

小特集 「AIトレンド・トップカンファレンス NIPS 2017」 報告会

# 機械学習による意思決定

## Machine Learning for Decision Making

小宮山 純平

東京大学生産技術研究所

Junpei Komiya

Institute of Industrial Science, The University of Tokyo.

junpei@komiya.info, [http://www.tkl.iis.u-tokyo.ac.jp/~jkomiya/index\\_jp.html](http://www.tkl.iis.u-tokyo.ac.jp/~jkomiya/index_jp.html)**Keywords:** machine learning, decision making, sequential learning, fairness.

### 1. 機械学習による意思決定

今日の人工知能は、「人工知能」という名前が想起させるような知的な処理のすべてを代替する、という壮大な夢のすべてを実現することはできていないかもしれない。しかし、大量のデータを正確に処理するという点においては、人間の能力をはるかに超えた処理能力を発揮することができる。人類の知性の象徴とされているチェスでは、1997年に人間のチャンピオンとの対局で人工知能が勝ち越した [Pandolfini 97]。今は多くの人が利用している検索エンジンは数百億~数兆の Web ページをクロールしてクエリに関連するページを表示することができるようになった。機械学習の最先端会議である NIPS は、機械学習にできることとできないことの境界線を探る場所として考えることができるだろう。特に、NIPS の発表で、機械学習における意思決定に筆者は注目した。意思決定 (decision making) は、複数の代替案から最善の解を求めることを指す。機械学習や、より広い意味での人工知能のゴールの一つは、人間の意思決定を支援したり、あるいは代替することである。機械学習のもつ意思決定能力の最も重要な論点として、以下の二つを考えた。

- 強化学習・オンライン学習：強化学習やオンライン学習は、逐次的な (繰返し) 意思決定を行うシステムをモデル化することができる。これらの学習の最も成功した例としては、先ほど例にあげたチェスのようなアブストラクトゲームがあげられる。アブストラクトゲームは、局面に応じた着手の選択と、その善悪の判断をどうするかという問題がある程度わかりやすい形で判断可能な意思決定問題だと考えることができる。ここ数年のアブストラクトゲームで最も華々しい例は、囲碁とポーカーである。囲碁は英 Deepmind 社によるプログラムが 2016 年に人間のトッププレーヤーに勝ち越した結果 [Silver 16] が話題となった。また、同年に Heads-up Hold'em (二

人ポーカー) で初めて人間トップレベルに達した [Brown 17a]。後者に関連して、NIPS 2017 における best paper のうち 1 本がこのポーカープログラムの探索アルゴリズムに関するものであった。2 章ではこの NIPS best paper とポーカープログラムの技術について解説する。

- 施策・政策の公平性：私企業の施策、あるいは公的機関の政策に機械学習をもち込むというアイデアは、特にデータが十分に集まるドメインにおいては十分現実的と考えられるようになった。特に、雇用、昇進、入試、ローンなどの与信情報などの関わる判断は極めて重要度が高いが、このような人間に関する意思決定は一定の公平性を担保する必要がある。国内でも、新卒採用の説明会で特定の Web システムが男女別で座席数を管理しているのではないかとという問題が話題となったが\*1、特定のグループ (例えば、性別、人種、年齢など) に対して、他のグループと比較して不当に不利な扱いをするのは、法律的にも倫理的にも問題となり得る。特に、過去のデータからの学習を行う機械学習では、過去にあった差別的なデータを学習してしまう可能性がある。このような問題に配慮した公平性配慮型の機械学習について、3 章にて解説する。

### 2. ポーカー AI における意思決定

チェス、将棋・囲碁といったゲームは着手によって後の局面が変化していく展開型ゲームであり、深く正確な読みが要求される。これらのゲームは実力がはっきりわかりやすく、ある種の人間の知性の象徴とされてきた。ポーカーはこれらのゲームとは一線を画した、ともすればギャンブル的なゲームであるというイメージはないだろうか。ポーカーにはいくつかのルールがあるが、最もメジャーなものは、手札の 2 枚と共有札の 5 枚を合わせ

\*1 <https://togetter.com/li/1204933>



図1 テキサスホールデムでは、2枚の手札と5枚の共有札を合わせて良い役をつくる。共有札は3枚、1枚、1枚と順番に公開されるが、図中では3+1枚が公開された状態である。図は <http://science.sciencemag.org/content/347/6218/122.full> より引用

て役の強さを競うテキサスホールデムである(図1)。共有札は最初は未公開だが、チップを賭け合い場面が進行するとだんだん共有札が公開されていき、最後までチップを賭け続けたプレーヤどうして強い役の持ち主がチップを取る。ホールデムでは相手の手札が非公開情報であり、この手札を推定し、相手が強いときに少ないチップを払うことによっており、相手が弱いときに大きくチップを得るのが戦略の基本となる。習熟度が低いプレーヤの間では、相手の顔やしぐさに対する自信の表れを見て相手の手札の強弱を推定するブラフゲーム的な要素が強いかもしれない。しかし、トッププレーヤの間で行われるテキサスホールデムは、ブラフ性がしだいに少なくなっていく<sup>\*2</sup>、正確な確率計算とゲーム理論のナッシュ均衡戦略が基本となる。すなわち、初手が完全に一様な分布だとすると、相手のチップを賭ける戦略が既知とした場合、各局面での相手の手札の確率分布を得ることができる。二人ゼロ和ゲームでは相手の最も強い行動に対して最善の戦略を考えるミニマックス戦略がゲーム理論的均衡になるため、このような戦略を考えるのが二人ポーカーの「神の一手」であるといえる。そして、二人ポーカーの神の一手を考えるために必要な局面の数は、10の161乗であり、チェスの局面数よりはるかに多い(!)。このため、ナッシュ均衡戦略を正確に得ることを目的にすると、二人ポーカーはチェスよりはるかに難しいゲームであるといえる。局面数だけに依存しないさらなる難しさとしては、チェスのように今の局面での情報がすべて与えられている完全情報ゲームと異なり、ポーカーでは今の情報セット(局面のうち、プレーヤに見える部分=自分の手札、場のチップの集まった手順、共有札)において相手の手札が単一に定まらず戦略に応じた確率分布を保持しなければならない。そして、相手のすべての戦略を考えなければならないのは、ひたすら広大な地平を旅するようすら見える……この問題

\*2 つまり、顔などから情報を得ることは高いレベルのプレーヤどうしては難しくなる。顔がないポーカープログラムはこの極致かもしれない。

に対する優秀なアルゴリズムとして、反実後悔最小化(counterfactual regret minimization)と呼ばれる逐次アルゴリズムが投入できる。反実後悔最小化では、最初に適当な戦略を用意し、この戦略を1ステップごとに局所改善したときに利得が増えるかどうかで自分の戦略の改善を行う。反実後悔最小化における戦略の改善は二人プレーヤでは、無限の回数を経ればナッシュ均衡戦略に収束する。しかし、反実後悔最小化の収束までに要する時間は、ゲーム木のサイズに依存する。そのため、ゲーム木をある種粗視化する「青写真戦略」(blueprint)を用意し、また深くなるにつれ指数的に難しくなるゲーム後半(4枚目、5枚目カードの公開と関連したチップのベット)では、部分ゲーム解決(subgame solving)という青写真の精緻化をリアルタイムで行うのが常道であった。部分ゲーム解決の手法はここ2~3年でいくつか提案されていたが、結果を悪化させてしまい安全でない(unsafe)か、もしくは相手の戦略に過学習気味なものであった。[Brown 17b]のNIPSベストペーパーは、悪化がないと保証される安全な(safe)部分ゲーム解決を行い、ゲーム後半での強さを保証するものであり、同研究チームがBrains vs. AI challengeにおいてポーカーのトッププロに大きく勝ち越す[Brown 17a]技術の基盤となるものであった。おおまかにいうと、NIPSベストペーパーの部分ゲーム解決は、ある戦略に対して、いつでも離脱するオプションがあると考え、その戦略にあえて入ったことに対する利得を加算し、青写真における後半の戦略を精密化する。

最後に、彼らポーカーの研究者らは、これらの人工知能技術の応用としてセキュリティの問題をあげた。例えば、空港の警備などで攻撃者(テロリスト)の取り得る攻撃経路が複数あり、ある程度しか情報がわからない場合は、それに対して最も頑健な防御者(警備員)の戦略はナッシュ均衡となるであろう。もう一つ、はっきり困難な問題としては、三人以上のポーカー<sup>\*3</sup>ではナッシュ均衡はもはや単一ではなく、反実後悔最小化で得られる戦略が実際のトッププロの戦略にどの程度合致するかも曖昧である。三人以上のポーカーにおける研究は例えば[Gibson 13]などがあるが、未知の広大な領域が広がっているといえる。ある意味では、攻撃者と防御者のセキュリティゲームより、多国間の利益解決ゲームのほうが難しいのかもしれない。

### 3. 機械学習における意思決定の公平性

人間に関わる意思決定(入試、雇用、昇進、ローンの与信など)を行う際、意思決定者にはある種の公平性が求められる。例えば、ローンの与信を特定の団体に対し

\*3 通常、テキサスホールデムは最大十人程度の多くのプレーヤで行われ、二人ゲームのほうが特殊ケースである。

て非常に甘く査定していたら、同じローンを負担している人は不利益を被るであろう。人間が意思決定を行うのと同程度か、あるいはそれ以上に機械学習による意思決定には公平性が気になる。機械学習における公平性は、個人レベルの公平性とグループレベルの公平性があるが、本稿では簡単のため後者について解説する。

機械学習における公平性というテーマは、ここ数年でその注目を大きく集めるようになった。機械学習・データマイニング学会のワークショップとして、Fairness, Accountability, and Transparency in Machine Learning (FATML)\*<sup>4</sup>があるが、このワークショップでの論文発表数は2014年：0件（招待講演のみ）、2015年：4件、2016年：13件、2017年：29件と指数的に増加している。また、2018年には公平性を扱う会議であるFAT\*\*<sup>5</sup>がFATMLと独立した学会として開かれた。

NIPS 2017では、機械学習の公平性に関する招待講演・チュートリアルがあった。BarocasとHardtによるチュートリアル“Fairness in Machine Learning”<sup>6</sup>は、この分野のある種の概観となっている。機械学習における公平性は、アメリカにおける人種差別の問題に大きく影響を受けている。1955年、デューク・エナジー社（電力会社）が社員の昇進に高校学位もしくは特定の資格保持を必須要件にしたが、これらの要件は人種のプロキシであった。つまり、白人社員と黒人社員の間に高校学位：34%（白人）・12%（黒人）、資格保持者：58%（白人）・6%（黒人）と大きな開きがあった。この取扱いはアメリカにおいて裁判になったが、1971年に結審し、最終的に「職業でのパフォーマンスに明確な関係のない資格や学位」は実際の差別意図があるかないかにかかわらず無効とされた。つまり、明示的に差別的な扱い（adverse treatment）だけでなく、間接的に他の変数が与える差別的な扱い（adverse impact）も不当であるとされたのである。機械学習における公平性は、これらの要件を数式に落としたものとなっている。つまり、グループへの帰属を表す公平性に配慮したセンシティブ属性を $s$ （例えば、人種や性別）と置いた場合に、昇進、雇用などの結果 $y$ を $s$ に依存しないようにしなければいけない。つまり、 $y$ と $s$ の独立性

$$y \perp\!\!\!\perp s$$

が要求される。これを、民主的公平性（demographic parity）と呼び、機械学習における公平性の代表的な要件としている。上記の例では、白人社員と黒人社員を同じ割合で昇進させなければならない<sup>7</sup>。一方、 $s$ 以外

の公平性に配慮を要しない属性（上記の例では学位や資格）を $x$ と書くと、仮に $s$ を用いなくても $x$ を用いて $y$ を学習した場合、 $y$ の予測は $x$ に依存するようになってしまう（adverse impact）。そのため、 $x$ と $s$ の間の相関や依存性を排除するような学習を行う。統計的機械学習における学習器はマジョリティを重視する「バンドワゴン効果」があるため、グループの比率が著しく異なる場合は注意が必要である。機械学習における代表的なタスク分類（例えば、[Feldman 15, Fukuchi 15, Kamiran 10]）、回帰[Fukuchi 15, Komiyama 18]、推薦[Kamishima 12]などにおいて、民主的公平性を保つことができる学習アルゴリズムが提案されている。一方で、あるグループの社員と別のグループの社員の間の職業でのパフォーマンスの違いである変数 $p$ を測定できる場合は、均等オッズ（equalized odds）条件

$$y \perp\!\!\!\perp s | p$$

を満たすことが公平性とされる[Hardt 16]。これは、パフォーマンスの高さによって社員を分け、それぞれのクラスの社員をセンシティブ属性によらず同等に扱うことに対応する。

Crawfordによる招待講演“the trouble with bias”は、機械学習におけるバイアスを紹介し、問題提起したものであった。具体的な問題としては、

- 顔認識の精度は人種ごとに差があること
- Google検索で“CEO”のクエリを入れると、スーツを着た男性ばかりがでてくること。また、職業を表す単語が暗黙のうちに性別を含意していること
- アフリカ人女性のSweeneyさんが、Googleの広告により事実と異なり犯罪者扱いされたこと（図2）

などの問題を紹介した。Crawfordはバイアスを「事実と異なる既成観念や偏見」と定義した。機械学習が引き起こすバイアスの代表的な問題を定義した後で、機械学習における意思決定がどのような利益をあげられるのか、

Ad related to latanya sweeney ⓘ

**Latanya Sweeney Truth**  
www.instantcheckmate.com/  
Looking for Latanya Sweeney? Check Latanya Sweeney's Arrests.

Ads by Google

**Latanya Sweeney, Arrested?**  
1) Enter Name and State. 2) Access Full Background Checks Instantly.  
www.instantcheckmate.com/

**Latanya Sweeney**  
Public Records Found For: Latanya Sweeney. View Now.  
www.publicrecords.com/

**La Tanya**  
Search for La Tanya Look Up Fast Results now!  
www.ask.com/La+Tanya

図2 Googleで“Latanya Sweeney”と検索したとき、Sweeneyさん自身の逮捕歴がないにもかかわらず、“Arrested?”（逮捕）と出る広告があった。図は[Sweeney 13]より引用

\*4 <https://www.fatml.org/>

\*5 <https://www.fatconference.org/> 学会の動画が <https://www.youtube.com/channel/UCs16j6ot-CYq-ZqYpO-vqMg/videos> で閲覧可能である。

\*6 スライドがNIPS公式サイトからのリンクで入手可能である。

\*7 職業でのパフォーマンスに明確な違いがない場合。

それによって起こり得るバイアスは誰が被害を受けるのかを考えなければならないことを伝えた。また、機械学習における公平性は法学的な (forensics) 側面, 倫理的な側面も重要であり, また学際 (interdisciplinary) 領域の研究を考えなければならないという内容であった。

### ◇ 参 考 文 献 ◇

- [Brown 17a] Brown, N. and Sandholm, T.: Superhuman AI for heads-up no-limit poker: Libratus beats top professionals, *Science*, Vol. 359, No. 6374, pp. 418-424 (2018)
- [Brown 17b] Brown, N. and Sandholm, T.: Safe and nested subgame solving for imperfect-information games, *NIPS 2017*, pp. 689-699 (2017)
- [Feldman 15] Feldman, M. and Friedler, S. A., Moeller, J., Scheidegger, C. and Venkatasubramanian, S.: Certifying and removing disparate impact, *ICDM 2015*, pp. 259-268 (2015)
- [Fukuchi 15] Fukuchi, K., Kamishima, T. and Sakuma, J.: Prediction with model-based neutrality, *IEICE Trans.*, Vol. 98-D, No. 8, pp. 1503-1516 (2015)
- [Gibson 13] Gibson, R.: Regret minimization in non-zero-sum games with applications to building champion multiplayer computer poker agents, arXiv 1305.0034 (2013)
- [Hardt 16] Hardt, M., Price, E. and Srebro, N.: Equality of opportunity in supervised learning, *NIPS 2016*, pp. 3315-3323 (2016)

- [Kamiran 10] Kamiran, F., Calders, T. and Pechenizkiy, M.: Discrimination aware decision tree learning, *ICDM 2010*, pp. 869-874 (2010)
- [Kamishima 12] Kamishima, T., Akaho, S., Asoh, H. and Sakuma, J.: Enhancement of the neutrality in recommendation, *Decisions@RecSys* (2012)
- [Komiya 18] Komiyama, J., Takeda, A., Honda, J. and Shima, H.: Nonconvex optimization for regression with fairness constraints, *ICML 2018* (2018)
- [Pandolfini 97] Pandolfini, B.: Kasparov and Deep Blue: *The Historic Chess Match Between Man and Machine*, Fireside Chess Library, ISBN 9780684848525 (1997)
- [Silver 16] Silver, D., et al.: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol. 529, p. 484-489 (2016)
- [Sweeney 13] Sweeney, L.: Discrimination in online ad delivery, *Commun. ACM*, Vol. 56, No. 5, pp. 44-54 (2013)
- [Zinkevich 07] Zinkevich, M., et al.: Regret minimization in games with incomplete information, *NIPS 2007*, pp. 1729-1736 (2007)

2018年6月15日 受理

### —— 著 者 紹 介 ——

小宮山 純平は, 前掲 (Vol. 33, No. 5, p. 631) 参照.