

## 論文の内容の要旨

論文題目      Excess Risk Transfer and Learning Problem Reduction  
                         towards Reliable Machine Learning  
                         (信頼性の高い機械学習を目指した剰余リスク転移と  
                         学習問題の帰着)

氏      名      包      含

現代社会ではウェブの出現や計算資源の急速な発達によって莫大なデータの収集が可能になり、データ駆動型の知識発見において機械学習が重要な要素技術の一つとなつつある。特に統計的機械学習に基づく帰納推論は目覚ましい成功を収めており、演繹推論に基づく旧来の人工知能に置き換わっている。データが大量にかつ自動的に蓄積されるような応用領域では、帰納推論に基づく知識システムが従来に比べてより柔軟で応用しやすくなっており、科学や工学などの水準が急速に向上している。

実用的には成功しているものの、演繹推論とは異なり帰納推論は実例を集めて結果の正しさを蓋然的に保証することしかできない。仮に統計的推論がリスクの大きい応用領域にて用いられる場合、有限の観測が与えられた下で未知の入力に対して予測がどの程度正しいかを知ることは肝要である。汎化の統計的学習理論はこの課題意識に根ざしている。汎化理論の中心的な目的は、経験リスク（有限のデータから計算された予測と観測の差）と期待リスク（未知のデータから計算された予測と観測の差）の違いを調べることである。過去数十年の間、当該分野は統計学、計算機科学、情報理論の各分野に支えられて大きな発展を遂げてきた。大まかに言えば、汎化理論によって、一般的には十分量の観測さえ与えられれば学習器は期待される出力を正しく予測することが可能であることがわかっている。それでは、信頼できる学習システムを構築するためには、できるだけ多くのデータを学習アルゴリズムに対して流し込むだけで果たして十分なのであろうか。

本博士論文では、汎化理論とは趣を異にする剰余リスク転移の学習理論を構築し、信頼性の高い学習アルゴリズムの設計や学習のメカニズムの解明においてリスク関数の設計が重要であるという見方を示す。剰余リスク転移の重要性は、二十一世紀初頭に行われてきた代理リスクに関する一連の研究を通して認知されはじめた。代理リスクは一

一般的な機械学習手法が最適化する学習基準であり、最終的な評価指標である目的リスクと異なっていたとしても、損失が適切に設計されていれば両者には単調な関係が存在することが明らかにされた。この単調性によって剰余リスク転移、すなわち代理リスクの剰余量の最小化によって目的リスクの剰余量の最小化が実現されるという関係性を得ることができる。結果的に、一般的に離散構造を持つが故に最適化が容易ではない目的リスクを用いる代わりに、最適化しやすい代理リスクを用いることが正当化される。

これまでの研究が主に二値分類や多値分類といった古典的な学習問題の解析を対象としてきた一方で、近年の学習問題はより複雑な構造を持つことが多い。本博士論文では、次の二つの観点から剰余リスク転移の理論を古典的な設定から拡張する。第一に、剰余リスク転移を予測に対するより多様な制約や性質と関連付ける。この関連付けは、従来理論が予測の正しさにのみ着目してきた点と大きく異なる。第二に、一見異なる二つの学習問題の間に関連性を見出すために剰余リスク転移を利用する。従来は目的リスクが最終的な対象であり代理リスクはあくまで補助的な量とみなされてきたが、ここでは二つの学習問題の困難性を比較するのに剰余リスク転移が利用できることを確認する。まとめると、剰余リスク転移によってより多種多様な制約を含む多くの学習問題が扱えるようになり、さらに学習問題の困難性に基づいて問題同士の関連性を明らかにすることができるようになる。したがって、より広いクラスの学習問題の構造や学習問題からどのような知識を獲得できるかが明らかになり、我々の最終目的に対して必要最小限な学習問題を設計する指針を得ることができる。補足すると、ここで考察している剰余リスク転移の理論は汎化理論とは概念的に独立しており、両者を統合することが可能である。以下、本博士論文の構成の詳細を示す。

第一章では、機械学習と統計的学習理論の歴史を振り返り、本学位論文の貢献をまとめる。

第二章では、教師付き学習の背景知識とその理論をまとめる。教師付き分類を定式化した後、汎化理論と代理誤差の古典的な解析方法に触れる。また、分類適合的損失や proper loss といった関連する概念を最近の関連研究とともに紹介する。

第三章では、F値やJaccard指標といった複雑な評価指標で分類器を評価する際に適切な代理目的関数の設計を行う。これらの複雑な評価指標は、近年情報検索やセマンティック・セグメンテーションにおいてクラス不均衡に対処するためによく用いられているが、線形分数型指標と呼ばれるクラスに属しており、古典的な代理損失を適用しにくい分解不可能性という性質を持つ。ここでは目的となる評価指標の代理となる目的関数を適切に設計することで、代理目的関数を最適化した際に評価指標が最適化されるための十分条件を導出する。ベンチマークデータを用いた数値実験では、特にデータ数が小さい状況下において、クラス事後確率に基づくプラグイン分類器と比較して設計した代理目的関数が良い性能を発揮することが確認できた。この結果は提案手法の有用性を検証するだけでなく、プラグイン分類器に対する優位性を知る上でも重要な手がかりとなる。

第四章では、敵対的攻撃に対して頑健な分類、すなわち敵対者がテスト入力に対して加える摂動に影響されにくい分類器の学習を考える。この学習問題は、摂動の大きさに対する制約下での最悪の分類誤差値を目的誤差とする、ミニマックス最適化として最適化されることが多い。従来の研究ではこの最適化を緩和することに主眼が置かれてきたが、目的リスクの観点から緩和問題が正当化できるかどうかは知られていない。そのため、剰余リスク転移解析を用いて頑健な分類誤差に対して適合している代理損失を調べ、結果として線形モデルの仮定の下では真に頑健な解が得られる凸な代理損失が存在しないことを明らかにしたと同時に、非凸損失を設計するために有用な指針が得られた。この結果は、敵対者の導入によって凸代理損失が適合的にならなくなるという事実に意外性があり、また適合的代理損失の理論が予測の正しさ以外に予測が備えるべき性質に対しても適用可能であるということが明らかになった点が重要である。

第五章では、類似度学習と分類の関係性を議論する。類似度学習は二つのデータ間の関係性を予測することでデータの有用な表現を得るための枠組みであり、距離学習や対比学習といった前処理問題を包摂する。学習した表現を用いて構成した分類器によって分類性能が向上されることが期待されているが、これまで理論的な背景はほとんど知られてこなかった。類似度学習がどのように分類性能を向上させるかを解明するため、我々は類似度学習の特定の定式化が分類誤差と密接に関係していることを示した。剰余リスク転移の観点から、この関係性によって類似度予測リスクの剰余量を最小化することで分類剰余リスクが最小化できることが説明できる。その結果、類似度学習は内部的には二値分類境界を学習しているということがわかった。剰余リスク転移の考え方をを用いると、例えば二値分類と類似度学習のように、二つの異なる学習問題をそれぞれ目的問題と代理問題とみなすことができ、学習問題間の関係性を調べたり、ある問題が別の問題に帰着したりすることができるかどうかを特徴づけることができるようになる可能性が秘められている。

第六章では、本学位論文の結論とこれからの課題を述べる。本論文では剰余リスク転移の見方から学習理論の新しい展望を提示し、学習器が期待される性質を満たしているかどうかを解析したり、学習問題を相互に帰着したりできるようになった。