博士論文

ガンマ線スペクトルデータを用いた ウラン廃棄物ドラム缶の迅速な分類への 機械学習法の適用性に関する研究

2017年 3月

秦 はるひ

岡山大学大学院 自然科学研究科

目次

第1章	序論	1
1.1	本研究の背景.............................	1
1.2	本研究の目的..............................	18
1.3	本論文の構成と内容.......................	19
第2章	SVM を用いたウラン廃棄物ドラム缶の分類	21
9 1	新言 新言	 91
2.1 9.9	福口	21 99
2.2 9.3	题/// / · · · · · · · · · · · · · · · · ·	22
2.0	胖竹子順	20
2.4	枯米と考☆	28
2.5	柏 言 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	31
第3章	RF を用いたウラン廃棄物ドラム缶の分類	33
3.1	緒言	33
3.2	適用データ	34
3.3	解析手順	35
3.4	結果と考察	41
3.5	結言	48
第4章	3 クラス分類に対する SVM と RF の分類性能の比較	49
4.1	緒言	49
4.2	適用データ	49
4.3	解析手法と手順	52
4.4	結果と考察	61

i

4.5	結言	67
第5章	結論	69
5.1	本研究の成果.............................	69
5.2	今後の展開	70
付録A	対象データの主成分分析	73
付録 B	対象データのニューラルネットワークによる分類	81
付録C	対象データのクラスター分析	85
付録 D	SVM におけるデータの前処理方法の比較	93
参考文献		97
謝辞		105
本論文を	構成する発表論文等	107

第1章

序論

1.1 本研究の背景

我が国は,主なエネルギー源の化石燃料を輸入に頼っており,国内外の状況の 変化に影響を受けやすい.そのため,原子力エネルギーを重要なエネルギー源と して技術開発を行い,商用利用してきた.2011年に発生した東京電力福島第一原 子力発電所事故を契機に,原子力エネルギーのあり方が見直されてきた.2014年 に閣議決定したエネルギー基本計画では,原発依存度を可能な限り低くするとし ているが,安全性,安定供給,経済効率性及び環境適合の総合的観点から依然とし て原子力は重要なベースロード電源として位置づけられている[1].

我が国では,2010年に19.5% あったエネルギー自給率が,東日本大震災以降 6%まで落ち込み,OECD加盟国34カ国中2番目に低い水準で,他の非資源産出 国のスペイン(26.7%),イタリア(20.1%),韓国(17.5%)と比較しても極めて低 く,エネルギーの調達先の多角化や国際資源の開発が進められているところであ る.また,原子力発電所の停止に伴い,火力発電への依存度が増え,温室効果ガス 排出量が増加していることからも,温室効果ガスを排出しない原子力エネルギー は重要な位置づけとなっている.2030年度のエネルギー需要は,徹底した省エネ ルギーを前提として原油換算で326 Gℓと見込まれている.政府は,その際のエ ネルギー自給率を23.4%まで引き上げることを目標としており,エネルギー構成 において原子力エネルギーは20~22%としている[2].

原子力エネルギーを利用するには、原子力発電だけでなく核燃料サイクルと呼ばれる一連の活動が必要である(Fig. 1.1). ウラン鉱石の採鉱から始まり、ウラ

ンの製錬・転換,濃縮を経て,核燃料に加工する.発電に使用した核燃料には,ま だ燃料として使えるウランとプルトニウムが残る.核燃料からウランとプルトニ ウムを取り出して新しい燃料(MOX 燃料)を作り、原子力発電所で再利用するプ ルサーマルと呼ばれるサイクルと,ウランを取り出し,再び濃縮し核燃料に加工 し原子力発電所で再利用するサイクルがある.原子力施設の運転に伴って,それ ぞれの工程で放射性廃棄物が発生し,放射能濃度レベルに応じた処理・処分が行 われる.



Fig.1.1 Nuclear fuel cycle.

原子力施設の運転に伴い発生する放射性廃棄物だけでなく,原子力発電所およ び核燃料施設を解体する際にも,放射性廃棄物が発生する.このとき発生する廃 棄物の多くは,低レベル放射性廃棄物と呼ばれる放射能濃度が比較的低いもので ある.また,東京電力福島第一原子力発電所の廃炉においても放射能濃度が低い ものから高いものまで幅広いレベルの放射性廃棄物が発生する.含まれる放射性 核種やその濃度(放射能インベントリ)に違いはあるものの,大量の放射性廃棄物 の放射能インベントリ情報を把握し適切に処分することは共通の課題である. 1.1.1 放射性廃棄物の分類と処分方法の概要

1.1.1.1 国際的な放射性廃棄物の分類と処分方法

国際原子力機関(IAEA)の放射性廃棄物の分類方法では,含まれる核種の半減 期と量によって Fig. 1.2 のように分類される [4]. 放射能濃度レベルが高くても 核種の半減期が短いために低いレベルの廃棄物として分類されるものや放射能濃 度レベルが低くても核種の半減期が長いために高いレベルの廃棄物として分類さ れるものなどもある. この分類に対応して処分方法が変わり,クリアランスや浅 地中処分から地層処分と多岐にわたる.



Fig.1.2 Conceptual illustration of the waste classification scheme [4].

1.1.1.2 我が国における放射性廃棄物の分類と処分方法

我が国では、国際的な放射性廃棄物の処分方法を運用するため、「原子力政策大綱」(平成17年10月14日閣議決定)において、放射性廃棄物の処理・処分の考 え方を示した.そこでは、「放射性廃棄物は、4つの原則(発生者責任の原則、放 射性廃棄物最小化の原則、合理的な処理・処分の原則、国民との相互理解に基づく

発生施設		放	(射性廃棄物の種類	埋設処分区分
百加珊旋恐	高レベル放射性廃棄物		レベル放射性廃棄物	地層処分
中规理工作			超ウラン核種を含む	放射能濃度レベルに応
				じて
MOX 燃料	1	放射	†性廃棄物(TRU 廃棄物)	浅地中~地層処分
加工施設	低			
ウラン取扱	レベ		ウラン廃棄物	未定(放射能濃度レベ
施設	ル			ルに応じて浅地中~地
	放 射			層処分)
	3性廃棄物		放射能濃度レベルの比	余裕深度処分(L1)
		廃棄物 発電所廃棄物	較的高い廃棄物	
原子力発電			放射能濃度レベルの比	浅地中ピット処分
所			較的低い廃棄物	(L2)
			放射能濃度レベルの極	浅地中トレンチ処分
			めて低い廃棄物	(L3)

 Table1.1
 Classification of radioactive waste in Japan.

Table1.2 Disposal methods depending on the classification of radioactive waste.

法令区分	方式	処分区分	処分方法
第一種廃棄物埋設	隔離型	地層処分	地下 300m より深い地層中
			に処分
		余裕深度処分	一般的な地下利用に対して
		(L1)	十分余裕を持った深度(地
			下 50~100m)への処分
第二種廃棄物埋設	管理型	浅地中ピット処	コンクリートピットを設け
		分 (L2)	た浅地中への処分
		浅地中トレンチ	人工構築物を設けない浅地
		処分(L3)	中埋設処分

実施の原則)の下で,その影響(放射性廃棄物中の放射性核種による被ばく)が有 意でない水準にまで減少するには超長期を要するものも含まれるという特徴を踏 まえて適切に区分を行い,それぞれの区分毎に安全に処理・処分することが重要」 としている.

この考えをもとに,原子力施設から発生する放射性廃棄物は発生施設と放射能 濃度レベルに応じて分類される(Table 1.1).隔離型と呼ばれる人間の生活環境か ら空間・物理的に隔離した第一種廃棄物埋設と管理型と呼ばれる比較的浅いとこ ろに処分するため空間・物理的隔離が十分ではなく制度的管理をしなければなら ない第二種廃棄物埋設がある.第二種廃棄物埋設では,放射性廃棄物の放射能濃



Fig.1.3 Concept of classification of radioactive waste depending on radioactivity in Japan.

度レベルによって処分深度や処分場のバリアの考え方が異なる(Table 1.2). 放射 性廃棄物の放射能濃度レベルによる区分の概念を Fig.1.3 に示す.また,核燃料物 質又は核燃料物質によつて汚染された物の第二種廃棄物埋設の事業に関する規則 (平成 27 年 8 月 31 日改正)および核原料物質,核燃料物質及び原子炉の規制に関 する法律施行令(平成 28 年 3 月 30 日改正)で定められる放射能濃度の上限値を Table 1.3 に示す.

現在, 浅地中トレンチ処分としては, 国立研究開発法人日本原子力研究開発機構 (JAEA) において動力試験炉 (JPDR)の解体に伴い発生したコンクリート等廃棄 物のトレンチ処分実地試験が行われている. 浅地中ピット処分については, 日本 原燃株式会社六ヶ所低レベル放射性廃棄物埋設センター(青森県六ヶ所村)にて 埋設事業が行われている. 余裕深度処分については, 現在原子力規制委員会の廃

放射性	濃度上限値(Bq/t)		
核種	浅地中トレンチ処分	浅地中ピット処分	余裕深度処分
$^{14}\mathrm{C}$	_	100 ギガ	10 ペタ
^{36}Cl	_	_	10 テラ
60 Co	10 ギガ	1ペタ	_
⁶³ Ni	_	10 テラ	_
90 Sr	10 メガ	10 テラ	_
$^{99}\mathrm{Tc}$	_	1 ギガ	100 テラ
129 I	_	_	1 テラ
$^{137}\mathrm{Cs}$	100 メガ	100 テラ	_
アルファ線を放出 する放射性物質	_	10 ギガ	100 ギガ

Table 1.3 Upper limit of radioactive concentration.

炉等に伴う放射性廃棄物の規制に関する検討チームで考え方[6]を打ち出したが、 具体的な規準等はまだ決まっておらず、処分実施には至っていない.地層処分に ついては、処分地の選定および同意形成を国が主体となって行っているところで ある.

1.1.1.3 クリアランスおよびクリアランスの類似概念

クリアランス (clearance) は、届出または認可されている施設や行為内の放射 性物質・物体を規制機関による放射線防護を目的とした規制管理から外すことで ある [7]. 人工核種について種々の被ばく経路を想定したシナリオで被ばく線量 が年間 10 µ Sv (一部発生頻度の少ないシナリオでは年間 1mSv) 程度になるよ うに放射能濃度レベル (クリアランスレベル) が設定され、天然核種については 自然界の存在範囲を考慮してクリアランスレベルが設定される [8]. 我が国では、 クリアランスレベル以下であることを国が確認することで規制対象から外れる (クリアランス制度). 2005 年に制定された原子炉の廃棄物に対するクリアラン スレベルを Table 1.4 に、2011 年に制定されたウラン加工施設等の廃棄物に対す るクリアランスレベルを Table 1.5 に示す [9, 10]. ウランの加工施設等の廃棄物 については、クリアランスレベルが設定されているものの現行では金属のみの制 度となっている. これまでに、原子炉では、日本原電株式会社東海発電所、中部 電力株式会社浜岡原子力発電所 5 号機のタービン、JAEA JRR-3 のコンクリー

,					
放射性物質	放射能濃度	放射性物質	放射能濃度	放射性物質	放射能濃度
の種類	(Bq/g)	の種類	(Bq/g)	の種類	(Bq/g)
³ H	100	⁶³ Ni	100	¹²⁹ I	0.01
^{14}C	1	⁶⁵ Zn	0.1	^{134}Cs	0.1
³⁶ Cl	1	90 Sr	1	^{137}Cs	0.1
⁴¹ Ca	100	94 Nb	0.1	133 Ba	0.1
46 Sc	0.1	95 Nb	1	152 Eu	0.1
^{54}Mn	0.1	99 Tc	1	154 Eu	0.1
55 Fe	1000	106 Ru	0.1	$^{160}\mathrm{Tb}$	1
59 Fe	1	^{108m} Ag	0.1	182 Ta	0.1
58 Co	1	^{110m} Ag	0.1	²³⁹ Pu	0.1
⁶⁰ Co	0.1	124 Sb	1	241 Pu	10
⁵⁹ Ni	100	123m Te	1	^{241}Am	0.1

Table1.4 Clearance level for nuclear power plant[9, 10].

Table1.5 Clearance level for uranium processing plant[9, 10].

放射性物質の種類	放射能濃度 (Bq/g)
$^{232}\mathrm{U}$	0.1
$^{234}\mathrm{U}$	1
$^{235}\mathrm{U}$	1
$^{236}\mathrm{U}$	10
$^{238}\mathrm{U}$	1

トがら,原子炉等廃止措置研究開発センターの金属などがクリアランスされており,ウランの加工施設等では,JAEA 人形峠環境技術センター(以下,人形峠センター)の金属がクリアランスされている. RI 汚染物についてはクリアランス 制度はあるもののクリアランスされた実績はまだない.

クリアランスの類似概念として規制免除(exemption)という考え方がある. ある線源または行為に起因する被ばくが非常に小さく,規制管理の適用が正当化 できない場合や規制上の管理を適用しなくとも実際の線量やリスクレベルに関係 なく最適であると考えられる場合に,その線源または行為について規制上の管理 を免除するという規制機関の決定のことである [7].校正用線源のように放射線 量が少量のものなどが該当する.

さらに、地球表面での宇宙放射線、人体内の⁴⁰K および IAEA 安全基準で示される基準値を下回るような自然起源放射性核種(naturally occurring radioactive

material: NORM) などの規制手段で管理ができないものや管理しても効果がないものなどは、規制除外 (exclusion) として取り扱われる [7].

これらの考え方とは別に,我が国では,2008年に出された原子力施設における 「放射性廃棄物でない廃棄物」の取り扱いに関するガイドライン[11]で,原子力施 設から発生する廃棄物であるものの,放射性物質で汚染された履歴のない「放射 性廃棄物でない廃棄物(Non Radioactive waste: NR)」について,一般産業廃棄 物と同様の取り扱いをするとの考え方が示された.

1.1.2 ウラン廃棄物の特徴

1.1.2.1 ウラン廃棄物の定義とその現状

原子力委員会が平成12年に取りまとめた「ウラン廃棄物の処理処分の基本的な 考え方について」[12] にて、「原子炉施設の運転に使用されるウランは、その原料 となるウラン鉱石から、製錬、転換、濃縮、再転換、成形加工等の工程を経て製造 される. これらの各工程を行う施設の運転・解体に伴い放射性廃棄物が発生する. これらの放射性廃棄物をウラン廃棄物という.」とされており、我が国固有の考え 方である. 我が国では、ウラン鉱石の採鉱からウラン濃縮までの研究開発を行っ てきた JAEA や民間の濃縮、再転換および成形加工施設があり、保管されている ウラン廃棄物の物量は200ℓドラム缶換算で約11万4千本である(Fig. 1.4).



Fig.1.4 Amount of uranium bearing waste in Japan.



Fig.1.5 Amount of waste generation until 2050[14].

2050 年頃までに我が国で発生するウラン廃棄物の最終処分体である廃棄体お よびクリアランス対象物(金属)の重量の試算結果を Fig. 1.5 に示す. 埋設処分 対象物は約5万t となり, 200ℓ ドラム缶に換算すると最大約25万本発生する.

1.1.2.2 ウラン廃棄物の処分における課題

ウランを含む放射性廃棄物であるウラン廃棄物(Uranium bearing waste)は、 我が国では原子炉から発生する放射性廃棄物とは分けて分類している(Table 1.1).原子力委員会が2000年に出した「ウラン廃棄物処理処分の基本的考え方 について」において、ウランは長半減期核種であるため、放射能の減衰に応じて 管理を行う浅地中処分の考え方が適用できないとされた[12].一方で他の研究施 設などから発生する研究施設等廃棄物と一緒に浅地中処分する方策も示されてい る[13].課題として、以下の3点が挙げられる[14].

- ウランおよびその子孫核種は、自然に存在する放射性物質であるため自然 放射能との関連の考慮
- ウランの崩壊に伴い生成する子孫核種であるラドンに起因する被ばくへの
 対応
- 半減期が極めて長く減衰に期待できないことを踏まえ、長期事象を踏まえ
 た考え方の整理が必要

海外では、ウラン廃棄物という枠組みはなく、ウラン廃棄物は他の放射性廃棄 物と同様に処分されている [15, 16, 17]. 旧原子力安全委員会は、長半減期核種で も濃度が低い場合には原子炉廃棄物と同様の処分を行うことが可能であるという 考え方を示している [13].

1.1.2.3 ウランの特徴

²³⁴U, ²³⁵U および ²³⁸U は天然に存在し,長寿命の放射性元素である.²³⁸U を 親核種とする放射壊変系列はウラン系列 (Fig. 1.7) と呼ばれ,途中,被ばく評価 上影響の大きい気体の ²²²Rn を経由する.通常,²³⁴U は ²³⁸U の子孫核種として 放射平衡状態で存在する (Table 1.6).

ウラン鉱石では²²⁶Ra 以降の子孫核種を伴っているが,原子炉の燃料を作る工程で,これらの不純物などが除去される.そのため,²²⁶Ra 以降の子孫核種の有意な生成は,数十万年を要する(Fig. 1.6).

	半減期	天然存在比	比放射能	天然の 1gU の放射能
	(y)	(at%)	(Bq/g)	(Bq/g)
²³⁴ U	2.455×10^5	0.0054	2.3×10^8	1.2×10^4
²³⁵ U	7.038×10^8	0.7204	8.0×10^4	5.7×10^2
²³⁸ U	4.468×10^{9}	99.2742	1.2×10^4	1.2×10^4

Table1.6 Characters of natural uranium[18].



Fig.1.6 Time dependence for radioactivity of refined natural uranium[19].



一度,原子炉の燃料として使用されたウランは再処理後,回収されたウランと して流通する.このウランの組成は,天然の精製されたウランとは異なり,²³²U, ²³³U,²³⁶Uが出現する.一例として,比出力が35 MW/t,燃焼度が28 GWD/tの 原子炉で使用した後,JAEA 核燃料サイクル工学研究所東海再処理施設で再処理 してから,20 年経過したと仮定した時のウラン同位体の重量比を,燃焼計算コー ドORIGEN2.2 で計算した.ここで,東海再処理施設にて再処理されてから20 年 程度経過しているウラン廃棄物が多いため20 年と設定した.ウラン同位体の重 量比は Table 1.7 に示す通りである.また,²³²U,²³⁶U はトリウム系列に合流し, 多くの子孫核種を伴って存在する (Fig. 1.8).

回収ウランの放射性核種の濃度組成 (核種組成)は、ウラン同位体の違いだけで なく、⁹⁹Tc などの核分裂生成物、²³⁷Np などの超ウラン元素も計算上は微量では あるが存在する.

この研究では、天然の鉱石から精製されたウランを天然ウラン(Natural Uranium: NU),再処理された使用済み燃料から回収されたウランを回収ウラン(Reprocessed Uranium: RU)という.いずれも燃料とするために²³⁵Uの濃縮を行 う.天然の²³⁵U 濃度は 0.711wt% であり、これを超えるものを濃縮ウラン(Enriched Uranium: EU),下回るものを劣化ウラン(Depleted Uranium: DU)とい う.原子炉の燃料として使用されるのは 4wt% 程度の NU や RU であり、使用済 み燃料の再処理後は 1wt% 程度の RU となる.濃縮工場で発生する劣化ウランは 0.45% 程度の NU や RU である.

	比放射能	重量比
	(Bq/g)	$(\mathrm{wt}\%)$
$^{232}\mathrm{U}$	$8.3 imes 10^{11}$	2.8×10^{-8}
²³³ U	$3.6 imes 10^8$	1.7×10^{-7}
$^{234}\mathrm{U}$	$2.3 imes 10^8$	$2.9 imes 10^{-2}$
$^{235}\mathrm{U}$	$8.0 imes 10^4$	1.6
²³⁶ U	2.4×10^6	0.47
²³⁸ U	1.2×10^4	98

Table1.7 Weight ratio of uranium isotopes in reprocessed uranium.



1.1.3 本研究に関する既存技術および既存研究

1.1.3.1 低レベル放射性廃棄物の放射能濃度推定方法

処分場の安全評価で重要となる核種(重要核種)には,ガンマ線を放出せず,廃棄 物の外側から非破壊測定で定量することが困難な難測定核種がある.一般に,難 測定核種の放射能濃度評価には,サンプリングを行い,核種分離のための化学的 な前処理が必要であるが,大量の廃棄物に対して行うのは現実的でない.そのた め,平均放射能濃度法と呼ばれる代表サンプルの放射化学分析等の測定結果から 得られる平均的な放射性物質濃度を用いる方法やガンマ線測定で廃棄物中の放射 能濃度を推定するスケーリングファクター(SF)法などを使って放射能インベン トリを推定評価する手法などが用いられている [21, 22, 23, 24].

日本原燃株式会社六ヶ所低レベル放射性廃棄物埋設センターでは,廃棄確認の 一項目である廃棄体の放射能濃度確認において,難測定核種で,理論計算法,平均 放射能濃度法および SF 法が採用されている [25, 26, 27].

■1.1.3.1.1 理論計算法 廃棄体化する前の廃棄物(原廃棄物)ごとに,計算条件 を設定して放射能濃度を推定する手法である.放射化量もしくは ORIGEN など の原子炉燃焼計算によって理論的に求められる値を用いる [23].ただし,理論計 算法が適用できる条件は限られており,放射化による汚染であれば中性子照射に より生じた核種のみ,付着等の汚染であれば生成メカニズムおよびプラント内の 物理化学的挙動が同一とみなせる場合のみに限定される.

■1.1.3.1.2 平均放射能濃度法 ガンマ線測定が容易である放射性核種(Key 核 種)の濃度によらず,難測定核種の濃度が一定範囲にあるときに適用される手法で ある.適切なサンプリングを行い,化学分析による放射能濃度の分布を調査し,平 均放射能濃度を求める.個々の廃棄体にその平均放射能濃度値を与えるため,埋 設処分の安全評価から求められる最大放射能濃度に対して裕度がある場合に限ら れる [23]. ■1.1.3.1.3 SF法 事前に設定した廃棄物グループから代表サンプルをとり,化 学分析を行い,ガンマ測定が容易な核種(Key 核種)と難測定核種の相関関係を 求め,Key 核種のガンマ線測定結果から難測定核種の量を推定する方法である. Fig. 1.9 に,サンプルを化学分析し,難測定核種である⁶³NiとKey 核種の⁶⁰Co の放射能濃度の分布を求めた例[22]を示す.両者ともに対数正規分布を示し,そ の組成比についても対数正規分布であり,サンプルの代表性に問題がないことが いえる.⁶³Niと⁶⁰Coの放射能濃度をプロットすると,対数スケールで線形性が 見られる(Fig. 1.10).⁶³Niと⁶⁰Coは,生成メカニズムと汚染状況が似ている ため,このケースでは,SF法が適用可能であると判断される.

このように、求めた相関関係と廃棄体ごとのガンマ線測定結果から放射能イン ベントリを推定するため、計算値から一律に推定する理論計算法や平均的な放射 能濃度で一律に推定する平均放射能濃度法に比べ,実態に即した推定方法である.



Fig.1.9 Exaples of distribution characteristic of radionuclide concentration and composition ratio[22].

(a) Radioactivity concentration distribution of ⁶³Ni, (b) r adioactivity concentration distribution of ⁶⁰Co, (c) composition ratio distribution (⁶³Ni/⁶⁰Co).



Fig.1.10 Scatter diagram for radioactivity concentration of ⁶³Ni and ⁶⁰Co on a logarithmic scale[22].

1.1.3.2 機械学習法

近年,統計的な手法を用いた機械学習法の研究が進み,ビックデータなどの大 規模なデータを解析する手法の研究がなされている.これらの手法を用いれば, 核種組成グループのデータに特徴的な量を総合的に統計的手法で自動的に抽出 し,人間では曖昧だった基準を明確にして瞬時に分類することができる.

本研究では、ガンマ線スペクトルデータに対して機械学習を行い、短時間で自動的に核種組成グループへ分類する方法について検討する.機械学習として、文字や画像などのデータを学習して特徴を抽出し、未知のデータを識別して既知の クラスに分類するパターン認識と呼ばれる方法を用いる.

■1.1.3.2.1 ニューラルネットワーク 形式ニューロンと呼ばれる判別分析をす る判別モデルを階層的に結合させたパターン認識モデルである [28, 29]. スペク トルデータの解析においては,正解が既知であるデータ集合(訓練データ)を学 習し分類モデルを作成する教師付き学習法を用いるのが一般的である.

初めて放射線スペクトルに対し、機械学習法を適用したのは、Olmos らである [30, 31]. これらの研究は、ガンマ線スペクトルデータに対し、ニューラルネット ワークを用いて核種同定を行った研究である. 訓練データには標準線源が用いら れており、数種の核種をさまざまな割合で混合した線源から得られるガンマ線ス ペクトルデータに対し、核種の混合比を推定している. その後、さらに NaI (Tl) シンチレーション検出器で取得したスペクトルデータの校正ずれなどによるドリ フト問題を改良している [32]. また、Kangas らは、アメリカのハンフォードサイ トで発生する放射性廃棄物に対し、迅速な汚染の同定を目的に、ガンマ線および アルファ線スペクトルデータへニューラルネットワークを適用した [33]. いずれ も全チャネルのスペクトルデータを入力し解析を行っているため、階層を深くす ることが困難であった.

その後,ニューラルネットワークに全チャネルのスペクトルデータを入力する のではなく,何らかの処理で次元削減したデータを入力する研究が出てきた.こ れにより,ニューラルネットワークの中間層の計算ができるようになり,ニュー ラルネットワークの階層性という特徴が発揮できるようになった.前処理とし て,主成分分析による次元削減が多くの研究で見られる [34, 35, 36]. その他に も,エネルギーピークを抽出して解析する方法 [37],最小2乗法で求めたピーク 面積を入力する方法 [38] などが研究されている.

ニューラルネットワークを用いた分類の応用として、盗取や混入などの警報シ ステムへの活用も考えられている [39].近年では、環境放射線の調査に応用され る例が増えてきている [40, 41, 42, 43, 44].ただし、放射性廃棄物ドラム缶の放射 能濃度の推定への応用は検討されていない.

■1.1.3.2.2 *k* 近傍法 ベクトル空間上の訓練データの計数値などの特徴量を表 す点の集合と、未知のデータの点の距離で分類する手法である. *k* は、未知の データの近傍にある訓練データの数を表し、*k* 個のデータがある高次元球の中に どのラベルのものが一番多く存在するかで分類する.

Dragović らの研究と Bellinger らの研究は, *k* 近傍法, ニューラルネットワーク などの比較を行っており, ニューラルネットワークが一番成績が良かったとして いる [45, 46]. Wei らは, 埋立られた廃棄物の検知や埋めた深さの分類に対し, *k* 近傍法を用いる際の最適な前処理方法について検討している [47].

k 近傍法の研究例は少ないが、スペクトルデータの機械学習ではニューラルネ ットワークの方が優れていると考えられる.

■1.1.3.2.3 サポートベクターマシン(SVM) 訓練データからサポートベクター と呼ばれる境界面からの距離が最大になるようなベクトルを導出し,2 クラスに 分けるための境界面を作成するアルゴリズムである [48]. 導出には,ラグランジ ュの未定乗数法を用いた最適化問題を解くことになるため,局所最適解が必ず大 局的最適解となる.

SVM の適用可能性については, 化学分析でのスペクトルデータに対する先行研 究が多くあり, マススペクトル法 [49, 50, 51], フーリエ変換赤外 (FT-IR) 分光法 [52], 紫外可視近赤外 (NIR) 分光法 [53, 54, 55, 56] およびラマン分光法 [57, 58, 59] などがある.

放射線スペクトルデータへの適用については, 前述の [41, 42] があるが, 前処理

方法で次元削減を行っているため、全チャネルを入力する一般的な SVM を使っ たスペクトル解析手法の適用可能性については不明である.

■1.1.3.2.4 ランダムフォレスト(RF) 単独では精度の低い決定木と呼ばれる 分類モデルを多数組み合わせることで精度を上げるというアンサンブル学習のひ とつである[60]. 訓練データをランダムに再構成するため外れ値に強く, さらに, 学習するパラメータもランダムに抽出するためノイズに強く計算速度も速いとさ れている. 化学分析では, マススペクトル法[51,61], NIR 分光法[56], FT-IR 分 光法[62] などで適用されている. しかし, 放射線スペクトルの分野では研究事例 がなく, 適用可能性が明らかになっていない.

1.2 本研究の目的

1.1.3 で示した通り, 原子力施設の操業および廃止措置における放射性廃棄物の 処理計画の策定や放射性廃棄物処分場での確認などにおいて, 放射能インベント リの情報が重要となる. 放射能インベントリの情報として把握すべき放射性核種 には, ガンマ線による非破壊測定が困難な難測定核種が存在している. 難測定核 種の推定には, SF 法の Key 核種と難測定核種の間の相関関係が利用されること がある. 先行事例として, 発電所廃棄物については SF 法の適用可能性について 検討が行われ, 原子炉ごとに得られる SF で放射能インベントリを評価してきた. 一方, ウラン廃棄物では, NU や RU など核種組成が異なった物質が同じ核燃料 サイクルを通る. 核種組成に応じたさまざまなガンマ線が放出されるため, 発生 履歴だけでなく, ガンマ線スペクトルデータから得られる核種組成グループ情報 を付与することで SF 法での放射能インベントリ評価の精度向上が期待される.

1.1.2 で示した通り,ウラン廃棄物は,我が国において現在 200ℓ ドラム缶換算 で11 万本以上保管されており,2050 年頃までにさらに約 25 万本発生すると見込 まれている. JAEA 人形峠センターでは,工程の運転履歴から NU, RU を分類 し,判断がつかないものについては安全側に RU に分類してきた.しかし,放射 性廃棄物を処理し,最終処分体である廃棄体を作製する際に放射能インベントリ を過大評価していれば,処理にかかる費用も増大する.実際の放射能インベント リにあった処理・処分をすることにより, 合理的な処分が実現するため, 核種組成 グループごとに精度良く分類することが重要である.

ガンマ線スペクトルデータに基づいて複数の判定者によって核種組成グループ に分類する場合,特定のピークの有無などで判定することになる.このとき,境 界領域にあるデータは,ピークの有無などの分類根拠が判定者により曖昧となり, 結果が一律にならないことがあるため,明確な分類根拠が求められている.機械 学習を用いれば,グループに特徴的な量を総合的に統計的手法で自動的に抽出し, 分類根拠を数値化して瞬時に分類できる.機械学習では,放射線エネルギーと核 種の組み合わせのライブラリー情報を使わずに,スペクトルデータのエネルギー ピークのフィッティングをしないまま,チャネルと計数値の情報のみで分類モデ ルを構築することができるという利点もある.

ニューラルネットワークには,原理的に初期値によって最適解が変化するという局所解の問題があった.そこで,2クラス分類のための機械学習法として,SVM というパターン認識アルゴリズムが発達してきた.SVMは,局所最適解が必ず大 局的最適解となるため,ニューラルネットワークの弱点を克服した分類手法とし て注目されている.さらに,決定木と呼ばれる精度の低い分類モデルを多数組み 合わせて精度を高める RF と呼ばれるアンサンブル学習法が画像認識などで発展 してきた.計算速度も速く,ノイズや外れ値に強いとされている.

このようなことから,放射線スペクトルに対してよく研究されてきたニューラ ルネットワークではなく,パターン認識手法である SVM と RF の適用性につい て検討する.本研究では, NaI (Tl) シンチレーション検出器による非破壊測定で 得られた低解像度ガンマ線スペクトルデータ用いて,ウラン廃棄物の分類を行う.

1.3 本論文の構成と内容

第1章では,放射性廃棄物の分類と処分方法の概要,ウラン廃棄物の特徴を述 ベ,これまでに行われてきた放射能インベントリ推定方法とスペクトルデータに 対する機械学習を用いた研究事例について取り上げ,本研究の目的について説明 した.

第2章では、最初のステップとして、NUとRUへの2クラス分類問題のSVM

の適用可能性について評価した. ウラン廃棄物の中でも同一の廃液処理工程から 発生する核種組成の異なるドラム缶のガンマ線スペクトルデータを用いて検証し た. 使用したスペクトルデータについて, 測定方法と NU と RU での特徴的なエ ネルギーピークの違いを取り上げた. SVM の考え方について解説し, 重要なパ ラメータについてパラメータサーベイを行った. SVM を実施した結果について 考察し, 適用可能性があることを示した.

第3章では、ノイズに強い機械学習のアルゴリズムである RF を使った分類に ついて、比較するデータの前処理方法と具体的なアルゴリズムについて検討した. 前処理方法ごとの精度について考察を行い、分類する際に最適な前処理方法を提 案し、RF について適用可能性があることを示した.

第4章では、実廃棄物の分類において2クラスより複雑な多クラス分類になる ことを想定し、ラジウムが多く含まれるウランを追加し、3クラスへの分類につい てSVMとRFの性能比較を行った.試験対象のドラム缶の性状について紹介し、 3クラスへ適用する際のSVMとRFのアルゴリズムについて検討した.試験結 果の考察から、SVMとRFの分類性能の特徴を明らかにした.

第5章では,第2章から第4章までの成果をまとめ,今後の展開について言及 した.

付録として,従来のニューラルネットワーク手法を踏襲し,付録 A でウラン廃 棄物の主成分分析,付録 B で主成分分析結果をニューラルネットワークに入力し 分類を行ったが,結果は好ましくなかった.また,付録 C では教師なし機械学習 法のひとつであるクラスター分析を行った.付録 D では,SVM におけるデータ 前処理方法の比較を行い,RF と同様の結果を得た.

第2章

SVM を用いた

ウラン廃棄物ドラム缶の分類

2.1 緒言

ウラン廃棄物の放射能インベントリの推定では、核種組成の違いから NU と RU に分けて考える必要がある.本章では、ウラン廃棄物ドラム缶から得られる 低解像度ガンマ線スペクトルデータに対し、SVM を使った分類が適用可能であ るか検討する.本研究の解析では、事前にガンマ線のエネルギーピークの情報が 必要なく、訓練データのみを使って、ガンマ線スペクトルデータを分類できると いう利点がある.

SVM は、パターン認識の機械学習法として、Cortes と Vapnik によって開発さ れた2クラス分類方法である [48]. SVM は、訓練データからサポートベクターと 呼ばれる境界面からの距離が最大になるようなベクトルを導出し、2 クラスに分 けるための境界面を作成するアルゴリズムである. 導出には、ラグランジュの未 定乗数法を用いた最適化問題を解くことになるため、局所最適解が必ず大局的最 適解となり、高い汎化能力を持つ.

本章では、SVMのパターン認識手法が放射線スペクトルに対しても適用可能で あるか、JAEA 人形峠センターの製錬転換施設の廃水処理工程において発生した ドラム缶のガンマ線スペクトルデータを用いて検討する.

2.2 適用データ

967 点の廃棄物ドラム缶のガンマ線スペクトルデータを用いて, SVM の適用可 能性について検討を行った. これらのドラム缶は, 1986 年~1996 年に JAEA 人 形峠センターの製錬転換施設の廃水処理工程から発生した中和澱物を収納してい る. ドラム缶には, 廃棄物発生日を元に NU または RU のラベルが付けられてお り, NU ドラム缶は 475 点, RU ドラム缶は 492 点である. ガンマ線測定は, 縦長 の 130 (W) × 80 (D) × 400 (L) mm の NaI(Tl) シンチレーション検出器 2 本, 鉄製の遮蔽容器, ドラム缶の回転台から構成される Q² system (CANBERRA 社 製)を使用し, 測定時間 300 秒で, ドラム缶を回転させながら測定した.

Fig. 2.1 に、Q² system で測定した典型的な NU と RU のガンマ線スペクト ルデータの例を示す. 解像度は低いものの、両方のスペクトルに ^{234m}Pa 由来の 1001 keV, 766 keV と ²³⁵U 由来の 186 keV のピークが見られ、RU のスペクトル では、原子炉内で生成される ²³²U, ²³⁶U の子孫核種であるトリウム系列の ²¹²Pb と ²⁰⁸Tl 由来の 239 keV, 511 keV, 583 keV のピークが見られる (1 チャネル= 約 4keV). 1~497 チャネル (約 0–2000 keV 相当)の計数値を使用した.

なお、訓練データには、NU 8 点と RU 4 点、計 12 点を用いた. 訓練データの ガンマ線スペクトルデータを Fig. 2.2、Fig. 2.3 に示す. これらの訓練データに は、運転期間ごとに一番初めに発生したドラム缶を選んだ. 分類対象のドラム缶 は、訓練データ 12 点を除いた 955 点である. データの前処理として、チャネルご とに計数値の対数をとった後、計数値の対数の全チャネルの総和で割って規格化 した(式 (2.1)). ここで x_n は、n 番目のチャネルの計数値を表す.

$$\boldsymbol{d} = \left\{ \frac{\log x_n}{\sum_{n=1}^{497} \log x_n} \right\}_{n=1}^{497}$$
(2.1)



Fig.2.1 Typical NU and RU spectra with NaI(Tl) detectors on Q^2 system.



Fig.2.2 Spectra of NU training data.



Fig.2.3 Spectra of RU training data.

2.3 解析手順

SVM は、Cortes と Vapnik によって開発された、2クラス分類のためのアルゴ リズムである [48]. 訓練データからサポートベクターと呼ばれる境界面からの距 離 (マージン)が最大になるようなベクトルを導出し、2クラスに分類するための 境界面を作成する (Fig. 2.4). 本研究で用いたカーネル法を使った SVM は、入力 データ d を高次元の特徴空間 H ~写像し、分類性能を向上させる (Fig. 2.5). こ こで求める境界面は、特徴空間の中で、2次計画法の最適解となる超平面である. さらに、本研究ではマージンの内側に訓練データが入り込むことを許容するソフ トマージン法を使うことで、汎化能力を上げる.

SVM で解くべき 2 次計画法問題を式(2.2)に示す.497 個の説明変数からな る i 番目の入力データ d_i とクラス値 $y_i \in \{+1, -1\}$ をもつ 12 個の訓練データ $S = \{d_i, y_i\}_{i=1}^{12}$ を用いてこの問題を解く.



Fig.2.4 Support vectors of 2 class data.



Fig.2.5 Concept of SVM using kernel method.

max
$$Q(\boldsymbol{\alpha}) = \sum_{i=1}^{12} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{12} \alpha_i \alpha_j y_i y_j K(\boldsymbol{d}_i, \boldsymbol{d}_j)$$

s.t.
$$\sum_{i=1}^{12} y_i \alpha_i = 0, \qquad 0 \le \alpha_i \le C \qquad (2.2)$$

ここで, α_i はラグランジュの未定乗数である. C はコストパラメータと呼ば れ,マージンの内側に訓練データが入り込む許容量を示すものである. C が高け れば高いほど,複雑な超平面を用いて訓練データを厳密に分けようとし,低けれ ば低いほど,マージンの内側にデータが入り込むことを許容する.

マッピングのためのカーネルは、ラジアル基底関数カーネル $K(d_i, d) = \exp(-\gamma \| d_i - d \|^2)$ を用いた、 γ は、ガウス分布の半径をコントロールするパラ メータである、 γ が大きくなると d_i のカバーする領域が狭くなるため、サポー トベクターの数が増え、 γ が小さくなると d_i のカバーする領域が広くなるため、 サポートベクターの数が減る.

次に,式(2.3)で定義される,未知のデータ*d*に対する決定値*D*と呼ばれる分 類の指標を導入する.なお,*D*=0の超平面が境界面となる.



Fig.2.6 Result of parameter survey of γ , C by grid search.

$$D(\boldsymbol{d}) = \sum_{i \in I} \alpha_i y_i K(\boldsymbol{d}_i, \boldsymbol{d}) + b$$
(2.3)

ここで,集合 I は,訓練データから得られるサポートベクターのインデックス を示し, b はバイアス項である [63].

NU とラベリングされている訓練データの場合は、クラス値 $y_i = 1$ とし、 RU とラベリングされている訓練データの場合は、クラス値 $y_i = -1$ とし た. グリッドサーチによるパラメータサーベイを行ったところ、 $\gamma = 10^{-3.1}$ で $C = 10^1$, 10^2 , 10^3 のとき、訓練データに対するモデルエラー値が最小となっ た (Fig. 2.6). 本研究では、 $\gamma \ge 10^{-3.1}$, C は、汎化能力を重視し、マージン の内側への訓練データの入り込みを一番許容する 10 と設定した.

なお、これらのデータは統計解析言語 R[64] で解析した. SVM については、 e1071 パッケージ [65] の svm 関数を使用した. パラメータ設定をした後、SVM の計算を行うには、0.35 秒かかり、955 点のデータに適用して決定値 D を求め るには、0.57 秒かかった.

2.4 結果と考察

Fig. 2.7 に、ドラム缶発生時期と式 (2.3) から導いた決定値を示す. D = 0 よ り上のエリアは、NUと分類されたデータの決定値を示し、下のエリアは、RUと 分類されたデータの決定値を示す. 初期値として、訓練データにおける分類クラ ス値に、NUは1、RUは -1を設定した. そのため、式 (2.3) から導出される決 定値が、1 や -1に近いものほど、より訓練データの特徴に近く、分類の確実性が 高い. 一方、分類の境界である0に近いものほど、分類の確実性が低い.

Fig. 2.8 に,元々のラベルと違う分類結果となったガンマ線スペクトルデータ を示す. Fig. 2.8 の A と B は NU と分類されたが,ラベル上では RU のドラ ム缶であった.しかし,RU 由来のピークがはっきりとはしていない.同じグラ フにおける C は,140 チャネル付近に,不明瞭なピークが出ているが,110 チャネ ルと 130 チャネル付近に 2 本のピークがないため,²⁰⁸Tl 由来のものとは認めら れない.Fig. 2.8 の D は,NU ラベルではあったが,RU に分類された.110 チャ ネルと 130 チャネル付近にピークがあるため,元々 RU のドラム缶だったと考え られる.

A, B, Cのドラム缶が発生した時期は, 製錬転換工程で, NUと RU が交互に 供給されていた.また, RUの製錬転換工程は, 試験段階であったことと乾式方法 を採っていたことより, 廃水処理による中和澱物の発生量が少ない.加えて, NU と RUの供給の切り替え時に発生したドラム缶については,工程のクリーンアッ プを行わなかったため,安全側に RU とラベリングしていた.そのため, SVM に よる分類結果は, 科学的には適正な分類結果であったが, 操業中の慣例にならった 分類結果にはならなかった.対照的に,Dのドラム缶は発生当時,ウランの製錬 転換は行われておらず,メンテナンスで発生したドラム缶と考えられ,付けられ た元々のラベルが間違いであったと考えられる.以上のことより,この SVM の 分類は,正しく行われたと考えられる.









2.5 結言

廃水処理工程から発生したウラン廃棄物ドラム缶のガンマ線スペクトルデータ 967点のうち,訓練データとして12点を機械学習し,残りの955点を用いて,SVM の適用可能性について検討を行った.その結果,4点のデータが元々付けられて いたラベルとは違う分類結果となった.うち1点は,元のラベルが間違っており, 3点は,操業中の慣例にならった保守的なラベルの付け方がされていた.元のラ ベルと SVM の分類結果は異なったもののいずれも科学的には正しい分類であっ た.SVM は,大量のデータを短い時間で分類することができ,決定値と呼ばれる 数値を伴った指標も提供できる.この手法は元のラベルの確認を行う補助ツール として活用でき,放射能インベントリの推定に役立つと期待できる.本章では,発 生日をもとに12点の訓練データを選んだが,最適な訓練データの選び方は今後の 課題である.
第3章

RF を用いた

ウラン廃棄物ドラム缶の分類

3.1 緒言

第2章では、低解像度ガンマ線スペクトルデータの2クラス分類に対し、SVM が適用可能であるとの結果が得られた.本章では、RF と呼ばれる画像認識や ビッグデータ解析などでよく用いられる手法を使った低解像度ガンマ線スペク トルデータの2クラス分類について検討する.RFは、計算速度も速く、ノイズ や外れ値に強いとされているため、放射線スペクトルにも適していると考えら れる.

Breiman が提案したパターン認識手法である RF は、与えた訓練データセット からランダムにサンプリングしたデータで多数の訓練データセットを再構成し、 学習により探知した分岐条件で、決定木と呼ばれる樹木状の分類モデルを多数作 成して分類するアンサンブル学習アルゴリズムである [60]. アンサンブル学習と は、単独では精度の低い分類モデルを多数組み合わせることによって精度を上げ る学習法である.

本章では,第2章で用いたデータから,メンテナンスで発生したドラム缶を除いたデータを使用し,RFの適用可能性について検討する.

3.2 適用データ

JAEA 人形峠センターの製錬転換施設の廃水処理工程で発生した中和澱物を封 入したドラム缶 954 点 (NU 469 点, RU 485 点)のガンマ線スペクトルデータを 使用した.第2章で用いたデータからメンテンナンスで発生したドラム缶を除外 したものである.工程で使用されたウランの起源によって NU または RU にラベ ル付けをしている.ガンマ線スペクトルデータは,495 チャネルまでに限定した.

効果的な前処理方法を検討するため、次に示す前処理方法の違う6種類のデー $g d_1 \sim d_6$ を用意した式(3.1) ~式(3.6). 計数値を規格化したデータ(d_1), 計数値 の少ない領域でピークを強調するため、計数値の対数をとって規格化したデータ (d_2),スペクトルの変化を強調するため計数値の差分をとって規格化したデータ (d_3),計数値の少ない領域での変化をより強調するため、計数値の対数をとり、さ らにその差分をとって規格化したデータ(d_4)、変化を強調するため計数値の差分 をとり、さらにその差分をとって規格化したデータ(d_5)、計数値の少ない領域で の変化をより強調するため、計数値の対数の差分をとり、さらにその差分をとっ て規格化したデータ(d_6)を比較した.

d₁:計数値を規格化したもの

$$\boldsymbol{d}_{1} = \left\{ \frac{x_{i}}{\sum_{i=1}^{495} x_{i}} \right\}_{i=1}^{495}$$
(3.1)

d₂:計数値の対数を規格化したもの

$$\boldsymbol{d}_{2} = \left\{ \frac{\log x_{i}}{\sum_{i=1}^{495} \log x_{i}} \right\}_{i=1}^{495}$$
(3.2)

d₃:計数値の差分を規格化したもの

$$\boldsymbol{d}_{3} = \left\{ \frac{x_{i+1} - x_{i}}{\sum_{i=1}^{494} (x_{i+1} - x_{i})} \right\}_{i=1}^{494}$$
(3.3)

d₄:計数値の対数の差分を規格化したもの

$$\boldsymbol{d}_{4} = \left\{ \frac{\log x_{i+1} - \log x_{i}}{\sum_{i=1}^{494} \left(\log x_{i+1} - \log x_{i}\right)} \right\}_{i=1}^{494}$$
(3.4)

d₅:計数値の差分の差分を規格化したもの

$$\boldsymbol{d}_{5} = \left\{ \frac{x_{i+2} - 2x_{i+1} - x_{i}}{\sum_{i=1}^{493} (x_{i+2} - 2x_{i+1} - x_{i})} \right\}_{i=1}^{493}$$
(3.5)

d₆:計数値の対数の差分の差分を規格化したもの

$$\boldsymbol{d}_{6} = \left\{ \frac{\log x_{i+2} - 2\log x_{i+1} - \log x_{i}}{\sum_{i=1}^{493} \left(\log x_{i+2} - 2\log x_{i+1} - \log x_{i}\right)} \right\}_{i=1}^{493}$$
(3.6)

3.3 解析手順

3.3.1 RF の手順

訓練データセット S から,重複を許したサンプリングを行うブートストラップ法 [60, 66] を用いて B 点のデータを抽出し,これを T 回繰り返して,新たな 訓練データセット S_1 , S_2 , …, S_T として再構成する.この新たな訓練データ セットを使い,ルートノードと呼ばれる出発点からある条件を用いて分岐させ る.Fig. 3.1 に示した〇は,ノード(分岐点)で,条件によって下位のノードに 振り分ける.この分岐を繰り返し,ノードのデータが単一のクラスとなれば,そ れを終端ノード(□)としてラベルを付け,決定木を完成させる.この決定木に, 前処理したデータ d を入力して,辿り着いた終端ノードのラベルをその決定木の 分類結果とする.このようにして得られた T 個の決定木の分類結果の多数決に より分類クラスを確定する.

単独の決定木の具体的な作成手順を①~⑤に示す.

①ランダムに選んだ m 個のチャネルまたは相当する序数について, B 点 の訓練データの規格化した計数値や差分などを昇順に並び替える. 隣り合 う 2 つの値の中央値をしきい値の候補とし, $(B-1) \times m$ 個用意する (Fig. 3.2).

②ジニ係数の減少度より、 $(B-1) \times m$ 個のしきい値候補を評価する.

③ジニ係数の減少度が最大となるチャネルまたは相当する序数としきい値 の組を求め,分岐条件とする.

- ④あるノードに属する訓練データと③で求めた分岐条件のしきい値を比較 して, 左右の下位のノードに分ける.分岐したノードについても, それぞれ 同様に分岐を繰り返す.
- ⑤ノードに属するデータのクラスが NU または RU のみになったら,そのノードを終端ノードとし,分岐を止める.終端ノードに行き着いた訓練データのクラスをその終端ノードの分類クラスとし,ラベルを付ける.

①では、NUと RU の違いが計数値やその差分に出ると考えられるため、NU と RU を分類するしきい値の候補として、隣接する値の中央値を用いている.

②,③で選定する最適なしきい値は、分岐前の t 番目のノード t のジニ係数 GI(t) から左右の下位のノードへの分岐時のジニ係数 $GI(t_{\text{Left}}), GI(t_{\text{Right}})$ との 差である減少度 $\Delta GI(t)$ が最大となる序数としきい値の組み合わせであり、ジニ 係数とその減少度は以下の式(3.7),式(3.8) で導かれる.

$$GI(t) = 1 - \sum_{y \in \{\text{NU, RU}\}} p(y|t)^2$$
(3.7)

$$\Delta GI(t) = GI(t) - (p_{\text{Left}}GI(t_{\text{Left}}) + p_{\text{Right}}GI(t_{\text{Right}}))$$
(3.8)

 $p_{\text{Left}}, p_{\text{Right}}$ は、あるノードに属するデータ数に対する、左または右の下位の ノードに属するデータ数の割合である。あるノード t におけるクラス C の事後 確率 p(C|t)は、そのノードに属するデータ数に対するクラス $y = \{\text{NU, RU}\}$ に 属するデータ数の割合 $N_y(t)/N(t)$ で求められる。







ジニ係数が小さいほど、他クラスの要素が少なく純度の高い状態を示し、1に近 づくほど他クラスの要素が多く混ざり合い不純物が多い状態を示す.つまり、あ るノードから下位のノードに分岐する際のジニ係数の減少度が大きいほど、より 分岐先の純度が高くなり、効率よく分類ができる.

分岐条件の探索では、事前に想定していない特徴量が「潜在的な支配パラメー タ」として抽出される可能性もあるため、事前のチャネルの絞り込みは行わない. 全体のフローを Fig. 3.3 に示した.

本解析では,統計解析言語 R[64] にて行い, RF については R の randomForest パッケージ [67] の randomForest 関数を使用した.

3.3.2 検討の手順

全データ 954 点の 1/3 程度を目安に,NU,RU それぞれ 150 点ずつ 300 点の データをランダムに選び,残りの 654 点で分類精度を評価した.300 点は, $d_1 \sim d_6$ の前処理をしたデータに NU または RU のクラス情報を組み合わせ,訓練デー タセット S を作成するためのデータとした.残りの 654 点は, $d_1 \sim d_6$ のデータ の前処理をして,RF の分類結果を実際のクラスと比較して正答率を求めるため のデータとした.

本研究では、訓練データ数によって、分類精度がどう変化するか前処理方法ご とに調べた.事前に準備した 300 点のデータから、2~300 点(1~150 組)まで 訓練データ数を変化させた.それぞれの訓練データ数で、50 回ずつ、訓練データ セット S をランダムに作成し、分類を繰り返して、正答率を平均した.また、訓 練データセット S からブートストラップ法でサンプリングするデータ数 B を訓 練データセット S のデータ数と同数、決定木の個数をT = 500、ランダムに選ぶ チャネルまたは相当する序数の数 m をすべてのチャネルまたは相当する序数の 個数の平方根 $\sqrt{495}$, $\sqrt{494}$, $\sqrt{493}$ より m = 22 と設定した.



3.4 結果と考察

6 種類の前処理方法による正答率を Fig. 3.4~Fig. 3.9 に示す.計数値の対数 を規格化する方法 (d_2)を除き,すべての方法において,平均の正答率は使用する 訓練データ数の増加に伴い収束している. d_1, d_3, d_4, d_5 では,99%以上の平均正 答率となり,NUまたは RU への分類に適用可能であることが確認できた.特に, 計数値の対数の差分を規格化する方法 (d_4)では,訓練データ数が 186 点以上で, 正答率がすべて1となり,6 種類の前処理方法の中で最適な前処理方法といえる.



Fig.3.4 Accuracy with d_1 .



Fig.3.6 Accuracy with d_3 .



Fig.3.8 Accuracy with d_5 .



Fig.3.9 Accuracy with d_6 .

300 点の訓練データを用いて RF を実行したときのジニ係数の減少度の平均を 相対度数 [68] で Fig. 3.10~Fig. 3.15 に示す. 差分を取らない $d_1 \ge d_2$ は, 400 チャネル以降 (1600keV 相当) に減少度の大きいチャネルがある. これらの前処 理方法では、少ない訓練データ数で他の方法に比べて正答率が低い. ウラン量が 少ないウラン廃棄物ドラム缶のデータでは、ウラン由来の主要なピークが小さく なり、相対的に、440 チャネル付近の ²¹⁴Bi など、バックグラウンド由来のピー クが大きく見える. そのため、ウラン由来のピーク以外に 400 番目以降のチャネ ルに減少度が大きいものが出現し、正答率も低くなったと考えられる. d_3, d_4 で は、400 番目以降のチャネルの減少度の大きいものがなくなっており、差分を取る ことでバックグラウンドの影響が抑えられることがわかった. 正答率が一番高い d_4 では、他の方法に比べ、511 keV、583 keV などの RU に特有のエネルギーピー ク付近で減少度が大きいところがあるため、他の方法より核種に応じた分類にな り、正答率が上がったと考えられる.

また, Fig. 3.13 における d₄ のジニ係数の減少度の平均を見ると, 50 チャネル付

近 (200keV 相当) より前のピークも重要であると認識していることがわかる.想定していたトリウム系列だけでなく,特性 X 線のピークも分類に使われており, 事前に想定していなかった「潜在的な支配パラメータ」が抽出されている.



Fig.3.10 Mean decrease in Gini index with d_1 .



Fig.3.11 Mean decrease in Gini index with d_2 .



Fig.3.12 Mean decrease in Gini index with d_3 .



Fig.3.13 Mean decrease in Gini index with d_4 .



Fig.3.14 Mean decrease in Gini index with d_5 .



Fig.3.15 Mean decrease in Gini index with d_6 .

3.5 結言

ウラン廃棄物ドラム缶の2クラス分類について,低解像度ガンマ線スペクトル データに対する RF の適用可能性を検討した.前処理方法が異なる6種類のデー タセットに対して RF を実行し,最適な前処理方法を検討した.その結果,計数 値を規格化する方法,計数値の差分を規格化する方法,計数値の差分の差分を規 格化する方法,計数値の対数の差分を規格化する方法の4種類の方法で,50回試 行の平均の正答率が99%以上に収束し,2クラス分類が可能であることが確認で きた.特に,計数値の対数をとり,前後の差分を取る前処理方法が一番高い正答率 となった.この方法では,訓練データ数が186点以上のとき,正答率はすべて1 になり,6種類の前処理方法の中で最適な前処理方法といえる.SVMに比べ,訓 練データ数が多く必要となるものの,RF でも2クラス分類が可能であると考え られる.

第4章

3 クラス分類に対する SVM と RF の分類性能の比較

4.1 緒言

第2章,第3章では、ウラン廃棄物ドラム缶からの低解像度ガンマ線スペクト ルデータを用いた分類について SVM と RF の適用可能性を確認した.これまで の検討では、ウランの起源によりウラン鉱石由来の NU と回収燃料に由来する RU の 2 クラスに分類した.同一工程から発生したドラム缶のガンマ線スペクト ルデータを機械学習させることで、ガンマ線エネルギーの事前知識、ピークサー チや遮蔽計算などを必要とせずに分類できることを示した.

本章では、実際の廃棄物管理において3クラス以上への分類が考えられるため、 これまでのNUおよびRUの2クラスに加え、イエローケーキからウランを精製 する際に不純物として発生したラジウムを含むドラム缶も分類対象とし、3クラ スの分類についてRFとSVMの性能比較を行う.

4.2 適用データ

JAEA 人形峠センター製錬転換施設の廃水処理工程で発生した中和澱物を封入 したドラム缶とイエローケーキ溶解工程で発生した珪藻土の入ったドラム缶のガ ンマ線スペクトルデータを用いた.廃水処理工程で発生したドラム缶は,第3章 で用いたデータと同一であり,ウランの起源によってNUまたはRUにラベル付 けをしているものである.珪藻土は,イエローケーキを酸で溶解した溶液の不純 物を吸着させるために使われたもので、その核種組成は NU をベースとしている ものの、ラジウムも多く含む. Table 4.1 に対象ドラム缶の詳細データを示す.便 宜上、分類のラベルとして、ラジウムが多く含まれている NU を Ra とする.

Class	Process	Main com-	Moisture	Mean gross	Number of
		ponent	content	weight^*	drums
NU	Waste water	CaF_2	50 %	195 kg	469
	treatment				
RU	Waste water	CaF_2	50~%	228 kg	485
	treatment				
Ra	Dissolving of	SiO_2	45 %	174 kg	83
	yellow cake				

Table4.1 Characteristics of uranium-bearing waste drums used for tests.

^{*}including drum weight

 Q^2 system で取得したガンマ線スペクトルデータの NU, RU および Ra の典 型例を Fig. 4.1 に示す. 3 クラス共通で, ^{234m}Pa 由来の 1001 keV, 766 keV と ²³⁵U 由来の 186 keV が見られる. RU のスペクトルでは, 原子炉内で生成される ²³²U, ²³⁶U の子孫核種であるトリウム系列の ²¹²Pb と ²⁰⁸Tl 由来の 239 keV, 511 keV, 583 keV, 861 keV のピークが見られる. Ra のスペクトルでは, ²²⁶Ra の子 孫核種である ²¹⁴Pb と ²¹⁴Bi 由来の 352 keV, 609 keV, 1120 keV, 1765 keV の ピークが確認できる. また, ²¹⁴Pb 由来の 242 keV, 295 keV についてはピーク が明瞭には確認できないものの,約 55~80 チャネルの範囲に NU および RU と は異なる形状が観察できる.



Fig.4.1 Typical spectra of NU, RU and Ra with NaI(Tl) detectors on Q^2 system.

4.3 解析手法と手順

4.3.1 入力データの前処理方法

RF への入力データ d は、Q² system から得られるチャネル (1~495 チャネ ル)と計数値を第3章で最適とされた方法 (式 (4.1))で前処理したものである. SVM についても、NU と RU への 2 クラス分類において、第3章と同じ 6 種類 の前処理方法に対し訓練データ数を変えながら 50 回ずつ分類し、比較した.6種 類の前処理方法とは、1)計数値を規格化、2)計数値の対数を規格化、3)計数値 の差分を規格化、4)計数値の対数の差分を規格化、5)計数値の差分の差分を規格 化、6)計数値の対数の差分の差分を規格化する方法である。各方法を比較した結 果、NU と RU への 2 クラス分類に対して、RF での結果と同様に、4)計数値の 対数の差分を規格化する方法 (式 (4.1))で一番精度が良くなることを確認したた め、SVM でも同様の前処理を行った (付録 D 参照).

$$d = (x_1, x_2, \dots, x_{494})$$

$$= \left\{ \frac{\log v_{n+1} - \log v_n}{\sum_{n=1}^{494} (\log v_{n+1} - \log v_n)} \right\}_{n=1}^{494}$$
(4.1)

ここで、 v_n はn番目のチャネルの計数値を表し、 x_n は規格化された計数値の 対数の差分値を表す.

4.3.2 訓練データ

訓練データには、入力データ d にクラス情報 $y \in \{NU, RU, Ra\}$ を組み合わせ る. NUと RU への2クラス分類に対する前処理方法の比較の際、SVM では各ク ラス5点、計 10 点で正答率が安定し、RF では各クラス 25点、計 50 点程度で正 答率が安定していく傾向が見られた(Fig. 4.2). そのため、それぞれのクラスか らランダムに 25 点のデータをサンプリングし、残りのデータ(NU 444 点、RU 460 点、Ra 58 点)で正答率を評価する.



Fig.4.2 Variation in the mean of accuracies depending on the number of training data.

4.3.3 RF の手順

RFは、与えた訓練データセットから、ランダムに重複を許してサンプリングし たデータで多数の訓練データセットを再構成し、それぞれの訓練データセットに 対し、機械学習で求めた分岐条件で、次々に分岐を繰り返し、決定木と呼ばれる樹 木状の分類モデルを作成して分類する.単独の決定木では分類精度は低いが、複 数の決定木の多数決をとることによって、分類精度を高める手法である[60,68]. ガンマ線スペクトルデータに適用する際の詳細な方法については第3章を参照さ れたい.

分岐条件算出に使用される3クラス分類のためのジニ係数の定義を式(4.2)に,

上位から下位へ分岐したときの減少度を式(4.3)に示す.ジニ係数が小さいほど 純度の高い状態を示し、単一クラスに近づくため、ジニ係数の減少度が大きいほ ど効率の良い分岐条件となる.なお、t はノード(分岐点)の番号を表し、t_{Left} お よび t_{Right} は左右の下位のノードを示す.

$$GI(t) = 1 - \sum_{y \in \{NU, RU, Ra\}} p(y|t)^2$$
 (4.2)

$$\Delta GI(t) = GI(t) - \left(p_{\text{Left}}GI(t_{\text{Left}}) + p_{\text{Right}}GI(t_{\text{Right}})\right)$$
(4.3)

 $p_{\text{Left}}, p_{\text{Right}}$ は、左右の下位のノードに属するデータ数/ノードtに属するデータ数で求められる。p(y|t)は、ノードtに属するデータ数に対するクラスyに属するデータ数の割合 $N_y(t)/N(t)$ で求められる。

3クラスの単独決定木の作成例を Fig. 4.3 に示す.まず,訓練データをルート ノード(出発点)に入力する.ノードtにおいて,データの序数nをランダムに複 数選ぶ.選んだ序数ごとに,差分値 x_n を昇順に並べ替える.並べ替えた後に,隣 り合う差分値の中間値をしきい値候補とし,ジニ係数の減少度が一番大きくなる しきい値 a_t と n_t との組み合わせを分岐条件とする.その序数 n_t での差分値 x_{n_t} について, a_t と比較して下位の左右ノードに分ける.単一クラスになるまでこの 分岐を続け,単一クラスとなったノードを終端ノードとし,分岐を止める.この作 業をT回繰り返し,T本の決定木を作成する.

T本の決定木を作成した後,すべての決定木に分類対象のデータ *d* を入力し, 行き着いた終端ノードのクラスの多数決によって分類する (Fig. 4.4).



Fig.4.3 Example of single decision tree growth for 3 classes (NU, RU, RU, Ra). Training data are divided into 2 child nodes by comparing between the variable x_{n_t} and the threshold a_t on each node until the node reaches the terminal node shown by squares.



4.3.4 SVM の手順

SVM は, パターン認識の機械学習法として, Cortes と Vapnik によって開発された, 2 クラス分類のためのアルゴリズムである [48]. 詳細は, 第 2 章を参照されたい.

SVM で解くべき 2 次計画法問題を式 (4.4) に示す. n 個の説明変数からなる i番目の訓練データ d_i とクラス値 $y_i \in \{+1, -1\}$ をもつ M 個の訓練データセット $S = \{d_i\}_{i=1}^M$ を用いて,目的関数 $Q(\alpha)$ を最大にする α を求める.

$$Q(\boldsymbol{\alpha}) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} \alpha_i \alpha_j y_i y_j K(\boldsymbol{d}_i, \boldsymbol{d}_j)$$
$$\sum_{i=1}^{M} y_i \alpha_i = 0, \qquad 0 \le \alpha_i \le C \qquad (4.4)$$

s.t.

ここで、 α_i はラグランジュの未定乗数であり、C は訓練データのマージンへの 入り込みの許容量を調整するパラメータである。カーネルは、ラジアル基底関数 カーネル $K(d_i, d) = \exp(-\gamma || d_i - d ||^2)$ を用いた。なお、 γ は、ガウス分布の 半径をコントロールするパラメータであり、サポートベクターのカバーする領域 の広さを調整するパラメータである。

次に,式(4.5)で定義される,入力データ*d*に対する決定値*D*と呼ばれる分類の指標を導入する[48,69].

$$D(\boldsymbol{d}) = \sum_{i \in I} \alpha_i y_i K(\boldsymbol{d}_i, \boldsymbol{d}) + b$$
(4.5)

ここで,集合 I は,訓練データから得られるサポートベクターのインデックス を示し, b はバイアス項である.

分類対象のデータ d を式(4.5)に入力し決定値 D を求める. D > 0ならば、 「+1」、D < 0ならば「-1」のクラスに分類される. なお、D=0の超平面が境界 面となる.

3 クラスへの分類に対し、2 クラス分類の SVM をそのまま使うことはできない。 多クラス分類の SVM として、One versus One (1 vs 1) と呼ばれる 2 クラ スずつ分類する方法(以下、SVM (1 vs 1)) と One versus Rest (1 vs R) と呼 ばれる1クラスと残りのクラスに分類する方法(以下, SVM (1 vs R))があり, 両方法の比較を行う.本研究において, SVM (1 vs 1)では, NUと RU, NUと Ra, RUと Ra について決定値を求め,絶対値が最大となる決定値が含まれる対 における2クラスの分類結果から分類クラスを求めた.SVM (1 vs R)では, NU と残りのクラス (RUと Ra), RUと残りのクラス (NUと Ra)と Raと残りのク ラス (NUと RU) について,訓練データの単独クラス側のクラス値を1とし,決 定値を求めた.3値を比較し決定値が最大になるクラスを分類クラスとした.

4.3.5 解析パラメータ

RF におけるブートストラップ数 *B* を訓練データセット *S* のデータ数と同数と し、ランダムに選ぶ序数の数 *m* をすべての序数の数の平方根 $\sqrt{494}$ より *m* =22 とし、決定木の本数 *T* をパッケージのデフォルト値より *T* = 500 とした.

SVM における γ, C について, $\gamma = 10^{-5} \sim 10^5, C = 10^{-2} \sim 10^5$ の範囲でパラ メータサーベイを行った (Fig. 4.5). このパラメータサーベイでは, Leave-One-Out 交差検証法 [29, 48] を用いて, エラー率を求め, 最小となるパラメータの組 み合わせを最適なパラメータとした. $\gamma = 0.001, C = 100$ のとき, エラー率が 0.00416 と最小となり, この γ, C の組み合わせで SVM の学習を行った.

サンプリングした訓練データを RF ではブートストラップ法で再構成して学習 し, SVM ではそのまま学習し,残りのデータを分類するという作業を繰り返した. 正答率は,クラスごとに機械学習による分類が正解したデータ数/元ラベルのデー タ数で算出した.繰り返し回数による正答率の平均の変化を Fig. 4.6~Fig. 4.8 に示す. 6000 回を超えると正答率が落ち着いたため,繰り返し回数 10000 回での 結果で評価することとした.

本解析は,統計解析言語 R[64] にて行い, RF では randomForest パッケージ [67], SVM では e1071 パッケージ [65] を使用した. なお, 1回の分類モデルの作 成および分類実行にかかる時間は, RF では 0.8 秒, SVM (1 vs 1) では 1.1 秒, SVM (1 vs R) では 1.0 秒となりいずれも迅速に分類できた.



Fig.4.5 Result of parameter survey for SVM. The dark area is the high error rate area, and the white area is the low error rate area. The circle at $\gamma = 0.001$, C = 100 shows the lowest error rate, which is 0.00416.



Fig.4.6 Variation in the mean of accuracies depending on repeat times ((a) NU).



Fig.4.7 Variation in the mean of accuracies depending on repeat times ((b) RU).



Fig.4.8 Variation in the mean of accuracies depending on repeat times ((c) Ra).

4.4 結果と考察

10000回繰り返し分類した際の正答率の平均値,中央値,標準偏差および最小値 を Table 4.2 に,正答率の分布を Fig. 4.9~Fig. 4.11 に示す.また,この分布を踏 まえ,分類クラスごとに等分散性を仮定しない2群の検定であるウェルチの t 検 定を行った.その結果,すべての比較において p 値が有意水準 5% 下回り,有意 な差があった.

NUへの分類については, SVM (1 vs 1)の正答率の中央値が 1.000 と最も高く, 標準偏差も SVM (1 vs 1) が最も小さいが,最小値は RF の方が 0.016 良い. NU への分類では, SVM (1 vs 1) もしくは RF の正答率が良いと考えられる. RUへ の分類については, RF と SVM (1 vs 1)の正答率の中央値が 1.000 と最も高く, 標準偏差は RF が最も小さかった.また,最小値も RF が一番良かったため, RU への分類では, RF の正答率が良いと考えられる.Raへの分類については,SVM (1 vs R)の正答率の中央値が 0.993 と一番大きく,標準偏差も一番小さいが,最 小値は RF の方が 0.028 良い.Raへの分類は, RF もしくは SVM (1 vs R)の 正答率が良いと考えられる.以上の結果をまとめると,Table 4.3 の通りである. 機械学習による分類を 10000 回実行したうち,データごとに誤分類となった回

Class	Method	Accuracy			
		Mean	Median	Standard	Minimum
				Deviation	
NU	RF	0.998	0.998	0.002	0.966
	SVM $(1 \text{ vs } 1)$	1.000	1.000	0.001	0.950
	SVM $(1 \text{ vs } R)$	0.983	0.983	0.019	0.845
RU	RF	0.999	1.000	0.002	0.989
	SVM (1 vs 1)	0.996	1.000	0.007	0.921
	SVM $(1 \text{ vs } R)$	0.996	0.983	0.007	0.926
Ra	RF	0.978	0.983	0.025	0.845
	SVM $(1 \text{ vs } 1)$	0.975	0.983	0.024	0.810
	SVM $(1 \text{ vs } R)$	0.988	0.993	0.016	0.817

Table 4.2 Accuracy of each classification method.



Fig.4.9 Distribution of the classification accuracies. White dots show the medians of accuracies, black boxes show the quantiles, and gray areas show the densities of accuracies. ((a) NU)



Fig.4.10 Distribution of the classification accuracies. ((b) RU)



Fig.4.11 Distribution of the classification accuracies. ((c) Ra)

Method	NU	RU	Ra
RF	А	А	В
SVM(1 vs 1)	А	В	С
SVM(1 vs R)	С	С	В

Table4.3 Evaluation of comparison with 3 methods.

A: Excellent, B: Fair, C: Not good

数を数え,1回以上誤分類となったデータのヒストグラムを Fig. 4.12~Fig. 4.14 に示す.このヒストグラムでは、右へ向かうほど同じデータが繰り返し誤分類さ れ、左に向かうほど誤分類の偶然性が増すと考えられる.1~250回誤分類となっ た左端の区間で,RFの誤分類データ数が他に比べて半分以下となっている.一度 でも誤分類となったデータの総数は、RF では54 点、SVM (1 vs 1) では110 点、 SVM (1 vs R) では245 点となり、RF が総数でも一番少なく、他の方法に比べ 安定して分類できると考えられる.



Fig.4.12 Distribution of the number of misclassified times (excluding the data that were not misclassified). The right top numbers show the total number of data misclassified. ((a) RF)



Fig.4.13 Distribution of the number of misclassified times. ((b) SVM (1 vs 1))



Fig.4.14 Distribution of the number of misclassified times. ((c) SVM (1 vs R))

一度でも誤分類されたデータについて、3 法でどのような関係性があるかをベン図 (Fig. 4.15) に示した.何度も繰り返し誤分類されたデータには、手法ごとの誤分類の傾向や特徴が現れると考え、100 回以上誤分類されたデータに絞り込み、データを分析した.

3 法共通で誤分類となった領域 D は,²²⁶Ra の子孫核種由来の微弱なピークが 含まれるデータであった。10 点が NU から Ra に誤分類となり,8 点が Ra から NU に誤分類となった*.訓練データに代表性がないものが混入しても通常は淘 汰され分類モデルに影響はでないが,今回の検討では訓練データ数が少なかった ため,NU の訓練データの中に入り込んだ²²⁶Ra の子孫核種由来の微弱なピーク の影響があった場合に,Ra から NU もしくは NU から Ra に誤分類されたと考 えられる.

領域 B の 42 点中 18 点も ²²⁶Ra の子孫核種由来の微弱なピークが含まれるデー タで,13 点が NU から Ra に,5 点が Ra から NU に誤分類となった.SVM で誤 分類となった領域 A の 39 点および領域 B の 42 点中 23 点は,RU から NU また は Ra に誤分類となった.これらのデータは,1001 keV に相当するチャネルが他

^{* 8} 点中 1 点については RF で 1 回のみ NU から RU に誤分類されたが,頻度が少なく理由は不 明である.



Fig.4.15 Classification of misclassified data.

のスペクトルデータに比べて右に 10 チャネル程度シフトしていた (Fig. 4.16). SVM の両法はチャネルすべてを使用するため、チャネルシフトの影響が大きく、 RF はランダムに選んだチャネルの組み合わせを使うため、チャネルシフトの影 響が小さかったと考えられる. RF のみで誤分類となった領域 F の 5 点のうち、 元ラベルが RU のデータ 4 点はいずれも ²³²U, ²³⁶U の子孫核種由来のピークが 弱いため、NU または Ra に誤分類したと考えられる.

領域 B の 1 点および領域 F の 1 点は, ²²⁶Ra 量が多く,明瞭なピークが多 数出たため,Ra から RU に誤分類したと推察される.なお,RF と SVM (1 vs R)に共通して誤分類となったデータは存在しなかった (領域 G). 誤分類の特徴 から,SVM ではチャネルシフトの影響があり,特に,SVM (1 vs R) では影響が 大きかった.通常の測定では定期的にゲイン調整することでチャネルシフトを防 ぐが,万が一チャネルシフトがあった場合でも,RF では,チャネルシフトに影響 されず,誤分類データ数が少ない安定した分類をし,人間が付けたラベルに近い 分類を行う傾向が見られた.



Fig.4.16 Example of shifted spectrum.

4.5 結言

ウラン廃棄物ドラム缶 (NU 469 点, RU 485 点, Ra 83 点)の低解像度ガンマ線 スペクトルデータを用いて, RF と SVM (1 vs 1 および 1 vs R) による3クラス 分類の性能比較を行った. SVM の両法は, 分類モデルの性質からチャネルシフト の影響を受けやすいことがわかった. 今回の訓練データ数は少なかったが繰り返 し計算することにより, RF でより安定した分類をする傾向を見ることができた. 今後は, 放射性核種が微量でも正確に分類できる手法を訓練データの選び方も含 めて検討する必要がある.
第5章

結論

5.1 本研究の成果

本研究では、ウラン廃棄物ドラム缶を NaI (Tl) シンチレーション検出器で測定 した低解像度ガンマ線スペクトルデータの機械学習による分類手法について検討 した.機械学習法のうちパターン認識手法として発展してきた SVM と RF をガ ンマ線スペクトルデータに適用する手法を提案し、性能について評価した.本研 究の方法は、学習するための訓練用のスペクトルデータを用意するのみで、ウラ ン廃棄物ドラム缶の内容物や収納状況、放出されるガンマ線エネルギーなどの事 前情報を必要としないという利点がある.

最初の段階として,まず,内容物の主成分が同様で,含まれる放射性核種の組成 が異なるウラン廃棄物ドラム缶のガンマ線スペクトルデータを NU と RU の 2 ク ラスへ分類する方法について検討した.廃水処理工程から発生するウラン廃棄物 ドラム缶 967 点のうち 12 点を訓練データとして,SVM の分類モデルを作成した. 残りの 955 点について,作成した分類モデルで分類したところ,4 点のデータが 元々付けられていたラベルとは違う分類結果となった.うち1 点は,元のラベル が間違っており,3 点は,操業中の慣例にならった保守的なラベルの付け方がされ ていた.元のラベルと SVM の分類結果は異なったもののいずれも科学的には正 しく分類できており,NU と RU の 2 クラスへの分類に対し,SVM が適用可能で あることを確認した.

次に,同様の分類について, RF の適用を検討した.メンテナンスで発生したド ラム缶を除いたウラン廃棄物ドラム缶 954 点について 300 点を訓練データとして ガンマ線スペクトルデータの前処理方法が異なる6種類のデータセットに対して RFを実行し,最適な前処理方法を検討した.残りの654点で正答率を算出した結 果,計数値の対数をとり,前後の差分を取る前処理方法が一番高い正答率となり, 6種類の前処理方法の中で最適な前処理方法となった.SVMに比べ,訓練データ 数が186点と多く必要になるものの,RFでもNUとRUの廃棄物への分類が可 能であることを確認した.

最後に、NU、RU およびラジウムが多く含まれる廃棄物について、SVM と RF の分類性能を評価した.その結果、いずれの方法でもラジウムが微量に含まれる 場合に NU とラジウムの誤分類が多く、SVM は分類モデルの性質からチャネル シフトの影響を受けやすいことが判明した.一方、RF はチャネルシフトに対して も安定した分類をすることができるものの、微量に含まれる核種による誤分類問 題の解決が今後の課題となった.

これらの手法は, SVM から得られる決定値や RF から得られる多数決の値など の明確な数値的根拠をもって分類することができ, SF 法の補助ツールとして期待 できる. どちらの方法でも分類速度は, ドラム缶 1000 本当たり数秒程度であるた め, 大量の廃棄物の分類に対して有用である.

5.2 今後の展開

放射性廃棄物の処理計画策定や放射性廃棄物処分場での確認において,核種の 種類や濃度などの放射能インベントリ情報が重要となる.すべてを化学分析する ことができないため,原子力施設の廃止措置で放射能インベントリの推定は,どの ラインで発生した廃棄物なのか,どのような運転で発生した廃棄物なのかといっ た,廃棄物の発生履歴を根拠としたものが主である.ガンマ線スペクトルデータ はドラム缶に廃棄物が収納された状態であれば,容易に取得が可能である.発生 履歴にガンマ線スペクトルデータから得られた核種組成グループ情報を付与する ことができれば,発生履歴に頼ったエビデンスに比べ,より客観的なエビデンス を提供できることが期待できる.

本研究で行った3クラス分類では, 訓練データをランダムに選んだため, 適切で ない訓練データが混ざっており分類精度が落ちた. 事前に訓練データからミスラ ベルデータや外れ値データを排除することによってより高い精度の分類ができる と考えられる.訓練データの決定値や多数決などの値の分布を求め,訓練データ を再抽出する方法など,外れ値を検出するアルゴリズムを用いて事前に訓練デー タの妥当性評価をする必要がある.

本研究の手法は,原理的にウラン廃棄物に限らず一般の放射線スペクトルの分 類に応用できる.研究施設等から発生する放射性物質などについても適用が期待 できる.また,事前に放射線スペクトルの機械学習をすることによって,不明な核 物質の発生元を特定する核鑑識への応用も期待できる.

付録A

対象データの主成分分析

機械学習法によっては、データの次元が少ない方が良い場合もある. 放射性ス ペクトルデータに対し、ニューラルネットワークで機械学習させる研究では、次 元を縮約して学習させることが多い. ここでは、主成分分析を用いて次元の縮約 について検討する. 主成分分析では、共分散行列の固有値問題を解き、大きい固 有値を持つ固有ベクトルで張られる部分空間を構成する [28, 70].

第4章で用いたスペクトルデータについて,そのデータの傾向を調べるために 主成分分析を行った.14点のスペクトルデータを主成分分析し,NU 469点,RU 485点,Ra 83点について射影したものを示す.

主成分分析には、統計解析言語 R[64] の標準組み込みパッケージ stats 中の princomp を使用した.また、カーネル主成分分析については、kernlab パッケー ジ [71] 中の kpca 関数を用いた.なお、カーネルにはラジアル基底関数を用い、 分散 $\sigma = 10^1, 10^2, 10^3, 10^4, 10^5$ とした.結果を Fig. A.1~Fig. A.6 に示す.NU は青、RU は赤、Ra は緑としてプロットした.

主成分分析とカーネル主成分分析の $\sigma = 10, 10^2$ は似た傾向を示すが、 σ が大 きくなるにつれて、広がりが小さくなっている。Fig. A.1 における PC1 vs PC3、 PC2 vs PC3 では、NU、RU、Ra が比較的クラスごとに分かれてプロットされて いる。また、Fig. A.7 に主成分分析によって主成分に射影する際の各チャネルの 値(主成分スコア)を示す。PC1 は NU のスペクトルを示し、PC2、PC3 で、主 に ²³²U、²³⁶U、²²⁶U 由来のスペクトルを拾っている。 $\mathbf{73}$





-0.1 0.0 0.1 0.2 0.3



8.0 4.0 0.0 4.0-



0 τ Ţ-







Fig.A.7 Principal component score $(\mathrm{PC1}{\sim}\mathrm{PC4})$.

付録 B

対象データのニューラルネット ワークによる分類

ここではニューラルネットワークを用いて分類を試みる. ニューラルネット ワークとは,神経細胞を模した形式ニューロンと呼ばれる判別器を階層的に結合 させたパターン認識モデルである [28]. 付録 A で行った主成分分析で次元削減 したデータに対し,統計解析言語 R[64] の標準組み込みパッケージ nnet を使用 し分類を行った.



Fig.B.1 Result of parameter survey for neural network.

 $\mathbf{81}$

分類を行う前に、ランダムに選んだ各クラス 80 個ずつのデータを用いて、10 分割交差検証法を 10 回ずつ繰り返し、最適なパラメータを求めた.入力層と出力層 の間にある隠れユニットが 4, weight decay が 10^{-4} のとき正答率が一番高くなった (Fig. B.1).

パラメータサーベイ結果を用い,第4章と同一の訓練データで,300回繰り返 し分類した. 正答率の分布を Table B.1 および Fig. B.2~Fig. B.4 に示す.

これらの結果より, 主成分分析を適用したデータに対し, ニューラルネットワークにて分類した場合, 正答率が低いことがわかった. NU, RU, Ra への分類では, ニューラルネットワークは適していないと考えられる.

Class -	Accuracy					
	Average	Median	Variance	Minimum		
NU	0.678	0.702	0.040	0.090		
RU	0.896	0.951	0.017	0.041		
Ra	0.794	0.810	0.012	0.069		

TableB.1 Accuracies of all class.



Fig.B.2 Distribution of NU accuracy by neural network.



Fig.B.3 Distribution of RU accuracy by neural network.



Fig.B.4 Distribution of Ra accuracy by neural network.

付録C

対象データのクラスター分析

ここでは、教師なし学習法のひとつであるクラスター分析を行う. データの分 布が正規分布の足し合わせに従うと仮定する混合正規分布モデルを用いてクラス ター分析を行う. このモデルに用いる正規分布のパラメータや正規分布の混合比 は EM アルゴリズムを用いて推定する [29, 70, 72]. 解析には統計解析言語 R[64] の mclust パッケージ [73] を使用した. 付録 A で行った主成分分析, カーネル主 成分分析 ($\sigma = 10^5$)の結果に対し,混合正規分布モデルを用いたクラスター分析 を行った結果を示す. 最適な混合正規分布モデルのモデルとクラスター数を求め るため, それぞれの情報量 BIC (Bayesian Information Criterion)を計算した結 果を Fig. C.1 (主成分分析), Fig. C.2 (カーネル主成分分析)に示す.

主成分分析については、分散に制約がなく、すべての要素モデルで形が同じとな る VEV モデルで、クラスター数6のとき BIC が最大となった.比較のため、実 際のクラス数3の場合についても、VEV モデルでクラスター分析を行った. PC2 と PC3、PC3 と PC4 での分布を Fig. C.3~Fig. C.8 に、対応表を Table C.1 およ び Table C.2 に示す.カーネル主成分分析については、分散に制約がなく、すべて の要素モデルで体積・向きが等しい EVE モデルで、クラスター数2のとき BIC が 最大となった. PC1 と PC2 での分布について元のラベルの分布を Fig. C.9 に、 クラスター分析の結果を Fig. C.10 に、対応表を Table C.3 に示す.

以上の結果より, 主成分分析を前処理とすると, RU とそれ以外への分類は比較 的うまくいくが, NU と Ra の分類が困難であることがわかった. カーネル主成分 分析の場合は, さらに NU と RU の分離も難しいことがわかった.



Fig.C.1 BIC of Gaussian mixture model for the result of PCA.



Fig.C.2 BIC of Gaussian mixture model for the result of kernel PCA.



Fig.C.3 Distribution of PCA data with original label (PC2, PC3).



Fig.C.4 Result of cluster analysis for PCA (PC2, PC3) (Cluster number 6).



Fig.C.5 Result of cluster analysis for PCA (PC2, PC3) (Cluster numbe 3) .



Fig.C.6 Distribution of PCA data with original label (PC3, PC4).



Fig.C.7 Result of cluster analysis for PCA (PC3, PC4) (Cluster number 6) .



Fig.C.8 Result of cluster analysis for PCA (PC3, PC4) (Cluster number 3).



Fig.C.9 Distribution of kernel PCA data with original label (PC1, PC2).



Fig.C.10 $\,$ Result of cluster analysis for kernel PCA (PC1, PC2).

		Estimated class					
		1	2	3	4	5	6
Original	Ra	0	0	5	0	0	78
label	NU	166	0	8	0	270	25
	RU	0	186	15	284	0	0

TableC.1 Result of cluster analysis for PCA (Cluster number 6) .

TableC.2 Result of cluster analysis for PCA (Cluster number 3) .

		Estimated class		class
		1	2	3
Original	Ra	0	83	0
label	NU	0	148	321
	RU	485	0	0

TableC.3 Result of cluster analysis for kernel PCA (Cluster number 2) .

		Estimated class	
		1	2
Original	Ra	22	61
label	NU	415	54
	RU	0	485

付録 D

SVM におけるデータの 前処理方法の比較

第3章で RF の最適な前処理方法について比較検討を行った. 同様に SVM に 対しても最適な前処理方法について比較した. 4.3.1 で述べたように, $d_1 \sim d_6$ に ついて比較を行った.

SVM は, e1071 パッケージ [65] の svm 関数を使用し, カーネルにラジアル基 底関数を指定し, $\gamma = 0.001$, C = 10 として計算を行った. 訓練データセットと して事前に準備した 300 点のデータから 2~300 点 (1~150 組) まで変化させた. それぞれ, 50 回ずつ, 訓練データセットをランダムに作成し, 分類を繰り返して, 正答率を平均した. 結果を Fig. D.1~Fig. D.6 に示す.

d₄では, 訓練データ数が10組(20点)以上のところですべて正答率が1となった. この結果より, SVM においても RF と同様に, 計数値の対数をとり, 前後の 差分を取る前処理方法が一番適していると考えられる.



Fig.D.2 Accuracy with d_2 .



Fig.D.4 Accuracy with d_4 .



Fig.D.6 Accuracy with d_6 .

参考文献

- [1] 閣議決定. エネルギー基本計画, 2014.
- [2] 経済産業省. 長期エネルギー需給見通し, 2015.
- [3] 日本原子力文化財団. 原子力・エネルギー図面集, 2016.
- [4] International Atomic Energy Agency. Classification of Radioactive Waste. IAEA Safety Standards for protecting people and the environment, General Safety Guide, No. GSG-1, 2009.
- [5] 長崎晋也,中山真一. 原子力教科書 放射性廃棄物の工学. オーム社,東京, 2011.
- [6] 原子力規制委員会. 炉内等廃棄物の埋設に係る規制の考え方について, 2016.
- [7] International Atomic Energy Agency. IAEA Safety Glossary: 2016 Revision. Vienna, 2016.
- [8] International Atomic Energy Agency. Radiation Protection and Safety of Radiation Sources: International Basic Safety Standards. IAEA Safety Standards for protecting people and the environment, General Safety Requirements Part 3, No. GSR-3. Vienna, 2014.
- [9] 経済産業省令第112号. 製錬事業者等における工場等において用いた資材その他の物に含まれる放射性物質の放射能濃度についての確認等に関する規則, 2005.
- [10] 文部科学省令第 49 号. 試験研究の用に供する原子炉等に係る放射能濃度に ついての確認等に関する規則, 2005.
- [11] 経済産業省原子力安全・保安院.原子力施設における「放射性廃棄物でない 廃棄物」の取り扱いに関するガイドライン,2008.
- [12] 原子力委員会. ウラン廃棄物処理処分の基本的考え方について, 2000.

- [13] 原子力安全委員会.研究所等から発生する放射性固体廃棄物の浅地中処分の 安全規制に関する基本的考え方,2006.
- [14]「東京電力福島第一原子力発電所事故以降の低レベル放射性廃棄物処理処分の在り方」特別専門委員会.低レベル放射性廃棄物処分におけるウランの扱いについて-浅地中トレンチ処分に係る規制への提言-平成 26 年度報告書. 日本原子力学会, 2015.
- [15] 長谷川信, 齋藤龍郎, 財津知久, 佐藤和彦, 坂井章浩, 麓弘道. 北米地域のウラン廃棄物処分に関する調査; 米国ユタ州, テキサス州およびカナダオンタリオ州における処分および規制の現状. JAEA-Review 2013-043. JAEA, 2013.
- [16] 佐藤和彦,坂井章浩,秦はるひ,麓弘道,川越浩,齋藤龍郎,長谷川信.欧州地域のウラン廃棄物処分に関する調査;フランス,英国及びスウェーデンにおける処分及び規制の現状.JAEA-Review 2014-006. JAEA, 2014.
- [17] 齋藤龍郎,坂井章浩,佐藤和彦,八木直人,秦はるひ,麓弘道,川越浩,長谷川 信.ウラン廃棄物に着目した諸外国の放射性廃棄物処分の情報整理.JAEA-Review-2014-021. JAEA, 2014.
- [18] 国立天文台. 理科年表 平成 23年. 丸善株式会社, 東京, 2010.
- [19] 日本原燃株式会社,独立行政法人日本原子力研究開発機構,株式会社グローバル・ニュークリア・フュエル・ジャパン,三菱原子燃料株式会社,原子燃料工業株式会社,株式会社ジェー・シー・オー.ウラン廃棄物の処分およびクリアランスに関する検討書.2006.
- [20] 日本アイソトープ協会. アイソトープ手帳. 丸善株式会社, 東京, 11版, 2011.
- [21] International Organization for Standardization. Nuclear energy Nuclear fuel technology — Scaling factor method to determine the radioactivity of low- and intermediate-level radioactive waste packages generated at nuclear power plants. ISO 21238:2007. 2007.
- [22] International Atomic Energy Agency. Determination and Use of Scaling Factors for Waste Characterization in Nuclear Power Plants. IAEA Nuclear Energy Series NW-T-1.18. Vienna, 2009.
- [23] 日本原子力学会. ピット処分及びトレンチ処分対象廃棄物の放射能濃度決定 に関する基本手順:2011. AESJ-SC-F022:2011. 2012.

- [24] 原子力規制委員会. 廃棄物確認に関する運用要領, 2014.
- [25] 原子力安全基盤機構. 均質・均一固化体の廃棄のための確認方法 (一部改正). JNES-SS-0801. 2008.
- [26] 原子力安全基盤機構.固体状廃棄物の分析による充填固化体に対するスケー リングファクタ等の継続使用の確認方法について. JNES-RE-2013-2028. 2014.
- [27] 日本原燃株式会社. 六ヶ所低レベル放射性廃棄物埋設センター廃棄体埋設確 認申請書, 2015.
- [28] 金明哲. R によるデータサイエンス. 森北出版, 東京, 2007.
- [29] 竹之内敬文, 村田昇. R で学ぶデータサイエンス 5:パターン認識. 共立出版, 東京, 2009.
- [30] P. Olmos, J.C. Diaz, J.M. Perez, P. Gomez, V. Rodellar, P. Aguayo, A. Bru, G. Garcia-Belmonte, and J.L. de Pablos. A new approach to automatic radiation spectrum analysis. *IEEE Trans. Nucl. Sci.*, Vol. 38, No. 4, pp. 971–975, 1991.
- [31] P. Olmos, J. C. Diaz, and J. M. Perez. Application of neural network techniques in gamma spectroscopy. Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip., Vol. 312, No. FEBRUARY 1992, pp. 167–173, 1992.
- [32] P. Olmos, J.C. Diaz, J.M. Perez, P. Aguayo, P. Gomez, and V. Rodellar. Drift problems in the automatic analysis of gamma-ray spectra using associative memory algorithms. *IEEE Trans. Nucl. Sci.*, Vol. 41, No. 3, pp. 637–641, jun 1994.
- [33] L.J. Kangas, G.L. Troyer, P.E. Keller, S. Hashem, and R.T. Kouzes. Alpha spectral analysis via artificial neural networks. In *Proc. 1994 IEEE Nucl. Sci. Symp. - NSS'94*, Vol. 1, pp. 418–421. IEEE, 1994.
- [34] V. Vigneron, J. Morel, M.C. Lépy, and J.M. Martinez. Statistical modelling of neural networks in γ-spectrometry. Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip., Vol. 369, No. 2, pp. 642–647, 1996.

- [35] V. Pilato, F. Tola, J.M. Martinez, and M. Huver. Application of neural networks to quantitative spectrometry analysis. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, Vol. 422, No. 1, pp. 423–427, 1999.
- [36] L. Chen and Y. X. Wei. Nuclide identification algorithm based on K-L transform and neural networks. Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip., Vol. 598, No. 2, pp. 450–453, 2009.
- [37] E. Yoshida, K. Shizuma, S. Endo, and T. Oka. Application of neural networks for the analysis of gamma-ray spectra measured with a Ge spectrometer. Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip., Vol. 484, No. 1, pp. 557–563, 2002.
- [38] S. Dragović, A. Onjia, S. Stanković, I. Aničin, and G. Bačić. Artificial neural network modelling of uncertainty in gamma-ray spectrometry. Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip., Vol. 540, No. 2-3, pp. 455–463, mar 2005.
- [39] L. J. Kangas, P. E. Keller, E. R. Siciliano, R. T. Kouzes, and J. H. Ely. The use of artificial neural networks in PVT-based radiation portal monitors. Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip., Vol. 587, No. 2-3, pp. 398–412, 2008.
- [40] A. Samolov, S. Dragović, M. Daković, and G. Bačić. Analysis of ⁷Be behaviour in the air by using a multilayer perceptron neural network. J. Environ. Radioact., Vol. 137, pp. 198–203, 2014.
- [41] A. Varley, A. Tyler, L. Smith, P. Dale, and M. Davies. Remediating radium contaminated legacy sites: Advances made through machine learning in routine monitoring of "hot" particles. *Sci. Total Environ.*, Vol. 521-522, pp. 270–279, 2015.
- [42] A. Varley, A. Tyler, L. Smith, and P. Dale. Development of a neural network approach to characterise ²²⁶Ra contamination at legacy sites using gamma-ray spectra taken from boreholes. J. Environ. Radioact., Vol. 140,

pp. 130-140, 2015.

- [43] A. Varley, A. Tyler, L. Smith, P. Dale, and M. Davies. Mapping the spatial distribution and activity of ²²⁶Ra at legacy sites through Machine Learning interpretation of gamma-ray spectrometry data. *Sci. Total Environ.*, Vol. 545, pp. 654–661, 2016.
- [44] M.R. Einian, S.M.R. Aghamiri, and R. Ghaderi. Application of neural network method to detect type of uranium contamination by estimation of activity ratio in environmental alpha spectra. J. Environ. Radioact., Vol. 151, pp. 75–81, 2016.
- [45] S. Dragović and A. Onjia. Classification of soil samples according to geographic origin using gamma-ray spectrometry and pattern recognition methods. *Appl. Radiat. Isot.*, Vol. 65, No. 2, pp. 218–224, 2007.
- [46] C. Bellinger, N. Japkowicz, and C. Drummond. Synthetic Oversampling for Advanced Radioactive Threat Detection. In 2015 IEEE 14th Int. Conf. Mach. Learn. Appl., pp. 948–953. IEEE, dec 2015.
- [47] W. Wei, Q. Du, and N. H. Younan. Optimized Spectral Transformation for Detection and Classification of Buried Radioactive Materials. *IEEE Trans. Nucl. Sci.*, Vol. 59, No. 4, pp. 1702–1710, aug 2012.
- [48] C. Cortes and V. Vapnik. Support-vector networks. Mach. Learn., Vol. 23, pp. 273–297, 1995.
- [49] R. L. Somorjai, B. Dolenko, and R. Baumgartner. Class prediction and discovery using gene microarray and proteomics mass spectroscopy data: curses, caveats, cautions. *Bioinformatics*, Vol. 19, No. 12, pp. 1484–1491, 2003.
- [50] M. Bern, D. Goldberg, W. H. McDonald, and J. R. Yates. Automatic Quality Assessment of Peptide Tandem Mass Spectra. *Bioinformatics*, Vol. 20, No. suppl 1, pp. i49–i54, 2004.
- [51] M. Hilario, A. Kalousis, C. Pellegrini, and M. Müseller. Processing and classification of protein mass spectra. *Mass Spectrom. Rev.*, Vol. 25, No. 3, pp. 409–449, 2006.

- [52] M. F. Ferrão, S. C. Godoy, A. E. Gerbase, C. Mello, J. C. Furtado, C. L. Petzhold, and R. J. Poppi. Non-destructive method for determination of hydroxyl value of soybean polyol by LS-SVM using HATR/FT-IR. Anal. Chim. Acta, Vol. 595, No. 12, pp. 114–119, 2007.
- [53] F. Chauchard, R. Cogdill, S. Roussel, J. M. Roger, and V. Bellon-Maurel. Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemom. Intell. Lab. Syst.*, Vol. 71, No. 2, pp. 141–150, 2004.
- [54] Q. Chen, J. Zhao, H. Zhang, and X. Wang. Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration. *Anal. Chim. Acta*, Vol. 572, No. 1, pp. 77–84, 2006.
- [55] O. Devos, C. Ruckebusch, A. Durand, L. Duponchel, and J. P. Huvenne. Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. *Chemom. Intell. Lab. Syst.*, Vol. 96, No. 1, pp. 27–33, 2009.
- [56] S. Chakraborty, B. S. Das, M. N. Ali, B. Li, M.C. Sarathjith, K. Majumdar, and D.P. Ray. Rapid estimation of compost enzymatic activity by spectral analysis method combined with machine learning. *Waste Manag.*, Vol. 34, No. 3, pp. 623–631, 2014.
- [57] P. Rösch, M. Harz, M. Schmitt, K. D. Peschke, O. Ronneberger, H. Burkhardt, H. W. Motzkus, M. Lankers, S. Hofer, H. Thiele, and J. Popp. Chemotaxonomic Identification of Single Bacteria by Micro-Raman Spectroscopy: Application to Clean-Room-Relevant Biological Contaminations. *Appl. Environ. Microbiol.*, Vol. 71, No. 3, pp. 1626–1637, 2005.
- [58] K. Gaus, P. Rösch, R. Petry, K.D. Peschke, O. Ronneberger, H. Burkhardt, K. Baumann, and J. Popp. Classification of lactic acid bacteria with UVresonance Raman spectroscopy. *Biopolymers*, Vol. 82, No. 4, pp. 286–290, 2006.

- [59] M. Sattlecker, C. Bessant, J. Smith, and N. Stone. Investigation of support vector machines and Raman spectroscopy for lymph node diagnostics. *Analyst*, Vol. 135, No. 5, pp. 895–901, 2010.
- [60] L. Breiman. Random Forests. Mach. Learn., Vol. 45, No. 1, pp. 5–32, 2001.
- [61] G. A. Satten, S. Datta, H. Moura, A. R. Woolfitt, M. da G. Carvalho, G. M. Carlone, B. K. De, A. Pavlopoulos, and J. R. Barr. Standardization and denoising algorithms for mass spectra to classify whole-organism bacterial specimens. *Bioinformatics*, Vol. 20, No. 17, pp. 3128–36, nov 2004.
- [62] B. H. Menze, W. Petrich, and F. A. Hamprecht. Multivariate feature selection and hierarchical classification for infrared spectroscopy: serum-based detection of bovine spongiform encephalopathy. *Anal. Bioanal. Chem.*, Vol. 387, No. 5, pp. 1801–1807, jan 2007.
- [63] 阿部重夫. パターン認識のためのサポートベクトルマシン入門. 森北出版, 東京, 2011.
- [64] R Core Team. R: A language and environment for statistical computing., 2014.
- [65] D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch. Misc functions of the Department of Statistics (e1071), TU Wien, 2014.
- [66] T. Hastie, R. Tibshirani, and J. Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. Springer, second edition, 2011.
- [67] A. Liaw and M. Wiener. Classification and Regression by randomForest. *R news*, Vol. 2, No. December, pp. 18–22, 2002.
- [68] 下川敏雄, 杉本知之, 後藤昌司. R で学ぶデータサイエンス 9: 樹木構造接近法. 共立出版, 東京, 2013.
- [69] 辻谷將明, 竹澤邦夫. R で学ぶデータサイエンス 6:マシンラーニング. 共立出版, 東京, 2009.
- [70] 平井有三. はじめてのパターン認識. 森北出版, 東京, 2012.
- [71] A. Karatzoglou, A. Smola, K. Hornik, and A. Zeileis. kernlab An S4 Package for Kernel Methods in R. J. Stat. Softw., Vol. 11, No. 9, pp. 1–20,

2004.

- [72] C. M. Bishop. Pattern Recognition and Machine Learning. Springer, New York, 2006.
- [73] C. Fraley, A. E. Raftery, T. B. Murphy, and L. Scrucca. mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. *Technical Report, Department of Statistics, University of Washington*, No. 597, 2012.
謝辞

本研究は,著者が岡山大学大学院自然科学研究科産業創成工学専攻に在学中に 実施したものです.在学中は,指導教官である同専攻教授 鈴木 和彦博士ならびに 研究員 麓 敦子氏にはあたたかいご指導・アドバイスを賜りましたことを心より感 謝いたします.

本論文の執筆にあたって,主査の岡山大学大学院自然科学研究科産業創成工学 専攻准教授 佐藤 治夫博士には,詳細にわたって丁寧な助言と熱心な指導をいた だき大変感謝いたします.副査の同専攻教授 柳瀬 眞一郎博士ならびに同研究科 生命医用工学専攻教授 五福 明夫博士におかれましても丁寧な助言をいただきま したことを感謝いたします.

また,国立研究開発法人日本原子力研究開発機構人形峠環境技術センターのみ なさまの多大なご支援があり本研究をやり遂げることができました.特に,環境 保全技術開発部研究主席石森有博士の根気強く,あたたかく,時に厳しいご指 導のおかげで論文をまとめることができました.心より深く感謝いたします.同 僚の横山 薫博士には,適切な助言をいただきましたことを感謝いたします.廃棄 物の履歴調査におきましては,同部処理技術開発課課長田中祥雄氏ならびに同 課金田 弘司氏にご協力いただきましたことを感謝いたします.

本論文を構成する発表論文等

1. 参考論文

(1) <u>Haruhi Hata</u>, Kaoru Yokoyama, Yuu Ishimori, Yoshiyuki Ohara, Yoshio Tanaka, Noritake Sugitsue. Application of support vector machine to rapid classification of uranium waste drums using low-resolution γ -ray spectra, *Applied Radiation and Isotopes*, Vol. 104, pp. 143–146, 2015. (第2章)

(2) <u>秦 はるひ</u>, 石森 有. ランダムフォレスト法による γ 線スペクトルを用いた放射性廃棄物ドラム缶の分類, *Radioisotopes*, Vol. 66, No. 1, pp. 1–10, 2017. (第3章)

 (3) <u>秦 はるひ</u>, 石森 有. ウランを含む廃棄物ドラム缶の γ 線スペクトル データに対するランダムフォレスト法とサポートベクターマシン法の分類 性能の比較, *Radioisotopes*, Vol. 66, No. 4, pp. 149–158, 2017 (in press).
(第4章)

- 2. その他の論文等
- 1) 査読付き論文

(1) 横山 薫, <u>秦 はるひ</u>, 小原 義之, 杉杖 典岳, CANBERRA 製. Qualitative and Quantative (Q²) 装置による廃棄物ドラム缶中のウラン放射能定量試験, *Radioisotopes*, Vol. 62, No. 11, pp. 833–840, 2013.

(2) 横山 薫, 秦 はるひ, 小原 義之, 長沼 政喜, 杉杖 典岳. 廃棄物 γ 線分析シ

ステムによるウラン 238 定量結果の検証, *Radioisotopes*, Vol. 63, No. 12, pp. 559–566, 2014.

(3) 横山 薫, <u>秦 はるひ</u>, 長沼 政喜, 小原 義之, 石森 有. 放射性廃棄物中のウラン 238 量のコンプトン散乱を用いた定量手法の検証, *Radioisotopes*, Vol. 64, No. 11, pp. 687–696, 2015.

2) 査読付き国際会議のプロシーディングス

(1) <u>Haruhi Hata</u>, Kaoru Yokoyama, Noritake Sugitsue. Systematic chemical decontamination using IF₇ gas, Proceedings of 14th International Conference on Environmental Remediation and Radioactive Waste Management (ICEM 2011) (CD-ROM), 105–110, Reims, 2011.

(2) <u>Haruhi Hata</u>, Kaoru Yokoyama, Noritake Sugitsue. Chemical decontamination using IF₇ gas, Proceedings of 20th International Conference on Nuclear Engineering and the ASME 2012 Power Conference (ICONE-20 & POWER 2012) (DVD-ROM), 7 Pages, Anaheim, 2012.