

Training Neural Nets with Congestion Control

Nume: Stoian Robert Alin

Grupa: 464

1. Articole existente si contextul "Arheologic"

Proiectul isi propune sa testeze o ipoteza inedita: aplicarea algoritmului AIMD (Additive Increase Multiplicative Decrease), fundamental pentru stabilitatea internetului (protocolul TCP), in antrenarea retelelor neuronale.

- **Referinta Arheologica:** Algoritmul AIMD (Chiu & Jain, 1989) garanteaza convergenta si echitatea in retelele de calculatoare prin cresterea liniara a ratei de transmisie si scaderea multiplicativa la congestie
- **Aplicabilitate AI:** Vom adapta acest mecanism pentru a ajusta dinamic Rata de Invatare (Learning Rate). Consideram "cresterea loss-ului" ca fiind echivalentul unei "congestii", declansand o scadere brusca a ratei de invatare pentru a stabiliza antrenarea

Exemplu pentru a vizualiza functionarea algoritmului, putem compara procesul de antrenare cu un vehicul pe o autostrada . Analogia Traficului :

- Additive Increase (Accelerare Prudenta): Cat timp drumul este liber (eroarea scade), soferul creste viteza treptat, liniar, pentru a ajunge la destinatie cat mai repede, fara a risca.
- Multiplicative Decrease (Franare de Urgenta): In momentul in care apare un obstacol sau o aglomerare (eroarea creste brusc), soferul nu doar ia piciorul de pe acceleratie, ci franeaza drastic, reducand viteza la jumatate.

2. Datele folosite si Analiza Exploratorie (EDA)

Vom utiliza setul de date **CIFAR-10**, un benchmark standard în Computer Vision.

- **Descriere:** Setul contine 60.000 de imagini color (32x32 pixeli), impartite echilibrat in 10 clase (avioane, masini, animale, etc.)
- **Analiza:** Imaginele au o rezolutie scazuta, ceea ce face task-ul dificil pentru modele simple, necesitand straturi convolutionale. Mai jos sunt prezentate esantioane aleatoare din setul de antrenare:



In concluzie , datele sunt corecte (echilibrate), dar sunt grele (neclare), asa ca am nevoie de un model mai avansat (CNN) ca sa le rezolve .

3.Ce modele am folosit si necesarul de calcul

Pentru a izola efectul algoritmului AIMD, folosim o arhitectura standard, usor de reprodus:

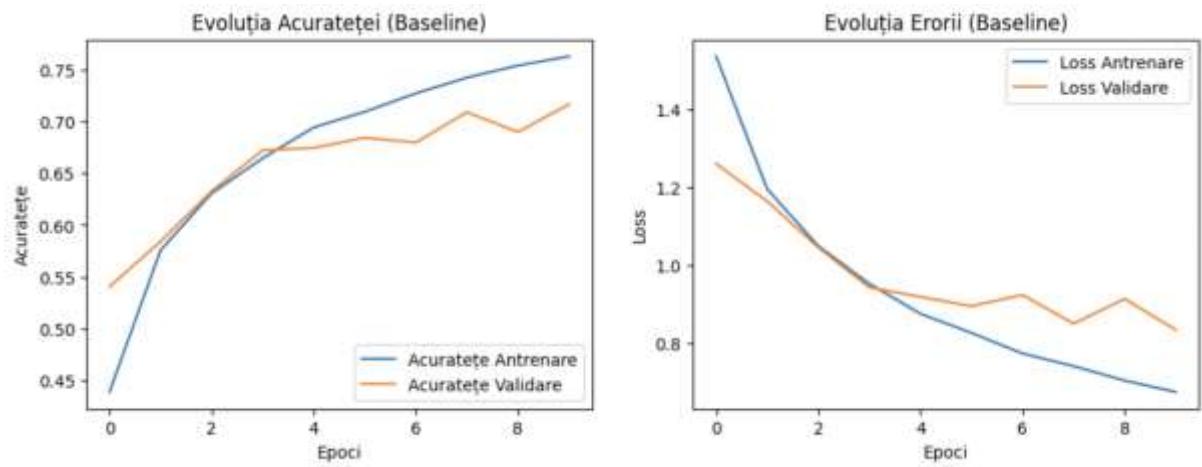
- Model: O retea neuronala convolutionala , Convolutional Neural Network (CNN) cu 3 straturi de convolutie (Conv2D), straturi de MaxPooling pentru reducerea dimensiunii si straturi Dense la final pentru clasificare
- Resurse: Antrenarea necesita accelerare GPU. Folosim mediul Google Colab (T4 GPU), unde un ciclu de antrenare pentru baseline dureaza sub 2 minute, ceea ce permite iterarea rapida a experimentelor

4.Ce metode de evaluare si comparatie se folosesc

Metodologia de evaluare este comparativa. Se va antrena acelasi model CNN in trei scenarii:

1. Baseline (Realizat): Learning Rate fix/standard.
2. Benchmark: Scheduler modern (ReduceLROnPlateau).
3. Experimental: Scheduler-ul propus AIMD.

Rezultate Intermediare (Baseline): Am implementat si rulat cu succes modelul Baseline. Asa cum se observa in graficele de mai jos, modelul converge corect, atingand o acuratete satisfacatoare in primele epoci, fara semne majore de overfitting. Aceasta este linia de baza pe care algoritmul AIMD trebuie sa o depaseasca.



5. Referinte Bibliografice

1. Sursa Arheologica (Fundamentul AIMD): Chiu, D. M., & Jain, R. (1989). Aceasta este studiul fundamental care a demonstrat matematic de ce algoritmul AIMD asigura stabilitatea in retele, concept pe care l-am adaptat in proiect. Link: https://www.cs.wustl.edu/~jain/papers/ftp/cong_av.pdf
2. Context Modern (Deep Learning): Smith, L. N. (2017). Cyclical Learning Rates for Training Neural Networks. arXiv preprint. Lucrare de referinta pentru metodele dinamice de ajustare a ratei de invatare, care serveste drept context pentru experimentul nostru. Link: <https://arxiv.org/abs/1506.01186>
3. Setul de Date: Krizhevsky, A. (2009). Learning Multiple Layers of Features from Tiny Images (CIFAR-10 Dataset). University of Toronto. Sursa oficiala a setului de date utilizat pentru antrenarea si validarea modelului. Link: <https://www.cs.toronto.edu/~kriz/cifar.html>