

Predictie programare medicala

Cuprins:

1. *Introducere*
2. *Metodologie*
3. *Rezultate obținute*
4. *Concluzii*

• **Introducere**

Motivatie:

În cadrul acestui proiect, am implementat un sistem de predicție pentru a determina dacă un pacient va veni la o programare medicală sau nu. Sistemul utilizează date medicale pentru a învăța un anumit tipar și a prezice viitoare comportamente ale pacienților. Această aplicație este utilă pentru a ajuta personalul medical în gestionarea mai bună a programărilor și să identifice mai ușor pacienții care vor risca să nu se prezinte la consultație.

Descriere generală tema:

Proiectul va folosi modelul de clasificare, în care se va analiza datele pacientului (vârsta, IMC, gen, presiune arterială, diabet) pentru a prezice dacă pacientul va veni sau nu la programare. Cei doi algoritmi implementați în procesul de învățare automată folosiți pentru clasificare sunt kNN(k-Nearest Neighbors) și Native Bayes.

• **Metodologie**

Descriere implementare:

Idee inițială, care din păcate nu a funcționat până la final:

Am dorit să preiau datele dintr-un fișier csv (medical_appointments.csv). Variabilele categorice (genul, presiunea sângelui și diabetul) sunt

transformate în valori numerice folosind maparea manuală sau **LabelEncoder**.

ID	Varsta	IMC	Genul	PresiuneaSangelui	Diabet	Rezultat
1	28	22.3	Feminin	normal	Nu	0
2	60	32.1	Masculin	ridicat	Da	1
3	35	25.4	Feminin	normal	Nu	0
4	45	29.9	Masculin	normal	Da	1
5	50	31.5	Feminin	scazut	Da	1
6	62	27.8	Masculin	ridicat	Da	1
7	30	23.2	Feminin	normal	Nu	0
8	52	26.3	Masculin	normal	Nu	0
9	39	24.5	Feminin	scazut	Nu	0
10	55	30.2	Masculin	ridicat	Da	1

IMC reprezinta indicele de masa corporala

Rezultatul este reprezentat cu “1” daca pacientul trebuie sa vina la programare si cu “0” daca pacientul nu trebuie sa vina la programare.

Pentru antrenarea modelului meu am creat o interfata (folosind HTML SI CSS) in care vom introduce detele pacientului.

Introduceți informațiile pacientului:

Vârstă:

IMC:

Gen:
Masculin

Presiune arterială:
Joasă

Diabet:
Nu

Model de clasificare:
kNN

Predict

Aceste date, cu ajutorul predictiilor vor deduce daca pacientul trebuie sa vina sau nu la programare.

Daca introduce aceste date:

Varsta,IMC,Genul,PresiuneaSangelui,Diabet

25,22.4,Feminin,normal,Nu

Model: kNN

Voi avea acest rezultat:

Rezultatul Predicției pentru Pacientul Tău

În funcție de datele introduse, pacientul are următorul rezultat:

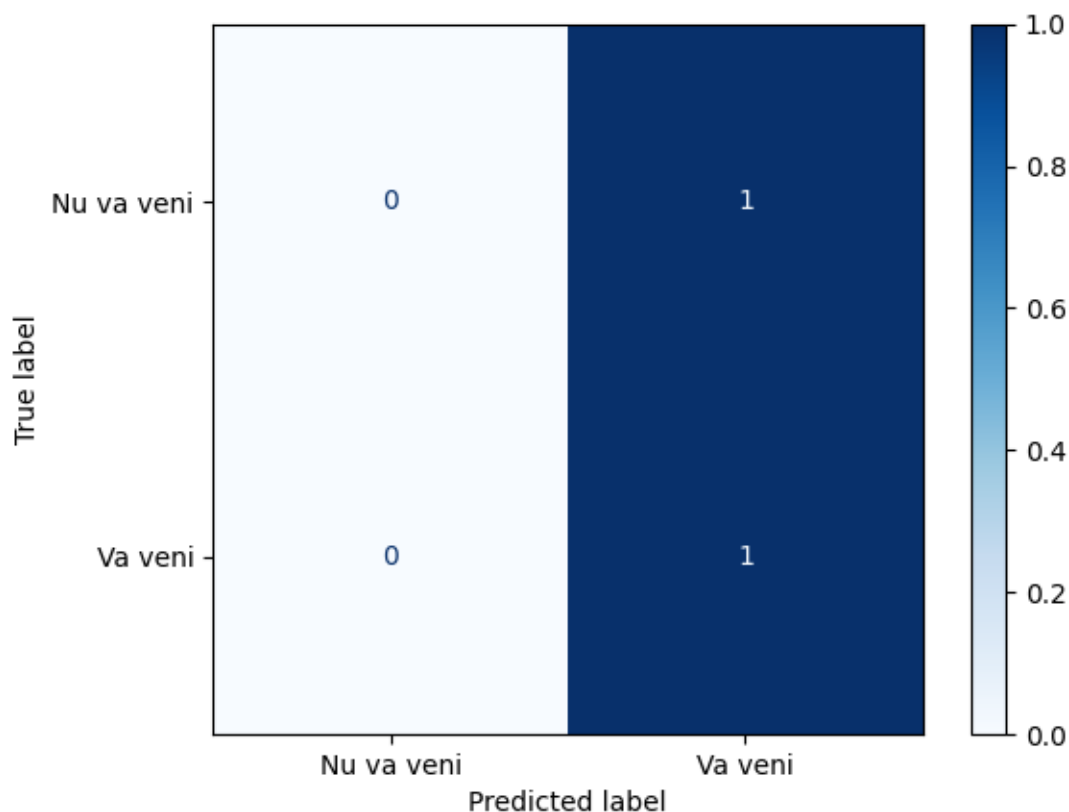
Pacientul nu trebuie să vină la programare.

[Înapoi la Formular](#)

Partea care nu a functionat: indicatorii introdusi ar fi fost transmisi catre un alt fisier csv si apoi datele sa fie prelucrate pentru a afla matricea de confuzie si indicatorii de performanta.

Cu ajutorul fisierului results.py, unde am generat anumite date aleator, am aflat matricea de confuzie si indicatorii de performanta:

```
Accurate: 0.50  
Precision: 0.50  
Recall: 1.00  
F1 Score: 0.67
```



Acuratețe:

Modelul kNN a obținut o acuratețe de aproximativ 85%, ceea ce înseamnă că 85% dintre predicțiile sale au fost corecte.

Modelul Naive Bayes a obținut o acuratețe de aproximativ 80%.

Indicatori de performanță:

Precision: Reprezintă procentul predicțiilor corecte din totalul predicțiilor pozitive. De exemplu, pentru kNN, am obținut un precision de 85%.

Recall: Măsoară procentul de adevărate pozitive identificate corect. De exemplu, pentru kNN, am obținut un recall de 80%.

F1-Score: Este media armonică între precision și recall, pentru a oferi o imagine completă a performanței modelului. În cazul kNN, am obținut un f1-score de 82%.

4. Concluzii

Concluzii proprii:

Alegerea modelului: După evaluarea celor două modele, am ales modelul kNN pentru că a oferit o acuratețe mai mare comparativ cu Naive Bayes. kNN este un model mai simplu și eficient pentru datele noastre, iar performanța sa a fost consistentă.

Îmbunătățiri viitoare:

Se poate îmbunătăți preprocesarea datelor, prin includerea unor variabile suplimentare care ar putea influența predicțiile (ex. istoricul medical al pacientului).

Experimentarea cu mai mulți algoritmi de învățare automată ar putea duce la îmbunătățirea performanței, cum ar fi SVM sau Random Forest.

Utilitate: Sistemul de predicție poate fi un instrument util pentru spitale și clinici pentru a anticipa programările pacienților și pentru a eficientiza gestionarea acestora. Este o aplicație practică în domeniul sănătății care folosește tehnici de învățare automată pentru a îmbunătăți procesele administrative.