

Optimizarea rețelelor neuronale pentru analiza riscului de diabet

Mădălina-Ioana Palade

April 10, 2025

1 Descrierea proiectului

Acest proiect se concentrează pe dezvoltarea unui model de învățare aprofundată pentru clasificarea riscului de diabet, utilizând o rețea neuronală cu un singur strat ascuns. Scopul principal este de a analiza indicatorii de sănătate și stilul de viață al pacienților pentru a prezice dacă aceștia sunt diabetici, pre-diabetici sau sănătoși, pe baza unui set de date ce include 35 de caracteristici medicale și demografice.

Modelul va implementa optimizarea neconstrânsă, aplicând cel puțin două metode de optimizare studiate la curs, precum Metoda Gradientului și Metoda Newton. Vor fi analizate performanțele fiecărei metode în termeni de timp, convergență și eficiență în cadrul antrenării modelului.

Prin utilizarea funcției de activare personalizată

$$g(z) = \cos(z) - z \quad (1)$$

, proiectul va explora comportamentele rețelei neuronale în raport cu diferitele tipuri de optimizare, în scopul de a maximiza acuratețea clasificării și a înțelege relația dintre stilul de viață și riscul de diabet.

2 Baza de date

Numele Variabilei	Rol	Tip	Descriere	Unități
ID	ID	Întreg	ID-ul pacientului	nu
Diabetes_binary	Țintă	Binare	0 = nu diabet, 1 = prediabet sau diabet	nu
HighBP	Caracteristică	Binare	0 = nu tensiune arterială mare, 1 = tensiune arterială mare	nu
HighChol	Caracteristică	Binare	0 = nu colesterol mare, 1 = colesterol mare	nu
CholCheck	Caracteristică	Binare	0 = nu test colesterol în ultimii 5 ani, 1 = test colesterol în ultimii 5 ani	nu
BMI	Caracteristică	Întreg	Indicele de masă corporală	nu
Smoker	Caracteristică	Binare	0 = nu, 1 = da	nu
Stroke	Caracteristică	Binare	0 = nu, 1 = da (accident vascular cerebral)	nu
HeartDiseaseorAttack	Caracteristică	Binare	0 = nu, 1 = da (boli cardiace sau infarct)	nu
PhysActivity	Caracteristică	Binare	0 = nu, 1 = da (activitate fizică în ultimele 30 de zile)	nu
Fruits	Caracteristică	Binare	0 = nu, 1 = da (consum fructe 1 sau mai multe ori pe zi)	nu
Veggies	Caracteristică	Binare	0 = nu, 1 = da (consum legume 1 sau mai multe ori pe zi)	nu
HvyAlcoholConsump	Caracteristică	Binare	0 = nu, 1 = da (consum excesiv alcool)	nu
AnyHealthcare	Caracteristică	Binare	0 = nu, 1 = da (acoperire de asigurare de sănătate)	nu
NoDocbcCost	Caracteristică	Binare	0 = nu, 1 = da (nu s-a putut consulta un medic din cauza costurilor)	nu
GenHlth	Caracteristică	Întreg	1 = excelent, 2 = foarte bun, 3 = bun, 4 = corect, 5 = slab	nu
MentHlth	Caracteristică	Întreg	Număr de zile cu sănătate mentală proastă în ultimele 30 de zile	nu
PhysHlth	Caracteristică	Întreg	Număr de zile cu sănătate fizică proastă în ultimele 30 de zile	nu
DiffWalk	Caracteristică	Binare	0 = nu, 1 = da (dificultăți mari la mers sau urcat scări)	nu

Numele Variabilei	Rol	Tip	Descriere	Unități
Sex	Caracteristică	Binare	0 = femeie, 1 = bărbat	nu
Age	Caracteristică	Întreg	Categorie de vârstă (13 nivele)	nu
Education	Caracteristică	Întreg	Nivel de educație (scale 1-6)	nu
Income	Caracteristică	Întreg	Venit (scale 1-8)	nu

3 Results

Present the results or findings of your work. Include any relevant tables, graphs, or figures.

4 Conclusion

Summarize your work and discuss any conclusions or insights you gained.