Decision-Centric Intelligent Marketing Analytics

Vaibhavi Umesh Navander
Team Leader

Prof. J.R. Mankar
Guide Name

Janhavi Kishor Patil
Devashri Dattatray Patil
Disha Ashok Kharat

Department of Computer Engineering
K.K. Wagh Institute of Engineering Education and Research,
Nashik, Maharashtra, India

ABSTRACT

The fast growth of digital marketing channels has resulted in large amounts of interaction data about customers that provides both an opportunity and a problem for companies looking to turn such data into valuable marketing intelligence. Even though collecting data has become much easier, making good marketing decisions based on such data has not yet become a trivial task. Standard analytics platforms focus only on descriptive analysis of past campaign results and do not have a capacity to forecast future behavior of customers or suggest appropriate marketing actions. In this paper, we describe the Decision-Centric Intelligent Marketing Analytics System (DECIMA) for fashion e-commerce based on supervised machine learning techniques for predicting customer purchase conversion behavior. Apart from standard prediction models, our DECIMA system is able to estimate the uncertainty of predictions through a risk score and analyze the combined effect of multiple campaigns at once. It will also be equipped with a model monitoring capability that will track model performance degradation and evolution of customer behavior distributions.

Keywords: Digital Marketing, Consumer Behavior Prediction, Machine Learning, Cross-Campaign Analysis, Predictive Analytics, Confidence-Based Risk Scoring, Model Monitoring, Marketing Decision Support System.

I. INTRODUCTION

Fashion e-commerce globally is a rapidly growing segment in digital retail industry owing to higher internet penetration, mobile commerce, and changes in consumer behavior in terms of purchasing. Digital marketing is described as a process of marketing products or services using internet-based methods such as e-commerce, social networks, email, and search engines [1], [4]. The sheer volume of data created by consumer interaction on these media including click-stream data, purchase history, marketing campaign effectiveness, and browsing patterns creates the empirical basis for predicting the effectiveness of the marketing campaign based on data analysis.

Marketing has been revolutionized from an intuitive practice to an informational science. Digital marketing advertisements are employed to target certain market segments, increase engagement rate, and maximize ROI (return on investment). Unfortunately, in spite of having sufficient data, the majority of organizations find themselves unable to utilize data in order to make marketing decisions [2], [5]. The main problem is that marketing effectiveness is controlled by a number of interrelated factors that cannot be analyzed separately, namely: demographic characteristics of customers, buying experience, channel of advertising, product type, seasonality, and rivalry.

Traditional methods of evaluating marketing performance are based on descriptive statistical analysis and judgment, which do not allow one to reflect the non-linear interrelation between such variables as customer engagement, cost of the campaign, rate of conversion and income [6], [14]. Consequently, all actions and decisions are reactive rather than proactive: one studies the results of the previous campaigns in order to conduct the next one, instead of predicting the future behaviour of customers prior to conducting the marketing campaign. It is especially critical for the fashion industry due to its short cycle of trends, large variety of products and limited time period, during which the campaign can be delivered.

In the realm of fashion e-commerce, however, the problem is further aggravated by the heterogeneity of customers. A Luxury Seeker customer behaves in a totally different manner when buying products than an Essentials Buyer, and trying to target both customers with the same campaign always results in failure for both [17], [19].

Corporate-scale artificial intelligence technologies like Vue.ai and Zalando's recommendation engine take care of the heterogeneity problem, but need substantial investment and are not affordable for small and medium sized (SMB) fashion companies. Thus, there exists a market structure where small and medium sized fashion companies lack access to AI-based personalization tools, unlike their bigger competitors.

Predictive analytics and machine learning present a way out of this dilemma. With the help of classification models trained on historical customer and campaign information, predictions can be made concerning how probable is the conversion of any customer for a particular campaign, the confidence factor, and the risks associated. But existing marketing systems that use ML do not go beyond making such predictions and do not give the marketing directions [3], [9]. Predictions that say the probability is 0.72 are not enough to act on.

In this paper, we introduce DECIMA, the Decision-Centric Intelligent Marketing Analytics System, for fashion E-commerce. DECIMA enhances the conventional ML pipeline with three levels of downstream intelligence, namely, a risk-scoring mechanism for prediction uncertainty, a dual-level recommendation engine that interprets predictions to suggest fashion-centric marketing decisions on customer segments and portfolios of campaigns, and a Weighted Sum Model (WSM)-based ranking system for evaluating the relative quality of all campaigns in terms of four performance criteria.

The system relies solely on open-source Python packages (scikit-learn, Flask, pandas, Postgres) and ingests data through a CSV file upload process, thus being deployable without the need for cloud services or data science knowledge. Such approach was chosen deliberately, allowing SMB fashion retailers to have access to decision-oriented marketing analytics previously available only to enterprises.

The rest of the thesis is organized as follows. Section II highlights literature review and research gaps. Section III discusses system methodology, mathematical modeling, and algorithms. Section IV gives details of the experimental setup. Section V discusses the results. Section VI concludes the thesis, and Section VII identifies future directions for research.

II. Review of Literature

Machine learning techniques have been applied in the field of marketing analytics and customer behavior prediction for more than three decades and involve a wide range of techniques, application areas, and evaluation approaches. The literature survey can be grouped into four clusters: customer behavior prediction, campaign evaluation and optimization, decision support systems, and domain-specific fashion analytics. Prediction of Customer Behaviour: The usefulness of data mining in business marketing has been confirmed by [1], proving that classification methods such as decision trees and logistic regression are the most popular tools used for customer conversion analysis. This result was later confirmed by [3], who investigated various machine learning techniques for market segmentation and showed that logistic regression is the most interpretable baseline and random forest is the most powerful ensemble technique for structured tabular data in marketing.

Combining RFM (Recency, Frequency, Monetary) analysis with classification algorithms was investigated in [6], where a substantial boost in repurchase prediction accuracy of customers was achieved by the combination of behavioral and demographic factors. An adaptive analytics methodology for customer retention based on ensemble learning and continuous retraining in case of concept drift has been developed in [7]. Big data techniques for predicting marketing models have been introduced by [2], [19], where the multidimensional approach to customer prediction significantly outperforms descriptive statistics with only one metric.

Campaign Evaluation and Optimisation: Machine learning algorithms for predicting ROI and campaign budget optimisation were suggested by [8], thus offering a methodological precursor to the WSM-based ranking of campaigns in DECIMA. Analysis of multi-channel marketing using anomaly detection has been studied by [9], [18], showing that performance patterns specific to each channel can be found with the help of clustering and statistical methods of finding outliers.

Multi-channel attribution modeling to enhance ROI in emerging markets was examined by [15], which showed that single-channel attribution is prone to significant underestimation of the share of organic and referral channels to the conversion process. Selection of optimal target segments for managing marketing campaigns based on effectiveness-driven models has been examined in [12], thus laying conceptual grounds for the segment-level recommendation engine in DECIMA. Customer Lifetime Value (CLV) prediction using probabilistic machine learning models was suggested by [16]. Applications of AI in Digital Marketing: The broader domain of optimization of marketing strategy based on AI technology was studied in [10], [17], wherein it was noted that elements of AI, such as natural language processing, computer vision and predictive analysis, have been implemented in recommendation systems, content personalization, and dynamic pricing.

Data clustering-based user segmentation approach in the e-commerce context was described in [11] in order to give the background for the persona segmentation (Luxury Seeker, Trendsetter, Seasonal Shopper, Essentials Buyer) implemented in this study. Frameworks of intelligent, data-driven precision marketing for enterprise use were studied in [19], and the issues of predictive analysis in e-commerce personalization were explored in [14]. The problem of concept drift detection based on digital twin-based anomaly detection techniques related to the model monitoring task of DECIMA was considered in [13].

Despite the large amount of research done in this area, four major gaps are still left untouched. First, current machine learning-based marketing systems concentrate only on the accuracy of predictions and lack a decision component. A model that gives a probability score needs an expert in the field to make sense of it, after which only can the marketing action be performed; non-tech-savvy marketing professionals cannot take actions based on a probability score [2], [5], [7]. Second, all existing approaches for comparing campaigns use single criterion evaluation, which is usually either ROI or conversion rate; however, real campaign portfolios are multidimensional objects [8], [12], [15].

Third, there is no existing approach in fashion retailing analytics that explicitly quantifies the prediction uncertainty

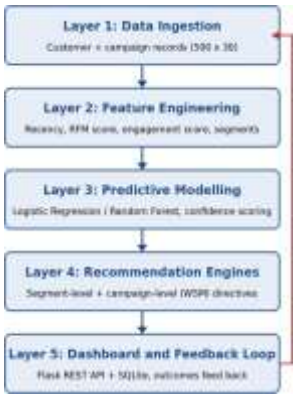


Fig. 1: *DCIMAS system architecture — five-layer pipeline from data ingestion to dashboard.*

and exposes it to the user as part of the output; confidence and risk are implicitly embedded in probabilities [3], [9], [13].

Finally, AI tools for fashion retailing are enterprise-level and out of reach for SMBs [17], [19].

The main technical challenge in this area is class imbalance: in actual fashion e-commerce data sets, conversion rates are generally between 15% and 35%, resulting in data sets where there are more non-conversions than conversions. Accuracy is a poor measure when class imbalance exists: a simple algorithm which always predicts non-conversion has excellent accuracy but does nothing.

The second challenge is heterogeneous features: demographic information (age, income) as well as behavioral data (recency, frequency, spend) and campaign data (channel, type, cost) exist in very different scales. The third challenge is generalization: an algorithm must continue to be predictive despite changes in customer behavior and campaigns.

Previous research has shown that ML-based marketing analytics is possible and valuable; however, these studies have not considered the entire process chain from raw data to actionable marketing directives.

The importance of DECIMA is due to its integration: prediction, uncertainty quantification, recommendation, and cross-campaign ranking have all been integrated into one framework with a shared data layer. The limitations of previous research are exactly what drive each component of our solution.

III. METHODOLOGY

A. Block Diagram and System Description

Fig. 1 illustrates the overall architecture of the DECIMA system. The system functions as a five-layer pipeline consisting of the following components: (1) Data Input Layer, which supports CSV uploads of customer and campaign data; (2) Preprocessing and Feature Engineering Layer; (3) Predictive Modelling and Scoring Layer; (4) Cross-Campaign Analysis and Recommendation Layer; and (5) Web Dashboard Layer. Each layer communicates with its neighboring layers through a Flask REST API.

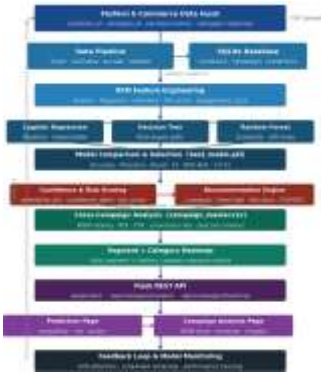


Fig. 2: *Workflow of DECIMA*

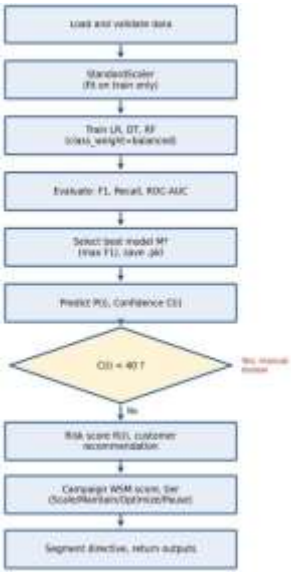


Fig. 3: *DCIMAS algorithm flowchart showing prediction, confidence routing, and recommendation logic*

Fig. 3 presents the comprehensive flowchart of the DECIMA algorithm, illustrating the different decision-making stages involved in the prediction and recommendation process.

B. Mathematical Model

The mathematical formulations below describe the principal computations performed by DECIMA. Each equation is numbered for reference in the algorithm.

1) Feature Engineering Using RFM: The engagement score of customer i is calculated as

$$E_i = \frac{30}{R_i + 1} + 40\frac{F_i}{F_{\max}} + 30c_i \qquad (1)$$

where R_i is the number of days since the customer's last purchase, F_i is purchase frequency, $F_{\max}$ is the maximum purchase frequency in the dataset, and $c_i \in \{0, 1\}$ represents the campaign response.

The RFM score is defined as

$$\text{RFM}_i = \frac{r_i + f_i + m_i}{3} \quad (2)$$

where r_i, f_i, and m_i are the recency, frequency, and monetary decile scores, respectively, each ranging from 1 to 5.

2) Logistic Regression Model: The probability of conversion for customer i is estimated as

$$P_i = \sigma(z_i) = \frac{1}{1 + \exp(-z_i)} \quad (3)$$

where

$$z_i = \beta_0 + \sum_{j=1}^{n} \beta_j x_{ij} \quad (4)$$

Here, x_{ij} is the value of feature j for customer i, β_j is its learned coefficient, and n is the number of features.

The model is trained by minimizing the regularized binary cross-entropy loss:

$$L(\boldsymbol{\beta}) = -\sum_{i=1}^{m} [y_i \log P_i + (1 - y_i) \log(1 - P_i)] + \lambda \|\boldsymbol{\beta}\|_2^2 \quad (5)$$

where $y_i \in \{0, 1\}$ is the actual conversion label, m is the number of training instances, and λ is the L2 regularization coefficient.

3) Confidence Score: The prediction confidence is defined as the scaled distance of the predicted probability from the decision boundary:

$$C_i = |P_i - 0.5| \times 200 \quad (6)$$

The confidence score ranges from 0 to 100. A score of 100 indicates maximum confidence under this definition, whereas a score of 0 indicates maximum ambiguity at $P_i = 0.5$.

The confidence label is assigned as follows: High if $C_i \geq 70$, Medium if $40 \leq C_i < 70$, and Low if $C_i < 40$.

4) Risk Score: The campaign investment risk score for customer i is

$$R_i = (1 - P_i) \times 100 + \delta_R + \delta_M - \delta_c \quad (7)$$

where δ_R is the recency penalty, δ_M is the RFM penalty, and δ_c is the campaign-response bonus. The RFM penalty is computed as

$$\delta_M = \max(0, (3.0 - \text{RFM}_i) \times 4) \quad (8)$$

The response bonus is $\delta_c = 8$ when the customer has responded to the campaign and zero otherwise. The recency penalty δ_R increases with the number of days since the last purchase. The resulting risk score is clipped to the interval $[0, 100]$.

The risk category is Low Risk for $R_i \leq 30$, Medium Risk for $30 < R_i \leq 70$, and High Risk for $R_i > 70$.

5) Weighted Sum Model: The composite score of campaign j is calculated using the Weighted Sum Model (WSM):

$$S_j = w_1 ROI'_j + w_2 CR'_j + w_3 CTR'_j + w_4 CpC'_j \quad (9)$$

where the criterion weights are $w_1 = 0.40$, $w_2 = 0.30$, $w_3 = 0.15$, and $w_4 = 0.15$. Here, ROI'_j, CR'_j, CTR'_j, and CpC'_j are normalized campaign metrics.

For benefit criteria such as ROI, conversion rate (CR), and click-through rate (CTR), min-max normalization is used:

$$x'_j = \frac{x_j - x_{\min}}{x_{\max} - x_{\min}} \quad (10)$$

For the cost criterion, cost per conversion (CpC), the normalization is inverted:

$$CpC'_j = \frac{CpC_{\max} - CpC_j}{CpC_{\max} - CpC_{\min}} \quad (11)$$

The return on investment for campaign j is

$$ROI_j = \frac{Revenue_j - Budget_j}{Budget_j} \quad (12)$$

where

$$Revenue_j = Conversions_j \times AvgOrderValue_j \quad (13)$$

6) Confidence-Weighted Segment Probability: For each style-segment s and fashion-category c combination, the confidence-weighted probability is

$$W_{s,c} = \bar{P}_{s,c} \left(\frac{\bar{C}_{s,c}}{100}\right) \quad (14)$$

where $\bar{P}_{s,c}$ and $\bar{C}_{s,c}$ are the mean predicted conversion probability and mean confidence score, respectively, for customers belonging to segment s and category c.

C. *Algorithm*

Algorithm presents the complete DCIMAS prediction and recommendation pipeline. The input consists of the preprocessed customer feature matrix $X \in \mathbb{R}^{m \times n}$ and the master campaign dataset containing k campaigns.

Algorithm 1 *DECIMA Prediction and Recommendation Pipeline*

Require: Customer feature matrix X and campaign dataset *Campaign_Master*
Ensure: Prediction records, campaign tiers, and heatmap W
Validate X and handle missing critical values
Split data into training (80%) and testing (20%) sets
Fit StandardScaler on training data only
Transform training and testing data
Train LR, DT, and RF using balanced class weights
Evaluate models using F1-score, recall, and ROC-AUC
Select best model $M^* = \arg\max_M F1(M)$
Save M^* as `best_model.pkl`
for each customer $i = 1, \ldots, m$ **do**
    Calculate prediction probability P_i
    Calculate confidence $C_i = |P_i - 0.5| \times 200$
    if $C_i < 40$ **then**
        Flag customer for manual review
    else
        Calculate risk score R_i
        Assign risk category
        Generate customer recommendation
        Store customer prediction record
    end if
end for
for each campaign $j = 1, \ldots, k$ **do**
    Normalize ROI, CR, CTR, and CpC
    Calculate campaign score
    $S_j = 0.40ROI'_j + 0.30CR'_j + 0.15CTR_j + 0.15CpC'_j$
    if $S_j \geq 0.65$ **then**
        Assign tier: Scale
    else if $S_j \geq 0.45$ **then**
        Assign tier: Maintain
    else if $S_j \geq 0.25$ **then**
        Assign tier: Optimize
    else
        Assign tier: Pause
    end if
    Generate campaign directive using ROI and CR
end for
Compute segment-category heatmap $W_{s,c}$
Assign segment directives using probability and confidence
Output prediction records, campaign tiers, and heatmap

IV. Experimental Setup

A. Description of the Dataset

The experimental dataset was designed to simulate a realistic fashion e-commerce retail scenario and consists of two connected datasets. The customer dataset contains 500 customer profiles with 30 features, while the campaign master dataset contains 50 campaigns (FCMP001–FCMP050), each with 20 features, including the Weighted Sum Model (WSM) metrics.

The datasets are connected through the common `campaign_id` attribute, resulting in 500 customer–campaign interaction records. Each record is labeled using `conversion_status`, where 1 indicates a purchase and 0 indicates no purchase. The dataset contains 142 positive conversion records and 358 negative records, corresponding to a positive-class ratio of 28.4%.

The customer interactions span five marketing channels: Instagram (26%), Facebook Ads (21%), Email (21%), Influencer Marketing (19%), and Google Ads (13%). The product categories include Western Wear (25%), Ethnic Wear (24%), Footwear (21%), Accessories (15%), and Winter Wear (14%). Based on spending behavior and shopping patterns, customers are categorized into four segments: Seasonal Shopper (38%), Trendsetter (25%), Essentials Buyer (24%), and Luxury Seeker (13%).

B. Feature Set

The model training process uses 20 input features selected from four feature categories: customer demographic and profile attributes, behavioral and engagement attributes, campaign-related attributes, and purchase-history attributes. Categorical features are encoded using `LabelEncoder`, while numerical features are standardized using `StandardScaler`. To prevent data leakage, the en-

TABLE I: *Complete Feature Set for Predictive Modelling*

Feature	Group	Type	Description
age	Demographics	Integer	Customer age in years (18–65)
income	Demographics	Integer	Annual income in INR
gender	Demographics	Encoded	Male/Female/Other → {0,1,2}
education	Demographics	Encoded	High School → PhD → {0,1,2,3}
marital_status	Demographics	Encoded	Single/Married/Divorced → {0,1,2}
fashion category	Behaviour	Encoded	5 categories → {0..4}
style segment	Behaviour	Encoded	4 personas → {0..3}
purchase_frequency	Behaviour	Integer	Number of purchases in period
avg_order_value	Behaviour	Float	Mean spend per order (INR)
total spend	Behaviour	Float	Lifetime spend (INR)
website visits	Behaviour	Integer	Monthly visits to platform
campaign type	Campaign	Encoded	5 types → {0..4}
channel	Campaign	Encoded	5 channels → {0..4}
campaign_response	Campaign	Binary	Responded to campaign: 1/0
recency	RFM	Integer	Days since last purchase
frequency	RFM	Integer	Same as purchase frequency
monetary	RFM	Float	Same as total spend
rfm_score	RFM	Float	Mean of R,F,M ordinal scores [Eq.2]
engagement_score	RFM	Float	Composite 0–100 score [Eq.1]
customer_segment	RFM	Encoded	High/Mid/Low spend band → {0,1,2}

coders and scaler are fitted using the training data only and subsequently applied to the test data. The fitted preprocessing objects are saved using Joblib to support consistent inference on new customer records.

C. *Performance Measures*

Five conventional classification metrics are used to evaluate the models on the test dataset, which contains 100 customer records. These metrics capture different aspects of model performance, particularly under class imbalance.

1) Accuracy: Accuracy measures the proportion of correctly classified records and is defined as

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

where TP , TN , FP , and FN denote true positives, true negatives, false positives, and false negatives, respectively. Accuracy can be misleading for imbalanced datasets because a classifier that always predicts non-conversion would achieve 71.6% accuracy on the overall dataset while failing to identify conversions.

2) Precision: Precision measures the proportion of predicted conversions that correspond to actual conversions:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{16}$$

In marketing applications, low precision may result in campaign expenditure being directed toward customers who do not convert.

3) Recall: Recall, also known as sensitivity, measures the proportion of actual conversions correctly identified by the model:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{17}$$

Low recall may indicate missed conversion opportunities and potential revenue loss.

4) F1-score: The F1-score is the harmonic mean of precision and recall:

$$F_1 = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{18}$$

The F1-score is used as the primary model-selection metric because it balances precision and recall. Model selection is performed using stratified five-fold cross-validation on the training data.

5) ROC-AUC: The Receiver Operating Characteristic–Area Under the Curve (ROC-AUC) measures the model's ability to distinguish between conversion and non-conversion across classification thresholds. An AUC of 0.5 indicates performance equivalent to random ranking, whereas an AUC of 1.0 indicates perfect discrimination. ROC-AUC complements the F1-score by evaluating ranking performance independently of a single classification threshold.

For campaign evaluation, four metrics are incorporated into the Weighted Sum Model: Return on Investment (ROI), with a weight of 0.40; Conversion Rate (CR), with a weight of 0.30; Click-Through Rate (CTR), with a weight of 0.15; and Cost per Conversion (CpC), with a weight of 0.15. These weights represent the relative importance assigned to each metric in campaign evaluation.

D. *Software Requirements*

The software requirements used to implement the proposed DCIMAS framework are summarised in Table II.

TABLE II: *Software Requirements*

Component	Tool / Version	Purpose
Language	Python 3.10+	Core implementation
ML Library	scikit-learn 1.3+	LR, DT, RF, preprocess-ing, and metrics
Data Handling	pandas 2.0+, NumPy 1.24+	Data manipulation and feature engineering
Visualization	Matplotlib 3.7+, Seaborn 0.12+	Charts and plots
Web Framework	Flask 3.0+	REST API and server
Database	Postgres 18.6	Customer and prediction storage
Model Storage	Joblib 1.3+	Saving the trained model and scaler
IDE	Jupyter Notebook 7.x	Development and analysis
Environment	Windows/Linux/macOS	Cross-platform support

V. Results and Discussion

A. *Model Comparison Results*

Table III presents the evaluation metrics for the three models on the held-out test set (n = 100). All models were evaluated using the same stratified 80/20 train-test split, random seed (42), and `class_weight='balanced'` configuration.

TABLE III: *Model Evaluation Results on Test Set (n = 100)*

Model	Accuracy	Precision	Recall	F1	ROC-AUC
Logistic Regression	0.61	0.55	0.82	0.54	0.69
Decision Tree	0.68	0.48	0.52	0.48	0.69
Random Forest	0.69	0.49	0.50	0.49	0.68

Logistic Regression achieved the highest F1-score (0.54) and recall (0.82), indicating its ability to identify a large proportion of actual purchases. Random Forest achieved the highest accuracy (0.69). However, accuracy alone may be misleading because non-conversions represent 71.6% of the overall dataset. A model that predominantly predicts the majority class may achieve relatively high accuracy while failing to identify potential buyers.

The recall of 0.82 indicates that Logistic Regression identified approximately 82% of the actual conversions in the test set. Given the reported class distribution, the test set contains approximately 28–29 positive records; the exact number should be verified against the test labels. Missed conversions represent potential revenue opportunities, while false positives may result in campaign expenditure on customers who do not purchase. This trade-off supports the use of recall and F1-score alongside accuracy.

The five-fold stratified cross-validation F1-scores further support the observed model ranking. Logistic Regression achieved 0.56, compared with 0.48 for Random Forest and 0.44 for Decision Tree. These results suggest that Logistic Regression performed consistently better under the selected evaluation procedure. The ROC curves for all three models are presented in Fig. 4.

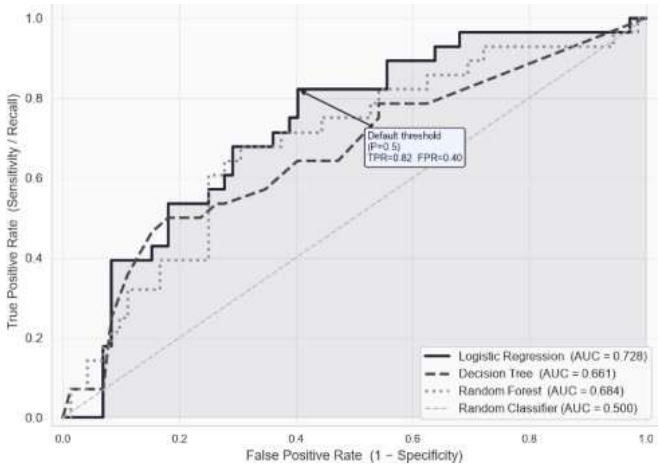


Fig. 4: *ROC curves for all three classifiers. Logistic Regression and Decision Tree achieve equal AUC (0.69) but Logistic Regression dominates on recall*

B. *WSM Campaign Ranking Results*

The Weighted Sum Model (WSM) composite scores were computed for all 50 campaigns using the normalization and scoring equations described in Section III. Fig. 5 presents the ten highest-ranked campaigns, while Fig. 6 illustrates the score distribution across the complete campaign portfolio.

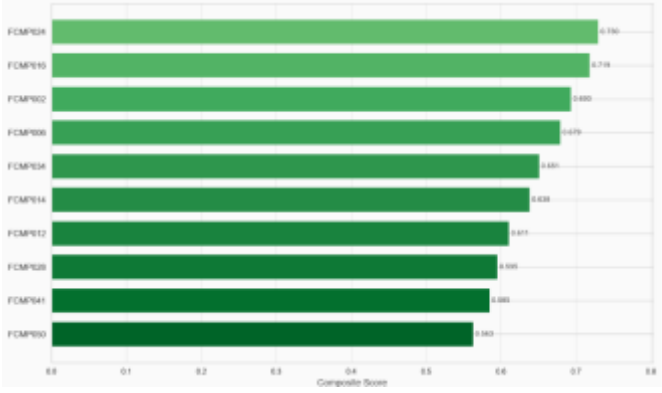


Fig. 5: *Top 10 campaigns by WSM composite score*

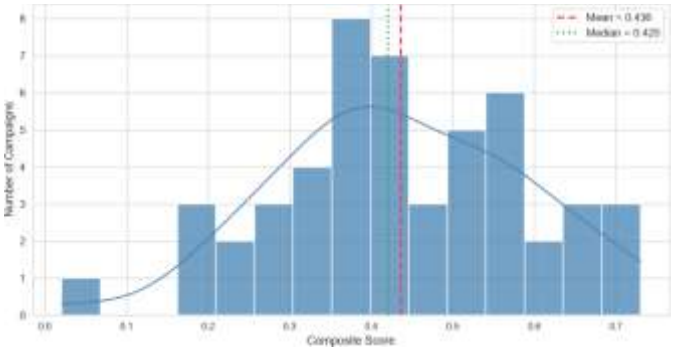


Fig. 6: *Distribution of composite scores across all 50 campaigns (Min=0.021, Max=0.730, Mean=0.436). Right-skew indicates a small elite of high-performing campaigns*

FCMP024, associated with Instagram, Ethnic Wear, and the festive season, achieved the highest composite score of 0.730. It was followed by FCMP016, associated with Influencer Marketing, at 0.720, and FCMP002, associated with Facebook Ads, at 0.700.

The score distribution was right-skewed, with a reported mean of 0.436 and median of 0.420. This suggests that a relatively small subset of campaigns achieved higher scores than the majority of campaigns. Instagram and Influencer Marketing recorded the highest channel-specific scores in the evaluated portfolio. FCMP041, the highest-ranked Google Ads campaign, achieved a score of 0.590. These findings describe the observed dataset and should not be interpreted as proof that one marketing channel will universally outperform another.

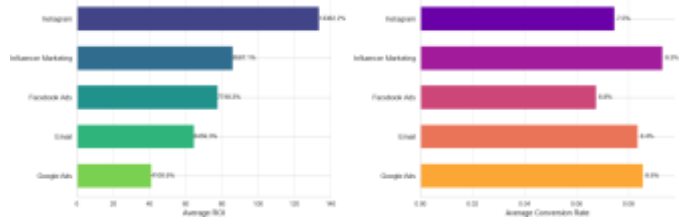


Fig. 7: *Average ROI (left) and average conversion rate (right) by marketing channel. Instagram leads both metrics*

C. *Campaign Recommendation Engine Findings*

Table IV presents the portfolio-level distribution of recommendation tiers generated for all 50 campaigns.

TABLE IV: Campaign Recommendation Tiers

Tier	Campaigns (*n*)	WSM Score Range	Campaign-Level Directive
Scale	5 (10%)	0.65–0.730	Increase budget — portfolio leaders
Maintain	18 (36%)	0.45–0.649	Hold spend, re-check next cycle
Optimize	20 (40%)	0.25–0.449	Revise creative or tighten targeting
Pause	7 (14%)	0.021–0.249	Reallocate budget to Scale campaigns

The Scale tier contained 10% of campaigns (n = 5), indicating that these campaigns were identified as candidates for increased budget allocation. The Optimize tier was the largest group, containing 40% (n = 20), suggesting that

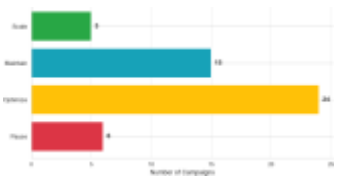


Fig. 8: *Campaign portfolio tier distribution.Scale tier represents 10tier 14*

many campaigns may benefit from changes to creative content, targeting, or allocation. The Pause tier accounted for 14% ($n = 7$), identifying campaigns that received low composite scores under the proposed scoring framework.

Examples of the generated campaign directives are as follows:

- **FCMP024:** "Scale budget on Instagram — ROI 148% and conversion 9.2% are among the best in the portfolio."
- **Lowest-performing campaign:** "Pause or reallocate budget—poor performance (ROI 2.1%, conversion 1.8%)."

These concise directives translate campaign metrics into action-oriented recommendations, helping marketing managers interpret the results and identify possible budget adjustments.

D. Segment-Level Recommendation Results

The segment-level recommendation engine evaluated all 20 combinations of customer style segment and fashion category using the weighted probability defined in Section III.

Luxury Seeker–Winter Wear, with a weighted probability of 64.0%, and Trendsetter–Footwear, with a weighted probability of 62.4%, received the "Safe to Scale" directive. Essentials Buyer–Ethnic Wear and Seasonal Shopper–Western Wear received "Revise or Pause" directives, with reported weighted probabilities below 35%.

The heatmap indicates that Winter Wear and Footwear were the strongest categories for Luxury Seekers and Trendsetters, respectively, with reported values of 80%. These findings can help prioritize segment-specific campaign experiments and budget allocation. However, the recommendations should be validated through controlled campaign testing before substantial budget changes are made.

VI. Conclusion

In this paper, a Decision-Centric Intelligent Marketing Analytics System (DCIMAS) for fashion e-commerce has been presented. The proposed system addresses three important research gaps: the gap between descriptive machine learning predictions and prescriptive decision-making, the lack of uncertainty quantification for marketing managers, and the limited accessibility of AI-driven decision-support systems for small and medium-sized fashion businesses.

The comparison of three supervised learning models on a fashion e-commerce dataset containing 500 customers with a positive-class ratio of 28.4% demonstrated that Logistic Regression achieved the highest F1-score (0.54) and recall (0.82). These results highlight the importance of selecting evaluation metrics appropriate for imbalanced binary classification. For the evaluated dataset, the interpretable linear model outperformed the more complex ensemble approach in terms of F1-score and recall. Selecting a model based solely on accuracy would have favored Random Forest, potentially overlooking the stronger conversion-identification performance of Logistic Regression.

The proposed confidence-based risk-scoring approach introduces uncertainty quantification into the marketing decision-making process. The confidence score, $C_i = |P_i - 0.5| \times 200$, and the composite risk score R_i, which incorporates recency, RFM, and campaign response, provide actionable information for customer targeting. According to the experimental results, 24.2% of customers were classified as having low confidence and were flagged for manual review. This mechanism can help reduce the risk of making campaign decisions based on uncertain predictions near the model's decision boundary.

A key contribution of DCIMAS is its dual-layer recommendation approach. At the customer level, the recommendation engine generates fashion-oriented actions, including Premium Lookbook, Flash Sale, Win-back Offer, Discount Coupon, and Social Awareness Post, according to customer segments and recency. At the campaign level, the system assigns one of four operational tiers—Scale, Maintain, Optimize, or Pause—to each of the 50 campaigns based on its Weighted Sum Model (WSM) composite score. It also generates concise, human-readable recommendations incorporating campaign-specific ROI and conversion rate.

The WSM-based campaign-ranking mechanism enables multicriteria evaluation using return on investment (ROI), conversion rate (CR), click-through rate (CTR), and cost per conversion (CpC). The evaluation identified five campaigns (10%) in the Scale tier and seven campaigns (14%) in the Pause tier, providing a basis for prioritizing budget allocation. The right-skewed distribution of campaign scores further indicates that a relatively small subset of campaigns achieved higher composite scores than the majority of the evaluated portfolio.

Finally, the proposed open-source implementation uses Python, Flask, scikit-learn, and SQLite to provide an accessible solution for fashion small and medium-sized businesses. Its CSV-based data workflow and portable file handling through Python's `pathlib` reduce deployment complexity and eliminate the need for cloud infrastructure. Overall, DCIMAS integrates predictive analytics, uncertainty-aware customer assessment, campaign ranking, and actionable recommendations into a unified decision-support framework for fashion e-commerce marketing.

VII. Future Scope

Several future research directions can be explored to address the limitations of the current DCIMAS framework and improve its predictive and decision-making capabilities.

Deep Learning for Nonlinear Prediction:The current system primarily relies on Logistic Regression for customer

conversion prediction. Future work can explore deep learning architectures such as TabNet and Multilayer Perceptrons (MLPs) to capture nonlinear relationships among customer behavior, engagement, and campaign features. Evaluating these models on larger production-scale datasets containing 50,000 or more customer records may help determine whether they improve recall and F1-score compared with conventional machine learning models. However, such improvements must be validated experimentally.

Reinforcement Learning for Recommendation: The customer-level and campaign-level recommendation engines currently rely on predefined rules and thresholds. Future work can investigate reinforcement learning to learn adaptive marketing policies from historical campaign interactions. In this framework, the agent observes the customer or campaign state, selects a marketing action, and receives a reward based on outcomes such as conversions, revenue, or campaign costs. A carefully designed reward function could enable the recommendation engine to adapt to changing customer preferences and campaign performance with less manual rule adjustment. Offline evaluation and controlled experiments would be required before deploying learned policies in live campaigns.

Insight Narration Using Generative AI:Integrating a Large Language Model (LLM) could provide a natural-language insight narration layer for DCIMAS. The model could receive structured information, including the highest-ranked campaign, the best-performing marketing channel, and segment-level heatmap results, and generate concise summaries for marketing managers. Such summaries could explain campaign rankings, highlight promising customer segments, and describe potential budget reallocation opportunities. To improve reliability, the narration layer should use verified system outputs and avoid generating unsupported performance claims.

Real-Time Streaming and Concept Drift:The current architecture processes static CSV datasets in batches. Future development could integrate Apache Kafka to process customer interaction events in real time. Incremental learning and concept-drift detection methods, including approaches related to the work in [13], could help identify changes in customer behavior and campaign performance over time. This would enable DCIMAS to update predictions as new interactions occur and support more responsive marketing recommendations. Further work would be needed to address event processing, model monitoring, data quality, and the computational requirements of a continuous deployment environment.

REFERENCES

[1] "Research on the Application of Data Mining Technology in Enterprise Marketing," ACM, 2024.
[2] "The Role of Big Data in Modern Marketing Decisions," IEEE, 2024.
[3] "Machine Learning-Based Market Segmentation and Consumer Behavior Prediction Models," ACM, 2024.
[4] "Data Mining and Intelligent Decision-Making in Digital Accounting Information Management Systems," ACM, 2024.
[5] "Machine Learning-Driven Consumer Behavior Analysis: An AI Approach for Intelligent Marketing," IEEE, 2025.
[6] "Enhancing Customer Repurchase Prediction: Integrating Classification Algorithms with RFM Analysis for Precision and Actionable Insights," ScienceDirect, 2025.
[7] "An Adaptive Analytics Framework for Customer Retention Through Integrative Feature Optimization and Ensemble Learning," ScienceDirect, 2025.
[8] "An Effective Machine Learning Model for Accurate ROI Prediction and Optimization in Marketing Strategies," IEEE, 2025.
[9] "Intelligent Analysis Methods for Multi-channel Marketing Data Based on Anomaly Detection Algorithms," ACM, 2025.
[10] "AI-Driven Marketing Strategy Optimization in the Digital Economy: A Machine Learning Approach," ACM, 2025.
[11] "Data-Driven Clustering Analysis in E-Commerce User Segmentation," 2025.
[12] "An Optimal Effectiveness-Driven Target Segment Selection Modeling Approach for Marketing Campaign Management," 2025.
[13] "Digital Twin-Based Anomaly Detection Under Concept Drift," 2025.
[14] "The Challenges and Opportunities of Implementing Predictive Analytics in Marketing Strategies and E-Commerce Personalisation," 2025.
[15] "Advances in Multi-Channel Attribution Modeling for Enhancing Marketing ROI in Emerging Economies," 2025.
[16] "A Data-Driven Approach to Customer Lifetime Value Prediction Using Probability and Machine Learning Models," 2025.
[17] "Research on the Application of Artificial Intelligence in Enterprise Digital Marketing," 2025.
[18] "SABIG Data-Driven Predictive Modeling for Enhancing BI and Strategic Decision-Making," 2026.
[19] "Intelligent Data-Driven Precision Marketing," 2026.