Intelligent analysis of the causes of the Challenger space shuttle disaster

Igor Petrov[4], Vladislav Mykhailenko[1], Roman Kharchenko[1], Yurii Gunchenko[2], Aleksandr Kochetkov[3] and Oksana Zui[2]

[1] National University "Odessa Maritime Academy", Didrikhson str.8, Odessa, 65029, Ukraine
[2] I.I. Mechnikov National University, Dvorianska Str., 2, Odessa, 65025, Ukraine
[3] Odessa National Maritime University, Mechnikova str.8, Odessa, 65029 Ukraine
[4] National University "Odessa Maritime Academy", Didrichson street 8, Odessa, 65029, Ukraine

Abstract

The publication presents an intelligent analysis of data indicating the causes of the Challenger space shuttle disaster. Data analysis was performed using a specialized Orange software package. Statistical indicators of failure of the fuel tank sealing rings, as well as the corresponding air temperature on the day of the spacecraft launch, were used as initial data. Data analysis using logistic regression indicates that at a temperature of 36 degrees Celsius there is an 89% chance of O-ring failure, which indicates a significant risk of defects. Overall, the average probability of failure for all five rings of the spacecraft was 56%, which confirmed the conclusions of the investigative commission about the causes of the disaster. The research methods used were approaches of regression analysis of data and the method of clustering data K – average. The degree of adequacy of the mathematical model implementing the method of logistic regression was 74%, which indicates an acceptable result of the obtained data. The proposed model allows predicting with a high degree of accuracy the possible failure of fuel tank elements of modern spacecraft.

Keywords

Intelligent Analysis, Space Shuttle Challenger, O-Ring, Logistic Regression, k-means method

1. Introduction

The prototype of the Starship spacecraft, developed by Elon Musk's SpaceX for flights to Mars and the Moon, landed for the first time during flight tests on March 3. As soon as the applause died down - it was the first successful landing - the device exploded. Questions for Musk only increased: why continue tests that always end in an explosion on the ground or burning up in the air? Meanwhile, another Starship prototype is scheduled to launch in the coming days, and on March 18, a critical test of its main competitor, the super-heavy rocket Space Launch System of Boeing, was carried out. After several decades, when super-heavy rockets were not developed because there were no suitable tasks for them, they are in demand again - a new technological competition for the creation of space rockets has begun. The start of Starship testing can be counted from the first successful flights of the very first demonstrator with a single engine, Starhopper: on April 5, 2019, it jumped one meter, then made two more jumps of 20 and 150 meters, and now decorates the test site, forming a backdrop for testing new rocket prototypes.

Their fate turned out to be more difficult. The next Starship prototypes — Mk1 (the same one against the backdrop of which Elon Musk held his historic press conference on September 28, 2019), SN1 and SN3 exploded during pressure tests of the tanks. Each such accident led to a change in the design of the rocket and launch facilities.

Already in August and September 2020, the flight demonstrators SN5 and SN6 made their first successful ascents on one engine to 150 meters and landed softly without any emergency situations. On February 2, 2021, the next flight demonstrator, SN9, was launched. It routinely ascended to an altitude of 10 kilometers and descended from it, but during the landing approach itself, one of the two required engines did not turn on. As a result, the prototype crashed into the ground again with a powerful explosion[1].

[1]ICST-2024: Information Control Systems & Technologies, September , 23 – 25, 2024, Odesa, Ukraine
firmn@gmail.com (I. Petrov); vlamihailenod@gmail.com (V. Mykhailenko); romannn30@gmail.com (R. Kharchenko); gunchenko@onu.edu.ua (Y. Gunchenko); 0679016767@ukr.net (A. Kochetkov); oks.zuj@gmail.com (O. Zui)
0000-0002-8740-6198 (I. Petrov); 0000-0003-2793-8966 (V. Mykhailenko); 0000-0003-3051-7513 (R. Kharchenko); 0000-0003-4423-8267 (Y. Gunchenko); 0000-0001-8200-4897 (A. Kochetkov); 0000-0001-9520-4441 (O. Zui)

Thus, it can be stated that the study of the causes of spacecraft accidents is a relevant task. And to understand the cause-and-effect relationships of the accidents that have occurred, it is necessary, according to the authors, to analyze in detail the history of unsuccessful spacecraft launches using the Shuttle as an example.

The Challenger disaster occurred on January 28, 1986, when the shuttle was destroyed at the very beginning of the mission as a result of a fuel tank explosion 73 seconds into the flight, which led to the death of all 7 crew members (Fig.1).

The disaster occurred over the Atlantic Ocean off the coast of the central part of the Florida Peninsula, USA. According to the conclusions of the commission to investigate the causes of the disaster, the destruction of the spacecraft was caused by damage to the sealing ring of the right solid rocket booster during launch. Damage to the ring caused a hole to burn out in the side of the accelerator, from which a jet stream flowed towards the external fuel tank. This led to the destruction of the tail mount of the right solid rocket booster and the supporting structures of the external fuel tank. Elements of the complex began to shift relative to each other. The destruction of the external fuel tank led to the ignition of fuel components and the explosion of the ship [2].

Figure 1: Photo of the Shuttle Challenger before launch [1]

Taking into account the global nature of the disaster, the task of intellectual analysis of data on the cause-and-effect factors that led to the disaster is of scientific interest. It is known that Data mining (Russian: data mining, data mining, in-depth data analysis) is a collective name used to denote a set of methods for detecting previously unknown, non-trivial, practically useful and accessible interpretation of knowledge in data necessary for decision-making in various fields human activity. Data mining methods can be used both for working with large data and for processing relatively small amounts of data (obtained, for example, from the results of individual experiments on the example of the failure of the sealing rings of a spacecraft [2].

2. Analysis of the design features of the spacecraft

To launch each shuttle, two solid fuel accelerators were used, consisting of seven sections, six of which were connected in pairs at the manufacturing stage. The resulting four parts were assembled together at the cosmodrome of the Space Center John F. Kennedy in the Vertical Assembly Building. The factory connections of the sections were covered with an asbestos-silicate coating, and the connections made at the cosmodrome were closed with two rubber sealing rings (based on the results of the investigation into the causes of the disaster, the number of rings was increased to three) [2]. The coating was required in order to prevent the breakthrough of high-temperature gases and ensure normal operation of the accelerator during the entire acceleration stage.

Engineers at the Marshall Space Center pointed out that the proposed connection of parts of the Thiokol accelerator was unacceptable. One of the engineers suggested that the electrical connections made using a second O-ring completely useless, but solid rocket booster project manager George Hardy did not attach much importance to this and did not submit the reports to the engineers for approval. As a result, connecting parts using O-rings was approved for flights in 1980 [2]. Even before the disaster, it was determined that the seals were subjected to significantly more stress after increasing pressure, resulting in O-ring failure after assembly (from 50 to 100 and then to 200 psi). The reason for the burning of the rings was the resulting holes in the putty, through which, after launch, concentrated gas flows appeared, destroying the rings. Even though it is known that holes in the putty cause erosion, and that higher leak flow pressure causes these holes to enlarge, the Thiokol expands and NASA decided to apply increased pressure to ensure that the ring joint actually passed the test for tightness [2]. Thiokol provides complete information thanks to the effectiveness of the O-rings depending on the initial temperature of the assembly. Its engineers found that at 100°F the O-ring maintained a seal, at 75°F the seal failed in 2.4 seconds, and at 50°F the O-ring was completely ineffective. The peculiarity was that during thermal elongation of a working accelerator ring, it is necessary to take into account the diameter angle following the diameter of the connection, and when cold enough, the rings became insufficiently elastic for this. However, due to pressure from their management, who were afraid of transferring the contract for the supply of accelerators to another company, the engineers were not able to convey all serious problems to NASA. Figure 2 shows a diagram of the connection between two sections of the shuttle solid rocket booster: A - 12.7 mm thick steel wall, B - second o-ring, C - backup o-ring, D - stability, E - accelerator section, F - thermal insulation , G - coating, H - sealing paste, I - rocket fuel. The night of January 27-28 was unusually cold by Florida standards. Temperatures in the Cape Canaveral area dropped to -1°C. Some of the specialists sounded the alarm. During one of the previous launches, which took place after the shuttle sat on the launch pad during a cold night, particularly severe damage to the joint was noted. It was suggested that low temperatures reduced the flexibility of the vulcanized sealing material, making it more brittle and vulnerable to stress. It is known that on the morning of January 28, several Thiokol engineers demanded the cancellation of the upcoming launch.

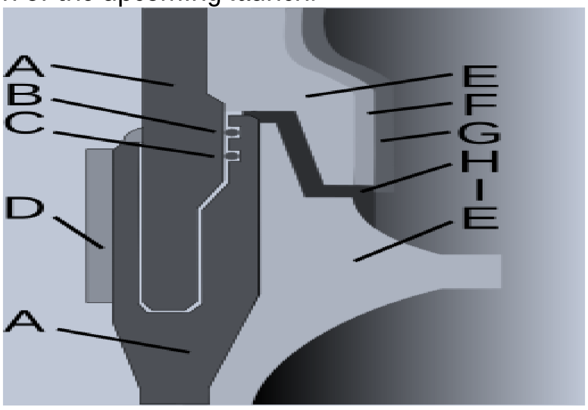

Figure 2: Connection diagram of two sections of the shuttle solid propellant accelerator

But, not wanting to disrupt the planned flight schedule, NASA officials opposed it. As for the Challenger crew, they did not even know about the existence of any problems with the boosters. After Challenger lifted off from the launch pad, the ship was doomed. Soon a stream of hot gases made its way through the sealing ring that had lost its elasticity and began to hit the fuel tank. She punched a hole in it and also burned out the connection to the accelerator itself. As a result, the accelerator simply came off and hit the fuel tank, which led to its instant destruction [3]. Based on O-ring observations, it is known that 24 flights of the spacecraft were successful. Also, the engineers calculated that a disaster could occur if all five main rings were broken, i.e. their seal will be compromised. At the same time, according to statistics, there were flights in which there was not a single breakage of the rings, but there were flights in which from one to three rings broke. Also, all previous flights of the ship before the disaster were at higher temperatures. On the day of the disaster, the air temperature was 36 degrees Fahrenheit and analysis of rings at this temperature had not previously been carried out. Based on the importance of analyzing the main causes of the disaster and preventing such events in the future, the scientific and practical task of analyzing the creation of a system for predicting the probability of aircraft accidents is relevant. To solve the problem, data mining methods (DMA) were used [4]. IAD or (Data Mining) is a decision support process based on searching for hidden patterns (information patterns) in data. In this case, the accumulated information is automatically generalized to information that can be characterized as knowledge. In general, the IAD process consists of three stages [5]: identifying patterns (free search); using identified patterns to predict unknown values (predictive modeling); exception analysis, designed to identify and interpret anomalies in the found patterns.

Sometimes an intermediate stage of checking the reliability of the found patterns between their discovery and use (validation stage) is explicitly identified. To analyze data on the spacecraft's O-rings and compare the analysis results with the conclusions of the investigative commission, the specialized Orange program was used [6].

3. Research methodology

The methods used by the authors to study the causes of failure of the sealing rings of the spacecraft fuel tank were methods from a modern scientific direction - intelligent data analysis (IDA). In this case, the IDA technology contains a number of stages: identifying patterns (free search); using the identified patterns to predict unknown values (predictive modeling); analysis of exceptions designed to identify and interpret anomalies in the found patterns. Based on the available set of open experimental data on spacecraft flights by year and the corresponding number of failures of the sealing rings of the fuel tank under different external conditions, the authors used regression analysis - logistic regression, as well as the k-means data clustering method. These methods have proven themselves to be successful in conducting statistical studies [7].

4. Development of a prediction model for O-ring failures in the Orange program

The Orange analytical system is a program with open source code for machine data processing and data visualization, which has a large set of advanced functions. The Orange software product, which is developed by the Laboratory of Bioinformatics at the University of Ljubljana, is intended for intelligent data analysis, statistical research and data visualization [8]. The method of modeling is the creation of an expert system for predicting the probability of ring failure at given outside temperatures, presented before analysis, based on criteria known from additional statistical data of failures during previous flights. At the initial stage, the system that is fragmented uses logistic regression algorithms [8,9]. We begin by adding a statistics file to the program (Fig. 3, a). The main goal is to predict the probability of malfunctions, i.e. ring breakage. To achieve this, we will use the variable Y, which we will define as the target (breakdown prediction). It should be noted that a value of 0 indicates that the ring has no breaks, while a value of 1 indicates that it is broken. Next comes the stage of connecting to File in our Data Table widget project for convenient display and analysis of available data (Fig. 3). Based on the results of the data analysis of the table, it can be seen that the first five rows (five rings) show the results of the first experiment (flight), identifying the ring that was damaged at a temperature of 53 degrees (three rings). The next five rows presented the results of other experiments conducted at 57 degrees (1 ring), and so on. Added a new module Distributions for visualization, which will allow us to avoid that a larger number will end up intact (Fig. 4). In Fig. 5 it is shown as a variable Y (value 0 – 110 rings are intact, 10 rings are broken during the entire flight).

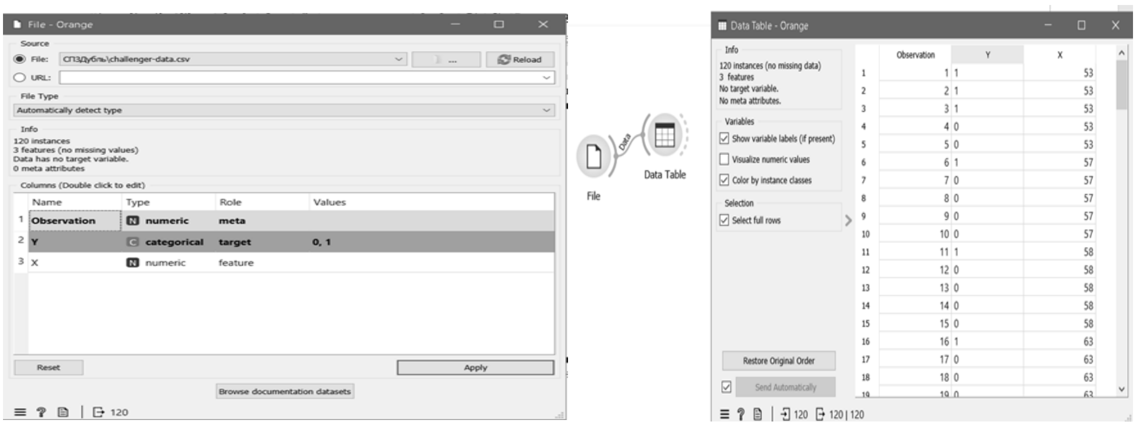

Figure 3: Adding the first file with statistical data and connecting a statistical data table to a file in the Orange program

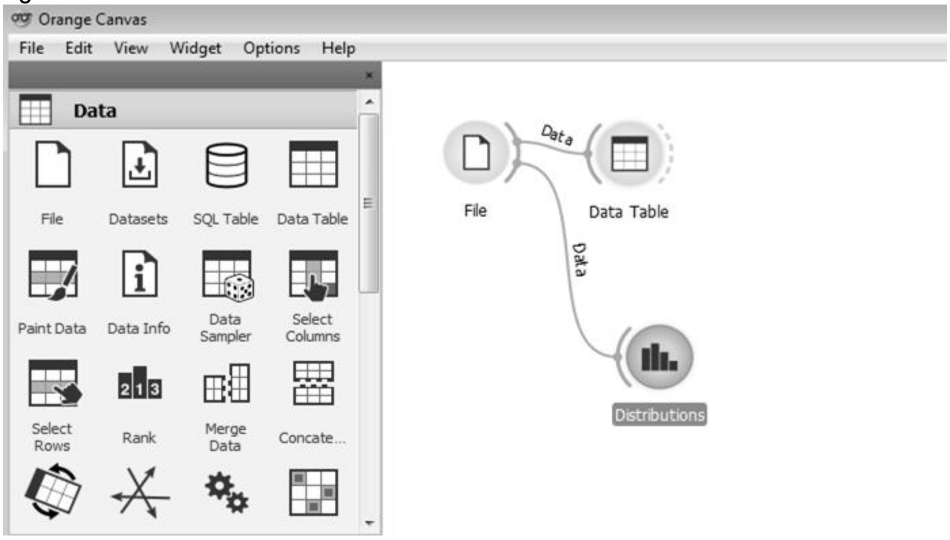

Figure 4: Adding a data visualization and analysis module Distributions

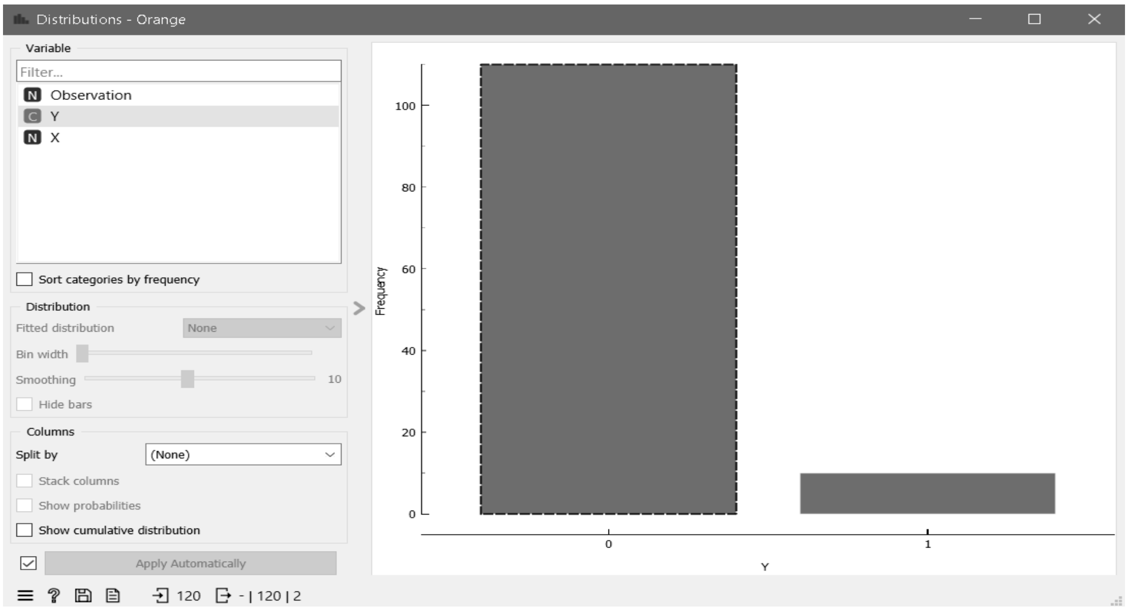

Figure 5: Y graph

It can be noted that previous flights were conducted at temperatures ranging from 50 to 90 degrees Fahrenheit, analyzed using the X variable (Figure 6). To investigate changes in the probability of failure occurrence as a function of temperature, temperature can be classified based on failure categories (Fig. 7). We will connect the Logistic Regression and Testing and Evaluation modules together. The Logistic Regression module is connected to the primary file (Fig. 8). The adequacy indicator of the model AUC = 0.74 (see Fig. 7). The AUC indicator (Area Under the ROC Curve) is a measure that allows you to summarize the performance of the model in a single number, measuring the area under the ROC curve. AUC ranges from 0 to 1, where a higher value of AUC indicates a higher performance of the obtained mathematical model [10].

It is known that the *logistic regression* or *logit model* used in the program is a statistical model used to predict the probability of an event occurring by comparing it with a logistic curve. This regression produces an answer in terms of the probability of a binary event (1 or 0). Logistic regression is used to predict the probability of an event occurring based on the values of a set of attributes. To do this, a so-called dependent variable is introduced, which takes only one of two values - as a rule, these are the numbers 0 (the event did not occur) and 1 (the event did occur), and a set of independent variables (also called signs, predictors or regressors) - real, on based on the values of which it is necessary to calculate the probability of accepting a particular value of the dependent variable [11].

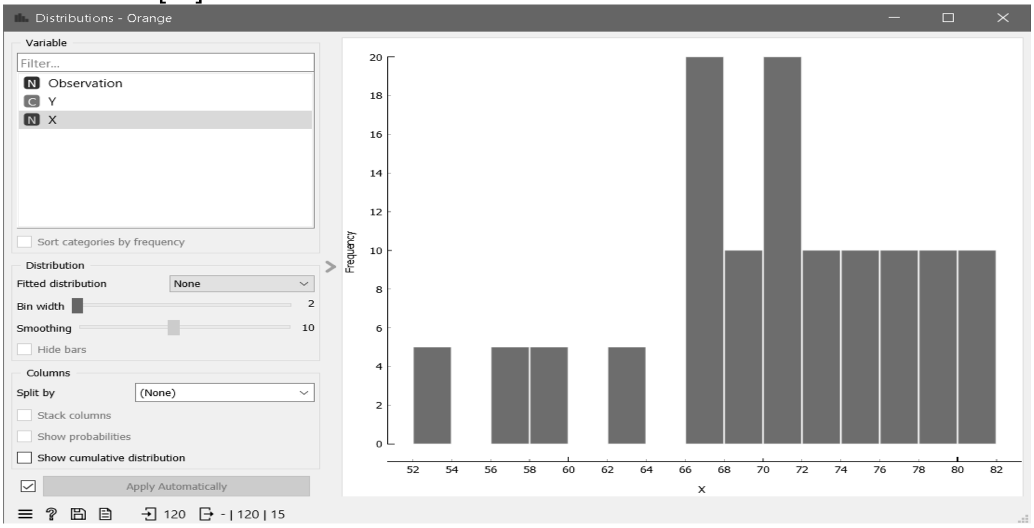

Figure 6: Graph by X (air temperature at shuttle launch)

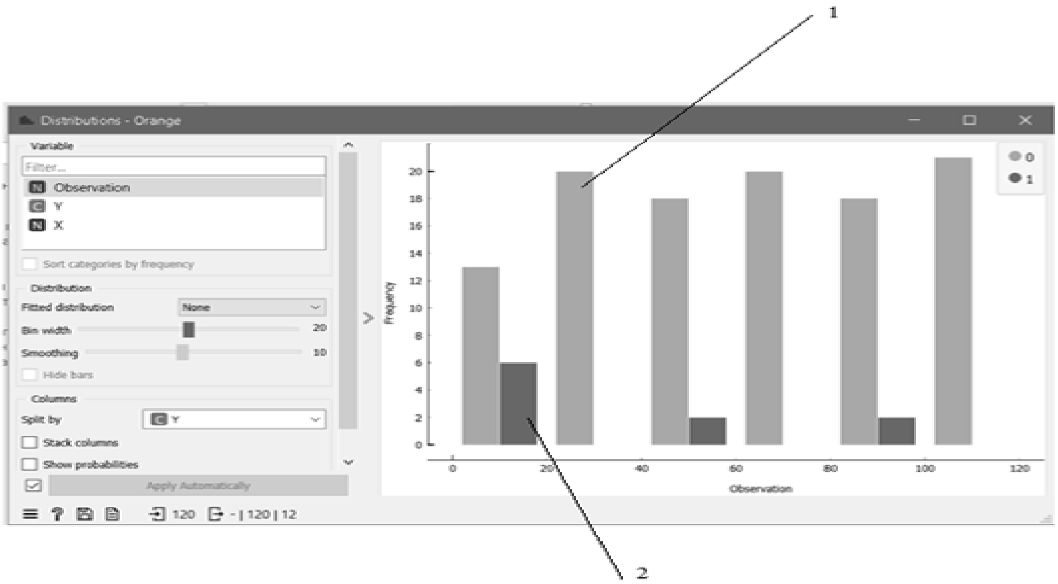

Figure 7: Relation of faults to temperature (1 – ring is working, 2 – faulty)

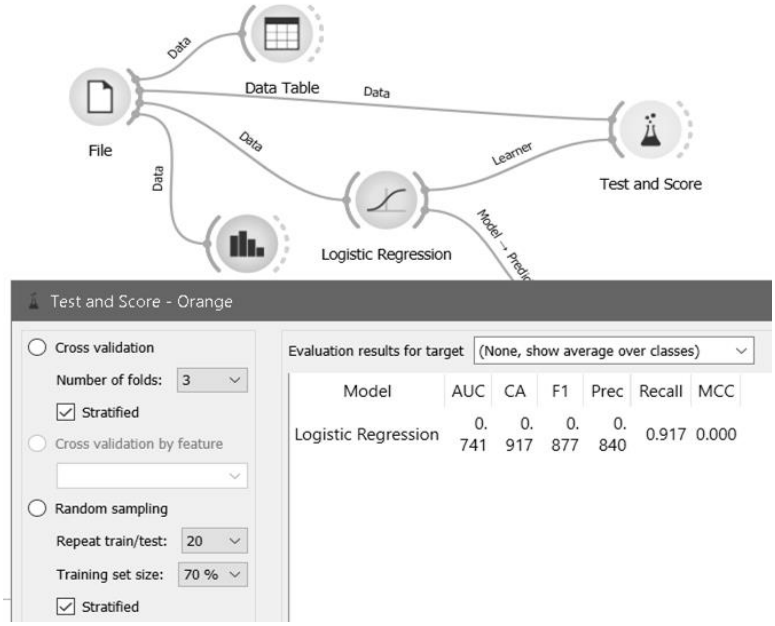

Figure 8: Testing and evaluations

The next stage is the inclusion of a file with temperature data on the day of launch to the program and the setting of the necessary parameters (Fig. 9). In this file, the parameters at a temperature of 36 degrees are specified, which are used to predict the probability of ring breakage.

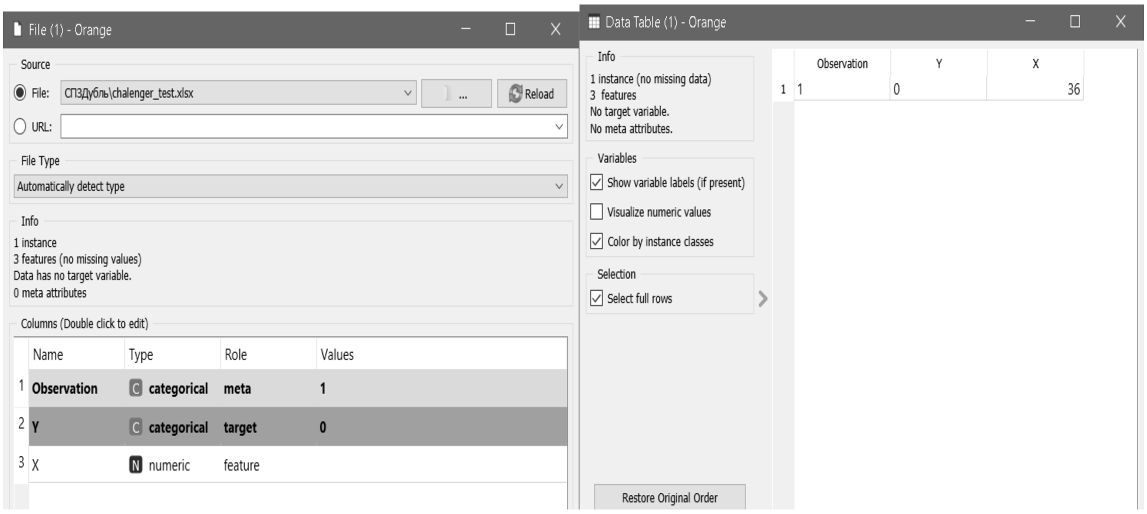

Figure 9: Files with new data

Display of predictions, or "Predictions", will allow us to predict the probability of failure by analyzing the data from the file using the results obtained from the regression analysis and result of predicting the failure of one ring (Chances of ring failure at 36 degrees Fahrenheit) is shown in Fig.10.

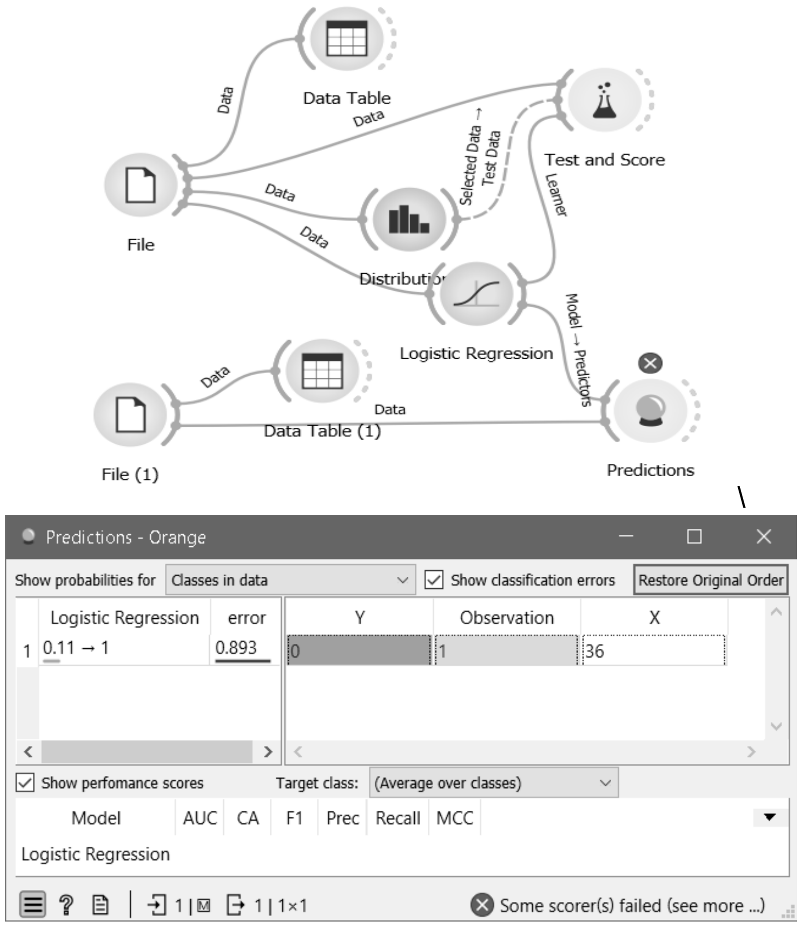

Figure 10: General connection of widgets in the program and chances of ring failure at 36 degrees Fahrenheit

The analysis showed that at a temperature of 36 degrees, the probability of failure of one ring is 89%, and the probability of failure for the entire set of five rings is 56% ($0.892 \approx 0.56$). That is, the expert system does not recommend launching the shuttle at this temperature. The obtained result fully confirms the conditions of the launch and the subsequent disaster. Thus, it is known that the morning of the next day, January 28, turned out to be unusually cold - the temperature dropped to −1 °C, the minimum acceptable for launch permission. The previous coldest start was carried out at a temperature of 12 °C. The low temperature caused concern among Thiokol engineers. During a closed televised conference between Thiokol and NASA management, they expressed concerns that such extreme conditions could adversely affect the elasticity of the solid rocket booster O-rings, since the connections had not been tested at temperatures below 12 °C, and recommended delaying the launch [12, 13]. At the same time, the problem of the O-rings had not yet been solved and was of the highest criticality, engineers doubted that both rings would be able to maintain the tightness of the connection at low temperatures. Thiokol engineer Roger Beaujoli insisted on the immediate cancellation of the launch. Thiokol management supported its engineers, but due to constant delays, NASA management was categorically against it. They dismissed concerns about the inelasticity of the ring, groundlessly arguing that if the main ring could not provide a seal, the backup ring would do so [14, 15, 16]. Seeing NASA's categorical attitude, the Thiokol management gave in and gave permission for the launch, which led to the subsequent destruction of the rings, fuel leakage and fire of the ship's fuel tank [17, 18, 19].

For additional analysis of the causes of the disaster, the authors used an additional data mining method – k-means. The k-means method is the most popular clustering method. It was invented in the 1950s by mathematician Hugo Steinhaus [20] and almost simultaneously by Stuart Lloyd [21]. It gained particular popularity after the work of MacQueen [22].

The action of the algorithm is such that it seeks to minimize the total squared deviation of the cluster points from the centers of these clusters:

$$V = \sum_{i=1}^{k} \sum_{x \in S_i}) \, (x - \mu_i)^2$$

where k is the number of clusters, S_i are the obtained clusters, $i = 1, 2, ..., k$, and μ_i are the centers of mass of all vectors x from cluster S_i

The k-means method is used for data clustering based on an algorithm for partitioning a vector space into a predetermined number of clusters k. The algorithm is an iterative procedure in which the following steps are performed:

1. The number of clusters k is selected.
2. From the initial data set, k observations are randomly selected to serve as the initial cluster centers.
3. For each observation of the initial set, the closest cluster center is determined (distances are measured in the Euclidean metric). In this case, the records “attracted” by a certain center form the initial clusters.
4. Centroids are calculated — the centers of gravity of the clusters. Each centroid is a vector whose elements are the average values of the corresponding features calculated for all records of the cluster.
5. The cluster center is shifted to its centroid, after which the centroid becomes the center of a new cluster.
6. The 3rd and 4th steps are repeated iteratively. Obviously, at each iteration, the boundaries of the clusters change and their centers shift. As a result, the distance between elements within the clusters is minimized and the intercluster distances increase.

The algorithm stops when the cluster boundaries and centroid locations do not stop changing from iteration to iteration, i.e., at each iteration, each cluster will contain the same set of observations. In practice, the algorithm usually finds a set of stable clusters in several dozen iterations.

To implement the presented algorithm, the K-means widget is used (Fig. 11).

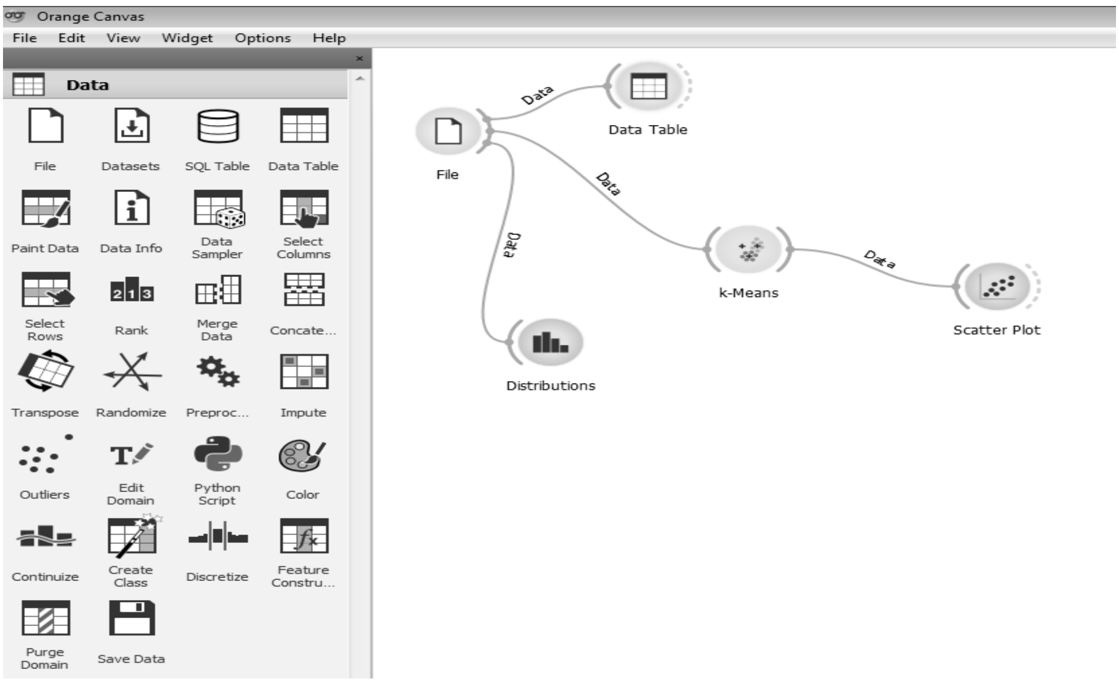

Figure 11: Software implementation of the k-means method

The window for demonstrating the result of the method is shown in Fig. 12.

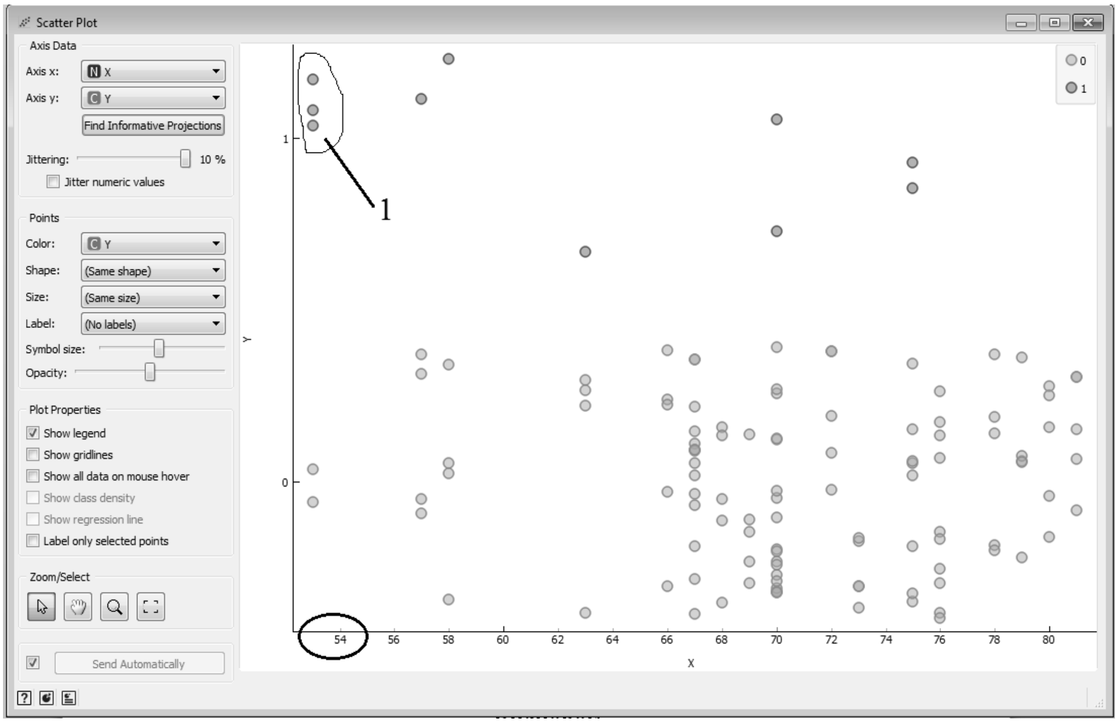

Figure 12: The result of the data classifier analysis method

Visual analysis of the data shows the highest number of ring failures (marked 1 in Fig. 12) at low temperatures, i.e. below 54 degrees Fahrenheit (12 degrees Celsius). Thus, it can be concluded that low air temperatures significantly affected the elasticity of the fuel tank rings.

5. Conclusions

The performed analysis of the use of regression modeling to predict the probability of failure of the space shuttle Challenger sealing rings emphasizes the high efficiency of this method for predicting potential emergency situations. During the research, tools such as Data Table, Distributions, Logistic Regression, and Test and Score were used, which allowed to systematically analyze the provided data.

The degree of adequacy of the mathematical model implementing the logistic regression method was 74%, which indicates an acceptable result of the data obtained.

The obtained results indicate that at a temperature of 36 degrees Celsius there is an 89% probability of ring failure, which indicates a significant risk of defects. Overall, the average probability of failure for all five rings is 56% ($0.89^2 \approx 0.56$), which indicates risky conditions for such a configuration in flight conditions. The limitation of the proposed approach, according to the authors, is the need for a large array of statistical experimental data on defects in the fuel elements of spacecraft.

The application of regression analysis in aerospace programs can serve as an important tool for reducing risks, increasing the level of safety of space missions, and contributing to the development of national and international space initiatives [23]. The proposed model allows predicting with a high degree of accuracy the possible failure of fuel tank elements of modern spacecraft. As a prospect for further research, it is possible to analyze the causes of accidents of experimental spacecraft for flights to Mars of the Elon Max company in the presence of accessible statistical data using the universal mathematical models proposed by the authors implementing the IAD methods.

References

[1] Why hasn't SpaceX landed humans on the Moon yet?. URL: https://www.quora.com/Why-hasnt-SpaceX-landed-humans-on-the-Moon-yet#:~:text=It%20builds%20rockets%20for%20money,could%20land%20on%20the%20Moon.

[2] Shuttle_Challenger_disaster, 2024. URL: //https://en.wikipedia.org/wiki/Space_Shuttle_Challenger_disaster.
[3] V.S. Mikhailenko, R.Yu. Kharchenko, Analysis of traditional and neuro-fuzzy adaptive systemof controlling the primary steam temperature in the direct flow steam generators in TPS. Automatic Control and Computer Sciences 48 6 (2014) 334–344. doi: 10.3103/S0146411614060066.
[4] V.S. Mikhaylenko, R.Y. Kharchenko, V.A. Shcherbinin, Analysis of the Predicting Neural Network Person Recognition System by Picture Image. Automatic Control and Computer Sciences, 54 3 (2020) 249–258. doi:10.3103/s0146411620030037.
[5] V.S. Mikhailenko, R.Yu. Kharchenko, V.A. Shcherbinin, V.V. Leshchenko, Using Neural Network Technologies to Simulate the Working Processes of Ship Steam Boilers CEUR Workshop Proceedingsthis link is disabled, urn:nbn:de:0074-3126-7, 2021, 3126, pp. 367–373.
[6] Data Mining Fruitful and Fun. URL: https://orangedatamining.com/
[7] K-means. URL: https://en.wikipedia.org/wiki/K-means_clustering
[8] R.Yu. Kharchenko, A.V. Kochetkov, V.S. Mikhaylenko, Analysis of methods for automated research of dc voltage converters of modular structure. Radio Electronics, Computer Science, Control 3 7 (2022). doi:10.15588/1607-3274-2022-3-1.
[9] R.Yu. Kharchenko, V.S. Mykhailenko, A.V. Kochetkov, Development of a Neuro-Fuzzy Intelligent Network for Monitoring and Control of Microclimate Systems. Automatic Control and Computer Sciences, 57 1 (2023) 27–36. doi:10.3103/s0146411623010066.
[10] SAP ANALYTICS CLOUD, 2024. URL: //https://help.sap.com/docs/SAP_ANALYTICS_CLOUD
[11] V. Mezhuyev, Y. Gunchenko, V. Mikhailenko, R. Kharchenko, V. Leshchenko, Mathematical Simulation Process Increasing Energy Efficiency of Ship Steam Boilers., in: Proceedings of the 11-th International Conference "Information Control Systems & Technologies" Odesa, Ukraine, September 21–23, 2023, pp. 141-152.
[12] A. Z. Spector, Achieving application requirements, in: S. Mullender (Ed.), Distributed Systems, 2nd. ed., ACM Press, New York, NY, 1990, pp. 19–33. doi:10.1145/90417. 90738.
[13] The Space Shuttle Challenger Disaster. Risk Revealed, (2024) 13–23. doi:10.1017/9781009299794.004
[14] J.M. Logsdon, Truths, Lies, and O-Rings: Inside the Space Shuttle Challenger Disaster. Space Policy, 25 4 (2009) 269–269. doi:10.1016/j.spacepol.2009.09.002
[15] Challenger missions 1983–1986. (n.d.). Springer Praxis Books, (2007) 273–282. doi:10.1007/978-0-387-49679-5_7
[16] The Space Shuttle Challenger Disaster. Saving Human Lives, (2005) 107–153. doi:10.1007/1-4020-2980-2_7
[17] Space Shuttle Challenger. Springer Praxis Books. Springer New York, (2007). doi:10.1007/978-0-387-49679-5
[18] J. Garten, The Space Shuttle Challenger Accident. (2011). doi:10.4135/9781526406958
[19] S. Pomeroy, Truth, Lies, and O-Rings: Inside the Space Shuttle Challenger Disaster. Technology and Culture, 51 4 (2010) 1038–1040. doi:1353/tech.2010.0077
[20] Recovery after Challenger. (n.d.). Springer Praxis Books in Space Exploration, 99–146. doi:10.1007/978-0-387-73972-4_3
[21] H. Steinhaus On the division of material bodies into parts. Bull. Acad. Polon. Sci., C1. III IV (1956): 801—804.
[22] S. Lloyd, Least square quantization in PCM's. Bell Telephone Laboratories Paper. (1957).
[23] J. MacQueen Some methods for classification and analysis of multivariate observations. in: Proc. 5th Berkeley Symp. on Math. Statistics and Probability, 1967, pp.281-297.
[24] V. Vychuzhanin, N. Rudnichenko, Z. Sagova, M. Smieszek, V. V. Cherniavskyi, A. I. Golovan, M. V. Volodarets, Analysis and structuring diagnostic large volume data of technical condition of complex equipment in transport. IOP Conference Series: Materials Science and Engineering, Volume 776, 24th Slovak-Polish International Scientific Conference on Machine Modelling and

Simulations - MMS 2019, 3-6 September 2019, Liptovský Ján, Slovakia, 2019 pp.1-11. DOI:10.1088/1757-899X/776/1/012049.