Management Report: Bank Marketing “Term Deposit” Predictor (end-to-end DSM)

This project transitions bank marketing from random calling to a data-driven strategy. By eliminating "data leakage" (removing call duration) and using SMOTE to balance the data, we created a model that predicts subscriptions based solely on client profiles.

Model Performance & Strategy

Model	Strategy	Accuracy	Recall	Business Impact
Dummy	Baseline	88.3%	0.00	Failed; identified zero subscribers.
Log. Reg.	Linear	88.0%	0.08	Weak; misses almost all potential leads.
CART	Conservative	76.2%	0.53	Efficient; best for limited call capacity.
Random Forest	Aggressive	44.5%	0.87	Optimal for growth; captures 87% of market.

Key Findings (EDA & SMOTE)

- **The 3-Call Limit:** Subscription probability drops drastically after the 3rd contact. Further attempts yield diminishing returns.
- **Target Profile:** Students, retirees, and individuals without existing loans are the most receptive segments.
- **Recall Optimization:** By adjusting the threshold to 0.3, Random Forest identified 1,386 subscribers, missing only 200 cases (False Negatives).

Strategic Recommendations

Objective	Recommended Model	Actionable Insight
Aggressive Growth	Random Forest (0.3)	Captures 87% of subscribers; prioritizes market share over call costs.
Cost Efficiency	CART	Maintains 76% accuracy; minimizes calls to uninterested leads.
Operational	Call Capping	Terminate contacts after the 3rd attempt to save Call Center resources.

Conclusion : Random Forest (Threshold 0.3) is the superior path for maximizing deposits, as the value of a new subscriber outweighs the cost of additional calls.

Project Team: Aleks Kadëna, Dea Xhanari, Ambra Lino