Recommending for Influencer Marketing: A Customer-Aware Deep Learning Approach with a Field Experiment

Motivation and Challenges

Influencer marketing is an emerging marketing strategy where brands collaborate with individual social media users—known as *influencers*—to promote their products or services [1]. The global industry of influencer marketing has experienced explosive growth, expanding from $9.2 billion in 2020 to nearly $24 billion in 2024 [2]. With the expanding market size, there is a growing demand from influencer platforms such as TikTok Shop, Amazon Live, and RedNote for algorithmic recommendation tools that intelligently match influencers with products or brands to enhance influencer endorsements [3, 4, 5]. However, challenges remain with this work. A recent survey revealed that 61% of marketers who collaborate with influencer platforms have difficulty identifying the best-fit influencers for their product campaigns [6]. Meanwhile, our industry collaborator, a leading influencer platform in Asia, reports that a significant number of influencers on their platform struggle to find suitable products for endorsement.

This prevalent friction highlights the need for a recommendation system (RS) that aligns with the following unique characteristics of influencer marketing and effectively addresses the associated unsolved challenges:

1. ***Influencer-Mediated Marketplace***. Unlike traditional marketplaces where recommendations are made directly to users, in an influencer marketing setting, influencers act as intermediaries— endorsing and guiding customers to find suitable products. As a result, existing RS may not be directly applicable to the influencer endorsement situation, where the influencer is not the end consumer of the products. Latent representations as commonly used in RS methods [7, 8] do not effectively reflect the preference and demand of the customers who follow influencers or purchase from them. To the best of our knowledge, none of the existing RS are designed to consider customer preferences in influencer-product matching for influencer marketing. This is one of the key motivations of our study.
2. ***Social Influence***. A survey found that 69% of customers trust influencer recommendations more than brand advertisements, demonstrating the social influence of influencers on customer purchasing decisions [9]. Social influences, which come from followership and product-adoption dynamics between influencers and their audiences, play a crucial role in shaping these decisions and can provide valuable insights for influencer-product recommendations. However, how to effectively incorporating these social influence factors into algorithmic recommendations remains an unresolved challenge.
3. ***Fast-Paced and No-replicated Endorsement***. On influencer platforms, endorsements typically have short time windows and are rarely repeated, leading to a short product lifespan[1]. This results in severe data sparsity and cold start issues, which significantly hinder the accuracy and timeliness of recommendation performance. Designing a specific recommendation model that accounts for these unique characteristics and technical issues during training and deployment, ensuring its practical applicability is still challenging.

[1]For example, for our industry collaborator, over 85% of the endorsed products have a sales history of less than one month.

Methodology

We propose a customer-aware attention network model to tackle these challenges in influencer-product matching for influencer marketing. An illustration of the structure is provided in Figure 1. The model has the following novel modules compared to existing state-of-the-art models:

1. Our model incorporates an independent customer representation learning module that uses feed-forward neural networks to pre-train customer-influencer followership, customer-influencer adoption relationships, and customer-product adoption relationships. This module ensures that the customer embedding vectors effectively capture the social influence from influencers as well as the customer's own purchase preferences.
2. We use Graph Neural Networks [10] to aggregate customer embeddings into influencer- and product-level representations. This aims to tackle the main challenge in the influencer-mediated marketplace: the recommendation to influencers should not only meet the their own preferences, but also the preferences of customers who follow or purchase from them.
3. We use a self-attention module from Transformers [11] to encode non-linear interactions and dependence among features and latent representations of customers, influencers, and products. Additionally, we use negative sampling to overcome the challenges of data sparsity and employ K-Nearest Neighbors imputation method [12] to address the cold start issues.

Field Experiment and Main Findings

To evaluate the impact of our proposed model on influencer and customer behavior, we partnered with one of the main influencer platforms in Asia, which specializes in the endorsement of beauty and cosmetic products.

This platform provides a large dataset for model training and testing, featuring 257,275 customers, 5,170 influencers, and 33,136 cosmetic products across 52 categories, with detailed records of 345,381 sales, 132,282 followership interactions, and 53,414 influencer endorsements.

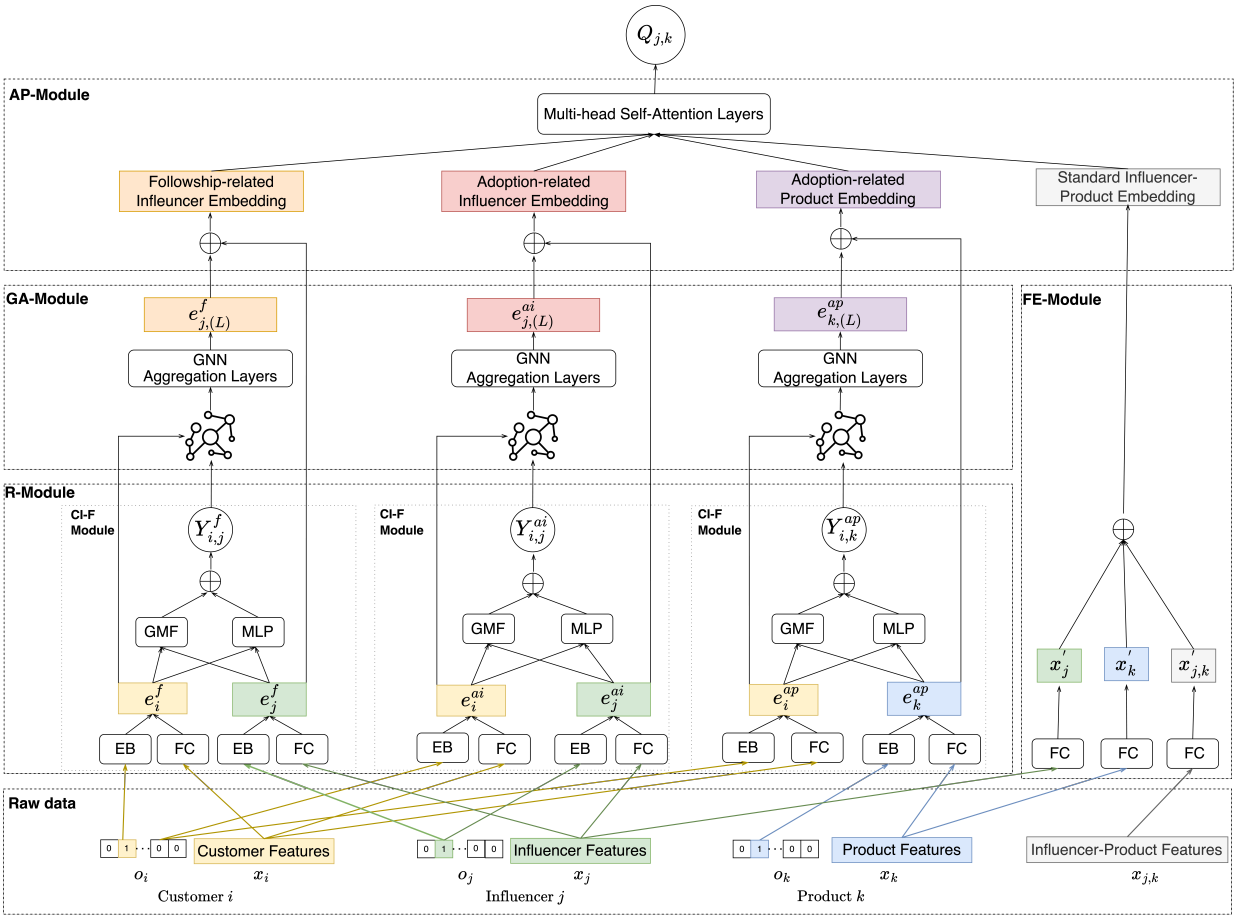
We conducted a field experiment from September to November 2024, randomly assigning 632 influencers into five groups. Each group received product recommendations for endorsement, generated by a distinct recommendation method. Details about the groups can be found in Table 1. We tracked their subsequent endorsement actions, as well as customer purchase and post-purchase behaviors.

Table 2 illustrates how different recommendation models align with influencers' preferences and how their sales performance is impacted by the recommendations. Compared to the state-of-the-art recommendation model (Group C), our model (Group E) enhances influencer endorsement engagement, resulting in a 53.0% higher endorsement rate. Additionally, it can drive 2.6 more in sales quantity and generate an additional $31.6 in revenue per endorsement. Table 3 shows how different recommendation models align with customers' preferences. Compared to all existing models, our model achieves at least a 24.5% increase in customer purchase likelihood, a 42.5% increase in review submission rates, and a 0.019-point improvement in product ratings, confirming its capability to enhance customer adoption and post-purchase satisfaction. The comparison between Group D and Group E in Tables 2 and 3 also demonstrates the effectiveness of our proposed customer representation learning and aggregation model structure.

In conclusion, we designed a customer-aware attention network model for influencer marketing, featuring novel modules that capture customer preferences and the influencers' social influence. The field experiment demonstrates the model's significantly positive impact on both influencer and customer utility, promising a more profitable RS for influencer platforms.

Appendix

Figure 1: Model Structure



Note: The model structure of the proposed method is presented. The input consists of CustomerIDs, InfluencerIDs, ProductIDs, and their associated features, while the output is the predicted sales quantity $Q_{j,k}$ for each product k endorsed by each influencer j. The R-Module (Representation learning Module) captures customer representation using three relational components: Customer–Influencer Followership (CI-F), Customer–Influencer Adoption (CI-A), and Customer–Product Adoption (CI-P). Within each component, IDs and raw features pass through an Embedding Layer (EB) and a Fully Connected Layer (FC), respectively, followed by Generalized Matrix Factorization (GMF) and a Multi-Layer Perceptron (MLP). The GA-Module (Graph Aggregation Module) leverages a Graph Neural Network to aggregate customer embeddings into influencer- and product-level representations. The FE-Module (Feature Embedding Module) uses fully connected layers to process features for influencers and products. Finally, the AP-Module (Attention-based Prediction Module) applies a self-attention mechanism to predict product quantity given the product–influencer pair and their aggregated embeddings. Here $\oplus$ denotes the concatenation of embeddings.

Table 1: Experimental Groups and Descriptions

Group	**Description**	**Causal Effect** (compared with treatment Group)
Group A	No product recommendations	The impact of deploying our model online for the platform
Group B	Recommendations generated by randomization.	Our Model - recommendation "nudge" effect
Group C	Recommendations generated by a DeepFM-based algorithm.	Improvement compared to existing SOTA RS model
Group D	Recommendations generated by our proposed partial model, which excludes user embeddings.	Effectiveness of customer representation learning modules
Group E	Treatment Group: Recommendations generated by our proposed full model	-

Table 2: Influencer Response and Sales Performance (Group E as Baseline)

	IfEndorse	**Sales Quantity**	**Sales Revenue**
	logistic	*OLS*	*OLS*
	(1)	(2)	(3)
Group A	-2.182***	-3.709***	-69.285***
	(0.553)	(0.368)	(7.851)
Group B	-1.439***	-3.848***	-76.069***
	(0.426)	(0.356)	(7.583)
Group C	-0.754**	-2.650***	-31.655***
	(0.351)	(0.355)	(7.559)
Group D	-0.511	-1.185***	-11.006
	(0.336)	(0.360)	(7.676)
Controls	Yes	Yes	Yes
Observations	632	528	528
R^2		0.261	0.247
Log Likelihood	-208.380		
Akaike Inf. Crit.	434.759		

Table Notes: Standard errors are in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Customer Response (Group E as Baseline)

	IfPurchase	**IfReview**	**Rating**	**Order Quantity**	**Order Spending**
	logistic	*logistic*	*OLS*	*OLS*	*OLS*
	(1)	(2)	(3)	(4)	(5)
Group A	-0.945***	-0.866***	-0.025***	-0.079***	-1.415***
	(0.094)	(0.254)	(0.007)	(0.007)	(0.136)
Group B	-0.973***	-0.808***	-0.025***	-0.077***	-1.308***
	(0.096)	(0.249)	(0.007)	(0.007)	(0.137)
Group C	-0.281***	-0.554**	-0.019***	-0.051***	-0.911***
	(0.080)	(0.238)	(0.007)	(0.007)	(0.139)
Group D	-0.172**	-0.112	-0.006	-0.027***	-0.448***
	(0.076)	(0.204)	(0.007)	(0.007)	(0.136)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	26,643	26,643	26,643	26,643	26,643
R^2			0.001	0.009	0.012
Log Likelihood	-5,427.512	-1,049.913			
Akaike Inf. Crit.	10,883.020	2,127.826			

Table Notes: Standard errors are in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

References

[1] Gimmick Marketing Marketing. *Influencer marketing*. 2012.

[2] Statista. *Global Influencer Market Size 2015-2025*. 2024. URL: https://www.statista.com/statistics/1092819/global-influencer-market-size/.

[3] Rheace Jones. *AI and Influencer Marketing: How Algorithms are Matching Brands with the Right Influencers*. 2024. URL: https://www.linkedin.com/pulse/ai-influencer-marketing-how-algorithms-matching-brands-rheace-jones/.

[4] inBeat. *The Use of AI in Influencer Marketing - Key Trends, Tools, and Strategies to Succeed*. 2024. URL: https://inbeat.agency/blog/ai-in-influencer-marketing.

[5] Jessie Liu and Yi Liu. "Asymmetric impact of matching technology on influencer marketing: Implications for platform revenue". In: *Marketing Science* 44.1 (2025), pp. 65–83.

[6] AtisfyReach. *When Influencer Marketing Goes Wrong and How to Avoid It*. 2024. URL: https://resources.atisfyreach.com/when-influencer-marketing-goes-wrong-and-how-to-avoid-it/.

[7] Xiaoyuan Su and Taghi M Khoshgoftaar. "A survey of collaborative filtering techniques". In: *Advances in artificial intelligence* 2009.1 (2009), p. 421425.

[8] Robin Van Meteren and Maarten Van Someren. "Using content-based filtering for recommendation". In: *Proceedings of the machine learning in the new information age: MLnet/ECML2000 workshop*. Vol. 30. Barcelona. 2000, pp. 47–56.

[9] Shopify. *28 Important Influencer Marketing Statistics To Know in 2025*. 2024. URL: https://www.shopify.com/blog/influencer-marketing-statistics.

[10] Franco Scarselli et al. "The graph neural network model". In: *IEEE transactions on neural networks* 20.1 (2008), pp. 61–80.

[11] A Vaswani. "Attention is all you need". In: *Advances in Neural Information Processing Systems* (2017).

[12] Utomo Pujianto, Aji Prasetya Wibawa, Muhammad Iqbal Akbar, et al. "K-nearest neighbor (k-NN) based missing data imputation". In: *2019 5th International Conference on Science in Information Technology (ICSITech)*. IEEE. 2019, pp. 83–88.