

テキスト要約技術と韻律情報を用いた オンデマンド講義における重要箇所推定手法の比較

竹内 美綺¹ 伊藤 彰則¹ 能勢 隆¹

概要：新型コロナウイルスの流行に伴い教育のオンライン化が急速に発展した。オンデマンド型授業では受講生の集中力が続かないことが問題となる。この問題に対し、講義で重要なポイントを受講生に教えるなどの対策が考えられる。本研究では、テキスト要約技術や韻律情報を用いて講義の重要な文を抽出する手法について比較した。テキスト要約技術による文抽出方法と声の大きさや声の高さなどの韻律情報による文抽出方法を組み合わせることでより人手での重要文抽出結果に近づくことがわかった。さらに他の韻律情報を用いることでより人手での重要文抽出結果に近づけられるだろう。

キーワード：テキスト要約、韻律情報、オンデマンド授業

1. はじめに

新型コロナウイルスの感染拡大により、2020年から大学の講義のオンライン化が急速に進んだ。感染の収束に伴い、今後は対面授業が徐々に増えていくと思われる。しかし、オンラインでの講義が成立することが分かったため、これからも教育のオンライン化が進むと考えられる。

オンラインでの講義には様々な形態があり、ポピュラーな形態としては「リアルタイム型」と「オンデマンド型」の2つが挙げられる。「リアルタイム型」とは、Zoom^{*1}やTeams^{*2}などのWeb会議サービスを用いて講義を生配信するものであり、「オンデマンド型」とは事前に収録された講義動画を受講者が視聴するものである。リアルタイム型の講義では教授や他の受講者とリアルタイムでのやりとりがあることも多く、適度な緊張感を持って集中することができる。一方、オンデマンド型の講義では受講者の好きなタイミングで受講できる反面、他の受複数の講者と相互に作用しながら動画を視聴するのは難しく、一人で自宅で聞くだけになり、集中力を保つのが難しい。そのため、オンデマンド講義をより有意義にするには受講者の注意を引くものが必要であると考えられる。

1.1 関連研究

関連研究として、Robalらによる、ビデオ講義において集中力の低下を検知し、警告するシステム“IntelliEye”の開発 [1] がある。受講者の集中力を保つために、“IntelliEye”は受講者の注意力低下を、受講者を写したカメラから顔の存在の有無やマウスの挙動などから検知し、動画の一時停止、警報音の再生、ビデオの周囲に赤いボーダーを繰り返し点滅させることで、リアルタイムで警告する。システムの評価実験では、受講者はそれぞれ3種類の警告の内、ランダムに選ばれた1種類の警告のみ受けた。

実験の結果、動画の一時停止と警報音での警告を受けた受講者は、赤いボーダーの警告を受けた学習者よりも集中力の低下した時間が著しく短かったが、警告数と不注意の持続時間は減少せず、学習効果には結び付かなかった。その原因としては、「集中度」を測る方法が十分に検討されていないこと、講義中は常に集中しなければならないという過度に厳しい前提やシステムと受講者のインタラクションが限定的であることがあげられる。

1.2 研究方針

1.1章で述べた先行研究の問題点を解消するために、以下の4つの条件を満たすシステムの開発があげられる。

- 受講生の集中度を高精度に推定することができる
- 講義の重要な部分を自動的に推定し、受講生にアドバイスする
- システムが受講生に能動的に働きかける
 - 受講生の集中度をモニタリングする

¹ 東北大学工学研究科
Sendai, Miyagi 980-8579, Japan

^{*1} <https://zoom.us/>

^{*2} <https://products.office.com/ja-jp/microsoft-teams/group-chat-software>

- 重要な部分について受講生の集中力が高まるように促す
- その他、システムとの楽しいインタラクションを提供することにより、受講生のモチベーションを高める

上に述べた4つの条件を満たすような「オンライン学習補助システム」を、受講生と同時にオンデマンド授業動画を視聴する対話エージェントを用いて開発することを目指す。また、この対話エージェントを「バーチャル・クラスメイト」と名づける。本研究では特に講義動画の内容の重要箇所を推定する手法の検討を目的とする。

1.3 検討手法

音声から重要文を推定する手法としては、Furuiら[2]が行なったような音声要約や、Horiら[3]や藤井ら[4]のように、講義音声を自動音声認識技術によってテキスト化し、テキスト要約の技術を利用して重要文の推定を行うものもある。しかし、音声には韻律などのパラ言語情報が含まれており、情報の重要度を表す一つの指標となり得ると考えられる。そのため、本研究では、講義音声を自動でテキスト化し要約する技術の他にパラ言語も考慮した重要部分の推定技術を開発する。

まずは抽出すべき重要な部分とはどのようなものかについて調査した。

2. 実験1: 人手による重要文評価

最初に、講義の受講者がどのような部分を重要と思っているのか、また、文面と動画では重要だと感じる部分に違いがあるかについて調査した。

2.1 実験手法

講義内容を理解している情報系の学生5人に、講義動画の書き起こしテキストとスライドなどの講義資料を見て、一文ごとに「授業中に重要だと指摘して欲しい度合い」として5段階で評価してもらった。また、書き起こしテキストから各文を評価した後に、講義動画を見て評価に違いがあればその文を再評価してもらった。

今回の実験では、講演者による趣向の偏りを防ぐために、東北大学2020年度前期講義「応用数学B第1回」、東北大学2020年度前期講義「電気回路学第1回」、および東北大学2021年度前期講義「計算機学第2回」の異なる3人の講演者の講義を使用した。

評価に用いる書き起こしテキストは、音声認識による書き起こしを人手で修正した。この際、誤字脱字を修正し、フィラーを削除した。

2.2 実験結果

人手による評価の結果、動画を見て文面での評価と変わった箇所は書き起こしテキスト全体に対して全体の5%～

7%と多くなかった。

評価が変わった箇所としては、まず講義動画の中でスライドへの書き込みやポインタを用いて説明している部分があげられる。また、講演者の声のトーンやジェスチャーに特徴のある部分についても評価が変わった人が多かった。文面では重要と評価しても、声がよく聞き取れないと重要と感じない、また、文面では低い評価でも、声が大きいと重要だと再評価されている箇所があった。具体的には、一例として、応用数学の講義において、意味のある動きではないが、手を大きく動かしながら、他の発言より大きな声でゆっくりと説明している部分では5人中4人が文面での評価より重要だと判定した。さらに、冗長な喋り方の教授の授業では、話題転換部分を重要であると指摘して欲しいという評価があった。

2.3 考察

スライドへの書き込みやポインタを用いた部分で評価が変わるのは、こういった説明では指示語が多くなり書き起こしテキストからは読み取りきれないからであると考えられる。

講演者の喋り方によって文面と動画で評価が変わったことから、自動で重要部分を抽出する際も文面だけではなくパラ言語情報も考慮する必要があると考えられる。しかし、テキスト情報のみからでも講義の重要箇所は概ね推定できそうであることからまずは重要文抽出手法として、既存の抽出型文書要約モデルを用いた場合の精度について調査した。

3. 抽出型文書要約

要約された文書には元文書の重要な文が含まれていると考えられる。このことから、既存の抽出型文書要約モデルを用いた重要文抽出を試みた。この章では4つの既存のモデルそれぞれのアルゴリズムについて紹介する。

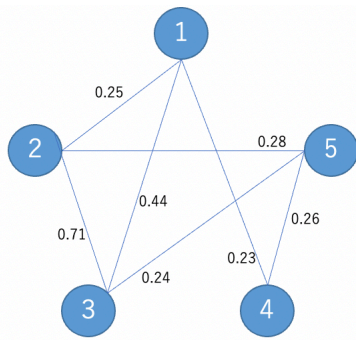
3.1 Luhn[5]

Luhn[5]による単語の出現頻度に基づく抽出方法である。文書の中での頻出単語を重要単語として選出し、これらが各文のどこに位置するかを把握する。文中の重要単語間の距離が5以内であれば、それらを1つのクラスタとする。クラスタの重要単語数の二乗をクラスタの最初と最後の距離で割った値をスコアとし、最終的に各クラスタのスコアの最大値をその文のスコアにする。

3.2 TextRank[6]

Michalceaら[6]によるgraph-based ランキングアルゴリズムを自然言語テキストに適応させる方法である。

多くの文と結びつきがある文は重要な文で、重要な文と結びつきが強い文は重要な文であるという仮定の元、テキ



1. 2016年の参議院選挙では自民党と公明党の連立与党が勝利した。
2. 2016年にトランプ氏が大統領に選ばれた。
3. 2020年の大統領選挙ではバイデン氏がトランプ氏より多くの票を集めた。
4. 野党が内閣不信任案を提出したが与党は否決する方針だ。
5. アメリカには大統領に対する不信任決議権がない。

図 1 TextRank のイメージ図

ストに関連付けしたグラフを作成する。この際グラフの頂点を文とし、グラフの頂点間の接続の強さを「類似性」で重み付けする。ここで「類似性」は内容の重なりとして2文の共通のトークン数で決定し、長文を避けるために2文の重なりを各文章の長さで割ることで正規化している。文を、 N_i 語の集合: $S_i = w_1^i, w_2^i \dots w_{N_i}^i$ として表すとき、2文 S_i と S_j の類似度は以下のように定義される。

$$\text{Similarity}(S_i, S_j) = \frac{|\{w_k | w_k \in S_i \& w_k \in S_j\}|}{\log(|S_i|) + \log(|S_j|)} \quad (1)$$

$G = (E, V)$ を頂点集合 V と辺集合 E を持つグラフとし、与えられた頂点 V_i に対して V_i へ向かう頂点の集合を $In(V_i)$ 、 V_i が指す頂点の集合を $Out(V_i)$ とする。このとき頂点 V_i のスコアを $WS(V_i)$ とする。グラフの頂点のスコアを計算する際にエッジの重みを考慮した graph-based ランキングの式は以下のものである。

$$WS(V_i) = (1-d) + d \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} WS(V_j) \quad (2)$$

d は減衰係数で 0.85 に設定されている。収束するまで graph-based ランキングアルゴリズムを繰り返し、最終的なスコアに基づいて頂点をソートする。

3.3 LexRank[7]

G.Erkan ら [7] による TextRank と同様に文書から頂点を文、辺を類似度とするグラフ構造を作り出してランキングを作成する方法である。

まず、それぞれの文について TF-IDF を計算する。TF, IDF は以下の式で表され、TF-IDF は TF の値と IDF の値を掛け合わせたものである。

$$TF = \frac{\text{文書 A における単語 X の出現頻度}}{\text{文書 A における全単語の出現頻度の和}} \quad (3)$$

$$IDF = \log\left(\frac{\text{全文書数}}{\text{単語 X を含む文書数}}\right) \quad (4)$$

それぞれの文同士の類似度をコサイン類似度を用いて計算し閾値を設定して隣接行列に変換する。

$$\begin{aligned} & \text{idf-modified-cosine}(x, y) \\ &= \frac{\sum_{w \in x, y} tf_{w, x} tf_{w, y} (idf_w)^2}{\sqrt{\sum_{x_i \in x} (tf_{x_i, x} idf_{x_i})^2} \times \sqrt{\sum_{y_i \in y} (tf_{y_i, y} idf_{y_i})^2}} \quad (5) \end{aligned}$$

これを状態遷移行列とし、確率遷移行列に変換する。確率遷移行列から固有値、固有ベクトルを求めて、固有値 1 の場合の固有ベクトルの値をスコアとする。この値に基づいて頂点をソートする。

3.4 BERTSUM[8]

Yang Liu[8] による「BERTSUM」は初めて BERT を事前学習モデルとして使った抽出型文章要約のモデルである。

BERTSUM のデータの入力形式は本来の BERT と比べて少し違いがある。BERT は文章単位で特徴を出力するため文章の先頭のみ [CLS] トークンがつけられているが、抽出型要約では文単位の予測となるため、各文の先頭に [CLS] トークンをつけることで文を表現するベクトルを取得している。さらに、Segment Embedding を入力し、文境界を識別している。Segment Label 入力 2 文目以後は、文章の奇数文と偶数文により違うラベル EA(奇数文) と EB(偶数文) で表示している。

BERT から文のベクトルを得た後、Decoder にはランダムに初期化された 6 層の Transformer を使用したものが最も精度が高かったため、今回の実装では Transformer を用いている。出力データは各文に対して、文章全体に対する重要性として 0 から 1 のスコアにして出力する。例えば文章 [文 1, 文 2, 文 3, 文 4, 文 5] に対して、出力スコアは [0.4, 0.2, 0.8, 0.1, 0.5] となる。スコアが高ければ文章中の重要度が高いということになる。

学習用の目標 (ground truth) は各文が文章において重要であるか否かで 0 と 1 で事前ラベリングする。例えば文章 [文 1, 文 2, 文 3, 文 4, 文 5] に対して、文 3 のみが重要であれば [0, 0, 1, 0, 0] をラベルとして入力し Loss 計算をする。

4. 実験 2: 機械的手法による重要文抽出実験

第 3 章で紹介した機械的手法を用いて実際にオンデマンド講義のテキストから要約を生成し、重要文抽出を行った。また、生成した文を人手評価による抽出文との比較を行なった。

4.1 実験手法

人手評価の実験と同じ講義の書き起こしテキストを使用

表 1 電気回路学

電気回路	ROUGE-1	ROUGE-2	ROUGE-L
Luhn	29.1	9.8	17.0
TextRank	30.2	9.0	18.9
LexRank	23.3	6.9	13.7
BERTSUM	34.1	14.0	22.7

表 2 計算機学

計算機学	ROUGE-1	ROUGE-2	ROUGE-L
Luhn	42.2	19.0	32.9
TextRank	41.6	19.6	37.1
LexRank	42.4	21.1	26.6
BERTSUM	42.5	22.0	32.1

表 3 応用数学

応用数学	ROUGE-1	ROUGE-2	ROUGE-L
Luhn	39.0	18.1	29.3
TextRank	37.0	21.1	32.6
LexRank	49.6	27.0	46.0
BERTSUM	19.6	3.3	10.9

し、それぞれの手法で各講義について 30 文を抽出した。

今回用いた BERTSUM では、BERT の事前学習モデルとして京都大学黒橋河原研究室の日本語のモデルを使用した BERTSUM の学習済み非公式実装^{*3}を用いた。この事前学習モデルでは、入力テキストは日本語 Wikipedia 全て (約 1800 万文)、形態素解析には Juman++ を用いている。また、BERTSUM の学習データには日本語のニュース記事とその要約を約 1 万個ほど用いている。また、BERT の入力長は固定の 512 トークンであるが講義内容は 512 トークンに収まらないので、長文の文章を語句単位で 512 トークン以下に切って別々の文章として要約を行った。

4.2 実験結果

自動要約の評価には主に ROUGE[9] が用いられている。ROUGE は 2 つの文章の類似度をいくつかの基準によって 0 から 100 の値の範囲で評価しており値が大きいほど類似度が高いとされる。今回は要約の正解を 5 人の人手評価の合計スコア上位 30 文として、各手法による要約の品質を ROUGE-1, ROUGE-2 および ROUGE-L を用いて評価した。

電気回路学、計算機学、応用数学それぞれの講義において、各手法による要約結果を ROUGE で評価したものを表 1～3 にまとめた。

Luhn を用いた要約結果では、文中に頻出単語が多くあるほど抽出されるため情報量が多い長めの文が多く抽出された。TextRank を用いた結果では文の長さを考慮しているが、情報量の多い長い文が多く抽出された。また、Luhn と TextRank ではどちらも頻出単語に重きが置かれるため、

^{*3} <https://github.com/neilctwu/YouyakuMan>

抽出された文の約 75 % が同じ文であった。LexRank を用いた結果では、各講義におけるキーワード (計算機学の場合、「入力」、「論理演算」、「and」など) が含まれる文が多く抽出されていた。BERTSUM を用いた結果では、512 トークンごとに一定の文が抽出されているため、全体から満遍なく文が抽出されていた一方で範囲内に重要文がない場合も何らかの文が抽出された。

4.2.1 考察

LexRank では、ROUGE 値が電気回路学では低く、応用数学では高かった。これは、電気回路学では専門的な単語が多く、この回の講義でのキーワードが埋められるという特徴があったのに対し、応用数学ではキーワードが限定的で、その他の説明には簡単な単語を多く用いるという特徴があったことと、人手評価の結果では各講義のキーワードが入っている文が評価が高い傾向にあったことが原因と考えられる。

BERTSUM による抽出文の ROUGE 評価値は、応用数学では他の講義に比べて低くなっていた。これは、学習データは書き言葉であるのに対し、応用数学の書き起こしテキストは本実験で用いた講義の書き起こしテキストの中で最も口語的であったために、モデルが適切な文を抽出できなかったためであると考えられる。

これらを考慮して、機械的手法によってより重要度の高い文を多く抽出するためには BERTSUM で抽出された文に他の手法で抽出された文を追加、削除する方法が考えられる。情報量の少ない文を省くためには、重要語の出現頻度によって抽出文が決まる Luhn の結果において重要度が低いと判定された文を BERTSUM の結果から省く方法が考えられる。書き起こしテキストと動画視聴での評価の差を埋めるために、音声の抑揚や音の高さと文の重要度の相関を調べ、それを考慮したモデルを検討する。

5. 実験 3: 既存の方法を組み合わせた重要文抽出実験

実験 2 の結果を考慮して、機械的手法によってより重要度の高い文を多く抽出するためには BERTSUM で抽出された文に他の手法で抽出された文を追加、削除する方法が考えられる。

5.1 実験手法

まず、情報量の少ない文を省くためには、重要語の出現頻度によって抽出文が決まる Luhn の結果において重要度が低いと判定された文を BERTSUM の結果から省く方法が考えられる。実験 3 では、実験 2 の BERTSUM で抽出された文から、Luhn での評価の下位 10 %、20 %、30 % の文があればそれぞれ削除した。

また、より口語的な講義に対しても重要な文を抽出するために、韻律情報を用いて文を抽出した。小林ら [10] の手

法に基づき、パワーの平均と標準偏差を求め、”平均+標準偏差”以上の文を抽出した。また、声の高さの情報として F0 を用いた。サンプリング周波数は 10kHz とし、F0 の上位 50 個が含まれる文を抽出した。

5.2 実験結果

それぞれの抽出文を実験 2 と同様にして人手評価の上位 40 文を正解としてそれぞれを ROUGE で評価した。その結果を表 4~6 に示す。

この際、表の手法は以下の通りである。

- BERT+pow: BERTSUM による重要文抽出結果にパワーが”平均+標準偏差”以上であった文を追加したもの
- BERT-L 下位○%: BERTSUM による重要文抽出結果から Luhn での結果の下位 10%, 20%, 30% の文があった場合、それらを削除したもの
- BERT+pow-L 下位○%: BERT による要約結果から Luhn での結果の下位 10%, 20%, 30% の文を削除し、パワーが”平均+標準偏差”以上であった文を追加したもの
- BERTSUM+F0: BERT による要約結果に F0 の値上位 50 個が含まれていた文を追加したもの (サンプリング周波数 10kHz)
- BERT+pow+F0: BERT による要約結果にパワーが”平均+標準偏差”以上であった文と、F0 の値上位 50 個が含まれていた文を追加したもの
- BERTSUM+F0-Luhn: BERT による要約結果に F0 の値上位 50 個が含まれていた文を追加し、Luhn での結果の下位 10%, 20%, 30% の文があった場合、それらを削除したもの
- B+pow + F0-Luhn: BERT による要約結果にパワーが”平均+標準偏差”以上であった文と、F0 の値上位 50 個が含まれていた文を追加した文をを追加したものから Luhn での結果の下位 10%, 20%, 30% の文を削除したもの

BERTSUM+power では電気回路学と応用数学で値が良くなった。特に全体的に抑揚が大きいと感じていた応用数学について値がかなりよくなっていた。計算機学では元々 BERTSUM の値がかなり良かったのもあり、パワーでの抽出文を加えると値がかなり下がってしまった。BERTSUM-Luhn では僅かながらもほとんどが値が良くなった。BERTSUM のみと BERTSUM+F0 の結果を比べると、計算機学では BERTSUM のみの方が値がよかったが、電気回路学と応用数学では F0 による抽出文を加えた方がわずかながら値が向上した。また、計算機学では F0 によって抽出された文のほとんどが正解の文ではなかったこと、抽出された文が多かった (24 文) ことから、ROUGE 値も大きく下がった。電気回路学、応用数学は正解の文が

表 4 計算機学

手法	文数	R-1	R-2	R-L
BERTSUM	40	50.4	25.0	32.0
BERT+pow	69	42.5	21.1	29.4
BERT-L 下位 10%	37	52.0	26.2	32.9
BERT-L 下位 20%	34	52.1	25.8	32.6
BERT-L 下位 30%	34	52.1	25.8	32.6
BERT+pow-L 下位 10%	66	44.6	21.8	29.3
BERT+pow-L 下位 20%	63	45.2	21.4	29.7
BERT+pow-L 下位 30%	63	45.2	21.4	29.7
BERTSUM+F0	54	40.5	20.0	28.9
BERT+pow+F0	81	35.8	17.7	27.0
BERT+F0-L 下位 10%	49	42.8	20.8	30.3
BERT+F0-L 下位 20%	45	43.8	20.9	29.7
BERT+F0-L 下位 30%	41	44.4	20.9	30.1
B+pow+F0-L 下位 10%	76	37.5	18.2	28.1
B+pow+F0-L 下位 20%	72	38.2	18.1	28.0
B+pow+F0-L 下位 30%	70	38.4	18.1	28.2

表 5 電気回路学

手法	文数	R-1	R-2	R-L
BERTSUM	40	31.1	11.6	24.0
BERT+pow	94	32.1	16.2	27.7
BERT-L 下位 10%	39	31.4	11.8	24.4
BERT-L 下位 20%	37	31.8	11.9	24.7
BERT-L 下位 30%	35	32.3	12.1	25.1
BERT-L 下位 10%+pow	93	32.5	16.4	28.0
BERT-L 下位 20%+pow	91	32.7	16.5	28.3
BERT-L 下位 30%+pow	89	33.1	16.7	28.6
BERTSUM+F0	48	33.3	13.0	26.9
BERT+pow+F0	96	28.4	14.7	24.7
BERT+F0-L 下位 10%	46	33.9	13.3	27.3
BERT+F0-L 下位 20%	44	34.3	13.4	27.6
BERT+F0-L 下位 30%	40	35.2	13.8	28.4
B+pow+F0-L 下位 10%	94	32.2	16.2	27.8
B+pow+F0-L 下位 20%	92	32.4	16.4	28.0
B+pow+F0-L 下位 30%	90	32.8	16.5	28.4

いくつか抽出されたこと、不正解でもスコアの高めの文が抽出されたことから ROUGE 値が若干ではあるものの上があった。

5.3 考察

計算機学と電気回路学では、BERTSUM+power + F0 の結果より BERTSUM+power, BERTSUM+F0 の方が値がよかった。この原因としては、正解の文に対して抽出した文の数が増えすぎたことが考えられ、単に抽出方法を増やし、組み合わせるだけでは値が向上しないことがわかった。一方で応用数学では BERTSUM+power + F0 の結果が BERTSUM+F0 より向上し、BERTSUM+power と同程度であった。書き起こしテキストのみからの文抽出がうまくいっていないものには多くの方法を組み合わせること

表 6 応用数学

手法	文数	R-1	R-2	R-L
BERTSUM	40	20.9	1.0	14.0
BERT+pow	77	26.6	9.4	23.1
BERT-L 下位 10%	39	21.1	1.0	14.0
BERT-L 下位 20%	35	20.2	1.0	14.7
BERT-L 下位 30%	33	18.7	1.0	15.0
BERT-L 下位 10%+pow	76	26.7	9.4	22.0
BERT-L 下位 20%+pow	72	27.6	9.7	24.0
BERT-L 下位 30%+pow	70	26.7	9.8	24.2
BERTSUM+F0	46	22.6	1.1	16.1
BERT+F0+pow	84	25.4	8.9	22.1
BERT+F0-L 下位 10%	45	22.8	1.0	16.2
BERT+F0-L 下位 20%	41	22.0	1.0	16.9
BERT+F0-L 下位 30%	37	19.5	1.0	15.9
B+pow+F0-L 下位 10%	83	26.7	8.9	23.3
B+pow+F0-L 下位 20%	79	27.6	9.3	24.1
B+pow+F0-L 下位 30%	77	26.6	9.3	24.3

によって値が向上する可能性があることがわかった。また、講義によって評価が変わるため、講義のタイプによる抽出方法の最適化について検討する必要があると考えられる。さらに、power も F0 も値が改善されても同程度であるため、発話速度などの他の韻律を用いた抽出方法について考える必要があると考えられる。

6. 終わりに

今回の発表では、書き起こしテキストと動画では重要と感じる部分に差があることを確認、受講者の「授業中重要であると指摘して欲しい部分」について調査や、各自動要約アルゴリズムによる要約の性能の比較、人手評価との比較、また、既存の方法を組み合わせることで精度が向上する可能性があることについての報告を行なった。

今後の予定としては、韻律を用いた手法の更なる調査を行い、BERTSUM の抽出結果に他の機械的手法を用いることでさらに人手評価に近いの抽出を実現するモデルの検討を行う予定である。

参考文献

- [1] Tarmo Robal, Yue Zhao, Christoph Lofi, Claudia Hauff.: IntelliEye:Enhancing MOOC Learners' Video Watching Experience through Real-Time Attention Tracking, HT '18: Proceedings of the 29th on Hypertext and Social Media, pp106-114 (2018).
- [2] Sadaaki Frui, Tomonori Kikuchi, Yousuke Shinnaka, Chiori Hori.: Speech-to-text and speech-to-speech summarization of spontaneous speech, IEEE Transactions on Speech and Audio Processing Volume:12, Issue:4, pp401-408 (2004).
- [3] Chiori Hori, Sadaaki Furui.: A new approach to automatic speech summarization, IEEE Transactions on Multimedia Volume:5, Issue:3, pp368-378 (2003).
- [4] 藤井 康寿, 山本 一公.: 重要文抽出に基づく講義音声の自動要約, 情報処理学会論文誌 51 巻, 3 号, pp1094-1106 (2010).

- [5] H.P.Luhn.: The Automatic Creation of Literature Abstracts, IBM Journal of Research and Development Volume:2, Issue:2, pp159-165 (1958).
- [6] Rada Mihalcea, Paul Tarau.: TextRank: Bringing Order into Texts, IEEE Transactions on Speech and Audio Processing Volume:12, Issue:4, pp404-411 (2004).
- [7] G.Erkan,D.R.Radev.: LexRank: Graph-based Lexical Centrality as Salience in Text Summarization, Journal of Artificial Intelligence Research Volume:22, pp457-479 (2004).
- [8] Yang Liu.: Fine-tune BERT for Extractive Summarization, arXiv preprint arXiv:1903.10318 (2019).
- [9] Chin Yew Lin.: ROUGE:A package for automatic evaluation of summaries, In Text Summarization Branches Out:Proceeding of the ACL-04 Workshop, pp.74-81 (2004).
- [10] 小林 聡, 山口 優, 中川 聖一.: 表層的言語情報と韻律情報を用いた講演音声の重要文抽出, 自然言語処理 12 巻, 6 号, pp.3-24 (2005).