COMP3202 Assignment 3

Report

Group 28

Preface

As the note on grading states that the models in Q4 will be ranked based on the balanced accuracy of the model, we have chosen balanced accuracy as our score for the refit parameter of the GridSearchCV function. Additionally, the performance reporting convention states that we must always report both sensitivity and recall, despite these being two words that describe the same performance metric (i.e. TP/(TP+FN)). We have followed the instructions and for each performance stated the value for both sensitivity and recall.

Abstract

We were given the task of training a support vector machine classification model (SVM) to attempt to discriminate between individuals with and without Alzheimer's Disease through the glucose metabolism levels of various regions of the brain. We trained and tested three SVM models to determine which would be best suited to this task. We measured the models' performance in accuracy, sensitivity/recall, specificity, precision, and balanced accuracy. We found that the linear SVM performed the worst at the task in all metrics, and that the non-linear SVM with a polynomial kernel and the non-linear SVM with a radial basis function kernel performed roughly equivalently, with the former having marginally better performance. We recommend training the latter two models on more data to more concretely determine which is better suited to the given task.

Q1.

We began by training a linear SVM for the task. The range of possible values for the regularization parameter C for a SMV is $(0, \infty)$. As such, the first task was to limit the scope of C values we would consider. According to Chi-Wei Hsu et al. “trying exponentially growing sequences of [your parameters] is a practical method to identify good parameters” (2003, p.5). As such, we chose the sequence 2^x where x ranges from -5 to 15 as our search space for C. After using a cross-validation based grid search of the aforementioned search space, we found that 2^4, or 16, produced the best balanced accuracy. We then re-trained our linear SVM on the entire “training” dataset with regularization parameter C set to 16 to achieve our final linear SVM classification model. Next we estimated the performance of the model (“Err”) by using it to predict the outcome of a previously unseen “test” dataset and comparing the model’s predictions to the labeled outcomes. The results of this estimation can be seen in Figure 1.

```
Generating results for Q1...
Performing classification...
Best C found to be 16
Linear SVM - Accuracy Score: 0.8343
Linear SVM - Sensitivity Score: 0.9426
Linear SVM - Specificity Score: 0.8024
Linear SVM - Precision Score: 0.5838
Linear SVM - Recall Score: 0.9426
Linear SVM - Balanced Accuracy Score: 0.8725
```

Figure 1. “Err” of linear SVM on “test” dataset, C = 16

The accuracy of the model was found to be 0.8343, the sensitivity/recall was 0.9426, the specificity was 0.8024, the precision was 0.5838, and the balanced

accuracy was 0.8725. These scores show that the model is “good” in every metric except precision, which indicates that the model is overly likely to predict that a given individual is Alzheimer’s Disease positive. Figures 2 - 7 show the performances of the models explored with regards to C. They show accuracy, sensitivity, specificity, precision, recall, and balanced accuracy respectively. Note that the x-axis of the plots has a logarithmic scale with base 2.

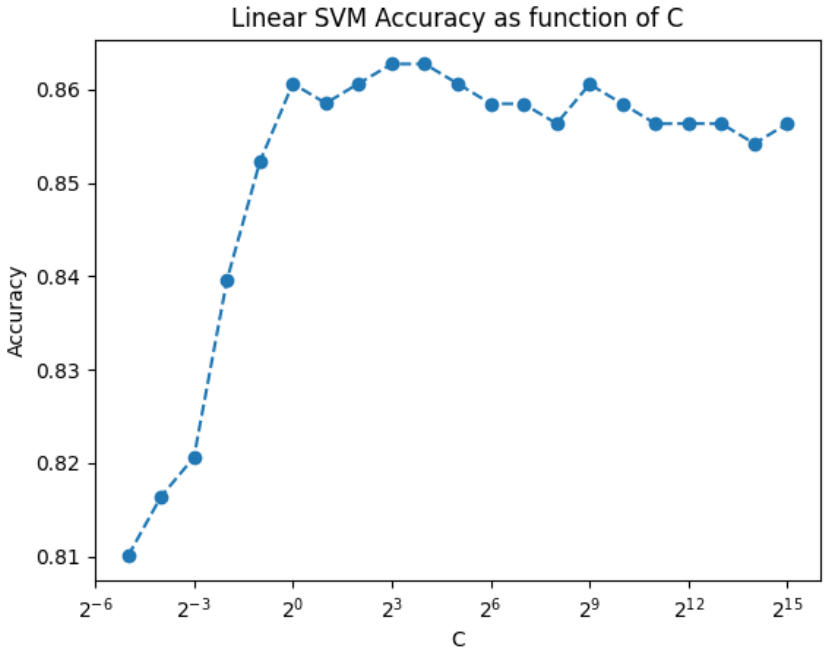

Figure 2. Accuracy of explored linear SVMs with respect to regularization parameter C

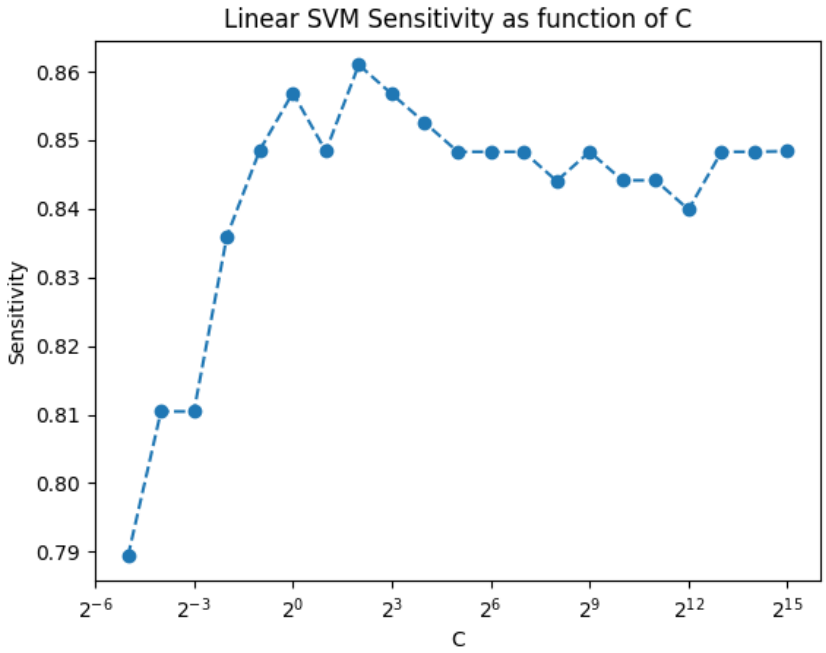

Figure 3. Sensitivity of explored linear SVMs with respect to regularization parameter C

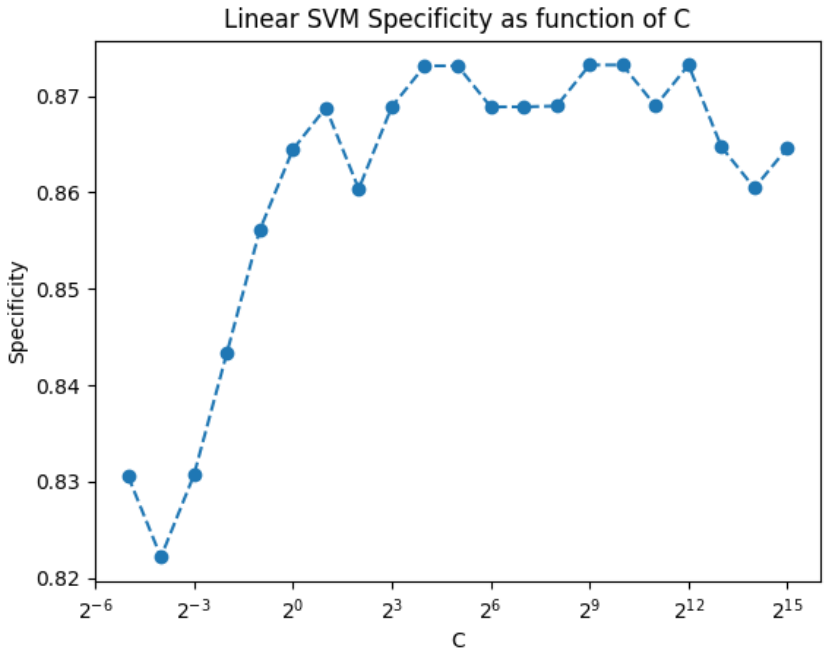

Figure 4. Specificity of explored linear SVMs with respect to regularization parameter C

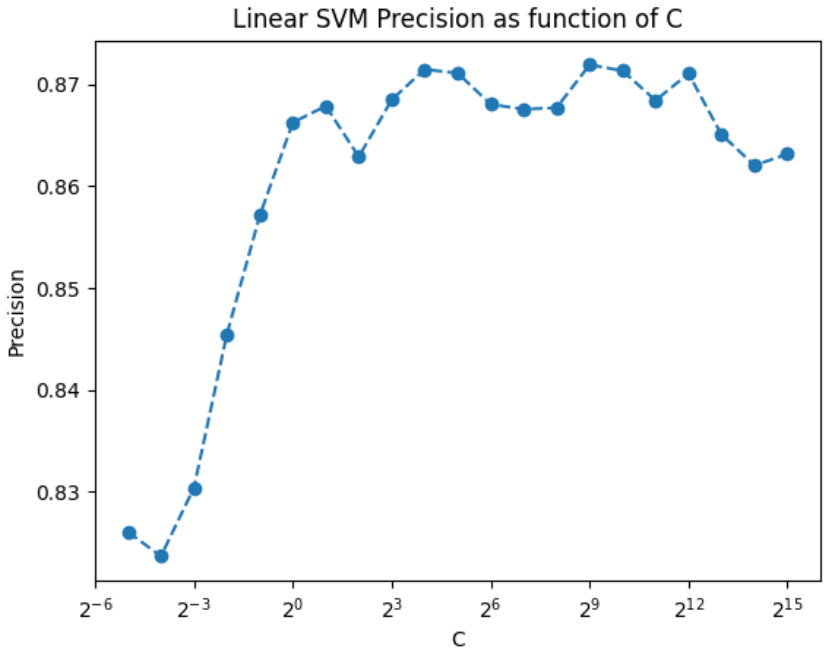

Figure 5. Precision of explored linear SVMs with respect to regularization parameter C

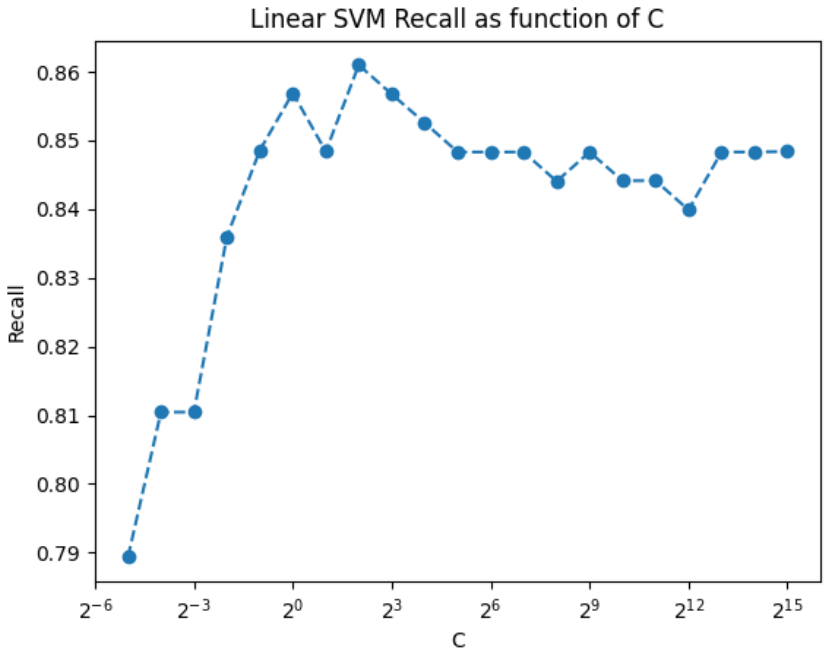

Figure 6. Recall of explored linear SVMs with respect to regularization parameter C

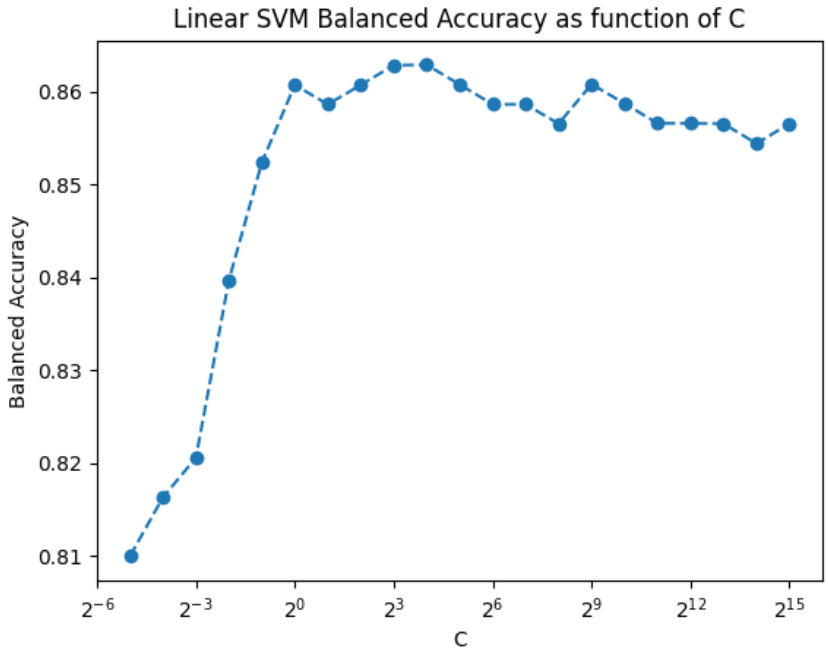

Figure 7. Balanced accuracy of explored linear SVMs with respect to regularization parameter C

The plots reveal a general trend in the performance metrics wherein the metrics are relatively low for small values of C (< 1) and tend to peak between C values of 4 and 32. The metrics then converge around 0.85 to 0.87, trending slightly downward for higher and higher values of C. This indicates that the ideal regularization parameter for attempting Alzheimer's Disease diagnoses through the given glucose metabolism measures falls within the low tens for a linear SVM.

Q2.

We next trained a non-linear SVM with a polynomial kernel (polySVM) for the same task. We again used a cross-validation based grid search to try and find the optimal values for the regularization parameter C and the degree parameter d. As with Q1, we used the sequence 2^x where x ranges from -5 to 15 as our search space for C. For the search space of d, it was found that for each higher degree added to the search space, exponentially longer processing times were required for the grid search. We were therefore forced to limit our search space to the values from 0 to 4 inclusive. The best parameters found by our search were a C value of 1 and a d value of 3. We then re-trained our polySVM with these parameters and estimated the "Err" using a previously unseen "test" dataset. The results of this can be seen in Figure 8.

```
Generating results for Q2...
Performing classification...
Best C found to be 1
Best d found to be 3
Polynomial kernel SVM - Accuracy Score: 0.8734
Polynomial kernel SVM - Sensitivity Score: 0.9508
Polynomial kernel SVM - Specificity Score: 0.8506
Polynomial kernel SVM - Precision Score: 0.6517
Polynomial kernel SVM - Recall Score: 0.9508
Polynomial kernel SVM - Balanced Accuracy Score: 0.9007
```

Figure 8. "Err" of non-linear SVM with a polynomial kernel on "test" dataset; C = 1, d = 3

The accuracy of the polySVM was found to be 0.8734, the sensitivity/recall was 0.9508, the specificity was 0.8506, the precision was 0.6517, and the balanced accuracy was 0.9007.

When comparing the “Err” of the polySVM to the linear SVM we found that the polySVM produced better results in every performance metric. The smallest improvement was observed in the sensitivity/recall where only an increase of 0.0082 was seen. The largest improvement was observed in precision with an increase of 0.0679. The improvements in accuracy, specificity, and balanced accuracy were 0.0391, 0.0482, and 0.0282 respectively. We consider all of these significant improvements in the efficacy of the model and conclude that between the linear SVM and the polySVM, the polySVM is better suited for attempting Alzheimer’s Disease diagnoses through the given glucose metabolism measures.

Q3.

Lastly, we trained a non-linear SVM with a radial basis function kernel (rbfSVM) for the task. Once again we utilized a cross-validation based grid search to attempt to find the best values for the rbfSVM’s regularization parameter, C, and gamma parameter, γ. We used the same range of values for C as in Q1 and Q2. For the values of γ we followed the recommendation on Chi-Wei Hsu et al. and used the sequence 2^x for x in range -15 to 3 (2003, p.5). The result of the grid search indicated that the best parameters for our rbfSVM were a C value of 32 and a γ value of 0.50. We then re-trained our model with these parameters and used a previously unseen “test” dataset to estimate the “Err” of our model; the results of this can be seen in Figure 9.

```
Generating results for Q3...
Performing classification...
Best C found to be 32
Best γ found to be 0.5000
RBF kernel SVM - Accuracy Score: 0.8715
RBF kernel SVM - Sensitivity Score: 0.9426
RBF kernel SVM - Specificity Score: 0.8506
RBF kernel SVM - Precision Score: 0.6497
RBF kernel SVM - Recall Score: 0.9426
RBF kernel SVM - Balanced Accuracy Score: 0.8966
```

Figure 9. “Err” of non-linear SVM with a radial basis function kernel on “test” dataset; C = 32, γ = 0.50.

We found the accuracy of the rbfSVM to be 0.8715, the sensitivity to be 0.9426, the specificity to be 0.8506, the precision to be 0.6497, the recall to be 0.9426, and the balanced accuracy to be 0.8966.

When compared to the linear SVM, the rbfSVM performs markedly better in every metric except sensitivity/recall where it performs exactly the same. The increase was 0.0372 for accuracy, 0.0482 for specificity, 0.0659 for precision, and 0.0241 for balanced accuracy. The increase in the metrics other than sensitivity/recall are enough for us to be confident in asserting that the rbfSVM is the better model for this task.

When comparing the rbfSVM to the polySVM, we find that in every metric except specificity, the rbfSVM performs marginally worse than the polySVM and in the case of specificity the models performed the same. The decrease is 0.0019 in accuracy, 0.0082 in sensitivity/recall, 0.0020 in precision, 0.0041 in balanced accuracy. While this seems to indicate that the polySVM is the better model for the task, the difference in the performances of the models is low enough that we aren’t confident in stating that this is a definite fact, and we believe that testing the models on more data is necessary to make a certain statement.

That being said, we have decided to use a polySVM for our model in Q4 as while we do not think the results are conclusive in and of themselves, they are the only results available for us to make our decision on.

References

Hsu, Chih-wei & Chang, Chih-chung & Lin, Chih-Jen. (2003). A Practical Guide to Support Vector Classification. Retrieved from https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf on April 17, 2023.