



SS

8BY8

RioTinto

30%

Malfunctioning
Sensors

\$5.1

Billion EBITDA

800

Trucks currently
in use

630

Thousand
tonnes lost
capacity

Fuel Consumption



Efficiency



Time

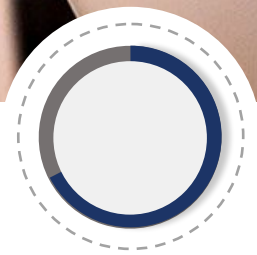




Novel Insight

Original approach to big data
issues

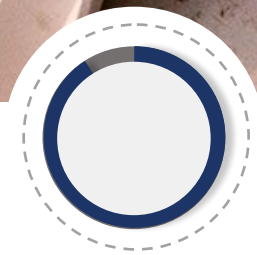
+



Relevant Data

Relevant, real-time data feed

+



Innovative Model

Innovative and original approach
to big data issues

=



Accurate Prediction

Innovative and original approach
to big data issues

1,200km

Huge Range

96%

Accurate
Predictions

01 Preprocessing

```
import xgboost as xgb
from sklearn.metrics import roc_auc_score
from xgboost.sklearn import XGBClassifier
from sklearn import cross_validation, metrics
from sklearn.grid_search import GridSearchCV
import numpy as np
from functools import partial
from sklearn.base import BaseEstimator
import gc

class SelfTunedXgboost(BaseEstimator):

    """
    """
    def get_params(self, deep=True):
        return self.current_params

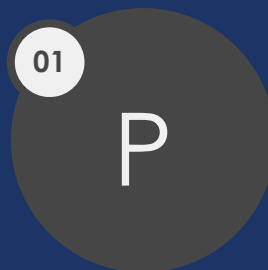
    """
    """
    Grid = partial(GridSearchCV, scoring='roc_auc')

    initial_parameters = {
        'learning_rate': 0.1,
        'n_estimators': 300,
        'max_depth': 3,
        'min_child_weight': 6,
        'gamma': 0,
        'subsample': 0.8,
        'colsample_bytree': 0.8,
        'objective': 'binary:logistic',
        'nthread': 1,
        'scale_pos_weight': 1
    }

    param_tune_list = [
        {'max_depth': list(range(4, 11)),
         'min_child_weight': list(range(3, 7)),
         'gamma': [i/10.0 for i in range(1, 10)],
         'subsample': [i/10.0 for i in range(1, 10)],
         'colsample_bytree': [i/10.0 for i in range(1, 10)]
        }

    def __init__(self, current_params = None, verbose = 0):
        self.verbose = verbose
        self.current_params = current_params
```

02 Point in Time



Point in Time

Identifying instantaneous probability of carry back



03 Aggregate

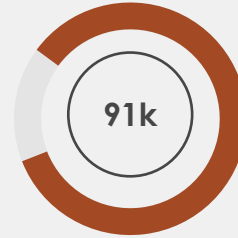


Sliding Window

A robust method for aggregating individual decisions.

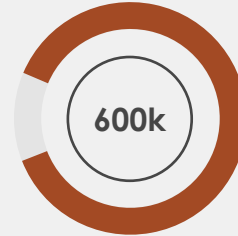
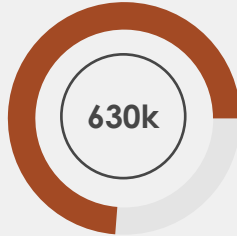


Carryback
On average for Site 1.



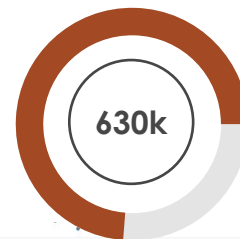
Trips Without Sensors
Just under 26% of all trips lacked sensor information.

Tonnes Lost Capacity
Expected amount of carryback not identified by current system.

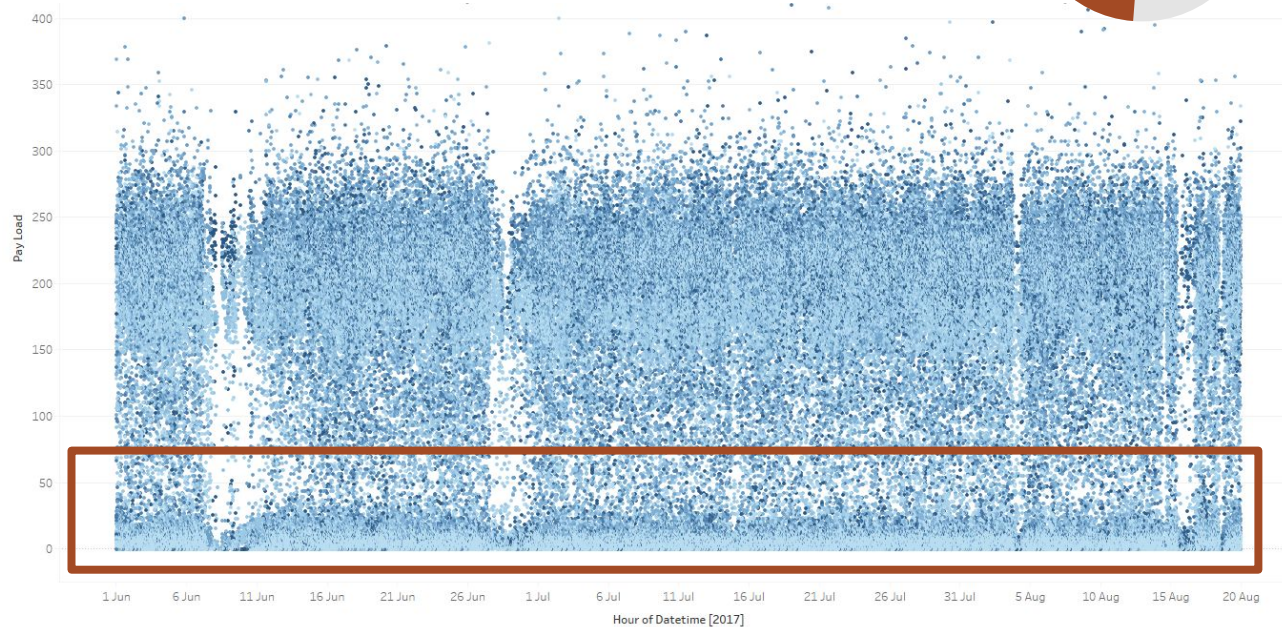


Additional Tonnes of
Carryback

Before

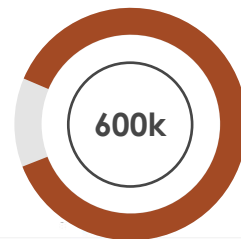


Tonnage Lost Capacity

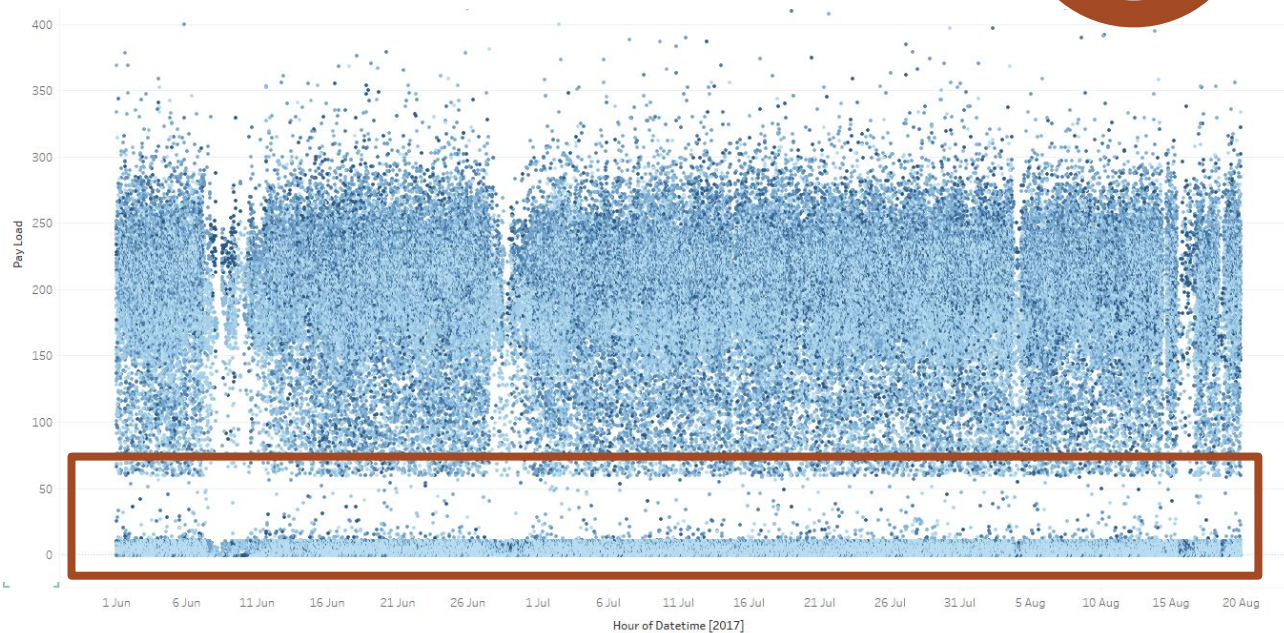


January - August
2017 Site 2

After



Identified Carry Back



January - August
2017 Site 2



HQ Feedback

Direct implementation into HQ systems



Using Existing Sensors

No new sensor installations needed, works with existing infrastructure



Real Results

High degree of accuracy in a huge network of vehicles



Ready Already

Prepared for integration into Rio Tinto's network



THANK YOU



DataRobot

RioTinto

