# Dropout әдісінің нейрондық желінің өнімділігіне әсері (MNIST мысалында)

## 1. Жұмыстың мақсаты

Бұл жұмыстың мақсаты — Dropout және басқа да регуляризация әдістерінің (L2, Max-Norm) нейрондық желілердің артық үйренуін (overfitting) азайтудағы тиімділігін зерттеу. Эксперимент Srivastava et al., 2014 мақаласындағы идеяға сүйенеді, бірақ жеңілдетілген архитектурада (шағын модель және CPU-де орындалатын нұсқа) жүзеге асырылды.

## 2. Модель архитектурасы

Барлық тәжірибелерде келесі қарапайым толық байланысқан нейрондық желі қолданылды:  
• Кіріс қабаты: 784 (28×28 пиксель)  
• 1-жасырын қабат: 512 нейрон, ReLU  
• 2-жасырын қабат: 512 нейрон, ReLU  
• Шығыс қабаты: 10 нейрон (Softmax)  
  
Оптимизатор: SGD (momentum = 0.9)  
Регуляризация: Dropout(p), L2 (weight decay), Max-Norm  
Фреймворк: PyTorch  
Деректер жиыны: MNIST (60 000 сурет – оқыту, 10 000 – тест)

## 3. Гиперпараметрлерді таңдау

Төрт түрлі оқу жылдамдығы тексерілді: lr = [0.001, 0.01, 0.05, 0.1]. Нәтижесінде ең тиімді learning rate = 0.01 екені анықталды.

## 4. Регуляризация әдістерін салыстыру

|  |  |  |  |
| --- | --- | --- | --- |
| Әдіс | Dropout | L2 | Max-Norm |
| L2 | Жоқ | Иә | Жоқ |
| Dropout | Иә | Жоқ | Жоқ |
| Dropout + L2 | Иә | Иә | Жоқ |
| Dropout + MaxNorm | Иә | Жоқ | Иә |

Нәтижелер бойынша Dropout + L2 комбинациясы ең жоғары дәлдік көрсетті (≈98%).

## 5. Dropout коэффициентінің әсері

Dropout параметрі p = [0.0, 0.2, 0.5, 0.8] мәндері бойынша тексерілді. Оптималды нәтиже p = 0.2 кезінде байқалды.

## 6. Оқыту жиынының өлшемінің әсері

Оқыту деректерінің көлемі өзгертілгенде Dropout әдісінің әсері айқын көрінді. 60 000 мысал кезінде ең жоғары дәлдік 98.1% болды.

## 7. Финалдық модель

Финалдық модель параметрлері:  
• Dropout (p): 0.2  
• Learning rate: 0.01  
• Momentum: 0.9  
• L2 weight decay: 1e-4  
• Эпох саны: 10  
  
Финалдық дәлдік: ≈98.2%, оқыту уақыты ≈150 секунд (CPU).

## 8. Қорытынды

1. Dropout + L2 комбинациясы ең тұрақты нәтиже берді.  
2. Оптималды p мәні 0.2–0.5 аралығында.  
3. Оқыту деректері көбейген сайын Dropout әсері күшейеді.  
4. Алынған нәтиже Srivastava et al. (2014) мақаласындағы нәтижелерге ұқсас.