Active Learning

Project 2

Agenda

- Introduction
- Datasets Description
- Pretrained Models
- Classifiers
- Query Strategies
- Visualization

Introduction

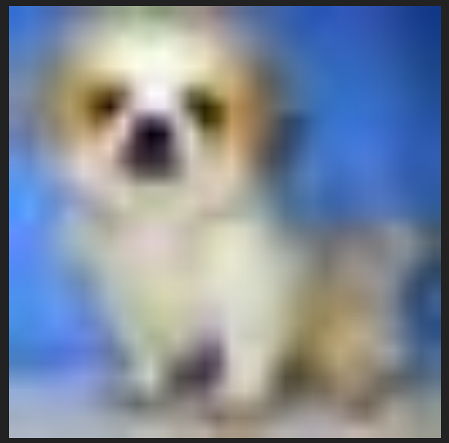
In this project, we aim to explore the effectiveness of active deep learning techniques in comparison to traditional approaches on popular benchmark datasets such as CIFAR-10 and Flowers Recognition.

We will employ a diverse set of classifiers and query strategies to conduct a comprehensive evaluation of the performance before and after the integration of active learning.

CIFAR-10

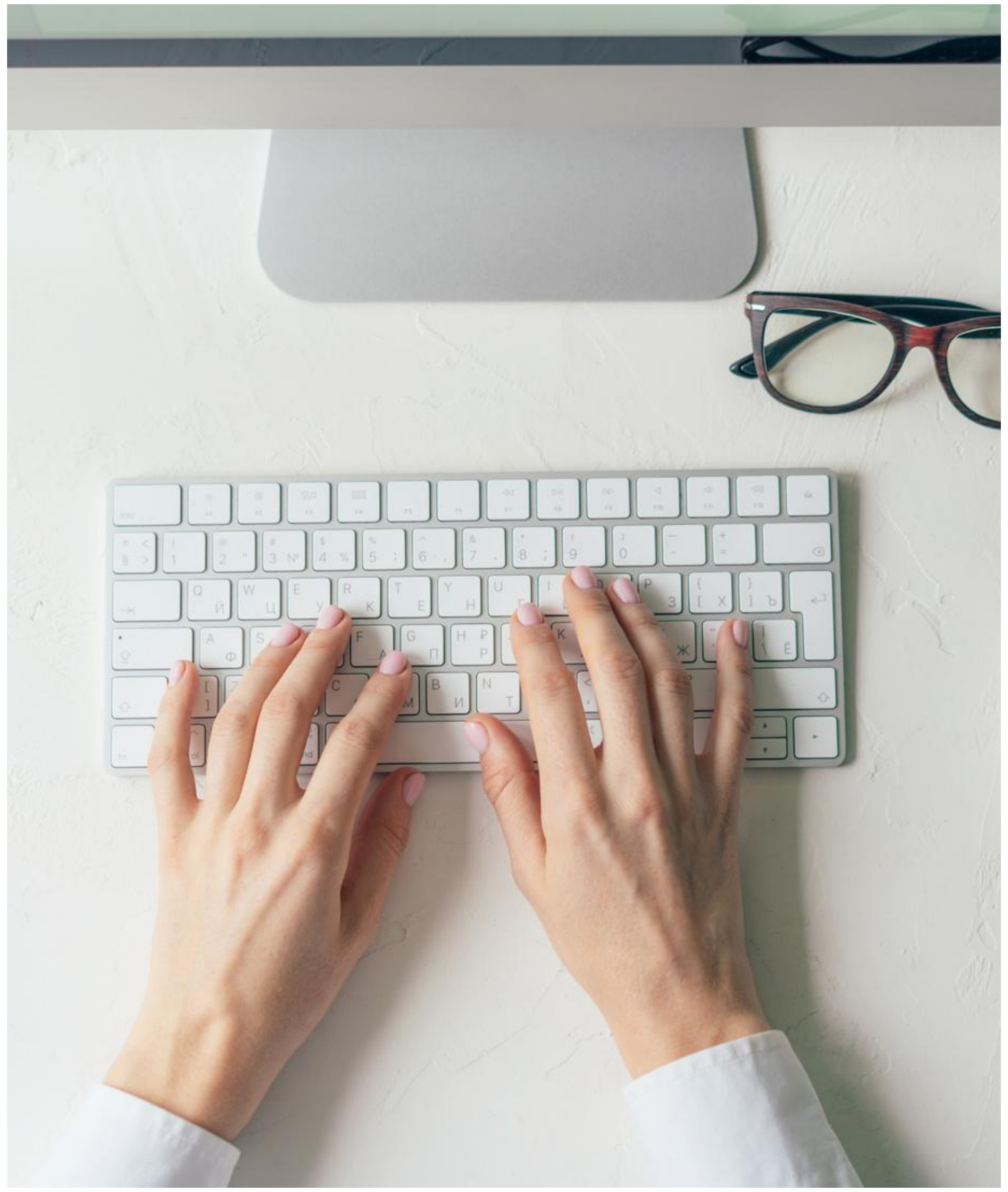
Dataset Description

The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.
the classes are: airplane, truck, ship, horse, frog, dog, deer, cat, bird, automobile.

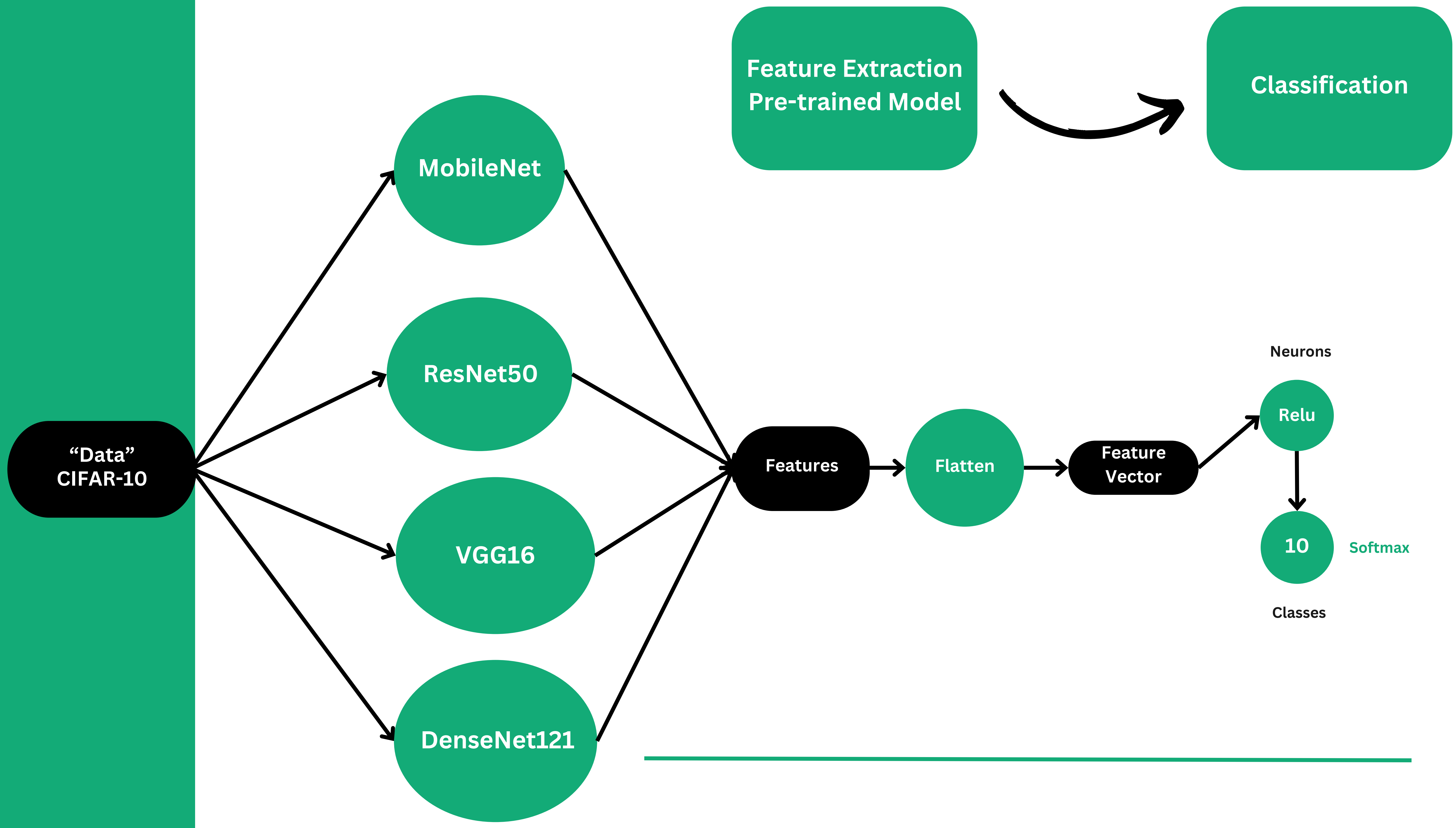


Experiments

- Pretrained Models (Feature Extraction)
- Custom Model
- Classifiers
- Query Strategies
- Results & Visualization



Pre-trained Models



Pre-trained Models

1

MobileNet Model

Testing Accuracy : 0.1902- Loss: 2.2073

2

DenseNet121 Model

Testing Accuracy : 0.3754 - Loss: 1.91

3

VGG16 Model

Testing Accuracy : 0.5344 -Loss: 2.1391

5

EfficientNetB0 Model

Testing Accuracy : 0.2677 - loss: 1.6011

Custom Model 

Structure:

- **Input Layer:** Accepts images with dimensions 32 × 32 pixels and three-color channels (RGB).
- **Convolutional Layer 1:** 32 filters, kernel size (3, 3), ReLU activation.
- **MaxPooling Layer 1:** Pool size (2, 2).
- **Convolutional Layer 2:** 64 filters, kernel size (3, 3), ReLU activation.
- **MaxPooling Layer 2:** Pool size (2, 2).
- **Convolutional Layer 3:** 64 filters, kernel size (3, 3), ReLU activation.
- **Flatten Layer:** Converts 3D feature maps to a 1D vector.
- **Dense Layer 1:** 64 units, ReLU activation.
- **Dense Layer 2:** 10 units (for CIFAR-10), softmax activation.

Accuracy: 53.44%

Loss: 2.1391

F1 Score = 0.534

Classifiers

Neural Network Architectures:

Using a pretrained model for extracting features and following it with dense layers for classification, we used a model that consists of 3 dense layers, (100) neurons -> (10) neurons with softmax.

We attempted to evaluate the performance by comparing with different activation functions as:

- **ReLU**
- **Tanh**

Note: The “RELU” activation function gives better performance.

Query Strategies

Random Sampling

Random sampling is the simplest and most straightforward query strategy in active learning. It involves randomly selecting instances from the unlabeled dataset for labeling.

Entropy

It measures the entropy of the predicted class probabilities for each instance. Higher entropy indicates higher uncertainty, suggesting that the instance is more informative for labeling.

Margin Sampling

It focuses on instances that lie close to the decision boundary, where the classifier is relatively uncertain. The decision boundary is typically defined by the margin between the probabilities assigned to the top two predicted classes.

Uncertainty Sampling

Uncertainty sampling is a query strategy in active learning where the model selects instances with the highest uncertainty, often measured by low confidence or high entropy, to improve its performance by focusing on challenging examples.

Consensus Entropy Sampling

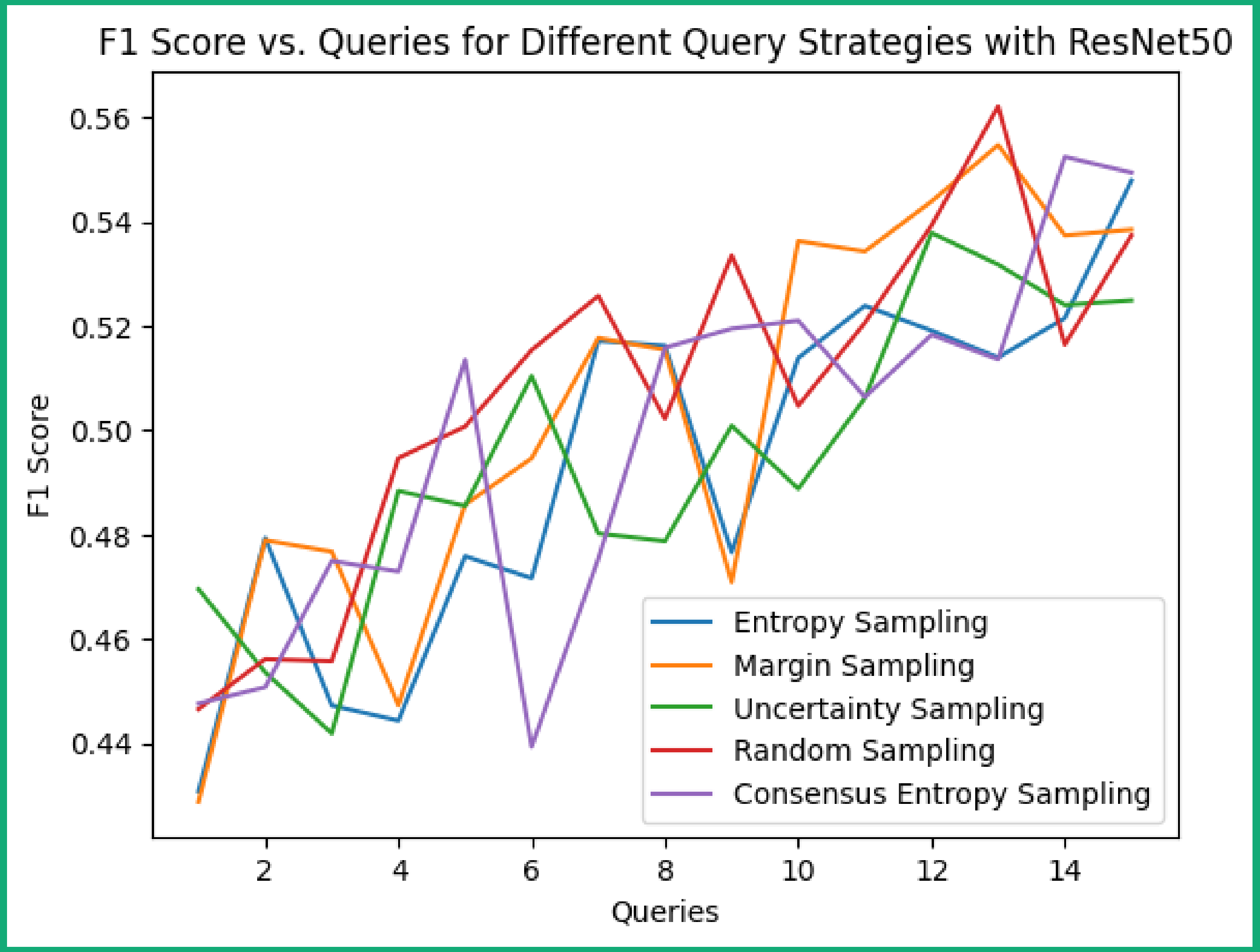
It combines predictions from multiple models and selects instances with high uncertainty, often measured by entropy or disagreement among model predictions. It leverages diverse model perspectives to prioritize labeling challenging data points, enhancing the learning process in active learning scenarios.

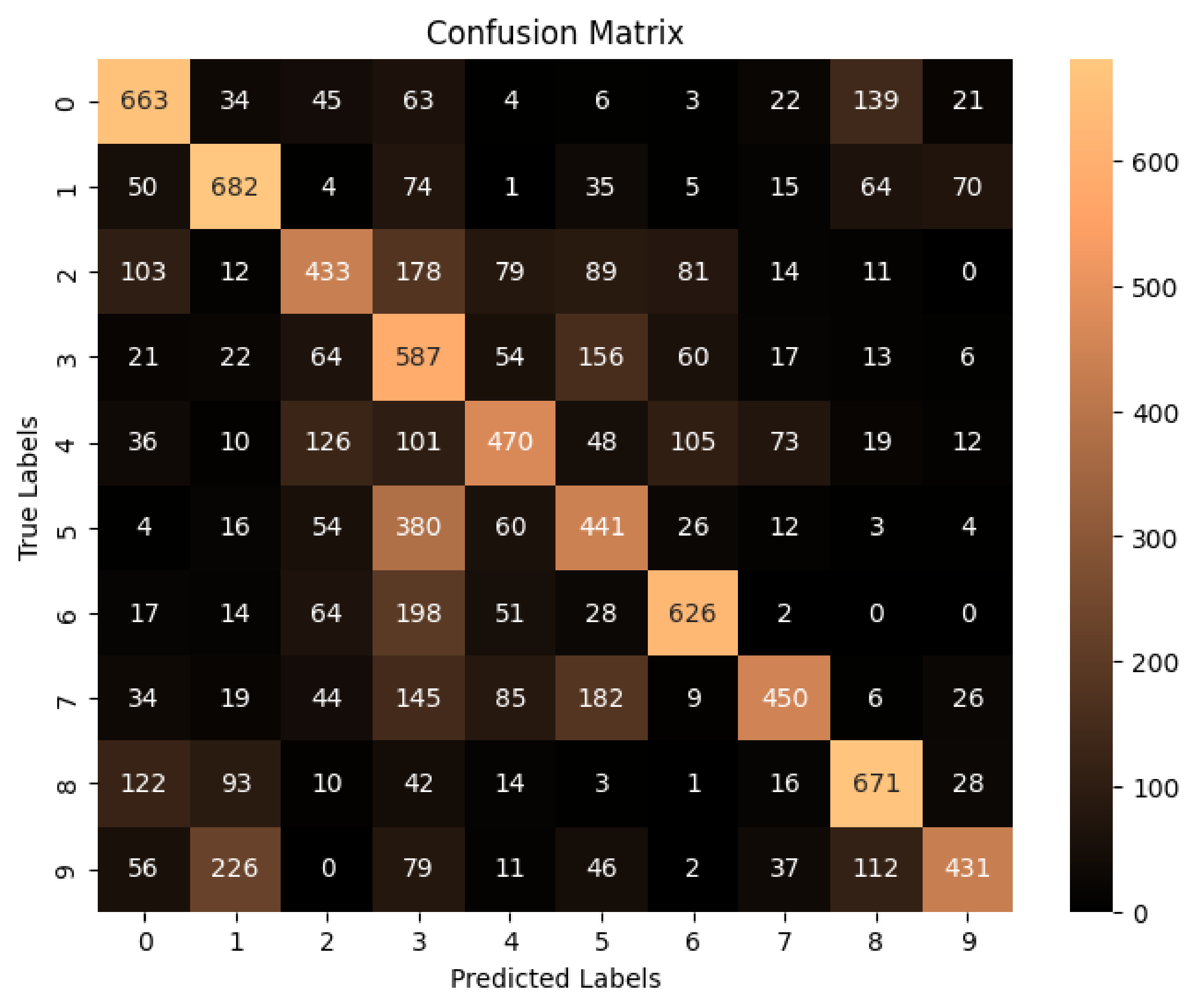
Results

NN × ResNet50 Results

Query Strategy	F1 score
Random Sampling	0.5373
Entropy Sampling	0.5478
Margin Sampling	0.5384
Uncertainty Sampling	0.5248
Consensus Entropy Sampling	0.5494

NN × ResNet50 Visualization





Conclusion

CIFAR-10 Dataset

We have gone through rigorous experiments to pick up the most effective architecture, and our winner was the ResNet50 model with: # epochs / queries = 15, number of instances = 2500 and a combination with Marginal sampling, which produced a test accuracy and an f1 score of 55.07% and 0.5136 respectively.

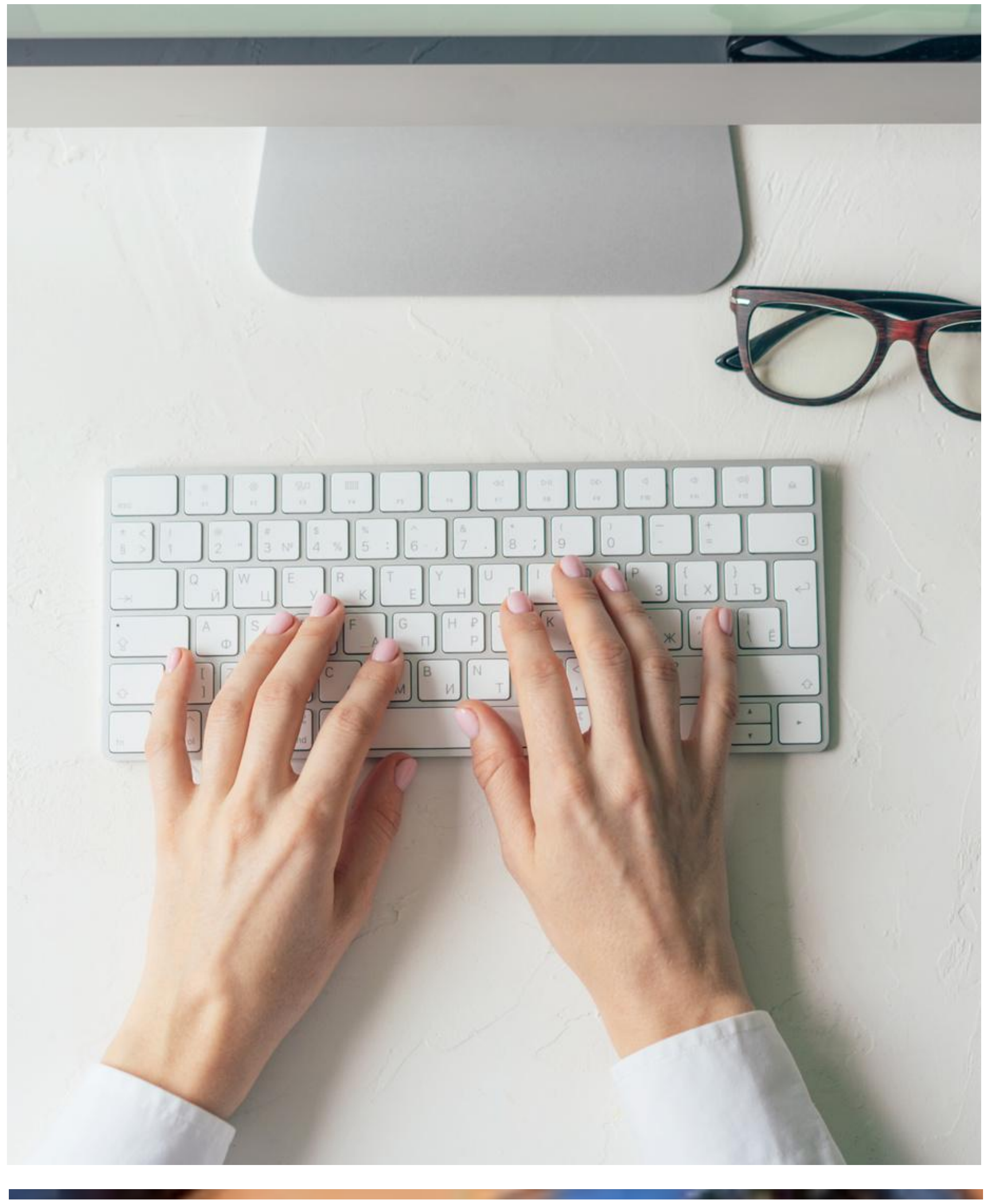
Flowers Recognition

Dataset Description

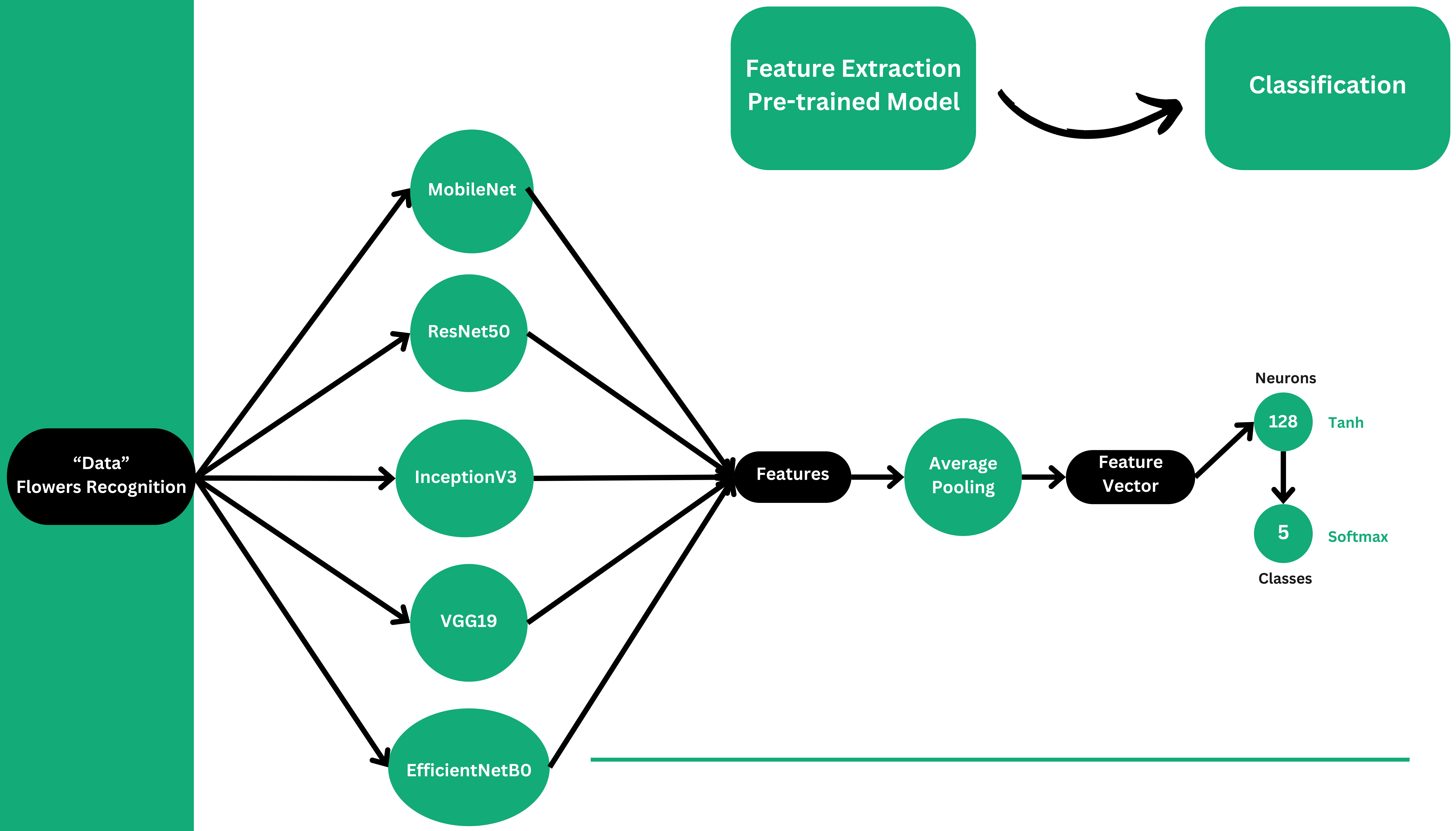
This dataset contains 4242 images of flowers.
The images are divided into 5 classes: Chamomile, Tulip, Rose, Sunflower and Dandelion.

Experiments

- Pretrained Models (Feature Extraction)
- Classifiers
- Query Strategies
- Visualization (Using PCA)
- Different initial splits of the dataset



Pre-trained Models



Pre-trained Models

1

MobileNet Model

Testing Accuracy : 0.9184 - loss: 0.3062

2

InceptionV3 Model

Testing Accuracy : 0.8741 - loss: 0.3864

3

VGG19 Model

Testing Accuracy : 0.7535 - loss: 0.7732

4

ResNet50 Model

Testing Accuracy : 0.5009 - loss: 1.2682

5

EfficientNetB0 Model

Testing Accuracy : 0.2677 - loss: 1.6011

Classifiers

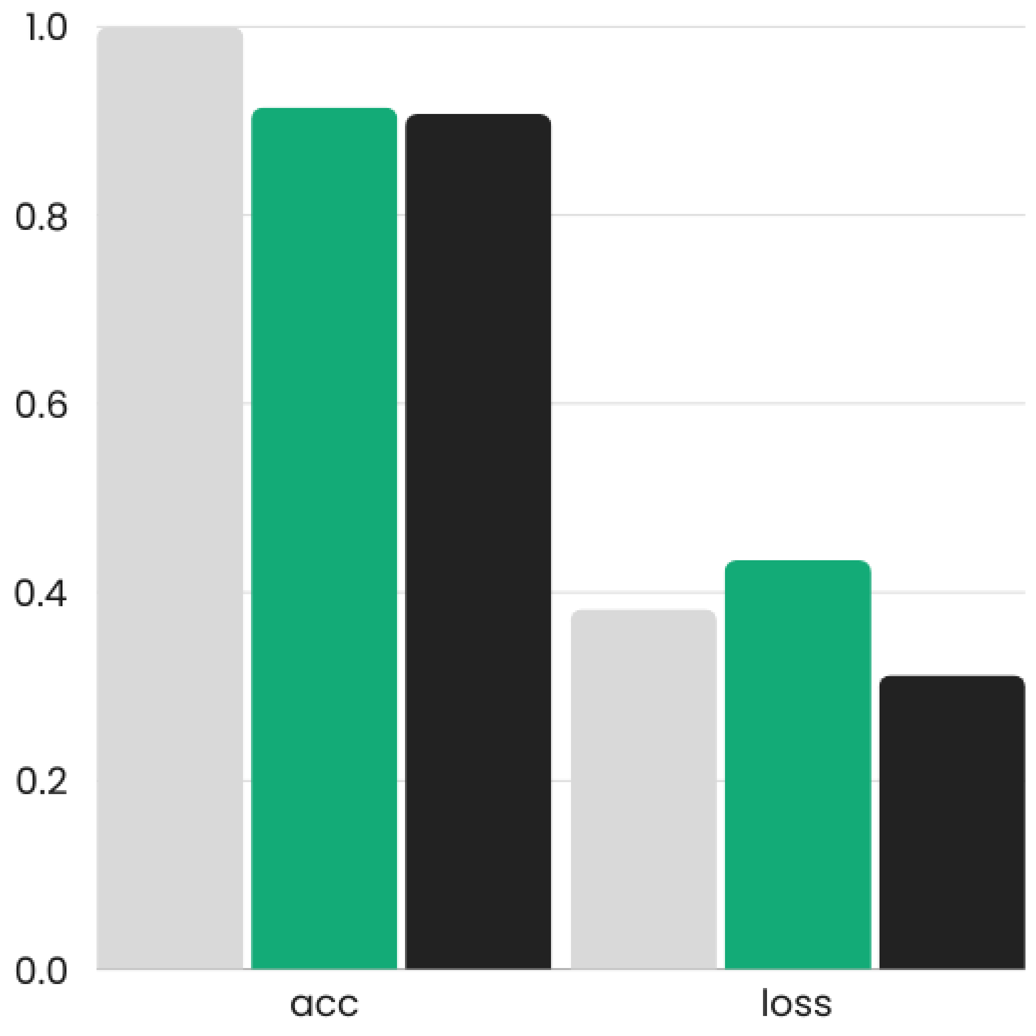
Neural Network Architectures:
Using a pretrained model for extracting features and following it with dense layers for classification.

1- First Architecture:
Consists of 2 dense layers, (128) neurons -> (5) neurons with softmax.

2- Second Architecture:
Consists of 4 dense layers, (512) neurons -> (128) neurons -> (64) neurons -> (5) neurons with softmax and we use dropout layers to avoid overfitting.

- We attempted to evaluate the performance by comparing different activation functions:

- **ReLU:** acc: 0.9167 - loss: 0.3821
- **Tanh:** acc: 0.9143 - loss: 0.4340
- **Sigmoid:** acc: 0.9073 - loss: 0.3119



Classifiers

- **SVM:** Aims to find an optimal hyperplane that separates the instances of different classes with the maximum margin. It is effective in handling high-dimensional data and can handle both linearly separable and non-linearly separable datasets using kernel functions.

- **Random Forest:** Constructs multiple decision trees and combines their predictions to make final decisions. Each decision tree is trained on a random subset of the training data and features. The predictions of individual trees are then aggregated to generate the final prediction.

Query Strategies

Random Sampling

Random sampling is the simplest and most straightforward query strategy in active learning. It involves randomly selecting instances from the unlabeled dataset for labeling.

Entropy

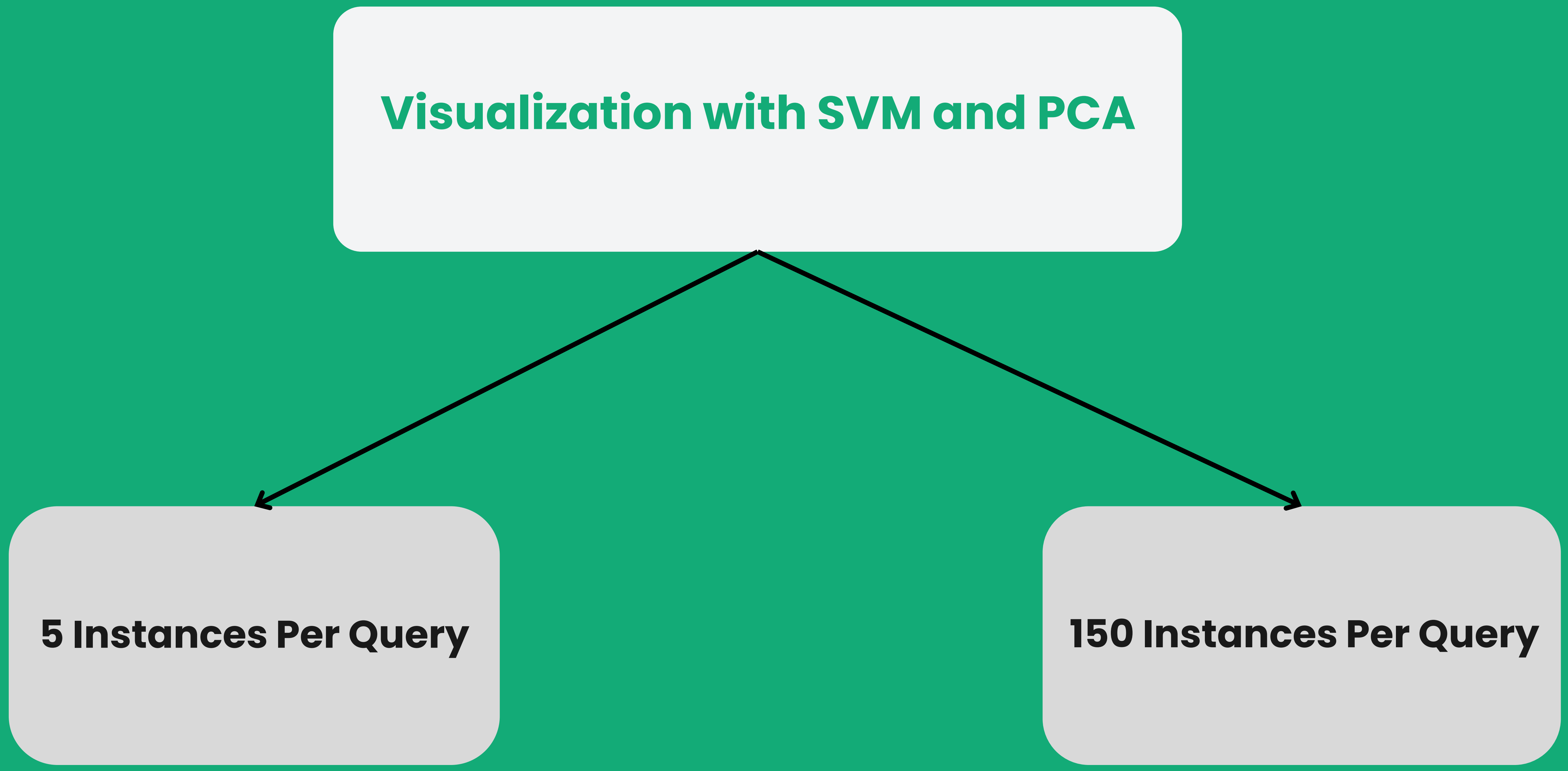
It measures the entropy of the predicted class probabilities for each instance. Higher entropy indicates higher uncertainty, suggesting that the instance is more informative for labeling.

Margin Sampling

It focuses on instances that lie close to the decision boundary, where the classifier is relatively uncertain. The decision boundary is typically defined by the margin between the probabilities assigned to the top two predicted classes.

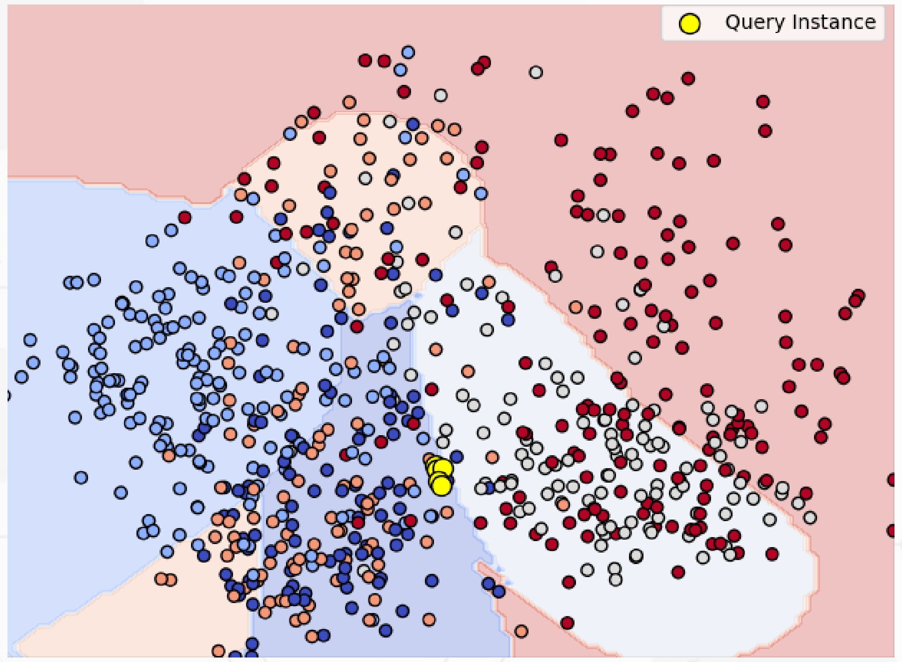
Query-By-Committee

Is a query strategy that involves maintaining a committee of diverse classifiers and selecting instances that elicit disagreement among them.

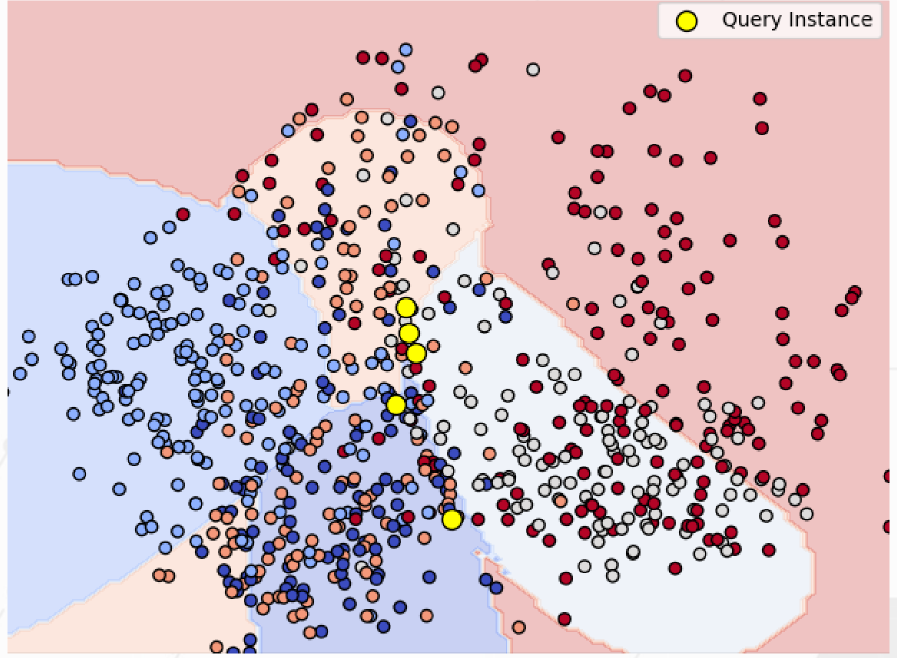

Visualization with SVM and PCA
5 Instances Per Query
150 Instances Per Query

Visualization with 5 Instances

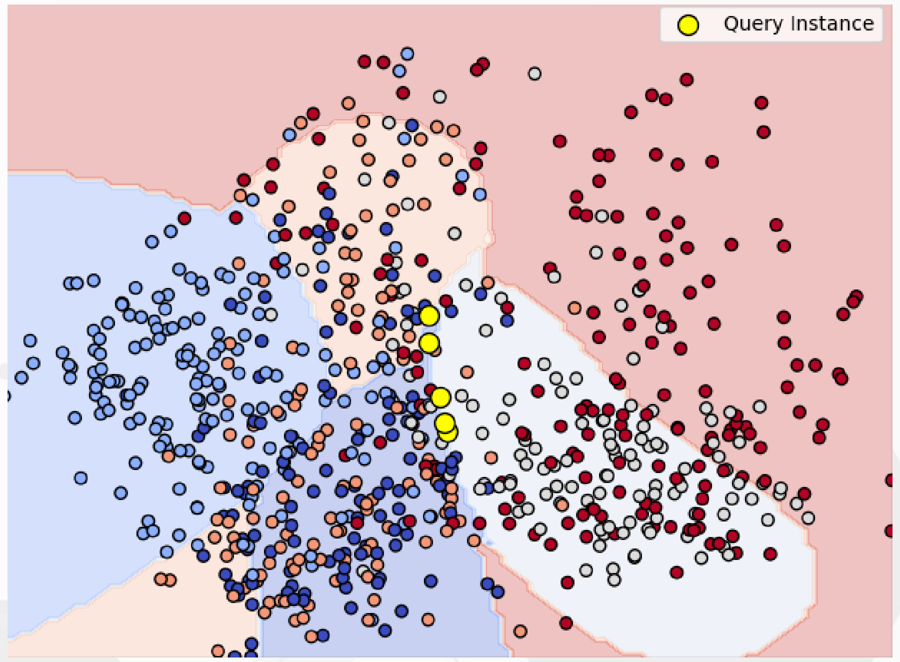
Iteration 1



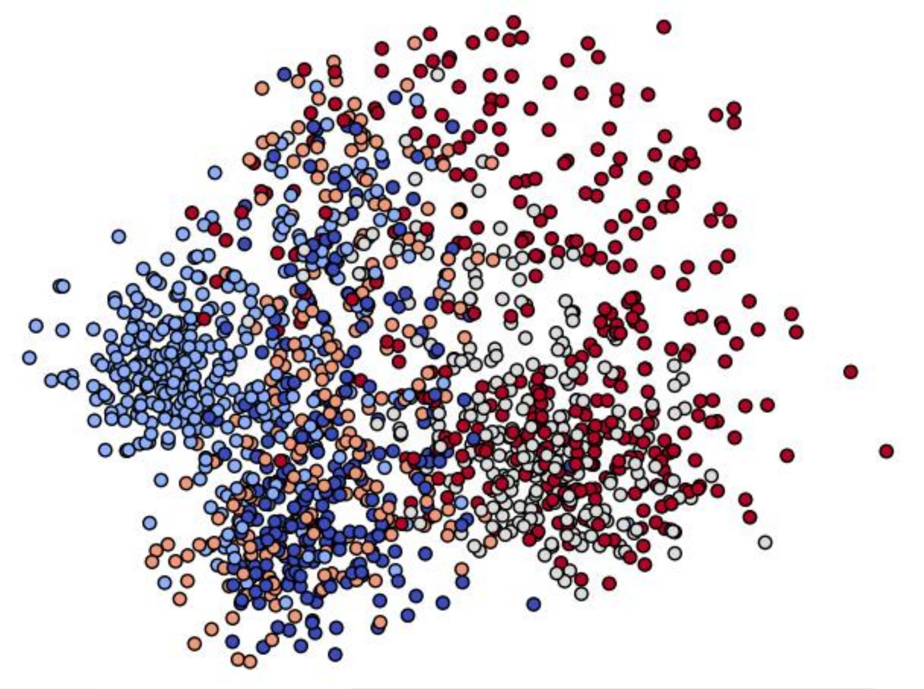
Iteration 7



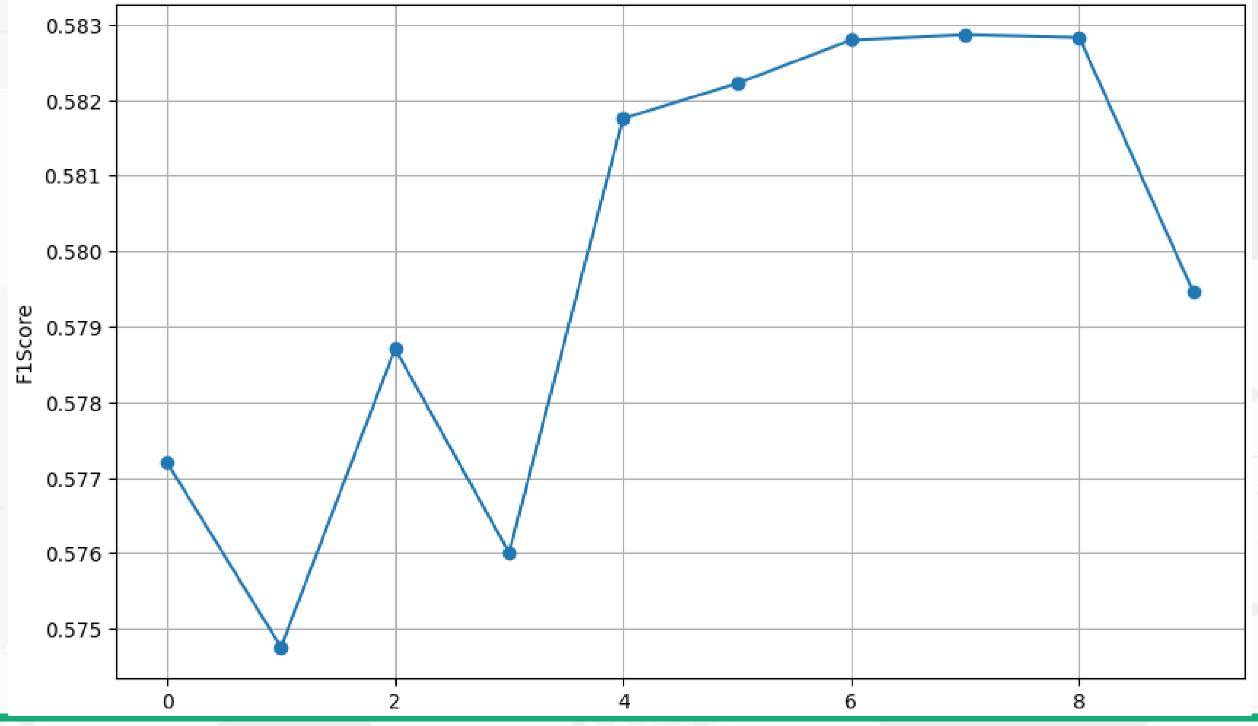
Iteration 10



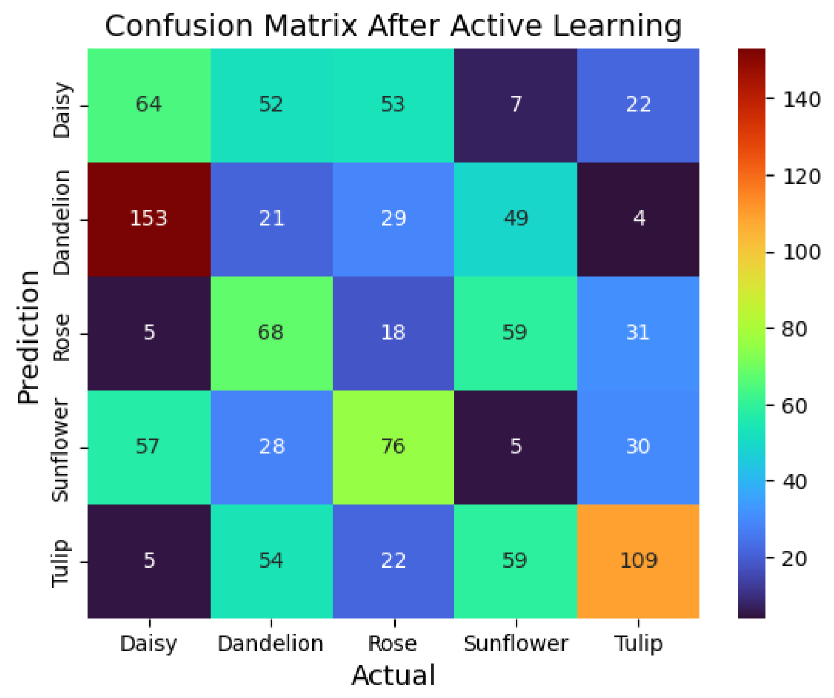
Initial data



F1 scores over iterations

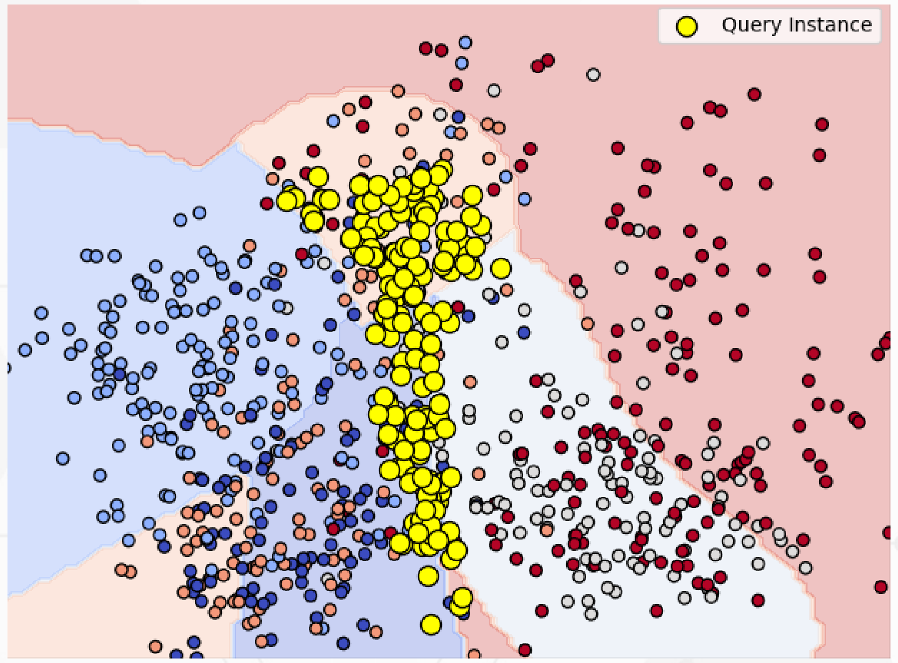


Confusion Matrix on test data

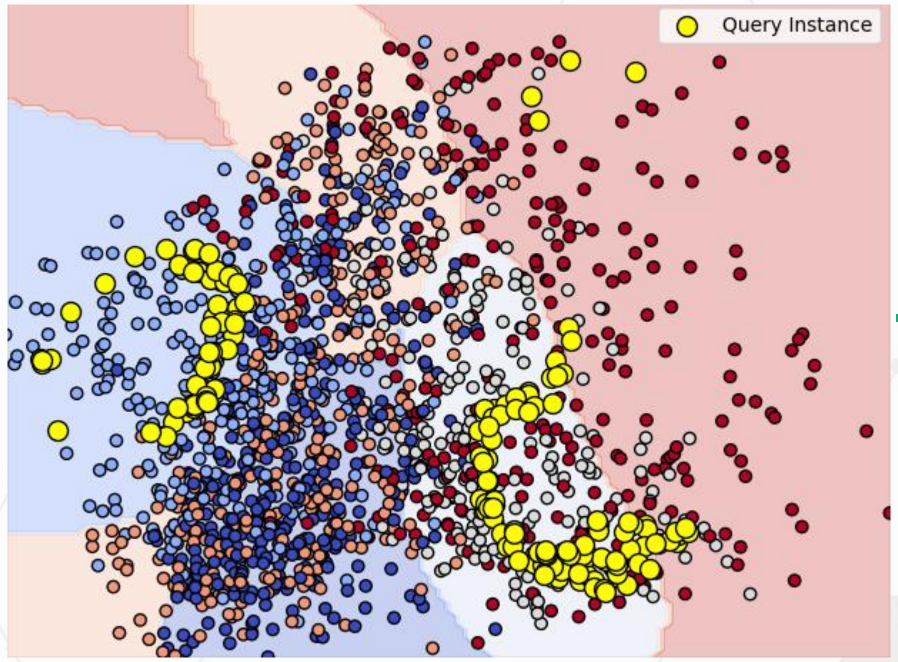


Visualization with 150 Instances

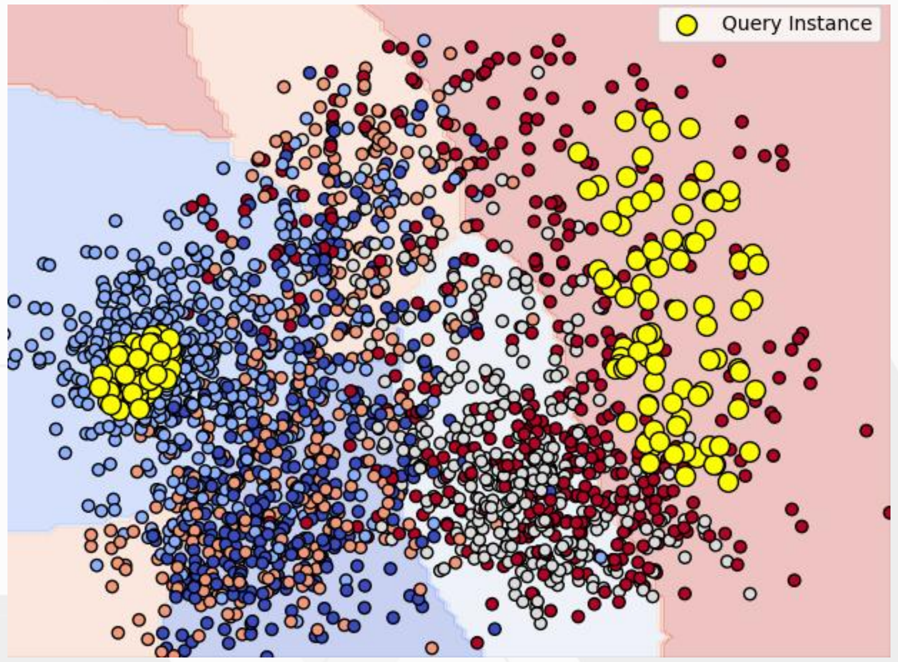
Iteration 1



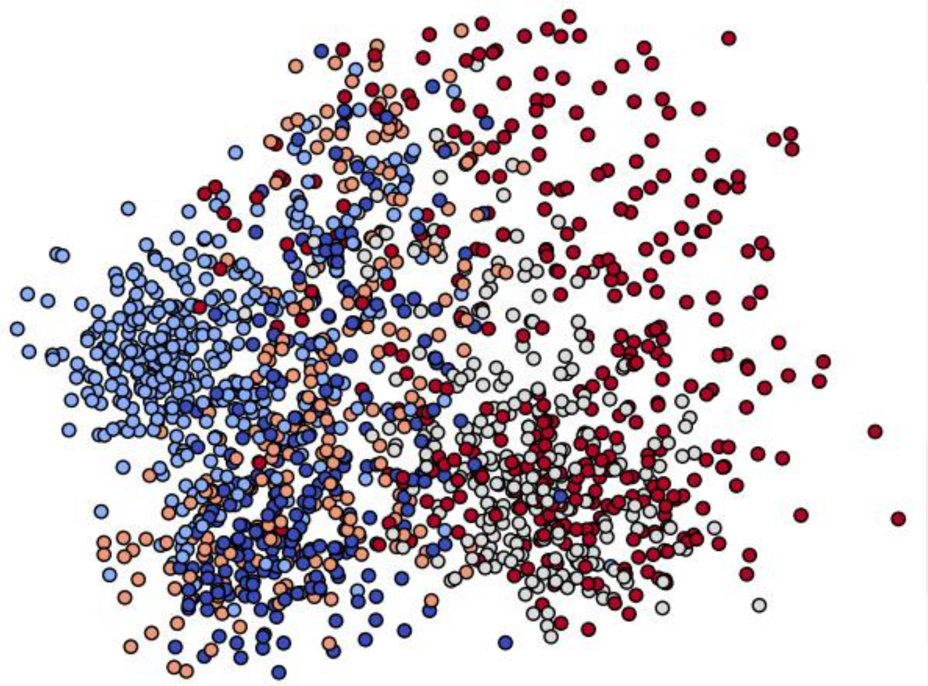
Iteration 7



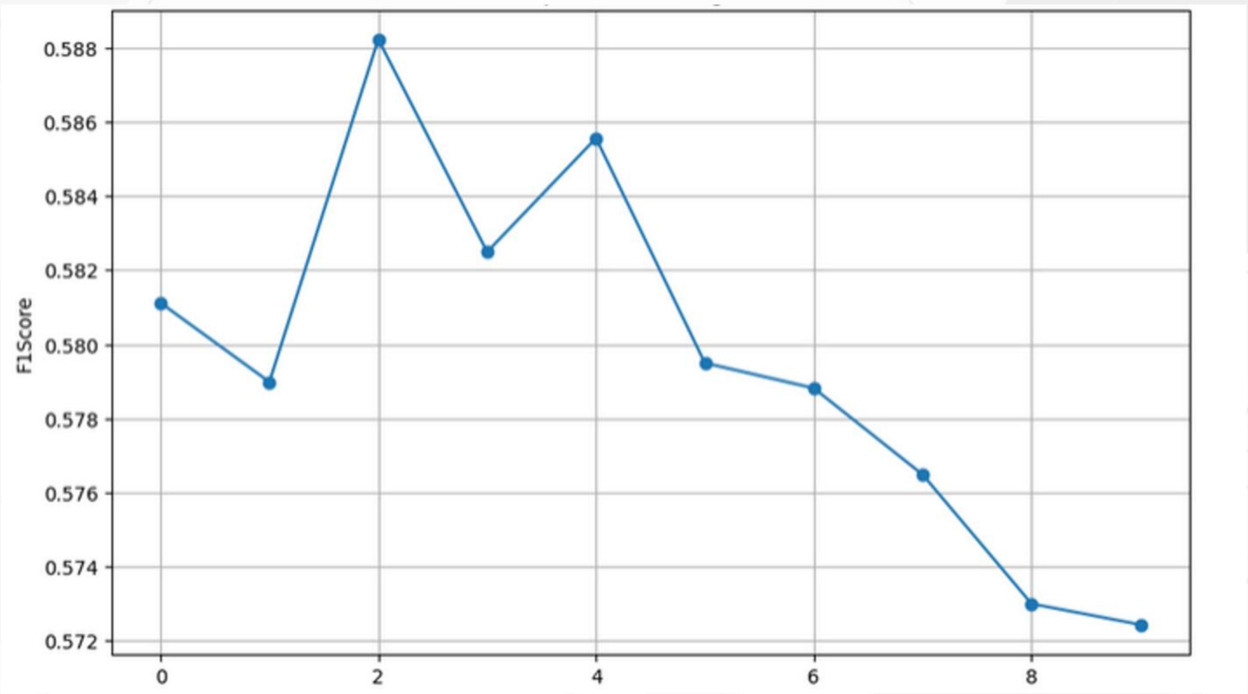
Iteration 10



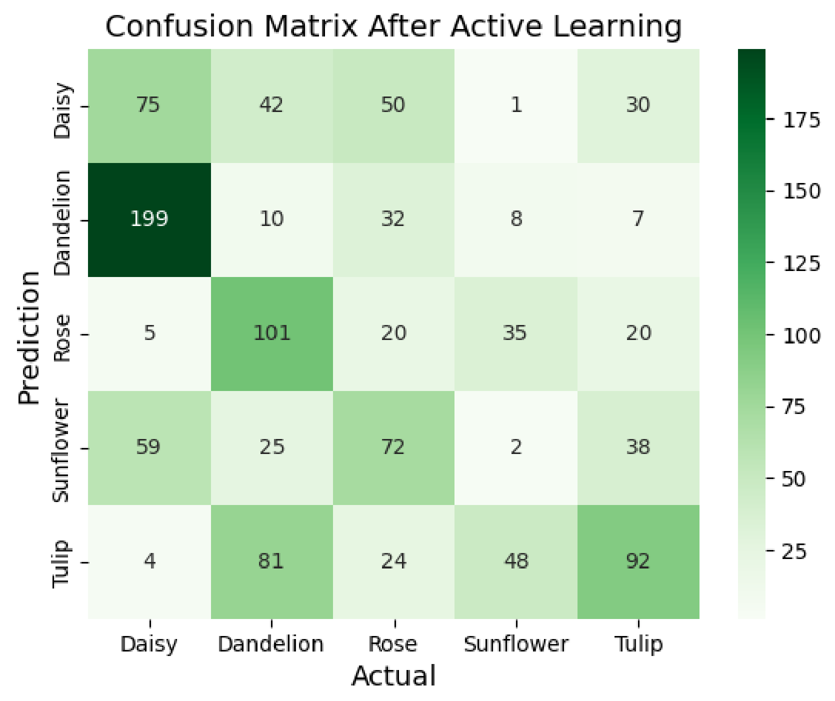
Initial data



F1 scores over iterations



Confusion Matrix on test data



Random Forest

173	12	6	3	4
6	223	4	15	8
1	5	142	1	32
6	9	10	156	15
2	10	20	4	213

SVM

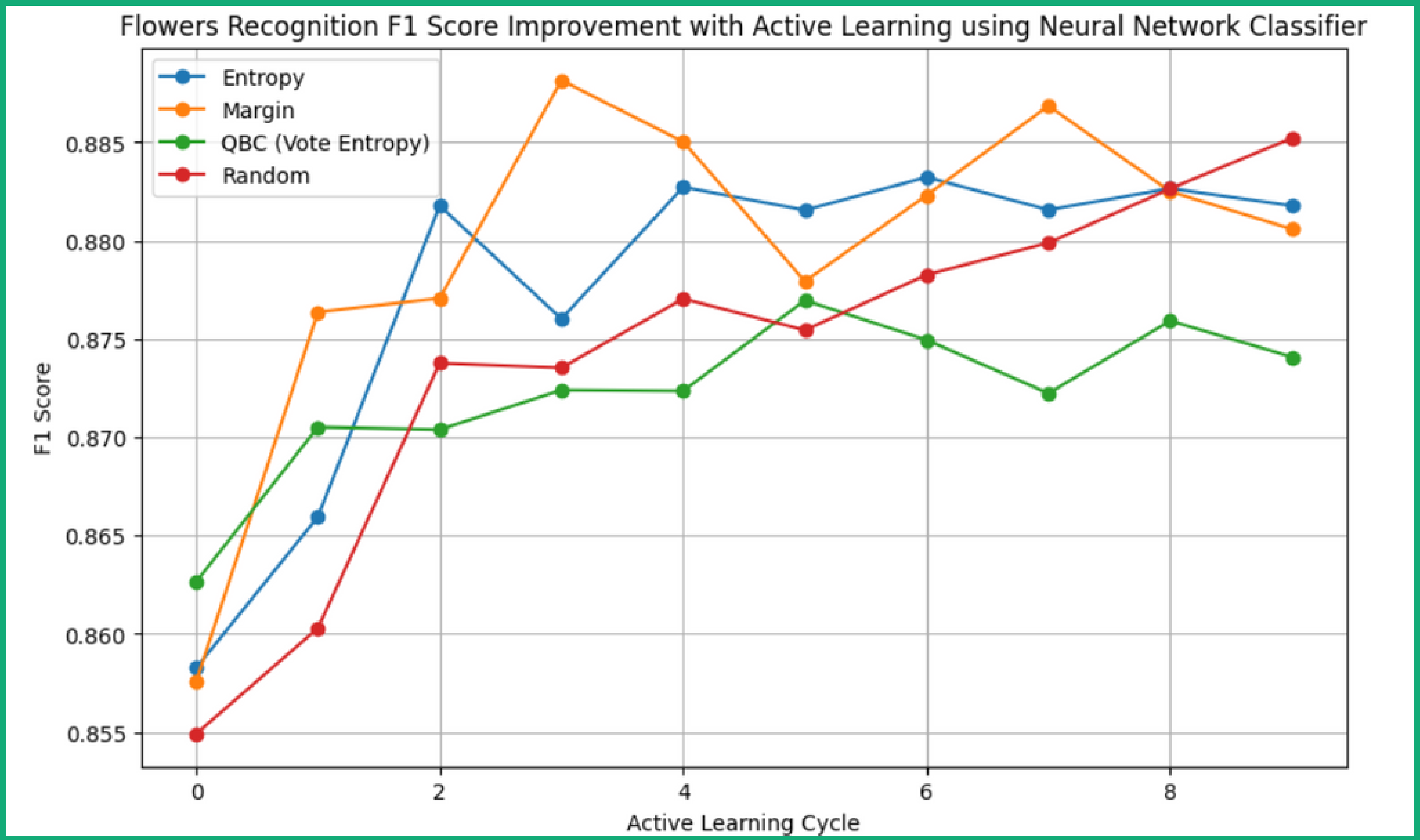
173	9	3	4	9
8	236	3	6	3
1	1	156	4	19
2	7	4	174	9
1	6	20	7	215

NN

173	8	6	5	6
10	228	4	11	3
2	1	155	4	19
2	2	9	175	8
2	3	24	9	211

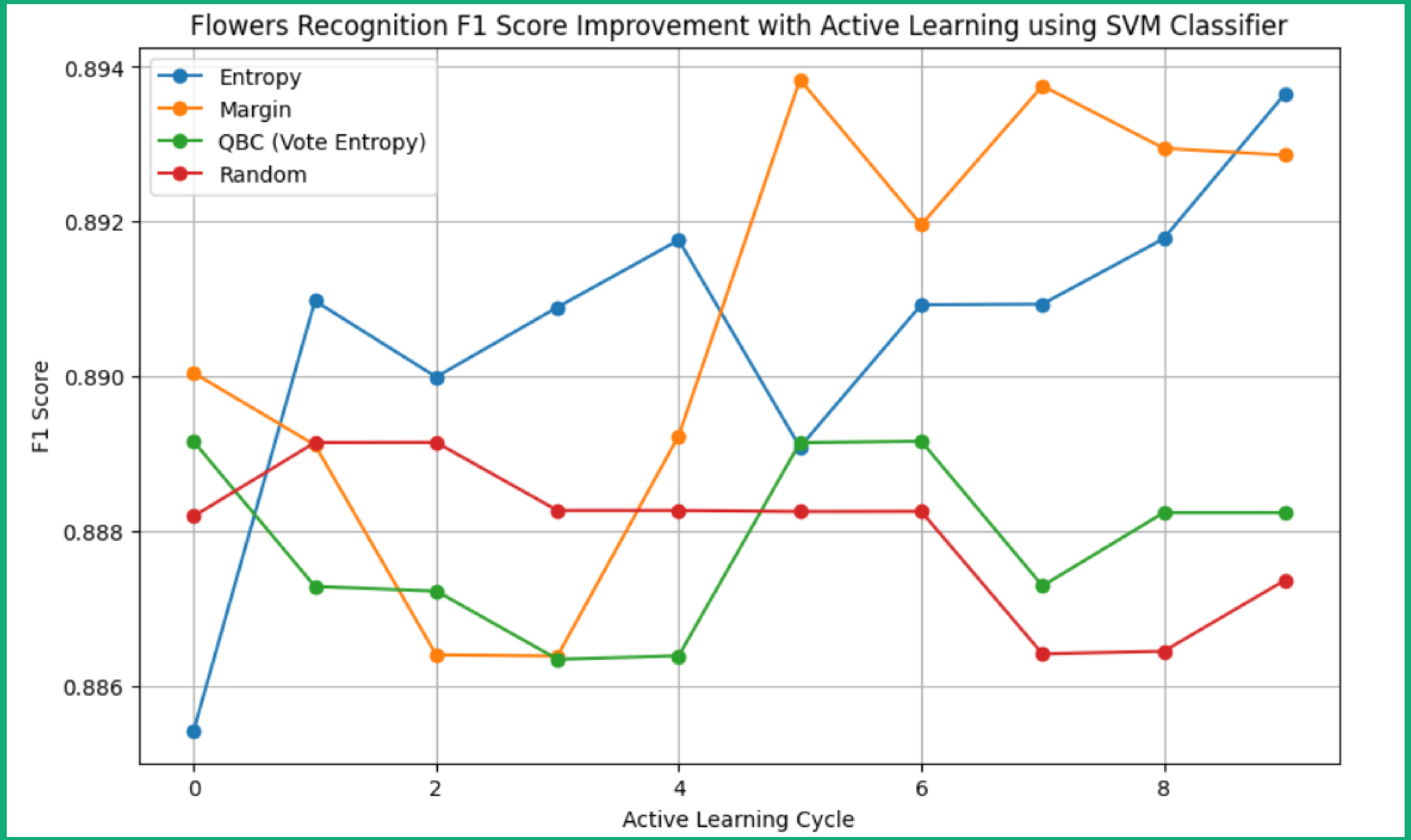
NN Results

Query Strategy	F1 Score
Entropy Sampling	0.87
Margin Sampling	0.88
Query-by-Committee	0.87
Random Sampling	0.88



SVM Results

Query Strategy	F1 Score
Entropy Sampling	0.88
Margin Sampling	0.89
Query-by-Committee	0.88
Random Sampling	0.88



Conclusion

Flowers Recognition Dataset

Through rigorous experimentation, we found the MobileNet model to be highly effective for flower recognition, achieving a testing accuracy of 90.7%, while the Support Vector Machine (SVM) classifier attained an F1 score of 88% on testing data. Additionally, the F1 scores achieved by different strategies were very close to each other, yet the Margin Sampling strategy achieved the highest F1 score of 89% on test data.

Our findings underscore the critical role of thoughtful model selection, activation function optimization, and active learning strategy in achieving robust performance for flower recognition tasks, with implications for future machine learning endeavors.

Flowers Recognition

“Imbalanced”

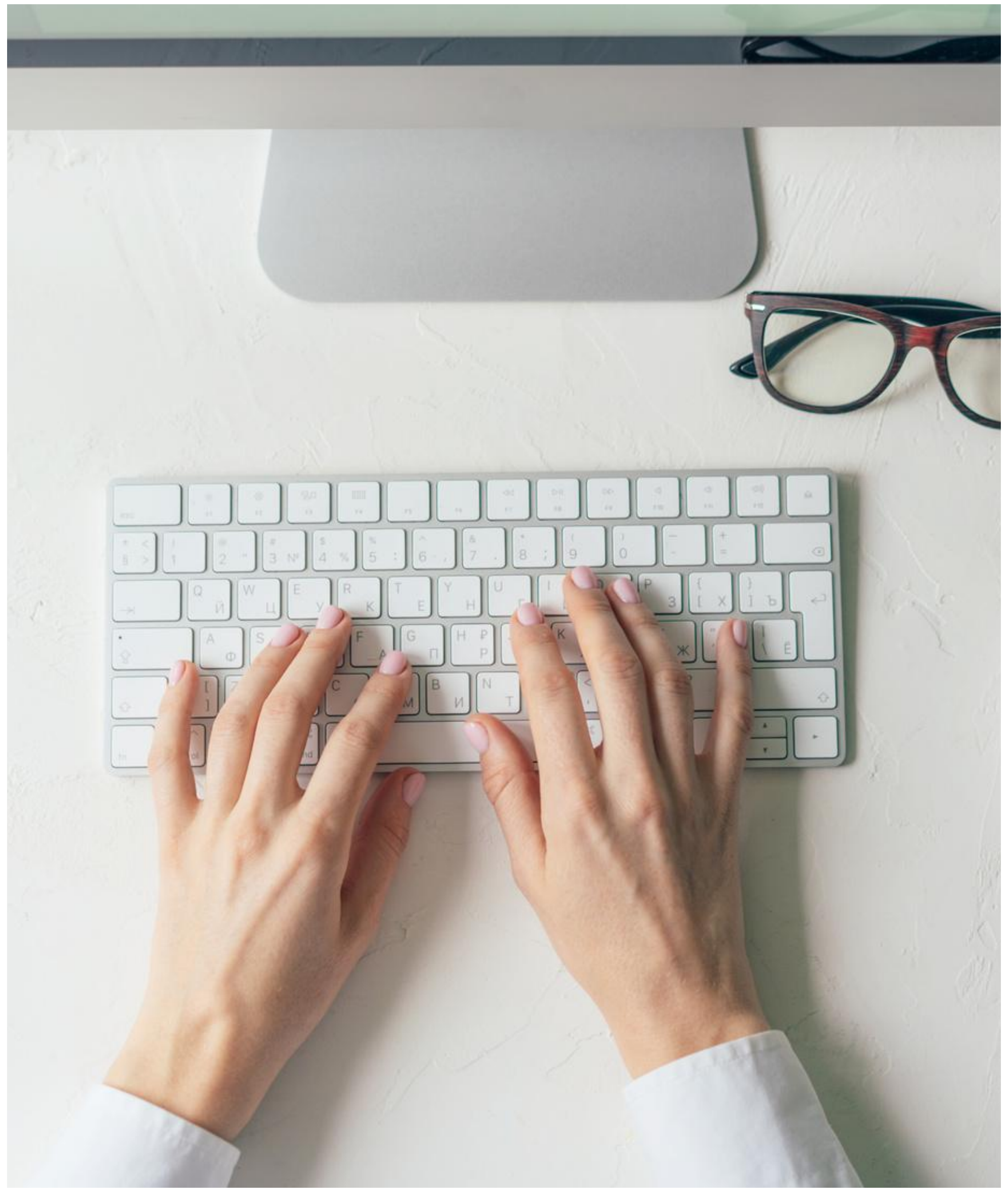
Dataset Description

The dataset is intentionally unbalanced, with Class 0 having most of the samples and the other classes having fewer samples.

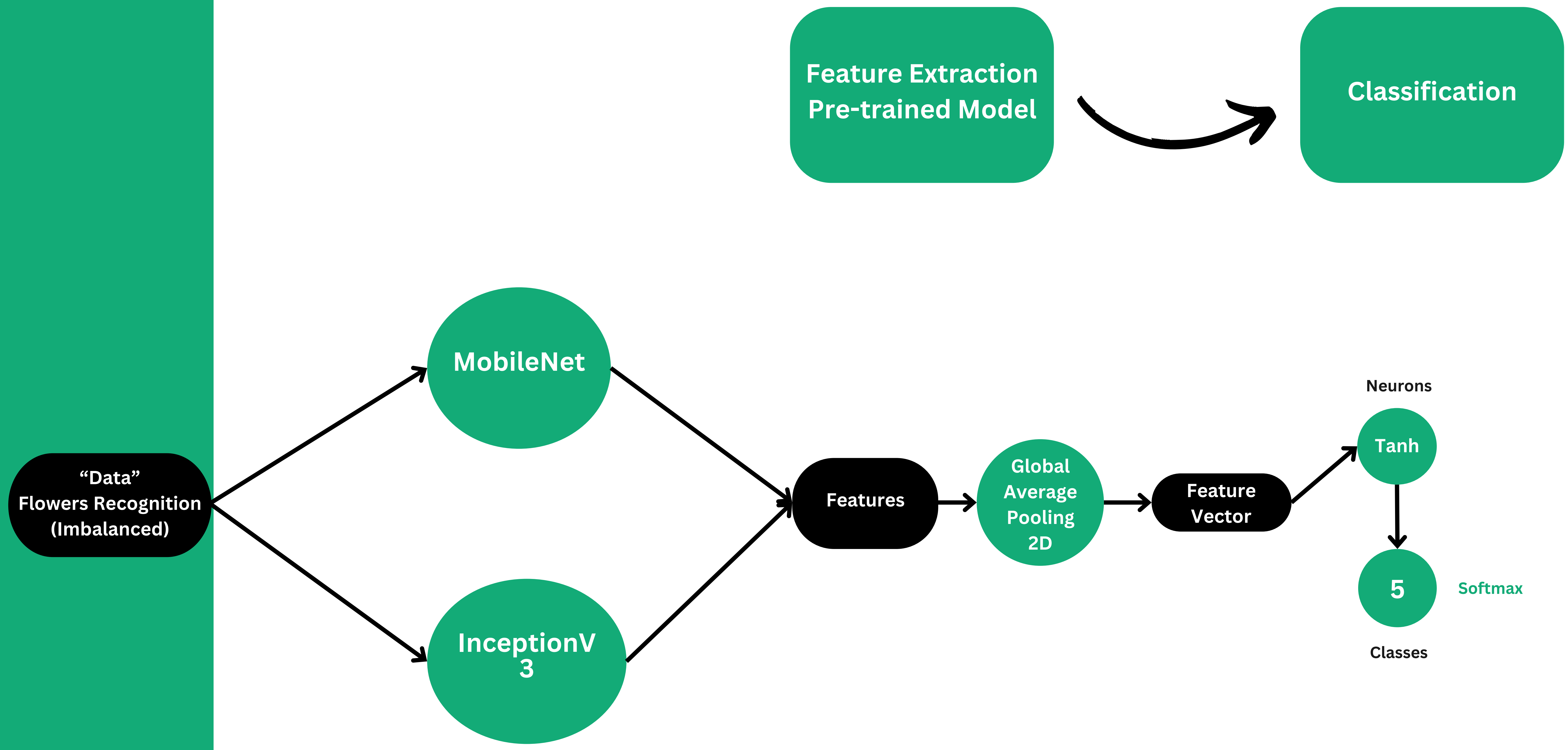
Total Samples: 1202
Class 0: 1052 samples
Class 1: 50 samples
Class 2: 20 samples
Class 3: 30 samples
Class 4: 50 samples

Experiments

- Pretrained Models (Feature Extraction)
- Classifiers
- Query Strategies
- RESULTS & visualization



Pre-trained Models



Pre-trained Models

1. **MobileNet Model**

   F1 Score: 0.8423

2. **InceptionV3 Model**

   F1 Score: 0.8174

Classifiers

Neural Network Architecture:

Using a pretrained model for extracting features and following it with dense layers for classification, a neural network model has been used for classification using: global average pooling, dense layers with tanh activation function, and a softmax output layer for 5-class classification.

Query Strategies

Random Sampling

Random sampling is the simplest and most straightforward query strategy in active learning. It involves randomly selecting instances from the unlabeled dataset for labeling.

Entropy

It measures the entropy of the predicted class probabilities for each instance. Higher entropy indicates higher uncertainty, suggesting that the instance is more informative for labeling.

Margin Sampling

It focuses on instances that lie close to the decision boundary, where the classifier is relatively uncertain. The decision boundary is typically defined by the margin between the probabilities assigned to the top two predicted classes.

Uncertainty Sampling

Uncertainty sampling is a query strategy in active learning where the model selects instances with the highest uncertainty, often measured by low confidence or high entropy, to improve its performance by focusing on challenging examples.

Consensus Entropy Sampling

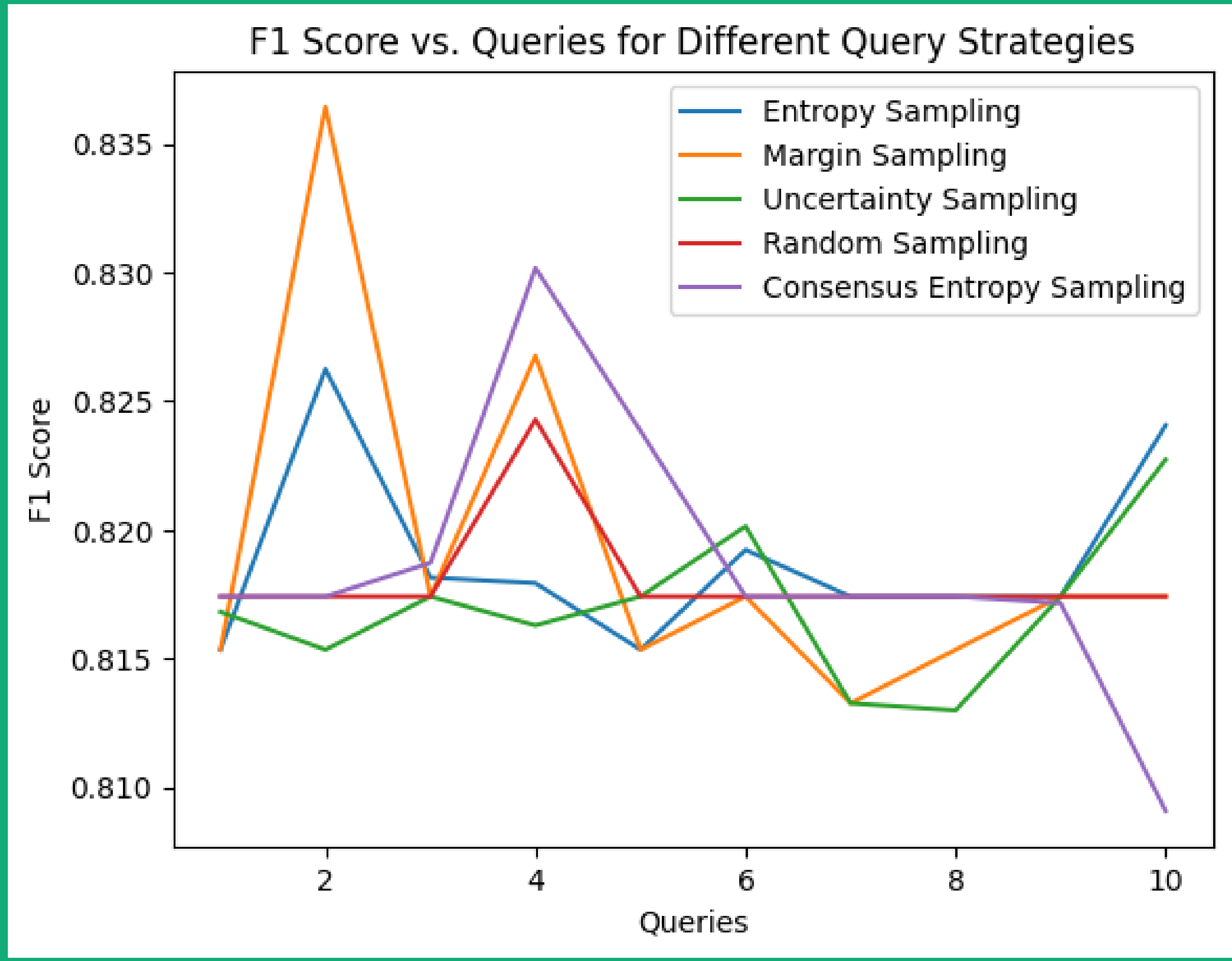
It combines predictions from multiple models and selects instances with high uncertainty, often measured by entropy or disagreement among model predictions. It leverages diverse model perspectives to prioritize labeling challenging data points, enhancing the learning process in active learning scenarios.

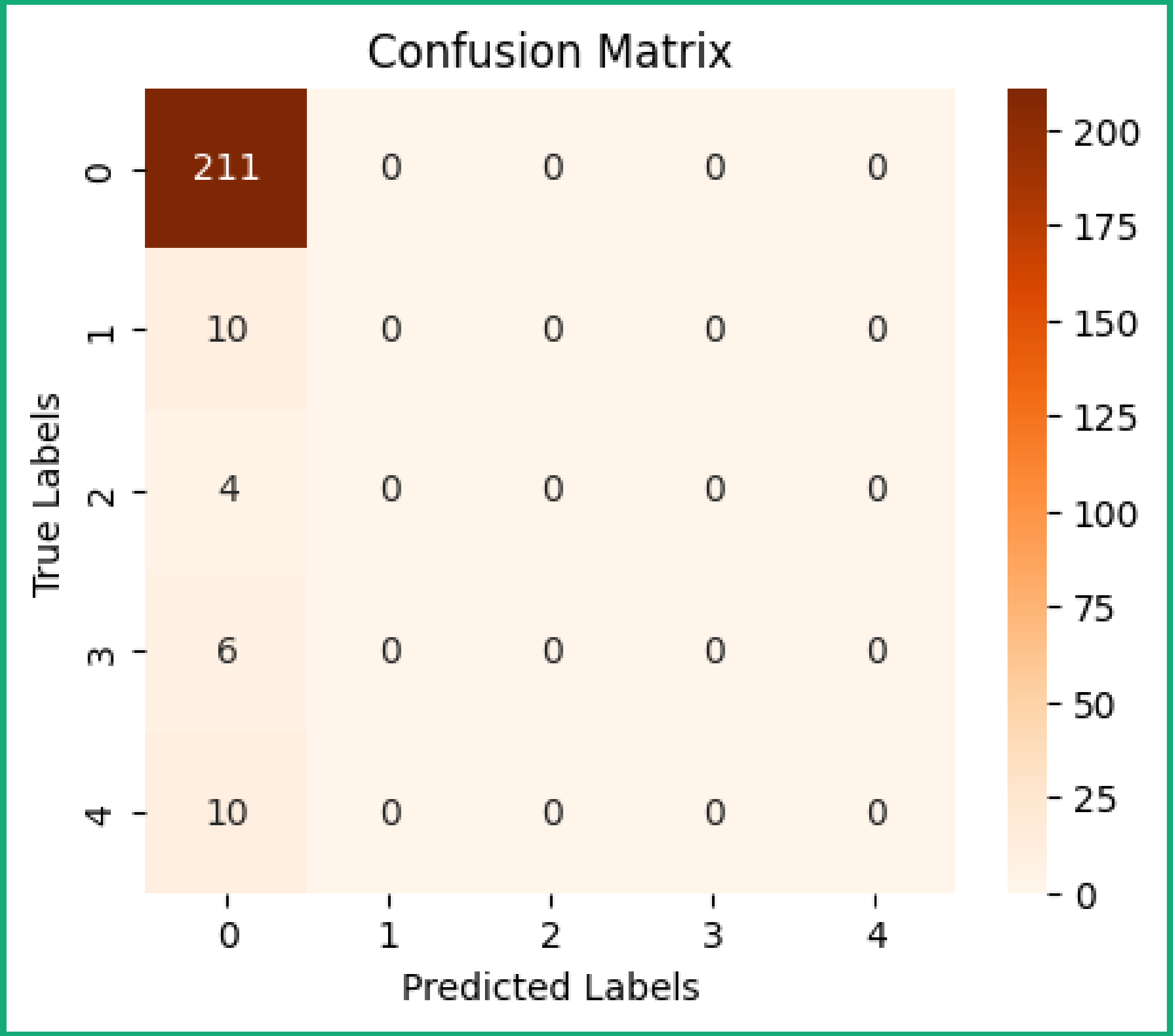
Results

NN × MobileNet Results

Query Strategy	F1 score
Random Sampling	0.8174
Entropy Sampling	0.8240
Margin Sampling	0.8174
Uncertainty Sampling	0.8227
Consensus Entropy Sampling	0.8090

NN × MobileNet Visualization





Conclusion

Flowers Recognition Dataset “Imbalanced”

We have gone through rigorous experiments to pick up the most effective architecture, and our winner was the MobileNet model with: number of epochs / queries = 10, number of instances = 40 and a combination with Uncertainty Sampling, which produced an f1 score of 0.8423.

Our Team

Arwa Sallam
20200067

Mahmoud Wael
20200505

Natalie Monged
20200586

Aly Walid
20200336

Lilian Stephanos
20200404

Thank You

arwasallam6@gmail.com

NatalieMonged20@gmail.com

aly.walidaly@yahoo.com

mahmoudwael677@gmail.com

liliansteven420@gmail.com