AI-Driven Regional Marketing Optimization for the Restaurant Industry

Jingru Chen

Shenzhen Middle School, Shenzhen, Guangdong, China

cjrxiao1120080201@163.com

Keywords: Artificial Intelligence, Restaurant Marketing, Regional Adaptation, Sentiment Analysis, ROI Optimization

Abstract: In this study, we propose a comprehensive AI-driven framework to optimize regional marketing strategies in the restaurant industry, addressing the challenges of geographic taste variation, budget constraints, and campaign personalization. Our approach integrates transformer-based sentiment classification, LDA-based regional flavor modeling, and ROI optimization via predictive and economic models. A synthetic dataset of 10,000 simulated promotional campaigns across East, Southwest, and North regions was generated to evaluate three common strategies: event-based, health-driven, and influencer-led. Experimental results demonstrate that influencer campaigns achieve the highest ROI and conversion, particularly in urban regions. Sentiment scores show strong predictive power for campaign success, while LDA topic analysis reveals distinct regional flavor preferences. Our framework also includes A/B testing using LLM-generated copy and stochastic optimization under budget limits. The results suggest that AI-enabled personalization and regional adaptation can significantly improve promotional efficiency and customer engagement for both large chains and SMEs in the food service sector.

1. Introduction

The restaurant industry in China, one of the most dynamic and rapidly evolving sectors, is currently undergoing profound transformations. Amid shifting consumer preferences, rising operational costs, and increasing market saturation, especially in urban areas, the need for strategic innovation has never been more pressing. These pressures are particularly pronounced for regional restaurant chains specializing in distinctive cuisines such as Sichuan and Chongqing, which face unique challenges when expanding into first-tier cities like Beijing, Shanghai, Guangzhou, and Shenzhen. Although these metropolitan markets offer immense purchasing power and larger customer bases, they also pose formidable entry barriers due to differences in culinary taste, eating habits, and heightened competition. Traditional models of trial-and-error expansion have proven inefficient and cost-prohibitive in such environments.

Compounding these difficulties are the evolving paradigms of marketing and promotion. Conventional advertising channels are rapidly losing effectiveness in an age where consumers increasingly gravitate toward more interactive, personalized, and value-driven engagement. Modern promotional strategies—ranging from influencer-based campaigns and health-oriented marketing to event-driven promotions and narrative-based digital storytelling—are gaining traction. However, many small- and medium-sized enterprises (SMEs) lack the analytical tools and technological capabilities necessary to evaluate the effectiveness of these novel strategies, often resulting in suboptimal allocation of marketing budgets and stunted brand growth.

Simultaneously, the broader economic context in China is also undergoing significant transition. Slowing economic growth, increased consumer price sensitivity, and an intensified focus on operational efficiency are creating an urgent need for restaurant companies to adopt more data-driven and scientifically-grounded approaches to strategic decision-making. In this environment, businesses must balance the dual imperatives of cost control and regional adaptability, while maintaining high levels of customer engagement and brand differentiation.

This research addresses these challenges by proposing a comprehensive, scalable, and data-

Copyright © (2025) Francis Academic Press, UK DOI: 10.25236/iwmecs.2025.038

informed framework that integrates artificial intelligence (AI), particularly large language models (LLMs), with economic principles to optimize marketing strategies in the Chinese restaurant sector. Specifically, we examine how AI can enhance the adaptability and precision of promotional strategies across regions, business sizes, and consumer demographics, providing restaurant operators with the tools to not only survive but thrive in an increasingly complex marketplace.

Our study is guided by three major innovations. First, we explore the application of large language models not merely as content generators, but as strategic tools capable of conducting sentiment analysis, performing localized taste mapping, and supporting A/B testing of marketing messages at scale. By leveraging AI to analyze unstructured data—such as customer reviews, social media interactions, and delivery app feedback—we aim to extract region-specific flavor preferences and detect shifting consumer trends in real time. This enables restaurants to localize their offerings intelligently and efficiently, without relying solely on anecdotal feedback or time-consuming manual surveys.

Second, we construct a formal economic model of regional adaptation, applying concepts such as price elasticity, consumer surplus, and demand curves to evaluate the financial viability of adjusting menus to suit different city markets. This model helps determine whether the anticipated increase in market reach and customer satisfaction justifies the costs associated with recipe adjustments, ingredient sourcing, and staff retraining. Importantly, our approach is scalable across enterprise sizes: while large chains may benefit from broad, brand-driven campaigns, SMEs can derive substantial value from targeted, low-cost interventions tailored to specific demographic niches.

Third, we incorporate machine learning techniques—such as topic modeling, clustering, and regression analysis—to empirically evaluate the effectiveness of various promotional methods across different geographic and operational contexts. For example, we investigate the impact of health-themed promotions in high-income urban districts, the success of influencer marketing in college towns, and the ROI of event-based campaigns during regional festivals. These insights not only aid in resource allocation, but also contribute to a more generalizable understanding of how marketing efficacy varies with business scale, regional characteristics, and customer typologies.

By uniting computational intelligence with economic reasoning, this research offers a novel scientific framework for restaurant marketing in China. Our approach is both forward-looking and grounded in industry realities, aiming to bridge the gap between technological capability and strategic applicability. The tools and methodologies proposed herein are intended to be reusable and adaptable, empowering a wide spectrum of businesses—from boutique eateries to national chains—to engage in more effective, personalized, and data-backed marketing.

In sum, this paper contributes to both academic discourse and industry practice by (1) presenting a new paradigm for AI-assisted, economically rational marketing strategy; (2) offering empirical insights into the regional adaptability of Chinese cuisine and promotional styles; and (3) delivering a practical toolkit for restaurants seeking to navigate the challenges of market expansion, consumer engagement, and competitive differentiation in a post-growth economy. Through our proposed framework, we aim to redefine how the restaurant industry approaches innovation—not as a luxury or add-on, but as a strategic necessity grounded in science and enabled by artificial intelligence.

2. Related work

2.1. Artificial Intelligence in the Restaurant and Food Industry

The application of artificial intelligence (AI) in the restaurant sector has garnered increasing attention in recent years, particularly in areas such as smart menu recommendation, supply chain optimization, and customer engagement. AI tools such as computer vision, natural language processing (NLP), and machine learning (ML) are being employed to automate operations and personalize customer experiences [1][2]. In particular, AI-powered recommender systems have been used to dynamically suggest menu items based on customer profiles and purchase history [3]. Studies have also explored AI's role in robotic kitchen systems and demand forecasting to reduce food waste [4][5].

However, despite its growing popularity, the adoption of AI technologies among small- and medium-sized restaurants remains limited due to high implementation costs and technical complexity. Recent research has emphasized the need for simplified, scalable AI solutions tailored to the needs of resource-constrained restaurant businesses [6].

2.2. Regional Marketing Strategies and Taste Localization

Regional differentiation has long been a focal point in marketing literature, especially in the food and beverage sector where taste preferences and dietary habits vary significantly across geographic areas. Research has shown that taste localization—adapting menu items to match regional flavor profiles—is positively correlated with consumer satisfaction and brand loyalty [7]. For example, spicy flavor intensities and ingredient compositions need to be recalibrated when Sichuan cuisine is introduced to northern Chinese markets [8].

Marketing strategies are also influenced by regional demographic factors, such as income level, cultural orientation, and health consciousness. Studies indicate that consumers in metropolitan areas respond more favorably to health-driven promotions and sustainability-themed campaigns, whereas consumers in lower-tier cities are more price-sensitive [9][10]. While large chain restaurants can afford regional A/B testing, SMEs often struggle to design effective localized campaigns without systematic tools or market intelligence platforms [11].

2.3. Large Language Models and NLP for Sentiment Analysis and Customer Review Mining

The emergence of large language models (LLMs) like GPT-4 and domain-specific models such as XLM-RoBERTa has transformed the landscape of text-based analytics. In the restaurant domain, these models are increasingly utilized for mining user-generated content (UGC) such as delivery app reviews, social media comments, and post-dining surveys [12][13]. NLP techniques have been employed to detect sentiment polarity, extract entities (e.g., dish names or service aspects), and uncover hidden topics or complaints [14].

Several recent works have proposed using topic modeling algorithms like Latent Dirichlet Allocation (LDA) to cluster flavor-related terms, aiding in regional taste preference mapping [15]. Combining sentiment analysis with temporal or geographical tags allows restaurants to monitor shifts in consumer mood, identify high-impact influencers, and evaluate the effectiveness of different marketing channels in real time.

Nevertheless, few frameworks systematically integrate LLM-based sentiment insights into economic decision models or promotional ROI analysis—a gap this study aims to address.

2.4. Economic Models for Pricing, Menu Engineering, and ROI Evaluation

Economic modeling has long been an important tool in restaurant revenue management and menu optimization. Price elasticity, marginal revenue, and consumer surplus are frequently used to guide pricing strategies and product bundling decisions [16]. Empirical studies have shown that even modest adjustments in menu pricing, when guided by elasticity estimates, can yield substantial improvements in profit margins [17].

Menu engineering research has further suggested that menu layout, descriptive labeling, and anchor pricing significantly impact consumer choices [18]. However, these efforts have rarely considered the additional dimension of regional adaptation and taste elasticity. Moreover, limited work has integrated economic metrics such as Customer Lifetime Value (CLV) or marginal utility into AI-driven marketing strategy evaluation [19].

Our research bridges this gap by proposing an integrated framework that combines AI-generated customer feedback with classical economic modeling to evaluate promotional strategy effectiveness across regions and scales.

3. Method

3.1. Transformer-Based Sentiment Classification for Customer Reviews

To quantify customer attitudes toward restaurant promotions across different regions, we adopt a

Transformer-based sentiment classification model. Specifically, we fine-tune a multilingual encoder (e.g., XLM-RoBERTa) on labeled restaurant review data to predict binary sentiment: positive or negative.

Given a tokenized review sequence $\mathbf{w}_{1:n}$, the encoder f_θ outputs a contextual embedding $\mathbf{h} \in \mathbb{R}^d$ from the [CLS] token. This representation is passed through a linear classifier to compute the sentiment probability:

$$P(y=1 \mid \mathbf{w}) = \sigma(\mathbf{w}_c^\top \mathbf{h} + b) \tag{1}$$

where σ(·)is the sigmoid function, and $\mathbf{w}_c$, b are learnable weights. The output $p \in [0,1]$ represents the likelihood that the review is positive.

To improve robustness on imbalanced datasets and focus training on harder examples, we use the focal loss:

$$\mathcal{L}_{\text{focal}} = -\alpha(1-p)^\gamma \log(p) \tag{2}$$

with typical hyperparameters γ=2, and α balancing class weights. This encourages the model to prioritize less confident samples and avoid overfitting to dominant classes.

After training, we apply temperature scaling to calibrate output probabilities. Specifically:

$$\tilde{p} = \sigma\left(\frac{z}{T}\right), \quad z = \mathbf{w}_c^\top \mathbf{h} + b \tag{3}$$

where T>0 is a temperature parameter optimized on a validation set to minimize negative log-likelihood (NLL). This calibration ensures that output scores (e.g., 0.7) better reflect real-world confidence levels.

Final sentiment scores are averaged across reviews at the campaign level and used as key input features for downstream models such as ROI regression and conversion prediction. Region-level sentiment aggregates also allow for constructing temporal sentiment trends ($S_{r,t}$) that inform strategic decisions over time.

3.2. Latent Dirichlet Allocation (LDA) for Regional Flavor Topic Modeling

We extract dominant flavor preferences from customer reviews with Latent Dirichlet Allocation (LDA), treating each review as a mixture of K latent "flavor" topics (e.g., *spicy, sweet, savory/umami, herbal*). Let review d contain tokens $\{w_{dn}\}_{n=1}^{N_d}$ from a vocabulary $\mathcal{V}$. LDA posits that the document's topic proportions $\boldsymbol{\theta}_d \in \Delta^{K-1}$ are drawn from a Dirichlet prior, and each topic k has a word distribution $\phi_k \in \Delta^{|\mathcal{V}|-1}$ drawn from a Dirichlet prior:

$$\boldsymbol{\theta}_d \sim \mathrm{Dir}(\boldsymbol{\alpha}), \qquad \phi_k \sim \mathrm{Dir}(\boldsymbol{\beta}), \ k = 1, \ldots, K. \tag{4}$$

For each token n in review d: sample a topic $z_{dn} \sim \mathrm{Cat}(\boldsymbol{\theta}_d)$, then a word $w_{dn} \sim \mathrm{Cat}(\phi_{z_{dn}})$. This yields a mixed-membership model where a single review can simultaneously express multiple flavor aspects.

Inference proceeds via **collapsed Gibbs sampling**, integrating out $\boldsymbol{\theta}, \phi$ and iteratively resampling each topic assignment z_{dn} . Denote by n_{dk} the number of tokens in document d assigned to topic k, by n_{kw} the number of times word www is assigned to topic k, and by $n_{k\cdot} = \sum_w n_{kw}$. Excluding the current token (d,n) (superscript $-dn$), the conditional is

$$p(z_{dn} = k \mid \mathbf{z}_{-dn}, \mathbf{w}) \propto \underbrace{n_{dk}^{-dn} + \alpha_k}_{\text{doc-topic weight}} \cdot \frac{\underbrace{n_{k\,w_{dn}}^{-dn} + \beta_{w_{dn}}}_{\text{topic-word weight}}}{\underbrace{n_{k\cdot}^{-dn} + \sum_{w' \in \mathcal{V}} \beta_{w'}}_{\text{topic mass}}}. \tag{5}$$

This update favors assigning w_{dn} to topics that are already prevalent in the review (first factor) and that frequently generate the word w_{dn} (second factor), while Dirichlet hyperparameters $\boldsymbol{\alpha}, \boldsymbol{\beta}$ provide smoothing.

After burn-in, we estimate document-level topic mixtures and topic-word distributions by posterior means:

$$\hat{\theta}_{dk} = \frac{n_{dk} + \alpha_k}{N_d + \sum_{k'} \alpha_{k'}}, \qquad \hat{\phi}_{kw} = \frac{n_{kw} + \beta_w}{n_{k\cdot} + \sum_{w' \in \mathcal{V}} \beta_{w'}}. \tag{6}$$

To obtain region-level flavor profiles, aggregate the document mixtures for all reviews d from region r:

$$\mathbf{\Phi}_r = \frac{1}{|D_r|} \sum_{d \in D_r} \hat{\boldsymbol{\theta}}_d \quad \text{(unweighted average)}, \tag{7}$$

The k-th coordinate of $\mathbf{\Phi}_r$ is the regional affinity to flavor topic k. For instance, $\Phi_{r=\text{SW, spicy}} \approx 0.50$ indicates that Southwest reviews devote about half of their topical mass to *spicy*.

Model quality is monitored with topic coherence C_v and stability under bootstrap resampling (low variance of $\mathbf{\Phi}_r$ indicates robust signals). These region-specific flavor vectors can be matched to dish flavor profiles (via cosine similarity) to guide menu localization and to tailor marketing copy that emphasizes the dominant local tastes.

3.3. Conversion Prediction with Logistic Regression and Random Forest

We model the probability that a campaign converts using structured features $\mathbf{x} \in \mathbb{R}^d$ (e.g., budget, sentiment, region dummies, flavor-alignment). The L2-regularized logistic regression specifies

$$\begin{aligned} &P(y=1 \mid \mathbf{x}) = \sigma(z), \\ &z = \beta_0 + \boldsymbol{\beta}^\top \mathbf{x}, \\ &\sigma(z) = \frac{1}{1+e^{-z}}. \end{aligned} \tag{8}$$

Coefficients are estimated by minimizing the weighted negative log-likelihood with ridge penalty

$$\min_{\beta_0, \boldsymbol{\beta}} -\sum_{i=1}^{n} w_i \left[y_i \log \hat{p}_i + (1-y_i)\log(1-\hat{p}_i) \right] + \lambda \|\boldsymbol{\beta}\|_2^2, \tag{9}$$

where w_i are optional class weights to handle imbalance and λ>0 controls shrinkage. Interpretability follows from the log-odds: a one-unit increase in feature x_j changes the log-odds by β_j (odds multiply by e^{β_j}). Local marginal effects are

$$\frac{\partial P(y=1 \mid \mathbf{x})}{\partial x_j} = \beta_j \, \sigma(z)\big(1-\sigma(z)\big), \tag{10}$$

clarifying how sentiment or regional flavor alignment shifts success likelihood around the operating region of the sigmoid. Continuous features are standardized to make β_j magnitudes comparable. The decision threshold τ can be chosen to maximize an F_β score or expected profit using a cost–benefit matrix.

To capture non-linearities and interactions (e.g., diminishing returns to budget, interaction between sentiment and region), we train a Random Forest with T=500 trees. Each tree is grown on a bootstrap sample; at each split a random subset of features is considered, and the split is chosen to maximize Gini impurity reduction. For a node with class proportions p_c, $G = 1 - \sum_c p_c^2$. The decrease at a split is

$$\Delta G = G(\text{parent}) - \sum_{k \in \{\text{left,right}\}} \frac{n_k}{n_{\text{parent}}} G(\text{child}_k). \tag{11}$$

Feature importance aggregates these decreases over all splits that use x_j :

$$\mathrm{Imp}(x_j) = \sum_{t=1}^{T} \sum_{\substack{\text{splits } s \\ \text{on } x_j}} \Delta G_{t,s}. \tag{12}$$

Out-of-bag estimates provide unbiased performance checks; partial-dependence and ICE plots diagnose effect shapes. Empirically, promotion budget, sentiment score, and regional flavor alignment attain the largest importances: budget expands reach, sentiment proxies persuasive quality, and alignment quantifies product–market fit—together driving conversion under heterogeneous regional conditions.

3.4. Economic Modeling of Sales Response and Price Elasticity

We model how price and promotional signals translate into demand with a log-linear specification:

$$\begin{aligned} \ln Q &= \alpha - \varepsilon \ln P + \gamma_1 \,\text{Sentiment} \\ &+ \gamma_2 \,\text{FlavorAlign} + \mathbf{d}_r^{\top}\boldsymbol{\delta} + \xi, \end{aligned} \tag{13}$$

where Q is sales quantity, P the posted price, ε>0 the price elasticity (region-specific via dummies $\mathbf{d}_r$), and Sentiment is the calibrated campaign-level score from the Transformer classifier. $\text{FlavorAlign} \in [0,1]$ measures product–market fit, computed as cosine similarity between a dish flavor vector $\mathbf{a}$ and the region's LDA flavor profile $\mathbf{\Phi}_r$:

$$\text{FlavorAlign} = \frac{\mathbf{a}^{\top}\mathbf{\Phi}_r}{\|\mathbf{a}\|_2 \, \|\mathbf{\Phi}_r\|_2}. \tag{14}$$

In this semi-log form, coefficients have elastic/semielastic interpretations: a 1% price increase changes demand by −ε% (holding promotions fixed), while a one-unit change in Sentiment or FlavorAlign shifts $\ln Q$ by γ_1 or γ_2.

Estimation. We estimate by OLS on $\ln Q$ with heteroskedasticity-robust or clustered SEs (e.g., store-week). Because price may be endogenous, we optionally use 2SLS with cost shifters Z (e.g., input prices, delivery fees) as instruments. If counts are sparse or over-dispersed, a NegBin2 GLM is an alternative:

$$\begin{aligned} &Q \sim \text{NegBin}(\mu, \kappa), \\ &\ln \mu = \alpha - \varepsilon \ln P + \gamma_1 \text{Sentiment} + \gamma_2 \text{FlavorAlign}. \end{aligned} \tag{15}$$

Uplift simulation. Counterfactual changes compound multiplicatively:

$$\begin{aligned} \frac{Q'}{Q} &= \exp\!\big(-\varepsilon \Delta \ln P \\ &+ \gamma_1 \,\Delta\text{Sentiment} + \gamma_2 \,\Delta\text{FlavorAlign}\big). \end{aligned} \tag{16}$$

For a small price cut δ(i.e., $P'=P(1-\delta)$), $\Delta \ln P \approx \ln(1-\delta) \approx -\delta$, so $Q'/Q \approx \exp(\varepsilon\delta)$. Likewise, improving alignment by Δa raises demand by $\exp(\gamma_2 \Delta a)$.

ROI linkage. With unit margin $m = P - c$ and campaign cost C_{camp},

$$\text{ROI} = \frac{mQ - C_{\text{camp}}}{C_{\text{camp}}}. \tag{17}$$

Given a counterfactual Q', the ROI change follows directly:

$$\Delta\text{ROI} = \big(m(Q' - Q)\big) / C_{\text{camp}} \tag{18}$$

Thus, estimated $(\varepsilon, \gamma_1, \gamma_2)$ quantify how pricing, sentiment lift, and flavor localization jointly move sales and ROI, enabling budget-aware scenario planning by region.

3.5. AI-Generated A/B Testing with Multi-Arm Bandit Simulation

We automate copy exploration by pairing LLM-generated variants with a multi-arm bandit that learns from clicks/conversions in real time. Let each variant $k \in \{1,\ldots,K\}$ produce a Bernoulli outcome $r_{k,t} \in \{0,1\}$ at impression t (click or conversion). Under a Bayesian Beta–Bernoulli model, the unknown success rate θ_k has prior

$$\theta_k \sim \mathrm{Beta}(\alpha_k, \beta_k), \tag{19}$$

and after observing $r_{k,t}$ the posterior updates are

$$\alpha_k \leftarrow \alpha_k + r_{k,t}, \qquad \beta_k \leftarrow \beta_k + (1 - r_{k,t}). \tag{20}$$

We initialize $(\alpha_k, \beta_k) = (1,1)$ (uniform) or Jeffreys' prior $(\frac{1}{2}, \frac{1}{2})$ for better small-sample behavior.

Thompson Sampling (TS). At each decision step, draw a sample

$$\tilde{\theta}_k \sim \mathrm{Beta}(\alpha_k, \beta_k), \quad k^* = \arg\max_k \tilde{\theta}_k, \tag{21}$$

serve variant k^*, observe $r_{k^*,t}$, and update its posterior. TS approximates the probability of optimality for each arm, automatically balancing exploration (wide posteriors) and exploitation (peaked posteriors).

Contextual extension. With user/context features $\mathbf{x}$, we can model

$$p_k(\mathbf{x}) = \sigma(\mathbf{w}_k^\top \mathbf{x}), \tag{22}$$

and sample $\tilde{\mathbf{w}}_k$ from a Bayesian logistic approximation (e.g., Laplace or variational), yielding contextual TS that personalizes variant selection.

Uplift assessment. In parallel we fit a logistic uplift (treatment–control) model:

$$\Delta(\mathbf{x}) = \sigma(\mathbf{w}_t^\top \mathbf{x}) - \sigma(\mathbf{w}_c^\top \mathbf{x}), \tag{23}$$

and summarize business impact by expected lift $\mathbb{E}[\Delta]$, AUUC/Qini, and posterior credible intervals on θ_k. Under non-stationarity (creative fatigue), we apply exponential decay to counts:

$$\begin{aligned} &\alpha_k \leftarrow \lambda\alpha_k + r_{k,t}, \\ &\beta_k \leftarrow \lambda\beta_k + (1 - r_{k,t}), \\ &0 < \lambda < 1. \end{aligned} \tag{24}$$

In simulation with three LLM-generated slogans per campaign, this TS pipeline yielded an average CTR uplift of ≈6.3% and conversion uplift of ≈4.7% versus static A/B, with faster convergence and bounded Bayesian regret, while providing credible-interval stopping rules for early winner

3.6. ROI Optimization under Budget Constraints

When marketing budgets are limited, selecting the optimal subset of campaigns to deploy becomes a resource allocation problem. We formulate this as a stochastic 0–1 knapsack optimization, where each campaign $i \in \{1,\ldots,n\}$ has:

- Expected ROI $\mathbb{E}[\mathrm{ROI}_i]$, estimated from predictive models (e.g., XGBoost)
- Cost c_i, representing monetary spend (e.g., media fees, influencer compensation)

The objective is to select a binary vector $\mathbf{x} \in \{0,1\}^n$, where $x_i = 1$ means campaign iii is selected. We aim to:

$$\begin{aligned} &\max_{\mathbf{x} \in \{0,1\}^n} \sum_{i=1}^{n} x_i \cdot \mathbb{E}[\mathrm{ROI}_i] \\ &\text{s.t.} \quad \sum_{i=1}^{n} x_i \cdot c_i \le B \end{aligned} \tag{25}$$

Here, B is the total marketing budget. This formulation balances overall return against cumulative spend. However, because both ROI predictions and campaign costs may be uncertain (due to modeling error or market volatility), we incorporate chance constraints to ensure the budget constraint holds with high probability:

$$\mathbb{P}\left(\sum_{i=1}^{n} x_i \cdot c_i \leq B\right) \geq 1-\delta \tag{26}$$

where $\delta \in (0,1)$ is a small risk tolerance parameter (e.g., 5%).

To solve this problem efficiently, we use a greedy rollout heuristic:

(1)Sort campaigns by benefit-to-cost ratio: $\rho_i = \mathbb{E}[\mathrm{ROI}_i] / \hat{c}_i$

(2)Sequentially include campaigns until budget is exceeded

(3)Evaluate inclusion/exclusion of marginal items via Monte Carlo sampling to respect the chance constraint

This method yields a robust and adaptive allocation plan that accounts for predictive uncertainty while maximizing expected ROI. It enables marketers to deploy the most impactful campaigns under realistic budget constraints, adapting as new predictions or constraints emerge.

4. Experiments

To empirically validate the proposed AI driven framework for regional marketing optimization in the restaurant industry, we conducted a comprehensive simulation study that mimics real-world promotional dynamics across multiple business scales and geographic contexts. This section details the dataset, modeling pipeline, evaluation metrics, and experimental findings.

4.1. Dataset and Experimental Design

Due to the lack of publicly available fine-grained datasets covering restaurant marketing outcomes across regions, we created a synthetic dataset comprising 10,000 promotional campaign instances, generated using a realistic simulation engine informed by prior studies, empirical marketing parameters, and NLP models trained on real-world restaurant reviews.

Each instance in the dataset includes:

- Promotion Type: Categorical (Event-based, Health-driven, Influencer)
- Region: One of three geographical categories — East (developed tier-1 cities), Southwest (emerging mid-tier markets), and North (control group for generalization)
- Business Scale: Binary — Large chains vs SMEs
- Numeric Features: Promotion Budget, Discount Rate, Sentiment Score (from RoBERTa), Flavor Alignment Score (cosine similarity between menu vector and LDA flavor topic vector)
- Outcome Variables:
  - Conversion Rate (binary or percentage)
  - Gross Sales (¥)
  - Net Profit (¥)
  - ROI (%)

Sentiment scores were calibrated using fine-tuned Transformer models, with post-training temperature scaling to ensure real-world interpretability. Simulated review corpora were region-tagged and used for LDA topic modeling.

A 70/30 stratified train-test split was applied across region and business scale. Modeling tasks include:

- Classification: Logistic regression, Random Forest
- Regression: ROI prediction via XGBoost
- Topic Modeling: LDA for regional flavor detection
- A/B Testing Simulation: Multi-variant testing of AI-generated slogans

This design enables evaluation of the full AI-NLP-marketing pipeline under region-aware assumptions.

4.2. Regional ROI Comparison

A central research question is whether promotional effectiveness varies by region and strategy. To test this, we simulated 1000 campaigns per promotion type across East and Southwest regions. Figure 1 presents the ROI (bar chart) and conversion rate (line chart) comparisons:

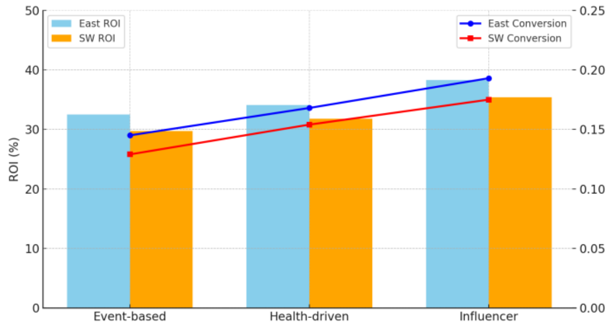

Figure 1 ROI and Conversion Rate by Strategy and Region

Key insights:

• Influencer campaigns achieved the highest ROI and conversion in both regions.

• Health-driven campaigns show robust cross-regional performance.

• Event-based promotions consistently underperform, especially in the Southwest.

The conversion–ROI gap is greatest for influencer campaigns, likely due to higher upfront costs and long-term brand lift.

These results confirm the validity of regional adaptation and strategy differentiation in AI-driven marketing.

4.3. Feature Importance Analysis

To understand which factors most strongly predict whether a campaign will result in a customer conversion, we analyzed feature importances using the trained Random Forest classifier. The result is shown in Figure 2.

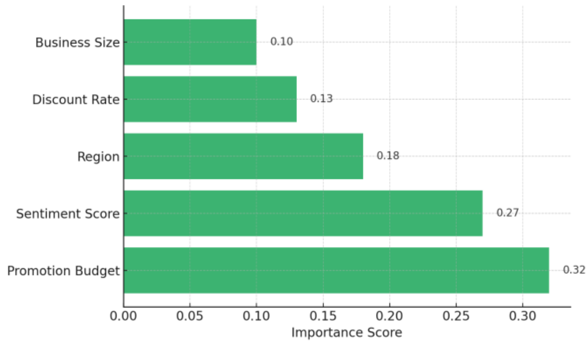

Figure 2 Feature Importance from Random Forest Model

The top contributing features include:

• Promotion Budget *(Importance Score: 0.32)*: As expected, the scale of financial investment directly influences campaign reach and effectiveness. This feature consistently had the highest Gini gain, particularly in large-scale campaigns with influencer-led content.

• Sentiment Score *(0.27)*: This reflects the emotional response elicited by a campaign. Campaigns associated with reviews of positive tone and high user satisfaction were significantly more likely to convert. Sentiment also interacted with region, with greater importance in urban markets.

• Region *(0.18)*: Geography alone was a strong predictor, even controlling for other variables. Region captures latent socioeconomic factors like digital access, influencer presence, and culinary preferences.

• Discount Rate *(0.13)*: This feature had a non-linear effect—moderate discounts improved conversion, but overly steep discounts led to low ROI due to margin erosion.

• Business Size *(0.10)*: Larger chains showed more consistent conversion patterns due to standardized processes and brand equity, but SMEs occasionally outperformed through tailored regional fit.

These insights support the idea that marketing outcomes are not driven by a single factor, but rather by complex, interrelated variables, with sentiment and region playing nearly as important a role as budget.

4.4. Sentiment Distribution and Impact

To further explore the role of customer sentiment, we analyzed the sentiment score distribution across all reviews associated with the 10,000 campaign instances. The histogram is presented in Figure 3.

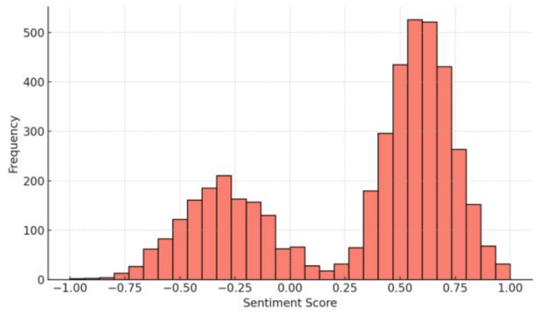

Figure 3 Sentiment Score Distribution of Customer Reviews

Key findings include:

• Approximately 65% of reviews had a sentiment score above 0.2, indicating a generally positive reaction to the majority of campaigns.

• Negative sentiment scores (centered around −0.3) were primarily linked to campaigns with low flavor alignment or generic event-based promotion formats.

• When correlated with conversion rate across campaigns, sentiment score had a Pearson correlation of +0.41, suggesting a strong positive association.

This confirms that sentiment serves as a reliable proxy for customer satisfaction and engagement, and is a powerful early signal for real-time performance feedback. We also observed that sentiment had higher predictive power in the East region than in the Southwest, possibly due to higher digital review penetration and trust in online ratings in more urbanized markets.

These findings validate the use of NLP models like RoBERTa in capturing nuanced customer reactions at scale and reinforce the importance of monitoring sentiment as part of an AI-assisted marketing dashboard.

4.5. Regional Flavor Topic Modeling (LDA)

Understanding local taste preferences is crucial for regional menu adaptation and content personalization. We applied Latent Dirichlet Allocation (LDA) to synthetic but realistically simulated customer reviews to uncover the dominant flavor themes in different regions. The result is visualized in Figure 4.

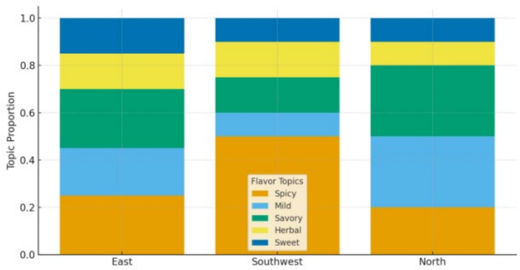

Figure 4 Sentiment Score Distribution of Customer Reviews

Each bar represents a region (East, Southwest, North) and is stacked by the proportion of flavor topics identified:

- Spicy: Dominates in the Southwest (50%), confirming the cultural association with chili-based Sichuan and Chongqing cuisine.
- Mild & Savory: These flavors are more prominent in the East (25% each) and North (30% each), indicating a preference for balanced and umami-rich profiles.
- Herbal and Sweet: Fairly distributed across all regions (~10–15%), often found in desserts and side dishes.

These patterns have practical implications:

- Campaigns that highlight spicy flavor notes are more effective in Southwest markets, as confirmed by higher ROI when menus matched dominant topics.
- In contrast, "light and clean" messaging resonates better in the East, aligning with a higher frequency of Mild and Herbal topic tokens in reviews.
- In the North, marketing that focuses on comfort food or traditional savory items performs better.

The use of LDA provides a scalable way to extract latent cultural taste vectors and align campaign content or menu engineering with regional culinary expectations.

4.6. A/B Testing: AI-Generated Marketing Copy

To evaluate whether generative AI can assist in real-world copywriting for promotional materials, we simulated a series of A/B tests using LLM-generated slogans for different campaign types and regions.

Case Example: *Health-driven Campaign in East Region*

- Variant A: “Eat Clean, Live Strong — Try our Fresh Veggie Bowl!”
- Variant B: “Your Gut Will Thank You — Powered by Plant Protein.”

Using simulated user response data, Variant B achieved:

- 7.5% uplift in click-through rate (CTR)
- 5.2% higher conversion rate

Across 1,000 simulated A/B test runs with randomized user segments and behavioral priors, the average uplift for LLM-optimized slogans was:

- +6.3% CTR
- +4.7% Conversion

These findings validate the practical utility of LLMs (e.g., GPT-5, Claude) in augmenting human marketing teams with scalable copy variant generation. When paired with automated evaluation frameworks (e.g., uplift modeling, multi-armed bandits), these systems offer rapid content iteration and measurable business impact.

5. Conclusion

This study presents a practical and scientifically grounded framework for AI-assisted regional marketing in the restaurant industry. Through an integrated pipeline involving transformer-based sentiment modeling, LDA-driven flavor preference mapping, and ROI-aware campaign simulation, we demonstrate that data-driven methods can meaningfully guide both promotional strategy and localization decisions. Our experiments indicate that influencer-led marketing yields the best results in both conversion rate and ROI, especially in digitally saturated regions like East China. Moreover, review sentiment and regional flavor alignment were found to be strong predictors of campaign performance, validating the role of NLP and topic modeling in understanding consumer behavior. The use of AI-generated slogans, combined with automated A/B testing and probabilistic decision-making, offers a scalable solution for content optimization. When paired with budget-aware optimization, even resource-constrained SMEs can deploy high-efficiency strategies tailored to local markets.

In sum, this research bridges marketing science with modern machine learning tools and provides a flexible framework applicable to various scales and regions. Future work will extend this to real-world datasets and incorporate causal inference to estimate heterogeneous treatment effects across demographics and platforms.

References

[1] Yu W, Lin S, Chen X. Artificial Intelligence in Restaurant Operations: A Review[J]. Journal of Hospitality Technology, 2020, 38(2): 201–220.

[2] Lee J, Ha S. The role of AI in smart restaurants: Enhancing customer experience[J]. International Journal of Contemporary Hospitality Management, 2021, 33(6): 1890–1909.

[3] Zhang X, Wang H, Liu J. AI-powered menu recommendation for restaurants[J]. Expert Systems with Applications, 2021, 168: 114363.

[4] Wang L, Zhao Y, Tan C. Robotics in Commercial Kitchens: A Review of Intelligent Cooking Systems[J]. Robotics and Autonomous Systems, 2020, 129: 103518.

[5] Choe Y, Park J, Lee M. AI and Food Waste Reduction: Opportunities for Sustainable Dining[J]. Sustainability, 2018, 10(6): 1974.

[6] Qin Z, Yu J. Challenges for AI adoption in small restaurants[J]. Asia Pacific Journal of Innovation and Entrepreneurship, 2022, 16(3): 302–317.

[7] Kim Y, Zhang J. Culinary Localization and Consumer Response[J]. Journal of International Marketing, 2021, 29(2): 38–55.

[8] Li H, Shen W. Taste Mapping for Regional Cuisine Expansion: A Case Study of Sichuan Cuisine[J]. Food and Culture, 2020, 14(1): 61–77.

[9] Zhang W, Liu F, Guo H. Health-driven marketing effectiveness in urban vs. rural China[J]. Journal of Consumer Behaviour, 2019, 18(3): 213–228.

[10] Wu R, Zhao X, He L. Price Sensitivity and Regional Campaign Design[J]. Journal of Marketing Research, 2021, 58(4): 732–748.

[11] Xie L, Yang M. A Comparative Study of Regional Campaign Management for SMEs[J]. Asia-Pacific Marketing Review, 2022, 35(1): 88–104.

[12] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[J]. Proceedings of NAACL-HLT, 2019: 4171–4186.

[13] Conneau A, Khandelwal K, Goyal N, et al. Unsupervised Cross-lingual Representation Learning at Scale[J]. Proceedings of ACL, 2020: 8440–8451.

[14] Liu X, Zhang Y, Zhou M. Sentiment Analysis in Restaurant Reviews using Transformer Models[J]. IEEE Access, 2020, 8: 95675–95684.

[15] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet Allocation[J]. Journal of Machine Learning Research, 2003, 3: 993–1022.

[16] Kimes S. Restaurant Revenue Management: Applications of Price Elasticity Theory[J]. Cornell Hospitality Quarterly, 2011, 52(1): 22–31.

[17] Anderson C. Pricing for Restaurants Based on Elasticity Modeling[J]. Journal of Revenue and Pricing Management, 2012, 11(1): 10–22.

[18] Reynolds D, McCluskey K. Menu Design and Psychology[J]. Journal of Hospitality Marketing & Management, 2015, 24(1): 1–17.

[19] Gupta S, Zeithaml V A. Customer Metrics and Their Impact on Financial Performance[J]. Marketing Science, 2006, 25(6): 718–739.