Deep Learning for Space (T708-609OS)

ESA/ESOC – 16/06/2023

Agenda

- Welcome and Study Presentation

- Activities Summary
  - OPS/AIT Use Cases Selection & Technology Mapping
  - Use Cases Implementation & Results:
    - Fluctuation on Malargüe Ka-band signal amplitude
    - Mars Express Thermal Power Consumption
    - Device-Under-Test Events Investigation
    - Device-Under-Test Anomaly Detection

- Assessment, Lessons Learned & Recommendations

Study Overview

Study Overview

- The objective of the activity
  - to assess the applicability and impact of deep learning in space applications
  - to identify which space applications will benefit the most from deep learning
  - to map which deep learning methods (e.g. architectures) work best for each identified application

- Examples of deep learning applications to be considered includes: early anomaly detection, contextual anomaly detection, diagnosis, prediction, knowledge discovery and more

- Assess the impact of using deep learning in each application (e.g. cost reduction, risk mitigation, enabling functionality, increased science return)
  - For Space Operations (OPS)
  - For Assembly, Integration, and Test (AIT)

OHB ATG europe Hes·so SOLENIX

Consortium

Consortium

SOLENIX

Prime Contractor

Hes·so

Subcontractor

ATG europe

accelerating innovation

Subcontractor

OHB

Subcontractor

Study Team

Deep Learning Project Team

Alex Baumgartner
Contract Manager

Simone Fratini
Study Manager (Solenix)

Hatem Ghorbel
Subcontractor Study Manager (HES-SO)

Samo Simonian
Subcontractor Study Manager (ATG)

Pascal Vermeer
Subcontractor Study Manager (OHB)

HES-SO
Emmanuel de Salis
Célien Donzé
Farid Abdalla

ATG
Arturo González-Llana
Israel Pons

OHB
Hannes Schnisa

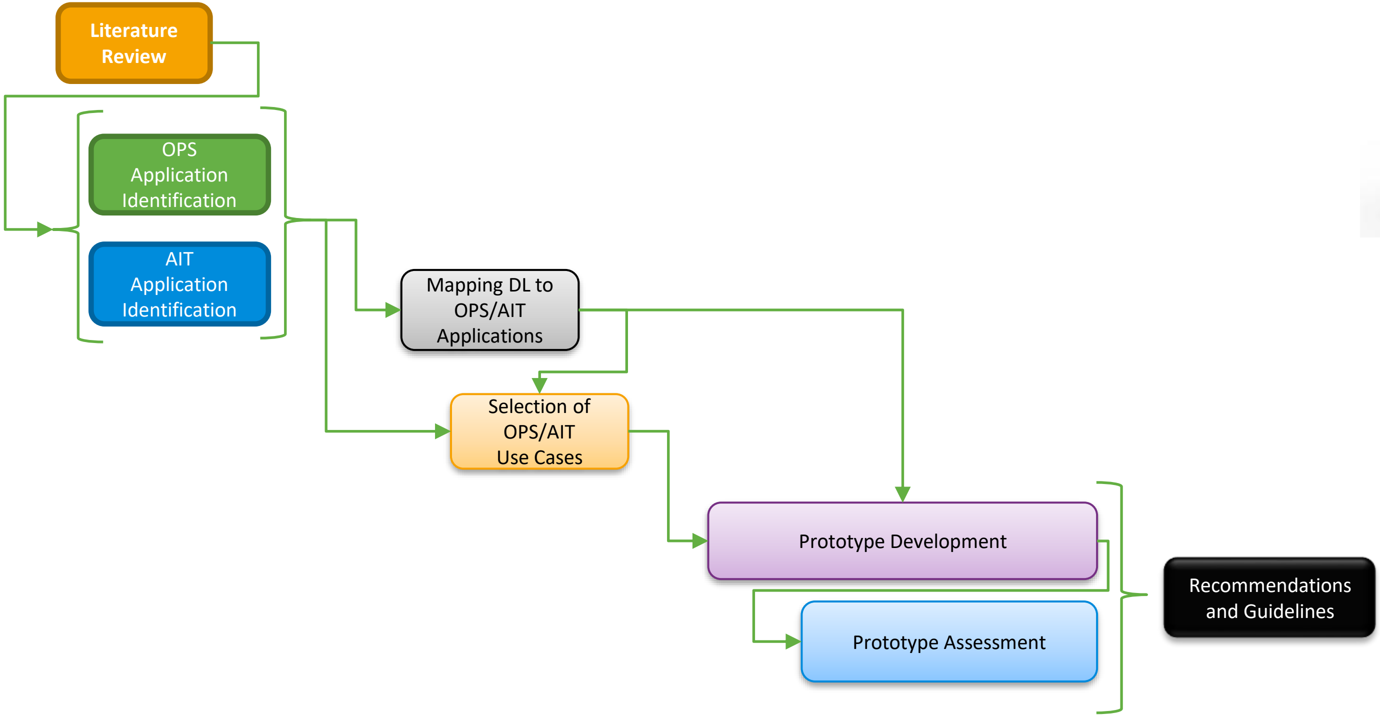
Solenix
Kevin Rojczyk
Cláudio Coelho
José Martinez

ESA / ESOC

Contract Officer

Technical Officer
Gabriele De Canio

Study Logic


Literature Review
OPS Application Identification
AIT Application Identification
Mapping DL to OPS/AIT Applications
Selection of OPS/AIT Use Cases
Prototype Development
Prototype Assessment
Recommendations and Guidelines


PLAN

Project Activities

AI, Machine Learning & Deep Learning Overview

- Artificial Intelligence is the general concept of intelligent programs
- Machine Learning (ML) is the usage of algorithms to create programs that can learn from data relationships
- Deep Learning is a subset of ML relying on deep Artificial Neural Networks and vast amounts of data to learn more complex relationships

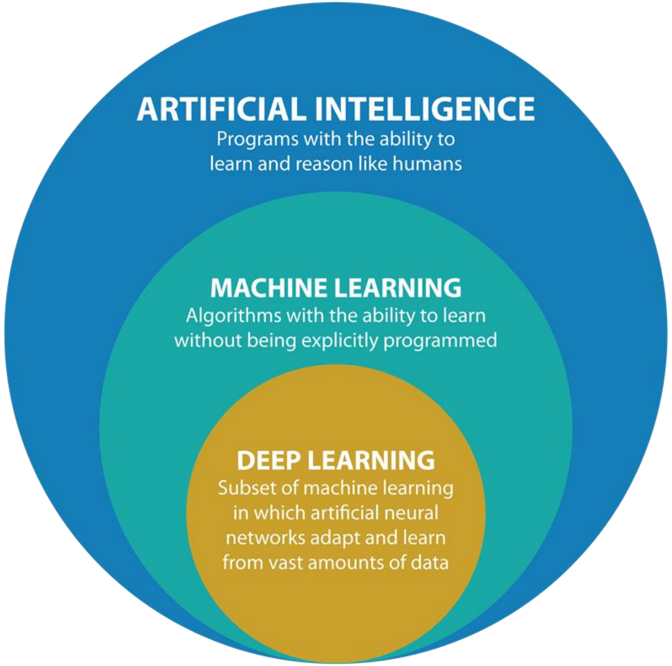
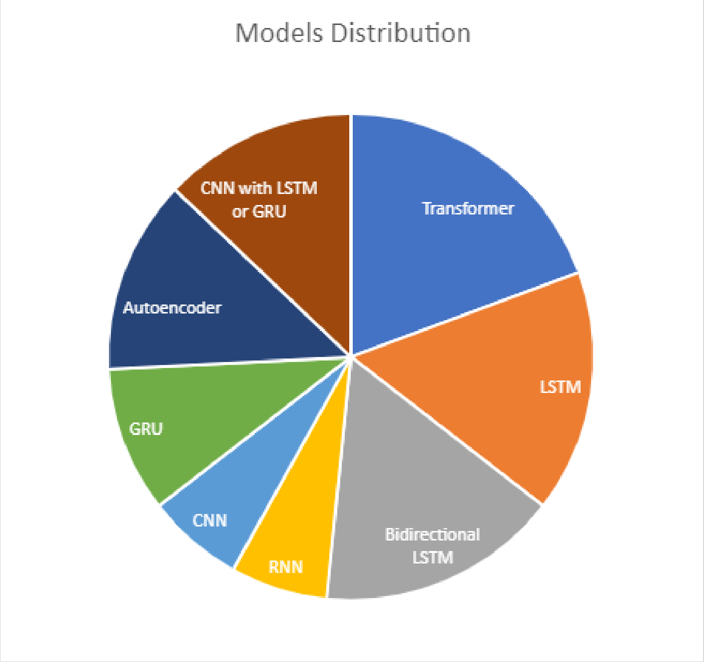


Image from: https://www.argility.com/argility-ecosystem-solutions/industry-4-0/machine-learning-deep-learning/

Literature Review

Latest Deep Learning technologies

- Usually studied conjointly with classical ML
- Advanced architectures are sometimes mix of ML and DL
- 711 publications were identified in the latest 5 years
- 20 use-cases were identified across those articles



OPS & AIT Application Identification

- 2 Workshops organized with ESA (ESOC/ESTEC)
- Viable Use Cases Identified :
  - 9 for OPS
    - **Mars Express Thermal Power Consumption**
    - Integral Radiation Belt Entry & Exit Prediction
    - Predicting the Impact of the Wind in Deep Space Antenna Pointing
    - Surrogate Models for High Computation Demanding Tasks (e.g. SIMULUS Simulators)
    - Find Optimal Policies with Reinforcement Learning
    - Anomaly Detection, Contextual Anomaly Detection, Anomaly Investigation
    - **Fluctuation of signal in Ka-bands links**
  - 7 for AIT
    - **Device-Under-Test Events Investigation**
    - Device-Under-Test Anomaly Detection
    - **Complex Systems Anomaly Detection**
    - EGSE Environment Event Investigator
    - EGSE Environment Anomaly Early Detection
    - Complex Systems Dynamics analyzer
    - Facilities maintenance and health monitoring

OHB ATG europe Hes·so SOLENIX

Selection of OPS & AIT Use Cases

- Data Availability – Do we have access to the data?
- Data Richness – Volume big enough?
- Data Quality – How many data errors ?
- Feasibility – Is the problem solveable by DL?
- Automatic Solution Evaluation – Can we check a solution without human interactions?
- New Paradigm – Novel approach or just DL applied?
- Scalability – Can it be adapted to other use cases?
- Relevance – What would be the impact?
- Timeliness – How urgent is a solution required?

OHB ATG europe Hes·so SOLENIX

OPS Prototypes

Fluctuation on Malargüe Ka-band signal amplitude

Summary of the use case

- Malargüe Ka-band signal link suffers from amplitude fluctuations
- Goal: understand what causes fluctuations (e.g., troposphere, wind, antenna position, etc.)

Fluctuation on Malargüe Ka-band signal amplitude

- Framed as a **nowcast** problem

TDCS (radiometer in tracking mode)
- IWV
- LWP
- ATT32000
- BTM31400

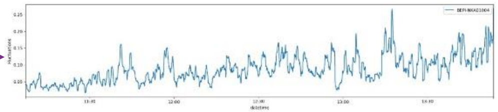
Weather Station
- Rain Intensity
- Wind Speed
- Wind Direction

Antenna Unit Control
- Azimuth + Elevation pos, average pos error, peak pos error

Feature Engineering

Deep Learning (Machine Learning)

Current Fluctuations



Fluctuation on Malargüe Ka-band signal amplitude

- Framed as a **nowcast** problem

TDCS (radiometer in tracking mode)
- IWV
- LWP
- ATT32000
- BTM31400

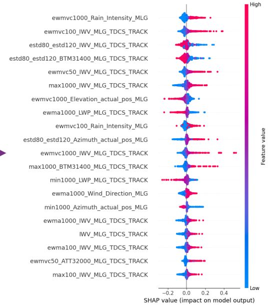
Weather Station
- Rain Intensity
- Wind Speed
- Wind Direction

Antenna Unit Control
- Azimuth + Elevation pos, average pos error, peak pos error

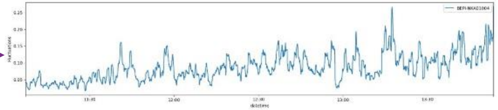
Feature Engineering

Deep Learning (Machine Learning)

Explainable AI



Current Fluctuations

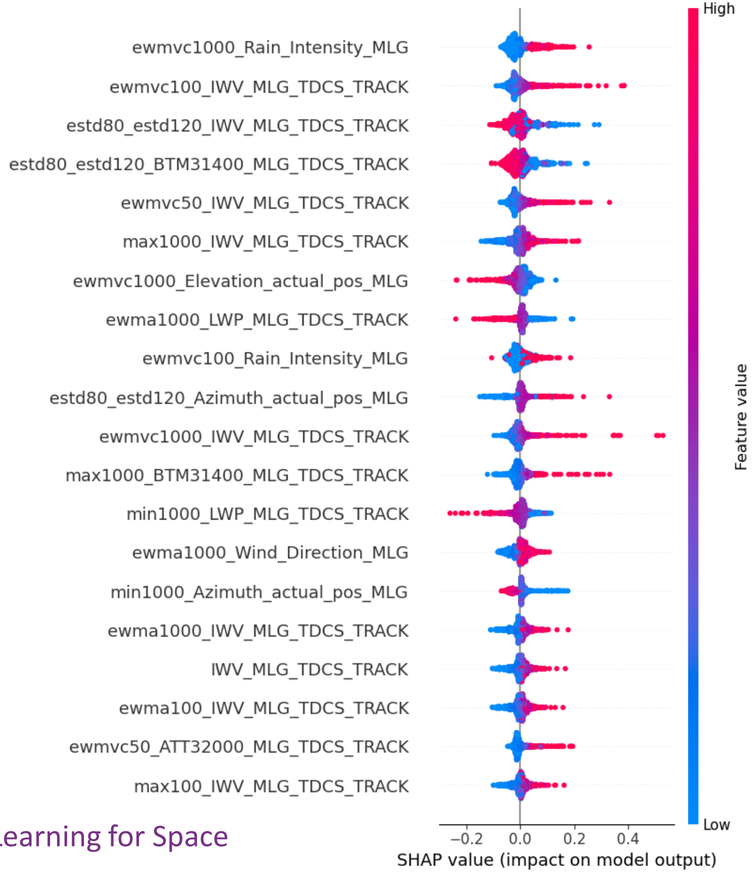


Fluctuation on Malargüe Ka-band signal amplitude

- Scenarios:
  - Uplink (KaT)
  - Downlink
  - Radio Science downlink

- Technologies
  - Classical ML (Gradient Boosting Trees) + Deep Learning
  - Deep Learning: Dense, 1D Convolutional, Recurrent
  - Explainable AI

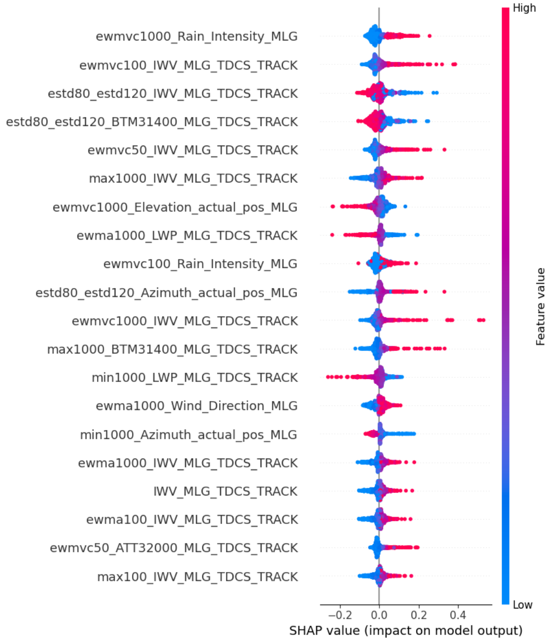
How to read SHAP summary plots



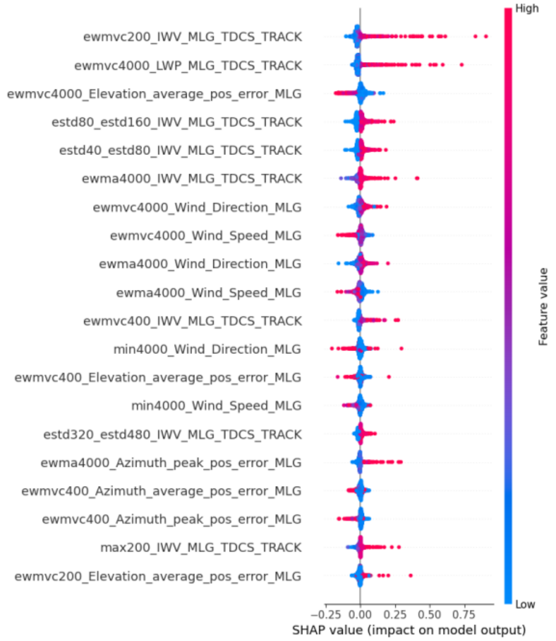
KaT Deep Learning Regression

ATG europe Hes·so SOLENIX

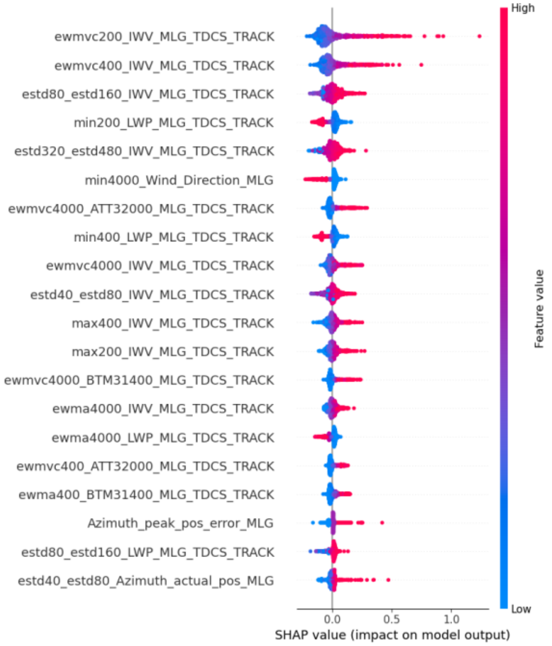
Fluctuation on Malargüe Ka-band signal amplitude



KaT Deep Learning Regression



Downlink Deep Learning Regression



Radio Science Deep Learning Regression

Malargüe Fluctuations use case - Conclusions

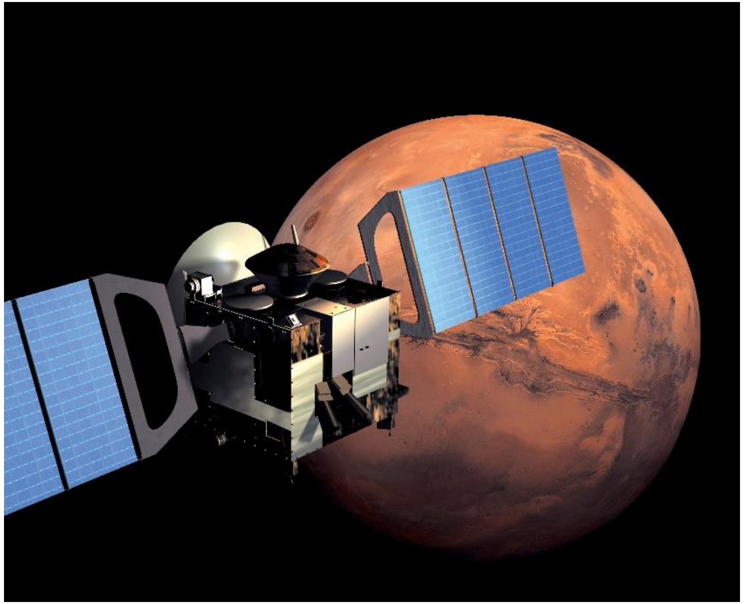
- Usually, DL and ML are used to predict the future
  - In this use case we used Deep Learning to nowcast
  - Explainability was the goal

- Explainability comes in terms of the provided features
  - Features need to carry predictive power
  - Features need to be easy to understand by domain experts

- Explanations from Deep Learning models found to be more useful than those provided by classical Machine Learning

Mars Express

- Mars Express orbiter
- Launched in 2003
  - Data available from 2008 to 2016
- Competition launched in 2016
  - To predict the power used by the thermal subsystem



Mars Express - Data

- 6 data types are available:

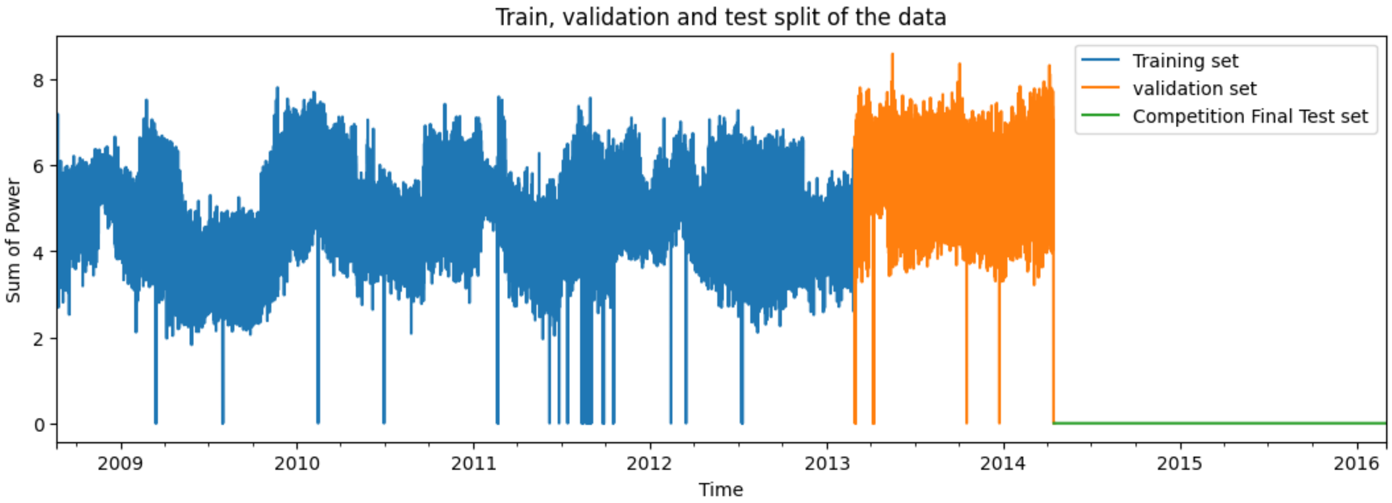
Input
1. **SAAF**: Solar Aspect Angles.
2. **LTDATA**: Long term data such as the sun-mars distance.
3. **DMOP**: Detailed Mission Operations Plan.
4. **FTL**: Flight dynamics TimeLine
5. **EVTF**: Other events (including eclipses).

Output
6. **POWER**: Electric current of 33 thermal power lines.

- Data is Downloaded from the competition website: https://kelvins.esa.int/mars-express-power-challenge/

Mars Express - Data

- Data splitting
- Operation changed in 2013
    - Validation set match Competition test set



Mars Express - Metric

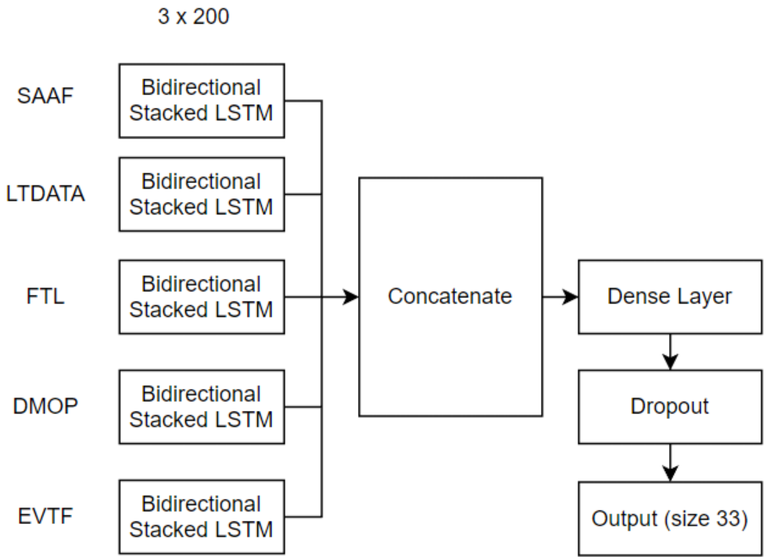
- Metric used: RMSE
  - Lower is better

$$\epsilon = \sqrt{\frac{1}{NM} \sum (c_{ij} - r_{ij})^2}$$

- Baseline score
  - Using mean values for each of the 33 power lines
  - Achieve RMSE of **0.138** on test set
- Best score of the competition is a RMSE of **0.08**

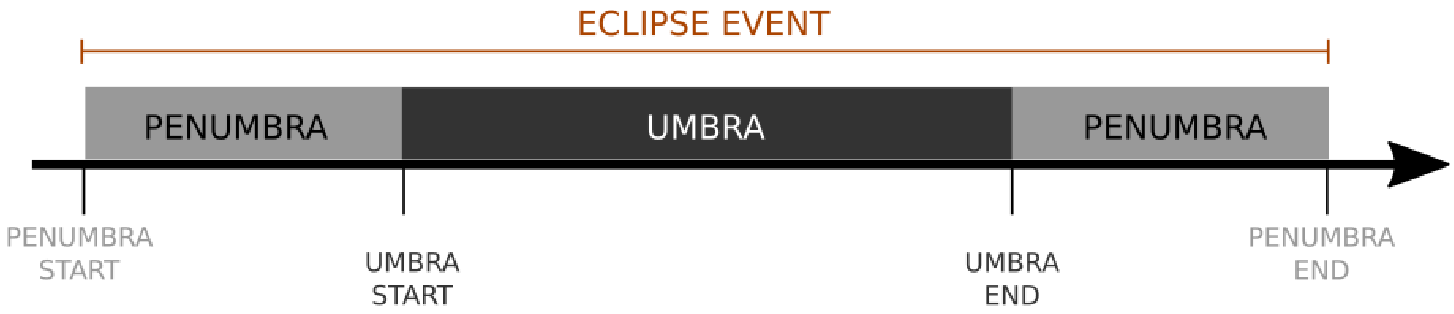
Mars Express - Models

- Data of different sampling rates
- Deep learning Models
  - Dense (Fully connected)
  - CNN
  - LSTM
  - Transformers (TST)
- 10 different models tested
  - With hyperparameter tuning



Mars Express - Models

- Adding CNN help the models
- Extracting Eclipses from EVTF file improve score



Mars Express - Results

- DL models are better than Baseline
  - CNN is the best architecture
- Not able to beat best of competition
  - Feature engineering helps a lot
- Scores are still in top 30 of competition

Model Version	Baseline (mean value)	V7 (CNN-LSTM)	V8 (DENSE only)	V9 (CNN only)	V10 (TST)
RMSE (best 0.08)	0.138	0.113	0.118	0.111	0.124
Training time	-	82 minutes	81 minutes	84 minutes	87 minutes

AIT Prototypes

AIT Overview

- AIT consists in a systematic step by step verification via testing from component level to fully integrated system
  - Focus of this project in electrical functional testing
  - Large amounts of data are systematically archived
- The spacecraft correct integration is checked
- The spacecraft software is verified based on housekeeping
  - Applications valid for AIT should be transferable to OPS

AIT Data for the Use Cases

- Complete E-AIT phases of the Meteosat Third Generation (MTG) project
  - Electrical functional test in every step of the integration
  - Full spacecraft tests
  - From 5 satellites
- Focus on housekeeping telemetries from the spacecraft software

AIT: Device-Under-Test Events Investigation

- Objective:
  - Go back to the root causes of an anomaly when an event happens.
- Type of Data:
  - Time Series:
    - Satellite Telemetry
    - Events of 4 different severities
  - The events can be used as labels for classification.

AIT: Device-Under-Test Events Investigation

- Methodology:
  - Supervised classification of events.
  - TSAI has been used to try multiple timeseries models
  - Use XAI (LIME) to find the importance of each input.
  - The model use a window of time before an event.
  - In simulated real time, when an event occurs the DL model prediction is compared with the real event. If identical, XAI can show the root cause.

AIT: Device-Under-Test Events Investigation

Severity 3

Classical ML (Random Forest)

	precision	recall	f1-score	support
0	0.98	0.99	0.99	139
1	0.00	0.00	0.00	2
2	0.00	0.00	0.00	1
3	0.94	0.85	0.89	40
4	0.00	0.00	0.00	1
5	0.98	1.00	0.99	81
6	0.00	0.00	0.00	1
micro avg	0.97	0.95	0.96	265
macro avg	0.41	0.41	0.41	265
weighted avg	0.95	0.95	$\boxed{0.95}$	265
samples avg	0.97	0.96	0.96	265

DL (Minirocket)

	precision	recall	f1-score	support
0	0.98	0.94	0.96	139
1	0.00	0.00	0.00	2
2	0.00	0.00	0.00	1
3	0.86	0.93	0.89	40
4	0.33	1.00	0.50	1
5	0.91	1.00	0.95	81
6	0.00	0.00	0.00	1
micro avg	0.88	0.94	0.91	265
macro avg	0.44	0.55	0.47	265
weighted avg	0.92	0.94	$\boxed{0.93}$	265
samples avg	0.91	0.94	0.92	265

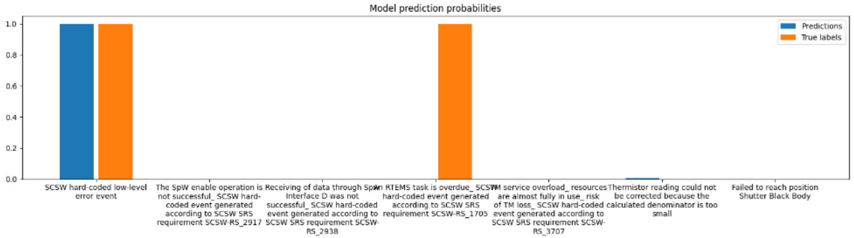
AIT: Device-Under-Test Events Investigation

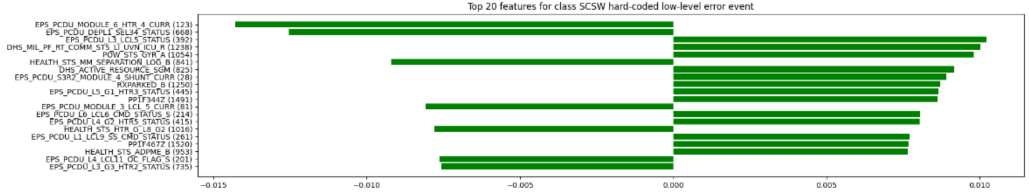
- Data for severity 3 events is quite small (~265 different events for the test set -> ~1060 for training, 7 classes of events).
- Performance for DL and ML methods is comparable
- DL however is able to find some classes that only have a few samples in the test set.
- For severity 4, DL seems promising,
  more data/ better filtering might improve the classifier.

	Baseline	ML (RF)	DL (MiniRocket)
Severity 4	0.01	0.04	0.21
Severity 3	0.56	0.95	0.93

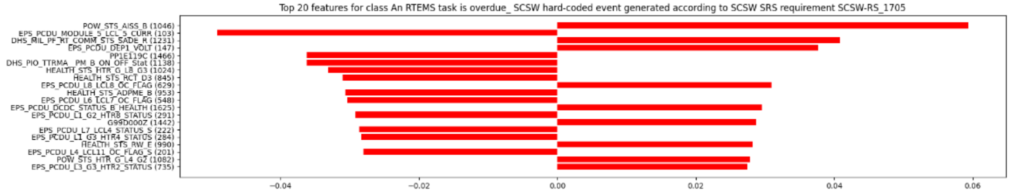
AIT: Device-Under-Test Events Investigation

- Example of explainability graphs





...



AIT: Device-Under-Test Events Investigation

- Multiple explainability graphs have been given to AIT experts
  - Physically meaningful relationships between inputs and events identified
  - Indirect dependencies found, compensating the lack of directly related signals in the inputs given
  - Only SCSW engineers with hands on experience on the mission could exploit fully the report → Need to involve the right personnel
- Explainability provides a good starting point to know which HK signals to focus on.

AIT: Complex Systems Anomaly Detection

- Objective:
  - Filter false alarms raised by the FDIR system by using a DL model to detect real anomalies.
- Type of Data:
  - Satellite Parameter Telemetry (Time Series)
  - High severity events
- Methodology:
  - Unsupervised Anomaly Detection
  - Events are used for fuzzy evaluation based on Precision@N, …
  - Evaluation of flagged events is done manually by AIT experts

AIT: Complex Systems Anomaly Detection

- Data Preparation (without final TM parameter selection)

[AIT SQL files]
260GB

[TM Data Sheet]
Sample Rate: 10s
Rows: 322k
Parameters 2.6k
Sessions: 279

AIT: CSAD – Benchmark Options

- Benchmark options:
  1. Algorithm: Isolation Forest *(baseline, classic ML)*, AutoEncoder, DeepSVDD, ~~Variational Autoencoder~~, ~~AnoGAN~~
  2. Contamination rate: (0.0, 1.0)
  3. Aggregated windows or sliding windows (with *n*-windows)
  4. Train on last session or on all previous sessions

AIT: CSAD - Key insights of the fuzzy evaluation

- Note: **Test set was too small** to provide secure results –
  only 111 high-severity events in total due filtering of specific subsystems
- **Isolation Forest seems to perform better if high precision is required** but
  only by sacrificing the recall
  (Prec. 0.18%, Recall 0.98%)
- The **deep learning models seem to have an advantage in a ranking scenario**, they lead the table ordered by Precision@N
  (DeepSVDD 0.55% vs 0.45% IForest)
- **DeepSVDD was our fastest algorithm**
  (DeepSVDD 40sec vs 140sec IForest)

AIT: CSAD – Manual Evaluation by AIT Experts

Anomaly data: TM Parameter, value and importance (based on LIME) which lead to the classification

AIT event data: TM Parameter and value which triggered the event

1.1.1 Finding #0 for timestamp 2021-07-27 10:49:30

Table with top most influencing parameters for the finding.

	importance	value
PP1E010K__0	0.025677	33.384617
PP1F742Z_OFF__0	0.024498	1.000000
PP1E129C__-3	0.022606	-0.000161
PP1E208V__-2	0.021499	0.019531
PP1E065C__0	0.017381	0.017425
PP1E080C__-1	0.016648	0.002070
PP1E124C__-3	0.016302	0.018332
PP1E022C__0	0.016034	0.015436

Table with information about the event

spid	name	rawValue	engValue
188880736	P88F50ZX	269942483	FCISTC_RP2L
	P88F554X	153	*NO VAL*
	P88F555X	656408980	PP1E252X
	P88F556X	1	EXPECTVALUE
	P88F557X	128	*NO VAL*
	P88F668X	0	*NO VAL*
	P88F669X	128	*NO VAL*
	P88F560X	0	EV_WL
	P88F561X	4	UV_BL
	P88F562X	9853035634518	1970-01-01T02:44:13.035
	P88F606X	9853	*NO VAL*
	P88F614X	597848	*NO VAL*

- Results:
  - Algorithm identified a valid high severity event in the dataset
  - Related signals provided to the tester had physical relationship with event.

Assessment, Lessons Learned & Recommendations

Lessons Learned - OPS

- Malargüe Use Case:
  - DL explanations better compared to explanations coming from classical ML
  - The XAI explanations come in terms of features. Craft features that are:
    - Predictive
    - Easy to understand by domain experts
- Mars Express Use Case:
  - Deep Learning algorithms could not catch some dependencies already known by experts, but still ranking good compared to other approaches
  - The absence of features engineering work to be considered in the balance

Lessons Learned - AIT

- The biggest challenge was related with data selection, preparing and processing

- Data structure, standard and practices of AIT are not defined with the purpose of being integrated in a ML/DL loop

- Once the data is available and processed, DL and classical ML algorithms don't show major differences

- AIT data availability for OPS shows great potential for DL/ML application

Lessons Learned – DL vs Classical ML

- It is not conclusive if Deep Learning is always the right choice for any problem
- Classical ML methods and newer DL should be considered as complementary tools of a toolbox
- Deep Learning vs Classical Machine Learning
  - Deep Learning gives much more possibilities than Classical ML
  - Need a lot of independent data for Deep Learning
  - Classical ML works better with smaller datasets
  - The community is working more on Deep Learning now
- Anomaly Detection
  - Deep Learning is not always superior to Classical ML
  - Deep Learning require more resources
- Conclusion: **Do not put all your eggs in one basket**

Recommendations

- Data Quality & Labeling
  - Investment should focus on data preparation and labeling to improve the data quality and usability
  - the cleaner and better documented the data is, the easier and more efficient it becomes for data-driven projects

- Explanation
  - Create good features if the explainability is the main focus
  - Features engineered need not only to carry predictive power but also to be intuitive to domain experts

- A New Mindset
  - Instore a DL/ML mindset earlier in the process
  - Foster cross-functional synergy between experts and industry during all the phases of the mission

Conclusions

- A review of Deep Learning literature oriented to Space Applications
- An analysis of possible applications in OPS and AIT that could benefit from DL
- An analysis of most promising DL technologies for space applications
- Evaluation of the impact of Explainable AI (XAI) to understand how DL technologies are taking decisions
- 4 prototypes to show and study applicability of Deep Learning in Operations and in Assembly, Integration & Tests
- Assessment, Lessons Learned and Recommendations for the applicability and benefits of DL for space applications

Thank You

Solenix Engineering GmbH
Spreestrasse 3
64295 Darmstadt
Germany

info@solenix.de

www.solenix.de

Discussion