Spacecraft System and Subsystems models optimization by AIT/AIV and Operations big data analysis (SMO)

Space system development lifecycle foresees a number of activities to be executed in order to checkout a system or subsystem. During these activities, a huge quantity of data is generated and mainly used for checking the outcome of tests, resulting in not fully exploiting the data for the potentially strategic optimization of the system. SMO (Spacecraft system and subsystems models optimization by Assembly, Integration & Test/Verification (AIT/AIV) and operations big data analysis) is a system that introduces a new systematic approach in the space system development lifecycle based on the exploitation of data generated from the design, integration, validation and operating phases, with the scope to create accurate ML models for each system and subsystem lifecycle phases.

To improve the data exploitation, the usage of big data and ML techniques is useful in order to reveal new findings on the system behaviour, starting from the analysis of system telemetry and/or test results.
This approach gives the possibility to compare engineering models and as-designed models, i.e. the ones coming from the early phases of a mission, with data-driven models, based on as-built, as-run, as-flown systems, which involve testing and operating mission phases instead.

The knowledge extracted by both AIT/AIV and operating data can be used to support and optimize the specification, the evolution and the maintenance of system/subsystem the data refer to.
The creation of space systems data-driven models is extensively used along the system development lifecycle for supporting the design, functional verification, validation and domain specific analyses.
Moreover, the most significant novelty consists of having a software system that enables both data exploration capabilities and a selection of ML models generation capabilities for spacecraft engineers, who are not data scientists.

Five use cases and four different end-users have been involved in the framework development to demonstrate the system goals in real scenarios: Euclid for spacecraft under integration and PMM and XMM as spacecrafts in operations.

At the end of the project the main achievements and lesson learned are the following:

- the big data techniques and technologies, as already demonstrated in other big data platform in the space domain, are mature to implemented a so complex multi-mission and multi-users data system;
- several machine and deep learning technologies have been integrated on top of the big data platform to have effective data exploitation offering several ML and DL frameworks to data scientists;
- AIT/AIV data are often insufficient to train ML model of several subsystems therefore data scientists could also exploit simulated data alone or in combination with real data to improve the accuracy of ML models;
- the availability of operational dataset is the most favourable scenario to develop data driven models although not all bounded system and subsystems conditions are present in the historical telemetry;
- the role of data scientist is central also in a data system as SMO aiming at leveraging the ML technologies to not experts, in the SMO workflow data scientists are facilitators that through SMO capabilities prepare the ML tools for the S/C Engineer;
- the comparison between as-designed models and data-driven models (as-built/as-run/as-flown models) is possible but it is not an immediate task nerveless the availability of data comparison tool. It is needed that the S/C Engineer brings into SMO the data knowledge, shares with data scientists the as-design model know-how and then Data Scientist can prepare and validate the model comparison component for a specific use case based on the available comparison tools;
- the reuse of data preparation, feature extractor and training and validation module has been demonstrated both within the investigation of the same spacecraft system/subsystems and across different spacecraft system/subsystems;
- ML model re-train capability and workflow have been defined and implemented to allow authorised S/C Engineer to retrain machine learning model in the cases where is valuable and feasible.