

Автоматизована генерація описів для будівель на основі зображень

Виконала
Кузнецова Анна, БП КН-4

Керівник
Іванюк А. О.

Чому ця робота важлива?



Знищена садиба Зеленських

Щороку унікальні пам'ятки історичної архітектури Києва опиняються під загрозою знесення - лише за 2024–2025 роки десятки об'єктів культурної спадщини були на межі знищення.

Однією з головних проблем є потреба в регулярному моніторингу, щоб вчасно виявляти ті будівлі, що знаходяться під загрозою. Оновлювати облік без автоматизованих інструментів вимагає занадто багато людського ресурсу.

Мета роботи

Розробити систему автоматичної генерації текстових описів архітектурних будівель Києва, дослідити та порівняти ефективність мультимодальних моделей для цієї задачі на реальних даних культурної спадщини міста.

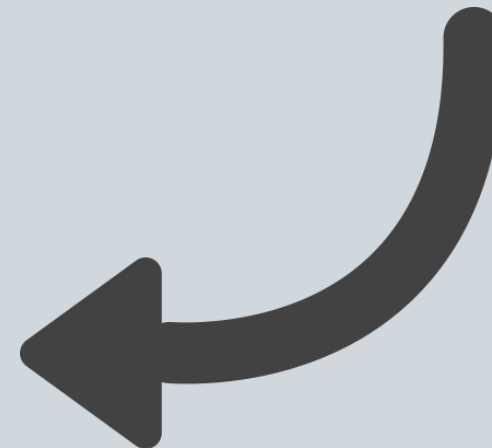
Результати роботи спрямовані на спрощення документування архітектурних пам'яток та підтримку процесу їхньої каталогізації.

Приклад роботи системи

- {
- "building_type": "house",
- "alternate_building_type": "apartments",
- "building_age": "1950",
- "floors": 2,
- "surface_material": "plaster",
- "alternate_surface_material": "concrete",
- "construction_material": "wood",
- "alternate_construction_material": "brick"
- }



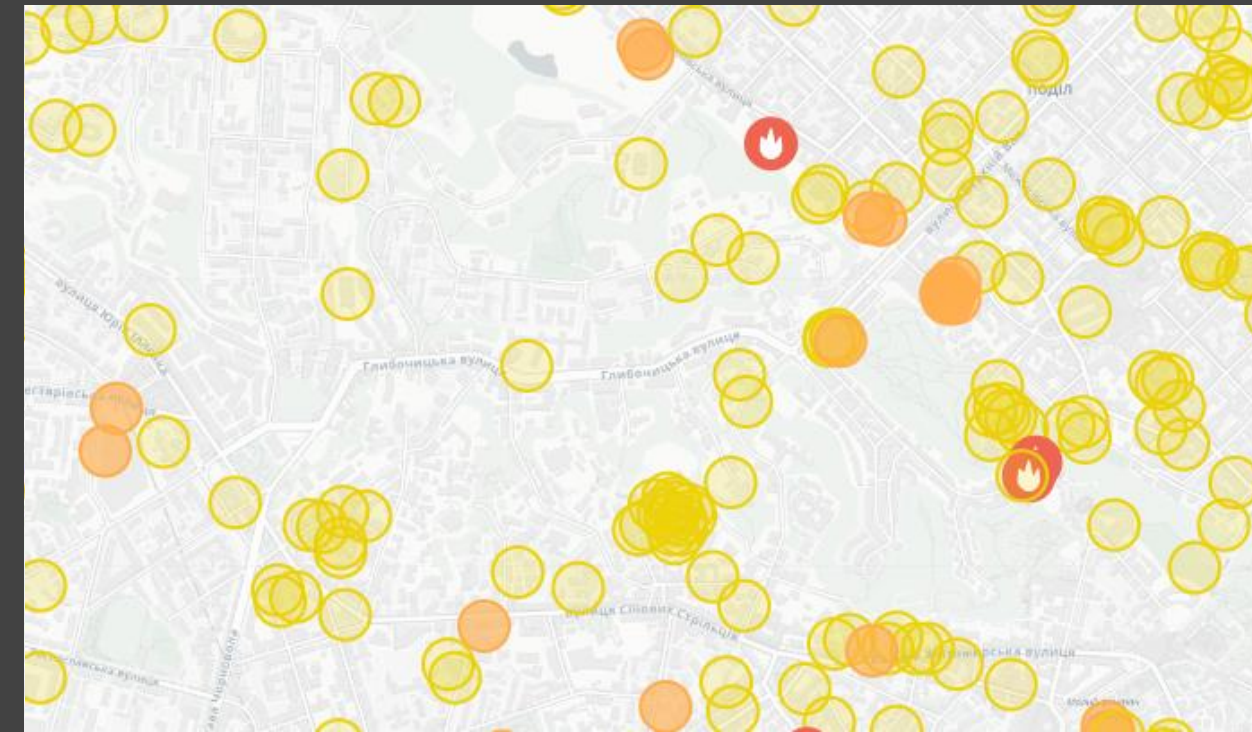
ВХІД



Вихід
Д

❁ Збір даних

1. Джерело 1: Сайт інтерактивної мапи, розробленої ГО “Мапа Реновацій”



Інтерактивна мапа ГО “Мапа Реновацій”

1. Джерело 2: Кадастровий реєстр від ДОКС

— — **Реєстр пам’яток** — —

Координати — —

Фотографії — —

Датасет — —

Результат: понад 1700 фотографій із структурованими характеристиками:

- Тип будівлі
- Рік побудови
- Категорія архітектури

Приклад запису в датасеті

Зображення та відповідні метадані



object_id	name	address	district	type	year	year_precision	status	lat	lon	num_images	image_paths
1	Навчальна будівля	Амосова, 9	солом	architecture	1896	exact	M	50.4245305	30.5038401	4	dataset/images/obj_1
2	Релігійна будівля	Андріївська, 1/1 (ріг)	под	architecture_urbanis	1891	range	M	50.460596	30.520113	4	dataset/images/obj_2
3	Прибуткова / комер	Андріївська, 2/12	под	architecture_urbanis	1896	range	Щ	50.4600838	30.520612	4	dataset/images/obj_3
4	Археологічна споруд	Андріївська, 7	под	archaeology	1150	century	M	50.4615412	30.5219917	4	dataset/images/obj_4
5	Прибуткова / комер	Андріївська, 9	под	architecture_urbanis	1910	exact	M	50.4618482	30.522495	4	dataset/images/obj_5

Методи оцінки експериментів

Як бейслайн метрики, було використано:

- BLEU-1
- BLEU-4
- METEOR
- ROUGE-L

Додатково для BLIP:

- bertscore_f1_free
- bertscore_f1_prompted

Через стилістичну невідповідність між референсними та згенерованими описами n-gram метрики давали занижені результати. Тому було розроблено **атрибутивний пайплайн оцінки**: автоматична екстракція 5 архітектурних атрибутів у JSON з подальшим порівнянням референсу та генерації.

Атрибути скалярні:

- к-ть поверхів
- епоха будівництва
- технічний стан

Accuracy

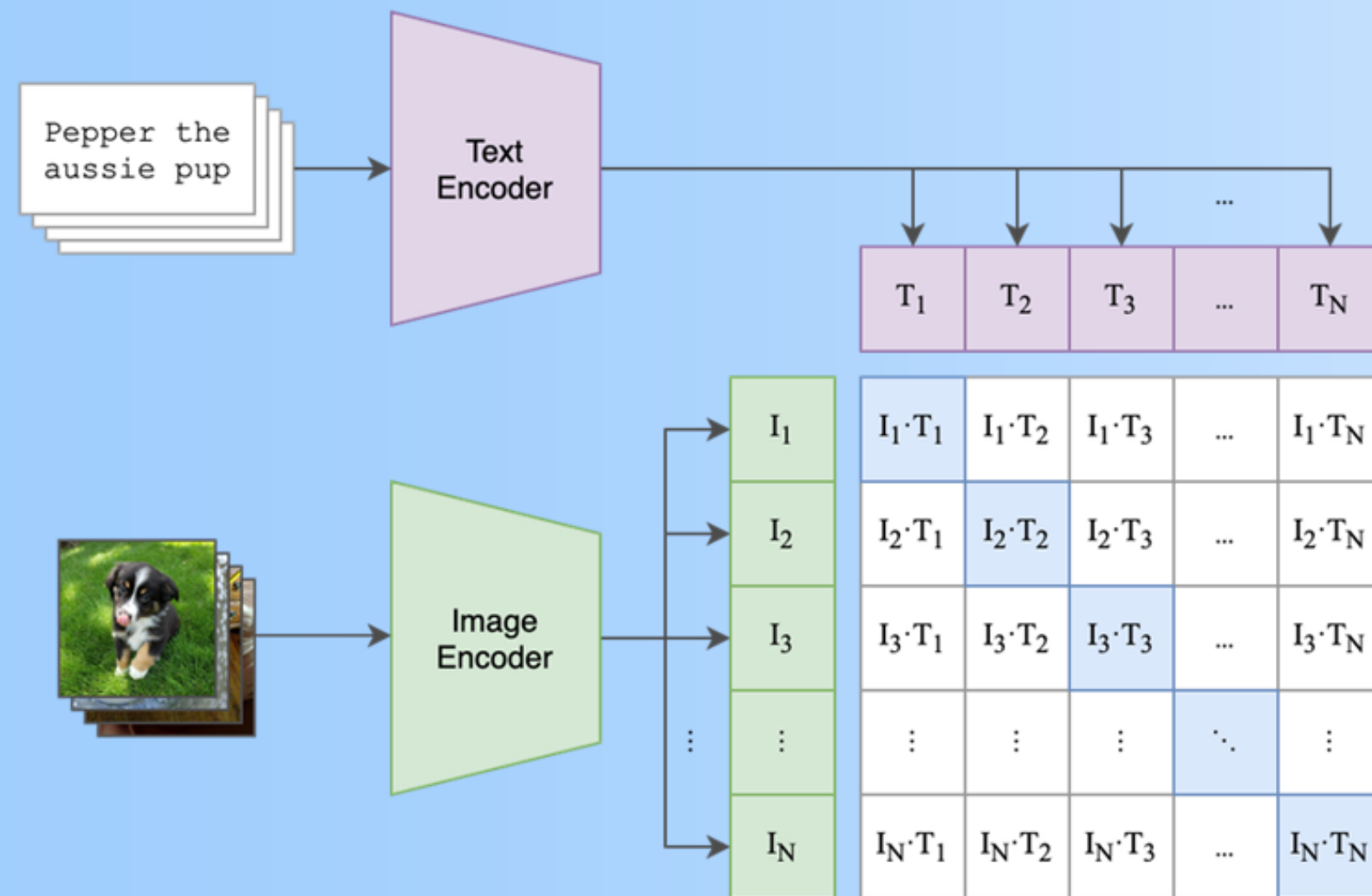
Атрибути листові:

- матеріал
- архітектурний стиль

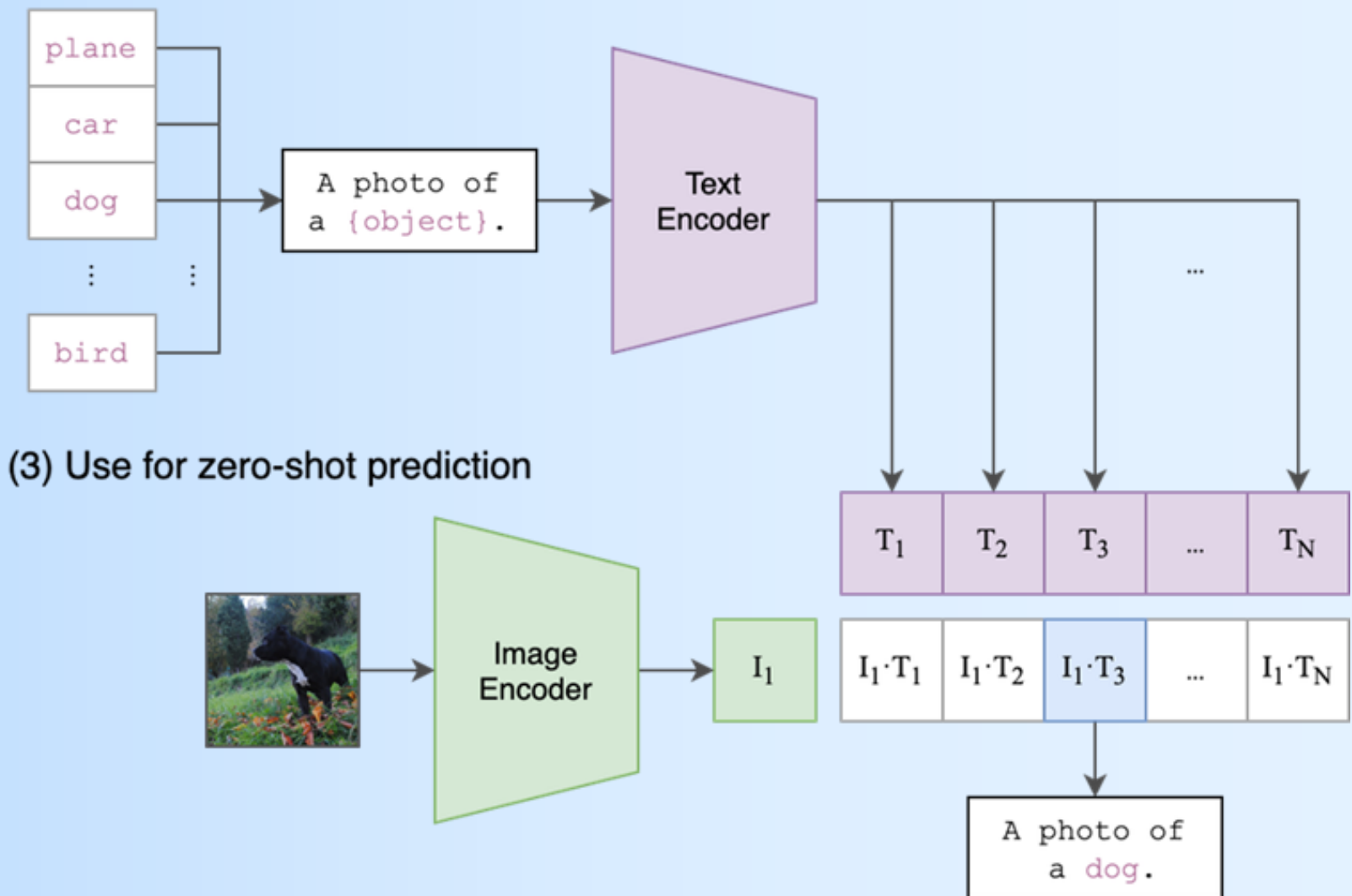
F1-score з підрахунком
TP/FP/FN.

Архітектура Clip

(1) Contrastive pre-training

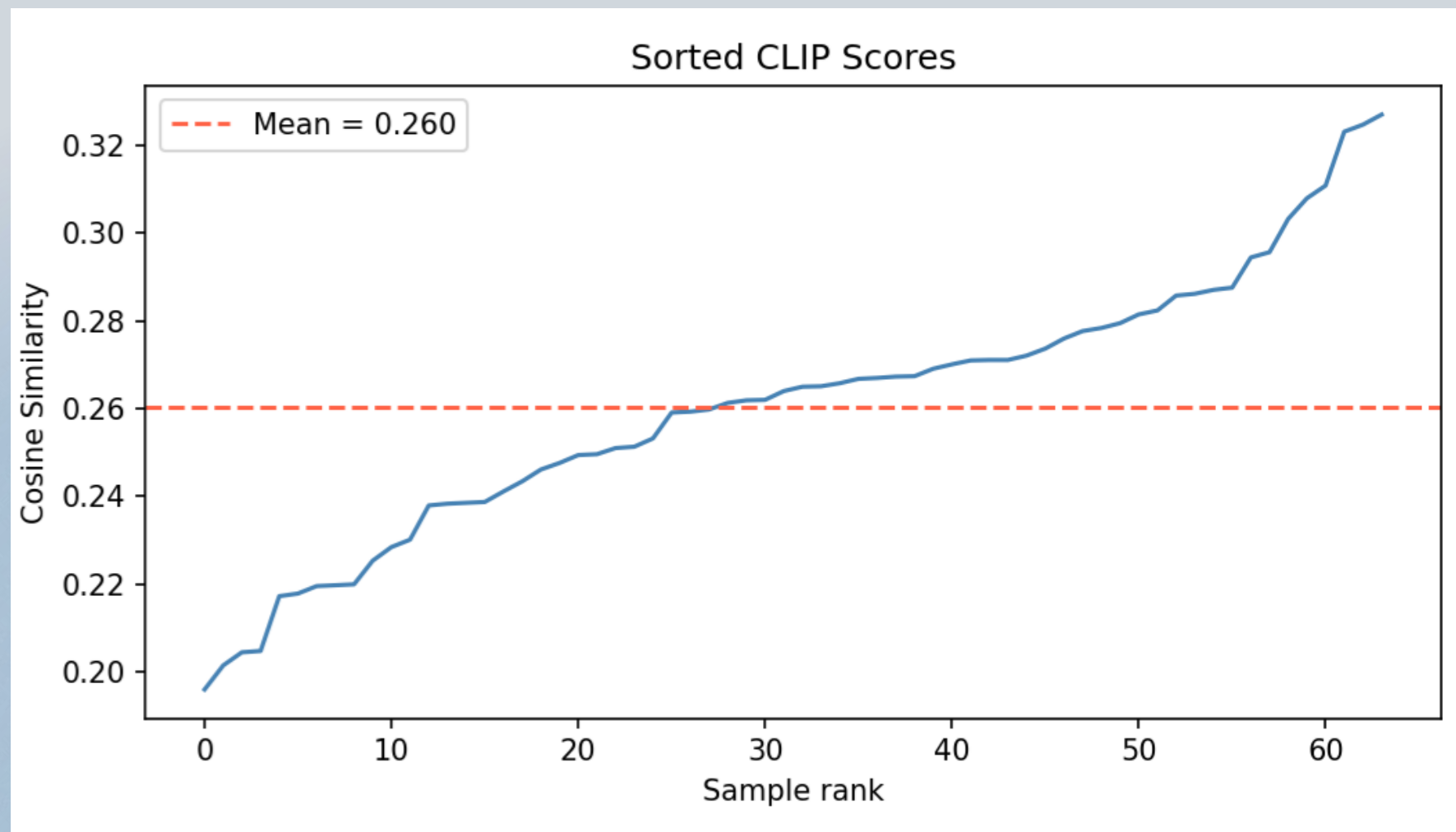


(2) Create dataset classifier from label text



(3) Use for zero-shot prediction

Експеримент 1: Clip



- Адаптація даних відповідно до характеристик моделі: україномовні описи перекладено англійською та скорочено до 77 tokenів.
- Оцінка косинусної подібності між візуальними та текстовими ембедингами на 64 тестових зразках.
- Виявлення ключових обмежень моделі в контексті досліджуваної задачі

Експеримент 2: Порівняння мультимодальних моделей

	stories	era	condition	material			style			macro avg
				precision	recall	F1	precision	recall	F1	
Blip	0.0	0.0	0.1429	0.1667	0.1667	0.1667	0.0	0.0	0.0	0.0619
Qwen3-VL	0.5882	0.25	0.2273	0.619	1.0	0.7381	0.05	0.05	0.05	0.3707
Llava-1.5	0.2353	0.2542	0.2727	0.625	0.6875	0.6458	0.1667	0.1667	0.1667	0.315
InternVL3.5	0.7778	0.4167	0.3043	0.5625	0.75	0.625	0.25	0.3125	0.2708	0.4789

Атрибутивна оцінка всіх досліджених
архітектур

Експеримент 3:

Донавчання моделі

методом низькорангової

адаптації (LoRA)



	floors	build age	building type			surface material			construction material			macro avg
			precision	recall	F1	precision	recall	F1	precision	recall	F1	
Донавчена Qwen3-VL	0.9167	0.3167	0.2909	0.5152	0.3648	0.3889	0.7778	0.5185	0.4375	0.875	0.5833	0.54

❁ Висновки

- Дотримання моделі зданого шаблону і формату
- Вказування неможливості встановлення атрибуту
- Підвищена точність для атрибутів:
 - кількість поверхів (0.5882→0.9167)
 - приблизному році побудови (0.25→0.3167)
- Підвищене значення **macro average**
 - з 0.3707 → 0.54

Прикладне застосування даної роботи можливе в інтеграції розробленого підходу з реєстровими системами ДОКС для автоматизованого поповнення облікової документації пам'яток.



Презентувала
Кузнецова Анна

Питання та Зауваження

 Дякую за приділений час