

Pre-training Acquisition Functions for Fixed Budget Active Learning

日野 英逸 モデリング研究系 教授

【概要】

データに対するラベル付けが可能な回数が固定されている状況での能動学習において、ラベル付けを要求するサンプルを選択する獲得関数を深層強化学習によって事前に学習する方法を提案した。これにより、一定のラベル付け回数で予測モデルの精度を向上させる必要がある状況に適した能動学習が実現できた。

【動機と意義】

- ・能動学習(Active Learning)は予測モデル(ここでは判別問題を考える)の予測精度を向上するために適した学習サンプルを逐次的に選択する手法である。特に、少数のラベル付き学習データ集合と、多数のラベルなしデータ集合(プール)が与えられ、プールの中からラベルを付けるべきサンプルを選択する問題設定を考える。
- ・現状の予測モデルとラベル付き学習データ集合の情報から、次にどのサンプルのラベル付けするかは、獲得関数によって決定する。
- ・様々な獲得関数の設計方法が提案されているが、アドホックな方法が多い。

【アプローチ】

- ・近年の能動学習の研究では、データ駆動型の獲得関数の設計に焦点が当てられているが、予算が固定された状況に合わせた獲得関数はまだ研究されていない。固定予算の能動学習問題に適した獲得関数の学習という問題に取り組んだ。
- ・強化学習を用いて固定予算の獲得関数の学習を実現した。具体的にはDQNを用いて、能動学習の運用段階に先立って獲得関数を学習する。強化学習を利用することで、利用可能なサンプル数が固定されている場合に適切なサンプルを選択するように獲得関数の学習を行うことができる。

【強化学習, DQN: Deep Q-Network】

強化学習とは、エージェントが置かれた状態で行動することにより、報酬を最大化しようとする機械学習の一分野である。強化学習により事前に学習された獲得関数を使用した能動学習を考える。強化学習の実装には様々な方法があり、代表的な強化学習法としてQ学習がある。Q学習はある時点でエージェントがある状態にいるとき、ある行動をとることに対する価値を計算する関数を得る方法である。従来のQ学習では、「状態・行動」に対する報酬をテーブルとして学習する。最近の有望なアプローチの一つであるDeep Q-Network(DQN)ではこのテーブルを連続的な関数として扱い、それを深層学習モデルで近似する。連続したデータ取得プロセスは「時間」の概念に対応しており、DQNは能動学習のための獲得関数の学習に既に利用されている。

【提案手法】

能動学習は一般的にデータが少ない状況で用いられることが多く、少数のデータで学習した獲得関数がうまく動作することは期待できない。提案手法では、他のドメインから収集したデータや人工的に作成した大量のデータを用いて、獲得関数(損失低減の予測値)をDQNの枠組みの中でモデル化して学習する。

強化学習における状態と行動を以下のように設計する:

- ・状態: 予測モデルを表すパラメータと学習データを表すパラメータ(例: 回帰係数、プール内の他のデータとの平均距離など)。
- ・行動: どのデータを選択するかを決定するパラメータ(例: 現在の予測モデルで評価した予測の不確実性)。

具体的には、予測モデルとして用いたランダムフォレストのOOB誤差や、ランダムフォレストを構成する木の分岐パターン、ラベル付けされたデータ集合の特徴量ベクトルからなるデザイン行列の固有値などを「状態」とした。また、不確実性サンプリングの考えに基づき、事後確率の最大値が最小となるサンプル選択を「行動」とした。新たに追加したラベル付けされたサンプルを含めて学習した場合の予測精度と、含めない場合の予測精度の差を、DQNの即時報酬とした。

このように状態と行動及び報酬を設計することで、現在の予測モデルとプールデータ(状態)からラベルのないデータ(行動)を選択してテスト損失の削減量(報酬)を予測するQ関数を学習することができる。

【人工データを用いた獲得関数の事前学習】

提案手法では、他領域のデータセットを用いて獲得関数を事前に学習する。本研究では、既存研究で提案されている手順に従い、2クラス判別のための6次元のデータセットを作成した。能動学習実験では、1,000個のデータを検証データとして使用し、残りの9,000個のデータを学習データとプールデータに分けた。学習時にはDQNの学習データ数をランダムに変化させ、様々な状況に対応できるようにした。また、能動学習の予算を100サンプルとし、100サンプル取得した時点でDQNの学習を終了した。

【実データを用いた評価実験】

使用したデータセットのプロファイルを表1にまとめた。比較手法として、代表的な能動学習アルゴリズムであるUncertainty Sampling (US), Query by Committee (QC), 及び強化学習を用いずに獲得関数を人工データから学習する方法であるLearning Active Learning (LAL)を用いた。プールデータ集合から取得するデータ数(予算)は100とした。

表1: 利用したデータのプロファイル

Dataset	Dimension	# of initial samples	# of test samples	# of pools	Attribute type
Googletrip	23	10	1000	4446	Quantitative
Tripadvisor	10	10	1000	31551	Quantitative
Wine white	11	200	1000	3698	Mixed
Wine red	11	100	500	999	Mixed
Car	6	10	500	1218	Quantitative
Adult	14	10	1000	47742	Mixed

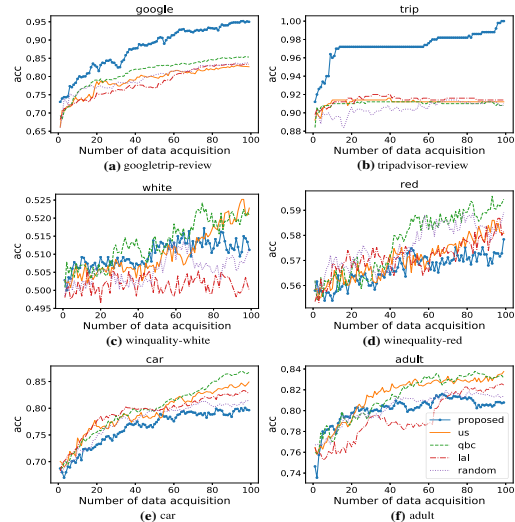


図1: 能動学習による判別精度

この設定で実験を行った結果を図1に示す。提案手法は、googletrip-reviewおよびtripadvisor-reviewデータセットにおいて、他の手法と比較して非常に良い結果を示し、win-whiteおよびadultデータセットにおいては、他の手法と比較して同等の結果を示した。6つのデータセットのうち、提案手法が他の手法と比較して良好な結果を得られなかったのは、Wine-RedデータセットとCarデータセットである。表から、これらの2つのデータセットは比較的小さいプールデータ集合であり、我々の提案手法は実際のデータセットと事前学習データセットが十分に似ていることを保証するために、他の手法よりも大きなプールデータを必要としている可能性がある。

提案手法が予算が固定された状況での最適なデータ選択を実現していることを示すために、能動学習手法とオラクルのデータ選択手法を比較する。ここで、オラクルとは、最も適切なデータセットを選択する手法のことである。組合せ計算を実行可能にするために、プールデータセットのサイズを25、獲得数(予算)を5にした。人工データセットを用いて獲得関数を学習し、tripadvisor-reviewデータセットを用いてテストした。この実験では、最終的なモデルの分類精度と、選択したサブセットのデータのマッチング率を比較する。表2は、5-foldクロスバリデーションの結果を示している。

提案手法によって選択されたデータの照合率は32%であり、ランダムサンプリングよりも高い。これは、文脈を考慮することで、オラクルに近い組み合わせが得られる確率が高くなることを示している。プールデータの数は25個と非常に少ないが、データ取得の組み合わせ数($25_C_5=53,130$)は十分に多い。無作為にデータを取得した場合、選択した5つのデータすべてがオラクルのデータと一致する確率は0.000019%なので、提案手法で得られる結果はランダムに取得した場合の期待値よりも高い。

【今後の課題】

データセット間の性能差は、事前学習に用いたデータセットと能動学習に用いたデータセットの類似性に起因すると考えられる。また、事前学習用データセットとの特徴分布の類似性が、提案手法の性能に最も重要な要素であると推測される。この特徴の類似性を調査し、獲得関数の学習に最適なデータセットを選択することが、今後の重要な課題である。

表2: 最適サンプル系列選択確率

Method	Accuracy mean±std	Index match rate
Oracle	93.59 ± 2.792	-
Proposed	93.59 ± 2.792	32.0%
Random	91.60 ± 2.939	16.0%

本研究は筑波大学山口優氏、亀山啓輔氏との共同研究に基づくものです。