Summary:

| **Key Term** | **Definition** |
| --- | --- |
| C | Also referred to as the C hyper-parameter - determines how flexible we are willing to be with the points that fall on the wrong side of our dividing boundary. Large C is mostly the classification error. Small C is mostly margin error. |
| Classification error | An error in prediction caused by a false negative response or a false positive response. |
| Error function | Classification error + margin error. Used to minimize SVG |
| Gamma | A hyperparameter that we tune during training to alter the type of curve from wide ones to very narrow. |
| Kernal Trick | A more efficient and less expensive way to transform data into higher dimensions, especially in the SVM algorithm |
| Marthin error | The distance between the 2 boundary lines and prediction line. Use to minimize gradient descent |
| Ploynomial Kernel | A kernel function commonly used with SVM that allows learning of non-linear models. |
| RBF | Radial Basis Functions are used in machine learning to find a non-linear classifier or regression lines, especially in SVM. |

Extra Resources:

Page 337 in this book: <https://static1.squarespace.com/static/5ff2adbe3fe4fe33db902812/t/6062a083acbfe82c7195b27d/1617076404560/ISLR%2BSeventh%2BPrinting.pdf>

From page 11 in Andrew Ng Notes:

<http://cs229.stanford.edu/notes2020fall/notes2020fall/cs229-notes3.pdf>