

(12) CERERE DE BREVET DE INVENȚIE

(21) Nr. cerere: a 2021 00563

(22) Data de depozit: 20/09/2021

(41) Data publicării cererii:
30/05/2022 BOPI nr. 5/2022

(71) Solicitant:
• UNIVERSITATEA POLITEHNICA DIN
BUCUREȘTI, SPLAIUL INDEPENDENȚEI
NR.313, SECTOR 6, BUCUREȘTI, B, RO

(72) Inventatori:
• RUȘEȚI ȘTEFAN, STR.DOAMNA GHICA,
NR.32B, BL.T3, AP.608, SECTOR 2,
BUCUREȘTI, RO;

• DĂSCALU MIHAI, STR.STOICA
LUDESCU, NR.61, ET.1, AP.7, SECTOR 1,
BUCUREȘTI, B, RO;
• GUȚU-ROBU MARIUS-GABRIEL,
STR.SG.MAJ.VASILE TOPLICEANU, NR.15,
BL.P42B, SC.3, ET.1, AP.68, SECTOR 5,
BUCUREȘTI, B, RO

(54) METODĂ DE RECOMANDARE DE CĂRȚI PE BAZA
RECENZIILOR FOLOSIND UN SPAȚIU DE REPREZENTARE
LATENT MULTI-ASPECT

(57) Rezumat:

Invenția se referă la o metodă de recomandare de cărți având la bază o codificare a recenziilor primite de la utilizatori, în funcție de mai multe aspecte detectate automat. La nivel de intrări, metoda folosește un graf de cunoștințe centrat pe cărți din care sunt derivate reprezentări vectoriale aferente fiecărei cărți, alături de descrieri textuale sau conținutul efectiv al cărților, precum și de contextul social derivat din recenziile utilizatorilor. Metoda conform invenției cuprinde două faze, o fază de antrenare și o fază de testare. În faza de antrenare, textul aferent cărții și recenziilor este transformat în reprezentări vectoriale folosind modele de limbă preantrenate, cum ar fi BERT, iar metoda propusă extrage dinamic aspecte latente la nivelul cărților care sunt aplicate automat pe fraze din recenziile utilizatorilor, în vederea identificării aspectelor dominante. Profilul utilizatorului este modelat sub forma unor măști individuale pentru fiecare aspect, care sunt utilizate în predicția notei unei recenzii. În faza de testare sunt furnizate recomandări din prisma aspectelor considerate relevante de utilizator în alte cărți evaluate.

Revendicări: 5
Figuri: 2

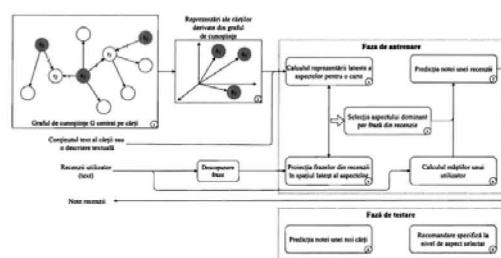


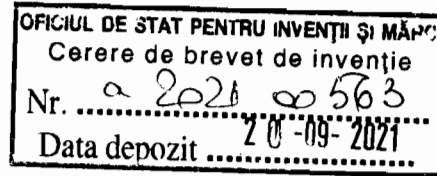
Fig. 1

Cu începere de la data publicării cererii de brevet, cererea asigură, în mod provizoriu, solicitantului, protecția conferită potrivit dispozițiilor art.32 din Legea nr.64/1991, cu excepția cazurilor în care cererea de brevet de invenție a fost respinsă, retrasă sau considerată ca fiind retrasă. Întinderea protecției conferite de cererea de brevet de invenție este determinată de revendicările conținute în cererea publicată în conformitate cu art.23 alin.(1) - (3).



Metodă de recomandare de cărți pe baza recenziilor folosind un spațiu de reprezentare latent multi-aspect

Descriere



Prezentarea domeniului de aplicare

Invenția se referă la o metodă de recomandare de cărți care îmbină reprezentări derivate dintr-un graf de cunoștințe cu descrieri textuale (conținutul efectiv al cărților), dar și cu contextul social derivat din recenziile utilizatorilor, prin utilizarea unui spațiu de reprezentare latent de aspecte de interes. Aspectele vizează, fără a se limita la, elemente despre tematică, stil de scriere, dificultate de înțelegere a textului, acțiune / personaje, și reprezintă elemente caracteristice ale textului, învățate automat de către metodă.

Domeniul de aplicare al metodei este vast, aceasta contribuind la nivelul tehnicilor de prelucrare a limbajului natural prin îmbunătățirea algoritmilor ce vizează furnizarea de recomandări semantice prin utilizarea unor meta-reprezentări într-un spațiu latent de aspecte. Metoda poate fi aplicată pentru a determina cărți de interes din prisma unui utilizator, atât la nivel general, cât și la nivel specific din prisma anumitor aspecte considerate ca fiind cele mai relevante. Totodată, metoda poate fi generalizată și aplicată în contexte similare de recomandări de diferite tipuri de resurse pentru care sunt furnizate ca intrare: 1) un graf de cunoștințe centrat pe un tip de entitate (spre exemplu, film, melodie), 2) descrieri textuale aferente fiecărei entități (spre exemplu, rezumat film, versuri melodie) și 3) recenzii furnizate de utilizatori, care includ atât un text scris, cât și o notă acordată de către utilizator entității vizate.

Stadiul tehnicii

Brevetul CN106202184B (<https://patents.google.com/patent/CN106202184B>) se axează pe recomandări colaborative de cărți și identifică problema necesarului de stocare a unui volum mare de date, ceea ce duce la probleme de scalabilitate atunci când se realizează interogări. În privința sistemelor de recomandare colaborativă, acestea de multe ori folosesc matrice rare (en., *sparse matrix*), iar utilizatorii noi se lovesc de problema „cold start” în care nu primesc recomandări adaptate din cauza absenței istoricului de interacțiune cu sistemul. Brevetul propune implementarea unei matrice de preferințe a utilizatorului, în care se consideră un scor de acord între utilizator și categoria din care o carte face parte. Se folosește, de asemenea, o matrice în care cititorii se regăsesc pe linii, iar atributele pe coloane; intersecția dintre o linie și o coloană reprezintă valoarea atributului pentru cititorul ales. Similaritatea între utilizatori este dată de graful ponderat între cititori și cărți (respectiv alt tip de produse utilizate). Ulterior se folosește un algoritm de cluster-izare pe mulțimea de cititori, iar la final cărțile recomandate sunt furnizate pe baza relației dintre cititor și clusterul de clasificare. Dezavantajul este că această metodă nu ia în considerare descrieri de tip text, care ar putea fi folosite pentru furnizarea de recomandări semantice.

Brevetul US20110282758A1 (<https://patents.google.com/patent/US20110282758A1>) este unul dintre brevetele deținute de Amazon pentru dezvoltarea unui serviciu de recomandări de produse. Acesta furnizează recomandări pe baza relațiilor dintre comportamentul utilizatorului (spre exemplu, acțiunea de a cumpăra) și produsele corespunzătoare. Se folosesc, de asemenea, relații între produse bazate pe conținutul acestora. Pornind de la o listă de produse recomandate pentru utilizator, se furnizează pentru fiecare dintre ele o listă de recomandări de alte produse similare cu cele inițial recomandate, acestea din urmă fiind ponderate cu preferința utilizatorului pentru produsul de interes inițial. În final, produsele rezultate sunt îmbinate într-o colecție de produse recomandate. Brevetul prezintă metoda de bază de recomandare, însă dezavantajul său este că nu sunt utilizate recenziile utilizatorilor care pot include informații esențiale cu privire la preferințele și opiniile acestora.

Brevetul US10120908B2 (<https://patents.google.com/patent/US10120908B2>) sugerează ideea utilizării primului scor oferit de un utilizator asupra unui document pentru a identifica alți posibili candidați pentru a fi recomandați. Se folosesc clustere de meta-date, în cadrul cărora fiecare lucrare are o pondere. Pentru fiecare dintre categoriile cluster-ului se caută clustere de concepte similare, precum și clustere de tendințe similare. Metoda consideră similarități semantice determinate folosind tehnici de procesare a limbajului natural, dar rezultatul final se limitează la gruparea de documente, fără personalizare raportat la preferințele utilizatorului.

În [Shu, K., Wang, S., & Liu, H. (2019, January). *Beyond news contents: The role of social context for fake news detection*. In *WSDM '19, Session 6: Networks and Social Behavior*. Melbourne, VIC, Australia:ACM <https://dl.acm.org/doi/10.1145/3289600.3290994>] este considerată problema știrilor apărute odată cu transformarea social media într-un mediu de informare folosit la scară largă și este accentuată importanța contextului social. Studiul analizează relațiile pe 3 dimensiuni între editor, știrile publicate de aceștia și cititori. Un caz extrem al acestor relații face referire la editori susținători de idei fixe, ce au tendința de a publica știri ce favorizează fațeta cu care aceștia se identifică. De cealaltă parte, un alt caz extrem arată că persoane puțin credibile au șanse mai mari să împărtășească cu ceilalți utilizatori știri false. Lucrarea propune explorarea contextului social pentru detectarea știrilor false raportat la triada editor-știre-utilizator folosind multiple reprezentări vectoriale (en., *embedding-uri*): 1) spațiul latent al conținutului știrilor; 2) caracteristicile utilizatorilor derivate din relațiile lor sociale; 3) reprezentarea relației dintre utilizatori și știri; 4) relații între editori și știri; 5) o componentă de clasificare semi-supervizată care învață o funcție de clasificare pentru a prezice categoria unei știri. Problema se reduce la una de optimizare a funcției obiectiv (en., *loss*) ce îmbină cele 5 componente. Sunt utilizate diverse metrice pentru analiza conținutului text, iar în final se demonstrează importanța relațiilor triadei editor-știre-utilizator. O abordare similară pentru reprezentarea unor triade *autor_recenzie-recenzie-cititor_recenzie* ar putea fi derivată din studiul inițial și aplicată într-un sistem de recomandare pe baza recenziilor. Totuși, articolul are limitări din prisma reprezentărilor semantice utilizate care nu pornesc de la modele de limbă pre-antrenate.

Node2vec [Palumbo, E., Rizzo, G., Troncy, R., Baralis, E., Osella, M., & Ferro, E. (2018, June). *Knowledge graph embeddings with node2vec for item recommendation*. In *The Semantic Web (ESWC 2018), Satellite Events*, pp. 117-120, Heraklion, Greece: Springer



https://link.springer.com/chapter/10.1007/978-3-319-98192-5_22] folosește grafuri de cunoștințe, care s-au dovedit a fi eficiente pentru integrarea de noi produse, respectiv considerarea problemei de raritate a datelor. Strategia se bazează pe reprezentări vectoriale ale caracteristicilor produselor, acestea putând fi comparate pentru recomandarea de produse similare. În plus, așa cum se arată în lucrare, modelul semantic node2vec surclasează alți algoritmi de filtrare colaborativă. Caracteristicile produselor sunt concrete, știute de la început, iar această limitare reprezintă principalul dezavantaj al metodei pentru sistemele în care aceste caracteristici nu sunt cunoscute inițial.

O abordare similară este cea bazată pe entity2rec [Palumbo, E., Monti, D., Rizzo, G., Troncy, R., & Baralis, E. (2020). *entity2rec: Property-specific knowledge graph embeddings for item recommendation*. *Expert Systems with Applications*, 151, 113235. <https://www.sciencedirect.com/science/article/abs/pii/S0957417420300610>], care folosește tot o reprezentare de tip graf de cunoștințe, dar care însă poate genera relații nu atât de evidente între produse. Algoritmul s-a dovedit eficient atunci când setul de date cuprinde puține caracteristici ale produselor, rezolvând parțial problema node2vec descrisă anterior. Cu toate acestea, metoda înclină spre produse cu popularitate mare, celelalte fiind rar recomandate. Entity2rec învață relații între produse pe baza caracteristicilor extrase din subgrafuri bazate pe proprietăți specifice. Strategia duce la o acuratețe ridicată, respectiv are abilitatea de a integra și un caracter de noutate, dar și un grad de recomandări cu caracter aleatoriu. Dezavantajul metodei este dat de faptul că aceasta nu consideră descrieri textuale sau recenzii.

În [Liu, D., Bai, T., Lian, J., Zhao, X., Sun, G., Wen, J. R., & Xie, X. (2019). *News Graph: An Enhanced Knowledge Graph for News Recommendation*. In *KaRS 2019, Beijing, China: ACM*. <https://www.microsoft.com/en-us/research/publication/news-graph-an-enhanced-knowledge-graph-for-news-recommendation/>] este introdusă o strategie nouă de construire a unui graf de cunoștințe pentru un domeniu. Autorii arată faptul că majoritatea algoritmilor care folosesc grafuri de cunoștințe se bazează pe date așa cum provin ele din sistemul de recomandare. Strategia din lucrare constă în trei etape: 1) introducerea unui nou grup de entități pentru considerarea informației de context / topice; 2) adăugarea legăturilor colaborative bazate pe interacțiunea utilizatorilor cu platforma; 3) eliminarea muchiilor irelevante din graful de cunoștințe. Autorii sugerează folosirea unei strategii similare și în cadrul altor sisteme de recomandare bazate pe știri sau articole. Metoda consideră reprezentări vectoriale pentru utilizatori și articole, dar, similar cu metodele anterioare, are dezavantajul că nu include recenzii.

În [Rana, C., & Jain, S. K. (2012). *Building a Book Recommender system using time based content filtering*. *WSEAS Transactions on Computers*, 11(2), 2224-2872 <http://www.wseas.us/journal/pdf/computers/2012/54-571.pdf>] este prezentată o metodă pentru un magazin online de cărți care ia în considerare posibilitatea de schimb a preferințelor utilizatorilor în timp. Se conturează astfel dimensiunea temporală, care cuantifică de câte ori un utilizator apreciază un produs de-a lungul timpului. Un contor intern este folosit pentru a fi incrementat de fiecare dată când un utilizator marchează un produs ca fiind favorit. Metoda folosește și preferințele utilizatorului pentru filtrarea bazată pe conținut în situațiile în care doar o anumită categorie de cărți ar fi pretabilă pentru recomandări. Articolul argumentează că



abordarea bazată pe schimbări în preferințe de-a lungul timpului ajută prin faptul că atât preferințele utilizatorului, cât și reprezentările interne ale sistemului se schimbă. Dezavantajul metodei rezidă în reprezentările interne ale cărților care nu consideră *embedding*-uri derivate din modele de limbă pentru o contextualizare adecvată.

Prezentarea sintetică a invenției

Problema tehnică pe care invenția își propune să o rezolve vizează limitările metodelor actuale de recomandare, acestea fiind centrate pe fațete individuale și nu pe furnizarea de recomandări sinergic. Metoda propusă folosește atât grafuri de cunoștințe ce modelează relațiile dintre cărți, cât și conținut textual sau descrierile cărților, alături de dimensiunea socială care consideră recenziile propriu-zise ale utilizatorilor. Astfel, prin transpunerea într-un spațiu latent multi-aspect, metoda vizează realizarea unei potriviri optime între interesele utilizatorilor și cărți prin utilizarea unor așa numite aspecte. Aspectele vizează, fără a se limita la, elemente despre tematică, stil de scriere, dificultate a textului, acțiune / personaje, și sunt extrase în mod automat de către rețeaua neurală propusă. Un beneficiu major al acestei invenții îl reprezintă posibilitatea de fuziona cunoștințe reflectate în meta-informații despre cărți cu descrieri textuale, dar și cu recenzii text ale utilizatorilor.

Avantaje

Conform invenției, metoda prezintă următoarele avantaje:

- Extragerea dintr-o recenzie a mai multor aspecte latente ale unei cărți care evidențiază diferite fațete specifice identificate automat;
- Unificarea modului de calcul al aspectelor pentru întreaga colecție de cărți;
- Reprezentarea intereselor unui utilizator prin aplicarea unor măști vectoriale peste codificarea aspectelor;
- Posibilitatea de furnizare de recomandări prin reprezentările latente ale aspectelor, invariant de existența recenziilor pentru o anumită carte;
- Introducerea unei reprezentări generice pentru utilizatorii noi, obținută prin prezicerea scorului mediu al unei cărți, eficientă în adresarea problemei „cold start”.

Prezentarea figurilor

În continuare, invenția va fi descrisă în detaliu, cu referire la Figura 1, care prezintă schema bloc a metodei propuse care utilizează graful de cunoștințe (1) din care sunt derivate reprezentări vectoriale ale cărților (2), conținutul textual al cărții sau o descriere textuală, precum și liste de recenzii ale utilizatorilor cu note aferente. Pornind de la intrările anterioare, metoda calculează în faza de antrenare reprezentări latente ale cărților în spațiul de aspecte (3), proiectează frazele în spațiul latent (4) și determină aspectele dominante (5) care, înmulțite cu măștile vectoriale ale utilizatorului (6), sunt folosite în predicția notelor recenziilor (7). Ulterior, în faza de testare sunt prezise note pentru noi cărți (8), respectiv sunt furnizate recomandări specifice folosind distanțe pe aspecte de interes (9).



Figura 2 introduce o modelare formală a transformărilor care au loc în rețeaua neurală antrenată end-to-end în cadrul căreia ponderile sunt actualizate prin backpropagation. Etapele detaliate din Figura 2 sunt aliniate la nivel de indecși primari cu etapele din schema bloc de ansamblu (Figura 1).

Descrierea detaliată a invenției

Metoda propusă presupune codificarea cărților pe baza recenziilor primite de la utilizatori în funcție de mai multe aspecte detectate automat. Această codificare poate fi folosită pentru recomandarea de cărți care să țină cont de aspectele menționate de utilizator în recenziile sale, dar și de părerea sa privitoare la acele aspecte.

În *faza de antrenare* presupunem că metoda de codificare a cărților are la dispoziție următoarele intrări:

- Un graf de cunoștințe G centrat pe cărți care modelează meta-informațiile acestora, inclusiv: autor, an publicare, domeniu, gen, editură, etc;
- Conținutul text al cărților sau o descriere textuală a acestora;
- O listă de recenzii ale utilizatorilor, care conțin un text și o notă referitoare la o carte din colecție;
- Un model de limbă pre-antrenat de tip *Transformer* [Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. In *31st Conf. on Neural Information Processing Systems (NIPS 2017)* (pp. 5998–6008). Long Beach, CA, USA].

Graful de cunoștințe G (1) este centrat pe cărți la nivel de clasă principală și sunt modelate inter-dependențe între indivizi din prisma multiplelor relații posibile (carte – autor, carte – editură, carte – gen) sub formă de triplete $\langle h, r, t \rangle$ în RDF [Beckett, D., & McBride, B. (2004). *RDF/XML syntax specification (revised)*. *W3C recommendation*, 10(2.3)]. Spre exemplu, în *Figura 1*, nodurile cu gri închis marchează cărți (h_1, h_2, h_3), nodurile cu gri deschis sunt autori (t_1, t_2), iar relația r modelează legături de tipul „are autor” între cărți și diverși autori.

Conceptualizarea este la liberă alegere conform specificului domeniului. Suplimentar, modelarea poate prelua clase și relații cheie din Dublin Core [DC; Weibel, S., Kunze, J., Lagoze, C., & Wolf, M. (1998). *Dublin core metadata for resource discovery*. *Internet Engineering Task Force RFC*, 2413(222), 132], spre exemplu DC:Title, DC:Publisher sau DC:Creator.

Pornind de la graful G (1) se va genera o reprezentare fixă a unei cărți de dimensiune c (2) care reprezintă un *embedding* al nodului de tip carte din graful de cunoștințe și care poate fi obținută prin diverse metode — spre exemplu, dar fără a se limita la TransE [Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). *Translating embeddings for modeling multi-relational data*. *Advances in neural information processing systems*, 26], node2vec

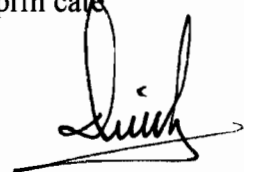


[Palumbo, E., Rizzo, G., Troncy, R., Baralis, E., Osella, M., & Ferro, E. (2018, June). Knowledge graph embeddings with node2vec for item recommendation. In *The Semantic Web (ESWC 2018), Satellite Events*, pp. 117-120, Heraklion, Greece: Springer https://link.springer.com/chapter/10.1007/978-3-319-98192-5_22], respectiv orice altă metodă prezentată în studiul [Wang, Q., Mao, Z., Wang, B., & Guo, L. (2017). Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering*, 29(12), 2724-2743]. În cazul în care graful de cunoștințe nu este disponibil, se pot utiliza *embedding*-uri de dimensiune c ale cărților care sunt învățate la antrenare.

Codificarea aspectelor se va face independent de părerea utilizatorilor, însă aceasta va fi modificată de recenzii prin *backpropagation*, împreună cu metoda de calcul a acesteia. Astfel, metoda de calcul a codificării aspectelor unei cărți va folosi codificarea cărții descrisă anterior sub forma unui vector de dimensiune c și reprezentarea textului cărții sau a unei descrieri aferentă sub forma unui vector de dimensiune e , vector construit printr-un model pre-antrenat care folosește arhitecturi de tip *Transformer*, precum BERT [Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2018)*. Minneapolis, Minnesota, USA: ACL]. Conform Figurii 2, aceste două reprezentări vectoriale sunt concatenate și trecute printr-un strat dens pentru a obține reprezentarea cărților (3.1) prin matricea de transformare T de dimensiune $[(e + c) \times d]$, urmat de câte încă un strat dens pentru fiecare aspect (A_1, A_2, \dots, A_k) care va fi reprezentat sub forma unui vector de dimensiune d – reprezentarea aspectelor unei cărți (3.2). Numărul de aspecte k pentru fiecare carte trebuie stabilit apriori, iar aceleași ponderi pentru straturile dense vor fi comune tuturor cărților.

O recenzie scrisă de un utilizator poate face referire la mai multe aspecte legate de o carte. Pentru identificarea acestor aspecte, este utilizată o metodă similară cu ProtoryNet [Hong, D., Baek, S. S., & Wang, T. (2020). Interpretable sequence classification via prototype trajectory. *arXiv preprint arXiv:2007.01777*], în care fiecare propoziție din recenzie este descompusă în fraze care sunt codificate cu un model pre-antrenat de tip *Transformer* (4.1). Reprezentările frazelor unei recenzii (4.1) de dimensiune e sunt transpuse în reprezentări ale frazelor în spațiul latent al aspectelor (4.2) prin matricea R de dimensiune $[e \times d]$ și comparate cu reprezentările unitare ale aspectelor. Modelul este forțat să aleagă un singur aspect per frază prin aplicarea unui strat de tip softmax peste similaritățile dintre fiecare aspect și fraza curentă (5.1), înmulțită cu o constantă mare γ . Astfel, un singur aspect va avea o pondere mare pentru o anumită frază, iar funcția rămâne diferențiabilă și permite aplicarea *backpropagation* prin toate straturile rețelei neurale.

Pornind de la un utilizator codificat inițial sub forma unui vector de dimensiune u , propria percepție a acestuia asupra unui aspect va fi reprezentată printr-un vector de aceeași dimensiune cu codificarea aspectelor d , care va fi aplicat sub forma unei măști de dimensiune d (6.2) prin înmulțire element cu element cu reprezentarea unitară a aspectului (7.1). Astfel, reprezentarea unui anumit aspect al unui cărți va fi independentă de utilizatori, dar percepția utilizatorului este modelată latent prin aplicarea măștii corespunzătoare acestuia. Fiecare utilizator va avea un număr de măști egal cu numărul de aspecte al unui cărți k ; măștile vor fi calculate prin câte



un strat dens, pornind de la codificarea inițială a utilizatorului de dimensiune u . Această codificare va permite gruparea utilizatorilor similari într-un spațiu vectorial, astfel încât măștile utilizatorilor similari să fie la rândul lor similare.

Astfel de măști au mai fost folosite pentru modificarea unor *embedding*-uri dintr-o perspectivă dată [Ruseti, S., Rebedea, T., & Trăușan-Matu, Ș. (2016, August). *Using embedding masks for word categorization. In Proceedings of the 1st Workshop on Representation Learning for NLP (pp. 201-205). Berlin, Germany: ACL*]. Acestea au avantajul unui număr mult mai mic de ponderi comparativ cu folosirea a k straturi dense diferite pentru fiecare utilizator, care ar fi dificil de învățat având la dispoziție un număr limitat de recenzii ale fiecărui utilizator.

Reprezentările unitare ale tuturor aspectelor (7.1) vor fi modificate prin aplicarea măștii proprii a utilizatorului corespunzătoare fiecărui aspect (6.2) prin înmulțire element cu element, iar întregul text este rezumat printr-o medie ponderată (7.2) a reprezentării aspectelor modificate, folosind scorurile calculate prin softmax. Astfel, metoda propusă asigură interpretabilitate prin proiecția fiecărei propoziții în spațiul latent al aspectului dominant.

Reprezentarea finală a recenziei (7.3) va fi folosită pentru predicția notei corespunzătoare recenziei prin intermediul unui strat dens (7.4). Toate componentele descrise vor fi conectate în aceeași rețea neurală reprezentată în Figura 2, iar ponderile vor fi actualizate prin *backpropagation*. Textul propriu-zis al unei recenzii este folosit doar pentru detecția aspectelor importante, dar nu este inclus în mod direct în predicție, deoarece el nu este prezent în etapa de recomandare.

Obiectivul de antrenare este bazat pe eroarea de predicție a notei fiecărei recenzii, dar include și o componentă care măsoară diversitatea aspectelor unei cărți, forțând modelul să codifice k aspecte diferite pentru fiecare carte. Diversitatea va fi măsurată prin distanța minimă dintre 2 aspecte diferite ale aceleiași cărți d_{min} , care va fi comparată cu o distanță minimă dorită δ .

$$loss = \frac{1}{N} \sum_{i=1}^N (predicted - target)^2 + \sigma(\delta - d_{min})$$

Pentru a putea codifica și utilizatori care nu au scris nicio recenzie, o reprezentare universală este învățată prin care se calculează măști ale aspectelor prin aceeași metodă. Această reprezentare este actualizată tot în faza de antrenare, prin prezicerea mediei recenziilor fiecărei cărți. Acesta reprezintă un mecanism eficient pentru adresarea problemei „cold start”.

Faza de testare este centrată pe recomandarea de cărți pentru un utilizator având la bază două strategii, ambele folosind o rețea neurală antrenată pentru predicția notei. În primul rând, modelul prezice nota aferentă unei cărți noi pentru un anumit utilizator (8), iar recomandările sunt sortate descrescător raportat la nota prezisă pentru un utilizator dat. În al doilea rând, metoda poate oferi recomandări specifice pentru un anumit aspect al unei cărți (9), comparând codificările aceluși aspect în spațiul latent pentru toate cărțile din colecție. Astfel, sunt furnizate recomandări pornind de la un aspect al unei cărți pe baza distanței minime față de același aspect determinat la nivelul celorlalte cărți disponibile în colecție.



Raportat la dimensiunea temporară, un utilizator nou pornește cu o codificare inițializată cu o reprezentare universală, iar pe măsură ce scrie recenzii întreaga rețea este antrenată în continuare pentru îmbunătățirea reprezentărilor aspectelor și a măștilor. Recenziile vechi sunt și ele folosite periodic la antrenare, astfel ca modelul să nu uite informațiile învățate cu mai mult timp în urmă.



Revendicări

1. Metodă de recomandare de cărți pe baza recenziilor folosind un spațiu de reprezentare latent multi-aspect, cuprinzând o fază de antrenare și o fază de testare, **caracterizată prin aceea că**, în *faza de antrenare*, se aplică un algoritm bazat pe rețele neurale pentru realizarea unei potriviri optime între interesele utilizatorilor și cărți prin utilizarea unei transpuneri într-un spațiu de reprezentare latent multi-aspect,

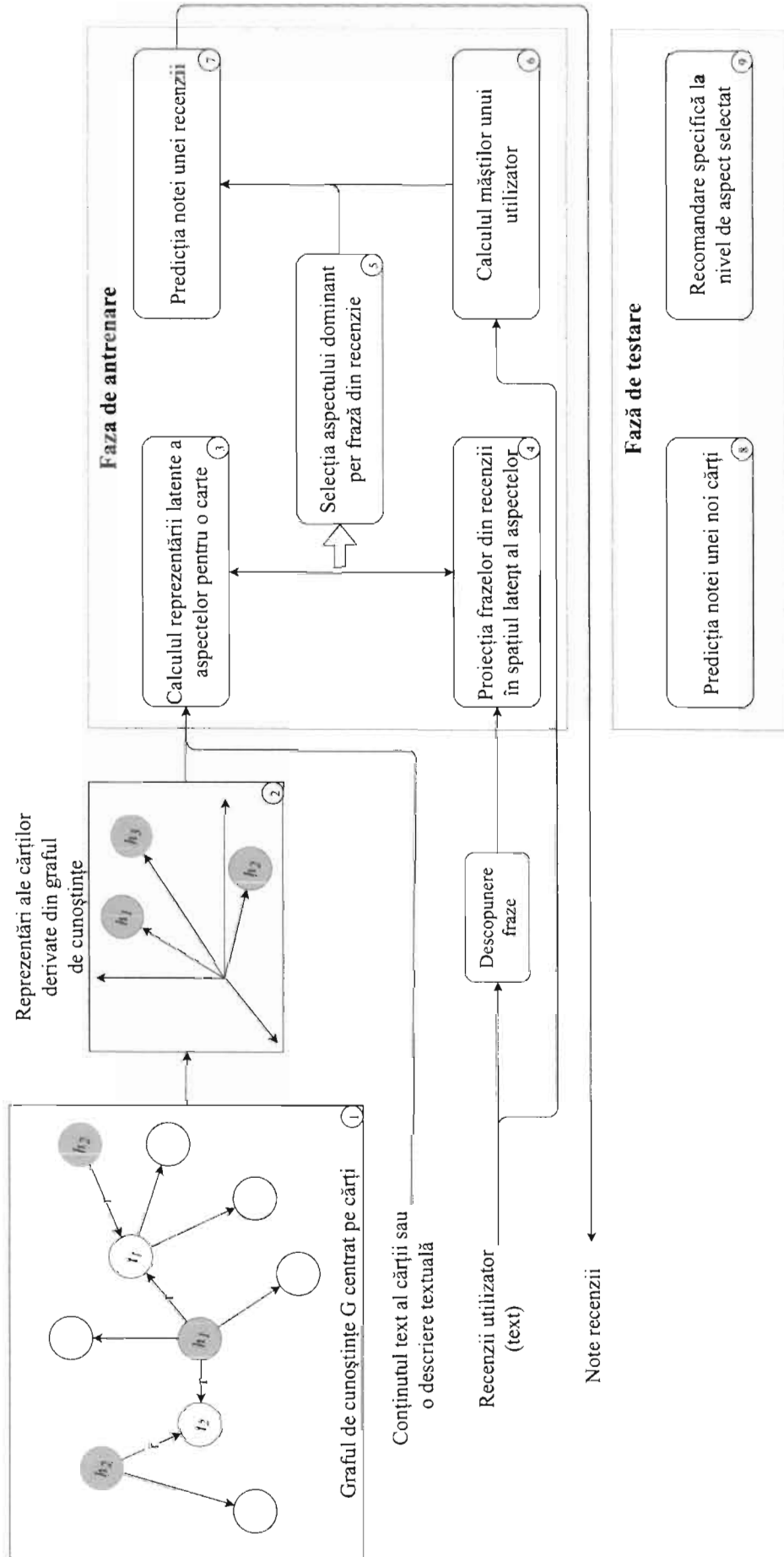
2. Metodă de calcul **conform revendicării 1, caracterizată prin aceea că** algoritmul aplicat cuprinde o primă etapă de calcul a reprezentărilor vectoriale pentru toate nodurile din graful de cunoștințe care modelează meta-informații din colecția de cărți; într-o a doua etapă se calculează reprezentările latente ale aspectelor fiecărei cărți folosind codificarea cărții din graful de cunoștințe și codificarea textului sau descrierii aferente acesteia; într-o a treia etapă se proiectează fiecare frază din recenziile utilizatorilor în spațiul latent al aspectelor cărții respective; într-o a patra etapă se selectează aspectul dominant per frază din recenzie; într-o a cincea etapă se calculează măști proprii utilizatorului pentru modificarea reprezentării fiecărui aspect; într-o a șasea etapă se prezice nota recenziei folosind spațiul latent multi-aspect mascat per utilizator și se actualizează ponderile prin *backpropagation* pentru întreaga rețea aferentă etapelor 2-6,

3. Metodă de calcul **conform revendicării 1, caracterizată prin aceea că** repetarea etapelor 2-6 ale algoritmului se realizează până la finalizarea antrenării și convergența rețelei, și se trece la o *fază de testare* în care sunt recomandate noi cărți în ordinea descrescătoare a notelor prezise pentru un utilizator, respectiv sunt oferite recomandări specifice pentru un anumit aspect al unei cărți comparând codificările acelu aspect în spațiul latent pentru toate cărțile din colecție.

4. Metodă **conform revendicării 1, caracterizată prin aceea că** reprezentări derivate din modele de limbă pre-antrenate bazate pe arhitectura de tip *Transformer* aplicate asupra descrierilor textuale ale cărților sau conținutul acestora, precum și asupra contextului social derivat din recenziile utilizatorilor, sunt combinate cu reprezentări derivate dintr-un graf de cunoștințe centrat pe cărți, printr-o proiecție într-un spațiu de reprezentare latent multi-aspect.

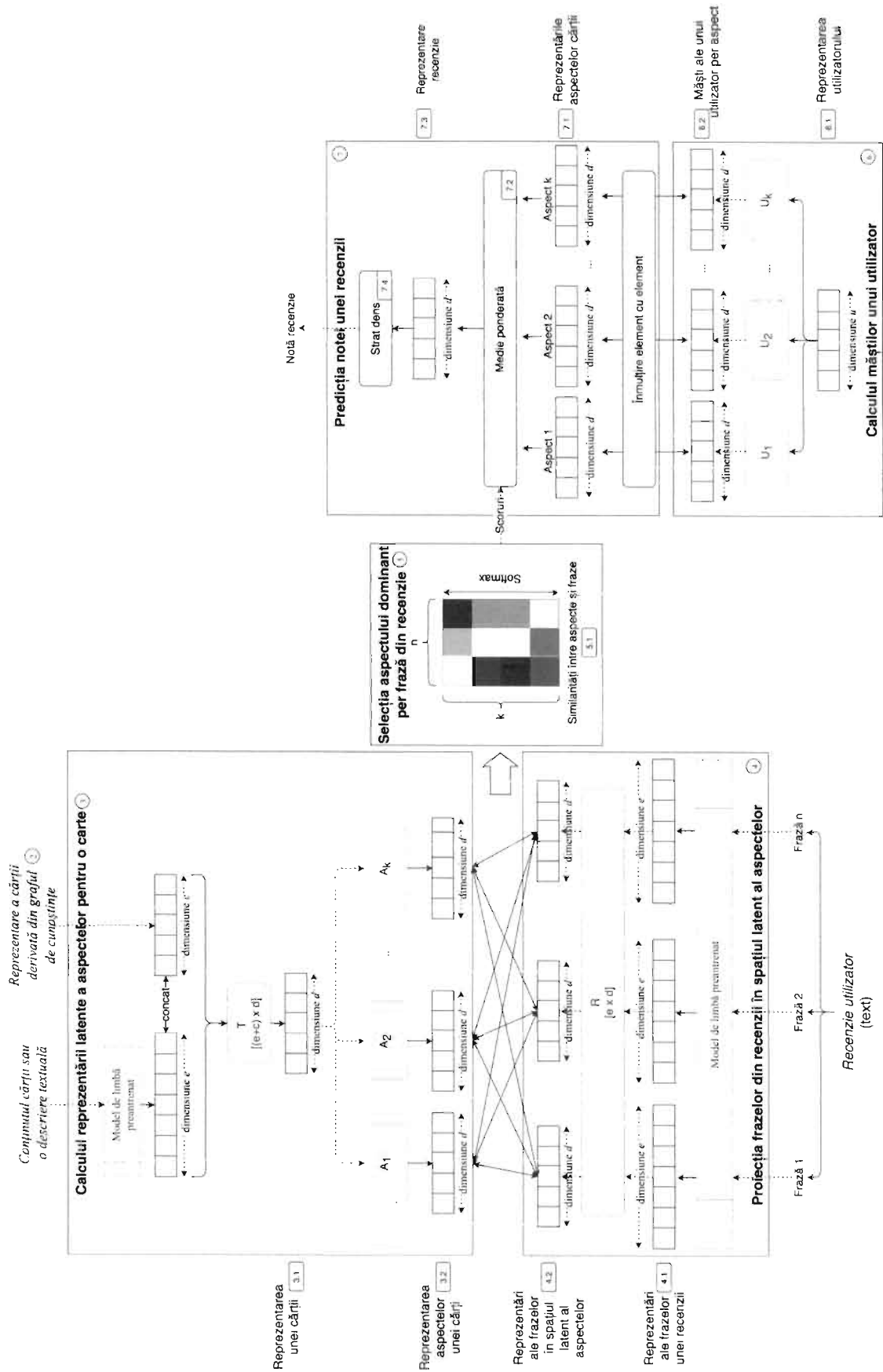
5. Metodă **conform revendicării 1, caracterizată prin aceea că** problema frecventă de „cold start” din sisteme de recomandare este adresată prin introducerea unei reprezentări generice pentru utilizatorii noi, obținută prin prezicerea scorului mediu al unei cărți.

Figura 1. Schema bloc a metodei de recomandare de cărți pe baza recenziilor folosind un spațiu de reprezentare latent multi-aspect



Liuh

Figura 2. Reprezentarea formală a transformărilor din etapa de antrenare



Signature