

機械学習を用いた環境分野に興味を持つ学生の意識分析 —AIはどこまで分析に役立つのか?—

家 富 誠 敏

【要旨】

本稿では、近年注目を集めている人工知能（AI）による機械学習（Machine Learning）[1][2]を用いて、アンケート調査結果の分析を行う。機械学習を用いることで、人間では発見困難なデータ間の因果関係を分析できる可能性がある。しかし、全く現実を反映していない間違った分析結果が得られることもある。いくつかの意識調査アンケートの結果を機械学習で分析し、有用な結果が得られるか否かを実験する。

キーワード：機械学習、意識調査、学生アンケート、傾向分析、環境分野

1. 本研究の目的

2017年にインターネット囲碁サイトに現れたアカウント「master」は、瞬く間にトッププロ達を相手に無敗の60連勝をあげ、その正体は誰なのかと話題を呼んだ。その正体がGoogle社の開発した囲碁AIであることが公表され、社会に衝撃が走ったことは記憶に新しい[3]。

囲碁AIの戦い方は、人間とは異なる部分が多かった。人間には理解不能な戦略で勝利を続けると「もはや人間ではAIに太刀打ちできない」と悲観する人もいれば、逆に「従来にない新たな戦略を検討するチャンスが生まれた」と歓喜する人もいた。良くも悪くも、AIが新しい風を棋界に呼び込んだのである。

さて、AI（人工知能）の活躍は、囲碁の世界だけではない。たとえば「機械の電子制御」「人間との対話」「データの分析」等、さまざまな分野で人間の知能を代替する活躍が考えられる。とはいえ、現在のAI技術では出来ないことも沢山ある。AI技術は、果たして「どのような分野」で「どこまでの活躍」が期待できるのだろうか。

本研究では、意識調査アンケートの結果分析にAIを利用し、既存のAI技術でどこまで有益な分析結果を抽出できるかを実験する。アンケート調査の結果分析は、調査対象によっては膨大な量のデータを処理する必要がある。人手での情報抽出には、多大な負荷や作業時間を要することが懸念される。もしAIに任せることが出来れば、たとえ大規模な調査であっても、負荷の軽減が大いに期待できるだろう。

また機械学習によって、これまで人間が見落としていた知見を発掘できる可能性も期待できる。しかし、既存の AI 技術では、それほど有益な分析結果が得られないことも十分考えられる。実験によって「どこまで分析できるのか」を検証することが本研究の目的である。

2. 決定木を予測モデルとした機械学習の仕組み

本研究では、AI を用いてアンケート調査結果の分析を行うが、AI と言っても様々な種類がある。ここでは本実験に用いる「決定木を予測モデルとした機械学習」について解説する。

2.1 機械学習と非機械学習

まず AI を大別すると「機械学習」と「非機械学習」に分類される。機械学習のパイオニアであるアーサー・サミュエルは、機械学習を「明示的にプログラムしなくても学習する能力をコンピュータに与える研究分野 (The field of study that gives computers the ability to learn without being explicitly programmed.)」と定義している[4]。

すなわちデータ分析を行う際、分析方法を人間が明示的に与えずとも「コンピュータが自分で分析方法を学習する方式」が機械学習である。実は、冒頭で述べた囲碁 AI も、機械学習方式である。その戦略は人間がプログラムしたものではなく、コンピュータ自身が学習によって必勝法を編み出している。そのため、人間では考えつかない妙手を取ることもあれば、学習が不完全な段階では、人間なら絶対しない誤手もあったという。

対して、コンピュータが自分で学習しない方式、言い換えれば予め人間が明示的に分析方法等をプログラミングして、その通りに計算するものが非機械学習である。家電等に内蔵されて電子制御を行う AI は、殆どが非機械学習方式である。

本研究では、人手では得られない知見を新たに発見できる可能性に期待して、機械学習方式を用いて、アンケート結果の分析を行う。

2.2 決定木を用いた機械学習

機械学習にも、学習方法等の違いによって様々な種類がある。本研究では「決定木学習 (Decision Tree Learning)」[5]という方式を用いる。決定木学習は、コンピュータが学習した結果を「決定木 (Decision Tree)」というデータ形式で表現する。決定木の例を図 1 に示す。

決定木は、木を模した構造となっており、節点 (ノード) を枝 (エッジ) で結んだ形で表現される。節点は「変数 (質問)」に、節点から下に伸びる枝は「変数値 (回答) による条件分岐」を表現している。決定木を参照する際は、最上段の節点から、枝による条件分岐に従って、末端の節点まで辿っていく形となる。決定木学習では、与えられたデータを基に学習した結果を、この決定木を用いて表現する。決定木は学習結果であると同時に、次に新たなデータが与えられた際にコンピュータが判断に用いる思考手順でもある。

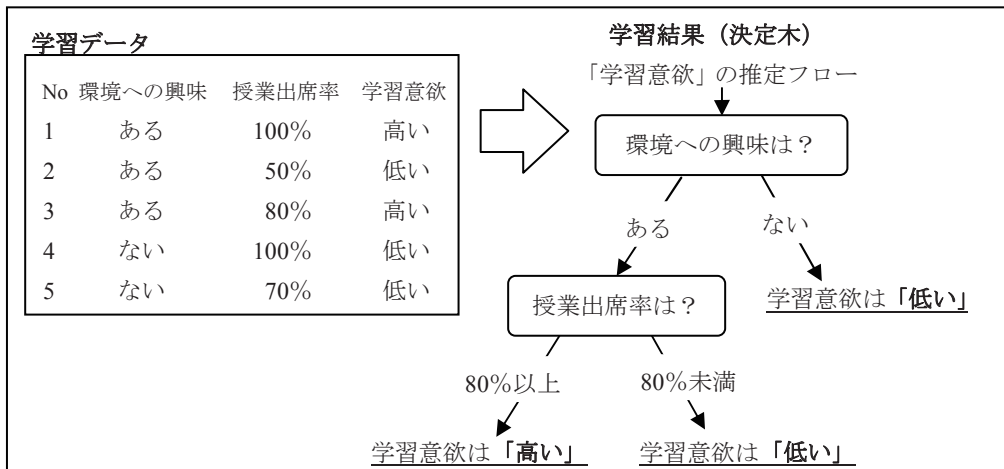


図1 決定木の例

決定木学習の特徴は、機械学習した結果が決定木で表現されるため、コンピュータの学習結果が人間にも理解しやすいことである。実は、冒頭でふれた囲碁 AI は、学習結果が非常に複雑なデータ構造（Deep Learning：多段階層ニューラルネットワーク方式[6]）であるため、コンピュータの学習結果を人間が見ても如何なる戦略を学習しているのか理解できないという特徴を持っていた。

本稿ではアンケート結果の機械学習を行うが、コンピュータにデータ傾向や相関等を分析させたとして、分析結果を人間が把握できないのでは、活用範囲が限られてしまう。対戦に勝てば人間に理解不能な戦略でも十分に役立つ囲碁 AI とは異なり、分析結果を人間が有効活用するためには、人間が解る形で結果を提示させる必要がある。そのため今回は、決定木を用いた機械学習で、アンケート結果を分析することとした。勿論、アンケート結果の傾向をコンピュータに学習させ、次にアンケートを実施した際の「回答予測」をするだけならば、（囲碁 AI のような）人間には理解不能なデータ構造で学習結果を表現する方式でも十分に役立つ可能性はあるが、本稿では研究対象外とした。

2.3 決定木を用いた機械学習の仕組み

決定木学習において、学習用データを基に決定木を作成する流れは、次の通りである。

- 1) 最も不純なデータが少なくなる分割方法を探す
- 2) 分割方法を決定木の節点および枝として加える
- 3) 分割後のデータについて、1)～2)を繰り返し、決定木の節点および枝を伸ばす
- 4) 十分に分割できたら終了する

以降、各部について詳しく説明する。

2.3.1 データの不純度

最も不純なデータが少なくなる分割方法を探す際、指標となるのが「データの不純度」である。たとえば、以下のような列 A と列 B が存在したとする。

列 A = <○、○、△、○、×> ○が占める割合：3/5、○以外が占める割合 2/5
列 B = <○、○、×、○、○> ○が占める割合：4/5、○以外が占める割合 1/5

列 A および B では、共に最も数が多い要素は○である。これが列を構成する「主たる要素」となっている。逆に○以外の要素は、不純物と考えられる。集合や列において、不純物（主たる要素以外）が占める割合を「不純度」と呼ぶ。上記の例では、列 A の不純度は 2/5、列 B の不純度は 1/5 と計算される。

2.3.2 不純度を用いた分割

決定木学習では、学習用データに対して様々な分割方法を試し、最も不純度が低くなる分割方法を探す。先ほどの列 A を例にすれば、列 A を 2 つに分割する方法は以下の組み合わせが考えられる。

分割方法 1：<○> <○、△、○、×>	左列不純物 0 + 右列不純物 2 = 計 2
分割方法 2：<○、○> <△、○、×>	左列不純物 0 + 右列不純物 2 = 計 2
分割方法 3：<○、○、△> <○、×>	左列不純物 1 + 右列不純物 1 = 計 2
分割方法 4：<○、○、△、○> <×>	左列不純物 1 + 右列不純物 0 = 計 1

列を分割した後の不純度は「分割後の各列における主たる要素」を考え、それ以外（不純物）の要素数を各々カウントして合計し、全体要素数で割れば算出できる。

たとえば「分割方法 1」の場合、左列<○>の主たる要素は○であり不純物は 0 個、右列<○、△、○、×>の主たる要素は○であり不純物は 2 個となっているため、合計不純物数は 0+2=2 となる。これを左列+右列の合計要素数 5 で割ると、不純度は 2/5 と計算される。ちなみに「分割方法 2」の右列<△、○、×>の場合、各要素が同数存在するが、何れを主たる要素と考えてもそれ以外の要素数は 2 となるため不純物数は 2 となる。

分割方法 1~4 を比較すると、分割方法 1~3 は不純度 2/5、分割方法 4 は不純度 1/5 であるため、最も不純度が低い分割方法は「分割方法 4」となる。これにより分割方法 4 が採用され、決定木に分割方法 4 を反映した図 2 のような節点および枝が追加される。

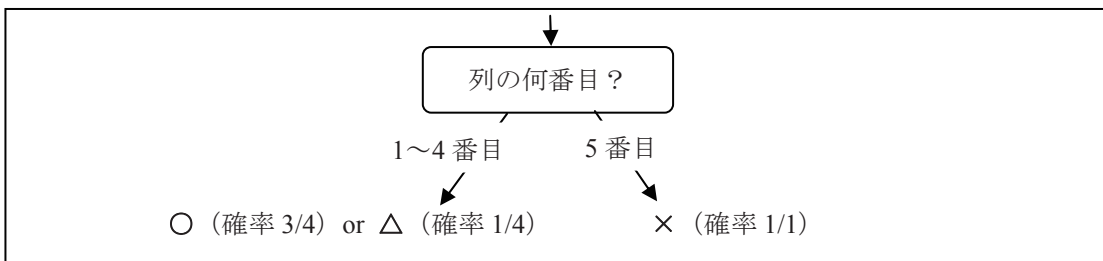


図 2 分割方法 4 を反映した決定木

図2を見ると、右枝側は不純度0、すなわち確率100%でデータを確定できているが、左枝側は不純物があり未確定である。この場合、分割方法4の左列<○、○、△、○>に対して、再び最も不純度が低くなる分割を探すことになる。不純度が十分に低くなるまで分割を繰り返し、それに従い、決定木の節点および枝を伸ばしていく。

また複数の属性、たとえばアンケートで複数の質問項目 Q1、Q2、Q3…等がある場合には、Q1 についての分割以外にも、Q2、Q3…に関する分割も試し、それらの中から最も不純度が低くなる分割を探して、決定木に反映する。これら試行には膨大な組み合わせの探索が必要となる。大量かつ複雑な計算が求められる機械学習は、まさにコンピュータ技術の発展により可能となった分野であると言えよう。

2.3.3 剪定による決定木の最適化

列 A を不純度が 0、すなわち値同定確率 100%になるまで決定木の節点および枝を伸ばすと、最終的に図3のようになる。

図3の決定木は、学習データを完全に再現しているため、一見、確度が高いように思える。しかし、もし学習データにノイズが含まれていた場合には、ノイズも完全再現している可能性が考えられる。これをオーバーフィッティング (Over Fitting) と呼ぶ。オーバーフィッティングを抑え、有意な学習結果を得るためには、ノイズを反映している節点および枝を削除、すなわち「剪定」することが求められる。

しかし、学習データのうち何れがノイズに相当するのかを判別することは難しい。そのため剪定する節点および枝の抽出には、K-分割交差検証法 (K-fold Cross-validation) が用いられることが多い。

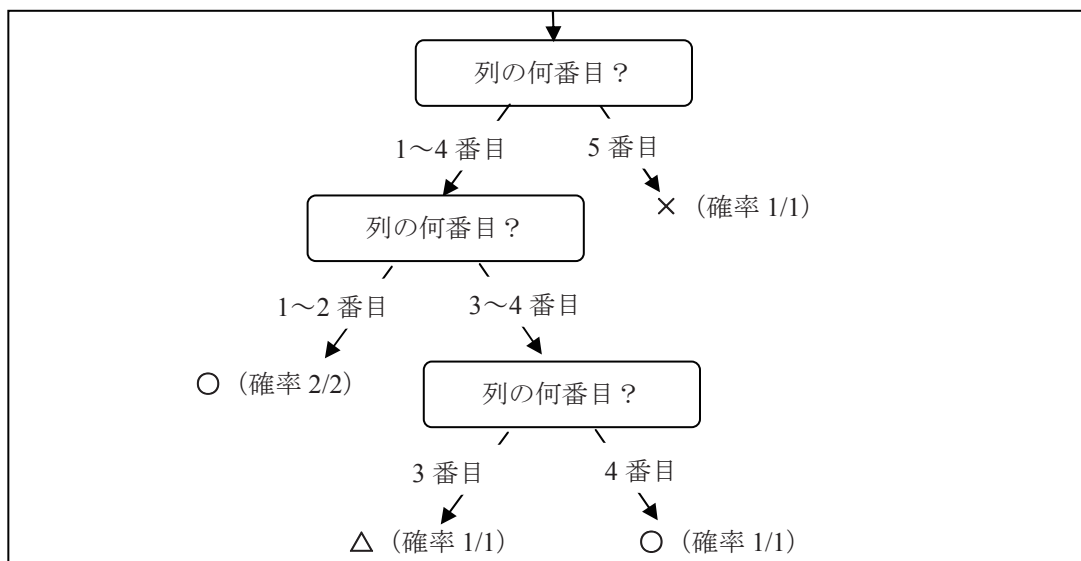


図3 不純度0になるまで伸ばした決定木

K-分割交差検証法を用いた剪定は、以下の手順で行われる。

- 1) 学習用データを K 個のグループに分割する
- 2) うち (K-1) 個のグループを選んで学習データとして用い、決定木を作成する
- 3) 作成した決定木を、残った 1 グループで検証しながら、剪定を行う
 - 決定木を用いて値の推定を行い、正しく推定できるか確認する (推定誤りを検出する)
 - 決定木の末端から剪定を行い、推定誤りが減少すれば、剪定を反映する
- 4) 2) ~3) の試行を学習データの組み合わせを変えて、K 回行う
- 5) K 回を平均して、推定誤りが全体的に減少する剪定法を選択する

3. 機械学習を用いた意識調査アンケート結果の分析

本研究では、機械学習を用いてアンケート結果の分析を行った。分析には「機械学習ソフトウェア Weka (Waikato Environment for Knowledge Analysis)」[7]を使用した。また学習アルゴリズムは、決定木学習で著名な「C4.5」[8]を選択し、剪定等のパラメータ算出には「10-分割交差検証 (K-分割交差検証を K=10 で実施)」を用いて機械学習を行った。

以下、いくつかのアンケート結果に対して、機械学習の分析を行った結果を紹介する。

3.1 卒業予定者に対する意識調査アンケートの分析 ～4年間満足の要因は何か～

最初に紹介するのは 2016 年に環境社会学部で実施した卒業予定者 (4 年生) に対する意識調査アンケートの分析結果である。本アンケートの実施概要を以下に示す。

調査日	: 2016 年 2 月下旬
対象	: 環境社会学部卒業予定者 回答者数 : 26 名
	※質問項目は「付録 A」参照

この分析では、アンケート項目のうち「環境社会学部で 4 年間学んで良かったと思いますか」に対して、「〇4 年間に満足している (①かなりそう思う or ②そう思う)」と「×4 年間に満足していない (③どちらともいえない or ④あまりそう思わない or ⑤全くそう思わない)」とに解答した 2 つのグループに分け、他の質問項目から何れのグループに属するのかを機械学習で推定させた。図 4 に学習結果の決定木を示す。

1) 「自分のしたい勉強が出来たか」が第 1 要因

まず図 4 の最上段の分岐を見ると、「〇4 年間に満足している」or「×4 年間満足していない」かの分岐は、質問項目「自分のしたい勉強が出来たか」に大きく依存していることがわかる。「自分のしたい勉強が出来たか」の質問に「④それほど満足していない」および「⑤全く満足していない」と回答した人は「×4 年間満足していない」を選ぶ傾向が見られる。

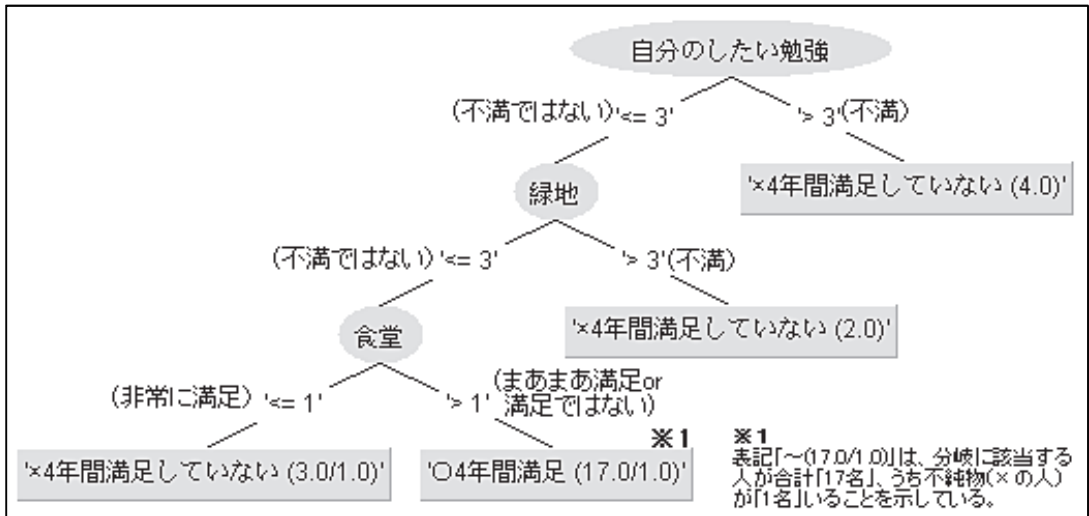


図4 大学4年間を満足している人の傾向（機械学習の結果）

2) 「キャンパス環境」関連が第2要因

対して「自分したい勉強が出来たか」に「①非常に満足」or「②まあまあ満足」or「③どちらでもない」を選んだ場合、「キャンパスの緑地（外構の緑地・オープンスペース）」に不満があると「×4年間満足していない」を選ぶ傾向が見られる。元データを検証してみると「キャンパスの緑地」への満足度は「図書館」「グラウンド」等の満足度とも強い相関があり、総じて「キャンパス環境」関連に不満を感じていると、「自分がしたい勉強が出来た」としても「×4年間満足していない」との回答につながっていると考えられる。

3) 「キャンパス施設を利用しない人」は自前で満足？

最後に「自分がしたい勉強」「キャンパスの緑地」の両方を満足している学生は「食堂」への満足度によって分岐している。不思議な結果であるが「食堂」に「①非常に満足」を選ぶと「×4年間満足していない」につながっており、一見逆転した相関である。

学習データのノイズや偶然の合致等の可能性も考えられるが、敢えて仮説を考えてみよう。食堂に対する満足度が低い場合には「そもそも食堂を利用していない」場合が含まれているかもしれない。たとえば「弁当等を持参する人」「学外の飲食店を利用する人」等である。食堂の満足度が低いにも関わらず「○4年間満足している」の回答につながっている人は、食事を自前で充足していた可能性が考えられる。

逆に、食堂に対する満足度が高いが「×4年間満足していない」という回答の人は、食堂以外に不満足の原因があると考えられる。食堂以外の施設や学生生活等について、各人が別々の不満を抱えていたため、決定木には単一の分岐が現れなかった可能性が考えられる。

機械学習では、表層のみを分析するため、こういった不可解な結果が現れることもある。

3.2 在学生に対する意識調査アンケートの分析 ～後輩に学部を勧める人の傾向～

次に紹介するのは、環境社会学部在学生に対する意識調査アンケートの分析結果である。本アンケートの実施概要を以下に示す。

調査日 : 2016年3月上旬
 対象 : 環境社会学部在学生 回答者数 : 94名
 ※質問項目は「付録B」参照

この分析では、アンケート項目のうち「あなたは、高校などの後輩から進路相談を受けたら、環境社会学部を勧めますか」に対して、「○後輩に勧める」と「×後輩に勧めない」とに回答した2つのグループに分け、他の質問項目から何れのグループに属するのかを機械学習で推定させた。図5に学習結果の決定木を示す。

1) 「自分が満足していること」が前提条件

まず最上段の分岐は「学部を選んで良かったか」の回答で分岐している。「学部を選んで良かったか」に対して「④それほど思わない」「⑤全く思わない」を選んだ人は「×後輩に勧めない」になっている。

次の分岐は「国籍」である。本アンケート結果の特有傾向かもしれないが、中国からの留学生は「○後輩に勧める」を選んでいる。以下、国籍が「日本」(左側の枝)の場合について詳しくみていこう。

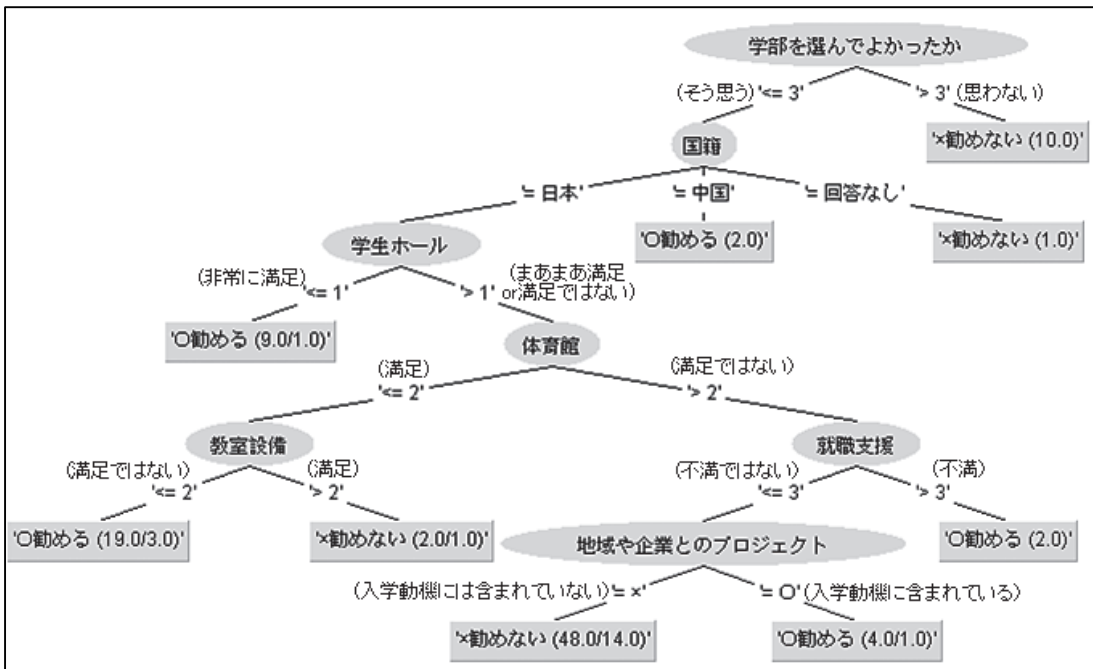


図5 高校の後輩に学部を勧める要因 (機械学習の結果)

2) 「キャンパス環境」に満足しているので後輩に勧める場合

「学生ホール」に「①非常に満足」の場合、それ以外でも「体育館」および「教室設備」に「①非常に満足」or「②まあまあ満足」の場合に「○後輩に勧める」を選ぶ傾向がある。自分は満足していても、後輩が満足できるかは解らない。キャンパス環境は、客観的な指標として他人も満足するだろうと予測できるためか、後輩に勧める要因となっていると考えられる。特に、休み時間や放課後の憩いの場となる「学生ホール」は「①非常に満足」が求められており、要因として重視されていることが伺える。

3) 「就職支援および企業との連携」に満足しているので後輩に勧める場合

「学生ホール」「体育館」等のキャンパス環境に満足していない場合でも、「就職支援に満足(3以下)」かつ「地域や企業と連携したプロジェクトに期待している(入学動機)」の場合には「○後輩に勧める」が選ばれている。

「地域や企業と連携したプロジェクト」は、大学で学んだ知識を活かして、地域や企業で環境活動を実践する教育プログラムである。学生のうちに社会での実践経験を積むことが期待できる。就職支援とあわせて、卒業後を見据えた教育こそが、学生達が大学に期待する役割であり、後輩に勧める上でも重要な要因であることが予想される。

4) 「キャンパス設備」や「就職支援」に満足していないのに後輩に勧める場合

図5の決定木右下には「学生ホール」「体育館」「就職支援」のすべてに満足していないが「○後輩に勧める」と回答している分岐がある。この学生は、如何なる理由で「○後輩に勧める」を選んだのであろうか。

決定木は、条件分岐(この場合は「○後輩に勧める」と「×後輩に勧めない」の分岐)で枝を増やすため、分岐せず一意に定まる場合には省略されてしまうことがある。元データを確認すると、この分岐に相当する学生は共通して「実習系の科目」に満足していると回答していることが確認できた。しかし他の条件によって、すでに「○後輩に勧める」と「×勧めない」に完全分離できたため、「実習系の科目」の分岐は省略されてしまったと考えられる。

人間が決定木を見ると、通常と逆の相関であるため不思議に感じる分岐であるが、機械学習では決定木の省略(剪定等)の結果、こういった学習結果が生成されることも起こりうるのである。

3.3 オープンキャンパス来場者アンケートの分析 ～実際に出願した高校生の特徴～

最後に紹介するのは、オープンキャンパス来場者(環境社会学部の体験授業イベントに参加した高校生対象)のアンケート結果分析である。本アンケートの実施概要を以下に示す。

調査日 : 2017年 3月～12月

対象 : オープンキャンパス来場者(高校生) 回答者数 : 78名

※質問項目は「付録C」参照

この分析では、入試出願リストと照合して、実際の入試に「○出願した」と「×出願しなかった」との2つのグループにオープンキャンパス来場者を分け、アンケート回答項目から何れのグループに属するのかを機械学習で推定させた。図6に学習結果の決定木を示す。

1) 「環境について興味関心がある分野」に対する回答と出願の関係

まず最上段の節点は「学年」の分岐であり、2年生以下の場合、必然的に「×出願しなかった」となっている。

次の節点では「環境のどのような分野に関心がありますか（複数可）」の質問に対して、「再生エネルギー」「農業と安全な食」「リサイクル」を選択したか否かで分岐している。まず左側の分岐を見ると「再生エネルギー」および「農業と安全な食」の両方に興味を持つと回答した者は「○出願した」につながっている。また右側の分岐を見ると「再生エネルギー」に興味が無くとも、「リサイクル」に興味があれば「○出願した」につながっている。

おそらく「再生エネルギー」等の「工学系環境技術」に興味を持つ人は本学部を敬遠しており、「農業と安全な食」「リサイクル」等の生活に関連した「環境活動の推進」に興味を持つ人が好んでいるのではないかと考えられる。

2) 「参加した時期」と出願の関係

興味がある領域で「再生エネルギー」にも「リサイクル」にも興味がないと回答した人でも、最終的に「○出願した」につながるパスが存在する。たとえばオープンキャンパスに「10月に参加」した場合には「○出願した」になっている。また「9月に参加した」場合でも条件次第（後述：授業内容が○のとき）で「○出願した」につながっている。

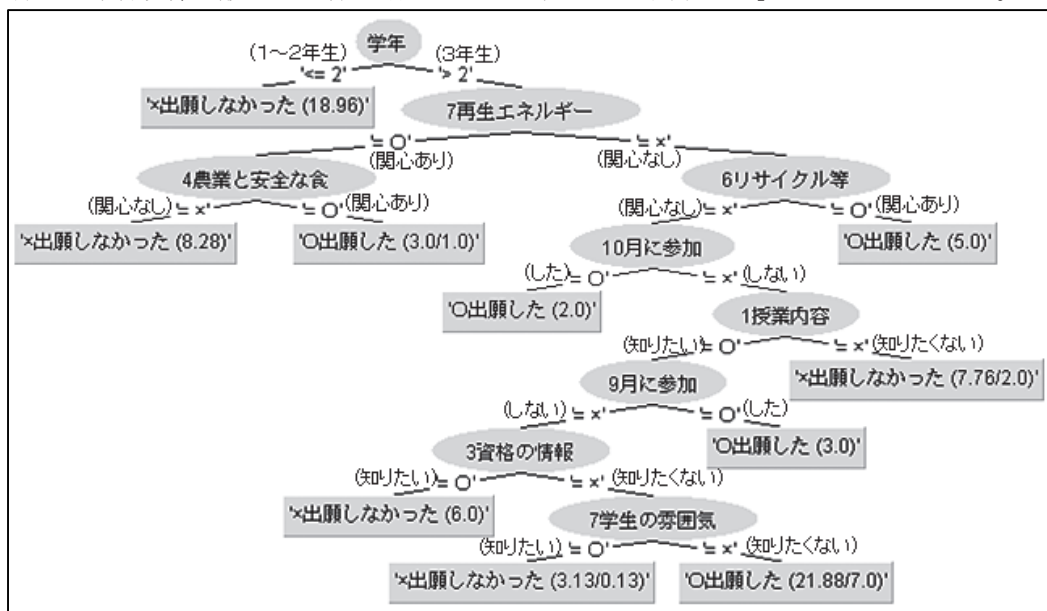


図6 オープンキャンパス来場者アンケートにおける入試出願の要因（機械学習の結果）

9月や10月は推薦入試が始まる時期であるため、この時期にオープンキャンパスに参加する者は、出願を決めた上で参加する傾向が高いのではないかと考えられる。ただし、この分岐は1)で述べた「興味ある領域」の下位に存在する。すなわち「農業や安全な食」および「リサイクル」に興味がある場合には、9～10月以外（8月以前）の参加が見られ、興味ある領域がそれ以外（たとえば「地球温暖化」「自然保護」等）の場合にのみ、9～10月に参加して出願する傾向が見られるということになる。おそらく「興味ある領域」によっては競合校が存在し、その結果、出願先の選定に時間を要するため、オープンキャンパス参加が遅い時期になっているのではないかと考えられる。

3)「参加した目的」に対する回答と出願の関係

その他の分岐は「オープンキャンパスに参加した目的（知りたいこと）」の回答結果と関連している。10月参加者以外では「学部の教育（授業）内容について知りたい」を選ばなかった人は「×出願しなかった」になっている。また9～10月参加者以外で「資格の取得について知りたい」を選ぶと「×出願しなかった」になっている。最後に「学生の雰囲気を知りたい」を参加目的に選んだ人も「×出願しなかった」になっている。

まず「教育内容について知りたい」を選ばなかった人を考えると、初めから本学部への進学を考えていなかった可能性がある。たとえば「体験授業イベントを楽しむためだけ」「友人の付き添い」等が目的だったと考えれば、出願しなかった結果も頷ける。

次に「資格の取得について知りたい」を選ぶと、何故か「×出願しなかった」結果になっているのは、もともと資格に興味があったが、説明を聞いて関心が高まらなかった可能性が考えられる。もっと説明を充実させる必要があったのかもしれない。

最後に「学生の雰囲気を知りたい」を参加目的に選んだ場合も「×出願しなかった」になっている。しかし逆に「学生の雰囲気を知りたい」を選ばなかった人は「○出願した」になっている。両者をあわせて考えると「学生の雰囲気を知りたい」と回答した人は、他の目的（入試方式、就職先、キャンパス施設について知りたい等）を回答した人より、出願傾向が薄くなっている可能性が考えられる。雰囲気重視で進学先を選ぶ人からは、実際の出願があまり期待できないのかもしれない。

勿論、これらは本アンケート結果特有の傾向であり、一般的傾向ではない可能性があるのは言うまでもない。今回は敢えて、結果を説明するための理由付けや仮説を検討してみたが、その中で、今まで着目していなかった観点からも試行錯誤することとなった。これも機械学習の面白さと言えるのではないだろうか。

5. まとめ

本研究では、機械学習を用いて3つのアンケート結果の分析を行った。果たして、機械学習の結果は有益であったと言えるだろうか。

機械学習が人間に匹敵するかという視点で論じれば、未だ機械学習は人間の分析を代替できるとは言えず「有益ではない」と言えるかもしれない。分析精度が甘いということもあるが、最終的に結果を活用するのが人間である以上、人間の思考プロセスに沿って分析できる方が有利であるだろう。機械学習の分析はあくまで機械的に行われるため、分析結果の意図が掴みづらいのが難点である。

対して、分析速度という視点で論じれば、機械学習の処理は圧倒的に速く「有益である」と言えるかもしれない。速い処理は、時間短縮だけでなく、繰り返しや多面的な分析を可能とする効果が期待できる。幅広い視点から多くの分析結果が得られるのが特徴である。

結論としては、月並みだが、機械学習によるアンケート分析は使い次第と言えるだろう。今回の実験においても、想定外の多様なデータ相関が機械学習により発見できた半面、分析結果を読み解くために大変に頭を悩ませられた。そのため、一手間かかったとしても、新しい視点や切り口を見つけたいときには、有効な手段であると考えられる。

文中でも述べたように、機械学習の結果は当該データ特有の傾向、すなわち一般性がなく、偶々であるかもしれない。機械学習の分析結果を読み解いても、必ず正しい知見が得られるとは限らない。しかし読み解くための考察の中から新たなアプローチを閃くことには期待できる。特に、思考に煮詰まった時には有用であるだろう。AIをパートナーとして、ブレインストーミングを行うイメージに近いかもしれない。

本研究では、機械学習でアンケート結果の分析を行ったが、想定外の結果も提示される中、その解釈を考えることで、興味深い知見が得られたと思う。これを機に、皆様にも機械学習に興味を持っていただき、様々な分野で活用いただければ幸いである。

【引用文献】

- [1] Tom M. Mitchell, “Machine Learning”, McGraw Hill, ISBN 0-07-042807-7, 1997.
- [2] T Menzies, Y Hu, “Data mining for very busy people”, IEEE Computer, October 2003.
- [3] “Google DeepMind 社のデミス・ハサビス氏のツイート（謎の棋士 master の正体を発表）”,
URL: twitter.com/demishassabis/status/816660463282954240/photo/1, 2017.
- [4] Arthur Samuel, “Some Studies in Machine Learning Using the Game of Checkers”, IBM Journal of Research and Development (Vol 44, Issue 1.2), January 2000.
- [5] J. R. Quinlan, “Induction of decision trees”, Machine learning (Vol 1, Issue 1), 1986.
- [6] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, ” Deep learning”, Nature (Vol 521), May 2015.

- [7] “WEKA: The Waikato Environment for Knowledge Analysis”, University of Waikato, Hamilton.
 URL: www.cs.waikato.ac.nz/ml/weka/, 1993-.
- [8] J. R. Quinlan, “C4.5: Programs for Machine Learning”, Morgan Kaufmann Publishers, 1993.

付録A 「卒業予定者対象 意識調査アンケート 質問項目（抜粋）」

質問項目	回答選択肢（マークシート方式）
性別	①男性 ②女性
国籍	①日本 ②中国 ③韓国 ④その他
環境社会学部で 4 年間学んで良かったと思いますか。	①かなりそう思う ②そう思う ③どちらでもない ④あまりそう思わない ⑤全くそう思わない
環境社会学部の学習環境に満足していますか。以下の各項目について、それぞれ該当する番号①～⑤を1つ選んでマークしてください。 ア. 教室の設備・施設 イ. 授業科目 ウ. 外構の緑地・オープンスペース エ. スポーツアリーナ オ. グラウンド カ. 図書館 キ. 学生ホール ク. 実習施設 ケ. 学生食堂 コ. 売店 サ. 書店 シ. スクールバス	①非常に満足している ②まあまあ満足している ③どちらでもない ④それほど満足していない ⑤全く満足していない
JIU の学生生活は満足していますか。以下の各項目について、それぞれ該当する番号①～⑤を1つ選んでマークしてください。 ア. 友人関係 イ. クラブ・サークル活動 ウ. ボランティア活動 エ. 自分のしたい勉強 オ. 奨学金制度 カ. オフィスアワー キ. 就職支援 ク. 学生相談室（カウンセラー制度）	①非常に満足している ②まあまあ満足している ③どちらでもない ④それほど満足していない ⑤全く満足していない

付録B 「在学生対象 意識調査アンケート 質問項目（抜粋）」

質問項目	回答選択肢（マークシート方式）
学年	①1年生 ②2年生 ③3年生 ④4年生
性別	①男性 ②女性
国籍	①日本 ②中国 ③韓国 ④その他
あなたは、環境社会学部を選んで良かったと思いますか。	①大変そう思う ②そう思う ③どちらでもない ④それほど思わない ⑤全く思わない
あなたが、環境社会学部を進学先として選択した理由をお聞かせ下さい。（複数選択可）	①文系・理系の枠組みを超えて環境のことが幅広く学べる ②実習系の科目と施設が充実している ③キャリア教育等の就職支援がしっかりしている ④国際大学として留学や海外研修が充実 ⑤地域や企業と連携したプロジェクト型教育に重点を置いている
あなたは高校などの後輩から進路相談を受けたら環境社会学部を勧めますか。	①はい ②いいえ
環境社会学部の学習環境に満足していますか。以下の各項目について、それぞれ該当する番号①～⑤を1つ選んでマークしてください。 ア. 教室の設備・施設 ： ※卒業予定者対象アンケートと同項目 シ. スクールバス	①非常に満足している ②まあまあ満足している ③どちらでもない ④それほど満足していない ⑤全く満足していない
総合的にみて、JIUの学生生活は満足していますか。以下の各項目について、それぞれ該当する番号①～⑤を1つ選んでマークしてください。 ア. 友人関係 ： ※卒業予定者対象アンケートと同項目 ク. 学生相談室（カウンセラー制度）	①非常に満足している ②まあまあ満足している ③どちらでもない ④それほど満足していない ⑤全く満足していない

付録C 「オープンキャンパス来場者アンケート 質問項目（抜粋）」

質問項目	回答選択肢（マークシート方式）
オープンキャンパスに参加した日付 (集計時に自動入力)	①1月 ②2月 ③3月 ④4月 ⑤5月 ⑥6月 ⑦7月 ⑧8月 ⑨9月 ⑩10月
学年	①高校1年 ②高校2年 ③高校3年
性別	①男性 ②女性
本日のオープンキャンパスに参加された目的(知りたいこと)をお聞かせ願います(複数可)	①学部の教育(授業)内容 ②入試方式 ③資格の取得 ④卒業後の就職先 ⑤キャンパスの雰囲気 ⑥教員の雰囲気 ⑦学生の雰囲気
あなたは、「環境」のどのような分野に関心がありますか(複数可)	①温暖化等地球環境問題 ②自然環境の保全 ③大気・水等生活環境の保全 ④農業・安全な食問題 ⑤園芸と健康 ⑥リサイクル等資源の循環 ⑦再生エネルギー
環境社会学部で奨励している以下の資格に興味のある資格はありますか(複数可)。	①eco(エコ)検定 ②生活園芸士 ③園芸療法士 ④ビオトープ管理士 ⑤アロマコーディネーター
学部のオープンキャンパスの印象はいかがでしたか(複数可)。	①学部としてよくまとまっていると思う ②先生方の対応が丁寧である ③学生たちの対応が親切で良い ④学部の説明が長すぎる ⑤説明より体験型を中心にしてほしい ⑥特になし
【後日に学部側で追加したデータ】 オープンキャンパス参加後、出願したか	①出願した ②出願しなかった

Awareness Analysis of Students interested in Environmental Studies using Machine Learning

- How useful is AI? -

Masatoshi Ietomi

Abstract

In this paper, we analyze the results of questionnaire survey using Machine Learning in Artificial Intelligence. Machine learning can analyze data trends that people do not notice. However sometimes we may get different results from reality. Analyze the results of several questionnaire surveys and verify that Machine Learning is useful.

Keywords: Machine Learning, Awareness Survey, Student Questionnaire, Trend Analysis,
Environmental Studies