Advances in Intelligent Systems and Computing 1357

Katsutoshi Yada · Daisuke Katagami · Yasufumi Takama · Takayuki Ito · Akinori Abe · Eri Sato-Shimokawara · Junichiro Mori · Naohiro Matsumura · Hisashi Kashima *Editors*

Advances in Artificial Intelligence

Selected Papers from the Annual Conference of Japanese Society of Artificial Intelligence (JSAI 2020)

Springer

Advances in Intelligent Systems and Computing

Volume 1357

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.

More information about this series at http://www.springer.com/series/11156

Katsutoshi Yada · Daisuke Katagami ·
Yasufumi Takama · Takayuki Ito ·
Akinori Abe · Eri Sato-Shimokawara ·
Junichiro Mori · Naohiro Matsumura ·
Hisashi Kashima
Editors

Advances in Artificial Intelligence

Selected Papers from the Annual Conference of Japanese Society of Artificial Intelligence (JSAI 2020)

Springer

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

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

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

Takayuki Ito
Department of Social Informatics
Kyoto University
Kyoto, Japan

Akinori Abe
Division of Behavioral Science,
Faculty of Letters
Chiba University
Chiba, Chiba, Japan

Eri Sato-Shimokawara
Department of Computer Science,
Graduate School of System Design
Tokyo Metropolitan University
Hino, Tokyo, Japan

Junichiro Mori
Mathematics and Informatics Center
The University of Tokyo
Tokyo, Japan

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

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

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-030-73112-0 ISBN 978-3-030-73113-7 (eBook)
https://doi.org/10.1007/978-3-030-73113-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains expanded versions of research papers presented at the international sessions of Annual Conference of the Japanese Society for Artificial Intelligence (JSAI), which was held online on June 2020. 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. In recent years, AI research has proved of great interest to business people. The event draws both more and more presenters and attendees every year, including people of diverse backgrounds such as law and the social sciences, in additional to artificial intelligence. We are extremely pleased to publish this collection of papers as the research results of our international sessions.

The 2020 JSAI Annual Conference drew a record number of contributions, including 85 submissions for the international session, surpassing the number for the previous year. Contributors selected topics from the five categories of machine learning, human interface and education aid, knowledge engineering, agents, and robots and real worlds, with half of these pertaining to machine learning. The papers are selected for presentation at the international sessions over a period of two months, which at least two reviewers screen each paper. In the second phase of the review process, the chosen papers are once again evaluated for originality, significance, and quality. The final twenty-four are included in this book. All are original and high quality, and represent key contributions to AI research.

We would like to extend our deepest appreciation to President Itsuki Noda, Vice President Koji Morikawa (General Chair of JSAI2020), Ex-President Naohiko Uramoto, and Executive Committee Chair Akisato Kimura, 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.

January 2021

Katsutoshi Yada
Chief Editor

Daisuke Katagami
Program Chair, JSAI 2020

Takayuki Ito
Yasufumi Takama
Eri Sato-Shimokawara
Akinori Abe
Junichiro Mori
Naohiro Matsumura
Hisashi Kashima
International Program Committee, JSAI 2020

Organization

The 34th Annual Conference of the Japanese Society for Artificial Intelligence, 2020

一般社団法人 人工知能学会
The Japanese Society for Artificial Intelligence

Conference Chair

Koji Morikawa	Connect Japan

Executive Committee Chairs

Akisato Kimura	NTT, Japan
Takeo Hosomi (vice)	NEC, Japan
Takuma Otsuka (sub)	NTT, Japan

Program Committee Chair

Daisuke Katagami	Tokyo Polytechnic University, Japan
Mitsunori Matsushita (vice)	Kansai University, Japan
Takahito Tomoto (sub)	Tokyo Polytechnic University, Japan

Program Committee

Masahiro Ito	Toshiba, Japan
Hiroto Yonenoh	Nagoya University, Japan
Tomohiro Harada	Tokyo Metropolitan University, Japan
Satoshi Yoshida	NEC, Japan
Ken Tsutsuguchi	Sojo University, Japan

Syuuichi Ojima	Sojo University, Japan
Noriko Horibe	Sojo University, Japan
Motoki Amagasaki	Kumamoto University, Japan
Masato Kiyama	Kumamoto University, Japan
Gou Koutaki	Kumamoto University, Japan
Yuzuru Okajima	NEC, Japan
Koh Mitsuda	NTT, Japan
Kohei Miyaguchi	IBM Research, Tokyo, Japan
Ryo Iwaki	IBM Research, Tokyo, Japan
Junpei Komiyama	New York University, USA
Hidekazu Oiwa	Google, USA
Takashi Nishibayashi	Voyage Group, Japan
Takehiro Terada	Takushoku University, Japan
Kyota Tsutsumida	NTT, Japan
Yoichi Sasaki	NEC, Japan
Jun Ichikawa	Kanagawa University, Japan
Keisuke Fujii	Nagoya University, Japan
Junpei Tsuji	Toyo University, Japan
Tomoyuki Kajiwara	Osaka University, Japan
Dai Hasegawa	Hokkai Gakuen University, Japan
Kazuya Narita	Insight Tech, Japan
Hiromi Narimatsu	NTT, Japan
Naoko Hayashida	Fujitsu Laboratories, Japan
Masakazu Hirokawa	Tsukuba University, Japan
Yuki Ogawa	Ritsumeikan University, Japan
Megumi Kurayama	National Institute of Technology, Hakodate College, Japan
Jun Sugiura	Yahoo Japan, Japan
Masashi Okada	Panasonic, Japan
Shun Tanaka	Future Corporation, Japan
Satoshi Oyama	Hokkaido University, Japan
Taiki Todo	Kyushu University, Japan
Tomoki Aburatani	Osaka Prefecture University, Japan
Kento Koike	Tokyo Polytechnic University, Japan
Yasufumi Takama	Tokyo Metropolitan University, Japan
Eri Shimokawara	Tokyo Metropolitan University, Japan
Takayuki Ito	Nagoya Institute of Technology, Japan
Akinori Abe	Chiba University, Japan
Katsutoshi Yada	Kansai University, Japan
Junichiro Moi	Tokyo University, Japan
Naohiro Matsumura	Osaka University, Japan
Hisashi Kashima	Kyoto University, Japan
Lieu-Hen Chen	National Chi Nan University, Taiwan
Yihsin Ho	Takushoku University, Japan
Yuya Yoshikawa	Chiba Institute of Technology, Japan
Hirotoshi Taira	Osaka Institute of Technology, Japan

Contents

A Node Classification Approach for Dynamically Extracting the Structures of Online Discussions 1
Shota Suzuki, Takayuki Ito, Ahmed Moustafa, and Rafik Hadfi

Visualizing Road Condition Information by Applying the AutoEncoder to Wheelchair Sensing Data for Road Barrier Assessment 13
Goh Sato, Takumi Watanabe, Hiroki Takahashi, Yojiro Yano, Yusuke Iwasawa, and Ikuko Eguchi Yairi

On the Legal Revision in PROLEG Program 25
Wachara Fungwacharakorn, Kanae Tsushima, and Ken Satoh

Viewpoint Planning Based on Uncertainty Maps Created from the Generative Query Network 37
Kelvin Lukman, Hiroki Mori, and Tetsuya Ogata

Active Learning-Based Data Collection in Crowd Replication 49
Lulu Gao and Shin'ichi Konomi

Mining in Discharge Summaries 61
Shusaku Tsumoto, Tomohiro Kimura, and Shoji Hirano

Identifying Snowfall Clouds at Syowa Station, Antarctica via a Convolutional Neural Network 73
Kazue Suzuki, Masaki Shimomura, Kazuyuki Nakamura, Naohiko Hirasawa, Hironori Yabuki, Takashi Yamanouchi, and Terumasa Tokunaga

Context-Aware Dialogue Response Generation Integrating Causal Relation 84
Bo Yang, Jianming Wu, and Gen Hattori

Impact of Domain Knowledge Quality on Inverse Reinforcement Learning 97
Reed Sogabe, Dinesh Bahadur Malla, Masaru Sogabe, Kitsuyoshi Sakamoto, and Tomah Sogabe

Intrinsically Motivated Lifelong Exploration in Reinforcement Learning 109
Nicolas Bougie and Ryutaro Ichise

A Preliminary Analysis of Offensive Language Detection Transferability from Social Media to Video Live Streaming Platforms 121
Zhiwei Gao, Shuntaro Yada, Shoko Wakamiya, and Eiji Aramaki

BERT-Based Dialogue Evaluation Methods with RUBER Framework 133
Khin Thet Htar, Yanan Wang, Jianming Wu, Gen Hattori, and Aye Thida

The Morandi Room: Entering the World of Morandi's Paintings Through Machine Learning 145
Shigeru Kobayashi, Ryota Kuwakubo, Shigeru Matsui, Yoshiyuki Otani, Xinqi Zhang, and Daisuke Niizumi

STBM: Stochastic Trading Behavior Model for Financial Markets 157
Masanori Hirano, Hiroyasu Matsushima, Kiyoshi Izumi, and Hiroki Sakaji

Gated Extra Memory Recurrent Unit for Learning Video Representations 166
Daria Vazhenina and Atsunori Kanemura

Balancing Policy Improvement and Evaluation in Risk-Sensitive Satisficing Algorithm 175
Hiroaki Wakabayashi, Takumi Kamiya, and Tatsuji Takahashi

Designing Lightweight Feature Descriptor Networks with Depthwise Separable Convolution 183
Yeo Ree Wang and Atsunori Kanemura

Hybrid Quantum-Classical Dynamic Programming Algorithm 192
Chih-Chieh Chen, Kodai Shiba, Masaru Sogabe, Katsuyoshi Sakamoto, and Tomah Sogabe

A Shift from Material to Informational Aspects of Shared Resources can Promote Language Evolution 200
Hiroto Yonenoh, Reiji Suzuki, and Takaya Arita

A System for Generating Audio Influenced by Audience Evaluation Using Interactive Genetic Algorithm 208
Maho Taniguchi, Kense Todo, Shoya Yasuda, and Masayuki Yamamura

Strategy and Spatial Representations for Improving Navigation Ability Without Vision Using Sensory Substitution Devices 216
Kouta Akimoto, Ayako Fukawa, and Ikuko Eguchi Yairi

A Politeness Control Method for Conversational Agents Considering Social Relationships with Users 224
Tomoki Miyamoto, Daisuke Katagami, and Mayumi Usami

An Experimental Framework for Exploring the Social Dynamics of a Real-Time Physical Network Composed of Human Participants Using Beacon Technology 232
Zineb Elhamer, Reiji Suzuki, and Takaya Arita

New Approach to Estimating Mental Health Score Using a Communication Agent 239
Satoshi Nakagawa, Shogo Yonekura, Hoshinori Kanazawa, Satoshi Nishikawa, and Yasuo Kuniyoshi

Author Index 249

A Node Classification Approach for Dynamically Extracting the Structures of Online Discussions

Shota Suzuki[1(✉)], Takayuki Ito[2], Ahmed Moustafa[1], and Rafik Hadfi[1]

[1] Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya-shi, Aichi-ken 466-8555, Japan
suzuki.shota@itolab.nitech.ac.jp

[2] Kyoto University, Yoshida Honmachi, Sakyo-ku, Kyoto-shi, Kyoto-fu 606-8501, Japan
http://www.itolab.nitech.ac.jp/itolab_hp_2017/

Abstract. Online discussion platforms require extracting the discussion structure in order to support understanding the flow of these discussions. Towards this end, this paper proposes an approach that performs node classification which is the first step towards extracting the structure of online discussions. In this regard, the proposed approach employs a graph attention network (GAT) in order to directly learn the discussion structure. In specific, the GAT, which is a type of graph neural networks (GNNs), encodes the graph structures directly. In addition, the GAT, which is based on attention architecture, is able to deal with different graph structures. In order to evaluate the proposed approach, we have conducted a set of experiments on the persuasive essay dataset that is styled using the issue-based information system (IBIS). The experimental results show that the proposed approach is able to classify the nodes in online discussion structures accurately.

Keywords: Agent · Deep learning · Facilitation · Natural language processing · Online discussion

1 Introduction

Online discussion support systems have received great attention because they are the next generation approaches for open and public citizen forums [10–12,17,22,28]. In this regard, intelligent consensus-building support systems, namely, COLLAGREE [10,11,22,28] and D-Agree [12] are developed and employed in several real-world experiments. In this context, COLLAGREE and D-Agree had provided support functions for human facilitators who coordinate, lead, integrate, classify, and summarize discussions in order to reach consensus. In their experiments, the consensus-building support systems successfully gathered many more opinions than comparable meetings and made the participants recognize the importance of facilitators.

K. Yada et al. (Eds.): JSAI 2020, AISC 1357, pp. 1–12, 2021.
https://doi.org/10.1007/978-3-030-73113-7_1

In the fields of online discussions and facilitation, our ultimate goal is to create an automated facilitation agent that takes the role of human facilitators. This agent has several benefits such as reacting quickly, facilitating the ongoing discussions in 24 h a day, and making fewer mistakes compared to human facilitators. Towards this end, we have developed the automated facilitation agent that can facilitate online discussions in COLLAGREE and D-Agree.

The automated facilitation agent needs a function that extracts the structures of online discussions in order to understand the situations of these discussions. Understanding the situations of online discussions enables the automated facilitation agent to post facilitations that coordinate and enhance the discussions. Although human facilitators can understand the situations by witnessing the discussions, the automated facilitation agent cannot witness the discussions. Hence, we address the problem of extracting the structures of online discussions.

The task of extracting online discussion structures consists of two subtasks which are node classification and link prediction [27]. The first subtask of node classification means classifying sentences which compose the submissions in online discussions. On the other hand, the second subtask of link prediction means predicting the relationships amongst classified nodes. In this paper, we challenge the first subtask of node classification because node classification with high accuracy results in link prediction with high accuracy.

Towards this end, we propose an approach that directly learns the graph structure for node classification as opposed to the existing feature-based approaches. The proposed approach employs a graph attention network (GAT) [29] that is a kind of graph neural networks (GNNs). In this regard, GAT and GNNs encode the graph structure directly using a neural network model. In particular, the GAT can deal with different graph structures because it employs attention mechanisms [1] unlike other GNNs. In this respect, the proposed approach can address many kinds of graph structures created from natural language texts.

In order to evaluate the proposed approach, we conduct a set of experiments on the persuasive essay dataset [26] formatted into the issue-based information system (IBIS) [15]. The IBIS is an approach for structuring online discussions. The elements of the IBIS are issues that are stated in the form of a controversial question, positions that are responses to an issue, and arguments that are evidence offered to support or oppose a position. The experimental results show achieving high F1 scores when classifying nodes into issues, positions, and arguments. As a result, we conclude that the proposed approach which directly learns the graph structure performs node classification well.

The remainder of this paper is organized as follows. In Sect. 2, we introduce the related works in the fields of the GNNs and the GAT. In Sect. 3, we detail the proposed approach which directly learns the graph structure with a GAT. In Sect. 4, we explain the experimental setup on the dataset of persuasive essays. In Sect. 5, we evaluate the experimental results. In Sect. 6, we discuss the classification performance of the proposed approach. In Sect. 7, we conclude this paper.

2 Related Works

2.1 Graph Neural Networks

GNNs encode graph structures directly using a neural network model. In this regard, GNNs had originated from [24] which showed that neural networks can deal with structured domains such as medical and technical diagnosis, molecular biology, chemistry, speech, text processing, and many others. In particular, the research [24] pointed out that standard neural networks and feature-based approaches are usually believed to be inadequate for dealing with complex structures represented as lists, trees, and graphs of variable sizes and complexities. Also, the research [24] mentioned that the current neural networks do not allow an efficient classification of structures of different sizes. In the attempts to adopt those complex structures, [8] presented the first GNN that is capable of directly processing graphs as opposed to feature-based approaches. In this respect, the research [8] introduced that GNNs can extend recursive neural networks by applying it on directed, undirected, labeled, and cyclic graphs. Particularly, the research indicated that valuable information is lost when traditional machine learning techniques try to cope with graphical data structures by a preprocessing procedure that transforms the graphs into simpler representations such as vectors or sequences of reals. Due to the lack of valuable information, these traditional machine learning techniques may suffer from poor performance and generalization. Besides, many approaches based on the GNNs such as a recurrent GNN [21], a convolutional GNN [3,9], a variational graph auto-encoder [14], and a spatio-temporal graph neural network [23] were proposed in order to promote performance and generalization in many fields of applications. In this regard, these approaches based on the GNNs have proven to learn many kinds of complex graph structures.

2.2 Graph Attention Network

The GAT [29] can deal with different graph structures because of its attention-based architecture. On the other hand, the usual GNNs cannot deal with different graph structures because they compute the eigendecomposition of the graph Laplacian, and their learned filters depend on the Laplacian eigenbasis in encoding graph structures [13]. As an advantage of dealing with different graph structures, the GAT achieved state-of-the-art performance across four established graph-based benchmark datasets of Cora, Citeseer, Pubmed, and protein-protein interaction (PPI). In particular, the results on the PPI dataset showed a significant improvement in the performance. In this regard, the research pointed out that the GAT has large predictive power by observing the entire neighborhood.

Although the existing research [29] adopted the tasks in the fields of citation network and biochemistry, the proposed approach tackles component classification in the field of argument mining. Therefore, it is rich in originality that the proposed approach performs a subtask in the field of argument mining. Besides,

Table 1. A text formatted into the IBIS

Issue1	Are cultural identities issues for immigrants?
Issue2	What is required of immigrants to the country which adopted them?
Position1	Keeping the cultural traditions in the destination countries is tremendous important
Argument1	They need a connection back to their country as well as teach their children their value of origin
Argument2	Fail to create this familiarity makes them felt isolated, in the extreme can lead to social disorder like autism
Issue3	What is required of immigrants to the country which adopted them?
Position2	Immigrants should follow the local customs in order to integrate into their adopted countries' cultures

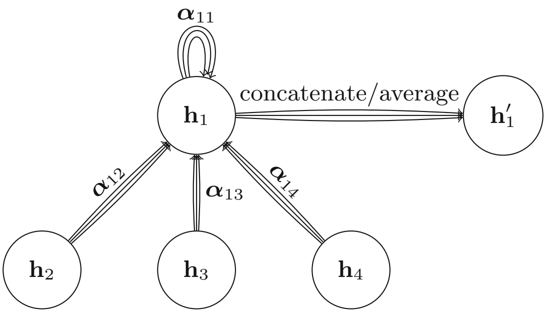


Fig. 1. The single layer of the proposed approach

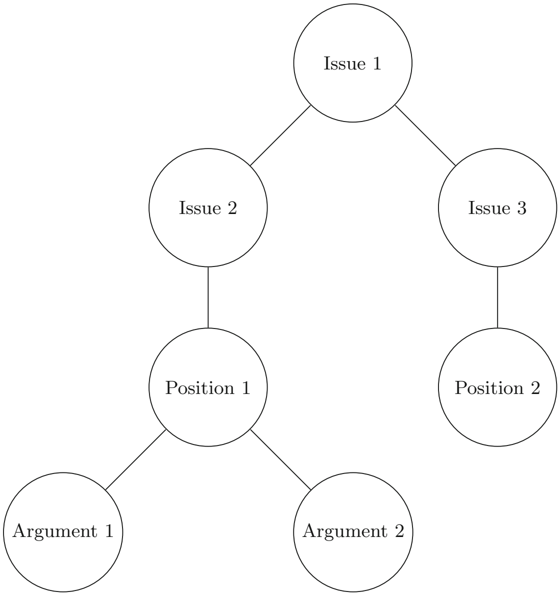
though the existing research [29] applied the GAT for semi-supervised learning, the proposed approach applies the GAT for supervised learning for the component classification. Hence, it has great value that the results show that the proposed approach based on the GAT is also suitable for supervised learning.

3 The Proposed Approach

3.1 The Single GAT Layer

The GAT is a kind of graph neural networks (GNNs) which encode the graph structure directly using a neural network. In particular, the GAT can deal with different graph structures because of its attention mechanism. Following the original study [29], the layer compute a linear combination of the features in order to get the output of final features for every node:

$$\mathbf{h}'_i = \sigma\left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \mathbf{W} \mathbf{h}_j\right)$$



$$G = (V, E)$$

$$\begin{aligned} V = \{&Issue\,1, Issue\,2, Issue\,3, Position\,1, Position\,2, \\ &Argument\,1, Argument\,2\} \\ E = \{&(Issue\,1, Issue\,2), (Issue\,2, Position\,1), \\ &(Position\,1, Argument\,1), (Position\,1, Argument\,2), \\ &(Issue\,1, Issue\,3), (Issue\,3, Position\,2) \end{aligned}$$

Fig. 2. The graph converted from the text

where σ is an activation function, $\mathcal{N}_i$ is first-order neighbors of node i including a self loop, α_{ij} is a element of an attention weight vector, $\mathbf{W}$ is a weight matrix for shared linear transformation, and $\mathbf{h}_j$ is features of node j. α_{ij} is calculated by the softmax function [7]:

$$\alpha_{ij} = \frac{\exp(a(\mathbf{W}\mathbf{h}_i, \mathbf{W}\mathbf{h}_j))}{\sum_{k \in \mathcal{N}_i} \exp(a(\mathbf{W}\mathbf{h}_i, \mathbf{W}\mathbf{h}_k))}$$

where $a : \mathbb{R}^{F'} \times \mathbb{R}^{F'} \rightarrow \mathbb{R}$ is an attention mechanism which calculates the importance of node j's features to node i's features. For the final layer, the proposed approach apply averaging of K independent attention mechanisms:

$$\mathbf{h}'_i = \sigma\left(\frac{1}{K}\sum_{k=1}^{K}\sum_{j \in \mathcal{N}_i} \alpha_{ij}^k \mathbf{W}^k \mathbf{h}_j\right).$$

As a result, the single layer of the proposed approach is illustrated as shown in Fig. 1.

3.2 Node Classification with Graph Attention Network

We propose an approach that directly learns the graph structure with the GAT. First, the proposed approach converts a text that consists of its sentences and their relations into a graph. Therefore, nodes of a graph correspond to sentences of a text, links of a graph correspond to relations of a text. For instance, there is a text formatted as the IBIS as shown in Table 1. In addition, the sentences in the text have relations such as {Issue 1 ↔ Issue 2, Issue 2 ↔ Position 1, Position 1 ↔ Argument 1, Position 1 ↔ Argument 2, Issue 1 ↔ Issue 3, Issue 3 ↔ Position 2}. And then, the proposed approach converts this text into the graph as shown in Fig. 2.

Second, the proposed approach embeds each node into a vector in order to build the features. In this regard, the proposed approach employs the below features as referred to in study [26]:

- Lexical features are binary unigram and 2K most frequent dependency word pairs.
- Indicator features are forward, backward, thesis, and rebuttal Indicators. In addition, the features contain binary values if first-person Indicators are present.
- Embedding features are word embeddings with the bidirectional encoder representations from transformers (BERT) [5].

At last, the proposed approach directly learns the structure of the graph G with supervised learning. In this regard, the proposed approach minimizes the cross-entropy loss in order to tune the weights of the hidden states of the model according to:

$$\mathcal{L} = -\sum_{i \in N} \sum_{c \in C} t_{ic} \ln h'_{ic}$$

where N is the number of sentences in an text, C is the set of classes of nodes in the IBIS ({Issue, Position, Argument}), e.g. $\{MajorClaim, Claim, Argument\}$, t_{ic} is ground truth of the i-th node and h'_{ic} is the output for the i-th node.

4 Experiments

4.1 Dataset

In order to evaluate the proposed approach, we conduct a set of experiments on the dataset of persuasive essays [26] formatted into the IBIS. In this regard, the dataset consists of 90 English texts. In order to evaluate the generalization performance, we randomly split the dataset into training data, validation data, and test data. The number of texts, the number of sentences, and the number of tokens are shown in Table 2. In addition, the classes of the sentences in this dataset are issues, positions, and arguments.

Table 2. The number of texts, number of sentences, and number of tokens in training/validation/test datasets.

	Training	Validation	Test
The number of texts	54	18	18
The number of sentences	1,238	420	413
The number of tokens	15,359	5,216	5,009
Average sentences/text	22.93	23.33	22.94
Average tokens/sentence	12.41	12.42	12.13

Table 3. The parameters for the proposed approach

The number of the GAT layers	2
The number of a layer dimensions	256
The dropout [25] rate of a GAT layer	0.1
The number of hidden attention heads	4
The number of output attention heads	4
The dropout [25] rate of an attention layer	0.1

4.2 Model Settings

In order to perform node classification with directly learning graph structures, we implement the proposed approach in Python. In this regard, we employ the PyTorch [18] which is an open-source machine learning framework that accelerates the path from research prototyping to production deployment. Besides, we adopted the deep graph library [30] which is a Python package built for easy implementation of graph neural network models family, on top of the existing deep learning frameworks such as PyTorch, MXNet, Gluon, and so on. In addition, we adopt the stochastic gradient descent [20] as an optimizer and the stochastic gradient descent with warm restarts (SGDR) [16] as a method for adjusting the learning rate. Moreover, the representative parameters of the model are shown in Table 3, and the representative parameters for training the model are shown in Table 4.

Table 4. The parameters for training the proposed approach

Epochs	2000
Batch size	18
Learning rate	0.001
Momentum for the SGD	0.9
The number of first restart for the SGDR	200

Table 5. The results of experiments

	Issues			Positions			Arguments		
	P	R	F	P	R	F	P	R	F
SVM	1.000	0.630	0.773	1.000	0.196	0.327	0.680	1.000	0.809
RF	0.967	0.967	0.967	0.897	0.380	0.534	0.798	0.983	0.881
GB	0.968	1.000	0.984	0.737	0.609	0.667	0.860	0.908	0.883
MLP	0.979	1.000	**0.989**	0.654	0.761	0.704	0.906	0.838	0.871
The proposed approach	0.936	0.957	0.946	0.910	0.880	**0.895**	0.978	0.983	**0.980**

4.3 Baselines

The proposed approach is compared with several baselines of the existing feature-based approaches. Particularly, we perform the same task with four machine learning approaches such as support vector machine (SVM) [4], random forest (RF) [2], gradient boosting (GB) [6], and multilayer perceptron (MLP).

4.4 Evaluation Methods

We evaluate the proposed approach and the baselines with the values of precision, recall, and F1 score [19]. In this regard, F1 score is the harmonic average of precision and recall:

$$F1\ Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}.$$

5 Results

The results of the experiments are shown in Table 5. The results show that the proposed approach achieves F1 score values of 0.946 in issues classification, 0.895 in positions classification, 0.980 in arguments classification.

To summarize, the results show that the proposed approach which directly learns the graph structure performs node classification well. Particularly, the results of classifying nodes into positions and arguments show the higher F1 score value than the baselines. In this regard, this result is attributed to the fact that the proposed approach learns the graph structures. Particularly, the proposed approach is useful for classifying nodes into positions because the nodes of positions have relations with both of the nodes of issues and arguments. On the other hand, the results of classifying nodes into issues show a little lower F1 score value than the baselines of the RF, GB, MLP. Therefore, we should employ features that are efficient for the proposed approach in order to classify nodes into issues more accurately.

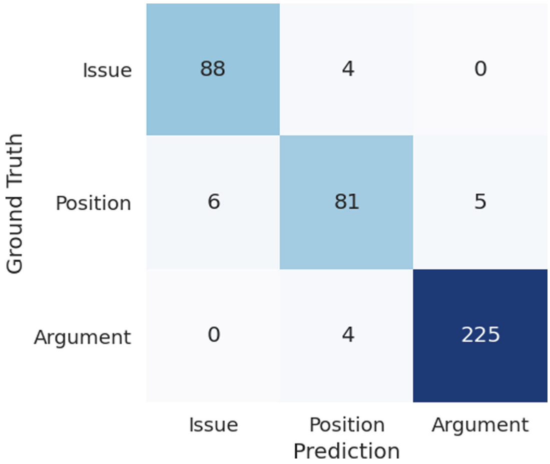


Fig. 3. Confusion matrix.

Table 6. The result of feature ablation.

Features	Issues			Positions			Arguments		
	P	R	F	P	R	F	P	R	F
All	0.936	0.957	0.946	0.910	0.880	0.895	0.978	0.983	0.980
All - Indicator	0.945	0.935	0.940	0.880	0.880	0.880	0.974	0.978	0.976
All - Lexical	0.926	0.946	0.935	0.886	0.848	0.867	0.970	0.978	0.974
All - Embedding	0.659	0.630	0.644	0.500	0.163	0.246	0.742	0.956	0.836

6 Discussion

6.1 Performance for Classifying Each Type of Node

In order to analyze the classification performance of the proposed approach, Fig. 3 shows the confusion matrix of the experimental results in Sect. 5. Figure 3 indicates that the proposed approach classifies issue and argument nodes almost exactly. In contrast, there is room for improvement in classifying claim nodes. In order to improve the performance more, it is required for the proposed approach to employ additional efficient features for node classification.

6.2 Feature Ablation Study

We examine which feature is efficient for node classification with a set of feature ablation experiments. In this respect, we except one of three kinds of features and conduct experiments with the rest features. The results of the ablation experiments are shown in Table 6. Table 6 demonstrates that the embedding feature is the most efficient one of three kinds of features. On the other hand, the indicator feature and the lexical feature improve the classification performance of the proposed approach a little.

Table 7. An example text and the outputs of the proposed approach.

Text	Prediction	Ground truth
What is the role of technology in tradition?	Issue	Issue
How can technology help traditional cultural heritage?	Issue	Issue
Technology enriches the way of displaying traditional cultural heritage, making it more vivid and appealing	Position	Position
China has successfully promoted traditional techniques in the Shanghai World Expo, using a variety of state-of–the-art technological methods, such as robots and LED screens	Argument	Argument
Many tourists around the world marveled at the perfect synthesis of the modern technology and the traditional culture	Argument	Argument
Technology plays a key role in promoting traditional techniques and lifestyles	Position	Position
What is the possible problem that technology bring for tradition?	Issue	Issue
Technology may have a detrimental effect on the traditional lifestyles	Position	Position
People, especially youngsters, are crazier about fresh and advanced things, such as digital products, thus becoming indifferent to traditional techniques	Argument	Argument
People tend to follow the trend back to the tradition in the recent years	Argument	Argument
What can technology do for tradition techniques and lifestyles?	Issue	Issue
Technology supports the preservation and promotion of traditional techniques and lifestyles	Position	Position
Only when traditional culture is integrated with the modern technology can it be developed in the long run	Argument	Argument
What do you think about the opinion that technology will destroy the tradition?	Issue	Issue
Their concern is groundless	Issue	Position

6.3 Case Study

In order to validate the proposed approach, we observe the outputs of the proposed approach. Table 7 shows an example text and the outputs of the proposed approach. We confirmed that the proposed approach classified all issue nodes and all argument nodes accurately. On the other hand, the proposed approach classified a position node of "Their concern is groundless." into an issue node. The additional efficient features would help the proposed approach classify the node accurately.

7 Conclusion and Future Work

This paper proposes a novel approach that classifies discussion nodes via directly learning the graph structures. In this regard, the proposed approach employs a special type of graph neural network, namely, graph attention network. In order to evaluate the proposed approach, a set of experiments has been conducted on the persuasive essay dataset that was formatted into IBIS structure. The experimental results show higher F1 scores when classifying nodes into positions and arguments as compared to other approaches. Future work is set to improve the accuracy of the issue nodes classification. In addition, devising efficient approaches for link prediction is also set for future work.

Acknowledgements. This work was supported by JST CREST Grant Number JPMJCR15E1 including AIP challenge program, Japan.

References

1. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473 (2014)
2. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
3. Bruna, J., Zaremba, W., Szlam, A., Lecun, Y.: Spectral networks and locally connected networks on graphs. In: International Conference on Learning Representations (ICLR 2014), CBLS, April 2014 (2014)
4. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
5. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
6. Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232 (2001)
7. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge (2016)
8. Gori, M., Monfardini, G., Scarselli, F.: A new model for learning in graph domains. In: Proceedings of the 2005 IEEE International Joint Conference on Neural Networks, vol. 2, pp. 729–734. IEEE (2005)
9. Henaff, M., Bruna, J., LeCun, Y.: Deep convolutional networks on graph-structured data. arXiv preprint arXiv:1506.05163 (2015)
10. Ito, T., Imi, Y., Ito, T., Hideshima, E.: COLLAGREE: a faciliator-mediated large-scale consensus support system. Collective Intell. **2014**, 10–12 (2014)
11. Ito, T., Imi, Y., Sato, M., Ito, T., Hideshima, E.: Incentive mechanism for managing large-scale internet-based discussions on COLLAGREE. Collective Intell. **2015** (2015)
12. Ito, T., Suzuki, S., Yamaguchi, N., Nishida, T., Hiraishi, K., Yoshino, K.: D-agree: crowd discussion support system based on automated facilitation agent. In: AAAI, pp. 13614–13615 (2020)
13. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
14. Kipf, T.N., Welling, M.: Variational graph auto-encoders. Stat **1050**, 21 (2016)

15. Kunz, W.: Issues as elements of information systems. Institute of Urban & Regional Development, University of California Working Paper, Berkeley 131 (1970)
16. Loshchilov, I., Hutter, F.: SGDR: stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983 (2016)
17. Malone, T.W.: Superminds: the surprising power of people and computers thinking together. Little, Brown Spark (2018)
18. Paszke, A., et al.: Pytorch: an imperative style, high-performance deep learning library. In: Wallach, H., Larochelle, H., Beygelzimer, A., d' Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems 32, pp. 8024–8035. Curran Associates, Inc. (2019). http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf
19. Powers, D.M.: Evaluation: from precision, recall and f-measure to ROC, informedness, markedness and correlation (2011)
20. Robbins, H., Monro, S.: A stochastic approximation method. Ann. Math. Stat. 400–407 (1951)
21. Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G.: The graph neural network model. IEEE Trans. Neural Netw. **20**(1), 61–80 (2008)
22. Sengoku, A., et al.: Discussion tree for managing large-scale internet-based discussions. Collective Intell. **2016** (2016)
23. Seo, Y., Defferrard, M., Vandergheynst, P., Bresson, X.: Structured sequence modeling with graph convolutional recurrent networks. In: Cheng, L., Leung, A.C.S., Ozawa, S. (eds.) ICONIP 2018. LNCS, vol. 11301, pp. 362–373. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04167-0_33
24. Sperduti, A., Starita, A.: Supervised neural networks for the classification of structures. IEEE Trans. Neural Netw. **8**(3), 714–735 (1997)
25. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**(1), 1929–1958 (2014)
26. Stab, C., Gurevych, I.: Parsing argumentation structures in persuasive essays. Comput. Linguist. **43**(3), 619–659 (2017)
27. Suzuki, S., et al.: Extraction of online discussion structures for automated facilitation agent. In: Ohsawa, Y., et al. (eds.) JSAI 2019. AISC, vol. 1128, pp. 150–161. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39878-1_14
28. Takahashi, K., et al.: Incentive mechanism based on qualit of opinion for large-scale discussion support. Collective Intell. **2016**, 16 (2016)
29. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)
30. Wang, M., et al.: Deep graph library: towards efficient and scalable deep learning on graphs. arXiv preprint arXiv:1909.01315 (2019)

Visualizing Road Condition Information by Applying the AutoEncoder to Wheelchair Sensing Data for Road Barrier Assessment

Goh Sato[1](✉), Takumi Watanabe[1], Hiroki Takahashi[1], Yojiro Yano[1], Yusuke Iwasawa[2], and Ikuko Eguchi Yairi[1]

[1] Graduate School of Science and Engineering, Sophia University, 7-1 Kioicho, Chiyoda-ku, Tokyo 102-8554, Japan
sato.g@yairilab.net

[2] Graduate School of Technology Management for Innovation, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Abstract. Providing accessibility information about sidewalks for people with difficulties with moving is an important social issue. Visualizing road surface conditions to show the accessibilities of the road is effective for this issue. However, conventional methods of collecting huge area accessibility information are based on manpower and have the problem of the large costs of time and money. To solve this problem, we have been proposing and implementing a system for estimating road surface conditions by machine learning with measured values of an acceleration sensor attached to wheelchairs. This paper examined the appropriateness of reconstruction errors which are calculated by Convolutional Variational AutoEncoder to assess the degree of road burden suitable for each user. The evaluation was conducted by calculating reconstruction errors from the traveling data of 14 wheelchair users and creating a map that reflects the information of the calculated errors. The evaluation results suggest that reconstruction errors can reflect the degree of the burden on each wheelchair user during traveling.

Keywords: Variational autoencoder · Convolutional neural network · Sidewalk accessibility · Human activity recognition

1 Introduction

Ensuring the accessibility of the sidewalk is an important social issue for supporting the mobility of the elderly and the mobility impaired, such as wheelchair users. One solution to this issue using information and communication technologies is to develop an accessibility map as an extensive geographic information system (GIS) that provides accessibility information on sidewalks [1, 2]. The personalized accessibility map (PAM) was also proposed as an optimal route navigation system for each user with difficulties in moving [3]. The basic idea of these systems is collecting and digitizing road conditions as accessibility information and visually providing the collected information to users. This paper, hereafter, calls whole process of this idea 'visualization'.

K. Yada et al. (Eds.): JSAI 2020, AISC 1357, pp. 13–24, 2021.
https://doi.org/10.1007/978-3-030-73113-7_2

The main problem to be solved for the visualization is a large-scale data collection of sidewalk accessibility information. Conventional methods of the data collection are as follows: experts evaluate each case from images taken of sidewalks [4], and cooperator evaluate sidewalks by crowdsourcing [5, 6]. These methods are based on manpower and have a problem that collecting large-scale information is difficult because of the time and money costs. Recently, due to the further development of intelligent gadgets, such as smartphones and wristwatch-shaped vital sensors, the ability to sense human activities is improving [7, 8]. To solve the problems above, we have proposed and implemented a system for estimating and classifying road surface conditions by using machine learning from the measured values of acceleration sensors attached to wheelchairs [9–13]. These methods have the difficulty of preparing enough labeled acceleration data and collecting many acceleration data for various road surface irregularities which are difficult for wheelchair users to travel, such as curbs and tactile blocks on the road. This paper uses Convolutional Variational AutoEncoder (CVAE), which has the advantage to detect anomalies even when labeled training data is none or low and anomaly data is low and imbalance. We examine whether the reconstruction error calculated by CVAE can be used as the degree of road burden on wheelchair users and try to provide road information suitable for each user.

2 Related Work

High-resolution satellite images were used for land cover classification according to the physical condition of the ground surfaces such as agricultural land, grazing land, and barren areas [14–16]. Eriksson et al. applied a simple machine learning approach to detect potholes on road surfaces using accelerometer and geographical positioning system (GPS) sensor data of taxies [17]. Allouch et al. applied machine learning to detect depressions in road surfaces using the accelerometer and gyroscope data of vehicles [18]. Abnormal traffic conditions in cities were detected using rich sensor data such as accelerometer, GPS, and voice data from smartphones [19, 20]. These methods can recognize large areas such as agricultural lands or roadways, but it is difficult to recognize conditions on sidewalks.

AutoEncoder (AE) is used for anomaly detection such as detection of anomaly events in video data [21], detection of breast cancer nuclei in medical images [22], and classification of human activities by using data with smartphones and wearable sensors [23, 24]. Generative Adversarial Network (GAN) is used for anomaly detection such as time-series sensor data [25] and real-world cyber-physical systems [26]. Metric Learning is used for anomaly detection from hyperspectral images [27]. Various models are used for anomaly detection, and Variational AutoEncoder (VAE) is used as well. VAE is good at modeling data uncertainty and detecting anomalies by mapping the characteristics of training data to the distribution of latent variables. VAE is used for anomaly detection such as daily activities [28], electrocardiogram [29], and multimodal sensory signals in combination with LSTM [30]. To the best of our knowledge, no research examines that the reconstruction error calculated by CVAE can be used as the degree of road burden for wheelchair users.

3 Overview of the Proposed Method

3.1 Acquisition of Wheelchair Sensing Data

Wheelchair sensing data used in this study is acceleration data of wheelchair users traveling about 1.4 km around Yotsuya Station in Tokyo (Route 1). The wheelchair users who participated in the experiment were six manual wheelchair users and three electric wheelchair users. The acceleration sensor (iPod touch) was attached under the wheelchair seat. Wheelchair users traveled Route 1 three laps. On the second lap, they traveled in the opposite direction from the first and third laps. Acceleration data were sampled in three directions of x, y, and z-axis at 50 Hz, and acceleration samples of about 7.5 h were obtained. To confirm the situations of the location where the acceleration samples were obtained, videos of the condition of wheelchair users and the road were taken during the experiment.

3.2 AutoEncoder and the Reconstruction Error

The AE compresses input data into small feature vectors and outputs reconstructed data which are similar to the input data. AE learns to minimize the errors between the reconstructed data and the input data. Data that resembles a lot of learned data can be accurately reconstructed. Conversely, data that does not resemble a lot of learned data are not accurately reconstructed. CVAE is a network in which the input data is convoluted into a latent variable z by the encoder; z is made into a standard normal distribution and deconvoluted by the decoder.

In this paper, the kernel size (number of feature maps) of the encoder is 5×1 (50), 5×1 (40), 3×1 (20), and the maximum pooling is 1×2, 1×2, 1×1. The kernel size of the decoder is 3×1 (20), 5×1 (40), 5×1 (50), 3×3 (1), and the upsampling is 1×1, 1×2, 1×2. The activation function is ReLU function. In the output layer, the reconstructed acceleration data is output linearly by using a Linear function. Adam is used for training the network, and the mean squared error is used as the loss function.

The CVAE model is created for each user. The acceleration data of eight of nine users are used for the model training, and the reconstruction error is calculated from the acceleration data of the remaining one. As a result, reconstruction errors corresponding to each of the nine wheelchair users are obtained. The reconstruction error is the mean squared error between the input and output data.

3.3 Visualizing Road Information

By using the reconstruction error and the corresponding location information, the information of the reconstruction error is reflected on Google Map (hereinafter referred to as the reconstruction error map). The color of plot points is determined by the ratio of each reconstruction error value to the highest reconstruction error value in three laps. Following the color bar shown in Fig. 1, the higher the ratio, the closer to red, and the lower the ratio, the closer to blue.

Reconstruction error maps we created are as follows: a reconstruction error map for each user (hereinafter referred to as map for each user) and the maximum reconstruction