

# Doporučovací systém Goodreads

Tým: Recommenders

Členové: Jan Lorenc, Marek Hlavačka, Jaromír Wysoglad

Datum: 2.5.2022

# Úkol



**Dopočítat matici hodnocení**



**Dataset: Goodreads**

Komiksy a grafické knihy  
Goodbooks-10 (Kaggle)

# Zpracování dat

## Interakce:

- user\_id, book\_id, rating

## Kniha:

- Číselné: id, průměrné hodnocení, počet hodnocení, recenzí, stran
- Textové: titul, série, autor (autoři), vydavatel, žánry, popis

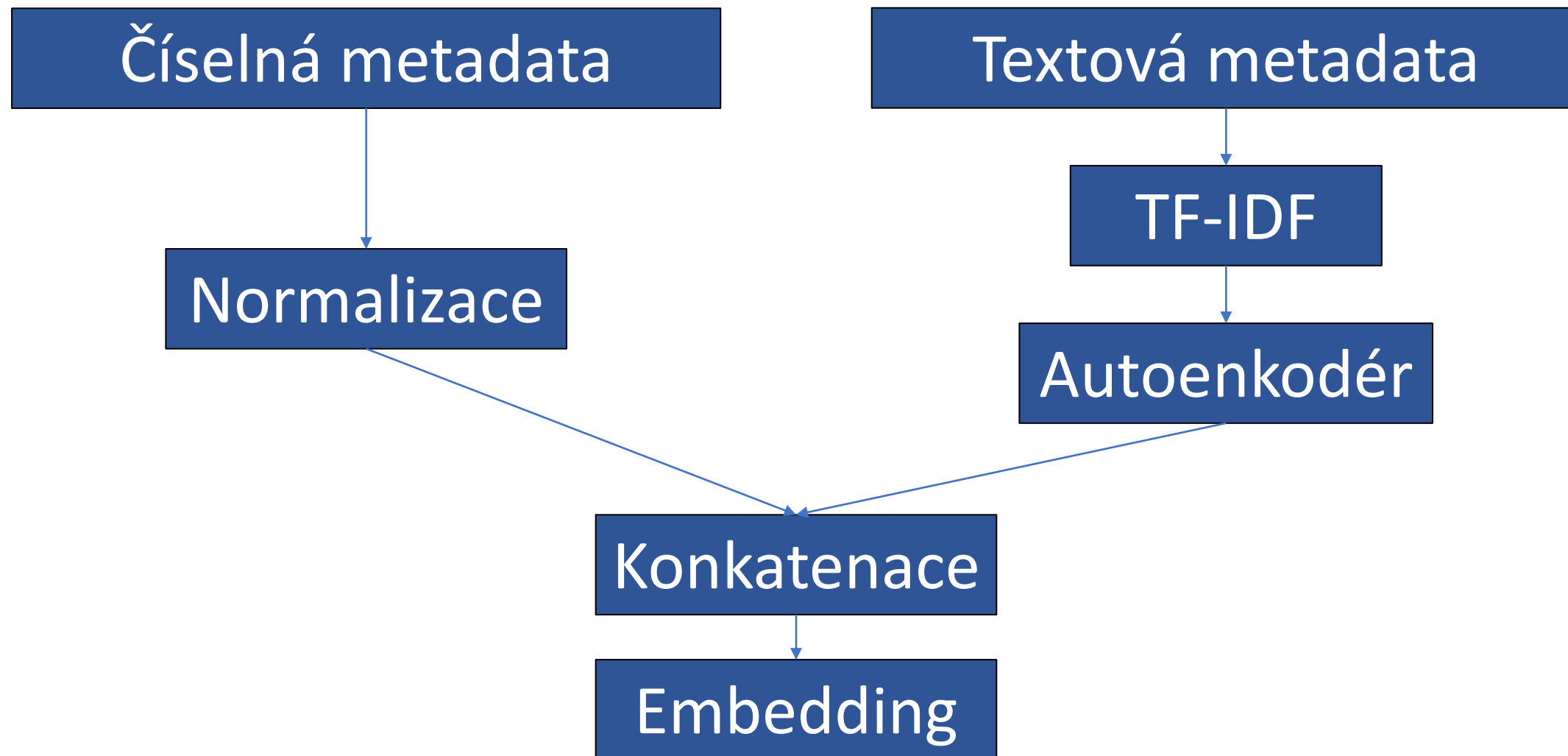
## Komiksy

- 7 347 630 interakcí
- 89 411 knih

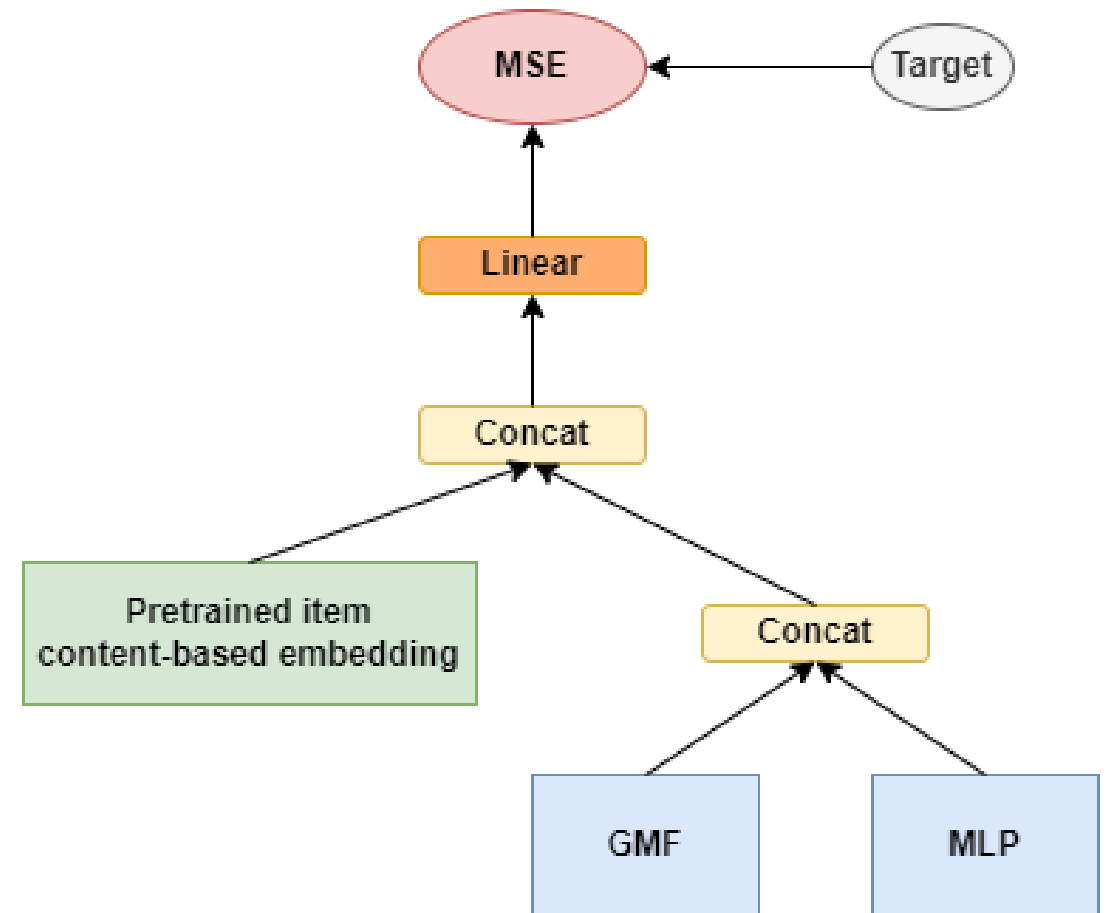
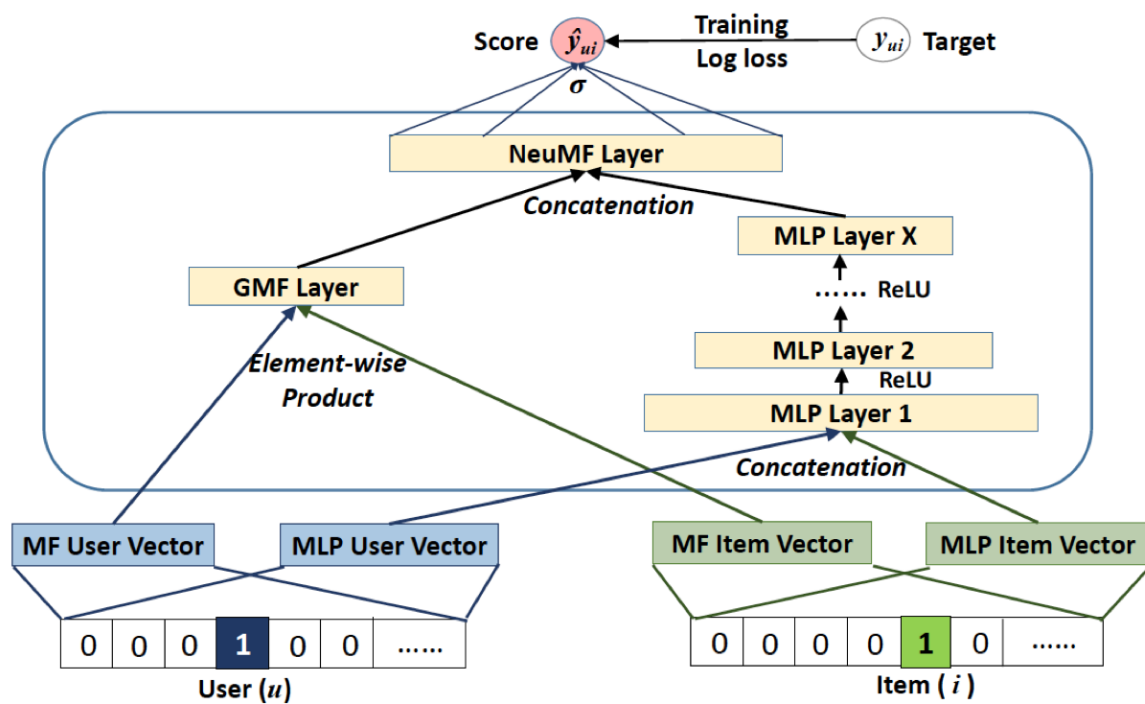
## Goodbooks-10k

- 981 756 interakcí
- 10 000 knih

# Embeddingy knih



# Model



# Experimenty

Komiksy

Model	MSE	Hits [%]
GMF	0.58789	52.078
NCF	0.54061	53.545
Hybrid	0.52594	55.853

Goodbooks

Model	MSE	Hits [%]
GMF	1.166043	38.422
NCF	0.732287	43.63
Hybrid	0.713658	46.03

Srovnání

Model	MSE
MF	2.944
<b>GMF</b>	<b>1.166043</b>
NN_MF	0.9804
Surprise SlopeOne	0.848977
Spark MF ALS	0.80393
Surprise KNNBasic	0.799593
Surprise CoClustering	0.778982
NN_1	0.766163
NN_2	0.750598
<b>NCF</b>	<b>0.732287</b>
Surprise FunkSVD	0.718087
<b>Hybrid</b>	<b>0.713658</b>