Last year we wrote a paper regarding artificial intelligence (AI) on how it had made, and was continuing to make its way into the terminal industry. We paid specific attention to how machine learning (ML) as a branch of AI could be implemented. In a paper that likened AI’s current state to that of Frankenstein, the article closed by saying that AI was coming, and that as an industry we can either be prepared or caught off-guard when it does. For INFORM, as a leading AI solution provider, the question wasn’t how to prepare for AI, but rather, how we can leverage the promise of ML and build it into our core AI driven solution.

As such in 2018, INFORM undertook an ML assessment project looking at maritime container terminals and how ML could be used to improve optimization and operational outcomes. The assessment aimed to achieve two results. Firstly, could INFORM’s broader ML algorithms, developed for use in other industries such as finance, be applied to our Optimization Modules that are used in terminals around the world. Secondly, if they could be, does that mean we can apply them to real-world terminal data and identify areas where improvements could be made to parameters that influence the optimization calculations of INFORM’s add-on Optimization Modules.

Working with a randomized sample of 1 million containers handled in the 2017 calendar year with 50 data variables (explanatory variables) at a selected terminal, we set off to answer these two questions. The dataset was split further; using a time slicing method, a training dataset (75% of the dataset) and a testing dataset (25% of the dataset) were created in accordance with good ML practices. Further, we worked with a human expert to review and identify variables amongst the 50 explanatory variables that would prove meaningful in the assessment. We identified 16 variables that have been used to build the random forests ML models...
presented later in this paper (see figures 2 and 6).

In short, the answer to the first question was a resounding yes; INFORM’s ML algorithms could be applied to work with our solution for the maritime terminal industry. To answer part two, it is worth exploring some of the areas of terminal optimization where it was identified that we could further improve our solution offering through the implementation of ML. While we identified many areas, we will focus on container outbound mode of transport and dwell time in this paper.

MORE INFORM TECHNICAL PAPERS

PREDICTING OUTBOUND MODE OF TRANSPORT

We started with the outbound mode of transport predictions. The mosaic plot (figure 1) shows what was expected based on the TOS information received upon container arrival versus how the container actually departed the terminal. The areas represented are proportional to the volume of containers handled. “NA” areas in grey, were unknown to the TOS when the container arrived. The data the TOS is configured to use to make decisions is accurate in only 62.9% of the total 1 million boxes sampled.

A random forest with 500 trees ML model was trained using the training dataset and subsequently tested against the testing dataset. The forest was tasked with identifying the importance of the pre-identified 16 variables (see figure 2).

Random forest ML models are very flexible algorithms that produce balanced results for classification and regression tasks. They are created by building a multitude of decision trees, in our case 500 trees, and then outputting the mode (classification) or mean (regression) of the individual trees. Importantly, random forests correct for a single decision trees’ tendency to ‘overfit’ to the training data. Finally, random forests improve upon the predictive power of single decision trees by making clever use of random chance.

THE FINDINGS

Using the revised ML generated prediction model opposed to the data available from the TOS upon container arrival would increase prediction accuracy to 83.6% - corresponding to a relative improvement in prediction accuracy of 33%. Figure 3 maps the improvement in accuracy from the TOS to the ML model against each outbound mode of transport (OMT). Looking into the data more closely, the TOS had a good accuracy at predicting OMT for containers leaving by ship (81.9%), average accuracy...
for truck OMT (65.5%), and poor accuracy for feeder (42.4%) and rail (4.7%). In comparison, the accuracy for all OMT shifts to average or much better: ship (94.3%), truck (87.2%), feeder (76.3%), and rail (53.0%).

**PREDICTING CONTAINER DWELL TIME**

Container dwell time is used within INFORM’s Optimization Modules to assist with container yard positioning calculations. The basic logic is straight-forward: when building stacks in your yard, place containers with longer dwell times at the bottom and containers with shorter dwell times at the top. In this way, you minimize the number of rehandles needed to retrieve containers for their outbound journey.

Given its relevance, dwell time is a central variable in optimizing the placement of containers in one’s terminal. However, the data point used to calculate dwell time — expected departure time — is frequently missing. In our dataset, 47% of containers were missing an expected departure time. This is visually represented in figure 1; data was available for ship (purple) and feeder (blue). Traditionally, for these instances, our optimization modules use a strategically calculated and pre-configured dwell time variable as a stand-in.

Working with the dataset, we drew up an empirical model to determine the mean dwell time for loaded containers where there was no expected departure time upon container arrival. The mean was calculated at 84 hours (see figure 4). Not too far off of the selected systems pre-configured 96 hour stand-in variable, so using basic statistical modelling, we can already achieve a small improvement.

From there, we decided to see what would happen when we factor in the dwell time versus the expected departure mode. A different picture emerged, see figure 5.
Containers leaving by ship stay longer, while containers leaving the terminal by truck and rail have, on average, a significantly shorter dwell time. There was a correlation between the complimentary storage duration offered by the terminal and the associated outbound mode of transport; that said, it was not the aim of this assessment to evaluate this finding and further assessment is needed to confirm causation vs association.

Working from the hypothesis that a better stand in dwell time could be predicted if the system took into consideration the OMT, a random forest with 500 trees ML model was trained on the training dataset and subsequently tested against the testing dataset. Again, the forest was tasked with identifying the importance of the pre-identified 16 variables (see figure 6).

Interestingly, weekday proved to be highly relevant. The ML model found that containers arriving on Thursday or Friday were likely to remain longer than those arriving Saturday through Wednesday. Our human expert attributed this to reduced operational hours over the weekend period.

Using the revised ML generated prediction model for dwell time instead of the standard dwell time variable resulted in relative improvement in prediction accuracy of 26.8%.

Figure 6

OPPORTUNITIES FOR FURTHER ASSESSMENT

Future assessments should aim to review available data from 2018 and run it against the same process to assess whether the findings from the 2017 data are consistent with more current data or what alternative patterns are seen. Further, considering external data sets could add additional insights beyond that of the core container data. For instance, vessel ETA vs ATA and ETD vs ATD differences could reveal additional patterns that would improve the dwell time prediction model. It is reasonable to assume that seasonal (monthly or quarterly) differences are also plausible due to container traffic patterns or the impact of weather.

APPLICATION IN TERMINAL OPERATIONS

It is expected that the ML models will lead to improved container location selection (OMT) and improved precise stack location (dwell time) both resulting in fewer re-handles. Let’s assume you can reduce re-handles by 1% in our example dataset of 1 million containers; that is 10,000 fewer moves. Further, let’s assume that the re-handle cost in a typical EU terminal is approximately 80 Euros, this would result in an annual savings of 800,000 Euros (US$900,000) for every 1% decrease in re-handles.

Using the Machine Learning Module within INFORM’s optimization solution, the application of ML to review and improve predicted outbound mode of transport, or OMT, and container dwell times should be run on a regular basis. Further, the output should be reviewed by a human expert before being used to modify optimization parameters. As noted in our previous paper, this expert discussion review process will assist operators in understanding the changes made inside of the optimization systems. Secondly, it will allow human operators to learn from the output as well as gain confidence in the ML Module’s ability to recommend appropriate parameter improvements for future use.

REFERENCES


ABOUT THE AUTHORS

Dr. Eva Savelsberg is Senior Vice President of INFORM’s Logistics Division. She specializes in Agile Optimization Software that renders a wide range of terminal processes more productive, agile, and reliable. Eva is also lecturer at the University of Aachen (RWTH), where she received her PhD in mechanical Engineering in 2002. Eva has published 5 books and over 35 papers on innovation in freight transportation.

Dr. Ulrich Dorndorf, Chief Technical Officer at INFORM, has many years of practical experience in developing and implementing optimization algorithms for transport, logistics, supply chain management, and production planning.

Matthew Wittemeier is responsible for Marketing and Sales of INFORM’s Logistics Division. He brings over twelve years of marketing experience from a breadth of industries including aviation, creative, finance, and software services. He holds a Bachelor of Management and Professional Studies from Southern Cross University in Australia.

ABOUT THE ORGANIZATION

INFORM specializes in Agile Optimization Software to improve operational decision making. Based in Aachen, Germany, the company has been in the optimization business for nearly 50 years and serves a wide span of logistics industries including maritime and intermodal terminals.

ENQUIRIES

www.inform-software.com