

**DEPARTAMENTUL DE INGINERIE ELECTRICĂ, ELECTRONICĂ ȘI CALCULATOARE  
PROGRAMUL DE STUDII CALCULATOARE**

**APLICAȚIE MOBILĂ PENTRU PREDICTION  
CREDITĂRII ÎN DOMENIUL IMOBILIAR**

LUCRARE DE LICENȚĂ

Absolvent: **Giulia-Ana POP**

Coordonatori: **Prof. dr. Habil. Ing. Oliviu MATEI**  
științifici: **Drd. Ing. Erdei RUDOLF**

**2024**

**DEPARTAMENTUL DE INGINERIE ELECTRICĂ, ELECTRONICĂ ȘI CALCULATOARE**  
**PROGRAMUL DE STUDII CALCULATOARE**

DECAN,  
Conf. dr. ing. ec. Dinu DARABĂ

DIRECTOR DEPARTAMENT,  
Sl. dr. ing. Claudiu LUNG

Absolvent: **Giulia-Ana POP**

**APLICAȚIE MOBILĂ PENTRU PREDICTIA CREDITĂRII ÎN DOMENIUL  
IMOBILIAR**

1. **Enunțul temei:** *Dezvoltarea unei aplicații mobile care utilizează un sistem de predicție a creditării ipotecare bazat pe un algoritm de învățare automată pentru a îmbunătăți acuratețea și eficiența procesului de aprobată a împrumuturilor bancare.*
2. **Conținutul lucrării:** *Lucrarea conține următoarele capitole: Introducere, Obiectivele proiectului, Studiu bibliografic, Analiză și fundamentare teoretică, Proiectare de detaliu și implementare, Testare și validare, Concluzii și Bibliografie*
3. **Locul documentării:** Centrul Universitar NORD din Baia Mare
4. **Consultanți:**
5. **Data emiterii temei:** 1 Noiembrie 2023
6. **Data predării:** 10 iulie 2024

Absolvent: Giulia-Ana POP

Coordonatori Prof. dr. Habil. Ing. Oliviu MATEI  
științifici: Drd. Ing. Erdei RUDOLF

# Cuprins

<b>Capitolul 1 Introducere</b>	<b>1</b>
1.1 Contextul proiectului . . . . .	1
1.2 Încadrarea problemei. Motivare . . . . .	3
<b>Capitolul 2 Obiectivele proiectului</b>	<b>4</b>
2.1 Scopul proiectului . . . . .	4
2.2 Obiectivele proiectului . . . . .	4
2.3 Indicatori de performanță . . . . .	5
<b>Capitolul 3 Studiu bibliografic</b>	<b>6</b>
3.1 Introducere . . . . .	6
3.2 Studii anterioare . . . . .	6
3.2.1 Studiu privind cunoștințele financiare și interesul pentru educația financiară . . . . .	6
3.2.2 Studii despre utilizarea algoritmilor de învățare automată în predicția acordării împrumuturilor ipotecare . . . . .	8
3.3 Aplicații și platforme web . . . . .	9
3.3.1 Mint . . . . .	9
3.3.2 Credit Karma . . . . .	9
3.3.3 NerdWallet . . . . .	10
3.3.4 Zillow . . . . .	10
3.3.5 Rocket Mortgage . . . . .	10
3.3.6 Loan Calculator . . . . .	11
3.3.7 HDFC Bank Home Loans . . . . .	11
3.4 Provocări și oportunități . . . . .	12
3.4.1 Provocări . . . . .	12
3.4.2 Oportunități . . . . .	12
3.5 Concluzii . . . . .	12
<b>Capitolul 4 Analiză și fundamentare teoretică</b>	<b>13</b>
4.1 Introducere . . . . .	13
4.2 Evaluarea Riscului de Credit . . . . .	13
4.3 Pregătirea Datelor . . . . .	14
4.3.1 Curățarea Datelor . . . . .	14
4.3.2 Transformarea Datelor . . . . .	14
4.4 Hiperparametri în Modelele de Machine Learning . . . . .	14
4.4.1 Numărul de Estimatori ( <i>n_estimators</i> ) . . . . .	14
4.4.2 Adâncimea Maximă ( <i>max_depth</i> ) . . . . .	14
4.4.3 Rata de Învățare ( <i>learning_rate</i> ) . . . . .	15
4.5 Algoritmi de Machine Learning . . . . .	15
4.5.1 Decision Tree . . . . .	16
4.5.2 Random Forest . . . . .	17
4.5.3 K-Nearest Neighbors (K-NN) . . . . .	17
4.5.4 XGBoost . . . . .	18

---

4.6	Formule Utilizate pentru Rambursarea Împrumuturilor cu Rată Fixă și Variabilă . . . . .	19
4.6.1	Rata Fixă . . . . .	20
4.6.2	Rata Variabilă . . . . .	20
4.7	Cont de Economii . . . . .	20
4.8	Rambursare Optimală . . . . .	20
4.8.1	Calcularea Plății Lunare a Împrumutului . . . . .	20
4.8.2	Calcularea Dobânzii Totale a Împrumutului . . . . .	21
4.8.3	Calcularea Contribuției Lunare pentru Acoperirea Dobânzii . . . . .	21
4.9	Prezentarea sursei și descrierea setului de date . . . . .	21
4.9.1	Descrierea sursei . . . . .	21
4.9.2	Calitatea datelor . . . . .	21
4.9.3	Descrierea atributelor . . . . .	22
<b>Capitolul 5</b>	<b>Proiectare de detaliu și implementare</b>	<b>23</b>
5.1	Definirea problemei și a soluției alese . . . . .	23
5.2	Arhitectura aplicației . . . . .	23
5.2.1	Utilizator . . . . .	24
5.2.2	Frontend (Android) . . . . .	24
5.2.3	Backend (Python) . . . . .	24
5.2.4	Algoritmul ML XGBoost . . . . .	24
5.2.5	Fluxul de date între Front-end și Back-end . . . . .	24
5.3	Analizarea setului de date . . . . .	25
5.4	Pregătirea datelor și corelația dintre ele . . . . .	26
5.4.1	Curățarea datelor . . . . .	26
5.4.2	Prelucrarea datelor . . . . .	26
5.4.3	Etichetarea datelor . . . . .	27
5.4.4	Matricea de corelație . . . . .	27
5.5	Antrenarea modelelor . . . . .	28
5.5.1	Eliminarea atributelor nefolositoare și împărțirea setului de date .	28
5.5.2	Antrenarea modelelor . . . . .	29
5.6	Comparație între algoritmii de învățare automată . . . . .	29
5.7	Implementarea aplicației mobile . . . . .	30
5.7.1	Arhitectura MVC . . . . .	30
5.7.2	Cazuri de utilizare . . . . .	31
5.7.3	Designul și implementarea paginilor principale . . . . .	32
5.7.4	Interfața utilizatorului . . . . .	35
5.7.5	Navigarea în aplicație . . . . .	35
5.8	Concluzie implementare . . . . .	35
<b>Capitolul 6</b>	<b>Testare și validare</b>	<b>36</b>
6.1	Evaluarea algoritmilor . . . . .	36
6.1.1	Decision Tree . . . . .	36
6.1.2	Random Forest . . . . .	36
6.1.3	k-Nearest Neighbors (k-NN) . . . . .	37
6.1.4	XGBoost . . . . .	37
6.2	Testare SonarQube . . . . .	38
6.2.1	Probleme de securitate . . . . .	38

6.2.2	Probleme de mentenabilitate . . . . .	39
6.2.3	Probleme de consistență . . . . .	41
6.2.4	Rezultate finale . . . . .	42
6.3	Validare prin feedback-ul utilizatorilor . . . . .	42
6.3.1	Selectarea utilizatorilor . . . . .	42
6.3.2	Implementarea îmbunătățirilor . . . . .	43
<b>Capitolul 7</b>	<b>Manual de instalare și utilizare</b>	<b>44</b>
7.1	Resurse software și hardware necesare . . . . .	44
7.1.1	Resurse hardware . . . . .	44
7.1.2	Resurse software . . . . .	44
7.2	Instalarea aplicației . . . . .	44
7.2.1	Instalare pe smartphone Android . . . . .	44
7.2.2	Instalare backend pe server local . . . . .	44
7.3	Utilizarea aplicației . . . . .	45
7.3.1	Predictia aprobării împrumutului . . . . .	45
7.3.2	Calculator de rambursare a împrumutului . . . . .	48
7.3.3	Calculator de economii . . . . .	49
7.3.4	Optimizarea rambursării împrumutului . . . . .	50
7.4	Concluzie . . . . .	50
<b>Capitolul 8</b>	<b>Concluzii</b>	<b>51</b>
8.1	Dezvoltări ulterioare . . . . .	51
<b>Bibliografie</b>		<b>53</b>

# Capitolul 1. Introducere

## 1.1. Contextul proiectului

În cadrul industriei financiare, creditele ipotecare joacă un rol esențial în asigurarea accesului la locuințe pentru indivizi și familii. Această activitate nu doar facilitează realizarea visului de a detine o casă, ci și influențează în mod semnificativ evoluția pieței imobiliare și stabilitatea economică generală. Cu toate acestea, procesul de acordare a creditelor ipotecare este unul complex și plin de riscuri pentru instituțiile financiare.

Conform celor mai recente statistici ale Băncii Naționale a României (BNR), numărul creditelor ipotecare a crescut semnificativ în ultimii ani, reflectând o creștere a cererii de locuințe și o accesibilitate sporită a finanțării. În 2023, volumul creditelor ipotecare noi acordate în lei în primele 9 luni a crescut constant în comparație cu aceeași perioadă a anului precedent. Motivul acestei creșteri este datorat ofertelor tot mai atractive de dobândă din partea băncilor, de care mulți români au profitat pentru achiziționarea unei locuințe. În total, volumul creditelor ipotecare a crescut cu 31% în anul 2023<sup>1</sup>, cum se poate vedea și în figura 1.1<sup>2</sup>.

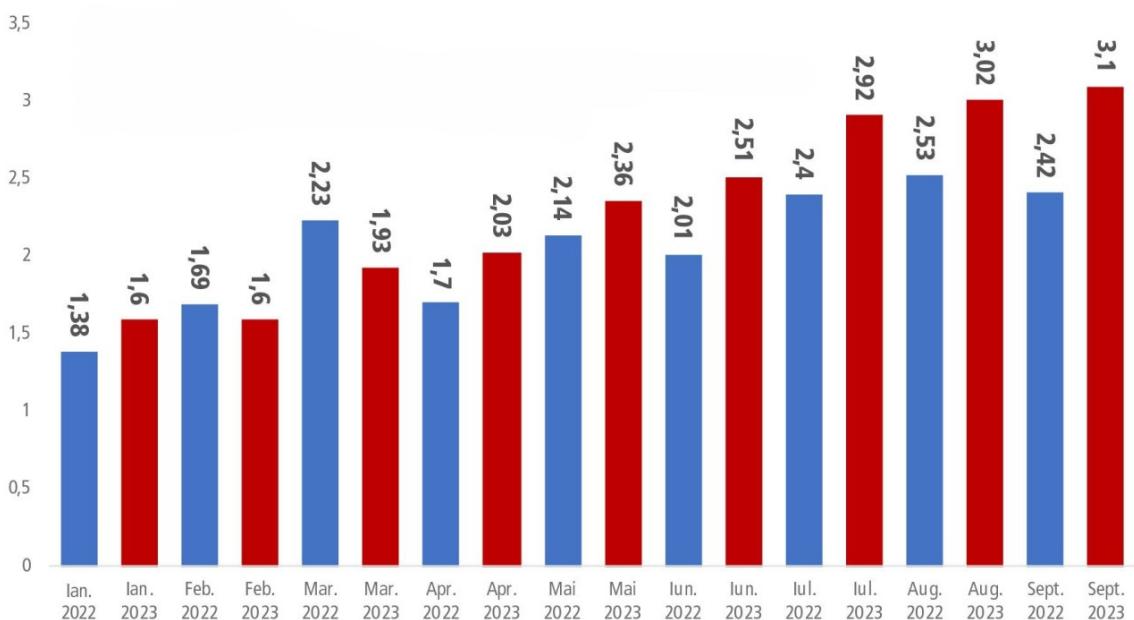


Figura 1.1: Statistică a evoluției creditelor ipotecare în 2023 comparativ cu 2022

Într-un mediu economic în continuă schimbare și în fața unei piețe imobiliare volatile, instituțiile financiare sunt presate să eficientizeze procesele de evaluare a riscului și să reducă expunerea la pierderi. Evaluarea inexactă a riscului poate duce la acordarea de credite către debitori nepotriviți, expunând instituțiile financiare la riscuri financiare.

<sup>1</sup><https://www.zf.ro/banci-si-asigurari/>, accesat: 15 ianuarie 2024 [Online]

<sup>2</sup><https://storage0.dms.mpinteractive.ro/media/1/1481/22466/22192109/1/5-credite-ipotecare.jpg>, accesat: 15 ianuarie 2024 [Online]

și operaționale considerabile[1]. Prin urmare, există o nevoie crescută de instrumente și tehnologii care să ajute instituțiile financiare în procesul lor de luare a deciziilor și să optimizeze procesul de acordare a creditelor ipotecare. Aceste instrumente trebuie să fie precise, eficiente și adaptate contextului economic actual, pentru a asigura stabilitatea financiară și a preveni pierderile[2].

Într-o lume în care deciziile financiare devin din ce în ce mai complexe, educația financiară a populației joacă un rol crucial. Mulți indivizi se confruntă cu dificultăți în a înțelege conceptele de bază ale creditării și economisirii, ceea ce poate duce la decizii financiare nesăbuite. Lipsa educației financiare poate avea consecințe negative asupra gestionării finanțelor personale și asupra stabilității economice generale[3]. Este important ca indivizii să dobândească cunoștințe financiare pentru a putea lua decizii informate și pentru a-și îmbunătăți situația materială.

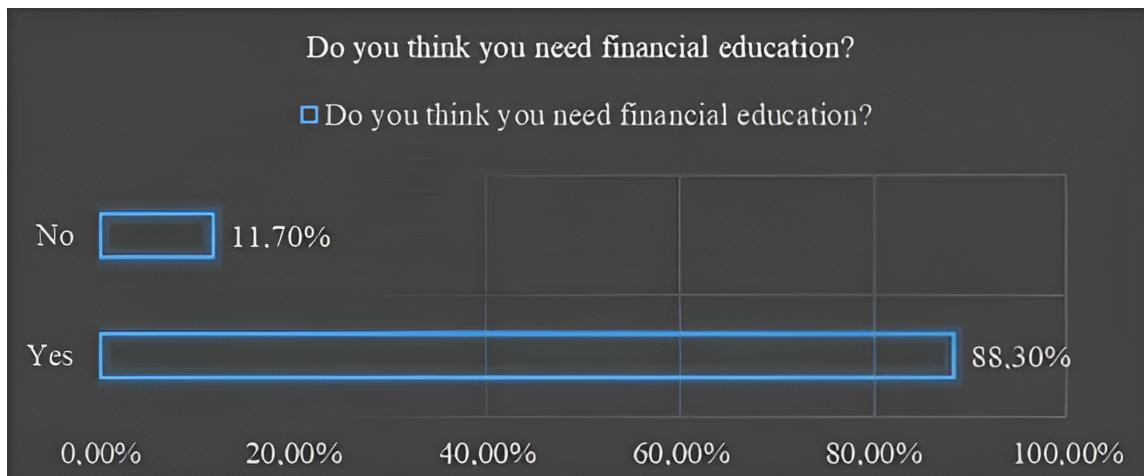


Figura 1.2: Nevoia pentru educație financiară în România[4]

Figura 1.2 ilustrează rezultatele unui studiu efectuat în 2023 asupra adulților din România, care a dezvăluit că un procent semnificativ (88.3%) dintre respondenți consideră că au nevoie de educație financiară, în timp ce doar 11.7% nu văd această necesitate[4]. Aceste statistici evidențiază necesitatea unor soluții inovatoare și eficiente pentru înțelegerea mai clară a finanțelor personale. Educația financiară nu doar că îi ajută pe indivizi să își gestioneze mai bine resursele financiare, dar contribuie și la stabilitatea economică generală, prevenind decizii financiare nesăbuite care ar putea avea un impact negativ atât la nivel personal, cât și la nivel macroeconomic.

Într-un astfel de context, este esențial ca instituțiile financiare să își asume un rol activ în promovarea educației financiare. Acestea ar trebui să dezvolte și să ofere resurse educaționale accesibile și relevante pentru publicul larg, inclusiv materiale educative, seminarii și aplicații interactive. De asemenea, colaborarea cu instituții educaționale și organizații non-guvernamentale poate amplifica impactul acestor inițiative, asigurând o transmitere mai largă a informațiilor și o implicare comunitară sporită. Prin educația financiară, indivizii pot dobândi competențele necesare pentru a naviga în peisajul finanțier complex și pentru a lua decizii informate, ceea ce poate conduce la o societate mai bine informată și mai responsabilă din punct de vedere financiar. Prin urmare, aplicația dezvoltată nu doar că poate sprijini instituțiile financiare în evaluarea riscului de credit, dar servește și ca un instrument educațional pentru utilizatori, contribuind astfel la o mai bună înțelegere și gestionare a finanțelor personale.

## 1.2. Încadrarea problemei. Motivare

Motivația din spatele acestui proiect de licență derivă din necesitatea de a dezvolta o soluție inovatoare și eficientă pentru evaluarea riscului și facilitarea accesului la creditare ipotecară. Într-o lume în care datele și tehnologiile avansate devin din ce în ce mai predominante, este esențial să se exploreze modalități noi și precise de evaluare a riscului și de luare a deciziilor financiare în sectorul imobiliar.

Într-un mediu economic dinamic și cu o piată imobiliară în continuă schimbare, tehnologiile precum învățarea automată și analiza datelor oferă instrumente puternice pentru a îmbunătăți procesele de luare a deciziilor financiare. Aceste tehnologii permit instituțiilor financiare să analizeze și să extragă informații valoroase din volume mari de date, identificând modele și tendințe ascunse care pot influența deciziile de creditare[5].

Prin dezvoltarea și implementarea unui model predictiv robust, acest proiect își propune să ofere o soluție practică și eficientă pentru instituțiile financiare care acordă credite ipotecare. Modelul predictiv va analiza datele relevante ale solicitantului și ale proprietății și va estima probabilitatea acordării unui credit pentru fiecare solicitant în parte. Această estimare va permite instituțiilor financiare să evaluateze mai precis și mai rapid riscul asociat fiecarui împrumut și să ia decizii informate și eficiente.

Un alt factor important care motivează acest proiect este creșterea competitivității în sectorul finanțier. Instituțiile financiare se află sub o presiune constantă pentru a inova și a oferi servicii și produse mai bune și mai eficiente[6]. Dezvoltarea unui model predictiv precis pentru creditarea ipotecară ar putea reprezenta un avantaj competitiv semnificativ pentru instituțiile financiare, permitându-le să ofere produse și servicii mai personalizate și adaptate nevoilor clienților lor.

În plus, aplicația dezvoltată le oferă utilizatorilor posibilitatea de a simula diverse scenarii finanțieri, de a înțelege impactul ratelor dobânzii și al termenelor de rambursare asupra finanțelor lor personale și de a-și planifica mai bine viitorul finanțier. Aplicația poate funcționa și ca un ghid în procesul de economisire, ajutând utilizatorii să-și stabilească obiective de economisire pe termen scurt sau lung și să estimateze cât de mult vor economisi într-o anumită perioadă, cum ar fi un an. Aplicația include, de asemenea, un calculator care ajută utilizatorii să determine cât ar trebui să economisească lunar pentru a recupera suma plătită ca dobândă până la sfârșitul perioadei de rambursare a împrumutului ipotecar.

În cele din urmă, trebuie subliniată importanța responsabilității și a eticii în utilizarea tehnologiilor pentru evaluarea riscului în creditare. Orice soluție tehnologică dezvoltată în acest domeniu trebuie să fie transparentă, echitabilă și nediscriminatorie și să respecte toate normele și reglementările privind protecția datelor și confidențialitatea informațiilor personale[7]. În acest sens, proiectul de licență își propune să răspundă acestei nevoi, oferind o soluție practică care să contribuie la creșterea stabilității și a performanței sectorului imobiliar în ansamblu.

## **Capitolul 2. Obiectivele proiectului**

### **2.1. Scopul proiectului**

Scopul principal al acestui proiect de licență este dezvoltarea unei aplicații mobile care să permită evaluarea eligibilității acordării unui credit ipotecar și să ofere funcționalități extinse pentru utilizatori, inclusiv calculatoare de rambursare a creditelor și economiei. Aceasta are scopul de a facilita luarea deciziilor financiare informate atât pentru instituțiile financiare, cât și pentru utilizatori, contribuind astfel la creșterea stabilității și performanței pieței imobiliare.

### **2.2. Obiectivele proiectului**

Obiectivele proiectului sunt gândite pentru a asigura dezvoltarea unei soluții eficiente și accesibile care să răspundă atât nevoilor instituțiilor financiare, cât și ale utilizatorilor individuali. Fiecare obiectiv contribuie la realizarea scopului general al proiectului, asigurând funcționalități esențiale și performanțe ridicate. În continuare sunt prezentate detaliat obiectivele acestui proiect:

#### **O1 Aplicație mobilă, accesibilă tuturor**

- Dezvoltarea unei aplicații mobile intuitive și ușor de utilizat, accesibilă pe diferite platforme mobile. Obiectivul este de a crea o interfață atractivă și simplă, care să faciliteze utilizarea eficientă de către toți utilizatorii, indiferent de experiența lor tehnică.

#### **O2 Utilizarea unui algoritm de învățare automată pentru predicția acordării creditelor ipotecare**

- Implementarea unui algoritm de învățare automată care să analizeze datele utilizatorilor și să estimeze probabilitatea acordării unui credit ipotecar. Acest obiectiv vizează integrarea tehnologiilor avansate pentru a oferi estimări precise și rapide, contribuind la eficientizarea procesului de creditare.

#### **O3 Validarea sistemului în condiții reale**

- Validarea aplicației prin testarea în condiții reale de utilizare este esențială pentru a asigura funcționalitatea și eficiența atât a predicțiilor, cât și a aplicației în general. Acest obiectiv urmărește verificarea funcționalității și utilității aplicației în scenarii reale și adunarea de informații valoroase pentru îmbunătățiri ulterioare.

#### **O4 Implementarea unui calculator de rambursare a creditelor**

- Dezvoltarea unei funcționalități complete care să permită utilizatorilor să calculeze rambursările lunare ale creditelor, oferind un plan detaliat de rambursare. Acest instrument va ajuta utilizatorii să înțeleagă mai bine cum se desfășoară procesul de rambursare și să își planifice finanțier viitorul.

#### **O5 Implementarea unui calculator de economii**

- Crearea unui simulator precis care să ajute utilizatorii să estimeze acumularea de economii prin depuneri regulate, oferind o imagine clară asupra economiilor pe termen lung. Acest calculator va încuraja utilizatorii să economisească și să-și gestioneze mai eficient resursele financiare.

#### O6 Sugestie personalizată de gestionare financiară

- Oferirea de recomandări personalizate pentru gestionarea eficientă a bugetului personal, bazate pe analiza datelor financiare ale utilizatorilor și obiectivele lor economice. Acest obiectiv urmărește să sprijine utilizatorii în a lua decizii financiare mai informate și să își îmbunătățească situația financiară.

### 2.3. Indicatori de performanță

Pentru a măsura succesul obiectivelor proiectului, au fost definiți următorii indicatori de performanță. Acești indicatori vor ajuta la evaluarea eficienței și impactului aplicației dezvoltate, asigurând că proiectul își atinge scopurile propuse:

#### R1 O aplicație mobilă funcțională și intuitivă

- Aplicația trebuie să fie complet funcțională, să ofere o experiență de utilizare plăcută, fără erori tehnice majore, și să fie ușor de navigat pentru orice utilizator, indiferent de nivelul său de cunoștințe tehnice. Acest rezultat va confirma că aplicația este bine concepută și eficientă în utilizarea de zi cu zi.

#### R2 Un algoritm de învățare automată

- Algoritmul de învățare automată trebuie să fie capabil să analizeze datele și să ofere predicții precise privind acordarea creditelor ipotecare. Implementarea unui astfel de algoritm va demonstra capacitatea tehnologică a aplicației de a lua decizii informate.

#### R3 O acuratețe de cel puțin 75%

- Algoritmul implementat trebuie să atingă o acuratețe de cel puțin 75%, asigurând astfel fiabilitatea și utilitatea sa în deciziile de creditare. Acest indicator va valida eficiența tehnologică a aplicației în predicția corectă a eligibilității creditului.

#### R4 Sondaj cu minim 5 persoane pentru testarea interactivității aplicației

- Eficiența și interactivitatea aplicației trebuie validate prin sondaje și feedback de la cel puțin 5 utilizatori, asigurând astfel că aplicația răspunde nevoilor reale ale utilizatorilor. Acest rezultat va ajuta la identificarea și rezolvarea eventualelor probleme, îmbunătățind astfel experiența utilizatorilor.

#### R5 O funcționalitate completă pentru calcularea rambursărilor lunare

- Calculatorul de rambursare trebuie să ofere utilizatorilor un plan detaliat și precis de rambursare a creditelor, cu toate informațiile necesare pentru gestionarea corectă a finanțelor. Acest indicator va confirma utilitatea practică a aplicației în planificarea financiară a utilizatorilor.

#### R6 O simulare precisă a acumulării de economii prin depunerile regulate

- Simulatorul de economii trebuie să ofere estimări precise ale acumulării de economii, permitând utilizatorilor să își planifice economiile pe termen scurt și lung în mod eficient. Acest rezultat va demonstra funcționalitatea completă și acuratețea simulatorului de economii.

#### R7 O recomandare eficientă pentru gestionarea bugetului personal

- Aplicația trebuie să ofere recomandări personalizate pentru gestionarea bugetului, bazate pe analiza datelor financiare ale utilizatorilor, contribuind astfel la o gestionare financiară mai responsabilă și informată. Acest indicator va asigura că utilizatorii primesc sfaturi valoroase pentru îmbunătățirea situației lor financiare.

## Capitolul 3. Studiu bibliografic

### 3.1. Introducere

În ultimii ani, avansurile tehnologice au transformat semnificativ domeniul creditării ipotecare. Aplicațiile mobile și platformele web au devenit esențiale pentru instituțiile financiare și pentru utilizatori, oferind soluții inovatoare pentru evaluarea riscului, gestionarea creditelor și optimizarea finanțelor personale. Algoritmii de învățare automată au jucat un rol crucial în accelerarea și îmbunătățirea proceselor de predicție a acordării creditelor, contribuind la decizii mai rapide și mai precise. Acest capitol examinează starea actuală a domeniului, evidențiind principalele aplicații și site-uri web care oferă funcționalități similare cu cele ale acestui proiect și trecând în revistă studii anterioare despre nivelul de cunoștințe financiare al populației.

### 3.2. Studii anterioare

#### 3.2.1. Studiu privind cunoștințele financiare și interesul pentru educația financiară

Un studiu recent realizat în România a arătat nivelul scăzut de educație financiară și interesul populației pentru subiectele financiare. Conform studiului[8], 92% din populația României este considerată financiar analfabetă, ceea ce subliniază necesitatea unor aplicații eficiente de educație financiară.

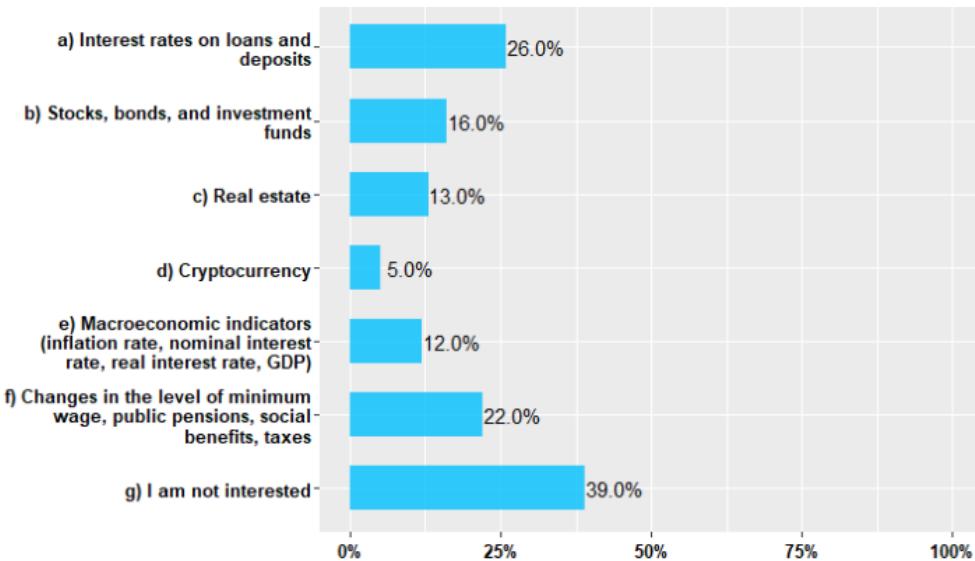


Figura 3.1: Interesul populației pentru concepte financiare[8]

Studiul a relevat că 39% dintre respondenți nu doresc să obțină sfaturi financiare, în timp ce majoritatea celor interesați (26%) ar dori să afle despre dobânzile creditelor și depozitelor (figura 3.1). Aceasta evidențiază un interes destul de mare pentru credite și economii, dar și un dezinteres general pentru educația financiară, care ar putea fi legată de neîncrederea în instituțiile financiare și de lipsa resurselor financiare.

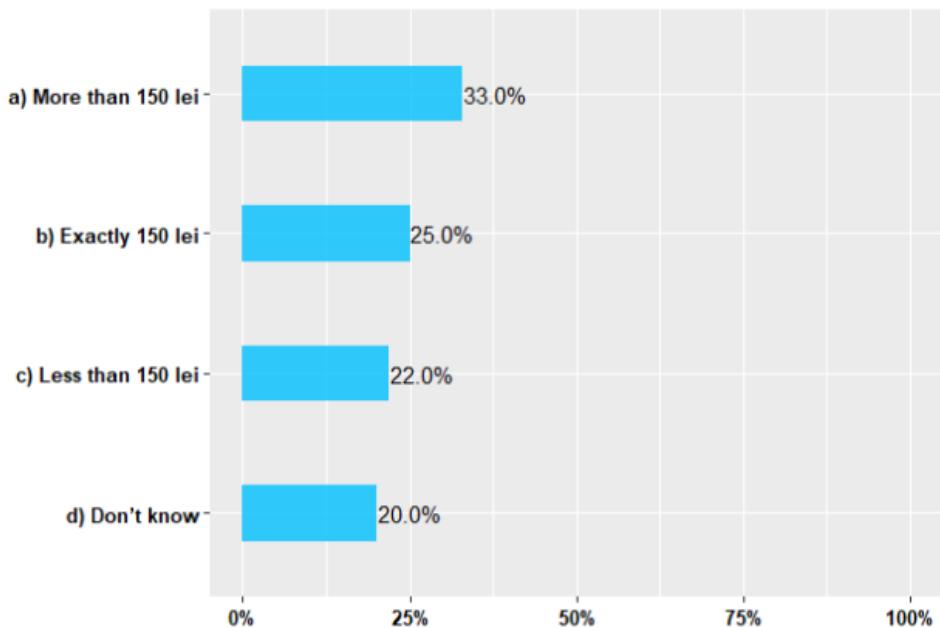


Figura 3.2: Să presupunem că aveți 100 LEI într-un cont de economii, iar rata dobânzii este de 10% pe an. După 5 ani, cât credeți că veți avea în cont dacă lăsați banii să crească? [8]

Într-o altă întrebare, *Să presupunem că aveți 100 LEI într-un cont de economii, iar rata dobânzii este de 10% pe an. După 5 ani, cât credeți că veți avea în cont dacă lăsați banii să crească?*, studiul a arătat că doar 33% dintre respondenți au înțeles corect efectul dobânzii asupra economiilor pe termen lung (figura 3.2). Aceste rezultate subliniază nevoia de aplicații care să îmbunătățească înțelegerea conceptelor financiare de bază.

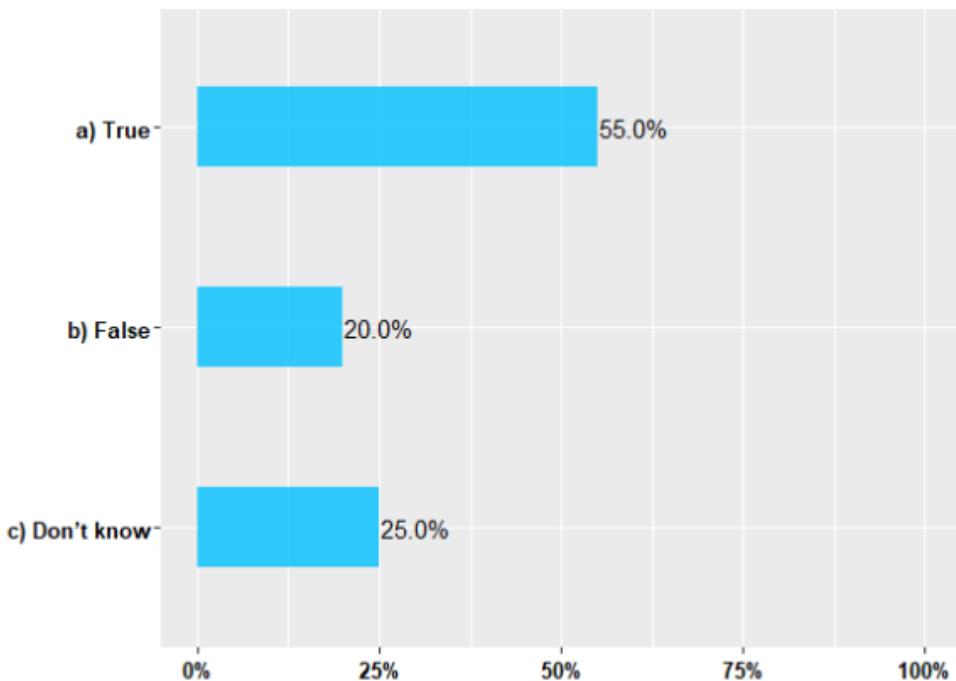


Figura 3.3: Adevărat sau fals: O ipotecă pe 15 ani necesită de obicei plăti lunare mai mari decât una pe 30 de ani, dar dobânda totală pe durata împrumutului va fi mai mică [8]

Alegerea făcută dintre adevărat sau fals pentru enunțul: *O ipotecă pe 15 ani necesită de obicei plăți lunare mai mari decât una pe 30 de ani, dar dobânda totală pe durata împrumutului va fi mai mică*, arată că doar 55% dintre respondenți au răspuns corect, înțelegând că o ipotecă pe 15 ani necesită plăți lunare mai mari decât una pe 30 de ani, dar cu dobândă totală mai mică pe durata împrumutului (figura 3.3). Acest lucru sugerează că există lacune semnificative în înțelegerea conceptelor de bază ale creditelor ipotecare în rândul populației din România.

### 3.2.2. Studii despre utilizarea algoritmilor de învățare automată în predicția acordării împrumuturilor ipotecare

Un număr semnificativ de studii recente au explorat utilizarea algoritmilor de învățare automată pentru predicția acordării împrumuturilor ipotecare, evidențiind eficacitatea acestor metode în reducerea riscului și optimizarea deciziilor de creditare.

Oualid, Hansali, Balouki și Moumoun prezintă în articolul [9], o analiză detaliată a diverselor tehnici de învățare automată aplicate în gestionarea riscului de credit. Studiul subliniază avantajele și provocările asociate utilizării acestor tehnici, precum și eficiența lor în predicția riscului de credit. Autorii au comparat metode precum regresia logistică, rețele neuronale și arborii de decizie, concluzionând că modelele bazate pe învățare automată pot îmbunătăți semnificativ acuratețea predicțiilor și pot reduce rata creditelor neperformante. Studiul subliniază avantajele și provocările asociate utilizării acestor tehnici, precum și eficiența lor în predicția riscului de credit. Autorii au comparat metode precum regresia logistică, rețele neuronale și arborii de decizie, concluzionând că modelele bazate pe învățare automată pot îmbunătăți semnificativ acuratețea predicțiilor și pot reduce rata creditelor neperformante.

Gupta, Pant, Kumar și Bansal în studiul [10], au demonstrat eficiența utilizării tehnicielor de învățare automată pentru predicția acordării împrumuturilor bancare. Folosind algoritmi precum Logistic Regression și Random Forest, cercetătorii au dezvoltat un sistem de predicție a împrumuturilor. Studiul a inclus analiza unui set de date extins, aplicând diverse tehnici de preprocesare a datelor și evaluând performanța modelelor pe baza mai multor metrii de acuratețe. Rezultatele au arătat că utilizarea tehnicielor de învățare automată poate îmbunătăți semnificativ procesul de evaluare a creditelor.

Sharma și Kumar în lucrarea [11], au explorat diferite metode de învățare automată pentru a analiza un set de date complex. Autorii au evaluat performanța fiecărui model în funcție de acuratețea, precizia și sensibilitatea acestora, concluzionând că modelele de învățare automată pot oferi predicții mai precise și mai rapide decât verificarea fiecărui aplicant de către angajații băncilor.

Dansana, Patro, Mishra, Prasad, Razak și Wogasso în articolul [12], au analizat impactul caracteristicilor împrumuturilor asupra predicției acestora folosind algoritmul Random Forest. Studiul a examinat variabile precum venitul solicitantului, scorul de credit și istoricul de rambursare, evaluând influența acestora asupra probabilității de aprobare a împrumutului. Rezultatele au arătat că Random Forest poate oferi predicții precise și fiabile, optimizând procesul decizional al băncilor și reducând riscurile.

Dutta în studiul [13], a examinat modul în care algoritmii de învățare automată pot îmbunătăți predicția împrumuturilor. Cercetarea a inclus o comparație între diferite modele de învățare automată, cum ar fi Naive Bayes, Decision Trees și Gradient Boosting, analizând performanța acestora în funcție de diverse seturi de date. Studiul a evidențiat beneficiile utilizării acestor tehnici pentru creșterea acurateței și fiabilității predicțiilor de împrumut, subliniind importanța preprocesării adecvate a datelor și a selecției caracte-

risticilor relevante.

Chakraborty și Joseph în lucrarea [14], au investigat aplicarea algoritmilor de învățare automată la nivelul băncilor centrale pentru predicția riscului de credit și gestionarea politicilor financiare. Studiul a analizat utilizarea tehnicielor avansate precum Deep Learning și Reinforcement Learning pentru monitorizarea stabilității financiare și optimizarea politicilor monetare. Autorii au subliniat importanța adoptării acestor tehnologii pentru îmbunătățirea stabilității financiare și reducerea riscului sistemic.

Anand, Velu și Whig în articolul [15], au explorat utilizarea modelelor de învățare automată pentru predicția comportamentului împrumuturilor în scopul securizării operațiunilor bancare. Studiul a folosit tehnici precum rețele neuronale și modele de ansamblu pentru a analiza seturi de date mari și complexe, evaluând performanța acestora în funcție de precizie și rata de fals pozitive. Rezultatele au arătat că modelele de învățare automată pot îmbunătăți semnificativ securitatea și fiabilitatea predicțiilor de împrumut, contribuind la reducerea riscului operațional pentru bănci.

Aceste studii demonstrează eficacitatea și relevanța algoritmilor de învățare automată în predicția acordării împrumuturilor ipotecare, subliniind importanța utilizării acestor tehnologii pentru îmbunătățirea proceselor decizionale ale instituțiilor financiare și reducerea riscului de credit.

### 3.3. Aplicații și platforme web

În această secțiune, vom explora principalele aplicații și platforme web utilizate pentru creditarea în domeniul imobiliar. Aceste instrumente sunt esențiale atât pentru instituțiile financiare, cât și pentru utilizatori, oferind soluții eficiente pentru gestionarea riscului, optimizarea procesului de creditare și îmbunătățirea înțelegerii financiare a utilizatorilor.

#### 3.3.1. Mint

Mint este o aplicație de gestionare a finanțelor personale. Oferă utilizatorilor posibilitatea de a-și urmări cheltuielile, de a-și crea bugete și de a-și monitoriza scorul de credit. Aplicația oferă, de asemenea, recomandări personalizate pentru economisire și investiții, bazate pe analiza comportamentului financiar al utilizatorului [16].

Avantajele aplicației Mint includ interfață intuitivă și ușor de utilizat, precum și gama largă de funcții financiare disponibile într-un singur loc.

Dezavantajul aplicației este că nu oferă funcționalități specifice pentru calcularea rambursărilor de credite ipotecare și poate fi copleșitoare pentru utilizatorii noi datorită multitudinii de opțiuni. De asemenea, Mint nu include funcționalități pentru predicția aprobării creditului ipotecar.

Aplicația propusă în acest proiect, pe lângă funcțiile de bază de gestionare a finanțelor, va include un calculator de rambursare a creditului ipotecar și predicția aprobării creditului ipotecar, funcții esențiale care lipsesc în Mint.

#### 3.3.2. Credit Karma

Credit Karma este o platformă care oferă utilizatorilor acces gratuit la scorul lor de credit și la rapoartele de credit. Pe lângă aceste servicii, Credit Karma oferă și recomandări personalizate pentru îmbunătățirea scorului de credit și pentru găsirea celor mai bune oferte de credite și împrumuturi. Platforma utilizează algoritmi avansați pentru a analiza datele financiare ale utilizatorilor și pentru a oferi recomandări precise [17].

Avantajele Credit Karma, includ accesul gratuit la scorul și rapoartele de credit, precum și recomandările personalizate pentru îmbunătățirea scorului de credit.

Dezavantajul platformei este limitarea la piața americană, cu funcționalități specifice doar pentru această regiune, și lipsa calculatoarelor pentru rambursarea creditelor ipotecare. De asemenea, Credit Karma nu oferă funcționalități pentru predictia aprobării creditului ipotecar.

Aplicația propusă va oferi nu doar monitorizarea scorului de credit, ci și calculatoare detaliate pentru rambursarea creditelor ipotecare.

### 3.3.3. NerdWallet

NerdWallet este un site web<sup>1</sup>, care oferă informații detaliate despre diverse produse financiare, inclusiv credite ipotecare, carduri de credit și conturi de economii. Utilizatorii pot compara ofertele diferitelor instituții financiare și pot primi recomandări personalizate pe baza profilului lor finanțier. NerdWallet folosește algoritmi de învățare automată pentru a analiza datele și pentru a oferi cele mai bune recomandări pentru utilizatori[18].

Avantajele NerdWallet includ comparațiile detaliate ale produselor financiare și recomandările personalizate bazate pe profilul finanțier al utilizatorilor.

Dezavantajul este că interfața web poate fi mai puțin accesibilă decât o aplicație mobilă și nu oferă un calculator detaliat de rambursare a creditelor ipotecare.

Aplicația noastră va include funcționalități mobile accesibile, calculatoare detaliate pentru credite ipotecare și economii, și predictia aprobării creditului ipotecar, aducând astfel valoare adăugată utilizatorilor care doresc să gestioneze finanțele de pe dispozitivele lor mobile.

### 3.3.4. Zillow

Zillow este o platformă online dedicată pieței imobiliare, oferind utilizatorilor informații despre proprietăți de vânzare și de închiriat. Pe lângă funcționalitățile de bază, Zillow include și un calculator de credite ipotecare, care permite utilizatorilor să estimeze costurile de rambursare a creditelor pe baza parametrilor specifici ai împrumutului [19].

Avantajele platformei Zillow includ informațiile detaliate despre proprietăți imobiliare și includerea unui calculator de credite ipotecare.

Dezavantajul este că funcționalitățile de gestionare a finanțelor sunt limitate în comparație cu alte aplicații, iar interfața web poate fi mai puțin convenabilă decât o aplicație mobilă. De asemenea, Zillow nu oferă funcționalități pentru predictia aprobării creditului ipotecar.

Aplicația noastră va oferi calculatoare de credite ipotecare, economii și predictia aprobării creditului ipotecar într-un format mobil accesibil, facilitând utilizatorilor gestionarea finanțelor de oriunde.

### 3.3.5. Rocket Mortgage

Rocket Mortgage este o platformă online de creditare ipotecară, care permite utilizatorilor să aplice pentru credite ipotecare și să obțină aprobarea rapid și ușor. Platforma utilizează algoritmi de învățare automată pentru a evalua riscul de credit și pentru a oferi aprobări instantanee. Rocket Mortgage se remarcă prin procesul său simplificat și prin

---

<sup>1</sup>NerdWallet,<https://www.nerdwallet.com/>, accesat: 20 martie 2024 [Online].

interfață prietenoasă pentru utilizatori.<sup>2</sup>

Avantajele Rocket Mortgage includ procesul de aplicare simplificat și rapid, precum și evaluarea automată a riscului de credit cu ajutorul algoritmilor de învățare automată.

Dezavantajul este că platforma este limitată la piața americană și nu include funcționalități pentru gestionarea bugetelor și economiilor.

Aplicația noastră va oferi nu doar evaluarea riscului de credit, ci și funcționalități extinse pentru gestionarea finanțelor personale, inclusiv calculatoare de rambursare și economii.

### 3.3.6. Loan Calculator

Loan Calculator este o aplicație disponibilă pe Google Play, care oferă utilizatorilor o gamă largă de calculatoare financiare pentru diferite tipuri de împrumuturi. Utilizatorii pot calcula plățile lunare, costurile totale ale împrumutului și alte aspecte financiare relevante pentru creditele ipotecare.<sup>3</sup>

Avantajele acestei aplicații includ varietatea de calculatoare financiare oferite și interfață simplă și ușor de utilizat.

Dezavantajul aplicației este că nu oferă funcționalități de gestionare a bugetului personal și nu include evaluări de risc de credit sau recomandări personalizate. De asemenea, aplicația nu include predicția aprobării creditului ipotecar.

Aplicația noastră va oferi o gestionare financiară prin care utilizatorul poate afla suma necesară pe care trebuie să o pună în contul de economii pentru acoperirea dobânzii la finalul împrumutului ipotecar.

### 3.3.7. HDFC Bank Home Loans

HDFC Bank Home Loans este o aplicație disponibilă pe Google Play<sup>4</sup> care permite utilizatorilor să aplice pentru credite ipotecare, să calculeze plățile lunare și să gestioneze detaliile împrumuturilor. Aplicația oferă informații detaliate despre ratele dobânzilor și termenii împrumuturilor, facilitând procesul de aplicare pentru credite [20].

Avantajele HDFC Bank Home Loans includ posibilitatea de a aplica direct pentru credite ipotecare și de a calcula plățile lunare, oferind totodată informații detaliate despre împrumuturi.

Dezavantajul este că aplicația este limitată la utilizatorii din India și nu include funcționalități de gestionare a bugetului personal sau evaluări de risc de credit. De asemenea, aplicația nu include predicția aprobării creditului ipotecar.

Aplicația noastră va oferi nu doar funcționalități pentru aplicarea și gestionarea împrumuturilor, ci și instrumente pentru gestionarea bugetului personal, evaluări de risc de credit și predicția aprobării creditului ipotecar, adresând astfel nevoi variate ale utilizatorilor.

Aceste aplicații și platforme oferă o varietate de funcționalități utile, fiecare având propriile sale avantaje și dezavantaje. Aplicația propusă în acest proiect își propune să combine cele mai bune caracteristici ale acestor instrumente existente și să adauge funcționalități suplimentare pentru a crea o soluție completă și eficientă pentru utilizatori.

<sup>2</sup>Rocket Mortgage,<https://www.bankrate.com/mortgages/reviews/rocket-mortgage/#pros-and-cons>, accesat: 20 martie 2024 [Online].

<sup>3</sup>Loan Calculator,[https://play.google.com/store/apps/details?id=au.com.scbrokers.calc&pcampaignid=web\\_share](https://play.google.com/store/apps/details?id=au.com.scbrokers.calc&pcampaignid=web_share), accesat: 20 martie 2024 [Online].

<sup>4</sup>HDFC Bank MobileBanking App,[https://play.google.com/store/apps/details?id=com.snapwork.hdfc&pcampaignid=web\\_share](https://play.google.com/store/apps/details?id=com.snapwork.hdfc&pcampaignid=web_share), accesat: 20 martie 2024 [Online].

### **3.4. Provocări și oportunități**

#### **3.4.1. Provocări**

Una dintre principalele provocări în domeniul creditării ipotecare este evaluarea precisă a riscului de credit. În ciuda avansurilor tehnologice, există încă dificultăți în colectarea și analiza datelor financiare relevante, precum și în integrarea acestora în modelele predictive [21]. Alte provocări includ calitatea datelor, unde datele incomplete sau inexakte pot afecta semnificativ acuratețea modelelor predictive, și heterogenitatea datelor, unde sursele variate de date pot introduce inconsistențe și pot complica procesul de integrare și analiză. Reglementările stricte privind conformitatea cu reglementările financiare și de protecție a datelor pot limita utilizarea unor tehnici de analiză a datelor și pot adăuga complexitate procesului de evaluare a riscului. Resursele limitate sunt o altă provocare, deoarece implementarea și întreținerea unor sisteme avansate de învățare automată necesită resurse semnificative, inclusiv timp, bani și expertiză tehnică.

#### **3.4.2. Oportunități**

Dezvoltarea continuă a tehnologiilor de învățare automată și a analizelor de date deschide noi oportunități pentru îmbunătățirea proceselor de creditare. Aplicațiile mobile și platformele web pot deveni instrumente esențiale pentru utilizatori și instituții financiare, facilitând evaluarea riscului și optimizând gestionarea finanțelor personale [22]. Alte oportunități includ personalizarea ofertelor de credit prin utilizarea datelor colectate, permitând instituțiilor financiare să răspundă mai bine nevoilor și profilului de risc al fie cărui client. Automatizarea proceselor de evaluare a riscului și de luare a deciziilor poate reduce semnificativ timpul și costurile operaționale. Aplicațiile mobile și platformele web fac serviciile financiare mai accesibile pentru o gamă largă de utilizatori, inclusiv pentru cei din zonele rurale sau cu acces limitat la sucursale bancare fizice.

### **3.5. Concluzii**

Starea actuală a domeniului creditării ipotecare este caracterizată de o evoluție rapidă și de integrarea tehnologiilor avansate. Aplicațiile mobile și platformele web joacă un rol crucial în acest context, oferind soluții inovatoare pentru evaluarea riscului și gestionarea finanțelor personale. Proiectul nostru își propune să contribuie la această evoluție, oferind o aplicație mobilă completă care să răspundă nevoilor utilizatorilor și să faciliteze luarea deciziilor financiare informate.

Prin integrarea funcționalităților de calculare a rambursărilor de credite ipotecare, economiei și predicția aprobării creditului ipotecar, aplicația noastră va oferi utilizatorilor un instrument puternic pentru gestionarea eficientă a finanțelor personale. În plus, utilizarea algoritmilor de învățare automată va asigura predicții precise și personalizate, contribuind astfel la optimizarea procesului de creditare.

## Capitolul 4. Analiză și fundamentare teoretică

### 4.1. Introducere

În lumea financiară modernă, creditarea ipotecară joacă un rol crucial în accesul la proprietăți imobiliare. Evaluarea eligibilității pentru credite ipotecare și gestionarea acestora au devenit mai sofisticate datorită avansurilor tehnologice, în special, prin utilizarea algoritmilor de învățare automată. Acest capitol își propune să analizeze detaliile teoretice și practice ale utilizării acestor algoritmi în predicția creditelor ipotecare și să exploreze factorii relevanti în dezvoltarea unei aplicații mobile eficiente pentru gestionarea și predicția creditelor ipotecare.

### 4.2. Evaluarea Riscului de Credit

Evaluarea riscului de credit este crucială în contextul creditării ipotecare. Acest proces implică analiza capacitatei unui solicitant de credit de a-și onora angajamentele financiare pe termen lung[23]. Principalele criterii utilizate în evaluarea riscului includ:

- **Istoricul de credit al solicitantului:** Analiza istoricului de credit implică examinarea comportamentului financiar anterior al solicitantului. Aceasta include evaluarea plășilor trecute, întârzierilor în rambursare și oricăror incidente de neplată. Un istoric de credit pozitiv, cu plăști efectuate la timp și fără incidente majore, indică o probabilitate mai mare ca solicitantul să își îndeplinească obligașii viitoare. De asemenea, un scor de credit ridicat reflectă un comportament financiar responsabil și reduce percepția riscului pentru creditori.
- **Veniturile:** Evaluarea veniturilor lunare și anuale ale solicitantului este esențială pentru a determina capacitatea acestuia de a rambursa creditul. Creditorii analizează sursele de venit, stabilitatea acestora și istoricul de angajare. Venituri stabile și suficiente pentru a acoperi plășile lunare ale creditului sunt un indicator pozitiv al capacitatei de rambursare.
- **Datoriile existente:** Verificarea datoriilor actuale ale solicitantului ajută la evaluarea gradului de îndatorare și a capacitatei acestuia de a face față noilor obligașii financiare. Un nivel ridicat de datori existente poate crește riscul de neplată, deoarece solicitantul ar putea întâmpina dificultăți în gestionarea mai multor obligașii financiare simultan. Creditorii calculează de obicei un raport al datoriilor în funcție de venit pentru a evalua capacitatea solicitantului de a prelua noi datori[24].
- **Valoarea proprietății:** Estimarea valorii proprietății pentru care se solicită creditul este crucială pentru a evalua riscul colateral în caz de neplată. Proprietatea servește ca garanție pentru credit, iar valoarea acesteia trebuie să fie suficientă pentru a acoperi suma împrumutată în caz de executare silită. Evaluatorii profesioniști sunt adesea angajați pentru a furniza o estimare precisă a valorii proprietății, luând în considerare factori precum locația, dimensiunea, starea și piata imobiliară locală.

## 4.3. Pregătirea Datelor

Pregătirea datelor este un pas esențial în procesul de dezvoltare a unui model de învățare automată eficient. Acest proces implică mai multe etape, inclusiv curățarea, transformarea și preprocesarea datelor, pentru a asigura calitatea și relevanța acestora pentru modelul ales [25].

### 4.3.1. Curățarea Datelor

Curățarea datelor presupune identificarea și corectarea sau eliminarea erorilor și a valorilor lipsă din setul de date. Acest pas este crucial pentru a evita antrenarea unui model pe date corupte sau incomplete, ceea ce ar putea cauza o performanță deficitară și rezultate inexacte [26].

- **Identificarea și tratarea valorilor lipsă:** Valorile lipsă pot fi gestionate prin eliminarea observațiilor incomplete sau prin imputarea acestora folosind metode statistice (mediană, medie, mod) sau modele de predicție [27].
- **Detectarea și corectarea erorilor:** Datele eronate sau anomalii trebuie identificate și corectate. Acest lucru poate implica verificarea datelor împotriva unor reguli de validare sau a surselor externe de referință [28].

### 4.3.2. Transformarea Datelor

Transformarea datelor implică conversia datelor brute în formate adecvate pentru analiză[29]. Aceasta poate include normalizarea sau standardizarea valorilor numerice și transformarea datelor categorice în reprezentări numerice.

- **Normalizarea și standardizarea:** Aceste tehnici sunt utilizate pentru a aduce toate valorile numerice într-un interval similar, reducând astfel influența disproporțională a unor variabile asupra modelului.
- **Transformarea datelor categorice:** Datele categorice pot fi transformate în valori numerice pentru a putea fi utilizate de algoritmii de învățare automată care acceptă doar numere.

## 4.4. Hiperparametri în Modelele de Machine Learning

Hiperparametrii sunt variabile configurabile ale unui algoritm de învățare automată, ajustate pentru a optimiza performanța modelului. Alegerea corectă a hiperparametrilor este esențială pentru a obține rezultate precise și robuste. Printre cei mai importanți hiperparametri pentru algoritmii de machine learning se numără numărul de estimatori (*n\_estimators*), adâncimea maximă (*max\_depth*) și rata de învățare (*learning\_rate*)[30].

### 4.4.1. Numărul de Estimatori (*n\_estimators*)

*n\_estimators* reprezintă numărul de arbori de decizie care vor fi construși în ansamblul modelului. Un număr mai mare de arbori poate îmbunătăți acuratețea modelului, dar crește și complexitatea și timpul de calcul. Creșterea numărului de estimatori reduce variabilitatea predicțiilor, ducând la o performanță mai stabilă și mai robustă a modelului.

### 4.4.2. Adâncimea Maximă (*max\_depth*)

*max\_depth* definește adâncimea maximă a fiecărui arbore de decizie. Limitarea adâncimii poate ajuta la prevenirea supraînvățării, asigurându-se că modelul generali-

zează bine pe date noi. Cu toate acestea, o adâncime prea mică poate duce la subînvățare, unde modelul nu capturează suficient de bine complexitatea datelor. Alegerea unei adâncimi optime este esențială pentru a asigura echilibrul între supraînvățare și subînvățare.

#### 4.4.3. Rata de Învățare (*learning\_rate*)

*learning\_rate* controlează rata cu care modelul se ajustează la erorile din fiecare iterare a antrenamentului. O rată de învățare mai mică poate conduce la o convergență mai lentă, dar mai stabilă, în timp ce o rată de învățare mai mare poate accelera antrenamentul, dar cu riscul de a sări peste soluțiile optime. Ajustarea corespunzătoare a ratei de învățare este crucială pentru obținerea unui model eficient și precis.

## 4.5. Algoritmi de Machine Learning

În prezent, domeniul creditării ipotecare beneficiază de avansurile tehnologice oferite de algoritmii de învățare automată. Aceste tehnologii permit o evaluare mai precisă și mai rapidă a riscului de credit, contribuind la optimizarea procesului decizional și la reducerea riscului pentru instituțiile financiare. Algoritmii de învățare automată sunt capabili să analizeze volume mari de date și să identifice tipare care ar fi dificil de detectat prin metode tradiționale. Astfel, acești algoritmi ajută la îmbunătățirea acurateței predicțiilor și la eficientizarea procesului de acordare a creditelor. Utilizarea eficientă a algoritmilor de învățare automată poate reduce semnificativ timpul de procesare și poate oferi o mai mare fiabilitate în evaluarea riscului asociat cu împrumuturile ipotecare.

Gupta et al. au subliniat în articolul [31], importanța utilizării algoritmilor de învățare automată pentru predicția eligibilității împrumuturilor. Studiul lor a demonstrat că utilizarea acestor algoritmi crește acuratețea și eficiența în predicția aprobării împrumuturilor bancare. Algoritmii testați în studiu, inclusiv Random Forest și Logistic Regression, au arătat performanțe superioare față de metodele tradiționale, reducând erorile de predicție și timpul necesar pentru evaluarea aplicațiilor de credit. Aceasta subliniază importanța utilizării tehniciilor moderne de învățare automată în sectorul bancar pentru a îmbunătăți procesul decizional și a minimiza riscurile asociate cu acordarea creditelor.

În mod similar, Arutjothi și Senthamarai au demonstrat, în studiul [32], că algoritmii de învățare automată sunt extrem de eficienți în determinarea eligibilității împrumuturilor bancare, îmbunătățind procesul decizional al băncilor. Studiul a evidențiat faptul că algoritmii precum K-Nearest Neighbors și Support Vector Machines oferă rezultate fiabile și rapide, contribuind la un proces decizional mai informat și mai transparent. Implementarea acestor algoritmi în sistemele bancare poate conduce la o mai bună gestionare a riscului și la o creștere a încrederii în capacitatea instituțiilor financiare de a evalua corect solicitările de credit.

Diversitatea algoritmilor de învățare automată oferă soluții adaptabile pentru diferite seturi de date și cerințe specifice. Algoritmi precum Decision Tree, Random Forest, K-Nearest Neighbors și XGBoost sunt frecvent utilizați în acest domeniu datorită capacității lor de a gestiona complexitatea și volumul mare de date. Fiecare dintre acești algoritmi are caracteristici și avantaje specifice, făcându-i potriviti pentru diverse aplicații, precum evaluarea riscului de credit și predicția acordării împrumuturilor. Alegerea algoritmului potrivit depinde de natura datelor disponibile și de obiectivele specifice ale analizei.

#### 4.5.1. Decision Tree

Algoritmul Decision Tree (Arborele de Decizie) este unul dintre cei mai utilizati algoritmi de clasificare și regresie datorită clarității și ușurinței de interpretare. Acesta funcționează prin împărțirea setului de date în subseturi mai mici, folosind reguli de decizie bazate pe caracteristicile datelor. Fiecare nod intern al arborelui reprezintă un atribut de date, fiecare ramură rezultatul unui test aplicat pe acel atribut, iar fiecare frunză reprezintă o etichetă de clasă sau o valoare numerică în cazul regresiei [33].

Algoritmul începe de la un nod rădăcină și utilizează criterii de impuritate, cum ar fi Gini și entropia, pentru a selecta atributele care oferă cea mai bună separare a datelor. Indicele Gini măsoară impuritatea unui nod, indicând probabilitatea ca un element aleatoriu să fie clasificat incorect dacă ar fi etichetat aleatoriu conform distribuției de etichete din acel nod. Practic, un indice Gini mai mic indică un nod mai pur, adică un nod în care majoritatea elementelor aparțin unei singure clase. Entropia, pe de altă parte, măsoară incertitudinea sau dezordinea dintr-un sistem, fiind calculată pe baza probabilităților de apariție a fiecărei clase. Un sistem cu entropie mare este unul în care datele sunt distribuite uniform între clase, ceea ce indică o mare incertitudine, pe când un sistem cu entropie mică este unul în care datele sunt concentrate într-o singură clasă, reducând astfel incertitudinea [34]. Procesul de împărțire continuă până când toate datele sunt clasificate sau se îndeplinește un criteriu de oprire, cum ar fi atingerea unei adâncimi maxime a arborelui sau a unui număr minim de elemente într-un nod.

Acest algoritm prezintă numeroase avantaje, care îl fac extrem de atractiv pentru diverse aplicații:

- **Ușor de înțeles și interpretat:** Structura arborelui este intuitivă, putând fi explicată cu ușurință chiar și celor care nu sunt experți. Fiecare decizie poate fi vizualizată sub forma unui grafic de tip arbore, ceea ce facilitează înțelegerea procesului de clasificare. Acest aspect este deosebit de valoros în domeniul unde interpretabilitatea este crucială, cum ar fi medicina sau finanțele [35].
- **Capabil să gestioneze date numerice și categorice:** Algoritmul poate fi aplicat pe diverse tipuri de date fără prea multe modificări, fiind astfel foarte flexibil și versatil în diferite aplicații. Aceasta înseamnă că poate manipula date care variază de la numere întregi și reale la categorii distincte, fără a necesita preprocesări complexe [36].
- **Captează relații complexe și non-liniare:** Permite modelarea interacțiunilor complexe dintre variabile, ceea ce îl face util în situațiile unde aceste interacțiuni sunt prezente. Poate identifica și reprezenta relații non-liniare între variabile, oferind o capacitate predictivă robustă, chiar și în problemele unde modelele liniare simple ar eșua [37].

Totuși, deși algoritmul oferă multiple avantaje, există și dezavantaje semnificative care trebuie luate în considerare pentru a asigura utilizarea sa eficientă și pentru a evita problemele comune care pot apărea în implementarea acestuia:

- **Supraînvățare (overfitting):** Dacă adâncimea arborelui nu este limitată, acesta poate memora datele de antrenament, performând slab pe datele noi. Aceasta se întâmplă deoarece arborele poate deveni prea complex și adaptat specificităților setului de antrenament. Supraînvățarea poate fi redusă prin utilizarea tehnicilor de tăiere (pruning), care elimină ramurile inutile sau redundante [38].
- **Sensibil la variațiile din date:** Mici modificări în setul de date pot duce la structuri complet diferite ale arborelui. Aceasta înseamnă că modelul este instabil și poate genera rezultate diferite cu seturi de date ușor variate [39].

- **Necesitate tăiere (pruning):** Pentru a combate supraînvățarea, este adesea necesară tăierea arborelui, un proces complex și consumator de timp. Pruning-ul implică eliminarea ramurilor care au o contribuție mică sau nulă la puterea predictivă a modelului, reducând astfel complexitatea acestuia și îmbunătățind generalizarea pe date noi [40].

Algoritmul Decision Tree este folosit pe scară largă datorită ușurinței de interpretare și aplicabilității sale versatile în diverse domenii, inclusiv creditarea ipotecară. Capacitatea sa de a modela relații complexe și de a oferi o interpretare clară a deciziilor îl face un instrument valoros în evaluarea riscului de credit și în alte aplicații financiare.

#### 4.5.2. Random Forest

Acest algoritm este o extensie a arborelui de decizie și este utilizat frecvent pentru clasificare și regresie datorită robustei și preciziei sale. Random Forest funcționează prin construirea unui număr mare de arbori de decizie în timpul antrenamentului. În cazul clasificării, algoritmul determină clasa finală prin agregarea predicțiilor tuturor arborilor, iar pentru regresie, prin calcularea mediei predicțiilor arborilor [41].

Fiecare arbore din pădure este construit folosind o selecție aleatorie de date din setul de antrenament. În plus, la fiecare nod de decizie, un subset aleatoriu al caracteristicilor este selectat pentru a determina cea mai bună diviziune. Această metodă reduce varianța modelului și ajută la prevenirea supraînvățării (overfitting) [42].

Acest algoritm are numeroase avantaje semnificative, ceea ce îl face extrem de valoros și util pentru diverse aplicații, precum:

- **Robust împotriva supraînvățării:** Prin construirea mai multor arbori și combinarea predicțiilor lor, Random Forest reduce riscul de supraînvățare prezent în arborii de decizie simpli [43].
- **Capabil să gestioneze date mari și complexe:** Algoritmul poate gestiona seturi de date mari și complexe cu multe caracteristici și instanțe, fiind foarte eficient în probleme de clasificare și regresie.
- **Flexibil și versatil:** Random Forest poate fi utilizat atât pentru probleme de clasificare, cât și pentru probleme de regresie. De asemenea, poate gestiona datele lipsă și poate estima importanța variabilelor [44].

Cu toate acestea, există și dezavantaje importante care trebuie avute în vedere pentru a utiliza algoritmul eficient și a evita problemele comune:

- **Consum mare de resurse:** Random Forest necesită mai multe resurse computaționale pentru antrenarea și predicția modelului, datorită numărului mare de arbori care trebuie construși și evaluati [45].
- **Complexitate crescută:** Deși fiecare arbore individual poate fi simplu, combinația de mulți arbori poate duce la un model complex care este greu de interpretat [46].
- **Sensibil la datele cu zgomot:** În prezența datelor cu zgomot, performanța Random Forest poate fi afectată, deoarece zgomotul poate influența construcția arborilor individuali [47].

Algoritmul Random Forest este folosit pe scară largă datorită capacitatei sale de a gestiona date complexe și variate, asigurând în același timp o performanță predictivă robustă.

#### 4.5.3. K-Nearest Neighbors (K-NN)

Algoritmul K-Nearest Neighbors (K-NN) este un model de învățare automată supervizată utilizat pentru probleme de clasificare și regresie. Acesta funcționează pe principiu

piul similarității, presupunând că instanțele care sunt apropiate în spațiul caracteristicilor vor avea clase similare. Algoritmul K-NN nu implică un proces de antrenament propriu-zis, ci toate datele de antrenament sunt stocate și utilizate în timpul predicției. Pentru a prezice clasa unei instanțe necunoscute, algoritmul calculează distanțele între aceasta și toate instanțele din setul de antrenament, selectând cele mai apropiate  $K$  instanțe (vecini). Clasa majoritară dintre acești vecini este atribuită instanței necunoscute în cazul clasificării, sau se calculează media valorilor în cazul regresiei [48].

Procesul K-NN implică mai multe etape. În primul rând, se calculează distanța dintre instanța necunoscută și toate instanțele din setul de antrenament, utilizând metriki de distanță precum distanța Euclidiană, distanța Manhattan sau distanța Minkowski. Distanța Euclidiană, de exemplu, este calculată ca rădăcina pătrată a sumei pătratelor diferențelor dintre valorile corespunzătoare ale caracteristicilor. După calcularea distanțelor, se selecteză cele mai apropiate  $K$  instanțe și se determină clasa majoritară sau media valorilor, în funcție de tipul de problemă (clasificare sau regresie) [49].

Algoritmul K-NN prezintă numeroase avantaje, care îl fac o alegere populară pentru diverse aplicații:

- **Simplicitate și Intuitivitate:** K-NN este ușor de înțeles și implementat, fiind un model simplu care nu necesită o fază de antrenament propriu-zisă. Este ideal pentru problemele în care interpretabilitatea este importantă și unde relațiile dintre date sunt evidente [50].
- **Performanță bună pentru seturi de date mici:** K-NN poate oferi rezultate precise și eficiente atunci când este aplicat pe seturi de date de dimensiuni reduse, unde relațiile dintre date sunt clare și bine definite.
- **Flexibilitate:** Poate fi utilizat atât pentru probleme de clasificare, cât și pentru probleme de regresie, fiind un model versatil care se adaptează ușor la diferite tipuri de probleme [51].

Totuși, există și dezavantaje semnificative care trebuie luate în considerare pentru a asigura utilizarea eficientă a acestui algoritm:

- **Cost Computațional Ridicat:** Calcularea distanței pentru toate instanțele din setul de antrenament poate fi costisitoare din punct de vedere computațional, mai ales pentru seturi de date mari. Acest lucru poate duce la un timp de predicție mare, ceea ce îl face ineficient pentru aplicațiile în timp real [52].
- **Dependență de Metrică:** Performanța algoritmului depinde în mare măsură de alegerea metricii de distanță, care poate influența acuratețea rezultatelor. Diferite metriki pot da rezultate diferite, iar alegerea metricii potrivite este esențială pentru obținerea unor predicții precise [53].
- **Sensibilitate la Zgomot:** K-NN poate fi sensibil la zgomot și la datele nerelevante, ceea ce poate duce la clasificări incorecte. În prezența zgomotului, performanța algoritmului poate scădea semnificativ, necesitând metode suplimentare de preprocesare a datelor pentru a elimina zgomotul [54].

K-Nearest Neighbors este un algoritm robust și flexibil, care, în ciuda dezavantajelor, rămâne o alegere populară pentru multe aplicații datorită simplității și eficienței sale în problemele de clasificare și regresie.

#### 4.5.4. XGBoost

XGBoost (Extreme Gradient Boosting)[55] este un algoritm de învățare automată dezvoltat pentru a optimiza performanța modelelor de clasificare și regresie prin utilizarea tehnicii de boosting. Acest algoritm a fost conceput pentru a fi extrem de eficient și

scalabil, fiind capabil să proceseze seturi de date mari și complexe cu rapiditate. XGBoost este recunoscut pentru viteza și performanța sa ridicată în competițiile de știință datelor, cum ar fi cele organizate pe platforma Kaggle, și este larg utilizat în industrie datorită capacitații sale de a gestiona date variate și complexe.

Algoritmul XGBoost funcționează prin crearea unui ansamblu de arbori de decizie, fiecare arbore fiind construit pentru a corecta erorile făcute de arborele precedent. În boosting, fiecare arbore nou este adăugat în mod secvențial, cu scopul de a reduce erorile de predicție ale modelului existent. XGBoost utilizează tehnici pentru a controla complexitatea modelului și pentru a preveni supraînvățarea (overfitting), ceea ce îl face foarte eficient pentru date zgomotoase și seturi mari de date.

XGBoost are numeroase caracteristici care îl fac o alegere de top pentru o varietate de aplicații:

- **Performanță ridicată:** Datorită optimizațiilor sale, XGBoost este foarte rapid și eficient, putând procesa seturi mari de date într-un timp scurt. Aceasta îl face ideal pentru aplicațiile în care timpul de calcul este critic.
- **Precizie mare:** Algoritmul oferă o acuratețe excelentă datorită metodei sale de corectare iterativă a erorilor și utilizării regularizării pentru a preveni supraînvățarea.
- **Flexibilitate:** XGBoost poate fi utilizat pentru o varietate de sarcini de clasificare și regresie și poate gestiona date eterogene, fiind capabil să proceseze atât date numerice, cât și categorice [56].
- **Robust la date zgomotoase:** Utilizarea regularizării și a tehnicii de ensemble face ca XGBoost să fie robust la date zgomotoase și la variațiile din seturile de date [57].

Totuși, algoritmul are și o serie de dezavantaje care trebuie luate în considerare înainte de a decide dacă este potrivit pentru cazul propriu. Aceste aspecte trebuie evaluate cu atenție pentru a asigura că XGBoost este cea mai bună alegere pentru problema specifică abordată:

- **Complexitate ridicată:** Implementarea și ajustarea modelului XGBoost pot fi complexe și necesită o înțelegere profundă a parametrilor și a hiperparametrilor pentru a obține performanță optimă.
- **Consum mare de resurse:** XGBoost poate consuma multe resurse de calcul, mai ales atunci când se lucrează cu seturi de date mari, care pot duce la tempi lungi de antrenament și pot necesita echipamente hardware mai performante.
- **Necesitatea ajustării atenției la hiperparametri:** Pentru a preveni supraînvățarea și a optimiza performanța, este esențial să se ajusteze cu atenție hiperparametrii modelului, ceea ce poate fi un proces consumator de timp și resurse.

XGBoost este un algoritm de învățare automată extrem de puternic și versatil, care, în ciuda complexității și a cerințelor mari de resurse, oferă performanțe remarcabile în problemele de clasificare și regresie, făcându-l ideal pentru aplicații complexe, inclusiv în evaluarea riscului de credit și alte domenii financiare.

### 4.6. Formule Utilizate pentru Rambursarea Împrumuturilor cu Rată Fixă și Variabilă

Pentru a dezvolta funcționalitățile de rambursare a creditului, se utilizează diverse formule matematice esențiale pentru modelarea comportamentului financiar al indivizilor și instituțiilor. Aceste formule permit calcularea plășilor necesare pentru rambursarea creditelor în condiții variabile sau fixe.

#### 4.6.1. Rata Fixă

Formula utilizată pentru calcularea plății lunare a unui credit cu rată fixă este:

$$M = \frac{P \cdot r \cdot (1 + r)^n}{(1 + r)^n - 1} \quad (4.1)$$

unde:

- $M$  este plata lunară,
- $P$  este principalul împrumutului,
- $r$  este dobânda lunară,
- $n$  este numărul total de plăți.

Formula (4.1) este un exemplu de utilizare a capital budgeting în contextul împrumuturilor, după cum este discutat în detaliu de Kohn [58].

#### 4.6.2. Rata Variabilă

Pentru un credit cu rată variabilă, plățile lunare pot fi recalculare periodic pe baza ratelor actualizate ale dobânzii, utilizând aceeași formulă, dar ajustând rata dobânzii în funcție de intervalele de ajustare specificate[58].

### 4.7. Cont de Economii

Formula pentru calcularea valorii viitoare din contul de economii, cu o depunere inițială și depuneri lunare regulate, este:

$$FV = P_0 \cdot (1 + r)^n + P \cdot \frac{(1 + r)^n - 1}{r} \quad (4.2)$$

unde:

- $FV$  este valoarea viitoare din contul de economii,
- $P_0$  este suma inițială depusă,
- $P$  este plata lunară,
- $r$  este dobânda lunară,
- $n$  este numărul total de luni.

Această formulă este derivată și explicată în detaliu în diverse surse academice, inclusiv articole online<sup>1</sup> și studii financiare [59].

### 4.8. Rambursare Optimală

Pentru a optimiza rambursarea unui credit, utilizăm o serie de formule care ne permit să calculăm plățile lunare, dobânda totală și contribuția lunară necesară pentru acoperirea dobânzii. Aceste calcule sunt esențiale pentru a stabili o strategie de rambursare eficientă și pentru a minimiza costurile totale ale împrumutului.

#### 4.8.1. Calcularea Plății Lunare a Împrumutului

Am calculat plata lunară a împrumutului folosind formula (4.1), prezentată în subcapitolul referitor la rata fixă. Această formulă este crucială pentru determinarea exactă a sumei lunare care trebuie plătită, asigurând astfel un plan de rambursare bine structurat și predictibil.

---

<sup>1</sup>[https://en.wikipedia.org/wiki/Compound\\_interest](https://en.wikipedia.org/wiki/Compound_interest), accesat: 6 martie 2024 [Online]

#### 4.8.2. Calcularea Dobânzii Totale a Împrumutului

Formula pentru calcularea dobânzii totale a unui împrumut este:

$$\text{Dobânda Totală} = M \cdot n - P \quad (4.3)$$

Această formulă calculează dobânda totală plătită pe durata împrumutului, scăzând suma principalului din produsul dintre plata lunată și numărul total de plăti.

#### 4.8.3. Calcularea Contribuției Lunare pentru Acoperirea Dobânzii

În cadrul cercetării, am dezvoltat o formulă care determină contribuția lunată  $C$  necesară într-un cont de economii pe durata împrumutului, astfel încât suma acumulată să acopere integral dobânda totală a împrumutului la sfârșitul perioadei. Formula, care asigură reducerea costurilor totale asociate împrumutului, este definită astfel:

$$C = \frac{\text{Dobânda Totală} - S_0 \cdot \left(1 + \frac{r_s}{12}\right)^n}{\left(1 + \frac{r_s}{12}\right)^n - 1} \cdot \frac{r_s}{12} \quad (4.4)$$

unde:

- Dobânda Totală este dobânda totală a împrumutului calculată anterior,
- $S_0$  este suma inițială din contul de economii,
- $r_s$  este dobânda pentru economii, exprimată ca fractie pentru dobânda lunată ( $\frac{r_s}{12}$ ),
- $n$  este numărul total de luni în care am pus bani în contul de economii.

### 4.9. Prezentarea sursei și descrierea setului de date

Setul de date utilizat pentru dezvoltarea modelului predictiv provine de pe platforma Kaggle, fiind intitulat *Loan Prediction Problem Dataset*<sup>2</sup>. Acesta a fost colectat și preprocesat pentru a servi ca bază în dezvoltarea și testarea algoritmilor de învățare automată pentru predicția riscului de credit.

#### 4.9.1. Descrierea sursei

Platforma Kaggle este cunoscută pentru furnizarea unor seturi de date variate și de înaltă calitate, utilizate în competiții de data science și în proiecte academice și industriale. Setul de date *Loan Prediction Problem Dataset* a fost creat pentru a ajuta cercetătorii și practicienii să dezvolte modele eficiente de predicție a acordării împrumuturilor, oferind date relevante despre solicitanții de credit ipotecar.

#### 4.9.2. Calitatea datelor

Acest set de date a fost ales datorită completitudinii și diversității atributelor sale. Cu toate acestea, datele brute necesită preprocesare pentru a asigura consistența și integritatea acestora. Acest proces include tratarea valorilor lipsă, eliminarea sau corectarea erorilor și transformarea datelor categorice în formate numerice adecvate.

---

<sup>2</sup><https://www.kaggle.com/datasets/altruistdelhite04/loan-prediction-problem-dataset>, accesat: 14 noiembrie 2023 [Online]

#### 4.9.3. Descrierea atributelor

Setul de date include următoarele atribute relevante pentru evaluarea riscului de credit:

- **Gender (Gen):** Analizarea genului solicitantului poate oferi informații despre potențiale diferențe de comportament în ceea ce privește gestionarea finanțelor și accesul la credite. Deși genul nu ar trebui să influențeze decizia de acordare a creditului, este util pentru a identifica eventuale disparități de gen în accesul la finanțare[60].
- **Married (Căsătorit):** Statutul marital al solicitantului poate indica stabilitatea financiară. Persoanele căsătorite pot avea un sprijin financiar suplimentar și pot prezenta un risc mai scăzut de neplată comparativ cu cele necăsătorite<sup>3</sup>.
- **Education (Educație):** Nivelul de educație este un indicator important al veniturilor și stabilității financiare. Persoanele cu un nivel de educație mai ridicat tend să aibă venituri mai mari și mai stabile, reducând astfel riscul de neplată a creditului.
- **Self Employed (Antreprenor):** Antreprenorii pot avea venituri variabile, ceea ce poate crește riscul de neplată.
- **Property Area (Zona Proprietății):** Zona în care se află proprietatea poate afecta valoarea acesteia și riscul colateral. Proprietățile din zonele urbane sau semiurbane tend să aibă valori mai mari și o lichiditate mai bună în caz de execuție silită.
- **Loan Status (Statutul Împrumutului):** Statutul împrumutului indică dacă un împrumut a fost aprobat sau respins. Acest atribut este esențial pentru dezvoltarea modelului predictiv, deoarece reprezintă variabila țintă.
- **Loan ID (ID Împrumut):** Identifierul unic pentru fiecare împrumut este folosit pentru a diferenția între înregistrări, asigurând integritatea datelor și evitarea dupliilor.
- **Dependents (Dependenți):** Numărul de persoane aflate în întreținerea solicitantului poate influența cheltuielile lunare și capacitatea de rambursare.
- **ApplicantIncome (Venitul Solicitantului):** Venitul lunar al solicitantului este un parametru esențial pentru evaluarea capacitații de rambursare a creditului. Veniturile stabile și suficiente indică o capacitate mai mare de a face față obligațiilor financiare.
- **CoapplicantIncome (Venitul Cosolicitantului):** Venitul lunar al cosolicitantului poate îmbunătăți capacitatea totală de rambursare a creditului, oferind un sprijin financiar suplimentar în cazul în care solicitantul principal întâmpină dificultăți.
- **LoanAmount (Suma Împrumutată):** Suma solicitată pentru împrumut este direct legată de capacitatea de rambursare și de riscul asumat de creditor. Sumele mai mari necesită o analiză mai detaliată a capacitații financiare a solicitantului.
- **Loan Amount Term (Termenul Împrumutului):** Durata împrumutului influențează mărimea plăților lunare. Un termen mai lung poate reduce plățile lunare, dar poate crește suma totală platită în dobânzi, afectând riscul de neplată.
- **Credit History (Istoric de Credit):** Istoricul de credit al solicitantului este un indicator foarte important al riscului de neplată. Un istoric de credit bun sugerează o probabilitate mai mare de rambursare la timp a împrumutului[61].

---

<sup>3</sup><https://refi.com/how-marital-relationship-status-affects-mortgage/>, accesat: 18 noiembrie 2023 [online]

## Capitolul 5. Proiectare de detaliu și implementare

Acest capitol cuprinde etapele de proiectare și implementare ale proiectului, oferind o imagine clară a etapelor de dezvoltare și a soluțiilor tehnice utilizate. Bazându-ne pe fundația menționată în capitolele anterioare, unde am stabilit scopul și obiectivele, am studiat diverse cercetări și am analizat aspectele teoretice necesare, vom detalia modul în care aceste concepte au fost puse în practică. Vom prezenta în mod structurat proiectarea aplicației, descrierea componentelor și procesul de implementare, evidențiind funcționalitățile cheie și metodologiile utilizate.

### 5.1. Definirea problemei și a soluției alese

În creditarea ipotecară modernă, evaluarea riscului și optimizarea procesului de acordare a creditelor sunt provocări majore. Sistemul tradițional este complex și adesea îndelungat, descurajând debitorii și generând ineficiențe. De asemenea, utilizatorii întâmpină dificultăți în gestionarea informațiilor financiare, planificarea și monitorizarea economiilor și plășilor.

Aplicația propusă reduce timpul de procesare prin utilizarea unui algoritm de învățare automată (Machine Learning - ML) pentru prezicerea aprobării unui împrumut. De asemenea, aceasta oferă utilizatorilor instrumente practice pentru gestionarea finanșelor, inclusiv calculatoare de rambursare și economii, contribuind astfel la o mai bună planificare și control financiar.

### 5.2. Arhitectura aplicației

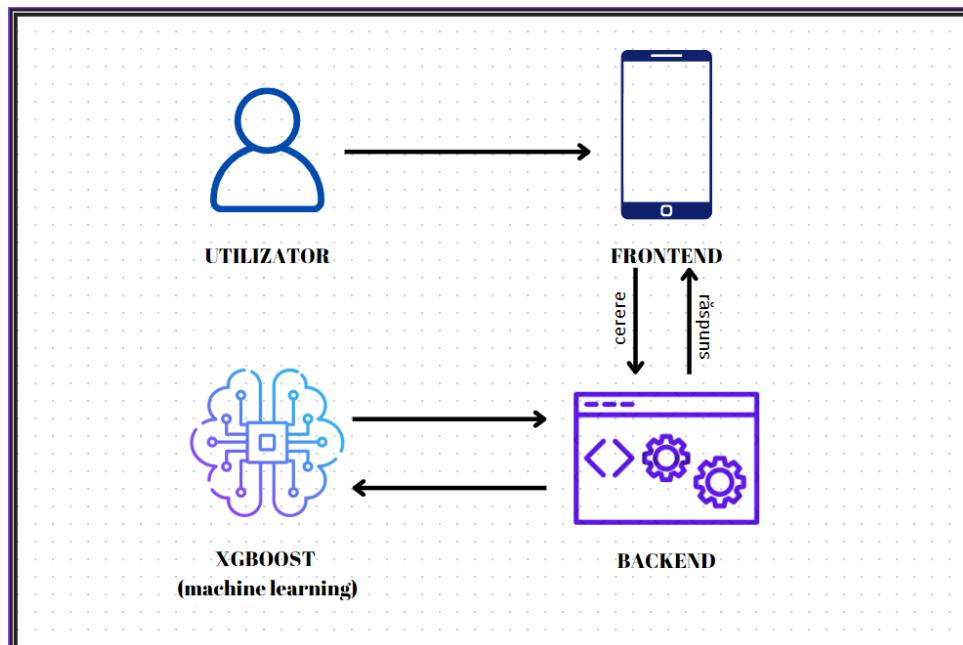


Figura 5.1: Diagrama de arhitectură a aplicației

Diagrama de arhitectură din figura 5.1 a aplicației ilustrează interacțiunea dintre utilizator, frontend, backend și algoritmul de învățare XGBoost, evidențiind fluxul de date și interacțiunea între componente. Această diagramă arată că utilizatorul interacționează cu frontend-ul aplicației Android, care trimite datele către backend-ul Python pentru procesare. Backend-ul utilizează algoritmul de învățare automată XGBoost pentru a efectua predicțiile și returnează rezultatele către frontend pentru a fi afișate utilizatorului.

#### 5.2.1. Utilizator

Utilizatorul interacționează cu aplicația Android prin intermediul unei interfețe grafice prietenoase. Aceasta introduce datele necesare pentru evaluarea riscului de credit, precum veniturile, educația și alte informații relevante.

#### 5.2.2. Frontend (Android)

Frontend-ul aplicației Android este responsabil pentru colectarea datelor de la utilizator și trimiterea acestora către backend. De asemenea, frontend-ul primește și afișează rezultatele predicțiilor realizate de backend. Comunicarea între frontend și backend se realizează prin cereri HTTP POST, folosind JSON pentru schimbul de date. Componente principale ale frontend-ului sunt:

- **PredictionActivity:** Permite utilizatorilor să introducă datele necesare pentru evaluarea riscului de credit și să vizualizeze predicțiile privind aprobarea creditului.
- **LoanCalculatorActivity:** Permite utilizatorilor să calculeze plățile lunare pentru împrumuturi ipotecare, folosind datele introduse de aceștia.
- **SavingsActivity:** Permite utilizatorilor să își planifice economiile și să estimeze valoarea viitoare a acestora.
- **OptimalRepaymentActivity:** Oferă utilizatorilor un calculator pentru a determina contribuția lunară pe care trebuie să o facă în contul de economii în fiecare lună pentru acoperirea dobânzii împrumutului la finalul acesteia.

#### 5.2.3. Backend (Python)

Backend-ul aplicației este implementat folosind Python și Flask. Acesta primește cererile de la frontend, procesează datele și utilizează modelul ML XGBoost pentru a efectua predicțiile. Backend-ul este responsabil pentru gestionarea logicii aplicației și a fluxului de date între componente. Codul backend-ului include endpoint-ul de predicție care primește datele de la frontend, le preprocesează și returnează rezultatele predicțiilor.

#### 5.2.4. Algoritmul ML XGBoost

Algoritmul ML XGBoost este utilizat pentru a analiza datele utilizatorului și pentru a realiza predicțiile privind aprobarea creditelor ipotecare. Modelul de învățare automată a fost antrenat pe un set de date, permitându-i să facă estimări precise pe baza noilor date furnizate de utilizator prin intermediul frontend-ului.

#### 5.2.5. Fluxul de date între Front-end și Back-end

- **Trimiterea datelor:** Aplicația Android trimite datele utilizatorilor (de exemplu venitul, suma împrumutului, termenul împrumutului) către backend sub formă de cerere POST în format JSON.
- **Procesarea datelor:** Backend-ul primește cererea, preprocesează datele și utilizează modelul XGBoost pentru a face o predicție asupra aprobării creditului.

- Returnarea rezultatului:** Backend-ul trimite rezultatul predicției înapoi la aplicația Android în format JSON. Aceste rezultate sunt apoi afișate utilizatorilor în interfața aplicației.

Arhitectura aplicației este concepută pentru a oferi o experiență simplă și intuitivă utilizatorilor, utilizând puterea algoritmilor de învățare automată pentru a evalua riscul de credit și a furniza instrumente financiare utile. Această structură modulară asigură flexibilitate și scalabilitate, permitând extinderea și îmbunătățirea funcționalităților aplicației în viitor.

### 5.3. Analizarea setului de date

Setul de date utilizat în acest proiect conține 612 înregistrări și 13 atribute. Aceste atribute sunt utilizate pentru a evalua eligibilitatea solicitării unui credit ipotecar. În continuare, vom descrie fiecare atribut și valorile pe care le poate lua.

Denumire	Descriere
<b>Loan_ID</b>	Identifier unic pentru fiecare împrumut
<b>Gender</b>	Genul solicitantului - cuprinde valori categorice: Male, Female
<b>Married</b>	Statutul marital - format din valori categorice: Yes, No
<b>Dependents</b>	Numărul de persoane aflate în întreținerea solicitantului - valori categorice: 0, 1, 2, 3+
<b>Education</b>	Nivelul de educație - cuprinde valori categorice: Graduate, Not Graduate
<b>Self_Employed</b>	Statutul de angajare - format din valori categorice: Yes, No
<b>ApplicantIncome</b>	Venitul lunar al solicitantului - valori numerice în euro
<b>CoapplicantIncome</b>	Venitul lunar al cosolicitantului - valori numerice în euro
<b>LoanAmount</b>	Suma solicitată pentru împrumut - valori numerice în mii de euro
<b>Loan_Amount_Term</b>	Durata împrumutului - cuprinde valori numerice în luni: 360, 120, 240, etc.
<b>Credit_History</b>	Istoricul de credit - format din valori numerice: 0 (rău), 1 (bun)
<b>Property_Area</b>	Zona proprietății - cuprinde valori categorice: Rural, Semiurban, Urban
<b>Loan_Status</b>	Statutul împrumutului - format din valori categorice: Y (aprobat), N (respins)

Tabela 5.1: Descrierea atributelor din setul de date

Cu ajutorul acestui set de date, aplicația are capacitatea de a analiza numeroase variabile și de a identifica modele relevante care pot influența aprobaarea unui împrumut ipotecar. Prin aplicarea unui algoritm de învățare automată asupra acestor date, aplicația va putea să determine corelațiile și tendințele semnificative, permitând astfel predicția cu acuratețe a statutului de aprobaare a împrumutului.

## 5.4. Pregătirea datelor și corelația dintre ele

Setul de date utilizat pentru acest proiect conține inițial 612 înregistrări și 13 attribute. Pentru a asigura integritatea și calitatea datelor necesare pentru antrenarea modelului de învățare automată, este necesară parcurgerea mai multor etape de pre-procesare. Aceste etape includ curățarea datelor, prelucrarea acestora și transformarea atributelor categorice în attribute numerice, după cum se vede și în figura 5.2.

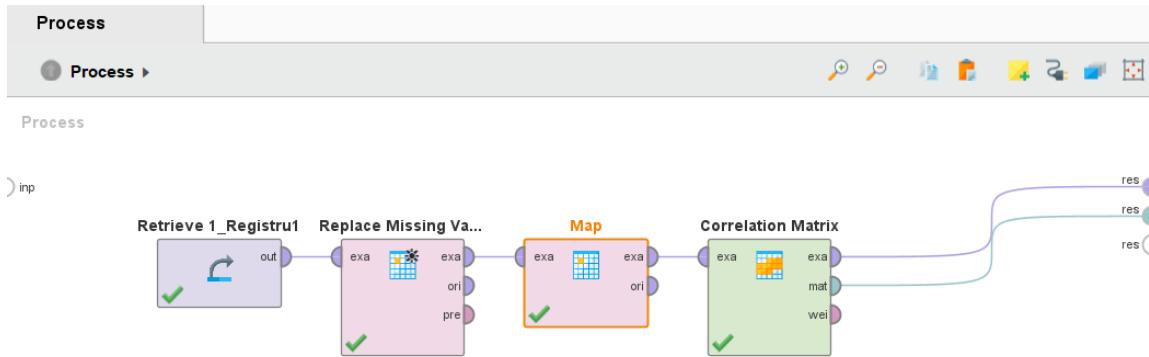


Figura 5.2: Pregătirea datelor în RapidMiner

### 5.4.1. Curățarea datelor

În primul rând, setul de date a fost analizat pentru a identifica valorile lipsă și cele incorecte. Din cele 612 înregistrări, anumite attribute aveau valori lipsă care trebuiau tratate pentru a nu afecta performanța modelului de învățare automată. Atributele cu valori lipsă și numărul de valori lipsă pentru fiecare sunt următoarele:

- **Gender:** 13 valori lipsă
- **Married:** 3 valori lipsă
- **Dependents:** 15 valori lipsă
- **Self\_Employed:** 32 valori lipsă
- **LoanAmount:** 22 valori lipsă
- **Loan\_Amount\_Term:** 14 valori lipsă
- **Credit\_History:** 50 valori lipsă

Pentru a gestiona valorile lipsă, s-a folosit operatorul *Replace Missing Values* din RapidMiner. Valorile lipsă au fost înlocuite fie cu mediana, fie cu modulul valorilor corespunzătoare fiecărui atribut. De exemplu, valorile lipsă din coloana *LoanAmount* au fost înlocuite cu mediana valorilor existente.

### 5.4.2. Prelucrarea datelor

După curățarea datelor, a fost necesară prelucrarea acestora pentru a le aduce într-un format adecvat modelului de învățare automată. Aceasta etapă este esențială pentru a asigura că modelul poate învăța eficient și pentru a îmbunătăți performanța predictiilor.

Atributele categorice au fost transformate în attribute numerice utilizând operatorul *Map*. Aceasta transformare este necesară deoarece majoritatea algoritmilor de învățare

automată nu pot lucra direct cu date categorice. Astfel, valorile textuale precum *Gender*, *Married*, *Education*, *Self\_Employed* și *Property\_Area* au fost convertite în valori numerice.

De exemplu, valorile *Male* și *Female* din coloana *Gender* au fost transformate în 0 și 1, respectiv, iar valorile *Urban*, *Rural* și *Semiurban* din coloana *Property\_Area* au fost transformate în 0, 1 și 2.

#### 5.4.3. Etichetarea datelor

Pentru a pregăti datele pentru antrenarea modelului, am etichetat coloana *Loan\_Status* cu *label*. Acest pas a fost esențial pentru a defini variabila țintă care va fi utilizată în procesul de antrenare a modelului.

#### 5.4.4. Matricea de corelație

Pentru a înțelege mai bine relațiile dintre diferitele atribute din setul de date, a fost generată o matrice de corelație. Matricea de corelație este un instrument statistic care măsoară gradul în care două variabile sunt lineare corelate. Aceasta ajută la identificarea legăturilor dintre variabile și poate indica care dintre acestea sunt mai relevante pentru modelul de predicție.

Attributes	Gender	Married	Educati...	Self_E...	Propert...	Loan_St...	Loan_ID	Depend...	Applicant...	Coappli...	LoanAm...	Loan_A...	Credit_...
Gender	1	-0.364	-0.046	-0.000	?	0.017	?	?	-0.060	-0.084	-0.109	0.072	0.009
Married	-0.364	1	0.014	0.005	?	-0.090	?	?	0.053	0.077	0.149	-0.099	-0.010
Education	-0.046	0.014	1	-0.011	?	0.085	?	?	-0.142	-0.063	-0.168	-0.080	0.073
Self_Emp...	-0.000	0.005	-0.011	1	?	0.003	?	?	0.127	-0.017	0.114	-0.036	0.001
Property_...	?	?	?	?	1	?	?	?	?	?	?	?	?
Loan_Sta...	0.017	-0.090	0.085	0.003	?	1	?	?	0.004	0.058	0.035	0.019	0.540
Loan_ID	?	?	?	?	?	?	1	?	?	?	?	?	?
Depende...	?	?	?	?	?	?	?	1	?	?	?	?	?
Applicantl...	-0.060	0.053	-0.142	0.127	?	0.004	?	?	1	-0.117	0.565	-0.047	0.018
Coapplica...	-0.084	0.077	-0.063	-0.017	?	0.058	?	?	-0.117	1	0.187	-0.061	-0.012
LoanAmmo...	-0.109	0.149	-0.168	0.114	?	0.035	?	?	0.565	0.187	1	0.036	0.000
Loan_Am...	0.072	-0.099	-0.080	-0.036	?	0.019	?	?	-0.047	-0.061	0.036	1	-0.002
Credit_Hi...	0.009	-0.010	0.073	0.001	?	0.540	?	?	0.018	-0.012	0.000	-0.002	1

Figura 5.3: Matricea de corelație a atributelor din setul de date

În figura 5.3 este prezentată matricea de corelație pentru setul de date utilizat. Fiecare celulă din matrice reprezintă coeficientul de corelație dintre două variabile. Valoarea coeficientului variază între -1 și 1, unde:

- **1** sugerează o corelație pozitivă perfectă.
- **-1** sugerează o corelație negativă perfectă.
- **0** sugerează lipsa corelației.

Analizând matricea de corelație, putem observa că anumite atribute au corelații notabile cu variabila țintă *Loan\_Status*, care indică dacă un împrumut a fost aprobat sau respins. Dintre acestea, cele mai semnificative sunt:

- **Credit\_History:** Acest atribut are cea mai puternică corelație pozitivă cu *Loan\_Status*. O valoare pozitivă a corelației sugerează că un istoric de credit bun (indicat de o valoare mare) este asociat cu aprobatarea împrumutului.
- **LoanAmount:** Există o corelație negativă moderată între suma împrumutată și *Loan\_Status*, indicând că împrumuturile mai mari au o probabilitate mai mică de a

fi aprobate.

- **ApplicantIncome:** Acest atribut are o corelație pozitivă moderată cu *Loan\_Status*, sugerând că venitul solicitantului influențează pozitiv aprobarea împrumutului.

Alte atrbute, cum ar fi *CoapplicantIncome*, *Dependents*, *Education* și *Married*, au corelații mai slabe cu *Loan\_Status*, ceea ce înseamnă că impactul lor asupra deciziei de aprobare a împrumutului este mai redus.

Prin utilizarea matricei de corelație, se pot identifica rapid relațiile importante dintre variabile și se poate determina care dintre acestea merită o atenție specială în procesul de dezvoltare a modelului de predicție. Acest pas este esențial pentru a îmbunătăți acuratețea și robustețea modelului, permitându-ne să facem predicții mai precise privind aprobarea împrumuturilor.

Astfel, prin preprocesarea și analizarea setului de date, am pregătit datele necesare pentru antrenarea modelului de învățare automată, asigurându-ne că sunt curate, relevante și adecvate pentru a furniza predicții precise.

## 5.5. Antrenarea modelelor

Pentru a antrena modelul de predicție a statutului împrumutului ipotecar, a fost necesar să eliminăm coloanele nefolositoare și să selectăm algoritmii de învățare automată potriviti. În figura 5.4, este prezentat procesul de eliminare a coloanei *Loan\_ID* și antrenare a modelului realizat în RapidMiner.

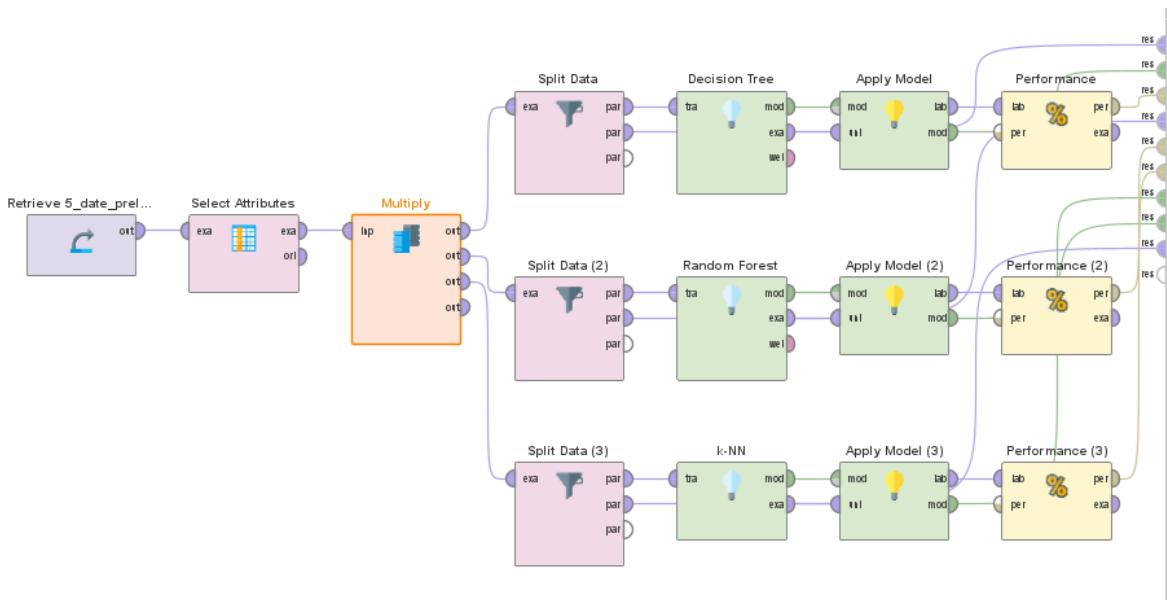


Figura 5.4: Algoritmii utilizați în RapidMiner

### 5.5.1. Eliminarea atributelor nefolositoare și împărțirea setului de date

Atributul *Loan\_ID* a fost eliminat utilizând operatorul *Select Attributes*, deoarece nu aduce informații relevante pentru prediciția statutului împrumutului.

Setul de date a fost împărțit în două subseturi: 80% pentru antrenament și 20% pentru testare, utilizând operatorul *Split Data*. Această împărțire a fost realizată pentru a evalua performanța modelului pe un set de date necunoscut în timpul antrenamentului.

### 5.5.2. Antrenarea modelelor

Pentru antrenarea modelului de predicție, au fost selectați patru algoritmi de învățare automată: *Decision Tree*, *Random Forest* și *k-Nearest Neighbors (k-NN)* în RapidMiner, și *XGBoost* în Python. Fiecare algoritm a fost aplicat pe datele de antrenament, iar performanța acestora a fost evaluată pe setul de date de testare.

## 5.6. Comparație între algoritmii de învățare automată

În RapidMiner am folosit trei algoritmi de învățare automată pentru a antrena modelul: *Decision Tree*, *Random Forest* și *k-Nearest Neighbors (k-NN)*. Fiecare algoritm a fost aplicat pe datele de antrenament, iar performanța a fost evaluată pe datele de testare. Rezultatele obținute sunt prezentate în Tabelul 5.2.

Algoritm	Acuratețe (%)
Decision Tree	78.69
Random Forest	77.87
k-Nearest Neighbors (k-NN)	68.85

Tabela 5.2: Comparația performanței algoritmilor de învățare automată din RapidMiner

Rezultatele din Tabelul 5.2 arată că modelul *Decision Tree* a obținut o acuratețe de 78.69%, modelul *Random Forest* a obținut o acuratețe de 77.87%, iar modelul *k-Nearest Neighbors* a obținut o acuratețe de 68.85%. Din aceste rezultate se poate observa că dintre algoritmii implementați în RapidMiner, *Decision Tree* este cel mai performant pentru problema noastră.

După antrenarea modelelor în RapidMiner, a fost implementat algoritmul *XGBoost* în Python, deoarece acest algoritm de învățare automată nu este disponibil în RapidMiner. Algoritmul *XGBoost* (Extreme Gradient Boosting) este cunoscut pentru performanța sa ridicată în competițiile de machine learning, fiind eficient în gestionarea datelor eterogene și a seturilor de date complexe.

Pentru implementarea *XGBoost*, setul de date preprocessat a fost utilizat și s-au aplicat tehnici de optimizare a hiperparametrilor pentru a identifica cei mai buni parametri pentru model. A fost folosită metoda *GridSearchCV*, o tehnică de validare încrucișată care permite explorarea combinatorică a diferenților hiperparametri și selectarea celor mai performanți. Hiperparametrii considerați pentru optimizare au fost *learning\_rate*, *max\_depth* și *n\_estimators*.

```
# Definirea grilei de hiperparametri pentru XGBoost
hyperparameters = {
    'n_estimators': [30, 50, 100],
    'max_depth': [3, 5, 7],
    'learning_rate': [0.01, 0.05, 0.1]
}
```

Figura 5.5: Definirea grilei de hiperparametri pentru XGBoost

În figura 5.5 se pot observa valorile selectate pentru hiperparametri:

- *n\_estimators*: [30, 50, 100]
- *max\_depth*: [3, 5, 7]
- *learning\_rate*: [0.01, 0.05, 0.1]

```
Run train_model x
C:\Users\popgi\PROJECTPython\venv\Scripts\python.exe C:\Users\popgi\PROJECTPython\train_model.py
Best Parameters: {'learning_rate': 0.01, 'max_depth': 5, 'n_estimators': 100}
Acuratețe cu cei mai buni hiperparametri: 81.30081300813008
Process finished with exit code 0
```

Figura 5.6: Performanța modelului XGBoost

După rularea *GridSearchCV*, au fost identificați cei mai buni hiperparametri pentru modelul *XGBoost*, prezentați în figura 5.6.

- *learning\_rate*: 0.01
- *max\_depth*: 5
- *n\_estimators*: 100

Acești parametri au fost utilizati pentru antrenarea finală a modelului *XGBoost*. Rezultatele obținute după antrenarea modelului *XGBoost* sunt prezentate, de asemenea, în figura 5.6. Acestea arată că modelul *XGBoost* a obținut o acuratețe de 81.30%, superioară acurateții obținute de modelul *Decision Tree*.

Astfel, implementarea algoritmului *XGBoost* în Python a demonstrat că acest model oferă o performanță superioară comparativ cu modelele antrenate anterior în Rapid-Miner. Utilizarea *XGBoost* s-a dovedit a fi eficientă în prediciția aprobării împrumuturilor ipotecare, oferind o acuratețe îmbunătățită și o mai bună capacitate de generalizare.

## 5.7. Implementarea aplicației mobile

Aplicația mobilă este dezvoltată utilizând limbajul Kotlin în Android Studio. Aceasta a fost concepută folosind arhitectura MVC (Model-View-Controller). Această arhitectură separă logica de business (Model), interfața utilizator (View) și controlul fluxului de aplicație (Controller), facilitând astfel dezvoltarea și întreținerea codului.

### 5.7.1. Arhitectura MVC

- **Model:** Reprezintă datele aplicației și logica de business. În acest proiect, modelul include clasele care gestionează datele despre împrumuturile ipotecare și datele despre utilizatori.
- **View:** Reprezintă interfața utilizatorului și modul în care sunt prezentate datele. Aplicația Android are diverse activități și fragmente care afișează formularele de intrare a datelor, rezultatele predicțiilor și alte informații relevante.
- **Controller:** Controlează fluxul aplicației și interacțiunile dintre model și view. În Android, activitățile și fragmentele acționează ca controlere, gestionând evenimentele utilizatorilor și actualizând modelul și view-ul în consecință.

### 5.7.2. Cazuri de utilizare

Pentru a asigura o utilizare eficientă și intuitivă a aplicației mobile, este esențial să înțelegem și să definim cazurile de utilizare. Cazurile de utilizare reprezintă scenarii specifice în care utilizatorii interacționează cu sistemul pentru a-și atinge obiectivele. În contextul acestei aplicații, utilizatorii au diverse necesități legate de gestionarea împrumuturilor ipotecare și a economiilor personale.

Diagrama de cazuri de utilizare din Figura 5.7 oferă o reprezentare vizuală a principalelor interacțiuni dintre utilizator și aplicație. Aceasta ilustrează modul în care utilizatorii pot introduce datele financiare personale, cum pot afla dacă li se va acorda un împrumut ipotecar, cum pot calcula plata lunară a împrumutului și cum pot estima economiile viitoare. De asemenea, utilizatorii pot determina contribuția lunară necesară pentru a acoperi costurile împrumutului ipotecar prin economii.

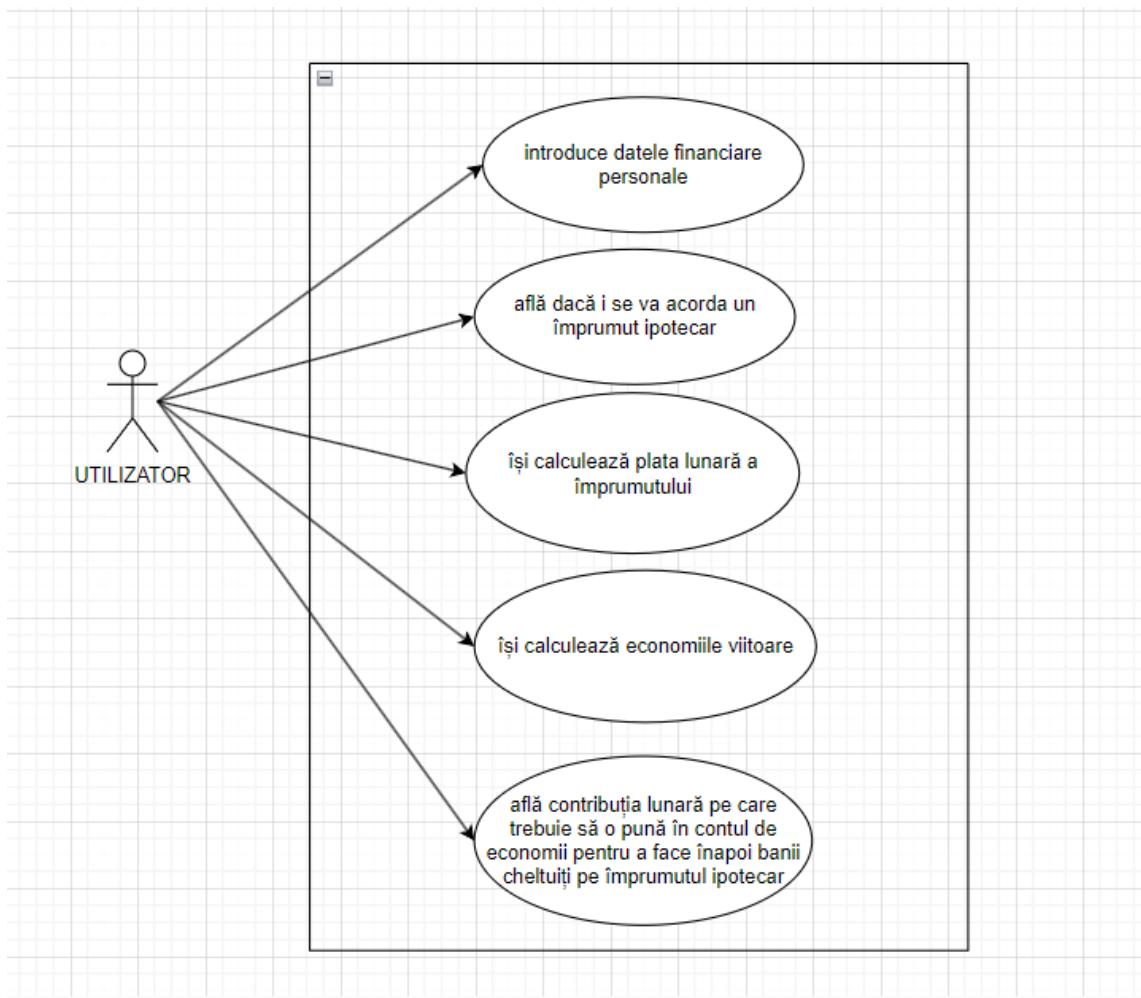


Figura 5.7: Cazuri de utilizare

Aceste cazuri de utilizare sunt necesare pentru a oferi utilizatorilor un set complet de instrumente pentru gestionarea eficientă a finanțelor personale și pentru a facilita procesul de solicitare a unui împrumut ipotecar.

### 5.7.3. Designul și implementarea paginilor principale

Paginile principale ale aplicației includ activități pentru predicția împrumuturilor, calculul împrumuturilor, planificarea economiilor și determinarea rambursării optime. Fiecare activitate a fost proiectată pentru a oferi o interfață utilizator intuitivă și pentru a asigura funcționalitatea corectă a aplicației.

#### **PredictionActivity**

PredictionActivity permite utilizatorilor să introducă datele necesare pentru evaluarea riscului de credit și să vizualizeze predicțiile privind aprobarea creditului. Această activitate include mai multe funcționalități cheie:

- **Formular de introducere a datelor:** Am utilizat diverse componente de interfață Android precum `EditText` pentru introducerea veniturilor și sumei împrumutului și `RadioButton` pentru selectarea genului și statutului marital. Acestea permit utilizatorilor să introducă informațiile necesare rapid și eficient.
- **Validarea datelor:** Înainte de a trimite datele către backend, am implementat validate pentru a asigura că toate câmpurile sunt completate corect. Aceasta ajută la prevenirea erorilor și la îmbunătățirea calității datelor.
- **Trimiterea datelor către backend:** Datele validate sunt trimise către backend-ul implementat în Python folosind cereri HTTP POST în format JSON. Backend-ul procesează aceste date și returnează predicțiile privind aprobarea creditului.

Designul acestor pagini este centrat pe utilizator, asigurându-se că toate informațiile necesare sunt introduse într-un mod organizat și clar, cum se poate vedea și în figura 5.8.

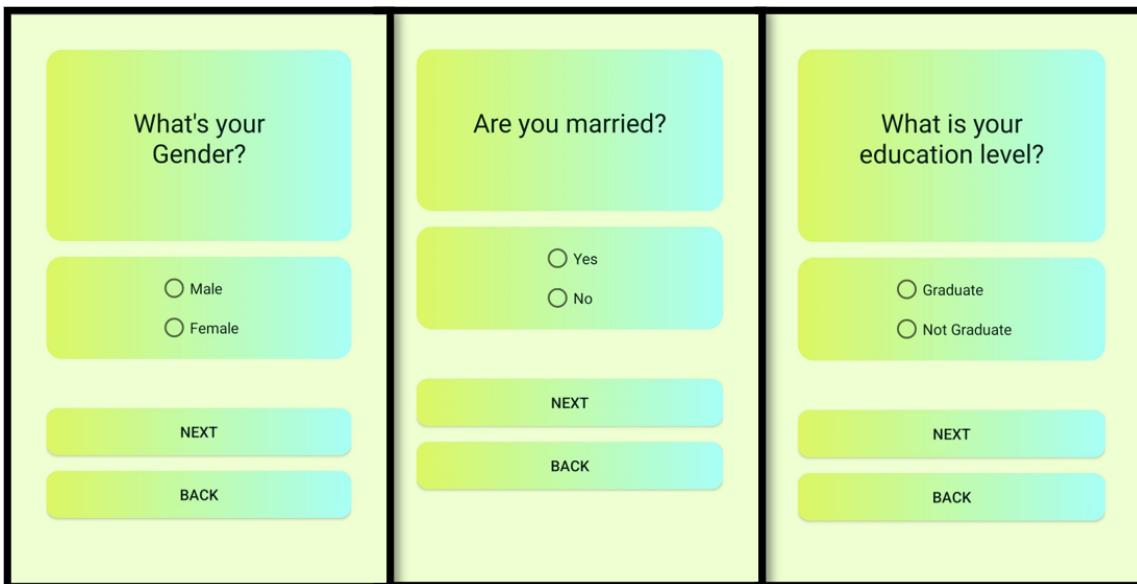


Figura 5.8: Pagini despre datele personale ale utilizatorului

#### **LoanCalculatorActivity**

LoanCalculatorActivity permite utilizatorilor să calculeze plățile lunare fixe sau variabile pentru împrumuturi ipotecare, oferindu-le astfel flexibilitate și opțiuni clare de gestionare a finanțelor personale. Această activitate include următoarele funcționalități:

- **Selectarea parametrilor de împrumut:** Am utilizat componente precum SeekBar pentru a permite utilizatorilor să selecteze rata dobânzii împrumutului. Această componentă oferă o modalitate interactivă de a ajusta parametrii și de a vedea instantaneu efectele asupra plății lunare.
- **Afișarea rezultatelor:** Rezultatele calculelor pentru plățiile împrumutului sunt afișate folosind componente TextView, care sunt actualizate după ce utilizatorul apasă pe butonul *CALCULATE*.
- **Calcularea plăților lunare:** Am implementat logica de calcul folosind formule financiare standard pentru a determina plățile lunare și am integrat aceste calcule în codul activității.

Designul acestei pagini este conceput pentru a fi ușor de utilizat și pentru a furniza informații clare și precise, cum se poate vedea și în figura 5.9.

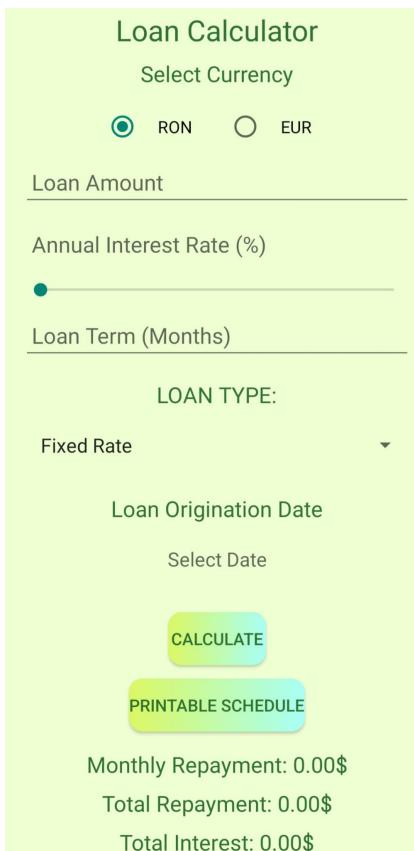


Figura 5.9: Interfața LoanCalculatorActivity

## SavingsActivity

SavingsActivity permite utilizatorilor să își planifice economiile și să estimeze valoarea viitoare a acestora. Această activitate include următoarele funcționalități:

- **Introducerea datelor economice:** Utilizatorii introduc date precum suma inițială, contribuția lunară și rata dobânzii utilizând componente EditText.
- **Calcularea economiilor viitoare:** Logica de calcul se bazează pe formula dobânzii compuse pe care am implementat-o în activitate.
- **Afișarea rezultatelor:** Rezultatele sunt afișate utilizatorului într-un format clar,

folosind componente `TextView`, care arată economiile viitoare și câștigurile din dobânzi.

Designul acestei pagini este realizat pentru a fi simplu și eficient, cum se poate vedea și în figura 5.10.



Figura 5.10: Interfața SavingsActivity

### OptimalRepaymentActivity

OptimalRepaymentActivity oferă utilizatorilor un calculator pentru a determina contribuția lunără necesară în contul de economii pentru acoperirea dobânzii împrumutului ipotecar. Această activitate include următoarele funcționalități:

- **Introducerea detaliilor împrumutului:** Utilizatorii introduc date precum suma împrumutului, rata dobânzii, durata împrumutului și economiile inițiale utilizând componente `EditText`.
- **Calcularea contribuției lunare optime:** Am implementat logica de calcul folosind formule financiare pentru a determina contribuția lunără necesară și am integrat aceste calcule în codul activității.
- **Afișarea rezultatelor:** Rezultatele sunt afișate utilizatorului folosind componente `TextView`, arătând contribuția lunără necesară și economiile finale.

Designul acestei pagini este orientat spre ușurința utilizării și claritatea informațiilor, cum se poate vedea și în figura 5.11.

Monthly Savings Contribution for Loan Interest Repayment

Select Currency

RON  EUR

Loan Amount

Annual Interest Rate (%)

Loan Term (Months)

Initial Savings

Annual Savings Interest Rate (%)

CALCULATE

Total Interest: 0.00

Monthly Contribution: 0.00

Figura 5.11: Interfața OptimalRepaymentActivity

#### 5.7.4. Interfața utilizatorului

Interfața utilizatorului este concepută pentru a fi intuitivă și ușor de utilizat. Fiecare activitate utilizează elemente standard de interfață, cum ar fi **RadioButton**, **Edit-Text**, **SeekBar** și **Button**, pentru a colecta datele de la utilizatori și a afișa rezultatele într-un mod clar și accesibil. Designul este realizat pentru a oferi o experiență plăcută utilizatorilor și pentru a facilita navigarea prin diferitele funcționalități ale aplicației.

#### 5.7.5. Navigarea în aplicație

Navigarea între activități este gestionată prin intenții (Intents) care permit transferul de date între diferitele ecrane ale aplicației. Fiecare activitate inițiază următoarea activitate pe baza acțiunilor utilizatorului și a datelor introduse, asigurând un flux logic și continuu al interacțiunilor utilizatorului cu aplicația.

Acest design modular și bine structurat asigură că aplicația este robustă, ușor de întreținut și extensibilă, permitând adăugarea de noi funcționalități în viitor.

## 5.8. Concluzie implementare

Prin implementarea acestor funcționalități, aplicația mobilă nu doar simplifică procesul de solicitare a unui credit ipotecar, ci și oferă utilizatorilor un set complet de instrumente pentru gestionarea eficientă a finanțelor personale.

# Capitolul 6. Testare și validare

Vom descrie testarea și validarea aplicației dezvoltate, împreună cu modelele utilizate. Scopul testării este de a asigura faptul că aplicația funcționează corect, îndeplinește cerințele specificate și oferă o experiență de utilizare satisfăcătoare.

## 6.1. Evaluarea algoritmilor

Performanța fiecărui model a fost evaluată utilizând mai multe metri, printre care acuratețea, precizia și recall-ul. În figurile 6.1, 6.2 și 6.3, sunt prezentate rezultatele evaluării pentru fiecare model.

### 6.1.1. Decision Tree

accuracy: 78.69%			
	true Y	true N	class precision
pred. Y	81	23	77.88%
pred. N	3	15	83.33%
class recall	96.43%	39.47%	

Figura 6.1: Performanța modelului Decision Tree

După cum se poate observa în figura 6.1, modelul *Decision Tree* a obținut o acuratețe de 78.69%. Matricea de confuzie arată următoarele:

- 81 de împrumuturi au fost corect clasificate ca fiind aprobate (true Y).
- 15 împrumuturi au fost corect clasificate ca fiind respinse (true N).
- 23 de împrumuturi respinse au fost clasificate incorrect ca fiind aprobate (false Y).
- 3 împrumuturi aprobate au fost clasificate incorrect ca fiind respinse (false N).

### 6.1.2. Random Forest

accuracy: 77.87%			
	true Y	true N	class precision
pred. Y	80	23	77.67%
pred. N	4	15	78.95%
class recall	95.24%	39.47%	

Figura 6.2: Performanța modelului Random Forest

După cum se poate observa în figura 6.2, modelul *Random Forest* a obținut o acuratețe de 77.87%. Matricea de confuzie arată următoarele:

- 80 de împrumuturi au fost corect clasificate ca fiind aprobate (true Y).

- 15 împrumuturi au fost corect clasificate ca fiind respinse (true N).
- 23 de împrumuturi respinse au fost clasificate incorrect ca fiind aprobată (false Y).
- 4 împrumuturi aprobată au fost clasificate incorrect ca fiind respinse (false N).

### 6.1.3. k-Nearest Neighbors (k-NN)

accuracy: 68.85%			
	true Y	true N	class precision
pred. Y	72	26	73.47%
pred. N	12	12	50.00%
class recall	85.71%	31.58%	

Figura 6.3: Performanța modelului k-Nearest Neighbors (k-NN)

După cum se poate observa în figura 6.3, modelul *k-Nearest Neighbors* a obținut o acuratețe de 68.85%. Matricea de confuzie arată următoarele:

- 72 de împrumuturi au fost corect clasificate ca fiind aprobată (true Y).
- 12 împrumuturi au fost corect clasificate ca fiind respinse (true N).
- 26 de împrumuturi respinse au fost clasificate incorrect ca fiind aprobată (false Y).
- 12 împrumuturi aprobată au fost clasificate incorrect ca fiind respinse (false N).

### 6.1.4. XGBoost

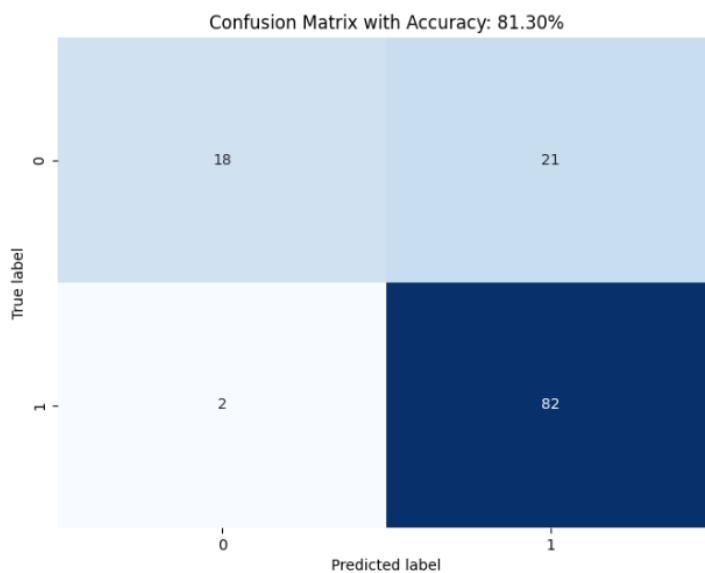


Figura 6.4: Performanța modelului XGBoost

După cum se poate observa în figura 6.4, modelul *XGBoost* a obținut o acuratețe de 81.30%, demonstrând că acest model oferă o performanță superioară comparativ cu modelele antrenate anterior în RapidMiner. Matricea de confuzie arată următoarele:

- 82 de împrumuturi au fost corect clasificate ca fiind aprobată (true Y).
- 18 împrumuturi au fost corect clasificate ca fiind respinse (true N).

- 21 de împrumuturi respinse au fost clasificate incorrect ca fiind aprobată (false Y).
- 2 împrumuturi aprobată au fost clasificate incorrect ca fiind respinse (false N).

## 6.2. Testare SonarQube

Pentru a asigura calitatea codului și a detecta eventualele probleme de securitate și mențenabilitate, am folosit SonarQube pentru analiza statică a codului sursă. SonarQube este un instrument de analiză a calității codului care ajută la identificarea problemelor precum bug-uri, vulnerabilități și code smells. Acesta efectuează o scanare detaliată a codului și oferă feedback automatizat asupra aspectelor critice, permitând dezvoltatorilor să corecteze rapid problemele și să îmbunătățească calitatea generală a codului.<sup>1</sup>

SonarQube integrează diverse reguli și practici de codare standardizate, oferind o evaluare cuprinzătoare a codului din perspectiva securității, fiabilității și mențenabilității. Prin utilizarea acestui instrument, ne-am asigurat că problemele comune de codare sunt abordate proactiv, iar codul nostru respectă standardele industriei.

În această secțiune, vom prezenta rezultatele obținute în urma analizei SonarQube, discutând despre problemele identificate și vom oferi sugestii pentru remedierea acestora. Analiza detaliată a fost esențială pentru a ne asigura că aplicația finală este robustă, securizată și ușor de întreținut.

După efectuarea analizei SonarQube, am obținut următoarele rezultate generale privind calitatea codului. Figura 6.5 prezintă o privire de ansamblu a rezultatelor analizei.

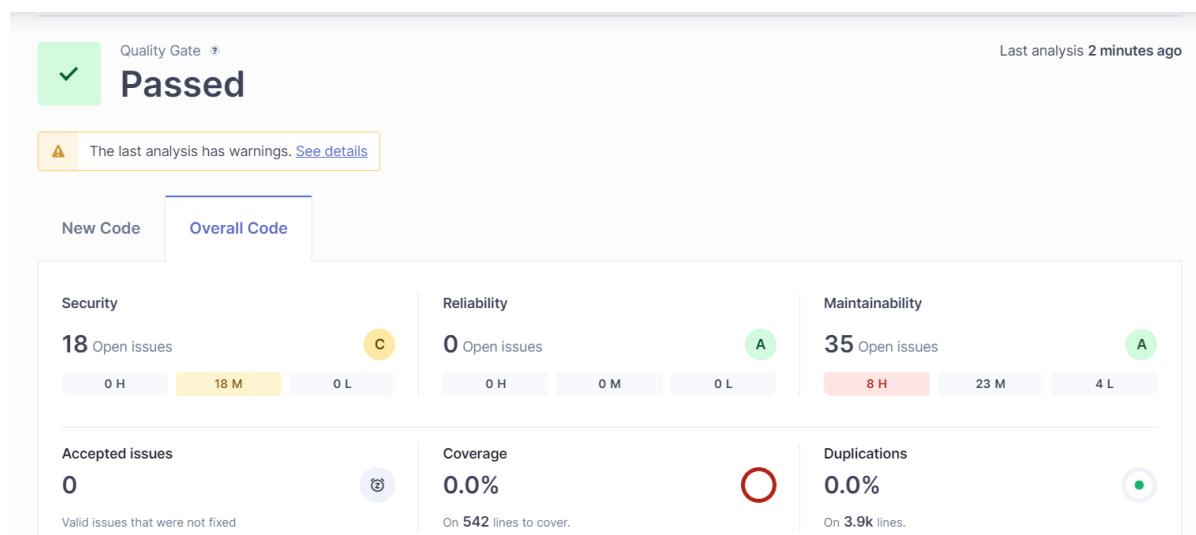


Figura 6.5: Rezultatele generale ale analizei SonarQube

Conform figurii 6.5, codul a trecut de pragul de calitate, însă au fost identificate mai multe probleme, inclusiv probleme de securitate și mențenabilitate. În continuare, vom detalia câteva dintre aceste probleme și cum le-am rezolvat.

### 6.2.1. Probleme de securitate

SonarQube a identificat mai multe probleme de securitate, printre care se numără și problema legată de permisiuni pe componente exportate. Figura 6.6 prezintă această problemă.

<sup>1</sup><https://www.sonarsource.com/>, accesat: 21 mai 2024 [Online]

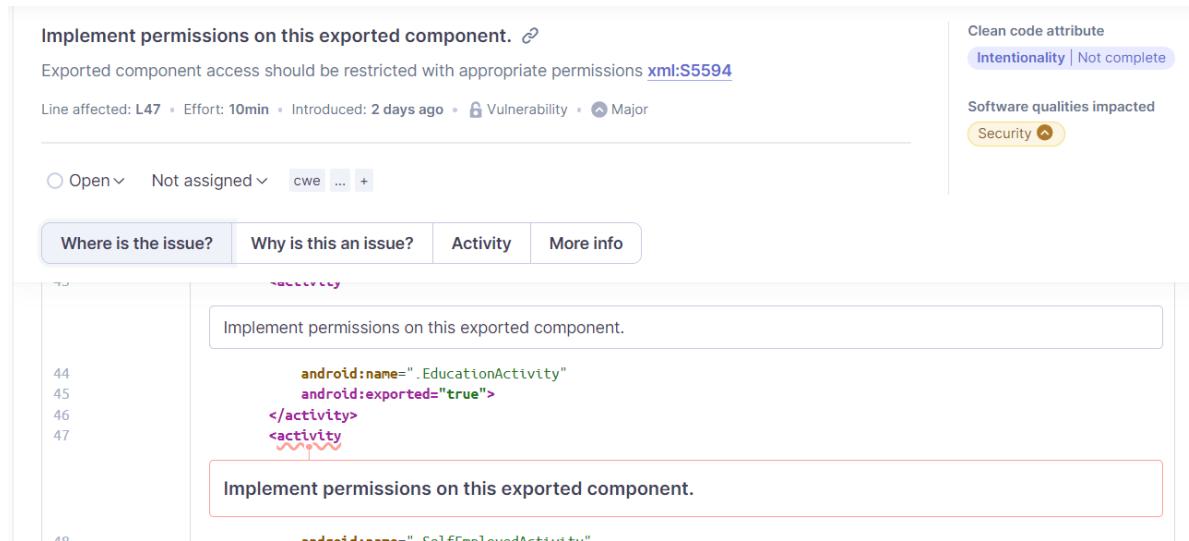


Figura 6.6: Implementarea permisiunilor pe componentelete exportate

*Implementare permisiunilor pe componentelete exportate:* Accesul la componentelete exportate trebuie să fie restrictionat prin permisiuni adecvate. Acest lucru este esențial pentru a preveni accesul neautorizat și a proteja datele sensibile. Pentru a rezolva această problemă, am adăugat permisiuni specifice pentru a limita accesul la componentelete exportate și am setat atributul *android:exported* la *false* pentru activitățile care nu trebuie accesate din afara aplicației.

### 6.2.2. Probleme de menținabilitate

SonarQube a identificat mai multe probleme de menținabilitate, inclusiv folosirea numerelor de versiune hardcodate. Figura 6.7 prezintă această problemă.

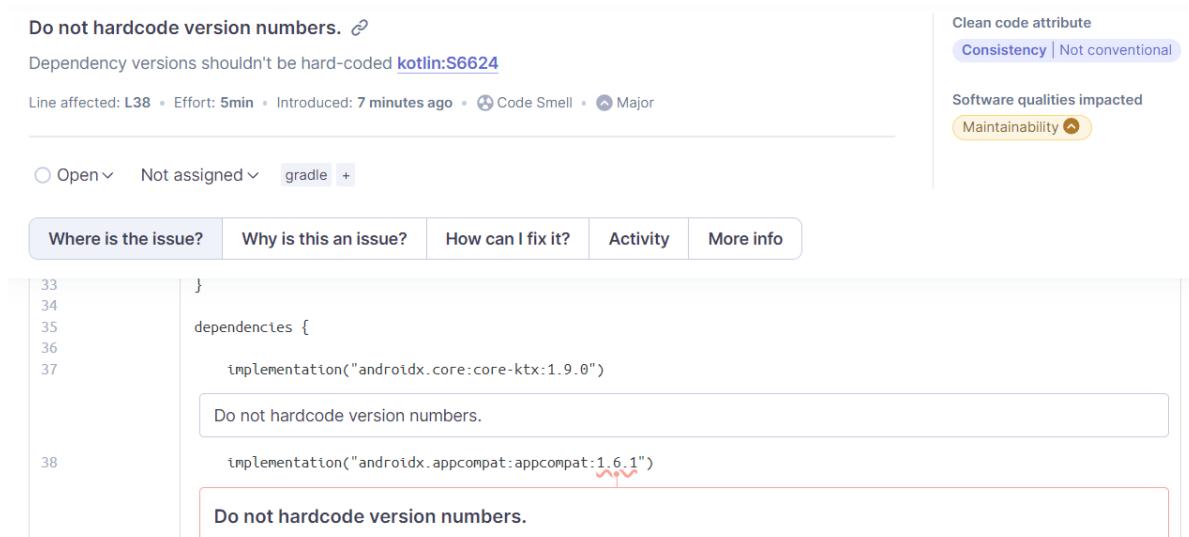


Figura 6.7: Problema de menținabilitate - Nu hardcodați numerele de versiune

*Nu hardcodați numerele de versiune:* Hardcodarea numerelor de versiune în cod este o practică nepotrivită care afectează menținabilitatea. Pentru a remedia această problemă, am definit variabile pentru versiunile dependențelor, facilitând astfel actualizarea

și menținerea consistenței pe întregul proiect.

SonarQube a identificat importuri neutilizate în cod. Figura 6.8 prezintă această problemă.

The screenshot shows a SonarQube analysis result for a Kotlin file named ApiService.kt. The code editor highlights the line 'import com.google.gson.annotations.SerializedName' with a red squiggle. A tooltip box appears over this line with the text 'Remove this unused import.' The top right corner of the interface shows 'Clean code attribute' and 'Intentionality | Not clear'. Below that, 'Software qualities impacted' shows 'Maintainability' as the selected option. The bottom of the interface has tabs for 'Where is the issue?', 'Why is this an issue?', 'How can I fix it?', 'Activity', and 'More info'.

```

1 package com.androidcodr.loan_prediction
2
3 import com.google.gson.annotations.SerializedName
4
5 import retrofit2.Call
6 import retrofit2.http.Body
7 import retrofit2.http.POST

```

Figura 6.8: Problema de menținabilitate - Eliminarea importurilor neutilizate

*Eliminarea importurilor neutilizate:* Importurile neutilizate în cod afectează menținabilitatea și pot duce la confuzii. Pentru a rezolva această problemă, am revizuit și am eliminat toate importurile care nu sunt utilizate în fișierele de cod. Acest proces ajută la menținerea codului curat și ușor de întreținut, reducând în același timp complexitatea acestuia.

SonarQube a identificat și o problemă legată de interfețele funcționale. Figura 6.9 prezintă această problemă.

The screenshot shows a SonarQube analysis result for a Kotlin file named ApiService.kt. The code editor highlights the line 'interface ApiService {' with a red squiggle. A tooltip box appears over this line with the text 'Make this interface functional or replace it with a function type.' The top right corner of the interface shows 'Clean code attribute' and 'Intentionality | Not clear'. Below that, 'Software qualities impacted' shows 'Maintainability' as the selected option. The bottom of the interface has tabs for 'Where is the issue?', 'Why is this an issue?', 'How can I fix it?', 'Activity', and 'More info'.

```

6 import retrofit2.http.POST
7
8 // Interfață pentru serviciul Retrofit
9 interface ApiService {
10
11     @POST("predict")
12     fun predictLoanStatus(@Body requestBody: RequestBody): Call<ApiResponse>
13 }

```

Figura 6.9: Problema de menținabilitate - Interfață funcțională

O interfață care conține o singură metodă abstractă ar trebui să fie marcată ca fiind funcțională pentru a clarifica intenția și a îmbunătăți lizibilitatea codului. Pentru a rezolva această problemă, am adăugat adnotarea `@FunctionalInterface` la interfață

menționată. Acest lucru asigură că interfața este tratată corespunzător și ajută la prevenirea adăugării accidentale de noi metode abstracte în viitor.

SonarQube a identificat și alte probleme de menținabilitate, cum ar fi parametrii lambda neutilizați în cod. Figura 6.10 prezintă această problemă.

The screenshot shows a SonarQube code review interface. At the top, there's a message: "Use "\_" instead of this unused lambda parameter "group". ⓘ". Below it, "Unused function parameters should be removed [kotlin:S1172](#)". The code affected is line 19: "radioGroupCreditHistory.setOnCheckedChangeListener { group, checkedId ->". A red callout box highlights this line with the text "Use "\_" instead of this unused lambda parameter "group".". The code continues with "val intent = Intent(this@CreditHistoryActivity, PredictionActivity::class.java)" and "intent.putExtras(intent.extras ?: Bundle()) // Preserve existing extras". On the right side, there are sections for "Clean code attribute" (Intentionality | Not clear) and "Software qualities impacted" (Maintainability).

Figura 6.10: Problema de menținabilitate - Parametri lambda neutilizați

*Utilizarea lui \_ pentru parametrii lambda neutilizați:* Parametrii lambda neutilizați în cod pot duce la confuzii și îngreunează menținabilitatea. Pentru a rezolva această problemă, am înlocuit parametrii neutilizați cu \_ în funcțiile lambda pentru a clarifica că aceștia nu sunt utilizati. Aceasta ajută la menținerea codului curat și la evitarea ambiguităților.

### 6.2.3. Probleme de consistență

SonarQube a identificat o problemă de consistență în denumirea variabilelor locale. Figura 6.11 prezintă această problemă.

The screenshot shows a SonarQube code review interface. At the top, there's a message: "Rename this local variable to match the regular expression ^`?[a-zA-Z][a-zA-Z0-9]\*`?\$ ⓘ". Below it, "Local variable and function parameter names should comply with a naming convention [kotlin:S117](#)". The code affected is line 59: "val M = loanAmount \* monthlyInterestRate / (1 - (1 + monthlyInterestRate).pow(-loanTerm))". A red callout box highlights this line with the text "Rename this local variable to match the regular expression ^`?[a-zA-Z][a-zA-Z0-9]\*`?\$". The code continues with "val totalInterest = M \* loanTerm - loanAmount". On the right side, there are sections for "Clean code attribute" (Consistency | Not identifiable) and "Software qualities impacted" (Maintainability).

Figura 6.11: Problema de consistență - Denumirea variabilelor locale

*Denumirea variabilelor locale:* Utilizarea convențiilor de denumire adecvate pentru variabilele locale este esențială pentru menținerea consistenței și lizibilității codului. Pentru a rezolva această problemă, am redenumit variabila *M* în *monthlyRepayment*, conform regulii de denumire care specifică utilizarea literelor mici și a unui nume descriptiv. Aceasta îmbunătățește claritatea și intenționalitatea codului.

#### 6.2.4. Rezultate finale

După implementarea modificărilor sugerate de SonarQube, am reușit să rezolvăm majoritatea problemelor identificate inițial. Numărul total de erori a scăzut de la 53 la 6, așa cum se poate observa în Figura 6.12.

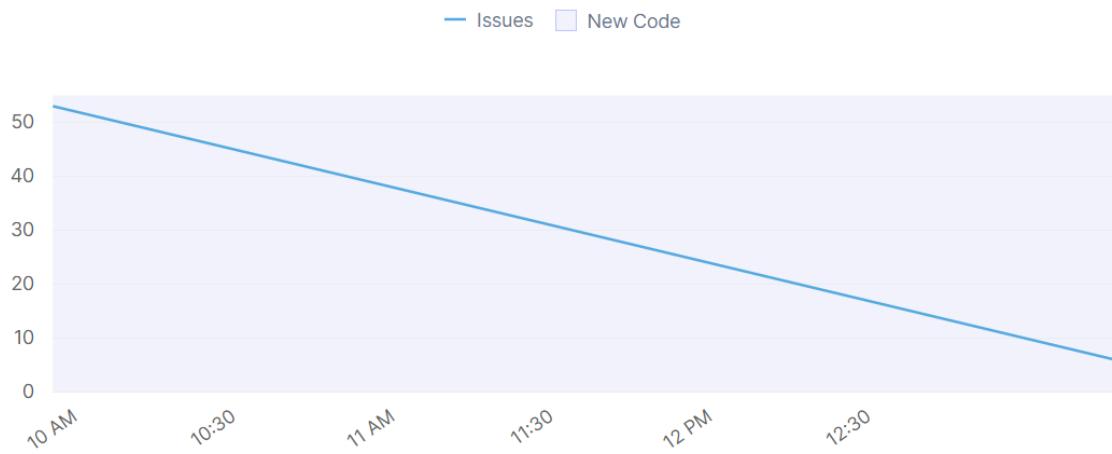


Figura 6.12: Evoluția numărului de erori în urma remedierii problemelor

Acest proces a fost esențial pentru a îmbunătăți calitatea codului nostru, asigurându-ne că aplicația finală este robustă, securizată și ușor de întreținut. Analiza detaliată și feedback-ul automatizat oferit de SonarQube ne-au permis să identificăm și să corectăm prompt problemele critice, contribuind astfel la dezvoltarea unui produs final de înaltă calitate.

### 6.3. Validare prin feedback-ul utilizatorilor

Pentru a evalua experiența utilizatorilor și a identifica posibilele îmbunătățiri, am realizat o fază de testare beta, în care un grup selectat de utilizatori reali au fost invitați să utilizeze aplicația și să ofere feedback. Această validare a implicat următorii pași:

#### 6.3.1. Selectarea utilizatorilor

Am selectat un grup divers de utilizatori beta, reprezentând diferite categorii demografice și niveluri de experiență cu tehnologia. Această diversitate a fost esențială pentru a obține o perspectivă cât mai completă asupra aplicației și a problemelor pe care utilizatorii le-ar putea întâlni.

Feedbackul primit de la cei 5 utilizatori a fost următorul:

##### 1. Utilizator 1:

“Mi-a plăcut aplicația! Funcția de predicție a împrumuturilor a fost foarte intuitivă. Am apreciat cât de ușor a fost să introduc datele mele personale

și să primesc o predicție clară asupra stării împrumutului meu. Procesul pas cu pas a asigurat că nu am ratat nicio informație importantă. Totuși, ar fi util să existe mai multe informații explicative pentru fiecare pas, astfel încât să înțeleg mai bine de ce sunt necesare anumite date.”

## 2. Utilizator 2:

“Aplicația este foarte folositoare. Am găsit activitatea de calcul a împrumutului extrem de utilă, iar calculatorul a oferit detalii precise despre ratele lunare și costul total al împrumutului. Interfața ar putea fi mai clară și să aibă un aspect mai plăcut.”

## 3. Utilizator 3:

“Am aflat cu ajutorul aplicației că aş avea șanse să primesc un împrumut. Îmi place ideea de a folosi o aplicație ca să afli acest lucru decât să stai la coziile de la bancă și după să afli că nu ești un candidat bun pentru ei. Îmi place și opțiunea de a genera un program de plată, ceea ce mă ajută să văd cum se va desfășura rambursarea împrumutului în timp, dar am întâmpinat o problemă deoarece nu am avut un buton de *back* ca să pot ieși din acest program.”

## 4. Utilizator 4:

“Mi-a plăcut aplicația deoarece am reușit să aflu rapid șansele mele de a primi un împrumut ipotecar. De asemenea am apreciat funcționalitatea aplicației de a lua în considerare diferite tipuri de împrumuturi și rate ale dobânzii. Aceasta face ca aplicația să fie versată și utilă pentru diferite scenarii financiare. Mi-a plăcut și partea de *Savings*.”

## 5. Utilizator 5:

“Aplicația e utilă dar s-ar mai putea adăuga lucruri. Secțiunea de economii este foarte informativă. Calculul câștigurilor brute și al impozitului pe venit reținut a fost clar și m-a ajutat să înțeleg mai bine cum să-mi gestionez economiile. Totuși, mi-ar plăcea să văd mai multe opțiuni care să mă ajute să îmi gestionez banii.”

### 6.3.2. Implementarea îmbunătățirilor

Pe baza analizei feedback-ului, am implementat următoarele îmbunătățiri:

- **Îmbunătățiri ale interfeței:** Ajustarea designului și navigării pentru a oferi o experiență mai intuitivă și cu un aspect mai plăcut.
- **Corectarea bug-urilor:** Rezolvarea problemelor raportate legate de funcționalitate și utilizare.

Acest proces de validare prin feedback-ul utilizatorilor a fost esențial pentru a asigura că aplicația dezvoltă oferă o experiență de utilizare satisfăcătoare și răspunde nevoilor utilizatorilor țintă. Prin implicarea activă a utilizatorilor în procesul de testare și îmbunătățire, am reușit să dezvoltăm un produs de înaltă calitate, care să fie bine primit pe piață.

# **Capitolul 7. Manual de instalare și utilizare**

## **7.1. Resurse software și hardware necesare**

Pentru a instala și rula aplicația, sunt necesare următoarele resurse:

### **7.1.1. Resurse hardware**

- Smartphone cu sistem de operare Android (versiunea 5.0 sau mai nouă).

### **7.1.2. Resurse software**

- Sistem de operare Windows pentru dezvoltarea aplicației.
- Android Studio (pentru dezvoltare și debugging).
- Python 3.7+ (pentru backend).
- Biblioteci Python necesare: Flask, XGBoost, scikit-learn, pandas, joblib.

## **7.2. Instalarea aplicației**

### **7.2.1. Instalare pe smartphone Android**

1. Obțineți fișierul APK al aplicației de la dezvoltator sau de pe dispozitivul de stocare extern (USB, card de memorie etc.).
2. Transferați fișierul APK pe telefonul dvs. prin intermediul unui cablu USB, Bluetooth sau orice altă metodă preferată de transfer de fișiere.
3. Deschideți managerul de fișiere pe telefonul dvs. și navigați la locația unde ați salvat fișierul APK.
4. Deschideți fișierul APK.
5. Dacă este prima dată când instalați aplicații din surse necunoscute, activați opțiunea "Allow from this source" din setările telefonului.
6. Continuați instalarea conform instrucțiunilor de pe ecran.
7. După finalizarea instalării, deschideți aplicația din meniul principal al telefonului.

### **7.2.2. Instalare backend pe server local**

1. Descărcați și instalați Python 3.7+.
2. Descărcați codul sursă al backend-ului primit de la dezvoltator.
3. Navigați în directorul unde ați descărcat codul sursă și instalați bibliotecile necesare:

```
pip install flask xgboost scikit-learn pandas joblib
```

4. Rulați serverul Flask:

```
python app.py
```

5. Asigurați-vă că adresa IP și portul specificat în aplicația Android se potrivesc cu cele pe care rulează serverul Flask.

## 7.3. Utilizarea aplicației

Aplicația are mai multe funcționalități utile pentru gestionarea împrumuturilor și economiilor. În continuare, vom descrie modul de utilizare a fiecărei funcționalități principale, folosind capturi de ecran și explicații pas cu pas.

### 7.3.1. Predicția aprobației împrumutului

- Deschideți aplicația și accesați secțiunea *Predict Your Loan Approval* din meniu principal (Figura 7.1).

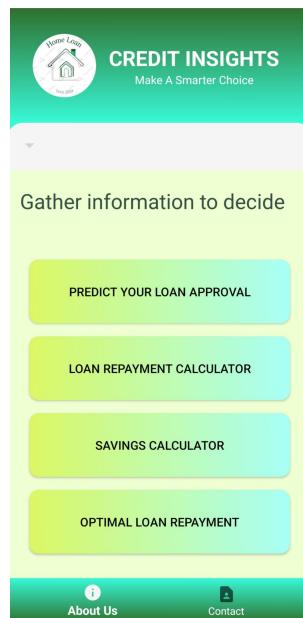


Figura 7.1: Meniul principal al aplicației

- Citiți introducerea, după care apăsați pe butonul *START* (Figura 7.2).



Figura 7.2: Enter Caption

3. Răspundeți la întrebarea despre vârstă selectând o opțiune (Figura 7.3).

Are you at least 18 years old?

Yes  
 No

NEXT  
BACK

Figura 7.3: Vârsta

4. Răspundeți pe rând la întrebările despre gen, statut marital și nivel de educație din paginile prezentate în figura 7.4.

What's your Gender?

Male  
 Female

NEXT  
BACK

Are you married?

Yes  
 No

NEXT  
BACK

What is your education level?

Graduate  
 Not Graduate

NEXT  
BACK

Figura 7.4: 3 pagini din aplicație: Gen, Statut marital, Educație

5. Răspundeți pe rând la întrebările despre angajare, zona proprietății și dependenți din paginile prezentate în figura 7.5.

Are you self-employed?

Yes  
 No

NEXT  
BACK

What is your property area?

Urban  
 Semi-Urban  
 Rural

NEXT  
BACK

How many dependents do you have?

0  
 1  
 2  
 3+

NEXT  
BACK

Figura 7.5: 3 pagini din aplicație: Angajare, Zonă proprietate, Numărul de dependenți

6. Selectați dacă doriți să introduceți suma în roni sau euro. După aceea, mutați cursorul pentru a selecta venitul solicitantului (în prima pagină din figura 7.6) și venitul cosolicitantului (în a doua pagină din figura 7.6). Sub cursor, va apărea suma selectată de dumneavoastră.

The figure consists of two side-by-side screenshots of a mobile application interface. Both screens have a light green header and footer. The left screen asks "What is your applicant income?" and shows two radio buttons: "RON" (selected) and "Euro". Below the buttons is a text input field containing "Selected Income (RON): 498" and "Income in Euro: 100". At the bottom are "NEXT" and "BACK" buttons. The right screen asks "What is your coapplicant income?" and also has "RON" selected. It shows a text input field with "Selected Income (Ron): 0". At the bottom are "NEXT" and "BACK" buttons.

Figura 7.6: Venitului solicitantului și cosolicitantului

7. Selectați termenul împrumutului cu ajutorul *spinner-ul* (în prima pagină din figura 7.7). Apoi, selectați dacă doriți să introduceți suma în roni sau euro. După aceea, mutați cursorul pentru a selecta suma împrumutului pe care doriți să îl faceți (în a doua pagină din figura 7.7).

The figure consists of two side-by-side screenshots of a mobile application interface. Both screens have a light green header and footer. The left screen asks "What is your loan amount term (in months)?". It features a dropdown spinner with "12" selected. At the bottom are "NEXT" and "BACK" buttons. The right screen asks "What is your loan amount?". It shows two radio buttons: "RON" (selected) and "Euro". Below the buttons is a text input field containing "Selected Loan Amount (RON): 5". At the bottom are "NEXT" and "BACK" buttons.

Figura 7.7: Suma și termenul împrumutului

8. Răspundeți la întrebarea despre istoricul de credit (Figura 7.8).

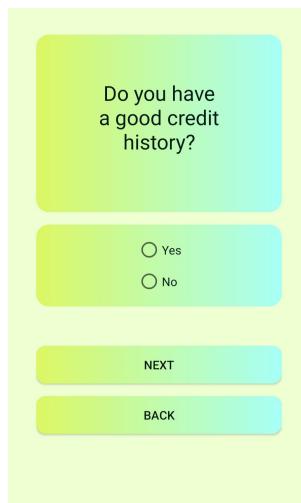


Figura 7.8: Istoricul de credit

9. După completarea tuturor informațiilor, va apărea pagina cu rezultatul predicției împrumutului. Apăsați *BACK TO MAIN* pentru a ajunge înapoi la pagina principală (Figura 7.9).

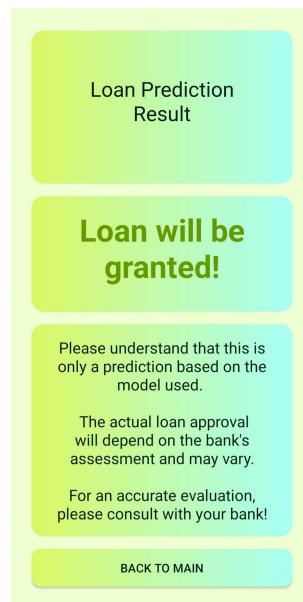


Figura 7.9: Rezultatul predicției împrumutului

#### 7.3.2. Calculator de rambursare a împrumutului

1. Alegeti opțiunea "Loan Repayment Calculator" din meniul principal (Figura 7.1).
2. Selectați între roni și euro. Introduceți detaliile împrumutului, cum ar fi suma, rata dobânzii, termenul în luni, tipul împrumutului(fix sau variabil), data începerii împrumutului (prima imagine din figura 7.10).
3. Apăsați butonul "CALCULATE" pentru a vedea detaliile rambursării, inclusiv plătile lunare și dobânda totală.

4. Dacă apăsați butonul *PRINTABLE SCHEDULE* se va deschide programul de plată lunar al împrumutului ipotecar (a doua imagine din figura 7.10). Acesta începe cu prima lună din care ati ales să înceapă împrumutul.

ID	Date	Rate	Payment	Balance
1	Jun 2024	5.50%	8819.13	202516.88
2	Jul 2024	5.50%	8819.13	193697.75
3	Aug 2024	5.50%	8819.13	184878.61
4	Sep 2024	5.50%	8819.13	176059.48
5	Oct 2024	5.50%	8819.13	167240.35
6	Nov 2024	5.50%	8819.13	158421.22
7	Dec 2024	5.50%	8819.13	149602.09
8	Jan 2025	5.50%	8819.13	140782.96
9	Feb 2025	5.50%	8819.13	131963.83
10	Mar 2025	5.50%	8819.13	123144.70
11	Apr 2025	5.50%	8819.13	114325.56
12	May 2025	5.50%	8819.13	105506.43
13	Jun 2025	5.20%	8792.20	96714.23
14	Jul 2025	5.20%	8792.20	87922.03
15	Aug 2025	5.20%	8792.20	79129.82
16	Sep 2025	5.20%	8792.20	70337.62
17	Oct 2025	5.20%	8792.20	61545.42
18	Nov 2025	5.20%	8792.20	52753.22
19	Dec 2025	5.20%	8792.20	43961.01
20	Jan 2026	5.20%	8792.20	35168.81
21	Feb 2026	5.20%	8792.20	26376.61
22	Mar 2026	5.20%	8792.20	17584.41
23	Apr 2026	5.20%	8792.20	8792.20
24	May 2026	5.20%	8792.20	0.00

Figura 7.10: Calculatorul de rambursare a împrumutului și programul de plată

### 7.3.3. Calculator de economii

- Selectați opțiunea "Savings Calculator" din meniul principal (Figura 7.1).
- Introduceți detalii precum suma inițială, depozitul lunar, rata dobânzii și numărul de luni (Figura 7.11).
- Apăsați butonul "CALCULATE" pentru a vedea economiile viitoare și câștigurile nete.

Figura 7.11: Calculatorul de economii

#### 7.3.4. Optimizarea rambursării împrumutului

1. Accesați opțiunea "Optimal Loan Repayment" din meniu principal (Figura 7.1).
2. Introduceți detalii precum suma împrumutului, dobânda anuală și termenul împrumutului (Figura 7.12).
3. Apăsați butonul "CALCULATE" pentru a obține informații despre contribuția lunăriă optimă pentru a acoperi dobânda împrumutului.

The screenshot displays a mobile application interface for calculating monthly savings contributions. At the top, it says "Monthly Savings Contribution for Loan Interest Repayment". Below that, it asks "Select Currency" with options "RON" (selected) and "EUR". There are five input fields: "Loan Amount", "Annual Interest Rate (%)", "Loan Term (Months)", "Initial Savings", and "Annual Savings Interest Rate (%)" (all empty). A large blue "CALCULATE" button is centered below these fields. At the bottom, the results are shown: "Total Interest: 0.00" and "Monthly Contribution: 0.00".

Figura 7.12: Optimizarea rambursării împrumutului

## 7.4. Concluzie

Acest manual oferă ghidul necesar pentru instalarea și utilizarea aplicației. Utilizând acest ghid, oricine poate instala și utiliza aplicația pentru a gestiona și optimiza împrumuturile și economiile. Aplicația oferă o interfață prietenoasă și funcționalități esențiale pentru o gestionare financiară eficientă.

## **Capitolul 8. Concluzii**

Proiectul de față își propune să ușureze procesul de evaluare a creditelor ipotecare prin utilizarea unei aplicații mobile care integrează algoritmi de învățare automată pentru a prezice aprobarea împrumuturilor și a îmbunătăți gestionarea financiară personală. Scopul principal al acestei lucrări a fost să dezvolte un instrument eficient și accesibil pentru utilizatori, care să ofere evaluări rapide și precise, contribuind astfel la o mai bună planificare financiară și reducerea riscului de credit.

Această lucrare a abordat tema propusă din perspective teoretice și practice. Din punct de vedere teoretic, s-au explorat diverse aspecte ale învățării automate aplicate în contextul evaluării riscului de credit, inclusiv utilizarea algoritmilor *Decision Tree*, *Random Forest*, *k-Nearest Neighbors* și *XGBoost*. Analiza detaliată a acestor algoritmi a permis selectarea celui mai performant model, *XGBoost*, care a obținut o acuratețe de 81.30%.

Din punct de vedere practic, s-a realizat dezvoltarea unei aplicații mobile intuitive, care oferă funcționalități precum predicția aprobării împrumuturilor, calculul plășilor lunare pentru împrumuturi și planificarea economiilor. Aplicația permite utilizatorilor să-și gestioneze eficient plășile lunare și să reducă costul total al împrumuturilor prin strategii optime de rambursare. De asemenea, calculatorul de economii ajută utilizatorii să-și atingă obiectivele financiare pe termen lung, oferind estimări precise ale sumelor economisite pe baza depunerilor regulate și a dobânzii.

Un element de inovație adus de această aplicație este funcționalitatea de calcul a contribuției lunare pe care utilizatorul trebuie să o facă în contul de economii pentru a acoperi costul dobânzii împrumutului ipotecar. Această funcționalitate ajută utilizatorii să-și planifice mai eficient economiile și să reducă povara financiară pe termen lung.

Rezultatele obținute au demonstrat că s-au îndeplinit obiectivele și cerințele dezvoltate în capitolul 2. Modelul predictiv a avut o performanță bună, iar funcționalitățile aplicației au fost apreciate de utilizatori. Feedback-ul utilizatorilor a fost esențial în rafinarea interfetei și a navigării aplicației, contribuind la o experiență de utilizare îmbunătățită. Utilizatorii au apreciat în mod special partea de predicție a acordării împrumuturilor, calculul plășilor lunare și planificarea economiilor.

Cu toate acestea, există încă loc pentru îmbunătățiri. Utilizarea unor seturi de date mai mari și mai diverse ar putea crește acuratețea modelului predictiv și ar îmbunătăți generalizarea acestuia. Adițional, integrarea unor surse de date externe, precum informații economice sau demografice, ar putea oferi o perspectivă mai completă și ar spori capacitatea modelului de a face predicții precise.

### **8.1. Dezvoltări ulterioare**

Pentru a îmbunătăți semnificativ aplicația, viitoarele dezvoltări ar trebui să se axeze pe câteva aspecte esențiale. Utilizarea unui set de date mai amplu și diversificat, care să includă informații din diferite regiuni geografice și demografice, va spori acuratețea și capacitatea de generalizare a modelului de învățare automată. Acest lucru va permite aplicației să ofere predicții mai precise și relevante pentru un spectru mai larg de utilizatori, reflectând mai bine diversitatea situațiilor financiare reale.

O altă îmbunătățire valoroasă ar fi implementarea unei funcționalități care să permită utilizatorilor să descarce rapoartele de plată în format PDF direct pe telefon. Aceasta va oferi o modalitate convenabilă și eficientă pentru utilizatori de a gestiona documentele financiare și de a păstra evidența plășilor. În plus, rapoartele PDF ar putea include grafice detaliate ale plășilor, ratei dobânzii și evoluției soldului împrumutului, oferind utilizatorilor o înțelegere vizuală mai bună a situației lor financiare.

Integrarea aplicației cu servicii de banking va permite utilizatorilor să gestioneze plășile și să aibă o imagine de ansamblu clară asupra situației lor financiare. Această integrare va oferi o experiență de utilizare mai coerentă și mai simplificată, facilitând gestionarea eficientă a finanșelor personale. Utilizatorii vor putea să-și monitorizeze conturile bancare, să programeze plăști automate pentru împrumuturi și să primească notificări despre tranzacțiile importante direct din aplicație.

Crearea unui server dedicat: Pentru a rezolva problema IP-ului dinamic și a asigura accesul constant la funcționalitățile de predicție, ar fi necesară crearea unui server dedicat care să ruleze permanent algoritmul Python. Acest lucru ar permite utilizatorilor să utilizeze aplicația pentru predicții oricând, fără întreruperi.

Extinderea aplicației pentru a oferi recomandări personalizate bazate pe comportamentul și preferințele utilizatorilor va transforma aplicația într-un veritabil asistent financiar virtual. Acest lucru va ajuta utilizatorii să ia decizii informate și să își gestioneze resursele financiare mai eficient.

În concluzie, dezvoltarea acestei aplicații mobile pentru prezicerea aprobării împrumuturilor ipotecare a demonstrat potențialul semnificativ al tehnologiei învățării automate în sectorul financiar. Implementarea viitoarelor îmbunătățiri va aduce un plus de valoare, facilitând accesul la credite și îmbunătățind gestionarea finanșelor personale. Astfel, aplicația se poziționează ca un instrument esențial pentru utilizatorii care doresc să afle mai ușor dacă vor primi un împrumut și să înțeleagă mai bine procesul de rambursare a creditelor și gestionarea finanșelor.

## Bibliografie

- [1] K. Kazbekova, A. Adambekova, S. Baimukhanova, G. Kenges, and D. Bokhaev, “Bank risk management in the conditions of financial system instability,” *Entrepreneurship and Sustainability Issues*, vol. 7, no. 4, p. 3269, 2020.
- [2] S. Bhatore, L. Mohan, and Y. R. Reddy, “Machine learning techniques for credit risk evaluation: a systematic literature review,” *Journal of Banking and Financial Technology*, vol. 4, no. 1, pp. 111–138, 2020.
- [3] A. Lusardi, “Financial literacy and the need for financial education: evidence and implications,” *Swiss Journal of Economics and Statistics*, vol. 155, no. 1, pp. 1–8, 2019.
- [4] V. L. Sidor and D. Manate, “Financial education among adults in romania,” *Journal of Financial Studies*, vol. 8, no. 14, pp. 165–180, 2023.
- [5] D. Broby, “The use of predictive analytics in finance,” *The Journal of Finance and Data Science*, vol. 8, pp. 145–161, 2022.
- [6] V. Murinde, E. Rizopoulos, and M. Zachariadis, “The impact of the fintech revolution on the future of banking: Opportunities and risks,” *International review of financial analysis*, vol. 81, p. 102103, 2022.
- [7] M. Favaretto, E. De Clercq, and B. S. Elger, “Big data and discrimination: perils, promises and solutions. a systematic review,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–27, 2019.
- [8] M. Nițoi, D. Clichici, C. Zeldea, M. M. Pochea, and C. Ciocîrlan, “Bunăstarea financiară și alfabetizarea financiară în românia (financial well-being and financial literacy in romania),” *Institutul de Economie Mondială Working Paper*, 2022.
- [9] A. Oualid, A. Hansali, Y. Balouki, and L. Moumoun, “Application of machine learning techniques for credit risk management: a survey,” in *Advances in Information, Communication and Cybersecurity: Proceedings of ICI2C'21*. Springer, 2022, pp. 180–191.
- [10] A. Gupta, V. Pant, S. Kumar, and P. K. Bansal, “Bank loan prediction system using machine learning,” in *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*, 2020, pp. 423–426.
- [11] A. Sharma and V. Kumar, “An exploratory study-based analysis on loan prediction,” *Inventive Communication and Computational Technologies: Proceedings of ICICCT 2022*, pp. 423–433, 2022.
- [12] D. Dansana, S. G. K. Patro, B. K. Mishra, V. Prasad, A. Razak, and A. W. Wodajo, “Analyzing the impact of loan features on bank loan prediction using random forest algorithm,” *Engineering Reports*, vol. 6, no. 2, p. e12707, 2024.

- [13] P. Dutta, "A study on machine learning algorithm for enhancement of loan prediction," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 3, 2021.
- [14] C. Chakraborty and A. Joseph, "Machine learning at central banks," 2017.
- [15] M. Anand, A. Velu, and P. Whig, "Prediction of loan behaviour with machine learning models for secure banking," *Journal of Computer Science and Engineering (JCSE)*, vol. 3, no. 1, pp. 1–13, 2022.
- [16] G. A. Perry and M. Krantz, *Mint. com for Dummies*. John Wiley & Sons, 2010.
- [17] E. Nagypal and J. Tobacman, "Credit characteristics, credit engagement tools, and financial well-being," *Consumer Financial Protection Bureau Office of Research Reports Series*, no. 19-7, 2019.
- [18] J. Karotkin, "Money life hacks: Optimizing your budget danielle molinar, nathan seibert."
- [19] B. A. Manko, "Teaching user-friendly web design: A case study on zillow. com in the real estate industry," *Journal of Information Technology Teaching Cases*, vol. 12, no. 1, pp. 35–42, 2022.
- [20] S. Joseph and M. A. Bhuva, "A study on customer satisfaction towards mobile banking app of hdfc bank in mumbai region."
- [21] D. J. Hand, "Construction and assessment of classification rules," *Wiley Series in Probability and Statistics*, 1997.
- [22] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*. Pearson, 2016.
- [23] R. Mileris, "Macroeconomic determinants of loan portfolio credit risk in banks," *Engineering Economics*, vol. 23, no. 5, pp. 496–504, 2012.
- [24] K. Rodrigo, T. Sandanayake, and A. Silva, "Personal loan default prediction and impact analysis of debt-to-income ratio," in *2023 8th International Conference on Information Technology Research (ICITR)*. IEEE, 2023, pp. 1–6.
- [25] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Data preprocessing for supervised learning," *International Journal of Computer Science*, 2006.
- [26] E. Rahm and H. H. Do, "Data cleaning: Problems and current approaches," *IEEE Data Engineering Bulletin*, 2000.
- [27] R. J. Little and D. B. Rubin, *Statistical Analysis with Missing Data*. John Wiley & Sons, 2014.
- [28] D. Pyle, *Data Preparation for Data Mining*. Morgan Kaufmann, 1999.
- [29] K. S. Sree, J. Karthik, C. Niharika, P. Srinivas, N. Ravinder, and C. Prasad, "Optimized conversion of categorical and numerical features in machine learning models," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*. IEEE, 2021, pp. 294–299.

- [30] L. Yang and A. Shami, “On hyperparameter optimization of machine learning algorithms: Theory and practice,” *Neurocomputing*, vol. 415, pp. 295–316, 2020.
- [31] A. Gupta, V. Pant, S. Kumar, and P. K. Bansal, “Bank loan prediction system using machine learning,” in *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*. IEEE, 2020, pp. 423–426.
- [32] G. Arutjothi and C. Senthamarai, “Prediction of loan status in commercial bank using machine learning classifier,” in *2017 international conference on intelligent sustainable systems (ICISS)*. IEEE, 2017, pp. 416–419.
- [33] L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications*. World Scientific, 2014.
- [34] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Elsevier, 2014.
- [35] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Elsevier, 2011.
- [36] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Routledge, 2017.
- [37] S. K. Murthy, “Automatic construction of decision trees from data: A multidisciplinary survey,” *Data Mining and Knowledge Discovery*, vol. 2, pp. 345–389, 1998.
- [38] G. M. Weiss, “Maximize what matters: Predicting customer value through decision trees and decision rules,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 593–602, 2008.
- [39] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, pp. 123–140, 1996.
- [40] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.
- [41] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [42] A. Liaw and M. Wiener, “Classification and regression by randomforest,” *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [43] T. K. Ho, “Random decision forests,” *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, pp. 278–282, 1995.
- [44] G. Biau and E. Scornet, “Random forests and permutation importance: Theoretical and empirical insights,” *Test*, vol. 25, pp. 197–227, 2016.
- [45] S. Bernard, L. Heutte, and S. Adam, “The influence of hyperparameters on random forest accuracy,” *The International Journal of Forest Research*, 2009.
- [46] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot, “Variable selection using random forests,” *Pattern Recognition Letters*, vol. 31, no. 14, pp. 2225–2236, 2010.

- [47] A. Verikas, A. Gelzinis, and M. Bacauskiene, “The random forest classifier: Overview, features, and application in bioinformatics,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 3, pp. 291–300, 2011.
- [48] N. S. Altman, “An introduction to kernel and nearest-neighbor nonparametric regression,” *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.
- [49] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley & Sons, 2001.
- [50] T. Cover and P. Hart, “Nearest neighbor pattern classification,” *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [51] L. E. Peterson, *K-nearest neighbor*. Springer, 2009.
- [52] E. Fix and J. L. Hodges, “Discriminatory analysis: Nonparametric discrimination: Consistency properties,” 1951.
- [53] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.
- [54] T. M. Mitchell, *Machine Learning*. McGraw-Hill, 1997.
- [55] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.
- [56] T. Chen and T. He, “Xgboost: extreme gradient boosting,” *R package version 0.4-2*, vol. 1, no. 4, pp. 1–4, 2015.
- [57] J. H. Friedman, “Greedy function approximation: a gradient boosting machine,” *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [58] R. Kohn, “A capital budgeting model of the supply and demand of loanable funds,” *Journal of Macroeconomics*, vol. 12, no. 3, pp. 427–436, 1990.
- [59] M. A. Kumar, A. Kumar, and H. P. P. H. Constable, “Power of compounding in systematic investment plan (sip).”
- [60] T. T. Lin, C.-C. Lee, and C.-H. Chen, “Impacts of the borrower’s attributes, loan contract contents, and collateral characteristics on mortgage loan default,” *The Service Industries Journal*, vol. 31, no. 9, pp. 1385–1404, 2011.
- [61] J. L. Negrin *et al.*, “The importance of borrowers’ history on credit behavior: The mexican experience,” Econometric Society, Tech. Rep., 2004.