Experiment Report: Observations and Analysis of SVM Kernel and Parameter Settings

Objective:

The goal of this experiment was to investigate the performance of Support Vector Machines (SVMs) in the context of a classification job using various kernels and parameter settings. The linear, radial basis function (RBF), and polynomial (poly) kernels were the primary focus of this experiment, each with three alternative sets of parameters.

Kernel and Parameter Settings:

For the linear kernel, three different settings of the regularization parameter (C) were used: 1, 10, and 100.

For the RBF kernel, three different settings of C and the gamma parameter were explored: (C: 1, gamma: 0.001), (C: 10, gamma: 0.001), and (C: 100, gamma: 0.01).

For the polynomial kernel, three different settings of C and the degree were considered: (C: 1, degree: 3), (C: 1, degree: 5), and (C: 10, degree: 3).

Results:

Linear Kernel:

```
Kernel: Linear, Parameters: {'C': 1}
Training Error: 0.2010
Cross-Validation Error: 0.2270
Test Error: 0.2133

Kernel: Linear, Parameters: {'C': 10}
Training Error: 0.2103
Cross-Validation Error: 0.2460
Test Error: 0.2288

Kernel: Linear, Parameters: {'C': 100}
Training Error: 0.2163
Cross-Validation Error: 0.2490
Test Error: 0.2242
```

Fig.1 Linear Kernel Result

With the 'C' parameter set to 1, the linear kernel model achieved a training error of 0.2010, a cross-validation error of 0.2270, and a test error of 0.2133. As the value of 'C' increased, so did the training and cross-validation errors, indicating overfitting.

When compared to the 'C=1' scenario, the model with 'C' set to 10 performed somewhat worse, with a greater training error (0.2103) and cross-validation error (0.2460).

Increasing 'C' to 100 resulted in increased training error (0.2163), cross-validation error (0.2490), and test error (0.2242). The overfitting trend continued.

RBF Kernel:

```
Kernel: RBF, Parameters: {'C': 1, 'gamma': 0.001}
Training Error: 0.0000
Cross-Validation Error: 0.4880
Test Error: 0.4895

Kernel: RBF, Parameters: {'C': 10, 'gamma': 0.001}
Training Error: 0.0000
Cross-Validation Error: 0.4880
Test Error: 0.4850

Kernel: RBF, Parameters: {'C': 100, 'gamma': 0.01}
Training Error: 0.0000
Cross-Validation Error: 0.4880
Test Error: 0.5014
```

Fig 2. RBF Kernel Result

The model exhibited 0% training error for the RBF kernel with 'C' set to 1 and 'gamma' set to 0.001, indicating that it perfectly suited the training data. The model's cross-validation error, however, was high at 0.4880, indicating poor generalization to unknown data. The test error was likewise high, coming in at 0.4895.

Increasing 'C' to 10 while maintaining the same 'gamma' produced a similar pattern of zero training error, high cross-validation error (0.4880), and somewhat reduced test error (0.4850).

When 'C' was set to 100 and 'gamma' was set to 0.01, the model still had 0% training error. The cross-validation error, however, remained high at 0.4880, while the test error climbed dramatically to 0.5014. This implies that raising 'C' and 'gamma' resulted in overfitting, in which the model fit the training data too closely yet performed badly on new data.

Polynomial Kernel:

```
Kernel: Polynomial, Parameters: {'C': 1, 'degree': 3}
Training Error: 0.2770
Cross-Validation Error: 0.3250
Test Error: 0.2862

Kernel: Polynomial, Parameters: {'C': 1, 'degree': 5}
Training Error: 0.2887
Cross-Validation Error: 0.3570
Test Error: 0.2990

Kernel: Polynomial, Parameters: {'C': 10, 'degree': 3}
Training Error: 0.2362
Cross-Validation Error: 0.2700
Test Error: 0.2534
```

Fig 3. Polynomial Kernel Result

The model had a training error of 0.2770, a cross-validation error of 0.3250, and a test error of 0.2862 for the polynomial kernel with 'C' set to 1 and 'degree' set to 3. Even though the training and test errors were quite close, the model demonstrated moderate generalization ability.

Both training and cross-validation errors rose when 'degree' was increased to 5 but 'C' remained at 1. With a training error of 0.2887, a cross-validation error of 0.3570, and a test error of 0.2990, the model overfitted.

Setting 'C' to 10 and a 'degree' of 3 resulted in better performance than 'C=1, degree=5.' The error in training was 0.2362, the error in cross-validation was 0.2700, and the error in testing was 0.2534. This implies that increasing the 'C' value improves the model's performance and generalization while still exhibiting some overfitting tendencies.

Overall Observations:

Among the investigated kernels and parameter settings, the linear kernel outperformed the RBF and polynomial kernels and was less prone to overfitting.

Overfitting was seen in the RBF kernel, particularly at larger 'C' and 'gamma' values.

The performance of the polynomial kernel varied with 'C' and 'degree.' A greater 'C' value resulted in better generalization, whereas a higher 'degree' resulted in overfitting.

Conclusions:

In terms of generalization, 'C=1' offered the best results for the linear kernel.

To avoid overfitting, extremely high 'C' and 'gamma' values are required for the RBF kernel.

In the case of the polynomial kernel, the combination of 'C=10' and 'degree=3' seems to provide a superior balance of training and test performance.

Experimenting with other parameter values and maybe different kernel types could assist achieve even better model performance. It is critical to fine-tune hyperparameters based on the unique dataset and task requirements.