Automated Classification of Radio Galaxies Using Ensemble Learning and SVD-based Feature Extraction

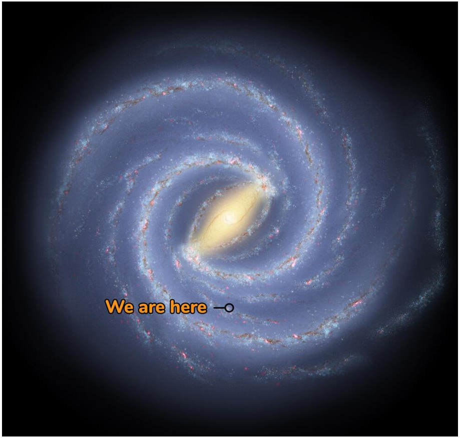
T. Ansah-Narh

December 23, 2024

Outline

Galaxies

Galaxies I



The Milky Way. *Image credit*: Ology website

- A **galaxy** is a massive collection of stars, gas, and dust bound together by gravity.
- Galaxies come in various shapes and sizes, with the Milky Way, our home galaxy, being a barred spiral galaxy.
- Spiral galaxies have a flat, rotating disk with spiral arms extending outward from a central bulge. These arms are filled with gas and dust, which is where new stars are born.

Galaxies II

(a) Elliptical; Credit: Wiki website

(b) Irregular, Credit: ESA/Hubble & NASA

- **Regular optical images**, like the ones used to identify galaxy shapes, reveal the distribution of stars, dust, and gas based on the visible light they emit or reflect.

Galaxies III

- However, **radio observations** reveal a different aspect of galaxies. Here's how they differ:
  - **Wavelength:-** Visible light has a much shorter wavelength compared to radio waves. This means it interacts differently with matter in galaxies. Radio waves can penetrate dust clouds much more effectively than visible light, allowing us to see deeper into the galaxy and potentially observe cool gas and cold dust that wouldn't be visible otherwise.
  - **Emission Mechanisms:-** The light we see from stars in galaxies is due to their internal thermonuclear reactions.

Galaxies IV

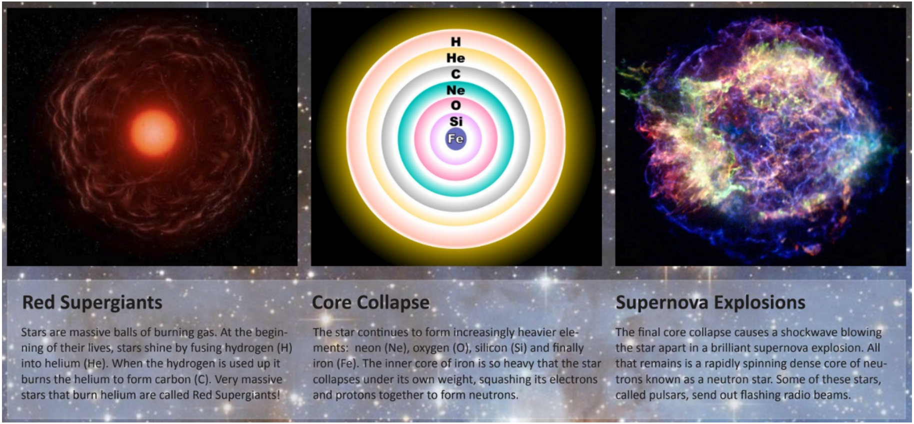
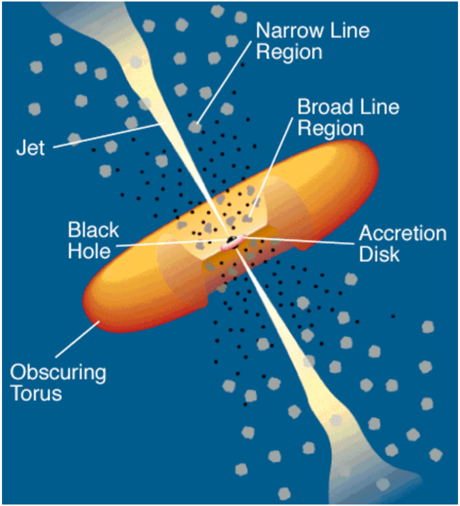


Figure: Birth and Death of a Massive Star. Created by Marisa & Elmarie

Galaxies V



An active galactic nuclei model. *Image credit*: C.M. Urry & P. Padovani.

- In contrast, radio emission from galaxies can come from various processes, including:
  - *Synchrotron radiation*: This occurs when high-energy electrons spiral around magnetic fields within the galaxy, emitting radio waves. This is often seen in the **jets** and **lobes** of active galaxies.
  - *Thermal radiation*: Collisions between gas particles can generate radio waves, revealing the presence of cool gas within the galaxy.

Galaxies VI

- **Structures Revealed**: Radio observations can highlight features not readily apparent in visible light. For example:
  - *Active Galactic Nuclei (AGN)*: These supermassive black holes at the center of some galaxies can launch powerful jets that emit strongly in radio wavelengths. These jets may not be readily visible in optical images.
  - *Spiral Arm Structure:* The distribution of gas and dust within the spiral arms of galaxies can be traced by radio observations, as these components emit radio waves.
  - *Hydrogen Gas*: Radio observations can detect the signature of neutral hydrogen gas (HI regions) within galaxies, which is an important component for star formation but may not be prominent in optical images.

Why Classify Radio Galaxies?

Classification of radio galaxies I

Morphological classification of radio galaxies is crucial for several reasons:

- **Understanding Galaxy Evolution:** Radio galaxies exhibit a wide range of morphologies, which can provide insights into the evolutionary processes of galaxies. Classifying them helps astronomers understand how galaxies form, evolve, and interact with their environments over cosmic time.
- **Probing Active Galactic Nuclei (AGN) Physics**: Radio galaxies often host AGN, which are powered by accretion onto supermassive black holes at their centers. Different morphological types can indicate different stages or modes of AGN activity, shedding light on the physics of black hole accretion and feedback mechanisms.

Classification of radio galaxies II

- **Environmental Influence:** The morphology of a radio galaxy can be influenced by its surroundings, such as interactions with neighboring galaxies, the intracluster medium, or cosmic filaments. Classifying morphologies helps discern these environmental impacts and their effects on galaxy evolution.
- **Cosmic Structures and Large-Scale Distribution:** Radio galaxies are often found in clusters and large-scale structures of the universe. Studying their morphologies helps map out these structures and understand the cosmic web's formation and evolution.

Classification of radio galaxies III

- **Technological Innovation:** Research into radio galaxies and other astronomical objects often drives technological advancements. For instance, innovations in radio astronomy, imaging techniques, data processing, and computational methods developed for studying radio galaxies can have spin-off applications in fields such as telecommunications, imaging technology, and signal processing. These innovations can lead to improved technologies that benefit various industries and society as a whole.

Problem Definition and Research Objective

Problem Definition and Research Objective I

- Traditional classification methods in radio astronomy have long relied on **manual** inspection and the **expertise** of human astronomers.
- The era of **Big Data**, spurred by advancements in next-generation radio telescopes, is set to generate unprecedented volumes of data.

(a) SKA; Credit: SKAO website

(b) LOFAR, Credit: LOFAR website

Problem Definition and Research Objective II

THE ASTROPHYSICAL JOURNAL
SUPPLEMENT SERIES

Classifying Radio Galaxies with the Convolutional Neural Network

A. K. Aniyan[1,2] and K. Thorat[1,2]

Published 2017 June 13 • © 2017. The American Astronomical Society. All rights reserved.

The Astrophysical Journal Supplement Series, Volume 230, Number 2

Citation A. K. Aniyan and K. Thorat 2017 *ApJS* **230** 20

DOI 10.3847/1538-4365/aa7333

Article PDF

Article ePub

Problem Definition and Research Objective III

JOURNAL ARTICLE

Radio Galaxy Zoo: compact and extended radio source classification with deep learning FREE

V Lukic, M Brüggen, J K Banfield, O I Wong, L Rudnick, R P Norris, B Simmons

Monthly Notices of the Royal Astronomical Society, Volume 476, Issue 1, May 2018, Pages 246–260, https://doi.org/10.1093/mnras/sty163

Published: 26 January 2018 **Article history** ▾

PDF Split View Cite Permissions Share ▾

Problem Definition and Research Objective IV

JOURNAL ARTICLE

Fanaroff–Riley classification of radio galaxies using group-equivariant convolutional neural networks

FREE

Anna M M Scaife, Fiona Porter

Monthly Notices of the Royal Astronomical Society, Volume 503, Issue 2, May 2021, Pages 2369–2379, https://doi.org/10.1093/mnras/stab530

Published: 26 February 2021 **Article history** ▾

PDF Split View Cite Permissions Share ▾

Problem Definition and Research Objective V

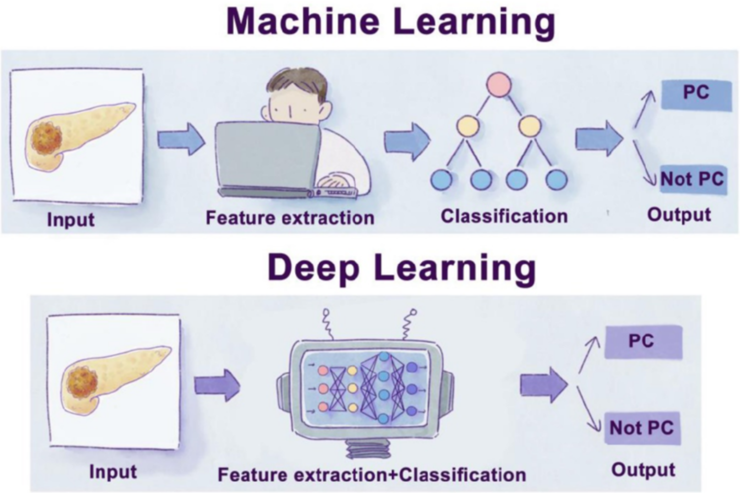


Figure: Credit: B. Huang et al. 2022

Problem Definition and Research Objective VI

- Even though, recent advancements in machine learning, such as transfer learning, and deep learning, have shown promise in automating the classification process, these methods often face limitations in handling the intricacies of vast datasets and evolving classification standard.
- Deep learning models require large labelled datasets for training, which are often not available in sufficient quantity and quality in radio astronomy.
- Additionally, these models can be computationally intensive, posing challenges for real-time processing of large datasets generated by modern radio telescopes.
- Therefore, there remains a pressing need for robust, adaptable approaches capable of effectively processing and classifying radio galaxies.

Radio Galaxy Dataset

Radio Galaxy Dataset I

1995ApJ...450.

The Astrophysical Journal, 450:559–577, 1995 September 10
© 1995. The American Astronomical Society. All rights reserved. Printed in U.S.A.

THE FIRST SURVEY: FAINT IMAGES OF THE RADIO SKY AT TWENTY CENTIMETERS

Robert H. Becker
Department of Physics, University of California, Davis; and Institute of Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, Livermore, CA 94450
Richard L. White
Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218
and
David J. Helfand
Department of Astronomy and Columbia Astrophysics Laboratory, 538 West 120th Street, New York, NY 10027
Received 1994 December 14; accepted 1995 March 20

ABSTRACT

The FIRST survey to produce Faint Images of the Radio Sky at Twenty centimeters is now underway using the NRAO Very Large Array. We describe here the scientific motivation for a large-area sky survey at radio frequencies which has a sensitivity and angular resolution comparable to the Palomar Observatory Sky Survey, and we recount the history that led to the current survey project. The technical design of the survey is covered in detail, including a description and justification of the grid pattern chosen, the rationale behind the integration time and angular resolution selected, and a summary of the other considerations which informed our planning for the project. A comprehensive description of the automated data analysis pipeline we have developed is presented.

We also report here the results of the first year of FIRST observations. A total of 144 hr of time in 1993 April and May was used for a variety of tests, as well as to cover an initial strip of the survey extending between 07^h15^m and 16^h30^m in a 2°8 wide declination zone passing through the local zenith ($28.2 < \delta < 31.0$). A total of 2153 individual pointings yielded an image database containing 1039 merged images $46'.5 \times 34'.5$ in extent with 1''.8 pixels and a typical rms of 0.13 mJy. A catalog derived from this 300 deg^2 region contains 28,000 radio sources. We have performed extensive tests on the images and source list in order to establish the photometric and astrometric accuracy of these data products. We find systematic astrometric errors of $<0''.05$; individual sources down to the 1 mJy survey flux density threshold have 90% confidence error circles with radii of $<1''$. CLEAN bias introduces a systematic underestimate of point-source flux densities of ~ 0.25 mJy; the bias is more severe for extended sources. Nonetheless, a comparison with a published deep survey field demonstrates that we successfully detect 39/49 sources with integrated flux densities greater than 0.75 mJy, including 19 of 20 sources above 2.0 mJy; the sources not detected are known to be very extended and so have surface brightnesses well below our threshold.

- The **FIRST Survey**: Faint Images of the Radio Sky at 20 cm is a significant astronomical survey conducted using the Very Large Array (VLA) radio telescope in New Mexico, USA.
- It aimed to create a detailed radio map of the northern sky at a frequency of 1.4 GHz (20 cm wavelength).
- The survey covered about 10 575 square degrees of the sky, identifying and cataloging over 800 000 sources of radio emission.

Radio Galaxy Dataset II

Class	Morphology	Radio Lobe Properties	Jets
Compact	Unresolved or slightly extended core	Not applicable	Not prominent or absent
FR-I	Fan-like with weak double lobes	Low power, faint, and short extensions	Weak and slow-moving
FR-II	Double-lobed with bright and well-defined lobes	High power, prominent, and extended lobes	Powerful and fast-moving
Bent	Distorted or asymmetric double-lobed structure	Asymmetry may reflect interaction with environment or intrinsic properties	Can be weak or strong depending on the origin of the bend

Radio Galaxy Dataset III

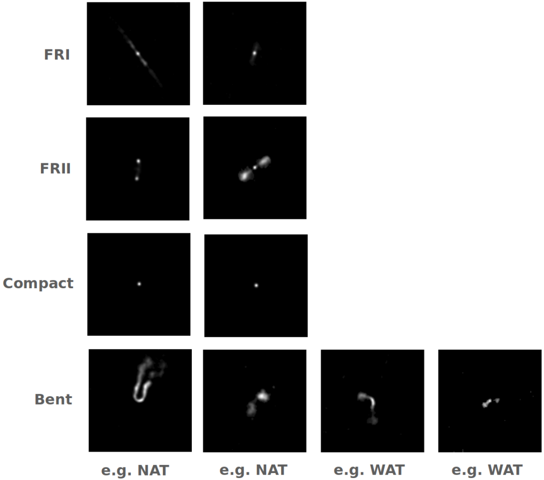


Figure: Class definitions of FR-I, FR-II, Compact and Bent. Florian Griese GitHub

Radio Galaxy Dataset IV

Table: The dataset has the following total number of samples per class.

Classes/Split	FRI	FRII	Compact	Bent	Total
Total	495	924	391	348	2158

Methodology

Singular Value Decomposition (SVD) I

- SVD is a fundamental matrix factorization technique in linear algebra.
- It decomposes a given matrix X into three other matrices, revealing many useful properties and structures of the original matrix.
- Given a matrix $X \in \mathbb{R}^{m \times n}$, the SVD is defined as:

$$X = U \Sigma V^T,$$

where:
  - $U \in \mathbb{R}^{m \times m}$ is an orthogonal matrix whose columns are the left singular vectors of X.
  - $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix with non-negative real numbers on the diagonal known as the singular values.
  - $V \in \mathbb{R}^{n \times n}$ is an orthogonal matrix whose columns are the right singular vectors of X.

Singular Value Decomposition (SVD) II

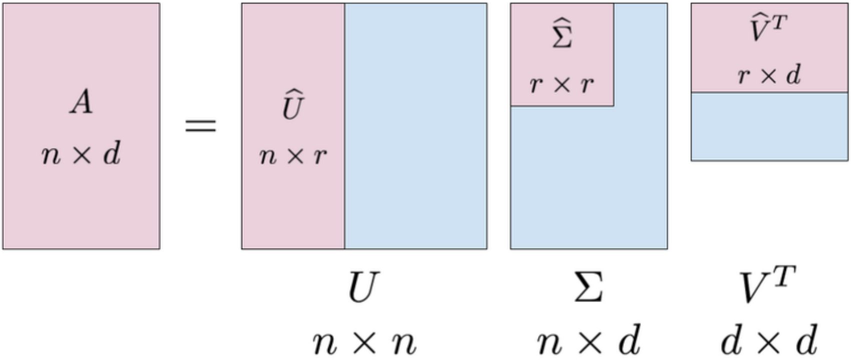
- The matrices U and V form two sets of orthonormal bases for the row space and column space of X, respectively. The singular values on the diagonal of Σ are sorted in descending order.
- The SVD can also be expressed as a sum of outer products, highlighting the rank-r approximation of X:

$$X = \sum_{i=1}^{r} \sigma_i \mathbf{u_i} \mathbf{v_i}^T,$$

where:

- $r = \text{rank}(X)$ is the number of non-zero singular values.
- σ_i are the singular values.
- $\mathbf{u_i}$ and $\mathbf{v_i}$ are the i-th columns of U and V, respectively.

Singular Value Decomposition (SVD) III



Singular Value Decomposition (SVD) IV

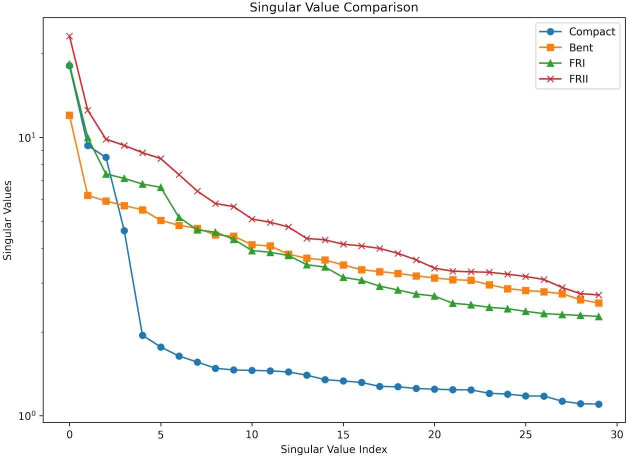


Figure: Spectral profile of truncated SVD.

Singular Value Decomposition (SVD) V

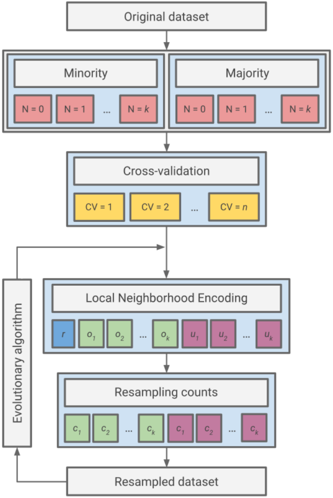


Image credit: Koziarski, M., & Woźniak, M. (2024)

- The original dataset is divided into bags based on the number of same class nearest neighbors, and afterwards into several cross-validation folds.
- Then, evolutionary algorithm is used to optimize Local Neighborhood Encodings (LNEs) coding the number of observations with specific number of same class nearest neighbors that will be over- and undersampled. Finally LNE resamples the original dataset.

Singular Value Decomposition (SVD) VI

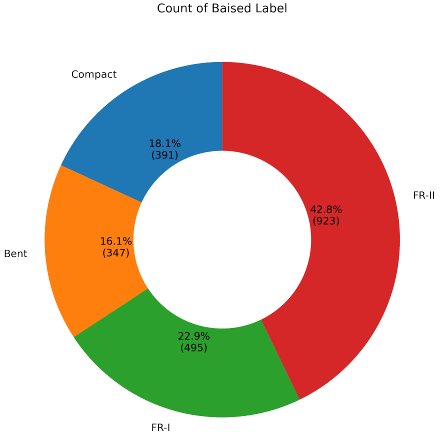


Figure: Balancing the data with Local Neighborhood Encodings method.

Ensemble Learning I

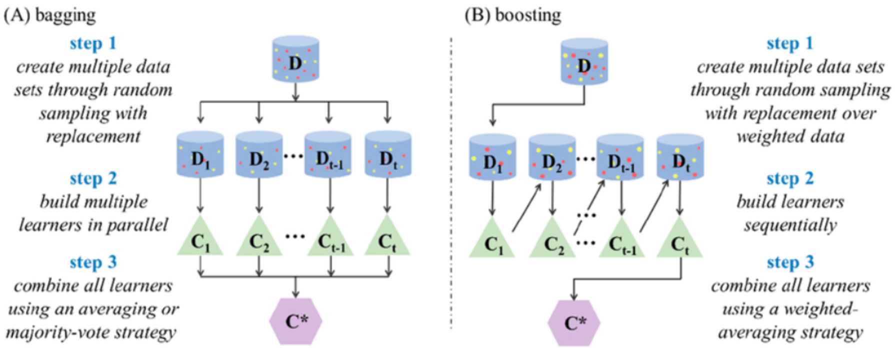


Figure: Illustrations of (A) bagging and (B) boosting ensemble algorithms. (Yang et al. 2019)

Ensemble Learning II

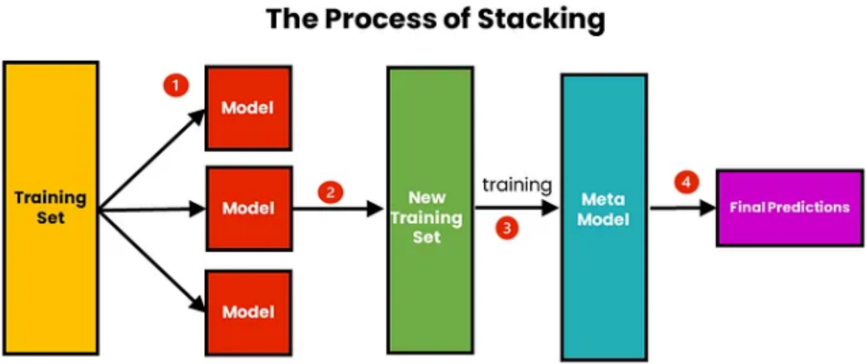
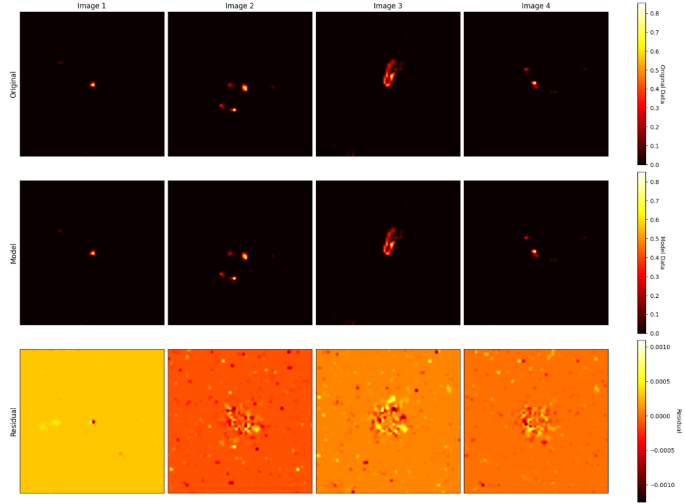


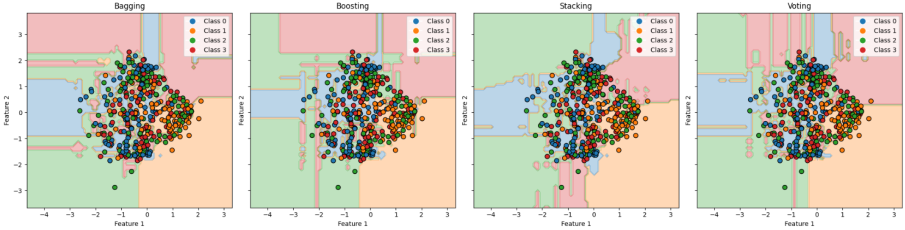
Image by Google

Results & Conclusion

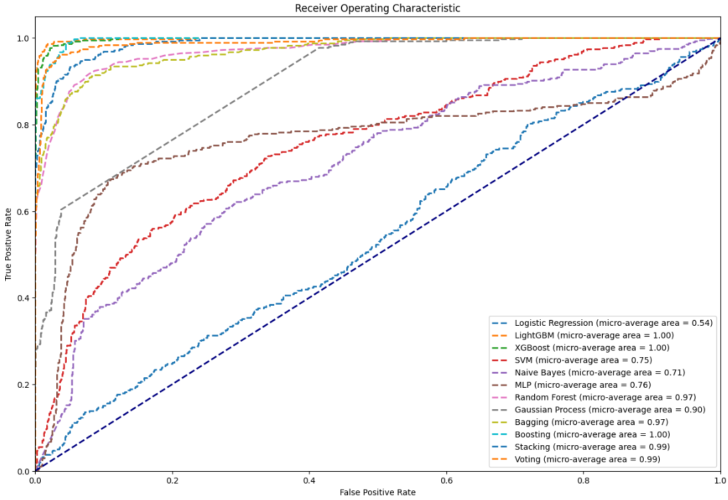
Results & Conclusion I



Results & Conclusion II



Results & Conclusion III



Results & Conclusion IV

- Out of an image size of 256×256, SVD was able to reconstruct the radio images using 120 top singular values, making it a good model feature extraction.
- The accuracy values for the ensemble models outperformed previous results.
- Further studies is to enrich the SVD features with Bayesian finite mixture models, GMM, GaMM, and BMM.
- a hybrid model of Variational AutoEncoder and GMM can considered as an alternative model for feature extraction.
- This is still work in progress . . . .

That's All Folks