



Search For Displaced Leptons in $\sqrt{s} = 13.6$ TeV pp Collisions with the ATLAS Detector

REU Program at Columbia University - Nevis Labs

Jessi Jha¹

¹*University of Pennsylvania, Philadelphia, PA, 19104*

August 5, 2023

Abstract

A search is presented for displaced leptons originating from the decay of long-lived scalar supersymmetric lepton partners (sleptons). This analysis uses the Run 3 data set of proton-proton collisions taken at the LHC at $\sqrt{s} = 13.6$ in 2022 and recorded by the ATLAS detector. By using reconstructed trajectories of electrons and photons in the ATLAS liquid argon calorimeter, this analysis searches for electron and photon signals that have a delayed time of arrival at the calorimeter. We use a gauge-mediated supersymmetry breaking model in which pair-produced sleptons decay into their standard model partner lepton and a gravitino which escapes direct detection and gives rise to missing transverse momentum. This paper outlines the implementation of a boosted decision tree machine learning algorithm that can be used to distinguish standard model background from the displaced leptons signal. The model is trained to separate standard model background from simulated Monte Carlo events based on the gauge-mediated supersymmetry model we are considering in this analysis.

Contents

1	Introduction	3
1.1	The Standard Model	3
1.2	Supersymmetry	5
1.3	Instrumentation	6
1.3.1	The Large Hadron Collider	6
1.3.2	The ATLAS Detector	7
2	The Search For Displaced Leptons	9
2.1	Motivation	9
2.2	Boosted Decision Trees in the Displaced Leptons Search	10
2.2.1	BDT Introduction	10
2.2.2	Signal Regions	11
2.2.3	Preselections	11
2.2.4	1eBDT	12
2.2.5	2eBDT	15
2.3	Event Level vs Displaced Object Tagger BDT	17
2.3.1	Applying the 1eBDT on a 2e Signal Region	18
2.3.2	Signal and Background Efficiency Comparisons	18
3	Conclusions and Next Steps	20
4	Acknowledgements	20

1 Introduction

1.1 The Standard Model

Decades of study on the nature of matter have led to the understanding that there are four fundamental forces governing our universe, and all matter can be broken down into building blocks known as fundamental particles. The Standard Model represents our best theory encapsulating the interactions between these fundamental particles and three of the four fundamental forces. CERN (2023b)

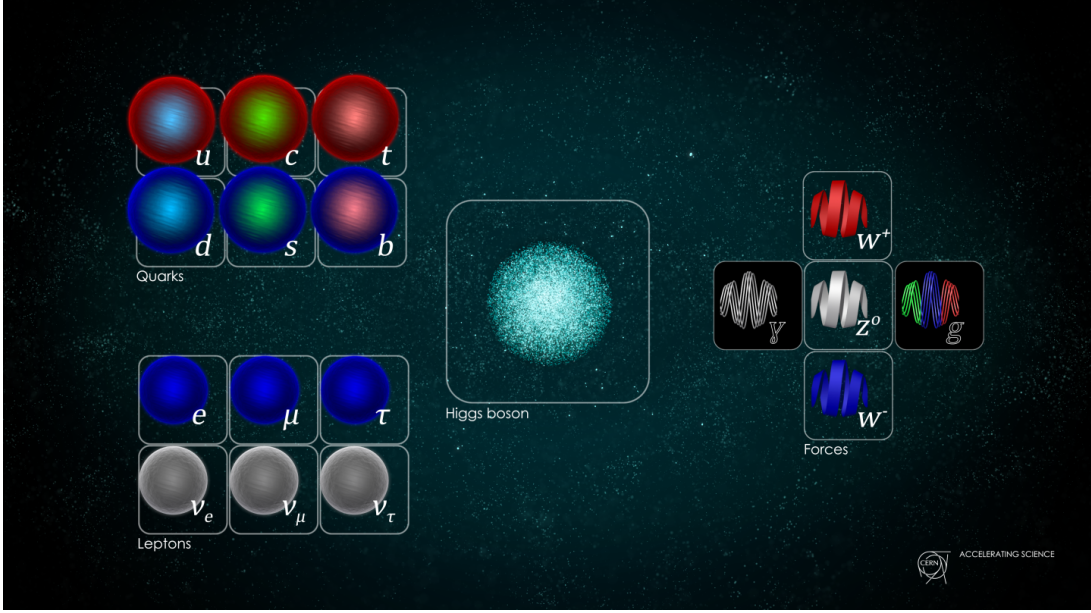


Figure 1: Particles of the Standard Model (Image: Daniel Dominguez/CERN)

Our universe is divided into two groups of particles: matter particles called fermions and force particles called bosons. There are four elementary forces: electromagnetic force, strong force, weak force, and gravity. The fundamental theory of electromagnetism, or Quantum Electrodynamics (QED), is a field theory that describes the interactions between electrically charged particles and light (photons). The strong force, which manifests as the binding of protons and neutrons in the nucleus, is described by the paradigm of Quantum Chromodynamics (QCD). Fundamental particle interactions also are mediated by the weak force, which is responsible for nuclear fusion properties. Lastly, gravity, the force most noticeable in our macroscopic world, is an attractive force that dictates the structure of the large-scale universe. (Thomson, 2013; CERN, 2023b)

The standard model theory incorporates three forces into its mathematical paradigm: the strong, weak, and electromagnetic forces. However, gravity has not been integrated into the Standard Model framework. Each standard model force can be described by a unique quantum field theory (QFT). These QFTs correspond to the interactions of a spin 1 force-carrying particle called a “gauge boson.” The transfer of virtual spin 1 gauge bosons corresponds to the angular momentum transfer between matter particles, which then manifests as a force. Gauge bosons, or force mediators, are governed by Bose-Einstein statistics and can be described using the same quantum numbers. All bosons have integer multiple spins, however, gauge bosons all have spin 1. There are five discovered bosons, including the W boson, the Z boson, the photon, the gluon, and the Higgs boson. The W^+ and W^- gauge bosons are carriers of the weak force and are responsible for nuclear β -decay that produces

radioactive isotopes and nuclear fusion. The Z boson is another carrier of the weak force and is the gauge boson responsible for weak current (the weak analog to electric charge current). The gluon is responsible for the strong force, and the photon is the electromagnetic force carrier; both of these gauge bosons are massless. The properties of the standard model forces are shown in 1. Finally, the Higgs Boson, which has spin 0, imparts mass to the particles that interact with the Higgs field. Without the Higgs Boson, all matter would be massless, and each fundamental particle would propagate at the speed of light. The Higgs boson has a mass of approximately 125 GeV. In QFT terms, the Higgs boson can be thought of as an interaction, or excitation, of the non-zero scalar Higgs field, which gives mass to initially massless particles. (Thomson, 2013; CERN, 2023b)

Force	Strength	Boson	Spin	Mass/GeV
Strong	1	Gluon (g)	1	0
Electromagnetism	10^{-3}	Photon (γ)	1	0
Weak	10^{-8}	W Boson (W^{pm})	1	80.4
Weak	10^{-8}	Z Boson (Z)	1	Row 4, 91.2

Table 1: *Particle Properties: Gauge Bosons. The strengths listed are approximate values for two fundamental particles at a distance of $10^{-15}m$. (Thomson, 2013)*

Matter particles, known as fermions, have dynamics that follow the Dirac relativistic wave equation. The Dirac wave equation states that for each fermion, there is an antiparticle with the same mass characteristics but with a different charge. Additionally, Fermions adhere to the Pauli Exclusion Principle, meaning that two fermions cannot possess the same quantum numbers. In other words, two fermions cannot occupy the same place simultaneously. All elementary fermions have a spin of $1/2$, an intrinsic property measured in the same units as angular momentum. There are two groups of fermions known as leptons and quarks; there are six leptons and six quarks, making for a total of 12 fermions. Each fermion experiences weak force interactions, however, neutrinos, which are electrically neutral particles, are the only fermions that do not participate in QED. Only quarks carry a color charge (the QCD equivalent of an electric charge) and interact with the strong force. Leptons and quarks are divided into three generations of particles, which are differentiated only by their masses. The lightest and most stable particles constitute the majority of matter in the universe; heavier particles decay into lighter ones. Since larger particles rapidly decay into smaller ones, scientists use high-energy collisions to study the properties of higher-mass particles. The three generations of quarks consist of the “up” and “down” quarks, the “charm” and “strange” quarks, and the “top” and “bottom” quarks. Due to the nature of QCD and strong force interaction, quarks cannot exist independently and are always bound to each other in configurations determined by a unique property called “color.” Configurations of quarks are composed such that each quark arrangement results in a colorless final object; these colorless arrangements are called hadrons (the proton and neutron are examples of hadrons). Lepton generations include electrons and electron neutrinos, muons and muon neutrinos, and tau and tau neutrino particles. Electrons, muons, and tau particles each carry a charge, whereas the electron neutrinos, muon neutrinos, and tau neutrinos are neutral with very small masses. Leptons do not require binding together and can exist as single particles. The matter content of Fermions is summarized in Tab. 2. (Thomson, 2013; CERN, 2023b)

Leptons				Quarks		
	Particle	Q	Mass/GeV	Particle	Q	Mass/GeV
First Generation	Electron	-1	0.000511	Up Quark	-1/3	0.003
	Electron Neutrino	0	$> 10^{-9}$	Down Quark	+2/3	0.005
Second Generation	Muon	-1	0.106	Strange Quark	-1/3	0.1
	Muon Neutrino	0	$> 10^{-9}$	Charm Quark	+2/3	1.3
Third Generation	Tau	-1	1.78	Bottom Quark	-1/3	4.5
	Tau Neutrino	0	$> 10^{-9}$	Top Quark	+2/3	174

Table 2: Particle Properties: Leptons and Quarks (Thomson, 2013)

1.2 Supersymmetry

Although the standard model has guided our understanding of the interactions between elementary particles, physicists acknowledge that this theory is incomplete. Supersymmetry (SUSY), an extension of the standard model, was developed to fill the gaps in our understanding. The supersymmetric theory suggests that there are partner particles to the standard model, called sparticles, which could “cure” several issues within the Standard Model. Most notably, SUSY phenomenology has the potential to address the “lightness” of the Higgs boson mass, it proposes well-motivated dark matter candidates and it lends itself to an elegant connection, or “unification” of the strong, weak, and electromagnetic interactions. (CERN, 2023c; Charitos, 2016; O’Keefe, 2020)

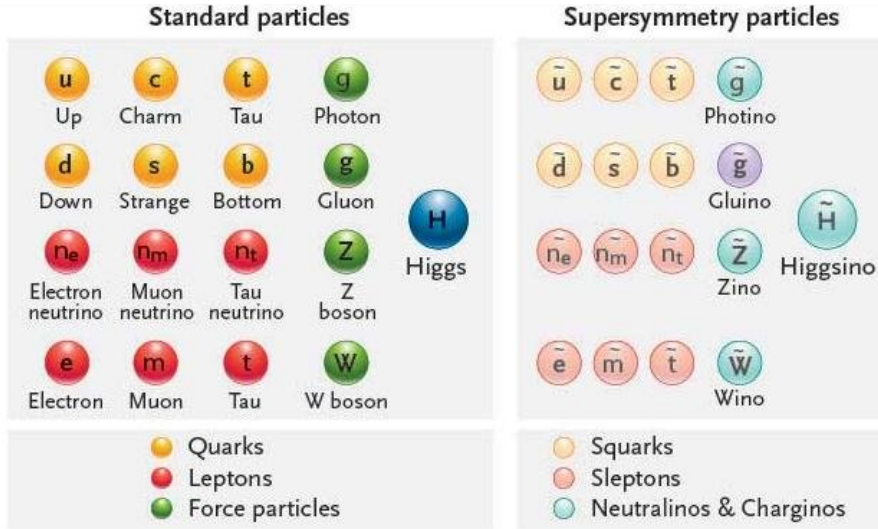


Figure 2: Supersymmetric Particles (Image: Daniel Murnane/CERN)

The mass of the Higgs Boson is approximately 125 GeV or about 130 times a proton mass. However, this mass is deemed extremely light, as the Higgs mass is contrived using two inputs - the “bare” Higgs mass, which we have not observed, and “quantum corrections”, which are the contributions made by other standard model processes. The value of the quantum corrections is measured to be largely negative - in the magnitude of negative 10^{18} GeV. This suggests that the “bare” Higgs mass is nearly opposite to the value of the quantum corrections - an unsettling result for many particle physicists. These large “radiative” corrections, which present a weakness of the standard model, can be explained by new symmetries - or in other words, SUSY mechanisms. Under

certain circumstances, SUSY presents a suitable dark matter candidate. Under “R-Parity” conserving SUSY models, standard model particles and SUSY particles are produced in pairs, implying that the lightest SUSY particle (LSP) cannot decay into standard model particles. A stable, neutral, and weakly interacting LSP could act as a great dark matter candidate - moreover, it’s possible that this SUSY particle is produced abundantly enough to match the required amount of dark matter that has been determined from previous experimental data. SUSY also has the potential to contribute to a “grand unification” of stand, weak, and electromagnetic forces. Although current observations suggest that the electroweak and strong forces come close, but do not quite, converge at high energies, SUSY manifestations at energies of around 1 TeV could provide modifications that lead to the realization of grand unification. (CERN, 2023c; Charitos, 2016; O’Keefe, 2020)

In nature, SUSY would manifest as each standard model particle having a nearly identical supersymmetric particle partner; supersymmetric particles are hypothesized to differ from standard model particles by their spin values. For example, leptons (which have spin $1/2$) would have a supersymmetric spin 0 “slepton” and for every spin $1/2$ quark there would exist a supersymmetric spin 0 “squark”. The theorized SUSY bosons would act similarly, with the exception that these particles would be a combination of states that include the Higgs SUSY partner. Due to observed natural phenomena, such as the electron being the lightest possible scalar spin $1/2$ partner, we know that SUSY particles cannot have equal masses to their partner particles. This means that SUSY must be a “broken” symmetry; all SUSY particles are theorized to be much heavier than their partner particles. The mode by which the symmetry “breaks” differs between SUSY models. However, within supersymmetric models, various parameters, such as the mass spectrum and lifetime of the particles, are unknown. These mass and lifetime parameters govern the nature of SUSY particle decay methods, which are used to reconstruct the SUSY signature. Thus, researchers must test a wide range of experimental signatures in the search for SUSY particles. (CERN, 2023c; Charitos, 2016; O’Keefe, 2020)

1.3 Instrumentation

1.3.1 The Large Hadron Collider

The Large Hadron Collider (LHC), located at CERN (European Council for Nuclear Research) in Geneva, Switzerland, is the world’s largest, most powerful particle accelerator. As the name suggests, the LHC collides hadrons, or more specifically, protons and heavy lead ions. It is structured as a 27-kilometer loop of superconducting magnets that utilizes accelerators used to increase the energy of the particles during runs. The LHC starts off by using electric fields to strip hydrogen atoms of their electrons to produce single protons. This process is monitored to produce particles in “bunches,” allowing for a 25-nanosecond “break” between collisions. After the protons are produced, metallic chambers, or radio frequencies (RF) cavities, resonate radio waves to transmit energy to the particle bunches. To ensure that the particles do not “hit” stray gas molecules within the accelerator, the beam is contained inside a vacuum called the beam pipe. Various magnets within the accelerator act as guides to bend particles in the direction of the curved loop, while other magnets serve to focus the beam and gather particles closer. The guiding magnets produce massive field strengths of 8.3 Teslas which are maintained by superconducting magnets (kept at temperatures of -271.3°C and cooled by systems of super fluid helium). Finally, the beams within the LHC collide at one of four locations (particle detectors) within the ring. These detectors correspond to different experiments, which include ATLAS, CMS, ALICE, and LHCb; there are also three smaller, more specialized detectors including LHCf, TOTEM, and MoEDAL. The primary focus of the LHC is to investigate unsolved aspects of particle physics and use high-energy collision “probe” answers from the most fundamental

particles of our universe. Fig. 3 shows the layout of the LHC complex, complete with the path of the accelerated particles and each of the detectors. CERN (2023a)

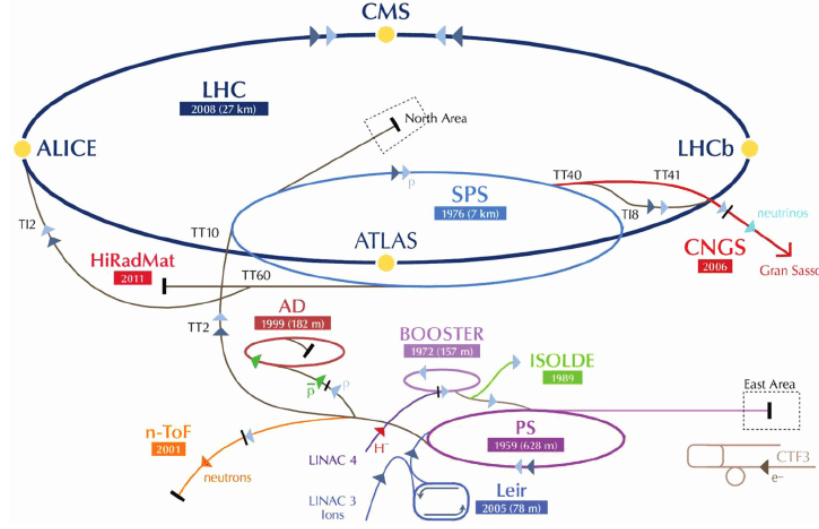


Figure 3: The LHC complex and detectors. (Thomason et al., 2013)

1.3.2 The ATLAS Detector

The ATLAS Detector, or the A Toroidal LHC ApparatuS, is a general-purpose detector, with a variety of uses, ranging from precision measurements of the Higgs Boson Mass to searches for beyond standard model (BSM) physics. When sufficiently high-energy particles collide, the energy of the interaction can be converted into matter - more specifically, this energy can be used to produce the fundamental particles that existed in the early universe. The goal of the ATLAS detector is to detect these high-energy particles by recording values such as particle energy, trajectory, and momentum. However, due to the volume of collisions that occur within the detector - about 1.7 billion proton-proton collisions per second - ATLAS does not record each measurement in memory. Instead, a Trigger and Data Acquisition system uses tailored algorithms to ensure that optimal data-collection of events with interesting characteristics. There are six unique subsystems within the ATLAS detector, each of which plays a critical role in correctly identifying and measuring particle data. Fig 4 displays four of the main systems; ATLAS is designed such that each detector wraps concentrically around the point of collision. (ATLAS Collaboration, 2023b,0,0)

The four main subsystems include the inner track detector, a superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer. These detectors are used to reconstruct signals left by particles that interact with each system. (ATLAS Collaboration, 2023b,0)

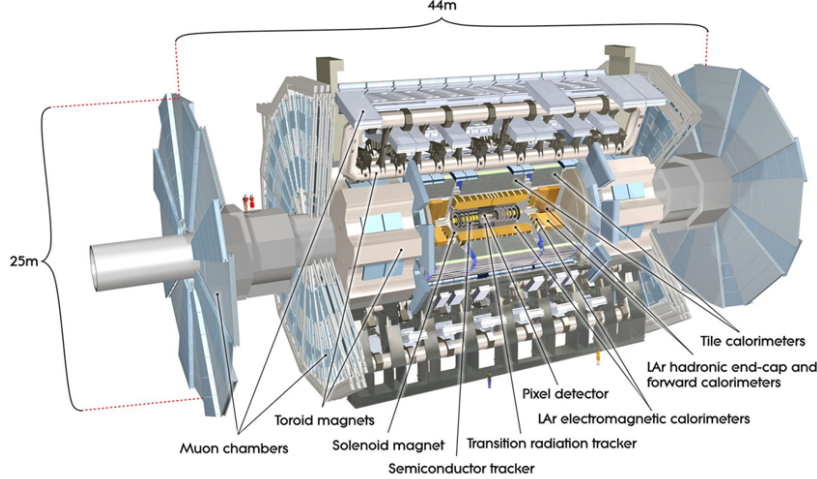


Figure 4: The ATLAS Detector featuring the Inner Detector, Calorimeter, Muon Spectrometer, and Magnet System.

1. The Inner Detector: The Inner Detector is the closest subsystem to the initial decay products of the collisions. It consists of the Semiconductor Tracker, Pixel Detector, and Transition Radiation Tracker, which work together to measure the direction, momentum, and charge (of electrically-charged particles) that result from each proton-proton collision.(ATLAS Collaboration, 2023a)
2. Calorimeter: The Calorimeter is used to measure the energy that each particle loses as it passes through the detector. The calorimeter absorbs the energy of particles produced in a collision, thereby stopping particles from passing through. The two components of the Calorimeter include the Liquid Argon (LAr) Calorimeter, which measures the energy of electrons and photons as they interact with matter, and the Tile Hadronic Calorimeter, which measures the energy of Hadrons as they interact with atomic nuclei. The LAr Calorimeter, which surrounds the Inner Detector, features a layer of metal that absorbs incoming particles and converts them into “showers” or lower energy particles. These lower-energy “showers” produce an electric current by ionizing the liquid argon located between the metal layers. Hadrons, which do not deposit all of their energy within the LAe Calorimeter, will deposit their remaining energies in the Tile Hadronic Calorimeter. Particles will hit the steel layers of the calorimeter, leading to the production of more showers. These showers are in turn converted to an electric current proportional to the initial energy of the hadron. (ATLAS Collaboration, 2023a)
3. Muon Spectrometer: The outermost (and largest) layer of the detector consists of muon detectors, which measure the momentum of muons. Muons are able to pass through both calorimeters as a result of their mass (which is 200 times larger than electron mass). The components of the muon detectors include Thin Gap Chambers, Resistive Plate Chambers, Monitored Drift Tubes, Small-Strip Thin-Gap Chambers, and Micromegas. (ATLAS Collaboration, 2023a)
4. Magnet System: The magnet system modifies the trajectories of charged particles through the usage of solenoidal and toroidal superconducting magnet systems. By bending the particle trajectories, ATLAS can accurately measure the charge and momentum of particles that pass through the detector. The solenoid magnet sits at the core of the experiment and provides a two-Tesla magnetic field. There are three toroids - one that surrounds the detector, and two

which are located at the ends of the experiment. These toroids provide a 3.5 Tesla magnetic field and are used in measuring the momentum of muons. (ATLAS Collaboration, 2023a)

2 The Search For Displaced Leptons

2.1 Motivation

Many beyond standard model theories, including various SUSY theories, feature long-lived particles as manifestations of their models. However, the detectors of the LHC were initially designed to identify and measure particles with short lifetimes (less than a few picoseconds) with decay products that originate near the initial point of collision. Long-lived particles (LLPs), or particles that have lifetimes between 0.1 ns to 100 ns, produce decay products that do not trace back to the original interaction point. This means SUSY model searches that require the presence of particles with large lifetime and displacement parameters, such as searches involving displaced leptons, require the identification of a decay vertex apart from the initial collision point. (Aad and et al.* , ATLAS Collaboration)

In well-motivated, gauge-mediated (GMSB), R-Parity conserving SUSY models, the lightest SUSY particle (LSP), the gravitino, has a small gravitational coupling with the next-to-lightest SUSY particle (NLSP). This means that the NLSP becomes long-lived; in these SUSY models, the NLSP is considered to be a slepton ($\tilde{l}$). Two long-lived sleptons are pair-produced from the initial proton-proton collision, with decay signatures of an undetectable gravitino and a lepton with the same flavor as the LLP. Fig. 5 shows the Feynman Diagram of the resulting pair-produced sleptons and their decay products. (Aad and et al.* , ATLAS Collaboration)

Due to the long-lived, pair-produced slepton, the resulting lepton originates from a different location than the original interaction vertex. In this analysis, we utilize the displaced/delayed nature of the lepton decay product to identify events as potential signal, as opposed to prompt standard model background. Moreover, in proton-proton collisions, transverse momentum is conserved, meaning the gravitino decay product, which is invisible to our detector, leads to the case of “missing” momentum. Taken together, these features, in addition to several other identification variables, provide the basis for “tagging” a displaced lepton. The appearance of displaced leptons beyond stand model background could indicate that SUSY processes are in effect.

The Columbia effort is focused on the case where the pair-produced sleptons from the original collision selectrons, therefore an electron decay product is the object of our search.

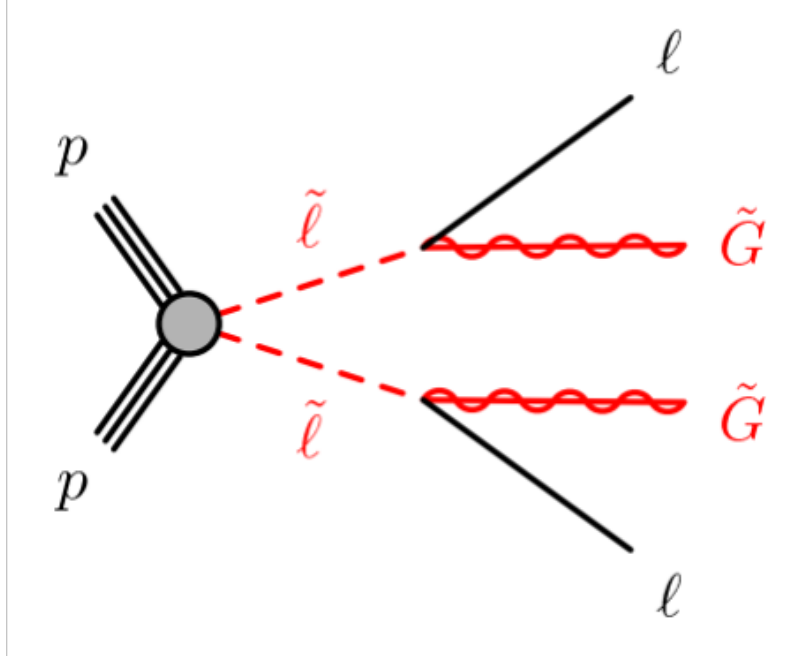


Figure 5: Feynman Diagram of the pair-produced sleptons from the initial proton-proton collision and their decay into a displaced lepton (ℓ) and a gravitino ($\tilde{G}$).

2.2 Boosted Decision Trees in the Displaced Leptons Search

2.2.1 BDT Introduction

In this analysis, we used a machine-learning approach, namely, we applied a Boosted Decision Tree (BDT) to separate Monte-Carlo generated signal events from ATLAS 2022 data background. The framework used for this BDT comes from the Toolkit for Multivariate Analysis (TMVA) Root-integrated environment, which is used to train and test BDTs in addition to mapping a classification function. (Albertsson et al., 2020)

Decision trees are a type of algorithm that utilize supervised learning - that is, events are classified based on known parameters (called input variables) and category labels (in our case, whether a lepton fits into the “signal” or “background” class). The TMVA classification analysis consists of a training phase, where our BDT is trained on an independent set of events, and an application stage, where the BDT is run for the classification problem it has been developed for. To ensure that these processes are orthogonal to one another, we use a technique called k-fold cross-validation. To perform cross-validation, we split our training sample L into k subsets of equal size, called L_k ; in other words, $L = \cup_{i=1..k} L_i$. For each L_i , we train a decision tree, T_i , using the subset of L that excludes L_i . T_i is then tested on the set L_i . Once this process is repeated for each of the k subsets, we are left with k classifiers (decision trees), that we then use to extract the mean performance and uncertainty. Using this cross-validation step, we choose the best-performing model to be retrained on the full training set. (Albertsson et al., 2020; Coadou, 2022)

Decision Tree Algorithms function similarly to rooted binary trees (at least in our implementation, where we only have two class labels, signal, and background). Decision trees start with two branches stemming from a root node, which are then recursively split into two more nodes per branch. Given a set of signal (s) and background (b) events, that taken together form the set X , we can struct the initial, or root node of the decision tree. The decision tree algorithm is then run as follows: (Coadou, 2022)

1. Step 1: If a node satisfies an exit condition (for example, it is n number of branches away from the root), end the algorithm. (Coadou, 2022)
2. Step 2: Sort each element of X for each input variable. (Coadou, 2022)
3. Step 3: Find the input variable that best splits X into most signal vs mostly background events; if there is not a splitting that improves the signal vs background separation, exit the algorithm. (Coadou, 2022)
4. Step 4: Using the variable that leads to the best separation of signal and background, split the data at the value that best separates the events. (Coadou, 2022)
5. Step 5: Recursively apply the algorithm starting from Step 1. (Coadou, 2022)

Boosting is one way to optimize performance and increase the stability of a decision tree. At a high level, boosting uses classification errors from previously made trees to develop a new tree that avoids those pitfalls. The most commonly used Boosting algorithm, called AdaBoost, reweights events based on previous decision tree errors and sums the reweighted results of individual classifiers (previous decision trees) to produce a new classifier tree. AdaBoost is the boosting type that the trees in this analysis use. (Albertsson et al., 2020; Coadou, 2022)

2.2.2 Signal Regions

Since the sleptons that are pair-produced in the in the models we are considering are long-lived, they can travel partway through the detector, leading to decay products (a lepton and gravitino) that have a significant displacement from the initial proton-proton collision. This means the electrons of our search are displaced. The ATLAS detector does not always reconstruct the displaced electrons correctly; sometimes, the displaced electrons are reconstructed as photons. The displaced leptons analysis considers five potential final states that could be the resulting decay products of the slepton. These final states are: 1 electron ($1e$), 1 photon (1γ), 2 electrons ($2e$), 2 photons (2γ), and 1 photon and 1 electron ($e\gamma$). This paper outlines BDT implementations specific to the $1e$ and $2e$ regions, although analyses of this type are not specific to the $1e$ and $2e$ signal regions.

2.2.3 Preselections

The data used in this classification analysis comes from Run 3 ATLAS data taken in 2022 and from Monte-Carlo signal-generated events. Before we run our BDT on signal and background events, we apply several cuts to the inputs to ensure we are only considering events which are relevant to the goals of the classification analysis. For background events, we require the electron time of arrival to be less than 0, which ensures we are blinding our data and looking at a region of space with very few expected signal events. The remaining background after the timing cuts are referred to as control region, or CR data. For signal regions, we apply the following cuts: p_T + trigger cuts (see Tab. 3 and Tab. 4 for trigger and kinematic requirements applied to the $1e$ and $2e$ plots, respectively), cuts on η (excludes crack values), ID and isolation cuts, and a veto on W mass.

Trigger Requirements	Kinematic Requirements
$e30LRT + TM_e + p_{T,e} > 40GeV$ OR $g140 + p_T > 150GeV$	$m_T < 60GeV$ OR $m_t > 100GeV$

Table 3: Cuts applied to signal $1e$ MC. For definitions of the variables in this tables, see Sec.2.2.4.

Trigger Requirements	Kinematic Requirements
$e30LRT + TM_e + p_{T,e} > 40\text{GeV}$ OR	$m_T < 80\text{GeV}$ OR
$g140 + p_T > 150\text{GeV}$ OR	$m_t > 100\text{GeV}$
$2g50 + p_T > 60\text{GeV}$ OR	

Table 4: Cuts applied to signal 2e MC. For definitions of the variables in this tables, see Sec.2.2.4 and Sec.2.2.5

2.2.4 1eBDT

For the BDT applied to the 1e signal region (1eBDT), we use the following event variables:

- $|\eta|$: Pseudorapidity, a spatial coordinate that specifies the angle of a particle relative to the beam axis. Defined as $-\ln(\tan(\frac{\theta}{2}))$
- Maximum ECell Energy: Energy deposited in the calorimeter cell that received the most energy from this electron.
- Displaced p_T : Percent difference between transverse momentum of electron track and electron object
- χ^2/Dof : Goodness of fit parameter for electron track form fit to the hits that form the track divided by the parameter degrees of freedom.
- nPix: Number of pixel layers crossed by the electron track.
- nMiss: Number of missing layers after the electron track's first hit.
- Cluster Energy: Energy deposited by the electron in the calorimeter
- Z_0 : Distance of closest approach from extrapolated electron track to the primary vertex in the longitudinal plane
- $t_{electron}$: Electron time of arrival at calorimeter
- D_0 : Distance of closest approach from extrapolated electron track to the primary vertex in the transverse plane

The plots for each of these variables in the 1 electron BDT region are shown in Fig. 6; the signal values are in blue while the background values are red. Fig. 7 ranks the significance of each input variable; in other words, this plot ranks the weight that the BDT prescribes to each variable when classifying an event as signal vs. background. In this plot, significance is measured as $s/\sqrt{b}$, where s is the number of signal events left after the BDT splits over that variable, and b is the amount of background left after the BDT splits over that variable.

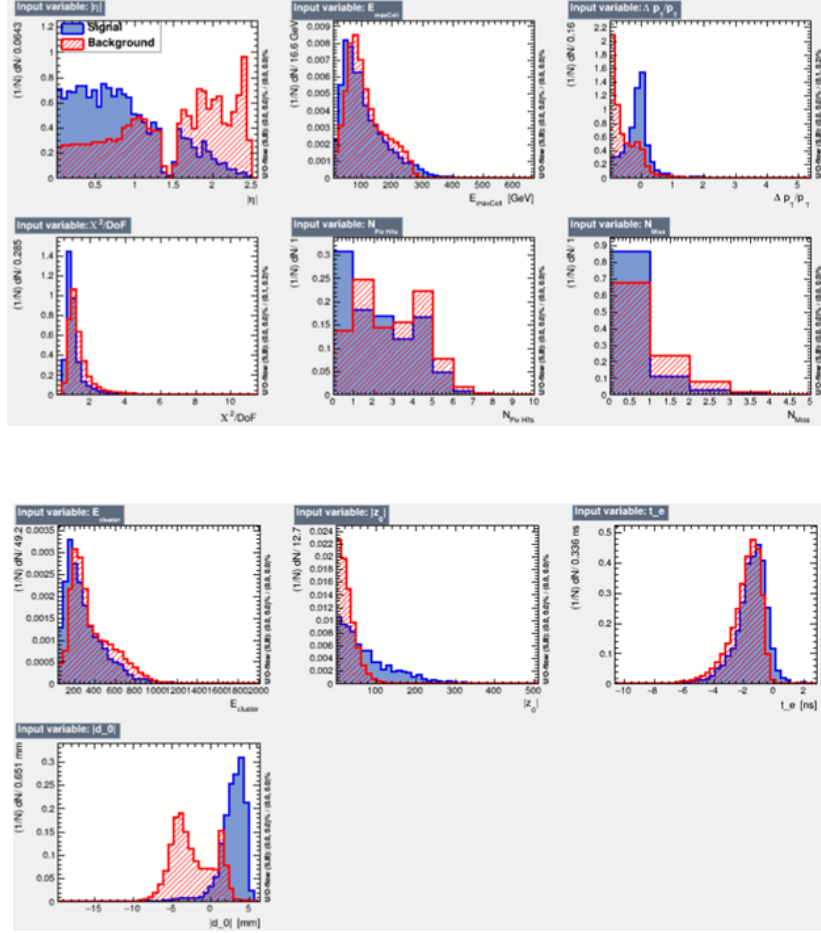


Figure 6: Input variables plots for the 1eBDT. When taken separately, these variables have a very weak ability to distinguish signal and background regions, however when input into our BDT, we find that the separation power increase dramatically, as shown in the ROC and output curves below.

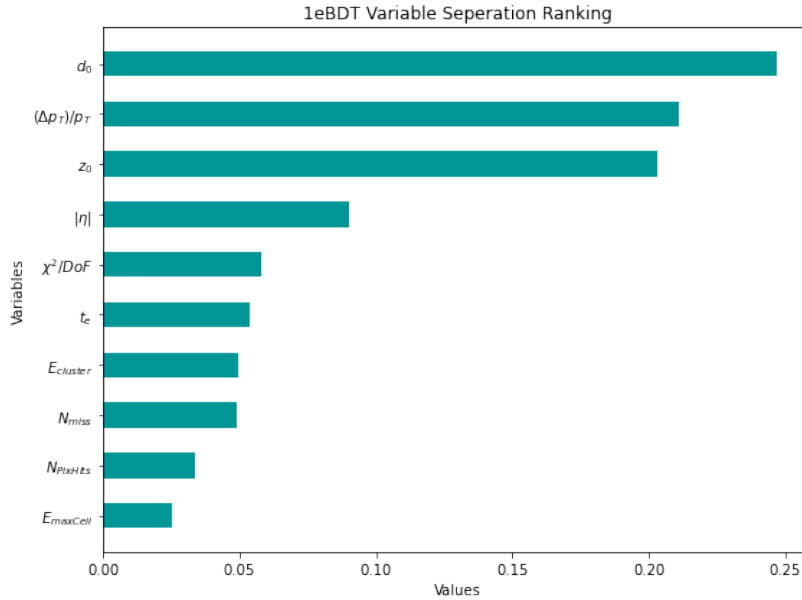


Figure 7: Input variable significance ranking for the 1eBDT; significance is measured as $s/\sqrt{b}$.

The following output curves are used to measure the performance of the 1eBDT. Fig. 8 shows the output plot, which used BDT weights to give each event a score regarding how signal-like vs background-like an event is. BDT scores that are large and positive are more signal-like, whereas values that are large and negative are more background-like. Fig. 9 displays the Receiver Operating Characteristic (ROC) curve, which is a representation of a binary classifier; this curve plots the signal efficiency vs the background rejection. The area under the curve tells us the probability that an event is correctly identified as signal or correctly rejected as background.

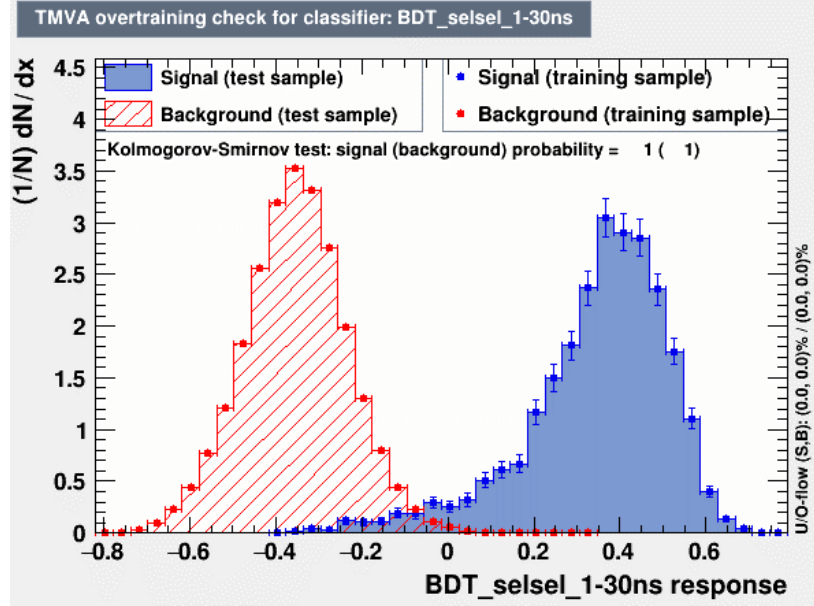


Figure 8: Output characteristics for the 1eBDT. The blue curve represents signal while the red curve displays control region data. Larger BDT score values indicate a more signal-like event. These curves provide a visual representation of the separating power of a BDT.

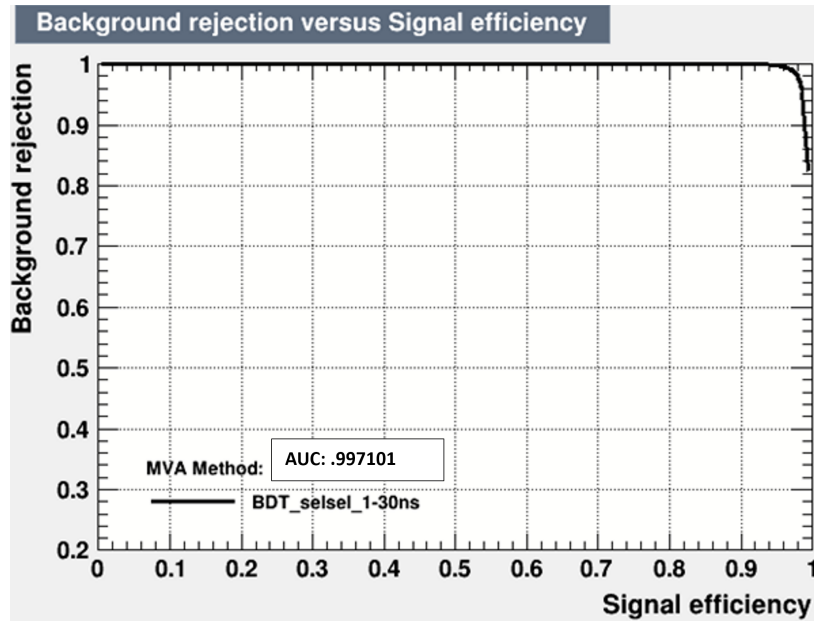


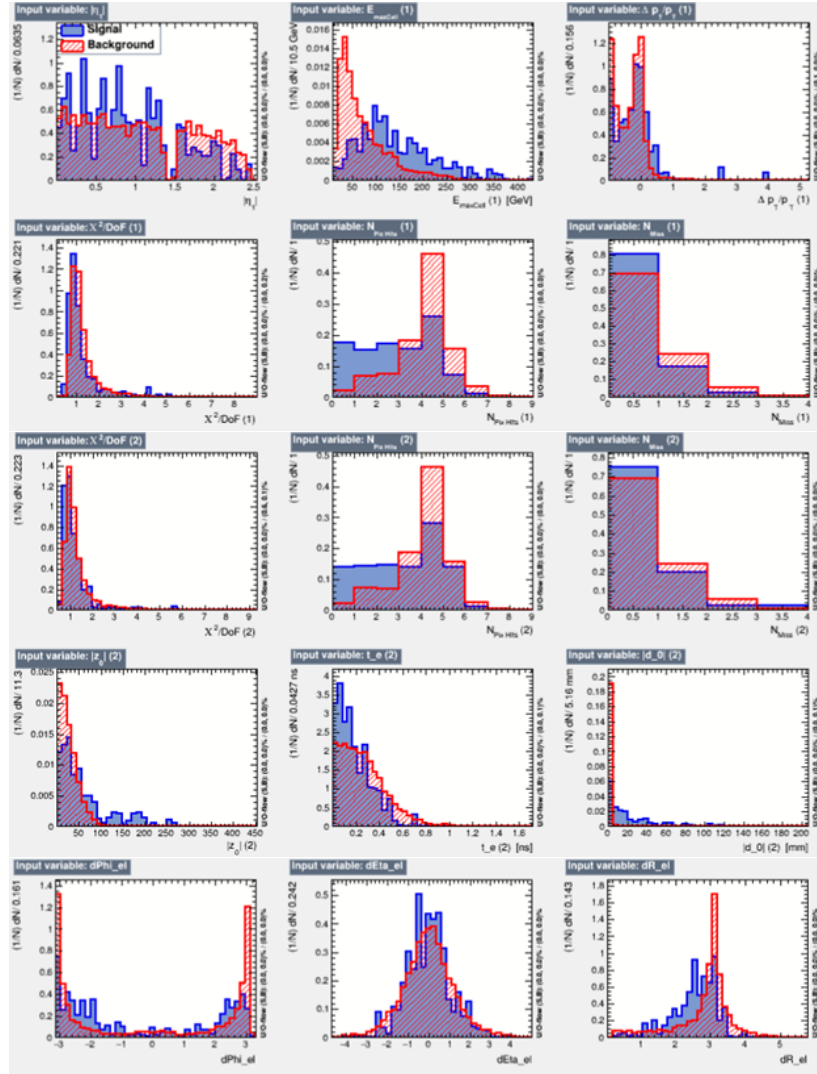
Figure 9: 1eBDT Receiver Operating Characteristic (ROC) curve, which plots the signal efficiency vs the background rejection. The area under the curve represents the probability that an input event is classified correctly as signal or background.

2.2.5 2eBDT

These are the corresponding input variable plots, input variable rankings and output plots for the BDT applied to the 2e signal region (2eBDT), which consists of a leading electron (electron with greater p_T) and a subleading electron. Most of the inputs are the same variables used to create the 1eBDT, just applied to both the leading and subleading electrons. However, there are three newly added correlation variables, which include:

- $d\phi$: angle between the leading and subleading electron
- $d\eta$: difference in η between the leading and the subleading electron
- dR : $\sqrt{(d\phi)^2 + (d\eta)^2}$

Fig. 10 show the input variable plots of the 2eBDT, Fig. 11 ranks the significance of each input variable, Fig. 12 shows the 2eBDT output plot, and Fig. 13 displays the ROC curve.



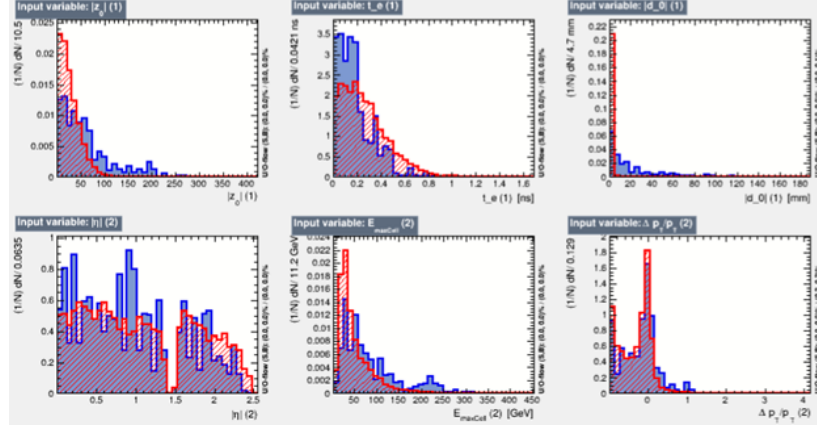


Figure 10: Input variables plots for the 2eBDT. hen taken separately, these variables have a very weak ability to distinguish signal and background regions, however, when input into our BDT, we find that the separation power increase dramatically, as shown in the ROC and output curves below.

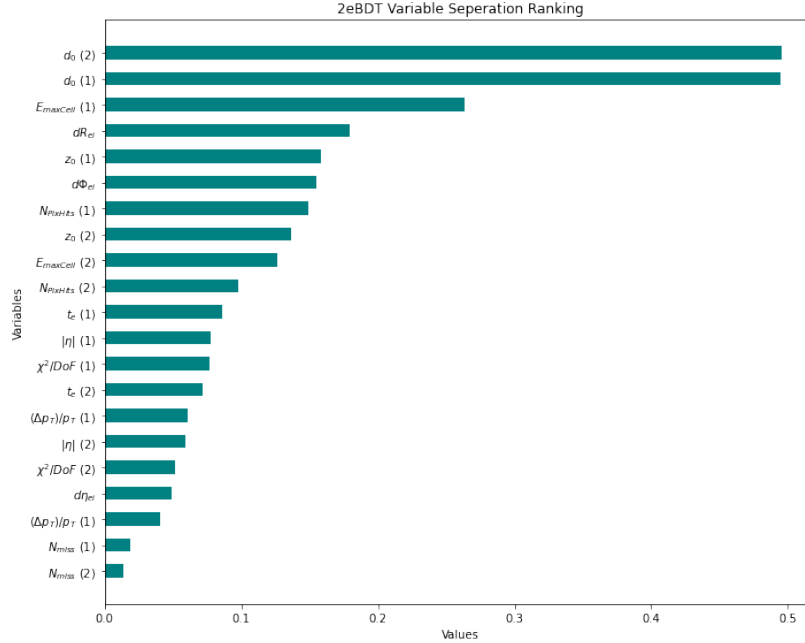


Figure 11: Input variable significance ranking for the 2eBDT; significance is measured as $s/\sqrt{b}$. Like the 1eBDT, d_0 (for both the leading and subleading variables) are ranked very highly. This is expected as one of the features of displaced leptons are high d_0 values. However, we also see d_R , one of the correlation variables as highly ranked.

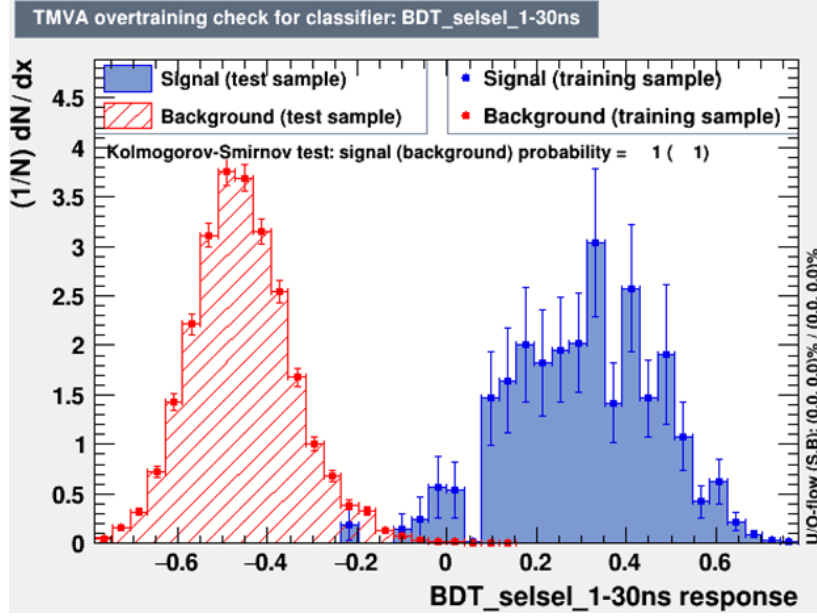


Figure 12: Output characteristics for the 2eBDT. The blue curve represents signal while the red curve displays control region data. Larger BDT score values indicate a more signal-like event. These curves provide a visual representation of the separating power of a BDT.

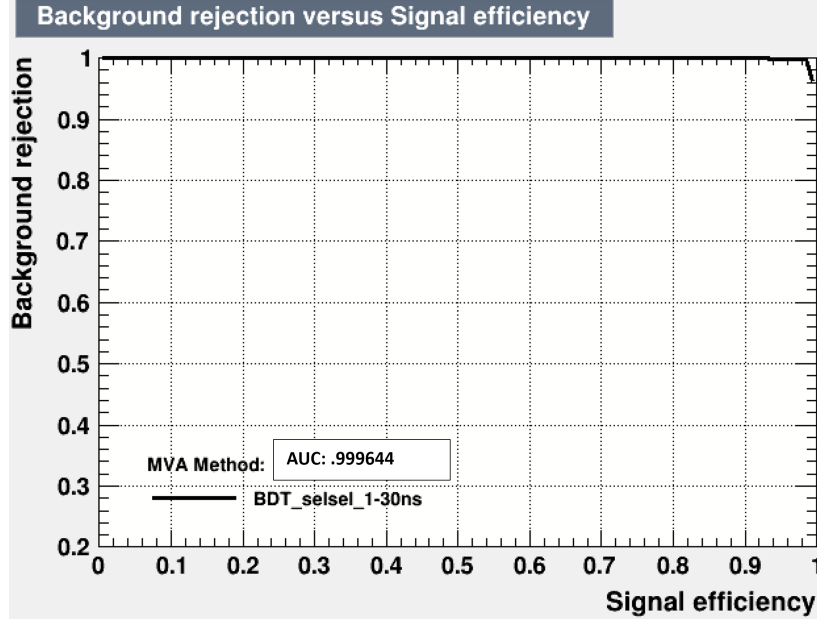


Figure 13: ROC for the 2eBDT. We can see from the AUC score that the 2eBDT performs a bit better than the 1eBDT, however, this is not surprising since there are more input variables in the 2eBDT.

2.3 Event Level vs Displaced Object Tagger BDT

The two BDT implementations show in Sec. 2.2.4 and Sec. 2.2.5 are examples of what we call an "event-level" BDT. This means that each of these BDTs are only intended to be used on the signal region for which they have been implemented; in other words, these BDTs are useful to us only

when we know the topology (signal region) of the objects we want to classify. While each of these BDTs separate signal from background well in their respective signal regions, we need to know an object's specific topology to perform classification which means that the event-level BDTs are more model-dependent, and thus have fewer use cases. A more useful BDT could classify events in various topologies - we call this type of BDT a displaced object "tagger" BDT.

2.3.1 Applying the 1eBDT on a 2e Signal Region

In this analysis, we implemented a displaced tagger BDT by extending the use cases of the 1eBDT. Specifically, we applied the 1eBDT in our $2e$ signal region twice - once for the leading electron and once for the subleading electron. This procedure is equivalent to checking if the leading electron is displaced and then checking if the subleading electron is displaced; both electrons will receive an individual BDT score, as opposed to both electrons receiving a score for the event in which they are both displaced.

2.3.2 Signal and Background Efficiency Comparisons

After implementing the object tagger BDT in the $2e$ signal region, we wanted to compare its performance to the well-performing 2eBDT (which is an event level BDT). We considered two types of object tagger BDTs: an tagger BDT that only considers the events where both electrons in the $2e$ region received high BDT scores as signal (Object Tagger 1) and a tagger BDT which considered the events where at least one of the two electron had a high bdt score as signal events (Object Tagger 2). Tab.5 shows the background efficiency comparisons between the two implementations of the displaced object tagger BDT in addition to the event level BDT when we keep 99 percent of the signal inputs. The first two rows show the efficiencies of Object Tagger 1; a quick analysis of the chart shows that it performs the worst, as it kept 75 percent of the background data while keeping 99 percent of the signal events, which shows very little separating power. Additionally, the $s/\sqrt{b}$ metric for the entire tree was the lowest in the table. Next, we have the 2eBDT, or the event-level (most model dependent) BDT; it performed considerably better than the Object Tagger 1 BDT. The event-level BDT only kept 12 percent of the background events at 99 percent signal efficiency, and the $s/\sqrt{b}$ score more than doubled for this classifier. However, the best distinguishing BDT was the Object Tagger 2 BDT, which kept only about 5 percent of background at 99 percent signal efficiency. The $s/\sqrt{b}$ scores also support this assessment. Given the signal efficiencies and $s/\sqrt{b}$ measures, we find the Object Tagger 2 BDT is the most appropriate model to use for this classification analysis.

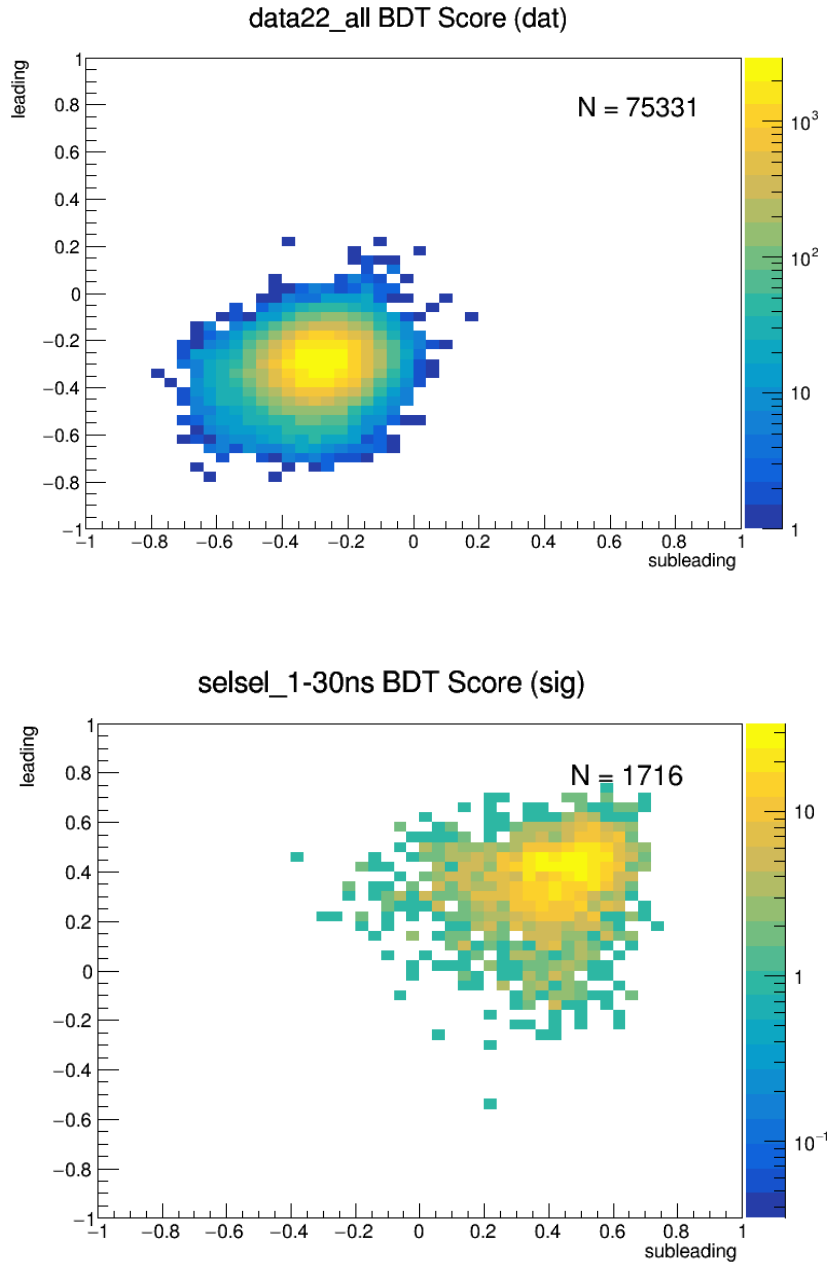


Figure 14: These graphs show use the BDT scores applied to every leading and subleading electron based on the $1e$ BDT weights. As we can see, the background inputs bdt scores (top plot) trend toward the bottom left, where as the signal inputs bdt scores (bottom plot) tend to the top right. This is a visual indication of the distinguishing ability of the displaced tagger BDT; from a qualitative perspective, we can see there is very good separation between signal and background.

	BDT Type	Total	Remaining	Efficiency	$s/\sqrt{b}$
Signal	Object Tagger 1	362.4	361.2	0.99	3.19
Background	Object Tagger 1	17162	12810	0.75	
Signal	Event Level	362.4	359.5	0.99	7.89
Background	Event Level	17162	2080	0.12	
Signal	Object Tagger 2	362.4	359.7	0.99	12.31
Background	Object Tagger 2	17162	853	0.049	

Table 5: Table displaying the performances of the three BDTs when the amount of signal that we are keeping is 0.99 (we keep this the same across all of the models that we are comparing so that our comparison is apples to apples.)

3 Conclusions and Next Steps

Throughout this analysis, we have completed a baseline BDTs for both the 1e and 2e regions, and compared event-level BDTs to a displaced object “tagger” BDT. We have concluded that by requiring at least one electron to be “flagged” as displaced, we find that the object-tagger BDT is the more appropriate classifier for the displaced leptons analysis. Next steps for the analysis include hyperparameter optimization of the 1eBDT, by changing variables such as boosting type (AdaBoost vs Gradient Boosting), tree depth, number of trees used in boosting, Node Size, etc. Moreover, we want to apply this BDT on unblinded data through a procedure called shapefit. If the shapefit results show excess of displaced electrons beyond standard model background, this could indicate SUSY phenomena. This means that our the 1eBDT implementation has the potential to be used as the main discrimination factor in the 1e (and maybe 2e and $e\gamma$) regions!

4 Acknowledgements

Thank you to all the individuals and organizations that made this research possible, including: the ATLAS Collaboration and the Columbia ATLAS group, everyone at Nevis who made this REU possible, with special thanks to John Parsons, Georgia Karagiorgi, and Amy Garwood, and my fellow REUs who I’m grateful I have gotten to know over the past 10 weeks. This material is based upon work supported by the National Science Foundation under Grant No. PHY/1950431.

References

- Aad, G. and et al.* (ATLAS Collaboration) (2021). Search for displaced leptons in $\sqrt{s} = 13$ tev pp collisions with the atlas detector. *Physical Review Letters*, 127:051802. Accessed: July 25, 2023.
- Albertsson, K., Gleyzer, S., Hoecker, A., Moneta, L., Speckmayer, P., Stelzer, J., Therhaag, J., von Toerne, E., Voss, H., and Wunsch, S. (2020). Tmva 4: Toolkit for multivariate data analysis with root. CERN. Accessed: July 25, 2023.
- ATLAS Collaboration (2021). Cern brochure 2021-004. <https://cds.cern.ch/record/2809109/files/CERN-Brochure-2021-004-Eng.pdf>. Accessed: July 25, 2023.
- ATLAS Collaboration (2023a). Atlas detector. <https://atlas.cern/Discover/Detector>. Accessed: July 25, 2023.
- ATLAS Collaboration (2023b). Search in diphoton and dielectron final states for displaced production of higgs or z bosons with the atlas detector in $s\sqrt{s} = 13$ tev pp collisions. *arXiv preprint*. Accessed: July 25, 2023.
- CERN (2023a). Large hadron collider. <https://home.cern/science/accelerators/large-hadron-collider>. Accessed: July 25, 2023.
- CERN (2023b). The standard model of particle physics. <https://home.cern/science/physics/standard-model>. Accessed: July 25, 2023.
- CERN (2023c). Supersymmetry. <https://home.cern/science/physics/supersymmetry#:~:text=Supersymmetry%20is%20an%20extension%20of,mass%20of%20the%20Higgs%20boson>. Accessed: July 25, 2023.
- Charitos, P. (2016). Where is supersymmetry? <https://ep-news.web.cern.ch/content/where-supersymmetry>. Accessed: July 25, 2023.
- Coadou, Y. (2022). Boosted decision trees. In *Artificial Intelligence for High Energy Physics*, pages 9–58. WORLD SCIENTIFIC.
- O’Keefe, M. (2020). Fine-tuning versus naturalness. <https://www.symmetrymagazine.org/article/fine-tuning-versus-naturalness>. Accessed: July 25, 2023.
- Thomason, J., Garoby, R., Gilardoni, S., Jenner, L., and Pasternak, J. (2013). Proton driver scenarios at cern and rutherford appleton laboratory. *Physical Review Special Topics - Accelerators and Beams*, 16.
- Thomson, M. (2013). *Modern Particle Physics*. Cambridge University Press.