**Citation :**

TY - BOOK, AU - Santis, Rodrigo ,AU - P. de Aguiar, Eduardo,AU - Goliatt, Leonardo

PY - 2017/11/08

T1 - Predicting Material Backorders in Inventory Management using Machine Learning

DO - 10.1109/LA-CCI.2017.8285684

**Data Description**

A backorder is the order which could not be fulfilled by the company. Due to high demand of a product, the company was not able to keep up with the delivery of the order. The backordering can lead to upsetting customer as they couldn't get what they ordered and the loyalty will decrease.

Also, company cannot overstock every product in their inventory to avoid such situation.

There has to be a way for the company to know for which products they can face this problem.

So, the company has shared a data file with different input features for each product and it hopes to find a pattern inside this data which can give them some insight.

The data file contains the historical data for some weeks prior to the week we are trying to predict.

The data has 23 columns including 22 features and one target column.

To model and predict the target, we’ll use the features columns, which are:

**sku** – Random ID for the product

**national\_inv** – Current inventory level for the part

**lead\_time** – Transit time for product (if available)

**in\_transit\_qty** – Amount of product in transit from source

**forecast\_3\_month** – Forecast sales for the next 3 months

**forecast\_6\_month** – Forecast sales for the next 6 months

**forecast\_9\_month** – Forecast sales for the next 9 months

**sales\_1\_month** – Sales quantity for the prior 1 month time period

**sales\_3\_month** – Sales quantity for the prior 3 month time period

**sales\_6\_month** – Sales quantity for the prior 6 month time period

**sales\_9\_month** – Sales quantity for the prior 9 month time period

**min\_bank** – Minimum recommend amount to stock

**potential\_issue** – Source issue for part identified

**pieces\_past\_due** – Parts overdue from source

**perf\_6\_month\_avg** – Source performance for prior 6 month period

**perf\_12\_month\_avg** – Source performance for prior 12 month period

**local\_bo\_qty** – Amount of stock orders overdue

**deck\_risk** – Part risk flag

**oe\_constraint** – Part risk flag

**ppap\_risk** – Part risk flag

**stop\_auto\_buy** – Part risk flag

**rev\_stop** – Part risk flag

**went\_on\_backorder** – Product actually went on backorder. This is the target value.

**Problem Statement**

1. Do a complete EDA in the python notebook file
2. Build a solution design architecture for end to end solution starting from data ingestion to deployment with a detail documentation.
3. Deploythe end to end automated solution to AWS.
4. Create a user interface for bulk testing uploaded through excel sheet and for a single record entry both.
5. Maintain log for each and every prediction request into any database
6. Create a low level documentation for end to end solution and deployment
7. Define a retraining approach in your documentation
8. Create an end to end video of the working of the project

**Evaluation Criteria**

1. Solution will be evaluated based on the best architecture design of the entire solution
2. Solution will be evaluated based on the best EDA
3. Solution will be evaluated based on the best optimization and model accuracy on unseen dataset
4. Solution will be evaluated based on final deployment and testing with a bulk dataset as well as data with single entry
5. Solution will be evaluated based on the best readability and proper commenting of the code.
6. Solution will be evaluated based on best interactive user interface

On the basis of the given data in the features column and target column,

you have to build a model which will be able to predict whether an order for a given product can go on backorder or not.