



**UNIVERSITATEA BABEȘ-BOLYAI**  
Facultatea de Matematică și Informatică



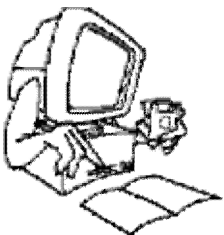
# **Proiecte**

## Metode Inteligente de Rezolvare a Problemelor Reale (MIRPR)

Laura Dioșan

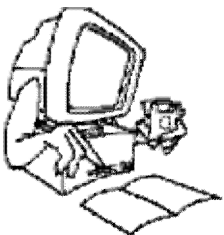
CUPRINS

<b>1. <u>TEME DE PROIECT</u></b>	<b>3</b>
<b>2. <u>CERINTE PENTRU REALIZAREA PROIECTELOR</u></b>	<b>4</b>
<b>3. <u>DETALII DESPRE PROIECTE</u></b>	<b>6</b>



## 1. Teme de proiect

1. [Rezolvarea problemelor de planificare](#)
2. [Identificarea pietonilor in imagini](#)
3. [Recunoasterea pietonilor in imagini](#)
4. [Identificarea produselor falsificate cu ajutorul imaginilor](#)
5. [Modelarea psihologica](#)
6. [Recunoasterea de obiecte in imagini](#)
7. [Gruparea imaginilor care contin diferite obiecte](#)
8. [Recunoasterea de fete in imagini](#)
9. [Recunoaşterea scrisului de mână](#)
10. [Previziunea vremii](#)



## 2. Cerințe pentru realizarea proiectelor

Proiectul pe care trebuie să-l realizați este o oportunitate de a explora o problemă din domeniul *Machine Learning (ML)* în contextul unor date reale. Proiectul va fi evaluat la finalul semestrului, dar pe parcursul semestrului fiecare echipă va trebui să furnizeze cadrului didactic îndrumător aplicația realizată și raportul aferent ei, în următoarea ordine:

- a. propunerea – 1 pagină (10% din nota finală)
- b. prima parte din raport – 3 – 4 pagini (20% din nota finală)
- c. raportul complet – 6-8 pagini (40% din nota finală)
- d. prezentare (30% din nota finală)

Proiectul implică rezolvarea unei probleme prin 2 metode diferite și analizarea rezultatelor obținute.

Citiți lista de date disponibile și proiecte potențiale. Se recomandă folosirea acestor date, însă este posibilă folosirea și altor surse (care trebuie să fie disponibile la momentul livrării propunerii de proiect).

### a. Propunerea

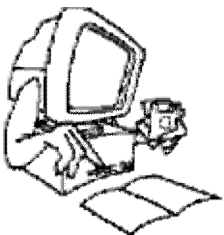
Propunerea proiectului trebuie să fie redactată pe maxim o pagină și să conțină:

- titlul proiectului și datele de identificare ale echipei
- datele de test
- ideea proiectului (aproximativ 2 paragrafe)
- lucrările care ar trebui citite (documentație) – bineînțeles că anumite lucrări ar trebui citite înainte de a redacta această propunere ☺
- care sunt obiectivele pe care vi le propuneți să le realizați până la următorul deadline (prima parte din raport)

### b. Prima parte din raport

Este un raport similar celui final, dar mai puțin complet. Trebuie să aibă aceeași structură cu raportul final:

- introducere și motivație
- definirea precisă a problemei
- abordări înrudite
- metoda de lucru
  - o de ce a fost aleasă această metodă
  - o descrierea modului de lucru (algoritmului)
- experimentele
  - o descrierea datelor (sursa datelor, clasificarea/tipologia datelor) și a modului de testare (parametrii, mod de testare)
  - o care sunt întrebările la care ar trebui să răspundă experimentele efectuate
  - o rezultatele obținute (măsurile de performanță calculate ca urmare a aplicării celor 2 algoritmi de ML pentru rezolvarea problemei)
- concluziile.



Unele secţiuni sau subsecţiuni sunt încă „under construction”. Secţiunile *Introducere*, *Definirea problemei* şi *Abordări înrudite* trebuie să fie în forma finală. Secţiunile *Metoda de lucru* şi *Experimentele* trebuie să fie schiţate în linii mari.

Evaluarea acestei forme a raportului:

- 70% pentru introducere (cu motivaţie cu tot) şi abordări similare
- 30% pentru metoda propusă şi schiţarea ei

**c. Raportul complet**

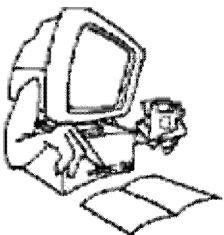
Completarea tuturor secţiunilor cu informaţiile necesare. Descrierea şi prezentarea detaliată a modului de lucru şi a rezultatelor obţinute.

Evaluarea acestei forme a raportului:

- 70% pentru partea experimentală
- 30% pentru concluzii şi direcţii viitoare

**d. Prezentarea**

Trebuie realizată o prezentare (PowerPoint sau alt tip) a muncii realizate de-a lungul semestrului. Structura prezentării urmează îndeaproape structura raportului, cu un accent deosebit pus pe rezultatele obţinute. Prezentarea trebuie să conţină aproximativ 10 slide-uri şi să dureze maxim 10 minute. Expunerea trebuie să includă şi prezentarea aplicaţiei realizate şi folosite pentru efectuarea experimentelor.



### 3. Detalii despre proiecte

#### PROIECTUL 1

Problema planificării (Scheduling)



#### Obiective

Planificarea eficientă a resurselor



#### Ideea proiectului

Asignarea unor resurse limitate unui set de entităţi (sarcini, evenimente publice, vehicule, indivizi) de-a lungul unei perioade astfel încât să se respecte anumite condiţii. Problema mai este cunoscută şi ca:

- a. Course timetabling problem
- b. Railway Scheduling Problems
- c. Scheduling Aircraft Landings
- d. Resource Constraint Project Scheduling Problem
- e. Job shop scheduling

Un exemplu concret de problema se poate găsi în fişierul `schedulingProblem.pdf`



#### Lista de *TO DO*-uri

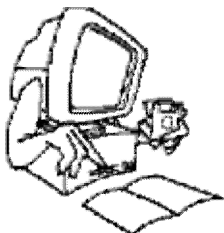
1. Alegerea unui set  $S$  de instanţe pentru problema dată.
2. Stabilirea unei reprezentări a soluţiei pentru problema dată.
3. Stabilirea unei funcţii de cost (care urmează să fie optimizată). Funcţia de cost poate fi una sau multe obiective.
4. Utilizarea unui algoritm de optimizare (EA, ACO, PSO, etc.) pentru a stabili valoarea optimă a funcţiei obiectiv.
5. Găsirea şi prezentarea soluţiei.



#### Date şi bibliografie

- a. <http://www.inf.ufpr.br/aurora/disciplinas/topicosia2/downloads/artigos/hyperEA.pdf>
- b. <http://www.cardiff.ac.uk/carbs/quant/rhyd/TTSurvey.pdf>
- c. <http://www.ascent.com/>
- d. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.38.7543&rep=rep1&type=pdf>
- e. <http://users.dsic.upv.es/grupos/gps/papers/Genetic-Railway.pdf>
- f. <http://www.stottlerhenke.com/products/aurora/?gclid=COi4-679xaUCFc4H3wodF2kWYQ>
- g. <http://whentowork.com/?gclid=CPCvIbD9xaUCFZQK3wodqzCNXw>
- h. <http://www.sciencedirect.com/science/article/B6VCT-4CK1XR8-2/2/72a80155c587e28e67dadb028ac71491>
- i. [http://www.dbcenter.cise.ufl.edu/seek/Publications/RCPSP\\_YGO.pdf](http://www.dbcenter.cise.ufl.edu/seek/Publications/RCPSP_YGO.pdf)

&



## PROIECTUL 2

Detectia de pietoni in imagini



### Obiective

Identificarea pietonilor in imagini cu ajutorul unui model de clasificare învăţat pe un set de imagini deja adnotate.



### Ideea proiectului

Se dă un set de imagini captate din trafic cu o camera positionata pe un vehicul (imaginile contin pietoni). Se cere să se identifice in imaginile date pozitia (coordonatele) pietonilor.



### Lista de *TO DO*-uri

1. Alegerea unui set  $S$  de imagini (o parte adnotate – pentru antrenament - SA, o parte neadnotate – pentru testare - ST; ambele subseturi trebuie sa contina atat imagini cu pietoni, cat si imagini fara pietoni).
2. Extragerea unui set de caracteristici din fiecare imagine din  $S$
3. Invatarea unui model de clasificare prin rulara unui algoritm de invatare automata (ANN, SVM, EAs, etc) pe caracteristicile imaginilor din SA.
4. Detectarea pietonilor in imaginile din ST pe baza modelului anterior invatat.

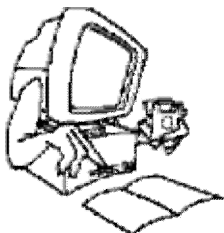


### Date și bibliografie

&



<http://www.lookingatpeople.com/download-daimler-ped-det-benchmark/index.html>  
<http://www.lookingatpeople.com/download-daimler-stereo-ped-det-benchmark/index.html>  
<http://www.lookingatpeople.com/download-daimler-stereo-ped-det-benchmark/index.html>  
<http://www.lookingatpeople.com/download-daimler-ped-det-benchmark/index.html>  
<http://www.pedestrian-detection.com/>  
<http://www.gavrila.net/Publications/door2door01.pdf>  
[http://www.cvc.uab.es/~dgeronimo/publications/geronimo\\_pami2010.pdf](http://www.cvc.uab.es/~dgeronimo/publications/geronimo_pami2010.pdf)  
<http://ebookbrowse.com/survey-of-pedestrian-detection-for-advanced-driver-assistance-pdf-d264642098>  
<http://www.vision.caltech.edu/publications/dollarCVPR09pedestrians.pdf>



**PROIECTUL 3**

Recunoaştere de pietoni in imagini

**Obiective**

Recunoaşterea pietonilor in imagini cu ajutorul unui model de clasificare învăţat pe un set de imagini deja adnotate.

**Ideea proiectului**

Se dă un set de imagini captate din trafic cu o camera positionata pe un vehicul (unele imagini contin pietoni, altele nu contin pietoni). Unele imagini au fost deja clasificate (adnotate) ca imagini cu pietoni. Se cere să se identifice etichetele (cu/fara pieton) corespunzatoare imaginilor ne-adnotate.

**Lista de *TO DO*-uri**

1. Alegerea unui set  $S$  de imagini (o parte adnotate – pentru antrenament - SA, o parte neadnotate – pentru testare - ST; ambele subseturi trebuie sa contina atat imagini cu pietoni, cat si imagini fara pietoni).
2. Extragerea unui set de caracteristici din fiecare imagine din  $S$
3. Invatarea unui model de clasificare prin rulara unui algoritm de invatare automata (ANN, SVM, EAs, etc) pe caracteristicile imaginilor din SA.
4. Recunoasterea pietonilor in imaginile din ST pe baza modelului anterior invatat.

**Date și bibliografie**

&amp;

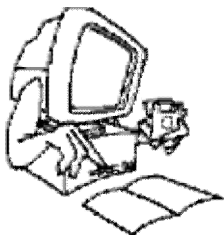


<http://www.lookingatpeople.com/download-daimler-ped-class-benchmark/index.html>

<http://www.lookingatpeople.com/download-daimler-ped-mcucoccl-class-benchmark/index.html>

<http://www.gavrila.net/pami06.pdf>

[http://www.gavrila.net/cvpr10\\_occlusion.pdf](http://www.gavrila.net/cvpr10_occlusion.pdf)





**PROIECTUL 4**

## Recunoaştere de produse falsificate

**Obiective**

Recunoaşterea produselor originale si suspectarea produselor false cu ajutorul unui model de clasificare învăţat pe un set de imagini deja adnotate.

**Ideea proiectului**

Se dă un set de imagini (unele imagini contin produse originale, altele contin produse contrafacute). Unele imagini au fost deja clasificate (adnotate) ca imagini de produse originale/contrafacute, iar alte imagini nu au fost etichetate. Se cere să se identifice etichetele (original/contrafacut) corespunzatoare imaginilor ne-adnotate.

**Lista de *TO DO*-uri**

1. Alegerea unui set  $S$  de imagini (o parte adnotate – pentru antrenament - SA, o parte neadnotate – pentru testare - ST; ambele subseturi trebuie sa contina atat imagini ale unor produse originale, cat si imagini ale produselor contrafacute).
2. Extragerea unui set de caracteristici din fiecare imagine din  $S$
3. Invatarea unui model de clasificare prin rulara unui algoritm de invatare automata (ANN, SVM, EAs, etc) pe caracteristicile imaginilor din SA.
4. Recunoasterea produselor originale in imaginile din ST pe baza modelului anterior invatat.

**Date și bibliografie**

[????](#)

&



**PROIECTUL 5**

Modelare de date psihologice (bodymedia)

**Obiective**

Modelare comportamentului uman și luarea unor decizii pe baza unor măsurători deja efectuate.

**Ideea proiectului**

Clasificarea comportamentului uman pe baza anumitor măsurători privind:

characteristic1	age
characteristic2	handedness
sensor1	gsr_low_average
sensor2	heat_flux_high_average
sensor3	near_body_temp_average
sensor4	pedometer
sensor5	skin_temp_average
sensor6	longitudinal_accelerometer_SAD
sensor7	longitudinal_accelerometer_average
sensor8	transverse_accelerometer_SAD
sensor9	transverse_accelerometer_average

Scopul este identificarea genului (feminin sau masculin) diferiților indivizi pe baza acestor informații. În fapt problema se reduce la o clasificare binară.

**Lista de TO DO-uri**

1. Alegerea unui set  $S$  de date (o parte din date sunt cupluri (intrari, iesiri) – pentru antrenament - SA, o parte sunt doar intrari– pentru testare - ST).
2. Invatarea unui model de predictie  $f$  (iesire =  $f$ (intrare)) prin rularea unui algoritm de invatare automata (ANN, SVM, EAs, etc) pe setul de date SA.
3. Extrapolarea modelului invatat pentru datele din ST.

**Date și bibliografie**

A se consulta directorul: Data/psiho

&

[http://www.inf.ed.ac.uk/teaching/courses/dme/html/datasets0405.ht](http://www.inf.ed.ac.uk/teaching/courses/dme/html/datasets0405.html)

[ml](#)



**PROIECTUL 6**

Recunoaşterea de obiecte in imagini

**Obiective**

Recunoaşterea (clasificarea) unui nou obiect pe baza unui model de clasificare învăţat pe un set de obiecte deja adnotate.

**Ideea proiectului**

Se dă un set de imagini colectate de pe Google Image (unele adnotate, altele nu) şi se cere să se identifice categoriile corecte pentru imaginile ne-adnotate.

**Lista de *TO DO*-uri**

1. Alegerea unui set  $S$  de imagini (o parte adnotate – pentru antrenament - SA, o parte neadnotate – pentru testare - ST; ambele subseturi trebuie sa contina atat imagini cu obiectul respectiv, cat si imagini fara obiectul respectiv).
2. Extragerea unui set de caracteristici din fiecare imagine din  $S$
3. Invatarea unui model de clasificare prin rulara unui algoritm de invatare automata (ANN, SVM, EAs, etc) pe caracteristicile imaginilor din SA.
4. Recunoasterea obiectului in imaginile din ST pe baza modelului anterior invatat.

**Date şi bibliografie**

&amp;



A se consulta directorul: Data/objRecogn

[http://www.vision.caltech.edu/Image\\_Datasets/Caltech256](http://www.vision.caltech.edu/Image_Datasets/Caltech256)



**PROIECTUL 7**

Gruparea imaginilor care contin diferite obiecte

**Obiective**

Etichetarea unui nou obiect pe baza includerii lui în grupul cel mai apropiat lui pe baza unui model învățat pe un set de obiecte.

**Ideea proiectului**

Se dă un set de imagini colectate de pe Google Image (neadnotate) și se cere să se grupeze aceste imagini în grupuri cât mai omogene și să se eticheteze aceste grupuri. În fapt, problema constă în clusterizarea acestor imagini.

**Lista de *TO DO*-uri**

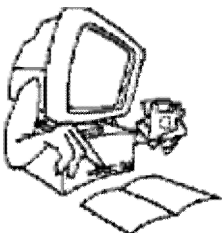
1. Alegerea unui set  $S$  de imagini (o parte adnotate – pentru antrenament - SA, o parte neadnotate – pentru testare - ST; ambele subseturi trebuie sa contina imagini etichetate ca facand parte din diferite grupuri).
2. Extragerea unui set de caracteristici din fiecare imagine din  $S$
3. Invatarea unui model de clusterizare prin rulara unui algoritm de invatare automata pe caracteristicile imaginilor din SA.
4. Etichetarea imaginilor din ST pe baza modelului anterior invatat.

**Date și bibliografie**

&



A se consulta directorul: Data/objRecogn  
[http://www.vision.caltech.edu/Image\\_Datasets/Caltech256](http://www.vision.caltech.edu/Image_Datasets/Caltech256)



**PROIECTUL 8**

## Recunoaştere de feţe in imagini

**Obiective**

Clasificarea unei noi imagini reprezentând o faţă (asocierea faţă-persoană) pe baza unui model deja învăţat pe un set de imagini deja adnotate.

**Ideea proiectului**

De-a lungul ultimilor ani, recunoaşterea feţei a devenit o zonă populară de cercetare şi una dintre aplicaţiile cele mai de succes de analiză şi înţelegere a imaginii. Natura problemei a determinat ca nu doar cercetătorii din informatică să fie interesaţi de aceasta, ci şi neurologii şi psihologii. Progresele informatice în acest domeniu vor oferi perspective utile pentru neurologi şi psihologi în modul în care funcţionează creierul uman, şi vice-versa.

O definiţie generală a problemei de recunoaştere a feţei (în computer vision) poate fi formulată astfel: dându-se imaginile (statice sau video) ale unei scene, să se identifice sau să se verifice (cu ajutorul unei baze de imagini) dacă una sau mai multe personaje apar în scenă.

**Lista de *TO DO*-uri**

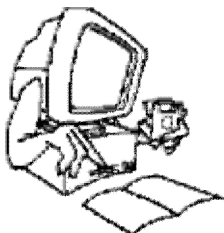
1. Alegerea unui set  $S$  de imagini (o parte adnotate – pentru antrenament - SA, o parte neadnotate – pentru testare - ST; ambele subseturi trebuie să conţină atât imagini cu feţe, cât şi imagini fără feţe).
2. Extragerea unui set de caracteristici din fiecare imagine din  $S$
3. Învăţarea unui model de clasificare prin rularea unui algoritmul de învăţare automată (ANN, SVM, EAs, etc) pe caracteristicile imaginilor din SA.
4. Recunoaşterea feţelor în imaginile din ST pe baza modelului anterior învăţat.

**Date şi bibliografie**

&amp;



A se consulta directorul: Data/faceRecognition  
<http://vis-www.cs.umass.edu/fddb/index.html>



**PROIECTUL 9**

## Recunoașterea scrisului de mână

**Obiective**

Clasificarea unei noi imagini reprezentând o cifra/literă scrisă de mână pe baza unui model învățat pe un set de imagini deja adnotate.

**Ideea proiectului**

Recunoașterea scrisului de mână este capacitatea unui computer de a interpreta inteligibil intrări scrise de mână provenite de la surse precum documente de hârtie, fotografii, *touch-screen*-uri și alte dispozitive. Imaginea textului scris poate fi percepută *off-line* de pe o bucată de hârtie prin scanare optică (recunoaștere optică a caracterelor) sau prin recunoaștere inteligentă a cuvintelor. Alternativ, mișcările de peniță pot fi percepute *on-line*, de exemplu, de pe o suprafață digitală în momentul trasării liniilor cu creionul optic.

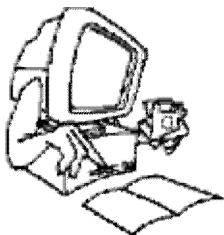
Recunoașterea scrisului de mână presupune în principal recunoașterea optică a caracterelor. Cu toate acestea, un sistem complet de recunoaștere a scrisului de mână, trebuie să țină cont și de formatarea textului, de segmentarea corectă a lui în cuvinte și caractere.

**Lista de *TO DO*-uri**

1. Alegerea unui set  $S$  de imagini (o parte adnotate – pentru antrenament - SA, o parte neadnotate – pentru testare - ST; ambele subseturi trebuie sa contina imagini cu toate cifrele).
2. Extragerea unui set de caracteristici din fiecare imagine din  $S$
3. Invatarea unui model de clasificare prin rularea unui algoritm de invatare automata (ANN, SVM, EAs, etc) pe caracteristicile imaginilor din SA.
4. Recunoasterea cifrelor in imaginile din ST pe baza modelului anterior invatat.

**Date și bibliografie**

A se consulta directorul: Data/handWritingRecogn  
<http://yann.lecun.com/exdb/mnist/>



**PROIECTUL 10**

Precipitații

**Obiective**

Previziunea vremii

**Ideea proiectului**

Previziunea vremii (a nivelului de precipitații și/sau a temperaturii) – problemă de regresie. Au fost înregistrate date privind nivelul precipitațiilor și cel al temperaturii de-a lungul unei perioade date și se dorește previziunea vremii (temperatură și precipitații) pentru o dată ulterioară.

**Lista de *TO DO*-uri**

1. Alegerea unui set  $S$  de date (o parte din date sunt cupluri (intrari, iesiri) – pentru antrenament - SA, o parte sunt doar intrari– pentru testare - ST).
2. Invatarea unui model de predicție  $f$  (iesire =  $f$ (intrare)) prin rularea unui algoritm de invatare automata (ANN, SVM, EAs, etc) pe setul de date SA.
3. Extrapolarea modelului invatat pentru datele din ST.

**Date și bibliografie**

A se consulta directorul: data/precipitation.

&amp;

[http://www.jisao.washington.edu/data\\_sets/widmann](http://www.jisao.washington.edu/data_sets/widmann)



[http://climate.geog.udel.edu/~climate/html\\_pages/download.html#ghcn\\_T\\_P\\_clim](http://climate.geog.udel.edu/~climate/html_pages/download.html#ghcn_T_P_clim)

