

多値分類問題における指標の再検討：F 値を中心に

林賢一（慶應義塾大学 理工学部）

要旨：分類問題は、医学的診断研究や機械学習など、多くの分野に現れる問題である。最も一般的な分類評価指標である正答率 (accuracy) は、クラス割合が不均衡なデータセットに対して、少数派のクラスの分類性能の改善を適切に反映しないことがしばしば指摘される。二値分類問題においては、正答率に代わる指標として F 値 (F-measure/F-score) や Matthews 相関係数 (MCC) などが用いられる。本研究では、これらを多値分類問題に拡張した指標の問題点とそれらを克服するためのアプローチについて議論する。