IDAO2020Team-256 Logout

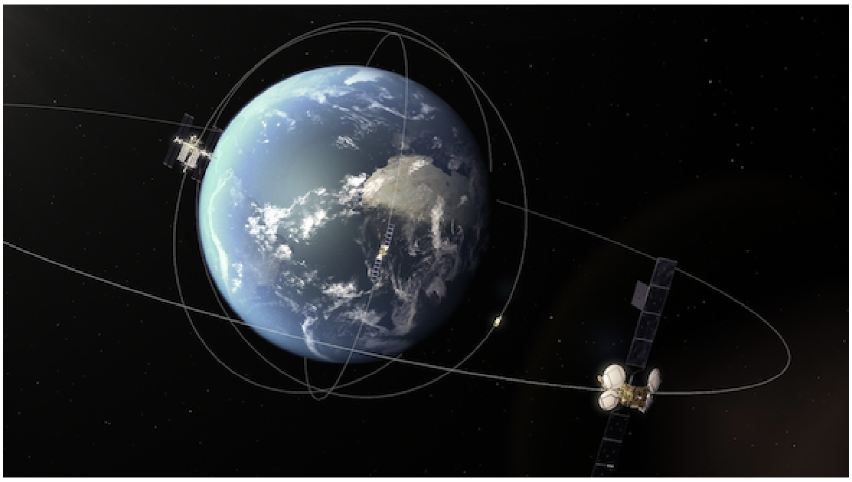
IDAO 2020

23 Jan 2020, 08:31:40
start: 15 Jan 2020, 16:00:00
finish: 12 Feb 2020, 03:59:59
till the end: 19d. 19h.
start: 15 Jan 2020, 16:00:00
end: 12 Feb 2020, 03:59:59
duration: 27d. 11h.

A. IDAO 2020 Track 1

Objective.

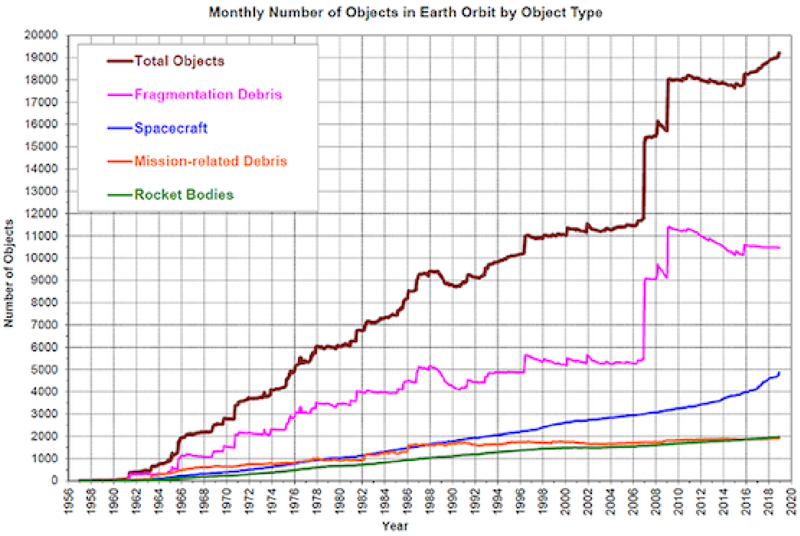
Build a model that would predict the position of space objects using simulation data.



Picture source

Background information.

Predicting the position of satellites is one of the most important tasks in astronomy. For example, information on the exact position of satellites in orbit is necessary to avoid extremely dangerous satellite collisions. Each collision leads not only to satellites destruction, but also results in thousands of space debris pieces. For instance, Iridium-Coscos collision in 2009 increased number of space debris by approximately 13%. Further collisions may result in Kessler syndrome and the inaccessibility of outer space. Also, a more accurate prediction of satellite position will help calculate more efficient maneuvers to save propellant and extend satellite life in orbit.



Picture source

At the same time, it is difficult to predict the exact position of a space object, since the form and other parameters might be unknown. Furthermore, mathematical-physical models cannot fully take into account many factors, such as the Earth's Surface Roughness and Solar pressure. One of the partners of the Olympiad and the author of the task is the Russian Astronomical Science Center (ASC). ASC is engaged in the creation and operation of automated optoelectronic systems for observing space objects and automated systems for processing information about space objects. This problem has special relevance to them. For example, maintaining the Catalog of Earth Satellite Orbits requires repeated observations of space objects. To do such observations, it is necessary to know where to target the telescopes (video).

Picture source

IDAO participants are asked to clarify the prediction of a Simplified General Perturbations-4 (SGP4) model. SGP4 is able to predict a lot of effects but is applied to near-Earth objects with an orbital period of fewer than 225 minutes, while high orbit space objects have orbital periods up to 200 hours. For the true position of the satellites, the position obtained using a more accurate simulator will be taken. Subsequently, the obtained models will be applied to real classified data and will help to predict the positions of these space objects.

Examples of SOTA — Improving Orbit Prediction Accuracy through Supervised Machine Learning, Machine Learning Approach to Improve Satellite Orbit Prediction Accuracy Using Publicly Available Data.

Quality metric.

Predictions of satellites' coordinates and velocities made by participants will be compared to the ground truth using SMAPE score.

More specifically, for a satellite and points in time, all the submitted predictions and true values are concatenated into 1-d arrays, then SMAPE is computed by the equivalent of the following code:

```
def smape(satellite_predicted_values, satellite_true_values):
    # the division, addition and subtraction are pointwise
    return np.mean(np.abs((satellite_predicted_values - satellite_true_values)
        / (np.abs(satellite_predicted_values) + np.abs(satellite_true_values))))
```

After that SMAPE is averaged over the satellites. Formula:

$$MEAN\,(SMAPE) = \frac{1}{n_{sat}} \sum_{sat=1}^{n_{sat}} \left[\frac{1}{6points_{sat}} \sum_{point=1}^{points_{sat}} \sum_{k=1}^{6} \frac{\left|pred^k_{sat,point} - true^k_{sat,point}\right|}{\left|pred^k_{sat,point}\right| + \left|true^k_{sat,point}\right|} \right],$$

where:
n_{sat} —number of satellites for averaging,
sat —index of satellite,
$points_{sat}$ — number of measurements (points in time) for satellite sat,
$point$ — index of the point,
$pred_{sat,point}, true_{sat,point}$ — predicted and true vectors of satellite coordinates and velocities (x, y, z, V_x, V_y, V_z) for point $point$ of satellite sat
.

Formula for the leaderboard score: $SCORE = 100 * (1 - MEAN\,(SMAPE))$. This score should be maximized by the competitors.

Input format

Train dataset for both tracks and test dataset for this track can be found in the data folder.

Dataset fields description:

- id — integer measurement id
- $epoch$ — datetime in "%Y-%m-%dT%H:%M:%S.%f" format (like 2014-02-01T00:44:57.685)
- sat_id — integer satellite id
- $x, y, z, x\_sim, y\_sim, z\_sim$ — real (obtained with the precise simulator) and simulated (obtained with SGP4-simulator) coordinates of the satellite (km)

- Vx, Vy, Vz, Vx_sim, Vy_sim, Vz_sim — real (obtained with the precise simulator) and simulated (obtained with SGP4-simulator) velocities of the satellite (km/s)

Train dataset contains real and simulated coordinates and velocities for 600 satellites, for one month. Test dataset contains only simulated coordinates and velocities for 300 of the satellites from train dataset, for the next month. Participants should predict real coordinates and velocities for these 300 satellites.

Output format

Please upload your predictions into the system in the .csv format. Your submission file must contain seven columns: id, x, y, z, Vx, Vy, Vz. Each line in the submitted file must correspond to the measurement in test dataset. A sample submission can be found in the data folder.

Public leaderboard is based on 50% (150 satellites) of the test data. Private leaderboard is based on the remaining 50% and will be revealed after the end of competition.

Only the last successful submission is used to compute public and private leaderboard scores. You should make sure, that at the end of competition the last successful submission is actually your very best attempt!

Notes

You may use additional information, data, advice on your own risk. All solutions will be checked by the jury to guarantee fair play. Please ask us in any vague or unclear case.

If you face any problems or have questions, please contact us via e-mail: hello@idao.world.

Select File is not chosen

Submit

Next