Natural Language Explanation for Recommendations and Beyond

LI Lei

A thesis submitted in partial fulfilment of the requirements

for the degree of

Doctor of Philosophy

Principal Supervisor:

Prof. SURNAME Name (Hong Kong Baptist University)

August 2025

DECLARATION

I hereby declare that this thesis represents my own work which has been done after registration for the degree of PhD at Hong Kong Baptist University, and has not been previously included in a thesis or dissertation submitted to this or any other institution for a degree, diploma or other qualifications.

I have read the University's current research ethics guidelines, and accept responsibility for the conduct of the procedures in accordance with the University's Research Ethics Committee (REC). I have attempted to identify all the risks related to this research that may arise in conducting this research, obtained the relevant ethical and/or safety approval (where applicable), and acknowledged my obligations and the rights of the participants.

Signature:_____

August 2025

ABSTRACT

Your abstract

Keywords: Explainable Recommendation, Explainable Artificial Intelligence, Recommender Systems

ACKNOWLEDGEMENTS

I would like to thank Dr. Guozhong Li for the helpful suggestion to develop this template and make it open source.

Acknowledgements for subsequent updates:

- Li Lei (May 2022): https://github.com/lileipisces/HKBUthesis

- OnlyBelter (August 2025): https://github.com/OnlyBelter

This template is now maintained by OnlyBelter at: HKBUthesis. Pull requests are welcome; let us keep it up to date together!

Changes

30 Aug 2025

- Centered all chapter titles.
- Fixed incorrect Table of Contents links.
- Added the file `abbreviations.tex`.

Table of Contents

List of Tables

List of Figures

List of Algorithms

List of Abbreviations

RNN recurrent neural network.

Chapter 1

Introduction

A brief introduction to this chapter.

1.1 Motivation

Sample citations: pre-trained GPT-2[1] [1], Transformer [2], recurrent neural network (RNN) [3]

1.2 Outline

1.3 Contributions

Use the following command to shorten the captions shown in List of Tables/Figures:

```
\caption[shorter caption in List of
Tables/Figures]{Real capation above or below the table/figure}
```

See Fig. 1.1 for example.

[1]Codes available at https://github.com/lileipisces/PEPLER

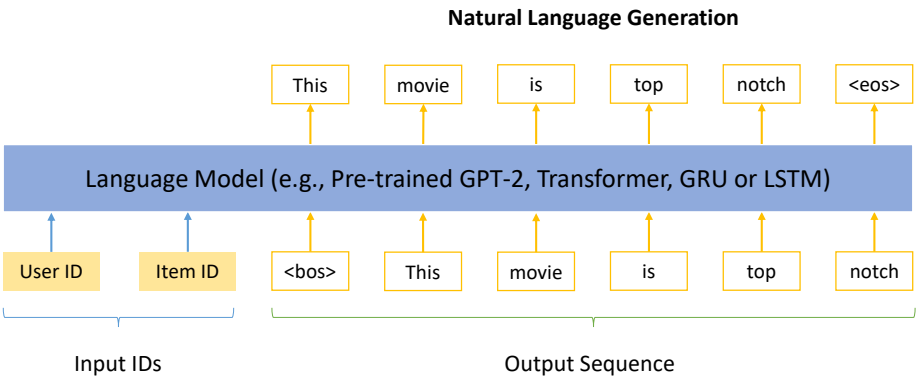

Figure 1.1. Overview of recommender systems-based natural language generation. In the case of recommendation explanation generation, the model is instructed to generate a word sequence for explaining why an item is recommended to the user.

Chapter 2

Literature Survey

2.1 Explainable Recommendation

2.2 Context-aware Recommendation

2.3 Natural Language Generation

2.4 Learning to Rank

Chapter 3

Natural Language Explanation Generation

3.1 Background

3.2 Problem Formulation

3.3 Model Description

3.3.1 Input Representation

3.3.2 Transformer and Attention Masking

3.3.3 Explanation and Recommendation

Explanation Generation:

Context Prediction:

Rating Prediction:

Multi-task Learning:

Algorithm 1 Sentence Grouping via Locality-Sensitive Hashing (LSH)

Input: shingle size n, similarity threshold t, minimum group size g
Output: explanation set $\mathcal{E}$, groups of sentences $\mathcal{M}$

Pre-process textual data to obtain the sentence collection $\mathcal{S}$
$lsh \leftarrow MinHashLSH(t)$, $\mathcal{C} \leftarrow \varnothing$
for sentence s in $\mathcal{S}$ **do**
  $m \leftarrow MinHash()$ // create MinHash for s
  for n-shingle h in s **do**
    $m.update(h)$ // convert s into m by encoding its n-shingles
  end for
  $lsh.insert(m)$, $\mathcal{C}.add(m)$ // $\mathcal{C}$: set of all sentences' MinHash
end for
$\mathcal{M} \leftarrow \varnothing$, $\mathcal{Q} \leftarrow \varnothing$ // $\mathcal{Q}$: set of queried sentences
for m in $\mathcal{C}$ **do**
  if m not in $\mathcal{Q}$ **then**
    $\mathcal{G} \leftarrow lsh.query(m)$ // $\mathcal{G}$: ID set of duplicate sentences
    if $\mathcal{G}.size > g$ **then**
      $\mathcal{M}.add(\mathcal{G})$ // only keep groups with enough sentences
      $\mathcal{E}.add(\mathcal{G}.get())$ // keep one explanation in each group
    end if
    for m' in $\mathcal{G}$ **do**
      $lsh.remove(m')$, $\mathcal{Q}.add(m')$ // for efficiency
    end for
  end if
end for

Table 3.1. Statistics of the three datasets.

	Yelp	Amazon	TripAdvisor
#users	27,147	7,506	9,765
#items	20,266	7,360	6,280
#records	1,293,247	441,783	320,023
#features	7,340	5,399	5,069
#records / user	47.64	58.86	32.77
#records / item	63.81	60.02	50.96
#words / explanation	12.32	14.14	13.01

3.4 Experimental Setup

3.4.1 Datasets

3.4.2 Evaluation Metrics

3.4.3 Compared Methods

3.4.4 Implementation Details

3.5 Results and Analysis

3.5.1 Quantitative Analysis on Explanations

3.5.2 Qualitative Case Study on Explanations

3.5.3 Recommendation Performance

3.5.4 Ablation Study

3.6 Summary

Table 3.2. Ablation study on the smallest dataset TripAdvisor. Arrows ↑ and ↓ respectively denote the performance increase and decrease compared with PETER.

	Explainability			Text Quality			Recommendation	
	FMR	FCR	DIV	USR	BLEU-1	BLEU-4	RMSE	MAE
Disable $\mathcal{L}_c$	0.06 ↓	0.03 ↓	5.75 ↓	0.01 ↓	15.37 ↓	0.86 ↓	0.80 ↑	0.61 ↑
Disable $\mathcal{L}_r$	0.07	0.14 ↑	2.90 ↑	0.10 ↑	16.16 ↑	1.15 ↑	3.23 ↓	3.10 ↓
Left-to-Right Masking	0.07	0.15 ↑	2.68 ↑	0.12 ↑	15.73 ↓	1.11	0.87 ↓	0.68 ↓
PETER	0.07	0.13	2.95	0.08	15.96	1.11	0.81	0.63

Chapter 4

Conclusion and Future Work

4.1 Conclusion

4.2 Future Work

Bibliography

1. Li, L., Zhang, Y. & Chen, L. Personalized Prompt Learning for Explainable Recommendation. *arXiv preprint arXiv:2202.07371* (2022).

2. Li, L., Zhang, Y. & Chen, L. *Personalized Transformer for Explainable Recommendation* in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (Association for Computational Linguistics, Online, 2021), 4947–4957.

3. Li, L., Zhang, Y. & Chen, L. *Generate Neural Template Explanations for Recommendation* in *Proceedings of the 29th ACM International Conference on Information and Knowledge Management* (2020), 755–764.

List of Publications

1. Lei Li, Yongfeng Zhang, Li Chen, *Personalized Prompt Learning for Explainable Recommendation*, **ACM Transactions on Information Systems**, 2022. [*submitted*]

2. Lei Li, Yongfeng Zhang, Li Chen, *Personalized Transformer for Explainable Recommendation*, **Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing**, pages 4947-4957, Online, Thailand, August 1–6, 2021. (**oral paper**)

3. Lei Li, Li Chen, Ruihai Dong, *CAESAR: Context-Aware Explanation based on Supervised Attention for Service Recommendations*, **Journal of Intelligent Information Systems**, volume 57 (1), pages 147-170, August 2021.

CURRICULUM VITAE

Academic qualifications of the thesis author, Mr. LI Lei:

- Received the degree of Bachelor of Engineering from Shenzhen University, June 2017.
- Received the degree of Bachelor of Science from Shenzhen University, June 2017.
- Received the degree of Master of XX from XX University, MM YYYY.

August 2025