Estimation of Electron Trajectory Parameters Using Supervised Machine Learning

Christoph Bernkopf,
Master Student

Problem

- Electron trajectory reconstruction algorithms operate at **limited accuracy**.

Problem - Kalman Filter

Seed

Look for compatible hits

Update state vector with candidates

Chi Square ranking

Take m best

Repeat

Problem - Gaussian Sum Filter

- Based on Kalman Filter.
- State vector is **convoluted** with each component of **Gaussian mixture**.

Solution

- At each iteration, **track information is lost**.
- **Neural Networks** can be used to regain parts of the lost information.

Track Parameters

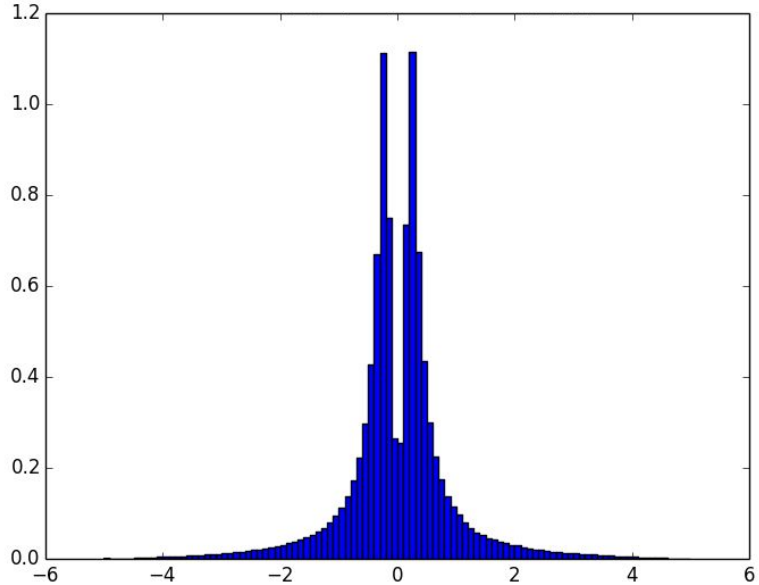
- signed inverse momentum **q/p**
- azimuth angle **phi**
- helix - beam line angle **theta** (also expressed as pseudorapidity)
- offset / impact parameter **d_0**
- impact parameter **d_z**

Track Data

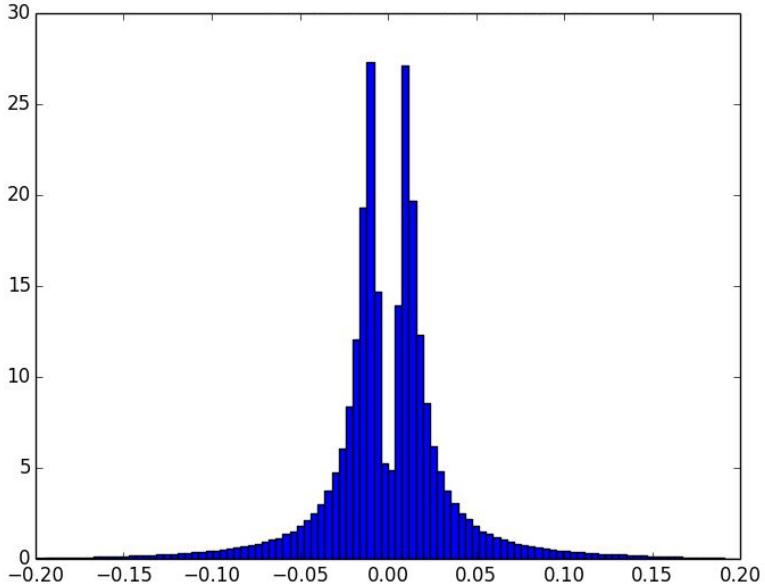
- 12 sets of **5 GSF reconstructed parameters**
- 1 set of **5 Geant4 simulated parameters**
- Data reconstruction & extraction from Root with **CMSSW** (C++)
- 2,3M Events

Example: Parameter q/p

Inputs (Reconstructed)



Targets (Simulated)



Baseline

- To compare with Machine Learning results.
- Calculated using the standard deviation of the residuals calculated using the mean.

$\frac{q}{p}$	$\frac{dx}{dz}$	$\frac{dy}{dz}$	x	y
0.0264	0.0102	0.0149	0.1479	0.3075

- Trace of the covariance matrix: 0.1175

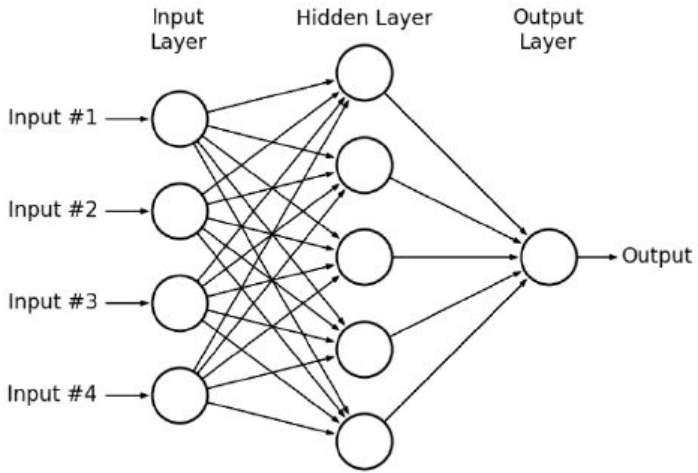
Neural Networks

output = input * weights

error = target - output

new_weights =
weights - error* learning_rate

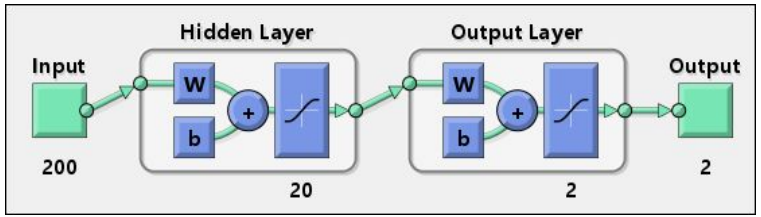
repeat



Libraries

TensorFlow

MATLAB



Implementation

ROOT

CMSSW Associator Reconstruction (on LXPLUS / Grid)

Extraction to training machine

Preprocessing

Neural Network

Experimentation

- Various activation functions.
- Dropout & Max Pooling.
- SGD, SGDm, Nesterov, Levenberg-Marquardt and more.
- Regularisations (L1 & L2)
- Hyperparameter Optimisation (Ada-Grad/Delta)
- Data Formatting (Conditioning, 3 or Single Parameter, etc.)

Results

1.

Events in Million	0 - 0.5	0.5 - 1	1 - 1.64
Correct Results	94.67 %	93 %	100 %

2.

Parameter	1	2	3	Average
Improvement	0.4682	0.4544	0.4386	0.4537

3.

Residual	Trace of the covariance matrix
Mixture Mean	0.1175
Residuals Test 26	$1.3386 * 10^{-4}$

Results

- Best algorithm: **Levenberg-Marquard** (damped least-squares).
- Network **overfits** after ~**20-50 epochs** (~20 minutes).
- Only parameters 1-3 were trained as targets of 4 & 5 == Zero.

Outlook

- Deep Learning.
- Library Switch: PyTorch, Fast.ai, Keras.
- GPU Servers.