Shallow and deep machine learning applications in KM3NeT

Stefan Geißelsöder – ECAP, University of Erlangen

2nd ASTERICS-OBELICS Workshop

16-19 October 2017, Barcelona, Spain.

H2020-Astronomy ESFRI and Research Infrastructure Cluster
(Grant Agreement number: 653477).

KM3NeT

Selected Machine Learning tasks

- Up/Down classification
- Track/Shower classification
- Selectfit
- Particle identification

Deep Learning

KM3NeT

KM3NeT

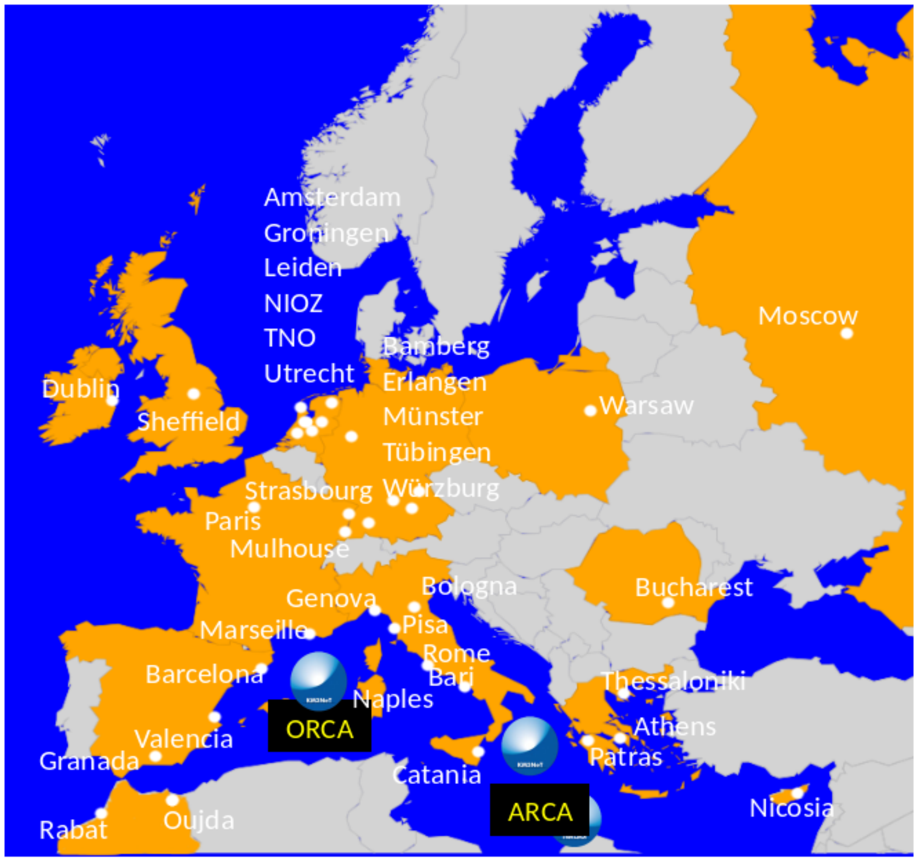
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

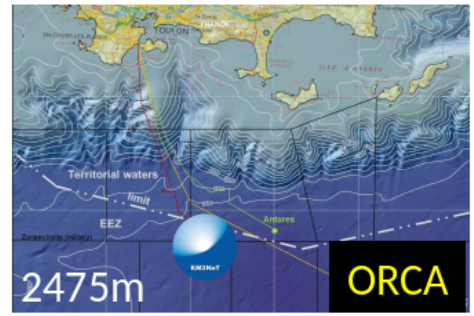
NATURWISSENSCHAFTLICHE FAKULTÄT

KM3NeT collaboration

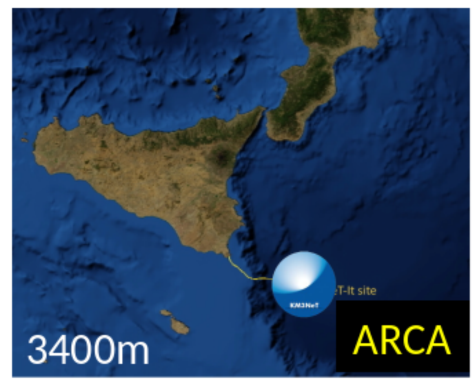
Single collaboration, single technology, multi-site infrastructure



+Nantes, Johannesburg, Marrakech, Tbilisi

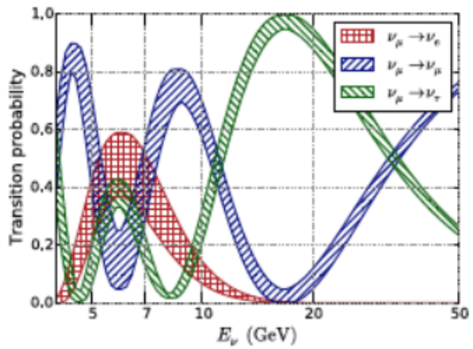


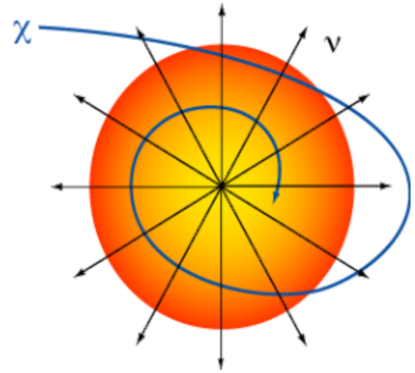
Oscillation **R**esearch with **C**osmics In the **A**byss

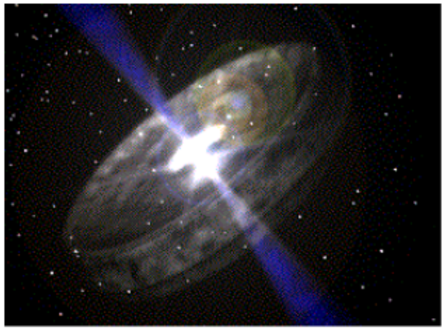


Astroparticle **R**esearch with **C**osmics In the **A**byss

KM3NeT science scope







Low Energy	Medium Energy	High Energy
MeV < E_ν < 100 GeV	10 GeV < E_ν < 10 TeV	E_ν > 10 TeV

ν Mass hierarchy

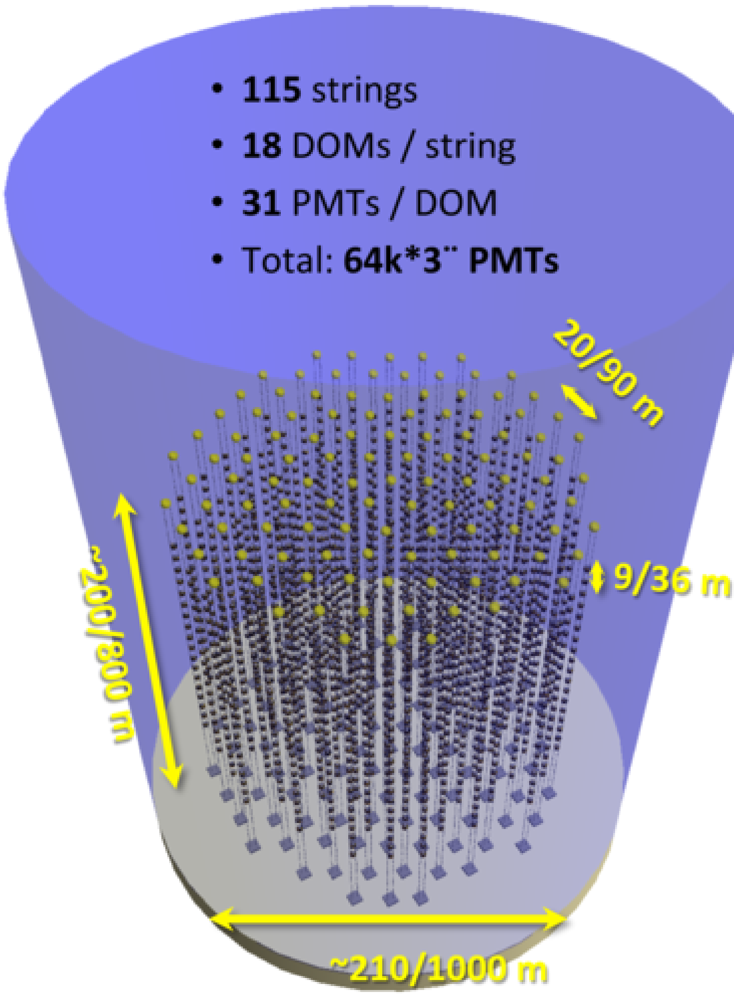
ν Oscillations

Cosmic ν

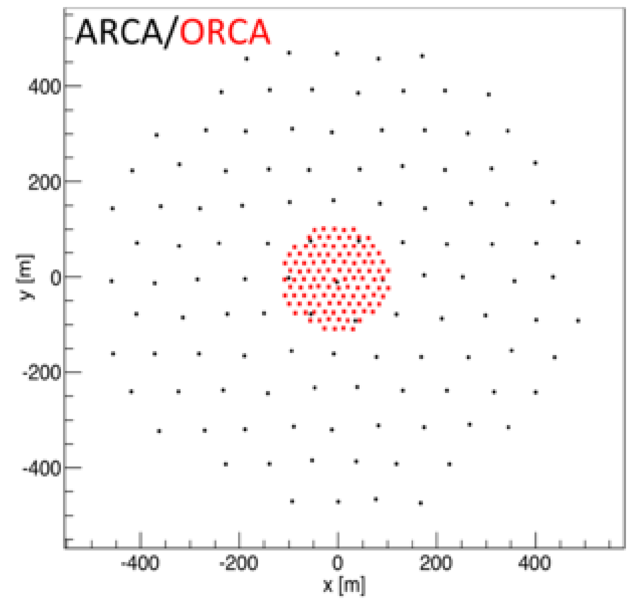
Origin and production mechanism of HE CR

KM3NeT-ORCA ANTARES KM3NeT-ARCA

KM3NeT building block

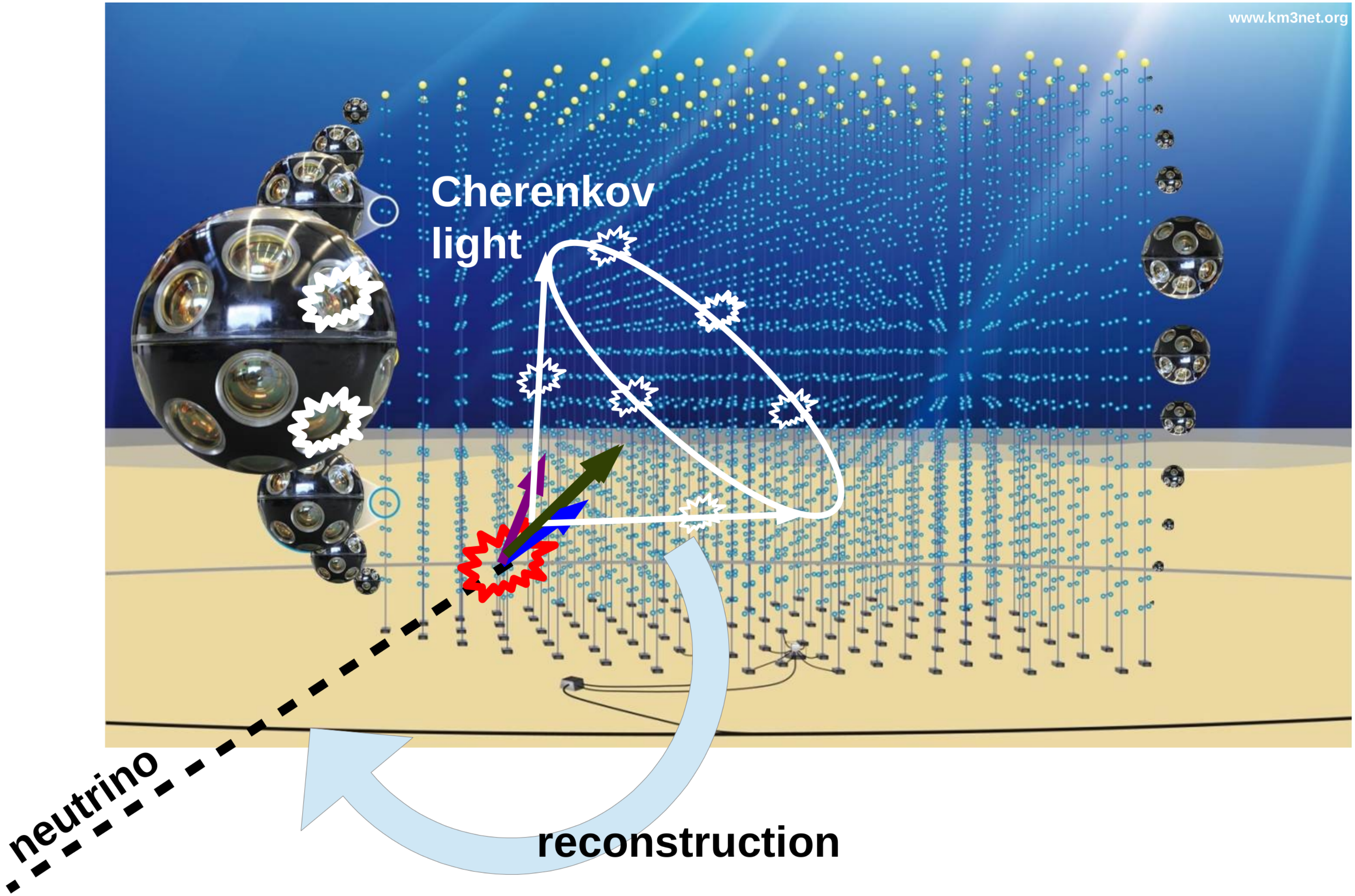


(For comparison, not build within)



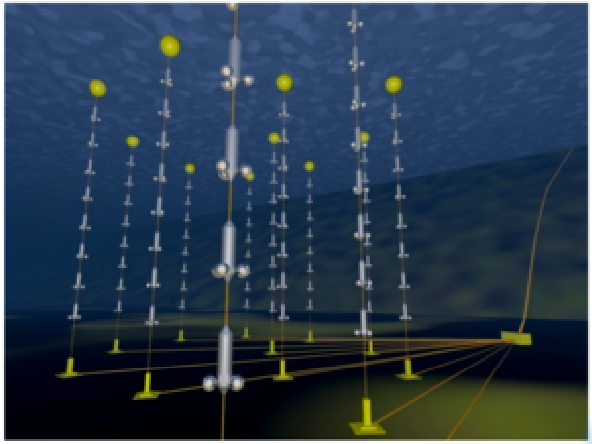
	ORCA	ARCA
String spacing	20 m	90 m
OM spacing	9 m	36 m
Depth	2470 m	3400 m
Instrumented mass	5.7 Mton	2*500 Mton

Detection principle



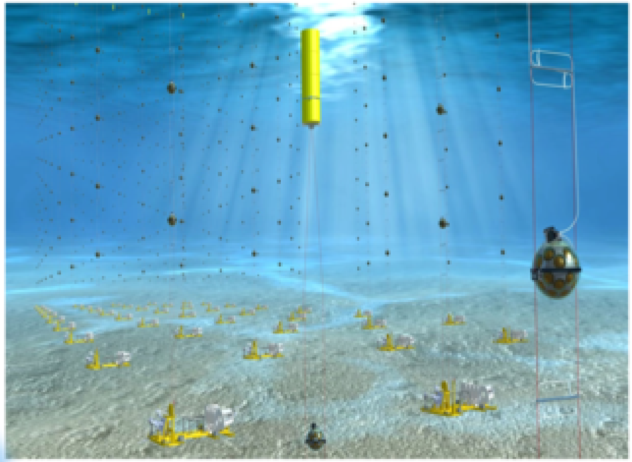
Hardware evolution

12 lines



ANTARES

3 Building Blocks (3*115 lines)



- Uniform angular coverage
- Directional information

KM3NeT

Selected Machine Learning tasks

KM3NeT

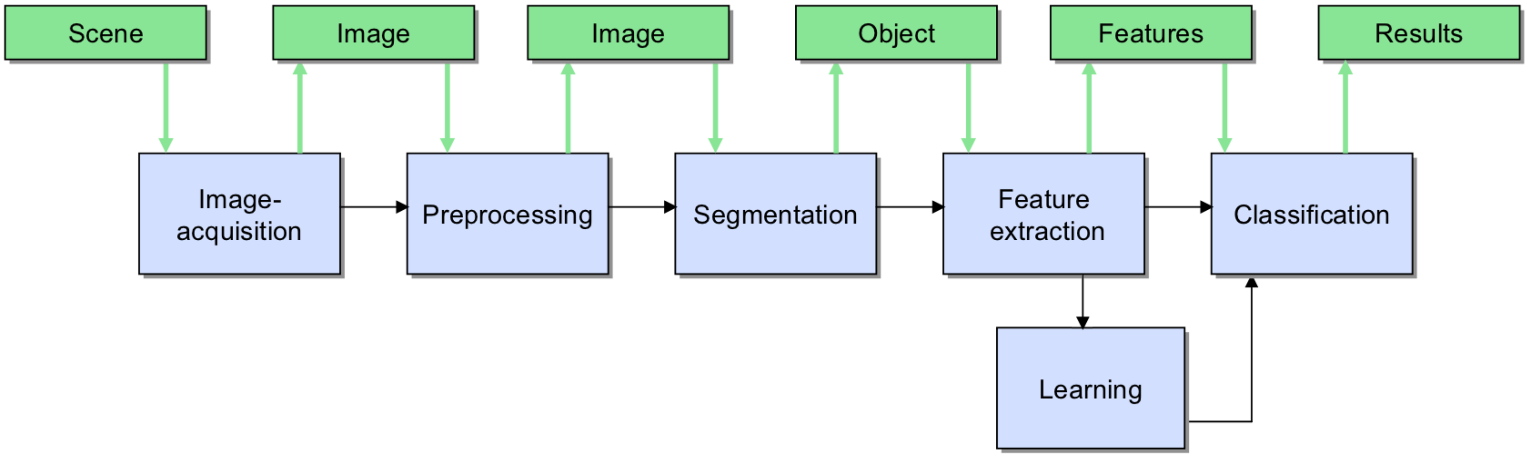
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

NATURWISSENSCHAFTLICHE FAKULTÄT

Pattern recognition pipeline

“Classical” pipeline for images:

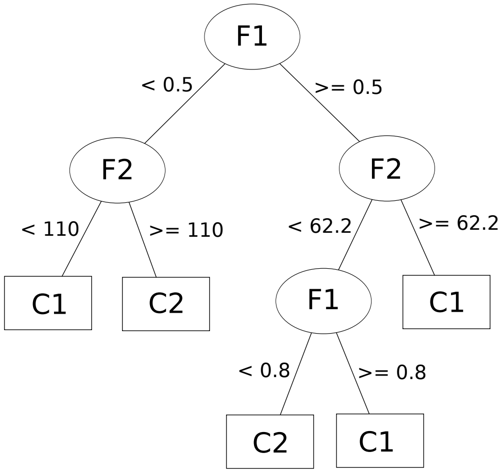


Machine Learning algorithms

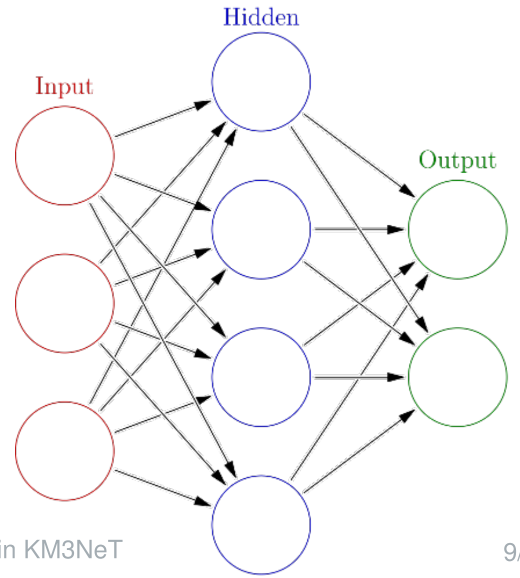
Machine learning:
Algorithm uses experience to improve
(= minimize error function on examples)

- Random Decision Forests
- Boosted Decision Trees
- Artificial Neural Networks
- Convolutional Neural Networks
- ...

Decision tree

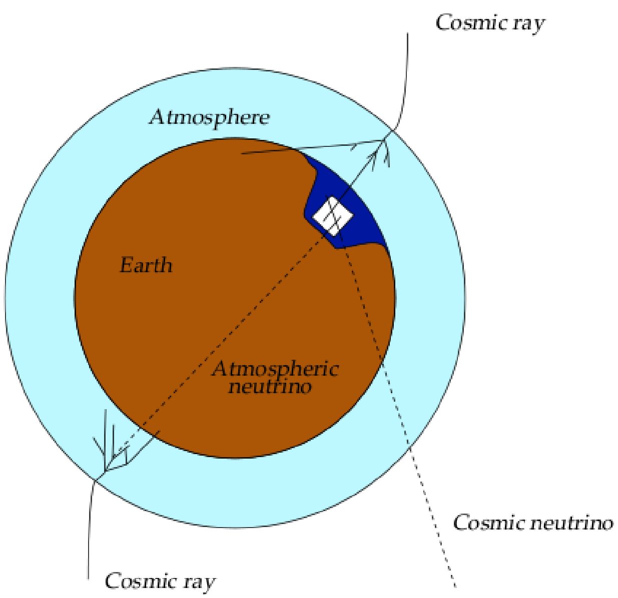


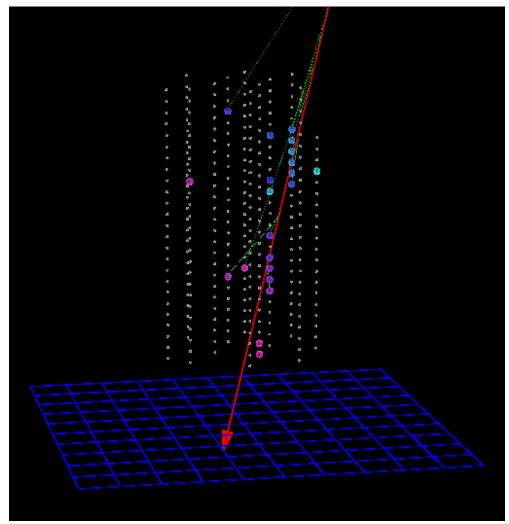
Feed forward neural network:

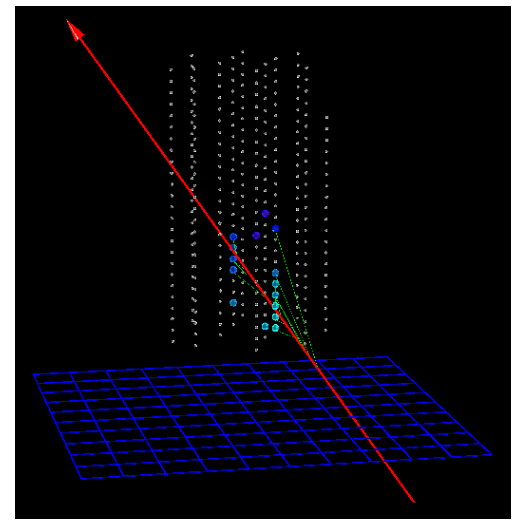


Up/Down classification

Example from ANTARES



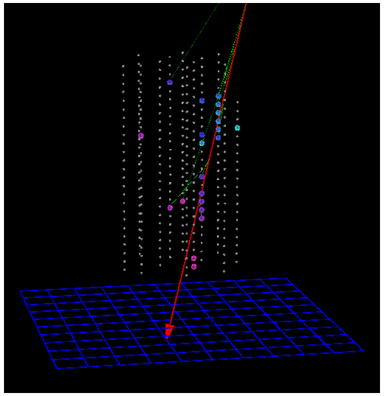


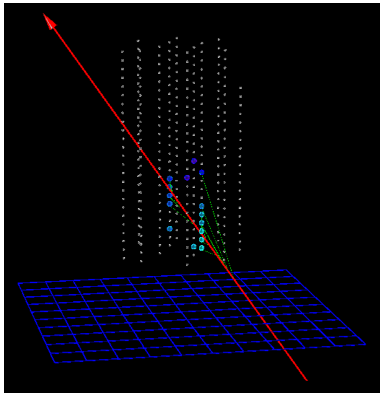


$\approx 10^6$ more atmospheric muons than neutrinos, all from above

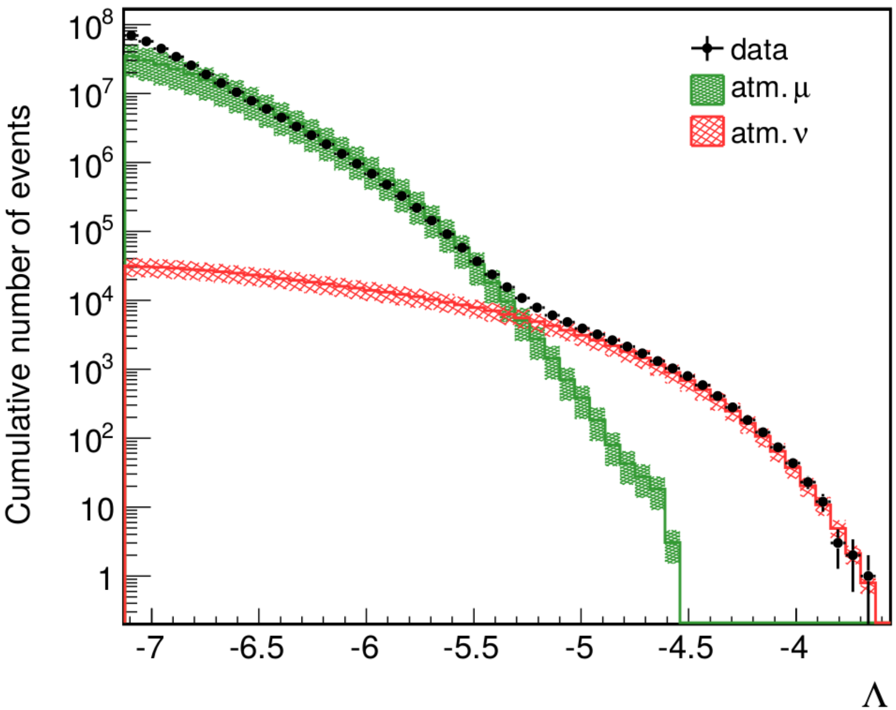
$\rightarrow$ For several analyses: Up/Down = Signal/background classification

Up/Down classification

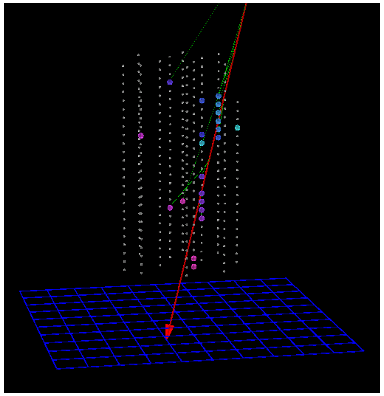


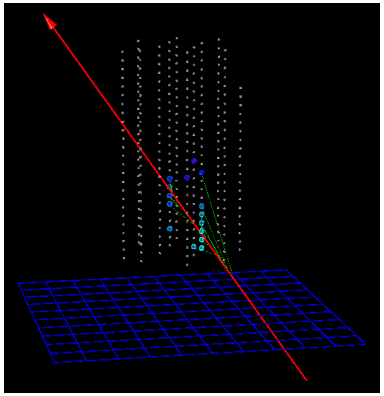


- Obvious solution: cuts on quality parameters of direction reconstruction

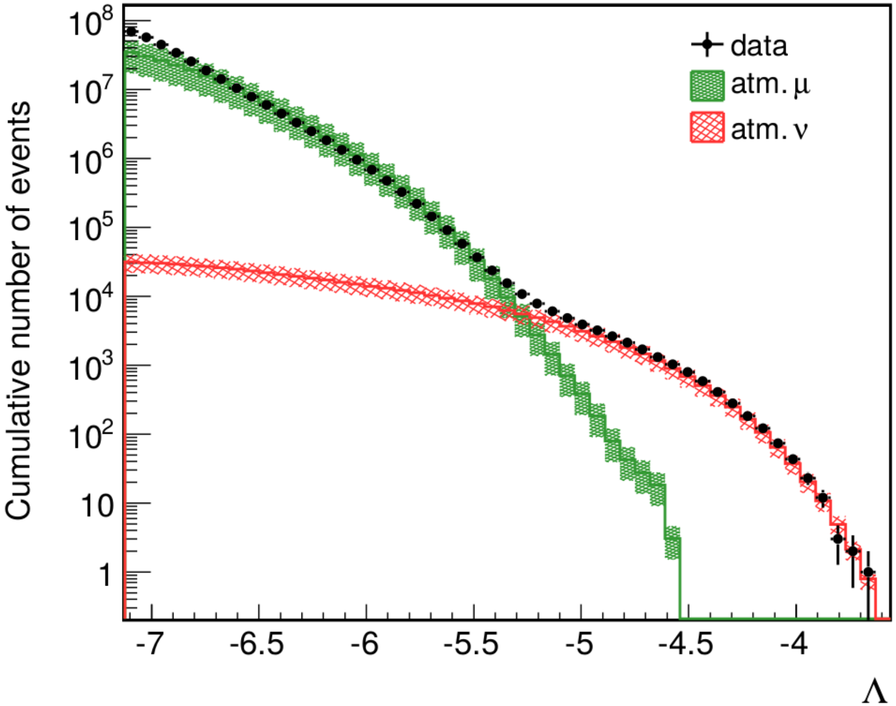


Up/Down classification

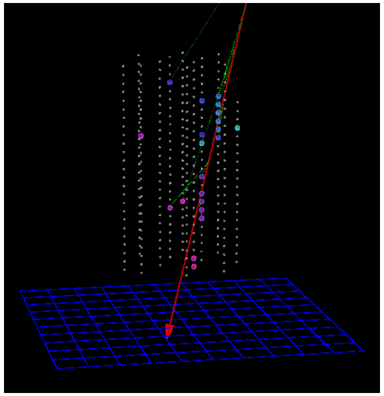


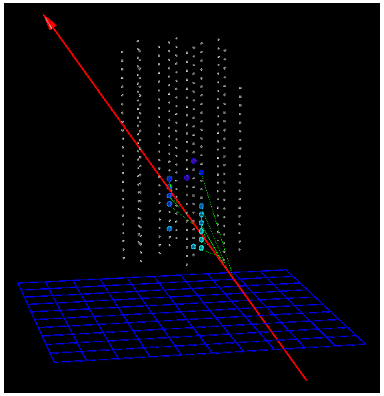


- Obvious solution: cuts on quality parameters of direction reconstruction
- Result:
  Loose most neutrinos!

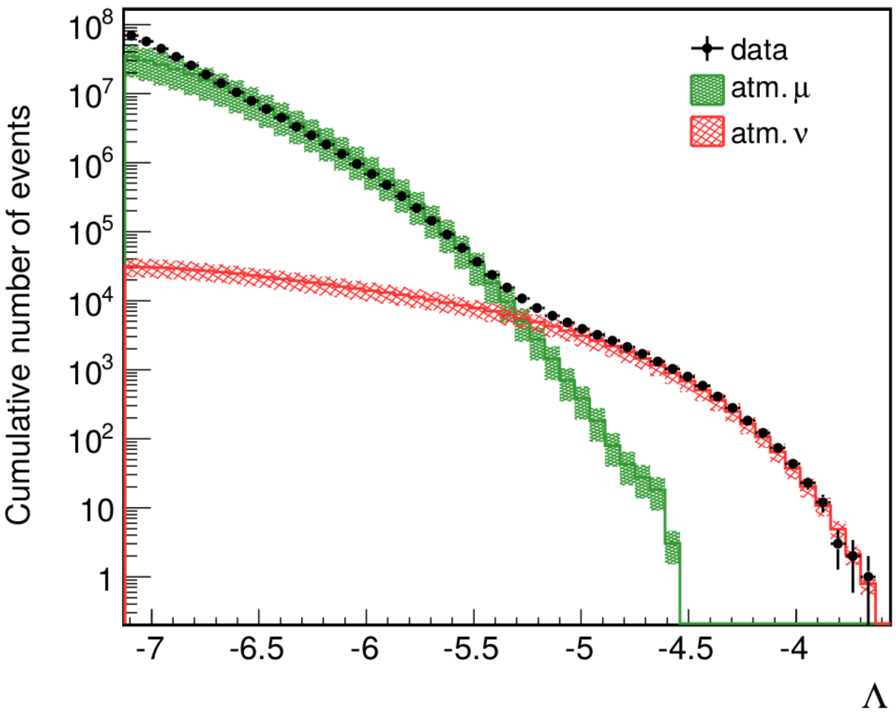


Up/Down classification



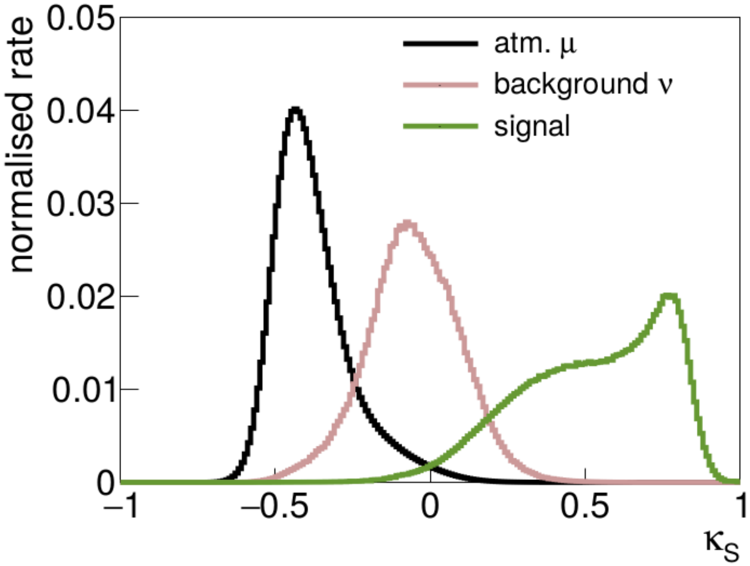


- Alternative: Random Decision Forest using more parameters (reconstruction, specifically designed and other algorithms)
- $\approx$ 99% muon suppression
- Not enough alone, but allows less strict cuts $\rightarrow$ more signal



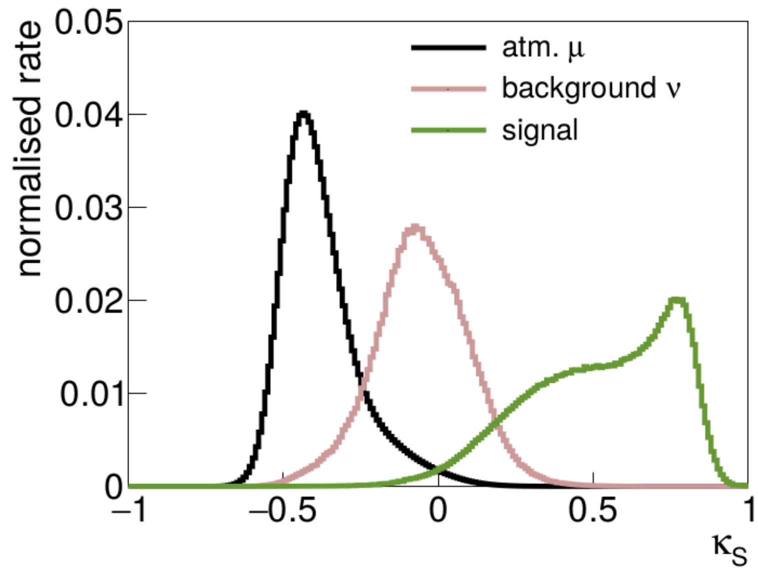
Track/Shower classification

- Different signatures requires different reconstructions
- Classes are:
  - Tracks a)
  - Showers c) and d)
- 11 features and a BDT



Track/Shower classification

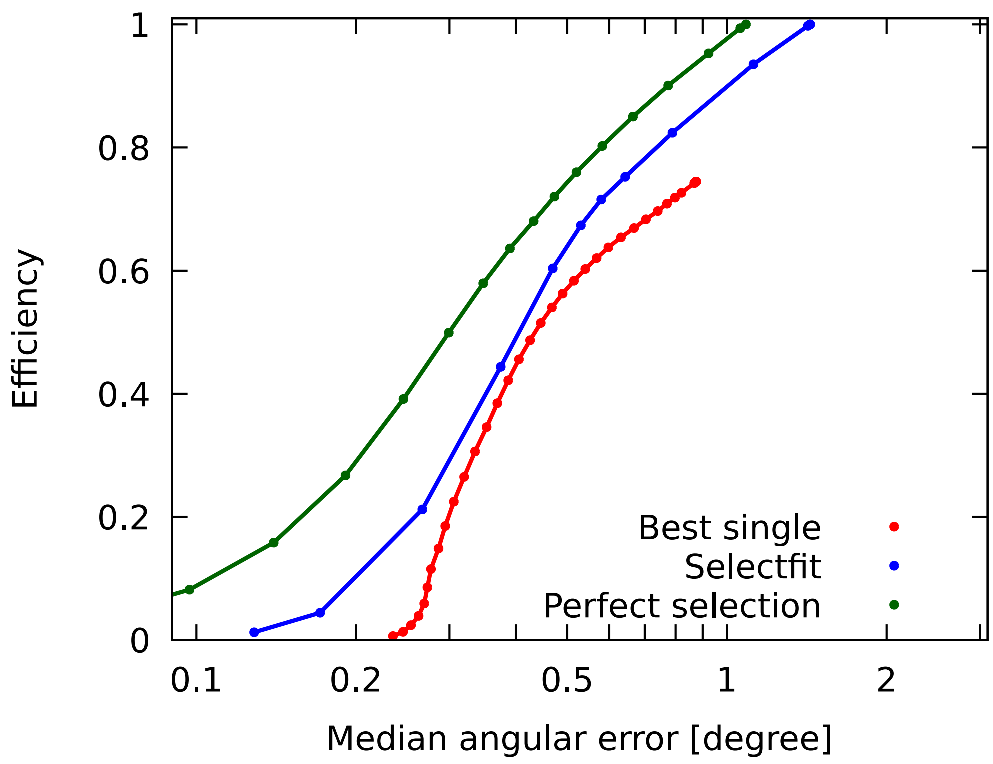
- Different signatures requires different reconstructions
- Classes are:
  - Tracks a)
  - Showers c) and d)
- 11 features and a BDT
- More efficient selection than using quality cuts for track and shower reconstructions



Selectfit

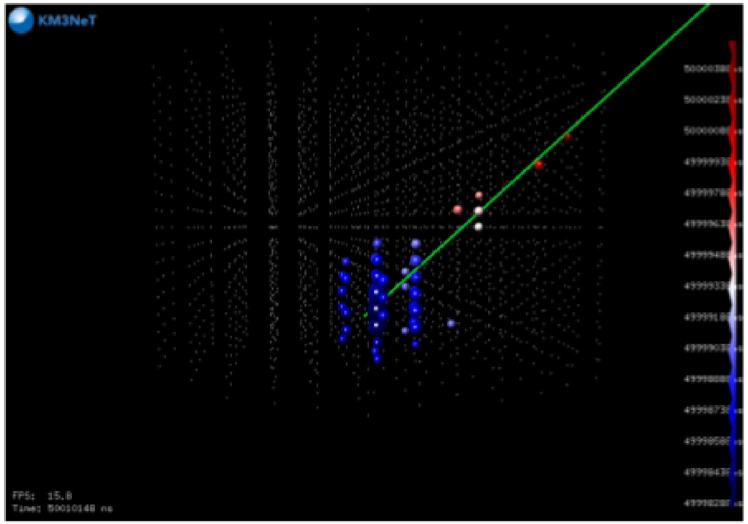
Example from ANTARES

- Multiple direction reconstructions available
- Combine them by RDF with reconstruction results and quality parameters
- Improves reconstruction accuracy/efficiency
- Allows combining different topologies

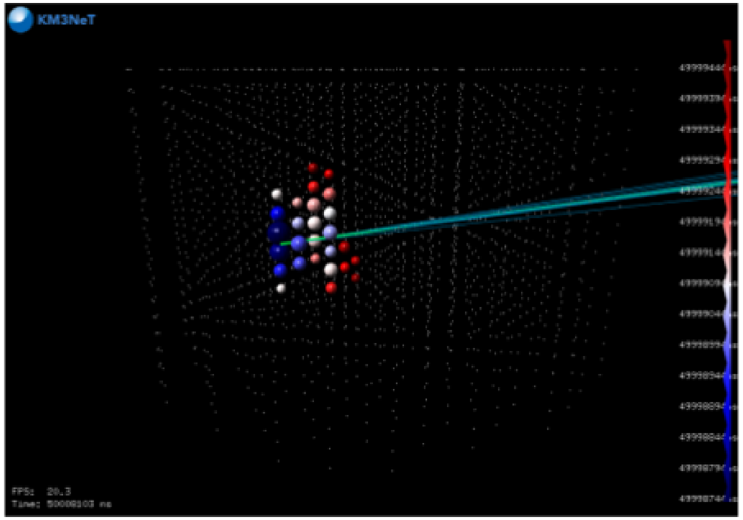


Particle identification

- Incorporate more knowledge e.g. for flavour composition
- Distinguish
  - Down-going tracks
  - Up-going tracks
  - Starting tracks
  - Cascade events
  - ν_τ double bang
- Again: more efficient than just stacking cuts
- Optimizing features and classification becomes complex



(d) starting track (ν_μ-CC, E_{ν_μ}=91 TeV, y=0.8)



(e) cascade (ν_μ-NC, E_{ν_μ}=1.06 PeV, y=0.6)

Deep Learning

KM3NeT

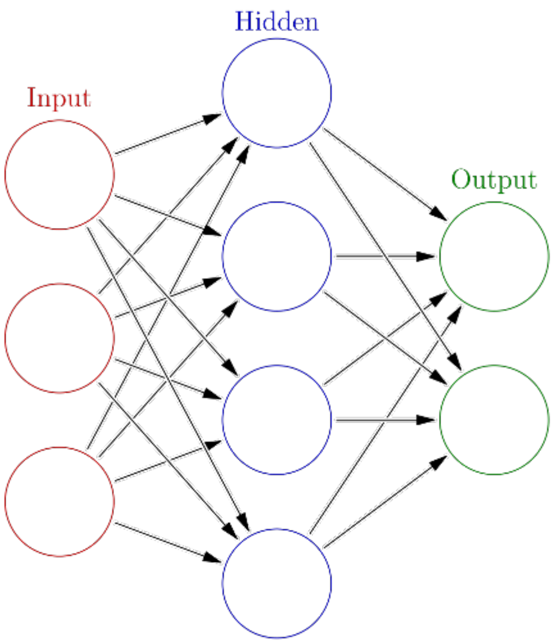
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

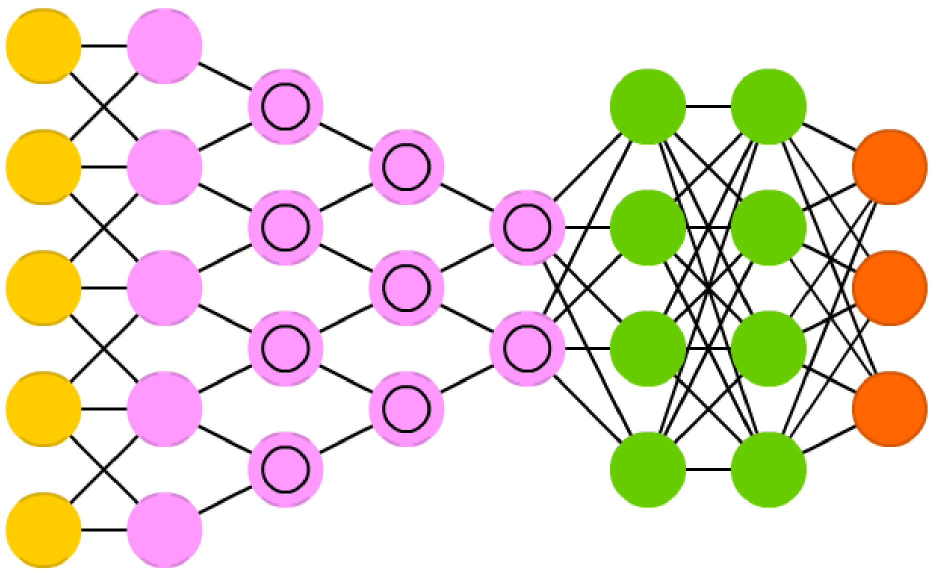
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

NATURWISSENSCHAFTLICHE FAKULTÄT

Deep Learning

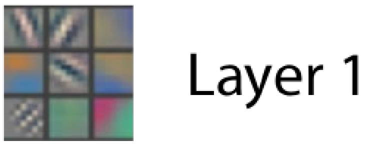
- In many architectures deep means many layers

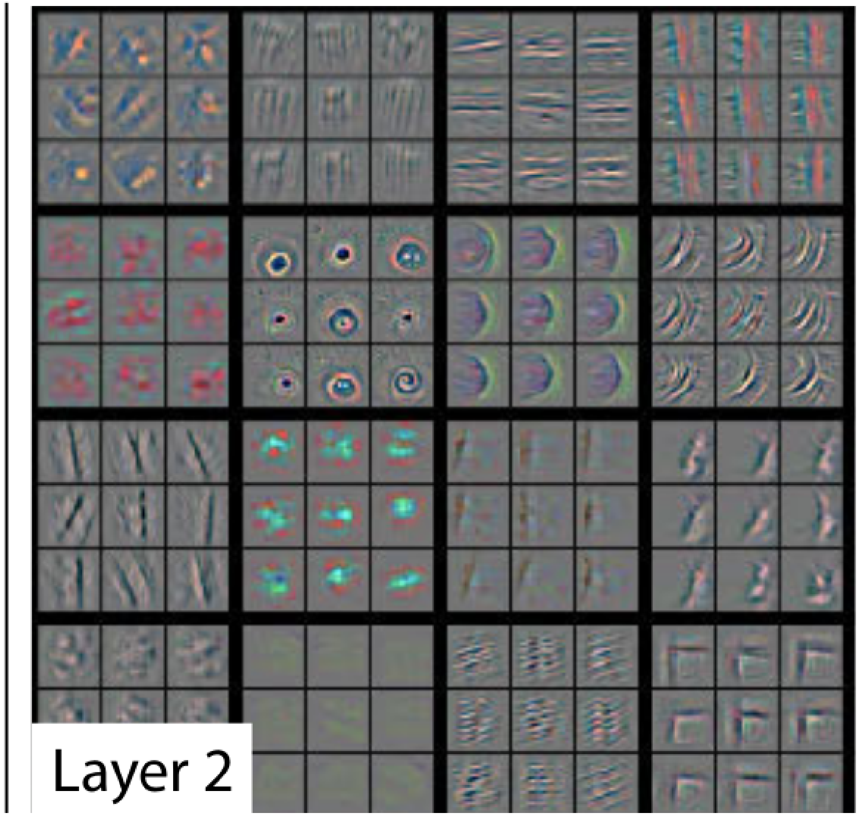




Deep Learning

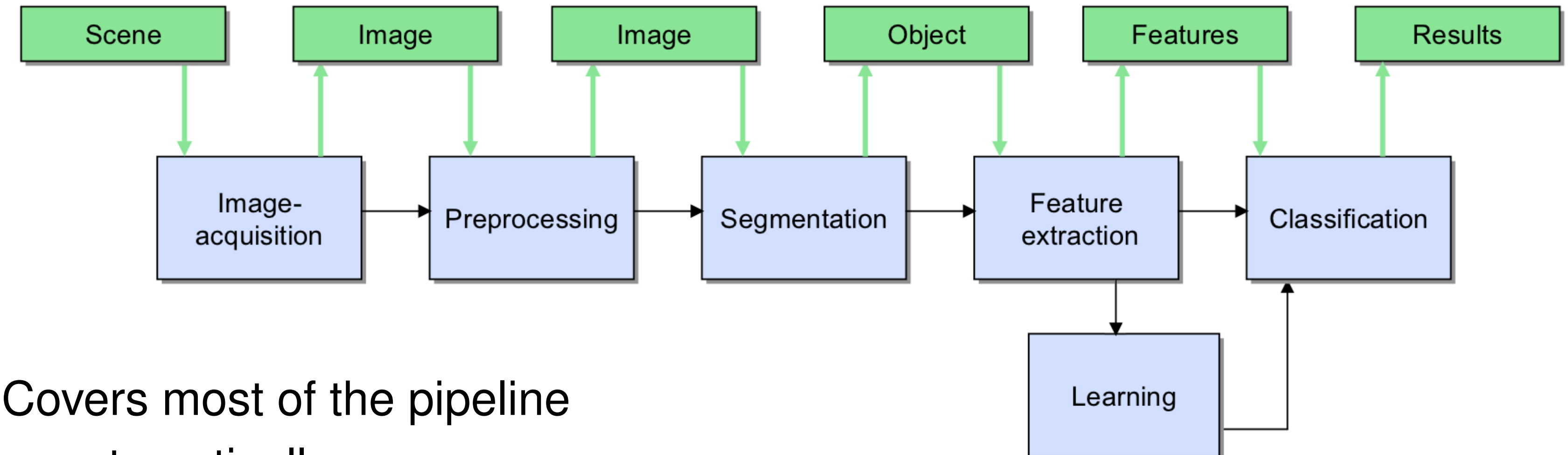
- In many architectures deep means many layers
- Can abstract from simple input
  $\rightarrow$ learns its own features ("Representation learning")





Deep Learning

- In many architectures deep means many layers
- Can abstract from simple input
  $\rightarrow$ learns its own features ("Representation learning")

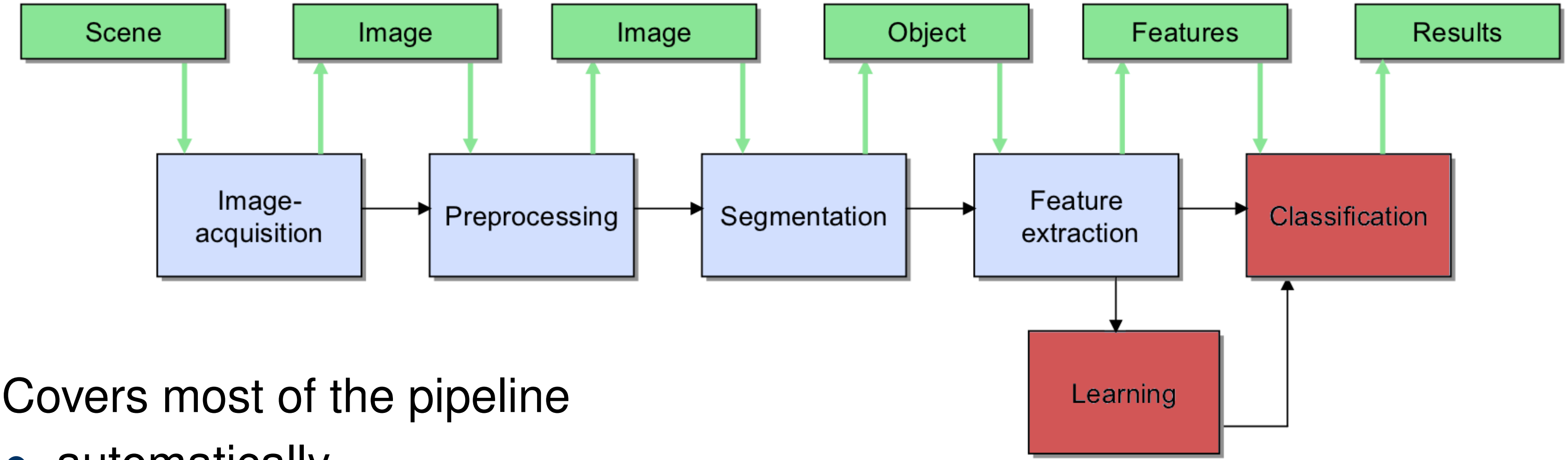


Covers most of the pipeline

- automatically
- sometimes better than experts

Deep Learning

- In many architectures deep means many layers
- Can abstract from simple input
  $\rightarrow$ learns its own features ("Representation learning")

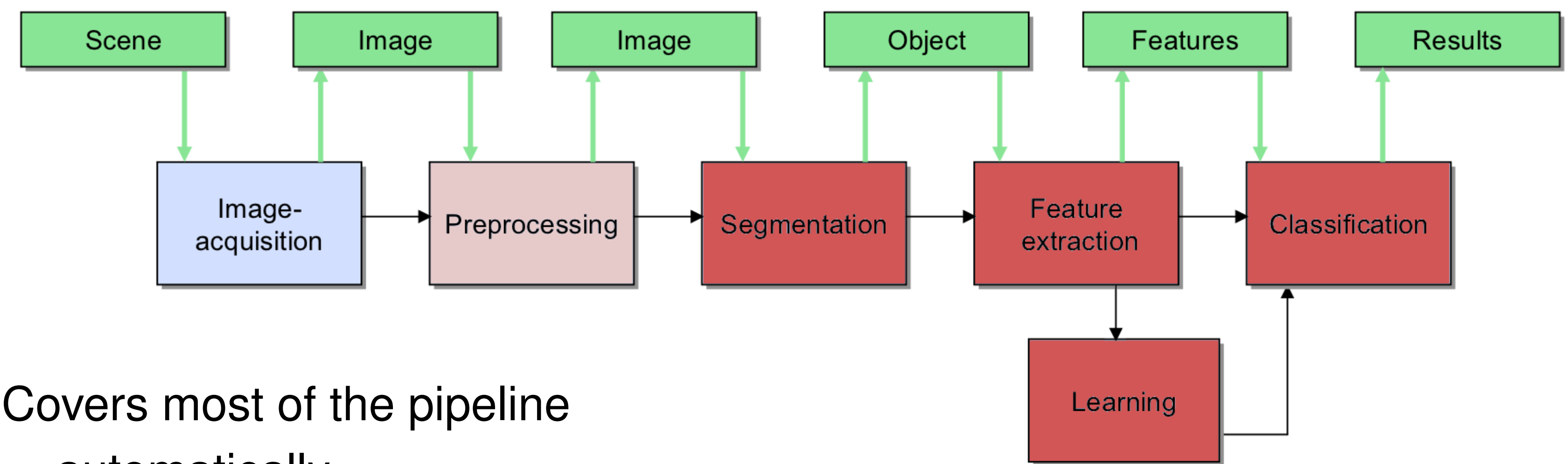


Covers most of the pipeline

- automatically
- sometimes better than experts

Deep Learning

- In many architectures deep means many layers
- Can abstract from simple input
  $\rightarrow$ learns its own features (“Representation learning”)

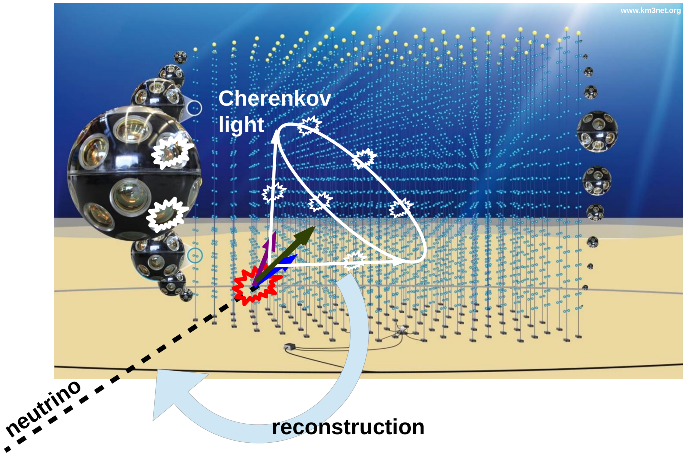


Covers most of the pipeline

- automatically
- sometimes better than experts

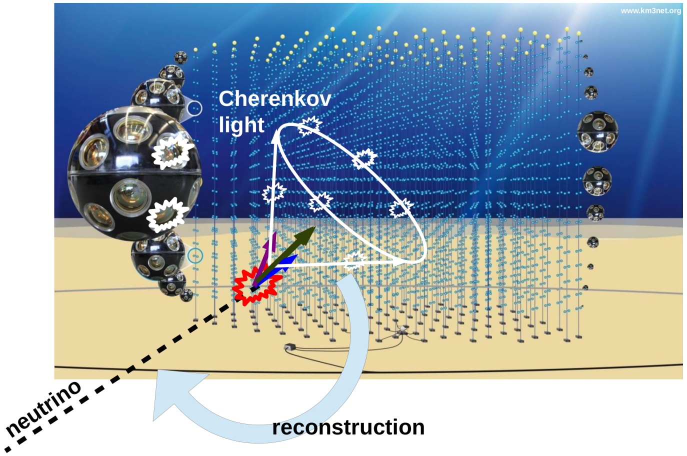
Deep Learning

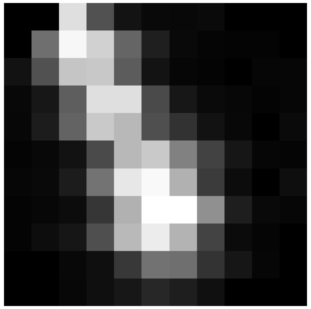
- Deep Learning frameworks optimized for 2D data
- Neutrino telescopes produce 4D data
- KM3NeT produces 6D data
- How to use Deep Learning?



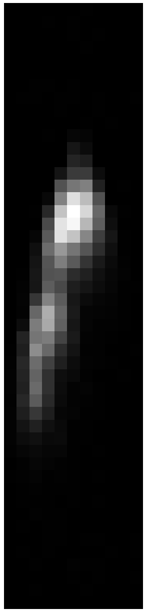
Deep Learning

- Deep Learning frameworks optimized for 2D data
- Neutrino telescopes produce 4D data
- KM3NeT produces 6D data
- How to use Deep Learning?
- Use projections?
- 2D, 3D, 4D and 3.5D
- Direct and transformed input (non-cartesian, time residuals)





x - y



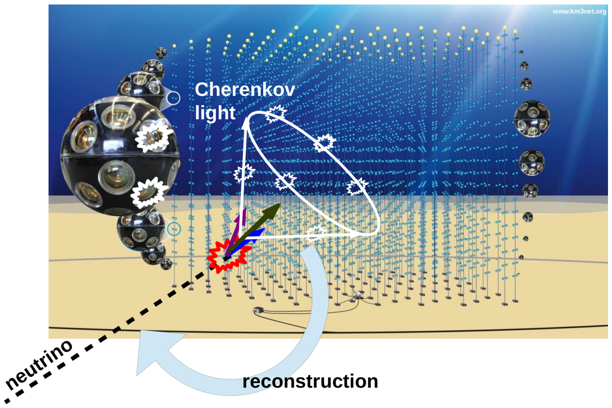
x - t

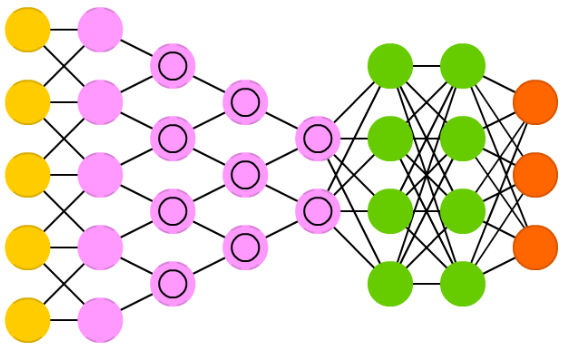
Deep Learning

- Testing CNNs, Residual nets, LSTMs with Keras, Tensorflow, CNTK
- Training benefits from as much data as possible
- Training takes hours or days
- No manual feature design $\rightarrow$ can be trained for all previous tasks plus new ones (regression of direction, bjorken y)
- Best results so far with CNNs / Residual nets with direct 3.5D
- Preliminary & where compared already: at least as good as previous machine learning
- Doesn't (yet) surpass algorithms that can incorporate a single, well computable physics scenario (e.g. track direction)
- Outlook: Investigate 6D data, Autoencoders, combination of CNN and LSTM, separated convolutions

Summary

- Neutrino astronomy often deals with high backgrounds and low signal statistics
- Machine Learning is more efficient than plain cuts on variables
- Deep Learning allows to tackle tasks hardly possible before
- KM3NeT investigates these new techniques to enhance the sensitivity achievable by various analyses





Thank you for your attention