Version:1.0 StartHTML:0000000156 EndHTML:0000020901 StartFragment:0000000469 EndFragment:000002089

## HematoVision: Advanced Blood Cell Classification Using Transfer Learning

HematoVision aims to develop an accurate and efficient model for classifying blood cells by employing transfer learning techniques. Utilizing a dataset of 12,000 annotated blood cell images, categorized into distinct classes such as eosinophils, lymphocytes, monocytes, and neutrophils, the project leverages pre-trained convolutional neural networks (CNNs) to expedite training and improve classification accuracy. Transfer learning allows the model to benefit from pre-existing knowledge of image features, significantly enhancing its performance and reducing computational costs. This approach provides a reliable and scalable tool for pathologists and healthcare professionals, ensuring precise and efficient blood cell classification.

**Scenario**

1: Automated Diagnostic Systems for Healthcare

Integrating HematoVision into automated diagnostic systems in clinical settings can revolutionize blood analysis. By using transfer learning, the system quickly adapts to the specifics of blood cell classification, capturing images of blood samples, classifying the cells in real-time, and generating detailed reports. This automation reduces the manual workload on pathologists, speeds up diagnostic processes, and ensures high accuracy in results, ultimately improving patient care and treatment efficiency.

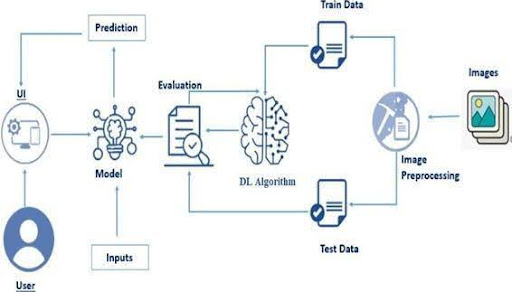
2: Remote Medical Consultations

HematoVision can be employed in telemedicine platforms to enhance remote consultations and diagnostics. With transfer learning, the model's ability to accurately classify blood cells from diverse sources is improved, allowing healthcare providers to upload blood cell images for automated analysis. This enables timely and accurate assessments without the need for in-person visits, facilitating better access to specialized medical expertise and improving healthcare delivery in remote or underserved areas.

3: Educational Tools for Medical Training

HematoVision's transfer learning-based classification model can be integrated into educational tools for medical training. By incorporating this advanced technology into interactive learning platforms, students and laboratory technicians can upload and analyze blood cell images to receive instant feedback. This hands-on learning experience enhances their understanding of blood cell morphology and classification, providing practical skills and knowledge that are crucial for accurate diagnostic practice and medical training.

**Architectecture**



**Prerequisites**

To complete this project, you must require the following software, concepts, and packages

Anaconda Navigator:

Refer to the link below to download Anaconda Navigator

Python packages:

Open anaconda prompt as administrator

Type “pip install numpy” and click enter.

Type “pip install pandas” and click enter.

Type “pip install scikit-learn” and click enter.

Type ”pip install matplotlib” and click enter.

Type ”pip install scipy” and click enter.

Type ”pip install seaborn” and click enter.

Type ”pip install tenserflow” and click enter.

Type “pip install Flask” and click enter.

**Prior Knowledge**

You must have prior knowledge of the following topics to complete this project.

DL Concepts

Neural Networks:: https://www.analyticsvidhya.com/blog/2020/02/cnn-vs-rnn-vs-mlp-analyzing-3-types-of-neural-networks-in-deep-learning/

Deep Learning Frameworks:: https://www.knowledgehut.com/blog/data-science/pytorch-vs-tensorflow

Transfer Learning: https://towardsdatascience.com/a-demonstration-of-transfer-learning-of-vgg-convolutional-neural-network-pre-trained-model-with-c9f5b8b1ab0a

VGG16: https://www.geeksforgeeks.org/vgg-16-cnn-model/

Convolutional Neural Networks (CNNs): https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/ s://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

Overfitting and Regularization:

https://www.analyticsvidhya.com/blog/2021/07/prevent-overfitting-using-regularization-techniques/

Optimizers: https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-deep-learning-optimizers/

Flask Basics: https://www.youtube.com/watch?v=lj4I\_CvBnt0

**Project Objectives**

By the end of this project, you will:

Know fundamental concepts and techniques used for Deep Learning.

Gain a broad understanding of data.

Have knowledge of pre-processing the data/transformation techniques on outliers and some visualization concepts.

⁷

**Project Flow**

The user interacts with the UI (User Interface) to choose the image.

The chosen image is analyzed by the model which is integrated with the flask application.

Once the model analyses the input the prediction is showcased on the UI

To accomplish this, we have to complete all the activities listed below,

Data Collection: Collect or download the dataset that you want to train.

Data pre-processing

Data Augmentation

Splitting data into train and test

Model building

Import the model-building libraries

Initializing the model

Training and testing the model

Evaluating the performance of the model

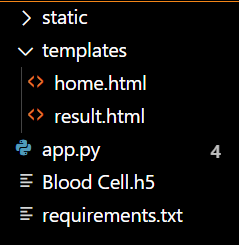
Save the model

Application Building

Create an HTML file

Build python code

**Project Structure**

Create the Project folder which contains files as shown below

We are building a Flask application with HTML pages stored in the templates folder and a Python script app.py for scripting.

Blood Cell.h5 is our saved model. Further, we will use this model for flask integration.

**Data Collection and Preparation**

ML depends heavily on data. It is the most crucial aspect that makes algorithm training possible. So, this section allows you to download the required dataset.

**Collect the dataset**

Download the dataset

There are many popular open sources for collecting the data. Eg: kaggle.com, UCI repository, etc.

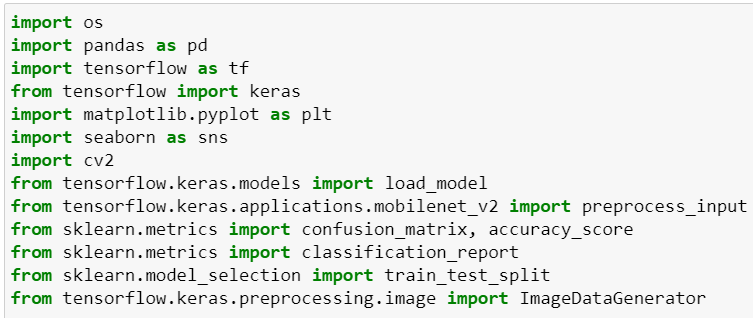
This dataset contains 12,500 augmented images of blood cells (JPEG) with accompanying cell type labels (CSV). There are approximately 3,000 images for each of 4 different cell types grouped into 4 different folders (according to cell type). The cell types are Eosinophil, Lymphocyte, Monocyte, and Neutrophil.

Link: https://www.kaggle.com/datasets/paultimothymooney/blood-cells/data

As the dataset is downloaded. Let us read and understand the data properly with the help of some visualization techniques and some analyzing techniques.

Note: There are several techniques for understanding the data. But here we have used some of it. In an additional way, you can use multiple techniques.

Activity 1.1: Importing the libraries:



Import the necessary libraries as shown in the image.

Activity 1.2: Read the Dataset:

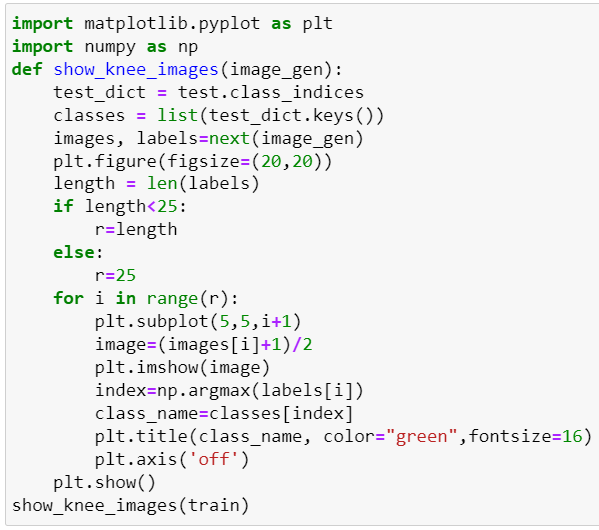
Our dataset format might be in .csv, excel files, .txt, .json, or zip files, etc. We can read the dataset with the help of pandas.

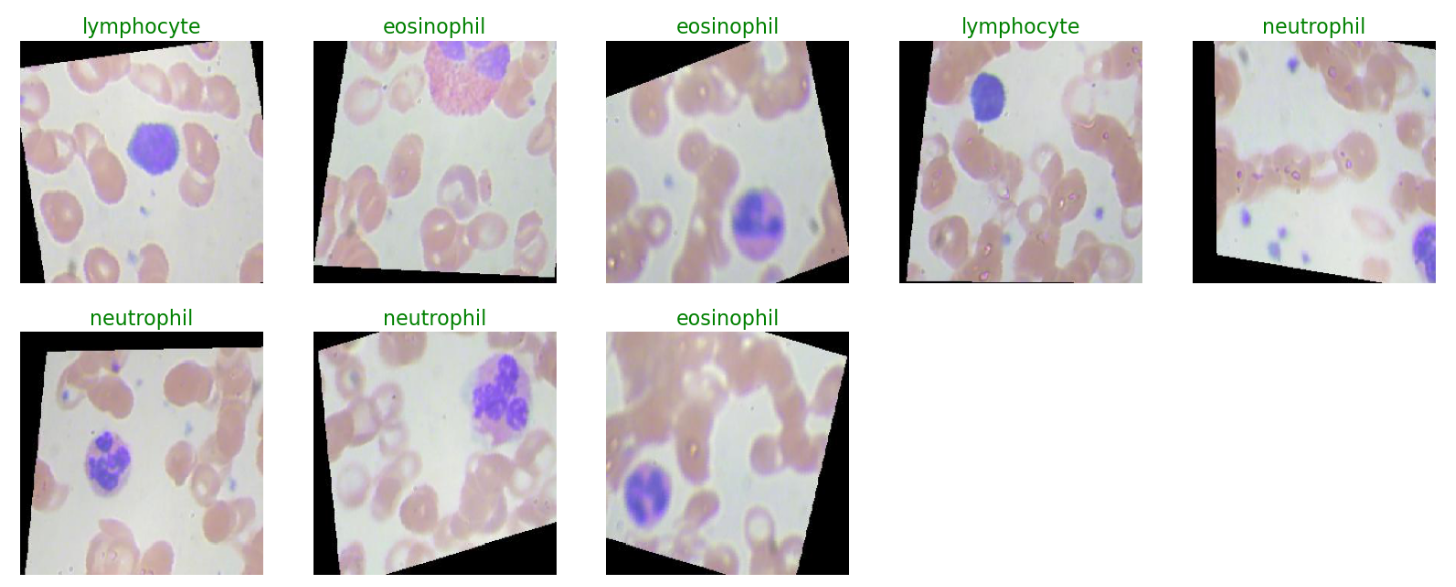
At first, unzip the data and convert it into a pandas data frame.



**Data Visualization**

The provided Python code imports necessary libraries and modules for image manipulation. It selects a random image file from a specified folder path. Then, it displays the randomly selected image using IPython's Image module. This code is useful for showcasing random images from a directory for various purposes like data exploration or testing image processing algorithms.





In the above code, I used class ace of diamond for prediction, This code randomly selects an image file from a specified folder (folder\_path) containing JPEG, PNG, or JPEG files, and then displays the selected image using IPython's display function. It utilizes Python's OS and random modules for file manipulation and random selection, respectively. And It has predicted correctly as ace of diamond.

**Data Augmentation**

Data augmentation is a technique commonly employed in machine learning, particularly in computer vision tasks such as image classification, including projects like the BloodCells Classification . The primary objective of data augmentation is to artificially expand the size of the training dataset by applying various transformations to the existing images, thereby increasing the diversity and robustness of the data available for model training. This approach is particularly beneficial when working with limited labeled data.

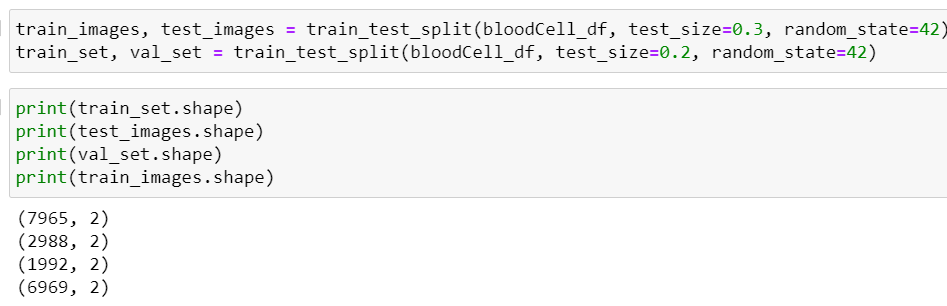
In the context of the 53 class Classification, data augmentation can involve applying transformations such as rotation, scaling, flipping, and changes in brightness or contrast to the original images of fossils. These transformations help the model generalize better to variations and potential distortions present in real-world images, enhancing its ability to accurately classify unseen data.

This is a crucial step but this data is already cropped from the augmented data so. this time it is skipped accuracy is not much affected but the training time increased.

**Split Data and Model Building**

Train-Test-Split:

In this project, we have already separated data for training and testing.



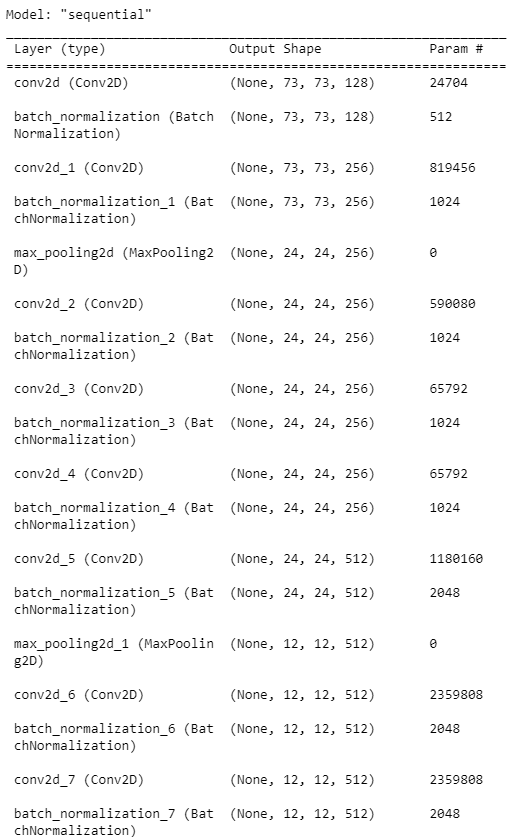


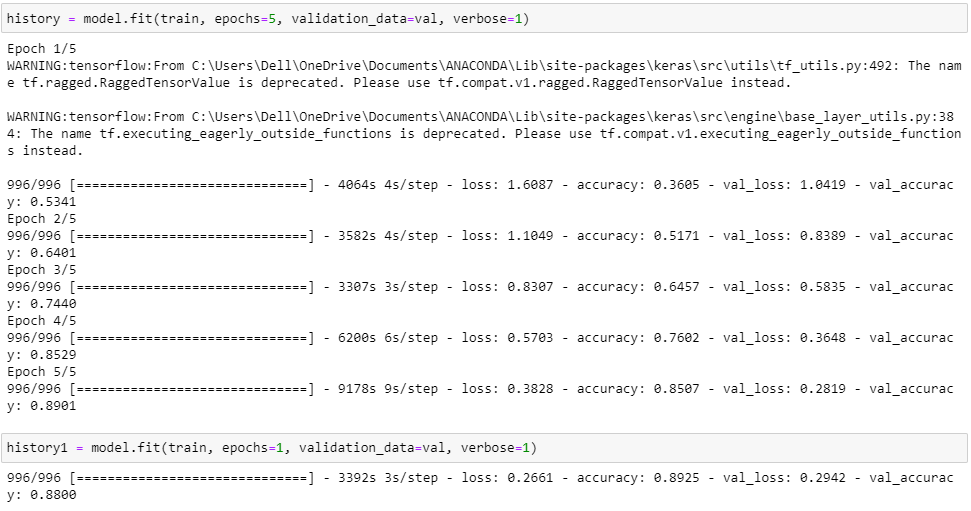
**Model Building:**

Mobilenet V2 Transfer-Learning Model:

The MobileNetV2-based neural network is created using a pre-trained MobileNetV2 architecture with frozen weights. The model is built sequentially, incorporating the MobileNetV2 base, a flattening layer, dropout for regularization, and a dense layer with SoftMax activation for classification into four categories of blood cells. The model is compiled using the Adam optimizer and categorical cross-entropy loss. During training, which spans 5 epochs, a generator is employed for the training data, and validation is conducted with callbacks such as Model Checkpoint and Early Stopping. The best-performing model is saved as "blood\_cell.h5" for future use. The model summary provides an overview of the architecture, showcasing the layers and parameters involved.

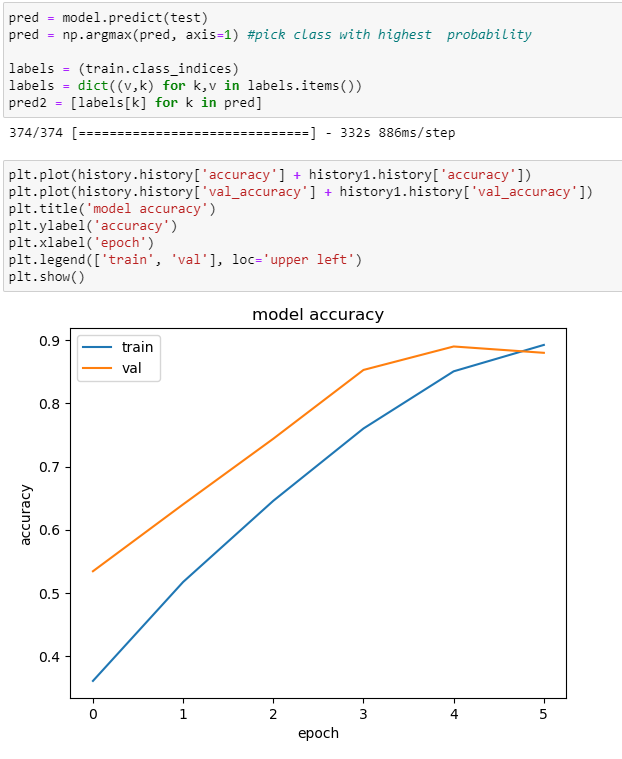






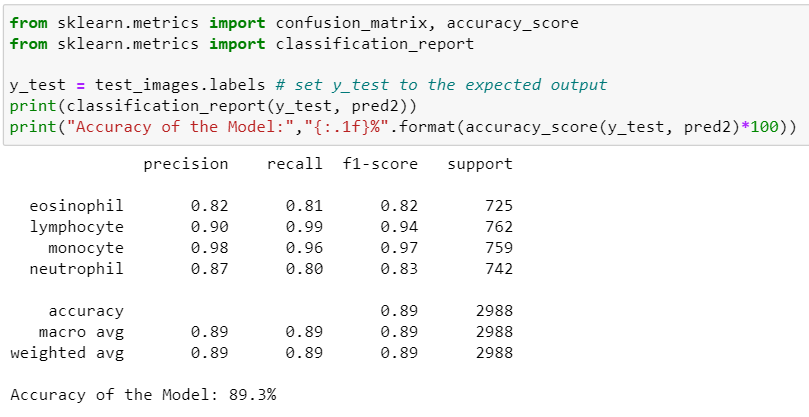
**Testing Model & Data Prediction**

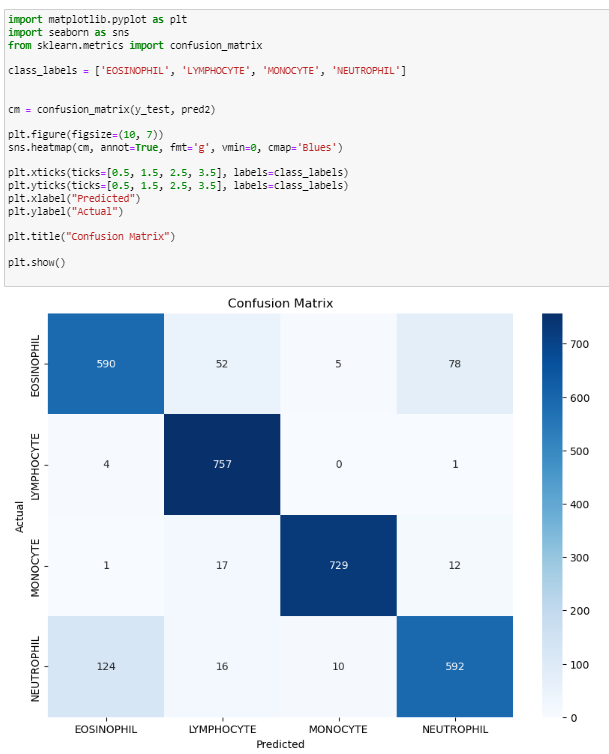
Evaluating the model



Here we have tested with the Mobilenet V2 Model With the help of the predict () function.







**Saving the model**

Finally, we have chosen the best model now saving that model



Application Building

In this section, we will be building a web application that is integrated into the model we built. A UI is provided for the uses where he has to enter the values for predictions. The enter values are given to the saved model and prediction is showcased on the UI.

This section has the following tasks

Building HTML Pages

Building server-side script

Building HTML Pages:

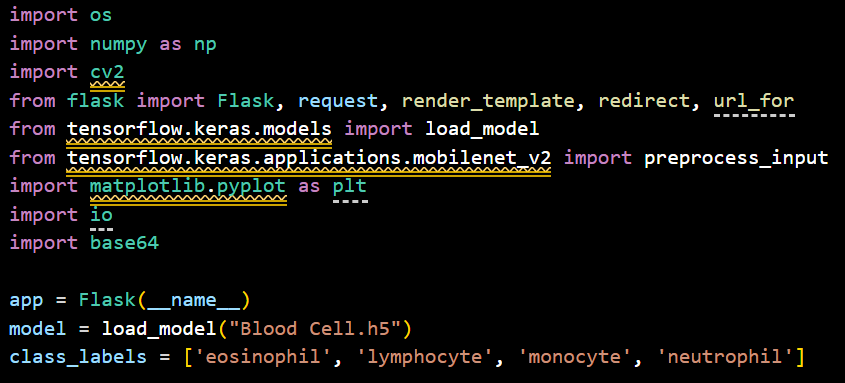
For this project create three HTML files namely

home.html

result.html

Build Python code:

Import the libraries

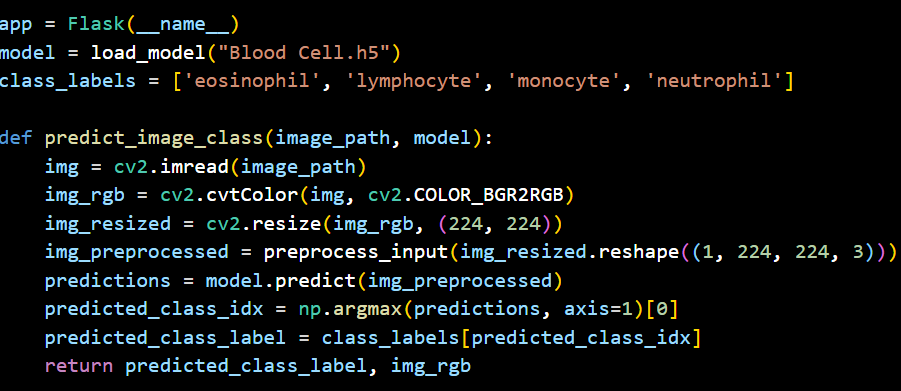


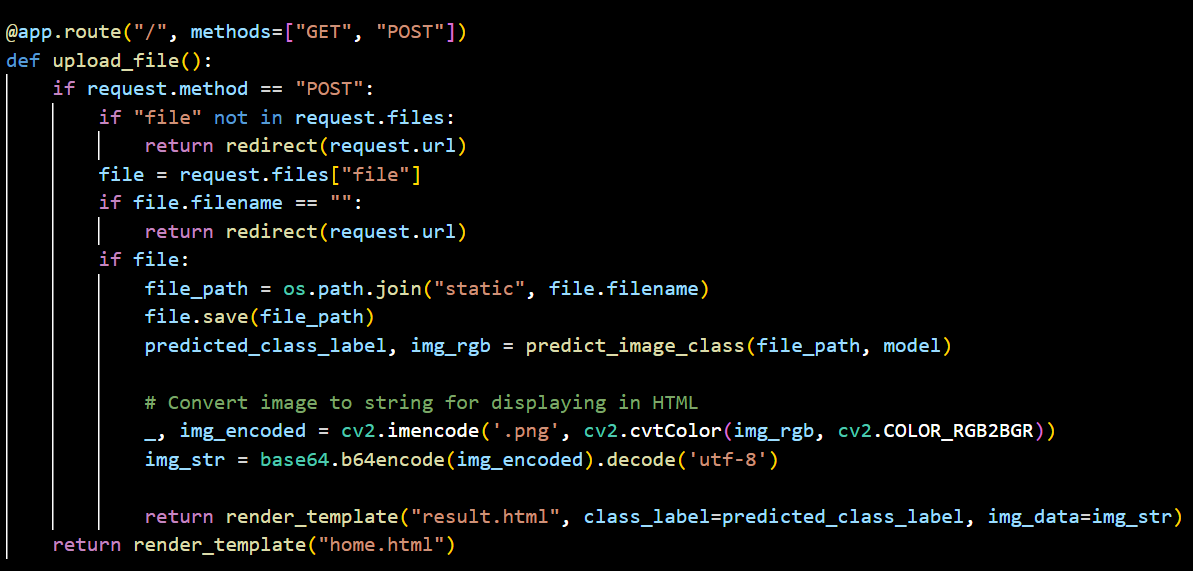
Load the saved model. Importing the Flask module in the project is mandatory. An object of the Flask class is our WSGI application. The Flask constructor takes the name of the current module (\_\_name\_\_) as argument.

Here we will be using the declared constructor to route to the HTML page which we have created earlier.

In the above example, the ‘/’ URL is bound with the index.html function. Hence, when the index page of the web server is opened in the browser, the html page will be rendered. Whenever you enter the values from the html page the values can be retrieved using POST Method.

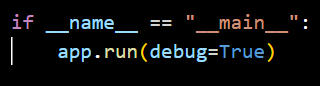
Retrieves the value from UI:





Here we are routing our app to the output() function. This function retrieves all the values from the HTML page using a Post request. That is stored in an array. This array is passed to the model. Predict () function. This function returns the prediction. This prediction value will rendered to the text that we have mentioned in the output.html page earlier.

Main Function:



**Run the web application**

: Run the application

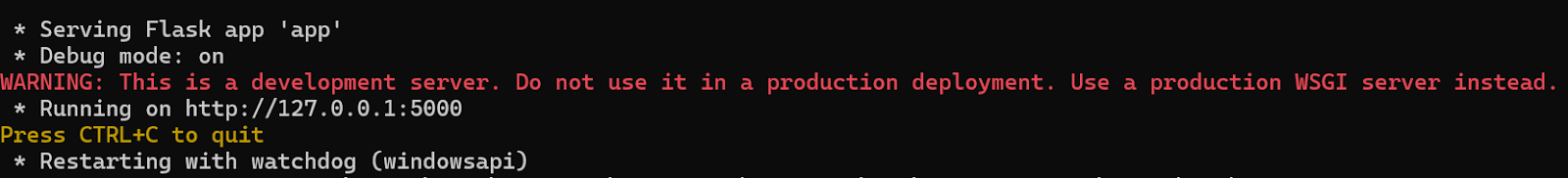
Open Anaconda prompt from the start menu

Navigate to the folder where your Python script is.

Now type the “app.py” command

Navigate to the local host where you can view your web page.

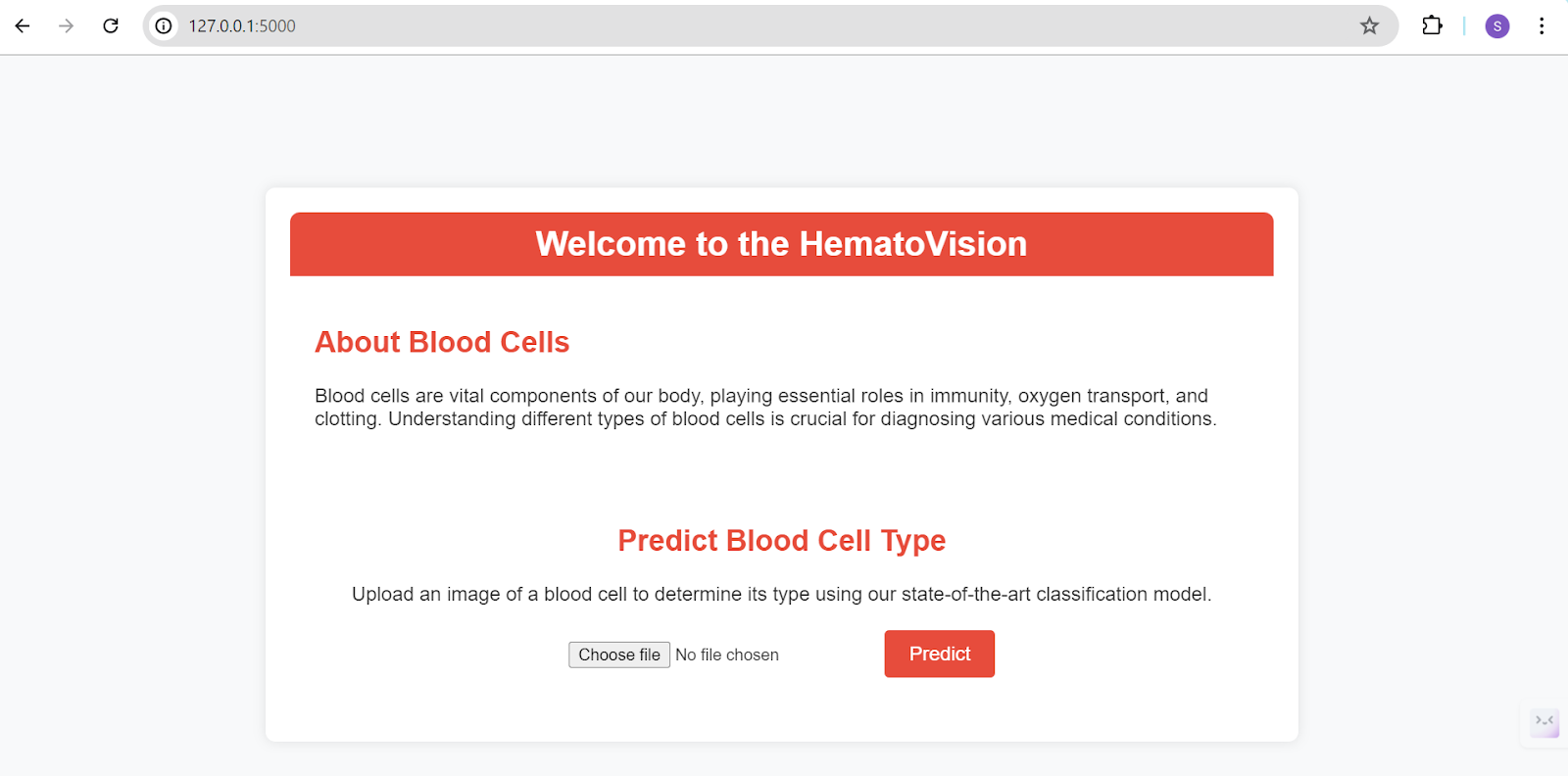
Click on the inspect button from the top right corner, enter the inputs, click on the predict button, and see the result/prediction on the web.



Now, Go the web browser and write the localhost url (http://127.0.0.1:5000) to get the below results

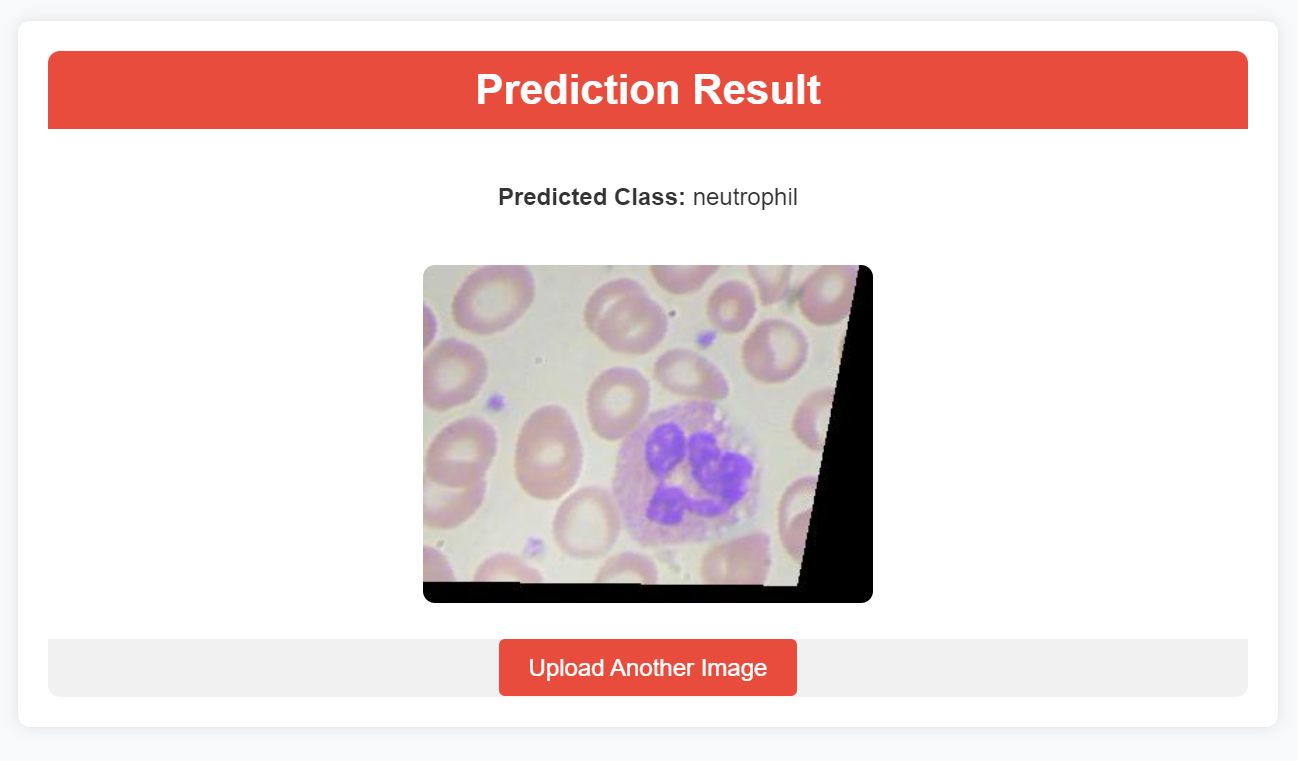
UI Image preview:

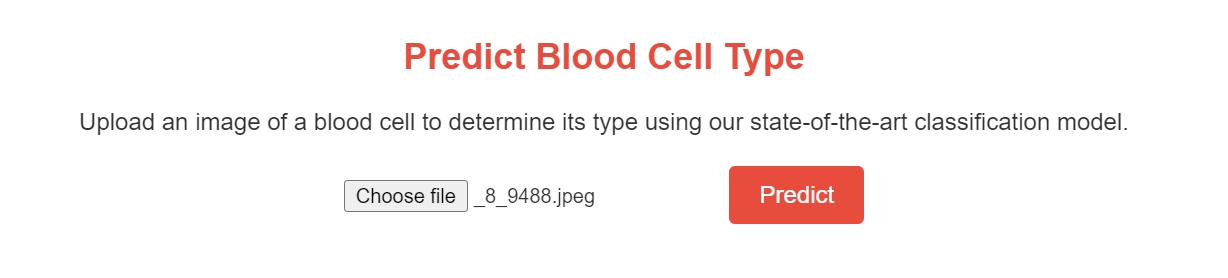
Let’s see what our index.html page looks like:



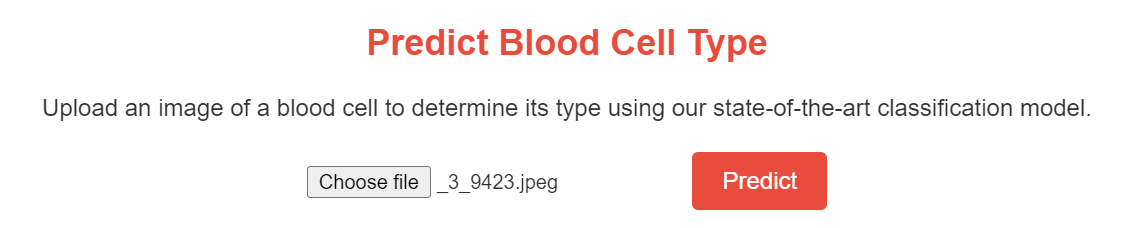
By clicking on choose file it will ask us to upload the image , then by clicking on the predict button , it will take us to the result.html

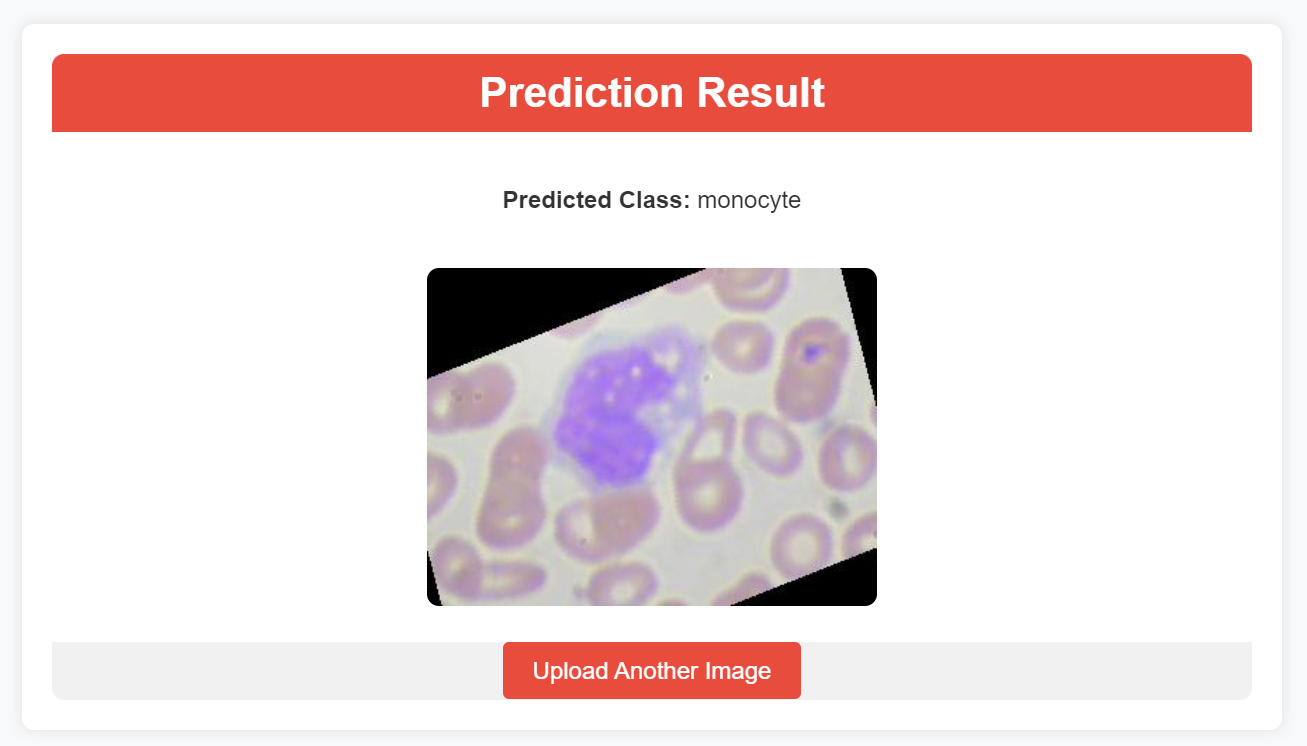
Test For Class-1 : Neutrophil





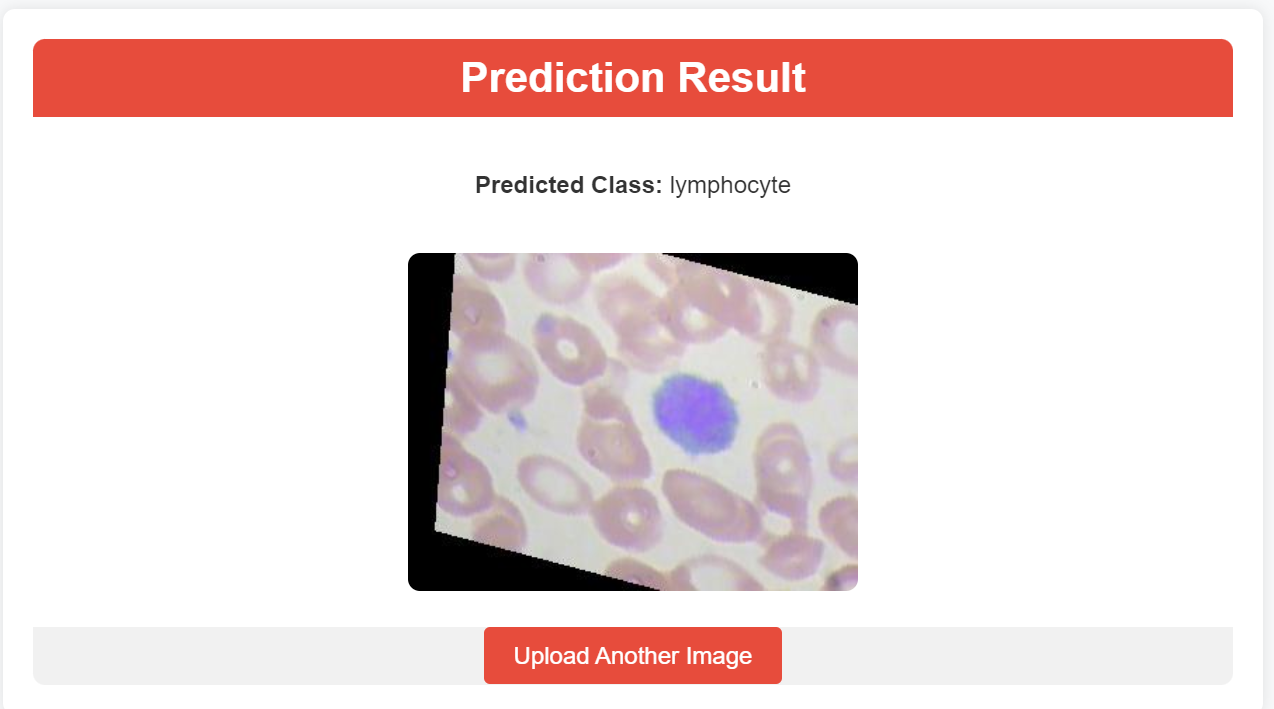
Test For Class-2 : Monocyte





Test For Class-3 : Lymphocyte





Test For Class-4 : Eosinophil

