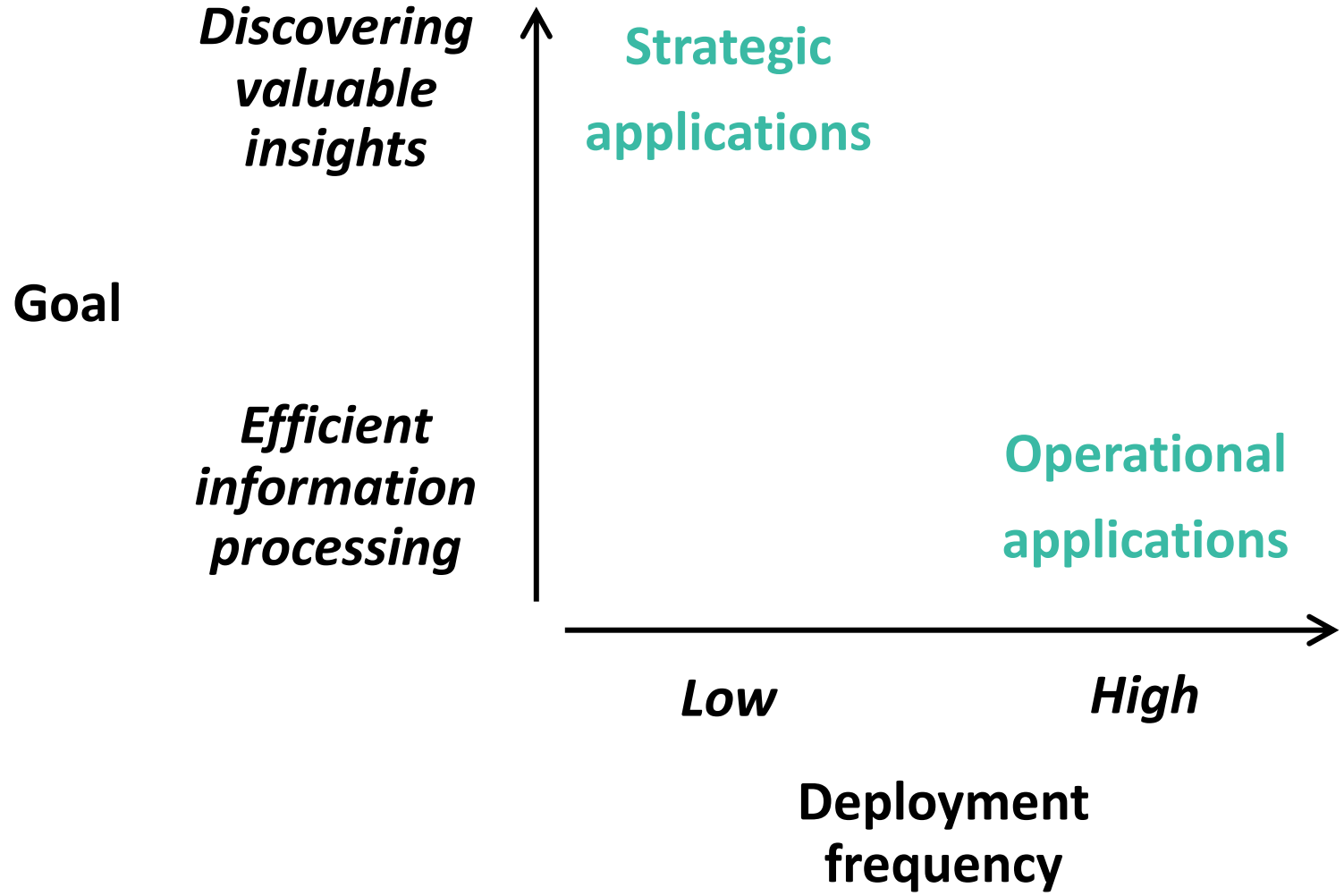
NLP for Business and Financial Analysts: Scope of Application

Goals

1. Appreciating the gamut of NLP applications
2. Distinguishing different types of NLP applications
3. Familiarising with selected examples of NLP applications

Broad categories of NLP applications



Operational applications

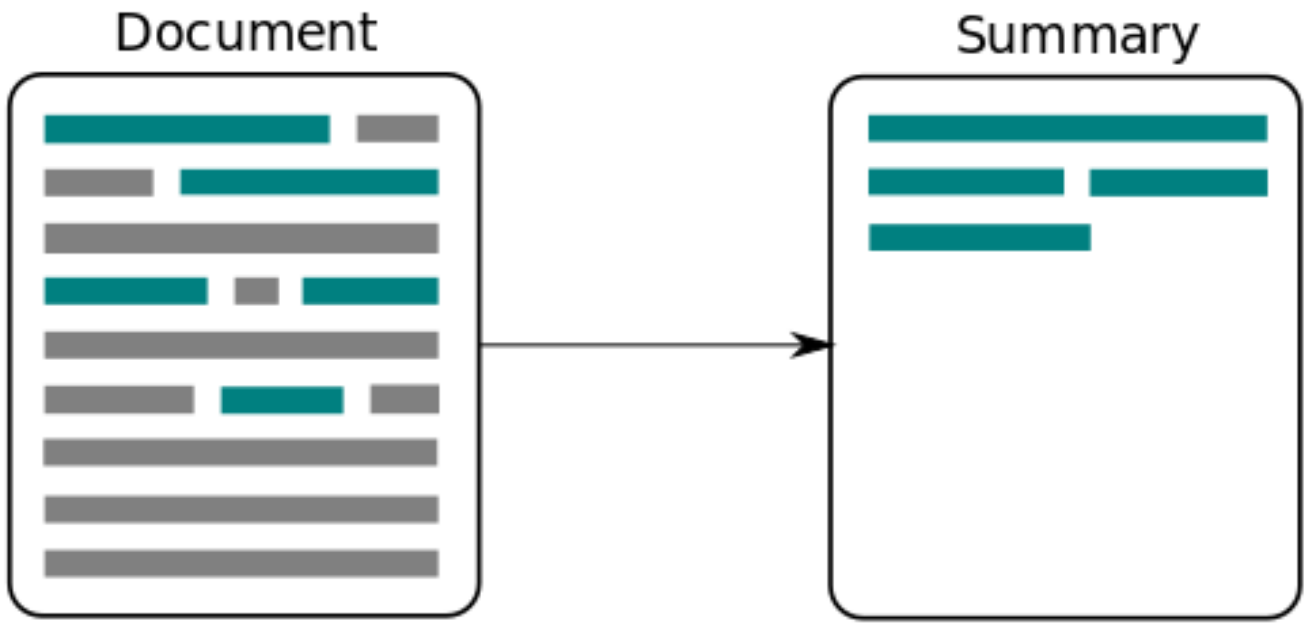
Examples of operational applications

- Text summarisation
- Speech recognition and synthesis
- Dialogue systems and chatbots
- Question answering
- Text translation
- Phonetics

Text summarization

Creating concise summaries out of long documents (e.g., legal documents) is a vital NLP task when actors must make choices efficiently.

Naïve text summarizers are easy to build with Python's NLTK library.



Chatbots

Chatbots play an increasingly central role in customer relationship management systems.

Simple chatbots can be designed with bare-bone Python code or NLTK.

The various ‘assistants’ we interact with are powered by Deep Learning and advanced NLP.

Operational applications are popular amongst software engineers — let us focus on strategic applications, closer business and financial analysts' work.

Strategic applications

Examples of strategic applications

- Vector semantics and embeddings
- Topic modelling
- Sentiment analysis
- Span/text categorization
- Information extraction

Embeddings

Embeddings have been used to appreciate how YouTube videos shape users' medical information.

Knowing YouTube users' health-related vocabulary can help improve the communication effectiveness of future medical campaigns.

MIS Quarterly

SPECIAL ISSUE

GO TO YOUTUBE AND CALL ME IN THE MORNING: USE OF SOCIAL MEDIA FOR CHRONIC CONDITIONS[1]

Xiao Liu
Department of Information Systems, W. P. Carey School of Business, Arizona State University,
Tempe, AZ 85287 U.S.A. {xiao.liu.10@asu.edu}

Bin Zhang
Eller College of Management, University of Arizona,
Tucson, AZ 85721-0108 U.S.A. {binzhang@arizona.edu}

Anjana Susarla
Eli Broad College of Business, Michigan State University,
East Lansing, MI 48824 U.S.A. {asusarla@msu.edu}

Rema Padman
Heinz College of Information Systems and Public Policy, Carnegie Mellon University,
Pittsburgh, PA 15213-3890 U.S.A. {rpadman@cmu.edu}

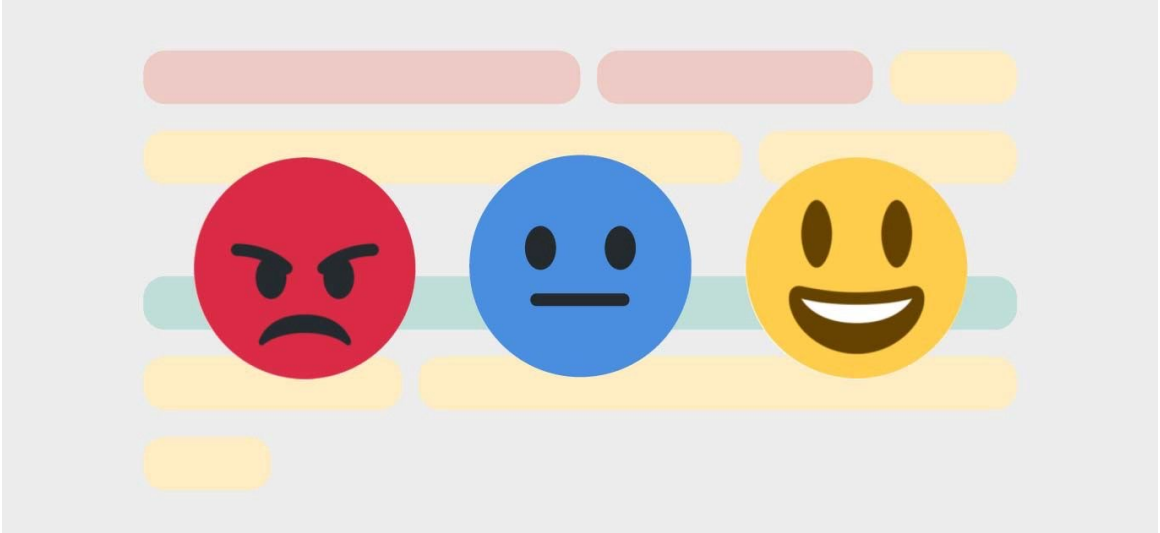
Video sharing social media platforms, such as YouTube, offer an effective way to deliver medical information. Few studies have identified evidence-backed digital therapeutics with technology-enabled interventions to improve the ease with which patients can retrieve medical information to manage chronic conditions. We propose an interdisciplinary lens that synthesizes deep learning methods with themes emphasized in Information Systems and Healthcare Informatics research to examine user engagement with encoded medical information in YouTube videos. We first use a bidirectional long short-term memory method to identify medical

Sentiment analysis

Social media — the main digital data source — show a polarising effect.

Sentiment analysis summarises actors' posture toward offerings (being sneakers, investment opportunities, or political leaders).

Hence, it offers critical inputs to profile consumers and sustain new product launch campaigns.

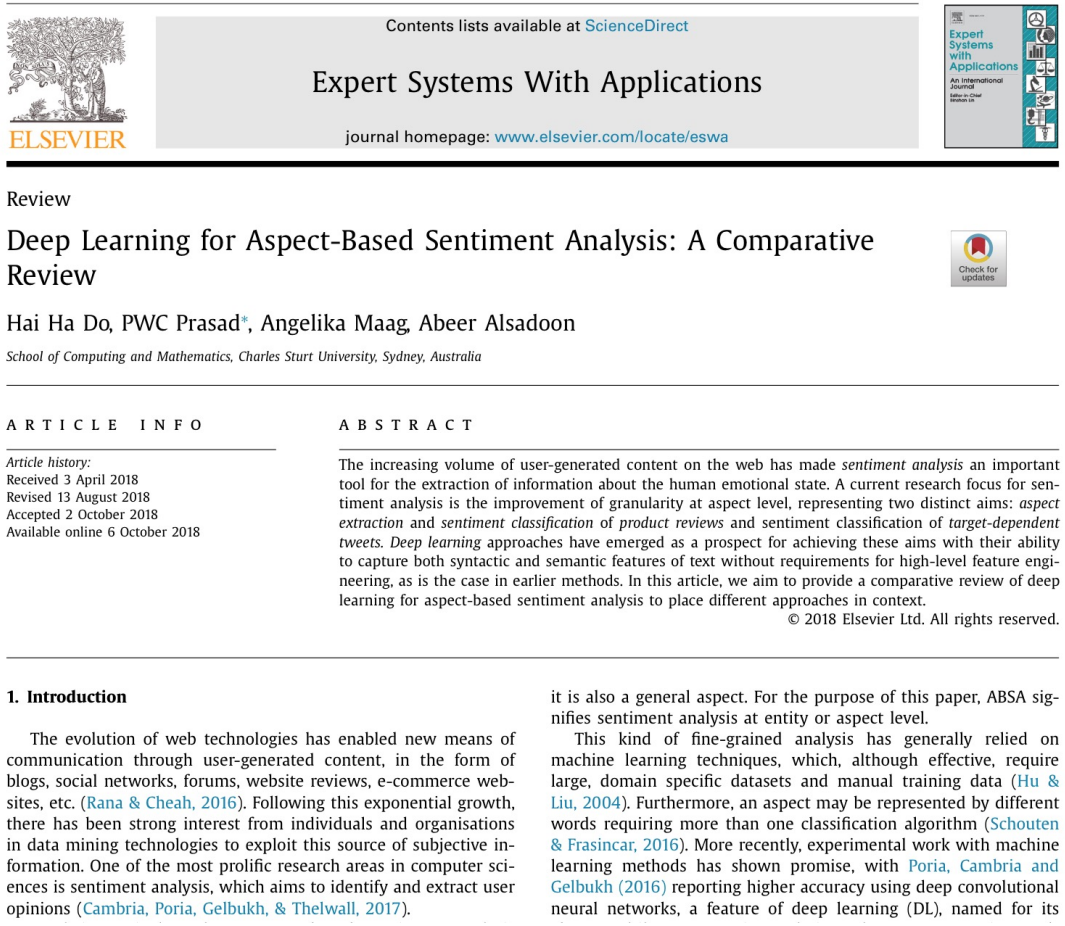


Aspect-based sentiment analysis

Sentiment analysis has been extensively applied to product review data.

Recently, analysts have used Deep Learning to break down sentiment analysis into three sub-tasks:

- Opinion target extraction (e.g., a restaurant service level)
- Aspect category detection (e.g., service speed or courtesy)
- Sentiment polarity (e.g., service speed's perceived adequacy).



Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Review

Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review

Hai Ha Do, PWC Prasad*, Angelika Maag, Abeer Alsadoon

School of Computing and Mathematics, Charles Sturt University, Sydney, Australia

ARTICLE INFO

Article history:
Received 3 April 2018
Revised 13 August 2018
Accepted 2 October 2018
Available online 6 October 2018

ABSTRACT

The increasing volume of user-generated content on the web has made *sentiment analysis* an important tool for the extraction of information about the human emotional state. A current research focus for sentiment analysis is the improvement of granularity at aspect level, representing two distinct aims: *aspect extraction* and *sentiment classification* of *product reviews* and sentiment classification of *target-dependent tweets. Deep learning* approaches have emerged as a prospect for achieving these aims with their ability to capture both syntactic and semantic features of text without requirements for high-level feature engineering, as is the case in earlier methods. In this article, we aim to provide a comparative review of deep learning for aspect-based sentiment analysis to place different approaches in context.

1. Introduction

The evolution of web technologies has enabled new means of communication through user-generated content, in the form of blogs, social networks, forums, website reviews, e-commerce websites, etc. (Rana & Cheah, 2016). Following this exponential growth, there has been strong interest from individuals and organisations in data mining technologies to exploit this source of subjective information. One of the most prolific research areas in computer sciences is sentiment analysis, which aims to identify and extract user opinions (Cambria, Poria, Gelbukh, & Thelwall, 2017).

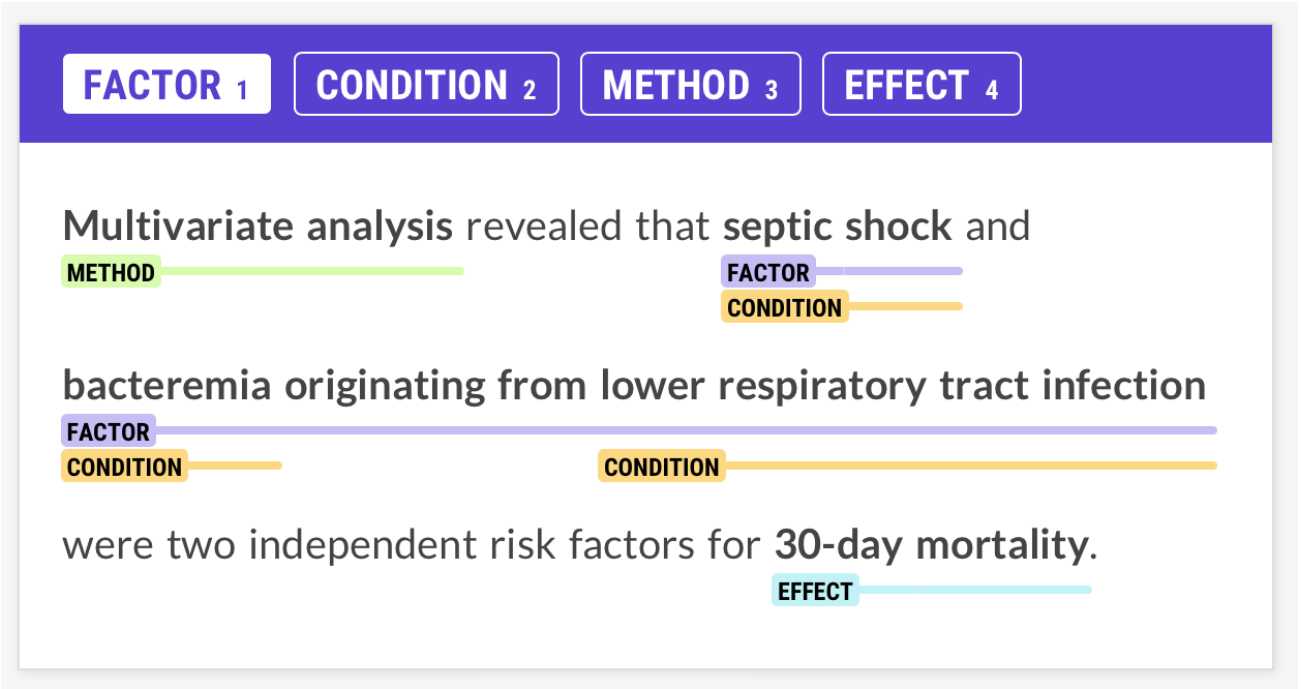
it is also a general aspect. For the purpose of this paper, ABSA signifies sentiment analysis at entity or aspect level.

This kind of fine-grained analysis has generally relied on machine learning techniques, which, although effective, require large, domain specific datasets and manual training data (Hu & Liu, 2004). Furthermore, an aspect may be represented by different words requiring more than one classification algorithm (Schouten & Frasincar, 2016). More recently, experimental work with machine learning methods has shown promise, with Poria, Cambria and Gelbukh (2016) reporting higher accuracy using deep convolutional neural networks, a feature of deep learning (DL), named for its

Span/text categorization

Analysts can reveal nuanced meanings emerging from text spans or a broader text corpus by drawing on ML-powered classifiers.

For example, categorising financial analyst reports' spans may reveal the foundations of stock market price predictions.



Topic modeling

Topic modelling is an unsupervised ML tool that reveals the themes included in text corpora.

For example, topic modelling can reveal the attention structure of financial analysts, who base their recommendations on subsets of earnings conference calls.

informs
http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 64, No. 6, June 2018, pp. 2833–2855
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach

Allen H. Huang,[a] Reuven Lehavy,[b] Amy Y. Zang,[a] Rong Zheng[c]

[a] Department of Accounting, Hong Kong University of Science and Technology, Kowloon, Hong Kong; [b] Ross School of Business, University of Michigan, Ann Arbor, Michigan 48109; [c] Department of Information Systems, Business Statistics, and Operations Management, Hong Kong University of Science and Technology, Kowloon, Hong Kong

Contact: allen.huang@ust.hk (AHH); rlehavy@umich.edu, http://orcid.org/0000-0003-3875-8848 (RL); amy.zang@ust.hk (AYZ); rzheng@ust.hk (RZ)

Received: September 25, 2015
Revised: July 25, 2016; November 6, 2016
Accepted: December 18, 2016
Published Online in Articles in Advance: June 13, 2017

https://doi.org/10.1287/mnsc.2017.2751

Abstract. This study examines analyst information intermediary roles using a textual analysis of analyst reports and corporate disclosures. We employ a topic modeling methodology from computational linguistic research to compare the thematic content of a large sample of analyst reports issued promptly after earnings conference calls with the content of the calls themselves. We show that analysts discuss exclusive topics beyond those from conference calls and interpret topics from conference calls. In addition, we find that investors place a greater value on new information in analyst reports when managers face greater incentives to withhold value-relevant information. Analyst interpretation is particularly valuable when the processing costs of conference call information increase. Finally, we document that investors react to analyst report content that simply confirms managers' conference call discussions. Overall, our study shows that analysts play the information intermediary roles by discovering information beyond corporate disclosures and by clarifying and confirming corporate disclosures.

History: Accepted by Suraj Srinivasan, accounting.
Funding: A. Huang, A. Zang, and R. Zheng express thanks for financial support provided by the Hong Kong University of Science and Technology. R. Lehavy expresses thanks for financial support from the Harry Jones Endowment Fund.
Supplemental Material: The Internet appendix is available at https://doi.org/10.1287/mnsc.2017.2751.

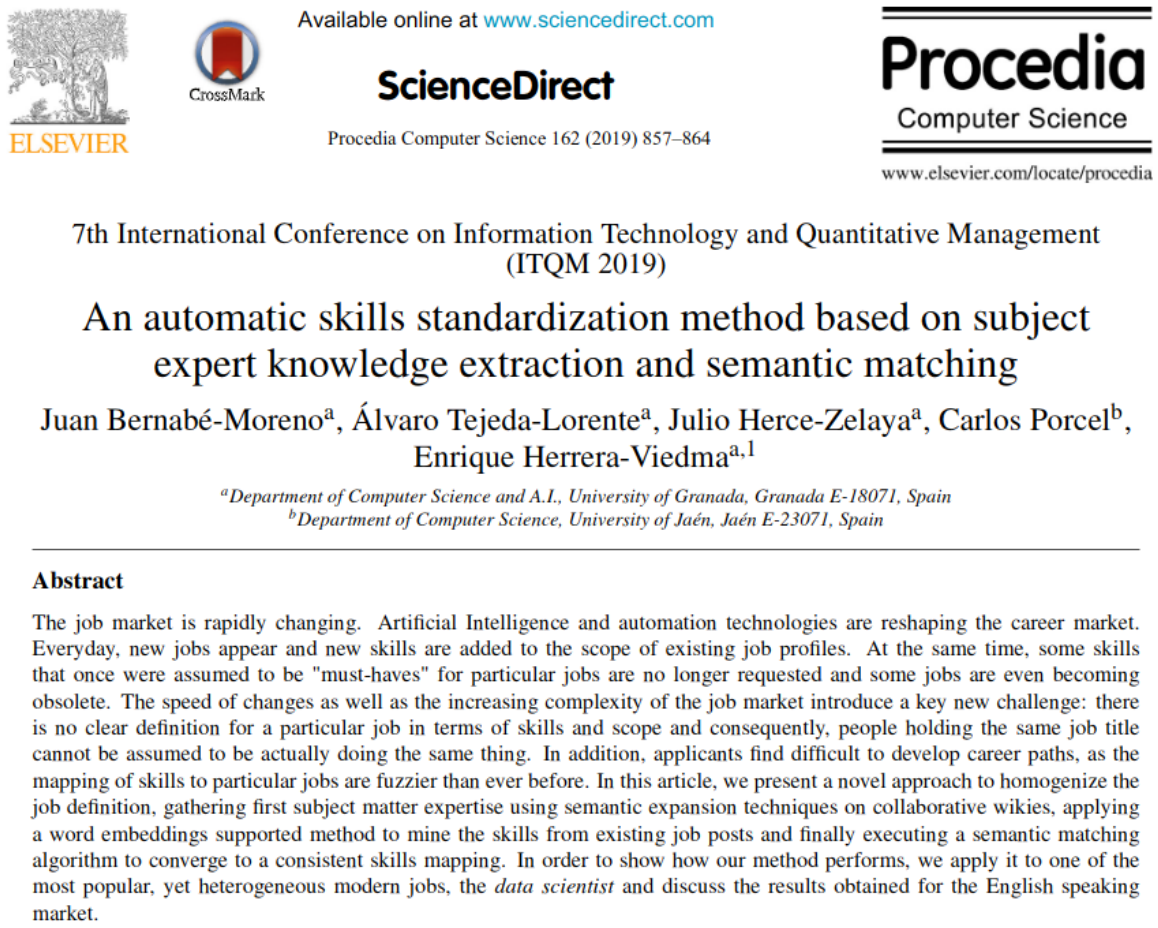
Keywords: analysts • discovery • interpretation • topic modeling • latent Dirichlet allocation

Information extraction

Documents contain structured information that can be complex to identify and retrieve.

Information extraction tools allow analysts to trace entities, events, dates, and relationships involving these elements.

Recently, information extraction tools have been used to recognise ad hoc entities such as ‘skills’ included in job posts.



ELSEVIER

CrossMark

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 162 (2019) 857–864

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Information Technology and Quantitative Management (ITQM 2019)

An automatic skills standardization method based on subject expert knowledge extraction and semantic matching

Juan Bernabé-Moreno[a], Álvaro Tejeda-Lorente[a], Julio Herce-Zelaya[a], Carlos Porcel[b], Enrique Herrera-Viedma[a,1]

[a]*Department of Computer Science and A.I., University of Granada, Granada E-18071, Spain*
[b]*Department of Computer Science, University of Jaén, Jaén E-23071, Spain*

Abstract

The job market is rapidly changing. Artificial Intelligence and automation technologies are reshaping the career market. Everyday, new jobs appear and new skills are added to the scope of existing job profiles. At the same time, some skills that once were assumed to be "must-haves" for particular jobs are no longer requested and some jobs are even becoming obsolete. The speed of changes as well as the increasing complexity of the job market introduce a key new challenge: there is no clear definition for a particular job in terms of skills and scope and consequently, people holding the same job title cannot be assumed to be actually doing the same thing. In addition, applicants find difficult to develop career paths, as the mapping of skills to particular jobs are fuzzier than ever before. In this article, we present a novel approach to homogenize the job definition, gathering first subject matter expertise using semantic expansion techniques on collaborative wikies, applying a word embeddings supported method to mine the skills from existing job posts and finally executing a semantic matching algorithm to converge to a consistent skills mapping. In order to show how our method performs, we apply it to one of the most popular, yet heterogeneous modern jobs, the *data scientist* and discuss the results obtained for the English speaking market.

Wrap-up

Main points

- There are two categories of NLP tools: operational and strategic tools.
- This module emphasises strategic tools such as ‘embeddings’, ‘sentiment analysis’, ‘span/text classification’, ‘topic modelling’, ‘information extraction’, which can significantly expand the decision making of business and financial analysts.