POLAR background prediction

S. Nguyen, N. Produit

University of Geneva

stephane.nguyen@etu.unige.ch

nicolas.produit@unige.ch

November 2, 2023

Overview

Gamma Ray Bursts (GRBs)

- What are GRBs?

[1] https://www.astro.unige.ch/polar/grb-light-curves

Gamma Ray Bursts (GRBs)

- What are GRBs?
  - Intense, energetic, and extra-galactic explosions

[1] https://www.astro.unige.ch/polar/grb-light-curves

Gamma Ray Bursts (GRBs)

- What are GRBs?
  - Intense, energetic, and extra-galactic explosions
  - Bursts of high-energy photons, gamma rays

[1] https://www.astro.unige.ch/polar/grb-light-curves

Gamma Ray Bursts (GRBs)

- What are GRBs?
  - Intense, energetic, and extra-galactic explosions
  - Bursts of high-energy photons, gamma rays
- Lead to spikes in photon counts/rates

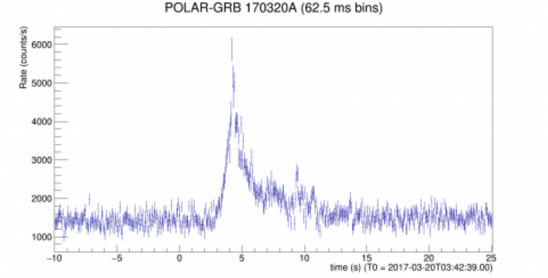


Figure: A light-curve containing a Gamma Ray Burst[1]

[1] https://www.astro.unige.ch/polar/grb-light-curves

What are astrophysicists interested in?

- Better understand GRBs

What are astrophysicists interested in?

- Better understand GRBs
- For this purpose, there's a need for:
  - Data taking $\rightarrow$ POLAR or POLAR-2 detector

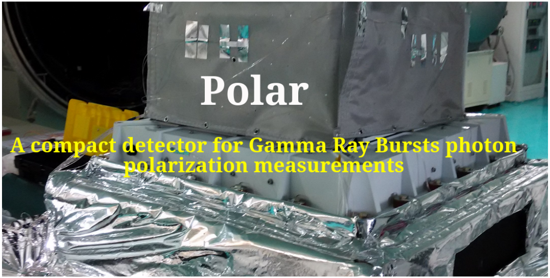


Figure: Picture taken from `https://www.astro.unige.ch/polar/`.

What are astrophysicists interested in?

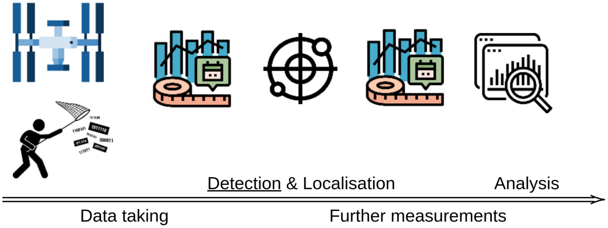


Figure: Process[2].

[2] Icons were taken from https://www.flaticon.com/free-icons/. More accurate attributions in the credits.

Related work

- Koziol's master thesis [1]: HAGRID[3]

[3] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[3]
  - Fast real-time GRB detection and localization using ML

[3] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[3]
  - Fast real-time GRB detection and localization using ML
  - In particular, for GRB detection, he used:

[3] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[3]
  - Fast real-time GRB detection and localization using ML
  - In particular, for GRB detection, he used:
    - Labeled simulated data with simulated GRBs

[3] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[3]
  - Fast real-time GRB detection and localization using ML
  - In particular, for GRB detection, he used:
    - Labeled simulated data with simulated GRBs
    - From past 60 [s] simulated light-curves + energy info.

[3] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[3]
  - Fast real-time GRB detection and localization using ML
  - In particular, for GRB detection, he used:
    - Labeled simulated data with simulated GRBs
    - From past 60 [s] simulated light-curves + energy info.
    - Predict the occurrence of a GRB at time t

[3] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[4]
  - Possible data distribution mismatch

[4] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[4]
  - Possible data distribution mismatch
  - But low miss-detection: correctly detected all real GRBs within 10 days of POLAR data.

[4] High Accuracy GRB Rapid Inference with Deep Learning in Space

Related work

- Koziol's master thesis [1]: HAGRID[4]
  - Possible data distribution mismatch
  - But low miss-detection: correctly detected all real GRBs within 10 days of POLAR data.
  - Does not account for other measurements

[4] High Accuracy GRB Rapid Inference with Deep Learning in Space

Our work

Would like to

- Train our model from real POLAR data

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels
  - $\rightarrow$ Background model

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels
  - $\rightarrow$ Background model
- Explicitly account for other measurements

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels
  - $\rightarrow$ Background model
- Explicitly account for other measurements
  - $\rightarrow$ Predict photon rates from diverse measurements

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels
  - $\rightarrow$ Background model
- Explicitly account for other measurements
  - $\rightarrow$ Predict photon rates from diverse measurements
- Reduce complexity barrier to model interpretability

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels
  - $\rightarrow$ Background model
- Explicitly account for other measurements
  - $\rightarrow$ Predict photon rates from diverse measurements
- Reduce complexity barrier to model interpretability
  - $\rightarrow$ Multi-Layer Perceptron (MLP)

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels
  - $\rightarrow$ Background model
- Explicitly account for other measurements
  - $\rightarrow$ Predict photon rates from diverse measurements
- Reduce complexity barrier to model interpretability
  - $\rightarrow$ Multi-Layer Perceptron (MLP)
- Try to interpet our trained model

Our work

Would like to

- Train our model from real POLAR data
  - No longer have access to labels
  - $\rightarrow$ Background model
- Explicitly account for other measurements
  - $\rightarrow$ Predict photon rates from diverse measurements
- Reduce complexity barrier to model interpretability
  - $\rightarrow$ Multi-Layer Perceptron (MLP)
- Try to interpet our trained model
  - $\rightarrow$ Gradients w.r.t input variables

Methodology

- Train background model **without** 25 known GRBs[5].

[5] 25 of 55 known GRBs [2] happening within our data

[6] Icons taken from various sources, especially from https://pypi.org/

Methodology

- Train background model **without** 25 known GRBs[5].
- Apply our model to everything

[5] 25 of 55 known GRBs [2] happening within our data

[6] Icons taken from various sources, especially from https://pypi.org/

Methodology

- Train background model **without** 25 known GRBs[5].
- Apply our model to everything
- Extract time intervals based on the residuals

[5] 25 of 55 known GRBs [2] happening within our data

[6] Icons taken from various sources, especially from https://pypi.org/

Methodology

- Train background model **without** 25 known GRBs[5].
- Apply our model to everything
- Extract time intervals based on the residuals
- Peek at model interpretability using Jacobian matrices

[5] 25 of 55 known GRBs [2] happening within our data

[6] Icons taken from various sources, especially from https://pypi.org/

Methodology

- Train background model **without** 25 known GRBs[5].
- Apply our model to everything
- Extract time intervals based on the residuals
- Peek at model interpretability using Jacobian matrices
- A few technologies[6]:

[5] 25 of 55 known GRBs [2] happening within our data

[6] Icons taken from various sources, especially from https://pypi.org/

Background model

- Multi-Layer Perceptron (MLP) background model[7]

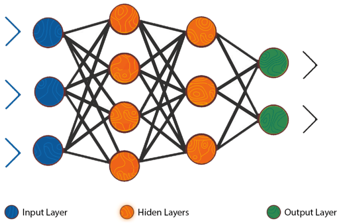


Figure: Multi-Layer Perceptron

- $f_\theta : \mathbb{R}^{18} \to \mathbb{R}^{6}$, 3×200 hidden units with ReLUs.

[7] Picture taken from https://medium.com/unpackai/from-anns-artificial-neural-networks-to-rnns-recurrent-neural-networks-93b638772fd1

Background model

- Multi-Layer Perceptron (MLP) background model[7]

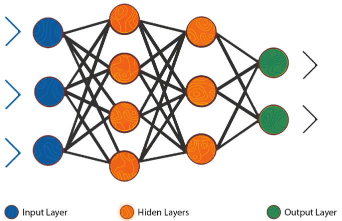


Figure: Multi-Layer Perceptron

- $f_\theta : \mathbb{R}^{18} \to \mathbb{R}^6$, 3×200 hidden units with ReLUs.
- Trained to minimize a weighted MSE

[7] Picture taken from https://medium.com/unpackai/from-anns-artificial-neural-networks-to-rnns-recurrent-neural-networks-93b638772fd1

Poor predictions

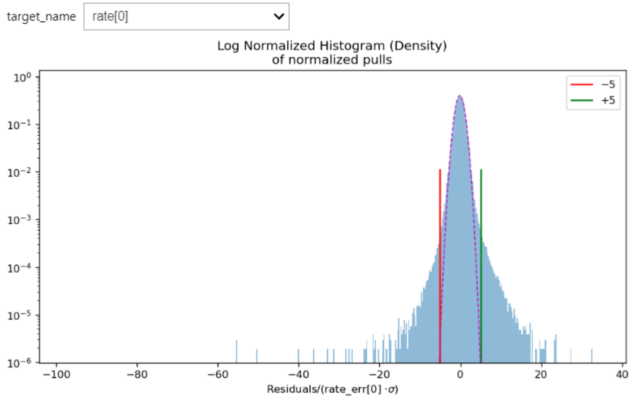


Figure: Normalized pulls for rate[0]

Cluster and cluster intersections

- Aggregating extracted data-points $\rightarrow$ clusters

Cluster and cluster intersections

- Aggregating extracted data-points $\rightarrow$ clusters
- GRBs should be visible in different energy bins

Cluster and cluster intersections

- Aggregating extracted data-points $\rightarrow$ clusters
- GRBs should be visible in different energy bins
- Cluster of data points extracted using different energy bins:
  - Each from the same set of energy bins
  - Or each satisfying same conditions (e.g. number of energy bins)

Results & discussion: Known GRBs

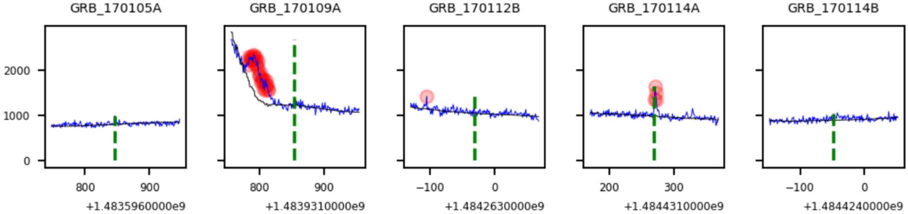


Figure: Energy bin 0, time windows of ± 100 [s] around 5 known GRB trigger times [2], based on pulls with $k = 5$. Predictions are in black.

Results & discussion: Known GRBs

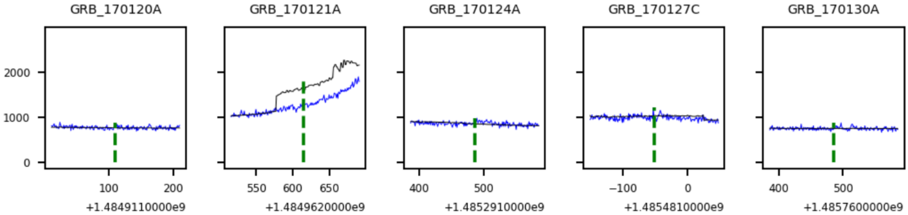
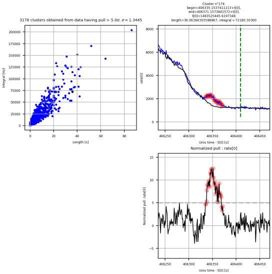
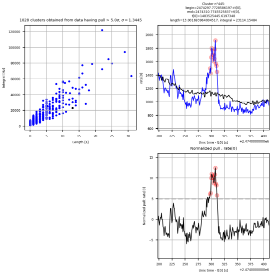


Figure: Energy bin 0, time windows of ± 100 [s] around 5 other known GRB trigger times [2], based on pulls with $k = 5$.

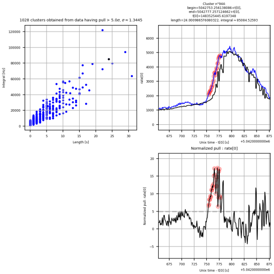
Results & discussion: Cluster intersections



(a) Cluster intersection with more the 3 energy bins; Solar flare confirmed by Prof. Nicolas Produit

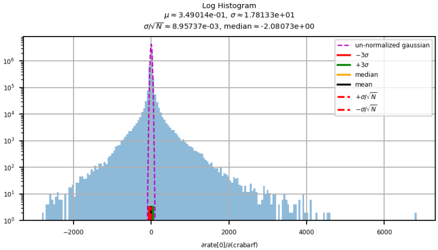


(b) Undetermined cluster intersection with all energy bins and with a 30-second discard window.

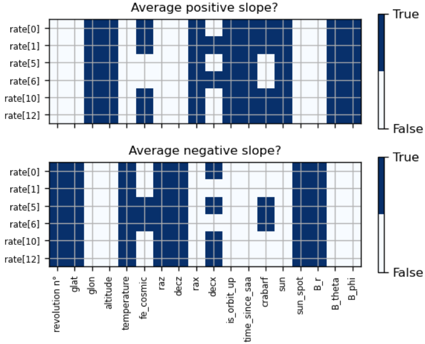


(c) Undetermined cluster intersection with all energy bins and with a 30-second discard window.

Results & discussion: Model interpretability

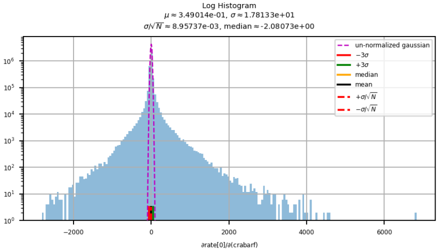


(a) Histogram of $\frac{\partial \text{rate}[0]}{\partial \text{crabarf}}$ (un-normalized features)

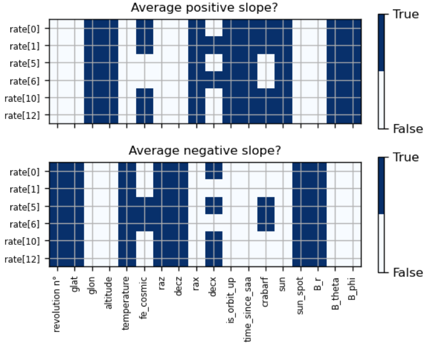


(b) Avg slope μ greater (lesser) than its (negative) standard error $\frac{\sigma}{\sqrt{N}}$?

Results & discussion: Model interpretability



(a) Histogram of $\frac{\partial \text{rate}[0]}{\partial \text{crabarf}}$ (un-normalized features)



(b) Avg slope μ greater (lesser) than its (negative) standard error $\frac{\sigma}{\sqrt{N}}$?

- Need for further analysis

Results & discussion

Table 6: Mean ± standard deviation (over 10 different seeds) number of positive/negative cluster intersections for different sets of energy bins bs or conditions[29]. This time, there's a discard window of 30 seconds.

bs or condition	negative	positive	bs or condition	negative	positive	bs or condition	negative	positive
# inter > 0	5762.1 ± 965.708	6487.6 ± 826.131	0,2	6.9 ± 2.726	10.7 ± 4.111	1,3	15.6 ± 6.398	51.2 ± 42.695
# inter > 1	4802.1 ± 822.34	5450.6 ± 752.954	0,2,3	6.0 ± 3.432	7.1 ± 1.37	1,3,4	1.0*	–
# inter > 2	3493.8 ± 610.247	4083.5 ± 564.85	0,2,3,4	8.0*	–	1,3,4,5	1.333 ± 0.577	1.5 ± 0.577
# inter > 3	3238.4 ± 602.769	3821.5 ± 538.957	0,2,3,4,5	212.2 ± 42.158	242.7 ± 65.109	1,3,5	–	5.5 ± 6.364
# inter > 4	1150.2 ± 232.379	1595.0 ± 202.499	0,2,3,5	1.0 ± 0.0	3.5 ± 3.536	1,4	1.0*	3.0*
# inter > 5	791.3 ± 165.14	1135.2 ± 149.199	0,2,4	1.5 ± 0.707	1.0*	1,4,5	2.125 ± 1.356	5.5 ± 3.742
0	23.8 ± 8.23	71.6 ± 18.368	0,2,4,5	5.8 ± 4.367	15.8 ± 5.095	1,5	–	2.0 ± 0.0
0,1	16.6 ± 6.059	62.3 ± 45.631	0,2,5	1.0 ± 0.0	2.5 ± 2.38	2	734.8 ± 169.59	939.8 ± 235.236
0,1,2	4.0 ± 2.0	15.2 ± 9.739	0,3	1.0*	3.0 ± 1.414	2,3	2459.9 ± 493.897	3265.5 ± 779.499
0,1,2,3	12.4 ± 3.806	30.5 ± 17.219	0,3,4,5	1.667 ± 1.155	1.75 ± 0.957	2,3,4	114.8 ± 335.669	11.75 ± 12.395
0,1,2,3,4	23.0*	2.5 ± 0.707	0,4	2.0*	–	2,3,4,5	3156.9 ± 657.884	4049.1 ± 676.033
0,1,2,3,4,5	791.3 ± 165.14	1135.2 ± 149.199	0,4,5	2.25 ± 1.753	11.6 ± 3.062	2,3,5	60.444 ± 11.479	225.0 ± 392.74
0,1,2,3,5	1.0 ± 0.0	3.667 ± 4.899	0,5	1.0*	1.75 ± 1.5	2,4	18.333 ± 38.563	2.333 ± 1.751
0,1,2,4	1.0*	3.0*	1	52.6 ± 15.16	104.0 ± 118.135	2,4,5	212.3 ± 59.913	244.9 ± 73.942
0,1,2,4,5	2.8 ± 1.751	11.2 ± 7.315	1,2	3.444 ± 1.424	8.5 ± 4.143	2,5	9.8 ± 6.356	55.5 ± 127.411
0,1,2,5	1.0*	2.667 ± 2.082	1,2,3	56.1 ± 19.221	115.1 ± 41.391	3	1206.3 ± 282.876	1633.0 ± 282.032
0,1,3	1.25 ± 0.5	4.875 ± 7.2	1,2,3,4	14.75 ± 27.5	1.4 ± 0.548	3,4	6.25 ± 9.215	1.0 ± 0.0
0,1,3,4,5	2.5 ± 2.811	2.25 ± 0.5	1,2,3,4,5	413.5 ± 141.535	714.4 ± 132.802	3,4,5	19.7 ± 8.166	16.5 ± 4.503
0,1,3,5	–	1.0*	1,2,3,5	2.167 ± 0.753	9.667 ± 19.268	3,5	1.889 ± 0.782	21.0 ± 58.877
0,1,4	1.0*	6.0*	1,2,4	–	4.0*	4	9.333 ± 20.174	3.833 ± 4.167
0,1,4,5	2.875 ± 1.959	5.9 ± 4.202	1,2,4,5	3.143 ± 2.673	4.889 ± 3.333	4,5	123.5 ± 30.395	124.0 ± 31.383
0,1,5	2.0*	2.0*	1,2,5	–	3.0*	5	7.3 ± 3.302	88.1 ± 234.779

Results & discussion

Table 6: Mean $\pm$ standard deviation (over 10 different seeds) number of positive/negative cluster intersections for different sets of energy bins bs or conditions[29]. This time, there's a discard window of 30 seconds.

bs or condition	negative	positive	bs or condition	negative	positive	bs or condition	negative	positive
# inter > 0	5762.1 ± 965.708	6487.6 ± 826.131	0,2	6.9 ± 2.726	10.7 ± 4.111	1,3	15.6 ± 6.398	51.2 ± 42.695
# inter > 1	4802.1 ± 822.34	5450.6 ± 752.954	0,2,3	6.0 ± 3.432	7.1 ± 1.37	1,3,4	1.0^*	–
# inter > 2	3493.8 ± 610.247	4083.5 ± 564.85	0,2,3,4	8.0^*	–	1,3,4,5	1.333 ± 0.577	1.5 ± 0.577
# inter > 3	3238.4 ± 602.769	3821.5 ± 538.957	0,2,3,4,5	212.2 ± 42.158	242.7 ± 65.109	1,3,5	–	5.5 ± 6.364
# inter > 4	1150.2 ± 232.379	1595.0 ± 202.499	0,2,3,5	1.0 ± 0.0	3.5 ± 3.536	1,4	1.0^*	3.0^*
# inter > 5	791.3 ± 165.14	1135.2 ± 149.199	0,2,4	1.5 ± 0.707	1.0^*	1,4,5	2.125 ± 1.356	5.5 ± 3.742
0	23.8 ± 8.23	71.6 ± 18.368	0,2,4,5	5.8 ± 4.367	15.8 ± 5.095	1,5	–	2.0 ± 0.0
0,1	16.6 ± 6.059	62.3 ± 45.631	0,2,5	1.0 ± 0.0	2.5 ± 2.38	2	734.8 ± 169.59	939.8 ± 235.236
0,1,2	4.0 ± 2.0	15.2 ± 9.739	0,3	1.0^*	3.0 ± 1.414	2,3	2459.9 ± 493.897	3265.5 ± 779.499
0,1,2,3	12.4 ± 3.806	30.5 ± 17.219	0,3,4,5	1.667 ± 1.155	1.75 ± 0.957	2,3,4	114.8 ± 335.669	11.75 ± 12.395
0,1,2,3,4	23.0^*	2.5 ± 0.707	0,4	2.0^*	–	2,3,4,5	3156.9 ± 657.884	4049.1 ± 676.033
0,1,2,3,4,5	791.3 ± 165.14	1135.2 ± 149.199	0,4,5	2.25 ± 1.753	11.6 ± 3.062	2,3,5	60.444 ± 11.479	225.0 ± 392.74
0,1,2,3,5	1.0 ± 0.0	3.667 ± 4.899	0,5	1.0^*	1.75 ± 1.5	2,4	18.333 ± 38.563	2.333 ± 1.751
0,1,2,4	1.0^*	3.0^*	1	52.6 ± 15.16	104.0 ± 118.135	2,4,5	212.3 ± 59.913	244.9 ± 73.942
0,1,2,4,5	2.8 ± 1.751	11.2 ± 7.315	1,2	3.444 ± 1.424	8.5 ± 4.143	2,5	9.8 ± 6.356	55.5 ± 127.411
0,1,2,5	1.0^*	2.667 ± 2.082	1,2,3	56.1 ± 19.221	115.1 ± 41.391	3	1206.3 ± 282.876	1633.0 ± 282.032
0,1,3	1.25 ± 0.5	4.875 ± 7.2	1,2,3,4	14.75 ± 27.5	1.4 ± 0.548	3,4	6.25 ± 9.215	1.0 ± 0.0
0,1,3,4,5	2.5 ± 2.811	2.25 ± 0.5	1,2,3,4,5	413.5 ± 141.535	714.4 ± 132.802	3,4,5	19.7 ± 8.166	16.5 ± 4.503
0,1,3,5	–	1.0^*	1,2,3,5	2.167 ± 0.753	9.667 ± 19.268	3,5	1.889 ± 0.782	21.0 ± 58.877
0,1,4	1.0^*	6.0^*	1,2,4	–	4.0^*	4	9.333 ± 20.174	3.833 ± 4.167
0,1,4,5	2.875 ± 1.959	5.9 ± 4.202	1,2,4,5	3.143 ± 2.673	4.889 ± 3.333	4,5	123.5 ± 30.395	124.0 ± 31.383
0,1,5	2.0^*	2.0^*	1,2,5	–	3.0^*	5	7.3 ± 3.302	88.1 ± 234.779

- Need an analysis of clusters' stability due to initial seeds

Conclusion: Summary

- Built a model of the background (light curve) using data collected from the POLAR detector

Conclusion: Summary

- Built a model of the background (light curve) using data collected from the POLAR detector
- Used poor predictions to extract time intervals

Conclusion: Summary

- Built a model of the background (light curve) using data collected from the POLAR detector
- Used poor predictions to extract time intervals
- Brief peek at model interpretability using partial derivatives of output w.r.t. input

Conclusion: Possible improvements

- Future analysis of our extracted time intervals is required
  - Categorize them into solar events, GRBs and so on

Conclusion: Possible improvements

- Future analysis of our extracted time intervals is required
  - Categorize them into solar events, GRBs and so on
  - Assess the clusters' stabilities due to initial seeds

Conclusion: Possible improvements

- Future analysis of our extracted time intervals is required
  - Categorize them into solar events, GRBs and so on
  - Assess the clusters' stabilities due to initial seeds
- Other possible improvements:

Conclusion: Possible improvements

- Future analysis of our extracted time intervals is required
  - Categorize them into solar events, GRBs and so on
  - Assess the clusters' stabilities due to initial seeds
- Other possible improvements:
  - Dive deeper into model interpretability

Conclusion: Possible improvements

- Future analysis of our extracted time intervals is required
  - Categorize them into solar events, GRBs and so on
  - Assess the clusters' stabilities due to initial seeds
- Other possible improvements:
  - Dive deeper into model interpretability
  - Change the methodology, e.g. use sequential models

Conclusion: Possible improvements

- Future analysis of our extracted time intervals is required
  - Categorize them into solar events, GRBs and so on
  - Assess the clusters' stabilities due to initial seeds
- Other possible improvements:
  - Dive deeper into model interpretability
  - Change the methodology, e.g. use sequential models
  - Improve code quality, GPU usage and decrease memory footprint.

Conclusion: Possible improvements

- Future analysis of our extracted time intervals is required
  - Categorize them into solar events, GRBs and so on
  - Assess the clusters' stabilities due to initial seeds
- Other possible improvements:
  - Dive deeper into model interpretability
  - Change the methodology, e.g. use sequential models
  - Improve code quality, GPU usage and decrease memory footprint.
- Manual inspection of our clusters by experts can give crucial information telling us what to do next.

The End

Credits

- Space station icons created by Freepik - Flaticon: `https://www.flaticon.com/free-icons/space-station`
- Collecting icons created by Leremy - Flaticon: `https://www.flaticon.com/free-icons/collecting`
- Detection icons created by Freepik - Flaticon: `https://www.flaticon.com/free-icons/detection`
- Analysis icons created by monkik - Flaticon: `https://www.flaticon.com/free-icons/analysis`
- Measure icons created by GOWI - Flaticon: `https://www.flaticon.com/free-icons/measure`

References

Figures without references come from author of the slides.

[1] Gilles Koziol, (2023)
HAGRID in Space. *Master's thesis*, University of Geneva.
https://cernbox.cern.ch/s/X1OvZ4iY19vewcV. Accessed: 2023-09-22.

[2] François Fleuret, (2023)
Deep Learning Course.
https://fleuret.org/dlc/. Accessed: 2023-09-22

[2] Shaolin Xiong, Yuanhao Wang, Zhengheng Li, Jianchao Sun, Yi Zhao, Hancheng Li, and Yue Huang, (2017)
Overview of the GRB observation by POLAR.
ICRC2017. page 640