

## 営業活動の意思決定プロセス強化における 環境モデルに基づくアプローチ

中山義人<sup>†1 †2</sup> 森雅広<sup>†2</sup> 斎藤忍<sup>†3</sup> 成末義哲<sup>†1</sup> 森川博之<sup>†1</sup>

**概要:** 営業活動における意思決定から、営業担当者個人の経験や直感といった属人的要素を取り除くことにより、営業活動を大幅に効率化するための手段が求められている。筆者らはこの課題に対し、機械学習モデルを用いた業務意思決定支援システムの構築を試みている。これまで、営業活動の意思決定プロセスに強化学習を適用することで、受注確率の高い営業プロセスの規則性を抽出することができた。その際、営業をエージェント、顧客を環境と位置づけ、その間のやり取りをシミュレータで自動実行する案件シミュレータを開発することで、学習に必要な十分な案件データ量を確保することができた。しかしこの案件シミュレータのモデリングでは、シミュレータのパラメータ設定範囲が固定化されており、さらには営業エージェントが環境に依存して自由度が制限されるという課題がある。そこで本稿では、案件の背景、要件、顧客のパーソナリティといった環境自体を、シミュレータで作成された案件データを利用して深層学習し、最適な価値関数や方策をプランニングするための環境モデルを構築する。さらにこの構築された環境モデルを利用して、営業エージェントとの間のシミュレーションをセルフプレイで行うことで営業エージェントの学習モデルを強化する。これにより、営業活動場面でより現実に近い理想的なリコメンドが可能となることが想定される。

**キーワード:** 営業活動, 意思決定支援, 環境モデル, プロセスマイニング, 人間参加型深層学習, プロセス強化

### An approach based on environment model for strengthening the decision-making process of sales activity

YOSHIHITO NAKAYAMA<sup>†1 †2</sup> MASAHIRO MORI<sup>†2</sup>  
SHINOBU SAITO<sup>†3</sup> YOSHIAKI NARUSE<sup>†1</sup> HIROYUKI MORIKAWA<sup>†1</sup>

**Abstract:** In the decision-making process of sales activities, the result depends greatly on the judgment of sales personnel. Therefore, the means for establishing the efficiency of sales activities throughout the organization are required by eliminating individuals' factors such as experience and intuition from the decision-making process. In order to solve this problem, we are developing the business decision support system using a machine learning model. Previous studies of applying reinforcement learning to the decision making process of sales activities have produced the extraction of regularity in sales processes with high order acceptance probability. In these studies, it was possible to secure sufficient amount of data to be required for learning by developing a simulator that positions sales as agents and customers as environment and automatically executing exchange between them. However, in modeling of the simulator, there are issues that the parameter setting range of the simulator is fixed and furthermore, the agent depends on the environment. In this paper, we construct an environmental model for planning optimal value functions and policies by deep learning of customer's environment such as background, requirements and customer's personalities. Furthermore, learning model is strengthened by performing self-play simulation with agent using this constructed environmental model. As a result, it is assumed that an ideal recommendation that is closer to reality is possible in the sales activity.

**Keywords:** Sales activity, Decision making support, Process mining, Human-in-the loop, Process enhancement

## 1. はじめに

営業活動における意思決定プロセスには予め明確なルールが存在せず、営業担当者の経験や判断にもとづいて実施されるため、その成果も自ずと属人的なものとなる。

筆者らは、機械学習モデルを用いることによって、属人

性を意思決定プロセスから除外し、受注確率の高い営業活動を組織全体で実施することが可能となるような業務意思決定支援システムの構築を目指している。

これまで、営業日誌などの非構造化データを入力情報として、予め明確なルールが存在しない非定型プロセスから、受注確率の高い営業プロセスの規則性を発見する学習モデ

†1 東京大学大学院工学系研究科  
School of engineering, The University of Tokyo  
†2 (株) NTT データイントラマート  
NTT DATA INTRAMART Corporation

†3 日本電信電話 (株)  
Nippon Telegraph and Telephone Corporation

ルの構築を実現することができた[1].

従来のプロセスマイニング手法[2][3]に前処理として、営業日報がどのアクティビティに属するのかを判別する「(1)アクティビティ推定」と、それに続いてそこから受注確率の高い営業プロセスを強化学習で推定する「(2)プロセス推定」の2つのステップを追加する応用アプローチとした(図1).

特に「(2)プロセス推定」では、学習アルゴリズムとして大量の教師データを必要とする隠れマルコフモデルを採用するために、案件シミュレータを開発した(図2).

この案件シミュレータにより、営業をエージェント、顧客のパーソナリティを環境と位置付け、その間のやり取りをシミュレータで自動実行することにより学習に必要な十分な案件データ量を確保することができた.

窓口となる顧客担当者のパーソナリティを通して、案件の背景や要件が提示される状況が環境であり、またこの環境から営業にフィードバックが返ることになる.

しかし案件シミュレータを利用した学習モデルの作成では、精度を上げるための案件シミュレータのモデリング手法に課題が残った(図3).

具体的には以下の2つの課題となる.

(a) 顧客のパラメータ設定の可変性.

シミュレータによる学習過程では、パラメータにより学習モデルの精度が変わるにも関わらず、パラメータの設定範囲が数種類に固定化されていた.

そのため、顧客も営業も設定範囲の中で行動することになり、すべての選択肢から最適方策をプランニングできない.

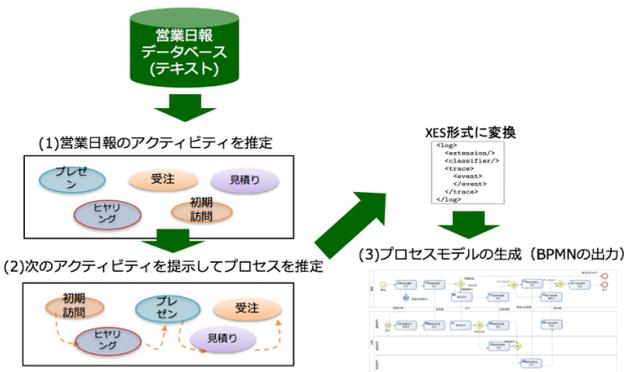


図1：プロセス発見の3ステップ

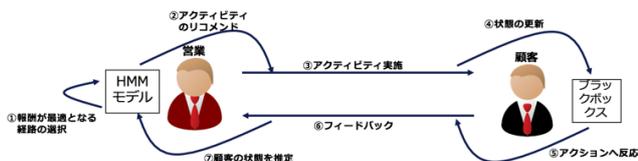


図2：案件シミュレータによる案件データの生成

(b) 営業エージェントの環境からの独立

本来、営業エージェントは環境から独立している必要があるが、このシミュレータではパラメータに環境の情報も含まれているため環境との間に依存性が出てしまい、パラメータ選択に環境の上限値が設定されてしまうことで自由度が制限されてしまう.

そこで本稿では、強化学習における環境自体を学習してそれを学習に活用するという環境モデルの構築を提案する.

具体的には、案件シミュレータで作成された案件データから「状態/アクティビティ/受注/顧客反応」を入力情報として、深層学習による環境モデルを構築する.

顧客の最終状態は、「受注/失注/辞退/強制終了」のいずれかになる.

これにより、パラメータの設定範囲が可変化し、最適化された状態へと改良される.

さらにこの作成された環境モデルを利用して、営業エージェントとの間のシミュレーションをセルフプレイで行うことで、ポリシーと価値を試行錯誤の探索で獲得し、営業エージェントの学習モデルを強化する.

これにより、営業活動場面でより現実の営業活動に近いリコメンドが可能となる手法を確立できることになる(図3).



図3：案件シミュレータを利用した学習モデルの課題

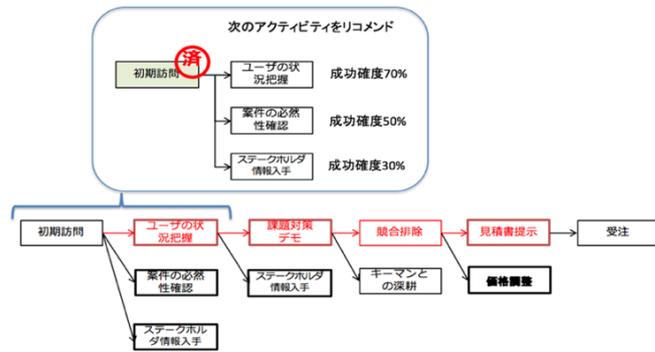


図4：次のアクティビティのリコメンド

## 2. 関連研究

強化学習には大きく2つの種類があり、モデルフリー強化学習では、環境モデルを学習するという事はせず、経験やサンプルから最適な価値関数や方策を学習する[4].

一方で、現実世界等の複雑な環境においては環境のモデルが未知であることが多いため、事前に環境モデルが判明していない状態から環境のモデルを学習するモデルベース強化学習の研究が行われている[5]. 環境モデルを経験から学習し、最適な価値関数や方策をプランニングする.

またこのプランニングによって次のアクションを決定しその経験から学習を繰り返す. つまり、教師あり学習への手法を使って環境モデルを効率的に学習できることと、構築した学習モデルの進展状況が把握できることがメリットとなる.

反面、モデルベース強化学習は環境モデルを推測し、さらに推測したモデルを元にして方策を推測するため、誤差が大きくなる傾向がある[6].

先述のように、営業エージェントと環境(顧客)の間の強化学習の過程を、開発した案件シミュレータで隠れマルコフモデルのアルゴリズムに基づきシミュレーションを行い学習モデルを構築した.

そのため、シミュレータの初期パラメータの設定精度を上げるには、すでに生成されている案件データを教師データとして利用して、環境モデルのパラメータをより最適にする必要があり、それにはモデルベースの強化学習である環境モデルの学習が適している.

## 3. 環境モデルの構築

営業活動における環境は複雑であり、また状態数が多いことから、古典的なQ-Learningでは状態は離散値となるため、状態をある程度量子化しないと計算することができない[5].

また、予定報酬をテーブルで表現するため、状態数が多いとテーブルの要素数が発散してしまう. つまり、通常の強化学習ではモデルの学習が発散する可能性があるため、深層学習アルゴリズムを用いたQ-Learning (DQN) で環境モデルを作成する[7].

その後、作成された環境モデルを利用して、営業エージェントとの間のシミュレーションをセルフプレイで行うことで営業エージェントの学習モデルを強化する. ここではゲーム分野で用いられる深層学習のアルゴリズムであるMCTS (モンテカルロ木探索) アルゴリズム[8][9]を使う.

ゲーム分野での深層学習では、対戦相手がどのような相手でも勝てる確率が高い手を探索することができる.

具体的には、顧客と営業の対戦という形で双方を対戦させ、学習モデルの精度を向上させる. このセルフプレイを繰り返すことでより強い対戦相手をシミュレートすること

ができ、結果として作成された学習モデルはMCTSによる探索で理想的なリコメンドが可能な仕組みとなる(図5).

またこれにより環境に関連する情報は、全て営業エージェントから取り除かれることから、最適方策を探るための制限をなくすることができる.

作成された学習モデルは、実際の営業に適用されることによりフィードバックデータで更新される. 更新した学習モデルを用いて、さらにセルフプレイで学習モデルの精度を向上させる.

これを繰り返してブラッシュアップしていくことで、実際の営業や顧客の動きに近くなるようなアクティビティをリコメンドできる.

## 4. 営業活動の意思決定プロセス強化に向けた環境モデルに基づくアプローチの課題と対応策

営業活動における意思決定に環境モデルによる強化を適用するには営業活動特有の課題がある. 以下ではその課題について述べるとともにその対応策についても記述する.

### 4.1 明確なルールがないセルフプレイ

囲碁などのセルフプレイにおいて深層学習を適用する事例が増えているが[10], 囲碁ではルールがあらかじめ明確である. それに対して営業活動では選択できるアクティビティが無限であり明確なルールがないため、MCTSによる探索が発散してしまい、セルフプレイの適用が不可能である.

そこで、事前に意思決定のプロセスである営業ステップごとにアクティビティを分類しておき、その制約の中からアクティビティを選択することでこの課題を解決する.

この分類制約にはBPMS (Business Process Management System) ツールを用いる. BPMSツールにはプロセス・モデルの機能が包含されており、特定目的のために実施されるアクティビティの関連セットを順序だったステップごとに定義することができる[11].

企業向けの営業活動を例にすると、図6に示すように、初期訪問から始まり、ヒヤリング、プレゼンテーション、見積り、受注へと基本的な意思決定のプロセスは進展していくものの、それぞれのステップの中は様々なアクティビティから構成されている.

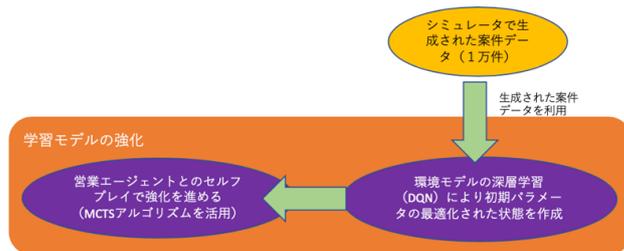


図5：環境モデルの構築と強化

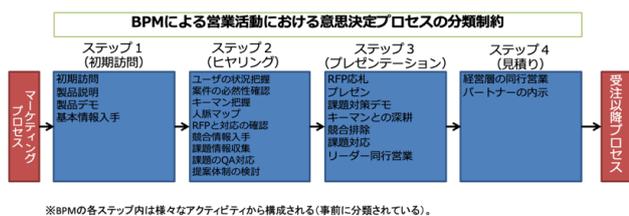


図6：BPMによる営業活動のアクティビティの分類

そのためBPMSツールにより、営業ステップに従ってアクティビティを事前に定義することからヒューリスティックな制約となり、MCTSの発散がなくなるため探索経路が有限になる。

## 4.2 MCTSの適切なアルゴリズム選定

MCTSによる強化学習は Value-Base, Actor-Critic など多数のアルゴリズムが存在するため[8], 今回のモデルに適したアルゴリズムを判定する必要がある。特に、ゴールにたどり着くまで報酬が確定しないような強化学習では、ポリシーネットワークやバリューネットワークを利用して、ゴールにたどり着く可能性が一番高くなるよう探索アルゴリズムを工夫する必要がある。

## 4.3 ゲームと異なるセルフプレイのゴール設定

セルフプレイで顧客と営業を対戦させるには、ゲームの勝敗のつけかたが課題となる。従来のゲームでは勝敗が対となる事象であり、片方が勝利条件に該当する場合、他方は敗北に該当する。

それに対し、営業活動における営業と顧客は共通利益の関係になり、その中でも営業側の活動はより良い条件で受注することがゴールとなっている。一方で、顧客側の活動としては、営業を失注させるのではなく、必要な情報を提供することで、要件にあった提案や見積を得ることが目的となる。

このセルフプレイには、営業活動の内容について吟味する機能はなく、アクティビティの実施有無と顧客の反応で、その有効性を判断している。そのため、理想的な営業活動としては、営業と顧客との意思疎通が十分であり、結果として共通利益が実現できることを目的にする。つまり、可能な限り多くの情報を収集し、顧客の良好な反応を得た上で受注することを目指した学習モデルの構築をおこなう。

## 4.4 教師データが全ての判断パターンを包含できない

環境モデルの深層学習において、教師データとなる案件データは十分なデータ量を確保しているもののすべての判断パターンを包含しているわけではない。

そこで、深層学習による環境モデルの作成時に、人間参加型深層学習 (human-in-the-loop) [12][13]を取り入れることで、より人間と機械知能を組み合わせ、効果的な機

械学習アルゴリズムを生成する。教師あり学習と能動学習の組み合わせとなる。

人間がアルゴリズム構築の学習段階とテスト段階の両方に関わり、絶え間ないフィードバックループができるため、回を重ねるごとにアルゴリズムの結果は向上する。

今回の研究ではフィードバックループの「人間によるアノテーション」を「営業人材によるアノテーション」に置き換えて適用を行い、環境モデルを構築、改善する(図7)。具体的には、学習段階における人間の役割は、結果を確認しながら違和感のある顧客反応をその場で修正する。テスト段階と評価段階においても同様に人間の確認を入れる。

## 4.5 競合は一つではない

実際の営業活動の場面では競合は不特定多数存在することも課題となる。そこで環境モデルを学習する際に、競合を環境と考えてモデル化する。

具体的には、環境モデルの特徴量として競合相手の情報を与えることを検討する。

## 5. 構築された学習モデルの評価計画

上述の手法で構築された強化学習の仕組みを評価する方法について述べる。

ここでは構築された学習モデルからリコメンドされたアクティビティの推移を、弊社の営業経験10年以上の熟練営業担当者10名にアンケートを実施することで、リコメンドの納得感を測定する予定である。これにより、プロセスの強化学習の進展を評価することができる。

また今回の仕組みにより、営業エージェントは環境から独立することが可能となるため、この学習モデルは営業プロセス以外の他領域へも適用することが可能となる。実際にいくつかの適用領域を選定し、評価を進める予定である。

## 6. 結論

従来人の判断が中心であった営業活動の意思決定の過程から、属人性を除外することで、受注確率の高い営業活動を組織全体で実施できるような意思決定支援システムが求められている。

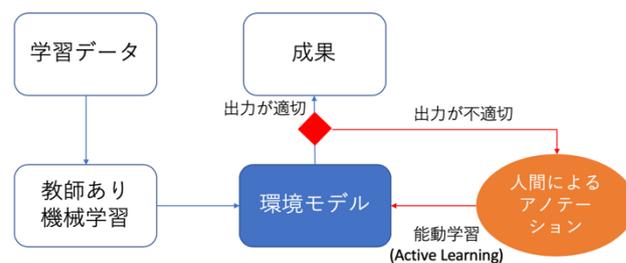


図7：人間参加型深層学習 (human-in-the-loop)

本稿では、受注確率の高い意思決定プロセスの学習モデル構築時に開発された案件シミュレータの課題を、環境モデルを深層学習することで解決する手法を提案した。さらにこの作成された環境モデルを利用して、営業エージェントとの間のシミュレーションをセルフプレイで行うことで、営業エージェントの学習モデルを強化することも期待できる。

しかしその手法の適用には営業活動特有の課題があるため、それらについての解決の方向性についても述べた。具体的には、明確なルールがない状況でのセルフプレイ適用時のBPMS ツールの活用、MCTS の適切なアルゴリズム選定、ゲームと異なるセルフプレイのゴール設定、教師データが全ての判断パターンを包含していないことに対する人間参加型深層学習の適用、さらに競合は一つではないことに対する考慮である。

現在、これら解決策の検証を実施しており、最終的に、効果的な営業活動の意思決定プロセス強化の手法が確立できれば、今後の営業活動の意思決定プロセスの標準化と効率化への寄与が期待できる。

## 参考文献

- [1] 中山 義人, 森 雅広, 成末 義哲, 森川 博之, ” 営業活動における意思決定のプロセス発見手法: プロセスマイニングの応用アプローチ”, 情報システム学会誌 Vol.14, No.1, 2018, pp. 26-38.
- [2] W.M.P. van der Aalst, ” Process Mining: Discovery, Conformance and Enhancement of Business Processes,” Springer, 2011, <http://www.springer.com/jp/book/9783642193446>, <http://www.processmining.org/book/start>.
- [3] 飯島 正, 田端 啓一, 斎藤 忍, “プロセスマイニング・サーベイ(第 01 回: 概要と基本概念),” 情報システム学会誌 Vol.11, No.2, 2017, pp.20-22.
- [4] Marco Wiering, Martijn van Otterlo, ” Reinforcement Learning State-of-the-Art”, Springer-Verlag Berlin Heidelberg 2012, <https://doi.org/10.1007/978-3-642-27645-3>.
- [5] Richard S Sutton. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In Machine Learning Proceedings 1990, pp. 216-224. Elsevier, 1990.
- [6] David Silver, Hado van Hasselt, Matteo Hessel, Tom Schaul, Arthur Guez, Tim Harley, Gabriel Dulac-Arnold, David Reichert, Neil Rabinowitz, Andre Barreto, et al. The predictron: End-to-end learning and planning. In Proceedings of the International Conference on Machine Learning (ICML), pp. 3191-3199, 2017.
- [7] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. NIPS Deep Learning Workshop, 2013.
- [8] Abramson, Bruce (1987). *The Expected-Outcome Model of Two-Player Games* (PDF). Technical report, Department of Computer Science, Columbia University. Retrieved 23 December 2013.
- [9] Bradberry, Jeff (2015-09-07). "Introduction to Monte Carlo Tree Search".
- [10] Abramson, Bruce (1987). *The Expected-Outcome Model of Two-Player Games* (PDF). Technical report, Department of Computer Science, Columbia University. Retrieved 23 December 2013.
- [11] P.H. Feiler and W.S.Humphrey. (1993). Software Process Development and Enactment: Concepts and Definitions, Proc. 2nd Int. Conf. on "Software Process"
- [12] DoD Modeling and Simulation (M&S) Glossary", DoD 5000.59-M, DoD, January 1998 "Directives Division" (PDF). Archived from the original (PDF) on 2007-07-10. Retrieved 2009-04-22.
- [13] Karwowski, Waldemar, *International encyclopedia of ergonomics and human factors*, ISBN 0-415-30430-X, 9780415304306, CRC Press, 2006