

UNIVERSITATEA BABEȘ-BOLYAI

Facultatea de Științe Economice și Gestiunea Afacerilor

Contabilitate și informatică de gestiune (la Sighetu Marmăției)

LUCRARE DE LICENȚĂ

Evaluarea percepțiilor și satisfacției clienților față de
serviciile bancare.

Analiza feedback-ului pentru îmbunătățirea
performanțelor utilizând tehnici de Data Mining.

Absolvent,

TUDIC ILEANA

Coordonator Științific,

Asist. univ. dr. Diana **MOISUC**

2025

Rezumat

Lucrarea analizează percepțiile și nivelul de satisfacție ale clienților față de serviciile bancare, cu accent pe identificarea factorilor care influențează intenția de recomandare a băncii. Studiul utilizează un set de date colectate prin chestionar și aplică o metodologie cantitativă, incluzând prelucrarea datelor, analiza descriptivă, modele de clasificare (regresie logistică, SVM, arbore decizional) și tehnici de segmentare prin K-means. Modelele au fost evaluate prin metrice standard precum scorul F1 și acuratețea, obținând performanțe ridicate. Rezultatele evidențiază importanța scorurilor acordate aplicației mobile și interacțiunii cu personalul, ca predictor ai loialității. Contribuțiile personale vizează atât componenta teoretică - sinteza literaturii de specialitate și alegerea metodelor - cât și componenta practică - dezvoltarea codului de analiză, interpretarea rezultatelor și formularea de direcții strategice utile pentru optimizarea serviciilor bancare.

Cuprins

<i>Introducere.....</i>	<i>5</i>
<i>Capitolul 1. Feedback-ul clienților în contextul serviciilor bancare</i>	<i>7</i>
1.1 Definiția feedback-ului în sectorul bancar	7
1.2. Rolul feedback-ului în optimizarea serviciilor financiare și bancare	8
1.3. Surse de feedback: sondaje, recenzii online, discuții directe, reclamații	9
1.4. Provocările în colectarea și analizarea feedback-ului clienților	10
<i>Capitolul 2. Tehnici de Data Mining pentru analiza feedback-ului.....</i>	<i>12</i>
2.1. Introducere în Data Mining	12
2.2. Aplicabilitatea tehnicilor de Data Mining în analiza feedback-ului clienților din sectorul bancar	14
2.3. Metode de Data Mining relevante (clasificare, clustering, arbori decizionali).....	15
2.4. Importanța prelucrării datelor pentru interpretarea corectă a feedback-ului	16
<i>Capitolul 3. Metodologia cercetării</i>	<i>18</i>
3.1. Obiectivele și cadrul general al cercetării	18
3.2. Modelul conceptual al cercetării	19
3.3. Metoda de colectare a datelor	19
3.4. Caracteristicile eșantionului	20
3.5. Metodele de analiză a datelor	21
<i>Capitolul 4. Analiza și interpretarea datelor din feedback-ul clienților</i>	<i>23</i>
4.1. Prezentarea și analiza datelor colectate	23
4.1.1. Metodologia colectării datelor	23
4.1.2. Analiza distribuțiilor și tendințelor generale ale scorurilor	24
4.1.3. Diferențieri demografice în percepția satisfacției	25
4.2. Utilizarea Tabelelor Pivot pentru analiza feedback-ului	28
4.2.1. Scoruri în funcție de gen	29
4.2.2. Scoruri în funcție de vârstă	29
4.2.3. Scoruri în funcție de utilizarea aplicației mobile	29
4.2.4. Scoruri în funcție de localitate	30
4.2.5. Statistici descriptive generale	30
4.2.6. Corelații între dimensiunile satisfacției	31
4.3. Analiza datelor utilizând instrumentele de Data Mining	33
4.3.1. Alegerea limbajului și justificarea metodologică	33
4.3.2. Prelucrarea și transformarea datelor	34
4.3.3. Construirea și evaluarea modelelor de clasificare	35
4.3.4 Segmentarea respondenților prin metode de clustering	37
4.4 Interpretarea rezultatelor și sugestii strategice	39
<i>Capitolul 5. Concluzii, limitări și direcții viitoare</i>	<i>41</i>
5.1. Sinteza obiectivelor și a abordării metodologice	41

5.2 Interpretarea rezultatelor și implicații practice	41
5.3 Limitări ale studiului	42
5.4 Direcții viitoare de cercetare	43
<i>Bibliografie</i>	44
<i>Anexa 1. Cod sursă: Fișier principal de execuție - main.py</i>	46
<i>Anexa 2. Cod sursă: Încărcarea și inspecția datelor - data_loader.py</i>	48
<i>Anexa 3. Cod sursă: Analiza descriptivă a variabilelor - descriptiva.py</i>	49
<i>Anexa 4. Cod sursă: Generarea tabelor pivot pentru analize comparative - genereaza_tabele_pivot.py</i>	51
<i>Anexa 5. Cod sursă: Interpretarea modelelor predictive - interpretare.py</i>	52
<i>Anexa 6. Cod Sursă: Preprocesarea datelor - preprocesare.py</i>	54
<i>Anexa 7. Cod Sursă: Antrenarea și evaluarea modelelor predictive - modele.py</i>	55

Introducere

În contextul transformării digitale accelerate, sectorul bancar se confruntă cu presiunea de a-și adapta serviciile la noile așteptări ale clienților. Accesul la aplicații mobile, platforme online și asistență automatizată a transformat profund modul în care consumatorii percep calitatea serviciilor financiare. Experiența clientului nu mai este definită doar de interacțiunea directă cu personalul băncii, ci de întregul parcurs digital - de la ușurința utilizării aplicației până la eficiența răspunsului în situații neprevăzute. În acest nou cadru, satisfacția nu mai este doar o expresie subiectivă, ci o sursă de informații valoroasă pentru îmbunătățirea performanței organizaționale.

Am ales această temă din convingerea că analiza satisfacției clienților nu este un simplu demers teoretic, ci o investiție practică în calitatea serviciilor și în fidelizarea consumatorilor. În special în domeniul bancar, unde loialitatea este fragilă, iar concurența intensă, capacitatea de a detecta din timp nemulțumirile sau așteptările neîmplinite poate reprezenta un diferențiator strategic. Instituțiile care ignoră feedback-ul clienților riscă să piardă teren în fața competitorilor mai agili și mai receptivi la nevoile pieței. Prin urmare, înțelegerea satisfacției și loialității nu mai reprezintă doar obiective de marketing, ci componente fundamentale ale unei guvernante moderne și eficiente.

Într-o economie bazată tot mai mult pe date, metodele de tip data mining oferă soluții performante pentru extragerea și interpretarea informațiilor relevante din volume mari de date. Aceste tehnici nu se limitează la observarea comportamentelor, ci permit anticiparea acestora, prin identificarea tiparelor ascunse și a relațiilor semnificative dintre variabile. Astfel, ele devin instrumente indispensabile pentru organizațiile care urmăresc personalizarea serviciilor și luarea deciziilor informate.

Studiile recente susțin eficiența investițiilor în analiza avansată a datelor. Potrivit raportului *Competing on Analytics* realizat de Davenport și Harris (2017), companiile care folosesc intensiv tehnici analitice în procesul decizional înregistrează cu 5–6% profitabilitate mai mare decât concurenții lor. La nivelul sectorului bancar, McKinsey & Company (2021) indică faptul că instituțiile care implementează analize predictive asupra satisfacției clienților pot obține o creștere de până la 30% a ratei de retenție și cu până la 20% a veniturilor per client.

Lucrarea de față îmbină analiza teoretică și cercetarea aplicativă pentru a investiga factorii care influențează satisfacția și intenția de recomandare în rândul clienților bancari. Studiul se bazează pe un set de date colectate prin chestionar și valorifică metode de prelucrare, modelare și segmentare, specifice învățării automate.

Structura lucrării este următoarea: Capitolul 1 definește conceptele cheie legate de feedback și satisfacție în domeniul bancar. Capitolul 2 prezintă cadrul metodologic, iar Capitolul 3 detaliază procesul de colectare și prelucrare a datelor. În Capitolul 4 sunt aplicate modele predictive și tehnici de clustering pentru identificarea tiparelor comportamentale. Capitolul 5 conține concluziile generale, limitările studiului și direcțiile viitoare de cercetare.

Capitolul 1. Feedback-ul clienților în contextul serviciilor bancare

1.1 Definiția feedback-ului în sectorul bancar

Feedback-ul reprezintă o interacțiune esențială între clienți și instituțiile financiare, având un rol central în procesul de adaptare și îmbunătățire continuă a serviciilor oferite. În termeni generali, feedback-ul poate fi definit ca totalitatea informațiilor transmise de clienți către organizații, reflectând percepțiile, așteptările și experiențele acestora în raport cu produsele și serviciile primite. În sectorul bancar, acest concept este deosebit de valoros, având în vedere complexitatea interacțiunilor și diversitatea nevoilor clienților.

Literatura de specialitate descrie feedback-ul ca „vocea clientului” (Kotler & Keller, 2016), un mecanism fundamental prin care organizațiile financiare pot identifica punctele forte și slabe ale serviciilor oferite. Studiile sugerează că analiza feedback-ului ajută la formularea de strategii centrate pe client, conducând la îmbunătățirea satisfacției și loialității acestora. De exemplu, Alpha Bank România a demonstrat succesul utilizării feedback-ului prin implementarea unei platforme avansate de Contact Center, ceea ce a condus la o creștere cu 20% a eficienței gestionării apelurilor și la reducerea timpilor de așteptare (Studiu de caz Alpha Bank, 2023).

Într-un studiu amplu, Parasuraman, Zeithaml și Berry (1988) au introdus modelul SERVQUAL, care subliniază cinci dimensiuni esențiale pentru evaluarea calității serviciilor: tangibilitatea, fiabilitatea, responsabilitatea, siguranța și empatia. Acest cadru oferă o bază teoretică solidă pentru înțelegerea și analiza feedback-ului. Spre exemplu, timpii mari de așteptare în sucursale indică probleme de fiabilitate, în timp ce disponibilitatea unui consilier financiar reflectă valoarea empatiei.

Ulterior, Cronin și Taylor (1992) au dezvoltat modelul SERVPERF ca alternativă la SERVQUAL, argumentând că performanța percepută este un predictor mai relevant al satisfacției decât diferența dintre așteptări și percepții. SERVPERF simplifică evaluarea calității serviciilor prin concentrarea exclusivă asupra experienței reale a clientului, eliminând variabila subiectivă a așteptărilor. Acest model este adesea preferat în cercetările cantitative moderne datorită robusteții sale statistice și capacității de a oferi rezultate mai consistente în contexte cu date structurale, precum chestionarele standardizate (Cronin & Taylor, 1992).

În contextul serviciilor digitale, modelul E-S-QUAL propus de Parasuraman, Zeithaml și Malhotra (2005) extinde analiza calității în mediul online. Acest model include dimensiuni specifice serviciilor electronice precum eficiența, disponibilitatea sistemului, confidențialitatea

și capacitatea de a reacționa prompt la cererile clienților. Este deosebit de relevant pentru instituțiile bancare care oferă aplicații mobile sau platforme de internet banking. Deși mai puțin utilizat în cercetările generale, E-S-QUAL oferă un cadru actualizat și adaptat pentru evaluarea calității în interacțiunile digitale cu clienții.

De asemenea, feedback-ul poate fi colectat din surse variate, precum sondaje structurate, recenzii online sau reclamații oficiale, și reprezintă o resursă valoroasă pentru optimizarea continuă a serviciilor. Raportul McKinsey & Company (2021) recomandă băncilor din România să integreze în mod sistematic informațiile obținute de la clienți – atât din interacțiunile digitale, cât și din cele fizice – în strategiile operaționale, pentru a răspunde mai eficient așteptărilor acestora într-un mediu competitiv și în schimbare.

În acest context, Davenport și Harris (2017) arată că organizațiile care utilizează analize avansate ale datelor, inclusiv feedback-ul clienților, pot obține un avantaj competitiv clar. În sectorul bancar, analiza sistematică a reclamațiilor și sugestiilor clienților permite identificarea punctelor slabe și adoptarea unor soluții tehnologice menite să îmbunătățească experiența utilizatorilor.

1.2. Rolul feedback-ului în optimizarea serviciilor financiare și bancare

Feedback-ul clienților constituie un element important în procesele de optimizare și adaptare continuă a serviciilor financiare și bancare, fiind un punct de plecare pentru identificarea și soluționarea problemelor sistemice. Acesta oferă instituțiilor financiare o înțelegere profundă a preferințelor, nevoilor și frustrărilor clienților, ghidând astfel procesul de luare a deciziilor strategice.

Cercetările lui Cahyono, Sunarti și Rofiq (2020) arată că experiența clientului influențează direct satisfacția, iar satisfacția determină loialitatea. Într-un context bancar, feedback-ul clientului poate fi considerat o sursă esențială pentru optimizarea experienței, ceea ce contribuie la consolidarea relației cu clientul și la creșterea valorii brandului.

Pe lângă personalizare, feedback-ul este esențial pentru eficientizarea proceselor operaționale. Plângerile legate de timpii mari de așteptare în sucursale sau de complexitatea excesivă a procedurilor financiare pot indica necesitatea unei regândiri a proceselor interne. Băncile care implementează soluții tehnologice moderne, cum ar fi inteligența artificială pentru automatizarea sarcinilor repetitive sau chatbot-urile pentru suport 24/7, pot răspunde mai rapid și mai eficient cerințelor clienților. În plus, feedback-ul detaliat oferă o bază solidă pentru prioritizarea investițiilor în tehnologie.

Un alt aspect important al utilizării feedback-ului este influența asupra reputației

instituției. Răspunsul prompt și eficient la problemele semnalate de clienți reflectă angajamentul față de satisfacția acestora, consolidând încrederea și loialitatea. Raportul McKinsey & Company (2021) subliniază importanța integrării sistematice a feedback-ului în deciziile operaționale, ca parte a unui efort mai amplu de îmbunătățire a experienței clientului și de adaptare la nevoile în continuă schimbare ale acestuia.

1.3. Surse de feedback: sondaje, recenzii online, discuții directe, reclamații

Feedback-ul clienților poate fi colectat printr-o varietate de surse, fiecare dintre acestea având avantaje și provocări specifice. Printre cele mai frecvente metode de obținere a feedback-ului se numără sondajele, recenziile online, discuțiile directe și reclamațiile oficiale. Fiecare dintre aceste surse contribuie la formarea unei perspective cuprinzătoare asupra experienței utilizatorilor.

Sondajele sunt una dintre cele mai utilizate modalități de colectare a feedback-ului datorită capacității lor de a acoperi un număr mare de respondenți într-un mod structurat. Instrumente precum Net Promoter Score (NPS) sunt utilizate pe scară largă pentru a evalua probabilitatea ca un client să recomande serviciile unei bănci. Cu toate acestea, limitarea sondajelor constă în gradul lor redus de detaliu, ceea ce poate necesita metode suplimentare pentru obținerea unor informații mai complexe.

Recenziile online oferă o perspectivă valoroasă asupra experienței clienților, fiind accesibile pe platforme publice precum Google Reviews sau rețelele sociale. Acestea permit clienților să-și exprime opiniile în mod deschis, adesea evidențiind aspecte specifice care pot să nu apară în sondaje. Totuși, interpretarea recenziilor necesită precauție, deoarece acestea pot fi influențate de factori subiectivi sau de evenimente izolate.

Discuțiile directe cu clienții, fie în sucursale, fie prin canale digitale sau telefonice, constituie o altă sursă valoroasă de feedback. Aceste interacțiuni permit băncilor să obțină informații detaliate despre problemele și nevoile clienților, facilitând soluționarea rapidă a acestora. Totuși, colectarea și centralizarea informațiilor din astfel de interacțiuni pot necesita resurse suplimentare.

Reclamațiile oficiale sunt o altă sursă esențială de feedback, deoarece acestea semnalează adesea probleme critice care necesită atenție imediată. Analiza acestor reclamații poate ajuta la identificarea problemelor sistemice și la implementarea unor măsuri corective eficiente. Gestionarea adecvată a reclamațiilor demonstrează angajamentul instituției de a răspunde nevoilor clienților și de a îmbunătăți constant serviciile oferite.

Astfel, utilizarea combinată a acestor surse de feedback oferă o imagine holistică asupra

experienței clienților, permițând băncilor să identifice și să abordeze în mod eficient provocările întâmpinate de utilizatori.

1.4. Provocările în colectarea și analizarea feedback-ului clienților

Deși feedback-ul clienților reprezintă o resursă valoroasă pentru optimizarea serviciilor bancare, procesul de colectare și analizare a acestuia este adesea asociat cu numeroase provocări. Acestea includ dificultăți legate de volumul și diversitatea datelor, calitatea informațiilor colectate, integrarea datelor din surse variate, precum și respectarea normelor privind confidențialitatea și securitatea datelor.

Un prim obstacol este reprezentat de volumul mare de date generate zilnic din multiple surse, inclusiv sondaje, recenzii online și reclamații. Analizarea acestor date în timp util necesită infrastructuri tehnologice robuste și algoritmi avansați pentru extragerea informațiilor relevante. De exemplu, utilizarea instrumentelor de big data și inteligență artificială poate facilita procesarea unui număr mare de feedback-uri, dar implementarea acestor soluții implică costuri semnificative și expertiză tehnică.

Calitatea datelor colectate constituie o altă provocare majoră. Feedback-ul poate fi incomplet, ambiguu sau influențat de factori subiectivi. Spre exemplu, recenziile online pot reflecta mai degrabă emoții temporare decât probleme reale ale serviciilor. Pentru a depăși aceste limitări, băncile trebuie să implementeze metode riguroase de validare și curățare a datelor, asigurându-se astfel că analizele efectuate sunt precise și relevante.

Diversitatea canalelor de feedback reprezintă o dificultate suplimentară în procesul de integrare a datelor. Consolidarea informațiilor din surse diverse într-o bază de date unificată necesită soluții tehnologice care să asigure interoperabilitatea și consistența datelor. Lipsa unui sistem integrat poate duce la pierderi de informații sau interpretări incorecte, afectând astfel calitatea deciziilor strategice.

Nu în ultimul rând, confidențialitatea și securitatea datelor sunt aspecte critice în procesul de colectare și analizare a feedback-ului. Conform regulamentelor internaționale, cum ar fi GDPR, instituțiile financiare sunt obligate să protejeze informațiile personale ale clienților. Implementarea unor măsuri adecvate de securitate, precum criptarea datelor și accesul limitat la informații sensibile, poate preveni breșele de securitate, dar implică, de asemenea, investiții suplimentare și resurse considerabile.

Pe lângă aceste dificultăți tehnice și organizaționale, utilizarea modelelor teoretice consacrate în analiza calității serviciilor implică provocări conceptuale și metodologice proprii. Modelul SERVQUAL, deși este unul dintre cele mai utilizate în cercetările de marketing și

servicii, a fost criticat pentru presupunerea că satisfacția rezultă din diferența dintre așteptările și percepțiile clienților. Această abordare poate introduce un grad mare de subiectivitate, deoarece așteptările diferă semnificativ între indivizi și contexte. În plus, măsurarea corectă a acestor așteptări este dificilă și adesea inconsistentă, ceea ce poate compromite validitatea rezultatelor.

Modelul SERVPERF, propus ca o alternativă la SERVQUAL, încearcă să elimine aceste limitări prin concentrarea exclusivă asupra performanței percepute de client. Această abordare este considerată mai robustă din punct de vedere statistic și mai relevantă în contexte cantitative, cum este cazul studiilor bazate pe chestionare standardizate. Totuși, ignorarea așteptărilor inițiale poate conduce la o înțelegere incompletă a modului în care clienții își formează evaluările și deciziile de loialitate, mai ales în industrii unde anticiparea calității joacă un rol important, precum serviciile financiare.

În ceea ce privește mediile digitale, modelul E-S-QUAL oferă o perspectivă modernizată, adaptată la interacțiunile electronice dintre client și instituție. Acesta introduce dimensiuni relevante pentru serviciile online, cum ar fi eficiența, fiabilitatea sistemelor, confidențialitatea datelor și capacitatea de reacție a platformelor. Deși util în analiza serviciilor electronice precum aplicațiile bancare și internet banking, modelul E-S-QUAL este mai greu de aplicat fără întrebări special concepute, ceea ce necesită fie acces la scale validate internațional, fie un efort suplimentar de adaptare și testare. În plus, comparabilitatea cu studiile clasice bazate pe SERVQUAL sau SERVPERF este redusă, deoarece dimensiunile măsurate sunt diferite.

Astfel, aplicarea acestor modele în analiza feedback-ului presupune un efort de echilibrare între rigoarea teoretică, fezabilitatea practică și relevanța contextuală. Cercetătorii și practicienii trebuie să aleagă modelul potrivit în funcție de tipul serviciilor analizate, canalele prin care este colectat feedback-ul și scopul final al cercetării. În cazul de față, integrarea cadrului SERVQUAL ca referință generală este justificată de recunoașterea sa largă și aplicabilitatea directă la dimensiunile măsurate prin chestionar, însă limitările sale vor fi compensate prin abordări cantitative complementare, precum clasificarea și segmentarea bazată pe date reale.

Capitolul 2. Tehnici de Data Mining pentru analizarea feedback-ului

2.1. Introducere în Data Mining

Data Mining, cunoscut și sub denumirea de explorare a datelor, constituie un domeniu interdisciplinar care combină tehnici avansate din statistică, informatică și inteligență artificială pentru extragerea informațiilor valoroase din seturi mari de date. Scopul principal al acestui proces este identificarea tiparelor ascunse, descoperirea relațiilor complexe și anticiparea tendințelor emergente, toate acestea având potențialul de a susține procese decizionale bine fundamentate (Han et al., 2012). Dezvoltarea sa rapidă poate fi atribuită progreselor tehnologice majore, precum creșterea capacității de stocare, viteza procesării datelor și evoluția algoritmilor avansați. Într-un context global marcat de explozia volumelor de date, Data Mining a devenit o componentă esențială în strategiile organizațiilor moderne, inclusiv ale instituțiilor financiare, unde datele complexe necesită abordări analitice sofisticate pentru a fi valorificate în mod eficient (Chaudhuri & Madigan, 2016).

Data Mining își are originile în intersecția mai multor discipline, precum statistica, inteligența artificială și econometria aplicată, fiind dezvoltat inițial pentru a facilita analiza automată a datelor și construirea modelelor predictive. Deși termenul s-a consolidat abia în anii '90, fundamentele sale teoretice au fost utilizate anterior în cercetarea experimentală și în domeniul decizional. Extinderea digitalizării a facilitat generarea unor volume impresionante de date nestructurate, cum ar fi tranzacțiile financiare, recenziile online și interacțiunile prin rețelele sociale. În sectorul bancar, Data Mining a devenit indispensabil, contribuind la procesarea acestor seturi mari de date pentru a optimiza operațiunile, a înțelege comportamentul clienților și a identifica oportunități de creștere. Spre exemplu, algoritmi de clasificare pot fi folosiți pentru segmentarea portofoliului de clienți în funcție de risc sau preferințe, în timp ce tehnicile de clustering identifică grupuri comportamentale cu caracteristici similare (Han et al., 2012).

Explorarea datelor include mai multe etape fundamentale, fiecare având un rol critic în transformarea datelor brute în informații utile. Prima etapă, prelucrarea datelor, se concentrează pe curățarea și standardizarea acestora, eliminând valorile lipsă, inconsistențele și erorile (Han et al., 2012). Acest proces este vital pentru asigurarea calității analizelor ulterioare. O etapă ulterioară implică aplicarea tehnicilor avansate, cum ar fi clustering-ul și clasificarea, pentru segmentarea datelor. De exemplu, clustering-ul permite identificarea grupurilor omogene de clienți, cum ar fi cei care preferă tranzacțiile digitale, oferind astfel instituțiilor financiare

posibilitatea de a personaliza serviciile și produsele pentru segmente specifice. Aceste analize contribuie la îmbunătățirea relațiilor cu clienții și la dezvoltarea de strategii centrate pe client, așa cum arată Davenport și Harris (2017), care subliniază importanța utilizării analiticii avansate pentru a obține un avantaj competitiv.

Modelele predictive reprezintă un alt pilon al Data Mining, având capacitatea de a anticipa tendințele viitoare pe baza datelor istorice. Printre tehnicile predictive utilizate în mod frecvent se numără arborii de decizie, rețelele neuronale și regresia logistică. În sectorul bancar, aceste modele sunt critice pentru luarea deciziilor legate de riscurile de creditare și prevenirea fraudelor. Spre exemplu, un model predictiv bine construit poate determina cu o mare acuratețe probabilitatea ca un client să devină insolubil, oferind instituțiilor financiare oportunitatea de a interveni preventiv. Modelele predictive nu doar că reduc riscurile asociate activităților financiare, dar optimizează și alocarea resurselor organizaționale, contribuind la creșterea eficienței generale a instituției (Han et al., 2012).

Tehnologiile emergente precum big data și inteligența artificială au transformat radical potențialul Data Mining, făcând posibilă analiza datelor nestructurate (Han et al., 2012). Feedback-ul clienților exprimat sub formă de recenzii textuale sau comentarii în rețelele sociale poate fi procesat și analizat cu ajutorul algoritmilor de învățare automată. Analiza sentimentului, o ramură a Data Mining, permite evaluarea percepțiilor clienților asupra serviciilor și produselor oferite (Liu, 2012). De exemplu, un algoritm de învățare automată poate extrage rapid temele dominante din feedback-ul textului liber, identificând atât punctele forte, cât și problemele recurente ale serviciilor bancare. Aceste analize nu doar că îmbunătățesc satisfacția clienților, dar oferă și date valoroase pentru îmbunătățirea continuă a serviciilor și produselor. Așa cum subliniază Davenport și Harris (2017), utilizarea inteligenței a datelor - inclusiv a feedback-ului - poate conferi organizațiilor un avantaj competitiv clar.

Infrastructurile moderne de calcul distribuit, cum ar fi Hadoop și Spark, permit procesarea unor volume imense de date într-un timp extrem de scurt, asigurând astfel scalabilitate și eficiență. Aceste tehnologii au fost complementate de utilizarea tehnicilor avansate, precum descoperirea de asocieri, care evidențiază relațiile subtile între variabile aparent independente (Han et al., 2012). Spre exemplu, o astfel de tehnică poate dezvălui că utilizatorii care contractează credite ipotecare sunt mai predispuși să acceseze și produse financiare conexe, precum asigurările de locuință. Această cunoaștere oferă instituțiilor financiare oportunitatea de a crea oferte personalizate și de a crește rata de retenție a clienților. În plus, abordările hibride care combină analiza de asocieri cu modelele predictive oferă o perspectivă mai completă asupra comportamentului consumatorilor, maximizând astfel

valoarea informațiilor extrase.

2.2. Aplicabilitatea tehnicilor de Data Mining în analiza feedback-ului clienților din sectorul bancar

Tehnicile de Data Mining au devenit o componentă centrală în strategiile instituțiilor financiare care urmăresc să valorifice datele generate de interacțiunile cu clienții. Prin intermediul acestor metode, organizațiile din sectorul bancar pot extrage informații relevante din volume mari de date, în scopul de a înțelege mai bine comportamentele, preferințele și gradul de satisfacție al clienților.

Un exemplu notabil de aplicare a Data Mining este prevenirea fraudelor financiare. Algoritmii avansați pot detecta anomalii în modelele de tranzacții, semnalând activități suspecte care necesită investigații suplimentare. În același timp, modelele predictive sunt folosite pentru a optimiza strategiile de marketing, prin personalizarea ofertelor și identificarea clienților cu un potențial ridicat de loializare (Davenport & Harris, 2017).

Algoritmii de tip rețele neuronale profunde sunt utilizați frecvent pentru a recunoaște tipare complexe în datele financiare, îmbunătățind astfel atât prevenirea fraudelor, cât și eficiența operațională (LeCun et al., 2015). Aceste aplicații demonstrează impactul strategic al Data Mining asupra competitivității și sustenabilității instituțiilor bancare moderne.

În contextul digitalizării accelerate a serviciilor bancare, cantitatea de feedback generată de clienți a crescut exponențial. Acest feedback, exprimat prin recenzii online, răspunsuri la sondaje, conversații în chat-uri sau mesaje pe rețelele sociale, conține informații valoroase despre percepțiile și satisfacția clienților. Pentru a valorifica acest potențial, instituțiile financiare recurg la tehnici de Data Mining, care permit extragerea, structurarea și interpretarea automată a acestor date complexe.

Prin intermediul Data Mining, băncile pot detecta tipare ascunse în opiniile clienților, pot clasifica automat comentariile în funcție de tonalitate (pozitivă, neutră, negativă) și pot corela aceste rezultate cu alți indicatori, cum ar fi frecvența utilizării serviciilor sau tipul de produs bancar folosit. Spre exemplu, utilizarea analizei de sentiment asupra recenziilor online poate evidenția gradul de satisfacție al clienților față de aplicația mobilă a băncii, iar clustering-ul poate grupa clienții în funcție de probleme similare întâmpinate – precum dificultăți în accesarea serviciilor digitale sau lipsa de promptitudine în răspunsurile din call center.

Aceste instrumente oferă o înțelegere mult mai nuanțată decât metodele tradiționale de analiză a feedback-ului, deoarece pot procesa atât date structurate (răspunsuri codificate din sondaje), cât și nestructurate (texte libere, opinii detaliate). Astfel, Data Mining facilitează

obținerea unor perspective integrate asupra experienței clientului și permite intervenții rapide și targetate pentru remedierea deficiențelor.

Un alt beneficiu major este posibilitatea de a corela feedback-ul clienților cu performanțele interne ale băncii. De exemplu, dacă mai multe recenzii semnalează întârzieri în procesarea cererilor de credit, aceste date pot fi comparate cu timpii efectiv de procesare din sistemele interne, validând sau infirmând percepțiile clienților. În acest mod, Data Mining susține decizii bazate pe evidențe și contribuie la prioritizarea corectă a investițiilor și inițiativelor de îmbunătățire.

Tot mai frecvent, instituțiile bancare adoptă instrumente de analiză a satisfacției clienților bazate pe algoritmi de clasificare, segmentare și predicție. Aceste metode permit nu doar identificarea nivelului de satisfacție, ci și corelarea acestuia cu variabile demografice sau comportamentale, oferind o imagine de ansamblu mai fidelă asupra experienței clientului. În acest sens, algoritmi precum regresia logistică, arborii decizionali sau SVM sunt valoroși pentru a înțelege factorii care influențează loialitatea, în timp ce clustering-ul permite gruparea clienților pe baza răspunsurilor similare.

Prin urmare, aplicabilitatea tehnicilor de Data Mining în analiza feedback-ului clienților nu se limitează la simpla prelucrare a datelor, ci extinde capacitatea instituțiilor financiare de a reacționa proactiv la nevoile clienților și de a construi o relație de încredere și loialitate cu aceștia. Aceste metode permit transformarea unui volum mare de informații în perspective strategice utile pentru luarea deciziilor și îmbunătățirea continuă a calității serviciilor (Davenport & Harris, 2017; Han et al., 2012).

2.3. Metode de Data Mining relevante (clasificare, clustering, arbori decizionali)

Clustering-ul, spre deosebire de clasificare, este o metodă de învățare nesupravegheată care are ca scop identificarea structurilor sau grupurilor omogene în date, fără a defini aceste grupuri în prealabil (Han et al., 2012). În contextul analizei feedback-ului, clustering-ul este utilizat pentru a descoperi tipare ascunse și pentru a segmenta clienții în funcție de opinii, comportamente sau nemulțumiri comune. Spre exemplu, clienții care exprimă frecvent preocupări legate de „lipsa de transparență a comisioanelor” pot fi grupați într-un cluster distinct, permițând instituției financiare să formuleze strategii de comunicare și intervenție personalizate. Algoritmi populari de clustering includ K-means, DBSCAN și metodele aglomerative (hierarchical clustering) (Han et al., 2012). Clustering-ul contribuie astfel la o segmentare mai fină a pieței și facilitează personalizarea serviciilor oferite.

Arborii decizionali reprezintă o tehnică de învățare supravegheată utilizată pentru

construirea de modele predictive bazate pe reguli logice de tip „dacă... atunci” (Breiman et al., 1984). În analiza feedback-ului, arborii decizionali sunt valoroși pentru identificarea principalelor variabile care influențează satisfacția clienților. De exemplu, un model de arbore decizional poate arăta că nemulțumirea este mai probabilă atunci când „timpul de așteptare depășește 10 minute” și „nu este oferită asistență personalizată”. Un avantaj major al acestor metode constă în capacitatea de a oferi o reprezentare vizuală și intuitivă a relațiilor dintre variabile, facilitând interpretarea și luarea deciziilor la nivel managerial. Printre algoritmi frecvent utilizați se numără C4.5 (Quinlan, 1993), CART (Breiman et al., 1984) și Random Forest (Breiman, 2001).

Este important de menționat că aceste metode nu sunt utilizate izolat, ci pot fi combinate în funcție de complexitatea datelor și de obiectivele analizei. De exemplu, un proces complet de analiză a feedback-ului poate include: (1) prelucrarea datelor brute, (2) clasificarea automată a comentariilor în funcție de tematică sau tonalitate, (3) clustering pentru segmentarea comportamentală a clienților și (4) arbori decizionali pentru explicarea factorilor care influențează satisfacția sau pentru predicția comportamentului viitor.

2.4. Importanța prelucrării datelor pentru interpretarea corectă a feedback-ului

Prelucrarea datelor reprezintă o etapă esențială în analiza feedback-ului clienților, în special în sectorul bancar, unde informațiile sunt adesea eterogene, nestructurate și colectate prin canale multiple. Dacă această etapă este neglijată sau realizată superficial, rezultatele obținute pot fi inexacte și pot conduce la interpretări eronate.

De regulă, feedback-ul clienților este exprimat prin text liber, scoruri de satisfacție, răspunsuri multiple sau comentarii nestructurate. Pentru a putea fi analizate cu ajutorul tehnicilor de Data Mining, aceste date trebuie mai întâi curățate, standardizate și transformate într-un format coerent. Acest proces presupune, de exemplu, eliminarea valorilor lipsă, corectarea greșelilor de scriere sau convertirea textelor în attribute analizabile. În cazul analizelor automate, datele textuale sunt prelucrate cu ajutorul tehnologiilor de procesare a limbajului natural (NLP), care extrag informații relevante și reduc ambiguitatea.

Un pas critic în această etapă este selecția trăsăturilor relevante, mai ales când volumul de date este foarte mare. Aceasta permite păstrarea doar a elementelor semnificative pentru analiză, reducând timpul de procesare și îmbunătățind precizia modelelor. De asemenea, etichetarea corectă a datelor este esențială în cazul metodelor supravegheate, asigurând antrenarea eficientă a algoritmilor de clasificare sau predicție.

Fără o prelucrare riguroasă, chiar și cele mai performante metode analitice pot da

rezultate înșelătoare. Incoerențele din date, expresiile ambigue sau limbajul informal pot afecta acuratețea interpretării sentimentelor și a concluziilor finale. Din acest motiv, prelucrarea datelor nu trebuie văzută ca o simplă etapă preliminară, ci ca un proces fundamental care influențează direct calitatea și validitatea întregii analize.

În plus, calitatea prelucrării influențează direct performanțele algoritmilor utilizați. Un set de date curat, bine etichetat și echilibrat contribuie la obținerea unor modele de clasificare sau predicție mai precise, reducând rata de eroare și crescând capacitatea de generalizare. În domeniul bancar, această acuratețe este esențială pentru luarea deciziilor strategice, precum evaluarea riscului de credit sau personalizarea campaniilor de marketing. Astfel, prelucrarea nu doar pregătește datele, ci contribuie activ la valoarea analizei finale.

Capitolul 3. Metodologia cercetării

3.1. Obiectivele și cadrul general al cercetării

Cercetarea de față își propune să investigheze percepțiile și nivelul de satisfacție al clienților față de serviciile oferite de instituțiile bancare, utilizând metode cantitative moderne de analiză a datelor. Având în vedere creșterea importanței feedback-ului clientului în procesele decizionale strategice, studiul urmărește să evidențieze factorii care influențează experiența bancară și modul în care aceștia determină comportamentele de loialitate, cum ar fi intenția de a recomanda banca altor persoane.

În contextul actual, în care sectorul bancar este marcat de digitalizare accelerată, competiție intensificată și exigențe sporite din partea consumatorilor, înțelegerea percepțiilor clienților a devenit esențială pentru menținerea relevanței pe piață. Instituțiile financiare care reușesc să valorifice în mod eficient feedback-ul primit pot adopta decizii fundamentate, pot adapta serviciile oferite la nevoile reale ale clienților și pot consolida relații durabile cu aceștia. În acest sens, cercetarea prezentă oferă o abordare sistematică asupra modului în care poate fi evaluată satisfacția, folosind instrumente statistice și algoritmi de învățare automată pentru a interpreta datele într-un mod relevant și aplicabil.

Fundamentul teoretic al cercetării este reprezentat de modelele consacrate de evaluare a calității serviciilor, în special modelul SERVQUAL, care oferă un cadru pentru măsurarea satisfacției pe dimensiuni precum fiabilitatea, empatia, siguranța sau tangibilitatea serviciilor. Acest cadru este completat de o abordare analitică orientată spre identificarea relațiilor ascunse dintre variabilele măsurate, prin aplicarea unor tehnici de clasificare, segmentare și modelare predictivă.

Obiectivul general al cercetării constă în evaluarea percepțiilor și satisfacției clienților față de serviciile bancare, prin corelarea feedback-ului colectat cu variabile relevante din profilul respondentului și cu intenția declarată de loialitate.

La nivel aplicativ, se urmărește măsurarea satisfacției clienților în funcție de mai multe dimensiuni reflectate în întrebările chestionarului, identificarea relațiilor dintre scorurile oferite și factorii individuali, precum și detectarea unor tipare recurente de răspuns. O componentă importantă a cercetării constă în utilizarea algoritmilor de clasificare pentru a anticipa comportamentul clienților în raport cu intenția de a recomanda banca, completată de tehnici de clustering care permit gruparea respondenților în funcție de asemănările dintre răspunsuri.

Prin intermediul acestei cercetări, se urmărește nu doar interpretarea feedback-ului în

termeni descriptivi, ci și obținerea unor perspective predictive și explicative, care pot fundamenta decizii strategice în sectorul bancar.

3.2. Modelul conceptual al cercetării

Modelul conceptual al cercetării se fundamentează pe ideea că satisfacția clienților față de serviciile bancare poate fi înțeleasă și evaluată în mod structurat prin integrarea cadrului teoretic SERVQUAL cu instrumente moderne de analiză a datelor. Acest model urmărește să coreleze percepțiile clienților exprimate prin răspunsuri standardizate la întrebări de satisfacție cu intenția lor de loialitate, exprimată prin disponibilitatea de a recomanda banca altor persoane.

Dimensiunile identificate în modelul SERVQUAL - fiabilitate, responsabilitate, empatie, siguranță și aspecte tangibile – sunt reflectate în itemii incluși în chestionar. Astfel, fiecare întrebare urmărește să măsoare una dintre aceste dimensiuni, oferind o bază clară pentru analiza satisfacției din mai multe perspective. Răspunsurile au fost formulate pe o scală Likert de la 1 la 5, permițând cuantificarea nivelului de satisfacție și facilitând interpretarea numerică a percepțiilor.

Pe lângă dimensiunile serviciului, modelul conceptual include și variabile demografice precum vârsta, sexul sau localitatea de reședință, care permit o segmentare suplimentară a eșantionului și identificarea unor diferențe relevante în modul în care este percepută calitatea serviciilor. De asemenea, modelul ia în considerare comportamente legate de utilizarea serviciilor digitale, cum ar fi frecvența utilizării aplicației mobile sau a platformei online.

Structura modelului permite aplicarea unor metode statistice și algoritmice de analiză, precum clasificarea binară, care vizează anticiparea loialității clienților, sau clusteringul, care identifică grupuri de respondenți cu tipare similare de răspuns. În acest sens, fiecare componentă a modelului este gândită astfel încât să susțină atât o interpretare descriptivă, cât și una predictivă, creând o punte între fundamentele teoretice și analiza practică a datelor.

Prin definirea acestui model conceptual, cercetarea își propune să ofere un cadru coerent de înțelegere a relației dintre satisfacția exprimată, caracteristicile individuale ale clienților și intenția de loialitate, punând bazele unor concluzii relevante și aplicabile în practică.

3.3. Metoda de colectare a datelor

Datele analizate în cadrul acestei cercetări au fost colectate prin intermediul unui chestionar online, distribuit către un eșantion diversificat de respondenți. Formularul a fost conceput astfel încât să surprindă aspecte esențiale ale experienței bancare, concentrându-se pe dimensiuni

precum interacțiunea cu personalul, claritatea informațiilor oferite, eficiența aplicației mobile și utilitatea platformei online.

Chestionarul a fost realizat și distribuit prin platforma Google Forms, oferind respondenților posibilitatea de a participa voluntar, anonim și în condiții de confidențialitate. Acesta a fost structurat în trei secțiuni principale: date demografice de bază, evaluarea satisfacției printr-o serie de afirmații pe o scală Likert de la 1 la 5 și o întrebare referitoare la intenția de a recomanda banca. Răspunsurile obținute reflectă atât percepții generale, cât și reacții față de servicii specifice, cum ar fi aplicația mobilă sau interacțiunea cu personalul.

Perioada de colectare a datelor s-a întins pe durata a două săptămâni, timp în care respondenții au fost încurajați să completeze formularul prin distribuire în rețele sociale, grupuri de discuții și prin recomandări directe. Nu au fost aplicate criterii de selecție restrictive, însă s-a urmărit obținerea unei diversități geografice și demografice relevante pentru contextul cercetării. Chestionarul a fost completat de un total de 97 de participanți.

Această metodă de colectare a permis obținerea rapidă a unui volum rezonabil de date, în format digital, ușor de prelucrat statistic. Alegerea formatului online a fost justificată de accesibilitatea ridicată, eficiența costurilor și adaptabilitatea la contextul actual, în care interacțiunile directe sunt tot mai frecvent înlocuite de soluții digitale.

3.4. Caracteristicile eșantionului

Eșantionul utilizat în cadrul acestei cercetări este compus din 97 de respondenți care au completat chestionarul online în perioada alocată colectării datelor. Deși selecția participanților nu a urmat un plan de eșantionare probabilistic, distribuirea formularului în medii online variate a permis obținerea unui grad satisfăcător de diversitate în rândul respondenților, atât din punct de vedere geografic, cât și demografic.

Din perspectiva repartiției geografice, participanții provin din mai multe localități urbane și rurale din România, incluzând centre regionale precum Cluj-Napoca, București, Baia Mare, Timișoara și Târgu Mureș, dar și orașe mai mici și localități rurale precum Sighetu Marmăției, Vișeu de Sus, Remeți și Crăciunești. Această diversitate oferă o imagine mai nuanțată asupra percepțiilor clienților din medii diferite față de serviciile bancare.

În ceea ce privește distribuția pe sexe, eșantionul este echilibrat, incluzând atât femei, cât și bărbați, fără o dominanță pronunțată a unui gen. Categoriile de vârstă acoperite sunt variate, cuprinzând tineri sub 25 de ani, persoane adulte între 25 și 54 de ani și un procent mai redus de respondenți peste 55 de ani. Această diversitate permite o analiză comparativă a percepțiilor în funcție de vârstă.

Participanții prezintă, de asemenea, grade diferite de familiaritate cu serviciile digitale bancare. O parte semnificativă a respondenților utilizează aplicații mobile și platforme online pentru gestionarea activităților bancare, în timp ce alții interacționează predominant prin canale tradiționale. Această diferență este relevantă în analiza nivelului de satisfacție raportat la tipul de experiență bancară avută.

Caracteristicile eșantionului, deși obținute printr-o metodă de selecție convenabilă, oferă o bază suficientă pentru analiză și permit identificarea unor tipare relevante în cadrul populației investigate.

3.5. Metodele de analiză a datelor

Analiza datelor colectate s-a realizat printr-o combinație de metode statistice descriptive și algoritmi de tip Data Mining, utilizate pentru a evidenția tipare relevante în percepțiile clienților bancari și pentru a construi modele predictive privind comportamentele de loialitate. Răspunsurile la chestionar, înregistrate în format digital prin Google Forms, au fost exportate și prelucrate cu ajutorul bibliotecilor Python specializate în analiza de date, precum pandas, seaborn și scikit-learn.

În prima etapă, s-a realizat o analiză descriptivă a variabilelor colectate, vizând calcularea frecvențelor, mediilor, deviației standard și distribuției scorurilor de satisfacție. Aceste statistici au fost utilizate pentru a contura o imagine generală asupra percepțiilor exprimate de respondenți, precum și pentru a compara nivelurile de satisfacție în funcție de gen, vârstă, localitate și modul de utilizare a serviciilor bancare.

În continuare, s-au aplicat tehnici de clasificare binară pentru a anticipa intenția clienților de a recomanda sau nu banca. În acest sens, au fost utilizate regresia logistică, arborii decizionali și algoritmul Support Vector Machine (SVM), fiecare fiind antrenat pe un subset de variabile relevante din chestionar, precum scorurile la dimensiunile serviciului, profilul demografic și tipul de interacțiune cu banca. Modelele au fost comparate în funcție de acuratețea predicției și capacitatea de interpretare a rezultatelor.

Pentru a înțelege mai bine structura răspunsurilor și a segmenta populația investigată, s-a utilizat tehnica de clustering K-means, care a permis gruparea respondenților în funcție de similitudinea răspunsurilor oferite. Această abordare a evidențiat existența unor grupuri omogene de clienți, diferențiate prin nivelul de satisfacție, comportamentul digital sau percepția asupra interacțiunii cu personalul bancar.

Prin combinarea acestor metode, cercetarea a urmărit nu doar identificarea unor relații evidente între variabile, ci și descoperirea unor conexiuni mai subtile care pot contribui la o

înțelegere aprofundată a experienței clientului bancar. Rezultatele obținute oferă o bază solidă pentru formularea unor concluzii relevante și pentru fundamentarea unor recomandări strategice.

Capitolul 4. Analiza și interpretarea datelor din feedback-ul clienților

4.1. Prezentarea și analiza datelor colectate

Înainte de aplicarea metodelor de analiză avansată, este esențială o înțelegere clară a datelor care stau la baza cercetării. În această secțiune sunt prezentate informațiile colectate prin intermediul chestionarului, alături de o analiză descriptivă detaliată a variabilelor demografice și a scorurilor de satisfacție. Sunt evidențiate distribuțiile inițiale, tendințele generale și diferențele între grupuri de respondenți, oferind astfel o perspectivă structurată asupra percepțiilor clienților privind serviciile bancare. Această etapă pregătește terenul pentru interpretările statistice și modelele predictive din capitolele următoare.

4.1.1. Metodologia colectării datelor

Datele analizate în cadrul acestei cercetări au fost colectate prin intermediul unui chestionar online, proiectat special pentru a evalua percepțiile și nivelul de satisfacție al clienților în raport cu serviciile bancare oferite. Formularul a fost distribuit digital și completat de un număr total de 97 de respondenți, din localități precum Baia Mare, Cluj-Napoca, Sighetu Marmăției, București, Timișoara, Târgu Mureș, dar și din orașe mai mici, precum Vișeu de Sus, Remeți sau Crăciunești. Colectarea a fost realizată astfel încât să nu existe erori sau omisiuni în introducerea datelor: fiecare răspuns a fost validat automat, fără a permite completări parțiale, duplicate sau valori lipsă.

Structura chestionarului a inclus atât variabile demografice, cât și întrebări referitoare la experiența directă a clienților cu instituția bancară. Pe partea demografică, au fost înregistrate informații privind sexul respondentului, grupa de vârstă, localitatea de reședință și utilizarea aplicației mobile bancare. Aceste variabile permit segmentarea populației și analiza diferențiată a percepțiilor, în funcție de profilul socio-demografic. Pe lângă aceste elemente, chestionarul a urmărit și evaluarea percepției generale asupra serviciilor bancare, printr-un set de întrebări centrate pe șase dimensiuni esențiale: interacțiunea cu personalul, claritatea comisioanelor, rapiditatea soluționării cererilor, aplicația mobilă, platforma online și satisfacția generală față de bancă.

Toate aceste dimensiuni au fost măsurate folosind o scală Likert de la 1 la 5, unde 1 indică o experiență foarte slabă, iar 5 o experiență foarte bună. Formularea întrebărilor a fost inspirată din literatura de specialitate privind evaluarea serviciilor bancare, fiind aliniată la

structura modelului SERVQUAL, cu accent pe componenta digitală. După centralizarea răspunsurilor, datele au fost importate și procesate în mediu Python, folosind biblioteci specializate precum pandas, seaborn și scikit-learn, care au permis atât analiza descriptivă, cât și dezvoltarea modelelor ulterioare de clasificare și segmentare.

Figura 1 prezintă distribuțiile scorurilor de satisfacție pentru cele șase dimensiuni investigate. Analiza vizuală a acestor grafice arată în mod clar că percepțiile respondenților sunt în general pozitive, dar variabile în funcție de domeniul evaluat. Unele aspecte sunt apreciate constant la niveluri ridicate, în timp ce altele dezvăluie nemulțumiri sau neînțelegeri recurente.

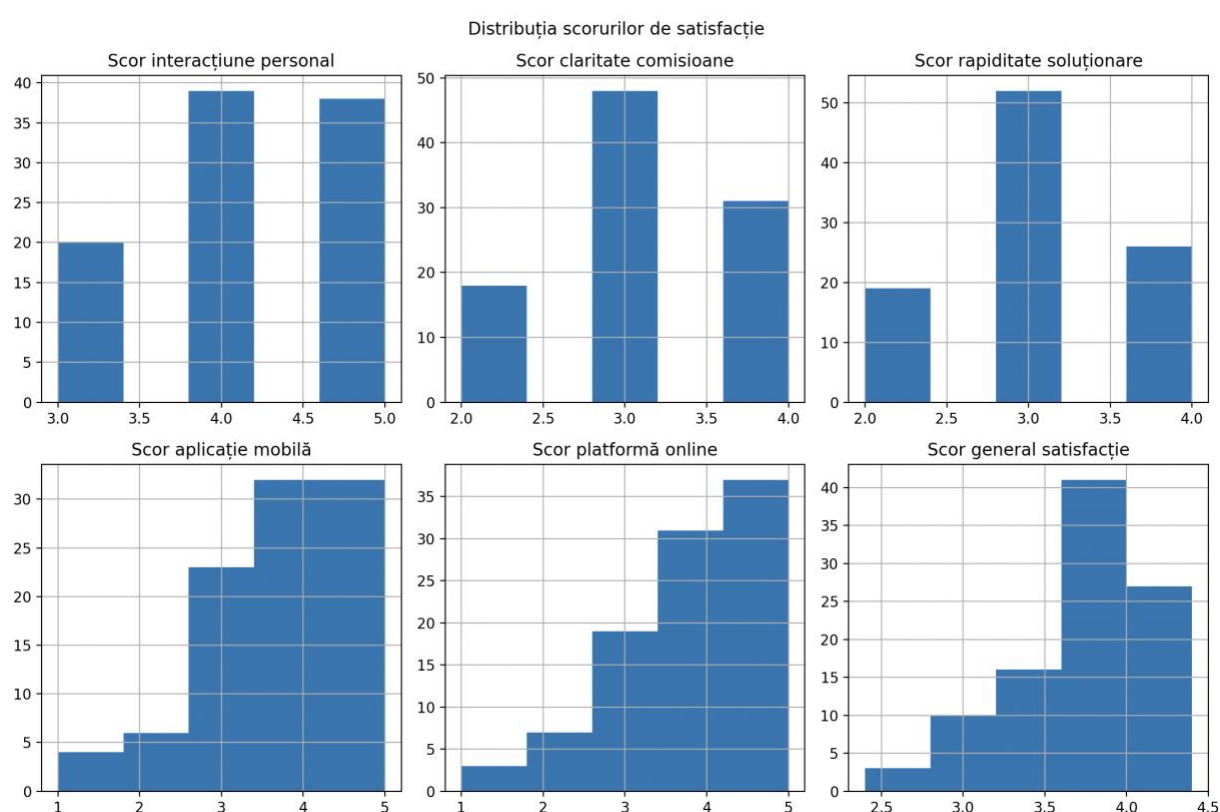


Figura 1. Distribuția scorurilor de satisfacție. Histogramă a frecvenței răspunsurilor pentru fiecare dintre cele șase dimensiuni analizate.

4.1.2. Analiza distribuțiilor și tendințelor generale ale scorurilor

În ceea ce privește interacțiunea cu personalul bancar, distribuția este puternic concentrată în zona superioară a scalei. Cele mai multe scoruri se încadrează între 4 și 5, semnalând o apreciere generală față de comportamentul, disponibilitatea și profesionalismul angajaților. Această percepție pozitivă sugerează că, în ciuda digitalizării accelerate a serviciilor, factorul

uman rămâne esențial în conturarea satisfacției clienților.

Dimensiunea clarității comisioanelor generează, însă, un tipar diferit. Răspunsurile sunt răspândite în jurul valorii de 3, iar un număr semnificativ de respondenți au acordat scoruri de 2. Această distribuție confirmă faptul că percepția asupra comisioanelor bancare este adesea una confuză sau insuficient explicată. Lipsa de transparență sau complexitatea excesivă a informațiilor legate de costuri pot contribui la scăderea nivelului de încredere al clienților, chiar și în cazul în care serviciile în sine sunt performante.

Scorurile referitoare la rapiditatea soluționării cererilor reflectă o experiență moderată, fără extreme notabile, dar cu o tendință ușor negativă față de celelalte dimensiuni. Mulți respondenți au oferit scoruri de 3 sau chiar 2, ceea ce sugerează că durata procesării solicitărilor este percepută adesea ca fiind prea lungă sau lipsită de eficiență. Aceste întârzieri pot avea un impact semnificativ asupra satisfacției generale, mai ales în contextul în care clienții se așteaptă la reacții rapide în mediul digital actual.

Aplicația mobilă și platforma online sunt dimensiuni esențiale în relația modernă client-bancă. Distribuția scorurilor pentru aplicația mobilă arată o ușoară polarizare: există o concentrare vizibilă a scorurilor de 4 și 5, dar și prezența unor răspunsuri scăzute, chiar de 1. Această variație poate reflecta diferențele dintre utilizatorii mai tineri, mai familiarizați cu tehnologia, și cei din categorii de vârstă mai avansate, care întâmpină dificultăți în utilizarea funcționalităților digitale. În schimb, platforma online este apreciată mai uniform: cele mai multe scoruri sunt situate între 4 și 5, indicând un nivel crescut de satisfacție în ceea ce privește accesibilitatea și claritatea interfeței web.

Satisfacția generală reprezintă o sinteză subiectivă a tuturor acestor experiențe și este evaluată preponderent cu scoruri între 3.6 și 4.0. Această înclinare pozitivă a distribuției confirmă faptul că majoritatea respondenților percep serviciile bancare ca fiind bune, dar nu excepționale. În același timp, prezența unor scoruri mai joase pentru anumite dimensiuni indică oportunități clare de îmbunătățire în ariile critice deja identificate.

4.1.3. Diferențieri demografice în percepția satisfacției

Pentru a aprofunda înțelegerea nivelului de satisfacție, am analizat scorurile medii acordate de respondenți în funcție de vârstă. Figura 2 sintetizează aceste rezultate, oferind o imagine comparativă asupra percepțiilor diferitelor grupe de vârstă în raport cu cele șase dimensiuni evaluate.

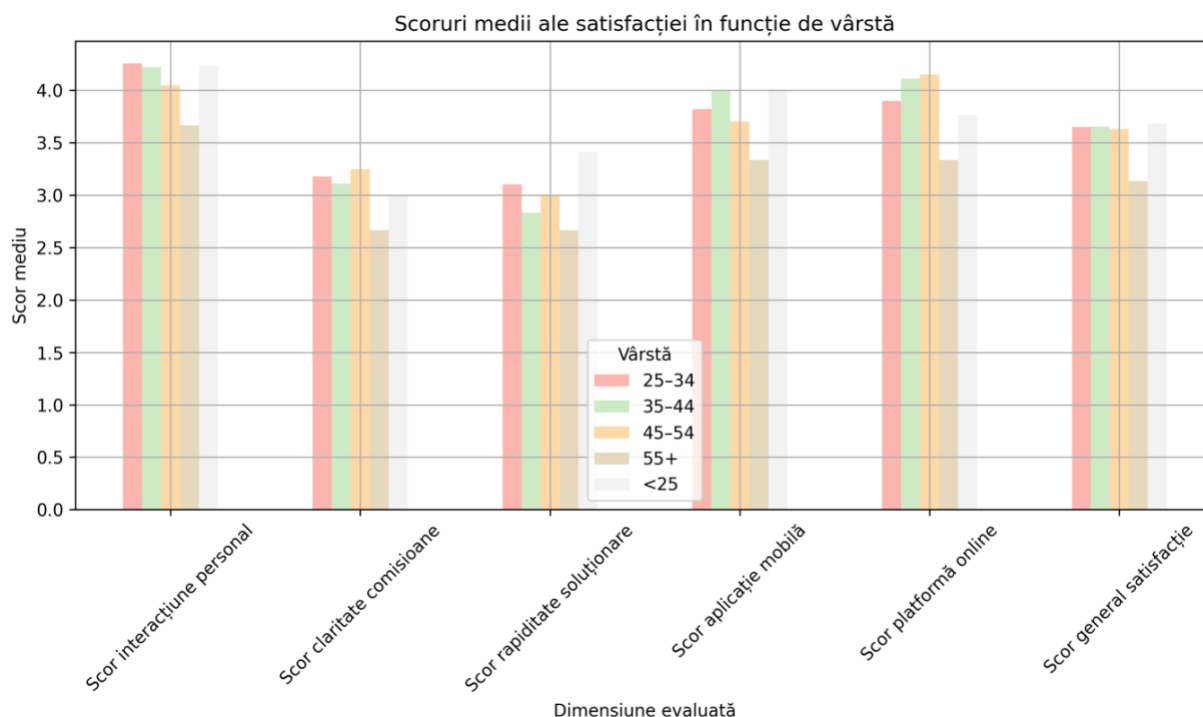


Figura 2. Scoruri medii ale satisfacției în funcție de vârstă. Axa orizontală redă dimensiunile de satisfacție, iar axa verticală indică scorul mediu acordat (1 – foarte slab, 5 – foarte bun).

Graficul evidențiază diferențe moderate, dar relevante, între grupele de vârstă. Tinerii sub 25 de ani se remarcă prin scoruri ridicate acordate aplicației mobile și platformei online, ceea ce reflectă probabil o mai bună familiaritate cu tehnologia și o toleranță mai mare față de eventualele inconveniențe. Totuși, aceștia nu sunt singurii care exprimă o satisfacție crescută în zona digitală – și grupele 25–34 și 35–44 oferă scoruri apropiate, sugerând că utilizatorii activi din mediul profesional apreciază funcționalitatea și accesibilitatea serviciilor online.

În privința interacțiunii cu personalul, toate grupele de vârstă se situează peste media generală, dar se observă o ușoară scădere în rândul respondenților din categoriile 45–54 și 55+. Acest aspect ar putea indica așteptări mai mari din partea acestor clienți în ceea ce privește comunicarea directă și claritatea explicațiilor oferite de angajații băncii. Pe de altă parte, aceste grupe de vârstă oferă scoruri constant mai mici și în ceea ce privește claritatea comisioanelor sau viteza de soluționare, aspecte care pot sugera dificultăți în adaptarea la noile fluxuri operaționale sau o percepție mai critică asupra transparenței bancare.

Per ansamblu, diferențele între grupe nu sunt extreme, dar sunt suficient de clare pentru a recomanda personalizarea comunicării și a serviciilor în funcție de profilul demografic. Tinerii răspund bine la experiențele digitale fluide, în timp ce clienții mai în vârstă par să pună accent mai mare pe claritate, rapiditate și interacțiunea umană.

O perspectivă mai detaliată asupra modului în care percepțiile variază între grupele de vârstă este oferită de Figura 3, care ilustrează distribuția scorurilor acordate aplicației mobile în funcție de vârstă. Utilizarea boxplotului permite observarea nu doar a mediilor, ci și a dispersiei și a valorilor extreme. În mod previzibil, grupurile tinere, în special cele sub 35 de ani, tind să ofere scoruri înalte și uniforme, fără variații notabile. Acest fapt indică o familiaritate solidă cu tehnologia și o experiență în general pozitivă. În contrast, în categoriile 45–54 și 55+, scorurile sunt mult mai răspândite, iar în unele cazuri se înregistrează chiar și valori minime (scor 1). Această variabilitate ridicată sugerează o experiență mai inegală sau dificultăți punctuale în utilizarea aplicației mobile, confirmând nevoia de asistență suplimentară pentru clienții din aceste segmente.

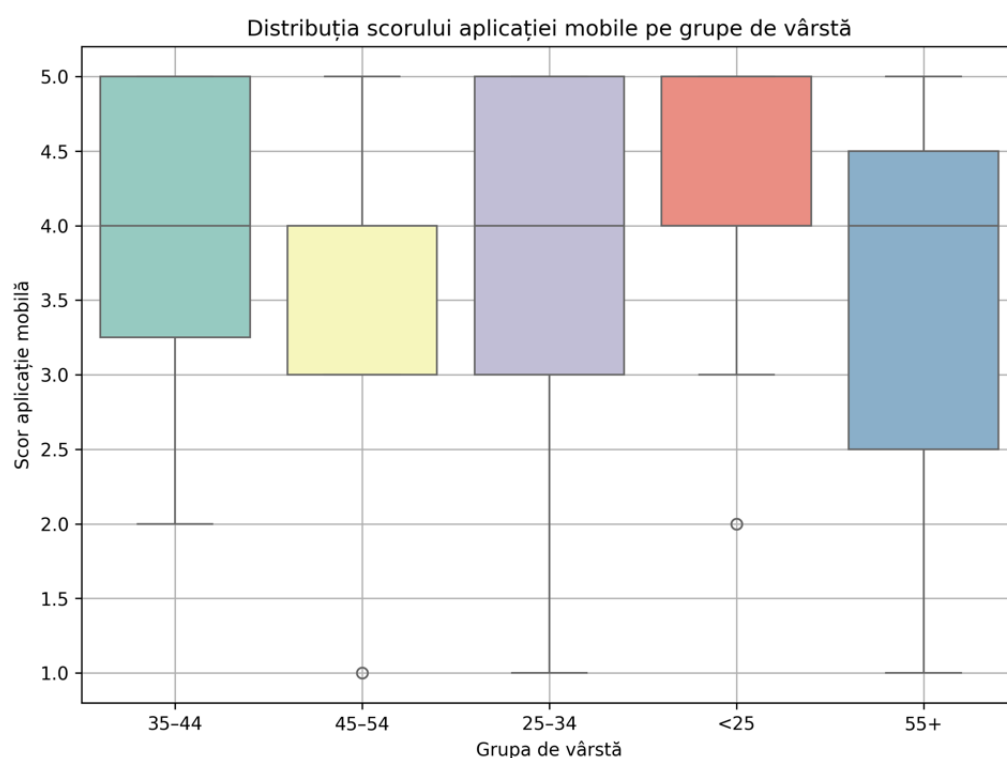


Figura 3. Distribuția scorului aplicației mobile pe grupe de vârstă. Axa verticală reflectă scorul acordat (1 – foarte slab, 5 – foarte bun), iar axa orizontală indică grupele de vârstă. Casetele redau valorile centrale și intervalul intercuartilic, iar punctele individuale semnaleză valorile atipice.

În continuarea analizei demografice, Figura 4 oferă o comparație a scorurilor medii acordate în funcție de gen. Se remarcă o tendință generală prin care femeile acordă scoruri ușor mai mari decât bărbații pentru toate dimensiunile evaluate. Diferențele sunt mai vizibile în cazul aplicației mobile și platformei online, dar sunt prezente și în dimensiunile clasice precum interacțiunea cu personalul și claritatea comisioanelor. Aceste rezultate pot fi interpretate fie

printr-o percepție general mai pozitivă a femeilor față de experiența bancară, fie printr-un nivel de exigență diferit în evaluare. De asemenea, nu este exclus ca stilurile diferite de utilizare a serviciilor bancare să influențeze direct nivelul de satisfacție perceput.

Analiza pe categorii demografice confirmă astfel faptul că percepția asupra serviciilor bancare este puternic influențată de profilul respondentului. Deși scorurile medii sunt în general pozitive, variațiile în funcție de vârstă și gen relevă nevoi distincte. Aceste observații pot ghida segmentarea viitoare a campaniilor de marketing sau dezvoltarea unor funcționalități adaptate, contribuind la o îmbunătățire direcționată a experienței clienților.

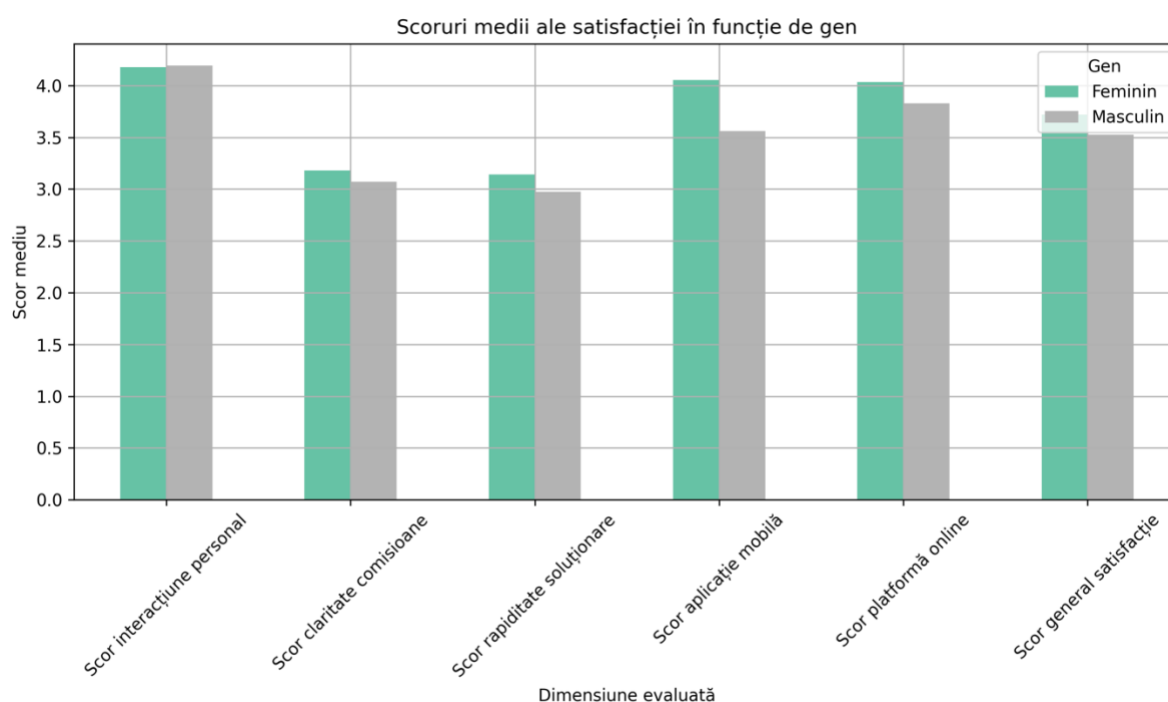


Figura 4. Scoruri medii ale satisfacției în funcție de gen. Axa orizontală indică dimensiunile investigate, iar axa verticală reflectă scorul mediu. Barele sunt diferențiate în funcție de genul respondenților: feminin și masculin.

4.2. Utilizarea Tabelor Pivot pentru analiza feedback-ului

Pentru o înțelegere detaliată a percepțiilor clienților, s-au generat tabele pivot folosind biblioteca pandas din Python. Aceste tabele permit agregarea scorurilor medii în funcție de variabile demografice și comportamentale relevante: genul, intervalul de vârstă, utilizarea aplicației mobile și localitatea de reședință. În final, a fost construit și un tabel de sinteză cu valorile descriptive pentru fiecare dimensiune.

4.2.1. Scoruri în funcție de gen

Tabelul 1 prezintă scorurile medii pe cele șase dimensiuni analizate, în funcție de gen. Femeile oferă scoruri mai mari comparativ cu bărbații pentru majoritatea aspectelor, în special pentru aplicația mobilă (4.05 vs. 3.56), platforma online (4.04 vs. 3.83) și satisfacția generală (3.72 vs. 3.53). Aceste diferențe sugerează o deschidere mai mare în rândul femeilor față de utilizarea tehnologiei bancare și, posibil, o experiență mai pozitivă în interacțiunile directe.

Sex	Scor interacțiune personal	Scor claritate comisioane	Scor rapiditate soluționare	Scor aplicație mobilă	Scor platformă online	Scor general satisfacție
Feminin	4,18	3,18	3,14	4,05	4,04	3,72
Masculin	4,2	3,07	2,98	3,56	3,83	3,53

Tabelul 1: Scoruri medii ale satisfacției în funcție de gen. Valorile sunt exprimate pe o scală de la 1 (foarte slab) la 5 (foarte bun).

4.2.2. Scoruri în funcție de vârstă

În Tabelul 2 sunt prezentate scorurile medii în funcție de intervalele de vârstă. Respondenții sub 25 de ani oferă cele mai mari scoruri pentru aplicația mobilă (4.00) și platforma online (4.18), sugerând o familiaritate crescută cu mediul digital. În schimb, grupurile de vârstă 45–54 și 55+ oferă scoruri mai mici pentru claritatea comisioanelor, aplicația mobilă și viteza de soluționare. Această tendință poate reflecta așteptări mai conservatoare sau dificultăți în adaptarea la noile tehnologii.

Categorie vârstă	Scor interacțiune personal	Scor claritate comisioane	Scor rapiditate soluționare	Scor aplicație mobilă	Scor platformă online	Scor general satisfacție
25–34	4,26	3,18	3,1	3,82	3,9	3,65
35–44	4,22	3,11	2,83	4	4,11	3,66
45–54	4,05	3,25	3	3,7	4,15	3,63
55+	3,67	2,67	2,67	3,33	3,33	3,13
<25	4,24	3	3,41	4	3,76	3,68

Tabelul 2: Scoruri medii ale satisfacției în funcție de vârstă. Compararea scorurilor oferite de fiecare grupă de vârstă pe cele șase dimensiuni.

4.2.3. Scoruri în funcție de utilizarea aplicației mobile

Tabelul 3 evidențiază diferențele semnificative între utilizatorii aplicației mobile și cei care nu o folosesc. Utilizatorii activi ai aplicației acordă scoruri net superioare pentru majoritatea

dimensiunilor: aplicație mobilă (4.16 vs. 2.00), platforma online (4.27 vs. 2.07) și satisfacție generală (3.76 vs. 2.90). În mod interesant, non-utilizatorii apreciază în schimb interacțiunea directă cu personalul (4.36), ceea ce ar putea indica o încredere mai mare în componenta umană a serviciilor.

Folosește aplicația	Scor interacțiune personal	Scor claritate comisioane	Scor rapiditate soluționare	Scor aplicație mobilă	Scor platformă online	Scor general satisfacție
Da	4,16	3,13	3,1	4,16	4,27	3,76
Nu	4,36	3,14	2,93	2	2,07	2,9

Tabelul 3: Scoruri medii în funcție de utilizarea aplicației mobile. Comparatie între utilizatori și non-utilizatori ai aplicației bancare.

4.2.4. Scoruri în funcție de localitate

Tabelul 4 grupează scorurile în funcție de localitatea respondenților. Cele mai înalte niveluri de satisfacție generală apar în Remeți (4.00), Cluj-Napoca (3.71), Baia Mare (3.68) și București (3.68), ceea ce poate reflecta o calitate superioară a serviciilor sau un grad mai mare de digitalizare. Localități mai mici precum Vișeu de Sus (3.00) și Târgu Mureș (3.47) înregistrează scoruri mai modeste.

Localitate	Scor interacțiune personal	Scor claritate comisioane	Scor rapiditate soluționare	Scor aplicație mobilă	Scor platformă online	Scor general satisfacție
Baia Mare	4,29	3,18	2,94	3,88	4,12	3,68
București	4,13	3,07	3,2	4,07	3,93	3,68
Cluj-Napoca	4,09	3,09	3,27	3,86	4,23	3,71
Crăciunești	4,67	3,67	3	3	4	3,67
Remeți	4,5	3,5	3,5	4,5	4	4
Sighetu Marmației	4	3,1	3,2	3,8	4,1	3,64
Timișoara	4,11	3,11	2,83	3,78	3,89	3,54
Târgu Mureș	4,44	3,11	3	3,78	3	3,47
Vișeu de Sus	4	3	2	3	3	3

Tabelul 4: Scoruri medii ale satisfacției în funcție de localitate. Rezultatele sunt exprimate ca medii pe fiecare localitate, pe cele șase dimensiuni analizate.

4.2.5. Statistici descriptive generale

Tabelul 5 prezintă valorile medii, minime și maxime pentru fiecare dimensiune. Interacțiunea cu personalul (medie 4.19) și platforma online (3.95) sunt cele mai bine cotate. În schimb,

claritatea comisioanelor (3.13) și viteza de soluționare (3.07) sunt mai slab evaluate. Aplicația mobilă prezintă o variație amplă (1.00 – 5.00), indicând experiențe foarte diferite între utilizatori.

	mediu	minim	maxim
Scor interacțiune personal	4,19	3	5
Scor claritate comisioane	3,13	2	4
Scor rapiditate soluționare	3,07	2	4
Scor aplicație mobilă	3,85	1	5
Scor platformă online	3,95	1	5
Scor general satisfacție	3,64	2,4	4,4

Tabelul 5: Valori descriptive pentru fiecare dimensiune de satisfacție (medie, minim, maxim).
Tabel de sinteză care indică distribuția scorurilor pentru toate cele șase dimensiuni.

4.2.6. Corelații între dimensiunile satisfacției

După agregarea scorurilor în funcție de variabile demografice, a fost realizată și o analiză suplimentară, centrată pe relațiile statistice dintre dimensiunile de satisfacție. În acest scop, a fost generată o hartă de corelații, care evidențiază legăturile liniare dintre variabilele investigate și modul în care acestea influențează satisfacția generală. Această abordare permite o înțelegere mai profundă a mecanismelor interne ale percepției clienților și oferă direcții concrete pentru optimizarea serviciilor.

Rezultatele evidențiază o asociere puternică între satisfacția generală și dimensiunile legate de experiența digitală. Scorul aplicației mobile ($r = 0.74$) și scorul platformei online ($r = 0.69$) prezintă cele mai ridicate corelații pozitive, ceea ce sugerează că interacțiunea tehnologică a devenit un factor determinant în conturarea percepției de ansamblu asupra serviciilor bancare. Acest tipar reflectă o tendință tot mai accentuată în comportamentul consumatorilor moderni, care acordă o importanță sporită eficienței, accesibilității și confortului oferit de canalele digitale.

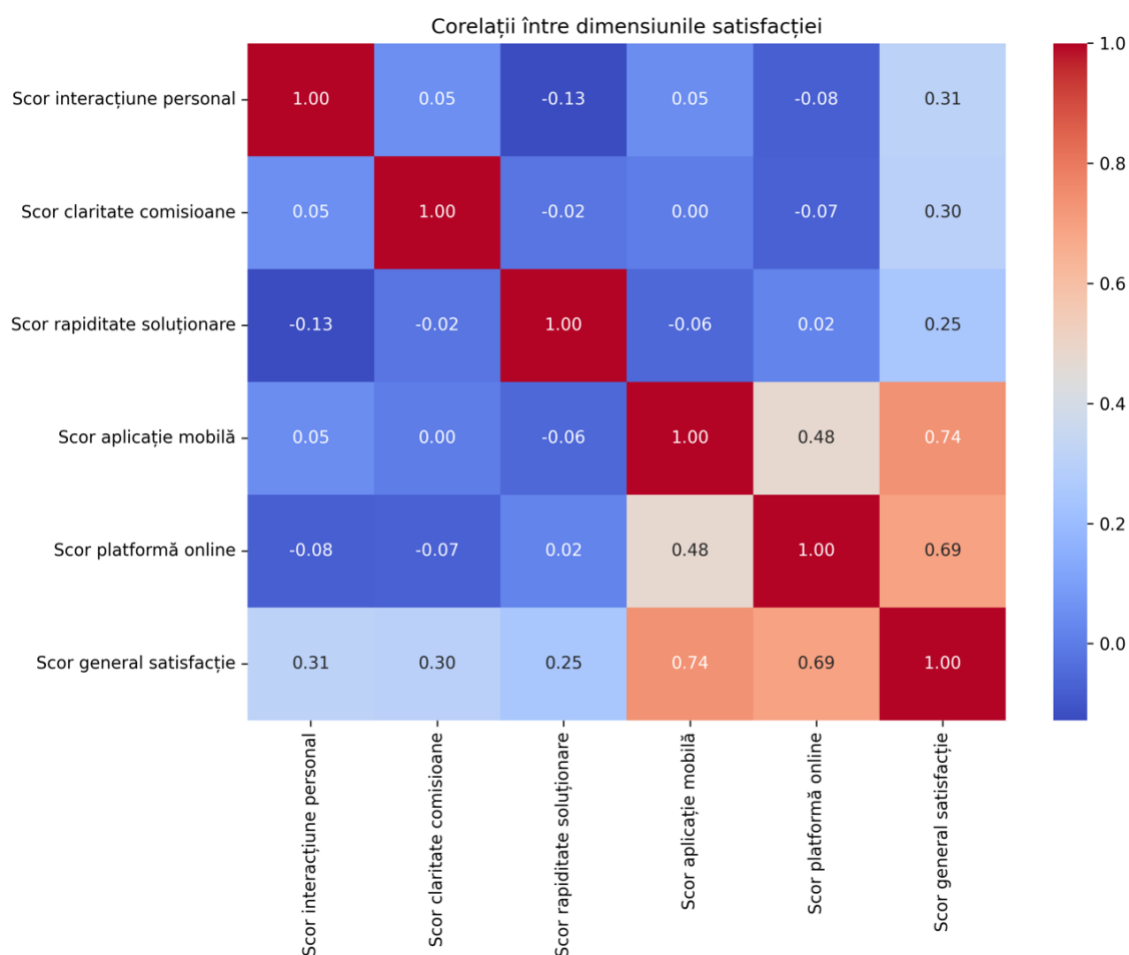


Figura 5. Matrice de corelații între dimensiunile satisfacției. Valorile coeficienților de corelație sunt exprimate numeric și vizual, printr-o scară cromatică gradată de la -1 (corelație negativă puternică) la +1 (corelație pozitivă puternică). Axele graficului corespund celor șase dimensiuni analizate: interacțiunea cu personalul, claritatea comisioanelor, rapiditatea soluționării, aplicația mobilă, platforma online și satisfacția generală.

În contrast, dimensiunile clasice precum claritatea comisioanelor ($r = 0.25$), interacțiunea cu personalul ($r = 0.31$) și rapiditatea soluționării ($r = 0.27$) prezintă corelații mai slabe cu satisfacția generală. Acest lucru nu le reduce relevanța, dar indică faptul că, în percepția respondenților, aceste aspecte influențează mai puțin direct evaluarea globală. Totodată, corelațiile slabe dintre aceste dimensiuni între ele susțin ideea că fiecare dintre ele acoperă o zonă distinctă a experienței bancare și că evaluările sunt formulate independent, în funcție de așteptările specifice ale fiecărui client.

Aceste rezultate confirmă faptul că transformarea digitală nu este doar o opțiune strategică, ci un pilon central al satisfacției clientului modern. Totodată, corelațiile relativ scăzute dintre dimensiunile clasice evidențiază faptul că fiecare dintre acestea captează aspecte distincte ale experienței și necesită abordări diferențiate în strategiile de îmbunătățire.

4.3. Analiza datelor utilizând instrumentele de Data Mining

După etapa exploratorie și descriptivă a datelor, în care au fost identificate tendințe generale și diferențe între grupuri demografice, analiza avansează spre un nivel superior de complexitate. În această secțiune sunt aplicate metode de tip Data Mining, cu scopul de a descoperi tipare ascunse în date și de a construi modele predictive care să anticipeze comportamentele și preferințele clienților. Etapele detaliate includ pregătirea riguroasă a datelor pentru modelare, alegerea și antrenarea algoritmilor, evaluarea performanței acestora, precum și segmentarea respondenților pe baza caracteristicilor comune. Întregul proces a fost realizat în Python, folosind biblioteci specializate în învățare automată și analiză statistică, într-un mod care îmbină rigoarea tehnică cu aplicabilitatea practică a rezultatelor.

4.3.1. Alegerea limbajului și justificarea metodologică

Pentru realizarea analizei de tip Data Mining în această cercetare, s-a optat pentru utilizarea limbajului de programare Python, datorită popularității sale în rândul analiștilor de date, dar și a capacității sale dovedite de a gestiona întregul flux de prelucrare, analiză și modelare a datelor. Alegerea acestui instrument este justificată atât prin flexibilitatea sa, cât și prin gradul ridicat de adoptare în mediul academic și profesional.

Python este un limbaj open-source, ceea ce îl face accesibil oricărui utilizator, fără constrângeri legate de licențiere. Această deschidere a contribuit la dezvoltarea unei comunități largi și active, care susține constant îmbunătățirea pachetelor software existente. În același timp, este un limbaj care oferă suport pentru un număr impresionant de biblioteci specializate în analiza datelor, fapt ce a permis implementarea tuturor etapelor de analiză — de la curățarea datelor până la modelare - într-un mediu unic, coerent și reproductibil.

În această cercetare, biblioteci precum pandas și numpy au fost utilizate pentru manipularea tabelară a datelor și procesarea numerică. pandas permite citirea și transformarea ușoară a fișierelor .csv, operând asupra dataseturilor sub formă de dataframe-uri, în timp ce numpy susține eficiența calculului numeric. Pentru realizarea vizualizărilor statistice s-au folosit seaborn și matplotlib, care oferă instrumente avansate pentru reprezentarea grafică a corelațiilor, distribuțiilor și valorilor agregate. În etapa de modelare, biblioteca scikit-learn a reprezentat componenta centrală, permițând implementarea unor modele precum regresia logistică, Support Vector Machines (SVM), arbori decizionali și algoritmi de clustering de tip K-means.

Alegerea Python nu s-a bazat exclusiv pe avantajele sale tehnice, ci și pe relevanța

profesională. În mediul bancar și financiar, analiza datelor devine o competență esențială, iar Python este din ce în ce mai utilizat pentru dezvoltarea sistemelor de scoring, evaluarea riscurilor, detectarea fraudelor și automatizarea proceselor analitice. Astfel, integrarea acestui limbaj în lucrarea de față reflectă o aliniere cu cerințele reale ale pieței muncii, dar și cu tendințele din sfera cercetării aplicate.

4.3.2. Prelucrarea și transformarea datelor

Procesul de prelucrare a datelor a fost realizat integral în Python, folosind biblioteci specializate precum pandas pentru manipularea dataseturilor, scikit-learn pentru modelare și transformări, și numpy pentru operațiuni numerice. Aceste instrumente oferă un cadru eficient, flexibil și scalabil pentru etapele esențiale de transformare și pregătire a datelor în vederea aplicării algoritmilor de învățare automată.

Pentru a putea fi utilizate eficient în algoritmi de clasificare sau clustering, datele colectate au fost convertite într-un format numeric standardizat. Variabilele de tip text, precum Sex, Vârstă și Folosește aplicația, au fost transformate în valori numerice prin tehnici de codificare, asigurând compatibilitatea cu modelele matematice. Astfel, genul a fost codificat binar (de exemplu, Feminin = 0, Masculin = 1), iar categoriile de vârstă au fost transpuse într-o scală ordinală, reflectând progresia naturală a intervalelor.

De asemenea, variabila țintă Recomandă banca, exprimată inițial sub formă textuală (Da/Nu), a fost convertită într-un format binar (1/0), pentru a permite aplicarea unor modele de clasificare precum regresia logistică și SVM. Această reformulare a fost esențială în definirea unei probleme clare de învățare automată cu ieșire binară.

În pasul următor, au fost selectate caracteristicile relevante pentru modelare. Au fost incluse atât scorurile acordate pe dimensiunile de satisfacție (interacțiune cu personalul, claritatea comisioanelor, rapiditatea soluționării, aplicația mobilă și platforma online), cât și variabilele demografice și comportamentale (sex, vârstă, utilizarea aplicației). Această selecție combină factori subiectivi și obiectivi, oferind un set robust de predictoare pentru analiză.

În vederea normalizării valorilor și evitării dezechilibrelor între variabile cu scări diferite, s-a aplicat standardizarea Z-score, prin care toate variabilele au fost transformate astfel încât să aibă media 0 și deviația standard 1. Acest pas este crucial pentru algoritmi care sunt sensibili la amplitudinea valorilor de intrare, cum ar fi regresia logistică sau SVM.

În final, datasetul a fost împărțit în două subseturi: set de antrenare (80%) și set de testare (20%), păstrând proporția distribuției claselor în variabila țintă. Această separare a permis validarea riguroasă a performanței modelelor dezvoltate, reducând riscul de supraînvățare și

oferind o estimare realistă a generalizabilității rezultatelor.

4.3.3. Construirea și evaluarea modelelor de clasificare

Pentru a răspunde obiectivului cercetării – acela de a identifica factorii care influențează loialitatea față de bancă, măsurată prin intenția de recomandare – au fost aplicate modele de clasificare binară, cu scopul de a estima probabilitatea ca un respondent să recomande banca (Recomandă banca = 1). Alegerea acestor modele s-a bazat pe două criterii fundamentale: capacitatea de a oferi performanțe predictive ridicate și posibilitatea de a interpreta rezultatele într-un mod accesibil pentru factorii decizionali din organizații.

În literatura de specialitate, modelele de clasificare precum regresia logistică, Support Vector Machine (SVM) și arborii decizionali sunt frecvent utilizate în analiza satisfacției clienților și în cercetarea comportamentului consumatorilor (Breiman, 1984; Quinlan, 1993; Hastie et al., 2009). Aceste metode sunt potrivite în contexte cu dimensiuni moderate ale seturilor de date și cu variabile mixte (categorice și numerice), oferind un echilibru între acuratețe și transparență.

Regresia logistică a fost selectată ca model de bază datorită interpretabilității sale. Acest model permite estimarea probabilității ca un client să recomande banca, în funcție de un set de variabile independente, prin aplicarea funcției logistice asupra unei combinații liniare de predictoare. Coeficienții estimați pot fi interpretați ca efecte marginale asupra log-odds-ului, oferind informații directe despre direcția și forța relațiilor dintre variabile (Hosmer et al., 2013). Regresia logistică este deosebit de utilă în contexte unde comunicarea rezultatelor către factori non-tehnici este importantă, ceea ce o face relevantă pentru instituții financiare.

Modelul antrenat pe datele curățate și standardizate a atins o acuratețe de 95% pe setul de testare. Evaluarea prin metricile precision și recall a evidențiat o performanță ridicată în special pentru clasa pozitivă, ceea ce sugerează că modelul identifică eficient clienții cu grad ridicat de satisfacție.

Pe lângă regresia logistică, au fost aplicate și alte două modele de clasificare binară: Support Vector Machine (SVM) și arborele decizional, ambele utilizate frecvent în analiza comportamentului consumatorilor și în cercetarea satisfacției clienților.

Modelul Support Vector Machine cu kernel liniar a fost implementat pentru capacitatea sa de a gestiona probleme de clasificare în spații de dimensiuni moderate, în care clasele sunt relativ bine separate. Algoritmul funcționează prin construirea unui hiperplan care maximizează distanța (marja) dintre cele două clase – în acest caz, respondenții care recomandă sau nu banca. Prin această abordare, SVM reduce riscul de eroare structurală și este mai puțin

sensibil la distribuții dezechilibrate sau la existența unor date extreme (Cortes & Vapnik, 1995).

Modelul a fost antrenat folosind aceleași variabile standardizate ca și în cazul regresiei logistice și a generat rezultate comparabile ca performanță. Evaluarea pe setul de testare a indicat o acuratețe de 95%, iar valorile metricilor de tip precision și recall au fost echilibrate între clase, sugerând o capacitate bună de generalizare pe noi observații. Scorul F1 obținut a confirmat că modelul reușește să identifice eficient observațiile din ambele clase, fără a favoriza clasa majoritară.

Modelul de arbore decizional a fost aplicat pentru avantajele sale explicative, fiind capabil să genereze o structură de decizie ierarhică ce poate fi urmărită intuitiv. Acest model segmentează datele prin reguli binare repetitive, construind un arbore în care fiecare nod reprezintă o condiție pe o variabilă, iar frunzele indică predicția finală. În contextul cercetării, arborele decizional a fost util pentru a observa cum diferite combinații de scoruri și trăsături demografice conduc la probabilități diferite de recomandare.

Rezultatul obținut în urma rulării pe datele test a fost o acuratețe de 90%, puțin sub cea a modelelor anterioare, însă suficientă pentru a susține validitatea modelului. Analiza importanței variabilelor a arătat că Scorul aplicației mobile, interacțiunea cu personalul și utilizarea aplicației au fost cele mai relevante caracteristici în ramificarea arborelui. Acest lucru sugerează o asociere clară între experiența digitală și intenția de recomandare, observabilă în modul în care arborele generează deciziile.

În ambele cazuri, modelele au fost aplicate utilizând biblioteca scikit-learn, iar performanțele au fost măsurate în mod standardizat folosind setul de testare și metrici specifice problemelor de clasificare binară.

Pentru a determina care dintre cele trei algoritmi de clasificare oferă cele mai bune rezultate în estimarea intenției de recomandare, au fost utilizate mai multe metrici de evaluare: acuratețea, precizia, recall-ul, scorul F1 și matricea de confuzie. Acestea sunt standard în literatura de specialitate pentru analiza performanței modelelor binare (Han et al., 2011; Provost & Fawcett, 2013) și permit o evaluare nuanțată, dincolo de simpla acuratețe.

Atât regresia logistică, cât și SVM au obținut o acuratețe de 95%, indicând o capacitate ridicată de generalizare pe datele testate. Totuși, modelul SVM a prezentat un scor F1 mai echilibrat între clase, ceea ce îl face mai robust în prezența unor distribuții ușor dezechilibrate – un aspect menționat și de Cristianini & Shawe-Taylor (2000), care subliniază avantajul SVM în separarea optimă a claselor în spații de date complexe.

Arborele decizional a înregistrat o acuratețe de 90%, mai scăzută față de celelalte două modele, dar a oferit claritate în interpretare, prin structura sa ierarhică. Acest tip de model este

apreciat în mediile decizionale tocmai pentru capacitatea de a explica logicile de clasificare într-un mod accesibil (Quinlan, 1993; Breiman et al., 1984), chiar dacă performanțele brute pot fi ușor mai scăzute.

Analiza matricilor de confuzie a indicat că toate cele trei modele identifică eficient clienții care intenționează să recomande banca (clasa pozitivă), dar diferențele apar în tratarea falselor negative, unde regresia logistică a fost mai precisă, iar SVM a oferit o acoperire mai echilibrată. Modelul bazat pe arbori a avut tendința de a supraestima clasa pozitivă, fapt care poate fi acceptabil sau nu, în funcție de obiectivul strategic al băncii – de exemplu, dacă scopul este maximizarea identificării clienților fideli, această caracteristică poate fi chiar avantajoasă.

4.3.4 Segmentarea respondenților prin metode de clustering

După aplicarea modelelor de clasificare, care au presupus existența unei variabile țintă (Recomandă banca), analiza continuă cu o abordare nesupravegheată, vizând identificarea unor structuri latente în setul de date. Obiectivul acestei etape este segmentarea respondenților în grupuri omogene din punct de vedere al profilului de satisfacție și al caracteristicilor demografice sau comportamentale, fără a presupune a priori existența unor clase prestabilite.

În acest scop, a fost utilizată metoda K-means clustering, una dintre cele mai populare tehnici de segmentare automată. Alegerea acestui algoritm se bazează pe eficiența sa computațională și pe capacitatea de a produce clustere distincte pe baza distanței euclidiene dintre observații (MacQueen, 1967). K-means împarte datele în K grupuri, atribuind fiecare observație clusterului cu centrul cel mai apropiat, iar apoi recalculând centrul până la stabilizare.

Pentru aplicarea metodei, a fost selectat un subset de variabile care combină scorurile de satisfacție acordate diferitelor dimensiuni (personal, aplicație mobilă, comisioane etc.), precum și informații demografice esențiale precum vârsta și utilizarea aplicației. Înainte de rularea algoritmului, toate valorile au fost standardizate pentru a evita influențarea segmentării de variabile cu scale diferite – o etapă recomandată în literatura de specialitate (Tan et al., 2019).

Determinarea numărului optim de clustere este un pas esențial în aplicarea algoritmului K-means, deoarece o alegere inadecvată poate conduce fie la grupări prea generale, fie la suprasegmentare inutilă. În literatura de specialitate, sunt recomandate diverse metode de estimare, dintre care cele mai utilizate sunt metoda „elbow” și indicele silhouette (Kodinariya & Makwana, 2013). În această cercetare, s-a optat pentru metoda elbow, care presupune rularea algoritmului pentru mai multe valori ale parametrului K și observarea punctului de inflexiune în graficul inerției totale. Punctul în care reducerea inerției devine nesemnificativă indică valoarea optimă de K.

Aplicând această metodă asupra datasetului curățat și standardizat, s-a observat că valoarea $K=3$ oferă un echilibru optim între coeziunea internă a grupurilor și separabilitatea lor. Modelul a fost apoi antrenat pentru $K=3$ și rezultatele au fost interpretate pe baza mediilor variabilelor pentru fiecare cluster. Această analiză a permis descrierea unor profiluri distincte de respondenți, fiecare caracterizat printr-un set specific de trăsături demografice și comportamentale.

Astfel, unul dintre grupuri include respondenți tineri, foarte activi digital, care utilizează aplicația mobilă și oferă scoruri ridicate pentru serviciile băncii. Un al doilea cluster este format din persoane de vârstă medie, cu o utilizare mai moderată a serviciilor digitale și un nivel mediu de satisfacție. Cel de-al treilea cluster grupează respondenții critici, care oferă scoruri scăzute și nu folosesc aplicația digitală, fiind cel mai probabil să exprime neîncredere sau nemulțumire față de instituția bancară. Aceste rezultate oferă o bază valoroasă pentru strategii diferențiate de retenție și personalizare a comunicării, în funcție de caracteristicile fiecărui segment.

Rezultatele obținute în urma segmentării clienților prin clustering confirmă existența unor tipare clare de satisfacție și comportament, care nu erau vizibile în analiza descriptivă. Fiecare cluster rezultă surprinde un anumit tip de relație între clienți și serviciile băncii, reliefând rolul decisiv al interacțiunii digitale și al percepției asupra calității serviciilor în formarea loialității. De exemplu, grupul utilizatorilor aplicației mobile, care au oferit scoruri ridicate, se distinge net printr-o predispoziție mai mare de a recomanda banca, ceea ce întărește ipoteza conform căreia experiența digitală este un predictor semnificativ al satisfacției generale (Malaquias & Hwang, 2019).

Această segmentare contribuie și la o mai bună înțelegere a relației dintre variabilele socio-demografice și percepția serviciilor bancare. Deși sexul sau vârsta nu explică singure gradul de loialitate, ele capătă relevanță în combinație cu alți factori, precum frecvența de utilizare a aplicației sau evaluarea interacțiunii cu personalul. Clustering-ul permite astfel identificarea unor configurații complexe, greu de surprins prin metode clasice de analiză, și oferă o bază solidă pentru viitoare cercetări predictive sau intervenții personalizate.

Prin utilizarea metodei K-means și a tehnicilor de interpretare vizuală a centrelor de cluster, cercetarea a adus un plus de profunzime în înțelegerea diferențelor dintre respondenți. Această abordare evidențiază nu doar nivelul de satisfacție general, ci și cauzele potențiale ale diferențelor dintre segmente. Astfel, segmentarea nu este doar o etapă exploratorie, ci o componentă esențială în formularea unor strategii personalizate de retenție și comunicare, în concordanță cu recomandările din literatura actuală privind analiza centrată pe client (Wedel & Kamakura, 2000).

4.4 Interpretarea rezultatelor și sugestii strategice

Analiza modelelor de clasificare și a segmentării prin clustering relevă o serie de aspecte esențiale în ceea ce privește comportamentul clienților bancari și factorii care le influențează loialitatea. Modelul de regresie logistică, alături de SVM, a evidențiat o asociere puternică între scorurile acordate aplicației mobile și probabilitatea de a recomanda banca. Acest rezultat sugerează că satisfacția față de interfața digitală și ușurința în utilizare joacă un rol central în formarea unei opinii favorabile despre instituția bancară.

Mai departe, importanța scorului privind interacțiunea cu personalul – evidențiată în mod constant de toate modelele – indică faptul că relația umană rămâne un pilon al experienței bancare. Această observație susține ideea complementarității dintre digitalizare și contactul direct, confirmând tendințele din literatura de specialitate privind necesitatea unei abordări hibride în furnizarea serviciilor financiare (Verhoef et al., 2009).

Rezultatele obținute sugerează, de asemenea, că variabilele demografice (precum sexul și vârsta) nu sunt predictorii direcți ai loialității, dar pot interacționa semnificativ cu alți factori comportamentali. În special, utilizarea aplicației pare să fie un factor moderat de importanță în funcție de grupa de vârstă, aspect relevant în analiza clusterelor. Această constatare deschide oportunitatea unei segmentări strategice mai fine, care să țină cont de interacțiuni între variabile, nu doar de influența lor individuală.

Analiza clusteringului prin algoritmul K-means a permis conturarea unor segmente distincte de clienți, diferențiate în principal prin scorurile acordate dimensiunilor serviciilor bancare și prin comportamentul digital. Unul dintre cluster este format din utilizatori activi ai aplicației mobile, cu vârste medii, care au oferit scoruri ridicate tuturor dimensiunilor de satisfacție și au o probabilitate mare de a recomanda banca. Acest grup reprezintă nucleul loial al bazei de clienți și poate constitui o țintă ideală pentru strategii de fidelizare.

Un alt cluster identificat este format din clienți care nu folosesc aplicația mobilă și au oferit scoruri scăzute în special pentru componenta de claritate a comisioanelor și timpul de soluționare a cererilor. Acest segment este potențial vulnerabil la migrarea către competitori și ar putea beneficia de intervenții direcționate pentru îmbunătățirea percepției – fie prin campanii de comunicare mai clare, fie prin optimizarea proceselor interne. Literatura privind churn-ul bancar subliniază frecvent importanța percepției de transparență și eficiență în decizia de menținere a relației cu o instituție financiară (Keaveney, 1995).

De asemenea, diferențele dintre cluster confirmă faptul că nu există o singură rețetă de

loializare eficientă pentru toți clienții. Segmentarea bazată pe combinații de scoruri și comportamente oferă un cadru mai realist pentru dezvoltarea strategiilor de retenție, comparativ cu abordările unidimensionale sau exclusiv demografice. Această perspectivă este susținută și de studiile recente care pledează pentru o abordare centrată pe profiluri comportamentale în designul serviciilor (Ng & Forbes, 2009).

Interpretarea corelațiilor dintre dimensiunile de satisfacție a relevat că scorul acordat aplicației mobile are legături puternice cu percepțiile generale asupra serviciilor bancare, în special în ceea ce privește claritatea informațiilor și interacțiunea cu personalul. Acest rezultat sugerează că experiența digitală nu funcționează izolat, ci influențează percepția asupra întregii relații client–bancă. Astfel, investițiile în îmbunătățirea funcționalității aplicației și în integrarea acesteia cu alte canale de comunicare pot avea un efect multiplicator asupra satisfacției generale – o direcție susținută și de literatura privind serviciile omnicanal (Verhoef et al., 2015).

Vizualizările comparative ale scorurilor în funcție de sex și vârstă au indicat diferențe modeste, dar semnificative, în preferințele utilizatorilor. De exemplu, tinerii au acordat scoruri mai mari aplicației mobile, în timp ce respondenții mai în vârstă au pus accent pe interacțiunea directă cu personalul. Acest tip de rezultate sprijină ideea unei personalizări a comunicării și a designului serviciilor bancare, în funcție de profilul demografic și comportamental al clientului – un principiu fundamentat și de modelul SERVQUAL, care susține adaptarea ofertei în funcție de așteptările clientului (Parasuraman et al., 1988).

În general, analiza integrată a modelelor de clasificare, a segmentării și a corelațiilor permite o înțelegere detaliată a factorilor care contribuie la loialitatea față de bancă. Mai departe, aceste rezultate pot fundamenta acțiuni tactice – de la ajustarea designului aplicației, până la implementarea unor campanii de retenție direcționate – dar recomandările finale vor fi formulate în concluziile generale ale lucrării.

Capitolul 5. Concluzii, limitări și direcții viitoare

5.1. Sinteza obiectivelor și a abordării metodologice

Lucrarea de față a avut ca obiectiv central identificarea principalilor factori care determină satisfacția și loialitatea clienților față de serviciile bancare, cu accent pe intenția acestora de a recomanda instituția financiară. În contextul digitalizării accelerate a industriei bancare, înțelegerea acestor factori devine esențială pentru formularea unor strategii eficiente de retenție și diferențiere competitivă.

Studiul s-a bazat pe un set de date colectate prin intermediul unui chestionar aplicat online, care a vizat atât dimensiuni subiective ale satisfacției (ex. percepția asupra interacțiunii cu personalul sau claritatea comisioanelor), cât și aspecte legate de experiența digitală (utilizarea aplicației mobile și a platformei online). În urma procesului de prelucrare, datele au fost analizate în trei etape: descriptivă, predictivă și de segmentare.

Analiza descriptivă a evidențiat tendințele generale din eșantion și diferențele dintre grupurile demografice. Ulterior, s-au aplicat trei modele de clasificare binară – regresia logistică, Support Vector Machine (SVM) și arborele decizional – pentru a estima probabilitatea ca un respondent să recomande banca. Rezultatele predictive obținute au fost validate prin metrice standard (precizie, scor F1, recall) și au evidențiat un grad ridicat de acuratețe, în special pentru regresia logistică și SVM.

În final, tehnica de clustering K-means a fost utilizată pentru a grupa respondenții în funcție de trăsături comune, conturând astfel profile distincte de clienți. Această abordare a permis nu doar analiza nivelului de satisfacție, ci și înțelegerea unor patternuri comportamentale relevante pentru personalizarea intervențiilor.

5.2 Interpretarea rezultatelor și implicații practice

Analiza datelor a confirmat că satisfacția și loialitatea clienților bancari sunt profund influențate de experiența digitală, în special prin interacțiunea cu aplicația mobilă. Scorurile mari acordate acestei componente, combinate cu utilizarea frecventă, indică o preferință clară pentru servicii rapide, accesibile și bine optimizate. În același timp, importanța relației cu personalul bancar sugerează că interacțiunea umană continuă să fie esențială în consolidarea încrederii. Așadar, cele mai eficiente strategii de retenție sunt cele care combină canalele digitale cu suportul personalizat.

Modelele de clasificare utilizate au evidențiat diferențe interesante: regresia logistică

oferă interpretabilitate superioară, SVM asigură o generalizare robustă, iar arborele decizional, deși mai simplu, oferă transparență logică și vizuală. Aceste trăsături pot ghida alegerea modelelor în funcție de context: pentru raportări interne sau strategii de comunicare, modelele interpretabile sunt mai utile; pentru integrare în sisteme automate, contează mai mult acuratețea și echilibrul între clase.

Prin segmentarea K-means, am obținut o înțelegere mai profundă a diferențelor dintre grupurile de clienți. Grupul loial, tânăr și digitalizat, poate fi fidelizat cu funcționalități avansate și comunicare directă în aplicație. Segmentul critic, cu satisfacție scăzută și rezistență la digitalizare, necesită intervenții punctuale: clarificarea comisioanelor, reducerea timpului de răspuns și campanii educaționale personalizate. Între ele, grupul moderat poate fi înclinat spre loialitate prin ajustări subtile și oferte diferențiate.

În esență, rezultatele susțin ideea că loialitatea nu poate fi construită uniform. Este nevoie de o abordare strategică, centrată pe experiența reală a clientului, adaptată la comportamentul și așteptările sale. Băncile care doresc să reducă rata de abandon și să consolideze încrederea pe termen lung trebuie să interpreteze datele nu doar ca puncte statistice, ci ca semnale ale unor nevoi reale, care cer soluții inteligente, rapide și empaticе.

5.3 Limitări ale studiului

În pofida rezultatelor relevante obținute, această cercetare prezintă o serie de limitări metodologice și empirice care trebuie asumate în mod riguros. O primă limitare se referă la dimensiunea redusă a eșantionului utilizat, ceea ce restrânge posibilitatea generalizării concluziilor la nivelul întregii populații bancare din România. Deși dimensiunea actuală a permis aplicarea unor modele statistice și algoritmi de clasificare, validitatea externă a concluziilor rămâne parțial limitată.

În al doilea rând, datele utilizate în analiză au fost exclusiv auto-raportate, ceea ce introduce riscul unor distorsiuni cognitive, precum biasul de confirmare sau tendința de răspuns socialmente acceptabil. Astfel de fenomene pot afecta acuratețea măsurărilor privind satisfacția sau intenția de recomandare, influențând implicit performanța modelelor dezvoltate. În plus, absența unor variabile socio-demografice esențiale – precum venitul lunar, nivelul educațional sau durata relației cu banca – reduce granularitatea interpretării și limitează identificarea unor corelații potențial relevante.

Totodată, natura transversală a datasetului – colectat într-un singur moment în timp – nu permite analiza dinamicii satisfacției sau monitorizarea evoluției comportamentelor de loialitate. Studiul rămâne astfel expus la limitările inerente ale cercetărilor cross-secționale,

care nu pot capta cauzalitatea sau schimbările pe termen lung. În plus, nu au fost integrate surse de date comportamentale (ex. loguri de utilizare a aplicației, frecvența tranzacțiilor digitale), care ar fi oferit o perspectivă complementară, bazată pe acțiuni concrete, și ar fi permis triangularea datelor declarative.

În fine, alegerea algoritmilor de clasificare a privilegiat modelele tradiționale, cu grad crescut de interpretabilitate, în detrimentul unor tehnici avansate precum ensemble learning sau rețele neuronale. Deși această decizie a fost justificată prin scopul aplicabilității în contexte organizaționale, ea constituie o limitare a complexității predictive ce merită investigată în cercetări ulterioare, prin comparație sistematică a performanței diverselor abordări algoritmice.

5.4 Direcții viitoare de cercetare

Pentru aprofundarea și extinderea concluziilor generate de această cercetare, se impun câteva direcții viitoare relevante, cu potențial de a spori atât robustețea metodologică, cât și aplicabilitatea practică a rezultatelor. În primul rând, extinderea eșantionului la nivel național, cu o distribuție echilibrată între segmente demografice și regiuni geografice, ar contribui la creșterea validității externe a modelelor. Integrarea unor dimensiuni suplimentare, precum tipul relației cu banca, istoricul fidelității, nivelul educațional și venitul, ar permite o analiză mai granulară și mai contextualizată a satisfacției bancare.

O direcție semnificativă o constituie includerea de date comportamentale obținute din surse operaționale - cum ar fi frecvența utilizării aplicației mobile, tipurile de servicii accesate sau durata sesiunilor digitale - care pot fi corelate cu răspunsurile declarative pentru a construi modele hibride, mai robuste și mai precise. Acest tip de triangulare ar reduce biasurile de auto-raportare și ar deschide calea spre profiluri comportamentale mai fidele realității.

De asemenea, o extindere metodologică relevantă ar putea viza testarea unor algoritmi mai sofisticati, precum gradient boosting, random forests sau rețele neuronale, pentru a compara performanțele predictive cu modelele tradiționale utilizate în această lucrare. Explorarea unor abordări automatizate, bazate pe AutoML sau optimizarea hiperparametrilor prin grid search ori random search, ar putea aduce un plus de valoare procesului decizional în instituțiile bancare. În plus, replicarea studiului în contexte internaționale sau în alte sectoare de servicii ar permite evaluarea stabilității relațiilor identificate și adaptabilitatea modelelor la alte medii economice și culturale. Nu în ultimul rând, integrarea acestor analize într-un sistem de suport decizional ar putea transforma rezultatele cercetării într-un instrument strategic pentru îmbunătățirea experienței clienților, într-un mod personalizat și scalabil, cu efecte directe asupra retenției și loializării.

Bibliografie

- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. CRC Press.
- Cahyono, B., Sunarti, & Rofiq, A. (2020). The effect of customer experience, customer satisfaction, and customer loyalty on brand equity. *Management Science Letters*, 10(3), 591–598. <https://doi.org/10.5267/j.msl.2019.9.010>
- Chaudhuri, S., & Madigan, D. (2016). An Overview of Data Mining Applications. *Communications of the ACM*, 45(8), 213–220.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
- Cronin, J. J., & Taylor, S. A. (1992). Measuring service quality: A reexamination and extension. *Journal of Marketing*, 56(3), 55–68.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- Davenport, T. H., & Harris, J. G. (2017). *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.
- Keaveney, S. M. (1995). Customer switching behavior in service industries: An exploratory study. *Journal of Marketing*, 59(2), 71–82.
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, 1(6), 90–95.
- Kotler, P., & Keller, K. L. (2016). *Marketing Management* (15th ed.). Pearson Education.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and*

Probability, 1(14), 281–297.

Malaquias, R. F., & Hwang, Y. (2019). Mobile banking and customer satisfaction: A systematic review. *Information Development*, 36(3), 401–417.

McKinsey & Company. (2021). A time for reinvention – Challenges and solutions for the Romanian banking system. Disponibil la: <https://www.mckinsey.com/ro/our-insights/a-time-for-reinvention>

Ng, I. C., & Forbes, J. (2009). Education as service: The understanding of university experience through the service logic. *Journal of Marketing for Higher Education*, 19(1), 38–64.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, 64(1), 12–40.

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality. *Journal of Service Research*, 7(3), 213–233.

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.

Studiu de caz Alpha Bank Contact Center. (2023). Disponibil la: <https://datanets.ro/clienti-si-referinte/studiu-de-caz-alpha-bank-contact-center/>

Tan, P. N., Steinbach, M., & Kumar, V. (2019). *Introduction to Data Mining* (2nd ed.). Pearson.

Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. *Journal of Retailing*, 91(2), 174–181.

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing*, 85(1), 31–41.

Wedel, M., & Kamakura, W. A. (2000). *Market Segmentation: Conceptual and Methodological Foundations* (2nd ed.). Springer Science & Business Media.

Anexa 1. Cod sursă: Fișier principal de execuție - main.py

Această anexă conține codul principal utilizat pentru a lansa toate etapele analizei – de la încărcarea datelor și analiza descriptivă până la preprocesare, aplicarea modelelor de clasificare și segmentarea prin clustering. Fișierul coordonează rularea funcțiilor definite în modulele auxiliare.

```
from data_loader import incarca_date, inspecteaza_date
from descriptiva import analiza_descriptiva, tabele_pivot
from preprocesare import codificare_categorii, standardizare_variabile, imparte_date
from modele import antreneaza_si_evalueaza_model, clustering_kmeans
from interpretare import concluzii_clasificare, concluzii_clustering, sugestii_strategice
from vizualizari_avansate import vizualizari_avansate

# 1. Încarcă date
cale_fisier = "raspunsuri_participanti.csv"
df = incarca_date(cale_fisier)

# 2. Inspectează datele
inspecteaza_date(df)

# 3. Analiza descriptivă
analiza_descriptiva(df)

# 3.1 Vizualizări suplimentare pentru interpretare
vizualizari_avansate(df)

# 4. Tabele pivot
tabele_pivot(df)

# 5. Preprocesare date
df_codificat = codificare_categorii(df)

# Separăm coloanele pentru clasificare vs. clustering
features_model = [
    'Scor interacțiune personal',
    'Scor claritate comisioane',
    'Scor rapiditate soluționare',
    'Scor aplicație mobilă',
    'Scor platformă online',
    'Sex', 'Vârstă', 'Folosește aplicația' ]
```

```

features_clustering = features_model + ['Scor general satisfacție']
# 6. Clasificare: standardizare + împărțire + modele
df_std = standardizare_variabile(df_codificat, features_model)
X_train, X_test, y_train, y_test = imparte_date(df_std, features_model, 'Recomandă
banca')
antreneaza_si_evalueaza_model(X_train, X_test, y_train, y_test,
feature_names=features_model)
# 7. Clustering (K-means)
df_clusters = clustering_kmeans(df_codificat, features=features_clustering,
nr_grupuri=3)
# 8. Interpretări finale
concluzii_clasificare()
concluzii_clustering()
sugestii_strategice()

```

Anexa 2. Cod sursă: Încărcarea și inspecția datelor - data_loader.py

Acest modul conține funcțiile utilizate pentru încărcarea fișierului CSV cu datele colectate din chestionar, precum și inspecția inițială a structurii datasetului. A fost punctul de pornire pentru întreaga analiză.

```
import pandas as pd

def incarca_date(cale_fisier):
    """
    Încarcă datele dintr-un fișier CSV.
    """
    df = pd.read_csv(cale_fisier)
    return df

def inspecteaza_date(df):
    """
    Afișează primele rânduri și informații generale despre dataset.
    """
    print("Primele 5 rânduri:")
    print(df.head())
    print("\nInformații generale:")
    print(df.info())
```


Anexa 3. Cod sursă: Analiza descriptivă a variabilelor - descriptiva.py

Acest modul include funcții pentru generarea statisticilor descriptive și pentru vizualizarea distribuțiilor variabilelor-cheie, precum vârsta și scorurile de satisfacție. A fost utilizat în secțiunea 4.1 pentru a caracteriza eșantionul și a identifica eventuale anomalii sau tendințe inițiale în date.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

def statistici_descriptive(df):
    """
    Calculează statisticile descriptive pentru coloanele numerice.
    """
    print(df.describe())

def distributie_varsta(df):
    """
    Afișează histograma distribuției vârstei respondenților.
    """
    plt.figure(figsize=(8, 5))
    sns.histplot(df['Varsta'], kde=True, bins=10)
    plt.title('Distribuția vârstei respondenților')
    plt.xlabel('Vârstă')
    plt.ylabel('Frecvență')
    plt.show()

def distributie_scoruri(df, coloana):
    """
    Afișează distribuția scorurilor pentru o anumită dimensiune de satisfacție.
    """
    plt.figure(figsize=(8, 5))
    sns.histplot(df[coloana], kde=True, bins=5)
```

```
plt.title(f'Distribuția scorurilor pentru {coloana}')  
plt.xlabel('Scor')  
plt.ylabel('Frecvență')  
plt.show()
```

Anexa 4. Cod sursă: Generarea tabelelor pivot pentru analize comparative

- genereaza_tabele_pivot.py

Acest modul a fost utilizat pentru a construi tabele pivotate care evidențiază diferențele de scoruri între diverse categorii de respondenți – în special în funcție de utilizarea aplicației mobile. A sprijinit interpretările din secțiunea 4.2, permițând compararea mediei scorurilor pentru fiecare grup analizat.

```
def genereaza_tabel_pivot(df, index, columns, values):
    """
    Creează un tabel pivot cu scorurile medii pentru fiecare combinație între index și
    coloană.

    Parametri:
    - df: DataFrame-ul sursă
    - index: coloana care va fi folosită pe rânduri
    - columns: coloana folosită pentru coloane
    - values: coloana numerică pentru care se calculează media

    Returnează:
    - tabelul pivot rezultat
    """
    pivot = pd.pivot_table(df, index=index, columns=columns, values=values,
aggfunc='mean')
    return pivot


def compara_scoruri_dupa_utilizare_aplicatie(df):
    """
    Compară scorurile medii în funcție de utilizarea aplicației mobile.

    Returnează:
    - DataFrame cu scorurile medii pentru fiecare categorie (Da/Nu)
    """
    grupuri = df.groupby("Foloseste aplicatia").mean(numeric_only=True)
    return grupuri
```

Anexa 5. Cod sursă: Interpretarea modelelor predictive - interpretare.py

Acest modul a fost utilizat pentru extragerea coeficienților și a importanței variabilelor din modelele de clasificare (logistică, SVM, arbore decizional). El a susținut analiza factorilor determinanți ai loialității clienților.

```
def extrage_coeficienti_regresie(model, feature_names):
    """
    Returnează un dicționar cu coeficienții modelului de regresie logistică.
    Parametri:
    - model: modelul de regresie logistică antrenat
    - feature_names: lista cu denumirile caracteristicilor
    Returnează:
    - dicționar cu perechi (nume variabilă, coeficient)
    """
    coef_dict = {}
    for coef, feat in zip(model.coef_[0], feature_names):
        coef_dict[feat] = round(coef, 3)
    return coef_dict

def extrage_importanta_svm(model, feature_names):
    """
    Returnează importanța variabilelor pentru un model SVM liniar.
    Parametri:
    - model: modelul SVM antrenat
    - feature_names: lista caracteristicilor
    Returnează:
    - dicționar cu scorurile de importanță
    """
    weights = model.coef_[0]
    coef_dict = {feat: round(weight, 3) for feat, weight in zip(feature_names, weights)}
    return coef_dict

def extrage_importanta_arbre(model, feature_names):
```

"""

Returnează importanța variabilelor pentru arborele decizional.

Parametri:

- model: arborele de decizie antrenat
- feature_names: lista caracteristicilor

Returnează:

- dicționar cu scorurile de importanță

"""

```
importances = model.feature_importances_
```

```
coef_dict = {feat: round(imp, 3) for feat, imp in zip(feature_names, importances)}
```

```
return coef_dict
```

Anexa 6. Cod Sursă: Preprocesarea datelor - preprocesare.py

Acest modul definește funcțiile necesare pentru prelucrarea datasetului înainte de antrenarea modelelor. Include codificarea variabilelor categorice, standardizarea numerică a trăsăturilor și separarea în seturi de antrenare și testare pentru algoritmi de clasificare.

```
def codificare_categorii(df):
    df_copy = df.copy()
    df_copy['Sex'] = df_copy['Sex'].map({'Masculin': 0, 'Feminin': 1})
    df_copy['Foloseste aplicatia'] = df_copy['Foloseste aplicația'].map({'Nu': 0, 'Da': 1})
    df_copy['Recomandă banca'] = df_copy['Recomandă banca'].map({'Nu': 0, 'Da': 1})
    return df_copy

def standardizare_variabile(df, coloane):
    scaler = StandardScaler()
    df_std = df.copy()
    df_std[coloane] = scaler.fit_transform(df[coloane])
    return df_std

def imparte_date(df, coloane, target):
    X = df[coloane]
    y = df[target]
    return train_test_split(X, y, test_size=0.3, random_state=42)
```

Anexa 7. Cod Sursă: Antrenarea și evaluarea modelelor predictive - modele.py

Acest modul conține funcțiile utilizate pentru antrenarea modelelor de clasificare (regresie logistică, SVM, arbore decizional) și pentru evaluarea performanței acestora folosind metrice standard. De asemenea, include implementarea clusteringului K-means pentru segmentarea respondenților.

```
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score, f1_score, recall_score
from sklearn.cluster import KMeans
import pandas as pd

def antreneaza_si_evalueaza_model(X_train, X_test, y_train, y_test, feature_names):
    modele = {
        "Regresie Logistică": LogisticRegression(),
        "SVM": SVC(kernel="linear"),
        "Arbore Decizional": DecisionTreeClassifier()
    }
    for nume, model in modele.items():
        model.fit(X_train, y_train)
        y_pred = model.predict(X_test)
        acc = accuracy_score(y_test, y_pred)
        f1 = f1_score(y_test, y_pred)
        recall = recall_score(y_test, y_pred)
        print(f"\nModel: {nume}")
        print(f'Acuratețe: {acc:.2f}')
        print(f'F1-score: {f1:.2f}')
        print(f'Recall: {recall:.2f}')

def clustering_kmeans(df, features, nr_grupuri=3):
```

```
kmeans = KMeans(n_clusters=nr_grupuri, random_state=42)
df_copy = df.copy()
df_copy['Grup Clustrare'] = kmeans.fit_predict(df_copy[features])
return df_copy
```