Project: Predicting the backorder in Global Supply Management

Material backorder is a common supply chain problem, impacting an inventory system service level and effectiveness. Identifying parts with the highest chances of shortage prior its occurrence can present a high opportunity to improve an overall company’s performance. In this competition, you will train classifiers to predict future backordered products and generate predictions for a test set.

The evaluation metric for this competition is [AUC (area under the curve)](https://www.kaggle.com/c/www.dataschool.io/roc-curves-and-auc-explained). The AUC (Area Under Curve) is the area enclosed by the ROC curve. A perfect classifier has AUC = 1 and a completely random classifier has AUC = 0.5. Usually, your model will score somewhere in between.

**Data Description**

File descriptions

* train.csv - the training set
* test.csv - the test set
* sampleSubmission.csv - a sample submission file in the correct format

Data fields

* sku -sku code
* national\_inv- Current inventory level of component
* lead\_time -Transit time
* in\_transit\_qtry - Quantity in transit
* forecast\_x\_month - Forecast sales for the net 3, 6, 9 months
* sales\_x\_month - Sales quantity for the prior 1, 3, 6, 9 months
* min\_bank - Minimum recommended amount in stock
* potential\_issue - Indictor variable noting potential issue with item
* pieces\_past\_due - Parts overdue from source
* perf\_x\_months\_avg - Source performance in the last 6 and 12 months
* local\_bo\_qty - Amount of stock orders overdue
* X17-X22 - General Risk Flags
* went\_on\_back\_order - Product went on backorder
* Validation - indicator variable for training (0), validation (1), and test set (2)

**Description**:

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.