Combining Machine Learning With Human Knowledge for Delivery Time Estimations

Markus Lochbrunner, Hans Friedrich Witschel

FHNW University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, CH-4600 Olten

Abstract

Although machine learning algorithms outperform humans in many predictive tasks, their quality depends much on the availability of sufficient and representative training data. On the other hand, humans are capable of making predictions based on "spontaneous" transfers of knowledge from other domains or situations in cases where no directly relevant experiences exist. This can be seen very well in the task of predicting lead times in goods transport, where sudden disruptions or shortages may occur that are not reflected in historical data, but known to a well-informed human. If the variation can be anticipated and more accurate lead times estimated, proactive measures can be taken to decrease the impact.

Therefore, we describe three novel approaches for delivery time predictions, combining a machine learning model with human input. The proposed logic covers two phases, learning based on actual delivery data and capturing human knowledge to cover exceptional situations not reflected in historical data. The proposed models and the resulting estimates were evaluated using deliveries from a retail company. It was found that the pure machine learning model delivers better results than a combination of humans and machines. On the one hand, this is caused by the complexity of incorporating human knowledge into the algorithm in a suitable way. On the other hand, it is also due to the tendency of humans to over-generalise the impact of certain events. Thus, although the pure machine learning model delivers superior estimation accuracy than the human-machine combination, our systematic qualitative analysis of the results presents insights for future development in this area.

Keywords

lead time estimation, regression, machine learning, knowledge engineering

CEUR-WS.org/Vol-3121/paper7.pdf

1. Introduction

Machine Learning (ML) techniques have been used for a long time and with great success for the prediction of both categorical and numerical outcomes in a wide range of application areas [1]. The respective algorithms identify patterns in historical data and extrapolate them to the future. For several types of patterns, it has been shown that ML systems are able to outperform humans, see e.g. [2].

However, ML approaches tend to provide rather poor results when only insufficient training data is available [3] – this can happen for various reasons, e.g. because gathering such data is expensive, data is skewed and thus under-representing certain rare events or because data

In A. Martin, K. Hinkelmann, H.-G. Fill, A. Gerber, D. Lenat, R. Stolle, F. van Harmelen (Eds.), Proceedings of the AAAI 2022 Spring Symposium on Machine Learning and Knowledge Engineering for Hybrid Intelligence (AAAI-MAKE 2022), Stanford University, Palo Alto, California, USA, March 21–23, 2022.

markus.lochbrunner@students.fhnw.ch (M. Lochbrunner); hansfriedrich.witschel@fhnw.ch (H. F. Witschel)
https://www.fhnw.ch/en/people/hans-friedrich-witschel (H. F. Witschel)
0000-0002-8608-9039 (H. F. Witschel)

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

has to be gathered in dynamic environments where patterns change quickly. The concept of *transfer learning* [4] can partially alleviate this problem by adapting data from related domains to new problems.

However, although ML approaches can outperform humans in many tasks, they do not have the ability to do *spontaneous* transfers: one has to carefully plan and prepare transfer learning - whereas humans are great in spontaneously transferring experiences between situations and domains.

One particular strength of humans is the ability to recognize when historical patterns are becoming invalid. In this paper, we will use the transport of goods and the prediction of delivery times as an example of an environment where patterns and contexts can change dynamically: although delivery times depend on a number of factors that machines can learn, there may be sudden disruptions, congestions or staff shortages that invalidate the learned patterns. We hypothesise that humans are able to pick these up by reading news (consider e.g. the blockage of the Suez channel in spring 2021), being in contact with other logistics providers to know about potential staff shortages etc.

Several scholars have proposed combinations of machine learning with human knowledge (see [3] for an overview), in order to combine the strengths of both. In our work, we want to apply some of these combinations to the problem of estimating delivery times in goods transportation. Our main focus will be on a systematic qualitative evaluation of the problems and benefits that such combinations imply. Although our results are not consistently favourable for an ML-Human combination, we believe that a better understanding of the respective strengths and weaknesses of both will contribute to developing better combinations in the future.

2. Background and Related Work

2.1. Transport Lead Time Context

The bridging of space for goods is referred to as goods transport. A transport system consists of the goods to be transported, the means of transportation and the associated process [5]. The process can be differentiated into internal (transport within a factory) and external transport (transport from one geographical location to the other) [5]. In the present case, only the latter will be dealt with.

According to [6], the physical process of transporting goods between countries by truck, ship, rail, or intermodal is referred to as international freight transportation. Freight transportation can include several stakeholders for each movement, including one or more shippers, carriers, forwarders, third-party logistics providers, and customs for international flows. Table 1 shows a selection of work in this area. Each transport mode has its challenges, especially regarding the number of different transport legs and the connected uncertainties.

2.2. Machine Learning in Lead Time Estimation

As the overall accessibility to data has increased over the last couple of years, researchers and commercial companies spotted the opportunity of benefits that machine learning models could bring for goods transportation.

Transport Mode	Reference	Input Data	Model
Road	[7]	Historical travel time data, weather data, traffic data	Decision trees, clustering techniques, support vector machines
	[8]	Historical travel time data, google maps data	Gradient boosting regression tree
Intermodal	[9]	Historical travel time data, weather data, traffic data	Random forest
	[10]	Terminal data	Adapted queuing theory
Ocean	[11]	Historical travel time data, real-time tracking data	Classification and regression tree algorithm
	[12]	Historical travel time data, weather data	Neural network, support vector regression

Table 1
Overview of Lead Time Estimation approaches

Algorithm Type according to [13]	Algorithm Name	Examples in Delivery Time Estimation
White-Box	Decision Trees	[14, 15]
	Bayesian Networks	[16]
Black-Box	K-Nearest Neighbours	[17, 18]
	Artificial Neural Networks	[19]
	Support Vector Machines	[20]

Table 2
Machine Learning Paper Overview

In our context, the aim of machine learning is to predict the delivery time in days, a regression task. White-box models, such as decision trees, are self-explanatory with regard to their mechanisms and the decisions they make – something that makes a combination with human knowledge easier. With black-box models such as neural networks, it is usually not possible to understand the model due to their interdependence and complexity [13]. Table 2 shows an overview of the different algorithms and a categorisation according to [13], whether they are white or black-box models and thus allow direct interaction with humans. Furthermore, related works are shown that have applied these algorithms in the field of delivery time estimation.

2.3. Combining Machine Learning and Human Knowledge

According to [21], systems that can learn from end users have rapidly gained popularity. Until recently, this development was primarily fuelled by the importance of domain knowledge for setting up machine learning algorithms. However, an increasing number of researchers recognise that it is not only the feature selection and construction of machine learning models that human input can significantly benefit the performance of systems. The authors also mention that plenty of systems that transform data into computational models only involved domain experts during the development phase. Nevertheless, there are also examples where the human factor is directly integrated into the algorithm. [3] describe how existing knowledge can be integrated into the machine learning process. The model designed is called «Informed Machine Learning», and it describes how existing knowledge can be incorporated into the design of the machine learning pipeline itself, feature engineering, or the pre-processing of the training data. [3] start with the knowledge source, which qualifies the origin of the knowledge. Figure 1 describes different ways of representing human knowledge, as well as its possible integration into the machine learning algorithm. The latter can take place in four different ways. For certain use cases, also combinations of these are possible [3]:

1. Influencing the training data (feature engineering or additional data sets).
2. Hypothesis Sets: The integration takes place via selecting the appropriate hyper parameters or selecting suitable algorithms.
3. Learning Algorithm: A loss function is integrated, which inserts the existing knowledge about algebraic equations.
4. Final Hypothesis: The prediction output of a learning pipeline, also called the final hypothesis, can be compared to established information. Predictions that do not comply with established constraints, for example, may be discarded or marked as suspect, ensure the findings are compatible with prior knowledge [3], e.g., for simulations with constraints, the final output can be cross-checked against an established rule set to comply with certain rules and standards.

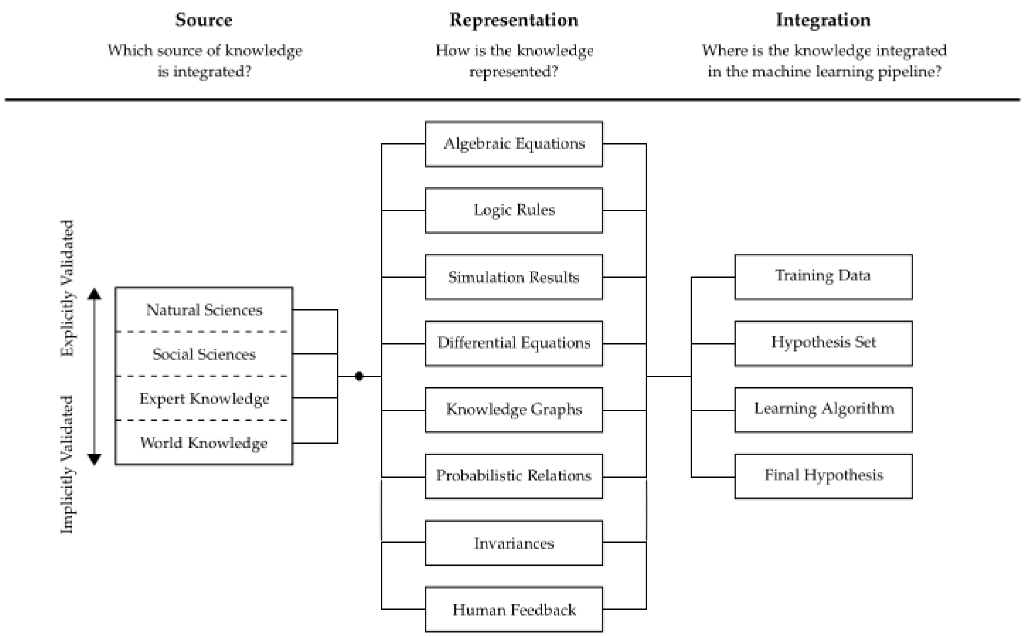


Figure 1: Taxonomy of Informed Machine Learning [3]

2.4. Contribution

The research regarding transport lead time estimations has shown that a large part of the work deals with last-mile deliveries. In general, the literature focuses on the transport mode "road" in the urban sector, while the authors [9] and [10] deal with intermodal transport chains. An investigation of multi-link transport using a combination of machine learning and human knowledge has not been undertaken in the literature to date. The estimation methods used are widely spread in the different papers studied. The literature focuses on the use of historically collected internal and external company data, which is used as the basis for machine learning. The knowledge collected by humans, which has often been learned over many years by experts in everyday life, is usually only used to develop machine learning algorithms. In contrast, the present work considers an extended approach, which also wants to use this knowledge during the actual prediction.

Besides, we see the systematic qualitative analysis of errors made by both ML models and humans as a key contribution of our work (see Section 5.2): it will enable us to understand how future ML-human combinations should be designed in order to avoid the imprecisions that our approaches still exhibit. Also, the approach is not limited to one transport mode but includes all different land, sea and air options.

3. Manual Lead Time Estimation Process

We describe several possible combinations of ML and human knowledge, all of which will be applied to the data of goods transports issued by a large retailer. To understand the current process of lead time estimation, we have conducted interviews and observations.

Since the challenge of delivery time estimates lies in multi-stage transport chains, the components and their challenges are examined more closely in this section, before describing the observed estimation procedures. For the delivery, the goods are loaded into the loading units on paper or Euro pallets and then transported by truck, vessel, train or, in very exceptional cases, by air. Consequently, with the exception of direct trucks from a sender to the end-receiver, the goods are handled in ports, consolidation points and transhipment terminals. Therefore, the delivery lead time can be divided into transport and handling times, as shown in Figure 2.

The co-worker responsible for the delivery estimation needs to gather the data from different systems and stakeholders in the company. Afterwards, the co-workers need to average the times of sub-components (valid for the various legs of the transported route), including historical waiting and handling times and the nomination shares of carriers with different lead times.

In the next step, the co-workers were observed while making the actual prediction. Compared to the instructions, the co-workers' actual approach was different. Instead of using the contractually agreed lead time components as the basis for calculating the lead time, the average of the last weeks is used as the basis for the estimate. Furthermore, the current process of information gathering is manual. Human mistakes can significantly impact estimations,

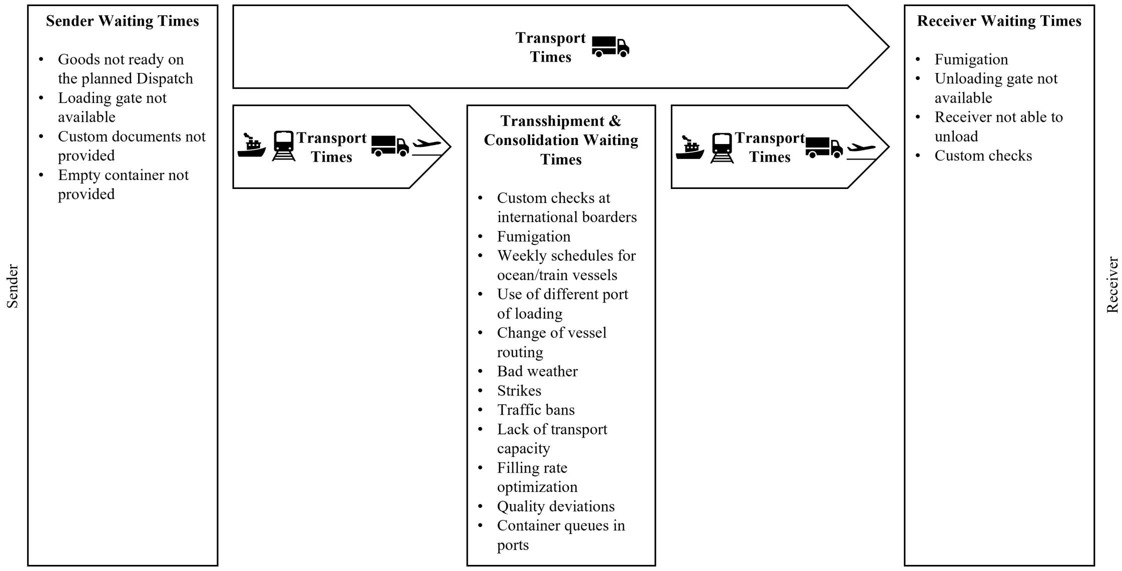


Figure 2: Waiting and Transport Times in the Supply Chain (derived from internal documents of the use case company)

especially on complex routes with different transport modes and waiting times. Despite having experienced co-workers in the team, the sheer number of senders and receivers makes it difficult to verify each route manually. Therefore, the goal is to develop a model which combines machine learning-based historical learning with the pro-active knowledge of the co-workers in the next chapter.

4. Models Combining Machine Learning and Human Knowledge

4.1. Data Preparation

Through the findings gained from observations and data mining results, waiting times and distance have the most significant impact on the estimated time and were, among others, used to extract features, which can be seen in Table 3. The source of the data is indicated in the feature source column. In the present work, all possible features were included in the experiments initially, and then Sequential Backward Elimination (SBE) [22] was applied. To further increase the accuracy, Sequential Forward Selection (SFS) was performed additionally. Thus, the data shown in Table 3 represents the final features for the lead time estimation model.

Over 60 different runs were performed in Azure Machine Learning. First, all regression algorithms available in Azure Machine Learning (see [23]) were included in the experiment. This was followed by an optimization of the parameters in order to adapt the model to the sample data provided by the retailer. By parallelising the runs, a large number of different algorithms could be tested per run. The best performing model was the XGBoost regression algorithm, which belongs to the family of tree algorithms.

Numerical or Categorical	Feature Source	Feature Name
Numerical	Consignment	Distance
	Shipment	Mid Receiver close to Sender Waiting Time
	Shipment	Port of Loading Waiting Time
	Shipment	Port of Discharge Waiting Time
	Shipment	Delivery Lead Time
Categorical	Consignment	Mid Receiver close to Receiver
	Consignment	Dispatch Month
	Shipment	Mid Receiver close to Sender
	Shipment	Main Carrier
	Shipment	Port of Loading
	Shipment	Port of Discharge

Table 3
Extracted Features

4.2. Model Design

In the next step, possibilities were evaluated as to whether the machine learning pipeline delivers better results when enriched with human knowledge. The manual lead time estimation described in section 3 and the pure ML model from the previous section provides the baseline, which we compare to three different alternative models described in this section. The basic approach of the procedure is shown in Figure 3.

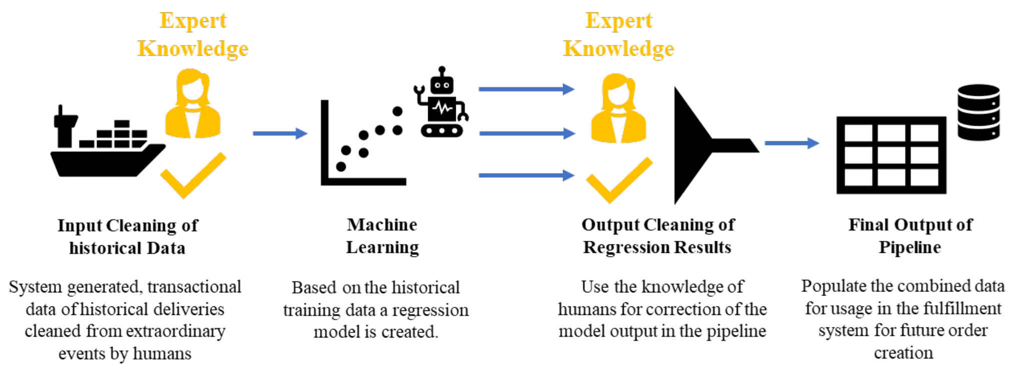


Figure 3: Pipeline Overview for the Lead Time Estimation Model

The lead time estimation model to be developed consists of two main components: A machine learning algorithm based on cleaned historical deliveries and an additional expert input that uses human knowledge to improve the results of the machine estimation. The approach involves modifying the machine learning output since the chosen algorithms are not directly interpretable by humans but only via explainability packages like SHAP [24] or LIME [25]. In the present case, SHAP was chosen as an explainer due to its good integration and strength to explain the tree based XGBoost algorithm.

In the following three sections, three different approaches are presented in more detail.

4.2.1. Model A

In the first approach, only the values of certain features are modified by humans (see Table 3).

More precisely, the XGBoost algorithm is trained with the historical shipment and consignment data, with two possible modifications by human transport planners:

- *Before training the model*, training data can be manually cleaned of unusual events (e.g. a significant disruption in international container traffic). For example, some historical delivery times were significantly larger than usual because ships could not call at the port in Shenzhen (Yantian) because of tropical cyclone "Kompasu" [26]. Removing those training data prevents one-off events from falsely influencing future estimates – with a risk of removing events that are actually not one-off (e.g. cyclones happening in almost regular intervals in certain regions).
- *Before applying the trained model*, the values of the port waiting times can be manually corrected. For the port waiting times, an external data set with current port waiting times serves as a support tool, but it only gives a direction and is subject to change by the transport planners to reflect current trends or events.

4.2.2. Model B

In the second approach, as in Model A, the adjusted waiting time is also included, and in addition, deviations known to the transport planner are recorded in an exception table. The exception form allows to capture lead time corrections for specific sender-receiver combinations. Optionally, transport planners can specify a reason for the correction, e.g. "Driver shortage for Heavy Goods Vehicles (HGV) in the UK" – which is not used by the regression algorithm.

The exceptions are applied at prediction time, i.e. corrections are added to or subtracted from the output of the regression algorithm , thus including any anomalies that are not reflected in the historical data.

4.2.3. Model C

Like Model A and B, the third model includes modification of waiting times. However, as opposed to Model B, the information is not captured in advance in an exception form. Instead, the machine estimate is compared to the average last four weeks' actual delivery times. A flag is set in cases of large deviations, allowing the planner to look at the sender-receiver connection in more detail, adjust the estimated value if necessary, and note a reason for the adjustment in the output. It should be mentioned that this approach identifies many exceptional cases only with a certain time lag. For example, a railroad company could announce construction work on a route, which would extend the delivery time due to a train detour. This proactive input would only be considered several days or even weeks later (maybe too late) – namely, when there are already recent deliveries that show a significant deviation from the machine-estimated value. However, letting transport planners examine and correct *all* connections (not just the "suspicious" ones) is not feasible because of their large number.

5. Evaluation

5.1. Quantitative evaluation

While we trained our model on historical data, the tests were performed on “live deliveries”: we asked transport planners to share their delivery estimates for 425 sender-receiver relations, which they made according to their usual procedure, see Section 3, thus forming our “Model Baseline”. We then ran all other models, again asking the transport planners for their input to Models A, B and C, and recorded their predictions. Finally, in the period from August 15 to September 15, we were then able to observe 5843 shipments on 330 of the 425 sender-receiver relations. Their delivery times formed the ground truth for our evaluation.

Figure 4 shows the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) of the purely human estimates (“Model Baseline”), the pure ML model and the three combined models introduced in the previous section.

	Model Baseline	Model ML	Model A	Model B	Model C
MAE	3.770	3.513	3.663	4.025	5.031
RMSE	5.347	5.161	5.112	5.574	7.313

Figure 4: Evaluation of the different models using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)

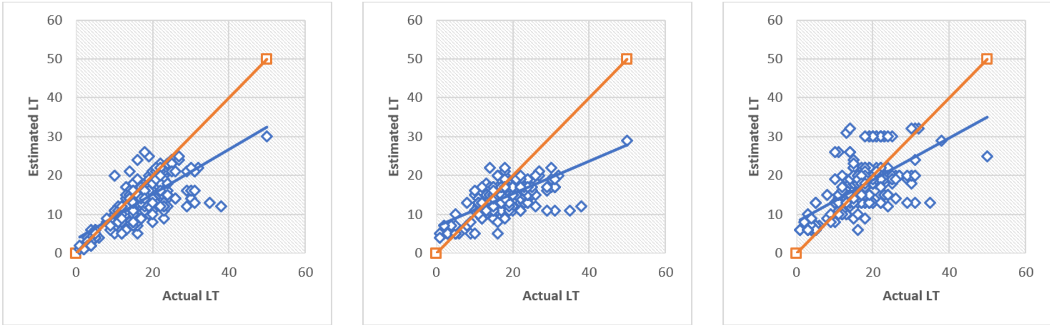
We additionally performed a Wilcoxon signed rank test to determine statistical significance between each combined model and the two baselines. The result showed statistically significant deviations between all models – A, B and C – and the human estimates (Model Baseline), whereas all differences between the three models and the ML baseline were not statistically significant at a confidence level of 95%. This is probably because estimates of the combined models largely follow the primary direction of the ML estimates.

Thus, both the ML approach and Model A outperform a purely manual estimation, whereas Model B and C lead to a significant deterioration of estimation results. All differences are rather small when averaged. However, we will see later that rather large differences can occur locally.

Looking more closely into the types of errors that each method makes, Figure 5 shows actual vs. estimated lead times for the two baselines and the most successful combination, Model A. Data points below the orange line indicate an underprediction of lead times.

We can see that these occur more frequently for the two baselines, whereas Model A does not have such bias in its errors. After considering the underlying training data, the bias of the ML model could be caused by increasing delivery times over the last months. Since the model learns from historical data and the delivery times were shorter in the beginning of the year, they are also estimated too short. However, the current situation implies longer delivery times. Therefore, the development can be seen as a COVID-19 consequence in 2021.

In the combinations of humans and machines, attempts are made to counteract this bias. Later, we will see that there is a risk of over-generalisation in the human interventions. Specifically, the high deviation in models B and C is also because planners grouped certain sender-receiver



(a) (b) (c)

Figure 5: Scatter plot of actual vs. estimated lead time for (a) Model Baseline, (b) ML and (c) Model A. The orange line has a slope of 1, i.e. represents correct predictions.

relationships in a lump-sum way. For example, during the estimation, it was assumed that a rail strike in the transit country Germany would affect the flow between Italy and Sweden. Therefore, an increase in the delivery time by five days was calculated. In fact, only a few suppliers were affected for a short time by the strike and not the entire flow, which affected the estimate's accuracy.

Another systematic mistake that the ML model makes is the overestimation of very short actual lead times – since these are exclusively Chinese senders and receivers, it can be assumed that the model has another bias, as for most destinations, suppliers from China have a longer delivery time than the rest of the world.

5.2. Qualitative analysis

For a qualitative analysis of errors, we used stratified sampling, binning errors with a step size of 2 days, and then taking random samples of connections, with a size according to the number of connections in the bin, considering only bins with absolute errors of at least 4 days. We then analysed the root cause of the sampled errors.

5.2.1. Errors of the ML model

In a nutshell, the ML model suffered from the following categories of problems (numbers in brackets indicate the number of connections in our sample affected by the problem):

- (5) Short-noticed deviation from the planned route due to operational circumstances, e.g. in case of spontaneous bottlenecks
- (3) Missing feature: this deviation mainly concerns routes where a so-called short-sea carrier is used where waiting times can vary greatly depending on the time the ships arrive in a port. The ML could not learn according patterns since only the distance was available as a feature.
- (1) Incorrect training data

One can see that the detected errors are ones that humans should be able to correct.

5.2.2. Errors of Models A, B and C

However, the combined models have other problems that cause them to perform suboptimally. These problems can be described as follows:

- **Lack of complexity:** sometimes, adjustments were not possible in a sufficiently fine-grained way. This concerned e.g. port waiting times in Model A where the transport planners could register only one waiting time per port. However, some ports (e.g. Shanghai) have a fast lane for ships arriving from a local port, i.e. one would need to differentiate waiting times into national and international arrivals. Although this is possible and would increase the precision of predictions, such exceptions and their complexity could also quickly lead to an excessive workload for knowledge engineering. A possible solution to this dilemma could be to check whether capturing exceptions will affect a sufficiently large number of deliveries – assuming that an 80:20 rule applies, i.e. capturing only the 20% most relevant exceptions should cover 80% of all deliveries.
- **Lack of precision:** similarly, adjustments that transport planners were asked to make for flagged connections in Model C seemed to generate an excessively high workload – we found that several errors were made because transport planners were overwhelmed with the many flagged sender-receiver relations and tried to save time by applying generalised adjustments to e.g. all relations involving the same sender and receiver countries. There does not seem to be an obvious solution to this problem – but it confirms the result that Model C is not a useful approach.
- **Wrong extrapolation of trends:** it was observed that humans tend to “extend” long-lasting trends beyond their actual duration. In some cases, there was a clear and long-lasting temporal trend of increasing delivery times – however, although the peak was reached at some point and the curve went down again, humans still assumed that it kept increasing.
- **Wrong incentives:** in the specific case we studied, transport planners were evaluated by availability of goods – and thus had an incentive to overestimate lead times, ensuring very timely delivery. Thus, planners inserted some “buffers” in their estimates in many cases. Of course, overestimation also has negative consequences, most notably warehouse capacities may be exceeded when goods arrive too early. However, warehouse capacity issues do not have direct negative impact on the evaluation of transport planners in the studied company. An obvious solution is to avoid bias by either not using any key performance indicator (KPI) for the evaluation of transport planners, or to add a KPI measuring storage costs.

6. Conclusion

In this study, we have described three possible ways of incorporating human knowledge into a machine learning approach to estimating delivery lead times. We performed a quantitative evaluation that did not show any notable improvements of such combined ML-human models.

However, we have been able to gain insights through our qualitative analysis that could lead to better combinations: we observed that both machine learning models and humans have their typical imperfections – as hypothesised, we see that machines tend to perform poorly when situations change in such a way that historical patterns become invalid. While humans are theoretically good in compensating for that, they tend to over-simplify certain issues, mostly to keep efforts and complexity manageable. Another significant problem with human estimates was caused by wrong incentives: humans had something to gain from overestimating lead times.

Based on these insights, we believe that more successful combinations of ML and human knowledge are possible in the future. We have learned from this study that their design needs to ensure that humans are not overwhelmed with feedback provision and that they do not have incentives for misprediction. We also believe that an improved integration of explainability algorithms and a closer knowledge integration through algorithms like Bayesian Networks could be addressed in future work to handle the discovered challenges.

References

[1] M. Paliwal, U. A. Kumar, Neural networks and statistical techniques: A review of applications, Expert systems with applications 36 (2009) 2–17.

[2] J. B. Hirschberg, F. Enos, S. Benus, R. L. Cautin, M. Graciarena, E. Shriberg, Personality factors in human deception detection: Comparing human to machine performance (2006).

[3] L. von Rueden, S. Mayer, K. Beckh, B. Georgiev, S. Giesselbach, R. Heese, B. Kirsch, J. Pfrommer, A. Pick, R. Ramamurthy, M. Walczak, J. Garcke, C. Bauckhage, J. Schuecker, Informed Machine Learning – A Taxonomy and Survey of Integrating Knowledge into Learning Systems (2019) 1–20. URL: http://arxiv.org/abs/1903.12394. arXiv:1903.12394.

[4] K. Weiss, T. M. Khoshgoftaar, D. Wang, A survey of transfer learning, Journal of Big data 3 (2016) 1–40.

[5] H. Gleissner, J. C. Femerling, Logistics, Springer Texts in Business and Economics, Springer International Publishing, Cham, 2013. URL: http://link.springer.com/10.1007/978-3-319-01769-3. doi:10.1007/978-3-319-01769-3.

[6] L. Barua, B. Zou, Y. Zhou, Machine learning for international freight transportation management: A comprehensive review, Research in Transportation Business and Management 34 (2020). URL: https://doi.org/10.1016/j.rtbm.2020.100453. doi:10.1016/j.rtbm.2020.100453.

[7] S. van der Spoel, C. Amrit, J. van Hillegersberg, Predictive analytics for truck arrival time estimation: a field study at a European distribution centre, International Journal of Production Research 55 (2017) 5062–5078. URL: https://www.tandfonline.com/action/journalInformation?journalCode=tprs20https://www.tandfonline.com/doi/full/10.1080/00207543.2015.1064183. doi:10.1080/00207543.2015.1064183.

[8] X. Li, R. Bai, Freight vehicle travel time prediction using Gradient Boosting Regression Tree, Proceedings - 2016 15th IEEE International Conference on Machine Learning and Applications, ICMLA 2016 (2017) 1010–1015. doi:10.1109/ICMLA.2016.101.

[9] A. Balster, O. Hansen, H. Friedrich, A. Ludwig, An ETA Prediction Model for Intermodal

Transport Networks Based on Machine Learning, Business and Information Systems Engineering 62 (2020) 403–416. doi:10.1007/s12599-020-00653-0.

[10] R. C. Leachman, P. Jula, Estimating flow times for containerized imports from Asia to the United States through the Western rail network, Transportation Research Part E: Logistics and Transportation Review 48 (2012) 296–309. doi:10.1016/j.tre.2011.07.002.

[11] S. Kim, H. Kim, Y. Park, Early detection of vessel delays using combined historical and real-time information, Journal of the Operational Research Society 68 (2017) 182–191. URL: https://www.tandfonline.com/action/journalInformation?journalCode=tjor20https://www.tandfonline.com/doi/full/10.1057/s41274-016-0104-4. doi:10.1057/s41274-016-0104-4.

[12] I. Parolas, ETA prediction for containerships at the Port of Rotterdam using Machine Learning Techniques, in: Transportation Research Board 96th Annual Meeting, 2017.

[13] O. Loyola-Gonzalez, Black-box vs. White-Box: Understanding their advantages and weaknesses from a practical point of view, IEEE Access 7 (2019) 154096–154113. doi:10.1109/ACCESS.2019.2949286.

[14] Y. Zhang, A. Haghani, A gradient boosting method to improve travel time prediction, Transportation Research Part C: Emerging Technologies 58 (2015) 308–324. doi:10.1016/j.trc.2015.02.019.

[15] G. Leshem, Y. Ritov, Traffic flow prediction using adaboost algorithm with random forests as a weak learner, PROCEEDINGS OF WORLD ACADEMY OF SCIENCE, ENGINEERING AND TECHNOLOGY VOLUME 21 1 (2007) 193–198.

[16] A. Prokhorchuk, J. Dauwels, P. Jaillet, Estimating Travel Time Distributions by Bayesian Network Inference, IEEE Transactions on Intelligent Transportation Systems 21 (2020) 1867–1876. URL: https://www.researchgate.net/publication/331541393_Estimating_Travel_Time_Distributions_by_Bayesian_Network_Inference. doi:10.1109/TITS.2019.2899906.

[17] J. Zhao, Y. Gao, J. Tang, L. Zhu, J. Ma, Highway Travel Time Prediction Using Sparse Tensor Completion Tactics and K -Nearest Neighbor Pattern Matching Method, Journal of Advanced Transportation 2018 (2018). doi:10.1155/2018/5721058.

[18] H. Chang, D. Park, S. Lee, H. Lee, S. Baek, Dynamic multi-interval bus travel time prediction using bus transit data, Transportmetrica 6 (2010) 19–38. URL: http://www.tandfonline.com/doi/abs/10.1080/18128600902929591. doi:10.1080/18128600902929591.

[19] P. S. Deshmukh, Travel Time Prediction using Neural Networks: A Literature Review, in: 2018 International Conference on Information, Communication, Engineering and Technology, ICICET 2018, Institute of Electrical and Electronics Engineers Inc., 2018. doi:10.1109/ICICET.2018.8533762.

[20] C. H. Wu, J. M. Ho, D. T. Lee, Travel-time prediction with support vector regression, in: IEEE Transactions on Intelligent Transportation Systems, volume 5, 2004, pp. 276–281. doi:10.1109/TITS.2004.837813.

[21] S. Amershi, M. Cakmak, W. B. Knox, T. Kulesza, Power to the people: The role of humans in interactive machine learning, AI Magazine 35 (2014) 105–120. doi:10.1609/aimag.v35i4.2513.

[22] S. Khalid, T. Khalil, S. Nasreen, A survey of feature selection and feature extraction techniques in machine learning, Proceedings of 2014 Science and Information Conference,

SAI 2014 (2014) 372–378. doi:10.1109/SAI.2014.6918213.
[23] Microsoft, Algorithm component reference for Azure Machine Learning designer, 2021. https://docs.microsoft.com/en-us/azure/machine-learning/component-reference/component-reference, visited 2022-01-23.
[24] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st international conference on neural information processing systems, 2017, pp. 4768–4777.
[25] M. T. Ribeiro, S. Singh, C. Guestrin, " why should i trust you?" explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1135–1144.
[26] Bloomberg, Kompasu Typhoon Warning Closes China's Shenzhen Port, Hitting Supply Chains - Bloomberg, 2021. https://www.bloomberg.com/news/articles/2021-10-12/worst-ship-traffic-jam-since-august-outside-top-chinese-port, visited 2021-11-02.