

深層学習を用いた 学術用語解説ウェブページの分かり易さ・見易さの自動評定

塩川 隼人[†] 岡田心太郎^{††} 廣花 智遥[†] 韓 炳材[†] 宇津呂武仁^{†††}

河田 容英^{††††} 神門 典子^{†††††}

[†] 筑波大学 大学院システム情報工学研究科 〒 305-8573 茨城県つくば市天王台 1-1-1

^{††} 筑波大学 理工学群 工学システム学類 〒 305-8573 茨城県つくば市天王台 1-1-1

^{†††} 筑波大学 システム情報系 知能機能工学域 〒 305-8573 茨城県つくば市天王台 1-1-1

^{††††} (株) ログワークス 〒 151-0053 東京都渋谷区代々木 1-3-15 天翔代々木ビル 6F

^{†††††} 国立情報学研究所 〒 101-8430 東京都千代田区一ツ橋 2-1-2

あらまし 本論文では、ウェブ上で学術用語を解説するページ群を対象として、それらのページ群における用語解説の分かり易さおよび見易さを自動評定する手法を提案する。理工系学術分野を対象とし、学術用語を検索クエリとする検索上位ウェブページを収集し、各ページに対し人手により参照用データの作成を行う。本論文では、特に、用語解説ウェブページの見易さを自動評定する CNN、および、用語解説テキストの分かり易さを自動評定する LSTM を導入し、その有効性を示す。

キーワード 学術用語解説ウェブページ, 深層学習, 分かり易さ, 見易さ, 検索エンジン

Measuring Beginner Friendliness / Visual Intelligibility of Web Pages explaining Academic Concepts by Deep Learning

Hayato SHIOKAWA[†], Shintaro OKADA^{††}, Chihiro HIROHANA[†], Bingcai HAN[†], Takehito

UTSURO^{†††}, Yasuhide KAWADA^{††††}, and Noriko KANDO^{†††††}

[†] Grad. Sch. of Systems and Information Engineering, University of Tsukuba, Tsukuba, 305-8573, Japan

^{††} College Eng. Sys., School Sci. and Eng., University of Tsukuba, Tsukuba, 305-8573, Japan

^{†††} Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, 305-8573, Japan

^{††††} Logworks Co., Ltd., Tokyo, 151-0053, Japan

^{†††††} National Institute of Informatics, Tokyo 101-8430, Japan

1. はじめに

インターネットが普及した現在、ウェブ上には多くの学術関連のコンテンツが存在する。このコンテンツを使って学術用語を学ぶことは、学校で使用される教科書・参考書による学習とは別に、学習手段の一つとして度々用いられている。インターネットで学術用語を学ぶ際、ウェブ検索に頼れば関連ページは容易に見つかる。しかし、「分かり易い用語解説」を見つけるには、検索上位ページを一つずつ見比べ読み進める非効率な作業が必要となる。実際に、図 1 に示すように、検索上位ページにおける「分かり易い」学術用語解説ウェブページの割合は多くても半数程度であり、「分かり易い」ページを十分多く収集で

きているとは言えない。そこで、本論文では、ウェブ閲覧者にとって見易く、分かり易く、かつ、役に立つページであるために充足すべき因子を特定し、「解説型ウェブページ」の分かり易さと見易さを自動評定する仕組みを創り上げる (図 2)。そして、ウェブページの見易さ、及び文章の分かり易さを手がかりとして、深層学習による学術用語解説ウェブページの良否判定を自動で行う手法を構築する。特に、理工系学術用語を検索クエリとして収集した用語解説ウェブページを対象として行った評価結果をふまえて、提案手法の有効性を示す。研究の全体像を図 3 に示す。本論文では、学術用語解説ウェブページの良否を判定するために「ウェブページの見易さ」と「文章の分かり易さ」をふまえて「全体評定」の判定を行う。学術用語解説ウェブペー

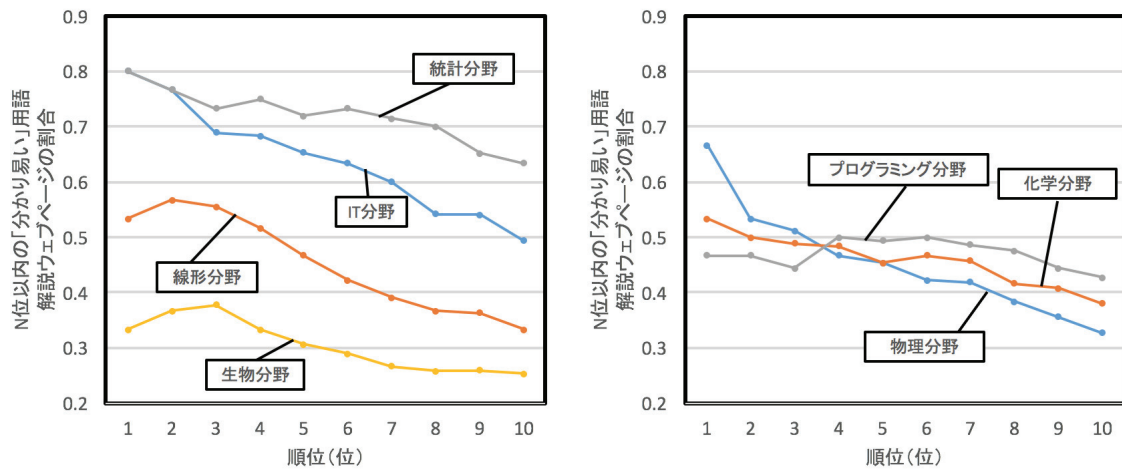


図1 ウェブ検索結果 N 位以内の「分かり易い」用語解説ウェブページの割合

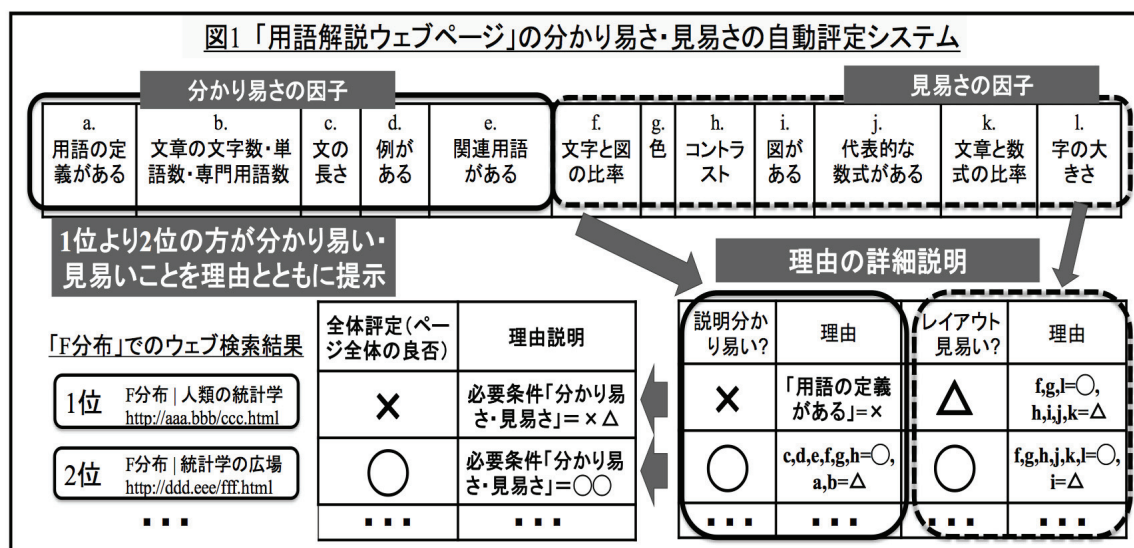


図2 用語解説ウェブページの分かり易さ・見易さの自動評価システム

ジを収集し、2.3.1節、および、2.3.2節で述べる基準に基づいて人手で評価を行う。次に、評価を行なったウェブページ集合を、分野毎に訓練用、開発用、評価用に分割する。訓練用のウェブページ集合を用いて、4.1節、3.1節に示すモデルを訓練した後、評価用ウェブページ集合を用いて評価を行う。

2. 参照用学術用語ウェブページ集合の作成

2.1 対象学術分野および用語

「文章の分かり易さ」、「レイアウトの見易さ」、および、「全体評価」の評価基準の傾向が類似している理工系学術分野を対象とし、特に、物理・統計・IT・生物・線形代数・プログラミング・化学分野を対象分野とした。次に、主に高校3年生または大学レベルの学術用語を対象として選定し、各分野において表1に示す学術用語を評価対象とする。

2.2 データセット作成手順

前節で収集した学術用語に対して用語解説ウェブページを収集し、表2(a)、および、表2(b)で述べる基準に基づきウェブ

ページの見易さ・文章の分かり易さ・全体評価を人手で判定した事例を蓄積する。まず、各用語を検索クエリとして検索エンジンによってウェブページを順位付けした検索の結果上位10件のウェブページを集めた。その際、HTMLファイル収集プログラムではアクセスできないページは参照用ページ集合に加えない。また、用語を収集した全ての学術分野で出現し、かつ学術用語ウェブページとして相応しくないと判断される Wikipedia^(注1)、コトバンク^(注2)、Weblio^(注3)、Yahoo!知恵袋^(注4)の4種類のサイトに含まれるページは全て除外する。そして、2.3.3節の基準に基づき全体評価を人手で行い、その判定結果を付与した事例を蓄積する。以上の手順によって、全体評価が「充足する」と判定された事例を正例、「充足しない」と判定された事例を負例として、各分野ごとに表1に示す数の正例・負例を収集し

(注1) : <https://ja.wikipedia.org/>

(注2) : <https://kotobank.jp/>

(注3) : <https://ejje.weblio.jp/>

(注4) : <https://chiebukuro.yahoo.co.jp/>

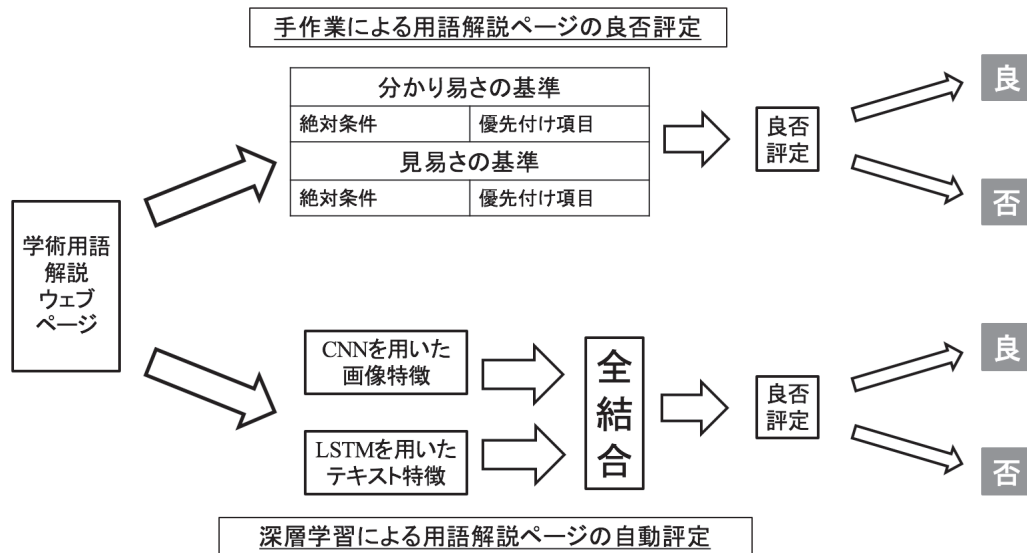


図3 研究の全体像

た。この全7分野のうち、物理、統計、IT、生物の4分野を訓練用、線形代数を開発用、プログラミング、化学の2分野を評価用とする。

2.3 人手評価基準

2.3.1 用語解説ウェブページの分かり易さ

「用語解説ウェブページの分かり易さ」の人手評価においては、表2(a)の評価基準を用いた。基準には絶対条件と優先付け項目を設け、絶対条件を満たし、かつ、複数の優先付け項目を満たすウェブページを、文章が分かり易い学術用語解説ウェブページであると評価した。

2.3.2 用語解説ウェブページの見易さ

「用語解説ウェブページの見易さ」の人手評価においては、表2(b)の評価基準を用いた。基準には絶対条件と優先付け項目を設け、絶対条件を満たし、かつ、複数の優先付け項目を満たすウェブページを、見易い学術用語解説ウェブページであると評価した。

2.3.3 全体評価

全体評価においては、文章の分かり易さ、および、ウェブページの見易さの両方の評価結果に基づいて、ウェブページ全体の良否を評価する。ただし、文章の分かり易さ、もしくは、ウェブページの見易さのどちらかが評価基準に達しなかった場合でも、表2(a),あるいは、表2(b)に示す優先付け項目の一部を満たしていれば、全体評価としては「良」と評価することができる、という形での運用を行った。

3. 文章の分かり易さをふまえた学術用語ウェブページの全体評価

3.1 LSTM

本論文では、HTMLファイルからHTMLタグを除去し抽出したテキストを深層学習モデルに入力し、文章の分かり易さをふまえて全体評価を行う。その際、LSTM (Long Short-Term Memory) を用いたモデルを基盤の特徴抽出器として用いる。

LSTMはテキストのような時系列データを扱うモデルで、同じく時系列データによく用いられるRNN (Recurrent Neural Network) の一種である。RNNの特徴として、以前に入力された情報を記憶することができるという点が挙げられるが、RNNは長期記憶を行う事が難しく、文全体の情報を記憶しなければならないテキストデータに適応し難い、という問題がある。これに対しLSTMは、単純なRNNと比べて長期記憶が可能なモデルとなっており、テキスト分類のタスクで一般的に用いられる。本論文では、特に、双方向LSTMを用い、複数のLSTM層と全結合層から構成されるモデル(図7(a))を利用する。使用するLSTM層の数は、収集したウェブページ集合の中から最も文章の行数が多いウェブページを抜き出し、抜き出したウェブページの行数と同じ数になるよう決定する。

3.2 評価

3.2.1 評価手順

評価においては、2.3.3節で述べた全体評価における良否判定をクラスとし、モデルが自動で分類した結果を評価する。訓練データを用いてモデルの訓練を50エポック行い、50エポック中開発データに対するaccuracyが最大となった訓練モデルを評価用モデルとして用いた。

3.2.2 評価結果

文章の分かり易さをふまえた学術用語解説ウェブページの全体評価結果を図8「テキスト特徴を用いた全体評価」のプロットに示す。ベースラインとしては、検索エンジンを用いて評価用2分野の各用語の検索を行った検索結果の上位10件以内のウェブページのうち、「分かり易い」用語解説ウェブページであると人手で判定したページの割合を用い、これを再現率・適合率曲線にプロットする。この結果から分かるように、テキスト特徴を用いたLSTMにより、ベースラインのプロットを上回る性能を達成している。

表 1 学術分野ごとの対象学術用語一覧

	分野	学術 用語数	対象とする学術用語	正例数および負例数		
				正例	負例	合計
訓練 事例	物理	15	電気力線, 張力, 慣性の法則, 遠心力, 電波, 電流, 万有引力, 交流, 音波, ホイートストンブリッジ, 反発係数, 相互誘導, 正電荷, 速度, 変圧器 (計 15 語)	48	89	137
	統計	15	事後確率, 事前分布, 正規分布, F 分布, 信頼区間, 標準偏差, 主成分分析, 回帰分析, 自己回帰モデル, ポアソン分布, マルコフ連鎖モデル, 確率密度関数, コーシー分布, 分散, パレート分析 (計 15 語)	91	51	142
	IT	15	API, DBMS, HTML, IP アドレス, JDBC, RDB, SDK, SQL, Unicode, URL, スコープマネジメント, ステークホルダーマネジメント, タイムマネジメント, ナレッジマネジメント, リスクマネジメント (計 15 語)	72	72	144
	生物	15	DNA, ショウジョウバエ, 原核生物, 減数分裂, 光合成, 細胞, 葉緑体, ミトコンドリア, 遺伝子, タンパク質, マクロファージ, RNA, ヘモグロビン, ナトリウムポンプ, コケ植物 (計 15 語)	37	98	135
開発 事例	線形代数	15	階数, 共役勾配, 行列式, クラメルの公式, クロネッカーのデルタ, 三角行列, 正規直交基底, 対角化, 直交行列, 特性多項式, 二次形式, ノルム, メネラウスの定理, ヤコビ行列, 内積 (計 15 語)	44	84	128
評価 事例	プログラミング	15	C 言語, Java, エスケープシーケンス, コマンドライン引数, スコープ, フィールド値, ポインタ, メソッド, 繰り返し処理, 構造体, 算術演算子, 条件分岐, 配列変数, 文字列, 論理演算 (計 15 語)	60	82	142
	化学	15	イオン結合, エステル, カルボン酸, ケトン, 化学反応式, 化学平衡, 共有結合, 合成高分子, 酸化還元, 遷移元素, 典型元素, 天然高分子, 燃料電池, 物質の三態, 芳香族有機化合物 (計 15 語)	57	86	143
合計		105	—	409	562	971

4. ウェブページの見易さをふまえた学術用語ウェブページの全体評価

4.1 ResNet

近年, 画像認識の分野では, 畳み込みニューラルネットワーク (CNN) を ImageNet などの大規模なデータセットに適用することにより, 様々なタスクにおいて高い性能を達成している. また, ImageNet のような大規模なデータセットを用いて訓練した CNN のパラメータは, その汎用性の高さから, 高性能な特徴抽出器として異なるドメインのタスクにおいても活用できることが知られている. そこで本論文では, ウェブページを画像化し CNN に入力することにより, ウェブページの見易さをふまえた全体評価を行う. その際, ResNet-50 モデル [2] を基盤の特徴抽出器として用いる. ResNet は, ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2015 [9] の「画像分類タスク (Image Classification)」, 「物体検出タスク (Object detection)」, および, 「位置特定タスク (Single-object localization)」において優勝したモデルである. このモデルは, 49 層の畳み込み層と 1 層の最大プーリング層, 平均プーリング層,

および, 1 層の全結合層 (1,000 値分類用) から成り立っている (図 7(b)). 訓練済みの ResNet-50 としては, ImageNet2015 のデータセットによって 1,000 値分類用に訓練済みのモデルが一般に公開されており, この訓練済みモデルは他のタスクにも広く転用可能なことが知られている. 本論文では, Python の深層学習ライブラリである Pytorch^(注5) 用に公開されているモデル^(注6) を利用して評価実験を行なった. 学術用語解説ウェブページの画像に対する良否評価タスクにおいて ResNet-50 を適用する際には, ResNet-50 から 1,000 値分類用の全結合層を取り除き, 2 値分類用の全結合層を新たに追加したモデルを用いた.

4.2 評価

4.2.1 評価手順

評価においては, 2.3.3 節で述べた全体評価における良否判定をクラスとし, モデルが自動で分類した結果を評価する. 収集した学術用語解説ウェブページのトップ画面を画像化し作成

(注5): <https://pytorch.org/>

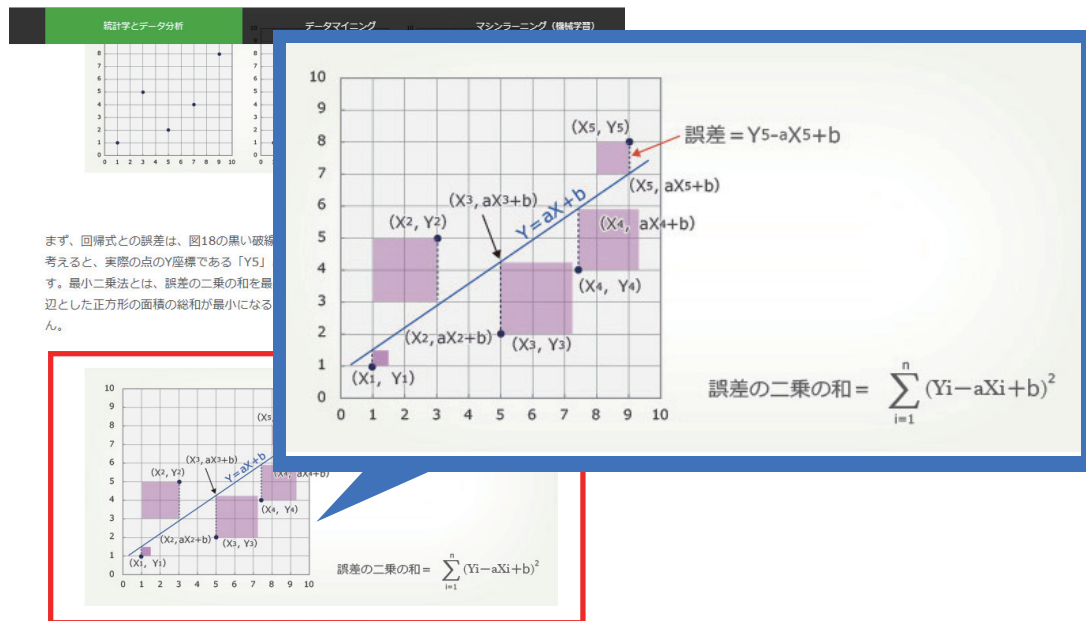
(注6): <https://github.com/pytorch/vision/blob/master/torchvision/models/resnet.py>

表 2 用語解説ウェブページの分かり易さ・見易さの評定基準
(a) 分かり易さの基準

絶対条件	a	用語の定義がある	対象の学術用語に関する定義の説明がウェブページ内に含まれている。
優先付け項目	b	文章の文字数・単語数・専門用語数	ウェブページに含まれる、文章の文字数・単語数・専門用語数が極端に多すぎず、かつ少なすぎない。
	c	文の長さ	ウェブページに含まれる一文の長さが極端に長すぎないもしくは短すぎない。
	d	例がある	図 6 の例に示すように、対象の学術用語に関する例題がウェブページ内に含まれているかどうか
	e	関連用語がある	対象の学術用語に関連する用語がウェブページ内に含まれているかどうか

(b) 見易さの基準

絶対条件	f	文字と図の比率	表や図に対して極端に文字が多くない。もしくは極端に文字が少くない。
	g	色	黒や原色など、文章を読む際に読みづらいと感じるような背景色を使用していない。
	h	コントラスト	白の背景に、黄色の文字等、文章を読む際に読みづらいと感じるような配色をしていない。
優先付け項目	i	図がある	図 4 の例に示すように、用語解説や例示のために使用されている図を一つ以上含んでいる。文字のみの表などは対象としない。
	j	代表的な数式がある	図 5 の例に示すように、一般的に数式と呼べるものを含んでいる。アルファベット一文字などは含めない。
	k	文章と数式の比率	文に対して極端に数式が多くないかどうか。
	i	字の大きさ	文章を読む際に、読みづらいと感じるほど字が小さくない。



回帰係数はどのように求めるか

回帰分析は予測をすることが目的のひとつでした。身長から体重を予測する、母親の身長から子供の身長を予測するなどです。相関関係を「 $Y = aX + b$ 」の一次方程式で表せたとすると、定数の a (傾き)

図 4 用語解説ウェブページにおける図の例 (文献 [11] より引用)

したデータセットを用いて、ResNet-50 の訓練を行う。50 エポック中で開発データに対する accuracy が最大となった訓練モデルを評価用モデルとして用いた。

● 対角行列 A の導出

一般に、対角化された行列は、対角成分がもとの行列の固有値になることが知られている。よって、A の固有値を求めて、対角成分に並べれば、対角化した行列 A が得られる。
A の固有値 λ を求めるには、固有方程式

$$|\lambda I - A| = 0 \quad (1)$$

を λ について解けばよい。左辺は3行3列の行列式であるので、

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda & 1/2 & 3/2 \\ -1 & \lambda - 3/2 & -3/2 \\ 1 & -1/2 & \lambda - 1/2 \end{vmatrix} \\ &= \lambda \left(\lambda - \frac{3}{2} \right) \left(\lambda - \frac{1}{2} \right) + \frac{1}{2} \left(-\frac{3}{2} \right) 1 + \frac{3}{2} (-1) \left(-\frac{1}{2} \right) \\ &\quad - \frac{3}{2} \left(\lambda - \frac{3}{2} \right) 1 - \frac{1}{2} (-1) \left(\lambda - \frac{1}{2} \right) - \lambda \left(-\frac{3}{2} \right) \left(-\frac{1}{2} \right) \\ &= \lambda^3 - 2\lambda^2 - \lambda + 2 \end{aligned}$$

である。よって、(1) は、

$$\lambda^3 - 2\lambda^2 - \lambda + 2 = 0$$

と置かれる。3次方程式であるので、解くのは簡単では

$$(\lambda + 1)(\lambda - 2)(\lambda - 1) = 0$$

となるため、解は

$$\lambda = -1, 1, 2$$

である。固有値が求められたので、対角行列 A は、

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

となる。

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda & 1/2 & 3/2 \\ -1 & \lambda - 3/2 & -3/2 \\ 1 & -1/2 & \lambda - 1/2 \end{vmatrix} \\ &= \lambda \left(\lambda - \frac{3}{2} \right) \left(\lambda - \frac{1}{2} \right) + \frac{1}{2} \left(-\frac{3}{2} \right) 1 + \frac{3}{2} (-1) \left(-\frac{1}{2} \right) \\ &\quad - \frac{3}{2} \left(\lambda - \frac{3}{2} \right) 1 - \frac{1}{2} (-1) \left(\lambda - \frac{1}{2} \right) - \lambda \left(-\frac{3}{2} \right) \left(-\frac{1}{2} \right) \\ &= \lambda^3 - 2\lambda^2 - \lambda + 2 \end{aligned}$$

図 5 用語解説ウェブページにおける式の例 (文献 [11] より引用)

よって、 $P^{-1}AP = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$
(ただし、 D は第 i 成分が y_i の対角行列)

対角化の具体例

実際に 2×2 行列を対角化して固有値、固有ベクトルが計算できる。

例題
 $A = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$ を対角化せよ。

解答
(A の固有値、固有ベクトルは固有値、固有ベクトルの定義と具体的な計算方法の例題で求めた)

固有値 1 に対応する固有ベクトル $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ 、
固有値 4 に対応する固有ベクトル $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ より、
 $P = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}$ 、 $D = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ とおくと、
 $P^{-1}AP = D$ となる (実際、簡単な計算で確認することもできてちょっと感動する)。

例題

$A = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$ を対角化せよ。

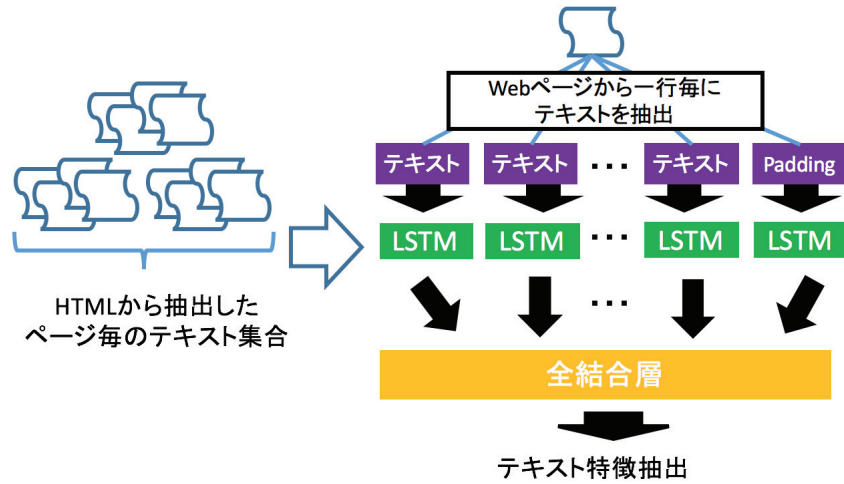
図 6 用語解説ウェブページにおける例示の例 (文献 [11] より引用)

4.2.2 評価結果

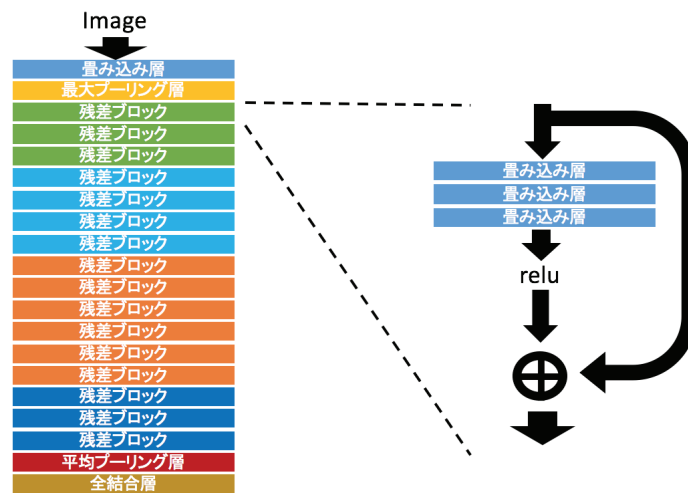
ウェブページの見易さをふまえた学術用語解説ウェブページの全体評価結果を図 8「画像特徴を用いた全体評価」のプロットに示す。ベースラインとしては、検索エンジンを用いて評価用 2 分野の各用語の検索を行った検索結果の上位 10 件以内のウェブページのうち、「分かり易い」用語解説ウェブページであると人手で判定したページの割合を用い、これを再現率・適合率曲線にプロットする。この結果から分かるように、画像特徴を用いた LSTM により、ベースラインのプロットを上回る性能を達成している。

5. 関連研究

本論文の関連タスクとして、文献 [1, 6] においては、HTML 構造上の特徴を利用した全体評価の自動評価手法を提案している。この方式と、テキスト特徴・画像特徴を利用する本論文の手法との併用により、異種特徴量を有効に利用した体系的方式の確立が可能になると考えられる。その他、文献 [3, 4] においては、「分かり易さ」と「見易さ」のうち、「分かり易さ」のみが充足され、「見易さ」が充足されない場合、および、逆に、「見易さ」のみが充足され、「分かり易さ」が充足されない場合に焦点を当て、それぞれ「見易さ」あるいは「分かり易さ」を損なう因子群を網羅的に分析している。さらに、文献 [3, 7] では、深



(a) テキスト特徴量を用いた全体評定用 LSTM モデル



(b) 画像特徴量を用いた全体評定用 ResNet モデル

図 7 用語解説ウェブページの全体評定用モデル

層学習によって用語解説の見易さを自動評定する手法において、見易さ自動評定結果の理由を提示する方式を提案している。また、本論文の学術用語解説ウェブページの分かり易さ自動評定タスクに関連して、コミュニティ型質問応答の分野においては、回答者による回答の良質さを自動評定する手法についての研究が行われている [5, 10]。その他、本論文における学術用語解説ウェブページの見易さの自動評定に関する関連タスクとして、文献 [8] においては、本論文とほぼ同様の深層学習手法によって、プレゼンテーションスライドの画像情報に対する印象を予測する手法を提案している。

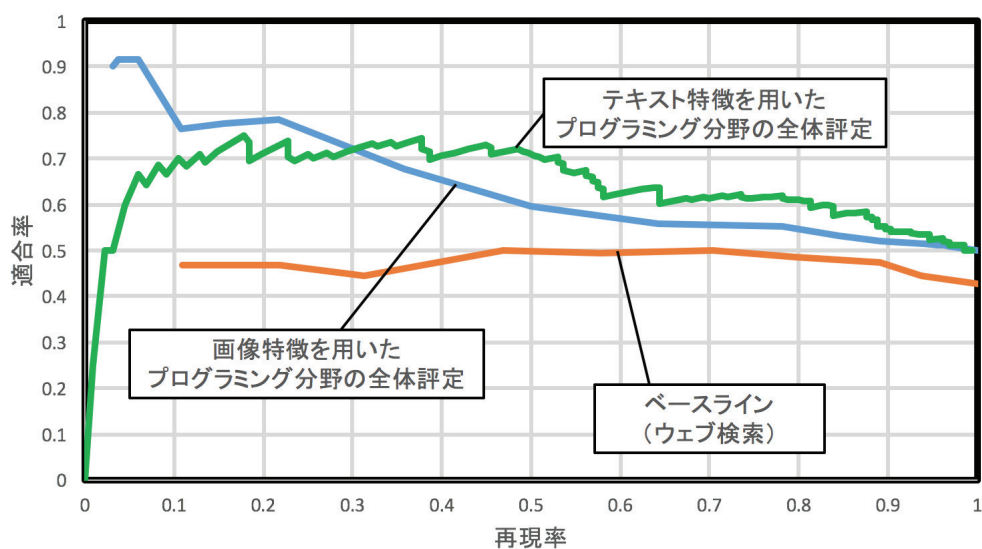
6. おわりに

本論文では、学術用語解説ウェブページに対して、「文章の分かり易さ」および「ウェブページの見易さ」の観点から人手評定を行い、深層学習を用いて、学術用語解説ウェブページの自動評定を行なった。画像特徴のみを利用して深層学習による全体評定を行なった場合、および、テキスト特徴のみを利用して深層学習による全体評定を行なった場合の両方において、ベ-

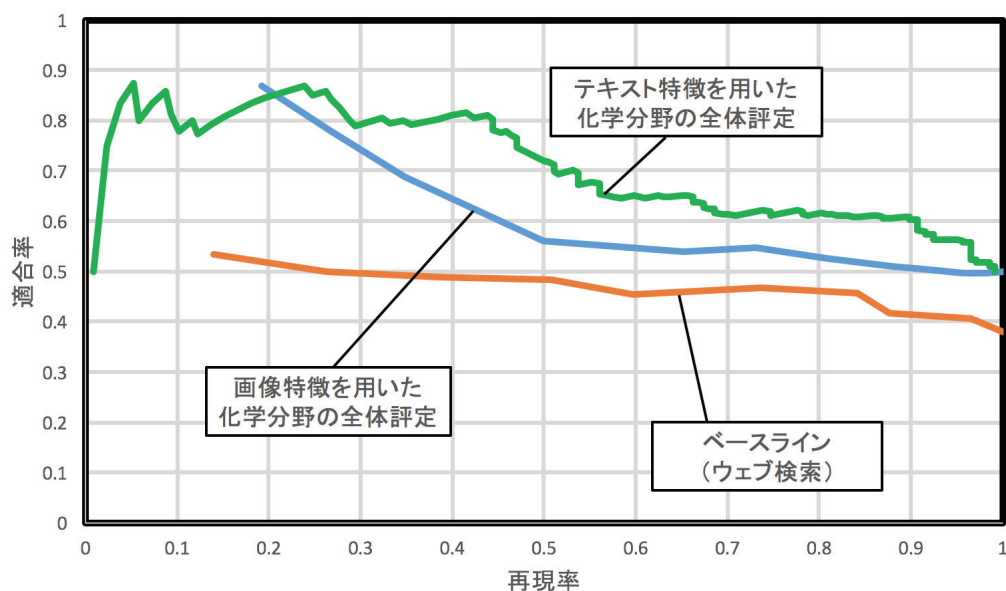
スラインである検索エンジンを利用した場合と比較して、高い精度で学術用語解説ウェブページの自動評定が行えることを示した。

文 献

- [1] B. Han, H. Shiokawa, K. Kawaguchi, T. Utsuro, and Y. Kawada. Measuring beginner friendliness of Chinese Web pages explaining academic concepts using HTML structures. 第 32 回人工知能学会全国大会論文集, 2018.
- [2] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proc. CVPR*, pp. 770–778, 2016.
- [3] 廣花智遥, 岡田心太朗, 塩川隼人, 韓炳材, 宇津呂武仁, 河田容英, 神門典子. 学術用語解説ウェブページの分かり易さ・見易さ因子分析および見易さ自動評定結果の理由提示. 情報処理学会研究報告, Vol. 2019-IFAT-134/2019-DC-112, , 2019.
- [4] 廣花智遥, 岡田心太朗, 宇津呂武仁, 河田容英, 神門典子. 学術用語解説ウェブページの良否評定のための分かり易さ・見易さ因子の分析. 第 33 回人工知能学会全国大会論文集, 2019.
- [5] 石川大介, 酒井哲也, 関洋平, 栗山和子, 神門典子. コミュニティ QA における良質回答の自動予測. 情報知識学会誌, Vol. 21, No. 3, pp. 362–382, 2011.
- [6] 春日孝秀, 塩川隼人, 韓炳材, 宇津呂武仁, 河田容英. HTML 構造上の特徴を利用した学術用語解説ウェブページの分かり易さの自動評定. 第 10 回 DEIM フォーラム論文集, 2018.



(a) プログラミング分野



(b) 化学分野

図 8 学術用語ウェブページの全体評価結果

- [7] 岡田心太郎, 塩川隼人, 韓炳材, 廣花智遥, 宇津呂武仁, 河田容英, 神門典子. 深層学習による学術用語解説ウェブページの見易さ自動評価結果の理由提示. 第 11 回 DEIM フォーラム論文集, 2019.
- [8] 大山真司, 山崎俊彦, 相澤清晴. プレゼンテーションスライドの客観評価と印象予測. 第 16 回 FIT 講演論文集, 第 3 巻, pp. 45–52, 2017.
- [9] 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 L. Fei-Fei. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, Vol. 115, No. 3, pp. 211–252, 2015.
- [10] T. Sakai, D. Ishikawa, N. Kando, Y. Seki, K. Kuriyama, and C.-Y. Lin. Using graded-relevance metrics for evaluating community QA answer selection. In *Proc. 4th WSDM*, pp. 187–196, 2011.
- [11] 塩川隼人, 春日孝秀, 韓炳材, 宇津呂武仁, 河田容英. 深層学習を用いた学術用語解説ウェブページの見易さの自動評価. 第 10 回 DEIM フォーラム論文集, 2018.