Machine learning for ASTRA simulations

DESY summer student project

Sergei Kladov Novosibirsk, 02.09.21





The problem and the goal

Problem

- Simulations take a lot of time
- Necessity of scanning a wide range of parameters
- Plain simulations slow down the optimization processes
- Precize results can be obtained only by these time-consumung simulations

Goal

• Write a software able to make a decent estimation of the output parameters based on the input parameters in a faster way





The purpose of ML

How to fit simulation data?

Why ML?

- Usual fit when we know the function
- ML fit when we don't know the function

The goal for our ML

- Sparse dataset
- Every point matters
- Small simulation errors



Methods and datasets

Methods:

- Neural network
- Decision tree
- Gradient tree boosting
- Kernel ridge, linear models
- Voting (averaging of others)



Datasets

- All are for a simple gun with booster lattice
- 1d initial charge scan
- 2d charge and solenoid field grid-like scan
- 2d random scan
- 3d charge, solenoid and electron gun fields scans



A simple tree model example

Fit results

General performance

Performance

- Qualitative match
- Red dots simulation scan points
- Can be plotted for each of models

Models

- Linear models and kernel ridge are not implemented for >1 dimension datasets
- NN fit on the top, Tree fit on the bottom



Fit results

Comparison of the models

Fit comparison

- Training accuracy of all models ~100%*
- Rigidity of the tree models
- NN + Boost their averaging

Relative deviation

- Tree errors up to 40%*
- Best model average errors ~5%
- NN + Boost is usually the best choice



Java automation program

Send a simulation task

- Convenient way to perform ASTRA scans
- Two possible scan mods

- Predict
- Train different models
- Visualize the training results
- Set the input parameters and obtain the estimation

🕌 Simulatio	on Estimator			—		👫 Simul	ation Estima	ator		—	
run.in	Load in file			Mak	e a prediction	/ 10	ad simulati	on results		Make a simula	tion dataset
parameter	lower value	upper value	step (for ste	parameter	Value	para	meter	value	estimate?	parameter	prediction
BUNCH	0.2	0.2	1.0	RUN	1.0			·	V	X pos	0.0
IAXB	359.99965	450	1.0	MAX_STEP	200000.0				V	Y pos	0.0
			_	ZSTART	0.0				V	Z pos	0.0
				ZSTOP	5.277					average kin	0.0
				ZEMIT	400.0				V	alfa X	0.0
				ZPHASE	1.0				V	alfa Y	0.0
				SCREEN(1)	2.4				V	Charge	0.0
				SCREEN(2)	5.277				¥	sigZ	0.0
				NRAD	50.0				V	sigX	0.0
				NLONG_IN	500.0				V	sigY	0.0
				N_MIN	800.0				V	energy spre	0.0
				MAX_SCALE	0.05				V	emittance Y	0.0
				MAX_COUNT	20.0				V	emittance X	0.0
				S_POS(1)	0.0						
				MAXE(1)	60.5						
				MAXE(2)	16.0						
				C_POS(1)	0.0						
				C POS(2)	2.675						
				NUE(1)	1.3						
				NUE(2)	1.3						
				PHI(1)	0.0						
				PHI(2)	0.0						
			->								
				<-							
				samples	number:						
		S	end stepped jo	b Send rando	m-sampled job	predict	train and	predict choose model	NN+Boost	~	

GUI visualization of the trained model

3d plots for 2d datasets (NN fit)





- Input parameters scan times can be significantly reduced with the usage of ML
- Multiple ML algorithms are implemented
- The models performance is tested on various datasets
- An error in the worst case is ~40%, but the usual relative divergence of the best model is ~5%
- The Java GUI program is written for ease of train, prediction, and scans

Thank you

Contact

DESY. Deutsches	Sergei Kladov
Elektronen-Synchrotron	PITZ
	s.kladov@g.nsu.ru
www.desy.de	+7 923 239 70 24
	https://kirikaueno.github.io/SanaFanSite/