Advances in Intelligent Systems and Computing 1423

Yasufumi Takama · Naohiro Matsumura · Katsutoshi Yada · Mitsunori Matsushita · Daisuke Katagami · Akinori Abe · Hisashi Kashima · Toshihiro Hiraoka · Takahiro Uchiya · Rafal Rzepka *Editors*

Advances in Artificial Intelligence

Selected Papers from the Annual Conference of Japanese Society of Artificial Intelligence (JSAI 2021)

Springer

Advances in Intelligent Systems and Computing

Volume 1423

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Indexed by DBLP, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST).

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

More information about this series at https://link.springer.com/bookseries/11156

Yasufumi Takama · Naohiro Matsumura ·
Katsutoshi Yada · Mitsunori Matsushita ·
Daisuke Katagami · Akinori Abe ·
Hisashi Kashima · Toshihiro Hiraoka ·
Takahiro Uchiya · Rafal Rzepka
Editors

Advances in Artificial Intelligence

Selected Papers from the Annual Conference of Japanese Society of Artificial Intelligence (JSAI 2021)

Springer

Editors
Yasufumi Takama
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan

Katsutoshi Yada
Faculty of Business and Commerce
Kansai University
Osaka, Japan

Daisuke Katagami
Department of Applied Computer Science
Tokyo Polytechnic University
Atsugi, Kanagawa, Japan

Hisashi Kashima
Department of Intelligence Science
and Technology
Kyoto University
Kyoto, Japan

Takahiro Uchiya
Department of Computer Science
and Engineering
Nagoya Institute of Technology
Nagoya, Japan

Naohiro Matsumura
Graduate School of Economics
Osaka University
Toyonaka, Osaka, Japan

Mitsunori Matsushita
Faculty of Informatics
Kansai University
Takatsuki, Osaka, Japan

Akinori Abe
Division of Behavioral Science
Chiba University
Chiba, Japan

Toshihiro Hiraoka
Institute of Industrial Science
The University of Tokyo
Meguro-ku, Tokyo, Japan

Rafal Rzepka
Graduate School of Information Science
and Technology Sapporo Hokkaido Japan
Hokkaido University
Sapporo, Hokkaido, Japan

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-030-96450-4 ISBN 978-3-030-96451-1 (eBook)
https://doi.org/10.1007/978-3-030-96451-1

Preface

This book contains extended versions of research papers presented at the international sessions at the 35th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI2021), which was held online in June 2021. The JSAI annual conferences are considered key events for our organization, and the international sessions held at these conferences play a key role for the society in its efforts to share Japan's research on artificial intelligence with other countries. With the increased impact of AI on the world, the JSAI annual conferences have recently attracted many participants not only from academia but also from industries. JSAI2021 was held online, and various research ideas were shared through active discussion with many participants. We are extremely pleased to publish this collection of papers as the research results of our international sessions.

The topics of the international sessions cover five categories: knowledge engineering, machine learning, agents, robots and real worlds, and human interface and education aid, which have been important topics for the JSAI community. A total of 57 papers were selected for presentation at the international sessions, among which 35 papers submitted to a regular category were handled as the candidate papers for this post-proceedings. Each of the candidate papers was reviewed by at least two experts from the viewpoint of originality (the novelty of the paper), significance (impact on sciences and business), and quality of presentation. As a result of this review process, 25 papers were selected and invited to submit their extended version from papers presented at JSAI2021. After checking the quality of submitted extended versions of papers by JSAI2021 International Program Committee Members, 19 papers are included in this book. All of the papers are original and high quality, representing key contributions to AI research.

We would like to extend our deepest appreciation to President Itsuki Noda, Vice President Ryutaro Ichise (General Chair of JSAI2021), and Executive Committee Chair Takeo Hosomi, as well as to the JSAI administrative staff and Springer

publishing staff for their tremendous assistance on this project. We would also like to thank the authors of the papers contained in this book and all international session contributors.

Yasufumi Takama
Chief Editor

Mitsunori Matsushita
Program Chair, JSAI 2021

Naohiro Matsumura
Vice Program Chair, JSAI 2021

Daisuke Katagami
Takahiro Uchiya
Akinori Abe
Katsutoshi Yada
Rafal Rzepka
Toshihiro Hiraoka
Hisashi Kashima
International Program Committee, JSAI 2021

Organization

The 35th Annual Conference of the Japanese Society for Artificial Intelligence, 2021

Conference Chair

Ryutaro Ichise	National Institute of Informatics, Japan

Executive Committee Chairs

Takeo Hosomi	NEC Corporation, Japan
Hiroyuki Toda (vice)	NTT Corporation, Japan

Program Committee Chairs

Mitsunori Matsushita	Kansai University, Japan
Naohiro Matsumura (vice)	Osaka University, Japan
Yoko Nishihara (sub)	Ritsumeikan University, Japan

Program Committee

Tomohiro Harada	Tokyo Metropolitan University, Japan
Satoshi Yoshida	NEC Corporation, Japan
Munehiko Sasajima	University of Hyogo, Japan
Eiichi Sunagawa	Toshiba Corporation, Japan
Koh Mitsuda	NTT Corporation, Japan
Daniel Andrade	NEC Corporation, Japan
Ryo Iwaki	IBM Research, Tokyo, Japan
Yuki Nakaguchi	NEC Corporation, Japan
Hidekazu Oiwa	Google, USA
Takashi Nishibayashi	Voyage Group, Japan

Koh Takeuchi	Kyoto University, Japan
Kyota Tsutsumida	NTT Corporation, Japan
Toshikazu Samura	Yamaguchi University, Japan
Takahito Tomoto	Tokyo Polytechnic University, Japan
Kazuya Kakizaki	NEC Corporation, Japan
Keisuke Fujii	Nagoya University, Japan
Masatoshi Yamada	Tokoha University, Japan
Dai Hasegawa	Hokkai-Gakuen University, Japan
Kazuya Narita	Insight Tech, Japan
Junichi Ochiai	Mirai Share Co., Ltd., Japan
Yo Ehara	Tokyo Gakugei University, Japan
Yuki Ogawa	Ritsumeikan University, Japan
Tatsuya Otoshi	Osaka University, Japan
Shun Tanaka	Future Corporation, Japan
Masashi Okada	Panasonic Corporation, Japan
Miho Matsunagi	Yahoo Japan Corporation, Japan
Koichi Ota	Japan Advanced Institute of Science and Technology, Japan
Taiki Todo	Kyushu University, Japan
Masahiro Suzuki	University of Tokyo, Japan
Kento Koike	Tokyo Polytechnic University, Japan
Yotaro Fuse	Kansai University, Japan
Yasufumi Takama	Tokyo Metropolitan University, Japan
Daisuke Katagami	Tokyo Polytechnic University, Japan
Takahiro Uchiya	Nagoya Institute of Technology, Japan
Akinori Abe	Chiba University, Japan
Katsutoshi Yada	Kansai University, Japan
Rafal Rzepka	Hokkaido University, Japan
Toshihiro Hiraoka	University of Tokyo, Japan
Hisashi Kashima	Kyoto University, Japan
Tadahiko Murata	Kansai University, Japan
Yihsin Ho	Takushoku University, Japan
Lieu-Hen Chen	National Chi Nan University, Taiwan
Yuya Yoshikawa	Chiba Institute of Technology, Japan
Yuko Sakurai	National Institute of Advanced Industrial Science and Technology, Japan
Hirotoshi Taira	Osaka Institute of Technology, Japan

Contents

Machine Reading Comprehension of News on Stock Price Changes

Seiji Suzuki, Gakuto Tsutsumi, and Takehito Utsuro(✉)

Graduate School of Science and Technology, University of Tsukuba, 1-1-1, Tennodai, Tsukuba, Ibaraki 305-8573, Japan
utsuro@iit.tsukuba.ac.jp

Abstract. One of the reasons why stock prices fluctuate greatly is because of IR announcements, which are information for investors, and news reports on events that are closely related to the companies. When a stock price change occurs, news sites for investors may report the stock price change and the reason for the change. However, such articles report only a certain portion of overall events that are closely related to reasons for stock price changes. Thus, in order to provide investors with information on those reasons for stock price changes, it is necessary to develop a system to collect information on events that could be closely related to the stock price changes of certain companies from the Internet. As the first step towards developing such a system, this paper takes an approach of employing a BERT-based machine reading comprehension model, which extracts causes of stock price changes from news reports on stock price changes. Those extracted reasons are intended to be further used to train a system to collect information on events that could be closely related to the stock price changes of certain companies from the Internet.

1 Introduction

One of the reasons why stock prices fluctuate greatly is because of IR announcements, which are information for investors, and news reports on events that are closely related to the companies. When such information is delivered, as shown in Fig. 1, the trading volume of the companies increase and the a stock price change occurs. When a stock price change occurs, news sites for investors may report the stock price change and the reason for the change as shown in Fig. 2. However, such articles report only a certain portion of overall events that are closely related to reasons for stock price changes. Thus, in order to provide investors with information on those reasons for stock price changes, it is necessary to develop a system to collect information on events that could be closely related to the stock price changes of certain companies from the Internet. As the first step towards developing such a system, this paper takes an approach of employing a BERT [5]-based machine reading comprehension model [14], which extracts reasons for stock price changes from news reports on stock price changes

Y. Takama et al. (Eds.): JSAI 2021, AISC 1423, pp. 1–12, 2022.
https://doi.org/10.1007/978-3-030-96451-1_1

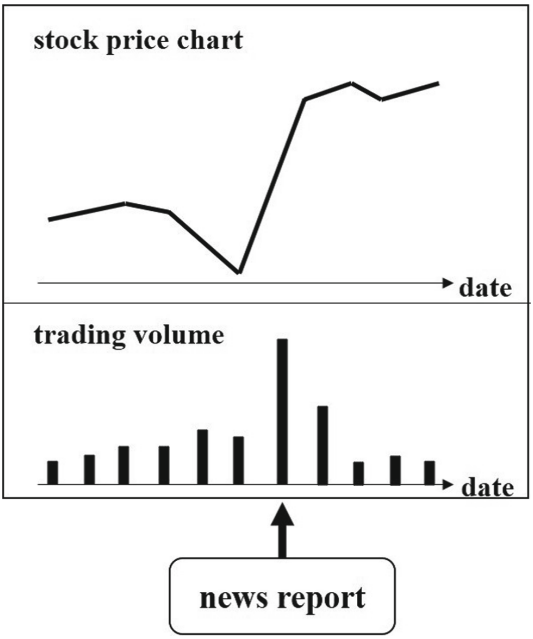


Fig. 1. Relation of stock price changes and trading volume per day and news report

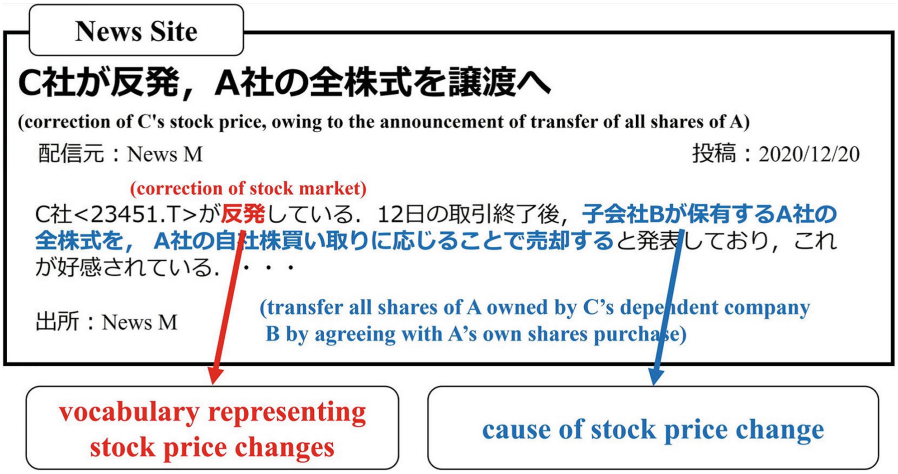


Fig. 2. The vocabulary representing stock price changes and a cause of stock price change: an example

(Fig. 3). Those extracted reasons are intended to be further used to train a system to collect information on events that could be closely related to the stock price changes of certain companies from the Internet.

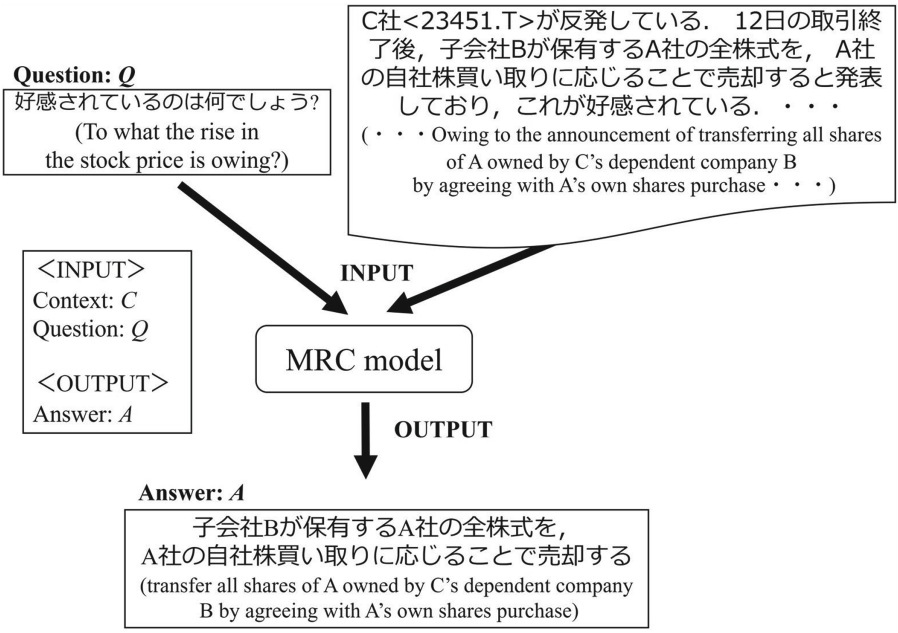
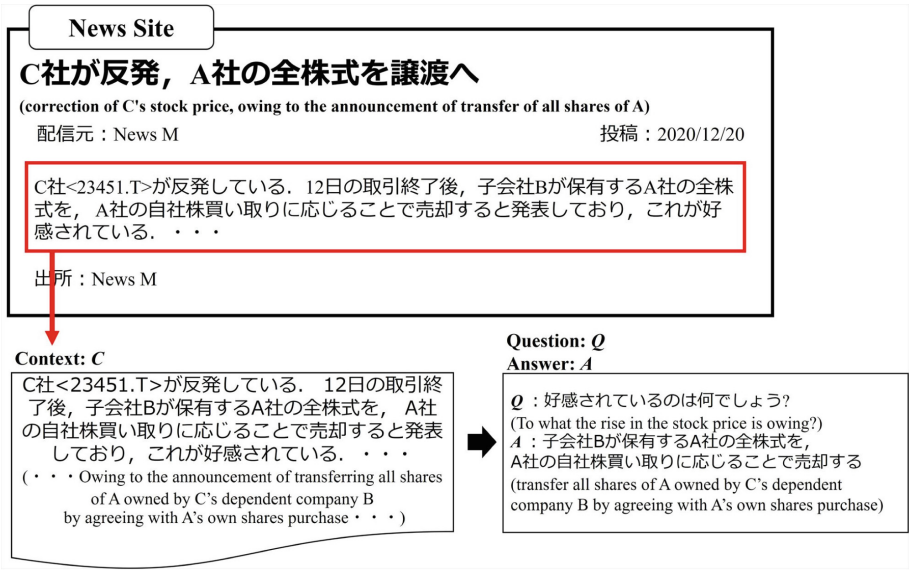


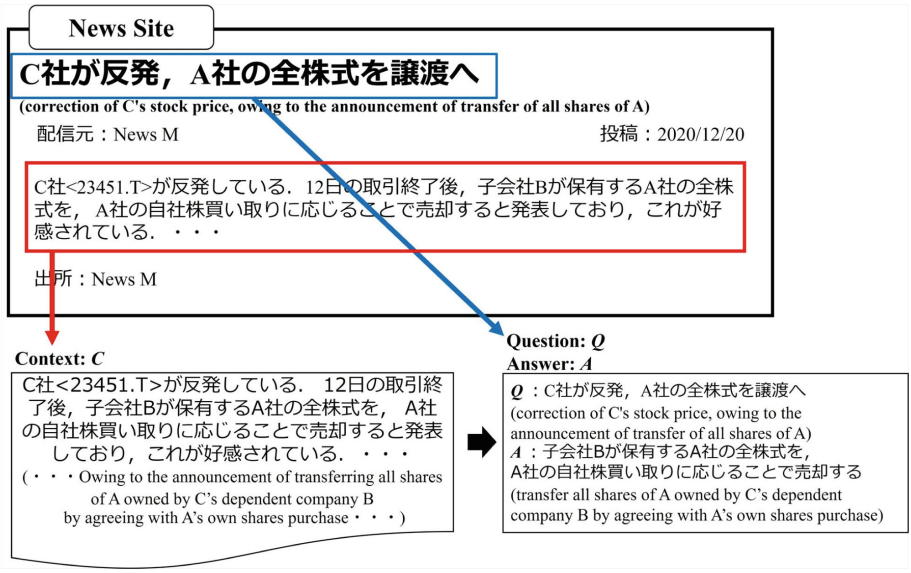
Fig. 3. The framework of machine reading comprehension of causes of stock price changes

The task of machine reading comprehension is formalized as below: given a question and a context text where the answer to the question is stated, the task is to extract the answer from the context text. The task of machine reading comprehension is divided into factoid and non-factoid machine reading comprehension sub-tasks. In factoid machine reading comprehension, it is known that the performance by a machine is over that by a human[1]. The task of machine reading comprehension of causes of stock price changes is regarded as a non-factoid machine reading comprehension sub-task. Out of such non-factoid machine reading comprehension sub-tasks, Chen et al. [4] studied the task of how-to tip machine reading comprehension. Chen et al. [4] explored a way to develop a dataset for training Japanese how-to tip QA models, and it applied BERT [5] to a Japanese how-to tip QA dataset. Evaluation results of Chen et al. [4] revealed that the how-to tip machine reading comprehension performance was almost comparative with that of the factoid machine reading comprehension even with the training data size reduced to around 4% of the factoid machine reading comprehension.

[1] When evaluated with the SQuAD [14] (https://rajpurkar.github.io/SQuAD-explorer/).



(a) When the Question is Manually Generated



(b) When the Title of the Article is used as the Question

Fig. 4. The procedure of developing an example of machine reading comprehension of causes of stock price changes

Table 1. Statistics of the Categories of 100 Articles delivered from "みんなの株式 (minna-no kabu-shiki)"

Category	# of articles
News on stock price changes and their causes	25
News on companies such as the announcements on New products	7
News on domestic equities	24
News on foreign equities	9
News on exchange market	7
News on bond market	2
News on financial market	2
News for individual investors	21
Economic indicator	3
Total	100

2 Stock Price News of "みんなの株式 (minna-no kabu-shiki)"

In this paper, the news site from which we collect the stock price news is minkabu.jp[2], where we used 3,300 articles[3] delivered from "みんなの株式 (minna-no kabu-shiki)." minkabu.jp includes about 260,000 articles (as of January 2021) that are delivered from "みんなの株式　(minna-no kabu-shiki)". Table 1 shows the statistics of the categories of 100 articles randomly sampled from the collected 3,300 articles. Based on this statistics, out of the overall 260,000 articles delivered from "みんなの株式　(minna-no kabu-shiki)", it is estimated that the number of articles on "news on stock price changes and their causes" amount to 65,000 (25%). Thus, "みんなの株式　(minna-no kabu-shiki)" can be considered as a resource with a sufficient number of articles for developing a dataset for the examples of machine reading comprehension of causes of stock price changes.

Out of the "rise" and "decline" of stock prices, this paper focuses on "rise" of stock prices and develop a dataset of examples of machine reading comprehension of causes of stock price changes. First, we collect 11 (and several more) vocabularies listed in Table 2 representing "rise" in stock prices and randomly selected 627 articles that include one of those "rise" vocabularies, out of the overall 3,300. This is the result of discarding 14 articles that are not appropriate for developing examples of machine reading comprehension of causes of stock price changes. This is also the result of discarding 35 articles including another 13 vocabularies representing "decline" in stock prices. Those 13 vocabularies include

[2] https://minkabu.jp/.
[3] Delivered from October 30th to December 3rd, 2020.

“反落 (reactionary fall)”, “下落 (decline)”, “続落 (continued fall)”, “急落 (fall rapidly)”, “売りに押され (drop)”, and “安値 (low price)”.

From those 627 articles, we develop the dataset of examples of machine reading comprehension of causes of stock price changes.

Table 2. # of Occurrences and their Ratio (%) of Vocabularies representing Rise in Stock Prices among 627 examples of machine reading comprehension of causes of stock price changes

Vocabulary	#	Ratio
反発 (correction)	176	19.7
続伸 (continued rise)	172	19.3
高値 (high price)	115	12.9
カイ気配 (bid price)	87	9.7
大幅高 (large rise)	66	7.4
上昇 (rise)	56	6.3
ストップ高 (hit limit high)	54	6.0
急伸 (rise rapidly)	49	5.5
連騰 (winning streak)	40	4.5
堅調 (increase steadily)	38	4.3
急騰 (sharp rise)	35	3.9
other	5	0.5
total	893	100

3 Generating Examples of Machine Reading Comprehension of Causes of Stock Price Changes from Stock Price News Articles

In the procedure of generating examples of machine reading comprehension of causes of stock price changes from stock price news articles, we take the following two approaches: (a) The question manually generated from the article is used as the question Q of the machine reading comprehension. (b) The title of the article is used as the question Q of the machine reading comprehension. In both cases, the whole article is used as the context C of the machine reading comprehension.

(a) This case is described in Fig. 4(a). The question manually generated from the article is used as the question Q of the machine reading comprehension.
(b) This case is described in Fig. 4(b). The title of the article is used as the question Q of the machine reading comprehension.

Table 3. # and Ratio of Manually Generated Questions Q among 627 Examples of Machine Reading Comprehension of Causes of Stock Price Changes

Question Q	#	ratio
材料視されているのは何でしょう? (What is the good offer in front of the market?)	130	20.7
好感されているのは何でしょう? (To what the rise in the stock price is owing?)	122	19.5
寄与したのは何でしょう? (What contributes to the rise in the stock price?)	61	9.7
要因は何でしょう? (What causes the rise in stock prices?)	48	7.7
発表したのは何でしょう? (What is announced?)	38	6.1
好調なのは何でしょう? (What is fare well?)	29	4.6
利益を押し上げたのは何でしょう? (What boosted the profits?)	25	4.0
期待されているのは何でしょう? (What is expected regarding the rise in the stock price?)	14	2.2
背景は何でしょう? (What is the background of the rise in the stock price?)	12	1.9
other (each # $\leq$ 10)	148	23.6
合計	627	100

Within sentences including the answer span, a vocabulary indicating the "cause" of the rise of the stock price (henceforth referred to as the "cause" vocabulary) usually appears. The overall number of those "cause" vocabularies is 62, which include the followings:

> "材料視 " (good offer in front of the market), "好感 " (owing to), "期待 " (expect), "発表 " (announce), "株価を刺激 " (motivates the rise in stock prices), "要因 " (cause), "引き上げ " (raised), and "追い風 " (positive).

The case (a) is quite straightforward in that we manually generate question sentences from those "cause" vocabularies as shown in the statistics of Table 3[4]. When beyond the machine reading comprehension setting, where only the question Q is available and the candidates of context C have to be automatically collected from a large pool of documents [3,9], we first identify those "cause"

[4] As an alternative, we also examine the approach of using a single general question sentence such as below:「A 社の株価が変動した理由はなんでしょう？」However, this approach does not perform well in the evaluation compared with the case (a) in this section.

vocabularies within the collected candidates of contexts, then the question Q is automatically switched according to the identified "cause" vocabulary and then the machine reading comprehension model is applied to the tuple of the question Q and the context C. The case (b), on the other hand, has an advantage in that we can omit the procedure of manually generating the question from the article.

4 Evaluation

4.1 Evaluation Procedure

As the version of BERT [5] implementation which can handle a text in Japanese, the TensorFlow version[5] and the Multilingual Cased model[6] were used as the pre-trained model. Before applying BERT modules, MeCab[7] was applied with mecab-ipadic-NEologd dictionary[8] and the Japanese text was segmented into a morpheme sequence. Then, within the BERT fine-tuning module, the WordPiece module with 110k shared WordPiece vocabulary was applied, and the Japanese text was further segmented into a subword unit sequence. Finally, the BERT fine-tuning module for machine reading comprehension[9] was applied as well as the fine-tuned model. In the fine-tuning procedure, the BERT pre-trained model was first fine-tuned with 12,000 factoid machine reading comprehension training examples[10], and then, the fine-tuned model was further fine-tuned with the training examples of machine reading comprehension of causes of stock price changes developed in the previous section.

4.2 Evaluation Results

In the evaluation, we examined the following three cases:

(a) This case is described in Fig. 4(a) and as the case (a) in Sect. 3.
(b) This case is described in Fig. 4(b) and as the case (b) in Sect. 3.
(c) The question Q is the same as the case (a). The title of the article is concatenated to the context of the cases (a) and (b) and then used as the context C.

Figure 5 shows the evaluation results of 5-fold cross validation. Although the case (b) is beneficial in that it can avoid the manual procedure of generating a question from the article, the best performing ones are the case (a) or the case (c), where the performance of the case (b) is about 76% for the exact match,

[5] https://github.com/google-research/bert.
[6] Trained with 104 languages, available from https://github.com/google-research/bert/blob/master/multilingual.md.
[7] http://taku910.github.io/mecab/ (in Japanese).
[8] https://github.com/neologd/mecab-ipadic-neologd.
[9] run_squad.py, with the number of epochs as 2, batch size as 8, and learning rate as 0.00003.
[10] http://www.cl.ecei.tohoku.ac.jp/rcqa/ (in Japanese).

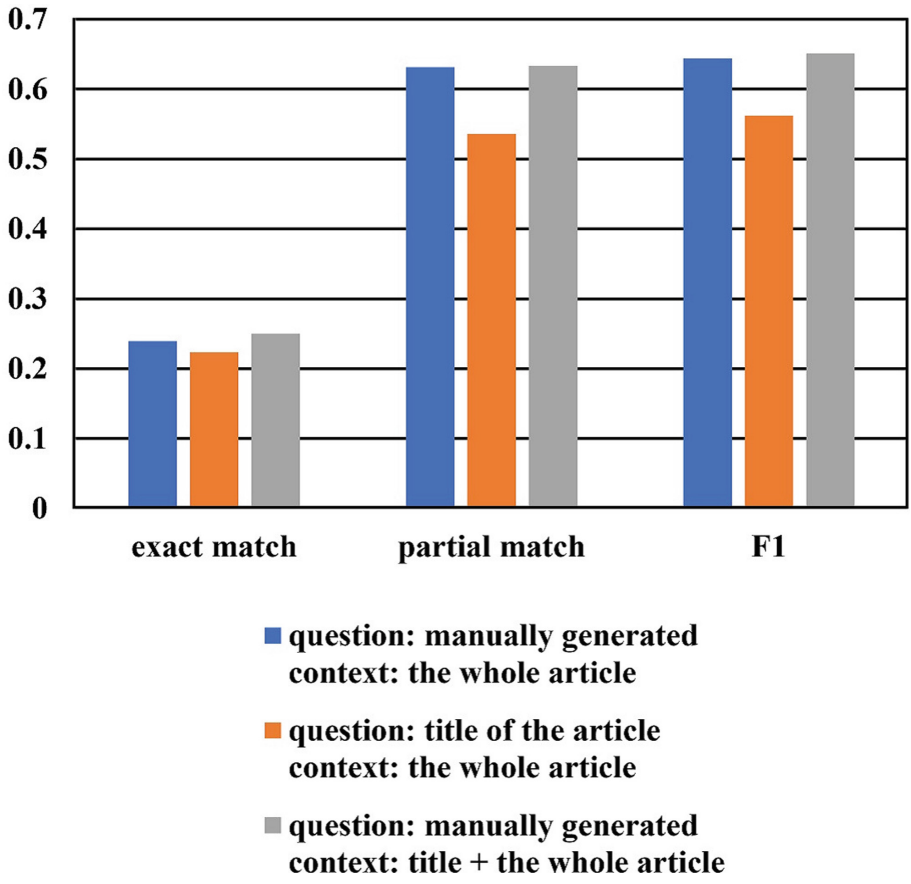


Fig. 5. Evaluation results

53% for the partial match, and 78% for the F1 score. For the case (b), we also examine the difference between the title of the article and the answer span. Then, out of the overall 627 examples for evaluation, for 496 examples (about 80 %), answer span does include additional information that are not included in the title. So, this means that for some of those 80% cases, hopefully the model of the case (b) successfully identifies some of those additional answer spans that are not included in the title of the article without manually creating the questions. Actually, as the result of manually examining the exact and partial matches of the case (b), for about 70–90% of the exact and partial matches, we observe that the model successfully identify those additional answer spans that are not included in the question.

5 Related Work

Chen et al. [4] reported that, in the framework of the fine-tuning of BERT [5], non-factoid how-to tip machine reading comprehension model fine-tuned with 2,576 training examples outperformed factoid machine reading comprehension model fine-tuned with 12,000 training examples (see footnote 10). Based on Chen et al. [4], we are planning to develop 1,000 training examples including "rise" vocabularies, 1,000 including "decline" vocabularies, and another 1,000 including both "rise" and "decline" vocabularies, and then fine-tune and evaluate the machine reading comprehension model of causes of stock price changes.

As other related work, Sakai et al. [15] studied to extract causal and complementary information for generating market analysis comments from articles describing the market conditions of the Nikkei Stock Average, where the proposed method achieved the 90% precision and 61% recall. Compared to Sakai et al. [15], we focus on the machine reading comprehension of the causes of the stock price changes of individual issues rather than the market analysis comments. It is also interesting to apply the approach that is similar to Sakai et al. [15] in our task focusing on individual issues.

In additon, Liu et al. [10] also studied the issue of pre-trained financial language model for financial text mining, where the task of general question answering in the financial domain is examined. They studied FiQA[11] Task 2 "Opinion-based QA over financial data". Compared to Liu et al. [10], this paper studies the issue of answering the causes of the rise of stock price, but not the general question answering in the financial domain. Also, Mariko et al. [11] organized the Financial Document Causality Detection Shared Task (FinCausal 2020), where the tasks such as detection of causes and effects in the general financial domain are studied [1,2,6–8,12,13,16]. This paper, on the other hand, concentrates on the issue of answering the causes of the rise of stock price, but not on the causes and effects in the general financial domain.

6 Conclusion

This paper took an approach of employing a BERT [5]-based machine reading comprehension model [14], which extracts reasons for stock price changes from news reports on stock price changes. Overall, the approach of using the title of the article as the question Q of the machine reading comprehension is recommended in that the procedure of manually generating the question from the article can be omitted, although its performance is slightly lower than the approach of manually creating questions. Future work includes incorporating not only the "rise" of stock prices but also the "decline" of them. Another future work is to increasing the number of training examples and to examine whether those increased training examples contribute to improving the machine reading comprehension performance. It also includes scaling up into beyond the machine reading comprehension setting where only the question Q is available and the

[11] https://sites.google.com/view/fiqa/home.

candidates of context C have to be automatically collected from a large pool of documents [3,9].

References

1. Becquin, G.: GBe at FinCausal 2020, task 2: span-based causality extraction for financial documents. In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 40–44 (2020)
2. Chakravarthy, S., Kanakagiri, T., Radhakrishnan, K., Umapathy, A.: Domino at FinCausal 2020, task 1 and 2: causal extraction system. In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 90–94 (2020)
3. Chen, D., Fisch, A., Weston, J., Bordes, A.: Reading Wikipedia to answer open-domain questions. In: Proceedings of 55th ACL, pp. 1870–1879 (2017)
4. Chen, T., Li, H., Kasamatsu, M., Utsuro, T., Kawada, Y.: Developing a how-to tip machine comprehension dataset and its evaluation in machine comprehension by BERT. In: Proceedings of 3rd FEVER, pp. 26–35 (2020)
5. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional Transformers for language understanding. In: Proceedings of NAACL-HLT, pp. 4171–4186 (2019)
6. Imoto, T., Ito, T.: JDD @ FinCausal 2020, task 2: financial document causality detection. In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 50–54 (2020)
7. Ionescu, M., Avram, A.M., Dima, G.A., Cercel, D.C., Dascalu, M.: UPB at FinCausal-2020, tasks 1 & 2: causality analysis in financial documents using pre-trained language models. In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 55–59 (2020)
8. Kao, P.W., Chen, C.C., Huang, H.H., Chen, H.H.: NTUNLPL at FinCausal 2020, task 2: improving causality detection using Viterbi decoder. In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 69–73 (2020)
9. Lee, K., Chang, M.W., Toutanova, K.: Latent retrieval for weakly supervised open domain question answering. In: Proceedings of 57th ACL, pp. 6086–6096 (2019)
10. Liu, Z., Huang, D., Huang, K., Li, Z., Zhao, J.: FinBERT: a pre-trained financial language representation model for financial text mining. In: Proceedings of 29th IJCAI, pp. 4513–4519 (2020)
11. Mariko, D., Abi-Akl, H., Labidurie, E., Durfort, S., De Mazancourt, H., El-Haj, M.: The financial document causality detection shared task (FinCausal 2020). In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 23–32 (2020)
12. Özenir, G., Karadeniz İ.: ISIKUN at the FinCausal 2020: linguistically informed machine-learning approach for causality identification in financial documents. In: Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 85-89 (2020)
13. Pielka, M., et al.: Fraunhofer IAIS at FinCausal 2020, tasks 1 & 2: using ensemble methods and sequence tagging to detect causality in financial documents. In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 64–68 (2020)

14. Pranav, R., Jian, Z., Konstantin, L., Percy, L.: SQuAD: 100,000+ questions for machine comprehension of text. In: Proceedings of EMNLP, pp. 2383–2392 (2016)
15. Sakai, H., Sakaji, H., Izumi, K., Matsui, H., Irie, K.: Extraction of causal and complementary information for generating market analysis comments by automatic generation of training data. In: Proceedings of 34th Annual Conference JSAI (2020). (in Japanese)
16. Szántó, Z., Berend, G.: ProsperAMnet at FinCausal 2020, task 1 & 2: modeling causality in financial texts using multi-headed transformers. In: Proceedings of 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp. 80–84 (2020)

A Deep Fusion Model Combining News Content and Historical Prices for Stock Trend Prediction

Jenq-Haur Wang(✉) and Hsin-Wen Liu

National Taipei University of Technology, Taipei, Taiwan
jhwang@ntut.edu.tw

Abstract. Stock markets are often influenced by various factors which makes it very challenging to predict. Machine learning and deep learning models are often used to predict stock trends from its historical prices. Since there's a lot of online information available in addition to stock prices, including technical indicators, news reports, and social information, we intend to combine news content for improving the performance. In this paper, we propose a deep fusion model for stock trend prediction combining news content with historical stock prices. Firstly, we utilize multi-layered Long Short-Term Memory (LSTM) to learn sequential information from stock prices. Then, we adopt Hybrid Attention Networks (HAN) which include both sentence-level and temporal attention to discover the relative importance of words from news reports. Finally, we compare early and late fusion models to improve stock trend prediction. The experimental results show that the best macro-F1 score of 79.0% can be achieved when we use late fusion to aggregate the prediction results of news content using 2-layer LSTM and that of historical prices in a 5-day window using HAN. As compared to individual models, the performance improvement of up to 40% can be obtained. This shows the potential of our proposed approach. Further investigation is needed for stock trend prediction in different markets.

Keywords: Stock trend prediction · Deep learning · Semantic analysis

1 Introduction

Stock markets are usually influenced by many different factors, including company revenues, industry growth, global economics, and political situations, to name a few. This makes it very challenging to predict the trend. Conventionally, it's common to conduct technical analysis based on stock prices and basic analysis based on company revenues for stock trend prediction in the industry. Machine learning and deep learning models are often used to predict stock trends given huge amount of data. Nowadays, there's a lot more online information available in addition to stock prices and technical indicators, including company revenues, news reports, and social information. Among these types of information, we intend to explore the effects of news content on stock trend prediction, since news reports might contain useful insights into the developments and perspectives of the corresponding company or industry. Given the rapid changes

Y. Takama et al. (Eds.): JSAI 2021, AISC 1423, pp. 13–26, 2022.
https://doi.org/10.1007/978-3-030-96451-1_2