



HybridBERT4Rec: A Hybrid Recommender System Based on BERT

Sequential Content-Based and Collaborative Filtering

Leon Knorr | November 6, 2023





Table of Contents

1. An Introduction to Sequential Content Recommendation

2. HybridBERT4Rec

- Architecture
- Strengths & Weaknesses

3. Model Performance & Experiments

4. Applicability to E-Learning

An Introduction to Sequential Content Recommendation
oo

HybridBERT4Rec
oooooo

Model Performance & Experiments
o

Applicability to E-Learning
o

Traditional CBF VS Sequential CBF



Target user
(Alice)



Figure 1: Example history for Alice in traditional CBF [1]

- models **general** user preference

Traditional CBF VS Sequential CBF



Target user
(Alice)



Figure 1: Example history for Alice in traditional CBF [1]

- models **general** user preference
- **BUT:** User preferences change over time! [4]

Traditional CBF VS Sequential CBF



Target user
(Alice)



Figure 1: Example history for Alice in traditional CBF [1]

- models **general** user preference
- **BUT:** User preferences change over time! [4]



Figure 2: Example history for Alice in sequential CBF [1]

- Considers the **order** of historical interactions
- Allows the modelling of “temporary spikes” of interests, as well as the general preferences [4]

A Common Approach to Sequential modelling

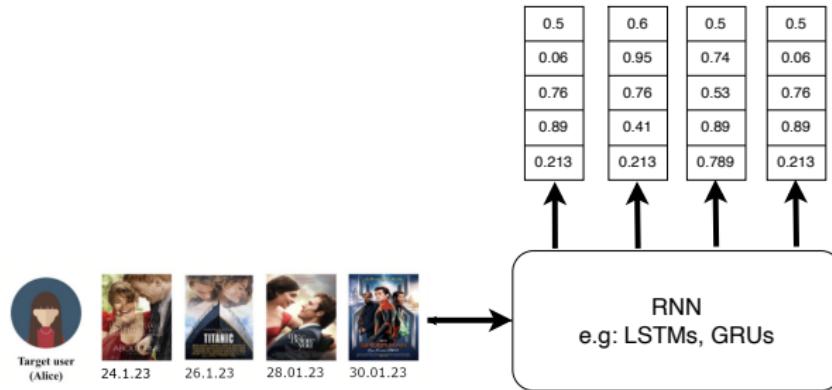


Figure 3: Sequential Content Recommendation with RNNs. [5]

- Suffers from common RNN problems! Especially: **Catastrophic forgetting**, uni-directionality [1]

High Level Overview

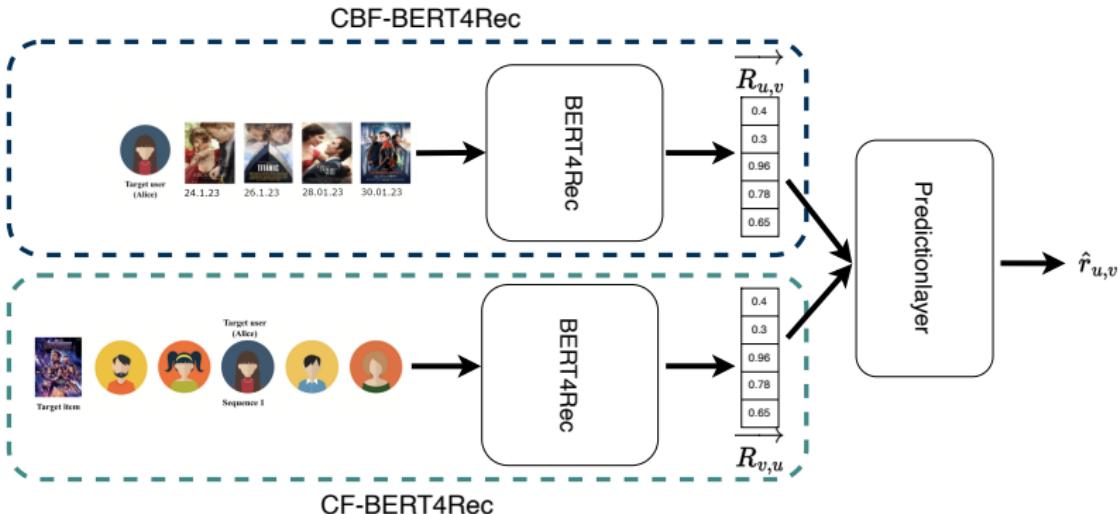


Figure 4: High level overview of HybridBERT4Recs Architecture. [1]

BERT4Rec [3]

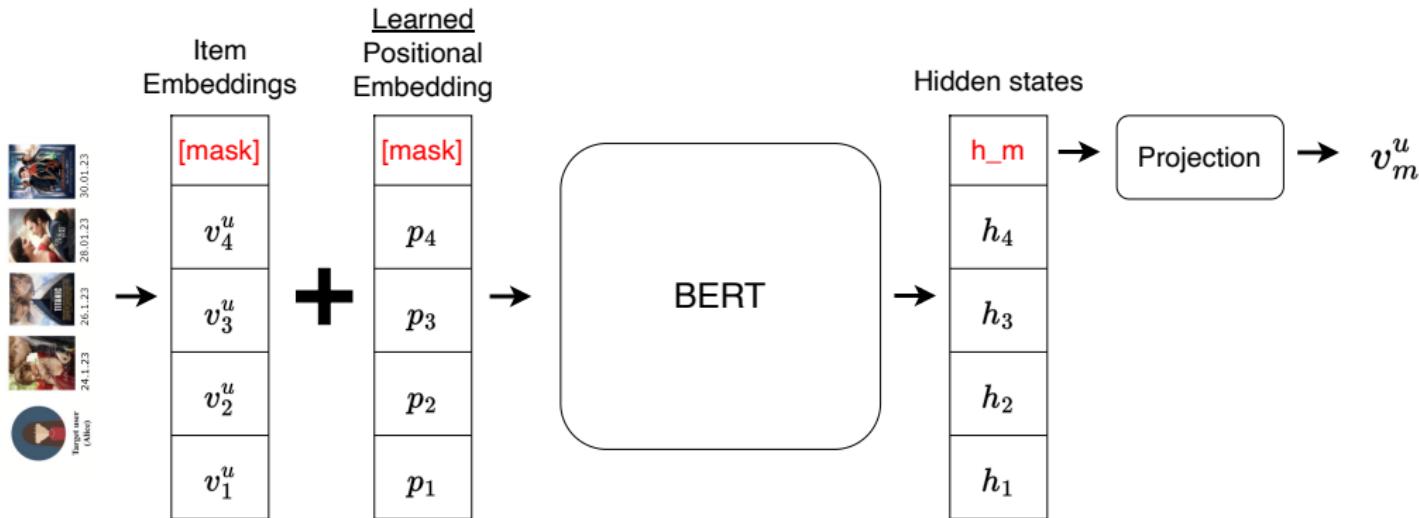


Figure 5: BERT4Rec Architecture, taking item embeddings v_t^u from user u history as input and predicts the next item v_m^u , u is likely to interact with. [3]

CF-HybridBERT4Rec

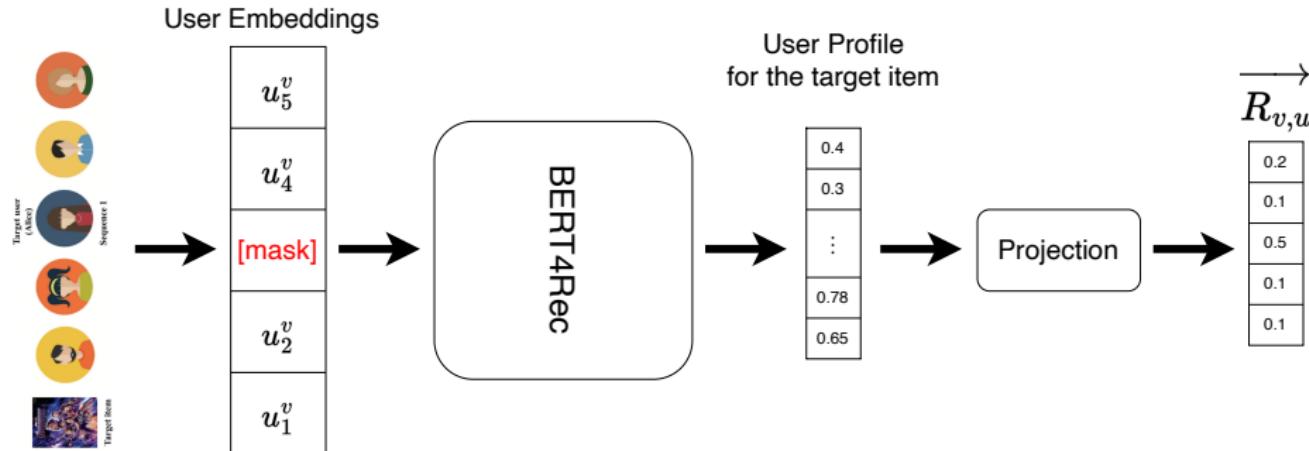


Figure 6: CF-HybridBERT4Rec Architecture, taking user embeddings from all users that have rated the target item v as input and predicts the “target item profile $\overrightarrow{R_{v,u}}$ ”. [1]

CBF-HybridBERT4Rec

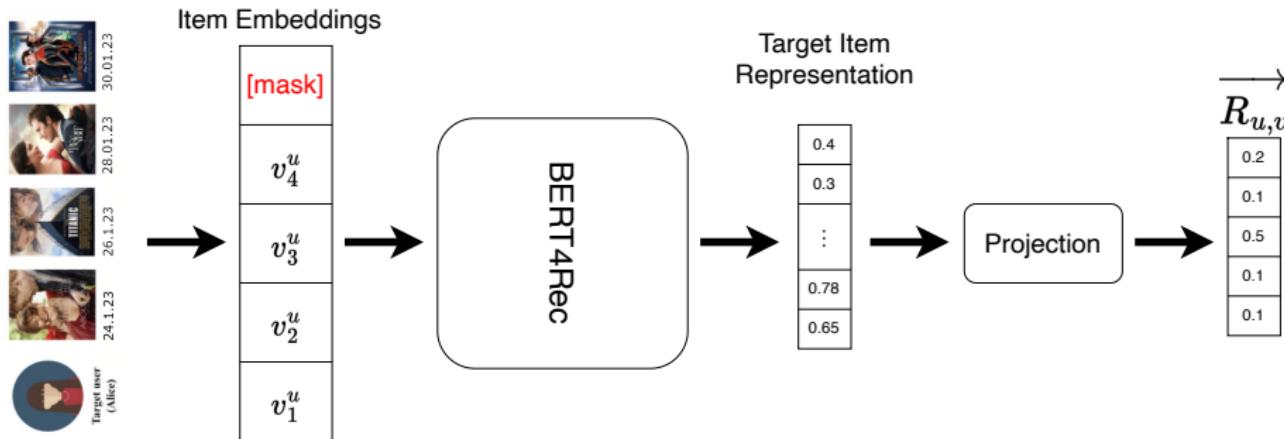


Figure 7: CBF-HybridBERT4Rec Architecture, taking item embeddings from a user u history as input and predicts a “target user profile $\vec{R}_{u,v}$ ”. [1]

Prediction Layer: Combining CF & CBF

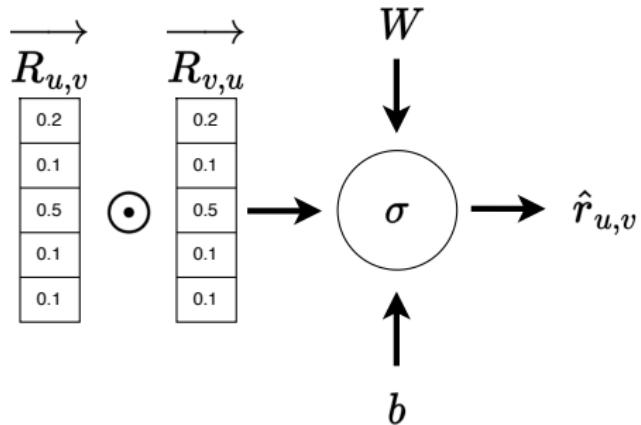


Figure 8: Schematic of HybridBERT4Recs Prediction layer, which uses a generalization of Matrix Factorization based on Neural Networks with Sigmoid activations to predict the rating $\hat{r}_{u,v}$ user u would assign to item v . [1]

Strengths & Weaknesses

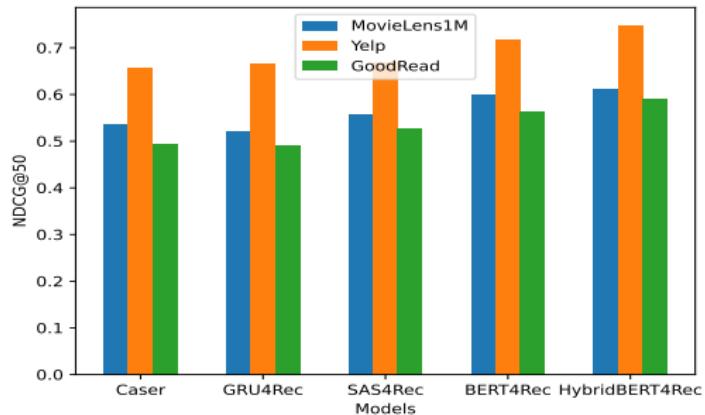
Strengths:

- Easy to parallelize
- Bi-directional
- Sequential
- CBF-, CF-model & Prediction layer can be executed independently
- Intermediate results can be cached

Weaknesses:

- Sequence length is limited
- Needs lots of processing power and memory
- Only uses rating information

Results



Critique:

Figure 9: Performance comparison of different recommender models on three datasets as published by the authors of HybridBERT4Rec. [1]

Results

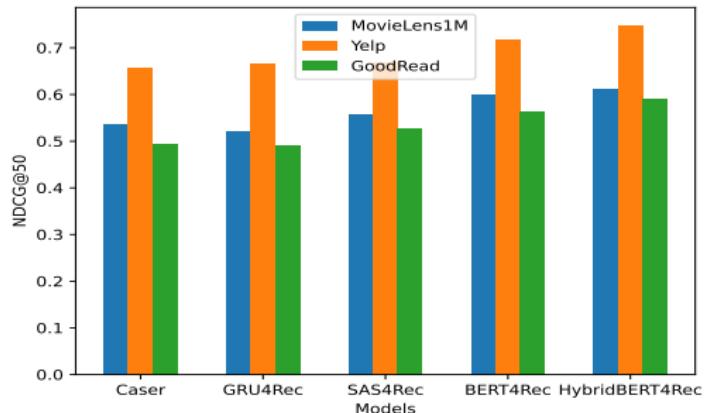
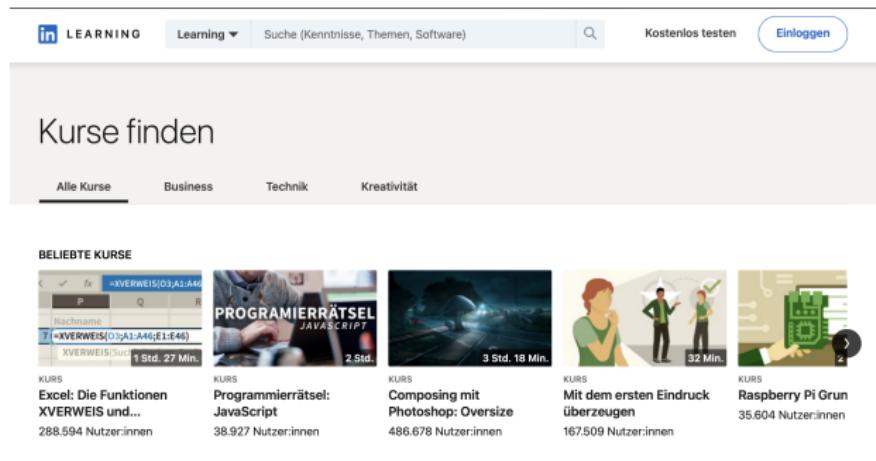


Figure 9: Performance comparison of different recommender models on three datasets as published by the authors of HybridBERT4Rec. [1]

Critique:

- No Hybrid model was evaluated!
- No information about data partitioning
- Generalization performance & real world Applicability is unknown

Applicability to E-Learning



Kurse finden

All Kurse Business Technik Kreativität

BELIEBTE KURSE

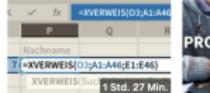
 KURS Excel: Die Funktionen XVERWEIS und... 288.594 Nutzer:innen	 KURS Programmierrätsel: JavaScript 38.927 Nutzer:innen	 KURS Composing mit Photoshop: Oversize 486.678 Nutzer:innen	 KURS Mit dem ersten Eindruck überzeugen 167.509 Nutzer:innen	 KURS Raspberry Pi Grun 35.604 Nutzer:innen
---	---	--	---	---

Figure 10: LinkedIn Learning landing-page. [2]

An Introduction to Sequential Content Recommendation
OO

HybridBERT4Rec
oooooooo



in LEARNING Learning ▾ Suche (Kenntnisse, Themen, Software) Suche (Kenntnisse, Themen, Software)

Lösungen für: Unternehmen Bildungseinrichtungen

Alle Themen / Technik / Softwareentwicklung / Programmiersprachen

PROGRAMMIERRÄTSEL JAVASCRIPT Vorschau

Programmierrätsel: JavaScript

Mit Thomas Rose · 552 Mitgliedern gefällt das
Dauer: 2 Std. · Niveau: Einsteiger:innen · Veröffentlicht am: 20.12.2019

Figure 11: LinkedIn Learning course overview. [2]

Model Performance & Experiments
O

Applicability to E-Learning
●

References

- [1] Chanapa Channarong et al. "HybridBERT4Rec: A Hybrid (Content-Based Filtering and Collaborative Filtering) Recommender System Based on BERT". In: *IEEE Access* 10 (2022), pp. 56193–56206. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2022.3177610. (Visited on 11/02/2023).
- [2] *LinkedIn Learning mit Lynda: Onlinekurse aus dem Business-, Technik- und Kreativbereich.* <https://de.linkedin.com/learning/>. (Visited on 11/03/2023).
- [3] Fei Sun et al. *BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer*. Aug. 2019. DOI: 10.48550/arXiv.1904.06690. arXiv: 1904.06690 [cs]. (Visited on 11/02/2023).
- [4] Shoujin Wang et al. "Sequential Recommender Systems: Challenges, Progress and Prospects". In: (2019), pp. 6332–6338. (Visited on 11/02/2023).
- [5] Feng Yu et al. "A Dynamic Recurrent Model for Next Basket Recommendation". In: *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '16. New York, NY, USA: Association for Computing Machinery, July 2016, pp. 729–732. ISBN: 978-1-4503-4069-4. DOI: 10.1145/2911451.2914683. (Visited on 11/02/2023).

References
ooo

Matrix Factorization

References
● ○ ○

BERT

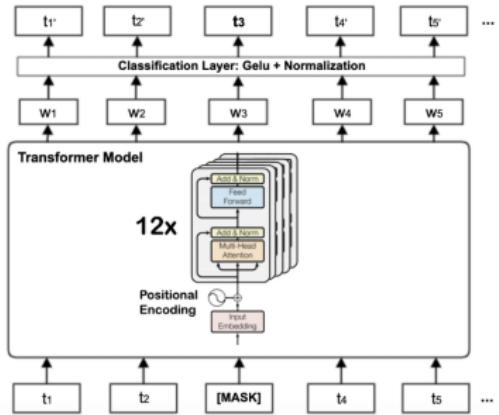


Figure 12: Schematic of BERTs Architecture ¹

¹source:

https://www.researchgate.net/figure/The-Transformer-based-BERT-base-architecture-with-twelve-encoder-blocks_fig2_349546860

References



Normalized Discounted Cumulative Gain (NDCG)

- Measures the rating lists quality by accessing the items rank and rating

$$DCG_k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i + 1)}$$

$$IDCG_k = \sum_{i=1}^{|REL_k|} \frac{rel_i}{\log_2(i + 1)}$$

$$NDCG_k = \frac{DCG_k}{IDCG_k}$$

source: https://en.wikipedia.org/wiki/Discounted_cumulative_gain; [1]

References

