# Predictive Analytics Report

Project Title: Predicting Client Subscription to Term Deposit

Prepared by: Daniel Appiok

Date: June 14,2025

## 1. Executive Summary

This project aimed to build a data-driven model to predict whether a client would subscribe to a term deposit based on historical marketing campaign data.   
The outcome variable (y) represents client subscription: 'yes' or 'no'.  
  
The objective was twofold:  
1. Develop a reliable predictive model  
2. Generate actionable insights to support marketing strategy  
  
After evaluating multiple models, Logistic Regression was chosen for its balance between performance and interpretability — making it the most suitable for business decisions.

## 2. Dataset Overview

- Source: Direct marketing campaign data from a banking institution  
- Total Records: ~41,000 clients  
- Features: 20 input variables + 1 target variable (y)  
- No missing values; data quality was high  
  
Key Data Attributes:  
Client Info: Age, job, marital status, education, etc.  
Campaign Info: Call duration, contact type, day/month of last contact, etc.  
Historical Outcomes: Previous campaign results, number of contacts, etc.  
Target: y: whether the client subscribed (yes or no)

## 3. Methodology

The project followed a typical end-to-end data science workflow:  
  
1. Exploratory Data Analysis (EDA)  
2. Preprocessing  
3. Modeling Techniques Compared:  
 - Logistic Regression (Interpretable baseline)  
 - Random Forest (Ensemble of decision trees)  
 - XGBoost (Boosted gradient trees)  
4. Evaluation Metrics: Accuracy, Precision, Recall, F1 Score

## 4. Model Performance Comparison

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Accuracy | Precision | Recall | F1 |
| Logistic regression | 83 | 88 | 85 | 83 |
| Random forest | 89 | 88 | 89 | 88 |
| XGBoost | 90 | 88 | 90 | 89 |

Logistic Regression offered comparable F1 score to the more complex models but with clear interpretability — making it ideal for business use.

## 5. Key Insights & Recommendations

Top Influential Features:  
- Duration of last call  
- Contact method (cellular)  
- Month of contact  
- Outcome of previous campaigns  
  
Client Characteristics:  
- Engaged longer in calls  
- Responded positively to past campaigns  
- Contacted via cellular  
- Avoiding low-conversion months  
  
Recommendations:  
- Prioritize longer conversations  
- Avoid calls in May and August  
- Focus on clients with prior success  
- Emphasize cellular over telephone contacts

## 6. Conclusion

This project delivered a robust, interpretable model to guide marketing outreach strategy.   
It provided actionable insights based on real behavior — enabling data-driven targeting and efficiency gains in campaign performance.