



Universitatea Națională de Știință și Tehnologie

POLITEHNICA București

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE

Splaiul Independenței 313, Sector 6, 060042- București

Web [www.acs.pub.ro](http://www.acs.pub.ro)

Tel +40 021 410 104, Fax +40 021 318 1014, Email [decanat@acs.pub.ro](mailto:decanat@acs.pub.ro)

## Programul de master

### Analiza avansată de date pentru afaceri

### Advanced Analytics for Business

#### Descrierea disciplinelor din planul de învățământ

#### Fundamente Data Science în Afaceri și Inginerie

#### Fundamentals of Data Science in Business and Engineering

Acest curs este o condiție prealabilă pentru majoritatea cursurilor de analiză statistică pentru analiza avansată de date. Cursul va oferi studenților o viziune integrată asupra metodelor statistice și a utilității acestora în diverse industrii. Urmând acest curs, studenții vor învăța cum să devină utilizatori de statistici descriptive și inferență statistică, regresie liniară, regresie logistică și să genereze statistici descriptive, să exploreze date cu ajutorul graficelor. Urmând acest curs, studenții vor obține următoarele cunoștințe: concepte de baza în statistică (statistica descriptivă și inferență statistică, populație statistică, eșantioane; variabile independente / dependente; tipuri de variabile: cantitative / categorice, scala de măsură, etc.); principalele noțiuni din statistica descriptivă (descrierea datelor, măsuri de localizare și variabilitate); reprezentarea statistică a datelor, distribuții de probabilitate; inferență statistică: intervale de încredere și elemente de matematică statistică utilizate în testarea ipotezelor; ANOVA test și regresie liniară simplă / multiplă; cum să aplice mai multe tehnici de comparație, să efectueze în mod practic regresii liniare și să evalueze ipotezele, să utilizeze tehnici de selecție a modelelor de regresie pentru a ajuta la alegerea variabilelor predictorului în regresie multiplă; să lucreze în mod real cu statistici de diagnosticare pentru a evalua ipotezele statistice și pentru a identifica potențialele valori aberante în regresia multiplă.

#### Fundamente de programare în procesarea datelor

#### Programming Essentials for Data Processing

Acest curs oferă posibilitatea studenților să învețe să scrie programe software pentru a accesa, explora, și pregăti analiza avansată de date. Cursul este structurat în două părți: elemente esențiale de programare în prelucrarea datelor; manipularea datelor și introducerea în limbaje de programare care susțin organizarea informațiilor structurate și nestructurate. A doua parte a cursului include cunoștințe despre: algoritmi și înțelegerea bazelor prelucrării datelor; evaluarea modului în care pot fi utilizate procesele avansate de analiză pentru a aborda problemele de afaceri; accesarea și crearea structurilor de date; gestionarea datelor, generarea de rapoarte și gestionarea erorilor; învățarea programării SAS pentru *data science* vs. programare în R / Python.

#### Tehnici de vizualizare și analiză complexă

#### Visual Analytics Techniques

Acest curs prezintă principalele noțiuni referitoare la pregătirea și explorarea vizuală a datelor, pentru a furniza o percepție asupra naturii datelor și pregătire a acestora pentru analiză avansată ulterioară. Tehnicile de vizualizare de date prezentate în curs sunt utilizate pentru extragerea unor perspective relevante asupra datelor (clustere, excepții, outliers, corelații între date), fără prelucrări complexe, în special pentru mari volume de date caracterizate prin volum (cantitate mare de date), varietate (date structurate, semistructurate, nestructurate) și viteză (viteza cu care sunt colectate datele și viteza cu



Universitatea Națională de Știință și Tehnologie  
POLITEHNICA București

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE

Splaiul Independenței 313, Sector 6, 060042- București

Web [www.acs.pub.ro](http://www.acs.pub.ro)

Tel +40 021 410 104, Fax +40 021 318 1014, Email [decanat@acs.pub.ro](mailto:decanat@acs.pub.ro)

care se modifică). Studenții vor învăța: ce înseamnă pregătirea datelor, analiză și crearea de rapoarte folosind software avansat de analiză; ce înseamnă restructurarea și structurarea datelor pentru analiză avansată (analiză geografică, prognoză, analiză de rețea, analiză cale, analiză text); să aplice proceduri complexe de structurare date pentru analiza avansată; ce înseamnă să „curețe” date (data cleansing); să înțeleagă caracteristicile avansate ale pregătirii, analizei și creării de rapoarte de date; să creeze rapoarte personalizate pentru a le utiliza în managementul deciziilor; să creeze rapoarte de date și filtre avansate și rapoarte interactive avansate cu parametri.

### Programare Python în business intelligence

#### Python programming for business intelligence

Acest curs este un curs *project-based*, care se bazează pe crearea de aplicații practice, urmărind dezvoltarea abilităților de programare ale studenților în limbajul de programare Python și programarea duală Base SAS / Python. Cursul se concentrează pe: instalarea și configurarea software-ului într-un mediu de programare; descrierea conceptelor generice de limbaj de programare; programare Python și integrare în Base SAS; utilizarea funcțiilor Python în Base SAS (programare SAS native); prezentarea generală a aplicațiilor specifice din domeniul machine learning și exemple de aplicații de învățare automată (clasificare, regresie, supervised / unsupervised learning); învățarea limbajului Python prin utilizarea de algoritmi specifici domeniului machine learning; introducerea componentelor de bază ale construirii și aplicării unei game de metode de învățare automată bazate pe modele și algoritmice (regresie, arbori de clasificare, Bayes Naive, random forest etc.), cu accent pe aplicații practice (seturi, teste, rate de eroare, etc.); metode inteligente de evaluare a performanței; descărcarea rezultatelor și folosirea sintaxei Python nativă pentru a compara modelele.

### Programare R în business intelligence

#### R programming for business intelligence

Acest curs este un curs *project-based*, care se bazează pe crearea de aplicații practice, urmărind dezvoltarea abilităților de programare ale studenților în limbajul de programare statistică R. Cursul se concentrează pe: introducerea limbajului de programare R; utilizarea R API pentru a prelua controlul acțiunilor Cloud Analytic Services (CAS) de la Jupyter Notebook; încărcarea datelor în mediul distribuit în memorie (CASLib), analiza datelor, crearea modelelor predictive în CAS folosind funcționalitatea R prin intermediul pachetului SWAT (SAS Wrapper for Analytics Transfer); învățarea modului de descărcare a rezultatelor clientului și folosirea sintaxei R native pentru a compara modelele; prezentarea generală a aplicațiilor specifice din domeniul machine learning, cu accent pe exemple de aplicații de învățare automată (clasificare, regresie, supervised / unsupervised learning).

### Introducere în analiza exploratorie de date

#### Introduction to Exploratory Data Analysis

Acest curs prezintă principalii pași de preprocesare a datelor, imediat după colectarea acestora din diverse surse, prin simpla vizualizare, reprezentare grafică, fără ipoteze, pentru a ajuta la evaluarea a priori a calității datelor. Studenții vor învăța cum să utilizeze tehnicile EDA pentru a înțelege, rezuma și investiga datele; examineze modul în care sunt distribuite datele colectate, să determine outliers și anomalii în prezentarea datelor, pentru a informa în mod corect formarea și explorarea ulterioară de ipoteze; înțeleagă cât mai bine cum se analizează structura unei baze de date pentru analiză de date; descopere pattern-uri în date; sintetizeze și să vizualizeze caracteristicile importante ale unui set de date; înțeleagă structura și variabilele adiacente datelor, cum să extragă și să creeze variabile relevante pentru crearea de modele de predicție și analiză avansată; își dezvolte intuiția despre setul de date, să ia în considerare modul în care a apărut acel set de date și să decidă cum poate fi investigat prin metode statistice mai formale; să efectueze



Universitatea Națională de Știință și Tehnologie  
POLITEHNICA București

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE

Splaiul Independenței 313, Sector 6, 060042- București

Web [www.acs.pub.ro](http://www.acs.pub.ro)

Tel +40 021 410 104, Fax +40 021 318 1014, Email [decanat@acs.pub.ro](mailto:decanat@acs.pub.ro)

o analiză vizuală a datelor pentru o mai bună înțelegere a problemelor de calitate (*data quality*); prelucreze, să aplice proceduri de "curățare" date (*data cleansing*).

### Modelare predictivă

#### Predictive modeling

Acest curs prezintă principalele noțiuni referitoare la domeniul analizei predictive, ca metodă statistică ce utilizează algoritmi specifici de tip învățare automată și data mining, pentru a prezice rezultate viitoare pe baza datelor istorice și a celor actuale. Studenții vor învăța: care sunt pașii esențiali din procesul de modelare predictivă, colectarea de date, formularea modelului statistic, evaluarea modelului, analiza rezultatelor, rafinarea modelului de predicție pe măsură ce sunt colectate noi date; înțelegerea specificului analizei predictive în afaceri și a principalelor provocări; să aplice modele predictive în diverse studii de caz; ce înseamnă analiza datelor istorice și actuale și generarea modelelor de predicție; principalele modele utilizate în modelare predictivă; ce înseamnă analiză predictivă; evaluarea modelelor predictive; evaluarea factorilor care pot afecta precizia rezultatelor (eficiența predictorului).

### Fundamente rețele neuronale

#### Neural Networks Essentials

Cursul combină teoria și practica pentru a aprofunda conceptele de bază ale modelelor de rețea neuronală și practicile esențiale pentru utilizarea acestora în aplicații din lumea reală. Cursul abordează următoarele aspecte: principalele paradigme de construcție a modelelor ANN (învățare supervizată / nesupervizată / reinforcement learning); alegerea parametrilor esențiali ai modelului în diverse studii de caz; principalele modele de rețele neuronale utilizate în aplicații practice (inclusiv perceptroni multistrat și rețele RBF); cum se alege o arhitectură de rețea neuronală pentru o anumită aplicație și metoda de antrenare corespunzătoare; cum se evită fenomenul de overfitting; introducerea deep learning; interpretarea predicțiilor realizate cu modele ANN; crearea modelelor și evaluarea ANN în mediul de lucru, implementarea modelelor ANN.

### Analiză vizuală avansată de text

#### Advanced visual text analytics

Acest curs prezintă tehnici specifice de descoperire a informațiilor, a pattern-urilor ascunse în seturi de date nestructurate, folosind puterea combinată a procesării limbajului natural, a învățării automate și a regulilor lingvistice. Sunt prezentate cele cinci componente caracteristice analizei vizuale de text: parsing, derivarea conceptului, derivarea topicii, clasificarea textului și analiza sentimentelor. Este prezentat modul de analiză a documentelor (date nestructurate) pentru a determina temele dominante din colecții de documente. Este prezentat modul de construcție a interogărilor lingvistice complexe pentru extragerea și structurarea informațiilor. Urmând acest curs, studenții vor învăța să: exploreze colecții de documente de tip text pentru a descoperi topicile importante; identifice în mod automat topicile textuale cheie din colecții mari de documente; creeze modele robuste de clasificare a conținutului în funcție de cerințele specifice ale organizației; creeze reguli de tip Boolean pentru clasificarea documentelor, în funcție de variabilele țintă categoricale; extragă scorul sentimentului la nivel de document; utilizeze practic cunoștințele de procesare automată de texte în aplicații de afaceri, e.g. chat-bots; utilizeze aceste abilități pe scară largă ca parte a analizei de marketing, a cercetării avansate de piață.

### Tehnici vizuale avansate de data mining

#### Advanced visual data mining techniques

Cursul utilizează o abordare interactivă pentru a prezenta vizualizarea, evaluarea și implementarea modelelor bazate pe date. Studenții vor aborda din punct de vedere practic diverse tehnici de învățare



Universitatea Națională de Știință și Tehnologie  
POLITEHNICA București

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE

Splaiul Independenței 313, Sector 6, 060042- București

Web [www.acs.pub.ro](http://www.acs.pub.ro)

Tel +40 021 410 104, Fax +40 021 318 1014, Email [decanat@acs.pub.ro](mailto:decanat@acs.pub.ro)

automată și studii de caz. Cursul prezintă o perspectivă vizuală complexă asupra principalelor principii de exploatare a datelor (*data mining*). Principalele componente ale cursului vizează: prezentarea, antrenarea și evaluarea performanțelor diverselor arhitecturi: rețele Bayesiene, modele de tip forest, modele predictive obținute prin metoda gradient boosting, rețele neuronale artificiale, rețele SVM (Support Vector Machine), algoritmi cu învățare supervizată de tip general (factorization machines); compararea diverselor modele obținute; transferarea modelelor analitice în *data pipelines*.

### Metode de colectare și analiza statistică a datelor Research Design

Acest curs prezintă principalele metode analiză calitativă, cantitativă și combinată, care implică colectarea primară și prelucrarea de date. Aceste metode vor fi folosite de către studenți pentru a putea formula ipoteze (întrebări) în proiecte de cercetare, incluzând analiza datelor reale colectate prin intermediul chestionarelor și testarea ipotezelor în procese de luare de decizii. Studenții vor învăța să: înțeleagă ce înseamnă analiza cantitativă de date și analiza calitativă de date în proiecte de business și de cercetare; înțeleagă care sunt principalele metode de colectare date (interviuri, chestionare, etc.), avantajele și dezavantajele utilizării acestora în diverse abordări (i.e. action research, studii de caz, surveys, etc.); creeze chestionare (surveys) complexe în mediu de afaceri și să analizeze aceste date colectate din studii de caz reale; înțeleagă ce înseamnă analiza statistică pentru date calitative și alegerea testelor statistice; cum se definește un model de cercetare (format din variabile și ipoteze); aplicarea scalei Likert pentru măsurarea variabilelor asociate modelului de cercetare.

### Fundamentele ingineriei programelor Software Engineering Essentials

Această disciplină are ca obiective generale însușirea cunoștințelor, principiilor și metodelor fundamentale ale "Ingineriei Programelor", înțelegerea complexității și provocărilor dezvoltării de software la scară industrială, precum și familiarizarea studenților cu lucrul în echipă în proiecte software. Înțelegerea specificului ingineriei software, comparativ cu alte discipline ingineresti și tehnice, este esențială în prezent, deoarece, chiar și dincolo de industria ICT, în numeroase ramuri industriale, aproape toate proiectele au și o componentă software, importanța acesteia fiind cu atât mai mare în proiectele axate pe date, pe colectarea, analiza și utilizarea lor. Vor fi studiate următoarele aspecte: activitățile tehnice uzuale din dezvoltarea de software - pentru a înțelege specificului și provocările fiecăreia; activitățile conexe de management și asigurare a calității în dezvoltarea de software; diversele modele uzuale ale ciclului de viață software (SDLC), pentru a înțelege specificul, avantajele, limitările și provocările fiecăreia, și pentru a dezvolta capacitate de alegere a modelului optim per proiect, sau integrare optimă în echipe ce folosesc anumite modele; principii pragmatice de modelare software și limbajul UML; numeroase alte subiecte punctuale, conexe, prin intermediul unor prezentări realizate de studenți.

### Prognoză și sisteme de suport a deciziilor în afaceri Forecasting and Decision Support Systems

Acest curs prezintă principalele concepte referitoare la modelele de prognoză care utilizează algoritmi complecși și la modul în care aceste prognoze pot fi folosite pentru susținerea deciziilor în mediul de afaceri. Cursul descrie principalii pași pentru prognoză și utilizarea acestora în procese de decizie, incluzând pregătirea și încărcarea datelor în memorie, modelarea și vizualizarea seriilor de timp, definirea variabilelor de tip atribut pentru realizarea vizualizărilor complexe. De asemenea, cursul prezintă principalele noțiuni referitoare la utilizarea *data pipelines* pentru generarea prognozelor și selectarea celui mai bun *pipeline* într-un proiect. Acest curs va oferi studenților posibilitatea să: folosească analiza datelor în scopuri practice; înțeleagă modul de utilizare a modelelor complexe pentru realizarea prognozelor;



Universitatea Națională de Știință și Tehnologie

POLITEHNICA București

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE

Splaiul Independenței 313, Sector 6, 060042- București

Web [www.acs.pub.ro](http://www.acs.pub.ro)

Tel +40 021 410 104, Fax +40 021 318 1014, Email [decanat@acs.pub.ro](mailto:decanat@acs.pub.ro)

Învețe modul în care sistemele de suport decizional (DSS), adică sistemele interactive de luare de decizii bazate pe software, pot asista viitorii manageri de proiecte în luarea deciziilor prin accesarea unor volume mari de informații generate de diverse sisteme informatice conexe, implicate în procesele de afaceri organizaționale; să includă practici de prognoză la scară largă în proiectul de prognoză.

### Concepte de optimizare în Data Science

#### Optimization Concepts for Data Science

Acest curs se concentrează pe descrierea aplicabilității practice a diverselor concepte de optimizare (liniară, neliniară și liniară mixtă), va prezenta și va exemplifica diverse studii de caz. Cursul abordează: o scurtă introducere a tehnicilor de optimizare; matematica de optimizare; o scurtă introducere metode de optimizare: linear / nonlinear / integer and mixed integer linear programming; identificarea și formularea abordării celei mai utile pentru rezolvarea diverselor probleme de optimizare; eficientizarea formulărilor; demonstrații cu exemple de formulare și optimizare a portofoliului; proceduri specifice și instrumentele open source pentru a formula și rezolva problemele de optimizare; modele de optimizare utilizate în mod obișnuit în industrie.

### Etica utilizării datelor și proiectarea serviciilor smart

#### Data Ethics and smart service design

Acest curs abordează aspecte referitoare la modul în care metodele moderne de colectare și analiză a unor cantități din ce în ce mai mari de date pot ridica probleme de etică în societate. Cursul descrie problema de proiectare utilizând date în diverse formate, ceea ce presupune înțelegerea modului în care se abordează în mod etic un proiect bazat intensiv pe date. Etica utilizării datelor are adresabilitate în toate etapele de dezvoltare a unui astfel de proiect: *administrarea datelor* (Stewarding data) – colectare, gestionare, partajare; *extragerea informației din date* – sub forma de servicii, produse, analize, perspective, vizualizări și data dashboards; *luarea de decizii* – informarea decidenților prin agregarea informației din multiple surse, experiență analiză. Studenții vor învăța cum: să definească diferența între morală, etică, lege; să explice de ce *data ethics* se aplică într-un context mult mai larg decât în utilizarea datelor; ce înseamnă *responsible data science* – asigurarea la nivelul design-ului de serviciu sau produs a corectitudinii (fairness), acurateței (accuracy), confidențialității (confidentiality) și transparenței (transparency) prin proiectare; să identifice și să învețe să gestioneze diverse aspecte de etică în diverse etape ale unui proiect; înțeleagă diversele framework-uri de utilizare etică a datelor; să aplice cunoștințele acumulate în proiectarea serviciilor smart care utilizează intensiv date.