A Comparative Machine Learning Study using Decision Trees and Ensemble Methods

Bank Telemarketing Profit Optimization

Course: Machine Learning

Date: June 10, 2025

Team Members

- **Laggah Julius**
  ID: 2120246034
- **Jusu Abdul Karim**
  ID: 2120246025
- **Sandar Win**
  ID: 2120246058
- **Turay Adamsay**
  ID: 2120246055

Introduction & Motivation

Introduction & Motivation

- **Background:** Financial institutions invest heavily in telemarketing for term deposit subscriptions.
- **Challenge:** Maximizing customer acquisition while minimizing wasted calls (high cost: $15/call, high CLV: $450/subscriber).
- **Problem Statement:** Balancing predictive accuracy with profitability, especially with asymmetric misclassification costs.
- **Profit Equation:**

$$Profit = (TP \times CLV) - ((TP + FP) \times Cost_{call})$$

(TP: True Positives, FP: False Positives, CLV: Customer Lifetime Value, $Cost_{call}$: Cost per call)

Key Contributions

Key Contributions

- Demonstrated effectiveness of **Cost-Complexity Pruning** in improving decision tree performance and interpretability.
- Introduced a **profit-centric evaluation framework** for robust model comparison.
- Extracted **actionable business rules** from optimized models for targeted outreach.

Dataset and Preprocessing

Dataset & Preprocessing

- **Dataset:** UCI Bank Marketing Dataset (41,188 instances, 20 features).
    - Numerical: Age, duration, campaign, euribor3m, etc.
    - Categorical: Job, marital, education, contact, etc.
    - Target: 'y' (binary: yes/no, 11.3% positive class).
- **Preprocessing Pipeline:**
    - Handling "unknown" values (mode imputation).
    - One-Hot Encoding for categorical features.
    - Feature reduction (highly correlated features removed).
    - Stratified 80/20 Train/Test Split.

Model Development

Model Development

- **Baseline Decision Tree:**
  - Unpruned CART classifier with Gini impurity.
  - Test ROC-AUC: 0.745.
- **Cost-Complexity Pruning (CCP):**
  - Minimized $C_\alpha(T) = R(T) + \alpha|T|$ to mitigate overfitting.
  - Hyperparameter grid search (max_depth, min_samples_split, ccp_alpha).

Comparative Models:

- Random Forest, Gradient Boosting, Logistic Regression

Performance Comparison

Performance Comparison

Table 1: Model Performance Comparison

Model	Accuracy	Precision	Recall	ROC-AUC
Baseline DT	0.897	0.543	0.549	0.745
Pruned DT	0.919	0.562	0.570	**0.943**
Random Forest	0.917	0.678	0.533	0.946
Logistic Regression	0.862	0.444	**0.902**	0.942
Gradient Boosting	**0.924**	**0.699**	0.566	**0.955**

- Pruned DT: Significant improvement in ROC-AUC (0.745 $\rightarrow$ 0.943).
- Gradient Boosting: Highest ROC-AUC (0.955), best overall discriminative power.
- Logistic Regression: Highest Recall (0.902), crucial for minimizing false negatives.

Model Performance

Model Performance: Discriminative Power

- We evaluated four models: **Pruned Decision Tree**, **Random Forest**, **Gradient Boosting**, and **Logistic Regression**.
- Our analysis included traditional metrics and profit-centric evaluation.

Discriminative Power (ROC-AUC):

- **Gradient Boosting**: ROC-AUC of **0.9551**.
- **Pruned Decision Tree**: Improved from 0.7449 to **0.9435** ($\uparrow$ **26.6%**), with tree size reduced by **94%** (3000 $\rightarrow$ 183 nodes).

💵 Model Performance: Campaign Profitability

Campaign Profitability:

- **Logistic Regression**: Highest profit of **$348,405**, driven by high recall (**90.19%**).
- **Gradient Boosting**: **$224,985**
- **Pruned Decision Tree**: **$214,980**

Model Interpretability

Pruned Decision Tree Structure & Interpretability

Pruned Decision Tree Structure

- The pruned decision tree offers significant interpretability, allowing us to extract actionable rules for targeted outreach.
- Its simplified structure effectively captures key decision paths relevant to customer subscription.

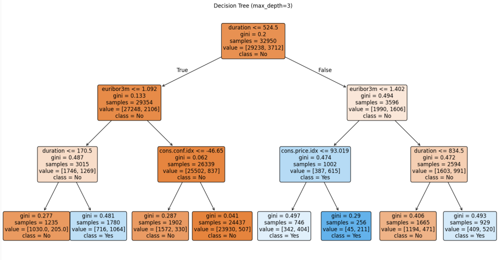

Figure 3: Decision Tree Structure (max depth = 3)

Feature Importance

Feature Importance & Interpretable Rules

Top 3 Feature Importances:

- **duration:** 0.3507 (most powerful predictor)
- **euribor3m:** 0.2197 (macroeconomic influence)
- **age:** 0.0980 (client propensity)

Interpretable Rules:

- **Rule 1:** If duration $\leq$ 524.5s AND euribor3m $\leq$ 1.402 $\implies$ 54.2% conversion.
- **Rule 2:** If 12.5 $<$ pdays $\leq$ 15.5 $\implies$ 59.8% conversion.

- Actionable insights for campaign managers.

Discussion

Discussion

- **Pruning Effectiveness:** 94% reduction in tree size (over 3000 to 183 nodes), 26.6% ROC-AUC improvement for Decision Trees.
- **Model Selection Trade-offs:** Profit optimization may favor simpler models (e.g., Logistic Regression) over statistically superior ones (e.g., Gradient Boosting) when false negatives are more costly.
- **Practical Implications:**
  - Interpretability of Decision Trees for actionable rules.
  - Feature importance for resource allocation.
  - Profit-driven evaluation aligns with business objectives.

Conclusion and Future Work

Conclusion & Future Work

- **Conclusion:** Pruned decision trees balance performance and interpretability. Profit optimization is crucial and can lead to selecting models with higher recall over higher overall accuracy.
- **Future Work:**
  - Experiment with entropy as splitting criterion.
  - Deeper feature analysis using SHAP.
  - Explore dynamic profit-aware learning algorithms.
  - Conduct real-world A/B testing.

References i

- Dahlén, M., Lange, F., & Rosengren, S. (2020). *Marketing Communications: A Brand Narrative Approach*. Sage Publications.
- Lessmann, S., Casu, B., & Reis, D. (2015). Predicting consumer churn in the mobile telecommunications industry: An application of a new hybrid approach. *Journal of Business Research*, *68*(4), 799–807.
- Verbeke, W., Martens, D., Baesens, B., & Van den Poel, D. (2012). Building profit functions for cost-sensitive learning in customer churn prediction. *Expert Systems with Applications*, *39*(1), 271–283.
- Molnar, C. (2020). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.). Leanpub.

References ii

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, F., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81–106.

Thank You for Your Attention!

Questions or feedback are welcome.