Economic Causal Chain Presentation System Considering Output Diversity to Search Economic Ripple Effects

Ryotaro Kobayashi
The University of Tokyo
Tokyo, Japan
b2022rkobayashi@socsim.org

Yuri Murayama
The University of Tokyo
Tokyo, Japan
murayama.yuri@sys.t.u-tokyo.ac.jp

Kiyoshi Izumi
The University of Tokyo
Tokyo, Japan
izumi@sys.t.u-tokyo.ac.jp

ABSTRACT

This study focuses on discovering the ripple effects on firms of significant external events such as pandemics, resource price spikes, and natural disasters. For this purpose, we constructed a system that presents a chain of other economic events derived from one event by extracting descriptions of causal relationships from a large amount of text data. The objective of this study is to construct a causal chain presentation system that takes into account the diversity of outputs to discover broader economic ripple effects. We developed a new algorithm that uses Maximal Marginal Relevance to represent causal chains. Evaluation experiments using financial statement summaries confirm that this approach produces a greater variety of outputs without sacrificing accuracy.

CCS CONCEPTS

• **Information systems → Decision support systems**; **Information retrieval**; **Information extraction**.

KEYWORDS

causal network, causality extraction, financial text mining

ACM Reference Format:
Ryotaro Kobayashi, Yuri Murayama, and Kiyoshi Izumi. 2023. Economic Causal Chain Presentation System Considering Output Diversity to Search Economic Ripple Effects. In *Proceedings of The 4th Workshop on Knowledge Discovery from Unstructured Data in Financial Services (SIGIR '23 KDF).* ACM, New York, NY, USA, 9 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

The economic activities of companies are significantly affected by external events such as pandemics, spikes in resource prices, and natural disasters. However, the impact of external events on financial markets is not fixed. The COVID-19 pandemic, for example, has thrust numerous small and medium-sized enterprises (SMEs), especially those in the food service and lodging industries, into unpredictable circumstances, leading to reduced performance and cash flow difficulties.

Not only for financial market analysis but also for policy-making, it is imperative to comprehend the ripple effects of such events to pinpoint the enterprises that need assistance accurately. The 2011 Great East Japan Earthquake demonstrated this need, as its impact was felt by companies not directly affected by the disaster, resulting in disrupted supply chains, reputational damage from the nuclear accident, decreased consumer spending, and loss of sales channels.

The outcomes of diverse events within the socioeconomic landscape are shaped by individuals' perceptions of the causal factors and the actions they decide to undertake based on these perceptions. Understanding the causal relationship among these events is critical for analyzing their ripple effects. In this study, we analyze a large volume of text data about the economy, automatically extracting descriptions of perceived cause-effect relationships. This approach enables us to construct a database reflecting the causal relationships among various economic events. Subsequently, we establish a system that uses specific event keywords as input, presenting a network of related economic events representing causal chains. Identifying these causal links among different economic events, facilitated by social and cognitive causal relations gleaned from the text, allows for recognizing ripple effects. This understanding of the effects passed down to companies due to socioeconomic changes is essential for decision-making by market analysts and policymakers.

More than focusing on the accuracy of a causal chain search system's output is required to discover ripple effects. The ripple effects under investigation should encompass a broad range of topics. Outputs leaning heavily towards a specific topic could cause cognitive biases and result in judgment errors. Consequently, this research aims to develop a system for retrieving causal chains that consider output diversity, facilitating the discovery of a broader spectrum of economic ripple effects. To this end, we devised a novel algorithm for presenting causal chains, utilizing the Maximal Marginal Relevance (MMR) [5].

The main contributions of this research are as follows:

- We developed a method to search for economic causal chains from textual data (unstructured data) considering the output diversity.
- This method enables us to predict possible ripple effects of external events by tracing a causal sequence from a text representing an event of interest.
- The proposed method prevents the results of the causal chain search from being biased toward only specific topics and enables the search for a broader range of ripple effects.

2 RELATED WORK

In recent years, much research concerning causal information extraction from natural language has been based on neural networks. For example, Dasgupta et al. [6] proposed a method for extracting causal information using Long short-term memory (LSTM)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SIGIR '23 KDF, July 27, 2023, Taipei, Taiwan
© 2023 Association for Computing Machinery.
ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

architecture. Furthermore, concerning English causal information extraction, various methods were proposed in the Financial Narrative Processing Workshops (FNP), which is a workshop of Colling 2020 because it is included in the Shared Task FinCausal 2020 workshop. Two types of tasks were set in the shared task: extracting causal information sentences and extracting causal-effect expressions from causal information sentences. Most proposed methods for extracting causal information sentences are based on BERT consisting of Transformer and achieved high performance [7, 8, 10]. The BERT-based method has also been proposed in the extraction causal expressions task [9].

Studies concerning the construction of causal chains, such as Ishii et al. [11], Alashri et al. [3], and Zhao et al. [21] exist. Ishii et al. proposed constructing causal networks by extracting causal expressions from newspaper articles and combining them with SVO based on the hypernym-hyponym relation dictionary. Alshri et al. proposed a concept-based causal chain construction method. Zhao et al. proposed a method for constructing causal chains that considers the cause-to-effect and effect-to-cause paths. In addition, various applications of causality other than building causal chains are expected. For example, in the world of robotics, Causal World [2], a new benchmark that considers causality has been proposed.

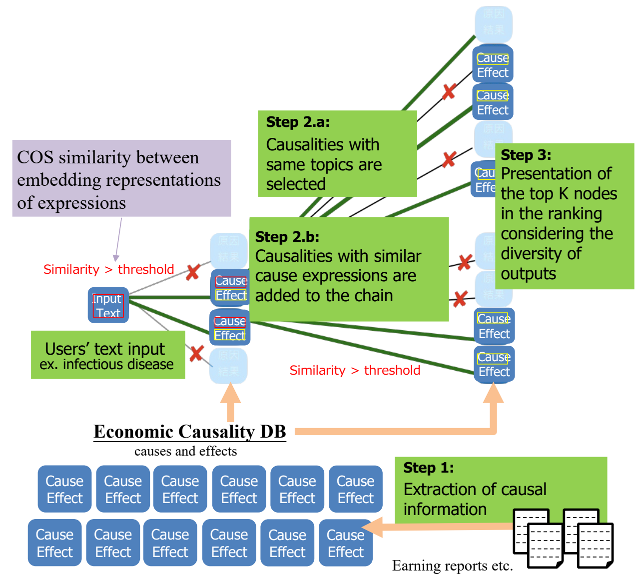

Figure 1: Overview of the causal chain presentation system.

3 METHOD

3.1 Overview

Figure 1 provides a schematic representation of the causal chain presentation system. Building the causal chain presentation system can be summarized in three main steps.

Step.1 Establishes a database of causal relations in economic events by identifying sentences containing causal relations in texts and extracting expressions indicating causes (cause expressions) and effects (effect expressions) as expression pairs.

Step.2 Constructs causal chains of economic events as a network by connecting nodes, pairs of cause-and-effect expressions, with a certain degree of similarity in a chain-like manner.

Step.3 Ranks the candidate nodes (pairs of cause-and-effect expressions) added in the previous step and presents the top K nodes considering the diversity of outputs.

The following sections describe each step.

3.2 Establishment of Causal Relation Database

Firstly, we require technology to extract human-recognized causal relationships from text data. We focus on phrases that provide clues to causal relationships and use the BERT-based machine learning method to determine whether sentences with clue phrases contain causal relationships [14]. Then, we extract pairs of cause-and-effect expressions based on syntactic patterns [17].

This research analyzes Japanese economic texts containing such relationships and builds a causality database specific to the economic domain. We extracted causal relationships from financial statement summaries, which listed companies regularly publish to disclose their business performance and financial status.

- Text data used: Approximately 20,000 financial summary texts were issued by approximately 2,300 companies between October 2012 and September 2022.
- Extracted causal relationships: 889,771

The causal relationships extracted were stored in the database, accompanied by the date of the financial summary's issuance and the entity that issued it.

3.3 Construction of Causal Chains

From the causal relation database established in the previous section, we construct causal chains by linking nodes, which are pairs of cause-and-effect expressions with a certain degree of similarity to form a tree structure-type network. Each node is considered to represent a specific economic event. When computing node similarity, we compare the effect expression of one node with the cause expression of another node. This enables linking other economic events (nodes) that extend from one event, creating a causal sequence of economic events based on recognized causal relationships.

Initially, a node with an empty cause expression is prepared using the user-entered keywords as the effect expression. The first network expansion is then performed with this node as the terminal node. We construct causal chains by repeating the following procedures N times; users can specify the number of times.

Step.2.a Selects nodes in the database that contain a cause expression with the same topic with the effect expression of the terminal node as possible nodes which has a causal relationship.

Step.2.b From the nodes selected in the previous step, adds nodes whose similarity with the effect expression of the terminal node based on the word embedding representations is greater than a threshold to the causal chain.

In Step 2.a, firstly, a topic analysis using Latent Dirichlet Allocation (LDA) [4] is performed on the nouns in all the cause-and-effect expressions in the database of causal relations to extract a determined number of topics. Secondly, the topic distribution is calculated for each node's cause/effect expression, and topics with an affiliation probability greater than a threshold are assigned to each expression. Each cause/effect expression may have multiple topics or no topics. Then, nodes in the database that has the cause expression with one or more topics in common with the topic of the effect expression of the terminal node are determined to be possible nodes added to the causal chain.

In Step 2.b, the nodes (pairs of cause-and-effect expressions) are considered similar if $Similarity$ is greater than or equal to a threshold θ, and $Similarity$ is calculated as follows using a configurable parameter ω that indicates the degree of context-sensitivity consideration.

$$Similarity = (1-\omega)\cdot\cos(\bar{\boldsymbol{w}}_{\boldsymbol{cause}},\ \bar{\boldsymbol{w}}'_{\boldsymbol{effect}}) + \omega\cdot\cos(\bar{\boldsymbol{w}}_{\boldsymbol{cause}},\ \bar{\boldsymbol{w}}'_{\boldsymbol{cause}}) \tag{1}$$

where $\bar{\boldsymbol{w}}_{\boldsymbol{cause}}$ is the average vector of embedding representations calculated by the word2vec model [16] for the nouns in the cause expression of the nodes to be judged. $\bar{\boldsymbol{w}}'_{\boldsymbol{cause}}$ is the average vector of the cause expression of the terminal node and $\bar{\boldsymbol{w}}'_{\boldsymbol{effect}}$ is the average vector of the effect expression of the terminal node in network expansion. Besides, $\cos(\bar{\boldsymbol{w}},\bar{\boldsymbol{w}}')$ computes the cosine similarity of two vectors as follows:

$$\cos(\bar{\boldsymbol{w}},\bar{\boldsymbol{w}}') = \frac{\bar{\boldsymbol{w}}\cdot\bar{\boldsymbol{w}}'}{\|\bar{\boldsymbol{w}}\|\|\bar{\boldsymbol{w}}'\|} \tag{2}$$

In Equation 1, $Similarity$ is calculated by adding not only the similarity between the cause-effect expressions in two nodes, but also the similarity between the cause-cause expressions in two nodes by the ratio of the context-sensitivity ω, thus considering the context leading up to the event indicated in the effect expression in new nodes.

3.4 Causal Chains Presentation Considering Diversity

In this study, we apply Maximal Marginal Relevance (MMR) [5], a method for generating diversity-aware rankings in information retrieval and recommendation systems, to construct a new causal chain presentation algorithm. Using this algorithm, we can improve the diversity of the top K nodes presented to users in our causal chain presentation system.

MMR is a Re-Ranking type algorithm that re-ranks the list of candidate items generated based on the user's queries or preferences, thereby increasing the $Diversity$ while maintaining accuracy. For a list of candidate items C generated based on a query or user interest, the list of items R returned as the final output result is determined by greedily adding items from C that maximize the following objective function.

$$f_{obj}(i,R) = (1-\alpha)\cdot rel(i) + \alpha\cdot\frac{1}{|R|}\sum_{j\in R} dist(i,j) \tag{3}$$

where α is a parameter that represents the degree of diversity consideration and $dist(i,j)$ is the distance (dissimilarity) between items i and j. The MMR method maintains a good fit $rel(i)$ with the user's query or preferences while adding items to R such that the variation among items increases according to a configurable parameter α, considering the output's diversity.

In the following, we describe a new causal chain presentation algorithm incorporating the MMR technique. Let C be the list of candidate cause-effect pairs (nodes) obtained as a result in 3.3 section, and R be the list of nodes that are finally presented to the user. The output, R, is obtained by computing the following for the set of last terminal nodes V that have been obtained by the previous iteration of network expansion.

$$R = \sum_{n\in V} R_n \tag{4}$$

where R_n is the top K nodes connected to and presented to the nth terminal node.

The algorithm for causal chain presentation using MMR performs the process shown in Figure 2 for each terminal node. Here, C_n is the list of candidate nodes at the nth terminal node, and the algorithm greedily adds nodes that maximize the objective function $f'_{obj}(i,R_n)$ among C_n, generating a diversity-aware ranking. The objective function $f'_{obj}(i,R_n)$ is calculated as follows:

$$f'_{obj}(i,R_n) = (1-\alpha)\cdot Similarity + \alpha\cdot\frac{1}{|R_n|}\sum_{j\in R_n} dist(i,j) \tag{5}$$

where $Similarity$ is computed as Equation 1. In this study, the distance function $dist(i,j)$ is calculated as follows:

$$dist(i,j) = 1 - \cos(\bar{\boldsymbol{w}}^{\boldsymbol{i}}_{\boldsymbol{effect}},\ \bar{\boldsymbol{w}}^{\boldsymbol{j}}_{\boldsymbol{effect}}) \tag{6}$$

where $\bar{\boldsymbol{w}}^{\boldsymbol{i}}_{\boldsymbol{effect}}$ is the average vector of the embedding representations of the i-th connected node and $\cos(\bar{\boldsymbol{w}},\bar{\boldsymbol{w}}')$ is the cosine similarity of the two vectors calculated as Equation 2.

Algorithm : **Causal-chain presentation using MMR**

Input: K; a set of candidate nodes C_n, s.t. $|C_n| > K$
Result: result list R_n, s.t. $|R_n| = K$

```
  R_n ← φ
  while |R_n| < K do
    i* ← argmax_{i ∈ C_n \ R_n} f'_obj(i, R_n)
    R_n ← R_n ∪ {i*}
  end while
```

Figure 2: Algorithm for causal chains presentation using Maximal Marginal Relevance (MMR) for each terminal node.

When the parameter α, representing the degree of diversity consideration, is 0.00, the second term on the right side of Equation 5 is ignored, and the ranking algorithm of the proposed method is equivalent to the "present in order of $Similarity$" algorithm of the existing method [12].

4 EXPERIMENTS

In this chapter, we describe an evaluation experiment for a system presenting diverse causal chains developed using a database built from causal relationships extracted from Japanese financial statement summaries to depict economic event sequences.

4.1 Pre-experiment: Determining the Degree of Diversity Consideration

In this section, we implement the causal chain presentation algorithm discussed in the previous section to Sakata et al.'s dataset [18]. We aim to assess whether the algorithm can generate diverse outputs without sacrificing accuracy by examining the trade-off between output accuracy and diversity evaluation. Additionally, we explore the optimal value for the degree of diversity consideration α in Equation 5. The annotated dataset facilitates automatic accuracy computation and experiments with varying degrees of diversity consideration α.

4.1.1 ***Experimental Setting.*** The dataset details are as follows: Sakata et al. created a dataset to assess a "frequently asked question" search task, which aims to yield relevant question-answer pairs from a database in response to a user's query. The dataset comprises 1,786 historical question-answer pairs from Amagasaki City's administrative page in Hyogo Prefecture and 784 inquiries (Query) to administrative municipalities collected through crowdsourcing.

This dataset comprises question-answer pairs, not cause-effect expression pairs from the text. Thus, applying our causal chain presentation algorithm does not yield a causal chain. Nevertheless, the annotation criteria for causal chain accuracy align with Sakata et al.'s dataset labeling standards. Hence, this dataset is suitable for our preliminary experiments. In the dataset created by Sakata et al., the degree of association between each query and QA pair is labeled according to the criteria in the table1.

Table 1: Relevance labeling criteria for each QA pair to a Query in the dataset created by Sakata et al. [18].

A : Contains correct information
B : Contains related information
C : Has the same topic but contains no related information
D : Not relevant

We applied our causal chain presentation algorithm to the dataset, evaluating question-answer pairs linked by a single network expansion with each Query. These QA pairs mimic the causal relation database in our proposed method, where the question signifies the cause expression, and the answer represents the effect expression.

Table 2 presents parameters for the preliminary experiment. Parameters θ and ω are employed to refine the candidate question-answer pairs, and the ensuing candidate connections list is reorganized with diversity in mind to display the top K candidate connections.

Table 2: Parameters of the causal chain presentation algorithm in the pre-experiment.

	Value	Description
N	1	Number of network expansion iterations
K	5	Maximum number of connected nodes per node
θ	0.40	Threshold for node-to-node similarity
ω	0.0	Degree of context-sensitivity consideration

We used MeCab [15] as the morphological analysis tool for extracting nouns in Question-Answer pairs and Queries. The dictionary used was mecab-ipadic-NEologd [19], which supports new words. The number of topics in the LDA was set to 100, and the topic probability threshold for classification was set to 0.2. The word2vec model was trained using Thomason Reuters news articles published from 2012 to 2017, and the number of dimensions was set to 200.

We evaluated the trade-off between output accuracy and diversity of the causal chain presentation algorithm by iteratively adjusting the degree of diversity consideration α from 0.00 to 1.00 in 0.01 increments and consistently assessing the outcomes.

4.1.2 ***Evaluation Metrics.*** We use Discounted Cumulative Gain (DCG) [13] to assess the output's accuracy. DCG is a metric used in information retrieval and recommendation systems to evaluate ranking performance, assuming each item has a multi-valued rating or gain.

The DCG metrics that focuses up to the Kth of the model's output ranking is called *DCG@K* and is computed as follows:

$$DCG@K = r_1 + \sum_{i=2}^{K} \frac{r_i}{\log_2 i} \tag{7}$$

where r_i is the gain of the i-th item in the ranking.

In the causal chain presentation system, the nodes connected from each terminal node are determined by greedily adding nodes for which the objective function in Equation 5 is maximal. Thus, the connected nodes for each terminal node are obtained as a ranking up to the K-th node. This study used the average of *DCG@K* for all terminal nodes as the output accuracy evaluation value.

In the preliminary experiments, the gain is given to each connected node (QA pair) is assumed to be 3 for A, 2 for B, 1 for C, and 0 for D, according to Table 1.

On the other hand, the diversity of the output is evaluated by calculating the following values, using the *Diversity* index [20] proposed in the research of the recommendation system.

$$Diversity(R) = \frac{\sum_{i\in R}\sum_{j\in R\setminus i} 1-\cos(\bar{\boldsymbol{w}}^{\boldsymbol{i}}_{\boldsymbol{effect}}, \bar{\boldsymbol{w}}^{\boldsymbol{j}}_{\boldsymbol{effect}})}{|R|(|R|-1)} \tag{8}$$

where R is all the nodes newly connected by a single network expansion.

In the preliminary experiments, instead of $\bar{\boldsymbol{w}}_{\boldsymbol{effect}}$, we use $\bar{\boldsymbol{w}}_{\boldsymbol{answer}}$, the average vector of embedding representations for the nouns in the answer sentences of the QA pairs.

4.1.3 ***Results and Discussion.*** Five queries were randomly selected from the inquiries (Query) in the dataset, and from these five queries, a network expansion was performed once. We evaluated the accuracy and diversity of the results for the $5K$ Question-Answer pairs obtained. This trial was repeated 300 times for each degree of diversity consideration α, and the average of the result scores was recorded.

The relationship between the measured accuracy evaluation values and the diversity evaluation values is shown in Figure3. For comparison, the average of the *Similarity* values calculated between the Query and QA pairs in the set of connected nodes R, calculated as Equation 1, is also shown. Here, both the average of *DCG@K* and the average of *Similarity* are plotted as normalized values with 1.00 as the score when $\alpha = 0.00$.

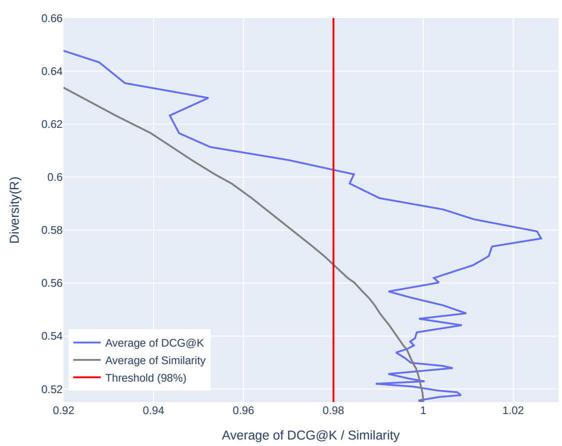

Figure 3: The trade-off relationship between accuracy and diversity evaluation of causal chain presentation algorithm.

Figure 3 is a line graph connecting the plots of the evaluation values when the diversity consideration α is changed from 0.00 to 1.00 in 0.01 increments. The accuracy evaluation is deficient as in the case of *Similarity* (or more rapidly) when the diversity consideration α is larger than a certain value. However, within a certain small value of α, the behavior is almost random. In the case of this experiment, there is no immediate loss of accuracy in exchange for the increase in diversity. On the other hand, *Similarity* shows a monotonous decay in trade-off with the increase in diversity.

To determine the diversity consideration α, the following optimization problem was solved.

$$\begin{aligned} &\text{maximize} \quad \text{Diversity(R)} \\ &\text{subject to:} \quad \text{Total DCG@K} \geq \beta \cdot \text{Total DCG@K}^* \end{aligned}$$

where R is the list of connected QA pairs and $Diversity(R)$ is the diversity evaluation for R. The $TotalDCG@K$ is the sum of the accuracy evaluation values calculated for each query (terminal node), and $TotalDCG@K^*$ is the sum of the accuracy evaluation values calculated for the case of the existing method ($\alpha = 0.00$). β is a parameter that indicates the tolerance for accuracy loss and is generally set to 90% or 95% depending on the purpose and strategy when used in the construction of information retrieval systems or recommendation systems [1]. In this study, we assumed that a decrease in accuracy of up to 2% is acceptable and solved for $\beta = 98\%$. As a result, we obtained the solution that can maximize the diversity when $\alpha = 0.40$ without losing accuracy beyond the threshold value.

4.2 Evaluation Experiment: Measuring the Effectiveness of Considering Output Diversity

In this section, we describe an evaluation experiment of the causal chain presentation system that considers the diversity of outputs, which was conducted using the causal relation database from financial statement summaries, as shown in section 3.2. We evaluated the output of the causal chain presentation system using as input the four cases shown in Table 3. To consider the nature of the financial statement summaries, we added the operation of narrowing down the "target industry" in the first network expansion. The industry classification of the firms followed the medium type of the Japan Exchange Group's industry classification.

Table 3: Four cases tested in the evaluation experiment of the causal chain presentation system.

	Input Keywords	Target Industry
Case1	Infectious disease, Lodging	Services
Case2	Weak yen, Exports	Textiles and Apparels
Case3	Wheat, Price	Foods
Case4	Typhoon, Damage	Warehousing and Transportation

4.2.1 ***Experimental Setting.*** Table 4 presents parameters for the preliminary experiment. We used MeCab and mecab-ipadic-NEologd in the morphological analysis. Configuration parameters in topic classification by LDA and word2vec model were the same as in the preliminary experiment. To evaluate the accuracy, we use the average value for all terminal nodes of *DCG@K* calculated for the nodes connected from each terminal node. In the assessment of diversity, we use the $Diversity(R)$ calculated as Equation 8, where R is all nodes connected by a single network extension.

Table 4: Parameters of the causal chain presentation algorithm in the evaluation experiment to measure the effectiveness of considering output diversity.

	Value	Description
N	3	Number of network expansion iterations
K	5	Maximum number of connected nodes per node
θ	0.40	Threshold for node-to-node similarity
ω	0.1	Degree of context-sensitivity consideration

4.2.2 ***Evaluation Metrics.*** The accuracy and diversity of the system's output are evaluated. The evaluation method is the same as in the preliminary experiment. To evaluate the accuracy, we use the average value for all terminal nodes of DCG@K calculated for the nodes connected from each terminal node. In the assessment of diversity, we use the $Diversity(R)$ calculated as Equation 8, where R is all nodes connected by a single network extension.

A multi-valued rating was manually assigned to each output connection node for accuracy evaluation. Accuracy in a causal sequence is considered to be the degree of association between connected nodes. If we recognize an actual association between connected nodes based on topic classification or word embedding, we can say that the causal sequence is valid. Therefore, in this study, the outputs were labeled according to the criteria shown in table 5 to evaluate their accuracy. The numbers from 0 to 3 in the table indicate the gain in $DCG@K$ calculation.

Table 5: Labeling criteria for each connected node used to evaluate the accuracy of the causal chain presentation system.

3 : Equivalent to the terminal node
2 : Contains related information to the terminal node
1 : Has the same topic but contains no related information to the terminal node
0 : Not relevant to the terminal node

4.2.3 *Results and Discussion*.

The evaluation of the proposed method with $\alpha = 0.40$, which is the optimal value obtained in the preliminary experiment in the previous section, was compared with that of the existing method (equal to the proposed method with $\alpha = 0.00$) [12]. The evaluation was performed for each iteration of network expansion, and the average of the evaluations was used as the evaluation of the entire causal chain presentation system. The evaluation results are shown in Figure 4 and Figure 5, where the total is the average of the four cases.

As shown in Figure 1, the superiority of the accuracy ratings varied from case to case. These results suggest that the proposed method only partially compromises output accuracy at the trade-off of increased diversity.

Regarding diversity evaluation values, the proposed method outperforms the existing methods in all cases. In particular, a significant increase in diversity was observed in Case 1. The input keywords in Case 1 include "infectious diseases," and it is expected that the topic of infectious diseases will appear in many financial reports published by companies in a wide range of industries. Therefore, the number of connection candidate nodes in Case 1 is more significant, and as a result, the impact of diversity-aware ranking on improving the diversity of the output results may have been greater.

The above results show that the proposed method can produce *more diverse output without sacrificing accuracy* compared to existing methods when diversity is considered at $\alpha = 0.40$ in Equation 5.

5 RIPPLE EFFECT ANALYSIS: CASE STUDIES

This section examines the proposed method's effect on considering output diversity by analyzing actual output cases. We compare the outputs for the same event with ($\alpha = 0.40$) and without ($\alpha = 0.00$) diversity in the ripple effect analysis.

The target industry in the first network expansion was "Services," with "infectious disease" as the input keyword. To improve

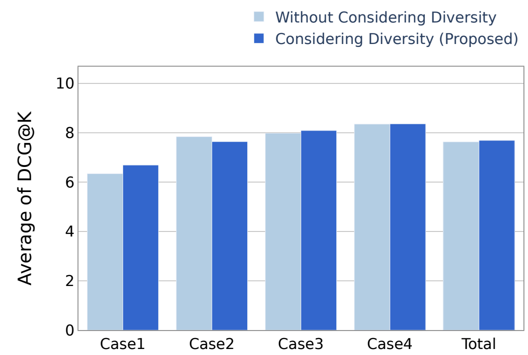

Figure 4: Comparison of evaluated values of accuracy for each case.

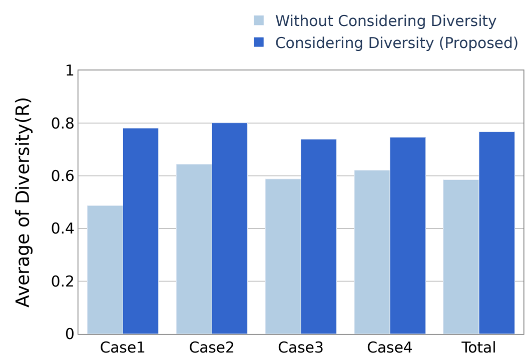

Figure 5: Comparison of evaluated values of diversity for each case.

readability, we examined the output cases with a more limited number of nodes than in the case of the evaluation experiment in section 4.2. We set $N = 2$, $K = 3$, $\theta = 0.40$, and $\omega = 0.3$. The data period was from January 1, 2020, to December 31, 2021. In addition to the output using financial statement summaries, the network was extended by one additional layer using news article texts. By integrating news articles, which are less specialized than financial reports and describe a wide range of topics, we expect to find a broader range of ripple effects. We extracted and used 93,467 causal relationships from Nihon Keizai Shimbun articles (including Nikkan Kogyo Shimbun) published from January 1, 2020, to December 31, 2021. The maximum number of connected nodes per node in the news article text layer $K' = 2$, and the threshold of inter-node similarity $\theta' = 0.60$ were set.

The output examples are shown in Figure 6 and Figure 7. The input keywords are displayed at the top. The causal chain presentation algorithm expands the network toward the bottom of the figure. The dashed arrows indicate the connections between causal relations. The solid arrows in the figure indicate causal relations

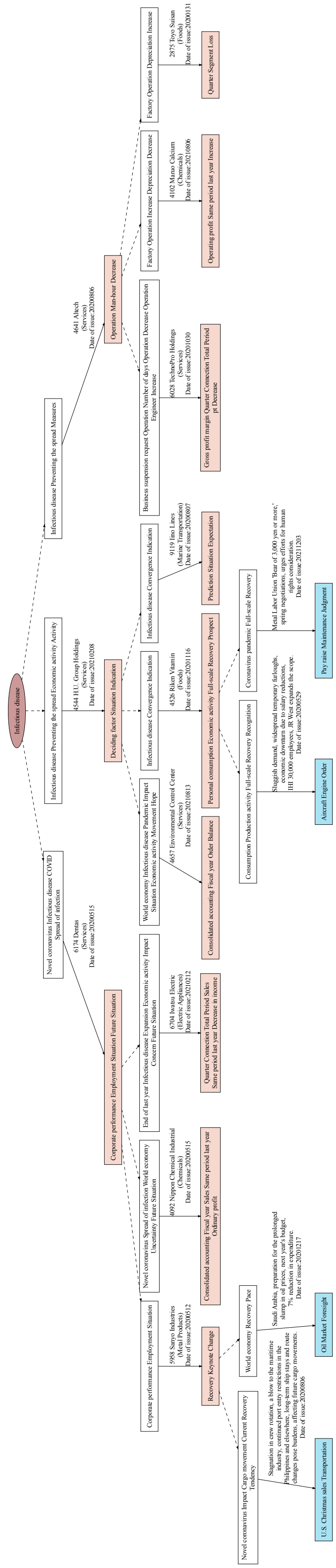

Figure 6: Example output of causal chain presentation system without diversity consideration ($\alpha = 0.00$)

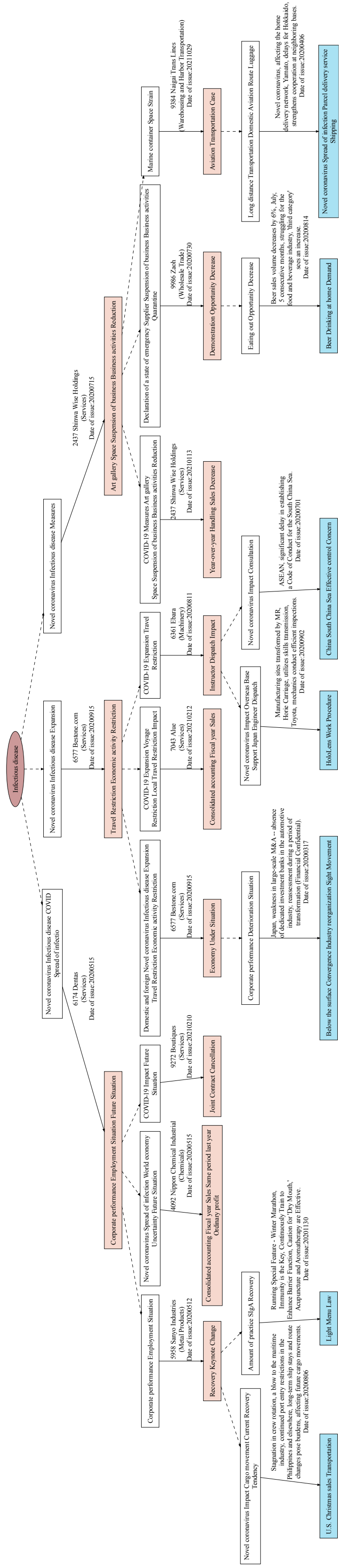

Figure 7: Example output of causal chain presentation system with diversity consideration (diversity consideration $\alpha = 0.40$)

that are connected from cause expressions to effect expressions. The words displayed are nouns in the cause/effect expression.

In the first level, the effect expression for the overall macroeconomic trend, "The outlook for corporate performance and the employment situation remains uncertain in the future situation," was commonly obtained by the existing and proposed methods. Similar topics regarding macroeconomic trends were continuously generated in the existing method shown in Figure 6. On the other hand, in the proposed method shown in Figure 7, such nodes were excluded from the outputs. Instead, nodes containing keywords such as "Travel restriction," "Suspension of business," and "Business activities reduction" were obtained. In this manner, the proposed method can present a more comprehensive range of topics by eliminating similar nodes.

At the third level, the effectiveness of considering diversity was evident. In the existing method, only causal relations were obtained by connecting with the keyword "Recovery" of the economy in Figure 6. On the other hand, the proposed method was able to connect causal expressions in news articles starting from a broader range of topics in Figure 7. From the effect expression in the financial statement summaries, "Dispatching instructors overseas has an impact," a node such as "Conveying work procedures through a HoloLens" was connected, and a novel event was obtained rather than a macroeconomic topic such as the promotion of the use of mixed reality at manufacturing sites. In addition, the topic of "Decreasing opportunities" presented an increase in "Demand for beer at home". Thus, the proposed method presents a broader range of topics by fusing news article texts.

6 CONCLUSIONS

In this study, we have constructed a causal chain presentation system that considers the diversity of output to discover broader economic ripple effects. Evaluation experiment using financial statement summaries has shown that the proposed method could produce more diverse outputs without sacrificing accuracy. The specific effects of considering output diversity have been confirmed through case studies. Future work includes verifying output accuracy more reliably and extending the model to multiple languages.

REFERENCES

[1] Deepak K. Agarwal and Bee-Chung Chen. 2016. *Statistical Methods for Recommender Systems.* Cambridge University Press. https://doi.org/10.1017/CBO9781139565868

[2] Ossama Ahmed, Frederik Träuble, Anirudh Goyal, Alexander Neitz, Yoshua Bengio, Bernhard Schölkopf, Manuel Wüthrich, and Stefan Bauer. 2020. CausalWorld: A Robotic Manipulation Benchmark for Causal Structure and Transfer Learning. arXiv:2010.04296 [cs.RO]

[3] Saud Alashri, Jiun-Yi Tsai, Anvesh Reddy Koppela, and Hasan Davulcu. 2018. Snowball: Extracting Causal Chains from Climate Change Text Corpora. In *2018 1st International Conference on Data Intelligence and Security (ICDIS).* IEEE, 234–241. https://doi.org/10.1109/ICDIS.2018.00045

[4] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *The Journal of Machine Learning Research* 3 (2003), 993–1022.

[5] Jaime Carbonell and Jade Goldstein. 1998. The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'98).* ACM, New York, NY, USA, 335–336. https://doi.org/10.1145/290941.291025

[6] Tirthankar Dasgupta, Rupsa Saha, Lipika Dey, and Abir Naskar. 2018. Automatic Extraction of Causal Relations from Text using Linguistically Informed Deep Neural Networks. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue (SIGDIAL).* ACL, Melbourne, Australia, 306–316. https://doi.org/10.18653/v1/W18-5035

[7] Denis Gordeev, Adis Davletov, Alexey Rey, and Nikolay Arefiev. 2020. LIORI at the FinCausal 2020 Shared task. In *Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation (FNP2019).* COLING, Barcelona, Spain (Online), 45–49.

[8] Sarthak Gupta. 2020. FiNLP at FinCausal 2020 Task 1: Mixture of BERTs for Causal Sentence Identification in Financial Texts. In *Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation (FNP2019).* COLING, Barcelona, Spain (Online), 74–79.

[9] Toshiya Imoto and Tomoki Ito. 2020. JDD @ FinCausal 2020, Task 2: Financial Document Causality Detection. In *Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation (FNP2019).* COLING, Barcelona, Spain (Online), 50–54.

[10] Marius Ionescu, Andrei-Marius Avram, George-Andrei Dima, Dumitru-Clementin Cercel, and Mihai Dascalu. 2020. UPB at FinCausal-2020, Tasks 1 & 2: Causality Analysis in Financial Documents using Pretrained Language Models. In *Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation (FNP2019).* COLING, Barcelona, Spain (Online), 55–59.

[11] Hiroshi Ishii, Qiang Ma, and Masatoshi Yoshikawa. 2012. Incremental Construction of Causal Network from News Articles. *Journal of Information Processing* 20, 1 (2012), 207–215. https://doi.org/10.2197/ipsjjip.20.207

[12] Kiyoshi Izumi and Hiroki Sakaji. 2020. Economic Causal-Chain Search using Text Mining Technology. In *Artificial Intelligence. IJCAI 2019 International Workshops.* Springer International Publishing, Macao, China, 23–35.

[13] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated Gain-Based Evaluation of IR Techniques. *ACM Transactions on Information Systems* 20, 4 (2002), 422–446. https://doi.org/10.1145/582415.582418

[14] Ryotaro Kobayashi, Hiroki Sakaji, and Kiyoshi Izumi. 2023. Determining Sentences Containing Causal Relationships in Financial Text Using BERT and GAT (in Japanese). In *Proceedings of the Twenty-nine Annual Meeting of the Association for Natural Language.* The Association for Natural Language Processing, 2709–2713.

[15] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. 2004. Applying Conditional Random Fields to Japanese Morphological Analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP).* ACL, Barcelona, Spain, 230–237.

[16] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13).* Curran Associates Inc., Red Hook, NY, USA, 3111–3119.

[17] Hiroki Sakaji, Risa Murono, Hiroyuki Sakai, Jason Bennett, and Kiyoshi Izumi. 2017. Discovery of Rare Causal Knowledge from Financial Statement Summaries. In *Proceedings of the 2017 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr).* IEEE, Honolulu, USA, 602–608. https://doi.org/10.1109/SSCI.2017.8285265

[18] Wataru Sakata, Tomohide Shibata, Ribeka Tanaka, and Sadao Kurohashi. 2019. FAQ Retrieval Using Query-Question Similarity and BERT-Based Query-Answer Relevance. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'19).* ACM, Paris, France, 1113–1116. https://doi.org/10.1145/3331184.3331326

[19] Toshinori Sato, Taiichi Hashimoto, and Manabu Okumura. 2017. Implementation of a Word Segmentation Dictionary Called mecab-ipadic-NEologd and Study on How to Use It Effectively for Information Retrieval (in Japanese). In *Proceedings of the Twenty-three Annual Meeting of the Association for Natural Language.* The Association for Natural Language Processing, 875–878.

[20] Saúl Vargas and Pablo Castells. 2014. Improving Sales Diversity by Recommending Users to Items. In *Proceedings of the 8th ACM Conference on Recommender Systems (RecSys '14).* ACM, Silicon Valley, USA, 145–152. https://doi.org/10.1145/2645710.2645744

[21] Sendong Zhao, Quan Wang, Sean Massung, Bing Qin, Ting Liu, Bin Wang, and ChengXiang Zhai. 2017. Constructing and Embedding Abstract Event Causality Networks from Text Snippets. In *Proceedings of the 10th ACM International Conference on Web Search and Data Mining (WSDM).* ACM, New York, NY, USA, 335–344. https://doi.org/10.1145/3018661.3018707