**Name: Ahmed Khalaf Mohammed ID: 20190028**

**Name: Youssef Ahmed Hamed ID: 20190627**

**Report**

**Introduction & setup:**

Active learning (AL) is a learning model that helps improve model accuracy by selecting data points for labeling, rather than relying on randomly selected data points. AL algorithms aim to increase the accuracy of models by repeatedly selecting the most informative samples from a large, unlabeled dataset, which are then classified by an expert or the model itself. In this report, we work on four strategies for AL: random sampling, margin sampling, entropy sampling, and least confidence sampling in four datasets: MNIST, Cifar10, Fashion MNIST and Street View House Numbers (SVHN).

* MNIST dataset:

The MNIST dataset consists of 70,000 grayscale images of handwritten numbers, divided into 60,000 training and 10,000 test samples. In this experiment, we initialize a small, labeled set of 50 data points and set the number of instances per query to 50. We then evaluate the three AL strategies in this dataset.

* Cifar10 dataset

The Cifar10 dataset contains 60000 training images and 10000 test images, each of which is a 32x32 color image of one of ten classes. In this experiment, we initialize a small, labeled set of 50 data points and set the number of instances per query to 50. We then evaluate the three AL strategies in this dataset.

# Fashion MNIST Dataset (unbalanced Dataset)

The Fashion-MNIST dataset contains 60,000 training images and 10,000 test images, each of which is a 28x28 grayscale image of a piece of clothing. The dataset is divided into 10 classes, each of which contains 6,000 training images and 1,000 test images. In this experiment, we initialize a small, labeled set of 50 data points and set the number of instances per query to 50. We then evaluate the three AL strategies in this dataset

* SVHN dataset (unbalanced dataset)

SVHN dataset is a set of labeled images of house numbers taken from Google Street View. It is a popular dataset for training machine learning models for digit recognition. It contains 73,257 training images and 26,032 test images, each of which is a 32x32 pixels in size, taken from a variety of angles and lighting conditions, and it is labeled with the corresponding house number. The SVHN dataset is a challenging dataset, but it is a valuable resource for training machine learning models for digit recognition. . In this experiment, we initialize a small, labeled set of 150 data points and set the number of instances per query to 60. We then evaluate the three AL strategies in this dataset.

**Results:**

# At MNIST dataset with 3 Active Learning Strategy

# *Random sampling* Strategy

# 

# Margin Sampling Strategy

# 

# Entropy Sampling Strategy

# 

# Least Confident Sampling Strategy

# 

# Plotting the results

# Training accuracy

# 

# Testing accuracy

# 

# 

* **At Cifar10 dataset with 3 Active Learning Strategy**

# *Random sampling* Strategy

# 

# Margin Sampling Strategy

# 

# Entropy Sampling Strategy

# 

# Least Confident Sampling Strategy

# 

# Plotting the results

# Training accuracy

# 

# Testing accuracy

# 

# At Fashion-MNIST dataset (unbalanced dataset) with 3 Active Learning Strategy

# *Random sampling* Strategy

# 

# Margin Sampling Strategy

# 

# Entropy Sampling Strategy

# 

# Least Confident Sampling Strategy

# 

# Plotting the results

# Training accuracy

# 

# Testing accuracy

# 

# At SVHN dataset (unbalanced dataset) with 3 Active Learning Strategy

# *Random sampling* Strategy

# 

# Margin Sampling Strategy

# 

# Entropy Sampling Strategy

# 

# Least Confident Sampling Strategy

# 

# Plotting the results

# Training accuracy

# 

# Testing accuracy

# 

# Discussion of the results:

# At MNIST dataset

# As shown in the results of MNIST dataset, Margin Sampling performed the best among the four strategies in the MNIST dataset, with a test accuracy of 95.9% after 50 rounds of querying. It selects samples close to the decision boundary between classes, this helps the classifier to learn the boundaries between classes more accurately. Random sampling, Entropy sampling and least confidence sampling also performed well on this dataset, with test accuracy of 94.1%, 95% and 95.7%, respectively, after 50 rounds of querying. Entropy Sampling selects samples with high uncertainty, while Least Confident Sampling selects samples that have the lowest confidence scores, and Random sampling selects a subset of the samples from each class. Margin Sampling was able to identify the most informative samples for labeling, which led to the best performance.

# At *Cifar10* dataset

# As shown in the results of Cifar10 dataset, Least Confident Sampling performed the best among the four strategies, with a test accuracy of 49.1% after 50 rounds of querying. It selects samples that have the lowest confidence scores. Margin Sampling, Entropy Sampling and Random sampling were performed poorly on this dataset, with a test accuracy of 40.5%, 45.1%, 46.1%, respectively, after 50 rounds of querying. Entropy Sampling selects samples with high uncertainty, while Random sampling selects a subset of the samples from each class, and Margin Sampling selects samples that are close to the decision boundary between classes.

# At *Fashion-MNIST dataset (unbalanced dataset)*

# As shown in the results of Fashion-MNIST dataset, Margin Sampling performed the best among the four strategies, with a test accuracy of 74.1% after 50 rounds of querying. It selects samples close to the decision boundary between classes, this helps the classifier to learn the boundaries between classes more accurately. Random sampling, Entropy sampling and least confidence sampling performed poorly on this dataset, with test accuracy of 73.3%, 68.6% and 66.1%, respectively, after 50 rounds of querying. Entropy Sampling selects samples with high uncertainty, while Least Confident Sampling selects samples that have the lowest confidence scores, and Random sampling selects a subset of the samples from each class.

# At *SVHN dataset (unbalanced dataset)*

# As shown in the results of SVHN dataset, Entropy Sampling performed the best among the four strategies, with a test accuracy of 31.2% after 60 rounds of querying. It selects samples with high uncertainty. Margin Sampling, Least Confident Sampling and Random sampling were performed poorly on this dataset, with a test accuracy of 31%, 30.3%, 19.5%, respectively, after 60 rounds of querying. Least Confident Sampling selects samples that have the lowest confidence scores, while Random sampling selects a subset of the samples from each class, and Margin Sampling selects samples that are close to the decision boundary between classes.

# Conclusion:

# In this report, we evaluated four strategies for AL: Random sampling, margin sampling, entropy sampling, and least confidence sampling in four datasets: MNIST, Cifar10, Fashion MNIST and Street View House Numbers (SVHN). The results indicate that the effectiveness of these strategies varies across the dataset. Margin sampling performed best on the MNIST dataset and on the Fashion-MNIST dataset, least confidence sampling performed best on the Cifar10 dataset, and Entropy Sampling performed best on the SVHN dataset.

# Overall, this paper emphasizes how crucial it is to choose an AL approach based on the data set being employed. All four methods work well at finding the most useful samples for labelling, although depending on the dataset, their effectiveness may differ. To improve the quality of their models, researchers should thoroughly assess various AL techniques on their datasets.