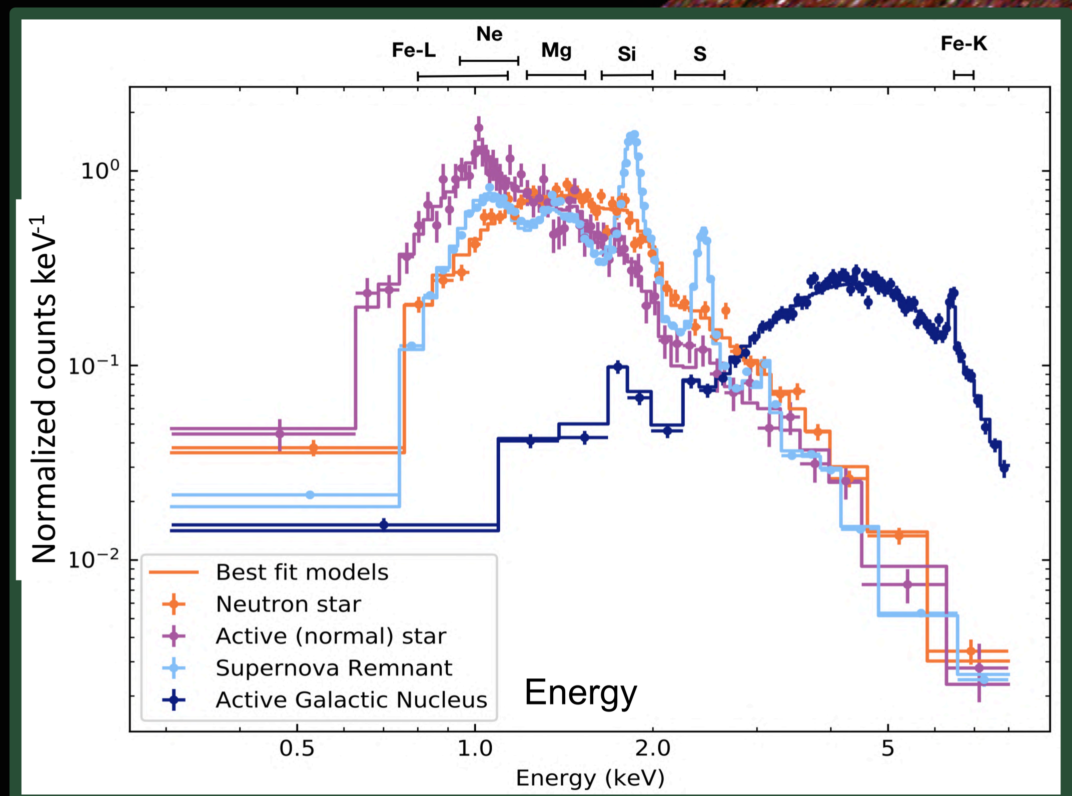
Data Driven Methods to Classify X-ray Sources

Pavan R Hebbar, Craig O. Heinke
University of Alberta

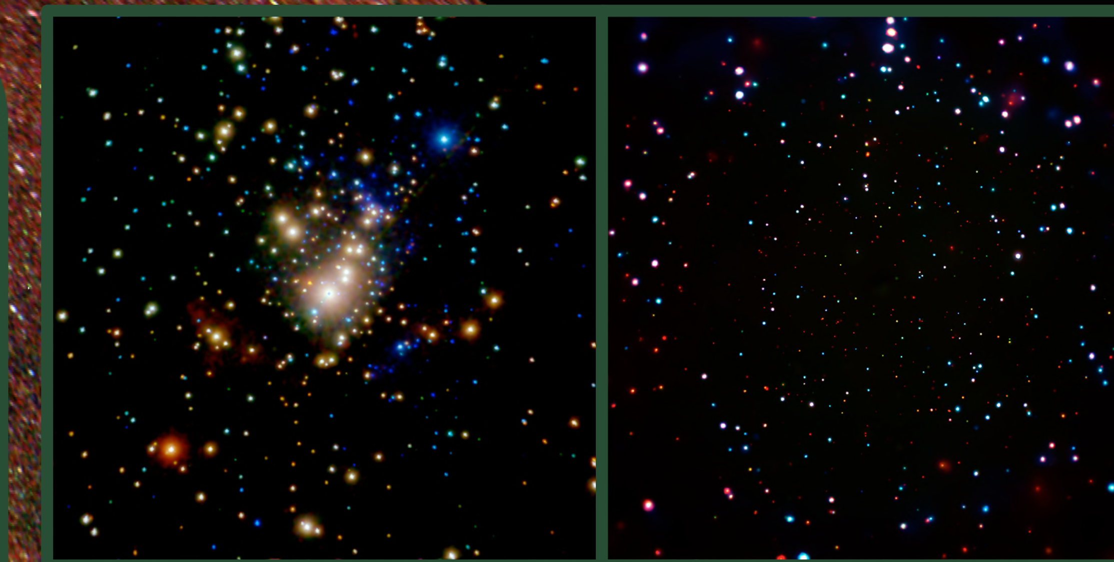


Sample spectra of neutron star (orange), active star, SNR, and AGN.

The X-ray sky

Modern X-ray telescopes have detected hundreds of thousands of X-ray sources (e.g. Evans et al. 2010). These include active galactic nuclei (AGN), supernova remnants (SNRs), active stars, and compact neutron star and black hole binaries. Common methods to distinguish these sources are modelling their X-ray spectra or using hardness ratios (HRs) to estimate the source properties.

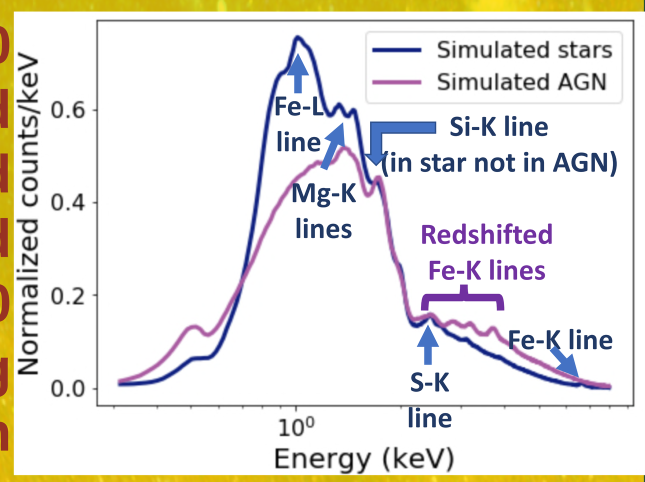
Detailed X-ray spectroscopy of thousands of sources is tedious, and HRs could fail to accurately categorize faint objects. These methods fail to employ the power of CCD detectors to identify emission lines in X-ray spectra. Machine-learning (ML) tools that can differentiate line- and continuum-dominated X-ray spectra will be accurate and efficient in classifying X-ray spectra. Here, we demonstrate the use of artificial neural networks (ANNs) to distinguish active stars in Orion nebula from AGN in Chandra Deep Field South.



X-ray image of (*Left*) Stars in Chandra Orion Ultradeep Project (COUP, Credits: CXC/ Fiegelson et al.) (*Right*) AGN in Chandra Deep Field South (CDFS, Credits: CXC/Luo et al.)

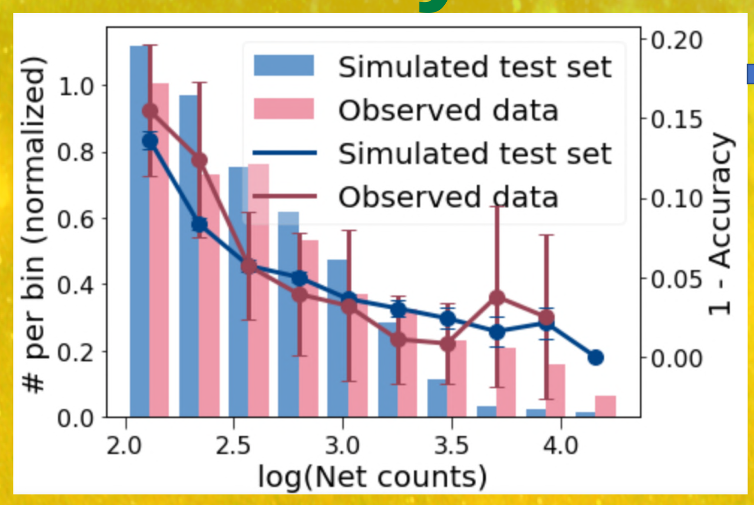
AGN vs. Active star

We simulated 100,000 spectra of AGN and stars from COUP and CDFS properties, and used spectra with >100 X-ray counts for training and testing a 1 hidden layer, 8-node ANN.

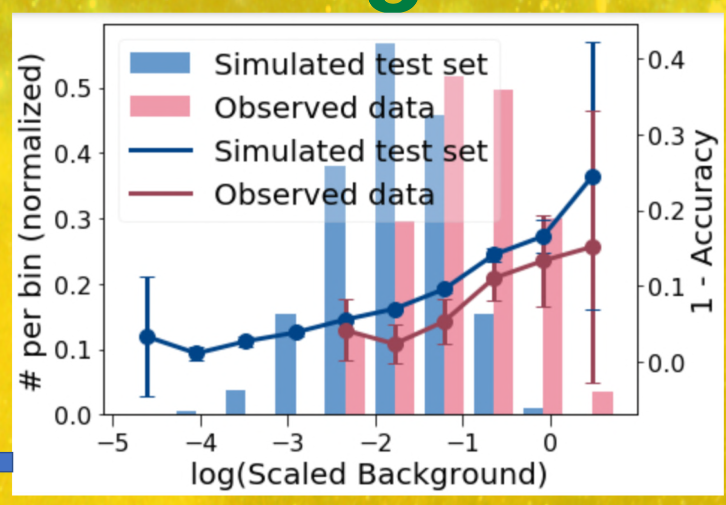


Stars show prominent emission lines from Mg, Si, S & Fe. AGN have mainly continuum spectra.

Accuracy w.r.t net counts & background

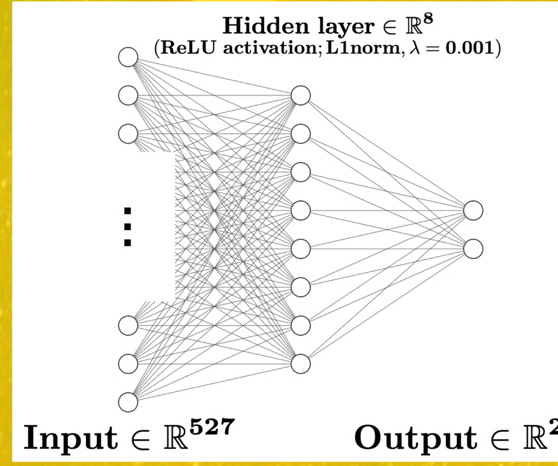


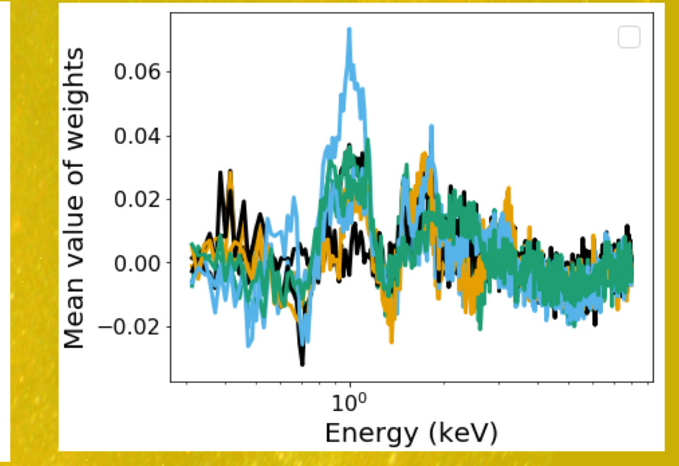
Accuracy > 90% for net counts > 300, & for background contribution < 5%



Net classification accuracy of 92% on simulated test set, and 91% on observed COUP and CDFS spectra. Our algorithm performs best on sources with net counts greater than 300 and background contribution less than 5%.

ANN model





(Left) ANN model architecture used. It outputs the probability of source being star or AGN. (Right) The weights of the trained ANN show that our model picks out the prominent emission lines for classification.

Role of source properties

- $N_H > 10^{24}$ absorbs X-rays with energy, $E < 2$ keV, leading to difficulties in detection of Fe-L, Mg & Si lines and thus identifying stars.
- Stars that were fit with temperatures, $T > 2 \times 10^7$ K are continuum-dominated and hence not recognized properly by our model.
- Our model tends to identify harder AGN with, power-law index, $\Gamma < 2$ better.
- AGN with strong Fe-K line are detected more efficiently.

Conclusions & Future

- Tested the use of NN as an accurate and efficient tool to differentiate AGN and active star spectra.
  - 100,000 spectra classified in a few minutes with ~92% accuracy.
  - Model works best for sources > 300 counts & background contribution < 5%
- Test other ML algorithms to increase accuracy and reduce size of training dataset.
- Apply algorithm on other types of sources.

References

- Evans I. N. et al. 2010, ApJS, 189, 37
- Getman K. V. et al 2005, ApJS, 160, 319
- Hebbar P. R. et al 2019, MNRAS, 485, 5604
- Park T. et al. 2006, ApJ 652, 610
- Tozzi P. et al. 2006, A&A, 451, 457

Acknowledgements

NSERC CRSNG

Pavan R Hebbar
hebbar@ualberta.ca
Submitted article:
sites.ualberta.ca/~hebbar/pdfs/Hebbar_stars_vs_AGN.pdf