PREDICTION OF THE PARTS THAT MAY LEAD TO DELAY DURING THE PRODUCTION PROCESS WITH MACHINE LEARNING

Buray TÜRKMEN // Ramis Akın KORKMAZ // Fatma BULUT
Academic Advisor: Prof. İbrahim AKGÜN // Company Advisor: Berk KARAHAN

2017-2018 Academic Year

ABDULLAH GÜL UNIVERSITY

DEPARTMENT OF INDUSTRIAL ENGINEERING

ABSTRACT

This project, which is jointly executed with FNSS Defense Systems Inc., aims to avoid problems such as urgent shipping charges, operational problems and late deliveries of the final product of the company, due to the insufficient inventory for the parts which are necessary for the production of the ordered products and risky in the sense of supply.

Complex, irregular and large data sets have been pre-processed and made meaningful. It is determined that the company's dataset does not repeat each other and that it does not follow a certain pattern and it is foreseen that it will be wrong to make a statistical analysis. For this reason, classification models of machine learning techniques are used as a solution method. From these models, the Random Forest model estimated the decision variable with 97.86% success rate and improved the current estimation ratio by 75%.

Finally, a risk matrix has been created by using the lead times and the probability of shortages of the parts. As a result of these study, the production planning process of the company became transparent and the average cycle time and overtime costs decreased.

This project is supported by the application number of 1139B411702824 within the scope of TÜBİTAK's "2209-B Sanayiye Yönelik Lisans Bitirme Tezi Destekleme Programı".

PROBLEM DEFINITION

The main problem addressed in this project is defined as the lack of sufficient quantities in the inventory that are needed to produce the ordered product and that are risky in the sense of the duration of the supply. This problem also leads to different problems. Some of the emerged problems because of the main problem are; emergency shipping charges, operational disruptions resulting from attempting to supply parts quickly, and loss of reputation in the sector resulting from the company delivering its final product late.

SYSTEM DESCRIPTION

The system under consideration is summarized in the infographics below.

METHODOLOGY

The problem mentioned is to be solved by classifying the parts as high risk and low risk according to thepossibility that they are not enough for production and supply duration. In this context, firstly the dataset provided by the firm is defined, classification models are created and various operations to applied on the datasets to make them suitable for use in classification models.

Dataset Description

109 report files from the company's MRP (Manufacturing Resource Planning) program covering November 2016 - November 2017 arethe basis of this work.

Receive Information - Inventory Information - Order Information - Vendor Information - Characteristics of the Parts - MRP Reports (11/2016 – 11/2017 109 Reports)

OVER 90	OVER 18	OVER 122	OVER 31
UNSTRUCTURED VARIABLE	GB DATA	CSV FILE	MILLION ROWS

Data Preprocessing

109 files containing MRP output were first transferred to the Big Query database running on Google's servers. At this stage, a binary variable is created, which indicates whether the part is available in sufficient quantity in the company, by comparing the recorded status of the part with the corresponding quantities of the part on the date and unique part codes of the MRP output.

Operations are done using the "pandas, numpy, csv, glob and os" packages in the Python programming language. As a result of the code operations, a table composed of 410,956 rows with 131 binary variables and 10 numerical variants and missing information was obtained as an input of machine learning models.

MACHINE LEARNING CLASSIFICATION MODELS

It has been determined that the dataset of the study company does not repeat each other and that the dataset does not fit in a certain pattern. As a result, it is predicted that making a statistical analysis will not give healthy results. For this reason, classification models were established from machine learning techniques in the study.

RESULTS

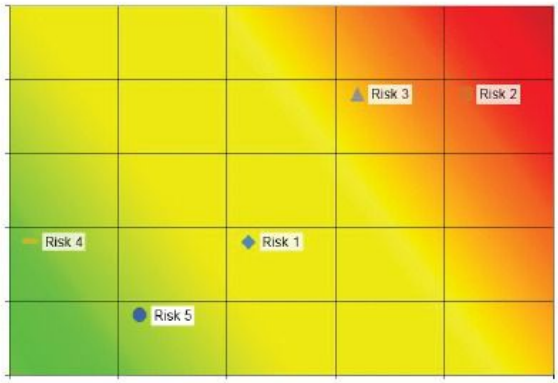
After applying the specified models to the regulated data set, the success rates were obtained. It is showed that Random Forest, Decision Trees and K-Nearest Neighbor models have resulted in the best estimation

Logistic Regression Analysis	K Nearest Neighbor	Support Vector Machine (Kernel)	Naive Bayes	Decision Trees	Random Forest	Artificial Neural Networks
%73.5	%96.7	%77.1	%67.7	%97.5	%97.9	%97.9

As a result of these studies, the company can estimate the availability of the part in the inventory with a success rate of 97.86%. According to these results, the company will achieve 75% improvement over the previous estimation success (56%).

A risk matrix is created by combining the probability that the part obtained are not in sufficient quantity in the inventory and the average procurement times.

The production planning process has become transparent and the physical and characteristic properties of the risky parts have been correlated after that this project. With the findings, it is expected that the reduction of the cycle time and the cost of the emergency cargo methods used for the project will decrease.



REFERENCES

Alsubaey, M., Asadi, A., ve Makatsoris, M. 2015. "A machine learning approach for predicting delays in construction logistics", International Journal of Advanced Logistics, 4(2), 115-130.

Bose, I., & Mahapatra, R. K., 2011. "Business data mining-a machine learning perspective", Information & Management, 39(3), 211-22.

Breiman, L. 2001." Machine Learning", 45(5).

Cortes, C., ve Vapnik, V., 1995. "Support-vector network", Machine Learning, 20(3), 273–297.

FNSS Savunma Sanayi. 2018. "FNSS Firma Profili", FNSS.

Martinez, V. 2012. "Flight delay prediction", ETH Zürich, Department of Computer Science.

Pearce, J., & Ferrier, S., 2000. "Evaluating the predictive performance of habitat models", Ecological Modelling , 225-245.

SPSS, 1999. "AnwerTree Algorithm Summary", SPSS White Paper, USA.

INDUSTRIAL ENGINEERING

FNSS

TÜBİTAK