RecSys – DK

Johannes Kruse

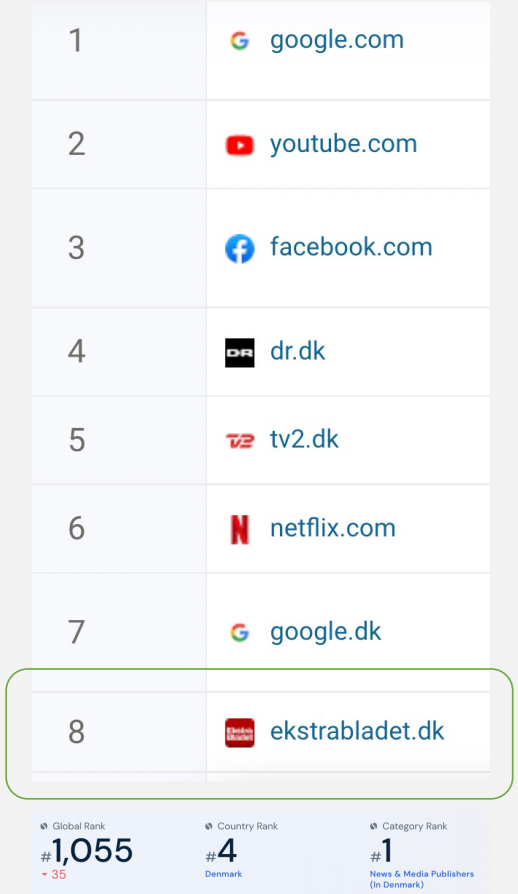
@EkstraBladet

Agenda

- Ekstra Bladet & PIN
- Challenges for News Recommendations
- Project: RRS-DK
  - *Responsible Recommender Systems for Danish News Publishing*
- Latest work
- Current work
- Importance of beyond accucary / metrics

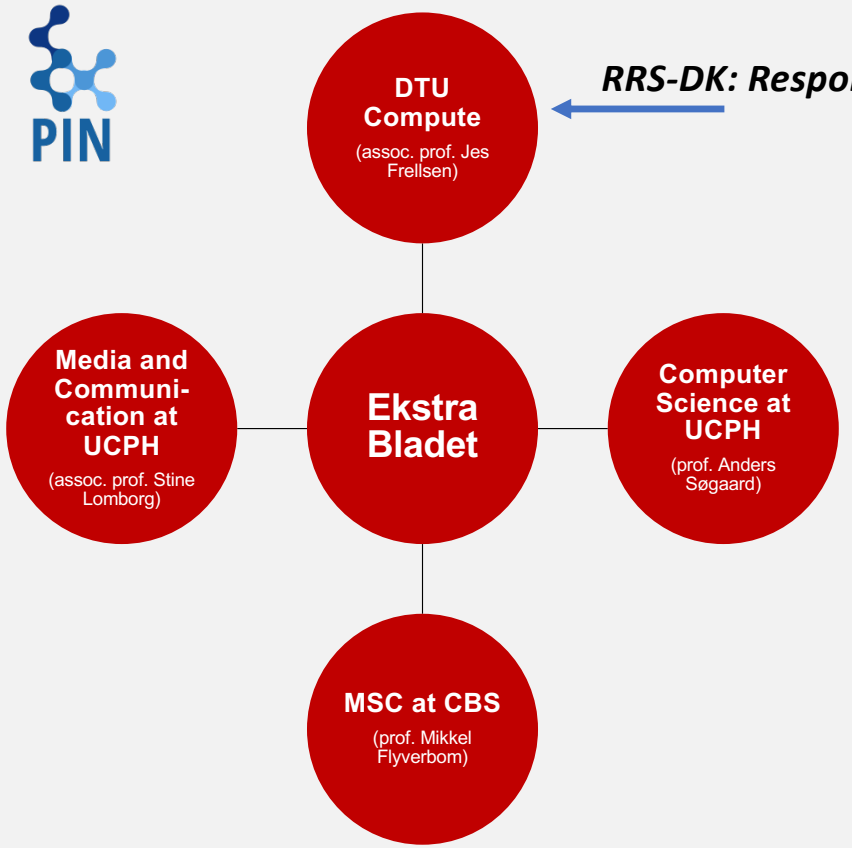
Ekstra Bladet

- Danish News Publisher
  - Text-driven; starting to do audio & video
- 8th Most Visited Websites by Traffic in the Denmark[1]
  - 1 million daily users
  - 450+ mio pageviews per month



1 https://www.similarweb.com/top-websites/denmark/ (07/06/22)

Platform Intelligence News



- Collaboration with leading researchers at Danish universities
- Ekstra Bladet defines and lead the research projects
- Strives for a holistic approach (full picture)
- Fund: IFD's Grand Solutions program (1,5 mil euro) and Industrial PhD programme (0,5 mil. euro)

Research Questions

Ongoing

1. Baseline

RQ1: How well does state-of-the-art RS for news from the perform in a Ekstra Bladet setting?

2. Improve

RQ2: By how much can the SOTA RS for Ekstra Bladet news?

3. Explainability

RQ3: How can the predictions of RS based on deep neural nets best be explained?

RSS-DK – Supervisors

Jes Frellsen
Assoc. Prof.
jefr@dtu.dk

Kasper Lindskow
Head of Research and Innovation at EB
kasper.lindskow@eb.dk

Michael Riis Andersen
Assoc. Prof.
miri@dtu.dk

Lars Kai Hansen
Prof.
lkai@dtu.dk

News Recommendations – Challenges

New news articles are posted continuously and expires fast

- Severe cold-start problem (e.g. Ekstra Bladet publish up to 350 articles daily)
- Survival time (time interval between first and last appearance) of more than 84.5% news articles in MIND is less than two days

User preferences and indirect and (might) change with the news agenda

- Click behaviors in an implicit proxy in regards of user interest
- User interests are usually diverse and dynamic (mood, time a day, device)

News articles contain rich textual information such as title and body

- Requires extensive pre-processing (IDs not appropriate), e.g., NLP techniques

AI alignment challenge

- Existing RS models are not necessarily aligned with news publishers' values and goals

Technical

- There are no general Recommender System framework that everyone is following

MIND Challenge

Download

The MIND dataset is free to download for research purposes under Microsoft Research License Terms. Before you download the dataset, please read these terms and click below button to confirm that you agree to them.

☑ I agree to the Microsoft Research License Terms.

This dataset can support many researches on news recommendation, and can be downloaded at:

Training Set

Validation Set

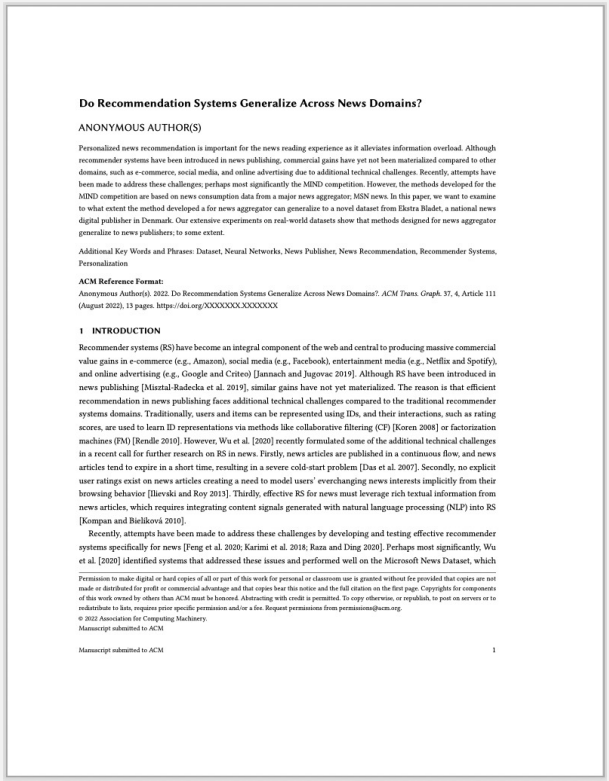
Test Set

Leaderboard

Rank	Team	AUC	MRR	nDCG@5	nDCG@10
1 OCT. 05, 2021	UniUM-Fastformer-Pretrain	0.7304	0.3770	0.4180	0.4718
2 SEPT. 02, 2021	MINER	0.7275	0.3724	0.4102	0.4661
3 AUG. 08, 2021	UniUM-Fastformer	0.7268	0.3745	0.4151	0.4684
4 SEPT. 14, 2022	pengwj	0.7256	0.3720	0.4101	0.4660
5 FEB. 10, 2022	Riiid	0.7252	0.3709	0.4089	0.4647

Do Recommendation Systems Generalize Across News Domains?

- ***Benchmark paper***
  - Baseline for future work
- *Submitted to RecSys 2022 (Seattle)*
  - *Not Accepted*



Do Recommendation Systems Generalize Across News Domains?

ANONYMOUS AUTHOR(S)

Personalized news recommendation is important for the news reading experience as it alleviates information overload. Although recommender systems have been introduced in news publishing, commercial gains have yet not been materialized compared to other domains, such as e-commerce, social media, and online advertising due to additional technical challenges. Recently, attempts have been made to address these challenges; perhaps most significantly the MIND competition. However, the methods developed for the MIND competition are based on news consumption data from a major news aggregator, MSN news. In this paper, we want to examine to what extent the method developed a for news aggregator can generalize to a novel dataset from Ekstra Bladet, a national news digital publisher in Denmark. Our extensive experiments on real-world datasets show that methods designed for news aggregator generalize to news publishers, to some extent.

Additional Key Words and Phrases: Dataset, Neural Networks, News Publisher, News Recommendation, Recommender Systems, Personalization

ACM Reference Format:

Anonymous Author(s). 2022. Do Recommendation Systems Generalize Across News Domains?. *ACM Trans. Graph.* 37, 4, Article 111 (August 2022), 13 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Recommender systems (RS) have become an integral component of the web and central to producing massive commercial value gains in e-commerce (e.g., Amazon), social media (e.g., Facebook), entertainment media (e.g., Netflix and Spotify), and online advertising (e.g., Google and Criteo) [Jannach and Jugovac 2019]. Although RS have been introduced in news publishing [Misztal-Radecka et al. 2019], similar gains have not yet materialized. The reason is that efficient recommendation in news publishing faces additional technical challenges compared to the traditional recommender systems domains. Traditionally, users and items can be represented using IDs, and their interactions, such as rating scores, are used to learn ID representations via methods like collaborative filtering (CF) [Koren 2008] or factorization machines (FM) [Rendle 2010]. However, Wu et al. [2020] recently formulated some of the additional technical challenges in a recent call for further research on RS in news. Firstly, news articles are published in a continuous flow, and news articles tend to expire in a short time, resulting in a severe cold-start problem [Das et al. 2007]. Secondly, no explicit user ratings exist on news articles creating a need to model users' everchanging news interests implicitly from their browsing behavior [Ilievski and Roy 2013]. Thirdly, effective RS for news must leverage rich textual information from news articles, which requires integrating content signals generated with natural language processing (NLP) into RS [Kompan and Bieliková 2010].

Recently, attempts have been made to address these challenges by developing and testing effective recommender systems specifically for news [Feng et al. 2020; Karimi et al. 2018; Raza and Ding 2020]. Perhaps most significantly, Wu et al. [2020] identified systems that addressed these issues and performed well on the Microsoft News Dataset, which

Manuscript submitted to ACM

Do Recommendation Systems Generalize Across News Domains?

- **Assumption**
  - Methods developed for the MIND competition are based on news consumption data from a major news aggregator; MSN news.
- **Motivation**
  - Examine to what extent the method developed a for news aggregator (and other purposes) can generalize to a novel dataset from Ekstra Bladet
- **Finding**
  - *“The main finding of our research is that methods designed for a news aggregator generalize to news publishes; to some extent.”*

General Recommendation Methods

STACKED

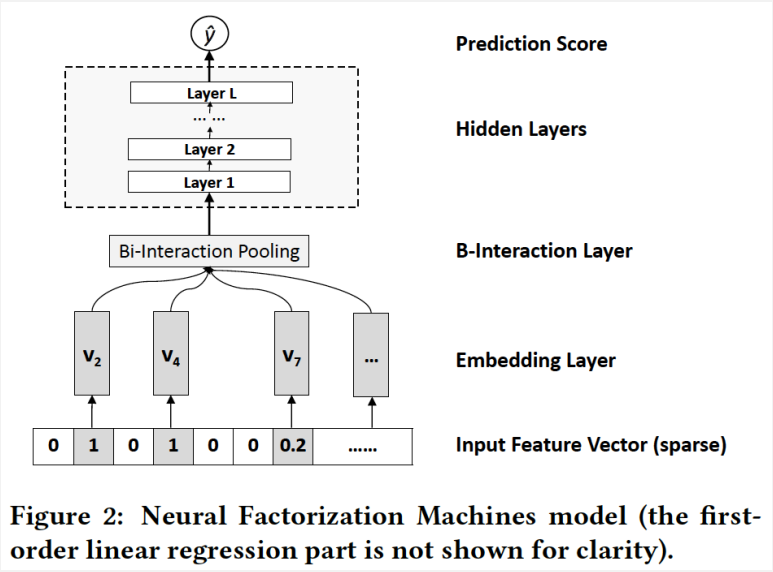
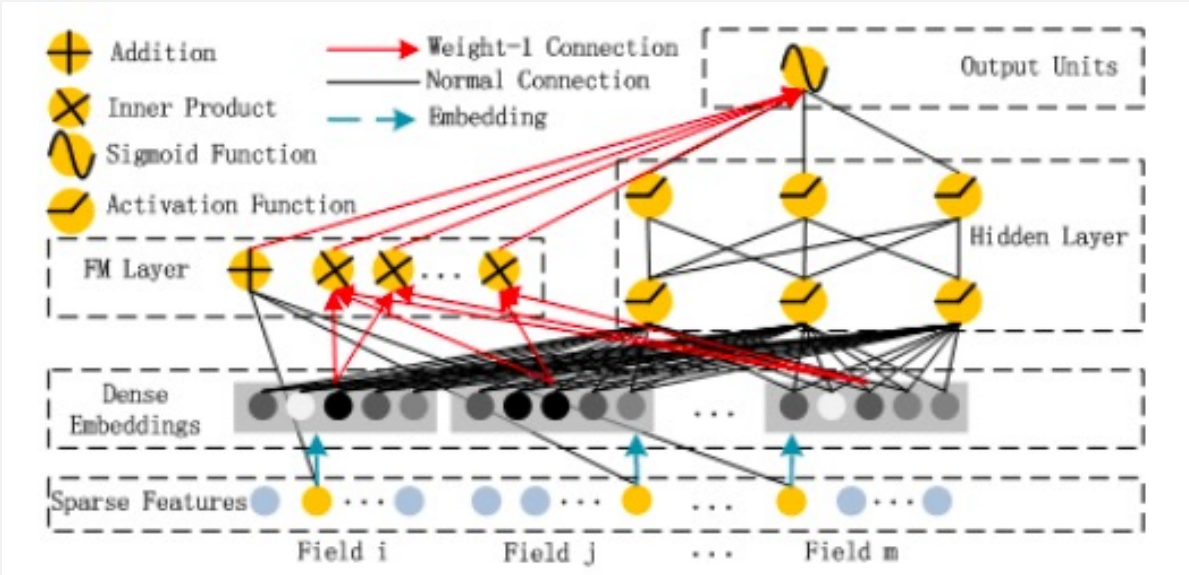


Figure 2: Neural Factorization Machines model (the first-order linear regression part is not shown for clarity).

NFM (He et al. 2017)

PARALLEL



DeepFM (Guo et al. 2017)

Rely on multiple features to make recommendations

https://github.com/shenweichen/DeepCTR-Torch

News Recommendation Methods

(-) LONG-TERM EMBEDDING

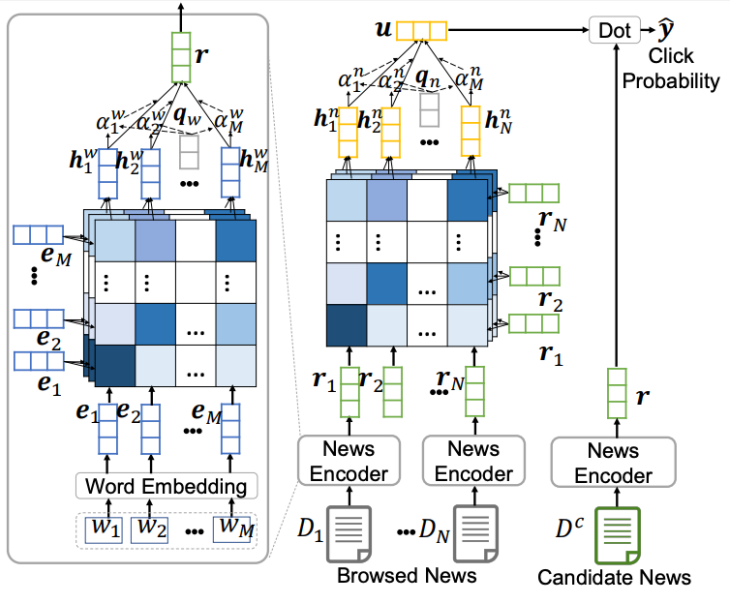
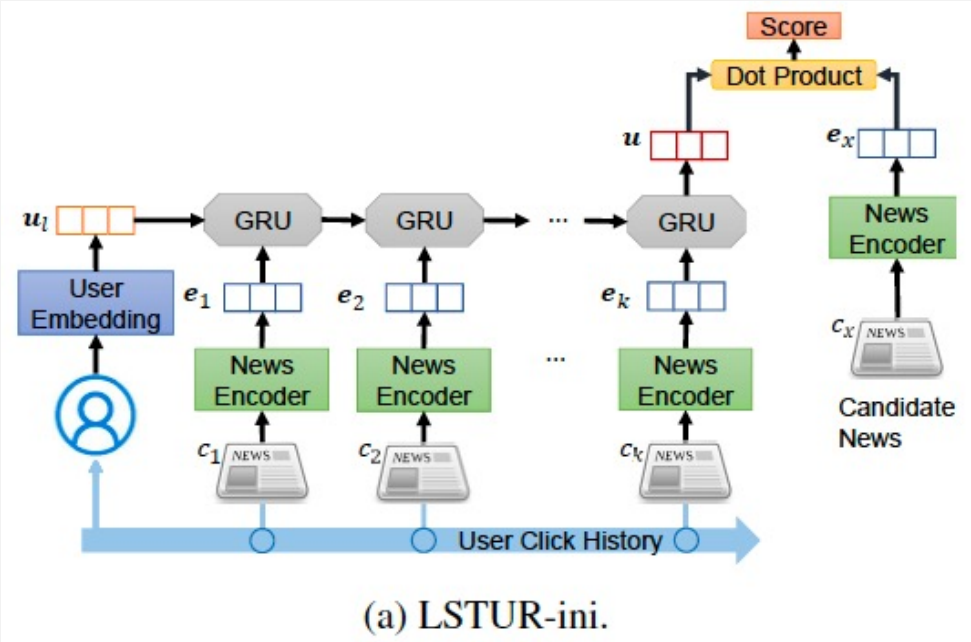


Figure 2: The framework of our *NRMS* approach.

Wu et al. 2019

(+) LONG-TERM EMBEDDING



(a) LSTUR-ini.

An et al. 2019

Developed based on news consumption data from a major news aggregator; MSN news https://github.com/microsoft/recommenders

Have we missed something?

Today…

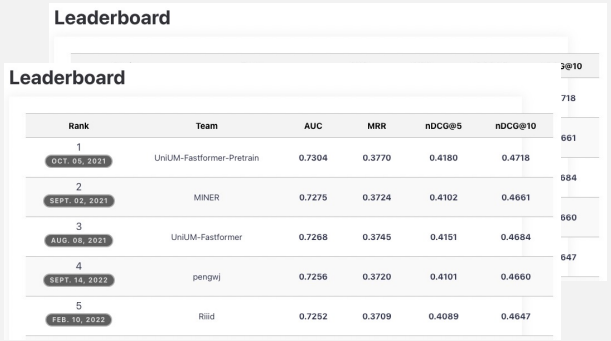
Open-source Dataset

Release: 2023

"EB-News" (dataset)

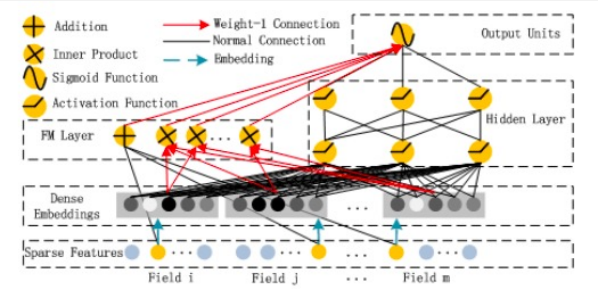
Format…

EB-Competition



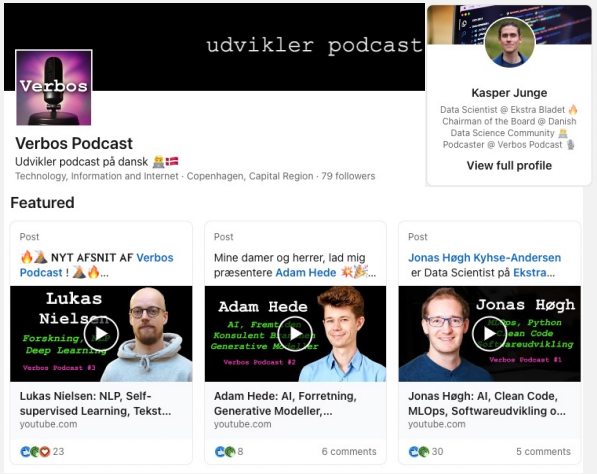
Evaluation Framework

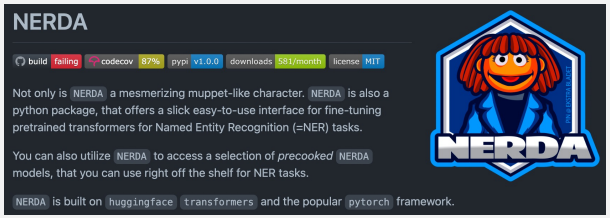
- Beyond Accuracy
  - *Diversity*
  - *Serendipity*
  - *Coverage*
  - *Popularity*
  - *Others?*

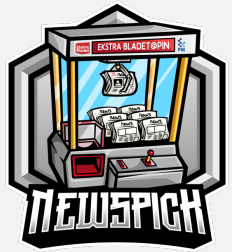


DeepFM (Guo et al. 2017)

Transparency…







Accountability…

"In the case, the Gonzalez family argues Google should be liable for the promotion of an Islamic State recruitment video by its algorithms. The video is allegedly tied to a 2015 terror attack in Paris that killed 130 people, including 23-year-old Nohemi Gonzalez."

US Supreme Court to consider recommender algorithms in key internet shield case

The U.S. Supreme Court agreed to hear a case centered on Section 230, a legal shield that protects internet platforms from civil and criminal liability for user content.
Source: U.S. Supreme Court

4 Nov, 2022 – link