

Predicting Courier Activity for Optimal Operations

Edris Hakimi



Executive Summary



Forecast the number of online courier partners for the next 30 days, evaluating XGBoost and Prophet to determine the most accurate model for capturing demand trends.

The analysis showed an **upward trend** in courier demand, **weekly fluctuations**, and **seasonal variations**. **Outliers were detected and corrected** using the KNN method, improving data reliability.

Prophet outperformed XGBoost ($R^2 = 0.832$ vs. **0.432**), providing more accurate courier demand forecasts. It predicted an **average of 72 couriers/day** over 30 days.



Given Data and Preprocessing Summary

1

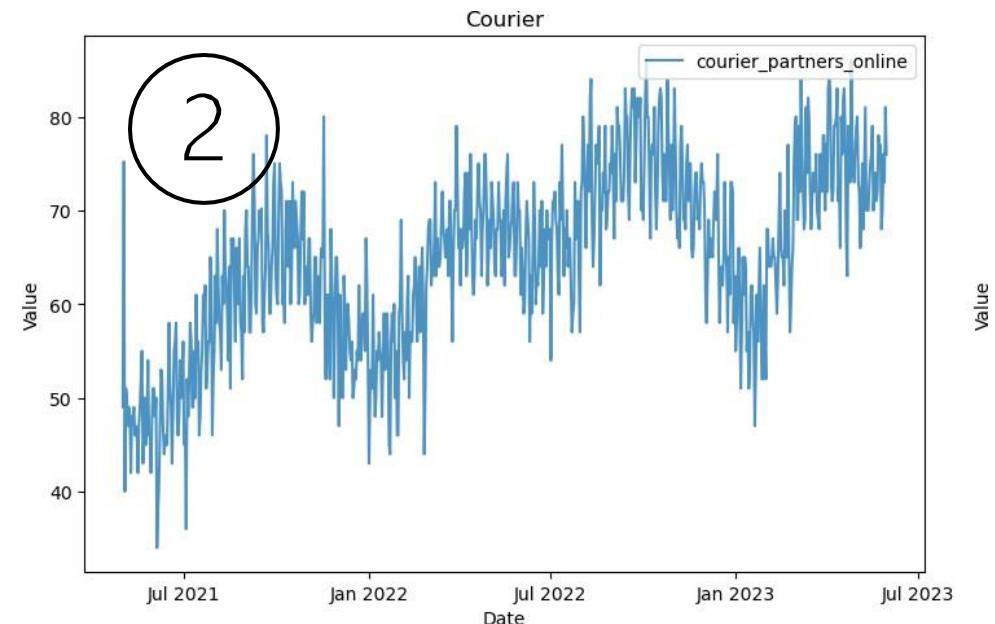
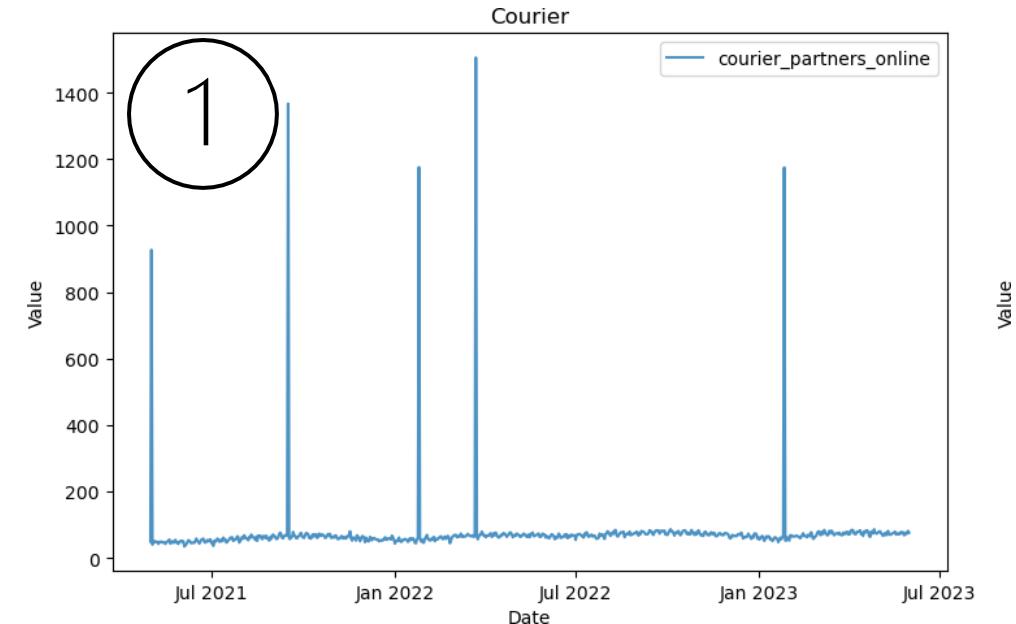
Before processing, it was evident that the courier dataset contained potential outliers. Further investigation confirmed their presence.



Clear outliers were identified in the courier data and addressed by filtering and imputing values using the KNN method.

2

After processing, the dataset has been cleaned, and the outliers have been effectively handled, resulting in improved data quality and reliability.



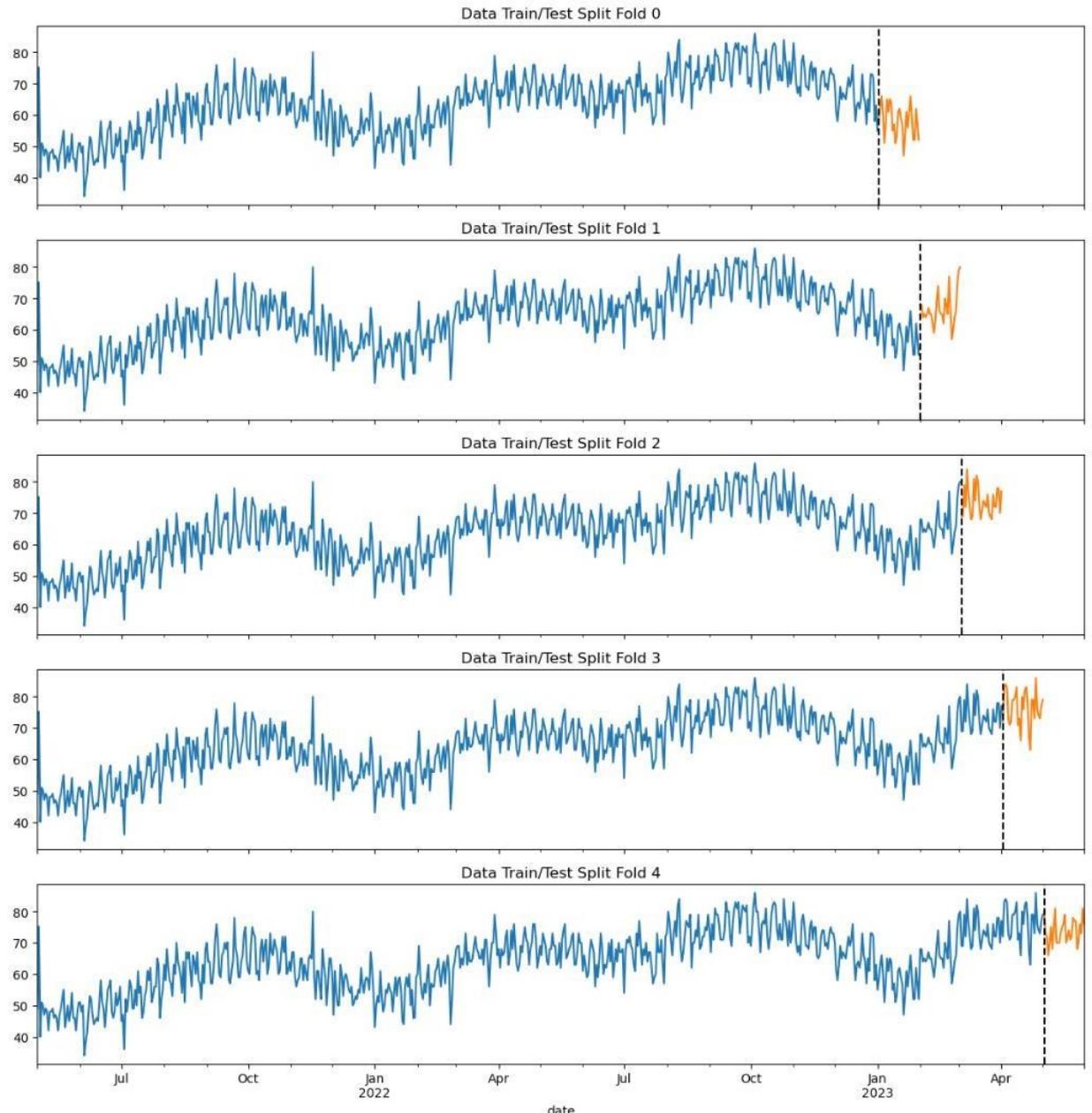
XGBoost Model

XGBoost is efficient, handles missing values and outliers, and captures complex patterns.

Rolling window cross-validation was used to prevent overfitting, with four-week time windows for this time-series task.

Features used: **Day of the year, Day of the week, Quarter, Month, Year, Lag1, Lag2, Lag3** (e.g., Lag1 = data from one year ago).

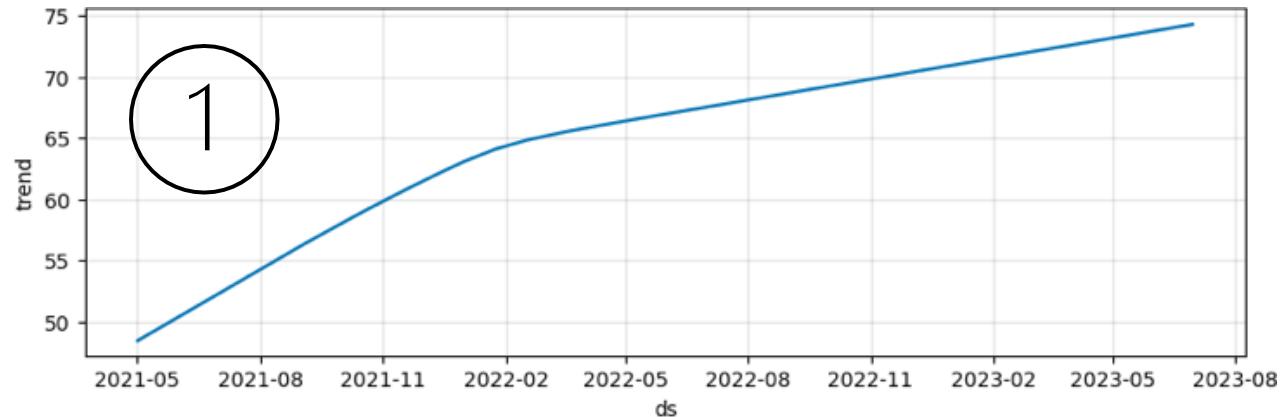
Folds (0-4): Each fold trains on earlier data and tests on the next four-week period, ensuring realistic time-series predictions.



Prophet Model

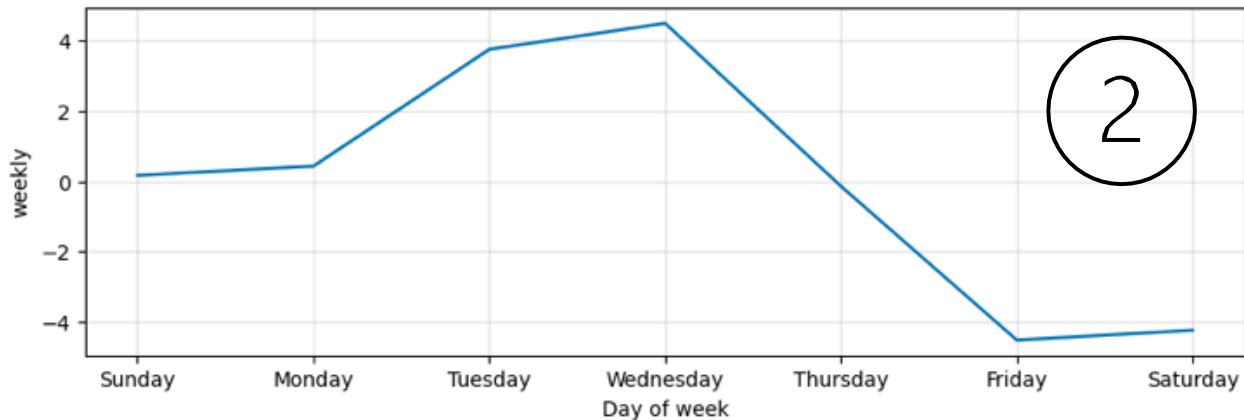
1

The time series shows an upward trend, indicating an increase in orders over time as Wolt's usage grows.



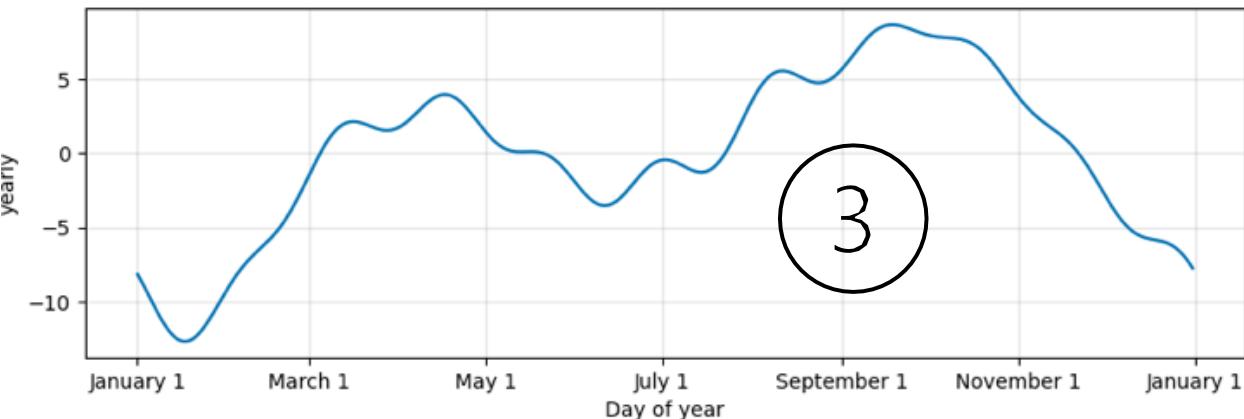
2

The graph shows that courier demand varies by day of the week, with fewer couriers needed toward the end of the week. This could be due to people dining out more on weekends instead of ordering food at home.



3

The graph shows seasonal variation, with fewer orders in summer and early winter due to outdoor dining, travel, holidays, and New Year's resolutions.

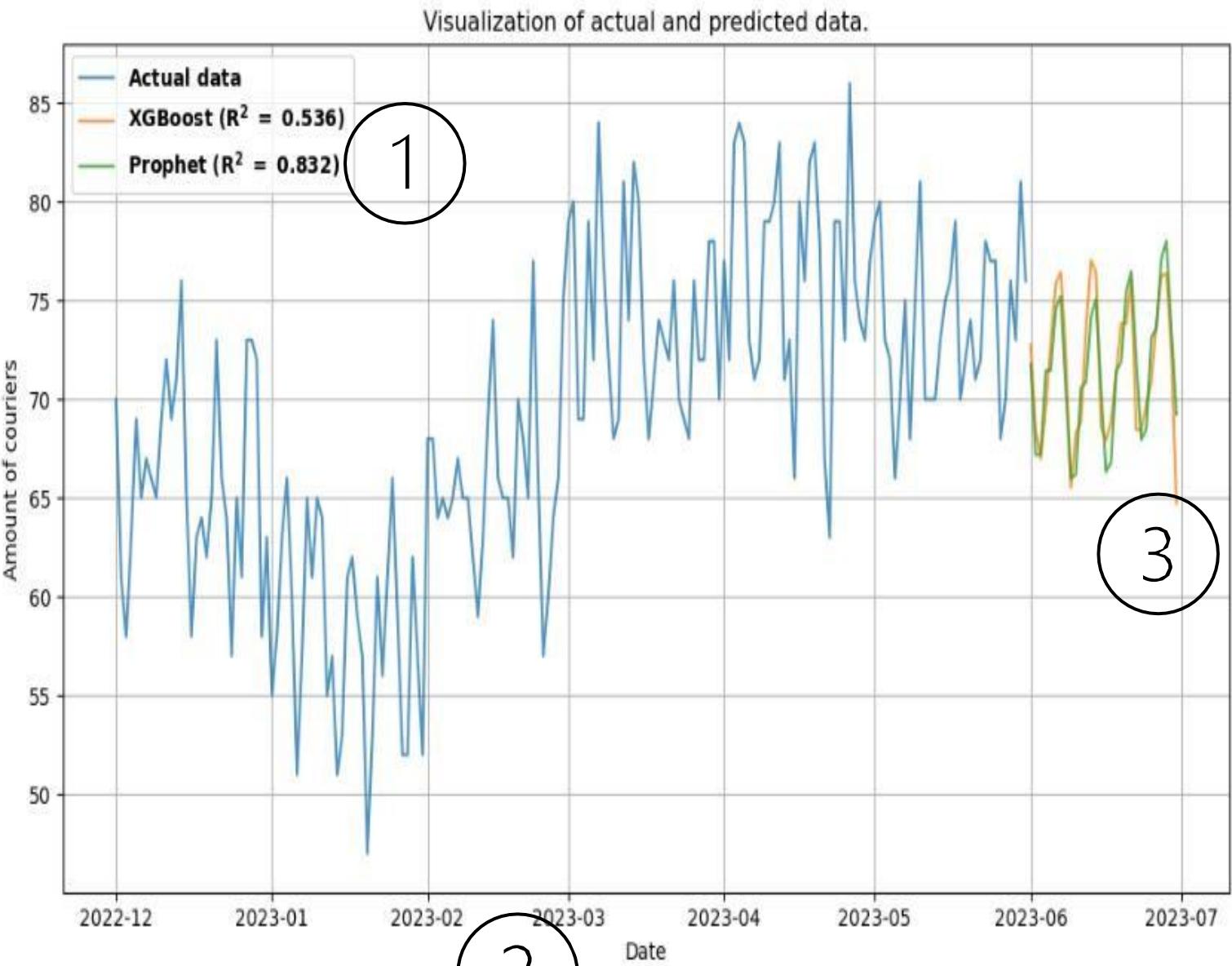


Results

- 1 Prophet achieves a higher R^2 (0.833) compared to XGBoost (0.432), making it the better model for capturing trends and forecasting.
- 2 The graph visualizes actual and predicted courier demand, focusing on the last 6 months for clarity.
- 3 Both models provide similar forecasts for the upcoming two weeks, with Prophet predicting an average of 72 couriers/day.

Why This Matters for Wolt:

Accurate predictions like those from the Prophet model enable Wolt to optimize resource allocation, ensuring the right number of couriers are available to meet customer demand effectively. This helps improve operational efficiency, reduce costs, and enhance customer satisfaction.



Optimizing Courier Demand Forecasting

	Enhance data inputs	Explore alternative models	Extend prediction horizons	Use an ensemble approach
Description	Incorporate external factors like public holidays, promotions, and local traffic to improve forecasting accuracy.	Test LSTMs for deep learning-based time-series analysis and ARIMA/SARIMA for capturing seasonality.	Move beyond 30-day forecasts to 3–6 months for better long-term planning and strategic decision-making.	Combine Prophet for trend forecasting and XGBoost for short-term fluctuations to improve overall prediction accuracy.
Motivation	Courier demand is influenced by external factors such as public holidays, promotions, and local traffic events. Incorporating these variables can improve prediction accuracy.	While Prophet and XGBoost performed well, other models may capture patterns differently . LSTMs are powerful for sequential data, and ARIMA/SARIMA specialize in time-series seasonality.	The current 30-day forecast is useful for short-term operations, but longer-term planning (3–6 months) enables better staffing, budgeting, and strategy development.	Each model has strengths: Prophet is better for trend detection, while XGBoost captures short-term fluctuations. Combining them can leverage their advantages .
Impact	Helps models capture real-world demand shifts, leading to more precise resource allocation and better operational planning .	Could improve accuracy in long-term trends and short-term variations , providing a more comprehensive forecasting solution.	Improves business decision-making , allowing Wolt to anticipate demand changes and adjust resources accordingly .	Results in higher forecasting accuracy , reducing uncertainty and ensuring more reliable courier availability predictions .



Thank You

