Predicting Estimated Arrival Times in Logistics Using Machine Learning

P

Peter Poschmann
Technische Universität Berlin, Germany

Manuel Weinke
Technische Universität Berlin, Germany

Frank Straube
Technische Universität Berlin, Germany

INTRODUCTION

ML techniques offer great potentials to support decision-making processes in companies by allowing the prediction of unknown information. The ability to extract and approximate system relationships from training data without explicit a priori knowledge makes them highly suitable for modeling highly complex and dynamic real-world systems. Due to the learning capability of ML-based systems, problems can be solved more flexibly, with less effort and with higher accuracy, and there is the potential to automate decisions (Wahlster, 2017). Against this background, ML-based applications are of particular importance, especially in the field of logistics. However, while the majority of logistics companies already use technologies for real-time visibility like Track & Trace, ML-based decision support systems have so far rarely been applied in logistics (Straube, 2019). The goal of the chapter is to demonstrate the application of ML methods on a significant use case in logistics practice: the prediction of ETA in intermodal transport networks as a basis for the detection of process disruptions.

The maritime transport chain serves as the practical use case considered in the chapter. International container transports by ship are handled via complex transport networks involving a large number of logistics actors. The implementation requires the interaction of numerous, closely timed and interdependent sub-processes. At the same time, the execution of the processes is influenced by a variety of impact factors such as resource availability, weather and the human factor. In addition, there is often no complete transparency between the actors, as information on process planning, status and disruptions has so far only been exchanged insufficiently and often manually between the involved logistics companies. Many of the decisions are therefore made under high uncertainty and rather reactively. As a result, according to Poschmann et al. (2019) decisions are often not optimal in regard to the entire chain and lead to high economic and ecological disadvantages in the form of unpunctual deliveries, resources not optimally utilized, cost-intensive special processes and unnecessary risk buffers.

Against this background, early prediction of arrival times and possible delays is highly important. Thus, ETA information enables logistics actors to identify and deal with possible process disruptions at an early stage by initiating appropriate measures. Furthermore, information on arrival times is an important basis to ensure a demand-oriented capacity planning with regard to material stocks, personnel and infrastructure (Walter, 2015). ML-based ETA predictions can thus make an important contribution to

DOI: 10.4018/978-1-7998-9220-5.ch160

improving today's logistics networks, which are affected by increasing customer requirements in terms of reliability, transparency and sustainability and cost efficiency (Handfield et al., 2013).

The chapter aims to demonstrate the application of ML for ETA prediction in logistics. A detailed insight is given into the results of a research project whose objective was to develop an ETA prediction for combined road-rail traffic in the port hinterland. The chapter is organized as follows: In the first step, a general methodology is presented, which contains all essential subphases of the development, starting from the requirements analysis and data collection up to the IT integration. Subsequently, the conception of an approach for the above-mentioned use case is presented and ML approaches for three selected sub-processes are prototypically implemented and evaluated.

BACKGROUND

The following is a brief description of the fundamentals relevant to this chapter. First, a definition of ML is given, followed by an overview of the state of the art in ETA prediction research.

Machine Learning

ML is a sub-domain of Artificial Intelligence (AI) and comprises various methods that enable computer systems to independently extract patterns from extensive data (Murphy, 2012). ML thus enables computer systems to learn inductively. (Nilsson, 2010) This means that inference takes place on the basis of hypothetical correlations that a learning algorithm has acquired in the course of a training process by adapting to observations and generalizing patterns contained therein. (Awad & Khanna, 2015) This automatic extraction of patterns from data enables the recognition of complex relationships that are not recognizable to humans, or only with great effort, and an industrial use, for example, for segmenting and predicting information, deriving rules and solving optimization problems. (Alpaydın, 2010; Döbel, et al., 2018). Approaches to ML can be roughly divided into the three main types supervised, unsupervised and reinforcement learning. (Russell & Norvig, 2010). For ETA prediction, supervised and unsupervised learning are of particular importance, as learning is mainly based on historical transport data.

Prediction of Estimated Times of Arrival

With regard to the existing approaches to ETA prediction in the literature, a basic distinction can be made between model-based approaches, which are based on simulations or analytical models, and data-based approaches, according to Wen et al. (2017). The application of ML for ETA prediction can be regarded as a sub-group of the data-based approaches. Model-based approaches have been widely used in the literature for delay prediction in rail networks. They can be found in Berger et al. (2011) as well as Büker and Seybold (2012), for example. However, their disadvantage lies in the complex modeling and the low adaptability to changing operational conditions. Despite the great potential of ML for ETA prediction, its application has been explored only selectively and with a strong focus on passenger transport. Existing approaches to logistics are usually only related to isolated sub-processes and specific modes of transport.

Initial approaches already exist for maritime transport, which represents the main leg in international maritime transport chains. Parolas et al. (2016) predict the arrival time of ocean vessels at the port of Rotterdam started about 120 hours before arrival using Artificial Neural Networks (ANN) and Support Vector Machines (SVM). Bodunov et al. (2018) and Lechtenberg et al. (2019) developed models for the prediction of ship arrival times at a specific port or in a specific destination region, as well as for the prediction of the destination port and port turnaround time. ANN, SVM as well as ensemble methods such as Extreme Gradient Boosting (XGB) were tested.

Existing approaches to rail transport are almost related to passenger transport. Appropriate concepts are developed, for example, in Oneto et al. (2018), Huang et al. (2020) and Marković et al. (2015), predominantly using ANNs. However, those approaches are only applicable to rail freight transport to a limited extent as passenger transport differs from freight transport by rail in respect of the operational processes and influencing factors. An approach for freight trains, but solely with statistical methods (linear regression), can only be found in Gorman (2009).

For road transport, existing approaches also focus primarily on passenger transport. For example, Fan and Gurmu (2015) use ANN to predict travel times in public bus transport. Due to the different conditions (e.g. stronger timetable dependency, shorter transport distances), these are not suitable for road freight transport. Only Li and Bai (2016) deal with the development of an approach for road freight transport using XGB, where only temporal characteristics are considered.

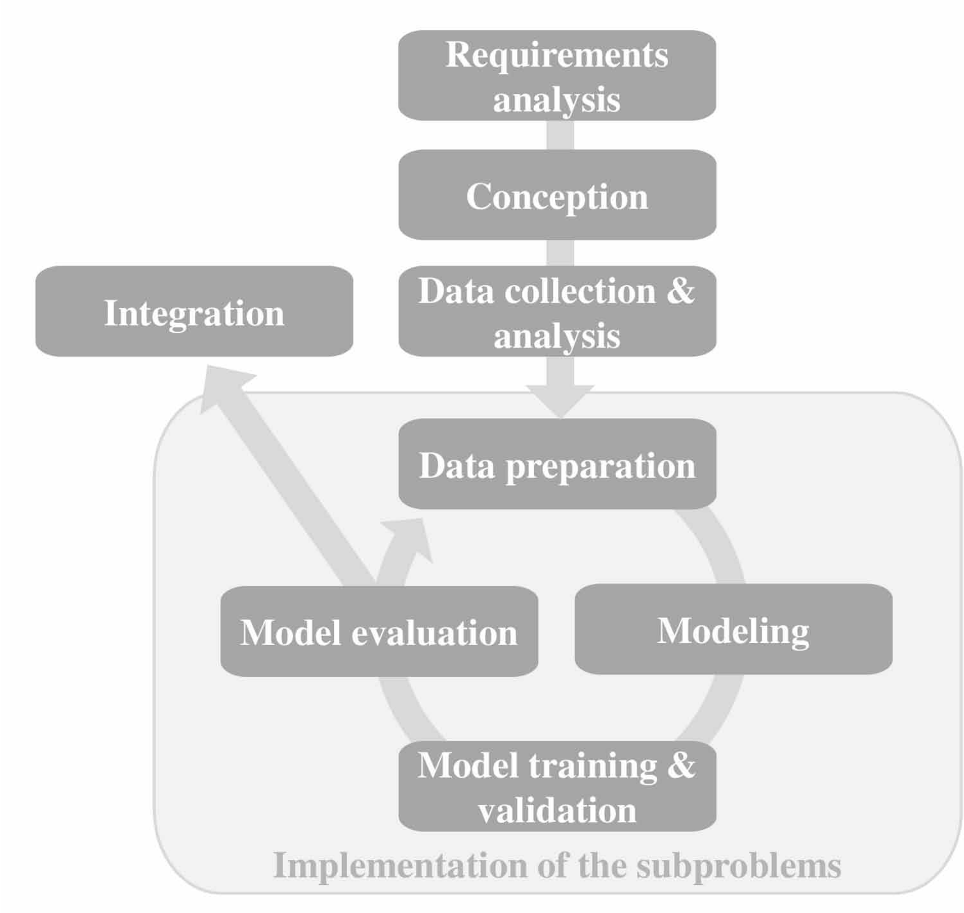
Except for the port turnaround time prediction by Lechtenberg et al. (2019), no ML-based approaches for the prediction of processes in logistical nodes such as inland terminals and marshalling yards could be found. Similarly rare are approaches for the prediction of more complex transport chains, which are, however, of fundamental importance for the implementation of ETA predictions for logistics. The only application of ML for the prediction of intermodal transport chains was found in Servos et al. (2020), who developed an ETA prediction for container transports within the maritime transport chain. This was done using ensemble methods such as Adaptive Boosting and SVM. The approach only incorporates GPS data, but no other data, so operational factors that influence arrival time are not taken into account.

In summary, there are currently only a few approaches to ETA prediction for both individual logistics processes and complex transport chains. Furthermore, there is especially a lack of a well-proven methodology on how to proceed methodically in the development of corresponding approaches. To close this research gap, in the following section, a general methodology is presented on how to develop ETA predictions for such application cases by using the example of combined road-rail transport based on a practical use case.

METHODOLOGY

The methodology for the development of the ETA prediction was based on the industry-wide standard CRISP-DM for conducting data-based projects (Wirth & Hipp, 2000). CRISP-DM is known to be a generic process model without reference to a specific use case. However, the development of a cross-actor ETA prediction requires certain sub-phases which do not emerge sufficiently concretely from this model. For example, different stakeholders need to be involved in the development process and multiple sub-models need to be developed to predict complex transportation chains. To take these aspects into account, the CRISP-DM was extended and refined for the problem of ETA prediction. The derived methodology is briefly presented below (see Figure 1). Subsequently, the results of these methodical steps are presented in particular sections. In the case of model development, some steps are summarized for the individual sub-problems. The activities of the integration will not be discussed according to the focus of the paper. The presented procedure can serve as a reference for similar problems in this application area.

Figure 1. Methodology for development of ETA prediction



Requirements analysis: To create a practice-oriented solution, the first step was to collect and assess crucial requirements on the solution, similar to the 'Business Understanding' in CRISP-DM. The phase included a process analysis, disruption analysis, and the identification of use cases and requirements for the ETA prediction in practice. As part of the process analysis, the logistics actors to be involved and the relevant logistics processes of the transport chain were first investigated. The goal of the subsequent disruption analysis was to identify all disruption reasons and other influencing factors that affect the process duration and must be considered as features in the models. Due to the high complexity of today's transport networks, it is probably not possible to incorporate all process configurations. Another step was therefore to prioritize use cases for the ETA and to define required prediction horizons, prediction qualities reference objects and events to which an ETA must relate. The requirements analysis was carried out in the form of expert interviews, comprising also a Failure Mode and Effect Analysis (FMEA).

Conception: Due to the complexity of transportation chains, a single model is usually not sufficient for ETA prediction. Rather, the problem must be decomposed into several problems according to the sub-processes and the requirements. Each sub-process is then represented by one or more specific regression or classification models with individual features. Rule-based systems (or simulation methods) can be used later to link the sub-models according to predefined rules and to incorporate prior knowledge. The definition of these sub-problems and the overall logic of the system were done in the conception phase.

Data collection & analysis: An essential precondition for the desired ML approach was the availability of historical data. To ensure a sufficient data basis, various measures were taken to obtain suitable data sources (similar to the phase 'Data Understanding' in CRISP-DM). The data obtained was processed and analyzed. This included both the analysis of process characteristics and the identification of possible factors influencing process times. Data visualization and methods from the field of unsupervised learning like cluster analysis and sequence mining were used for this purpose.

Data preparation: Similar to CRISP-DM, the next step was to properly prepare the data for each sub-problem. Feature engineering is a central task in this process. Here, suitable input variables (features) were designed that represent the identified influencing factors and disruptions as well as possible. This involves using existing variables as well as creating new variables, for example by combining and transforming the existing variables. This phase also included standard ML tasks such as outlier handling, missing value handling and feature encoding.

Modeling: For each sub-problem, a suitable forecasting approach is designed. Supervised learning methods for regression or classification were used, depending on the sub-problem. Aspects such as the type of labels and features, data quantity and data quality were taken into account to choose a proper approach.

Model training & validation: The selected ML methods were finally trained with historical training data for each sub-problem, using Grid Search and Cross-Validation (CV) for hyperparameter tuning. The statistical programming language R and further open source packages were used for implementation and evaluation.

Model evaluation: To determine the achievable prediction quality of the respective trained models, they were applied to independent test data. Various quality measures such as Mean Absolute Error (MAE) were used for regression problems and accuracy, Cohens kappa and no-information rate for classification problems. In addition, other problem-specific metrics were calculated for better external understanding.

Integration: Once sufficient prediction quality had been achieved for each sub-problem, the model was integrated into an overall system (similar to the 'Deployment' phase in CRISP-DM). This enables the models to be linked in the sense of a "door-to-port" prediction. In addition, a rule-based decision support system was developed based on the ETA to detect disruptions in the chain and recommend appropriate measures. The developed prototype is available online at the link www.smecs-eta.de.

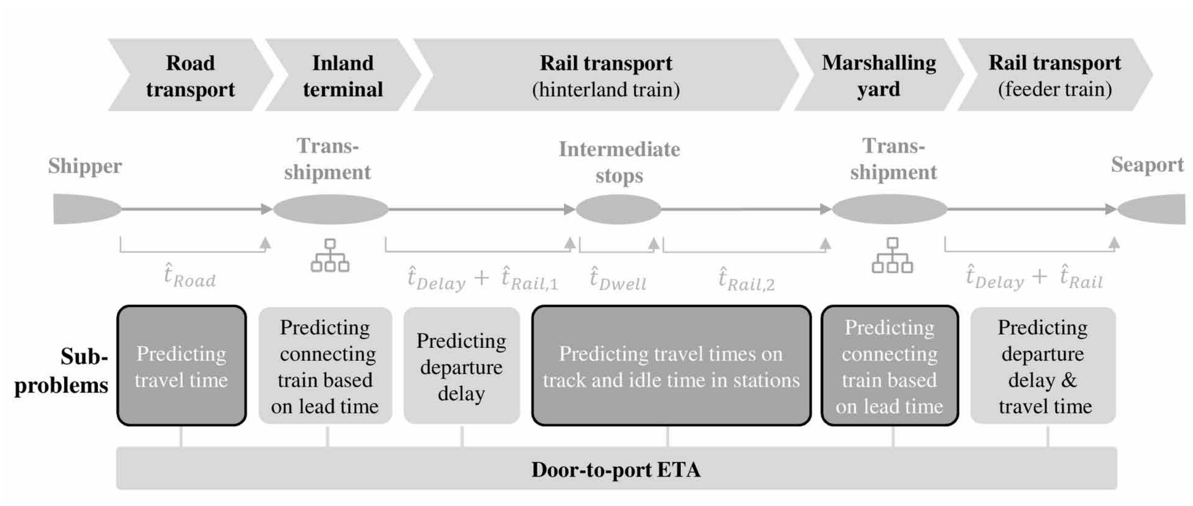
REQUIREMENTS ANALYSIS & CONCEPTION

Transport of sea freight containers in combined road-rail transport in the port hinterland (export) was chosen as the use case under investigation. The transport chain consists of five major sub-processes, which are shown in Figure 2. The first sub-process consists of road transport of the container between a shipper and the hinterland terminal. At the hinterland terminal, the container is transferred to rail and then carried by a hinterland train to a marshalling yard. There, the train is broken up and the wagons are distributed to newly formed feeder trains, which finally are moving to the seaport.

Several potential users and use cases could be identified for an ETA prognosis for the described process chain. This resulted in different reference points of the ETA (e.g. arrival of container at hinterland terminal, arrival of train at marshalling yard), reference objects (e.g. container, train, truck) and prediction times. In the disruption analysis, a large number of influencing factors were identified for each of the aforementioned sub-processes, which form the basis for feature engineering. The detailed results of the requirements analysis can be found in Poschmann et al. (2019).

On the basis of the process analysis and real-world requirements, six sub-problems were finally defined, which, together, enable a door-to-port prognosis (see Figure 2). The overall prediction is achieved by a logical connection of all individual sub-models in the sense of a process chain prediction. The output of a model serves as one input information for the subsequent prediction models. The development of the highlighted sub-problems will be discussed in detail in this chapter. In addition, representative transport relations (so-called pilot relations) were defined for the development of prototypes. The pilot relations represent container shipments from a variety of shippers (origin) via three hinterland terminals (transshipment road to rail) and a marshalling yard (transfer from hinterland to feeder train) to a seaport (destination).

Figure 2. Process chain and defined sub-problems for Door-to-port ETA prediction



DATA COLLECTION & ANALYSIS

To carry out the development, four years of historical data were obtained from various logistics and transport companies such as a rail transport company, a rail infrastructure operator and a CT operator from the project consortium. The data was obtained from a total of 16 IT systems, in particular booking, scheduling and enterprise resource planning (ERP) systems. In addition, further external data sources were obtained, comprising information about external influencing factors like weather conditions and vacations. Not all data sources were available for the entire time period. Figure 3 provides an overview of the available data for each of the three sub-problems. The data can be roughly divided into two types of information: Information on the physical processes (time stamps and geopositions) and information on the conditions and influencing factors, serving as the basis for feature engineering.

For road transport, a total of about 47,000 truck trips (after data cleansing) were available. The process data includes information about the actual time of departure and arrival as well as information about the locations (manually entered address information). For the construction of possible features, additional data on the load (dangerous goods status, container weight and length), weather conditions (air temperature, precipitation, wind speed and direction, snow) as well as vacation periods and public holidays were available. No information was available on planned times, sub-processes and the route.

The highest data coverage was provided for rail transports. For this sub-process, data on about 2,100 hinterland train journeys (after data cleansing) were available. Unlike for road transport, the process data included detailed route information in the form of waypoints as well as additional planned times from the timetable. In addition, information on planned sub-processes (e.g. locomotive crew changes) was available for each train journey. As a basis for feature construction, data sources on train and locomotive characteristics (e. g. train weight and type of traction unit), construction works (location and period of road works on route and consequences), weather conditions as well as vacation periods and public holidays were available.

For the marshalling yard, wagon-based transport plans for about 16,000 wagon movements (after data cleansing) were available. The dataset used contained information about the planned and actual inbound and outbound trains of each wagon as well as the corresponding arrival and departure times. In addition, information about the timetables of all outbound feeder trains, construction works, weather conditions and vacation periods and public holidays could be considered. Data about the sub-processes within the marshalling yard was not available, so this had to be treated as a black box.

Figure 3. Overview of the used information in the raw data set

	Road transport	Rail transport	Marshalling yard
Process data	• Origin and destination • Actual time of departure and arrival	• Planned route over several waypoints (incl. geopositions) • Actual and planned times of departure and arrival and further timestamps for each waypoint • Planned sub-processes	• Planned and actual times of arrival and departure for inbound and outbound trains • Planned and actual inbound / outbound trains of each wagon • Schedules of alternative trains
Data on the conditions	• Container weight and length • Dangerous goods status • Weather conditions • Vacations and public holidays	• Construction works on track • Train characteristics • Locomotive characteristics • Weather conditions • Vacations and public holidays	• Construction works inside the marshalling yard • Weather conditions • Vacations and public holidays

IMPLEMENTATION OF THE SUB-PROBLEMS

In the following section, the stages from data preparation to model evaluation are described in detail for each of the three selected sub-problems Road Transportation, Rail Transportation and Marshalling Yard.

ROAD TRANSPORT

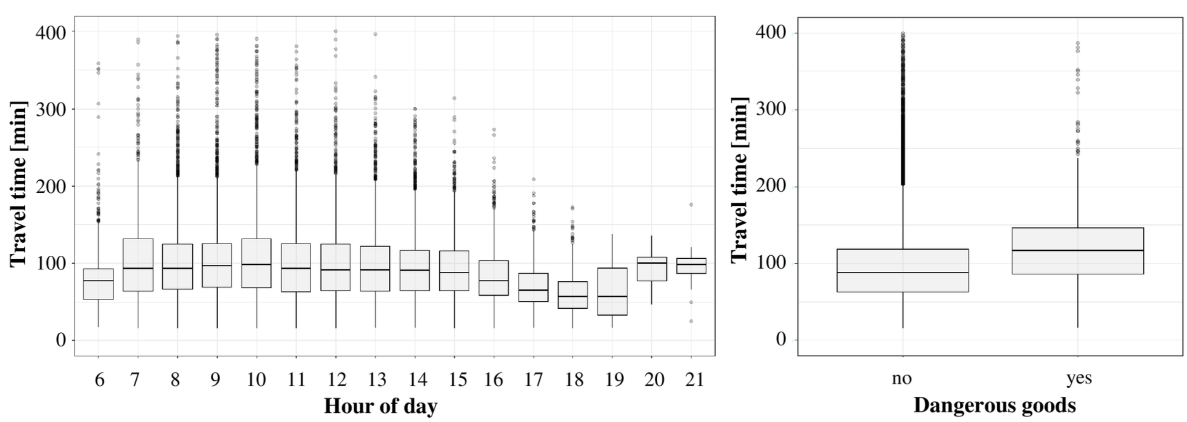
The first sub-problem relates to the road transport of containers between shippers and the three hinterland terminals on the pilot relations. The considered journeys mainly take place in the near region of the hinterland terminals and therefore have short travel times of only a few hours. The aim of the ML models to be developed is to predict the expected arrival time of a container at the hinterland terminal. In addition to the driving time, the process duration can also include times for rest periods and intermediate parking of the container, which are not known beforehand. Among other things, the process duration is strongly influenced by the traffic situation (congestion) and weather conditions.

During the data cleansing process, journeys were removed that had no plausible or missing time and location information (e.g. journeys with journey times of less than 15 minutes or more than 10 hours). Some of the journeys related to the collection of the empty container were also removed. After data cleaning, 47,278 valid trips remained for the model development.

In the next step, the actual journey times were determined for all journeys from the departure and arrival times. In addition, source-destination relations were defined from the trips and each trip was assigned to a relation. The formation of the relations was based on the postal code area and the respective inland terminal, i.e. all trips from the same postal code area to the same inland terminal belong to the same relation. This resulted in a total of 1,020 possible source-destination relations.

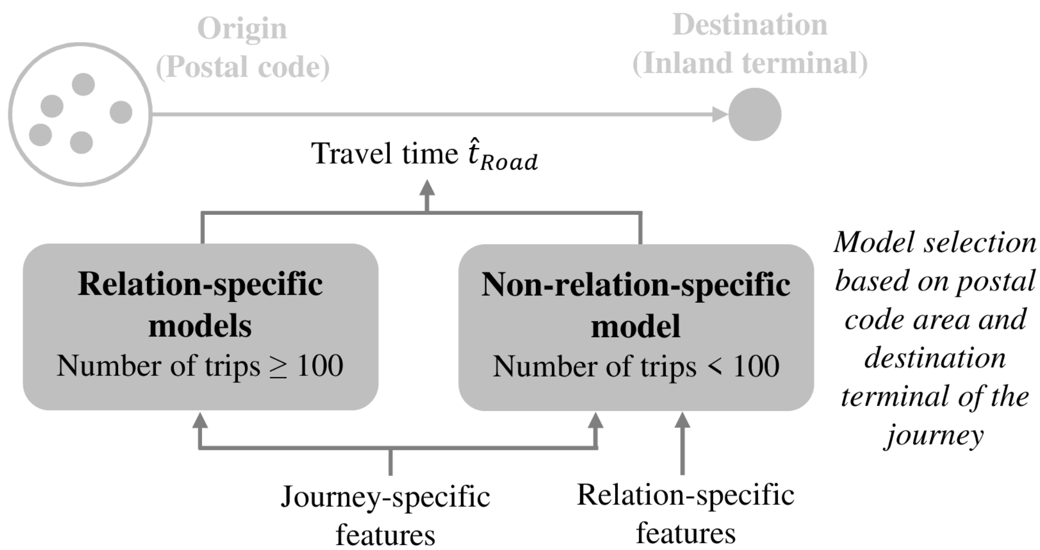
In order to determine the most influencing factors, correlations between travel times and various variables were investigated. Figure 4 shows the correlations between the hour of departure and the dangerous goods status and the travel time. Significant correlations can be seen for time of day, which can be caused by traffic congestion or breaks in the journey. There is also a slight correlation between travel time and the dangerous goods status of the container.

Figure 4. Correlations between travel time and hour of day / dangerous goods status



In the next step, an approach for predicting the arrival times of trucks and containers at the hinterland terminals was developed using supervised learning methods. The task is treated as a regression problem with the travel time as target variable. The arrival time (ETA) is determined from the planned or actual departure time and the predicted travel time. Since many trips occur on the same source-destination relations, two types of features can be distinguished. Journey-specific features represent conditions that only refer to a specific journey (e.g. the container weight). Relation-specific features, on the other hand, contain information that is valid equally for all trips of a relation (e.g., the aerial distance). In order to enable the best possible learning of both journey- and relation-specific features, a hybrid approach with several sub-models was chosen (Figure 5). For frequently travelled relations with more than or equal 100 available historical journeys, a relation-specific model was trained in each case. For relations with less than 100 past journeys, a non-relation-specific model was trained on the whole training set, as the amount of training data for relation-specific models is too small. Only relation-specific characteristics are included in the relation-specific sub-models, whereas relation-related characteristics are also taken into account in the non-relation-specific model.

Figure 5. Technical approach for prediction of road travel times



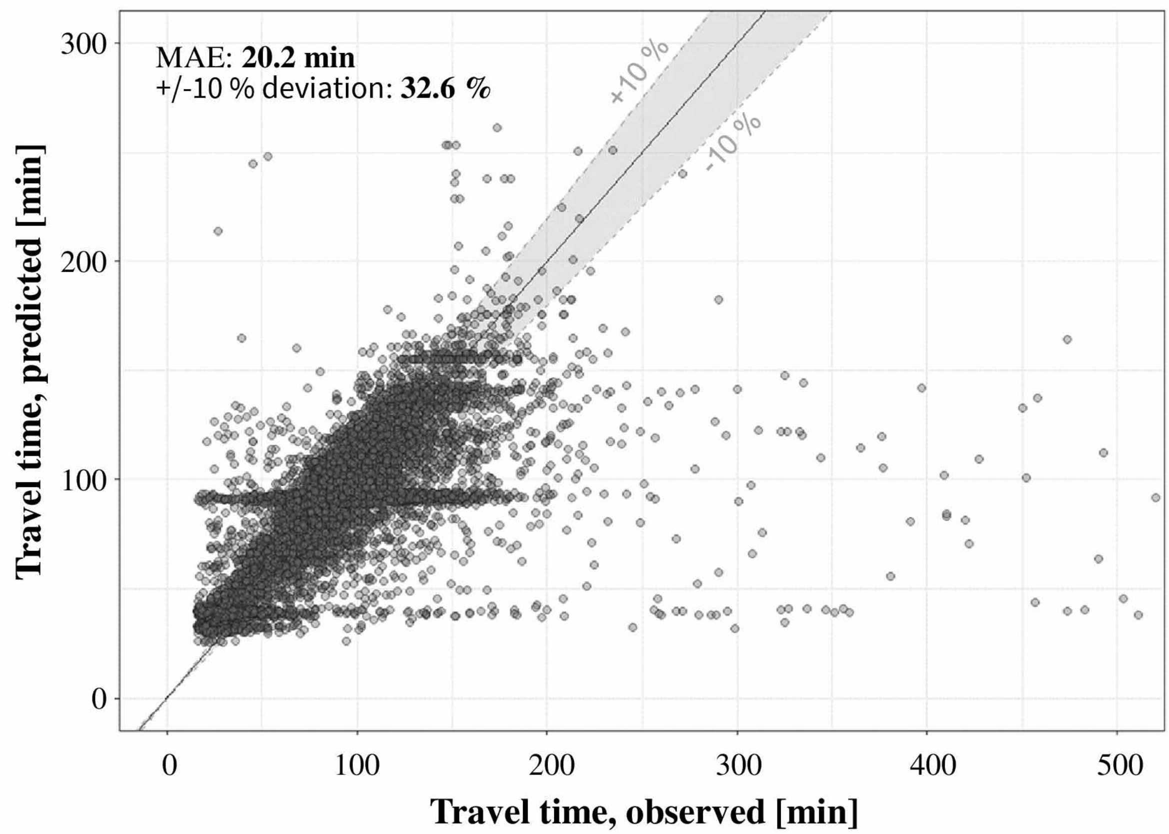
From the available data, eleven journey- and relation-specific features were selected. Trip-specific features include time-related features (hour of departure, weekday, month, holiday density), shipment-related features (weight and length of the container, dangerous goods status) and weather-related features (temperature, precipitation). As relation-specific features, based on the training data, the median travel time on each relation as well as the linear air distance between the starting point (center of the postal code area) and the destination (geo-coordinates of the inland terminal) were calculated. Categorical characteristics (e.g. weekday) were transformed into binary variables by one-hot encoding.

Different ML methods were tested. For the implementation of the relation-specific models, the highest prediction quality was achieved with XGB, whereas a Linear Regression Tree described in Zeileis et al. (2008) was identified as the optimal method for the non-relation-specific model, as it allows for an optimal consideration of the continuous, relation-specific features.

For the purpose of evaluation, the data was split in a ratio of 80% / 20% into a training / validation dataset and a test dataset. For the model selection, hyperparameter tuning was performed using grid search. Finally, a final evaluation of the trained models was carried out using the test set. The MAE is 20.2 minutes. 32.6% of the predicted journeys are within a maximum deviation interval of +/- 10% of the travel time observed. The evaluation results are presented in Figure 6.

It can be seen that the majority of journeys can be predicted with a relatively high accuracy. Nevertheless, there are also cases with larger prediction errors, which can be caused, for example, by additional travel interruptions (e.g. rest periods), waiting times at hinterland terminals and various previously unforeseeable disruptions (e.g. construction sites, traffic jams, technical failures). Also, documentation errors due to the manual data entry of departure and arrival times by the driver can be the cause of existing variations. By improving the data quality and considering further data sources, a further increase in the quality of the prognosis can be expected. Possible additional data sources include, in particular, more detailed information on the planned routes, processes (rest periods, intermediate stops), the traffic situation on the route, planned construction sites and waiting times at the inland terminal.

Figure 6. Evaluation results for prediction of road travel times



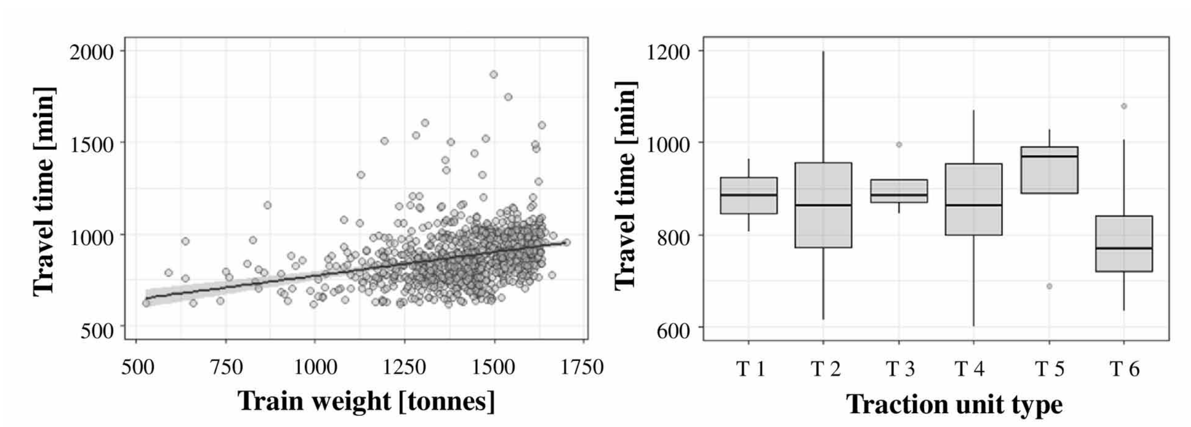
RAIL TRANSPORT

As a further sub-problem, the prediction of container transport by rail should be discussed. The focus is on hinterland trains between hinterland terminals and marshalling yards. Train runs on three source-destination relations are considered (starting from three different hinterland terminals to a marshalling yard). For each relation, however, there are a variety of possible routes through the rail network that connect the respective hinterland terminal with the marshalling yard. Compared to road transport, rail transport is subject to a timetable that clearly defines the spatial and temporal course of a journey via discrete waypoints in the rail network. In addition to traction, rail transport includes intermediate stops, where, among other things, locomotive crew changes and locomotive changes are carried out. Rail transport is characterized by numerous different disruptive and influencing factors that can lead to timetable deviations and must therefore be integrated into the model as input variables. These include, for example, train path conflicts, delays in personnel changes, weather influences, construction sites and technical disruptions. By developing suitable features, it has been attempted to take these influencing factors into account as far as possible.

In the initial processing, non-representative journeys (outliers) were removed. This also included journeys that predominantly ran on rarely used route sections. Due to the relatively small number of train journeys, missing values were replaced as far as possible instead of removing the cases. A total of 2,107 journeys remained in the data set (Relation A: 870 journeys, Relation B: 470 journeys, Relation C: 767 journeys). For relation C, some data sources were not available so that certain features could not be generated.

Subsequently, a comprehensive data analysis was carried out to identify factors influencing the journey time. Figure 7 shows an example of the correlations between influencing factors of different cause groups and the journey times of the trains on the basis of scatterplots and boxplots. The visualization shows that the train weight has a positive influence on the travel time. Furthermore, correlations between the traction unit type and travel time can be observed. The factors shown are therefore to be included in the model as features along with others.

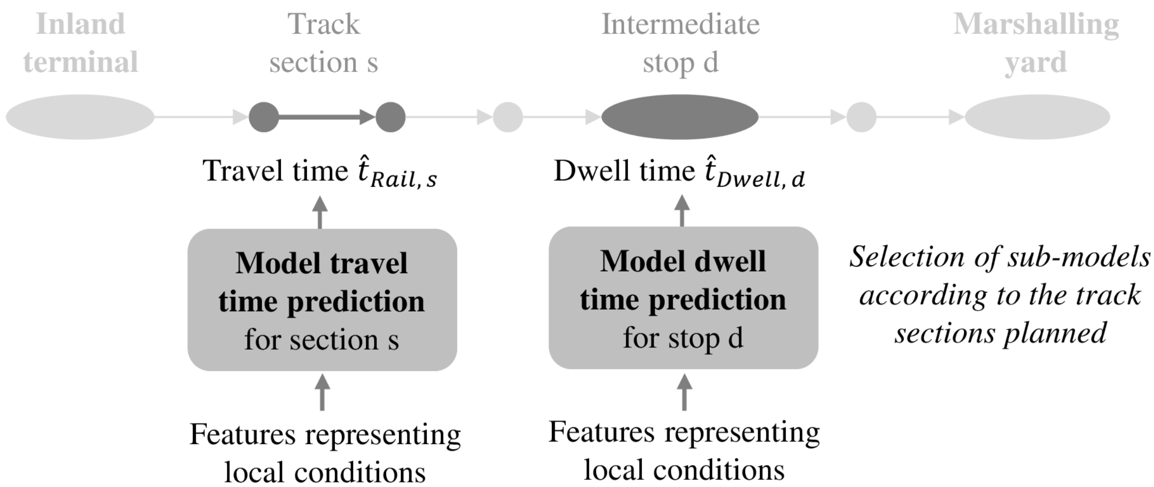
Figure 7. Correlations between travel time and train weight / traction unit type



As described at the beginning, each train journey is made up of a sequence of partial sections and intermediate stops, each with fixed planned times. In order to consider the different route configurations in the prediction, an approach was chosen in which several sub-models are trained (see Figure 8). Each sub-model represents a specific route section or a station. In each sub-model, local influencing factors were considered (e.g. local construction works and weather conditions). The target variable (labels) of each sub-model is the travel time on the specific route section or the dwell time at a station. The prediction of the ETA at the destination results from the actual or predicted departure time and the sum of the sub-process times over all planned route sections or stations. A model was estimated for a section or station if at least 30 historical journeys were available for training, otherwise the planned times from the timetable were used for this section, as the amount of data was not sufficient for training a model.

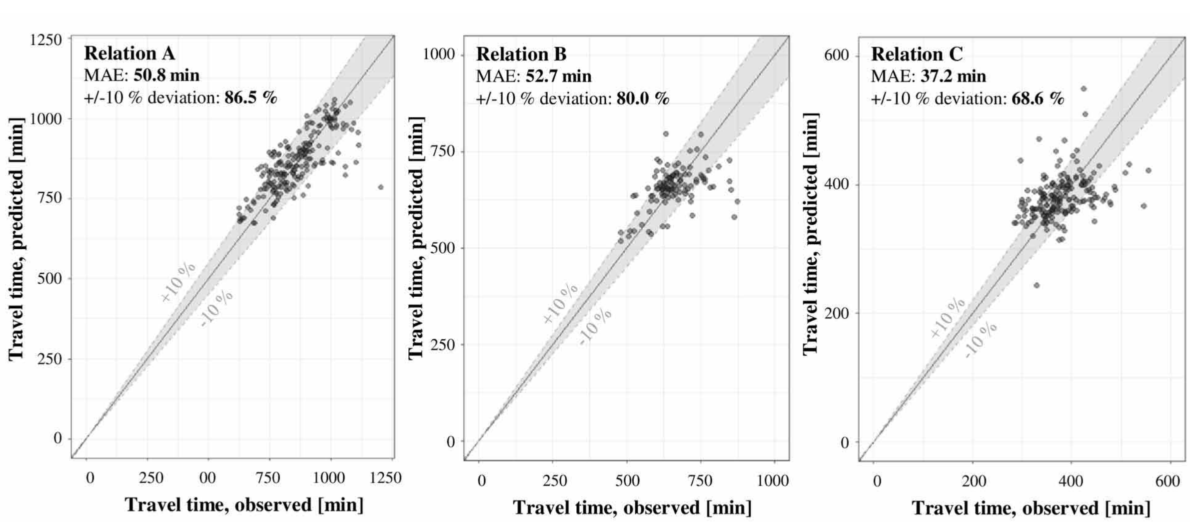
To implement the approach, the data set was transformed and aggregated accordingly, so that each train journey is mapped as a sequence of the planned sub-sections. Subsequently, a total of 21 features were generated and selected. These include information from the timetable, train and locomotive characteristics, staff availability, construction sites, weather conditions and temporal factors. Due to the large number of sub-models, the final features considered in each sub-model were selected as part of the automatic feature selection of the training procedure. Different ML methods were tested for the sub-models described. The highest prediction quality was achieved with the Random Forest method for travel time prediction and XGB for dwell time prediction.

Figure 8. Technical approach for prediction of rail process times



For evaluation the data was divided into a training / validation data set and a test data set in a ratio of 75% / 25%. Grid Search and CV were used to tune the hyperparameters. The evaluation was done separately for each of the three relations. The results of the evaluation are shown in Figure 9.

Figure 9. Evaluation results for prediction of rail travel times



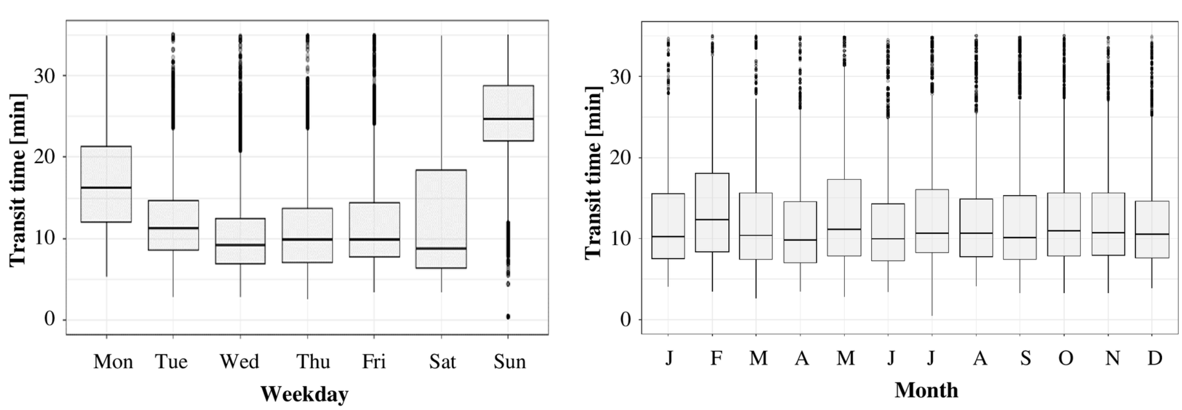
The highest prediction quality was achieved for Relation A. The MAE on this route was 50.8 minutes with a mean total travel time of approx. 900 minutes. 86% of the predictions were within a deviation interval of max. +/-10% compared to the actual travel time. The lowest prediction quality (+/-10%) was achieved on route C with 68%, which can be explained in particular by the poorer data basis. On route B, a high prediction quality was also achieved with a quality of 80%. Overall, it can be concluded that the selected approach allows to predict rail process time with a high precision. However, some influencing factors, in particular technical disruptions, traffic-related influences could not yet be sufficiently considered and require the integration of further data sources.

MARSHALLING YARD

To illustrate the prediction of logistical nodes, the marshalling yard is considered as a third sub-problem. The hinterland trains are broken up in this and the wagons are allocated to multiple feeder trains which serve the seaport terminals. The trains between the marshalling yard and the seaport move several times a day according to a timetable. The process consists of several sub-processes such as train splitting, marshalling and train formation. The allocation of wagons to trains is carried out according to the "first-in-first-out" principle. The transit time depends on various influencing factors to be considered as features in the model, e.g. the timetable of outgoing trains, the degree of utilization of infrastructure and trains, road works, but also planned buffer times.

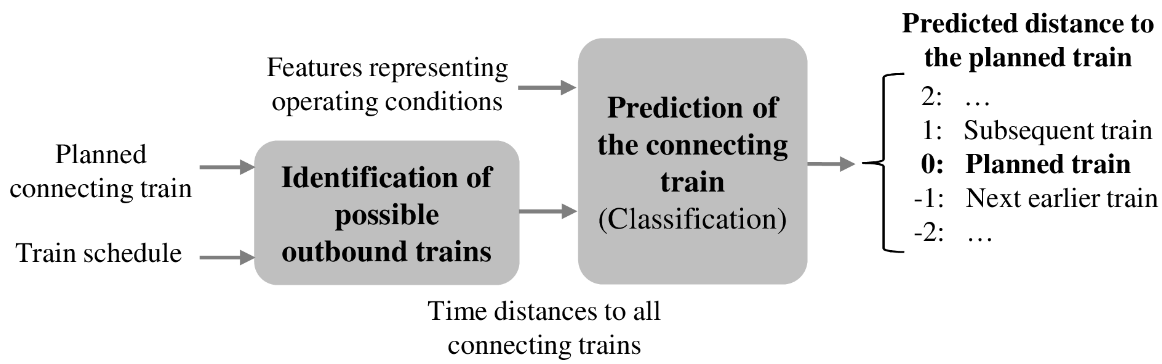
In the data preparation phase, all complete cases were first determined on the pilot relations. Each case represents the passage of a wagon through the marshalling yard, starting with the arrival of the hinterland train and ending with the departure of the feeder train. The original data set contained many identical cases in which wagons have the same inbound and outbound train. As these duplicates do not provide any gain of information, the data set was first cleansed so that only one case with an identical inbound and outbound train is available. The cleaned data included 15,765 cases. In the following data analysis, factors influencing the turnaround time were investigated. Figure 10 shows an example of the correlations between the weekday and month and the transit time of a wagon. The weekday has a high influence on the throughput time, whereas only minor differences can be observed for the month.

Figure 10. Correlations between transit time and weekday / month in marshalling yard



Since the objective is to predict the connecting train of a wagon, the process was treated as a classification problem. The approach used is illustrated in Figure 11. The basis was the connecting train planned for a particular wagon according to the transport plan. With the help of timetable data, further alternative connecting trains to the planned destination were subsequently determined before and after the planned train. These were then mapped to a discrete variable with a total of nine possible states (<-3, -3, -2, -1, 0, 1, 2, 3, >3). Category "0" corresponds to the planned initial train, while "-1", for example, corresponds to the next earlier train moving to the same destination. Categories "> 3" and "< 3" correspond to a connecting train departing at least four trains before or after the scheduled train. The restriction was introduced because a classification implies a restriction on the number of classes. In order to be able to apply supervised learning, the actual connecting trains were also determined for all cases and assigned to the categories.

Figure 11. Technical approach for prediction of connecting trains in marshalling yard

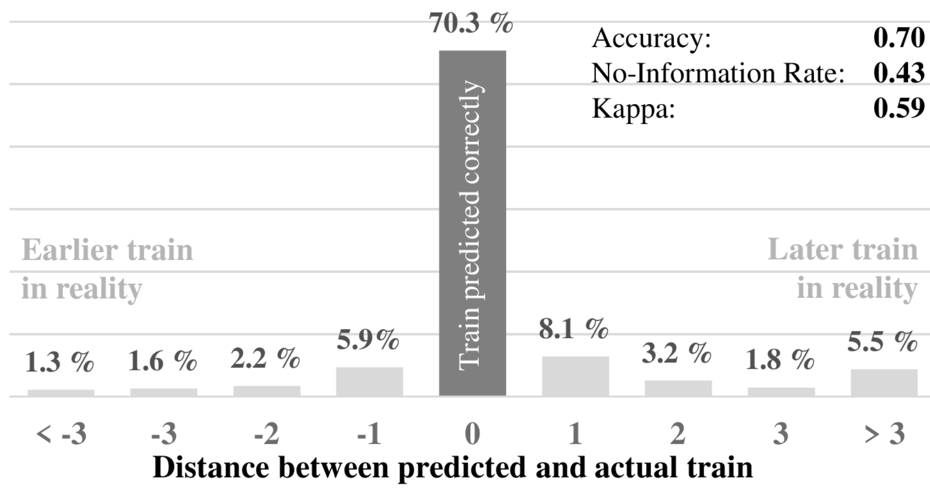


As part of the feature engineering 18 features were designed. These include the available lead times to all possible connecting trains of a wagon as well as temporal features and features on train characteristics, capacity utilization and construction work. Due to the ordinal character of the target variable (labels), the Ordinal Forest model based on Hornung (2020) was used as ML method.

Finally, the developed model was evaluated with a ratio of 90% / 10% for training / validation set and test set. For the test set an accuracy of 0.70 and a kappa coefficient of 0.59 could be achieved. The no-information rate was 0.43. Therefore, the developed model has a clear added value compared to a naive estimator whose output is always the main class of training data (in this case the planned connecting train). In order to better assess the dispersion of the prediction, Figure 12 shows the observable prediction error as the distance between the predicted and actual connecting train. It can be seen that in 84.3% of the cases the model forecasts the connecting train exactly or only has an error of +/- one train. Only in 16% of the cases higher deviations from the actual train were present.

Various factors can be the cause of the existing prediction error. On the one hand, connection scheduling in the marshalling yard is a complex decision-making situation that is carried out manually and often at short-term depending on current operating conditions. An improvement of the prediction can therefore be achieved by integrating further information into the model, e.g. on planned sub-processes within the yard as well as on the infrastructure and train utilization. Furthermore, it seems useful to take into account the closing times of a container's ocean-going vessel, which may also influence transit time in the marshalling yard. Another option for improvement may be the choice of alternative class definitions, which, for example, are based on the time of arrival of the wagon and not on the planned outbound train.

Figure 12. Evaluation results for prediction of connecting trains in marshalling yard



FUTURE RESEARCH DIRECTIONS

In the previous sections, approaches were presented on how ML can be used to predict complex transport chains in logistics. The development and evaluation of these approaches were only carried out under laboratory conditions with historical data and for selected sub-processes, so that no final statements can be made about the feasibility in practice. In addition, only a very limited number of ML methods, data sources, and modeling opportunities could be tested. This results in various directions for future research.

On the first hand, it is important to develop approaches for other logistics processes. These include, for example, other modes of transport such as inland waterways and air transport, but also further approaches for logistical nodes such as the seaport. In addition, testing with more comprehensive training data sets and further characteristics is useful. For example, numerous influencing factors have not yet been sufficiently considered within the tested models. These include, in particular, serious and rare events (e.g., heavy weather, accidents), whose prediction poses a particular challenge due to the limited availability of data. To take into account as many effects on arrival time as possible, a large number of characteristics are often needed. Against this background, the testing of novel feature encoding techniques can also be of great importance, e.g. to reduce the number of features.

In addition to alternative forms of modeling, it appears promising to test other ML methods. For example, deep learning approaches and recurrent neural networks offer high potential for ETA prediction and should be tested. A high potential also arises in the combination with other ML and AI fields. Especially unsupervised learning clustering methods could be used for data aggregation and feature extraction. Symbolic AI approaches (e.g. rule-based systems and fuzzy logic) can be useful to incorporate expert knowledge into the forecast or to implement recommendations for actions based on the ETA in the sense of prescriptive analytics. Another direction for future research concerns approaches for linking the individual sub models in order to be able to predict even complex transport chains with varying process configurations. In the presented use case, this was only done in a simplified way with logical operators. Here, the combination of ML and simulation methods could offer a high potential. For a high user acceptance in real-time operation, the comprehensibility of the predictions will be of great importance. Against this backdrop, approaches of Explainable AI can be highly valuable. Finally, the testing of ETA predictions under real conditions represents another important step in future research. In this phase, further technical and organizational challenges arise, such as the handling of missing data, the re-training of the models and also the user acceptance of ML-based information.

CONCLUSION

In this chapter, a methodology and ML-based approach for implementing ETA predictions for complex transport chains in logistics was presented. The evaluation results based on a real use case show that supervised ML techniques offer a high potential to accurately predict the duration of logistic processes and possible disruptions. However, this quality is fundamentally influenced by the availability and scope of training data fundamentally. Particularly in transport, where numerous companies are often involved, the cross-actor data integration poses a great challenge for ML projects. The optimal ML method and the suitable features are not known in advance and have to be determined in numerous development and evaluation runs. This leads to a highly iterative process, whereas the involvement of domain experts was recognized as an important criterion for success. The provided results and the shown challenges of using ML in the operational environment could be used for further activated in this application field. In general, it can be concluded that the use of ML for logistics represents a high potential for improving process flows through intelligent decision support systems. This applies not only to the presented use case of ETA prediction, which makes an important contribution to increasing the efficiency, reliability and flexibility of logistics networks by higher transparency about process times of transport orders. It also concerns other areas of application of high uncertainty and complexity, most of which have not yet been researched. Interdisciplinary research between ML and logistics will therefore play an important role in the future.

ACKNOWLEDGMENT

The authors thank all participants for their cooperation in the project as well as their confidence in providing comprehensive data and further company-related information. Special thanks apply to the employees of DB Cargo AG and the representatives of the other involved subsidiaries of Deutsche Bahn AG.

The results of this paper are based on the research project SMECS (Smart Event Forecast for Seaports) funded by the German Federal Ministry of Transport and Digital Infrastructure (IHATEC, 2017).

REFERENCES

Alpaydin, E. (2010). *Introduction to machine learning*. The MIT Press.

Awad, M., & Khanna, R. (2015). *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*. Apress Open. doi:10.1007/978-1-4302-5990-9

Berger, A., Gebhardt, A., Müller-Hannemann, M., & Ostrowski, M. (2011). Stochastic delay prediction in large train networks. In Caprara, A., & Kontogiannis, S. (Eds.), *11th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems* (pp. 100–111). OASICS. 10.4230/OASIcs.ATMOS.2011.100

Bodunov, O., Schmidt, F., Martin, A., Brito, A., & Fetzer, C. (2018). Real-time destination and ETA prediction for maritime traffic. In *DEBS '18: Proceedings of the 12th ACM International Conference on Distributed and Event-based Systems* (pp. 198–201). New York: ACM. 10.1145/3210284.3220502

Büker, T., & Seybold, B. (2012). Stochastic modelling of delay propagation in large networks. *Journal of Rail Transport Planning & Management*, *2*(1-2), 34–50. doi:10.1016/j.jrtpm.2012.10.001

Döbel, I., Leis, M., Molina Vogelsang, M., Welz, J., Neustroev, D., Petzka, H., Riemer, A., Rüping, S., Voss, A., & Wegele, M. (2018). *Maschinelles Lernen: Eine Analyse zu Kompetenzen, Forschung und Anwendung*. Fraunhofer-Gesellschaft.

Fan, W., & Gurmu, Z. (2015). Dynamic travel time prediction models for buses using only GPS data. *International Journal of Transportation Science and Technology*, *4*(4), 353–366. doi:10.1016/S2046-0430(16)30168-X

Gorman, M. F. (2009). Statistical estimation of railroad congestion delay. *Transportation Research Part E, Logistics and Transportation Review*, *45*(3), 446–456. doi:10.1016/j.tre.2008.08.004

Handfield, R., Straube, F., Pfohl, H.-C., & Wieland, A. (2013). *Trends and strategies in logistics and supply chain management: Embracing global logistics complexity to drive market advantage*. DVV Media Group.

Hornung, R. (2020). Ordinal Forests. *Journal of Classification*, *37*(1), 4–17. doi:10.100700357-018-9302-x

Huang, P., Wen, C., Fu, L., Peng, Q., & Tang, Y. (2020). A deep learning approach for multi-attribute data: A study of train delay prediction in railway systems. *Information Sciences*, *516*, 234–253. doi:10.1016/j.ins.2019.12.053

IHATEC. (2017). *SMECS - Smart Event Forecast for Seaports*. https://www.innovativehafentechnologien.de/wpcontent/uploads/2017/09/IHATEC_Projektsteckbrief_SMECS_formatiert.pdf

Lechtenberg, S., Siqueira Braga, D., & Hellingrath, B. (2019). Automatic Identification System (AIS) data-based ship-supply forecasting. In *Proceedings of the Hamburg International Conference of Logistics 2019* (pp. 2–24). Berlin: Epubli. 10.15480/882.2487

Li, X., & Bai, R. (2016). Freight vehicle travel time prediction using gradient boosting regression tree. In *15th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1010–1015). New York: IEEE. 10.1109/ICMLA.2016.0182

Marković, N., Milinković, S., Tikhonov, K. S., & Schonfeld, P. (2015). Analyzing passenger train arrival delays with support vector regression. *Transportation Research Part C, Emerging Technologies*, *56*(1), 251–262. doi:10.1016/j.trc.2015.04.004

Murphy, K. P. (2012). *Machine Learning: A probabilistic perspective*. The MIT Press.

Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press.

Oneto, L., Fumeo, E., Clerico, G., Canepa, R., Papa, F., Dambra, C., Mazzino, N., & Anguita, D. (2018). Train delay prediction systems: A big data analytics perspective. *Big Data Research*, *11*(3), 54–64. doi:10.1016/j.bdr.2017.05.002

Parolas, I., Tavasszy, L., Kourounioti, I., & van Duin, R. (2016). Prediction of vessel's Estimated Time of Arrival (ETA) using machine learning: A port of Rotterdam case study. In *Transportation Research Board 96th Annual Meeting Compendium of Papers*. Washington, DC: Transportation Research Board.

Poschmann, P., Weinke, M., Balster, A., Straube, F., Friedrich, H., & Ludwig, A. (2019). Realization of ETA predictions for intermodal logistics networks using artificial intelligence. In Clausen U., Langkau S., & Kreuz F. (Eds), *Advances in production, logistics and traffic: Proceedings of the 4th interdisciplinary conference on production logistics and traffic 2019* (pp. 155–176). Cham: Springer. 10.1007/978-3-030-13535-5_12

Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: A modern approach* (3rd ed.). Prentice-Hall.

Servos, N., Liu, X., Teucke, M., & Freitag, M. (2020). Travel time prediction in a multimodal freight transport relation using machine learning algorithms. *Logistics*, *4*(1), 1. Advance online publication. doi:10.3390/logistics4010001

Straube, F. (Ed.). (2019). *Trends and Strategies in Logistics: Pathway of Digital Transformation in Logistics Best Practice Concepts and Future Developments*. Universitätsverlag der TU Berlin.

Wahlster, W. (2017). Künstliche Intelligenz als Grundlage autonomer Systeme. *Informatik-Spektrum*, *40*(5), 409–418. doi:10.100700287-017-1049-y

Walter, F. (2015). *Informationsaustausch in der maritimen Transportkette*. Springer. doi:10.1007/978-3-658-09661-8

Wen, C., Lessan, J., Fu, L., Huang, P., & Jiang, C. (2017). Data-driven models for predicting delay recovery in high-speed rail. In *4th International Conference on Transportation Information and Safety (ICTIS)* (pp. 144–151). Piscataway, NJ: IEEE. 10.1109/ICTIS.2017.8047758

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining* (pp. 29–39). Academic Press.

Zeileis, A., Hothorn, T., & Hornik, K. (2008). Model-Based Recursive Partitioning. *Journal of Computational and Graphical Statistics*, *17*(2), 492–514. doi:10.1198/106186008X319331

ADDITIONAL READING

Alessandrini, A., Mazzarella, F., & Vespe, M. (2019). Estimated Time of Arrival Using Historical Vessel Tracking Data. *IEEE Transactions on Intelligent Transportation Systems*, *20*(1), 7–15. doi:10.1109/TITS.2017.2789279

Cerreto, F., Nielsen, B. F., Nielsen, O. A., & Harrod, S. S. (2018). Application of Data Clustering to Railway Delay Pattern Recognition. *Journal of Advanced Transportation*, *2018*, 1–18. Advance online publication. doi:10.1155/2018/6164534

Oneto, L., Buselli, I., Lulli, A., Canepa, R., Petralli, S., & Anguita, D. (2020). A dynamic, interpretable, and robust hybrid data analytics system for train movements in large-scale railway networks. *International Journal of Data Science and Analytics*, *9*(1), 95–111. doi:10.100741060-018-00171-z

Poschmann, P., Weinke, M., Straube, F., Kliewer, J., & Gerhardt, F. (2022). Künstliche Intelligenz in der Binnenschifffahrt: Steigerung der Zuverlässigkeit von Binnenschifftransporten durch datenbasierte Ankunftszeitprognosen. *Internationales Verkehrswesen, 74*(2).

Straube, F., Poschmann, P., Weinke, M., Friedrich, H., Ludwig, A., & Balster, A. (2020). *Smart Event Forecast for Seaports (SMECS) – Schlussbericht*. Technische Universität Berlin, Fachgebiet Logistik.

Strottmann Kern, C., de Medeiros, I. P., & Yoneyama, T. (2015). Data-driven aircraft estimated time of arrival prediction. In *2015 Annual IEEE Systems Conference (SysCon) Proceedings* (pp. 727–733). 10.1109/SYSCON.2015.7116837

Wang, R., & Work, D. B. (2015). Data driven approaches for passenger train delay estimation. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems* (pp. 535–540). 10.1109/ITSC.2015.94

Weinke, M., Poschmann, P., & Straube, F. (2021). Decision-making in Multimodal Supply Chains using Machine Learning. In Adapting to the future: how digitalization shapes sustainable logistics and resilient supply chain management (pp. 301–326). Berlin: Epubli. doi:10.15480/882.3991

KEY TERMS AND DEFINITIONS

Accuracy: Error measure for assessing the quality of a prediction (classification) which corresponds to the proportion of correctly predicted test cases to all test cases.

Estimated Time of Arrival: Expected arrival time of a vehicle, container or shipment at a defined location considering the current conditions.

Feature: Input variable of an ML model, containing formalized and known information about the problem to be learned.

Intermodal Transport: Transport chain which comprises multiple modes of transport (e.g. rail, road).

Label: Target variable of a supervised ML model that is usually known in the training process and is to be predicted when the model is applied.

Machine Learning: Sub-field of Artificial Intelligence which comprises various methods that enable computer systems to extract patterns from data.

Marshalling Yard: Railroad yard used for separating and sorting railroad cars and forming freight trains.

Mean Squared Error: Error measure for assessing the quality of a prediction (regression) which is determined as the mean value of the squared deviations between the predicted and actual values.

Transport Chain: Sequence of several transport and transshipment processes for the shipment of goods from an origin to a destination.