Using Neural Networks to Separate Signal from Background with Real Missing Transverse Energy

APS NW Meeting 2016
Jannicke Pearkes
UBC Engineering Physics
Supervised by: Dr. Robert Kowalewski
Motivation: High Level Trigger

ATLAS Detector

Reduces data rate by factor of $\sim 10^5$!

Lots of data!

Trigger

Storage

Missing Transverse Momentum (MET) Trigger

Much less data…
Motivation: Missing Transverse Momentum

When a particle escapes the detector undetected

MET: Negative vector sum of transverse energies
Sum ET: Scalar sum of transverse energies

Neutrino / Dark Matter?
Monte Carlo Samples

Signal: Associated Higgs Production, Pythia 8+EvtGen
Background: Minimum bias (random trigger), Pythia 8

Image: CDF Experiment
MET Threshold

70 GeV Cut: 99% background rejection but only 41% efficient
Can We Use Other Variables to Do Better?

High Level Inputs

- **Total Transverse Energy [GeV]**
  - Events vs. Total Transverse Energy [GeV]
  - Red: Signal (2vqbb), Blue: Background (Minbias)
  - ATLAS Work in progress

- **ET Maximum Calorimeter Tower [GeV]**
  - Events vs. ET Maximum Calorimeter Tower [GeV]
  - Red: Signal (2vqbb), Blue: Background (Minbias)
  - ATLAS Work in progress

Low Level Inputs

- **Transverse Energy [GeV]**
  - Eta vs. Phi [rad]
  - ATLAS Work in progress
  - Single Background Event
  - Single Signal Event

- **Transverse Energy [GeV]**
  - Eta vs. Phi [rad]
  - ATLAS Work in progress
  - Single Background Event
  - Single Signal Event
Neural Networks Trained on High Level Inputs

5 Different High Level Input Variables:

- MET: Missing Transverse Momentum
- Sum ET: Total Transverse Momentum
- ET Max Tower: Energy deposited in the calorimeter tower with the maximum energy
- Eta Max Tower: Eta of the calorimeter tower with the maximum energy
- NPV: Number of primary vertices

Implemented in Theano

http://deeplearning.net/software/theano/
Results: High Level Inputs

ROC Curve obtained on High Level Inputs

ATLAS Work in progress
Results: High Level Inputs

ROC Curve obtained on High Level Inputs

**ATLAS** Work in progress

- MET
- MLP
- Sum ET
- Max Tower
What if we use low level inputs?

### Signal event

![Signal event](image)

**ATLAS Work in progress**

**Transverse Energy [GeV]**

**Entries**

- Mean x: 0.3669
- Mean y: 0.2119
- RMS x: 1.708
- RMS y: 1.087
What if we use low level inputs?

Background event

<table>
<thead>
<tr>
<th>Phi [rad]</th>
<th>calo_towers_y_prime</th>
<th>Entries</th>
<th>Mean x</th>
<th>Mean y</th>
<th>Std Dev x</th>
<th>Std Dev y</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.125</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>-2.5</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>-2</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>-1.5</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>1.5</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>2.5</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>594</td>
<td>0.1365</td>
<td>0.02916</td>
<td>1.728</td>
<td>1.406</td>
</tr>
</tbody>
</table>

Transverse Energy [GeV]
Pre-Processing

Signal

Background

Raw

Processed

Sum over 5000 events

Align MET vector by translating along Phi, then flip about Eta=0 and Phi=0 to position the tower with the most energy deposited in upper right quadrant
Convolutional Neural Network

Image: Theano
Neural Network Performance

**ROC Curve obtained with Convolutional Network Comparison**

**ATLAS** Work in progress

- **MET**
- **MLP (HIGH)**
- **COVNET**
Signal Event Animation

Training iteration: 99 Prediction: 0.429 Truth: 1.0

*ATLAS* Work in progress
Background Event Animation

*ATLAS* Work in progress
Conclusion

• Performed studies of multivariate methods on high and low level inputs
• Area under curve obtained with neural networks is higher than with MET threshold
• However, neural networks do not necessarily perform as well at the high rejection rates required for the trigger
• More work needs to be performed to determine suitability for trigger applications
Thank you!
Preprocessing

**Phi [rad]**

**Eta**

**calo_towers_phi**

Entries: 276
Mean x: 0.1093
Mean y: 0.3303
RMS x: 1.74
RMS y: 1.447

**calo_towers**

Entries: 276
Mean x: 0.07075
Mean y: 0.3303
RMS x: 1.744
RMS y: 1.447

*ATLAS Work in progress*
Preprocessing

ATLAS Work in progress
Preprocessing

\[ \Phi \] (rad)

\[ \eta \]

\[ \text{calo} \text{towers}_{\phi} \]

Entries

Mean \text{x}

Mean \text{y}

RMS \text{x}

RMS \text{y}

\[ \Delta \phi \]

\[ \text{ATLAS Work in progress} \]
ROC Curves: High Level Inputs

ATLAS Work in progress
Weights Animation

Network Weights - Training Iteration: 99 Layer: 1

*ATLAS* Work in progress
Effect of Training on Weights

Layer 0

Before Training

After Training

Layer 1

Before Training

After Training

ATLAS Work in progress
Effect of Training on Signal Event

Training iteration: 99 Prediction: 0.450 Truth: 1.0

Training iteration: 11499 Prediction: 0.990 Truth: 1.0
Effect of Training on Background Event

ATLAS Work in progress

ATLAS Work in progress
Motivation: Missing Transverse Energy

\[ E_T^{\text{miss}} = \sqrt{\left( \sum_i E_x \right)^2 + \left( \sum_i E_y \right)^2} \]

Neutrino? Dark Matter?
Details:

• Signal Monte Carlo samples generated with Pythia 8+EvtGen, background Monte Carlo samples generated with Pythia 8

• Used 100,000 signal and 100,000 background events for training, validation and test sets (80-10-10 % split).
Convolution Step

Center element of the kernel is placed over the source pixel. The source pixel is then replaced with a weighted sum of itself and nearby pixels.

Source pixel

Convolution kernel (emboss)

New pixel value (destination pixel)

(4 x 0)
(0 x 0)
(0 x 0)
(0 x 0)
(0 x 1)
(0 x 1)
(0 x 0)
(0 x 1)

+ (-4 x 2)
-8

Image: Mac Developer Library
Identity

\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
\end{bmatrix}
\]

Sharpen

\[
\begin{bmatrix}
0 & -1 & 0 \\
-1 & 5 & -1 \\
0 & -1 & 0 \\
\end{bmatrix}
\]

Edge Detection

\[
\begin{bmatrix}
1 & 0 & -1 \\
0 & 0 & 0 \\
-1 & 0 & 1 \\
\end{bmatrix}
\]

Blur

\[
\frac{1}{16}
\begin{bmatrix}
1 & 2 & 1 \\
2 & 4 & 2 \\
1 & 2 & 1 \\
\end{bmatrix}
\]

Image: OpenCV Documentation