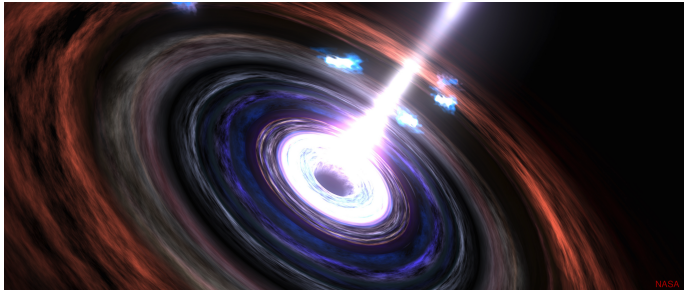
technische universität
dortmund

lehrstuhl
physik e5



Study on the classification of blazars with machine learning techniques using multiwavelength information

Konstantin Pfrang

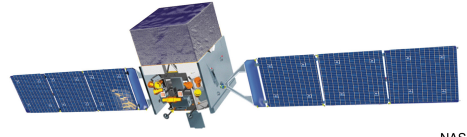
October 8, 2016

School for Astroparticle Physics 2016

Motivation

3FGL catalog

- *Fermi* LAT
- 3033 γ-ray sources
- Energies from 100 MeV to 300 GeV



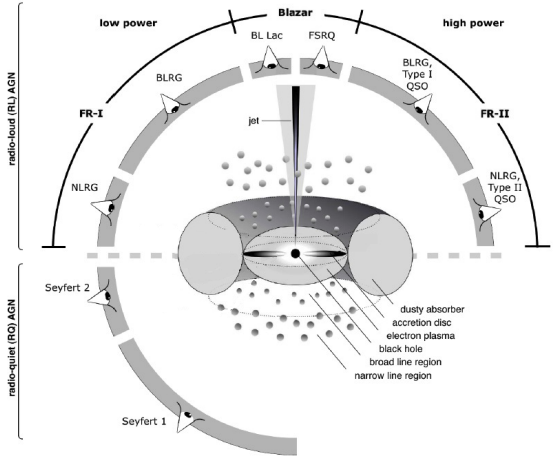
NASA

	# objects
AGNs	1744
non-AGNs	279
unassociated	1010
AGNs of uncertain type	573

$\Rightarrow$ Classification using machine learning

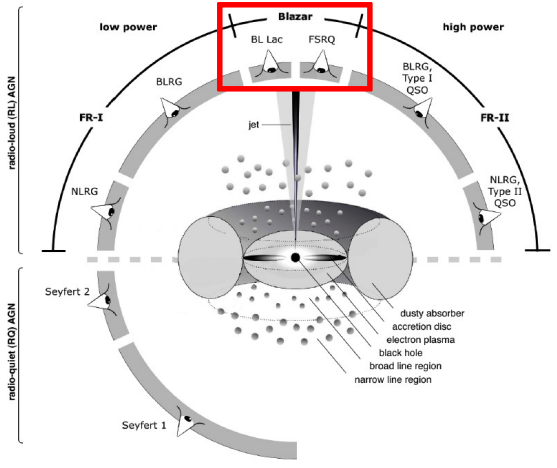
How is the performance of the classification using MWL information?

Active Galactic Nuclei



Beckmann & Shrader (2012)

Active Galactic Nuclei



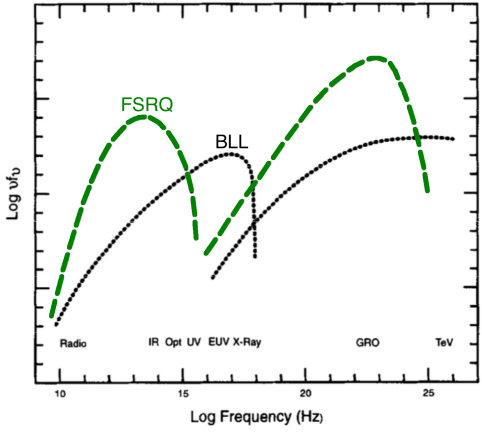
Beckmann & Shrader (2012)

Blazars

- Classification in BLLs and FSRQs
- FSRQs more luminous

Differentiation based on fluxes

Spectral Energy Distribution (SED)

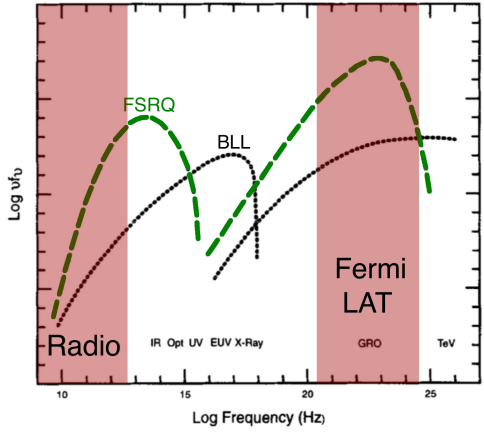


Multiwavelength approach for AGN classification

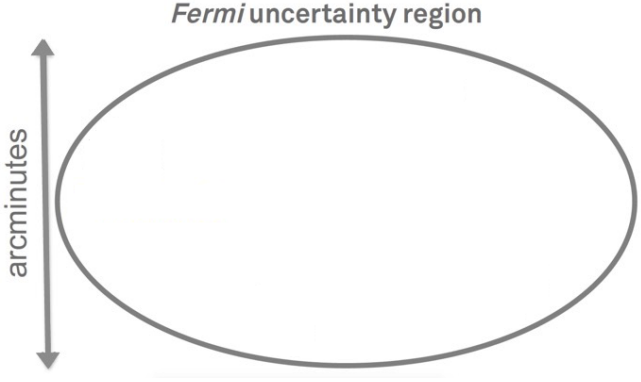
- Experiments not able to record whole SED
- Remove restriction to one energy level

$\Rightarrow$ Search for counterparts

Spectral Energy Distribution (SED))

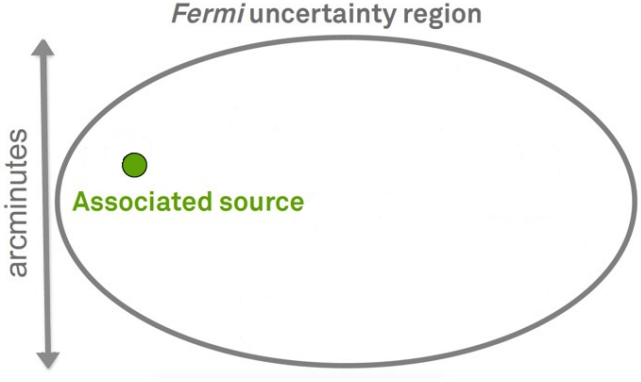


Search for counterparts



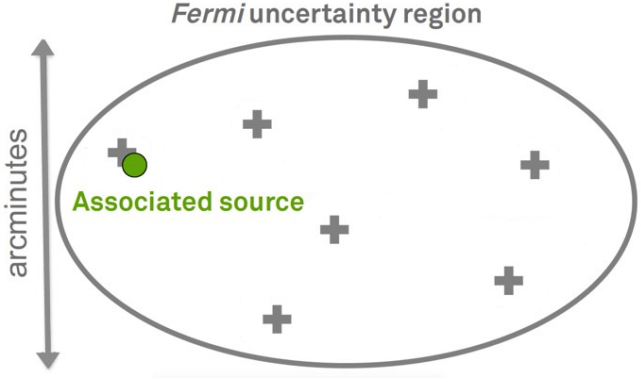
1. **catalog:** 3FGL
   - 95% uncertainty region

Search for counterparts



1. **catalog:** 3FGL
   - 95% uncertainty region
   - associated source in other catalog

Search for counterparts

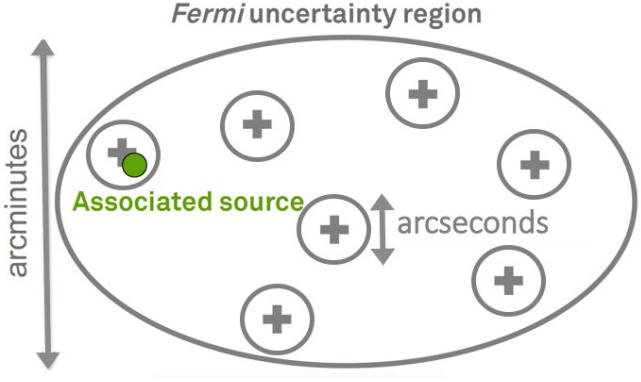


radio sources

2. catalog: NVSS

- NRAO VLA Sky Survey
- 1.8 million sources at 1.4 GHz

Search for counterparts

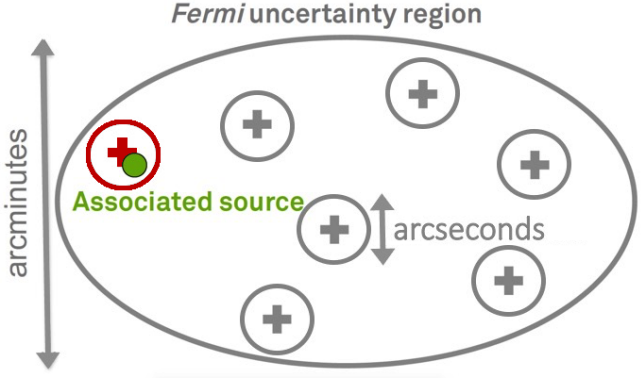


radio sources

2. catalog: NVSS

- NRAO VLA Sky Survey
- 1.8 million sources at 1.4 GHz
- Uncertainties significant smaller

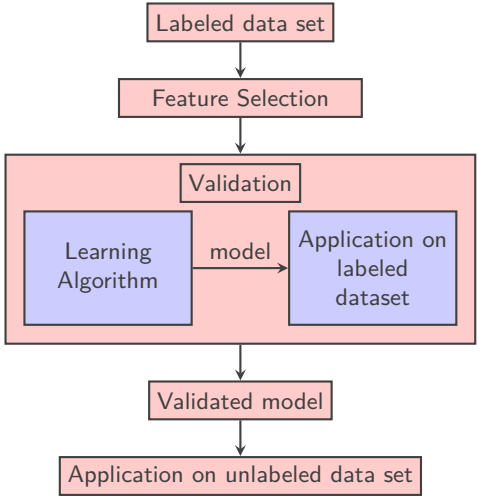
Search for counterparts



Association of 3FGL to NVSS counterpart!

Classification using data mining

Validation of model essential!



Preparation of data set

- Counterpart search

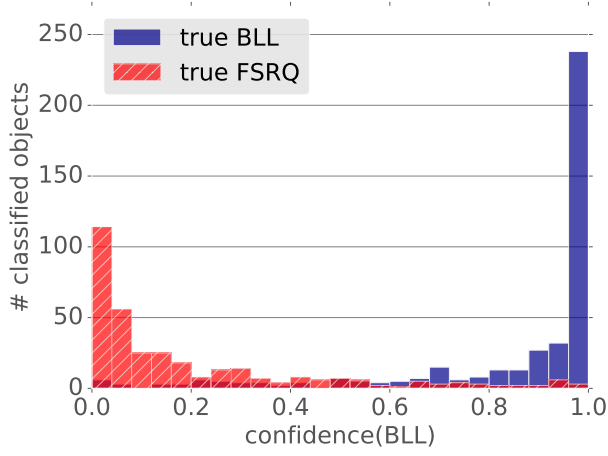
total Blazars	**766**
BLLs	422
FSRQs	344

- Feature generation through combination of different wavelengths
- Feature selection to reduce number of attributes

total attributes	**345**	
3FGL	28	
NVSS	2	Selected attributes
combined	20	

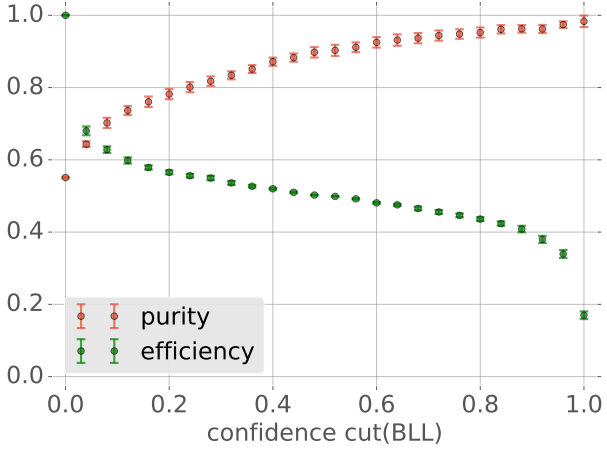
Performance estimation

- Confidence cut:
  least score for classification
- Performance heavily dependent on
  choice of confidence cut



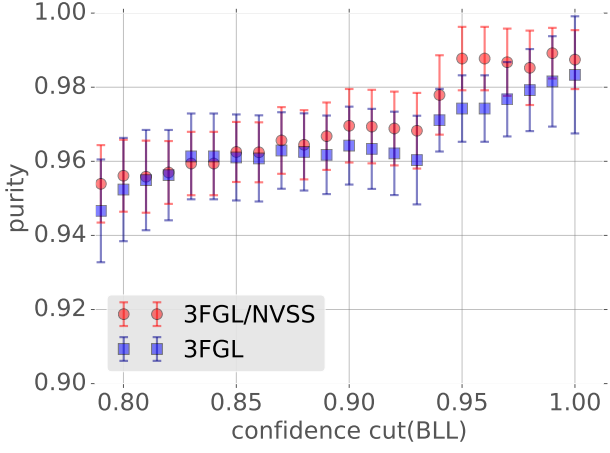
Performance estimation

- Purity:
  Ratio of correctly and falsely classified objects
- Efficiency:
  Ratio correctly classified events to total number of the class



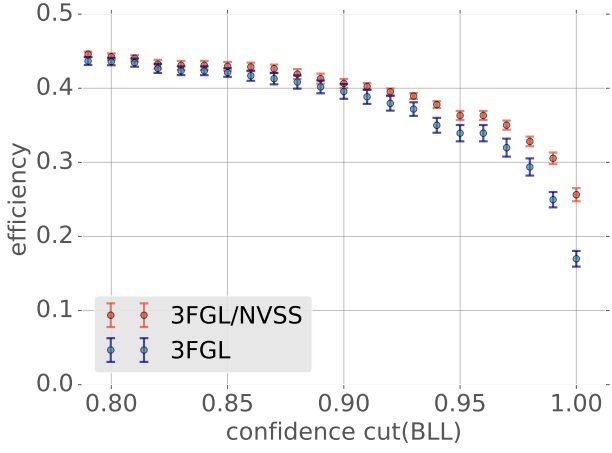
⇒ Trade off between efficiency and purity

Performance estimation



Performance estimation

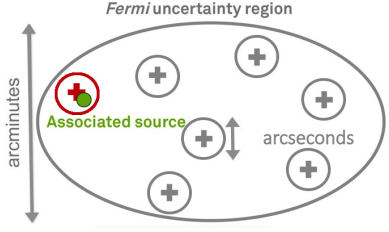
	3FGL	3FGL/NVSS
purity	≥ 0.97	
# BLLs	226 ± 2	283 ± 3
# FSRQs	68 ± 1	94 ± 1

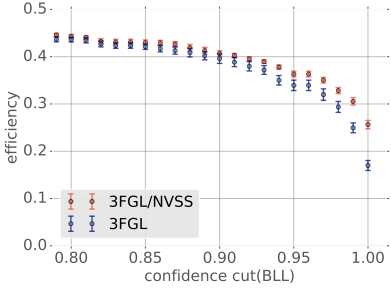


$\Rightarrow$ Tremendous improvement using MWL information!

Conclusion

- Method for counterpart search to 3FGL objects
- Classification of blazars using machine learning
- Improved efficiency using MWL information

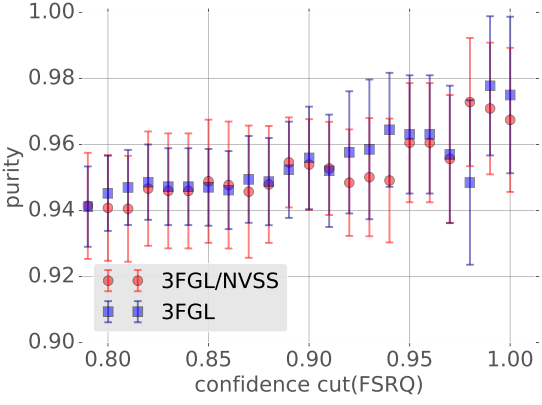


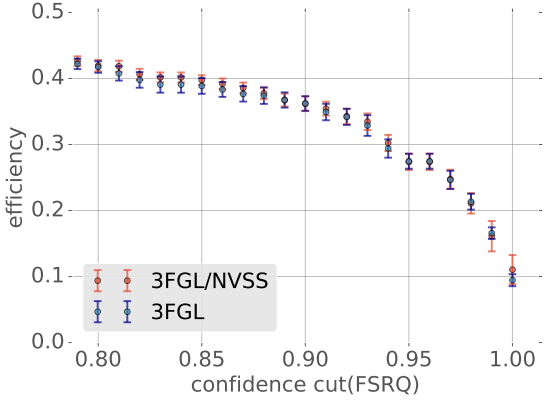


Outlook

- Extension to more wavelengths
- Possibility for counterpart search for unassociated sources

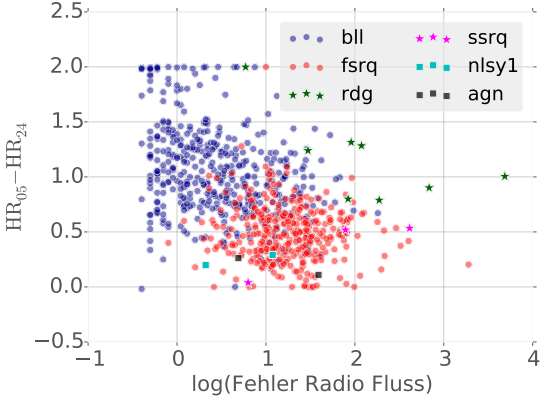
Backup - quality FSRQ-classification

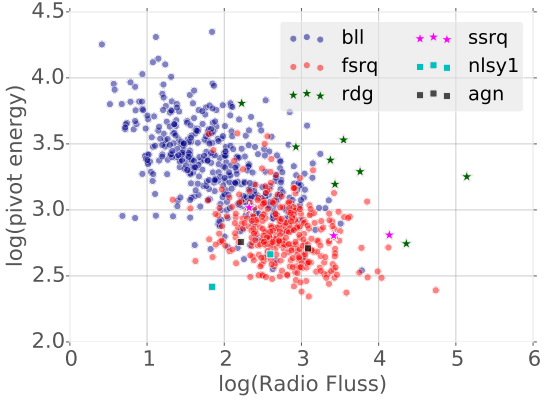




- FSRQ classification not as efficient and pure as for BLLs
- Possible reason: subclasses of BLLs (LBL's SED very similar to FSRQ's!)

Backup


bll
fsrq
rdg
ssrq
nlsy1
agn
$HR_{05} - HR_{24}$
log(Fehler Radio Fluss)


bll
fsrq
rdg
ssrq
nlsy1
agn
log(pivot energy)
log(Radio Fluss)