

## **Artificial intelligence for last-mile logistics – Procedures and architecture**

**André Rosendorff**, University of Applied Sciences for Telecommunications (HfTL), Germany, [andre.rosendorff@web.de](mailto:andre.rosendorff@web.de)

**Alexander Hodes**, University of Applied Sciences for Telecommunications Leipzig (HfTL), Germany, [alexander.hodes@live.com](mailto:alexander.hodes@live.com)

**Benjamin Fabian**, Technical University of Applied Sciences Wildau (TH Wildau), Germany, [benjamin.fabian@th-wildau.de](mailto:benjamin.fabian@th-wildau.de)

### **Abstract**

*Artificial Intelligence (AI) is becoming increasingly important in many industries due to its diverse areas of application and potential. In logistics in particular, increasing customer demands and the growth in shipment volumes are leading to difficulties in forecasting delivery times, especially for the last mile. This paper explores the potential of using AI to improve delivery forecasting. For this purpose, a structured theoretical solution approach and a method for improving delivery forecasting using AI are presented. In doing so, the important phases of the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, a standard process for data mining, are adopted and discussed in detail to illustrate the complexity and importance of each task such as data preparation or evaluation. Subsequently, by embedding the described solution into an overall system architecture for information systems, ideas for the integration of the solution into the complexity of real information systems for logistics are given.*

**Keywords:** Supply chain management, logistics, artificial intelligence, machine learning, business intelligence.

### **Introduction**

A particular challenge in logistics is last-mile delivery, which is the journey of the package from the last delivery base to the final recipient (Hsiao et al., 2018). This part is considered one of the most expensive parts of the supply chain and is also difficult to plan (Hsiao et al., 2018). Challenging factors include the number of possible delivery scenarios, relevant parameters, and the growing influence of a package's recipient. For example, nowadays, the recipients want to indicate their availability for a shipment to be delivered at the desired time or to a desired neighbor. When a recipient is absent, the highest share of delivery costs is incurred in the last mile (Gevaers et al., 2009; 2014). The absence of a receiver is often caused by the lack of an accurate time forecast of the delivery (Digiesi et al., 2017). Gevaers et al. (2014) described the unsuccessful delivery scenario using the first-time hit rate (FTHR). This percentage rate indicates the average successful delivery attempts to recipients. In the absence of more accurate last-mile information, this rate is necessarily low (Gevaers et al., 2009; 2014).

To help parcel recipients adjust to the delivery time, service providers can offer forecasts of when the parcel will be delivered with a certain lead time. The accuracy of such forecasts depends on many influencing factors. The delivery time indicated by such a forecast also depends on the service provider. For example, at the time of writing, DPD in Germany gives a forecast for the day of delivery at intervals of one hour, while Hermes gives a forecast of two to four hours. A more accurate forecast, along with a shorter delivery time, is a customer requirement of growing importance. An industry study by DPD indicated that the majority of customers include an accurate delivery forecast in their purchase decision (DPDgroup, 2017). A study by Crockford et al. (2013) also found a strong desire for delivery within a specific time window.

This increasing business demand motivates our current study. In order to better plan supply chain processes and avoid incurring high costs, accurate prediction of delivery time is required. Based on the current importance and potential of the technology, our article explores how Artificial Intelligence (AI) and Machine Learning can be used to predict the duration of delivery of a shipment in the final part of the supply chain. We also developed a corresponding process and information system architecture. To this end, we define a general framework for last-mile logistics.

Our article is structured as follows. First, we describe the current state of research in Artificial Intelligence in logistics. Then, we describe the objectives and methods of our study. We then present how delivery forecasting can be improved, including a discussion of the different phases of optimization required to develop the solution. In the next section, we integrate this theoretical solution into a possible overall technical architecture, presenting the solution together with the individual systems of its environment. Finally, we discuss the contributions and limitations of our work and draw a conclusion.

## **Related Work**

Some previous studies discuss the use of Artificial Intelligence and Machine Learning in logistics and supply chain management. For example, Souza (2014) described the use of methods of regression analysis to make demand forecasts more accurate in planning. According to Souza (2014), several external factors influence demand, including demographic and technological trends and competitive information. Demand forecasts are used to better plan manufacturing and production, such as implementing pricing strategies. Döring (2010) described the use of Machine Learning for the automated and intelligent control of change plans in production networks. The material flow is automatically controlled by using clustering methods of unsupervised learning. Runkler et al. (2010) described neuronal forecasts and swarm-based optimization for the assignment of production lots to production units in production logistics. This replaces manual planning and rule-based systems.

An example of an industry case is Cargonexx (2019), which uses AI in logistics to organize the freight truck traffic more efficiently by better predicting cargo flows (Cichosz, 2018). As another use case, Singapore's SingPost has developed an application for optimizing vehicle routes using AI (Jucha, 2021; Urquhart, 2018). This is motivated by the problem of inefficiency in terms of resource use by vehicles and employees.

Furthermore, Ko et al. (2020) compared different last mile solutions that involve cooperation of individual actors in the delivery process. Lu et al. (2020) focused their work on driver-related issues to develop approaches for improving efficiency. Especially during the peak season from Black Friday to Christmas, many seasonal drivers are hired to handle the increased delivery volume (Soltes, 2014). From their point of view, such use of temporary workers is often not well enough organized and does not lead to a positive effect in terms of time and costs (Lu et al., 2020). Lu et al. (2020) proposed an approach of dividing routes into no-helper and with-helper routes. By a corresponding mathematical analysis, they find that parameters such as fuel costs and time can be reduced.

There are further ideas for using AI for last mile optimization with efficient vehicle routing, such as Vehicle Route Optimization (VRO). This optimization can lead to minimal cost and time improvements for visiting a range of locations (Open Data Science, 2019). We suggest implementing such models in three phases, starting with local improvements, followed by fleet management improvements. Finally, demand forecasting can be improved to support customer satisfaction (Open Data Science, 2019).

## **Goals and Methods**

A major goal of this research is to describe a structured approach (Peffer et al., 2012) to optimize last mile delivery forecasting by using AI. This optimization will be based on a framework for the development of AI projects in order to represent the individual steps in a comprehensible way. The developed optimization will then be represented in an overall information systems architecture, including interfaces and corresponding necessary data flows between the individual systems. This will serve as a first step for the development and evaluation of an all-encompassing prototype (Peffer et al., 2012).

A highly promising approach for optimization is the development of a Machine Learning model for predictive analysis of delivery time for the last part of the supply chain. This can be based on the historical and constantly growing database of service providers. For predictive analysis, methods of supervised learning are suitable, which involve the dedicated training of a model on historical data and then applying it to predict the target variable for current cases (Kelleher et al., 2015).

In our work, synthetically generated data is used for representation and visualization. This is aimed to be very similar to real data and should illustrate the complexity of a real use case. The data was generated using a data generator that we developed in Python. It reflects several fundamental assumptions on real-world logistics. The individual attributes of the data and their calculation can be found in Table 1. This data generator and all further scripts for data preparation and forecasting can be accessed via the following GitHub repository: <https://github.com/alexanderhodes/artificial-intelligence-for-last-mile-logistics>.

**Table 1.** Calculation of Attributes by the Data Generator

Variable	Calculation	Description
Region	Random number between 1 and 5	Representation of the different areas of delivery according to the number of stops and longer distances
Duration	A sum depending on region, weather extreme and traffic	Delivery time depending on region, weather extreme and traffic
Kilometers	Random number depending on region	Distance in kilometers to be covered for the delivery
Number of Stops	Random number depending on region	Number of stops that are present during delivery
Weather Extreme	Random value from none, rain and snow	Weather conditions that influence the duration of delivery
Traffic	Random value between 0 and 100	Traffic affecting the duration of delivery

For the practical realization of the optimization of the delivery forecast, the use of an established industry standard for data mining is an important way to simplify the integration, updating and maintenance of the system (Grossman et al., 2002). Machine Learning methods are a central component of data mining and AI. Using a process model that is established and accepted in the field leads to faster and more cost-effective implementation, as well as an increase in project success (Gordon, 2018). Various data mining standards and process models are available for projects, for an overview see (Kurgan & Musilek, 2006).

The widely established CRISP-DM process is used in our study. It has been proven in various applications and is sufficiently documented and tested (Garrido Azevedo et al., 2007). Due to the widespread use and popularity of the CRISP-DM process model, it leads to a good and common understanding of projects (Garrido Azevedo et al., 2007). In addition, it is considered well suited for newcomers, as well as specialists (Shearer, 2000). Due to its industry focus, it is also a well-understood approach for developing solutions in the service sector. An important prerequisite for a data-mining model is the existence of historical data records of last-mile shipments. The machine learning model to be developed will identify patterns and correlations between the data in order to forecast the delivery of future shipments.

### **Optimizing the Delivery Forecast**

**Business Understanding.** Designing the solution requires an understanding of the underlying logistics business and knowledge of the business processes in which the application is to be used. Based on this, the requirements and goals can be defined from a business perspective. Identifying the variable to be predicted by the Machine Learning model and finding suitable predictors form the central step of this phase. In the delivery prediction use case, the variable to be studied is the duration of a shipment on the last mile. This also requires determining the input variables. The machine learning model can learn patterns and dependencies between these variables, also called features, and delivery time as the target.

The three objectives of the variable selection (Guyon & Elisseeff, 2003) are divided as follows:

1. Improving the predictive performance of predictors.
2. Providing faster and more cost-effective predictors.
3. Providing a better understanding of the underlying business process that generates the data.

There are several methods for finding appropriate variables from a technical perspective (Hearty, 2016; Kelleher et al., 2015). Table 2 lists important input variables on which delivery time can be considered dependent.

**Table 2.** Input Variables with the Reason for Selection

Variable/ Predictor	Reason for selection
Destination	A destination address is required to determine the route that the carrier must cover.
Kilometer	The length of the route to be traveled influences the delivery time.
Number of stops	The number of stops that are necessary influences the delivery time.
Weather extremes	Weather conditions affect traffic.
Traffic load	The traffic load affects the delivery time.

It should be noted that there may be further factors that can affect the delivery time. These are to be determined in future work. Since the variable to be predicted is a continuous numerical value, regression techniques must be applied (Cuesta, 2013). Due to the dependency between the development of the models and the available data, it is recommended to identify appropriate data sources at this stage (Cuesta, 2013). One can obtain and merge the data for the Machine Learning algorithms from multiple sources (Prill et al., 2017).

A data warehouse system could be suitable for this purpose. Data on historical deliveries needs to be identified, including the time of delivery and other last mile events and timestamps. An example of this is the scan of the parcel when the delivery vehicle is loaded. One of the possible data sources could be a tracking system of the logistics service provider where historical shipments are stored. Other sources include weather systems that provide weather data for the time of delivery. In addition, traffic systems or mapping services must be included for route creation and planning. Traffic loads can also be derived from these.

**Data Understanding.** When developing a prediction model, it is necessary to recognize relationships in the data (Cuesta, 2013). To develop a general understanding of the data and its metadata, it is recommended to create data dictionaries with information about the schema and descriptions of the variables. This was also done for the supply forecasting use case based on (McDaniel, 1994), which is shown in Table 3.

**Table 3.** Example of a Data Dictionary for Delivery Forecasting

Variable	Data Type	Values	Description	Source
Duration	Float	$x > 0.0$	Time measurement in hours from the start of delivery to handover	Shipment tracking system
Region	Integer	$1 \leq x \leq 5$	Destination region	Shipment tracking system
Kilometers	Float	$x > 0.0$	Number of kilometers the carrier has to travel from the delivery base to the final destination	Traffic data system
Number of Stops	Integer	$x \geq 0$	Number of stops the carrier has to make in order to deliver packages	Shipment tracking system
Weather Extreme	String	none, rain, snow	Weather conditions at the time of delivery	Weather data system
Traffic	Float	$0.0 < x \leq 100.0$	Traffic load in percent at the time of the delivery	Traffic data system

After describing the data, it is necessary to further examine it (Kelleher et al., 2015). For this goal, the data can be visualized graphically. Such visualization is used to derive insights about the distribution and relationship of the data. Various techniques can be used for visualization. For optimization, a method of visualizing the relationships between pairs of features is helpful to examine which descriptive features may be useful for predicting delivery times. Kelleher et al. (2015) recommend creating a Scatter PLOt Matrix (SPLOM) to show the dispersion of the data (Kelleher et al., 2015). This allows the group characteristics of the data to be examined. Figure 1 shows an SPLOM for our generated data set.

**Data Preparation.** This phase includes all activities to create the final data set that will be used to train and evaluate the model. In practice, this does not take place in a fixed defined process, but often in several steps with no prescribed order. Techniques for preparing and processing the data are described in (Cooley et al., 1999; Kelleher et al., 2015). Examples of such techniques include the creation of new attributes and the formatting or transforming data (Kelleher et al., 2015; McLellan et al., 2003; Soibelman & Hyunjo, 2002). Single attribute transformations may be necessary to transform symbolic fields into numeric values (Kelleher et al., 2015; McLellan et al., 2003; Soibelman & Hyunjo, 2002). Modeling tools and Machine Learning algorithms often require this transformation (Kelleher et al., 2015; McLellan et al., 2003; Soibelman & Hyunjo, 2002). For this reason, such transformation is performed for non-numeric attributes of the example dataset. For the specified variables, only the categorical value of the weather conditions is affected. The transformation increases the number of dimensions by the number of categorical values of the weather conditions. Based on this, the models can be trained by algorithms in the next phase to provide predictions for previously unseen data.

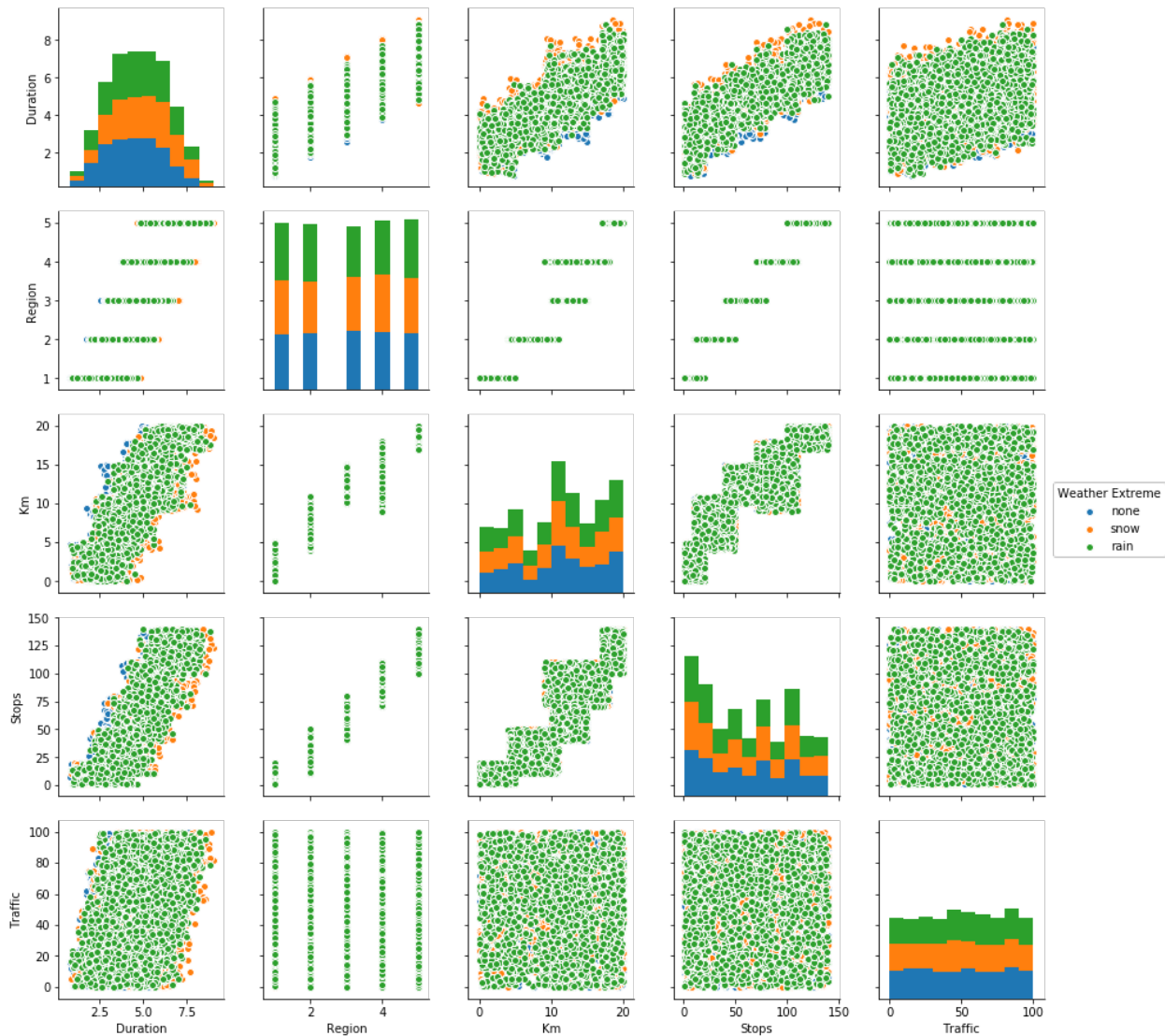


Figure 1. Scanner Plot Matrix for Example Data Set

**Modeling.** In the modeling phase, the training of the models is performed based on the processed data from the past. The goal is to obtain information about the behavior to be predicted from the historical data (Garrido Azevedo et al., 2007; Kurgan & Musilek, 2006). For the application of supervised learning, the delivery time must be predicted as an output variable Y from the input feature vector X established in the previous phases (Carrière et al., 2018). The steps of the modeling phase include selecting the modeling techniques, creating the test design, and creating the models with the selected Machine Learning algorithms (Shearer, 2000; Wirth & Hipp, 2000). Because some modeling techniques have specific data format requirements, it may be necessary to return to the previous phase of data preparation (Shearer, 2000; Wirth & Hipp, 2000).

Due to the a priori uncertainty of which technique is best suited for the problem, the selection may be based on multiple algorithms. The results of the learning algorithms can be examined and compared on an empirical basis in the next phase (Shearer, 2000; Wirth & Hipp, 2000). Finally, it can be determined which technique is most appropriate for the data set used.

For the evaluation, the data is divided into training and test sets. The prediction model is built with the training data. It thus contains the learned knowledge about the past that was extracted in the training procedure. The quality of the model is checked with the separate test data. It can be tested how well the model responds to unknown data. In addition, the model can be tested for underfitting and overfitting. Testing the models can help to select the model with the algorithm that gives the best results.

A simple and effective method for evaluation is cross-validation (Guyon et al., 2011). The reason for using this method is that it provides more representative estimates of the model quality. Alternatively, k-fold cross-validation can be used, in which the split is performed k times (Bengio & Chapados, 2003). Finally, the results of the k models are obtained. Splitting the data set into the three parts of training, parameter tuning, and test set is also possible.

At the end of the phase, the prediction models have to be built with the chosen learning algorithms. A key step here is the calibration of the hyperparameters to optimal values (Garrido Azevedo et al., 2007; Kurgan & Musilek, 2006). Hyperparameters are parameters to configure the learning algorithms. They must be optimized before the training is performed. Setting the parameters is crucial to evaluate the potential of a method and achieve good results. The available parameters differ from algorithm to algorithm. Techniques for optimizing hyperparameters are described in (Bergstra et al., 2013; Thornton et al., 2013). Grid search is one of the most commonly used strategies; due to its ease of implementation and widespread use, it has become the standard (Bergstra & Bengio, 2012).

**Evaluation.** After the prediction models have been developed, it is necessary to check and evaluate the quality of the prediction results before deployment. This supports the final selection of the learning algorithm. Our example evaluation is based on the metrics listed in Table 4, which are commonly used in practice to evaluate the performance of prediction models for continuous values. The model's predictions for the test set are compared to the actual target values, according to the test design created in the previous phase. The goal is to measure how accurately the predicted values match the correct ones (Kelleher et al., 2015).

Another way to compare Machine Learning algorithms is to plot a learning curve. The plot allows you to estimate, for each algorithm, how the size of the training set affects the accuracy of the predictions. In addition, the plot is helpful to detect underfitting or overfitting. For example, if the curve for training data increases while it decreases for test data, a loss of generalization ability of the model can be inferred. In such a case, it may be helpful to increase the number of data items in the training set (Bonoaccorso, 2018).



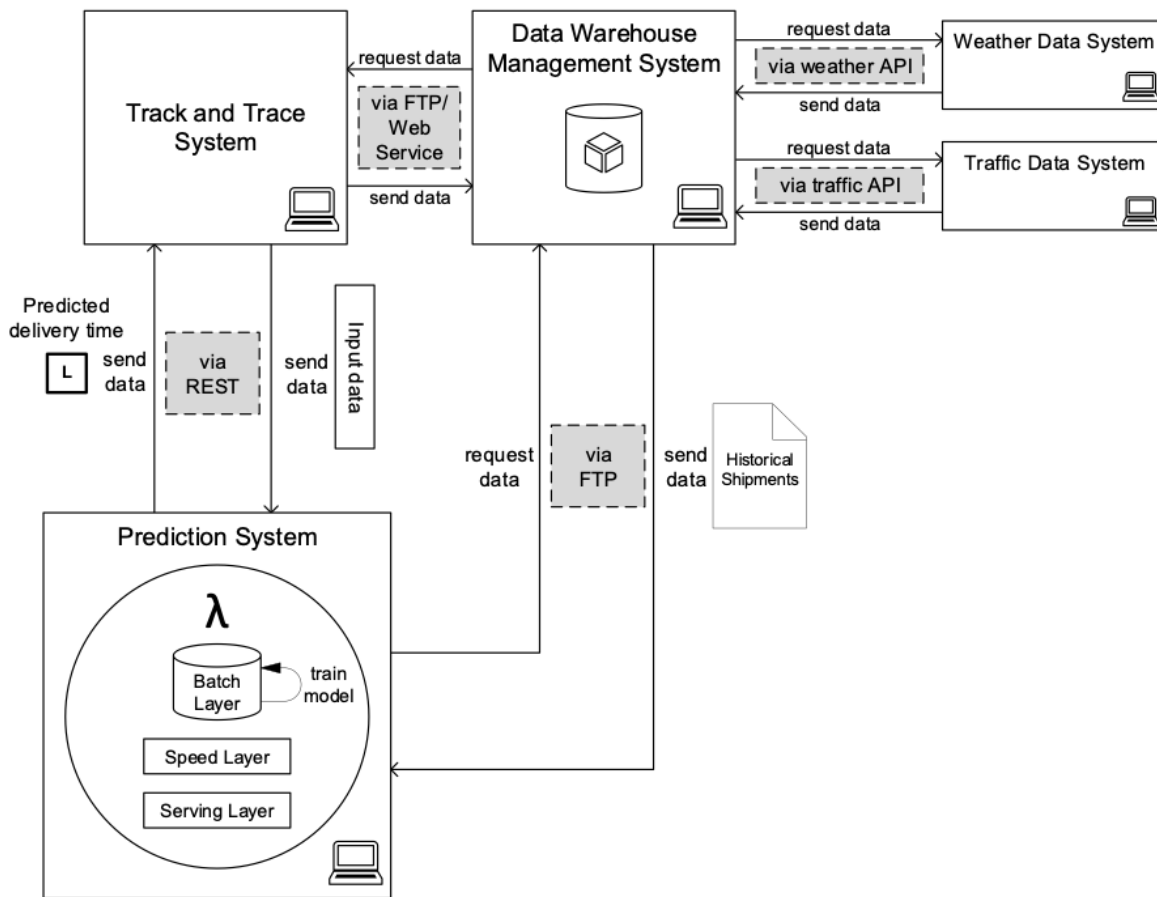
**Table 4.** Evaluation Metrics for the Predictions

Key number	Formula	Value range	Best value
R <sup>2</sup>	$\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$	[0;1]	1
MSE	$\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$	[0;∞]	0
RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$	[0;∞]	0
MAE	$\frac{1}{n} \sum_{i=1}^n  \hat{y}_i - y_i $	[0;∞]	0

**Deployment.** Since prediction models are based on the assumption that patterns learned from the training data are representative of future data, the implementation of an ongoing validation scheme must be planned during the deployment phase. For example, routes may change due to road reconstruction or other reasons. As a result, relationships between features that have been learned in advance no longer apply. Techniques to determine when the model begins to be outdated are described in (Géron, 2017; Kelleher et al., 2015). Another starting point is to retrain the model after a certain time interval. This can be done by incorporating the newly added data to keep the model up to date (Kelleher et al., 2015).

### **Integration into an Information Systems Architecture**

After the theoretical concept and procedures for predicting delivery time have been designed, the solution is concretized into a possible architecture. As a first consideration, the Business AI Framework (Gentsch, 2018) is adapted to our use case, not shown here for brevity. In addition to considering the architecture from a business perspective, there is also a technical perspective. A possible Big Data architecture model with the integration of a prediction system is shown in Figure 2. A more detailed view of the Lambda architecture is shown in Figure 3. The diagrams illustrate the relationships and communication between the individual systems. The central component is the forecasting system, which is directly connected to the shipment tracking system of the logistics service provider. The Machine Learning model is provided on this system.



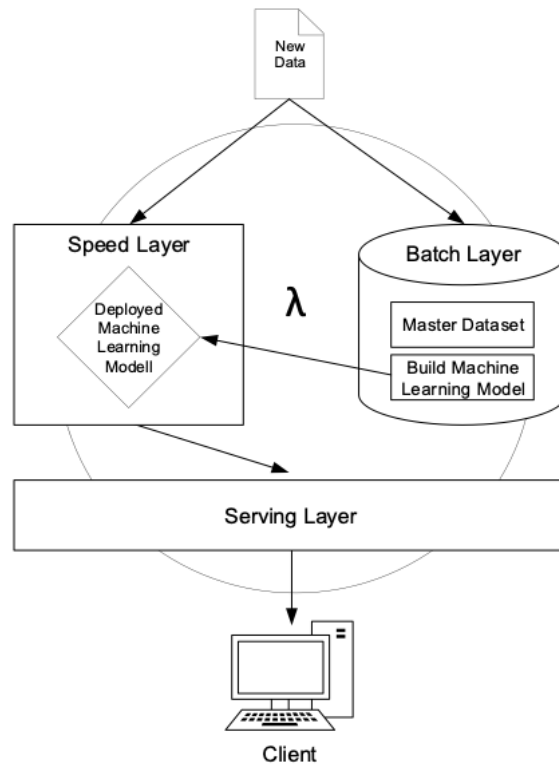
**Figure 2.** Architecture for Last-Mile Prediction

The Lambda architecture (Marz & Warren, 2015) is very promising the implementation of the forecasting system. It provides efficient solutions for processing large amounts of data. Positive characteristics of the concept are high robustness, fault tolerance, scalability, generalizability, and extensibility. In addition, such a system requires minimal maintenance. According to (Hasani et al., 2014), the Lambda architecture is the first approach that addresses the complexity of Big Data by defining clear principles.

On the batch layer, the historical data can be stored as a master dataset and then adapted to the Machine Learning model. After the model has been built, it can be made available on the Speed Layer. If the delivery time for a shipment is to be predicted, the input data is transferred from the tracking system to the Speed Layer to calculate the forecast. The result is then sent back to the client via the Serving Layer. If the model needs to be re-trained with new data, this can be done on the Batch Layer. During training, which can take several hours depending on the size of the database, further calculations can be performed on the Speed Layer.

For training of the Machine Learning model, data is stored centrally in a data warehouse where data from multiple sources is merged. For example, historical shipment data originates from the

shipment tracking system. This can be transferred to the data warehouse via File Transfer Protocol (FTP) or Web Services. External data sources are required for weather and traffic data.



**Figure 3.** Integration of the Machine Learning Model into the Lambda Architecture

Once the model has been fitted to the historical data, it can be made available for shipment tracking. When planning the last mile route, the tracking system passes the input data to the Machine Learning model to create the feature vector with the variables to predict the delivery time. The model calculates a forecast based on the learned relationships and sends the result back to the tracking system. A web service can be provided for data exchange between the tracking system and the Machine Learning model. An application programming interface (API) based on Representational State Transfer (REST) could be used for this purpose.

Once the delivery forecast is known, customers can be notified of the estimated time of arrival by SMS or e-mail. This allows them to better plan their presence at the time of delivery. If it is already clear that the customer cannot be present at the specified time, it must be possible to specify the absence and the desired delivery time and thus avoid an unsuccessful delivery attempt in order to save costs.

## Discussion

In this article, we have described the procedure for optimizing the delivery forecast using AI based on the CRISP-DM process model. The first contribution of our work is the development of a

theoretical model for the use of AI in logistics for delivery time prediction. Another contribution is the description of a structured process for developing this optimization. We have shown which data is necessary for the calculation of the prediction. With the description of each phase, the necessary steps to develop the solution and the corresponding procedures were shown. By generating data representing the individual influencing factors, the steps show how a delivery forecast can be generated from the data. Based on the theoretical model and the structured procedure, a foundation for using AI to optimize delivery forecasting has been established. This is kept general and comprehensible by using synthetically generated data. However, this should provide a basis for the development of a prototype that can be used to account for the complexity of real-world data.

In terms of constraints, our research is so far limited to theoretically structured aspects (Peffer et al., 2012) of improving last mile delivery prediction through AI. In future work, a practical implementation of the solution is planned, where a prototype will be developed based on the described theory and evaluated under real-world conditions. Another area of research could be the selection of additional suitable variables to predict delivery times. For example, weather conditions could be extended to include temperature, precipitation height, wind direction or wind speed. In addition to traffic density, the number of road intersections, the number of traffic lights or speed zones on the delivery route could also be determined.

Furthermore, it would be useful to analyze whether better forecasting results can be achieved by including days of the week and times of the day for deliveries. Suitable methods must be found for variable selection. In the future, it will also be necessary to integrate and test further methods of automatic learning for prediction. Since the current work only looked at last mile delivery time prediction, looking at the entire supply chain provides an additional area of research, in particular to investigate how AI can be used to predict the overall supply chain delivery time. In combination, this would allow us to look at implementing a comprehensive system for the individual supply chain players. If all partners provide selected data to a central system, a wide range of further applications for analyzing and evaluating the data will arise (Fabian et al., 2013). Some of these applications could also be developed by recent decentralized approaches such as Federated Learning. In addition, the combination and integration of other important technologies such as the Internet of Things or Blockchain must be investigated.

## **Conclusion**

In this article, we have provided a theoretical solution approach and information systems architecture for improving delivery forecasting using AI. Although AI is already being applied in other areas of logistics, there is a lack of concrete and operational frameworks for the last mile, especially for delivery time forecasting. Since it is important to adhere to industry standards in data science, an established process model should be applied. For this purpose, our work adopted the important phases of the CRISP-DM framework, which describe the individual steps of optimization.

Another important aspect is the integration of data from different heterogeneous source systems. Our examples demonstrate the complexity and importance of the individual tasks such as data preparation or evaluation. Furthermore, the importance of choosing appropriate metrics and key

performance indicators to evaluate the forecasts is highlighted. By embedding our process-based theoretical solution for last-mile delivery forecasts into an overall architecture, we have provided the first concrete steps towards integrating the solution into the complexity of real-world applications. The individual necessary data flows between the systems were visualized and described.

In the future, we aim to put our approach into practice to test and further optimize the solution in a real-world environment. This should involve iterative modeling and optimization using real-world data. At the same time, it would be very helpful to carry out pilot tests with practice partners to apply, evaluate, and further refine the processes and architecture proposed in this article.

## References

- Bengio, Y., & Chapados, N. (2003). Extensions to metric-based model selection. *Journal of Machine Learning Research*, 3, 1209–1227.
- Bergstra, J., Yamins, D., & Cox, D. (2013). Making a science of model search: Hyperparameter optimization. *Proceedings of Machine Learning Research*, 28(1), 115–123.
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(10), 281–305.
- Bonoaccorso, G. (2018). *Machine learning algorithms: Popular algorithms for data science and machine learning* (2nd ed.). Packt Publishing.
- Cargonexx. (2019). *Cutting-edge technologies for even more affordable and easier transports*. <https://www.cargonexx.com/en/shipper>
- Carrière, M., Mechel, B., & Oudot, S. (2018). Statistical analysis and parameter selection for mapper. *Journal of Machine Learning Research*, 19(12), 1–39.
- Cichosz, M. (2018). Digitalization and competitiveness in the logistics service industry. *E-mentor*, 5 (77), 73-82. <https://doi.org/10.15219/em77.1392>
- Cooley, R., Mobasher, B., & Srivastava, J. (1999). Data preparation for mining world wide web browsing patterns. *Knowledge and Information Systems*, 1(1), 5–32.
- Crockford, G., Ritschel, F., & Schmieder, U. M. (2013). *Handel in theorie und praxis*. Springer.
- Cuesta, H. (2013). *Practical data analysis: Transform, model, and visualize your data through hands-on projects, developed in open source tools*. Packt Publishing.
- Digiesi, S., Fanti, M. P., Mummolo, G., & Silvestri, B. (2017). Externalities reduction strategies in last mile logistics: A review. *Proceedings 2017 IEEE International Conference on Service Operations and Logistics, and Informatics*. IEEE.
- Döring, A. (2010). Intelligente Steuerung der Änderungsplanung in Produktionsnetzwerken der Serienfertigung. *KI - Künstliche Intelligenz*, 24(2), 165–168.
- DPDgroup. (2017). *E-shopper barometer: Länderreport Deutschland*. <https://www.dpd.com/de/en/2017/02/20/erstes-e-shopper-barometer-der-dpdgroup-deutschland-ist-europameister-im-online-shopping/>

- 
- Fabian, B., Kunz, S., Müller, S., & Günther, O. (2013). Secure federation of semantic information services. *Decision Support Systems*, 55(1), 385–398.
- Garrido Azevedo, S., Ferreira, J., & Leitão, J. (2007). The role of logistics' information and communication technologies in promoting competitive advantages of the firm. *Munich Personal RePEc Archive*, Paper 1359.
- Gentsch, P. (2018). *Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und Bots zu einem Algorithmic Business - Konzepte, Technologien und Best Practices*. Springer.
- Géron, A. (2017). *Hands-on machine learning with Scikit-Learn & TensorFlow*. O'Reilly.
- Gevaers, R., van de Voorde, E., & Vanellander, T. (2009). Characteristics of innovations in last-mile logistics – Using best practices, case studies and making the link with green and sustainable logistics. *Proceedings of the European Transport Conference 2009*. Association for European Transport.
- Gevaers, R., van de Voorde, E., & Vanellander, T. (2014). Cost modelling and simulation of last-mile characteristics in an innovative B2C supply chain environment with implications on urban areas and cities. *Procedia – Social and Behavioral Sciences*, 125, 398–411. <https://doi.org/10.1016/j.sbspro.2014.01.1483>
- Gordon, A. (2018). *Can AI really improve industrial production efficiency?* Forbes. <https://www.forbes.com/sites/forbestechcouncil/2018/07/02/can-ai-really-improve-industrial-production-efficiency/#2293e488145f>
- Grossman, R. L., Hornick, M. F., & Meyer, G. (2002). Data mining standards initiatives. *Communications of the ACM*, 45(8), 59–61. <https://doi.org/10.1145/545151.545180>
- Guyon, I., Cawley, G., Dror, G., & Saffari, A. (2011). *Hands-on pattern recognition: Challenges in machine learning, Volume 1*. Microtome Publishing.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157–1182.
- Hasani, Z., Kon-Popovska, M., & Velinov, G. (2014). Lambda architecture for real time big data analytic. *ICT Innovations Conference Web Proceedings*, (pp. 133–143). Association for Information and Communication Technologies.
- Hearty, J. (2016). *Advanced machine learning with Python*. Packt Publishing.
- Hsiao, Y. H., Chen, M. C., Lu, K. Y., & Chin, C. L. (2018). Last-mile distribution planning for fruit-and-vegetable cold chains. *The International Journal of Logistics Management*, 29(3), 862–886.
- Jucha, P. (2021). Use of artificial intelligence in last mile delivery. *Proceedings of the 20<sup>th</sup> International Scientific Conference Globalization and its Socio-Economic Consequences 2020, SHS Web of Conferences*, 92. <https://doi.org/10.1051/shsconf/20219204011>
- Kelleher, J. D., MacNamee, B., & D'Arcy, A. (2015). *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*. MIT Press.

- 
- Ko, S. Y., Sari, R. P., Makhmudov, M., & Ko, C. S. (2020). Collaboration model for service clustering in last-mile delivery. *Sustainability*, 12(14), 5844. <https://doi.org/10.3390/su12145844>
- Kurgan, L. A., & Musilek, P. (2006). A survey of knowledge discovery and data mining process models. *The Knowledge Engineering Review*, 21(1), 1–24.
- Lu, SH., Suzuki, Y., Clottey, T. (2020). The last mile: Managing driver helper dispatching for package delivery services. *Journal of Business Logistics*, 41(3), 206–221.
- Marz, N., & Warren, J. (2015). *Big data: Principles and best practices of scalable real-time data systems*. Manning Publications.
- McDaniel, G. (1994). *IBM dictionary of computing*. IBM.
- McLellan E., MacQueen K. M., & Neidig J. (2003). Beyond the qualitative interview: Data preparation and transcription. *Field Methods*, 15(1), 63-84.
- Open Data Science. (2019, October 14). *AI as the ultimate disrupter in logistics: How to manage last-mile costs?* Medium. <https://medium.com/@ODSC/ai-as-the-ultimate-disrupter-in-logistics-how-to-manage-last-mile-costs-c4874e8f2ea0>
- Peffer, K., Rothenberger, M., Tuunanen, T., & Vaezi, R. (2012). Design science research evaluation. *Proceedings of the International Conference on Design Science Research in Information Systems*, (pp. 398–410). Springer.
- Prill, D., Kranzer, S., & Merz, R. (2017). Improving maintenance processes with data science. In Haber, P., Lampoltshammer T., & Mayr, M. (Eds.), *Data Science – Analytics and Applications. Proceedings of the 1st International Data Science Conference (iDSC2017)*, (pp. 105–107). Springer.
- Runkler, T. A., Grothmann, R., & Bamberger, J. (2010). Optimierung industrieller Logistikprozesse mit Verfahren der Schwarmintelligenz und rekurrenten neuronalen Netzen. *KI – Künstliche Intelligenz*, 24(2), 149–152.
- Shearer, C. (2000). The CRISP-DM Model: The new blueprint for data mining. *Journal of Data Warehousing*, 5(4), 13–22. <https://mineracaodados.files.wordpress.com/2012/04/the-crisp-dm-model-the-new-blueprint-for-data-mining-shearer-colin.pdf>
- Soibelman, L., & Hyunjo, K. (2002). Data preparation process for construction knowledge generation through knowledge discovery in databases. *Journal of Computing in Civil Engineering*, 16(1), 39-48.
- Soltes, F. (2014). *Special delivery*. National Retail Federation. <https://nrf.com/blog/special-delivery>
- Souza, G. C. (2014). Supply chain analytics. *Business Horizons*, 57(5), 595–605.
- Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2013). Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In Y. Ding (Ed.), *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2013)*, (pp. 847–855). ACM Press.

Urquhart, D. (2018). *SingPost integrates AI into its Southeast Asia-wide logistics platform*. Asiarcargobuzz. <https://asiarcargobuzz.com/2018/12/11/singpost-integrates-ai-into-its-southeast-asiawide-logistics-platform/>

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

## Authors Biographies

**André Rosendorff, B.Sc.** After studying business informatics at the University of Applied Sciences for Telecommunications Leipzig, André Rosendorff began studying Data Science (M.Sc.) at the Albstadt-Sigmaringen University of Applied Sciences. In parallel, he works as a Data Science Consultant. His research interests are in the area of Artificial Intelligence, especially Computer Vision, MLOps and Natural Language Processing.



**Alexander Hodes, M.Sc.** Alexander Hodes currently works as a full-stack developer at a startup in Munich. Previously, he studied Business Informatics (B.Sc.) at the University of Applied Sciences for Telecommunications Leipzig and Distributed and Mobile Applications (M.Sc.) at the Wilhelm Büchner University of Applied Sciences. His research focuses on the development of mobile applications and possible uses of blockchain.



**Benjamin Fabian, Prof. Dr. habil.** Ben Fabian is currently Professor of E-Government, IT Security, and IT Management at the Technical University of Applied Sciences Wildau and Senior Researcher for Information Systems at Humboldt University of Berlin. Previously, he was Professor of Business Intelligence and Data Science at the University of Applied Sciences for Telecommunications Leipzig. His research interests are Data Science, Information Security, Privacy and Ethics, with a focus on design science research, software engineering, mathematical and empirical methods at the intersection of these areas.

