

# 少量の教師データでの効率的な 機械学習型予測器の構築

明治大学 総合数理学部  
ネットワークデザイン学科  
専任准教授 櫻井 義尚

2019年12月10日

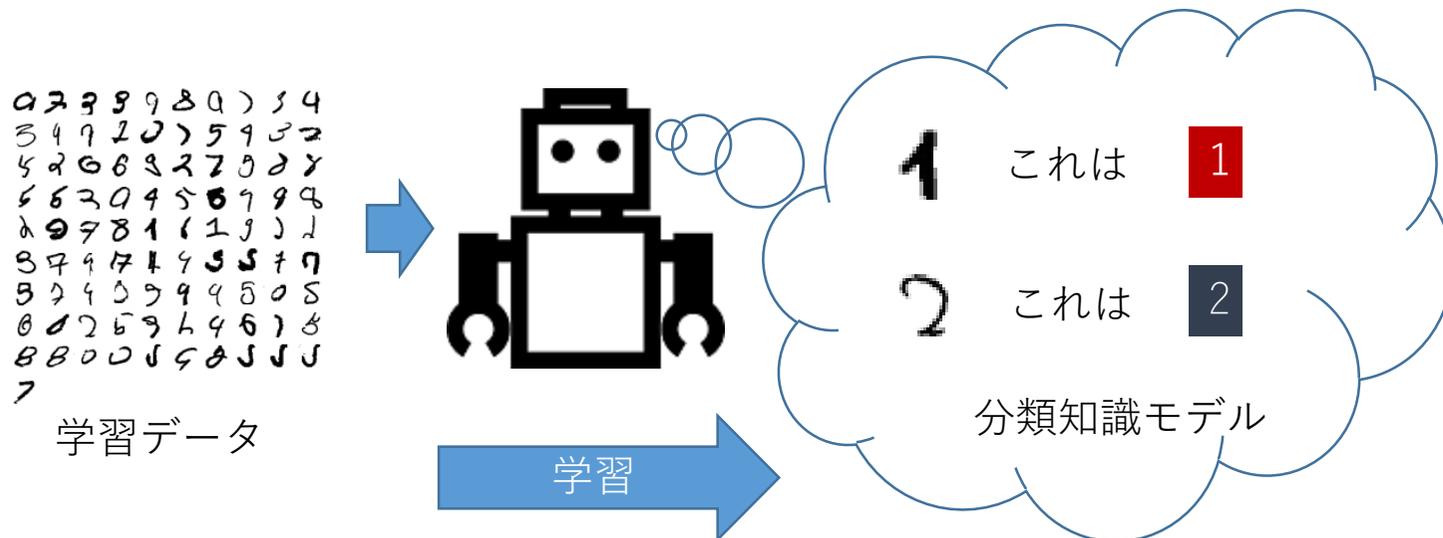


# 目次

- 教師あり機械学習とは
- 従来の機械学習システムの問題点
- 新技術の特徴・従来技術との比較
- 想定される用途
- 実用化に向けた課題
- 企業への期待

# 機械学習とは？

- 人の知能の一つ「学習」を機械で実現する「機械学習」に注目が集まっている。
- データから統計的手法などを用いて学習し、様々な知能を実現する
  - パターン認識（分類）など

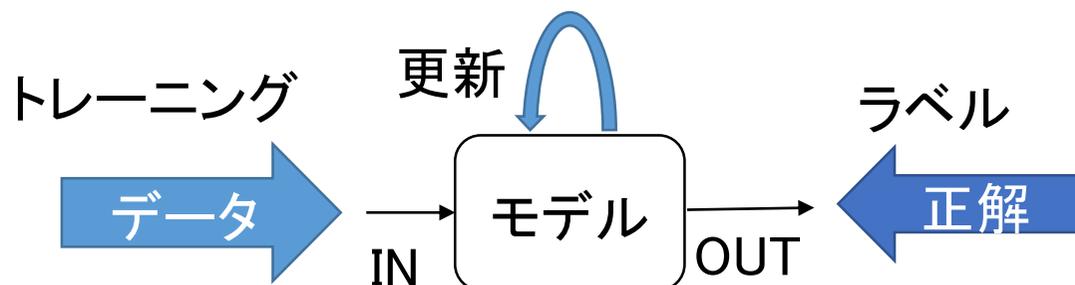


# 教師あり機械学習の学習手順

## ステップ①学習

(正解)事例データから学習する

- ・入力と出力がわかっている事例
- ・入出力のモデルを学習(推定)



## ステップ②推論

未知のデータに対して正解を推定・予測

- ・推定したモデルを使って正解を推定する



# 教師データとは

airplane



automobile



bird



cat



deer



dog



frog



horse



ship



truck



教師データ

(supervised data):

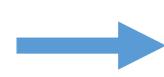
正解がラベル付けされたデータ。ラベル付きデータとも呼ばれる。

トレーニングデータ (training data):

学習に使う教師データ。訓練データ、学習データとも呼ばれる。

出典: <https://www.cs.toronto.edu/~kriz/cifar.html>

画像データ

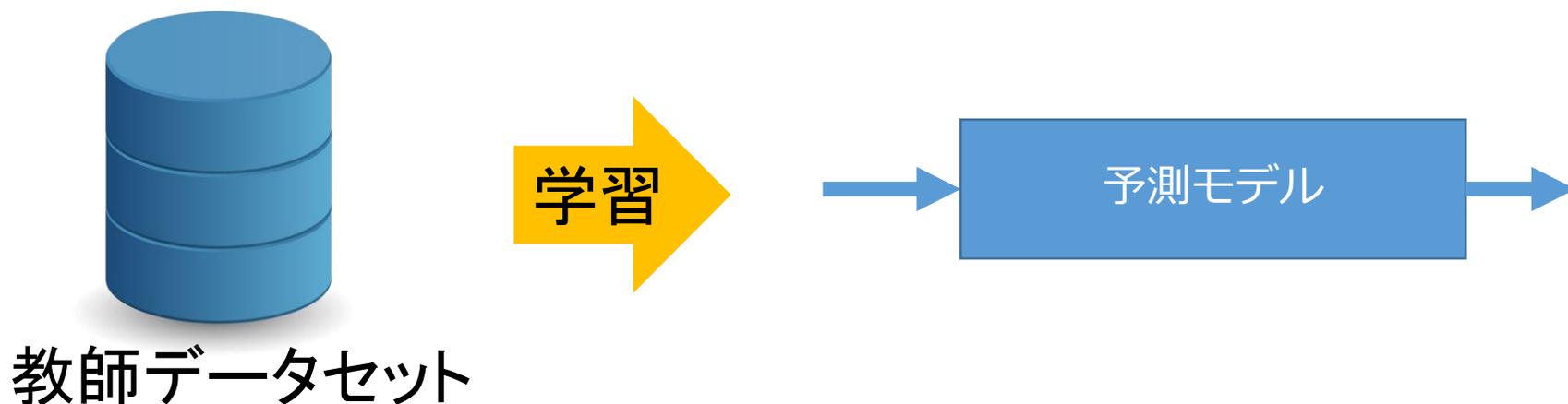


機械学習モデル



何の写真か

# 従来の機械学習システムとその問題点



ビジネス実務向けの精度の高い機械学習予測モデルの構築には**大量の教師データを必要**とする。

## 導入課題

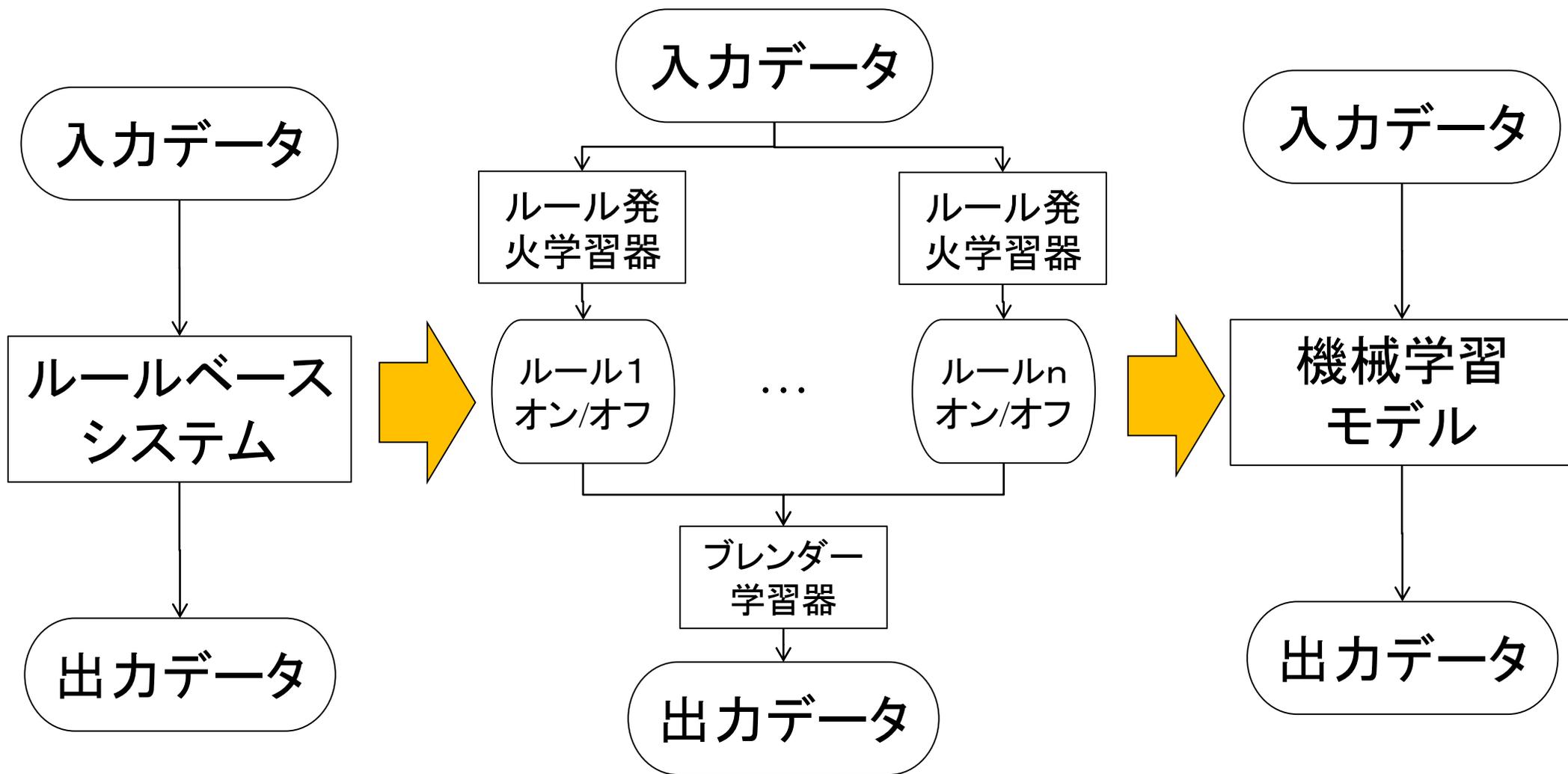
- ✓ 十分な教師データが用意できない
- ✓ 教師データを人手で作成する工数が膨大

# 新技術の特徴

- 教師データだけに頼らず、現場の知識やノウハウを活用することで、少ない教師データから精度の高いモデルを構築
  - 既存のモデルを拡張する形で機械学習を導入可能
  - モデルの蒸留と階層的アンサンブル学習を用いる事でルールベースシステムからの効率的な知識抽出が可能

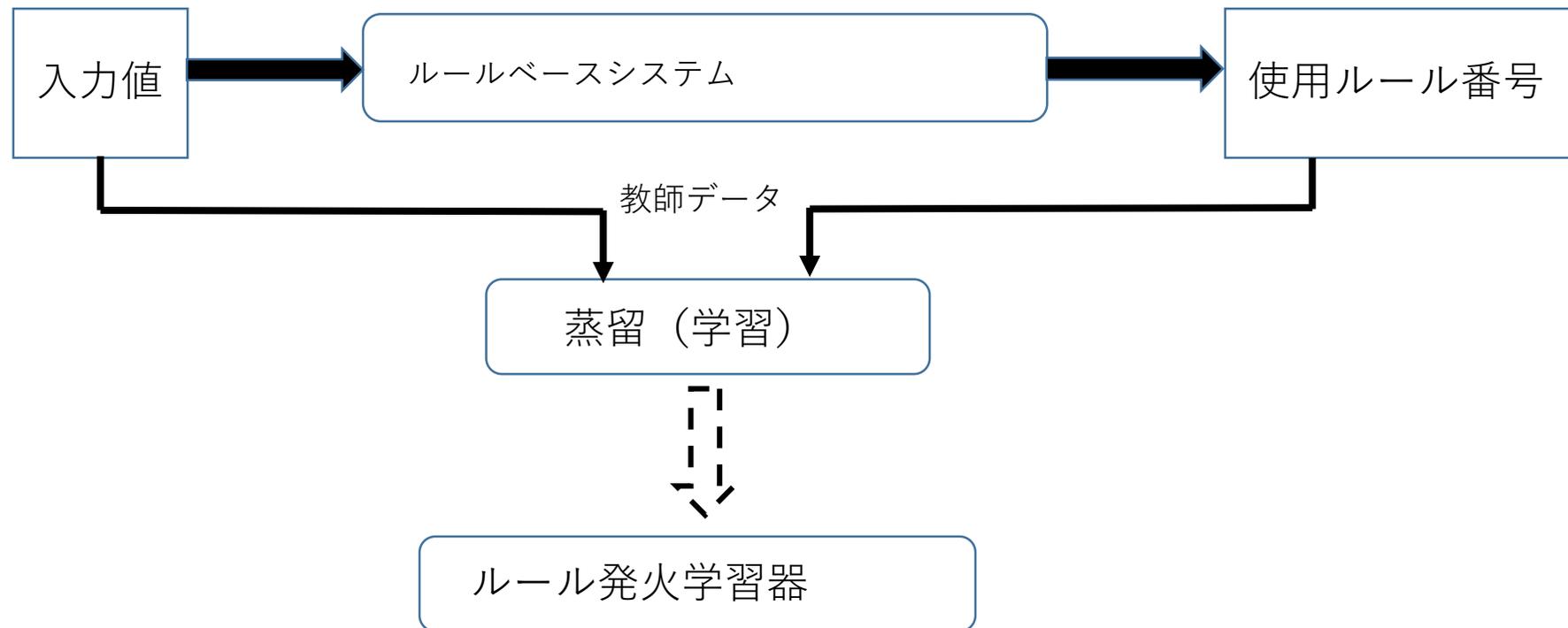
# 新技術の特徴

– 階層的にルールベースから蒸留モデルを作成



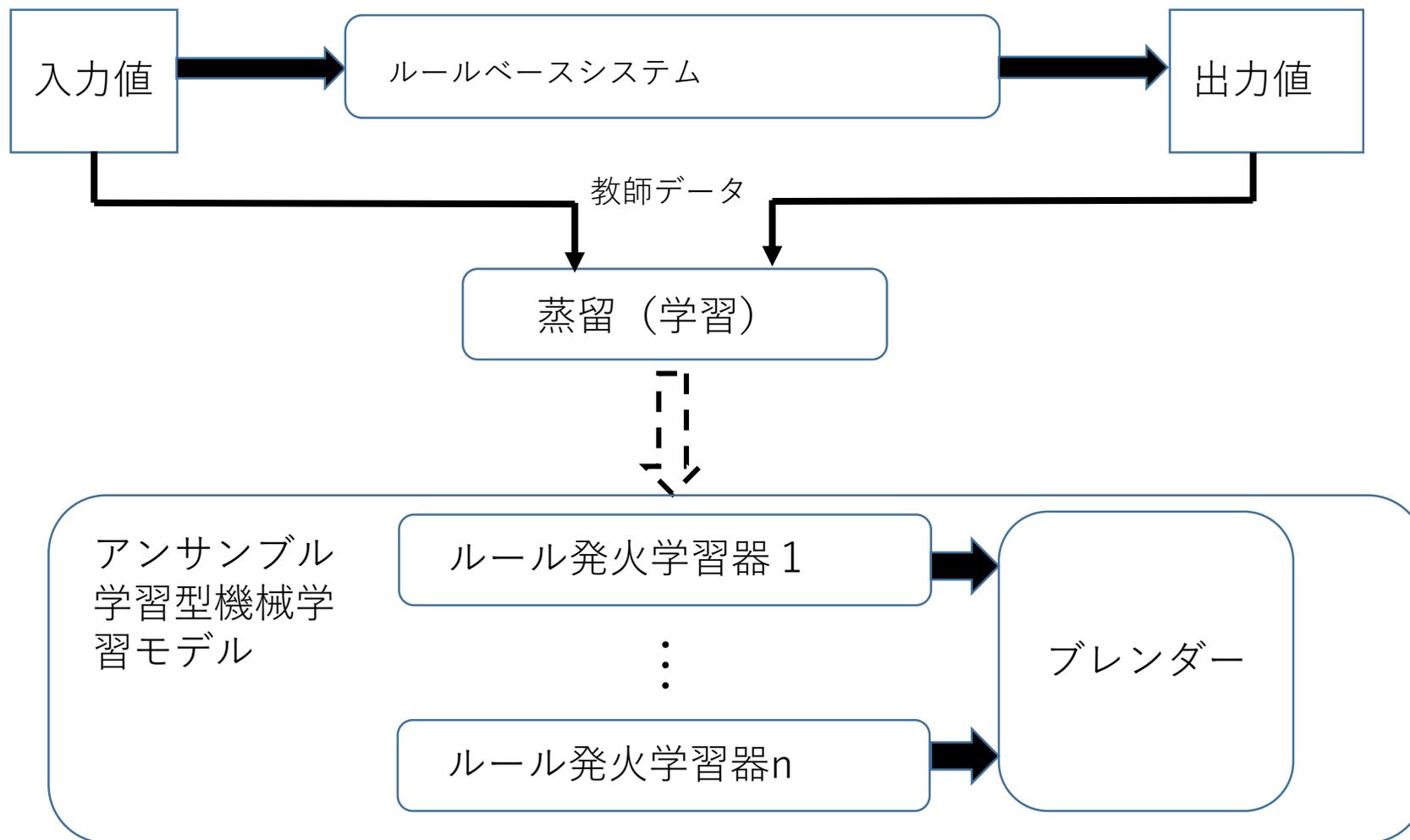
# 新技術の特徴

- まず、ルール発火(入力条件判定)を予測する機械学習モデルを蒸留する



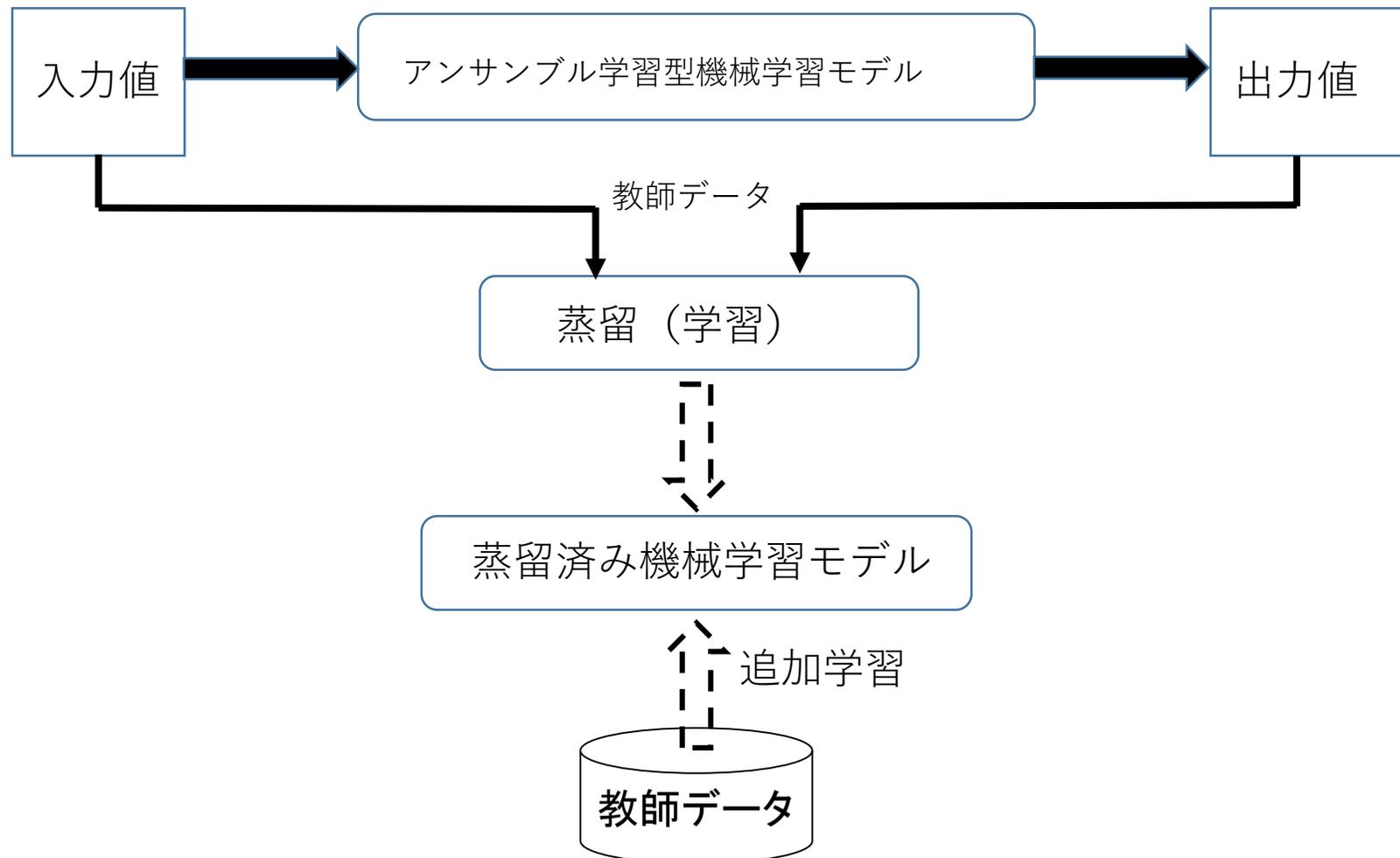
# 新技術の特徴

- ルール発火学習器による出力を特徴量として、最終的な出力値を予測する機械学習モデル(ブレンダー)をルールベースの入出力を教師データとして学習する



# 新技術の特徴

- アンサンブル学習型の機械学習モデルの蒸留モデルを構築し、教師データで追加学習を行うことで最終的な機械学習モデルを構築



# 従来技術との比較

- 従来、ニューラルネットワークからニューラルネットワークへの知識の蒸留はあったが、ルールベースから効率的に蒸留モデルを構築する技術は無かった。
- ルールベースを介して機械学習モデルを構築することで、ビジネスの現場に適したのモデルの構築方法を提供できるようになった。

# 想定される用途

- 大量の教師データが無いシステムの機械学習モデル構築で以下のどれかを満たす場合
  - 既にルールベースによるシステム運用をしている
  - 現場ノウハウからルールベースの構築が可能
- 更新が頻繁な現運用システムを機械学習モデル化したい場合

# 実用化に向けた課題

- 現在、簡易的な実験のみ実施中。実問題レベルの複雑・大規模なシステムでの検証が必要。
- 様々な問題に適用していくことで、現実的な精度や適した課題の検証、分析が必要。

## 企業への期待

- 様々な課題に対して機械学習モデルの構築が現実的な工数で可能になると考えている。
- ルールベースなどのレガシーシステムを機械学習モデルに置き換えていきたい企業との共同研究を希望。
- 文書のフィルタリングなど、自然言語処理分野での展開を考えている企業には、本技術の導入が有効と思われる。

# 本技術に関する知的財産権

- 発明の名称 : 学習モデル構築装置、学習モデル構築方法及びコンピュータプログラム
- 出願人 : 学校法人明治大学
- 発明者 : 櫻井義尚

# 産学連携の経歴

- 2018年-2019年 K社と共同研究実施
- 2019年- T社と共同研究実施

# お問い合わせ先

**明治大学**

**研究推進部 生田研究知財事務室**

**T E L 044-934-7639**

**F A X 044-934-7917**

**e-mail [tlo-ikuta@mics.meiji.ac.jp](mailto:tlo-ikuta@mics.meiji.ac.jp)**