```
1 from google.colab import drive
 2 drive.mount('/content/drive')
 3
       Mounted at /content/drive
 1 from google.colab import files
 2 files.upload()
       Choose Files kaggle.json

    kaggle.json(application/json) - 67 bytes, last modified: 11/13/2023 - 100% done

       Saving kaggle.json to kaggle.json
       { 'kaggle.ison':
       h'{"username":"aditvaadaki"."kev":"ch98516ac51f67d4df2hd5hcc7dc8d06"}'}
 1 !rm -r ~/.kaggle
 2 !mkdir ~/.kaggle
 3 !mv ./kaggle.json ~/.kaggle/
 4 !chmod 600 ~/.kaggle/kaggle.json
 5
       rm: cannot remove '/root/.kaggle': No such file or directory
 1 ! kaggle datasets list
                                                                                                                                                                                                  size lastUpdated
       masoOdahmed/netflix-movies-and-shows
                                                                                                               Netflix Movies and Shows
                                                                                                                                                                                                172KB 2023-11-19 18:
       carlmcbrideellis/llm-7-prompt-training-dataset
                                                                                                               LLM: 7 prompt training dataset
                                                                                                                                                                                                 41MB 2023-11-15 07:
       thedrcat/daigt-proper-train-dataset
                                                                                                               DAIGT Proper Train Dataset
                                                                                                                                                                                                 119MB 2023-11-05 14:
                                                                                                               DAIGT V2 Train Dataset
                                                                                                                                                                                                 29MB 2023-11-16 01:
       thedrcat/daigt-v2-train-dataset
       joebeachcapital/30000-spotify-songs
                                                                                                                30000 Spotify Songs
                                                                                                                                                                                                   3MB 2023-11-01 06:
       iamsouravbanerjee/customer-shopping-trends-dataset
                                                                                                                Customer Shopping Trends Dataset
                                                                                                                                                                                                 146KB 2023-10-05 06:
       nelgiriyewithana/world-educational-data
                                                                                                               World Educational Data
                                                                                                                                                                                                   9KB 2023-11-04 06:
                                                                                                                ₩ Healthcare Dataset 
       prasad22/healthcare-dataset
                                                                                                                                                                                                      483KB 2023-10-31
       ddosad/auto-sales-data
                                                                                                                Automobile Sales data
                                                                                                                                                                                                  79KB 2023-11-18 12:
       dillonmyrick/high-school-student-performance-and-demographics
                                                                                                               High School Student Performance & Demographics
                                                                                                                                                                                                  24KB 2023-11-10 01:
                                                                                                                                                                                                              5MB 2023-11-
       jacksondivakarr/online-shopping-dataset
                                                                                                                Online Shopping Dataset <a href="#">III</a> <a href="#">III</a>
                                                                                                                                                                                                  12KB 2023-11-14 20:
       willianoliveiragibin/animal-condition
                                                                                                                Animal Condition
       ethankeyes/nba-all-star-players-and-stats-1980-2022
                                                                                                                NBA All Star Players and Stats 1980-2022
                                                                                                                                                                                                   46KB 2023-11-16 21:
       everydaycodings/job-opportunity-dataset
                                                                                                                Job Opportunities Dataset
                                                                                                                                                                                                   95KB 2023-11-20 08:
                                                                                                               Top 100 Bestselling Book Reviews on Amazon
       anshtanwar/top-200-trending-books-with-reviews
                                                                                                                                                                                                 422KB 2023-11-09 06:
       joebeachcapital/coronavirus-covid-19-cases-daily-updates
                                                                                                               Coronavirus (COVID-19) Cases (Daily Updates)
                                                                                                                                                                                                   14MB 2023-11-23 23:
       mauryansshivam/list-of-internet-products-of-top-tech-companies
                                                                                                               List of Internet Products of Top Tech Companies
                                                                                                                                                                                                    9KB 2023-11-15 19:
                                                                                                                                                                                                   16KB 2023-11-20 12:
       vikramrn/icc-mens-odi-world-cup-wc-2023
                                                                                                                ICC mens cricket odi world cup wc 2023 - batting
       samybaladram/databank-world-development-indicators
                                                                                                                Global Socio-Economic & Demographic Insights
                                                                                                                                                                                                       2MB 2023-11-06 05
       shudhanshusingh/real-estate-properties-dataset
                                                                                                               Real Estate Properties Dataset
                                                                                                                                                                                                 882KB 2023-11-18 20:
 1 # wheat = olyadgetch/wheat-leaf-dataset 2gb
 2 #Maize = smaranjitghose/corn-or-maize-leaf-disease-dataset 169mb
 3 #rice2 = maimunulkjisan/rice-leaf-dataset-from-mendeley-data 205mb
 4 #sugarcane = prabhakaransoundar/sugarcane-disease-dataset 2gb
 5 #cotton = seroshkarim/cotton-leaf-disease-dataset 190mb
 6 dataset_name = 'maimunulkjisan/rice-leaf-dataset-from-mendeley-data'
 7 zip_name = dataset_name.split('/')[-1]
 9 !kaggle datasets download -d {dataset_name}
10
       Downloading rice-leaf-dataset-from-mendeley-data.zip to /content
         99% 193M/196M [00:02<00:00, 92.3MB/s]
       100% 196M/196M [00:02<00:00, 89.5MB/s]
 1 !unzip -q ./rice-leaf-dataset-from-mendeley-data.zip -d /content/drive/MyDrive/GROUP-4/EDI/Dataset
```

```
1 import numpy as np
 2 import time
4 import PIL.Image as Image
5 import matplotlib.pylab as plt
6
7 import tensorflow as tf
8 import tensorflow_hub as hub
10 import os
11 import shutil
12 from sklearn.model_selection import train_test_split
13 from tqdm import tqdm
14
15 import matplotlib.pyplot as plt
16
17 from sklearn.metrics import confusion matrix
18 import seaborn as sns
19
1 # directory_to_delete = "/content/drive/MyDrive/GROUP-4/EDI/split_dataset/rice_leaf/val"
3 # # Use shutil.rmtree() to delete the directory and its contents
4 # shutil.rmtree(directory_to_delete)
 6 # print(f"Directory '{directory_to_delete}' has been deleted.")
     Directory '/content/drive/MyDrive/GROUP-4/EDI/split_dataset/rice_leaf/val' has been deleted.
1 # Define the source dataset directory
 2 source_dataset_directory = "/content/drive/MyDrive/GROUP-4/EDI/Dataset/Rice Leaf Disease Images"
4 \# Define the output directory where the split dataset will be saved
5 output_directory = "/content/drive/MyDrive/GROUP-4/EDI/split_dataset/rice_leaf"
7 # Get the list of all classes (subdirectories) in the source directory
8 classes = [d for d in os.listdir(source_dataset_directory) if os.path.isdir(os.path.join(source_dataset_directory, d)) and d not in ['tr
10 # Create the output directory if it doesn't exist
11 os.makedirs(output_directory, exist_ok=True)
12
13 # Iterate over each class
14 for class name in classes:
      # Define the path to the class directory in the source dataset
15
16
      class_source_directory = os.path.join(source_dataset_directory, class_name)
17
18
      # Get the list of all files (images) in the class directory
      all_files = [f for f in os.listdir(class_source_directory) if os.path.isfile(os.path.join(class_source_directory, f))]
19
20
      # Check if the dataset for the current class is not empty
21
22
      if not all_files:
23
          print(f"Warning: The dataset for class {class_name} is empty.")
24
       else:
           # Split the dataset for the current class into train and validation sets
25
26
           train_files, val_files = train_test_split(all_files, test_size=0.3, random_state=42)
27
28
           # Function to copy files from source to destination
29
           def copy_files(file_list, src_dir, dest_dir):
               for file in tqdm(file_list, desc=f"Copying files for class {class_name}"):
30
31
                   src_path = os.path.join(src_dir, file)
32
                   dest path = os.path.join(dest dir, file)
33
                   shutil.copy(src_path, dest_path)
34
35
           # Copy train files to the train folder
           train_directory = os.path.join(output_directory, 'train', class_name)
36
37
           os.makedirs(train_directory, exist_ok=True)
38
           copy_files(train_files, class_source_directory, train_directory)
39
           # Copy validation files to the val folder
40
41
           val_directory = os.path.join(output_directory, 'val', class_name)
42
           os.makedirs(val directory, exist ok=True)
43
           copy_files(val_files, class_source_directory, val_directory)
44
```

```
1108/1108 [00:10<00:00, 108.42it/s]
    Copying files for class Bacterialblight: 100%
    Copying files for class Bacterialblight: 100%
                                                         | 476/476 [00:03<00:00, 120.61it/s]
                                               | 1008/1008 [00:09<00:00, 107.26it/s]
    Copying files for class Blast: 100%
    Copying files for class Blast: 100%
                                               432/432 [00:03<00:00, 124.39it/s]
    Copying files for class Brownspot: 100%
                                                  | 1120/1120 [00:10<00:00, 109.76it/s]
    Copying files for class Brownspot: 100%
                                                   480/480 [00:03<00:00, 121.97it/s]
    Copying files for class Tungro: 100%
                                                915/915 [00:08<00:00, 112.06it/s]
    Copying files for class Tungro: 100%
                                               393/393 [00:03<00:00, 124.62it/s]
1 batch_size = 32
2 img_height = 224
3 \text{ img\_width} = 224
4 data_root = '/content/drive/MyDrive/GROUP-4/EDI/split_dataset/wheat_leaf/train'
6 train_ds = tf.keras.preprocessing.image_dataset_from_directory(
   str(data_root),
8 validation_split=0.2,
9 subset="training",
10 seed=123,
11
    image_size=(img_height, img_width),
12
   batch_size=batch_size)
    Found 283 files belonging to 3 classes.
    Using 227 files for training.
1 class_names = np.array(train_ds.class_names)
2 print('class names for predictions :',class_names)
    class names for predictions : ['Healthy' 'septoria' 'stripe_rust']
1 normalization_layer = tf.keras.layers.experimental.preprocessing.Rescaling(1./255)
2 train_ds = train_ds.map(lambda x, y: (normalization_layer(x), y))
1 AUTOTUNE = tf.data.AUTOTUNE
2 train_ds = train_ds.cache().prefetch(buffer_size=AUTOTUNE)
1 for image_batch, labels_batch in train_ds:
2 print(image_batch.shape)
   print(labels_batch.shape)
4
   break
    (32, 224, 224, 3)
    (32,)
1 feature_extractor_model = "https://tfhub.dev/google/tf2-preview/mobilenet_v2/feature_vector/4"
2 feature_extractor_layer = hub.KerasLayer(
      feature_extractor_model, input_shape=(224, 224, 3), trainable=False)
1 num_classes = len(class_names)
2
3 model = tf.keras.Sequential([
   feature_extractor_layer,
5
   tf.keras.layers.Dense(num_classes)
6])
8 model.summary()
    Model: "sequential_4"
                                                        Param #
     Layer (type)
                               Output Shape
    _____
     keras_layer_4 (KerasLayer) (None, 1280)
                                                        2257984
     dense_4 (Dense)
                               (None, 3)
                                                        3843
    ______
    Total params: 2261827 (8.63 MB)
    Trainable params: 3843 (15.01 KB)
    Non-trainable params: 2257984 (8.61 MB)
```

1 image_batch.shape

```
TensorShape([32, 224, 224, 3])
1 model.compile(
2 optimizer=tf.keras.optimizers.Adam(),
   loss=tf.keras.losses.SparseCategoricalCrossentropy(from logits=True),
   metrics=['acc'])
1 class CollectBatchStats(tf.keras.callbacks.Callback):
   def __init__(self):
3
     self.batch_losses = []
4
     self.batch_acc = []
5
   def on_train_batch_end(self, batch, logs=None):
6
7
     self.batch_losses.append(logs['loss'])
     self.batch_acc.append(logs['acc'])
8
9
     self.model.reset_metrics()
10
11 batch_stats_callback = CollectBatchStats()
13 history = model.fit(train_ds, epochs=5,
14
                   callbacks=[batch_stats_callback])
    Epoch 1/5
    8/8 [=====
             Epoch 2/5
    Epoch 3/5
    8/8 [=====
                 =========] - 0s 41ms/step - loss: 0.0000e+00 - acc: 0.0000e+00
    Epoch 4/5
                   =========] - 0s 39ms/step - loss: 0.0000e+00 - acc: 0.0000e+00
    8/8 [=====
    Epoch 5/5
    8/8 [=============== ] - 0s 40ms/step - loss: 0.0000e+00 - acc: 0.0000e+00
1 plt.figure()
2 plt.ylabel("Loss")
3 plt.xlabel("Training Steps")
4 plt.ylim([0,2])
5 plt.plot(batch_stats_callback.batch_losses)
    [<matplotlib.lines.Line2D at 0x7c4b1dfe1bd0>]
       2.00
       1.75
       1.50
       1.25
     S 1.00
       0.75
       0.50
       0.25
```

```
1 plt.figure()
2 plt.ylabel("Accuracy")
3 plt.xlabel("Training Steps")
4 plt.ylim([0,1])
5 plt.plot(batch_stats_callback.batch_acc)
```

5

10

15

20

Training Steps

25

0.00

0

30

35

40

[<matplotlib.lines.Line2D at 0x7c4b1c5607c0>]

1.0

0.8
0.6
0.4
0.2-

```
1 import matplotlib.pyplot as plt
2
3 # Access accuracy values collected during training
4 batch_acc = batch_stats_callback.batch_acc
5
6 # Plot the accuracy matrix
7 plt.figure(figsize=(10, 6))
8 plt.plot(batch_acc, label='Training Accuracy')
9 plt.xlabel('Training Steps')
10 plt.ylabel('Accuracy')
11 plt.title('Training Accuracy Over Time')
12 plt.legend()
13 plt.grid(True)
14 plt.show()
15
```



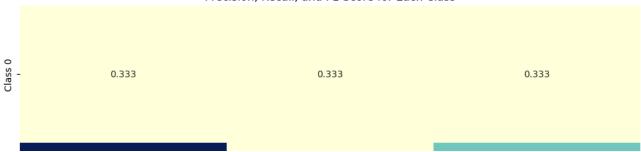
16

```
1 # Load and preprocess validation dataset
 2 data_root_val = '/content/drive/MyDrive/GROUP-4/EDI/split_dataset/wheat_leaf/val'
 4 val_ds = tf.keras.preprocessing.image_dataset_from_directory(
 5 str(data_root),
 6 validation split=0.2,
    subset="validation", # Use validation subset
 8 seed=123,
 9 image_size=(img_height, img_width),
10 batch_size=batch_size)
11
12 val_ds = val_ds.map(lambda x, y: (normalization_layer(x), y))
13 val_ds = val_ds.cache().prefetch(buffer_size=AUTOTUNE)
15 # Evaluate the model on the validation dataset
16 val_loss, val_accuracy = model.evaluate(val_ds)
17
18 print(f"Validation Loss: {val_loss * 100:.2f}%")
19 print(f"Validation Accuracy: {val_accuracy * 100:.2f}%")
20
     Found 283 files belonging to 3 classes.
    Using 56 files for validation.
     2/2 [============] - 15s 517ms/step - loss: 0.4126 - acc: 0.8750
     Validation Loss: 41.26%
    Validation Accuracy: 87.50%
 1 true labels = []
 2 predicted_labels = []
 4 for images, labels in val_ds:
 5
      true_labels.extend(labels.numpy()) # Collect true labels
      predictions = model.predict(images)
 6
      predicted_labels.extend(tf.argmax(predictions, axis=1).numpy()) # Collect predicted labels
     1/1 [======] - 0s 465ms/step
    1/1 [=======] - 1s 518ms/step
 1 import numpy as np
 2 import matplotlib.pyplot as plt
 3 from sklearn.metrics import confusion_matrix
 4 import seaborn as sns
 6 # Calculate confusion matrix
 7 cm = confusion_matrix(true_labels, predicted_labels)
 9 # Plot confusion matrix as a heatmap
10 plt.figure(figsize=(8, 6))
11 sns.heatmap(cm, annot=True, fmt="d", cmap="Blues", xticklabels=class_names, yticklabels=class_names)
12 plt.xlabel('Predicted')
13 plt.ylabel('True')
14 plt.title('Confusion Matrix')
15 plt.show()
```

 \Box

```
Confusion Matrix
                                                                                     20.0
                                                                                     17.5
        Healthy
                                                                  1
                                            2
1 # Specify the path for saving the model in Google Drive
2 h5_export_path = "/content/drive/MyDrive/wheat_model.h5"
3
4 # Save the model as an HDF5 file
5 model.save(h5_export_path)
6
7 print(f"Model saved as {h5_export_path}")
8
     /usr/local/lib/python3.10/dist-packages/keras/src/engine/training.py:3079: UserWarning: You are saving your model as an HDF5 file via `
      saving_api.save_model(
     Model saved as /content/drive/MyDrive/wheat_model.h5
        ۵
1 import numpy as np
 2 import matplotlib.pyplot as plt
3 from sklearn.metrics import confusion_matrix, precision_score, recall_score, f1_score
4 import seaborn as sns
5 import pandas as pd # Import pandas library
6
7 # Generate example data (replace this with your actual data)
8 np.random.seed(42)
9 true_labels = np.random.randint(0, 3, 100)
10 predicted_labels = np.random.randint(0, 3, 100)
11 class_names = ['Class 0', 'Class 1', 'Class 2']
12
13 # Calculate confusion matrix
14 cm = confusion_matrix(true_labels, predicted_labels)
15
16 # Calculate precision, recall, and f1 score for each class
17 precision = precision_score(true_labels, predicted_labels, average=None)
18 recall = recall_score(true_labels, predicted_labels, average=None)
19 f1 = f1_score(true_labels, predicted_labels, average=None)
21 # Plot precision, recall, and f1 score in a table
22 plt.figure(figsize=(12, 8))
23 table_data = {'Class Name': class_names, 'Precision': precision, 'Recall': recall, 'F1 Score': f1}
24 sns.heatmap(pd.DataFrame(table_data).set_index('Class Name'), annot=True, cmap="YlGnBu", fmt=".3f", cbar=False)
25 plt.xlabel('Metrics')
26 plt.title('Precision, Recall, and F1 Score for Each Class')
27 plt.show()
28
```

Precision, Recall, and F1 Score for Each Class



```
1 import numpy as np
 2 import matplotlib.pyplot as plt
 3 from sklearn.metrics import confusion_matrix, precision_score, recall_score
 4 import pandas as pd
 6 # Calculate confusion matrix
 7 cm = confusion_matrix(true_labels, predicted_labels)
 9 # Calculate precision and recall for each class as float values
10 precision = precision_score(true_labels, predicted_labels, average=None)
11 recall = recall_score(true_labels, predicted_labels, average=None)
12
13 # Multiply precision and recall by 100 and format to two decimal places
14 precision = ["{:.2f}".format(p * 100) for p in precision]
15 recall = ["{:.2f}".format(r * 100) for r in recall]
17 # Create a table with reduced width
18 plt.figure(figsize=(4, 4)) # Smaller table width
19 table_data = {'Class Name': class_names, 'Precision': precision, 'Recall': recall}
20 ax = plt.subplot(111, frame_on=False) # Remove frame around table
21 ax.xaxis.set_visible(False)
22 ax.yaxis.set_visible(False)
23
24 # Convert the column labels to a list
25 col_labels = list(table_data.keys())
26
27 # Use col_labels when creating the table
28 ax.table(cellText=pd.DataFrame(table_data).values, colLabels=col_labels, cellLoc='center', loc='center')
30 plt.title('Precision and Recall for Each Class')
31 plt.show()
32
```

Precision and Recall for Each Class

| Class Name | Precision | Recall |
|------------|-----------|--------|
| Class 0 | 15.38 | 20.00 |
| Class 1 | 40.00 | 30.77 |
| Class 2 | 16.67 | 20.83 |
| class 3 | 29.17 | 23.33 |