Combining Unsupervised Anomaly Detection Techniques In BSM Physics Searches At The LHC

Adam Leinweber

University of Adelaide

adam.leinweber@adelaide.edu.au

February 10, 2021

A Brief Introduction

A Brief Introduction

- Recent searches for BSM physics at the Large Hadron Collider have been unsuccessful in detecting any new physics.

A Brief Introduction

- Recent searches for BSM physics at the Large Hadron Collider have been unsuccessful in detecting any new physics.
- Partially because we do not know the particular standard model extension that nature has chosen.

A Brief Introduction

- Recent searches for BSM physics at the Large Hadron Collider have been unsuccessful in detecting any new physics.
- Partially because we do not know the particular standard model extension that nature has chosen.
- We propose using various unsupervised anomaly detection methods in order to identify signal regions of interest for further analysis. It is important to use unsupervised methods as we wish to perform analyses in a "signal agnostic" way.

The Method

The Method

We want to ask "how anomalous is a given event compared to the SM background?". The method is as follows:

The Method

We want to ask "how anomalous is a given event compared to the SM background?". The method is as follows:

1. Choose a final state and model the relevant background processes.

The Method

We want to ask "how anomalous is a given event compared to the SM background?". The method is as follows:

1. Choose a final state and model the relevant background processes.
2. Train the anomaly detection techniques on a subset of the background model.

The Method

We want to ask "how anomalous is a given event compared to the SM background?". The method is as follows:

1. Choose a final state and model the relevant background processes.
2. Train the anomaly detection techniques on a subset of the background model.
3. Feed through the remainder of the background model set and the data that you wish to analyse, obtaining anomaly scores for each event.

The Method

We want to ask "how anomalous is a given event compared to the SM background?". The method is as follows:

1. Choose a final state and model the relevant background processes.
2. Train the anomaly detection techniques on a subset of the background model.
3. Feed through the remainder of the background model set and the data that you wish to analyse, obtaining anomaly scores for each event.
4. Normalise and combine these anomaly scores for each event. In this presentation I will show AND, OR, sum, and product combinations.

The Method

We want to ask "how anomalous is a given event compared to the SM background?". The method is as follows:

1. Choose a final state and model the relevant background processes.
2. Train the anomaly detection techniques on a subset of the background model.
3. Feed through the remainder of the background model set and the data that you wish to analyse, obtaining anomaly scores for each event.
4. Normalise and combine these anomaly scores for each event. In this presentation I will show AND, OR, sum, and product combinations.
5. This new combined anomaly score can then be used in further analysis.

Anomaly Detection Algorithms

The anomaly scores used for these combinations are generated using the following method:

Anomaly Detection Algorithms

The anomaly scores used for these combinations are generated using the following method:

1. Train a variational autoencoder (VAE) on background events and use the reconstruction loss as an anomaly score.

Anomaly Detection Algorithms

The anomaly scores used for these combinations are generated using the following method:

1. Train a variational autoencoder (VAE) on background events and use the reconstruction loss as an anomaly score.
2. The VAE creates a dimensionally reduced space as part of its construction. This is referred to as the "latent space".

Anomaly Detection Algorithms

The anomaly scores used for these combinations are generated using the following method:

1. Train a variational autoencoder (VAE) on background events and use the reconstruction loss as an anomaly score.
2. The VAE creates a dimensionally reduced space as part of its construction. This is referred to as the "latent space".
3. We can pass events into the VAE and obtain a dimensionally reduced representation of each event, known as the latent space representation.

Anomaly Detection Algorithms

The anomaly scores used for these combinations are generated using the following method:

1. Train a variational autoencoder (VAE) on background events and use the reconstruction loss as an anomaly score.
2. The VAE creates a dimensionally reduced space as part of its construction. This is referred to as the "latent space".
3. We can pass events into the VAE and obtain a dimensionally reduced representation of each event, known as the latent space representation.
4. Training in the latent space significantly improves the performance of other algorithms, partially because the number of dimensions is lower.

Anomaly Detection Algorithms

The anomaly scores used for these combinations are generated using the following method:

1. Train a variational autoencoder (VAE) on background events and use the reconstruction loss as an anomaly score.
2. The VAE creates a dimensionally reduced space as part of its construction. This is referred to as the "latent space".
3. We can pass events into the VAE and obtain a dimensionally reduced representation of each event, known as the latent space representation.
4. Training in the latent space significantly improves the performance of other algorithms, partially because the number of dimensions is lower.
5. We then train three other algorithms on these latent space representations. The algorithms used are an Isolation Forest, a Gaussian Mixture Model, and a Static Autoencoder.

Anomaly Detection Algorithms

The anomaly scores used for these combinations are generated using the following method:

1. Train a variational autoencoder (VAE) on background events and use the reconstruction loss as an anomaly score.
2. The VAE creates a dimensionally reduced space as part of its construction. This is referred to as the "latent space".
3. We can pass events into the VAE and obtain a dimensionally reduced representation of each event, known as the latent space representation.
4. Training in the latent space significantly improves the performance of other algorithms, partially because the number of dimensions is lower.
5. We then train three other algorithms on these latent space representations. The algorithms used are an Isolation Forest, a Gaussian Mixture Model, and a Static Autoencoder.
6. After normalising to uniform background density, we have four anomaly scores for each event.

Combination Methods

We now combine these four algorithms in various ways. Here x_i denotes a vector containing the 4 anomaly scores for a given event.

Combination Methods

We now combine these four algorithms in various ways. Here x_i denotes a vector containing the 4 anomaly scores for a given event.

- Logical AND: $min(x_i)$

Combination Methods

We now combine these four algorithms in various ways. Here x_i denotes a vector containing the 4 anomaly scores for a given event.

- Logical AND: $min(x_i)$
- Logical OR: $max(x_i)$

Combination Methods

We now combine these four algorithms in various ways. Here x_i denotes a vector containing the 4 anomaly scores for a given event.

- Logical AND: $min(x_i)$
- Logical OR: $max(x_i)$
- Sum: $\sum_i x_i$

Combination Methods

We now combine these four algorithms in various ways. Here x_i denotes a vector containing the 4 anomaly scores for a given event.

- Logical AND: $min(x_i)$
- Logical OR: $max(x_i)$
- Sum: $\sum_i x_i$
- Product: $\prod_i x_i$

Some Results

The Signals

To assess the performance of these algorithms we test them on a handful of interesting signal models.

The Signals

To assess the performance of these algorithms we test them on a handful of interesting signal models.

- Supersymmetric gluino pair production. $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$

The Signals

To assess the performance of these algorithms we test them on a handful of interesting signal models.

- Supersymmetric gluino pair production. $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$
- Supersymmetric stop quark pair production. $\tilde{t}_1 \to t\tilde{\chi}_1^0$

The Signals

To assess the performance of these algorithms we test them on a handful of interesting signal models.

- Supersymmetric gluino pair production. $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$
- Supersymmetric stop quark pair production. $\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0$

Gluino signals 01-06 have gluino masses of 1-2.2 TeV in 200 GeV increments, and neutralino masses of 1 GeV. Stop signals 01-04 have stop masses of 220, 330, 400, and 800 GeV. Stop signal 01 has a neutralino mass of 20 GeV while the others have a mass of 100 GeV.

The Signals

To assess the performance of these algorithms we test them on a handful of interesting signal models.

- Supersymmetric gluino pair production. $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$
- Supersymmetric stop quark pair production. $\tilde{t}_1 \to t\tilde{\chi}_1^0$

Gluino signals 01-06 have gluino masses of 1-2.2 TeV in 200 GeV increments, and neutralino masses of 1 GeV. Stop signals 01-04 have stop masses of 220, 330, 400, and 800 GeV. Stop signal 01 has a neutralino mass of 20 GeV while the others have a mass of 100 GeV.

$$m_{eff} = E_T^{miss} + \sum p_T^{jets} \tag{1}$$

The Signals

To assess the performance of these algorithms we test them on a handful of interesting signal models.

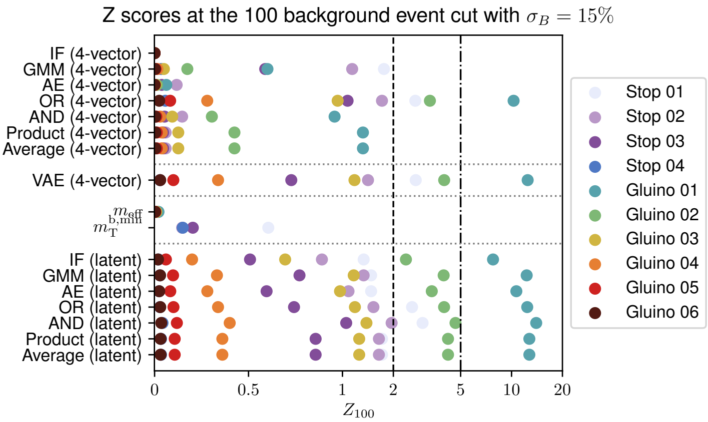
- Supersymmetric gluino pair production. $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$
- Supersymmetric stop quark pair production. $\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0$

Gluino signals 01-06 have gluino masses of 1-2.2 TeV in 200 GeV increments, and neutralino masses of 1 GeV. Stop signals 01-04 have stop masses of 220, 330, 400, and 800 GeV. Stop signal 01 has a neutralino mass of 20 GeV while the others have a mass of 100 GeV.

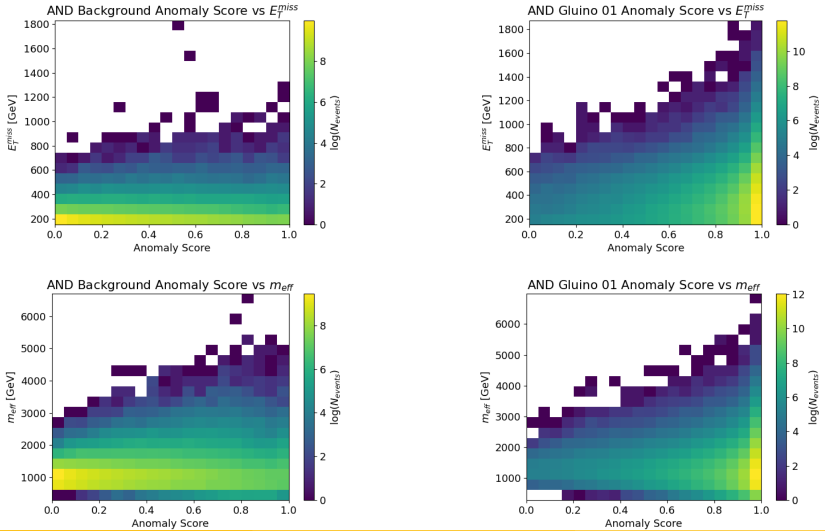
$$m_{eff} = E_T^{miss} + \sum p_T^{jets} \tag{1}$$

$$m_t^{b,min} = \sqrt{2p_T^b E_T^{miss}[1 - cos\Delta\phi(\boldsymbol{p}_T^b, \boldsymbol{p}_T^{miss})]} \tag{2}$$

Z-Scores For Each Algorithm With a 15% Systematic



2D Correlation Plots Between AND Anomaly Score and Physical Variables For Gluino 01



Summary

- Standard LHC search techniques have been unable to detect any new physics, partially due to the fact that you need to choose a signal to search for.

Summary

- Standard LHC search techniques have been unable to detect any new physics, partially due to the fact that you need to choose a signal to search for.
- We, have developed a technique that is able to differentiate signal from background while making minimal assumptions about the signal.

Summary

- Standard LHC search techniques have been unable to detect any new physics, partially due to the fact that you need to choose a signal to search for.
- We, have developed a technique that is able to differentiate signal from background while making minimal assumptions about the signal.
- This technique provides a powerful tool that can be used to determine signal regions to explore further.

Summary

- Standard LHC search techniques have been unable to detect any new physics, partially due to the fact that you need to choose a signal to search for.
- We, have developed a technique that is able to differentiate signal from background while making minimal assumptions about the signal.
- This technique provides a powerful tool that can be used to determine signal regions to explore further.
- Future work involves expanding this technique with more algorithms, optimising the combination methods, and applying this to more datasets.

End