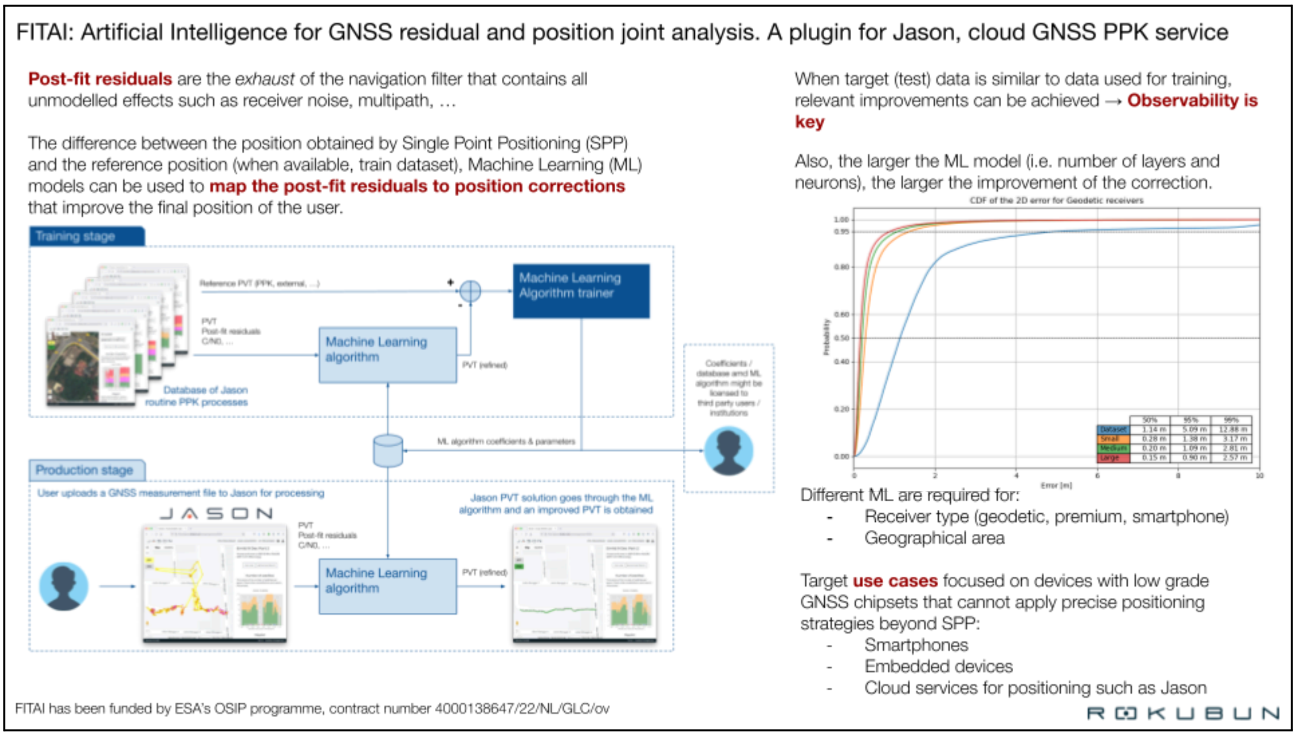

FITAI: Artificial Intelligence for GNSS post-fit residual and position joint analysis. A plugin for Jason, Rokubun cloud GNSS processing service

Executive Summary Report

Early technology development

OSIP Open Discovery Ideas Channel (Commercial and technical maturation of ESA's inventions)

Affiliation(s): Rokubun

Activity summary:

FITAI is a Machine Learning based technique that improves GNSS positioning engine accuracy by mapping signal strength and residuals into position errors. These inputs are fed into a neural network to learn and correct errors, offering a non-invasive solution that provides better location precision without relying on external sources like reference stations.

Publishing Date: 2024-03-15
Contract Number: 4000138647/22/NL/GLC/ov
Implemented as ESA Initial Support for Innovation

ESA Discovery & Preparation
From breakthrough ideas to mission feasibility. Discovery & Preparation is laying the groundwork for the future of space in Europe
Learn more on www.esa.int/discovery
Deliverables published on https://nebula.esa.int

→ THE EUROPEAN SPACE AGENCY

FITAI is a system that improves the accuracy of a positioning engine through deep learning techniques by using GNSS information, such as post-fit residuals, SNR, elevation, and azimuth, to train a deep neural network that corrects the positioning engine's outputs.

A positioning engine uses measurements from satellites or other sources to determine a user's position, velocity, and time. After the solution is computed, the residuals represent the difference between the actual measurements and the predicted measurements based on the solution. In particular, the post-fit residuals represent all the unmodelled errors that could not be absorbed by the parameters estimated by the positioning.

In FITAI, the neural network is trained using the post-fit residuals obtained from a Single Point Positioning (SPP) solution. The network makes use of a reference solution (PPK or Ground Truth) to calculate the correction in the East, North, Up (ENU) components relative to the SPP solution. This training process enables the model to predict the expected correction, ultimately refining the positioning result.

The general system overview of the FITAI architecture is shown in figure 1:

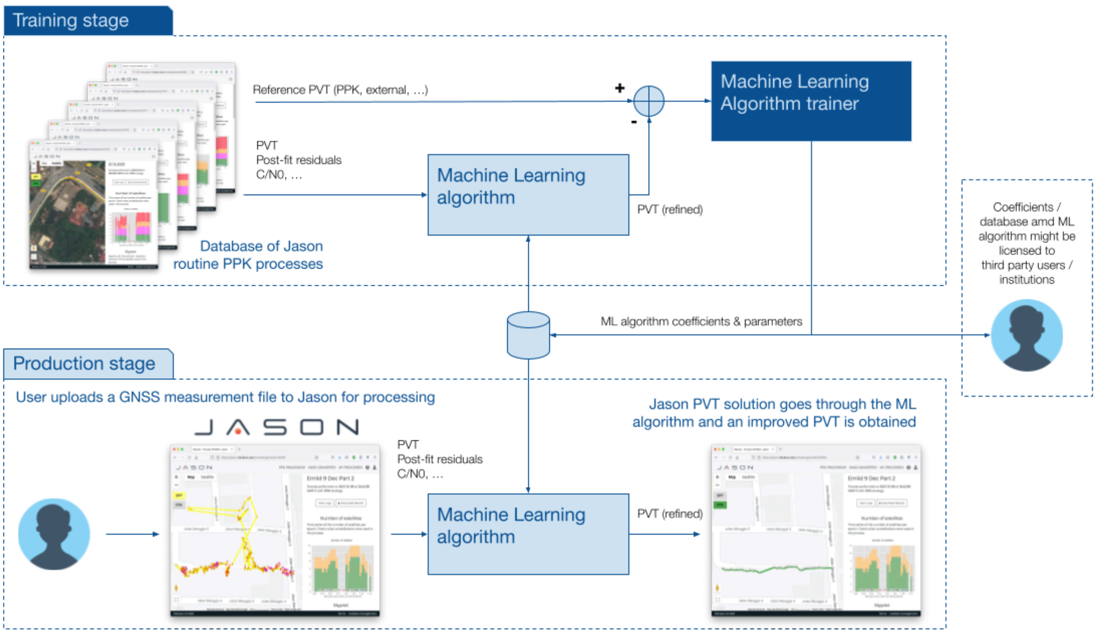

Figure 1: General overview of the FITAI architecture and data flow

Data overview

Data is one of the most crucial components of any machine learning problem because it forms the foundation on which the models are built. The quality of the data can have a significant impact on the accuracy and reliability of the resulting models. Unrepresentative data can limit the model's ability to generalize to new data, leading to poor performance in real-world scenarios.

In the context of FITAI, one of the main challenges that the project faced was having enough data to generalize to any possible scenario for a positioning engine regardless of the location of the data. In addition, the data needs to be properly tagged according to, e.g.: kind of environment (urban, non-urban, open-sky), quality of the receiver (geodetic, high-end, smartphone), number of satellites in view for the computation of the solution and its distribution in the sky, among others. FITAI uses the following sources of data to train the models:

- JASON: Rokubun's Cloud GNSS Processing service that allows users to upload GNSS data and performs the best possible solution based on the data available. It prioritizes performing PPK if a CORS base station is close enough with data available.

- Dedicated data campaigns: Rokubun performs regular data take campaigns to assess the performance of the algorithms being developed for different projects. These data takes are mainly recorded around the Barcelona area with multiple devices and ground truth to compare the performance with different receivers.

- Google Decimeter Challenge: FITAI uses the dataset provided by Google in the three Kaggle competitions. It contains GNSS data from a different set of smartphones (Google Pixel, Samsung Galaxy S20, Xiaomi Mi8) together with a truth obtained with a geodetic receiver.

FITAI also takes into account the grade of the receiver to have more information on the kind of data being ingested. Hence, three main categories have been considered in the project. Table 1 shows the number of samples per receiver category used in each of the datasets.

Type	Samples
Geodetic	3.3M
Premium	5M
Smartphone	1.6M
All	**9.9M**

Table 1: Number of samples per receiver category

The data workflow in a machine learning project involves organizing, storing, and processing the data used to train, validate, and test machine learning models. The workflow developed in FITAI is shown in Figure 2:

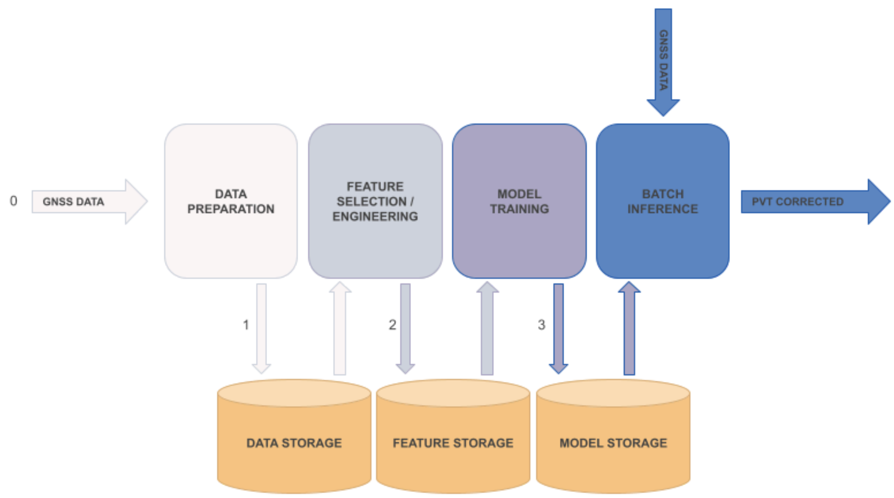

Figure 2: Data workflow in FITAI

Each of these stages can be summarized as follows:

- Data Preparation: Raw GNSS data (0) is received and processed in order to get all the required information to train the model. This data is stored (1) and used in the subsequent stages.
- Feature Selection and Engineering: The data created in the first stage (1) is used to create a set of features (2) that will depend on the type of model that is going to be trained.

- Model Training: The features generated are read (2) and preprocessed before training the model. For instance, the data is normalized, if required, and split in the training, validation, and test datasets with a proportion that will depend on the amount of data available. Once the model is trained it is stored (3) together with a `model.json` file that contains relevant information about the trained model, such as the number of neurons per hidden layer, number of hidden layers, variables used (post-fit residuals, SNR, etc), etc.
- Batch Inference: Once the model is trained, it can be deployed to production by loading the saved model (3) and performing the inference process on new GNSS data.

One key aspect of FITAI is the selection of the features as the input to the neural network. FITAI organizes the input layer based on the elevation of the satellite, with a fixed size of 27 bins each one of three degrees. Then the postfit residual for each satellite is assigned to one of those bins. One drawback is that there may be some data loss since more than one satellite can be assigned to the same elevation bin. Figure 3 show a representation of the Fully Connected Neural Network developed in FITAI.

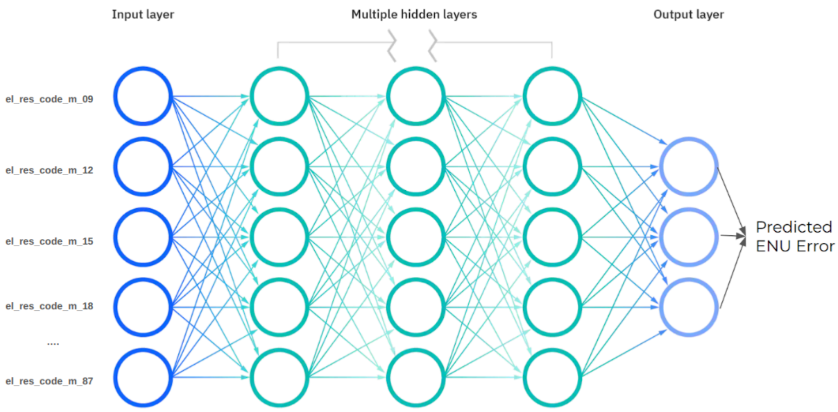

Figure 3: Neural Network architecture developed in FITAI

Results

The results show a potential improvement using FITAI models when the GNSS data is similar to the one used during model training. Figure 4 shows the cumulative distribution function (CDF) of the 2D error for the Test dataset a model trained with geodetic data:

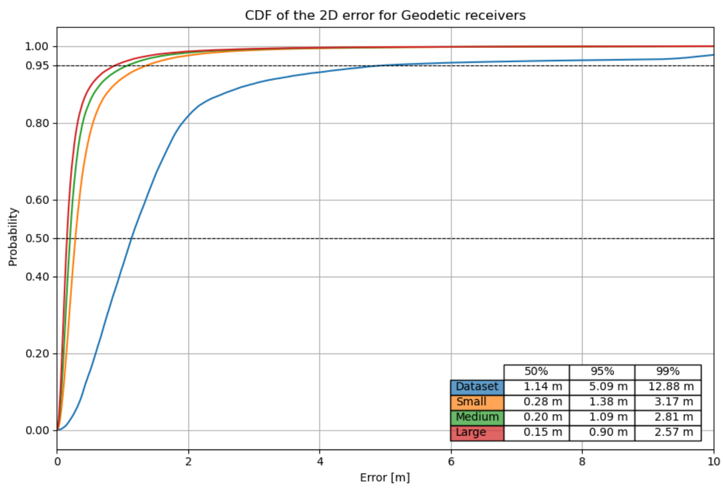

Figure 4: 2D error CDF for geodetic receivers

Similar results are obtained for premium devices, although the improvement in smartphones is smaller. However, when assessing the performance of the models against data with different locations and receivers from the one used, the improvement is smaller and in some cases can degrade the SPP solution.

Conclusions

Data is one of the most important factors when training models, although sometimes more does not mean better results. In particular, one of the key aspects seen during the project is that when creating models, it is important to focus on what problem needs to be solved and which kind of receiver is going to be running the navigation engine. Having models trained with different receivers, but then run them only in a subset of them may waste accuracy in the results. The behavior of the values of the post fit residuals for a geodetic receiver is different from the ones for a smartphone.

The results show that using specific models for a receiver category or even a receiver model can achieve better accuracy than a general one. This is important because it affects the planning to perform campaigns to obtain data. For instance, if the aim is to set a navigation platform in a car manufacturer, the focus should be to use data coming from the type of receiver that is going to run the model instead of trying to use data from smartphones.

This has been seen during the validation where models trained for datasets with a lot of diversity in the input do not improve the solution or it improves below 10%.

Another conclusion is that the size of the Neural Network (NN) improves up to a certain limit: differences from Small to Medium are noticeable, but from Medium to Large are not that high. Meanwhile the cost of storage and computational power increases by four. Therefore, going for larger NN topologies is not clearly justified considering the time it requires to train the model and predict the correction.

Our results demonstrate that the NN that exploits the residuals along the elevation profile can improve the positioning accuracy of the Single Point Positioning (SPP) solution. However, the findings do not provide a definitive conclusion regarding the network's generalizability.

However, one of the reasons to use these features is that the number of neurons in the input layer is smaller than doing it based on per Satellite and Channel. Supporting multiple constellations increases the size of the input layer to around a hundred features while the elevation strategy keeps the number of neurons fixed to the number of bins. The current implementation has a length of 27 neurons per property, which means that using code residuals and SNR keeps the input layer at 54 neurons.

Business model

Crucial to the successful launch of the FITAI solution is the selection of the right business model.

This concept will be commercialized in two different implementations:

1) FITAI offers a **cloud-based service** that allows users to query with their position and receive corrections to improve their accuracy.

This service could be made available as a subscription and could be commercially exploited as a B2B2C solution for mobility operators, car manufacturers, fleet operators, and other businesses seeking to enhance their users' (or assets) positioning.

2) FITAI can offer a **license for embedded software** that can be deployed in navigation devices to improve the accuracy of position information delivered by GNSS chips. This software could be commercially exploited through IPR licensing by charging a small fee for each customer device that includes FITAI's embedded software.