

# Сенімді модел құру, неге және қалай

---

## 2-бөлім: Сенімді қалыптастыру

**Data**

**Architecture**

**Tuning**

Нақты әлемдегі ML жүйелерінің өмірінде соңғы өнімділікті  
анықтаудың ең маңызды факторы қандай?

## 2-бөлім: Сенімді қалыптастыру



Data

1. Деректер ең маңызды фактор болып табылады. бірақ оны автоматтандыру қиын (деректер үнемі өзгереді, деректерде домен болжамдары бар, деректер сапасы жинаудың ең жақсы тәжірибелеріне байланысты және т.б.).
2. Архитектура барған сайын тауарлануда.
3. Баптау тұжырымдамалық тұрғыдан қарапайым, бірақ іс жүзінде қымбат болуы мүмкін..

## 2-бөлім: Сенімді қалыптастыру

Үш қадамдық жоспар:

- Үлгіңізге сену үшін деректеріңізге сену керек -> деректерді тексеру, тұтастықты тексеру, т.б.
- Модельге сену үшін жаттығу тәртібіне -> гипер баптау, экспериментті бақылау, сандық мақсат және т.б. сену керек.
- Модельге сену үшін оған шеткі жағдайларда немесе сізді ерекше қызықтыратын жағдайларда сену керек -> қателерді талдау, «қара жәшік» тестілеу және т.б.

Сіз деректеріңізге сенесіз бе?

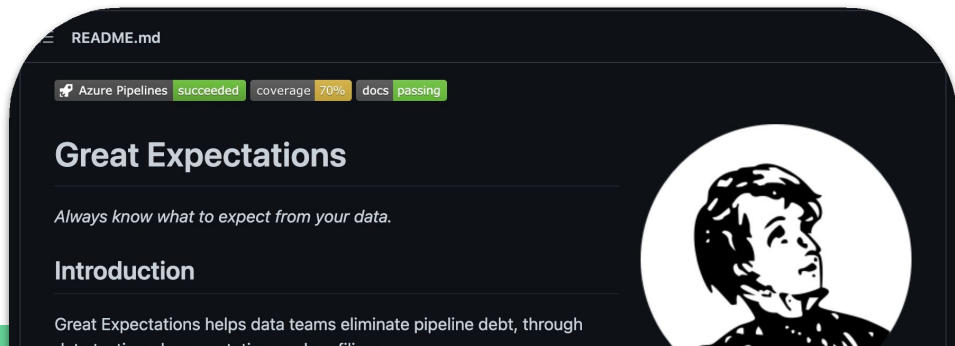
---

# Деректеріңізге сену

Модельге сену үшін деректеріңізге сену керек

Академиялық параметрлерде (және сіздің үй тапсырмаларыңызда!) деректер сізге жиі дайындалады, тазартылады және талдауыңыз үшін (бір нүктеге дейін) қалыпқа келтіріледі.

Бұл нақты әлемде болмайды: деректерді жинау өте күрделі процесс болуы мүмкін және ML жасамас бұрын, біздің «деректерді күту» орындалатынына көз жеткізу маңызды.



# Деректеріңізге сену

Модельге сену үшін деректеріңізге сену керек

Деректерімізді сұрағымыз келетін кейбір сұрақтар:

- Кейбір жетіспейтін құндылықтар бар ма? (Иә болса, онымен не істейміз?)
- Деректер жинағы теңгерімсіз бе? (Иә болса, онымен не істейміз?)
- $X$  мүмкіндігі үшін мән ауқымы орынды ма? Мысалы, біз «жас» бағанында 120-ға дейін тек оң мәндер болады деп күтеміз.
- $X$  мүмкіндігі үшін орташа мән/медиан негізді ме? Мысалы, егер деректер жиыны жалпы популяцияны көрсетсе, «IQ» бағанының шамамен 100 мәні болады деп күтеміз.

Сіз модел оқытуға сенесіз бе?

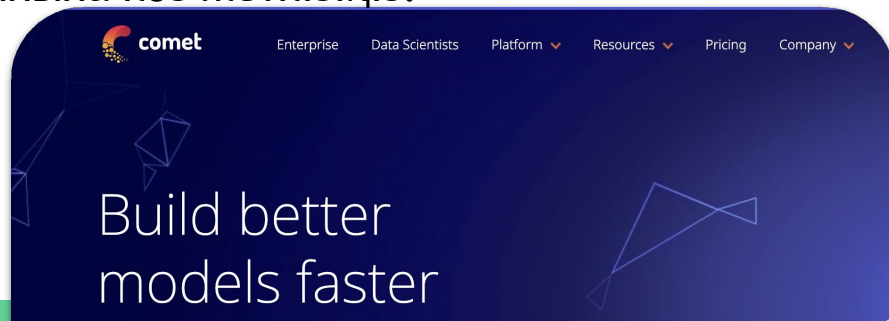
---



# Сіздің оқытуыңызға сену

## Модельге сену үшін ОҚЫТУ тәртібіне сену керек

- Пойыздың, валидацияның, сынақ бөлінісінің дұрыс екеніне көз жеткізіңіз (С: акциялардың тарихи бағалары туралы деректер жинағын қалай бөлеміз?)
- Сәйкес гиперпараметрлерді анықтап, оларды дұрыс оңтайландырыңыз: эксперименттерді бақылау және ұйымдастыру үшін экспериментті бақылау жүйесін пайдаланыңыз
- Нәтижелерді шығаруға болатындай артефактілерді (деректер, үлгілер) нұсқасына көз жеткізіңіз (С: кездейсоқтықпен қалай күресеміз?)
- Пайдалану жағдайын ескере отырып, сынақ жиынындағы соңғы көрсеткіштердің қанағаттандыратынына көз жеткізіңіз.



Сіз өзіңіздің бағалауыңызға сенесіз  
бе?

---

# Train vs Test: неге?

Сіз енді оқытудың негізгі ағынымен және үлгіні бағалаумен таныссыз:

- Пойыздағы деректер жиынын бөлу және сынақ
- Жаттығу жиынындағы пойыз үлгісі
- Сынақ жиынындағы сынақ үлгісі және есеп көрсеткіштері (мысалы, дәлдік және қайта шақыру)

Неліктен сынақ жиынындағы көрсеткіштер туралы есеп беру керек?

# Train vs Test: неге?

- Сіз енді оқытудың негізгі ағынымен және үлгіні бағалаумен таныссыз:
  - Пойыздағы деректер жиынын бөлу және сынақ
  - Жаттығу жиынындағы пойыз үлгісі
  - Сынақ жиынындағы сынақ үлгісі және есеп көрсеткіштері (мысалы, дәлдік және қайта шақыру)
- Неліктен сынақ жиынындағы көрсеткіштер туралы есеп беру керек?
  - Өйткені сынақ жинағы жалпылау үшін табиғи прокси болып табылады: модель жаңа кірістерге тап болған кезде қаншалықты жақсы жұмыс істейді?
    - Нақты өмірде бұл дегеніміз: біз оны тұтынушылар алдында орналастырған кезде модель қаншалықты жақсы жұмыс істейді?
  - Модель «біздің көрсеткіштерімізге сәйкес» жақсы жұмыс істеп тұрса, біз оны қолданған кезде жақсы жұмыс істейді деп күтеміз.

# Көрсеткіштер: неге?

- Сіз әртүрлі көрсеткіштер түрлерін кездестірдіңіз :

- Регрессия көрсеткіштері: мысалы,  $R^2$ , MSE т.б.

- Жіктеу көрсеткіштері: дәлдік, еске түсіру және т.б.

$$R^2 = 1 - \frac{SS_{Regression}}{SS_{Total}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

- Метрикалардың қолданылуы неде?

- Өте күрделі нысанды бірнеше сандармен қорытындылаңыз

- Үлгілерді салыстыруға болатын етіп жасаңыз

- Бізге жеке және өріс ретінде «ілгерілеу» сезімін беріңіз

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

Сіз өлшемейтін нәрсені жақсарта  
алмайсыз

# Көрсеткіштер: неге?

- Сіз әртүрлі көрсеткіштер түрлерін кездестірдіңіз :

- Регрессия көрсеткіштері: мысалы,  $R^2$ , MSE т.б.
- Жіктеу көрсеткіштері: дәлдік, еске түсіру және т.б.

$$R^2 = 1 - \frac{SS_{Regression}}{SS_{Total}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

- Метрикалардың қолданылуы неде?

- Өте күрделі нысанды бірнеше сандармен қорытындылаңыз
- Үлгілерді салыстыруға болатын етіп жасаңыз
- Бізге жеке және өріс ретінде «ілгерілеу» сезімін беріңіз

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

- Жақсы көрсеткіш дегеніміз не?

- Біз қызықтыратын шынайы құбылыспен байланысты
- Көрсеткіш жақсарған кезде «менің әлемім жақсырақ»
- Ойын қиын
- Түсінуге оңай

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

# Тәжірибедегі көрсеткіштер

- **TL;DR:** there is no a priori right / wrong metric for your project (as long as you're not using a regression metric for classification!) - even the same problem can be framed and evaluated differently depending on what we care about.
- **Example 1:**
  - Model: a binary classifier for healthcare
  - Input: a X-Ray
  - Output: cancer diagnosis
  - If I ask you to pick between reporting precision and recall, what would you do? **What is more important, to be always right or to not miss anything?**



# Тәжірибедегі көрсеткіштер

- TL;DR: жобаңыз үшін априорлы дұрыс/бұрыс метрика жоқ (жіктеу үшін регрессия метрикасын пайдаланбасаңыз!) - тіпті бірдей мәселе де біз ойлайтын нәрсеге байланысты әртүрлі негізделуі және бағалануы мүмкін..
- **2-мысал:**
  - Үлгі: құқық қорғау органдарына арналған екілік классификатор (ғылыми-фантастикалық мысал)
  - Кіріспе: қылмыстың дәлелі
  - Нәтижесі: кінәлі / кінәсіз үкім
  - Егер мен сізден есеп беру дәлдігі мен еске түсіру арасында таңдауды сұрасам, не істер едіңіз? Ең бастысы, әрқашан дұрыс болу немесе ештеңені жіберіп алмау?

«Сануға болатынның  
бәрі де саналмайды  
және есептелетіннің  
бәрі де саналмайды»  
(А. Эйнштейн)