

Customer Lifetime Value Prediction

Using Machine Learning to Predict Customer Value

The Starks

Othmane Zizi (261255341)

Fares Joni (261254593)

Tanmay Giri (261272443)

Context

What is Customer Lifetime Value (CLV)?

The total revenue a business expects from a customer over their entire relationship.

Why Does It Matter?

- ▶ Acquiring new customers costs **5-25x more** than retaining existing ones
- ▶ CLV-driven strategies improve acquisition efficiency by **5-8%**
- ▶ Enables smarter allocation of marketing budgets

Business Applications

- ▶ Identify high-value customers for VIP treatment
- ▶ Target at-risk customers with retention campaigns
- ▶ Segment customers for personalized marketing

Hypothesis

"Historical purchase patterns (recency, frequency, and monetary value) can accurately predict a customer's future lifetime value."

What We're Testing

- ▶ Can RFM features predict future customer spending?
- ▶ Which ML model performs best? (Linear vs. Random Forest vs. XGBoost)
- ▶ Can we create actionable customer segments from predicted CLV?

Success Metrics

RMSE

Prediction Error

R²

Model Fit

Lift

Top 20% Identification

Data

Online Retail II Dataset — UCI Machine Learning Repository

541K

Transactions

2 Years

Dec 2009 - Dec 2011

UK

Online Giftware Retailer

RFM Features

- ▶ **Recency** — days since last purchase
- ▶ **Frequency** — number of orders
- ▶ **Monetary** — total spend

Behavioral Features

- ▶ **Tenure** — customer age
- ▶ **Avg time** between purchases
- ▶ **Unique products** bought

Target Variable

CLV = Total customer spend in the next 6 months

GitHub Repository

othmane-zizi-pro / nwa

Type / to search

Code Issues 7 Pull requests Discussions Actions Projects 1 Wiki Security Insights Settings

nwa Public

Pin Watch 0 Fork 0 Star 0

main 1 Branch 0 Tags Go to file Add file Code

othmane-zizi-pro and claude Initial commit: CLV Prediction Project 729b93e · 17 minutes ago 7 Commits

| File | Commit Message | Time |
|-------------------|--|----------------|
| data | Initial commit: CLV Prediction Project | 17 minutes ago |
| notebooks | Initial commit: CLV Prediction Project | 17 minutes ago |
| reports/figures | Initial commit: CLV Prediction Project | 17 minutes ago |
| src | Initial commit: CLV Prediction Project | 17 minutes ago |
| .gitignore | Initial commit: CLV Prediction Project | 17 minutes ago |
| LICENSE | Create LICENSE | 8 hours ago |
| README.md | Initial commit: CLV Prediction Project | 17 minutes ago |
| presentation.md | Initial commit: CLV Prediction Project | 17 minutes ago |
| presentation.pptx | Initial commit: CLV Prediction Project | 17 minutes ago |
| requirements.txt | Initial commit: CLV Prediction Project | 17 minutes ago |

About

No description, website, or topics provided.

Readme MIT license Activity 0 stars 0 watching 0 forks

Releases

No releases published [Create a new release](#)

Packages

No packages published [Publish your first package](#)

GitHub Project Board

The screenshot shows the GitHub Project Board interface for the repository "othmane-zizi-pro / Projects / CLV Prediction Project".

Key elements of the interface include:

- Header:** Shows the repository path and a search bar with placeholder "Type / to search".
- Toolbar:** Includes buttons for "Add status update", "Insights", "Workflows 6", and a "View" dropdown.
- Section Header:** "CLV Prediction Project" with a "View 1" button and a "+ New view" link.
- Filter Bar:** "Filter by keyword or by field" and a "View" settings icon.
- Table:** A grid-based project board with columns for "Title", "Assignees", "Status", "Linked pull requests", and "Sub-issues progress".
- Data:** A list of 9 tasks:
 - 1 Initialize repository structure #1 (Done)
 - 2 Download and load Online Retail II dataset #2 (Done)
 - 3 Data cleaning and preprocessing #3 (In Progress)
 - 4 Feature engineering (RFM + behavioral) #4 (Todo)
 - 5 Exploratory Data Analysis #5 (In Progress)
 - 6 Train baseline model #6 (Todo)
 - 7 Train advanced models (RF, XGBoost) #7 (Todo)
 - 8 Customer segmentation #8 (Todo)
 - 9 Final presentation #9 (In Progress)
- Bottom Bar:** A note "+ You can use [Control + Space] to add an item".

Summary

- | **Problem:** Identify high-value customers
- | **Data:** 541K retail transactions over 2 years
- | **Approach:** RFM features → ML models → Segmentation
- | **Goal:** Enable targeted, data-driven marketing

Questions?