ПРОЕКТ НА ТЕМУ ВИКОРИСТАННЯ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ (CNN)

**Тема: Класифікація мінералів**

**Датасет:** <https://www.kaggle.com/datasets/asiedubrempong/minerals-identification-dataset>

**Хід роботи:**

Мій датасет містить 7 різновидів мінералів, а саме: Biotite, Bornite, Chrysocolla, Malachite, Muscovite, Pyrite, Quartz.

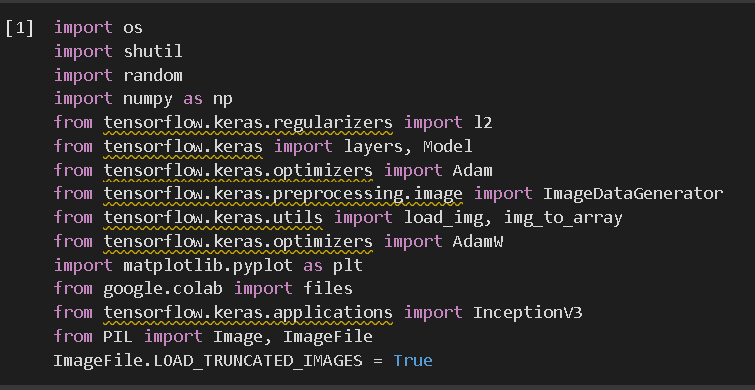
Всього в ході виконання проекту було збережено 5 експериментів. І останній файл з порівняннями експериментів:

* [minerals\_classification\_EXP1.ipynb](https://colab.research.google.com/drive/1JAxFzCA0vZb7Ts7TBrr7rjx9Sz4cmGef?usp=sharing)
* [minerals\_classification\_EXP2.ipynb](https://colab.research.google.com/drive/1FbSOpa10mSm1dOvFb2XpTD327lJ1dsPl?usp=sharing)
* [minerals\_classification\_EXP3.ipynb](https://colab.research.google.com/drive/17NHY_TLByQKmQXhXDGiHVo1KvUbsbSzr?usp=sharing)
* [minerals\_classification\_EXP4.ipynb](https://colab.research.google.com/drive/1-FouId3i7NJF823ARXr6SUI3Pl_Vjalb?usp=sharing)
* [minerals\_classification\_EXP5.ipynb](https://colab.research.google.com/drive/1hEI6DYeBWHlK3gB7BO8qBNcpKN0tSlew?usp=sharing)
* [minerals\_classification\_final.ipynb](https://colab.research.google.com/drive/1IF8bQUDXgbVBIrg8826D4zfU3Bpbwt1p?usp=sharing)

У цьому проекті створено та використано модель на основі попередньо навченої мережі InceptionV3, побудованої на базі «imagenet».

## ЕКСПЕРИМЕНТ №1

* [minerals\_classification\_EXP1.ipynb](https://colab.research.google.com/drive/1JAxFzCA0vZb7Ts7TBrr7rjx9Sz4cmGef?usp=sharing)

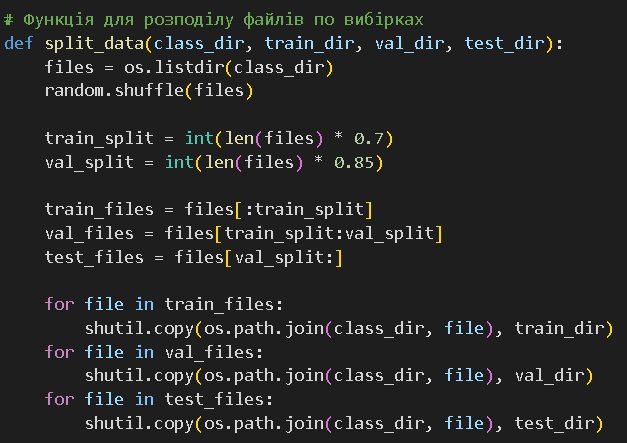
Імпорт необхідних бібліотек  


Розархівування архіву, в якому знаходиться датасет

Зображення, що містить текст, знімок екрана, монітор, Шрифт

Автоматично згенерований опис

Всі файли були розподілені по вибіркам (train, validation та test)

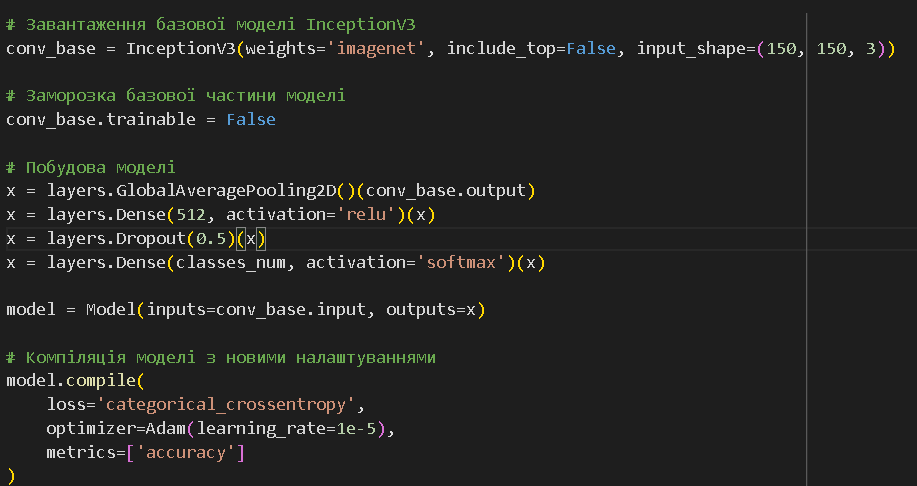


Створення генераторів даних для кожної вибірки

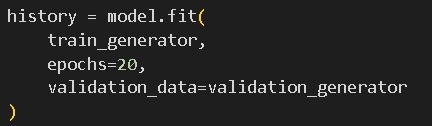
Зображення, що містить текст, знімок екрана, програмне забезпечення, Веб-сайт

Автоматично згенерований опис

Створення моделі



Після налаштування вибираємо параметри для навчання моделі, в моєму випадку я вибрав 20 епох



Результат навчання:

Зображення, що містить текст, знімок екрана, меню, Шрифт

Автоматично згенерований опис

Зображення, що містить Графік, ряд, текст, знімок екрана

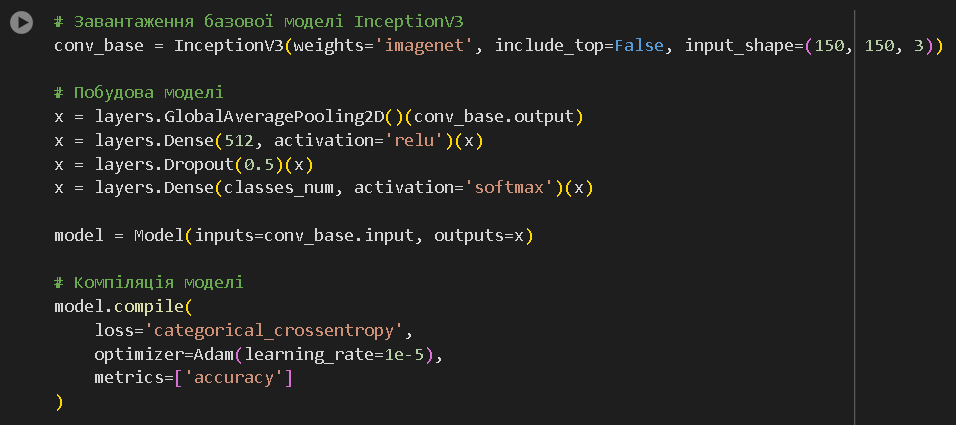
Автоматично згенерований опис

Досить погані результати. В наступних експериментах буду старатись їх покращити.

## ЕКСПЕРИМЕНТ №2

[minerals\_classification\_EXP2.ipynb](https://colab.research.google.com/drive/1FbSOpa10mSm1dOvFb2XpTD327lJ1dsPl?usp=sharing)

У цьому експерименті відбулась мінімальна кількість змін: в моделі провелось навчання без заморожування шарів



Зображення, що містить текст, знімок екрана, Шрифт, меню

Автоматично згенерований опис

Зображення, що містить текст, Графік, знімок екрана, ряд

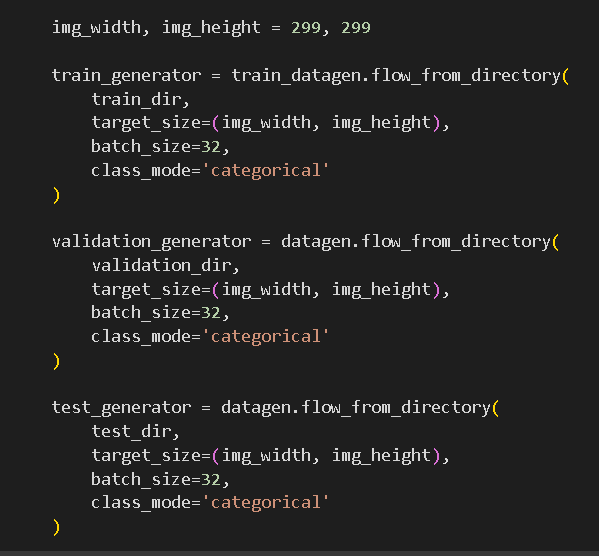
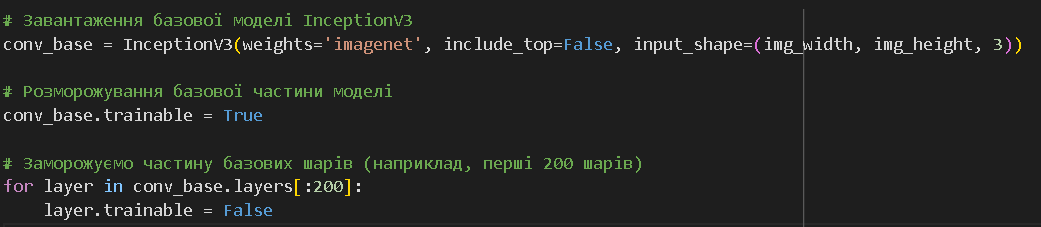
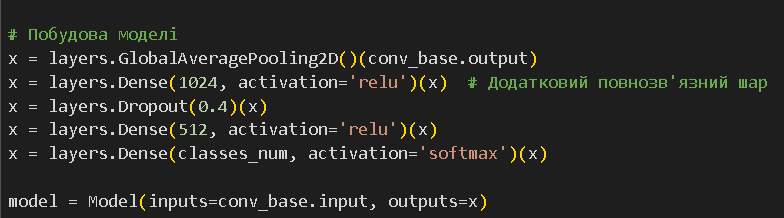
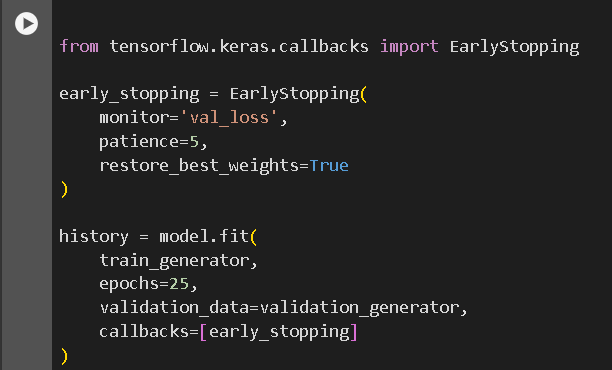
Автоматично згенерований опис

Це сприяло покращенню точності, але можна краще.

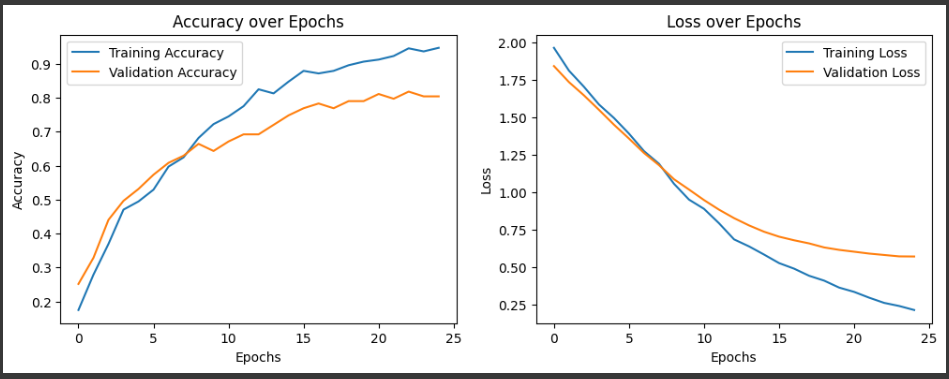
## ЕКСПЕРИМЕНТ №3

[minerals\_classification\_EXP3.ipynb](https://colab.research.google.com/drive/17NHY_TLByQKmQXhXDGiHVo1KvUbsbSzr?usp=sharing)

На цьому етапі були проведені наступні зміни:

* збільшення розміру вхідних зображень з 150х150 на 299х299
* Розмороження базової частини моделі та замороження х кількості базових шарів (в моєму випадку - 200шт)
* Інтеграція додаткового шару
* навчання моделі з частково розмороженою базовою частиною та з використанням ранньої зупинки

Результат навчання моделі: 

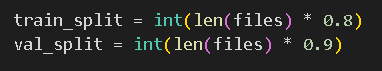
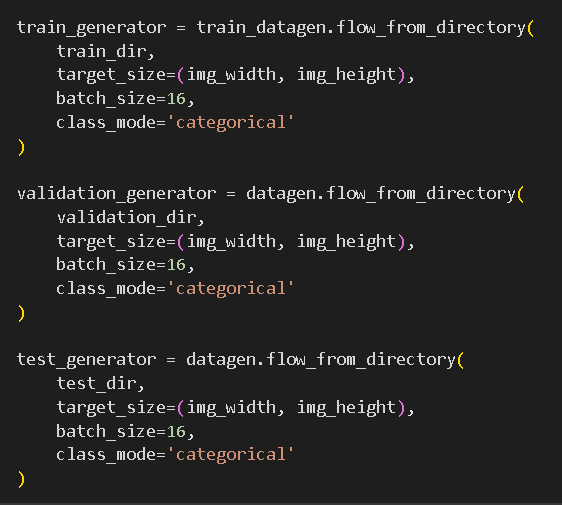
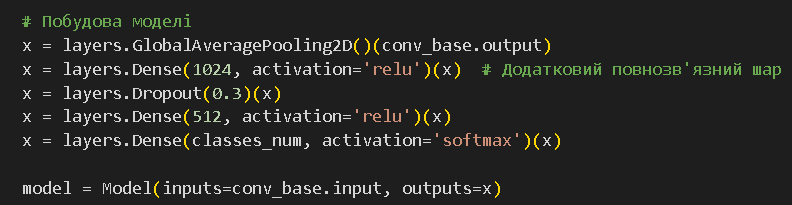
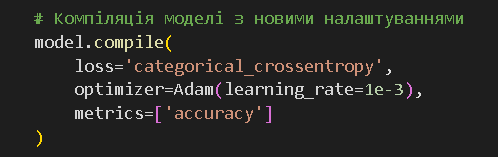
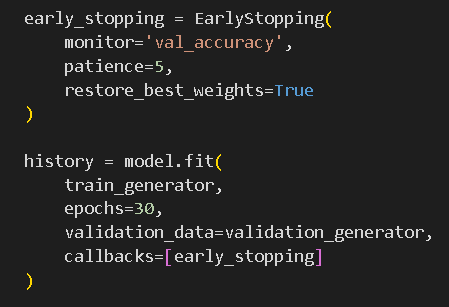


Ці зміни дали аж неочікувано гарну точність, але спробуємо покращити ще.

## ЕКСПЕРИМЕНТ №4

[minerals\_classification\_EXP4.ipynb](https://colab.research.google.com/drive/1-FouId3i7NJF823ARXr6SUI3Pl_Vjalb?usp=sharing)

У цій частині було збільшено загальну кількість зображень для тренувального та валідаційного наборів, зменшено параметр batch\_size у всіх генераторах, знижено значення Dropout, скориговано learning\_rate та додано механізм ранньої зупинки для запобігання перенавчанню (цей механізм припиняє навчання, якщо протягом 5 епох не спостерігається покращення на валідаційних даних).

* Кількість файлів, розрахованих для тренувального та валідаційних наборів (збільшено з 0.7 та 0.85 до 0.8 та 0.9)
* Зменшено значення batch\_size з 32 до 16
* Зменшення Dropout з 0.4 до 0.3
* Заміна значення learning\_rate з 1e-5 на 1e-3 щоб прискорити навчання
* Додавання раннього зупинення протягом 5 епох

Результат навчання моделі:

Зображення, що містить текст, знімок екрана, Шрифт

Автоматично згенерований опис

Зображення, що містить Графік, знімок екрана, ряд, текст

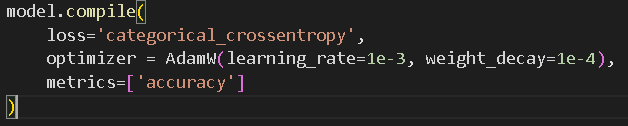
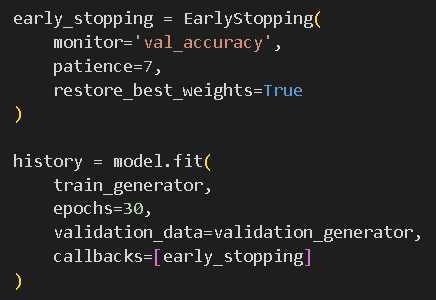
Автоматично згенерований опис

Можливо, точність моделі цього разу трохи знизилася, але ми уникнули перенавчання. До того ж процес навчання став швидшим, що покращує ефективність використання ресурсів.

## ЕКСПЕРИМЕНТ №5

[minerals\_classification\_EXP5.ipynb](https://colab.research.google.com/drive/1hEI6DYeBWHlK3gB7BO8qBNcpKN0tSlew?usp=sharing)

У фінальному експерименті були внесені незначні зміни, спрямовані на відтворення або покращення точності моделі №3 з одночасним уникненням перенавчання. Для цього значення patience було збільшено до 7, щоб запобігти передчасній зупинці навчання, і забезпечено збереження найкращої моделі завдяки параметру restore\_best\_weights=True. Крім того, заміна оптимізатора **Adam** на **AdamW** була спрямована на покращення регуляризації та зменшення ризику перенавчання, що сприяє більш стабільному процесу навчання моделі.



Результат навчання моделі: Зображення, що містить текст, знімок екрана, Шрифт

Автоматично згенерований опис

Зображення, що містить знімок екрана, Графік, текст, ряд

Автоматично згенерований опис

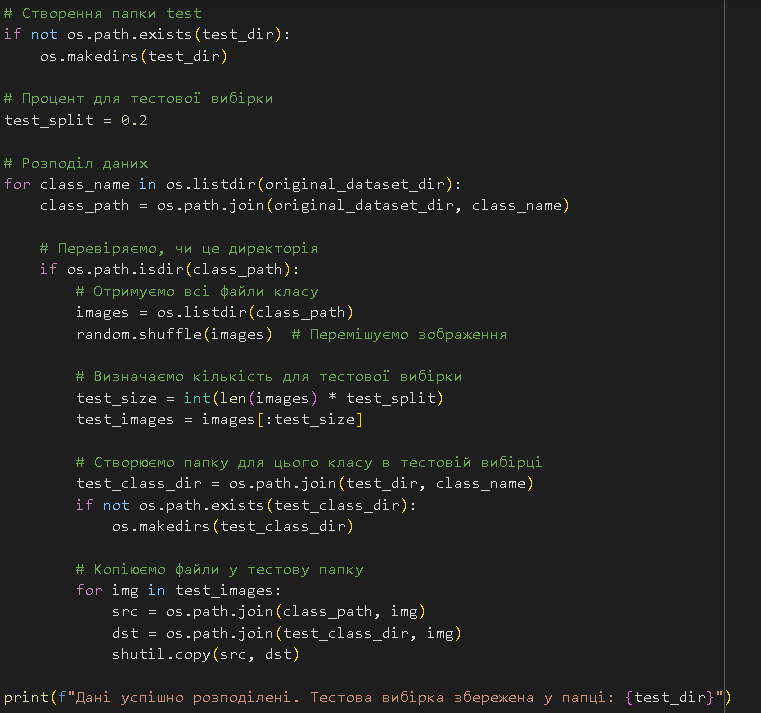
Тепер вдалось уникнути передчасного припинення та знизити ризик перенавчання.

## ПОРІВНЯННЯ МОДЕЛЕЙ

[minerals\_classification\_final.ipynb](https://colab.research.google.com/drive/1IF8bQUDXgbVBIrg8826D4zfU3Bpbwt1p?usp=sharing)

Останнім етапом проєкту є практичне порівняння моделей. Для цього було вибрано 100 фотографій мінералів із тестового набору, за якими здійснювалося оцінювання продуктивності всіх створених моделей. Результати порівнювалися на основі точності класифікації, часу обробки та ефективності виявлення складних прикладів.

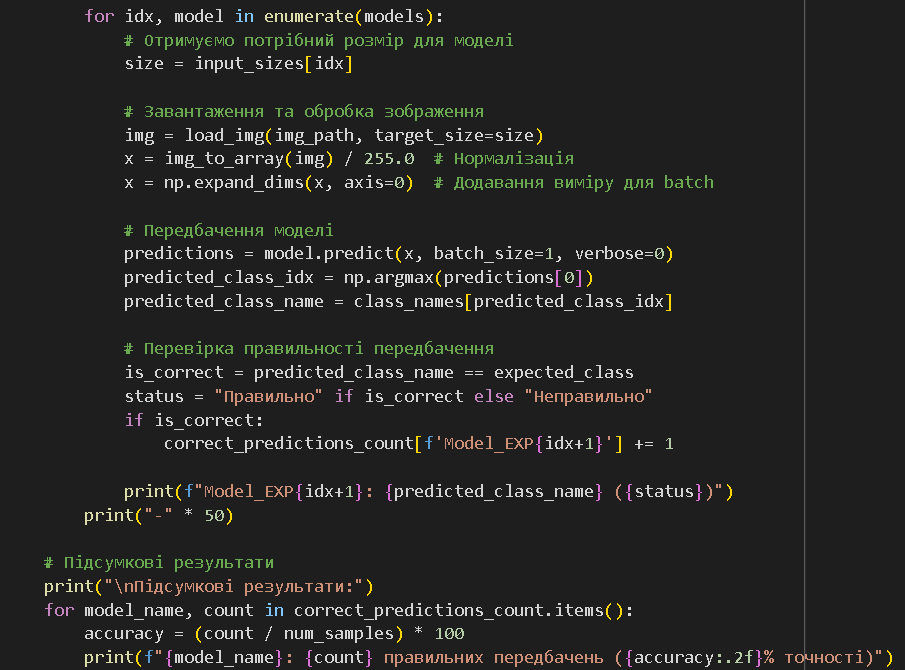
Попередньо розділяєм дані, виділяючи 20% на тестову вибірку (інші вибірки тут не потрібні).



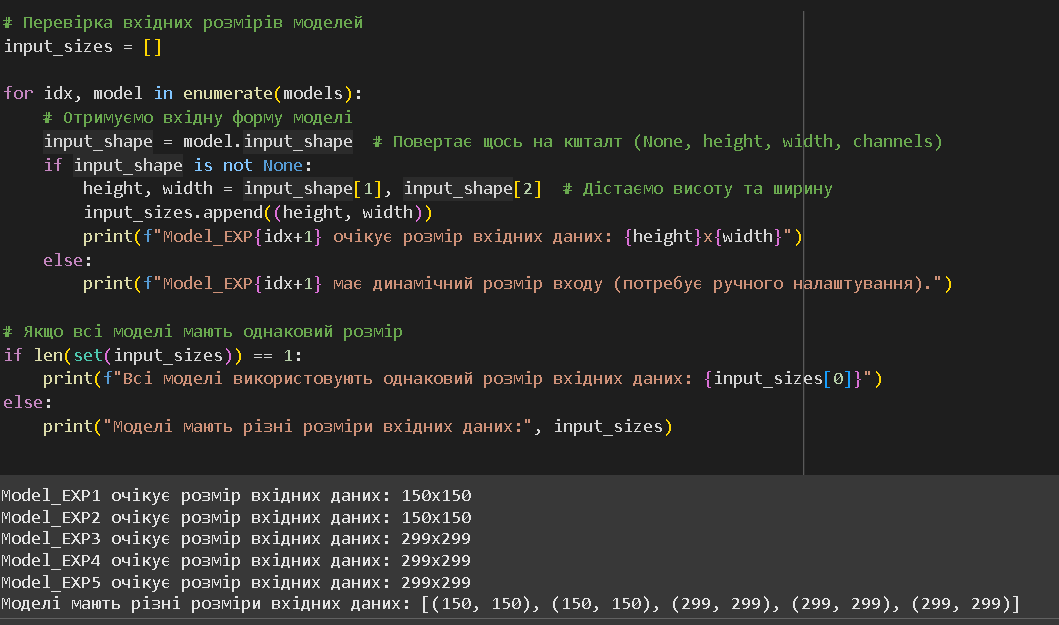
Створюємо список з видами мінералів, які є у датасеті. Потім відбувається завантаження моделей.

Зображення, що містить текст, знімок екрана, програмне забезпечення, Мультимедійне програмне забезпечення

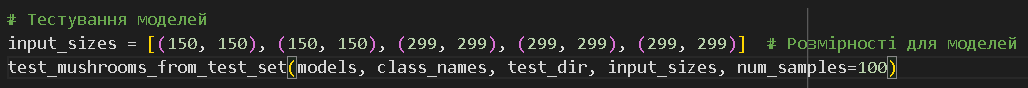
Автоматично згенерований опис



Визначаємо розмір вхідних даних для кожної моделі



Останній крок - запуск створеної функції для порівняння.



 Зображення, що містить знімок екрана, текст, Мінерал, скеля

Автоматично згенерований опис Зображення, що містить текст, Фармацевтичний препарат, таблетка, ліки

Автоматично згенерований опис Зображення, що містить риф, знімок екрана, текст, рослина

Автоматично згенерований опис Зображення, що містить знімок екрана, риф, текст, Бірюза

Автоматично згенерований опис

Зображення, що містить текст, знімок екрана, Шрифт

Автоматично згенерований опис

#### ВИСНОВКИ

У результаті п’яти експериментів була створена модель, яка продемонструвала високу точність як на тренувальних, так і на тестових даних. Модель правильно класифікувала 96 із 100 фото екземплярів, що є справді хорошим результатом.