Зертханалық жұмыс №9

Тақырыбы: Transformer моделін кескіндерді классификациялауға қолдану

Тапсырма:
Vision Transformer (ViT) моделін пайдаланып, CIFAR-10 деректер жиынтығындағы кескіндерді классификациялайтын жүйе құрыңыз.

Жұмысты орындау барысында:

1. CIFAR-10 деректер жиынтығын жүктеп, оның кластары мен құрылымын зерттеңіз.
2. Кескіндерді модельге сәйкес өлшемге келтіріп, қалыптандырыңыз.
3. Кескіндерді бірдей өлшемдегі бөліктерге- **patch** элементтеріне бөліңіз.
4. Алдын ала оқытылған Vision Transformer моделін жүктеңіз.
5. Модельдің шығыс қабатын CIFAR-10 жиынтығының 10 класына бейімдеңіз.
6. Модельді оқыту деректері арқылы оқытыңыз.
7. Тестілік деректердегі модель сапасын:
    - Accuracy;
    - Precision;
    - Recall;
    - F1-score метрикалары арқылы бағалаңыз.
8. Кемінде 10 тестілік кескін үшін нақты және модель болжаған класс атауларын көрсетіңіз.
9. Қате классификацияланған кескіндерді талдаңыз.
10. Орындалған жұмыс бойынша қорытынды жазыңыз.

Ұсынылатын құралдар: Python, Google Colab, PyTorch немесе TensorFlow, Hugging Face Transformers.

Күтілетін нәтиже: Vision Transformer моделі кіріс кескіндерін CIFAR-10 деректер жиынтығындағы тиісті кластарға жіктеуі керек.

Бақылау сұрақтары

1. Transformer архитектурасының негізгі жұмыс істеу принципі қандай?
2. Vision Transformer моделінде кескін неліктен patch элементтеріне бөлінеді?
3. Self-attention механизмі кескінді өңдеу кезінде қандай қызмет атқарады?
4. Vision Transformer мен CNN архитектураларының негізгі айырмашылықтары қандай?
5. Кескіндерді классификациялау сапасын бағалау үшін Accuracy, Precision, Recall және F1-score метрикалары қалай қолданылады?