



## СИСТЕМИ ЗА ПРЕПОРУКУ У *E-COMMERCE* ПЛАТФОРМАМА ЗАСНОВАНИ НА ГРАФ НЕУРОНСКИМ МРЕЖАМА

### RECOMMENDATION SYSTEMS IN E-COMMERCE PLATFORMS BASED ON GRAPH NEURAL NETWORKS

Матија Матовић, Јелена Сливка, *Факултет техничких наука, Нови Сад*

#### Област – РАЧУНАРСТВО И АУТОМАТИКА

**Кратак садржај** – У области е-трговине, велика количина новца се годишње улаже у интернет маркетинг. Циљ ових улагања је да се производи представе што ширем обиму корисника, како би они могли бити приказани што већем броју људи који би за те производе били заинтересовани. Употребом система за препоруку, ови трошкови би се могли смањити, а профит повећати, тако што би се сваком кориснику приказивале персонализоване препоруке, детаљно прилагођене његовом укусу. Са тим у вези овај рад истражује неколико различитих приступа имплементацији система за препоруку: традиционални приступ и приступи засновани на модерној архитектури граф неуронских мрежа. За учење су прикупљени подаци о интеракцијама корисника са ставкама, као и подаци о самим корисницима и ставкама. Коначан циљ овог рада је испитати квалитет модерних метода заснованих на граф неуронским мрежама, упоредити га са традиционалним методама и истражити њихове досад недовољно испитане предности.

**Кључне речи:** граф неуронске мреже, системи за препоруку, е-трговина, факторизација матрица.

**Abstract** – In e-commerce, a large amount of money is invested annually in internet marketing. These investments aim to present products to a broader range of users so that they can be shown to more people interested in those products. A recommendation system could reduce these costs and increase profit by displaying personalized recommendations tailored to each user's taste. In this regard, this paper explores several approaches to implementing a recommendation system: a traditional approach and approaches based on the modern architecture of graph neural networks. Data on user interactions with items and data on users and items themselves were collected for learning. This paper aims to examine the quality of modern methods based on graph neural networks, compare them with traditional methods, and explore their underexplored advantages.

**Keywords:** graph neural networks, recommendation systems, e-commerce, matrix factorization

#### НАПОМЕНА:

Овај рад произтекао је из мастер рада чији ментор је била др Јелена Сливка, ванр. проф.

#### 1. УВОД

У непрестано растућем пољу е-трговине, системи за препоруку су постали незаменљиви алати за унапређење корисничког искуства и оптимизацију пословних перформанси. Ови систем заузимају кључну улогу у вођењу корисника кроз обимну понуду производа и услуга, нудећи персонализоване препоруке које се уклапају са индивидуалним укусима и понашањем. Економски значај ефикасних система за препоруку се не сме занемарити, јер не само што доприносе ангажовању (*engagement*) и задовољству корисника, већ и значајно утичу на приходе пословања промовишући релевантне ставке и подстичући понављање трансакција и куповина корисника на датој платформи [1]. Са технолошким напретком и порастом доступности података, потенцијал за коришћење софистицираних алгоритама и иновативних архитектура, као што су граф неуронске мреже (GNN, *Graph neural networks*) [2] представља одличну прилику за значајно побољшавање тачности и корисности система за препоруку.

Мотивисан дубоким утицајем који системи за препоруку имају у пољу е-трговине, овај рад истражује и имплементира најновије приступе препоручивању, са посебним нагласком на GNN. GNN су добиле на значају због своје неупоредиве способности да науче сложене односе и зависности присутне у комплексним структурама података, нарочито у случајевима где се подаци и њихове интеракције могу ефикасно моделовати као графови. Искоришћавањем предности GNN-ова, циљ је унапредити тачност и ефикасност препорука инкорпорирањем нијансираног разумевања интеракција корисник-садржај у виду мреже (графа) интеракција. Овај рад истражује област система за препоруку, оцењује ефикасност GNN као префериране архитектуре и истражује њихов потенцијал да преобликује област препорука у савременој е-трговини. Кроз компаративну анализу са традиционалним методама, овај рад настоји да отвори пут за унапређење стратегије препорука које одговарају динамичним потребама савремених потрошача.

У раду се врши прикупљање података о трансакцијама и оценама ставки од стране корисника на неком од доступних скупова података и потом додатно обогаћивање података релевантим својствима. Потом се над тим подацима тренирају модели засновани на традиционалним методама, као и модели засновани на савременој GNN архитектури, како би се њихови резултати поредили, и анализирали предности и мане.

У наредном поглављу дат је преглед претходних решења која су се бавила системима за препоруку. У поглављу 3 су дати детаљи решења и имплементације система, а њихови резултати су евалуирани и поређени у поглављу 4. Коначно, 5. поглавље представља закључак рада.

## 2. ПРЕТХОДНА РЕШЕЊА

У контексту машинског учења, први приступ изради система за препоруку се може пронаћи у системима заснованим на садржају (*content-based filtering*) [3]. Њиховом појавом се десила велика промена у дотадашњем начину израде ових система, јер се први пут показало да рачунар може сам да научи или искористи корисничке укусе, без потребе за системима базираним на правилима која су писали експерти. Ови системи функционисали су тако што су анализирали својства и атрибуте ставки са којима је корисник интераговао у прошлости, и на основу њих налази сличне и препоручује их кориснику. Њихов главни проблем је био што нису могли да кориснику препоручују нешто изван онога на шта је он досад навикао, као и проблем издвајања оних својстава која би била најрелевантнија за предикцију. Такође, нису могле ни да искористе ставке које су се свиђале сличним корисницима.

Након њих се појављују системи засновани на колаборативном филтрирању и факторизацији матрица [4]. Ови системи анализирају матрицу интеракција корисник-ставка, и потом, нашавши по интеракцијама најсличније кориснике, текућем кориснику препоручују ставке које су слични корисници најбоље оценили.

У скорије време, до појаве система заснованих на дубоком учењу, најчешће су се користили хибридни модели, који су своје предикције износили комбинујући гласове више модела заснованих на претходне две методе, тиме надокнађујући недостатке система заснованих на садржају у виду неискоришћавања интеракција сличних корисника, као и недостатке система заснованих на колаборативном филтрирању у виду проблема са ново-додатим ставкама [5].

Појавом граф неуронских мрежа [2], брзо се увиђа корисност ове архитектуре у системима за препоруку, с обзиром на то да се интеракције корисника и ставки могу представити у виду бипартитног хетерогеног графа, где су корисници и ставке чворови са својим својствима, а интеракције између њих гране графа. Једна од ранијих и најуспешнијих примена је GraphSAGE модел [6], који се показао као јако ефикасно решење, и који користи *Pinterest*. NGCN (*Neural Graph Collaborative Filtering*) [7] је нешто сложенији модел који показује још боље резултате на свим релевантним скуповима података, али уз цену веће рачунарске сложености. Такође, значајан је и HGAT (*Heterogenous Graph Attention Network*) [8], који користи и мета-путање (два корисника су повезана уколико су интераговала са истом ставком нпр.), како би извукао додатне зависности и шаблоне понашања вишег нивоа, и који је резултате боље од NGCN на релевантним скуповима података.

Овај рад ће представити решење засновано на GraphSAGE због његове мале рачунарске сложености,

као и решење засновано на HGAT због побољшања које мета-путање пружају.

## 3. МЕТОДОЛОГИЈА

У наредном поглављу је описан скуп података коришћен у овом раду, начин на који је извршена експлоративна анализа података и архитектура испробаних модела за предикцију.

### 3.1. Скуп података

За евалуацију решења је коришћен *OList Brazilian E-commerce* [9] скуп података. Садржи информације о преко 110.000 трансакција, 32.000 производа и 99.000 корисника. Информације које садржи о производима су бразилска савезна држава и град у којој је корисник регистрован. Информације које садржи о производима укључују њихове димензије (дужина, ширина и висина), тежина, цена, категорија (којих има 71), као и дужина наслова и описа. Због анонимизације података, текст наслова и описа производа су замењени њиховом дужином. Такође, скуп података садржи и податке о свакој трансакцији: датум и време трансакције, количина сваког од производа, као и оцена коју је корисник дао својој наруџбини након њеног извршавања.

### 3.2. Експлоративна анализа података

Анализом података је утврђено да већина поруџбина садржи само једну ставку, односно, да је просечан број ставки по поруџбини 1,2. Такође, утврђено је да су оцене 4 и 5 несразмерно распрострањене у скупу података, односно, убедљиво највећи број поруџбина има додељене оцене 5 и 4.

Ове две ставке представљају отежавајућу околност за моделе који само користе интеракције корисник-ставка, као што су то управо модели засновани на факторизацији матрица. Разлог за ово је тај што нема довољно интеракција по особи како би се могло квалитетно закључивати, као и то што је већина оцена 5, па је скуп небалансиран. Модели засновани на факторизацији матрица искључиво користе матрицу интеракција (тј. оцене корисника), а како је квалитет ових података у скупу лош због поменутих разлога, ови модели ће имати лошије перформансе.

### 3.3. Претпроцесирање података

Подаци о корисницима су због анонимизације прилично шкрти. Због тога су додатно обогаћени метрикама које могу дати наговештај о имовинском стању и куповним навикама корисника. С обзиром да су били доступни подаци о географској локацији сваког корисника, подаци су обогаћени вредностима као што су: индекс људске развијености (HDI - *Human Development Index*), бруто домаћи производ по глави становника сваке државе у којој се корисник налази и густина насељености сваког од градова у коме се корисник налази.

Ови бројеви су добар показатељ куповне моћи и навика становника, јер указују на куповну моћ, развијеност инфраструктуре и доступност интернету и другим добрима [10]. Нумеричке вредности о БДП, HDI, цени производа и густини насељености су потом

нормализовани. Категорије производа су из текста категоризовани у бројеве.

Подаци о интеракцијама корисник-ставка су након иницијалне обраде претворени у формате које користе модели разматрани у овом раду. За модел заснован на факторизацији матрица подаци су претворени у матрицу интеракција корисник-ставка; редови су корисници, а колоне ставке, у пољима се налазе оцене које је корисник дао ставки.

За моделе засноване на GNN, подаци су претворени у бипартитни хетерогени граф; корисници и ставке су два типа чворова који садрже својства и информације, док су гране између њих оцена коју је корисник дао ставци.

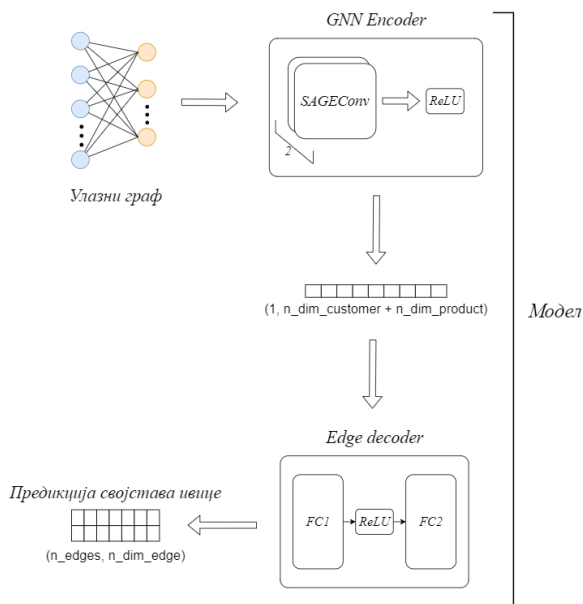
### 3.4. Архитектура модела

Модел заснован на факторизацији матрица се састоји од два дела: први део предвиђа основицу оцене на основу одступања просечне оцене корисника ( $b_x$ ) и просечне оцене ставке ( $b_i$ ) од глобалног просека ( $\mu$ ). Други део тражи по оценама најсличније ставке текућој, и рачуна отежињени просек тих ставки да би додатно поправио предикцију основице.

Отежињавање при израчунавању оцене се врши помоћу сличности две ставке ( $S_{ij}$ ), које се добија израчунавањем Жакаровог индекса, као што је описано у [11]. Коначна предикција се добија комбиновањем ових елемената као што је приказано у формули (1).

$$r_{xi} = \mu + b_x + b_i + \frac{\sum_{j \in N(x,i)} S_{ij} \cdot (r_{xj} - b_{xj})}{\sum_{j \in N(x,i)} S_{ij}} \quad \#(1)$$

У склопу овог рада имплементирана су два модела заснована на граф неуронским мрежама. Први модел се састоји од једног енкодера који прима текући чвор и матрицу суседства како би направило стабло израчунавања и на основу њега конструисало латентну репрезентацију чвора (Слика 1).

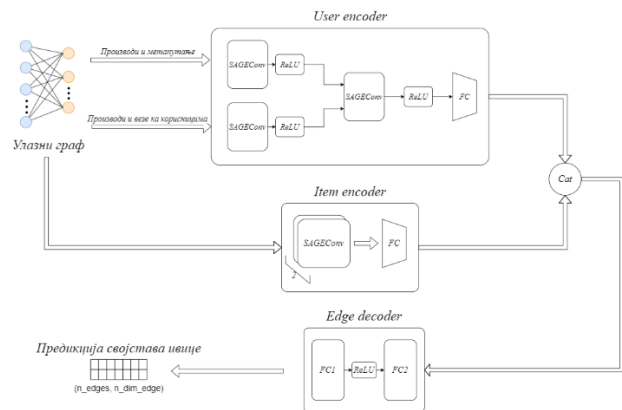


Слика 1. Дијаграм архитектуре модела са једним енкодером за обе врсте чворова

С обзиром на то да се састоји од два SAGEConv [12] слоја, стабло израчунавања је дубине два. Овај модел

има и декодер ивица. То је двослојни перцептрон који прима конкатениране латентне репрезентације производа и корисника и на основу њих даје предикцију оцене.

Други модел се састоји од два енкодера (Слика 2). С обзиром на то да у графу постоје мета-путање [8] између сродних производа, латентне репрезентације ова два типа чворова се морају другачије израчунати. Оба енкодера примају чвор и матрицу суседства, с тим што енкодер корисника додатно обрађује репрезентацију производа пропуштајући је кроз додатни граф-конволуциони слој са само подацима о мета-повезаности. Добијене латентне репрезентације користе на исти начин као и једно-енкодерски модел.



Слика 2. Дијаграм архитектуре модела са два енкодера и мета-путањима

## 4. РЕЗУЛТАТИ И ДИСКУСИЈА

Скуп података је подељен на тренинг и тест скуп у размери 80:20. Међутим, код система за препоруку, тренирање је најчешће трансдуктивно. То значи да је скуп чворова и у тренинг и у тест скупу исти, док се само интеракције деле у размери 80:20. Ово је реалнији пример примене, јер се системи за препоруку обично тренирају специфично за одређену платформу, на подацима који у њој већ постоје. Међутим, док модел заснован на факторизацији матрица пати од проблема увођења нових чворова, јер за њих нема довољно интеракција да би предвиђао оцене, модели засновани на граф неуронским мрежама немају већих проблема да нове податке пропусти кроз истрениране енкодере и добију латентне репрезентације.

Евалуација се вршила помоћу отежињене средње грешке квадрата (WMSE – *Weighted Mean Squared Error*) због чињенице да је скуп података изузетно небалансиран, те је потребно више кажњавати грешке на мање заступљеним оценама. Најбоље резултате су показали модели засновани на граф неуронским мрежама, што је и очекивано с обзиром на то да су посебно намењени да што боље уче над подацима структурираним у виду графа (Табела 1). Такође, за разлику од традиционалних модела способни су да користе и податке о тополошкој структури графа, као и податке које садрже сами чворови. Још једна предност коју ови модели имају над факторизацијом матрица је то, што је предикција много бржа. Док је

код факторизације матрица за сваку предикцију сваки пут потребно проћи кроз матрицу и наћи најсличније ставке, код GNN модела потребно је само корисника и ставку пропустити кроз већ истрениране енкодере како би се добила предикција.

Модел који користи мета-путање даје нешто боље резултате од обичног ГНН модела, с обзиром на то да може да искористи зависности вишег нивоа које се могу налазити у овим метапутањима.

Табела 1. Постигнути резултати са сваком од модела. Mat-Fac је модел са факторизацијом матрица, а 1-enc и 2-enc meta-path су модели са 1 и 2 енкодера

Модел	Test WMSE
Mat-Fac	1.97
1-enc	1.21
2-enc metapath	<b>0.98</b>

## 5. ЗАКЉУЧАК

Овај рад се бави анализом употребљивости нове архитектуре граф неуронских мрежа у проблемима са великом могућношћу примене као што су системи за препоруку, и пореди их са до сада често коришћеним, традиционалним приступима решавања ових проблема.

Мотивацију за овај рад је представљала чињеница да су системи за препоруку од великог значаја за било коју платформу засновану на садржају, поготово у области е-трговине, и да у великој мери доприносе њиховом профиту. Истраживање како нови приступи могу додатно побољшати препоруке се стога може сматрати посебно значајним.

У циљу тога, имплементирана су два модела различите сложености, заснована на граф неуронским мрежама, као и модел заснован на факторизацији матрица. Након тога, они су истренирани и њихови резултати поређени на истом скупу података. Скуп података чини податке о обављеним трансакцијама на бразилској OList веб продавници, обogaћен додатним подацима на основу географских локација корисника. Над овим подацима GNN модели су показали значајно боље резултате, као и дугорочну употребљивост у односу на факторизацију матрица.

Предмет даљег истраживања могли би бити сложенији модели, засновани на неким другим, рачунарски сложенијим варијацијама GNN архитектура (GAT [13], NGCN [7]). Осим тога, корисно би било испитати како би се ови модели понашали на бољим скуповима података, који садрже више информација о корисницима и производима.

## 6. ЛИТЕРАТУРА

[1] Miquido, „We know what you like! Recommendation Systems in business,“ 24 June 2020. [На мрежи]. Available: <https://medium.com/swlh/we-know-what-you-like-perks-of-recommendation-systems-in-business-5f227bb6d09>. [Последњи приступ 27 September 2023].

- [2] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner и G. Monfardini, „The Graph Neural Network Model,“ *IEEE Transactions on Neural Networks*, т. I, бр. 20, pp. 61-80, 2008.
- [3] M. Balabanovic и Y. Shoham, „Fab: content-based, collaborative recommendation,“ *Communications of the ACM*, т. III, бр. 40, pp. 66-72, 1997.
- [4] Y. Koren, R. Bell и C. Voinsky, „Matrix factorization techniques for recommender systems,“ *Computer*, т. 42, бр. 8, pp. 30-37, 2009.
- [5] Y. Koren, „The Bellkor Solution to the Netflix Grand Prize,“ у *Netlix Prize Documentation*, т. 81, 2009, pp. 1-10.
- [6] W. Hamilton, Z. Ying и J. Leskovec, „Inductive Representation Learning on Large Graphs,“ у *Advances in Neural Information Processing Systems*, 2017.
- [7] X. Wang, X. He, M. Wang и F. Feng, „Neural Graph Collaborative Filtering,“ у *42nd international ACM SIGIR conference on Research and development in Information Retrieval*, 2019.
- [8] W. Xiao, H. Ji, C. Shi, B. Wang, Y. Ye и C. Peng, „Heterogeneous Graph Attention Network,“ у *The world wide web conference*, 2019.
- [9] OList и A. Sionek, „OList Brazillian Ecommerce Dataset [Kaggle],“ [На мрежи]. Available: <https://doi.org/10.34740/KAGGLE/DSV/195341>. [Последњи приступ 12 July 2023].
- [10] M. Davis и M. Palumbo, „A primer on the economics and time series econometrics of wealth effects,“ 2001.
- [11] J. Leskovec, A. Rajaraman и J. Ullman D, *Mining of Massive Datasets*, Cambridge University Press, 2010.
- [12] T. Kipf и M. Welling, „Semi-supervised Classification with Graph Convolutional Networks,“ arXiv, 2016.
- [13] P. Velickovic, C. Guillem, A. Casanova и A. Romero, „Graph Attention Networks,“ arXiv, 2017.

## Кратка биографија:



**Матија Матовић** рођен је 1. јула 1999. године у Ужицу. Основне академске студије завршио је 2022. године на Факултету техничких наука, на ком брани и мастер рад 2023. године из области Електротехнике и рачунарства.  
контакт: [matija.matovic@uns.ac.rs](mailto:matija.matovic@uns.ac.rs)