

大阪科学・大学記者クラブ 御中  
(同時提供先：文部科学記者会、科学記者会)

2023年1月13日  
大阪公立大学

## マルチラベル識別問題においてデータとラベルの継続学習を実現！ 人工知能の新たなデータ学習手法を開発

### <ポイント>

- ◇マルチラベル識別問題において、データとラベルの継続学習が可能な手法を提案。
- ◇「識別性能」と「継続学習能力」を両立する手法を提案し、数値実験により従来手法に対する優位性を示す。
- ◇データの継続学習能力により、ビッグデータの前処理を行うツールとしての活用も可能。

### <概要>

大阪公立大学大学院 情報学研究科の増山 直輝准教授、能島 裕介教授らの研究グループは、複数の情報（ラベル）を持つデータの「識別性能」と「データを継続的に学習する能力」を両立した、新しい学習手法を開発しました。また、本手法が従来手法に対して優位性を示すことを、実世界マルチラベルデータを用いた数値実験により明らかにしました。

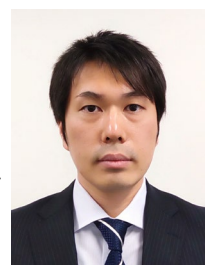
IoT技術の発展に伴い、容易に収集可能となったビッグデータの活用方法として人工知能技術が注目されています。多くの従来研究では、データとそのデータに対応する単一の情報（ラベル）の関係性の学習を主眼としています。しかし、実際的なデータ、例えば1枚の風景写真の中には「空」、「山」、「雲」といった複数のラベルが含まれますが、このようなデータと複数のラベルの関係性の学習を行う手法は十分に研究されていません。また、近年の人工知能には、継続的に入手した大量のデータを効率的に学習するために、過去に学習した知識を損なわずに新しい知識を継続的に学習できる能力も求められますが、これらを両立する手法はありませんでした。

本研究の提案手法は、アルゴリズムがシンプルなため、発展型の考案や他のアルゴリズムとの統合を容易に行うことができます。また、データ間の類似度に基づいてデータをグループ分けするクラスタリング法が根幹となっているため、継続的にビッグデータの前処理を行うツールとしての活用が期待できます。

本研究成果は、国際学術誌「IEEE Transactions on Pattern Analysis and Machine Intelligence」のオンライン速報版に、2022年12月19日に掲載されました。

IoT技術の発展に伴うビッグデータ社会において、人工知能によるデータの効率的な利活用の実現は非常に重要です。

本研究の提案手法は、データとラベル情報が一対多の関係にあるマルチラベルデータを継続的に学習可能であり、今後のさらなるビッグデータ社会において人工知能に求められる能力を有していると考えています。



増山 直輝准教授

## <研究の背景>

従来の人工知能によるデータの学習では、あるデータとそれに対応する現象や物体（ラベル情報）が一对一の関係にあるシングルラベル識別問題を主に扱っていました。しかし、実世界において多くの場合でデータとラベル情報が一对一の関係になることは稀です。そこで近年では、より実際的な問題としてデータとラベル情報が一对多の関係にあるマルチラベルデータを扱うマルチラベル識別問題に注目が集まっています。

IoT 技術の進歩により、私たちは大量の多種多様なデータを容易かつ継続的に入手することが可能となりました。人工知能がこのようなデータを効率的に学習するためには、過去に学習した知識を損なわずに新しい知識を継続的に学習する能力が重要になります。しかし、従来研究においてマルチラベル識別問題を扱う手法は提案されていますが、「識別性能」と「データを継続に学習する能力」を両立する手法はありませんでした。

## <研究の内容>

本研究の提案手法は、データの学習を脳の電気生理学的な知見に基づく学習理論である適応共鳴理論<sup>\*1</sup>によるクラスタリング手法によって行います。この手法は、逐次的に入力されるデータを自律適応的に集約し、識別器となるクラスタを継続的かつ効率的に生成することができます。また、各クラスタに付与するラベル情報を、ベイズの定理に基づく手法により継続的に学習します。このように、データの学習とデータに対応するラベル情報の学習を独立に、かつ継続的に行うことで高い識別性能と継続学習能力を両立しています。

以下に示すものは、2次元人工データに対する継続学習の一例です。図1は、学習するデータセットを示しています。図1(a)から(g)のデータ分布は、それぞれ順番に入力され、各データ分布は0または1で表現される複数のラベル情報を持ちます。図2は、本研究の提案手法の学習過程を示したものです。各データ分布が入力されるごとに、過去に学習した情報を保持しつつ、新しい情報を学習出来ている様子が見て取れます。図3は、比較手法による学習過程です。比較手法ではデータの学習は可能ですが、各データ分布のラベル情報の学習は適切に行えていないことが分かります。

論文内では16個の実世界マルチラベルデータを用いた識別実験を行い、最新の手法を含む6種類の比較手法に対する提案手法の優位性を示しました。

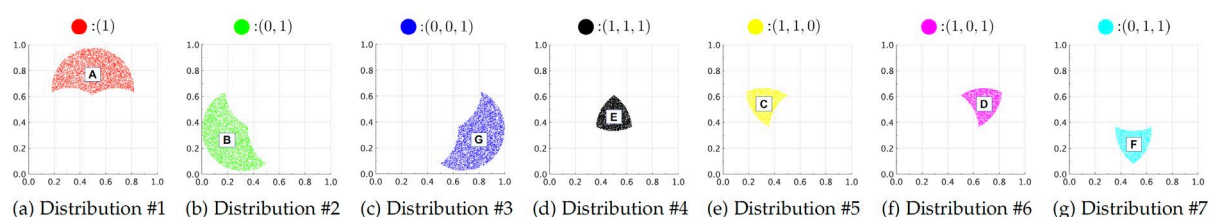


図1 学習する2次元人工データセットとラベル情報

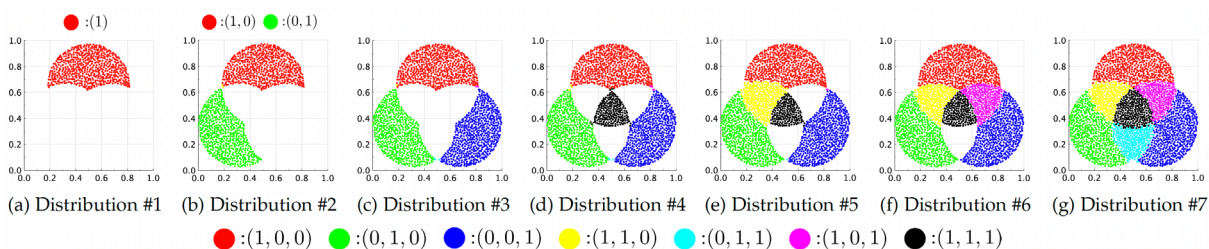


図2 本研究の提案手法によるデータの学習過程

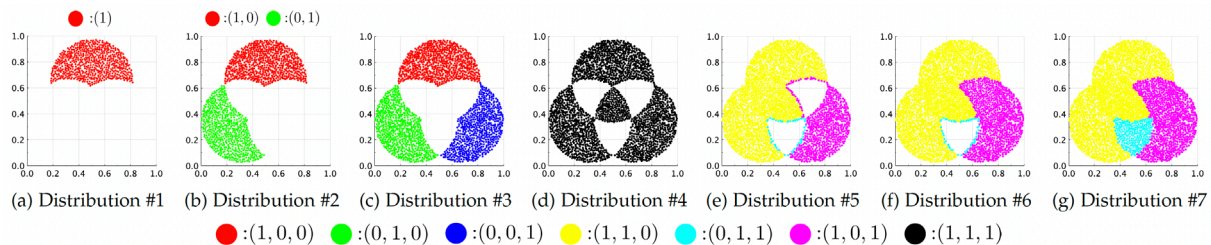


図3 比較手法によるデータの学習過程

### <期待される効果・今後の展開>

本研究の提案手法は、実世界の実際的なデータであるマルチラベルデータを扱うことが可能です。さらに、継続的な学習能力を持つため、今後の人工知能に求められる能力を有していると考えられます。

本研究の提案手法の根幹はクラスタリング手法のため、データの前処理として適用することができます。そのため、データを利用する様々な分野への応用が可能です。また、アルゴリズム自体が非常にシンプルで、さまざまな発展型の考案や他のアルゴリズムとの統合が容易なため、今後は、提案手法の拡張と実アプリケーションへの実装に取り組みます。

### <資金情報>

本研究の一部は、平成 29 年度文部科学省卓越研究員事業および日本学術振興会科学研究費助成事業（課題番号：JP22K12199）からの支援を受けて行われました。

### <掲載誌情報>

【発表雑誌】IEEE Transactions on Pattern Analysis and Machine Intelligence (IF=24.314)

【論文名】Multi-label Classification via Adaptive Resonance Theory-based Clustering

【著者】N. Masuyama, Y. Nojima, C. K. Loo, and H. Ishibuchi

【掲載 URL】<https://doi.org/10.1109/TPAMI.2022.3230414>

### <用語解説>

※1 適応共鳴理論…適応共鳴理論は脳の電気生理学的な知見に基づく学習理論であり、その基本的概念は 1976 年に S. Grossberg によって提案された。多くの機械学習手法において、過去に学習した知識を保持する能力（安定性）と新しい知識を学習する能力（可塑性）はトレードオフの関係を示す。一般的に、このトレードオフ関係は安定性と可塑性のジレンマと呼ばれ、適応共鳴理論はこのジレンマを解決するアプローチの 1 つとして知られている。

#### 【研究内容に関する問い合わせ先】

大阪公立大学大学院 情報学研究科  
准教授 増山 直輝 (ますやま なおき)  
TEL : 072-254-9825  
E-mail : [masuyama@omu.ac.jp](mailto:masuyama@omu.ac.jp)

#### 【報道に関する問い合わせ先】

大阪公立大学 広報課  
担当：竹内  
TEL : 06-6605-3411  
E-mail : [koho-list@ml.omu.ac.jp](mailto:koho-list@ml.omu.ac.jp)