**Research Plan on Causal Effects**

**LIU, Liangjie**

**1. Research Background**

In today's business environment, data-driven decision-making is profoundly impacting enterprises' marketing and user operation strategies. The widespread adoption of recommendation systems and precision marketing models has significantly enhanced user experience and corporate revenue. However, traditional recommendation models primarily focus on predicting users' purchase probability and revenue growth while often neglecting the identification of naturally converting users—those who would complete a desired action without intervention. This oversight leads to resource wastage and unnecessary marketing expenses.

Furthermore, merely considering purchase probability may not accurately measure the true value of marketing actions. For example, some users may complete purchases without marketing intervention, or those who received intervention might not contribute significantly to long-term revenue. Therefore, developing models that quantify the causal effects of interventions is crucial for guiding precision marketing.

Against this backdrop, this study aims to analyze the following key questions:

1. **Existence of causal relationships**: Whether a particular marketing action significantly impacts users' target behaviors (e.g., purchase, revenue increase).
2. **Magnitude of intervention effects**: The actual contribution of marketing actions to user behavior, including uplift probability and long-term effects.

Answering these questions is of great practical significance for optimizing resource allocation and operational decision-making. For example, by precisely identifying user groups sensitive to marketing interventions, businesses can reduce marketing costs while maximizing return on investment (ROI).

Possible research scenarios include:

* **Marketing and purchase relationships**: Analyzing whether marketing exposure increases user purchase intention and identifying sensitive user groups.
* **Recommendation-purchase and revenue relationships**: Examining whether purchasing recommended products leads to long-term revenue growth and evaluating potential negative effects (e.g., product returns or additional cost reductions).

**2. Literature Review**

Causal inference has been widely applied across various fields, particularly in economics, medical research, and marketing analytics. Below is a summary of related literature:

**2.1 Fundamental Methods of Causal Inference**

The study of causal inference originates from the Rubin Causal Model (RCM), which establishes the theoretical foundation for quantifying causal effects.

* **Propensity Score Methods**: Balances the distribution of covariates between treated and control groups, effectively reducing confounding bias.
* **Matching Methods**: Directly compares similar users by matching specific characteristics between the treatment and control groups.

**2.2 Advanced Methods for Causal Effect Estimation**

With the increasing complexity and scale of data, causal effect estimation methods have evolved:

* **Difference-in-Differences (DID)**: Estimates the effects of policies or marketing interventions by comparing changes over time.
* **Causal Forest**: Combines random forest techniques with causal inference to capture heterogeneous treatment effects.
* **Deep Learning-Based Causal Models**: Such as Counterfactual Regression Networks, which can handle high-dimensional features and nonlinear relationships.
* **Heterogeneous Treatment Effect Estimation**: Identifies differences in treatment effects across user groups, aiding personalized decision-making.

**2.3 Applications in Marketing**

Causal inference has been extensively studied in marketing analytics:

* **Marketing Uplift Modeling**: Gutierrez and Gérardy (2017) reviewed the practical application of causal inference in uplift modeling, focusing on quantifying the incremental value of marketing actions.
* **Personalized Treatment Effect Estimation**: Shalit et al. (2017) proposed deep learning-based methods for estimating individualized treatment effects, demonstrating their potential in precision marketing.
* **Customer Lifetime Value (CLV) Analysis**: Causal inference aids in analyzing CLV, providing theoretical support for long-term marketing strategies.

Additionally, the integration of causal inference and machine learning has facilitated large-scale data applications, such as causal graphical models and structural equation models, which are gaining research interest.

**3. Research Directions**

This study focuses on causal effects in user operations and marketing scenarios, exploring the following directions:

1. **Optimization of Causal Effect Analysis Methods**
   * Expanding model applications and improving computational efficiency for large-scale user data.
   * Integrating causal graphical models to construct a framework for analyzing complex user behavior pathways.
2. **Optimization of Personalized Marketing Strategies**
   * Using heterogeneous effect analysis to identify user segments sensitive to marketing interventions and quantify the potential benefits of personalized strategies.
   * Exploring dynamic intervention strategies to optimize marketing deployment over time.
3. **Analysis and Validation of Long-Term Effects (Optional)**
   * Investigating the impact of recommendation behavior on CLV and potential long-term behavioral shifts.
   * Validating the influence of interventions on user retention and brand loyalty to inform strategic marketing decisions.

**4. Expected Outcomes**

Through this research, the following goals are expected to be achieved:

1. Develop a causal effect analysis framework tailored to user operation scenarios, integrating mainstream methods while innovating in causal inference techniques.
2. Provide scientific support for precise marketing strategies, reducing wasted marketing expenditures and increasing CLV.

**5. Data Description**

The study will utilize the following data sources:

1. **User Feature Data**
   * Basic Information: Plan type, tenure, region.
   * Billing Information: Recharge records, balance, spending statistics.
   * Behavioral Information: Data usage, voice call duration, app usage frequency.
2. **Interaction and Engagement Data**
   * Contact Types: App, SMS, in-store interactions.
   * Event Records: Balance inquiries, marketing push notifications, product subscriptions.

**6. Timeline**

1. **October - November 2024: Literature Review & Method Exploration**
   * Systematic analysis of existing literature, identifying suitable causal inference methods for marketing scenarios.
   * Output literature review and refine research direction.
2. **December 2024 - January 2025: Research Plan Development**
   * Finalize research plan and design algorithm framework.
3. **February - March 2025: Experiments & Algorithm Optimization**
   * Implement and validate algorithms in real-world business scenarios.
   * Produce experimental reports and refine methodologies.
4. **April - May 2025: Summary & Report Writing**
   * Consolidate research findings and prepare final reports and presentation materials.

**7. References**

1. Gutierrez, P., & Jean-Yves Gérardy. (2017). Causal Inference and Uplift Modelling: A Review of the Literature. International Conference on Predictive Applications and APIs.
2. Shalit, U., Johansson, F. D., & Sontag, D. (2017). Estimating Individual Treatment Effect: Generalization Bounds and Algorithms.
3. Yao, L., Chu, Z., Li, S., Li, Y., Gao, J., & Zhang, A. (2020). A Survey on Causal Inference.