

# 分類の類似度指標

研究分野: 人工知能基礎

キーワード: 分類、等価性の類似、商集合の類似、F値

貢献できるSDGsの区分:



情報システム学部 情報セキュリティ学科 教授 山口 文彦

教員情報URL <https://sun.ac.jp/researchinfo/yamagu/>

## 研究概要

機械学習などによって対象の集合の分類を得るとき、何らかの方法で、その分類の正しさを評価することが必要になる。教師あり学習の場合には、教師として正解である分類が与えられるので、学習器が得た分類と正解との類似度によって、学習器の性能を評価することができる。分類するという行為を、有限な集合上に同値関係を導入して、その同値関係による商集合を得ることであると捉えれば、分類の類似を測るためには、商集合の類似を測ればよいと思える。

しかし、Jaccard係数のような従来の集合の類似度を商集合に適用すると、要素の類似が考慮されないために、精緻さに欠ける指標となってしまう。そこで、一方の分類が他方の分類よりも細かい分類となっているか否かを測る指標を提案したうえで、両方向の指標の調和平均によって分類の類似度を定義する提案を行った。

本研究室では、手書き記号を分類する問題について本手法を適用しているが、本手法は一般に有限な集合の分類について、二つの分類の間の類似度を定義するものである。

## 産学連携の可能性(アピールポイント)

ビッグデータ解析などにおいて、対象の分類が結果として得られることがある。二つのグループに分類する場合は、従来の集合の類似度を使って評価することが可能だが、分類後のグループ数が不特定な場合には、従来の類似度指標が適用しづらいことがある。

本研究の成果は、そのような場合の分類器の性能評価などに使うことができる。

## 外部との連携実績等

〈2012年4月1日～2016年3月31日〉 科学研究費補助金, 24500313

〈2015年4月1日～2019年3月31日〉 科学研究費補助金, 15K00472