

# 深発月震源分類に適した機械学習法の研究

菊池 栞<sup>\*1</sup> 加藤 広大<sup>\*1</sup> 山田 竜平<sup>\*2</sup> 山本 幸生<sup>\*3</sup> 廣田 雅春<sup>\*4</sup>  
横山 昌平<sup>\*5</sup> 石川 博<sup>\*1</sup>

## Study of the suitable machine learning techniques for classification of deep moonquake sources

Shiori Kikuchi<sup>\*1</sup>, Kodai Kato<sup>\*1</sup>, Ryuhei Yamada<sup>\*2</sup>, Yukio Yamamoto<sup>\*3</sup>, Masaharu Hirota<sup>\*4</sup>,  
Yokoyama Syohei<sup>\*5</sup>, Hiroshi Ishikawa<sup>\*1</sup>

### Abstract

Deep moonquake occurs at depth of about 1000km, and it is most frequent lunar seismic event. Due to considerable noises and low amplitudes in deep moonquake waveforms, we have some difficulties to locate the sources when we use conventional method. In this paper, we have investigated suitable machine learning methods to classify the deep moonquake sources based on similarity among the waveforms. The machine learning-based method is more useful to reduce the computational time to classify the sources compared with the conventional method, and it has also advantage to apply the some types of effective features such as power spectral density for the classification. We compared performances of some machine learning methods in order to find the suitable method. Then, the deep events misclassified from some machine learning methods has been also investigated, and we have analyzed causes of misclassification (e.g., mislabel, outlier) for improvement of the classification. It can be expected that the suitable machine learning method will enable us to create a new moonquake event catalog and identify the unclassified events.

**Key Words:** Moonquakes, Machine Learning, Neural Network

### 概 要

深発月震とは、深さ約1,000kmで発生し、観測された数が多い月震である。これまで深発月震の震源の推定が行われてきたが、深発月震の波形にノイズが多く含まれることと、振幅が小さいために、これまでの手法では、震源の推定が困難な月震も多く存在している。そこで、我々は、波形の類似性に基づいて、月震の震源推定を行うための、最適な機械学習法について研究を行なった。震源の分類に機械学習を用いることで、これまでの推定とは異なる新たな特徴を用いた震源の推定や、推定を行う際の解析処理時間の削減が期待できる。本研究では、深発月震の震源を高精度に推定する手法を発見するため、複数の機械学習の推定の性能を比較し、検討する。その後、分類性能の向上のために、誤って推定されたイベントに着目し、複数の手法や特徴量による震源の推定結果の比較を行なった。その結果から、月震ラベルの間違いや、震源における外れ値などを視野に入れながら、誤って推定される原因の分析を行った。本研究で選択した最適な機械学習法による分類を行うことで、新しい月震イベントカタログの作成や、これまで未分類であったイベントの同定が可能となることが期待される。

doi: 10.20637/JAXA-RR-16-007/0005

\* 平成28年11月24日受付 (Received November 24, 2016)

<sup>\*1</sup> 首都大学東京大学院システムデザイン研究科 (Graduate School of System Design, Tokyo Metropolitan University)

<sup>\*2</sup> 国立天文台 RISE 月惑星探査検討室 (Research of Interior Structure and Evolution of solar system bodies, National Astronomical Observatory of Japan)

<sup>\*3</sup> 宇宙航空研究開発機構宇宙科学研究所 (Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency)

<sup>\*4</sup> 大分工業高等専門学校情報工学科 (National Institute of Technology, Oita College)

<sup>\*5</sup> 静岡大学情報学部 (Faculty of Informatics, Shizuoka University)

## 1 はじめに

NASA のアポロミッションによって、Apollo Lunar Surface Experiments Package (ALSEP) と呼ばれる観測装置群が設置された。ALSEP を用いた実験のひとつである Passive Seismic Experiment (PSE) は、ALSEP に含まれる地震計を用いて月面上で月震の連続観測を行う実験である。この実験では、5つの地震計が月面に設置されており、その内、1-4台の地震計を使用して約7年半に渡る月震の記録が取得された。これらのデータは、長周期地震計の3軸成分と短周期地震計の上下動成分で構成されている。

PSE によって得られた月震の観測データは、現在も解析が行われており、月震の発生原因、活動度、および月の内部構造の予測について多くの知見が得られている [10],[8]。月震は、発生深度や、発生要因をもとに、熱月震、人工衝突の信号、隕石衝突、浅発月震、および深発月震の5種類に分けられている。

深発月震は、深さ約 1,000km で発生する月震であり、これまでに観測された数が最も多い月震である。また、深発月震は、同一震源から周期的に発生し、同じ震源で発生した月震の波形<sup>\*6</sup> は類似していることがわかっている [7],[2]。観測された深発月震の一部には震源を表すラベルが付与されており、月震イベントカタログ<sup>\*7</sup> として公開されている [17]<sup>\*8</sup>。ラベルは、A1, A2, . . . と名付けられており、109のグループに分類されている。特に Apollo では、複数の点で観測されたイベントが少なく、波形の振幅も小さくノイズも多いため、カタログでは、波形の類似性を用いた震源推定とラベルの付与も行われている [9]。一方、これまでの推定手法では、そのノイズの多さなどから、既存の震源に分類することが困難な月震も多数存在する。この問題に対して、震源が未知の深発月震の震源推定には、従来の方法で見えなかった、震源ごとに共通する新たな特徴の発見や、新たな分類手法が必要となる。そこで、本研究では、深発月震源の新たな分類手法として、機械学習に着目する。特に、ラベルが付与されていない大量のデータに人手でラベルを付与することは、解析処理時間が膨大にかかってしまうが、機械学習を用いることで、それが自動化できると考えられる。そこで、我々は、月震の震源推定を多クラス分類問題と捉え、従来の分類結果を学習データとして、機械学習の1つである教師あり学習を行うことで、自動的に深発月震の震源を分類し、分類結果をラベルとして深発月震に付与することを目指す。クラス分類とは、幾つかの震源を選択し、その震源で発生したイベントを学習器に学習させ、震源が未知であるイベントを、選択された震源のうちどの震源に属するのが適切かを予測することである。

教師あり学習を深発月震に適用した研究に、後藤ら [16] の研究が挙げられる。後藤ら [16] の研究では、Support Vector Machine (SVM)[13] を深発月震分類に用いており、その有効性が示されている。しかし、SVM 以外での分類が行われていないため、SVM 以外に深発月震分類に適した手法が存在する可能性がある。そこで本研究では、代表的な教師ありの学習の手法を用いて深発月震の分類を行い、分類性能を比較する。比較する手法は、Neural Network[5], SVM, ロジスティック回帰 [3], アンサンブル学習から Random Forest[1] とブースティングの5つを用いた。SVM では Linear カーネルと、RBF カーネルを、ブースティングでは AdaBoost[4] を用いた。そして、手法ごとに分類性能を比較することで、深発月震に適した機械学習手法を検討する。また、分類性能の向上を目指すために、本研究では誤分類されたイベントに着目した。誤分類される原因には、手法や特徴量が、震源の特徴をうまく捉えることができていないことや、ラベルの間違いなどが考えられる。このラベル間違いが疑われるイベントをデータセットから除くことで、分類性能が向上すると考えられる。そこで、複数の手法や特徴量で、誤分類されるイベントを比較し、誤分類される原因について考察を行なった。手法は分類性能の高い手法を用いて、特徴量は分類性能の比較で用いた特徴量と、時系列の月震波形を用いた。

本論文の構成は次の通りである。2章では、機械学習手法の分類性能を比較する手法とその結果を述べる。3章では、誤分類となる確率の高いイベントの決定を行う手法と、結果を述べる。4章では、2章と3章の結

表 1 実験に用いた震源ごとの月震イベントの数

震源名	A1	A6	A7	A8	A9	A10	A18	A20	A21	A33
イベント数	173	32	29	39	62	35	36	48	26	28

果から得られた考察を述べる。5章では、本論文のまとめと今後の課題について述べる。

## 2 機械学習手法の分類性能の比較

本章では、月震源分類における機械学習手法の分類性能の比較を行う。

### 2.1 実験手法

本章では、実験に用いるデータセット、特徴量と、比較する代表的な教師あり学習を行う、機械学習の手法と、その評価方法について述べる。

#### 2.1.1 データセット

本論文ではイベントの数が特に多い深発月震のみを扱う<sup>\*9</sup>。月震データは、X軸、Y軸、Z軸の3方向の成分で記録されている。本研究では、長周期地震計のデータを用いて、X軸、Y軸、Z軸の月震の3成分をそれぞれLPX, LPY, LPZとして表記する。また、月震の波形は、多くのノイズを含むため、本研究では、前処理を適用した波形を用いて震源の分類を行う。前処理として、平均引き、トレンド引きを行った後、0.3-1.5Hzのバンドパスフィルタを適用し、スパイク除去処理を行った。1つの波形に対して、3成分ごとに前処理を行った後の、月震波形の例を図1に示す。図1において、横軸は時間を示し、縦軸は振幅を示す。図1より、1つのイベントにおいて、成分の違いによって波形が異なることがわかる。本研究では、筆者らが行なった先行研究[15]より、LPZを用いた場合に各分類手法の分類性能が高かったため、本研究でもLPZのデータを使用する。

本研究では、最も観測期間が長く、イベント数が多いApollo12号地点で観測された月震データのみを用いる。分類の対象とする震源は、従来手法でラベルが付与されているイベント数が20以上ある10の震源を用いた。各震源のイベント数を表1に示す。また、震源によってイベント数が異なるが、筆者らの先行研究[15]より、イベント数にばらつきがある場合でも分類性能が高かったため、イベント数を均衡にするなどの前処理は行わない。

本研究では、分類に使用するイベントの、継続する長さを、筆者らの先行研究[15]において分類性能が高く、かつデータ量が少なくなる15分とした。

#### 2.1.2 比較する手法

イベントの特徴を学習するために用いる、代表的な教師あり学習を行う、機械学習手法について述べる。

#### Neural Network

Neural Networkは、様々な研究者達によって構築された機械学習のアルゴリズムである[5]。2012年

<sup>\*6</sup> 以後、特に明示のない場合は、波形は月震の波形を表す。

<sup>\*7</sup> 以下、カタログとする。また、このカタログに合わせて、発生した、深発月震が強く疑われる個々の現象を、イベントと呼称する。

<sup>\*8</sup> DARTS as ISAS/JAXA <http://www.darts.isas.jaxa.jp>

<sup>\*9</sup> 以後、特に明示のない場合は、深発月震を月震と呼称する。

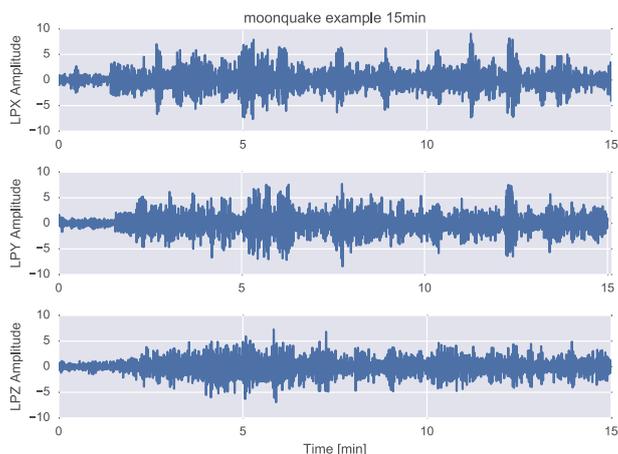


図1 月震波の例

の画像コンテスト<sup>\*10</sup>において、Hinton らがこの手法を用いて、その有用性に注目が集まった。本研究では、Neural Network の一種である、多層パーセプトロンを用いる。入力データを各ニューロンで、重みや活性化関数を用いながら値の変更を行っていく。そして、出力層での出力と、正解データとを比較し、その誤差を逆伝播させることで重みを更新し、学習を行う。多層パーセプトロンは、入力層と出力層の2層のみでは線形分類と等価であるが、中間層を追加すると、非線形分類が可能になる。本研究は、多クラス分類を行うため、出力層の誤差関数にはクロスエントロピーを用い、活性化関数には、ソフトマックス関数を用いた。ソフトマックス関数とは、出力の値の総和を1とすることで、分類されるクラスを確率分布にするための関数である。出力層のニューロン数は、本研究で行う分類のクラス数である10とした。

#### Random Forest

アンサンブル学習の一種である Random Forest は、2001年にLeo Breimanが提案した機械学習のアルゴリズムである[1]。Random Forest は、ブートストラップサンプルから複数の決定木で学習する。そして、それぞれの決定木による分類結果を統合し、最終的な分類結果とする。この手法の特徴として、学習が高速であることや、データの欠損や偏りがある場合でも性能の低下が少ないことが挙げられる。

#### SVM

SVM は1995年頃にVladimir N. Vapnikが提案したパターン識別用の機械学習のアルゴリズムである[13]。与えられた学習データから識別境界面を設定し、識別境界面に最も近いデータ点の識別境界面からの距離を最大化し、未知データがどちらに分類されるかを判定する。この手法の特徴として、データの次元数が大きい場合でも分類の性能が高いことや、設定するパラメータが比較的少ないことが挙げられる。本研究では、線形分類を行うために用いられる、線形カーネルと、非線形分類を行うために用いられる、RBFカーネルをそれぞれ用いた。本論文では、それぞれをSVM-LinearとSVM-RBFと表記する。

#### ロジスティック回帰

ロジスティック回帰はDavid Coxが提案したアルゴリズムである[3]。ロジスティック回帰は、ロジスティック関数を用いて、与えられた説明変数を基に、ある要素がどのカテゴリに属するかを条件付き確

<sup>\*10</sup> ImageNet Large Scale Visual Recognition Challenge  
<http://image-net.org/challenges/LSVRC/2012/index>

率を予測するアルゴリズムで、分類モデルとして利用することができる。出力が 0 から 1 の間であること以外は線形回帰と同じ振る舞いをするが、ロジスティック回帰では閾値を設けてクラスの予想を行っている。

### AdaBoost

ブースティングの一種である AdaBoost は、Yoav Freund と Robert E. Schapire が提案した機械学習のアルゴリズムである [4]。ブースティングとは、学習データを用いて学習を行い、その学習結果を踏まえて逐次に重みの調整を繰り返すことで複数の学習結果を求め、その結果を組み合わせることで性能を向上させる手法である。そのブースティングの代表的な手法である AdaBoost は、分類精度が 50 % を少し超える程度の、弱い分類器を大量に用意し、その 1 つ 1 つに重みを付けることで強い分類器を作成し、分類を行う。本研究では弱い分類器として決定木を用いた。また、本研究ではマルチクラス分類を行うため、AdaBoost のアルゴリズムとして、SAMME.R[14] を用いた。この手法の特徴として、ノイズが多いデータや異常値に影響を受けてしまうものの、他のブースティング手法より比較的過学習が起きにくいことが挙げられる。

#### 2.1.3 評価指標

手法の分類性能の評価を行うために、交差検定を行なう。まず使用するデータを学習データと評価データに分割する。学習データを用いて学習器のモデルを作成し、そのモデルを用いて評価データの分類を行う。その分類結果を用いて、本研究では、評価の指標として、適合率と再現率の調和平均である F 値を用いる。適合率とは、分類の結果の中に、どれだけ正解が含まれているかの割合を指し、再現率とは、分類の結果として出てくるべきもののうち、実際正しく分類されたものの割合を指す。この適合率と再現率はトレードオフの関係のため、両方の値をバランスよく評価するために、調和平均を計算した F 値を求め、それを用いて分類性能の評価を行う。例として、イベントを分類した際の、A1 の適合率、再現率、F 値は以下の式で定義される。

$$\begin{aligned} \text{適合率} &= \frac{A1 \text{ と予測して正解した数}}{A1 \text{ と予測した数}} \\ \text{再現率} &= \frac{A1 \text{ と予測して正解した数}}{\text{全体の } A1 \text{ の数}} \\ \text{F 値} &= \frac{2 * \text{再現率} * \text{適合率}}{\text{再現率} + \text{適合率}} \end{aligned}$$

他の震源においても以上の式を用いて、F 値を計算する。

#### 2.1.4 本研究で用いる特徴量

本研究では、月震の波形を用いて Power Spectral Density (PSD) を計算したものを特徴量として用いる。PSD は、周波数成分毎に振幅の強度を計算したものであり、時系列データの時間相関分析などに用いられる。本研究では、Welch 法を用いて、PSD の推定を行なっている。その際に、FFT を用いてスペクトルの計算を行なっているため、2 の累乗となるように、サンプリング数を決定する。本実験では、筆者らの先行研究 [15] により、分類性能が最も高かった 2048 をサンプリング数とした。

## 2.2 実験結果

本章では、2.1 節で述べたそれぞれの手法を用いて、月震の分類を行い、その分類性能を評価する。

表2 Neural Network のハイパーパラメータ

ニューロン数 (一層目)	ニューロン数 (二層目)	活性化関数	最適化関数	Dropout
60	40	tanh	Adam	あり

表3 各手法における適合率, 再現率, F 値

手法名	適合率	再現率	F 値
Neural Network	0.97	0.97	0.96
Random Forest	0.79	0.69	0.66
SVM-Linear	0.92	0.92	0.92
SVM-RBF	0.85	0.84	0.84
ロジスティック回帰	0.87	0.87	0.87
AdaBoost	0.13	0.34	0.19

### 2.2.1 ハイパーパラメータの決定

本研究で月震の分類に用いる各手法には、それぞれハイパーパラメータという任意に決定できる値が存在する。分類性能を上げるためにはこれらのハイパーパラメータを調節し、データセットに最適な分類器を構築する必要がある。

Neural Network については予備実験を行い、ハイパーパラメータである中間層のニューロン数、層の数、活性化関数、最適化関数、Dropout の有無を決定する。本研究では中間層のニューロン数、層の数、活性化関数、最適化関数、Dropout の有無を変更し分類を行い、それぞれの分類結果について、F 値を求めることで適切なハイパーパラメータを決定する。はじめに、1つ目の中間層のニューロン数を決定する。ニューロン数を増加させながら分類を行い、ニューロン数の増加によって、分類結果の F 値の変動がほとんど起こらなくなる時点でのニューロン数を本研究で用いるパラメータとする。

中間層の数と、追加される中間層のニューロン数を決定する。中間層の層の数を増加させ、分類結果の F 値の変化を確認し、F 値が上昇した場合、層の数を増加させる。このとき、新たに追加した層について、ニューロン数を1つ目の層の場合と同様に決定する。これを、分類結果の F 値が上昇しなくなるまで繰り返す。

各層に適用する各種関数を決定する。活性化関数 sigmoid, tanh, ReLU のうち、分類を行なった際に最も F 値が高くなる関数を各層に適用する。

Neural Network においては、学習中にランダムに選んだ一部のニューロンの働きを停止させることによって、頑健な学習を実現する、Dropout と呼ばれる手法の有効性が示唆されている。Dropout を無効にした状態で分類を行なった結果、過学習を起こしていることが観測されたため、Neural Network の中間層には Dropout を適用した。

最後に、最適化関数を決定する。最適化関数として Adam[6], AdaGrad, AdaDelta, SGD のそれぞれを用いて分類を行い、分類結果の F 値が最も高くなる最適化関数を本研究で用いるパラメータとする。チューニングを行なった結果の、Neural Network に適用するパラメータを表2に示す。

Neural Network については、Python のモジュールである Chainer[12] を用いた。その他の手法については scikit-learn[11] を用いて、グリッドサーチによるハイパーパラメータの決定を行なった。

### 2.2.2 各手法の分類結果

2.2.1 項のチューニング後、月震のデータセットに対して、6つの手法を用いて分類を行い、それぞれの分類性能の比較を行う。その分類の結果得られた適合率、再現率、F 値を表3に示す。これらの値は、震源ごとに

計算したものを平均した値である。表 3 より、Neural Network を用いて分類を行なった場合に、適合率、再現率、F 値共に最も高くなった。また、F 値は 0.96 となり、月震を高精度に分類することができた。その次に SVM-Linear、ロジスティック回帰となり、線形分類を行う手法の分類性能が高い結果となった。逆にアンサンブル学習である Random Forest と AdaBoost は低い分類性能となった。

### 3 誤分類となる確率の高いイベントの決定

分類性能の向上を目指すため、本章では誤分類されたイベントに着目する。誤分類にはいくつかの原因が考えられるが、ラベルの間違いなどが疑われるイベントをデータセットから除くことで、分類性能の向上が見込まれる。本章では、誤分類となる確率の高いイベントを発見し、その誤分類の原因についての分析を述べる。

#### 3.1 実験手法

分類手法や、特徴量などを変化させた複数の分類法で誤分類されたイベントを比較し、誤分類される確率の高いイベントの発見を行う。機械学習の分類性能を算出する際、全体のデータを学習データと評価データに分割するが、全ての分割パターンで分類を行うことは膨大な計算コストがかかるため、今回は学習データと評価データの分割をランダムに変更したもので、複数回分類を繰り返し、誤分類されたイベントをカウントした。本研究では、計算時間が長時間にならない 1500 回の分類を行なった。そして、各イベントの誤分類された回数を、そのイベントが評価データとして用いられた回数で割ることで、そのイベントがどの程度の確率で誤分類されるかを計算する。これらの実験に加え、誤分類されたイベントが、どういった原因で誤分類されるかを確かめるための実験も行なった。これまでの具体的な流れを以下に示す。

1. 複数の手法を用いた分類を複数回行う。その中で、イベントごとに誤分類された回数をカウントし、確率を計算する。
2. 複数の特徴量を用いた分類を複数回行う。その中で、イベントごとに誤分類された回数をカウントし、確率を計算する。
3. 各震源におけるイベントの外れ値を検出し、誤分類される確率の高いイベントが、その外れ値でないかを確認する。
4. 1, 2 より、手法と特徴量ごとに誤分類される確率の高いイベントをランキング形式にし、ランキング同士で相関係数を計算する。

1 は、手法ごとに誤分類されたイベントを比較することで、手法の違いに関わらず誤分類されるイベントを発見する。2.1.2 項の中で、F 値が 0.80 を超えた手法を用いて分類を行い、誤分類イベントをカウントする。

2 では、特徴量ごとに誤分類されたイベントを比較することで、特徴量の違いに関わらず誤分類されるイベントを発見する。特徴量には、PSD の他に、月震波形のデータを用いた。月震波形は、PSD と同様に LPZ を用いて、イベントの継続する長さを 15 分とした。

3 は、誤分類の要因が外れ値でないかを確認する。外れ値の確認は、マハラノビス距離を用いて、外れ値の判定を行う。マハラノビス距離とは、統計学で用いられる距離を表す手法の 1 つで、多次元のデータが相関を持つ場合に使用される。平均が  $\mu = \{\mu_1, \mu_2, \mu_3, \dots, \mu_p\}^T$ 、共分散行列が  $\Sigma$  であるような多変数ベクトル  $x = \{x_1, x_2, x_3, \dots, x_p\}^T$  で表される行列に対するマハラノビス距離は、下記の式のように定義される。

$$D_M(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}$$

具体的には、 $\mu$  には誤分類イベントの属する震源のイベントの平均を入力し、 $x$  には誤分類イベントを入力する。このマハラノビス距離を用いて、誤分類イベントの距離を各イベントのものと比較し、その誤分類イベン

表4 各手法と各特微量における誤分類確率の平均（上位5件）

(NN: Neural Network, SVM-L: SVM-Linear, SVM-R: SVM-RBF, LR: ロジスティック回帰)

イベント	PSD				月震波形				平均
	NN	SVM-L	SVM-R	LR	NN	SVM-L	SVM-R	LR	
A18-29	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000
A21-12	0.454	0.814	1.000	0.978	1.000	1.000	1.000	1.000	0.906
A20-33	0.285	0.994	0.977	1.000	0.843	1.000	1.000	0.895	0.874
A10-33	0.986	0.994	1.000	1.000	0.837	0.820	0.550	0.693	0.860
A21-4	0.230	0.732	0.891	0.841	0.995	1.000	1.000	1.000	0.836

表5 各手法と各特微量における分類同士の相関係数 (NN: Neural Network, SVM-L: SVM-Linear, SVM-R: SVM-RBF, LR: ロジスティック回帰)

特微量	手法	PSD				月震波形			
		NN	SVM-L	SVM-R	LR	NN	SVM-L	SVM-R	LR
PSD	NN		0.700	0.620	-0.662	-0.281	0.008	0.062	-0.240
	SVM-L			0.722	-0.860	-0.257	-0.015	0.174	-0.214
	SVM-RBF				-0.791	-0.079	0.153	0.397	-0.077
	LR					-0.217	0.083	0.162	-0.205
月震波形	NN						0.724	0.127	-0.779
	SVM-L							-0.309	-0.451
	SVM-RBF								-0.144
	LR								

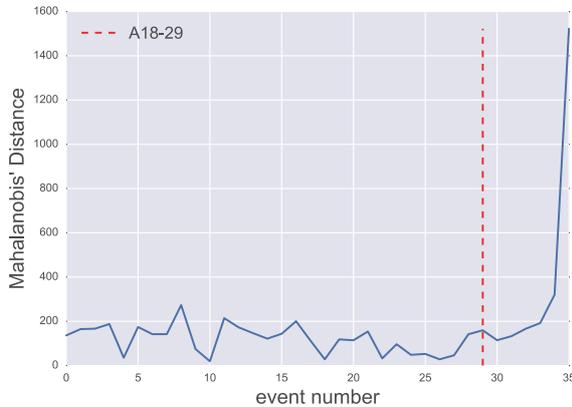


図2 A18における各イベントのマハラノビス距離

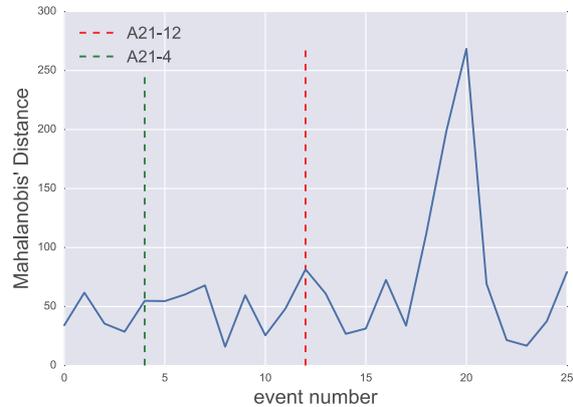


図3 A21における各イベントのマハラノビス距離

トが、属する震源の外れ値でないかを判定する。

4は、1と2による、特微量と手法ごとの誤分類確率による順位を用いて、各分類結果間の相関係数を計算する。この相関係数によって、どういった要因で誤分類されるイベントが異なるのかを考察する。相関係数には、ピアソン相関を用いた。

### 3.2 実験結果

月震分類における分類性能の高かった、Neural Network, SVM-Linear, SVM-RBF, ロジスティック回帰を用いて、誤分類イベントの決定を行う。特微量には、PSDと、PSDを計算していない時系列の月震波形を用いて分類を行なった。月震波形を用いた分類では、Neural Network, SVM-Linear, SVM-RBF, ロジス

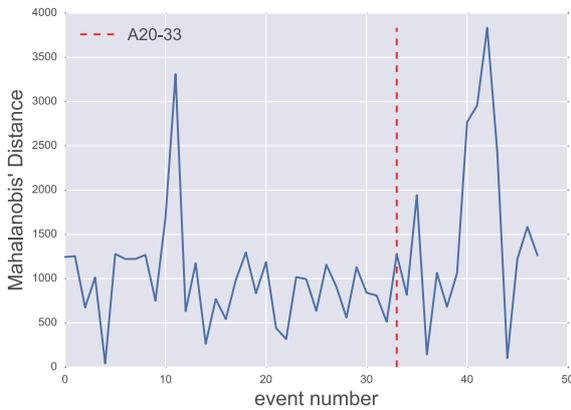


図4 A20における各イベントのマハラノビス距離

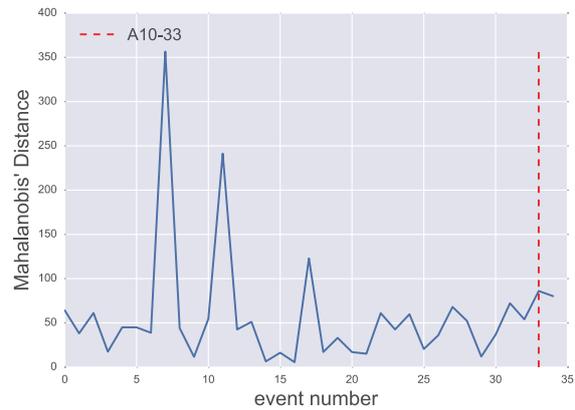


図5 A10における各イベントのマハラノビス距離

ティック回帰のどの手法においても F 値が 0.80 以上となり、PSD を用いた分類よりは低いものの、比較的高い F 値となった。これらの各手法と各特徴量を用いて、学習データと評価データの分割をランダムに行いながら 1500 回の分類を行い、各イベントの誤分類された確率の平均を求めた。その確率の高いイベントの上位 5 件の各特徴量と各手法による誤分類確率と、その平均を表 4 に示す。この順位の 6 位以下は、誤分類確率の平均は 0.8 以下となった。表 4 のイベントの番号は、震源の名前と、その震源ごとに観測された日時の早い方から番号を設定している。また、表 4 は、小数点第 3 位までを表示しているが、表中の誤分類確率が 1.000 のものは、1500 回の試行で一度も正しく分類されなかったことを意味している。

表 4 のイベントがその震源の外れ値でないかを確認するために、各震源における各イベントのマハラノビス距離を計算した。各震源における各イベントのマハラノビス距離を、図 2、図 3、図 4、図 5 に示し、表 4 のイベントをそれぞれの図中の縦線で表す。

PSD を用いて分類した誤分類イベントと、月震波形を用いて分類した誤分類イベントの誤分類確率による順位を用いて、各手法と各特徴量の相関を求めた結果を表 5 に表す。表 5 より、PSD を用いた分類では、相関係数が比較的高い結果となった。また、特徴量を変えた場合では、相関係数が低くなる傾向が見られた。

## 4 考察

本章では、2 章と 3 章の結果について考察について述べる。

### 4.1 機械学習手法の分類性能の比較の考察

表 3 より、Neural Network を用いて分類を行なった場合に、適合率、再現率、F 値共に最も高くなった。本研究で用いた Neural Network は中間層が存在し、非線形の分類を行うことのできるモデルで、その中間層において、適切に特徴量が抽出できたため、イベント数の少ない震源を含んだデータセットでも分類性能が高くなったと考えられる。

表 3 より、Neural Network の次に分類結果の F 値が高くなったのは、SVM-Linear を用いて分類を行なった場合であり、次に F 値が高くなったのは、ロジスティック回帰であった。SVM-Linear は、SVM で線形分類を行う場合である。同様に、ロジスティック回帰も線形分類を行う手法である。これらの手法は、非線形の分類を行う SVM-RBF よりも高い F 値を観測した。一般的に、次元数の高いデータにおける高次元空間では、データの分離に複雑な識別平面を要さない事が多いため、線形分類と非線形の分類性能に、あまり違いが

出ないことが知られている。しかし、線形分類を行う手法は、非線形分類を行う手法に比べて、境界面を決定する際のハイパーパラメータが少ないため、今回のような学習データが少ない場合に、分類性能が高くなったと考えられる。これらのことから、月震データは線形識別器でも高精度で分類可能であると考えられる。

表3より、Random ForestとAdaBoostは、他の4つの手法に比べ分類結果のF値の値が低くなった。決定木をベースとしたAdaBoostは決定木自体がデータ集合の細部に対して非常に敏感であり、過学習を起こしやすいことが知られている。また、決定木を組み合わせたRandom Forestは、学習データが少ない場合に過学習を起こしやすいことが知られている。これらの手法については、過学習が発生したため、分類性能が低くなったと考えられる。

これらより、Neural Networkが月震の分類において、分類性能が高い手法であることがわかった。

## 4.2 誤分類となる確率の高いイベントについての考察

表4より、誤分類される確率の高いイベントを示した。誤分類確率の平均が最も高いA18-29では、表4より、月震波形を用いたNeural Networkを除いた全ての分類方法で、正しく分類されることは一度もなかった。また、月震波形を用いたNeural Networkでも高い確率で誤分類していることから、A18-29は平均して0.9996という高い確率で誤分類された。また、図2より、他のイベントのマハラノビス距離と値が近いいため、外れ値が原因で誤分類が起きている可能性は低いと考えられる。よって、A18-29は手法、特徴量に関わらず、誤分類される確率が高いイベントであることが示唆された。このイベントが誤分類される理由としては、震源の主な特徴と類似しない波形であったため、または震源のラベルが間違っていることなどが挙げられる。

表4のA18-29以外のイベントについても考察する。図3、図4、図5より、A18-29と同様に、その震源における他のイベントのマハラノビス距離と値が近いいため、これらのイベントの誤分類は外れ値が原因で起きている可能性は低いと考えられる。A21-12は誤分類確率の平均が高かったものの、PSDを用いたNeural Networkでは、誤分類される確率が低かった。同様に、A20-33も特徴量に関わらず、誤分類される確率が高いが、PSDを用いたNeural Networkによる誤分類確率は低く、この分類法では誤分類されることが少なかった。よってA21-12とA20-33のイベントは、PSDを用いたNeural Networkによって正しく分類できる確率が高い。そのため、これらのイベントの誤分類確率が高い原因は、手法によるものだと考えられる。A10-33は、PSDを用いた分類では誤分類確率が1に近かったが、月震波形を用いた分類では誤分類確率は0.5から0.8程度となった。A21-4は、月震波形を用いた分類では誤分類確率が1に近かったが、PSDを用いた分類では誤分類確率が0.8程度となり、PSDを用いたNeural Networkでは、0.230となった。よって、A10-33とA21-4は、全体的に誤分類確率が高いものの、特徴量によって分類に成功する場合もある傾向となった。そのため、これらのイベントの誤分類確率が高い原因は、特徴量によるものだと考えられる。

表5では、PSDを用いた分類のそれぞれで相関が高い傾向となった。同様に、月震波形を用いた分類でも、PSDよりも少ないものの、同じ傾向が見られた。これにより、月震分類は、特徴が原因で誤分類が起こる可能性があることが示唆された。PSDを用いた場合と、月震波形を用いた場合の、手法ごとの相関を比較した場合では、あまり相関が見られなかった。これらのことより、月震分類において、学習のされ方と、どのイベントが誤分類されるかは、手法よりも特徴量に大きく依存すると考えられる。

## 5 まとめ

本論文では月震分類に適した分類手法を検討するため、機械学習の複数手法の分類性能の比較を行なった。その結果、Neural Networkによる分類の分類性能が、本研究で用いた手法の中では最も高くなった。この手法を用いる事で、ラベルが付与されていないイベントに高精度でラベルを付与できることが期待される。併せて、Neural Networkは分類のクラスが確率で出力されるため、確率付きの月震カタログを作成も可能である。

また、高頻度に誤分類されるイベントに着目し、その理由の究明を行い、震源の特徴に類似しないまたは、震源ラベルが間違っているイベントを示唆した。そして、分類性能は特徴量の種類に大きく起因していることも示唆された。今後の課題として、分類に適した特徴量の発見が挙げられる。これによって、さらに高精度な分類が可能になると考えられる。また、ノイズを多く含んだイベントを含めての実験も挙げられる。本実験はノイズの少ないイベントのみで実験を行なったため、ノイズを多く含んだイベントも含めることで、震源が未知のイベントの推定も可能になると考えられる。

## 謝辞

本稿の執筆にあたって多数の有益な助言を頂いた、首都大学東京特任助教の莊司 慶行氏に感謝いたします。本研究（の一部）は傾斜的研究（全学分）学長裁量枠戦略的研究プロジェクト戦略的研究支援枠「ソーシャルビッグデータの分析・応用のための学術基盤の研究」による。

## 参考文献

- [1] BREIMAN, L. Random forests. *Machine learning* 45, 1 (2001), 5–32.
- [2] BULOW, R. C., JOHNSON, C. L., BILLS, B. G., AND SHEARER, P. M. Temporal and spatial properties of some deep moonquake clusters. *Journal of Geophysical Research: Planets (1991–2012)* 112, E9 (2007).
- [3] COX, D. R. The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)* (1958), 215–242.
- [4] FREUND, Y., AND SCHAPIRE, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational learning theory* (1995), Springer, pp. 23–37.
- [5] HAYKIN, S., AND NETWORK, N. A comprehensive foundation. *Neural Networks* 2, 2004 (2004).
- [6] KINGMA, D., AND BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- [7] LAMMLEIN, D. R. Lunar seismicity and tectonics. *Physics of the Earth and Planetary Interiors* 14, 3 (1977), 224–273.
- [8] LOGNONNÉ, P., GAGNEPAIN-BEYNEIX, J., AND CHENET, H. A new seismic model of the moon: implications for structure, thermal evolution and formation of the moon. *Earth and Planetary Science Letters* 211, 1 (2003), 27–44.
- [9] NAKAMURA, Y., LATHAM, G., DORMAN, J., AND HARRIS, J. Passive seismic experiment long-period event catalog. *Galveston Geophysics Laboratory Contribution* 491 (1981), 314.
- [10] NAKAMURA, Y., LATHAM, G. V., AND DORMAN, H. J. Apollo lunar seismic experiment—final summary. *Journal of Geophysical Research: Solid Earth (1978–2012)* 87, S01 (1982), A117–A123.
- [11] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M., AND DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
- [12] TOKUI, S., OONO, K., HIDO, S., AND CLAYTON, J. Chainer: a next-generation open source framework for deep learning. In *LearningSys Workshop on Machine Learning Systems at Neural Information Processing Systems (NIPS)* (2015).

- [13] VAPNIK, V. N., AND VAPNIK, V. *Statistical learning theory*, vol. 1. Wiley New York, 1998.
- [14] ZHU, J., ZOU, H., ROSSET, S., AND HASTIE, T. Multi-class adaboost. *Statistics and its Interface* 2, 3 (2009), 349–360.
- [15] 菊池栞, 山田竜平, 山本幸生, 横山昌平, AND 石川博. 月震分類に適した機械学習手法の検討. **第8回データ工学と情報マネジメントに関するフォーラム**, E4-1 (2016).
- [16] 後藤康路, 山田竜平, 山本幸生, 横山昌平, AND 石川博. 波形の類似性を考慮した大規模月地震データの som による可視化システム. **第6回データ工学と情報マネジメントに関するフォーラム**, E1-4 (2014).
- [17] 山田竜平, 山本幸生, 桑村潤, AND 中村吉雄. アポロ月地震データ公開システムの開発. **宇宙科学情報解析論文誌 宇宙航空研究開発機構研究開発報告**, 1 (2012), 121–131.