Leveraging Machine Learning for Revenue Optimization via Strategic Couponing

DS405 Machine Learning Applications in Business and Economics | Winter Term 2024/2025 | Maurice Schmetzer

Contents

Model Performance & Revenue Impact

Model Performance

(Based on Validation Dataset)

- **Trained Models:** Random Forest, Adaboost, XGBoost
- XGBoost performed best with highest increase in revenue.
- Validation Set Accuracy: **68.26%**
- **Only 10.42%** of customers are predicted to not purchase again even though they would (False Negatives).

Will Customers Re-Order within 90 Days?

Reality		Model Prediction: No	Model Prediction: Yes
	No	60.01%	21.32%
	Yes	10.42%	8.25%

Revenue Impact of False Negatives is higher than False Positives!

Revenue Impact

(Based on Validation Dataset)

	Voucher for Everyone	Voucher Only for Model Predictions
Revenue Increase per Customer	0.08 €	0.23 €
Expected Revenue Increase of Test Set	2,699.33 €	7,429.42 €
Financial Benefit of Implementing Model	4,730.08 € (+ 175.23%)	

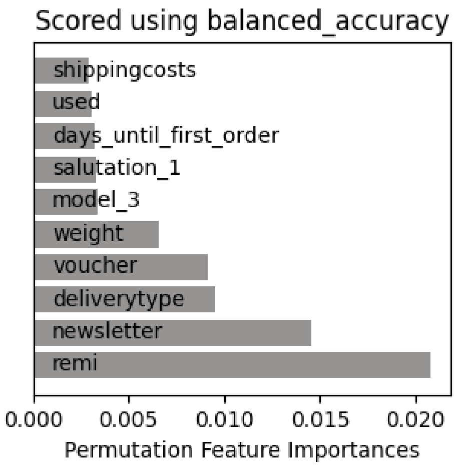
Criteria Driving Model Predictions

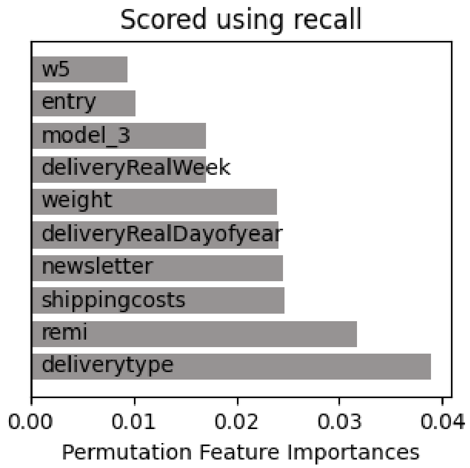
Features

- Features: Data associated with a customer's intital order.
- Removed Features: Customer Number, Title, Domain, Points (No Variation), Dates
- Added Features:

Features	Explanation
12 Date Features	Key-Date variables for dates including day of deek, day of year, indicator if it is month, quarter or year end, etc.
Delivery Delay Days	Variable that shows delivery delay by number of days.
Delivery Delay	Variable that indicates if there was any delivery delay.
Significant Delivery Delay Days	Variable that indicates if there was a delivery delay above 10 days.

Permutation Feature Importance (PFI)

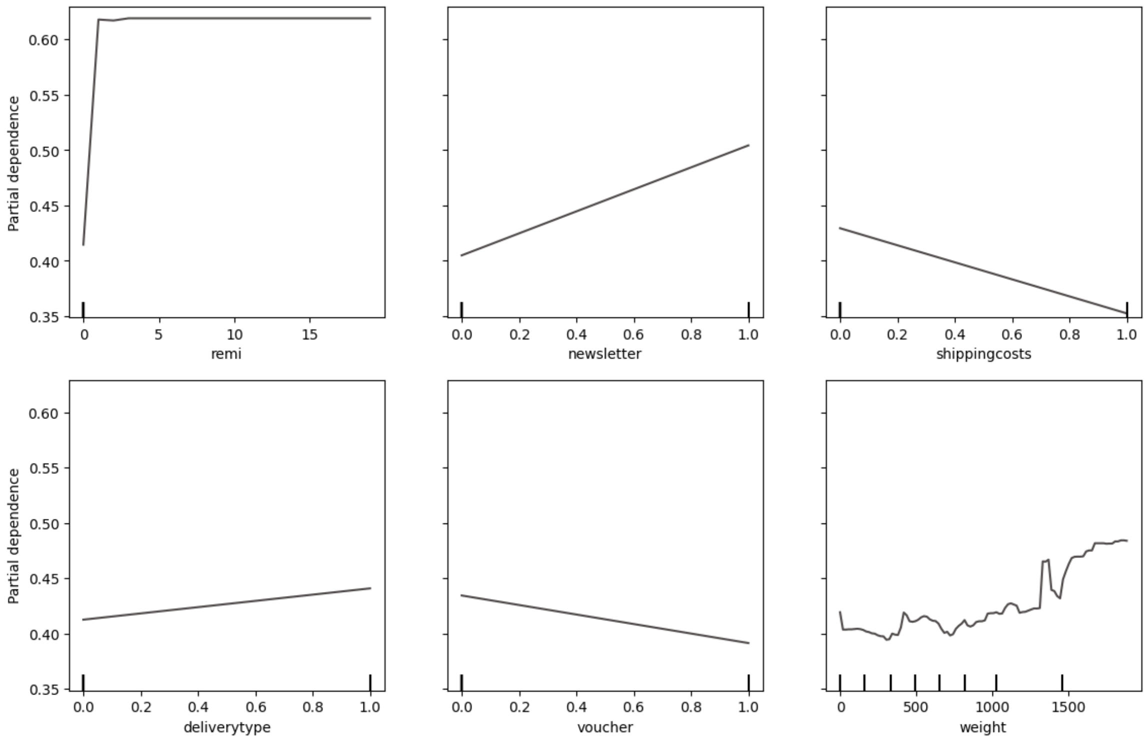




Most important features:

- Remi = Number of remitted / returned items
- Newsletter = Indicates if customer is subscribed to newsletter
- Shipping Costs = Indicates if customer paid shipping cost
- Delivery Type = Indicates if order was dispatched or collected
- Voucher = Indicates if customer redeemed a voucher
- Weight = Weight of Shipment

Partial Dependence Plot (PDP)

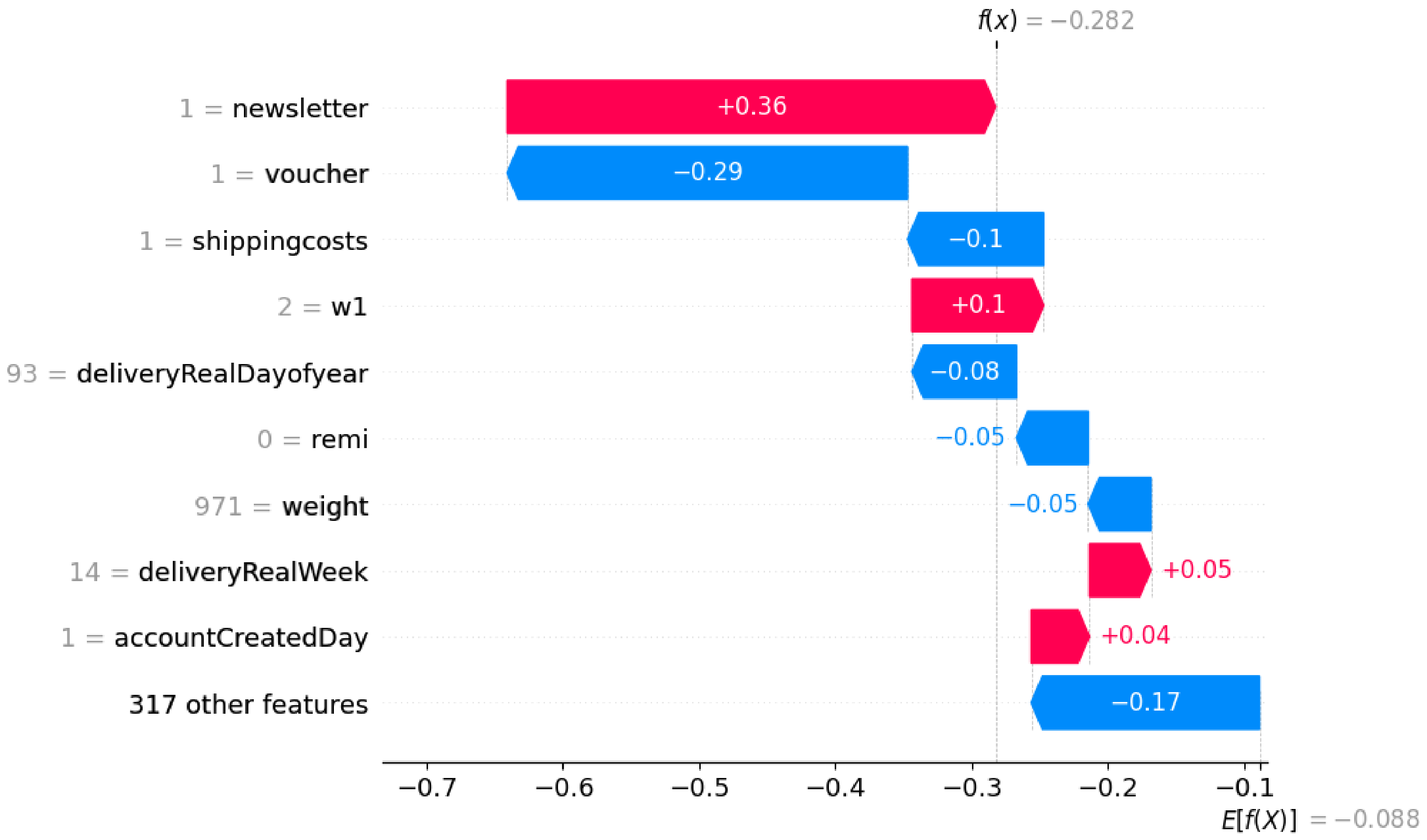


Features		Prob. to Buy Again
More Remitted Items	↑	Small number enough
Subscribed to Newsletter	↑	
Paid Shipping Costs	↓	
Order Collected	↗	Only small impact
Used Voucher	↓	
Weight	↗	Probability increase at larger weights

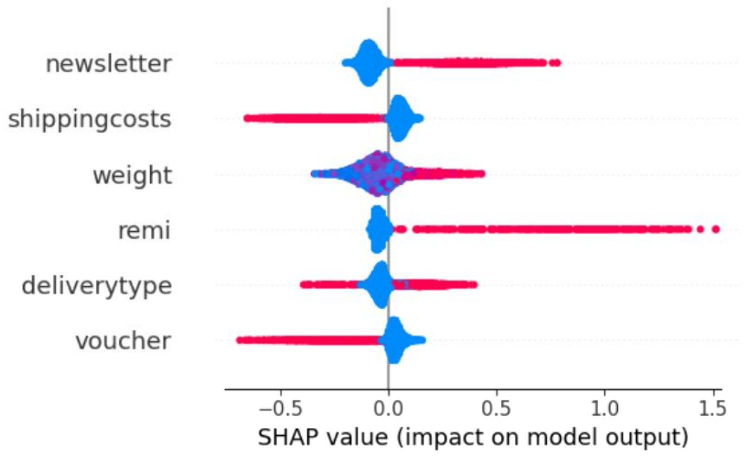
PDP – Possible Explanations

Features	Prob. to Buy Again	Possible Reasons
More Remitted Items	↑	Person that returned item possibly buys a different/new item.
Subscribed to Newsletter	↑	Person thinks more often about products due to newsletter & is reminded to buy again.
Paid Shipping Costs	↓	Customers might find other shops without shipping cost.
Order Collected	↗	Mixed experiences of picking up (Good & Bad)?!
Used Voucher	↓	Cost are generally too high & customer only bought last time due to voucher.
Weight	↗	Mixed relation. Unclear.

Shapley Additive Explanations (SHAP) - Local



Shapley Additive Explanations (SHAP) - Global



Similar observations:

- Newsletter unsubscribed customers have a big negative impact on retentions while subscribing has a positive impact.
- Paying Shipping Cost has negative impact on retentions.
- Large weight has positive impact while small weight has both positive and big negative impacts.
- Higher number of remitted items has positive impacts while low number has larger negative impacts.
- Delivery Type has mixed results. Order collection has both negative and positive impact while dispatched orders have mostly negative impact on retentions.
- Having used a voucher previously reduced likelihood of another purchase.

Recommendations

Recommendations

Immediate Steps

- Use the model & Send out vouchers!

  → Increases average revenue from 0.08 € per customer to 0.23 € (+175.23%)

Further Steps

- Increase number of customers subscribed to newsletter.
- Reduce shipping cost.

Thank You!