Using Deep Learning for Robust Classification of Fast Radio Bursts

Abstract

- The **true nature** of fast radio bursts (FRBs) are **unknown**.
- A Supervised Variational Autoencoder (sVAE) architecture was used to **classify FRBs** and **analyzed the latent space** of the **first CHIME catalog**.
- The model achieves **state-of-the-art classification accuracy** and reveals a clear separation between repeater and non-repeater populations.
- **Spectral index** and **spectral running** are identified as the **dominant features** distinguishing repeaters.
- **Four** previously labeled **non-repeating FRBs** are identified as **high-probability repeater candidates**.

Introduction

- **Fast Radio Bursts (FRBs)** are **brief, intense bursts of radio waves** lasting only a **few milliseconds**. [11]
- Detection has increased significantly with modern radio telescopes, yet their **physical origin remains largely unknown**.
- Observations indicate that **some FRBs repeat** while others appear **as isolated events**.
- Important question: **How can we find structure & patterns in FRB signals?**
- **Supervised Variational Autoencoders (sVAEs)** are deep generative models that learn to encode and classify data [3, 6]

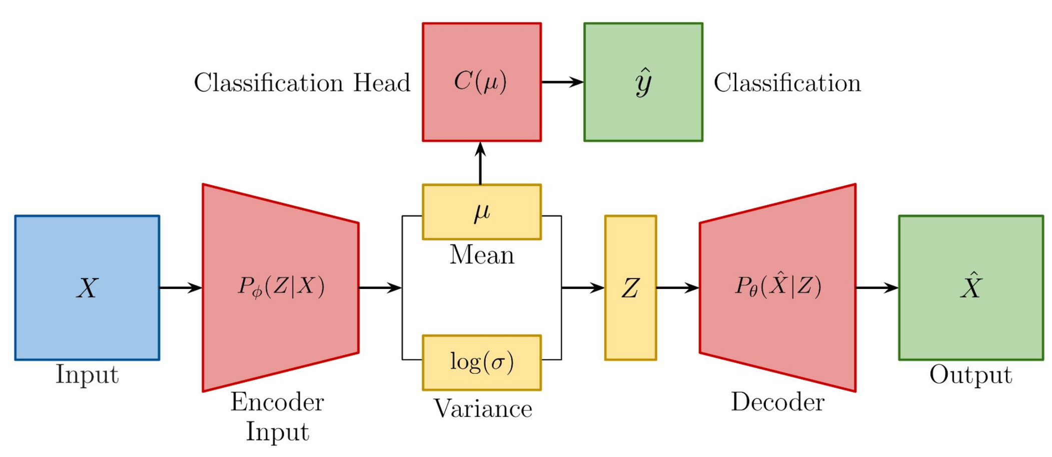

Figure 1. Visual representation of model architecture.
Graphic created by student researcher in Google Drawings, 2025.

Methods

Training & Evaluation

1. Processed the first CHIME FRB catalog for neural network usage [2].
2. Trained the sVAE using the **Adam optimizer** in **PyTorch** [9].
3. Performed **Bayesian optimization** via **Optuna** [1] to explore hyperparameters for the most accurate model.
4. Generated a **classification report** to observe how the optimized model performs on separating repeaters/non-repeaters

Dimensionality Reduction Techniques

1. Passed **data through the encoder** to generate **latent vectors**
   a. **Principal Component Analysis** (PCA): For **linear projection** and initial visualization of latent space [10]
   b. **t-SNE**: For **capturing non-linear structures** and visualizing clusters in two dimensions [10].

Latent Space Partial Dependence Analysis

1. Measured how **latent feature affects repeater probability** [8]
2. Chose the two most correlated data features with each latent feature
3. Measured how latent space feature **affects correlated physical feature value**
4. Found **high-impact latent features** and plot **how physical features correlate with repeater probability** in latent space

Results

Metric	Value
Accuracy	0.9807
Precision	0.9805
Recall	0.9807
F_2-Score	**0.9807**

Table 1. Classification performance of final model.
Table created by student researcher in Google Slides, 2026.

FRB	References?
FRB20181218C	Yes [7]
FRB20190122C	No
FRB20190221A	**Yes [4, 5, 7]**
FRB20190320A	No

Table 2. Model false positives.
Table created by student researcher in Google Slides, 2026.

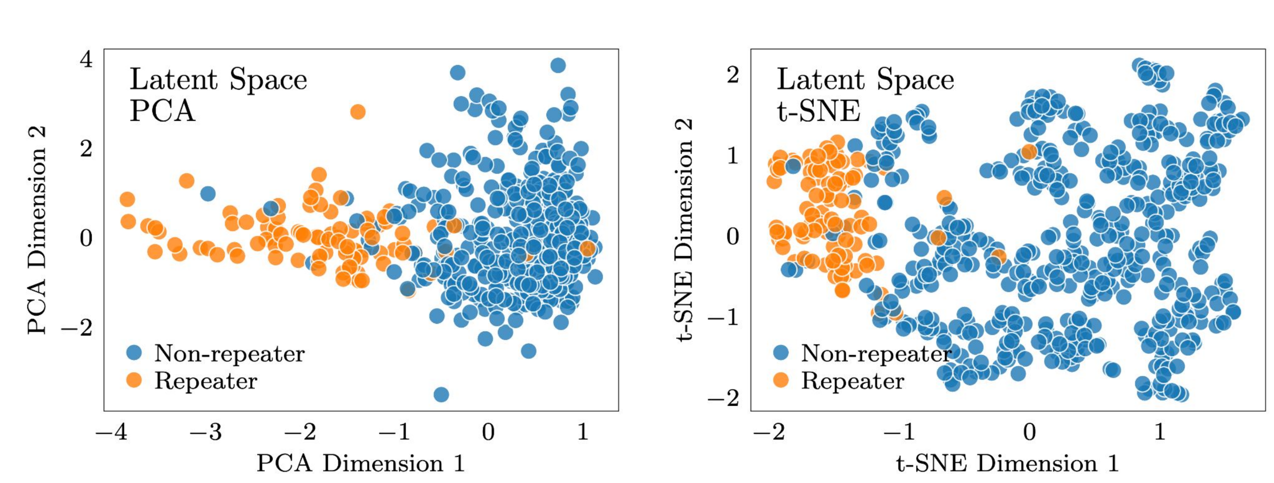

Figure 2. PCA and t-SNE of latent space representations, colored by repeater class.

Figure created by student researcher in Matplotlib, 2025.

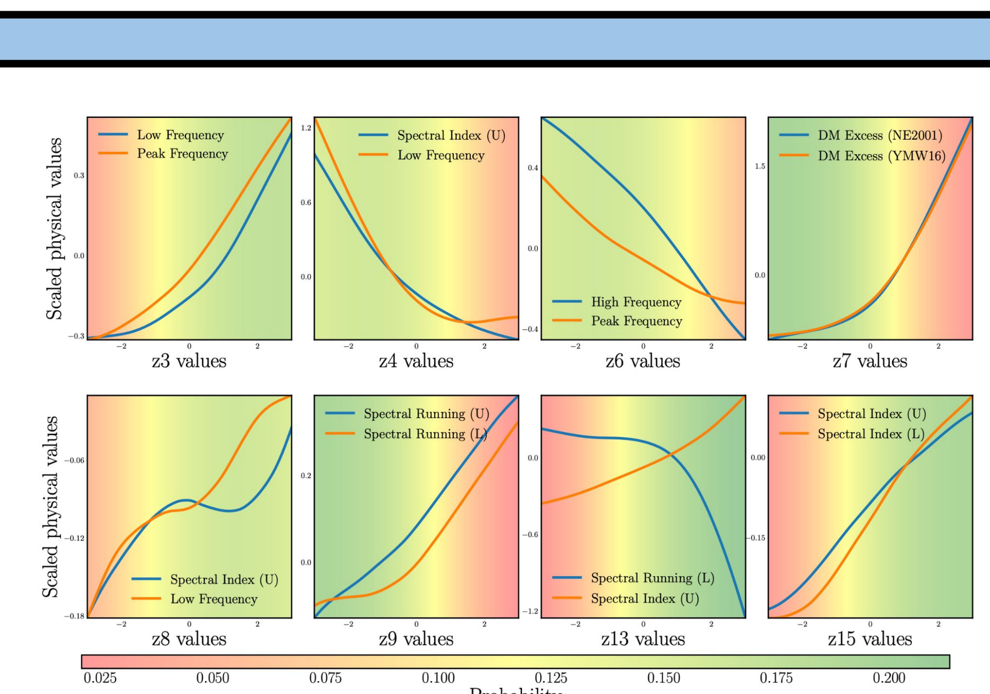

Figure 3. Latent space traversal for the most impactful features. Physical feature evolution is plotted, along with repeater probability as a color gradient.

Figure created by student researcher in Matplotlib, 2025.

Discussion

- Model achieves **state-of-the-art performance on classification benchmark** (0.9807 F2 score, **highest achieved**)
- The clear separation between populations in the sVAE latent space suggests the **distinct progenitor models** for repeaters and non-repeaters.
- Further validated by the fact that repeater likelihood is associated with **higher spectral indices** and **lower spectral running**.
- This supports the idea that repeating FRBs originate from **coherent magnetospheric processes** [12].
- This points toward **repeaters having high-energy environments** capable of sustained emission**, such as young, hyperactive magnetars**.
- Non-repeaters appear to arise from **cataclysmic, one-off events** like compact object mergers [12].
- The model identified **four false positives** (non-repeaters predicted as repeaters) for further follow up.
- Two false positives **have been previously identified**, especially FRB20190221A (marked by **three other papers** previously).
- The sVAE latent space is shown to be **physically interpretable** by using partial dependence analysis.

References

[1] Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.1907.10902 [2] Amiri, M., Andersen, B. C., Bandura, K., Berger, S., Bhardwaj, M., Boyce, M. M., ... Eadie, G. (2021). The First CHIME/FRB Fast Radio Burst Catalog. The Astrophysical Journal Supplement Series, 257(2), 59. https://doi.org/10.3847/1538-4365/ac33ab [3] Ji, T., Sri Theja Vuppala, Chowdhary, G., & Driggs-Campbell, K. (2020). Multi-Modal Anomaly Detection for Unstructured and Uncertain Environments. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.2012.08637 [4] Jia-Ming Zhu-Ge, Luo, J.-W., & Zhang, B. (2022). Machine learning classification of CHIME fast radio bursts – II. Unsupervised methods. Monthly Notices of the Royal Astronomical Society, 519(2), 1823–1836. https://doi.org/10.1093/mnras/stac3599 [5] Júnior, A. J. B., Fortunato, J. A. S., Silvestre, L. J., Alencar, T. V., & Hipólito-Ricaldi, W. S. (2026). Comparative analysis of machine learning techniques for feature selection and classification of Fast Radio Bursts. Journal of High Energy Astrophysics, 49, 100449. https://doi.org/10.1016/j.jheap.2025.100449 [6] Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.1312.6114 [7] Luo, J.-W., Zhu-Ge, J.-M., & Zhang, B. (2022). Machine learning classification of fast radio bursts: I. supervised methods. Monthly Notices of the Royal Astronomical Society, 518(2), 1629–1641. https://doi.org/10.1093/mnras/stac3206 [8] Molnar, C. (2025). Interpretable Machine Learning. Christoph Molnar. [9] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J. T., Chanan, G., ... Bai, J. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.1912.01703 [10] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, É. (2018, June 5). Scikit-learn: Machine Learning in Python. https://doi.org/10.48550/arXiv.1201.0490 [11] Petroff, E., Jason, & Lorimer, D. R. (2022). Fast radio bursts at the dawn of the 2020s. The Astronomy and Astrophysics Review, 30(1). https://doi.org/10.1007/s00159-022-00139-w [12] Sun, W.-P., Zhang, J.-G., Li, Y., Hou, W.-T., Zhang, F.-W., Zhang, J.-F., & Zhang, X. (2025). Exploring the Key Features of Repeating Fast Radio Bursts with Machine Learning. The Astrophysical Journal, 980(2), 185–185. https://doi.org/10.3847/1538-4357/adad6a