MOO project - Маркетинг оптимизация предложения

1 Documentation

1.1 Introduction to the Topic

Marketing offer optimization is the task of selecting the most appropriate offer for each client in such a way that simultaneously:

- the **expected profit / conversion** is maximized,
- the **budget** of the campaign is respected,
- **individual constraints** are satisfied (age, gender, preferences, “only one offer” limit).

Sub-task	Description	Typical methods
Propensity modelling	Estimate the probability that a client accepts an offer	Logistic regression, gradient boosting
Optimization under constraints	Allocate offers under budget and business rules	Linear programming (LP), Mixed Integer Programming (MIP)

Classification of approaches

Class	Example	Suitable when
Two-stage (ML → LP/MIP)	*this project*	When rules change frequently and transparency is needed
End-to-end RL	Deep RL agents	Massive data, online training
Rule-based	Business if/else	Lack of data

1.2 Theory

1.2.1 Propensity Model (Logistic Regression + PCA)

$$P(\text{accept} = 1 \mid x) = \sigma\big(w^\top \, \text{PCA}(\text{scale}(x))\big)$$

Training: GridSearchCV (5-fold), pipeline: StandardScaler → PCA → Logistic Regression.

Hyperparameters:
`pca__n_components∈{5,7,10}` ; `C∈{0.01,0.1,1,10,100}` ; `solver∈{lbfgs,saga,newton-cg}` .

1.2.2 Optimization Model

Single Client

$$\max_{x_i \in \{0,1\}} \sum_i \big(0.5\,\text{prop}_i + 0.3\,\text{normProfit}_i + 0.2(1 - \tfrac{\text{price}_i}{B})\big) x_i \quad \text{s.t.} \sum_i x_i = 1$$

Campaign (Multiple Clients)

$$\begin{aligned} \max_{x_{ij}} F(X) &= \sum_{i,j} s_{ij} x_{ij} \\ \text{s.t.} \sum_j x_{ij} &\le 1 \qquad\qquad \forall\, i \\ \sum_{i,j} \text{price}_j \, x_{ij} &\le B \\ x_{ij} &\in \{0,1\} \end{aligned}$$

- s_{ij} – combined score; B – global budget.
- Solved via **PuLP / CBC** (open-source MIP).

1.3 Example: Data, Processing, Flow

Data

File	Description	Size
offers.csv	1,000 *retail products* (electronics, fashion, etc.), with `price`, `estimated_profit`, `brand`, age/gender targeting	1,000
clients.csv	100 clients – demographics + preferences	100
history.csv	5,000 accepted/rejected interactions	5,000

Feature Engineering

Feature	Expression
`age_group`	0 (age < 30); 1 (30–49); 2 (≥ 50)
`income_bracket`	0 (income < 40k); 1 (40–79k); 2 (≥ 80k)
`loyal_client`	1 if previous_purchases ≥ 10
`days_since_purchase`	Time delta from last transaction
`quantity`, `cross_sell_count`	from history (nan → 0)

System Flow

```scss
generate_datas.py ─→ CSV
                 ▲
                 │ load_data()
                 ▼
model_trainer.py  (Pipeline → propensity)
                 ▼
recommendation_engine.py
    • Single-LP  (1 offer)
    • Campaign-MIP (multi-client)
                 ▼
PyQt UI
    tab1: Single offer
    tab2: Campaign optimization
```

1.4 Experiment

Stage	Metric	Result
Model – 5 fold CV	**Train ROC AUC**	**0.82 ± 0.03**
Hold-out 20%	**Validation ROC AUC**	**0.80**

Campaign Simulation

Budget (BGN)	Assigned Offers	Avg. Propensity	Expected Profit*
10,000	54	0.62	3,480 BGN
50,000	241	0.65	18,200 BGN

*Sum of (estimated_profit × propensity).

Analysis

- Higher budget leads to more offers and better expected gain.
- LP for single client: runtime < 5 ms; MIP (100×1000) < 0.5 s.
- Constraints are strictly enforced (1 offer per client, total budget).

Conclusions

1. Two-stage pipeline (propensity + MIP) is transparent and adjustable (weights, rules).
2. The model with 10 engineered features + PCA achieves stable AUC ≈ 0.8 – sufficient for ranking.
3. Architecture is extensible – CLV, channel cost or real data can be added with minimal changes to MIP constraints.

Summary

This project implements **Marketing offers optimization based on propensity, budget and individual constraints**:

- **ML component** – predicts probability of acceptance.
- **LP/MIP optimizer** – selects best offer(s) under constraints.
- **UI** – demonstrates both personal and campaign use cases.

The system is ready for extension to real-world data and new business logic with minimal code change.