

発火点決定ルールを調査する決定木作成の実行時間

林 亮子^{1,a)} 中田 侑江^{1,†1}

概要：発火点とは、可燃物を空気中で加熱すると自然発火する温度であり、工業的に重要な量である。発火点は主に実験的に調べられているが、危険を伴うために測定が困難な量である。また、発火は非正常な現象であるため、直接シミュレーションを用いて調べることは困難である。そこで、著者らは近年データマイニングを用いて発火点の予測を試みている。本稿では、データマイニングの手法の一つである決定木を用いて発火点の分類ルールを検討する際の実行時間を調査した結果を報告する。入力データは分子量、融点、沸点、炭素原子個数、酸素原子個数およびいくつかの官能基の有無、特徴的な結合の有無を含むものとし、統計プログラミング言語 R の決定木パッケージ rpart を用いる。240 件の実データを用いた実行時間はミリ秒スケールであり、十分短い実行時間である。次に、将来的に入力データが増えた場合の実行時間を予測するため、乱数データを組み合わせる手法と、既存データを繰り返し使う手法とを用いて擬似的にデータを増やして実行時間変化を調べた。その結果、今回調べた範囲では、実行時間はデータ量増加に対してはおおむね線形で増加し、さらに得られる決定木で分岐が生じるかどうか大きく依存することがわかった。

キーワード：データマイニング，決定木，R，性能評価

1. はじめに

近年ではデータマイニング技術 [1], [2], [3] が成熟し、誰でも簡単にデータマイニングを利用できるようになった。さらに、データマイニングの物質科学関連分野 [4], [5], [6], [7], [8] への応用が注目され、内外においてこの分野におけるプロジェクトがいくつか立ち上がっている。これまでに筆者らは統計プログラミング言語 R (以後、「R」と呼ぶものとする。) の決定木パッケージを用いて、実験およびシミュレーションが困難な量である発火点 [9] の予測を試験的に行った [10]。R はデータマイニングでよく用いられるものの一つであるが、性能については不明な点も多い。そこで本稿では、文献 [10] で使用している rpart パッケージの処理時間を調査した結果を報告する。

第 2 節以降の本稿の構成を述べる。第 2 節は類例研究を紹介し、本稿の位置づけを説明する。第 3 節は本稿で使用したデータと処理内容を説明し、本稿が用いる統計プログラミング言語 R と決定木を紹介する。第 4 節ではまず実データを用いて決定木を作成する際の処理時間を調査した結果を紹介する。次に、データを擬似的に増加させたとき

の処理時間を検討した結果を示す。第 5 節は本稿で得られた結果をまとめ、今後の課題を述べる。

2. 類例研究

化学分野においては、計算機の黎明期から計算機を利用して化学物質の性質を調べる分野が存在し、ケモインフィマティクス [6] と呼ばれている。ケモインフォマティクスでは、分子の性質を表す量を記述子と呼ぶ。記述子として何を用いると最適なものはまだ定説がなく、データ分析の目的によって必要な記述子を検討する必要がある。ケモインフォマティクスには定量構造活性相関と呼ばれる分野があり、分子の性質を示す情報から発火点を予測する試みがすでに行われている [7] [8]。

Tsai, Chen, Liaw は定量構造活性相関の手法を 820 件の分子データに用いて発火点予測を行った [7]。文献 [7] では 4 個の記述子とその線形結合式を用いて最大誤差 89K、平均誤差 36K で有機化合物の発火点を予測した。Tsai らは、同じ物質の分子力学計算結果に含まれる 3000 種類以上の記述子から、結果として同じ数値になるものを統合し、記述子相互の依存性が小さいものに限定したのちに記述子を 1 個ずつ加えて影響を検討するなどの作業を行って記述子を 4 個にしぼりこんでいる。そのため使用した記述子は、分子の具体的な構造を直接示すというよりは、理論的な量

¹ 金沢工業大学
KIT, 3-1 Yatsukahoh, Hakusan, Ishikawa 924-0838, Japan

^{†1} 現在、清水建設

^{a)} ryoko@neptune.kanazawa-it.ac.jp

である。

Shi, Chen, Chen は文字列長が最大 12 文字の分子フラグメント文字列を記述子に用いる手法により, 212 件の有機化合物データを使用して, 実際の発火点との誤差が 50K から 90K 程度で発火点を予測した [8]。Shi らの使用した記述子は, 分子を部分的に分割したフラグメントを使用するもので, 分子の具体的な構造に近い量である。

文献 [7] と文献 [8] は定量的に発火点を予測するという観点からは一定の成果を挙げている。そして記述子の影響を数値的に検討しているが, 使用した記述子が最適かどうかはまだ議論の余地がある。一方著者らの行った研究では決定木と呼ばれるデータマイニング手法を使用しており [10], 発火点を分類する過程で影響の大きい記述子から優先的に分岐条件に用いる。見方を変えると, 決定木を使用すると, 影響度が不明な大量の記述子をそのまま入力に用いても, 影響の大きい記述子の自動抽出ができる可能性がある。一方, データマイニングにおいては長い実行時間を要する場合も存在し, 使用するソフトウェアも十分最適化されていない場合もある。そのため, 本稿では今後の研究の展開に耐えられるかどうかという観点から, 現在使用しているソフトウェアの実行時間を調査する。

文献 [7] は, 得られた発火点予測精度の類例研究との比較を行っており, 類例研究で扱っている分子の件数が掲載されている。このような分野では, データを用意するプロセスが最も困難であり, 文献 [7] によると扱う分子の種類は 50 件程度が多く, 件数の多いもので数百件程度である。そのため, データマイニング応用分野の中では, 「ビッグデータ」というよりも, どちらかというデータ件数が少ない問題に分類される。本稿では 240 種類程度の分子を扱っており, 類例研究と比較して遜色ない量のデータは用意している。

3. 発火点を予測する決定木

3.1 本稿が行う処理内容

著者らが行った発火点予測の詳細は文献 [10] で報告したので, 本稿ではその要点のみを紹介する。本稿で使用するデータの中核部分は「国際化学物質安全性カード」[5]として公開されているものを使用する。データの概要は 3.2 で述べる。本稿では, 統計プログラミング言語 R[1], [2], [3](以後, 「R」と記す) を使用する。R はオープンソースであるため, これまで数多くのパッケージが開発され, 公開されている。特に近年はデータマイニング手法を実装したパッケージが数多く公開されている。本稿ではパッケージ rpart[1], [2] を用いて決定木を作成する。パッケージ rpart は決定木を作成する関数 rpart および関連するいくつかの関数を含む。R の処理内容は 3.3, 結果の概要を 3.4 で紹介する。

3.2 使用する実データの概要

本稿で使用する分子量, 融点, 沸点, 発火点データは, 「国際化学物質安全性カード」ウェブページ [5] から得た。このウェブページでは官能基がわかる構造式も掲載されており, 本稿で用いる分子は構造が簡単であるため, 文献 [4] も併用して構造式から炭素原子個数, 酸素原子個数, 官能基の個数, 各種結合の個数が得られる。

決定木関連分野では, 記述子は予測変数と呼ばれ, 予測対象は基準変数と呼ばれる。今回は「決定木を用いて発火点の決定ルールを調べる」という考え方であるため, 入力データには予測対象である発火点を含む。表 1 に決定木の作成で使用した化学物質データの一例を示す。表 1 に示すように, 本研究で使用する入力データは, 以下のものである。

分子の性質を表す連続値: 分子量, 沸点, 融点, 発火点。
分子の特徴的な原子個数 (0 または自然数): 炭素原子個数, 酸素原子個数。炭素原子は必ず存在するが, 酸素は存在しない場合があり, そのときは 0 とする。
特徴的な構造の個数 (0 または自然数): ベンゼン環, 二重結合, ヒドロキシル基, カルボキシル基, ケトン基, エステル結合, エーテル結合。なお, 存在しない場合は 0 とする。

以上のデータを各分子は欠損値なしで持つ。以後, 1 個の分子を 1 個の「オブザベーション」と呼ぶこととする。本稿では, 1 オブザベーションにつき連続値と離散値を含む 13 個のデータがある。そして, 基準変数は 1 個で発火点データ, その他の 12 個のデータを予測変数に用いる。今回使用する予測変数の一部は発火過程モデルを考慮して選択した [10] が, 発火点に影響することがあらかじめ予想できるのは沸点とベンゼン環個数であり, それ以外の予測変数が発火点に影響するかどうかは不明である。

本稿で扱う化学物質は, おおまかに以下のような分類が可能である。最後に記載した件数は, 本稿で決定木作成に

表 1 使用データ例

物質名	メタン CH_4	ベンゼン C_6H_6
オブザベーション番号	1	33
分子量	16.0	78.11
沸点	-161.49	80.1
融点	-182.48	5.5
炭素原子個数	1	6
酸素原子個数	0	0
ベンゼン環	0	1
炭素間二重結合	0	0
ヒドロキシル基	0	0
カルボキシル基	0	0
ケトン基	0	0
エステル結合	0	0
エーテル結合	0	0
発火点	537.0	498.0

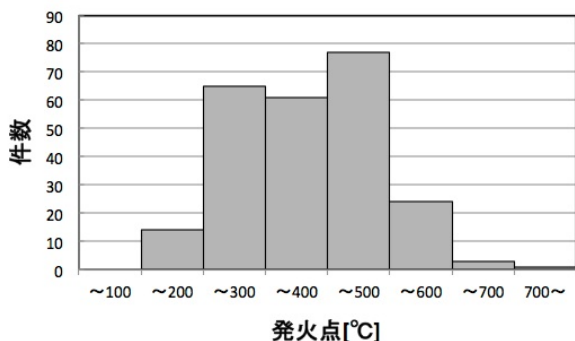


図 1 使用したオブザベーションにおける発火点の分布

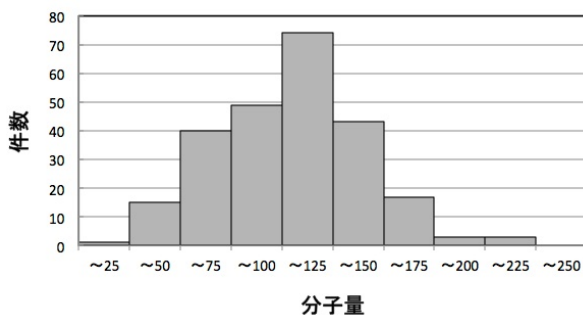


図 2 使用したオブザベーションにおける分子量の分布

使用したオブザベーション件数である。

鎖状飽和化合物 直鎖型で単結合のみで結合した化合物。129 件。

鎖状不飽和化合物 直鎖型で二重結合や三重結合を含む化合物。38 件。

環状飽和化合物 環状で単結合のみで結合した化合物。10 件。

環状不飽和化合物 環状で二重結合を含む化合物。10 件。

芳香族化合物 ベンゼン環を持つ化合物。41 件。

複素環状化合物 炭素と酸素（エーテル結合）から成る環状を持つ化合物。17 件。

これらのグループ間で分子量と発火点の関係を比較すると、芳香族化合物は他の化合物よりも発火点が高くなる傾向がある [10]。

今回用いたデータの性質を示すため、オブザベーションの発火点と分子量の分布を示す。まず発火点の分布を図 1 に示す。図 1 によると、最も温度の低い発火点でも 100 度以上であり、最高の発火点は 700 度台であった。ほとんどのオブザベーションは 200 度から 500 度の範囲にあり、類似研究の発火点分布とおおむね同様である。次に、分子量の分布を図 2 に示す。図 2 によると、今回使用した分子の分子量は最大 225 程度で、分子量 100 から 125 までの分子が最も多く 70 件程度あり、その階級から離れるにつれて、件数は減少する。

3.3 決定木パッケージ rpart の概要

本稿では、データマイニングの代表的な手法の一つである決定木を用いる。決定木とは、機械学習のアルゴリズム CART によって統計的な相関関係を分析し、結果を木構造で表した図である [1]。本稿では連続変数である発火点を予測するため、基準変数が連続変数の場合に用いられる回帰木を扱う。

回帰木においては以下のように分岐基準を選択する。回帰木では、平方和の分解を用いてデータを最大 2 つのグループに分類する。親ノード（回帰木の作成開始時は根ノードであり、全オブザベーションを含む）において、基準変数の偏差平方和 S は次式で得られる。

$$S = \sum_{i=1}^N (y_i - \bar{y})^2 \quad (1)$$

ここで、 N は親ノードが持つオブザベーション個数、 y は i 番目のオブザベーションが持つ基準変数値、 \bar{y} は親ノード内の基準変数の平均値である。次に、ある予測変数 T の特定の値を用いて 2 分岐する場合を考え、右子が N_R 個、左子が N_L 個のオブザベーションを含むものとする。このとき、右子の偏差平方和 S_R^T および左子の偏差平方和 S_L^T は次式で表せる。

$$S_R^T = \sum_{i=1}^{N_R} (y_i - \bar{y}_R)^2 \quad (2)$$

$$S_L^T = \sum_{i=1}^{N_L} (y_i - \bar{y}_L)^2 \quad (3)$$

ここで、 \bar{y}_R と \bar{y}_L は右子と左子それぞれのノード内での基準変数の平均値である。予測変数の値候補について、式 (1)-(式 (2)+式 (3)) に相当する

$$S - (S_R^T + S_L^T) \quad (4)$$

を計算し、式 (4) が最大となる予測変数を採用して分岐する。式 (1) すなわち S は、この分岐を検討する段階では定数であるため、 $S_R^T + S_L^T$ が最小となる予測変数 T の値が式 (4) を最大にする。「 T の値としてどの量のどの数値を用いるのか」は機械学習によって決定するため、得られた分岐ルールには、最適な量の最適な数値であることの保証はない。また、演算量の事前の推測は困難である。

次に、rpart パッケージの概要を紹介する。rpart パッケージは関連するいくつかの関数を含んでおり、決定木を作成する関数は rpart 関数である。rpart 関数の制御部分は R 言語を用いており、計算部分は C 言語で開発されたファイルといくつかのヘッダファイルから構成されている。決定木作成の中核部分プログラム rpart.c を用いて、分類条件の決定やノードの増加を行う。文献 [2] によると、関数 rpart は決定木作成時にデータセットをランダムに分割して交差検定も行い、デフォルトでは 10 分割交差確認を行

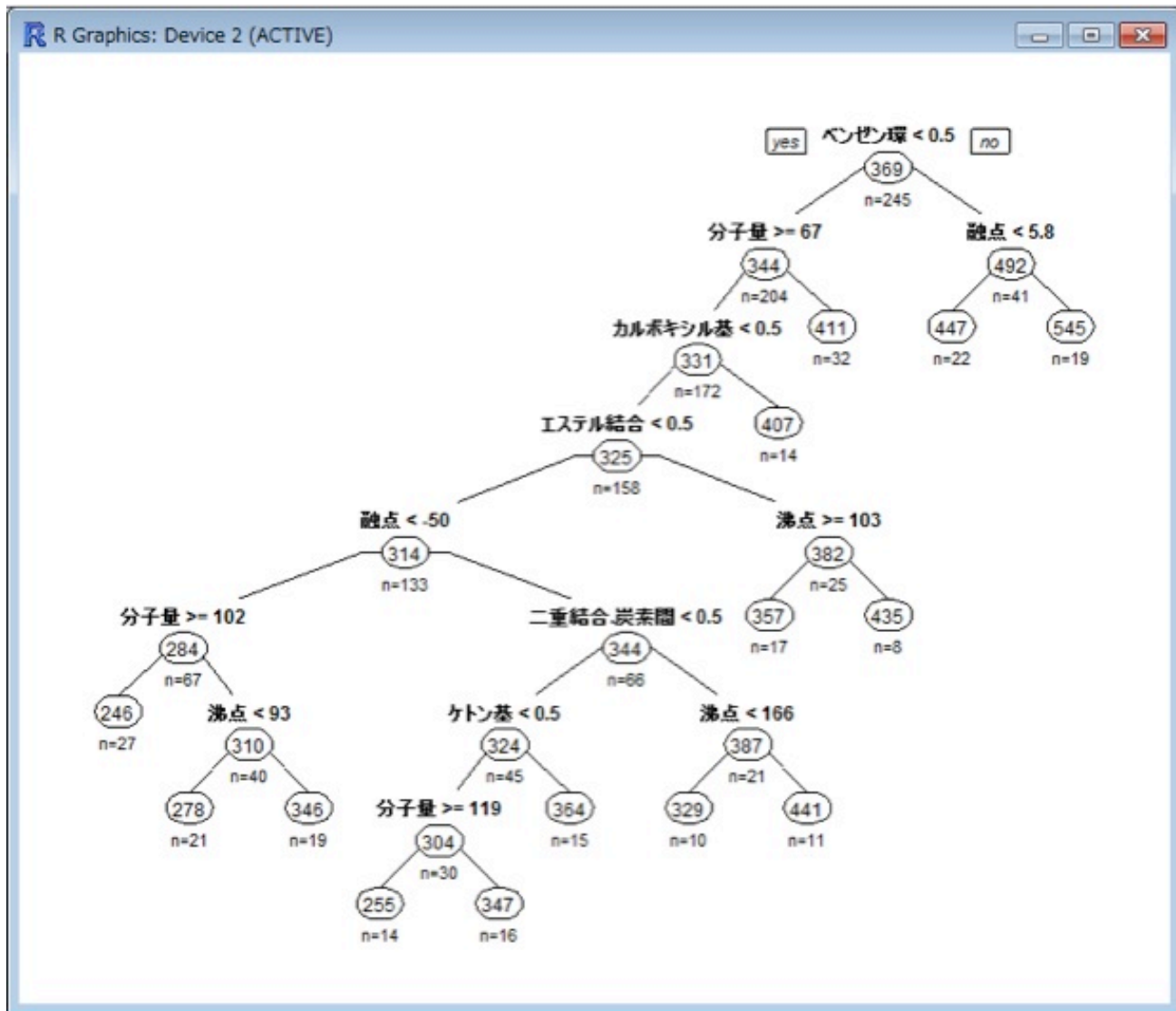


図 3 245 種類の分子における発火点の決定木

う。本稿では交差検定について明示的な設定は行っていないため、デフォルトの 10 分割交差確認を行うので、交差検定プロセスで処理時間が増加することが予想される。

3.4 決定木の作成例

次に、決定木の作成例を紹介する。245 件のオブザベーションを含むデータを関数 `rpart` に入力すると、図 3 のような決定木が得られる [10]。図 3 中の楕円は木のノードを表し、ノードに記入した数値は、そのノードに所属するオブザベーションの平均発火点を示す。丸の下の「n=」に続く数値は、そのノードに所属するオブザベーション個数である。各ノードの上に記入している式は、そこから分岐する際の分岐条件であり、分岐条件を満たすオブザベーションは次に左子へ、満たさないときは右子へ所属する。

図 3 は分岐が非常に多く、過学習が起こっている可能性があるが、その議論は文献 [10] に譲り、本稿では図 3 で得られたルールの主なものを紹介する。この木は過学習の傾

向があるため、根に近いルールのみに興味がある。ここでは根の分岐ルールのみを紹介する。

根の分岐ルール: ベンゼン環の有無で分岐し、ベンゼン環を持たない分子の平均発火点は 344 度 (根の左子の楕円内の数値)、ベンゼン環を持つ分子の平均発火点は 492 度 (根の右子の楕円内の数値) である。

3.2 で紹介したように、分子グループごとに分子量と発火点の関係を簡単に調べただけでも「ベンゼン環を持つ分子は他の種類の分子よりも発火点が高くなる」傾向がわかり、その傾向が重要なルールとして得られているため、図 3 の決定木はある程度適切に作成されているものと考えられる。その他のルールと図 3 の木を使用した発火点の予測については、文献 [10] を参照されたい。

4. 性能評価

4.1 実際の分子データを用いた性能評価

表 2 に本節で用いた計算機環境を示す。表 2 に示すよ

表 2 本稿で用いた計算機環境

機種名	Mac Pro
CPU	6-Core Intel Xeon, 3.06GHz, 2CPU, 二次キャッシュ(コア単位) 256KB 三次キャッシュ(プロセッサ単位) 12MB
メモリ	64GB 1333MHz DDR3
OS	Mac OS 10.8.5
R	Version 3.2.1(2015-06-18)
rpart パッケージ	rpart_4.1.10

うに, CPU はマルチコアの Intel Xeon であるが, 本稿で用いた R のパッケージは通常のものであるため, 基本的には単一コアでの処理時間と考えられる. 表 2 の計算機は, 通常使用するデスクトップ計算機であり, OS は Mac OS, R は Version 3.1.1, rpart パッケージのバージョンは rpart_4.1-10 である.

R では system.time 関数があり, これを用いると, 処理時間のユーザ CPU 時間, システム CPU 時間, 経過時間を計測することができる [3]. 時間の解像度はシステム依存であるが, 通常 1/100 秒である. 本稿ではユーザ CPU 時間を処理時間として用いる. 3.3 で述べたように, rpart 関数は内部で R 言語や C 言語で作成された他の関数を再帰的に呼び出すよう実装されており, 内部処理別の処理時間計測が困難であるため, 本稿では rpart 関数全体の処理時間を調べた結果を報告する.

最初に, 今回用意した実データを用いて処理時間を計測する. 今回用意した 245 件のオブザベーションと 12 個の予測変数では処理時間がミリ秒程度であり, 時間の解像度の影響が大きいと考えられる. そこで rpart 関数実行を 1000 回繰り返して数秒程度の処理時間にし, 計測結果を 1000 で除算することで, 1 回あたりの処理時間を見積もった.

実データを用いて実際に決定木を得る際の, オブザベーション数と処理時間の関係を図 4 に示す. 図 4 では, 245 件のデータからランダムに必要な個数のデータを抜き出した 10 種類のデータセットを作成して, 処理時間を計測した結果を用いる. 各データ量で 10 回の計測の最大値, 平均値, 最小値を示す. 図 4 によると, オブザベーション件数増加に対しておおむね線形で処理時間が増加する. そして, 240 件のオブザベーションでも 10m 秒程度で木が得られており, 十分対話処理可能な処理時間である.

図 4 では, 処理時間の最小値と最大値の差は平均値に対して十分小さく, 1 回の計測実行でおおむね実態を反映した結果が得られるものと考えられる. そのため, 以後は 1 回の計測実行で得られた結果のみを示す.

4.2 人工データを用いた性能評価

今回使用したデータでは, rpart パッケージで決定木を生成する際の処理時間は十分短く, 対話処理で結果が得られるが, データが大量になった場合に備えて, データを人

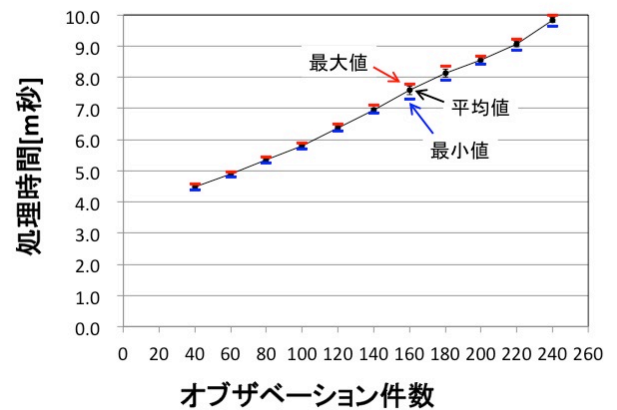


図 4 オブザベーション件数と処理時間の関係

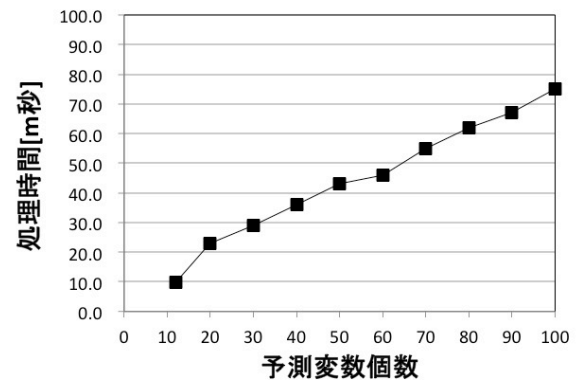


図 5 予測変数の個数と処理時間の関係 (オブザベーション 240 件)

工的に増加させて処理時間を調べる. 本稿ではデータ管理に Excel を用いているので, 乱数生成には Excel の乱数生成関数である rand, または範囲を指定して乱数を生成する randbetween 関数を用いる.

第 2 節で述べたように, 記述子として何を用いるかは議論があるが, 決定木を用いると, 分類に影響が大きい記述子が自動的に得られる可能性がある. 例えば, 文献 [7] で当初扱った 3000 個以上の記述子をそのまま決定木作成時に入力してしまい, 影響の大きい記述子を自動決定することも考えられる. そこで, オブザベーションが 240 件の場合に, 乱数を用いて予測変数を人工的に追加し, 処理時間を計測した. 予測変数個数と処理時間の関係を図 5 に示す. 図 5 の左端のデータは図 4 中のオブザベーション 240 件での平均値をそのまま使用している. 図 5 によると, 予測変数個数の増加に対しておおむね線形に処理時間も増加していることがわかる. 図 5 によると, 予測変数が 100 個のときに処理時間が 80m 秒程度であるため, 例えば予測

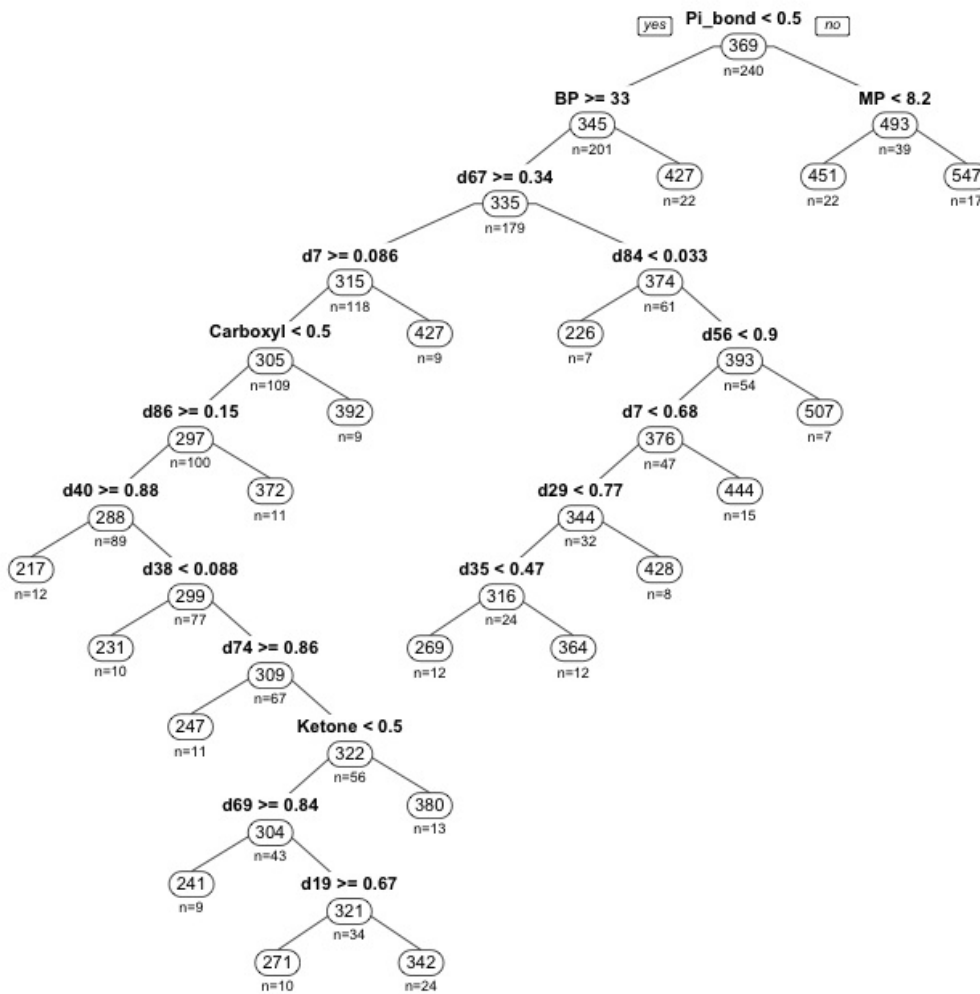


図 6 人工的に予測変数を追加した場合の決定木例 (オブザベーション 240 件, 予測変数 100 個, うち 12 個が実データ, 88 個は乱数.)

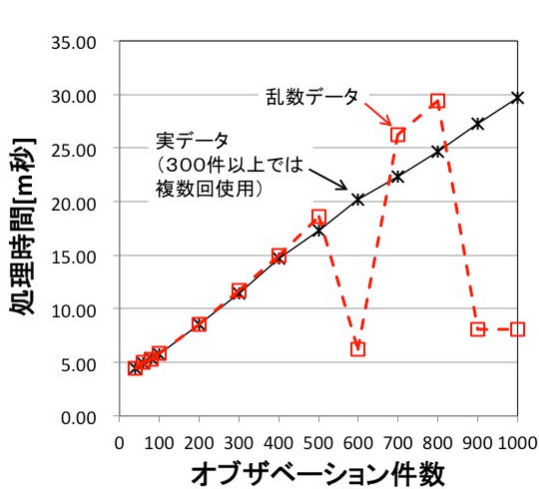


図 7 人工データを用いたオブザベーション件数と処理時間の関係

図 8 乱数を用いたオブザベーション 1000 件による決定木作成例

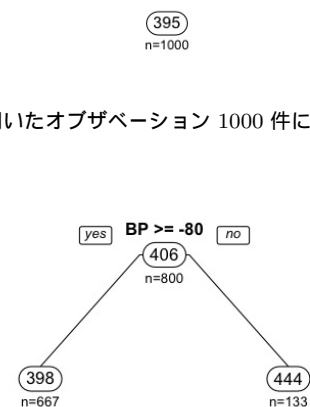


図 9 乱数を用いたオブザベーション 800 件による決定木作成例

変数が 3000 個のときは, おおむね 30 倍の 2.4 秒で計算できることとなる. この程度の処理時間であれば対話処理可能と考えられるため, この予測が正しいとすれば, 予測変

数 3000 個を持つ 240 件程度のオブザベーションでの決定木作成は、現在の計算機環境上で現在の rpart パッケージを用いて十分可能である。

参考までに、乱数を用いて予測変数を追加した際に得られた決定木を図 6 に示す。図 6 中の分岐条件において、不等号の左側に現れる「d」と 2 桁の数字を組み合わせた予測変数は人工的に追加した乱数であり、実際には発火点と何の関係もない偽物の記述子である。そのため、偽物の記述子を分岐条件に用いたルールは偽物のルールである。図 6 の深さ 1 までのノードは実データに含まれる記述子を用いており、真のルールである可能性が高い。一方深さ 2 の左端のノードでは偽物のルールで分岐しているために、過学習の結果現れたルールと考えるべきである。

次に、予測変数は 1 2 個のままオブザベーション件数をさらに増加させた場合の処理時間を調べる。オブザベーション件数を増やす際に、大量の実データを収集することは難しいので、まず乱数を用いて擬似的にオブザベーションを作成することを考えた。しかし、後述するように、それでは処理時間が実情にあわない可能性があるため、240 件の実データを繰り返し使用することで擬似的にオブザベーション件数を増やす方法も用いて処理時間を調べた。

2 種類の手法でデータを人工的に増加させ、計測した処理時間を図 7 に示す。図 7 において、正方形マークと破線は実データを用いずに乱数データのみを用いて用意したデータを使用した処理時間で、アスタリスクマークと実線は実データを繰り返し用いた処理時間である。なお、乱数データはオブザベーション 200 件以下でもすべて乱数を用いて生成している。また、実データを用いた図 7 中の計測結果は、200 件以下では図 4 中の平均値をそのまま使用している。図 7 によると、実データを用いた場合はオブザベーション件数の増加に対しておおむね線形で処理時間が増加している。一方で、乱数データでは 500 件までは実データと同程度の処理時間であるが、600 件、900 件、1000 件では大きく処理時間が減少しており、700 件、800 件では実データと同様の処理時間である。

図 7 において乱数データの処理時間が大きく増減している理由を調べるため、それぞれの処理時間計測時に処理結果として得られた木を確認する。乱数データでは基準変数と予測変数の間には全く関係がないため、ルールが得られるとしても偽物のルールであり、極端な場合はルールが全く得られず根ノードのみの木を作成している可能性がある。

処理時間が実データと同様であると考えられる 800 件のオブザベーションで得られた木を図 9 に、処理時間が実データよりも著しく小さくなっている 1000 件のオブザベーションで得られた木を図 8 に示す。図 9 と図 8 のいずれも簡素であるが、2 つの大きな違いは分岐の有無である。

3.3 で述べたように、分岐がある場合は交差検定が行わ

れ、計算量が大幅に増加する可能性がある。図 9 では分岐は 1 回のみであるが、それでも図 7 のように大幅に処理時間が変化することがわかった。また、分岐に関連する処理は主に数値計算を行っており、分岐があると処理時間が大幅に増加することから、本稿が本来目的とする処理を行う場合は、例えば入出力関連処理などよりも計算部分で実行時間を要することが間接的にわかる。

5. おわりに

本稿では発火点を予測することを目的として、R の rpart パッケージを用いた決定木の処理時間を検討した。その結果、240 種類程度の分子で 12 個程度の記述子を用いる場合はミリ秒程度の処理時間で決定木が得られることがわかった。そして、1000 種類程度の分子で 100 個程度までの記述子を用いても、データ量増加に対しておおむね線形で処理時間が増加することがわかった。一般にデータマイニングにおいては入出力処理も処理時間に影響する可能性があるが、本稿で行う処理の範囲では、処理時間は主に計算処理に依存することが間接的に示唆された。

今後の課題は以下の通りである。今回作成した決定木では発火過程モデルを考慮して選択した記述子を用いたが、今後はシミュレーションで得られる情報や記述子の相関なども考慮して、適切な記述子の発見手法を検討したい。本稿の結果、3000 個程度の記述子を扱って対話処理を行うことも可能であることが予想できたので、今後取り組んでいきたい。決定木はニューラルネット等よりも、分類で用いた規則が人間に理解しやすいが、得られたルールは最適である保証がないことが指摘されている [1], [2]。このような手法をどのように物質関連分野で利用していくのか、利用方法なども今後検討していきたい。

参考文献

- [1] 豊田秀樹：データマイニング入門，東京図書株式会社 (2008)。
- [2] 金 明哲：[連載] フリーソフトによるデータ解析・マイニング 第 19 回 R と樹木モデル (2)，入手先 (<https://www1.doshisha.ac.jp/mjin/R/19.html>) (2016.2.8)。
- [3] 福島真太郎：R によるハイパフォーマンスコンピューティング，ソシム株式会社 (2014)。
- [4] 吉村壽次代表編集：化学辞典 (第 2 版) 小型版，森北出版株式会社 (2009)。
- [5] 国立医薬品食品衛生研究所 (NIHS)：国際化学物質安全性カード (ICSC) 日本語版，入手先 (<http://www.nihs.go.jp/ICSC/>) (2015.12.20)。
- [6] J.Gasteriger, T.Engel 編集，船津公人，佐藤寛子，増井秀行訳：ケモインフォマティクス 予測と設計のための化学情報学，丸善株式会社 (2005)。
- [7] Tsai, F.-Y., Chen, C.-C., Liaw, H.-J.: *A model for predicting the auto-ignition temperature using quantitative structure property relationship approach*, *Procedia Engineering* 45, 512-517, (2012)。
- [8] Shi, J., Chen, L., Chen, W.: *Prediction on the auto-*

ignition temperature using substructural molecular fragments, *Procedia Engineering* 84, 879-886, (2014).

- [9] J. G. Quintiere, 大宮喜文, 若月薫訳: 基礎 火災現象原論, 共立出版 (2009).
- [10] 中田 侑江, 林 亮子: 炭化水素および類例分子の発火点決定木, MPS-2016-107-14, (2016).