KBQS – Capstone project

Naila Mendes
Professional Certificate in ML & AI - December 2022
UC Berkeley

1 Business understanding

Risk management (RM) applied to aviation maintenance is the key for improving performance, finding latent conditions, preventing human errors, and consequently improving safety. The focus on increasing safety should be to eliminate or prevent hazards, not to eliminate failures (Leveson, 2016). Hazard identification is at the core of mitigating risks, but its effectiveness can be easily affected by individuals conducting the risk assessment (Goh et al., 2010). Furthermore, either because of lack of experience or lack of time, hazard identification often fails to identify hazards associated with an activity (Goh et al., 2010). It is suggested that the use of artificial intelligence (AI) tools is one of the ways to deal with these potential problems and to facilitate the hazard identification process (Goh et al., 2010). AI has many benefits such as creating predictive algorithms (Boden, 1996) and is a point of interest to be applied in a socio-technical system (Mogles et al., 2018). AI is one of the tools that have been emerging in the past years and has been applied to risk management (Robinson, 2019; Ratnayake, 2016).

Presently human error models focus on predicting errors to assess the probability of human error (Claros et al., 2017), but the results can be sub-optimal due to data limitations related to the lack of consistency in incident classification (Claros et al., 2017). Accident causation models and new methods can be developed to predict and prevent accidents (Chen et al., 2021). A combination of multiple methods could be an excellent solution, where each method can overcome each other's weaknesses (Mogles et al., 2018), thus creating a more robust safety and quality system in aviation maintenance (Mendes et al., 2022).

The Knowledge Base Quality System (KBQS) methodology applied in aviation maintenance was developed to improve the overall safety in aviation by combining predictive and interactive methods (Mendes, 2022). This methodology is a combination of data collection and integration reported between different systems, data analysis to predict new risks, action prioritization on decision-making, active engagement of the work force by an active interface with system and process designers, managers, and operators (Mendes, 2022).

The goal of this Capstone project is to create a Machine Learning model focused on creating a predictive approach to the KBQS methodology.

KBQS intends to create a Risk Management methodology that combines predictive and interactive methods. A predictive measurement includes: use the data collected to predict and prevent negative outcomes applying artificial intelligence

AI and Machine Learning (ML) models and algorithms can be applied and explored on risk modeling and on risk management based on extensive literature review. AI field has been applied to risk management in the past years. Risk analysis can be performed using Linear regression to determine which factors are more impactful managing a quality and safety system. The algorithms can continuously monitor a maintenance service center and determine ways to operate in an optimized level. Risk management can also use classification models such as KNN or Logistic regression can be used to predict when a combination of factors will create a quality escape or an incident. Another example is Logistic Regression, that can also be applied on risk analysis during audits or on risk evaluation. Logistic regression is typically used for problems involving binary classification, which helps predict output variables that are discrete into two classes. The maintenance service centers can monitor their risk, allowing time to determine if it needs further action.

1.1 Problem Statement

Analyze historical data of quality escapes and incident and other types of hazard reports and quality reports to understand which understand which factors can help to predict and prevent quality escapes and safety incidents in maintenance aviation.

1.2 Business goal

Organizations need an early recognition of latent risks and offer timely and targeted countermeasures. Proactive and predictive countermeasure for eliminating threats or risks on high reliability organizations, specifically in aviation industry.

Goal: Predict the most relevant/risk factors on a maintenance service center as well as predict the overall risk using a machine learning model. Reduce the risk, increase customer satisfaction, protect the brand. It is vital to identifying which conditions will help to predict and improve the safety in aviation.

2 Data understanding

Different features can be collected during the data collection to determine the risk of the company in different areas: finance department (default, account receivables and payables, P&L), HR (turnover rate, training hours), Operations (quality escapes, on time delivery, loading), etc. Each area can be investigated to determine the risk level. The output training data can be selected associating if there was a negative or did not have a negative event depending on the area selected (quality escape, low customer survey, finding, high turnover rate, etc).

The Knowledge Base Quality System is composed by 3 essential elements: Data Collection and monitoring, Data Analysis and Processing, Corrective Action Implementation. The modeling will focus on the data collection and the data analysis process by creating a Machine Learning model.

Aviation maintenance consists of multiple processes interconnected, where each process supports the core process, that is the inspection and maintenance of the aircraft. Each maintenance process is essential to allow the successful completion of the next process without quality or safety issues. Common supporting processes in aviation maintenance are: Technical data; Tool Control; Material Control; Engineering; Inspection & Quality System; Record of Maintenance; Training; Technical Personnel; Operations; SMS; QMS; Subcontractors; Planning, Scheduling (AS9110). The control of processes in a maintenance organization is essential. Considering why each process is so crucial, these processes were grouped in Clusters. The system also collects additional parameters input from System Tracking such as man-power loading. The parameters that are monitored and indexed are hangar loading, the ratio between the manpower occupied by manpower available.

2.1 Gathering and Describing Data

Data source: Historical data from Quality System and Safety Management System

Data:

- Predict feature: quality escape (independent variable)
- Other features (dependent variables):
    - Categorical features:
        - Cluster: A group of similar process to track each area independently
        - Location: department, site, hangar
        - Team and shift
        - Type of report: QCAR: non-conformity or findings, QPY: throughput yield or internal quality escapes, small issues reported during the maintenance process by inspectors, or external quality escapes TPY), incident: safety related issue that is identified after the aircraft is released.
    - Numerical features:
        - Date and time
        - Risk Index: Associated with the impact of each Cluster and each type of report, for example: non-conformity, different level of findings based on their impact, or safety incidents.
        - Loading: Hangar loading calculated by labor hours used divided by available man power
        - Certified/non-certified: ratio between total technical personnel that are certified by non-certified

2.2 Early data exploration and Quality check

The data was prepared and explored. A review of the risk index was created below:

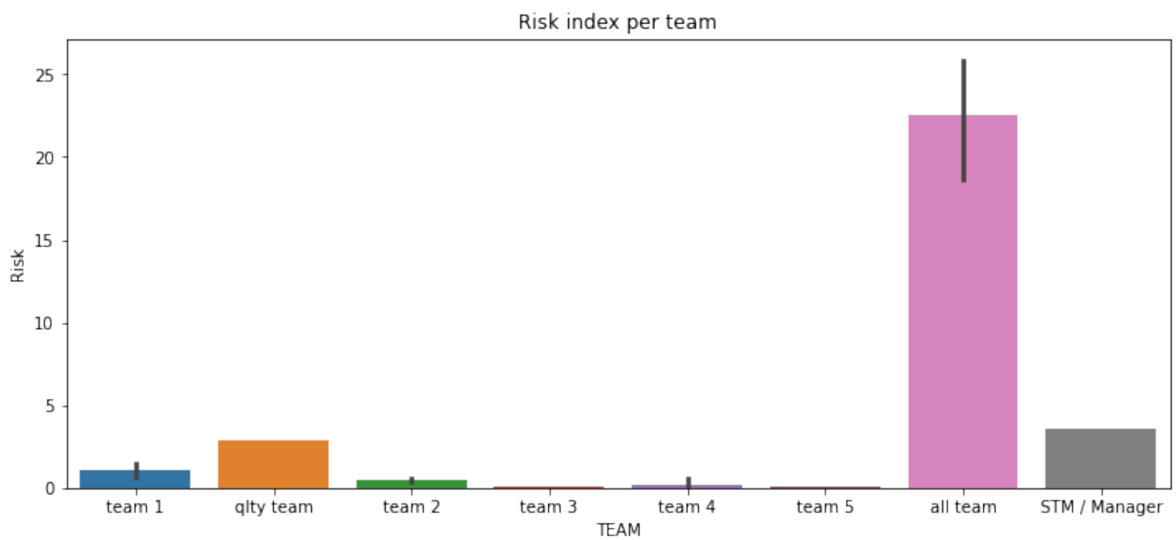
Risk Index Sum for each Cluster

Cluster	Risk
EDUCATION & TRAINING	161.002590
RECORD OF MAINTENANCE	106.922344
INSPECTION & QUALITY SYSTEM	79.690070
TOOL CONTROL & MANAGEMENT	76.466912

SAFETY	51.147868
QUALITY SUPPORT	37.255436
TECHNICAL DATA	26.968678
TECHNICAL PERSONNEL	26.833765
PARTS/MATERIAL MANAGEMENT & CONTROL	3.104565
OPERATIONS & FACILITY	0.161617

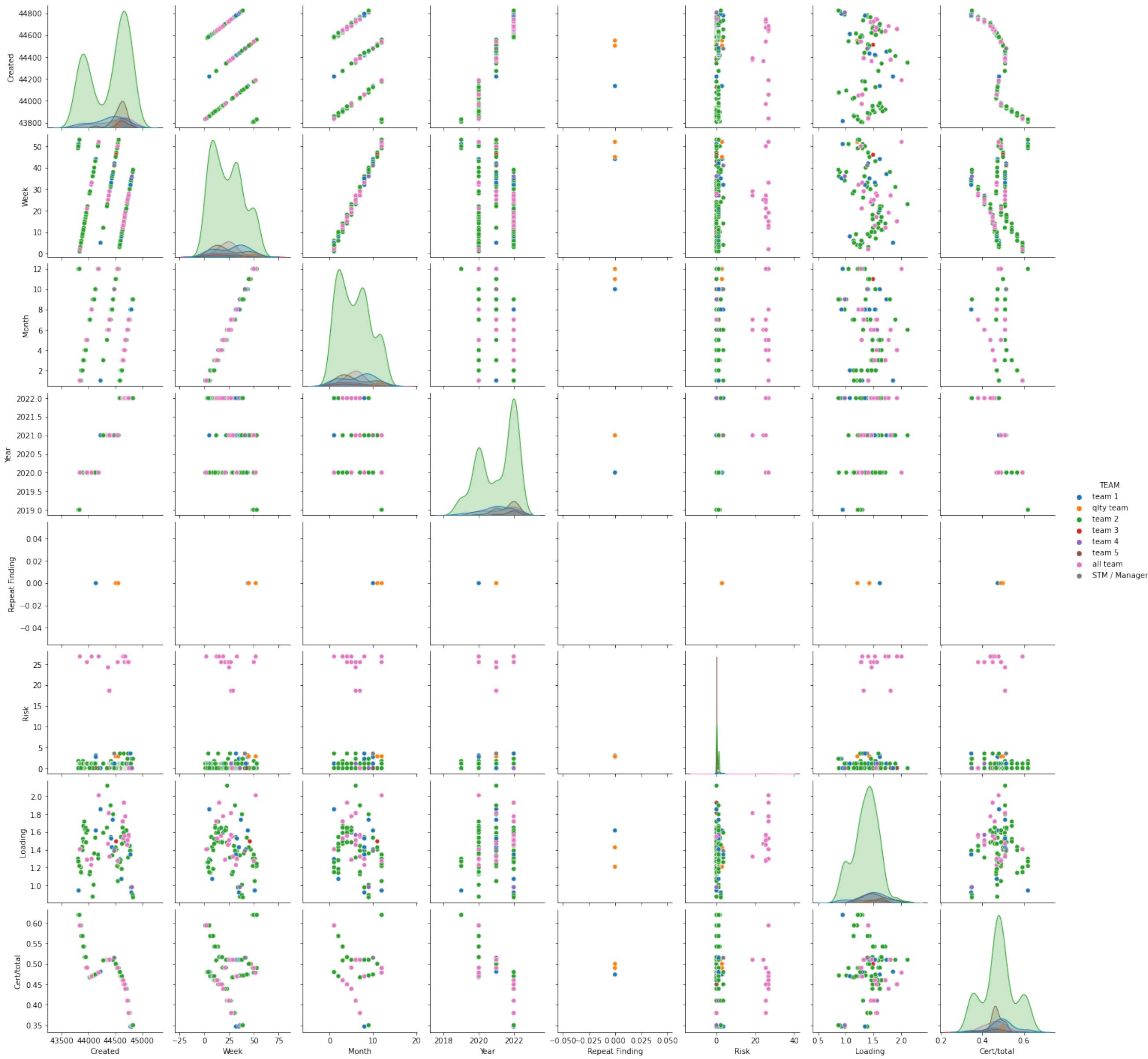
Risk Index per Team

Cluster	Risk
all team	428.970241
team 2	102.898682
team 1	25.007492
qlty team	5.776464
STM / Manager	3.577835
team 4	1.533562
team 5	1.469066
team 3	0.320504



Team 1 has reports that the total sum is higher. All team, STM/Manager, and quality team get many different reports associated with multiple teams.

Team



There are still many features that can be potentially dropped off from the analysis.

3 Data preparation

3.1 Data cleansing

There are some columns that had data missing such as Repeat Finding, Applicable Rules, Main root cause, and Additional Root cause, and Loading. Except for loading, most of these columns don't add value to this analysis. It may be interesting to use when applying NLP techniques.

The time related data was also removed for this analysis.

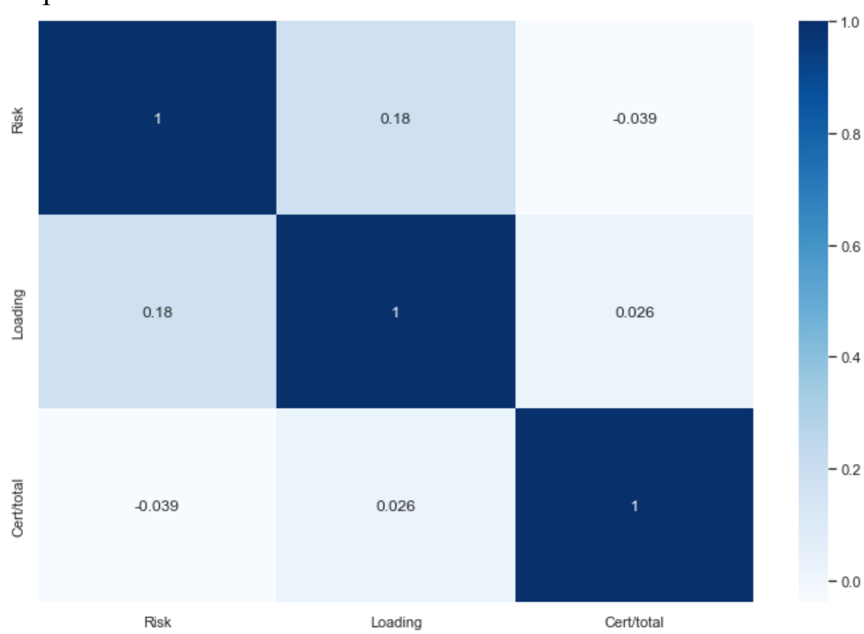
The finding code was also removed since it is directly associated to a cluster not adding value to this analysis.

Based on pre-evaluation:

3.2 Final data
Sample of the final features:

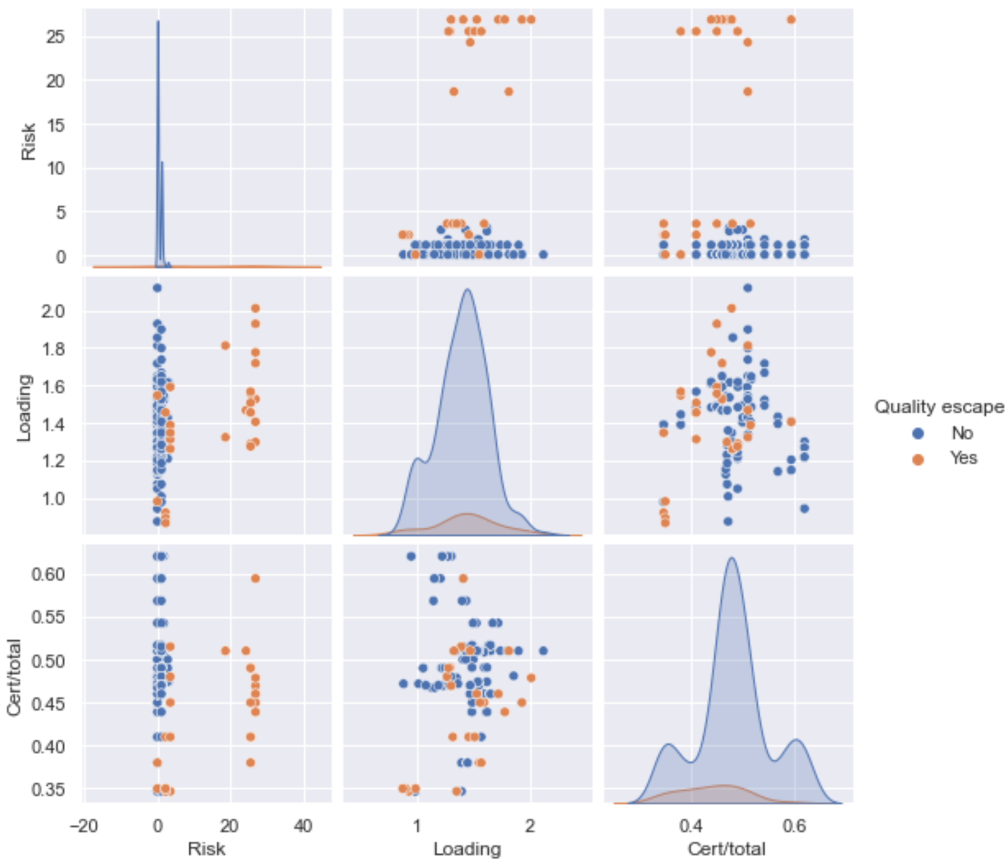
Cluster	Risk	Quality escape	Loading	Cert/total	
73	5	0	0	1.649113	0.460000
231	5	4	0	1.337504	0.510000
279	1	13	1	1.455548	0.410000
267	9	13	1	1.288590	0.490668
98	5	0	0	1.283031	0.470000
175	5	0	0	1.432404	0.568140
276	0	14	1	1.928229	0.450000

The heatmap below was created to ensure that no additional features need to be removed.



4 Data understanding

The final data selected has a total of 5 features including the quality escape. The relation of the loading, cert/ total and the risk are high with the quality escape. However, correlation is not causation. The factors were identified as the ones that will work better predicting a quality escape.



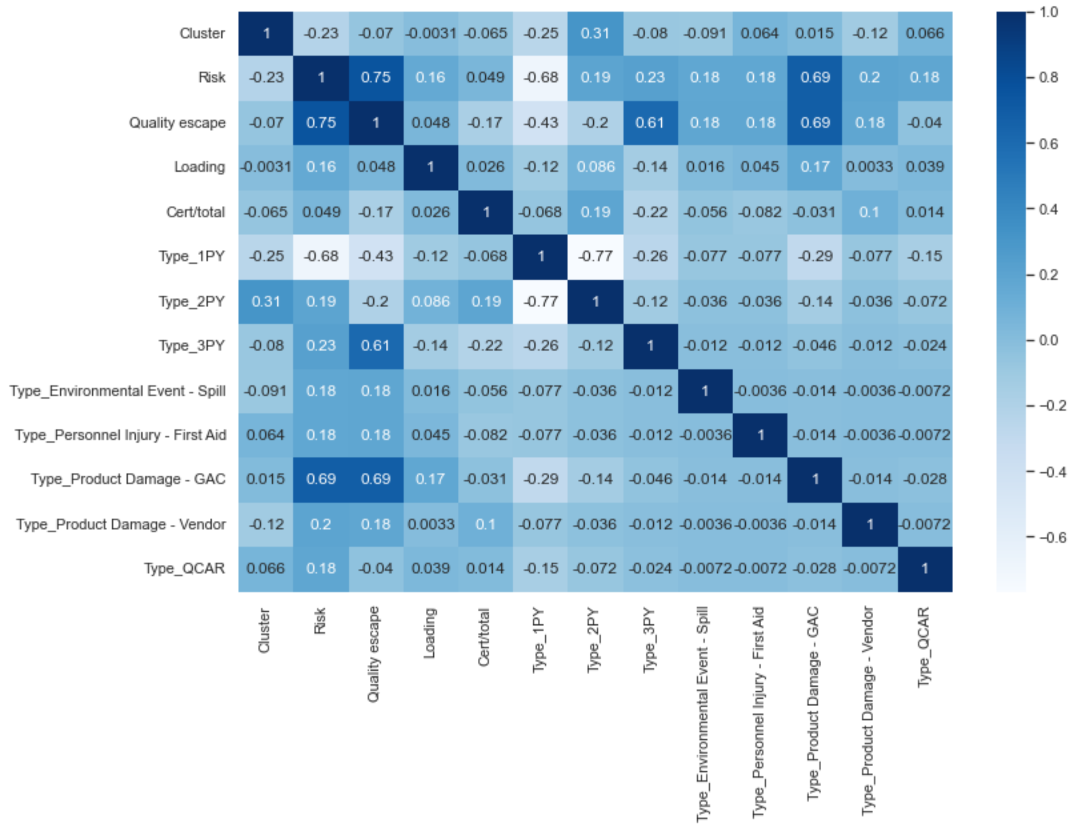
- Modeling Techniques that will be used are Linear regression to determine each feature impact and Classification Models.
- The Expected results: Create an accurate recommendation/classification model
- Identify the feature importance based on the model and selected data

The difference between Regression and Classification models, is that the first given a set of features, the goal is to predict a real-valued outcome, and the latter the goal is to predict what class the sample belongs to (AI/ML Berkeley, 2022). Create a regression model could be useful to create a comparison risk score between hangars, shops, and different site locations. This would be useful to monitor the risk based on the parameters collected on each Hangar. The metrics to measure Regression will be R_squared and MSE.

4.1 Exploratory Data Analysis

```
LinearRegression()
Baseline Score (R_squared): 1.0
```

```
Baseline (Coeficients): [-2.47337152e-16  8.51221502e-17 -7.21852782e-18  8
.96183286e-16
 -6.25000000e-01 -6.25000000e-01  3.75000000e-01  3.75000000e-01
  3.75000000e-01  3.75000000e-01  3.75000000e-01 -6.25000000e-01]
Baseline MSE (Training data): 4.862294448549749e-31
Baseline MSE (Test data): 3.428697816486572e-31
```



This data works best with a Classification model. Using Logistic Regression

```
Simple model Train score: 0.9904761904761905
Simple model Test score: 1.0
```

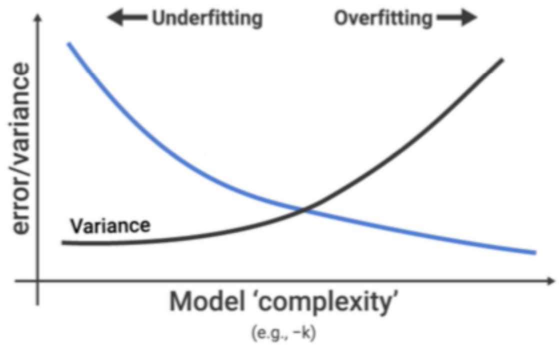
The accuracy is too high since the reports determine if there is a quality escape. The columns "Type" needs to be dropped from the data and then re-evaluated. Considering that the goal of the research is to help the service center by building a classification model that can predict quality escapes on the basis of present factors and historical risk factors. A Classification model, using accuracy and recall makes more sense.

5 Modeling

A classification model predicts if there will be a quality escape based on the data analysis of the features. A Classification model is "a supervised machine learning algorithm is one that relies on labeled input data to learn a function that produces an appropriate output when given new unlabeled data "(Berkeley AI, class 12).

- K-nearest neighbors

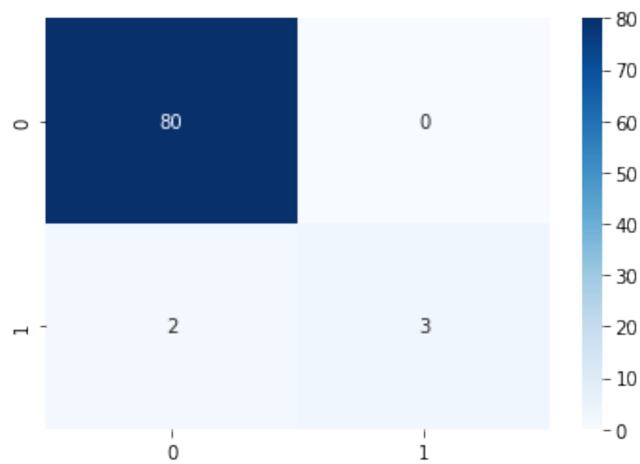
- “KNN algorithm is a non-parametric, supervised learning classifier, which uses proximitiy to make classifications or predictions about the grouping of an individual data point. It works off the assumption that similar points can be found near one another “(Berkeley AI, class 12)



As the misclassification rate (-k) decreases or accuracy increases, the error decreases, although the complexity of the model increases and also the chance of overfitting. (Berkeley AI, class 12).

Accuracy simply measures how often the classifier correctly predicts. We can define accuracy as the ratio of the number of correct predictions and the total number of predictions.

KNN Classifier Accuracy



5.2 Model fitting

GridSearchCV was used to identify the best k on KNN model with the lowest validation error applying cross-validation (Berkeley AI, class 12). The classification rate is minimized by maximizing accuracy. Also the data was standardized applying Scaler.

5.2 Model evaluation

5.2.1 Models Comparison and Overview

Comparing the performance of the Dummy model, Basic Model, Logistic Regression model, KNN Classifier, Decision Tree, and SVC models.

Models	Best Score	Mean Fit Time	Train Accuracy	Test Accuracy
Dummy Model	-	-	0.801	0.800
Basic Model	-	-	0.907	0.911
Logistic Regression	0.91	0.299	0.855	0.852
Decision Tree	0.914	0.058	1.000	0.889
KNN Classifier	0.905	0.041	0.926	0.902
SVC	0.908	23.74	0.920	0.910

Decision tree looks overfitted. KNN classifier performed slightly better than basic model. SVC had slightly better results, but the time was higher than the other models. KNN will be further improved.

5.2.2 Improving the model

Data metrics - Features The features should not be reduced, since it was already cleaned to keep the basic metrics from the Service Center

Optimizing KNN hyperparameter: Exploring additional hyperparameters of our models have additional hyperparameters to tune and explore. Update verbose to True on GridSearchCV Test another type of GridSearchCV and tweak the Number of neighbors in KNN.

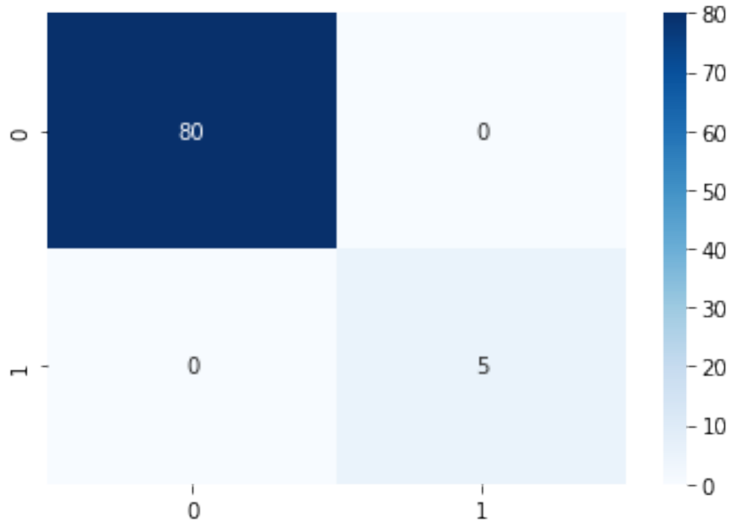
Results:

```
Fitting 5 folds for each of 144 candidates, totalling 720 fits
              precision    recall  f1-score   support

           0       0.98      1.00      0.99        80
           1       1.00      0.60      0.75         5

    accuracy                           0.98        85
   macro avg       0.99      0.80      0.87        85
weighted avg       0.98      0.98      0.97        85

KNN Classifier Best Params: {'knn__algorithm': 'auto', 'knn__n_neighbors': 10, 'knn__weights': 'uniform'}
KNN Classifier Accuracy: 0.9694871794871794
KNN Classifier Mean Fit Time: 0.00401986903614468
```

5.2.3 Adjusting scores

Changed verbose to True and n_neighbors to 13, final results were:

```
KNN Classifier Train score1: 1.0
KNN Classifier Test score1: 1.0
```



7 Conclusion

Risk management in aviation maintenance needs to evolve from reactive methods and proactive methods to predictive-interactive methods. The Knowledge-Based Quality System is a methodology designed to collect data, integrate, analyze, and process the data collected from the Quality System and Safety System, presenting adapted solutions for the issues identified on the processes. The methodology integrates man and machine, using man creativity capacity and the ability to implement abstracts ideas and the machine, using the data analysis processing capacity helping humans to on the decision-making process, mitigating risk and preventing errors.

The KNN Classifier provided a good way to perform data analysis and predict when there will be a quality escape based on the features monitored on the service center. The KNN had great results providing high accuracy and a low mean time to fit.

This model should be used to detect early signs and triggers of the system and indicate where the weaknesses of the system are based on the issues reported and collected. The KBQS methodology determines where the weaknesses and vulnerabilities of the system are that needs improvement and how to improve applying interactive methods with the help of the predictive model that was developed.

7.1 Future Research

Regarding future research, KBQS can be further refined and adjusted to improve predictive capability. Additional data inputs from the system can be added and correlated to data that was previously corrected further adjusting risk weights and improving predictability results.
Other techniques such as Case Based Reasoning, NLP, Ensemble techniques, and Deep Learning techniques can be applied to the model.

Reference

Boden, M.A. ed., 1996. Artificial intelligence. Elsevier.

Chen, Y., Feng, W., Jiang, Z., Duan, L., Cheng, S., 2021. An accident causation model based on safety information cognition and its application. Reliab. Eng. Syst. Saf. Vol, 207 pp.-.
Claros, B., Sun, C., Edara, P., 2017. Enhancing safety risk management with quantitative measures. Transp. Res. Rec. 2603, 1–12.
Goh, Y.M., Chua, D.K.H., 2010. Case-based reasoning approach to construction safety hazard identification: Adaptation and utilization. J. Constr. Eng. Manage. 136, 170–178.
Leveson, N.G., 2016. Engineering a safer world: Systems thinking applied to safety (p. 560). The MIT Press.
Mendes, N., 2022. Risk management in aviation maintenance, Dissertation. UFSCar
Mendes, N., Vieira, J.G.V. and Mano, A.P., 2022. Risk management in aviation maintenance: A systematic literature review. Safety Science, 153, p.105810.
Mogles, N., Padget, J., Bosse, T., 2018. Systemic approaches to incident analysis in aviation: Comparison of STAMP, agent-based modelling and institutions. Saf. Sci. 108, 59–71.
Ratnayake, R.M.C., 2016. Knowledge based engineering approach for subsea pipeline systems' FFR assessment. TQM J. 28, 40–61.
Robinson, S.D., 2019. Temporal topic modeling applied to aviation safety reports: A subject matter expert review. Saf. Sci. 116, 275–286.