

機械学習を使用した KEK Linac 加速器運転調整システムの
開発

(R & D of the KEK Linac accelerator tuning
using machine learning)

理学研究科

数物系専攻

城庵颯

概要

本研究では、高エネルギー加速器研究機構（KEK）における電子・陽電子入射器 Linac の加速器運転調整システムの開発を行った。具体的には、機械学習を用いた加速管の RF 位相調整及び、ビーム位置補正のためのステアリング電磁石の調整のための基礎開発を行った。開発を行うために、Linac 加速器運転データを収集し、そのデータを用いて加速器運転パラメータの最適化手法を開発を行った。Linac 加速器運転では、加速器を構成する数千点の機器パラメータを制御している。温度変化、振動等、環境の影響を受けて運転状況が変わるため、運転パラメータを常時調整している。運転調整の最適化、高速化を行い、入射効率を向上させるために、本研究では、機械学習を用いた運転パラメータ調整方法を開発した。機械学習は、複数のパラメータ同士の複雑なパターン認識を得意とするため、加速器運転調整に適用させることで調整時間の高速化や調整性能の向上、安定化が期待できると考えられる。本研究では、機械学習の中でもニューラルネットワークを使用した教師あり学習に注目した。ニューラルネットワークを使用した教師あり学習を用いて加速器パラメータを推定することで、加速器調整パラメータの高速化、最適化が可能になると考えられる。新たな加速器運転パラメータの調整方法として、ニューラルネットワークを用いた加速管の RF 位相調整及び、ステアリング電磁石の調整システムを開発し、その性能評価を行った。

目次

第 1 章	序論	1
第 2 章	加速器	3
2.1	加速器の基本性能	3
2.2	加速器の加速原理	4
2.3	加速器で使用される電磁石	6
2.4	Linac 加速器	7
2.5	ネットワーク分散型制御	9
第 3 章	機械学習	11
3.1	機械学習について	11
3.2	教師あり学習としてのニューラルネットワーク	12
3.3	動作環境	20
第 4 章	機械学習を使用した加速器運転パラメータ予測	22
4.1	機械学習を使用した加速器運転パラメータ予測の方針	22
4.2	加速器データの収集	22
4.3	加速器運転調整パラメータ予測の指標	23
4.4	ネットワークデザイン	24
4.5	データ 1 を使用した予測	25
4.6	データ 1 で学習したネットワークによるデータ 2 の評価	39
4.7	データ 1-1 で学習したネットワークによるデータ 1-2 の評価	44
第 5 章	まとめ	48
	参考文献	49
	謝辞	50

目次

2.1	静電場を使用した加速の例	5
2.2	1つの静電場を使用し繰り返し加速する例	5
2.3	ヴィデレー型加速器の例	6
2.4	加速電波による電子の加速の例	6
2.5	KEKにおける加速器の配置図	9
2.6	Linac 加速器の配置図	9
2.7	ネットワーク分散型制御の例	10
2.8	EPICS を使用した制御システムの例	10
3.1	ニューラルネットワークの例 (分類学習)	12
3.2	ニューラルネットワークの計算の様子	13
3.3	ReLU 関数	14
3.4	シグモイド関数	14
3.5	ディープニューラルネットワークの例	18
3.6	ドロップアウトの概念図	19
4.1	DNN を使用した加速器運転パラメータの予測の流れ	22
4.2	Linac 加速器 A Sector の配置図	24
4.3	Linac 加速器 5 Sector の配置図	24
4.4	DNN を使用した加速器運転パラメータの予測の流れ (Q_{ratio} を用いた場合)	24
4.5	作成したネットワークの概略図	25
4.6	Linac 加速器 B Sector の配置図	26
4.7	B Sector5 個目にあるクライストロン出力の RF 位相パラメータと Q_{ratio} の関係	26
4.8	データ 1 を使用した RF 位相パラメータの回帰予測の結果まとめ	28
4.9	パラメータを分類した際の B Sector5 個目にあるクライストロン出力の RF 位相パラメータと Q_{ratio} の関係	29
4.10	ステアリング電磁石の選別の例	31
4.11	Linac 加速器 3 Sector の配置図	32
4.12	Linac 加速器 4 Sector の配置図	32
4.13	ステアリング電磁石 (PY_32_4) パラメータと Q_{ratio} の関係	33
4.14	ステアリング電磁石 (PY_44_4) パラメータと Q_{ratio} の関係	33
4.15	データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ	35
4.16	データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの回帰予測の結果まとめ	35

4.17	パラメータを分類した際のステアリング電磁石 (PY_32_4) パラメータと Q_{ratio} の関係	36
4.18	パラメータを分類した際のステアリング電磁石 (PY_44_4) パラメータと Q_{ratio} の関係	37
4.19	データ 1 とデータ 2 での RF パラメータの分布比較	43
4.20	データ 1 とデータ 2 での電磁石パラメータの分布比較	43
4.21	データ 1-1 とデータ 1-2 での RF パラメータと電磁石パラメータの分布比較	44

表目次

3.1	動作環境のまとめ	20
3.2	ディープニューラルネットワークのライブラリのバージョン一覧	21
4.1	データ 1 を使用した RF 位相パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ	27
4.2	データ 1 を使用した RF 位相パラメータの回帰予測の結果まとめ	27
4.3	データ 1 を使用した RF 位相パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ	29
4.4	データ 1 を使用した説明変数 1 の場合の RF 位相パラメータの回帰予測の結果	30
4.5	データ 1 を使用した説明変数 2 の場合の RF 位相パラメータの回帰予測の結果	30
4.6	データ 1 を使用した説明変数 3 の場合の RF 位相パラメータの回帰予測の結果	30
4.7	データ 1 を使用した RF 位相パラメータの回帰予測の結果まとめ	31
4.8	データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ	34
4.9	データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ	34
4.10	データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ	35
4.11	データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの回帰予測の結果まとめ	35
4.12	データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ	37
4.13	データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ	38
4.14	データ 1 を使用した説明変数 1-1 の場合のステアリング電磁石 (PY_32_4) パラメータの分類予測の結果	38
4.15	データ 1 を使用した説明変数 1-2 の場合のステアリング電磁石 (PY_32_4) パラメータの分類予測の結果	38
4.16	データ 1 を使用した説明変数 1-3 の場合のステアリング電磁石 (PY_32_4) パラメータの分類予測の結果	39
4.17	データ 1 を使用した説明変数 2-1 の場合のステアリング電磁石 (PY_44_4) パラメータの分類予測の結果	39
4.18	データ 1 を使用した説明変数 2-2 の場合のステアリング電磁石 (PY_44_4) パラメータの分類予測の結果	40

4.19	データ 1 を使用した説明変数 2-3 の場合のステアリング電磁石 (PY_44_4) パラメータの分類予測の結果	40
4.20	データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ	40
4.21	データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの分類予測の結果まとめ	40
4.22	データ 1 を使用したネットワークでのデータ 2 の評価についての RF 位相パラメータの回帰予測の結果まとめ	41
4.23	データ 1 を使用したネットワークでのデータ 2 の評価についての RF 位相パラメータの分類予測の結果まとめ	41
4.24	データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ	42
4.25	データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_44_4) パラメータの回帰予測の結果まとめ	42
4.26	データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ	42
4.27	データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_44_4) パラメータの分類予測の結果まとめ	42
4.28	データ 1-1 を使用した RF 位相パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ	45
4.29	データ 1-1 を使用した RF 位相パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ	45
4.30	データ 1-1 を使用した RF 位相パラメータの回帰予測の結果まとめ	46
4.31	データ 1-1 を使用したネットワークでのデータ 1-2 の評価についての RF 位相パラメータの回帰予測の結果まとめ	46
4.32	データ 1-1 を使用した RF 位相パラメータの分類予測の結果まとめ	46
4.33	データ 1-1 を使用したネットワークでのデータ 1-2 の評価についての RF 位相パラメータの分類予測の結果まとめ	46
4.34	データ 1-1 を使用したステアリング電磁石パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ	47
4.35	データ 1-1 を使用したステアリング電磁石パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ	47
4.36	データ 1-1 を使用したでのステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ	47
4.37	データ 1-1 を使用したネットワークでのデータ 1-2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ	47
4.38	データ 1-1 を使用したステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ	47
4.39	データ 1-1 を使用したネットワークでのデータ 1-2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ	47

第 1 章

序論

本研究では、高エネルギー加速器研究機構 (KEK) における電子・陽電子入射器 Linac の加速器運転データを用いて、Linac 加速器の運転調整パラメータシステム開発を行った。具体的には、機械学習を用いた加速管の RF 位相調整及び、ビーム位置補正のためのステアリング電磁石の調整のための基礎開発を行った。開発を行うために、Linac 加速器運転データを収集し、そのデータを用いて加速器運転パラメータの最適化手法の開発を行った。

ここで Linac 加速器での加速器運転パラメータの調整には大きく以下の 2 種類がある。

1. Linac 加速器ビーム源からビームを下流まで通すための調整
2. ビームの入射効率を上げるための調整

1. の調整は Linac 加速器が長期の shutdown 時の、再立ち上げを行う際に行われる。頻度としては夏と冬の年 2 回行われるもので、調整を完了するのにかかる時間は 1 日程度である。2. の調整は Linac 加速器で高い入射効率を目指し行われている調整である。Linac 加速器運転では、加速器を構成する数千点の機器パラメータを制御しており、温度変化、振動等、環境の影響を受けて運転状況が変わるため、加速器運転パラメータを常時調整している。本研究では 2. の調整についての開発を行った。

現在の 2. の調整において以下のような問題がある。

1. 加速器の運転パラメータの数は多く、複雑なシステムであること
現在加速器調整は人の手で行っているため、このことよりパラメータの中から相関を見つけることが困難になり、調整する人によって最適化の速さ、調整精度が異なってしまう。
2. 加速器運転パラメータに入射効率の再現性がないこと
加速器には同じ加速器運転パラメータを入力しても同じ結果が得られないという特徴があり、加速器周囲の環境の変化によって、常に運転パラメータの最適化が必要になってしまう。

問題 2. については、加速器運転パラメータのみからの入射効率の再現性がなかったとしても、加速器周囲の環境パラメータを含めることで再現性が得られると考えられる。

そこで上記の問題点に対して、複数のパラメータ同士の複雑なパターン認識を得意とする機械学習に注目した。加速器運転調整に機械学習を適応することで、

1. 加速器調整時間の高速化
2. 加速器調整性能の向上、安定化

が期待できる。

機械学習の中でもニューラルネットワークを使用した教師あり学習に注目した。ニューラルネットワ

ークを使用した教師あり学習を用いて加速器運転パラメータを推定することで、Linac 加速器の運転調整パラメータの調整が行えると考えられる。

ニューラルネットワークを使用した教師あり学習を用いて加速器パラメータを推定することで、加速器調整パラメータの高速化、最適化が可能になると考えられる。新たな加速器運転パラメータの調整方法として、ニューラルネットワークを用いた加速管の RF 位相調整及び、ステアリング電磁石の調整システムを開発し、その性能評価を行った。

第 2 章

加速器

本章では加速器の基本原理 [1],[2],[3] 及び、KEK Linac 加速器について述べる。

加速器とは荷電粒子を加速する装置である。荷電粒子としては電子や陽子が良く用いられるが、陽電子や反陽子などの反粒子や、重陽子やアルファ粒子からウランに至る重イオンも加速されている。

2.1節で加速器の基本性能について、2.2節で加速器の加速について説明し、加速器で使用される電磁石について 2.3節で簡単に紹介する。2.4節で今回使用した電子・陽電子線形入射器である Linac 加速器について説明する [4],[5],[6],[7]。また 2.5節で本研究でデータ収集に使用したネットワーク分散型制御について紹介する。

2.1 加速器の基本性能

加速器に以下の 5 つの基本的な性能が必要である。

- 粒子源

粒子源とは加速器で加速する荷電粒子を生成する部分である。粒子源にはできるだけ強度が高く、広がり少ないビームをつくることが要求される。電子ビームの粒子源は電子銃、イオンビームの粒子源はイオン源と呼ばれる。以下で本研究で使用した電子ビームの中でも一般的な電子銃について的手法について解説する。

電子銃としては熱陰極からの熱電子放出による手法が一般的である。熱電子放出による電子銃では、金属片を熱することでその中に含まれる電子を不規則にゆさぶり、金属から飛び出す電子を使用する。またこの手法は 1 つの面に孔を開けた金属製の箱の中で行う。箱の中に置いた金属には負電荷を帯びさせることで、電子は反発を受けさらに飛び出しやすくなる。さらに箱の孔が開いた面に高い正の電氣的なポテンシャルをかけておくことで、電子をこの孔に向かって進みやすくすることができ、もちろん全ての電子が孔を通り抜けるわけではないが、孔から出てきたいくつかの電子がビームに使用される。

また電子銃やイオン源からビームを引き出すときには、50 ~100kV の高電圧をかけるが、電子や 1 価のイオンの場合、粒子源から出た段階で 50 ~100keV のエネルギーを粒子は得ることになる。

- ビームの加速

加速器での加速される粒子は電荷をもつ粒子に限られる。これは加速には電場が利用されるからである。加速の方法は 2.2節で詳しく述べるが、コンデンサのように 2 枚の電極間に発生する電場を使用する方法や、空洞共振器に発生する高周波電場や磁場が時間的に変化するとき発生する誘導起電力を使用する方法がある。

- ビームの収束

どのような加速器にも中心軌道というものがあり、理想的な条件の下では粒子はこの軌道上を運動するように設計される。またビームは多数の粒子が含まれている塊(バンチ)で出来ているが、粒子源でビームが作られたときから粒子は一塊となっているわけではなく、粒子の出る位置がばらついており運動の方向もそろっていない。したがってビームを進行方向と直角な面で切ると、その断面はある程度の大きさを持ってしまう。更にビームは進むにつれて断面は広がり、ビームに対して何らかの収束力が働かない限り発散してしまう。そこで加速器ではビームに収束力をかける必要があり、電場や磁場を用いて行われる。また 2.3節で加速器のビーム収束に使用される電磁石について紹介する。

- 真空

加速器ではビームの通る経路を真空に保つ必要がある。ビームが通る容器は真空ダクト(真空チェンバー)と呼ばれ、多くの場合は非磁性の金属(ステンレスやアルミ合金など)が使用される。真空ダクトの中に残留ガスがあると、ビームはこのガスと衝突することで軌道が外れ、その結果ビーム中の粒子数は減少してしまう。通常、ビーム中の粒子数が半分になる時間をビームの寿命と呼ぶので、ビームの寿命は真空ダクト中の残留ガスの圧力に反比例する。これより真空技術は加速器において重要な技術となっていることがわかる。

また真空技術が注意すべき加速器ならではの問題が多数存在している。加速器では真空ダクトの中を大強度のビームが走るため、ビームと真空ダクトの導体の間との電磁相互作用により大きな電磁場が誘起され、これによりビームの運動が乱される。また電子加速器の場合はビームを曲げると放射光が発生し、放射光が真空ダクトに当たると光脱離により吸着ガスが発生したり、放射光が真空ダクトに吸収されることで大きな熱が発生してしまう。このような問題があることから、加速器で使用される真空システムは通常の真空システムと異なり様々な問題に対応することが必要となる。

- 制御

加速器は上にあげたような様々な要素からなる複雑なシステムであり、その全ての動作を把握し、的確に制御を行う必要がある。制御には速く行わなければならない制御と、そこまでの速さを求めない制御がある。前者の制御に関してはハードワイヤーを用いた制御を行い、後者の制御に関してはネットワークを用いた制御システムが使用されることもある。後者の技術としてはネットワーク分散型制御システムがしばしば用いられる。このシステムについては 2.5節で解説を行う。

2.2 加速器の加速原理

加速器での加速にはいくつかの方法がある。本節では加速器で使用する加速方法のうち、もっとも単純な静電場を使用した加速と、現在幅広く使用されている高周波加速について説明する。

2.2.1 静電場での加速

加速器での1番簡単な荷電粒子の加速方法は、静電場を使用した加速である。図 2.1に静電場を利用した加速の例を示す。コンデンサのような電荷を帯電させた金属平面電極によるクローン力を使用する。

平面電極に孔をあけることによって、その孔に荷電粒子を通過させ、その際に平面電極間の電位差によって荷電粒子を加速させることができる。

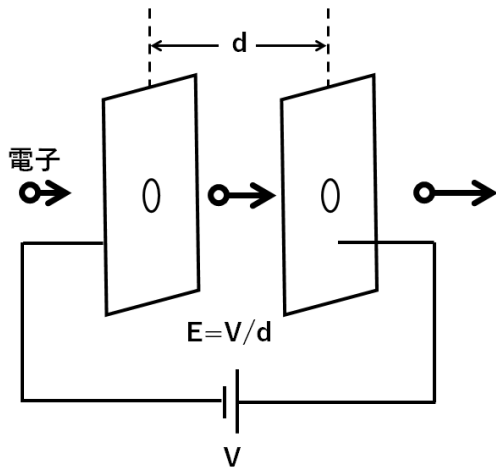


図2.1 静電場を使用した加速の例

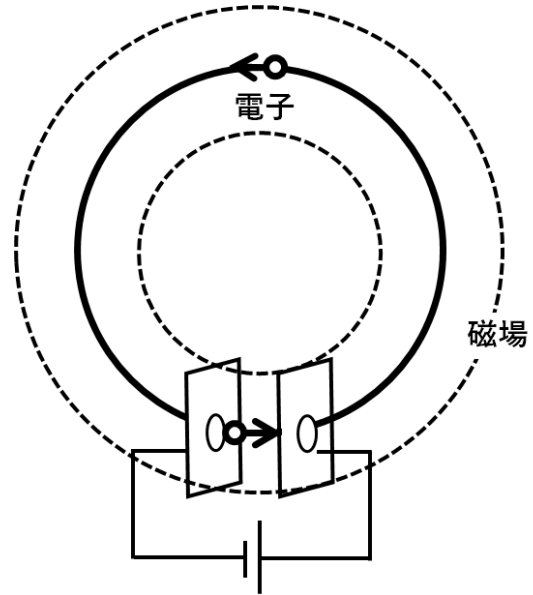


図2.2 1つの静電場を使用し繰り返し加速する例

しかし静電場での加速には問題がある。金属平面間の最大印加電圧には放電限界からの制約がある。そのため1度の金属平面間での加速には限界が出てしまう。この限界を考慮し、繰り返し荷電粒子を金属平面間に通すことで高エネルギーを得る手法も考えられる。しかしこれは不可能である。図2.2に1つの静電場を使用し繰り返し加速する例を示す。これからわかるように磁場を用いて荷電粒子を回転させると、左の電極から右の電極への間で加速されたエネルギーは、右の電極から左の電極への間で減速され、エネルギーは相殺されてしまう。そのため繰り返し1つの静電場で加速しても、粒子は結果的に加速できない。

このように静電場での加速には高いエネルギーを得ることができないという問題がある。この問題を解決するべく高周波加速が開発された。高周波加速については次節で説明する。

2.2.2 高周波 (RF) 加速

高周波加速はRF加速とも呼ばれる。また高周波加速の元となった加速方法に多重加速と呼ばれるものがある。多重加速では静電場による加速で使った電極を直線状に複数個設置することで、それぞれの電極間で繰り返し加速を行い高いエネルギーを得ることができる。またこの方法では電極間の電位差を比較的低い値にすることで、電極が放電を起こす問題も解決することができる。多重加速を応用した方法に高周波加速がある。高周波加速とは、多重加速での電極に加える電圧を周期的に変化させることで、加速器中の荷電粒子の運動とこの周期を一定に保ちながら加速させる方法である。円筒形の電極を多数一直線に並べ、電極間で加速を繰り返す。図2.3に高周波加速の簡単な例としてヴィデレー型加速器を示す。ヴィデレー型加速器では多数の電極をもつ装置を造り、電極を1つおきに接続し、一方は接地し他方に高周波電圧を与えている。このためヴィデレー型加速器ではこの高周波加速を用いることで、多重加速よりも高いエネルギーを得ることができた。

ここで本研究で使用した線形電子加速器での高周波加速の原理を説明する。電子の進行方向に電場の成分をもつ進行波の位相速度を、電子がいつも加速の位相にのって加速電場の作用を受けるように調整

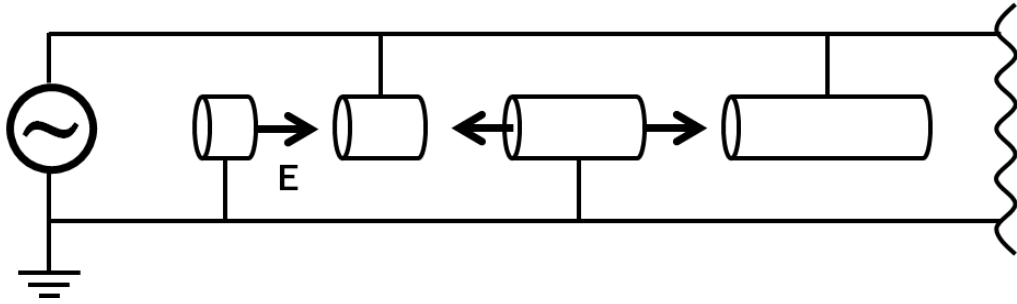


図2.3 ヴィデレー型加速器の例

することによって効率よく加速することができる。図 2.4 に加速電波による電子の加速の例を示す。ここで左から右へ位相速度 v_p ですむ電波を示す。A-B 間では電子が加速される位相、B-C 間では電子が減速される位相となる。ここで位相速度 v_p と電子の速度 v_e に大きな差があれば、電波に対する電子の位相 θ は時間とともに変化していくので電子は加速と減速を繰り返してしまい、結果として電子は加速も減速もされないままである。しかし位相速度 v_p と電子の速度 v_e に差がなければ、電子は常に電場から $E_z = E_0 \sin \theta$ の作用を受け、電波とともに距離 l を進めば、電子は $elE_0 \sin \theta$ の運動エネルギーを得ることになる。またここで加速を続けさせるために電子が加速するごとに、電波の位相速度も増やしていく必要があるが、電子の質量は軽いため、位相速度は一定 ($\approx c$) にすることができる。

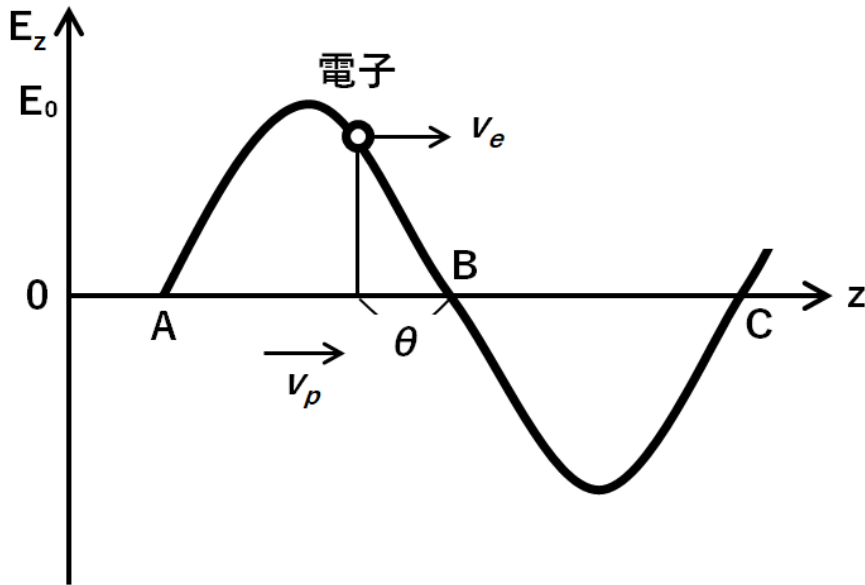


図2.4 加速電波による電子の加速の例

2.3 加速器で使用される電磁石

本節では加速器で使用される種々の電磁石について、その特徴と使用目的を紹介する。

- 偏向電磁石

偏向電磁石はビームの方向を変える電磁石である。荷電粒子は磁力線の中を通過するとフレミングの左手の法則による力を受けその進行方向が曲げられる。偏向電磁石では上下に磁場をかける

ことで荷電粒子の進行方向を左右に曲げることができる。また偏向電磁石を 90 度傾ければ、上下方向へも向きを曲げられる。線形加速器などの直線加速器ではあまり使用されないが、シンクロトロンなどの円形加速器ではこの偏向電磁石を適切に配置することで、ビームを曲げビームが円形に近い軌道を保つようにしている。円形加速器では一見、なめらかな円形のように見えるが、実際は偏向電磁石でビームを曲げている部分以外は直線となっている。

- 4 極電磁石

4 極電磁石は加速器で粒子ビームを細く絞るために使用される電磁石である。N 極の電磁石と S 極の電磁石それぞれ 2 つずつ、同じ極の電磁石が対面になるように配置する。ビームは同じ電荷を持つ荷電粒子が集まったバンチとなっているため、加速器を通過しているうちに、それぞれの粒子が電氣的に反発し合って発散してしまう。またビームは加速中あるいは輸送中にも発散する傾向にある。そのため 4 極電磁石はビームの発散を防ぎ、ビームを中心軌道周辺の狭い空間に保持するため、ビームを中心に集める働きをする。4 極電磁石は中心から遠ざかるに従って磁場が強くなるように製作されているため、発散して中心から外れた粒子をよりビーム軌道中心方向へ曲げて集めることができ、ビームを細く絞ることができる。光学系の凸レンズ、凹レンズと同じ役割を果たすが、四極電磁石 1 台では水平、垂直方向それぞれに逆のレンズとして働くため、2 台以上を組み合わせて使用する。

- 6 極電磁石

6 極電磁石はビームの運動量差に伴う閉軌道のズレや焦点位置のズレの補正に使用する電磁石である。ビームを加速している間の粒子の進行速度は完全に揃っているのではなくばらついており、速度分布をもっている。これより粒子の速さによって 4 極電磁石から受ける力が異なり、速度の速い粒子ほど受ける力が弱くなる。これによってビームが発散する方向に向かってしまうので、これを防ぎビームを収束させるために 6 極電磁石を使用する。またこの焦点位置のズレは光学系で色収差とよばれるものと同じである。

- 補正電磁石 (ステアリング電磁石)

補正電磁石は偏向電磁石や四極電磁石でのビーム軌道のズレの補正に使用する電磁石である。偏向電磁石や四極電磁石の電源は強磁場を作り出すために大電流を出せるが、そのため細かな磁場の微調整が苦手である。これらの電磁石が苦手としている調整を行い、また磁石の設置位置が理想的な軌道からどうしてもずれてしまっている誤差の吸収など行う。

2.4 Linac 加速器

2.4.1 線形加速器

線形加速器とは高周波加速を用いて加速を行う直線型の加速器である。高周波加速については 2.2 節で説明した。荷電粒子の走行時間に合わせて電極を並べて加速を行う発想は 1920 年代からあったが、当時は電波技術が未発達であったのと、イオン収束の研究が進んでいなかったため成功を収めていなかった。第 2 次世界大戦後に電波技術が急速に発展し、荷電粒子を加速することに成功した。

線形加速器では加速する荷電粒子により構造が異なる。ここでは本研究で使用した線形電子加速器について簡単に説明する。電子を線形加速器で加速する場合、80keV のエネルギーですでに $\beta = \frac{v}{c} = 0.5$

に達するので、電子銃から電子が射出された段階で電子は高速となっている。その後の高周波を用いた加速では、電子の進行方向に電場をもち、位相速度を電子の速さに合致させた導波管を用いる。円形導波管で、円形の「ひだ」を適当に作って、位相速度を小さくし、しかも電子の加速につれて位相速度が変化するようにする。高周波としては 1,000 ~10,000MHz のマイクロ波を用い、電源はクライストロンを使用する。

線形電子加速器は信頼度を高く、非常に多くの用途で使用されている。加速粒子のエネルギーが 1MeV 前後のものは医療、材料検査、放射線化学のために使用されることが多く、10MeV 以上のものは原子核物理学実験のために使用されることが多い、また数 GeV 以上のものが素粒子実験に使用されることが多い。

また線形加速器は直接加速器実験に使用されるのではなく、しばしば入射器として使用される。入射器の簡単な説明は次節で行う。

2.4.2 入射器

粒子のエネルギーが低いと、加速器内の残留気体による散乱、残留磁気による磁場の乱れの影響を受けやすく、また速度変化の幅が大きいため加速周波数の変調の幅が広がる。これらの理由からシンクロトロンでは、加速器に入射する前に荷電粒子のエネルギーを予めある程度の高さまで上げておく必要がある。このような目的で使われる前段の加速器を入射器と呼ぶ。一般的に線形加速器が用いられる。

2.4.3 Linac 加速器

Linac 加速器は KEK にある全長 600m の線形電子・陽電子加速器である [4],[5][6],[7]。入射器として使用され、電子を Photon Factory (PF)[8]、PF Advanced Ring (PF-AR)[9]、SuperKEKB 加速器電子リング (HER)[10] へ入射し、陽電子を SuperKEKB 加速器陽電子リング (LER)[10] へ入射する。そのため入射先の加速器の要請により、異なるエネルギーの粒子を異なる間隔で振り分け入射することが可能となっている。また運転性能としては最大 50Hz、2 バンチ (96ns 間隔) 運転が可能である。図 2.5 に Linac 加速器及び、Linac 加速器から入射される種々の加速器の配置図を示す。

本研究では、Linac 加速器の SuperKEKB 加速器 HER リングへの入射データを用いて開発を行った。また図 2.6 に Linac 加速器の全体配置図を示す。加速器の上流が A Sector となっており、一番下流が 5 Sector となっている。また Linac 加速器は線形加速器であるが、図のように厳密な直線ではなく J のような形をしている。建設当初の Linac 加速器では 1 Sector から 5 Sector までの直線状となっており、エネルギーも 2.5GeV であった。そこから super KEKB 加速器の前身となる KEKB 加速器の入射器としてエネルギーを 8GeV に upgrade した際に A Sector から C Sector が増設され、形状も J のような形になった。その歴史背景から厳密に直線でなくなった今でも線形加速器とされている。また super KEKB 加速器の入射器として upgrade した際に、陽電子のために 2 Sector と 3 Sector の間の D Sector に陽電子ビーム用にダンピングリングを増設した。

Linac 加速器にはビーム位置モニター (BPM) が 100 台、ステアリング電磁石が 200 台、RF モニターが 60 台が配置されている。また Linac 加速器では高周波源としてクライストロンを使用しており、クライストロン 1 本の出力 (50MW) を 4 つに分けて、4 本の加速管に高周波を供給している。

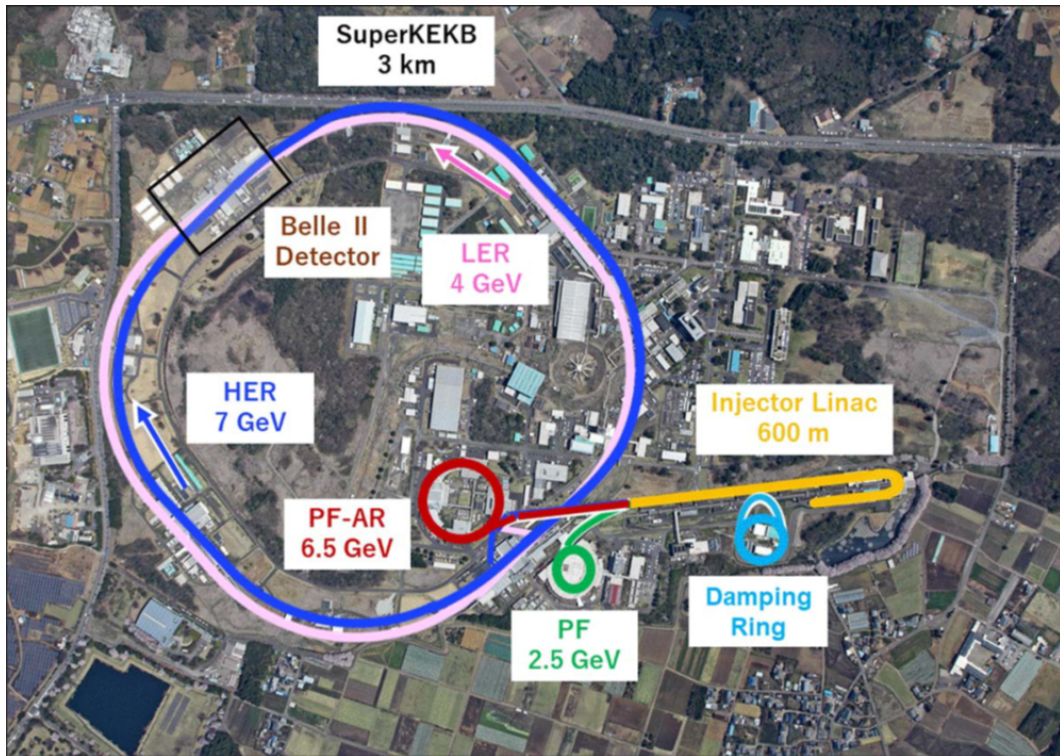


図2.5 KEK における加速器の配置図

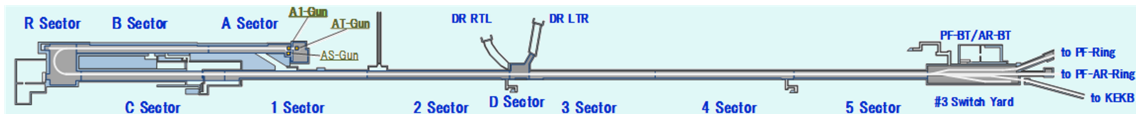


図2.6 Linac 加速器の配置図

2.5 ネットワーク分散型制御

一般的な機器制御では、以下のように行われる。例えば計測器で電圧または電流を出力した場合、その電気信号はAD変換を行いデジタル信号として測定装置から出力される。制御用装置と測定装置の間はデジタル信号でデータをやり取りされ、測定装置の制御が行われる。オペレーターをこの説明に加えると、オペレーターは制御用装置に命令し、その命令を受けて制御用装置が測定装置に命令する。その命令を受け、測定装置が動作され、その結果送信されたデータは制御用装置がオペレーターに表示される。

ここでネットワーク分散型制御の場合、ネットワークを介さない機器制御のように1つずつの機器を制御装置で制御するのではなく、システムを構成する各機器ごとに機器制御装置があり、その機器制御装置はネットワークで接続されて、相互に通信、監視を行う。複数の機器制御装置を一手に制御を行うことができるので、大規模実験などの多人数が様々な機器制御を行う際に有用である。図2.7でネットワーク分散型制御の例を示す。

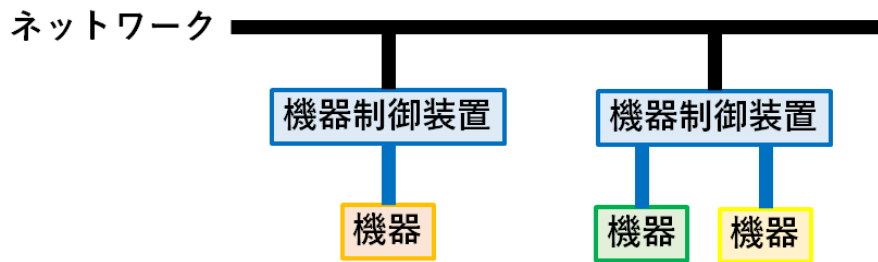


図2.7 ネットワーク分散型制御の例

2.5.1 EPICS の仕組み

EPICS とは Experiment Physics and Industrial Control System の略であり、国内外の粒子加速器、大規模実験、望遠鏡などで使用されている分散制御システムを構築するソフトウェアである [11]。高エネルギー物理学、原子核、加速器、宇宙、重力波実験分野で国際共同開発され使用されている。(例：KEKB、J-PARK、KAGRA など) このような分散制御システムは通常、数十台または数百台のコンピュータで構成される。EPICS はそれらの間の通信を可能にし、中央制御室やインターネットを介して遠隔での制御やフィードバックを行うことができる。図 2.8 に EPICS を使用した制御システムの例を示す。EPICS の特徴としては制御用装置が「IOC」と「OPI」の 2 重構造になっていることが挙げられる。IOC とは Input Output Controller の略で、測定装置とネットワークの間でデータをやり取りし、機器を制御する制御用装置となっている。OPI とは Operator Interface の略で、オペレーターとのインターフェースとして、オペレーターが入出力を行う制御装置となっている。これによってオペレーターは OPI を使用しネットワークを介して IOC に指示を出すことができる。

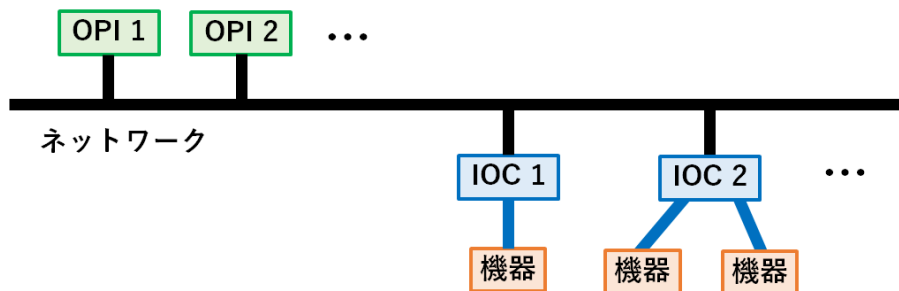


図2.8 EPICS を使用した制御システムの例

第 3 章

機械学習

本章では、3.1節で機械学習について簡単に紹介し、3.2節で今回使用したディープニューラルネットワークと前提知識となるニューラルネットワークの説明を行う [12]。また 3.3節で本研究で使用した機械学習の動作環境を示す [13],[14]。

3.1 機械学習について

機械学習とはその字が表す通り「機械（コンピュータ）が学習する」ことで、説明変数（入力データ）をもとに学習を行い、そのデータの持つパターンを抽出することで目的変数（正当値、出力データ）を予測するモデルを構築する手法である。また機械学習の種類は主に「教師あり学習」「教師なし学習」「強化学習」の3つに分類される。

- 教師あり学習

教師あり学習は入力データと正答値（ラベルと呼ばれることが多い）が既知のデータセットを用い学習を行い、入力データと正答値の関係からモデルを構築する手法である。あらかじめ正答値が付与されていることから「教師「あり」学習」と呼ばれる。一度モデルを構築すれば、未知の入力データに適用し、出力値として正答値の予測を行うことが可能である。またモデルが出力する値の種類によって「分類学習」と「回帰学習」がある。分類学習は学習の結果として離散値を出力する学習である。例えば mnist と呼ばれる手書き数字の画像認識が有名である。手書き数字の画像データを入力データとし、その画像が 0 から 9 までのどの数字を表すかを正答値として、画像データからそこに書かれている数字の予測を行うことで画像分類を行う。回帰学習は学習の結果として連続値を出力する問題である。例えばある土地の気温と湿度のデータがある場合、そのデータから学習を行ったモデルを使用し、入力データとして気温（あるいは湿度）を入力することで、湿度（あるいは気温）を連続値として出力し予測することが可能である。

- 教師なし学習

教師なし学習は入力データからそのデータの構造や特性などを学習し、特徴を見つけ出す手法である。教師あり学習と異なり、あらかじめ正答値が付与されていないことから「教師「なし」学習」と呼ばれる。また教師あり学習と同様に、一度モデルを構築すれば未知の入力データに適用し、以前に学習させたデータと同じ特徴を見つけ出すことが可能である。例えば、ある店舗での購入者の購入情報をもとに、購入者のグルーピングを行う際に使用される。またデータに正答値が含まれていないため、モデルの正当性を判断する際は、抽出したパラメータや相関から正当性の判断を行うことが可能である。

- 強化学習

強化学習とは環境駆動型の学習である。教師あり学習や教師なし学習のような固定的で明確なデータを基にした学習ではなく、プログラム自身が与えられた環境を観測し、各試行での行動の評価を行うことでモデルの更新を行い、連続した一連の試行の結果、「報酬が最大となる行動」を学習する手法である。例えば、倒立振り子（例：箒や定規を手の上で直立、維持させる遊び）について、倒立を維持し続けるモデルがある。これは振り子の傾きや加速度から支点の位置を調整し、倒立を維持できた場合に正の報酬を与え、倒立が崩れた場合に負の報酬を与えることで、振り子を維持し続けるモデルを構築することが出来る。社会での使用例としては、自動車の自動運転技術やゲームの自動操作がある。

機械学習の手法には様々なもの存在するが、ここでは今回使用したディープニューラルネットワークについて 3.2節で解説を行う。

3.2 教師あり学習としてのニューラルネットワーク

3.2.1 ニューラルネットワーク

ニューラルネットワークとは、脳神経細胞ニューロンを模した基本単位と、そのつながりを数式的なモデルで表現したものである。ニューラルネットワークを図で表すと、図 3.1 のようになる。

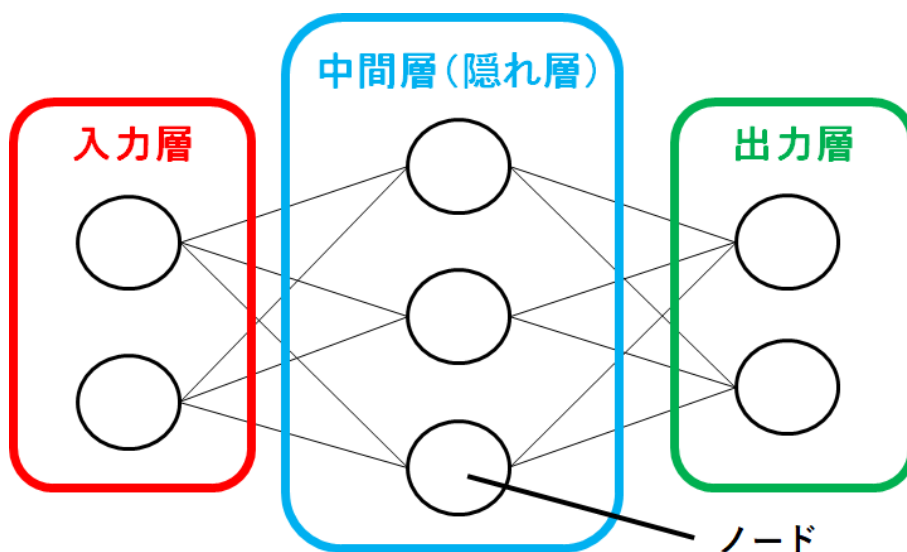


図3.1 ニューラルネットワークの例 (分類学習)

ここで一番左の列を入力層、一番右の列を出力層、中間の列を中間層と呼ぶ。また中間層は入力層や出力層と異なり、人の目には見えないことより、隠れ層と呼ばれることもある。また図 3.1 で丸で表されているものはノードと呼ばれており、入力信号を受け取った際にそれぞれに固有の重みが計算され出力する。計算の詳細は 3.2.2 節に記載する。ニューラルネットワークでは各ノードからの出力が次のノードへと伝播することで計算が進められる。またニューラルネットワークの層数やノードの数などはハイパーパラメータと呼ばれる。ハイパーパラメータの詳細は 3.2.7 節に記載する。

3.2.2 ニューラルネットワークの計算

ニューラルネットワークでどのように計算が行われるかを詳しく記載する。図 3.2はニューラルネットワークの中の1つのノードで行われる計算を図示したものである。

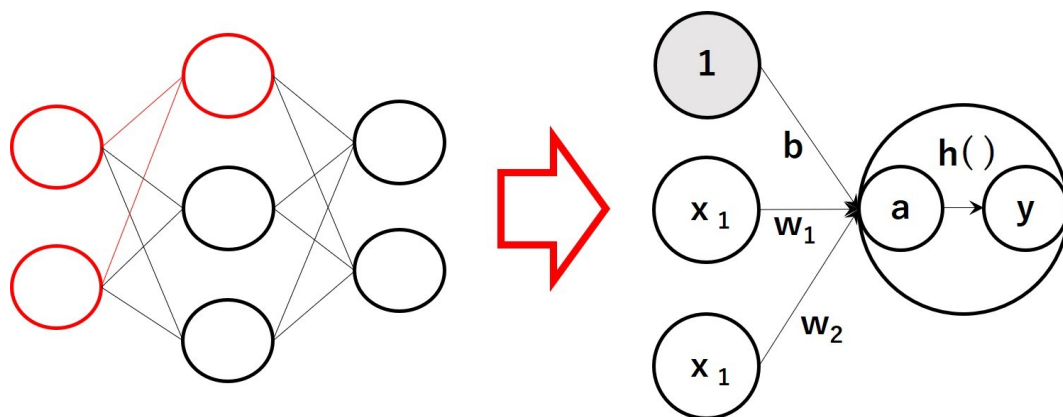


図3.2 ニューラルネットワークの計算の様子

左図はニューラルネットワークの全体図を、右図は左図の赤色の部分の拡大図を示している。ここで x_1 と x_2 と 1 の 3 つの信号がノードへの入力となっている。また b はバイアス、 w_1 と w_2 は各信号の重みを表している。ノードの中に書かれている a は重み付きの入力信号とバイアスの総和を計算したものであり、 $h()$ は活性化関数と呼ばれるものである。ここで x_1, x_2 の入力データが w_1, w_2 のフィルターを通して、次のノードへ伝わっていることがわかる。またバイアスとはノードの発火しやすさを制御するパラメータとなっている。これらが足し合わされ次のノードへの入力となっている。受け取った入力を数式で表すと式 3.1 となり、この入力を活性化関数を通して出力することで、次のノードへの入力となる。 y を出力とすると式 3.2 となる。

$$a = b + w_1x_1 + w_2x_2 \quad (3.1)$$

$$y = h(a) \quad (3.2)$$

活性化関数 $h(a)$ は様々なものが考案されているが、本研究で使用した関数を含めて代表的な関数を次節で紹介する。

3.2.3 活性化関数

活性化関数として本研究で使用した ReLU 関数と同じく代表的なシグモイド関数についての紹介する。

- ReLU 関数

ReLU (Rectified Linear Unit) 関数は入力が 0 を超えれば、その入力をそのまま出力し、0 以下ならば 0 を出力する関数である。入力データとして余分な情報 (ノイズ) が多いデータについて、負の値の出力を切り捨てることで、より特徴をつかみやすい効果を期待できる。

ReLU 関数を数式で表すと、以下のようになる。

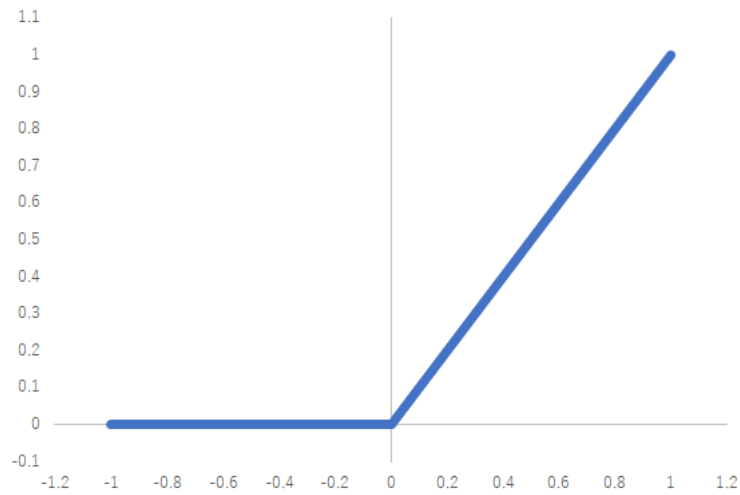


図3.3 ReLU 関数

$$h(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases} \quad (3.3)$$

- シグモイド関数

シグモイド関数は入力した値が大きければ大きいほど 1 に近づき、小さければ小さいほど 0 に近づく関数である。特徴として ReLU 関数と異なり滑らかな曲線となっていることである。これにより入力の値を全て使用することが可能となる。

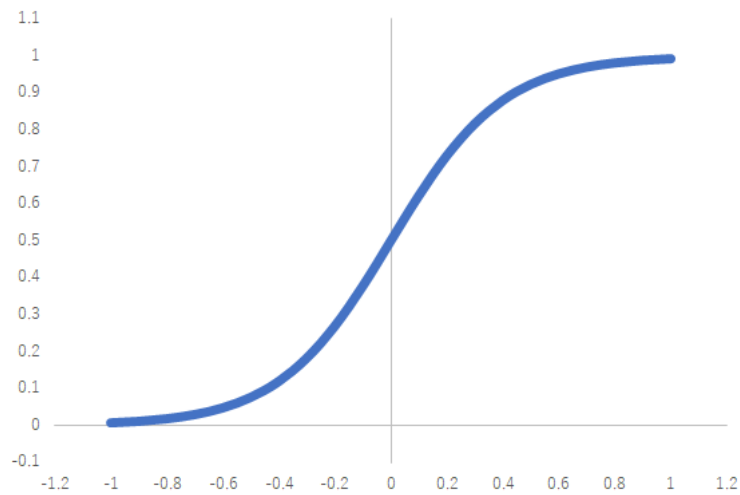


図3.4 シグモイド関数

シグモイド関数を数式で表すと、以下のようになる。

$$h(x) = \frac{1}{1 + \exp(-x)} \quad (3.4)$$

3.2.4 ニューラルネットワークの学習

最終的に出力されたデータは目的変数と比較される。そこで目的変数に対し誤差が大きい場合は、中間層での重み W が適切ではないということなので重み W の更新を行う。この重み W の調整を繰り返して、出力結果が正しい値となるようにする作業がニューラルネットワークの学習である。ここでニューラルネットワークの学習に重要な損失関数と勾配について説明する。

3.2.4.1 損失関数

損失関数 (loss function) とはニューラルネットワークの学習で用いられる指標であり、このニューラルネットワークからの出力と目的変数との誤差を用いて計算を行う。ニューラルネットワークの学習では損失関数から得られる値が小さくなるように重み W を更新する。一般的に回帰問題については 2 乗和誤差 (mean squared error) を、分類問題については交差エントロピー誤差 (cross entropy error) を用いられる。ここでは 2 乗誤差と交差エントロピー誤差について、正答値からの誤差の評価方法について簡単に説明する。

- 2 乗和誤差

2 乗和誤差は数式で次のように表される。

$$E = \frac{1}{2} \sum_k (y_k - t_k)^2 \quad (3.5)$$

ここで y_k はニューラルネットワークの出力、 t_k は目的変数、 k はデータの次元数を表す。

- 交差エントロピー誤差

交差エントロピー誤差は数式で次のように表される。

$$E = - \sum_k t_k \log y_k \quad (3.6)$$

ここで \log は底が e の自然対数を表し、 y_k はニューラルネットワークの出力、 t_k は目的変数、 k はデータの次元数を表す。また目的変数は正答値のインデックスのみが 1 でその他が 0 である、one-hot 表現を用いて表される。例えば、1,2,3 という目的変数があるとすると、one-hot 表現では $[1,0,0], [0,1,0], [0,0,1]$ となる。そのため式 3.6 は実質的に正解の目的変数が 1 に対応するインデックスの値のみ計算することになる。式 3.6 からわかるように y_k が 0 に近づくにつれ (正答値から離れるにつれ) E の値は大きくなり、1 に近づくにつれ (正答値に近づくにつれ) E の値が小さくなるのがわかる。

3.2.4.2 ミニバッチ学習

ニューラルネットワークが学習を行う際に、損失関数は全ての訓練データを対象として求める必要がある。そのため訓練データの数だけ損失関数が存在し、その和が学習の指標となる。学習に使用される際は、損失関数の和をその数で割った平均の損失関数が使用される。しかし一般的にニューラルネットワークに使用されるデータ数は膨大であり、その和を求めるには少々時間がかかってしまう。そこで訓練データの中から一部を選び出し、それを全体の近似として利用する。ここで選び出された塊をミニバッチと呼び、このような学習をミニバッチ学習と呼ぶ。例えば、訓練データが 10000 個あった場合、ミニバッチのサイズを 50 とすると 200 個のデータとなるので、学習時間の高速化が望める。またミニバ

タッチごとに少ないデータでまとめて学習を行いパラメータの更新を行うので、局所解にはまりにくく学習停滞が起きにくい、同時に訓練データにおいて外れ値がある場合に影響を受けやすいという特徴がある。

3.2.4.3 勾配法

勾配法とはニューラルネットワークにおいて重み w を更新する手法である。ニューラルネットワークにおいて最適な重み w は、損失関数の最小化に必要不可欠である。しかし損失関数は複雑であり、重み w のパラメータ空間も広大であるため、最小値を見つけ出すことは困難である。そこで勾配を利用し、関数の最小値を探し出す手法が勾配法である。各地点において関数の値を最も減らす方向を示すのが勾配とし、現在の位置から勾配方向に一定距離だけパラメータを更新する。この作業を繰り返すことで、関数の値を徐々に減らすとができる。

勾配法を数式で表すと以下ようになる。

$$w = w - \eta \frac{\partial f}{\partial w} \quad (3.7)$$

ここで η は更新の量、 f は最小にしたい関数を表す。ニューラルネットワークの学習においては学習率 (Learning rate) と呼ばれる。式 3.7 は 1 回の更新式を示しており、このステップを繰り返す行うことで目的の関数の値を徐々に減らしていくことができる。

また勾配法を行う最適化アルゴリズム (オプティマイザ) は多数存在する。以下に例としてニューラルネットワークでよく使用される SGD、RMSprop、本研究でを使用した Adam について説明する。

- SGD

SGD とは stochastic gradient descent (確率的勾配降下法) の略である。数式で表すと式 3.8 となる。

$$w^{t+1} \leftarrow w^t - \eta \frac{\partial f(w^t)}{\partial w^t} \quad (3.8)$$

ここで更新する重みパラメータを w 、 w に関する損失関数の勾配を $\frac{\partial f}{\partial w}$ である。また η は学習係数を表している。また式中の \leftarrow は右辺の値で左辺の値を更新するということを意味している。SGD は勾配方向へのある一定距離だけ進むという単純な方法である。また w の更新ごとにランダムにサンプルを選び直すことで局所的な最適解への収束を防いでいる。

- RMSprop

RMSprop とは Adagrad という学習率の変化を可能にしたオプティマイザの改善策である。数式で表すと式 3.9 となる。

$$\begin{aligned} h_t &= \alpha h_{t-1} + (1 - \alpha) \nabla f(w^t)^2 & (h_0 = 0) \\ \eta_t &= \frac{\eta_0}{\sqrt{h_t} + \epsilon} \\ w^{t+1} &\leftarrow w^t - \eta_t \nabla f(w^t) \end{aligned} \quad (3.9)$$

ここで更新する重みパラメータを w 、 η は学習係数を示し、式中の \leftarrow は右辺の値で左辺の値を更新するということを意味している。また α は過去の勾配の影響を抑えるパラメータ、 ϵ は正の微小量となっている。過去の勾配を記憶し、過去の勾配から学習係数の自動調整を行う。また過去の勾配の記憶は徐々に失われ、新しい勾配の情報が大きく反映されるようになっている。これにより学習が進むにつれ学習係数を次第に小さくできるので、学習の収束を早める効果が得られる。

- Adam

Adam とは Adaptive moment estimation の略である。数式で表すと式 3.10 となる。

$$\begin{aligned}
 m_{t+1} &= \beta_1 m_t + (1 - \beta_1) \nabla f(w^t) \\
 v_{t+1} &= \beta_2 v_t + (1 - \beta_2) \nabla f(w^t)^2 \\
 \hat{m} &= \frac{m_{t+1}}{1 - \beta_1^t} \\
 \hat{v} &= \frac{v_{t+1}}{1 - \beta_2^t} \\
 w^{t+1} &\leftarrow w^t - \alpha \frac{\hat{m}}{\sqrt{\hat{v} + \epsilon}}
 \end{aligned} \tag{3.10}$$

ここで更新する重みパラメータを w 、 α は学習係数を示し、式中の \leftarrow は右辺の値で左辺の値を更新するという意味している。また β は過去の勾配を抑えるパラメータ、 ϵ は正の微小量となっている。RMSprop と同様に過去の勾配を記憶し、過去の勾配から学習係数の自動調整を行う。RMSprop と異なり過去の勾配を指数関数的に減衰させることによって、学習の収束をさらに早めることができる。また既存のオプティマイザーの利点を集めているので、現在最も評価されているオプティマイザーとなっている。

3.2.5 ディープニューラルネットワーク

ここまでニューラルネットワークの計算方法や学習方法について述べてきた。ディープニューラルネットワークとは文字通り、ニューラルネットワークをベースに中間層を増やした、つまり層を深くしたネットワークである。中間層を増やすことで、ニューラルネットワークと比較し以下のような利点がある。

- ネットワークのパラメータ数を少なくすることができる。
中間層を増やすことは、ネットワークの中で中間層が 1 層のニューラルネットワークを繰り返すことと同じである。これにより 1 つのニューラルネットワークのパラメータ数を抑えることができる。またパラメータ数の差は中間層を深くするにつれて大きくなる。
- 学習の効率を上げることができる。
中間層が浅いネットワークと比較し、中間層が深いネットワークでは、学習データを少なくすることができ、高速で学習が行うことができる。例えば、「犬」を認識する問題について考えると、中間層が浅いネットワークでは、一度に「犬」の特徴を理解する必要がある。「犬」には様々な犬種があり、また撮影環境に見え方は依存します。そのため「犬」の特徴を理解するためには、多くのバリエーションに富んだ学習データが必要となるので、学習に時間がかかる。中間層が深いネットワークでは学習すべき問題について階層的に分解することが可能なので、それぞれの層が学習すべき問題は単純な問題に落とし込むことができ、学習の高速化が望める。
- 高度なパターン認識を行うことができる。
それぞれの中間層で階層的に情報を渡しているため、中間層から次の中間層への学習において、抽出した情報を効率的に使用することができる。それによってネットワークで学習したモデルが、より高度なパターンを表現することができる。

図 3.5 にディープニューラルネットワークの例を示す。図 3.1 と比較すると中間層が増え、ネットワー

クが深くなっていることがよくわかる。

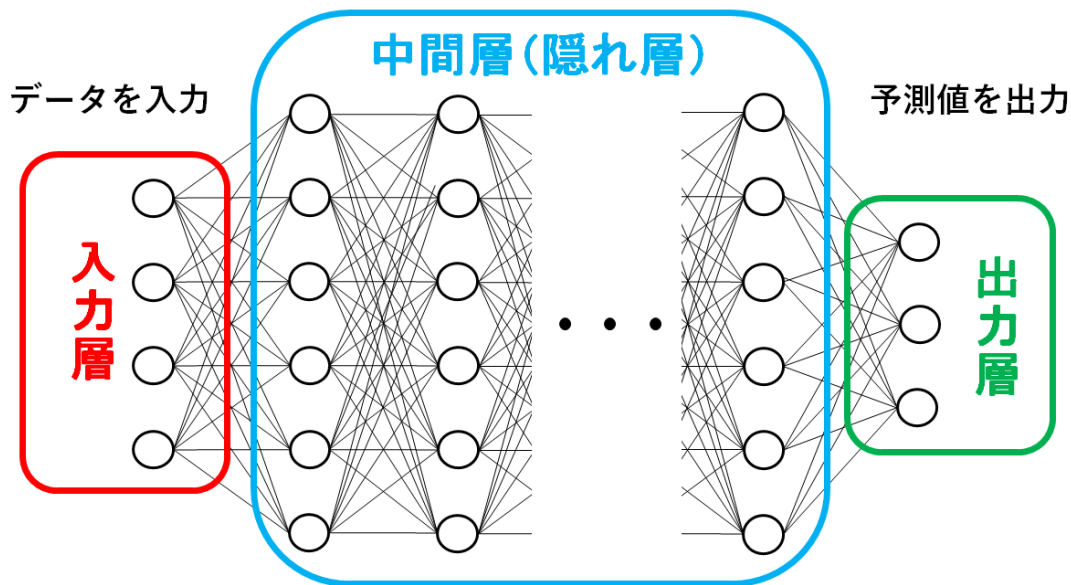


図3.5 ディープニューラルネットワークの例

本研究ではディープニューラルネットワークを使用した開発を行った。またディープニューラルネットワークを構成するハイパーパラメータについては、3.2.7で紹介する。

3.2.6 過学習

ニューラルネットワークの学習においての問題の1つに過学習がある。過学習とはニューラルネットワークが学習を進める際に、訓練データだけに適応しすぎてしまい、訓練データに含まれない他のデータにはうまく対応できない状態を言い、汎用性の高いモデルを目指す上で問題となる。また過学習が起きる原因としては主に以下が挙げられる。

- パラメータを大量に持ち、表現力の高いモデルであること
- 訓練データが少ないこと

またここで過学習を抑制する手法として一般的な手法である荷重減衰と、本研究で使用したドロップアウトを以下で説明する。

3.2.6.1 荷重減衰

荷重減衰 (Weight decay) とは過学習抑制において古くから使用される手法である。過学習は重み W が大きな値をとることで発生することが多いことから、学習の過程において大きな重み W を持つことに対してペナルティを課すことで過学習を抑制する。ニューラルネットワークの学習において、損失関数に重み W の2乗ノルム (L2 ノルム) である $\frac{1}{2}\lambda W^2$ を加算する。ここで λ は正則化の強さを表す。 λ を大きく設定すればするほど、大きな重み W に対し強いペナルティを課すことができる。

3.2.6.2 ドロップアウト

ドロップアウト (Dropout) とは中間層のノードをランダムに消去しながら学習する手法である。ニューラルネットワークの学習において、訓練時に中間層のノードをランダムに選出し、その消去を行う。

図 3.6左図のニューラルネットワークから、右図のようにノードは消去される。消去されたノードは図 3.6右図のように信号伝達が行われなくなる。テスト時には全てのノードに信号を伝達する。ノードをランダムに消去することから過学習を抑制することができる。

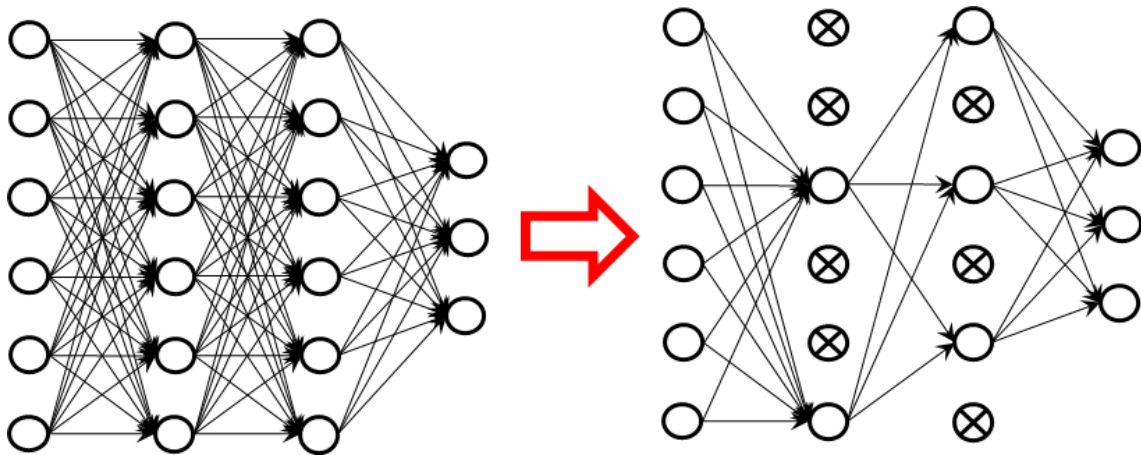


図3.6 ドロップアウトの概念図

3.2.7 ハイパーパラメータ

これまでに説明してきた通り、ハイパーパラメータとは機械学習アルゴリズムが力を発揮するために必要不可欠なものとなっている。しかし重み W などの学習アルゴリズムが自動で獲得できるパラメータと異なり、人の手で設定する必要があるパラメータのことである。ハイパーパラメータの値を適切な値に設定しなければ、性能の悪いモデルになり正しく学習を行うことができない。以下でハイパーパラメータの詳細について説明する。

またハイパーパラメータの設定は重要にもかかわらず、特に DNN ではハイパーパラメータの数は多く手動での調整にはかなりの時間を費やす作業となっている。そこで本研究では optuna というハイパーパラメータの最適化を自動化するためのソフトウェアフレームワークを使用し最適化を行った。

- エポック数

エポック (epoch) 数とは一度の学習において訓練データを使用した回数である。エポックとは単位を表しており、学習において訓練データ全てを 1 度使い切った時が 1 エポックとなる。エポック数が多いほど訓練データを繰り返し使用するので、学習に時間がかかる。またエポック数が少なすぎると学習を終えることができず、逆に多すぎると過学習を起こす。

- バッチサイズ

バッチサイズとはミニバッチ学習の際のミニバッチの大きさである。バッチサイズが大きいほど学習を行う回数が減るので時間を短縮できるが、学習が局所解にはまりやすくなる。逆にバッチサイズが小さいほど学習を行う回数が増えるので時間がかかるが、学習が局所解にはまりにくくなる。またバッチサイズが大きいほど一度の学習に使用するデータサイズが増える、パソコンのメモリに負荷がかかる、そのためバッチサイズの最大値はパソコンに依存する。

- ドロップアウト率

ドロップアウト率とは、過学習を学習する手法の一種であるドロップアウトを行う比率である。

ドロップアウト率が高いほど消去するノードが増えるので過学習を抑制することができる。しかしノードを消去してしまうので、ドロップアウト率が高いほど学習が進みにくくなる。

- オプティマイザーと学習率

オプティマイザーとは勾配法の最適化アルゴリズムであり、学習率とは一度の学習でどれだけの学習を行うか、どれだけのパラメータの更新を行うかを決定する割合である。オプティマイザーにより学習の収束の速さが異なる。また学習率は式 3.7において η で表され、学習率が大きいほど学習が早くなり、学習率が小さいほど学習が遅くなる。また学習の結果としてはエポック数と同じ様な効果が得られる。

- 中間層ノード数

中間層ノード数とは各中間層にあるノードの個数である。ノード数が大きいほど様々な特徴を抽出することができるが、余計な特徴まで拾い、またそれぞれノードごとの相互作用が小さくなり学習がうまく進まない場合もある。またノード数が小さいほど入力データの圧縮ができるが、本来必要な情報まで捨ててしまう可能性もある。

- 中間層数

中間層数とは中間層の層数である。中間層の層数が多いほど、より複雑なモデルを学習できる。しかしニューラルネットワークの学習は誤差の伝播により行われているので、層数が増えるほど正しく誤差を伝播していくことが困難になる。そのため過度に層数を増やすと、局所的な解につかまりやすくなり、過学習を起こす可能性もある。

3.3 動作環境

表 3.1に使用したコンピュータのスペックを示す。

OS	Ubuntu18.04.2 LTS
OS 種別	64 ビット
メモリー	31GiB
プロセッサ	Intel(R) Core(TM) i7-8700K CPU @ 3.70Ghz
グラフィックカード	NVIDIA Corporation GP104 [GeForce GTX 1080]

表3.1 動作環境のまとめ

表 3.2にディープニューラルネットワークを行う際に使用したモジュールのバージョン一覧を示す。表中の線より下は GPU を使用して学習を行う際に必要なモジュールである。

モジュール	バージョン
pip	19.2.3
h5py	2.9.0
numpy	1.16.4
scikit-learn	0.21.2
keras	2.2.4
tensorflow-gpu	1.13.1
optuna	0.17.0
cuda	10.0.130

表3.2 デープニューラルネットワークのライブラリのバージョン一覧

第 4 章

機械学習を使用した加速器運転パラメータ予測

本章では本研究で行った機械学習による加速器運転パラメータ予測について、その手法と結果を示す。予測結果については、全て DNN での試行を 3 回繰り返し、その平均の結果を載せている。

4.1 機械学習を使用した加速器運転パラメータ予測の方針

本研究では膨大なデータの中から相関を見つけることが得意とする DNN を使用し、過去の加速器運転パラメータを DNN に学習させ、そこから入射効率が高くなるような加速器運転パラメータの推定を行う。

DNN を使用した教師あり学習を用いた加速器運転パラメータの推定として、Linac 加速器の過去のデータから良いビームのデータを抽出し、その他のパラメータから、目的の加速器調整パラメータを予測できるネットワークをデザインすることを目標とする。

図 4.1 に DNN を使用した加速器運転パラメータの予測の流れを示す。入力パラメータとして加速器データを、入射効率が低いデータで学習した DNN に入力することで、入射効率が高くなるような加速器パラメータを出力パラメータとして得る。



図4.1 DNN を使用した加速器運転パラメータの予測の流れ

4.2 加速器データの収集

本研究では Super KEKB HER 用の 1nC 電子ビームにおいての、Linac 加速器運転データを 2018 年 11 月 12 日から 2019 年 11 月 39 日までのうち 120 日相当、1,389,000 shot を収集し研究に使用した。また開発を行うにあたり、2019 年夏季の長期 shutdown の前後でデータを分割し、以下のように名前をつけた。

データ1 2018年11月12日から2019年6月30日まで 1,180,500 shot

データ2 2019年10月17日から2019年11月30日まで 208,500 shot

ここでデータ1の中には1回分の冬季の shutdown が含まれているが、ここではまとめて shutdown 前としている。

また収集データに含まれるパラメータの種類と数は以下のとおりである。

- ビーム位置モニター (BPM) 情報 (ビームの位置や電荷量) : 4600 個
- RF 情報 (クライストロンの RF 位相とパワー) : 340 個
- 電磁石情報 (電磁石励起電流値読み返し値) : 811 個
- 環境情報 (環境温度や冷却水温など) : 732 個
- ユーティリティ情報 (機械室の冷却吹ポンプ圧力など) : 324 個

それぞれのパラメータの収集方法については、Linac 加速器の分散型ネットワーク制御システムの EPICS を使用し、自動で取得できるようにした。その際、各 shot でのパラメータの厳密な同期は行わなかった。EPICS を用いてパラメータを読み出す際にネットワークを介するので、各パラメータの収集時間の差が最大で数秒生じてしまう。そのため本研究で収集した 1 shot ごとのデータは本当の意味での 1 shot ごとのデータではない。しかし本研究ではその時間での細かい加速器パラメータの調整は行われず、同一の shot として問題ないとして扱っている。

4.3 加速器運転調整パラメータ予測の指標

本研究の最終的な加速器運転調整の目標は入射効率向上である。入射効率とは前段の加速器から後段の加速器にビームを入射する際、前段の加速器でのビームのうち、後段の加速器で入射されるビームの割合である。そのため入射効率は、ビームの位置やタイミング、粒子の大きさなど様々なパラメータに依存し、また 1shot ごとで決まる量でもない。このように入射効率は様々な要因を含んだ値であり、調整指標にするのは困難と考えられる。そこで本研究では入射効率に代わる調整の指標として、Linac 加速器の上流と下流のビーム位置モニター (BPM) における、電荷量の比 Q_{ratio} を用いた。

$$Q_{ratio} = \frac{\text{Linac 加速器下流の BPM(SP_58_0) での電荷量}}{\text{Linac 加速器上流の BPM(SP_A1_M) での電荷量}} \quad (4.1)$$

Q_{ratio} はビームが生成されたところでの電荷量から見た、Linac 加速器から出るところでの電荷量となっている。これより Q_{ratio} が高くなるように調整を行うことで、Linac 加速器の上流から下流までのビーム損失が少なくすることができる。

図 4.2 に Linac 加速器 A Sector の配置図、図 4.3 に Linac 加速器 5 Sector の配置図を示す。また図の赤い円で囲まれたものがそれぞれ使用した BPM(SP_A1_M, SP_58_0) である。

今回収集したデータの中に異なる電子銃を使用したデータが混在していたため、加速器上流の BPM として A Sector 最初の BPM を使用するのではなく、様々な電子銃からの経路が収束した点での BPM(SP_A1_M) を使用した。

Linac 加速器の最下流には Switch Yard が存在するが、経路がわかれており複雑であるため、Linac 加速器下流の BPM としては全てのビームが通る 5 Sector 最後の BPM(SP_58_0) を使用した。

図 4.4 に Q_{ratio} を調整の方針とした場合の、DNN を使用した加速器運転パラメータの予測の流れを示す。入力パラメータとして加速器データを、 Q_{ratio} が高いデータで学習した DNN に入力することで、 Q_{ratio} が高くなるような加速器パラメータを出力パラメータとして得る。今回は Q_{ratio} が 0.9 以上のデータを Q_{ratio} が高いデータとして扱った。

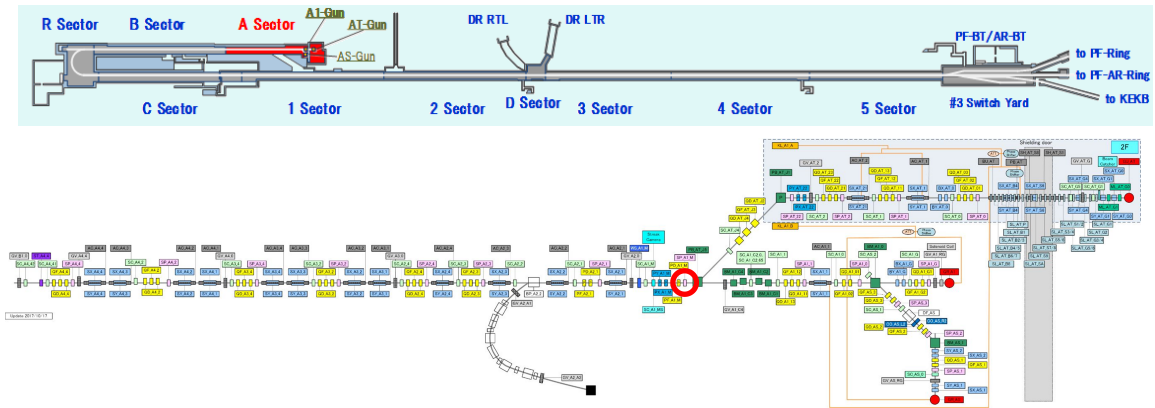


図4.2 Linac 加速器 A Sector の配置図

上図の赤色の部分が A Sector の位置であり、下図が A Sector の拡大図となっている。また下図の右方が Linac 加速器の上流側、左方が下流側である。

下図の赤い円で囲まれた BPM が加速器上流として使用した BPM(SP_A1_M) である。

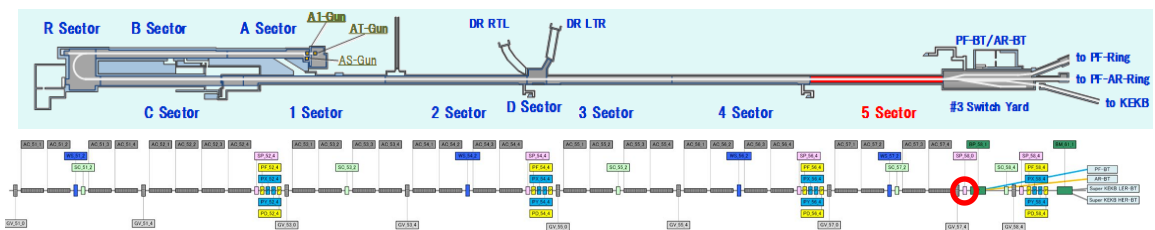


図4.3 Linac 加速器 5 Sector の配置図

上図の赤色の部分が 5 Sector の位置であり、下図が 5 Sector の拡大図となっている。また下図の左方が Linac 加速器の上流側、右方が下流側である。

下図の赤い円で囲まれた BPM が加速器下流として使用した BPM(SP_58_0) である。



図4.4 DNN を使用した加速器運転パラメータの予測の流れ (Q_{ratio} を用いた場合)

4.4 ネットワークデザイン

次に作成したネットワークについて簡単に説明する。図 4.5 に作成したネットワークの概略図を示す。赤色のひし形が入力層、青色の各丸四角形が中間層、紫色の三角形がドロップアウト層、緑色の円形が出力層を表している。このネットワークではドロップアウト層は中間層の間に挿入されており、図からわかる通り上流側から順番に、入力層、中間層が n 層 (n はネットワークにより異なる)、ドロップアウト層、中間層が 1 層、出力層という順番になる。

またネットワークには様々なハイパーパラメータが存在し、その値によってネットワークの性能が異なってしまう。本研究ではハイパーパラメータの最適化を行い、与えるデータによって最適な予測がで

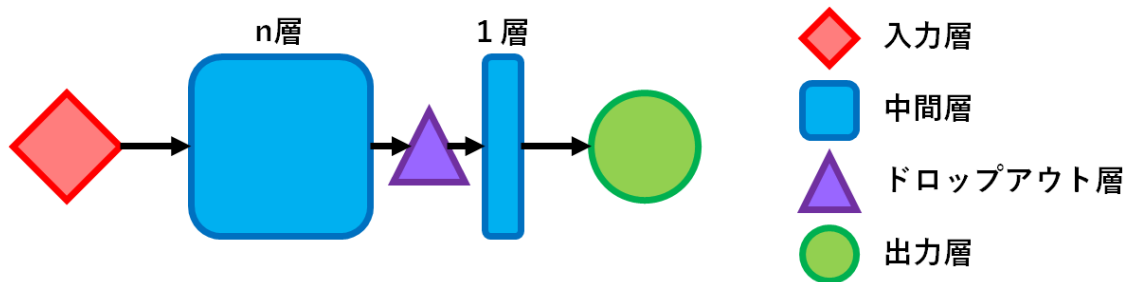


図4.5 作成したネットワークの概略図

きるようにした。最適化を行ったパラメータは以下のパラメータである。

- バッチサイズ
- 学習率
- 中間層のノード数
- 中間層数 (図 4.5の n 層部分)
- ドロップアウト率

各ハイパーパラメータの値については後の各予測について、それぞれで各説明変数ごとにまとめたものを表で示す(中間層数については図 4.5の n について示す)。またこの簡条書きにないハイパーパラメータについては本研究を通して以下のように固定し使用している。

- エポック数: 50
- オプティマイザー: Adam
- 活性化関数: Relu

4.5 データ 1 を使用した予測

本節では収集データのうちデータ 1 のみを使用した最適値の予測について示す。

4.5.1 RF パラメータの予測

最適な Q_{ratio} を得るための RF パラメータを予測する DNN の開発を行った。

4.5.1.1 データセットの作成

今回予測を行う RF パラメータを選ぶに際し、Linac 加速器の共同研究者の方に、入射効率の向上に効果的なパラメータをお聴きした。その中でもパラメータの値がばらけており、DNN での予測が可能でありそうなパラメータとして、「B Sector5 個目にあるクライストロン (KL_B5) が出力する RF 位相パラメータ」を選別した。図 4.6に B Sector の配置図を示す。上図の赤色の部分が Linac 加速器中の B Sector の位置、下図が B Sector の拡大図となっている。図の赤色で囲った加速管が RF 位相パラメータとして使用したクライストロン (KL_B5) が高周波を送っている加速管である。

図 4.7に選別した B Sector 5 個目にあるクライストロン出力の RF 位相パラメータと Q_{ratio} の関係を示す。縦軸が RF 位相パラメータ、横軸が Q_{ratio} となっている。また図中のオレンジ色の直線が $Q_{ratio} = 0.9$ を表している。

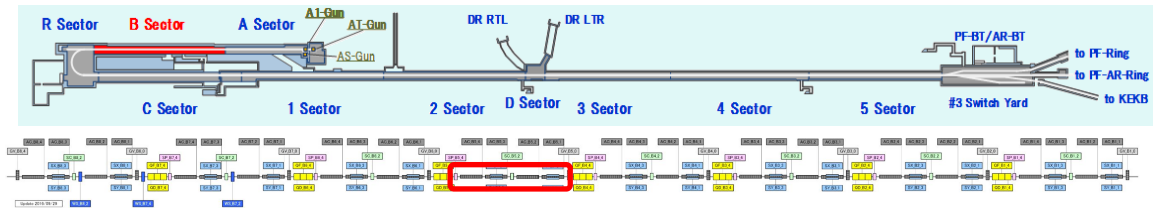


図4.6 Linac 加速器 B Sector の配置図

上図の赤色の部分が B Sector の位置であり、下図が B Sector の拡大図となっている。また下図の右方が Linac 加速器の上流側、左方が下流側である。

下図の赤い円で囲まれた加速管が RF 位相パラメータとして使用したクライストロン (KL_B5) が高周波を送っている加速管である。

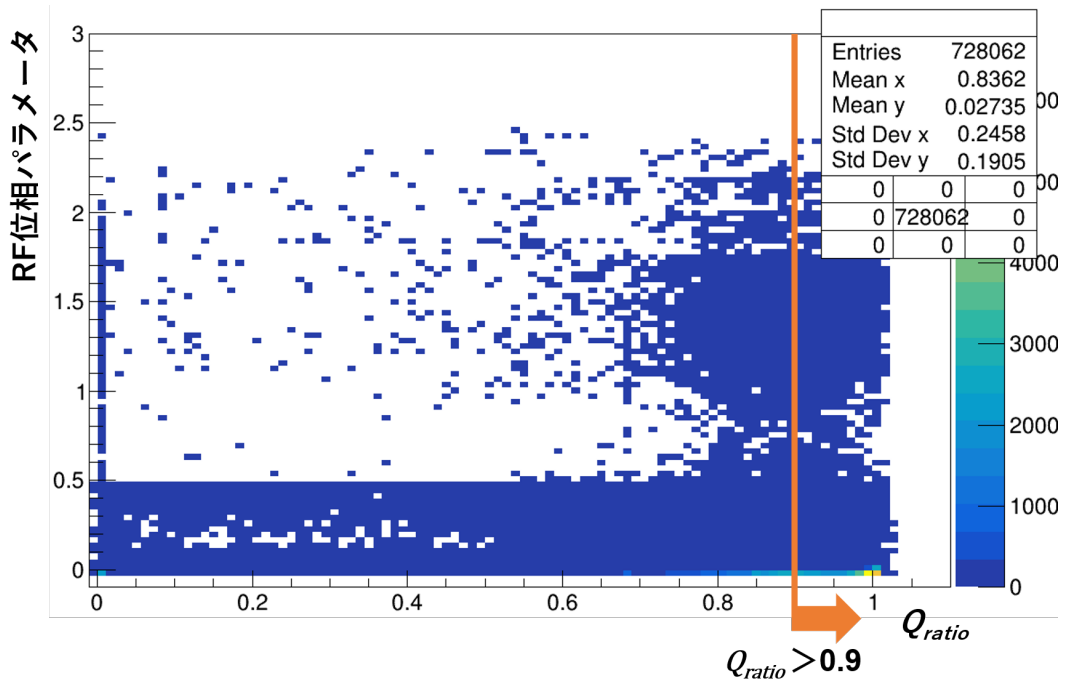


図4.7 B Sector5 個目にあるクライストロン出力の RF 位相パラメータと Q_{ratio} の関係

図から RF 位相パラメータと Q_{ratio} には直接的な相関がないことがわかる。収集データの中には Q_{ratio} が高いデータから低いデータまで、つまり良いビームから悪いビームまで存在していることがわかる。本研究では Q_{ratio} が高いときのパラメータ予測を行うため、 Q_{ratio} が 0.9 以上のデータを抽出して DNN での学習を行う。

これより DNN で使用するパラメータは以下のようになる。

- 目的変数
 - B Sector5 個目にあるクライストロン (KL_B5) が出力する RF 位相パラメータ
- 説明変数 1
 - クライストロン (KL_B5) のパラメータとクライストロン (KL_B6) のパラメータを除いた全 RF パラメータ
 - : 合計 328 パラメータ
- 説明変数 2

クライストロン (KL_B5) のパラメータとクライストロン (KL_B6) のパラメータを除いた全 RF パラメータ、全環境パラメータ

: 合計 1060 パラメータ

- 説明変数 3

クライストロン (KL_B5) のパラメータとクライストロン (KL_B6) のパラメータを除いた全 RF パラメータ、全電磁石パラメータ、全環境パラメータ

: 合計 1871 パラメータ

目的変数は今回予測をしたいパラメータ、説明変数は目的変数を説明できる可能性があるパラメータ群である。説明変数は 3 種類用意した。1 種類目は全 RF パラメータのうち、今回目的変数として使用したクライストロン (KL_B5) のパラメータと、加速器調整を行う際に KL_B5 と関連させ調整を行うクライストロン (KL_B6) のパラメータを除いたもの。2 種類目は 1 種類目のパラメータに、全環境パラメータを加えたもの。3 種類目は 2 種類目のパラメータに、全電磁石パラメータを加えたもの。このそれぞれの説明変数から目的変数を予測する DNN を設計し、比較を行う。

また抽出データは全体で 396,538 データである。そこからデータをランダムに 8:1:1(データ数 317,230:39,654:39,654) に分割し、学習用、検証用、テスト用とした。以下で扱う結果の値はテストデータを使用した結果となっている。

4.5.1.2 回帰学習を用いた予測結果

DNN の回帰学習を用いて、RF 位相パラメータの予測を行った。表 4.1 にそれぞれの説明変数について、最適化を行ったハイパーパラメータを示す。

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1	1000	0.001	100	10	0.14
説明変数 2	8000	0.0035	750	9	0.49
説明変数 3	4000	0.0077	850	9	0.16

表4.1 データ 1 を使用した RF 位相パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ

DNN の回帰学習の結果を以下に示す。図 4.8 にそれぞれの説明変数の場合の、RF 位相パラメータの正解値と予測値の関係を示す。それぞれの図について、横軸が RF 位相パラメータの正解値、縦軸がそれぞれの RF 位相パラメータでのネットワークが出力した予測値となっている。また左図が説明変数 1、中図が説明変数 2、右図が説明変数 3 を使用した場合の結果となっている。表 4.2 にそれぞれの説明変数の際の回帰学習の結果をまとめる。loss とは損失関数の値であり、予測値と正解値の乖離の大きさを表している。つまり loss 値が小さいほど良いネットワークを作成されたという意味になっている。loss の算出には平均二乗誤差 (mse) を使用している。

	loss
説明変数 1	0.00719
説明変数 2	0.0300
説明変数 3	0.0301

表4.2 データ 1 を使用した RF 位相パラメータの回帰予測の結果まとめ

説明変数 1 の場合は概ね学習が成功しているように思われるが、縦軸と横軸の 0 付近に点が固まって

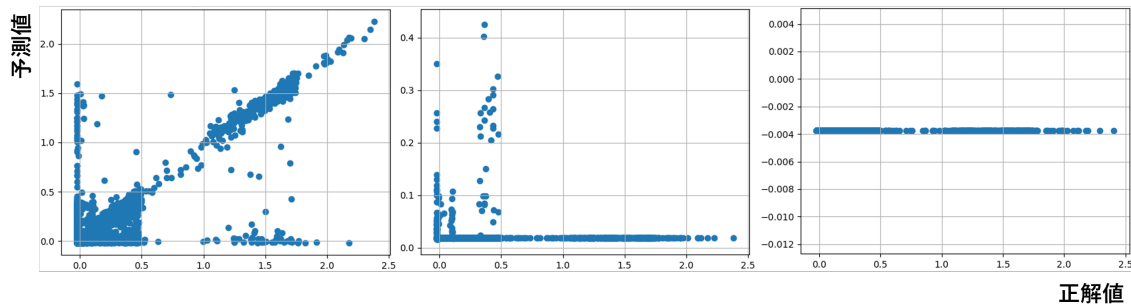


図4.8 データ1を使用したRF位相パラメータの回帰予測の結果まとめ
 正解値と予測値のパラメータの関係
 左図: 説明変数1の結果、中図: 説明変数2の結果、右図: 説明変数3の結果

おり、完全には学習ができていないことがわかった。また説明変数2と説明変数3の場合は学習がうまくできていないことがわかった。この理由としては次のような理由が考えられる。図4.7から、RF位相パラメータの値が局所的に固まっていることがわかる。今回入力データとして Q_{ratio} が範囲内のRF位相パラメータを全て入れたので、数が多いパラメータにネットワークの学習が引っ張られてしまい、学習が上手くいかなかったと考えられる。

そこでこの問題を解決するために、RF位相パラメータの値を分割し、それぞれのパラメータのデータ数を整え、ネットワークに入力する方法を考えた。次節でその方法として分類学習を用いた場合の方法と結果を示す。

4.5.1.3 分類学習を用いた予測

図4.9にRF位相パラメータの値を分割した際の Q_{ratio} とRF位相パラメータの関係を示す。図4.7と同様に、縦軸がRF位相パラメータ、横軸が Q_{ratio} となっており、灰色の直線が $Q_{ratio} = 0.9$ を表している。オレンジ色の直線は今回RF位相パラメータの値を分割した境界となっており、赤色の数字は赤色の吹き出しの部分に設定したパラメータとなっている。

以下に分割したそれぞれのパラメータのパラメータの範囲とデータ数を示す。

- パラメータ0 RF位相パラメータが、0以下の値：データ数 44120
- パラメータ1 RF位相パラメータが、0以上0.1以下の値：データ数 20000
- パラメータ2 RF位相パラメータが、0.1以上0.2以下の値：データ数 5196
- パラメータ3 RF位相パラメータが、0.2以上0.3以下の値：データ数 4625
- パラメータ4 RF位相パラメータが、0.3以上0.4以下の値：データ数 10760
- パラメータ5 RF位相パラメータが、0.4以上0.5以下の値：データ数 11592
- パラメータ6 RF位相パラメータが、0.5以上の値：データ数 3707

ここでデータ数はそれぞれのパラメータで異なりすぎないように調整している。またデータ数の合計は100,000データである。回帰学習の場合と同様にそこからデータをランダムに8:1:1(データ数80,000:10,000:10,000)に分割し、学習用、検証用、テスト用とした。以下で扱う結果の値はテストデータを使用した結果となっている。また説明変数1,2,3と目的変数は4.5.1.1節と同様である。

また表4.3にそれぞれの説明変数について、最適化を行ったハイパーパラメータを示す。

DNNの分類学習の結果を以下に示す。表4.4、表4.5、表4.6にそれぞれの説明変数の場合の、RF位相パラメータの正解値と予測値の関係を示す。それぞれの表について、各列がRF位相パラメータの正

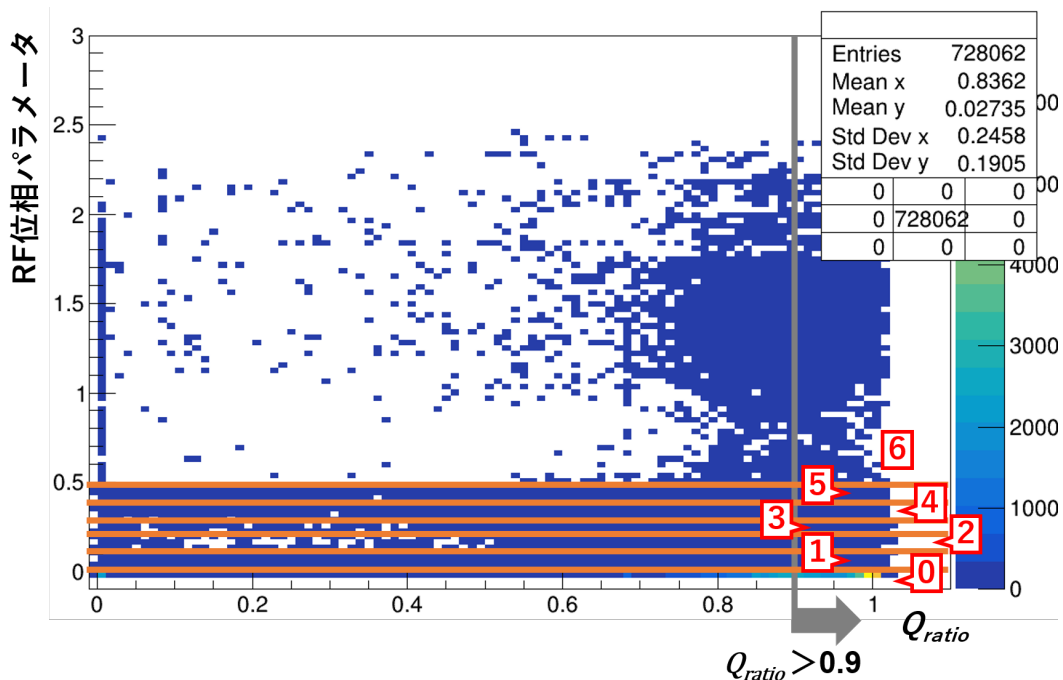


図4.9 パラメータを分類した際の B Sector5 個目にあるクライストロン出力の RF 位相パラメータと Q_{ratio} の関係

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1	7000	0.0031	100	8	0.29
説明変数 2	1000	0.007	100	4	0.13
説明変数 3	1000	0.054	100	9	0.29

表4.3 データ 1 を使用した RF 位相パラメータの分類予測で使用したネットワークのハイパーパラメーターまとめ

解値、各行がそれぞれの RF 位相パラメータでのネットワークが出力した予測値となっており、それぞれのセルが RF 位相パラメータの正解値に対してネットワークが出力した予測値のデータ数となっている。つまり左下から右上への対角のセルが正しく予測できたデータ数となっている。また表 4.4 が説明変数 1、表 4.5 が説明変数 2、表 4.6 が説明変数 3 を使用した場合の結果となっている。また表 4.7 にそれぞれの説明変数の際の回帰学習の結果をまとめる。ここで acc は正答率を表しており、loss は回帰学習と同様に予測値と正解値のズレの大きさを表している。loss の算出には多クラス交差エントロピー (categorical cross entropy) を使用している。

これより DNN の分類学習を使用すると、回帰学習の場合と異なり説明変数 1 に加え、説明変数 2 と説明変数 3 の全ての場合で正しく学習を行うことができた。全ての場合で正答率が 9 割に近い結果となったが、特に説明変数 2 の「RF パラメータと環境パラメータ」を使用した場合が最も良い結果が得られた。また表 4.4、表 4.5、表 4.6 からわかる通り、不正解のデータも正解値のセルの隣のセルに多く集まっていることがわかる。ここで各パラメータは自身で設定したものであり、本来のパラメータは連続値になっているので、隣のパラメータと不正解したとしても、完全に学習が誤りとは考えにくい。これより結果として既に高い正答率になっているが、この「不正解ではあるが、正解に近いデータ」を加味することで、DNN での学習は正しく行われていたと考えられる。

		正解値						
		0	1	2	3	4	5	6
予測値	6	0	0	14	4	3	3	344
	5	26	29	19	5	138	956	1
	4	11	9	21	34	810	114	1
	3	1	1	46	385	22	10	21
	2	2	11	428	39	12	23	9
	1	40	1812	30	2	13	15	0
	0	4352	105	6	1	38	34	0

表4.4 データ1を使用した説明変数1の場合のRF位相パラメータの回帰予測の結果

		正解値						
		0	1	2	3	4	5	6
予測値	6	0	0	5	8	3	5	299
	5	10	7	10	2	74	865	1
	4	13	16	20	20	930	172	2
	3	0	0	16	343	22	2	2
	2	1	8	461	88	16	14	26
	1	69	1910	36	2	22	27	0
	0	4332	60	3	0	29	28	0

表4.5 データ1を使用した説明変数2の場合のRF位相パラメータの回帰予測の結果

		正解値						
		0	1	2	3	4	5	6
予測値	6	0	0	2	3	2	0	303
	5	11	5	25	11	134	945	7
	4	34	8	1	19	850	143	0
	3	5	2	76	414	29	11	35
	2	0	0	314	19	12	21	18
	1	13	1859	33	3	32	25	0
	0	4327	178	11	0	27	33	0

表4.6 データ1を使用した説明変数3の場合のRF位相パラメータの回帰予測の結果

これよりDNNを使用したRF位相パラメータの予測は正しく行えたと考えられる。

4.5.2 ステアリング電磁石パラメータの予測

最適な Q_{ratio} を得るためのステアリング電磁石パラメータを予測するDNNを、RFパラメータと同様に開発を行った。

	acc	loss
説明変数 1	0.905	0.342
説明変数 2	0.915	0.272
説明変数 3	0.897	0.343

表4.7 データ 1 を使用した RF 位相パラメータの回帰予測の結果まとめ

4.5.2.1 データセットの作成

ステアリング電磁石パラメータについては RF パラメータと異なり、入射効率の向上に効果的なパラメータをお聴きできなかったため以下のような方法でパラメータの選別を行った。

図 4.10 にステアリング電磁石の選別の例を示す。ステアリング電磁石パラメータと Q_{ratio} の関係の例を表しており、縦軸がステアリング電磁石パラメータ、横軸が Q_{ratio} となっている。また灰色の破線は $Q_{ratio} = 0.3$ と $Q_{ratio} = 0.6$ を表している。ここで黒色の点はデータプロットを、青色、緑色、赤色の円はそれぞれ $Q_{ratio} < 0.3$ (低)、 $0.3 \leq Q_{ratio} < 0.6$ (中)、 $0.6 \leq Q_{ratio}$ (高) の範囲にあるデータの平均値の点を表している。この低と中、高の値が低 < 中 < 高、もしくは低 > 中 > 高となるステアリング電磁石パラメータを、 Q_{ratio} と相関があるパラメータと考え抽出した。

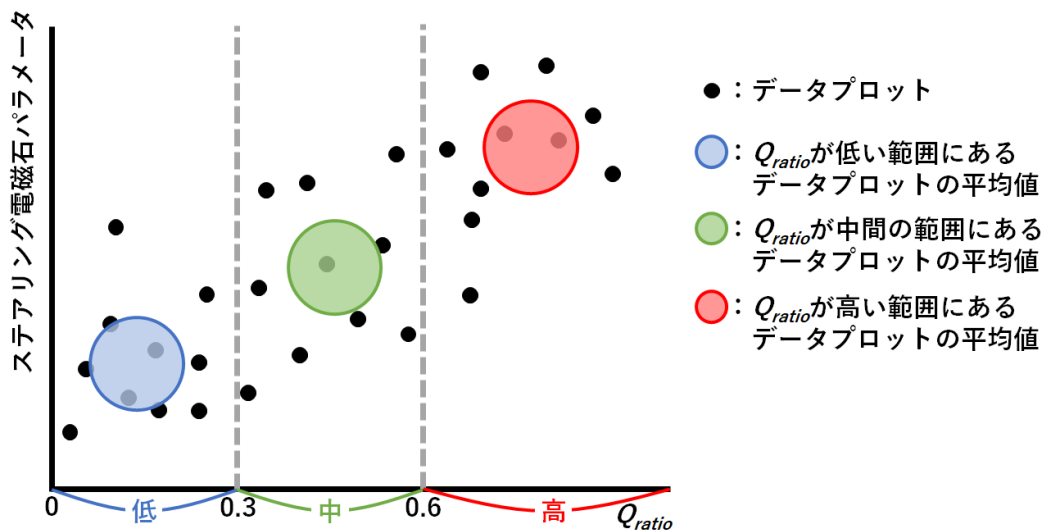


図4.10 ステアリング電磁石の選別の例

この方法を使用し、「3 Sector2 個目にあるステアリング電磁石の垂直方向 (PY_32_4) のパラメータ」と「4 Sector4 個目にあるステアリング電磁石の垂直方向 (PY_44_4) のパラメータ」を選別した。図 4.11 に 3 Sector の配置図を示す。上図の赤色の部分が Linac 加速器中の 3 Sector の位置、下図が 3 Sector の拡大図となっている。図の赤色の円で囲まれたステアリング電磁石が使用したステアリング電磁石 (PY_32_4) である。同様に図 4.12 に 4 Sector の配置図を示す。上図の赤色の部分が Linac 加速器中の 4 Sector の位置、下図が 4 Sector の拡大図となっている。図の赤色の円で囲まれたステアリング電磁石が使用したステアリング電磁石 (PY_44_4) である。

図 4.13 に選別した 3 Sector 2 個目にあるステアリング電磁石の垂直方向パラメータと Q_{ratio} の関係を、図 4.14 に選別した 4 Sector 4 個目にあるステアリング電磁石の垂直方向パラメータと Q_{ratio} の関係を示す。それぞれ縦軸がステアリング電磁石パラメータ、横軸が Q_{ratio} となっている。また図中のオ

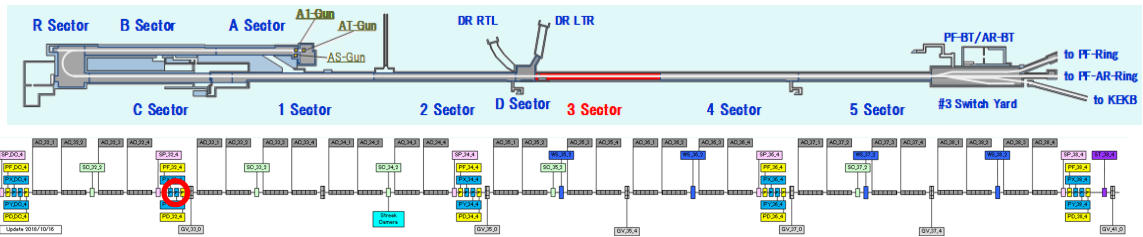


図4.11 Linac 加速器 3 Sector の配置図

上図の赤色の部分が 3 Sector の位置であり、下図が 3 Sector の拡大図となっている。また下図の左方が Linac 加速器の上流側、右方が下流側である。

下図の赤色の円で囲まれたステアリング電磁石が使用したステアリング電磁石 (PY_32_4) である。

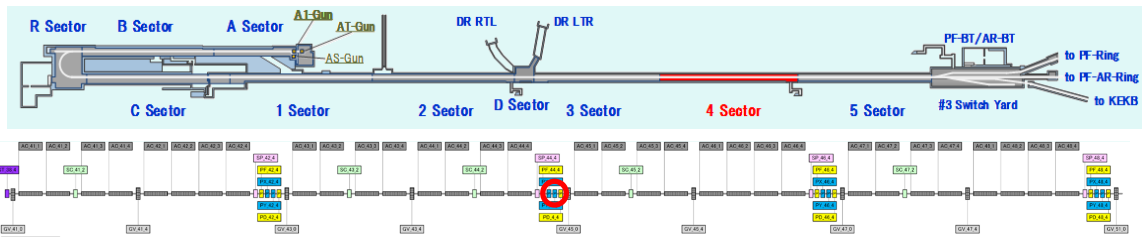


図4.12 Linac 加速器 4 Sector の配置図

上図の赤色の部分が 4 Sector の位置であり、下図が 4 Sector の拡大図となっている。また下図の左方が Linac 加速器の上流側、右方が下流側である。

下図の赤色の円で囲まれたステアリング電磁石が使用したステアリング電磁石 (PY_44_4) である。

レンジ色の直線が $Q_{ratio} = 0.9$ を表している。

図からそれぞれのステアリング電磁石パラメータと Q_{ratio} には直接的な相関がないことがわかる。収集データの中には Q_{ratio} が高いデータから低いデータまで、つまり良いビームから悪いビームまで存在していることがわかる。本研究では Q_{ratio} が高いときのパラメータ予測を行うため、 Q_{ratio} が 0.9 以上のデータを抽出して DNN での学習を行う。

これより DNN で使用するパラメータは以下のようになる。

- 目的変数 1
 Q_{ratio} が 0.9 以上のステアリング電磁石 (PY_32_4) パラメータ
- 目的変数 2
 Q_{ratio} が 0.9 以上のステアリング電磁石 (PY_44_4) パラメータ
- 説明変数 1-1
 ステアリング電磁石 (PY_32_4) パラメータを除いた全電磁石パラメータ
 : 合計 809 パラメータ
- 説明変数 1-2
 ステアリング電磁石 (PY_32_4) パラメータを除いた全電磁石パラメータ、全環境パラメータ
 : 合計 1541 パラメータ
- 説明変数 1-3
 ステアリング電磁石 (PY_32_4) パラメータを除いた全電磁石パラメータ、全 RF パラメータ、全環境パラメータ
 : 合計 1881 パラメータ
- 説明変数 2-1

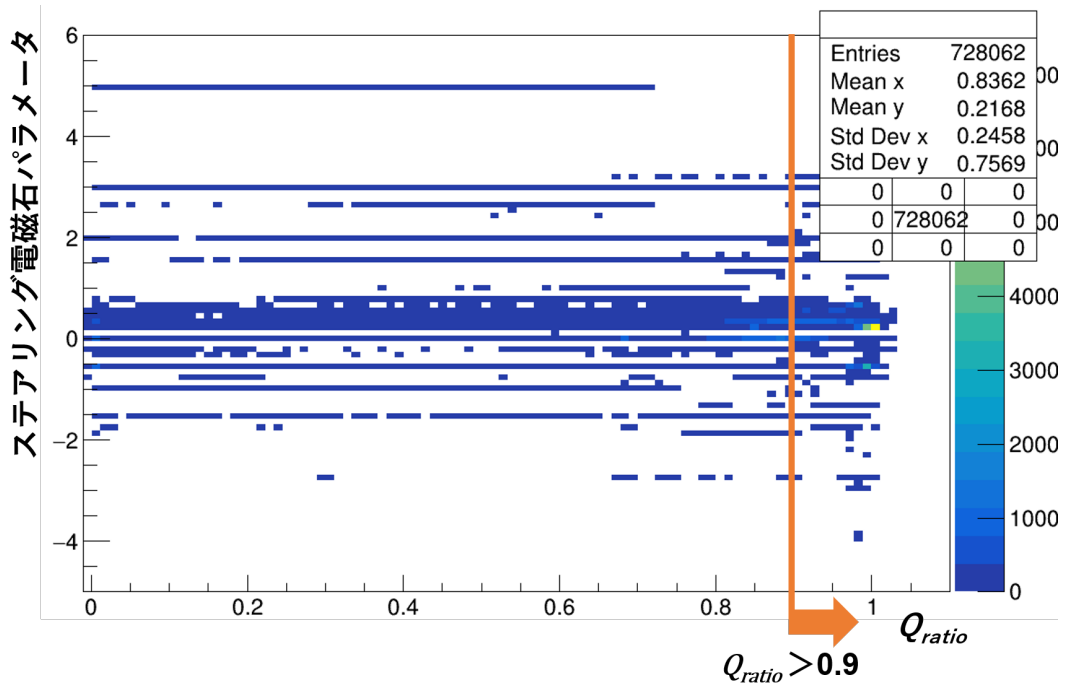


図4.13 ステアリング電磁石 (PY_32_4) パラメータと Q_{ratio} の関係

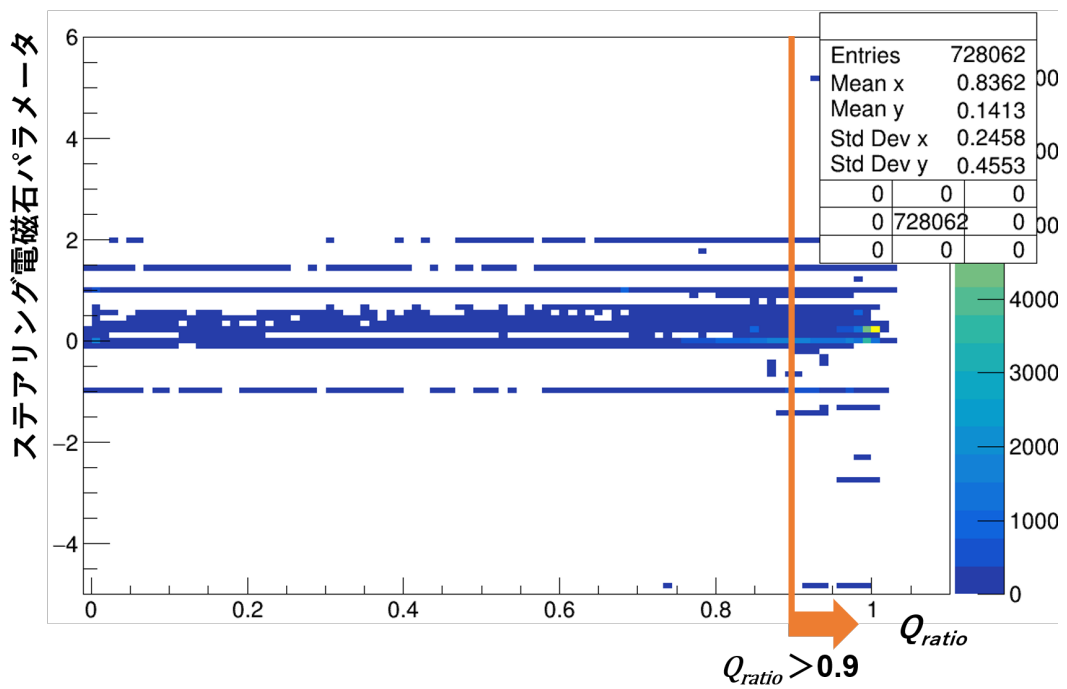


図4.14 ステアリング電磁石 (PY_44_4) パラメータと Q_{ratio} の関係

ステアリング電磁石 (PY_44_4) パラメータを除いた全電磁石パラメータ

: 合計 810 パラメータ

- 説明変数 2-2

ステアリング電磁石 (PY_44_4) パラメータを除いた全電磁石パラメータ、全環境パラメータ

: 合計 1542 パラメータ

- 説明変数 2-3

ステアリング電磁石 (PY_44_4) パラメータを除いた全電磁石パラメータ、全 RF パラメータ、全環境パラメータ

: 合計 1882 パラメータ

目的変数は今回予測をしたいパラメータ、説明変数は目的変数を説明できる可能性があるパラメータ群である。目的変数はステアリング電磁石 2 種類分、説明変数はそれぞれのステアリング電磁石に対して 3 種類用意した。目的変数の 1 種類目が 3 Sector2 個目にあるステアリング電磁石の垂直方向パラメータ (PY_32_4)、2 種類目が 4 Sector4 個目にあるステアリング電磁石の垂直方向パラメータ (PY_44_4) である。それぞれの説明変数の 1 種類目は全ステアリング電磁石パラメータのうち、今回目的変数として使用したそれぞれのステアリング電磁石パラメータを除いたもの。2 種類目は 1 種類目のパラメータに、全環境パラメータを加えたもの。3 種類目は 2 種類目のパラメータに、全 RF パラメータを加えたもの。このそれぞれの説明変数から目的変数を予測する DNN を設計し、比較を行う。

また抽出データは RF 位相パラメータと同様にどちらも全体で 396,538 データである。そこからデータをランダムに 8:1:1(データ数 317,230:39,654:39,654) に分割し、学習用、検証用、テスト用とした。以下で扱う結果の値はテストデータを使用した結果となっている。

4.5.2.2 回帰学習を用いた予測

DNN の回帰学習を用いて、ステアリング電磁石パラメータの予測を行った。表 4.8、表 4.9 にそれぞれの説明変数について、最適化を行ったハイパーパラメータを示す。

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1-1	1000	0.0015	100	4	0.015
説明変数 1-2	1000	0.0012	100	5	0.074
説明変数 1-3	1000	0.001	100	6	0.0036

表4.8 データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 2-1	6000	0.002	100	3	0.0063
説明変数 2-2	1000	0.0018	100	8	0.0048
説明変数 2-3	7000	0.002	100	7	0.12

表4.9 データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ

DNN の回帰学習の結果を以下に示す。図 4.15 と図 4.16 にそれぞれの目的変数の場合でそれぞれの説明変数の場合の、ステアリング電磁石パラメータの正解値と予測値の関係を示す。それぞれの図につい

て、横軸がステアリング電磁石パラメータの正解値、縦軸がそれぞれのステアリング電磁石パラメータでのネットワークが出力した予測値となっている。また左図が説明変数 1、中図が説明変数 2、右図が説明変数 3 を使用した場合の結果となっている。表 4.10と表 4.11にそれぞれの目的変数の場合でそれぞれの説明変数の際の回帰学習の結果をまとめる。loss とは予測値と正解値の乖離の大きさを表している。loss の算出には平均二乗誤差 (mse) を使用している。

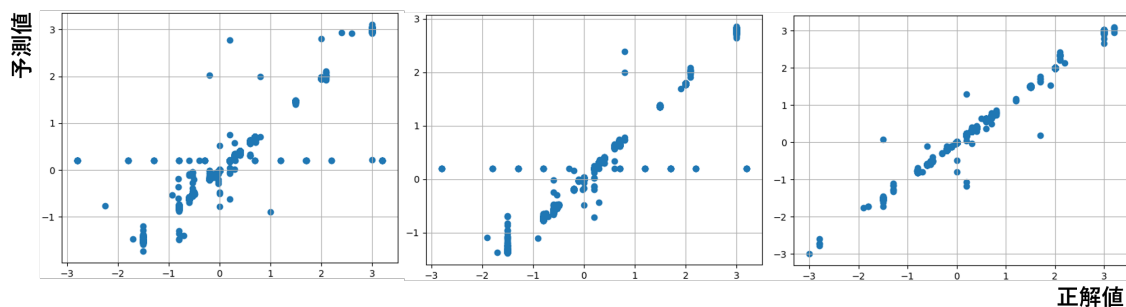


図4.15 データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ
 予測したパラメータと正解のパラメータの関係
 左図: 説明変数 1-1 の結果、中図: 説明変数 1-2 の結果、右図: 説明変数 1-3 の結果

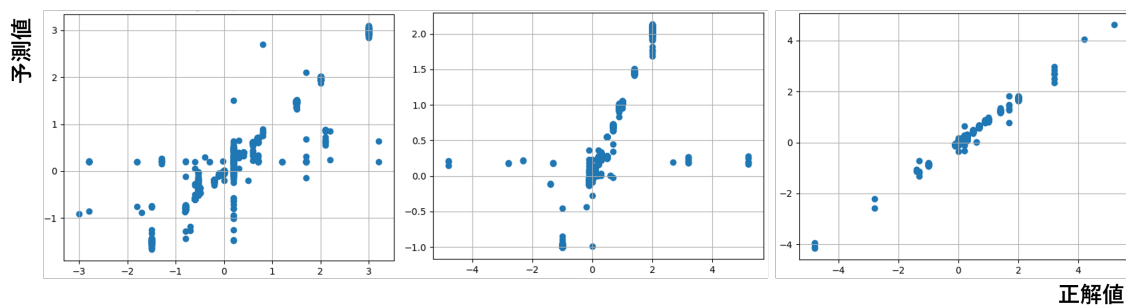


図4.16 データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの回帰予測の結果まとめ
 予測したパラメータと正解のパラメータの関係
 左図: 説明変数 2-1 の結果、中図: 説明変数 2-2 の結果、右図: 説明変数 2-3 の結果

	loss
説明変数 1-1	0.00552
説明変数 1-2	0.00414
説明変数 1-3	0.000281

表4.10 データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ

	loss
説明変数 2-1	0.00585
説明変数 2-2	0.00851
説明変数 2-3	0.00275

表4.11 データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの回帰予測の結果まとめ

説明変数 1、説明変数 2 と説明変数 3 の全ての場合で学習が成功している。目的変数がどちらの場合でも説明変数 3 の場合の学習が最も良い結果となった。また説明変数 1 と説明変数 2 の結果は、どちらの目的変数の場合でも学習の成功度に大差はなかった。

また RF 位相パラメータの場合と同様にステアリング電磁石の値を分割し、それぞれのパラメータの

データ数を整え、ネットワークに入力する方法を考えた。次節でその方法として分類学習を用いた場合の方法と結果を示す。

4.5.2.3 分類学習を用いた予測

図 4.17と図 4.18にそれぞれのステアリング電磁石パラメータの値を分割した際の Q_{ratio} とステアリング電磁石パラメータの関係を示す。図 4.13と図 4.14 と同様に、縦軸がステアリング電磁石パラメータ、横軸が Q_{ratio} となっており、灰色の直線が $Q_{ratio} = 0.9$ を表している。オレンジ色の直線は今回ステアリング電磁石パラメータの値を分割した境界となっており、赤色の数字は赤色の吹き出しの部分に設定したパラメータとなっている。

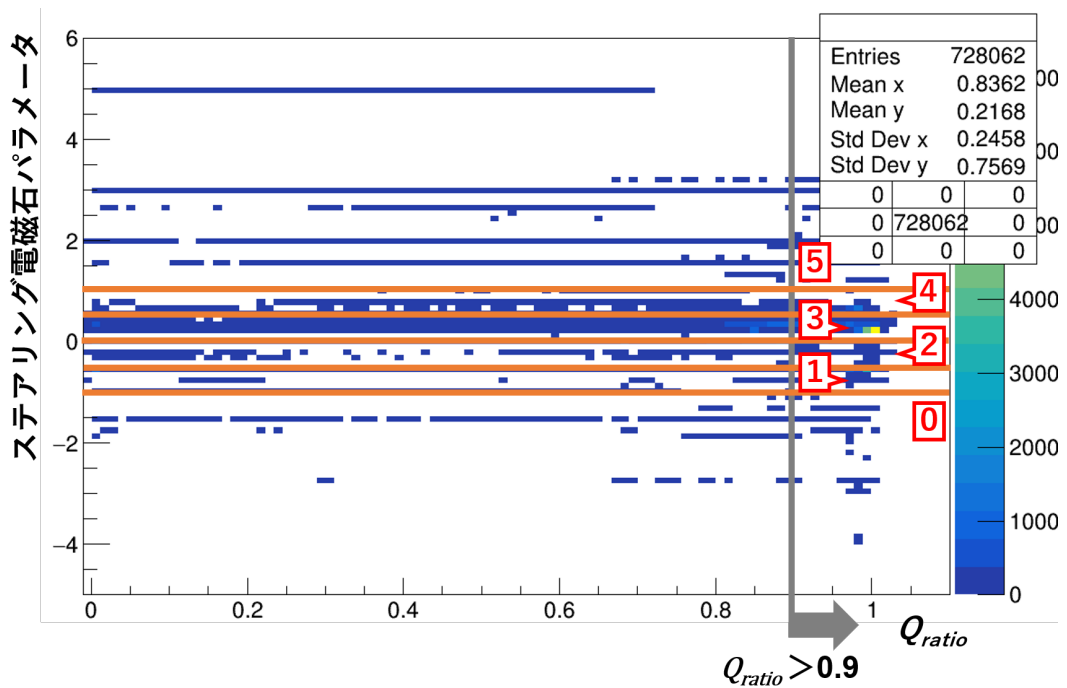


図4.17 パラメータを分類した際のステアリング電磁石 (PY_32_4) パラメータと Q_{ratio} の関係

以下に分割したそれぞれのパラメータのパラメータの範囲とデータ数を示す。

3 Sector2 個目にあるステアリング電磁石の垂直方向パラメータについて

- パラメータ 0 ステアリング電磁石パラメータが、-0.6 以下の値：データ数 6711
- パラメータ 1 ステアリング電磁石パラメータが、-0.6 以上 -0.4 以下の値：データ数 59933
- パラメータ 2 ステアリング電磁石パラメータが、-0.4 以上 0 以下の値：データ数 31060
- パラメータ 3 ステアリング電磁石パラメータが、0 以上 0.2 以下の値：データ数 24266
- パラメータ 4 ステアリング電磁石パラメータが、0.2 以上 0.4 以下の値：データ数 63492
- パラメータ 5 ステアリング電磁石パラメータが、0.4 以上 0.6 以下の値：データ数 31677
- パラメータ 6 ステアリング電磁石パラメータが、0.6 以上の値：データ数 32861

4 Sector4 個目にあるステアリング電磁石の垂直方向パラメータについて

- パラメータ 0 ステアリング電磁石パラメータが、-1 以下の値：データ数 20367

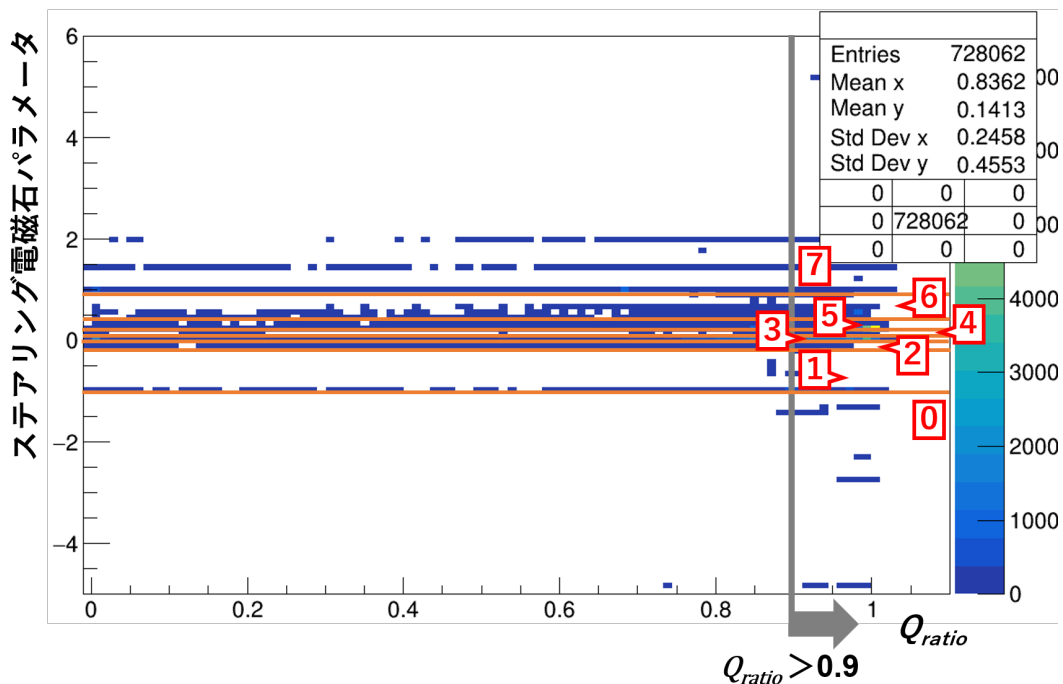


図4.18 パラメータを分類した際のステアリング電磁石 (PY_44_4) パラメータと Q_{ratio} の関係

- パラメータ 1 ステアリング電磁石パラメータが、-1 以上 -0.2 以下の値：データ数 5314
- パラメータ 2 ステアリング電磁石パラメータが、-0.2 以上 0 以下の値：データ数 50520
- パラメータ 3 ステアリング電磁石パラメータが、0 以上 0.1 以下の値：データ数 70000
- パラメータ 4 ステアリング電磁石パラメータが、0.1 以上 0.2 以下の値：データ数 73604
- パラメータ 5 ステアリング電磁石パラメータが、0.2 以上 0.4 以下の値：データ数 34545
- パラメータ 6 ステアリング電磁石パラメータが、0.4 以上 0.8 以下の値：データ数 18206
- パラメータ 7 ステアリング電磁石パラメータが、0.8 以上：データ数 27444

ここでデータ数はそれぞれのパラメータで異なりすぎないように調整している。またデータ数の合計はステアリング電磁石 (PY_32_4) パラメータ (目的変数 1) で 250,000 データ、ステアリング電磁石 (PY_44_4) パラメータ (目的変数 2) で 300,000 である。回帰学習と同様にこの場合そこからデータをランダムに 8:1:1(データ数 目的変数 1 の場合 200,000:25,000:25,000 目的変数 2 の場合 240,000:30,000:30,000) に分割し、学習用、検証用、テスト用とした。以下で扱う結果の値はテストデータを使用した結果となっている。また説明変数 1,2,3 と目的変数は 4.5.2.1 節と同様である。

また表 4.12 と表 4.13 にそれぞれの説明変数について、最適化を行ったハイパーパラメータを示す。

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1-1	8000	0.0033	100	3	0.13
説明変数 1-2	1000	0.0016	100	2	0.21
説明変数 1-3	1000	0.0012	100	4	0.23

表4.12 データ 1 を使用したステアリング電磁石 (PY_32_4) パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ

DNN の分類学習の結果を以下に示す。表 4.14, 表 4.15, 表 4.16 にそれぞれの説明変数の場合の、目的

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 2-1	1000	0.001	100	7	0.021
説明変数 2-2	1000	0.0012	100	7	0.19
説明変数 2-3	8000	0.001	450	7	0.22

表4.13 データ 1 を使用したステアリング電磁石 (PY_44_4) パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ

変数 1 のステアリング電磁石パラメータの正解値と予測値の関係を、表 4.17, 表 4.18, 表 4.19にそれぞれの説明変数の場合の、目的変数 2 のステアリング電磁石パラメータの正解値と予測値の関係を示す。それぞれの表について、各列がステアリング電磁石パラメータの正解値、各行がそれぞれのステアリング電磁石パラメータでのネットワークが出力した予測値となっており、それぞれのセルがステアリング電磁石パラメータの正解値に対してネットワークが出力した予測値のデータ数である。つまり左下から右上への対角のセルが正しく予測できたデータ数となっている。また表 4.14と表 4.17が説明変数 1、表 4.15と表 4.18が説明変数 2、表 4.16と表 4.19が説明変数 3 を使用した場合の結果となっている。表 4.20と表 4.21にそれぞれの説明変数の際の回帰学習の結果をまとめる。ここで acc は正答率を表しており、loss は同様に予測値と正解値の乖離の大きさを表している。loss の算出には多クラス交差エントロピー (categorical cross entropy) を使用している。

		正解値						
		0	1	2	3	4	5	6
予測値	6	22	22	2	9	0	0	3226
	5	0	0	0	0	0	3174	7
	4	0	0	0	50	6280	0	0
	3	1	0	95	2349	70	0	8
	2	1	6	3041	60	0	0	0
	1	0	5964	3	2	0	0	0
	0	622	0	3	3	0	0	0

表4.14 データ 1 を使用した説明変数 1-1 の場合のステアリング電磁石 (PY_32_4) パラメータの分類予測の結果

		正解値						
		0	1	2	3	4	5	6
予測値	6	3	0	0	0	0	1	3305
	5	0	0	0	0	0	3206	1
	4	0	0	0	48	6424	0	0
	3	0	0	89	2258	14	0	0
	2	0	0	3011	52	0	0	0
	1	1	5912	5	0	0	0	0
	0	651	1	4	1	0	0	13

表4.15 データ 1 を使用した説明変数 1-2 の場合のステアリング電磁石 (PY_32_4) パラメータの分類予測の結果

		正解値						
		0	1	2	3	4	5	6
予測値	6	0	0	0	3	1	2	3234
	5	0	0	0	0	0	3241	5
	4	0	0	0	47	6324	0	0
	3	0	1	110	2314	8	0	0
	2	0	0	2939	26	0	0	0
	1	2	6035	0	1	0	0	0
	0	707	0	0	0	0	0	0

表4.16 データ 1 を使用した説明変数 1-3 の場合のステアリング電磁石 (PY_32_4) パラメータの分類予測の結果

		正解値							
		0	1	2	3	4	5	6	7
予測値	7	4	0	0	0	1	0	0	2761
	6	0	0	0	0	0	0	1812	0
	5	0	0	0	0	312	2600	0	0
	4	3	0	0	0	7110	890	1	4
	3	0	0	790	6102	0	6	0	0
	2	5	0	4198	834	0	0	6	0
	1	124	191	0	0	0	0	0	0
	0	1889	341	3	0	0	0	0	8

表4.17 データ 1 を使用した説明変数 2-1 の場合のステアリング電磁石 (PY_44_4) パラメータの分類予測の結果

これより DNN の分類学習を使用すると説明変数 1、説明変数 2 と説明変数 3 の全ての場合で学習が正しくできることがわかった。また正答率も目的変数 1 では全ての場合で 9.9 割に近い結果であり、目的変数 2 では全ての場合で 8.8 割に近い結果となった。また回帰学習では説明変数 3 の場合が最も良い結果が得られたが、分類学習では異なる結果となった。また表 4.14、表 4.15、表 4.16、そして表 4.17、表 4.18、表 4.19 からわかる通り、RF 位相パラメータと同様に不正解のデータも正解値のセルの隣のセルに多く集まっていることがわかる。ここで各パラメータは自身で設定したものであり、本来のパラメータは連続値になっているので、隣のパラメータと不正解したとしても、完全に学習が誤りとは考えにくい。これより結果として既に高い正答率になっているが、この「不正解ではあるが、正解に近いデータ」を加味することで、DNN での学習は正しく行われていたと考えられる。これより DNN を使用したステアリング電磁石パラメータの予測は正しく行えたと考えられる。

4.6 データ 1 で学習したネットワークによるデータ 2 の評価

加速器の運用では 1 年に 2 回、夏と冬に長期の shutdown が存在しており、その shutdown の後に改めてビームを出し、調整をやり直している。加速器調整システムの構築を行う上では、shutdown 前のデータで作成したネットワークが shutdown 後に使用できるか重要である。

そこで本節では 4.5 節でデータ 1 から作成したネットワークを使用し、データ 2 のデータでの評価を

		正解値							
		0	1	2	3	4	5	6	7
予測値	7	0	0	0	0	0	0	0	2733
	6	0	0	0	0	0	0	1762	0
	5	0	0	0	0	357	2590	0	0
	4	2	0	0	0	7003	846	1	4
	3	0	0	905	6519	0	1	0	0
	2	3	0	4127	614	0	0	46	0
	1	26	46	0	0	0	0	0	0
	0	1940	467	2	0	3	0	2	1

表4.18 データ1を使用した説明変数2-2の場合のステアリング電磁石 (PY_44_4) パラメータの分類予測の結果

		正解値							
		0	1	2	3	4	5	6	7
予測値	7	0	0	0	0	1	0	0	2700
	6	0	0	1	0	0	0	1792	0
	5	0	0	0	0	381	2645	0	0
	4	0	0	0	0	6862	856	0	1
	3	0	0	989	6418	0	2	0	0
	2	2	0	4169	643	0	2	1	0
	1	206	226	0	0	0	0	0	0
	0	1828	272	1	0	1	0	1	0

表4.19 データ1を使用した説明変数2-3の場合のステアリング電磁石 (PY_44_4) パラメータの分類予測の結果

	acc	loss
説明変数 1-1	0.988	0.0413
説明変数 1-2	0.991	0.0328
説明変数 1-3	0.991	0.0351

表4.20 データ1を使用したステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ

	acc	loss
説明変数 2-1	0.886	0.231
説明変数 2-2	0.890	0.228
説明変数 2-3	0.875	0.297

表4.21 データ1を使用したステアリング電磁石 (PY_44_4) パラメータの分類予測の結果まとめ

行う。

4.6.1 RF パラメータの予測

B Sector5 個目にあるクライストロン (KL_B5) が出力する RF 位相パラメータについて、4.5.1.2節と 4.5.1.3節で作成したネットワークを使用し、データ2での検証を行った。表4.22に回帰学習の結果を、表4.23に分類学習の結果を示す。

また以下に分類学習で使ったパラメータと各パラメータのデータ数を示す。各パラメータの範囲に

については 4.5.1.3節と同様であるため、パラメータによってデータ数の偏りがある。

パラメータ 0 RF 位相パラメータが、0 以下の値：データ数 41104
 パラメータ 1 RF 位相パラメータが、0 以上 0.1 以下の値：データ数 5674
 パラメータ 2 RF 位相パラメータが、0.1 以上 0.2 以下の値：データ数 642
 パラメータ 3 RF 位相パラメータが、0.2 以上 0.3 以下の値：データ数 308
 パラメータ 4 RF 位相パラメータが、0.3 以上 0.4 以下の値：データ数 262
 パラメータ 5 RF 位相パラメータが、0.4 以上 0.5 以下の値：データ数 722
 パラメータ 6 RF 位相パラメータが、0.5 以上の値：データ数 1336

	loss
説明変数 1	0.385
説明変数 2	0.0178
説明変数 3	0.0265

表4.22 データ 1 を使用したネットワークでのデータ 2 の評価についての RF 位相パラメータの回帰予測の結果まとめ

	acc	loss
説明変数 1	0.369	8.79
説明変数 2	0.393	3.94
説明変数 3	0.250	12.0

表4.23 データ 1 を使用したネットワークでのデータ 2 の評価についての RF 位相パラメータの分類予測の結果まとめ

4.5.1節の結果と比較すると、回帰予測と分類予測のどちらの結果でも全ての説明変数で学習が成功していないことがわかる。これより RF 位相パラメータの予測は、ネットワークの作成と別時期のデータを使用すると、DNN での予測ができないことがわかった。

4.6.2 ステアリング電磁石パラメータの予測

3 Sector2 個目にあるステアリング電磁石の垂直方向パラメータ (PY_32_4) と 4 Sector4 個目にあるステアリング電磁石の垂直方向パラメータ (PY_44_4) について、4.5.2.2節と 4.5.2.3節で作成したネットワークを使用し、データ 2 での検証を行った。表 4.24と表 4.25に回帰学習の結果を、表 4.26と表 4.27に分類学習の結果を示す。

また以下に分類学習で使ったパラメータと各パラメータのデータ数を示す。各パラメータの範囲については 4.5.2.3節と同様であるため、パラメータによってデータ数の偏りがあったり、そもそもパラメータにデータが無いものも存在する。

ステアリング電磁石 (PY_32_4) パラメータについて

パラメータ 0 ステアリング電磁石パラメータが、-0.6 以下の値：データ数 34488
 パラメータ 1 ステアリング電磁石パラメータが、-0.6 以上 -0.4 以下の値：データ数 14822
 パラメータ 2 ステアリング電磁石パラメータが、-0.4 以上 0 以下の値：データ数 738

- パラメータ 3 ステアリング電磁石パラメータが、0 以上 0.2 以下の値：データ数 0
- パラメータ 4 ステアリング電磁石パラメータが、0.2 以上 0.4 以下の値：データ数 0
- パラメータ 5 ステアリング電磁石パラメータが、0.4 以上 0.6 以下の値：データ数 0
- パラメータ 6 ステアリング電磁石パラメータが、0.6 以上の値：データ数 0

ステアリング電磁石 (PY_44_4) パラメータについて

- パラメータ 0 ステアリング電磁石パラメータが、-1 以下の値：データ数 0
- パラメータ 1 ステアリング電磁石パラメータが、-1 以上 -0.2 以下の値：データ数 0
- パラメータ 2 ステアリング電磁石パラメータが、-0.2 以上 0 以下の値：データ数 0
- パラメータ 3 ステアリング電磁石パラメータが、0 以上 0.1 以下の値：データ数 0
- パラメータ 4 ステアリング電磁石パラメータが、0.1 以上 0.2 以下の値：データ数 0
- パラメータ 5 ステアリング電磁石パラメータが、0.2 以上 0.4 以下の値：データ数 0
- パラメータ 6 ステアリング電磁石パラメータが、0.4 以上 0.8 以下の値：データ数 29460
- パラメータ 7 ステアリング電磁石パラメータが、0.8 以上：データ数 20588

	loss
説明変数 1-1	41.5
説明変数 1-2	1.27
説明変数 1-3	278

表4.24 データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ

	loss
説明変数 2-1	2.32
説明変数 2-2	0.924
説明変数 2-3	42.4

表4.25 データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_44_4) パラメータの回帰予測の結果まとめ

	acc	loss
説明変数 1-1	0.0696	14.7
説明変数 1-2	0.0999	9.55
説明変数 1-3	0.000286	16.1

表4.26 データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ

	acc	loss
説明変数 2-1	0.0443	15.2
説明変数 2-2	0.0909	7.34
説明変数 2-3	0.00	16.1

表4.27 データ 1 を使用したネットワークでのデータ 2 の評価についてのステアリング電磁石 (PY_44_4) パラメータの分類予測の結果まとめ

4.5.2節の結果と比較すると、回帰予測と分類予測のどちらの結果でも全ての目的変数、説明変数で学習が成功していないことがわかる。これよりステアリング電磁石パラメータの予測は、ネットワークの作成と別時期のデータを使用すると、DNN での予測ができないことがわかった。

4.6.3 考察

RF 位相パラメータとステアリング電磁石パラメータの両方について、データ 1 を使用したネットワークでのデータ 2 の評価を行うことができない理由について考察する。ここで図 4.19と図 4.20に、RF パラメータと電磁石パラメータについてのそれぞれでデータ 1 とデータ 2 のパラメータの分布を比較し

たヒストグラムを示す。

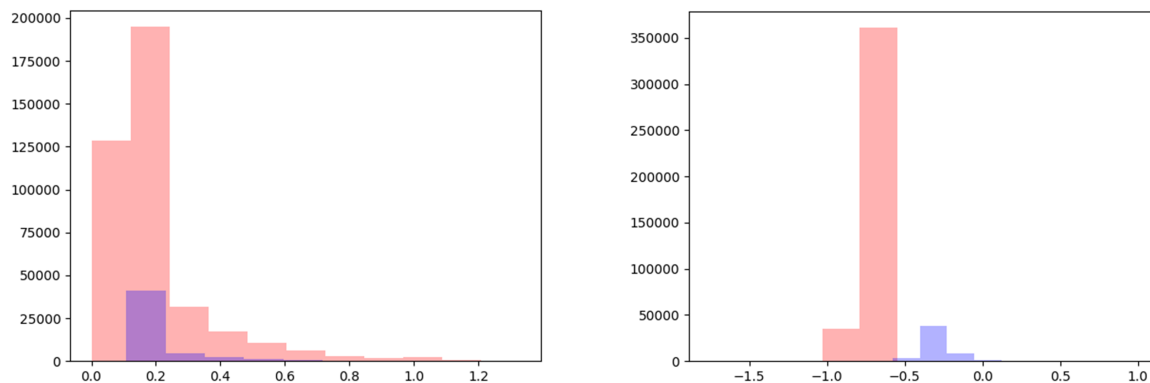


図4.19 データ 1 とデータ 2 での RF パラメータの分布比較

左図：3 Sector3 個目にあるクライストロン (KL_33) が出力する RF 位相パラメータについての比較

右図：4 Sector3 個目にあるクライストロン (KL_43) が出力する RF 振幅パラメータについての比較

赤色のヒストグラムがデータ 1 を、青色のヒストグラムがデータ 2 を示している。

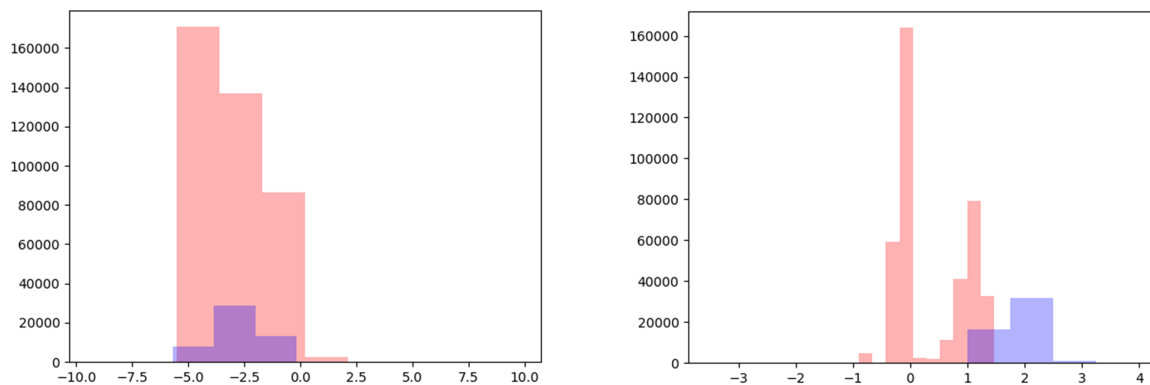


図4.20 データ 1 とデータ 2 での電磁石パラメータの分布比較

左図：5 Sector6 個目にあるステアリング電磁石 (PY_56_4) の垂直方向パラメータについての比較

右図：2 Sector4 個目にあるステアリング電磁石 (PX_24_4) の水平方向パラメータについての比較

赤色のヒストグラムがデータ 1 を、青色のヒストグラムがデータ 2 を示している。

それぞれの図について横軸は各パラメータの値である。赤色がデータ 1 のヒストグラムであり、青色がデータ 2 のヒストグラムである。RF パラメータと電磁石パラメータのそれぞれについて、左図がパラメータ分布の範囲がデータ収集時期に関係なく同じパラメータの例を表しており、右図がパラメータ分布の範囲がデータ収集時期によって異なるパラメータの例を表している。これからわかる通り、RF パラメータと電磁石パラメータの両方で shutdown の前後で、異なるパラメータ値を使用し、加速器調整の最適化を行う可能性があるということがわかった。

またここで DNN を使用した教師あり学習では、ネットワーク作成を行う際に学習用データとして使用したデータを基にして予測を行うので、学習用データにないパターンのデータからの予測は行うことが出来ない。これより今回のように、ネットワーク作成時と検証時でパラメータの範囲が異なってしまうと正しく学習ができないと考えられる。

4.7 データ 1-1 で学習したネットワークによるデータ 1-2 の評価

本節では 4.6 節での考察について、4.5 節で最適値の予測が可能だったデータ 1 を分割し検証を行う。データ 1 から以下のように新しいデータセットを作成する。

データ 1-1 2019 年 2 月 9 日から 2019 年 3 月 8 日まで 252,000 shot

データ 1-2 2019 年 3 月 9 日から 2019 年 6 月 30 日まで 928,500 shot

データ 1-1 とデータ 1-2 の間には shutdown が存在せず、連続したデータとなっている。

またここで図 4.21 に、RF パラメータと電磁石パラメータについてのそれぞれでデータ 1-1 とデータ 1-2 のパラメータの分布を比較したヒストグラムを示す。

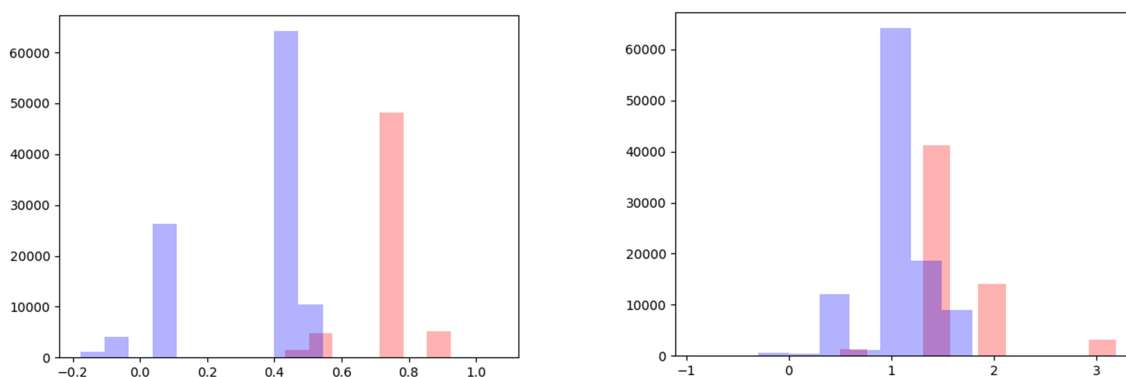


図4.21 データ 1-1 とデータ 1-2 での RF パラメータと電磁石パラメータの分布比較

左図：3 Sector4 個目にあるクライストロン (KL_34) が出力する RF 振幅パラメータについての比較

右図：3 Sector4 個目にあるステアリング電磁石 (PX_34_4) の水平方向パラメータについての比較

赤色のヒストグラムがデータ 1-1 を、青色のヒストグラムがデータ 1-2 を示している。

それぞれの図について横軸は各パラメータの値である。赤色がデータ 1-1 のヒストグラムであり、青色がデータ 1-2 のヒストグラムである。shutdown の有無にかかわらず RF パラメータと電磁石パラメータのそれぞれについて、パラメータ分布の範囲がデータ収集時期によって異なるものが存在することがわかる。

shutdown がないにもかかわらずパラメータ分布が異なるデータを使用することで、4.6 節の考察の検証を行う。本節ではデータ 1 の前半に当たるデータ 1-1 で DNN を学習させ、データ 1 の後半に当たるデータ 1-2 での評価を行う。

4.7.1 RF パラメータの予測

B Sector5 個目にあるクライストロン (KL_B5) が出力する RF 位相パラメータについて、データ 1-1 で作成したネットワークを使用し、データ 1-2 での検証を行った。

以下にデータ 1-1 の分類学習で使用したパラメータと各パラメータのデータ数を示す。

パラメータ 0 RF 位相パラメータが、0 以下の値：データ数 2068

パラメータ 1 RF 位相パラメータが、0 以上 0.1 以下の値：データ数 626

パラメータ 2 RF 位相パラメータが、0.1 以上 0.2 以下の値：データ数 272

パラメータ 3 RF 位相パラメータが、0.2 以上 0.3 以下の値：データ数 862
 パラメータ 4 RF 位相パラメータが、0.3 以上 0.4 以下の値：データ数 509
 パラメータ 5 RF 位相パラメータが、0.4 以上 0.5 以下の値：データ数 311
 パラメータ 6 RF 位相パラメータが、0.5 以上の値：データ数 352

ここでデータ数はそれぞれのパラメータで異なりすぎないように調整している。
 また以下にデータ 1-2 の分類学習で使用したパラメータと各パラメータのデータ数を示す。

パラメータ 0 RF 位相パラメータが、0 以下の値：データ数 92957
 パラメータ 1 RF 位相パラメータが、0 以上 0.1 以下の値：データ数 2682
 パラメータ 2 RF 位相パラメータが、0.1 以上 0.2 以下の値：データ数 2658
 パラメータ 3 RF 位相パラメータが、0.2 以上 0.3 以下の値：データ数 3079
 パラメータ 4 RF 位相パラメータが、0.3 以上 0.4 以下の値：データ数 858
 パラメータ 5 RF 位相パラメータが、0.4 以上 0.5 以下の値：データ数 479
 パラメータ 6 RF 位相パラメータが、0.5 以上の値：データ数 3355

またここでの説明変数 1,2,3 と目的変数は 4.5.1.1 節と同様である。

また表 4.28 と表 4.29 に回帰学習と分類学習についてそれぞれの説明変数について、最適化を行ったハイパーパラメータを示す。

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1	1000	0.0091	700	3	0.085
説明変数 2	1000	0.0015	100	5	0.17
説明変数 3	1000	0.0022	100	5	0.18

表4.28 データ 1-1 を使用した RF 位相パラメータの回帰予測で使用したネットワークのハイパーパラメータまとめ

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1	1000	0.0042	100	4	0.16
説明変数 2	1000	0.0070	100	4	0.20
説明変数 3	1000	0.0011	100	5	0.14

表4.29 データ 1-1 を使用した RF 位相パラメータの分類予測で使用したネットワークのハイパーパラメータまとめ

DNN での回帰学習と分類学習の結果を以下に示す。ここで acc は正答率を表しており、loss は予測値と正解値の乖離の大きさを表している。loss の算出には回帰学習では mse、分類学習では categorical cross entropy を使用している。

表 4.30 と表 4.32 がデータ 1-1 を使用した回帰予測と分類予測の結果を表す。また表 4.31 と表 4.33 がデータ 1-1 を使用したネットワークでのデータ 1-2 の評価の回帰予測と分類予測の結果を表す。データ 1-1 を使用した結果と比較し、データ 1-1 を使用したネットワークでのデータ 1-2 の評価の全ての結果で予測が成功していないことがわかる。これより RF 位相パラメータの予測では、ネットワークの作成時と評価時の DNN に入力するパラメータの分布が異なれば、DNN での予測ができないことがわかった。

	loss
説明変数 1	0.0112
説明変数 2	0.00666
説明変数 3	0.00928

表4.30 データ 1-1 を使用した RF 位相パラメータの回帰予測の結果まとめ

	loss
説明変数 1	0.0871
説明変数 2	11288
説明変数 3	3276

表4.31 データ 1-1 を使用したネットワークでのデータ 1-2 の評価についての RF 位相パラメータの回帰予測の結果まとめ

	acc	loss
説明変数 1	0.704	2.48
説明変数 2	0.701	1.20
説明変数 3	0.675	2.05

表4.32 データ 1-1 を使用した RF 位相パラメータの分類予測の結果まとめ

	acc	loss
説明変数 1	0.574	6.55
説明変数 2	0.634	3.35
説明変数 3	0.233	11.0

表4.33 データ 1-1 を使用したネットワークでのデータ 1-2 の評価についての RF 位相パラメータの分類予測の結果まとめ

4.7.2 ステアリング電磁石パラメータの予測

3 Sector2 個目にあるステアリング電磁石の垂直方向パラメータ (PY_32_4) について、データ 1-1 で作成したネットワークを使用し、データ 1-2 での検証を行った。ここで 4 Sector4 個目にあるステアリング電磁石の垂直方向パラメータ (PY_44_4) については、データ 1-1 でのパラメータのばらつきが少なく予測困難だと判断し行わなかった。

以下にデータ 1-1 の分類学習で使用したパラメータと各パラメータのデータ数を示す。

パラメータ 0 ステアリング電磁石パラメータが、0.1 以下の値：データ数 15600
 パラメータ 1 ステアリング電磁石パラメータが、0.1 以上 0.5 以下の値：データ数 17742
 パラメータ 2 ステアリング電磁石パラメータが、0.5 以上 1 以下の値：データ数 21615
 パラメータ 3 ステアリング電磁石パラメータが、1 以上の値：データ数 4717

ここでデータ数はそれぞれのパラメータで異なりすぎないように調整している。

また以下にデータ 1-2 の分類学習で使用したパラメータと各パラメータのデータ数を示す。

パラメータ 0 ステアリング電磁石パラメータが、0.1 以下の値：データ数 30709
 パラメータ 1 ステアリング電磁石パラメータが、0.1 以上 0.5 以下の値：データ数 75241
 パラメータ 2 ステアリング電磁石パラメータが、0.5 以上 1 以下の値：データ数 14
 パラメータ 3 ステアリング電磁石パラメータが、1 以上の値：データ数 104

またここで説明変数 1-1,1-2,1-3 と目的変数 1 は 4.5.2.1 節と同様である。

また表 4.34 と表 4.35 に回帰学習と分類学習についてそれぞれの説明変数について、最適化を行ったハイパーパラメータを示す。

DNN での回帰学習と分類学習の結果を以下に示す。ここで acc は正答率を表しており、loss は予測値と正解値の乖離の大きさを表している。loss の算出には回帰学習では mse、分類学習では categorical

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1-1	1000	0.0094	100	5	0.20
説明変数 1-2	1000	0.0060	100	5	0.043
説明変数 1-3	1000	0.0010	100	4	0.14

表4.34 データ 1-1 を使用したステアリング電磁石パラメータの回帰予測で使したネットワークのハイパーパラメータまとめ

ハイパーパラメータ	バッチサイズ	学習率	中間ノード数	中間層数	ドロップアウト率
説明変数 1-1	1000	0.0042	100	4	0.16
説明変数 1-2	1000	0.0070	100	4	0.20
説明変数 1-3	1000	0.0013	100	3	0.078

表4.35 データ 1-1 を使用したステアリング電磁石パラメータの分類予測で使したネットワークのハイパーパラメータまとめ

cross entropy を使用している。

	loss
説明変数 1-1	0.0110
説明変数 1-2	0.000845
説明変数 1-3	0.00437

表4.36 データ 1-1 を使用したでのステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ

	loss
説明変数 2-1	0.0871
説明変数 2-2	11288
説明変数 2-3	3276

表4.37 データ 1-1 を使用したネットワークでのデータ 1-2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの回帰予測の結果まとめ

	acc	loss
説明変数 1-1	0.988	0.194
説明変数 1-2	0.999	0.00263
説明変数 1-3	0.974	0.0589

表4.38 データ 1-1 を使用したステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ

	acc	loss
説明変数 2-1	0.326	10.9
説明変数 2-2	0.335	10.72
説明変数 2-3	0.723	4.46

表4.39 データ 1-1 を使用したネットワークでのデータ 1-2 の評価についてのステアリング電磁石 (PY_32_4) パラメータの分類予測の結果まとめ

表 4.36と表 4.38がデータ 1-1 を使用した回帰予測と分類予測の結果を表す。また表 4.37と表 4.39がデータ 1-1 を使用したネットワークでのデータ 1-2 の評価の回帰予測と分類予測の結果を表す。データ 1-1 を使用した結果と比較し、データ 1-1 を使用したネットワークでのデータ 1-2 の評価の全ての結果で予測が成功していないことがわかる。これよりステアリング電磁石パラメータの予測は、ネットワークの作成時と評価時の DNN に入力するパラメータの分布が異なれば、DNN での予測ができないことがわかった。

第 5 章

まとめ

本研究では、高エネルギー加速器研究機構 (KEK) における電子・陽電子入射器 Linac の加速器運転データを用いて、機械学習を用いた加速器運転調整システムの開発を行った。収集した Linac 加速器運転データのうち、RF データと電磁石データ、環境データを用いたディープニューラルネットワークを使用し、RF 位相パラメータとステアリング電磁石パラメータの調整システム開発を行い、その性能評価を行った。RF 位相パラメータとステアリング電磁石パラメータについて、RF パラメータ、電磁石パラメータと環境パラメータを使用することで加速器運転調整最適パラメータを予測を行うことが出来た。RF 位相パラメータについて、回帰予測では RF パラメータを使用することで、分類予測については RF パラメータと環境パラメータを使用することで、最も精度の良い予測を行うことが出来た。またステアリング電磁石パラメータについて、回帰予測では電磁石パラメータと RF パラメータ、環境パラメータを使用することで、分類予測については電磁石パラメータと環境パラメータを使用することで、最も精度の良い予測を行うことが出来た。

また作成したネットワークを使用し、別の時期に収集した Linac 加速器運転データを予測できるかの検証を行った。この結果、ネットワークの学習時に使用したデータのパラメータ分布と評価時に使用したデータのパラメータ分布が異なっている場合、最適なパラメータ値を予測できないことが明らかになった。これより評価用データに対応した教師用データでネットワークの学習を行うことの重要性がわかった。

今後の課題としては、刻々と変化するパラメータの調整を行うために環境駆動型機械学習として強化学習の適応が有効だと考えられるため、強化学習の加速器運転調整への適応を検討する。

参考文献

- [1] 亀井亨・木原元央,「パリティ物理学コース 加速器科学」 丸善株式会社 (1993)
- [2] 熊谷寛夫,「実験物理学講座 28 加速器」,共立出版株式会社 (1975)
- [3] 神谷幸秀,「加速器の原理 -シンクロトロン及びストレージング-」,OHO '84 高エネルギー加速器セミナー (1984) II-1
- [4] 小川雄二郎,「電子線形加速器の概要」,OHO '02 高エネルギー加速器セミナー (2002) 1-1
- [5] 杉村高志,「KEKB 入射器入門」,OHO '04 高エネルギー加速器セミナー (2004) 5-1
- [6] R. Zhang et al. 「KEK 電子陽電子入射器の現状」 in Proceedings of the 16th Annual Meeting of Particle Accelerator Society of Japan, Kyoto, Japan, Jul.
- [7] M. Sato, “Injector beam operation” ,KEKB Accelerator Review Committee (2019)
- [8] I. Sato, ”Accelerator structure and beam transport system for the KEK photon factory injector” Nucl. Inst. Meth 177, 1 (1980), 91
- [9] T. Ieiri et. Al., ”Bunch lengthening recently observed at PF-AR”, Proceedings of the 2003 Particle Accelerator Conference, (2003), 3077
- [10] Z. Dolezal and S. Uno ”Belle II Technical Design Report” High Energy Accelerator Research Organization (2010)
- [11] EPICS <https://epics.anl.gov/>
- [12] 斎藤康毅,「ゼロから作る Deep Learning -Python で学ぶディープラーニングの理論と実装」,株式会社オライリー・ジャパン (2016)
- [13] TensorFlow <https://www.tensorflow.org/>
- [14] Keras <https://keras.io/>

謝辞

本論文を作成するにあたり、ご指導を頂いた宇宙・素粒子実験物理学研究室の岩崎昌子先生、また同研究室の中野英一先生に心より感謝を致します。直接の指導を頂いた岩崎先生には大変お世話になりました。素粒子物理から加速器技術に至るまでの幅広い知見からご指導頂き、またご自身が大変お忙しい中でも質問などに丁寧に答えて下さいました。また中野先生には理学部物理学科在学中に卒業研究の指導を頂いた頃から高エネルギー物理学の基礎をご指導頂き、現在の本研究に至るまで様々な助言を頂きました。心より感謝を申し上げます。

Linac 加速器でのデータ収集に当たり、高エネルギー加速器研究機構 (KEK) の佐藤政則様、佐武いつか様に大変お世話になりました。佐藤様には私が加速器技術に疎い中、素朴な質問にも答えて頂くなど様々な面で助けて頂いたことに感謝致します。

大阪大学データビリティフロンティア機構 (IDS) の長原一先生、中島悠太先生、武村紀子先生、大阪大学核物理研究センター (RCNP) の中野貴志先生には大変お世話になりました。また本研究は、大阪大学核物理研究センター (RCNP) 研究プロジェクト、大阪大学データビリティフロンティア機構学際プロジェクトから、多大なご協力を頂きました。

また本研究の一部は、文部科学省「ダイバーシティ研究環境実現イニシアティブ (牽引型)」連携型共同研究助成、および、大阪市立大学戦略的研究 (基盤研究) の助成を受けて実施されました。感謝致します。

宇宙・素粒子実験物理学研究室の皆様には、研究の実施に際して多くの協力を頂きました。特に共に機械学習関連の研究を行った同期の岸田直也君とは、議論し合い切磋琢磨しながら研究を進めることができました。そのおかげで楽しく充実した研究生活を送ることができたと思います。本当にありがとうございました。

最後になりましたが、私がこのような素晴らしい研究に携わり、物理学の面白さを知ることができたのは支えて下さった皆様のおかげです。改めて私の研究生活を支えて下さった全ての方に感謝を致します。