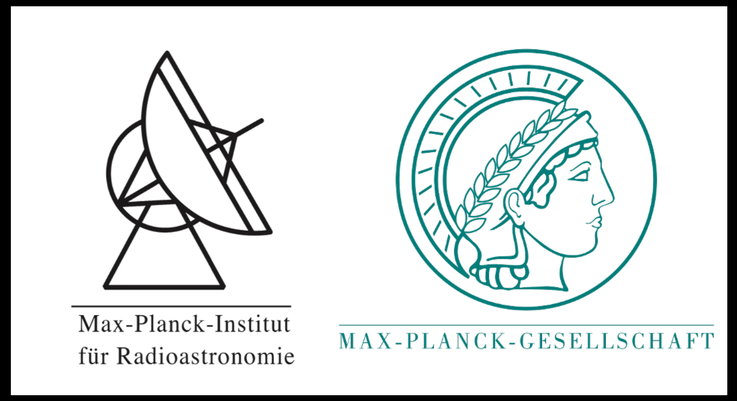
# Discovering New Pulsars with Machine Learning





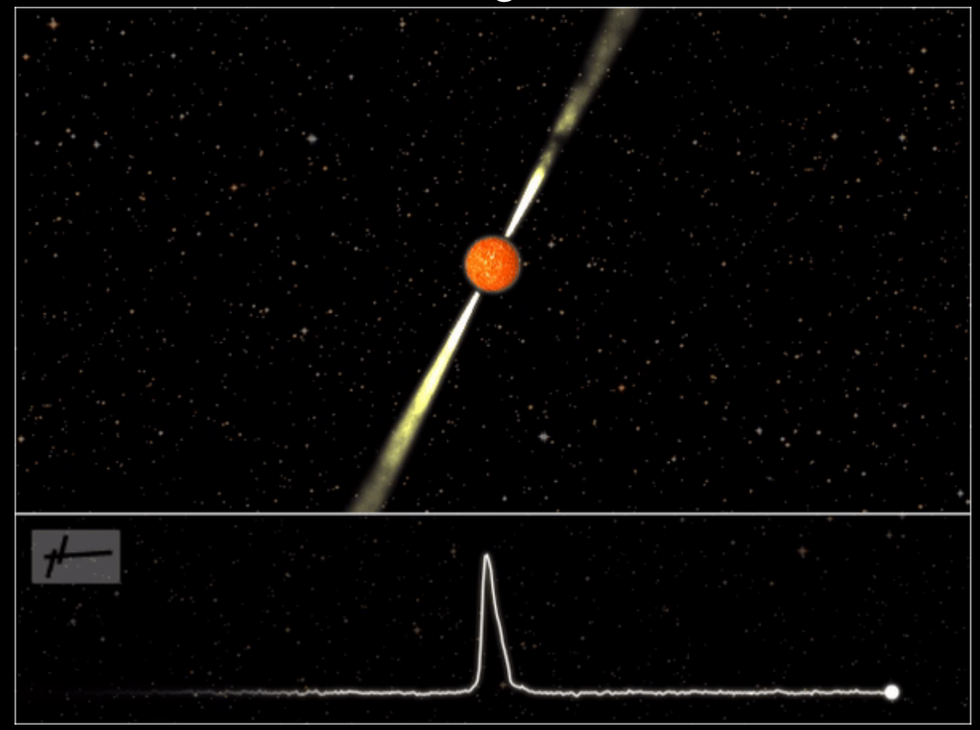
CS & Physics Meet-up, TU Dortmund, 30-11-2023

**Dr. Vishnu Balakrishnan, Postdoc, Max Planck Institute for Radio Astronomy**

# What is a Pulsar?

- **A pulsar is a highly magnetised rotating neutron star that emits twin beams of electromagnetic radiation.**
- **Spin Period ranges from 1.4 ms to 23.5 seconds.**
- **Mass Range** $\approx 1.2 - 2.0\,\mathrm{M}_{\odot}$ **, Radius** $\approx 10 - 15\,\mathrm{km}$
- **Inferred Magnetic Fields** $\approx 10^{8} - 10^{14}\,\mathrm{G}$


**Cosmic lighthouse**

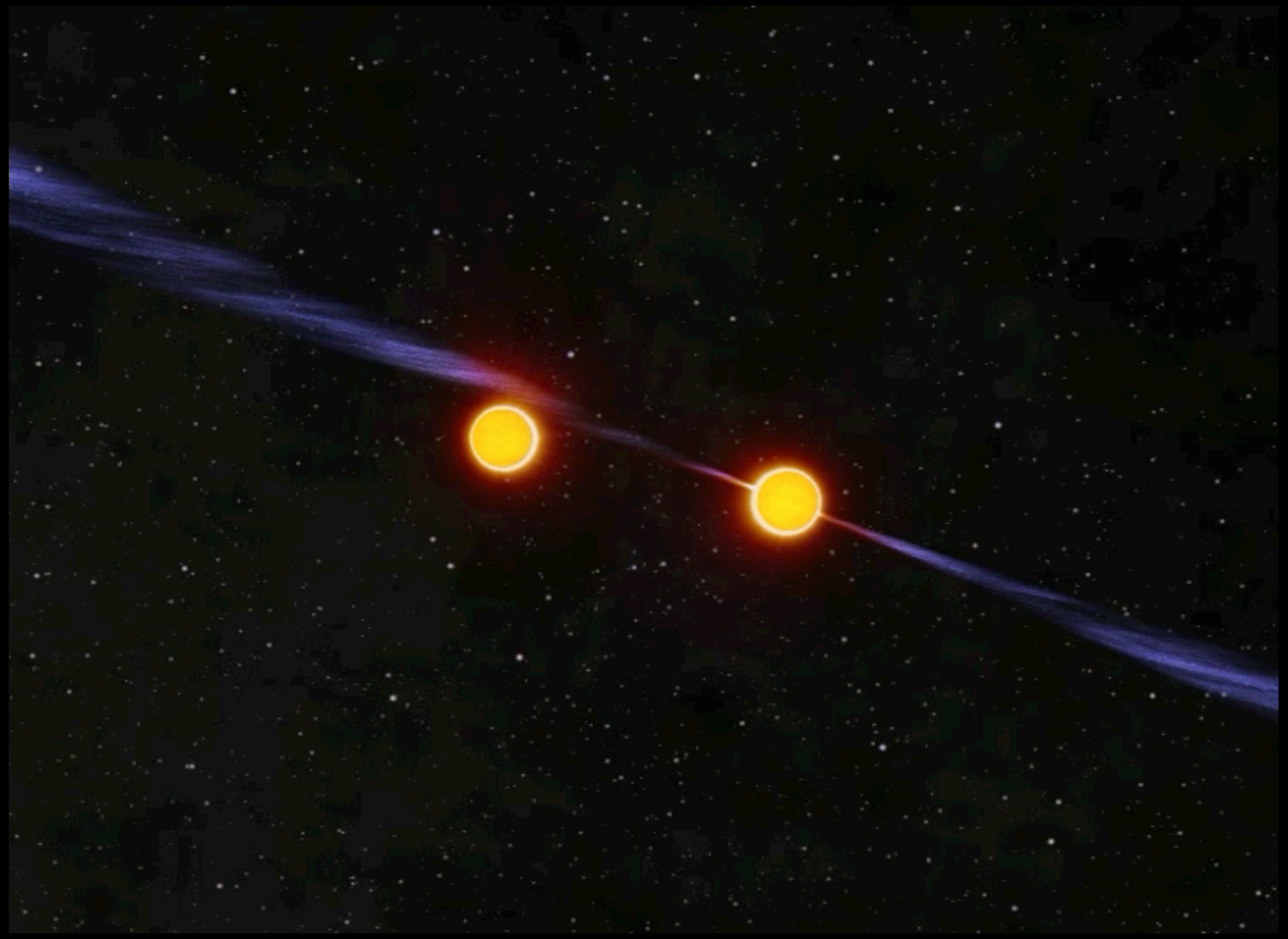


**Credit: Dr. Joeri Van Leeuwen**

Millisecond Pulsars are extremely stable clocks.

PSR J0437-4715: Spin Freq: 173.6879458121843(5) Hz

# Applications: Testing Einstein's theory of General Relativity

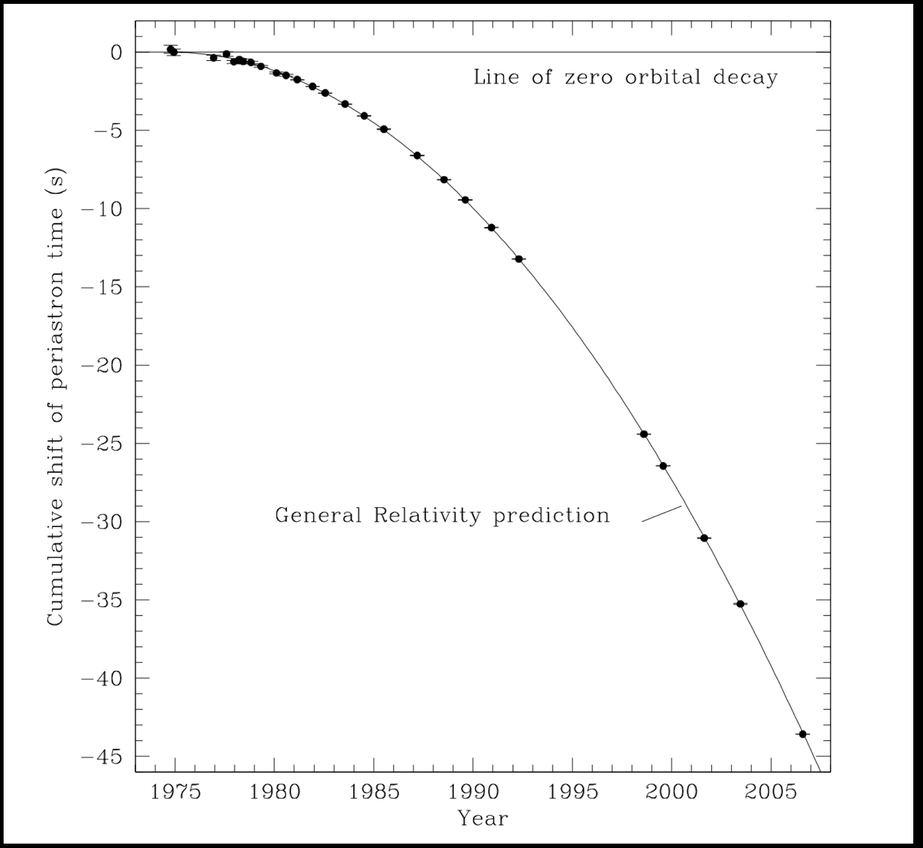


Orbits shrink over time due to emission of Gravitational Waves

Credits: John Rowe

## First Binary Pulsar Discovery

**Hulse-Taylor Binary**




**Hulse and Taylor were awarded the Nobel prize in Physics in 1993.**

# Current Data Challenges

- Pulsar Group (led by Prof. Kramer) in MPIfR, Bonn is conducting the most sensitive pulsar survey in the Southern Hemisphere using MeerKAT radio telescope.

- Quasi-real time survey. We record, analyse and delete 350 TB/week.

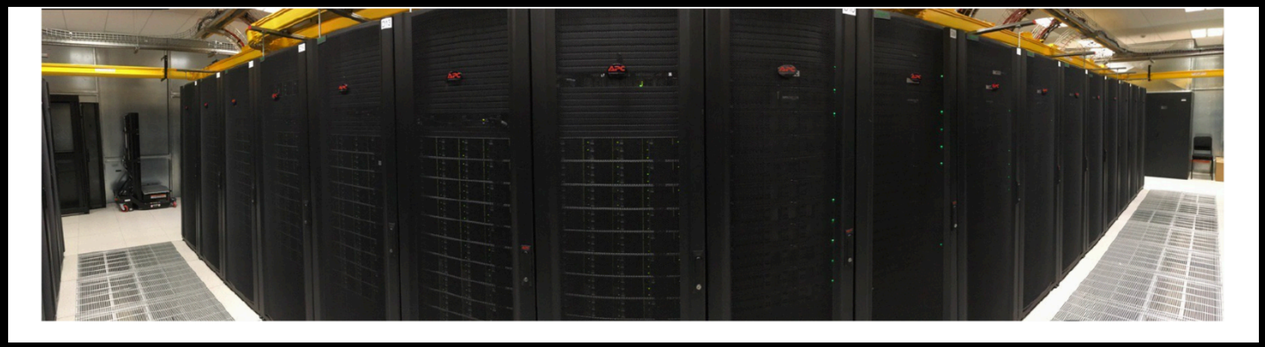
Prof. Michael Kramer

## ERC funded Project: COMPACT

- ERC starting grant awarded to Dr. Vivek Krishnan
- Record 1.6 PB raw data. Data transferred to Germany by airfreight.
- Processing Deadline: 2-months
- Reduced data product (time series, ~GB) transferred to over 7 HPC clusters located in Australia, Germany, Italy, UK, USA, SA.

PI of COMPACT:
Dr. Vivek Krishnan

MeerKAT Radio Telescope located in South Africa (Image courtesy: SAROO)

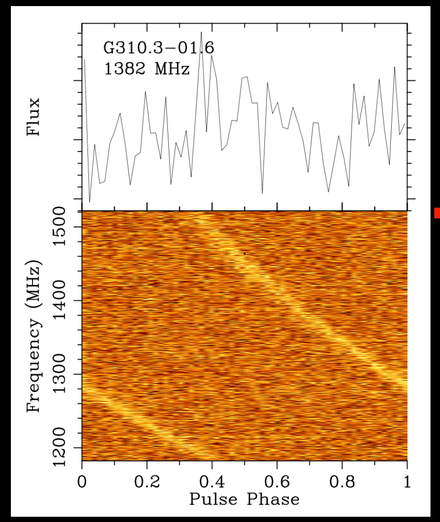


APSUSE High Performance Computing Cluster (Credits: E.Barr)

# Traditional Pulsar Search - Parameter estimation Problem

Credit: John Smith, CSIRO

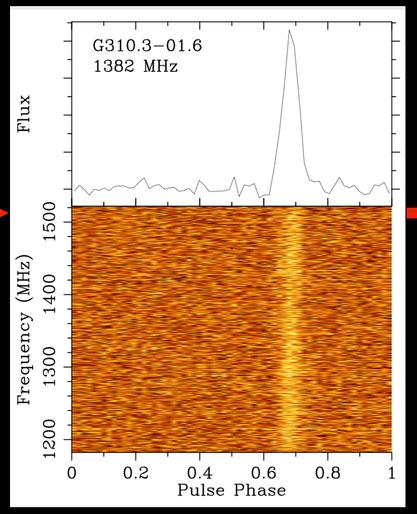
**Correct for Distance**



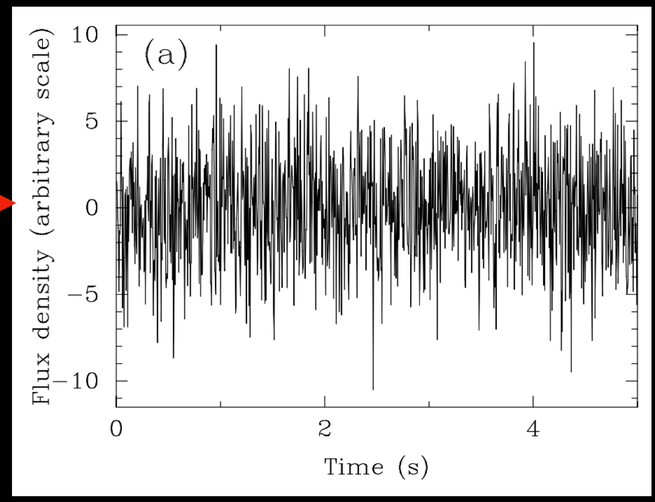

**Dedispersion**

**Search across several trials**

$\mathcal{O}(NlogN) \times 10^3$

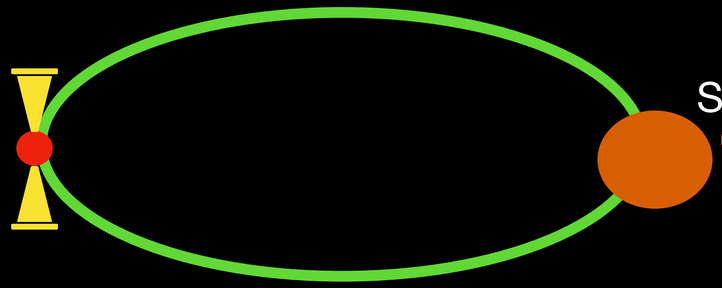



**Average over y axis to form timeseries**



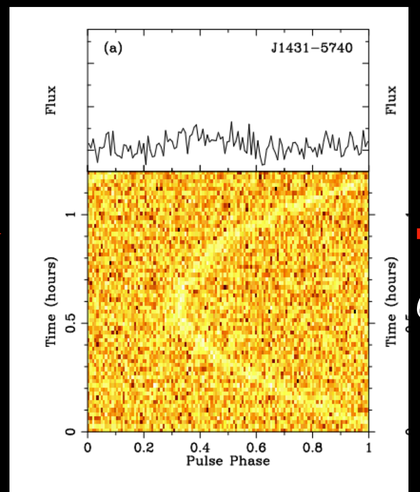

Size: 200 GB - 2 TB/Beam

**Correct for Doppler Modulation of Period**



Solve Keplerian Orbit

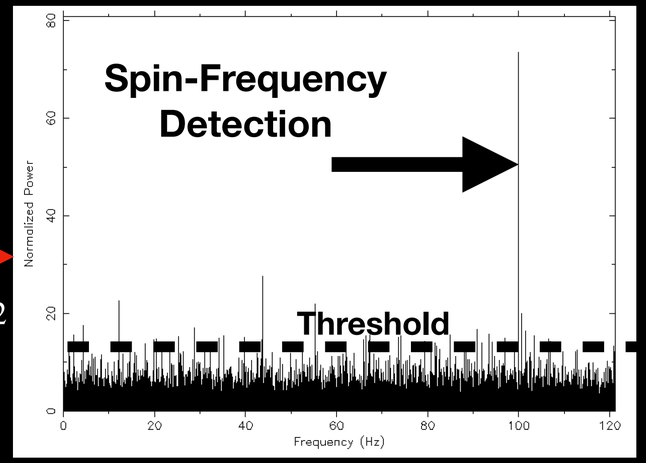
$\mathcal{O}(N) \times 10^{3-9}$




Find Spin Period

$\mathcal{O}(NlogN) \times 10^{12}$

**FFT**



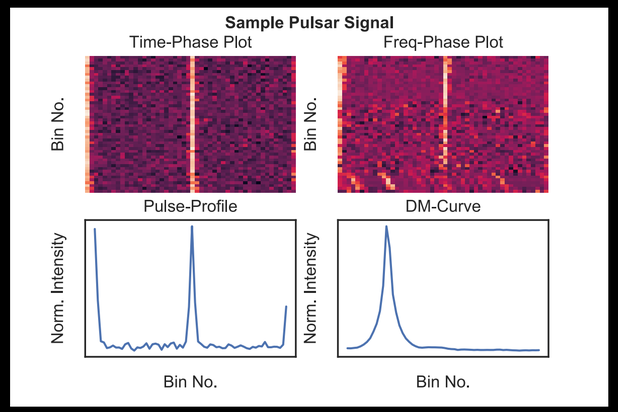

Five unknowns: $P_b, x, T0, e, \omega$

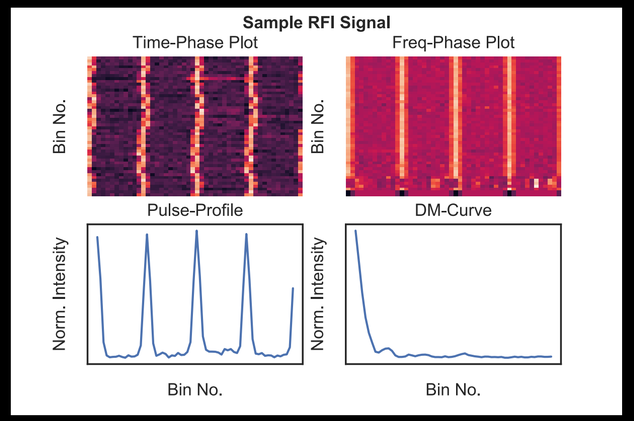
# ML in Traditional Pulsar Searches

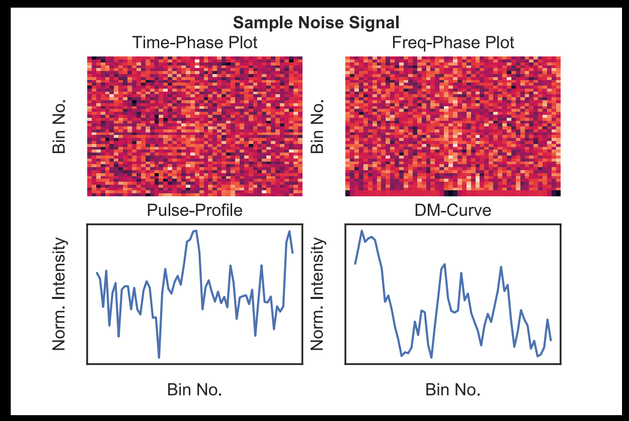
## What is a Pulsar Candidate?

**Four Dimensional Data cube consisting Frequency, time, rotational phase and power of signal.**

## Classification Problem (Solved!)




Several published papers (since 2010) attempting to solve this problem.

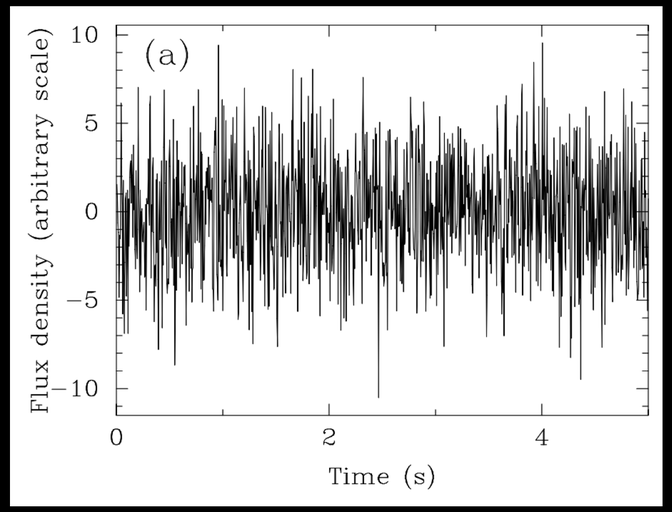
Architectures Used

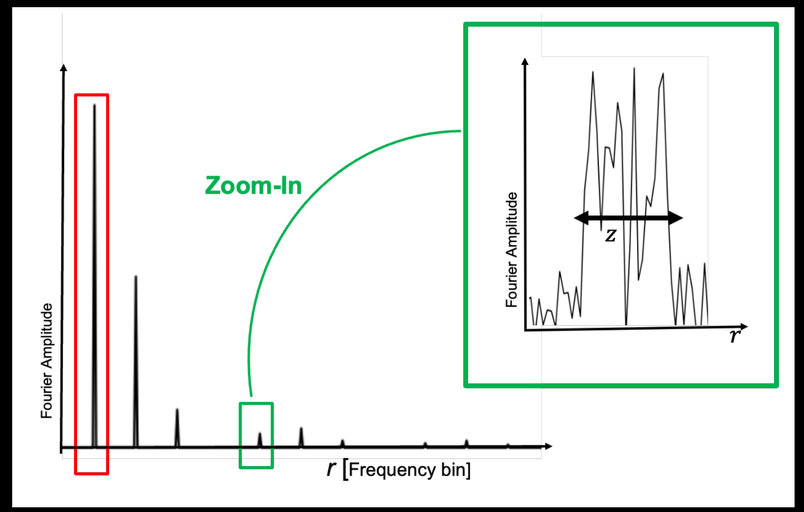
MLP, CNN, SVM, GANs, Decision Trees, Boosting Algorithms, Autoencoders, Ensemble Models

Recall rate close to 100 %! Already found more than 100 pulsars!

# Interesting Problems that ML can help with in Binary Pulsar Searches

## Parameter Estimation




Can an ML network read the smeared FFT signal and predict the orbital parameters?

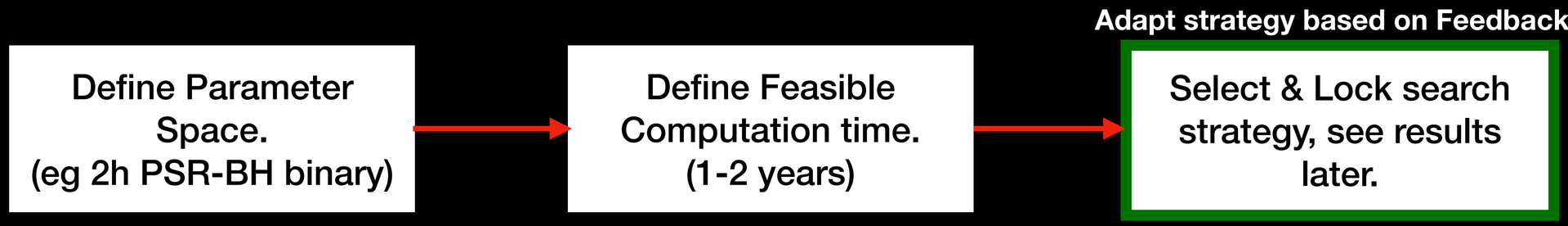
PulsarNet: Attention-based Neural Networks to Discover Pulsar Binaries. Check out his poster!

Much faster than classical approaches (simulated data in white-noise regime)

Mr. Abhinav Tyagi (Master's student)

# Interesting Problems that ML can help with in Binary Pulsar Searches

Classical Approach:




Modelling Pulsar Searching (Parameter Estimation) as a game.

Rules:

1. State S is the current position in the parameter space. Environment is our data stream.
2. Agent has multiple actions to choose from. Eg: Run classical algorithm x with some filter setting.
3. High Reward for finding a new exciting pulsar, moderate reward for a redetection of known pulsar.
4. No reward for interference signals. Add time penalty to save computation.
5. Isn't this a Markov Decision process (MDP)?
6. Task: Find a more optimal policy to maximise rewards than our current strategy.
7. Pros: Can adapt based on changes in environment. For eg: Increased levels of Radio frequency interference, bright transient events.

Can we use Reinforcement Learning for this problem?

# Thank You!