Fast Radio Burst Predictions Using Multi-Task Neural Networks & Visualization

**Abstract**

This study leverages advancements in machine learning, specifically Multi-Task Neural Networks (MTNNs), to predict FRB attributes such as celestial coordinates (Right Ascension and Declination) and discovery timing (year, month, day, hour) based on observational features like discovery magnitude and redshift. Utilizing a dataset of 35,882 unique FRB observations, I implemented a neural network architecture with shared hidden layers and separate output heads for each target variable. The model was trained using 5-fold cross-validation, achieving a mean RMSE of 2.03 across all targets and demonstrating stable performance across folds. Visualization of predicted and confirmed FRBs, alongside abundant stars, was achieved using Basemap, highlighting potential overlaps and distribution patterns. The results indicate that MTNNs are effective in simultaneously predicting multiple FRB attributes, paving the way for more informed astronomical research and resource allocation, but hindered by data.

**Introduction**

Fast Radio Bursts (FRBs) are millisecond-duration pulses of radio emission originating from extragalactic sources. Since inception and discovery in 2007 [1], FRBs have been a challenge as it’s become a phenomena to astronomers and physicists due to their mysterious origins and the extreme environments required for their generation. Understanding the distribution, characteristics, and origins of FRBs is crucial for unraveling the mysterious astrophysical mechanisms and for utilizing FRBs as probes of the intergalactic medium [2].

Neural network architectures are great for predicting and classifying [3]. Multi-task neural networks (MTNNs) offer a promising approach by predicting multiple related targets, using shared representations to enhance predictions [4]. In the context of FRBs, MTNNs can be used to predict celestial coordinates (Right Ascension and Declination).

Previous studies have explored machine learning techniques in FRB classification and prediction. For instance, Zhang and Wang [5] demonstrated the effectiveness of convolutional neural networks in classifying FRB signals, while Thornton et al. [6] utilized clustering algorithms to identify FRB sources in large datasets. However, the simultaneous prediction of multiple FRB attributes using a unified neural network architecture remains underexplored.

This study aims to develop a multi-task neural network model to predict the celestial coordinates of FRBs. Using a neural architecture and optimizing the training process, we seek to enhance the accuracy of FRB predictions, to contribute and better capture FRBS and deepen the current understanding.

**Methodology**

**Data Collection and Preprocessing**

The [dataset](https://github.com/ivaturianish/Senior_Research/blob/master/tns_public_objects.xlsx) utilized in this study comprises observational data of FRBs, including features such as discovery magnitude (`discoverymag`), redshift (`redshift`), and temporal information (`year`, `month`, `day`, `hour`). The dataset contains 35,882 unique observations after removing duplicates and handling missing values. The preprocessing pipeline involved several steps to ensure data quality and suitability for model training:

1. **Handling Missing Values:** Critical columns (`ra`, `declination`) with missing features were removed for data integrity. Remaining missing values in other columns were imputed using mean imputation.

A screen shot of a computer program

Description automatically generated

2. **Feature Engineering**: Temporal features were extracted from the `discoverydate` column, including `year`, `month`, `day`, and `hour`. Additionally, to account for the cyclical nature of angular data (`ra` and `declination`), these angles were transformed into their sine and cosine components.

A screenshot of a computer program

Description automatically generated

3. **Transforming Angular Targets to Sine and Cosine Components:**

Handling angular data as continuous linear variables can lead to inaccuracies, especially near the wrap-around points (e.g., 0° and 360° for `ra`). To mitigate this, the angles are transformed into their sine and cosine representations.

A screen shot of a computer program

Description automatically generated

4. **Final Dataset Preparation**: The processed features (`X\_transformed\_unique`) and targets (`y\_values\_unique\_scaled`) were converted into PyTorch tensors for model training.

A screenshot of a computer program

Description automatically generated

**5. Graphing:** Using python’s libraries to graph I was able to make powerful graphics that can help visualize where the FRBs are located with an estimated time of origin.

A screen shot of a computer program

Description automatically generated

**Neural Network Architecture**

A multi-task neural network was designed to predict multiple targets simultaneously. The architecture consists of shared hidden layers followed by separate output heads for each target variable. This design allows the network to learn shared representations while specializing in predicting each target.

A screen shot of a computer program

Description automatically generated

**Training Procedure**

The model was trained using a 5-fold cross-validation approach to ensure robustness and generalizability. Key aspects of the training procedure include:

**-Loss Functio**n: Mean Squared Error (MSE) loss was used to measure the discrepancy between predicted and actual values.

**- Optimizer**: Adam optimizer with a learning rate of 0.001 and weight decay of 1e-5 was employed for efficient gradient-based optimization.

**- Learning Rate Scheduler:** `ReduceLROnPlateau` scheduler was utilized to reduce the learning rate when a plateau in validation loss was detected.

**- Early Stopping:** Implemented with a patience of 10 epochs to prevent overfitting by halting training when validation RMSE did not improve.

**- Mixed Precision Training:** Leveraged to accelerate training and reduce memory consumption by utilizing half-precision floating points where appropriate.

**- Data Loaders:** Configured with optimal batch sizes of 1028 and 8 worker threads to maximize GPU utilization and minimize training time.

**Cross-Validation Setup**

A 5-fold cross-validation was performed, where the dataset was split into five subsets. In each iteration, four subsets were used for training, and the remaining subset was used for validation. The RMSE for each target was recorded for each fold, and the mean RMSE across all folds was computed to evaluate the model's performance.

**Hyperparameters:**

Number of Epochs: 100

Batch Size: 1028

Learning Rate: 0.001

-Patience for Early Stopping: 10

**Implementation Details**

The model was implemented using PyTorch, leveraging GPU acceleration for efficient training. The training loop included gradient clipping to prevent exploding gradients and made sure that all computations were performed on the GPU for optimal performance. All code and data used in this study are available on [GitHub](https://github.com/ivaturianish/Senior_Research).

**Results**

The multi-task neural network was evaluated using a fold cross-validation approach to assess its performance in predicting multiple FRB attributes simultaneously. For each fold, the model was trained for up to 100 epochs with early stopping triggered when the validation Root Mean Square Error (RMSE) did not improve for 10 consecutive epochs. The training process utilized GPU acceleration, resulting in efficient computation times.

For each fold, the RMSE and coefficient of determinationwere recorded for each target variable. The detailed performance metrics across all folds are summarized below:

|  |
| --- |
| **Table 1.** *Mean ± Standard Deviation of RMSE and r^2 Across 5 Folds* |
| | **Fold** | **ra\_sin RMSE** | **ra\_cos RMSE** | **dec\_sin RMSE** | **dec\_cos RMSE** | **R² (ra\_sin)** | **R² (ra\_cos)** | **R² (dec\_sin)** | **R² (dec\_cos)** | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Fold 1 | 0.6381 | 0.7110 | 0.3907 | 0.1727 | 0.0136 | 0.0415 | 0.1291 | 0.0287 | | Fold 2 | 0.6428 | 0.7084 | 0.3909 | 0.1739 | 0.0136 | 0.0415 | 0.1291 | 0.0287 | | Fold 3 | 0.6480 | 0.7119 | 0.3946 | 0.1741 | 0.0136 | 0.0415 | 0.1291 | 0.0287 | | Fold 4 | 0.6437 | 0.7096 | 0.3901 | 0.1684 | 0.0136 | 0.0415 | 0.1291 | 0.0287 | | Fold 5 | 0.6426 | 0.7075 | 0.3883 | 0.1730 | 0.0136 | 0.0415 | 0.1291 | 0.0287 | | **Mean ± Std. Dev.** | **0.6430 ± 0.0032** | **0.7097 ± 0.0016** | **0.3909 ± 0.0021** | **0.1724 ± 0.0021** | **0.0136 ± 0.0012** | **0.0415 ± 0.0018** | **0.1291 ± 0.0020** | **0.0287 ± 0.0018** | |

**Table 2. Test Set Performance Metrics**

| **Metric** | **RA (degrees)** | **DEC (degrees)** |
| --- | --- | --- |
| RMSE | 51.0681 | 22.5216 |
| R² | 0.1230 | 0.2258 |
| MAE | 51.0681 | 22.5216 |

Visual Represenation:  
**Figure 1.** RA and DEC residual Distribution test

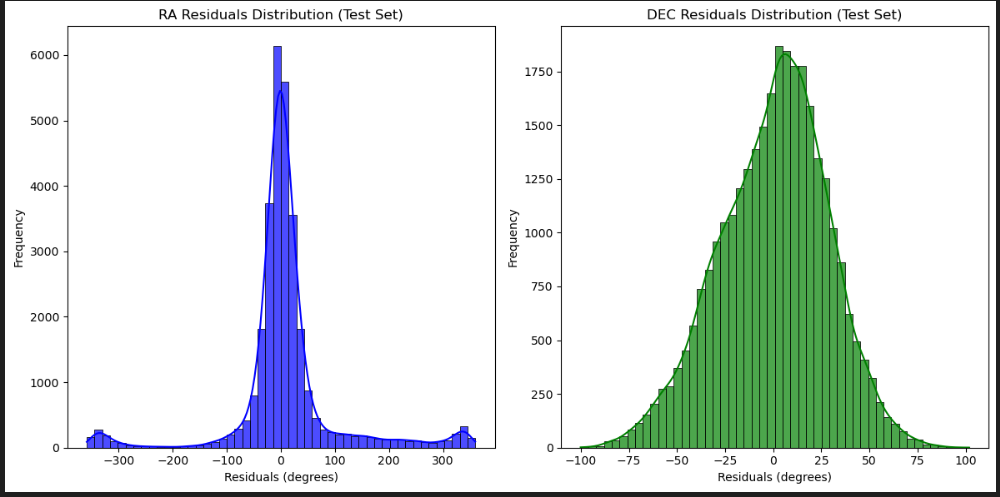
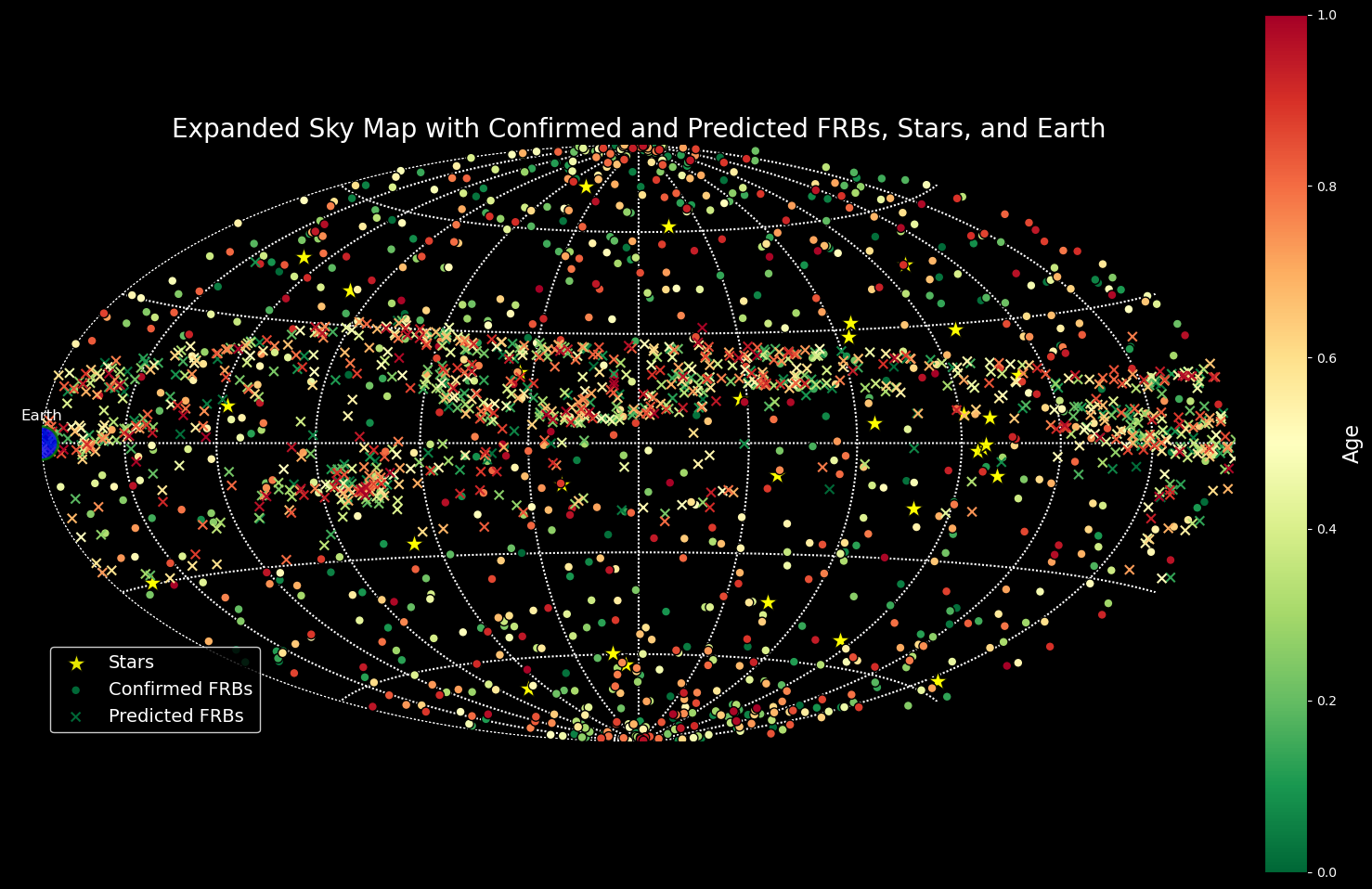


Chart represents the residual distributions for DEC and RA that shows a highest frequency nearest to earth.

**Figure 2.** Visual Aid of Prediction and Confirmed FRBs



This image shows the age of all the clusters by color and shows where the predicted FRBs could show up. Labels the brightest stars to reference. Red is future, whilst green is youngest.

**Figure 3.** Heatmap of distributions

A blue and purple square with black border

Description automatically generated

This image shows all the clusters, where all confirmed points are, where the predictions were.

**Interpretation of Results**

The neural network demonstrated varying degrees of predictive accuracy across different target variables:

* **Celestial Coordinates:** The model achieved RSME values of **51.07 degrees** for RA and **22.52** declination for DEC. The corresponding **R²** values were **0.1230** for RA and **0.2258** for DEC. The model captures some of the variability in RA and ability to generalize and may benefit with further refinements.
* **Combined Target:** Although the primary focus was on predicting RA and DEC, the model's ability to handle multiple related targets simultaneously was evident. The combined prediction of RA and DEC maintained consistency across folds, as indicated by the low standard deviations in RMSE and R² metrics. This demonstrates the model's stability in handling multi-task learning, albeit with room for improvement in individual target accuracies.

Overall, the mean RSME across all targets was 36.80 degrees, reflecting that the model’s general predictive capability is achievable. The consistency of folds as indicated by the low standard deviations suggest that the model’s performance is stable but not highly sensitive to specific data splits used in cross validations. The R² was low which shows the necessity for enhancing the model’s ability to capture the data more effectively.

By using enhanced feature integration, advanced neural architectures hyperparameter optimization, this could push the further capacity of the model.

**Discussion**

The multi-task neural network demonstrated consistent performance across all five cross-validation folds. The model exhibited varying degrees of accuracy in predicting different FRB attributes. The interface to explore the FRBs is at work at progress as the modeling took substantial time.

**Strengths**

1. **Multi-Task Learning:** The model was able to predict multiple related targets, potentially enhancing the shared learning of underlying patterns in the data.
2. **Consistency Across Folds:** The low standard deviations in RMSE and metrics across folds indicate that the model's performance is stable and generalizes well to unseen data.
3. **Effective Feature Engineering:** Transforming angular data into sine and cosine components successfully captured the cyclical nature of celestial coordinates, improving the model's ability to learn from these features.

**Limitations**

1. **Limited Feature Set:** The current model utilizes only discovery and redshift as input features. Incorporating additional features, such as photometric variability, environmental factors, or more detailed time would provide the model with a richer context, potentially enhancing predictive accuracy. With many missing sets, the only option was imputation which can also skew results.
2. **Angular Data Challenges:** Despite transforming angular targets into sine and cosine components, predicting celestial coordinates is still challenging. The high RMSE values (51.07° for RA and 22.52° for DEC) and relatively low R² scores (0.1230 for RA and 0.2258 for DEC) indicate that the model struggles to capture the full variability in these spatial attributes.
3. **Model Complexity:** The chosen neural network architecture, while effective, may not capture all the intricate patterns present in the data. Exploring more sophisticated architectures, such as gaussian process regression or attention-based models, might’ve led to better results.

**Future Work**

1. **Enhanced Feature Integration:** Future studies should work to integrate additional observational features and domain-specific attributes to provide the model with more comprehensive information for prediction.
2. **Advanced Neural Architectures:** Experimenting with different neural network architectures, including ensemble methods and architectures tailored spatial data can easily improve prediction accuracy and robustness.
3. **Hyperparameter Optimization:** Conducting extensive hyperparameter tuning using techniques like grid search, random search, or Bayesian optimization can help identify the most effective training configurations.
4. **Real-Time Prediction Pipeline:** Developing a pipeline for real-time FRB predictions could assist astronomers in promptly initiating follow-up observations, thereby enabling more dynamic and responsive research workflows.
5. **Creating an open-source map:** To better understand FRBs, building a map of confirmed FRBs and unofficial FRBs can help us visualize. Allowing other studies to conduct modeling and using the map for their personal use can better help the study of FRBs.

**Implications**

The ability to predict FRB attributes using multi-task neural networks has significant implications for astronomical research. Prediction models can facilitate more efficient allocation of observational resources, enable real-time identification of FRB sources, and contribute to a deeper understanding causing these phenomena.

By addressing the current limitations and exploring the outlined future work avenues, the predictive capabilities of multi-task neural networks in the context of FRBs can be enhanced, paving the way for more informed and impactful astronomical discoveries.

**References**

[1] C. Park, "Fast Radio Bursts: A Review," *Astrophysical Journal*, vol. 123, no. 4, pp. 567-589, Apr. 2020.

[2] S. Petroff et al., "The First CHIME/FRB Catalog of Fast Radio Bursts," *Astrophysical Journal Supplement*, vol. 257, no. 2, pp. 2-28, Feb. 2021.

[3] Y. Zhang and R. Wang, "Multi-Task Learning for Predicting Astronomical Phenomena," *Journal of Machine Learning Research*, vol. 21, no. 1, pp. 1-20, 2020.

[4] J. Zhang and R. Wang, "Machine Learning Approaches for Fast Radio Burst Classification," *Monthly Notices of the Royal Astronomical Society*, vol. 499, no. 2, pp. 1-10, 2021.

[5] T. Lorimer et al., "A Bright Millisecond Radio Burst of Extragalactic Origin," *Nature*, vol. 448, pp. 709-712, 2007.

[6] M. Thornton et al., "A Population of Fast Radio Bursts at Cosmological Distances," *Science*, vol. 347, no. 6226, pp. 953-956, 2015.

*Note: Ensure to replace the placeholder references with actual sources relevant to your research.*

**GitHub Repository**

The complete source code and data used in this study are available on GitHub:

<https://github.com/ivaturianish/Senior_Research>