

機械学習を用いた 加速器制御の開発

2021/8/24

岩崎昌子 (阪市大/NITEP, 阪大RCNP/IDS)



大型加速器を用いた素粒子実験

- ▶ 大型実験施設（巨大加速器、測定器）を使用
→ 実験装置の制御点数は数千-数万点以上
- ▶ 膨大量の実験データ
→ BelleII実験では、年間数十ペタバイトの予定
- ▶ 膨大な背景事象から極微の信号事象を抽出

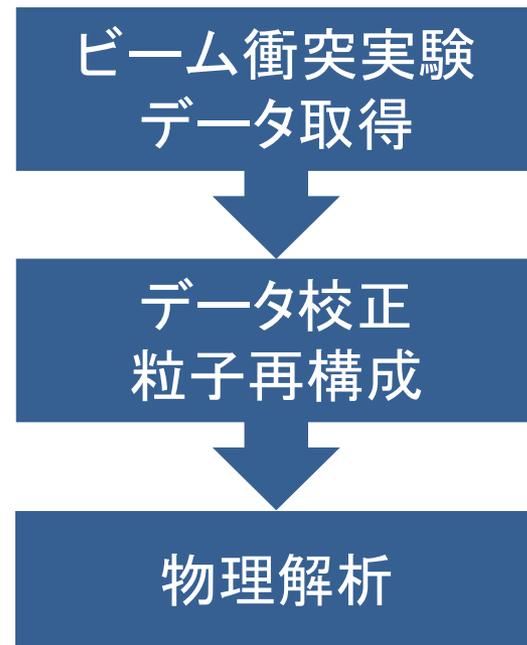
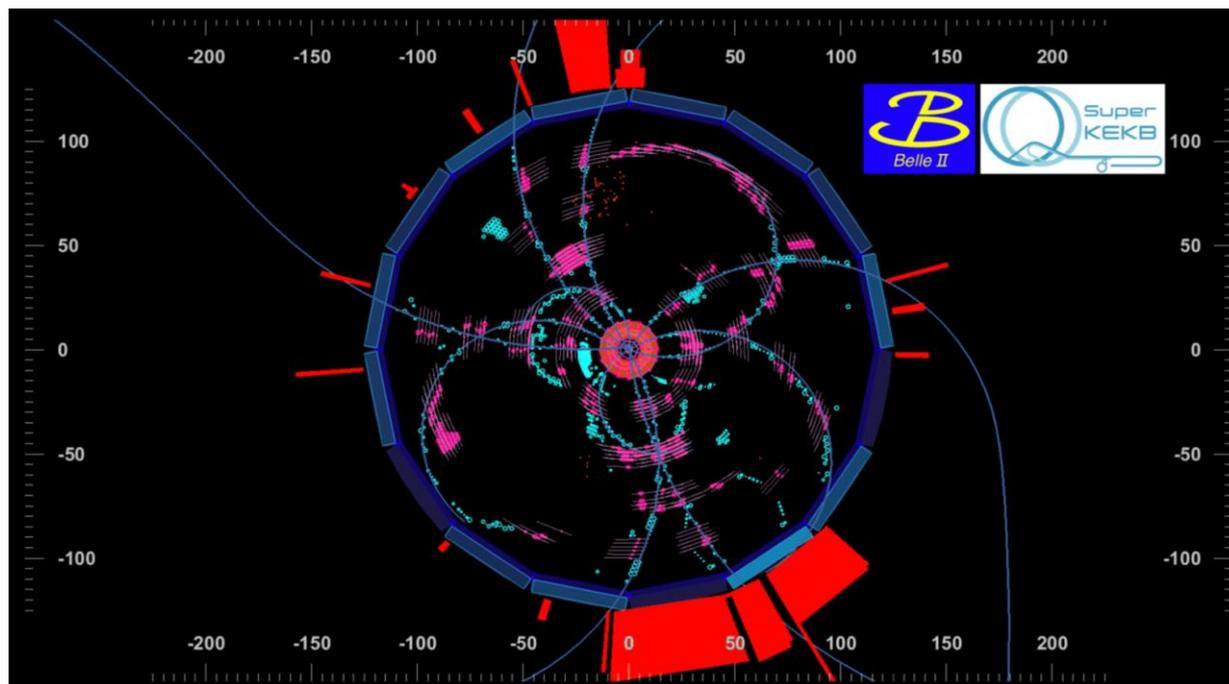
大量データを処理を
いかに、効率よく・精度よく行うか

大型加速器実験での データ処理

情報分野で開発された、**最先端機械学習技術**を
加速器実験でのデータ処理へ導入することで
従来よりも**高性能**、**高速**なデータ処理を期待

機械学習 → 加速器実験のビッグデータに対する
強力なデータ処理ツール

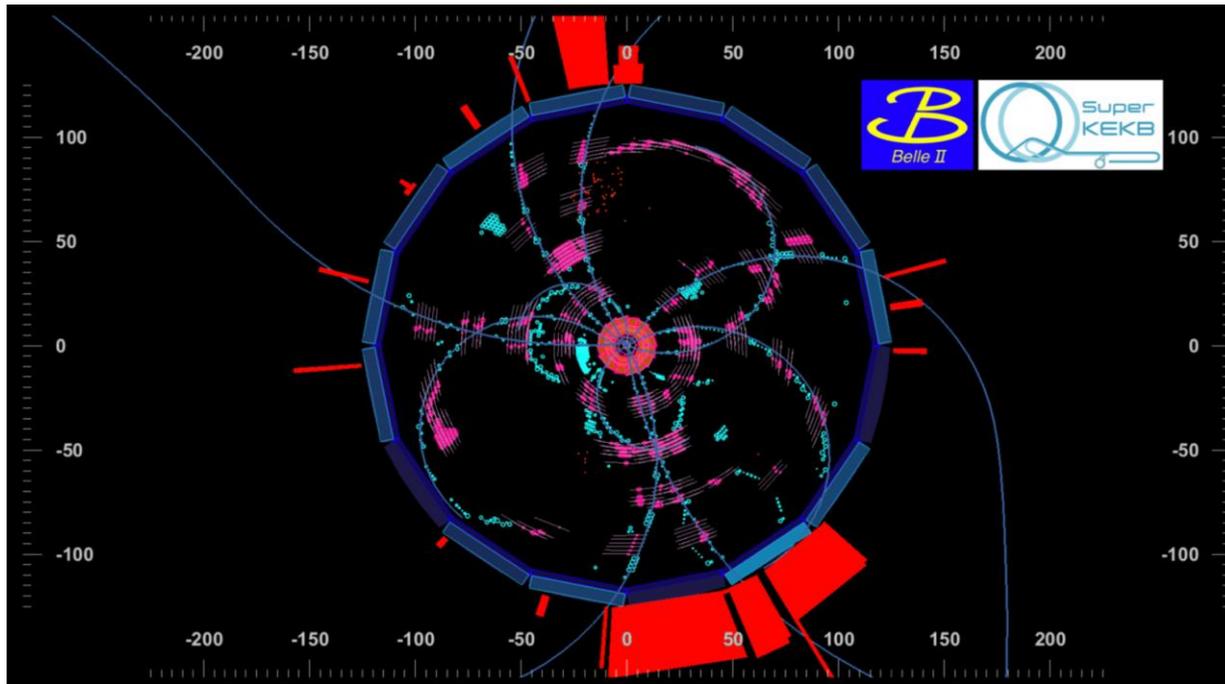
大型加速器実験 × 最先端情報学



大型加速器実験
大量データ 高額な運転経費

BelleII実験
生成データ量 年間~30ペタバイト
運転経費 年間数十億円

大型加速器実験 × 最先端機械学習の適用



機械学習の適用

ビーム衝突実験
データ取得

データ校正
粒子再構成

物理解析

大型加速器実験
大量データ 高額な運転経費

BelleII実験
生成データ量 年間~30ペタバイト
運転経費 年間数十億円

実験の効率化、大量データを効率的に学術的結果へ導くために、
最先端機械学習の適応が有効

例：10%の向上 → 年間数億円の費用効果

機械学習

Supervised Learning
教師あり学習

Task driven

Classification
分類・識別

Regression
回帰

Unsupervised Learning
教師なし学習

Data driven

Dimensionality Reduction
次元削減

Clustering

Reinforcement Learning
強化学習

Environment driven

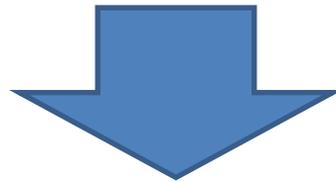
Algorithm learns to react to the environment

Real-time decisions
Game AI
Learning Tasks
Robot Navigator
....

いろいろなタイプの機械学習があります

加速器実験での機械学習

これまで加速器実験で行ってきた機械学習は
おもに特徴量(High-level data)を用いた
教師あり学習による分類・識別だった



計算機技術・機械学習技術の進展により
様々なタイプの機械学習が可能になった
低特徴データ(生データ、low-level data)を用いた
機械学習も可能になった

加速器実験への 最先端機械学習の適用

【キーワード】

- 低特徴データ (Low-level data, 生データ)
- 様々なタイプの機械学習
回帰、次元削減、強化学習、
再帰型ネットワーク、畳み込み型ネットワーク ...

“加速器実験および関連分野への 深層学習の適用と進化”

2018-23年度 阪大RCNP 研究プロジェクト/阪大IDS 学際プロジェクト
2021年度 JHPCN(学際大規模情報基盤共同利用・共同研究拠点) 萌芽型共同研究

原子核・素粒子物理研究者と情報科学研究者が
部局、大学、機関の枠を超えてグループを形成
加速器実験および関連分野へ深層学習を適用させ、
さらにその研究結果を発展させる

素粒子
原子核物理

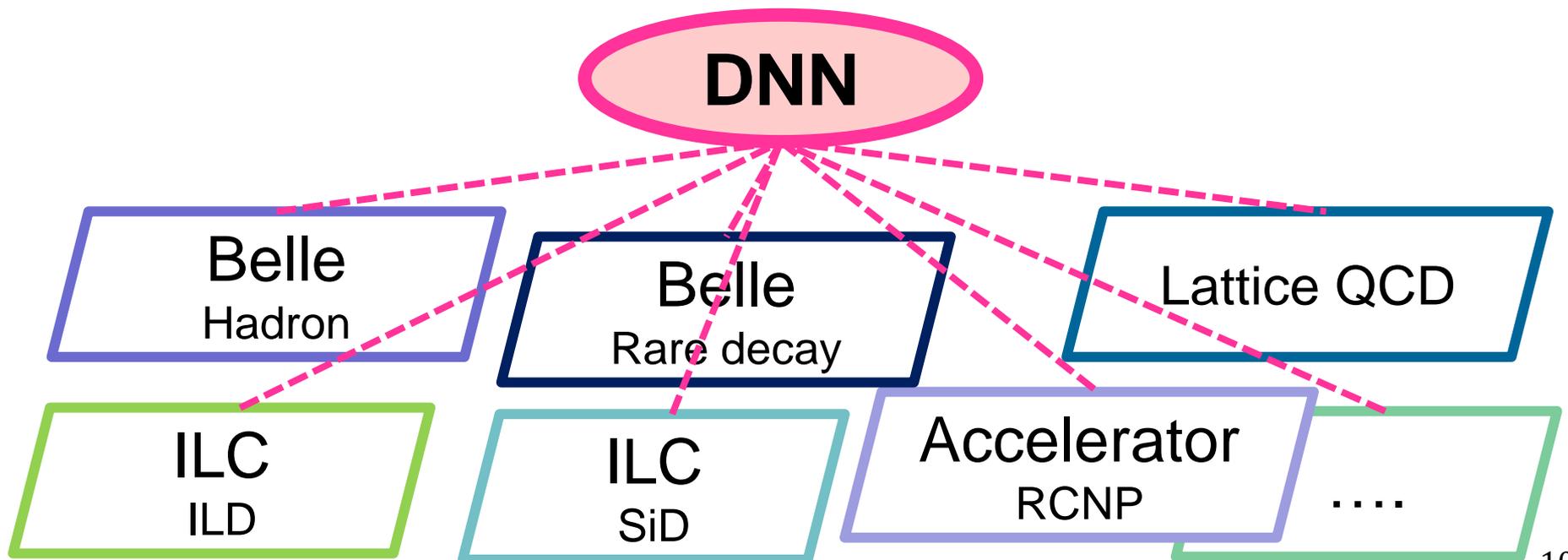


情報科学

“加速器実験および関連分野への 深層学習の適用と進化”

2018-23年度 阪大RCNP 研究プロジェクト/阪大IDS 学際プロジェクト
2021年度 JHPCN(学際大規模情報基盤共同利用・共同研究拠点) 萌芽型共同研究

プロジェクトではDNNの適用について議論 (ほぼ隔週でミーティング)



研究項目

- Belle・ILC実験でのフレーバー識別 (大阪市大, 阪大IDS・RCNP)
- ILC実験用電磁カロリメータエネルギー較正
(大阪市大, 阪大IDS, U. Oregon, SLAC)
- ILC実験でのジェットクラスタリング (九州大学 素粒子実験)
- ILC実験での崩壊点検出 (九州大学 素粒子実験)
- ILC実験でのジェットフレーバー識別 (九州大学 素粒子実験)
- KEK Linac加速器運転調整 (KEK, 大阪市大, 阪大IDS)
- ILC実験でのビームサイズ測定 (東北大学)
- RCNP イオン源制御 (阪大RCNP・IDS)
- 格子QCD計算への適用 (阪大RCNP・IDS)
- RIビーム粒子識別の自動最適化手法の開発
(九州大学 原子核実験, 東大CNS)

.....

主な研究項目

生データを用いた 機械学習手法開発

- 物理解析手法の開発
- 測定器較正
- Vertexing
- Clustering
- ...

加速器制御への 機械学習の導入

KEK Linac
RCNP
九州大学
...

その他

....

主な研究項目

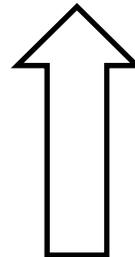
生データを用いた 機械学習手法開発

- 物理解析手法の開発
- 測定器較正
- Vertexing
- Clustering
- ...

加速器制御への 機械学習の導入

KEK Linac
RCNP
九州大学
...

その他
....



本講演で紹介します

機械学習を用いた加速器制御

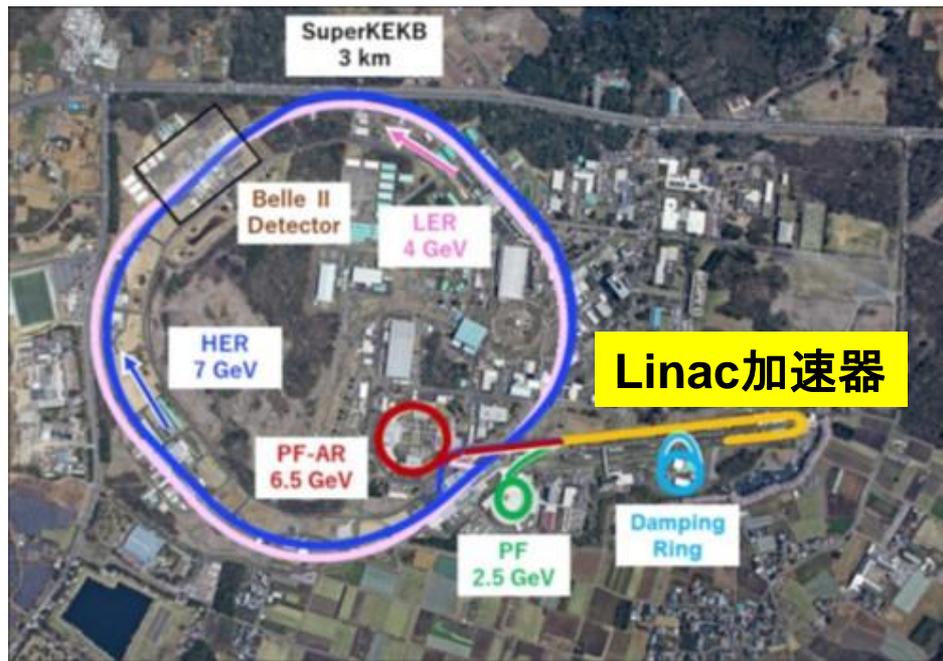
KEK電子陽電子入射器(Linac) 運転調整システムの開発

久野彰浩（阪市大理）, 岩崎昌子（阪市大理, 阪市大NITEP, 阪大RCNP, 阪大IDS）,
佐藤政則（KEK, 総研大）, 佐武いつか（KEK）, 中島悠太（阪大IDS）,
武村紀子（阪大IDS）, 長原一（阪大IDS）, 中野貴志（阪大RCNP, 阪大IDS）

機械学習を用いた加速器制御

KEK Linac加速器調整の開発

KEK Linac、大阪市大、阪大IDS



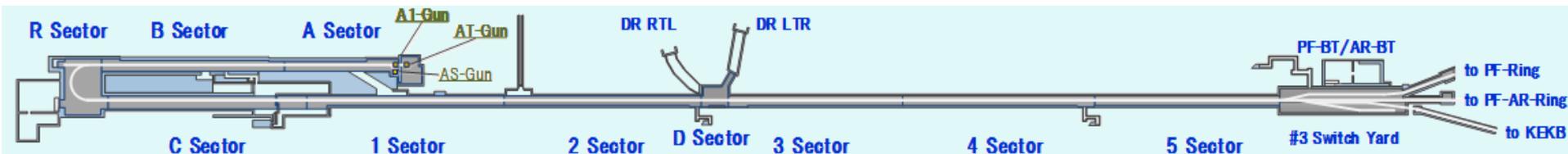
加速器制御の問題点

- 構成要素が多く、**複雑なシステム**
→ 調整速度が個人の技量に依存
- 温度変化等、周囲の**環境が変化する**
→ 常時調整が必要

入射効率を高めるために
効率のよい加速器調整が重要

RFモニター: 60台 ステアリング電磁石: 200台
ビーム位置モニター(BPM): 100台

機械学習を用いた加速器制御を検討



教師なし学習(VAE)による加速器状況の可視化

KEK Linac、大阪市大、阪大IDS

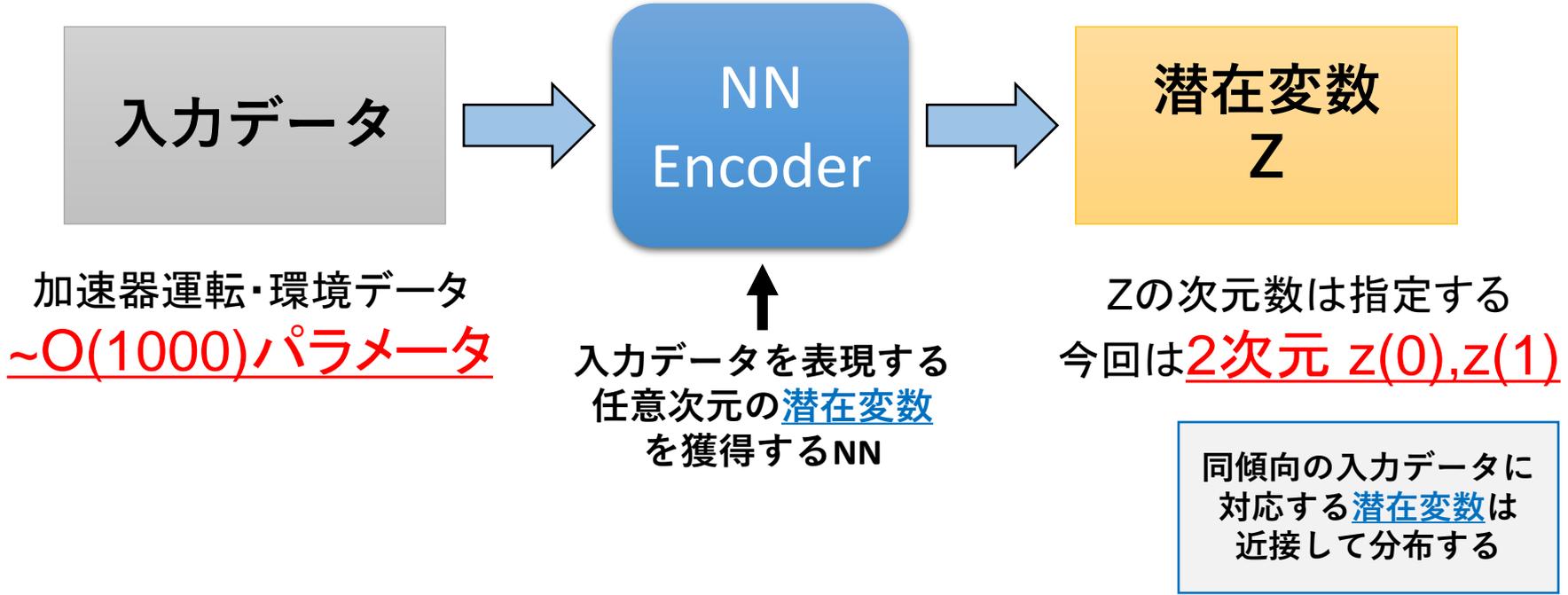
熟練の運転員

加速器の環境パラメータ（周囲の温度・湿度等）や運転パラメータをモニターし「加速器の状況」を判断して、最適な運転パラメータ設定値を予想する

KEK Linac : 環境パラメータ(冷却水温、湿度等) ~700 + 運転パラメータ ~500 → ~O(1000)

- 「加速器の状況」が可視化できないだろうか？
- 「加速器の状況」と「最適な設定値」の相関が得られないだろうか？

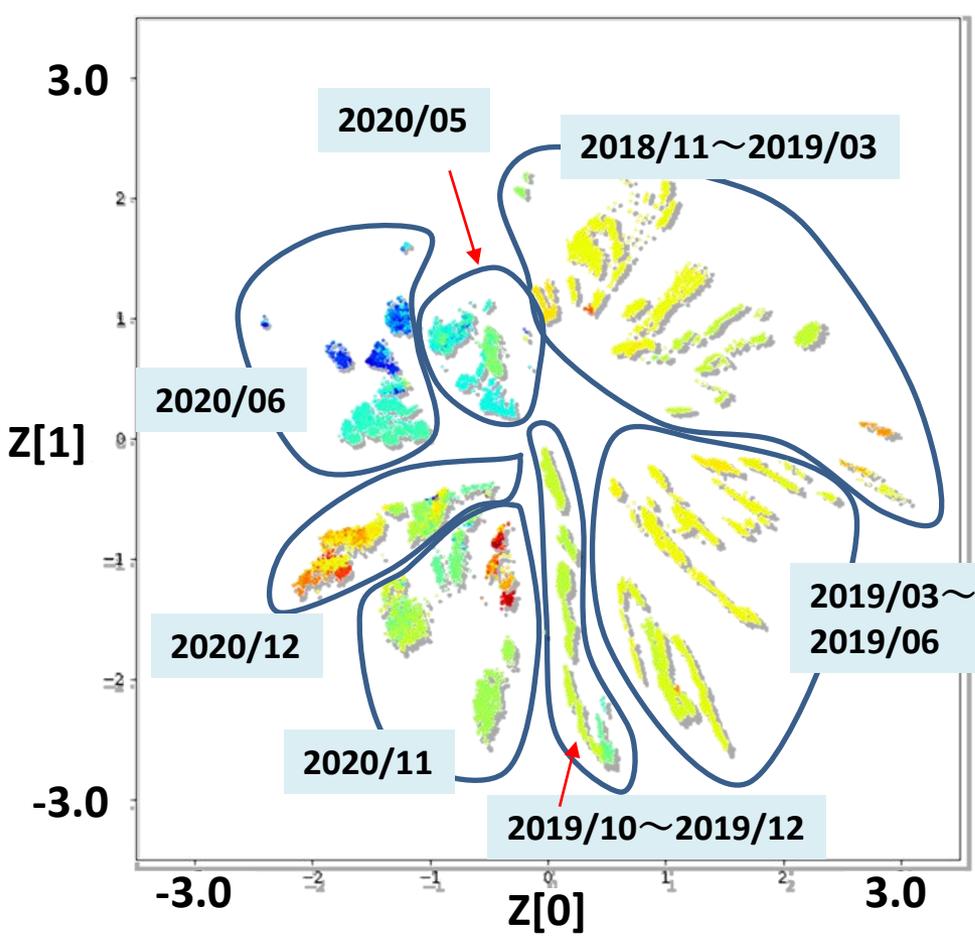
変分オートエンコーダ VAE(Variational Auto Encoder)による
加速器運転・環境パラメータの可視化



教師なし学習(VAE)による加速器状況の可視化

教師なし学習(VAE)による次元削減で
1296パラメータの加速器データを2次元で表現 → **可視化!**

2D Latent Variable Z(0, 1)



A.Hisano (Osaka-City U.)

Steering Magnet (3 sector 2nd)
Current value (normalized to 0-1)

Visualize the accelerator status with 2 parameters

In short term (~1month)
acc. status does not drastically change

In long term (> 6 month)
acc. parameters vary over a wide range

機械学習を用いた加速器運転調整：強化学習の準備

機械学習による運転調整で、環境変化に適応した加速器運転を目指したい
→ **強化学習**の導入が有効だと思われる

強化学習：周囲の環境変化に応じて、最適な行動を選択する機械学習
(報酬を最大化するような行動を選択)

→入射効率を報酬とした強化学習を導入して、入射効率の向上を目指す

ただし、入射効率を最大化するために
実際に加速器の設定値を変化させながら調整することは危険
(車の運転で、ハンドルを左右に切ってから進む方向を決めるのは危険)

加速器運転調整に強化学習を導入するためには以下が必要

1. 機械学習による **適切なパラメータ初期値の予測**
2. 予測精度向上のための **学習用加速器シミュレーター**

準備研究：機械学習による入射効率の予測

加速器パラメータと入射効率の関係がDNNで学習可能か検証

- 過去データを用いて、DNNで最適な初期値が設定可能か？
- 過去データをもとに、シミュレーターが構築可能か？

機械学習を用いた加速器運転調整：入射効率の予測

KEK Linac、大阪市大、阪大IDS

- 加速器パラメータから入射効率を予測するNNをデザインした

入力：環境パラメータ(冷却水温、湿度等) ~700 + 運転パラメータ ~500 → ~O(1000)

出力：入射効率(Q_{ratio})

入射効率の指標として、
加速器上流と下流のBPMにおける電荷量の比 Q_{ratio} を利用

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



加速器運転・環境データ
~O(1000)パラメータ

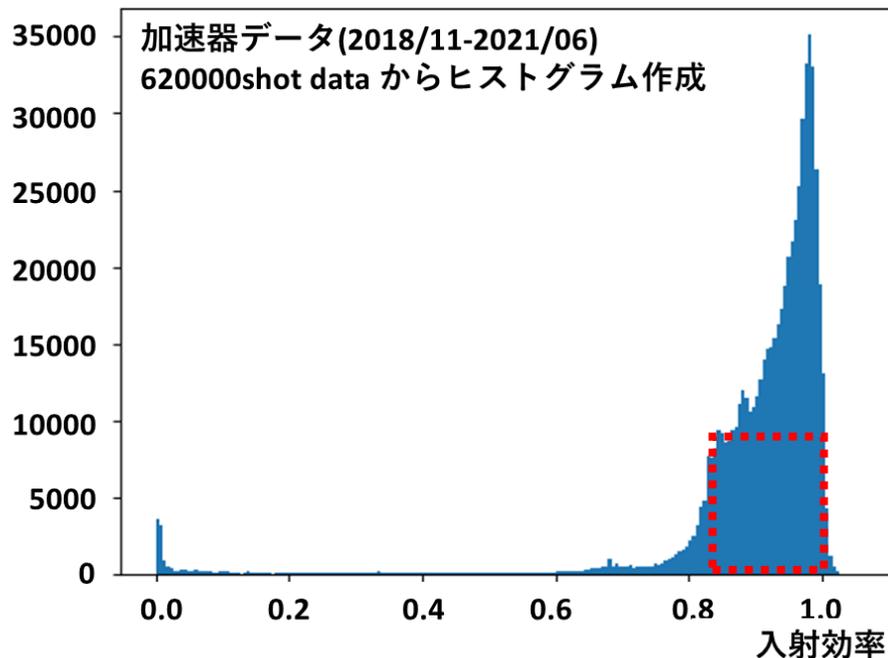
中間層4層
Optimizer: Adam
活性化関数: Relu

機械学習を用いた加速器運転調整：入射効率の予測

KEK Linac、大阪市大、阪大IDS

- NNの学習用データについて

データ数



学習データを均一化するため、入射効率が0.85以上のデータについて、以下3つの区間でデータ数が均一になるよう抽出を行った(赤枠)

- 入射効率 0.85以上0.90未満
- 入射効率 0.90以上0.95未満
- 入射効率 0.95以上

- 2018年11月から2021年6月までのLinac運転データ(SuperKEKB入射用電子ビーム)を用いて、NNの学習・評価を行った

機械学習を用いた加速器運転調整：入射効率の予測

KEK Linac、大阪市大、阪大IDS

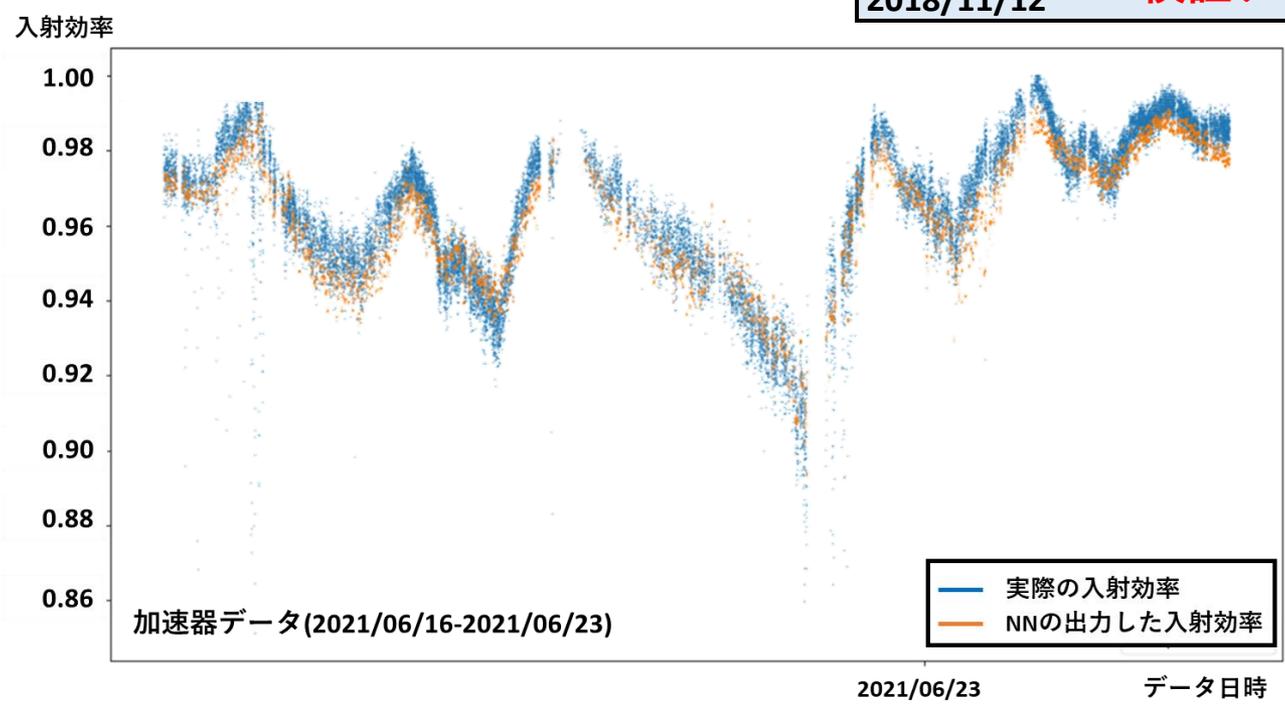
入射効率予測結果 その1

学習・検証データは、どちらもLinac運転データ全期間(2018年11月～2021年6月)を使用

学習用データは、15万イベント
検証用データは、学習に使用していないイベント

2018/11/12 **学習データ** 2021/06/23

2018/11/12 **検証データ** 2021/06/23



A.Hisano
(Osaka-City U.)

非常に良く一致している → DNNが加速器パラメータと入射効率の関係を学習可能

DNN が予測した入射効率と実際の入射効率の平均二乗誤差 = 0.00030

機械学習を用いた加速器運転調整：入射効率の予測

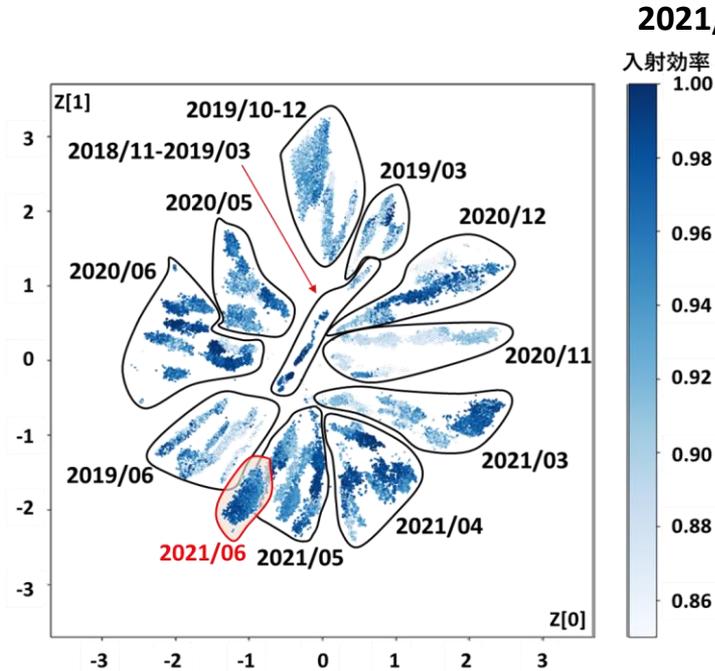
KEK Linac、大阪市大、阪大IDS

入射効率予測結果 その2

検証データは、2021年6月

学習用データは、以下の3パターンを使用

(学習用データは、それぞれ15万イベント)



VAEの出力をみると、
 2021年6月の加速器パラメータの状態は、
 2021年5月の加速器パラメータの状態と類似

→ 2021年5月のデータで学習すると、
 2021年6月の入射効率が予測可能？！

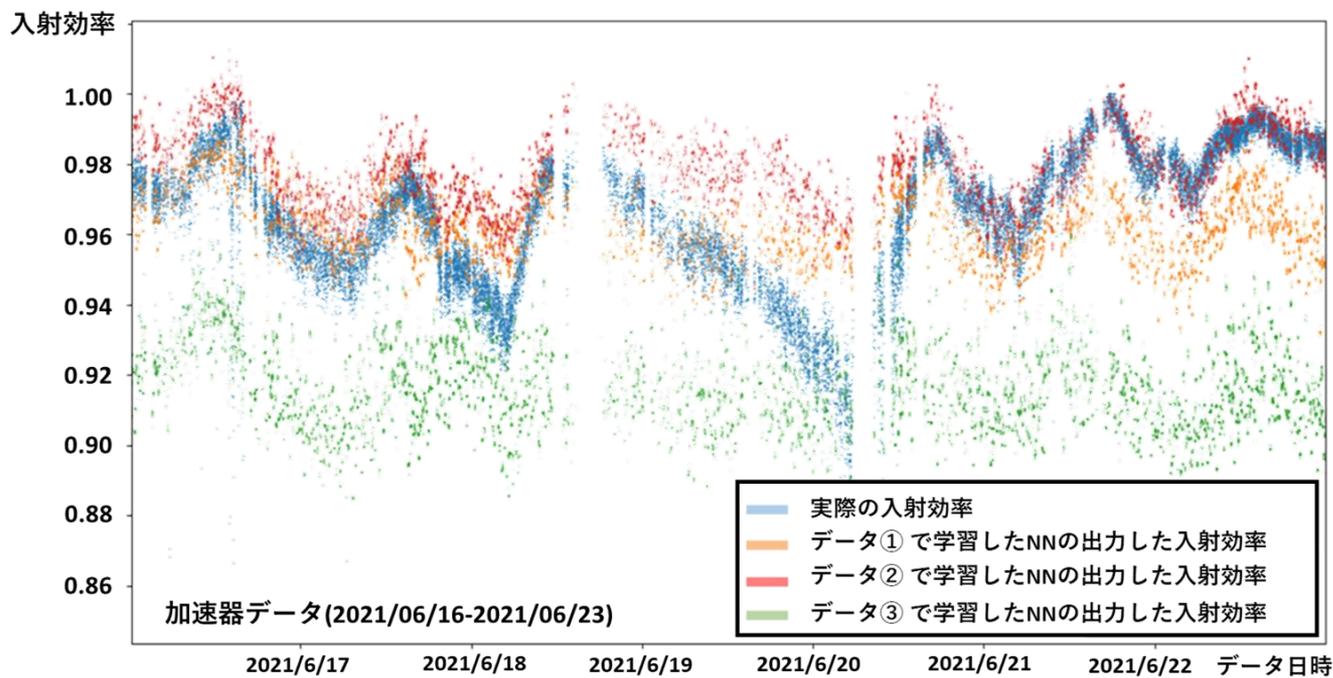
A.Hisano
 (Osaka-City U.)

機械学習を用いた加速器運転調整：入射効率の予測

KEK Linac、大阪市大、阪大IDS

入射効率予測結果 その2

検証データは、2021年6月
 学習用データは、右の3パターンを使用
 (それぞれ15万イベント)



A.Hisano
 (Osaka-City U.)

DNN が予測した入射効率と
 実際の入射効率の
 平均二乗誤差

データ①	Ave 0.00038
データ②	Ave 0.00038
データ③	Ave 0.00300

2021年5月を含むデータ(①②)で学習すると
 2021年6月の入射効率をほぼ予測可能

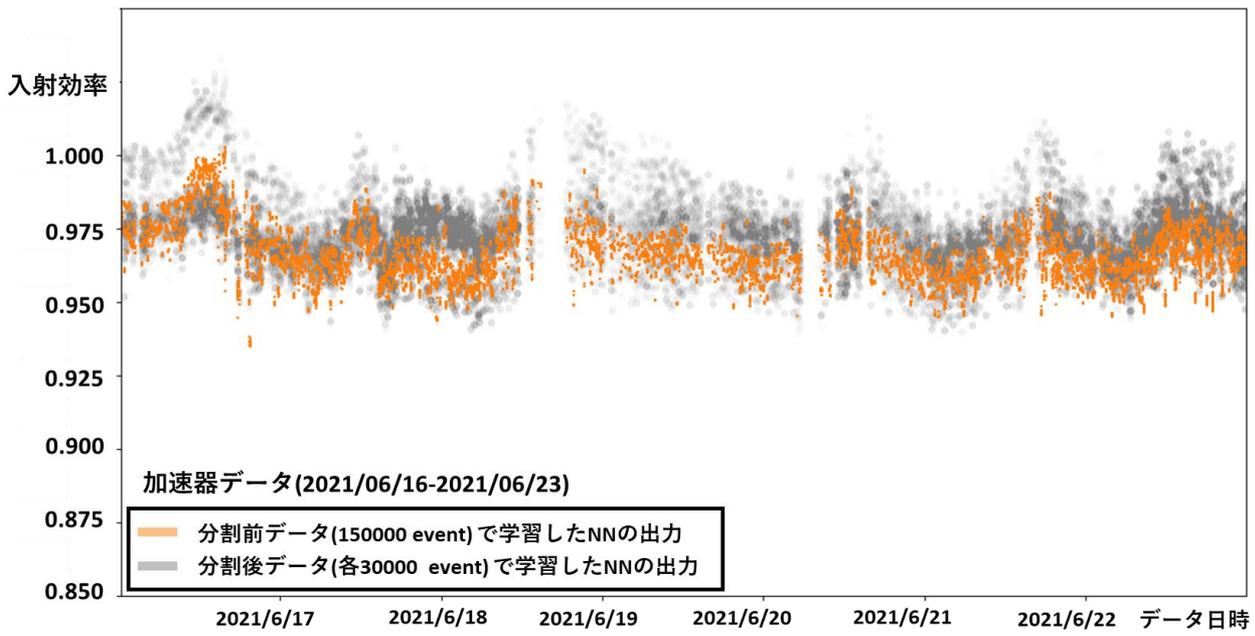
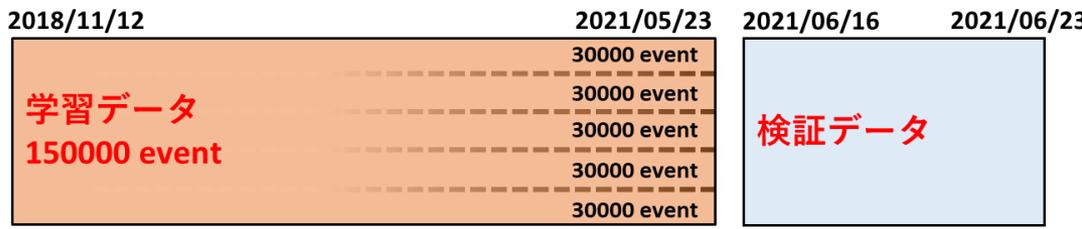
(全区間 0.00030)

機械学習を用いた加速器運転調整：入射効率の予測

KEK Linac、大阪市大、阪大IDS

入射効率予測結果 その3

検証データは、2021年6月
 学習用データは、学習データ①を5分割
 (それぞれ3万イベント)



A.Hisano
 (Osaka-City U.)

DNN が予測した入射効率と実際の
 入射効率の平均二乗誤差

- 分割前(150000 event) Ave 0.00036
- 分割後(30000 event) Ave 0.00043
- Ave 0.00036
- Ave 0.00045
- Ave 0.00067
- Ave 0.00038

学習用データの統計が高い方が入射効率の予測精度が高くなる

→ 加速器シミュレータで学習用データの統計量を確保することが有効?!

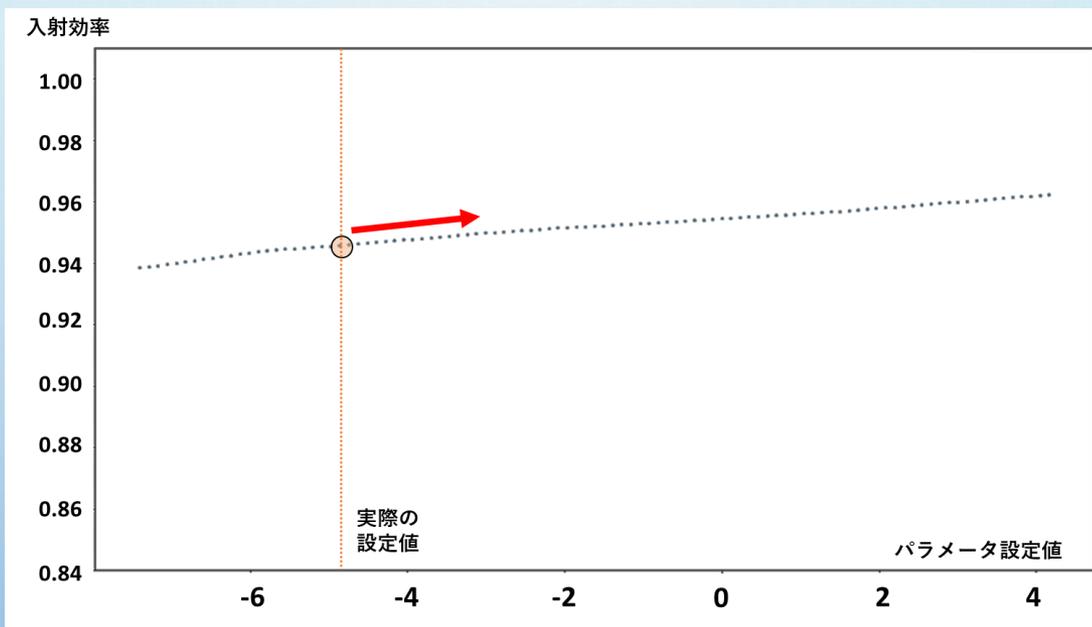
まとめ

機械学習を用いた加速器制御の開発を進めている

- 1) VAEを用いて、 $\sim O(1000)$ の加速器パラメータを可視化することができた
→ 複雑なシステムの可視化に成功
- 2) 機械学習により過去データから入射効率を予測できた
→ 加速器パラメータと入射効率の相関が学習可能であることが示された
 - 2-1: 加速器の状態が類似した過去データで学習すると予測可能になる
 - 2-2: 学習用データの統計が高いと、予測精度が向上する

今後の展望

- 1) 機械学習で得られた加速器パラメータと入射効率の相関から
運転パラメータの最適化を行う



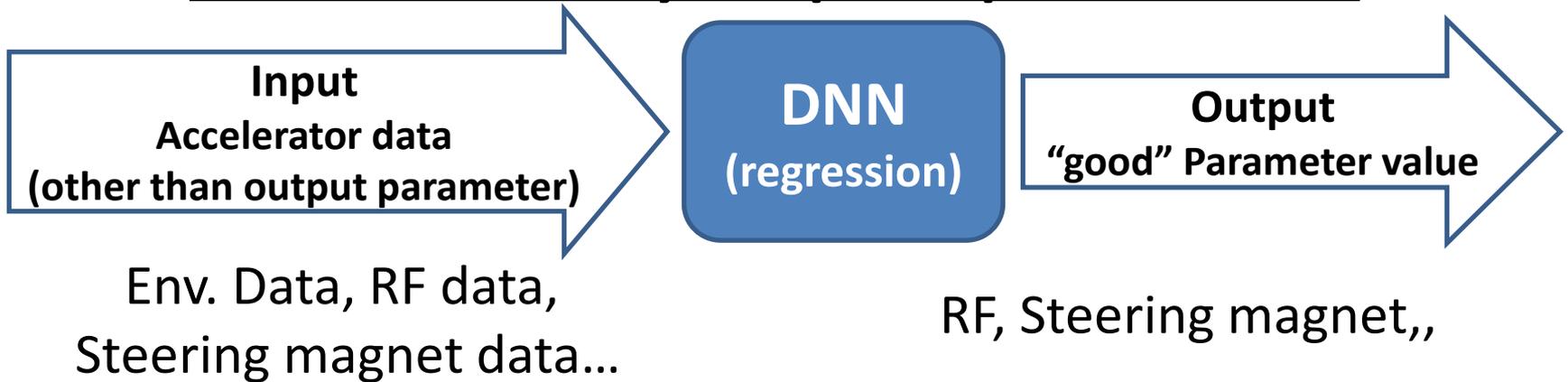
- 2) 機械学習を用いた加速器シミュレータの開発

開発状況について、物理学会で発表予定

Parameter Tuning using DNN

To get the “good” accelerator tuning value to achieve the high injection efficiency, we use DNN

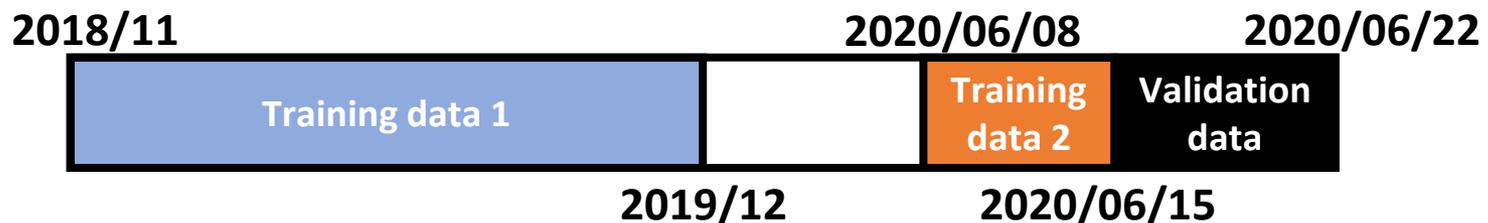
DNN is trained by the past operation data



Training data 1 = ~ 1.5 year ago

Training data 2 = ~ 1 week ago, just before the validation data

Training data 3 = continuously update (~ 1 day)



機械学習を用いた加速器調整

A.Hisano (Osaka-City U.)

2018/11

2020/06/08

2020/06/22



2019/12

2020/06/15

Steering Magnet
(3 sector 2nd)
Current value

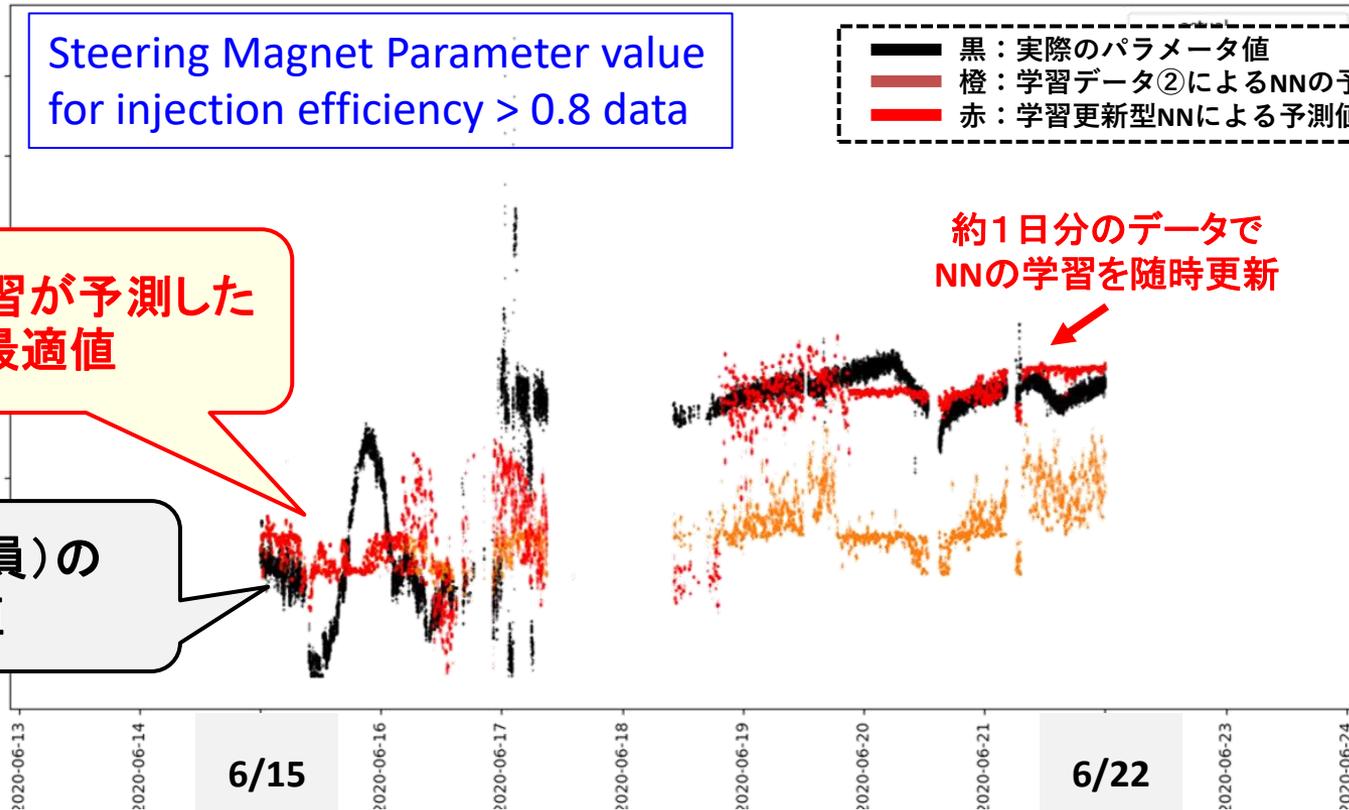
Steering Magnet Parameter value
for injection efficiency > 0.8 data

黒：実際のパラメータ値
橙：学習データ②によるNNの予測値
赤：学習更新型NNによる予測値

機械学習が予測した
最適値

人間(運転員)の
設定値

約1日分のデータで
NNの学習を随時更新



人間 vs 機械学習?!