Cognitive Intelligence and Robotics

Series editors

Cognitive Intelligence refers to the natural intelligence of humans/animals involving the brain to serve the necessary biological functioning to perform an intelligent activity. Although tracing a hard boundary to distinguish intelligent activities from others remains controversial, most of the common behaviors/activities of living organisms that cannot be fully synthesized by artificial means are regarded as intelligent. Thus the act of natural sensing and perception, understanding of the environment and voluntary control of muscles, blood-flow rate, respiration rate, heartbeat, and sweating rate, which can be performed by lower level mammals, indeed, are intelligent. Besides the above, advanced mammals can perform more sophisticated cognitive tasks, including logical reasoning, learning and recognition and complex planning/coordination, none of which could be realized artificially to the level of a baby, and thus are regarded as cognitively intelligent.

The series aims at covering two important aspects of the brain science. First, it would attempt to uncover the mystery behind the biological basis of cognition with special emphasis on the decoding of stimulated brain signals/images. The coverage in this area includes neural basis of sensory perception, motor control, sensory-motor coordination and also understanding the biological basis of higher-level cognition, such as memory and learning, reasoning and complex planning. The second objective of the series is to publish brain-inspired models of learning, perception, memory and coordination for realization on robots to enable them to mimic the cognitive activities performed by the living creatures. These brain-inspired models of machine intelligence would supplement the behavioral counterparts, studied in traditional AI.

The series includes textbooks, monographs, contributed volumes and even selected conference proceedings.

More information about this series at http://www.springer.com/series/15488

Abhijit Mishra · Pushpak Bhattacharyya

Cognitively Inspired Natural Language Processing

An Investigation Based on Eye-tracking

Springer

Abhijit Mishra
India Research Lab
IBM Research
Bangalore, Karnataka, India

Pushpak Bhattacharyya
Indian Institute of Technology Patna
Patna, Bihar, India

ISSN 2520-1956 ISSN 2520-1964 (electronic)
Cognitive Intelligence and Robotics
ISBN 978-981-13-1515-2 ISBN 978-981-13-1516-9 (eBook)
https://doi.org/10.1007/978-981-13-1516-9

Library of Congress Control Number: 2018947487

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Natural language processing (NLP) is the field of research concerned with endowing computers with language ability—of understanding and generating language. Linguistics is one of the oldest disciplines, purporting to place study of language on a scientific footing, discovering the laws governing language analysis and production. Discovery of computer and, later, the advent of Internet introduced a completely new dimension to the study of language, viz. *Computational Linguistics (CL)*. Like other fields of artificial intelligence, CL sought to sieve apart the drudgery component of language analysis and generation from the creative component and delegate the former to mechanical processing. That is why *CL* is also called natural language processing (NLP).

Availability of huge amount of textual data in electronic form engendered NLP-ML, the use of Machine Learning (ML) techniques to process language. Until huge amounts of text data in e-form became available, rules were embedded in computers to do, for example, translation from one language to another, automatic question answering, and summarization. Rule-based systems are characterized by high precision and low recall. They are brittle too. For example, the accuracy of traditional part-of-speech (POS) taggers falls by at least 20% when applied to noisy text like tweets and social media postings. The reasons for the failings of rule-based systems are not far to seek. Rules are products of human understanding of phenomena. When these phenomena are of language, arbitrariness and exceptions confront the rule-framers more often than not, derailing theories. So NLP—like many other fields of AI—is by necessity a mixture of neat rules and rote-learnt patterns, the latter many times overwhelming the former, so much so that NLP is sometimes uttered in the same breath as machine learning.

Extracting and weighing text patterns with probability is the *modus operandi* of ML-NLP. Parts of text (e.g., N-grams) and properties of text (e.g., parts of speech) provide features for machine learning systems to make decisions, for example, to decide whether the text contains positive, neutral, or negative opinion; or, for example, out of 25 answers to a question, which one is the most appropriate. Agreed that such decisions are shallow, agreed that they are based on assumptions of underlying distributions which may be completely off the mark, but such statistics-based approaches have "delivered" more often than not, worked by producing "something" at least, instead of coming a cropper. ML-NLP has been found to be useful and is here to stay.

However, the crux of ML-NLP is the set of "features," the driver of the learning machine. Features are needed to train a machine. They are again needed when a new input arrives. For POS tagging, for example, suffixes of words are one of the important features. Features are also uncovered by annotation, enrichment of text by metainformation. For POS tagging, annotation produces POS tags on training data. This POS-tagged data is used to do parameters setting in HMMs/MEMMs/CRFs/Neural-Nets. A huge amount of annotated data is typically needed for solving complex NLP tasks like machine translation or sentiment analysis.

Annotation-driven machine learning-based NLP is the backdrop against which the study of the current monograph has been conducted. The authors would like to call this kind of NLP, Cognitive NLP. The basic paradigm, of course, is ML-NLP. Text is processed by a learning machine to make decisions. But the input to the machine is not only text parts and its features, but also extra-textual and even extra-linguistic features. We are talking about capturing the behavior not only of the text but also of the reader of the text.

Why is this point of view a potent one? For one thing, text data is not just data! Text is a manifestation of thought and emotion that give rise to cognitive processes in the brain. When a reader reads a piece of text, she experiences emotions, stances, nuances, subtleties, inferences, suggestions, and much more. There is a give-and-take between the reader and the text, a synergy, or shall we say a teamwork? Text reveals its secrets to a willing reader, and the reader responds by moving or staying the eye, producing brain waves and making face and body movements, all of which are capturable by the modern-day technology of eye-trackers, EEGs, and MEGs.

The research reported in the current monograph originates in the question, "Cannot ML-NLP be made more effective by capturing the Reader's behaviour when she is reading the text?" Take, for example, the problem of sarcasm detection. The surface information on the text is exactly in contradiction with the intended meaning, "I love being ignored"—the mirthless utterance of a frustrated guest to her host is exactly in opposition with the intent, "I have been ignored, and I have NOT liked it." On the face of it, the text does not convey the frustration, but the body language of the speaker is a give-way; the eye is an indication from which a sensitive host will know that she has not exactly done her job. When a machine is called upon to decide whether a piece of text or utterance is sarcastic, it often fails badly. Traditional sentiment detection systems show at least 10% fall in accuracy when called upon to decide the sentiment polarity of a sarcastic text. The research reported in the monograph breaks new ground by harnessing clues from eye behavior—captured by eye-trackers—for doing automatic sarcasm detection. Traditional ML features like N-grams, POS tags are albeit used. But these features are augmented with features extracted from eye behavior expressed by what is called scanpath in the monograph, to decide whether the input text is sarcastic or non-sarcastic. Scanpath essentially is the path described by staying (called “fixation”) of the eye and its transitions (called “saccade”).

It is useful to note here the emphasis given to “reading and readability.” At the heart of the eye-tracking study is capturing reading behavior. In fact, the seminal work by Carpenter and Gail in 1980 placed eye tracking at the center of readability study which until then was measured by text properties alone, like number of words and clauses. This work has inspired the foray of the authors into Cognitive NLP, viz. deploying features of cognitive behavior for machine learning in NLP. Readability itself is better measured by properties of scanpaths, as this monograph demonstrates.

An offshoot of this research is a rationalization of pricing of annotation. ML-NLP is a non-starter without annotated data. The authors have, however, been doubtful of the way annotators are compensated for their work. Annotation payment is typically in terms of the length of the sentences, i.e., the number of words; payment increases linearly with the number of words. But which sentence is more difficult to annotate, “John is in a bus” or “John is in a soup,” say, for sentiment or for the creation of parallel translation or for question answer database? It is the second one, since it makes use of an idiom and therefore demands a higher level of language skill. This difficulty is actually observed in the eye-tracking behavior of the annotator when she annotates into a screen fitted with eye-tracker. “Soup” has a higher duration of fixation, as the brain grapples with cognitive load of the metaphor of “soup”. The research reported in the monograph depicts well-appreciated work on predicting annotation difficulty in case of creating parallel corpora for machine translation and sentiment marking. Quantification of annotation difficulty is done by proposing translation complexity index (TCI) and sentiment annotation complexity (SAC) which are predicted by a regressor using text and eye-tracking features.

In summary, this monograph is a depiction of trail-blazing work on the use of cognitive behavior in ML-NLP and should prove useful to students of natural language processing, machine learning, and cognitive science.

Bangalore, India — Abhijit Mishra
Patna, India — Pushpak Bhattacharyya
July 2018

Acknowledgements

First, we would like to thank Prof. Michael Carl, Center for Research and Innovations in Translation and Translation Technologies (CRITT), Copenhagen Business School (CBS), Denmark, who shared his knowledge, experience, and expertise in the field of psycholinguistics and translation process research that helped us incubate the core ideas and approaches in a relatively newer domain like Cognitive NLP.

Professor Azizdduin Khan from the Department of Humanities and Social Sciences, IIT Bombay, has wholeheartedly supported our research through his mentorship and guidance; our sincere thanks to Prof. Khan. We also thank Prof. Ganesh Ramakrishnan, Department of Computer Science and Engineering, and Prof. Malhar Kulkarni, Department of Humanities and Social Sciences, for their valuable inputs.

For effectively carrying out the research, we needed continuous access to good-quality eye-tracking machinery. Thanks to IRCC, IIT Bombay, for granting us funds to procure a state-of-the-art SR Research EyeLink 1000 Plus eye-tracker. We would like to thank the lexicographers of CFILT, Mrs. Jaya Jha, Mrs. Laxmi Kashyap, Mrs. Rajita Sukla, and other members of CFILT laboratory, and the students of IIT Bombay who participated in the eye-tracking experiments as annotators.

Finally, a special vote of thanks goes to our collaborators: Kuntal Dey and Seema Nagar from IBM Research, India; Dr. Aditya Joshi, Dr. Joe Cheri Ross, Dr. Anoop Kunchukuttan, and Diptesh Kanojia from IIT Bombay. It is because of such collaborators, our ideas could be refined and translated into high-quality publications and systems.

Abhijit Mishra
Pushpak Bhattacharyya

Contents

About the Authors

Abhijit Mishra is currently a part of IBM Research, Bangalore, India, where he serves as Research Scientist in the Department of Cognitive Solutions and Services. Prior to joining IBM Research, he was a Ph.D. student in the Department of Computer Science and Engineering, Indian Institute of Technology Bombay. He interned at the Center for Research and Innovation in Translation and Translation Technologies, CBS, Copenhagen, under the guidance of Prof. Michael Carl. He was also a part of "Developing Multilingual Resources for Indian Languages through Crowdsourcing," a project launched by the IIT Bombay in collaboration with Xerox Research Center India, Bangalore. The aim of the project was to build a system that helps NLP developers customize and float linguistic annotation tasks using popular crowdsourcing service providers (like Amazon's Mechanical Turk). He is currently involved in multiple projects based on natural language generation.

Prof. Pushpak Bhattacharyya is recent past President of the ACL (2016–2017). He is Director of the IIT Patna and Vijay and Sita Vashee Chair Professor in the Department of Computer Science and Engineering, IIT Bombay. He studied at IIT Kharagpur (B.Tech.), IIT Kanpur (M.Tech.), and IIT Bombay (Ph.D.) and has been Visiting Scholar and Faculty at MIT; Stanford; UT Houston; and University Joseph Fourier, France. His main research areas are natural language processing, machine learning, and artificial intelligence. He has published more than 250 research papers and led government and industry projects of international and national importance. He is Author of the textbook "Machine Translation," Fellow of the National Academy of Engineering, Eminent Engineer awardee of the Institute of Engineers, India, and Recipient of the Patwardhan Award (IIT Bombay) and VNMM Award (IIT Roorkee)—both for technology development—and faculty grants from IBM, Microsoft, Yahoo, and the United Nations.