AUTOMATION OF STANDBY DUTY PLANNING FOR RESCUE DRIVERS VIA A FORECASTING

CASE STUDY

Author: Jelson Lino

29 February 2024

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

1. 1. Business Understanding

1. 1. 1. Business Problem

The current standby-duty plan for Berlin's red-cross rescue service struggles with inefficiencies, leading to situations where there are not enough standby drivers or too many standby drivers are kept on hold. The HR planning department wants to improve the current planning logic by incorporating predictive models to estimate the daily number of standby rescue drivers more accurately.

1. 1. 2. Project Objectives

This project aims to develop a predictive system to estimate the number of standby rescue drivers required, improve the efficiency of the standby-duty plan by increasing the percentage of standby drivers being activated compared to the current approach of keeping 90 drivers on hold, reduce situations where there are not enough standby drivers or excess standby drivers, ensure that the new planning model can incorporate seasonal patterns and other factors affecting the number of required standby drivers, and meet the deadline for finishing the duty plan on the 15th of the current month for the upcoming month.

1. 1. 3. Success Criteria

The success criteria of this project include increasing the percentage of standby drivers being activated compared to the current approach, reducing the number of situations where there are not enough standby drivers or excess standby drivers, developing a predictive system with a reasonable level of accuracy in estimating the daily number of standby drivers required, and meeting the deadline for finishing the duty plan on the 15th of the current month for the upcoming month consistently.

Project Plan

The execution of this project followed the following steps:
- Understand the current standby-duty plan and its shortcomings.
- Understand and analyze the data provided.
- Develop predictive models to estimate the daily number of standby drivers required.
- Provide recommendations on the implementation of the selected predictive system into the planning process.

1. 1. 4. Resources

A Windows laptop with RAM and processor of 8.00 GB and Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz, respectively, was used to run Python and its libraries and frameworks to process the data and

train and evaluate the machine learning algorithms. In addition, Brisqi was used as a project management tool.

2. DATA UNDERSTANDING

The data was stored in a csv file. Pandas was used to load and understand the data. There were eight columns. None of the columns had any nun values nor what could be considered outliers. One of the columns had object-type data stored in it, while the remaining columns stored numerical data. The following description defines the nature of each column in the dataset.

• date: entry date

• n_sick: number of drivers called sick on duty

• calls: number of emergency calls

• n_duty: number of drivers on duty available

• n_sby: number of standby resources available

• sby_need: number of standbys, which are activated on a given day

• dafted: number of additional drivers needed due to not enough standbys

To improve readability and facilitate analysis, the columns were renamed based on the description given above. There was an Unnamed column, which was simply an index column and not relevant to data modeling. Therefore, it was removed from the data. The date column stores object data. However, the information stored in it is date data. Therefore, the format of this column was changed from object to datetime.

3. DATA PREPARATION

The seasonality claimed by the HR department was checked by plotting the data.

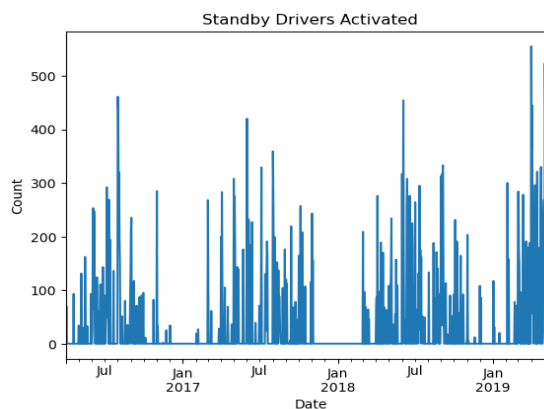Figure 1: Activated Standby Drivers

Figure 1 indicates a seasonal pattern in the number of activated standby drivers. The peak season is in the Summer, while the Winter registers the lowest numbers of activated standby drivers.

The data also suggest that there are days in which no standby driver is activated. To further understand the data, Figure 2 shows the relationship between the target variable and all the independent variables.

Figure 2: Relationship Between the Target Feature and Other Variables
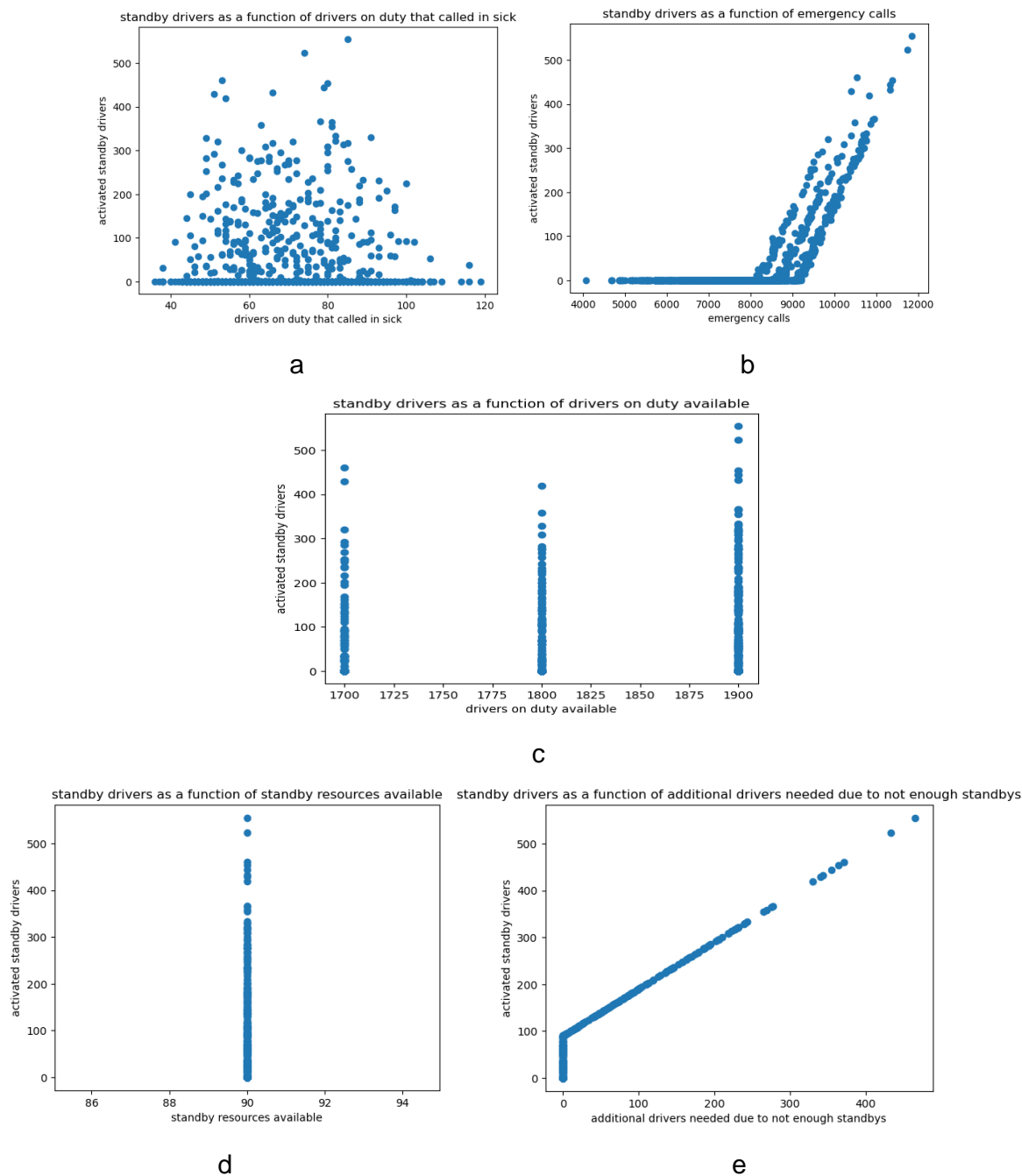

a

b

c

d

e

Figure 2a does not show a clear mathematical relationship between the target variable and sick drivers. Nevertheless, this variable still has an impact on the target variable. Figure 2b shows that there is a threshold after which the target variable and emergency calls are linearly related. The three lines that appear after the threshold are due to the three unique values found in the number of available drivers, as shown in Figure 2c. Figure 2c suggests that the number of drivers available increased twice

since 2016. Figure 2d shows that the number of standby resources available is always 90. Thus, this column should not be used in the training of the machine learning models. The constant value of this column is the reason why there is a threshold of 90 in Figure 2e. Figure 2e shows a linear relationship between the target variable and the number of additional drivers needed due to insufficient standby drivers after the threshold of 90 activated standby drivers is reached. This is because additional drivers are only needed when the 90 standby drivers cannot respond to all the calls that the on-duty drivers are unable to respond to. The following figures show the seasonal pattern of each feature.

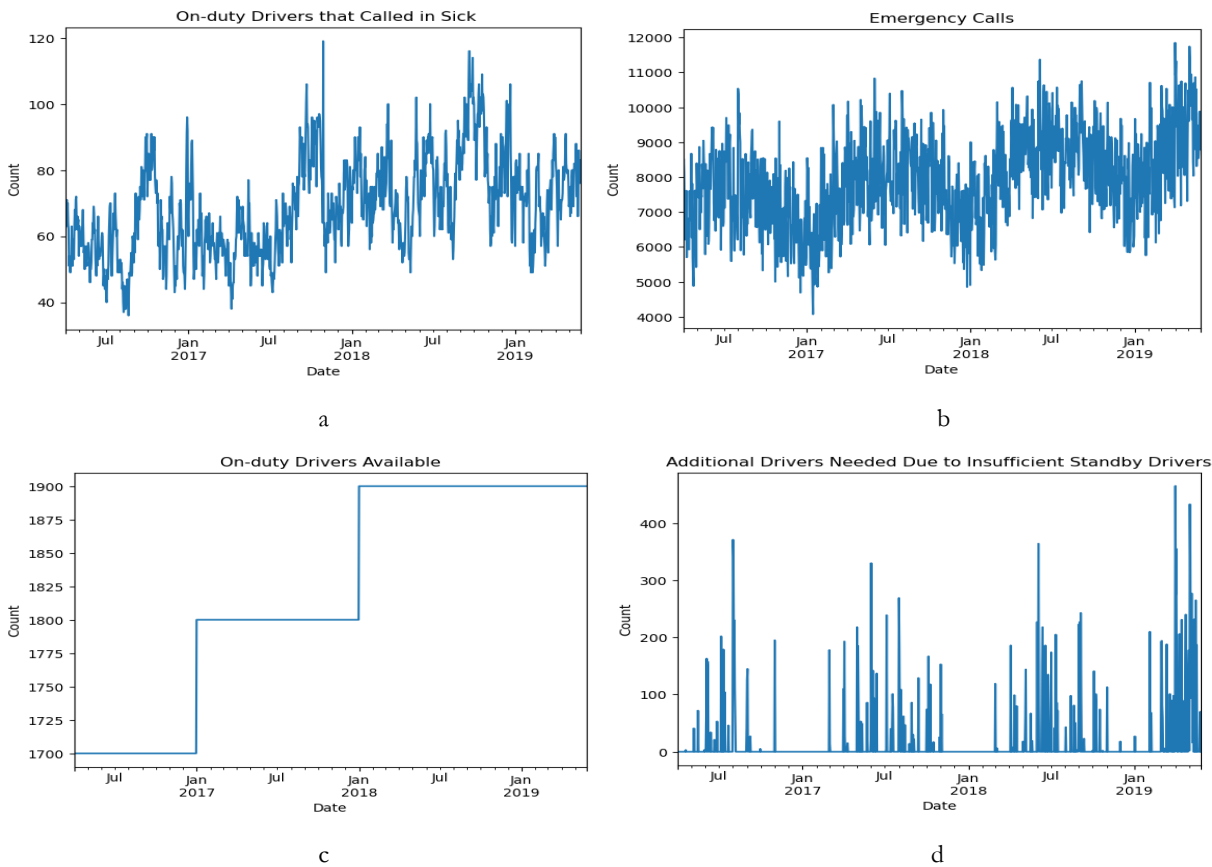Figure 3: Features Over Time

a

b

c

d

Figure 3 shows a clear seasonal pattern on most variables. The fact that, regardless of the number of on-duty drivers calling in sick, the number of on-duty drivers available is almost always constant raises concern about the quality and veracity of the data. However, we will assume that when n on-duty drivers call in sick, n standby drivers are put on duty to replace those who called in sick, and n non-standby drivers are put on standby to replace those who are now on duty.

A date is a data point that does not repeat itself; it only appears once in the dataset. Because we are interested in building a model that can grasp the seasonal aspect of the data, three more variables were created from the date column; they are month, day of month, and day of week. The values of these variables can appear more than once and can be extracted from any given month. Thus,

given the date whose activated standby drivers one wants to predict, the values of these columns are extracted and fed into the predictive system. Although the data plots shown above show the nature of each variable, to further understand the relationship between variables, a correlation heatmap is shown below.

Figure 4: Correlation Heatmap
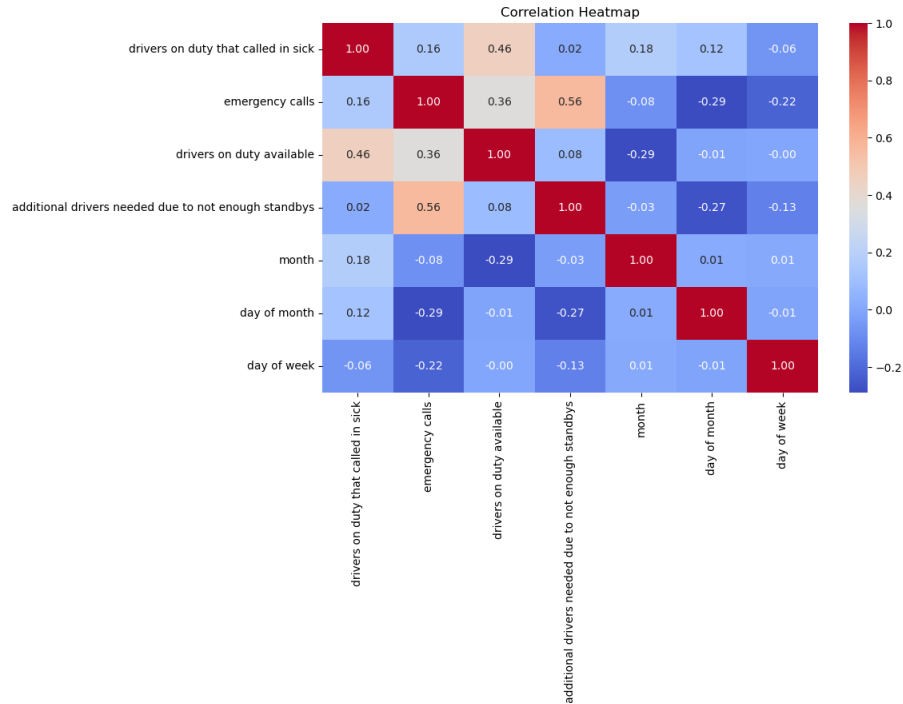
There is a moderate correlation between emergency calls and additional drivers needed due to not enough standby drivers. This correlation is expected because the higher the volume of calls, the higher the number of drivers activated and the higher the number of additional drivers needed due to not enough standby drivers. Because the variable standby resources available is always 90, the higher the number of calls, the higher the number of additional standby drivers needed (once the threshold of calls that the on-duty and the 90 standby drives can bear is reached). After the threshold of 90, these two variables are expected to be directly proportional. Nevertheless, the objective is to predict the number of standby drivers activated so that the standby drivers are set based on the prediction and ideally the number of additional drivers is kept at zero.

In addition, this variable (additional drivers needed due to not enough standby drivers) only exists because the number of standby drivers available is always 90. If we set the number of standby drivers available based on the prediction for activated standby drivers, then we will be able to keep additional drivers at zero (or close to zero). Therefore, the variable additional drivers needed due to not enough standby drivers was dropped from the analysis.

Based on the understanding we have of the data, the number of additional drivers needed due to not enough standbys + the number of standby resources available is equal to the number of standbys

activated, when the number of standbys activated is equal to or greater than 90. When the number of standby drivers activated is less than 90 (or the number of standby resources available), the number of additional drivers needed due to not enough standby drivers is zero. Because of this, added to the fact that it is a constant, the column standby resources available was dropped from the analysis.

4. MODELING

As demonstrated, there are temporal patterns in the data and there are independent features that do have some impact on the target feature. For example, the number of calls on a given day does have an impact on the number of standby drivers activated. Therefore, to predict the number of standby drivers activated, we need to first predict all the variables that have an impact on our target feature (standby drivers activated).

Based on Figure 13, to predict the target feature, the values of six independent variables are required (considering that additional drivers needed due to insufficient standby drivers is dropped). Three of those variables are obtained from the current date, that is the date whose standby drivers activated is being predicted. The variables obtained from the date are month, day of month, and day of week. The variables on-duty drivers that called in sick, emergency calls, and on-duty drivers that are available have temporal patterns, as shown in Figure 3. In addition, due to the nature of the variable being predicted (a numerical continuous variable), a supervised regression model is appropriate for this project. The three variables whose values cannot be extracted from date are predicted separately using different time series models. This means that the entire prediction system works based on four supervised regression models. Figure 5 illustrates how the system works.

Figure 5: Standby Drivers Activated Prediction System
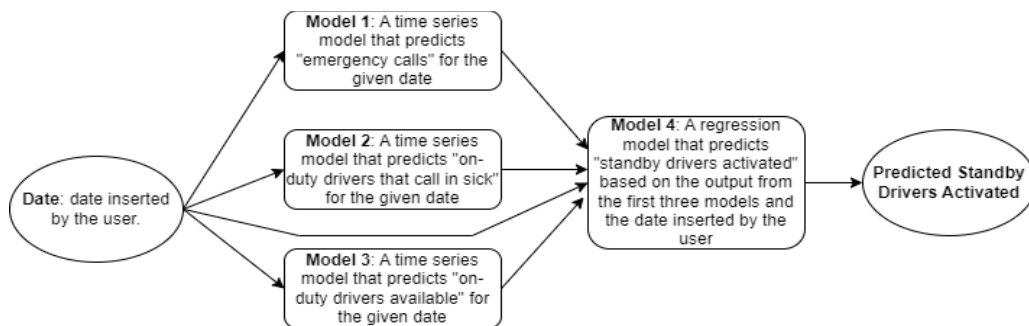
Three time-series models (models 1, 2, and 3) predict on-duty drivers that called in sick, emergency calls, and on-duty drivers that are available. The output from these models and the intended month, day of the week, and day of the month are fed into the main model (model 4) that predicts the number of standby drivers activated. The available data can be used to run the four models of the entire system. In this report, model 4 is presented first. However, during operation, the order presented in Figure 5 is kept.

4. 1. Model 4

As shown in Figure 5, model 4 is the main model of the system. In this section, the training and evaluation of model 4 are presented. Because the target variable (standby drivers activated) is a numerical continuous variable, regression is the appropriate machine learning method to be used in this project. In addition, the target variable shows a temporal pattern. Thus, a multivariate time series model could be used. However, due to extremely poor performance, multivariate time series models were put aside.

4. 1. 2. Splitting Data Vertically and Horizontally

The data was split into the target variable and independent variables. All variables, except standby drivers activated, were selected as independent variables, while the variable standby drivers activated was selected as the target feature. Furthermore, the data was also split into training and test sets, with 80% of the data being used for training and the remaining being used for testing.

A baseline model was developed, as shown below, against which all subsequent models were compared. Seven supervised regression models were trained using the same data. The objective was to evaluate various models and select the best one. The metrics used to evaluate the models were mean-squared error (MSE) and r-squared. MSE represents the average squared difference between the predicted values and the actual values. Lower MSE indicates better predictive performance, with values closer to zero suggesting a better fit. R2, also known as the coefficient of determination, measures the proportion of the variance in the target variable that is predictable from the independent variables. It ranges from 0 to 1, where 1 indicates a perfect fit and 0 indicates no linear relationship between the variables. In this project, any model with values worse than those of the baseline model was considered unsuitable. Models performing better than the baseline were compared against each other. Figure 15 shows the models and their metrics.
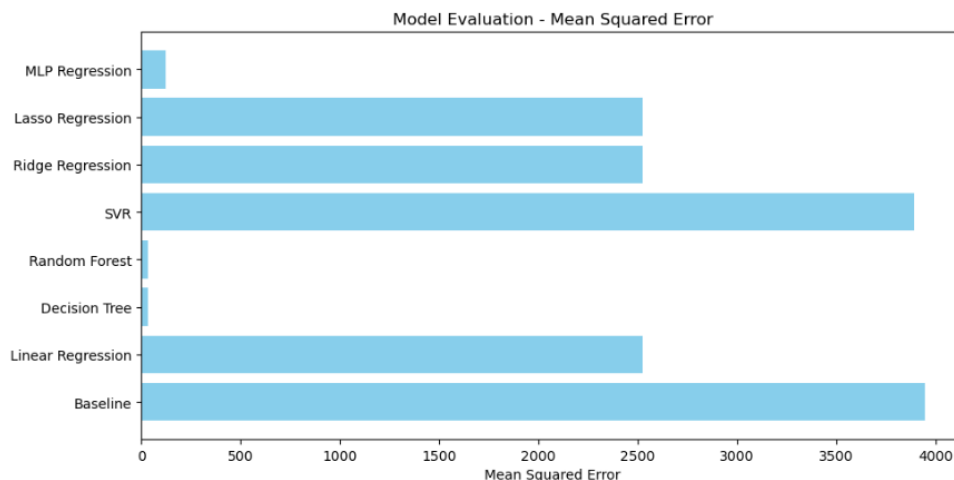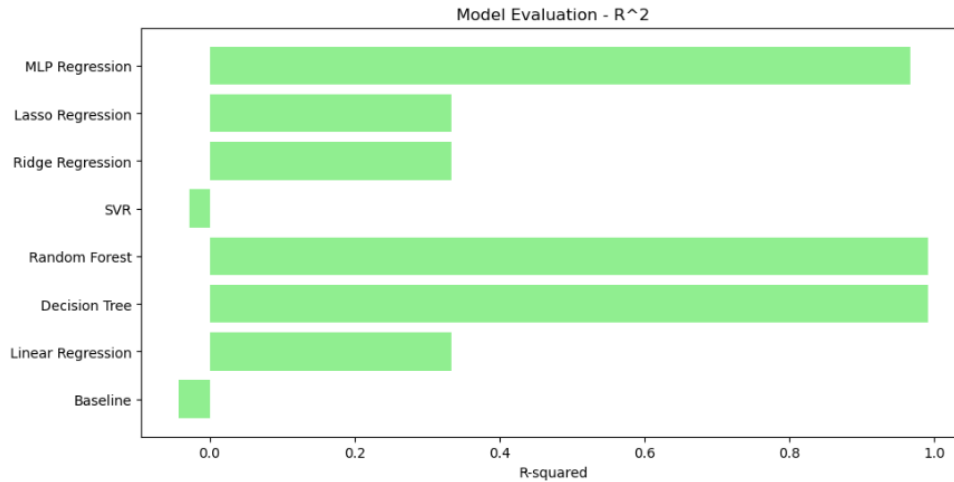
Figure 6a: Model Evaluation - MSE

Figure 6b: Model Evaluation - R2


Model Evaluation - R^2

Figures 6a and 6b show that the decision tree model is the best-performing model, as it has the best value in both metrics, and is closely followed by the random forest model. SVR (support vector regressor) is the worst-performing model. Therefore, the decision tree model is used as model 4 of Figure 5.

4. 2. Model 1

The data on the column emergency calls was split into 80% and 20% for training and testing, respectively, while the Date column was used as the index. Then, a SARIMAX model was trained. The model achieved an MSE of 254.5966. Figure 16 shows the performance of this model compared with the actual values.

Figure 7: Emergency Calls
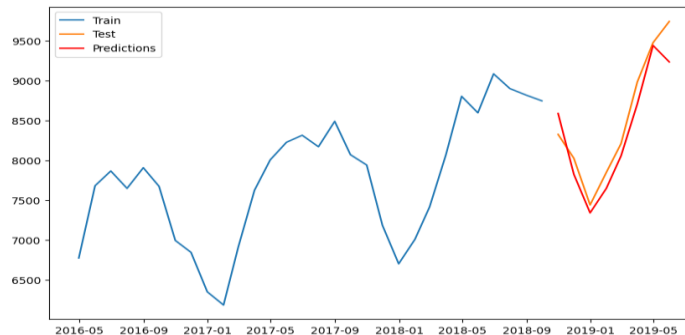
4. 3. Models 2 & 3

Models 2 and 3 were built using a convolutional neural network designed based on Keras with a TensorFlow backend. 916 data points were used as training data, while 115 and 114 data points were used as validation and testing data, respectively. In both models, 7 previous values are used to predict the next value. Models 2 and 3 had MSE values of 41.30 and 0.641, respectively. Figures 8a and 8b show the performance of models 2 and 3, respectively.
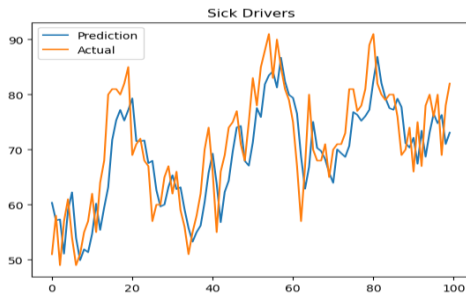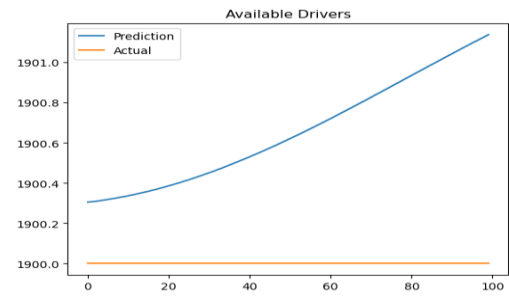
Figure 8a: Model 2 - Drivers That Call in Sick

Figure 8b: Model 3 - Drivers Available

5. EVALUATION

MSE indicates the average squared difference between the predicted and actual values. A lower MSE suggests better predictive performance. All seven machine learning models considered for model 4 have an MSE lower than that of the baseline model. However, the lowest MSE is that of the decision tree model. This means that the decision tree model has the smallest deviation from the true values. The MSE of the decision tree model is 32.5; this implies that the mean absolute deviation from the true values is approximately 5.7. Considering that the values of the variable that model 4 predicts range from zero to more than 500, an error of +/- 5.7 is reasonably acceptable. In practice, it means that when the model predicts, for example, 500 standby drivers activated, there could be an error of at least 5.7, making the true value between 494 and 506. However, as part of the data fed into model 4 is already a prediction from three models, the difference between the predicted and true values of model 4 can slightly increase because the output from the three models could already be slightly deviated from the true values. Models 1, 2, and 3 have MSE values of 254.49, 41.30, and 0.6411, respectively. As shown in Figure 9, the variable emergency calls is the most important variable in the prediction of standby drivers activated. This variable is predicted by model 1 and has an MSE of 254.59, which is acceptable, given that the values predicted can be higher than 9500. An MSE of 254.59 implies that the mean absolute difference between the predicted and true values is only 15.96.

Figure 9: Feature Importance
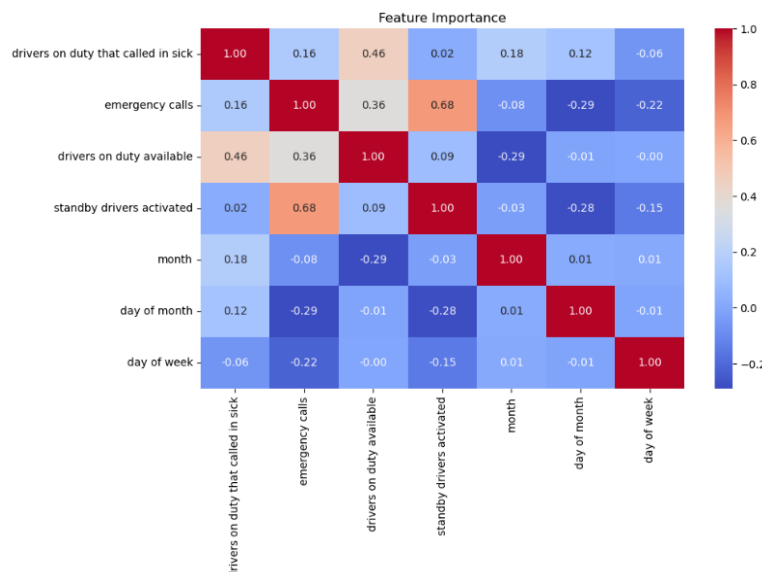
The results indicate that this system can help the HR department plan more accurately by minimizing situations where too many drivers are kept on hold. In addition, the number of additional drivers needed due to insufficient standby drivers will also be minimized; however, there is no guarantee that this number will always be zero.

6. DEPLOYMENT

As shown in Figure 5, the entire system uses four models. The system can be used by the HR team as a web app deployed on Azure or any other suitable cloud service provider. The web app can be used in two manners. The HR team can access the web app via URL (http triggered) and make predictions as often as necessary or the program can be set to automatically run (time triggered) and produce a report with the predictions on the 14th day of every month.

The first approach allows for experimentation, and, in case of any unexpected change or issue, the HR team can rerun the program as many times as necessary without involving the analytics/IT team. In the second approach, the process is completely automated, which means that the HR team does not interact directly with the program as the team simply receives an automatic report; this process allows for a completely automated scheduling service. However, in case of any unexpected change or issue, the analytics/IT team will have to respond to the change/issue.

Whatever the option used, the model's performance will have to be monitored regularly and the system should be retrained at least once every two months (because in two months enough new data is produced to retrain the system). However, if recent predictions deviate significantly from the true values, the model must be retrained immediately, regardless of whether or not two months have already passed from the last retraining.

7. CONCLUSION

In this project, a system was developed that predicts the number of standby drivers activated and automates and optimizes the standby-duty plan of the Red Cross in Berlin by increasing the percentage of standby drivers being activated, thus, reducing situations where there are too many or too few standby drivers. The system, which is composed of four supervised regression algorithms (one SARIMAX, two CNNs, and one Decision Tree), incorporates seasonal patterns. The system can be deployed on Azure, or any other cloud service provider, as a web app to allow the HR team to open it from their browser. Further studies can be conducted, once additional data is available, to improve the mean squared error of the SARIMAX model and the first CNN model. The predicted and actual values of every month should be recorded and saved for performance monitoring and retraining of the entire system.