Measuring the seismic parameters of subgiant using machine learning.

Presented by: Gursharan singh

Supervised by: Prof. Shravan Hanasoge

I would like to special thank **Dr. Meenakshi Gaira** for helping and discussing every step of the project.

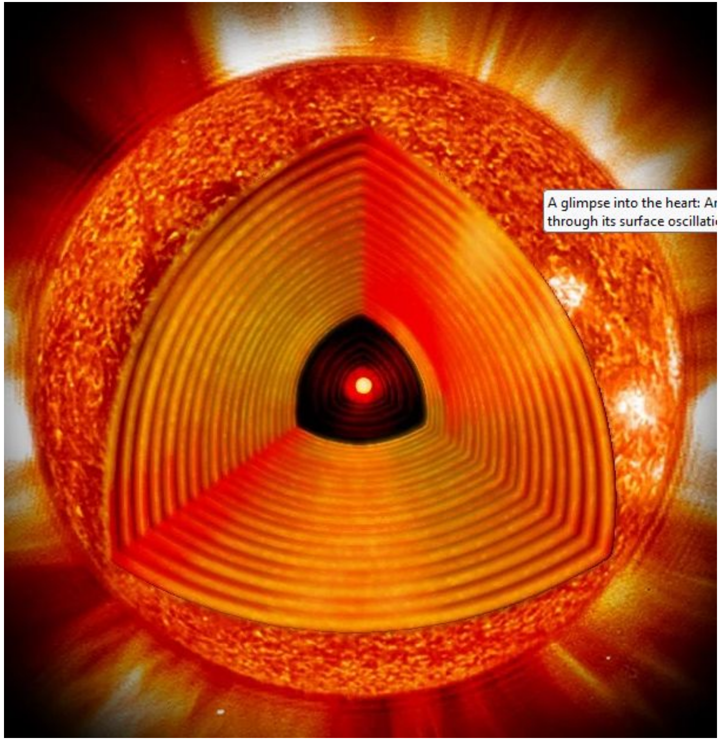
Tata Institute of Fundamental Research
Mumbai, India

Outline of Talk

- Asteroseismology
- Types of Modes
- Mix Modes
- Machine Learning
- Results

Asteroseismology

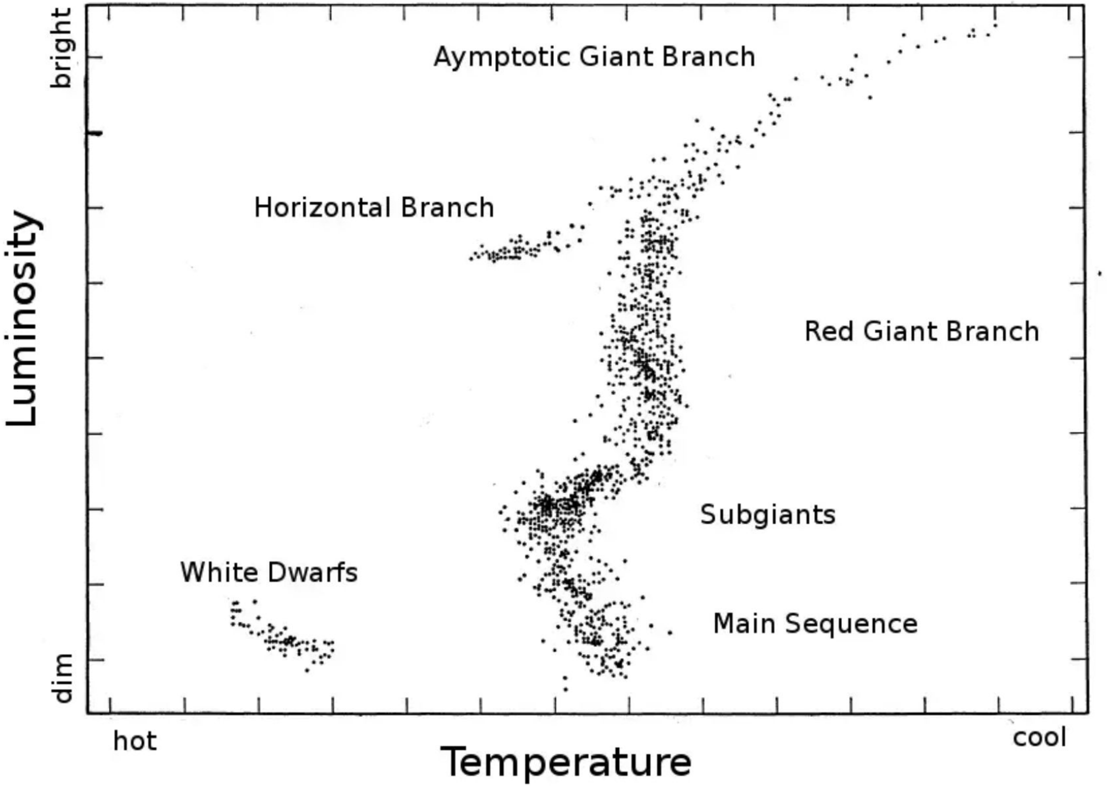
Asteroseismology is the study of the internal structures of stars through the analysis of their intrinsic global oscillations.



Reference: http: //www.physics.usyd.edu.au/~bedding/kepler/, Earl Bellinger / ESA

Why Subgiant ?

Studying subgiants is crucial to building our understanding of the evolutionary transition from the main sequence to the red giant phases.

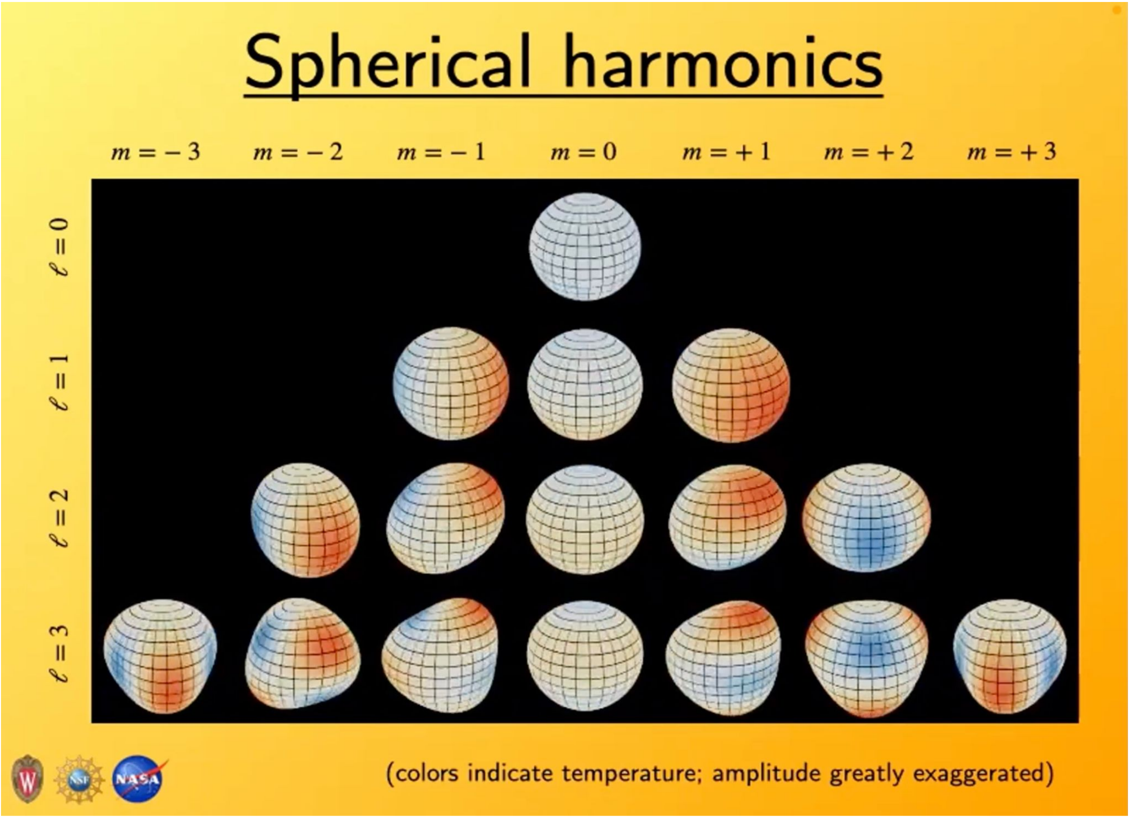


https://www.researchgate.net/figure/The-different-evolutionary-phases-of-low-to-intermediate-mass-stars-in-a_fig1_369189308

Oscillations

An Oscillation is characterized by:

- The radial order, n
- The angular degree, ℓ
- The azimuthal order, m



Type of Modes

Stellar **oscillations** are characterized by different **modes**, which describe how waves propagate inside a star.

Types of Oscillation Modes in Stars:

- P modes
- G modes

P modes

- Restoring force arises from the pressure gradient.
- Mainly confined in the convective envelope.

G modes

- Restoring force arises from the Buoyancy effect
- Mainly confined in the radiative region of the star.

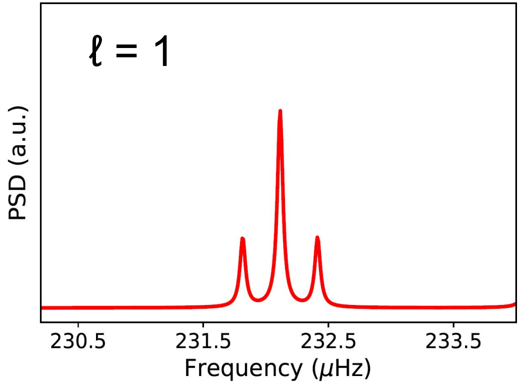
Effect due to Rotation

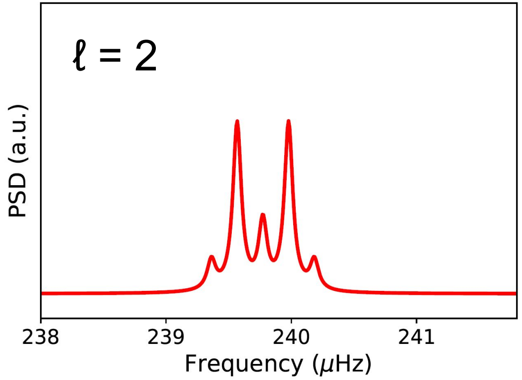
https://link.springer.com/article/10.1007/s00340-024-08190-4

Its breaks the spherical symmetry assumption and degeneracy in m [i.e. now for each ℓ we have $2\ell+1$ modes in m] .

Example:

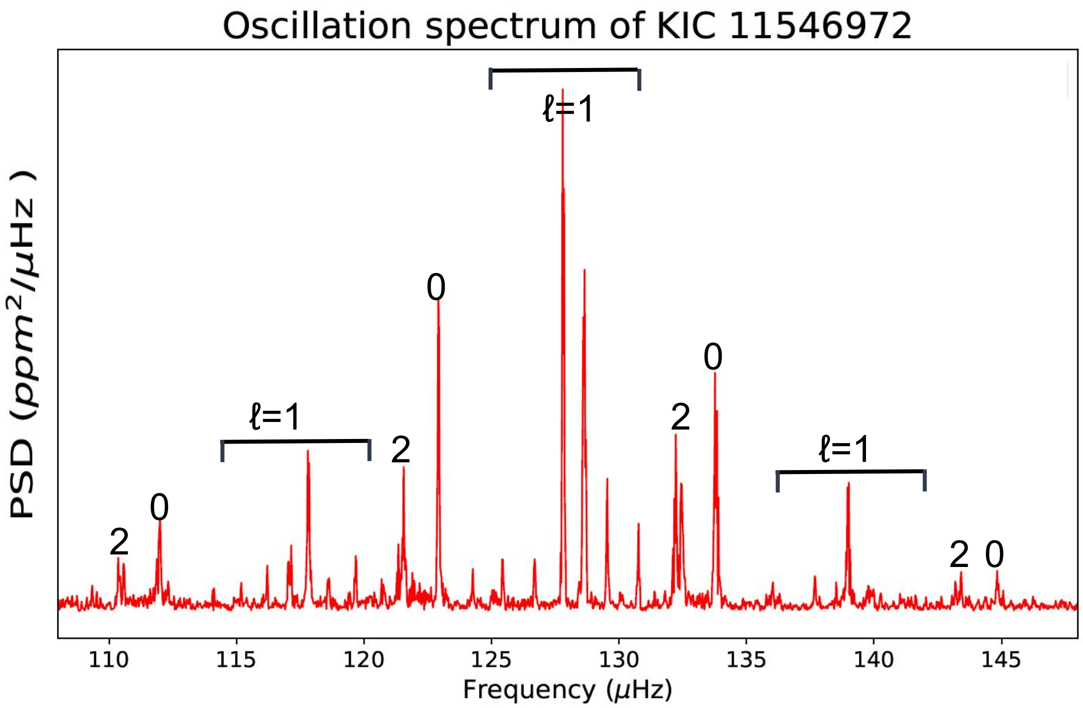
- ℓ = 1 -> m = -1 , 0 , 1
- ℓ = 2 -> m = -2 , -1 , 0 , 1 ,2





Mixed Modes

- Results from the coupling of p and g modes.
- In main sequence stars, p modes and g modes frequencies are far apart.
- **In sub-giants, core contracts and envelope inflates =>** p-modes and g-modes frequencies come closer
- Mixed modes reveal the internal structure and dynamics due to their g-mode behaviour.



Asymptotic theory of stellar oscillation

For **n** $\gg$ **l**, we observe the following characteristics:

- **P-modes (Pressure Modes)**
  - These modes exhibit a special characteristic: the spacing between adjacent peaks of same ℓ in the power spectrum is approximately equal to **Δv**, the large frequency separation.

- **G-modes (Gravity Modes)**
  - Unlike p-modes, g-modes of same ℓ are **equally spaced in period [1/v]**, meaning their separation in the period domain remains nearly constant rather than in frequency.

Asymptotic theory of stellar oscillation

- Frequencies of p modes

$$\nu_{p;n,l} = \left(n_p + \frac{l}{2} + \epsilon_p\right)\Delta\nu - d_{0l}\Delta\nu$$

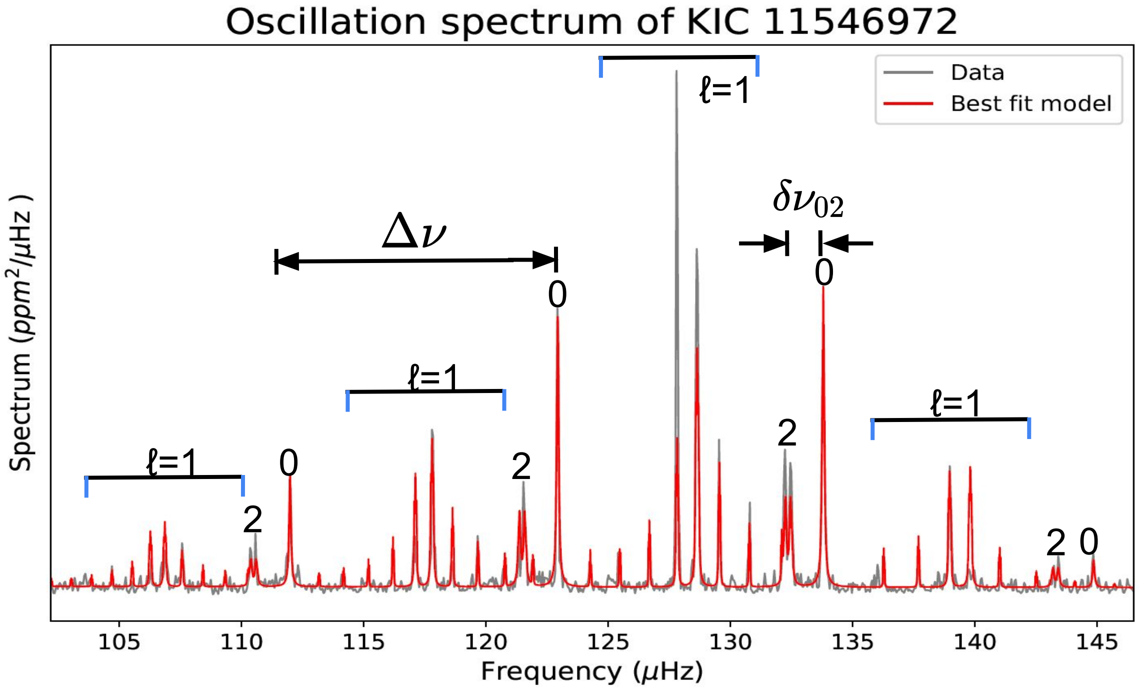
- Frequencies of g modes

$$\frac{1}{\nu_g} = (-n_g + \epsilon_g)\Delta\Pi$$

- Frequencies of Mixed modes

$$\tan\pi\frac{\nu - \nu_p}{\Delta\nu} = q\tan\frac{\pi}{\Delta\Pi}\left(\frac{1}{\nu} - \frac{1}{\nu_g}\right)$$

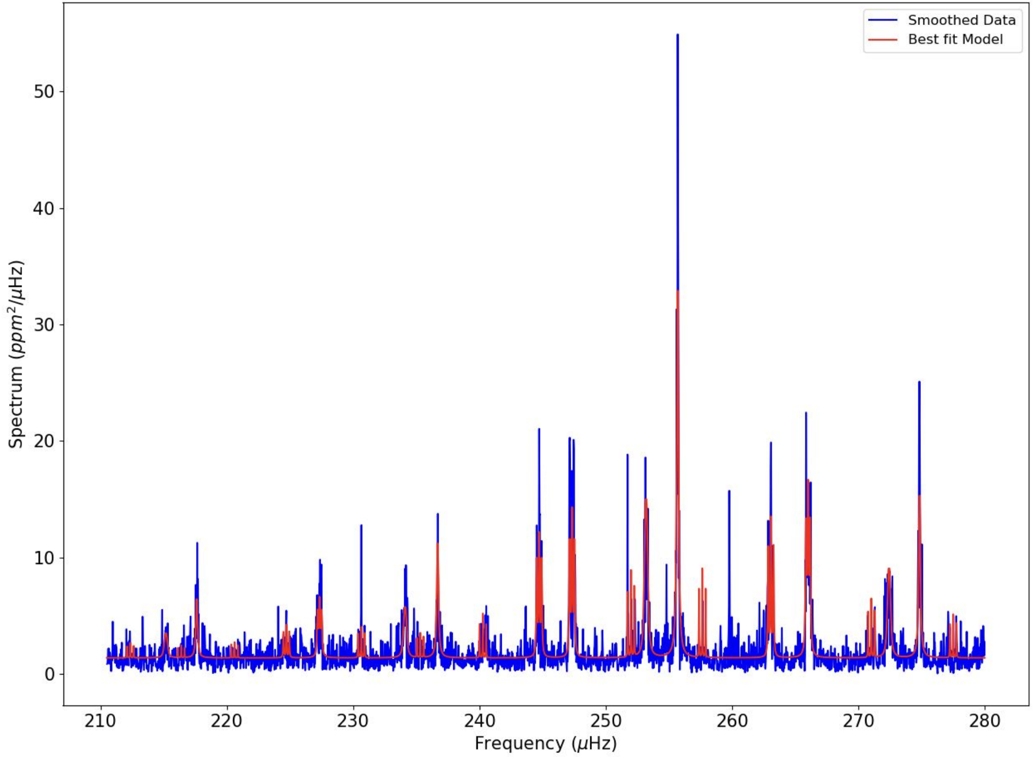
q is the coupling factor.
ΔΠ is the period spacing between pure dipolar g modes.



$$\delta\nu_{0,2} = d_{02}\Delta\nu$$

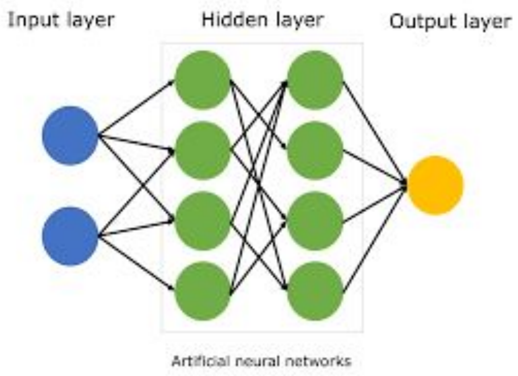
MCMC [Monte Carlo Markov Chain]

- Time consuming weeks.
- Computationally expensive.



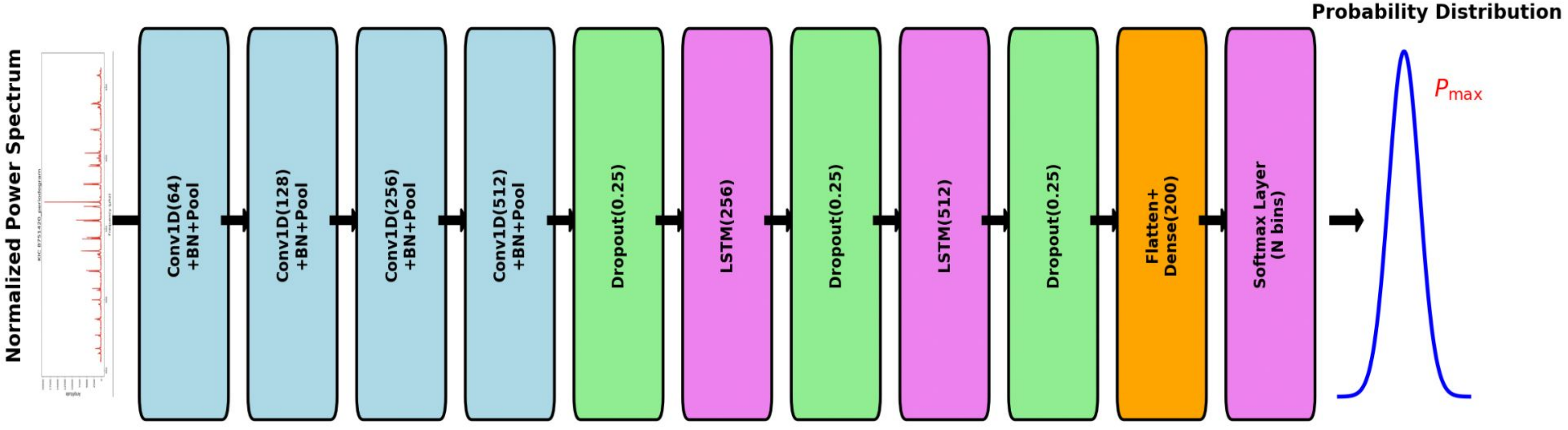
Machine learning

- Predictive Model.
- After Training it gives results in seconds.



https://www.google.com/imgres?q=neural%

Hybrid Model CNN-LSTM



CNN {convolutional neural network}

LSTM{LONG SHORT TERM MEMORY}

Synthetic Data

We have generated the synthetic Data using the using a spectra simulator available on GitHub https://github.com/OthmanB/Spectra-Simulator-C.

We have set the $\Delta\nu$ range between 18 to 75 μHz

And $\Delta\Pi$ between 60 to 600 sec

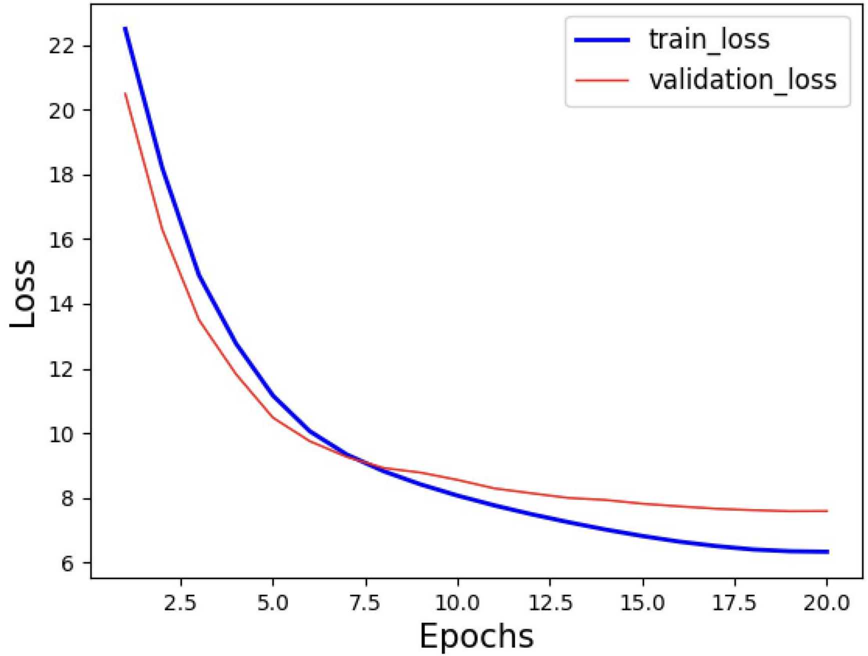
And coupling q between 0 to 0.7

All ranges are Uniform

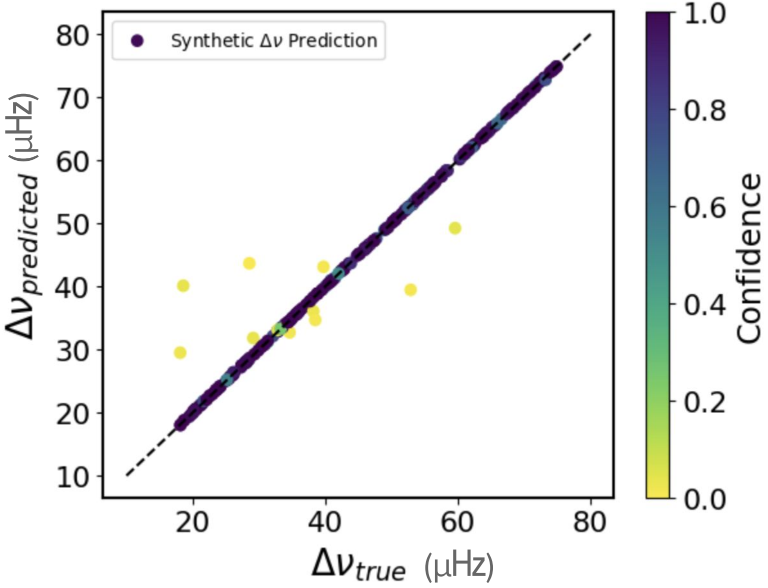
Training

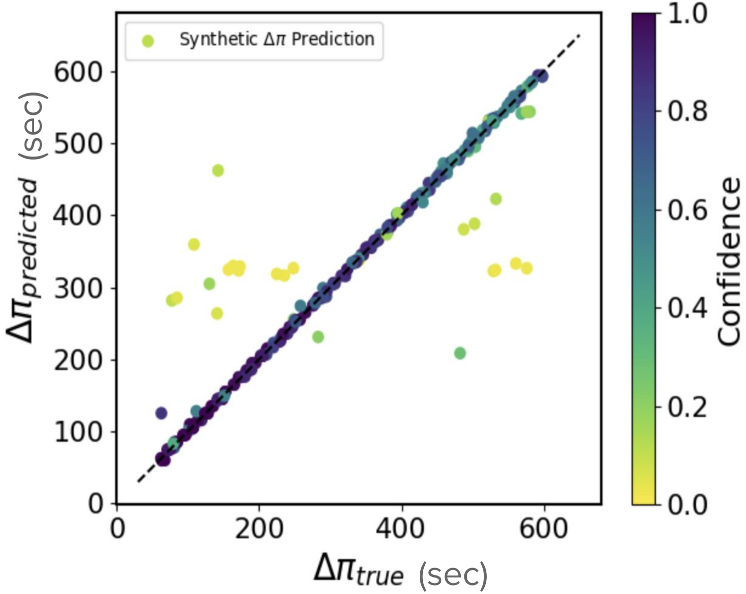
We have generated 1,024,000 samples of synthetic subgiant oscillation spectra.

We split the synthetics into 64000 instances for validation and 64000 for testing and use the remaining for training (896000).

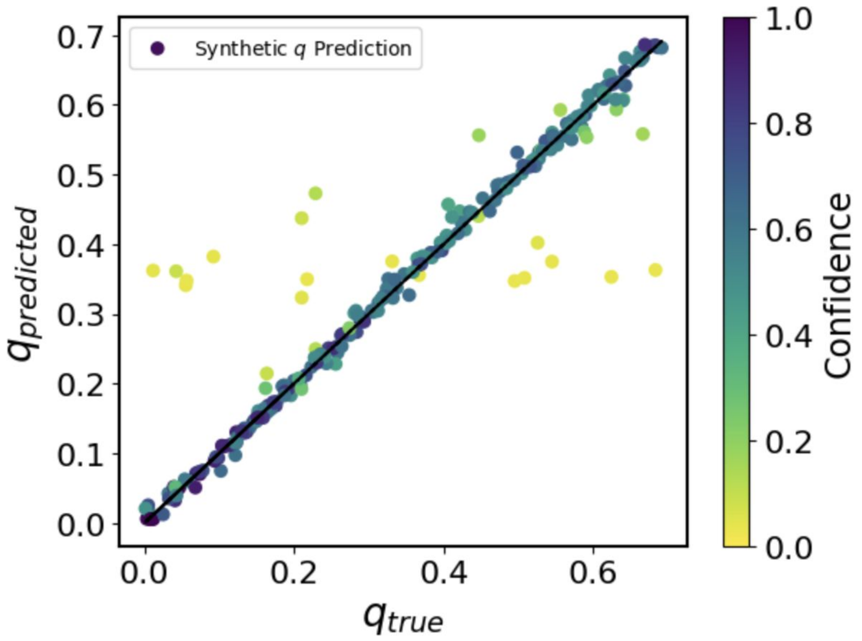


Synthetic Data Result

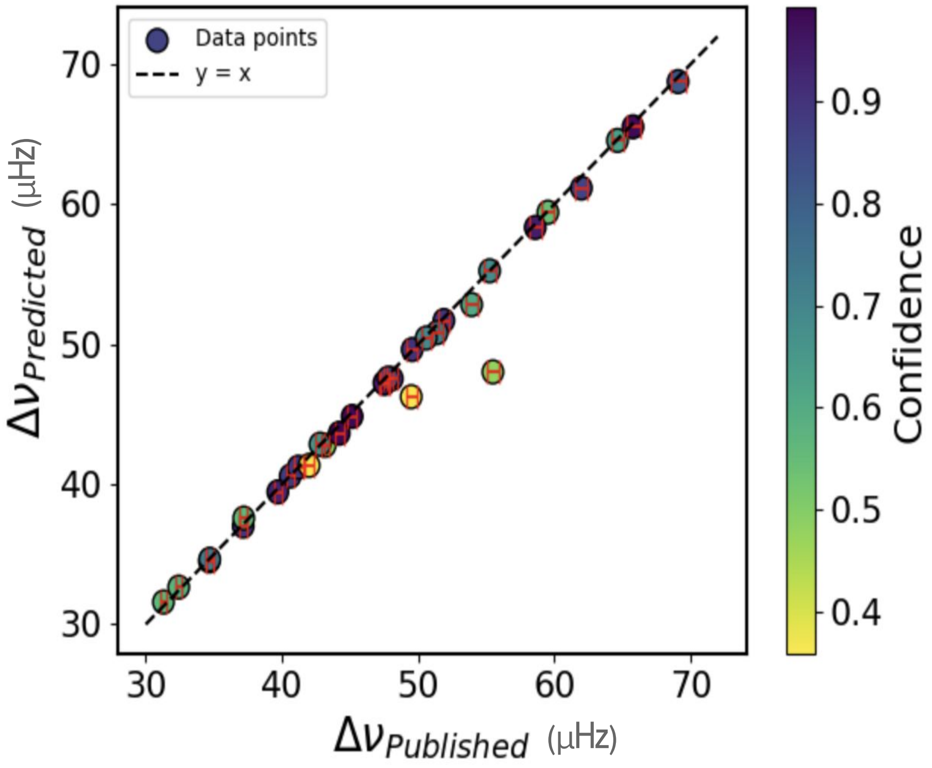


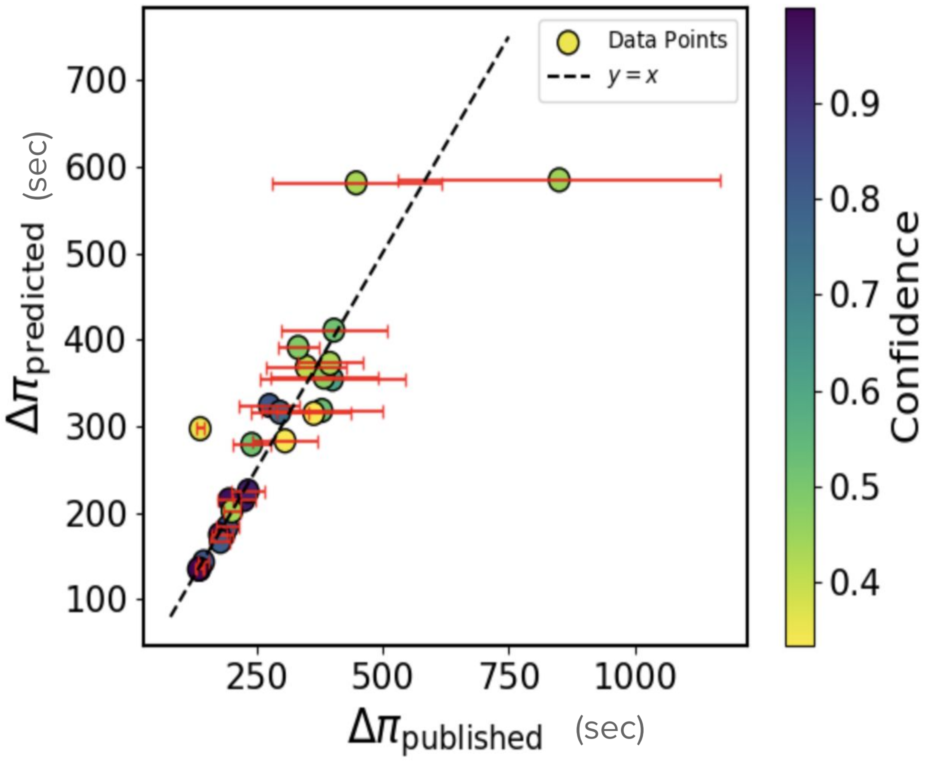


Synthetic Data Result

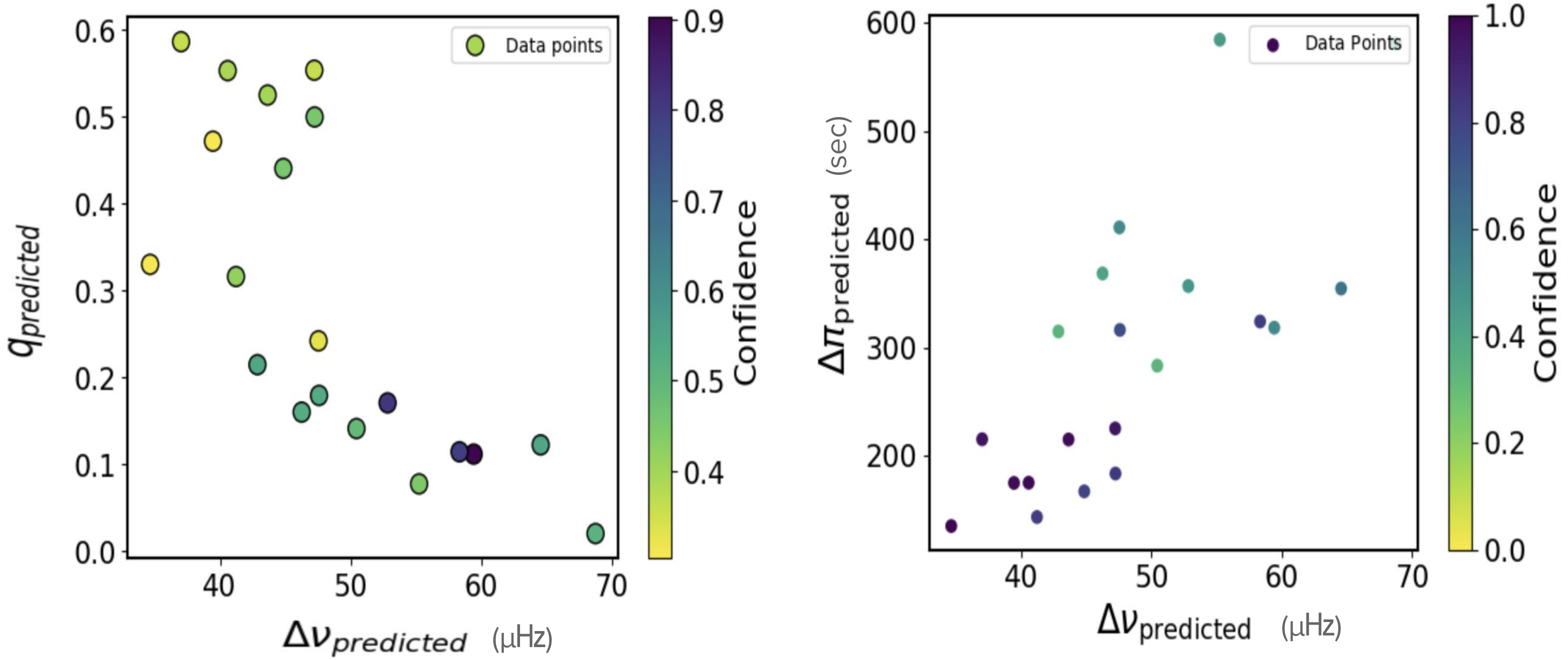


Published data results





Results



Data points
0.6
0.5
0.4
0.3
0.2
0.1
0.0
40
50
60
70
$q_{predicted}$
$\Delta\nu_{predicted}$ (μHz)
0.9
0.8
0.7
0.6
0.5
0.4
Confidence
Data Points
600
500
400
300
200
$\Delta\pi_{predicted}$ (sec)
$\Delta\nu_{predicted}$ (μHz)
1.0
0.8
0.6
0.4
0.2
0.0
Confidence

THANK YOU

References:

[1] O. Benomar, T. R. Bedding, D. Stello, S. Deheuvels, T. R. White, and J. Christensen-Dalsgaard. Masses of subgiant stars from asteroseismology using the coupling strengths of mixed modes. Astrophys. J. Lett., 745:L33, 2012.

[2] Tiago L. Campante. An Introduction to Data Analysis in Asteroseismology, page 55–74. Springer International Publishing, July 2017.

[3] Miguel Clara, Margarida S. Cunha, Pedro P. Avelino, Tiago L. Campante, S´ebastien Deheuvels, and Daniel R. Reese. On the seismic modelling of subgiant stars: testing different grid interpolation methods. 2025.

[4] Siddharth Dhanpal, Othman Benomar, Shravan Hanasoge, Abhisek Kundu, Dattaraj Dhuri, Dipankar Das, and Bharat Kaul. Measuring frequency and period separations in red-giant stars using machine learning. The Astrophysical Journal, 928(2):188, April 2022.

[5] Rafael A. Garc´ıa and J´erˆome Ballot. Asteroseismology of solar-type stars. Living Reviews in Solar Physics, 16(1):4, 2019.

[6] Saskia Hekker and Anwesh Mazumdar. Solar-like oscillations in subgiant and red-giant stars: mixed modes. Proceedings of the International Astronomical Union, 9(S301):325–331, 2013.

[7] Marc Hon, Earl P Bellinger, Saskia Hekker, Dennis Stello, and James S Kuszelewicz. Asteroseismic inference of subgiant evolutionary parameters with deep learning. Monthly Notices of the Royal Astronomical Society, 499(2):2445–2461, September 2020.

[8] B. Mosser, O. Benomar, K. Belkacem, M. J. Goupil, N. Lagarde, E. Michel, Y. Lebreton, D. Stello, M. Vrard, C. Barban, T. R. Bedding, S. Deheuvels, W. J. Chaplin, J. De Ridder, Y. Elsworth, J. Montalban, A. Noels, R. M. Ouazzani, R. Samadi, T. R. White, and H. Kjeldsen. Mixed modes in red giants: a window on stellar evolution. Astronomy amp; Astrophysics, 572:L5, November 2014.

[9] Tian, Zhijia, Bi, Shaolan, Bedding, Timothy R., and Yang, Wuming. Asteroseismic analysis of solar-mass subgiants kic 6442183 and kic 11137075 observed by kepler. AA, 580:A44, 2015

Back up slides

Coupling

The coupling strength **q** quantifies how strongly **p-modes (pressure modes)** and **g-modes (gravity modes)** interact in a star.

- **Higher coupling (larger q)** → Indicates that the **p-mode and g-mode cavities overlap more**, allowing stronger energy exchange between the two modes. This results in **mixed modes** that have characteristics of both p- and g-modes.

- **Lower coupling (smaller q)** → Means the p- and g-mode cavities are more **separated**, leading to weaker interaction and less mode mixing.

Scaling relations

$$R \approx R_\odot \left(\frac{\nu_{\max}}{\nu_{\max\odot}}\right)\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{1/2}$$

$$M \approx M_\odot \left(\frac{\nu_{\max}}{\nu_{\max\odot}}\right)^{3}\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-4}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{3/2}$$

Accuracy result

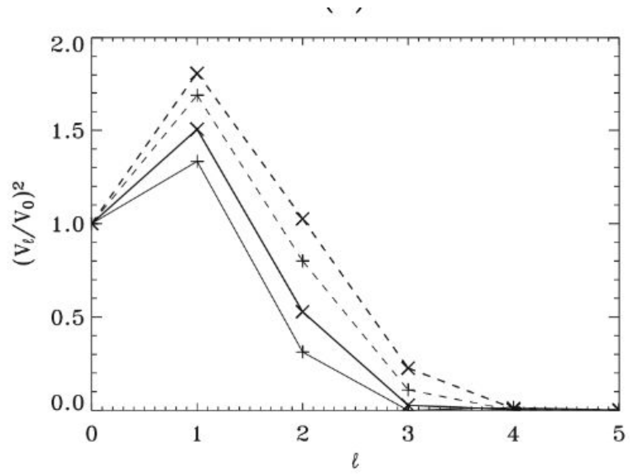
We have obtained a validation accuracy of 84.5%, 63.4%, and 54.1% for $\Delta\nu$, $\Delta\Pi$, q respectively.

Visibility

Modes of higher degree **ℓ** do not appear in the power spectrum as their amplitudes are significantly reduced due to **geometric cancellation** when observed over the entire stellar disk.

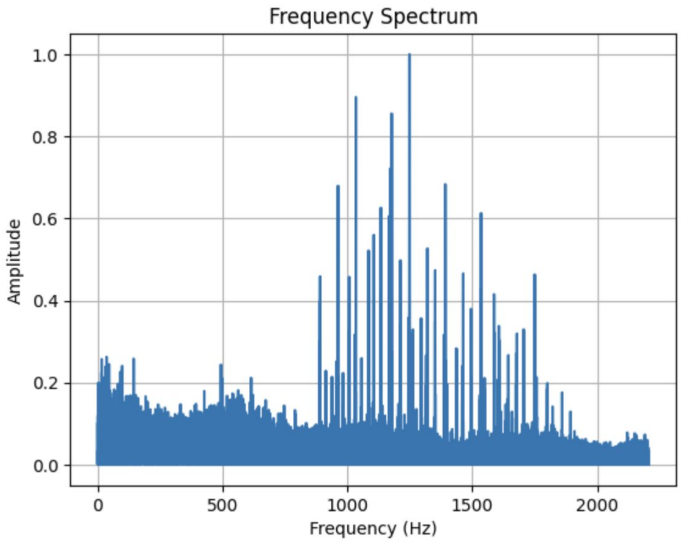
The observed amplitude may be expressed as:

$$a_{n,\ell,m} = r_{\ell,m}(i) V_\ell A$$

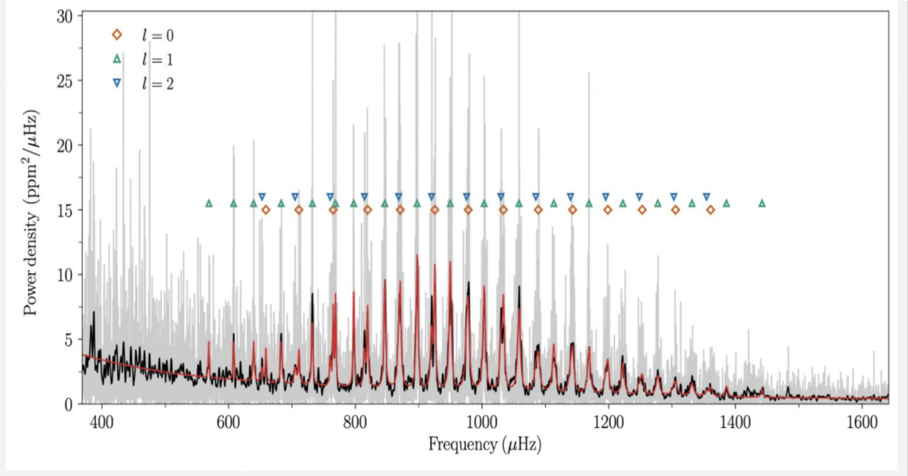


Synthetic vs original

Synthetic

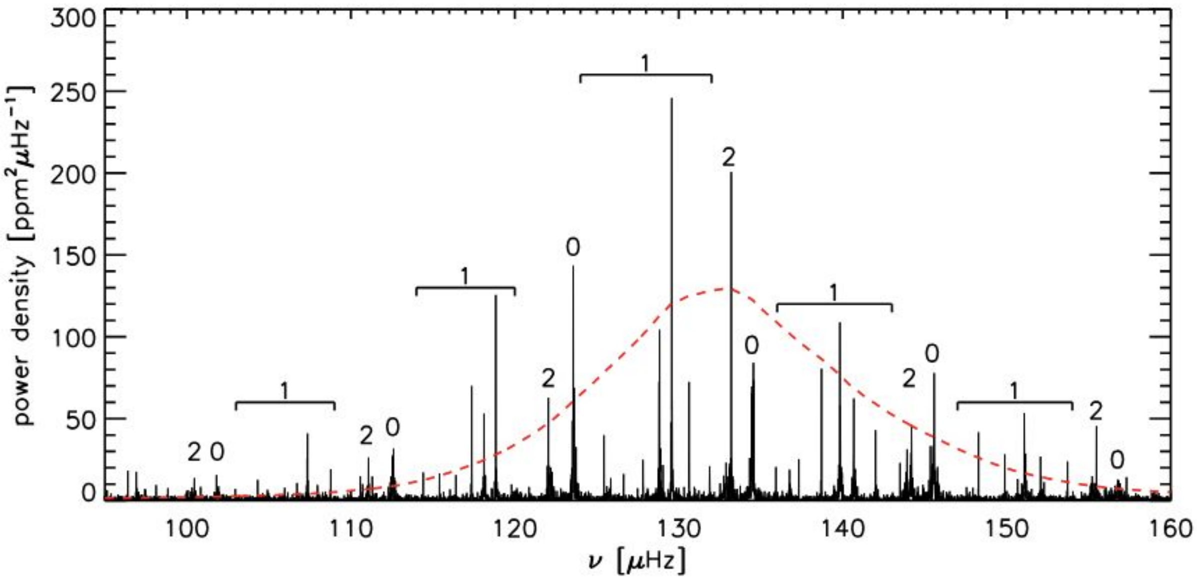


Original



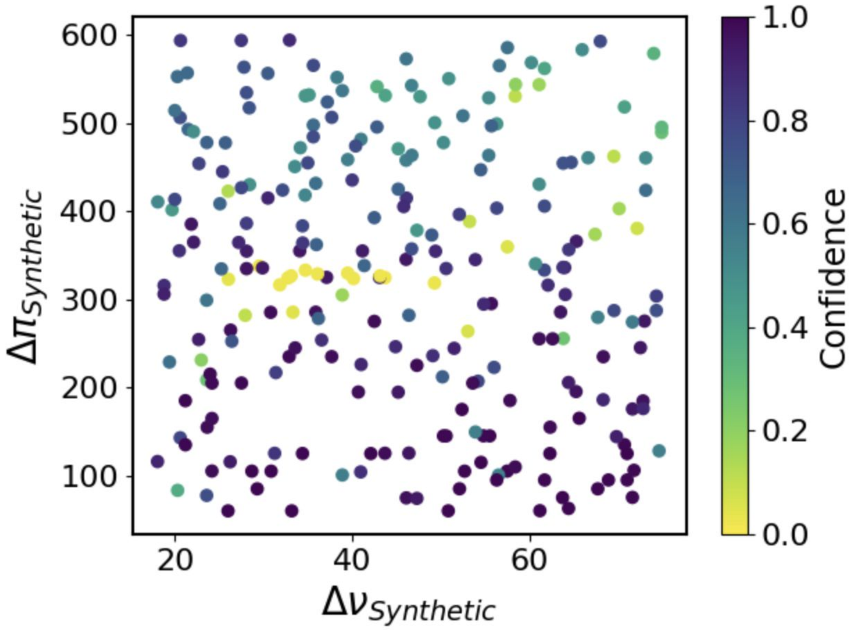
Power Spectrum

It is a absolute squared of the fourier transform of the light curve.

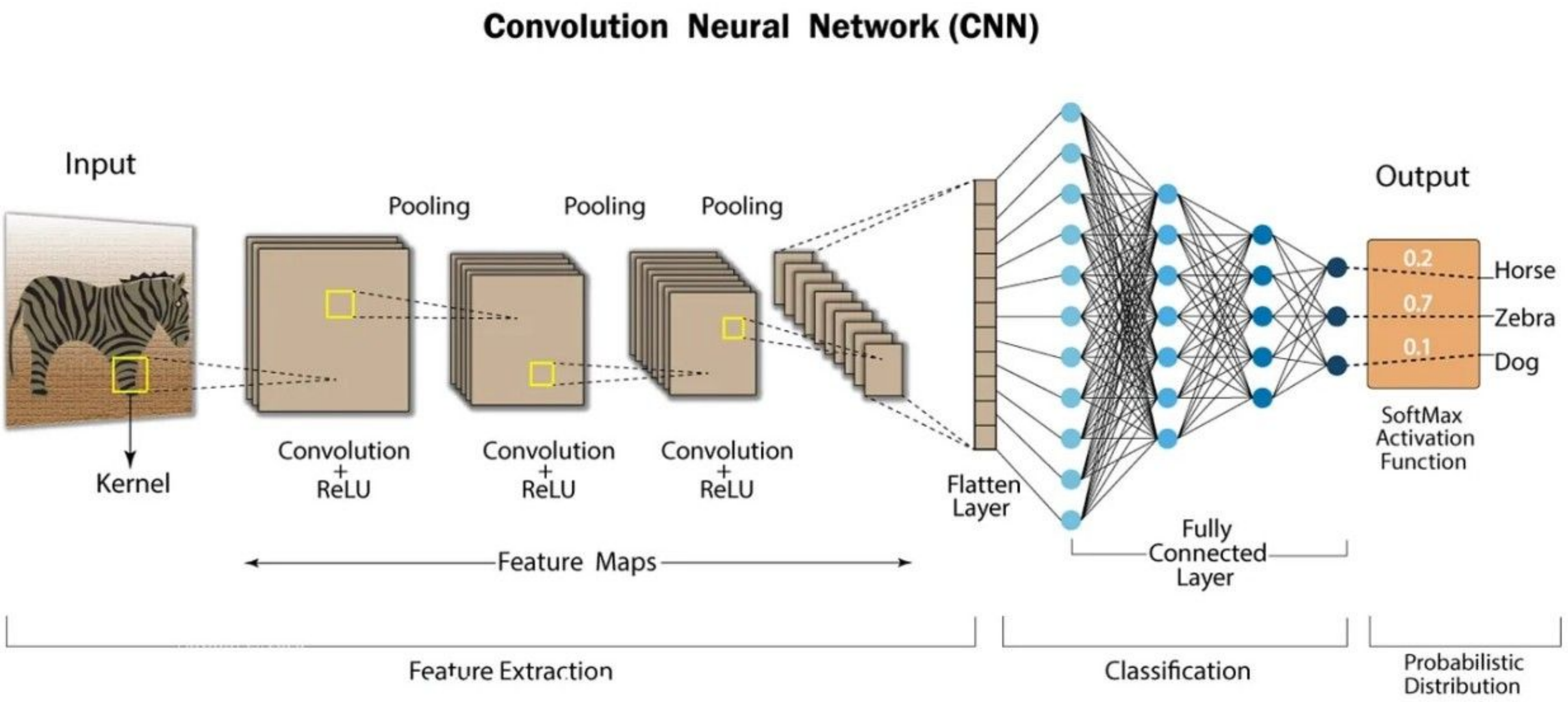


Power spectra of a red giant kic=

Machine predicted the uniform distribution

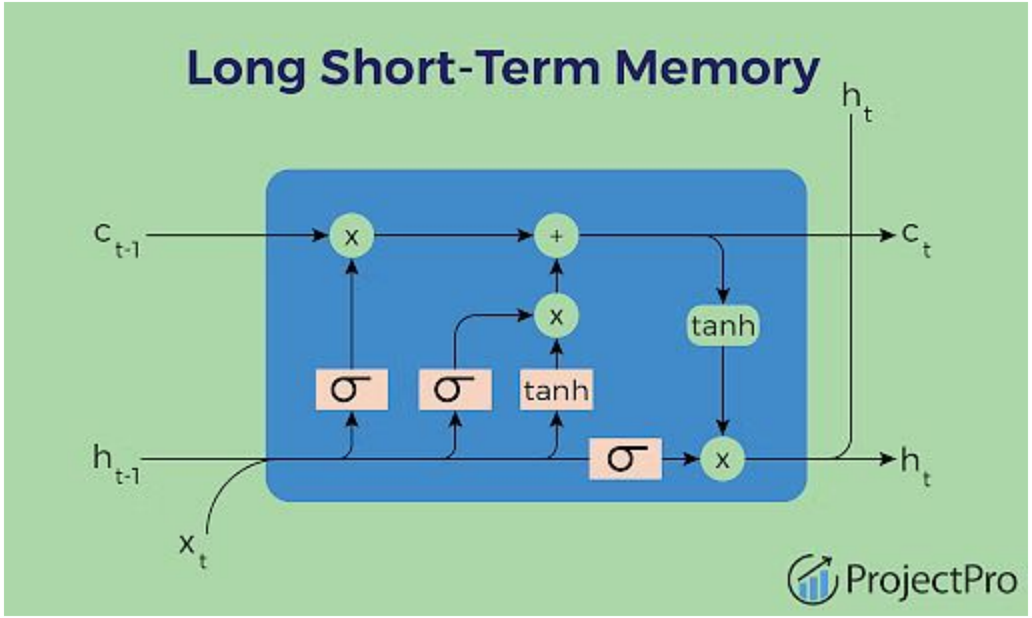


CNN



Convolution Neural Network (CNN)
Input
Pooling
Pooling
Pooling
Output
0.2
0.7
0.1
Horse
Zebra
Dog
SoftMax Activation Function
Kernel
Convolution + ReLU
Convolution + ReLU
Convolution + ReLU
Flatten Layer
Fully Connected Layer
Feature Maps
Feature Extraction
Classification
Probabilistic Distribution

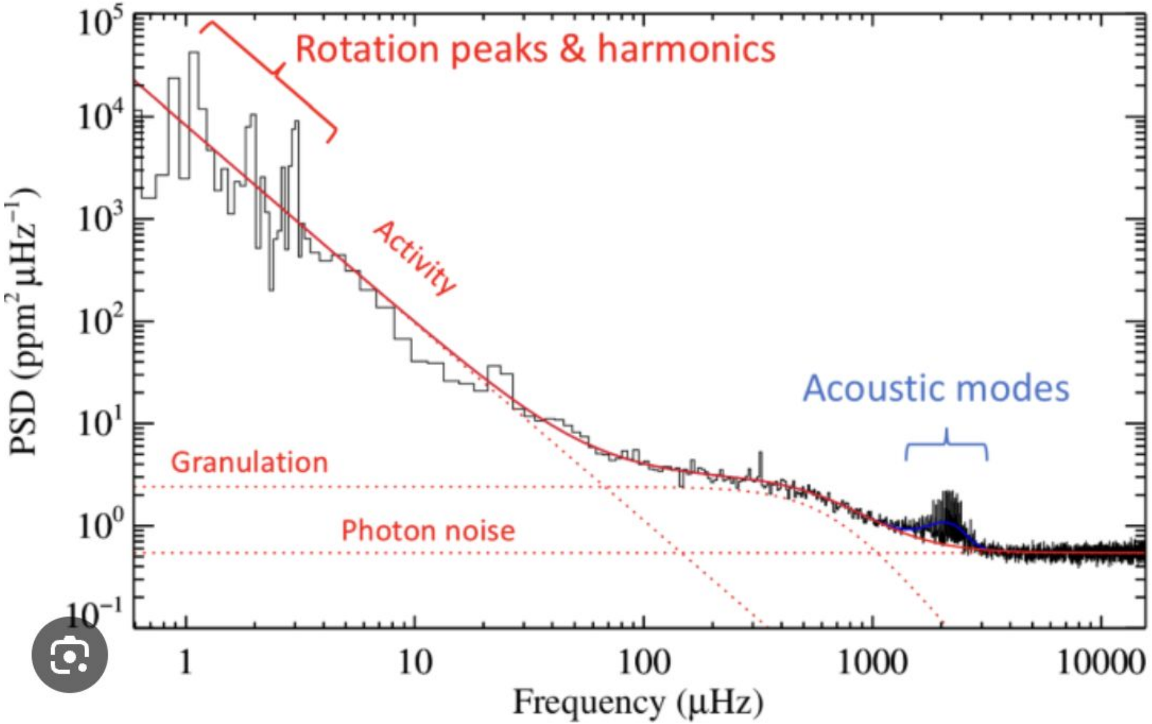
LSTM



Long Short-Term Memory
h_t
c_{t-1}
x
+
c_t
x
tanh
σ
σ
tanh
h_{t-1}
σ
x
h_t
x_t
ProjectPro

LSTM working

noise



Rotation causes brightness variations due to starspots moving in and out of view.

Activity" represents the **stellar magnetic activity** and **convection-related variability**.

Photon noise is the noise due to the **detector (telescope CCD noise) and photon counting statistics**.

The term $\nu_{\max}$ in Equation (A1) refers to the frequency corresponding to maximum amplitude. Observations (Stello et al. 2009) and scaling relations (Ulrich 1986) have demonstrated a strong interdependent relation between $\Delta\nu$ and $\nu_{\max}$. For our simulations, we choose $\nu_{\max}$ based on $\Delta\nu$ from the relation given in Stello et al. (2009), with 10% deviation, as follows:

$$\nu_{\max} = (\Delta\nu/0.263)^{1/0.772} \pm 0.1(\Delta\nu/0.263)^{1/0.772}. \qquad (A2)$$

Cadence -226sec short cadence

Tobs-2year

- The **Nyquist frequency** is the highest frequency that can be correctly sampled given the cadence.
  - **For Long Cadence (LC):** Nyquist frequency ≈ **283 μHz**
  - **For Short Cadence (SC):** Nyquist frequency ≈ **8500 μHz**
- If the oscillation frequency is above the Nyquist frequency, it can create **aliasing artifacts** in the spectrum.

Solar-like oscillations are pressure-driven oscillations (**p-modes**) observed in stars with convective envelopes. These oscillations are excited by **turbulent convection** and are stochastically driven and intrinsically damped, similar to oscillations in the Sun.

Yes, the **power spectrum** is indeed the **absolute square** of the **Fourier transform** of the time-series data.

Detailed Explanation:

The process to obtain the power spectrum from a time series, like a star's light curve, involves the following steps:

1. **Fourier Transform**

   The **Fourier transform** is used to convert the time-series data (which is in the **time domain**) into the **frequency domain**. This tells us how much of the signal (e.g., the light curve) exists at each frequency. Mathematically, the Fourier transform $\mathcal{F}(t)$ of a time series $f(t)$ is given by:

   $$\mathcal{F}(f(t)) = \int_{-\infty}^{\infty} f(t)e^{-i2\pi ft}dt$$

   where $f(t)$ is the time-series data and f is the frequency.

2. **Power Spectrum**

   The **power spectrum** is then calculated as the **absolute square** of the Fourier transform:

   $$P(f) = |\mathcal{F}(f(t))|^2$$

   where $P(f)$ represents the **power at each frequency** f.

Reconciliation of the Statements:

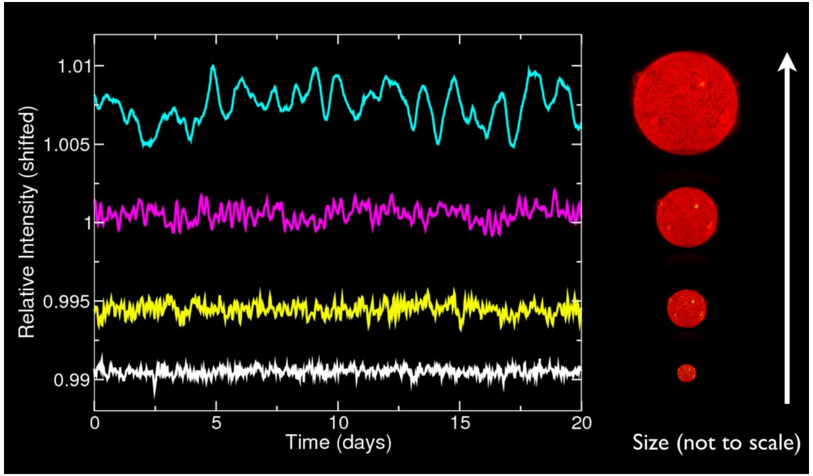
- **Shorter cadence** increases the Nyquist frequency, which allows you to **detect higher-frequency oscillations** (those occurring at higher frequencies, such as p-modes).
- **Longer cadence** reduces the Nyquist frequency, making it difficult to detect **high-frequency oscillations**, as they will be undersampled and lead to aliasing. The higher the frequency of oscillations, the **shorter the cadence** required to detect them without aliasing.

Summary of the Correct Understanding:

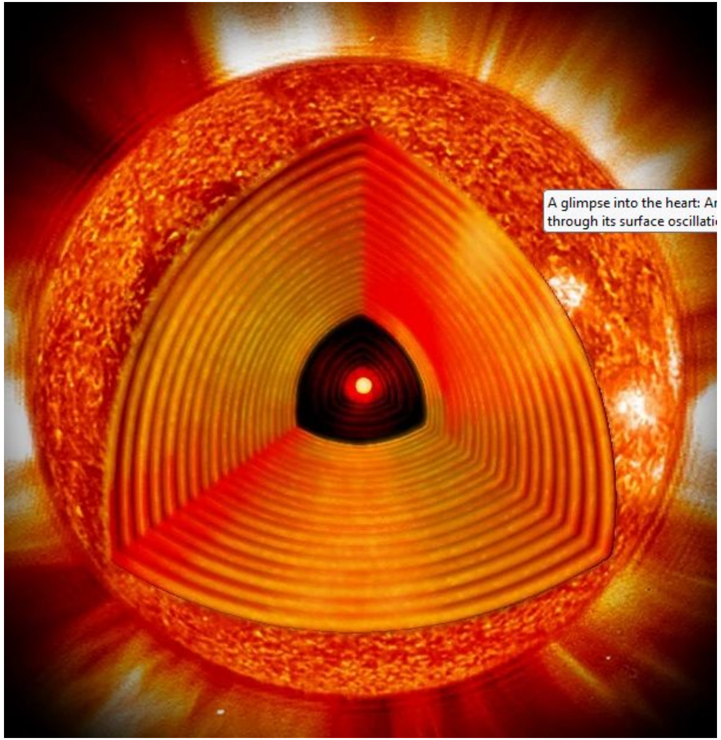
- **Shorter cadence → higher Nyquist frequency → can detect higher-frequency oscillations.**
- **Longer cadence → lower Nyquist frequency → limits detection of high-frequency oscillations.**

Asteroseismology

Study of the structure and dynamics of stars through their oscillations.



https://www2.ifa.hawaii.edu/research/Stars.shtml



Reference: http: //www.physics.usyd.edu.au/~bedding/kepler/, Earl Bellinger / ESA