of Technology, Poland
Mohand-Said Hacid	Université Claude Bernard Lyon 1, France
Hakim Hacid	Zayed University, UAE
Allel Hadjali	LIAS, ENSMA, France
Mirsad Hadzikadic	UNC Charlotte, USA
Ayman Hajja	College of Charleston, USA
Alois Haselboeck	Siemens AG, Austria
Shoji Hirano	Shimane University, Japan
Jaakko Hollmén	Aalto University, Finland

Andreas Holzinger	Medical University and Graz University of Technology, Austria
Andreas Hotho	University of Wüerzburg, Germany
Lothar Hotz	University of Hamburg, Germany
Dietmar Jannach	University of Klagenfurt, Austria
Adam Jatowt	Kyoto University, Japan
Roman Kern	Know-Center GmbH, Austria
Matthias Klusch	DFKI, Germany
Dragi Kocev	Jožef Stefan Institute, Slovenia
Roxane Koitz	Graz University of Technology, Austria
Bozena Kostek	Gdańsk University of Technology, Poland
Mieczysław Kłopotek	Polish Academy of Sciences, Poland
Dominique Laurent	Université de Cergy-Pontoise, France
Marie-Jeanne Lesot	LIP6, UPMC, France
Rory Lewis	University of Colorado at Colorado Springs, USA
Elisabeth Lex	Graz University of Technology, Austria
Antoni Ligeza	AGH University of Science and Technology, Poland
Yang Liu	Hong Kong Baptist University, Hong Kong
Jiming Liu	Hong Kong Baptist University, Hong Kong
Corrado Loglisci	Università degli Studi di Bari, Italy
Henrique Lopes Cardoso	University of Porto, Portugal
Donato Malerba	Università degli Studi di Bari, Italy
Giuseppe Manco	ICAR-CNR, Italy
Yannis Manolopoulos	Open University of Cyprus, Cyprus
Małgorzata Marciniak	Polish Academy of Science, Poland
Mamoun Mardini	University of Florida, USA
Elio Masciari	University of Naples, Italy
Paola Mello	University of Bologna, Italy
João Mendes-Moreira	University of Porto, Portugal
Luis Moreira-Matias	NEC Laboratories Europe, Germany
Mikolaj Morzy	Poznań University of Technology, Poland
Agnieszka Mykowiecka	IPI PAN, Poland
Tomi Männistö	University of Helsinki, Finland
Mirco Nanni	ISTI-CNR, Italy
Amedeo Napoli	LORIA, CNRS, Inria, Université de Lorraine, France
Pance Panov	Jožef Stefan Institute, Slovenia
Jan Paralic	Technical University Kosice, Slovakia
Ruggero G. Pensa	University of Torino, Italy
Jean-Marc Petit	Université de Lyon, INSA Lyon, France
Ingo Pill	Graz University of Technology, Austria
Luca Piovesan	DISIT, Università del Piemonte Orientale, Italy
Olivier Pivert	IRISA-ENSSAT, France
Lubos Popelinsky	Masaryk University, Czech Republic
Jan Rauch	University of Economics, Prague, Czech Republic
Marek Reformat	University of Alberta, Canada
Henryk Rybiński	Warsaw University of Technology, Poland

Hiroshi Sakai	Kyushu Institute of Technology, Japan
Tiago Santos	Graz University of Technology, Austria
Christoph Schommer	University of Luxembourg, Luxembourg
Marian Scuturici	LIRIS, INSA Lyon, France
Nazha Selmaoui-Folcher	University of New Caledonia, France
Giovanni Semeraro	Università degli Studi di Bari, Italy
Samira Shaikh	UNC Charlotte, USA
Dominik Slezak	University of Warsaw, Poland
Urszula Stanczyk	Silesian University of Technology, Poland
Jerzy Stefanowski	Poznań University of Technology, Poland
Marcin Sydow	PJIIT and ICS PAS, Poland
Katarzyna Tarnowska	San Jose State University, USA
Herna Viktor	University of Ottawa, Canada
Simon Walk	Graz University of Technology, Austria
Alicja Wieczorkowska	Polish-Japanese Academy of Information Technology, Poland
David Wilson	UNC Charlotte, USA
Yiyu Yao	University of Regina, Canada
Jure Zabkar	University of Ljubljana, Slovenia
Slawomir Zadrozny	Polish Academy of Sciences, Poland
Wlodek Zadrozny	UNC Charlotte, USA
Bernard Zenko	Jožef Stefan Institute, Slovenia
Beata Zielosko	University of Silesia
Arkaitz Zubiaga	Queen Mary University of London, UK

Additional Reviewers

Max Toller
Henryk Rybiński
Allel Hadjali
Giuseppe Manco
Aijun An
Michelangelo Ceci
Giovanni Semeraro
Michael Granitzer
Simon Walk

Publicity Chair

Müslüm Atas	Graz University of Technology, Austria

Publication Chair

Trang Tran	Graz University of Technology, Austria

Web and Local Committee

Jörg Baumann	Graz University of Technology, Austria
Elisabeth Orthofer	Graz University of Technology, Austria
Petra Schindler	Graz University of Technology, Austria

Invited Talks

Complementing Behavioural Modeling with Cognitive Modeling for Better Recommendations

Marko Tkalčič

University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technologies, Glagoljaška 8, SI-6000 Koper, Slovenia
marko.tkalcic@famnit.upr.si

Abstract. Recommender systems are systems that help users in decision-making situations where there is an abundance of choices. We can find them in our everyday lives, for example in online shops. State-of-the-art research in recommender systems has shown the benefits of behavioural modeling. Behavioural modeling means that we use past ratings, purchases, clicks etc. to model the user preferences. However, behavioural modeling is not able to capture certain aspects of the user preferences. In this talk I will show how the usage of complementary research in cognitive models, such as personality and emotions, can benefit recommender systems.

Keywords: Recommender systems · Behavioural modeling · Cognitive modeling

Fairness Is Not a Number: Methodological Implications of the Politics of Fairness-Aware Systems

Robin Burke

Department of Information Science, University of Colorado, Boulder
robin.burke@colorado.edu

Abstract. This invited talk will explore the methodological challenges for intelligent systems development posed by questions of algorithmic fairness. In particular, we will discuss the ambiguity of fairness as a concept and the gap between simple formalizations of fairness questions and the challenges of real applications, where there are a multiplicity of stakeholders who may perceive fairness in different ways.

Keywords: Fairness · Machine learning

1 Fairness in Machine Learning and Recommendation

The problem of bias and fairness in algorithmic systems generally and in machine learning systems in particular is a critical issue for our increasingly data-centric world. However, the emergence of this topic has thrust computer scientists into unfamiliar territory, especially with regard to system development. A substantial body of research on fairness in machine learning, especially in classification settings, has emerged in the past ten years, including formalizing definitions of fairness [1, 5, 7, 10] and offering algorithmic techniques to mitigate unfairness ∼ [8, 11–13]. However, as noted in [4], there is little published research that investigates the complexity of real-world practices around fairness and reports findings from practical implementations. As a consequence, results from fairness-aware machine learning research often lack relevance for developers of real systems, and methodologies are lacking.

Fairness may be thought of as a property of particular system outcomes, like other system characteristics, such as reliability, and like these properties has to designed into systems, accounted for in testing, and monitored in deployment. Fairness is also a product of a complex set of social norms and interactions. Fair outcomes will be achieved and sustained throughout a system's lifecycle, if and only if, the process or methodology by which the system is developed and maintained itself incorporates the

This work has been supported in part by the National Science Foundation under award no. IIS-1911025.

multiple viewpoints of different stakeholders and provides a mechanism for arbitrating between them.

Economists recognize four rubrics under which fairness can be defined [9]:

1. Fairness as exogenous right: The definition of and necessity for fairness is imposed from outside, by legal requirement, for example.
2. Fairness as reward: Fairness is served by providing extra benefits from a system for those who contribute more to it. For example, a salary bonus in recognition of excellent work.
3. Fairness as compensation: Providing a benefit to those otherwise disadvantaged. A handicap in golf, for example, helps players of different abilities compete on a level basis.
4. Fairness as fitness. A subtle efficiency-based condition that states that a fair distribution is one in which resources go to those best able to utilize them. Thus, in dividing an estate, the inheritance of a piano might be fairly given to the most musically-inclined family member.

It is worth noting that the definitions do not all agree with each other or even point in the same direction in particular cases, particularly #2 and #3.

Fairness in machine learning is most often framed in terms of rubric #1. If government regulations require non-discriminatory treatment, organizations must comply. Such a stance has the promotes a legalistic approach in which organizations treat algorithmic fairness as just another compliance concern along with environmental regulations or financial disclosures. Note that such regulations themselves are inevitably the product of societal embrace of one or more of the other fairness rubrics, as applied to a particular issue.

A legalistic approach to fairness sometimes results in a "head-in-the-sand" approach. An organization may avoid evaluating its systems for discriminatory outcomes, because such a finding would incur liability for remedying the problem. Costs for building and maintaining fairness properties in systems can thus be deferred or avoided by neglecting to test for them, leaving it to external parties (usually those facing discrimination) to identify unfair impacts and seek redress under the law. Such an approach may be practically appealing, but it is not ethical for system designers who can reasonably anticipate such harms not to seek to mitigate them in advance: note the assertion in the ACM's code of ethics that computing professionals should strive to "minimize negative consequences of computing."

In addition to leading organizations to avoid proactive approaches, the focus on legalistic, exogenous aspects of fairness tends to reduce the issue to a matter of meeting some specific court-established legal test. This hides something crucial about fairness: namely, its contested and inevitably political nature. Regulations around discrimination or other fairness concerns do not arise from the sage wisdom of legal scholars; they arise from multi-vocal contestation in a political sphere where claims to various kinds of rights are asserted, argued and eventually codified (or not) through regulatory action.

Thus, as designers and developers of intelligent systems, it is incumbent on us to be proactive in the incorporation of fairness into system design. In doing so, we will further need to recognize that our view of fairness will necessarily shift from a focus on

fixed constraints imposed from *without*, to an understanding of fairness as a nuanced, multiply-interpreted, and sometimes-contested construct derived from *within* a real-world organizational context, where different stakeholders understand fairness in different ways. Such a point of view is consistent with scholarship in sociology [3], organizational justice [2], welfare economics [9], and public administration [6].

As examples of this perspective in action, we can consider cases of systems operating in organizational contexts where fairness derives from organizational mission. In these cases, the legalistic framing is less salient and the contested nature of fairness is more evident. In this talk, we will look particularly at recommendation in two such contexts: the non-profit microlending site Kiva.org and a (planned) open news recommender site. We will examine the different claims to fairness that arise and how the design and implementation of intelligent systems in these areas can be open to the on-going dialog between them.

References

1. Chouldechova, A.: Fair prediction with disparate impact: a study of bias in recidivism prediction instruments. Big data **5**(2), 153–163 (2017)
2. Colquitt, J.A.: On the dimensionality of organizational justice: a construct validation of a measure. J. Appl. Psychol. **86**(3), 386 (2001)
3. Cook, K.S., Hegtvedt, K.A.: Distributive justice, equity, and equality. Ann. Rev. Sociol. **9** (1), 217–241 (1983)
4. Cramer, H., Holstein, K., Vaughan, J.W., Daumé III, H., Dudík, M., Wallach, H., Reddy, S., Garcia-Gathright, J.: Challenges of incorporating algorithmic fairness into industry practice (2019)
5. Dwork, C., Hardt, M., Pitassi, T., Reingold, O., Zemel, R.: Fairness through awareness. In: Proceedings of the 3rd Innovations in Theoretical Computer Science Conference. pp. 214–226. ACM (2012)
6. Frederickson, H.G.: Social Equity and Public Administration: Origins, Developments, and Applications: Origins, Developments, and Applications. Routledge (2015)
7. Hardt, M., Price, E., Srebro, N., et al.: Equality of opportunity in supervised learning. In: Advances in Neural Information Processing Systems. pp. 3315–3323 (2016)
8. Kamiran, F., Calders, T., Pechenizkiy, M.: Discrimination aware decision tree learning. In: 2010 IEEE 10th International Conference on Data Mining (ICDM), pp. 869–874. IEEE (2010)
9. Moulin, H.: Fair Division and Collective Welfare. MIT press (2004)
10. Narayanan, A.: Translation tutorial: 21 fairness definitions and their politics. In: Proceedings of the Conference on Fairness Accountability and Transparency, New York, USA (2018)
11. Pedreshi, D., Ruggieri, S., Turini, F.: Discrimination-aware data mining. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 560–568. ACM (2008)
12. Zemel, R., Wu, Y., Swersky, K., Pitassi, T., Dwork, C.: Learning fair representations. In: Proceedings of the 30th International Conference on Machine Learning (ICML-13), pp. 325–333 (2013)
13. Zhang, L., Wu, X. Anti-discrimination learning: a causal modeling-based framework. Int. J. Data Sci. Anal. **4**, 1–16 (2017). https://doi.org/10.1007/s41060-017-0058-x

Brevity

Kristina Gligorić[1] and Ashton Anderson[2] and Robert West[1]

[1] EPFL, Lausanne, Switzerland
{robert.west, kristina.gligoric}@epfl.ch
[2] University of Toronto, Toronto, Ontario, Canada
ashton@cs.toronto.edu

Abstract. In online communities, where billions of people strive to propagate their messages, understanding how wording affects success is of primary importance. In this work, we are interested in one particularly salient aspect of wording: brevity. What is the causal effect of brevity on message success? What are the linguistic traits of brevity? When is brevity beneficial, and when is it not? Whereas most prior work has studied the effect of wording on style and success in observational setups, we conduct a controlled experiment, in which crowd workers shorten social media posts to prescribed target lengths and other crowd workers subsequently rate the original and shortened versions. This allows us to isolate the causal effect of brevity on the success of a message. We find that concise messages are on average more successful than the original messages up to a length reduction of 30–40%. The optimal reduction is on average between 10% and 20%. The observed effect is robust across different subpopulations of raters and is the strongest for raters who visit social media on a daily basis. Finally, we discover unique linguistic and content traits of brevity and correlate them with the measured probability of success in order to distinguish effective from ineffective shortening strategies. Overall, our findings are important for developing a better understanding of the effect of brevity on the success of messages in online social media.

Keywords: Causal effects · Experimental methods · Linguistic style · Twitter · Crowdsourcing

1 Introduction

How to convey a message most successfully is an age-old question. In online communities, where billions of people consume and strive to propagate their messages, understanding how wording affects success is of primary importance. In this work, we are interested in one particularly salient aspect of wording: brevity, or conciseness. What is the causal effect of brevity on message success? What are the linguistic traits of brevity? When is brevity beneficial? To establish a causal link between brevity and success, one would need to compare posts that convey the exact same semantic

This is an abridged version of a longer paper with a longer title [4].

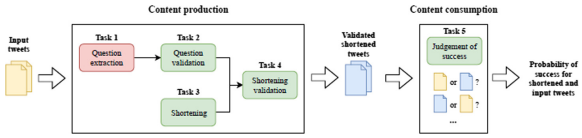

Fig. 1. Schematic diagram of the experimental design. The experiment consists of two parts, designed to replicate the *production* and *consumption* of textual content in online social media. The goal of the content production part (Tasks 1–4) is, for a given set of input tweets, to output shortened versions, having validated that the meaning is preserved. In the content consumption part of the experiment (Task 5), we show participants pairs of tweets, one treated short tweet, and the control long original tweet and ask which one will get more retweets. The outputs of the setup are binary votes (several per pair), based on which we compute the probabilities of success for each tweet version.

information and differ only in the number of characters used to express the fixed semantic content in a specific lexical and syntactic surface form.

Observational studies have striven to approximate this ideal goal by carefully controlling for confounding factors [2, 3, 6]. However, prior research arrived at contradictory conclusions and has not been able to guarantee that the semantic content of compared posts is identical, or that length is not confounded with other factors such as the inherent attractiveness of a message.

2 Experimental Setup

In order to overcome the aforementioned methodological hurdle inherent in observational designs and to more closely approximate the ideal of comparing two messages—one long, one short—expressing the exact same semantic content, we adopt an experimental approach instead, depicted schematically in Fig. 1.

Our experimental strategy consists of two steps: first, in the content production phase, we extract shortened versions of original tweets, and second, in the content consumption phase, we measure the quality of shortened versions compared with the unshortened version.

By ensuring (via additional checks) that length is the only difference between the unshortened tweets and the tweets shortened to prespecified lengths, we can attribute differences in quality to be causally related to brevity. Building on the fact that regular crowd workers can reduce the length of text by up to 70% without cutting any major content [1] and can accurately estimate which of two messages in a pair (for a fixed topic and user) will be shared more frequently [6], we deployed our experiment on Amazon Mechanical Turk.

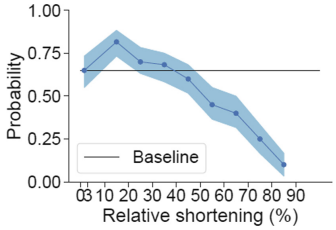

Fig. 2. The effect of conciseness as a function of the level of reduction, with 95% bootstrapped confidence interval.

3 Results

To the best of our knowledge, we are the first to study the effect of brevity in online social media in an experimental fashion. Applying our experimental framework, we collect short versions at 9 brevity levels for 60 original tweets, judged against the original tweets in a total of 27,000 binary votes. Based on this dataset, we address our research questions. We find that concise versions are on average more successful than the original messages up to a length reduction of 45%, while the optimal reduction is on average between 10% and 20% (resulting length 211–215 characters, down from the original 250) (Fig. 1). The observed effect is robust across different subpopulations of participants. Studying linguistic traits of brevity, we find that the shortening process disproportionally preserves verbs and negations—parts of speech that carry essential information—in contrast to, e.g., articles and adverbs. The shortening also preserves affect and subjective perceptions (quantified via the LIWC dictionary), and the effect is strongest for negative emotions. Addressing the question of when brevity is beneficial and when it is not, we find initial evidence that it is effective to omit certain function words and to insert commas and full stops, presumably as it increases readability by structuring or splitting long sentences. Ineffective editing strategies include deleting hashtags as well as question and exclamation marks, which have the potential to elicit discussion and reactions.

4 Discussion

In this work, our goals are threefold: to measure the effects of length constraints on tweet quality, to determine the linguistic traits of brevity, and to find when brevity is beneficial. To address these goals, we designed a large experiment that measured the quality of the shortened versions compared with the originals.

In brief, there are significant benefits of brevity. We observe that tweets can be successfully reduced up to 45% of their original length with no reduction in quality. The optimal range of shortening is consistently between 10% and 20% of the original length.

Practically, our findings are important for developing a better understanding of the effect of wording on the success of messages in online social media.

Methodologically, our experimental design highlights the power of a novel data collection paradigm, where crowd workers supply research data in the form of slightly edited text [1, 5, 7].

References

1. Bernstein, M.S., Little, G., Miller, R.C., Hartmann, B., Ackerman, M.S., Karger, D.R., Crowell, D., Panovich, K.: Soylent: a word processor with a crowd inside. Commun. ACM, **58**(8), 85–94 (2015)
2. Bramoulle, Y., Ductor. L.: Title length. J. Econ. Behav. Organ. **150**, 311–324 (2018)
3. Gligorić, K., Anderson, A., West, R.: How constraints affect content: the case of Twitter's switch from 140 to 280 characters. In: Proceedings of the International Conference on Web and Social Media (ICWSM) (2018)
4. Gligorić, K., Anderson, A., West, R.: Causal effects of brevity on style and success in social media. In: Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social (CSCW) (2019)
5. Ribeiro, M.H., Gligorić, K., West, R.: Message distortion in information cascades. In: Proceedings of the World Wide Web Conference (WWW) (2019)
6. Tan, C., Lee, L., Pang, B.: The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter. In: Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL) (2014)
7. West, R., Horvitz, E.: Reverse-engineering satire, or paper on computational humor accepted despite making serious advances. In: Proceedings of the AAAI Conference on Artificial Intelligence (AAAI) (2019)

Contents

Deep Learning and Embeddings

Digital Signal Processing

Short Papers