

# PREDICTING THE SUCCESS OF KICKSTARTER PROJECTS

## AUTHORS

Shoon Lae Than Tun, Enes Talha Ciftci, Abdullakh Ali Unal

## AFFILIATIONS

Our research is proudly supported by our group(Rangsit University), reflecting the commitment to fostering innovative and impactful academic projects.

## INTRODUCTION

KICKSTARTER CONNECTS CREATORS WITH BACKERS, OFFERING A PLATFORM FOR FUNDING INNOVATIVE IDEAS. HOWEVER, ONLY 40% OF PROJECTS SUCCEED. THIS STUDY INVESTIGATES THE KEY FACTORS THAT IMPACT PROJECT SUCCESS AND USES MACHINE LEARNING TO PREDICT OUTCOMES. WE AIM TO PROVIDE INSIGHTS TO IMPROVE CROWDFUNDING SUCCESS RATES.

## OBJECTIVE

To predict the success of Kickstarter projects using machine learning and identify actionable factors that creators can focus on to optimize their campaigns.

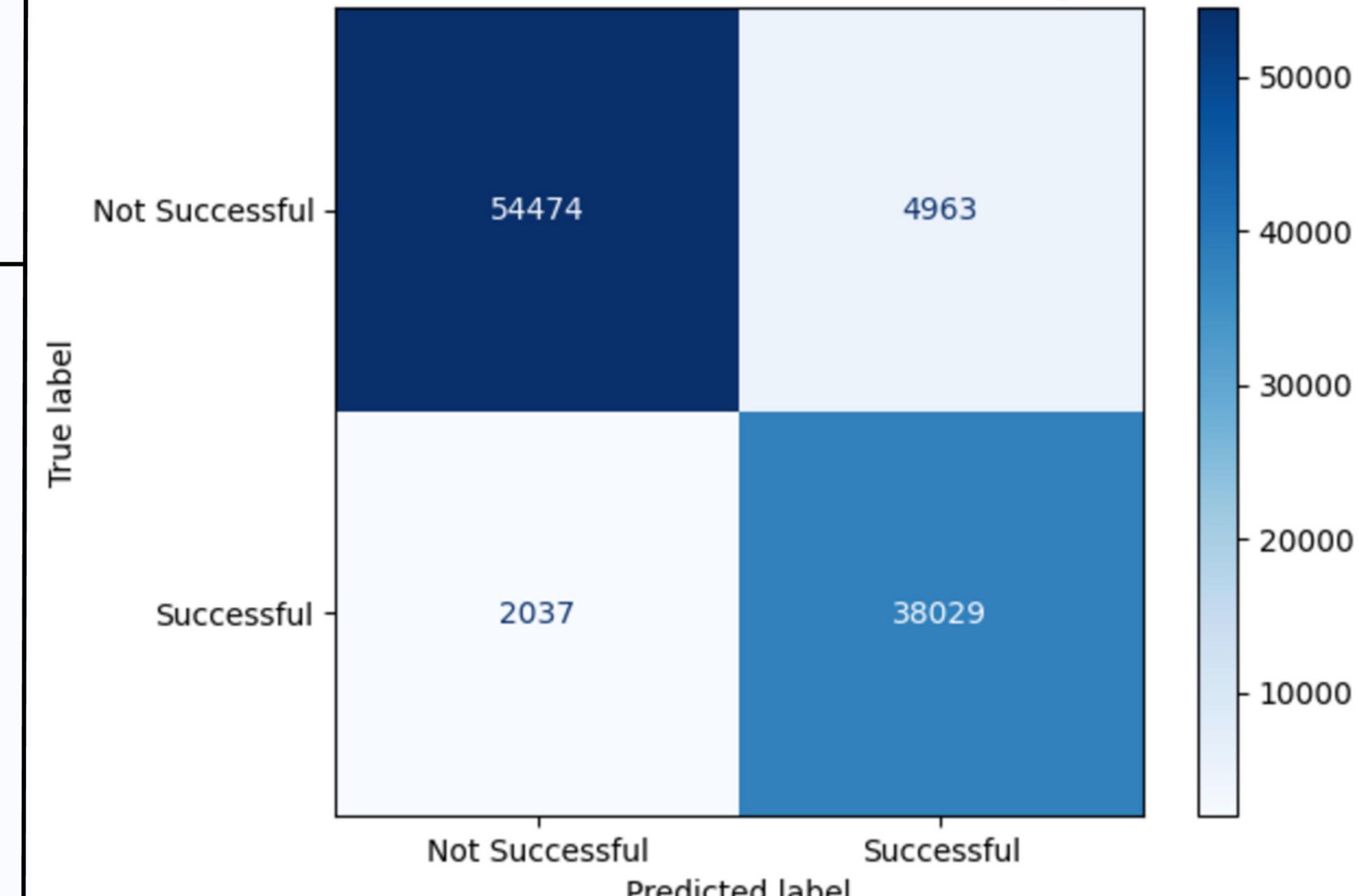
## METHODOLOGY

- Data Acquisition:  
Dataset sourced from Kaggle.
- Data Preprocessing:  
Cleaned and filtered by retaining only "Successful" and "Failed" in project states.
- Handling Missing values:
- Feature Engineering:  
Calculated campaign duration as the difference between deadline and launch date.
- Selected key features:  
project category, funding goal, and duration.
- Model Development:  
Implemented Logistic Regression and Gradient Boosting.  
Dataset split 70-30 for training and testing.  
Evaluated using accuracy, precision, recall, and F1 score.

## RESULTS/FINDINGS

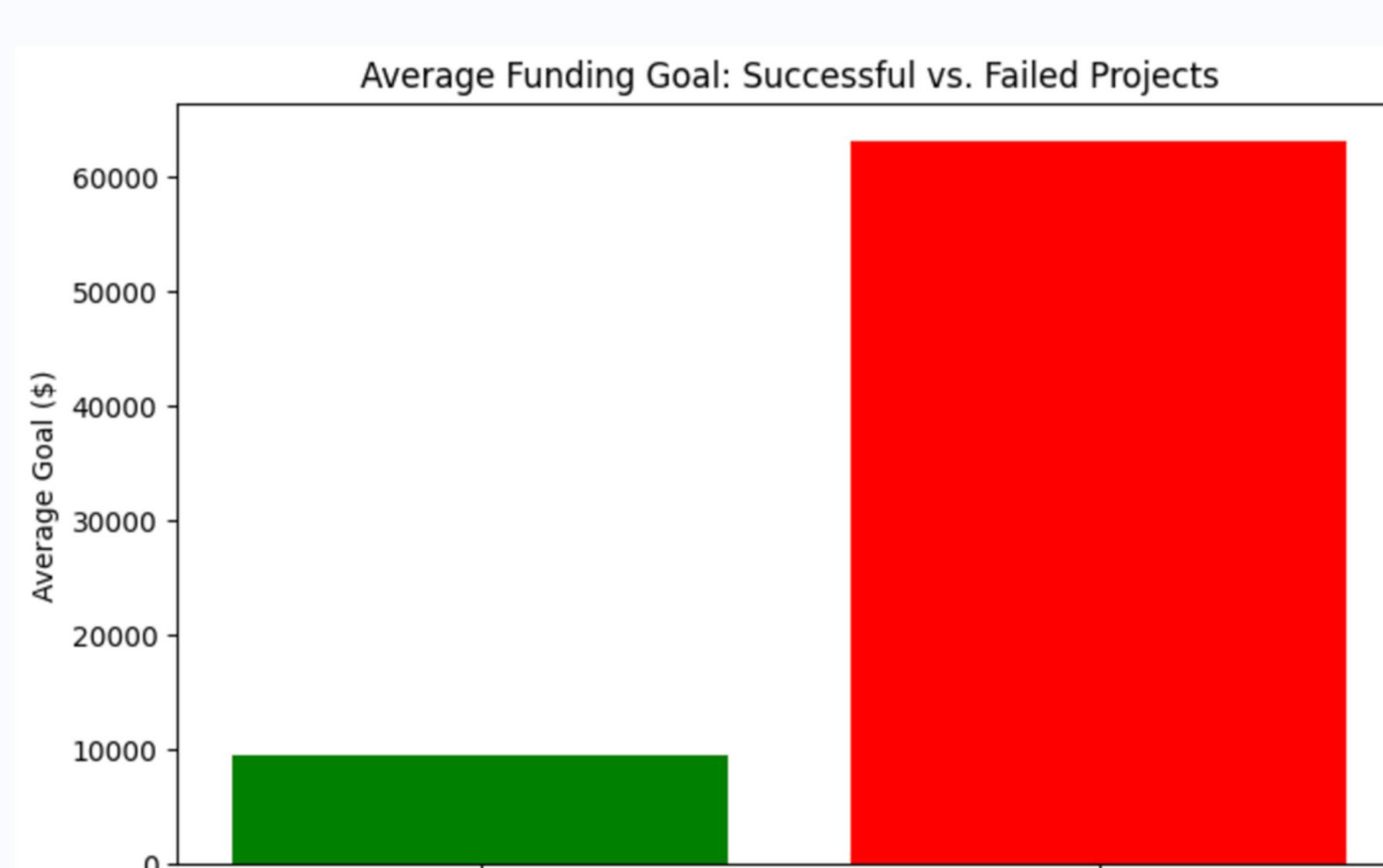
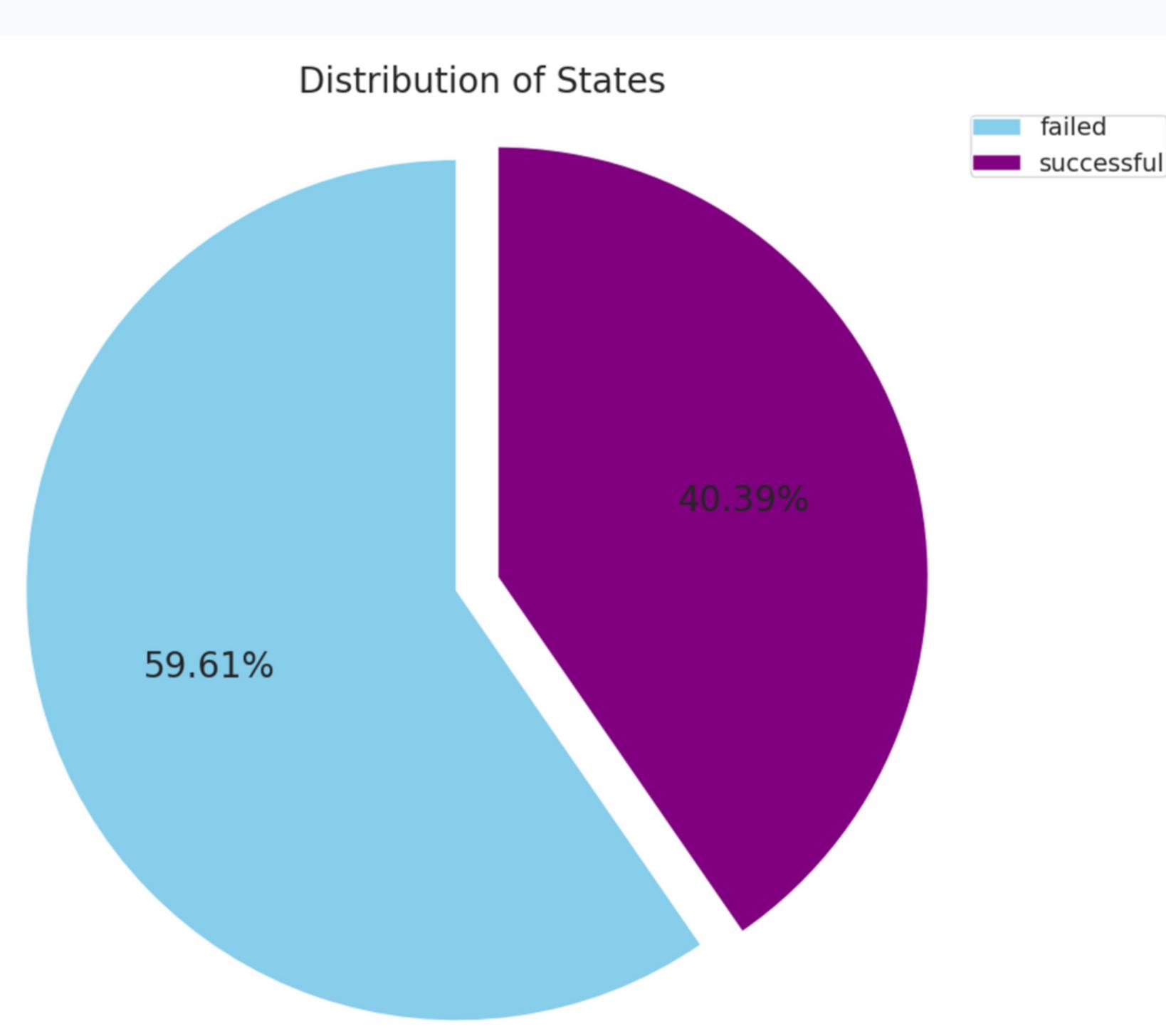
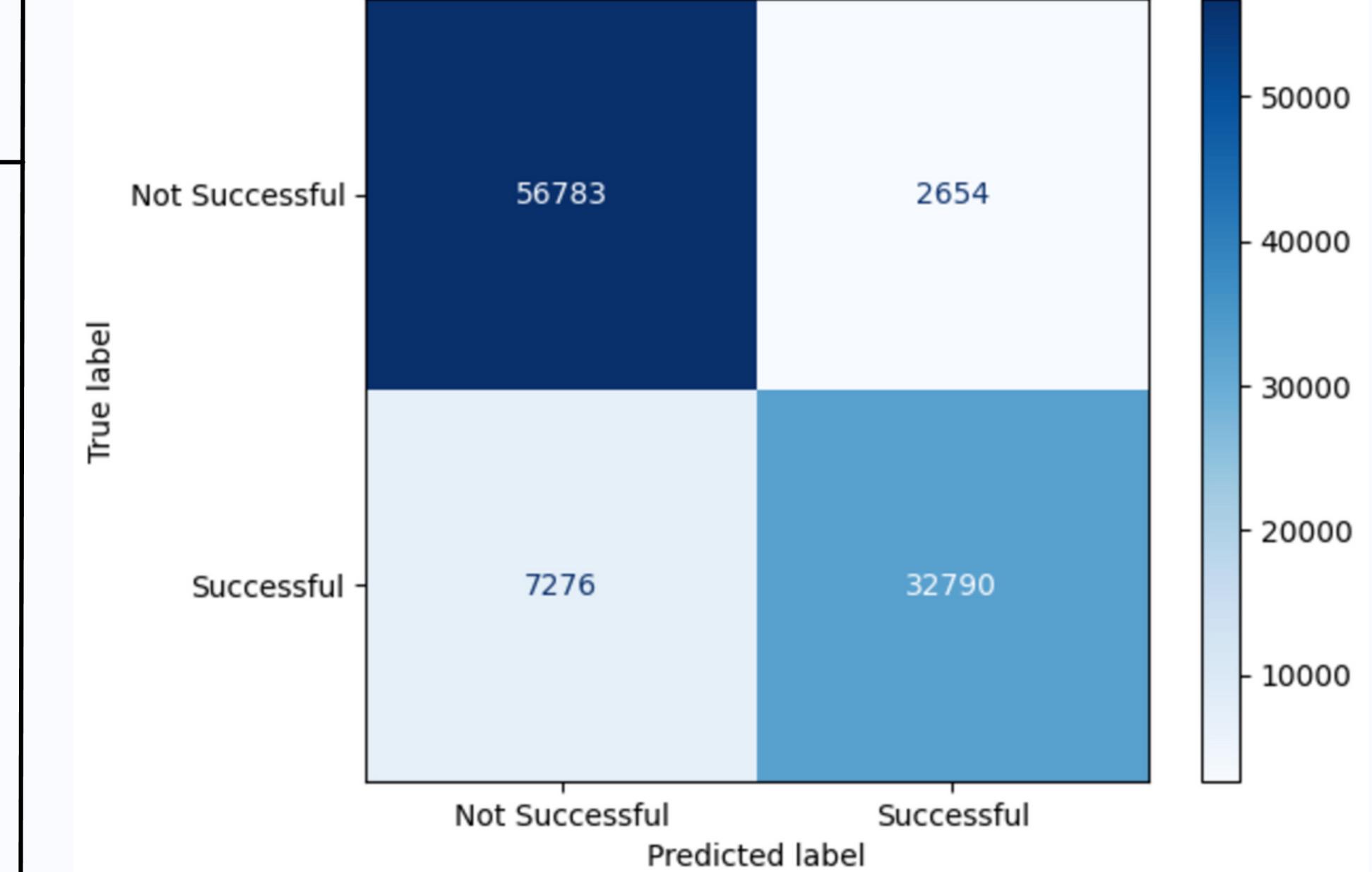
- Gradient Boosting achieved 92% accuracy, outperforming Logistic Regression (90%).
- Gradient Boosting had 95% recall for "Successful" projects but slightly lower precision (88%) while Logistic has 93%.
- Early backer engagement, manageable funding goals, and well-defined project descriptions were key success factors

Confusion Matrix - Gradient Boosting

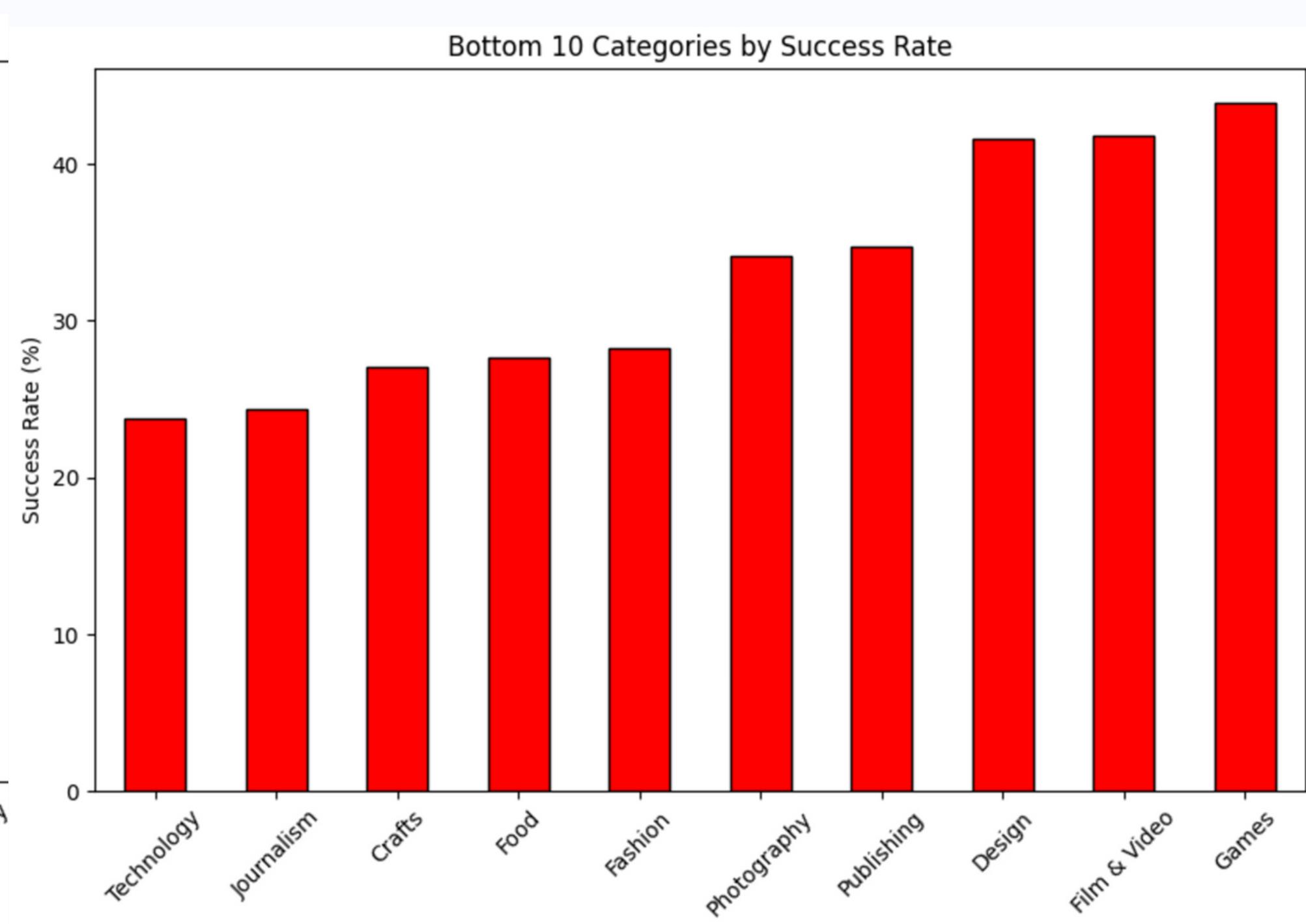
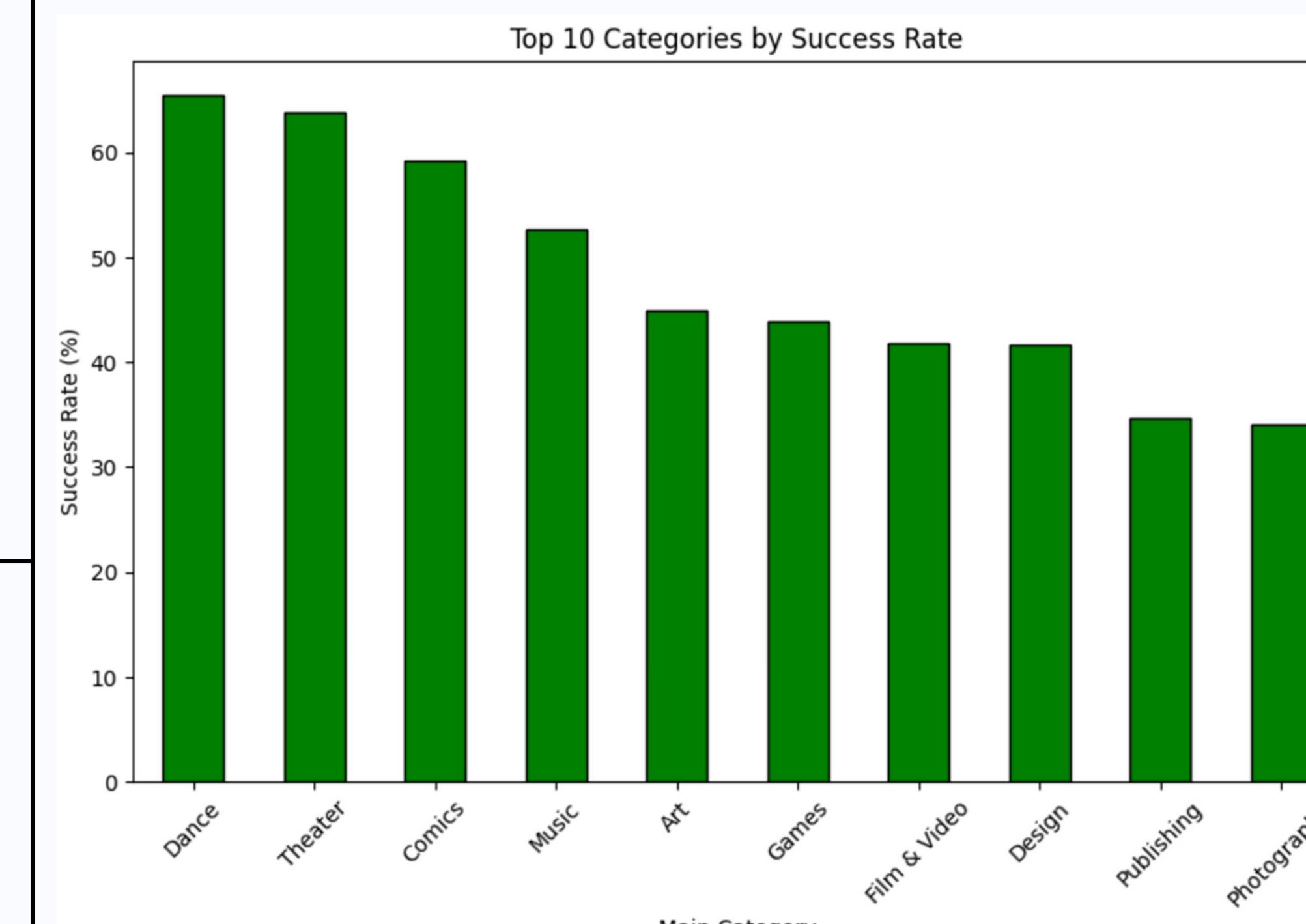


The model's performance shows that the model is most accurate at predicting the "Not Successful" class, with high classifications.

Confusion Matrix - Logistic Regression



Successful projects have lower funding goal compared to Failed projects.



## ANALYSIS

- Gradient Boosting effectively handled non-linear relationships, providing more reliable predictions, making it ideal for complex crowdfunding datasets.
- Logistic Regression offers simplicity and transparency but requires well-tuned features for optimal performance.
- Findings emphasize actionable strategies for creators, such as optimizing campaign duration and engaging early supporters.

## CONCLUSION

This study highlights the effectiveness of machine learning in predicting Kickstarter success, with Gradient Boosting providing the best balance of precision and recall. Future work could include unstructured data, like project descriptions, and explore ensemble methods to further enhance prediction accuracy.