BELLE II PIXELDETECTOR CLUSTER ANALYSES USING NEURAL NETWORK ALGORITHMS

STEPHANIE KÄS

ON THE BEHALF OF THE BELLE II GROUP AT THE UNIVERSITY OF GIESSEN

Results by: apl. Prof. Dr. J. S. Lange, K. Dort, S. Käs, M. Peter, I. Heinz, J. Bilk, J. Budak, P. Lehnhardt, F. Zorn
PROJECTS RELATED TO DATA SCIENCES AND AI….

- Classic MLP & Kohonen Maps
- Hopfield Networks, Voxels, Autoencoder...
- Networks on FPGA
BELLE II EXPERIMENT

Location: Tsukuba, Japan
High luminosity
BELLE II PIXELDETECTOR

- Innermost detector
- Pixelated silicon sensors (PXD)
- 2 layers of 40 sensors each
- 8 M pixels

Captures highly ionizing particles.
PXD clusters

Signal: Antideuterons

Background

9x9 matrix ADC values

Low ADC values

High ADC values

M. Peter, Unpublished
Cluster properties

- Total charge
- Minimum charge
- Maximum charge
- Total size
- Size in u
- Size in v
ANTIDUEUTERON DATASET

Goal
Differenciate between $\bar{d}$ and background
ANTIDEUTERON DATASET

Goal
Differenciate between $\bar{d}$ and background.

K. Dort, Search for Highly Ionizing Particles with the Pixel Detector in the Belle II Experiment.
PROJECTS RELATED TO DATA SCIENCES AND AI….

- Classic MLP & Kohonen Maps
- Hopfield Networks, Voxels, Autoencoder…
- Networks on FPGA
**KOHONEN MAPS**

- Unsupervised learning
- Self-organizing

Figure 6: The first step of a self-organizing map. One can see the data clouds in blue, the vectors of each neuron as black x’s and on the floor the U-matrix.
KOHONEN MAPS

Background-like nodes

Antideuteron-like nodes

S. Käs, Multiparameter Analysis of the Belle II Pixeldetector's Data.

Self-Organizing Maps Parameters

- 15 x 15 Nodes
- Neighborhood function: Gaussian
- Width of Gaussian: 7
- Learning Rate: 0.01
KOHONEN MAPS

**Efficiency**

\[
\text{Efficiency} = \frac{\text{number of correctly identified signals}}{\text{total number of signals}}
\]

**Background rejection**

\[
\text{Background rejection} = \frac{\text{number of correctly identified BG data}}{\text{total number of BG data}}
\]

Reminder: ROC-Curves

K. Dort, Search for Highly Ionizing Particles with the Pixel Detector in the Belle II Experiment.
COMPARISON: KOHONEN MAP & MLP

MLP performed better.

K. Dort. Search for Highly Ionizing Particles with the Pixel Detector in the Belle II Experiment.
PROJECTS RELATED TO DATA SCIENCES AND AI….

CLASSIC MLP & KOHONEN MAPS

HOPFIELD NETWORKS, VOXELS, AUTOENCODER…

NETWORKS ON FPGA

STEPHANIE KÄS – PXD CLUSTER ANALYSIS USING NNS - DPG-FRÜHJAHRSTAGUNG 2021
HOPFIELD NETWORK

- There is only 1 layer
  input layer = output layer
- Each node is connected to each node

*Six patterns are stored in a Hopfield network.*
Hopfield networks require binary data:

![Binary Data Table]

How can we represent other cluster properties (charge, shape, …)?
Hopfield networks require binary data:

How can we represent other cluster properties (charge, shape, …)?

Customized activation function
HOPFIELD NETWORK – CUSTOMIZED ACTIVATION FUNCTION

FIG. 11: Fit function for signal rate depending on (a) x-pixel-length $n_x$ and (b) y-pixel-length $n_y$. Fits were made for each parameter.

Function $f_i$
HOPFIELD NETWORK – CUSTOMIZED ACTIVATION FUNCTION

Activation function $f$

$\text{f}_1 \quad \text{f}_2 \quad \text{f}_3 \quad \ldots \quad \text{f}_{10}$

Individual functions $f_i$ are weighted by their rate range of possible domain.
HOPFIELD NETWORK – CUSTOMIZED ACTIVATION FUNCTION

Activation function $f$

$f_1$, $f_2$, $f_3$, ..., $f_{10}$

Individual functions $f_i$ are weighted by their rate range of possible domain.

Cluster shape added via Zernicke moments
HOPFIELD NETWORK – RESULTS

Efficiency

\[
\text{Efficiency} = \frac{\text{number of correctly identified signals}}{\text{total number of signals}}
\]

Efficiency 96.83%

Background rejection

\[
\text{Background rejection} = \frac{\text{number of correctly identified background}}{\text{total number of background}}
\]

BG rejection 98.49%
HOPFIELD NETWORK – RESULTS

Efficiency

\[
\text{Efficiency} = \frac{\text{number of correctly identified signals}}{\text{total number of signals}}
\]

Efficiency 96.83%

Background rejection

\[
\text{BG rejection} = \frac{\text{number of correctly identified background}}{\text{total number of background}}
\]

BG rejection 98.49%

Downsides?

• Requires preprocessing
• Static / not versatile
PROJECTS RELATED TO DATA SCIENCES AND AI…

- Classic MLP & Kohonen Maps
- Hopfield Networks, Voxels, Autoencoder…
- Networks on FPGA
VOXELS - SPHERICAL

- Ideal radius?
- Curse of dimensionality

92.38% coverage of data

Figure 19: Demonstrating what impact the input parameters have on finding the boundary and thus on the end result. J. Bilk & J. Budak, Detecting Clusters in Highdimensional Data
VOXELS

Spherical voxels

92.38% coverage of data

Cubic voxels

>99% coverage of data

J. Bilk & J. Budak, Detecting Clusters in Highdimensional Data
Problematic on overlapping data sets!
ONGOING PROJECTS

Elastic Matching

CNNs

…will be presented in 2022!
PROJECTS RELATED TO DATA SCIENCES AND AI….

Classic MLP & Kohonen Maps

Hopfield Networks, Voxels, Autoencoder…

Networks on FPGA
FPGA HARDWARE IMPLEMENTATION

Two approaches:

- Matrixmult. on FPGA (DSP Slices)
- High Level Synthesis (C++ on FPGA)

FOR FAST MATRIX OPERATIONS

FPGA Data Acquisition System of the PXD

Ongoing projects by

Peter Lehnhardt & Falk Zorn
PROJECTS RELATED TO DATA SCIENCES AND AI….

Classic MLP & Kohonen Maps

Hopfield Networks, Voxels, Autoencoder…

Networks on FPGA
WE HAVE A DATE TOMORROW!

18:00 - 18:15

Topic: Autoencoder

Identification of exotic highly ionising particles at the Belle II pixel detector using unsupervised autoencoders

By Katharina Dort
THANK YOU!

QUESTIONS?

STEPHANIE.KAES@PHYSIK.UNI-GIESSEN.DE
WHAT ARE HIGHLY IONIZING PARTICLES?

- Energy deposition of particles in matter increases with
- decreasing momentum of projectile
  Example: slow pions
- increasing mass of projectile
  Example: deuterons, heavy exotic particles
WHAT ARE HIGHLY IONIZING PARTICLES?

- Energy deposition of particles in matter increases with
  
  *decreasing momentum of projectile*
  
  **Example:** slow pions

- increasing mass of projectile
  
  **Example:** deuterons, heavy exotic particles

- non-Bethe-Bloch energy loss
  
  **Example:** magnetically charged particles
FEED-FORWARD NEURAL NETWORKS

- Supervised learning in order to separate HIPs from beam background*
- Implemented with PyTorch and trained on CPU and GPU
- Loss and accuracy is monitored during training (~8h)
- Cut on classification axis determines accuracy of neural network

*beam background
- Official mixed MC beam background
- Includes luminosity-dependent and beam-induced background

Feed-Forward Network Parameters
- 4 layers / 2 hidden
- > 50 nodes per layer
- ReLu Activation Function
- CrossEntropy Loss Function
- Stochastic Gradient Descent (SGD) Optimizer
- Batch Size: 256
- Learning Rate: 0.0001
- Momentum: 0.9

K. Dort, Belle II Pixel detector Cluster Analysis Using Neural Network Algorithms, Talk @ Vienna 13/01/20
HOPFIELD NETWORKS

- Recurrent binary neural network (associative memory)
- Network learns pattern associated with background/signal input vector -> stored in weight matrix
- Weight between neurons determines energy of the entire network
- Stable state is reached when energy of network is minimized (similar to spin-spin interaction in quantum mechanical many-body systems)
- Incomplete or distorted patterns are recognized
HOPFIELD NETWORK – USE OF ACTIVATION FUNCTION

\[
\begin{pmatrix}
  w_{11} & \cdots & w_{1n} \\
  \vdots & \ddots & \vdots \\
  w_{n1} & \cdots & w_{nn}
\end{pmatrix}
\]

- **Signal weight matrix**

\[
\begin{pmatrix}
  w_{11}^\text{signal} & \cdots & w_{1n}^\text{signal} \\
  \vdots & \ddots & \vdots \\
  w_{n1}^\text{signal} & \cdots & w_{nn}^\text{signal}
\end{pmatrix}
\]

- **Background weight matrix**

\[
\begin{pmatrix}
  w_{11}^\text{background} & \cdots & w_{1n}^\text{background} \\
  \vdots & \ddots & \vdots \\
  w_{n1}^\text{background} & \cdots & w_{nn}^\text{background}
\end{pmatrix}
\]

**Separately calculated**

\[
f \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad 1 - f
\]

\[
\begin{pmatrix}
  w_{11} & \cdots & w_{n1} \\
  \vdots & \ddots & \vdots \\
  w_{1n} & \cdots & w_{nn}
\end{pmatrix}
\]

**Weighted sum = full weight matrix**
HOPFIELD NETWORKS

- So far, Hopfield network is trained to store 4 patterns (2 background, 2 signal)
- Custom activation function (instead of binary activation) is used to feed additional information into the network
- Local, global properties and Zernike moments used in final analysis

Separation of magnetic monopoles from beam background

<table>
<thead>
<tr>
<th>3D</th>
<th>local prop.</th>
<th>global prop.</th>
<th>Zernike mom.</th>
<th>accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>63.0 %</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>76.5 %</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>95.7 %</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>97.7 %</td>
</tr>
</tbody>
</table>

Local prop - cluster size + cluster size in x/y + max. charge + cluster charge
Global prop - local prop + global position
Zernike mom. - global prop + Zernike moments

K. Dort, Belle II Pixel Detector Cluster Analysis Using Neural Network Algorithms, Talk @ Vienna 13/01/20
NEURAL NETWORKS ON FPGA

- Field programmable gate arrays (FPGAs) have prominent role in data acquisition for Belle II
- Highly parallel processing architecture make FPGAs ideal candidates for machine learning tasks
- DSP slices can be used for saving resources and speed up computation
  

- Communication with DSP slices and adaptation of neural network to FPGA architecture currently tested
- In future, neural networks could be directly integrated into the ONSEN which would also require online cluster finding on DHH or ONSEN (not implemented yet)
Hardware-Accelerated Neural Networks

- More performant and efficient neural network inference
- Accelerated AI application development on Xilinx FPGAs via High-Level Synthesis (HLS), FINN framework and Vitis AI development environment
- Attempt to allow real-time AI applications in time critical environments