15 ДӘРІС

Мультимодальды және ашық сөздік көру

Мәтін мен кескінді біріктіретін заманауи бағыттар

CSE8841

02

Мультимодальды модель

1. **Әртүрлі модаль бір мағыналық кеңістікке түседі**
2. Кескін encoder-і визуалды белгілерді береді
3. Мәтін encoder-і тілдік мағынаны береді
4. Fusion модальдар арасын байланыстырады

03 Vision language pretraining

01

Кескін мәтін жұптары оқыту дерегі болады

02

Contrastive мақсат сәйкес жұпты жақындатады

03

Captioning мақсат мәтін генерациясын үйретеді

04

Instruction tuning тапсырма орындауды бейімдейді

Тақырып элементтері бір жүйе ретінде қарастырылады.

04

Zero shot тану

Класс атауы мәтіндік prompt-қа айналады

Кескін embedding барлық prompt-пен салыстырылады

Ең жоғары ұқсастық болжам береді

Жаңа класс үшін қайта оқыту қажет емес

05

Zero shot классификация

$$v = E_{image}(x), \quad u_k = E_{text}(prompt_k)$$

$$s_k = \cos(v, u_k)/\tau$$

$$\hat{y} = \arg\max_k s_k$$

Класс атаулары мәтіндік prompt ретінде кодталады. Модель кескінге ең ұқсас мәтін embedding-ін таңдайды.

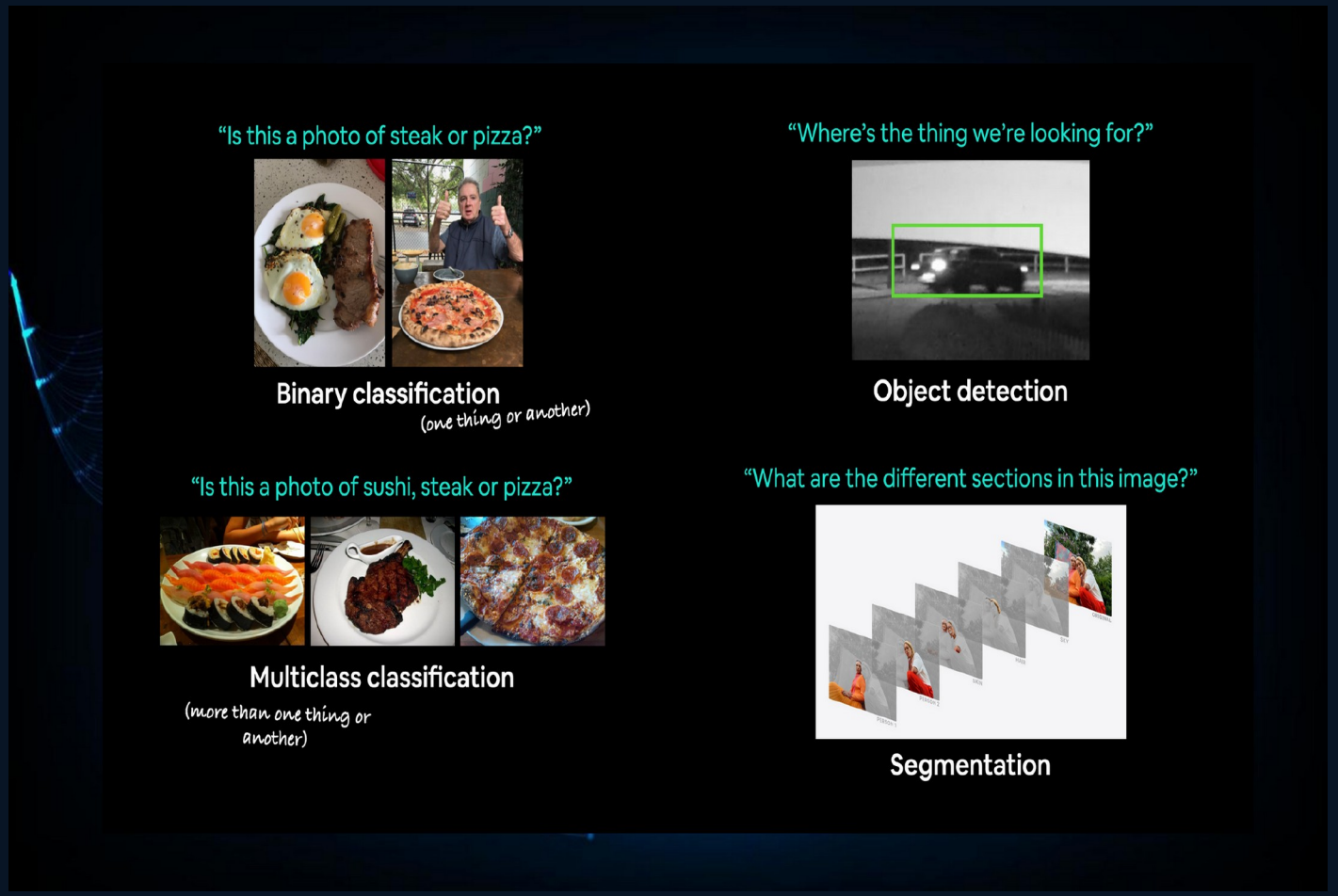
06

Визуалды сұраққа жауап

1. **Кескін контекст ретінде кодталады**
2. Сұрақ қажетті ақпаратты бағыттайды
3. Модель жауапты мәтін түрінде шығарады
4. Дәлелсіз жауап қаупі тексеруді қажет етеді

Модель нәтижесін түсіндіру

Ықтималдық сенімділікті көрсетеді.
Қате матрицасы шатастырылған кластарды ашады.
Визуал түсіндіру модель назарын тексереді.



Дереккөз: берілген дәріс материалынан бейімделген

Доменге бейімдеу

Prompt инженериясы терминдерді нақтылайды

Few shot мысалдар контекст береді

Adapter шағын параметр бөлігін үйретеді

Retrieval сыртқы білімді қосады

Нәтижені бағалау

```
from sklearn.metrics import confusion_matrix

pred = model(x).argmax(dim=1)
cm = confusion_matrix(y.cpu(), pred.cpu())
accuracy = (pred == y).float().mean()
```

Accuracy жалпы дәлдікті береді, қате матрицасы әр класс бойынша қатені көрсетеді.

Код дәріс материалындағы тәсілге сүйеніп бейімделді

ҚОРЫТЫНДЫ

Негізгі тұжырымдар

1. **Мультимодальды модель кескін мен тілді ортақ мағынада біріктіреді**
2. Open vocabulary тәсілі жаңа санатты мәтін арқылы сұрайды
3. Заманауи жүйеде сапамен бірге сенімділік те бағаланады

Өзіндік тексеру: негізгі формуланы түсіндіріп, әдістің бір практикалық қолданылуын атаңыз.