Assignment 3

$$KNN\ Implementation\ -\ Experiment\ and\ Result\ Discussion$$

The overall experiment included some steps:

1. Importing data from drive
2. Preprocessing done to enhance image features
    a. Sharpening Image
    b. Blur Image
    c. Histogram Equalization
    d. Thresholding Image
    e. Negative Image
3. Data Reshape and Label Encoding
4. Implementing KNN
    a. Different Metrics with fixed K value
    b. Cross Validation with fixed K value and different metrics
    c. Implementing Grid Search CV
        i. Odd numbers of neighbors from 3 to 19
        ii. Uniform and distance weights
        iii. Euclidean, Manhattan, Chebyshev Metrics

Result Discussion:

Preprocessing:

1. Tried to do different image preprocessing methods. Better results were found after histogram equalization, image sharpening and then negative image. This significantly improved the result.

	Before Preprocessing	After Preprocessing
f1	69%	78%

2. Implemented Grid Search
    a. Best Metric: Manhattan Distance
    b. Best F1: 78%
3. Implemented Cross Validation for improving result.

Drawbacks

1. Time consuming for higher collection of data.
2. No definite rule for determining k value other than conventional sqrt(n), when n is the number of samples.

Strength

1. Easy to implement.
2. Lazy learning.