

先端科学技術研究科 修士論文要旨

| | | | |
|---|-------------------------------|-----|--------------|
| 所属研究室 (主指導教員) | 数理情報学 (池田 和司 (教授)) | | |
| 学籍番号 | 2111190 | 提出日 | 令和 5年 1月 19日 |
| 学生氏名 | 西村 虎太郎ジェームス | | |
| 論文題目 | サンプル分布の局所密度に着目したカーネル法に基づく境界補正 | | |
| 要旨 | | | |
| <p>機械学習が著しい成功を収めた今日においても、実用ベースでは不均衡な分布を持つデータセットへの対処に多くのコストが割かれている。これは実世界でサンプリングされた多くのデータが多数派クラスと少数派クラスが存在する不均衡な分布を持つことに起因する。このような不均衡データに対して学習を行ったモデルは多数派クラスに過剰に適合する傾向がありクラス間のアンバランスが問題になる場合がある。すなわち、全体の精度は十分であるように見えるが、少数派クラスに対する精度は低いということが起こり得る。このような問題に対する対策は多く存在するが情報の欠損やデータ構造の破壊や、データの局所的分布の無視といった問題が存在する。そこで本研究ではサンプルの局所的な分布に着目した新たな手法の提案を行う。そして、人工データ及び実データから作成した不均衡データに対して実験を行い提案手法が既存手法に比べて有利な条件が存在することを示した。</p> | | | |