# Нейрондық желіні оқыту, валидациялау және сапасын бағалау

## Оқыту қатесінің төмен болуы модельдің жаңа деректерде жақсы жұмыс істейтінін білдіре ме?

### Дәрістің негізгі міндеттері

- Train, validation және test жиындарының рөлін ажырату.
- Жалпылауды, underfitting және overfitting құбылыстарын талдау.
- Регрессия мен жіктеу метрикаларын есепке сәйкес таңдау.

# Негізгі мақсат — жаңа деректерге жалпылау

### Оқытудағы нәтиже

Модель 10 000 оқыту кескінін жақсы жіктейді.

Бұл бұрын көрген мысалдардағы сапаны көрсетеді.

### Жаңа деректердегі нәтиже

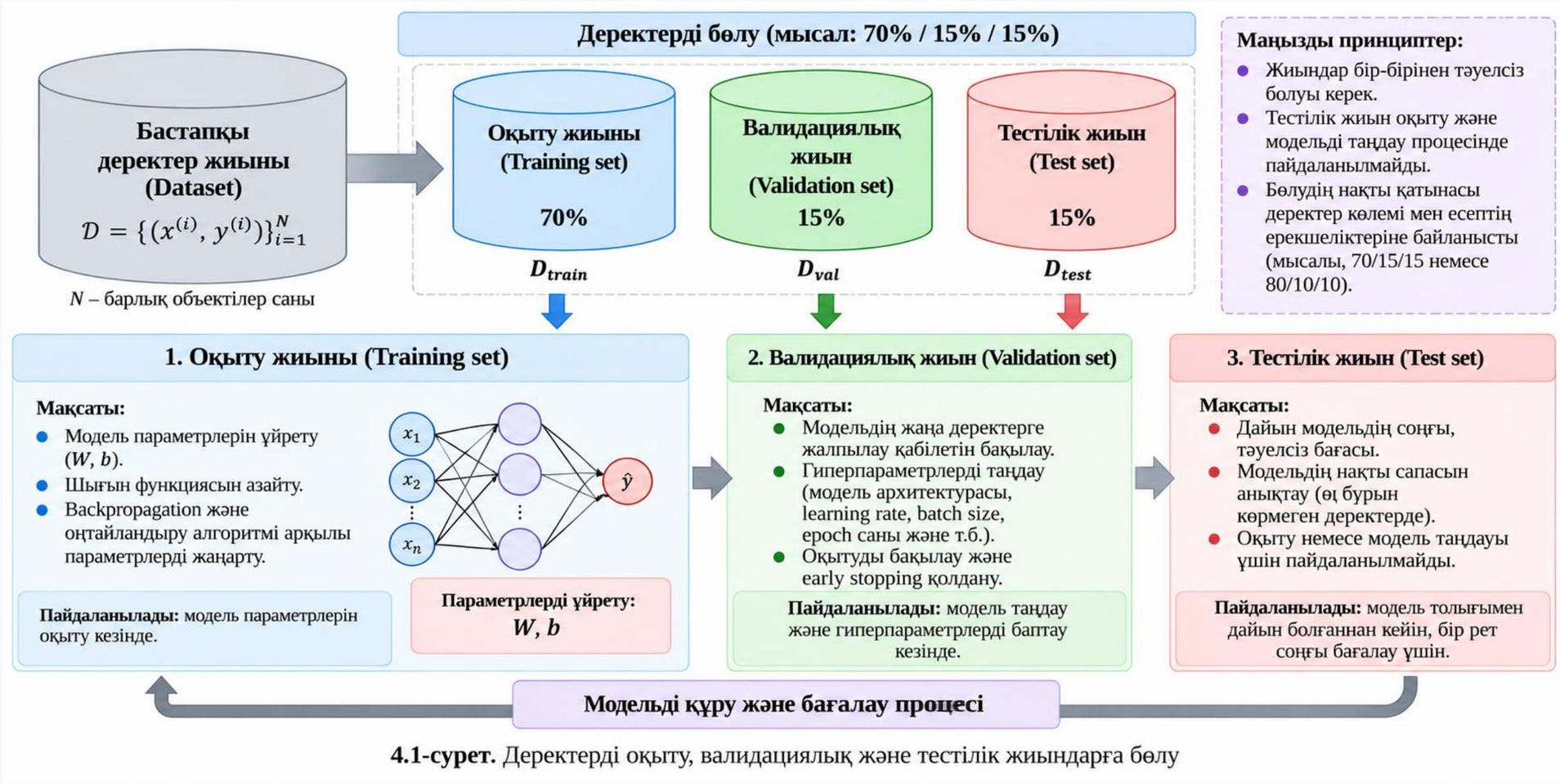
Модель жаңа кескіндерде жиі қателеседі.

Демек, жалпылау қабілеті жеткіліксіз болуы мүмкін.

**Жалпылау — оқытуға пайдаланылмаған деректерде дұрыс болжам жасау қабілеті.**

**Оқыту мен қорытынды бағалау бір деректерде жүргізілмеуі тиіс.**

# Train, validation және test жиындары




**4.1-сурет.** Деректерді оқыту, валидациялық және тестілік жиындарға бөлу

Дереккөз: 4-дәріс конспектісі, 3-бет, 4.1-сурет. Түпнұсқа.

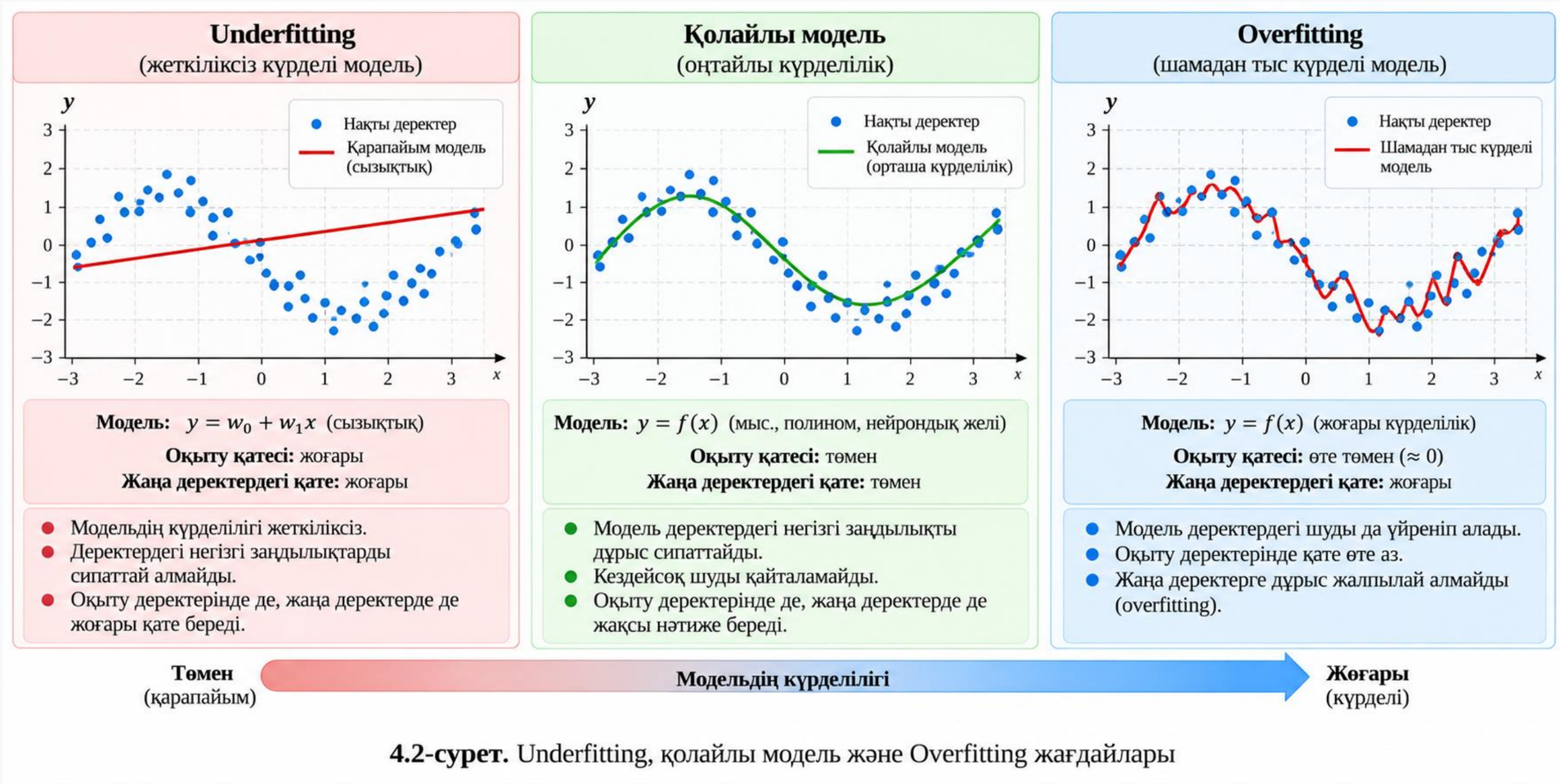
# Деректерді пайдаланудың дұрыс тәртібі

| Жиын | Негізгі қызметі | Не анықталады? |
|---|---|---|
| Train | Параметрлерді оқыту | Салмақтар W, ығысулар b |
| Validation | Баптау және модель таңдау | Архитектура, η, batch, epoch |
| Test | Дайын модельді соңғы бағалау | Тәуелсіз сапа көрсеткіштері |

Бөлу үлгілері: 70% / 15% / 15% немесе 80% / 10% / 10%.
Алдын ала өңдеу параметрлері train жиынында анықталады.

**Test нәтижесіне қарап модельді қайта баптау бағалаудың тәуелсіздігін бұзады.**

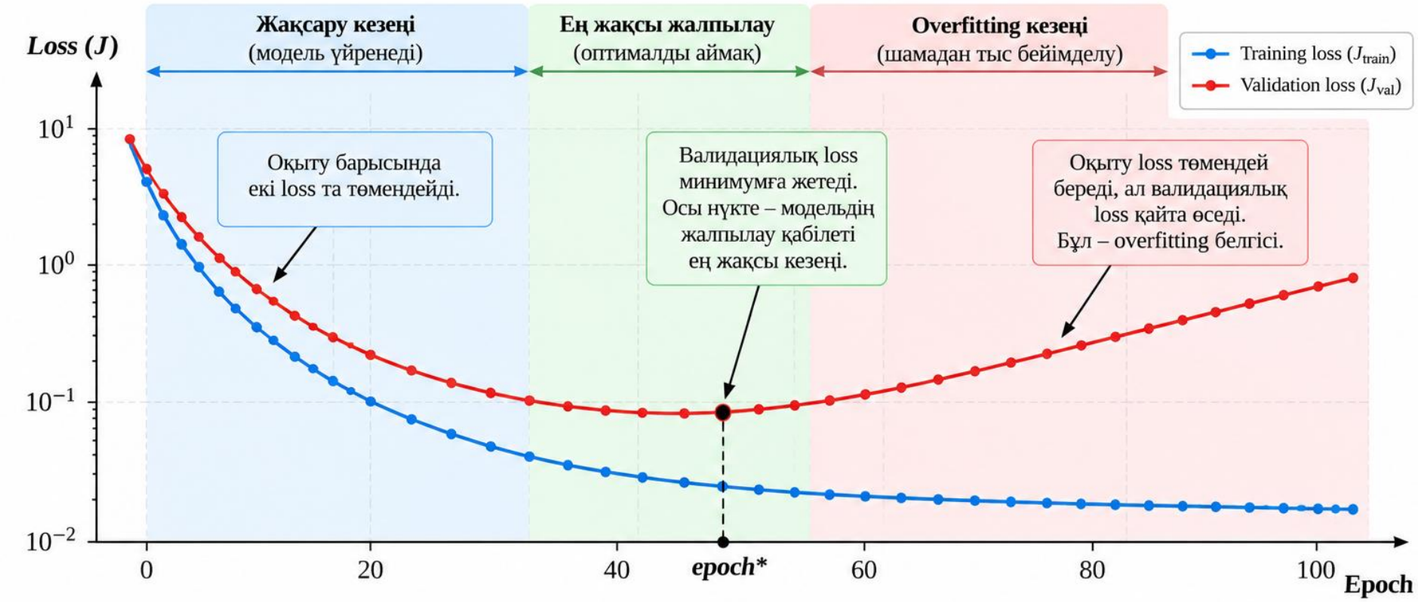
# Underfitting, қолайлы модель және overfitting




**4.2-сурет.** Underfitting, қолайлы модель және Overfitting жағдайлары

Дереккөз: 4-дәріс конспектісі, 4-бет, 4.2-сурет. Түпнұсқа.

# Training және validation loss қисықтарын талдау




**1. Жақсару кезеңі**

- Модель деректердегі заңдылықтарды үйрене бастайды.
- Training loss және Validation loss бірге төмендейді.
- Модельдің жалпылау қабілеті жақсарады.

**2. Оптималды аймақ**

- Validation loss ең кіші мәнге жетеді (epoch*).
- Бұл нүктеде модельдің жаңа деректерге жалпылау қабілеті ең жақсы.
- Практикада осы нүктедегі модель салмақтары сақталады (early stopping).

**3. Overfitting кезеңі**

- Модель оқыту деректеріне шамадан тыс бейімделе бастайды.
- Training loss төмендей береді.
- Validation loss қайта өседі.
- Жаңа деректер үшін модельдің сапасы нашарлайды.

**4.3-сурет.** Training loss және Validation loss өзгерісі арқылы overfitting-ті анықтау

Дереккөз: 4-дәріс конспектісі, 5-бет, 4.3-сурет. Түпнұсқа.

# Early Stopping: оқытуды қашан тоқтатамыз?

**Валидациялық сапа жақсармай қалғанда оқытуды тоқтату.**

| | |
|---|---|
| **1. Бақылау** | Әр epoch соңында validation loss есептеледі. |
| **2. Сақтау** | Ең жақсы көрсеткішке сәйкес модель параметрлері сақталады. |
| **3. Тоқтату** | Patience = 5: бес epoch бойы жақсару болмаса, оқыту тоқтайды. |

**Қорытынды модель ретінде ең жақсы validation нәтижесінің параметрлерін қалпына келтіреміз.**

# Шығын функциясы мен бағалау метрикасы

## Loss — оқыту үшін

Параметрлерді оңтайландыру мақсатын береді.

Мысалдар:
- регрессияда MSE;
- екілік жіктеуде binary cross-entropy.

## Metric — сапаны түсіндіру үшін

Нәтижені міндет тұрғысынан бағалайды.

Мысалдар:
- MAE, RMSE, $R^2$;
- Accuracy, Precision, Recall, F1.

**Бір көрсеткіш екі рөлде де қолданылуы мүмкін: мысалы, MSE.**

# Регрессия метрикалары: MAE, MSE және RMSE

**MAE**

$$\frac{1}{N}\sum_i |y_i - \hat{y}_i|$$

Орташа абсолюттік қате; мақсатты шаманың өлшем бірлігінде.

**MSE**

$$\frac{1}{N}\sum_i (y_i - \hat{y}_i)^2$$

Үлкен қателерге көбірек салмақ береді; өлшем бірлігі квадратталады.

**RMSE**

$$\sqrt{\text{MSE}}$$

Мақсатты шаманың өлшем бірлігінде; үлкен қателерге сезімтал.

# Детерминация коэффициенті $R^2$

$$R^2 = 1 - \Sigma_i(y_i - \hat{y}_i)^2 / \Sigma_i(y_i - \bar{y})^2$$

$\bar{y}$ — бағаланатын нақты мәндердің орташа мәні.

**$R^2 = 1$**

Қарастырылған деректерде дәл болжам.

**$R^2 = 0$**

Орташа мәнді тұрақты болжаумен бірдей квадраттық қате.

**$R^2 < 0$**

Орташа мәнді тұрақты болжаудан нашар нәтиже.

**Ескерту: нақты y мәндері тұрақты болса, формуладағы бөлім нөлге тең.**

# Регрессия: дәрістегі сандық мысал

| Нақты $y_i$ | Болжам $\hat{y}_i$ | $\lvert y_i - \hat{y}_i \rvert$ | $(y_i - \hat{y}_i)^2$ |
|---|---|---|---|
| 10 | 12 | 2 | 4 |
| 20 | 18 | 2 | 4 |
| 30 | 33 | 3 | 9 |

**MAE = 7/3 ≈ 2,33 MSE = 17/3 ≈ 5,67**

**RMSE = √(17/3) ≈ 2,38 $R^2$ = 1 − 17/200 = 0,915**

$R^2$ мәні дәрістегі осы деректер бойынша қосымша есептелді.

# Екілік жіктеу: қателер матрицасы

| | Болжам: оң | Болжам: теріс |
|---|---|---|
| Нақты: оң | TP — дұрыс оң | FN — жалған теріс |
| Нақты: теріс | FP — жалған оң | TN — дұрыс теріс |

**FP мысалы**

Қалыпты құрылғыны «ақаулы» деп белгілеу.

**FN мысалы**

Ақаулы құрылғыны өткізіп алу.

Оң класс ретінде «ақау бар» жағдайы алынған.

# Accuracy, Precision, Recall және F1

**Accuracy**

**(TP + TN) / (TP + TN + FP + FN)**

Барлық объектілер ішіндегі дұрыс болжамдар үлесі.

**Precision**

**TP / (TP + FP)**

Оң деп болжанғандардың ішіндегі шын оңдар үлесі.

**Recall**

**TP / (TP + FN)**

Барлық нақты оң объектілердің ішінен табылғандар үлесі.

**F1**

**2 × Precision × Recall / (Precision + Recall)**

Precision мен Recall-дың гармониялық ортасы.

# Неліктен Accuracy жалғыз өзі жеткіліксіз?

### Мысал 1: TP = 40, TN = 50

FP = 5, FN = 5, N = 100

Accuracy = 90%
Precision = 40/45 ≈ 0,889
Recall = 40/45 ≈ 0,889
F1 ≈ 0,889

### Мысал 2: барлық болжам — 0

950 теріс, 50 оң объект.

Accuracy = 95%, бірақ Recall = 0.
Модель бірде-бір оң объектіні таппады.

Оң болжам жоқ: Precision бөлгіші нөл.

**Метрика таңдауы класс теңгеріміне және FP/FN қателерінің салдарына тәуелді.**

# Бақылау сұрақтары

**01** Train, validation және test жиындарының қызметтері қалай ажыратылады?

**02** Параметрлер мен гиперпараметрлер қандай деректер бойынша анықталады?

**03** Loss қисықтары underfitting және overfitting туралы не көрсетеді?

**04** Patience не үшін қажет және қай checkpoint сақталуы тиіс?

**05** MAE, RMSE және $R^2$ қандай ақпарат береді?

**06** Қандай жағдайда 95% Accuracy нашар модельді жасыруы мүмкін?

**Қорытынды қағида: модельді таңдау — validation бойынша, соңғы бағалау — test бойынша.**

# Дереккөз және иллюстрациялар

## Негізгі дереккөз

«4-дәріс. Нейрондық желіні оқыту, валидациялау және сапасын бағалау». Пайдаланушы ұсынған конспект, 9 бет.

## Конспектіден алынған түпнұсқа суреттер

4.1-сурет — деректерді бөлу, 3-бет → 3-слайд.
4.2-сурет — модель күрделілігі және жалпылау, 4-бет → 5-слайд.
4.3-сурет — training/validation loss, 5-бет → 6-слайд.

Метрикалардың мысалдары конспект бойынша сақталды. Регрессия мысалына $R^2$ есебі және нөлдік бөлімге қатысты қысқа түсіндірмелер қосылды.