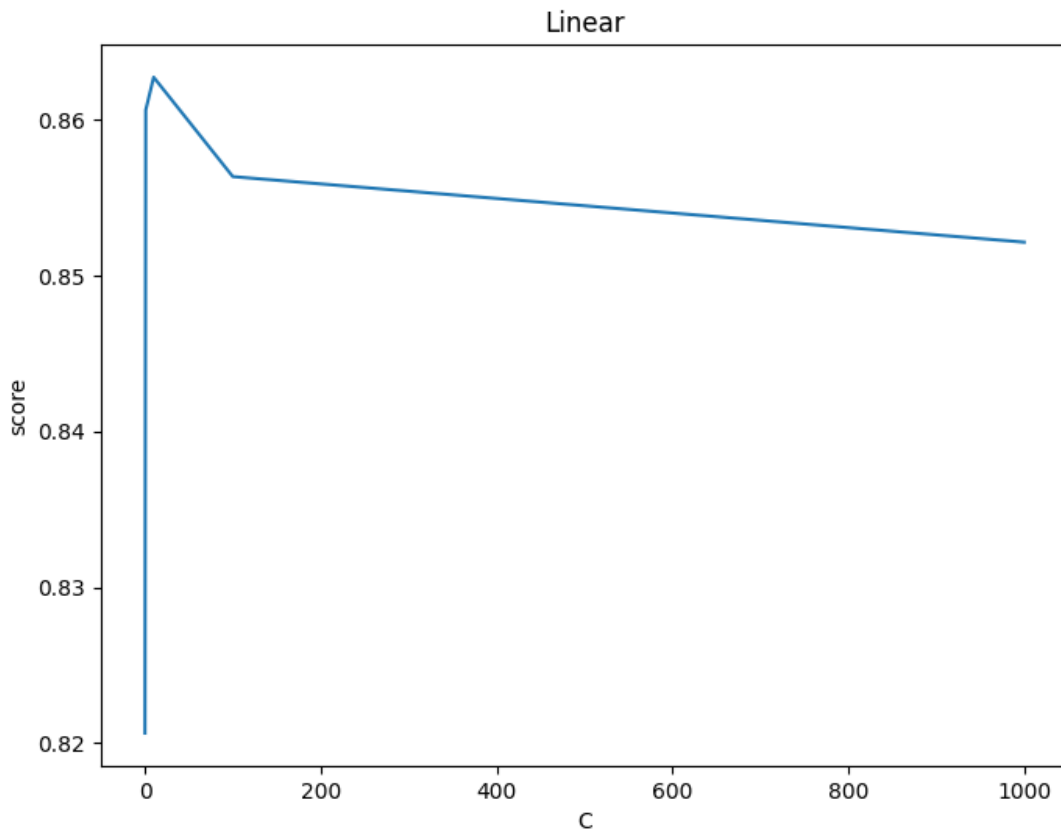# Question 1:

Figure 1 – Performance of *linear* SVM model as a function of *"C"*

As observed above, there is an extremely steep increase in the performance of the *linear* SVM model, from 0 to 8 then another increase at 10. From 10, the performance decreases fairly steeply again, and then tapers off as C approaches larger values.

Question 2:

```
kernel: linear
C: 10
accuracy: 0.8361266294227188
sensitivity: 0.9344262295081968
specificity: 0.8072289156626506
precision: 0.5876288659793815
recall: 0.9344262295081968
balanced accuracy: 0.8708275725854238
================================
kernel: poly
C: 100
d: 2
accuracy: 0.8715083798882681
sensitivity: 0.9426229508196722
specificity: 0.8506024096385543
precision: 0.6497175141242938
recall: 0.9426229508196722
balanced accuracy: 0.8966126802291132
```

Figure 2 – Performance comparison of *linear* SVC to *poly* SVC

From the figure above, there was an improvement in performance across all metrics, using the *poly* SVC model. This indicates that the domain problem (flagging of Alzheimer's) is either better described as a non-linear problem, or it just benefits in general from the enhanced computation that a *poly* model provides.

Question 3:

```
kernel: linear
C: 10
accuracy: 0.8361266294227188
sensitivity: 0.9344262295081968
specificity: 0.8072289156626506
precision: 0.5876288659793815
recall: 0.9344262295081968
balanced accuracy: 0.8708275725854238
================================
kernel: poly
C: 100
d: 2
accuracy: 0.8715083798882681
sensitivity: 0.9426229508196722
specificity: 0.8506024096385543
precision: 0.6497175141242938
recall: 0.9426229508196722
balanced accuracy: 0.8966126802291132
================================
kernel: rbf
C: 10
gamma: 1
accuracy: 0.8733705772811918
sensitivity: 0.9426229508196722
specificity: 0.8530120481927711
precision: 0.6534090909090909
recall: 0.9426229508196722
balanced accuracy: 0.8978174995062216
================================
```

Fig. 3 – Performance comparison of *linear*, *poly*, and *rbf* SVC models

As expected, the *rbf* model outperformed both the *linear* and the *poly* models, across all metrics. In fact, attempting to fine-tune *"C"* values revealed that the model's performance remains the same from about 8.6-10, indicating the strength of the *rbf* model.

Question 4 (diagnoseDAT)

In an attempt to improve the performance of our SVC model, we ran a grid search between our best kernel variation thus far, *rbf*, and a new kernel type, *sigmoid* (and its *coef0* hyper-parameter). The *rbf* kernel outperformed the sigmoid, so we retained it.

Subsequently, we tried additional hyper-parameters - *shrinking*, and *class_weight* - while continuously fine-tuning additional *"C"* and *"γ"*. Setting *"γ"* to 'scale'' yielded an improvement in the performance as indicated by the **accuracy**, **specificity**, **precision**, and **balanced accuracy** metrics of the model, when compared to the *rbf* model from Question 3.

```
kernel: rbf
C: 10
gamma: scale
accuracy: 0.8752327746741154
sensitivity: 0.9426229508196722
specificity: 0.8554216867469879
precision: 0.6571428571428571
recall: 0.9426229508196722
balanced accuracy: 0.89902231878333
```

Fig. 4 - model performance improvement