|  |  |
| --- | --- |
|  | Universitatea Tehnică a Moldovei |

**Asistent Virtual pentru Optimizarea Experienței Pacienților în Domeniul Sănătății**

|  |  |
| --- | --- |
| Student: | Ciobanu Cristalin |
|  |  |
| Coordonator: | Braniște Rodica,  lect. univ. |

**Chişinău, 2024**

CUPRINS

[INTRODUCERE 1](#_Toc164340810)

[1 ANALIZA PRODUSELOR PROGRAM EXISTENTE ÎN OPTIMIZAREA EXPERIENȚEI PACIENȚILOR ÎN DOMENIUL SĂNĂTĂȚII 4](#_Toc164340811)

[1.1 MyChart 5](#_Toc164340813)

[1.2 Zocdoc 8](#_Toc164340814)

[1.3 Teladoc Health 12](#_Toc164340815)

[1.4 Concluzii intermediare la primul capitol 14](#_Toc164340816)

[2 CADRUL FUNDAMENTAL TEORETIC PRIVIND DEZVOLTAREA UNUI ASISTENT VIRTUAL 16](#_Toc164340817)

[2.1 Inteligența artificială 16](#_Toc164340818)

[2.2 Aspecte etice în inteligența artificială 21](#_Toc164340819)

[2.3 Procesarea limbajului natural(NLP) 22](#_Toc164340820)

[2.4 Învățarea automată. 26](#_Toc164340821)

[2.5 Instrumente utilizate 29](#_Toc164340822)

[2.6 Modelarea conceptului sistemului 34](#_Toc164340823)

[2.7 Concluzii intermediare la al doilea capitol 40](#_Toc164340824)

[3 DEZVOLTAREA UNEI APLICAȚII PENTRU PREDICȚIA RISCULUI ATACULUI DE CORD 42](#_Toc164340825)

[3.1 Analiza distibuțiilor variabilelor 42](#_Toc164340826)

[3.2 Avantajele regresiei logistice 47](#_Toc164340827)

[3.3 Antrenarea modelului de regresie logistică 47](#_Toc164340828)

[3.4 Acurateția modelului 48](#_Toc164340829)

[3.5 49](#_Toc164340830)

[CONCLUZII 50](#_Toc164340831)

[BIBLIOGRAFIE 52](#_Toc164340832)

# INTRODUCERE

În contextul acestei lucrări de licență, am elaborat o aplicație, cu scopul de a automatiza procesul de prezicere a riscurilor apariției atacului de cord în baza unor date introduse de utilizator. Relevanța acestui subiect este evident în prezent, având în vedere că pe parcursul ultimelor decenii, cercetările în domeniul sănătății cardiovasculare au evoluat considerabil, beneficiind de avansurile tehnologice și de accesul sporit la date relevante. Atacul de cord (infarctul miocardic) rămâne una dintre cele mai presante probleme de sănătate la nivel mondial, exercitând o presiune semnificativă asupra sistemelor de sănătate și având consecințe serioase asupra calității vieții și a longevității populației.

În urma pandemiei de COVID-19, s-a observat o creștere semnificativă a afecțiunilor legate de inimă și a problemelor cardiovasculare. Conform unui studiu efectuat în Statele Unite ale Americii, bolile de inimă reprezintă principala cauză de deces pentru bărbați, femei și persoane din majoritatea grupurilor rasiale și etnice din Statele Unite. De exemplu, dacă luăm date mai generale legate de boli cardiovasculare, nu neapărat atac de cord, o persoană moare la fiecare 33 de secunde în Statele Unite. Aproximativ 695.000 de persoane din Statele Unite au murit din cauza bolilor de inimă în 2021 - adică 1 din 5 decese.

Bolile de inimă au costat Statele Unite aproximativ 239,9 miliarde de dolari în fiecare an, în perioada 2018-2019. Acest lucru include costul serviciilor de asistență medicală, al medicamentelor și al productivității pierdute din cauza decesului.

Dacă ne axăm pe atacuri de cord, în Statele Unite, cineva suferă un atac de cord la fiecare 40 de secunde. În fiecare an, aproximativ 805.000 de persoane din Statele Unite suferă un atac de cord. Dintre acestea, 605.000 reprezintă un prim atac de cord, iar altele 200.000 de cazuri se întâmplă la persoane care au avut deja un atac de cord. Aproximativ 1 din 5 atacuri de cord sunt silențioase - daunele sunt făcute, dar persoana nu este conștientă de ele[1].

În lumina acestor dovezi, actualitatea dezvoltării unui asistent virtual bazat pe un model de machine learning este pozitivă. Un asemenea model va putea ajuta utilizatorul să înțeleagă în baza indicatorilor furnizați dacă este expus unui eventual atac de cord sau nu.

Aceste funcționalități ale sistemului contribuie la accelerarea accesului pacienților la îngrijirea medicală necesară, oferind o soluție rapidă, simplificată și convenabilă.

**Scopul principal** al acestei teze este de a explora dezvoltarea și implementarea asistenților virtuali în contextul optimizării experienței pacienților în domeniul sănătății. Ne propunem să înțelegem impactul acestora asupra facilitării accesului oamenilor la medicină și să dezvoltăm un asistent virtual modern și eficient pentru optimizarea experienței pacienților în domeniul sănătății.

**Problemele existente** includ nevoia de a reduce munca manuală și de a oferi soluții personalizate pentru utilizatori. Asistenții virtuali pot reprezenta o modalitate de a răspunde la aceste probleme, dar există încă provocări tehnice și de implementare care trebuie abordate.

**Obiectivele** acestei teze sunt următoarele:

1. Să dezvoltăm un asistent virtual eficient care să poată automatiza și simplifica procesele umane în privința diagnosticării și oferirea de recomandări pacienților.
2. Să identificăm și să soluționăm provocările tehnice legate de implementarea asistenților virtuali în optimizarea experienței pacienților în domeniul sănătății.
3. Să explorăm aspectele teoretice legate de tehnologia asistenților virtuali și să le aplicăm în practică.

Pentru a oferi o soluție practică și funcțională în domeniul dezvoltării platforme din cadrul prezentei teze au fost folosite următoarele **metodologii de cercetare**:

1. Analiza documentelor și a literaturii existente despre asistenții virtuali, tehnologia lor și impactul asupra societății privind optimizărea experienței pacienților în domeniul sănătății.
2. Studii de caz pentru a investiga exemple concrete de asistenți virtuali și modul în care aceștia pot fi utilizați în domeniul optimizării experienței pacienților.
3. Dezvoltarea și testarea unui asistent virtual propriu pentru a înțelege procesul de implementare și provocările tehnice implicate.

Structura tezei de licență este compusă din trei capitole distincte, fiecare având un scop specific în abordarea dezvoltării și implementării unui asistent virtual în era digitală. Iată cum sunt organizate aceste capitole:

**Primul capitol** se axează pe analiza atentă a produselor program existente. Scopul acestui capitol este să identifice și să evalueze sisteme informatice disponibile pe piață. Prin intermediul unui studiu de piață riguros, vom investiga variate produse analizând caracteristicile și funcționalitățile acestora pentru a evita neajusurile sistemelor existente.

**Al doilea capitol** se concentrează pe furnizarea unui fundament teoretic și conceptual necesar pentru înțelegerea asistenților virtuali și a contextului lor în era digitală. În cadrul acestui capitol, ne propunem să explorăm mai detaliat conceptele cheie legate de asistenții virtuali, precum și să evidențiem importanța lor în societatea contemporană. De asemenea, vom investiga tehnologiile care stau la baza asistenților virtuali, cum ar fi procesarea de limbaj natural, învățarea automată. Totodată, se vor explora diverse etape utilizate în dezvoltarea unui asistent virtual.

**Al treilea capitol** se concentrează pe detaliile legate de proiectarea și implementarea unui asistent virtual propriu. Aici, vom defini obiectivele proiectului nostru și cerințele pe care le dorim să le îndeplinească asistentul virtual. Implementarea și testarea asistentului virtual sunt aspecte esențiale, iar acest capitol va explora cum acestea sunt realizate pentru a asigura funcționalitatea corectă a asistentului virtual dezvoltat.

# ANALIZA PRODUSELOR PROGRAM EXISTENTE ÎN OPTIMIZAREA EXPERIENȚEI PACIENȚILOR ÎN DOMENIUL SĂNĂTĂȚII

Un asistent virtual pentru stabilirea unui diagnostic prealabil în baza simptomelor și primirea unor sugestii practice este o inovație semnificativă care aduce numeroase beneficii atât pacienților, cât și pentru specialiștii din domeniul sănătații. Funcționalitățile sale cheie și importanța sa sunt esențiale în îmbunătățirea sistemului de asistență medicală.

Una dintre funcționalitățile esențiale ale acestui sistem este programarea și gestionarea programărilor. Pacienții pot utiliza această platformă pentru a programa sau anula programări la medici sau pentru servicii de laborator, ceea ce face procesul mai eficient și mai accesibil. Aceasta reduce nevoia de a aștepta la cozi sau de a contacta personalul medical prin intermediul telefonului, ceea ce aduce un beneficiu semnificativ în ceea ce privește gestionarea timpului și accesibilitatea serviciilor.

Asistentul virtual poate, de asemenea, să ofere răspunsuri la întrebările pacienților, oferind informații de bază despre afecțiuni, tratamente, medicamente și sănătate în general. Acest lucru contribuie la creșterea nivelului de înțelegere al pacienților cu privire la problemele lor de sănătate și la opțiunile disponibile pentru tratament.

Pacienții pot introduce simptomele pe care le simt, iar asistentul virtual poate oferi sugestii privind cauzele posibile și dacă ar trebui să solicite asistență medicală. Acest lucru poate contribui la identificarea precoce a problemelor de sănătate și la intervenția promptă a profesioniștilor medicali.

De asemenea, asistentul virtual poate oferi sfaturi privind sănătatea și prevenția bolilor, personalizate pentru fiecare pacient. Aceasta promovează un stil de viață sănătos și ajută la evitarea unor afecțiuni grave prin educație și conștientizare.

Un alt aspect important este accesul la dosarul medical electronic, care permite pacienților să aibă o imagine clară asupra istoricului lor medical și să îl împărtășească cu alți profesioniști din domeniul sănătății. Aceasta facilitează colaborarea între diferiți furnizori de servicii medicale și asigură furnizarea unor tratamente eficiente.

Asistentul virtual poate, de asemenea, să ofere asistență post-tratament și să monitorizeze evoluția pacienților după tratament sau internare. Acest lucru poate contribui la îmbunătățirea recuperării și la prevenirea recidivelor.

În ceea ce privește importanța acestui sistem informatic în domeniul sănătății, aceasta este evidentă. Acest instrument îmbunătățește accesibilitatea la serviciile medicale, crește eficiența în gestionarea programărilor și monitorizarea stării de sănătate a pacienților, reduce costurile și contribuie la îmbunătățirea satisfacției pacienților. În plus, asistentul virtual poate reprezenta o resursă valoroasă pentru profesioniștii din sănătate, ajutându-i să ofere servicii mai bune și să se concentreze asupra îngrijirii pacienților.



## MyChart

MyChart(*figura 1.1.1*) reprezintă o platformă inovatoare de gestionare a sănătății online, oferind pacienților acces facil și securizat la informațiile medicale personale. Această soluție tehnologică a fost dezvoltată pentru a îmbunătăți experiența pacientului, facilitând interacțiunea cu furnizorii de servicii medicale și permițând o gestionare mai eficientă a datelor de sănătate. În continuare, vom explora caracteristicile cheie ale platformei MyChart și modul în care aceasta a evoluat pentru a răspunde nevoilor în schimbare ale pacienților și profesioniștilor din domeniul sănătății.



**Figura 1.1.1***Logo-ul MyChart*

MyChart a luat naștere în contextul dezvoltării tehnologiei informației în domeniul sănătății, începând cu anii 1990, când primele eforturi au fost făcute pentru a oferi pacienților acces la informațiile lor medicale prin intermediul internetului. De-a lungul decadelor, platforma a evoluat semnificativ, adaptându-se la schimbările tehnologice și necesitățile crescânde ale utilizatorilor.

Orientată către inovație și adaptare, MyChart continuă să se dezvolte pentru a aborda provocările în creștere din domeniul sănătății. Avansurile în securitatea datelor, extinderea funcționalităților și integrarea tehnologiilor emergente, cum ar fi inteligența artificială, vor juca un rol crucial în modelarea viitorului acestor soluții, contribuind la îmbunătățirea continuă a îngrijirii pacienților și a eficienței în gestionarea sănătății personale[2].

### *Apariția și dezvoltare*

Apariția și dezvoltarea MyChart au fost influențate de avansurile tehnologice și de nevoile tot mai mari ale pacienților de a avea acces rapid și convenabil la informațiile lor medicale.

**Anii 1990 - *Începuturile:***

1. Ideea de a oferi pacienților acces online la informațiile lor medicale a început să prindă contur în anii 1990, odată cu creșterea utilizării internetului.
2. Unele spitale și sisteme de sănătate au început să dezvolte portaluri online pentru pacienți, oferind funcționalități precum programări online și acces la rezultatele testelor.

**Anii 2000 - *Extinderea funcționalităților:***

1. În această perioadă, platformele de gestionare a sănătății online au început să se dezvolte pentru a include mai multe funcționalități, cum ar fi accesul la istoricul medical, gestionarea rețetelor și comunicarea online cu furnizorii de servicii medicale.
2. Companii specializate au început să ofere soluții integrate care să permită pacienților să acceseze datele lor medicale de la diferite instituții medicale.

**Anii 2010 - *Mobilitate și interoperabilitate:***

1. Odată cu popularizarea dispozitivelor mobile, platformele MyChart au fost adaptate pentru a fi accesibile de pe smartphone-uri și tablete.
2. S-a acordat o atenție sporită interoperabilității, facilitând schimbul de informații medicale între diferite sisteme și instituții medicale.
3. Securitatea datelor a devenit o preocupare majoră, iar platformele MyChart au început să implementeze măsuri robuste de protecție a confidențialității pacienților.

**Anii 2020 -** ***Extinderea funcționalităților și adoptarea generalizată:***

1. MyChart și platforme similare au continuat să se extindă, oferind pacienților acces la mai multe informații, inclusiv rapoarte de imagistică medicală și informații despre tratamente.
2. Telemedicina a fost integrată în MyChart, permițând pacienților să aibă consultații medicale online.
3. Adoptarea generalizată a platformelor MyChart în sistemele de sănătate din întreaga lume a fost accelerată de pandemia COVID-19, care a evidențiat importanța accesului la servicii de sănătate online.

Astăzi, MyChart și platforme similare sunt utilizate în întreaga lume, facilitând interacțiunea pacienților cu furnizorii de servicii medicale și contribuind la gestionarea eficientă a datelor de sănătate. Dezvoltarea continuă a acestor platforme este probabil să continue, cu accent pe inovații tehnologice, securitatea datelor și extinderea funcționalităților pentru a răspunde nevoilor tot mai complexe ale pacienților și profesioniștilor din domeniul sănătății[3].

### Avantaje și dezavantaje

MyChart, precum și alte platforme de gestionare a sănătății online, aduc cu sine atât avantaje semnificative, cât și dezavantaje.

**Avantaje:**

1. **Acces la informații medicale**:

Pacienții au acces facil la informațiile lor medicale, inclusiv rezultatele testelor, istoricul medical și planurile de tratament.

1. **Comunicare eficientă**:

MyChart facilitează comunicarea între pacienți și furnizori de servicii medicale prin intermediul mesajelor online, facilitând întrebările, răspunsurile și programarea de consultații.

1. **Gestionare a sănătății personale**:

Pacienții pot monitoriza și gestiona propriile lor date de sănătate, inclusiv programări, medicamente și simptome, contribuind la o implicare mai activă în propriul lor proces de îngrijire.

1. **Telemedicină**:

Platforma MyChart permite consultări medicale online, ceea ce este benefic pentru pacienții care nu pot ajunge fizic la un cabinet medical.

1. **Interoperabilitate**:

Platforma MyChart au fost dezvoltată pentru a asigura interoperabilitatea cu diferite sisteme de sănătate, permițând schimbul eficient de informații medicale între instituții.

1. **Ușurință în gestionarea rețetelor**:

Pacienții pot gestiona rețetele online, inclusiv solicitarea de reînnoire și urmărirea medicamentelor prescrise.

**Dezavantaje:**

1. **Securitatea datelor**:

Îngrijorările privind securitatea datelor pot apărea, mai ales în contextul tot mai frecventelor încălcări de securitate cibernetică.

1. **Excluderea unor grupuri**:

Unele persoane, în special cele în vârstă sau cu acces limitat la tehnologie, ar putea fi excluse din beneficiile platformelor online, creând o discrepanță digitală.

1. **Probleme tehnice**:

Erori tehnice sau probleme de conectivitate pot afecta experiența utilizatorului și accesul la informații critice de sănătate.

1. **Lipsa de standardizare**:

Lipsa unei standardizări complete în domeniu poate duce la fragmentarea datelor medicale și dificultăți în schimbul eficient de informații între diferite sisteme.

1. **Confidențialitatea și etica**:

Există preocupări cu privire la confidențialitatea datelor și la modul în care informațiile medicale sunt utilizate și stocate, ridicând întrebări legate de etică și consimțământul pacientului.

În ciuda dezavantajelor, MyChart continuă să evolueze, abordând treptat aceste provocări și aducând îmbunătățiri constante pentru a oferi o experiență mai sigură și mai eficientă pacienților în gestionarea sănătății lor.

## Zocdoc

Zocdoc(*figura 1.2.1*) este o platformă online care facilitează programarea la medici și furnizori de servicii medicale. A fost fondată pentru a rezolva problema programărilor medicale și pentru a oferi pacienților un mod simplu și convenabil de a găsi și programa întâlniri cu medici.



**Figura 1.2.1***Logo-ul ZocDoc*

Prin eliminarea barierelor în procesul de programare și furnizarea de informații utile pentru pacienți, platforma a contribuit la creșterea accesibilității la îngrijiri medicale și la îmbunătățirea experienței pacientului. De asemenea, Zocdoc a evoluat pentru a răspunde la schimbările din industrie, incluzând telemedicina și continuând să-și extindă oferta de servicii[4].

### Istoria apariției

Zocdoc a fost fondată în anul 2007 de către Dr. Cyrus Massoumi, Dr. Oliver Kharraz și Dr. Nick Ganju. Ideea înființării acestei platforme a pornit de la o experiență personală a lui Cyrus Massoumi, care a avut o problemă medicală urgentă și a constatat cât de dificil poate fi să obțină o programare rapidă la un medic.

În timpul unei vizite în străinătate, Massoumi a avut un accident și a avut nevoie de îngrijiri medicale urgente. În căutarea unui medic, a întâmpinat dificultăți în a găsi disponibilitate rapidă și a observat că sistemul de programare medicală nu era la fel de eficient pe cât ar fi putut fi. Această experiență l-a inspirat să creeze o platformă care să faciliteze procesul de programare și să ofere pacienților acces rapid la îngrijiri medicale.

Astfel, în 2007, împreună cu Dr. Oliver Kharraz și Dr. Nick Ganju, Massoumi a fondat Zocdoc cu scopul de a conecta pacienții cu medicii și furnizorii de servicii medicale într-un mod mai eficient. Ei au dezvoltat o platformă online care permite pacienților să caute, să găsească informații relevante despre medici, să vadă disponibilitatea acestora și să programeze întâlniri rapid.

În timp, Zocdoc a evoluat, extinzându-și serviciile și adaptându-se la schimbările din industrie. Platforma a devenit cunoscută pentru abordarea sa inovatoare în simplificarea procesului de programare medicală și pentru facilitarea accesului pacienților la îngrijiri medicale de calitate într-un timp cât mai scurt posibil[5].

### Funcționalități

Zocdoc furnizează o gamă variată de funcționalități atât pentru pacienți, cât și pentru medici și furnizori de servicii medicale, având ca obiectiv îmbunătățirea procesului de programare și facilitarea experienței în domeniul asistenței medicale.

Utilizatorii au posibilitatea de a căuta medici și furnizori de servicii medicale în funcție de specializare, locație, asigurare medicală și disponibilitate. Procesul de programare este simplificat, permițând pacienților să programeze întâlniri online în timp real, alegând data și ora potrivită în funcție de disponibilitatea medicului.

Recenziile și evaluările oferite de alți pacienți joacă un rol esențial în luarea deciziilor informate privind alegerea unui medic. Zocdoc oferă notificări și amintiri pentru a ajuta pacienții să-și gestioneze programările și să reducă ratele de absență.

Informațiile despre asigurare sunt disponibile pe platformă, furnizând detalii despre planurile de asigurare acceptate de medicii înregistrați. Acest aspect ajută pacienții să înțeleagă costurile asociate cu vizitele la medic.

Zocdoc facilitează întregul proces pentru pacienți, permitându-le să completeze formularele necesare și să gestioneze aspecte legate de programări și pregătirea pentru vizitele la medic.

Platforma a evoluat pentru a include și servicii de telemedicină, oferind posibilitatea programării pentru consultații online. O altă caracteristică importantă este asistența multilingvă, care sprijină accesul la servicii medicale pentru o gamă diversă de utilizatori.

Interfața intuitivă și prietenoasă pentru utilizatori facilitează navigarea și utilizarea eficientă a platformei. Medicii beneficiază și ei de avantaje, având posibilitatea de a gestiona propriile profile, actualiza informații de contact, adăuga orar și gestiona programările.

În plus, Zocdoc se integrează adesea cu sistemele de asigurare medicală, contribuind la transparența în privința acoperirii asigurării pentru serviciile medicale. Toate aceste funcționalități converg către îndeplinirea viziunii Zocdoc de a facilita accesul la îngrijiri medicale și de a optimiza procesul de programare pentru beneficiul tuturor părților implicate[5].

### Analiza pro și contra

Zocdoc, o platformă inovatoare în domeniul asistenței medicale, aduce cu sine o serie de beneficii semnificative, evidențiate prin funcționalități și caracteristici proeminente. Aceste aspecte pozitive contribuie la îmbunătățirea experienței pacienților și la optimizarea procesului de acces la îngrijiri medicale. Să analizăm argumentele *pro* care fac din Zocdoc o resursă valoroasă în domeniul sănătății.

1. **Facilitarea programărilor**:

Zocdoc simplifică procesul de programare a întâlnirilor medicale, oferind pacienților posibilitatea de a găsi rapid medici disponibili și de a programa întâlniri online.

1. **Recenzii și evaluări**:

Funcția de recenzii și evaluări permite pacienților să ia decizii informate în privința alegerea medicilor, contribuind la transparența în ceea ce privește calitatea serviciilor oferite.

1. **Notificări și amintiri**:

Sistemul de notificări și amintiri contribuie la reducerea ratei de absență a pacienților, asigurându-se că aceștia își respectă programările.

1. **Gestionarea informațiilor de asigurare**:

Furnizarea detaliilor privind planurile de asigurare acceptate de medici aduce claritate asupra costurilor asociate vizitelor la medic.

1. **Telemedicină**:

Extinderea serviciilor pentru a include programări pentru consultații online se aliniază cu tendința crescândă a telemedicinii, oferind o opțiune convenabilă pentru pacienți.

1. **Gestionarea întregului proces**:

Zocdoc permite pacienților să completeze formulare și să gestioneze întregul proces într-un mod simplificat și eficient.

Atunci când analizăm platforma Zocdoc, este crucial să luăm în considerare și perspectivele critice care au fost exprimate în privința acestui serviciu. Chiar dacă Zocdoc a adus numeroase beneficii în ceea ce privește accesibilitatea la îngrijiri medicale, există argumente contra care subliniază potențialele neajunsuri sau aspecte care ar putea ridica îngrijorări pentru anumite categorii de utilizatori și pentru sistemul de sănătate în ansamblu. Aceste aspecte, care vor fi detaliate în continuare, includ preocupări cu privire la confidențialitatea datelor, accesibilitate limitată și posibile probleme tehnice, contribuind la o perspectivă mai cuprinzătoare asupra impactului și utilizării platformei Zocdoc.

1. **Accesibilitate limitată**:

Platforma poate exclude anumite categorii de pacienți care nu au acces facil la tehnologie sau care nu sunt confortabili cu utilizarea platformelor online.

1. **Potențiale probleme tehnice**:

Erori tehnice sau probleme de conectivitate pot afecta experiența utilizatorului și accesul pacienților la serviciile medicale.

1. **Posibile probleme de confidențialitate**:

Există preocupări cu privire la confidențialitatea datelor medicale, iar orice încălcare a securității poate ridica probleme serioase.

1. **Dependența de recenzii**:

În unele cazuri, pacienții pot să depindă prea mult de recenzii, care pot fi subiective și influențate de experiențe individuale.

1. **Costurile asociate telemedicinii**:

Deși telemedicina poate oferi o opțiune convenabilă, costurile asociate acestui tip de serviciu pot reprezenta o preocupare pentru unii pacienți.

În ansamblu, Zocdoc aduce numeroase beneficii în ceea ce privește facilitarea accesului la îngrijiri medicale, dar este important să se abordeze și să se depășească eventualele provocări legate de securitatea datelor, accesibilitate și integrare cu alte sisteme din domeniul sănătății[6].

## Teladoc Health

Teladoc Health (*figura 1.3.1*) este o companie de sănătate virtuală care furnizează servicii de telemedicină și soluții de îngrijire medicală la distanță. Fondată în 2002, Teladoc oferă pacienților acces la consultații medicale online prin intermediul unui portal digital, permițându-le să discute cu medici și profesioniști din domeniul sănătăți și fără a fi nevoie să se deplaseze fizic la un cabinet medical. Această formă de îngrijire medicală la distanță a devenit din ce în ce mai importantă, mai ales în contextul creșterii tehnologiei și al nevoii de soluții de sănătate mai flexibile.



**Figura 1.3.1***Logo-ul Teladoc Health*

Teladoc a evoluat de-a lungul anilor și a extins gama de servicii oferite, acoperind o varietate de specialități medicale, inclusiv medicină generală, psihiatrie, dermatologie și altele. Compania și-a propus să ofere acces mai ușor la îngrijirea medicală, eliminând barierele legate de distanță geografică sau de disponibilitatea fizică.

Cu o rețea extinsă de medici licențiați și profesioniști medicali, Teladoc Health a devenit o prezență notabilă în industria sănătății virtuale, contribuind la transformarea modului în care pacienții accesează și primesc îngrijirea medicală. Serviciile sale acoperă o gamă largă de nevoi medicale, de la consultații generale până la îngrijire specializată, oferind o alternativă comodă și eficientă pentru pacienții care caută acces la asistență medicală fără a se deplasa la un cabinet tradițional[7].

### Scurt istoric

Teladoc Health a fost fondată în anul 2002, iar de atunci, a devenit una dintre cele mai proeminente companii de sănătate virtuală și de telemedicină din lume.

**Anul 2002: *Fondarea Companiei***

Teladoc a fost fondată în anul 2002 cu scopul de a oferi acces la îngrijire medicală la distanță prin intermediul tehnologiei. Ideea a fost să furnizeze consultații medicale online pentru pacienți în diverse locații, eliminând necesitatea deplasării fizice la un cabinet medical.

**Anii 2000 - 2010: *Extindere și Dezvoltare***

În primii ani de activitate, Teladoc s-a concentrat pe dezvoltarea și extinderea serviciilor sale de telemedicină. A colaborat cu diverse organizații medicale și asigurători pentru a oferi servicii de sănătate virtuală.

**Anul 2015: *Listare la Bursa de Valori din New York (NYSE)***

Teladoc a devenit o companie publică prin listarea sa la Bursa de Valori din New York (NYSE) în 2015, consolidându-și poziția în piața de sănătate virtuală.

**Anii 2010 - Prezent**: ***Creștere Rapidă și Diversificare***

În perioada următoare, Teladoc a experimentat o creștere rapidă și a diversificat gama de servicii oferite. A achiziționat alte companii din domeniul sănătății digitale pentru a-și extinde portofoliul și a oferi o gamă mai largă de servicii, inclusiv îngrijirea psihiatrică și terapie online.

**Anul 2020: *Impactul Pandemiei de COVID-19***

Pandemia de COVID-19 a amplificat importanța telemedicinii, iar Teladoc a înregistrat o creștere semnificativă a cererii în acest context. Oamenii au căutat soluții de îngrijire la distanță, iar Teladoc a jucat un rol crucial în furnizarea de servicii medicale accesibile în timpuri de criză.

Teladoc Health a continuat să rămână un actor major în transformarea modului în care oamenii accesează și primesc îngrijire medicală. Cu tehnologii inovatoare și o rețea extinsă de profesioniști medicali, compania continuă să joace un rol important în evoluția sănătății digitale la nivel global[8].

### Analiza comparativă

Teladoc Health și Zocdoc sunt două platforme inovatoare care aduc contribuții semnificative în domeniul sănătății digitale, cu un accent diferit asupra experienței pacientului.

Teladoc Health se distinge prin oferirea de servicii extinse de telemedicină, facilitând consultații online cu profesioniști din diverse specializări medicale. Această platformă globală a devenit esențială în furnizarea de soluții de îngrijire medicală la distanță și colaborează activ cu companii și asigurători pentru a extinde accesul la îngrijirea medicală în rândul angajaților și asiguraților.

Pe de altă parte, Zocdoc se concentrează pe simplificarea procesului de programare și găsirea medicilor. Oferind o platformă pentru recenzii și evaluări, Zocdoc ajută pacienții să-și împărtășească experiențele și să ia decizii informate în privința alegerea medicului. În ciuda ofertei de servicii de telemedicină, Zocdoc păstrează o tradiție solidă în facilitarea programărilor pentru întâlniri medicale fizice.

Teladoc Health și Zocdoc prezintă diferențe notabile în privința priorităților. Teladoc se remarcă prin extinderea accesului la îngrijirea medicală la nivel global și focalizarea pe telemedicină, în timp ce Zocdoc pune accent pe gestionarea eficientă a programărilor și concentrarea pe experiența pacientului în contextul întâlnirilor fizice.

Alegerea între cele două platforme depinde de preferințele individuale ale pacienților și de nevoile specifice de îngrijire medicală, întrucât fiecare dintre ele aduce contribuții valoroase la evoluția sănătății digitale.

## Concluzii intermediare la primul capitol

O platformă de telemedicină oferă o serie de beneficii și funcționalități importante atât pentru pacienți, cât și pentru profesioniștii din domeniul medical. Prin intermediul acestei platforme, pacienții pot avea acces la consultații medicale online, eliminând necesitatea de a se deplasa fizic la un cabinet medical. Acest lucru este deosebit de util pentru cei aflați în zone geografice izolate sau pentru cei cu dificultăți de mobilitate.

Un alt avantaj al telemedicinii constă în posibilitatea de a efectua diagnosticări și tratamente online. Medicii pot evalua simptomele și starea de sănătate a pacienților și pot prescrie tratamente sau oferi recomandări adecvate, reducând astfel timpul de așteptare pentru programări și facilitând accesul la îngrijirea medicală.

Platformele de telemedicină permit, de asemenea, monitorizarea continuă a pacienților, astfel încât medicii să poată urmări evoluția stării lor de sănătate și să ofere intervenții sau ajustări la tratament în timp real. Acest lucru contribuie la îmbunătățirea managementului bolilor cronice și la prevenirea complicațiilor.

Gestionarea eficientă a dosarelor medicale este un alt beneficiu al telemedicinii. Pacienții pot avea acces la istoricul lor medical, rezultatele testelor și alte informații relevante într-un mediu sigur și centralizat. Aceasta facilitează comunicarea și colaborarea între profesioniștii din domeniul medical și îmbunătățește calitatea îngrijirii oferite.

Pe lângă beneficiile pentru pacienți, telemedicina aduce și economii de timp și costuri reduse. Pacienții nu mai trebuie să petreacă timp și bani pentru deplasările la clinici sau spitale, iar acest lucru poate fi deosebit de important în cazul persoanelor cu mobilitate redusă sau în situații de urgență.

Totuși, un asistentul virtual dezvoltat pe baza tehnologiilor de învățare automată și oferirea de sfaturi pentru a preveni atacul de cord ar ajuta un medic să se concentreze pe pacienți care se află într-o stare gravă. De asemenea, un astfel de asistent virtual se va remarca ca o soluție inovatoare în contextul sănătății digitale, oferind un spectru larg de funcționalități care îmbunătățesc nu doar monitorizarea sănătății, ci și calitatea vieții utilizatorilor. Utilizând tehnologii avansate precum modelele de învățare automată, asistentul este capabil să analizeze datele individuale și să prezică riscurile de apariție a unui atac de cord în baza datelor introduse de utilizator. Această capacitate de a oferi alerte timpurii și de a recomanda măsuri preventive reprezintă o valoare inestimabilă în gestionarea sănătății personale.

Asistentul ar putea oferi și sfaturi personalizate pentru îmbunătățirea sănătății, adaptate nevoilor și obiectivelor individuale ale utilizatorilor. Aceste sfaturi pot include recomandări legate de alimentație, exerciții fizice, gestionarea stresului sau alte aspecte relevante pentru un stil de viață sănătos.

Prin monitorizarea continuă a parametrilor de sănătate, asistentul contribuie la menținerea unei supravegheri constante asupra stării de sănătate a utilizatorilor. Această funcționalitate nu doar informează utilizatorii despre modificările semnificative în parametrii lor de sănătate, ci îi și motivează să rămână angajați în îngrijirea personală.

Interfața intuitivă și ușor de utilizat a asistentului face ca accesarea și interacțiunea cu informațiile de sănătate să fie convenabile și eficiente pentru utilizatori. Astfel, asistentul devine un partener de încredere în gestionarea sănătății personale, facilitând luarea deciziilor informate și promovând un stil de viață sănătos și preventiv.

Integrarea unui asistent virtual într-o platformă de telemedicină aduce un nivel superior de accesibilitate și eficiență în îngrijirea medicală. Acest asistent virtual servește ca un partener inteligent pentru utilizatori, oferind o gamă largă de funcționalități esențiale.

În concluzie, integrarea unui asistent virtual într-o platformă de telemedicină reprezintă o evoluție esențială în furnizarea îngrijirii medicale, oferind accesibilitate, eficiență și suport continuu pentru îmbunătățirea sănătății și a experienței utilizatorilor.

# CADRUL FUNDAMENTAL TEORETIC PRIVIND DEZVOLTAREA UNUI ASISTENT VIRTUAL

Dezvoltarea unui asistent virtual pentru nevoile zilnice ale utilizatorului implică mai multe aspecte teoretice și practice.

Asistenții virtuali se bazează pe tehnologii de Inteligență Artificială (IA), cum ar fi procesarea limbajului natural (NLP) și învățarea automată (Machine Learning). Înțelegerea acestor concepte de bază este esențială. Pentru a dezvolta un asistent util, trebuie să înțelegem nevoile, preferințele și comportamentul utilizatorului.

NLP este o componentă centrală a unui asistent virtual. Acesta implică înțelegerea limbajului uman, extragerea informațiilor și generarea de răspunsuri semnificative. Abordări precum recunoașterea vocii și analiza sentimentelor sunt importante în NLP. De asemenea, un asistent trebuie să aibă acces la baze de date relevante și să poată gestiona informațiile corect. Acest lucru implică cunoașterea tehnologiilor și formelor de stocare a datelor.

Dezvoltarea modelelor de învățare automată este esențială pentru îmbunătățirea funcționalităților asistentului, cum ar fi înțelegerea intențiilor utilizatorului și adaptarea la nevoile acestuia. Pentru a fi adaptabil la necesitățile utilizatorului, este important cum asistentul interacționează cu utilizatorul. Interfața utilizatorului (UI) și experiența utilizatorului (UX) trebuie să fie proiectate pentru a oferi o experiență plăcută și eficientă. În plus, întrucât trăim în era tehnologiei, iar datele noastre sunt necesare pentru a utiliza un anumit software, asigurarea securității datelor și a confidențialității este de o importanță deosebită, având în vedere că asistenții virtuali pot avea acces la informații sensibile, iar respectarea normelor etice și reglementărilor este vitală, mai ales în ceea ce privește colectarea și stocarea datelor utilizatorilor.

În pofida faptului că un asistent virtual nu poate ajunge la ideal instantaneu în momentul creării, el este în proces de dezvoltare continuă care necesită teste și evaluări constante pentru a asigura performanța și utilitatea asistentului.

Acest cadru teoretic reprezintă o bază solidă pentru dezvoltarea unui asistent virtual pentru nevoile zilnice ale utilizatorului. Este important să se combine cunoștințele teoretice cu abilitățile practice de dezvoltare software și de gestionare a proiectelor pentru a crea un asistent eficient și util.

## Inteligența artificială

Inteligenta artificiala (IA) este o ramură a informaticii care se concentrează pe dezvoltarea de sisteme și tehnologii capabile să simuleze și să execute sarcini care, atunci când sunt efectuate de către oameni, necesită inteligență. Scopul inteligenței artificiale este să creeze mașini și sisteme care pot învăța, raționa, rezolva probleme și să ia decizii într-un mod similar cu oamenii.

Conceptul de inteligență artificială include o varietate de tehnici și abordări, cum ar fi:

1. **Prelucrarea limbajului natural (NLP):** acest domeniu implică dezvoltarea de tehnologii care permit computerelor să înțeleagă, să proceseze și să genereze limbaj uman, permițând interacțiuni mai naturale cu sistemele informatice.
2. **Învățarea automată (Machine Learning):** învățarea automată este o subdomeniu a IA care se concentrează pe dezvoltarea de algoritmi care permit computerelor să învețe din date și să facă prognoze sau să efectueze sarcini fără a fi programate explicit.
3. **Învățarea profundă (Deep Learning):** acesta este un subdomeniu al învățării automate care se bazează pe rețele neuronale artificiale profunde pentru a rezolva probleme complexe, cum ar fi recunoașterea imaginilor sau prelucrarea limbajului natural.

Inteligența artificiala a avut un impact semnificativ în diferite aspecte ale vieții moderne, de la medicină la tehnologie, industrie și multe altele. Dezvoltarea și cercetarea continuă în acest domeniu deschid drumuri noi pentru inovație și îmbunătățirea continuă a tehnologiilor bazate pe IA[9].

### Apariția și dezvoltarea intelegenței artificiale

Conceptul de inteligență artificială (IA) are rădăcini care se întind pe mai multe decenii și a cunoscut o evoluție semnificativă pe parcursul timpului. Aici este o sinteză a apariției inteligenței artificiale:

În deceniile anilor '40-'50, pionieri precum Alan Turing au formulat conceptul fundamental al Mașinii Turing, deschizând uși pentru dezvoltarea de mașini capabile să simuleze procesele de gândire umană. În paralel, Warren McCulloch și Walter Pitts au elaborat modele teoretice de neuroni artificiali, deschizând calea pentru dezvoltarea rețelelor neuronale artificiale, fundamentale în învățarea profundă.

În anii '50-'60, au apărut primele programe pentru jocuri precum șahul, care au concurat cu jucătorii umani. *The Logic Theorist*, dezvoltat de Allen Newell și Herbert A. Simon, este un exemplu notoriu[9].

În perioada anilor '60-'70, limbaje de programare precum LISP au fost create pentru a susține dezvoltarea inteligenței artificiale. Proiecte notabile au inclus *Shakey the Robot*, unul dintre primii roboți mobili care putea naviga în medii controlate și efectua sarcini simple[10].

În anii '80-'90, interesul pentru inteligența artificială a crescut semnificativ, aducând tehnologii precum sistemele expert care au fost folosite în domenii precum diagnosticarea medicală. IBM a dezvoltat computerul *Deep Thought* în 1980, care a fost unul dintre primele calculatoare capabile să concureze la nivel înalt în șah[11].

Începând cu anii 2000 și continuând până în prezent, învățarea automată (machine learning) a devenit un domeniu-cheie al inteligenței artificiale. Acest lucru a dus la dezvoltarea algoritmilor și tehnologiilor care au avansat semnificativ: recunoașterea vocală, viziunea artificială și prelucrarea limbajului natural. Tehnologii precum rețelele neuronale profunde (deep learning) au transformat capacitatea sistemelor de a înțelege și interpreta date complexe.

Astăzi, inteligența artificială are un impact profund într-o varietate de domenii, de la asistenții virtuali în telefoanele mobile la vehicule autonome și medicină. Deși drumul inteligenței artificiale a fost lung și plin de provocări, cercetarea și dezvoltarea continuă deschizând noi orizonturi de inovație și extind limitele cunoașterii și tehnologiei.

### Testul Turing

Testul Turing, propus de Alan Turing în 1950 într-un articol intitulat *Computing Machinery and Intelligence*, este o metodă de evaluare a capacității unei mașini de a exhiba comportament inteligent similar cu cel al unui om. Scopul fundamental al testului este de a determina dacă o mașină poate să gândească într-un mod care să fie indistinguibil de gândirea umană în cadrul unei conversații.

Testul Turing se desfășoară sub forma unui joc cu trei participanți: un om (numit și "cunoscător"), un operator uman și o mașină. Testul presupune un schimb de întrebări și răspunsuri într-un format de text scris.

1. **Participanții:**

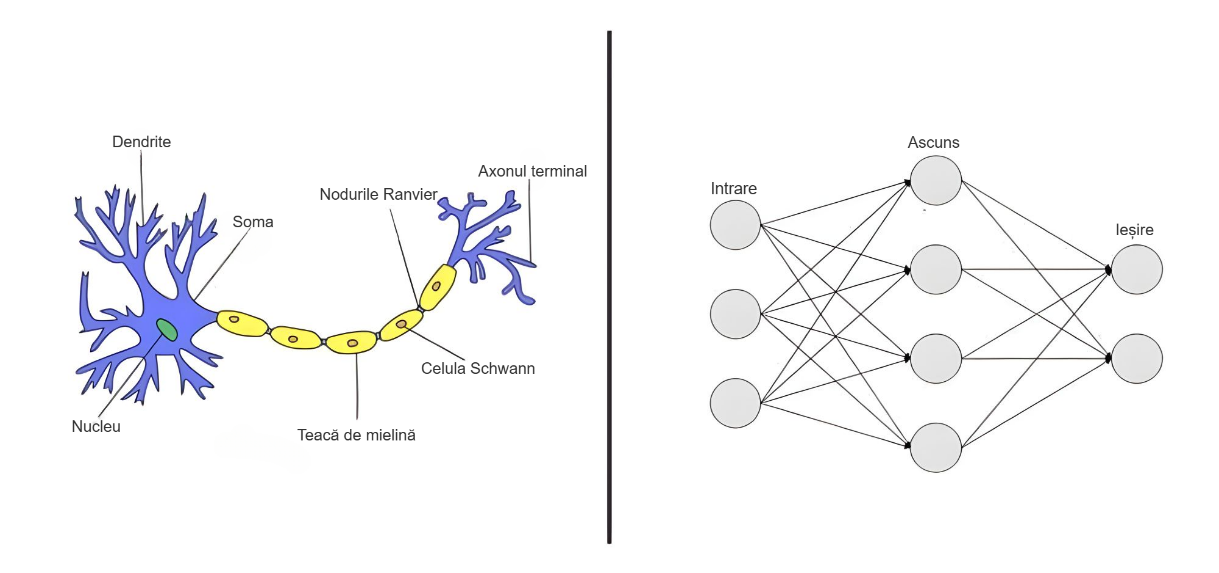
* *Cunoscătorul (evaluatorul)*: este un om care are rolul de a evalua conversația dintre operator și mașină.
* *Operatorul uman:* este un alt om care poate comunica cu cunoscătorul și cu mașina, dar este izolat de viziunea directă asupra mașinii. Operatorul are misiunea de a răspunde la întrebările cunoscătorului și de a se comporta în mod cât mai natural și inteligent posibil.
* *Mașina (subiectul testului):* mașina care este supusă testului trebuie să răspundă la întrebările și solicitările operatorului într-un mod care să o facă să pară cât mai umană posibil.

1. **Scopul testului**: scopul este de a evalua dacă mașina poate să răspundă la întrebările și solicitările operatorului astfel încât cunoscătorul să nu poată distinge între răspunsurile mașinii și cele ale unui om. Cu alte cuvinte, dacă mașina poate să înșele cunoscătorul, atunci poate fi considerată că a trecut testul și a demonstrat un comportament inteligent.
2. **Durata conversației**: testul se desfășoară într-o conversație de durată limitată, de obicei, în jur de cinci minute.
3. **Rezultate**: dacă cunoscătorul nu poate să distingă între răspunsurile mașinii și cele ale operatorului uman într-un procent semnificativ din conversații, se consideră că mașina a trecut testul.

Testul Turing a reprezentat o marcă importantă în dezvoltarea inteligenței artificiale, în special în dezvoltarea de agenți conversaționali sau asistenți virtuali. Cu toate acestea, trebuie să se sublinieze că testul Turing a fost criticat pentru că nu reprezintă neapărat o măsură adecvată a inteligenței și că există multiple limite și critici la adresa acestuia. Cunoașterea contextului și a abilităților specifice ale mașinii de către operator poate influența rezultatele testului. De aceea, de-a lungul timpului, au fost dezvoltate și alte metode de evaluare a inteligenței artificiale care trec dincolo de Testul Turing[12].

### Rețele neuronale

Rețelele neuronale artificiale (RNA), ca și oamenii, învață prin exemplu. O RNA este configurată pentru o aplicație specifică, cum ar fi recunoașterea modelelor sau clasificarea datelor, prin un proces de învățare. Învățarea în sistemele biologice implică ajustări ale sinapsei conexiuni sinaptice care există între neuroni. Acest lucru este valabil și în cazul RNA.



**Figura 2.1***Principiul de lucru a rețelei neuronale*

Dendritele sunt componente esențiale ale unui neuron(*figura 2.1*), reprezentând fibre ramificate care se extind din corpul celular sau soma. În corpul celular se găsesc nucleul și alte structuri cruciale pentru procesarea chimică și producția de neuro-transmițători. Aceste dendrite au rolul de a primi semnale sau impulsuri de la alți neuroni sau celule, formând astfel conexiuni sinaptice esențiale în rețeaua neuronală.

Pe de altă parte, axonul este o fibră singulară ce transportă informațiile de la soma la situsurile sinaptice ale altor neuroni, fie că acestea sunt dendrite sau soma ale altor celule. Cocoașa axonului reprezintă locul de însumare a informațiilor primite. Este important de menționat că influența colectivă a impulsurilor de la alți neuroni determină dacă un potențial de acțiune va fi inițiat la nivelul cocoașei axonale și apoi propagat de-a lungul axonului.

Mielina, alcătuită din celule grase, izolează axonul de radiațiile electrice, accelerând astfel transmiterea semnalelor. Nodurile lui Ranvier sunt spațiile dintre celulele tecii de mielină, unde semnalele săritor de la un spațiu la altul. Această structură a axonului, cu învelișul de mielină și nodurile lui Ranvier, contribuie la eficiența și rapiditatea transmiterii semnalelor electrice în cadrul sistemului nervos.

Sinapsa reprezintă punctul de legătură între doi neuroni sau între un neuron și o altă celulă, cum ar fi un mușchi sau o glandă. La nivelul acestor joncțiuni, are loc comunicarea electrochimică între neuroni, iar butoanele terminale ale axonului eliberează neurotransmițători, substanțe chimice care permit transmiterea semnalelor de la un neuron la altul.

Astfel, fluxul de informații într-o celulă neuronală implică interacțiuni complexe la nivelul dendritelor, soma, axonului și sinapselor, contribuind la funcționarea corectă a sistemului nervos și la realizarea proceselor cognitive și motorii în organism[13].

Neuronii artificiali sunt unitățile fundamentale ale rețelelor neuronale artificiale, modelate după funcționarea neuronilor biologici. În ciuda simplificărilor aduse, principiul de bază al acestora constă în prelucrarea informațiilor într-un mod similar cu neuronii naturali.

Acești neuroni artificiali au componentele lor esențiale, cum ar fi intrările, care corespund dendritelor neuronilor biologici, reprezentând semnalele sau datele de intrare. Funcția de activare decide, asemănător cu procesul de integrare din corpul celular al neuronului biologic, dacă neuronul artificial va genera un semnal de ieșire pe baza intrărilor primite. Greutățile, asemănătoare sinapselor, reprezintă ponderile asociate fiecărei intrări și sunt ajustate în timpul procesului de învățare al rețelei. Sumatorul ponderat calculează suma produselor dintre intrări și greutăți, similar procesului de însumare a semnalelor în cocoașa axonului neuronului biologic. Funcția de ieșire, asemănătoare cu procesul de eliberare a neurotransmițătorilor la nivelul butoanelor terminale, generează semnalul de ieșire al neuronului artificial.

În esență, neuronii artificiali sunt proiectați pentru a reproduce procesele de învățare și adaptare observate în creierul uman, facilitând astfel înțelegerea și procesarea eficientă a informațiilor în cadrul rețelelor neuronale artificiale.

### Tipuri de rețele neuronale

Rețelele neuronale sunt instrumente fundamentale în domeniul inteligenței artificiale, având diverse arhitecturi specializate pentru rezolvarea diferitelor tipuri de probleme.

FNN-urile, reprezintă cea mai simplă formă de arhitectură neuronală. Informația traversează aceste rețele într-o direcție unică, de la stratul de intrare, prin straturile ascunse, până la stratul de ieșire, fără a exista bucle sau conexiuni ciclice. Aceste rețele sunt des întâlnite în rezolvarea problemelor de clasificare și regresie, având ca obiectiv învățarea relațiilor dintre intrări și ieșiri.

RNN-urile se disting prin conexiunile ciclice între neuroni, permițând rețelei să memoreze informații într-un mod secvențial. Această arhitectură le face potrivite pentru manipularea datelor cu dependențe temporale sau secvențiale, cum ar fi limbajul natural sau serii temporale. Ele sunt folosite în aplicații precum traducerea automată, generarea de texte sau recunoașterea vorbirii.

CNN-urile sunt specializate în prelucrarea datelor spațiale, cum ar fi imaginile. Utilizează straturi de convoluție pentru a identifica caracteristici locale și păstrează aceleași ponderi pentru a identifica modele în întreaga imagine. Principalele domenii de aplicare includ clasificarea imaginilor, segmentarea semantică și recunoașterea de obiecte.

GAN-urile constau într-un generator care produce date noi și un discriminator care încearcă să distingă între datele generate și cele reale. Acest duel între generator și discriminator duce la o îmbunătățire continuă a calității datelor generate. Sunt utilizate pentru generarea de imagini realiste, sinteză de texte sau crearea de conținut multimedia, având un impact semnificativ în domeniul creativ.

LSTMs sunt o variantă specializată a RNN, proiectată pentru a gestiona problema dispariției gradienților întâlnită în rețelele neurale tradiționale. Datorită arhitecturii complexe, LSTMs sunt eficiente în gestionarea secvențelor de date lungi, fiind folosite în traducerea automată, generarea de texte și predicții bazate pe serii temporale.

Aceste tipuri de rețele neuronale operează pe principii specifice, adaptându-se nevoilor diverse din domeniul inteligenței artificiale.

## Aspecte etice în inteligența artificială

Aspectele etice ale inteligenței artificiale (IA) devin din ce în ce mai cruciale în contextul creșterii prevalenței acestei tehnologii în societate. În abordarea acestor aspecte, identificăm câteva puncte semnificative.

Un aspect important este confidențialitatea datelor, având în vedere capacitatea IA de a procesa și analiza cantități masive de informații. Acest fapt ridică întrebări fundamentale cu privire la protejarea datelor personale și necesitatea stabilirii unor standarde etice riguroase.

Discriminarea algoritmilor este o problemă sensibilă, întrucât algoritmii de IA pot reflecta și amplifica tendințele preexistente din datele pe care sunt antrenați. Combaterea acestor tendințe este crucială pentru asigurarea corectitudinii și imparțialității în rezultatele furnizate de sistemele bazate pe IA.

Impactul asupra locurilor de muncă este o preocupare etică majoră, deoarece automatizarea generată de IA poate afecta negativ locurile de muncă umane. Este esențial să se abordeze în mod etic tranziția către o economie digitală și să se sprijine adaptarea lucrătorilor la noile cerințe ale pieței.

Transparența și explicabilitatea în funcționarea algoritmilor de învățare automată reprezintă o altă dimensiune etică importantă. Utilizatorii trebuie să înțeleagă procesele decizionale ale acestor sisteme pentru a avea încredere în rezultatele și deciziile lor.

Securitatea cibernetică devine o prioritate, pe măsură ce IA devine parte integrantă în sistemele critice. Este esențial să se acorde atenție maximă securității pentru a preveni atacurile malefice asupra acestor sisteme.

Responsabilitatea și luarea deciziilor reprezintă un alt punct cheie, iar clarificarea responsabilității pentru deciziile luate de sistemele IA este esențială. Stabilirea de linii clare și modalități de aducere la răspundere a creatorilor și utilizatorilor acestor tehnologii este crucială într-un cadru etic.

Abordarea și soluționarea acestor aspecte etice sunt fundamentale pentru a asigura o dezvoltare și utilizare responsabilă a inteligenței artificiale, contribuind la beneficiul durabil al întregii societăți.

## Procesarea limbajului natural(NLP)

Preprocesarea textului reprezintă prima etapă a prelucrării limbajului natural (NLP) și este fundamentală pentru pregătirea și curățarea datelor textuale înainte de a putea fi analizate sau procesate de un model lingvistic. Ea este importantă deoarece contribuie la reducerea zgomotului și a ambiguității din datele textuale, permițând ulterior analiza și procesarea eficientă de către asistentul virtual. După ce textul a fost preprocesat, poate fi analizat sintactic, semantic, sau pot fi aplicate alte tehnici NLP pentru a extrage informații relevante și pentru a răspunde la cerințele utilizatorului în mod eficace.

### Curățarea Textului

Curățarea textului este o etapă esențială în preprocesarea datelor textuale și în procesarea limbajului natural (NLP). Această etapă implică eliminarea elementelor nedorite din text pentru a pregăti datele pentru analiză și interpretare. Eliminarea semnelor de punctuație, a caracterelor speciale și a simbolurilor, precum @, #, $, %, și altele, este o parte importantă a procesului. Aceste elemente, de obicei, nu adaugă informații semantice și pot crea zgomot în date.

De asemenea, cifrele pot fi eliminate, în special dacă analiza se concentrează pe text și nu pe valori numerice. Aceasta ajută la reducerea complexității și la uniformizarea textului. Simbolurile de linie nouă, spațiile multiple sau alte caractere de formatare pot fi înlocuite cu spații goale sau eliminate pentru a asigura uniformitatea și pentru a evita erorile în analiză. În unele cazuri, texte pot conține caractere Unicode neobișnuite, care pot fi înlocuite cu echivalente mai simple sau eliminate pentru o prelucrare mai eficientă.

Curățarea textului are ca scop reducerea dimensiunii datelor, îmbunătățirea calității analizei și asigurarea că datele sunt pregătite corespunzător pentru etapele ulterioare de prelucrare a limbajului natural. Aceasta contribuie la obținerea unor rezultate mai precise și semnificative în cadrul proiectelor NLP[14].

### Transformarea în Litere Mici

Transformarea textului în litere mici (minuscule) este o etapă esențială în preprocesarea datelor textuale și în procesarea limbajului natural (NLP). Această etapă are ca scop uniformizarea textului, astfel încât cuvintele să fie tratate ca fiind identice indiferent de cazul literelor folosite. Aceasta este o operațiune simplă, dar esențială în pregătirea datelor textuale pentru analiză și interpretare.

Aceasta ajută la eliminarea ambiguităților cauzate de utilizarea diferită a literelor mari și mici în cuvinte. De exemplu, cuvintele "casa" și "Casa" vor fi tratate ca fiind identice după transformarea în litere mici, eliminând astfel ambiguitățile.

De asemenea, permite compararea simplă a cuvintelor atunci când se realizează comparații sau căutări în text, deoarece nu se mai ține cont de cazul literelor. Această etapă are un impact semnificativ în reducerea redundanței în date, economisind spațiu de stocare și asigurând că analiza textului se face în mod eficient și coerent, indiferent de modul în care au fost introduse literele.

În final, transformarea textului în litere mici este o practică comună și utilă în NLP, având rolul de a pregăti datele textuale pentru analize semantice și sintactice mai precise, astfel încât asistentul virtual să poată înțelege și răspunde la cerințele utilizatorului în mod corespunzător[14].

### Eliminarea Stop Words

Eliminarea cuvintelor de oprire (stop words) reprezintă o etapă esențială în preprocesarea datelor textuale în procesarea limbajului natural (NLP). Cuvintele de oprire sunt cuvinte comune și frecvent utilizate în limbaj precum "și," "sau," "pentru," "cu," "la," etc., care, de obicei, nu adaugă valoare semantică sau semnificativă analizei textului. Eliminarea lor contribuie la reducerea dimensiunii datelor și la concentrarea asupra cuvintelor cheie.

Eliminarea cuvintelor de oprire se efectuează în text după ce acesta a fost curățat de semne de punctuație și alte elemente nedorite. Scopul acestei eliminări constă în reducerea zgomotului în date și concentrarea pe cuvintele care aduc semnificație semantică. Cuvintele de oprire, în general, nu contribuie semnificativ la înțelegerea conținutului textului și pot fi ignorate în majoritatea analizelor NLP.

Cu toate acestea, există excepții, și anume în situații specifice. De exemplu, în analiza de sentiment, cuvinte de oprire precum "nu" sau "foarte" pot avea un impact semnificativ asupra interpretării sentimentului și pot fi păstrate în unele situații pentru a asigura o analiză corectă.

Eliminarea cuvintelor de oprire este o etapă de bază în preprocesarea textului și depinde de specificul proiectului NLP. Aceasta ajută la pregătirea datelor pentru analize semantice mai precise și pentru ca asistentul virtual să poată înțelege și interpreta corect conținutul textului în contextul utilizatorului[14].

### Stemming și Lemmatizare

Procesarea limbajului natural (NLP) implică o serie de tehnici și etape cruciale pentru a pregăti textul pentru analiză și interpretare. În acest context, două tehnici esențiale sunt "stemming" și "lemmatizarea," care au ca scop simplificarea cuvintelor pentru a facilita analiza în cadrul proiectelor NLP.

**Stemming** este o tehnică care constă în reducerea unui cuvânt la forma sa de bază sau *stem*. De exemplu, cuvintele *alergam*, *alergați*, și *alergare* pot fi reduse la același *stem* de bază, *alerg*. Algoritmele de stemming utilizează reguli de tăiere pentru a realiza această simplificare. Stemming este util atunci când se dorește o preprocesare simplă a textului, dar poate duce la forme trunchiate sau nevalide ale cuvintelor în anumite contexte.

**Lemmatizarea**, pe de altă parte, este o tehnică mai sofisticată. Aceasta implică reducerea unui cuvânt la forma sa de bază sau *lemă*, care este un cuvânt valid din punct de vedere semantic. De exemplu, cuvintele *alergam*, *alergați*, și *alergare* pot fi reduse la forma lor de bază *alerga*. Pentru a realiza lemmatizarea, se folosesc resurse lexicale sau dicționare de cuvinte pentru a găsi forma corectă de bază a cuvântului. Lemmatizarea este preferată atunci când se dorește o preprocesare mai precisă a textului și conservarea semanticii cuvintelor. Aceasta evită producerea de forme trunchiate sau nevalide ale cuvintelor.

Alegerea între stemming și lemmatizare depinde de obiectivele specifice ale proiectului NLP. Stemming este mai simplu, dar mai puțin precis, în timp ce lemmatizarea este mai sofisticată și mai precisă, dar poate necesita resurse lexice. Ambele tehnici au locul lor în preprocesarea textului, în funcție de contextul și scopul proiectului[14].

### Eliminarea Elementelor Irrelevant

Eliminarea elementelor nedorite sau irelevante din text reprezintă o etapă esențială în preprocesarea datelor pentru proiectele de procesare a limbajului natural (NLP). Această etapă se adaptează în funcție de obiectivele specifice ale analizei și poate include eliminarea unor elemente precum adresele de e-mail, URL-urile sau numerele de telefon.

Pentru început, definirea clară a elementelor considerate irelevante este fundamentală. Adică, trebuie de stabilit care sunt elementele pe care dorim să le eliminăm din text în funcție de contextul și obiectivele analizei.

După ce definim elementele irelevante, următorul pas constă în identificarea și extragerea acestora din text. Acest lucru se poate face folosind tehnici precum expresiile regulate sau metode specifice de procesare de text. De exemplu, pentru găsirea adreselor de e-mail, putem crea șabloane de căutare care să identifice adresele de tipul "adresa@domeniu.com."

După identificarea acestor elemente irelevante, avem două opțiuni principale: să le înlocuim cu token-uri generice, precum "EMAIL" sau "URL," sau să le eliminăm complet din text. Alegerea între aceste două opțiuni depinde de contextul și scopul analizei.

Scopul final al eliminării elementelor irelevante este concentrarea asupra conținutului semnificativ din text. De exemplu, în analiza de sentiment, eliminarea adresele de e-mail sau URL-urile poate ajuta la focalizarea pe exprimarea sentimentelor în text.

Documentarea clară a etapelor de eliminare a elementelor irelevante și urmărirea modificărilor aduse textului sursă sunt esențiale pentru menținerea transparenței și reproducibilității analizei[14].

### Tokenizarea

Tokenizarea este o etapă fundamentală în procesarea limbajului natural (NLP) și reprezintă procesul de divizare a unui text în unități mai mici, numite "tokenuri." Aceste tokenuri pot fi cuvinte individuale, grupuri de cuvinte sau simboluri, și reprezintă unitățile de bază cu care se lucrează în analiza și procesarea textului.

Într-un text, fraza *Am citit o carte interesantă* ar fi tokenizată în următoarele tokenuri: ["Am", "citit", "o", "carte", "interesantă"].

Tokenizarea ia în considerare semnele de punctuație, cum ar fi virgulele sau punctele, și spațiile pentru a separa cuvintele și elementele semantice din text. Ea este importantă în NLP deoarece reprezintă prima etapă în preprocesarea textului. Tokenurile devin unitățile de bază pentru analize ulterioare, cum ar fi analiza sintactică, analiza semantică, generarea de text sau analiza sentimentelor.

Procesul de tokenizare poate varia în funcție de limba în care se lucrează, deoarece regulile de despărțire a cuvintelor și semnelor de punctuație sunt diferite. Sunt disponibile biblioteci și instrumente specifice pentru tokenizarea textului în mai multe limbi. Tokenizarea poate prezenta ambiguități în cazul cuvintelor compuse sau al limbajului figurat. De exemplu, cuvântul *mergerea* poate fi divizat în *merg* + *ea* sau *mergere* + *a*, în funcție de context. De obicei, cuvintele de oprire (stop words), precum *și*, *sau*, *pentru*, sunt eliminate în timpul tokenizării, deoarece nu adaugă semnificație semantică.

Tokenizarea reprezintă o etapă esențială în procesarea limbajului natural, deoarece transformă textul într-o structură mai ușor de analizat și de lucrat. Aceasta oferă baza pentru analiza și interpretarea semnificației textului și reprezintă un pas cheie în dezvoltarea asistenților virtuali și a altor aplicații NLP[14].

## Învățarea automată.

Învățarea automată (Machine Learning, ML) este o ramură a inteligenței artificiale (IA) care se concentrează pe dezvoltarea de algoritmi și modele care permit sistemelor de calculatoare să învețe și să se îmbunătățească în mod automat din experiență. Învățarea automată este utilizată într-o varietate de domenii, inclusiv procesarea limbajului natural (NLP), viziune artificială, recunoașterea vocală, recomandarea de conținut, diagnoza medicală și multe altele.

### Date și Experiență

Datele reprezintă materia primă a învățării automate. Acestea pot fi obținute dintr-o varietate de surse, inclusiv senzori, baze de date, imagini, texte și multe altele. Calitatea și cantitatea datelor sunt cruciale pentru antrenarea unui model precis. Datele trebuie să fie complete, corecte și relevante pentru problema pe care o soluționăm.

Experiența constă în procesul de antrenare a modelului de învățare automată pe baza datelor. Modelul învață să identifice pattern-uri, reguli și relații în date, ajustând parametrii săi pentru a se potrivi cât mai bine cu datele de antrenament. Evaluarea modelului se face prin compararea predicțiilor sale cu datele de testare sau de validare.

În funcție de performanța modelului, acesta poate necesita ajustări și optimizări pentru a atinge rezultatele dorite. Procesul de învățare automată este unul continuu, cu modele care se pot actualiza pe măsură ce apar date noi sau se schimbă cerințele.

### Antrenarea Modelului

Antrenarea modelului în cadrul procesului de învățare automată reprezintă un pas esențial. Acesta presupune utilizarea datelor pentru a instrui modelul astfel încât să poată învăța și să fie capabil să facă prognoze sau să efectueze sarcini specifice. Procesul de antrenare implică găsirea de pattern-uri, reguli sau relații în datele disponibile, precum și ajustarea parametrilor modelului pentru a se potrivi cu datele de antrenament.

În timpul antrenării, modelul învață să facă conexiuni între diferite caracteristici sau elemente din setul de date. Această înțelegere se bazează pe algoritmii specifici utilizați, care pot varia de la rețele neuronale profunde la arbori de decizie sau alți algoritmi de învățare automată.

De asemenea, este important să se menționeze că calitatea datelor de antrenament joacă un rol crucial în succesul procesului de antrenare. Datele de antrenament ar trebui să fie reprezentative pentru problema sau sarcina pe care modelul trebuie să o rezolve. Dacă datele de antrenament sunt incomplete sau incorecte, modelul poate să nu fie capabil să facă prognoze precise sau să ia decizii adecvate.

Pe măsură ce procesul de antrenare progresează, se urmărește obținerea unui model care să aibă o performanță cât mai bună pe datele de antrenament și să fie capabil să generalizeze la date noi, nevăzute anterior. Astfel, antrenarea modelului este o etapă esențială în dezvoltarea soluțiilor bazate pe învățarea automată și necesită expertiză în domeniul specific al problemei pe care modelul urmează să o rezolve.

### Deep Learning

Deep learning este un subdomeniu a învățării automate care se concentrează pe utilizarea rețelelor neuronale artificiale profunde pentru a rezolva probleme complexe. Aceste rețele, inspirate de funcționarea creierului uman, sunt alcătuite din straturi succesive de neuroni artificiali și au capacitatea de a învăța și extrage caracteristici din date prin intermediul acestor straturi.

Un aspect cheie al deep learning este capacitatea sa de a identifica și de a interpreta caracteristici complexe din datele de intrare. Acest lucru a făcut posibilă dezvoltarea aplicațiilor precum recunoașterea vocală, prelucrarea limbajului natural și multe altele. Prin intermediul acestor aplicații, sistemele pot analiza imagini, texte sau sunete și pot lua decizii sau face predicții într-un mod similar cu modul în care o persoană ar face-o.

Pentru antrenarea modelelor deep learning, este necesar un volum mare de date și resurse computaționale semnificative. Cu toate acestea, rezultatele obținute pot fi impresionante. De exemplu, în domeniul viziunii artificiale, rețelele neuronale profunde pot recunoaște obiecte în imagini cu o precizie notabilă, ceea ce le face utile în aplicații precum vehicule autonome sau securitatea video.

Deep learning este în continuă evoluție, cu cercetări și dezvoltări constante care vizează îmbunătățirea performanței și aplicabilității acestor rețele. Această tehnologie reprezintă un domeniu promițător în dezvoltarea inteligenței artificiale și are potențialul de a aduce inovații semnificative într-o gamă variată de domenii[15].

### Învățarea Continuă

Învățarea continuă reprezintă o parte fundamentală a dezvoltării în domeniul inteligenței artificiale și al învățării automate. Acest concept se referă la capacitatea sistemelor sau modelelor de a-și îmbunătăți performanța pe măsură ce intră în contact cu date noi și experimentează noi situații. Acest proces este esențial pentru menținerea relevanței și eficacității modelelor de învățare automată pe termen lung.

Inspirată de modul în care oamenii învață și se adaptează, învățarea continuă implică două aspecte cheie:

1. **Actualizarea modelului**: pe măsură ce noi date devin disponibile, modelele de învățare automată pot fi reantrenate sau actualizate pentru a integra aceste date noi. Acest proces ajută la menținerea actualității și a preciziei modelelor.
2. **Adaptarea la schimbări**: în anumite cazuri, mediul sau cerințele pot suferi modificări semnificative. Modelele de învățare automată trebuie să fie capabile să se adapteze la aceste schimbări pentru a continua să ofere soluții eficiente.

Un exemplu concret de învățare continuă este reprezentat de asistenții vocali precum Siri[16]. Aceste sisteme trebuie să fie actualizate în mod regulat pentru a răspunde la întrebări sau comenzi noi, să înțeleagă evoluțiile limbajului și să se adapteze la diverse contexte culturale și sociale.

Învățarea continuă este un element cheie în dezvoltarea modelelor de învățare automată, permițându-le să rămână relevante și eficiente în fața datelor și cerințelor în continuă schimbare. Acest proces esențial contribuie la asigurarea faptului că modelele de inteligență artificială pot funcționa eficient pe termen lung într-o varietate de aplicații și domenii. un rol crucial în dezvoltarea asistenților virtuali pentru necesitățile cotidiene ale utilizatorilor. Ea permite asistenților să înțeleagă și să răspundă la limbajul natural, să ofere recomandări personalizate și să îmbunătățească experiența utilizatorului în mod constant[17].

## Instrumente utilizate

Procesul de selectare a instrumentelor pentru dezvoltarea unui software constituie un pilon esențial pentru atingerea succesului proiectului și pentru obținerea unei aplicații de calitate. În acest sens, este crucial să avem o înțelegere clară a cerințelor proiectului, aceasta fiind baza unei alegeri informate. Definirea obiectivelor și funcționalităților necesare ne ghidează către instrumentele potrivite pentru a satisface optim aceste cerințe.

Tipul de aplicație reprezintă un alt aspect crucial, fie că este vorba de o aplicație web, mobilă sau desktop. Selectarea instrumentelor specializate pentru tipul respectiv de dezvoltare poate aduce beneficii semnificative în ceea ce privește eficiența și performanța proiectului. Comunitatea și suportul asociate instrumentelor sunt esențiale pentru a rezolva problemele întâmpinate și pentru a ține pasul cu noile dezvoltări. Instrumentele cu o comunitate activă și documentație bogată facilitează integrarea și rezolvarea problemelor în mod eficient.

Aspectele de performanță și scalabilitate sunt fundamentale în dezvoltarea software-ului, mai ales în cazul proiectelor cu potențial de creștere în complexitate și dimensiune. Instrumentele selectate trebuie să gestioneze cerințele de performanță și să ofere o scalabilitate adecvată pentru a face față evoluțiilor proiectului. Integrarea cu alte servicii sau instrumente este adesea necesară, iar asigurarea că instrumentele alese permit o integrare facilă și eficientă cu alte componente este esențială pentru funcționalitatea generală a aplicației.

În concluzie, fiecare proiect este unic, iar procesul de selecție a instrumentelor trebuie adaptat la cerințele specifice și la contextul dezvoltării software pentru a asigura succesul și durabilitatea proiectului.

### UML

În decursul anilor 1990, a apărut o necesitate crescândă de a avea o metodă standardizată pentru a modela și documenta sistemele software. Istoria UML implică colaborarea unor figuri-cheie din comunitatea ingineriei software.

În prima jumătate a anilor 1990, Grady Booch, James Rumbaugh și Ivar Jacobson dezvoltau independent propriile metode de modelare orientată pe obiect. Booch a creat Metoda Booch, Rumbaugh a dezvoltat Tehnica de Modelare a Obiectelor (OMT), iar Jacobson a creat Ingineria Software Orientată pe Obiect (OOSE).

În 1994, cei trei metodologi s-au unit sub auspiciile Rational Software Corporation, cu scopul de a crea un limbaj de modelare unificat care să combine cele mai bune aspecte ale abordărilor individuale. Rezultatul a fost Metoda Unificată (UM), o combinație a metodelor lui Booch, Rumbaugh și Jacobson.

În 1995, metoda unificată a evoluat, iar echipa a decis să o supună la standardizare către Object Management Group (OMG). OMG este un consorțiu internațional care dezvoltă și menține specificații pentru industria calculatoarelor.

În ianuarie 1996, OMG a adoptat Metoda Unificată ca bază pentru Limbajul Unificat de Modelare (UML). Prima versiune a UML, cunoscută sub denumirea de UML 1.0, a fost lansată.

În perioada 1997-1999, UML a trecut prin mai multe revizii pentru a aborda feedback-ul și a îmbunătăți limbajul. În acest interval, UML a câștigat o acceptare largă în industrie și a fost adoptată de diverse organizații și companii[18].

În 2005,a fost introdus UML 2.0, care a inclus caracteristici noi și îmbunătățiri. Noua versiune a avut ca scop îmbunătățirea clarității și preciziei diagramei UML și abordarea nevoilor în evoluție ale comunității de dezvoltare software.

Au fost lansate versiuni ulterioare, precum UML 2.1, UML 2.2, și altele, cu îmbunătățiri și adăugiri incrementale.

De-a lungul istoriei sale, UML a fost continuu rafinat și extins, cu contribuții din partea practicienilor și a cadrelor academice. OMG continuă să supravegheze evoluția UML, asigurându-se că rămâne relevant și eficient pentru modelarea și proiectarea sistemelor software. UML a devenit un standard larg adoptat în industria dezvoltării software, jucând un rol crucial în activitățile de comunicare și proiectare[19].

#### Diagrame structurale

Diagrama claselor este un element fundamental al sistemelor software orientate pe obiecte și este cea mai utilizată în UML. Ea descrie structura statică a unui sistem, prezentând clasele, metodele și atributele acestora, facilitând identificarea relațiilor dintre clase sau obiecte.

Diagrama structurii compozite este utilizată pentru a reprezenta structura internă a unei clase și interacțiunile sale cu alte părți ale sistemului. Ea evidențiază relațiile dintre părți și configurarea lor, influențând comportamentul clasificatorului.

Diagrama obiectului reprezintă o captură de ecran a instanțelor dintr-un sistem și relațiile dintre acestea. Acest tip de diagramă se concentrează asupra comportamentului la nivelul instanțelor de clasă, diferind de diagrama de clasă care descrie clasificatorii și relațiile lor.

Diagrama componentelor ilustrează modul în care sunt organizate componentele fizice ale unui sistem, oferind detalii de implementare și descriind relațiile structurale între elementele software.

Diagrama de implementare se referă la hardware-ul și software-ul unui sistem, evidențiind componentele hardware existente și componentele software care rulează pe acestea. Aceasta arată arhitectura sistemului și distribuția artefactelor software pe ținte distribuite.

Diagrama pachetelor descrie organizarea pachetelor și a elementelor acestora, prezentând dependențele dintre diverse pachete și compoziția internă a acestora. Pachetele ajută la organizarea diagramele UML în grupuri semnificative și facilitează înțelegerea diagramei[19].

#### Diagrame de comportament

O diagramă de stare este utilizată pentru a reprezenta starea unui sistem sau a unei părți a acestuia la anumite momente finite de timp. Este o modalitate de a ilustra comportamentul dinamic al unei clase în răspuns la schimbări de stimuli externi, folosind tranziții între stări.

Diagramele de activitate sunt utilizate pentru a ilustra fluxul de control într-un sistem și pot fi aplicate pentru a descrie pașii implicați în realizarea unui caz de utilizare. Aceste diagrame modelează activități secvențiale și concurente, concentrându-se pe fluxul și secvența acțiunilor care au loc. Ele oferă o reprezentare vizuală a fluxurilor de lucru și pot descrie cauzele evenimentelor.

Diagrama cazurilor de utilizare este folosită pentru a descrie funcționalitatea unui sistem sau a unei părți a acestuia. Aceasta evidențiază cerințele funcționale și interacțiunile cu agenți externi, cunoscuți sub numele de actori. Aceste diagrame oferă o imagine generală a modului în care sistemul poate fi utilizat, fără a intra în detalii de implementare.

Diagrama secvențială descrie interacțiunile dintre obiecte într-o ordine secvențială, evidențiind modul în care aceste interacțiuni au loc. Este utilizată pentru a documenta și înțelege cerințele pentru sisteme noi sau existente, concentrându-se pe relațiile și ordinea obiectelor.

Diagrama de comunicare (sau diagrama de colaborare în versiunile anterioare ale UML) se axează pe mesajele schimbate între obiecte, concentrându-se în primul rând pe obiecte și relațiile lor. Aceasta oferă o perspectivă asupra interacțiunilor dintre obiecte într-o formă mai liberă decât diagramele secvențiale.

Diagrama de sincronizare este o variantă specială a diagramei secvențiale, utilizată pentru a descrie comportamentul obiectelor într-un interval de timp. Ea evidențiază constrângerile de timp și durată care influențează stările și comportamentul obiectelor.

Diagrama de prezentare generală a interacțiunii modelează secvențe de acțiuni pentru a simplifica interacțiunile complexe în apariții mai simple, fiind o combinație între diagramele de activitate și cele secvențiale[19].

### Python

Python este un limbaj de programare creat de Guido van Rossum și lansat inițial în 1991. A fost conceput cu scopul de a oferi o sintaxă simplă și ușor de înțeles, încurajând dezvoltatorii să scrie cod clar și legibil. Datorită caracteristicilor sale, Python a devenit rapid unul dintre cele mai populare limbaje de programare din lume, fiind utilizat într-o varietate largă de domenii, inclusiv dezvoltarea web, analiza datelor, învățarea automată și inteligența artificială.

Ecosistemul Python este extrem de bogat, cu o multitudine de biblioteci și framework-uri care extind funcționalitățile sale în diverse domenii. De exemplu, Django este folosit pentru dezvoltarea rapidă a aplicațiilor web, NumPy și Pandas sunt esențiale pentru analiza datelor, în timp ce TensorFlow și PyTorch sunt preferate în domeniul învățării automate.

Cu toate acestea, Python nu este fără critici sau provocări. Una dintre criticile aduse este legată de performanța sa în sarcini intensive din punct de vedere al resurselor. Cu toate acestea, această problemă a fost abordată prin integrarea cu alte tehnologii și librării specializate.

Perspectivele pentru viitorul Pythonului sunt promițătoare, în special în ceea ce privește domeniile inteligenței artificiale și învățării automate. Dezvoltările continue în acest sens, împreună cu suportul unei comunități active și inovatoare, mențin relevanța și interesul pentru Python pe termen lung[20].

### Streamlit

Streamlit este o bibliotecă Python populară și eficientă pentru crearea rapidă a aplicațiilor web interactive. Lansată în 2019, această bibliotecă a devenit rapid preferata dezvoltatorilor datorită simplității și puterii sale în generarea interfețelor de utilizator intuitive pentru analiza datelor și vizualizarea rezultatelor.

Principala caracteristică a Streamlit este simplitatea utilizării sale. Dezvoltatorii pot crea aplicații web complexe fără a avea nevoie de cunoștințe avansate de dezvoltare web. Cu ajutorul Python și a unor funcții simple, se pot construi interfețe de utilizator interactive și atrăgătoare.

O altă calitate esențială a Streamlit este capacitatea sa de a afișa date și rezultatele analizelor în mod vizual. Graficele, tabelele și alte elemente vizuale sunt integrate ușor în aplicații, facilitând înțelegerea și explorarea datelor pentru utilizatori.

Streamlit oferă, de asemenea, o integrare fluentă cu alte biblioteci Python folosite în analiza datelor, precum Pandas, Matplotlib, Plotly sau Altair. Această integrare simplifică adăugarea funcționalităților complexe în aplicații și extinde posibilitățile de vizualizare și analiză a datelor.

Aplicațiile create cu Streamlit se actualizează dinamic pe măsură ce utilizatorii interacționează cu ele, fără a fi nevoie de refresh manual al paginii. Aceasta oferă o experiență interactivă plăcută și eficientă pentru utilizatori.

Cu un suport puternic pentru interacțiuni complexe, cum ar fi filtre, selecții multiple sau actualizări în timp real ale vizualizărilor, Streamlit este o soluție versatilă pentru dezvoltarea aplicațiilor web de analiză a datelor.

Având o comunitate activă de dezvoltatori și utilizatori, Streamlit continuă să se dezvolte și să se îmbunătățească, oferind soluții tot mai bune și mai flexibile pentru crearea de aplicații web interactive de succes în domeniul analizei datelor[21].

### Anaconda Navigator

Anaconda Navigator este o aplicație desktop dezvoltată de Anaconda, care oferă un mediu integrat pentru gestionarea mediilor de lucru Python și a pachetelor asociate. Prin intermediul său, utilizatorii pot crea, gestiona și lucra cu mediile de lucru Python, configurându-le cu pachete și biblioteci specifice pentru proiectele lor.

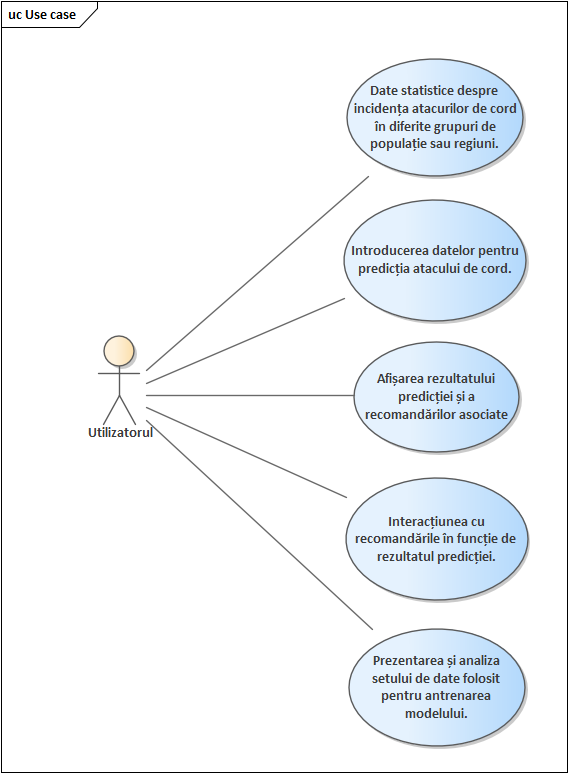
Unul dintre aspectele importante ale Anaconda Navigator este interfața grafică intuitivă și ușor de navigat. Aceasta face ca gestionarea mediilor de lucru și a pachetelor să fie accesibilă chiar și pentru cei mai puțin experimentați utilizatori. De asemenea, oferă acces rapid la diverse unelte și aplicații utile pentru dezvoltarea și analiza datelor în Python, precum Jupyter Notebook, Spyder sau RStudio.

Navigator-ul Anaconda permite, de asemenea, utilizatorilor să gestioneze kernel-urile asociate cu mediile de lucru Python, permițându-le să ruleze și să testeze codul în diferite configurații și versiuni ale limbajului. Astfel, utilizatorii pot lucra cu mai multe medi de lucru în paralel și pot testa codul în diferite contexte.

Prin actualizările regulate și suportul activ din partea comunității, Anaconda oferă utilizatorilor acces la cele mai recente funcționalități și îmbunătățiri. Este o unealtă esențială pentru dezvoltatorii și analiștii care lucrează cu Python și au nevoie de un mediu integrat și eficient pentru gestionarea mediilor de lucru și a pachetelor asociate.

## Modelarea conceptului sistemului

Diagrama *cazurilor de utilizare* în UML propusă prezintă o aplicație ce are un sistem de asistență virtuală conceput pentru optimizarea experienței pacienților în domeniul sănătății(*figura 2.2*). Această reprezentare vizuală a sistemului este crucială pentru înțelegerea interacțiunilor dintre actorul implicat și cazurile de utilizare asociate.



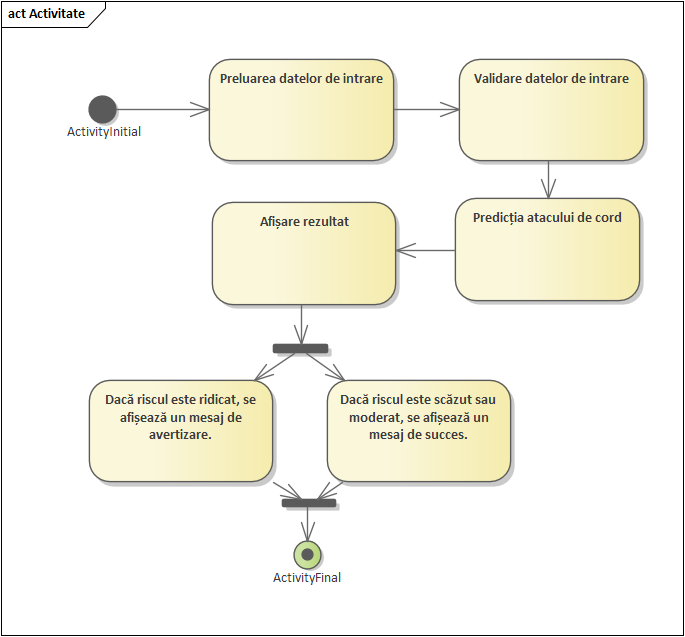
**Figura 2.2***Diagrama Use Case*

Examinând diagrama cazurilor de utilizare, actorul este desemnat ca utilizatorul. Cazurile de utilizare, care reprezintă acțiuni pe care sistemul le poate efectua pentru actor, sunt enumerate după cum urmează:

* Date statistice despre incidența atacurilor de cord în diferite grupuri de populație sau regiuni;
* Introducerea datelor pentru predicția atacului de cord;
* Afișarea rezultatului predicție și a recomandărilor asociate.
* Interacțiunea cu recomandările în funcție de rezultatul predicție;
* Prezentarea și analiza setului de date folosit pentru antrenarea modelului.

În esență, diagramele UML a cazurilor de utilizare stabilesc o bază pentru înțelegerea cerințelor sistemului prin reprezentarea funcționalităților sistemului din perspectiva actorilor săi. Acestea colaborează cu alte diagrame UML, cum ar fi diagramele de secvență și diagramele de activitate, pentru a oferi o imagine cuprinzătoare a comportamentului sistemului.

Diagrama UML propusă modelează un sistem de asistență virtuală bazat pe un model de învățare automată dedicat optimizării experienței pacienților în cadrul sistemului de sănătate(*figura 2.3*). Scopul fundamental al acestui sistem este de a sprijini pacienții în gestionarea eficientă a stării lor de sănătate, contribuind la crearea unei noi posibilități de diagnosticare.



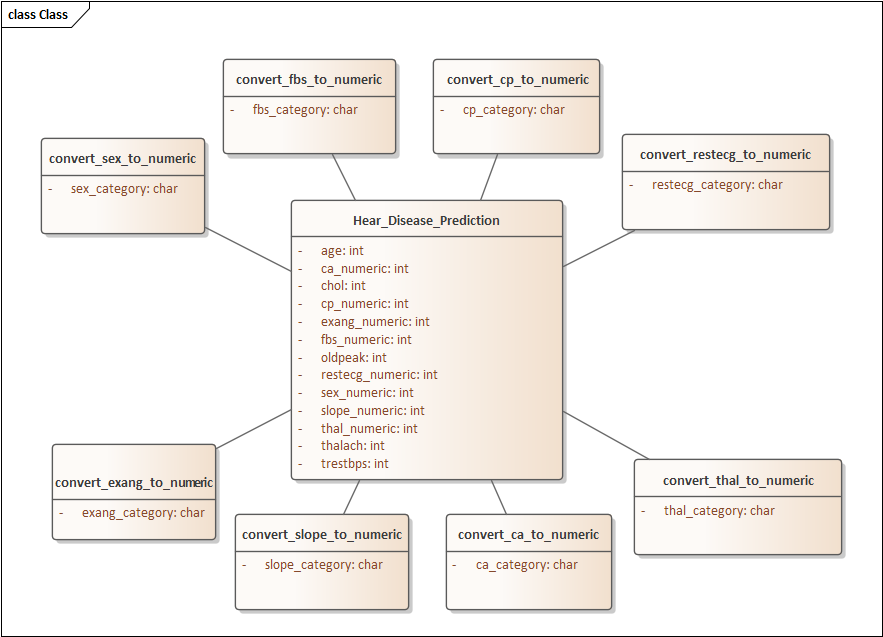
**Figura 2.3***Diagrama de activități*

Diagrama UML prezentată în figura 2.3 ilustrează fluxul de lucru al unui sistem de predicție a riscului de atac de cord. În continuare sunt prezentate activitățile:

* Preluarea datelor de intrare: această activitate colectează datele necesare pentru a calcula riscul de atac de cord. Datele sunt introduse manual de către utilizator.
* Validare date de intrare: asigură corectitudinea și completitudinea datelor introduse. Datele incomplete sau inconsistente pot duce la rezultate eronate.
* Predicția atacului de cord: această activitate utilizează un model de predicție pentru a calcula riscul de atac de cord al utilizatorului. Modelul este bazat pe algoritmi de învățare automată.
* Afișare rezultat: prezintă rezultatul predicției utilizatorului. Rezultatul este un mesaj de avertizare (risc ridicat) sau un mesaj de succes (risc scăzut sau moderat). Mesajul include, de asemenea, sfaturi pentru reducerea riscului de atac de cord sau menținerea unui stil de viață sănătos.

Diagrama de activități(*figura 2.3*) prezentată oferă o reprezentare clară și concisă a fluxului de lucru al unui sistem de predicție a riscului de atac de cord.

Diagrama *de clasă* în limbajul UML oferă o reprezentare structurată a unui sistem de asistență virtuală, concentrat pe îmbunătățirea experienței pacienților în cadrul domeniului sănătății(*figura 2.4*). Aceasta ilustrează clasa principală, relațiile dintre clasă și metode. De asemenea, ilustrează și modul în care aplicația interacționează cu utilizatorii.



**Figura 2.4***Diagrama de clasă*

Diagrama UML prezintă clasa numită HeartDiseasePrediction. Această clasă este responsabilă pentru prezicerea riscului de boli de inimă a unui utilizator.

Clasa are următoarele atribute:

* age: vârsta pacientului;
* sex\_category: sexul pacientului (bărbat sau femeie);
* cp\_category: tipul de durere toracică experimentată de pacient;
* trestbps: tensiunea arterială sistolică a pacientului în repaus;
* chol: nivelul de colesterol al pacientului;
* fbs\_category: nivelul de glucoză din sânge al pacientului;
* restecg\_category: rezultatul electrocardiogramei pacientului (normal sau anormal);
* thalach: frecvența cardiacă maximă a pacientului;
* exang\_category: tipul de angină experimentată de pacient (tipică, asimptomatică sau non-anginală);
* oldpeak: deplasarea segmentului ST a pacientului;
* slope\_category: panta segmentului ST al pacientului;
* ca\_category: numărul de vase coronare principale ale pacientului care sunt blocate;
* thal\_category: tipul de defect de umplere ventriculară al pacientului.

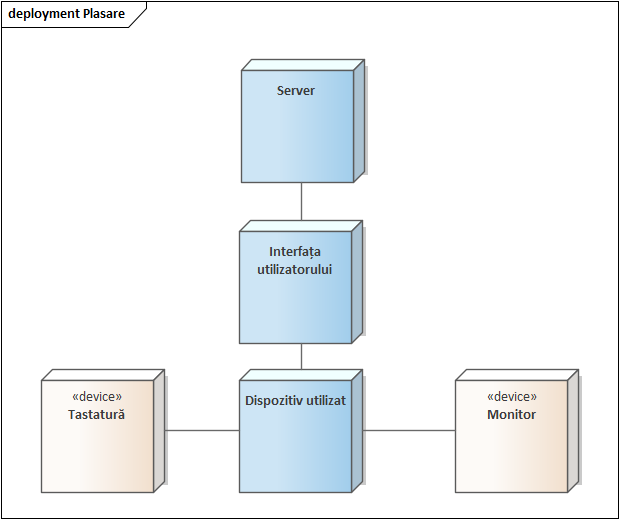
Clasa are următoarele metode:

* convert\_sex\_to\_numeric(): convertește categoria sexului pacientului într-un număr;
* convert\_cp\_to\_numeric(): convertește categoria durerii toracice a pacientului într-un număr;
* convert\_fbs\_to\_numeric(): convertește categoria nivelului de glucoză din sânge a pacientului într-un număr;
* convert\_restecg\_to\_numeric(): convertește categoria electrocardiogramei pacientului într-un număr;
* convert\_exang\_to\_numeric(): convertește categoria anginei pacientului într-un număr;
* convert\_slope\_to\_numeric(): convertește categoria pantei segmentului ST al pacientului într-un număr;
* convert\_ca\_to\_numeric(): convertește categoria vaselor coronare blocate ale pacientului într-un număr
* convert\_thal\_to\_numeric(): convertește categoria defectului de umplere ventriculară a pacientului într-un număr;
* display\_heart\_disease\_prediction(): afișează predicția riscului de boli de inimă a pacientului.

Diagrama UML prezentată în figura 2.5 este o diagramăde plasare, ilustrând modul în care componentele aplicații sunt distribuite. Diagrama cuprinde trei componente cheie:

* **Serverul:** această componentă are rolul de a stoca datele și de a furniza servicii web. De asemenea, este responsabil pentru administrare aplicației.
* **Interfața utilizatorului:** această componentă permite interacțiunea utilizatorilor cu aplicația, oferind funcționalități precum importanța diagnosticării timpurii a atacului de cord, predicția riscului apariției atacului de cord și o vizualizare a datelor pe baza căror a fost antrenat modelul de predicție.
* **Dispozitivele utilizate:** Această componentă include diverse dispozitive, cum ar fi computere desktop, laptopuri, telefoane mobile și tablete, pe care utilizatorii pot accesa interfața sistemului de vânzări.

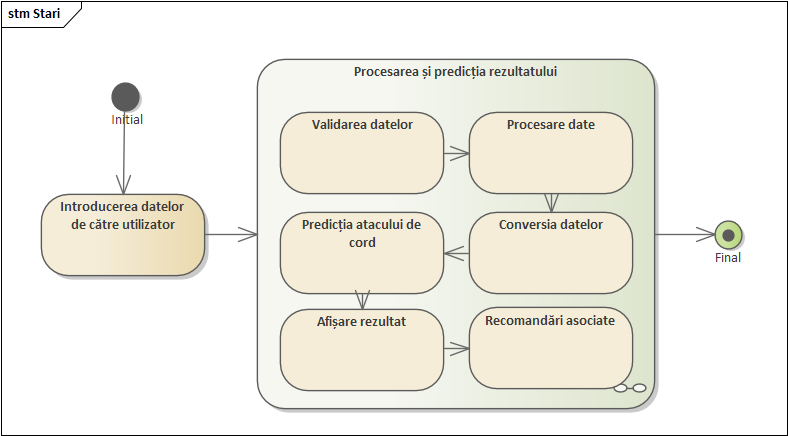
Diagrama sugerează o arhitectură centralizată cu serverul, situate pe un dispozitiv central, posibil un server web.



**Figura 2.5***Diagrama de plasare*

Această diagramă furnizează o perspectivă globală a modului în care componentele unei aplicații de predicție a riscului apariției atacului de cord sunt distribuite pe diverse dispozitive, contribuind la înțelegerea arhitecturii generale a sistemului.

Diagrama UML prezentată în figura 2.6 descrie un proces de procesare a datelor și predicție a rezultatului.

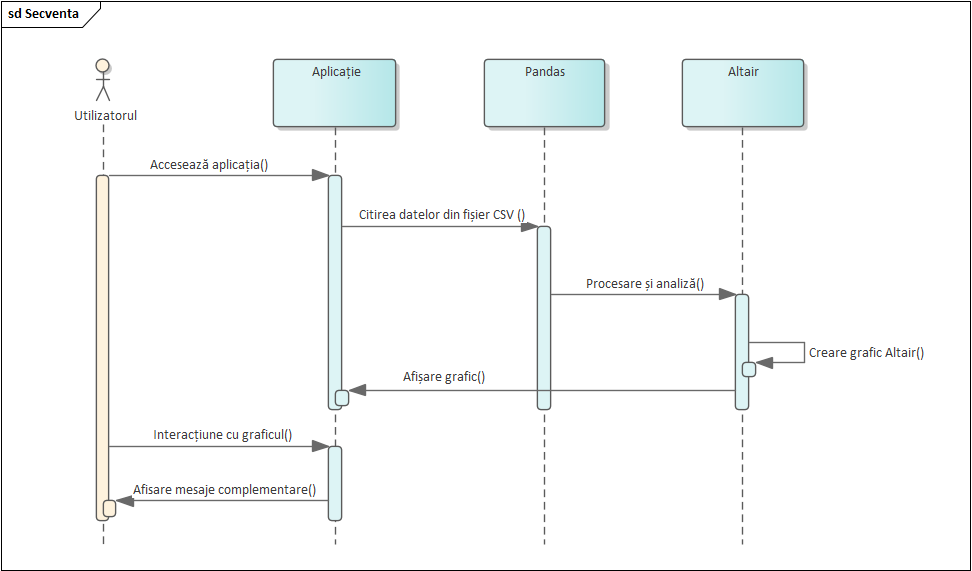


**Figura 2.6***Diagrama de stări*

Procesul este format din mai multe etape, reprezentate în figura 2.6, și are un flux definit, indicat de săgeți.

* Introducerea datelor de către utilizator: utilizatorul introduce datele inițiale necesare procesului.
* Validarea datelor: datele introduse sunt verificate pentru a se asigura că sunt corecte și complete.
* Conversia datelor: datele sunt convertite într-un format compatibil cu procesul de analiză.
* Procesare date: datele sunt procesate și analizate folosind algoritmi specifici.
* Predicția atacului de cord: se utilizează modele de predicție pentru a estima riscul de atac de cord.
* Afișare rezultat: rezultatul predicției este afișat utilizatorului.
* Recomandări asociate: în funcție de rezultatul predicției, pot fi afișate recomandări personalizate pentru reducerea riscului de atac de cord.

Diagrama UML oferă o reprezentare vizuală de ansamblu a procesului de predicție a riscului de atac de cord. Diagrama este utilă pentru a înțelege etapele principale ale procesului și fluxul de informații.



**Figura 2.7***Diagrama de secvență*

Diagrama de secvență din figura 2.7 prezintă interacțiunile dintre utilizator, aplicație, Pandas, Altair, elemente care sunt prezente într-un proces de analiză a datelor. Utilizatorul accesează aplicația, care apoi citește datele dintr-un fișier CSV și le procesează folosind Pandas pentru a identifica modele și a obține informații semnificative. Apoi, aplicația utilizează Altair pentru a crea un grafic bazat pe datele procesate, pe care îl afișează utilizatorului. Acesta poate interacționa cu graficul și primește mesaje complementare pentru a obține detalii suplimentare despre datele prezentate. Diagrama se oprește la acest nivel, sugerând posibile etape ulterioare, cum ar fi salvarea graficului sau exportarea datelor.

## Concluzii intermediare la al doilea capitol

În acest capitol au fost stipulate toate momentele teoretice ce țin de dezvoltarea produsului din cadrul acestei lucrări. Am explorat domeniul inteligenței artificiale și subdomeniile sale relevante, inclusiv procesarea limbajului natural (NLP) și învățarea automată. Am discutat despre apariția și dezvoltarea inteligenței artificiale, de la originile sale în anii '40-'50 până în epoca modernă a învățării profunde și a vehiculelor autonome.

De asemenea, am acoperit conceptele-cheie din procesarea limbajului natural, precum curățarea textului, transformarea în litere mici, eliminarea stop words, stemming și lemmatizare, eliminarea elementelor irelevante și tokenizare. Aceste etape sunt cruciale în prelucrarea eficientă a limbajului natural pentru a extrage informații semnificative.

În ceea ce privește învățarea automată, am discutat despre importanța datelor și a experienței în antrenarea modelelor de inteligență artificială. Procesul de antrenare a unui model a fost prezentat, evidențiind rolul seturilor de date și al algoritmilor în dezvoltarea modelelor eficiente. De asemenea, am abordat conceptul de învățare continuă și principiul de lucru a unei rețele neuronale, care sunt esențiale pentru menținerea relevanței și eficacității modelelor de inteligență artificială pe termen lung.

În plus, am realizat modelare sistemului folosind diagramele UML. Aceste diagrame au contribuit la o înțelegere mai clară a structurii și funcționării aplicației, evidențiind interacțiunile dintre actori și componentele sistemului, structura statică a acestuia, fluxurile de activitate și interacțiunile secvențiale între diferitele module. Această abordare a adus claritate în dezvoltarea și comunicarea aspectelor tehnice ale proiectului.

# DEZVOLTAREA UNEI APLICAȚII PENTRU PREDICȚIA RISCULUI ATACULUI DE CORD

Pe parcursul ultimelor decenii, cercetările în domeniul sănătății cardiovasculare au evoluat considerabil, beneficiind de avansurile tehnologice și de accesul sporit la date relevante. Cu toate acestea, există încă lacune semnificative în înțelegerea complexă a interacțiunilor dintre factorii de risc, precum vârsta, sexul, nivelurile de colesterol și presiunea arterială, care contribuie la apariția atacurilor de cord.

Atacul de cord (infarctul miocardic) rămâne una dintre cele mai presante probleme de sănătate la nivel mondial, exercitând o presiune semnificativă asupra sistemelor de sănătate și având consecințe serioase asupra calității vieții și a longevității populației.

Prin intermediul antrenării unui model ce a putut prezice apariția riscului atacului de cord, ne-am propus să aducem o contribuție semnificativă în domeniu. Am explorat relațiile dintre variabile, am dezvoltat modele logistice relevante și am evaluat impactul acestora asupra probabilității de atac de cord. De asemenea, am examinat modul în care factorii de risc identificați în acest studiu pot influența elaborarea unor strategii personalizate de prevenție și intervenție.

Seturile de date ample și accesibile precum "Heart Attack Analysis & Prediction" au oferit oportunități semnificative de a investiga profunzimea și amploarea problemelor de sănătate cardiovasculară[22]. În lumina acestui context, a fost desfășurat un studiu detaliat, utilizând acest set de date, pentru a analiza și a prezice riscul de atac de cord. Scopul principal a fost să se identifice factorii de risc cheie și să se dezvolte un model logistic eficient pentru a estima probabilitatea de a suferi un atac cardiac în viitor.

Într-un context mai larg, această cercetare poate contribui la avansarea cunoștințelor noastre despre riscul cardiovascular, oferind fundamentul pentru dezvoltarea unor metode de identificare și intervenție mai eficiente. Prin urmare, această investigație s-a aliniat eforturilor globale de combatere a bolilor cardiovasculare, reprezentând un pas esențial către îmbunătățirea sănătății publice și a calității vieții.

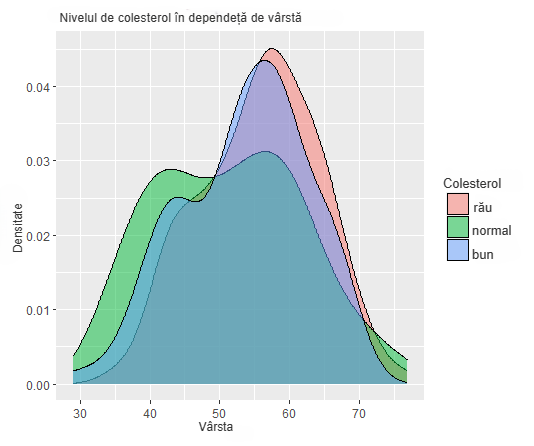
## Analiza distibuțiilor variabilelor

Analiza distribuțiilor variabilelor în contextul creării unui model de învățare automată pentru predicția riscului atacului de cord este esențială pentru a înțelege mai bine datele și pentru a asigura calitatea și eficacitatea modelului. Această analiză ne ajută să obținem o perspectivă profundă asupra distribuțiilor diferitelor caracteristici din setul de date.

Prin examinarea distribuțiilor variabilelor, putem identifica modele, tendințe și relații între diferitele caracteristici și variabile din setul de date. Această înțelegere este crucială pentru selectarea caracteristicilor relevante care vor fi utilizate în modelul nostru de predicție a riscului de atac de cord. Identificarea caracteristicilor semnificative ne permite să concentrăm eforturile pe dezvoltarea unor modele precise și eficiente.

Astfel, în încercarea de a înțelege relația dintre vârstă și nivelul colesterolului din sânge, am analizat setul de date. Una din presupuneri este dacă există o asociere semnificativă între înaintarea în vârstă și nivelurile de colesterol în sânge.

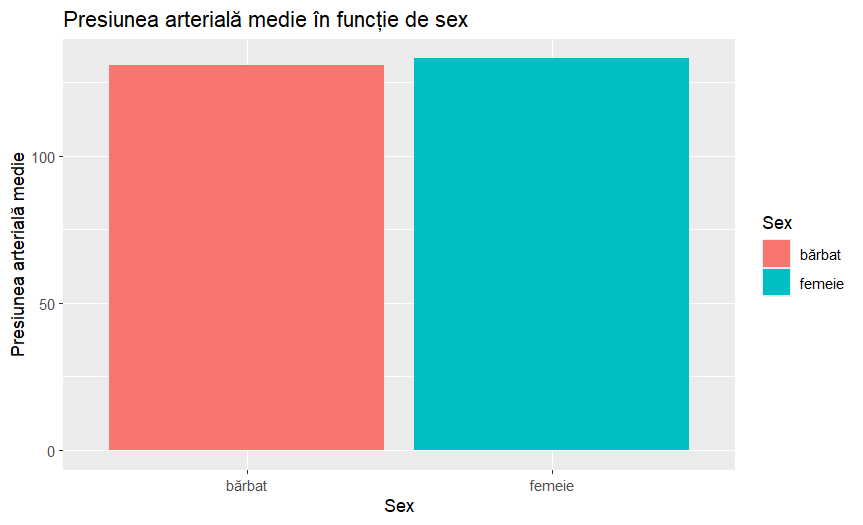
Prin urmare, ne-am concentrat pe variabilele *age* și *chol* pentru a explora trendurile posibile. Pentru a ilustra această relație și pentru a facilita înțelegerea rezultatelor noastre, am pregătit o figură informativă care sintetizează corelația între vârstă și nivelurile de colesterol în sânge(figura 3.1.1).



**Figura 3.1** Nivelul de colesterol raportat la vârstă

Plotul final este un grafic de densitate(figura 3.1.1) pentru a evidenția distribuția nivelurilor de colesterol în funcție de vârstă. Acesta utilizează curbe de densitate pentru fiecare nivel de colesterol. Putem observa că după vârsta de 40 de ani, nivelul colesterolului are tendința de a crește și a trece din categoria Normal în categoria Bad.

În efortul de a dezvălui posibile corelații între variabilele tensiunii arteriale și sexului, am analizat atent setul nostru de date. O presupunere pe care am examinat-o este dacă există o asociere semnificativă între sex și nivelurile tensiunii arteriale, cu o posibilă tendință de a observa niveluri mai ridicate la femei în comparație cu bărbații.

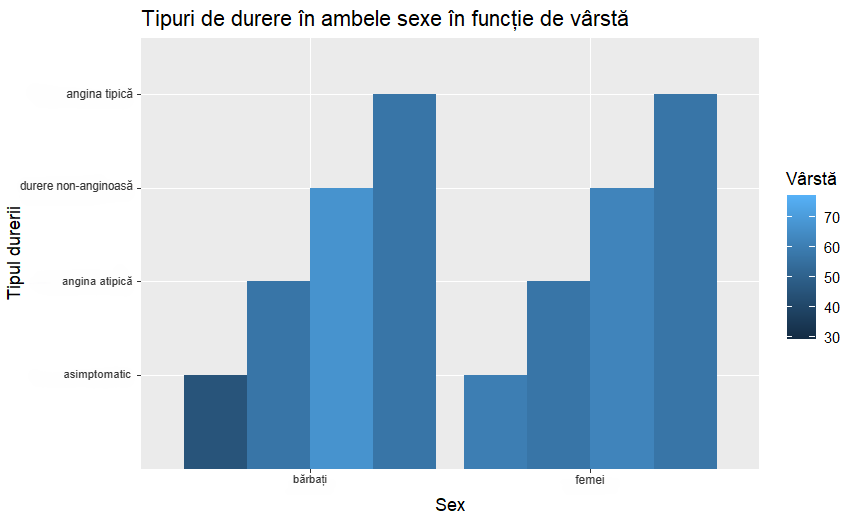


**Figura 3.2** Media tensiunii arteriale pentru ambele sexe

În urma analiza acestei presupuneri, am optat pentru vizualizarea cu ajutorul unui grafic de tip bar, reprezentat în figura 3.1.2, pentru a compara distribuțiile tensiunii arteriale în funcție de sex. Dar, spre surprindere presupunerea nu s-a adeverit, fiind depistată o diferență minoră în setul de date analizat.

O altă presupunere pe care am cercetat-o este dacă există vreo asociere semnificativă între vârstă și prevalența anumitor tipuri de durere, sugerând posibile schimbări în raportarea sau frecvența acestora odată cu înaintarea în vârstă.

Pentru a evidenția aceste potențiale corelații, am focalizat atenția asupra variabilelor relevante, respectiv 'age' și 'cp'.

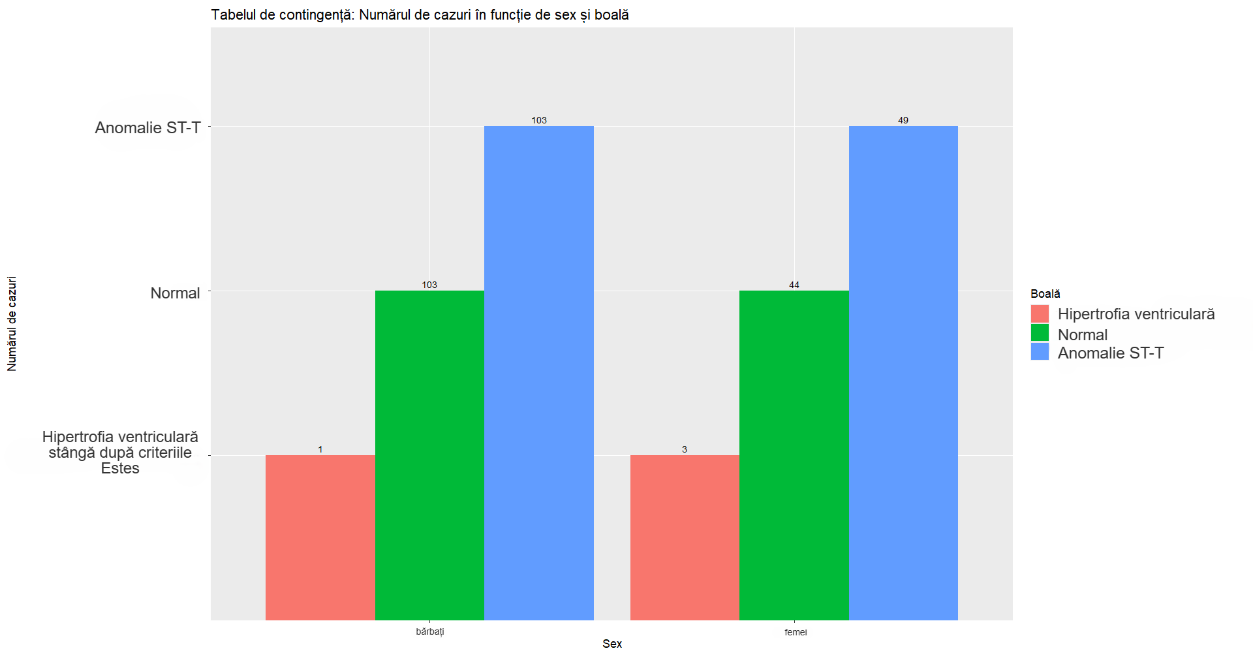


**Figura 3.3** Tipul de durere raportat la sex

Folosind un grafic de tip bar, am intenționat să evidențiem frecvența diferitelor tipuri de durere în grupuri de vârstă distincte. Astfel, putem observa în figura 3.1.3 că bărbații cu vârsta cuprinsă între 30 și 40 de ani nu prezintă dureri, iar la sexul opus diapazonul fiind mai mare, ajungând și la vârsta de 70 de ani în unele cazuri. Totuși, vârsta își ia a ei, de aceia, o dată cu înaintarea în vârstă sunt tot mai prezente diferite tipuri de dureri, făcându-și apariția unele doar la efort fizic continu, altele fiind permanente, mai ales în rândul bărbaților ce trec de vârsta de 50 de ani.

În procesul de cercetare a datelor, am analizat atent variabilele legate de sexul pacienților și numărul de cazuri de diferite boli pentru a evalua posibile diferențe semnificative în incidența acestora în funcție de sex. O presupunere pe care am investigat-o este dacă există vreo asociere semnificativă între sex și frecvența anumitor boli, sugerând că acestea pot avea o incidență mai ridicată la un anumit sex în comparație cu celălalt.

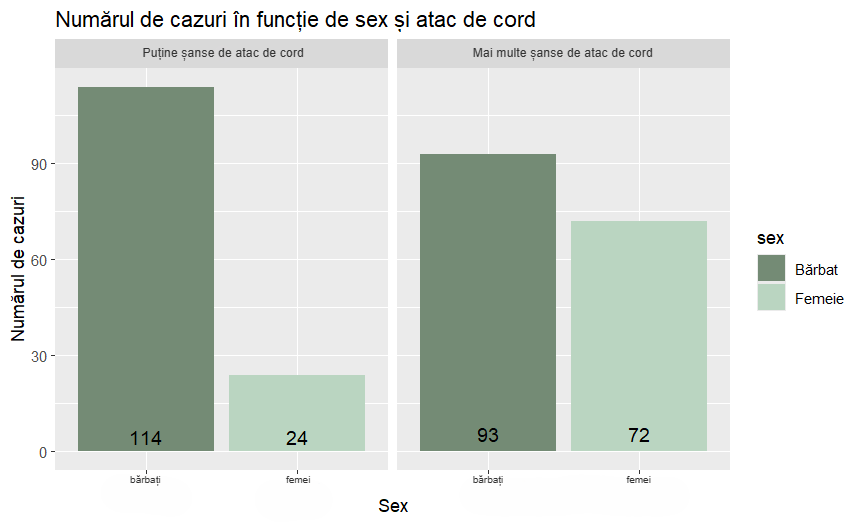
În vederea evidențierii acestor posibile diferențe, am direcționat atenția către variabilele relevante, cum ar fi 'sex' și 'restcg'.



**Figura 3.4** Afecțiuni ale cordului în dependență de sex

Cu ajutorul unui grafic de tip bar, am dorit să comparăm numărul de cazuri ale diferitelor boli între sexe. Plotul din figura 3.1.4 ne oferă claritate. În cazul hipertrofiei ventriculare stângă în rândul bărbaților este doar un singur caz, comparativ cu femeile unde sunt 3 cazuri. De asemenea, în rândul bărbaților au fost înregistrate 103 persoane la care nu au fost detectate afectiuni ale cordului, comparativ cu femeile la care această cifră este mult mai mică, fiind doar 44 de cazuri. Ce ține de a 3-a afecțiune, care poate fi depistată cu efectuarea doar a unei electrocardiograme, rezultatul este opus, fiind înregistrate 103 cazuri între rândurile bărbaților și 49 la femei. În urma analizei, putem afirma cu siguranță că unele afecțiuni ale cordului sunt prepoderente doar la un sex, celălalt sex fiind mai imun unor asemenea afecțiuni.

De asemenea, datele medicale arată că, în general, bărbații au o incidență mai mare de atacuri de cord decât femeile. Acest lucru contravine presupunerii inițiale, care sugera că femeile ar fi mai predispuse la atacuri de cord. Factori precum diferențele biologice, hormonale și comportamentale pot influența acest fenomen. Este esențial să subliniem că aceste concluzii se bazează pe tendințe generale și că fiecare individ poate prezenta riscuri specifice în funcție de istoricul său medical, stilul de viață și alți factori.



**Figura 3.5** Numărul de atac de cord pe sex

Similar cu presupunerea precedentă, vom utiliza un grafic de tip bar pentru a compara numărul de cazuri de atacuri de cord între sexe. Astfel, putem observa în figura 3.1.5 că din 207 bărbați și 96 femei, 114 bărbați și 24 de femei nu au suferit un atac de cord, pe când 93 de bărbați și 72 de femei au suferit un atac de cord. Putem concluziona că doar 25% din femei nu au suferit atact de cord, spre deosebire de bărbați unde raportul este 55%.

Astfel, putem afirma că analiza distribuțiilor variabilelor aduce claritate și consistență în procesul de pregătire a datelor pentru antrenarea modelului de învățare automată. Ea reprezintă un pas esențial în construirea unui model robust și precis care poate face predicții relevante și utile în identificarea riscului de atac de cord și în elaborarea unor strategii eficiente de prevenție și intervenție în domeniul sănătății cardiovasculare.

De asemenea, analiza distribuțiilor variabilelor ne ajută să gestionăm datele lipsă sau eronate, să normalizăm sau să standardizăm datele acolo unde este necesar și să detectăm dezechilibrele în date, cum ar fi dezechilibrul între clase. Aceste aspecte sunt critice pentru asigurarea calității datelor și pentru optimizarea performanței modelului.

## Avantajele regresiei logistice

În dezvoltarea unui model de predicție pentru riscul de atac de cord, am optat pentru utilizarea învățării automate și a regresiei logistice în locul rețelelor neuronale din mai multe motive importante.

Primul motiv esențial este interpretabilitatea și transparența oferite de regresia logistică. Această metodă permite o înțelegere clară a relațiilor dintre variabilele de intrare și rezultatele prezise, lucru crucial în contextul medical, unde este esențial să înțelegem cum anumiți factori de risc contribuie la probabilitatea unui atac de cord. Capacitatea de a interpreta și explica modelele este deosebit de valoroasă pentru elaborarea unor strategii preventive eficiente.

Un alt aspect important este eficiența și performanța regresiei logistice în situații în care datele sunt bine structurate și există o relație liniară între variabilele de intrare și variabila țintă. Având în vedere că anumiți factori de risc, cum ar fi vârsta, sexul și nivelul de colesterol, pot influența semnificativ riscul de atac de cord, regresia logistică poate oferi rezultate precise și rapide, fiind astfel o alegere eficientă și practică.

De asemenea, această metodă este mai ușor de implementat și de gestionat decât rețelele neuronale, care pot fi extrem de complexe și necesită o cantitate semnificativă de date pentru antrenare. Astfel, în situații în care există restricții de resurse computaționale sau dimensiunea setului de date este limitată, ca și în acest caz, regresia logistică este soluția cea mai eficientă.

Prin urmare, alegerea regresiei logistice pentru antrenarea modelului de predicție a riscului de atac de cord se datorează avantajelor sale în interpretabilitate, eficiență și gestionabilitate, ceea ce o face o opțiune relevantă și eficientă pentru acest scop specific.

## Antrenarea modelului de regresie logistică

Pentru a antrena un model eficient pentru predictia riscului de atac de cord, începem prin încărcarea setului de date dintr-un fișier CSV într-un DataFrame pandas. Acest pas este esențial pentru lucrul cu datele în analiza și antrenarea modelului:

*import pandas as pd*

*data = pd.read\_csv('heart\_disease\_data.csv')*

Următorul pas este pregătirea datelor, unde separăm caracteristicile de intrare (predictori) și variabila țintă (target) din setul de date:

*predictors = data.drop("target", axis=1)*

*target = data["target"]*

După pregătirea datelor, împărțim setul de date în seturi de antrenare și testare folosind funcția *train\_test\_split* din biblioteca *sklearn.model\_selection*:

*from sklearn.model\_selection import train\_test\_split*

*X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(predictors, target, test\_size=0.20, random\_state=0)*

Pentru antrenarea modelului de regresie logistică, importăm clasa *LogisticRegression* din *sklearn.linear\_model* și inițializăm un obiect LogisticRegression numit *lr* cu parametrii specificați, apoi îl antrenăm folosind datele de antrenare:

*from sklearn.linear\_model import LogisticRegression*

*lr = LogisticRegression(max\_iter=10000)*

*lr.fit(X\_train, Y\_train)*

După antrenare, folosim modelul pentru a face predicții pe setul de testare și calculăm scorul de acuratețe pentru evaluarea performanței modelului:

*from sklearn.metrics import accuracy\_score*

*Y\_pred\_lr = lr.predict(X\_test)*

*score\_lr = round(accuracy\_score(Y\_pred\_lr, Y\_test) \* 100, 2)*

*print("Scorul de acuratețe obținut cu ajutorul regresiei logistice este: " + str(score\_lr) + " %")*

În final, modelul antrenat este salvat folosind *joblib.dump* pentru a putea fi utilizat ulterior:

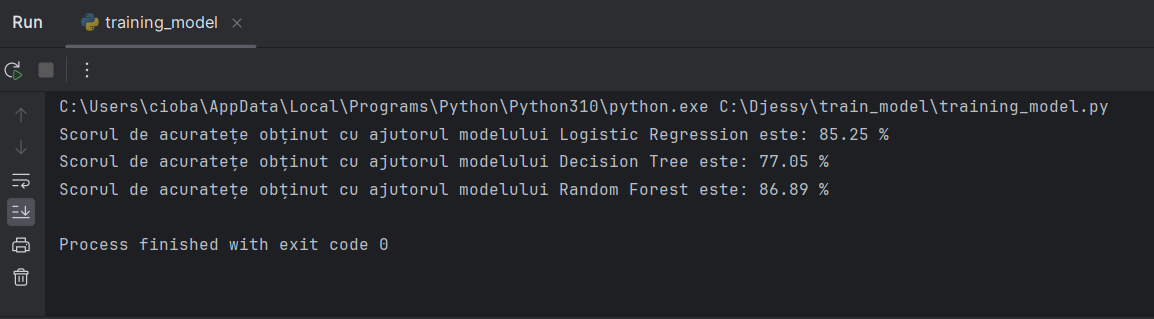
*import joblib*

*joblib.dump(lr, 'logistic\_regression\_model.sav')*

Acești pași reprezintă fluxul de lucru complet pentru încărcarea datelor, antrenarea și evaluarea unui model de regresie logistică pentru predictia riscului de atac de cord.

## Acurateția modelului

Acuratețea modelelor de predicție joacă un rol esențial în identificarea precoce a persoanelor cu risc crescut de a suferi un atac de cord, permițând intervenții prompte și personalizate. Îmbunătățirea performanței și acurateței modelelor de predicție este crucială pentru furnizarea unei asistențe medicale de calitate și pentru salvarea de vieți.



**Figura 3.6** Acurateția modelului antrenat

În figura 3.6 sunt prezentate rezultatele scorului de acuratețe a 3 modele de învățare automată supervisată.

*Logistic Regression* este o metodă de clasificare utilizată pentru a prezice probabilitatea asociată cu două sau mai multe clase. Funcționează prin găsirea celei mai bune linii de separare între clase în datele de antrenare, transformând scorurile brute în probabilități între 0 și 1 cu ajutorul unei funcții logistice. Este adesea utilizat pentru probleme de clasificare binară, precum prezicerea riscului de atac de cord. Logistic Regression a obținut un scor de acuratețe de 85.25%. Acest lucru înseamnă că modelul Logistic Regression a fost capabil să prezică corect clasa corectă în aproximativ 85.25% din cazuri pe setul de testare.

*Decision Trees* sunt structuri arborescente utilizate pentru a lua decizii pe baza caracteristicilor unui set de date. Fiecare nod al arborelui reprezintă o decizie care maximizează separarea datelor în grupuri omogene. Acestea sunt utile pentru clasificare și regresie, fiind eficiente în rezolvarea problemelor non-liniare și complexe prin construirea unor arbori de decizie adaptabili și interpretabili. Decision Tree a obținut un scor de acuratețe de 77.05%. Modelul Decision Tree a avut o performanță bună, dar puțin inferioară Logistic Regression, probabil din cauza complexității problemelor de clasificare din setul de date.

*Random Forest* este un algoritm de învățare supervizată care combină mai mulți arbori de decizie pentru a lua decizii mai precise. Prin construirea unui set de arbori pe baza unor subseturi aleatorii ale datelor și caracteristicilor, apoi combinând rezultatele lor prin vot majoritar, Random Forest oferă soluții eficiente pentru clasificare și regresie în probleme complexe cu multe caracteristici. Random Forest Algorithm a obținut cel mai bun scor de acuratețe, de 86.89%. Acest lucru indică faptul că algoritmul Random Forest a fost capabil să obțină o precizie mai mare în clasificare decât celelalte două modele, datorită capacității sale de a combina multiple arbori de decizie și de a reduce riscul de supradaptare.

## Interfața aplicației

# CONCLUZII

În concluzie, au fost analizate detaliat aspectele teoretice legate de dezvoltarea produsului abordat în această lucrare. Explorarea a inclus o călătorie prin evoluția inteligenței artificiale, începând de la originile sale în anii '40-'50 și până la perioada modernă a învățării profunde și vehiculelor autonome.

Au fost acoperite concepte esențiale din procesarea limbajului natural, precum curățarea textului, transformarea în litere mici, eliminarea stop words, stemming și lemmatizare, precum și eliminarea elementelor irelevante și tokenizare. Aceste etape sunt cruciale pentru a asigura prelucrarea eficientă a limbajului natural și pentru a extrage informații semnificative din texte.

A fost evidențiat procesul de antrenare a unui model, subliniind rolul seturilor de date și al algoritmilor în dezvoltarea modelelor eficiente. De asemenea, a fost abordat conceptul de învățare continuă și am expliact principiul de funcționare a unei rețele neuronale, elemente esențiale pentru menținerea relevanței și eficacității modelelor de inteligență artificială pe termen lung.

S-au explorat trei platforme inovatoare din domeniul sănătății - MyChart, Zocdoc și Teladoc Health - evidențiind evoluția lor, funcționalitățile cheie și impactul asupra îngrijirii medicale.

MyChart a evidențiat cum această platformă a luat naștere în contextul avansurilor tehnologice, oferind pacienților acces facil la informațiile medicale și evoluând constant pentru a răspunde nevoilor în schimbare ale utilizatorilor. Avantajele, precum accesul la informații medicale, comunicarea eficientă și gestionarea sănătății personale, au fost contrabalansate de dezavantaje precum preocupările privind securitatea datelor și excluderea unor grupuri.

Zocdoc, este o soluție simplă și convenabilă pentru programarea la medici, eliminând barierele din procesul de programare și facilitând accesul pacienților la îngrijirea medicală. Funcționalitățile variate, de la căutarea medicilor la gestionarea întregului proces, au fost evidențiate ca beneficii, dar au fost discutate și aspecte precum accesibilitatea limitată și potențialele probleme tehnice.

Teladoc Health, fondată în 2002, și-a extins serviciile și s-a adaptat la nevoile actuale, oferind pacienților acces la consultații medicale online în diverse specialități. Creșterea sa rapidă a fost accentuată de impactul pandemiei de COVID-19.

În ansamblu, aceste platforme reprezintă răspunsuri inovatoare la provocările din domeniul sănătății, aducând atât avantaje semnificative, cât și provocări. De la gestionarea datelor medicale personale la facilitarea programărilor la medici și până la furnizarea de îngrijire medicală la distanță, aceste soluții contribuie la transformarea modului în care pacienții interacționează cu sistemul de sănătate. Cu toate acestea, este important să se abordeze cu atenție provocările legate de securitatea datelor, accesibilitate și etică pentru a asigura o experiență optimă și sigură pentru toți utilizatorii.

# BIBLIOGRAFIE

1. CDC, resursă web a Centrului pentru Prevenirea și Controlul Bolilor, Departamentului de Sănătate Publică al SUA, ‘Heart Disease Facts’ [citat 02.01.2024]. Disponibil: <https://www.cdc.gov/heartdisease/facts.htm>
2. *MyChart by Epic,* Epic [online]. Disponibil: <https://www.mychart.org/About>
3. *What You Can Do With MyChart,* Epic [citat 04.01.2024]. Disponibil: <https://www.mychart.org/Features>
4. *The healthcare industry should work for patients,* ZocDoc [citat 06.01.2024]. Disponibil: <https://www.zocdoc.com/about/>
5. *Zocdoc’s founder on the No. 1 health-care consumer problem that never changes,* CNBC [citat 06.01.2024]. Disponibil: <https://www.cnbc.com/2022/09/29/zocdoc-on-the-no-1-health-care-consumer-problem-that-never-changes.html>
6. *Zocdoc,* saasgenius [citat 06.01.2024]. Disponibil: <https://www.saasgenius.com/reviews/zocdoc/>
7. *Teladoc Health,* Teladoc [citat 08.01.2024]. Disponibil: <https://www.teladochealth.com/about/>
8. *How the North Texas Telemedicine Revolution Began,* dmagazine [citat 08.01.2024]. Disponibil: <https://www.dmagazine.com/publications/d-ceo/2016/november/how-the-north-texas-telemedicine-revolution-began/>
9. *What is artificial intelligence (AI)?* IBM [citat 10.01.2024]. Disponibil: <https://www.ibm.com/topics/artificial-intelligence>
10. *Shakey the Robot,* SRI International [citat 15.01.2024]. Disponibil: <https://www.sri.com/hoi/shakey-the-robot/>
11. *Deep Blue,* IMB[citat 20.01.2024]. Disponibil: <https://www.ibm.com/ibm/history/exhibits/vintage/vintage_4506VV1001.html>
12. TURING, A.M.*, Computing Machinery and Intelligence,* [online]. Research-Gate, 1950 [citat 30.01.2024]. Disponibil: <https://phil415.pbworks.com/f/TuringComputing.pdf>
13. AGGARWAL, Charu, *Neural Networks and Deep Learning.* [online]. Research-Gate, 2018 [citat 10.02.2024]. Disponibil: <http://ndl.ethernet.edu.et/bitstream/123456789/88552/1/2018_Book_NeuralNetworksAndDeepLearning.pdf>
14. *A complete guide to Natural Language Processing,* DeepLearning.AI [citat 13.02.2024]. Disponibil: <https://www.deeplearning.ai/resources/natural-language-processing/>
15. *What is deep learning?,* IBM [citat 15.02.2024]. Disponibil: <https://www.ibm.com/topics/deep-learning>
16. *Deep Learning for Siri’s Voice: On-device Deep Mixture Density Networks for Hybrid Unit Selection Synthesis,* Machine Learning Research [citat 18.02.2024]. Disponibil: <https://machinelearning.apple.com/research/siri-voices>
17. *How to apply continual learning to your machine learning models,* Medium [citat 20.02.2024]. Disponibil: <https://towardsdatascience.com/how-to-apply-continual-learning-to-your-machine-learning-models-4754adcd7f7f>
18. NORAN, Ovidiu S., *Bussines Modeling: UML vs IDEF* [online]. Griffith Inoversity:School of computing and Information Technology [citat 21.02.2023]. Disponibil: <https://www.area-c54.it/public/business%20modelling%20-%20uml%20vs%20idef.pdf>
19. BELL, Donald, *An introduction to the Unified Modeling Language* [online]. New York: IBM, 2003 [citat 29.02.2024]. Disponibil: <https://developer.ibm.com/articles/an-introduction-to-uml/>
20. *Python® – the language of today and tomorrow,* Python Institute [citat 03.03.2024]. Disponibil: <https://pythoninstitute.org/about-python>
21. *Python Tutorial: Streamlit,* datacamp [citat 10.03.2024]. Disponibil: <https://www.datacamp.com/tutorial/streamlit>
22. Kaggle, platformă online pentru seturi de date, ‘Heart Attack Analysis & Prediction Dataset’ [citat 14.04.2024], disponibil: <https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset?resource=download>