Rata de invatare pe setul de invatare si pe setul de test au fost intotdeauna asemanatoare, asa ca am trecut o singura valoare.

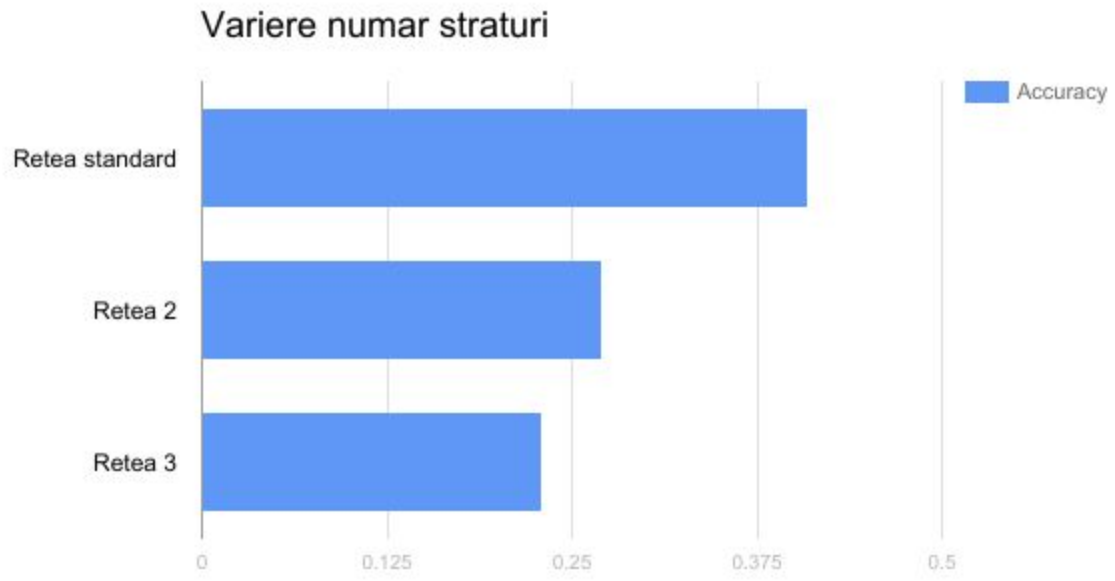
RETELE LINIARE

RETEA STANDARD

LinearizeLayer(), Layer(300), TanhLayer(), Layer(10) - **0.41**

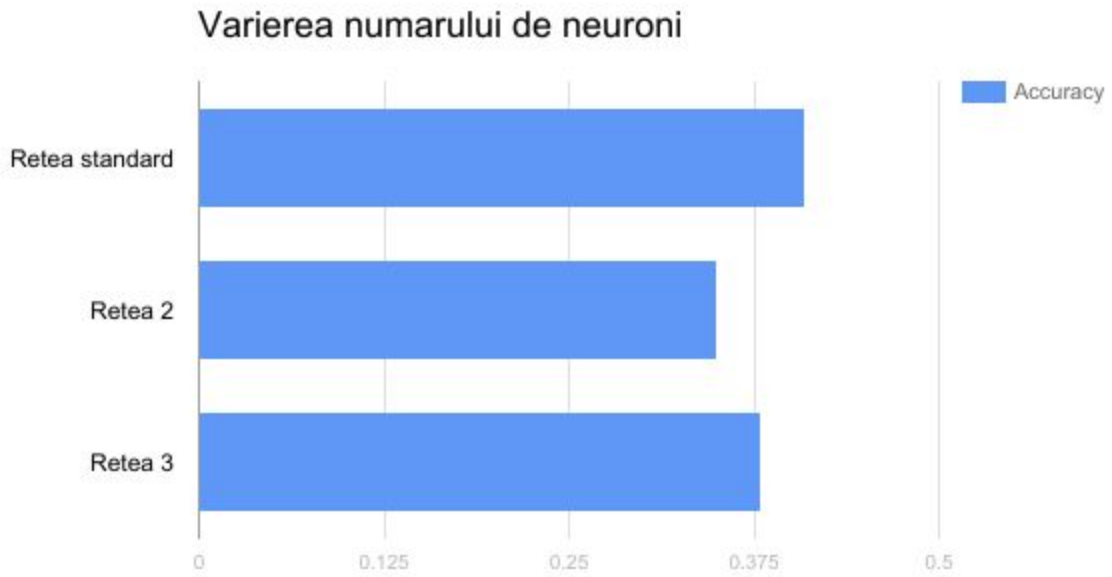
VARIERE STRUCTURA

Retea 2 - LinearizeLayer(), Layer(300), TanhLayer(), Layer(10), SoftMaxLayer() - **0.27**
Retea 3 - LinearizeLayer(), TanhLayer(), Layer(300), TanhLayer(), Layer(10), SoftMaxLayer() - **0.23**



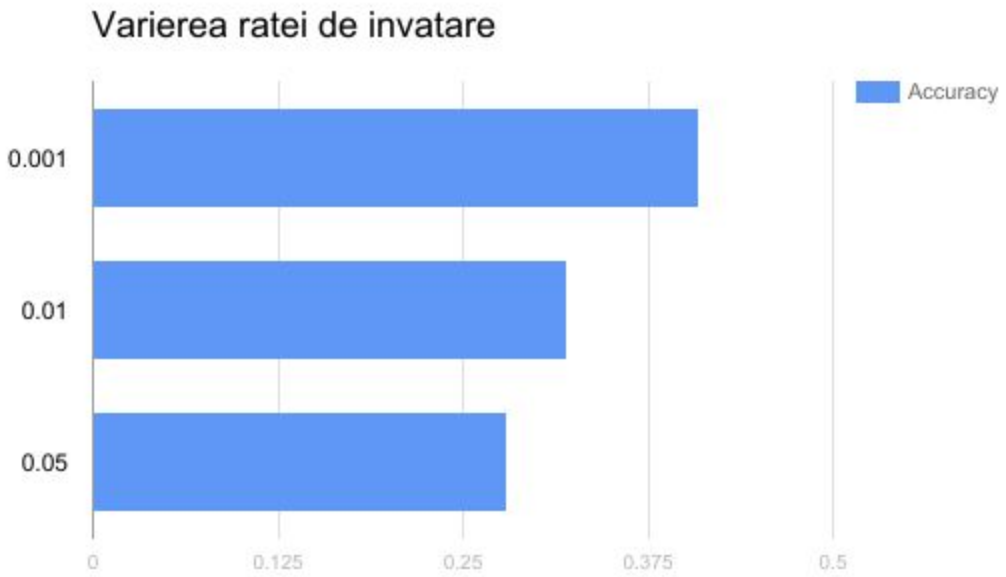
VARIERE NUMAR NEURONI

Retea 2 - LinearizeLayer(), Layer(200), TanhLayer(), Layer(10) - **0.35**
Retea 3 - LinearizeLayer(), Layer(400), TanhLayer(), Layer(10) - **0.38**



VARIERE LEARNING RATE

Retea 2 - LinearizeLayer(), Layer(300), TanhLayer(), Layer(10), LR = 0.01 - **0.32**
Retea 3 - LinearizeLayer(), Layer(300), TanhLayer(), Layer(10), LR = 0.05 - **0.28**



Am putut observa ca o rata de invatare isi atinge maximul mai repede, dar apoi descreste, intrucat deplasarea in directia scaderii gradientului are un pas foarte mare.

Retele convolutionale

RETEA STANDARD:

```
nn = FeedForward([
    ConvolutionalLayer(3, 32, 32, 6, 5, 1),
    MaxPoolingLayer(2),
    ReluLayer(),
    ConvolutionalLayer(6, 14, 14, 16, 5, 1),
    MaxPoolingLayer(2),
    ReluLayer(),
    LinearizeLayer(16, 5, 5),
    Layer(400, 300, relu),
    Layer(300, 10, identity),
    SoftmaxLayer()])
```

LEARNING RATE - 0.001

0.46

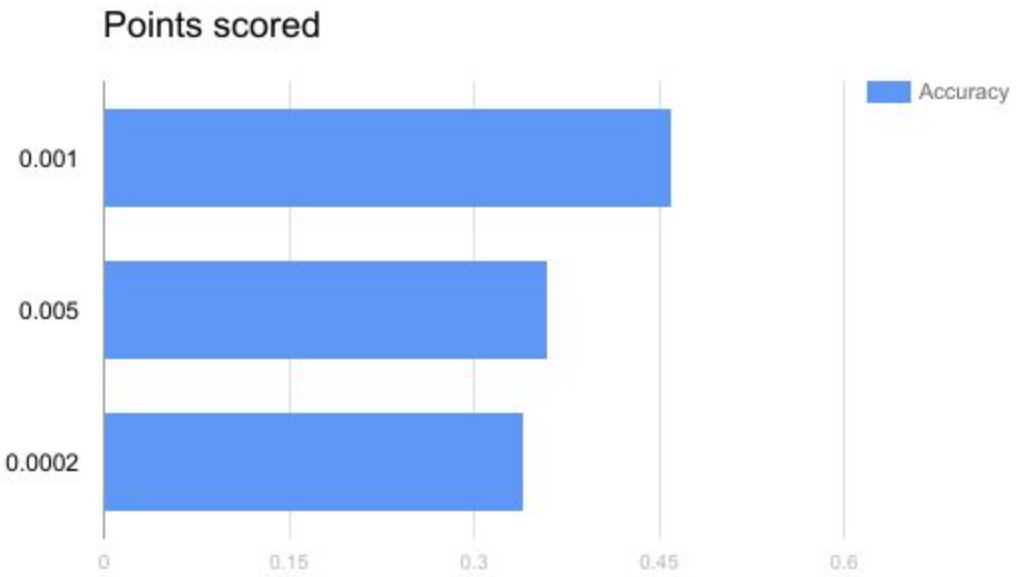
Variere learning rate

Lr = 0.005

0.36 - **CRESTERE MULT MAI RAPIDA**

Lr = 0.0002

0.34



Variere stride

Stride 2 ->

```
nn = FeedForward([
        ConvolutionalLayer(3, 32, 32, 6, 6, 2),
        ReluLayer(),
        ConvolutionalLayer(6, 14, 14, 16, 4, 2),
        ReluLayer(),
        LinearizeLayer(16, 6, 6),
        Layer(576, 300, relu),
        Layer(300, 10, identity),
        SoftmaxLayer()])
```

Accuracy - 0.43

Variere filter size

Filter size 7*7 ->

```
nn = FeedForward([
        ConvolutionalLayer(3, 32, 32, 6, 7, 1),
        ReluLayer(),
        ConvolutionalLayer(6, 26, 26, 16, 7, 1),
        ReluLayer(),
        ConvolutionalLayer(16, 20, 20, 20, 7, 1),
        MaxPoolingLayer(2),
        ReluLayer(),
        LinearizeLayer(20, 7, 7),
        FullyConnected(980, 100, logistic),
        FullyConnected(100, 10, identity),
        SoftMax()])
```

Accuracy - 0.31

Variere momentum

Folosirea impulsului nu a afectat acuratetea finala a retelei, dar a scazut considerabil timpul de atingere al maximului.