Project Presentation

Image classification using convolutional neural network

Presented by: Joshua Magboma

Problem statement

- The objective of this project is to develop a highly accurate image classification model using the CIFAR-100 dataset. The project aims to leverage advanced deep learning techniques to build and optimize a model that achieves superior classification accuracy and can classify any uploaded images into the defined classes.

The cifar100 dataset

The CIFAR-100 dataset is a collection of small, labeled images used to and test machine learning models

WHAT IS IN THE CIFAR-100 DATASET?

- TOTAL IMAGES: 60,000
- IMAGE SIZE: 32X32
- CLASSES: 100
- LABELED DATA: TAGS FOR EACH CLASSES

DATA COLLECTION

The data was downloaded using the TensorFlow inbuilt function call.

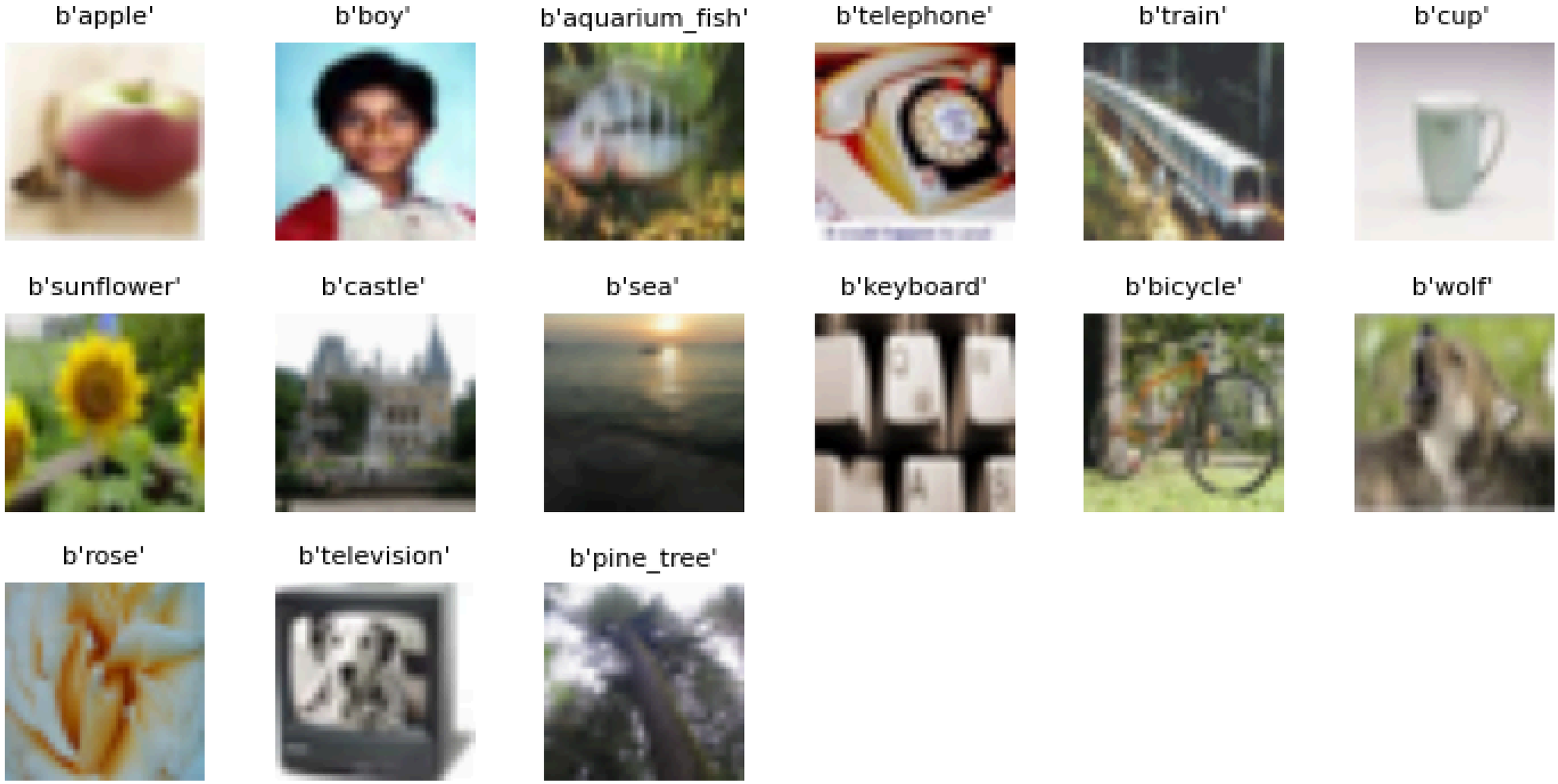
STRUCTURE:

- **X_TRAIN: 50,000 TRAINING IMAGES (32X32X3)**
- **Y_TRAIN: 50,000 CORRESPONDING LABELS 0-99**
- **X_TEST: 10,000 TEST IMAGES**
- **Y_TEST: 10,000 CORRESPONDING FOR TESTING**

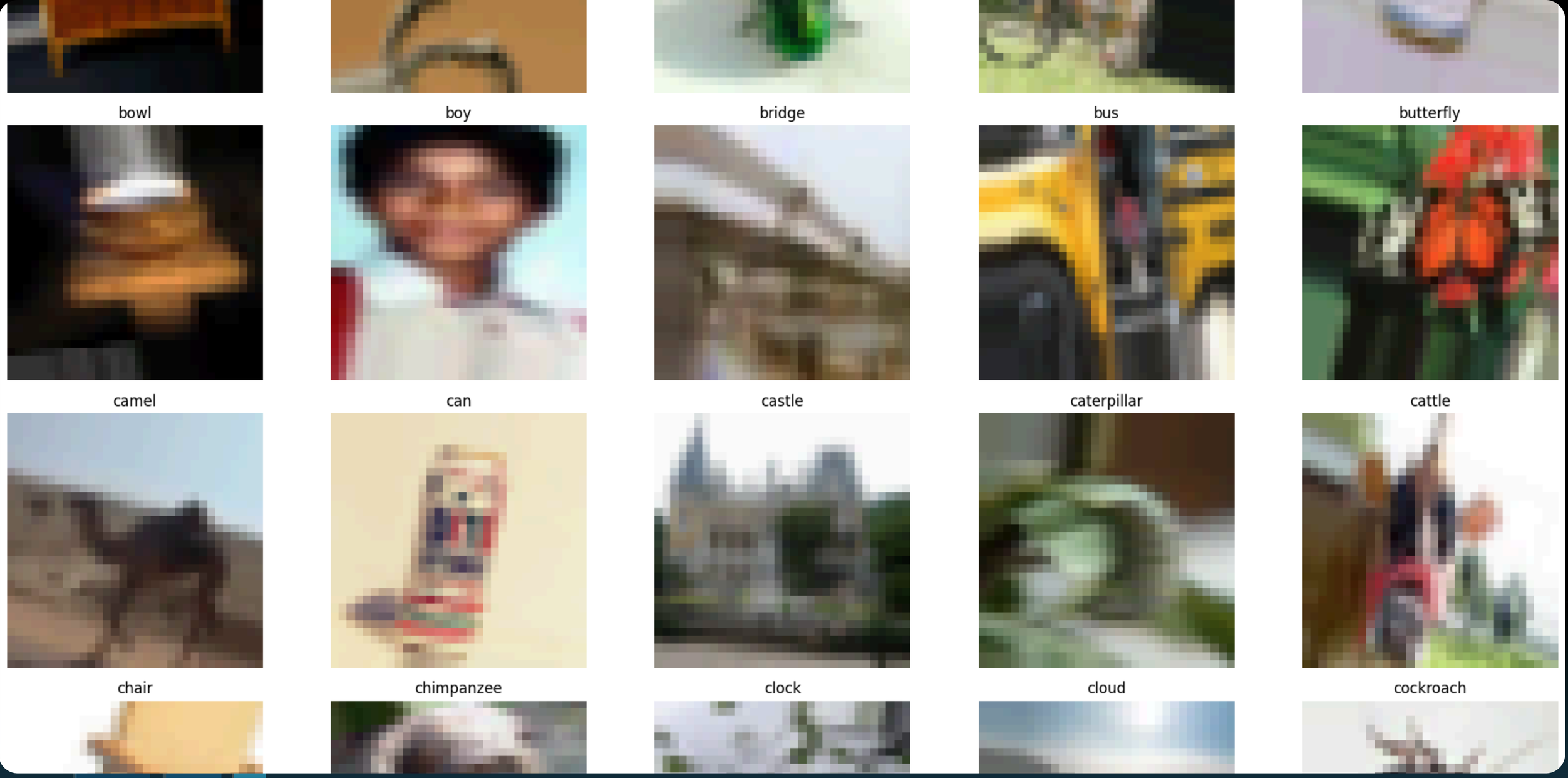
Data preprocessing

NORMALIZATION: IT IS THE PROCESS OF SCALING THE DATA SO THAT ALL THE VALUES CAN BE IN A SPECIFIC RANGE OF -1,1 OR 0,1

AUGMENTATION: IT IS THE PROCESS OF ADDING DIFFERENT VARIATIONS TO AN IMAGE WITHOUT INTRODUCING NEW IMAGES TO THE DATASET. THIS CAN INCLUDE, ZOOM, ROTATION, FLIPPING, BRIGHTNESS AND CONTRAST .



Sample Images from the cifar dataset



AUGMENTED SAMPLE IMAGE FROM CIFAR-100

Modelling

It is the process of creating a machine learning/Deep learning model to learn patterns in a dataset and make predictions

BASE MODEL

A baseline deep learning model is a simple, initial version of a deep learning model used as a benchmark to compare the performance of more complex models. It provides a reference point, helping you understand if more advanced models are actually improving performance.

Convolutional Neural Network

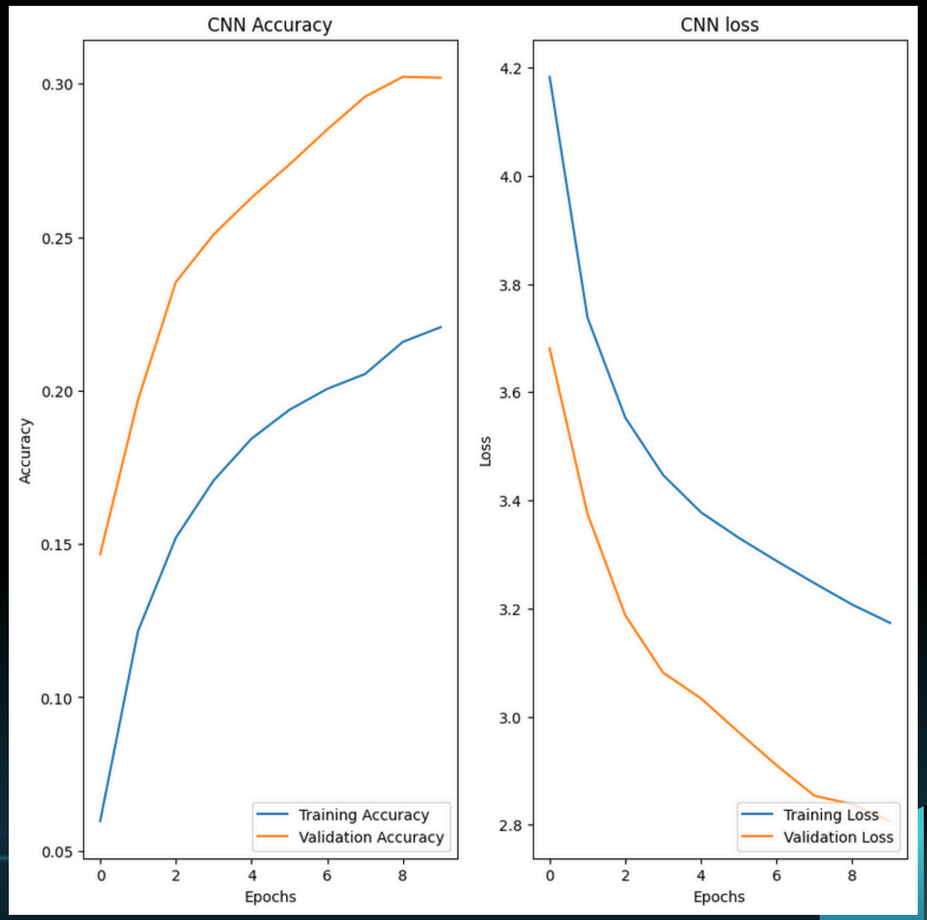
This consists of the following layers

DROPOUT	**RANDOMLY TURNS OFF SOME NEURONS DURING TRAINING**	**TURNS OFF 50% OF THE NEURONS DURING TRAINING**
CONV2D	IT EXTRACT THE FEATURES FROM THE IMAGES SUCH AS TEXTURE, PATTERN AND EDGES	IT HAS 32 FILTERS WITH A HEIGHT OF 2 AND WIDTH OF 2
MAXPOOLING2D	REDUCE THE SPATIAL DIMENSION OF THE FEATURE MAP	IT REDUCES THE WIDTH TO 2 AND HEIGHT TO 2
FLATTEN	CONVERT MULTI DIMENSIONAL INTO A SINGLE DIMENSIONAL VECTOR	
DENSE	IN THIS LAYER EVERY INPUT IS CONNECTED TO EVERY LAYER	

Convolutional neural network

Base model

ACCURACY SCORE	PRECISION SCORE	RECALL SCORES
30.20	29.40	30.20
EPOCHS=10		



TRANSFER LEARNING(RESNET50)

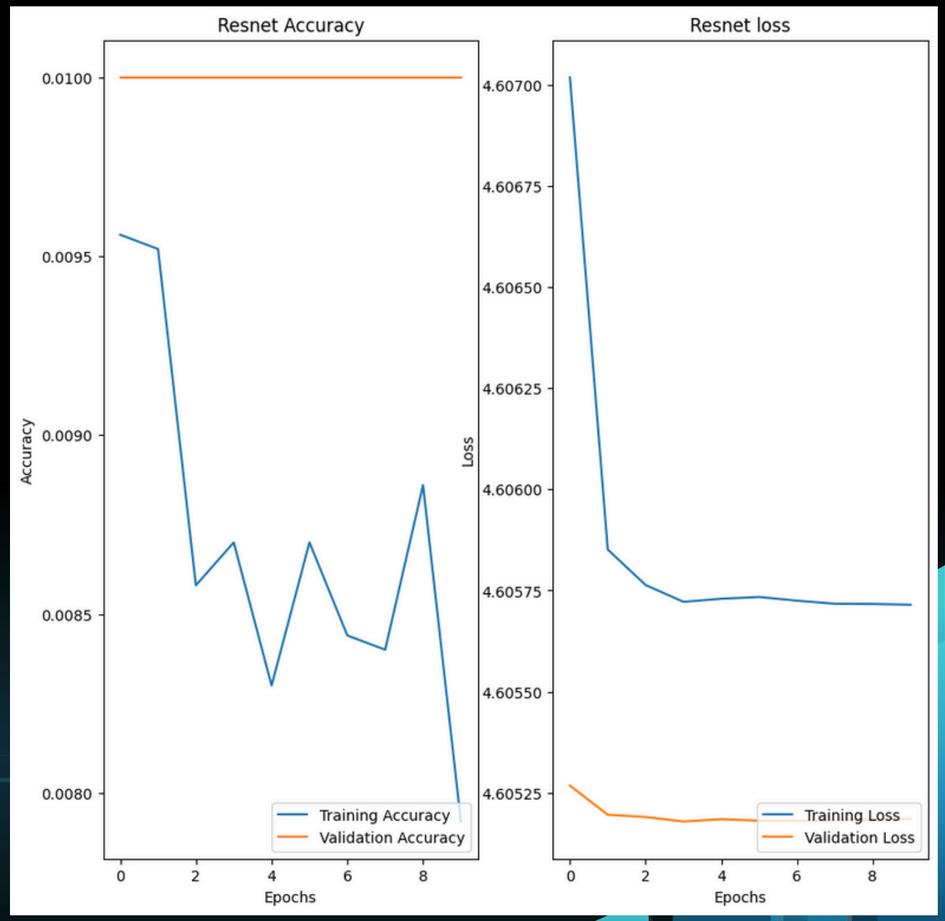
TRANSFER LEARNING IS THE PROCESS OF USING PRE-TRAINED MODEL TRAINED ON A LARGE DATASET IS USED AS THE STARTING POINT FOR A NEW TASK.

LAYERS	EXPLAINATION
EARLY LAYERS	LEARNS THE COLOR, TEXTURE OF THE IMAGE
MIDDLE LAYERS	LEARNS THE SHAPE AND STRUCTURE OF THE OBJECT IN THE IMAGE
DEEP LAYERS	RECOGNIZE THE WHOLE IMAGE E.G AN APPLE

RESNET50

LAYERS == FREEZE: THIS IMPLIES THAT THE LAYERS WILL NOT LEARN ANYTHING NEW DURING TRAINING, THE MODEL MAINTAINS THE KNOWLEDGEIN THIS LAYERS AND FOCUS ON THE NEW LAYERS THAT COME AFTERWARD

ACCURACY SCORE	PRECISION SCORE	RECALL SCORES
1.00 %	0.01	1.00
EPOCHS=10		



HYPER PARAMETER: THIS ARE PARAMETERS THAT CAN BE ADJUSTED BEFORE TRAINING A MACHINE LEARNING MODEL. THEY DETERMINE HOW WELL A MODEL LEARNS PATTERNS IN A DATA

HYPER PARAMETER TUNING: IT IS THE PROCESS OF FINDING THE BEST PARAMETERS THAT CAN HELP THE MODEL PERFORM AT ITS BEST

TECHNIQUES USED TO IMPLEMENT HYPERPARAMETER FINE -TUNING

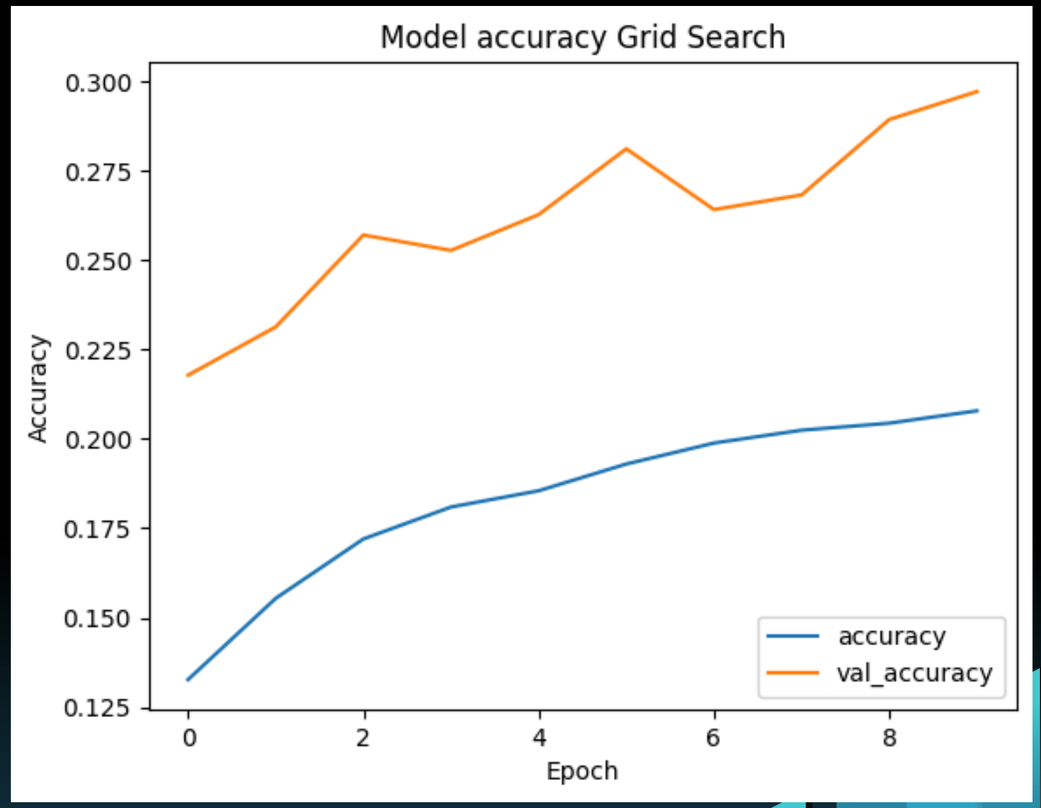
- GRID SEARCH: IT IS USED TO FIND THE BEST PARAMETERS FOR A MODEL. IT ACHIEVES THIS BY TRYING EVERY POSSIBLE COMBINATION AND THEN MEASURES THE MODEL PERFORMANCE ON EACH COMBINATION IN ORDER TO GET THE BEST FOR THE MODEL
- BAYESIAN OPTIMIZATION: THIS IS A TECHNIQUE THAT PREDICTS WHICH COMBINATIONS ARE BEST FOR THE MODEL AND TESTS ONLY THOSE

Grid search

DROPOUT RATE (0.4)	**RANDOMLY TURNS OFF SOME NEURONS DURING TRAINING**	**TURNS OFF 40% OF THE NEURONS DURING TRAINING**
FILTER(32)	IT EXTRACT THE FEATURES FROM THE IMAGES SUCH AS TEXTURE, PATTERN AND EDGES	IT HAS 32 FILTERS WITH A HEIGHT OF 2 AND WIDTH OF 2
KERNEL SIZE(3)	THE HEIGHT AND WIDTH OF THE FILTER	IT IS 3X3 WHICH REPRESENTS THE NUMBER OF PIXELS THE FILTER COVERS

GRID SEARCH RESULT

ACCURACY SCORE	PRECISION SCORE	RECALL SCORES
29.7 %	28.78 %	29.7 %
EPOCHS=10		

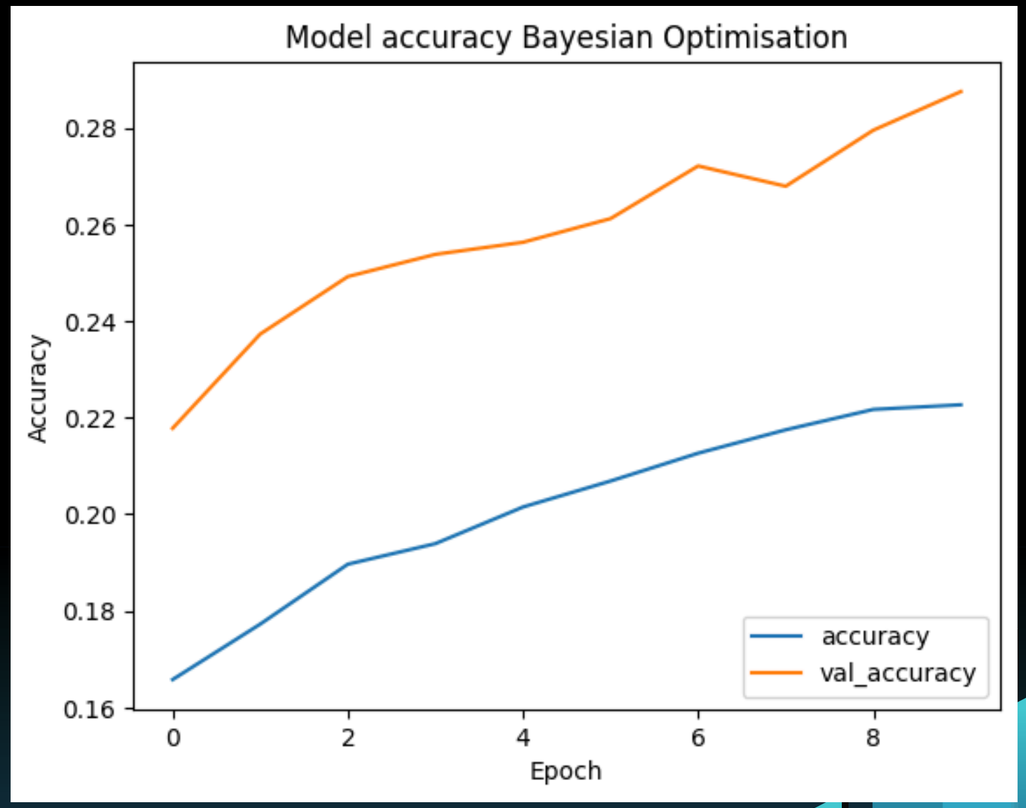


BAYESIAN SEARCH RESULT

DROPOUT RATE (0.295)	**RANDOMLY TURNS OFF SOME NEURONS DURING TRAINING**	**TURNS OFF 29.5% OF THE NEURONS DURING TRAINING**
FILTER(128)	IT EXTRACT THE FEATURES FROM THE IMAGES SUCH AS TEXTURE, PATTERN AND EDGES	IT HAS 32 FILTERS WITH A HEIGHT OF 2 AND WIDTH OF 2
LEARNING RATE (1.66E-05	HOW FAST THE MODEL LEARNS	

BAYESIAN OPTIMIZATION RESULT

ACCURACY SCORE	PRECISION SCORE	RECALL SCORES
0.98	0.1338	0.98
EPOCHS=10		



Hypothesis testing

It is a statistical method used to make informed decision using data. It helps us to determine if there is enough evidence to support a claim about a system.

- NULL HYPOTHESIS:IT REPRESENTS THE DEFAULT OR INITIAL ASSUMPTION, IT ASSUMES NO CHANGE OR EFFECT
- ALTERNATE HYPOTHESIS :IT WHAT HAS TO BE PROVED. IT ASSUMES A CHANGE, EFFECT OR DIFFERENCE.

Hypothesis testing

null hypothesis: Data augmentation has no effect on the model

alternative hypothesis: data augmentation has effect on the model

data collection: the data collected is the metrics of the models performance on augmented images and non augmented images.

T- TEST

IT IS A STATISTICAL METHOD USED TO COMPARE THE MEAN BETWEEN TWO GROUPS AND CHECK IF THE DIFFERENCE IS SIGNIFICANT.

TWO SAMPLE T- TEST: COMPARES THE MEANS OF TWO INDEPENDENT GROUPS.

DATA COLLECTION: THE DATA COLLECTED IS THE METRICS OF THE MODELS PERFORMANCE ON AUGMENTED IMAGES AND NON AUGMENTED IMAGES.

To make decisions we consider key values such as:

- significant value :this is the threshold for deciding if we rject the null hypothesis. It is usually set 0.05. 5% chance of making the wrong decision
- p- value; this tells us how likely the result occured by chance

If the p-value<0.05 the result is unlikely to have happened by chance so we reject the null hypothesis

if the p-value>0.05: fail to reject the null hypothesis

DATA COLLECTION:

THE DATA COLLECTED IS THE METRICS OF THE MODELS PERFORMANCE ON AUGMENTED IMAGES AND NON AUGMENTED IMAGES.

THE DATA COLLECTED WAS THE METRICS OF THE LOGISTICS REGRESSION AND DEEP LEARNING PERFORMANCE.

THE METRICS ARE USED TO FORM GROUP ONE AND TWO.

Hypothesis 1: deep learning models outperform traditional models

	CNN	Logistic regression
ACCURRACY	37.5%	7.74%
PRECISION SCORE	38.38%	7.72%
RECALL SCORE	37.5%	7.74%
P-VALUE=0.0001 SIGNIFICANT VALUE=0.05		

HO: DEEP LEARNING MODELS DO NOT OUTPERFORM LOGISTIC REGRESSION

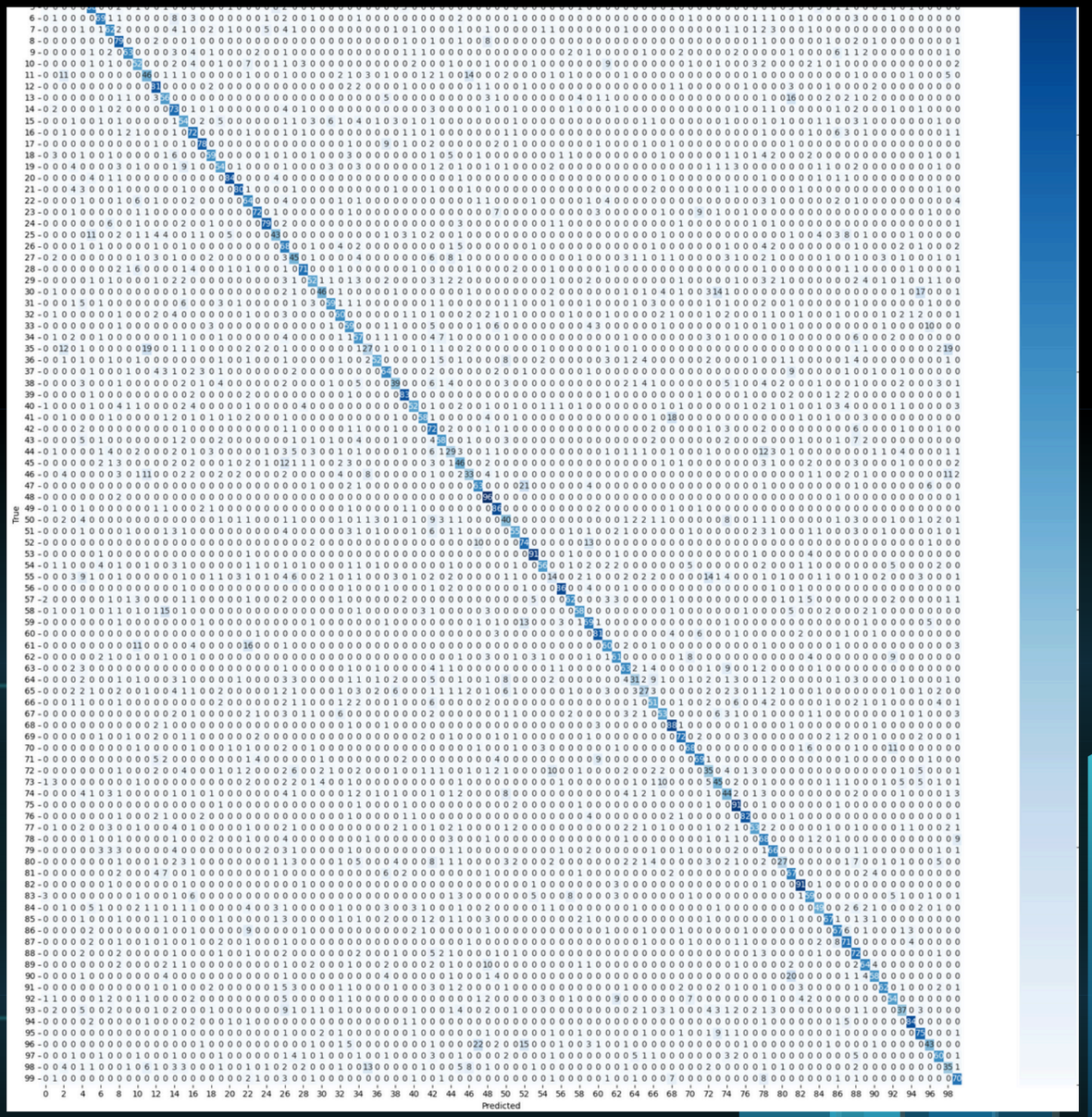
H1: DEEP LEARNING MODELS OUTPERFORM LOGISTIC REGRESSION MODELS RAPH TEXT

***DATA COLLECTION**: THE DATA COLLECTED IS THE METRICS OF THE MODELS PERFORMANCE*

THE P-VALUE IS LESS THAN THE SIGNIFICANT VALUE THEREFORE WE REJECT THE NULL HYPOTHESIS.

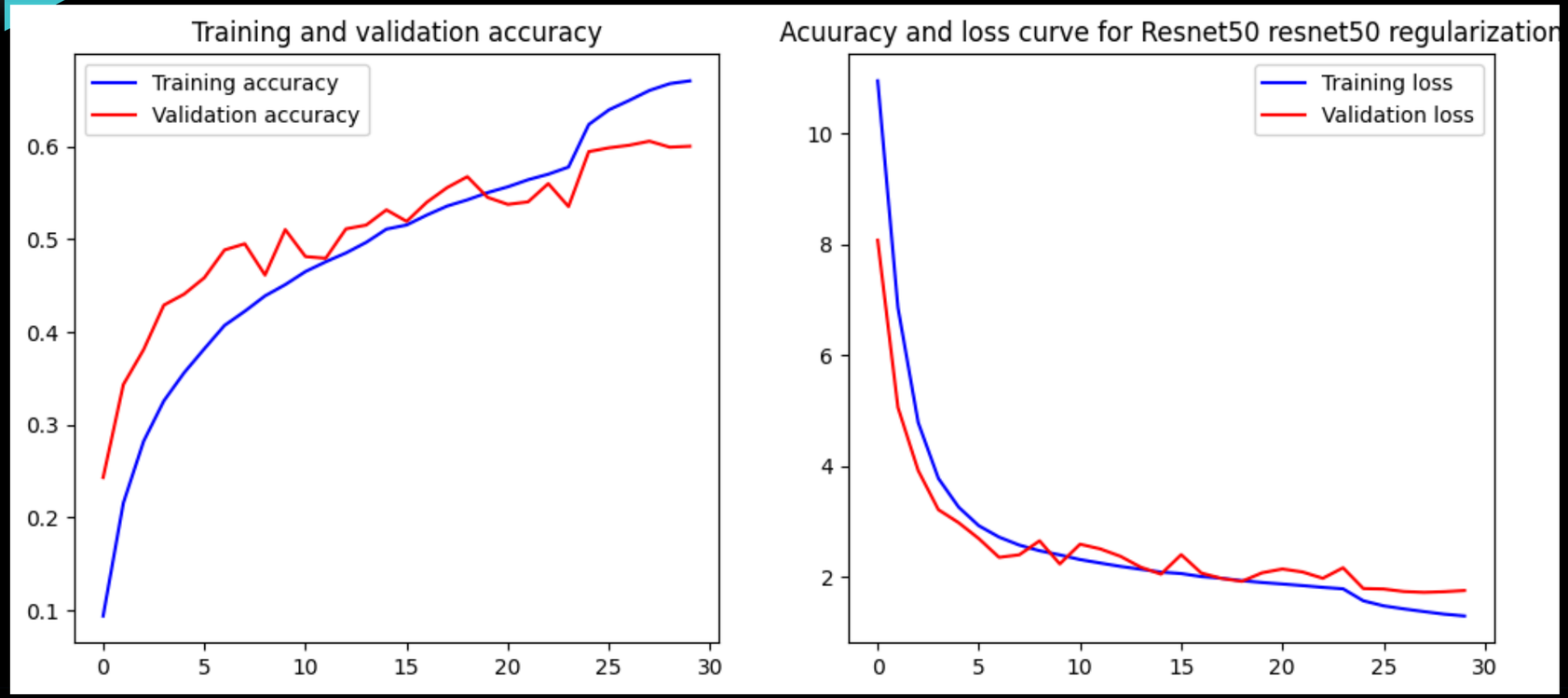
FINAL MODEL: RESNET50

ACCURACY	PRECISION SCORE	RECALL SCORE
60.58%	62.14%	60.58%
EPOCHS=30		
PATIENT=5		

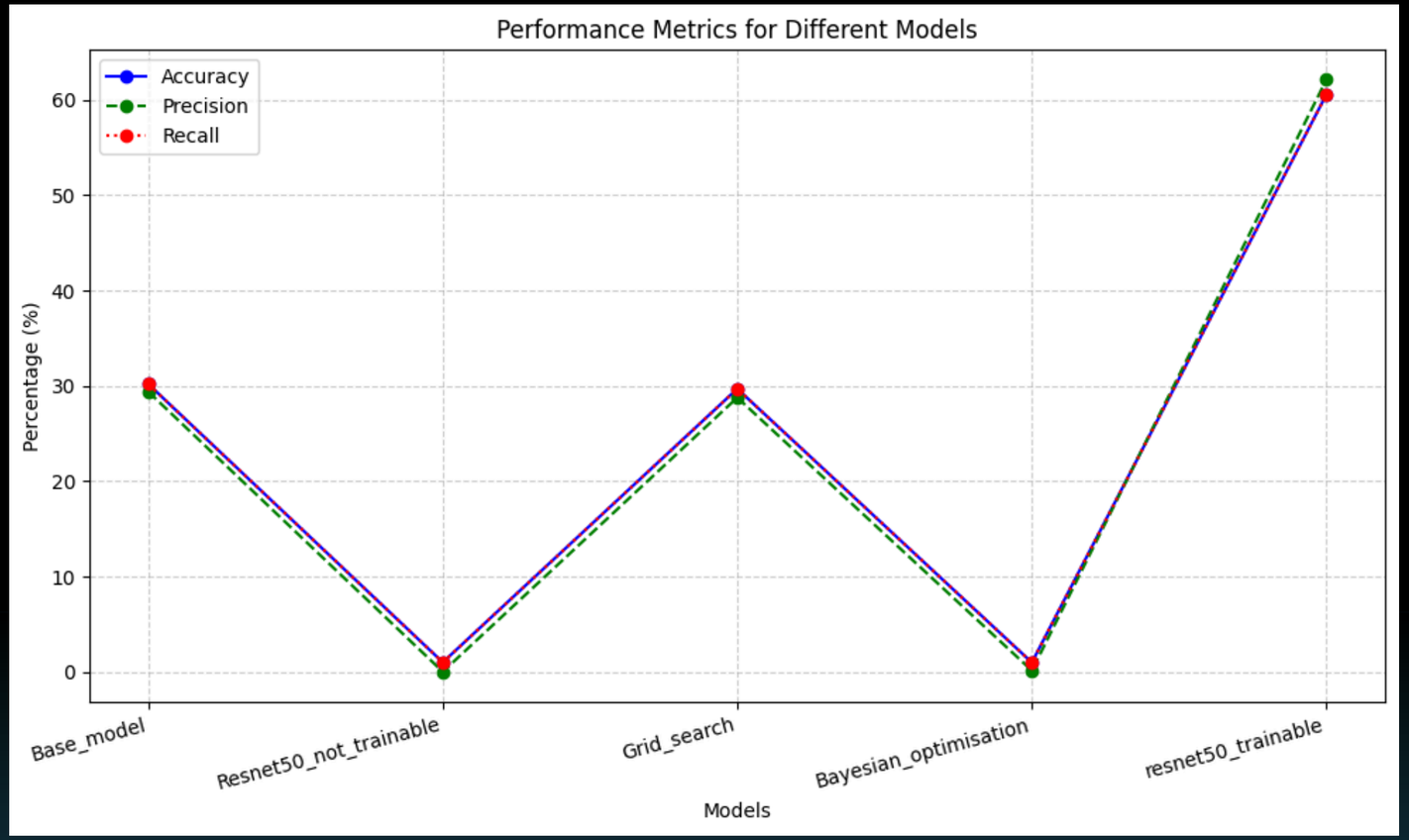


EARLY STOPPING ONCE THE MODEL PERFORMANCE ON THE VALIDATION SET STOPS IMPROVING THE TRAINING IS HALTED.

FINAL MODEL: RESNET50



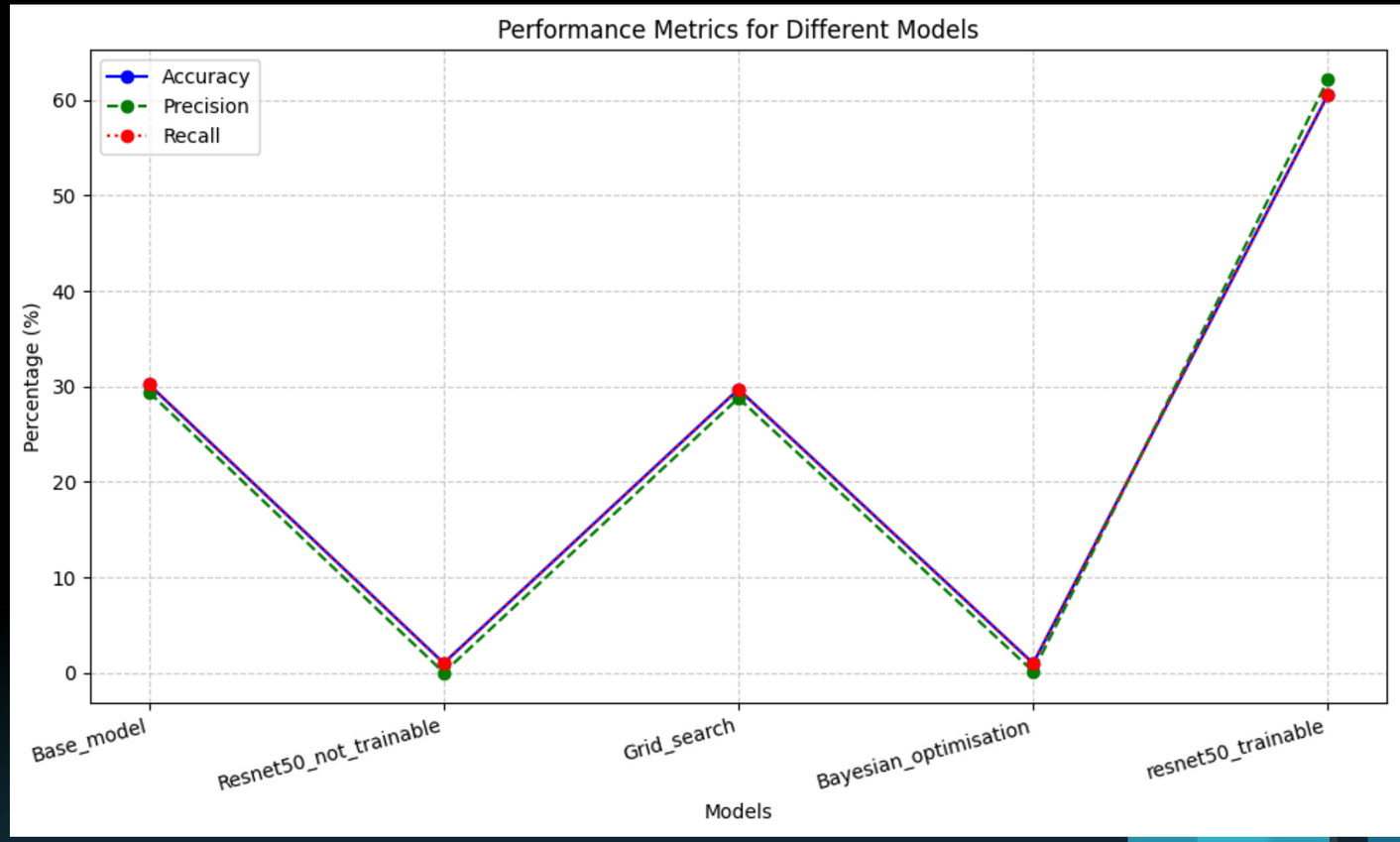
MODEL PERFORMANCE COMPARISON



Your paragraph text

Recommendation

- A MORE ADVANCED PRE-TRAINED MODEL CAN BE USED (VGG19)
- FURTHER EXPERIMENTATION CAN BE MADE WITH EARLY STOPPING PATIENT VALUE
- REDUCED LEARNING RATE CAN BE IMPLEMENTED
- MODELS CAN BE TRAINED FOR MORE EPOCHS
- REGULARIZATION TECHNIQUES CAN BE IMPLEMENTED



CONCLUSION

- DEEP LEARNING PERFORMS BETTER ON THE CIFAR-100 DATASET THAN REGULAR MACHINE LEARNING MODELS
- TRANSFER LEARNING IMPROVES MODEL PERFORMANCE SIGNIFICANTLY WHEN THE LAYERS ARE TRAINABLE

Thank You for listening