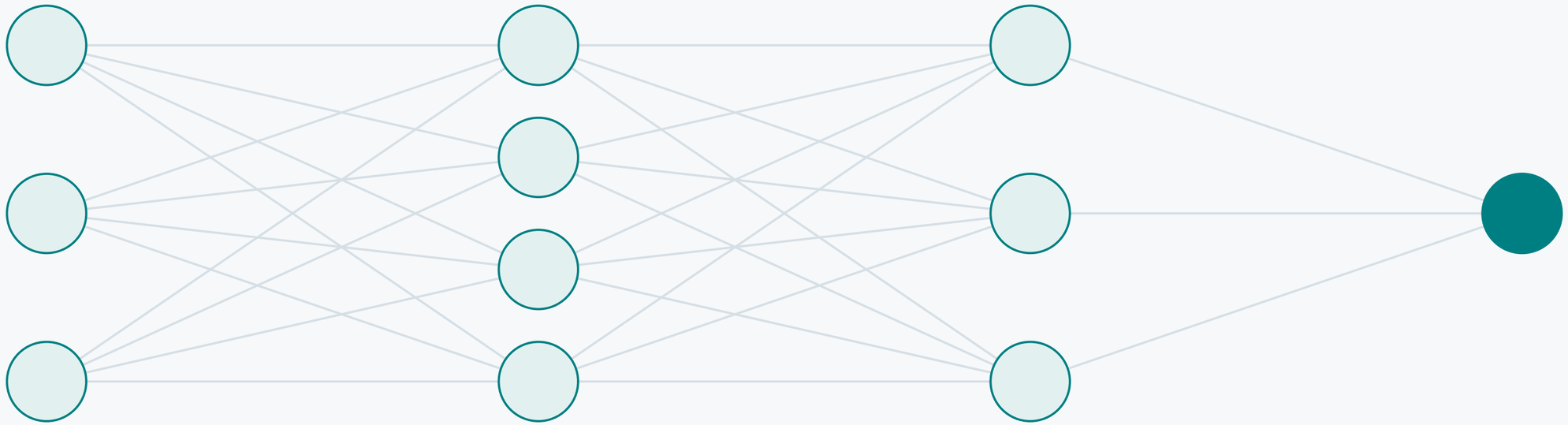
Көпқабатты персептрон

Нейрондық желінің архитектурасы

Қабаттар • салмақ матрицалары • ақпараттың алға таралуы



$$\mathbf{x} \rightarrow \mathbf{a}^{(1)} \rightarrow \mathbf{a}^{(2)} \rightarrow \hat{\mathbf{y}}$$

Дәрістің мақсаты және оқу нәтижелері

01 Сипаттау

MLP құрылымын және қабаттардың қызметін түсіндіру.

02 Жазу

Нейрон мен қабат теңдеулерін скалярлық және матрицалық түрде жазу.

03 Есептеу

Параметрлер санын және берілген желінің шығысын есептеу.

04 Негіздеу

Активация мен шығыс қабатын міндетке сәйкес таңдау.

Негізгі сұрақ: бірнеше нейрон бір модель ретінде қалай есептейді?

Еске түсіру: бір нейрон

$$z = w^{T}x + b \quad a = f(z)$$

z — активацияға дейін

Кірістердің салмақталған қосындысы және ығысу.

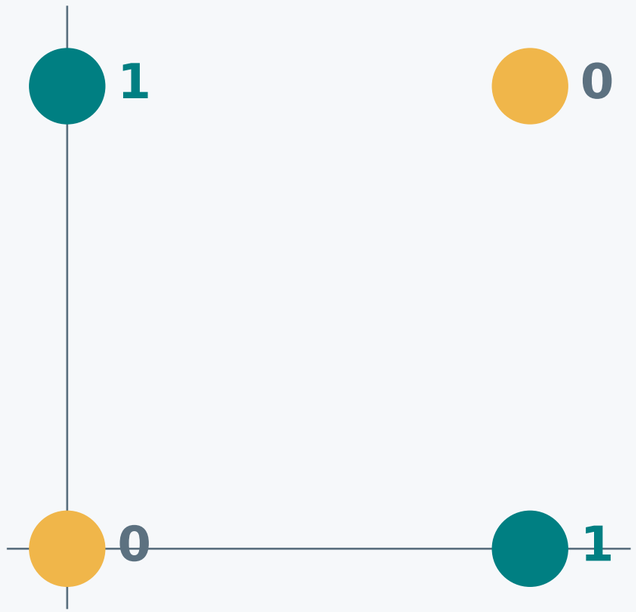
a — активациядан кейін

Осы нейронның шығысы; келесі қабаттың кірісіне айналады.

Көпқабатты желіде: $a^{(0)} = x$, ал соңғы активация $a^{(L)} = \hat{y}$.

Неліктен бір нейрон жеткіліксіз?

XOR: қарама-қарсы төбелер бір класқа жатады.



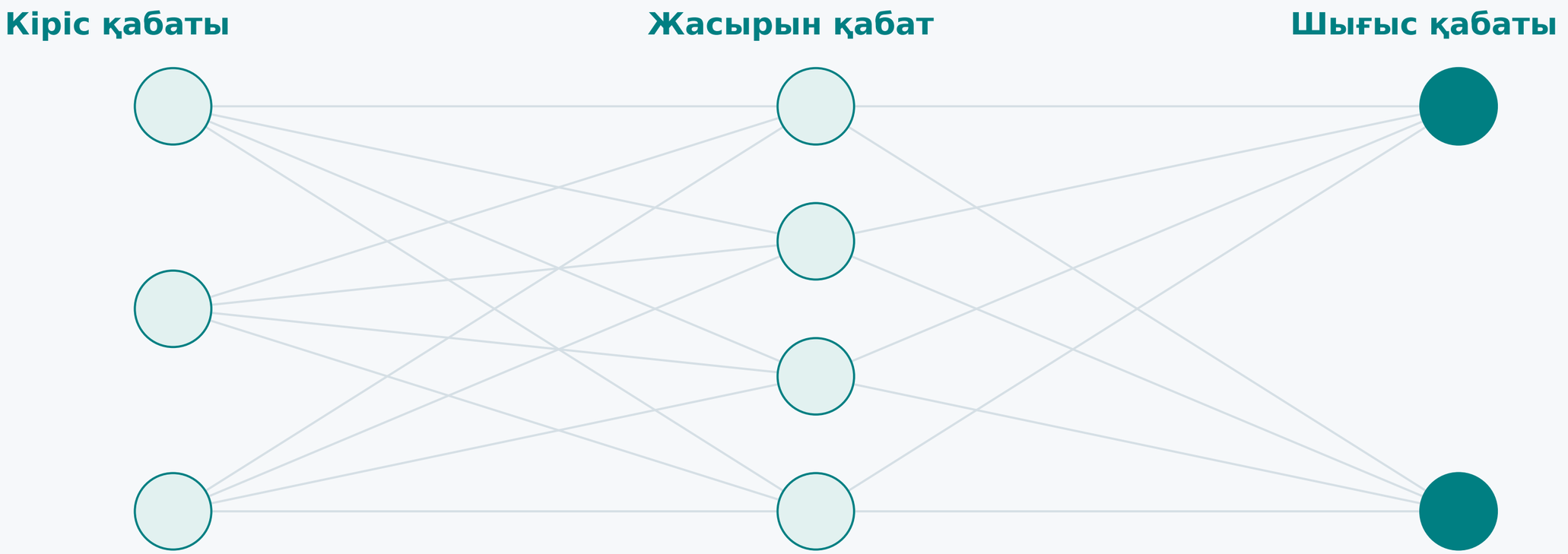
Бір түзу XOR кластарын ажыратпайды.

Жасырын қабаттар бастапқы белгілердің бейсызық комбинацияларын құрады.

Бұл күрделірек шешім шекарасын қалыптастыруға мүмкіндік береді.

Желінің мүмкіндігін қабаттармен бірге бейсызық активациялар арттырады.

MLP: тікелей таралатын желі



MLP — Multilayer Perceptron; ақпарат қалыпты есептеуде кірістен шығысқа өтеді.

Үш типтегі қабаттың қызметі

Кіріс қабаты

Бастапқы белгілерді қабылдайды.

Мысал: датчиктің үш өлшемі.

$a^{(0)} = x$

Жасырын қабат

Аралық көріністерді қалыптастырады.

Мәндері желі ішінде есептеледі.

$a^{(l)} = f^{(l)}(z^{(l)})$

Шығыс қабаты

Міндетке сәйкес болжам береді.

Сан, класс ұпайы немесе ықтималдық бағасы.

$\hat{y} = a^{(L)}$

Шығыс өлшемі мен функциясы шешілетін міндетке байланысты.

Архитектураны қалай оқимыз?

4 → 8 → 4 → 1

4 кіріс

Белгілер саны

8 нейрон

1-жасырын қабат

4 нейрон

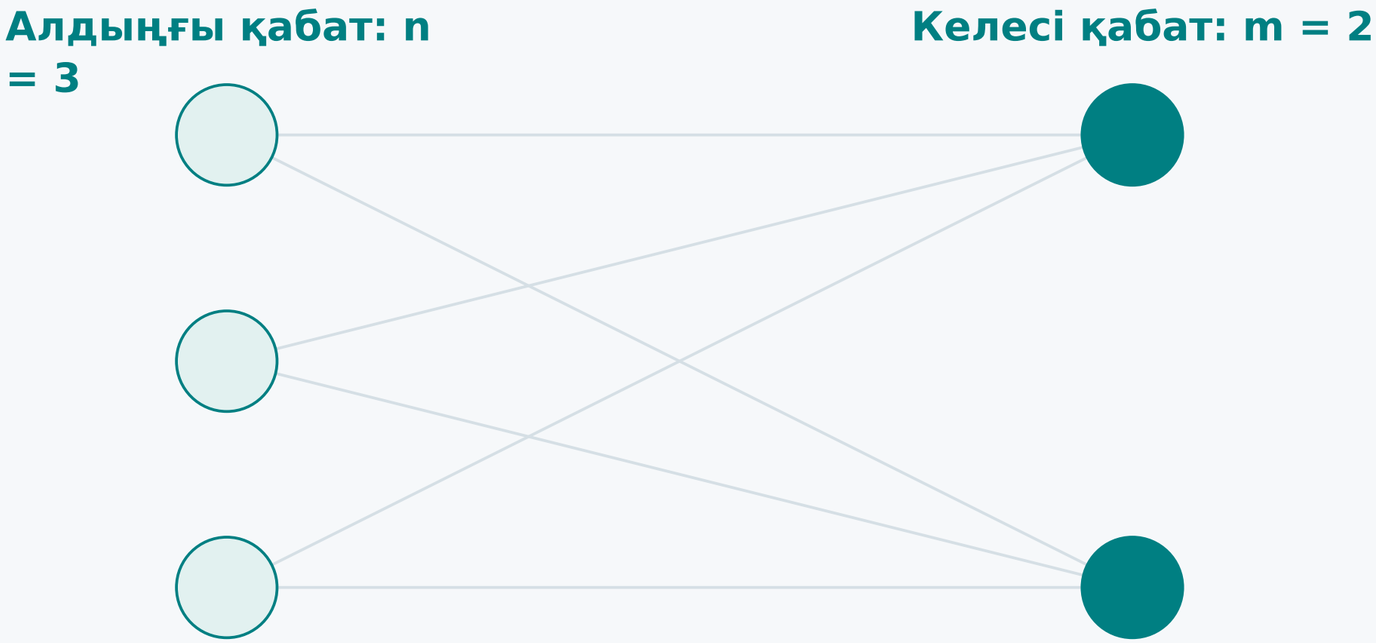
2-жасырын қабат

1 нейрон

Шығыс қабаты

Осы белгілеуде: 2 жасырын қабат және 3 есептеу қабаты (L = 3).

Толық байланысқан қабат — Dense



Әр кіріс әр шығыс нейронмен байланысады.

Салмақтар: 3 × 2 = 6

Ығысулар: 2

Барлығы: 8 параметр

Жалпы жағдайда: салмақтар саны n × m, ығысулар саны m.

Салмақ индексінің мағынасы

$$w_{ji}^{(l)} : (l-1)\text{-қабаттағы } i \rightarrow l\text{-қабаттағы } j$$

Төменгі индекстер

j — ағымдағы нейрон
i — алдыңғы қабаттың нейроны

Жоғарғы индекс

l — қабаттың нөмірі.
Бұл дәреже көрсеткіші емес.

Мысал: $w_{21}^{(1)}$ — x_1 кірісінен 1-жасырын қабаттың 2-нейронына баратын салмақ.

Қабаттағы бір нейронның теңдеуі

$$z_j^{(l)} = \Sigma_i\, w_{ji}^{(l)}\, a_i^{(l-1)} + b_j^{(l)}$$

$$a_j^{(l)} = f^{(l)}(z_j^{(l)})$$

$a_i^{(l-1)}$ — алдыңғы қабаттың шығысы
$b_j^{(l)}$ — ағымдағы нейронның ығысуы
$a_j^{(l)}$ — келесі қабатқа берілетін жаңа мән

Бүкіл қабатты бірден есептеу

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)}$$

$$a^{(l)} = f^{(l)}(z^{(l)})$$

W: m × n $a^{(l-1)}$: n × 1 b: m × 1

Нәтиже: z және $a^{(l)}$ векторларының өлшемі m × 1.

Өлшемдерді тексеру — формула мен кодтағы қателерді табудың алғашқы қадамы.

Қабаттағы параметрлер саны

$$N = n \times m + m = (n + 1)m$$

n × m салмақ

n кірістің әрқайсысы m нейронмен байланысады.

m ығысу

Әр шығыс нейронға бір жеке bias параметрі беріледі.

Кіріс түйіндерінің өзінде бұл модельде оқытылатын параметр жоқ.

Мысал: 4 → 8 → 4 → 1 желісі

4 → 8	4 × 8 + 8	**40**
8 → 4	8 × 4 + 4	**36**
4 → 1	4 × 1 + 1	**5**

Nжалпы = 40 + 36 + 5 = 81

Параметрлер және архитектура таңдауы

Оқытылатын параметрлер

Салмақ матрицалары $W^{(l)}$
Ығысу векторлары $b^{(l)}$

Деректер негізінде оқыту барысында жаңартылады.

Архитектуралық таңдаулар

Жасырын қабаттар саны
Қабаттардың ені
Активация функциялары

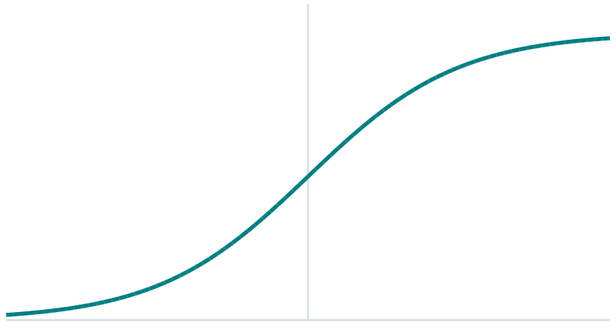
Кіріс пен шығыс өлшемі міндетке сәйкес анықталады.

Архитектураны таңдау мен берілген архитектураның параметрлерін оқытуды ажыратыңыз.

Активация функцияларын таңдау

Sigmoid

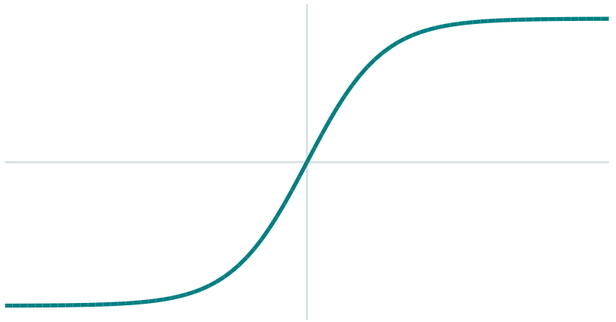
1 / (1 + exp(−z))



Шығыс: (0; 1)

Tanh

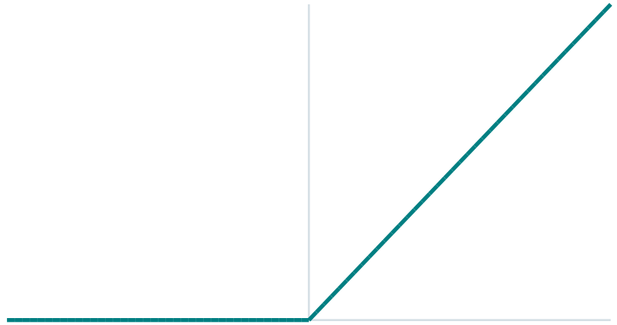
tanh(z)



Шығыс: (−1; 1)

ReLU

max(0, z)



Шығыс: [0; +∞)

Жасырын және шығыс қабаттарында әртүрлі активациялар қолданылуы мүмкін.

Бейсызықтық болмаса не болады?

$$h = W_1x + b_1 \quad \hat{y} = W_2h + b_2$$

$$\hat{y} = (W_2W_1)x + (W_2b_1 + b_2)$$

Бірнеше аффиндік қабат бір аффиндік түрлендіруге тең.

Күрделі бейсызық тәуелділіктер үшін қабаттар арасында бейсызық активация қажет.

Шығыс қабаты міндетке сәйкес таңдалады

Сандық регрессия

Мысал: температура.

Бір шығыс нейроны.

Жиі: $\hat{y} = z$
(тепе-теңдік функциясы).

Екілік жіктеу

Мысал: ақау / қалыпты.

Бір шығыс нейроны.

Sigmoid: $0 < \hat{y} < 1$.

K кластың біреуі

Мысал: 0–9 цифрлары.

K шығыс нейроны.

Softmax: оң мәндер,
қосындысы 1.

Sigmoid/Softmax нәтижесін ықтималдық бағасы ретінде қолдану тиісті оқытуды талап етеді.

Ақпараттың алға таралуы

Кіріс → жасырын қабаттар → болжам

$$a^{(0)} = x$$

$$z^{(l)} = W^{(l)}a^{(l-1)} + b^{(l)}$$

Әр қабатта: $a^{(l)} = f^{(l)}(z^{(l)})$, $l = 1, \ldots, L$

Forward pass — берілген параметрлер бойынша есептеу; параметрлер бұл кезеңде өзгермейді.

Сандық мысал: 2 → 2 → 1

Кіріс және құрылым

$x = (0{,}5;\ 1{,}0)^T$

2 кіріс → 2 жасырын → 1 шығыс

Жасырын: ReLU
Шығыс: Sigmoid

Жасырын қабат

1-нейрон:
$w = (0{,}4;\ 0{,}6)$, $b = 0{,}1$

2-нейрон:
$w = (-0{,}2;\ 0{,}8)$, $b = 0$

Шығыс нейроны

$v = (0{,}5;\ -0{,}4)$

$b_{шығыс} = 0{,}2$

Табу керек: $\hat{y} = ?$

Бұл — параметрлері алдын ала берілген желінің бір forward pass есебі.

1-кезең: жасырын қабатты есептеу

1-нейрон

$z_1 = 0{,}4 \times 0{,}5 + 0{,}6 \times 1{,}0 + 0{,}1$
$z_1 = 0{,}9$

$a_1 = \mathrm{ReLU}(0{,}9) = 0{,}9$

2-нейрон

$z_2 = -0{,}2 \times 0{,}5 + 0{,}8 \times 1{,}0 + 0$
$z_2 = 0{,}7$

$a_2 = \mathrm{ReLU}(0{,}7) = 0{,}7$

$$a^{(1)} = (0{,}9;\ 0{,}7)^T$$

2-кезең: шығыс қабатын есептеу

$$z_{\text{шығыс}} = 0{,}5 \times 0{,}9 - 0{,}4 \times 0{,}7 + 0{,}2 = 0{,}37$$

$$\hat{y} = 1 / (1 + \exp(-0{,}37)) \approx 0{,}591$$

Кіріс: (0,5; 1,0) → аралық көрініс: (0,9; 0,7) → 0,591

Егер шешім шегі 0,5 деп алынса, болжам 1-класқа жатқызылады.

Осы есептің NumPy арқылы орындалуы

```
import numpy as np

x = np.array([0.5, 1.0])
W1 = np.array([[0.4, 0.6], [-0.2, 0.8]])
b1 = np.array([0.1, 0.0])
W2 = np.array([[0.5, -0.4]])
b2 = np.array([0.2])

a1 = np.maximum(0, W1 @ x + b1)
z2 = W2 @ a1 + b2
y_hat = 1 / (1 + np.exp(-z2))
print(a1, y_hat) # [0.9 0.7] [0.59145898]
```

@ — матрицалық көбейту

maximum — ReLU

exp — экспонента

W және b өзгермейді.

Архитектураны таңдау: мүмкіндік және шектеу

Таңдау факторлары

Міндет пен деректер көлемі
Белгілердің сипаты
Қабаттар саны мен ені
Валидациядағы сапа
Жад пен есептеу уақыты

MLP шектеулері

Dense қабаттарда параметрлер тез өседі.

Кескіннің кеңістіктік және тізбектің уақыттық құрылымы арнайы ескерілмейді.

Күрделілік артық үйрену қаупін арттыруы мүмкін.

Қарапайым базалық модельден бастап, күрделілікті дәлелді түрде арттырыңыз.

Өзіңізді тексеріңіз

01 Кіріс, жасырын және шығыс қабаттарының қызметі қандай?

02 $w_{ji}^{(l)}$ индексіндегі i, j және l нені білдіреді?

03 Dense қабаты үшін параметрлер саны қалай есептеледі?

04 Неліктен бірнеше аффиндік қабат жеткіліксіз?

05 Forward pass кезінде W және b өзгере ме?

Шағын тапсырма

3 → 4 → 1 желісінде, әр есептеу қабатында bias болса,
барлығы қанша оқытылатын параметр бар?

Тексеру, бекіту және келесі дәріс

$$N = (3 \times 4 + 4) + (4 \times 1 + 1) = 21$$

MLP бірнеше қабат арқылы жаңа көріністер қалыптастырады.
Әр қабат: аффиндік түрлендіру + активация.
Архитектура параметрлер санына және модель мүмкіндігіне әсер етеді.
Forward pass соңғы болжамды есептейді.

Келесі дәріс: нейрондық желіні оқыту және қателерді кері тарату алгоритмі.

Дереккөздер және әдістемелік толықтырулар

Негізгі дереккөздер

1. «2-дәріс. Көпқабатты персептрон және нейрондық желінің архитектурасы». Дәріс конспектісі, 11 б.
2. «Нейрондық желілерге кіріспе». Интерактивті силлабус, 2026.

Силлабуста ұсынылған әдебиеттер

• Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016.
• Chollet F. Deep Learning with Python. 2nd ed. Manning, 2021.
• Prince S. J. D. Understanding Deep Learning. MIT Press, 2023.

XOR сызбасы, өлшемдерді тексеру, Softmax туралы түсіндірме және NumPy коды — конспектіні түсіндіретін әдістемелік толықтырулар. Барлық сызбалар осы презентация үшін әзірленді.