

RAPORT

ETAPA 1 - iulie 2023

Titlu Proiect: Soluții informatice pentru analiza impactului rețelelor de social media asupra instrumentelor investiționale cu grad ridicat de risc: cryptomonede și bursă

Domeniul științific: Științe economice

Membri:

prof.univ.dr. Adela Bâra,

drd. Costin Băroiu

În cadrul acestei etape, am diseminat rezultatele cercetării în următoarele lucrări:

- 1) ***Big data management and NoSQL databases***, Simona-Vasilica Oprea, Adela Bâra și Niculae Oprea, lucrare prezentată în cadrul celei de-a XX-a ediție a Conferinței Internaționale „Present Issues of Global Economys”, care a avut loc la Constanța în perioada 30 iunie – 1 iulie 2023. Articolul va fi publicat în volumul de lucrări al revistei „Ovidius University Annals. Economic Sciences Series”
- 2) ***A Descriptive – Predictive – Prescriptive Framework for the Social Media – Cryptocurrency Relationship***, Alexandru – Costin Băroiu și Adela Bâra. Articolul este în curs de pregătire și va fi transmis pentru publicare către jurnalul IEEE Access.

În lucrarea ***Big data management and NoSQL databases***, ne propunem să prezentăm contextul general, caracteristicile NoSQL, gestionarea datelor folosind baze de date (BD) NoSQL și un exemplu de stocare în MongoDB. Mai mult, organizarea și gestionarea unor volume mari de date (Big Data), tendința în domeniul sistemelor de baze de date avansate și al business intelligence vor fi descrise în această cercetare. Volumele mari de date din diverse domenii de activitate necesită un management adecvat pentru a obține un avantaj competitiv sau pentru a crește calitatea deciziilor. Diverse domenii generează volume mari de date, cum ar fi: rețelele sociale, criptomonede, bursa, consumerismul, dispozitivele electrice, senzorii, infrastructură inteligentă, cum ar fi contorizarea prin dispozitive smart, sistemele de securitate și supraveghere, sistemul de sănătate, transport, sectoarele de retail și cele industriale etc.

Evoluția managementului datelor a pornit de la fișiere (în anii 60) și s-a extins la baze de date. În anii ‘80 au apărut bazele de date relaționale, în anii ‘90 au apărut și bazele de date orientate obiect (Rashid & Sawyer, 2005). După anii 2010, a apărut o nouă categorie de baze de date, și anume baze de date non-relaționale sau baze de date NoSQL (Not only SQL) (Chen & Lee, 2019). Au apărut și bazele de date autonome (Curino et al., 2013). Toate coexistă și au particularități interesante care le fac potrivite pentru anumite aplicații (R. E. Schuler & Kessleman, 2019). Această evoluție a convins companiile să încerce una dintre cele mai recente baze de date (Bjeladinovic et al., 2020), (Bjeladinovic, 2018), dar au existat cazuri în care factorii de decizie au revenit la baze de date relaționale (<https://www.theguardian.com/info/2018/nov/30/postgo-helloby-post>). Mai multe lucrări de cercetare științifică au abordat evoluția în timp a bazelor de date (R. Schuler & Kesselman, 2021), (Domínguez et al., 2008), (Cleve et al., 2015), (Hillenbrand et al., 2022). Tendințele recente în tehnologia bazelor de date au fost investigate în (Lieponienė, 2021). Diverse rețele sociale și de cercetare, internetul dispozitivelor interconectate (Internet of Things sau IoT), comerțul electronic etc. necesită un management fiabil al datelor. Provocările acestor rețele sunt legate de volumul de date pe care le generează, viteza și varietatea datelor care trebuie stocate și procesate pentru a extrage informații utile și a obține avantaje competitive. (Lieponienė, 2021) a comparat BD relaționale și non-relaționale, luând în considerare alternative precum BD multi-model (Lu & Holubová, 2019) și BD cloud. Managementul volumelor mari de date în cloud a fost analizat în (Hameurlain & Morvan, 2016). Utilizarea Big Data Analytics în asistența medicală este investigată în (Batko & Ślęzak, 2022). Utilizarea sa s-a răspândit în numeroase domenii, cum ar fi: producție (Li et al., 2022), guvern (Long et al., 2021), energie (Zhou et al., 2016) etc. Aplicabilitatea bazelor de date multi-model pentru navigație a fost, de asemenea, subliniată în (Simon-Nagy et al., 2020).

Metodologia cercetării - Managementul datelor. O bază de date relațională (BD) reprezintă metoda tradițională de organizare a datelor în colecții omogene de date structurate. Ele sunt organizate după un model abstract care descrie modul în care datele sunt reprezentate și accesate, folosind formalizări matematice. Principalele elemente caracteristice ale modelului de date sunt următoarele: Structura datelor, Operatorii modelului, Constrângerile de integritate. Avantajele și dezavantajele BD relaționale sunt prezentate în Tabelul 1.

Tabel 1. Avantajele și dezavantajele BD relaționale

<p>Avantajele DB relațional sunt:</p> <ul style="list-style-type: none"> ✓ Organizarea unitară, coerentă și structurată a datelor după o schemă normalizată definită anterior (schemă on-write); ✓ Protecția datelor asigurată de Sistemul de Gestire al Bazei de Date (SGBD) vizează: securitatea și integritatea datelor; ✓ Partajarea datelor; ✓ Acces competitiv; ✓ Independența datelor față de aplicații ✓ Optimizarea accesului la date; ✓ Limbajele relaționale sunt standardizate (standard SQL). 	<p>Dezavantajele BD relaționale sunt:</p> <ul style="list-style-type: none"> ✓ Pierderea unor informații semantice utile prin normalizare; ✓ Operațiile relaționale, chiar și cele optimizate, sunt costisitoare din punct de vedere al resurselor de calcul; ✓ Limbile relaționale sunt limitate.
---	---

BD relaționale implementează proprietățile ACID: Atomicitate, Consistență, Izolație și Durabilitate.

Big Data are ca scop integrarea, organizarea, procesarea și stocarea unor volume mari de date într-un timp scurt. Reprezintă o suită de tehnologii utilizate pentru colectarea, procesarea, stocarea și analiza datelor caracterizate prin volum mare, varietate, viteză (viteză), veridicitate, așa cum este descris în Figura 1. Cele 10 Vs de Big Data (<https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx>) se referă la mai multe V-uri care includ vizualizare și valoare. Managementul unor volume mari de date include tehnologii orientate spre managementul datelor care permit integrarea, procesarea și analiza unor volume mari de date într-un timp scurt. Scopul este de a extrage informații utile din date și de a obține un avantaj competitiv. Sursele multiple de date includ aplicații și datele corespunzătoare din fișiere jurnal, senzori și date extrase în timp real, surse web și text sau date care reflectă activitatea pe rețelele sociale sau forumuri, articole etc.

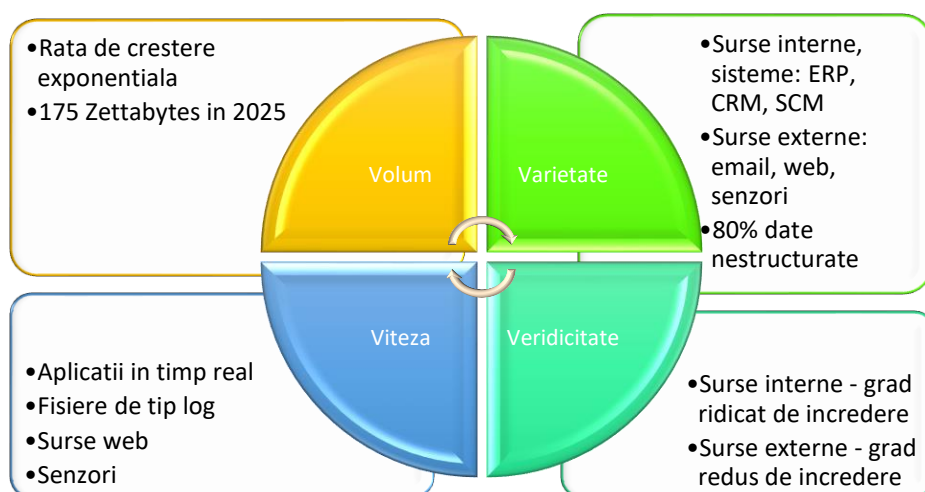


Figura 1. Caracteristicile big data

Începând cu anii 80, datele au fost stocate în mod tradițional în baze de date relaționale și depozitate de date. Cu toate acestea, BD relaționale nu au funcționalitatea de a organiza volume mari de date și cerințele de procesare legate de volum, varietate, veridicitate și viteza de colectare a datelor necesită BD non-relaționale sau NoSQL. Bazele de date NoSQL au următoarele avantaje

- ✓ Nu au o schemă (structură) fixă care să impună restricții privind stocarea datelor (schemă on-read);
- ✓ Permit stocarea și prelucrarea datelor semistructurate sau nestructurate;
- ✓ Lipsa necesității de normalizare a datelor și a relațiilor dintre tabele aduce performanțe sporite aplicațiilor;

- ✓ Oferă suport pentru replicare, distribuție, API (Application Programming Interface);
- ✓ Capacitatea de a scala orizontal sau de a replica și distribui date pe mai multe servere;
- ✓ Permite modificarea dinamică a structurii colecțiilor de date.

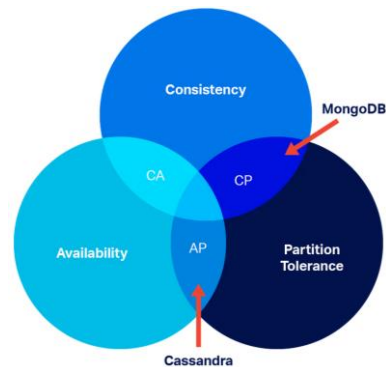
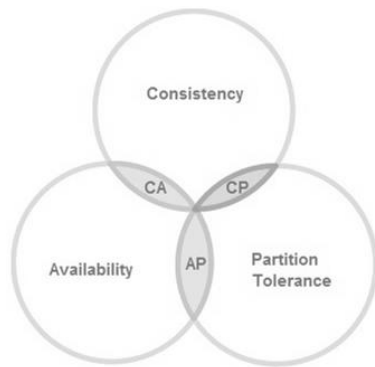
Baze de date NoSQL – dezavantaje:

- ✓ Nu implementează ACID;
- ✓ Lipsa standardizării limbii;
- ✓ Facilități de management reduse;
- ✓ Dinamica mai mare a versiunilor;
- ✓ Implementează parțial teorema CAP (Consistency, Availability, Partition tolerance).

Bazele de date NoSQL implementează teorema CAP care se referă la:

- ✓ Consecvența se referă la faptul că toți utilizatorii sistemului accesează aceleași date, indiferent de emoment;
- ✓ Disponibilitatea este o garanție că fiecare solicitare (query) va primi răspuns;
- ✓ Toleranța la partiționare este proprietatea sistemului de a funcționa în cazul în care anumite noduri din sistem nu mai sunt funcționale.

Conform teoremei CAP, o BD nu poate satisface simultan toate cele 3 constrângeri, dar poate excela la oricare dintre ele: CA, CP sau AP (ilustrat grafic în Figura 2). Prin urmare, există baze de date care implementează mai bine CP, în timp ce altele implementează mai bine AP. Două exemple de baze de date sunt MongoDB, care este mai bună la Partiționare și Cassandra, care este mai bună la Disponibilitate.

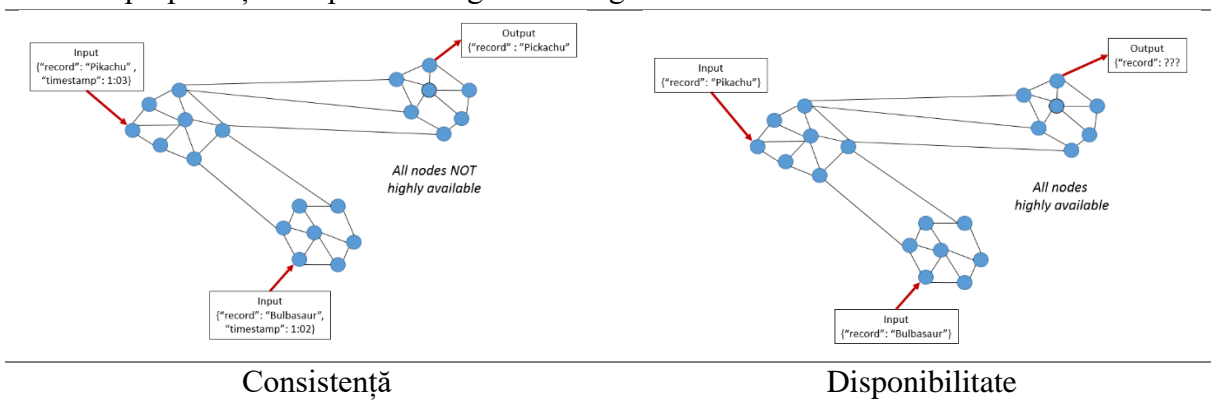


Teorema CAP – combinațiile posibile

BD NoSQL care implementează teorema CAP

Figura 2. Teorema CAP și combinația de 2 constrângeri: CA, CP sau AP

Cele trei proprietăți sunt prezentate grafic în Figura 3.



Consistență

Disponibilitate

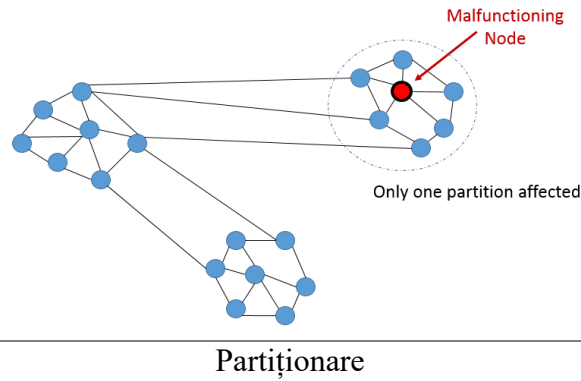


Figura 3. Teorema CAP: Consistență, Disponibilitate și Partiționare

Managementul datelor în BD NoSQL include următoarele metode de organizare a datelor, prin: Coloane; Valoare cheie; Document; Grafic. Pe coloane, datele sunt stocate în coloane (Suárez-Cabal et al., 2023), (Khourdifi et al., 2018). Următoarele baze de date sunt incluse în această categorie HBase, Cassandra, Hypertable, Accumulo, Amazon Simple DB, Cloudata, Cloudera. În ceea ce privește cheia-valoare, valorile sunt identificate printr-o cheie unică (Puangsaijai & Puntheeranurak, 2017). Modelul de date este simplu, scalabil, ușor de manipulat, fără o schemă predefinită (El Alami et al., 2018). Câteva exemple de astfel de BD includ: Redis, Riak, Oracle NoSQL Database, Tarantool, Tokyo Cabinet etc. Documentele sunt date semi-structurate și pot fi stocate sub diferite forme: JSON, BSON, XML etc. Documentele sunt stocate în BD printr-o cheie unică pentru a identifica acel document (Blanco et al., 2022). Couchbase și MongoDB sunt cele mai populare sisteme care implementează stocarea documentelor. Evaluarea performanței BD de documente NoSQL, cum ar fi Couchbase, CouchDB și MongoDB este analizată în (Carvalho et al., 2023). Graficele folosesc noduri, muchii/arce sau relații dintre ele (legături) și proprietățile nodurilor (Angles and Gutierrez, 2008). DB-urile bazate pe grafice nu folosesc indecși pentru a selecta o anumită valoare din BD, ci în schimb folosesc relațiile dintre noduri pentru a prelua elemente adiacente. Exemple de sisteme care implementează stocarea bazată pe grafice sunt: Allegro, Neo4J, OrientDB, Virtuoso. Un model eficient de baze de date grafice este furnizat în (Vyawahare et al., 2019). În plus, un exemplu de aplicație cu o bază de date care implementează grafuri BD este oferit de The Oracle of Bacon, care oferă link-uri interesante pentru filme (<http://oracleofbacon.org/movie/links.php>). Un rezumat al exemplelor este prezentat în Tabelul 2.

Tabel 2. Exemple de BD

Nr.	BD	Exemple
1	Pe coloane	HBase, Cassandra
2	Cheie-valoare	Oracle NoSQL Database, Riak, Redis
3	Document	CouchDB, Couchbase, and MongoDB
4	Grafuri	Allegro, Neo4J, OrientDB, Virtuoso
5	Multi-model	ArangoDB

BD sunt clasate în funcție de popularitatea lor. La scrierea acestei lucrări, Oracle și MongoDB sunt pe primul loc pentru BD relaționale și non-relaționale (<https://db-engines.com/en/ranking>), urmate de MySQL și Redis. 4.

Rezultate. MongoDB Atlas este un mediu care permite crearea unei baze de date non-relaționale (Figura 4 – Cluster0). Pe baza stringului de conexiune, un utilizator se poate conecta de la o aplicație (Figura 10).

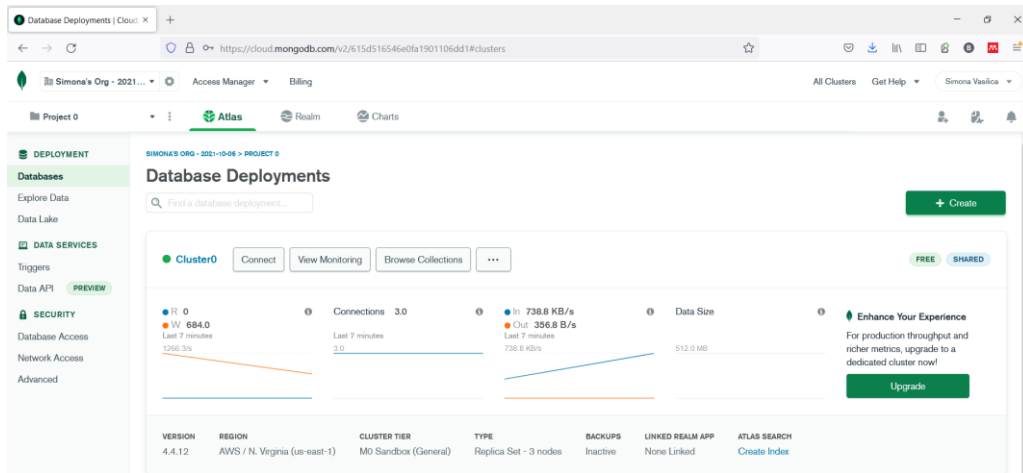


Figura 4. Atlas MongoDB

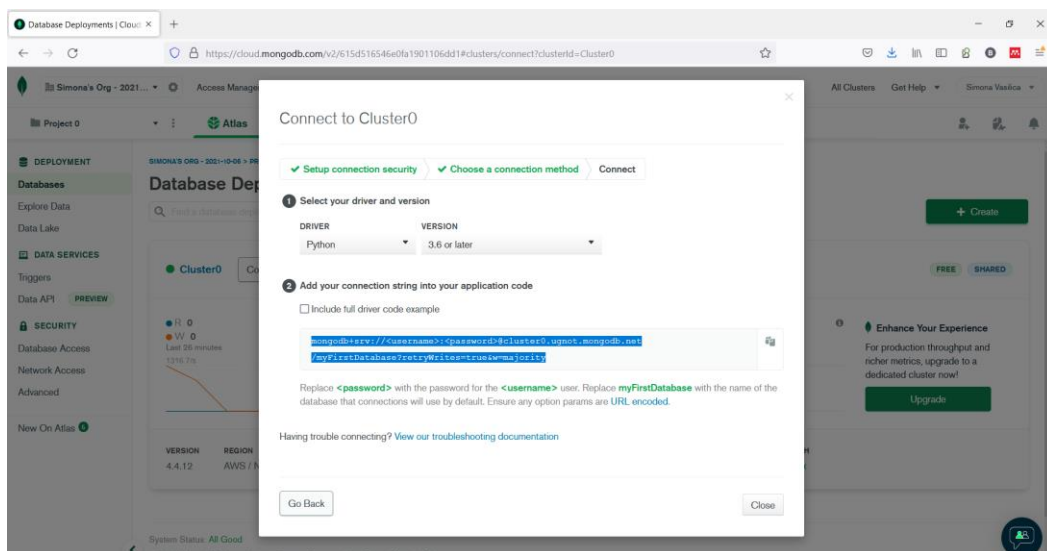


Figura 5. String pentru conectare la MongoDB

Folosind șirul de conexiune, un utilizator se conectează prin NoSQL Booster pentru Mongo la Cluster0 (Figura 6).

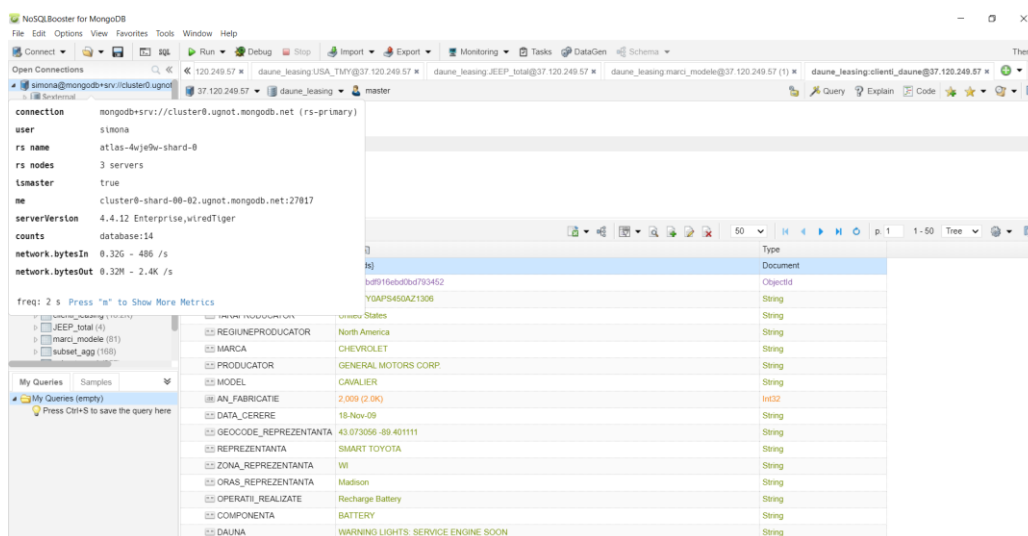


Figura 6. NoSQL Booster

În Figura 7, în NoSQL Booster pentru MongoDB, un nod este primar și două sunt secundare. Acestea asigură replicarea datelor în MongoDB.

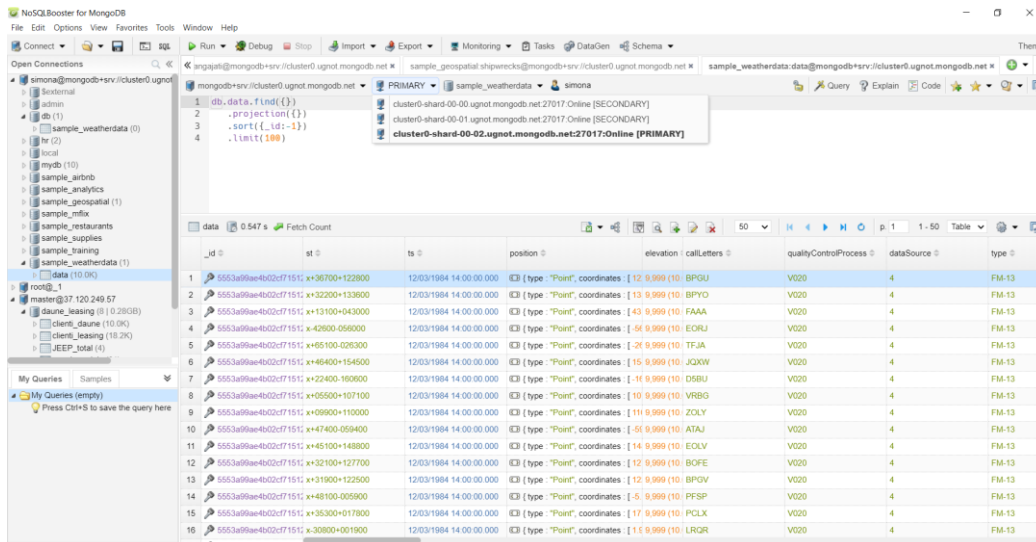


Figura 7. Nodurile Cluster0

MongoDB include metode care corespund instrucțiunilor SQL. De exemplu, instrucțiunea SQL SELECT poate fi comparată cu metoda MongoDB find(). Conducele (pipeline) ca în (Figura 8) pot fi setate să filtreze și să grupeze setul de date după diverse variabile.

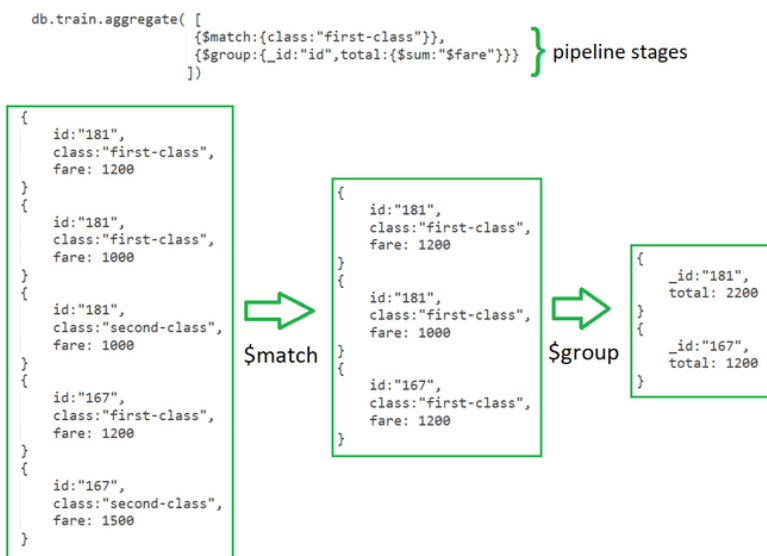


Figura 8. Pipeline de agregare în MongoDB

Concluzii. În această lucrare a fost evidențiată o scurtă evoluție a managementului datelor și a bazelor de date, împreună cu o comparație între BD relaționale și non-relaționale. Bazele de date NoSQL permit organizarea flexibilă a unor volume mari de date folosind scheme de date extensibile (grupuri de coloane, documente, grafice, cheie-valoare) în comparație cu modelarea datelor în scheme relaționale fixe. Sunt proiectate pentru scalare orizontală prin modele de distribuție a datelor cu mai multe noduri. Acceptă diferite interfețe pentru accesul la date și analiza avansată, așa cum este exemplificat în MongoDB.

În cadrul lucrării *A Descriptive – Predictive – Prescriptive Framework for the Social Media – Cryptocurrency Relationship* am subliniat faptul că Bitcoin și alte criptomonede au atras atenția multor oameni care căutau să investească și, mai important, să știe când să investească. Pentru a determina cel mai bun moment pentru a investi, este necesar să înțelegem cum evoluează prețul. Ca atare, predicția prețurilor criptomonedelor a devenit un domeniu foarte activ de cercetare în mediul academic și nu numai. Astfel, importanța unui cadru de analiză și de luare a deciziilor nu poate fi subestimată.

Un cadru descriptiv-predictiv-prescriptiv pentru relația social media-criptomonedă servește mai multor scopuri, cum ar fi îmbunătățirea nivelului de cunoaștere a influențelor rețelelor sociale pe piețele financiare, efectuarea unui efort de cercetare interdisciplinară și dezvoltarea modelelor avansate de predicție. Relația dintre rețelele sociale și bitcoin a fost cercetată anterior (Baroiu & Dobrita Ene, 2022), și s-au găsit dovezi ale unei conexiuni. Din punct de vedere academic, un astfel de cadru este de mare interes deoarece explorează domeniul relativ nou al influenței social media pe piețele financiare. Odată cu dezvoltarea tehnologiei digitale, metodele tradiționale de predicție și înțelegere a tendințelor financiare au fost modificate. Rețelele sociale joacă acum un rol semnificativ în stimularea opiniei publice și a sentimentului pieței. Înțelegând această relație, se poate obține o perspectivă asupra modului în care piețele financiare moderne sunt influențate și conduse de activitatea online și tendințele sociale.

Cadrul propus se află la intersecția diferitelor domenii, cum ar fi economie, știința datelor, IT, sociologie și comunicații. Un astfel de cadru ar putea stimula o mulțime de cercetări interdisciplinare, deoarece implică analiza și înțelegerea unor fenomene complexe, cum ar fi comportamentul rețelelor sociale, dinamica criptomonedelor și interacțiunea dintre ele. În cele din urmă, odată cu apariția Big Data și a tehnicilor avansate de învățare automată, există o căutare continuă pentru a construi modele predictive mai complexe. Acest cadru ar putea fi un plus valoros la acest domeniu de cunoaștere. Examinând modul în care tendințele rețelelor sociale pot prezice fluctuațiile prețurilor criptomonedelor, modelele predictive pot fi îmbunătățite și factorii din spatele performanței acestor modele pot fi mai bine înțeleși. O astfel de activitate are implicații potențiale pentru prognoza financiară, tranzacționarea algoritmică și managementul riscurilor.

STUDIUL LITERATURII

În ceea ce privește cadrele în contextul piețelor cripto, literatura este încă la început. Există studii care propun un cadru pentru identificarea principalelor driveri pentru investiția în criptomonede (Böyükaslan & Ecer, 2021). Modelul fuzzy Full Consistency Method-Bonferroni (FUCOM-F'B) este utilizat în această lucrare pentru a analiza motivele cheie ale investiției în criptomonede. Cei 23 de driveri au fost clasificați în cinci categorii: funcționalitate, finanțe, infrastructură juridică, tehnologie și securitate. Potrivit sondajului, cele mai importante elemente în alegerea unei criptomonede sunt criptarea electronică puternică și utilizarea semnăturilor digitale. Cercetarea validează fiabilitatea, aplicabilitatea și stabilitatea modelului sugerat, luând, de asemenea, în considerare ambiguitățile și subiectivitatea în procesul de luare a deciziilor. În cele din urmă, studiul servește ca instrument de luare a deciziilor pentru autorități de reglementare, politicieni, practicieni și investitori în criptomonede.

O altă lucrare din domeniu propune un cadru pentru partajarea sistemică a riscurilor criptomonedelor în criza COVID-19 (Akhtaruzzaman et al., 2022). Modelul Valoare la Risc Condițională (CoVaR) este utilizat în această lucrare pentru a crea un indice sistemic de contagiune (SCI) pentru criptomonede, cu scopul de a analiza efectele lor de propagare. În perioada de timp COVID-19, SCI a atins un vârf, indicând prezența căilor de contagiune induse de pandemie. În plus, creșterea interconexiunilor în rețelele sistemice de criptomonede în timpul epidemiei a evidențiat o creștere a numărului acestor căi de contagiune. Acest studiu oferă informații practice investitorilor, ajutându-i să înțeleagă vulnerabilitatea sistemică a anumitor criptomonede și să emită judecăți educate atât în situații de criză, cât și în afara crizei.

Cu toate acestea, majoritatea eforturilor de dezvoltare a cadrelor pentru criptomonede apar într-un context legislativ (European Commission, 2021), unde diverse instituții precum Comisia Europeană încearcă să încadreze aceste instrumente financiare într-un cadru legal. Astfel, domeniul cunoașterii este încă la început și se poate aduce o contribuție semnificativă. Cadrul propus în acest raport este, prin urmare, unul dintre primele cadre care se adresează domeniului criptomonedelor și este primul, după cunoștințele noastre, care abordează valorile în lanț, sentimentul rețelelor sociale și inteligența artificială în contextul criptomonedelor.

Această secțiune prezintă cadrul, oferind o descriere detaliată a fiecărei secțiuni și a fiecărui pas. Scopul principal al cadrului este de a prezenta un proces raționalizat de colectare a datelor, analiză a datelor, prognozarea a seriilor de timp și luare a deciziilor în contextul piețelor criptomonedelor. Pentru început, va fi prezentată secțiunea de analiză descriptivă.

În etapa inițială, datele sunt strânse din înregistrările istorice ale prețurilor, datele din lanț și conținutul rețelelor sociale. Aceste date brute sunt procesate pentru a fi pregătite pentru analiză. Tehnicile avansate de învățare automată și de învățare profundă sunt folosite pentru a transforma informațiile nestructurate în evaluări cantitative ale sentimentelor pentru analiza sentimentelor pe datele din rețelele sociale. Simultan, datele numerice, care includ prețurile istorice și datele din lanț, sunt normalizate pentru a asigura coerența. Componentele sezoniere din date sunt, de asemenea, recunoscute și eliminate pentru a evita distorsiunile în analiza ulterioară, asigurând o descriere precisă a tendințelor intrinseci.

În urma pregătirii inițiale a datelor, se efectuează o examinare aprofundată a interrelațiilor dintre variabile la diferite intervale de timp. Acest lucru se realizează prin utilizarea matricei de corelație, un instrument puternic care oferă înțelegere vizuală a potențialelor corelații dintre diferite variabile. Perspectivele obținute din această analiză servesc ca fundație pentru următorul pas în cadru, care este dezvoltarea modelului predictiv. Pe baza corelațiilor identificate, datele procesate anterior sunt utilizate pentru a construi un model de predicție care urmărește să prognozeze tendințele viitoare ale datelor, oferind o imagine mai cuprinzătoare a relației complexe Social Media - Criptomonedă.

Pe valoarea țintă, care în acest caz este prețul, sunt create mai multe praguri unice în timpul procesului de construire a modelului. Aceste praguri încorporează o varietate de concepte statistice și de învățare automată, inclusiv corelația Pearson, importanța caracteristicilor și pragurile de analiză a componentelor principale (PCA). Aceste valori ajută la identificarea celor mai importante valori care afectează variabila țintă. După finalizarea acestui proces de extracție bazat pe prag, datele sunt împărțite în subseturi de instruire, validare și testare. Setul de antrenament este folosit pentru a antrena modelul, setul de validare este folosit pentru a rafina modelul, iar setul de testare este utilizat pentru a evalua performanța modelului cu date necunoscute și pentru a asigura generalizarea modelului și capacitatea de a gestiona date viitoare, din lumea reală.

După separarea datelor, abordarea trece la pasul critic al optimizării hiperparametrilor. În această fază, un număr mare de modele este antrenat, fiecare cu un set unic de parametri. Aceste variabile includ dimensiunea stratului ascuns, rata dropout și rata de învățare. Performanța fiecărui model este evaluată folosind un set de metrici de eroare, oferind o evaluare amănunțită a puterii predictive și acurateței modelului. Eroarea pătratică medie (MSE), rădăcina din eroarea pătratică medie (RMSE), eroarea medie absolută (MAE) și eroarea procentuală medie absolută (MAPE) sunt valorile utilizate la acest pas. Aceste valori funcționează împreună pentru a selecta cel mai performant model, cel cu cele mai precise prognoze și cel mai mic nivel de eroare.

În urma etapei complete de evaluare, modelul cu cea mai bună performanță este ales pe baza pragului de date selectat și a combinației ideale de parametri. Acest model atinge echilibrul între precizie și generalizare. În viitor, metode mai avansate pentru reglarea hiperparametrului ar putea fi utilizate pentru a îmbunătăți cadrul. GridSearch, care examinează exhaustiv spațiul parametrilor, și RandomSearch, care prelevează aleatoriu spațiul parametrilor, ar putea oferi capacități de reglare mai mari și pot oferi modele și mai performante.

Pe lângă aceste tehnici, ar putea fi folosite și alte metode, cum ar fi algoritmi genetici. Selecția naturală a inspirat algoritmi genetici, care folosesc metode precum mutația, încrucișarea și selecția pentru a explora spațiul hiperparametric. Acest tehnici încep cu o populație de soluții potențiale și o rafinează iterativ pentru a identifica soluția cea mai potrivită. În acest exemplu, soluția cea mai potrivită ar fi modelul cel mai precis. Avantajul este capacitatea lor de a face față spațiilor mari și complexe de soluții, ceea ce le face utile pentru optimizarea hiperparametrilor în iterațiile viitoare ale cadrului. Aceste îmbunătățiri propuse ar putea crește puterea generală de predicție și acuratețe a analizei.

Procesul se încheie cu o etapă de analiză prescriptivă. Perspectivele culese din analiza descriptivă și analiza predictivă efectuată în etapele precedente sunt utilizate pentru a consilia și a conduce acțiunile viitoare și

procesele de luare a deciziilor în această etapă. Aceasta ar putea include sugestii pentru idei de investiții în criptomonede bazate pe tendințele observate în rețelele sociale sau pe indicatorii modului în care modificările datelor din lanț pot influența viitoarele fluctuații de preț. Cadrul complet cuprinde întregul proces, de la colectarea și preprocesarea datelor până la analiza descriptivă, predictivă și prescriptivă.

Dezvoltarea unui cadru descriptiv-predictiv-prescriptiv (DPP) pentru relația social media-criptomonede are o importanță academică și industrială substanțială. Din punct de vedere academic, acest cadru oferă o modalitate structurată de a înțelege și modela un domeniu complicat și puțin cercetat, cu perspective de cercetare și dezvoltare suplimentare. Acest aspect se adaugă la corpul de cunoștințe și stabilește bazele cercetărilor viitoare. Cadrul DPP poate fi un instrument puternic de luare a deciziilor pentru industrie, în special pentru cei interesați de investiții în bitcoin sau fintech. Părțile interesate pot obține informații de neprețuit asupra modelelor pieței, pot prezice cu exactitate mișcările viitoare ale pieței și pot concepe planuri solide prin valorificarea sentimentelor rețelelor sociale, a datelor istorice ale prețurilor și a datelor din lanț. Capacitatea cadrului de a conduce decizii strategice subliniază utilitatea sa practică într-o economie digitală în continuă schimbare.

CONCLUZII

Pentru a rezuma, cadrul propus în acest studiu oferă o abordare sistematică și ordonată pentru a explora relația dintre rețelele sociale și dinamica prețurilor bitcoin. Cadrul reprezintă un proces end-to-end, începând cu colectarea și preprocesarea datelor și progresând prin analiza exploratorie a datelor, dezvoltarea modelelor de predicție, până la etapa finală a luării deciziilor prescriptive. Acest cadru end-to-end asigură că fiecare pas al procesului este acoperit, culminând cu o examinare exhaustivă.

Această metodologie este concepută pentru a utiliza capacitățile învățării automate moderne și tehnicilor de învățare profundă. Aceste instrumente sunt folosite pentru a aduce lumină asupra interacțiunii multidimensionale și nuanțate dintre tendințele sociale media și fluctuațiile pieței bitcoin. Prin utilizarea cadrului propus, pot fi descoperite modele și tendințe ascunse care altfel ar putea trece neobservate.

În continuare, există oportunități promițătoare pentru cadrul DPP să fie îmbunătățit și perfecționat. Cercetările viitoare ar putea include adăugarea unor tehnici avansate de reglare a hiperparametrilor, cum ar fi GridSearch, RandomSearch sau algoritmi genetici, care ar putea duce la dezvoltarea unor modele predictive mai precise și mai rezistente. De asemenea, în cercetări viitoare pot fi introduse modelele mari de limbă, care au arătat rezultate foarte bune într-o varietate de domenii.

Scopul dezvoltării acestui cadru robust și cuprinzător este de a avea ramificații practice tangibile, precum și de a contribui la discursul savant și la înțelegerea acestor fenomene. Constatările din examinarea relației social media-criptomonede pot duce la metode de investiții mai complexe și mai performante. În plus, constatările pot ajuta factorii de decizie să negocieze topografia în schimbare rapidă a tărâmului digital, ghidând dezvoltarea cadrelor legislative și de reglementare adecvate.

REFERINȚE

- Akhtaruzzaman, M., Boubaker, S., Nguyen, D. K., & Rahman, M. R. (2022). Systemic risk-sharing framework of cryptocurrencies in the COVID–19 crisis. *Finance Research Letters*, 47, 102787. <https://doi.org/10.1016/j.frl.2022.102787>
- Baroiu, A.-C., & Dobrita Ene, G. (2022). Twitter Sentiment and Bitcoin Price – Is there a connection? 2022 *26th International Conference on System Theory, Control and Computing (ICSTCC)*, 258–262. <https://doi.org/10.1109/ICSTCC55426.2022.9931814>
- Batko, K., & Ślęzak, A. (2022). The use of Big Data Analytics in healthcare. *Journal of Big Data*. <https://doi.org/10.1186/s40537-021-00553-4>
- Bjeladinovic, S. (2018). A fresh approach for hybrid SQL/NoSQL database design based on data structuredness. *Enterprise Information Systems*. <https://doi.org/10.1080/17517575.2018.1446102>

- Bjeladinovic, S., Marjanovic, Z., & Babarogic, S. (2020). A proposal of architecture for integration and uniform use of hybrid SQL/NoSQL database components. *Journal of Systems and Software*. <https://doi.org/10.1016/j.jss.2020.110633>
- Blanco, C., García-Saiz, D., Rosado, D. G., Santos-Olmo, A., Peral, J., Maté, A., Trujillo, J., & Fernández-Medina, E. (2022). Security policies by design in NoSQL document databases. *Journal of Information Security and Applications*. <https://doi.org/10.1016/j.jisa.2022.103120>
- Böyükaslan, A., & Ecer, F. (2021). Determination of drivers for investing in cryptocurrencies through a fuzzy full consistency method-Bonferroni (FUCOM-F'B) framework. *Technology in Society*, 67, 101745. <https://doi.org/10.1016/j.techsoc.2021.101745>
- Carvalho, I., Sá, F., & Bernardino, J. (2023). Performance Evaluation of NoSQL Document Databases: Couchbase, CouchDB, and MongoDB. *Algorithms*. <https://doi.org/10.3390/a16020078>
- Chen, J. K., & Lee, W. Z. (2019). An introduction of NoSQL databases based on their categories and application industries. *Algorithms*. <https://doi.org/10.3390/a12050106>
- Cleve, A., Gobert, M., Meurice, L., Maes, J., & Weber, J. (2015). Understanding database schema evolution: A case study. *Science of Computer Programming*. <https://doi.org/10.1016/j.scico.2013.11.025>
- Curino, C., Moon, H. J., Deutsch, A., & Zaniolo, C. (2013). Automating the database schema evolution process. *VLDB Journal*. <https://doi.org/10.1007/s00778-012-0302-x>
- Domínguez, E., Lloret, J., Rubio, Á. L., & Zapata, M. A. (2008). MeDEA: A database evolution architecture with traceability. *Data and Knowledge Engineering*. <https://doi.org/10.1016/j.datak.2007.12.001>
- El Alami, A., Bahaj, M., & Khourdifi, Y. (2018). Supply of a key value database redis in-memory by data from a relational database. *19th IEEE Mediterranean Electrotechnical Conference, MELECON 2018 - Proceedings*. <https://doi.org/10.1109/MELCON.2018.8379066>
- European Commission. (2021, January 11). *Financial services – EU regulatory framework for crypto-assets*. European Commission.
- Hameurlain, A., & Morvan, F. (2016). Big data management in the cloud: Evolution or crossroad? *Communications in Computer and Information Science*. https://doi.org/10.1007/978-3-319-34099-9_2
- Hillenbrand, A., Störl, U., Nabiye, S., & Klettke, M. (2022). Self-adapting data migration in the context of schema evolution in NoSQL databases. *Distributed and Parallel Databases*. <https://doi.org/10.1007/s10619-021-07334-1>
- Khourdifi, Y., Bahaj, M., & Elalami, A. (2018). A new approach for migration of a relational database into column-oriented nosql database on hadoop. *Journal of Theoretical and Applied Information Technology*.
- Li, C., Chen, Y., & Shang, Y. (2022). A review of industrial big data for decision making in intelligent manufacturing. In *Engineering Science and Technology, an International Journal*. <https://doi.org/10.1016/j.jestch.2021.06.001>
- Lieponienė, J. (2021). Recent trends in database technology. *Baltic Journal of Modern Computing*. <https://doi.org/10.22364/BJMC.2020.8.4.06>
- Long, C. K., Agrawal, R., Trung, H. Q., & Van Pham, H. (2021). A big data framework for E-Government in Industry 4.0. In *Open Computer Science*. <https://doi.org/10.1515/comp-2020-0191>
- Lu, J., & Holubová, I. (2019). Multi-model Databases: A new journey to handle the variety of data. In *ACM Computing Surveys*. <https://doi.org/10.1145/3323214>
- Puangsaikai, W., & Puntheeranurak, S. (2017). A comparative study of relational database and key-value database for big data applications. *2017 International Electrical Engineering Congress, IEECON 2017*. <https://doi.org/10.1109/IEECON.2017.8075813>
- Rashid, A., & Sawyer, P. (2005). A database evolution taxonomy for object-oriented databases. *Journal of Software Maintenance and Evolution*. <https://doi.org/10.1002/smr.310>
- Schuler, R. E., & Kessleman, C. (2019). A High-level User-oriented Framework for Database Evolution. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3335783.3335787>

- Schuler, R., & Kesselman, C. (2021). CHiSEL: a user-oriented framework for simplifying database evolution. *Distributed and Parallel Databases*. <https://doi.org/10.1007/s10619-020-07314-x>
- Simon-Nagy, G., Fleiner, R., & Varkonyi-Koczy, A. (2020). Applicability of Multi-model Databases for Accessible Indoor Navigation. *IEEE Medical Measurements and Applications, MeMeA 2020 - Conference Proceedings*. <https://doi.org/10.1109/MeMeA49120.2020.9137149>
- Suárez-Cabal, M. J., Suárez-Otero, P., de la Riva, C., & Tuya, J. (2023). MDICA: Maintenance of data integrity in column-oriented database applications. *Computer Standards and Interfaces*. <https://doi.org/10.1016/j.csi.2022.103642>
- Vyawahare, H. R., Karde, P. P., & Thakare, V. M. (2019). An efficient graph database model. *International Journal of Innovative Technology and Exploring Engineering*. <https://doi.org/10.35940/ijitee.I7805.0881019>
- Zhou, K., Fu, C., & Yang, S. (2016). Big data driven smart energy management: From big data to big insights. In *Renewable and Sustainable Energy Reviews*. <https://doi.org/10.1016/j.rser.2015.11.050>