連載

5分で分かる!? 有名論文ナナメ読み

金崎朝子（東京工業大学）

Zhang, Qi and Goldman, Sally A :
EM-DD : An Improved Multiple-Instance Learning Technique

NIPS (2001)

Multiple-Instance 学習とは

弱教師あり学習という言葉を耳にしたことがあるだろうか．一般的な機械学習の多くは「教師あり学習」と呼ばれており，入力信号に対してシステムがどのような出力を出すべきか（教師信号）が学習データとして与えられる．これに対して，弱教師あり学習では弱い教師信号が与えられるのだ．問題を解くための解を与えるのが教師とするならば，ヒントを与えるのが弱教師といったところだ．ところで筆者は昔，ロボットのカメラ画像から物体を認識する研究をしていた．物体認識器を作るためにはロボットにさまざまな角度から物体を見せ，「これがコップだ」等と教える必要がある．このとき，画像にはコップ以外のさまざまな物体が写り込んでしまうので，画像編集ソフト等を起動して画像データを編集し，コップの領域を教師信号として塗りつぶす（手間のかかる）作業が必要であった．そこで筆者は，「物体領域の正解を与えなくとも，画像の中にその物体が写っているか否かという情報を与えるだけで，システムが勝手に正解領域を抽出して学習してくれないだろうか」と考えた．少し調べた後に，まさにそのような問題を解決するのが Multiple-Instance 学習であることが分かったのだ．

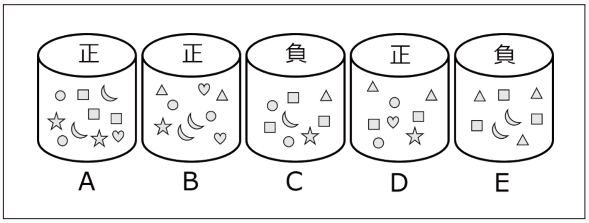

図 -1　Multiple-Instance 学習のサンプル例

Multiple-Instance 学習では複数のインスタンスからなる集合に対して正か負かの教師ラベルが与えられる．正の集合は少なくとも 1 つ以上の正のインスタンスを含み，負の集合は負のインスタンスのみを含む．このように定義された「集合のラベル」を手がかりとして，果たしてどのインスタンスが正のインスタンスであるかを突き止めるというのが Multiple-Instance 学習の醍醐味だ．図 -1 に例を示す．集合 A，集合 B，集合 D は正のラベルを持ち，集合 C と集合 E は負のラベルを持つ．この例では，正の集合内には 1 つ以上存在し，かつ負の集合内には 1 つもないインスタンス，すなわちハート型のインスタンスが正のインスタンスである．

Multiple-Instance 学習が初めて提唱されたのは薬物活性予測問題であった．薬物に含まれる分子はさまざまな低エネルギー構造体を取ることができるが，どの低エネルギー構造体が薬物活性に寄与しているのかが分からない．そこで，さまざまな分子構成の薬物を集め，それらが活性化するかしないかをテストし，その情報を手がかりとして目的の低エネルギー構造体を突き止めた．この問題では低エネルギー構造体がインスタンスであり，薬物がその集合であったが，筆者の研究の例では物体がインスタンスであり，画像がその集合というわけだ．このように，Multiple-Instance 学習が適用できる問題のシナリオは意外とさまざまな分野に散見されるのではないかと思われる．

多様性密度

Multiple-Instance 学習を解くとはどういうことなのだろうか．直感的には，先述の薬物活性予測問題のように何らかの事象の原因を突き止める行為に近い．たとえば，ドラッグつながりというわけではないが，ある薬物 X を使用していた人物が罪を犯したとしよう．この一例のみから薬物 X が危険だと判断するのは短絡的である．しかしながら，もし犯罪者の 90% 以上が薬物 X を使用していたとしたら，薬物 X が犯罪を増長させるかもしれないという疑いがもう少し確からしいものになる．さらには，もし善良な市民の誰もが薬物 X を使用した経験がないとしたら，「薬物 X は犯罪を促す」と言っても過言ではないのかもしれない（最後の条件を見落とすと，犯罪者が全員パンを食べていたという事実からパンを危険な食べ物だと見なすといった明らかな誤判断が許容されてしまうだろう）．

以上の推論を Multiple-Instance 学習の文脈で説明すると，「正のラベルを持つ集合には（多く）共通しており，負のラベルを持つ集合には（あまり）ない特徴が（ある事象の）原因である」と考えられる．ここで言う「原因」のことを Multiple-Instance 学習ではコンセプト点と呼ぶ．コンセプト点を発見するための指標として，特徴空間において「(A) 正の集合の交差から (B) 負の集合の和を引いた度合い」を表す Diverse Density（多様性密度）が提案された．ここで，特徴空間において 0 から 1 までの値を取る二点間の非線形距離を定義する．各点における多様性密度は，(A)「ある正集合に属するすべてのインスタンスとの距離を掛け合わせた値を 1 から引いた値」をすべての正の集合に対して掛け合わせた値と (B) すべての負集合に属するすべてのインスタンスとの距離を掛け合わせた値の掛け算で定義される．つまり，1 つ以上のインスタンスから近い正集合が多ければ多いほど，また負集合のすべてのインスタンスから遠ければ遠いほど多様性密度が高くなり，この指標が最大となる点がコンセプト点であるという考えである．

より，定義に忠実に

多様性密度はすべてのインスタンスとの距離に基づいて計算される．しかしながら，思い出してほしい．Multiple-Instance 学習の定義によれば「正の集合は少なくとも 1 つ以上の正のインスタンスを含む」とある．すなわち，集合の中でたった 1 つのインスタンスが正，つまりコンセプト点に近ければよく，他のすべての点はどうだってよいのだ．負の集合に関しても，集合の中で最もコンセプト点に近い点とコンセプト点との距離が十分に大きければ，すべての点がコンセプト点から遠いということが言える．以上をふまえて，コンセプト点との最近傍点のみを考慮するよう多様性密度の計算方法を修正したのが今回の紹介論文 EM-DD だ．EM-DD は EM アルゴリズムと呼ばれる反復解法を用いる．最初に (E ステップ)，現在の仮説におけるコンセプト点に最も近い点を各集合から抽出する．次に (M ステップ)，これらの点を用いて多様性密度を計算し，これが最大となる点をコンセプト点とする．これらのステップを収束するまで繰り返す．多様性密度の原論文と比較して，EM-DD は計算に用いるインスタンスの数が非常に少ないため，計算量が少ないアルゴリズムになっている．

こうして，EM-DD は Multiple-Instance 学習の定義に忠実かつ効率的な 1 つの解法を世に示したのだ．この記事を読んでいる読者の研究分野で役立つこともあるだろうか．あるいは，よりスマートな解法が思いついたならぜひとも論文を書いて筆者に教えてほしい．

(2020 年 11 月 2 日受付)

金崎朝子 kanezaki@c.titech.ac.jp

2010 年東京大学大学院情報理工学系研究科修士課程修了．2013 年同大同研究科博士課程修了，博士（情報理工学）．2020 年 4 月より東京工業大学准教授．機械学習，物体認識，ロボットビジョンの研究に従事．