ALY 6020:

PREDCTIVE ANALYTICS

Week 3: Logistic Regression Analysis: Subscription Behavior Prediction Models

Submitted To:
Prof. Chinthaka Pathum Dinesh Herath Gedara, Faculty Lecturer

Submitted By:
Abhilash Dikshit

Academic Term: Winter 2024
Northeastern University, Vancouver, BC,Canada
Master of Professional Studies in Analytics

January 27, 2024

Title: Subscription Behavior Prediction Models

I. Objective:

The goal of this analysis is to predict subscription behavior using logistic regression and support vector machine (SVM) models based on a dataset provided by a magazine company. The dataset contains information about customers, their demographics, and past interactions with the company.

II. Data Overview:

The dataset includes features such as customer demographics (age, income, education, marital status), past campaign acceptances, complaints, and various spending behavior metrics. The target variable, "Response," indicates whether a customer accepted the offer in the last campaign.

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines
0	5524	1957	Graduation	Single	58138.0	0	0	2012-09-04	58	635
1	2174	1954	Graduation	Single	46344.0	1	1	2014-03-08	38	11
2	4141	1965	Graduation	Together	71613.0	0	0	2013-08-21	26	426
3	6182	1984	Graduation	Together	26646.0	1	0	2014-02-10	26	11
4	5324	1981	PhD	Married	58293.0	1	0	2014-01-19	94	173
...	...	...	...	...	...	...	...	...	...	...
2235	10870	1967	Graduation	Married	61223.0	0	1	2013-06-13	46	709
2236	4001	1946	PhD	Together	64014.0	2	1	2014-06-10	56	406
2237	7270	1981	Graduation	Divorced	56981.0	0	0	2014-01-25	91	908
2238	8235	1956	Master	Together	69245.0	0	1	2014-01-24	8	428
2239	9405	1954	PhD	Married	52869.0	1	1	2012-10-15	40	84

2240 rows × 29 columns

III. Data Preprocessing:

Handled missing values using mean imputation.

Selected relevant features, including 'Year_Birth', 'Kidhome', 'Teenhome', 'Income', 'MntWines', 'NumWebVisitsMonth', and 'Recency'.

Converted categorical variables to numerical representations.

Split the data into training and testing sets.

IV. Modeling:

Two models were trained and evaluated: Logistic Regression and Support Vector Machine (SVM).

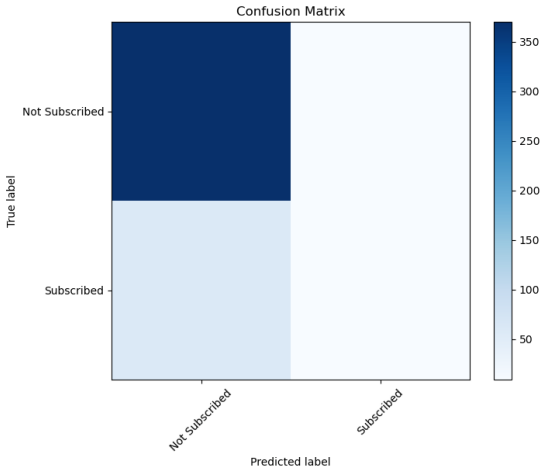
```
Metrics for Logistic Regression Model:
Accuracy: 0.8482
Precision: 0.5263
Recall: 0.1449
Confusion Matrix:
[[370   9]
 [ 59  10]]
Metrics for SVM Model:
Accuracy: 0.8594
Precision: 0.6364
Recall: 0.2029
Confusion Matrix:
[[371   8]
 [ 55  14]]
```

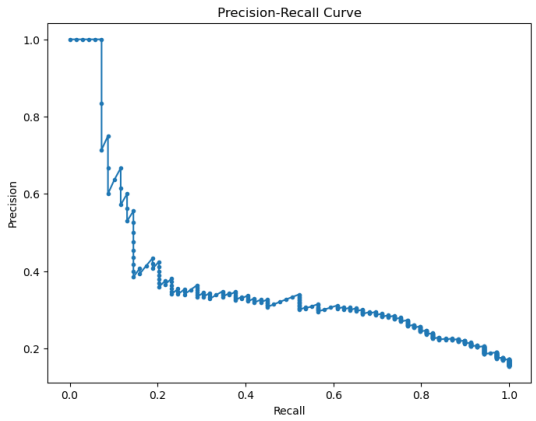
Logistic Regression:

Accuracy: 0.8482

Precision: 0.5263

Recall: 0.1449



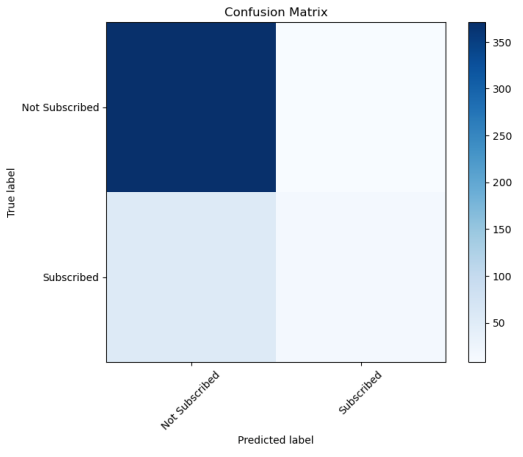


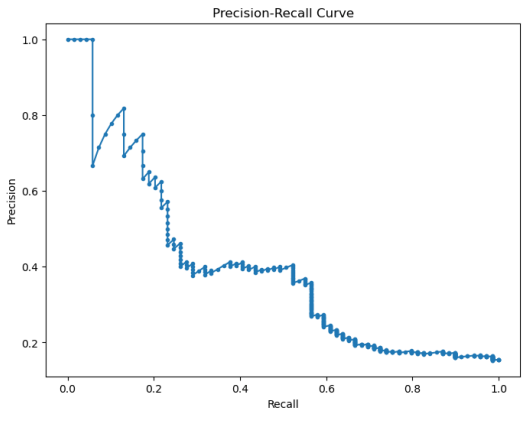
SVM:

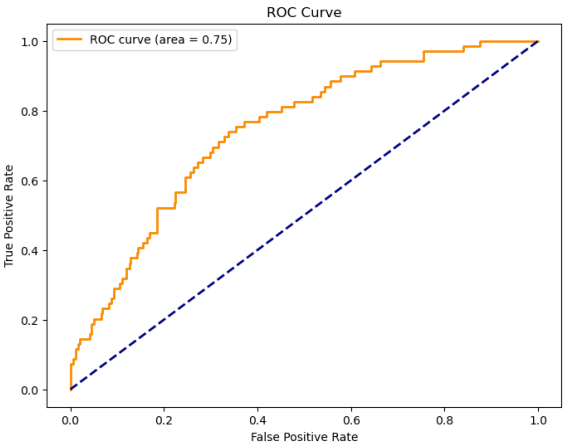
Accuracy: 0.8594

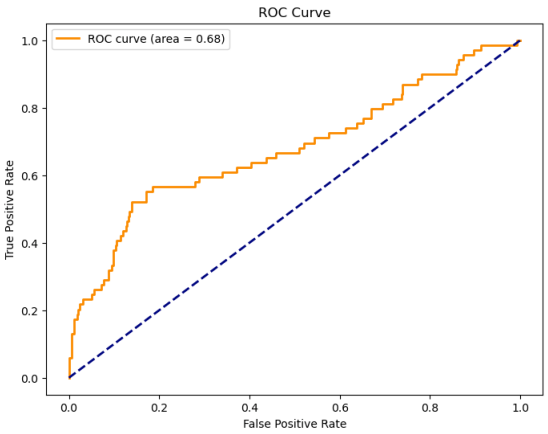
Precision: 0.6364

Recall: 0.2029









Performance Comparison:

Accuracy: Both models demonstrated similar overall accuracy, with SVM slightly outperforming Logistic Regression.

Precision: SVM exhibited higher precision, indicating a better ability to correctly identify true positives among predicted positives.

Recall: SVM had a higher recall, suggesting a better ability to capture positive cases among all actual positive cases.

V. Recommendation:

If precision and recall are equally important, or if there is a need for balance, the SVM model is recommended due to its higher precision and recall.

If precision is a critical factor, SVM may be preferred for its better performance in minimizing false positives.

Logistic Regression may be preferred for its interpretability, as it provides coefficients for each feature, aiding in understanding the impact of individual variables.

Next Steps:

Consider further hyperparameter tuning for both models.
Explore additional feature engineering to enhance model performance.
Evaluate models on a larger dataset if available for more robust conclusions.

VI. Conclusion:

The SVM model, with its slightly superior accuracy, precision, and recall, appears to be a promising choice for predicting subscription behavior. However, the final decision should be based on the specific business goals, interpretability needs, and the importance of precision and recall in the context of the magazine company's objectives.