

Universitatea “Babeş-Bolyai” Cluj-Napoca

Facultatea de Fizică

Specializarea Fizică Informatică

Lucrare de Licență

Coordonatori științifici,

Lect. dr. Iancu D. Ștefania

Prof. dr. Leopold Nicolae

Absolventa,

Türke Andreea

Universitatea “Babeş-Bolyai” Cluj-Napoca

Facultatea de Fizică

Specializarea Fizică Informatică

Lucrare de Licență

Determinarea concentrației de L-cisteină pe nanoparticule de aur, folosind algoritmi de învățare automată

Coordonatori științifici,

Lect. dr. Iancu D. Ștefania

Prof. dr. Leopold Nicolae

Absolventa,

Türke Andreea

Cuprins

INTRODUCERE.....	7
1. INSTRUMENTAȚIE MODERNĂ ÎN ANALIZA NANOMATERIALELOR ȘI BIOSENZORILOR.....	9
1.1. NANOPARTICULE. REZONANȚĂ PLASMONICĂ.....	9
1.2. BIOSENZORI OPTICI.....	12
1.3. SPECTROSCOPIA UV-VIS.....	13
1.3.1. <i>Tranziții electronice</i>	14
1.3.2. <i>Analiza calitativă</i>	16
1.3.3. <i>Analiza cantitativă</i>	19
1.3.4. <i>Spectrofotometrul ultraviolet-vizibil</i>	22
2. ALGORITMI DE ÎNVĂȚARE AUTOMATĂ UTILIZAȚI ÎN ANALIZA DATELOR	26
2.1. INTRODUCERE ÎN ÎNVĂȚAREA AUTOMATĂ.....	26
2.2. ALGORITMI UTILIZAȚI: PLS, RANDOM FOREST, SVM ȘI REȚELE NEURONALE.....	28
2.2.1. <i>Partial Least Squares (PLS)</i>	28
2.2.2. <i>Random Forest</i>	30
2.2.3. <i>Support Vector Machine (SVM)</i>	31
2.2.4. <i>Rețele Neuronale</i>	32
3. MATERIALE ȘI METODE.....	35
3.1. SINTEZA AuNP ȘI CARACTERIZAREA PRIN SPECTROSCOPIE UV-VIS.....	35
3.2. OPTIMIZAREA FUNCȚIONALIZĂRII AuNP CU L-CISTEINĂ.....	38
3.3. UTILIZAREA PLATFORMEI QUASAR PENTRU PREPROCESAREA ȘI TESTAREA MODELELOR DE ÎNVĂȚARE AUTOMATĂ.....	40
4. REZULTATE ȘI DISCUȚII.....	44
4.1. PROCESAREA ȘI ANALIZA DATELOR EXPERIMENTALE.....	44

4.1.1.	<i>Comportamentul AuNP la concentrații mari (3-120mg/mL) și concentrații foarte mici (0-0,5mg/mL) de l-cisteină.....</i>	44
4.1.2.	<i>Comportamentul AuNP la concentrații de l-cisteină în intervalul 3-0,5 mg/mL.....</i>	49
4.2.	DETERMINAREA GRADULUI DE ACOPERIRE	56
4.2.1.	<i>Concentrația maximă de acoperire.....</i>	57
4.2.2.	<i>Concentrația maximă de saturație.....</i>	58
CONCLUZII.....		62
BIBLIOGRAFIE.....		64

Abstract

Cysteine detection has high importance in biomedical and environmental applications. Its strong affinity for gold nanoparticles (AuNPs) via thiol-gold interactions induces measurable surface plasmon resonance (SPR) shifts. By leveraging these shifts, SPR-based AuNP biosensors enable highly sensitive, label-free detection, making them ideal for rapid biochemical sensing. This study optimizes the detection of cysteine using AuNP biosensors by testing different chemical conditions and implementing machine learning algorithms to enhance the accuracy and efficiency of biochemical sensing.

AuNPs were synthesized following the Turkevich method [1] by reducing a 1 mM HAuCl_4 solution with 77 mM sodium citrate. The SPR band was monitored in the presence of cysteine within a concentration range of 10^{-8} M to 10^{-3} M. To achieve this, UV-Vis absorbance spectra were recorded for AuNP-cysteine mixtures using a Jasco UV-Vis absorbance spectrometer. The diameter of nanoparticles was determined from the SPR band.

The biosensor's performance was optimized by adjusting key parameters such as nanoparticle size, pH, and temperature. AuNPs with a diameter of 23 nm exhibited no visible color change upon cysteine addition, whereas reducing their size to 13 nm led to SPR shifts upon cysteine interaction. Larger nanoparticles were less effective, as aggregation required significantly more time. Further optimization experiments were conducted by varying pH and temperature. At an acidic pH of 4, nanoparticles aggregated too quickly, having an uncontrolled aggregation, while at a basic pH of 8 aggregation did not occur at all. Heating the samples with different cysteine concentrations accelerated the reaction excessively, preventing the acquisition of reliable data, whereas lower temperatures slowed the process beyond the normal rate. Therefore, we adopted neutral pH and room temperature (25°C) as the optimized conditions, balancing reaction speed and data reliability.

Next, using the optimized protocol, we acquired UV-Vis absorbance spectra for mixtures of cysteine and AuNPs. An extensive experimental investigation was conducted, covering 69 samples corresponding to 58 distinct cysteine concentrations. A higher sampling density was applied within the critical range of 0.5–3.0 mg/mL, where the most relevant plasmonic variations were observed. This concentration interval proved optimal for studying the aggregation behavior of AuNPs.

Various regression machine learning models were evaluated, including Partial Least Squares (PLS), Support Vector Machines (SVM), Random Forest (RF), and neural networks. On the full dataset, PLS achieved high performance ($R^2 = 0.989$), followed by Random Forest ($R^2 \approx 0.940$), while

neural networks and SVM performed poorly. When focusing on the optimized interval (0.5–3.0 mg/mL), PLS demonstrated outstanding predictive capability ($R^2 = 0.998$), followed closely by Random Forest ($R^2 = 0.980$), with neural networks showing moderate performance ($R^2 = 0.960$)

Finally, by using 13 nm nanoparticles and using neutral pH and room temperature conditions, we were able to optimize the chemical conditions to accelerate the aggregation of AuNPs. With an R^2 value of 0.861 after evaluation on the test data, the RF model outperforms previous regression models and provides a reliable tool for accurately forecasting cysteine concentrations. The evaluated range of cysteine concentrations was 10^{-8} M to 10^{-3} M. This work demonstrates how AuNP biosensors and machine learning algorithms can be combined to enhance accuracy and sensitivity in biochemical monitoring.

Introducere

Nanoparticulele de aur (AuNP) atrag un interes științific considerabil datorită proprietăților lor fizico-chimice unice, în special a celor optice, care le fac ideale pentru aplicații în domenii precum biosenzorii, diagnosticarea medicală, terapia țintită sau remedierea mediului [2]. Una dintre caracteristicile esențiale ale AuNP este rezonanța plasmonică de suprafață (Surface Plasmon Resonance – SPR), care generează un maxim de absorbție în regiunea vizibilă a spectrului și se manifestă vizual prin culoarea intensă a suspensiilor coloidale. Această proprietate este sensibilă la starea de dispersie a nanoparticulelor și poate fi utilizată pentru detectarea modificărilor de agregare în prezența analiților [3].

L-cisteina (Cys), un aminoacid ce conține o grupare tiol (-SH), prezintă o afinitate puternică față de suprafața AuNP, prin formarea de legături stabile de tip Au-S. În funcție de pH-ul soluției și de concentrația de L-cisteină, aceasta poate exista sub formă neutră sau zwitterionică și se poate auto-asambla pe suprafața NP, influențând astfel starea lor coloidală [4]. Adsorbția moleculelor de cisteină poate induce forțe de atracție electrostatică între NP, conducând la agregare și, implicit, la o modificare a semnalului optic detectat prin spectroscopie UV-Vis [5].

În contextul dezvoltării senzorilor colorimetrici, aceste schimbări de absorbție pot fi corelate cu prezența și concentrația unor substanțe țintă, făcând din AuNP funcționalizate cu L-cisteină un sistem promițător pentru detecția rapidă a ionilor sau compușilor toxici (de exemplu, arsenic sau metale grele) [6].

Totuși, comportamentul acestui sistem este complex, iar variabilele implicate – precum concentrația de L-cisteină, dimensiunea nanoparticulelor, raportul molar și pH-ul – interacționează într-un mod dificil de descris prin metode clasice. În acest sens, algoritmi de învățare automată, oferă o metodologie eficientă pentru analiza și modelarea relațiilor neliniare dintre parametrii experimentali și comportamentul agregării al nanoparticulelor [7].

Scopul acestei lucrări este de a determina concentrația de L-cisteină legată pe suprafața AuNP, utilizând date experimentale obținute prin spectroscopie UV-Vis și metode avansate de învățare automată. Studiul își propune să compare performanța mai multor algoritmi de regresie (Partial Least Squares – PLS, Random Forest, Rețele Neuronale, Support Vector Machines – SVM) pentru a stabili cel mai eficient model în estimarea concentrației de cisteină. Prin această abordare, se urmărește obținerea unui model predictiv robust, care să contribuie la optimizarea procesului de funcționalizare

a AuNP și la dezvoltarea unor senzori optici simpli, portabili și eficienți, cu aplicații potențiale în monitorizarea mediului și domeniul medical.

1. Instrumentație Modernă în Analiza Nanomaterialelor și Biosenzorilor

1.1. Nanoparticule. Rezonanță Plasmonică

Nanoparticulele sunt particule cu dimensiuni în intervalul nanometric (de obicei, 1–100 nm) și prezintă proprietăți fizice și chimice unice, care diferă semnificativ de cele ale materialelor macroscopice. În special, nanoparticulele de metale nobile precum aurul (AuNP) sunt studiate intens datorită *proprietăților* lor *optice*, *electronice* și *catalitice extraordinare*. Metodele moderne de sinteză, precum metoda Turkevich sau metoda Brust-Schiffrin, permit obținerea unor particule cu dimensiuni controlate și o bună monodispersitate, esențiale pentru aplicații în domeniul biosenzorilor, imagisticii și terapiei fototermice [8].

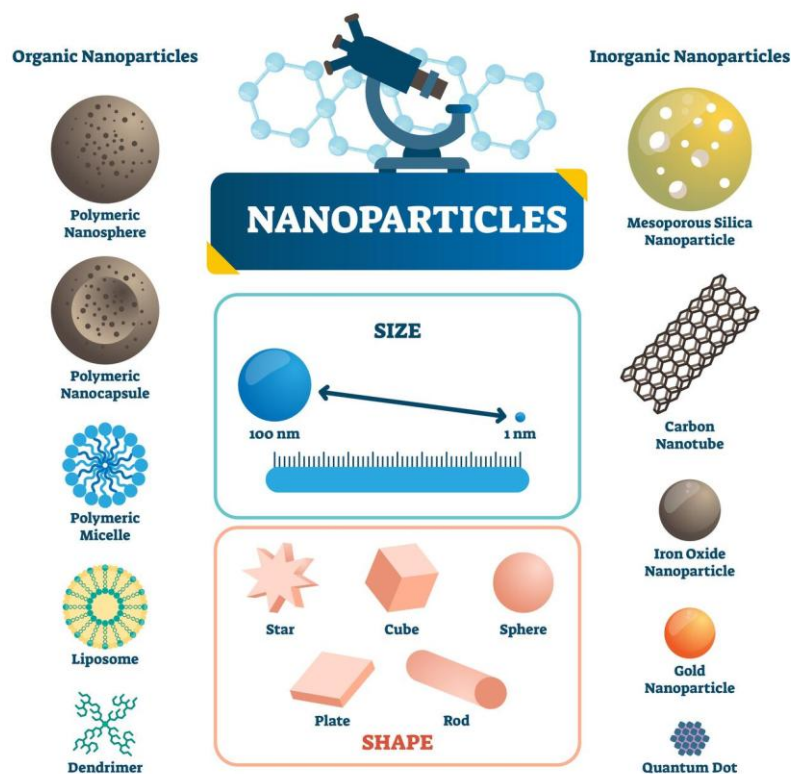


Figura 1.1. Tipuri de nanoparticule. Dimensiunea și diferitele forme ale acestora. [9]

Rezonanță Plasmonică este fenomenul prin care electronii de conducție ai metalelor oscilează colectiv în prezența unui câmp electromagnetic incident. În nanoparticulele de aur, aceste oscilații conduc la un maxim de absorbție și dispersie a luminii la o lungime de undă specifică, de regulă în

intervalul 520–530 nm pentru particulele sferice [10]. Această absorbție specifică este extrem de sensibilă la:

1. *Dimensiunea și forma nanoparticulelor*: Nanoparticulele sferice de aur prezintă un singur maxim plasmonic, în timp ce nanoparticulele anisotrope (ex. bastonașe, discuri) pot genera mai multe moduri plasmonice, fiecare cu o rezonanță specifică [10]. Aceste variații permit ajustarea spectrală a maximului de absorbție și sunt esențiale pentru aplicații personalizate.
2. *Mediul dielectric*: Rezonanța plasmonică este extrem de sensibilă la modificările din mediul înconjurător. Schimbările în indicelui de refracție al mediului înconjurător determină o deplasare a maximului de absorbție. Acest aspect este exploatat în senzorii optici, unde modificările locale ale mediului, prin adsorbția unui analit, se traduc în schimbări spectrale [11].
3. *Interacțiuni interparticulare*: Agregarea nanoparticulelor modifică distribuția câmpului electromagnetic în regiunea interparticulară, ceea ce duce la un *efect batocromic* (deplasare a benzii de absorbție spre lungimi de undă mai mari, denumită și “*red-shift*”). Această proprietate este deosebit de utilă pentru senzorii colorimetrici, unde schimbarea culorii soluției indică prezența unui analit [12].

Pentru nanoparticulele cu dimensiuni mult mai mici decât lungimea de undă a luminii incidente, *teoria Mie* oferă o descriere analitică a modului în care lumina este absorbită și dispersată/împrăștiată de particule sferice. Rezolvarea ecuațiilor Maxwell în condiții de frontieră, conform teoriei Mie, permite calculul eficienței de absorbție și dispersie, evidențiind dependența acestor mărimi de dimensiunea, compoziția și mediul particulelor.

Datorită sensibilității extreme la schimbări de mediu, rezonanța plasmonică este utilizată pe scară largă în *dezvoltarea biosenzorilor optici*. Senzorii care se bazează pe rezonanță plasmonică pot detecta analiți la concentrații foarte mici prin monitorizarea modificărilor spectrale, măsurate de tehnici precum spectroscopia UV-Vis. În plus, aceste nanoparticule sunt folosite în:

- a. *Senzorii colorimetrici*: Schimbările de culoare ale soluțiilor de AuNP sunt observabile cu ochiul liber sau cu instrumente optice, facilitând detectarea rapidă a analiților [13].
- b. *SERS (Surface-enhanced Raman scattering)*: Amplificarea radiației Raman datorită câmpurilor locale intense, generate de rezonanța plasmonică, este utilizată în detecția moleculelor la nivel single-molecule.

- c. **Terapia fototermică și imagistică:** Proprietățile plasmonice ale nanoparticulelor de aur permit absorbția eficientă a luminii și transformarea acesteia în căldură, utilizată în tratamentele fototermice pentru cancer, precum și în imagistica biomedicală.

Teoria AuNP și a rezonanței plasmonice implică o interacțiune complexă între proprietățile electronice ale metalului, geometria particulelor și mediul dielectric înconjurător. Înțelegerea acestor fenomene este crucială pentru proiectarea senzorilor optici cu performanțe ridicate și pentru aplicarea lor în diverse domenii, de la monitorizarea mediului până la diagnosticarea medicală [14].

Nanoparticulele foarte mici (<5 nm) au o suprafață specifică foarte mare, ceea ce duce la o energie superficială ridicată, făcându-le mai puțin stabile din punct de vedere termodinamic și susceptibile la agregare sau fuziune [15]. Pentru a preveni coalescența, sunt necesari agenți stabilizatori puternici, cum ar fi liganzi, polimeri sau surfactanți [16]. Datorită suprafeței extinse, aceste nanoparticule prezintă o reactivitate crescută, fiind folosite și pentru aplicații catalitice. Cu toate acestea, stabilitatea lor este limitată în absența unui control adecvat al mediului [17].

Nanoparticulele de dimensiuni medii (5-50 nm) prezintă o stabilitate coloidală mai bună, în special dacă sunt acoperite cu o sarcină electrostatică sau sterică adecvată [17]. Dimensiunea mai mare reduce tensiunea superficială relativă, crescând stabilitatea comparativ cu nanoparticulele mai mici [18]. Acestea sunt frecvent utilizate în imagistică biomedicală și vectorizarea de medicamente, oferind un echilibru optim între stabilitate și interacțiunea cu mediul biologic [16]. De asemenea, pot interacționa rapid cu proteine și alte biomolecule, ceea ce duce la formarea unei coroane proteice, un factor important în aplicațiile biomedicale [15].

Nanoparticulele mari (>50 nm), deși sunt mai stabile din punct de vedere termodinamic, pot prezenta o stabilitate coloidală redusă din cauza tendinței crescute de sedimentare în suspensii. Pentru a evita aglomerarea, sunt necesare metode eficiente de stabilizare [14] și un strat mai gros de stabilizatori [19]. Aceste nanoparticule sunt preferate în aplicații plasmonice datorită proprietăților optice bine definite, fiind utilizate pe scară largă în dispozitive optoelectronice și senzori [14].

1.2. Biosenzori Optici

Biosenzorii optici reprezintă sisteme analitice care combină elemente biologice de recunoaștere cu transductoare optice, permițând astfel *deteția specifică a analiților* prin măsurarea modificărilor în semnalele optice, de exemplu, absorbția, emisia sau refracția luminii [20]. Acești senzori sunt foarte apreciați datorită sensibilității ridicate, costurilor reduse și posibilității de integrare în dispozitive portabile pentru monitorizarea in-situ a mediului sau în aplicații biomedicale [21].

Un exemplu reprezentativ de biosenzor optic implică utilizarea *nanoparticulelor de aur* (AuNP), care, datorită proprietăților lor optice unice – în special fenomenul de rezonanță plasmonică de suprafață (SPR) – pot semnala prezența unui analit prin schimbări de culoare ale soluției [22]. Nanoparticulele de aur prezintă un maxim de absorbție caracteristic în intervalul 500–550 nm, iar orice modificare a dimensiunii sau agregării acestora conduce la deplasări ale acestui maxim, fenomen exploatat în senzorii colorimetrice [23].

Funcționalizarea suprafeței AuNP cu molecule biologice, cum ar fi aminoacizii, sporește selectivitatea biosenzorilor. În mod particular, L-cisteina (Cys), un aminoacid care se remarcă prin capacitatea sa de a se lega puternic de suprafața aurului datorită grupului tiol (-SH), formând legături covalente Au–S, care asigură atât *stabilizarea nanoparticulelor*, cât și posibilitatea *inducerii agregării controlate* în prezența analiților specifici [24]. Această agregare determină un *efect batocromic*, vizibil ca o deplasare a maximului de absorbție plasmonică, semnalizând astfel prezența analitului [25].

Biosenzorii optici prezintă avantaje remarcabile datorită caracteristicilor lor precum: simplitatea operării, portabilitatea și capacitatea de monitorizare în timp real, putând fi astfel utilizați pe scară largă în aplicații ce variază de la *deteția poluanților ambientali* până la *platforme de diagnostic medical* [9]. Mai mult, integrarea tehnicilor de inteligență artificială, în special a rețelelor neuronale artificiale, în analiza datelor obținute de la biosenzori, permite optimizarea și predicția comportamentului sistemelor, contribuind astfel la îmbunătățirea sensibilității și selectivității senzorilor [26].

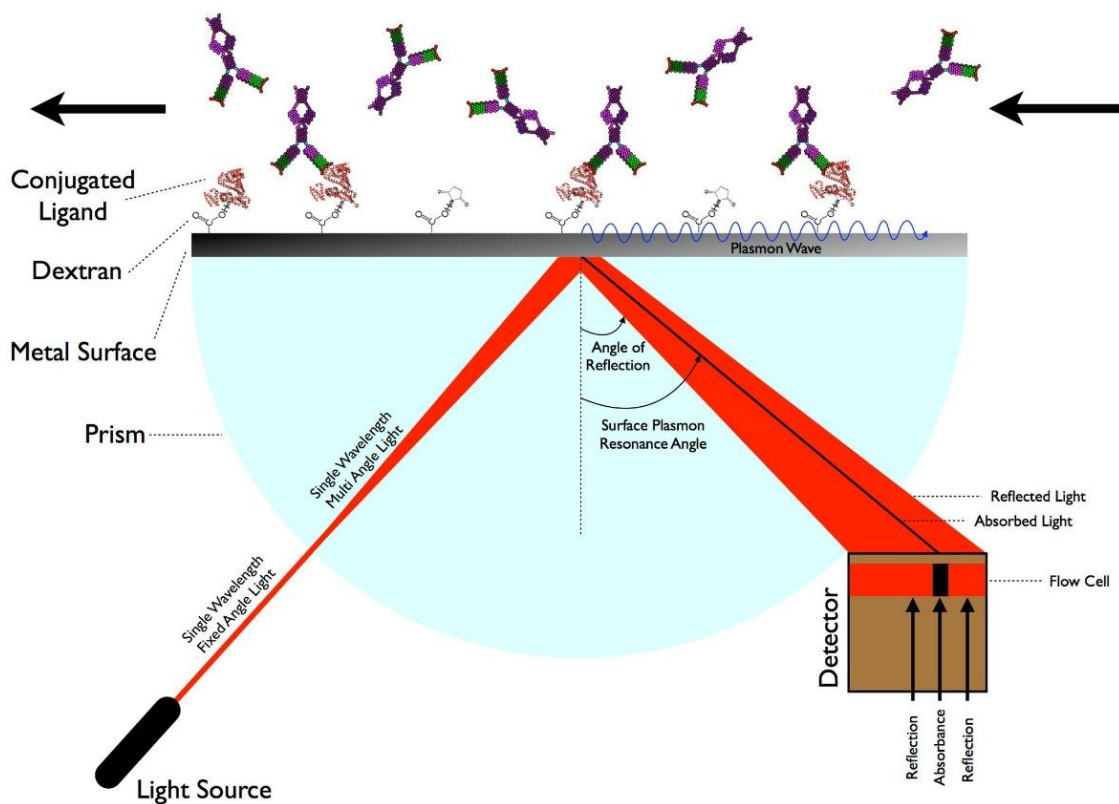


Figura 1.2. Principiul de funcționare al unui biosenzor optic bazat pe rezonanță plasmonică de suprafață (SPR). [27]

1.3. Spectroscopia UV-Vis

Spectroscopia UV-Vis se referă la spectroscopia de absorbție sau la spectroscopia de reflectanță în regiunea ultravioletă și vizibilă. Fiind relativ economică și ușor de implementat, această metodă este utilizată pe scară largă în diverse aplicații atât fundamentale, cât și aplicative. Singura cerință este ca proba să absoarbă în regiunea UV-Vis, adică să conțină un **cromofor**. Spectroscopia de absorbție completează spectroscopia de fluorescență. Parametrii de interes, pe lângă lungimea de undă de măsurare, sunt *absorbția* (A), *transmitanța* ($\%T$) sau *reflectanța* ($\%R$) și modificarea acestora în timp. [28]

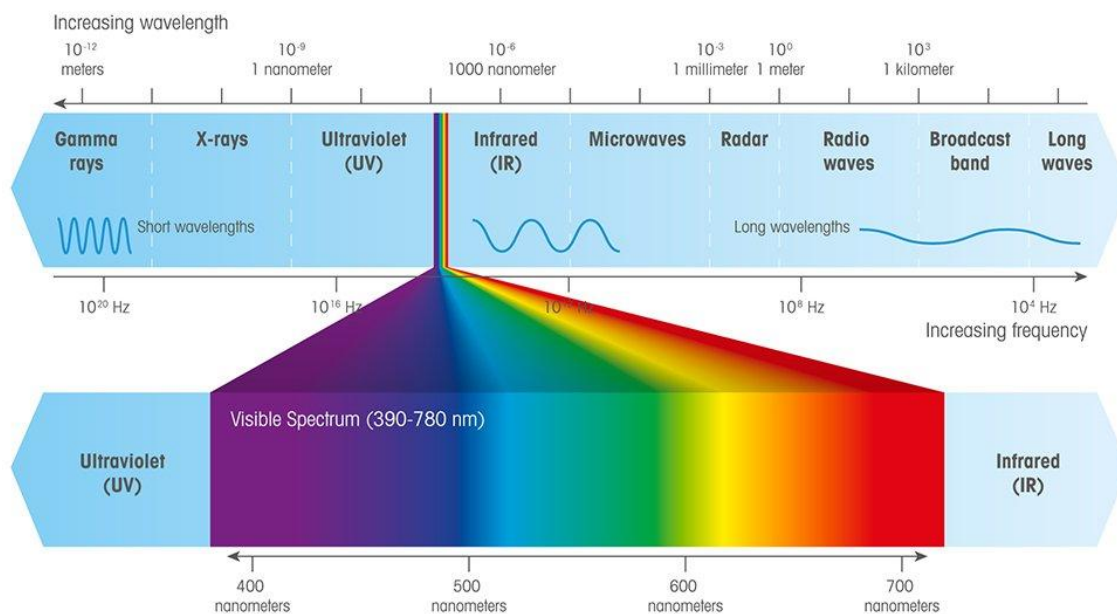


Figura 1.3. Spectrul electromagnetic cu domeniul UV și vizibil. [29]

1.3.1. Tranziții electronice

Majoritatea moleculelor și ionilor absorb energie în domeniul ultraviolet sau vizibil. Absorbția unui foton excită un electron din cromofor către un orbital molecular de energie mai mare, formându-se astfel o stare electronică excitată. Pentru cromoforii organici se disting patru tipuri principale de tranziții electronice [28]:

- $\pi \rightarrow \pi^*$: apar în compușii cu legături duble conjugate și sunt responsabile pentru absorbțiile în regiunea UV și vizibilă.
- $n \rightarrow \pi^*$: implică electroni neparticipanți (n) și legături π^* , apărând frecvent în compușii cu oxigen, azot sau halogeni.
- $\sigma \rightarrow \sigma^*$: sunt specifice legăturilor simple, dar necesită energii mari și apar, de obicei, în regiunea UV de vid.
- $n \rightarrow \sigma^*$: caracteristice compușilor saturați cu heteroatomi, situate în domeniul UV apropiat.

Complexele metalelor de tranziție sunt adesea colorate datorită prezenței stărilor electronice asociate cu orbitalii d incompleți [28].

Când radiația electromagnetică interacționează cu materia, pot avea loc mai multe procese, precum reflexia, dispersia/împrăștierea, absorbția, fluorescența/fosforescența sau o reacție fotochimică. În spectroscopia UV-Vis, **interesul principal este absorbția luminii** [30].

Deoarece lumina este o formă de energie, absorbția acesteia de către materie determină creșterea nivelului energetic al moleculelor sau atomilor. Energia potențială totală a unei molecule este suma **energiilor electronice, vibraționale și rotaționale** [30]:

$$E_{total} = E_{electronică} + E_{vibrațională} + E_{rotațională} \quad \text{Ecuație 1.1}$$

Cantitatea de energie pe care o posedă o moleculă în fiecare formă nu este continuă, ci este o succesiune de niveluri sau **stări discrete**. Diferențele de energie între aceste stări urmează ordinea [30]:

$$E_{electronică} > E_{vibrațională} > E_{rotațională} \quad \text{Ecuație 1.2}$$

În unele molecule și atomi, fotonii din domeniile UV și vizibili au suficientă energie pentru a produce **tranziții electronice** între diferitele niveluri energetice. Lungimea de undă a luminii absorbite este aceea care are energia necesară pentru a transfera un electron de pe un nivel energetic inferior pe unul superior. Aceste tranziții ar trebui să determine benzi de absorbție foarte înguste, la lungimi de undă caracteristice diferențelor dintre nivelurile energetice ale speciilor absorbante. Acest lucru este adevărat **pentru atomi** [30], așa cum este ilustrat în *Figura 1.4*.

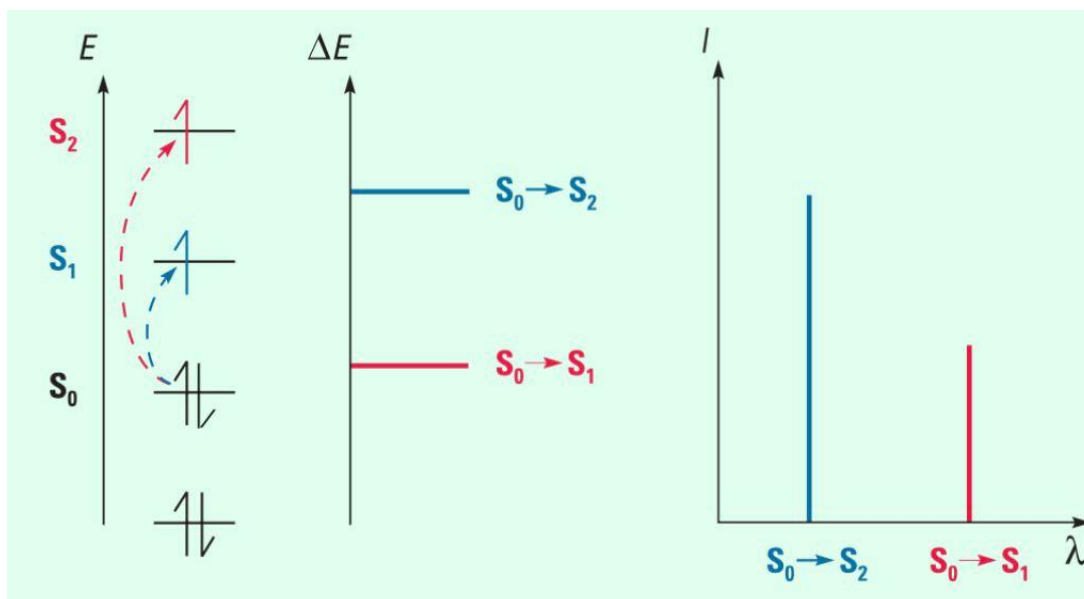


Figura 1.4. Tranziții electronice și spectrele în cazul atomilor. [30]

În cazul moleculelor, nivelurile de energie vibrațională și rotațională sunt suprapuse peste nivelurile de energie electronică. Deoarece pot avea loc multe tranziții cu energii diferite, benzile de absorbție devin mai late (Figura 1.5). Iar, lărgirea benzilor este și mai accentuată în soluții, din cauza interacțiunilor dintre solvent și solut [30].

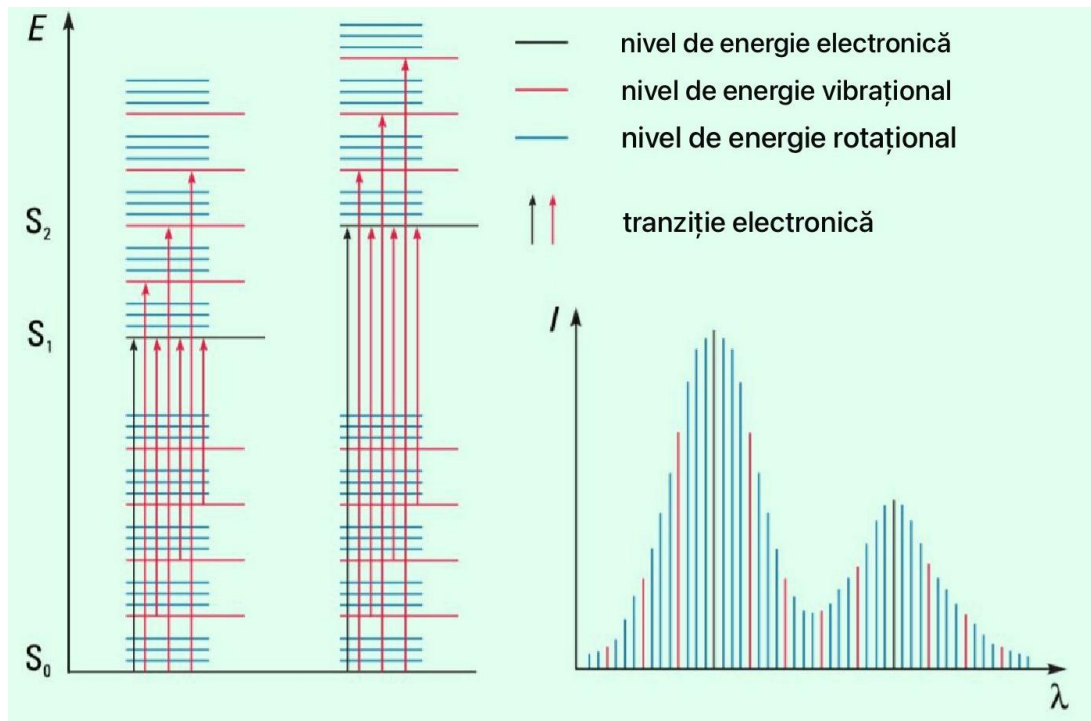


Figura 1.5. Tranziții electronice și spectre UV-Vis în molecule. [30]

1.3.2. Analiza calitativă

Spectrele obținute prin spectroscopie UV-Vis prezintă, de obicei, benzi late de absorbție. Spre deosebire de spectroscopia în infraroșu, care produce numeroase benzi înguste, spectroscopia UV-Vis furnizează informații calitative limitate. Majoritatea absorbțiilor compușilor organici sunt determinate de prezența legăturilor π (adică nesaturate). [30]

Un **cromofor** este un grup molecular care conține o legătură dublă conjugată. Inserarea acestuia într-o hidrocarbură saturată (care nu prezintă absorbție UV-Vis) conduce la apariția unui compus cu absorbție în intervalul 185-1000 nm. Prezența unei benzi de absorbție la o anumită lungime de undă este un indiciu clar al existenței unui cromofor. Poziția maximului de absorbție nu este fixă, ci depinde de mediul molecular al cromoforului și de solventul utilizat. Factori precum **pH-ul și temperatura pot influența intensitatea și poziția lungimii de undă a maximelor de absorbție**. [30]

Conjugarea unei legături duble cu alte legături duble determină creșterea atât a intensității, cât și a lungimii de undă a benzilor de absorbție. [30]

Ionii metalelor de tranziție prezintă niveluri de energie electronică care determină absorbția luminii în domeniul vizibil, între 400-700 nm. [30]

Confirmarea Identității. Spectroscopia UV-Vis nu permite o identificare absolută a unei substanțe, dar este adesea utilizată pentru compararea identității prin suprapunerea spectrelor. Când spectrele sunt foarte asemănătoare, analiza derivativă poate fi utilă. Spectrele derivate prezintă un număr crescut de benzi comparativ cu spectrele de bază, fapt ce poate fi valorificat în analiza calitativă pentru caracterizarea sau identificarea substanțelor. Rezoluția spectrală îmbunătățită prin utilizarea derivatelor permite separarea benzilor care, inițial, par suprapuse. În *Figura 1.6*, simulările arată că două benzi gaussiane cu o lățime naturală de bandă (NBW) de 40 nm, separate de 30 nm, apar ca o singură bandă în spectrul de absorbție obișnuit. În spectrul derivativ de ordinul patru, aceste două benzi sunt clar distinse, cu maxime apropiate de lungimile de undă individuale ale componentelor. [30]

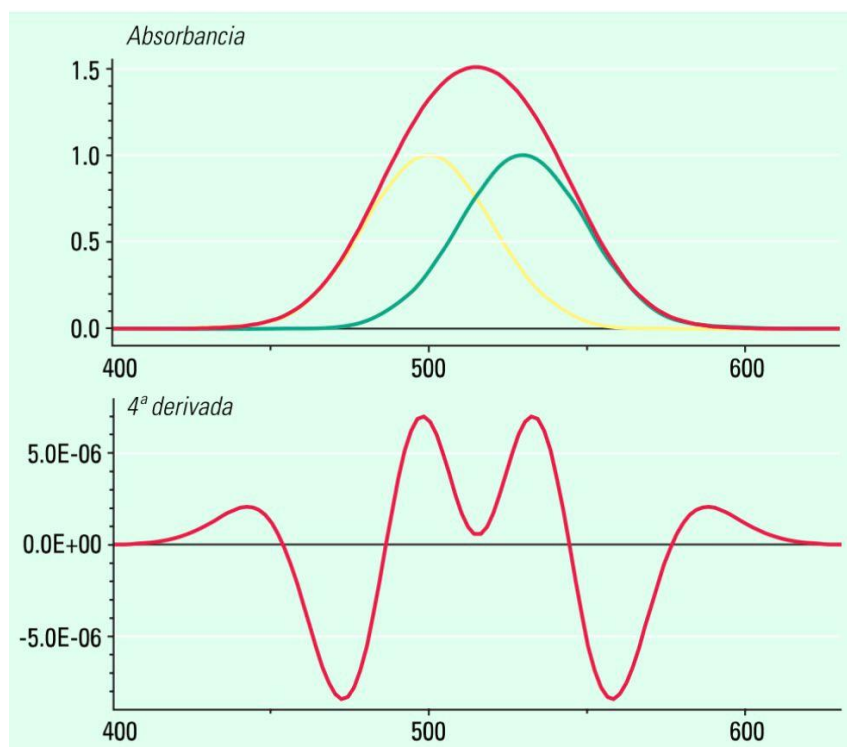


Figura 1.6. Îmbunătățirea rezoluției. [30]

Relația cu Culoarea. Culoarea este o proprietate importantă a unei substanțe, aceasta este legată de absorbția sau reflectanța luminii. Ochiul uman percepe culoarea complementară lungimii de undă absorbite (Figura 1.7 și Figura 1.8). [30]

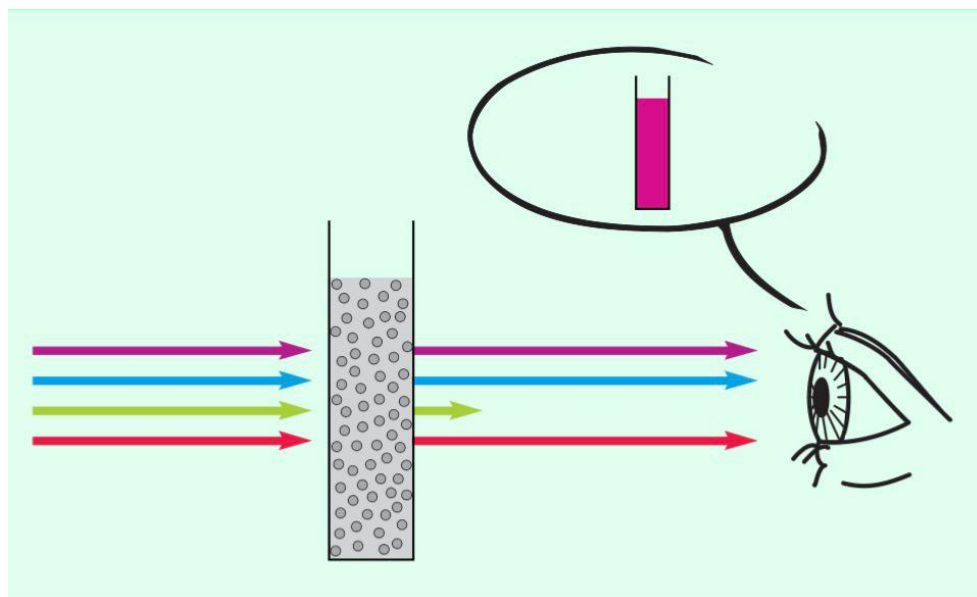


Figura 1.7. Transmisia culorii unei substanțe. [30]

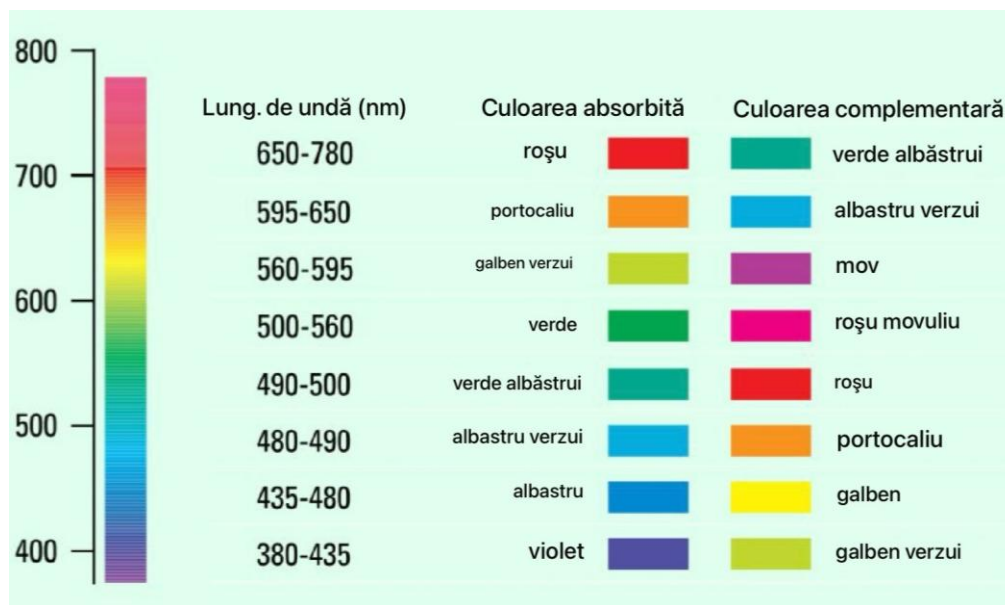


Figura 1.8. Absorbanța și culorile complementare. [30]

Percepția culorii este influențată de factori precum *spectrul sursei de lumină* și *structura suprafeței substanței*. Pentru caracterizarea exactă a culorii, se utilizează sisteme standardizate, precum CIE L^*a^*b , și instrumente specializate integrate cu software de analiză a culorii. Spectrofotometrele moderne pot fi echipate pentru măsurarea exactă a culorii. [30]

Spectroscopia UV-Vis oferă informații despre caracteristicile fizico-chimice ale compușilor, contribuind astfel la **determinarea identității** acestora. Printre acestea se numără *polaritatea*, *stabilitatea termică* și *comportamentul de disociere*. Pentru obținerea unor rezultate precise, acuratețea lungimii de undă absolute și a fotometriei absolute este esențială pentru identificarea compușilor necunoscuți. Compararea spectrelor obținute în momente diferite și pe instrumente variate presupune o calibrare riguroasă și utilizarea unei rezoluții instrumentale bine definite. [30]

1.3.3. Analiza cantitativă

Spectroscopia UV-Vis este utilizată frecvent în chimia analitică pentru **determinarea cantitativă** a diversilor analiți sau probe, cum ar fi ionii metalelor de tranziție, compușii organici cu grad ridicat de conjugare și macromoleculele biologice. Analiza spectrală se efectuează, de regulă, în soluții, deși se pot studia și solidele sau gazele. [28]

Compușii organici, în special cei cu un grad ridicat de conjugare, absorb și lumina în regiunile UV sau vizibile ale spectrului electromagnetic. Solvenții utilizați pentru aceste determinări sunt, de obicei, apa pentru compușii hidrosolubili sau etanolul pentru compușii solubili în solvenți organici. Solvenții organici pot avea o absorbție UV semnificativă; nu toți solvenții sunt adecvați pentru utilizarea în spectroscopia UV. De exemplu, etanolul absoarbe foarte slab la majoritatea lungimilor de undă. [28]

Spectroscopia UV-Vis este, de asemenea, utilizată și în *industria semiconductorilor* pentru măsurarea grosimii și a proprietăților optice ale filmelor subțiri depuse pe o plachetă (wafer). Spectrometrele UV-Vis sunt folosite pentru *măsurarea reflectanței luminii*, iar datele pot fi analizate prin ecuațiile de dispersie Forouhi-Bloomer pentru a determina indicele de refracție (n) și coeficientul de extincție (k) al unui film, pe baza gamei spectrale măsurate. [28]

Legea Beer-Lambert stabilește că absorbția unei soluții este direct proporțională cu concentrația speciilor absorbante din soluție și cu lungimea drumului optic. Prin urmare, pentru o

lungime de undă fixă, spectroscopia UV-Vis poate fi utilizată pentru *determinarea concentrației absorbantului într-o soluție*. Este necesar să se cunoască cât de rapid variază absorbția odată cu concentrația. Această informație poate fi obținută din referințe (tabele cu coeficienți de extincție molară) sau, mai exact, determinată prin intermediul unei curbe de calibrare. [28]

Metoda se folosește cel mai frecvent în mod cantitativ pentru determinarea concentrațiilor unei specii absorbante din soluție, utilizând **legea Beer-Lambert** [28]:

$$A = \log_{10} \frac{I_0}{I} = \epsilon c L \quad \text{Ecuație 1.3}$$

Unde:

A reprezintă *absorbția măsurată*;

I_0 este *intensitatea luminii incidente* la o anumită lungime de undă;

I este *intensitatea luminii transmise*;

L este *lungimea drumului optic prin probă*;

c este *concentrația speciei absorbante*.

Pentru fiecare specie și lungime de undă, ϵ este o constantă cunoscută ca și **coeficientul de absorbție molară sau extincție**, o proprietate moleculară fundamentală într-un anumit solvent, la o anumită temperatură și presiune, exprimată în $L \cdot mol^{-1} \cdot cm^{-1}$. Absorbția și extincția ϵ sunt uneori definite în termeni de logaritm natural în loc de logaritm în baza 10. [28]

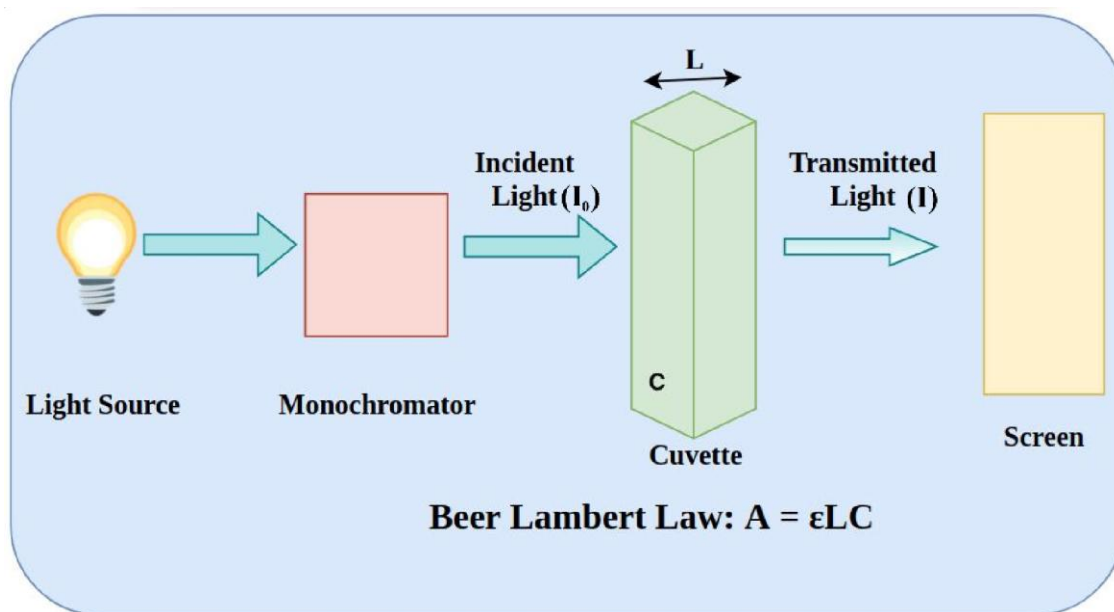


Figura 1.9. Reprezentarea legii Beer-Lambert. [31]

Legea Beer-Lambert este utilă pentru caracterizarea multor compuși, dar nu reprezintă o relație universală pentru concentrație și absorbție pentru toate substanțele. Uneori, pentru molecule foarte mari și complexe, cum ar fi coloranții organici (de exemplu, Xylenol Orange sau Neutral Red), se observă o relație polinomială de gradul al doilea între absorbție și concentrație. [28]

Spectrometrele UV-Vis sunt adesea integrate ca detectoare în cromatografia lichidă de înaltă performanță (HPLC). Prezența unui analit oferă un răspuns care se presupune a fi proporțional cu concentrația. Pentru rezultate precise, răspunsul instrumentului la analitul din proba necunoscută trebuie comparat cu răspunsul obținut pentru un standard; aceasta este foarte similară cu utilizarea curbelor de calibrare. Răspunsul (de exemplu, înălțimea peak-ului) pentru o anumită concentrație este cunoscut sub numele de *factor de răspuns*. [28]

Lungimile de undă ale peak-urilor de absorbție pot fi corelate cu tipurile de legături dintr-o anumită moleculă și sunt valoroase pentru *determinarea grupărilor funcționale* prezente în moleculă. **Regulile Woodward-Fieser** reprezintă un set de observații empirice folosite pentru a prezice λ_{\max} , lungimea de undă la care se produce cea mai intensă absorbție UV-Vis pentru compușii organici conjugați. Totuși, spectrul în sine nu reprezintă o dovadă specifică pentru o anumită probă. *Natura solventului, pH-ul soluției, temperatura, concentrațiile ridicate de electroliți și prezența altor substanțe interferente* pot influența spectrul de absorbție. De asemenea, variații experimentale precum *lățimea fantei* (lățimea de bandă efectivă) a spectrometrului pot altera spectrul. Pentru a aplica corect spectroscopia UV-Vis în analiză, aceste variabile trebuie controlate sau corectate pentru identificarea substanțelor prezente. [28]

Lățimea de bandă spectrală. Este important să se folosească o *sursă de lumină monocromatică* pentru a lumina celula probei. Monocromaticitatea se măsoară ca lățimea triunghiului format de vârful de intensitate la jumătate din intensitatea maximă. Un spectrometru are o lățime de bandă spectrală care indică cât de monocromatică este lumina incidentă. Dacă această lățime de bandă este comparabilă (sau mai mare) decât lățimea liniei de absorbție, coeficientul de extincție măsurat va fi incorect. În măsurătorile de referință, lățimea de bandă a instrumentului (a luminii incidente) se menține sub lățimea liniilor spectrale. Când se măsoară o probă, lățimea de bandă a luminii incidente trebuie să fie, de asemenea, suficient de restrânsă. Reducerea lățimii de bandă reduce energia ce ajunge la detector, necesitând astfel un timp de măsurare mai lung pentru a obține același raport semnal/zgomot. [28]

Eroarea de lungime de undă. În lichide, coeficientul de extincție variază lent cu lungimea de undă. Un peak al curbei de absorbție (o lungime de undă la care absorbția atinge un maxim) este locul unde rata de schimbare a absorbției cu lungimea de undă este minimă. De regulă, măsurătorile se efectuează la un peak pentru a minimiza erorile generate de posibilele erori de lungime de undă ale instrumentului (adică erori datorate unui coeficient de extincție diferit de cel presupus). [28]

Deviațiile de la legea Beer-Lambert. La concentrații suficient de mari, benzile de absorbție se vor satura și vor prezenta o aplatizare a absorbției. Peak-ul de absorbție pare aplatizat deoarece aproape 100% din lumină este deja absorbită. Concentrația la care acest fenomen apare depinde de compusul măsurat. O metodă de testare a acestui efect este variația lungimii traseului optic. Conform legii Beer-Lambert, variația concentrației și a lungimii drumului optic au un efect echivalent: diluarea unei soluții de 10 ori are același efect ca și reducerea lungimii drumului optic de 10 ori. Dacă sunt disponibile celule cu lungimi de drum diferite, verificarea acestei relații poate ajuta la identificarea aplatizării absorbției. [28]

Probele care nu sunt omogene pot prezenta deviații de la legea Beer-Lambert din cauza fenomenului de aplatizare a absorbției. Acest lucru se poate întâmpla, de exemplu, când substanța absorbantă este conținută în particule suspendate. Deviațiile vor fi mai evidente în condiții de concentrație scăzută și absorbție ridicată. Unele studii au propus metode pentru corectarea acestor deviații. [28]

Surse de incertitudine în măsurători. Factorii menționați mai sus contribuie la incertitudinea măsurătorilor obținute prin spectrofotometria UV-Vis. În analiza chimică cantitativă, rezultatele sunt, de asemenea, influențate de sursele de incertitudine care derivă din natura compușilor sau soluțiilor măsurate. Acestea includ *interferențele spectrale cauzate de suprapunerea benzilor de absorbție, decolorarea speciilor absorbante* (datorată descompunerii sau reacției) și eventualele *diferențe de compoziție între proba și soluția de calibrare*. [28]

1.3.4. Spectrofotometrul ultraviolet-vizibil

Instrumentul utilizat în spectroscopia ultravioletă-visible se numește *spectrofotometru UV-Vis*. Acesta măsoară intensitatea luminii după ce trece printr-o probă (I) și o compară cu intensitatea luminii înainte de trecerea prin probă (I_0). Raportul $T = \frac{I}{I_0}$ se numește **transmitanță** și se exprimă, de obicei, ca un procent (%T). **Absorbția** A, se determină pe baza transmitanței după formula [28]:

$$A = -\log \frac{\%T}{100\%} \quad \text{Ecuație 1.4}$$

Spectrofotometrul UV-Vis poate fi configurat și pentru măsurarea reflectanței. În acest caz, instrumentul măsoară intensitatea luminii reflectate de o probă (I) și o compară cu intensitatea luminii reflectate de un material de referință (I_0). Raportul $R = \frac{I}{I_0}$ se numește **reflectanță** și se exprimă, de obicei, ca procent (%R) [28].

Componentele de bază ale unui spectrofotometru sunt: o sursă de lumină, un suport pentru probă (de exemplu, o cuvă), un monocromator (rețea de difracție sau prismă) pentru separarea diferitelor lungimi de undă ale luminii și un detector. [28]

Sursa de radiație poate fi: un filament de tungsten (pentru intervalul 300–2500 nm), o lampă de arc de deuteriu, care furnizează o emisie continuă în regiunea ultravioletă (190–400 nm), o lampă de arc de xenon, care emite continuu din 160 până la 2000 nm, sau, mai recent, diode emițătoare de lumină (LED) pentru lungimile de undă vizibile. [28]

Detectorul poate fi un tub fotomultiplicator, o fotodiodă, o matrice de fotodiode sau un dispozitiv de încărcare cu cuplaj (CCD). Detectoarele cu o singură fotodiodă și tuburile fotomultiplicatoare sunt utilizate împreună cu monocromatoarele cu scanare, care filtrează lumina astfel încât doar lumina de o singură lungime de undă ajunge la detector. Monocromatoarele fixe sunt folosite împreună cu CCD-urile sau seturile de fotodiode. Aceste dispozitive, care sunt formate din numeroși detectori organizați în grupuri unidimensionale sau bidimensionale, pot colecta diferite lungimi de undă în câte un pixel sau grup de pixeli. [28]

Un spectrofotometru poate avea **un singur fascicul** sau **fascicule duble**. Într-un instrument cu un singur fascicul (precum Spectronic 20), toată lumina trece prin celula probei. Intensitatea I_0 se măsoară eliminând proba. Acesta este cel mai vechi design, încă folosit în laboratoarele didactice și în industrie. [28]

Într-un **spectrometru cu fascicule duble**, lumina este împărțită în două fascicule înainte de a ajunge la probă. Un fascicul este utilizat ca **referință**, iar celălalt trece prin **probă**. Intensitatea fasciculului de referință este considerată 100% transmitanță (sau 0 absorbție), iar valoarea afișată este raportul dintre intensitățile celor două fascicule. Unele instrumente cu fascicule duble au doi detectori (fotodiode), astfel încât proba și fasciculul de referință sunt măsurate simultan. În alte instrumente, cele două fascicule trec printr-un comutator, care blochează un fascicul la un moment dat. Detectorul alternează între măsurarea fasciculului de probă și a celui de referință, sincronizat cu comutatorul. Pot

exista, de asemenea, unul sau mai multe intervale „întunecate” în ciclul de măsurare, în care intensitățile sunt corectate prin scăderea intensității măsurate în intervalul întunecat înainte de calculul raportului. [28]

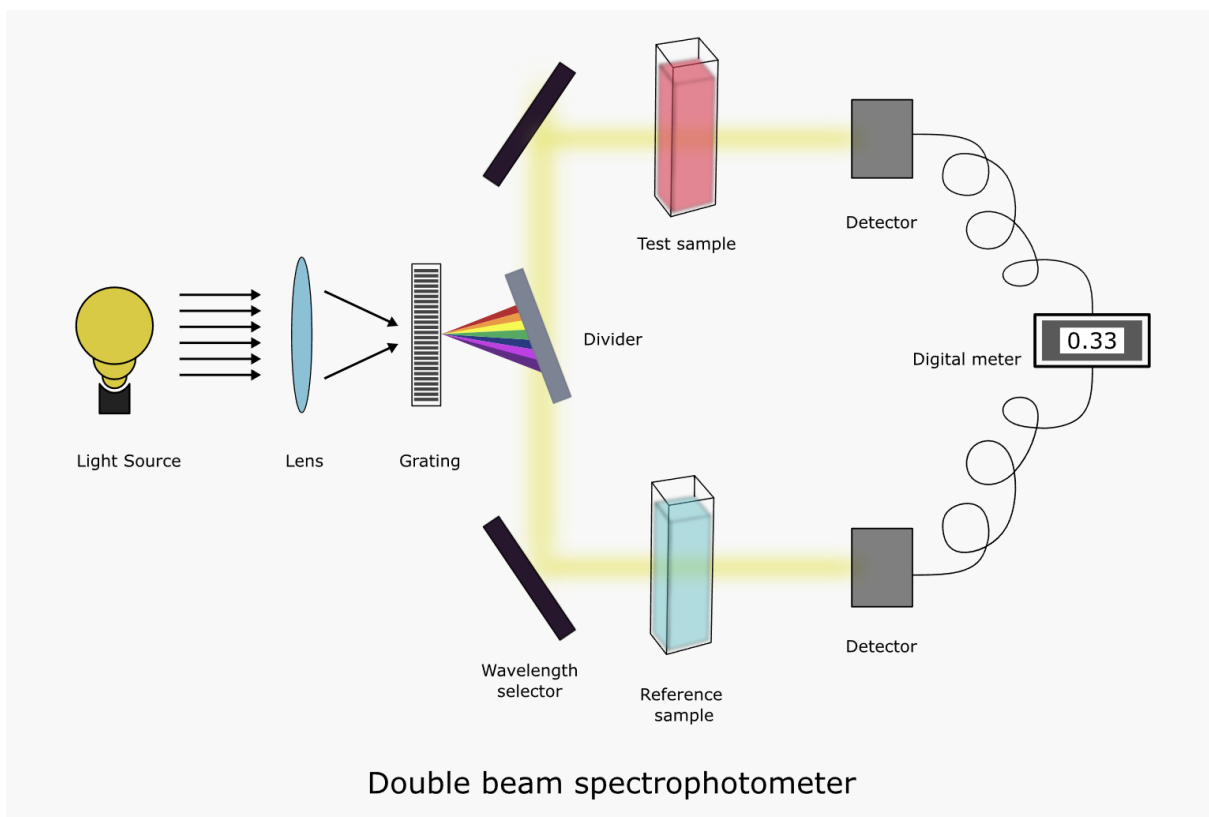


Figura 1.10. Schematizarea simplificată a unui spectrofotometru UV-Vis cu dublu fascicul. [32]

Într-un **spectrometru cu un singur fascicul**, se măsoară mai întâi o cuvetă care conține doar solventul. De exemplu, Mettler Toledo a dezvoltat un spectrofotometru cu un singur fascicul, care permite măsurători rapide și precise în domeniul UV-VIS. Sursa de lumină constă într-o lampă flash de xenon pentru regiunile ultraviolet (UV), vizibil (VIS) și infraroșu apropiat, acoperind un interval spectral de la 190 la 1100 nm. Radiația lămpii este focalizată pe o fibră optică, care direcționează fasciculul de lumină către o cuvetă ce conține soluția probei. Lumina trece prin probă, iar componentele acesteia absorb anumite lungimi de undă. Lumina reziduală este colectată după cuvetă printr-o fibră optică și transmisă către un spectrograf, care dispune lumina într-o rețea de difracție pentru a separa diferitele lungimi de undă și un senzor CCD pentru înregistrarea datelor. În acest fel, întregul spectru este măsurat simultan, permițând o înregistrare rapidă. [28]

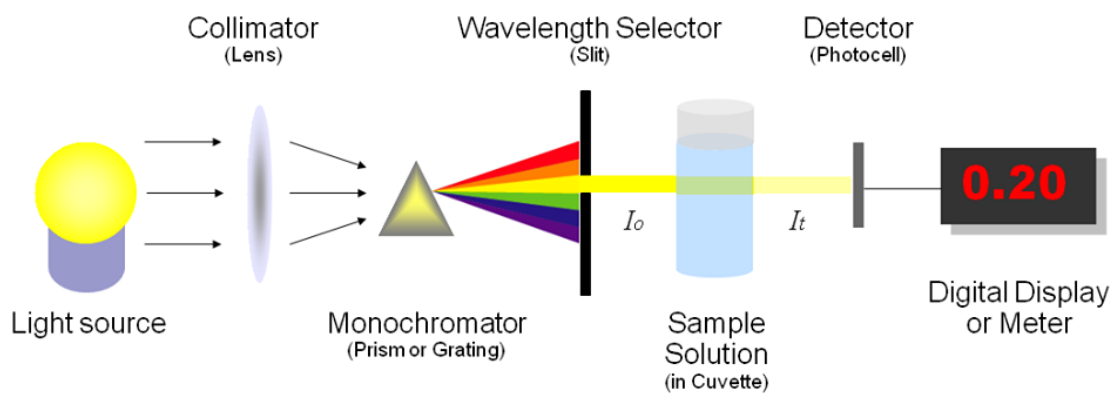


Figura 1.11. Schematizare simplificată a unui spectrofotometru UV-Vis cu un singur fascicul. [33]

Probele pentru spectrofotometria UV-Vis sunt, de regulă, *lichide*, deși se poate măsura absorbția și pentru *gaze* sau *solide*. Probele sunt plasate, de obicei, într-o celulă transparentă, cunoscută sub numele de *cuvă*. Cuvele sunt, în general, de formă dreptunghiulară, cu o lățime internă de 1 cm (această lățime reprezintă lungimea drumului optic, d , în legea Beer-Lambert). Tipul recipientului pentru probă trebuie să permită trecerea radiației pe regiunea spectrală de interes. Cuvele cele mai frecvent utilizate sunt fabricate din cuarț sau silica topită de înaltă calitate, deoarece sunt transparente în regiunile UV, vizibile și în infraroșu apropiat. Cuvele din sticlă și plastic sunt, de asemenea, frecvent întâlnite, însă sticla și majoritatea plasticelor absorb în UV, ceea ce limitează utilizarea lor la lungimi de undă vizibile. [28]

Adesea, un spectrofotometru mai sofisticat poate produce direct un spectru complet de absorbție pe toate lungimile de undă de interes. În instrumentele mai simple, absorbția se măsoară la o singură lungime de undă, iar operatorul compilează ulterior spectrul. Eliminând dependența de concentrație, coeficientul de extincție (ϵ) poate fi determinat în funcție de lungimea de undă. [28]

2. Algoritmi de Învățare Automată Utilizați în Analiza Datelor

2.1. Introducere în Învățarea Automată

Învățarea automată, cunoscută sub numele de *Machine Learning*, reprezintă o ramură a inteligenței artificiale care se ocupă cu *dezvoltarea de algoritmi capabili să învețe tipare din date și să facă predicții* sau să ia decizii fără a fi programați explicit pentru fiecare sarcină. Acest domeniu a cunoscut o evoluție semnificativă în ultimele decenii, fiind aplicat în diverse arii precum medicină, educație, finanțe și industrie [34].

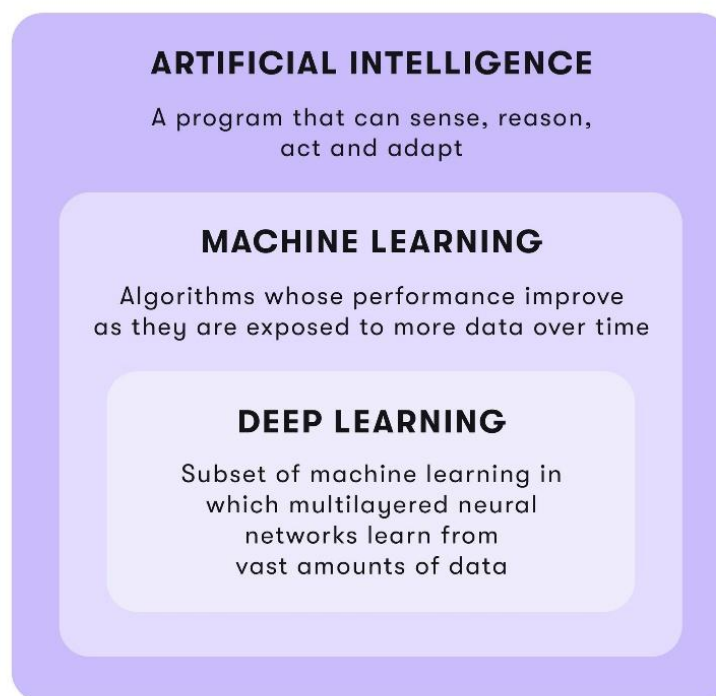


Figura 2.1. Relația între inteligența artificială, învățarea automată și învățarea profundă. [35]

Conceptul de bază al învățării automate constă în dezvoltarea unor modele matematice care analizează seturi mari de date, identificând relații și tipare ascunse. Aceste modele sunt utilizate ulterior pentru a face predicții asupra unor date noi, necunoscute anterior [36]. Spre deosebire de programarea clasică, unde instrucțiunile sunt definite în mod explicit de către programator, în învățarea automată, algoritmul este antrenat pe un set de date și învață singur relațiile dintre variabilele de intrare și rezultatul dorit.

Există *trei tipuri principale de învățare automată*:

- A. **Învățare supravegheată** (supervised learning): algoritmul este antrenat pe un set de date etichetate, unde rezultatul dorit este cunoscut. Exemple de probleme abordate sunt clasificarea și regresia [34].
- B. **Învățare nesupravegheată** (unsupervised learning): algoritmul analizează date neetichetate și descoperă structuri ascunse sau grupări naturale în date. Un exemplu comun este analiza clustering [36].
- C. **Învățare prin consolidare** (reinforcement learning): algoritmul învață prin interacțiuni cu un mediu, primind recompense pentru acțiunile corecte și penalizări pentru cele greșite [34].

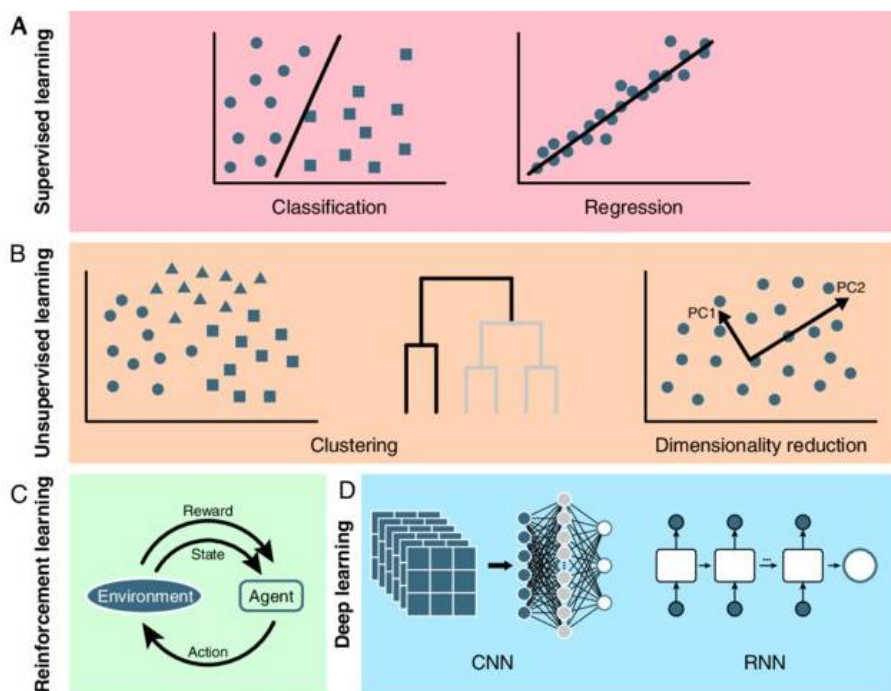


Figura 2.2. Modul de funcționare a celor trei tipuri de învățare automată. [37]

În domeniul științelor naturale, inclusiv în nanotehnologie și chimie, algoritmi de învățare automată sunt tot mai des folosiți pentru a analiza comportamentul AuNP și pentru a prezice procese precum agregarea acestora în prezența diferitelor substanțe, cum ar fi L-cisteina. Aceste aplicații permit optimizarea condițiilor experimentale și reducerea timpului necesar pentru obținerea rezultatelor [38].

Astfel, învățarea automată reprezintă un instrument puternic care facilitează analiza și interpretarea datelor complexe, contribuind la dezvoltarea unor soluții inovatoare în domeniul științific și industrial.

2.2. Algoritmi Utilizați: PLS, Random Forest, SVM și Rețele Neuronale

2.2.1. Partial Least Squares (PLS)

Metoda Partial Least Squares (PLS) este o tehnică statistică utilizată pentru modelarea relațiilor dintre un set de variabile independente (predictori) și un set de variabile dependente (răspunsuri). PLS se încadrează în categoria algoritmilor de *învățare supravegheată* (supervised learning), fiind utilizată în special pentru regresie și modelarea variabilelor latente. Este deosebit de eficientă atunci când datele prezintă multicolaritate și când numărul de predictori este mai mare decât numărul de observații [39].

Această metodă este utilizată frecvent în analiza regresiei, precum și în modelarea ecuațiilor structurale (PLS-SEM), fiind apreciată pentru capacitatea sa de a explica varianța variabilelor latente și pentru robustețea sa în fața datelor imperfecte [40].

Mod de funcționare. PLS combină caracteristicile analizei *componentelor principale* (PCA) cu *regresia multiplă*. Spre deosebire de regresia liniară clasică, PLS *reduce dimensionalitatea datelor* prin construirea unor componente latente (scoruri PLS), care reprezintă combinații liniare ale predictorilor inițiali. Aceste componente sunt alese astfel încât să maximizeze covarianța cu variabilele dependente [39].

1. Modelul pentru matricea variabilelor independente X este exprimat astfel:

$$\begin{array}{c} \text{Inputs} \\ \mathbf{X} \\ (n,p) \end{array} = \begin{array}{c} \text{PLS scores} \\ \mathbf{Z} \\ (n,k) \end{array} \begin{array}{c} \text{PLS loadings} \\ \mathbf{V}^T \\ (k,p) \end{array} + \begin{array}{c} \text{X-Residuals} \\ \mathbf{E} \\ (n,p) \end{array}$$

Figura 2.3. Diagrama matricei pentru datele de intrare. [41]

X – Matricea variabilelor independente (predictorilor), adică datele de intrare.

Z – Matricea componentelor PLS (scorurilor latente), reprezentând combinațiile liniare ale predictorilor care explică cel mai bine variabila dependentă.

V – Matricea încărcăturilor PLS (PLS loadings), care conține coeficienții de ponderare pentru reconstruirea variabilelor originale din componentele latente.

E – Matricea reziduurilor asociate variabilelor independente, adică diferența dintre valorile reale și cele estimate ale predictorilor.

2. Modelul pentru variabila dependentă Y este exprimat astfel:

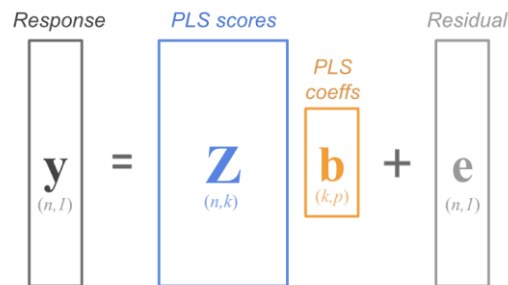


Figura 2.4. Diagrama matricii pentru răspunsuri. [41]

Y – Vectorul (sau matricea) variabilelor dependente (răspunsurilor), adică valorile pe care vrem să le prezicem.

b – Vectorul coeficienților de regresie PLS, care exprimă influența fiecărei componente latente asupra variabilei dependente.

e – Vectorul reziduurilor asociate variabilei dependente, adică diferența dintre valorile reale și cele estimate ale răspunsurilor.

3. Aproximarea pentru X și predicția pentru Y devin:

$$\hat{X} = ZV^T \quad \text{Ecuație 2.1}$$

$$\hat{y} = Zb^{pls} \quad \text{Ecuație 2.2}$$

unde b^{pls} indică faptul că coeficienții obținuți sunt asociați componentelor PLS.

Avantajul major al metodei PLS este că permite modelarea variabilelor latente și poate fi aplicată în situații cu eșantioane mici sau cu multicolaritate puternică în date [42].

Alegerea numărului optim de componente se face prin validare încrucișată: se antrenează modelul cu diferite valori pentru n componente și se selectează valoarea care minimizează eroarea de validare (ex. RMSE, MSE).

2.2.2. Random Forest

Random Forest este un algoritm de învățare automată bazat pe metode de ansamblu, care construiește mai mulți arbori de decizie și combină rezultatele acestora pentru a obține o predicție finală mai precisă și robustă [43]. Acest algoritm a fost dezvoltat pentru a remedia slăbiciunile arborilor de decizie simpli, care pot fi susceptibili la supra-adekvare (overfitting).

Random Forest se încadrează în categoria algoritmilor de *învățare supravegheată* (supervised learning), fiind utilizat atât pentru sarcini de clasificare, cât și pentru regresie [44].

Modul de funcționare:

1. ***Bootstrap Aggregation (Bagging)***: Se generează mai multe seturi de date de antrenament prin eșantionare cu înlocuire (*bootstrap*) din setul original. Fiecare sub-set este folosit pentru a antrena un arbore de decizie.
2. ***Crearea Arborilor de Decizie***: La fiecare nod al unui arbore, se alege un subset aleatoriu de caracteristici în loc să se folosească toate caracteristicile. Arborele crește până la o anumită adâncime sau până când frunzele conțin un număr minim de observații.
3. ***Averaging sau Voting***: Pentru **regresie**, predicțiile arborilor se mediează. Pentru **clasificare**, se face un vot majoritar între arbori.

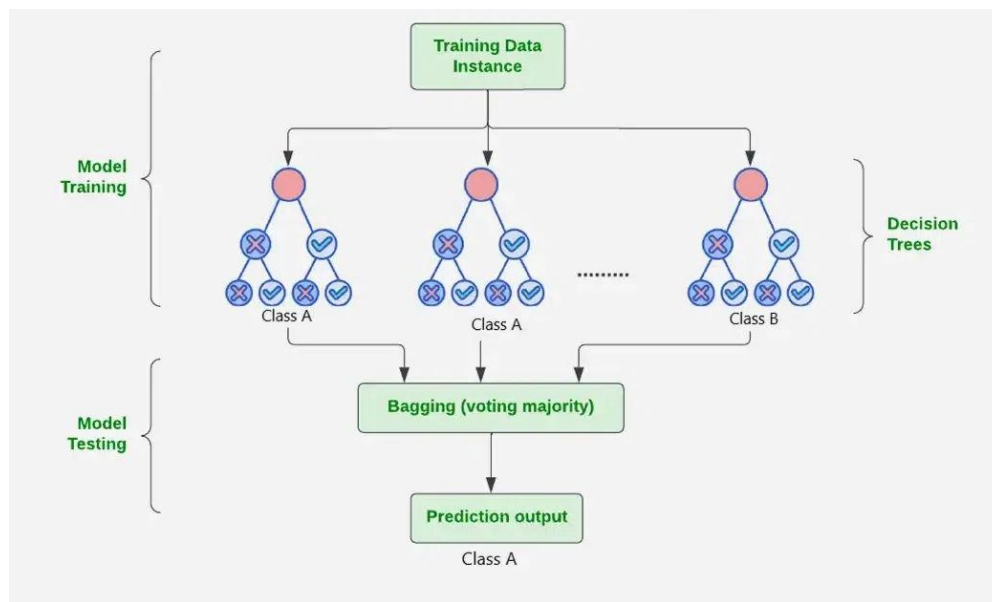


Figura 2.5. Modul de funcționare a algoritmului de învățare Random Forest. [45]

Caracteristici principale a algoritmului menționat:

- Reduce Overfitting-ul:** Arborii de decizie simpli tind să se suprapotrivească pe datele de antrenament, dar combinarea mai multor arbori reduce acest efect.
- Robust la zgomot:** Fiind bazat pe bootstrap și alegeri aleatorii de trăsături, este mai puțin sensibil la valori extreme și date zgomotoase.
- Capabil să gestioneze date mari:** Funcționează bine cu seturi de date mari și cu multe caracteristici.
- Poate lucra cu caracteristici corelate:** Nu e afectat sever de multicolaritate.
- Funcționează atât cu date numerice, cât și categorice.**

2.2.3. Support Vector Machine (SVM)

Algoritmul *Support Vector Machine* (SVM) reprezintă o metodă de clasificare și regresie utilizată în *învățarea automată supravegheată*, fiind recunoscut pentru eficiența sa în rezolvarea problemelor complexe cu date de tip nelinear sau de dimensiuni mari. SVM este utilizat într-o gamă largă de domenii, inclusiv recunoaștere de tipare, bioinformatică, chimie computațională și, mai recent, nanotehnologie.

Principiul de funcționare. Acest algoritm se bazează pe identificarea unui hiperplan optim care separă clasele de date astfel încât distanța dintre cele mai apropiate puncte (marginii) ale claselor să fie maximă[46,47]. În cazul în care datele nu sunt liniar separabile, SVM utilizează **funcții kernel** pentru a transforma datele într-un spațiu de dimensiuni superioare, unde acestea pot fi separate liniar [46]. Cele mai frecvente funcții kernel sunt: Kernel liniar, Kernel polinomial, Kernel RBF (Radial Basis Function), Kernel sigmoid.

Hiperplanul este determinat pe baza unui subset de puncte de antrenare numite **vectori suport**, care definesc granițele dintre clase. [48]

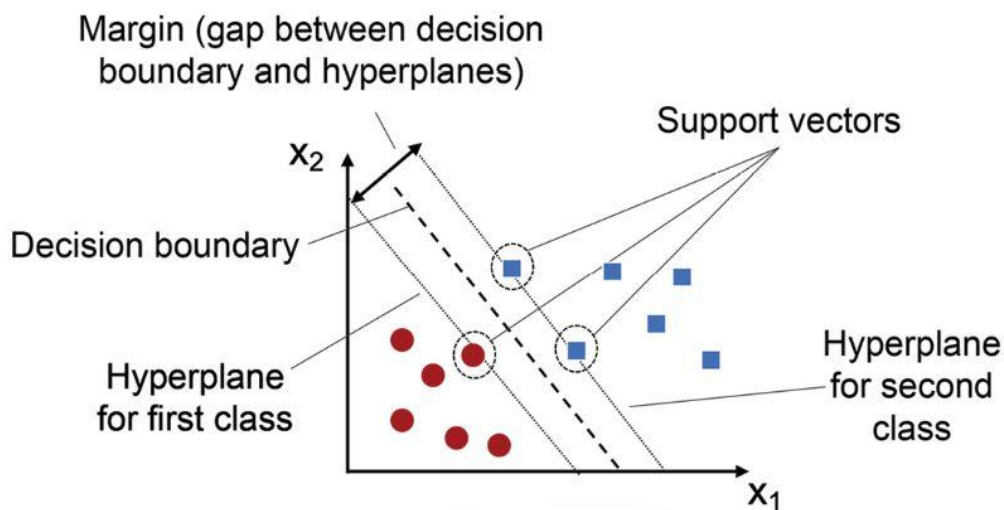


Figura 2.6. Schematizarea hiperplanului cu vectorii suport. [49]

2.2.4. Rețele Neuronale

Rețelele neuronale artificiale (RNA) reprezintă un model matematic inspirat din funcționarea creierului uman, utilizat pentru a rezolva probleme complexe din domeniul inteligenței artificiale și al învățării automate. Acestea constau într-o rețea de noduri, numite neuroni artificiali, interconectați prin conexiuni ponderate, care procesează și transmit informații [50]. Acestea constituie o componentă fundamentală a învățării automate moderne, contribuind semnificativ la dezvoltarea tehnologiilor de inteligență artificială, cu aplicabilitate extinsă în diverse domenii, precum viziunea computerizată, procesarea limbajului natural și robotica. [51]

Rețelele neuronale se încadrează în cadrul învățării automate (Machine Learning) ca o tehnică specifică de învățare profundă (Deep Learning). Acestea pot fi utilizate în trei tipuri principale de învățare:

- Învățare supravegheată: Rețeaua este antrenată pe un set de date etichetate, unde fiecare intrare are un răspuns corect asociat. Scopul este de a generaliza relația dintre intrări și ieșiri. [52]
- Învățare nesupravegheată: Datele nu sunt etichetate, iar rețeaua trebuie să identifice structuri și tipare din setul de date. [51]
- Învățare prin consolidare (Reinforcement Learning): Rețeaua învață prin interacțiunea cu un mediu, primind recompense pentru acțiunile corecte și penalizări pentru cele greșite. [50]

Funcționarea RNA se bazează pe transmiterea informației în rețeaua de **neuroni artificiali**. Fiecare neuron primește semnale de intrare ponderate, le combină și aplică o funcție de activare pentru a genera o ieșire. Aceasta este transmisă mai departe către alți neuroni până când rețeaua produce rezultatul final. Procesul de antrenare presupune ajustarea ponderilor conexiunilor utilizând algoritmi precum propagarea inversă a erorii (**Backpropagation**). [52]

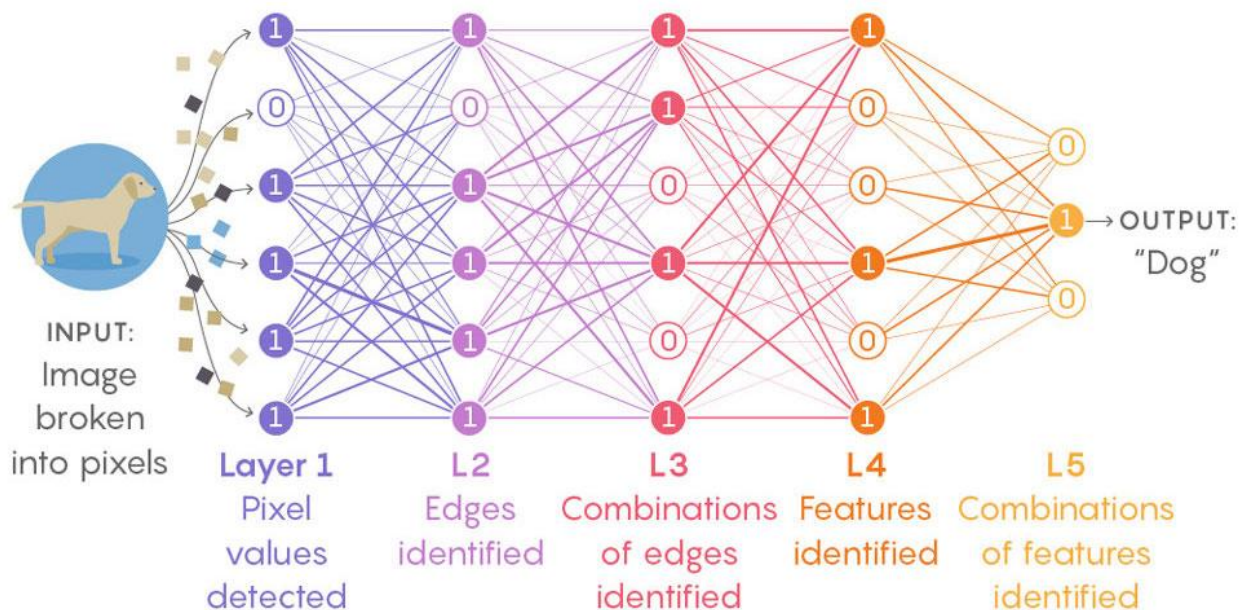


Figura 2.7. Exemplu de mod de funcționare a unei rețele neuronale artificiale. [53]

	<i>Avantaje</i>	<i>Dezavantaje</i>
PLS	<ol style="list-style-type: none"> Poate lucra cu date unde numărul variabilelor este mai mare decât numărul de observații. Este robust la multicolaritate (relații puternice între variabilele independente). Reduce dimensionalitatea datelor, făcând analiza mai eficientă. Integrează procesul de reducere a dimensiunii cu regresia, optimizând ambele simultan. 	<ol style="list-style-type: none"> Alegerea numărului de componente latente poate influența puternic performanța. Nu este cea mai precisă metodă pentru seturi de date mici sau necorelate. Interpretabilitatea componentelor poate fi uneori dificilă.
Random Forest	<ol style="list-style-type: none"> Robustețe la supra-adekvare datorită ansamblului de arbori [43]. Capacitate bună de generalizare pentru date necunoscute. Manevrarea eficientă a unui număr mare de variabile explicative și a interacțiunilor dintre acestea [54]. Posibilitatea de a estima importanța variabilelor utilizate în predicții. 	<ol style="list-style-type: none"> Complexitatea crescută și nevoia mai mare de resurse computaționale comparativ cu modele simple precum regresia logistică [44]. Interpretabilitate redusă față de modelele liniare sau arborii de decizie individuali.
SVM	<ol style="list-style-type: none"> Eficiența ridicată în probleme cu dimensiuni mari Performanță excelentă în cazul datelor neliniare datorită funcțiilor kernel Capacitate bună de generalizare pe date necunoscute 	<ol style="list-style-type: none"> Sensibilitate la alegerea parametrilor kernel și a parametrului de regularizare Consum ridicat de resurse computaționale pentru seturi mari de date Interpretabilitate redusă comparativ cu alte metode
Rețele neuronale	<ol style="list-style-type: none"> Capacitatea de a modela relații complexe în date. Adaptabilitate și capacitate de generalizare. Performanțe remarcabile în recunoașterea tiparelor, clasificare și predicții. 	<ol style="list-style-type: none"> Necesită volume mari de date pentru antrenare. Complexitate computațională ridicată. Opacitate (“cutie neagră”) – dificultatea de a interpreta modul în care rețeaua ajunge la o decizie.

Tabel 2.1. Avantajele și dezavantajele metodelor de învățare automată selectate.

3. Materiale și Metode

3.1. Sinteza AuNP și caracterizarea prin spectroscopie UV-Vis

Nanoparticulele de aur prezintă proprietăți optice excepționale în regiunea vizibilă a spectrului electromagnetic, datorită rezonanței plasmonice de suprafață. Aceste caracteristici fac din ele materiale de interes major pentru aplicații în biosenzori, imagistică și detecție chimică. Performanța acestor aplicații este determinată în mare măsură de dimensiunea, forma și distribuția NP, precum și de natura chimică a microambientului. Din acest motiv, sinteza controlată a AuNP este o etapă esențială în pregătirea sistemelor coloidale funcționale. [55]

Pentru sinteza AuNP, există mai multe opțiuni de sintetizare, însă una dintre cele mai folosite este **metoda Turkevich**. Aceasta presupune reducerea tetraclorurii de aur (HAuCl_4) în soluție apoasă, utilizând citrat de sodiu ($\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$) atât ca **agent reducător**, cât și ca **stabilizator**. Reacția are loc la temperaturi apropiate de punctul de fierbere al apei, sub agitare constantă, pentru a asigura o distribuție termică uniformă. În urma reacției, aurul trece din forma Au^{3+} în forma sa elementară, Au^0 , conducând la formarea de **nanoparticule sferice**. [55]

Relația dintre concentrația de HAuCl_4 și cea de citrat de sodiu este esențială pentru controlul dimensiunii nanoparticulelor. Variația raportului molar $R = [\text{Au}^{3+}]/[\text{Na}_3\text{-cit}]$ influențează direct diametrul mediu al particulelor obținute. Conform studiilor realizate de Frens, la valori mici ale lui R se obțin particule de dimensiuni reduse, cu o culoare roșu aprins, în timp ce la valori mari ale lui R se formează particule de până la 147 nm, care conferă soluției o nuanță violetă. *Tabel 3.1* ne arată aceste rezultate experimentale. [55]

$[\text{HAuCl}_4](\text{mM})$	$[\text{Na}_3\text{-cit}](\text{mM})$	R	d(nm)	Culoare
0,294	0,766	0,38	16	Roșu aprins
	0,575	0,51	24,5	Roșu
	0,383	0,77	41	Roșu
	0,230	1,28	71.5	Vișiniu
	0,161	1,83	97,5	Violet
	0,123	2,39	147	Violet

Tabel 3.1. Datele care indică concentrațiile de Au^{+3} și de citrat de sodiu pentru a varia R cu scopul de a forma AuNP, având diametre d diferite. [55]

Această metodă conduce la formarea de NP aproape monodisperse, cu o variație sub 10% în diametru, ceea ce este esențial pentru aplicațiile care necesită reproductibilitate ridicată. În plus, o mică proporție (~5%) din particule poate prezenta o formă triunghiulară, indicând o ușoară anizotropie. Moleculele de citrat formează un strat negativ încărcat la suprafața nanoparticulelor, ceea ce asigură o stabilitate coloidală ridicată, prevenind agregarea acestora (potențialul Z fiind de aproximativ -40 mV). [55]

În cadrul acestei lucrări, a fost utilizată această procedură standardizată pentru sinteza AuNP, în scopul funcționalizării acestora cu molecule de L-cisteină și al antrenării algoritmilor de învățare automată pentru detecția concentrației acestora. Au fost obținute trei seturi de coloizi auriferi prin reducerea unei soluții de HAuCl_4 1 mM cu citrat de sodiu în două concentrații diferite: 43 mM și 77 mM. Astfel, s-au obținut NP cu diametrul de aproximativ 23 nm (pentru concentrația de 43 mM) și două seturi cu dimensiuni de aproximativ 13 nm (pentru concentrația de 77 mM), care au fost utilizate ulterior în experimentele de funcționalizare și detecție.

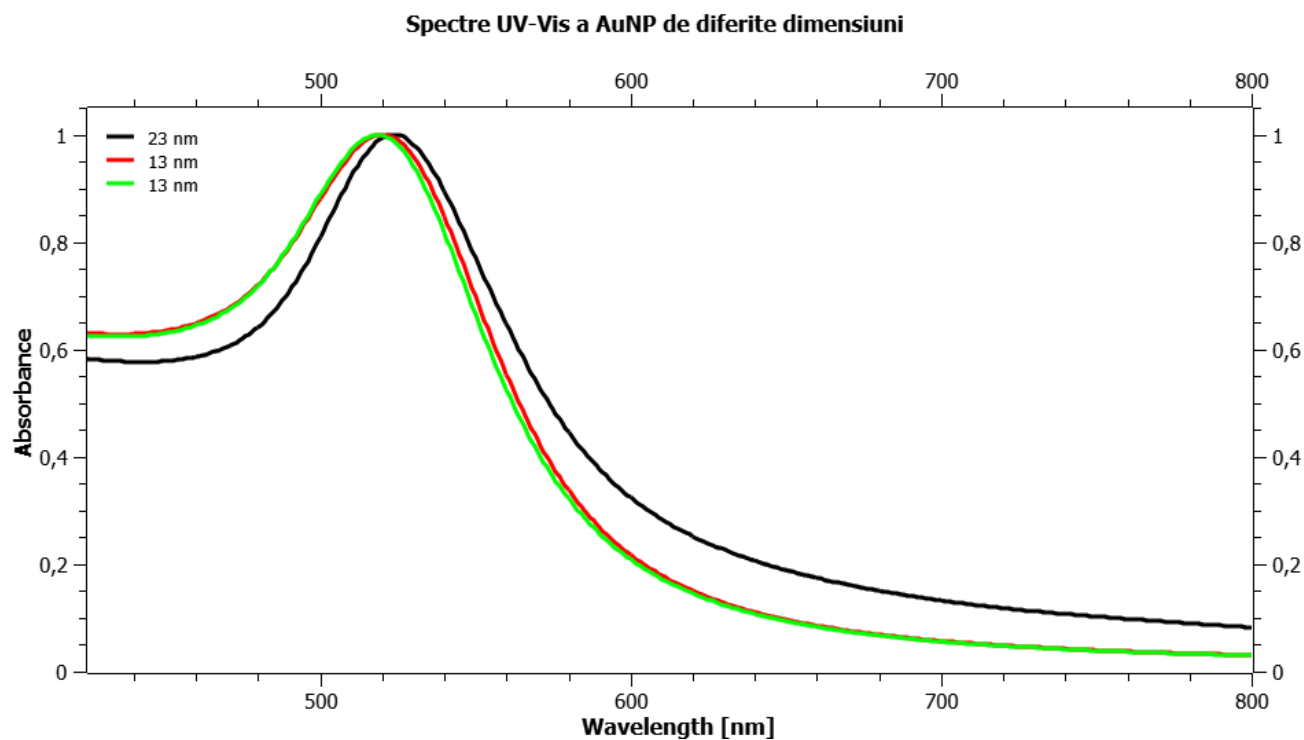


Figura 3.1. Spectrele celor trei seturi de nanoparticule sintetizate prin metoda Turkevich, având dimensiunile de 13nm, respectiv 23nm.

Caracterizarea optică a AuNP este esențială pentru validarea sintezei și determinarea dimensiunilor și concentrației acestora în soluție. Proprietățile optice ale AuNP în dispersii apoase sunt puternic dependente de mărimea, forma și distribuția particulelor, iar una dintre cele mai utile metode de caracterizare este spectrofotometria UV-Vis. Această metodă permite analiza benzilor de absorbție plasmonică, care apar ca urmare a rezonanței plasmonilor de suprafață SPR, fenomen specific nanoparticulelor metalice. [55]

În decursul lucrării, pentru achiziția spectrelor de absorbție UV-Vis ale nanoparticulelor sintetizate, s-a utilizat spectrometrul **Jasco V630**. Acesta a oferit măsurători precise și reproductibile în intervalul de lungimi de undă corespunzător benzilor plasmonice ale AuNP, în special în jurul valorii de 520 nm, specifică nanoparticulelor sferice. [55]

Pentru estimarea diametrului particulelor, s-a utilizat metoda propusă de Haiss *et al.* (2007) care permite determinarea dimensiunii medii a AuNP, pornind de la raportul dintre absorbția la maximul benzii plasmonice (A_{SPR}) și absorbția la 450 nm (A_{450}). Relația matematică utilizată este exprimată în următoarea ecuație:

$$d = \exp\left(3 \frac{A_{SPR}}{A_{450}} - 2,20\right) \quad \text{Ecuație 3.1}$$

Această metodă este validată pentru particule sferice cu diametrul sub 35 nm, obținând rezultate comparabile cu cele determinate prin microscopia electronică de transmisie (TEM). [55]

În continuare, cunoscând dimensiunea particulelor (d), absorbția la 450 nm și coeficientul de extincție molară (ϵ), se poate determina concentrația AuNP în soluția finală utilizând următoarele relații [55]:

$$\ln \epsilon_{450} = 3,0869 \ln d + 10,869 \quad \text{Ecuație 3.2}$$

$$C = \frac{A_{450}}{\epsilon_{450}} \cdot \text{DilutionFactor} \quad \text{Ecuație 3.3}$$

Aplicând ecuațiile 3.1 și 3.3 pentru soluțiile sintetizate, au fost obținute diametre medii de aproximativ 13 nm și 23 nm, în concordanță cu valorile estimate din raportul concentrațiilor reactivilor utilizați (R). Concentrațiile finale ale AuNP calculate au fost de $3,05 \cdot 10^{-9}$ M pentru nanoparticulele de 13 nm și $7,55 \cdot 10^{-10}$ M pentru cele de 23 nm.

3.2. Optimizarea funcționalizării AuNP cu L-cisteină

Pentru optimizarea funcționalizării AuNP cu L-cisteină, am preparat inițial o soluție de L-cisteină cu concentrația de 10^{-2} M. În cadrul sintezei prin aglutinare controlată, s-a utilizat un raport molar de 1:1 între AuNP și L-cisteină. Acest raport a condus la rezultate rapide și evidente, manifestate printr-o schimbare colorimetrică caracteristică, indicând agregarea nanoparticulelor.

S-a observat că dimensiunea NP influențează semnificativ cinetica procesului de agregare. Pentru AuNP cu diametrul de 23 nm, agregarea s-a produs într-un interval de aproximativ 30 de minute. În schimb, pentru NP de 13 nm, procesul de agregare devenea vizibil încă din primul minut. Această diferență semnificativă a condus la investigarea efectului altor parametri, precum pH-ul și temperatura, asupra eficienței funcționalizării.

În prima etapă, am variat pH-ul amestecului format din AuNP și L-cisteină. Soluția de pornire avea un pH neutru (~ 7). Pentru modificarea acestuia, s-a adăugat NaOH 1% pentru a obține un pH alcalin (~ 8) și HNO_3 pentru obținerea unui pH acid (~ 4). Spectrele UV-Vis obținute (Figura 3.2) arată comportamentul agregării AuNP în condiții de pH diferite. Rezultatele indică faptul că un pH alcalin încetinește agregarea, în timp ce un pH acid determină deplasări plasmonice excesive, ceea ce sugerează o agregare necontrolată. pH-ul neutru s-a dovedit a fi optim, oferind o deplasare plasmonică echilibrată și reproductibilă.

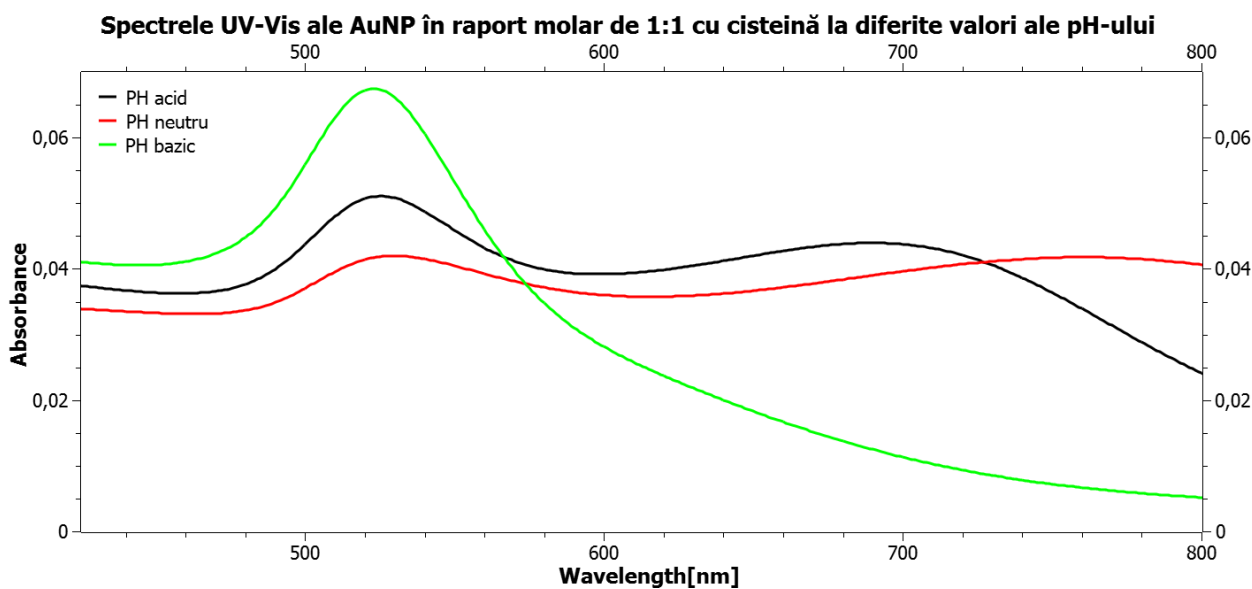


Figura 3.2. Comportamentul agregării AuNPs de 23 nm în prezența cisteinei la pH diferit.

În continuare, s-au investigat efectele temperaturii și ale concentrației de L-cisteină. Amestecurile de AuNP și cisteină, cu concentrații variind între 10^{-2} M și 10^{-9} M, au fost încălzite timp de 20 de minute la temperaturi ridicate ($\sim 60^\circ\text{C}$). Spectrele obținute (Figura 3.3) arată că la concentrații ridicate de cisteină au loc deplasări plasmonice excesive, însoțite de aplatizarea spectrului UV-Vis, indicând agregare rapidă și instabilă. La concentrații mai scăzute, schimbările plasmonice au fost mai controlate și mai progresive. Astfel, metoda de încălzire s-a dovedit inefficientă pentru analiză, din cauza pierderii de volum prin evaporare.

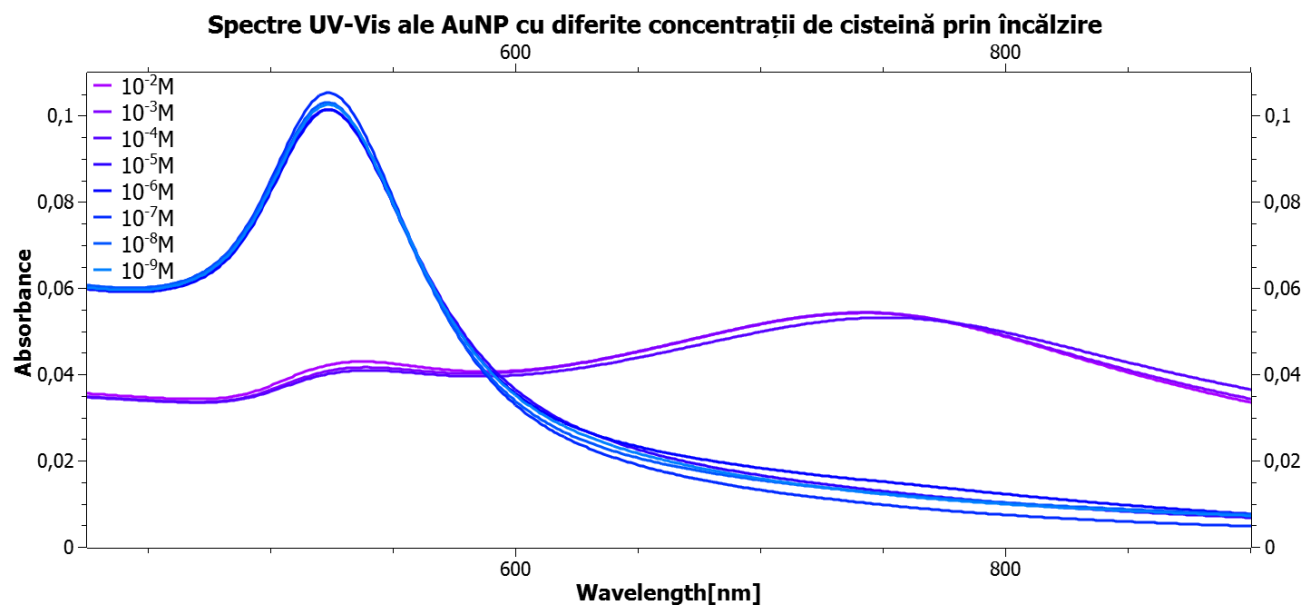


Figura 3.3. Comportamentul agregării AuNP de 23 nm în funcție de concentrația de cisteină, în condiții de încălzire.

Pe baza rezultatelor obținute, s-a concluzionat că metoda optimă pentru funcționalizarea AuNPs cu L-cisteină constă în menținerea amestecului la temperatura camerei (25°C), utilizarea unui pH neutru și preferința pentru nanoparticule de dimensiuni mai mici (13 nm), pentru a reduce timpul necesar agregării și pentru a obține rezultate reproductibile și stabile.

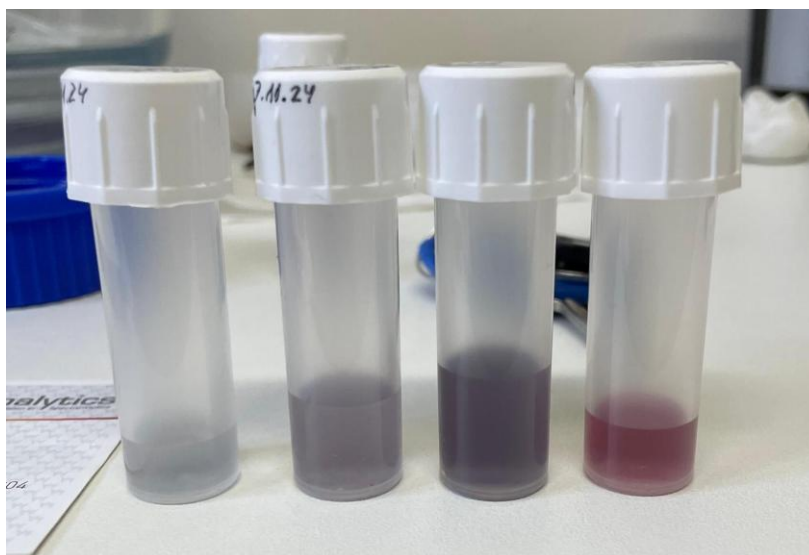


Figura 3.4. Aspectul vizual al soluțiilor de AuNP funcționalizate cu L-cisteină, obținute prin metode diferite: (1) încălzire timp de 20 minute, (2) condiții de pH acid (~4), (3) pH neutru (~7), și (4) pH bazic (~8). Se observă diferențe evidente de culoare, asociate cu gradul de agregare al NP.

3.3. Utilizarea platformei Quasar pentru preprocesarea și testarea modelelor de învățare automată

În această lucrare, pentru prelucrarea datelor spectroscopice și dezvoltarea modelelor de regresie, a fost utilizată platforma **Quasar**. Aceasta oferă un mediu vizual interactiv care permite construirea de fluxuri de analiză prin conectarea unor module funcționale, fără a necesita programare explicită. Astfel, Quasar este ideală pentru aplicații științifice precum spectroscopia, oferind atât flexibilitate, cât și control asupra etapelor de procesare.

Primul pas în cadrul fluxului a fost importul fișierelor cu spectre, realizat prin widget-ul **File**, care permite încărcarea datelor în format compatibil (.csv, .txt, .xls etc.). Ulterior, datele au fost trimise către componenta **Preprocess Spectra**, unde au fost aplicate etape esențiale de preprocesare, precum **normalizarea** spectrelor pentru a elimina variațiile de intensitate nedorite. Această etapă a asigurat crearea unui set de date uniform și coerent, pregătit pentru analiza ulterioară.

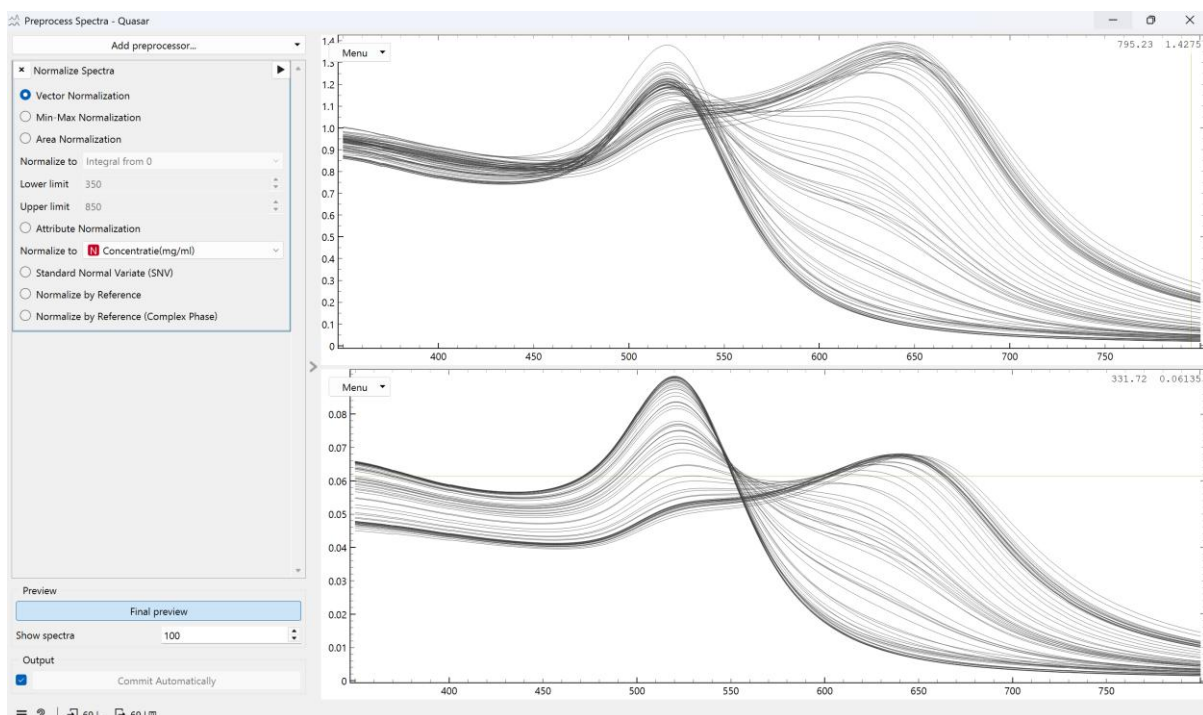


Figura 3.5. Preprocesarea spectrelor, folosind normalizarea vectorială.

După preprocesare, datele au fost transmise către widget-ul **Select Columns**, unde am identificat și selectat variabilele de intrare și variabila țintă, care în cazul nostru este concentrația de cisteină. Vizualizarea tabelară intermediară a fost asigurată de modulul **Data Table**, care a permis o inspecție rapidă a valorilor selectate.

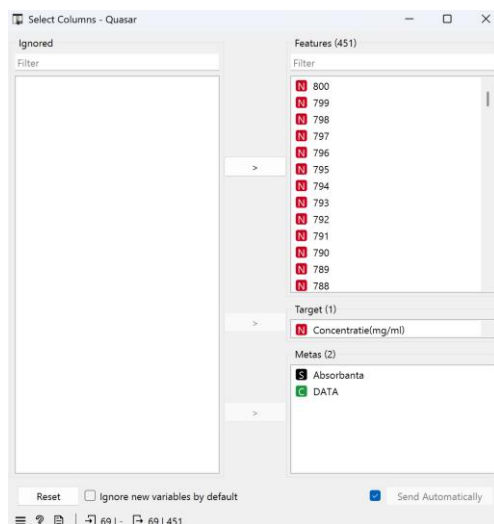


Figura 3.6. Interfața de selecție a caracteristicilor și a variabilei țintă. Cele 451 de lungimi de undă au fost selectate ca predictorii (features), iar concentrația de cisteină (Concentrație (mg/mL)) a fost setată ca variabilă țintă (target) pentru antrenarea modelelor de regresie.

Apoi, pentru a evalua performanța modelelor, am utilizat widget-ul **Data Sampler** pentru împărțirea datelor în seturi de antrenare și testare.

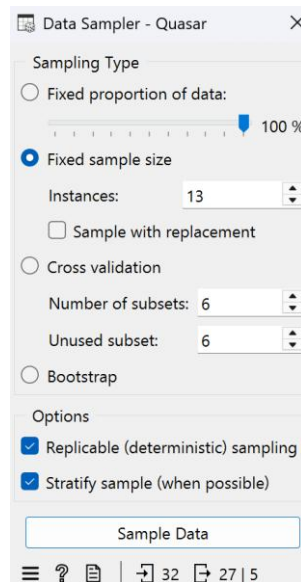


Figura 3.7. A fost ales un eșantion fix de instanțe pentru antrenare, cu opțiunile de eșantionare deterministă și stratificată activate. Restul probelor au fost utilizate pentru testare. În etapele ulterioare, mărimea setului de testare a fost crescută treptat pentru a evalua stabilitatea modelelor.

Datele astfel pregătite au fost introduse în patru algoritmi diferiți de învățare automată, implementați prin widget-urile dedicate din Quasar: **PLS**, **SVM**, **Random Forest** și **Neural Network**.

Fiecare model a fost conectat la același set de date de testare și antrenare, asigurând astfel un cadru de comparație uniform.

Modelele antrenate au fost conectate la widget-ul **Predictions**, care a combinat rezultatele obținute și le-a furnizat ca ieșire. Predicțiile au fost comparate cu valorile reale, iar evaluarea vizuală a fost realizată prin widget-ul **Scatter Plot**, care a permis inspecția corelației dintre predicții și datele reale.

Acest flux vizual (Figura 3.8) a permis nu doar testarea eficientă a diverselor modele, ci și o înțelegere intuitivă a procesului de învățare automată, într-un mod reproductibil și documentabil.

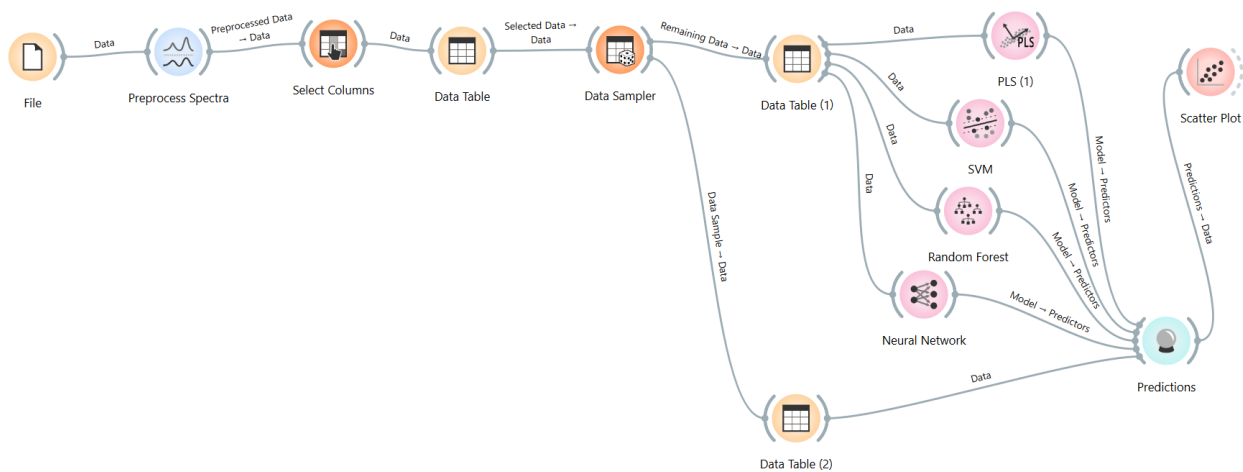


Figura 3.8. Diagrama fluxului de lucru utilizat în platforma Quasar pentru analiza spectrelor UV-Vis. Datele sunt prelucrate, apoi împărțite în seturi de antrenare și testare, urmate de antrenarea a patru modele de regresie: PLS, SVM, Random Forest și Neural Network. Predicțiile generate sunt comparate vizual în modulul de tip Scatter Plot.

4. Rezultate și Discuții

4.1. Procesarea și analiza datelor experimentale

Procesarea datelor a implicat atât modificarea, cât și analiza factorilor care pot influența rezultatele determinării concentrației de L-cisteină, precum porozitatea AuNP și domeniul de predicție utilizat în modelele de învățare automată. Pentru identificarea parametrilor optimi, s-au realizat modificări pas cu pas ale variabilelor implicate, urmărind obținerea celor mai bune rezultate din punct de vedere al preciziei și robusteții predicțiilor.

4.1.1. Comportamentul AuNP la concentrații mari (3-120mg/mL) și concentrații foarte mici (0-0,5mg/mL) de l-cisteină

S-a urmărit determinarea unității de măsură corespunzătoare pentru exprimarea concentrației, precum și verificarea comportamentului AuNP în prezența concentrațiilor ridicate de L-cisteină (3 – 120 mg/mL) și a celor foarte scăzute (0 – 0,5 mg/mL).

Pentru un set de 18 concentrații diferite de L-cisteină, am aplicat cele patru metode de învățare automată. În Tabel 4.1 sunt prezentate valorile concentrațiilor folosite de L-cisteină, exprimate atât în moli, cât și în mg/mL:

Nr. crt.	Concentrație (M)	Concentrație (mg/mL)
1	10^{-3}	121,160
2	$5,00 \cdot 10^{-4}$	60,580
3	$1,00 \cdot 10^{-4}$	12,120
4	$5,00 \cdot 10^{-5}$	6,0580
5	$3,25 \cdot 10^{-5}$	3,938
6	$2,44 \cdot 10^{-5}$	2,956
7	$1,63 \cdot 10^{-5}$	1,975
8	$8,13 \cdot 10^{-6}$	0,985
9	$2,96 \cdot 10^{-6}$	0,359
10	$2,22 \cdot 10^{-6}$	0,269

11	$1,48 \cdot 10^{-6}$	0,179
12	$7,39 \cdot 10^{-7}$	0,089
13	$2,70 \cdot 10^{-7}$	0,032
14	$1,34 \cdot 10^{-7}$	0,016
15	$6,83 \cdot 10^{-8}$	0,008
16	$2,60 \cdot 10^{-8}$	0,003
17	$1,95 \cdot 10^{-8}$	0,002
18	0	0

Tabel 4.1. Tabelul concentrațiilor de l-cisteină studiate pe AuNP.

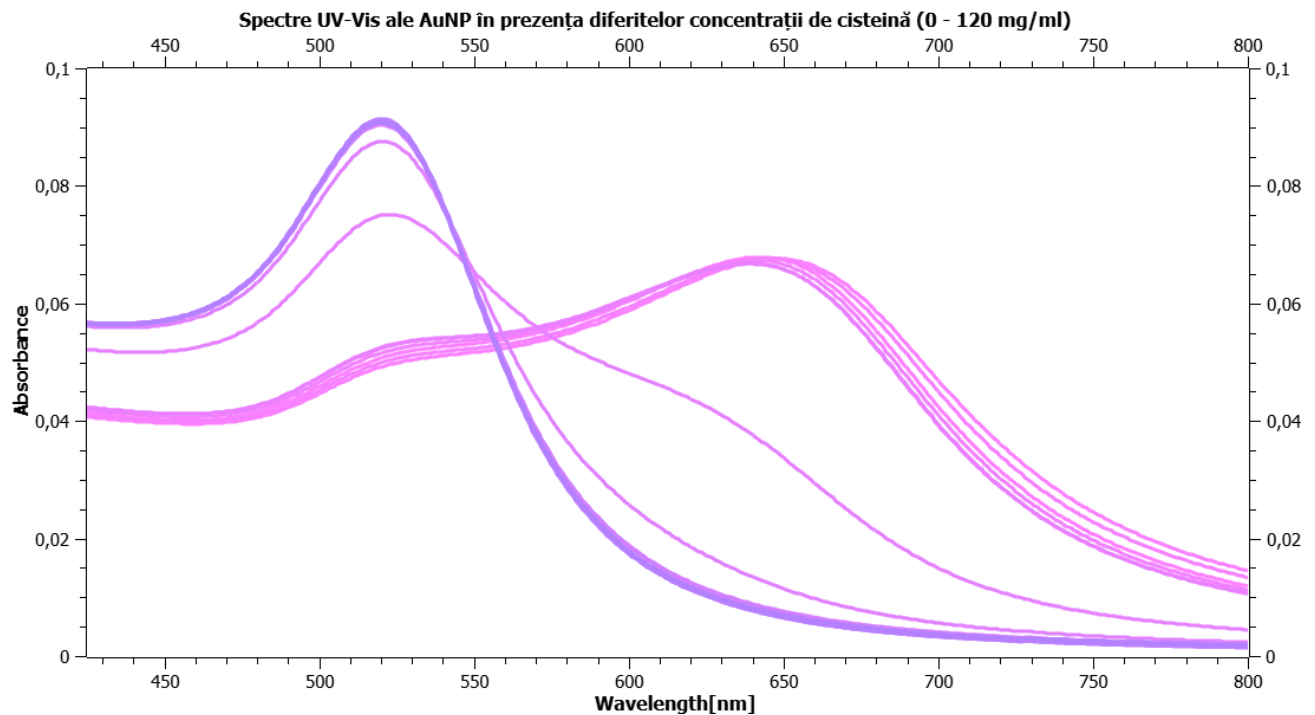


Figura 4.1. Spectrul soluțiilor de AuNP în prezența diferitelor concentrații de cisteină, asociate Tabel 4.1. Se observă un gol spectral între valorile mari (2-120mg/mL) și cele foarte mici (0-0,4mg/mL), unde se regăsesc probele cu concentrații intermediare de 1,975 mg/mL și 0,985mg/mL.



Figura 4.2. Aspectul vizual al soluțiilor din Tabel 4.1, unde se poate remarca variația culorii coloidului în funcție de concentrație: de la roșu rubiniu (pentru AuNP inițiale), trecând prin nuanțe de vișiniu și mov, până la albastru intens. Această schimbare cromatică indică trecerea de la NP dispersate la NP agregate, care ulterior se depun la baza recipientului.

Pentru analiza detecției de cisteină, datele obținute au fost introduse în platforma **Quasar**, unde au fost testate cele patru metode de învățare automată: **regresia PLS (Partial Least Squares)**, **Random Forest**, **SVM (Support Vector Machines)** și **rețele neuronale artificiale**.

Inițial, datele au fost exprimate în moli, dar, datorită valorilor foarte mici, modelele nu reușeau să genereze rezultate semnificative (valorile R^2 erau 0, iar RMSE nu putea fi calculat). Astfel, s-a ales continuarea procesării datelor în unități de **mg/mL**, care oferă o scală mai potrivită pentru antrenarea și testarea modelelor.

Pentru evaluarea performanței fiecărui model, s-au urmărit două metrici esențiale:

- **R^2** (coeficientul de determinare), care ar trebui să fie cât mai aproape de 1 pentru a indica o corelație puternică între valorile reale și cele prezise;
- **RMSE** (Root Mean Square Error), care trebuie să fie cât mai apropiat de 0, pentru a reflecta o eroare mică a predicției.

În prima etapă de testare, **nu s-a aplicat un data sampler** – toate datele au fost introduse în modele, iar predicțiile au fost realizate pe același set de date. Această abordare inițială permite verificarea comportamentului modelului fără influența variației introduse de selecția aleatorie a subseturilor de test. Ulterior, pentru validare, se va aplica și o împărțire a datelor în seturi de antrenare și test.

Pentru ajustarea modelelor, au fost realizați următorii pași:

- ♥ Pentru **regresia PLS**, s-a modificat **numărul de componente latente**, pentru a optimiza capacitatea de predicție. Iar, pentru a evita supraînvățarea (overfitting- modelul "învață" zgomotul

din date în loc să extragă relația reală dintre variabile), vom stabili ca numărul de componente să fie mai mic decât o treime din numărul total de probe

- ♥ În cazul **Random Forest**, s-a ajustat numărul de arbori de decizie (*trees*) pentru a îmbunătăți acuratețea și stabilitatea modelului. Pentru seturi de date reduse, cum este cel analizat în această lucrare, nu este necesar un număr mare de arbori. Am urmărit evitarea **sub-antrenării (underfitting)**, în cazul unui număr prea mic de arbori, precum și evitarea utilizării ineficiente de resurse, în cazul unui număr foarte mare de arbori, care nu ar aduce beneficii semnificative. Numărul optim de copaci a fost ales pe baza observației erorii RMSE și a valorii R^2 , în funcție de comportamentul modelului.
- ♥ Pentru **SVM** au fost păstrați parametrii prestabiliți (impliciti), întrucât scopul testării inițiale a fost acela de a observa comportamentul general al modelului fără optimizări suplimentare. Această alegere se justifică prin dorința de a evalua performanța de bază a metodei, urmând ca eventualele îmbunătățiri să fie aplicate ulterior.
- ♥ Pentru **rețelele neuronale**, a fost modificat numărul de neuroni din stratul ascuns pentru a observa impactul asupra performanței modelului. Ajustarea acestui parametru are un rol esențial în capacitatea de învățare a rețelei: un număr prea mic de neuroni poate conduce la sub-antrenare (underfitting), în timp ce un număr prea mare poate duce la supraînvățare (overfitting). Astfel, au fost testate mai multe configurații (ex. 5, 10, 15 neuroni în stratul ascuns), urmărind valorile R^2 și RMSE obținute. S-a ales varianta optimă în funcție de echilibrul între acuratețea predicției și stabilitatea modelului.

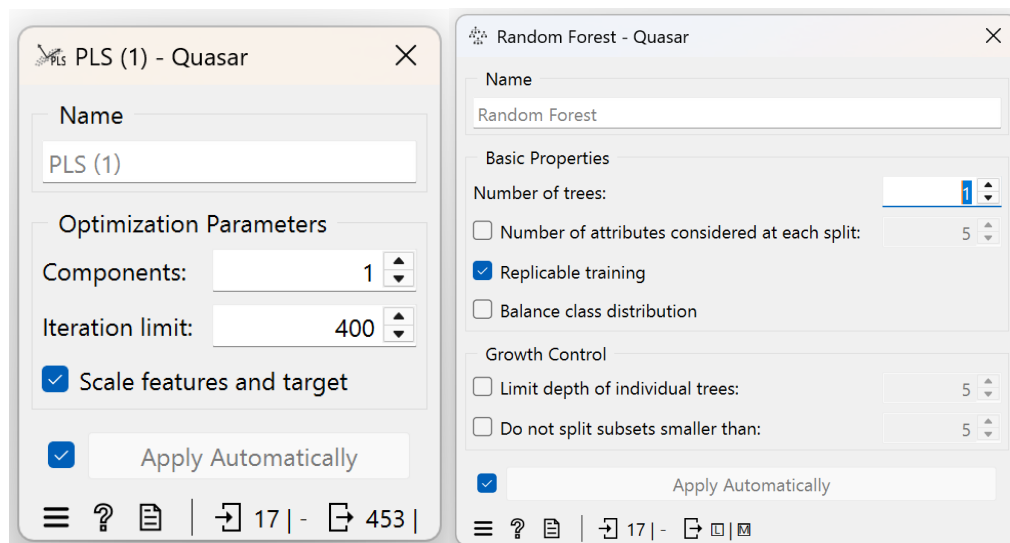


Figura 4.3. Parametri implicați de funcționare a PLS-ului și a Random Forest-ului în Quasar.

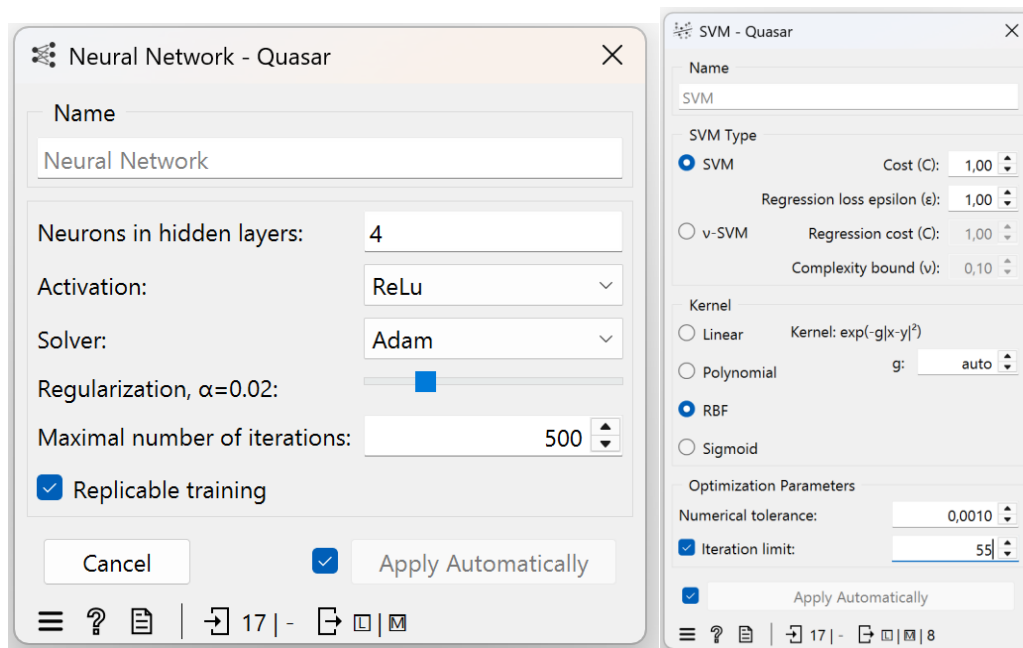


Figura 4.4. Parametri implicați de funcționare a SVM-ului și a rețelelor neuronale în Quasar.

În urma ajustărilor efectuate pentru fiecare model de învățare automată, s-au obținut următoarele rezultate:

- ♣ **Regresia PLS** a înregistrat cea mai bună performanță la un număr de **6 componente latente**, cu un coeficient de determinare $R^2 = 0,553$ și o eroare RMSE de **20**. Acest rezultat indică o capacitate moderată de predicție, specifică modelelor liniare aplicate pe seturi reduse de date.
- ♣ **Rețelele neuronale** au prezentat valori relativ constante ale metricilor de performanță, indiferent de numărul de neuroni din stratul ascuns (atâta timp cât acesta a fost mai mare de 5). Valorile R^2 au variat între **0,3 și 0,4**, iar **RMSE între 23 și 24**, ceea ce sugerează o capacitate predictivă limitată în contextul actual al datelor.
- ♣ **Modelul SVM** a avut rezultate slabe, cu **valori negative ale R^2** , ceea ce indică faptul că modelul nu reușește să surprindă relația dintre spectre și concentrațiile de cisteină.
- ♣ **Random Forest** a oferit cele mai bune rezultate dintre toate modelele testate. Pentru un număr de arbori mai mare de 10, s-au obținut valori ale R^2 între **0,8 și 0,9**, iar **RMSE a variat între 11 și 12**, demonstrând o capacitate ridicată de generalizare, chiar și în condiții de set de date redus.

Aceste rezultate sugerează că, pentru îmbunătățirea performanței modelelor, este necesară o extindere a setului de date. Din acest motiv vom crește numărul de probe analizate ca să avem o distribuție mai uniformă pe întregul interval de concentrații, pentru a elimina golurile spectrale și a asigura o mai bună acoperire a domeniului de predicție.

4.1.2. Comportamentul AuNP la concentrații de l-cisteină în intervalul 3-0,5 mg/mL

În această secțiune a lucrării, a fost acoperit întregul interval de concentrații de la cele foarte mari până la cele foarte mici, fiind preparate 69 de probe corespunzătoare unui număr total de 58 de concentrații distincte. Dintre acestea, 27 de concentrații au fost situate în intervalul critic 0,5–3,0 mg/mL, 16 concentrații au fost alese din domeniul 3–120 mg/mL, iar 15 din domeniul 0–0,5 mg/mL. Această abordare a permis o acoperire uniformă a domeniului de interes, cu o densitate mai mare a punctelor experimentale în zona critică, unde se așteaptă cele mai semnificative variații ale semnalului.

Nr. crt.	Concentrație (M)	Concentrație (mg/mL)	Nr. crt.	Concentrație (M)	Concentrație (mg/mL)
1	10^{-3}	121,160	30	$1,63 \cdot 10^{-5}$	1,975
2	$9,08 \cdot 10^{-4}$	110,000	31	$1,57 \cdot 10^{-5}$	1,900
3	$8,25 \cdot 10^{-4}$	100,000	32	$1,49 \cdot 10^{-5}$	1,800
4	$7,43 \cdot 10^{-4}$	90,000	33	$1,40 \cdot 10^{-5}$	1,700
5	$6,60 \cdot 10^{-4}$	80,000	34	$1,32 \cdot 10^{-5}$	1,600
6	$5,00 \cdot 10^{-4}$	60,580	35	$1,24 \cdot 10^{-5}$	1,500
7	$4,13 \cdot 10^{-4}$	50,000	36	$1,16 \cdot 10^{-5}$	1,400
8	$3,30 \cdot 10^{-4}$	40,000	37	$1,07 \cdot 10^{-5}$	1,300
9	$2,48 \cdot 10^{-4}$	30,000	38	10^{-5}	1,212
10	$1,65 \cdot 10^{-4}$	20,000	39	$9,08 \cdot 10^{-6}$	1,100

11	10^{-4}	12,116	40	$8,13 \cdot 10^{-6}$	0,985
12	$8,25 \cdot 10^{-5}$	10,000	41	$5,78 \cdot 10^{-6}$	0,700
13	$6,60 \cdot 10^{-5}$	8,000	42	$5 \cdot 10^{-6}$	0,606
14	$5,00 \cdot 10^{-5}$	6,058	43	$4,13 \cdot 10^{-6}$	0,500
15	$3,30 \cdot 10^{-5}$	4,000	44	$2,96 \cdot 10^{-6}$	0,359
16	$3,25 \cdot 10^{-5}$	3,938	45	$2,22 \cdot 10^{-6}$	0,269
17	$2,44 \cdot 10^{-5}$	2,956	46	$1,48 \cdot 10^{-6}$	0,179
18	$2,23 \cdot 10^{-5}$	2,700	47	10^{-6}	0,121
19	$2,19 \cdot 10^{-5}$	2,650	48	$7,39 \cdot 10^{-7}$	0,089
20	$2,15 \cdot 10^{-5}$	2,600	49	$5,00 \cdot 10^{-7}$	0,060
21	$2,10 \cdot 10^{-5}$	2,550	50	$2,70 \cdot 10^{-7}$	0,032
22	$2,06 \cdot 10^{-5}$	2,500	51	$1,34 \cdot 10^{-7}$	0,016
23	$2,02 \cdot 10^{-5}$	2,450	52	10^{-7}	0,012
24	$1,98 \cdot 10^{-5}$	2,400	53	$6,83 \cdot 10^{-8}$	0,008
25	$1,94 \cdot 10^{-5}$	2,350	54	$5,00 \cdot 10^{-8}$	0,006
26	$1,90 \cdot 10^{-5}$	2,300	55	$2,60 \cdot 10^{-8}$	0,003
27	$1,82 \cdot 10^{-5}$	2,200	56	$1,95 \cdot 10^{-8}$	0,002
28	$1,73 \cdot 10^{-5}$	2,100	57	10^{-8}	0,001
29	$1,65 \cdot 10^{-5}$	2,000	58	0	0

Tabel 4.2. Tabel cu **valori intermediare** astfel încât să existe o distribuție mai uniformă, noile valori sunt introduse pentru a acoperi zona critică (0,5-3mg/mL).

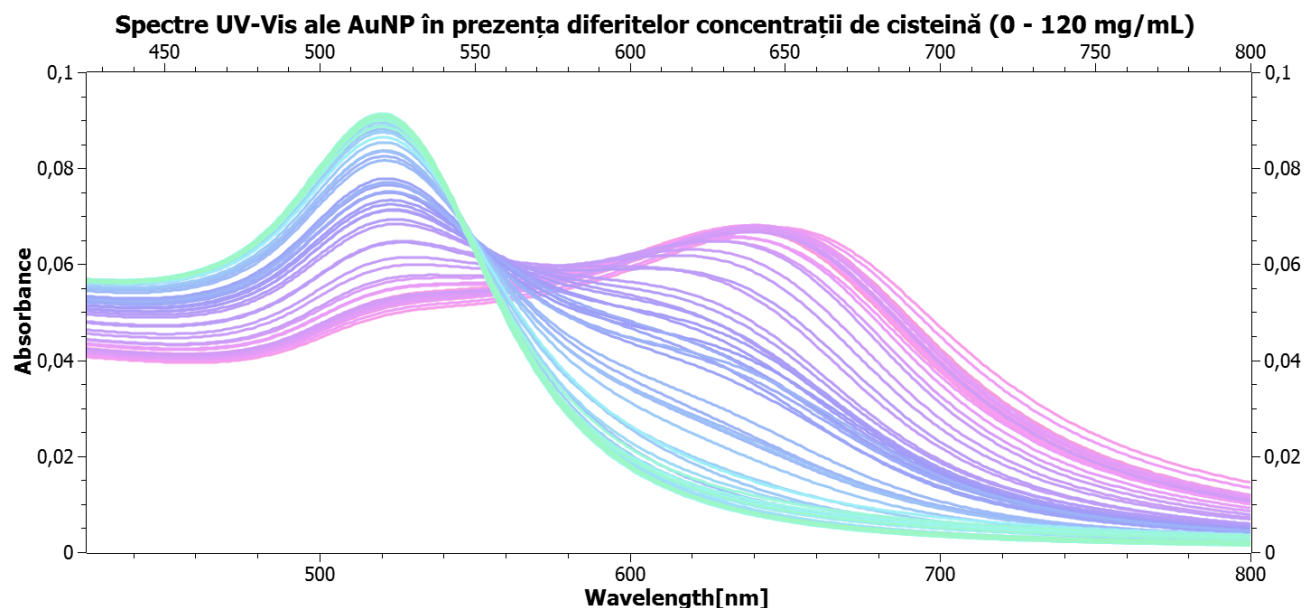


Figura 4.5. Spectrele optice ale soluțiilor de AuNP în prezența diferitelor concentrații de l-cisteină, așa cum sunt descrise în Tabel 4.2. Se observă o bună continuitate spectrală, care reflectă tranziția de la agregarea completă la absența agregării, evidențiind sensibilitatea semnalului plasmonic la modificările în concentrația cisteinei.

Setul extins de date a fost utilizat ulterior pentru antrenarea și testarea modelelor de învățare automată, evaluând capacitatea fiecărui model de a prezice concentrația de cisteină. Rezultatele pentru metoda **Partial Least Squares (PLS)** indică o îmbunătățire progresivă a performanței pe măsură ce a fost crescut numărul de componente latente. Astfel, coeficientul de determinare (R^2) a crescut până la 0,989, iar eroarea pătratică medie (RMSE) a scăzut la 3,412, ceea ce demonstrează o corelație puternică între spectre și valorile concentrației.

Modelul **Random Forest** a prezentat o performanță ridicată, atingând un R^2 de aproximativ 0,940 și valori ale RMSE în jur de 8, ceea ce confirmă robustețea metodei chiar și în prezența datelor provenite din întregul interval de concentrații. În schimb, metodele bazate pe rețele neuronale nu au reușit să ofere rezultate satisfăcătoare în această etapă, indiferent de numărul de neuroni ascunși utilizați, R^2 rămânând în jurul valorii de 0,420, cu un RMSE de aproximativ 25. Modelul **SVM** a continuat să producă rezultate nesatisfăcătoare, cu un R^2 aproape nul.

PLS		
Nr. de componente	R ²	RMSE
1	-	-
2	0,414	25,361
3	0,417	25,296
4	0,428	25,045
5	0,451	24,550
6	0,458	24,384
7	0,465	24,227
8	0,535	22,584
9	0,548	22,282
10	0,603	20,883
11	0,622	20,379
12	0,667	19,111
13	0,718	17,582
14	0,746	16,682
15	0,827	13,796
16	0,870	11,935
17	0,903	10,304
18	0,930	8,794
19	0,956	6,916
20	0,968	5,879
21	0,977	5,068
22	0,984	4,209
23	0,989	3,412

Random Forest		
Nr. de copaci	R ²	RMSE
1	-	-
2	0,871	11,918
3	0,874	11,767
4	0,882	11,373
5	0,877	11,638
6	0,907	10,114
7	0,919	9,408
8	0,925	9,055
9	0,937	8,317
10	0,946	7,690
11	0,936	8,378
12	0,933	8,573
13	0,934	8,509
14	0,938	8,244
15	0,944	7,846
16	0,940	8,129
17	0,939	8,187
18	0,935	8,460
19	0,937	8,298
20	0,937	8,292

Tabel 4.3. Rezultatele obținute în urma modificării nr. de componente a PLS-ului și a numărului de arbori în Random Forest.

Analiza distribuției spectrelor în funcție de concentrație a evidențiat o discontinuitate semnificativă în zonele de concentrații foarte mari și foarte mici. Acest lucru a fost reflectat și în performanța slabă a unor modele, fapt care a condus la decizia de a exclude aceste extreme din analiza finală și de a păstra doar intervalul cuprins între 0,5 și 3,0 mg/mL. Această selecție s-a bazat pe observația conform căreia spectrele în acest interval sunt mai consistente, iar diferențele de absorbție sunt corelate mai clar cu concentrația.

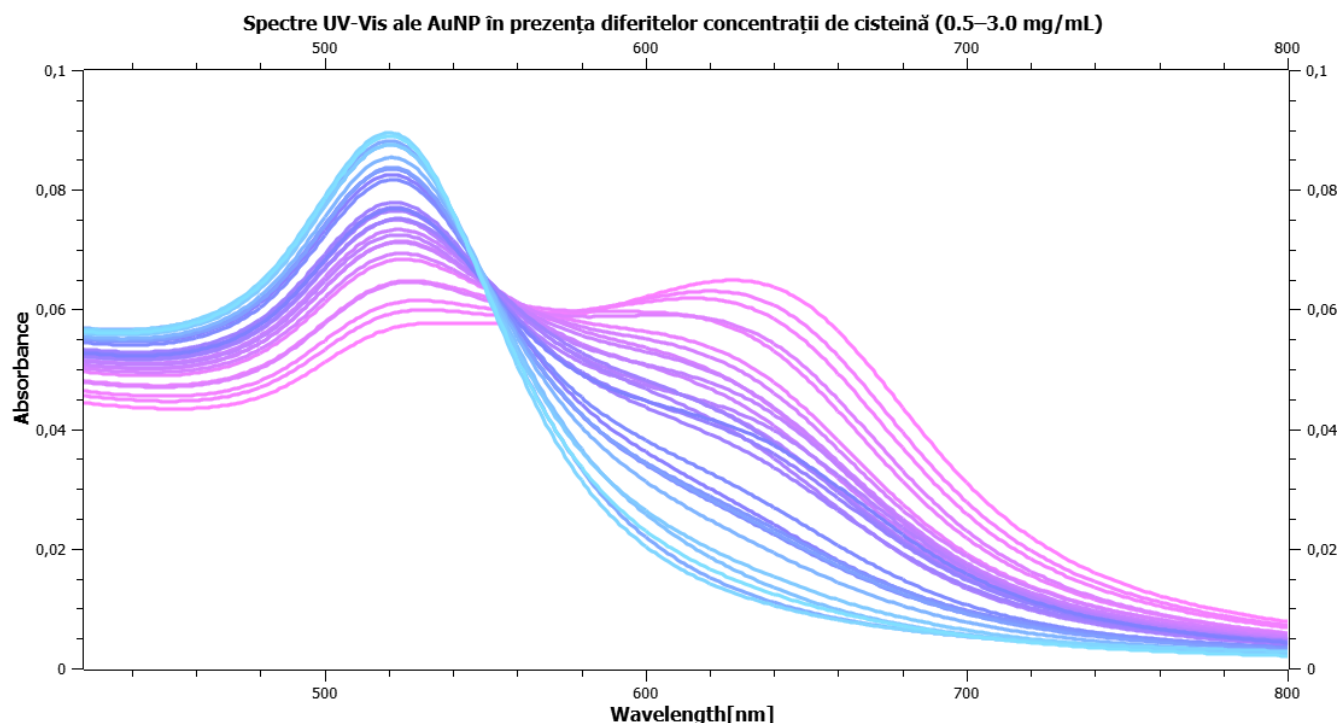


Figura 4.6. Spectrele soluțiilor de AuNP în prezența diferitelor concentrații de cisteină, asociate intervalului 0.5-3.0mg/mL.

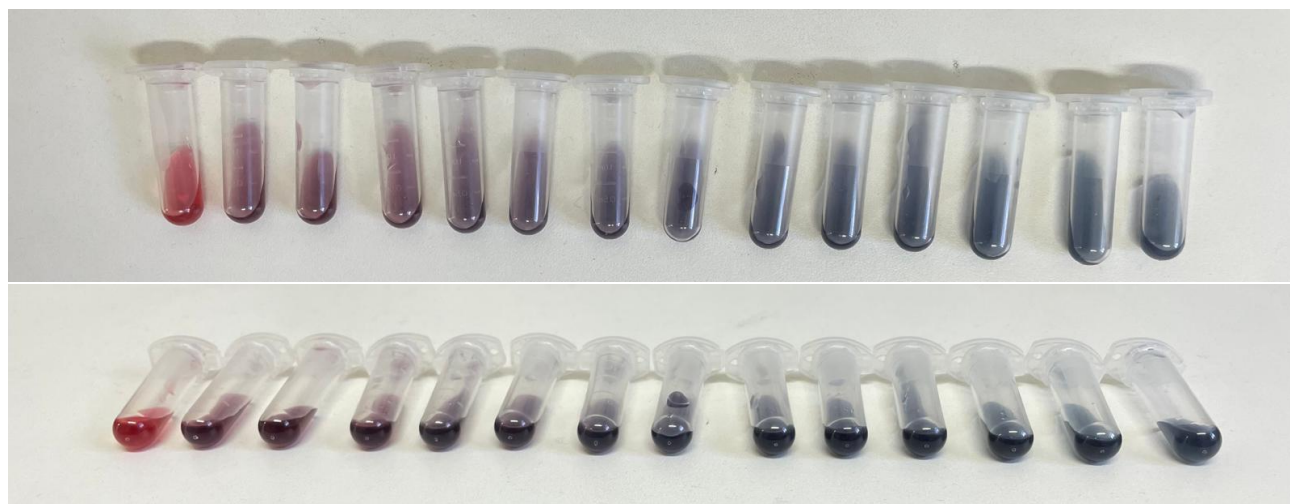


Figura 4.7. Aspectul vizual al soluțiilor cu concentrații între 0,5 și 3 mg/mL, agregarea nu este atât de agresivă ca și la cele de concentrații mai mari, ele nu reușesc să se depună pe fundul recipientului.

Utilizând cele 30 de probe selectate din intervalul 0,5–3,0 mg/mL, s-au reaplicat metodele de învățare automată pentru a evalua potențialul predictiv al modelelor în condiții optimizate. Modelul PLS a oferit cele mai bune rezultate, atingând un R^2 de 0,998 și un RMSE de doar 0,031, demonstrând

o capacitate excelentă de predicție. Modelul Random Forest a obținut performanțe similare, cu un R^2 cuprins între 0,970 și 0,980 și valori RMSE sub 0,136, confirmând adaptabilitatea metodei la seturi de date bine calibrate.

Rețelele neuronale, deși în general necesită un volum mai mare de date pentru a atinge performanțe optime, au reușit să ofere rezultate competitive în acest context. Valorile obținute pentru R^2 au variat între 0,769 și 0,960, iar RMSE între 0,123 și 0,324. Deși inferioare celor înregistrate de PLS și Random Forest, aceste rezultate sunt totuși acceptabile pentru un set relativ redus de probe. În contrast, SVM-ul a continuat să ofere performanțe nesatisfăcătoare, ajungând la un R^2 de 0,708.

PLS			Random Forest			Neural Networks		
Nr. de componente	R^2	RMSE	Nr. de copaci	R^2	RMSE	Nr. neuroni ascunși	R^2	RMSE
1	-	-	1	-	-	1	-	-
2	0,947	0,155	2	0,980	0,094	2	0,932	0,175
3	0,965	0,126	3	0,959	0,136	3	0,941	0,164
4	0,977	0,101	4	0,961	0,133	4	0,959	0,136
5	0,984	0,085	5	0,977	0,102	5	0,960	0,135
6	0,989	0,070	6	0,980	0,096	6	0,941	0,163
7	0,990	0,067	7	0,980	0,096	7	0,967	0,123
8	0,993	0,056	8	0,983	0,088	8	0,769	0,324
9	0,995	0,046	9	0,983	0,087	9	0,934	0,174
10	0,998	0,031	10	0,983	0,087	10	0,936	0,171
			11	0,981	0,094			

Tabel 4.4. Rezultatele obținute în intervalul 0,5-3,0mg/mL pentru modelele: PLS, Random Forest și Rețele Neuronale.

În urma studierii optimizărilor pentru fiecare model în parte, a fost realizată o etapă suplimentară de validare, menită să evalueze capacitatea definitivă de determinare a concentrației de

L-cisteină a modelelor. Spre deosebire de etapa anterioară, în care evaluarea performanței s-a realizat folosind întregul set de date ca și set de testare, în această fază, a fost rezervat exclusiv un subset de probe pentru testare, fiind exclus din procesul de antrenare. Această abordare permite o evaluare mai riguroasă și realistă a capacității de generalizare a fiecărui model.

Nr. probe pentru test	R^2			
	Rețele Neuronale	PLS	Random Forest	SVM
2	0,731	0,992	0,985	0,729
3	0,689	0,803	0,941	0,299
4	-	0,599	0,980	0,456
5	0,640	0,879	0,956	0,747
6	0,747	0,905	0,905	0,757
7	0,351	0,416	0,800	0,807
8	0,414	0,446	0,818	0,753
9	0,444	0,495	0,747	0,774
10	0,729	0,754	0,907	0,659
11	0,759	0,795	0,873	0,662
12	0,769	0,801	0,943	0,665
13	0,802	0,820	0,921	0,659
14	0,818	0,821	0,888	0,660
15	0,804	0,813	0,737	0,650
16	0,775	0,833	0,904	0,667
17	0,854	0,867	0,843	0,506
18	0,851	0,883	0,869	0,384
19	0,841	0,865	0,875	0,422
20	0,730	0,894	0,726	-
21	0,735	0,882	0,708	-
22	0,822	0,826	0,743	-

Tabel 4.5. Rezultatele obținute a coeficientului de determinare (R^2), în funcție de numărul de probe utilizate pentru testare.

Analiza performanței comparative a celor patru modele testate indică în mod clar superioritatea modelului Random Forest, care a obținut o valoare medie a coeficientului de determinare R^2 de **0,861** pe seturile de test. Acesta este urmat de modelul PLS, cu o medie de **0,780**, demonstrând de asemenea o capacitate bună de predicție în contextul datelor experimentale.

Rețelele neuronale, deși capabile să atingă performanțe competitive în etapa de validare internă, au prezentat o scădere a performanței în această etapă, cel mai probabil din cauza dimensiunii reduse a setului de antrenare. Cu o medie de **0,681**, acest model a demonstrat o sensibilitate crescută la variațiile din setul de test, fapt ce limitează aplicabilitatea sa fără un volum de date semnificativ mai mare.

Modelul SVM a continuat să ofere rezultate nesatisfăcătoare, cu o medie a R^2 de **0,536**, ceea ce sugerează o capacitate predictivă redusă în condițiile testate și o posibilă nepotrivire cu distribuția și complexitatea datelor utilizate.

4.2. Determinarea gradului de acoperire

Pentru a explica de ce determinarea cantitativă a cisteinei este validă doar într-un anumit interval de concentrații, presupunem că agregarea nanoparticulelor este determinată de gradul de acoperire al suprafeței AuNP cu molecule de cisteină.

Gradul de acoperire, notat cu θ , este definit ca raportul dintre numărul de molecule de cisteină care se leagă efectiv de suprafața unei nanoparticule de aur și numărul maxim posibil de molecule care pot fi adsorbite:

$$\theta = \frac{N_{cys}}{N_{max}} \quad \text{Ecuație 4.1}$$

Unde:

N_{cys} este numărul de molecule de cisteină legate de o AuNP,

N_{max} este numărul maxim de molecule de cisteină care se pot lega de o AuNP.

Comportamentul adsorbției este descris de izoterma Langmuir [56]:

$$N_{cys} = N_{max} \cdot \frac{K \cdot C_{cys}}{1 + K \cdot C_{cys}} \Rightarrow \theta = \frac{K \cdot C_{cys}}{1 + K \cdot C_{cys}} \quad \text{Ecuație 4.2}$$

Unde:

K este constanta de echilibru pentru adsorbția cisteinei,

C_{cys} concentrația de cisteină.

4.2.1. Concentrația maximă de acoperire

Atunci când gradul de acoperire $\theta = 1$, întreaga suprafață a nanoparticulei de aur (AuNP) este acoperită cu un **monostrat compact** de molecule de cisteină. Conform ecuației Langmuir (ecuația 4.1), această condiție corespunde:

$$1 = \frac{N_{cys}}{N_{max}} \Rightarrow N_{cys} = N_{max} \quad \text{Ecuație 4.3}$$

Pentru NP noastre de 13nm și concentrație de $C_{AuNP} = 3,05$ nM, putem determina numărul maxim de molecule de cisteină de pe suprafața unei nanoparticule N_{max} . Având o rază medie de 6,5nm, putem calcula aria A_{NP} , care este egală cu 531nm^2 .

Considerând că aria ocupată de cisteină $A_{cys,NP}$ variază între 0,2 și 0,3 nm^2 pe suprafața AuNP, atunci numărul maxim de molecule de cisteină pe suprafața AuNP este:

$$N_{max} = \frac{A_{NP}}{A_{cys,NP}} = 2124 \text{ molecule cys/AuNP} \quad \text{Ecuație 4.4}$$

Pentru volumele folosite în lucrarea noastră în fiecare recipient: $V_{coloid} = 0,9\text{mL}$ și $V_{cys} = 0,1\text{mL}$, calculăm numărul de NP în unitatea de volum:

$$C_{AuNP} \left(\frac{NP}{\text{mL}} \right) = C_{AuNP} \left(\frac{\text{mol}}{\text{L}} \right) \cdot N_A \cdot 10^{-3}$$

$$C_{AuNP} \left(\frac{NP}{\text{mL}} \right) = 1,84 \cdot 10^{12} \frac{NP}{\text{mL}} \quad \text{Ecuație 4.5}$$

Într-un volum de 0,9 mL de coloid se regăsesc:

$$N_{AuNP}(0,9\text{mL}) = C_{AuNP} \left(\frac{NP}{\text{mL}} \right) \cdot 0,9\text{mL} = 1,656 \cdot 10^{12} \text{ AuNP} \quad \text{Ecuație 4.6}$$

Rezultă că numărul total de molecule de cisteină necesar pentru acoperirea completă a tuturor nanoparticulelor din 0,9 mL de coloid este:

$$N_{max}(0,9mL) = N_{max} \cdot N_{AuNP}(0,9mL) = 3,517 \cdot 10^{15} AuNP \quad \text{Ecuație 4.7}$$

În plus, numărul de molecule de cisteină din 0,1mL poate fi scris ca:

$$N_{cys}(0,1mL) = C_{cys,max} \cdot 0,1mL \cdot N_A \quad \text{Ecuație 4.8}$$

Folosind ecuațiile 4.3, 4.7 și 4.8, obținem concentrația maximă de acoperire a suprafeței de AuNP cu cisteină $C_{cys,max}$:

$$C_{cys,max} = \frac{N_{max}(0,9mL)}{0,1mL \cdot N_A} = 5,84 \cdot 10^{-5} M \quad \text{Ecuație 4.9}$$

Pentru a obține o acoperire completă cu un monostrat de molecule de cisteină pe nanoparticulele de aur din 0,9 mL de coloid, este necesară adăugarea a 0,1 mL de soluție de cisteină cu o concentrație de $5,84 \times 10^{-5} M$, ceea ce corespunde unei concentrații finale în sistem de **0,700 mg/mL**.

4.2.2. Concentrația maximă de saturație

Gradul de acoperire al suprafeței nanoparticulelor de aur cu molecule de cisteină poate fi corelat direct cu deplasarea maximului benzii de absorbție plasmonice. Relația este exprimată prin ecuația:

$$\theta_{saturație} = \frac{\lambda - \lambda_0}{\lambda_{max,sat} - \lambda_0} \quad \text{Ecuație 4.10}$$

Unde:

λ este lungimea de undă maximă la o anumită concentrație,

$\lambda_{max,sat}$ este lungimea de undă la saturație (650 nm),

λ_0 este lungimea de undă fără cisteină (520 nm).

Această relație permite cuantificarea gradului de acoperire θ , pornind de la datele spectroscopice experimentale. Astfel, pentru fiecare spectru în care se observă o deplasare a plasmonului, a fost calculat θ și s-a asociat concentrația corespunzătoare de cisteină. Valorile obținute sunt prezentate în *Tabel 4.6*.

Lungimea de undă maximă la fiecare concentrație λ	Gradul de acoperire $\theta_{\text{saturație}}$	Concentrație cisteină (mg/mL)
637	0,914	121,160
636	0,906	121,160
642	0,953	121,160
637	0,914	110,000
637	0,914	100,000
637	0,914	90,000
638	0,921	80,000
638	0,921	60,580
648	1,000	60,580
639	0,929	50,000
637	0,914	40,000
641	0,945	30,000
635	0,898	20,000
645	0,976	12,116
642	0,953	10,000
640	0,937	8,000
629	0,851	6,058
641	0,945	6,058
639	0,929	4,000
638	0,921	3,938
639	0,929	2,956
620	0,781	2,700
627	0,836	2,700
615	0,742	2,650
527	0,055	2,600
525	0,039	2,550
524	0,031	2,500
526	0,046	2,500
524	0,031	2,450
524	0,031	2,400
523	0,023	2,350
523	0,023	2,300
522	0,015	2,200
522	0,015	2,100
522	0,015	2,000
522	0,015	2,000
523	0,023	1,975
522	0,015	1,900
521	0,008	1,800
522	0,016	1,700
521	0,008	1,600
521	0,008	1,500
521	0,008	1,400
521	0,008	1,300

Tabel 4.6. Valorile experimentale ale λ , gradul de acoperire la saturație $\theta_{\text{saturație}}$ și concentrația de cisteină. Se observă o creștere semnificativă pentru θ , începând cu concentrații de aproximativ 2,65 mg/mL, ceea ce sugerează că particulele nu mai sunt stabilizate individual, ci formează agregate.

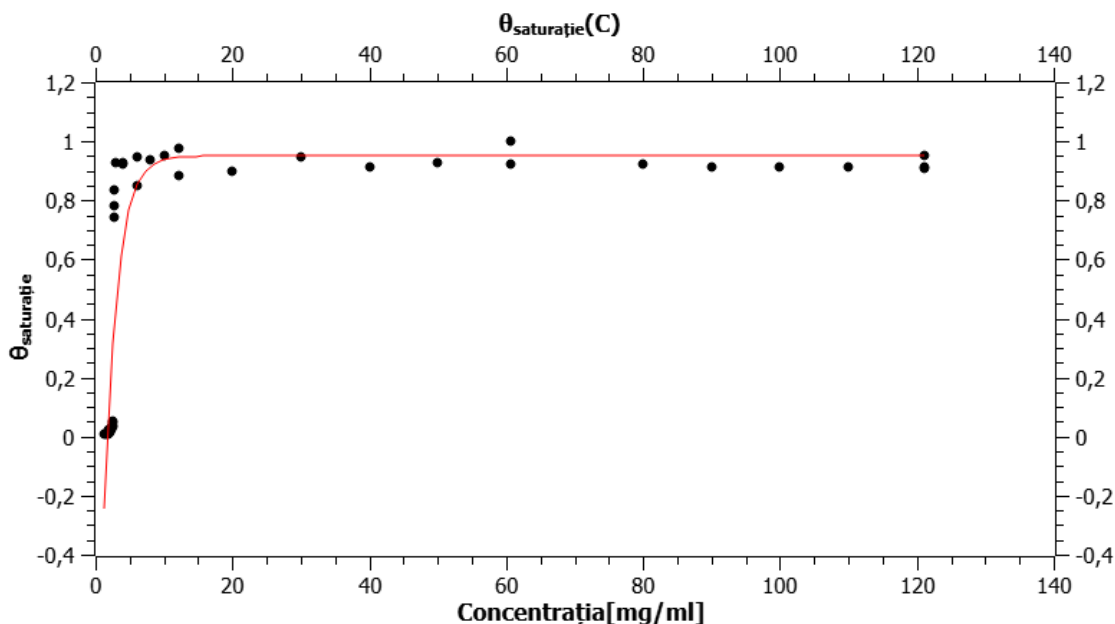


Figura 4.8. Reprezentarea ecuației lui Langmuir.

Tranziția evidențiată în spectrele plasmonice reflectă un comportament adsorbțiv de tip Langmuir, specific suprafețelor neporoase. Figura 4.8 prezintă o izotermă de adsorbție de tip I [56], caracteristică materialelor compacte, fără porozitate, așa cum sunt AuNP sintetizate prin metoda Turkevich. Aceste NP prezintă o suprafață continuă, acoperită cu citrat, fără defecte sau goluri structurale care să influențeze mecanismul de adsorbție.

Analiza spectrelor UV-Vis a evidențiat o creștere abruptă a gradului de acoperire $\theta_{\text{saturație}}$, începând de la concentrații de aproximativ **2,65 mg/mL**, ceea ce indică atingerea pragului de saturație al suprafeței și inițierea fenomenului de **agregare inter-particulară**. În această regiune, NP nu mai sunt complet stabilizate individual, ceea ce conduce la formarea de agregate vizibile spectroscopic prin deplasarea plasmonului.

Pe baza analizei spectroscopice și a calculelor teoretice, s-a stabilit că concentrația maximă de adsorbție, corespunzătoare acoperirii complete a suprafeței cu un monostrat de molecule de cisteină, este de 0,700 mg/mL, în timp ce concentrația de saturație experimentală, la care se observă inițierea agregării plasmonice, este de 2,650 mg/mL. Acest interval, delimitat de pragurile fizico-

chimice de adsorbție completă și de apariția agregării, se suprapune aproape perfect peste domeniul [0,5 – 3,0 mg/mL], identificat anterior ca interval optim pentru modelare și predicție.

Prin urmare, acest interval a fost adoptat ca zonă optimă pentru analiza plasmonică și modelarea procesului de adsorbție. La concentrații mai mari de 3,0 mg/mL se observă **agregare masivă**, ce distorsionează spectrul și reduce fiabilitatea analizei, iar la concentrații sub 0,5 mg/mL **semnalul plasmonic devine prea slab** pentru a permite o evaluare precisă a gradului de acoperire.

Concluzii

Acoperirea eficientă a intervalului de interes: A fost realizată o investigație experimentală extinsă, incluzând 69 de probe corespunzătoare la 58 de concentrații distincte, cu o densitate mai mare în zona critică 0,5–3,0 mg/mL, unde au fost observate cele mai relevante variații plasmonice. Această zonă s-a dovedit a fi optimă pentru analiza comportamentului de agregare a AuNP.

Evidențierea tranziției optice și comportamentului de agregare: Spectrele obținute au arătat o tranziție clară de la agregare completă (la concentrații peste 3mg/mL) la absența agregării (la concentrații sub 0,5mg/mL), cu o continuitate spectrală bună în intervalul 0,5–3,0 mg/mL. Aceasta reflectă o sensibilitate crescută a semnalului plasmonic în această zonă, în concordanță cu modificările adsorbției cisteinei la nivelul suprafeței nanoparticulelor.

Formarea monostratului de cisteină și efectele excesului: La concentrații moderate, moleculele de cisteină adsorb specific pe suprafața AuNP prin gruparea tiol (–SH), formând un **monostrat auto-asamblat stabil (SAM)**. După atingerea saturației suprafeței, moleculele de cisteină aflate în exces nu mai pot interacționa covalent cu aurul și rămân în soluție sau formează **straturi secundare instabile** prin interacțiuni nespecifice. Aceste interacțiuni slabe pot influența stabilitatea coloidală și pot contribui la agregare, în special în regimuri de concentrație mare, unde punțile slabe între nanoparticule devin posibile.

Determinarea intervalului valid pentru predicție: Datorită discontinuităților observate în spectrele provenite de la concentrații extreme, doar intervalul 0,5–3,0 mg/mL a fost reținut pentru analiza finală. În acest interval, spectrele sunt mai coerente și corelate cu variațiile concentrației.

Performanța modelelor de învățare automată (set complet vs. set restrâns):

1. Pe întregul set de date (0-120 mg/mL):

- ✓ *PLS*: $R^2 = 0,989$, $RMSE = 3,412$
- ✓ *Random Forest*: $R^2 = 0,940$, $RMSE = 8$
- ✓ *Rețele neuronale*: $R^2 = 0,420$, $RMSE = 3,412$
- ✓ *SVM*: $R^2 = 0$ - performanță nesatisfăcătoare

2. Pe intervalul optimizat (0,5–3,0 mg/mL):

- ✓ *PLS*: $R^2 = 0,998$, $RMSE = 0,031$ - cea mai bună performanță
- ✓ *Random Forest*: R^2 între 0,970–0,980, $RMSE$ sub 0,136 – performanță foarte bună

- ✓ *Rețele neuronale*: R^2 între 0,769–0,960, RMSE între 0,123–0,324 – performanță moderată
- ✓ *SVM*: $R^2 = 0,708$ – rezultat scăzut, dar mai bun decât în analiza globală

Evaluarea capacității de generalizare (validare pe subseturi separate):

- ☺ Random Forest a obținut cea mai bună performanță generală, cu o medie a R^2 de 0,861, demonstrând o robustețe ridicată și o capacitate bună de generalizare.
- ☺ PLS a avut o medie a R^2 de 0,780, performanță bună dar inferioară față de Random Forest.
- ☹ Rețelele neuronale au avut o medie a R^2 de 0,681, cu variații mari, indicând instabilitate în lipsa unui volum mare de date.
- ☹ SVM a continuat să se dovedească inefficient (medie $R^2 = 0,536$), fiind nepotrivit pentru acest tip de problemă.

Justificarea comportamentului sistemului prin modelul de adsorbție Langmuir:

Agregarea AuNP este determinată de gradul de acoperire θ al suprafeței NP cu molecule de cisteină. Acest grad urmează o **izotermă Langmuir**, unde concentrațiile mari conduc la saturația suprafeței, iar concentrațiile mici la o acoperire insuficientă ($\theta \ll 1$). Intervalul 0,5–3,0 mg/mL corespunde zonei în care variațiile θ sunt cele mai semnificative, ceea ce explică sensibilitatea crescută a semnalului plasmonic în acest interval. În plus, **formarea unui monostrat complet și stabil** asigură o relație predictibilă între concentrație și semnal, în timp ce **excesul de cisteină contribuie la variabilitate și agregare**.

Intervalul 0,5–3,0 mg/mL reprezintă zona optimă pentru determinarea cantitativă a cisteinei prin analiză plasmonică a AuNP. În acest interval, comportamentul sistemului este controlat de formarea monostratului de cisteină pe AuNP și este bine descris de modelul de adsorbție Langmuir. Modelele de învățare automată, în special PLS și Random Forest, oferă performanțe predictive excelente datorită coerenței spectrelor și a stabilității fizico-chimice. În afara acestui interval, interferențele datorate excesului de cisteină și formării de structuri nespecifice reduc semnificativ capacitatea de predicție.

Bibliografie

- [1] Kimling J, Maier M, Okenve B, Kotaidis V, Ballot H and Plech A 2006 Turkevich Method for Gold Nanoparticle Synthesis Revisited *J. Phys. Chem. B* **110** 15700–7
- [2] Daniel M and Astruc D 2004 Gold Nanoparticles: Assembly, Supramolecular Chemistry, Quantum-Size-Related Properties, and Applications Toward Biology, Catalysis, and Nanotechnology. *ChemInform* **35** chin.200416213
- [3] Jain P K, Huang X, El-Sayed I H and El-Sayed M A 2007 Review of Some Interesting Surface Plasmon Resonance-enhanced Properties of Noble Metal Nanoparticles and Their Applications to Biosystems *Plasmonics* **2** 107–18
- [4] Cao Y C, Jin R and Mirkin C A 2002 Nanoparticles with Raman Spectroscopic Fingerprints for DNA and RNA Detection *Science* **297** 1536–40
- [5] Saha K, Agasti S S, Kim C, Li X and Rotello V M 2012 Gold Nanoparticles in Chemical and Biological Sensing *Chem. Rev.* **112** 2739–79
- [6] Li B, Li X, Dong Y, Wang B, Li D, Shi Y and Wu Y 2017 Colorimetric Sensor Array Based on Gold Nanoparticles with Diverse Surface Charges for Microorganisms Identification *Anal. Chem.* **89** 10639–43
- [7] Anon 2007 Pattern Recognition and Machine Learning *J. Electron. Imaging* **16** 049901
- [8] Bohren, C. F. and Huffman, D. R., Absorption and Scattering of Light by Small Particles, Wiley, 1983.
- [9] Nanonetworks Nanoparticles Illustrations & Vectors. [Nanoparticles Labeled Infographic. Microscopic Element Vector Illustration Stock Vector - Illustration of liposome, element: 125797974](#). Dreamstime. Accessed: 4 March 2025
- [10] Jain, P. K. et al. “Calculated Absorption and Scattering Properties of Gold Nanoparticles of Different Size, Shape, and Composition: Applications in Biological Imaging and Biomedicine,” *J. Phys. Chem. B*, 2006.
- [11] Maier, S. A. Plasmonics: Fundamentals and Applications, Springer, 2007.

- [12] Link S. et al. “Localized Surface Plasmon Resonance Spectroscopy of Nanoparticles for Biomedical Applications,” *J. Phys. Chem. C*, 2006.
- [13] J. Homola, *Surface Plasmon Resonance Based Sensors*, Springer, 2006.
- [14] Podila R, Chen R, Ke P C, Brown J M and Rao A M 2012 Effects of surface functional groups on the formation of nanoparticle-protein corona *Applied Physics Letters* **101** 263701
- [15] Kister T, Monego D, Mulvaney P, Widmer-Cooper A and Kraus T 2018 On the colloidal stability of apolar nanoparticles: The role of particle size and ligand shell structure *ACS Nano* **12** 5969–77
- [16] Phan A D, Hoang T X, Nghiem T H L and Woods L M 2013 Surface plasmon resonances of protein-conjugated gold nanoparticles on graphitic substrates *Applied Physics Letters* **103** 163702
- [17] Berciaud S, Cognet L, Tamarat P and Lounis B 2005 Observation of intrinsic size effects in the optical response of individual gold nanoparticles *Nano Lett.* **5** 515–8
- [18] Murphy C J, Sau T K, Gole A M, Orendorff C J, Gao J, Gou L, Hunyadi S E and Li T 2005 Anisotropic Metal Nanoparticles: Synthesis, Assembly, and Optical Applications *J. Phys. Chem. B* **109** 13857–70
- [19] Pissuwan D, Valenzuela S M and Cortie M B 2006 Therapeutic possibilities of plasmonically heated gold nanoparticles *Trends in Biotechnology* **24** 62–7
- [20] D. D. Cunningham “Optical Biosensors: Fundamentals and Applications,” *Biosens. Bioelectron.*, vol. 26, pp. 123–132, 2010.
- [21] A. Smith et al. “Advances in Optical Biosensing for Environmental Monitoring,” *Anal. Chem.*, vol. 85, pp. 592–600, 2013.
- [22] B. Chen and K. A. Wilkins “Gold Nanoparticles in Optical Biosensors,” *J. Phys. Chem. C*, vol. 114, pp. 7071–7078, 2010.
- [23] M. J. Liz-Marzán “Tailoring Surface Plasmon Resonances of Metal Nanoparticles,” *Langmuir*, vol. 22, pp. 32–41, 2006.

- [24] S. A. Mirkin et al. “A DNA-Based Method for Rationally Assembling Nanoparticles into Macroscopic Materials,” *Nature*, vol. 382, pp. 607–609, 1996.
- [25] P. K. Jain et al. “Calculated Absorption and Scattering Properties of Gold Nanoparticles of Different Size, Shape, and Composition: Applications in Biological Imaging and Biomedicine,” *J. Phys. Chem. B*, vol. 110, pp. 7238–7248, 2006.
- [26] R. R. Naik et al. “Application of Neural Networks in the Design and Optimization of Biosensors,” *Biosens. Bioelectron.*, vol. 45, pp. 55–64, 2013.
- [27] Wikipedia contributors Resonancia de plasmones superficiales. [Resonancia de plasmones superficiales - Wikipedia, la enciclopedia libre](#). Wikipedia. Accessed 4 March 2025
- [28] Academia Lab (2025). Espectroscopia ultravioleta-visible. <https://academia-lab.com/enciclopedia/espectroscopia-ultravioleta-visible/>. Enciclopedia. Accessed: 11 February 2025
- [29] Stabilizzare la temperatura negli spettrometri [Stabilizzare la temperatura negli spettrometri - Caronte Consulting](#). Caronte Consulting. Accessed: 4 March 2025
- [30] Tony Owen “Fundamentos de la espectroscopía UV-visible moderna” Academia.edu ([55](#)) [Fundamentos de la espectroscopía UV-visible moderna](#). Accessed: 4 March 2025
- [31] Beer-Lambert Law [Beer Lambert Law - GeeksforGeeks](#). GeeksforGeeks. Accessed: 4 March 2025
- [32] Spectrophotometers <https://psiberg.com/uv-vis-spectroscopy/> Psiberg. Accessed: 4 March 2025
- [33] Guía 7 FII ByG Redes 2016. [Microsoft Word - Guia_7_FII_ByG_redes_2016.doc](#). Materia de Física 2 - FII UBA. Accessed: 4 March 2025
- [34] Alvarado A R V 2018 INTRODUCCIÓN AL MACHINE LEARNING
- [35] Machine Learning & Deep Learning — When Computers Learn Independently [Machine Learning & Deep Learning: Definition & More](#). Bots & People. Accessed: 4 March 2025

- [36] Forero-Corba W and Negre Bennasar F 2023 Técnicas y aplicaciones del Machine Learning e Inteligencia Artificial en educación: una revisión sistemática *RIED* **27** 209–53
- [37] Chen L, Liu G and Zhang T 2024 Integrating machine learning and genome editing for crop improvement *aBIOTECH* **5** 262–77
- [38] Alcobilla Ferrara E 2023 Introducción al machine learning en Senología *Revista de Senología y Patología Mamaria* **36** 100503
- [39] Rosipal R and Krämer N 2006 Overview and Recent Advances in Partial Least Squares *Subspace, Latent Structure and Feature Selection* Lecture Notes in Computer Science vol 3940, ed C Saunders, M Grobelnik, S Gunn and J Shawe-Taylor (Berlin, Heidelberg: Springer Berlin Heidelberg) pp 34–51
- [40] Martínez Ávila M and Fierro Moreno E 2018 Aplicación de la técnica PLS-SEM en la gestión del conocimiento: un enfoque técnico práctico / Application of the PLS-SEM technique in Knowledge Management: a practical technical approach *RIDE* **8** 130–64
- [41] Sanchez, G., Marzban, E. (2020) *All Models Are Wrong: Concepts of Statistical Learning*. [All Models Are Wrong: Concepts of Statistical Learning](#). Accessed: 10 March 2025
- [42] Norabuena Mendoza C H, Huamán Osorio A P and Ramirez Asis E H 2020 Modelo de Ecuaciones Estructurales (Con estimación PLS) basado en calidad de servicio y lealtad del Cliente de las Cajas Rurales Peruanas *CADM* 081
- [43] Cutler A, Cutler D R and Stevens J R 2012 Random Forests *Ensemble Machine Learning* ed C Zhang and Y Ma (New York, NY: Springer New York) pp 157–75
- [44] Couronné R, Probst P and Boulesteix A-L 2018 Random forest versus logistic regression: a large-scale benchmark experiment *BMC Bioinformatics* **19** 270
- [45] Random Forest Algorithm in Machine Learning [Random Forest Algorithm in Machine Learning - GeeksforGeeks](#). GeeksforGeeks. Accessed: 4 March 2025
- [46] Kecman V 2005 Support Vector Machines – An Introduction *Support Vector Machines: Theory and Applications* Studies in Fuzziness and Soft Computing vol 177, ed L Wang (Berlin, Heidelberg: Springer Berlin Heidelberg) pp 1–47

- [47] Jakkula V Tutorial on Support Vector Machine (SVM)
- [48] León E C Introducción a las máquinas de vector soporte (SVM) en aprendizaje supervisado
- [49] Classification Model - SVM Classifier Python Example [Support Vector Machine \(SVM\) Python Example - Analytics Yogi](#). Vitalflux. Accessed: 4 March 2025
- [50] Ruiz C A, Basualdo M S and Matich D J Redes Neuronales: Conceptos Básicos y Aplicaciones.
- [51] Ramos J M S and Montes A L GRADO EN INGENIERÍA MATEMÁTICA
- [52] Pose M G Introduccion a las Redes de Neuronas Artificiales
- [53] Vitor Maus A, Martins Razera T, Kuntz R, Linhares Junior V and Camillo Junior A 2021 Contagem e Classificação de Veículos por Visão Computacional *Anais do XII Computer on the Beach - COTB '21 Computer on the Beach* (Online: Universidade do Vale do Itajaí) pp 049–56
- [54] Medina-Merino R F and Ñique-Chacón C I 2017 Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python *Interfases* **0** 165
- [55] Varela G Nanosensores plasmónicos en la detección colorimétrica de algunas bases pirimidínicas y derivados
- [56] [Capitulo2 PDF | PDF | Adsorción | Carbón](#). Accessed: 17 June 2025

DECLARAȚIE PE PROPRIE RĂSPUNDERE

Subsemnata, TÜRKE ANDREEA, declar că Lucrarea de absolvire/Lucrarea de licență/Proiectul de diplomă/Lucrarea de disertație pe care o voi prezenta în cadrul examenului de finalizare a studiilor la Facultatea de FIZICĂ, din cadrul Universității Babeș-Bolyai, în sesiunea IULIE 2025, sub îndrumarea PROF. DR. LEOPOLD NICOLAE și LECT. DR. IANCU ȘTEFANIA, reprezintă o operă personală. Menționez că nu am plagiat o altă lucrare publicată, prezentată public sau un fișier postat pe Internet. Pentru realizarea lucrării am folosit exclusiv bibliografia prezentată și nu am ascuns nici o altă sursă bibliografică sau fișier electronic pe care să le fi folosit la redactarea lucrării.

Prezenta declarație este parte a lucrării și se anexează la aceasta.

Data,

18.06.2025

Nume,

Andreea Türke

Semnătură

