

Measuring Beginner Friendliness of Japanese Web Pages explaining Academic Concepts by Integrating Neural Image Feature and Text Features

Hayato Shiokawa, Kota Kawaguchi, Bingcai Han, Takehito Utsuro

Graduate School of Systems and Information Engineering,
University of Tsukuba, Tsukuba, 305-8573, JAPAN

Yasuhide Kawada
Logworks Co., Ltd.
Tokyo, 151-0051, JAPAN

Masaharu Yoshioka
Hokkaido University
Sapporo, 060-0808, JAPAN

Noriko Kando
National Institute of Informatics
Tokyo, 101-8430, JAPAN

Abstract

Search engine is an important tool of modern academic study, but the results are lack of measurement of beginner friendliness. In order to improve the efficiency of using search engine for academic study, it is necessary to invent a technique of measuring the beginner friendliness of a Web page explaining academic concepts and to build an automatic measurement system. This paper studies how to integrate heterogeneous features such as a neural image feature generated from the image of the Web page by a variant of CNN (convolutional neural network) as well as text features extracted from the body text of the HTML file of the Web page. Integration is performed through the framework of the SVM classifier learning. Evaluation results show that heterogeneous features perform better than each individual type of features.

1 Introduction

Search engine is a quite important tool for obtaining fundamental as well as practical knowledge when it comes to the study of academic concepts. However, when we intend to find beginner friendly Web pages through search engine, it is necessary to compare many pages by manual work. The reason of ineffective manual comparison is that there is no systematic criterion on measuring beginner friendliness of Web pages in the results of search engine. Therefore, it comes up with us to invent a technique of measuring beginner friendliness of Web pages explaining academic concepts automatically, and finally build a whole assisting system for promoting academic study using search engine, which would improve the effi-

ciency of Web learning.

More specifically, this paper proposes how to automatically measure beginner friendliness of Web pages explaining academic concepts. Before we formalize the framework of automatic measurement of beginner friendliness of Web pages explaining academic concepts, we examine how we manually measure beginner friendliness of those Web pages. The upper half of Figure 1 lists each individual factor that are supposed to be consulted when we judge the overall beginner friendliness of those Web pages. This paper, namely, considers that those individual factors include a) whether or not to contain definition of academic concepts, b) whether or not to contain formulas, c) whether or not to contain figures, d) whether or not to contain examples, e) beginner friendliness of the text of the Web page, and f) visual intelligibility of the Web page layout.

Figure 2(a) shows an example of beginner friendly Web page explaining an academic concept (“probability density function”) of the field of statistics. The Web page of Figure 2(a) can be judged as beginner friendly since it has a visually intelligible layout of the title of the page, the formula, the text of its explanation, and its figure. The text of its explanation is simple but easy to understand, while it has a reference for further studies in the bottom of the page. Figure 2(b), on the other hand, illustrates typical cases of Web pages explaining academic concepts that are not beginner friendly. The case 1 contains a sufficient definition of the academic concept, a figure, a formula, and an example, while its layout is not visually intelligible and its explanation text is not easy to understand. The case 2 is an opposite case, which has a visually intelligible layout as well as the explanation text which is easy to understand, while it lacks a figure nor an example, and having an insufficient definition of the academic concept.

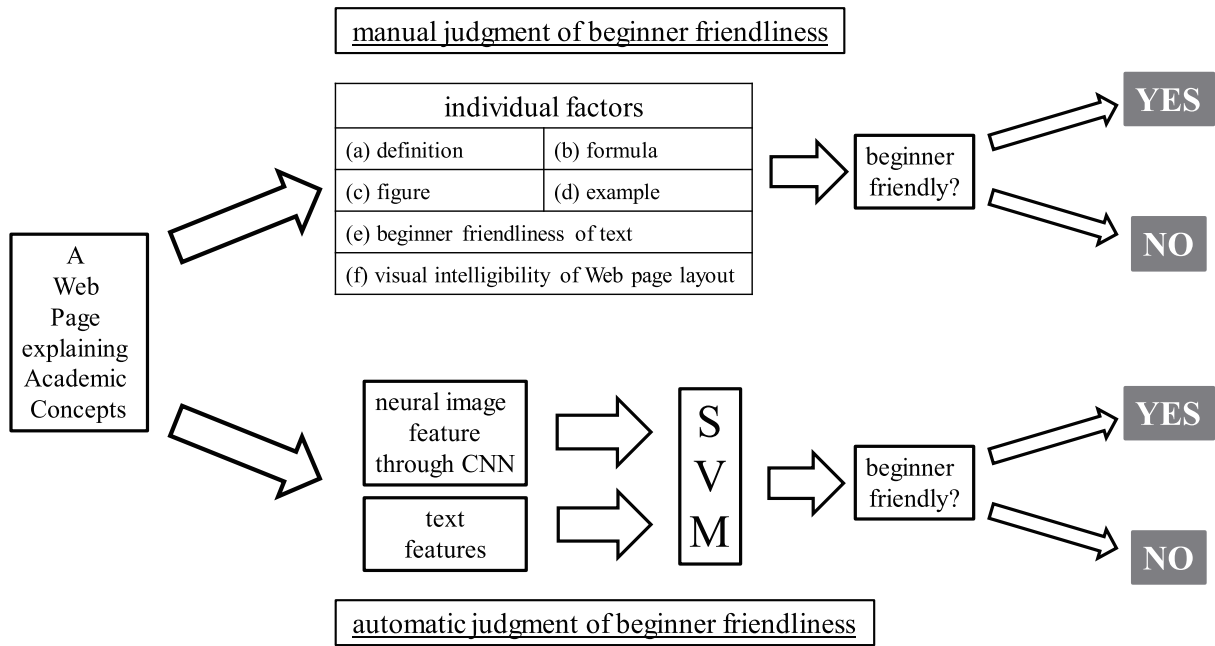


Figure 1: Beginner Friendliness of Web Pages explaining Academic Concepts: Manual Judgment vs. Automatic Judgment by SVM

“an example of probability density function (normal distribution)”


Its description is simple but easy to understand.

Its layout of the title, formula, description, and figure is visually intelligible.

supplementary note for beginners and those who hope to refer to details, introducing “fundamental issues of normal distribution”


Case 1

- having its definition, figure, formula, and example.
- insufficient in visual intelligibility nor in beginner friendliness of text.



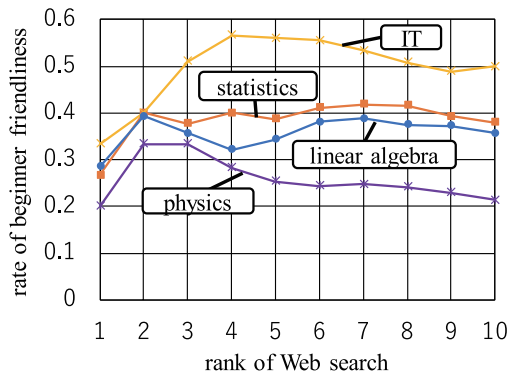
Case 2

- having visually intelligible layout and beginner friendly text.
- no figure / no example, insufficient definition

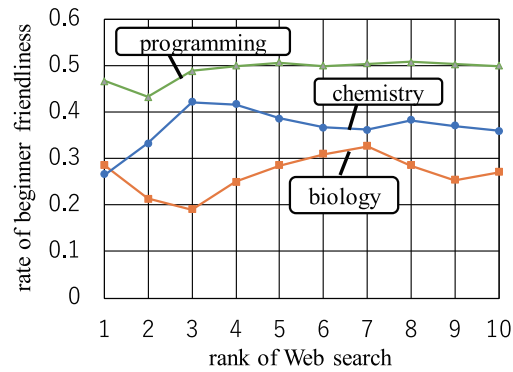


(a) An Example of Beginner Friendly Web Page (excerpt from <https://mathtrain.jp/pmitsudo>) (b) Typical Cases of Beginner Unfriendly Web Pages

Figure 2: Beginner Friendly/Unfriendly Web Pages explaining Academic Concepts



(a) linear algebra, physics, IT, and statistics



(b) biology, programming, and chemistry

Figure 3: Rate of Beginner Friendly Web Pages explaining Academic Concepts ranked at 10th or Higher in the Results of Search Engine

More importantly, when we intend to find beginner friendly Web pages explaining academic concepts through search engine, it is necessary to compare many pages by manual work. The reason of ineffective manual comparison is that there is no systematic criterion on measuring beginner friendliness of Web pages in the results of search engine. Figure 3 shows an evidence of non-existence of such systematic criterion on measuring beginner friendliness of Web pages ranked at 10th or higher by Google search engine in the case of the overall 96 queries of academic terms from the seven academic fields of linear algebra, physics, biology, programming, IT, statistics, and chemistry. The figure plots the rates of the beginner friendly Web pages among those ranked at N -th or higher ($N = 1, \dots, 10$), among which are mostly those explaining academic concepts of the query academic terms. This evidence supports the claim that there is no systematic criterion on measuring beginner friendliness of Web pages explaining academic concepts in the results of Google search engine.

Based on such observation as well as the motivation of finding beginner friendly Web pages explaining academic concepts, this paper studies how to automatically measure beginner friendliness of Web pages explaining academic concepts. As we formalize in the lower half of Figure 1, we integrate heterogeneous features such as a neural image feature generated from the image of the Web page by a variant of CNN (convolutional neural network) as well as text features extracted from the body text of the HTML file of the Web page. Among those individual factors above, the neural

image feature is mostly intended to cover f) visual intelligibility of the Web page layout, while it also partially covers a) definition of academic concepts, b) formulas, c) figures, and d) examples. The text features, on the other hand, are intended to cover d) examples, and e) beginner friendliness of the text of the Web page.

This paper formalizes to integrate those heterogeneous features through the framework of the SVM classifier learning. Evaluation results show that heterogeneous features perform better than each individual type of features.

2 Factors of Beginner Friendliness of Web Pages explaining Academic Concepts

This section describes details of individual factors of beginner friendliness of Web pages explaining academic concepts, as well as their correlations to the overall judgment of beginner friendliness of Web pages.

2.1 Individual Factors

As we describe in the previous section as well as in the upper half of Figure 1, we abstract six individual factors including definition, formula, figure, example, beginner friendliness of text and Web page layout. For each factor, the followings illustrate rough rules on how we manually measure each factor.

- (a) Definition: with this factor, it is examined whether the Web page contains correct and precise definition of the explained academic concept.

Table 1: Query Academic Terms

academic fields	# of queries	academic terms
linear algebra	14	階数 (rank), 共役勾配 (conjugate gradient), 行列式 (determinant), クラメル公式 (Cramer's rule), クロネッカーのデルタ (Kronecker delta), 三角行列 (triangular matrix), 正規直交基底 (orthonormal basis), 対角化 (diagonalization), 直交行列 (orthogonal matrix), 特性多項式 (characteristic polynomial), 二次形式 (quadratic form), ノルム (norm), メネラウスの定理 (Menelaus' theorem), ヤコビ行列 (Jacobian matrix)
physics	15	電気力線 (line of electric force), 張力 (tension), 慣性の法則 (Newton's first law), 遠心力 (centrifugal force), 電波 (radio wave), 電流 (electric current), 万有引力 (universal gravitation), 交流 (alternating current), 音波 (acoustic wave), ホイートストーンブリッジ (Wheatstone bridge), 反発係数 (coefficient of restitution), 相互誘導 (mutual induction), 正電荷 (positive electric charge), 速度 (velocity), 変圧器 (transformer)
biology	7	DNA, ショウジョウバエ (Drosophilidae), 原核生物 (Prokaryote), 減数分裂 (meiosis), 光合成 (photosynthesis), 細胞 (cell), 葉緑体 (Chloroplast)
programming	15	C 言語 (C), Java, エスケープシーケンス (escape sequence), コマンドライン引数 (command line arguments), スコープ (scope), フィールド値 (field value), ポインタ (pointer), メソッド (method), 繰り返し処理 (repetitive processing), 構造体 (structure), 算術演算子 (arithmetic operator), 条件分岐 (branch), 配列変数 (array variable), 文字列 (string), 論理演算 (logical operation)
IT	15	API, DBMS, HTML, IP アドレス (IP address), JDBC, RDB, SDK, SQL, Unicode, URL, スコープマネジメント (scope management), ステークホルダーマネジメント (stakeholder management), タイムマネジメント (time management), ナレッジマネジメント (knowledge management), リスクマネジメント (risk management)
statistics	15	事後分布 (posterior distribution), 確率 (probability), 相関係数 (correlation coefficient), 信頼区間 (confidence interval), k 平均法 (k-means clustering), 回帰分析 (regression), 十分統計量 (sufficient statistic), 確率密度関数 (probability density function), ガンマ分布 (gamma distribution), 事前分布 (prior distribution), 主成分分析 (principal component analysis), コーシー分布 (Cauchy distribution), f 分布 (F distribution), 自己回帰モデル (autoregressive model), 自己回帰 (autoregressive)
chemistry	15	イオン結合 (ionic bond), エステル (ester), カルボン酸 (carboxylic acid), ケトン (ketone), 化学反応式 (chemical equation), 化学平衡 (chemical equilibrium), 共有結合 (covalent bond), 合成高分子 (synthetic polymer), 酸化還元 (redox), 遷移元素 (transition elements), 典型元素 (main-group element), 天然高分子 (natural polymer), 燃料電池 (fuel cell), 物質の三態 (three states of matter), 芳香族有機化合物 (aromatic compounds)
total	96	—

- (b) Formula: with this factor, it is examined whether the Web page contains formula within its text or figures. The formulas should be relevant to the academic concept explained in the Web page.
- (c) Figure: with this factor, it is examined whether the Web page contains figures or pictures relevant to the academic concept explained in the Web page, except when the figure shows formula(s) only.
- (d) Example: with this factor, it is examined whether the Web page contains examples relevant to the academic concept explained in the Web page, including examples of application, proof, explanation and so on. When the examples are shown in figures, it is considered as positive for both of the factors figure and example.
- (e) Beginner friendliness of text: with this factor, it is examined whether the text of the Web page is beginner friendly. More specifically, the amount of information of the text content needs to be within a certain range. The beginner friendliness of the text is violated when too many occurrences of technical terms are observed in the text. It is also required that if too little or too much academic information is included in the text, then that is regarded as violating beginner friendliness of the text. Another criterion is to avoid that the text is too stiff.
- (f) Visual intelligibility of Web page layout: with this factor, it is examined whether the layout of the Web page is visually intelligible. More specifically, the topmost part of the Web page should not be only in text, but should also include figures. Also, the rate of the area of advertisements should be less than a certain upper bound. Furthermore, the background of the Web page should not be in dark color. It is recommended that the top

page has a menu bar as well as a table of contents.

2.2 Overall Measurement considering Individual Factors

When we manually judge the overall beginner friendliness of Web pages explaining academic concepts, there exist certain rules and each individual factor has a certain correlation to the overall judgment. Out of the a) to f) individual factors, the three factors a) definition, e) beginner friendliness of text, as well as f) visual intelligibility of Web page layout, are primary factors compared to the remaining other three factors. All the three two factors should be satisfied in order for the overall beginner friendliness to be satisfied. When all the three factors are satisfied, the overall beginner friendliness tends to be satisfied if at least one of the remaining three factors is satisfied. Out of the remaining other three factors, the more of them are satisfied, the more the overall beginner friendliness is satisfied.

3 Reference Data Set of Web Pages Explaining Academic Concepts

This section describes the details of how we collect the reference data set of Web pages explaining academic concepts as well as the procedure before we judge the overall beginner friendliness of each collected Web page explaining academic concepts according to the criterion discussed in the previous section.

3.1 Academic Fields and Concepts for Study

As for the academic fields for which we collect academic terms to be used as queries, we focus on those within science and technology academic fields, mainly because science and technology academic fields tend to have similar criterion on judging the beginner friendliness of text, the visual intelligibility of the Web page layout, and the overall beginner friendliness of the Web page itself. Out of those science and technology academic fields, we select the following seven for study: linear algebra, physics, biology, programming, IT, statistics, and chemistry. For each field, we select 15 or less academic terms as queries for academic concepts that are around the level of high school or university education, as listed in Table 1. Those query academic terms are selected under the criterion that certain number of Web pages ranked at

10th or higher by Google search engine are those explaining academic concepts.

3.2 Reference Data Set

For each academic term collected in the previous section, we collect the highest 10 Web pages ranked by the Google search engine when each academic term used as the query. In this procedure of collecting Web pages, we ignore Web pages whose HTML files can not be accessed. Then, the first author of this paper¹ judged the overall beginner friendliness as well as the visual intelligibility of the Web page layout of each collected Web page explaining academic concepts according to the criterion discussed in Section 2. Finally, in the procedure of fine-tuning the VGG16 model for judging visual intelligibility of the layout of the Web pages explaining academic concepts, we consider those Web pages which satisfy the visual intelligibility as positive samples while those which do not satisfy the visual intelligibility as negative samples, where their numbers are as shown in Table 2. Similarly, in the procedure of training the SVM classifier for judging the overall beginner friendliness of the Web pages explaining academic concepts, we consider those Web pages which satisfy the overall beginner friendliness as positive samples while those which do not satisfy the overall beginner friendliness as negative samples, where their numbers are also as shown in Table 2. Out of the total seven academic fields, we use the Web pages from five academic fields as training samples, while those from the remaining two as test samples.

4 Neural Image Feature

This section describes the procedure of transforming each Web page explaining academic concepts into its Web page layout image, and then of generating the neural image feature expression from each Web page layout image.

¹ In the preliminary study where two authors of this paper worked on developing reference data set and analyzed their agreement rate, it is discovered that the results of the task of judging the overall beginner friendliness of Web pages explaining academic concepts as well as the visual intelligibility of their Web page layout may vary according to the annotators' knowledge level as well as preferences. Thus, in this paper, in the procedure of developing reference data set, we prefer the consistency of the reference data and we decided to develop reference data set with only one annotator.

Table 2: Numbers of Positive/Negative Samples of Manual Judgment of Visual Intelligibility of Page Layout and Overall Beginner Friendliness of the Web Page

	academic fields	visual intelligibility		overall beginner friendliness		total
		positive	negative	positive	negative	
training	linear algebra	31	58	36	53	89
	physics	34	93	26	101	127
	biology	12	52	17	47	64
	programming	58	84	70	72	142
	IT	37	52	39	50	89
	total	172	339	188	323	511
test	statistics	58	53	49	62	111
	chemistry	50	83	49	84	133
	total	108	136	98	146	244
	total	280	475	286	469	755

4.1 VGG16 (Simonyan and Zisserman, 2015)

It has been well known that deep learning techniques have been applied to a number of tasks in a broad range of research fields and have achieved remarkable improvement over the state of the art baselines. In the domain of pattern recognition such as image recognition, especially, it is noted that convolutional neural networks (CNN) as well as a large scale image data set such as ImageNet (Russakovsky et al., 2014) greatly contribute to achieving high performance in various image recognition tasks. Furthermore, parameters of CNN pre-trained using a large scale general purpose data set of images (e.g. natural images) have been proved to be quite useful for extracting universal features that can be easily fine-tuned to image recognition tasks of certain specific domains such as the medical domain (Shin et al., 2016; Tajbakhsh et al., 2016).

Following those successes of the approach of fine-tuning of pre-trained general purpose CNN parameters for image recognition, this paper applies the approach to the task of automatic judgment of visual intelligibility of the layout of the Web pages explaining academic concepts. More specifically, we employ VGG16 model (Simonyan and Zisserman, 2015) as the general purpose CNN for extracting universal features. VGG16 model won second prize in the image classification task and first prize in the single-object localization task in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 (Russakovsky et al., 2014). Its neural net architecture consists of a stack of 13 convolutional layers with 5 intermediate max-pooling layers, followed by three fully-connected layers, among which the third layer performs 1000-way ILSVRC

classification with 1000 channels (one for each class). The final layer is the soft-max layer. The VGG16 model is pre-trained for the task of 1000-way ILSVRC classification with the ImageNet 2014 data set and is publicly available. It is also known that the pre-trained VGG16 model is widely transferable to other image recognition tasks through fine-tuning. In this paper, as one of the available versions of VGG16 model, we employ the one² available as a model within Keras³, an open source neural network library written in Python.

4.2 Feature of Visual Intelligibility of Web Pages explaining Academic Concepts

This section describes how to generate the neural image feature expression from the layout each Web page explaining academic concepts.

First, each Web page is transformed into its Web page layout image, to which the fine-tuned VGG16 model is applied so as to automatically judge the visual intelligibility of the Web page layout image.

Next, in the fine-tuning of the VGG16 model, its three fully-connected layers of 1000-way ILSVRC classification as well as the soft-max layer are replaced with another three fully-connected layers of binary classification (of judging visual intelligibility of the Web page layout image) as well as the soft-max layer. Throughout the fine-tuning, out of the overall 13 convolutional layers with 5 intermediate max-pooling layers, pre-trained parameters of 10 convolutional layers with 4 intermediate max-pooling layers are

² <https://github.com/keras-team/keras/blob/master/keras/applications/vgg16.py>

³ <https://keras.io/ja/>

kept unchanged, while the remaining three convolutional layers, one max-pooling layer, and the subsequent three fully-connected layers are fine-tuned with the reference training data set (i.e., from the five academic fields of linear algebra, physics, biology, programming, and IT) developed in Section 3.2. Those Web pages from the two remaining academic fields of statistics and chemistry are the reference test samples.

The actual feature values utilized in the subsequent classifier learning of judging the overall beginner friendliness of the Web page explaining academic concepts are the score of the softmax function, ranging within the interval of $[0,1]$, which can be regarded as the confidence of judging the visual intelligibility of the Web page layout.

More specifically, for the five training academic fields, each Web page is annotated with the neural image feature according to the following procedure: i.e., we fine-tune the VGG16 model with four training academic fields out of the total five, then, each Web page of the remaining one training academic field is annotated with the visual intelligibility judged by the VGG16 model fine-tuned with the other four training academic fields.

For the two test academic fields, on the other hand, first we fine-tune five VGG16 models each of which is fine-tuned with four out of five training academic fields. Then, for each test Web page explaining academic concepts, out of those five fine-tuned VGG16 models, one model is randomly selected and applied to the test Web page, where the test Web page is annotated with the visual intelligibility judged by the selected fine-tuned VGG16 model.

5 Text Features

Within the scope of this paper, as the text features for judging the beginner friendliness of the text of explaining academic concepts, almost low level features such as frequencies of character types, words/strings, and HTML tags for pagination functions are employed. The number of specific features among those three types of text features employed in this paper is ten in total. With a preliminary evaluation procedure, we examined much larger candidates list of text features including those ten features⁴, and then, we decided to

⁴ Actually, we examined in total 26 features, i.e., eight character type features, 16 word/string features, HTML tag

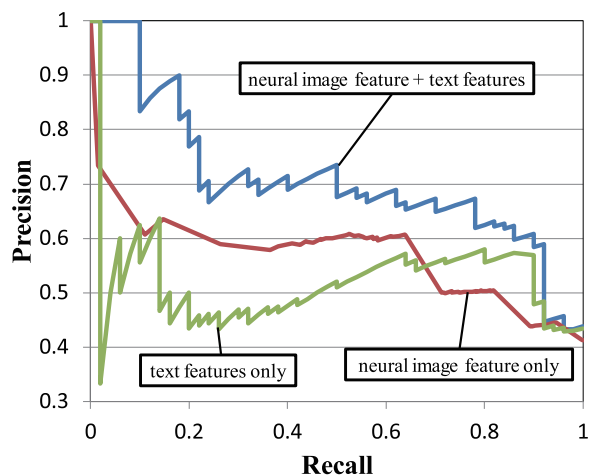


Figure 4: Evaluation Results

employ only those ten features.

5.1 Character Type Features

Japanese sentences are composed mostly of three types of characters, kanji, hiragana, and katakana. Kanji is Chinese characters. Hiragana and katakana are original Japanese characters, where hiragana character is used for Japanese words not covered by kanji and for grammatical inflections, while katakana character is used for transcription of foreign language words into Japanese and the writing of loan words, for emphasis, for onomatopoeia, for technical and scientific terms, and for names of plants, animals, minerals, and often Japanese companies. Following those situations of character types of Japanese sentences, as character type features, we use frequencies of those three character types, kanji, hiragana, and katakana.

5.2 Word/String Features

In this paper, we examined various words/strings as candidates of word/string features, where we finally decided to employ the following six Japanese words/strings and use the frequencies of those words/strings as word/string features.

- “利用規約”(terms of use)
- “相談”(consultation)
- “ノウハウ”(know-how)
- “法”(a constituent character of words such as “方法”(expedient) and “手法”(method))

features for pagination, and HTML tag features for images, out of which ten is selected as an optimal feature combination.

- “困”(a constituent character of a verb “困る”(get into a situation where one needs assistance), where it is intended to count the frequency of a phrase such as “こんな困ったことはありませんか?” (Do you have any experience of having a trouble like this?))
- Total frequencies of a word “例”(example) and symbols “Q0”, . . . , “Q9”, which are intended to count the frequencies of examples and questions.

5.3 Pagination Feature

This feature is introduced to detect paginated Web pages, where a Web page content is divided into a sequence of paginated numbered Web pages. More specifically, any digit sequence immediately after the HTML tag “>” and immediately before the HTML tag “<” is detected and their frequency is counted and used as the pagination feature.

6 Evaluation

6.1 Evaluation Procedure

In this paper, we apply the `sklearn.SVM.SVC` tool of `scikit-learn` (Pedregosa et al., 2011) package to the task of judging the overall beginner friendliness of the Web page explaining academic concepts. Here, for each Web page, the overall beginner friendliness of the Web page explaining academic concepts is used as the class value. We examined the following two approaches to binarizing features which take more than two discrete values or continuous values;

- Dividing the range of discrete values or the continuous values into a certain number of disjoint sub-ranges each of which is exclusive of other sub-ranges.
- Dividing the range of discrete values or the continuous values into a certain number of overlapping sub-ranges which share their lower bounds, i.e., those sub-ranges have exactly the same lower bound.

Through the preliminary evaluation, we employed the approach (b), where the ranges of discrete feature values or the continuous feature values are divided into 20 to 40 overlapping sub-ranges. As the kernel function of the SVM, we used the Radial Basis Function (RBF) kernel. A cost parameter (1 or 10) and a gamma parameter (0.01, 0.001, and 0.0001) of RBF kernel were set by grid search where the area of the ROC curve is optimized.

6.2 Evaluation Results

In the evaluation, we plot recall-precision curves by changing the lower bound of the confidence score of the SVM judgment. Figure 4 compares the performance the following three combinations of features:

- Both the neural image feature and the text features are used.
- Only the neural image feature is used.
- Only the text features are used.

The evaluation results clearly show that integrating the two types of features as in (i) outperform each individual feature(s) (ii) and (iii).

7 Related Work

No existing work studied the task of judging beginner friendliness of Web pages explaining academic concepts. As one of the related tasks, that of estimating presentation skills based on slides and audio features has been studied. For example, Lizard et al. (2014) applied machine learning methods, where the most relevant slide-based features are number of words, images, and tables as well as the maximum font size, while the most significant audio-based features are pitch and filled pauses related ones. Another related task is to evaluate community QA answers (e.g., Wang et al. (2009) and Sakai et al. (2011)). For example, Wang et al. (2009) studied how to rank community answers and evaluated the method using user-labeled “best answers” of Yahoo!Answers Web site as the gold standard positive examples. Compared to the task of ranking community answers, the current task of judging beginner friendliness of Web pages explaining academic concepts is different in that we examine neural image feature, while, in the community answer ranking task, they usually do not consider any image feature when ranking community answers. Also, approaches to text readability judgment (e.g., (Pitler and Nenkova, 2004; González-Garduño and Søggaard, 2017)) are closely related to the task of beginner friendliness of the text of the Web page and the features studied in those previous work need to be studied also in this paper.

8 Conclusion

This paper studied how to integrate heterogeneous features such as a neural image feature generated

from the image of the Web page by a variant of CNN as well as text features extracted from the body text of the HTML file of the Web page. Integration was performed through the framework of the SVM classifier learning. Evaluation results showed that heterogeneous features perform better than each individual type of features. We are now working on developing a reference data set where several annotators participate in the task of developing a reference data set, and then the inter-annotator agreement rate is examined.

Future work include introducing more sophisticated techniques of measuring beginner friendliness of text contents, where it is expected that features that are more semantics-based than frequencies of character types as well as words/strings frequencies contribute to measuring beginner friendliness. Another future work is to incorporate much more detailed list of HTML tags as features of SVM. Preliminary evaluation results indicate that those HTML tag features also contribute to judging beginner friendliness of Web pages explaining academic concepts. This is mainly because one who is capable of developing beginner friendly Web pages explaining academic concepts tends to use certain types of HTML tags frequently and this tendency helps judging beginner friendliness of those Web pages. We plan to report those results in other conferences.

References

- A. V. González-Garduño and A. Søgaard. 2017. Using gaze to predict text readability. In *Proc. 12th BEA*, pages 438–443.
- G. Lazard, B. Guamán, K. Chiluzza, G. Castells, and X. Ochoa. 2014. Estimation of presentations skills based on slides and audio features. In *Proc. MLA*, pages 37–44.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- E. Pitler and A. Nenkova. 2004. Revisiting readability: A unified framework for predicting text quality. In *Proc. EMNLP*, pages 186–195.
- O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. S. Bernstein, A. C. Berg, and F.-F. Li. 2014. ImageNet large scale visual recognition challenge. *CoRR*, abs/1409.0575.
- T. Sakai, D. Ishikawa, N. Kando, Y. Seki, K. Kuriyama, and C.-Y. Lin. 2011. Using graded-relevance metrics for evaluating community QA answer selection. In *Proc. 4th WSDM*, pages 187–196.
- H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers. 2016. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5):1285–1298.
- K. Simonyan and A. Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. In *Proc. 3rd ICLR*.
- N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang. 2016. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35(5):1299–1312.
- X.-J. Wang, X. Tu, D. Feng, and L. Zhang. 2009. Ranking community answers by modeling question-answer relationships via analogical reasoning. In *Proc. 32nd SIGIR*, pages 179–186.