**PROIECTE LA CURSUL BIG DATA**

**Algoritmul general pentru realizarea unui proiect:**

1. **Identificarea problemei și obiectivelor**:
   * Definirea problemei de cercetare sau a temei de proiect.
   * Stabilirea obiectivelor specifice și a rezultatelor așteptate.
2. **Colectarea datelor**:
   * Identificarea surselor de date (structurate sau nestructurate).
   * Colectarea și stocarea datelor relevante folosind baze de date SQL/NoSQL, API-uri, sau surse publice de date.
3. **Preprocesarea datelor**:
   * Curățarea datelor (eliminarea valorilor lipsă sau a zgomotului).
   * Transformarea și normalizarea datelor pentru a fi utilizabile în modelele predictive.
4. **Explorarea și analiza datelor**:
   * Explorarea statistică a setului de date pentru a înțelege corelațiile și tiparele.
   * Vizualizarea datelor pentru o înțelegere mai profundă.
5. **Dezvoltarea modelului**:
   * Alegerea algoritmului de Machine Learning sau Data Mining adecvat (regresie, clasificare, clustering etc.).
   * Antrenarea și validarea modelului pe seturi de date de antrenament și test.
6. **Evaluarea performanței**:
   * Evaluarea modelului folosind metrici de performanță (precizie, recall, acuratețe, MSE etc.).
   * Ajustarea parametrilor modelului pentru a optimiza performanța.
7. **Implementarea și integrarea**:
   * Implementarea soluției finale și integrarea acesteia într-un sistem de producție sau într-o platformă utilizabilă.
8. **Vizualizarea și interpretarea rezultatelor**:
   * Prezentarea rezultatelor folosind instrumente de vizualizare (Tableau, Power BI etc.).
   * Elaborarea unui raport detaliat cu concluzii și recomandări.

**Exemplu temă**

**Tema: Analiza Big Data pentru Dezvoltarea unei Pensiuni Turistice Sustenabile**

**Obiectiv:**

Realizarea unui sistem de analiză Big Data care să sprijine deciziile privind înființarea și dezvoltarea unei pensiuni turistice, folosind surse masive de date (social media, recenzii turistice, date meteo, trenduri economice).

**Tehnologii:**

* **Python** – pentru colectare și preprocesare date;
* **Hadoop / Spark** – pentru stocare și procesare distribuită a volumelor mari de date;
* **NoSQL (MongoDB, Cassandra, Neo4j)** – pentru stocarea recenziilor, relațiilor și trendurilor turistice;
* **APIs (TripAdvisor, Google Maps, Booking, Facebook, Instagram)** – colectarea de date despre recenzii și comportamentul turiștilor;
* **Machine Learning / NLP** – analiză recenzii și predicții de cerere;
* **Tableau / PowerBI / Python (Matplotlib, Seaborn)** – pentru vizualizarea datelor și rapoarte interactive.

**Conținut:**

1. **Colectarea datelor**
   * Recenzii și comentarii din social media (Facebook, Instagram).
   * Date de pe platforme turistice (Booking, TripAdvisor, Google Maps).
   * Date oficiale: statistici (statbank), meteo, infrastructură.
2. **Preprocesarea datelor**
   * Curățarea datelor text (eliminarea stop-words, stemming, lematizare).
   * Normalizarea datelor numerice (ocupare, tarife medii, sezonalitate).
   * Integrarea într-un Data Lake (HDFS + Hive/Spark).
3. **Analiza Big Data**
   * **Analiza sentimentelor** pe recenziile turiștilor (pozitiv/negativ/neutru).
   * **Predicția gradului de ocupare** în funcție de sezon, evenimente, vreme.
   * **Identificarea trendurilor**: preferințe turiști (familie, business, relaxare, academic).
4. **Modelare și Machine Learning**
   * Clasificare recenzii prin NLP (Naive Bayes, SVM, Rețele Neuronale).
   * Modele de regresie pentru previziunea veniturilor.
   * Recomandări pentru optimizarea tarifelor și serviciilor.
5. **Vizualizare și Raportare**
   * Dashboard interactiv cu indicatori: scor satisfacție, sezonalitate, tarife optime.
   * Hărți de tip **heatmap** pentru fluxurile turistice (Google Maps API + vizualizare Big Data).

**Rezultate așteptate:**

* O **hartă interactivă a preferințelor turiștilor** (în funcție de datele analizate).
* Un **raport cu scenarii predictive** (cât de viabilă este investiția într-o pensiune turistică într-o anumită locație).
* Un **model de clasificare a sentimentelor** aplicat pe recenziile turiștilor.
* Recomandări concrete pentru **strategia de marketing și poziționare pe piață**.

**TEME DE PROIECT**

**1. Analiza Sentimentelor pe Rețelele Sociale**

* **Obiectiv**: Determinarea opiniei publice față de un subiect specific folosind analiza sentimentelor pe date colectate de pe rețelele sociale.
* **Tehnologii**: Python, API-uri sociale (Twitter, Facebook), Natural Language Processing (NLP), Hadoop, Spark.
* **Conținut**: Colectarea datelor în timp real, preprocesarea datelor text, utilizarea tehnicilor de NLP și antrenarea unui model de clasificare a sentimentelor.

**2. Sistem de Recoamandare pentru Filme**

* **Obiectiv**: Construirea unui sistem care să recomande filme pe baza istoricului de vizionare al utilizatorilor.
* **Tehnologii**: Apache Spark, Hadoop, NoSQL, Machine Learning, algoritmi de filtrare colaborativă.
* **Conținut**: Colectarea datelor despre vizionările anterioare, antrenarea unui model de filtrare colaborativă, evaluarea performanței și optimizarea recomandărilor.

**3. Predicția Vânzărilor în Retail**

* **Obiectiv**: Crearea unui model predictiv pentru a estima vânzările viitoare pe baza datelor istorice.
* **Tehnologii**: Python, Hadoop, Spark, regresie liniară, Machine Learning.
* **Conținut**: Colectarea datelor de vânzări, antrenarea unui model de regresie pentru predicția vânzărilor și analiza variabilelor influențatoare.

**4. Detectarea Fraudelor în Tranzacții Financiare**

* **Obiectiv**: Implementarea unui sistem de detectare a fraudelor în tranzacții bancare folosind modele de clasificare.
* **Tehnologii**: Python, Spark, Random Forest, SVM, Apache Kafka pentru procesarea în timp real.
* **Conținut**: Colectarea și etichetarea datelor, antrenarea unui model de clasificare și implementarea unui sistem de alertare în timp real.

**5. Monitorizarea Traficului în Timp Real**

* **Obiectiv**: Dezvoltarea unei soluții pentru monitorizarea în timp real a traficului urban.
* **Tehnologii**: Apache Kafka, Spark Streaming, NoSQL, algoritmi de optimizare a fluxului de trafic.
* **Conținut**: Colectarea datelor din surse live (camere video, senzori), prelucrarea și optimizarea fluxului de trafic, vizualizarea în timp real a rezultatelor.

**6. Analiza Performanței Website-urilor**

* **Obiectiv**: Analiza performanței unui site web și detectarea posibilelor îmbunătățiri.
* **Tehnologii**: Python, Hadoop, Spark, Google Analytics API.
* **Conținut**: Colectarea datelor despre performanța site-ului, analizarea timpilor de răspuns și optimizarea acestora.

**7. Predicția Consumului de Energie**

* **Obiectiv**: Utilizarea datelor istorice pentru a crea un model care să prezică consumul de energie.
* **Tehnologii**: Hadoop, Spark, Machine Learning, modele de regresie.
* **Conținut**: Colectarea și analiza datelor, dezvoltarea unui model de predicție și evaluarea acestuia.

**8. Detectarea Anomaliilor în Rețele**

* **Obiectiv**: Detectarea anomaliilor în traficul din rețele pentru a preveni atacuri sau utilizări necorespunzătoare.
* **Tehnologii**: Hadoop, Spark, algoritmi de detectare a anomaliilor (Isolation Forest, Autoencoder).
* **Conținut**: Colectarea și analiza traficului de rețea, antrenarea unui model de detectare a anomaliilor, monitorizarea în timp real a traficului.

**9. Analiza Datelor Meteo pentru Predicții Climatice**

* **Obiectiv**: Construirea unui model pentru predicția condițiilor meteorologice viitoare.
* **Tehnologii**: Python, Spark, modele de regresie, Machine Learning.
* **Conținut**: Colectarea datelor meteo, preprocesarea datelor și antrenarea unui model pentru predicția climatului.

**10. Optimizarea Proceselor în Lanțul de Aprovizionare**

* **Obiectiv**: Optimizarea aprovizionării în funcție de cererea estimată utilizând Big Data.
* **Tehnologii**: Hadoop, Spark, NoSQL, Machine Learning.
* **Conținut**: Colectarea datelor despre cereri și stocuri, predicția cererii și optimizarea proceselor de aprovizionare.

**11. Analiza Comportamentului de Cumpărare al Utilizatorilor**

* **Obiectiv**: Identificarea tiparelor de cumpărare și clasificarea utilizatorilor în segmente.
* **Tehnologii**: Apache Spark, Hadoop, Machine Learning (clustering, k-means), NoSQL.
* **Conținut**: Colectarea și preprocesarea datelor de tranzacții, segmentarea utilizatorilor pe baza comportamentului și analizarea rezultatelor pentru îmbunătățirea strategiilor de marketing.

**12. Clasificarea Automată a E-mailurilor**

* **Obiectiv**: Crearea unui sistem automat de clasificare a e-mailurilor în diferite categorii (spam, non-spam).
* **Tehnologii**: Python, Natural Language Processing (NLP), Hadoop, Spark, Logistic Regression.
* **Conținut**: Colectarea datelor despre e-mailuri, preprocesarea textului și dezvoltarea unui model de clasificare bazat pe tehnici NLP.

**13. Analiza Big Data pentru Detectarea Timpurilor de Așteptare în Spitale**

* **Obiectiv**: Predicția și optimizarea timpului de așteptare pentru pacienți în spitale.
* **Tehnologii**: Hadoop, Spark, modele de regresie, Machine Learning.
* **Conținut**: Colectarea datelor istorice despre timp de așteptare, dezvoltarea unui model predictiv și implementarea soluției de optimizare.

**14. Optimizarea Managementului Inventarului folosind Big Data**

* **Obiectiv**: Optimizarea inventarului bazat pe datele de vânzări și cerere.
* **Tehnologii**: Apache Spark, NoSQL, Machine Learning (regresie).
* **Conținut**: Analiza datelor istorice despre stocuri și vânzări, dezvoltarea unui model pentru optimizarea inventarului în timp real.

**15. Analiza Tranzacțiilor pe Piața de Capital**

* **Obiectiv**: Analiza datelor istorice ale tranzacțiilor bursiere pentru a prezice tendințele pieței.
* **Tehnologii**: Hadoop, Spark, Machine Learning (time series forecasting), Python.
* **Conținut**: Colectarea datelor bursiere, prelucrarea și dezvoltarea unui model predictiv pentru prețurile acțiunilor.

**16. Analiza Preferințelor Utilizatorilor într-o Platformă de Streaming**

* **Obiectiv**: Descoperirea preferințelor utilizatorilor pe baza datelor de vizionare.
* **Tehnologii**: Hadoop, Spark, algoritmi de recomandare, NoSQL.
* **Conținut**: Colectarea și prelucrarea datelor de vizionare, antrenarea unui model de recomandare pentru a oferi utilizatorilor conținut personalizat.

**17. Automatizarea Răspunsului pentru Suport Tehnic folosind Chatbot-uri**

* **Obiectiv**: Dezvoltarea unui chatbot care să ofere răspunsuri automate la cererile de suport.
* **Tehnologii**: Python, NLP, Machine Learning, Apache Kafka.
* **Conținut**: Colectarea datelor din conversațiile anterioare, preprocesarea textului și antrenarea chatbot-ului folosind tehnici de NLP.

**18. Predicția Tendințelor în Fashion Retail**

* **Obiectiv**: Folosirea Big Data pentru a prezice tendințele în industria de modă.
* **Tehnologii**: Hadoop, Spark, Machine Learning (clustering, time series analysis).
* **Conținut**: Colectarea datelor de vânzări și feedback

**19. Analiza Datelor Despre Utilizarea Energiei pentru Eficiență Energetică**

* **Obiectiv**: Optimizarea consumului de energie în clădiri folosind date colectate de la senzori.
* **Tehnologii**: Hadoop, Spark, IoT, Machine Learning (regresie).
* **Conținut**: Colectarea datelor de la senzori, preprocesarea acestora și dezvoltarea unui model predictiv pentru optimizarea consumului de energie.

**20. Detectarea Schimbărilor Climatice prin Analiza Datelor Meteo**

* **Obiectiv**: Monitorizarea și detectarea schimbărilor climatice pe baza datelor istorice.
* **Tehnologii**: Hadoop, Spark, Machine Learning (time series), Python.
* **Conținut**: Colectarea datelor meteo din diverse surse, analizarea acestora și dezvoltarea unui model predictiv pentru schimbările climatice.

**21. Analiza Big Data pentru Optimizarea Rețelelor de Transport**

* **Obiectiv**: Optimizarea rutelor de transport public pe baza datelor istorice și în timp real.
* **Tehnologii**: Hadoop, Spark Streaming, Machine Learning, NoSQL.
* **Conținut**: Colectarea datelor de trafic și a rutelor, analizarea fluxurilor de transport și optimizarea rutelor pentru eficiență maximă.

**22. Predicția Eșecurilor în Producție cu Big Data**

* **Obiectiv**: Prevenirea eșecurilor în producție prin analizarea datelor istorice despre echipamente și producție.
* **Tehnologii**: Hadoop, Spark, Machine Learning (regresie), IoT.
* **Conținut**: Colectarea datelor despre echipamente, preprocesarea acestora și dezvoltarea unui model de predicție pentru eșecuri și mentenanță preventivă.

**23. Securitatea Datelor în Cloud prin Analiza Big Data**

* **Obiectiv**: Detectarea și prevenirea atacurilor cibernetice asupra datelor stocate în cloud.
* **Tehnologii**: Hadoop, Spark, Machine Learning (anomaly detection), NoSQL.
* **Conținut**: Colectarea datelor despre acces și securitate, dezvoltarea unui model pentru detectarea anomaliilor și prevenirea atacurilor.

**24. Optimizarea Campaniilor de Marketing prin Analiza Comportamentului Utilizatorilor**

* **Obiectiv**: Optimizarea campaniilor de marketing pe baza comportamentului de cumpărare și vizualizare al utilizatorilor.
* **Tehnologii**: Hadoop, Spark, algoritmi de clustering și regresie, Machine Learning.
* **Conținut**: Colectarea datelor despre utilizatori, segmentarea acestora și optimizarea campaniilor de marketing pe baza preferințelor și comportamentului lor.

**25. Analiza Social Media pentru Detectarea Tendințelor Virale**

* **Obiectiv**: Detectarea timpurie a trendurilor virale pe rețelele sociale.
* **Tehnologii**: Python, Spark, Machine Learning, NLP.
* **Conținut**: Colectarea datelor de pe rețelele sociale, utilizarea algoritmilor de Machine Learning pentru detectarea tiparelor și identificarea trendurilor emergente.

**26. Analiza Datelor de Sănătate pentru Predicția Bolilor Cronice**

* **Obiectiv**: Identificarea și predicția riscului de boli cronice pe baza datelor de sănătate ale pacienților.
* **Tehnologii**: Hadoop, Spark, Machine Learning (classification, clustering).
* **Conținut**: Colectarea datelor medicale, prelucrarea acestora și dezvoltarea unui model predictiv pentru riscul de boli cronice.

**27. Analiza Sentimentelor în Comentarii Online**

* **Obiectiv**: Detectarea sentimentului public față de produse sau servicii pe baza comentariilor online.
* **Tehnologii**: Python, NLP, Hadoop, Spark.
* **Conținut**: Colectarea și preprocesarea datelor textuale din recenzii, antrenarea unui model pentru clasificarea sentimentelor și analiza acestora.

**28. Detectarea Fraudelor în Asigurări prin Big Data**

* **Obiectiv**: Crearea unui sistem de detectare a fraudelor în asigurări folosind Big Data și Machine Learning.
* **Tehnologii**: Hadoop, Spark, Machine Learning (classification, anomaly detection).
* **Conținut**: Colectarea datelor despre tranzacții și revendicări, preprocesarea acestora și dezvoltarea unui model pentru detectarea fraudelor.

**29. Analiza Preferințelor de Călătorie și Predicția Destinațiilor Viitoare**

* **Obiectiv**: Descoperirea tiparelor de călătorie și predicția destinațiilor preferate ale utilizatorilor.
* **Tehnologii**: Hadoop, Spark, Machine Learning (clustering, recommendation systems).
* **Conținut**: Colectarea datelor despre călătoriile anterioare, segmentarea utilizatorilor și oferirea de recomandări personalizate.

**30. Optimizarea Prețurilor pe Platformele de E-Commerce folosind Big Data**

* **Obiectiv**: Optimizarea prețurilor produselor pe baza analizelor comportamentale și istoricului de vânzări.
* **Tehnologii**: Spark, Hadoop, algoritmi de Machine Learning (regression, classification).
* **Conținut**: Colectarea și analizarea datelor despre prețuri, vânzări și cerere, optimizarea algoritmilor de stabilire a prețurilor.

**31. Clasificarea Documentelor prin Tehnici de Text Mining**

* **Obiectiv**: Crearea unui sistem de clasificare automată a documentelor pe baza conținutului text.
* **Tehnologii**: Python, NLP, Hadoop, Spark.
* **Conținut**: Colectarea și preprocesarea documentelor, antrenarea unui model de clasificare a textului, evaluarea performanței și ajustarea parametrilor modelului.

**32. Analiza Comportamentului Utilizatorilor pentru Optimizarea Funcționalității Aplicațiilor Mobile**

* **Obiectiv**: Optimizarea funcționalităților aplicațiilor mobile pe baza comportamentului utilizatorilor.
* **Tehnologii**: Spark, Hadoop, Machine Learning (clustering, classification).
* **Conținut**: Colectarea datelor de utilizare a aplicației, segmentarea utilizatorilor și optimizarea experienței utilizatorilor prin funcționalități personalizate.

**33. Predicția Înclinației Clienților pentru Împrumuturi în Sistemul Bancar**

* **Obiectiv**: Dezvoltarea unui model pentru predicția șanselor ca un client să solicite un împrumut.
* **Tehnologii**: Hadoop, Spark, algoritmi de regresie logistică, Machine Learning.
* **Conținut**: Analiza datelor financiare ale clienților și antrenarea unui model de predicție pentru acordarea împrumuturilor.

**34. Analiza Rețelelor de Distribuție pentru Optimizarea Logisticii**

* **Obiectiv**: Optimizarea rețelelor de distribuție folosind Big Data pentru reducerea costurilor și creșterea eficienței.
* **Tehnologii**: Hadoop, Spark, Machine Learning (optimization algorithms).
* **Conținut**: Colectarea și analizarea datelor despre livrări și distribuție, optimizarea rutelor și a proceselor logistice.

**35. Optimizarea Căutării în Magazinul Online folosind Big Data**

* **Obiectiv**: Îmbunătățirea rezultatelor căutărilor în magazinele online pe baza datelor istorice și comportamentale.
* **Tehnologii**: Hadoop, Spark, algoritmi de recomandare, NLP.
* **Conținut**: Analiza datelor despre căutările utilizatorilor, optimizarea rezultatelor căutării și a sugestiilor de produse.

**36. Sistem de Recoamandare pentru Cărți în Biblioteci Digitale**

* **Obiectiv**: Construirea unui sistem de recomandare care să sugereze cărți utilizatorilor pe baza intereselor și lecturilor anterioare.
* **Tehnologii**: Hadoop, Spark, Machine Learning (collaborative filtering).
* **Conținut**: Colectarea datelor despre lecturile anterioare, antrenarea unui model de recomandare și evaluarea performanței acestuia.

**37. Predicția Consumului de Apă în Orașe Inteligente**

* **Obiectiv**: Predictarea consumului de apă pentru a optimiza resursele.
* **Tehnologii**: IoT, Hadoop, Spark, Machine Learning (time series analysis).
* **Conținut**: Colectarea datelor de consum de la senzori, analizarea acestora și crearea unui model pentru predicția consumului viitor.

**38. Optimizarea Planificării Resurselor în Companii**

* **Obiectiv**: Optimizarea resurselor umane și materiale în companii prin analiza datelor.
* **Tehnologii**: Hadoop, Spark, algoritmi de optimizare.
* **Conținut**: Colectarea datelor din procesele de planificare, dezvoltarea unui model pentru alocarea resurselor în mod optim.

**39. Sistem de Alertare pentru Managementul Dezastrelor Naturale**

* **Obiectiv**: Dezvoltarea unui sistem automat de alertare în caz de dezastre naturale bazat pe analiza datelor în timp real.
* **Tehnologii**: Hadoop, Spark Streaming, Machine Learning (anomaly detection).
* **Conținut**: Colectarea datelor de la diverse surse (senzori, sateliți), dezvoltarea unui model de detectare și implementarea unui sistem de alertă în timp real.

**40. Detectarea Incidentelor de Securitate în Rețele de Calculatoare**

* **Obiectiv**: Detectarea și prevenirea incidentelor de securitate în rețele de calculatoare folosind Big Data.
* **Tehnologii**: Hadoop, Spark, algoritmi de detecție a anomaliilor, Machine Learning.
* **Conținut**: Colectarea și analizarea datelor de trafic din rețea, identificarea anomaliilor și dezvoltarea unui model de prevenire a atacurilor cibernetice.

**41. Analiza Big Data pentru Îmbunătățirea Serviciilor în Industria Hotelieră**

* **Obiectiv**: Creșterea satisfacției clienților în industria hotelieră prin analiza feedback-ului și comportamentului.
* **Tehnologii**: Hadoop, Spark, Machine Learning (clustering, sentiment analysis).
* **Conținut**: Colectarea datelor de feedback și analizarea comportamentului clienților pentru personalizarea serviciilor și îmbunătățirea experienței acestora.

**42. Optimizarea Lanțului de Aprovizionare cu Big Data**

* **Obiectiv**: Reducerea costurilor și optimizarea timpilor de livrare în lanțul de aprovizionare.
* **Tehnologii**: Hadoop, Spark, Machine Learning (optimization algorithms).
* **Conținut**: Analizarea datelor despre furnizori și transporturi, optimizarea rutelor și gestionarea stocurilor.

**43. Predicția Riscului de Creditare pe Baza Datelor de Istoric Financiar**

* **Obiectiv**: Dezvoltarea unui model pentru a prezice riscul de creditare al clienților bancari.
* **Tehnologii**: Hadoop, Spark, Machine Learning (classification, regression).
* **Conținut**: Colectarea și preprocesarea datelor financiare ale clienților, dezvoltarea unui model de evaluare a riscului și aplicarea acestuia pentru decizii de creditare.

**44. Sistem de Recoamandare de Filme în Funcție de Comportamentul Utilizatorului**

* **Obiectiv**: Crearea unui sistem de recomandare care sugerează filme utilizatorilor pe baza vizionărilor anterioare și a preferințelor.
* **Tehnologii**: Hadoop, Spark, Machine Learning (collaborative filtering).
* **Conținut**: Colectarea datelor despre vizionări, antrenarea unui model de recomandare și optimizarea acestuia pentru a oferi sugestii personalizate.

**45. Predicția Defectelor de Producție în Industria Automotive**

* **Obiectiv**: Detectarea și prevenirea defectelor de fabricație în industria auto pe baza datelor istorice de producție.
* **Tehnologii**: Hadoop, Spark, Machine Learning (classification, anomaly detection).
* **Conținut**: Colectarea datelor de producție, prelucrarea acestora și dezvoltarea unui model pentru detectarea timpurie a defectelor.

**46. Analiza Datelor pentru Creșterea Conversiilor pe Magazinele Online**

* **Obiectiv**: Creșterea ratei de conversie a vizitelor în vânzări pe platformele e-commerce.
* **Tehnologii**: Hadoop, Spark, Machine Learning (clustering, regression).
* **Conținut**: Colectarea și analizarea datelor despre comportamentul vizitatorilor, segmentarea acestora și optimizarea campaniilor de marketing.

**47. Predicția Răspunsurilor Pacienților la Tratament prin Analiza Big Data**

* **Obiectiv**: Identificarea factorilor care influențează răspunsul pacienților la tratamente medicale.
* **Tehnologii**: Hadoop, Spark, Machine Learning (classification, clustering).
* **Conținut**: Colectarea datelor medicale și prelucrarea acestora pentru a dezvolta un model predictiv care să anticipeze răspunsul pacienților la tratamente.

**48. Optimizarea Energiei în Rețelele de Producție prin IoT și Big Data**

* **Obiectiv**: Crearea unui sistem pentru optimizarea consumului energetic în procesele de producție.
* **Tehnologii**: Hadoop, Spark, IoT, Machine Learning (optimization algorithms).
* **Conținut**: Colectarea datelor de la echipamente industriale, analizarea acestora și optimizarea consumului de energie pe baza modelelor predictive.

**49. Analiza Datelor Despre Comportamentul Școlar pentru Îmbunătățirea Educației**

* **Obiectiv**: Îmbunătățirea performanțelor academice prin analiza datelor despre comportamentul elevilor.
* **Tehnologii**: Hadoop, Spark, Machine Learning (clustering, classification).
* **Conținut**: Colectarea datelor despre participarea la cursuri, notele și comportamentul elevilor, dezvoltarea unui model pentru identificarea factorilor care influențează succesul școlar.

**50. Optimizarea Fluxurilor de Lucru în Industriile de Producție cu Big Data**

* **Obiectiv**: Creșterea eficienței fluxurilor de lucru în fabrici prin analiza și optimizarea proceselor de producție.
* **Tehnologii**: Hadoop, Spark, Machine Learning (optimization algorithms).
* **Conținut**: Colectarea datelor din procesele de producție, analizarea acestora și dezvoltarea unor modele pentru optimizarea fluxurilor de lucru și reducerea pierderilor.