Machine Learning at BESSY II

Alexander Schütt

Helmholtz-Zentrum Berlin

ACCLAIM Meeting 2023

Table of Contents

Table of Contents

Problem Description

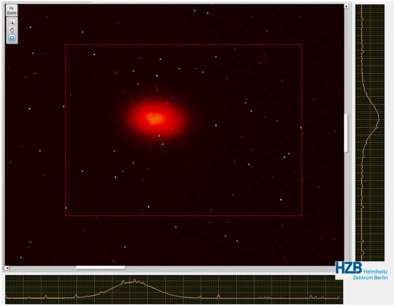


Figure 1: Beam image taken by a pinhole camera. Picture provided by the Helmholtz Zentrum Berlin.

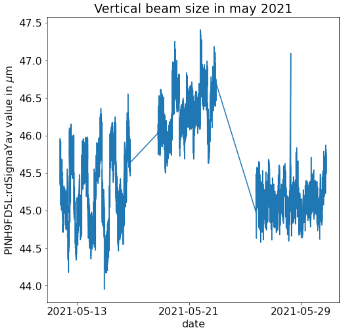


Figure 2: Vertical beam size of electron beam, measured using a pinhole camera.

Problem Description

- Task: Predict the vertical beam size of the electron beam
- Why?
  - Derive vertical beam size from device settings
  - Long term: Adjust noise generator for constant beamsize $\Rightarrow$ minimize impact from one user to another
  - Preparation for BESSY III: round beam

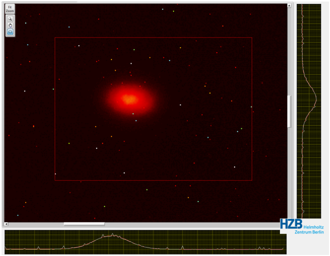
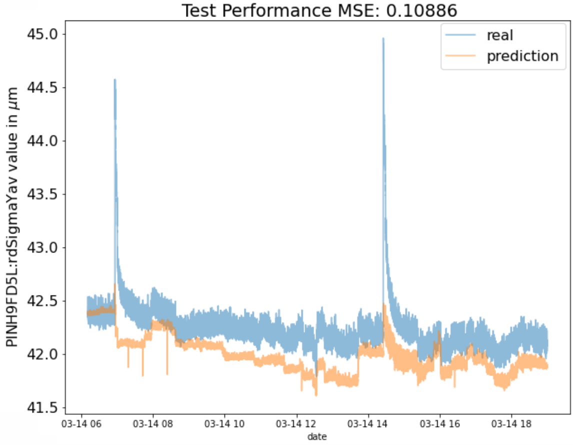


Figure 3: Beam image taken by a pinhole camera. Picture provided by the Helmholtz Zentrum Berlin.

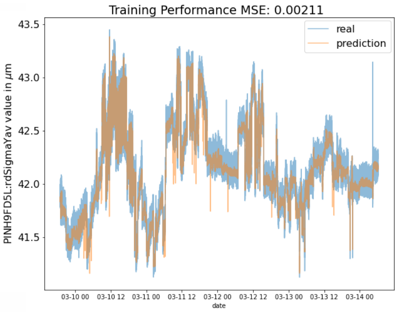
Model setting

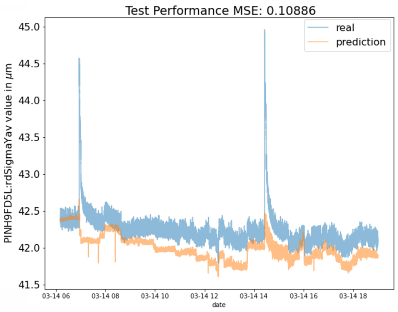
- How is the data saved?
  - time series data, i.e. time points $t_1, \ldots, t_N$ exist
  - 152 different BESSY II parameters, e.g. undulator settings, vacuum statuses, steerer settings
  - vertical beam size is given as real number
- In our case typically $t_i - t_{i-1} \approx \frac{1}{4}$ second and $N > 1.5 * 10^6$ per week
- In the following we have split the dataset $\mathcal{D}$ into $\mathcal{D}_{train}$ and $\mathcal{D}_{test}$

Neural Network results


Test Performance MSE: 0.10886
real
prediction
PINH9FD5L:rdSigmaYav value in μm
45.0
44.5
44.0
43.5
43.0
42.5
42.0
41.5
03-14 06
03-14 08
03-14 10
03-14 12
03-14 14
03-14 16
03-14 18
date

Performance Analysis





- Results look promising, only slight offset
  $\Rightarrow$ could be solved by more data
- In some sections visible: Time dependence
- Time dependence result by thermal effects

Long short-term memory (LSTM)

- Special RNN, enabling long-term dependencies
- Fundamental difference:
  - multiple input values are considered, i.e. y_t depends on $x_t, x_{t-1}, \ldots, x_{t-m}$
  - allows information to flow over several blocks
  - using a cell state

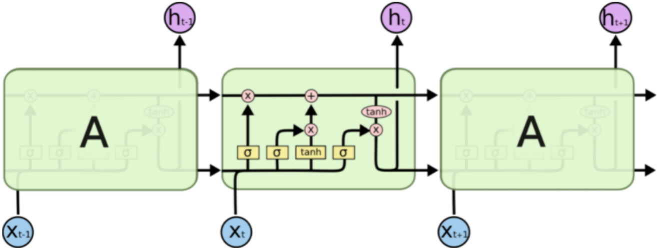
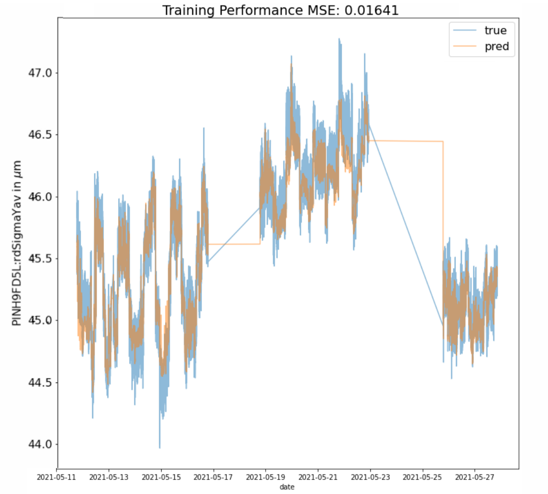
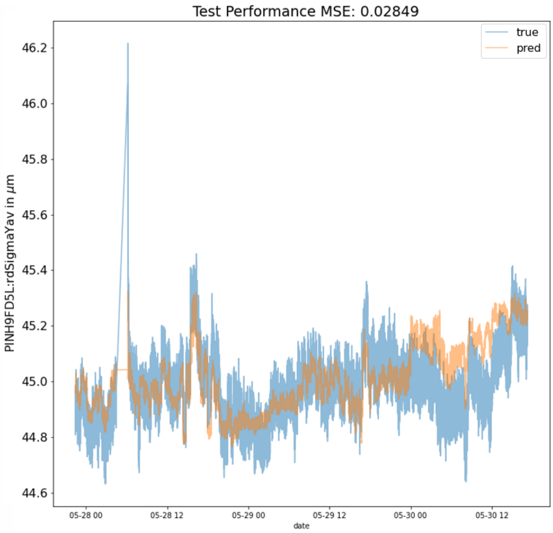


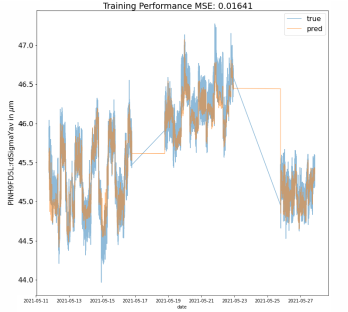
Figure 4: LSTM structure. Picture by C. Olah ([3])

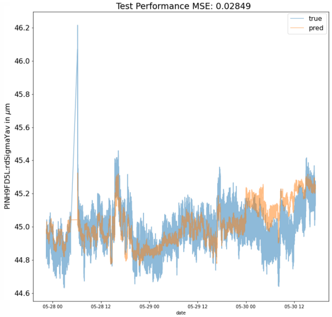
Recurrent neural network regression




Test Performance MSE: 0.02849
true
pred
46.2
46.0
45.8
45.6
45.4
45.2
45.0
44.8
44.6
PINH9FD5L:rdSigmaYav in μm
05-28 00
05-28 12
05-29 00
05-29 12
05-30 00
05-30 12
date

Performance Analysis

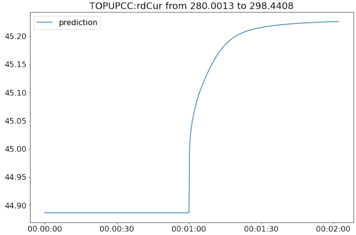


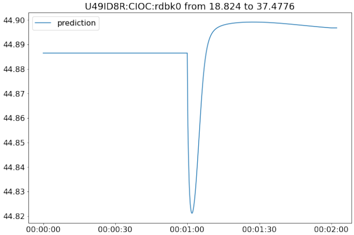


- Very promising results
- Issue: Training took 22h (CPU machine)

Decay Behaviour: Predicted

- Goal: Find influence of single parameters
- Created Artificial dataset
- After 1min: change parameter from min to max





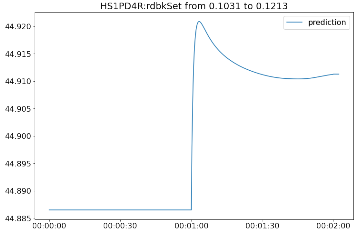


Table of Contents

Introduction: 4 Kicker bump

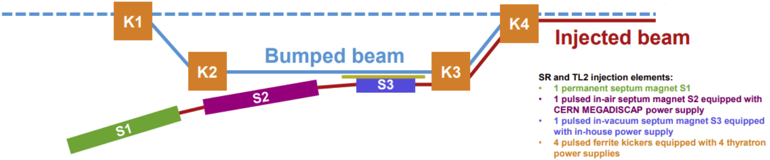


Figure 5: 4 Kicker bump

Disadvantages:

- affects stored beam
- might cause the loss of electrons

Introduction: Nonlinear Kicker bump

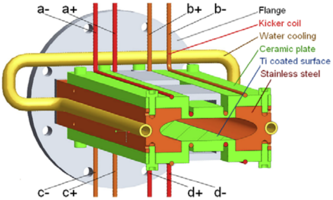


Figure 6: Kicker magnet structure by [1].

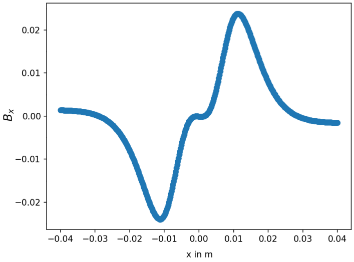


Figure 7: Kicker field strength plot.

Advantages:

- Only one kicker
- Only small effects on stored beam

Problem Description

- Task: Find optimal NLK settings
- Steps:
  - Chose steerer settings leading to NLK
  - Chose activation round and strength of NLK
- Goal:
  - Improve injection efficiency
  - Minimize impact on stored beam

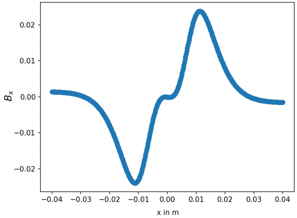


Figure 8: Kicker field strength plot.

Model setting

- In each round from $0, \ldots, 1000$:
  - Input: $x, \mathrm{p}x$ information of electrons
  - Chose if we want to activate NLK
  - Chose strength of NLK
- NLK can only be activated once
- Goal: Maximize number of electrons that survive 1000 rounds
- When is a electron lost?
  - If it hits the septum sheet, that lies at 15mm
- Modelable by Markov Decision Process

Model setting

- In each round from $0, \dots, 1000$:
    - Input: $x, \mathrm{p}x$ information of electrons
    - Chose if we want to activate NLK
    - Chose strength of NLK
- NLK can only be activated once

Definition (Markov Decision Process)

A Markov Decision Process is a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, where

- $\mathcal{S}$ state space (electron information + NLK activated?)
- $\mathcal{A}$ action space (NLK strength $\in [-1, 1]$)
- $\mathcal{P}$ probability transition, with $\mathcal{P}_a(s, s') = \mathbb{P}(S_{t+1} = s' | S_t = s, A_t = a)$
- $\mathcal{R}$ reward function and γ discount factor ($\mathcal{R}$ portion survived)

Intuition

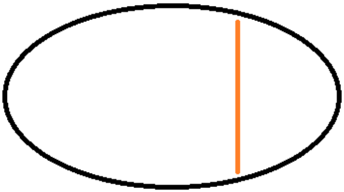


Figure 9: Figure showing the septum sheet at 15mm.

- Injected electrons survive multiple rounds without a kick
- But eventually crash into the septum sheet

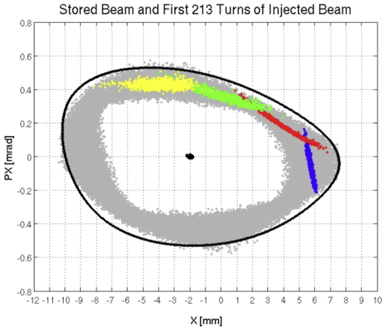
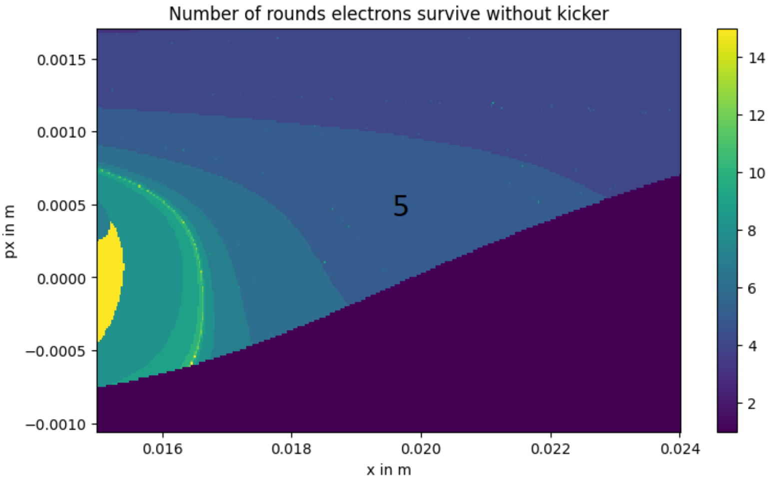


Figure 10: Example of phase space. Blue first round, red second, green third and yellow fourth round. Modified from Lin Liu et al. [2]

Without Kicker (deterministic model)



How to use the Kicker

- Kicker only changes flight direction
- Kicker strength scales magnetic field strength
- How to kick optimally?
    - x value of electron must be between -15mm and 15mm
    - Kicker strength must be chosen, s.t. new drift is equal 0
- Problem: If x value of electron is too close to 0

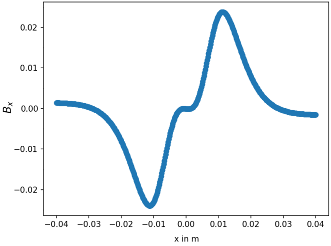
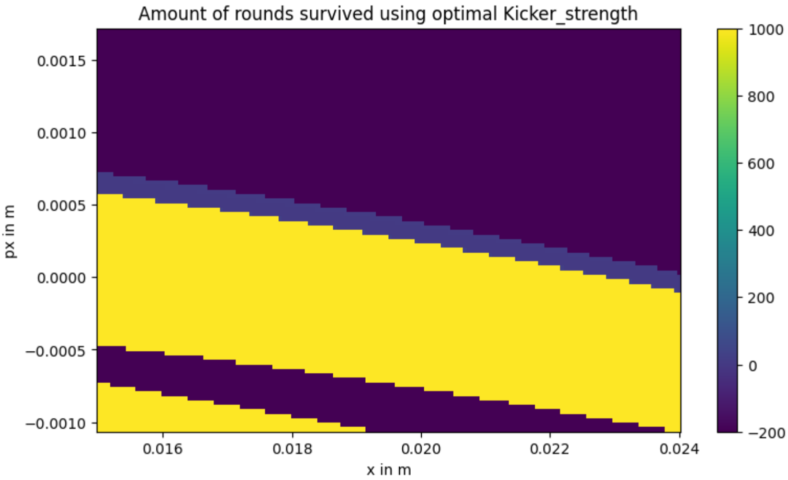
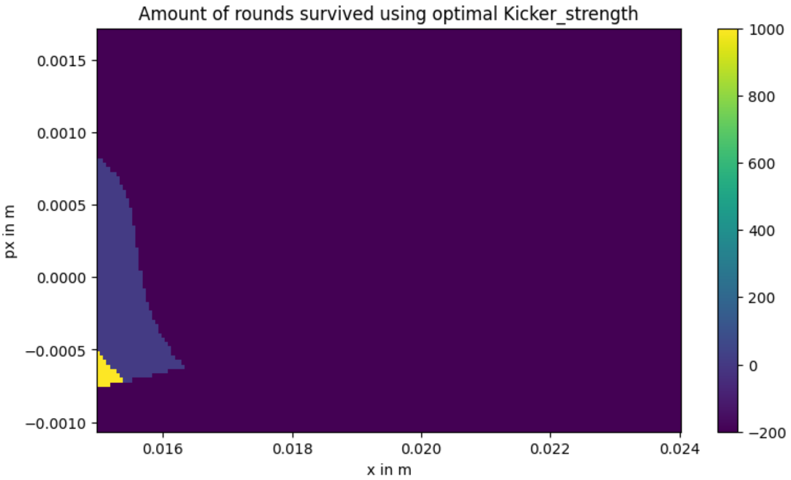


Figure 11: Kicker field strength plot.

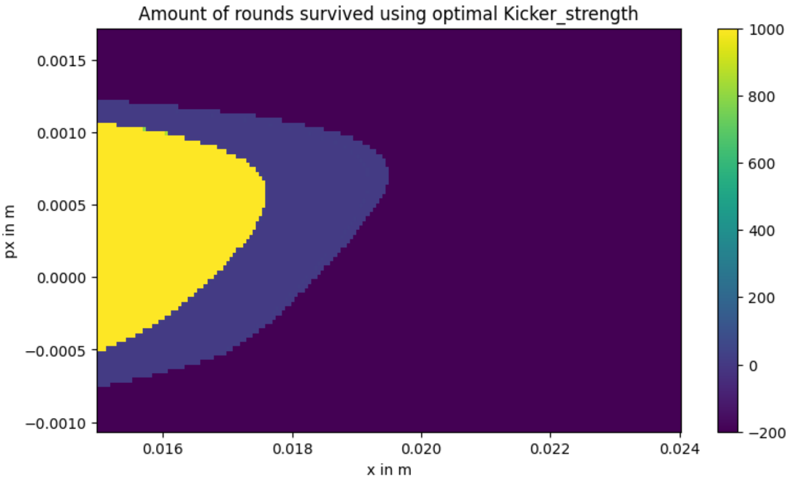
Programming Results: Round 0



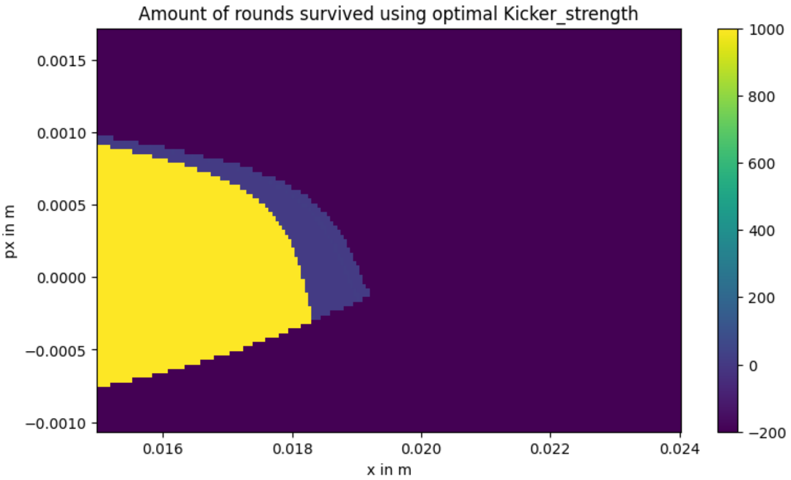
Programming Results: Round 1



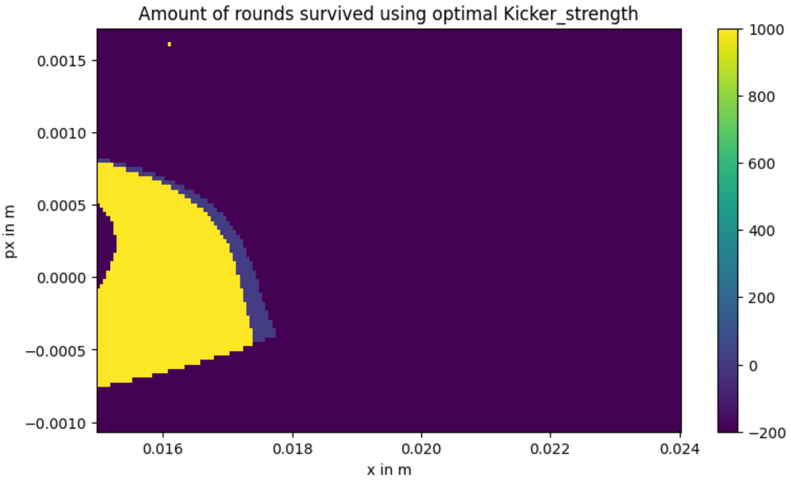
Programming Results: Round 2



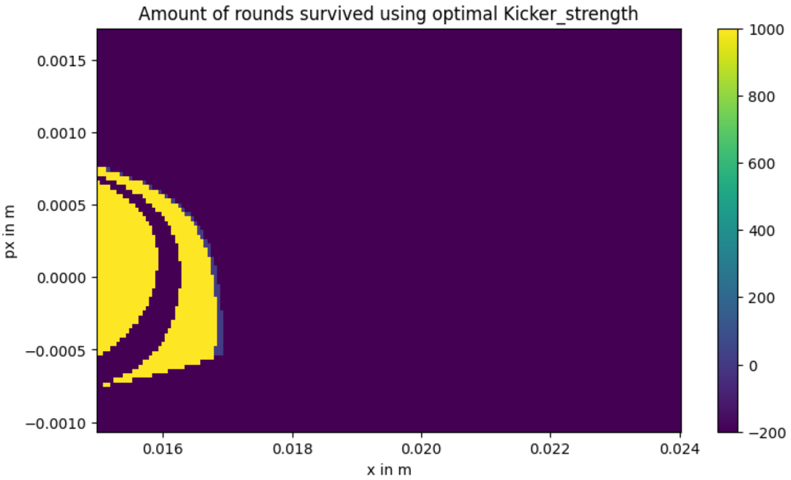
Programming Results: Round 3



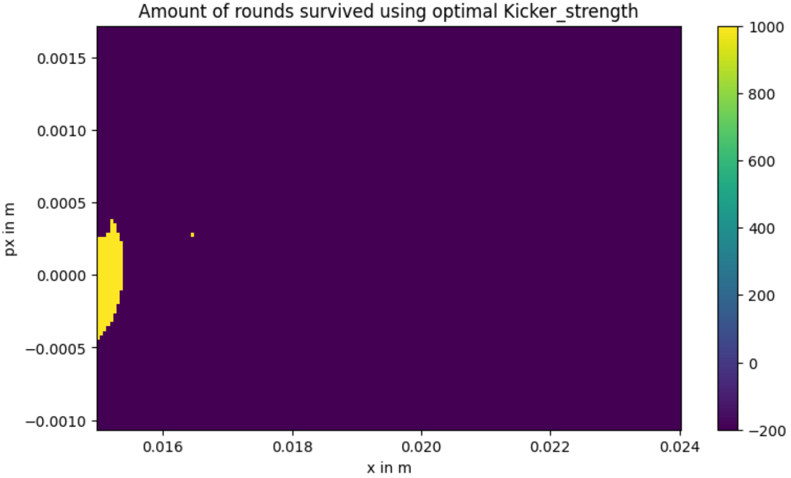
Programming Results: Round 4



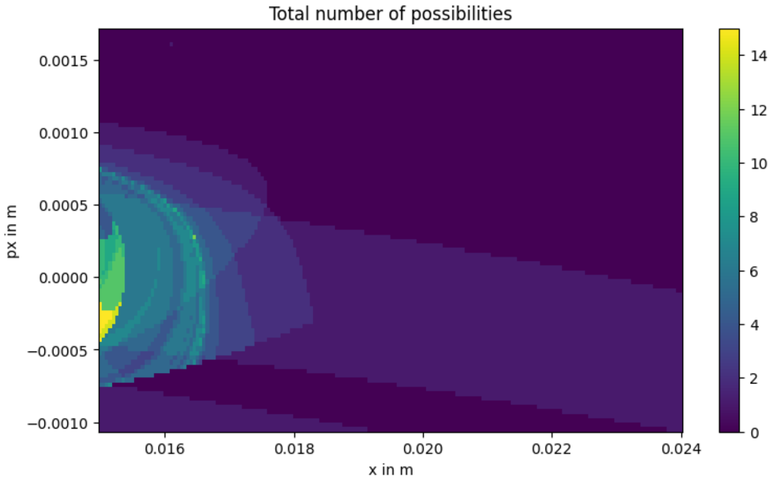
Programming Results: Round 5



Programming Results: Round 13



Programming Results



Programming Results

- We can see:
  - Using only Round 0 is not optimal
  - Lot of opportunities
- Left to do:
  - Find optimal NLK settings for distribution of electrons
  - Stochastic model
  - Find the optimal steerer settings before the NLK

Table of Contents

Summary

What have we seen?

- Vertical beam size prediction
  - Motivated choice of LSTMs
  - Compared performance for vertical beam size prediction
  - Analysed how different parameters change the beam size
- Nonlinear Kicker Injection
  - 4 kicker bump and NLK injection
  - Advantages using NLK injection
  - Intuition behind the NLK functionality
  - First programming results
  - What is left to do

References

[1] T Atkinson et al. "Development of a non-linear kicker system to facilitate a new injection scheme for the bessy ii storage ring". In: (Jan. 2011).

[2] Lin Liu et al. "Injection Dynamics for Sirius Using a Nonlinear Kicker". In: 2016.

[3] Christopher Olah. *Understanding LSTM Networks*. 2015. URL: https://colah.github.io/posts/2015-08-Understanding-LSTMs/ (visited on 11/04/2021).

[4] Alexander Schütt. *Bachelorthesis*. 2021. URL: https://tubcloud.tu-berlin.de/s/k4s33NQGxA6ad7X.