# Classifying Satellite Orbits

Richy Peterson

### Air Force Institute of Technology

### *I have color-coded this template to show which sections are required for homeworks 4, 5 & 6. When turning in HW4, please change your updated HW4 text to black, and retain the placeholder text for HW 5/6. Similarly, when turning in HW5, retain the placeholder text for HW6.*

* Homework 4
* Homework 5
* Homework 6

### Abstract — Describe and motivate the problem. Mention your approach and list your key findings or results, along with why they are important. Include 7 or more numbers, such as number of datapoints, modeling metrics, or the benefits of the model. Compare the performance of the best model to the trivial model.

*Keywords — blah, blah and blah*

### I. Introduction & Background

The Spectrum Program Executive Office[[1]](#footnote-1) (PEO Spectrum) of the Defense Information Systems Agency (DISA) maintains the Joint Spectrum Data Repository (JSDR). JSDR is the Joint Authoritative Data Source for all DoD spectrum-related data [1]. As such, it is a repository of spectrum management supporting data sources that provides access to frequency assignments, spectrum certifications, interference reports, engineering characteristics, and equipment data. One segment of the JSDR is the Joint Spectrum Center (JSC) Equipment, Tactical, and Space (JETS) database. The JETS database contains data for DoD, commercial, and coalition equipment, as well as unit-equipment and platform-equipment relationships.

The roots of JETS began in the 1970s and much of the data was manually intensive to maintain. The JETS database consists of over 50 tables with several tables consisting of over 100 fields for the tens of thousands of records. In the past decade several of the JETS data sources dried up and in 2018 the JETS databased was no longer maintained. In 2020 PEO Spectrum was tasked with updating and maintaining the JETS database [2].

Several years later the space component of the JETS database is still being updated. As is it updated it is important to include key data fields that enable simple and efficient data queries for users. A common field for satellite data is the orbit type. The four most common orbit types are low earth orbit (LEO), medium earth orbit (MEO), geosynchronous orbit (GEO), and highly elliptical orbit (HEO) [3]. These are categorized based on the mean motion and eccentricity of the satellite orbit. The mean motion is the number of revolutions the satellite makes in 24 hours. The eccentricity is a measure of how elliptical the orbit is. More specific orbit types (e.g., sun-synchronous, polar, geostationary, and Molniya) are grouped based on factors such as the orientation, direction, and inclination of the satellite orbit.

As the satellite data is updated in the JETS database, an automated process to validate or determine the orbit type is needed. In the case where the orbit type is not present in the originating data record, the process should determine the orbit type and apply the label to the record. When ingesting data with an orbit type already specified, the process should validate the orbit type and indicate when a different label should be applied. The purpose of this analysis is to determine how well machine learning and neural network algorithms can predict the satellite orbit type based on other orbital parameters in the data. To be applied to new data fed into the JETS database the model must generalize to unseen data.

While the motivation for this analysis, to update a database, is likely unique, other orbit type classification work has previously been performed for cislunar satellite orbits [4] and asteroid orbits [5]. Machine learning has also been used to classify types of satellite maneuvers in orbit [6], [7]. The similarity in all these is classifying a type of orbit or maneuver from orbital parameters, or from changes in parameters for the case of maneuvers.

1. *Data Acquisition*

The data for this analysis was retrieved from the JSDR[[2]](#footnote-2) using the JETS Satellite Query tool. To obtain the entire dataset no filters were specified in the query. The complete dataset contained 55 fields and 9,539 records. However, most of the data was either missing the orbital parameters or was labeled as “dummy” or “filler” records. Those records, along with all duplicated records, were removed. All controlled unclassified information was also removed. The cleaned dataset contained only 11 data fields and 1,022 records.

1. *Data Understanding*

The 11 data fields are listed and described in Table 1. The first 10 are the orbital parameters and will be the input variables for the model. The last field, Orbit Type, is the target variable the model is trained to predict. The distributions for the orbital parameters are shown in Fig. 1. The Argument of Perigee, Mean Anomaly, and Right Ascension of Ascending Node distributions somewhat resemble a uniform distribution. All other variables do not resemble common distributions but have discrete spikes at certain values which generally correlate to the various orbit types. The Eccentricity is an exception where there is a single spike and there appears to be no other values. There is a range of eccentricity values all the way up to 0.97 but the proportion compared to the number around zero is so low they are not visible in the plot. Similarly there are a few data points that have much higher values for the apogee distance, the period, and the satellite height but they are not visible in the figure.

Table 1. JETS Data Field Descriptions

|  |  |
| --- | --- |
| Variable | Description [8] |
| Apogee (km) | Distance from Earth center to the farthest point in the orbit. |
| Argument of Perigee (deg) | Angle between perigee and the ascending node. |
| Eccentricity | Distance between foci of the orbital ellipse normalized to the major axis. |
| Inclination Angle (deg) | Angle between the orbital plane and the Earths equatorial plane. |
| Mean Anomaly (deg) | Angle between perigee and the satellite position |
| Mean Motion (revs/day) | Number of revolutions in one day |
| Perigee (km) | Distance from Earth center to the nearest point in orbit. |
| Period (minutes) | Time of a single orbital revolution. |
| Right Ascension of Ascending Node (deg) | Angle between the vernal equinox and the point the orbit crosses the equator traveling north. |
| Satellite Height (km) | Height above the surface of the Earth. |
| Orbit Type | Type of Orbit |

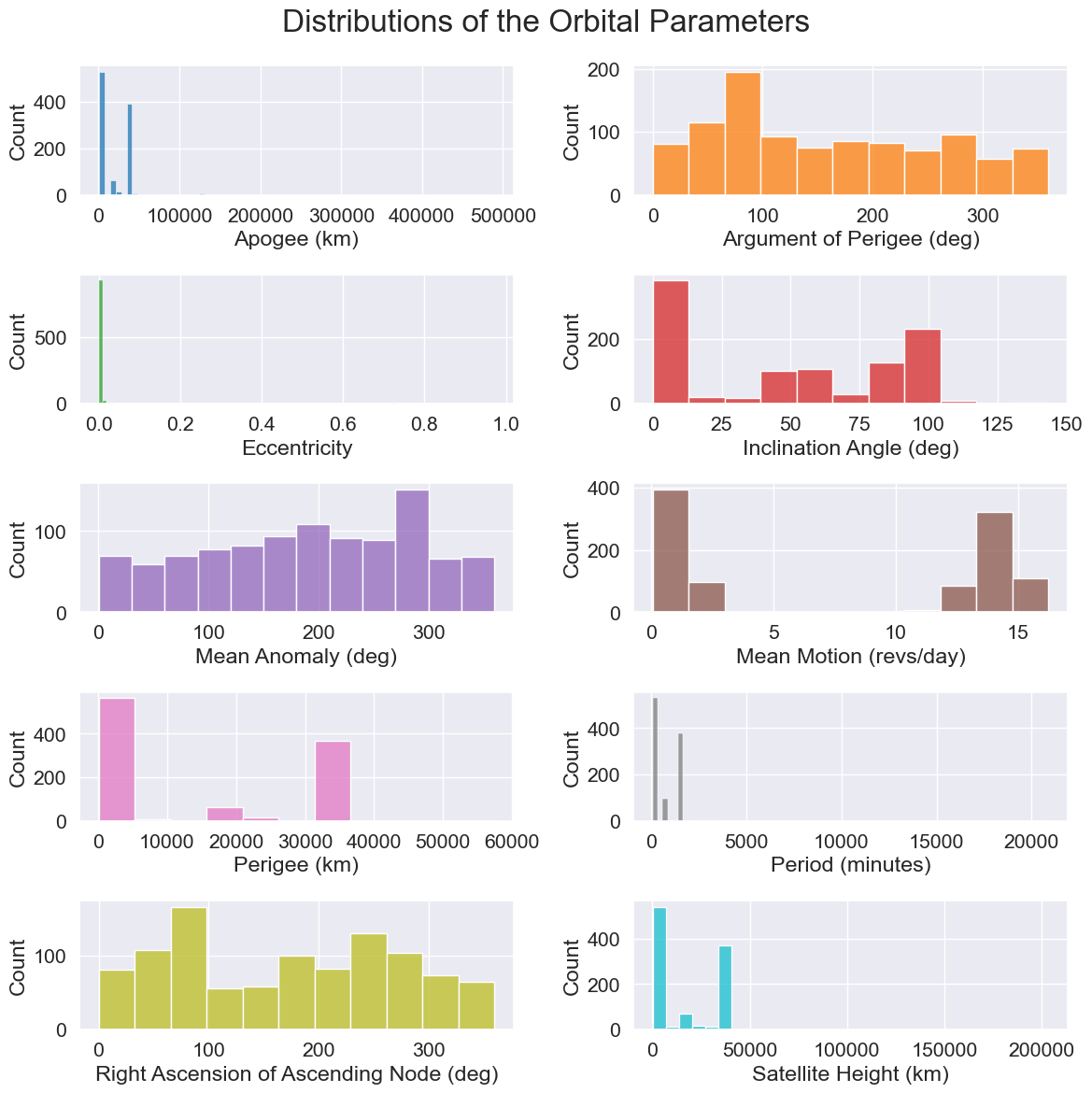


Fig. 1. Histograms for the orbital parameters.

The Orbit Type data field contained 16 different classes that included variations of LEO, GEO, MEO, and HEO satellite orbits. However, three orbit type classes contained over half the records and nine orbit type classes had five or fewer records. Due to concern of being able to develop a model that could accurately predict the detailed orbit types with so few data points, the data was grouped into five classes: the four main orbit types and an “Other” class for the rest. The “Other” class consisted of the following specific orbit types, with the number of records indicated in parenthesis: Deep Eccentric (11), Deep Space (2), and Non-GEO (1). The distribution of the data into the broader orbit type classes is shown in Fig. 2. The data is still not balanced as the LEO and GEO classes are much larger than the other classes.

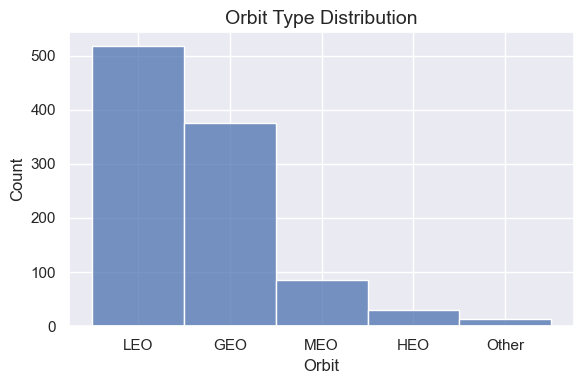


Fig. 2. Distribution of the Orbit Type classes.

With the orbit types grouped into the broader categories the mean motion and eccentricity should be main factors in distinguishing between the orbit types. Fig. 3 shows a scatterplot of the mean motion and eccentricity with the Orbit Type classes distinguished as separate colors. The LEO class has a distinguishable cluster with high mean motion and low eccentricity. The GEO class tends to have an eccentricity less than 0.3 and a mean motion of approximately one revolution per day. The MEO class also tends to have an eccentricity below 0.3, but the mean motion varies between around 2-12 rev/day. It is surprising to see a few data points belonging to the GEO and MEO classes that have eccentricities around 0.8.

The HEO class is distinguishable with eccentricity greater than 0.5 and mean motion around two revolution per day or higher. The Other class is predominately from the original Deep Eccentric orbit type, so it is no surprise that it tends to have high eccentricity and the lowest mean motion of all data points.

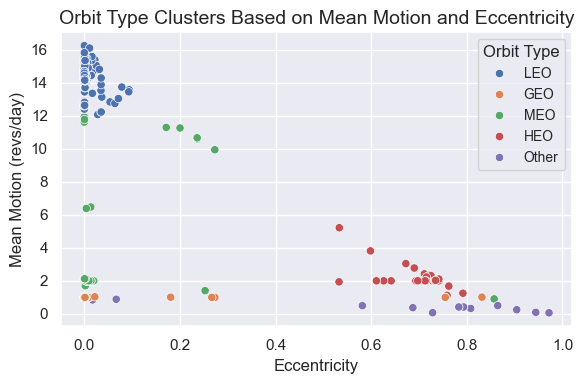


Fig. 3. Relationship between the mean motion and eccentricity for the various orbit types.

### II. Method

(use past tense for all, as if you have already done the work)

A few introductory sentences.

1. *Data Preparation*

Document the methods that you used, such as transformations or filtering. If applicable, include up to 2 before/after figures showing the results of your data preparation.

1. *Metrics*

This example is for classification: “In this analysis, the classification metrics of precision, recall and f1 are used to measure the performance of the classical and neural network models. In the dataset, the positive class is a food crisis. In the case of an actual food crisis, it is critical that a model prediction of no food crisis be avoided. As a result, recall is the most important metric, which helps to avoid false negatives. Precision was also selected, as accurate predictions of true positives are important, and f1 was chosen due to the unbalanced nature of the dataset where only 15.3% of labels are positive. These 3 metrics also facilitate a direct comparison to prior work. In certain cases, the area under the receiver operating characteristic curve (AUC) and accuracy are also presented to highlight differences between models.”

1. *Classical Modeling*

Justify your selection of modeling variations, and briefly describe the algorithms. Variations could include different algorithms, adding or removing input variables, or transformations.

Discuss the contributions of input variables to your model (z or p test).

Document how you monitor to prevent overfitting, such as a train/validate split, or cross-validation.

1. *Neural Network Modeling*

Justify your selection of optimization algorithm, loss function and regularization technique, based on the type of problem you are modeling (see concept map).

Document how you monitor to prevent overfitting, such as a train/validate split, or cross-validation. Justify your selection, based on your number of datapoints and number of input variables.

Discuss what hyperparameter sweeps you will perform, and what regularization technique you will explore.

### III. Analysis & Results

A few introductory sentences.

1. *Classical Modeling*

Document performance metrics for each modeling variation:

* R2, MSE and residual analysis for regression
* Accuracy, precision, recall and area under the ROC curve (AUC) for classification. For classification, if you have an unbalanced dataset it might be good to focus on the number of false negatives and false positives

Add a table that summarizes the metrics for your 3 variations.

* If regression, also calculate metrics that would result if your model always predicted the mean of the labels
* If classification, also calculate metrics that would result if your model always predicted “1” and if it predicted randomly (chance).

Justify your selection of the best model

Table 2. (example) Logistic Regression Model Results. (note to class: the P-value selection, RGE and Select K Best are just examples – you can choose the modeling variations that you try)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | P-val | | AUC | Precision | Recall | Accuracy |
| Full | | 0.00 | 0.58 | 0.73 | 0.18 | 0.86 |
| P-value Selection | | 0.00 | 0.58 | 0.81 | 0.16 | 0.86 |
| RFE | | 0.00 | 0.58 | 0.81 | 0.16 | 0.86 |
| Select K Best | | 0.00 | 0.59 | 0.75 | 0.20 | 0.87 |
| Chance | | -- | 0.50 | 0.15 | 0.50 | 0.50 |
| Always Predicts Crisis | | -- | -- | 0.15 | -- | 0.15 |
| Never Predicts Crisis | | -- | -- | -- | -- | 0.85 |
| Goal | | -- | -- | 0.85 | 0.75 | 0.90 |

1. *Neural Network Modeling*

Document model performance that results from a hyperparameter (HP) sweep on neurons per layer, number of layers and 2 other hyperparameters.

Investigate how 1 regularization technique affects your best model, and document the results.

Document performance metrics from your modeling similar to the classical section above. Include rows for:

* Preliminary NN model
* After HP sweep
* After regularization

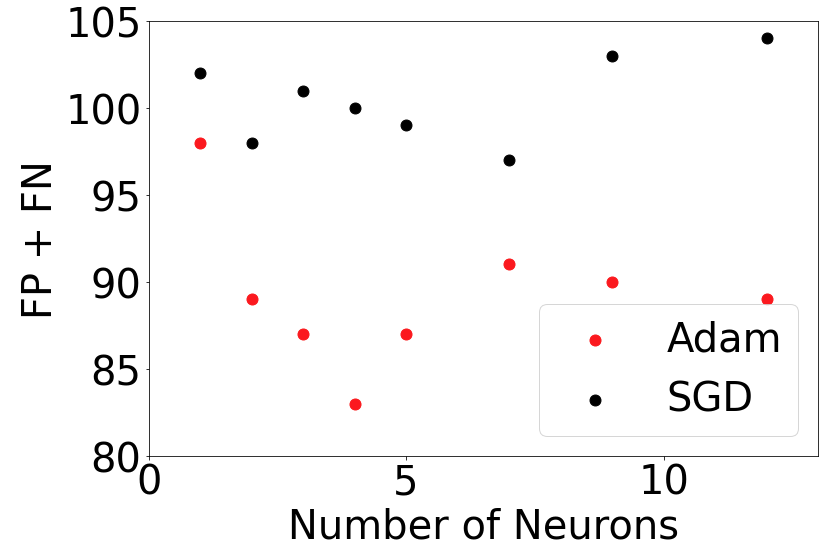


Figure 5. Model performance resulting from hyperparameter sweep.

1. *Model* *Evaluation*

Evaluate the results of your efforts using the criteria established in homework #4. Assess the value of this effort to your organization.

Discuss the final model selected, and why it is was selected.

Mention inferences that can be drawn from the modeling and from the data. Particularly insightful findings could be mentioned in your abstract and conclusion.

Include a figure that shows your model is not overfitting the data, or has an acceptable level of overfitting (<10%)

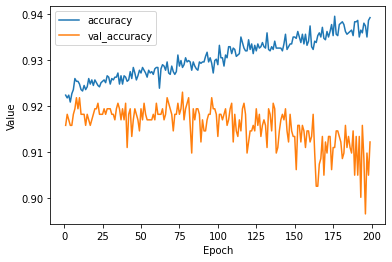


Figure 5. Neural network overfitting the dataset.

(student example) “With significant capacity built into the model, there is a significant divergence between the training accuracy and the validation accuracy. This model is clearly overfitting the data. A reduced capacity model was then implemented, and the number of epochs was reduced to 100. The results are shown in Figure 6.”

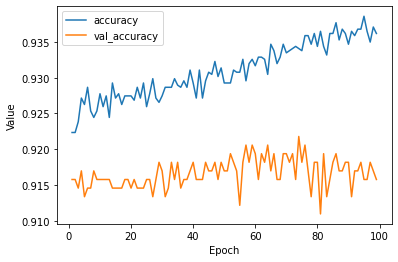


Figure 6. Overfitting corrected by XYZ.

1. *Model Application (a.k.a. deployment)*

Talk the audience through a simple application of your model.

How can it be used to further the mission of your organization? This could include:

* Qualitative applications:
  + mention how you could apply inferences from classical modeling - here is an example from an AF Drug Testing Lab student:
  + Inference “younger, less educated individuals who exhibit sensation seeking behavior and are open to experience tend to be at higher risk for THC use”
  + Application: we recommend a personality survey for new employees, and focus drug prevention efforts on those at highest risk for TCH use
* Quantitative applications:
  + how can you best use the prediction of the model to make a decision?
  + You could create an excel (or Python) analysis showing how a decision could be made
  + Be sure to mention the impact of that decision (i.e. we could save $500K, avoid 20% of space launch failures, or predeploy food aid to areas that have the highest risk for a food crisis

### IV. Conclusion

Restate main points, and avoid adding new information.

### References

Note to instructor: I am using the Word bibliography tool.

|  |  |
| --- | --- |
| [1] | U.S. Department of Defense, *DODI 8320.05 - Electromagnetic Spectrum Data Sharing,* Washington, DC: Chief Information Officer, 2017. |
| [2] | J. M. O'Hehir, *DSO Data Class: JSC Equipment, Tactical, Space (JETS),* 2023. |
| [3] | U.S. Department of Defense, *Joint Publication 3-14 - Joint Space Operations,* 2023. |
| [4] | G. Martin, C. J. Wetterer, J. Lau, J. Case, N. Toner, C. C. Chow and P. Dao, "Cislunar Periodic Orbit Family Classification from Astrometric and Photometric Observations Using Machine Learning," in *Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2020. |
| [5] | R. López, "Classify asteroid orbits using machine learning," Neural Designer, 31 August 2023. [Online]. Available: https://www.neuraldesigner.com/learning/examples/orbit-class/. [Accessed 6 November 2023]. |
| [6] | P. DiBona, J. Foster, A. Falcone and M. Czajkowski, "Machine learning for RSO maneuver classification and orbital patern prediction," in *Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2020. |
| [7] | T. G. Roberts and R. Linares, "Geosynchronous satellite maneuver classification via supervised machine learning," in *Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2021. |
| [8] | J. M. O'Hehir, Ed., *JETS Data Dictionary with codes and structure,* 2023. |

1. Prior to Oct 2023 PEO Spectrum was the Defense Spectrum Organization (DSO) [↑](#footnote-ref-1)
2. Accessible on NIPRNet at <https://gemsis.disa.mil>, access requires an account. [↑](#footnote-ref-2)