Multi-frequency variabilities: blazar classification and statistical properties

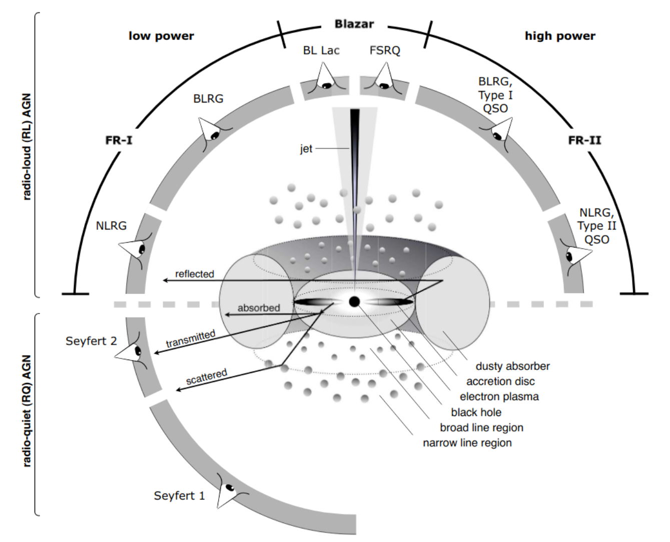
Rattanapong Yoyponsan, Utane Sawangwit, Siramas Komonjinda

What are blazars ?

Blazars are ...

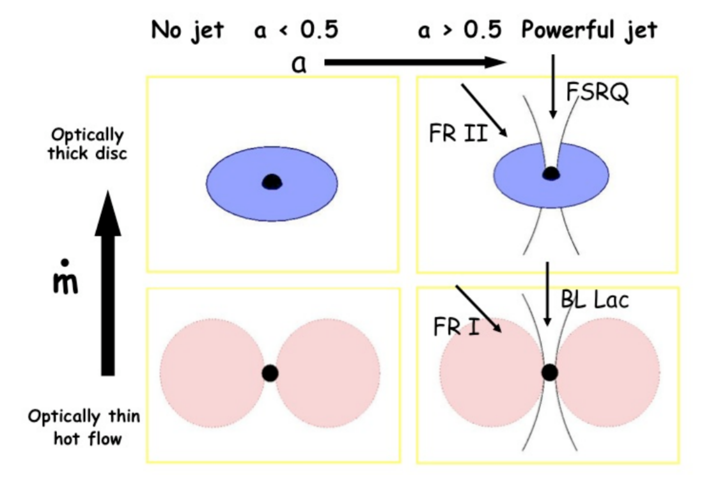
- The Active Galactic Nuclei (AGNs) which their jet direct at the observer’s line-of-sight
- Multi-frequency emitters of electromagnetic radiation (from radio to gamma-ray
- Highly variable in brightness

AGN unification scheme (Beckmann & Shrader 2012).



Blazars are classified into 2 classes based on emission lines

- **BL Lacs** have narrow emission lines (EW < 5 Angstrom) or no emission lines
- **FSRQs** have wider emission line (EW > 5 Angstrom)



Geometry of accretion disk and how it implicates blazar classification
(Maraschi et al. 2012)

- **BL Lacs** have optically thin geometrically thick hot flow disk
- **FSRQs** have optically thick geometrically thin disk

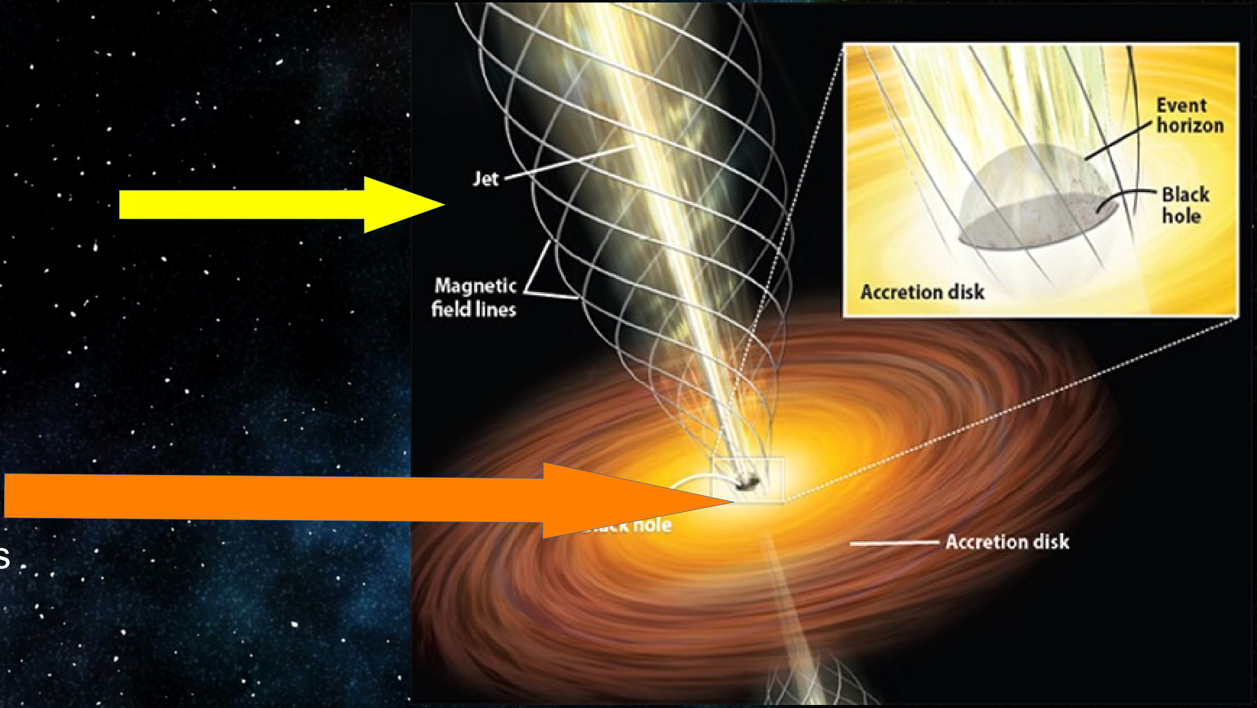
Radiation Mechanisms

Non-thermal

- Synchrotron radiation
- Inverse Compton scattering
- Synchrotron self-compton
- External compton

Thermal

- Black body radiation
- Emission from infalling particles

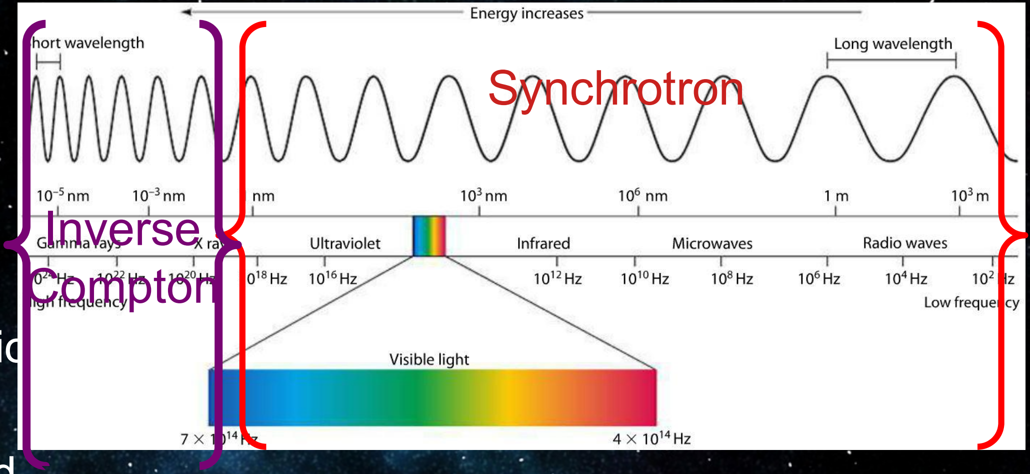


http://www.daviddarling.info/encyclopedia/E/event_horizon.html

Credit & ©: Astronomy / Roen Kelly

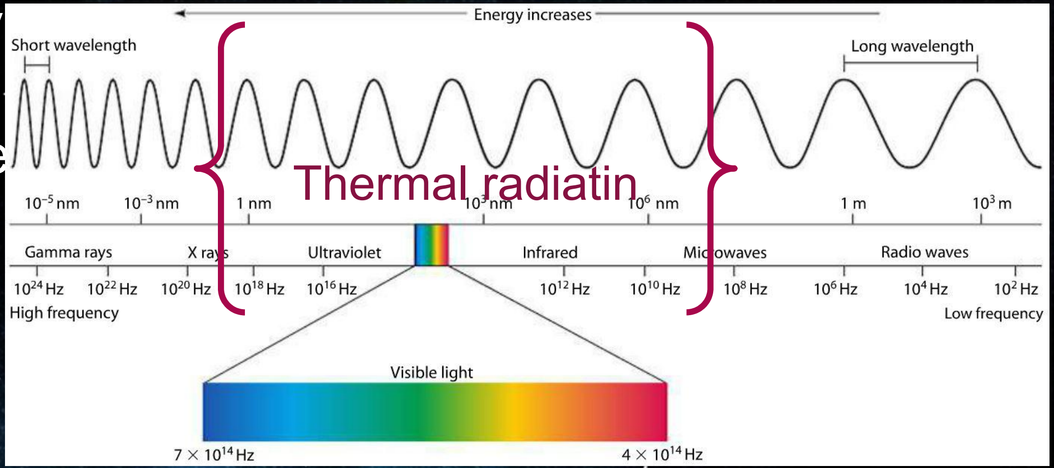
Radiation Mechanisms

- Non-Thermal
  - **Synchrotron** – electron accelerated in magnetic field
    - From radio to x-ray (<10^11 ~10^18)
  - **Inverse Compton Scattering** – photons are energized by high energy particles
    - **Synchrotron self-compton (SSC)**-- The seed photons are from synchrotron radiatio
    - **External compton (EC)** – The seed photons are from outside the jet, i.e. broad-line region, inter stellar medium etc.
    - From x-ray to gammaray (10^18 >)

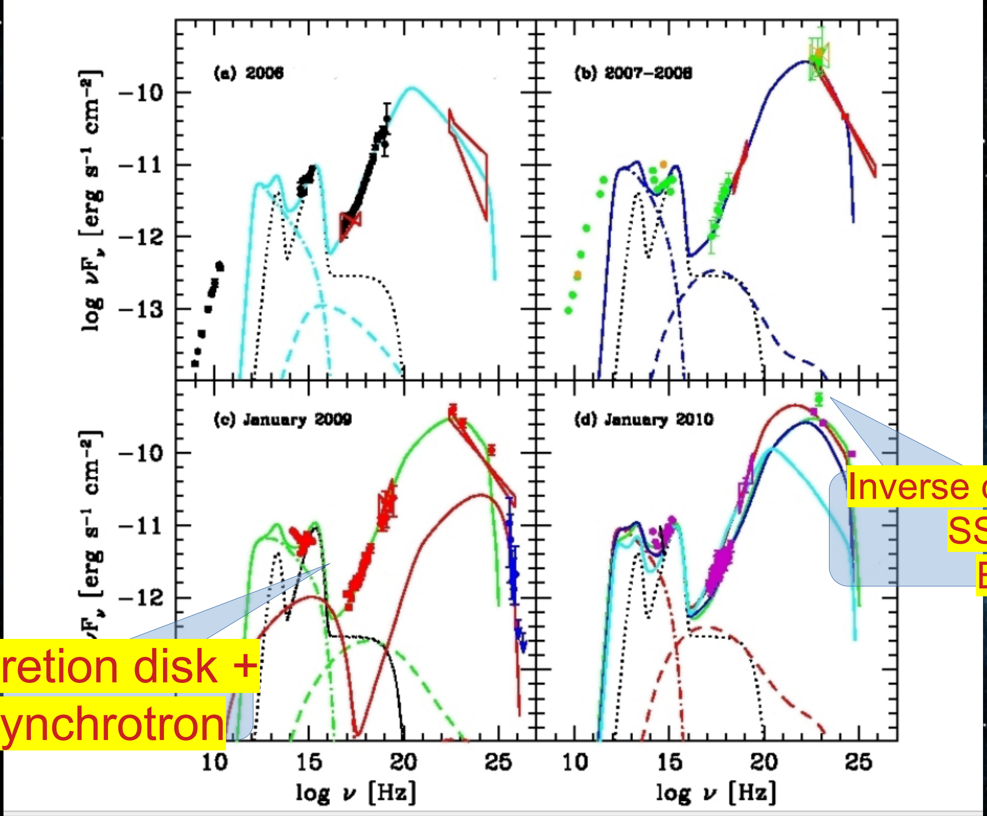


Radiation Mechanisms

- Thermal
  - **Black body radiation** – the effectiv temperature depended on mass transfer rate, mass of the black hole and the Schwarszchild radius. (Mineshige, Yonehara and Kawaguchi 1999)
  - Advective dominated accretion flow – radiation power affected by the advection in the disk. Also depends on the geometry of the disk



Identifying blazars from spectral energy distribution (SED)



Accretion disk + Synchrotron

Inverse compton + SSC + EC

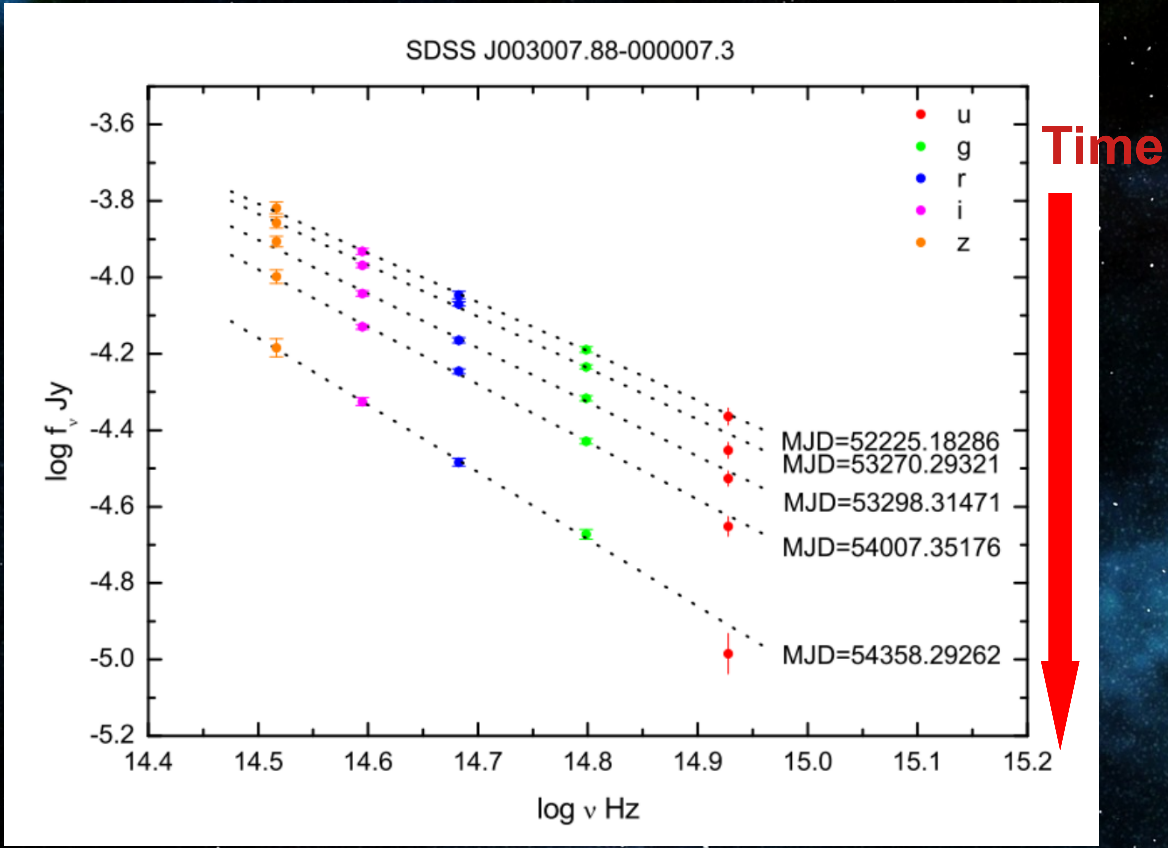
Simultaneous or quasi-simultaneous multi-wavelength spectra of PKS 1510–089 at different epochs in the rest frame

The shape of SED changes overtime, suggesting the disruption within radiative process

Castignani et al. 2017, Figure 7
A&A 601, A30 (2017)

DOI: 10.1051/0004-6361/201629775

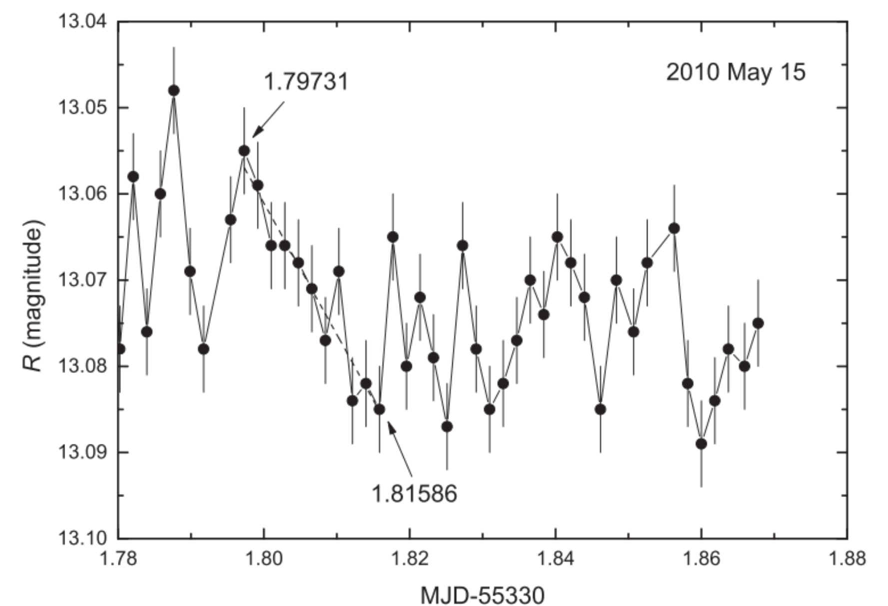
Variation of SED



The change of SED from early (the top most dotted line) to late (the bottom most dotted line) observation. Optical flux decreases across the whole spectrum. The flux of higher frequencies decrease faster than the lower frequency, suggesting that the object become more “red” as the overall flux decreases.

Figure 5: Example of the spectral energy distribution (SED) variability of SDSS

J003007.88−000007.3 with a selection of different continuum spectra. The dotted lines are the best fits to ugriz photometric data. (Mao & Zhang 2016).



Possible intra-day variability of Mrk 501 (Feng et al. 2017)Figure 6 https://doi.org/10.3847/1538-4357/aa9123

Intra-day magnitude fluctuation considered to be the result of shock in the jet. The model is based on the relativistic shock propagating down the jet and interacting with the nonuniform portion in the jet flow (Feng et al. 2017)

The internal shock of plasm in the process of synchrotron and inverse-compton radiation can also be the caused of the shock (Pian et al. 2007)

Questions and Motivation

Questions and motivations

- Blazars can be identified and classified through careful observation in multiple-frequency. Is it possible to do the same task with less information?
- Is it possible to use the data from all-sky surveys to perform data classification ?
- What information can we derived from the surveyed data ?

Aim of the research

- To derived statistical information from survey data
- To investigate the use of machine learning to classify blazars BL Lacs and FSRQs using the information extracted from the survey data

Blazar sample

- Roma-BZCAT 5th edition
  - updated in 2015
  - 3561 sources
  - multi-frequency catalog
  - 5 classes:

FSRQs	1909	• 53.6 %
BL Lac	1059	• 29.7 %
BL Lac-galaxy dominated	274	• 7.7 %
Blazar Uncertain type	227	• 6.4 %
BL Lac Candidate	92	• 2.6 %

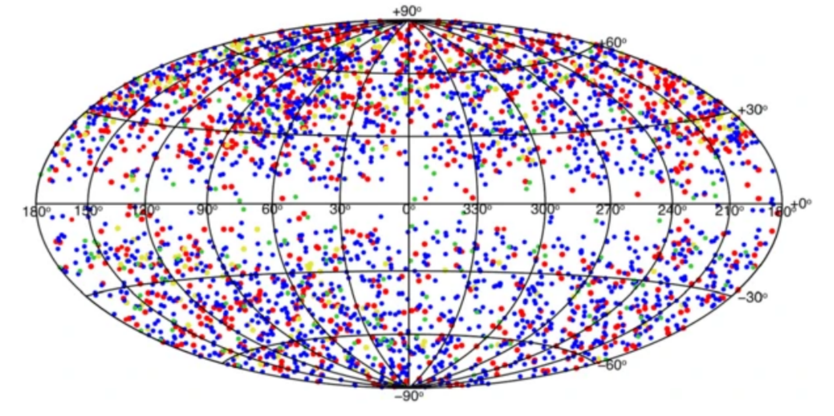


Figure 8 : Aitoff projection of blazars in Roma-BZCAT 5th edition (Massaro et al. 2015) arXiv:1502.07755v1

Zwicky Transient Facility (ZTF)

- transient-focused optical time-domain survey
- g, r, i band
- 3-night cadence
- 2019-2020 survey

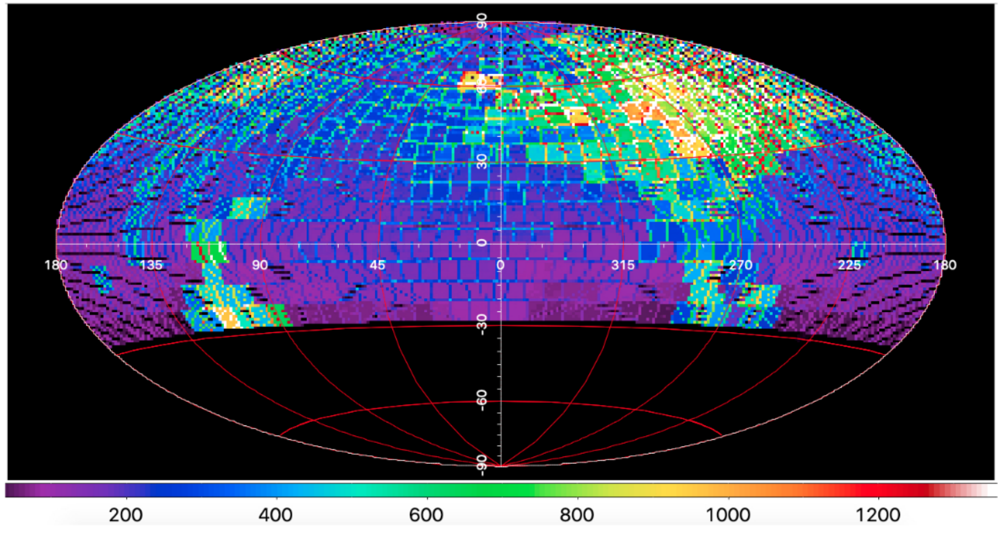


Figure 9: ZTF coverage map

Bellm et al 2019

https://iopscience.iop.org/article/10.1088/1538-3873/aaecbe

- Owens Valley Radio Observatory (OVRO)
  - blazar monitoring proram since 2008-2020
  - 40-m single-dish radio telescope
  - 15 GHz follow-up of Fermi-LAT
  - monitors 1,847 blazars

Richards et al. 2011 arXiv:1011.3111

Gravitation-wave Optical Transient Observer (GOTO)

- Optical telescope geared toward
- transient observation
- Prioritize signals triggered
- by gravitational wave detectors
- All-sky surveys repeating every 6 nights
- Data used in this project is reduced using
- LSST stack data reduction pipeline
- Calibrated to PanSTAR-1 L band
- (Mullaney et al. 2020, arXive:2010.15142v1) ,

Selection criteria

- Light curves must have more than 10 data points
- Cover more than 70 % of blazar samples

ZTF Data availability (N > 10)

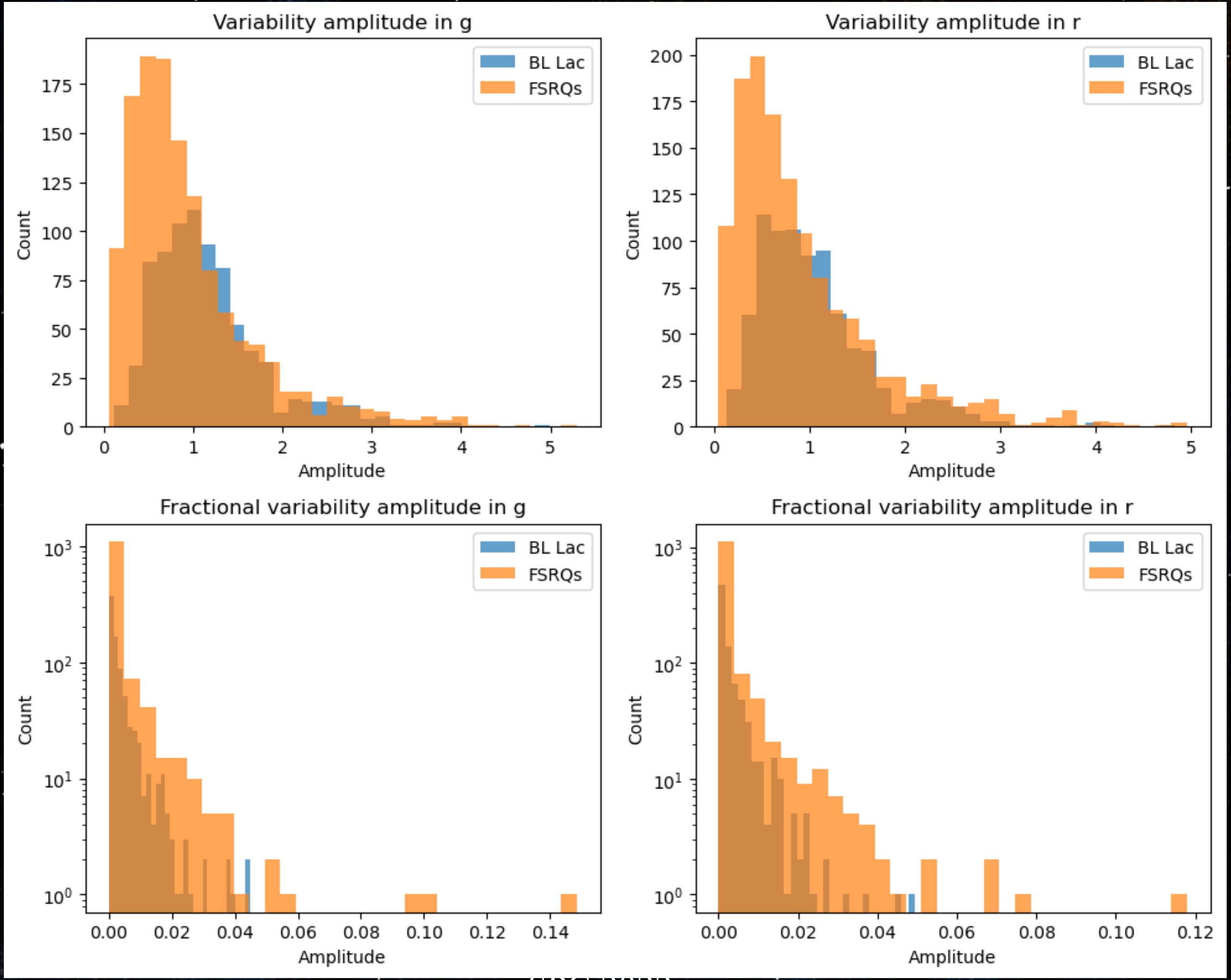
	g_band	r_band	i_band	• L_band	• 15 GHz
count	2499	2616	1087	• 1123	• 730
mean	179.87	237.76	36.92	• 27.28	• 565
std	171.86	206.79	25.13	• 14.63	• 93
min	10	10	10	• 10	• 356
25%	57	73	16	• 16	• 501
50%	126	184	31	• 23	• 557
75%	247	340.25	52	• 35	• 625
max	1436	1573	154	• 108	• 112

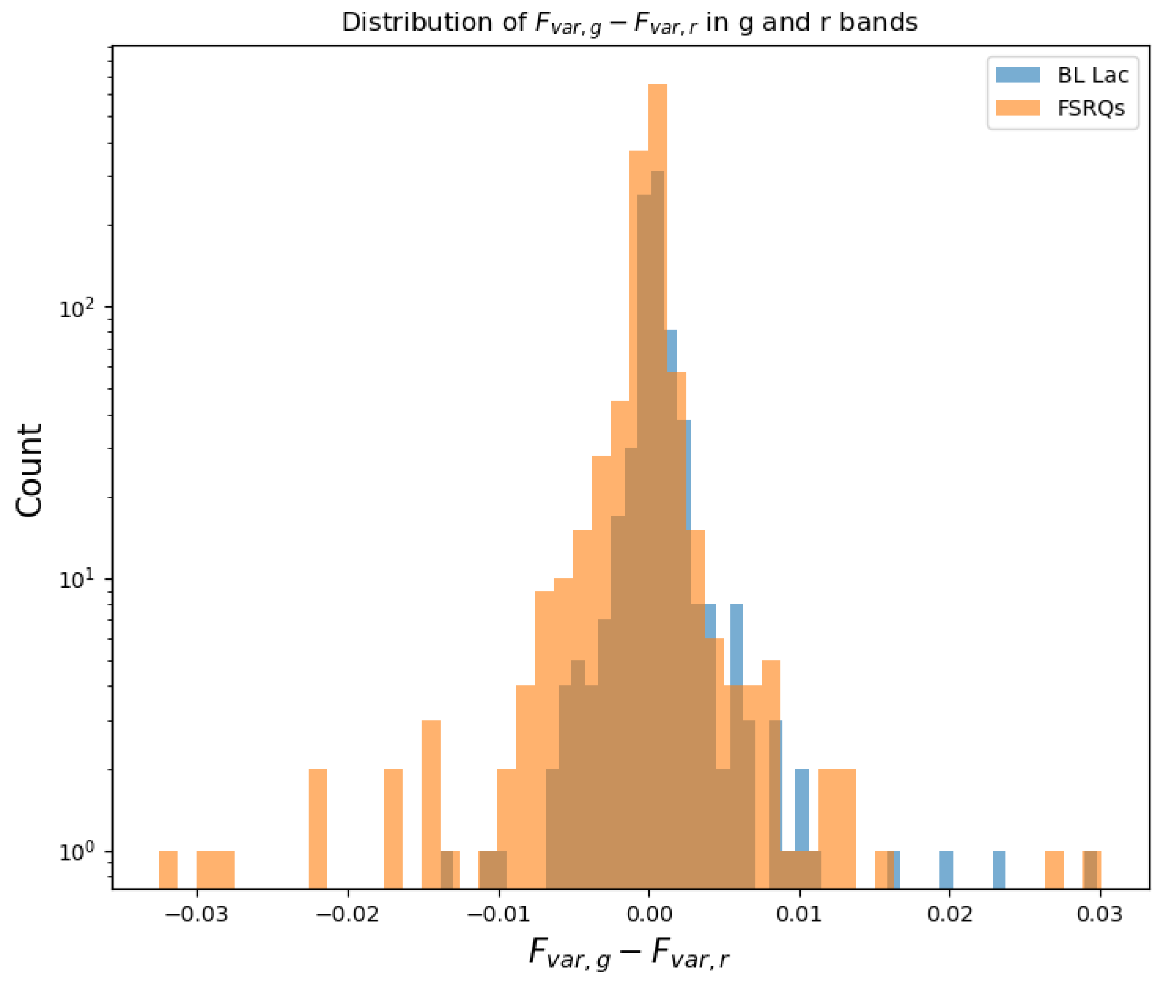
Attributes

- Brightness
  - Brightness Variability (Variability Amplitude)
  - Fractional Variability

$$\Psi_r = \sqrt{(A_{max} - A_{min})^2 - 2\sigma^2}$$

$$F_{var} = \sqrt{\frac{S^2 - \overline{\sigma_{err}^2}}{\overline{x^2}}}$$


Variability amplitude in g
Variability amplitude in r
Fractional variability amplitude in g
Fractional variability amplitude in r
BL Lac
FSRQs
Count
Amplitude
0
25
50
75
100
125
150
175
200
1
2
3
4
5
10^0
10^1
10^2
10^3
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14


Distribution of $F_{var,g} - F_{var,r}$ in g and r bands
BL Lac
FSRQs
Count
10^2
10^1
10^0
−0.03
−0.02
−0.01
0.00
0.01
0.02
0.03
$F_{var,g} - F_{var,r}$

Discrete Correlation Function

$$UDCF_{ij}(\tau) = \frac{(x_i - \bar{x})(y_j - \bar{y})}{\sigma_x^2 \sigma_y^2}$$

$$DCF(\tau) = \frac{1}{m} \sum_{k=1}^{m} UDCF_k$$

Discrete Correlation Function

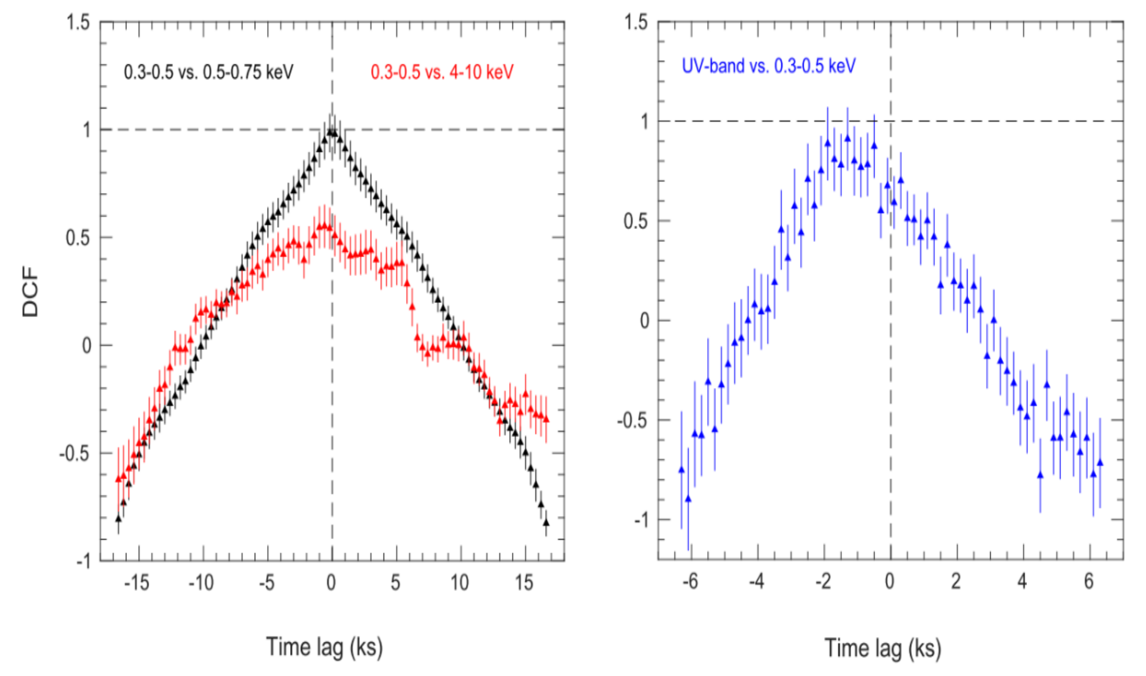
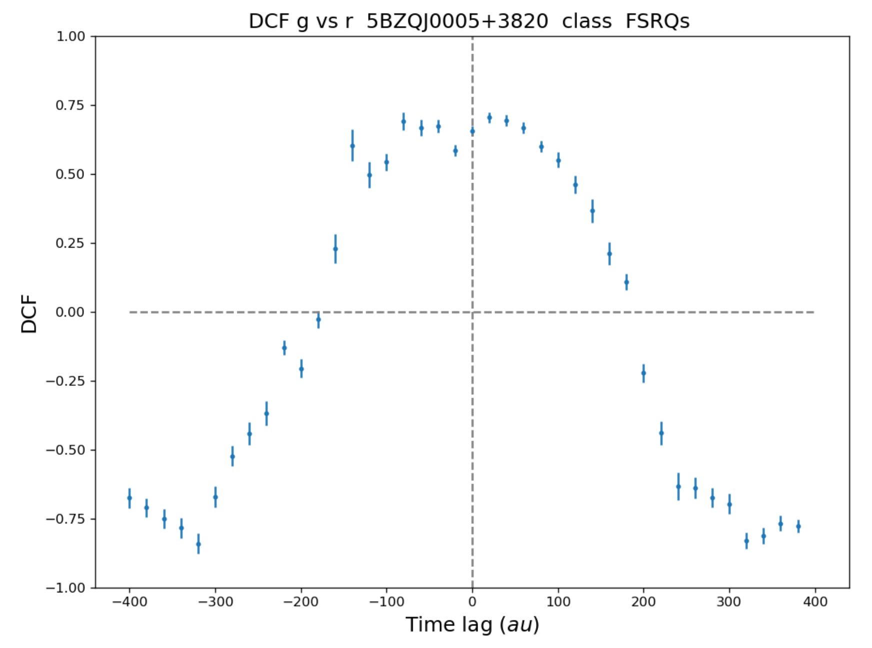
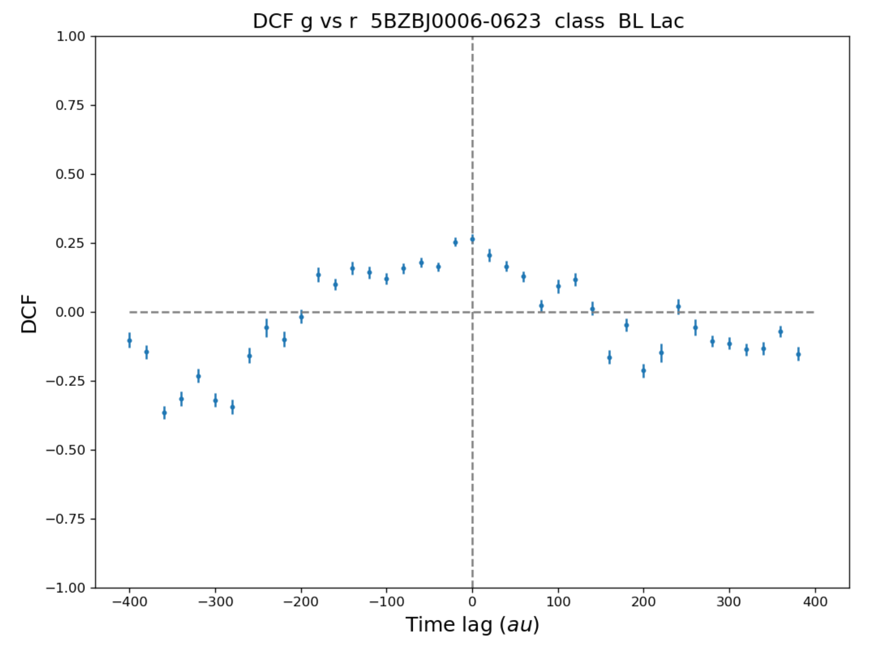


Figure 7: DCF of blazar ON 231. Left: DCF between different X-ray energy bands. A
soft lag of ~ -0.40 ks is detected between the 0.3-0.5 keV and 4-10 keV bands (red
curve). Right: DCF representation between UV and 0.3 - 0.5 keV X-ray bands, where
a soft lag around -1.25 ks is observed (Kalita et al. 2019)

Machine Learning





Machine Learning

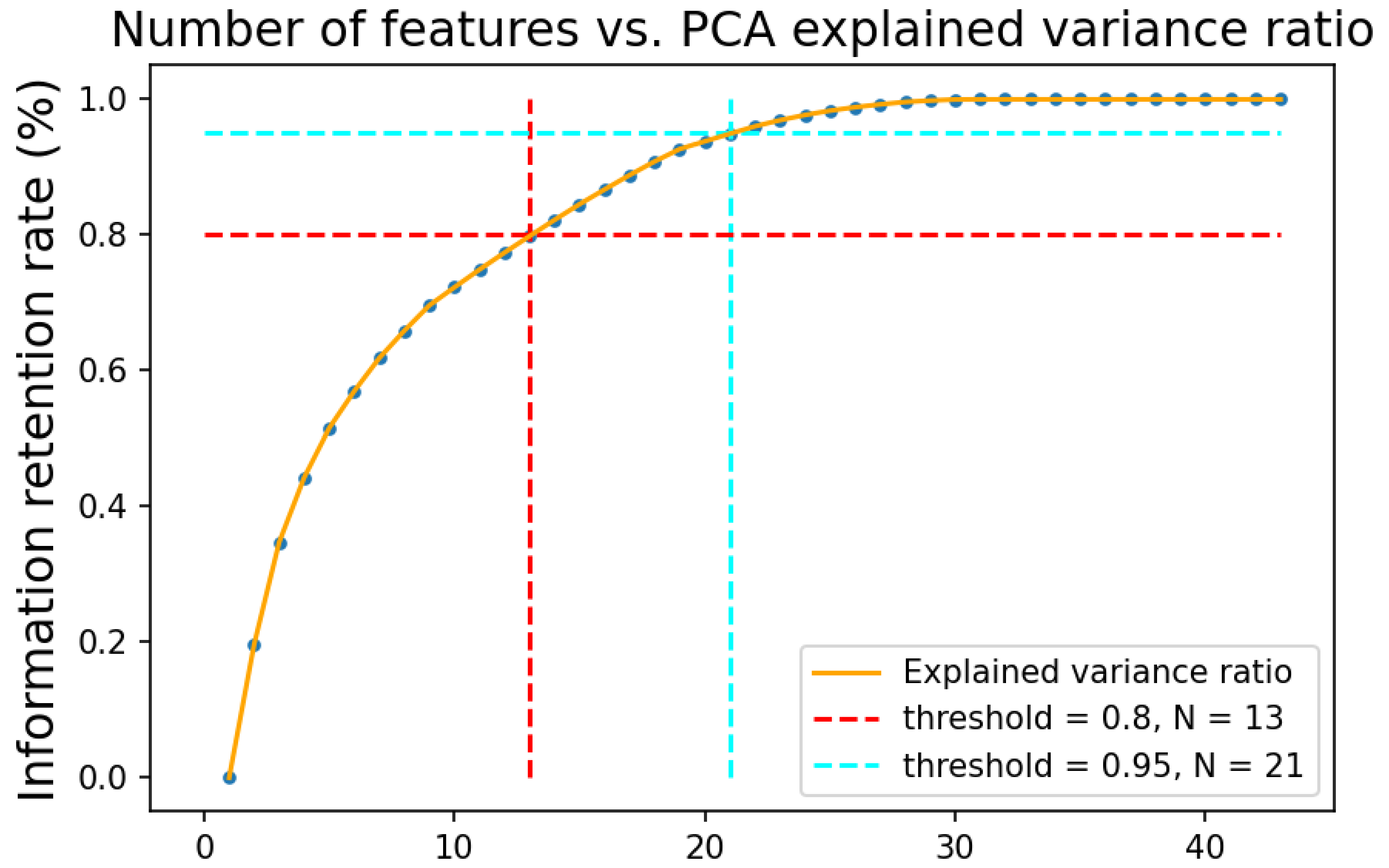
1. Selecting data with more than 70 % availability
2. Padding missing data with the mean of each features
3. Scale the value of each features to the maximum and minimum value
4. Do principle component analysis (PCA)
5. Training and testing machine learning algorithm

Feature extraction

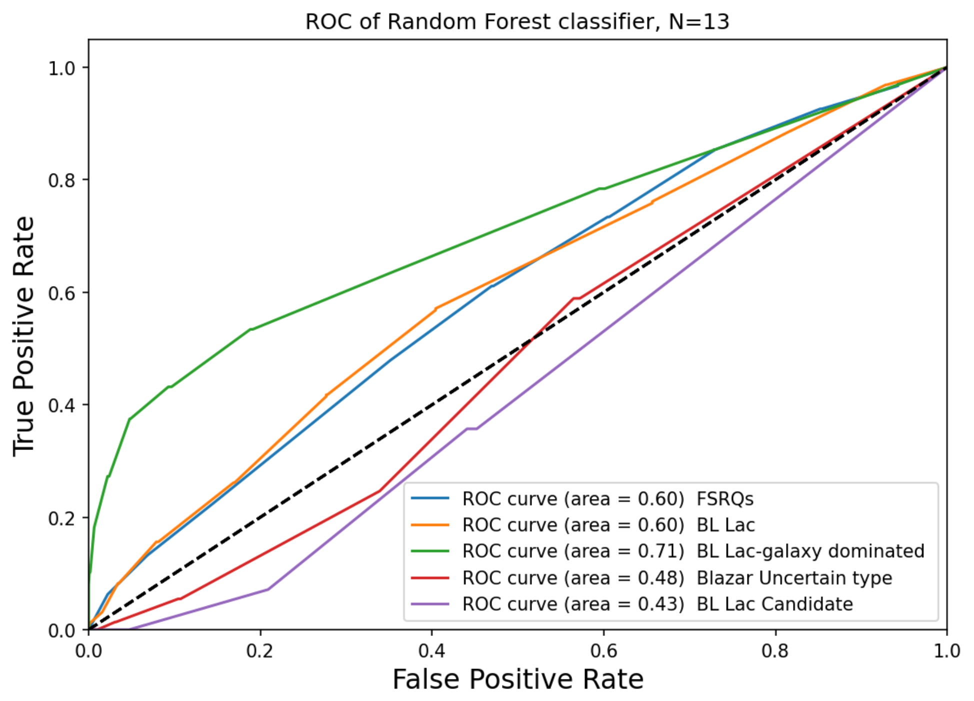
- Features can be extract from any form of time series analysis, such as, variability amplitudes, DCF values and time lags, color, etc.
- It can also be the parameters acquired by fitting mathematical models to data
  - Linear regression, sinusoidal fitting for light curve
  - Gaussain fitting, Z-score value for DCF
  - etc.

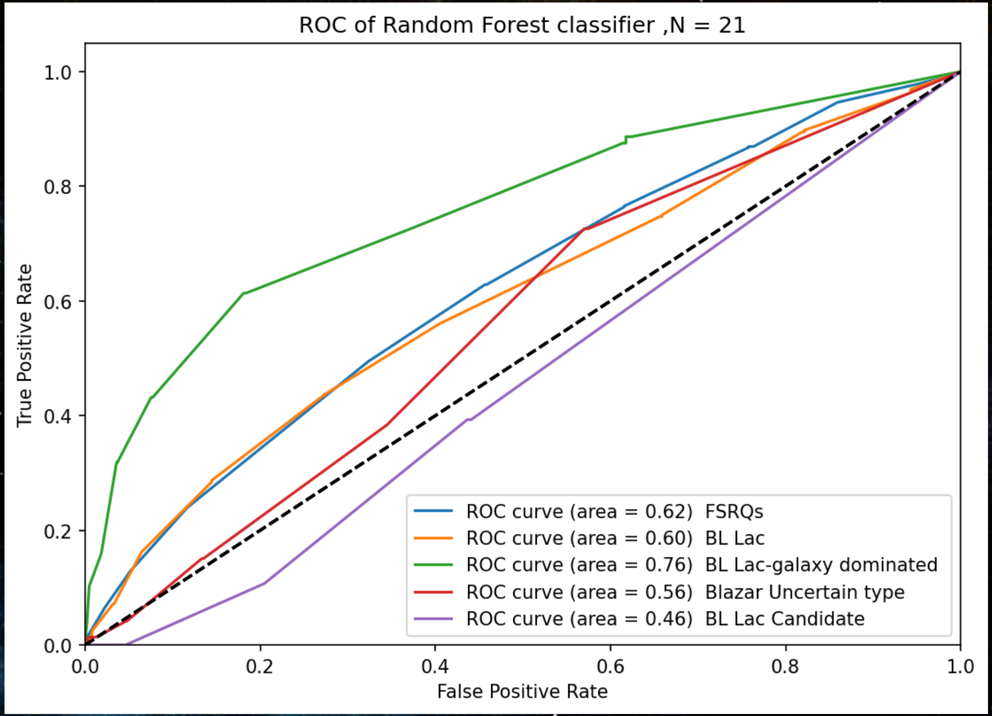
PCA

- Reducing dimension of large data sets into smaller one by containing most of the information
- Testing the reduction to N features, find the N where PCA explained variance ratio where the information is retained more than 80 %
- at this threshold the minimum number is 14


Number of features vs. PCA explained variance ratio
Information retention rate (%)
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Explained variance ratio
threshold = 0.8, N = 13
threshold = 0.95, N = 21

Testing with Random Forest Classifier





The number of features slightly affect the classification result.
The area under the curve (AUC): the closer to 1 the better

Summary

- For the primary analysis, we use the data with 70% completeness to avoid false information through data padding
- The features used in the test produce mediocre results
- The classification of galaxy-dominated BL Lacs out-perform the classification of FSRQs and BL Lacs

What next

- Experiment on adding data with less than 70 % availability (expand frequency ranges but smaller sample)
- Experiment on feature extraction (DW model, MA model, Encoder-Decoder method)
- Add data from Fermi-LAT

Acknowledgements

- The data from GOTO reduced using LSST data processing pipeline is accessible with the help of Dr. John Mullaney and Lydia Makrygianni from The University of Sheffield, under the NARIT-GOTO collaboration

References

Castignani, G. et al., Multiwavelength variability study and search for periodicity of PKS 1510–089, Atron. Astrophys., 601, A30, 2017

Mao, L. & Zhang, X., Long-term optical variability properties of blazars in the SDSS Stripe 82 Astrophys. Space Sci., 361, 345, 2016

Feng, H. et al., Search for Intra-day Optical Variability in Mrk 501, Astrophys. J., 849:161 2017

Pian, E. et al. Detection of Gamma Rays with E > 300 GeV from Markarian 501, Astrophys. J., 664, 106, 2007

Massaro, E. et al., The 5th edition of the Roma-BZCAT. A Short Presentation, arXiv:1502.0775v1

Bellm, E. C. et al., The Zwicky Transient Facility: System Overview, Performance, and First Results. Publ. Astron. Soc. Pac., 131,018002, 2019

Richards, J. L. et al., Blazars in the Fermi Era: The OVRO 40-m Telescope Monitoring program. Astrophys. J. S., 194, 99, 2011

Q / A