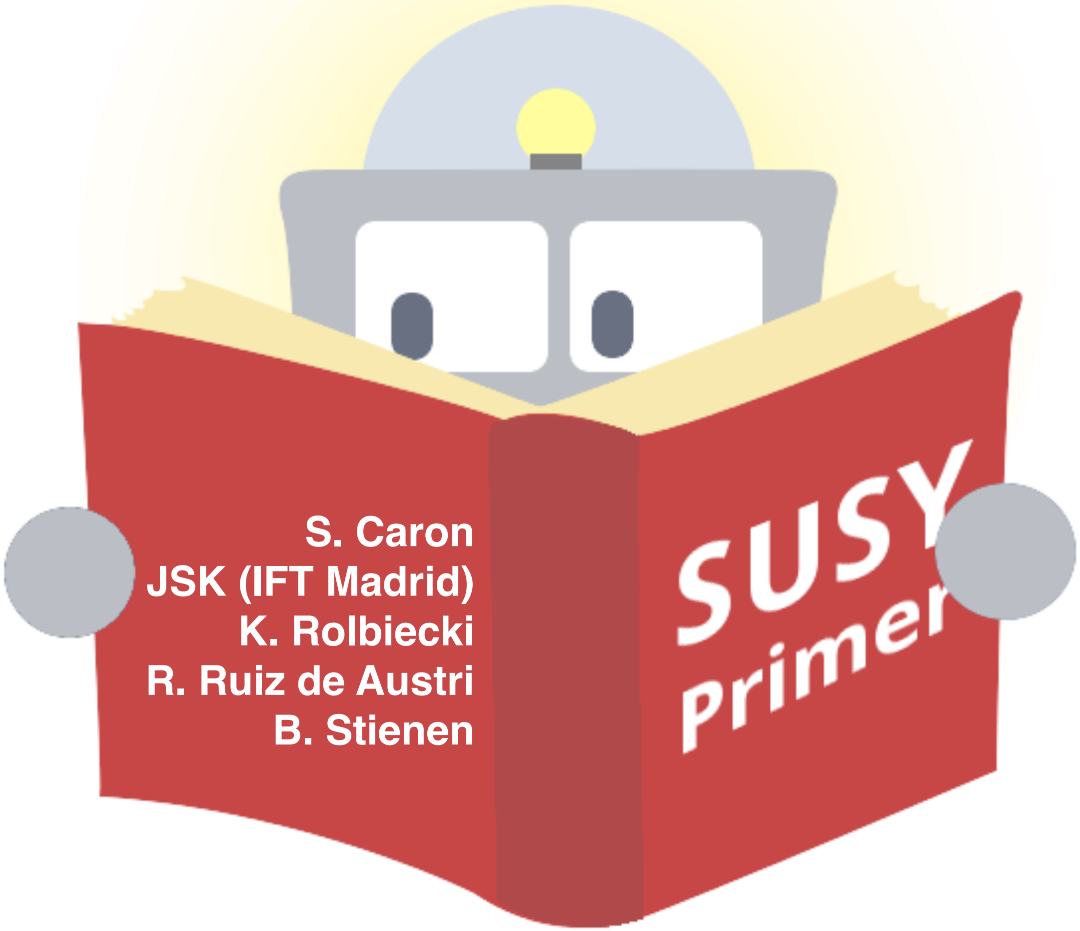
The BSM-AI Project: SUSY-AI



Reinterpreting the results of new physics searches at the LHC

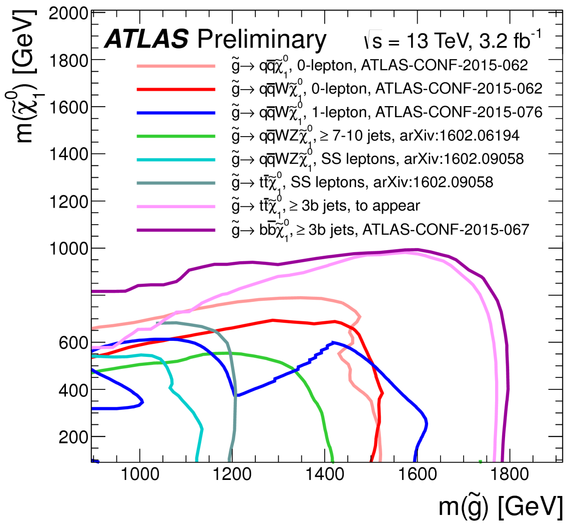
CERN, 15-17 June 2016

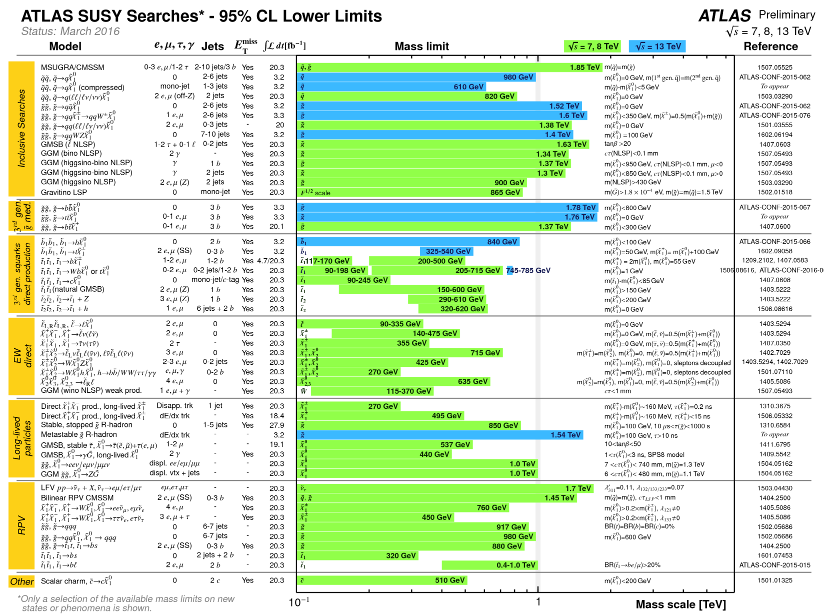
Motivation

LHC Run 2

- first Run 2 data has been analysed and no significant excess above SM expectation has been found apart from the 750 GeV resonance

- null results can be translated in strong limits on BSM scenarios





Constraining BSM

- experimentalists mostly interpret exclusion in simplified models assuming a few BSM particles accessible at the LHC

- in the past constrained models such as mSUGRA with its full parameter space were investigated

- the derived results are very impressive and many SUSY particles below the multi TeV scale are seemingly excluded!

Drawbacks…

- however, many simplified models are unrealistic
- constrained models have a small number of parameters and therefore their whole parameter space can be tested against all phenomenological constraints but the conclusions are not *universal*
- even slight generalisation of the model invalidates the limits
- how can we constrain the most general realization of a model against experimental data?

Derive my own limits…

- generate parton level events with Madgraph
- hadronize events with Pythia, Herwig, Sherpa
- calculate the NLO cross section
- simulate detector response with Delphes, PGS
- code up the relevant searches with hundreds of SRs
- validate the whole implementation

There are no free lunches

- validation can be really painful
- a huge number of ATLAS and CMS searches are on the market
- the implementation of all those searches is time consuming
- several attempts to recast LHC limits, e.g. Atom, CheckMATE, FastLim, MadAnalysis 5, SModels and many others

Simplified vs General

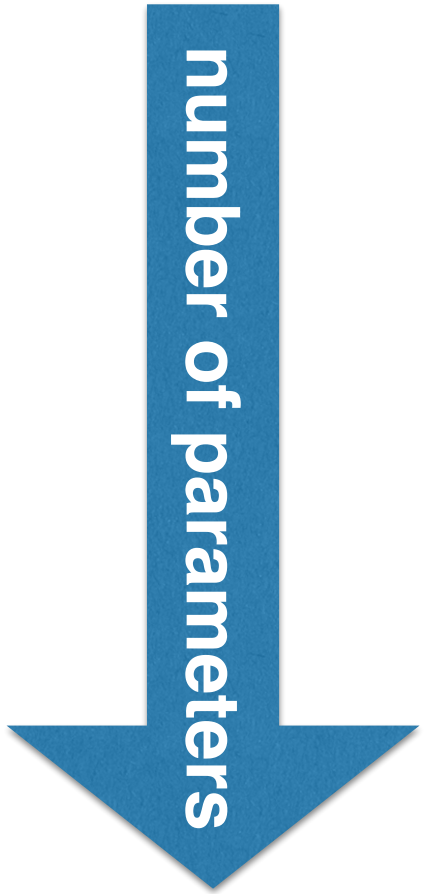
	Pro	Con
FastLim, SModels	very fast, very easy too use	only for simplified models
Atom, CheckMATE, MadAnalysis 5	are very generic	relatively slow since they depend on MC input

Model complexity

simplified models
(1-3 parameters)

constrained models, e.g. mSUGRA
(4-6 parameters)

general models, e.g. pMSSM
(7-20 parameters)

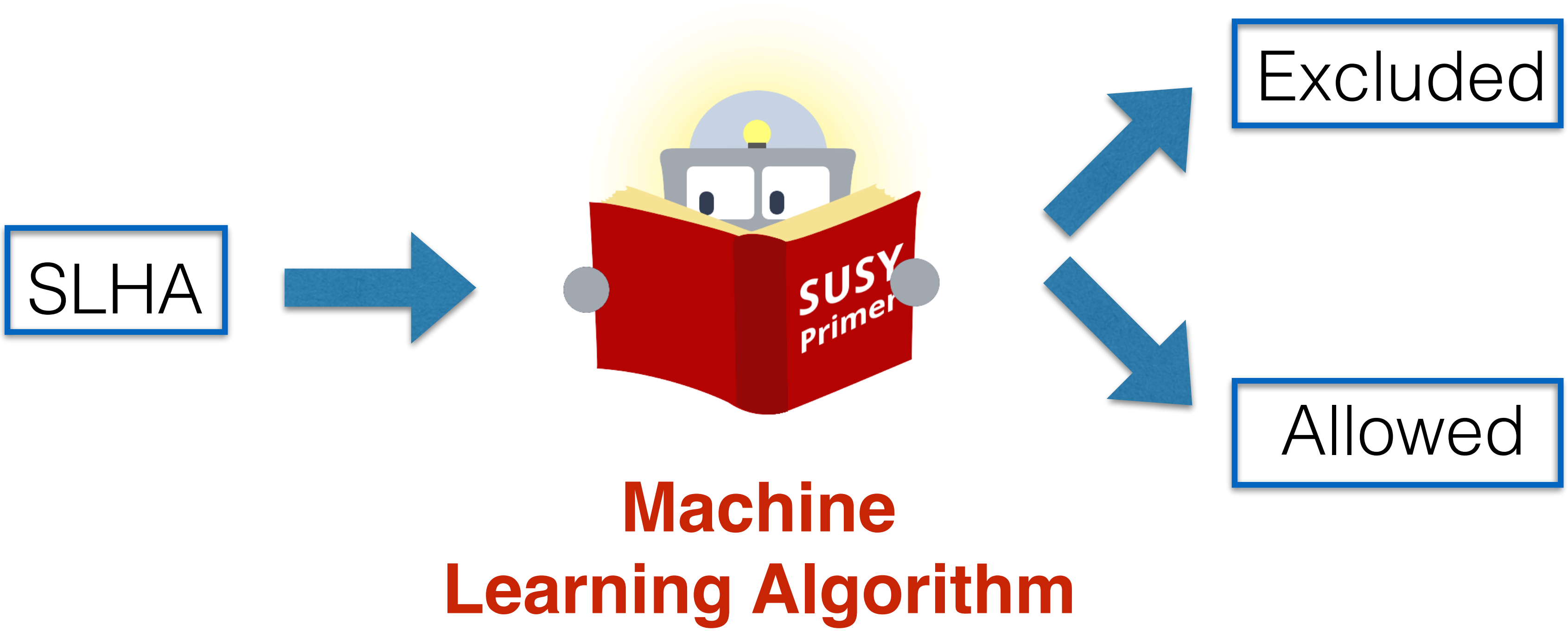


with increasing complexity, a general scan of parameter space becomes impractical (curse of dimensionality)

Global fits

- we want to test a model against EWPO, DD and ID DM, Higgs precision measurements, *b* physics,…
- however, testing a complex model against collider data is difficult if the parameter space has a high dimensional volume
- e.g., in arXiv:1507.07008, we presented a new global fit of the pMSSM-19 compatible with all DM and collider constraints while accomodating the gamma ray excess from the Galactic centre
- testing the collider constraints was computationally expensive
- how can we speed up collider tests for complex models?

Our idea: SUSY-AI



≈ 5000 predictions / CPU second

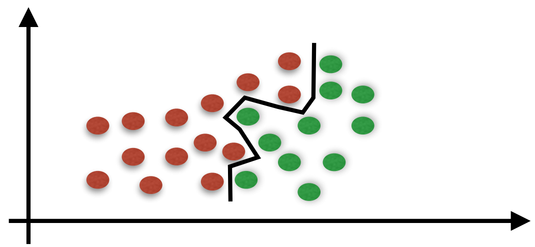
What is Machine Learning?

Machine Learning I

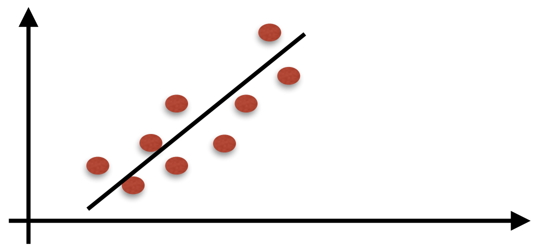
- a ML algorithm can improve its performance using training data
- the algorithm has a large number of fit parameters which can be determined by data
- ML is applied in situations which are very challenging, e.g. face or handwriting recognition

Machine Learning II

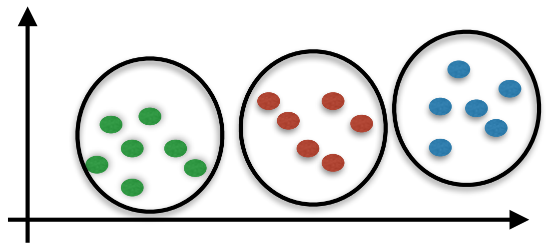
Classification - supervised



Regression - supervised

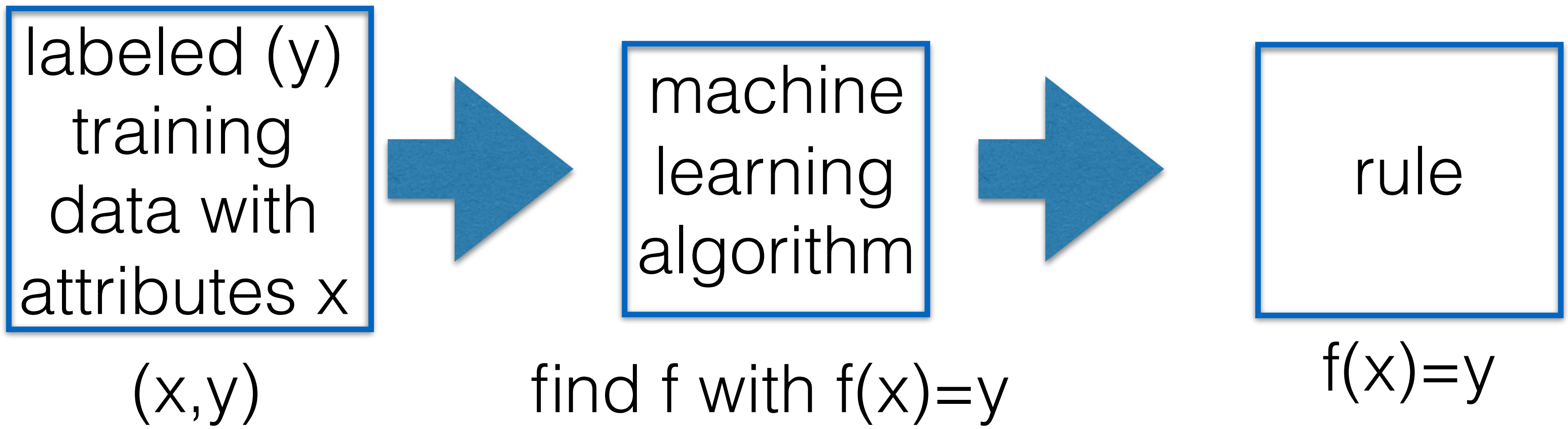


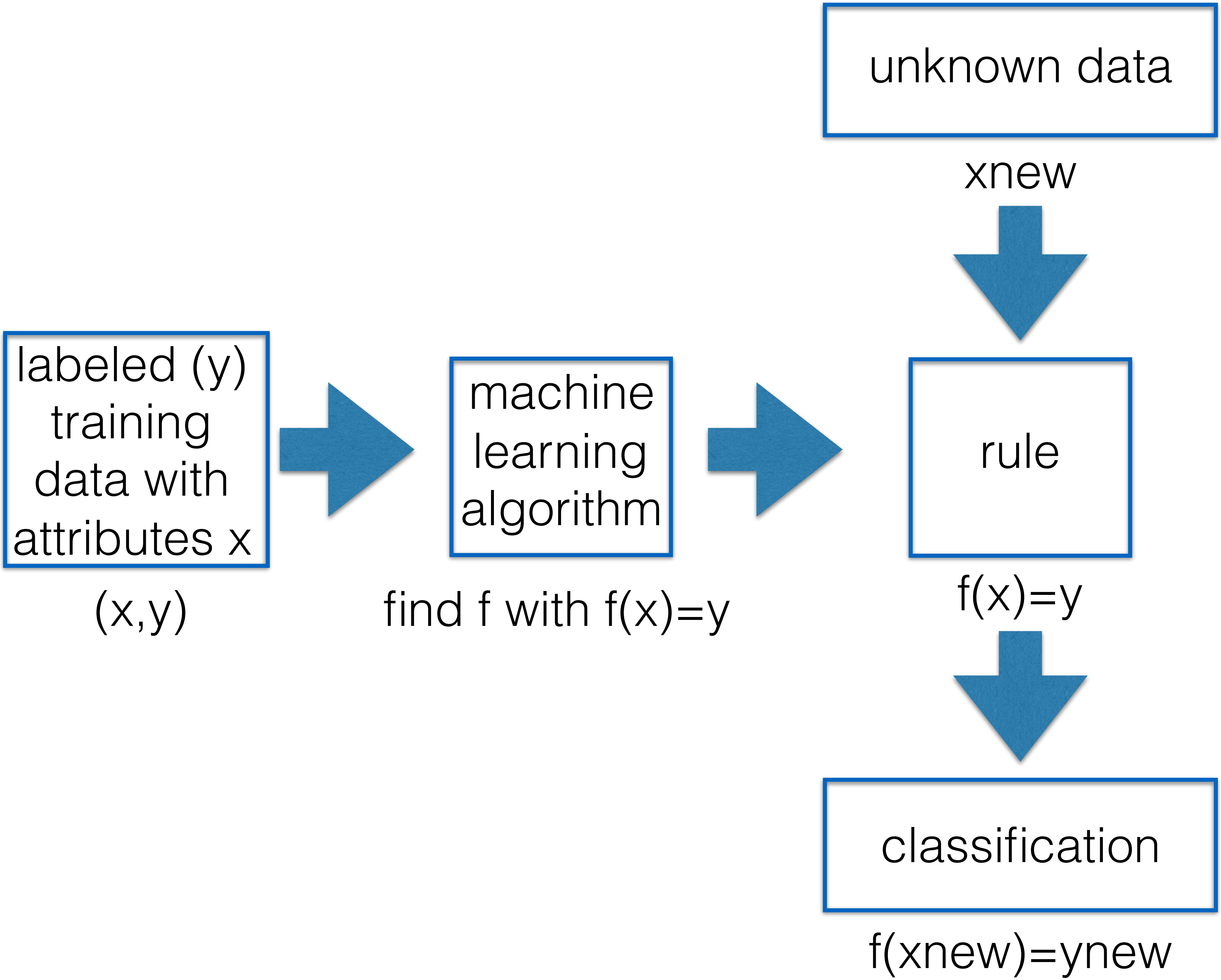
Unsupervised learning



Example

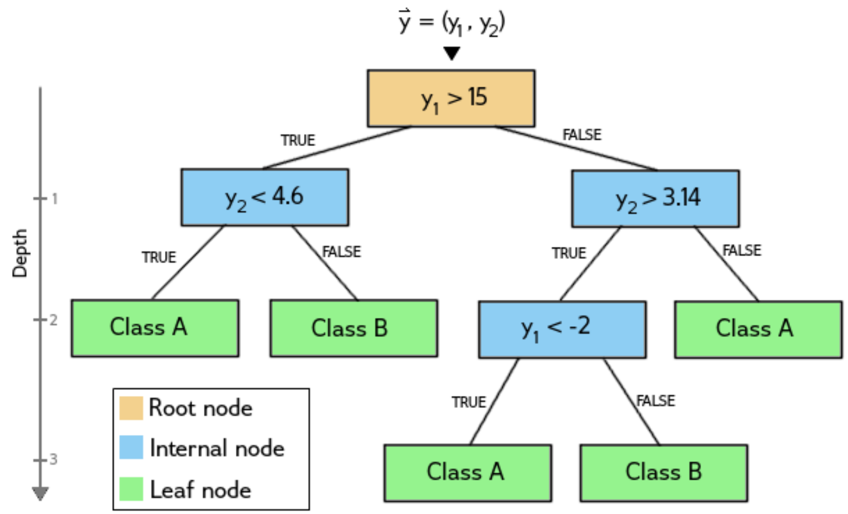
example	attribute 1	attribute 2	label
banana	…	…	1
tomato	…	…	0
cherry	…	…	1
apple	…	…	1
onion	…	…	0
cucumber	…	…	0
orange	…	…	1
water melon	…	…	?
turnip	…	…	?
maiz	…	…	?


labeled (y) training data with attributes x
(x,y)
machine learning algorithm
find f with f(x)=y
rule
f(x)=y


unknown data
xnew
labeled (y) training data with attributes x
(x,y)
machine learning algorithm
find f with f(x)=y
rule
f(x)=y
classification
f(xnew)=ynew

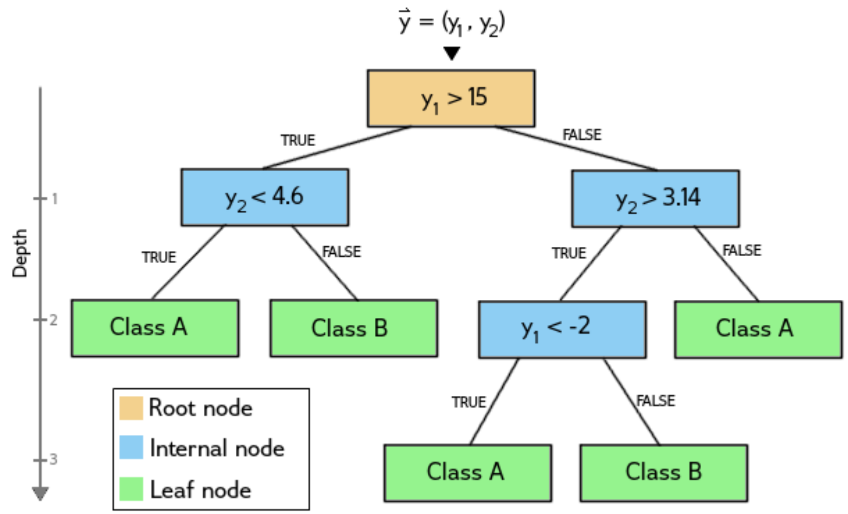
Decision Trees I

- a Decision Tree is a commonly used classification algorithm
- DT consists of several nodes and at each node a test is performed
- the attribute set moves down the tree until the final leaf node is reached
- at the final leaf node, a class label is assigned to the attribute set



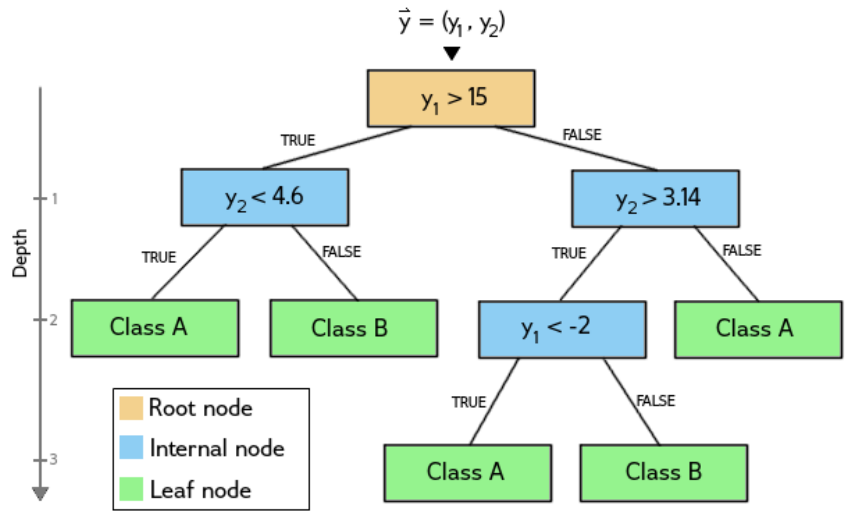
Decision Trees II

- The DT works on the whole attribute set
- every test corresponds to a cut in this parameter space
- a DT split the attribute set into disjunct regions
- disadvantage: tendency of overtraining, i.e. DT learns the noise



Decision Trees III

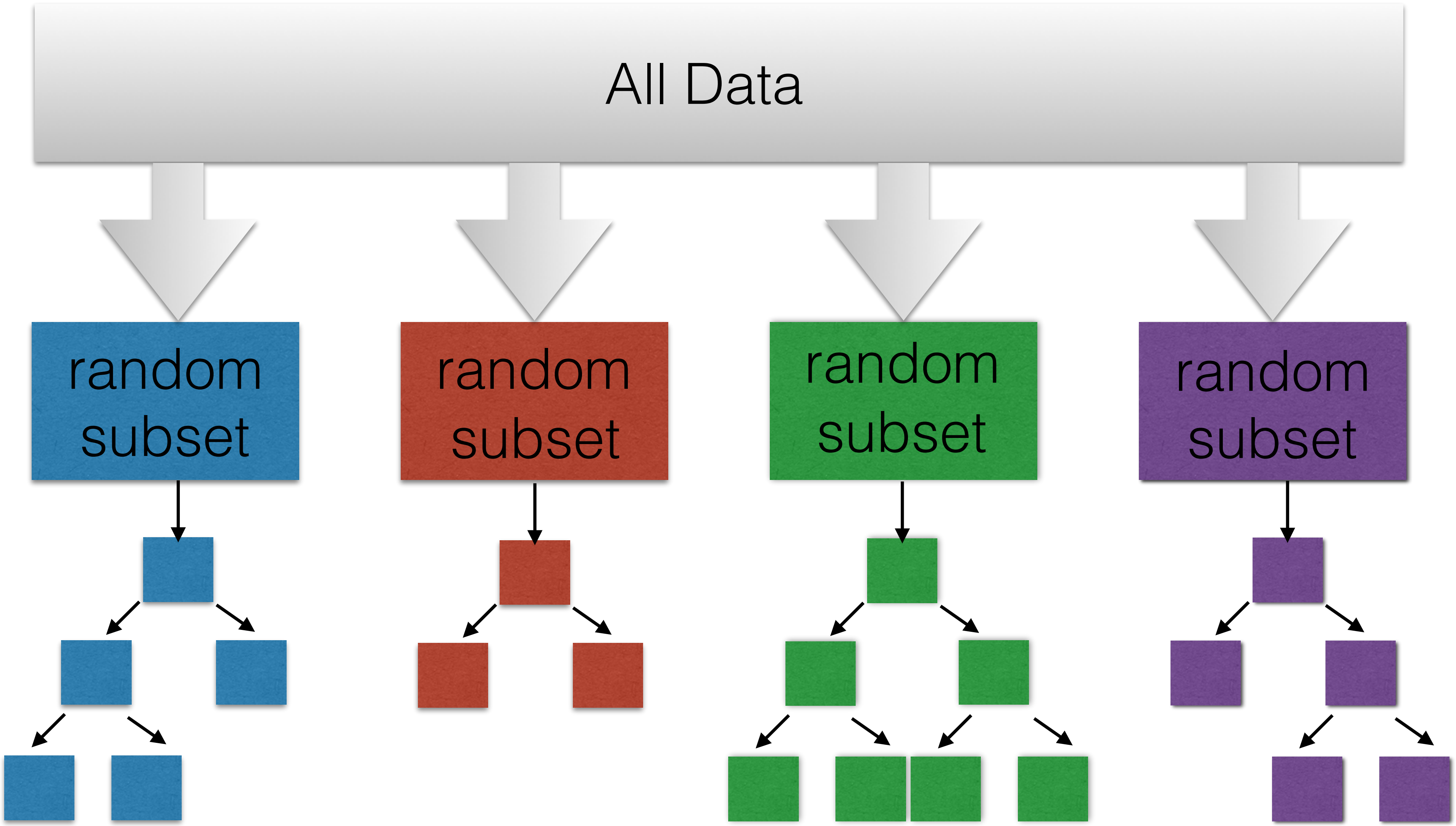
- methods to suppress overtraining
- pruning: training the entire tree but cut away all nodes beyond a certain depth
- boosting: combine multiple DTs into a single classifier
- bagging: for a group of DTs, each decision tree is trained on a random selection of a subset of the attribute set



Random Forest I

- trees are weak learners but a forest is a strong learner
- a random forest combines trees (boosting)
- draw *N* bootstrap samples from original sample
- fit a classification tree to each bootstrap sample
- randomly preselect *M* attribute variables at each node (bagging)

Random Forest II



at each node a random subset of the attribute set is chosen

Random Forest III

- Output the ensemble of trees
- *R*=(# trees prediction of class *C*) / (total number of trees)
- *R*=probability of attribute set belonging to *C*
- classification in RF is done by majority vote

pMSSM-19 & ATLAS

pMSSM-19 I

- the most general MSSM has a large number of input parameters, $\mathcal{O}(100)$ soft breaking parameters!
- it is unfeasible for a dedicated collider study
- assumptions on the soft breaking sector heavily reduces number of free parameters
- however, this approach might be too constraining
- consider a MSSM taking into account all constraints from particle physics experiments

pMSSM-19 II

- consider the most general and CP conserving MSSM

- assume minimal flavour violation

- demand that the lightest neutralino is the LSP

- require the first two generation sfermions are degenerate and decoupled

- 19 weak scale parameters = pMSSM-19

pMSSM-19 and ATLAS I

- ATLAS (arXiv:1508.06608) performed a study on the pMSSM-19
- ATLAS considered 5x10^8 model points based on arXiv:1206.4321
- all model points had to satisfy preselection cuts
- 310,327 model points satisfy all theoretical and experimental constraints

Parameter	Description	Scanned range
$m_{\tilde{L}_1}$	1st/2nd gen. $SU(2)$ doublet soft breaking slepton mass	[90 GeV, 4 TeV]
$m_{\tilde{E}_1}$	1st/2nd gen. $SU(2)$ singlet soft breaking slepton mass	[90 GeV, 4 TeV]
$m_{\tilde{L}_3}$	3rd gen. $SU(2)$ doublet soft breaking slepton mass	[90 GeV, 4 TeV]
$m_{\tilde{E}_3}$	3rd gen. $SU(2)$ singlet soft breaking slepton mass	[90 GeV, 4 TeV]
$m_{\tilde{Q}_1}$	1st/2nd gen. $SU(2)$ doublet soft breaking squark mass	[200 GeV, 4 TeV]
$m_{\tilde{U}_1}$	1st/2nd gen. $SU(2)$ singlet soft breaking squark mass	[200 GeV, 4 TeV]
$m_{\tilde{D}_1}$	1st/2nd gen. $SU(2)$ singlet soft breaking squark mass	[200 GeV, 4 TeV]
$m_{\tilde{Q}_3}$	3rd gen. $SU(2)$ doublet soft breaking squark mass	[100 GeV, 4 TeV]
$m_{\tilde{U}_3}$	3rd gen. $SU(2)$ singlet soft breaking squark mass	[100 GeV, 4 TeV]
$m_{\tilde{D}_3}$	3rd gen. $SU(2)$ singlet soft breaking squark mass	[100 GeV, 4 TeV]
A_t	Stop trilinear coupling	[−8 TeV, 8 TeV]
A_b	Sbottom trilinear coupling	[−4 TeV, 4 TeV]
A_τ	Stau trilinear coupling	[−4 TeV, 4 TeV]
$\|\mu\|$	Higgsino mass parameter	[80 GeV, 4 TeV]
$\|M_1\|$	Bino mass parameter	[0 TeV, 4 TeV]
$\|M_2\|$	Wino mass parameter	[70 GeV, 4 TeV]
M_3	Gluino mass parameter	[200 GeV, 4 TeV]
M_A	Pseudoscalar Higgs mass	[100 GeV, 4 TeV]
$\tan\beta$	Ratio of vacuum expectation values	[1, 60]

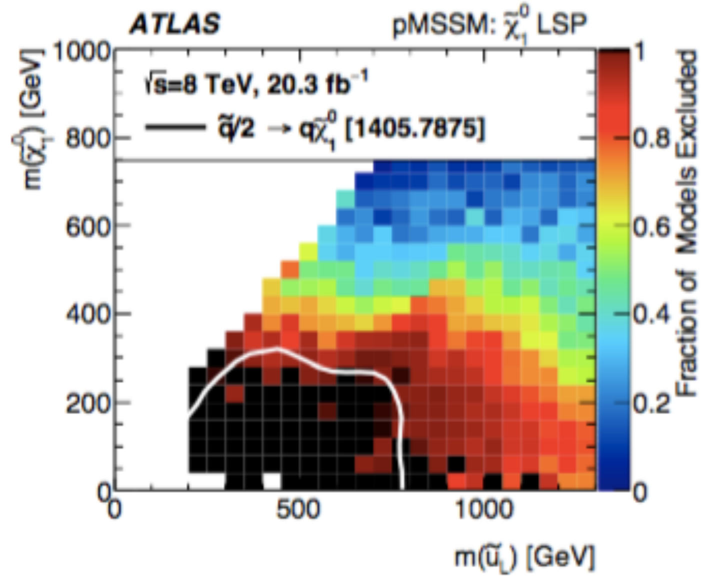
Parameter	Minimum Value	Maximum Value
$\Delta\rho$	-0.0005	0.0017
$\Delta(g-2)_\mu$	-17.7×10^{-10}	43.8×10^{-10}
$\mathrm{BR}(b \to s\gamma)$	2.69×10^{-4}	3.87×10^{-4}
$\mathrm{BR}(B_s \to \mu^+\mu^-)$	1.6×10^{-9}	4.2×10^{-9}
$\mathrm{BR}(B^+ \to \tau^+\nu_\tau)$	66×10^{-6}	161×10^{-6}
$\Omega_{\tilde{\chi}_1^0} h^2$	–	0.1208
$\Gamma_{\mathrm{invisible}}(Z)$	–	2 MeV
Masses of charged sparticle	100 GeV	–
$m_{\tilde{\chi}_1^\pm}$	103 GeV	–
m_h	124 GeV	128 GeV

pMSSM-19 and ATLAS II

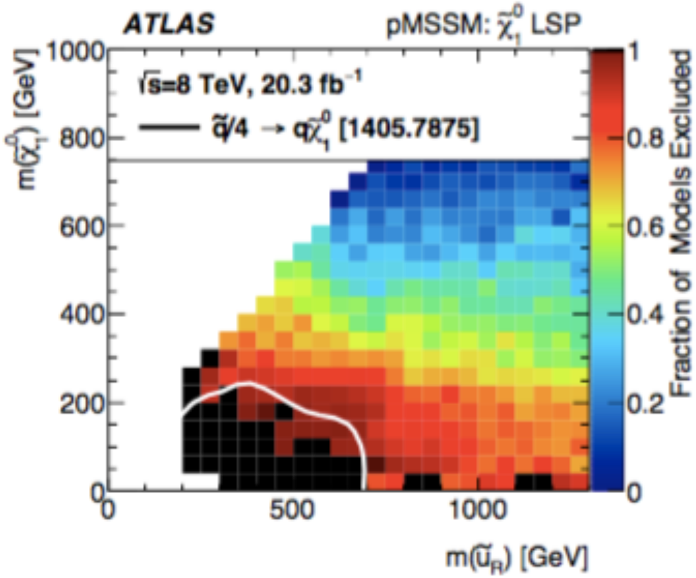
- ATLAS considered 22 separate analyses of Run 1
- a large number of final state topologies are covered
- all relevant processes were generated at truth level
- a fast detector simulation based on GEANT4 were performed

Reference	Final State	Category
[39]	0 lepton + $2-6$ jets + $\not\!\!E_T$	Inclusive
[40]	0 lepton + $7-10$ jets + $\not\!\!E_T$	
[41]	1 lepton + jets + $\not\!\!E_T$	
[42]	$\tau(\tau/\ell)$ + jets + $\not\!\!E_T$	
[43]	SS/3 lepton + jets + $\not\!\!E_T$	
[44]	b jets + 0/1 lepton + $\not\!\!E_T$	
[45]	monojet	
[46]	0 lepton stop search	Third generation squarks
[47]	1 lepton stop search	
[48]	2 lepton stop search	
[49]	monojet search	
[50]	stop search with Z in final state	
[51]	$2b$ jet sbottom search	
[4]	asymmetric stop search	
[52]	1 lepton plus Higgs final state	Electroweak
[53]	dilepton final state	
[54]	2τ final state	
[55]	trilepton final state	
[56]	four-lepton final state	
[57]	disappearing track	
[58, 59]	Long-lived particle search	Other
[60]	$H/A \to \tau\tau$ search	

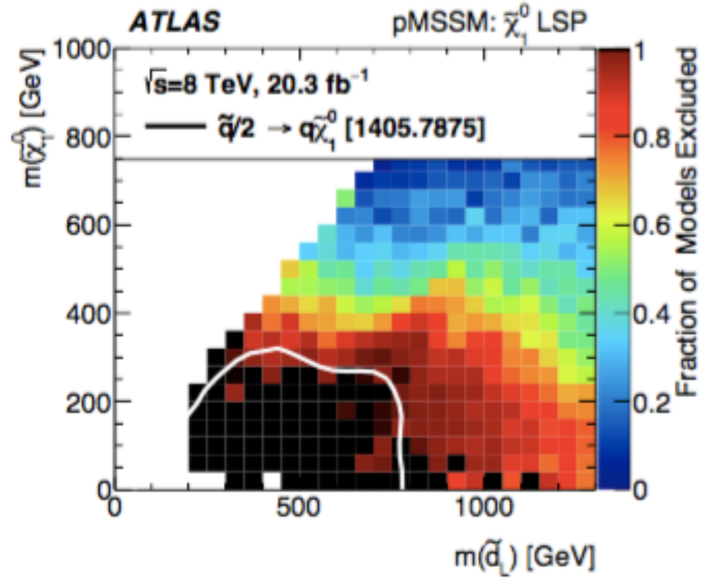
pMSSM-19 and ATLAS III

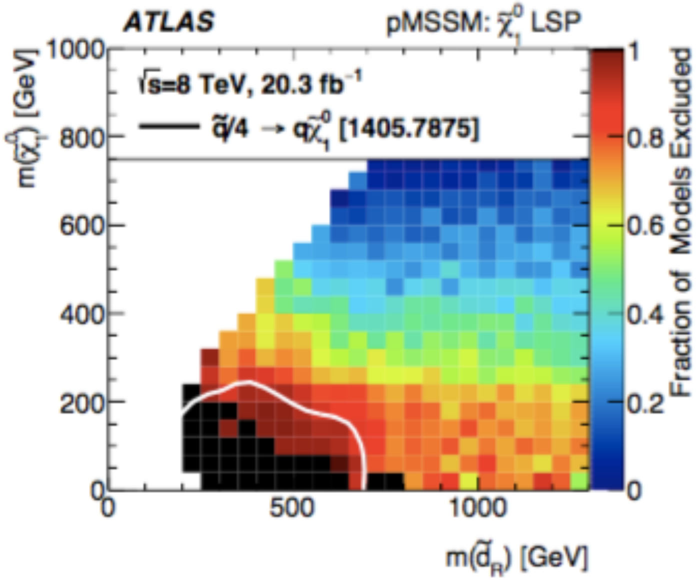


(a) Left up squark



(b) Right up squark





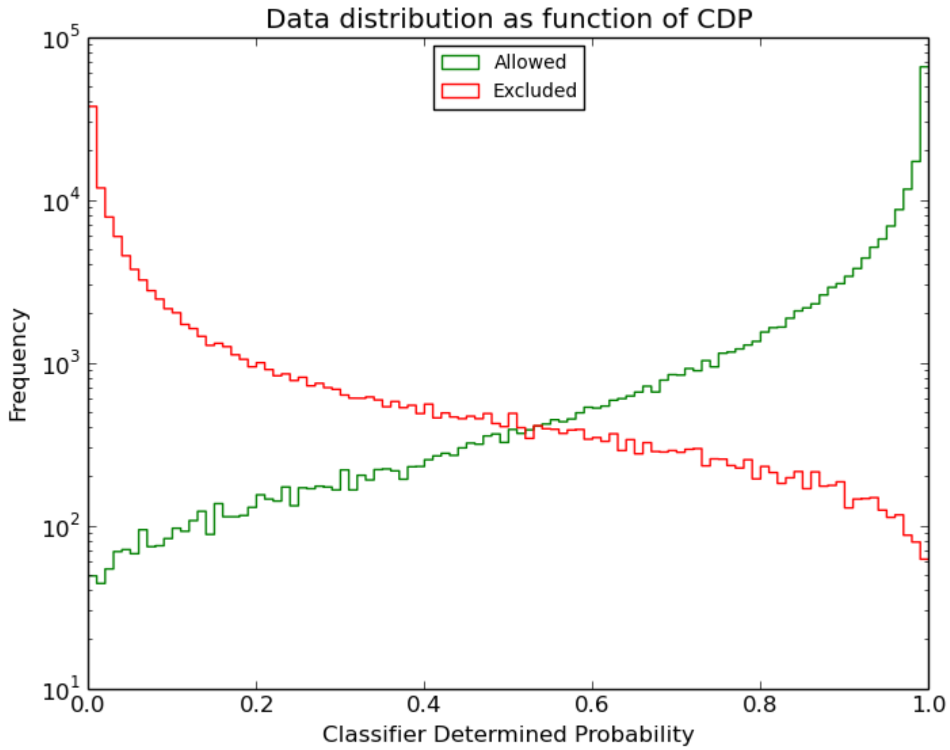
Training

Training of SUSY-AI I

- we used the Python package *scikit-learn-0.17.1*
- we trained our RF classifier with the ATLAS data points
- we determined the optimal classifier configuration in a grid search
- 900 DT with a maximal depth of 30 nodes and a maximum number of features considered at each node of 12

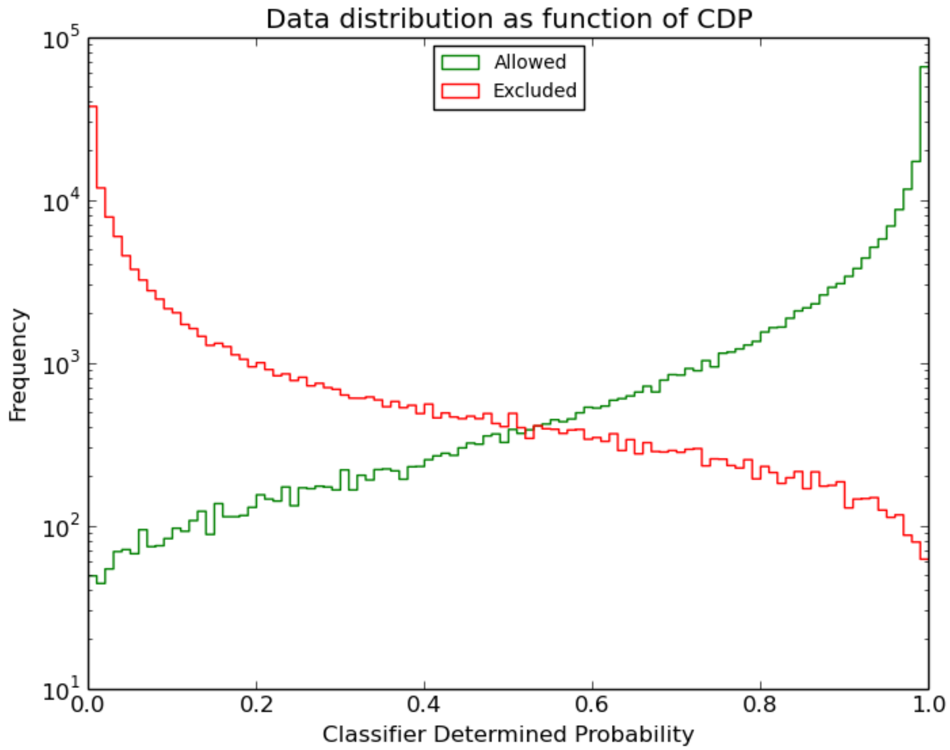
Training of SUSY-AI II

- all predicted data points are assigned with a classification probability by the RF classifier
- the green histogram includes all points which are truly allowed
- the red histogram includes all truly excluded points
- the x-axis corresponds to the classifier determined probability (CDP)



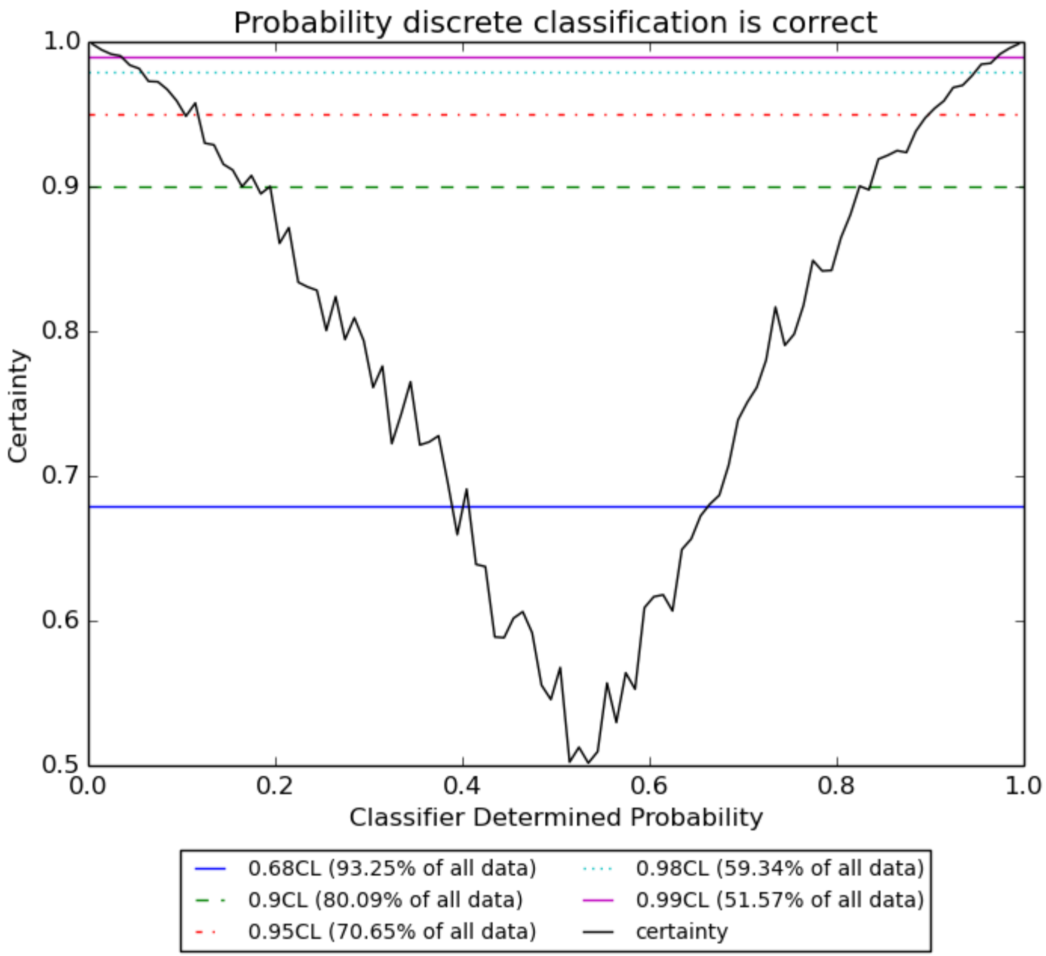
Training of SUSY-AI III

- CDP is the probability that the model point is allowed
- majority of points are correctly classified
- however, perfect classification is not possible
- a cut makes the classification binary, e.g. a cut at 0.5, i.e. for ≥0.5, point is allowed



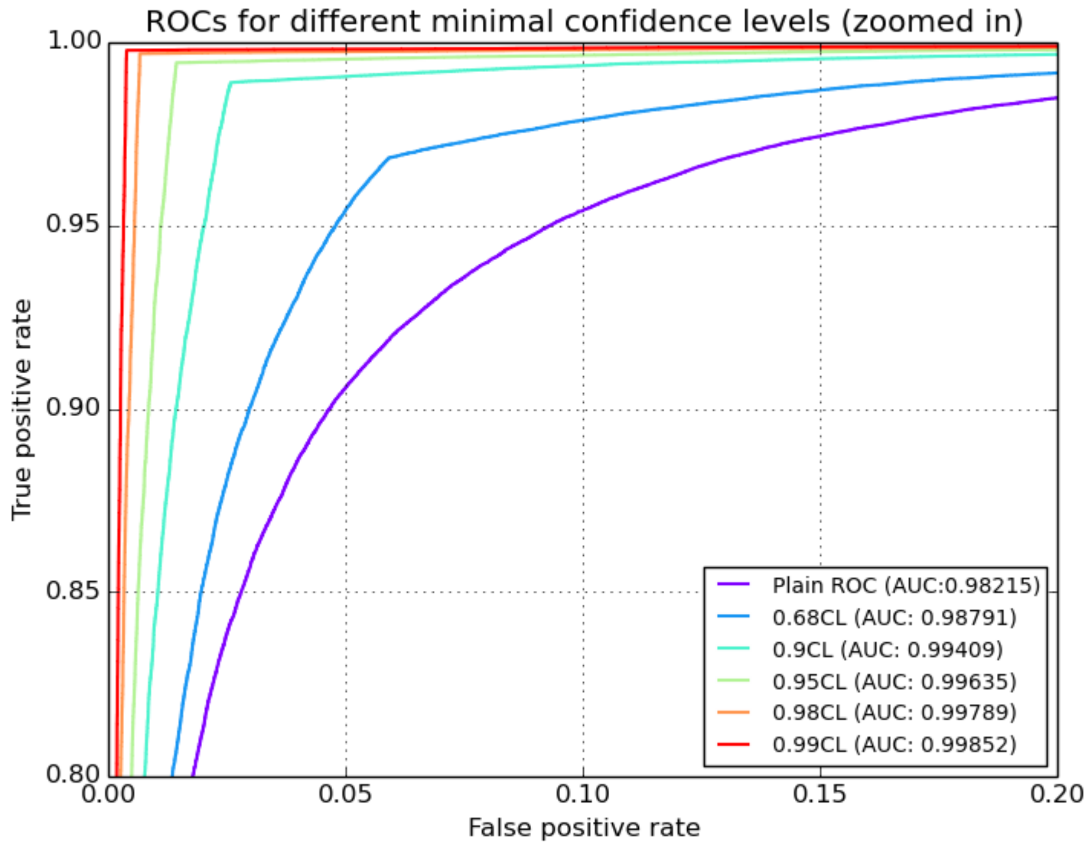
Training of SUSY-AI IV

- take ratio of upper histogram and total number of points in each bin
- it allows a frequentist confidence level that a point with given CDP is allowed or excluded
- e.g.: a CL of 98% corresponds to a CDP of below 0.05 or above 0.95
- a CL of 95% corresponds to predicted probabilities below 0.133 or above 0.9



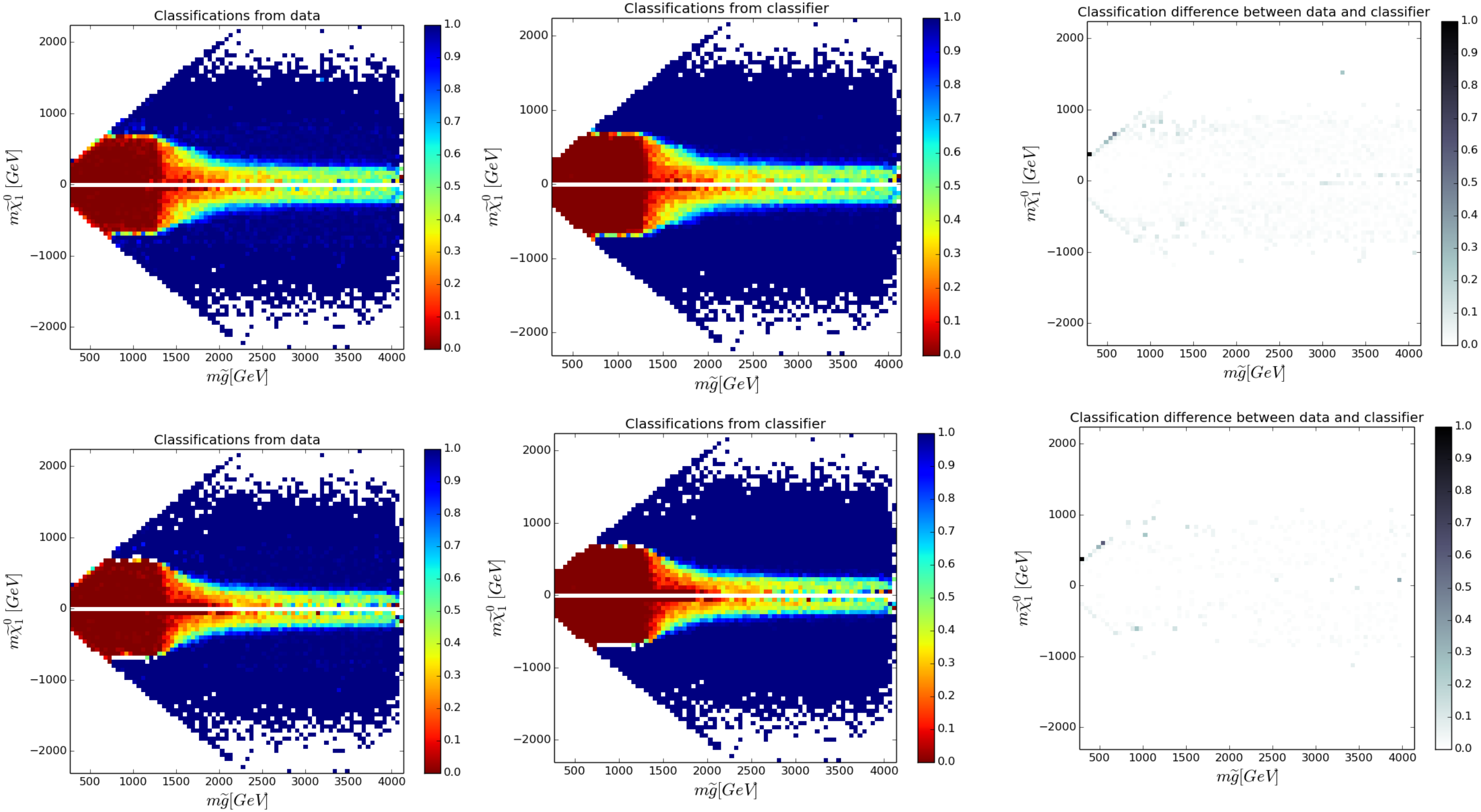
Training of SUSY-AI V

- a "harder" cut provides more reliable results for classification but larger number of points are removed
- the performance can be quantified by the ROC curve
- higher CL cuts increases the AUC

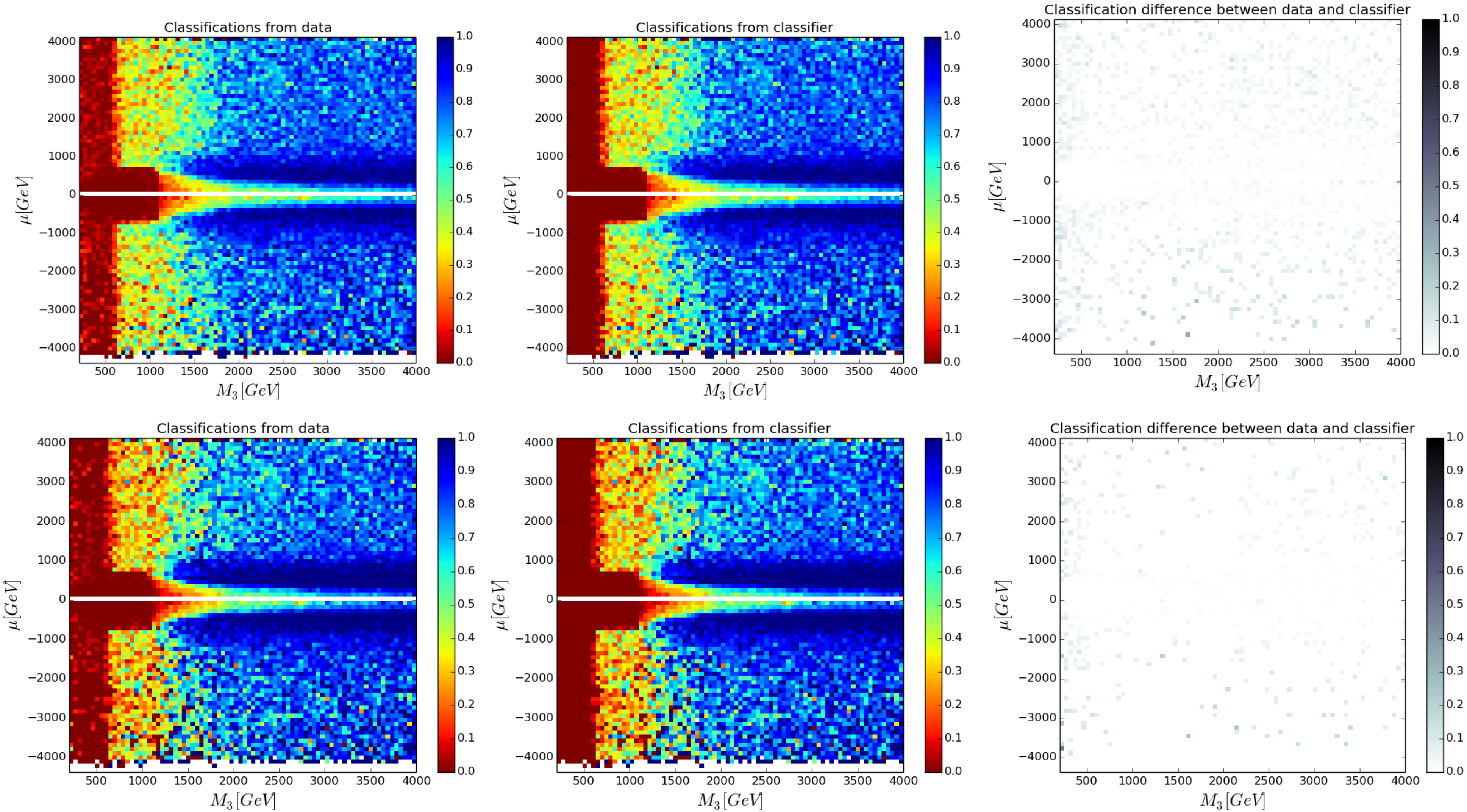


Performance

Performance of SUSY-AI I



Performance of SUSY-AI II



Applications

natural SUSY I

- a minimal natural SUSY scenario consists of light higgsinos, SU(2) doublet third generations squarks, a SU(2) singlet stop and multi-TeV gluinos

- the scenario consists of six input soft breaking parameters

- 22000 benchmark points were generated and the produced MC events analysed with CheckMATE

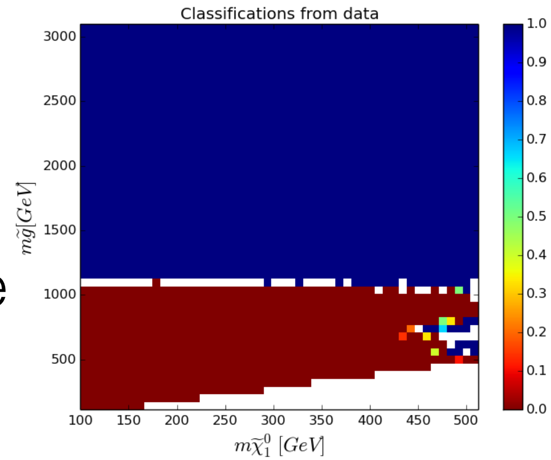
- we derived limits in minimal NSUSY parameter space (arXiv:1511.04461)

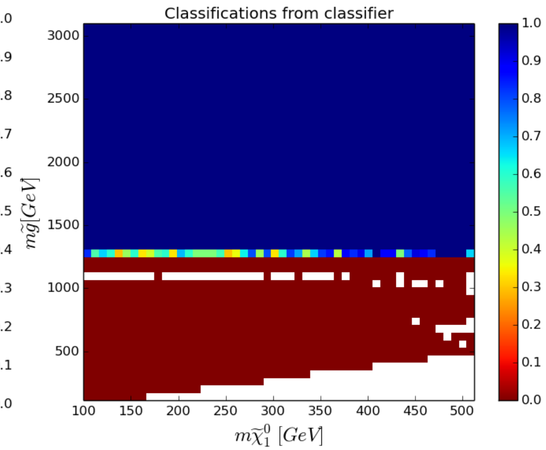
Parameter	Description	Scanned range
$m_{\tilde{Q}_3}$	3[rd] generation $SU(2)$ doublet soft breaking squark mass	[0.1 TeV, 1.5 TeV]
$m_{\tilde{U}_3}$	3[rd] generation $SU(2)$ singlet soft breaking squark mass	[0.1 TeV, 1.5 TeV]
M_3	Gluino mass parameter	[0.1 TeV, 3.0 TeV]
A_t	Stop trilinear coupling	[−3.0 TeV, 3.0 TeV]
μ	Higgsino mass parameter	[0.1 TeV, 0.5 TeV]
$\tan\beta$	Ratio of vacuum expectation values	[1, 20]

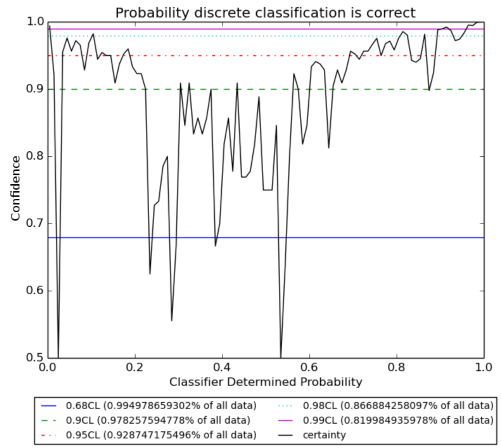
Reference	Final State	$\mathcal{L}$ [fb^{-1}]	#SR
1308.2631 (ATLAS) [51]	$0\ell+2b$ jets$+\not{E}_T$	20.1	6
1403.4853 (ATLAS) [48]	$2\ell+\not{E}_T$	20.3	12
1404.2500 (ATLAS) [43]	SS 2ℓ or 3ℓ	20.3	5
1407.0583 (ATLAS) [47]	$1\ell+(b)$ jets$+\not{E}_T$	20.0	27
1407.0608 (ATLAS) [49]	monojet$+\not{E}_T$	20.3	3
1303.2985 (CMS) [89]	α_T+b jets	11.7	59
ATLAS-CONF-2012-104 [90]	$1\ell+\geq 4$ jets$+\not{E}_T$	5.8	2
ATLAS-CONF-2013-024 [91]	$0\ell+6$ $(2b)$ jets$+\not{E}_T$	20.5	3
ATLAS-CONF-2013-047 [92]	$0\ell+2$-6 jets$+\not{E}_T$	20.3	10
ATLAS-CONF-2013-061 [93]	0-$1\ell+\geq 3b$ jets$+\not{E}_T$	20.1	9
ATLAS-CONF-2013-062 [94]	1-$2\ell+3$-6 jets$+\not{E}_T$	20.0	19
CMS-SUS-13-016 [95]	OS $2\ell+\geq 3b$ jets	19.7	1

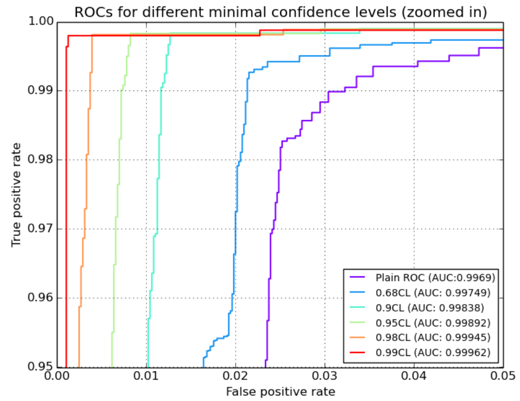
natural SUSY II

- we tested all 22k benchmark points with SUSY-AI
- we were able to reproduce the limits
- we derived somewhat better results since the procedure with CheckMATE was conservative
- there are wrong classifications but a confidence level cut provides reliable results



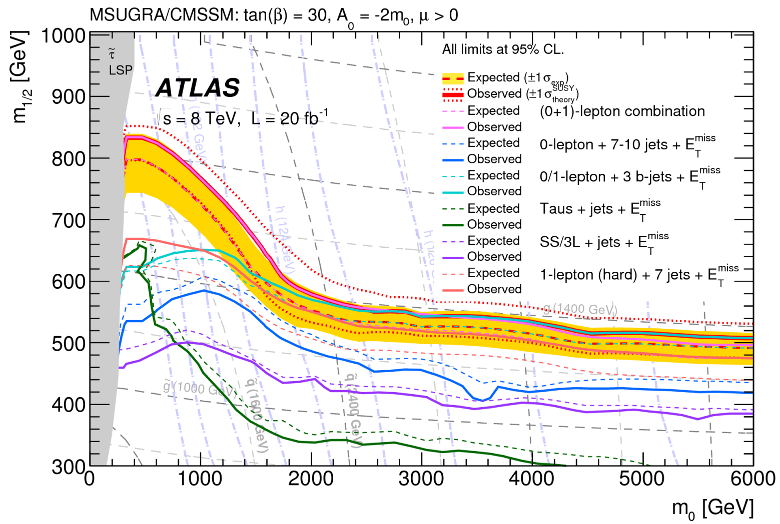


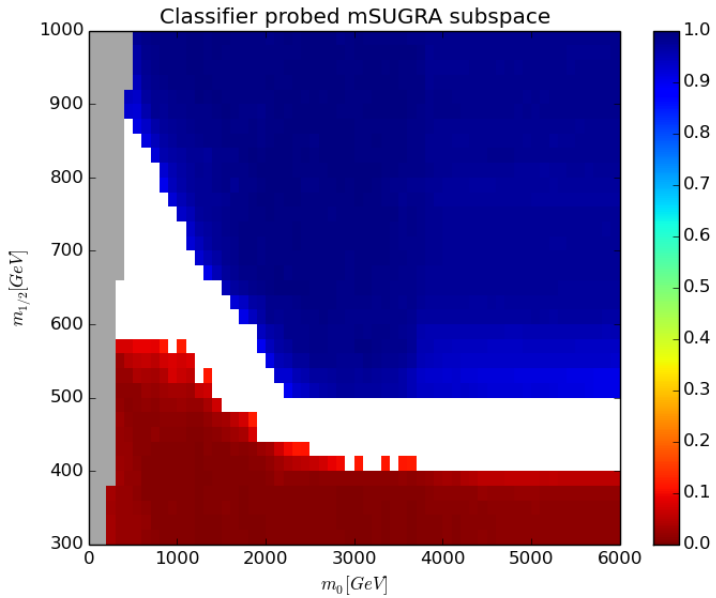




mSUGRA

- we performed a test with the constrained SUGRA model
- it has 4 and 1/2 parameters:
- m0, m12, A0, tan β and sign μ
- we set tan β=0 and A0=2m12
- all points outside of the sampling range were relocated into the sampling region





The Tool

The Tool

```
from susyai import susyai
import numpy as np

sa = susyai("susyai_classifier_python_v3.pkl")
data = np.array([[30, 4.0276e2, 7.3196e2, 2.1862e3, 1.0,
          4.0713e3, 4.4890e3, 4.4752e3, 4.4743e3, 2.8806e3,
          3.7855e3, 1.3240e3, 2.9076e3, 4.2226e3, 4.2056e3,
          3.4290e3, 3.8608e3, -4.3154e3, -8.1538e3, -7.3680e3]])
clas, pred, cert = sa.predict(data)
```

```
from susyai import susyai
import numpy as np

sa = susyai("susyai_classifier_python_v3.pkl")
sa.set_coordinate_selector(1)
sa.set_id_selector('filename')

files = ['spectrum1.slha', 'spectrum2.slha']
clas, pred, cert, coords, ids = sa.predict_files(files)
```

Outlook

BSM-AI and SUSY-AI

- we will provide classifiers for the MSSM and the NMSSM updated with 13 TeV data based on a larger training set
- we want to perform the difficult task of predicting the efficiencies/likelihoods (interesting for people performing a global fit)
- we want to include non collider constraints
- we work on providing classifiers for non-SUSY models
- ultimate goal is to consider a generic model independent approach

Conclusion

- we trained a RF classifier on over 310,000 data points of the pMSSM-19
- we used the results from the ATLAS (arXiv:1508.06608) pMSSM study
- we obtain the correct classification with an accuracy of at least 93.8%
- we will continuously update SUSY-AI with future LHC results
- we want to provide classifiers for other BSM

http://susyai.hepforge.org