Forecasting of Beam Interlocks in High Intensity (Hadron) Accelerators

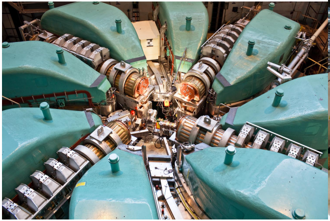
Andreas Adelmann & Jochem Snuverink

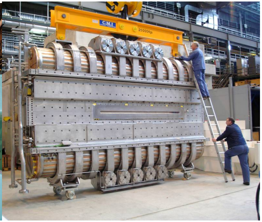
March 1, 2018

Modelling Challenges PSI

Consider a 0.59 GeV, 2.3 mA 1.4 MW (CW) Proton Cyclotron facility

- uncontrolled & controlled beam loss $\mathcal{O}(2\mu A = const)$ in large and complex structures

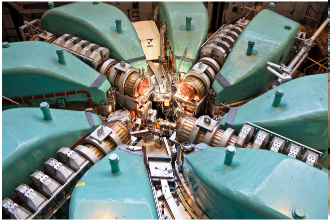


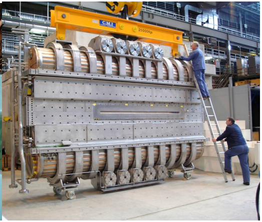


Modelling Challenges PSI

Consider a 0.59 GeV, 2.3 mA 1.4 MW (CW) Proton Cyclotron facility

- uncontrolled & controlled beam loss $\mathcal{O}(2\mu A = const)$ in large and complex structures
- PSI Ring: 99.98% transmission $\rightarrow \mathcal{O}(10^{-4}) \rightarrow 4\sigma$





Motivations for ML

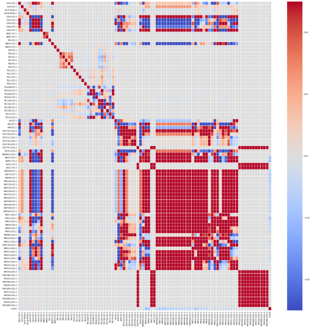
- Predict and counteract interlocks
  - aid/guide machine tuning
  - interlocks causes $\mathcal{O}(2\%)$ beam time loss
- Reduce controlled and uncontrolled beam loss
  - less activation
  - better machine protection
- Cheap to Evaluate Surrogate Models
  - ab-initio models are impracticable for operation support
  - precise on-line models
  - check poster **Surrogate Models based on Supervised Learning and Polynomial Chaos for the Argonne Wakefield Accelerator**

Challenges

Beam-time statistics for HIPA (Courtesy of A.Parfenova)		
	2017	**10 yr av.**
Scheduled user beam time	4838 hr	
Current integral (Ah)		
meson production targets	7.97	
SINQ	4.10	
UCN	0.064	
isotope production	0.013	
Outages	228 hr	
Availability	93.1%	90 %

Challenges

- rare events
- large variability of the data
- high dimensionality of the data



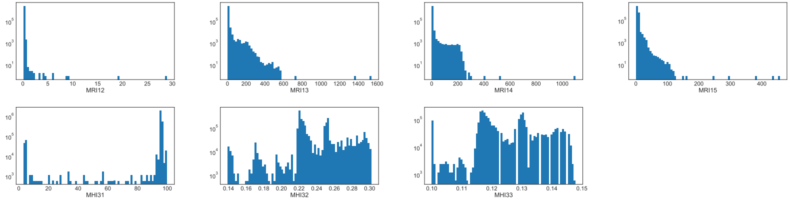
Real Dataset from HIPA

more on data: see talk J. Snuverink on Friday

- $\approx$ 90 values (magnet currents, non-intercepting diagnostics, etc.)
- cleaning and filtering: 57 significant parameters
- data from last two weeks of the 2017 run
  - 10 Hz (11 900 000 samples)
  - in total about 450 interlocks (very unbalanced!)

Data is Key (ionisation monitors)

- raw data
- almost the same machine setup



Interlock prediction: Binary Classifier I

- supervised leaning
- Theano `http://deeplearning.net/software/theano/`

Logistic regression:

- x: input variables $\dim x = 57$
- y: 0,1 (0: no interlock, 1: interlock)
- Logistic function:

$$P(y = interlock|x, w, b) = \frac{1}{1 + e^{-b-w*x}}$$

PAUL SCHERRER INSTITUT

Interlock prediction: Binary Classifier II

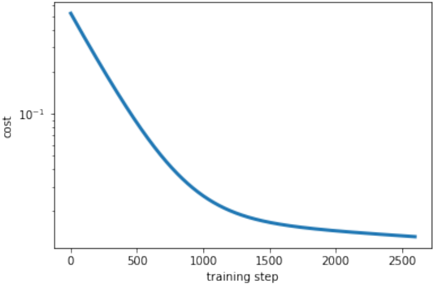
- Cost function c to be minimised (cross entropy loss + L_2-regularisation):

$$
\begin{aligned}
c = & -\frac{1}{N}\sum(y\log(P(y)) \\
& + (1-y)\log(1-P(y))) + \lambda * \Sigma w^2
\end{aligned}
$$

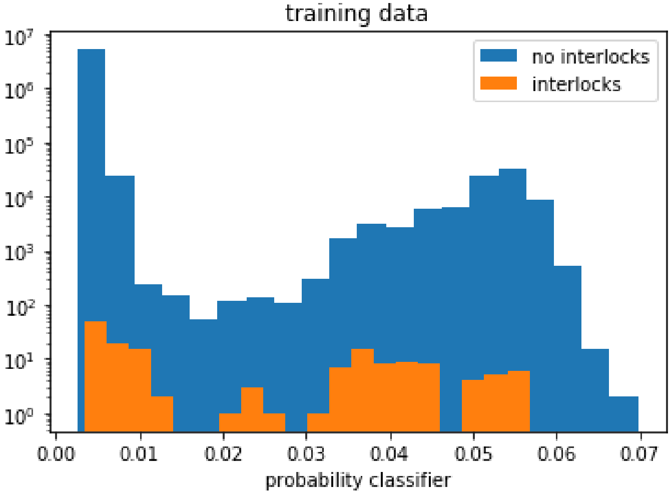
- Optimisation method: stochastic gradient descent
- Hyper parameters: regularisation factor λ (0.01), learning rate $\alpha = 0.1 \ldots 0.01$

Training

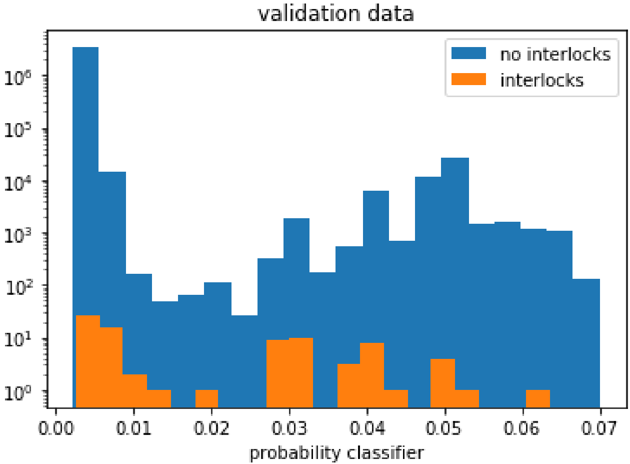
- 4 days of training data
- training until model error is acceptable (empirical tuning of α)
- 4 days of validation data



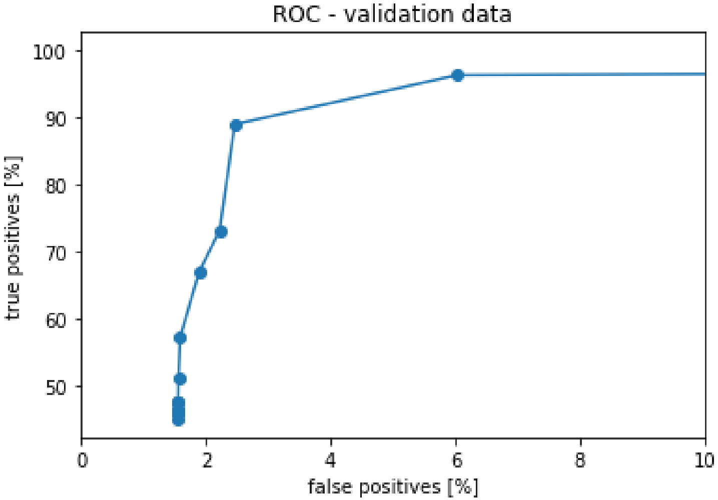
Training Results (super prelim.)



Validation (super prelim.)



Selecting what you want ...



An Example from Fusion I

Most critical problem for Magnetic Fusion Energy (MFE)

avoid/mitigate large-scale major disruptions

- Current Status: 8 years of R&D results using Support Vector Machine ML
  - reported success rates in mid-80% range for JET 30 ms before disruptions , BUT $> 95\%$ with false alarm rate $< 3\%$ actually needed for ITER (P. DeVries, et al. (2015)

Deep Recurrent Neural Networks (RNNs) I

- Deep
  - Hierarchical representation of complex data, building up salient features
  - Obviating the need for hand tuning, feature engineering, and feature selection
- Recurrent
  - Natural notion of time and memory: at every time-step, the output depends on $s(t-1)$ and $x(t)$
  - The internal state can act as memory and accumulate information of what has happened in the past

PAUL SCHERRER INSTITUT PSI

Summary and Outlook

- We used a very simplistic model for static prediction of interlocks (many thanks to Auralee for guidance)
- Reasonable results are encouraging

Most critical for high intensity accelerators

- Predict a rare event (interlock) $T_p = \mathcal{O}(x \times 10), x$ small [ms]
- Develop a surrogate model for measured losses

Summary and Outlook

- Use the simplistic regression model to properly select/normalise data
- Learn / understand quality requiements of the data x
- Follow the ITER route by using Deep Recurrent Neural Networks (RNNs)
- your input is very welcome ...