Space Missions ANN - Python & PHP

Majdi M. S. Awad
Back End Developer
Email: majdiawad.php@gmail.com | Mobile: +971 (055) 993 8785
Linkedin Account: https://www.linkedin.com/in/majdi-awad-aa2384317/
HackerRank Account: https://www.hackerrank.com/profile/majdiawad_php
Abu Dhabi, United Arab Emirates

Dataset

- The data set contains the information needed to create the predictive model.
- It comprises a data matrix.
- Its columns represent variables, and its rows represent samples.
- Variables in a data set can be of three types:
    - Inputs, which are the independent variables.
    - Targets, the dependent variables.
    - Unused variables, which will be neither targets nor inputs.
- Additionally, samples can be:
    - Training samples, used to build the model.
    - Selection samples, are used to select the model with the best generalization properties.
    - Testing samples, used to validate the accuracy of the model.
    - Unused samples, not used at all.

Data Preview Table

- The following table shows a preview of the data set obtained from the file space_missions.csv.
- Here, the number of columns is 12, and the number of samples is 150.
- Note that some columns are categorical.
- Therefore, the number of variables is greater than the number of columns, namely, 33.

	company	temperature_f	wind_speed_mph	humidity_pct	vehicle_type	liftoff_thrust_kn	payload_to_orbit_kg	rocket_height_m	...	miss
1	space_x	86	9	74	falcon	343	470	22.25	...	
2	space_x	NA	NA	NA	falcon	343	470	22.25	...	
...	...	...	...	...	...	...	...	...	...	...
150	boeing	50	0	100	delta	2533	2703	38.9	...	

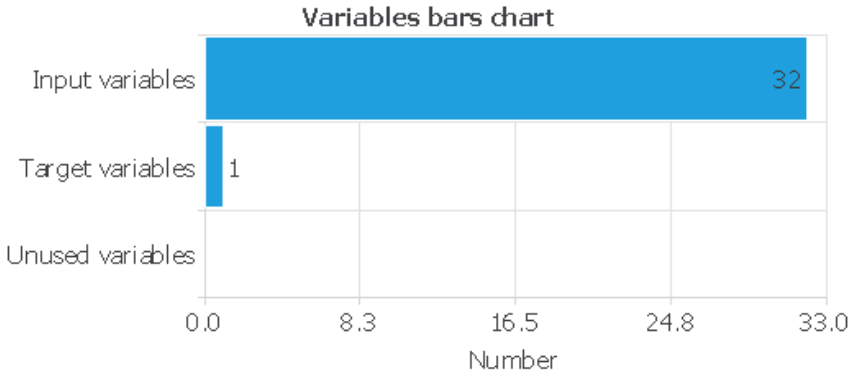
Columns Table

The following table depicts the names, types, and uses of all the columns in the data set. The data set has 32 inputs, 1 targets, and 0 unused variables.

	Name	Type	Use
1	company	Categorical	Input
2	temperature_f	Numeric	Input
3	wind_speed_mph	Numeric	Input
4	humidity_pct	Numeric	Input
5	vehicle_type	Categorical	Input
6	liftoff_thrust_kn	Numeric	Input
7	payload_to_orbit_kg	Numeric	Input
8	rocket_height_m	Numeric	Input
9	fairing_diameter_m	Numeric	Input
10	payload_mass_kg	Numeric	Input
11	payload_orbit	Categorical	Input
12	mission_status	Binary	Target

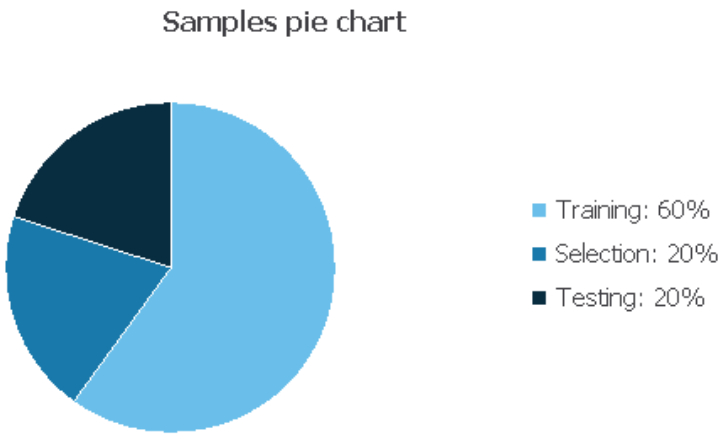
Variables Bars Chart

- The following chart illustrates the use of the variables.
- It depicts the numbers of inputs (32), targets (1), and unused variables (0).



Samples Pie Chart

- The following pie chart details the uses of all the samples in the data set.
- The total number of samples is 150.
- The number of training samples is 90 (60%), the number of selection samples is 30 (20%), and the number of testing samples is 30 (20%).



Missing Values Table

The number of missing values in the data set is 65. The missing values will be substituted by the mean of the corresponding variable when training. The following table lists the number of missing values for each variable.

	Missing values number
company	0
temperature_f	14
wind_speed_mph	14
humidity_pct	14
vehicle_type	0
liftoff_thrust_kn	0
payload_to_orbit_kg	0
rocket_height_m	0
fairing_diameter_m	4
payload_mass_kg	19
payload_orbit	0
mission_status	0

Data Statistics

- When designing a model, basic statistics are valuable since they might alert to spurious data.
- It is necessary to check for the correctness of every single variable's most critical statistical measures.

Data Statistics table

- The table below shows the minimums, maximums, means, and standard deviations of all the variables in the data set.
- Note that the following variables are categorical: company, vehicle_type, payload_orbit.
- As a consequence, no mean and standard deviation are shown for them.
- The minimum column shows the class of the categorical variable with the least number of appearances.
- In contrast, the maximum column shows the class of the categorical variable with the major number of appearances.

	Index	Minimum	Maximum	Mean	Deviation
company	-	-	-	-	-
temperature_f	2	37	93	70.4	12.7
wind_speed_mph	3	0	23	8.01	4.97
humidity_pct	4	37	100	76.1	15.7
vehicle_type	-	-	-	-	-
liftoff_thrust_kn	6	343	2.28e+04	5.67e+03	3.62e+03
payload_to_orbit_kg	7	380	6.38e+04	1.07e+04	9.5e+03
rocket_height_m	8	19	72	56.3	16.6
fairing_diameter_m	9	1	5.4	4.25	1.31
payload_mass_kg	10	0	1.56e+04	3.5e+03	3.39e+03
payload_orbit	-	-	-	-	-
mission_status	12	0	1	0.8	0.401

Positives () Statistics Table

The table below shows the minimums, maximums, means, and standard deviations of all the variables for a positive target corresponding to: .

	Minimum	Maximum	Mean	Deviation
company	martin_marietta	space_x	-	-
temperature_f	37	91	70	12.5
wind_speed_mph	0	23	7.74	4.95
humidity_pct	37	100	76	15.8
vehicle_type	titan	falcon	-	-
liftoff_thrust_kn	343	2.28e+04	6.1e+03	3.28e+03
payload_to_orbit_kg	470	6.38e+04	1.21e+04	9.64e+03
rocket_height_m	22.3	70	60.2	14.5
fairing_diameter_m	1.5	5.2	4.56	1.05
payload_mass_kg	0	1.56e+04	3.82e+03	3.54e+03
payload_orbit	high_earth_orbit	low_earth_orbit	-	-

Nigatives () Statistics Table

The table below shows the minimums, maximums, means, and standard deviations of all the variables for a negative target corresponding to:

	Minimum	Maximum	Mean	Deviation
company	martin_marietta	space_x	-	-
temperature_f	50	93	73.2	13.7
wind_speed_mph	0	20	9.68	4.9
humidity_pct	42	100	76.4	15.8
vehicle_type	vega	falcon	-	-
liftoff_thrust_kn	343	1.51e+04	3.93e+03	4.38e+03
payload_to_orbit_kg	380	2.84e+04	5.28e+03	6.65e+03
rocket_height_m	19	72	41.1	15.7
fairing_diameter_m	1	5.4	2.85	1.47
payload_mass_kg	0	5.5e+03	1.78e+03	1.72e+03
payload_orbit	sun_earth_orbit	low_earth_orbit	-	-

Scatter Charts

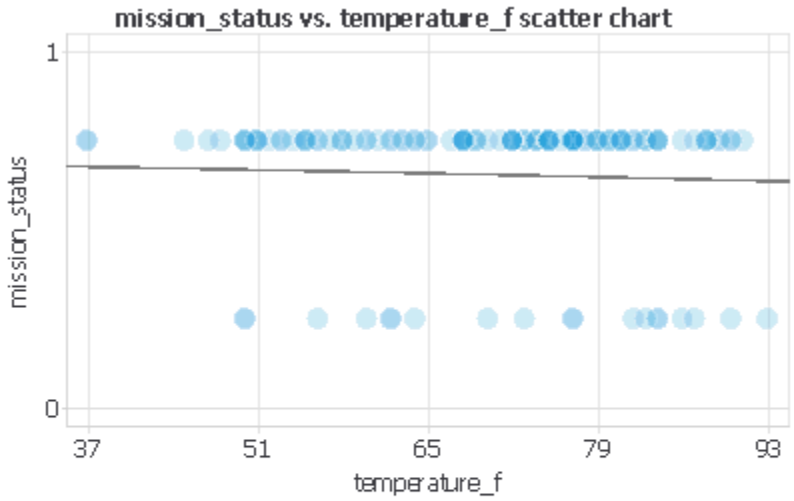
- This task plots all target variables versus all input variables in the data set.
- These charts might help to see the dependencies of the targets with the inputs.

mission_status VS company scatter charts

Input column is categorical and target column is binary. Therefore, no scatter chart is shown here.

mission_status VS temperature_f scatter charts

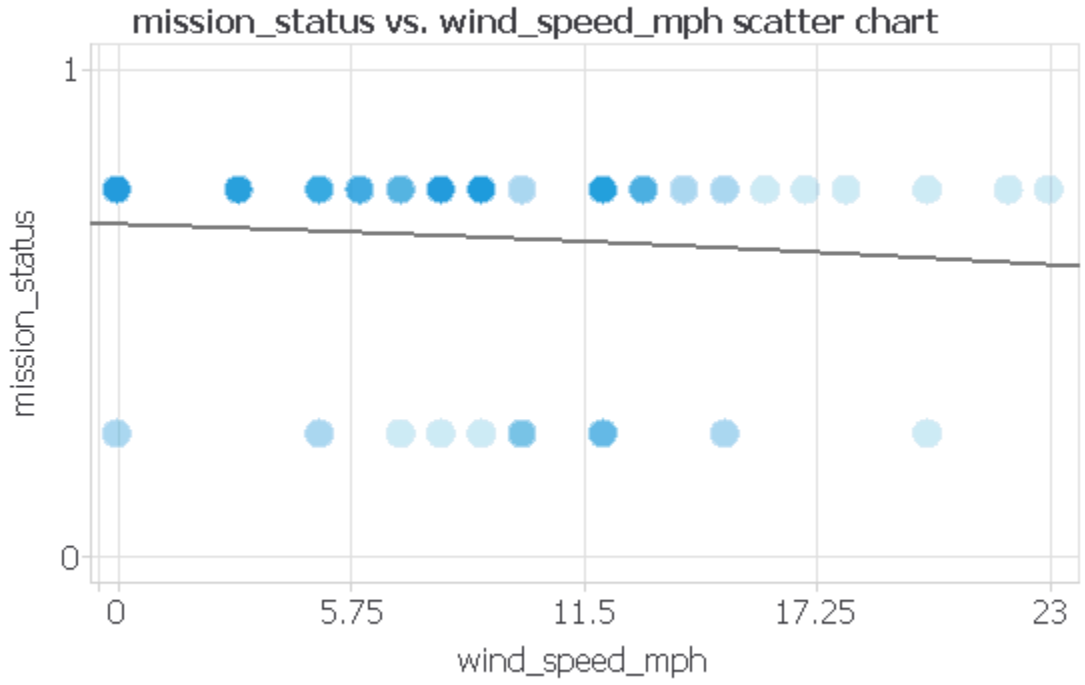
The following chart shows the input temperature_f and target mission_status scatter plot.



It also shows the regression line between both variables, of type logistic, with a correlation value of 0.091.

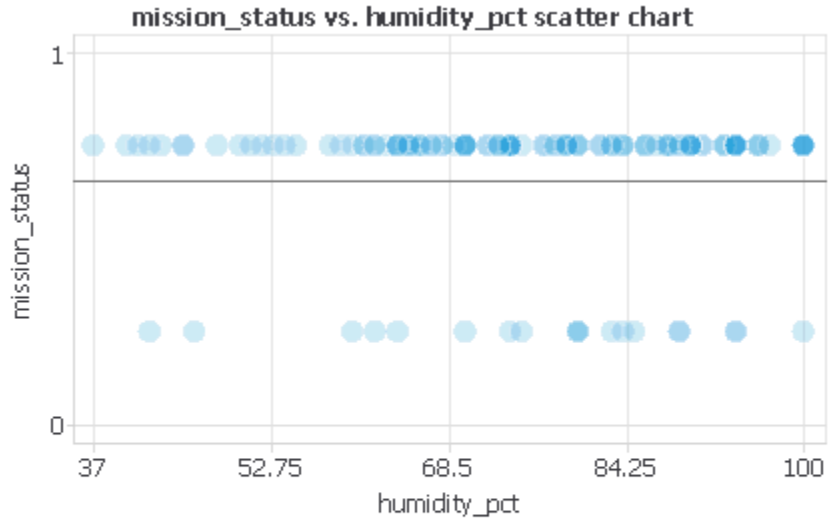
mission_status VS wind_speed_mph scatter charts

The following chart shows the input wind_speed_mph and target mission_status scatter plot.



mission_status VS humidity_pct scatter charts

The following chart shows the input humidity_pct and target mission_status scatter plot.



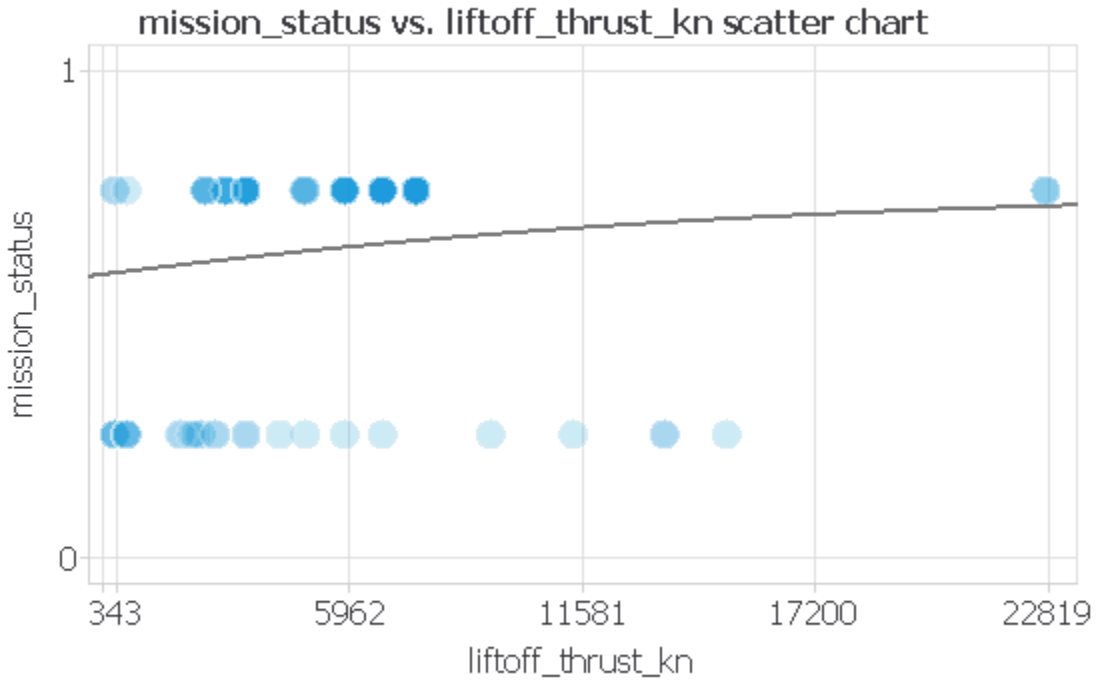
It also shows the regression line between both variables, of type logistic, with a correlation value of 0.008.

mission_status VS vehicle_typr scatter charts

Input column is categorical and target column is binary. Therefore, no scatter chart is shown here.

mission_status VS liftoff_thrust_kn scatter charts

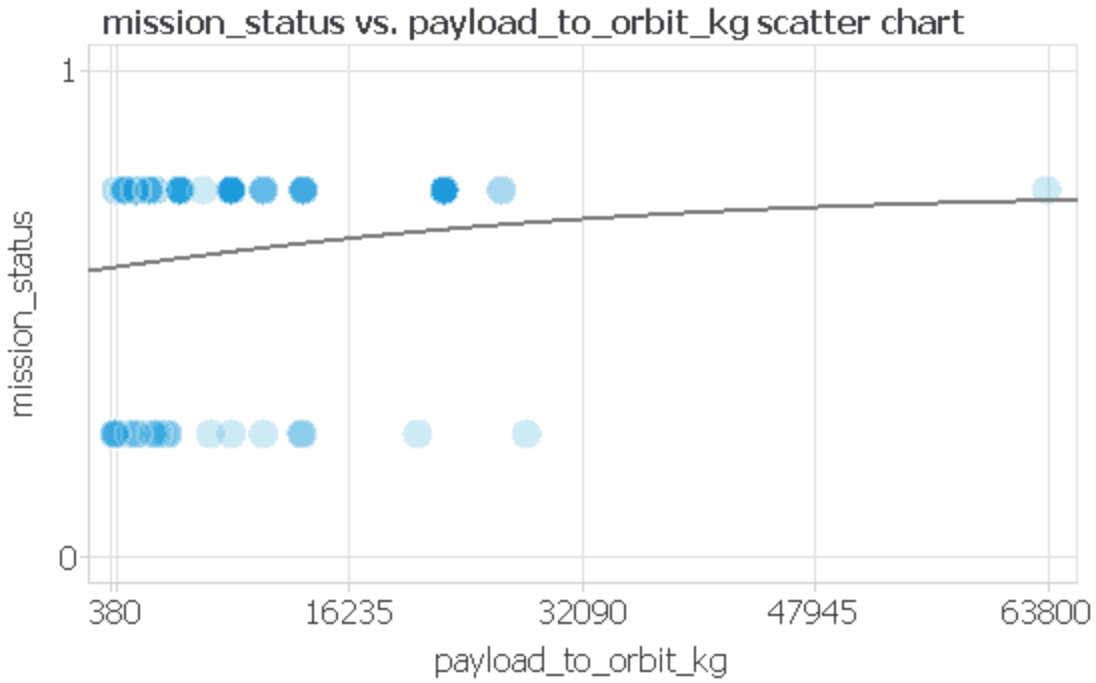
The following chart shows the input liftoff_thrust_kn and target mission_status scatter plot.



It also shows the regression line between both variables, of type logistic, with a correlation value of 0.322.

mission_status VS payload_to_orbit_kg scatter charts

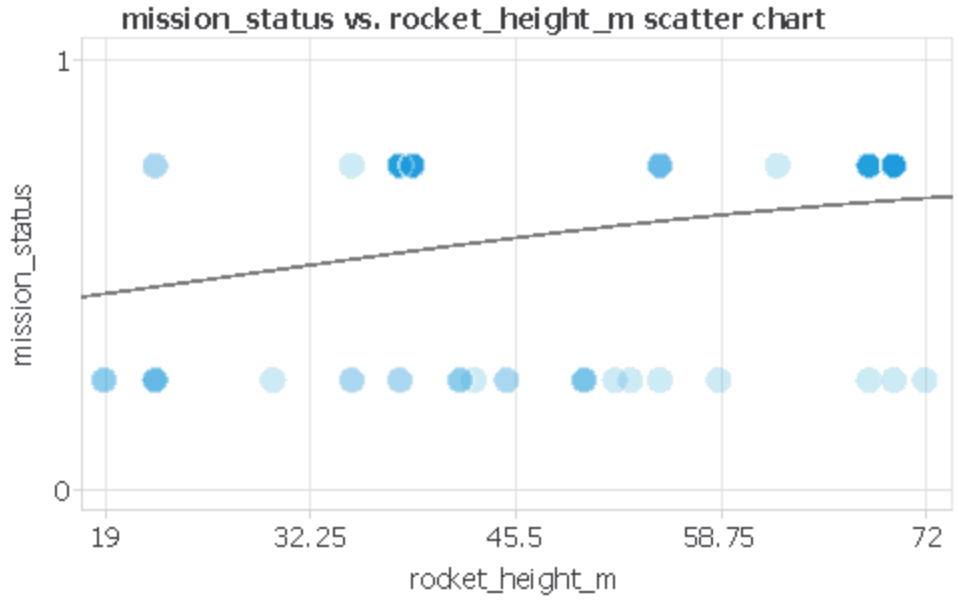
The following chart shows the input payload_to_orbit_kg and target mission_status scatter plot.



It also shows the regression line between both variables, of type logistic, with a correlation value of 0.315.

mission_status VS rocket_height_m scatter charts

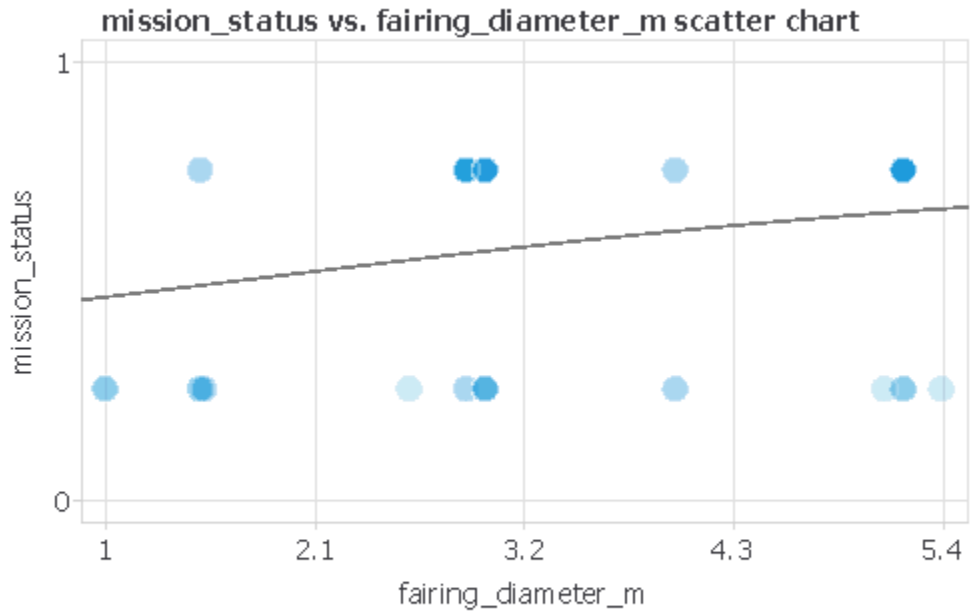
The following chart shows the input rocket_height_m and target mission_status scatter plot.



It also shows the regression line between both variables, of type logistic, with a correlation value of 0.459.

mission_status VS fairing_diameter_m scatter charts

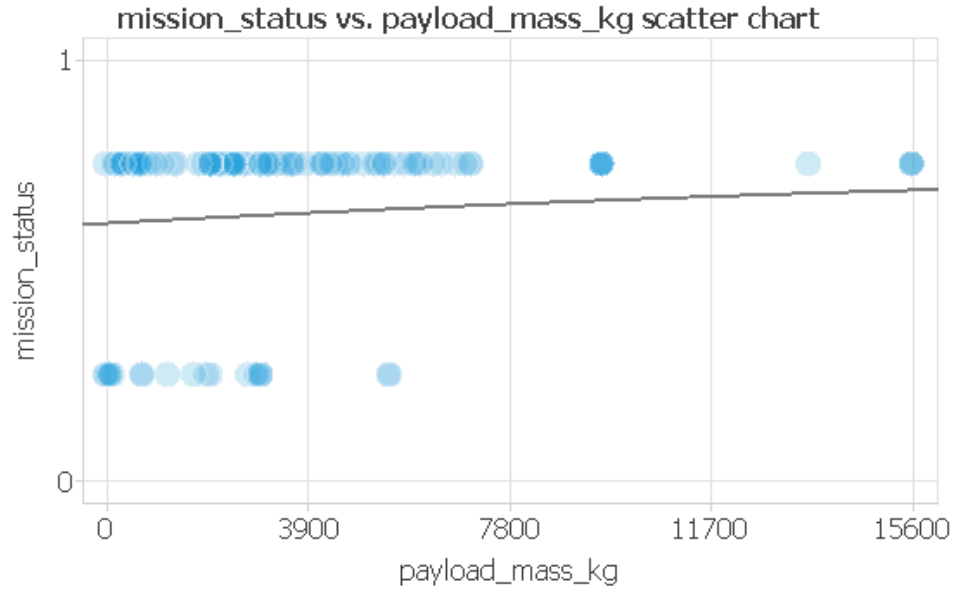
The following chart shows the input fairing_diameter_m and target mission_status scatter plot.



It also shows the regression line between both variables, of type logistic, with a correlation value of 0.515.

mission_status VS payload_mass_kg scatter charts

The following chart shows the input payload_mass_kg and target mission_status scatter plot.



It also shows the regression line between both variables, of type logistic, with a correlation value of 0.233.

mission_status VS payload_orbit scatter charts

Input column is categorical and target column is binary. Therefore, no scatter chart is shown here.

Inputs Correlations

- This task calculates the absolute values of the correlations among all inputs.
- These correlations can be of different types depending on the nature of the variables: linear, exponential, power, logarithmic, or logistic.
- The correlation is a numeric value between 0 and 1 that expresses the strength of the relationship between two input variables in the data set.
- When it is close to 1, it indicates a strong relationship and a value close to 0 indicates no relationship.

Maximal Inputs Correlations

The following table shows the value of the top 10 most significant input correlations.

	Input	Input	Correlation value
1	company	vehicle_type	0.971
2	rocket_height_m	fairing_diameter_m	0.949
3	payload_to_orbit_kg	rocket_height_m	0.873
4	liftoff_thrust_kn	rocket_height_m	0.869
5	payload_to_orbit_kg	fairing_diameter_m	0.869
6	liftoff_thrust_kn	fairing_diameter_m	0.829
7	company	liftoff_thrust_kn	0.822
8	liftoff_thrust_kn	payload_to_orbit_kg	0.822
9	company	rocket_height_m	0.809
10	vehicle_type	rocket_height_m	0.798

As we can see in the previous table, the three most relevant correlations of this model are the following:

- The correlation between the input company and vehicle_type is: 0.971.
- The correlation between the input rocket_height_m and fairing_diameter_m is: 0.949.
- The correlation between the input payload_to_orbit_kg and rocket_height_m is: 0.873.

The following table shows the value of the correlations among all input variables.

	company	temperature_f	wind_speed_mph	humidity_pct	vehicle_type	liftoff_thrust_kn	payload_to_orbit_kg	rocket_height_m	fairing_diameter_m	payload_mass_kg	payload_orbit
company	1	0.653	0.651	0.654	0.971	0.822	0.731	0.809	0.72	0.695	0.458
temperature_f	0.653	1	0.185	-0.078	0.64	0.074	-0.081	-0.092	-0.044	-0.061	0.455
wind_speed_mph	0.651	0.185	1	-0.518	0.642	-0.06	0.033	0.046	-0.069	0.051	0.441
humidity_pct	0.654	-0.078	-0.518	1	0.643	0.087	-0.082	0.016	0.027	-0.047	0.451
vehicle_type	0.971	0.64	0.642	0.643	1	0.694	0.652	0.798	0.781	0.707	0.458
liftoff_thrust_kn	0.822	0.074	-0.06	0.087	0.694	1	0.822	0.869	0.829	0.458	0.471
payload_to_orbit_kg	0.731	-0.081	0.033	-0.082	0.652	0.822	1	0.873	0.869	0.506	0.456
rocket_height_m	0.809	-0.092	0.046	0.016	0.798	0.869	0.873	1	0.949	0.539	0.494
fairing_diameter_m	0.72	-0.044	-0.069	0.027	0.781	0.829	0.869	0.949	1	0.496	0.483
payload_mass_kg	0.695	-0.061	0.051	-0.047	0.707	0.458	0.506	0.539	0.496	1	0.546
payload_orbit	0.458	0.455	0.441	0.451	0.458	0.471	0.456	0.494	0.483	0.546	1

The following table shows the value of the top 10 most significant Spearman input correlations.

	Input	Input	Correlation value
1	company	vehicle_type	0.972
2	rocket_height_m	fairing_diameter_m	0.866
3	liftoff_thrust_kn	rocket_height_m	0.864
4	payload_to_orbit_kg	rocket_height_m	0.844
5	payload_to_orbit_kg	fairing_diameter_m	0.829
6	liftoff_thrust_kn	payload_to_orbit_kg	0.829
7	company	fairing_diameter_m	0.828
8	vehicle_type	fairing_diameter_m	0.821
9	company	rocket_height_m	0.802
10	liftoff_thrust_kn	fairing_diameter_m	0.792

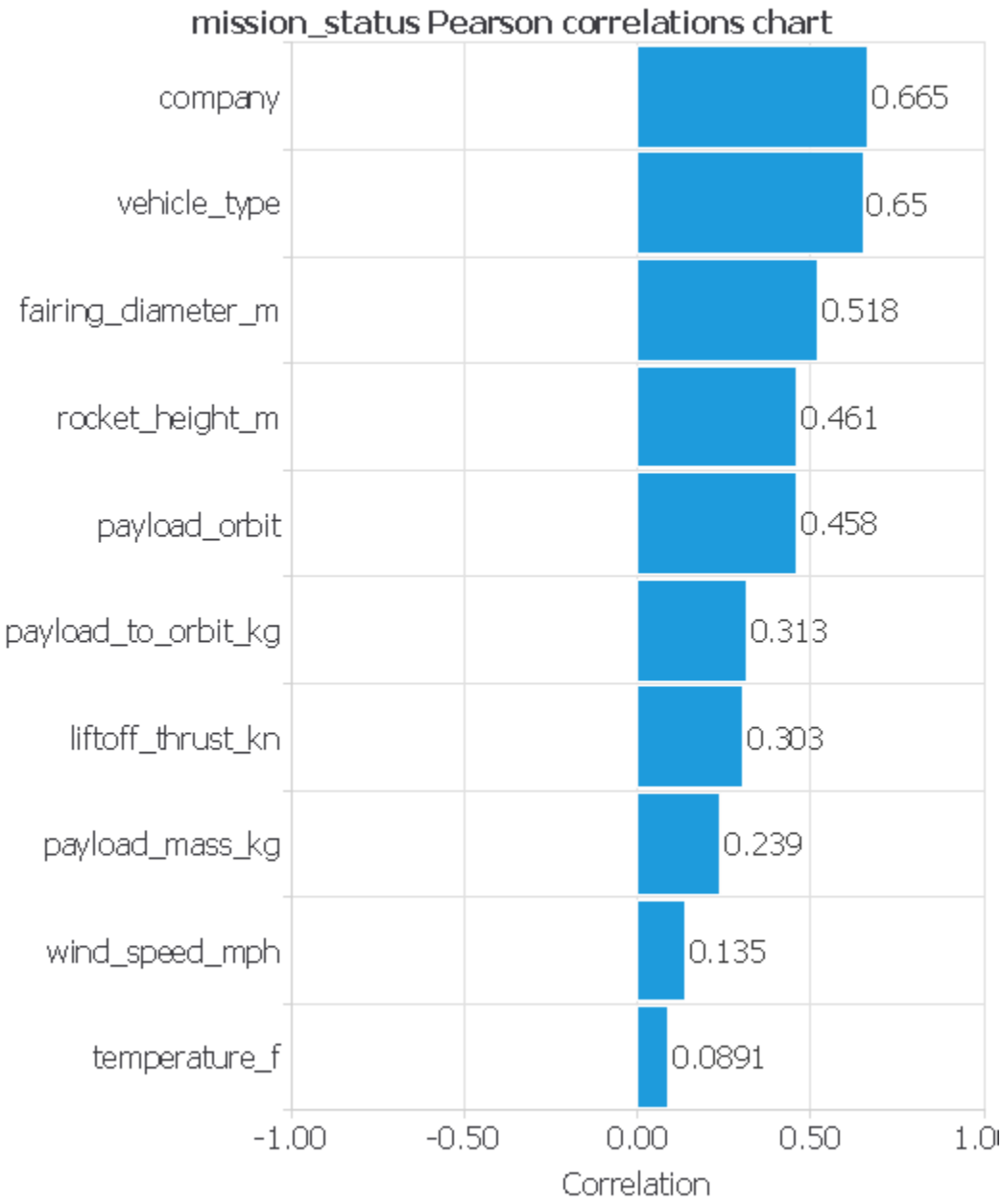
As we can see in the previous table, the three most relevant correlations of this model are the following:

- The correlation between the input company and vehicle_type is: 0.972.
- The correlation between the input rocket_height_m and fairing_diameter_m is: 0.866.
- The correlation between the input liftoff_thrust_kn and rocket_height_m is: 0.864.

The following table shows the value of the spearman correlations among all input variables.

	company	temperature_f	wind_speed_mph	humidity_pct	vehicle_type	liftoff_thrust_kn	payload_to_orbit_kg	rocket_height_m	fairing_diameter_m	payload_mass_kg	payload_orbit
company	1	0.65	0.654	0.651	0.972	0.621	0.754	0.802	0.828	0.703	0.458
temperature_f	0.65	1	0.231	-0.12	0.641	0.019	-0.051	-0.057	-0.027	-0.033	0.456
wind_speed_mph	0.654	0.231	1	-0.54	0.634	-0.046	-0.024	0.024	-0.095	0.072	0.441
humidity_pct	0.651	-0.12	-0.54	1	0.643	0.013	-0.057	0.013	0.067	-0.066	0.445
vehicle_type	0.972	0.641	0.634	0.643	1	0.693	0.727	0.781	0.821	0.734	0.458
liftoff_thrust_kn	0.621	0.019	-0.046	0.013	0.693	1	0.829	0.864	0.792	0.673	0.47
payload_to_orbit_kg	0.754	-0.051	-0.024	-0.057	0.727	0.829	1	0.844	0.829	0.546	0.455
rocket_height_m	0.802	-0.057	0.024	0.013	0.781	0.864	0.844	1	0.866	0.704	0.48
fairing_diameter_m	0.828	-0.027	-0.095	0.067	0.821	0.792	0.829	0.866	1	0.633	0.486
payload_mass_kg	0.703	-0.033	0.072	-0.066	0.734	0.673	0.546	0.704	0.633	1	0.579
payload_orbit	0.458	0.456	0.441	0.445	0.458	0.47	0.455	0.48	0.486	0.579	1

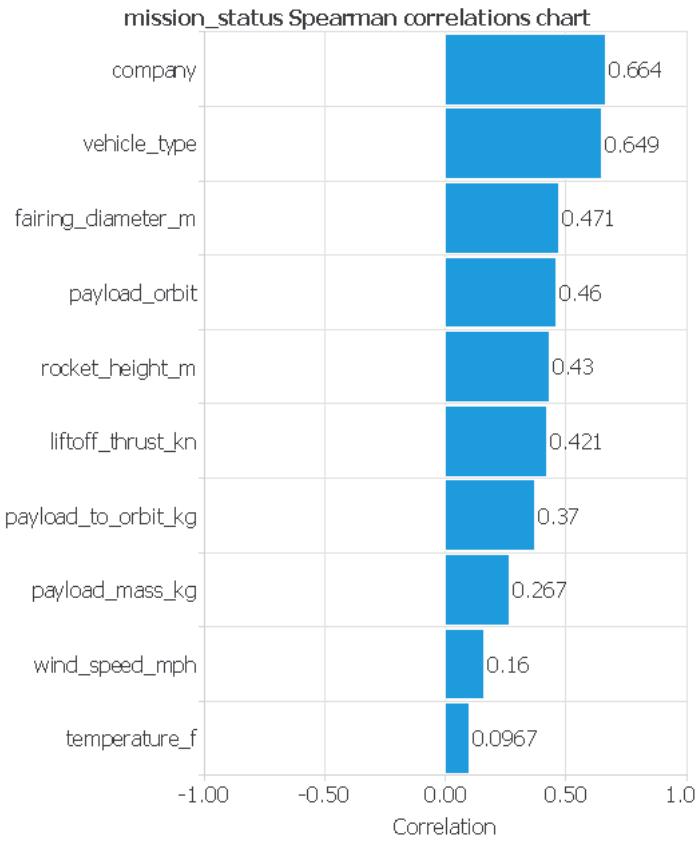
The following chart illustrates the target 'mission_status' dependency with the 10 input columns with greatest correlation in the data set.



The following table shows the value of all the Pearson correlations between input and target variables and the 95% confidence intervals.

	Index	Type	mission_status	Lower CI limit	Upper CI limit
company	1	logistic	0.664861	0.564778	0.745670
vehicle_type	5	logistic	0.650379	0.547147	0.734122
fairing_diameter_m	9	logistic	0.517702	0.387874	0.627387
rocket_height_m	8	logistic	0.460513	0.324177	0.578109
payload_orbit	11	logistic	0.457623	0.320895	0.575665
payload_to_orbit_kg	7	logistic	0.313495	0.161339	0.451093
liftoff_thrust_kn	6	logistic	0.302715	0.149717	0.441555
payload_mass_kg	10	logistic	0.238828	0.070177	0.394205
wind_speed_mph	3	logistic	0.134976	-0.034132	0.296571
temperature_f	2	logistic	0.089141	-0.080398	0.253668
humidity_pct	4	logistic	0.007747	-0.160795	0.175851

The following chart illustrates the target 'mission_status' dependency with the 10 input columns with greatest correlation in the data set.



The following table shows the value of all the Spearman correlations between input and target variables and the 95% confidence intervals.

	Index	Type	mission_status	Lower CI limit	Upper CI limit
company	1	logistic	0.663916	0.563624	0.744918
vehicle_type	5	logistic	0.648871	0.545316	0.732917
fairing_diameter_m	9	logistic	0.470978	0.334091	0.588407
payload_orbit	11	logistic	0.460025	0.323622	0.577696
rocket_height_m	8	logistic	0.429680	0.289343	0.551935
liftoff_thrust_kn	6	logistic	0.420954	0.279552	0.544483
payload_to_orbit_kg	7	logistic	0.369750	0.222684	0.500362
payload_mass_kg	10	logistic	0.266635	0.099668	0.418998
wind_speed_mph	3	logistic	0.160094	-0.008467	0.319808
temperature_f	2	logistic	0.096687	-0.072831	0.260776
humidity_pct	4	logistic	0.014119	-0.154582	0.182019

Uncorrelated Variables

- The less correlated variables with the targets can confuse the model with data that could not influence the predicted variable.
- Therefore, this task removes the inputs with a lower correlation than a given value.

Uncorrelated Columns table

- The number of uncorrelated variables in the data set is 4.
- Neural Designer set those variables to 'Unused'.
- The following table shows the uncorrelated columns in the data set.

	Variable name
1	temperature_f
2	wind_speed_mph
3	humidity_pct
4	payload_mass_kg

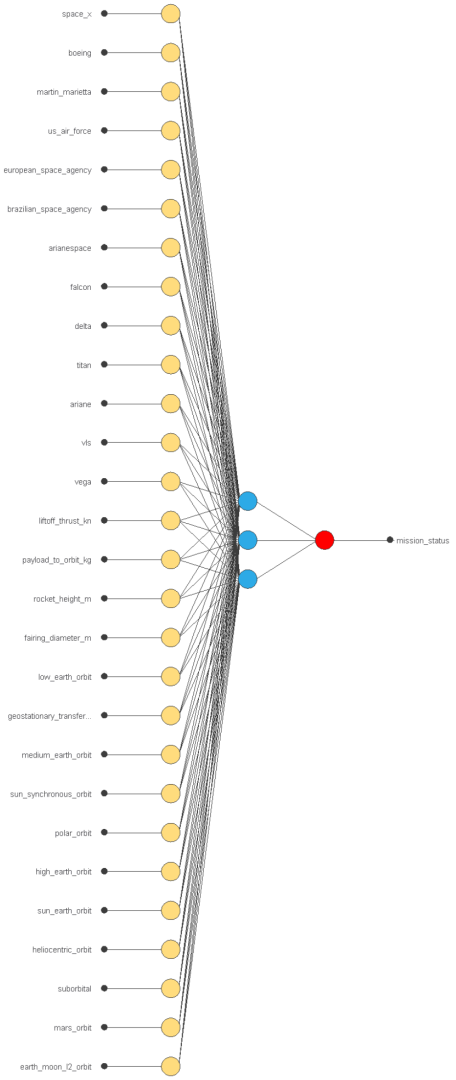
The Neural Network

- The neural network represents the predictive model.
- In Neural Designer, neural networks allow deep architectures, a class of universal approximators.

Network Architecture

The next figure depicts a graphical representation of the network architecture.
It contains the following layers:

- Scaling layer with 28 neurons (yellow).
- Perceptron layer with 3 neurons (blue).
- Probabilistic layer with 1 neurons (red).



As shown in the figure above, there are 28 inputs (columns used) to the neural network. This neural network has 1 potential outputs:

- Mission_status

Inputs

- The number of inputs is 28.
- The following table depicts the names of the inputs to the neural network.

	Name
1	space_x
2	boeing
3	martin_marietta
4	us_air_force
5	european_space_agency
6	brazilian_space_agency
7	arianespace
8	falcon
9	delta
10	titan
11	ariane
12	vls
13	vega
14	liftoff_thrust_kn
15	payload_to_orbit_kg
16	rocket_height_m
17	fairing_diameter_m
18	low_earth_orbit
19	geostationary_transfer_orbit
20	medium_earth_orbit
21	sun_synchronous_orbit
22	polar_orbit
23	high_earth_orbit
24	sun_earth_orbit
25	heliocentric_orbit
26	suborbital
27	mars_orbit
28	earth_moon_l2_orbit

Scaling Layer

- The size of the scaling layer is 28, the number of inputs.
- The following table shows the values used for scaling the inputs.
- They include the minimum, maximum, mean, and standard deviation.

	Minimum	Maximum	Mean	Deviation	Scaler
space_x	0	1	0.633	0.484	MeanStandardDeviation
boeing	0	1	0.252	0.435	MeanStandardDeviation
martin_marietta	0	1	0.007	0.082	MeanStandardDeviation
us_air_force	0	1	0.048	0.214	MeanStandardDeviation
european_space_agency	0	1	0.014	0.116	MeanStandardDeviation
brazilian_space_agency	0	1	0.02	0.142	MeanStandardDeviation
arianespace	0	1	0.027	0.163	MeanStandardDeviation
falcon	0	1	0.633	0.484	MeanStandardDeviation
delta	0	1	0.252	0.435	MeanStandardDeviation
titan	0	1	0.054	0.228	MeanStandardDeviation
ariane	0	1	0.034	0.182	MeanStandardDeviation
vls	0	1	0.02	0.142	MeanStandardDeviation
vega	0	1	0.007	0.082	MeanStandardDeviation
liftoff_thrust_kn	343	15120	5318.35	2683.75	MeanStandardDeviation
payload_to_orbit_kg	380	28370	10130.2	8307.18	MeanStandardDeviation
rocket_height_m	19	72	56.058	16.63	MeanStandardDeviation
fairing_diameter_m	1	5.4	4.233	1.297	MeanStandardDeviation
low_earth_orbit	0	1	0.429	0.497	MeanStandardDeviation
geostationary_transfer_orbit	0	1	0.299	0.46	MeanStandardDeviation
medium_earth_orbit	0	1	0.054	0.228	MeanStandardDeviation
sun_synchronous_orbit	0	1	0.061	0.241	MeanStandardDeviation
polar_orbit	0	1	0.082	0.275	MeanStandardDeviation
high_earth_orbit	0	1	0.014	0.116	MeanStandardDeviation
sun_earth_orbit	0	1	0.007	0.082	MeanStandardDeviation
heliocentric_orbit	0	1	0.034	0.182	MeanStandardDeviation
suborbital	0	1	0.007	0.082	MeanStandardDeviation
mars_orbit	0	1	0.007	0.082	MeanStandardDeviation
earth_moon_l2_orbit	0	1	0.007	0.082	MeanStandardDeviation

Python Code:

In the development of the neural network model, I undertook the task of designing and implementing a computational framework to predict outputs based on various input parameters relevant to space exploration and rocket technology. The model, instantiated as the NeuralNetwork class, incorporates a detailed input layer with twenty-eight features, encompassing parameters such as payload capacity, rocket height, and various orbital classifications. Each input is scaled appropriately to normalize the data, facilitating more accurate predictions. The core of the model consists of a single layer of perceptrons where I utilized the hyperbolic tangent function to compute the activation of each neuron. The weights and biases were meticulously calibrated to reflect the influence of each input feature on the final output. The calculate_outputs method executes the forward pass, applying the scaled inputs to the perceptrons and generating predictions. The main function serves as the user interface, prompting for input values and displaying the resultant predictions, thereby demonstrating the practical application of the neural network in a real-world context. Through this implementation, I have effectively translated theoretical neural network concepts into a tangible predictive tool.

Python code is available on my github account with MIT License:
https://github.com/majdi-php-sql/Space_Mission_ANN_Python

PHP Code:

I designed and developed this Rest API client-side demo using HTML, CSS, PHP, and JavaScript to demonstrate a rocket simulation prediction model. The interface leverages Bootstrap for a clean, responsive design, ensuring user-friendliness. The backend PHP script processes query parameters to validate input and execute a multilayer perceptron model, implemented as a series of scaled input transformations and hyperbolic tangent (tanh) activation functions. This script checks if the parameters are numeric, scales them appropriately, and calculates the perceptron outputs through successive layers. The final perceptron output is then rescaled to provide a prediction value between 0 and 1, which is returned as a JSON object along with a validation status message. The page provides instructions and feedback to the user, ensuring clarity and ease of use. I designed and developed the PHP backend for this rocket simulation prediction model, ensuring that it efficiently handles user inputs and performs necessary calculations. The PHP code processes the form data, interacts with the database to fetch or store information, and ensures secure data handling. By leveraging PHP's server-side capabilities, I was able to implement robust logic that supports the dynamic nature of the simulation, delivering real-time results and providing a seamless user experience. This integration of PHP is crucial for the application's functionality, as it bridges the client-side interface with the server-side processing, making the entire system cohesive and responsive.

PHP & JS code is available on my github account with MIT License:
https://github.com/majdi-php-sql/Space_Mission_ANN_PHP