Розділ 1: Статистична постановка задач розпізнавання
Домашнє завдання

Перш ніж почати

Вимоги

- якомога більше дій повинно бути описано словами: це дозволяє перевірити розуміння матеріалу та перевіряти роботи на унікальність;
- для кожної змінної, константи і функції повинно бути вказано множину (або область значень і визначення), якій вона належить: це дозволяє запобігати багатьох помилок та покращує зрозумілість роботи;
- письмові роботи виконуються індивідуально.

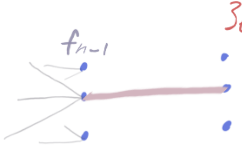
Рекомендації

- ретельно перевіряйте свою роботу перед тим, як відправляти її викладачеві;
- якщо завдання виконується на аркуші паперу і фотографується, намагайтеся задіяти достатньо освітлення і фотографувати якомога рівніше, щоб усі частини сторінки з розв'язком було добре видно.

Ілюстрації

Рекомендації до виконання письмових завдань на прикладах і картинках

Звідки такий висновок?



$$\Rightarrow d_{n-1} \in \underset{d}{\arg\min}\, f_{n-1}(d) = F_{n-1}(d)$$

колізія імені d

в якій множині шукати d?

Малюнок не є доведенням

$$\min_{d \in D}\left[f_{n-1}(d) + \lambda(d_n)\right] =$$

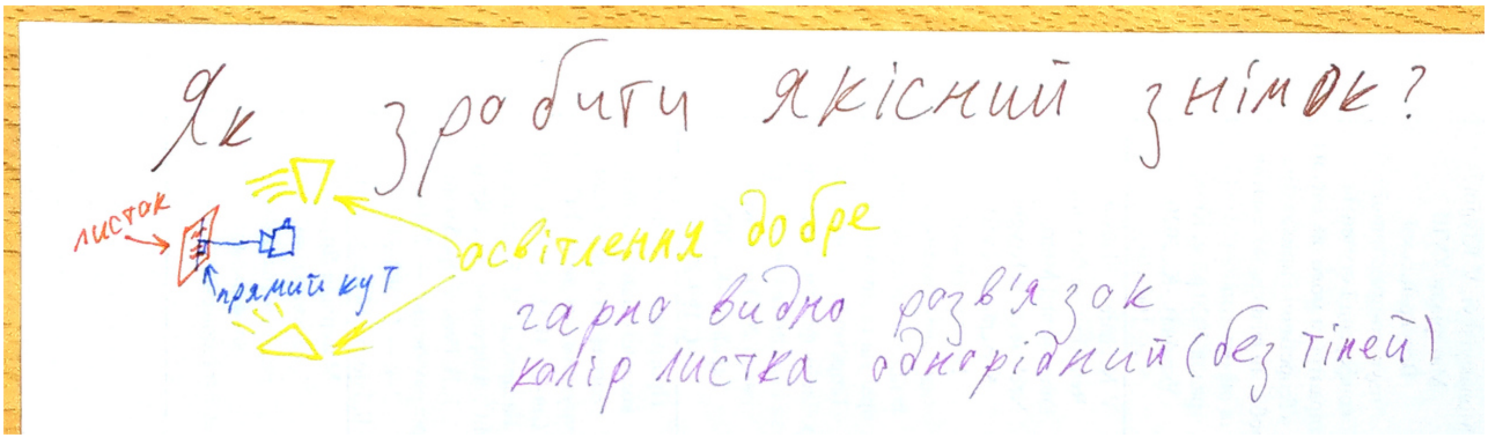
Оскільки $\lambda(d_n)$ не залежить від d, його можна винести як константу за межі мінімізації

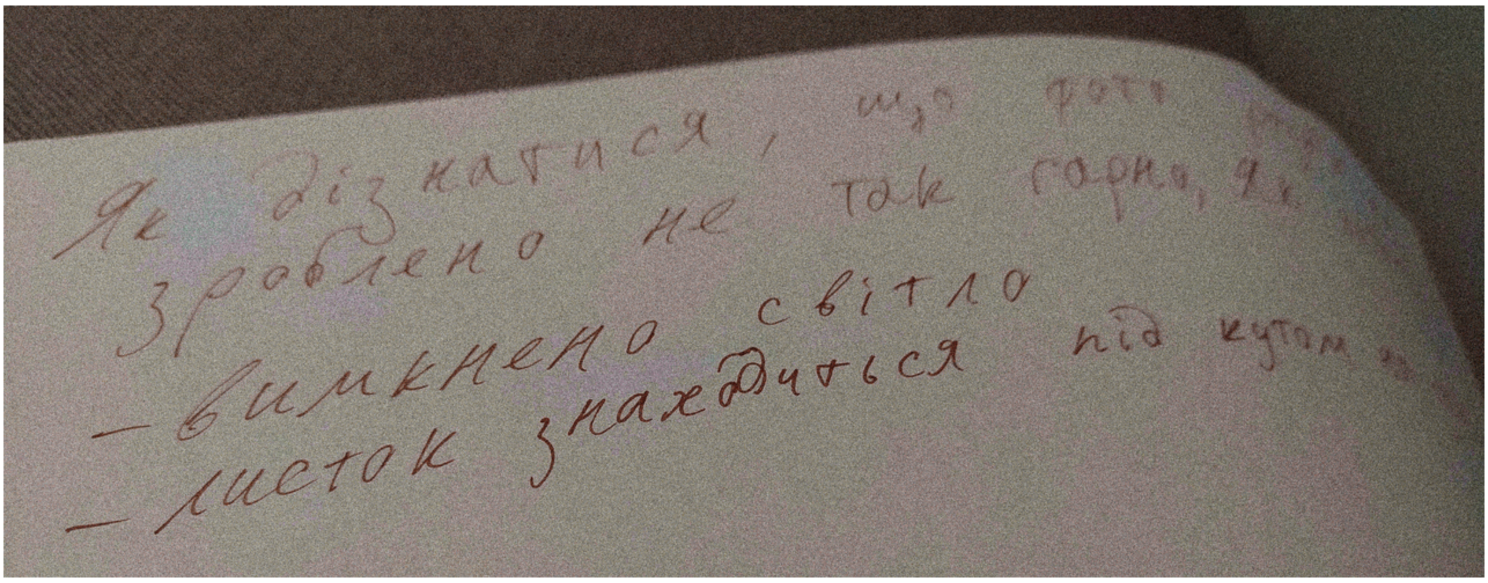
детальні коментарі

$$= \left[\min_{d \in D} f_{n-1}(d)\right] + \lambda(d_n), \ \forall d_n \in D, \forall n = N..1.$$

вказано множини

визначено всі змінні





1 Основне (2 бали)

1.1 Мінімізація часткового ризику

Показати, що для таких стратегій q^*, що

$$q^* \in \operatorname*{argmin}_{q:X\to D} \sum_{\substack{k\in K\\ x\in X}} w\left(q\left(x\right), k\right) \cdot p\left(x, k\right),$$

виконується

$$\forall x \in X : q^*\left(x\right) \in \operatorname*{argmin}_{d\in D} \sum_{k\in K} w\left(d, k\right) \cdot p\left(x, k\right).$$

Навіщо цей факт потрібен на практиці?

1.2 Розпізнавання тривимірних об'єктів

Розв'язати задачу 2.1 (додаткова задача цього розділу) з дещо іншими умовами:

- $K = \mathbb{R}^{100}$ — множина параметрів моделі;
- $X = \mathbb{R}^{10^6\times 3}$ — множина тривимірних моделей з мільйоном вершин;
- $\otimes = +$ — операція додавання матриць $10^6 \times 3$;
- розподіл шуму та параметрів береться за варіантами, що призначає викладач

Варіант	$p\left(k\right)$	$p\left(N\right)$
1	$\mathcal{N}$	$\mathcal{N}$
2	$\mathcal{N}$	$\mathcal{L}$
3	$\mathcal{L}$	$\mathcal{N}$
4	$\mathcal{L}$	$\mathcal{L}$

$\mathcal{N}$ — компоненти незалежні та розподілені за стандартним нормальним законом, $\mathcal{L}$ — компоненти незалежні та розподілені за законом Лапласа з масштабом 1 та зсувом 0.

Для допитливих

За посиланням `https://flame.is.tue.mpg.de/interactive-model-viewer` можна побавитися з параметрами k, щоб побачити $g\left(k\right)$. За посиланням `https://flame.is.tue.mpg.de` ви можете дізнатися

- що собою може являти g;
- звідки береться розподіл параметрів k;
- як це дозволяє розпізнавати тривимірні об'єкти.

2 Додаткове (3 бали)

2.1 Розпізнавання за допомогою породжуючої моделі

Дано

- групу $\langle X, \otimes \rangle$, де X — скінченна множина сигналів;
- скінченну множину станів K;
- множину можливих розв'язків $D = K$, що співпадає з множиною станів;
- штрафну функцію $w(d, k) = t \cdot [\![d \neq k]\!] + c$, де t — додатнє дійсне число, c — невід'ємне дійсне число, а $[\![\cdot]\!]$ — дужки Айверсона.

Статистичну модель $p(x, k)$ задано опосередковано: відомо, що $x = g(k) \otimes N$, де $N : \Omega \to X$ — шум з відомим розподілом, а $g : K \to X$ — відома бієктивна функція, що для кожного стану генерує певний "ідеальний" образ; апріорні ймовірності $p(k)$ теж відомі для всіх $k \in K$.

Знайти оптимальну стратегію q^* (якомога краще спростити вираз)

$$q^* \in \operatorname*{argmin}_{q:X \to D} \sum_{\substack{k \in K \\ x \in X}} w(q(x), k) \cdot p(x, k).$$

- Чи всі властивості групи є необхідними для використання отриманої стратегії?
- Чи є достатнім для функції g бути сюр'єктивною?
- Чи стане простішою задача, якщо $\forall k \in K : p(k) = \frac{1}{|K|}$?

3 Комп'ютерне (4 бали)

Перш ніж почати

- спочатку розв'язується теорія, а вже потім на її основі пишеться код: це дозволяє не починати роботу з реалізації алгоритму, що є невірнім;
- бажано перевірити алгоритм на папірці: отримані дані також можна буде використовувати в автотестах;
- мова програмування довільна;
- дозволяється написання коду в бригадах розміром до двох людей; за commit messages буде визначатися внесок кожного учаснику, і на основі цього робитися висновок щодо того, чи було роботу виконано бригадою чи однією людиною;
- бажано використовувати технології, що спрощують збірку й запуск коду: CMake для C, C++ і Fortran, setuptools для Python, npm для NodeJS, Maven/Ant для Java/Scala і так далі.

Критерії оцінювання

- 2 бали (обов'язкова умова): коректно працююча програма з кодом у доступному викладачеві репозитарії (наприклад, GitHub, GitLab, BitBucket) та наявність теоретичного розв'язку поставленої задачі;
- 1 бал: наявність автотестів (наприклад, doctest у Python, Boost.Test у C++, JUnit у Java);
- 1 бал: наявність змістовних commit messages та CI (наприклад, TravisCI, Jenkins, GitLab Runner).

Якщо не знаєте, що обрати для тестування або CI, або як писати тести чи прив'язати до свого репозитарію CI, зверніться до викладача.

3.1 Бінарна штрафна функція

Задача

На вхід програмі через WebSocket передаються бінарні зашумлені зображення відомих еталонів. Шум — набір незалежних однаково розподілених випадкових величин з розподілом Бернуллі і параметром $0 \leq p \leq 1$, який визначається користувачем. Шум застосовується до кожного пікселя за допомогою виключної диз'юнкції (XOR). Потрібно мінімізувати ризик байєсової стратегії для штрафної функції $w(k, d) = [\![k \neq d]\!]$.

Мета

Закріпити навички максимізації апостеріорної ймовірності.

Завдання

Завдання First з `https://sprs.herokuapp.com`. Програма має працювати за умов

1. рівень шуму 0 (без шуму всі відповіді мають бути вірними);
2. рівень шуму 1 (при повній інверсії всі відповіді мають бути вірними);
3. рівень шуму 0.4 і менше (а також 0.6 і більше) й масштабі 20 на 20 і більше, адже, коли картинка має розмір 100 на 60, ймовірність "перетворення" однієї цифри на іншу після накладання такого рівня шуму надто мала, тому відповідь майже завжди повинна бути вірною.

3.2 Штрафна функція L_1

Задача

На вхід програмі через WebSocket передаються ненормовані значення $p(k \mid x)$. Мінімізувати ризик байєсової стратегії для штрафної функції $w(k, d) = |k - d|$. Потужність множини $K = \{0, 1, \ldots, |K| - 1\}$ визначає користувач.

Мета

Засвоїти на практиці, що максимізація апостеріорної ймовірності не у всіх випадках є оптимальною стратегією.

Завдання

Завдання Second з `https://sprs.herokuapp.com`.