

Introduction

In an ever-changing business landscape, success hinges on effective marketing strategies. This research delves into gaps in existing literature, aiming to unravel the complex relationship between marketing techniques and sales performance. We focus on radio, TV, and social media efficacy, the impact of Instagram interactions, and the correlation between user demographics and marketing outcomes. Through a robust methodology, our study offers practical insights for businesses to refine approaches and elevate sales in the evolving marketing terrain.

Abstract

This study investigates the nuanced connection between marketing techniques and sales, addressing gaps in current literature. Focusing on radio, TV, and social media effectiveness, Instagram impact, and user demographics, the research employs a comprehensive methodology. The literature review covers existing research on marketing techniques, Instagram, and demographics, offering insights into optimal channels and strategies. Findings, rooted in rigorous data analysis, provide practical implications for businesses, enhancing our understanding of marketing dynamics.

Problem Definition

Optimal Marketing Platform:

Which marketing platform yields the highest efficacy for businesses in reaching their target audience and achieving desired outcomes?

Influence of Ad Features on User Response:

How do the various features of advertisements impact user response and engagement, and what attributes contribute to a more favorable audience reaction?

Impact of Pre-collected Data on Ad Targeting:

To what extent does leveraging pre-collected data for ad targeting influence user response and engagement, and how does this strategic use of data contribute to advertising effectiveness?

Algorithms

We used different machine learning algorithms to build the best models to asses our data.

1-Regression Models:

What is regression?

Linear regression: models the relationship between the dependent variable and one or more independent variables as a linear equation.

Polynomial regression: is a form of linear regression where the relationship is modeled as an nth-degree polynomial

2-Classification Models:

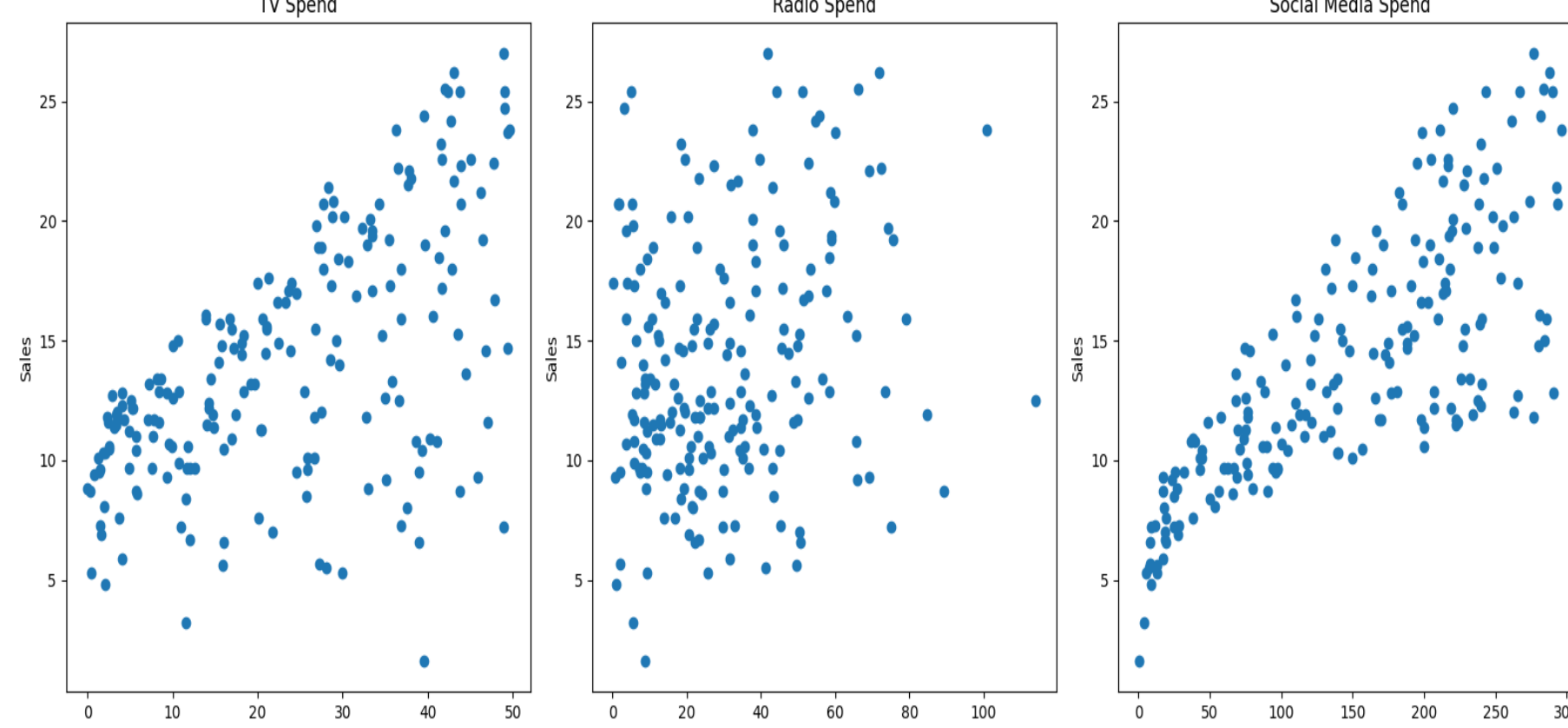
Classification Model	Pros.	cons
Logistic Regression	Probabilistic approach, gives information about statistical significance of features	The logistic regression assumptions
K-NN	Simple to understand, fat and efficient	Need to choose the number of neighbors k
SVM	Performant, not biased by outliers, not sensitive to overfitting	Not appropriate for non-linear problems, not the best choice for large number of features
Kernel SVM	High performance on non-linear problems ,not sensitive to overfitting	not the best choice for large number of features, more complex
Naïve Bayes	Efficient, not biased by outliers, works on non- linear problems, probabilistic approach	Based on the assumption that features have symmetrical relevance
Decision Tree	Interpretability, no need for feature scaling , works on both linear /non –linear problems	Poor results on too small datasets, overfitting can easily occur
Random Forests	Powerful and accurate, good performance on many problems, including non-linear	Not interpretability, overfitting can easily occur , need to choose the number of trees

Most Effective Marketing Technique

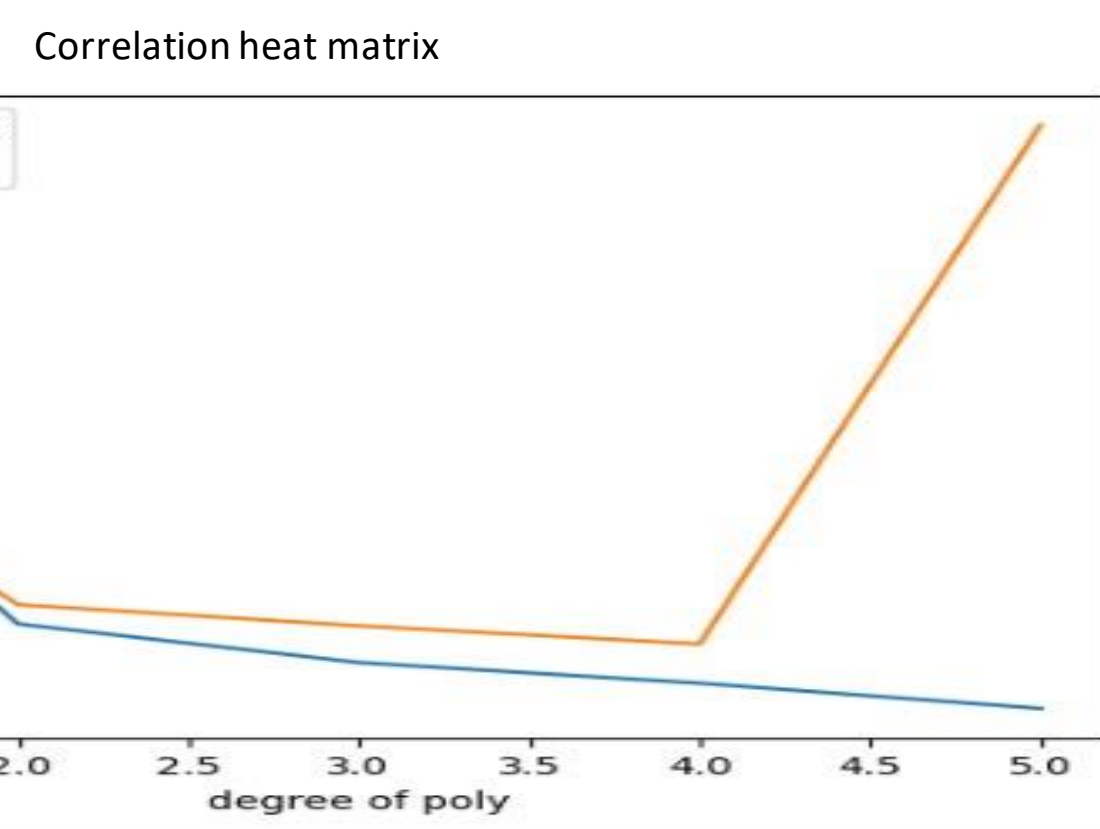
Goal: Determining the most effective marketing platform between TV, social media and radio.

Approach: Comparing turnover from investments in different platforms and constructing a model using regression.

Observations:



Scatter plot between different marketing techniques and sales



The correlation matrix indicates that social media yields the highest results. However, the linear regression equation did not accurately reflect this observation. Consequently, we chose to optimize the polynomial regression of the nth degree, discovering an improvement in error representation.

MAE (normalized)	0.0477852	MAE (normalized)	0.0154571
RMSE (normalized)	0.5969102	RMSE (normalized)	0.0138467
R2_Score_Train	0.8856655	R2_Score_Train	0.9952995
R2_Score_Test	0.9185781	R2_Score_Test	0.9908673

Linear regression errors

Polynomial regression errors

Findings:

There is a direct proportionality between Investment Capital and sales. The ranking of Marketing Techniques based on their impact on sales is as follows: Social Media, TV, and radio.

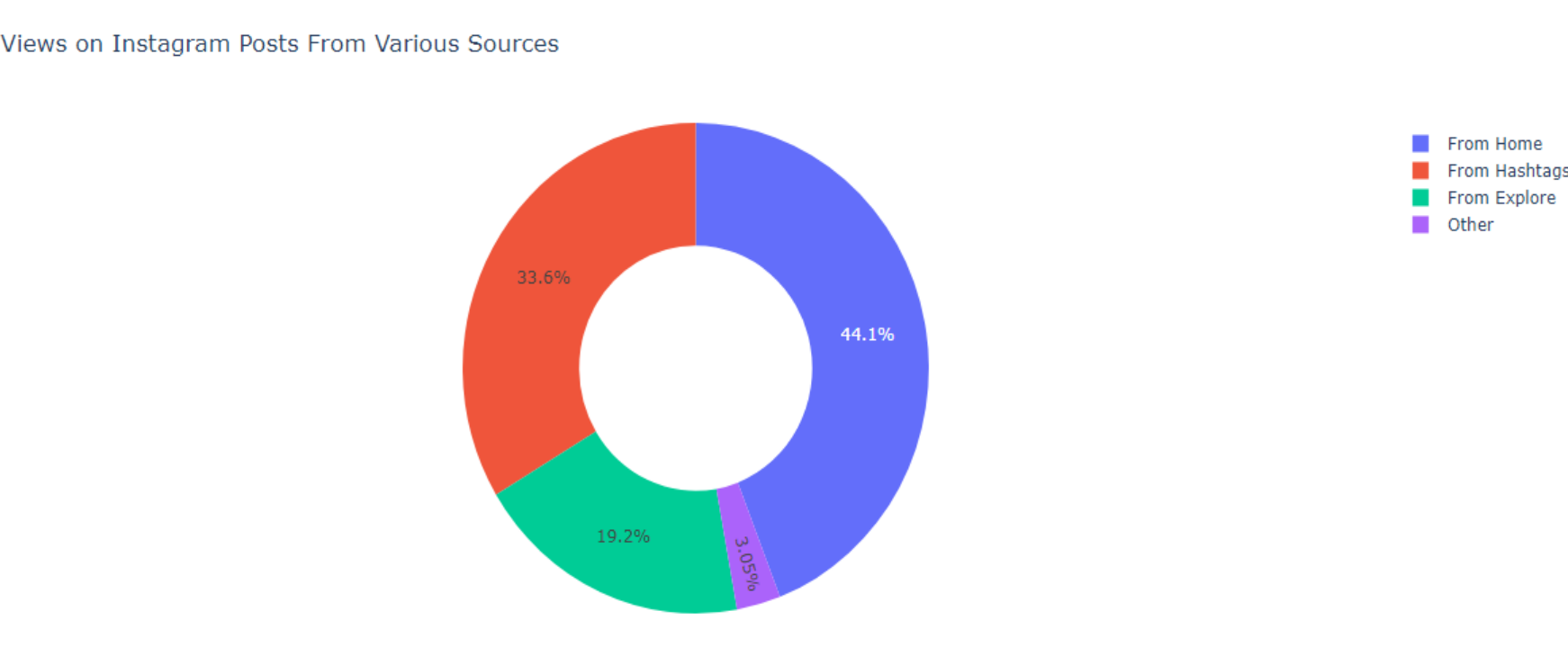
Ad Features Vs. Engagement

Goal:

Investigating the most effective strategies to enhance reach on an Instagram post.

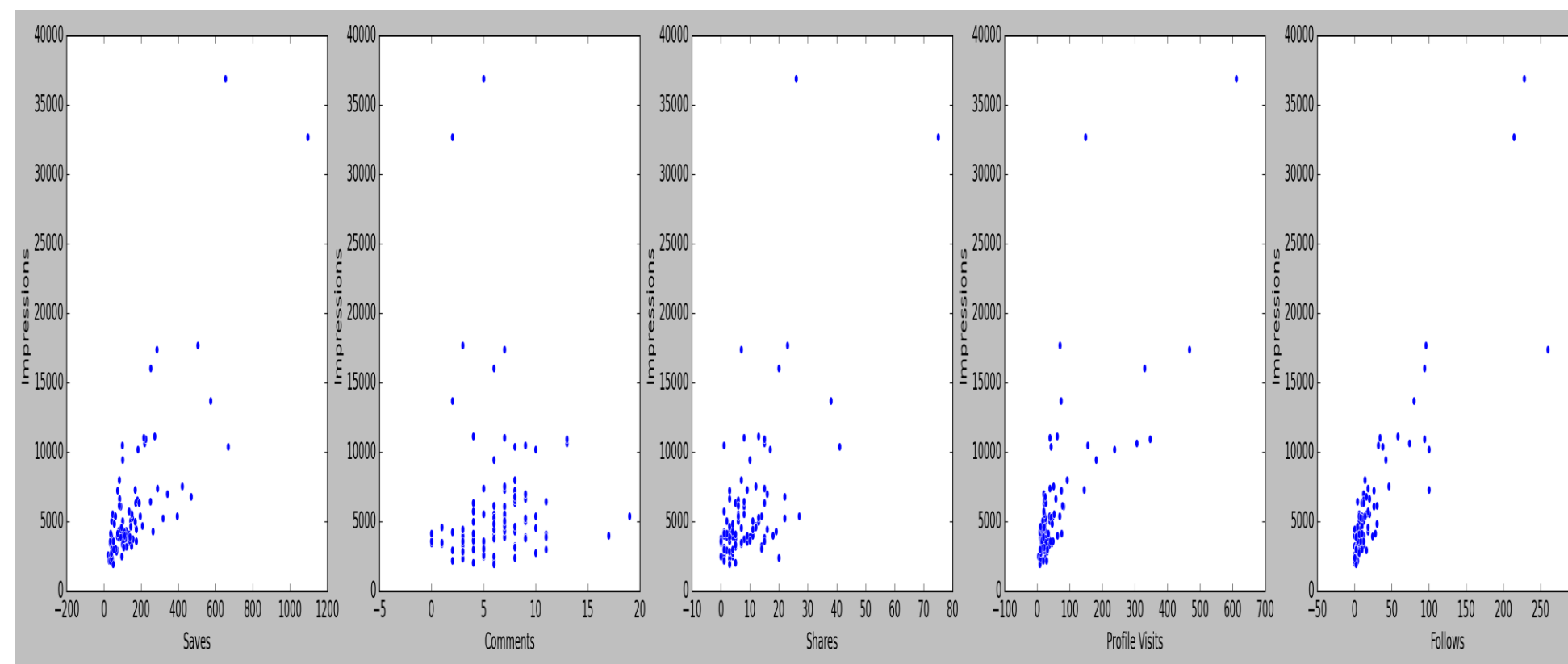
Approach: Assessing Post Reach across Home Pages, Hashtags, Explorer, and Other Platforms for Comparative Analysis.

Observations:



Views on Instagram posts from various sources

By applying linear regression we obtained:



Relation between different interactions and impressions

MAE	0.002195121345446386
RMSE (normalized)	0.00349616768504986
R2_Score_Train	0.9993693795723955
R2_Score_Test	0.9984465023274626

Linear regression errors

Findings:

Views from explore and follows are the most two factors that are best correlated to impressions on ADs.

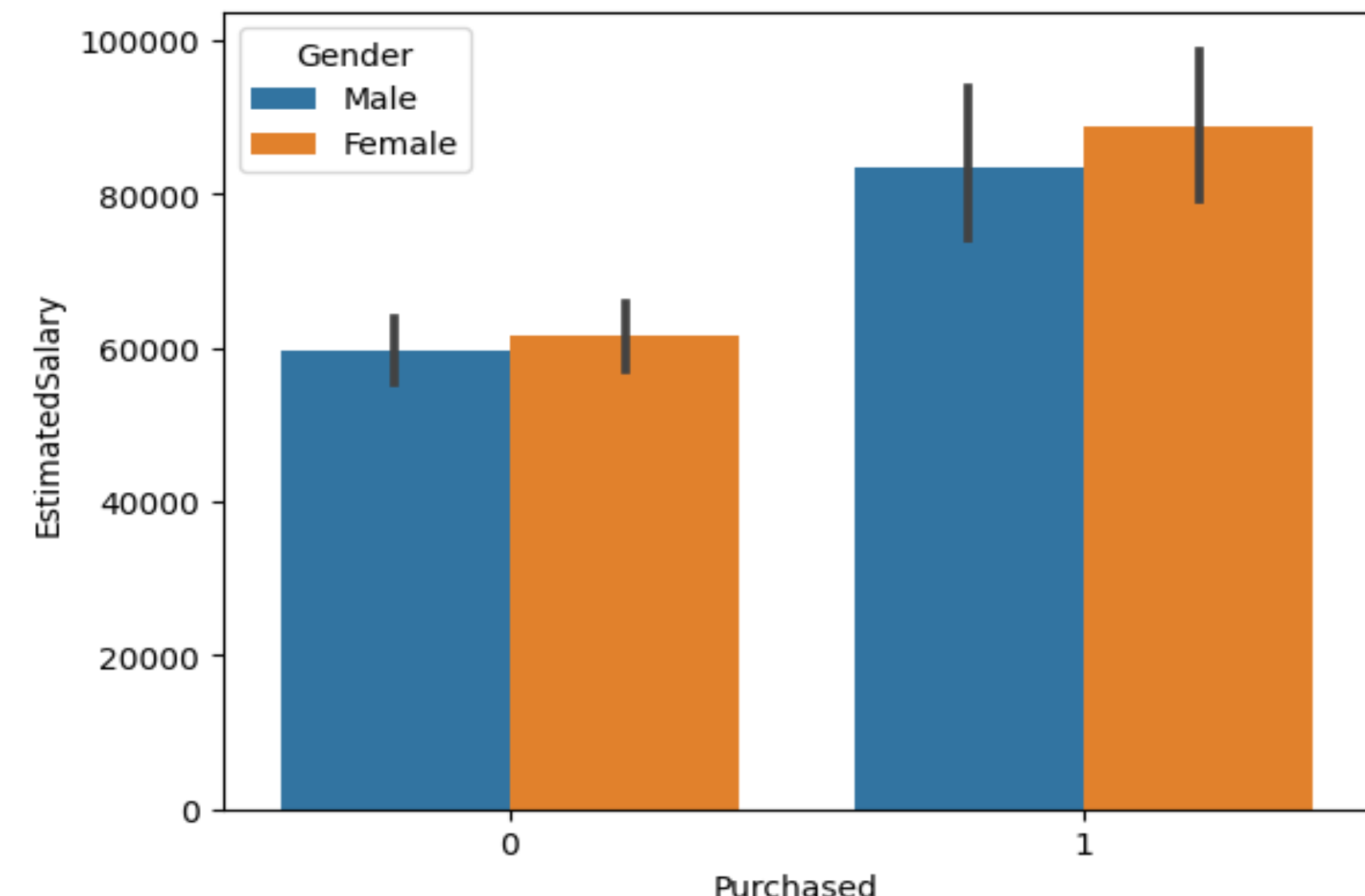
Impact of Ad Targeting on Sales Metrics

Goal: Assessing the Impact of Leveraging Social Media's Pre-stored User Data for Effective Ad Targeting on Sales Performance.

Approach:

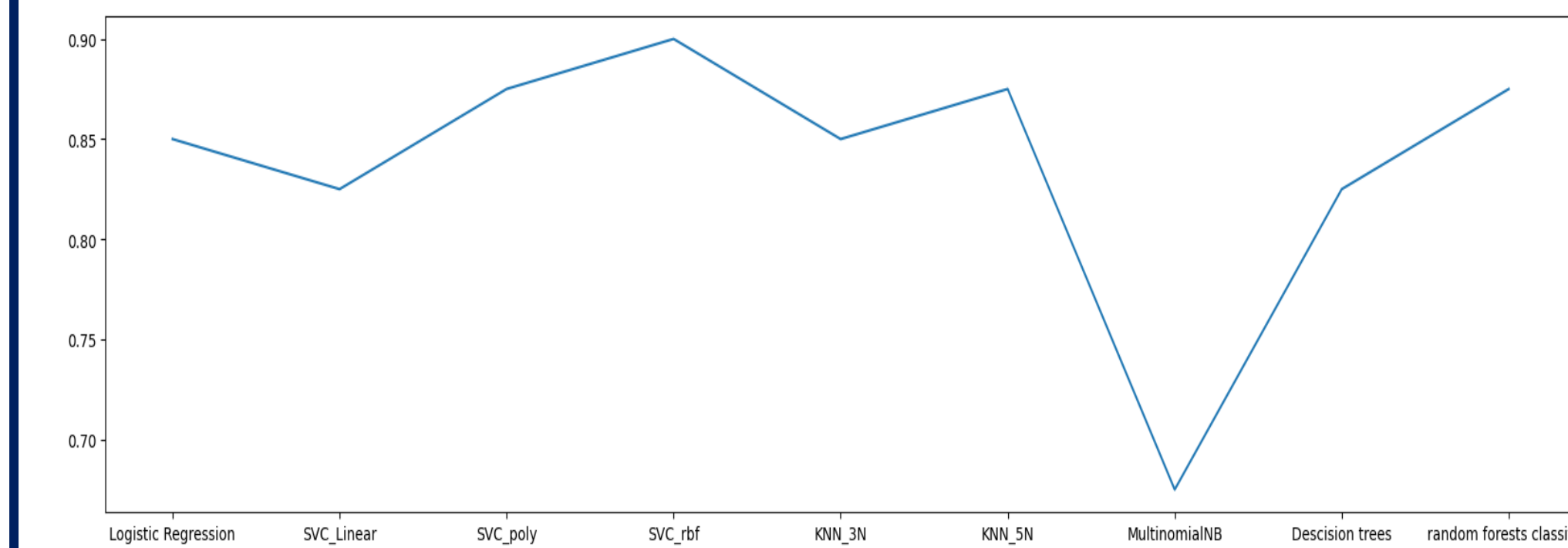
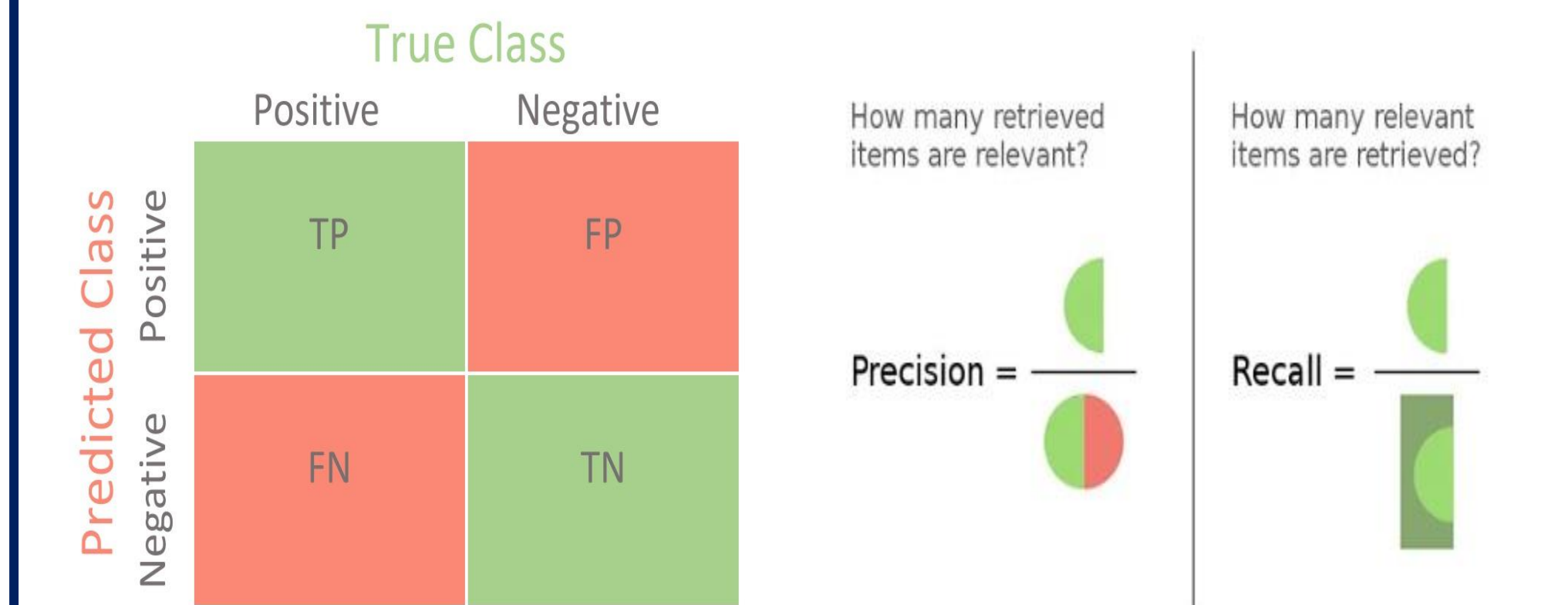
Evaluating the effectiveness of diverse classification methods in predicting outcomes. Employing a comprehensive analysis through confusion matrix and assessing metrics such as accuracy, recall, F1 score, and precision to determine the model with the most accurate predictions.

Observations:



Boxplot between gender and estimated salary after dropping outliers

Accuracy can mislead in imbalanced data. Metrics like precision, recall, and F1-score offer nuanced insights, ensuring a comprehensive evaluation of a model's performance.



effect of different models on the accuracy

Findings:

Support vector classifier (SVC) with RBF kernel is found to be the best model.

Conclusion and Future Work

Conclusion:

Our research underscores social media's advertising dominance, underscores the importance of impressions in impactful Instagram ads, and introduces a predictive model for targeted ad delivery using user data. These insights guide advertisers in the dynamic digital marketing landscape, shaping strategies in channels, content creation, and data-driven ad deployment.

Future Work:

Advanced Analytics: Explore sophisticated techniques like machine learning (neural networks, ensemble methods) for uncovering intricate advertising patterns.

Cross-Platform Analysis: Extend research to diverse advertising platforms, analyzing strategies and user engagement metrics across social media, search engines, and emerging channels.

Ethical Considerations: Address ethical implications of using user data for targeted advertising, ensuring compliance with evolving privacy regulations.

Consumer Behavior Analysis: Deepen understanding of consumer behavior in response to ads, considering emotional responses, brand perceptions, and cultural variables.

Global Market Analysis: Include a global perspective, studying cultural and regional variations in advertising effectiveness and user preferences.

References

- Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow by Aurélien Géron
- An Introduction to Statistical Learning: With Applications in R by Daniela Witten, Gareth M. James, Trevor Hastie, Robert Tibshirani

Authors

1. Bishoy Wadea Fathy
2. Eslam Wageh Samer
3. Ashraf Bahy Galal
4. Peter Safwat Sahyon
5. Sarah Ahdy Kamal