PROCEEDINGS OF SCIENCE

Investigating machine learning approaches for differentiating MGFs and SGRBs

D.J. Maheso[a,*] **and S. Razzaque**[b]

[a] *Centre for Space Research (CSR), North-West University Potchefstroom, South Africa*

[b] *Centre for Astro-Particle Physics (CAPP) and Department of Physics, University of Johannesburg, Aukland Park, South Africa*

E-mail: d.j.maheso@gmail.com, srazzaque@uj.ac.za

Magnetar giant flares (MGFs) and short gamma-ray bursts (SGRBs) are short gamma-ray transients (SGRTs) with overlapping temporal and spectral characteristics, making them challenging to distinguish, especially when their redshift is unknown. In this study, we apply supervised machine learning using a Support Vector Machine to classify MGFs and SGRBs. Temporal parameters (including pulse rise times from Norris function fits) and spectral features (derived from Comptonized model fits over the 10 keV–40 MeV range) are extracted as input features for classification. We analyse 15 MGF and 101 SGRB samples from 10 *Fermi*-GBM sources. Classifier performance is assessed using leave-one-source-out cross-validation. The model successfully classifies most SGRBs, but struggles with MGFs due to limited training data and significant feature overlap. This work highlights the potential and challenges of incorporating machine learning into the automated classification of these SGRTs.

High Energy Astrophysics in Southern Africa (HEASA2025)
16-20 September, 2025
University of Johannesburg, South Africa

*Speaker

 ISSN 1824-8039 . Published by SISSA Medialab.

PoS(HEASA2025)043

1. Introduction

Short gamma-ray bursts (SGRBs) are one of the most luminous events in the Universe. They are characterised by brief durations ($T_{90} < 2$ s), high isotropic energies ($E_{\rm iso} \sim 10^{51}$ erg), and cosmological origins [1], making them ideal cosmological probes. Their association with compact binary mergers—particularly neutron star (NS) mergers, confirmed by the detection of gravitational wave event GW170817 and SGRB 170817A [2] makes them important sources for multi-messenger studies.

Magnetar giant flares (MGFs), less energetic transients originating from highly magnetised NSs in nearby galaxies [3], share several observational characteristics with SGRBs despite different progenitors and emission mechanisms. MGFs produce intense, short-lived γ-ray emission followed by fading oscillatory pulses. However, MGFs can mimic the temporal and spectral signatures of SGRBs when observed at large distances, where oscillatory pulses are below detection sensitivity, leading to possible misclassification.

The observational similarities between the two classes pose challenges for population and cosmological studies. Misidentifying local MGFs as cosmological SGRBs can bias estimates of burst rates and binary merger statistics. As a result, developing reliable methods to distinguish the two transients is essential for accurate source population analyses and cosmological studies.

In this work, we explore the use of supervised machine learning (ML) algorithms to distinguish between these SGRTs based on their temporal and spectral characteristics. ML can reveal subtle correlations often missed by traditional methods. Specifically, we extract pulse parameters derived from Norris function fits and spectral parameters obtained from Comptonized model fits (see section 2). The classification framework and Support Vector Machine (SVM) implementation are detailed in section 3. By clustering the transients through ML approaches, we aim to improve source classification and better understand these SGRTs. The results are therefore discussed in section 4, followed by conclusions in section 5.

2. Methodology

2.1 Data selection

The transient sample was obtained from the *Fermi* Gamma-ray Burst Monitor (*Fermi*-GBM). The selected sources exhibit well-defined peaks and sufficient signal-to-noise for reliable temporal and spectral fitting. The sample includes eight confirmed SGRBs and two MGFs misclassified as SGRBs upon detection (see Table 1). The two MGFs detected by *Fermi*-GBM to date, designated with GRB naming conventions are GRB200415A [4] and GRB231115A [5]. Temporal and spectral features were extracted from background-subtracted light curves and time-resolved spectra using *XSPEC 12.15.0*, utilising data from the brightest sodium iodide (NaI) and bismuth germanate (BGO) detectors.

The NaI (10 keV–1 MeV) and BGO (200 keV–40 MeV) detectors were divided into six and four energy channels respectively, resulting in ten energy channels while ensuring adequate photon statistics (equivalent to $\geq 20 - 25$ photons per channel). Only the brightest detectors were selected

to maximise the signal-to-noise ratio and ensure accurate spectral fitting. For each source, each energy channel was treated as an individual sample, and for multi-peaked bursts each peak was analysed across all channels. This produced 15 MGF and 101 SGRB samples. To prevent data leakage, Leave-One-Out-Cross-Validation (LOOCV) was utilised (see section 3).

GRB	T_{90} (s)	z	Detector(s)	Samples
090510	0.960 ± 0.138	0.903	n6 + b1	29
100206A	0.176 ± 0.072	0.41	n5 + b0	7
100816A	2.045 ± 0.229	0.8049	n7	12
111117A	0.432±0.082	2.211	n6	8
131004A	1.152 ± 0.590	0.717	na	2
200415A	**0.144±0.036**	**0.02**	**n3 + b0**	**9**
200826A	1.136 ± 0.332	0.7481	n7 + b1	22
201221D	0.144 ± 0.066	1.046	n7	5
210323A	0.960 ± 0.781	0.42	n5 + b0	16
231115A	**0.032 ± 0.036**	–	**n7 + b1**	**6**

Table 1: SGRB sample including the MGFs (in bold) used in temporal and spectral analysis. The 'Samples' column indicates the number of data points extracted per source across all energy channels and peaks.

2.2 Feature extraction

Two feature categories were used for classification: temporal and spectral. Temporal features were obtained by modelling the transients profiles with the Norris function [6]. The function describes the pulse asymmetry in GRB light curves and is defined in Equation 1. Here, A is the pulse count rate, t_{peak} is the time of maximum intensity, t_{rise} and t_{fall} are the rise and decay times. The parameters ν_1 and ν_2 control the pulse asymmetry. This model describes the fast-rise–exponential-decay behaviour of SGRTs. Figure 1a shows an example of a pulse fit using the function.

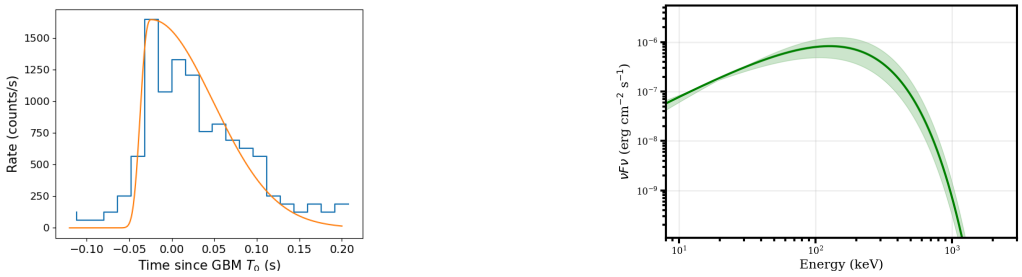
$$I(t) = \begin{cases} A \exp\left[-\left(\frac{|t-t_{peak}|}{t_{rise}}\right)^{\nu_1}\right]; & t < t_{peak} \\ A \exp\left[-\left(\frac{|t-t_{peak}|}{t_{fall}}\right)^{\nu_2}\right]; & t > t_{peak} \end{cases} \tag{1}$$

Time-resolved spectra were fit using the Comptonised (comp) spectral model (Equation 2), which is commonly used to fit SGRB spectra [7]. The comp model (Figure 1b) characterises the emission through the low-energy photon index, α, and the spectral peak energy, E_{peak}. Energy flux, F, and photon flux, P, were obtained by integrating the spectral model over the observed energy range (Equation 3). Fluence, S, and isotropic energy, E_{iso}, are calculated using the burst duration, T_{90}, luminosity distance, D_L, bolometric fluence, S_{bol}, and redshift, z (Equation 4). Although z correlates with E_{iso}, it was excluded to enable classification of transients with unknown z, which is common for newly detected events where spectroscopy follow-up is unavailable.

$$N(E) = \left(\frac{E}{100\,\mathrm{keV}}\right)^{\alpha} \exp\left(-\frac{E(\alpha+2)}{E_{peak}}\right) \tag{2}$$

$$F = \int_{E_{\min}}^{E_{\max}} E\,N(E)\,dE, \quad P = \int_{E_{\min}}^{E_{\max}} N(E)\,dE. \tag{3}$$

$$S = F\,T_{90}, \quad E_{\mathrm{iso}} = 4\pi D_L^2\,S_{\mathrm{bol}}\,(1+z). \tag{4}$$



(a) Norris function fit to the main pulse.

(b) Comptonised spectral model fit.

Figure 1: Temporal and spectral analysis of GRB201221D. **(a)** Light-curve fit highlights the pulse morphology. **(b)** The Comptonised model describes the observed νF_ν spectrum.

In summary, three temporal and six spectral features were fed into the classification algorithm, resulting in nine-dimensional (9D) feature space for each sample. The combined set of temporal and spectral features provides comprehensive characterisation. The features used are: t_{rise}, t_{fall}, A, α, E_{peak}, F, P, S, and E_{iso}.

3. Support Vector Machine

The Support Vector Machine (SVM) is a supervised learning algorithm that creates an optimal decision boundary (hyperplane) to separate classes in high-dimensional feature spaces. SVMs are well-suited for small to medium-sized datasets with limited samples, making them appropriate for this study. Here, the SVM was used to distinguish between the SGRT classes. All features were scaled to prevent the algorithm from being biased towards large values such as E_{iso} during training. Moreover, not all detectors had sufficient photon counts (i.e $\geq 20-25$ photons per channel) across the full ten-energy channel range. In cases where data were missing due to insufficient photon counts, missing feature values were filled with the mean value of that feature computed from the same source. The SVM was implemented with a radial basis (RBF) kernel, and the hyperparameters (C, γ) were optimised via nested grid search cross-validation.

Misclassification is controlled by the regularisation hyperparameter, C, while the curvature of the decision boundary is determined by γ. The algorithm was evaluated using LOOCV, where all samples from one source were grouped and withheld for validation while the algorithm was trained on the remaining sources. This ensures the algorithm never sees validation data during training.

LOOCV addresses pseudo-replication, where correlated measurements are treated as independent observations and different time intervals from the same source are treated as independent samples. LOOCV reduces pseudo-replication by excluding all observations from the same source during each validation fold. Therefore, forcing the algorithm to learn general patterns rather than source-specific signatures. This tests whether the classifier can distinguish between the two SGRTs utilising only intrinsic physical properties rather than source dependent characteristics.

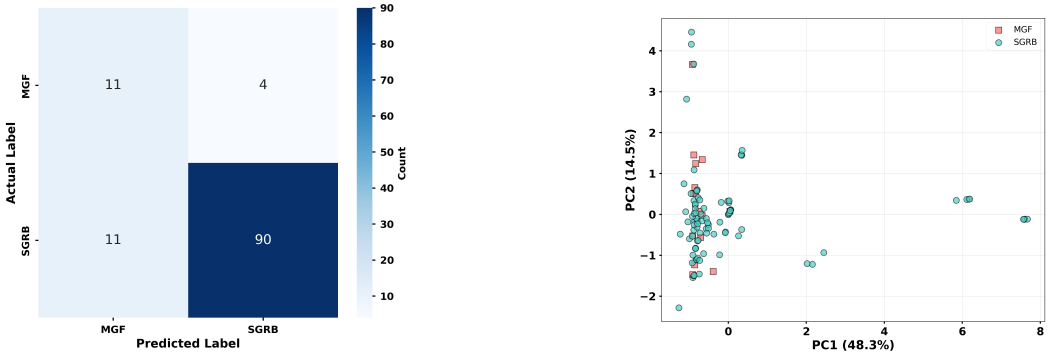
To address class imbalance, balanced class weights were applied. This approach assigns high importance (more weight) on the minority MGF class and lower weight to the majority SGRB

class. Increasing the penalty for misclassifying MGFs. However, with only two confirmed MGFs, MGF diversity is still insufficient for the weighting to fully counteract the imbalance. Once optimal hyperparameters were obtained, the final classification model was trained on all data. The model was then applied to classify candidate MGFs identified by Burns et al. [8].

4. Results and discussion

Following hyperparameter optimisation via grid search (see section 3). The confusion matrix in Figure 2a indicates that the model correctly classifies 90 out of 101 SGRBs while 11 out of 15 MGFs are misclassified due to the limited number of training data and significant feature overlap between the two SGRT classes.

To visualise the nine-dimensional (9D) feature space, we performed Principal Component Analysis (PCA), reducing the data to two-dimensions (2D) (Figure 2b). The first two principal components (PC) together explain 62.8% of the variance (PC1: 48.3%, PC2: 14.5%), which is a relatively moderate value of the variance. This shows the complexity of the feature space, with discriminative information spread across multiple dimensions, hence explaining the substantial overlap. The transients occupy overlapping regions particularly due to their spectral features (see Table 2). The spectral parameters: flux, fluence, E_{iso} and E_{peak} strongly correlate with the PCA components. However, α, moderately correlates with the PC2 component but in the negative direction (loading ≈ -0.6). This anti-correlation may reflect the differences in the hardness of the transients spectra. MGFs typically exhibit harder low-energy spectra (more negative α) due to rapid magnetic energy dissipation in the magnetar magnetosphere, while SGRBs show softer low-energy indices consistent with synchrotron emission from relativistic jets. Temporal features on the other hand, show negligible contribution to the principal components, leading us to conclude that spectral features best classify the transients.



(a) Confusion matrix of the training data.

(b) PCA 2D projection of the training data.

Figure 2: Comparison of MGF and SGRB features. **(a)** The confusion matrix shows the classification performance of the SVM using LOOCV. **(b)** The PCA projection illustrates the distribution of both classes in the reduced 2D feature space, where PC1 and PC2 depict the total variance as 48.3% and 14.5% respectively. The variance shows how much of the original 9D feature information is retained.

Candidate MGFs from Burns et al. [8] were classified using the trained model to evaluate its performance on unseen data. The predictions are summarised in Table 3, where all events were

classified as SGRBs with moderate confidence (62.5-65.7%) indicating the classifier's bias towards the majority SGRB data.

Principal Component	Feature	Loading	Contribution
PC1 (48.3%)	Flux (erg/cm²/s)	0.9992	25.61%
	E_{iso} (erg)	0.9947	25.38%
	Fluence (erg/cm²)	0.9685	24.06%
PC2 (14.5%)	E_{peak} (keV)	0.8146	56.57%
	α	-0.6784	39.23%
	t_{fall} (ms)	-0.2112	3.80%

Table 2: Top features contributing to PC1 and PC2. Loadings show feature correlations, and contributions indicate their relative importance.

MGF candidate	Prediction	Confidence	P(MGF)	P(SGRB)
GRB070201	SGRB	65.7%	34.3%	65.7%
GRB051103	SGRB	62.7%	37.3%	62.7%
GRB070222	SGRB	62.7%	37.3%	62.7%
GRB200415A	SGRB	62.8%	37.2%	62.8%
GRB041227	SGRB	62.4%	37.6%	62.4%
GRB980827	SGRB	62.8%	37.2%	62.8%
GRB790305B	SGRB	62.5%	37.5%	62.5%

Table 3: MGFs misclassified as SGRBs with predicted labels, confidence, and probabilities.

The training data are heavily SGRB-dominated. Despite applying balanced class weights to penalise MGF misclassifications more heavily, the SVM still struggles to generalise to the under-represented MGF class. The main contributing factor is that both SGRT pulses show similar temporal and spectral features during their initial hard spikes, resulting in strong feature overlap as shown in Figure 2b. Recent unsupervised clustering studies (e.g. Misra et al. 2024 [9]) identify similar difficulty separating MGFs from SGRBs due to overlapping spectral properties, consistent with our PCA findings. This physical similarity arises because both transients involve a sudden release of electromagnetic energy. MGFs typically display rapid rise times (few milliseconds [8]) and steeper decay slopes consistent with the rapidly dissipated magnetic energy, while SGRBs from compact binary mergers, tend to exhibit longer decay phases and slightly higher $E_{\rm iso}$ values due to the relativistic jet emission at $\approx 10^{12}$ cm from the central engine. The misclassification of MGFs as SGRBs therefore reflects both the intrinsic similarity of their emission properties and the scarcity of known MGFs available for training [10].

The use of LOOCV is challenging for MGFs because there are only two MGFs: GRB200415A and GRB231115A, holding one out leaves only a single event for training, limiting the classifiers ability to capture MGF source diversity. Although GRB200415A is a confirmed MGF and appears in both our training set and the Burns et al. candidate list [8]. When held out during LOOCV and later classified as an SGRB by the trained model (Table 3), this represents a successful prevention of data leakage despite the misclassification due to insufficient MGF diversity in the training data.

Despite these challenges, the SVM demonstrates promising capability to distinguish between the two SGRTs. The substantial feature overlap suggests that the underlying emission mechanisms are closely related, consistent with models proposing magnetar as central engines of GRBs [11]. This physical connection, in addition to the limited training data and class imbalance, explains the limited separation between MGFs and SGRBs in our results.

5. Conclusions and future work

In this study, we applied a Support Vector Machine (SVM) to distinguish between SGRBs and MGFs using temporal and spectral parameters derived from *Fermi*-GBM observations. The results show that although both transients share overlapping characteristics, particularly during their initial hard spikes, spectral features such as flux, fluence, $E_{\rm iso}$, and $E_{\rm peak}$ are the most discriminative, as

demonstrated by the PCA results (Table 2) where these features show the strongest loadings on the PC. This overlap reflects their intrinsic physical similarity and a possible connection in their emission mechanisms. However, MGF scarcity and the resulting class imbalance bias the model toward SGRBs, indicating that more diverse MGF samples are needed as weighting alone cannot fully compensate for feature overlap. We demonstrate the potential of SVM for identifying patterns between the two SGRTs. Future work should expand the MGF sample, incorporate additional mission data, and refine feature selection to improve model robustness and improve our understanding of the relationship between SGRBs and MGFs.

Acknowledgements

We acknowledge financial support from North-West University and thank the *Fermi* Science Support Center for developing and maintaining the analysis tools used in this study.

References

[1] E. Nakar, Short-hard gamma-ray bursts, *Physics Reports* **442** (1–6) (2007) 166–236.

[2] B. P. Abbott, R. Abbott, et al., Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A, *The Astrophysical Journal Letters* **848** (2) (2017) L13.

[3] K. Hurley, S. E. Boggs, et al., An exceptionally bright flare from SGR 1806–20 and the origins of short-duration γ-ray bursts, *Nature* **434** (7037) (2005) 1098–1103.

[4] D. Svinkin, D. Frederiks, et al., A bright γ-ray flare interpreted as a giant magnetar flare in NGC 253, *Nature* **589** (7841) (2021) 211–213.

[5] A. C. Trigg, R. Stewart, et al., Extragalactic magnetar giant flare GRB 231115A: Insights from Fermi/GBM observations, *Astronomy & Astrophysics* **694** (2025) A323.

[6] J. P. Norris, R. J. Nemiroff, et al., Attributes of pulses in long bright gamma-ray bursts, *The Astrophysical Journal* **459** (1996) 393.

[7] L. Nava, G. Ghirlanda, et al., Spectral properties of 438 GRBs detected by *Fermi*/GBM, *Astronomy & Astrophysics* **530** (2011) A21.

[8] E. Burns, D. Svinkin, et al., Identification of a local sample of gamma-ray bursts consistent with a magnetar giant flare origin, *The Astrophysical Journal Letters* **907** (2) (2021) L28.

[9] K. Misra, K. G. Arun, et al., Diversity in Fermi/GBM Gamma-ray bursts: New insights from machine learning, *The Astrophysical Journal* **974** (1) (2024) 55.

[10] S. Mereghetti, M. Rigoselli, et al., A magnetar giant flare in the nearby starburst galaxy M82, *Nature* **629** (8010) (2024) 58–61.

[11] A. Rowlinson, P. T. O'Brien, et al., Signatures of magnetar central engines in short GRB light curves, *Monthly Notices of the Royal Astronomical Society* **430** (2) (2013) 1061–1087.