Probabilistic Forecasting Of Pedestrian Flow Using Machine Learning

Luisa Dosch

2777143

Bachelor Thesis

Chair for Information Systems and Business Analytics
University Würzburg

Supervisor: Prof. Dr. Christoph Flath
Assistant: Dr. Nikolai Stein

Würzburg, 31.10.2024

Contents

List of Figures

List of Tables

German Summary

Diese Bachelorarbeit befasst sich mit der probabilistischen Vorhersage des Fußgängerverkehrs in Würzburg unter Einsatz fortgeschrittener Methoden des maschinellen Lernens (ML) und Deep Learning (DL). Ziel ist es, Einzelhandelsunternehmen durch präzisere Vorhersagen in den Bereichen Bestandsverwaltung und Personaleinsatzplanung zu unterstützen. Zur Analyse der Fußgängerdaten auf unterschiedlichen Zeitskalen, einschließlich Kurzzeit- (2-Tage) und Langzeitprognosen (7-Tage), wurden Modelle wie CatBoost, LightGBM, Random Forest, Quantilregression und Long Short-Term Memory (LSTM) verwendet.

Die Analyse zeigt, dass zeitliche Faktoren, wie der Wochentag, einen entscheidenden Einfluss auf den Fußgängerverkehr haben. Samstags wurde der höchste und Sonntags der niedrigste Verkehr gemessen. Zudem wurden saisonale Schwankungen festgestellt, mit erhöhtem Fußgängerverkehr im Juli und Dezember. Stündlich betrachtet liegen die Verkehrsspitzen zwischen 11:00 und 13:00 Uhr. Das Modell CatBoost erzielte sowohl bei Kurz- als auch Langzeitvorhersagen die höchste Genauigkeit und übertraf dabei alle anderen Methoden.

Zu den wichtigsten Einflussfaktoren für die Prognosen zählen der 7-Tage-Lag der Fußgängerzahlen, der Wochentag und standortspezifische Unterschiede. Die Anwendung dieser Vorhersagemodelle ermöglicht Einzelhändlern eine verbesserte Planung, indem sie ihre Bestände und Schichtpläne besser auf die zu erwartenden Kundenströme abstimmen. Dies trägt zu einer effizienteren Ressourcennutzung und einer gesteigerten Kundenzufriedenheit im Einzelhandel bei.

Abstract

This thesis explores the probabilistic forecasting of pedestrian flow using advanced Machine Learning (ML) and Deep Learning (DL) models to predict foot traffic in Würzburg. The research aims to provide actionable insights for retail operations, focusing on improving inventory management and employee shift planning. Using models such as CatBoost, LightGBM, Random Forest (RF), Quantile Regression and Long Short-Term Memory (LSTM), pedestrian traffic data was analyzed across different temporal scales, including 2-day and 7-day forecasting horizons.

The findings reveal that pedestrian activity is strongly influenced by temporal factors, with Saturdays consistently recording the highest foot traffic and Sundays the lowest. Seasonal patterns were also identified, with traffic peaking in July and December, while hourly trends showed peak activity between 11:00 AM and 1:00 PM. The best-performing model was CatBoost, which achieved the highest accuracy in both the short-term (2-day) and long-term (7-day) forecasts, outperforming all other methods. LSTM was slightly less accurate overall, but showed a performance that can compete with tree-based models, especially for hourly forecasts.

Key predictive drivers include the 7-day lag of pedestrian counts, the day of the week, and location-specific factors. The integration of these forecasting models offers significant improvements in operational planning, helping retailers better align resources such as stock levels and staffing with customer demand patterns. The application of these models enables more efficient decision-making, resulting in enhanced operational efficiency and customer satisfaction in retail settings.

1 Introduction and Context of this Bachelor Thesis

1.1 Introduction

In 1897, the French artist Camille Pissarro captured scenes of Boulevard Montmartre, portraying the flow of pedestrians from his hotel window at various times of the day and in different weather conditions. Today, over a century later, researchers can study these same dynamic patterns of human movement, thanks to advanced digital tools and vast amounts of data (Fan and Loo, 2021). This evolution mirrors the way technology has transformed how we observe and analyze the world around us, particularly in the context of pedestrian traffic.

While Camille Pissarro observed the flow of pedestrians for artistic enjoyment, today accurate foot traffic forecasting has become a key element in optimizing business operations, especially for physical stores like restaurants and retail shops. The ability to predict foot traffic not only helps businesses anticipate customer demand, but it also enables more effective operations management, such as labor scheduling and stock control. The further into the future businesses can forecast pedestrian flow, the more efficiently they can manage resources, reduce waste, and ultimately increase profits. By adjusting staffing levels and managing inventory in response to predicted foot traffic, businesses can improve both efficiency and customer satisfaction (Abrishami, Kumar, and Nienaber, 2017).

With recent advancements in data collection technologies, such as the widespread use of wireless access points and sensors, businesses can now gather foot traffic data at scale. This data can be leveraged to make reliable predictions, allowing for smarter decision-making in areas like inventory management and workforce planning (Abrishami, Kumar, and Nienaber, 2017). The integration of these predictive tools empowers businesses to better prepare for fluctuating demand, ensuring smoother operations and higher profitability.

This bachelor thesis analyzes the probabilistic forecasting of pedestrian flow using ML and DL models. The goal is to predict pedestrian traffic distributions in Würzburg and provide actionable insights for retail businesses. These insights will support more effective inventory management and employee shift planning, ultimately enhancing operational efficiency.

1.2 Research Question

The primary research question for this thesis is:

How can machine learning be used to estimate conditional pedestrian traffic distributions to support decision-making in retail store operations?

This thesis focuses on developing probabilistic forecasting models for pedestrian traffic using ML techniques. The aim is to predict the distribution of foot traffic in Würzburg, providing actionable insights for optimizing key areas of retail operations such as inventory management and employee shift planning. This thesis centers on generating accurate forecasts, which can directly inform decision-making in these essential use cases:

1. **Inventory Management**: Utilizing both 2-day and 7-day pedestrian traffic forecasts allows retail managers to better anticipate customer demand over the short and medium term. These forecasts enable stores to adjust stock levels according to expected customer flows, reducing the likelihood of overstocking or stock shortages. The probabilistic models developed in this thesis offer valuable data to guide inventory decisions, ensuring readiness for fluctuating traffic patterns both in the immediate and upcoming week.
2. **Employee Shift Planning**: The precise hourly traffic forecasts, coupled with the 48 hour and 168 hour projections, are essential for optimizing employee shift planning. By predicting foot traffic trends throughout the day and extending forecasts over several days, retail managers can better allocate staff during peak and off-peak periods. This allows for effective workforce management, ensuring adequate coverage during busy times while controlling labor costs during slower hours.

For both inventory management and employee shift planning, the pedestrian traffic forecasts generated in this thesis are evaluated using two data sets with different levels of granularity. While the primary focus in this thesis is on accurate probabilistic forecasting, the application of these predictions requires combining them with optimization techniques. In practice, these forecasts are used to address stochastic optimization problems, enabling businesses to make informed decisions on resource allocation.

This work enhances the distribution of pedestrian traffic predictions, serving as essential input for optimization tasks. Techniques like Sample Approximation Algorithms (SAA) and Weighted Sample Approximation Algorithms (WSAA) transform the forecasts into actionable strategies, optimizing inventory levels and workforce planning.

In summary, this thesis develops and evaluates probabilistic models for pedestrian traffic, providing the foundational input necessary for solving real-world problems in inventory management and employee shift planning through stochastic optimization.

2 Related Work

DL has revolutionized time series forecasting, with models such as LSTM networks playing a significant role due to their ability to capture long-term dependencies in data. Recurrent Neural Network (RNN)s are often hindered by issues like vanishing gradients, which reduce their effectiveness in long-term forecasting tasks. LSTMs overcome this by incorporating memory blocks, allowing them to retain information over extended periods, making them particularly suitable for tasks involving temporal data (Zhao et al., 2017), such as traffic and pedestrian flow prediction. These models have gained widespread use in studies focused on modeling complex, non-linear temporal dynamics, as seen in applications like traffic and pedestrian flow forecasting. For instance, Harrou et al. (2022) introduced a hybrid model combining LSTM and Variational Autoencoders, further improving the accuracy of pedestrian and bicycle traffic flow predictions.

Despite their strengths, DL models like LSTM do not always outperform traditional ML methods in real-world applications. Research has demonstrated that ensemble-based models, such as Gradient Boosted Trees (GBT), LightGBM, and CatBoost, frequently achieve better results in time series forecasting. In a comparative study, Rady, Fawzy, and Abdel Fattah (2021) found that tree-based methods like RF and GBT provided superior accuracy compared to traditional models like AutoRegressive Integrated Moving Average (ARIMA), highlighting their ability to capture non-linear relationships and manage large feature sets effectively in complex forecasting tasks.

The effectiveness of tree-based models was further underscored in the M5 Uncertainty competition, which focused on probabilistic forecasting. The winning solutions employed LightGBM for quantile regression, demonstrating the model's ability to handle uncertainty across multiple quantiles. (Makridakis et al., 2022).

Uncertainty estimation is vital for prescriptive analytics techniques, which can help make better operational decisions based on available feature information. Schmidt and Pibernik (2023) used a random forest model to estimate the distribution of future demand, and then applied stochastic optimization to make dynamic inventory decisions over time. For shift planning, Notz, Wolf, and Pibernik (2023) introduced a new method that uses a random forest to solve the staffing problem for multiple shifts.

In conclusion, both tree-based ML models and DL techniques have proven effective in time series forecasting. While LSTM networks excel at capturing long-term dependencies, gradient boosting methods such as LightGBM and CatBoost often provide superior performance in practical forecasting scenarios due to their flexibility and predictive accuracy. This research builds on these advancements to forecast pedestrian flow, supporting improved demand prediction and operational planning.

3 Exploratory Data Analysis

This section presents an Exploratory Data Analysis (EDA) of the dataset to uncover patterns, relationships, and irregularities. Understanding these insights is crucial for selecting appropriate models and features and enhancing the accuracy of predictions.

3.1 Data Overview

The dataset contains time series data capturing pedestrian flow at both daily and hourly intervals, providing insights into foot traffic at different temporal resolutions. This data is essential for optimizing operational decisions in inventory management and staff scheduling.

3.1.1 Daily Data

The findings from the daily data set (Würzburg, 2024b) are especially relevant for inventory management. The dataset helps businesses forecast overall demand by analyzing trends at the daily level, which can inform stock level adjustments and improve efficiency.

3.1.2 Hourly Data

Hourly data (Würzburg, 2024a) provides detailed insights into intra-day patterns, crucial for shift planning and labor management. This data allows us to identify peak periods (e.g., morning, afternoon, and evening), enabling better alignment of staff schedules with customer demand.

3.2 Selected Features

The following features in Table 1 were selected from the dataset for analysis from both daily and hourly datasets:

Daily Data Features	**Hourly Data Features**
Weekday	Weekday
Month	Month
Temperature	Temperature
Location	Location
Weather Condition	Weather Condition
	Hour

Table 1: Selected features for daily and hourly data

3.3 Data Cleaning

To ensure a complete and consistent dataset, missing pedestrian counts were filled using the previous day's or hour's value, preserving the continuity of the time series. Additionally, categorical variables such as weekday, month, location, and hour were transformed using one-hot encoding. This technique converts each category into a separate binary feature, making them suitable for use in ML models.

3.4 Descriptive Analysis of Features

This section provides an overview of key features in the dataset, exploring their distribution and how they relate to pedestrian activity.

3.4.1 Distribution of Pedestrian Count

The pedestrian count distributions reveal significant differences across the three locations, as summarized in Figure 1.

Spiegelstraße has the lowest overall foot traffic, with a mean pedestrian count of 13,728 and a relatively small standard deviation of 4,112, indicating more consistent traffic compared to the other locations. The pedestrian counts at this location range from 0 to 29,672, with 50% of the values falling between 12,187 and 16,176. The Coefficient of Variation (CV) is 0.30, reflecting a relatively low variation in pedestrian numbers, suggesting a more stable and predictable flow of foot traffic.

In contrast, **Kaiserstraße** exhibits higher variability, with a mean pedestrian count of 20,362 and a standard deviation of 6,197. Pedestrian counts range from 0 to 34,682, indicating greater fluctuation. The Interquartile Range (IQR) is between 19,586 and 24,057, which shows a more concentrated flow during peak periods. The CV of 0.30 indicates that, despite higher absolute counts, the relative variability is comparable to that of Spiegelstraße, meaning the fluctuations around the mean are proportionally similar.

Schönbornstraße records the highest pedestrian traffic, with a mean of 23,444 and the widest range of counts, spanning from 1,258 to 62,555. The standard deviation of 10,992 is also the largest among the three locations, highlighting the significant variability in traffic. The median count is 24,952, and the upper quartile reaches 30,242, suggesting that this location often experiences substantial foot traffic. The CV of 0.47 reflects a much higher level of variation in pedestrian counts relative to the mean, indicating more extreme fluctuations and less predictability in traffic patterns.

Overall, while Spiegelstraße shows more consistent and stable traffic patterns, both Kaiserstraße and Schönbornstraße experience higher foot traffic with greater variability, with Schönbornstraße demonstrating the most extreme fluctuations in pedestrian counts.

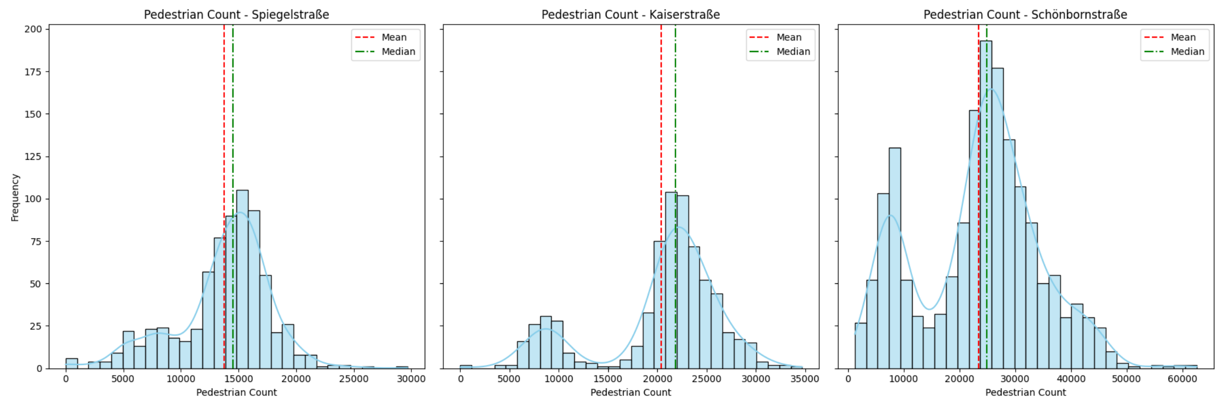


Figure 1: Pedestrian count distributions across different locations

3.4.2 Location Analysis

The pedestrian data is collected from three nearby locations: *Kaiserstraße*, *Spiegelstraße*, and *Schönbornstraße*, as shown in Figure 2.

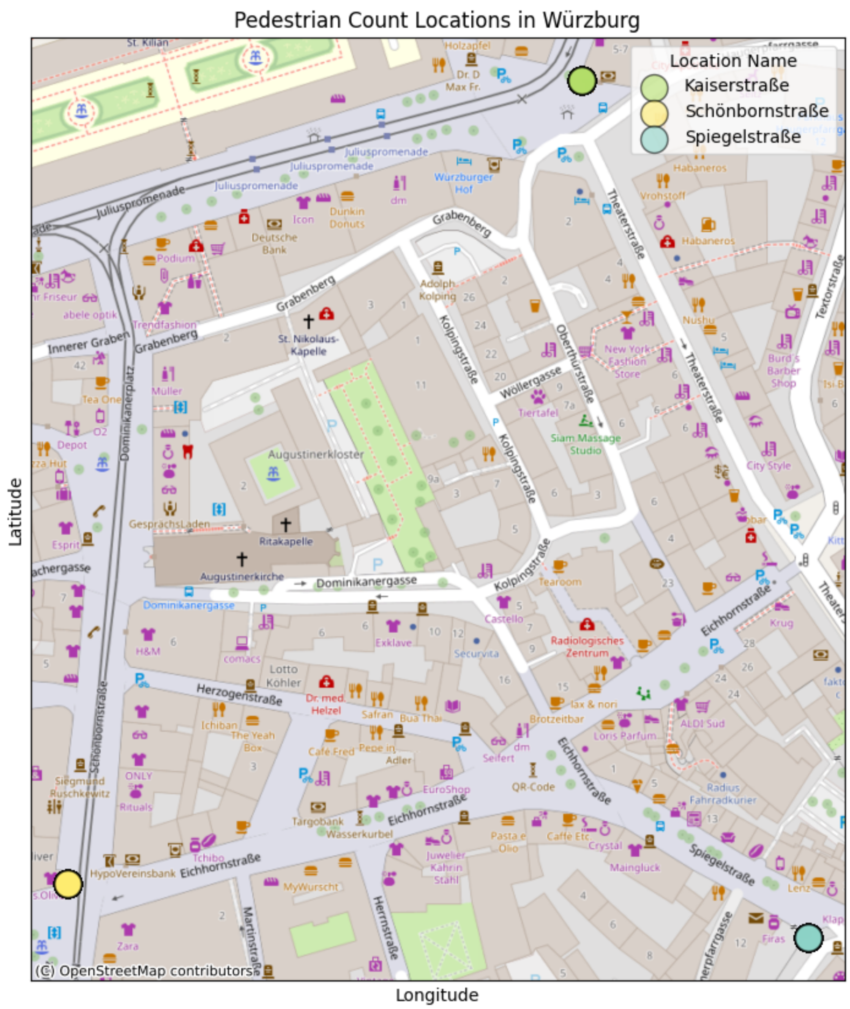


Figure 2: Map of pedestrian count locations

As seen in figure 3, the *Schönbornstraße* records the highest pedestrian count, likely due to its proximity to shopping centers and city center, followed by *Kaiserstraße*, which is near the train station and city centre, and lastly *Spiegelstraße*, which is also located in the city center.

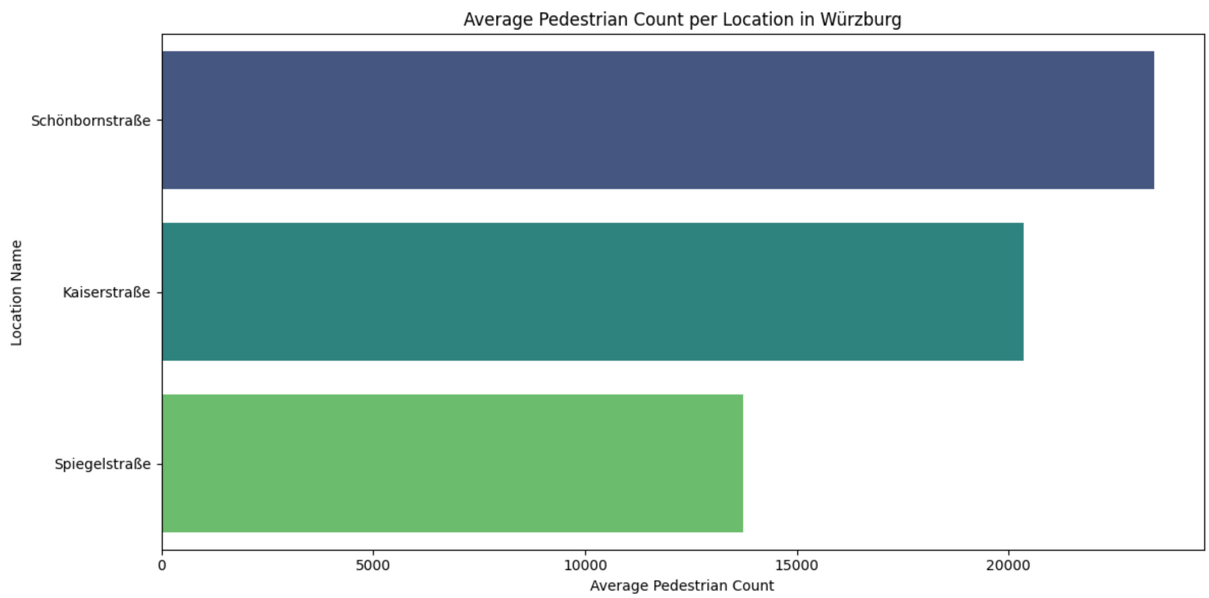


Figure 3: Daily average pedestrian count by location

The difference of peak days and hours in different locations is further analysed in 3.6.3 and 3.6.4

3.5 Relationship with Weather Variables

This section examines the influence of temperature and various weather conditions on pedestrian activity.

3.5.1 Impact of Temperature on Pedestrian Counts

Although the correlation between temperature and pedestrian count is weak (0.09), Figure 4 suggests higher pedestrian activity on warm days (10−30°C). Extreme temperatures, both hot and cold, see lower pedestrian numbers, likely due to discomfort caused by weather conditions.

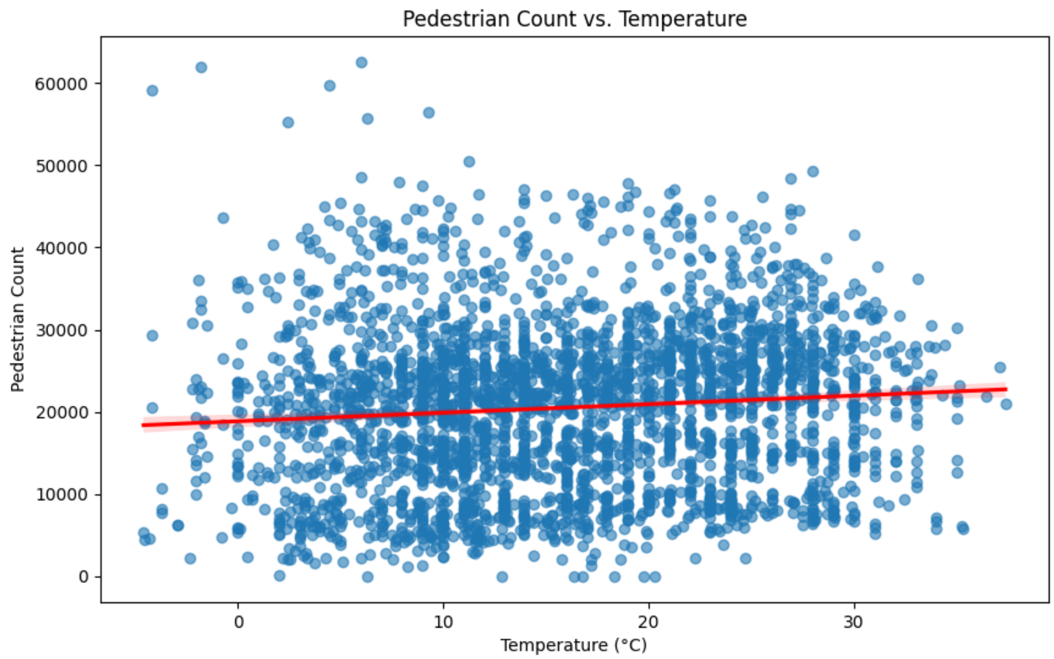


Figure 4: Pedestrian count vs. Temperature

3.5.2 Weather Conditions

As illustrated in Figure 5, there is no significant variation in pedestrian numbers across different weather conditions. However, windy and snowy days show the lowest counts, likely because adverse weather discourages outdoor activity. Due to the minimal variation in pedestrian traffic and the unpredictability of future weather conditions, weather conditions will be excluded from further analysis in the ML models.

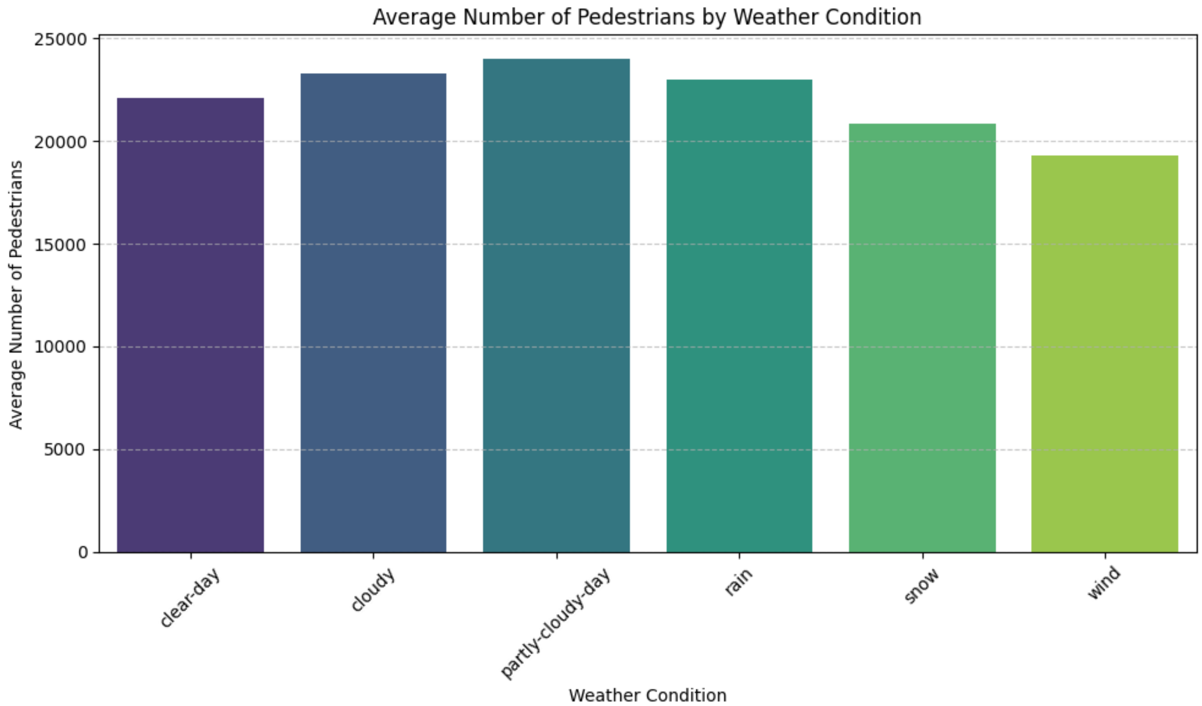


Figure 5: Average number of pedestrians by weather condition

3.6 Time-Based Patterns

This section explores how pedestrian activity varies across different time frames, including years, months, weekdays, and hours.

3.6.1 Yearly Pedestrian Patterns

As shown in figure 6, there is a significant drop in pedestrian traffic during the COVID-19 lockdowns. Due to the abnormal nature of this period, this data will be excluded from further analysis to maintain the dataset's relevance and accuracy.

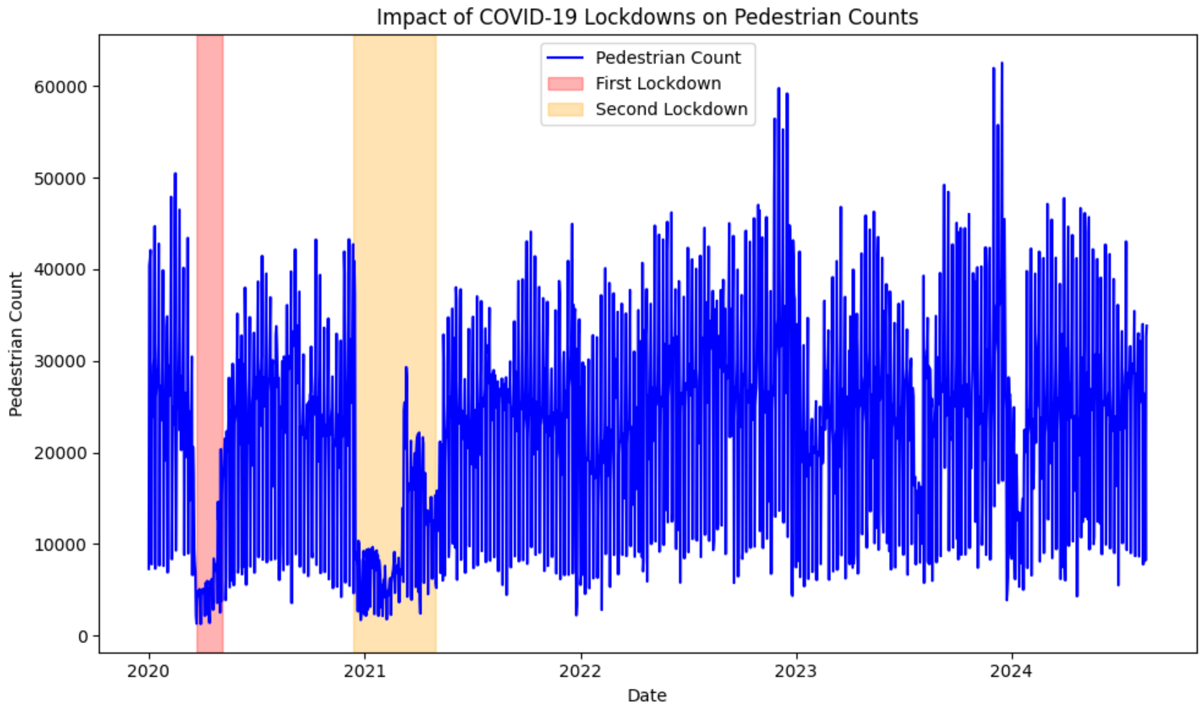


Figure 6: Impact of COVID-19 lockdowns on pedestrian counts

3.6.2 Monthly Pedestrian Trends

As seen in figure 7, pedestrian counts are highest in July, May, June, and December, coinciding with favorable weather, public events and seasonal shopping trends. The lowest counts are recorded in February, January, and September.

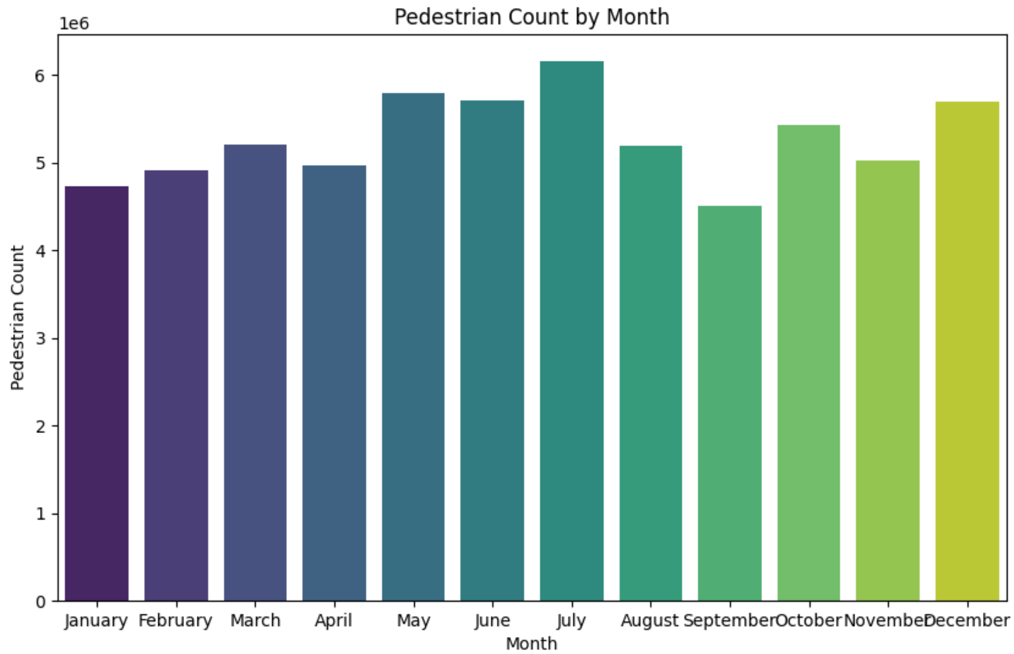


Figure 7: Average number of pedestrians per month

3.6.3 Weekly Patterns

As shown in Figure 8, foot traffic peaks on Saturdays, driven by shopping activities, while Sundays record the lowest pedestrian counts due to shop closures. Fridays also see relatively high numbers.

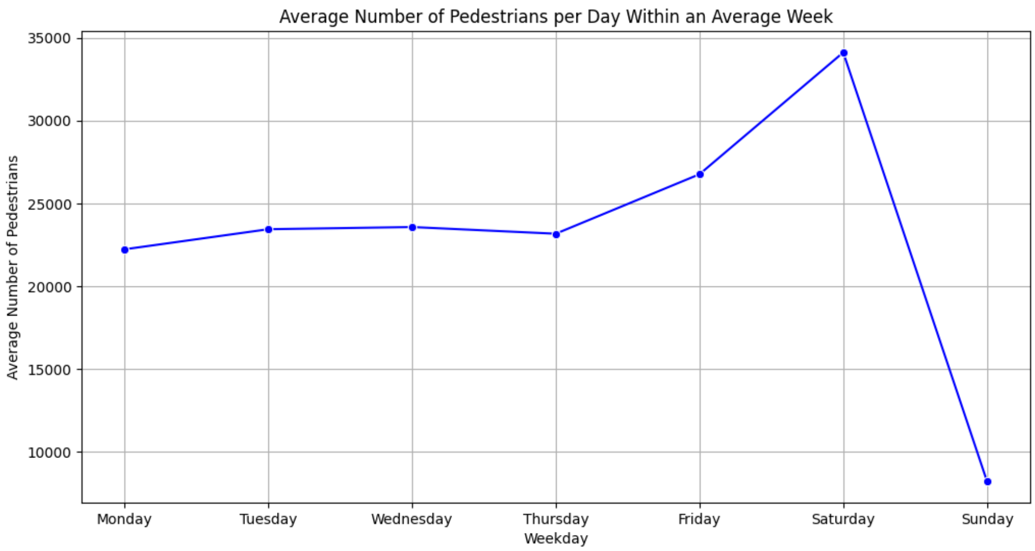


Figure 8: Average number of pedestrians per day within an average week

Table 2 highlights the percentage deviation of pedestrian counts from the weekly average at each location. A positive percentage indicates more foot traffic than the average, while a negative percentage reflects fewer pedestrians.

The data shows that Saturdays consistently experience the highest foot traffic, with *Schönbornstraße* seeing the most significant increase at 49% above the average. Conversely, Sundays show the sharpest decline across all locations, with foot traffic in *Schönbornstraße* dropping by 65%.

During the weekdays, both *Kaiserstraße* and *Spiegelstraße* exhibit relatively stable pedestrian activity. In contrast, *Schönbornstraße* maintains steady counts during the week but sees a notable spike on Saturdays, suggesting its pedestrian traffic is more influenced by weekend shopping and leisure activities.

Overall, the weekday patterns are fairly consistent across locations, but the most substantial variations occur on weekends, particularly in *Schönbornstraße*, which experiences both the largest increases on Saturdays and the steepest drops on Sundays.

Location	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
Kaiserstraße	-1%	4%	3%	6%	16%	27%	-56%
Schönbornstraße	-5%	1%	2%	1%	17%	49%	-65%
Spiegelstraße	-2%	6%	4%	6%	13%	18%	-45%

Table 2: Percentage deviation of pedestrian counts by day of the week and location

3.6.4 Hourly Patterns

Pedestrian traffic patterns in Würzburg show consistent peaks during midday hours, with some variation across the three locations. As shown in Figure 9, *Schönbornstraße* experiences the highest foot traffic at 2:00 PM, averaging 2,543 pedestrians. *Spiegelstraße* sees its peak at 1:00 PM with 1,700 pedestrians, while *Kaiserstraße* reaches its busiest hour at 11:00 AM, averaging 1,975 pedestrians.

The lowest activity is observed in the early morning (12:00 AM to 4:00 AM) across all locations, with pedestrian counts as low as 2 in *Spiegelstraße* at 1:00 AM. Evening hours (after 7:00 PM) also show a sharp decline, with all locations recording significantly lower counts compared to their midday peaks.

Overall, while foot traffic patterns are similar, *Schönbornstraße* consistently sees the highest pedestrian traffic, particularly in the afternoon. In contrast, *Kaiserstraße* experiences earlier peaks, likely due to its proximity to the train station. *Spiegelstraße* has lower overall counts, with a noticeable decline after 11:00 AM, possibly reflecting its location's reduced attraction for midday activities.

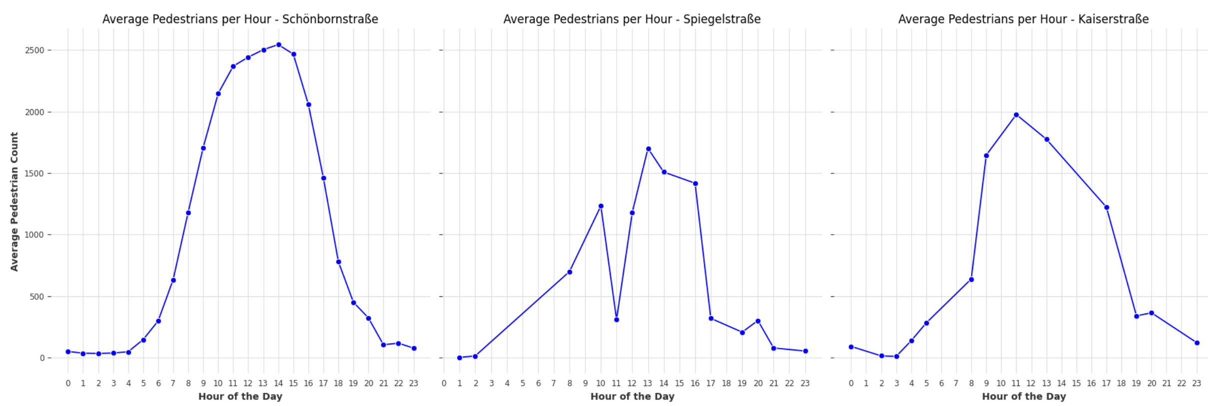


Figure 9: Average number of pedestrians per hour of the day across different locations

4 Feature Engineering

This section outlines the creation of additional features to improve model performance.

4.1 Time-Based Features

Time-based features are designed to capture patterns in pedestrian activity over different temporal intervals, such as special events, holidays, and weekly cycles.

4.1.1 Impact of Special Events on Pedestrian Count

Special events, such as the Christmas market and festivals, are well-known for attracting large crowds in Würzburg. For example, pedestrian traffic peaks on Saturdays during the Christmas market and on Sundays during events like Carnival and "Sunday Shopping."

As shown in Table 3, pedestrian counts generally increase on days with special events, particularly on weekends. Pedestrian numbers are notably higher on Saturdays and Sundays when special events occur. For instance, on Sundays with special events, the average pedestrian count increases from 7,996 to 10,831. Similarly, on Saturdays, foot traffic rises from 29,521 on regular days to 30,423 on days with special events.

However, the data indicates that the number of special events during the workweek (Monday to Wednesday) is relatively low, ranging from only 1 to 3 events per day. This suggests that the observed fluctuations in pedestrian counts on these days may lack statistical significance.

In contrast, Sundays have the highest number of special events (36), and these days consistently show a significant increase in foot traffic, suggesting that special events on Sundays have a strong impact on pedestrian movement. This highlights the importance of Sunday events, which draw more visitors despite the otherwise lower average foot traffic on Sundays.

While special events generally boost pedestrian numbers on weekends and specific weekdays like Saturday and Sunday, the effect is less pronounced or even negative on certain weekdays like Monday and Thursday. For example, Thursdays see a decrease in pedestrian traffic during special events (from 21,826 to 18,257), indicating that weekday events might not attract as many visitors due to the constraints of work schedules.

In summary, special events have a positive impact on pedestrian traffic, particularly on weekends, with the most significant increases observed on Saturdays and Sundays. On weekdays, however, the lower number of events combined with people's limited availability results in a less noticeable or even negative impact on pedestrian activity.

Day of Week	Is Special Event	Pedestrians Count	Number of Special Events
Monday	False	20697.30	2
	True	10975.00	
Tuesday	False	21619.80	1
	True	27053.00	
Wednesday	False	21837.48	3
	True	23037.29	
Thursday	False	21826.77	11
	True	18257.04	
Friday	False	24637.47	16
	True	24387.03	
Saturday	False	29521.59	13
	True	30423.25	
Sunday	False	7996.52	36
	True	10831.37	

Table 3: Pedestrians count based on day of week, special event status, and number of special events

To better capture the influence of these events, a new feature was created to flag special event days, which is detailed further in the Table 8.

4.1.2 Impact of Holidays on Pedestrian Count

As illustrated in Figure 10, pedestrian traffic is generally lower on public holidays in Würzburg. To account for this trend, a one-hot encoded feature, is holiday, was added to the dataset, covering bavarian public holidays.

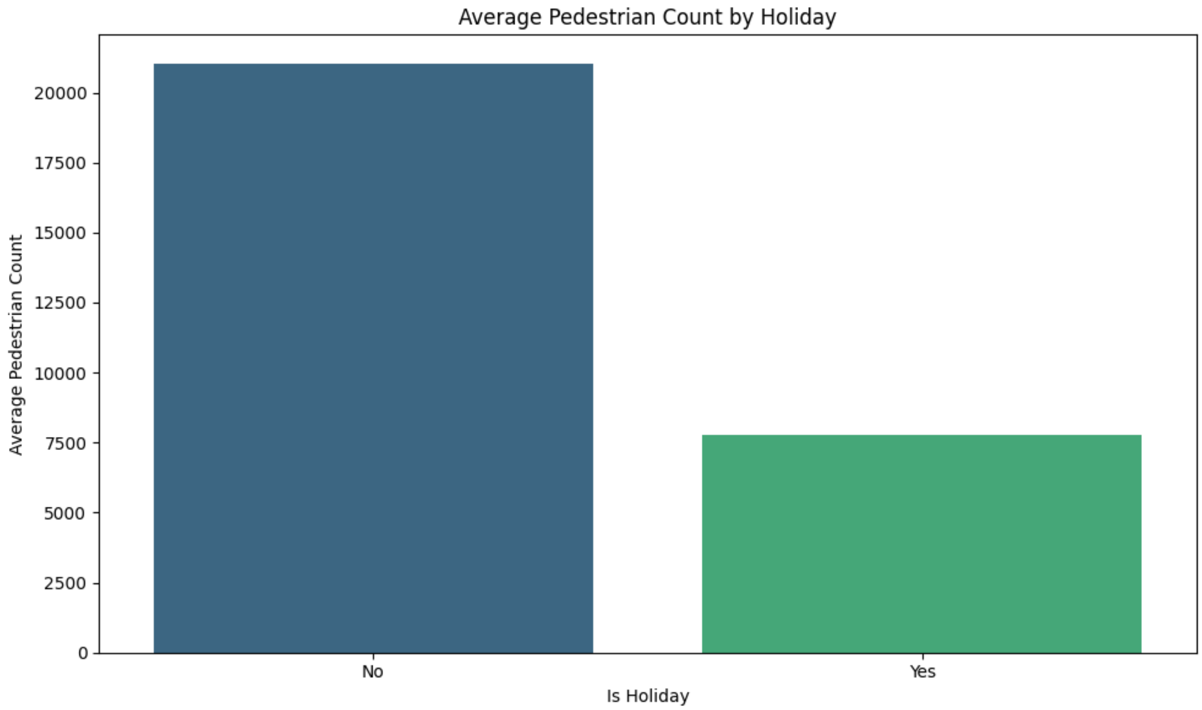


Figure 10: Average pedestrian counts on holidays

4.2 Lag-Feature

A lag feature was introduced to capture the pedestrian count from previous days, helping to model the weekly cyclical patterns in foot traffic, particularly in Würzburg's city center.

Figure 11 illustrates the autocorrelation of the lag feature, showing that the highest correlation occurs with a 7-day lag. This indicates a strong weekly pattern in pedestrian traffic, leading to the introduction of a 7-day lag feature in the model.

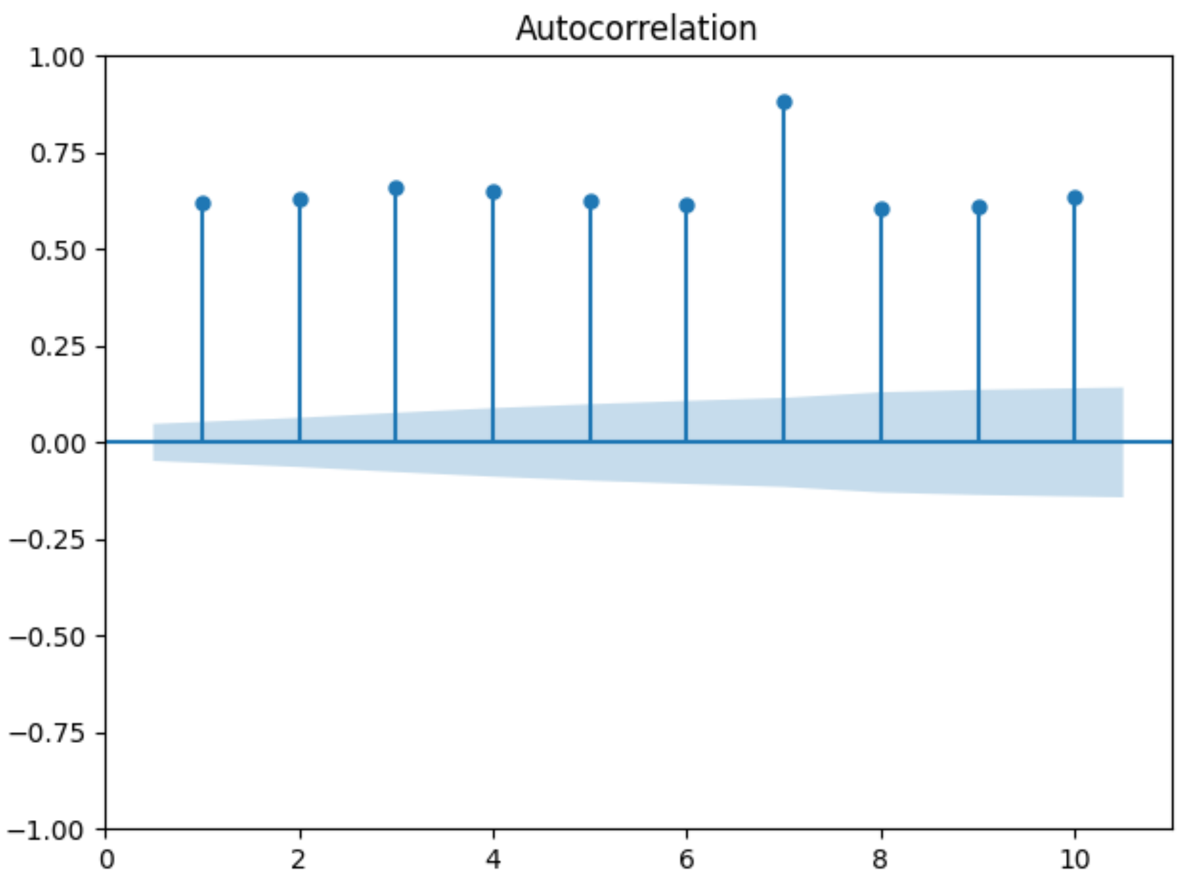


Figure 11: Autocorrelation of lag feature

4.3 Retrieval of Weather Forecast Data

To enhance the dataset, weather forecast data was retrieved using an Application Programming Interface (API) from Zippenfenig (2023). This data includes hourly measurements of temperature, relative humidity, and wind speed. The temperature forecast was integrated into the test set of the ML and DL models to improve prediction accuracy.

5 Model Development

This section outlines the development of forecasting models designed to predict pedestrian traffic, focusing on probabilistic approaches to account for uncertainty and variability in the data.

5.1 Probabilistic Forecast

In pedestrian traffic prediction, variability and uncertainty are inherent due to factors such as weather, special events, and holidays. These factors cause fluctuations in traffic that must be accounted for in any robust forecasting model.

Probabilistic forecasting is ideal for this task as it not only predicts an expected pedestrian count but also provides a distribution of potential outcomes. This method enables businesses to prepare for a range of scenarios, rather than relying on a single-point prediction. For instance, rather than forecasting just one number, probabilistic models provide a spectrum, showing the likelihood of different outcomes, allowing for better planning for both typical traffic days and unexpected surges (Gneiting, 2008).

Incorporating probabilistic models enhances decision-making in key areas such as inventory management and employee shift planning. These operations require flexibility, and probabilistic forecasts allow businesses to prepare for low, medium, and high traffic situations. By using these models, retailers can minimize risks associated with understocking or overstaffing, improving operational efficiency.

In this thesis, models are employed for their predictive accuracy and ability to provide probabilistic estimates via quantile regression. These techniques enable us to capture the range of possible pedestrian counts across various conditions, ensuring both precision and reliability. This probabilistic approach is particularly relevant to the challenges discussed in the research question, such as optimizing inventory and shift planning.

By embracing the uncertainty inherent in pedestrian flow, probabilistic forecasting offers the flexibility to prepare for a range of scenarios, making it a valuable tool for improving operational efficiency in retail environments.

5.2 Pedestrian Prediction Models

To accurately predict pedestrian traffic, various ML models are employed, each utilizing key input features that capture the temporal and contextual factors influencing foot traffic.

5.2.1 Input Features

The following Table 4 outlines the key input features used in the pedestrian prediction models, while temperature forecasts are only available in the test set.

Feature	Description
Timestamp	The date and time of the observation
Temperature	Current temperature during the observation
Pedestrian Count	The recorded pedestrian count at the given time
Location	The specific location in the city
Special Event	Whether a special event is occurring
Holidays	Indicates if the day is a holiday
Day of the Week	The day of the week (Monday-Sunday)
Hour	The hour of the day
Pedestrian Count 7d Lag	Pedestrian count from the same time 7 days earlier
Temperature	temperature for the observation period

Table 4: Input features for pedestrian prediction models

5.2.2 Models

Several models were employed to predict pedestrian counts, each selected for its capacity to handle the complexities of the dataset. These models range from simple baseline approaches to advanced ML techniques, each offering specific advantages in this scenario:

- **Empirical Distribution:** The baseline model generates quantile estimates based on the empirical distribution of pedestrian counts from the training data. This model represents what pedestrian predictions might look like if retailers were to make estimates based on historical patterns alone.
- **Quantile Regression:** The quantile regression model implemented with Pedregosa et al. (2011a) uses important input features such as the day of the week and the location . This model produces linear relationships between the quantile predictions and the input features.
- **LightGBM, CatBoost, and Quantile Random Forest:** These tree-based models are well-suited for capturing complex, non-linear relationships in the data. LightGBM (Ke et al., 2017), CatBoost (Prokhorenkova et al., 2018), and Quantile Random Forest (Johnson, 2024) were selected for their ability to model intricate interactions between features like special events and weekday.
- **LSTM:** The LSTM model, a DL technique, was developed to capture sequential dependencies in time series data but can also be extended with additional features. Implemented using Olivares et al. (2022), LSTMs are particularly effective at modeling long-term dependencies, making them ideal for time-series forecasting where past events significantly impact future pedestrian counts.

5.2.3 Post-Processing Quantile Predictions

Quantile regression models can sometimes generate inconsistencies where higher quantiles yield lower predictions than lower quantiles, violating the monotonicity property of the conditional quantile function.

This problem is often referred to as Quantile Crossing (QC), a term that has its origins in the linear quantile regression. Here the term QC illustrates the phenomenon that two regression lines for two different quantiles cross at some point because both lines have different slopes. This poses significant challenges in linear and nonlinear quantile regression models (Park et al., 2022). While Park et al. (2022) proposed a flexible and efficient distribution-free quantile estimation framework that resolves QC through a neural network layer, in this study, the issue was addressed using a simpler post-processing approach.

Figure 12 illustrates this problem using the inverse Cumulative Distribution Function (CDF) of the LightGBM model, where monotonicity is violated.

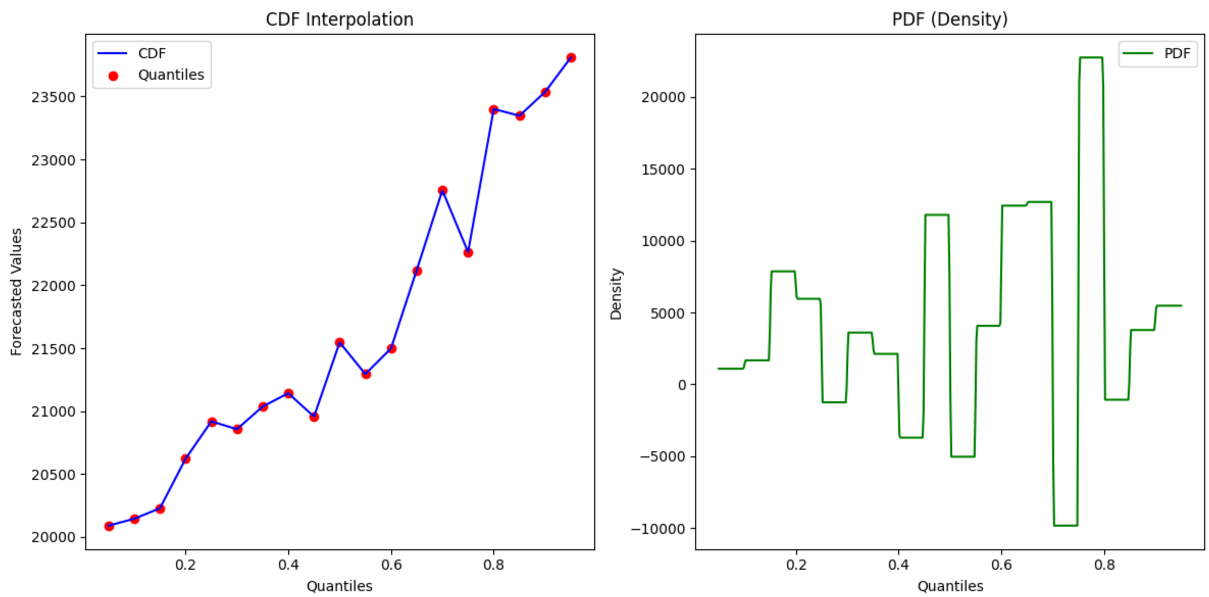


Figure 12: CDF Interpolation and PDF of LightGBM Model

To correct this, a post-processing step was applied to the output of the quantile regression model. This step ensures that the quantile predictions follow a monotonic relationship, meaning that higher quantiles always have greater or equal values compared to lower quantiles. As shown in Figure 13, this post-processing step successfully restores the consistency in the predicted distribution.

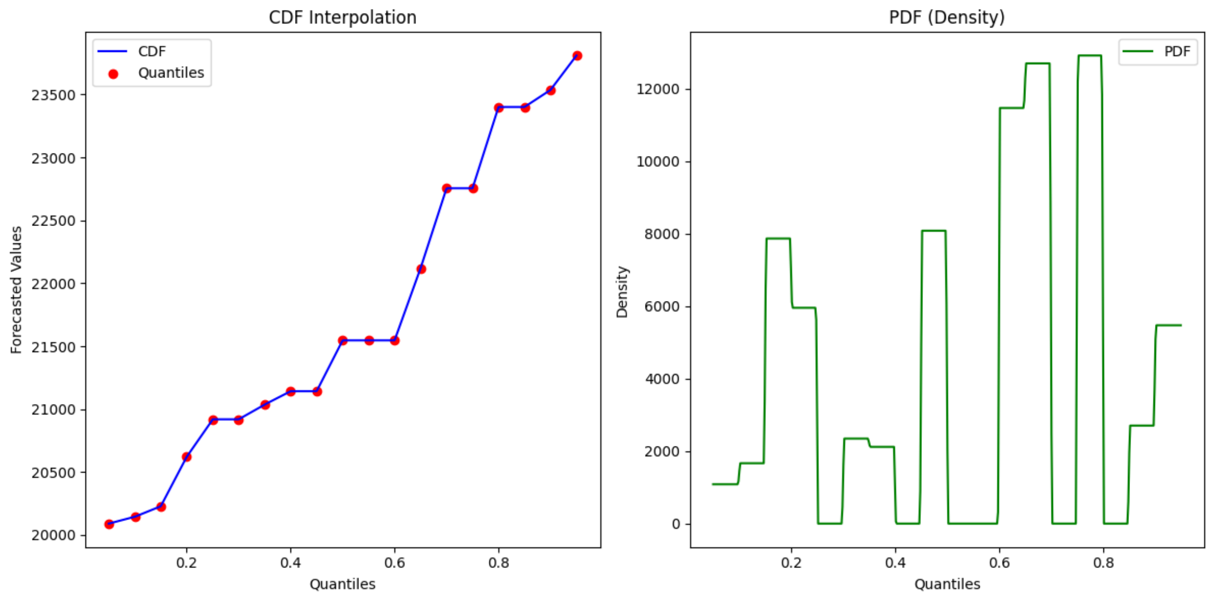


Figure 13: CDF Interpolation and PDF of LightGBM Model after post-processing

6 Model Evaluation

Evaluating the performance of probabilistic pedestrian traffic forecasting models requires a robust methodology to ensure reliable predictions across different time periods and scenarios. This section outlines the methods used to assess the accuracy and effectiveness of the models developed in this thesis

6.1 Train-Test Split

The time series split approach in this study is designed to evaluate model performance over both short-term and long-term forecasting horizons. Two different configurations were employed:

- **Short-term splits**: A total of 9 train-test splits were created, each covering a 2-day forecast period. These splits were spaced at least 2 days apart to ensure more frequent testing and provide a higher resolution for evaluating the model's ability to make short-term predictions. This setup allows for a more granular assessment of the model's performance in predicting immediate future pedestrian traffic patterns.
- **Long-term splits**: In contrast, 3 train-test splits were generated for a 7-day forecast period, with split dates spaced at least 7 days apart. These splits were designed to assess the model's ability to forecast over a longer horizon, offering insights into its performance in capturing broader, week-long pedestrian traffic trends.

However, the number of splits was limited by the availability of weather forecast data and gaps in the city's pedestrian traffic recordings. Only the latest weather forecasts could be retrieved from the freely available API we used. It was not possible to retrieve historical forecasts. This constraint prevented the creation of more splits, as the models require both weather predictions and consistent traffic data to ensure robust testing and reliable results. In each split, weather forecast data, including both daily and hourly temperature information, was integrated to account for real-world variables that influence pedestrian movement. This setup allowed the model to be tested under conditions that mimic future uncertainty, enhancing the realism of the predictions. By incorporating this weather data, the model's adaptability to changing environmental factors could be better assessed, offering a more thorough evaluation of its practical application in scenarios where external conditions, such as temperature fluctuations, play a significant role in pedestrian behavior.

6.2 Evaluation Metrics

The performance of the predictive models is assessed using a combination of probabilistic and point-estimate evaluation metrics. These metrics provide a comprehensive understanding of the model's ability to forecast pedestrian counts accurately across different quantiles, as well as to capture overall prediction errors.

6.2.1 Pinball Loss

To measure how well a quantile regression estimates a specific quantile q of the conditional distribution, the pinball loss is a useful tool. It is calculated as:

$$L_q(y_i - \hat{y}_i) = \begin{cases} q \cdot (y_i - \hat{y}_i), & \text{if } y_i > \hat{y}_i \\ (1-q) \cdot (\hat{y}_i - y_i), & \text{otherwise.} \end{cases}$$

where:

- y_i is the true value,
- $\hat{y}_i$ is the predicted value,
- q is the quantile level (e.g., 0.5 for the median),

Unlike traditional metrics like Mean Squared Error (MSE), the pinball loss penalizes overestimations and underestimations differently, which fits the needs of quantile regression.

6.2.2 Multi-Quantile Loss

To measure the accuracy of probabilistic forecasts, the Multi Quantile Loss (MQL) is used. While the pinball loss focuses on single point estimates, the MQL evaluates the model's performance across multiple quantiles, making it ideal for probabilistic forecasting (Pedregosa et al., 2011b). This allows to evaluate a range of outcomes, covering both common and more extreme pedestrian counts (Lewinson, 2024).

For our range of $m = 19$ quantiles $\{0.05, 0.1, 0.15, \ldots 0.9, 0.95\}$ the MQL is given by:

$$\text{MQL} = \frac{1}{n \cdot m} \sum_{i=1}^{n} \sum_{k=1}^{m} L_{q_k}(y_i, \hat{y}_i),$$

where

- y_i is the true pedestrian count,
- $\hat{y}_i$ is the predicted pedestrian count,
- $q_k = 0.05 \cdot k$ is the k-th of our 19 quantiles,
- $m = 19$ is the number of quantiles
- n is the number of observations.

where $q_k = 0.05 \cdot k$ is the k-th of our 19 quantiles.

The MQL averages the prediction accuracy calculated with the pinball loss across different quantiles and thus provides a broader overview of how well the model captures the uncertainty and variability in pedestrian traffic.

6.2.3 Mean Absolute Error

The Mean Absolute Error (MAE) is a standard metric that measures the average magnitude of errors between predicted and actual pedestrian counts. It is particularly intuitive as it is expressed in the same units as the original data, making it easy to interpret.

The formula for MAE is:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

where:

- y_i is the true pedestrian count,
- $\hat{y}_i$ is the predicted pedestrian count,
- n is the number of observations.

The MAE is a special case of the pinball loss when $q = 0.5$. At the 50th percentile (median), the pinball loss treats over-predictions and under-predictions equally, making it the mean pinball loss equivalent to the MAE. This means that the pinball loss with $q = 0.5$ corresponds exactly to the MAE, as both evaluate the average absolute error.

6.3 Results and Analysis

In this section, the performance of various models on forecasting pedestrian traffic for inventory management and employee shift planning is analyzed.

6.3.1 Inventory Management

When weather forecast was incorporated (as shown in Table 9), most models experienced a decline in performance. The primary reason for this is that the models were trained using actual historical temperature data, while the test set used forecasted weather data, which is inherently less accurate. (We were not able to train the model on historical weather forecasts, but only on the actual historical weather, as only the latest weather forecasts can be retrieved via the API. The historical weather forecasts for the training data simply could not be retrieved.)

In real-world applications, retailers rely on forecasted rather than historical weather data, but this uncertainty in weather forecasts reduced model reliability for pedestrian traffic predictions, highlighting a limitation when using forecasted weather variables.

Table 5 presents the results of different models for predicting pedestrian traffic over two-day and seven-day horizons, excluding weather data.

CatBoost and **LightGBM** emerged as the best-performing models for both the two-day and seven-day horizons, with MQL improvements of 65.91% and 65.67% in the two-day horizon and MAE improvements of around 68%. These models excel at capturing non-linear relationships and interactions between categorical variables, making them highly effective for predicting pedestrian traffic influenced by factors such as holidays, daily patterns, and special events.

Random Forest, while slightly behind CatBoost and LightGBM, demonstrated stable performance due to its robustness in handling varied datasets, though it did not reach the optimization level of gradient-boosting methods.

LSTM also performed well but was outperformed by tree-based models. This is likely due to the limited dataset size, as deep learning models like LSTM typically benefit from larger datasets to fully capture complex temporal dependencies and provide optimal results.

Quantile Regression struggled significantly, particularly in the seven-day horizon, showing a negative MQL improvement of -13.94%, which indicates limitations in capturing long-term variability in pedestrian traffic, making it less reliable for extended forecast periods.

While the MQL metric identifies CatBoost as the best model, if we consider MAE as a priority, LightGBM is the top performer, followed by CatBoost, LSTM, RF, and Quantile Regression, with Empirical Quantiles performing the worst. Thus, model rankings may vary depending on whether predictive accuracy in MQL or MAE is prioritized.

The MAE values, ranging from 2594 to 3252, represent a relatively small percentage of the average daily pedestrian count, approximately 20,525. For instance, an MAE of 2677 equates to an average error of around 13%, which indicates a reasonable level of accuracy. However, for practical inventory management, this margin of error could still lead to overstocking or stockouts, emphasizing the need for continued model refinement.

Method	MQL	MAE	Improvement MQL (%)	Improvement MAE (%)
Two Days Horizon				
CatBoost	1079.99	2700.78	65.91	68.90
LightGBM	1087.38	2677.33	65.67	69.17
RandomForest	1094.39	2771.80	65.45	68.08
LSTM	1216.66	2749.80	61.59	68.34
Quantile Regression	1910.13	3252.88	39.70	62.54
Empirical Quantiles	3167.88	8684.56	0.00	0.00
Seven Days Horizon				
CatBoost	1043.20	2667.56	65.81	67.95
LightGBM	1044.12	2594.40	65.78	68.83
RandomForest	1072.00	2709.75	64.87	67.44
LSTM	1206.81	2706.46	60.45	67.48
Quantile Regression	3476.67	3135.56	-13.94	62.33
Empirical Quantiles	3051.34	8323.53	0.00	0.00

Table 5: Evaluation of different methods of daily data with two and seven day horizons

T6 compares the MAE, CV and the relative error (MAE/mean) for the LightGBM predictions for the 50% quantile at the different locations. In both the two-day and seven-day horizons, **Kaiserstraße** consistently exhibited the lowest MAE, and relative error, indicating the model's high predictive accuracy and stability for this location. This suggests that foot traffic in Kaiserstraße is easier to forecast accurately.

In contrast, **Schönbornstraße** had the highest MAE, CV, and relative error across both horizons, indicating greater variability and a higher proportion of prediction error relative to the mean foot traffic in this area. The elevated relative error at Schönbornstraße highlights the challenge in accurately forecasting foot traffic in more dynamic or variable locations, where fluctuations are harder for the model to capture precisely.

Location	MAE	Mean	CV	Relative Error (MAE/Mean)
2-Day Horizon				
Schönbornstraße	4091.41	23444	0.47	0.17
Kaiserstraße	1794.92	20362	0.30	0.09
Spiegelstraße	2189.46	13728	0.30	0.16
7-Day Horizon				
Schönbornstraße	3830.97	23444	0.47	0.16
Kaiserstraße	1754.25	20362	0.30	0.09
Spiegelstraße	2204.27	13728	0.30	0.16

Table 6: MAE, Mean, CV, and Relative Error (MAE/Mean) of LightGBM model for different locations across 2-day and 7-day horizons

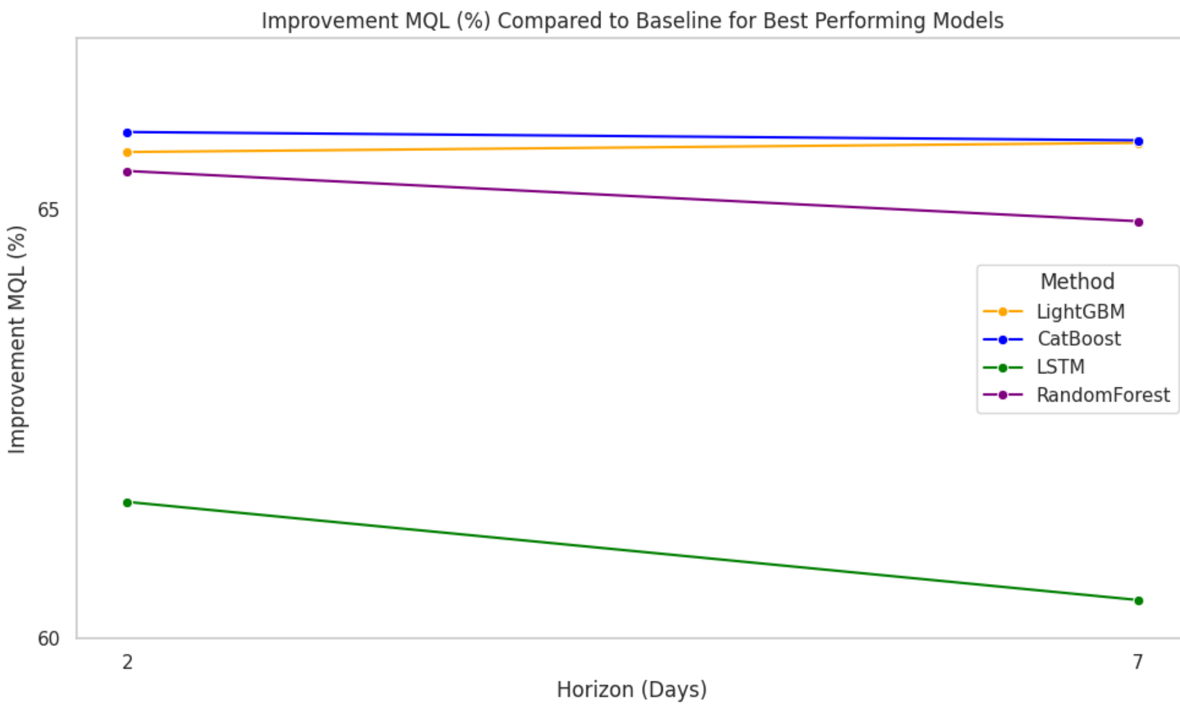


Figure 14: Average performance improvement in MQL of best performing models of daily data

In summary, both **CatBoost** and **LightGBM** consistently demonstrated the highest effectiveness for pedestrian traffic forecasting across two-day and seven-day horizons, underscoring their suitability for inventory management applications where accurate short-term predictions are essential. The results in Figure 14 further reveal that advanced machine learning models outperformed the baseline methods, achieving MQL improvements of 60–65%. This improvement enables retailers to adjust stock levels more precisely based on reliable foot traffic forecasts, minimizing the risks of overstocking or stockouts.

6.3.2 Employee Shift Planning

As observed in the inventory management analysis (6.3.1), including weather data led to a slight decline in model performance, which was also evident in the hourly forecast (see Table 10). This decline is likely due to the inherent unreliability of weather forecasts when compared to historical data used during training. Similar to the approach used in the inventory management section (6.3.1), weather data was included at hourly intervals.

Table 7 summarizes the performance of various models in predicting pedestrian traffic over 48-hour and 168-hour forecast horizons, excluding weather data.

CatBoost consistently outperformed other models for both the 48-hour and 168-hour horizons, achieving MQL improvements of 75.37% and 76.26%, with MAE improvements around

77%. These results highlight CatBoost's capability to effectively capture complex patterns in hourly pedestrian traffic, making it highly valuable for both short- and long-term shift planning. **LightGBM**, though slightly behind CatBoost, also demonstrated strong results.

LSTM, while slightly behind LightGBM in MQL, showed particularly strong performance in MAE, especially in the 48-hour forecast, where it achieved a 75 76.93% improvement and was the best model. This suggests that LSTM is well-suited for scenarios with larger datasets.

Random Forest, while stable, lagged behind due to its limited ability to capture complex temporal dependencies. **Quantile Regression**, in particular, struggled with long-term predictions, showing only a 2.28% MQL improvement in the 168-hour horizon, underscoring its limitations for extended-period forecasting.

Considering the average hourly pedestrian count is approximately 797.78, an MAE of 163.07 translates to an error of about 20.44%. This indicates that these models can provide relatively accurate forecasts, with error rates just above 20% of the hourly average.

Method	**MQL**	**MAE**	**Improvement MQL (%)**	**Improvement MAE (%)**
48 Hours Horizon				
CatBoost	67.02	167.96	75.37	76.90
LightGBM	74.55	188.93	72.61	74.01
LSTM	74.75	167.71	72.53	76.93
RandomForest	82.57	198.72	69.66	72.66
Quantile Regression	233.83	232.31	14.08	68.04
Empirical Quantiles	272.15	726.97	0.00	0.00
168 Hours Horizon				
CatBoost	64.49	163.07	76.26	77.68
LightGBM	71.21	181.53	73.79	75.15
LSTM	78.02	178.30	71.28	75.59
RandomForest	79.43	195.43	70.76	73.25
Quantile Regression	265.48	218.70	2.28	70.06
Empirical Quantiles	271.68	730.58	0.00	0.00

Table 7: Evaluation of different methods for 48-hour and 168-hour horizons

Although the ranking in Figure 15 is based on MQL, if MAE were prioritized, the order would shift to CatBoost, LSTM, LightGBM, RF, Quantile Regression, and finally, the Empirical Quantiles.

Overall, the substantial improvements in MQL and MAE across all models compared to baseline methods emphasize their potential for optimizing staffing costs and ensuring efficient workforce allocation

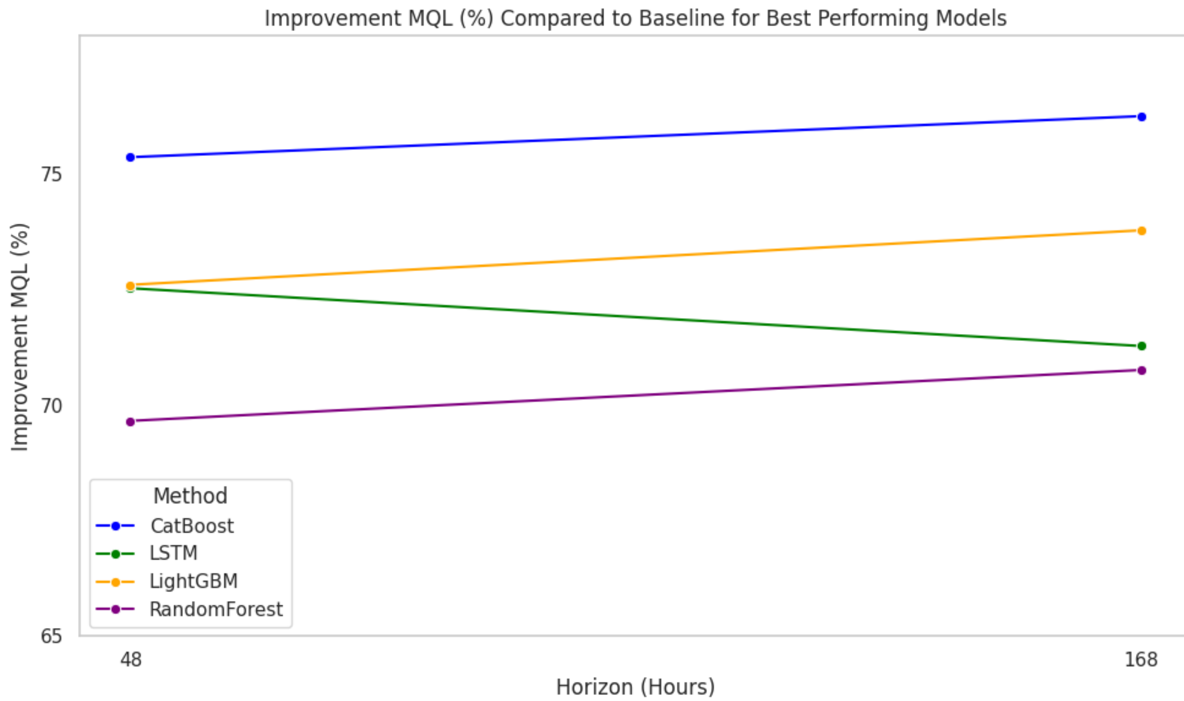
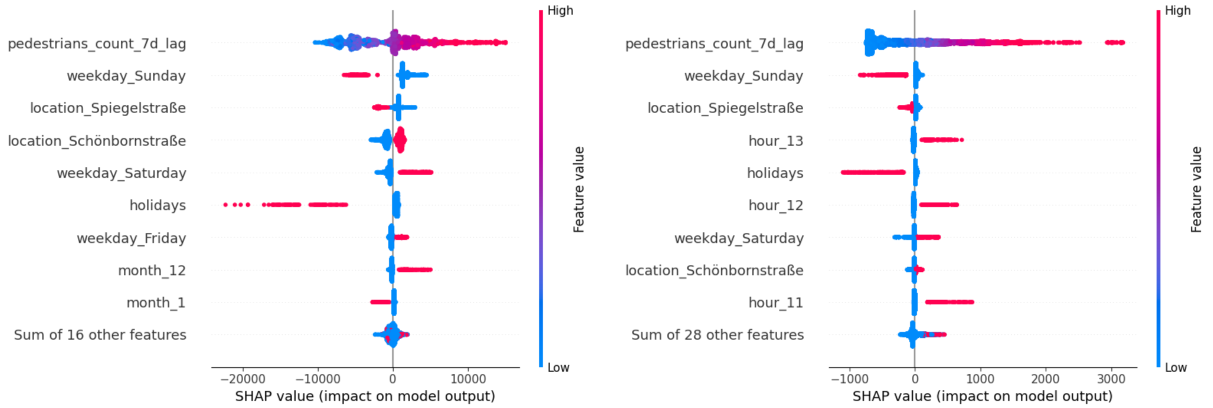


Figure 15: Average performance improvement in MQL of best performing models of hourly data

7 Identifying Key Drivers

Understanding the key factors influencing pedestrian traffic is crucial for enhancing model accuracy and making informed decisions in retail operations. The following summarizes the most significant variables identified from both the daily and hourly datasets.



(a) Feature importance of daily dataset

(b) Feature importance of hourly dataset

Figure 16: Feature importance of LightGBM model of the hourly and daily datasets

The feature importance analysis for the daily and hourly datasets, as shown in Figure 16, highlights the following key drivers:

- **7-Day Pedestrian Lag**: The most influential feature across both datasets, indicating that pedestrian traffic follows consistent weekly patterns. Retailers can leverage this to predict future demand by analyzing traffic from the previous week.
- **Day of the Week (Sunday and Saturday)**: Both days are significant. Sundays consistently show lower pedestrian traffic due to store closures, while Saturdays experience peaks due to increased shopping activity. These patterns are essential for adjusting inventory and staffing levels.
- **Location (Spiegelstraße and Schönbornstraße)**: Location-specific factors strongly influence foot traffic. High-traffic areas such as Spiegelstraße and Schönbornstraße see significantly higher pedestrian counts, underlining the importance of incorporating location data into forecasting models.
- **Hour of the Day**: In the hourly dataset, the time of day plays a critical role, with peak pedestrian traffic observed at 11:00 AM, 12:00 PM, and 1:00 PM. This information is essential for optimizing employee shift planning and ensuring adequate staffing during busy hours.
- **Holidays and Months**: Holidays and seasonal shopping trends (e.g., December) significantly impact pedestrian traffic. Public holidays tend to reduce foot traffic, while December sees a notable increase due to holiday shopping, and January generally experiences lower traffic. These temporal factors must be considered in retail operations for both staffing and inventory management.

8 Conclusion

In this study, various factors influencing pedestrian traffic were analysed and the performance of different machine learning models for forecasting were assessed, particularly in the context of inventory management and employee shift planning.

9 Summary of Findings and Implications for Retail Operations

The following key insights were derived from this thesis, providing important implications for retail operations:

1. **Temporal Factors:** Pedestrian activity is heavily influenced by time-based factors, with Saturdays consistently showing the highest traffic and Sundays the lowest across all locations. This suggests that retailers should adjust their operational strategies, such as inventory levels and staffing, to account for predictable variations in customer flow, especially ensuring adequate resources during weekends.
2. **Weather Influence:** Although extreme temperatures (both very high and very low) tend to reduce foot traffic, overall weather conditions have a relatively minor impact compared to other factors such as time of day and seasonality. Retailers can therefore rely more on temporal and seasonal trends for planning. However, they should remain adaptable during extreme weather events, as these can temporarily disrupt usual pedestrian patterns. Additionally, the inherent inaccuracy of weather forecasts means that incorporating weather predictions into machine learning models offers limited benefit for reliably enhancing forecast accuracy.
3. **Seasonal Patterns:** Pedestrian traffic follows clear seasonal trends, with higher activity in the months of July, June, May and December. Retailers should anticipate higher demand during these peak months and adjust inventory management and promotional strategies accordingly, ensuring they are prepared for both short-term and long-term fluctuations in customer traffic.
4. **Hourly Activity Trends:** The highest pedestrian traffic occurs between 11:00 AM and 1:00 PM, with early mornings and late evenings seeing the lowest activity. Retailers should align employee shift schedules with these peaks in pedestrian traffic, ensuring adequate staff availability during busy hours to improve customer service and reduce labor costs during quieter times.
5. **Model Performance:** CatBoost outperformed all other models in both single-day and multi-day forecasting, and most models showed substantial improvements over the baseline. The use of advanced machine learning models like CatBoost can significantly enhance forecasting accuracy, providing retailers with reliable tools for demand prediction and operational planning.
6. **Key Predictive Drivers:** The most important factors driving pedestrian traffic predictions include:
   - 7-day lag of pedestrian counts

- Day of the week, particularly the high traffic on Saturdays and low traffic on Sundays
- Location-specific factors
- Hour of the Day

Retailers should focus on these key predictive drivers when developing forecasting models, as they are essential for accurate predictions. Understanding the patterns of customer traffic based on these factors can help optimize inventory management, staffing, and marketing strategies.

These findings emphasize the importance of leveraging advanced machine learning models to improve retail operations. By integrating these insights into decision-making processes, retailers can better align their resources with actual pedestrian traffic patterns, resulting in improved operational efficiency and customer satisfaction.

9.1 Discussion

The results of this thesis align with broader trends in time series forecasting, where advanced ML models like CatBoost and LightGBM have demonstrated strong performance in handling complex datasets. In the M5 Uncertainty Competition, LightGBM was a top-performing model (Makridakis et al., 2022), and this thesis found similar results with CatBoost and LightGBM excelling in pedestrian traffic forecasting for inventory management. These tree-based models effectively capture non-linear patterns and interactions within the data, which is crucial for improving decision-making in areas like inventory management and employee shift planning.

However, unlike the dominance of tree-based models seen in some studies, this thesis found that LSTM was almost as effective as CatBoost and LightGBM in the hourly forecasts. This suggests that LSTM, with its ability to capture sequential dependencies in time series data, can be highly competitive. This finding is consistent with results from the M5 Uncertainty Competition, where many of the top-performing methods also utilized LSTM architectures (Makridakis et al., 2022). Also Harrou et al. (2022) demonstrated that hybrid models, such as those combining LSTM with Variational Autoencoders (VAE), can further improve prediction accuracy, particularly for traffic flow forecasting. This suggests that ensemble methods or hybrid models, which combine the strengths of deep learning and tree-based approaches, could further enhance predictive accuracy, especially in long-term forecasting tasks where sequential dependencies are more prominent.

Random Forest also played a role in forecasting tasks, as shown in the work of Schmidt and Pibernik (2023), where it was effective for time series forecasting. However, Random Forest lacks the boosting mechanisms that give CatBoost and LightGBM their edge in modeling non-linear relationships and complex interactions. The results from this thesis suggest that these advanced models would likely outperform Random Forest in time series forecasting.

The study from Cohen and Dalyot (2020) provides a different perspective by exploring how ML can support navigation systems for blind pedestrians. In this context, RF performed well, achieving 95% accuracy in predicting pedestrian traffic flow levels using spatial features like street geometry and points of interest. This highlights the utility of RF in applications where detailed temporal data might be limited, particularly in open mapping platforms like OpenStreetMap. However, it is likely that more advanced models like CatBoost or LightGBM could offer even greater accuracy in similar contexts by handling more complex data interactions. This could be explored in future research.

In summary, this thesis demonstrates that tree-based models like CatBoost and LightGBM perform exceptionally well in pedestrian traffic forecasting, but LSTM remains highly competitive, particularly for larger datasets like hourly forecasts. Studies like Cohen and Dalyot (2020) suggest that simpler algorithms like RF can still excel in specific contexts, especially for short-term predictions or spatially rich datasets.

9.2 Limitations and Future Work

While this thesis achieved significant advancements in pedestrian traffic forecasting, several limitations must be addressed to further improve model accuracy and applicability.

- **Limited Dataset:** The dataset used in this thesis was collected in Würzburg, which may limit the generalizability of the findings to other cities or regions. Future work could extend this research by collecting and analyzing pedestrian traffic data from multiple locations with different demographic and socioeconomic characteristics to enhance the robustness of the predictions.
- **Ensemble Methods:** The thesis focused on individual models like CatBoost, LightGBM, and LSTM, but ensemble methods were not explored. Combining models, such as an ensemble of LightGBM and LSTM, could potentially improve prediction accuracy by leveraging the strengths of each approach. Future research could experiment with ensemble techniques to explore their impact on pedestrian traffic forecasting.
- **Transformers and Foundational Models:** The use of transformers, which have shown strong performance in capturing sequential dependencies, could be explored for future pedestrian traffic forecasting. As foundational models become increasingly accessible, future research may benefit from assessing their suitability for this type of deep learning application.
- **Prediction Accuracy for Inventory and Shift Planning:** Despite improvements in MAE, prediction errors still hover around 13% for daily forecasts (used in inventory management) and 20% for hourly forecasts (used in employee shift planning). Further work could focus on reducing these errors to minimize the risks of overstocking, stockouts, or understaffing in retail settings.
- **Incorporating Retail Demand Data:** Including retailer-specific demand data could enable future studies to measure the direct impact of pedestrian traffic on actual retail demand, providing insights into how pedestrian volume translates into store visits and purchases. This would facilitate a more comprehensive analysis of the influence of foot traffic on retail operations.
- **Prescriptive Analytics with SAA:** This thesis primarily conducted predictive analysis.

Future research could benefit from applying prescriptive analytics, such as SAA, to translate predictions into actionable recommendations for inventory and workforce management.

This thesis demonstrates that advanced machine learning models, particularly CatBoost and LightGBM, serve as reliable tools for forecasting pedestrian traffic—an essential factor in optimizing inventory management and employee shift planning in retail environments. By leveraging predictive models, retailers can make data-driven decisions that align resources with actual customer traffic patterns, leading to more efficient operations and enhanced customer satisfaction.

To further enhance model performance, future research should consider utilizing larger datasets, foundational models, and prescriptive analytics.

References

Abrishami, Soheila, Piyush Kumar, and Wickus Nienaber (2017). "Smart Stores: A scalable foot traffic collection and prediction system". In: *Advances in Data Mining. Applications and Theoretical Aspects*, pp. 107–121.

Cohen, Achituv and Sagi Dalyot (2020). "Machine-learning prediction models for pedestrian traffic flow levels: towards optimizing walking routes for blind pedestrians". In: *Transactions in GIS* 24.5, pp. 1264–1279.

Fan, Zhuangyuan and Becky P.Y. Loo (2021). "Street life and pedestrian activities in smart cities: opportunities and challenges for computational urban science". In: *Computational Urban Science* 1.26.

Gneiting, Tilmann (2008). "Editorial: Probabilistic Forecasting". In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 171.2, pp. 319–321.

Harrou, Fouzi, Aissa Dairi, Amine Zeroual, and Ying Sun (2022). "Forecasting of Bicycle and Pedestrian Traffic Using Flexible and Efficient Hybrid Deep Learning Approach". In: *Applied Sciences* 12.9, p. 4482.

Johnson, Reid A. (2024). "quantile-forest: A Python Package for Quantile Regression Forests". In: *Journal of Open Source Software* 9.93, p. 5976. DOI: 10.21105/joss.05976. URL: https://doi.org/10.21105/joss.05976.

Ke, Guolin, Qiwei Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qi Ye, and Tie-Yan Liu (2017). "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett. Vol. 30. Curran Associates, Inc., pp. 3146–3154. URL: https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

Lewinson, Towards Data Science Eryk (Aug. 4, 2024). *An Introduction to Quantile Loss, a.k.a. the Pinball Loss*. URL: https://towardsdatascience.com/an-introduction-to-quantile-loss-a-k-a-the-pinball-loss-33cccac378a9 (visited on 10/02/2024).

Makridakis, Spyros, Evangelos Spiliotis, Vassilios Assimakopoulos, Zhi Chen, Anil Gaba, Ilia Tsetlin, and Robert L. Winkler (2022). "The M5 uncertainty competition: Results, findings and conclusions". In: *International Journal of Forecasting* 38.4, pp. 1365–1385.

Notz, Pascal M., Peter K. Wolf, and Richard Pibernik (2023). "Prescriptive analytics for a multi-shift staffing problem". In: *European Journal of Operational Research* 305.2, pp. 887–901.

Olivares, Kin G., Cristian Challú, Federico Garza, Max Mergenthaler Canseco, and Artur Dubrawski (2022). *NeuralForecast: User friendly state-of-the-art neural forecasting models*. PyCon Salt Lake City, Utah, US 2022. URL: https://github.com/Nixtla/neuralforecast.

Park, Youngsuk, Danielle Maddix, François-Xavier Aubet, Kelvin Kan, Jan Gasthaus, and Yuyang Wang (2022). "Learning Quantile Functions without Quantile Crossing for Distribution-free Time Series Forecasting". In: *Proceedings of The 25th International Conference on Arti-*

ficial Intelligence and Statistics. Ed. by Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel Valera. Vol. 151. Proceedings of Machine Learning Research. PMLR, pp. 8127–8150.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011a). *Scikit-learn: Machine Learning in Python*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.QuantileRegressor.html.

— (2011b). *Scikit-learn: Machine Learning in Python*. Accessed: 2024-10-02. URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_pinball_loss.html.

Prokhorenkova, Liudmila, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin (2018). "CatBoost: unbiased boosting with categorical features". In: *arXiv preprint arXiv:1706.09516*. URL: https://arxiv.org/abs/1706.09516.

Rady, EL Houssainy A., Haitham Fawzy, and Amal Mohamed Abdel Fattah (2021). "Time Series Forecasting Using Tree Based Methods". In: *Journal of Statistics Applications & Probability* 10.1, p. 121.

Schmidt, Felix and Richard Pibernik (2023). "Data-Driven Inventory Control For Large Product Portfolios: A Practical Application of Prescriptive Analytics". In: *SSRN Electronic Journal*. URL: https://ssrn.com/abstract=4629964.

Würzburg, Open Data (Sept. 29, 2024a). *Passantenzählung - Stundendaten*. URL: https://opendata.wuerzburg.de/explore/dataset/passantenzaehlung_stundendaten/ (visited on 09/29/2024).

— (Sept. 29, 2024b). *Passantenzählung - Tagesdaten*. URL: https://opendata.wuerzburg.de/explore/dataset/passantenzaehlung_tagesdaten/ (visited on 09/29/2024).

Zhao, Zheng, Weihai Chen, Xingming Wu, Peter C. Y. Chen, and Jingmeng Liu (2017). "LSTM network: a deep learning approach for short-term traffic forecast". In: *IET Intelligent Transport Systems* 11.2, pp. 68–75.

Zippenfenig, P. (2023). *Open-Meteo.com Weather API [Computer software]*. DOI: 10.5281/ZENODO.7970649. URL: https://doi.org/10.5281/ZENODO.7970649.

A Appendix

A.1 List of Abbreviations

API Application Programming Interface

ARIMA AutoRegressive Integrated Moving Average

CDF Cumulative Distribution Function

CV Coefficient of Variation

DL Deep Learning

EDA Exploratory Data Analysis

GBT Gradient Boosted Trees

IQR Interquartile Range

LSTM Long Short-Term Memory

MAE Mean Absolute Error

ML Machine Learning

MQL Multi Quantile Loss

MSE Mean Squared Error

QC Quantile Crossing

RF Random Forest

RNN Recurrent Neural Network

SAA Sample Approximation Algorithms

VAE Variational Autoencoders

WSAA Weighted Sample Approximation Algorithms

A.2 Tables

A.2.1 List of Special Events

Event	Dates
Carnival (Fasching)	2020-02-23, 2024-02-03, 2024-02-11, 2022-02-13, 2023-02-19
Shopping Sunday (Mandelsonntag)	2022-10-30, 2023-10-29, 2021-10-31
Christmas Market	2023-12-01, 2023-12-23, 2022-11-05, 2022-12-23, 2020-11-27, 2020-12-15
City Festival (Stadtfest)	2022-09-16, 2022-09-17, 2023-09-13, 2023-09-14
Residence Run (Residenzlauf)	2022-05-01, 2023-04-30, 2024-04-28
Mozart Festival (Mozartfest)	2020-06-09, 2020-06-27, 2021-05-28, 2021-06-27, 2022-05-20, 2022-06-19, 2023-06-02, 2023-07-02, 2024-05-24, 2024-06-23
Residence Wine Festival (Residenz Weinfest)	2024-06-28, 2024-07-07, 2023-06-30, 2023-07-09, 2022-07-10
Kiliani Fair (Kiliani)	2022-07-01, 2022-07-17, 2023-06-30, 2023-07-16, 2024-07-05, 2024-07-21
Spring Festival (Frühjahrsvolksfest)	2024-03-09, 2024-03-24, 2023-03-18, 2023-04-02, 2022-03-26, 2022-04-10
Africa Festival	2024-05-30, 2024-06-02, 2023-05-26, 2023-05-29, 2022-05-26, 2022-05-29
Street Food Festival	2022-05-28, 2022-05-29
All Saints Fair (Allerheiligenmesse)	2023-10-28, 2023-11-12, 2022-10-29, 2022-11-13
Harbor Summer (Hafensommer)	2024-07-19, 2024-08-04, 2023-07-21, 2023-08-06, 2022-07-22, 2022-08-07, 2021-08-07, 2021-08-08
Free & Outdoors (Umsonst & Draußen)	2024-06-20, 2024-06-23, 2023-06-08, 2023-06-11, 2022-06-16, 2022-06-19, 2021-09-02, 2021-09-05
Wine Village (Weindorf)	2022-05-25, 2022-06-06, 2023-05-26, 2023-06-04, 2024-05-29, 2024-06-09
Wine Parade (Weinparade)	2024-08-22, 2024-09-01, 2023-08-24, 2023-09-03, 2022-08-25, 2022-09-01

Table 8: Special events and their dates

A.2.2 Daily Forecasts with Temperature-Based Weather Predictions

Method	MQL	MAE	Improvement MQL (%)	Improvement MAE (%)
Two Days Horizon				
LSTM	1208.27	2834.14	61.86	67.37
LightGBM	1201.75	3070.48	62.06	64.64
RandomForest	1299.91	3223.70	58.97	62.88
Quantile Regression	1910.13	3289.70	39.70	62.12
CatBoost	1465.73	3593.60	53.73	58.62
Empirical Quantiles	3167.88	8684.56	0.00	0.00
Seven Days Horizon				
LSTM	1094.73	2554.08	64.12	69.31
LightGBM	1209.63	3022.49	60.36	63.69
Quantile Regression	3476.67	3159.31	-13.94	62.04
RandomForest	1317.56	3464.33	56.82	58.38
CatBoost	1478.93	3618.10	51.53	56.53
Empirical Quantiles	3051.34	8323.53	0.00	0.00

Table 9: Evaluation of different methods with two and seven day horizons, including weather forecast feature.

A.2.3 Hourly Forecasts with Temperature-Based Weather Predictions

Method	MQL	MAE	Improvement MQL (%)	Improvement MAE (%)
48 Hours Horizon				
LSTM	73.99	167.65	72.81	76.94
CatBoost	69.50	174.95	74.46	75.93
RandomForest	77.52	190.34	71.52	73.82
LightGBM	76.99	193.85	71.71	73.33
Quantile Regression	233.83	232.42	14.08	68.03
Empirical Quantiles	272.15	726.97	0.00	0.00
168 Hours Horizon				
LSTM	74.53	166.62	72.57	77.19
CatBoost	68.39	173.02	74.83	76.32
RandomForest	75.53	185.82	72.20	74.57
LightGBM	75.44	192.42	72.23	73.66
Quantile Regression	265.48	218.81	2.28	70.05
Empirical Quantiles	271.68	730.58	0.00	0.00

Table 10: Evaluation of different methods for 48-hour and 168-hour horizons, including weather forecast feature.

A.3 Code

The code corresponding to this bachelor thesis can be found online on Github, see:

`https://github.com/luisadosch/Code---BA`