A Supplementary Experiments

In this section, we provide more experimental results in terms of the details about our proposed method. Specifically, we investigate two more questions:

- **RQ5** Is it necessary to develop a different adoption of the MMI framework mentioned in Section 4.5 the multi-task model?
- **RQ6** Is the effectiveness of the MMI framework in terms of alignment with item features limited in a specific feature setting?

A.1 RQ5: The necessity of different adaptation of MMI framework

As mentioned before, to ensure the recommended performance of multi-task learning models, we combine the original model loss $L_{backbone}$ with policy-gradient loss from the MMI framework to fine-tune the backbone model. To investigate the benefits of such adaptation, we compare the performance of PETER with different training objectives on the TripAdvisor dataset. The results are shown in Table 10 Although directly replacing the original learning objective of the backbone model with MMI outperforms in terms of Rating Alignment, the recommendation performance decreases drastically. We argue that this is caused by the symmetry of the Mutual Information.

$$I(R;E) = H(R) - H(R|E) = H(E) - H(E|R) = I(E;R) \tag{20}$$

Different from post-hoc generation models, the variable R from multi-task learning models is not fixed. Thus, for them, maximizing mutual information $I(R;E)$ will take a risk of leading the rating prediction to align with an unqualified explanation. Hence, the combination of $L_{Backbone}$ and L_{RL} is a more suitable method for applying MMI framework on multi-task learning models.

A.2 RQ6: MMI for feature alignment under different feature settings

. To prove the success of the MMI framework for feature alignment is not contributed by the specific choice of feature assignment, we compare ApRef2Seq and ApRef2Seq + MMI under different feature settings: assigning Top 10/20/50/100 features based on estimation or directly extracting from corresponding user review. As shown in Figure 3, our MMI framework benefits backbone models under all feature settings. Moreover, the increase ratio of FMR is higher when the feature is assigned by estimation which reflects a similar scenario on a real recommendation platform.

Table 10: Performance of PETER on TripAdvisor dataset with different training objectives in terms of Recommendation performance and Alignment with Rating. Accuracy represents 5-class Sentiment Accuracy with percentage values.

	MAE	RMSE	$\frac{I(R;E)}{H(R)}$	Accuracy
PETER	**0.6236**	**0.7978**	0.18	47.69
PETER + MMI w/o $L_{Backbone}$	0.7000	0.9389	**0.71**	**78.12**
PETER + MMI	0.6343	0.8088	0.44	70.55

B Preliminary results on simultaneous alignment with ratings and features

Table 11: Result of Simultaneously aligning explanation with rating and feature on TripAdvisor dataset

	$\frac{I(R;E)}{H(R)}$	Sentiment Accuracy	$I(F;E)$	FMR
PETER+	0.25	54.51,88.96	1.40	55.90
PETER+ +MMI(F;E)+MMI(R;E)	0.40	66.36,91.27	2.62	69.21

To investigate whether the MMI framework can be extended to simultaneously align explanation with predicted rating and item feature, we try a direct adaptation that linearly combines $I(R;E)$ and $I(F;E)$ together as the new MI reward. To be more specific, The MI reward introduced in Section 4.2 is re-designed as $(1-\epsilon) * I_{\theta_{MIR}}(R;E) + \epsilon * I_{\theta_{MIF}}(R;F)$. The parameter ϵ controls the balance between rating alignment and feature alignment. We conduct a pilot experiment on TripAdvisor dataset based on backbone model PETER+ (PETER+ is a variation of PETER using feature as additional input, so it's suitable for both rating alignment and feature alignment . Please note that the performance on rating alignment of PETER+ is different from the one of PETER, and the parameter ϵ is set to 0.2, for we observed that the value of $I(R;E)$ is at a smaller scale compared with $I(F;E)$).

The experiment result is shown in Table 11. It shows the potential of our proposed method to simultaneously align generated explanations with both ratings and features. However, during our experiment, we do realize the relationship between rating alignment and feature alignment needs a more rigorous analysis and deserves more in-depth investigation. Due to the space limit of a single conference paper, we leave them for future work. We will continue to explore this direction and design more advanced methods like Conditional Mutual Information -based framework to further solve this problem. For instance, we can use a two-step optimizing method: the first step uses $I(F;E)$ as the primary reward to enhance alignment with features. The second step uses Conditional Mutual Information-based reward $I(R;E|F)$ to further enhance alignment with rating meanwhile maintaining the ability to align with features.

Figure 3: FMR Performnace of ApRef2Seq and ApRef2Seq + MMI under different feature assignment settings