Lecture Notes in Computer Science 12801

More information about this subseries at http://www.springer.com/series/7409

Elisabeth Métais · Farid Meziane ·
Helmut Horacek · Epaminondas Kapetanios (Eds.)

Natural Language Processing and Information Systems

26th International Conference on Applications
of Natural Language to Information Systems, NLDB 2021
Saarbrücken, Germany, June 23–25, 2021
Proceedings

Springer

Editors
Elisabeth Métais
Conservatoire National des Arts et Métiers
Paris, France

Farid Meziane
University of Derby
Derby, UK

Helmut Horacek
German Research Center for Artificial Intelligence
Saarbrücken, Germany

Epaminondas Kapetanios
University of Hertfordshire
Hatfield, UK

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-030-80598-2 ISBN 978-3-030-80599-9 (eBook)
https://doi.org/10.1007/978-3-030-80599-9

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

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 presented at NLDB 2021, the 26th International Conference on Applications of Natural Language to Information Systems held during June 23–25, 2021, as a video conference at the German Research Center for Artificial Intelligence in Saarbrücken, Germany. We received 82 submissions for the conference. Each paper was assigned to three reviewers, taking into account preferences expressed by the Program Committee members as much as possible. After the review deadline, Program Committee members were asked to complete missing reviews. On the basis of these reviews, the Conference Organization Committee members and the Program Committee Chair decided to accept papers with an average score closer to *acceptance* than *weak acceptance* as full papers and papers with an average score around *weak acceptance* as short papers. In borderline cases, credit was given to experimentally-oriented papers with novel and ambitious concepts.

The final acceptance rate counting the number of full papers according to NLDB tradition was 23 percent (19 out of 82), similarly competitive in comparison to previous years. In addition, 14 submissions were accepted as short papers, and no posters since NLDB had to be held as a video conference like last year. Full papers were allowed a maximum of 12 pages and short papers a maximum of 8 pages. Originally, one more long paper and three more short papers were accepted, but the authors preferred to retract their submissions for personal reasons.

Similar as last year, the popular topics of classification and sentiment analysis have been addressed by many papers, and successful tools are reused and adapted, such as the transformer BERT. Following the trends of previous years, there is more diversification in the topics and specific issues addressed even in comparison to NLDB 2020. Several papers focus on methodological issues per se, prominently on perspectives on learning.

In addition to the reviewed papers, there were two invited talks at NLDB 2021:

- *Manfred Stede*, University of Potsdam, Germany
- *Elke Teich*, Saarland University, Germany

Moreover, a panel was organized by the recently built consortium NL4XAI (Interactive Natural Language Technology for Explainable Artificial Intelligence) to introduce their activities.

The accepted contributions (long and short papers) covered a wide range of topics, which we classified in nine topic areas, each covering a section in this volume:

- The Role of Learning
- Methodological Approaches
- Semantic Relations
- Classification
- Sentiment Analysis
- Social Media

- Linking Documents
- Multimodality
- Applications

The Role of Learning

One long and two short papers were categorized in this section. The long paper addresses limitations of learning, which may arise due to insufficient conceptual coverage of the data. One of the short papers deals with the combination of learning models, and the other short paper shows a machine in a learning environment inspired by human learning.

Methodological Approaches

Four long and one short paper were categorized in this section. The first two long papers deal with making frequently used techniques more fine-grained, namely word embedding and auto encoding. The short paper emphasizes modularity in a speech-to-speech translation model. Among the other two long papers, one addresses a specific form of entities, when numerals are included, and the other one deals with transfer across domains.

Semantic Relations

Three long papers were categorized in this section. The first one concentrates on causality relations, elaborated for a medical subdomain. The second one shows how precision about the entities involved can improve recognition in business domains. The third paper also addresses the role of preciseness of information, to detect combinations of relations which constitute events.

Classification

This traditional topic constitutes the longest section in the proceedings, with three long and four short papers. Some of the papers feature technical enhancements, including the use of word embeddings, features of multiword expressions, and applying a BERT capsule model. The majority of the contributions is oriented on the needs of the area of application. Some cover certain kinds of speech, such as hate speech, figurative language - irony and sarcasm - and technical language. Others are dedicated to specific domains: law and COVID-19 messages.

Sentiment Analysis

This is another popular topic, providing the second largest section, with three long and three short papers. As in the preceding section, there are papers featuring methodological issues and others oriented on the intended application. Methodological enhancements include multi-step transfer learning, cross-lingual learning, and importance weighting. Application topics vary greatly, they range from low resource language, opinions about vaccines, and aspects of structured product reviews, to headline stance and even literary artifacts.

Social Media

Three long papers were categorized in this section. They cover quite diverse tasks in analyzing texts in social media. These tasks are the detection of claims and evidence in arguments, attribution of authorship, and prediction of mental disorders.

Linking Documents

One long and one short paper were categorized in this section. They both address linking of some sort of a master document to a set of enhancing documents. One paper

uses the supplementary material for background information in news, whilst the other one does this for citation context of scientific documents.

Multimodality

One long and one short paper were categorized in this section. The long one addresses the recognition of essential elements from an image and its textual description, and the short one enhances the categorization of writers by incorporating visual data about them in the categorization process.

Applications

Two short papers were categorized in this section. They both address human computer interaction aspects of some sort. One of them attempts to enhance intent classification by multiple models, whilst the other one aims at retrieving graphical user interface prototypes on the basis of natural language specifications.

The conference organizers are indebted to the reviewers for their engagement in a vigorous submission evaluation process. We would also like to thank, for the organization help, some members of the DFKI GmbH.

June 2021

Epaminondas Kapetanios
Helmut Horacek
Elisabeth Métais
Farid Meziane

Organization

Conference Organization

Elisabeth Métais	Conservatoire des Arts et Métiers, France
Farid Meziane	University of Derby, UK
Helmut Horacek	German Research Center for AI, Germany

Program Chair

Epaminondas Kapetanios	University of Hertfordshire, UK

Program Committee

Jacky Akoka	CNAM & TEM, France
Hidir Aras	FIZ Karlsruhe, Germany
Nicolas Béchet	IRISA, France
Gloria Bordogna	IREA, CNR, Italy
Sandra Bringay	LIRMM, France
Christian Chiarcos	University Frankfurt am Main, Germany
Raja Chiky	ISEP, France
Philipp Cimiano	University of Bielefeld, Germany
Mohsin Farid	University of Derby, UK
Vladimir Fomichov	Moscow Aviation Institute, Russia
Flavius Frasincar	Erasmus University Rotterdam, The Netherlands
André Freitas	The University of Manchester, UK
Yaakov Hacohen-Kerner	Jerusalem College of Technology, Israel
Michael Herweg	IBM, Germany
Helmut Horacek	German Research Center for AI, Germany
Dino Ienco	IRSTEA, France
Ashwin Ittoo	HEC, University of Liège, Belgium
Epaminondas Kapetanios	University of Hertfordshire, UK
Zoubida Kedad	UVSQ, France
Eric Kergosien	GERiiCO, University of Lille, France
Christian Kop	University of Klagenfurt, Austria
Valia Kordoni	Humboldt University Berlin, Germany
Jochen Leidner	Refinitiv Labs and University of Sheffield, UK
Deryle W. Lonsdale	Brigham Young University, USA
Cédric Lopez	Emvista, France
Natalia Loukachevitch	Moscow State University, Russia
John P. Mccrae	National University of Ireland, Ireland
Elisabeth Métais	Conservatoire des Arts et Métiers, France
Farid Meziane	University of Derby, UK

Luisa Mich	University of Trento, Italy
Balakrishna Mithun	Limba Corp., USA
Nada Mimouni	Conservatoire des Arts et Métiers, France
Rafael Muñoz	Universidad de Alicante, Spain
Karen O'Shea	University of Derby, UK
Davide Picca	Columbia University, USA
Mathieu Roche	Cirad, TETIS, France
Paolo Rosso	Technical University of Valencia, Spain
Bahar Sateli	Concordia University, Canada
Khaled Shaalan	The British University in Dubai, UAE
Lucia Siciliani	University of Bari Aldo Moro, Italy
Max Silberztein	Université de Franche-Comté, France
Vijay Sugumaran	Oakland University Rochester, USA
Maguelonne Teisseire	Irstea, TETIS, France
Maria Truşcă	Bucharest University of Economic Studies, Romania
Dan Tufis	Academia Romana, Romania
Nicolas Turenne	INRA UPEM, France
Luis Alfonso Ureña	Universidad de Jaén, Spain
Sunil Vadera	University of Salford, UK
Csaba Veres	UIB, Norway
Isabelle Wattiau	ESSEC & CNAM, France
Wlodek Zadrozny	UNCC, USA

Webmaster

Christian Willms	German Research Center for AI, Germany

Additional Reviewers

Rémy Decoupes	Martin Lentschat
Mohammed Erritali	Nils Müller
Javi Fernández Koustava Goswami	Atul Kr. Ojha
Seethalakshmi Gopalakrishnan	Priya Rani
Philipp Heinisch	Raniah Rawass
Jude Hemanth	Lucile Sautot
Roberto Interdonato	Rima Türker
Rodrique Kafando	Lei Zhang

Contents