

Maja Jabłońska

Rekurencyjne sieci neuronowe i systemy rekomendacji

Core #6



Tradycyjne podejścia

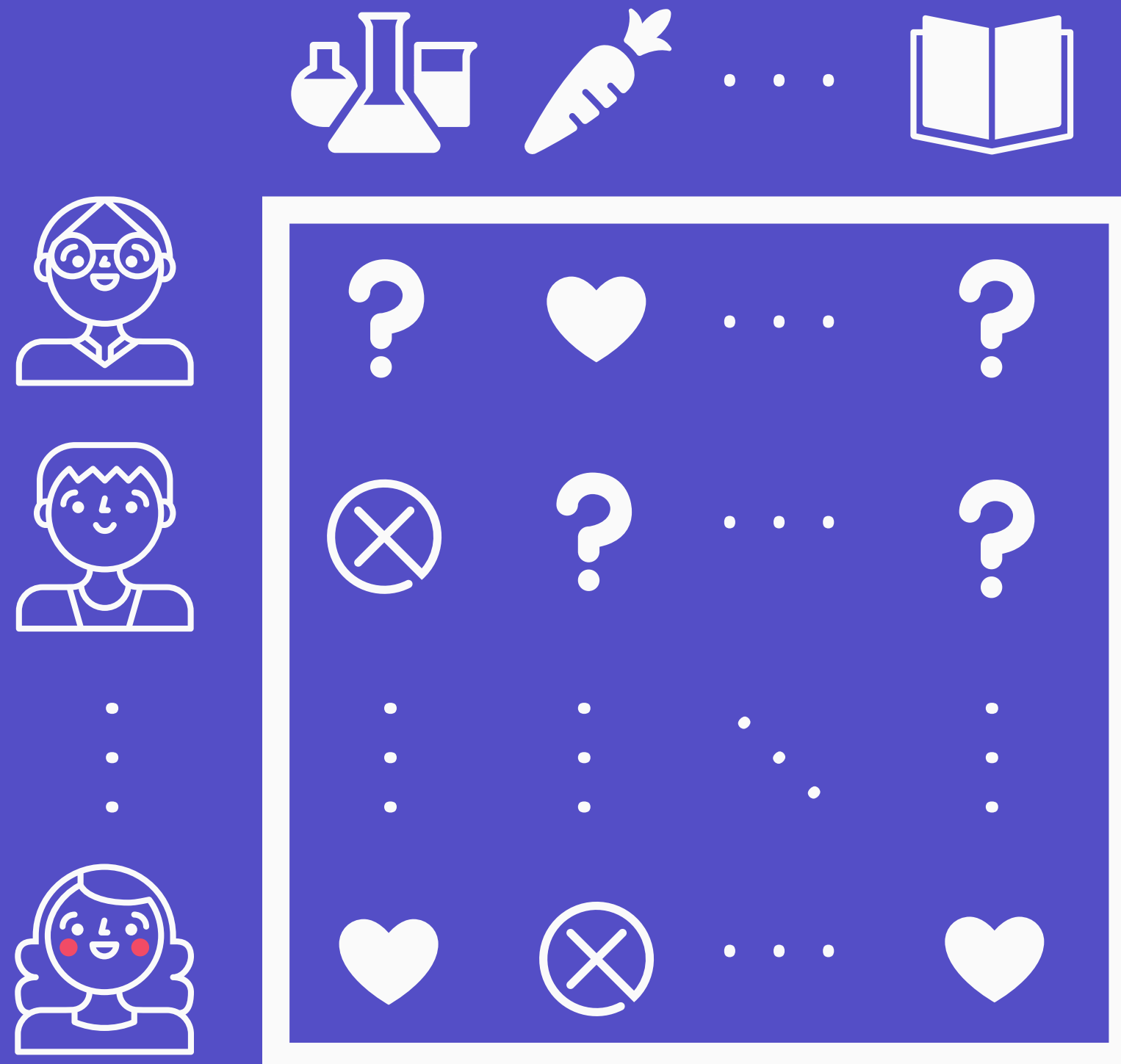
01

Collaborative Filtering

02

Content-based Filtering

Collaborative Filtering



Jedynymi danymi są
interakcje użytkowników i
produktów



Memory based

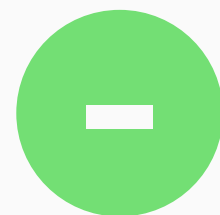
- Duże macierze rzadkie reprezentujące interakcje użytkownik-produkt
- np. KNN

Model based

- Ukryty model reprezentujący gusta użytkowników lub cechy produktów
- Model w postaci mniejszej, gęstszej macierzy
- np. faktoryzacja macierzy, FCNN

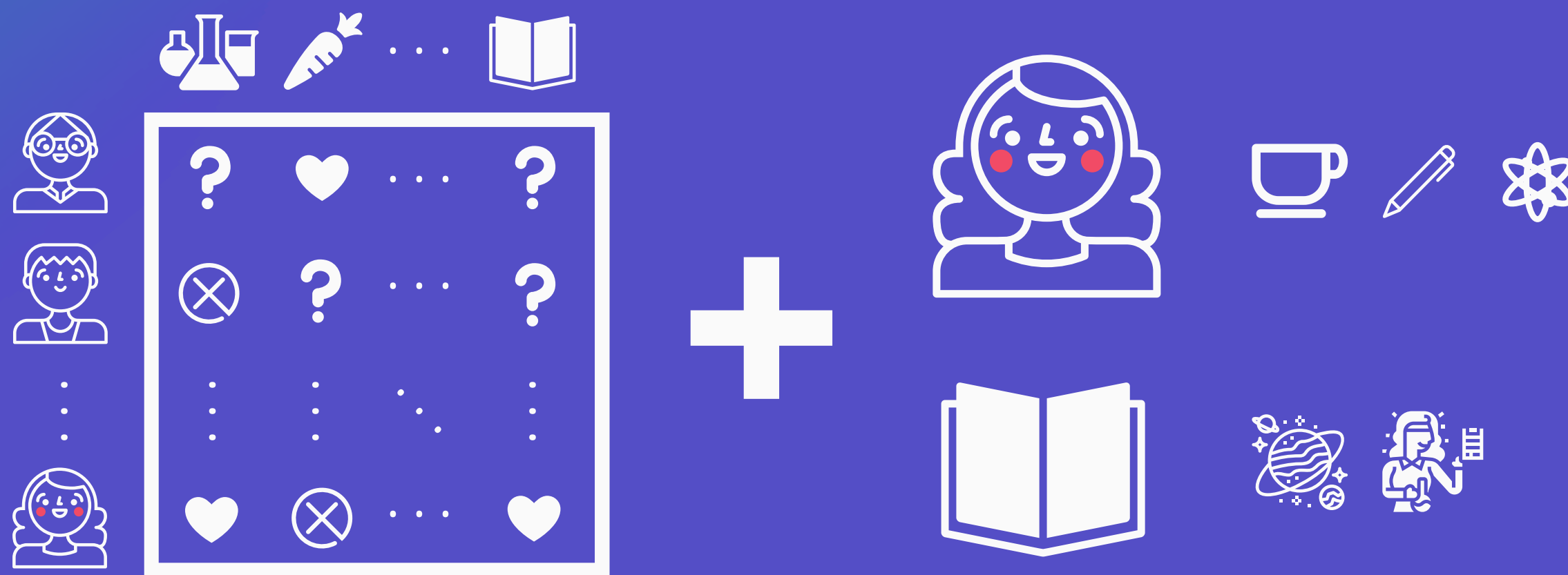


Nie potrzeba żadnych dodatkowych informacji poza historią interakcji



Problemy z wprowadzaniem nowych użytkowników lub produktów- brak danych o ich interakcjach

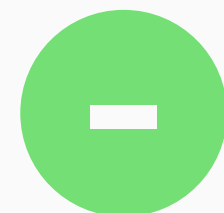
Content based



Dane o poprzednich interakcjach + cechy produktów i użytkowników

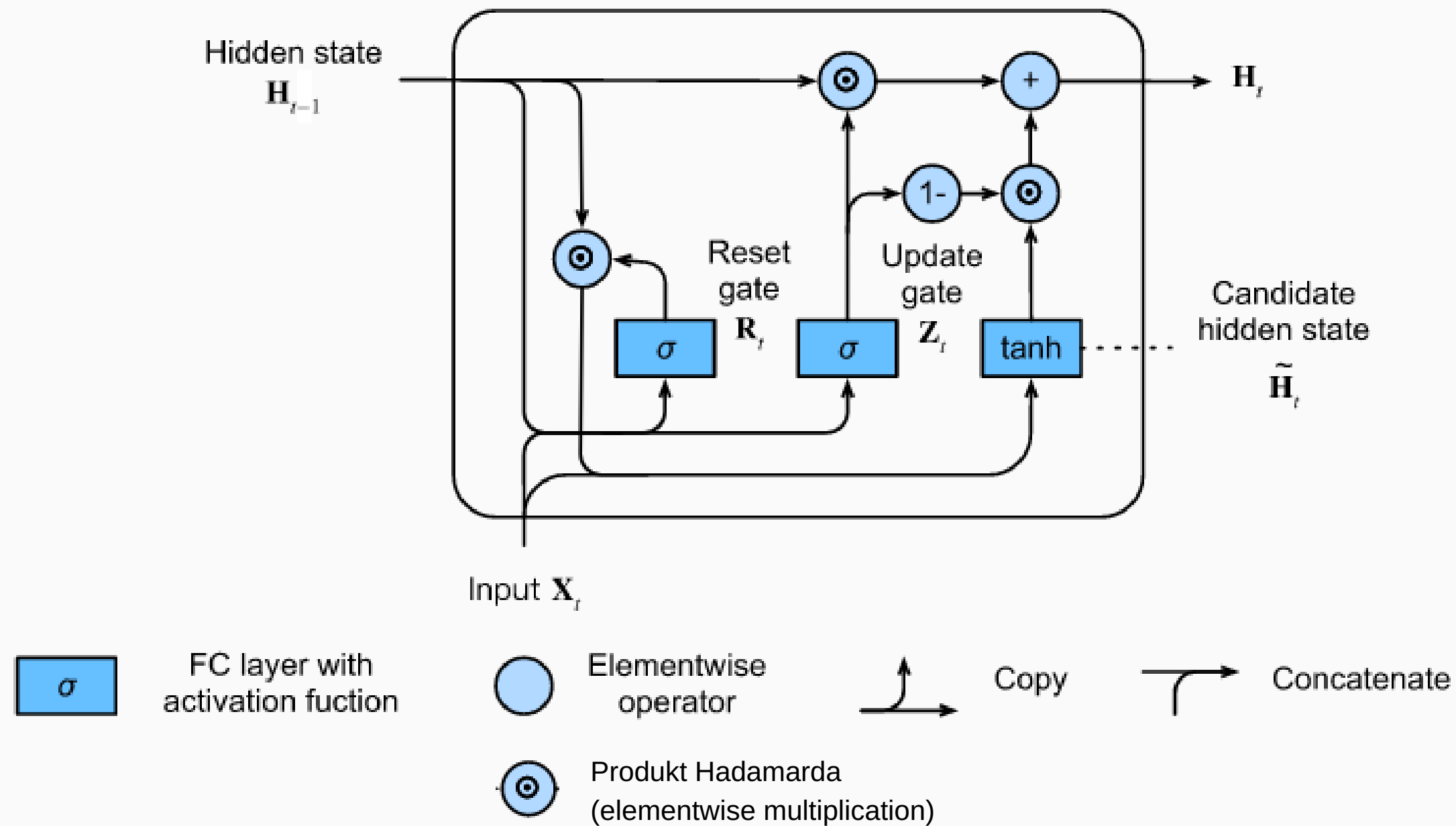


Dużo mniejszy niż w przypadku collaborative filtering problem z dodawaniem nowych użytkowników lub produktów



Ryzyko zbyt dużej specjalizacji (użytkownikowi nie są nigdy rekomendowane przedmioty spoza danej kategorii)

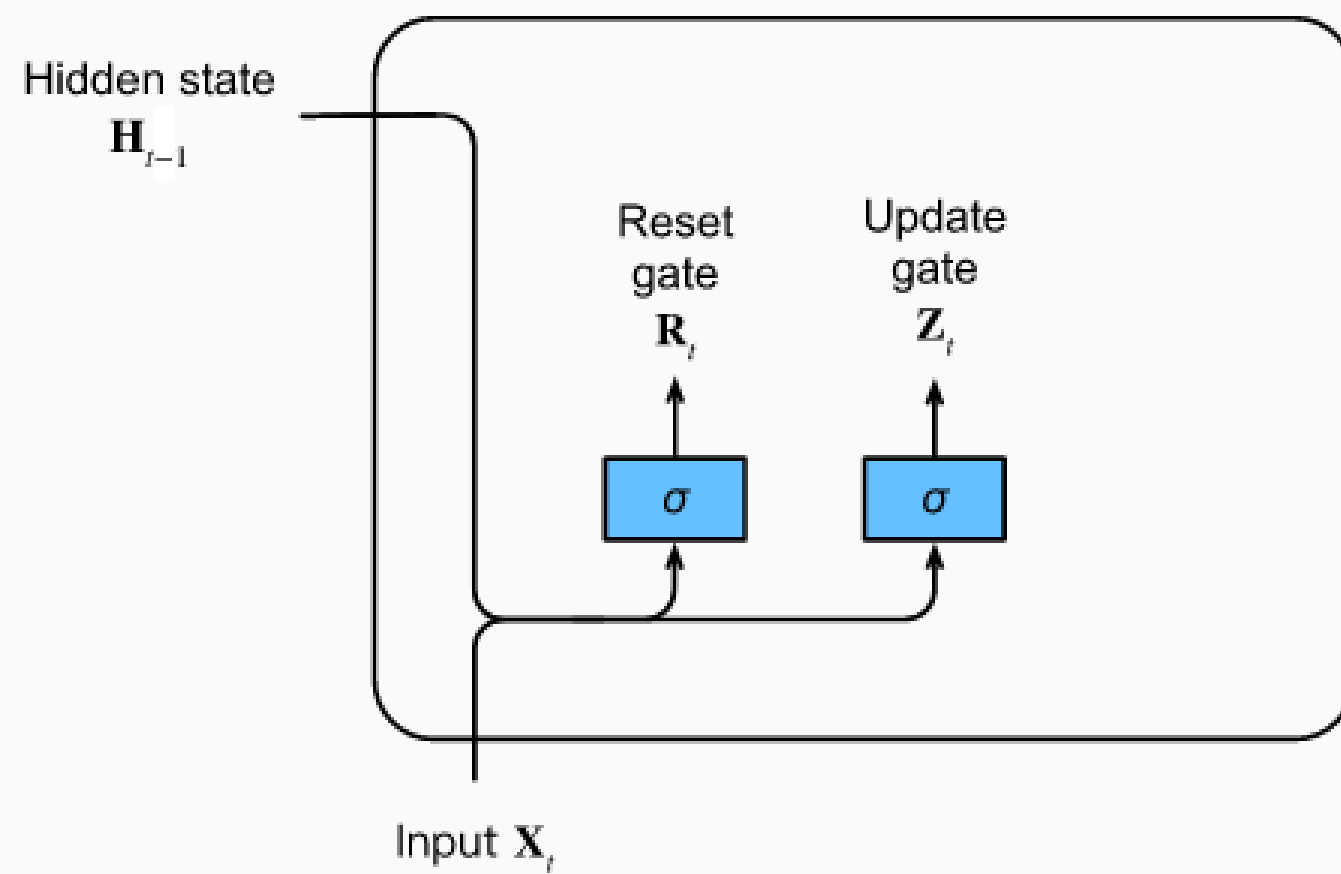
Gated Recurrent Unit



Gated Recurrent Unit

- Like LSTM, but smoler and faster
- ... ale mniej potężne
- Dwie bramki (gates):
 - Update gate kontroluje ile z obecnego stanu ukrytego zostanie przekazane do następnej komórki
 - Reset gate kontroluje, które dane z poprzednich kroków chcemy zapomnieć (kontroluje informacje wychodzące z pamięci)





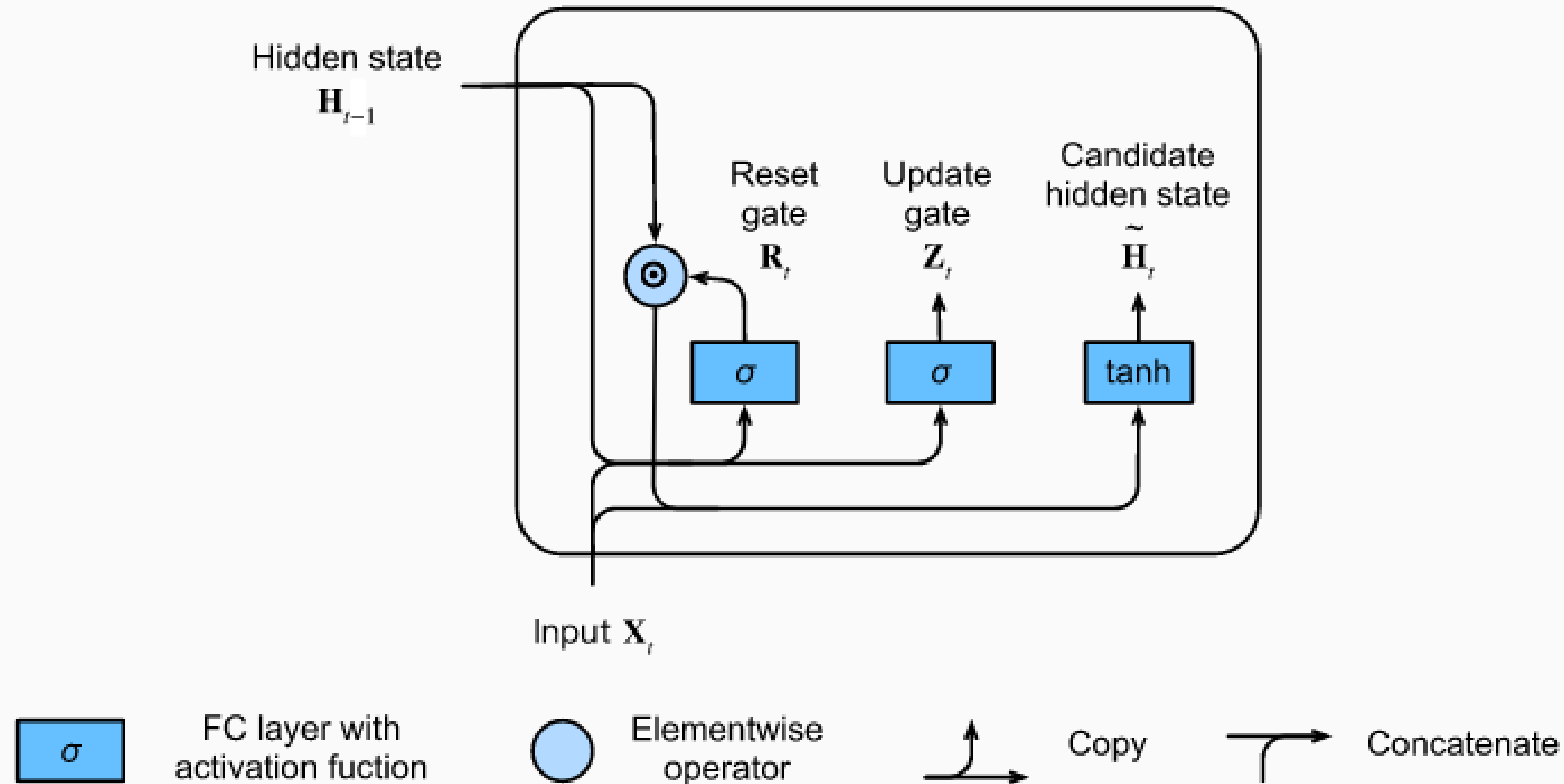
$$\mathbf{R}_t = \sigma(\mathbf{X}_t \mathbf{W}_{xr} + \mathbf{H}_{t-1} \mathbf{W}_{hr} + \mathbf{b}_r),$$

$$\mathbf{Z}_t = \sigma(\mathbf{X}_t \mathbf{W}_{xz} + \mathbf{H}_{t-1} \mathbf{W}_{hz} + \mathbf{b}_z).$$

$$\left. \begin{array}{l} \mathbf{W}_{xr}, \mathbf{W}_{xz} \in \mathbb{R}^{d \times h} \\ \mathbf{W}_{hr}, \mathbf{W}_{hz} \in \mathbb{R}^{h \times h} \end{array} \right\} \text{wagi}$$

$$\mathbf{b}_r, \mathbf{b}_z \in \mathbb{R}^{1 \times h} \quad \text{bias}$$

Reset gate



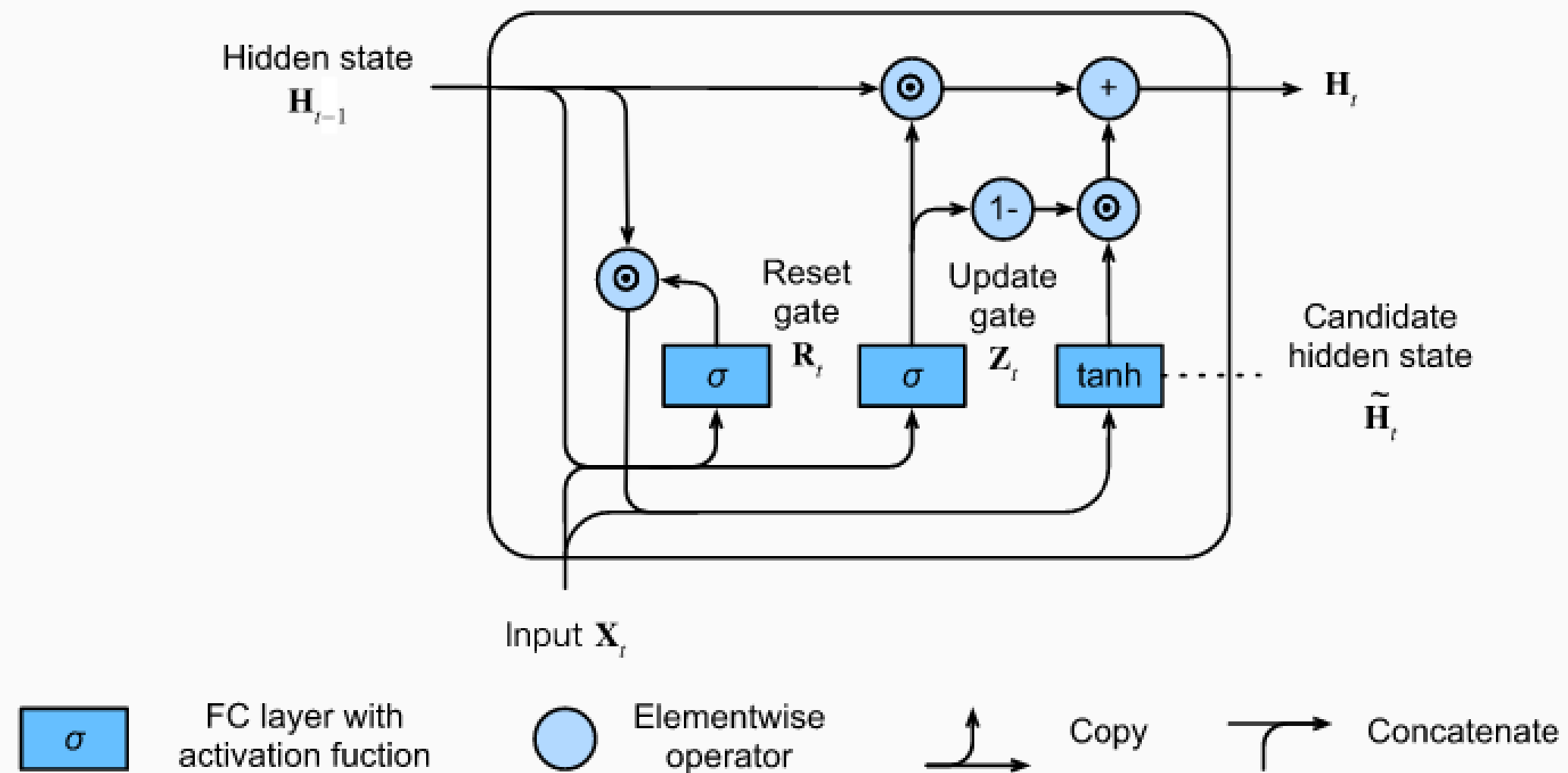
$$\mathbf{H}_t = \tanh(\mathbf{X}_t \mathbf{W}_{xh} + \mathbf{H}_{t-1} \mathbf{W}_{hh} + \mathbf{b}_h)$$

Nowy stan ukryty w konwencjonalnej RNN

$$\tilde{\mathbf{H}}_t = \tanh(\mathbf{X}_t \mathbf{W}_{xh} + (\mathbf{R}_t \odot \mathbf{H}_{t-1}) \mathbf{W}_{hh} + \mathbf{b}_h)$$

Kandydat na nowy stan ukryty (czeka go jeszcze przejście przez update gate)

Update gate



$$\mathbf{H}_t = \mathbf{Z}_t \odot \tilde{\mathbf{H}}_t + (1 - \mathbf{Z}_t) \odot \mathbf{H}_{t-1}$$

Nowy stan ukryty

Rekomendacje oparte na sesjach

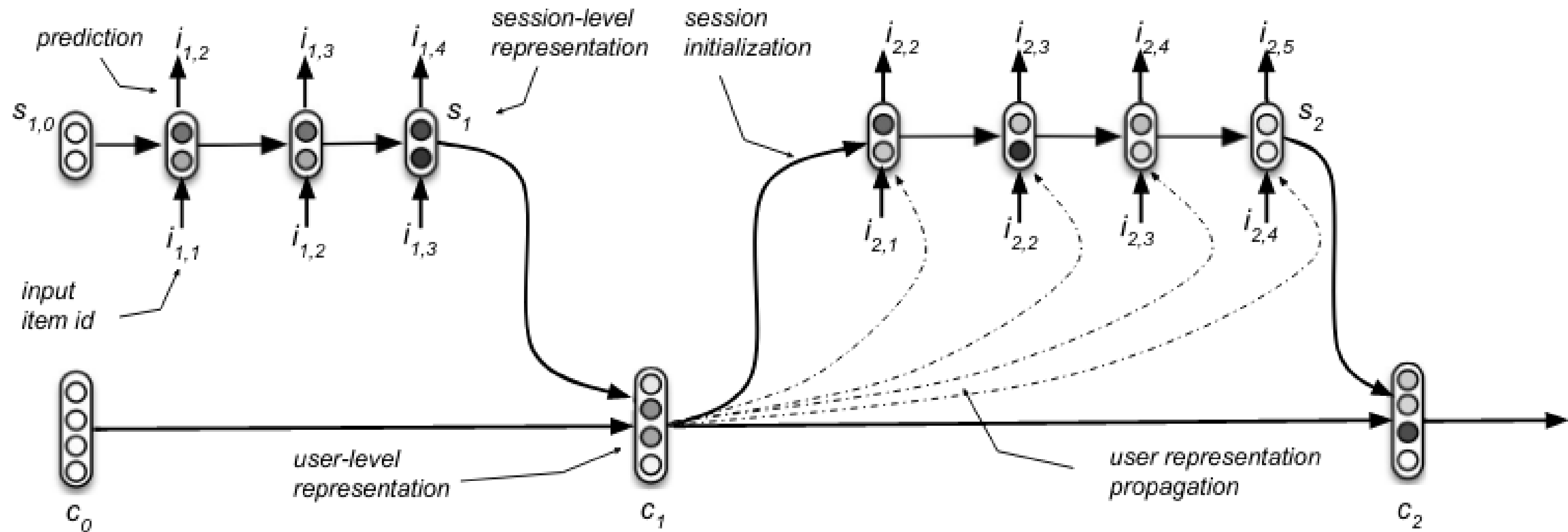
01

Session-based recommenders
użytkownik jest anonimowy

02

Session-aware recommenders
mamy dostęp do poprzednich
sesji użytkownika

Architektura sieci





HRNN Init

output warstwy użytkownika
wyłącznie inicjalizuje warstwę
następnej sieci

HRNN All

output warstwy użytkownika jest
propagowany do każdego kroku
w warstwie sesji

*(co skutkuje większą
złożonością modelu)*

Warianty

Zbiory danych



XING

interakcje z ofertami
pracy



VIDEO

historia oglądania
filmików (podobne do
Youtube)

Modele bazowe

01

PPOP (Personal POP)

02


Item-KNN

03

Session based RNN

04

Session based RNN Concat



● **Recall@5**
(hit-rate metric)
proporcja
przypadków w
których
relewantny
przedmiot znalazł
się w top 5 do
wszystkich
przypadków

● **Precision@5**
proporcja
prawidłowych
rekomendacji w
top 5

● **MRR@5**
ranking
relewantnego
przedmiotu
(odpowiednia miara
dla systemu w którym
kolejność
rekomendacji jest
istotna)

Wyniki

		XING			VIDEO		
		Recall@5	MRR@5	Precision@5	Recall@5	MRR@5	Precision@5
small	Item-KNN	0.0697	0.0406	0.0139	0.4192	0.2916	0.0838
	PPOP	0.1326	0.0939	0.0265	0.3887	0.3031	0.0777
	RNN	0.1292	0.0799	0.0258	0.4639	0.3366	0.0928
	RNN Concat	0.1358	0.0844	0.0272	0.4682	0.3459	0.0936
	HRNN All	<u>0.1334[†]</u>	0.0842	<u>0.0267[†]</u>	<u>0.5272</u>	<u>0.3663</u>	<u>0.1054</u>
	HRNN Init	<u>0.1337[†]</u>	0.0832	<u>0.0267[†]</u>	<u>0.5421</u>	<u>0.4119</u>	<u>0.1084</u>
	RNN	0.1317	0.0796	0.0263	0.5551	0.3886	0.1110
	RNN Concat	0.1467	0.0878	0.0293	0.5582	0.4333	0.1116
large	HRNN All	<u>0.1482[†]</u>	<u>0.0925</u>	<u>0.0296[†]</u>	0.5191	0.3877	0.1038
	HRNN Init	<u>0.1473[†]</u>	<u>0.0901</u>	<u>0.0295[†]</u>	<u>0.5947</u>	<u>0.4433</u>	<u>0.1189</u>

	VIDEOXXL		
	Recall@5	MRR@5	Precision@5
RNN	0.3415	0.2314	0.0683
RNN Concat	0.3459	0.2368	0.0692
HRNN All	0.3621	0.2658	0.0724
HRNN Init	0.4362	0.3261	0.0872

źródła

Personalizing Session-based Recommendations with Hierarchical Recurrent Neural Networks

<https://arxiv.org/pdf/1706.04148.pdf>

Gated Recurrent Units

https://d2l.ai/chapter_recurrent-modern/gru.html#reset-gates-and-update-gates

Understanding GRU Networks

<https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be>

Introduction to recommender systems

<https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada>

