Case Study: Classifying Snake Species in the DMV Area

Imagine you are working for a wildlife safety organization that aims to help residents of the DMV (District of Columbia, Maryland, and Virginia) quickly identify local snakes they encounter. Many snakes are harmless, but some—like copperheads—are venomous and pose serious risks if misidentified.

Your organization wants to create a **machine learning model** that can take a **photo of a snake** and **predict its species**. However, you face several real-world challenges:

- The dataset is relatively small.
- Some species look very similar to others.
- The photos vary in quality, lighting, and angle.

Your Mission: You will **preprocess** real-world image data and **train a Convolutional Neural Network (CNN)** model to classify local snake species from photos.
Your goal is to build a basic image classifier that can distinguish between different snake species **with strong accuracy** — and recommend next steps to improve it further.

You'll encounter challenges like **data augmentation**, **class imbalance**, and **model evaluation** — exactly the kinds of practical problems real-world data scientists face when building computer vision models.

At the end, you'll produce a short deliverable showing:

- How you prepared and organized the image data
- How you built and trained a CNN model
- How you evaluated your model's performance
  What improvements you would suggest for future iterations

Start your project here: https://github.com/Seminahn01/DS4002casestudy_snake