DEIM Forum 2021 E11-4

メタ特徴を用いた分類可能性予測

Predicting Classificatability Using Meta-Features

早川　雄登†　　新美　礼彦††

† 公立はこだて未来大学大学院 システム情報科学研究科　〒 041–8655 北海道函館市亀田中野町 116–2
†† 公立はこだて未来大学 システム情報科学部　〒 041–8655 北海道函館市亀田中野町 116–2
E-mail: †{g2120036,niimi}@fun.ac.jp

あらまし　近年，様々な分野でデータマイニングが活用されている．データ分析にはコストが生じることとなるが，そのデータを用いた分析で実際に有用なパタンを発見できるとは限らない．また，現時点ではデータ自体に分析する価値があるかどうかは，実際に分析を行う前に判断することは困難である．我々はこの問題の解決策のひとつとして，データ分析のうち分類タスクに着目し，分類タスク実行前にデータが目的変数の分類に寄与する情報を持っているかを確認する方法を検討している．そこで本稿では，データがある目的変数に対して分類タスクを実行するために必要な情報を持っているかを分類可能性という指標で定義し，メタ特徴を用いて同様のデータセットから分類可能性を予測する方法について論じる．

キーワード　データマイニング，機械学習，メタ学習

1 はじめに

近年，様々な領域でデータマイニングが応用されている．しかし，データマイニングを行った際に有用なパタンが得られるかどうかは分からないため，解析に必要となるコストに対して見合った結果が得られるかどうかは定かではないという問題が存在している．特に解析を望む主体がデータマイニングの専門家でない場合にこの問題が顕著に現れると考えられ，そもそもデータ自体に目的を達成するための情報が含まれているかがより判断し難い状況にあると言える．

また，近年のデータマイニング需要に対して Auto Machine Learning（AutoML）の研究・開発が盛んとなっている．いくつか例を挙げると，オープンソースのものであれば Auto-WEKA [1] や Auto-Sklearn [2]，企業が開発・提供しているものであれば Google AutoML Tables [3] や IBM AutoAI [4] などがある．これらの AutoML フレームワークを用いると，データマイニングの処理をある程度自動化することが可能となるが，やはりデータマイニングに関する知識が必要である．また，企業が開発・提供しているものに関しては利用するためのコストが必要となる．

我々はこれらの背景から，データに対しコストを投入する前に予めデータ自体が分析可能であるかを調べられるシステムが有用なのではないかと考えた．

そのため本稿では，いくつかの既存の分類器評価指標に着目して，データセットが特定の目的属性に対して分類タスクを実行するにあたって必要な情報を持っているかという期待度として分類可能性を定義し，それをデータセットから抽出したメタ特徴によって予測する方法について論じる．

2 先行研究

鳴海らは，データセットの性質から回帰分析により予測することが可能であるかという観点から，分類タスクにおいてデータセットが目的属性を予測するための情報を有しているかどうかをメタ特徴を用いて予測している [5]．その結果，複数のデータセットに対し決定木の一手法である CART と多層パーセプトロン（MLP）により構築した分類器の予測性能として F1 値に対する予測の決定係数 R^2 が 0.5 を上回り，良好な予測ができたと考えられる．

しかし，この研究ではいくつかの課題が残った．その中でも，メタ学習の目的属性に関するものとして，回帰によって予測された値の RMSE が目的属性の値域に対し大きいことと，どのようなアルゴリズムに対して応用することが可能かという検討を行う必要があるという 2 つの課題がある．

まず，予測結果の RMSE が値域に対し大きいという課題について考える．目的属性が取りうる本来の値域に対し RMSE が 15%程度生じており，この数値の利用方法次第では削減する必要が生じると考えられる．しかし，この研究の目的であったデータ自体に目的を達成するための情報が含まれているかを確かめるというタスクの上では，必ずしも厳密な実数値として予測する必要はないとも考えられる．

次に，応用可能な他のアルゴリズムを確かめるという課題について考える．この研究では回帰の目的属性として CART の Accuracy と F1 値，MLP の Accuracy と F1 値という 4 種類を用意し実験を行った．そして，CART と MLP それぞれによって分類を行った際の「データセット毎の F1 値の絶対差」とそれぞれの方法に対する「メタ学習器による F1 値の予測結果の絶対誤差」で相関二乗を取った結果から，CART と MLP

の分類アルゴリズムの違いに起因する予測性能の差による影響を受けていないと判断した．しかし，実際に他のアルゴリズムでこの方法が有用であるかどうかが確かめられていない．また全てのアルゴリズムについて実験により網羅することが必要であるかという疑問も残されている．

これら 2 つの課題は，前者は実数値として F1 値や Accuracy といった分類性能を予測していたこと，後者は特定のアルゴリズムによって構築された分類器に対する値を求めていたことに起因していると考えられる．従って，この方法を用いるのではなく，そのデータセットが目的属性に対して分類を行うための情報を持っているかという指標を定義し予測することにより解決できるのではないかと考えられる．

3 関 連 手 法

本稿ではデータセットの分類タスクに対する評価指標として分類可能性を定義するが，データセットや分類器に対する評価指標は様々なものが存在している．ここではその中でも，データセットの評価や分析としての統計モデリングと分類器評価指標について紹介する．

3.1 統計モデリング

データの性質を調べ評価するという観点では，データが正規分布に従っていると仮定して平均値や分散などの統計値を求める方法がある．しかし，全てのデータが正規分布に従っていると仮定することは困難であるため，他の確率分布やその組み合わせを用いて推定を行うことがある．このようにデータが何かしらの確率分布に従っていると想定しパラメタを推定する一連の試みは統計モデリングと呼ばれる．

しかし，統計モデリングを用いて各特徴量の確率分布やパラメタを求めることから，直接的に目的属性が説明属性によって説明しうるかという部分について考えるためには，統計的な学習と予測が必要となると考えられる．そのため，今回対象としている説明属性の潜在的な分類タスクに対する適性を計るためにはそのまま用いることは難しいと言える．

3.2 分類器評価指標

既存の分類器評価指標として，モデルのデータへの適合をそのまま評価するものとモデル選択基準を用いる方法がある．

まず，モデルのデータへの適合を評価する代表的なものとして，Accuracy や Recall，Precision，それらを組み合わせた F 値（特に F1 値）のほか，LogLoss や AUC(Area Under the Curve) といったものが挙げられる．これらは主に分類タスクで構築されたモデルに対しどの程度望ましい分類が行われているかを判断する指標として用いられている．

次に，モデル選択基準の代表的なものとして，赤池情報量基準（AIC）やベイズ情報量基準（BIC），最小記述長（MDL）が挙げられる．これらはモデルの複雑さとデータ適合度とのバランスを測る指標であり，なるべく少ないパラメタでより良い適合度を出すモデルを選択するために用いられる．

これらの評価指標は特定のモデルやそのパラメタとデータの相性を判断する指標にすぎないと考えられる．そのため，検証に用いたある分類器構築アルゴリズムを用いて上手く分析できた場合にはこれらの値が良好となると考えられる．しかし，それらの組み合わせ方次第では分類可能性として用いることができるのではないかとも考えられる．

4 分類可能性の定義

前述の各関連手法を踏まえて，我々はデータ自体が持っている潜在的な分類タスクに対する期待度というものが存在しているとして，その期待度を用いてデータを判別することが可能ならば，データに対し簡易的な分析を行うことで，そのデータに詳細分析をする価値があるかどうかの判断ができるのではないかと考えた．ここで，この期待度を分類可能性として定義する．

4.1 定義に用いるアルゴリズム

まず，一般的な分類タスクで用いられる分類器構築アルゴリズムについて考える．ここでは，Tavasoli による記事 [6] を基にアルゴリズムを選出する．この記事ではよく用いられている分類器構築アルゴリズムとして以下の 8 手法が挙げられている．

- Logistic Regression
- Decision Tree
- Support Vector Machine
- Naive Bayes
- K-Nearest Neighbors
- Random Forest
- Gradient Boost
- AdaBoost

これらのアルゴリズムはいずれも一般的な分類タスクに用いられているため，今回はこれら 8 つの分類器構築アルゴリズムを用いて分類可能性の定義を行う．

4.2 定義に用いる評価指標

次に，これらのアルゴリズムにより構築された分類器の評価について考える．前述のアルゴリズムはいずれも異なる方策による分類を行うため，これらのアルゴリズムで構築された分類器間で予測性能に差異が見られると考えられる．個別の分類器による適合率評価の比較によって分類タスクに対するそのデータセットの期待度は比較できると考えられる．今回は 8 つのアルゴリズムにより構築された分類器について，以下の 5 種類の各評価指標の最大値を用いて分類可能性の定義を行う．

- DACR (Default Accuracy)
- DPPV (Default Positive Predictive Value)
- DNPV (Default Negative Predictive Value)

- DTPR (Default True Positive Rate)
- DTNR (Default True Negative Rate)

これらの指標は Confusion Matrix ベースのものとして一般的であり，Seliya らはこれらの指標を含む 22 の性能指標を比較している [7]．結論において，評価指標の組み合わせの一例として，AUC (ROC–AUC)，BRI (Brier Inaccuracy)，KACR (Kolmogorov-Smirnov Statistics Accuracy) が挙げられている．

本研究では Seliya らの論文を基に分類可能性定義の基となる評価指標を選択したが，これらの選択はこの論文の結論に則っていないと思われる．しかし，これらの評価指標を用いたのには以下の 2 つの理由がある．

まず，今回用いた指標は前述の通り全て Confusion Matrix ベースのものである．Confusion Matrix はデータセットに対する分類器の適合を正例，負例に対する正解数としてカウントしている．そのため発想として簡潔でありかつ，重要であると言える．そして今回用いた 5 つの指標はそれぞれ Confusion Matrix を要約した指標であるとも言える．

以下に各指標の定義を示す．Confusion Matrix における各数値を，TP (True Positive, 真陽性)，TN (True Negative, 真陰性)，FP (False Positive, 偽陽性)，FN (False Negative, 偽陰性) とすると，

$$DACR = \frac{TP + TN}{TP + TN + FP + FN},$$

$$DPPV = \frac{TP}{TP + FP}, \quad DNPV = \frac{TN}{TN + FN},$$

$$DTPR = \frac{TP}{TP + FN}, \quad DTNR = \frac{TN}{TN + FP}.$$

ここで各指標の D (Default) は分類器の決定閾値 t が 0.5 であることを表す．式で示されている通り，DACR は総数に対する正解，それ以外は TP と TN の FP と FN との和に対する比率を表す．そのため，これら全ての値を考慮すると，結果的に分類器の直接的な性能評価指標となる．よって，1 つのデータに対し複数の分類器を構築した際に直接比較することが可能であり，複数のデータに対して同条件で分類器を構築し比較を行った場合はよりその分類器構築法に適合したデータであると考えることができる．これが 1 つ目の理由である．

次に，分類器の決定閾値 t を 0.5 としたことについて述べる．決定閾値を変化させた結果最も良い結果をそれぞれ BACR のように B (Best) を冠して示した場合，0.5 を用いた場合と比較すると，どの指標についても $BACR \geqq DACR$ が成り立つと言える．また，Kolmogorov-Smirnov 統計量を用いて閾値を決定したものについても同様に 0.5 を用いた場合以上の値となると言える．これは，Kolmogorov-Smirnov 検定を用いて 2 つの分布関数間の距離が最大になる閾値を採用するためであり，結果的にその分類器の性能は向上することとなる．

これらは分析対象データの価値を過剰評価に繋がることが考えられ，本研究のユースケースであるデータ分析を行うためのコストを投じる価値判断に用いるという観点から考えると，対象データに関して再考するという意味で適当でないと言える．これが 2 つ目の理由である．

4.3 定　　義

本稿で提案する “分類可能性” は，前述の 8 つの分類器構築アルゴリズムを用いて構築した分類器に対し 5 つの評価指標の各最大値が全て 0.75 を超えるか否かである．

ここで 0.75 という値について述べる．これは，一般的にこの値を超えていることが望ましいとされている値が定義できないため次の理由から決定した．上述の通りそれぞれの指標が Confusion Matrix 上の各値同士の比率となっていることから考えることができ，0.75 を用いると TP,TN が FP,FN に対し 3 倍以上となるような分類が可能となる．もちろん，最も理想的な分類器はいずれの指標も最大値である 1.0 であるが，実際に構築に用いるデータおよび今後観測されうるデータに対して 1.0 を実現するのは非常に困難である．そのため，ある程度の分類精度でありかつ現実的な値の一例として 0.75 を採用した．

また，各指標について 8 つの分類器構築アルゴリズムを用いて構築した際の最大値を採用したことについて述べる．これは，前述の通り，ある特定のアルゴリズムによって構築された分類器に対する分類性能ではなく，様々な方策によって分類される際に期待される分類性能を求めたいというのが理由である．

5 分類可能性予測手法の検討

ここまでで，データセットの潜在的な分類タスクに対する期待度の指標として分類可能性を定義した．この指標の利用方法としては，第 1 節において述べた通りデータセットのメタ特徴から予測することによって，データセットに対し分類器構築およびその評価を行うことなくそのデータセットが分類タスクを実行するに当たって必要な情報を持っているかどうかを簡易的に判断できるのではないかと考えられる．よって，ここではその方法について述べる．

本稿では，主に Hutter らの書籍 [8] を参考にメタ特徴の候補を選定した．実際に用いたメタ特徴は以下の通りである．ここで，統計的特徴量に関しては数値属性のみを対象として計算する．ここで数値属性としたものは，Stevens の尺度分類 [9] のうち数値尺度もしくは比率尺度を持つものであり，順序尺度や名義尺度のものは対象外とした．また，情報理論的特徴量においては，数値属性に関して最小記述長原理に基づく離散化 [10] を行ったものについて情報量を計算した．特筆なく平均値と記述しているものは全て算術平均である．

- 基本特徴量
  - （1） インスタンス数
  - （2） データセット次元性 [11], [12]
- 統計的特徴量
  - （3） 各属性の歪度の平均値
  - （4） 各属性の尖度の平均値

（5） 各属性間の相関行列の平均値
（6） 各属性内のクラス毎の重心距離の平均値 [13]
（7） 各属性間における ANOVA 検定の P 値の平均値
（8） 第一主成分の分散
（9） 第一主成分の歪度
（10） 元データの 95%を再現する主成分数

- 情報理論的特徴量
  （11） 目的属性のエントロピー
  （12） 最大エントロピーで正規化された各属性の情報エントロピーの平均値 [14]
  （13） 目的属性に対する相互情報量の平均値
  （14） 目的属性に対する相互情報量の最大値
  （15） 等価特徴数 [11]
  （16） 雑音信号比 [15]
- 複雑特徴量
  （17） Fisher 判別比の最大値 [16]
  （18） 主要 2 クラスにおける Volume of Overlap [16]

これらのメタ特徴を用いて前述の分類可能性を予測するために分類問題として考える．

6 実　　験

本実験では人工データを生成し，それらについて分類可能性とメタ特徴を求め，メタ学習を行う．それにより分類可能性の予測が可能であるかを確かめる．

また，分類可能性の定義において閾値を 0.75 としたが，複数の閾値について結果を比較し，考察する．

6.1 実 験 手 順

実験は以下の手順で実施する．なお，人工データの生成については付録にて述べる．

（1） 人工データ $\mathcal{D}$ を生成する．

（2） $\mathcal{D}$ のうち 9 割の事例を用いて構築した分類器 $\mathcal{M}$ を作成する．

（3） $\mathcal{M}$ に対して構築する際に用いなかった事例を用いて予測を行ったモデル適合度を求める．

（4） 分類可能性と $\mathcal{D}$ のメタ特徴セットを合わせてメタデータセットの 1 事例として保存する．

（5） (1) から (4) の手順を繰り返しメタデータセットを作成する．

（6） (3) で求めたモデル適合度から各事例に対する分類可能性を計算し目的属性とする．

（7） 作成したメタデータセットを用いて，分類可能性を予測するメタ学習器を構築・評価する．

表 1　Over–Sampling 前の各閾値における予測性能

閾値	負例	正例	正例率	Accuracy	F1(正)	F1(負)
0.60	4743	4265	47.3%	0.510	0.447	0.560
0.70	6718	2290	25.4%	0.698	0.150	0.816
0.75	7715	1293	14.4%	0.829	0.073	0.906
0.80	8299	709	7.8%	0.905	0.029	0.950
0.90	8723	285	3.2%	0.962	0.000	0.980

表 2　Over–Sampling 後の各閾値における予測性能

閾値	負例	正例	Accuracy	F1(正)	F1(負)
0.60	4743	4743	0.557	0.551	0.563
0.70	6718	6718	0.893	0.898	0.888
0.75	7715	7715	0.959	0.898	0.958
0.80	8299	8299	0.979	0.979	0.978
0.90	8723	8723	0.992	0.992	0.992

また，メタ学習器の構築においては，Random Forest により行った．メタ学習器の評価では分割数 10–Fold 交叉検証を行う．実験は scikit-learn [17] を用いて行い，Random Forest のパラメタは scikit-learn の標準パラメタを用いた．また，分類可能性計算のための各アルゴリズムに対するパラメタについては付録にて述べる．

6.2 実 験 結 果

まず，分類可能性の計算において複数の閾値を用いて 2 値化した際の目的属性内の割合，Accuracy，正例負例両方の F1 値を表 1 に示す．ここで負例の F1 値は NPV と TNR の調和平均とした．

次に，各閾値により目的属性を 2 値化したメタデータの minority クラスを Random Over–Sampling し再試行した結果を表 2 に示す．Over–Sampling は，今回生成した人工データ群から抽出した分類可能性が特に高閾値になるに従って予測を行いたい正例の予測性能の低下していることが，分類可能性の不均衡によるものでないかと考えられるため行った．

7 考　　察

本稿で実験に用いた人工データに対する予測結果を元に，本稿における分類可能性の定義と，本稿で提案するメタ学習を用いた分類可能性の予測手法について議論する．

まず，閾値と目的属性内のクラス割合，モデル適合率について述べる．人工データを用いたメタデータにおいて，Over–Sampling の有無に関わらず Accuracy は閾値の上昇と共に増加している．これは，5 つの評価指標全てで高い値を示すデータに特有の特徴があり，それをメタデータによってキャプチャできていることが考えられる．表 3 に閾値 $\{0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95\}$ における寄与度上位 3 位のメタ特徴を示す．ここで各属性名は，“MI max” は相互情報量の最大値，“MI avg” は相互情報量の平均値，“NSR” は雑音信号比，“EqFeats” は等価特徴数，“corr avg” は相関行列の平均値，“skew avg” は歪度の平均値である．

多くの閾値において，サンプリングの有無に関わらず相互情

表 3　各閾値における高寄与度メタ特徴

閾値	サンプリングなし			サンプリングあり		
	1 位	2 位	3 位	1 位	2 位	3 位
0.55	MI max	NSR	EqFeats	MI max	NSR	EqFeats
0.60	MI max	MI avg	NSR	MI max	EqFeats	NSR
0.65	MI max	EqFeats	NSR	MI max	EqFeats	NSR
0.70	MI max	NSR	EqFeats	MI max	NSR	EqFeats
0.75	MI max	NSR	MI avg	MI max	EqFeats	NSR
0.80	MI max	EqFeats	NSR	MI max	EqFeats	NSR
0.85	MI max	EqFeats	NSR	MI max	EqFeats	corr avg
0.90	EqFeats	NSR	MI max	corr avg	EqFeats	skew avg
0.95	MI max	EqFeats	MI avg	corr avg	EqFeats	NSR

表 4　Under–Sampling 後の各閾値における予測性能

閾値	負例	正例	Accuracy	F1(正)	F1(負)
0.60	4265	4265	0.504	0.502	0.505
0.70	2290	2290	0.491	0.485	0.498
0.75	1293	1293	0.486	0.479	0.494
0.80	709	709	0.499	0.496	0.501
0.90	285	285	0.504	0.494	0.513

報量，雑音信号比，等価特徴数が重要なメタ特徴となっている．これらは目的属性に対する各属性の情報エントロピーに基づいた値であり，特に評価指標で高い値を示すデータは綺麗に分離できるデータと言える．また，Over–Sampling している場合の 0.85，0.90，0.95 において相関行列の平均値や歪度の平均値といった統計量が高い寄与度を示しているが，3%程度の正例を負例に合わせて過剰に増加させているため，適切な分類ができていないと考えられる．

Over–Sampling の可否については，行っていない場合に正例の F1 値が非常に低く，行った場合に劇的に増加していることから，低い分類性能がクラスの不均衡によって生じていると考えられる．しかし，前述の通り majority クラスとの比率の差が大きい場合には適切な分類ができているとは言い難いと考えられるため，適切な閾値の範囲においては Over–Sampling を用いるべきであると言える．また，Random Under–Sampling も試行したが，特に不均衡となる高閾値について正例の F1 値を除く分類性能が低下した (表 4)．これは majority クラスに属す事例 (対象がメタデータであるため，事例は各人工データセット) を無作為に抽出した結果，データの性質が損なわれたことに起因すると考えられる．

閾値の決定については，Over–Sampling を行っている場合に関しては，閾値を上昇させることによりモデルの評価指標が増加している．しかし，過剰な Over–Sampling による問題に加えて，実際のデータに対して使う際に高い閾値による分類可能性予測が必要であるかを含めて検討する必要がある．

8 おわりに

本稿では，データセットが特定の目的属性に対して分類タスクを実行するために必要な情報を持っている期待度として分類可能性を定義し，それをメタ特徴によって予測する方法について論じた．分類可能性の定義とその予測に関する実験の結果，8 種類の分類器構築アルゴリズムによる分類性能として 5 つの評価指標を用いて定義した分類可能性を，提案したメタ特徴によって予測が可能であることが示唆された．

しかし，現状では本稿の実験に対して生成された人工データに関する実験しか行っておらず，この人工データが様々な実データを模倣しているものとは限らない．また，実際のデータに対して実用可能な性能を示し，目的に対し利用意義のある閾値を設定する必要がある．

そのため今後の方針として，この方法を実データに対して用いることの可否について検証するとともに，今回人工データで模倣できなかったデータへの対応策を検討する必要があると考えられる．

付　　録

人工データ生成

本実験で用いる人工データは以下の手順で生成した．

（1） N 個の事例を持つ M 個の属性 $\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_M\}, \boldsymbol{x}_i = \{x_{i1}, x_{i2}, \cdots, x_{iN}\}$ を互いに独立した標準正規分布で生成する

（2） M 個のうち L 個の属性を，標準正規分布から得られたそれぞれの属性で独立した閾値を用いて $\{0, 1\}$ の二値属性とする．

（3） $\boldsymbol{X'}$ の各属性の線形結合からなる系列を閾値 0 で離散化した $\boldsymbol{y} = \{y_1, y_2, \cdots, y_N\}$ を生成する．

（4） $\boldsymbol{X}$ から任意の属性・事例を削除し，$\hat{\boldsymbol{X}}$ を生成する．

（5） $\mathcal{D} = \{\hat{\boldsymbol{X}}, \boldsymbol{y}\}$ なる $\mathcal{D}$ を人工データセットとする．

ここで，(3) における線形結合の各属性に係る重みは，以下の 5 つの確率分布によって生成した．

- 標準正規分布
- ロジスティック分布（$\mu = 0, \sigma^2 = 1$）
- コーシー分布（$x_0 = 0, \gamma = 1$）
- ラプラス分布（$\mu = 0, b = 1$）
- 一様分布 $\mathcal{U}(-1, 1)$

また，(4) における属性・事例の削除について述べる．まず属性は，各属性について 50% の確率で選択を行った．次に事例は，割合 $\rho \sim \mathcal{N}(0.5, 0.2)$ を生成し，$\rho < 0$ の場合 $\rho = \rho + 1$，$\rho > 1$ の場合 $\rho = \rho - 1$ とした上で，事例数 N に対し ρ の割合でランダムサンプリングを行った．

各分類器構築のパラメタ

本稿で提案する分類可能性では，8 つの分類器構築アルゴリズムを用いる．その際の各アルゴリズムに対するパラメタは全データセットに対し固定であり以下の通りである．

- Logistic Regression
  - 正則化ペナルティ: L2 (係数 C:1.0)
  - 最適化問題解法: L-BFGS 法
- Decision Tree (CART)
  - 最大深度: 指定なし
- Suport Vector Machine
  - カーネル関数: RBF(係数 γ:$\frac{1}{\text{特徴数}\cdot\text{分散}}$)
  - 正則化ペナルティ: L2(係数 C:1.0)
- Naive Bayes (Gaussian)
  - 事前確率: 一様
- K-Nearest Neighbors (5-NN)
  - 重み: 一様
  - 距離計算: ユークリッド距離
- Random Forest
  - 各木の最大深度: 指定なし
  - 推定器の数: 100
- Gradient Boost
  - 各木の最大深度: 3
  - 推定器の数: 100
  - 誤差判定基準: Friedman MSE
  - サブサンプルサイズ (割合): 1.0
- AdaBoost
  - 基底推定器: CART
  - 推定器の数: 50
  - ブースティングアルゴリズム: SAMME.R

文　　献

[1] L. Kotthoff, C. Thornton, H.H. Hoos, F. Hutter, K. Leyton-Brown, "Auto-WEKA 2.0: Automatic Model Selection and Hyperparameter Optimization in WEKA", Journal of Machine Learning Research, Vol. 18, No. 1, pp. 1–5, 2017.

[2] M. Feurer, K. Eggensperger, S. Falkner, M. Lindauer, F. Hutter, "Auto-Sklearn 2.0: The Next Generation", arXiv:2007.04074, 2020.

[3] "Auto ML Tables |Google Cloud", `https://cloud.google.com/automl-tables` (Accessed Feb. 9, 2021).

[4] "IBM Watson Studio – AutoAI |IBM", `https://www.ibm.com/cloud/watson-studio/autoai`, (Accessed Feb. 9, 2021).

[5] 鳴海雄登，新美礼彦，"データの性質を用いた分類性能予測に関する検討"，研究報告データベースシステム（DBS），Vol. 2020-DBS-171，No. 3，pp. 1–7, 2020.

[6] S. Tavasoli, "Top 10 Machine Learning Algorithms List [2021 Updated]", `https://www.simplilearn.com/10-algorithms-machine-learning-engineers-need-to-know-article` (Accessed Feb. 9, 2021).

[7] N. Seliya, M. Taghi, J.V. Hulse, "A Study on the Relationships of Classifier Performance Metrics", 2009 21st IEEE international conference on tools with artificial intelligence, pp.59–66, 2009.

[8] F. Hutter, L. Kotthoff, J. Vanschoren, "Automated Machine Learning", Springer, 2019.

[9] S.S. Stevens, "On the Theory of Scales of Measurement.", Science, Vol.103, No.2684, pp.677-680, AAAS, 1946.

[10] U.M. Fayyad, K.B. Irani, "Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning", International Joint Conferences on Artificial Intelligence, Vol. 13, pp. 1022–1027, 1993.

[11] A. Filchenkov, A. Pendryak, "Datasets Meta-Feature Description for Recommending Feature Selection Algorithm", 2015 Artificial Intelligence and Natural Language and Information Extraction, Social Media and Web Search FRUCT Conference (AINL-ISMW FRUCT), pp. 11–18, 2015.

[12] G. Wang, Q. Song, H. Sun, X. Zhang, B. Xu, Y. Zhou, "A Feature Subset Selection Algorithm Automatic Recommendation Method", Journal of Artificial Intelligence Research, Vol. 47, pp. 1–34, 2013.

[13] S. Ali, K.A. Smith, "On Learning Algorithm Selection for Classification", Applied Soft Computing, Vol. 6, No. 2, pp. 119–138, 2006.

[14] C. Castiello, G. Castellano, A.M. Fanelli, "Meta-data: Characterization of Input Features for Meta-Learning", International Conference on Modeling Decisions for Artificial Intelligence, pp. 457–468, 2005.

[15] D. Michie, D.J. Spiegelhalter, C.C. Taylor, J. Compbell, "Machine Learning, Neural and Statistical Classification", Ellis Horwood, 1995.

[16] T.K. Ho, M. Basu, "Complexity Measures of Supervised Classification Problems", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 3, pp. 289–300, 2002.

[17] "Scikit-Learn: Machine Learning in Python – Scikit-Learn 0.24.1 Documentation", `https://scikit-learn.org/stable/` (Accessed Feb. 9, 2021).