

Title	勾配ブースティングの高度化と応用に関する研究
Sub Title	A study on the advancement and application of gradient boosting
Author	鈴木, 秀男(Suzuki, Hideo)
Publisher	慶應義塾大学
Publication year	2023
Jtitle	学事振興資金研究成果実績報告書 (2022.)
JaLC DOI	
Abstract	<p>教師あり学習のうちの一つである勾配ブースティングは、既に作成している学習器について、与えた損失関数によって得られる勾配を疑似的な残差として定義し、その残差に対して適合することでブースティングの構造を取る手法である。</p> <p>本研究では、勾配ブースティングに対して正則化アルゴリズムと、Momentum SGD (Stochastic Gradient Boosting)やAdadelata、Adamといった最適化アルゴリズムを導入することにより、複数の手法を提案する。正則化は、作成するモデルの自由度に制限を加えることにより過学習を抑える効果がある。正則化項（または罰則項）を作成し、その値をモデルの損失と同時に扱う。本研究では、決定木の全ての葉のスコアにおけるL1正則化項、L2正則化項を用いた。また、従来のSGDでは、学習データをシャッフルし、その中からランダムに1つを取り出して誤差を計算し、勾配法により損失関数が小さくなるようにパラメータを更新する。従来のSGDの問題点として、収束が遅く振動や鞍点に陥ることがある。SGDを改良したMomentum SGD等の手法では、既に得られている前の勾配情報を用いることで振動を抑制して従来のSDGの問題を緩和している。勾配ブースティングに対する正則化および最適化アルゴリズムの効果を検証するために、UCI Machine Learning Repositoryのいくつかのデータセットに対する予測精度と計算効率の指標の測定を行い、従来のSGD、SGD（正則化）、SGD（最適化アルゴリズム）、SGD（正則化+最適化アルゴリズムMomentum）の比較検討を行った。その結果、おおむね、予測精度および計算効率の観点から、SDG（正則化+Adam）が良いことが示された。正則化とMomentumの相乗効果があることが示唆される。</p> <p>今後、他の正則化項（例えば、Elastic NetやGrouped Lassoなど）、勾配ブースティングに特化した最適化アルゴリズムの提案が課題である。</p> <p>Gradient boosting, regarded as one of the machine learning methods, is a method to take the structure of boosting by defining the gradient obtained by the given loss function as a pseudo residual and fitting it to the residual.</p> <p>In this study, we propose several methods for gradient boosting by introducing regularization and optimization algorithms such as Momentum SGD (Stochastic Gradient Boosting), Adadelata, and Adam. Regularization has the effect of suppressing overfitting by constraining the degrees of freedom of the constructed model. The regularization term (or penalty term) is created, and its value is treated as the model loss. The regularization term is calculated numerically and added to the loss function obtained from the residual between the predicted and measured values. In this study, we use the L1 regularization term and the L2 regularization term in the scores of all leaves of the decision tree. In the conventional SGD, the training data is shuffled, one is randomly extracted from it, the error is calculated, and the parameters are updated so as to reduce the loss function by using the gradient method. The problem with the conventional SGD is that it converges slowly, causing vibrations and saddle points. The optimization algorithms, which are improved versions of the conventional SGD, suppress vibration by using the gradient information from the previous period, which enables us to alleviate the problems of the conventional SGD. In order to verify the effect of regularization and improved algorithms on gradient boosting, we measure the predictive accuracy and calculation efficiency indicators for several datasets of the UCI Machine Learning Repository, and compare those of the conventional SGD, SGD (regularization), SGD (improved algorithms) and SGD (regularization +improved algorithms). The result shows that SGD (regularization + Adam) is generally good in terms of prediction accuracy and calculation efficiency. It suggests that there is a synergistic effect of regularization and improved algorithms. In future studies, using other regularization terms (e.g., Elastic Net, grouped Lasso) and proposing a new algorithm customized for the gradient boosting are considered.</p>
Notes	
Genre	Research Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=2022000010-20220090

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

研究代表者	所属	理工学部	職名	教授	補助額	200 (B) 千円
	氏名	鈴木 秀男	氏名 (英語)	Hideo Suzuki		
研究課題 (日本語)						
勾配ブースティングの高度化と応用に関する研究						
研究課題 (英訳)						
A Study on the Advancement and Application of Gradient Boosting						
1. 研究成果実績の概要						
<p>教師あり学習のうちの一つである勾配ブースティングは、既に作成している学習器について、与えた損失関数によって得られる勾配を疑似的な残差として定義し、その残差に対して適合することでブースティングの構造を取る手法である。</p> <p>本研究では、勾配ブースティングに対して正則化アルゴリズムと、Momentum SGD (Stochastic Gradient Boosting)や Adadelta、Adamといった最適化アルゴリズムを導入することにより、複数の手法を提案する。正則化は、作成するモデルの自由度に制限を加えることにより過学習を抑える効果がある。正則化項(または罰則項)を作成し、その値をモデルの損失と同時に扱う。本研究では、決定木の全ての葉のスコアにおける L1 正則化項、L2 正則化項を用いた。また、従来の SGD では、学習データをシャッフルし、その中からランダムに1つを取り出して誤差を計算し、勾配法により損失関数が小さくなるようにパラメータを更新する。従来の SGD の問題点として、収束が遅く振動や鞍点に陥ることがある。SGD を改良した Momentum SGD 等の手法では、既に得られている前の勾配情報を用いることで振動を抑制して従来の SGD の問題を緩和している。</p> <p>勾配ブースティングに対する正則化および最適化アルゴリズムの効果を検証するために、UCI Machine Learning Repository のいくつかのデータセットに対する予測精度と計算効率の指標の測定を行い、従来の SGD、SGD (正則化)、SGD (最適化アルゴリズム)、SGD (正則化 + 最適化アルゴリズム Momentum) の比較検討を行った。その結果、おおむね、予測精度および計算効率の観点から、SGD (正則化 + Adam) が良いことが示された。正則化と Momentum の相乗効果があることが示唆される。</p> <p>今後、他の正則化項(例えば、Elastic Net や Grouped Lasso など)、勾配ブースティングに特化した最適化アルゴリズムの提案が課題である。</p>						
2. 研究成果実績の概要 (英訳)						
<p>Gradient boosting, regarded as one of the machine learning methods, is a method to take the structure of boosting by defining the gradient obtained by the given loss function as a pseudo residual and fitting it to the residual.</p> <p>In this study, we propose several methods for gradient boosting by introducing regularization and optimization algorithms such as Momentum SGD (Stochastic Gradient Boosting), Adadelta, and Adam. Regularization has the effect of suppressing overfitting by constraining the degrees of freedom of the constructed model. The regularization term (or penalty term) is created, and its value is treated as the model loss. The regularization term is calculated numerically and added to the loss function obtained from the residual between the predicted and measured values. In this study, we use the L1 regularization term and the L2 regularization term in the scores of all leaves of the decision tree. In the conventional SGD, the training data is shuffled, one is randomly extracted from it, the error is calculated, and the parameters are updated so as to reduce the loss function by using the gradient method. The problem with the conventional SGD is that it converges slowly, causing vibrations and saddle points. The optimization algorithms, which are improved versions of the conventional SGD, suppress vibration by using the gradient information from the previous period, which enables us to alleviate the problems of the conventional SGD. In order to verify the effect of regularization and improved algorithms on gradient boosting, we measure the predictive accuracy and calculation efficiency indicators for several datasets of the UCI Machine Learning Repository, and compare those of the conventional SGD, SGD (regularization), SGD (improved algorithms) and SGD (regularization + improved algorithms). The result shows that SGD (regularization + Adam) is generally good in terms of prediction accuracy and calculation efficiency. It suggests that there is a synergistic effect of regularization and improved algorithms. In future studies, using other regularization terms (e.g., Elastic Net, grouped Lasso) and proposing a new algorithm customized for the gradient boosting are considered.</p>						
3. 本研究課題に関する発表						
発表者氏名 (著者・講演者)	発表課題名 (著書名・演題)	発表学術誌名 (著書発行所・講演学会)	学術誌発行年月 (著書発行年月・講演年月)			
Nagomu Iwasa and Hideo Suzuki	The Proceeding of the 20th ANQ Congress 2022 (Beijing, China)	The Proceeding of the 20th ANQ Congress 2022 (Beijing, China)	2022年10月			
岩佐和、鈴木秀男	正則化と Adam を用いた勾配ブースティングの改良	(一社)日本品質管理学会第52回 年次大会研究発表会	2022年11月			