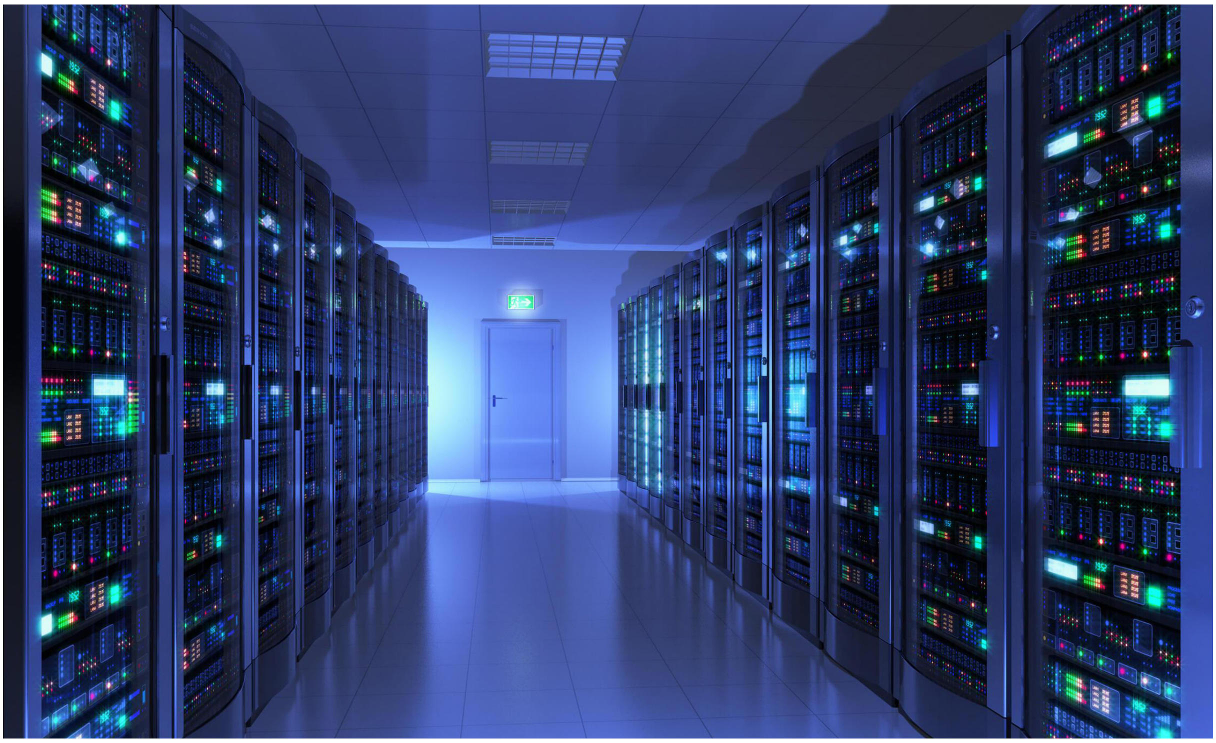
MACHINE LEARNING SOLUTIONS FOR EXCEPTION HANDLING

IBM is one of the biggest IT-companies worldwide. One of their activities is providing hardware solutions and offering service contracts to their customers. The Service Parts Operations department within IBM is responsible for keeping enough spare part inventory, to fulfil these contracts for the EMEA-region (Europe, Middle East and Africa). Their goal is to do this at minimal costs and to achieve this, many processes are automated by using the software tool Servigistics. This tool monitors the real-time status of the supply chain and automatically places orders, to keep the inventory level between certain levels, that are calculated based on historical demand. In the figure below, the black vertical line represents the current date and the other two vertical lines represent the lead times for different suppliers. The blue line is the Policy Safety Stock and the red line is the Maximum Inventory Level. An important function of this tool is to track whether the inventory level will be sufficient for the next two years.



SIMULATION MODEL FOR SCENARIO ANALYSES

When Servigistics is not able to keep the right amount of spare parts within the planning horizon, the exception handling process is triggered. Servigistics categorizes these exceptions into different review reasons and informs the planners by sending an alert. The planners have a look at the exceptions and decide what to do. In 2018 a total of about 35.000 exceptions were triggered for the EMEA-region. Currently, the performance of the decisions made by the planner on an exception is unknown. Justin Fennis, MSc student at the University of Twente, examined the possibilities of using (supervised) machine learning techniques to predict whether an action is required or not, with the goal to reduce the number of exceptions that planners usually handle.

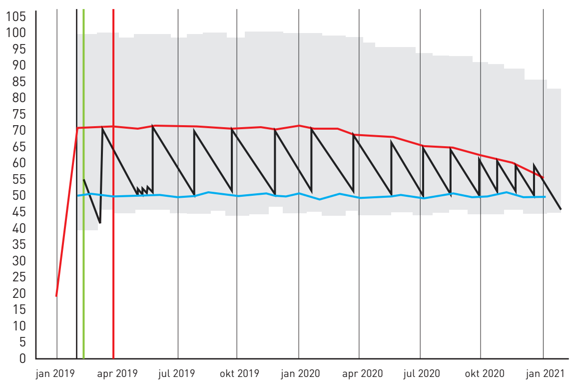


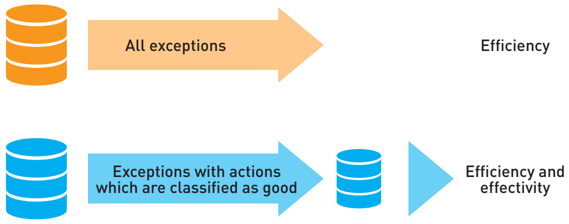
Figure 1

SOLUTION APPROACH

The choice was made to focus on four review reasons, which account for about 40% of the exceptions triggered per year. These review reasons are: projected stock out, stocked out, below must order point and projected inventory below must order point. The projected horizon is two years and the must order

point is the lower bound of the grey area, shown in the graph. For these review reasons, data the following were gathered over a period of about 2 months: the exceptions, the actions taken by the planners and the data on which the decisions have been taken.
From the original data, 43 characteristics (called features) were derived, which give information about the exception and the part connected to the exception. These features can be categorized as part characteristics, inventory information, order information, tactical settings and exception related.

Since the performance of the decisions of the planners is unknown, a performance indicator is introduced which qualifies the decision taken as good or bad, by looking at whether the initial decision by the planner resolved the exception. This performance indicator revealed that about 65% of the decisions taken are qualified as good. It is used to create two data sets: one containing all exceptions and one containing only the exceptions qualified as good. Using the first data set for building a model will only result in a more efficient process, whereas the second data set could also make it more effective.



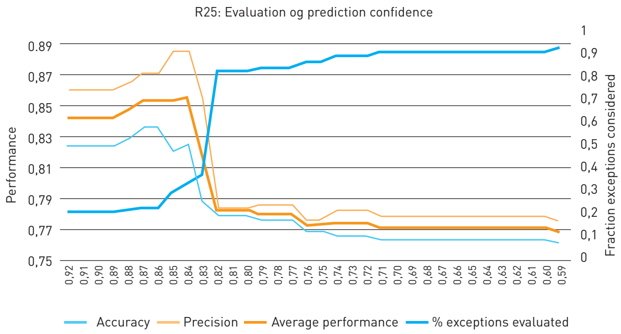
'Decision tree algorithms' appeared to be the most suitable machine learning technique for this research and a test revealed that the C5.0 algorithm available in SPSS Modeler performed best. To evaluate the performance, three performance indicators are used: (1) the accuracy, which measures the fraction or correctly classified instances, (2) the precision of the no action predictions, which is the fraction of correctly classified instances when no action is predicted and finally (3) the AUC which represents the ability of the model the make a distinction between classes. Since we are dealing with relatively small datasets, cross validation is used to evaluate the performance of the models. During cross validation, the original dataset is split into five folds, in which four folds are used to build a model and the remaining fold is used to evaluate the results. In this way, five models are built and the performance is evaluated based on the averages.

RESULTS

The average results per review reason are shown in the table below.

RR	Performance			
	Accuracy	AUC	Precision	S. Dev
Projected stock out	0,635	0,674	0,639	0,081
Stocked out	0,744	0,747	0,753	0,075
Below MOP	0,600	0,544	0,662	0,060
Projectes invetory below MOP	0,666	0,679	0,680	0,036

If we have a look at these results, we see that the results are promising, but the performance should increase before it can be a true planner assistant. Especially the limited size of the dataset has a significant impact on the results. This is observed in the stability of the results and the presence of some illogical patterns in the decision trees. However, if we use the performance indicator by only selecting the decisions which are qualified as good for the training set, the results are better. On top of that, if we use a threshold for the prediction confidence of the model, a significant improvement in the performance is achieved, which is shown in the graph below. Using a prediction thresholds leads to a lower fraction of exceptions considered (black line, right x-axis), but the performance increases (orange line, left x-axis).



FUTURE RESEARCH

Since the limited dataset has a high impact on the results, we first encourage to gather more data in order to create a bigger dataset. Next to that, we recommend involving the planners, which are located in India, in the process to give valuable information like important features that have an influence on their decision. Finally, an online learning would be suitable for this problem, which means that the model can be updated after every exception and interaction will be possible.

Topsector Logistiek

FACTS

Researcher	Justin Fennis
University	University of Twente
Supervisors	Dr. Matthieu van der Heijden
	Ir. J.P. Hazewinkel
Company	IBM

ProSeLoNext project – Powered by:

Gordian Logistic Experts
TU/e Technische Universiteit Eindhoven University of Technology
UNIVERSITEIT TWENTE.
Erasmus Universiteit Rotterdam
Tilburg University

This activity was co-financed by the Netherlands Organization for Scientific Research (NWO), on behalf of the Ministry of Infrastructure and Water Management.

6648