HOMEWORK ASSIGNMENT 3

SVM For image classification

1. CIFAR 10 dataset was imported and SVM was implemented in this dataset.

1.1 PCA was performed using sklearn and 90% of variance was retained on the dataset

retaining 90 % variance

```
[ ] # Calculating optimal k

    K = 0
    total = sum(pca.explained_variance_)
    current_sum = 0

    while(current_sum / total < 0.90):
        current_sum += pca.explained_variance_[K]
        K += 1
    print(K)
```

```
96
```

```
[ ] pca.explained_variance_
```

```
array([3.65294943e+06, 1.40842701e+06, 8.19020367e+05, ...,
       6.00529319e-02, 5.91042461e-02, 5.90691286e-02])
```

```
## Applying PCA with  calculated K where 90 % variance is retained

pca = PCA(n_components=K, whiten=True, svd_solver='randomized')

X_train = pca.fit_transform(X_train)
```

1.2 Using Grid search CV with 5 folds to find best hyperparameters for SVMs with linear (C) and RBF kernels (C, gamma).

Hyperparameter tuning with Grid search CV

```
[ ] param_grid = {'kernel': ('linear', 'rbf'),   ### defining  parameters
                  'C': [100, 10, 1]}
```

```
[ ] grid = GridSearchCV(SVC(), param_grid,refit = True, verbose = 3, n_jobs=8)
```

```
[ ] grid.fit(X_train, Y_train)  ##fitting the model for grid search
```

```
Fitting 5 folds for each of 6 candidates, totalling 30 fits
GridSearchCV(estimator=SVC(), n_jobs=8,
             param_grid={'C': [100, 10, 1], 'kernel': ('linear', 'rbf')},
             verbose=3)
```

```
[ ]
    # printing best parameter after tuning
    print(grid.best_params_)
```

```
{'C': 10, 'kernel': 'rbf'}
```

```
[ ] # printing the model looks after hyper-parameter tuning
    print(grid.best_estimator_)
```

```
SVC(C=10)
```

1.3 Use SVM with best hyperparameters from above to classify this dataset through one vs rest (OVR) approach. Report the accuracy over five folds along with mean accuracy. Also, report mean class accuracy.

OVO

Accuracy over 5 folds

```
scores
```

```
array([0.471, 0.463, 0.464, 0.471, 0.47 ])
```

Mean Accuracy

```
print('Mean Accuracy: %.3f ' % (mean(scores)))
```

```
Mean Accuracy: 0.468
```

Mean class accuracy

```
array([0.53731343, 0.61190965, 0.35852713, 0.33759843, 0.39139139,
       0.36392743, 0.48349515, 0.51948052, 0.56      , 0.58409786])
```

1.4 Use SVM with best hyperparameters from above to classify this dataset through one vs rest (OVR) approach. Report the accuracy over five folds along with mean accuracy. Also, report mean class accuracy.

OVR

Accuracy over 5 folds

```
scores
```

```
array([0.471, 0.463, 0.464, 0.471, 0.47 ])
```

Mean accuracy

```
print('Mean Accuracy: %.3f ' % (mean(scores)))
```

```
Mean Accuracy: 0.468
```

Mean class accuracy

```
array([0.53731343, 0.61190965, 0.35852713, 0.33759843, 0.39139139,
       0.36392743, 0.48349515, 0.51948052, 0.56      , 0.58409786])
```

The accuracy over five folds along with mean accuracy and mean class accuracy was found to be same for both OVO and OVR approach.

Theory Questions

1. Using a kernel $K(x,y) = \phi(x)^T \phi(y)$ we can transform data into a higher dimension for classifying inseperable data. ($\phi(x)$ is a feature mapping function for input x). Given two points $x = (x1, x2)$ and $y = (y1, y2)$, what would be the feature mapping function ϕ for the kernel $K(x,y) = (x^T y)^2$.

→ Given, feature mapping $\phi(x)$ for input x
the kernel is given by

$$K(x,y) = \phi(x)^T \phi(y)$$

we can replace the inner product
$\langle \phi(x), \phi(z) \rangle$ with $K(x,y)$ in svm algorith.

Given two points $x = (x1, x2)$ and $y = (y1, y2)$

$$x, y \in R^n, \quad K(x,y) = (x^T y)^2$$

$$K(x,y) = \left(\sum_i^n x_i y_i\right)\left(\sum_i^n x_i y_i\right)$$

$$= \sum_i^n \sum_j^n x_i x_j y_i y_j$$

$$\cancel{= \sum_{i,j}^n (x_i y_j)(y_i x_j\ (y_i y_j)}$$

$$= \sum_{i,j}^n (x_i x_j)(y_i y_j)$$

$= \phi(x^T)\phi(y)$

Here, ϕ is the feature mapping given by.

$$\phi(x) = \begin{bmatrix} x_1 x_1 \\ x_1 x_2 \\ x_2 x_1 \\ x_2 x_2 \end{bmatrix}$$

Here calculating the feature mapping is of complexity $O(n^2)$ and calculating $K(x, z)$ is of complexity $O(n)$ since it is a ~~A~~ inner product $x^T y$ which is then squared.

$K(x, z) = (x^T z)^2$