



AI を用いたアクティブ・ラーニングにおける 学生評価支援手法に関する研究

鞆 大 輔

概要 大学におけるアクティブ・ラーニングでの代表的な手法であるグループワークは個々の参加学生に対する評価が困難であることから大人数での講義では導入が困難であると考えられている。しかし学生間で相互評価を行わせることでワークの成果物に対する個々の学生の貢献度を推し量る事は可能である。本研究ではこの相互評価に公平性を担保するため、AI（人工知能）を用いた学生の貢献度を可視化・評価を行うシステムの構築を行い、その実証実験および成果測定を行った。

Abstract Active learning at the university is difficult to evaluate the participating students, so it is difficult to introduce in a large number of lectures. However, evaluation can be done by letting students perform mutual evaluation. In this research, we introduce AI to this mutual evaluation to visualize the degree of student contribution. And it did the demonstration experiment and the result measurement.

キーワード AI, アクティブ・ラーニング, ルーブリック評価, 産学連携

原稿提出日 2019年5月28日

は じ め に

昨今大学教育において学習者の能動的な参加を促すアクティブ・ラーニングの実施が強く求められている。アクティブ・ラーニングにおける代表的な手法の一つであるグループワークは能動的な学習において重要な位置を占める教育手法であるが、半面、成績評価の視点から考えればグループに所属する学生がグループの成果物生成に対してどの程度貢献しているのかを把握する事は困難である。特に近年では単位認定の厳格化から講義時間外での学習時間を確保することを求めていることもあり、反転授業や事前・事後学習のような講義時間外、学校・教室外での作業が行われるケースも増加している。しかしこれらの講義時間外の学習における各学生の活動を教員が正確に把握することは非常に困難であるとともに、教員の負担が増加する。このため、グループワークはアクティブ・ラーニングにおける有益な手段であるにもかかわらず、厳密な成績評価を求められる講義科目では採用が見送られるケースも多いと考えられる。

この問題を解消するため、グループワークに参加する学生同士のやりとりを人工知能（以下、AI）によって分析し、各学生が課題作成にどの程度貢献しているかを定数化して評価を行うことで、グループワークでの成績評価支援を行うシステムを構築することとした。

なお本研究は TDC ソフト株式会社との共同研究により行われたものであり、研究分担は評価手法の考案や解析用データの収集を筆者、AI システムおよびチャットプラットフォームの構築運用を TDC ソフト、教師データの構築および実証実験および成果検証を両者の共同で実施している。

1. 研 究 背 景

1-1. 対象講義の運営方法と問題点

AI の開発および実証実験は近畿大学経営学部1年生を対象とした情報系科目「情報倫理」を対象として行った。当該科目は大教室で実施される座学の科目であり、1クラスあたりの受講者数は平均300名前後と多く、受講生数に応じて6～10名程度のグループを編成して課題に取り組む講義スタイルを採用している。

当該科目は応用倫理学に属する科目であり、明確な正解のない問題に対して学生が自ら

正しさを探求するために必要となる知識と判断能力を養う目的としている。そのため成績評価に試験を採用せず、提示された課題に対する取り組み姿勢や成果物である課題の完成度によって評価を行う。課題は指定されたテーマについて調査、議論、検討し、その結果をドキュメントとしてまとめ、最終的にドキュメント提出と発表を行う形式で行われる。1つの課題に3コマ（課題提示回、課題作成回、発表回）を割り当て、受講生は1コマ目に提示された課題を3コマ目に発表・提出するが、課題の難易度や作業工数的に講義時間内のみの作業では課題を完結する事は不可能である。そのため受講生は各課題に対して1コマあたり3～5時間程度、1課題あたり6～10時間程度のオンラインおよびオフラインでの講義外グループワークが必須となる。

役職	学籍番号	署名	出席回数	配分/累計	除籍
オピニオンリーダー	181016001		2/3	12	
資料作成	181016002		3/3	20	
調査係	181016004		1/3	0	レ
調査係	181016011		3/3	12	
リーダー	181016014		3/3	12	
ご意見番	181016016		3/3	13	
プレゼンター	181016021		2/3	11	
サブリーダー	181016032		3/3	20	
グループの評価合計は上限		100	ポイントです		

図1-1. 相互評価シートと記入例

グループワークにおける成績評価は発表や提出されたドキュメントといった教員が評価可能な成果物に基づく「グループ単位での素点」を学生が自分たちで相互評価した各課題での「貢献ポイント」に基づいて配分する方式を採用している。課題開始時にそれぞれが担当する役職を決定し、課題提出の際にグループ内でそれぞれが役職に沿ってどの程度課題作成に貢献したのかを相互評価、貢献ポイントの配分を決定し、相互評価シート（図1-1）へ記入した上で課題と共に提出を行わせる。貢献ポイントは配分可能な得点の上限が定められており、グループメンバーにこれを配分すると共に著しく貢献度の低い「グループへのただ乗り」を目論む学生（以下、フリーライダーと呼称）については受講生の協議によりグループからの追放を可能としている。この仕組みは成績評価の公平性を担保しフリーライダーの発生を抑止する事を目的としており、課題に積極的に取り組む学生か

らは好意的に評価されている。

しかしこの手法も完全とは言い難く、毎年いくつかのグループから貢献ポイント配分を巡る諍いについて相談が寄せられている。相談の内容は概ね「声高に自分にも得点を分配しろと強弁する学生に対し、仕方なく配分を行ったがやはり納得できない」というものであり、多くの場合このような主張を行う学生はフリーライダーもしくはその傾向が強い不真面目な学生である。まじめな学生に対する公平な評価を実施するためにも何らかの対策が必要であることは明白であり、この問題を受けて2016年度からは評価の公平性と得点配分基準の明示を目的として学生の相互評価指針としてループリック評価による数値化を導入したが、効果は限定的であった。

1-2. 評価支援 AI 開発への経緯

2017年度はループリック評価をさらに推し進める形で AI による評価の導入を行うこととした。AI による評価を実施する契機となったのはループリック評価の導入を開始した2016年末、文教分野での知見蓄積を模索していた TDC ソフト株式会社から大学との共同研究に関する打診を受けたことによる。当初は大人数教室での講義における学生—教員間、学生—学生間の相互コミュニケーションを促進する、クリッカーや SNS、小テストシステムを統合した e—ラーニングプラットフォームの開発を想定していたが、これらの機能は Moodle 等の既存システムを併用することである程度対応可能なものであったため、研究および技術開発を目的に先端技術を用いた手法の研究へとテーマを変更。研究テーマに関する議論の中で前節に述べたグループワークにおける評価の公平性担保の問題について検討した結果、この問題を人工知能を用いて解決を試みる研究を行うこととなった。

研究テーマ決定後、2017年度上半期中にシステムデザインについての検討を実施。従来の学生による相互評価を用いた成績評価手法を基に AI が学生の貢献度を評価する手法を検討した結果、グループワークに関連するやりとりを Moodle 上のチャットで行わせ、その会話ログを AI が自然言語解析、これを基に貢献度を評価する手法を採用することとなった。

これらの経緯を経て、2017年9月より正式に共同研究案件として契約を締結、AI の開発に着手した。

2. AI による評価手法

2-1. 評価手法システムの概略

AI による評価を実装するにあたり検討を要した項目は次の2点である。まず AI が「貢献度」を評価する基準となる教師データをどのように構築するかと言う点。次に AI が評価を行う対象となるデータ、すなわちグループワークに参加している学生の取り組み姿勢をどのような形式でデータ化するかというデータソースに関する問題である。

教員やファシリテーターが同席してグループワークにおける参加者の貢献度を判定する場合、参加者同士の会話に注視する事が多いため、本来であれば AI による貢献度判定も口頭による会話、すなわち音声データを元に評価を行うことが望ましい。しかし現時点での音声認識、テキスト化技術ではディスカッション時の口語表現を正確に文字起こしする事が困難であり、実現化には長い道のりを要する事が想定された。そのため、学生の多くが課外時間におけるグループワークを SNS やメッセージ等ツールを用いたオンライン上でのチャットで実施している点に着目、チャット上でのテキスト（ログ）を AI のデータソースとすることとした。

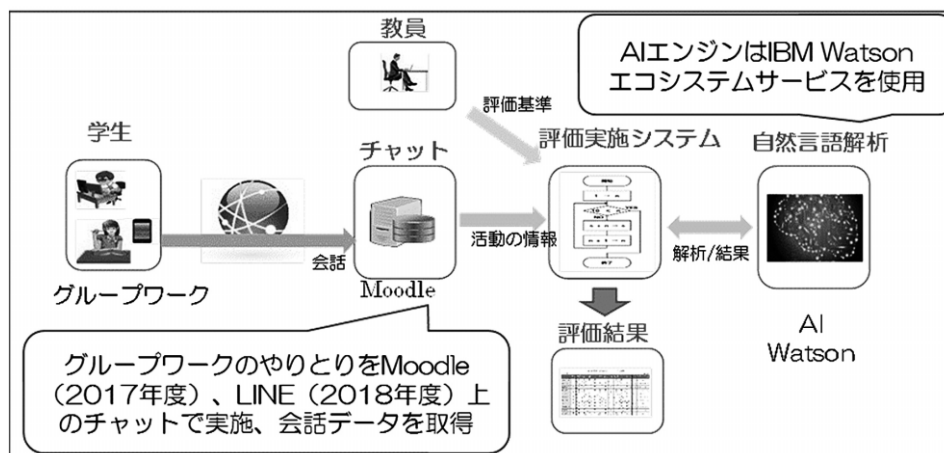


図 2-1. AI による評価システムの構成

また、貢献度の評価には紙ベースでの評価に用いているルーブリック評価と同じようにリーダー、ドキュメンタライター、プレゼンター等のグループワークにおける役割を元にした9つの評価クラスを設定。それぞれの役割でどのような活動が課題作成への貢献とな

るかについて0～+3までの4段階の評価項目を設定し、チャット上での発言を元に評価値を加算・減算する形で評点を決定する形とした。なおループリック評価の評価項目と内容、および実装状況については表2-1に詳細を記載する。

ログ採取用のチャットシステムについては2017年度の実証実験では標準eラーニングプラットフォームであるMoodleを使用した。学生にとっては馴染みのないシステムであることが起因しチャットの発言数が少なくなる傾向が見られたため、2018年度はLINEでのグループトークのデータを用いることとした。自然言語解析および評価のためのAIエンジンにはIBM WATSON（日本語版オンラインエコシステム）を使用し、学生がオンラインでの発言した会話ログを評価、結果を出力する仕組みとした。

なお評価システムの構築にあたり、AIによる評価は教員や学生による評価の代替物ではなく、あくまでも学生の相互評価を支援するツールとしての位置づけを心がけた。AIによる評価は採点結果を決定するものではなく、あくまでもグループワークにおける相互評価に公平性を担保する指標となることを期待したものである。そのため、AIを導入することによる具体的な効果としてはグループワークに参加している各学生の貢献度を可視化することで評価の目安とする事、グループワークに積極的に参加せずに成果のただ取りを目論むフリーライダーを検出、排除する根拠を与え公平な評価を行える環境を実現する事、もしくはフリーライドしづらい環境を実現する事を目的としている。

2-2. AIによる評価手法

先に述べたとおり、AIによる貢献度評価手法は人間の目視による処理をトレースする形を採用している。例えば図2-2のような会話が行われた場合、人間はチャットログの発言者である学生A、B、Cについて、学生Aはグループワークの進行を司っているリーダー的な学生、学生Bは割り当てられた課題を行っていないトラブルメーカー的な学生、学生Cは自分の課題以外に他者の課題をフォローするサポート的な立ち回りを行っている学生であるという印象を抱く。例のような短い会話ログからも学生AとCは課題作成に貢献しており、学生Bは貢献度が低い、もしくはグループ全体の作業を阻害しているという評価を行う事が可能であり、このようなそれぞれの「役割」に特徴的な発言をチャットログから抽出、出現回数や発言内容を評価する事で各学生の役割毎のスコアリングを行う。

例えばリーダーの役割であるグループ全体のマネジメントの評価項目については図2-3のようなルールで評価を行っている。ループリック評価におけるグループマネジメントの役割はグループ全体の総括や運営を司るものとしており、課題達成の可否、および課題へ

の全員参加が達成されているかと言う点を評価基準としている。これを教師データ化するにあたり、グループ全体の発言頻度による評価軸とリータとして相応しい発言を検出した回数を評価軸としている。ただしグループ全体の発言数は参加メンバー全員に共通するデータであるため、リーダーとしての発言が多い学生をリーダーもしくは副リーダーと認定し、これら役職についていると判定された学生にのみ上記発言数によるスコアリングを

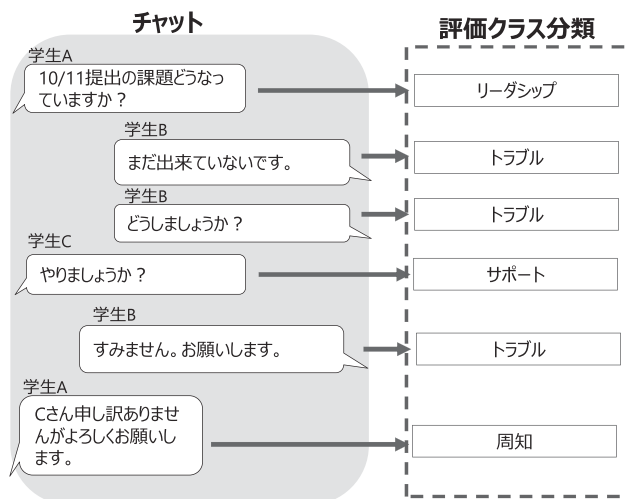


図 2-2. チャット内容の例と評価クラス分類

項目名称／対応役職	C(+0)	B(+1)	A(+2)	S(+3)
グループマネジメント ※リーダー	□求められた課題を達成出来なかった、もしくはこの項目に該当しない	□課題は完成したが、参加していないメンバーが多数いた	□ほとんどのメンバーが参加し、課題が完成した	□全員参加で課題を完成させた

評価したいこと	評価内容	取得すべき情報	具体的な会話の内容 (例)	メモ
グループ全体での課題達成	メンバーの会話数が偏りがなく多いこと。	メンバーの参加率 メンバーの会話数	-	同種の連続会話は1回の発言としてカウントし、グループ全体の会話数が多い程議論が活発と判断
プロデューサー、コーディネータとしての行動	リーダーシップに関する会話の数の評価	指示に関する会話	xxしてなあ xxしてください どうしますか？※1	メンバー学生の場合は参加熱意として評価
		決定に関する会話	xxでやります ではXXで	
		周知に関する会話	10/11 xx実施 xxしましょう xx追加します	その他「囃」のワードが含まれる場合も該当

図 2-3. 評価基準の例（グループマネジメント）

表2-1. ルーブリック評価と実装状況

項目名称／対応役職	概略	C(+0)	B(+1)	A(+2)	S(+3)
グループマネジメント ※リーダー	グループの統括・運営に関する評価項目	<input type="checkbox"/> 求められた課題を達成出来なかった、もしくはこの項目に該当しない役職である	<input type="checkbox"/> 課題は完成したが、参加していないメンバーが多数いた	<input type="checkbox"/> ほとんどのメンバーが参加し、課題が完成した	<input type="checkbox"/> 全員参加で課題を完成させた
トラブルシュート ※リーダー	グループ活動時のトラブル解決に関する評価項目	<input type="checkbox"/> トラブルが生じたが上手く解決出来なかった、もしくはこの項目に該当しない役職である	<input type="checkbox"/> トラブルが生じ、解決したが問題が残った	<input type="checkbox"/> トラブルが生じたが上手く解決した	<input type="checkbox"/> トラブルの発生を未然に防いだ
フォローアップ ※メンバー全員	グループメンバーとしての共同作業に関する評価項目	<input type="checkbox"/> 与えられた分担を完了できていない	<input type="checkbox"/> 与えられた分担を完了した	<input type="checkbox"/> 与えられた分担分以外にも自発的な作業を行った	<input type="checkbox"/> リーダーを補佐してグループ運営に貢献した
サポート ※メンバー全員	担当業務外で他者をサポートした事に対する評価項目	<input type="checkbox"/> 他のメンバーにサポートされた	<input type="checkbox"/> 他のメンバーのサポートは行っていないが、他人のサポートも受けていない	<input type="checkbox"/> 他のメンバーのサポートを行った	<input type="checkbox"/> 複数メンバーに対するサポートやフォローを行った
リサーチ ※メンバー全員	課題作成に必要な事項の調査活動に関する評価項目	<input type="checkbox"/> 調査活動に参加していない	<input type="checkbox"/> 調査活動に参加したが、締切は守れなかった	<input type="checkbox"/> 締切を守って調査活動を終えた	<input type="checkbox"/> 要求された以上の精度で調査を行った
アイデアクリエイション ※オピニオンリーダー	課題に対するアイデア・意見創出に関する評価項目	<input type="checkbox"/> アイデア創出に参加していない	<input type="checkbox"/> アイデア創出に参加したが、従属的な立場だった	<input type="checkbox"/> アイデア創出に積極的に参加した	<input type="checkbox"/> グループのアイデア創出を牽引する立場として活躍した
ドキュメントライティング ※ドキュメントライター	課題として提出するPowerPoint等の文書作成に関する評価項目	<input type="checkbox"/> 課題文書の作成に関与していない	<input type="checkbox"/> 文書作成に参加したが、締切は守れなかった	<input type="checkbox"/> 締切を守って文書を作成した	<input type="checkbox"/> 要求された以上の品質で文書を作成した
プレゼンテーション ※プレゼンター	発表に関する評価項目	<input type="checkbox"/> 発表には関与していない	<input type="checkbox"/> 発表リハーサルのサポート、もしくは発表申請を行った	<input type="checkbox"/> 発表者としてリハーサルを行ったが登壇はしていない	<input type="checkbox"/> 発表者として登壇した
クリティカルシンキング ※ご意見番	課題の完成度を高める批判的嗜好に関する評価項目	<input type="checkbox"/> 批判的視点からの意見を行わない、もしくは批判的な意見がグループの空気を悪くした	<input type="checkbox"/> 批判的視点からの意見したが改善への影響は無かった	<input type="checkbox"/> 批判的視点からの意見によって改善になんらかの影響を与えた	<input type="checkbox"/> 批判的視点からの意見で課題の完成度が高まった

背景白: AIによる評価が可能な項目

背景灰: 現時点ではAI評価困難だが改善可能な項目

背景黒: 発言数が少ない、もしくは現在のAI能力では評価不能な項目

適用する形としている。他の役割についても同様の手法で会話内容と補助的に算出可能なデータを併用する形でスコアリングを実施、項目毎の評価を実現した。

しかしいくつかの項目についてはチャットでの発言で判定が不能であったり、事例が不足していたりするため具体的な教師データの構築が行えなかった。例えばプレゼンターの評価項目である「発表者として登壇」したかどうかと言う判定は課題提出日当日でなければ判定出来ず、課題作成時のチャットログからは「登壇予定」を推測することしかできない。またトラブルシューティングに関する項目については、トラブルが発生した事例自体が少ない上に各事例での状況がまちまちであるため、トラブルの予兆を示す会話データを特定することが困難であった。

また AI エンジンである Watson の機能的な問題から、複数の話者による会話の流れを認識することが出来ず、それぞれの発言単体での評価を余儀なくされた点も教師データ構築の障害となった。特にサポートに関する会話はサポートする側とされる側を対として認識する必要があるが、これが実現出来ないため有効な教師データを構築することが不可能であった。

以上のような点を踏まえ、2017年度終了時点での教師データ実装状況が表 2-1 である。図中、白背景の項目は評価手法が確立できた項目、灰色背景の項目は精度に若干難があるもののある程度の評価が行える、もしくは今後改善が期待できる項目、黒背景の部分は先に述べた理由から評価手法が確立できていない項目となる。全36項目のうち17項目が白、8項目が灰色、11項目が黒となっており、約 3 割の項目で評価が困難な状態ではあったが評価手法そのものの有効性を検証する必要性があったことからこの実装状況で実証実験を行うこととした。

3. 実証実験による検証結果

3-1. 実証実験までの流れ

開発初年次である2017年度は AI を用いた評価そのものの実現可能性検証を主要目的都市、さらには AI を用いた評価の運用手法の確立、評価精度の測定を目的とした実証実験に取り組んだ。実証実験の対象科目である科目、情報倫理では1セメスターの講義期間中に4回の課題が実施されるため、1回目の課題で学生が発言したチャットログを採取、これを元に基礎的な教師データの作成を行った。教師データがある程度整った段階で2回目、3回目の課題でのチャットログを AI に試験的に評価させ、その結果を学生に提示して評

価精度の確認を行い、学生からの指摘に基づき教師データを改良する PDCA サイクルで開発を実施した。

またこの時点では AI による評価精度が実用的なレベルではないことが予想されたため、学生が行う相互評価に AI の誤評価が影響を及ぼすことを避けるため、学生がループリック評価法に基づく手動での相互評価を提出した翌週に参考データとして AI による評価結果を提示する方式を採用した。AI による評価結果と学生による相互評価結果との比較で教師データの修正を行った結果、3 回目の課題実施時に評価精度が一定のレベルに到達したと判断されたことから最終課題となる 4 回目の課題では学生が相互評価を行う際の参考データとして AI による評価結果の提示を行い、学生はこれを参照しながら相互評価を行った。この方式は当初から想定していた学生の自己評価を支援する指標としての AI 利用であり、開発初年度である2017年度中に AI による貢献度評価手法とその運用スタイルの確立に一定の道筋を付けることが出来た点は僥倖であった。

3-2. 実証実験の評価

4 回目の課題では運用方法の確立と同時に、AI による評価の妥当性を検証するため、受講生に対して「AI の評価に対するメタ評価」を行わせる形で検証を行った。

このメタ評価の結果が図 3-1 である。学生が自分たちで相互評価を行った結果（学生評価）と AI によるチャット分析結果（AI 評価）のどちらを最終的な自己評価の数値としたかを個人別／評価項目別で調査したものを「AI 評価の受容度」として表している。学生と AI の評価が合致した項目については「評価合致」とし、学生と AI の評価が異なった場合、最終的に提出した評価を決定する際に学生自身の評価結果を採用した（＝AI の評価を受け入れなかった）場合は「学生評価採用」、逆に AI の評価を採用して評価結果を修正した場合は「AI 評価採用」と分類した。この 3 分類のうち、「評価合致」および「AI 評価採用」は学生が AI の評価を受け入れていると判断し、「学生評価採用」については AI の評価内容に対する不満があると判断した。

評価項目毎に受容度のばらつきはあるものの、AI の評価を優先したという回答が全体の14.0%、学生評価と AI 評価が合致していたと言う回答が62.7%であり、全評価項目のうち76.7%が学生に受容されているとの回答を得た。この数値は教師データの実装状況から考えると十分な精度であると評することが出来る。

個別項目に関する分析として、興味深い点は AI による評価が十分に実装できていないサポートおよびトラブルシュートの両項目で学生の評価が分かれた点である。トラブル

シュートは本来トラブル收拾や未然防止による評価項目であるため、トラブルが発生しなかったグループについては全員が評価 0 となる。そのため、AI による評価が不能（＝ 0 評価）であっても、トラブルが発生しなかったグループについては学生の評価と相違しない事になり、結果として受容度が高まったと考えられる。一方でサポートについては会話ログを確認した際にグループワークの中で日常的に行われていた事が目視では確認されていたが、AI はこれを評価できていないため、評価精度が低いというメタ評価に繋がったと考えられる。

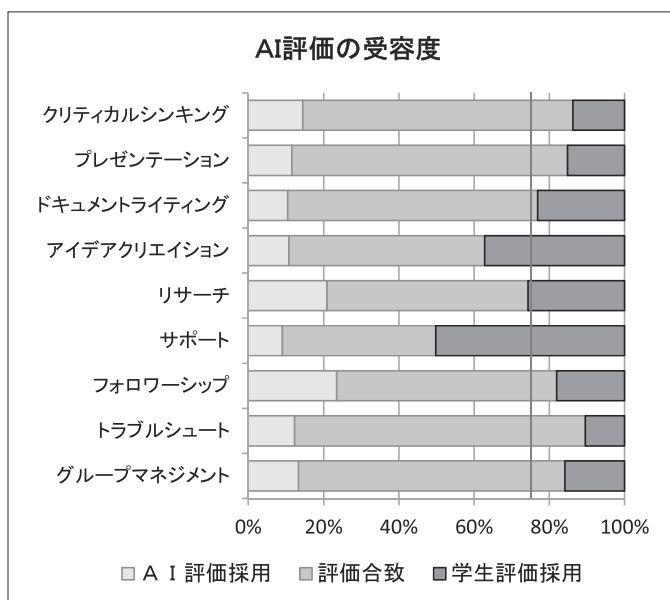


図 3-1. 学生の AI 評価受容度

また受容度が項目平均を下回ったアイデアクリエーションについては「アイデアが役だったかどうか」「積極的・従属的な立場かどうか」と言う、前後の会話の文脈を評価する必要がある評価基準であったことから評価精度が十分に高められなかったと考えられる。

これらの点から総合的に考えると2017年度の実証実験の結果は、AI による貢献度評価そのものは一定の精度で可能であると言う事、また教師データの生成方法や評価基準そのものの見直しなど、評価精度を高めるための工夫が必要という2点が結論となる。

次に AI による評価を導入することで期待できる効果として挙げられるフリーライダーの検出および排除、もしくは抑止の効果について検証したものが表 3-1 である。除籍率の数値は学生が相互評価の際にフリーライダーを検出、グループから除籍した割合を示す

表 3-1. フリーライダー排除率の推移

年度	除籍者数	受講者数	除籍率	備考
2017	31	428	7.2%	AI評価導入
2016	24	473	5.1%	ループリック導入
2015	22	605	3.6%	
2014	25	639	3.9%	
2013	27	673	4.0%	履修取消制度導入
2012	15	308	4.9%	
2011	41	956	4.3%	

数値であり、この数値がフリーライダーの排除率に相当する。

備考欄に記載したとおり、2011年から2017年までの間に3つの制度・運用上の変更点に加わり、この事がフリーライダーの排除に影響を及ぼしていることが見て取れる。2011年および2012年は GPA 制度や履修取り消し制度が導入される以前の年度であり、学習意欲の乏しい学生も最後まで履修を継続していたことからフリーライダーの割合が比較的多い年度であった。2013以降は履修取り消し制度が導入され、フリーライドを企む受講モチベーションの低い学生は早期に講義からドロップアウトしたためか除籍率は年々低下する傾向にある。2016年および2017年はフリーライダーの検出強化を行った年度であり、ループリック評価の導入（2016）および AI による評価の導入（2017）はそれぞれ単年度のデータではあるものの、除籍率を向上させており、フリーライダーの検出と排除に一定の効果が出ていることが推測される。このことから各種取り組みによるフリーライダーの排除は成績評価の公平性を担保する意味では当初の目標通りの成果が出ていると言えるが、一方で本来大学での講義のあり方としてはフリーライダーを発生させないが理想であることは言うまでも無い。そのため、今後は検出と排除ではなくフリーライダー発生を予測し、受講モチベーションの下がった学生を事前に注意、指導することで受講態度の改善を求める仕組みを構築することが必要である。

4. 今後の展望

現状の AI 評価における問題としては、AI エンジンである WATSON の機能的な制約から連続した文章の流れを評価出来ず、会話の流れを正確に追いつけていない箇所がある点が挙げられる。これが原因となりループリック評価で設定した評価が適正に判断できておらず、特に他人の発言との関係性を評価する必要があるサポートの項目において顕著な影響が見うけられる。また感情の変遷についても現状の AI では解析できないため、トラ

AI を用いたアクティブ・ラーニングにおける学生評価支援手法に関する研究（輅）

ブルの徴候や発生についても検出率が困難な状況となっている。

これらの問題は連文節解析や感情認識と言った AI の拡張 API を導入、あるいは他の AI エンジンを用いる事で解消が可能であると考えられ、2019年度の実証実験に向けた新しいアプローチを模索中である。

その他改善点として学生から Moodle のチャットが使用しづらいと言う声が多数寄せられたため、2018年度からは学生が使い慣れたメッセージングアプリ LINE へ運用環境を変更した他、現段階のテキストベースのログ解析では拾いきれない口頭会話も評価対象とすべく、スマートスピーカーやモバイルデバイスによる音声入力・テキスト化を用いて、オフラインでのワークについても研究を行っている。

また現状では AI による評価はフリーライダーの検出と排除に有効ではあるが、教育的な側面から言えばフリーライダーが発生しにくい環境、すなわち参加学生全員がグループワークに取り組める環境を構築する事が望ましい。そのため将来的には AI による評価を基に各メンバーの能力適性ややる気を測定し、グループワークを円滑に進めるためのアドバイスやサポートを行う機能をパッケージングしたアダプティブ・ラーニング（個人適応型学習）用のツールとして、AI パーソナルチューター的なシステムへの発展を計画している。

参 考 文 献

- Google. (日付不明). Google Cloud Natural Language API. 参照日：2019年5月3日，参照先：
<https://cloud.google.com/natural-language/?hl=ja>
- IBM. (日付不明). IBM Watson Natural Language Classifier - Japan. 参照日：2019年5月3日，参照先：<https://www.ibm.com/watson/jp-ja/developercloud/nl-classifier.html>
- 文部科学省. (2012年8月28日). 新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～（答申）. 参照日：2019年5月3日，参照先：http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1325047.htm