

Topical Lectures on Machine learning

TMVA tutorials

Rabah Abdul Khalek

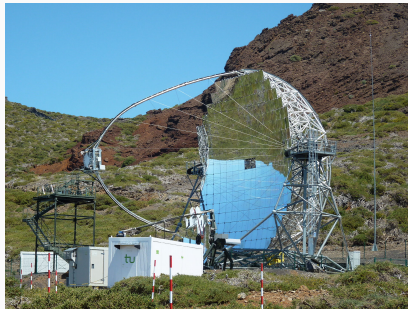
April 6, 2018

Nikhef

MAGIC Telescope

Detect and study primarily photons coming from:

- Growing black holes in active galactic nuclei.
- Supernova remnants, due to their interest as sources of cosmic rays.
- Other galactic sources such as pulsar wind nebulae or X-ray binaries.
- Unidentified EGRET or Fermi sources.
- Gamma ray bursts.
- Annihilation of dark matter.

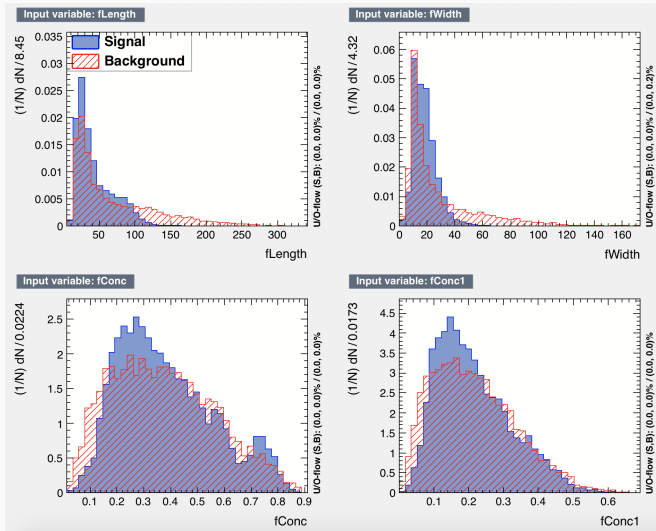


MAGIC data

The data are MC generated to simulate registration of high energy gamma particles in a ground-based atmospheric Cherenkov gamma telescope using the imaging technique.

index	variable	description
1	fLength	major axis of ellipse
2	fWidth	minor axis of ellipse
3	fSize	10-log of sum of content of all pixels
4	fConc	ratio of sum of two highest pixels over fSize
5	fConc1	ratio of highest pixel over fSize
6	fAsym	distance from highest pixel to center
7	fM3Long	3rd root of third moment along major axis
8	fM3Trans	3rd root of third moment along minor axis
9	fAlpha	angle of major axis with vector to origin
10	fDist	distance from origin to center of ellipse
	class	1 for gamma (signal), 0 for hadron (background)

MAGIC data



Summary of the steps

The main source of background are the hadronic showers initiated by cosmic rays in the upper atmosphere.

1. Partitioning data.

Total set = 19020 events.

Training set (60%) = $8632(s) + 4681(b) = 11099$.

Testing set (30%) = $2467(s) + 1339(b) = 6020$.

Validation set (10%) = 1901.

2. Build the **NN/MLP** Classifier, train it and test it.
3. Apply the trained classifier on the validation set.
4. Compare the performance of different algorithms.

Start by cloning the repo:

```
git clone  
https://github.com/rabah-khalek/TMVA_tutorials.git  
cd TMVA_tutorials
```

Hands-on coding

Start by cloning the repo:

```
git clone  
https://github.com/rabah-khalek/TMVA_tutorials.git  
cd TMVA_tutorials
```

Step1: Load the shower data into TTrees:

```
git checkout step1
```

You should be able to create three files in data/: `signal.root`, `background.root` and `validation.root`.

Hands-on coding

Start by cloning the repo:

```
git clone  
https://github.com/rabah-khalek/TMVA_tutorials.git  
cd TMVA_tutorials
```

Step1: Load the shower data into TTrees:

```
git checkout step1
```

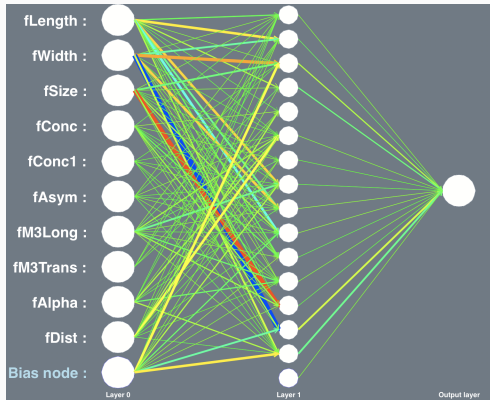
Step1: Check the solution:

```
git checkout step1_solution
```


Hands-on coding

Step2: Build the Classifier, train it and test it.

```
git checkout step2
```



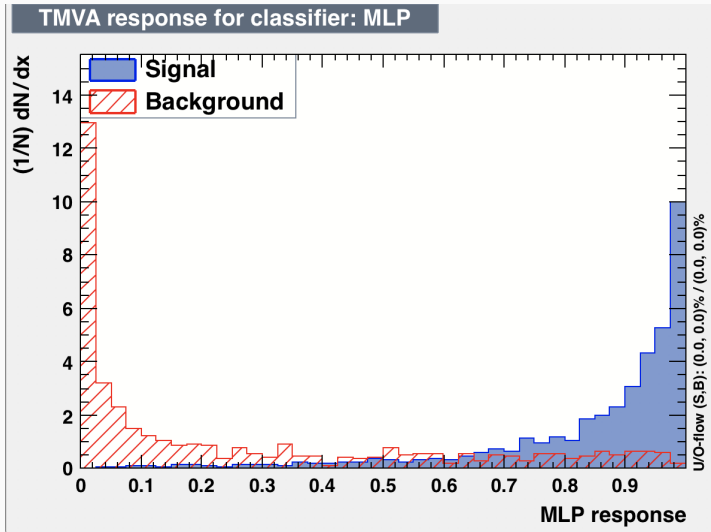
Step2: Build the Classifier, train it and test it.

```
git checkout step2
```

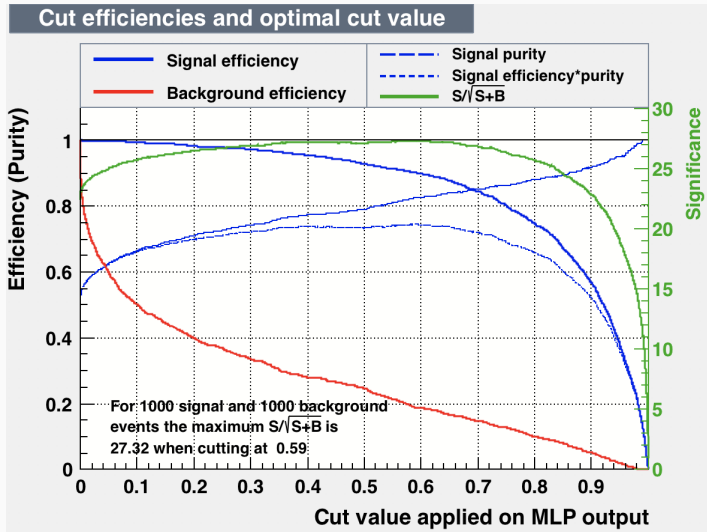
Step2: Check the solution

```
git checkout step2_solution
```

classifier output on testing set



Cuts on the classifier



Step3: Apply the trained classifier on the validation set.

```
git checkout step3
```

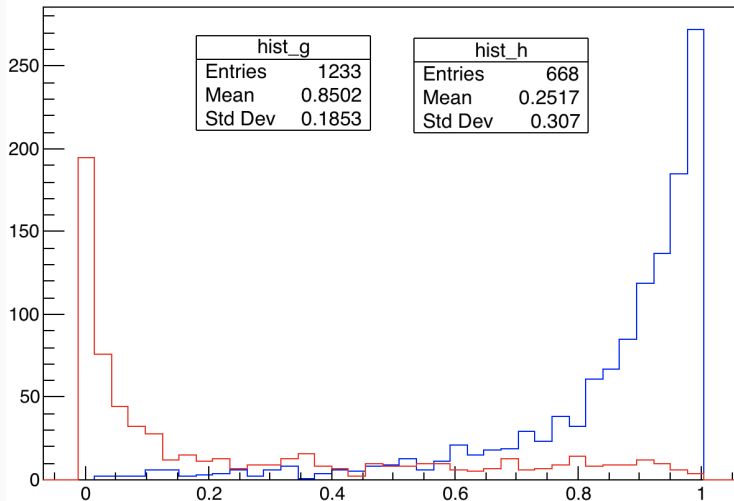
Step3: Apply the trained classifier on the validation set.

```
git checkout step3
```

Step3: Check the solution

```
git checkout step3_solution
```

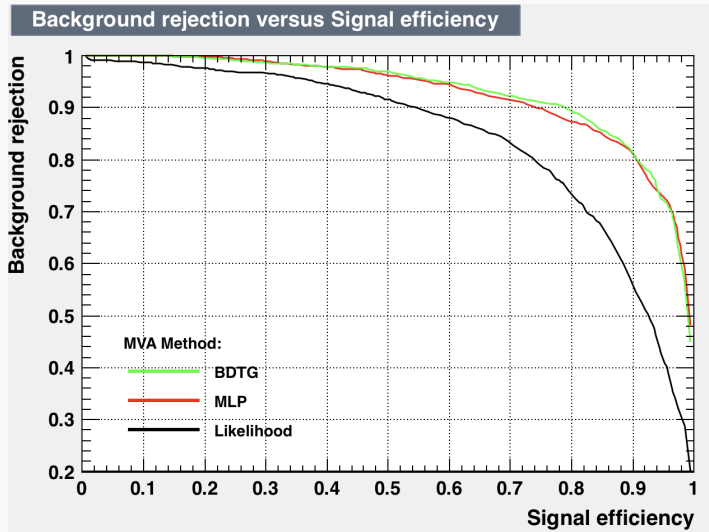
Results - Application on validation set



Step4: Compare the performance of different algorithms.

```
git checkout full_version
```


ROC curve - the classifier performance



Thank you

Thank you!