



CS6316 Machine Learning Practical Project

Key Results and Improvements

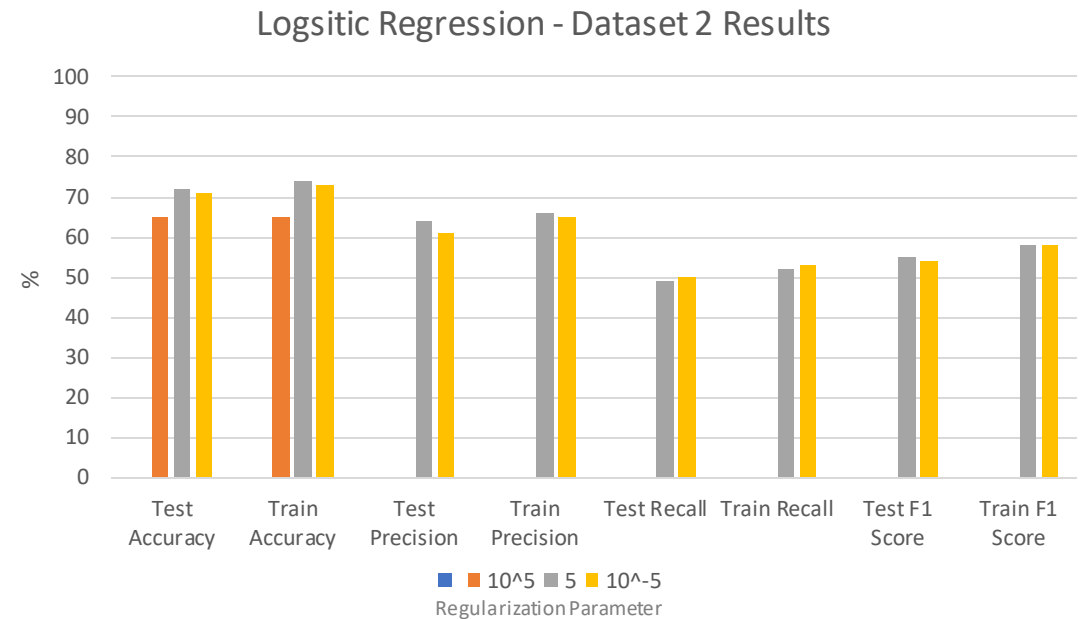
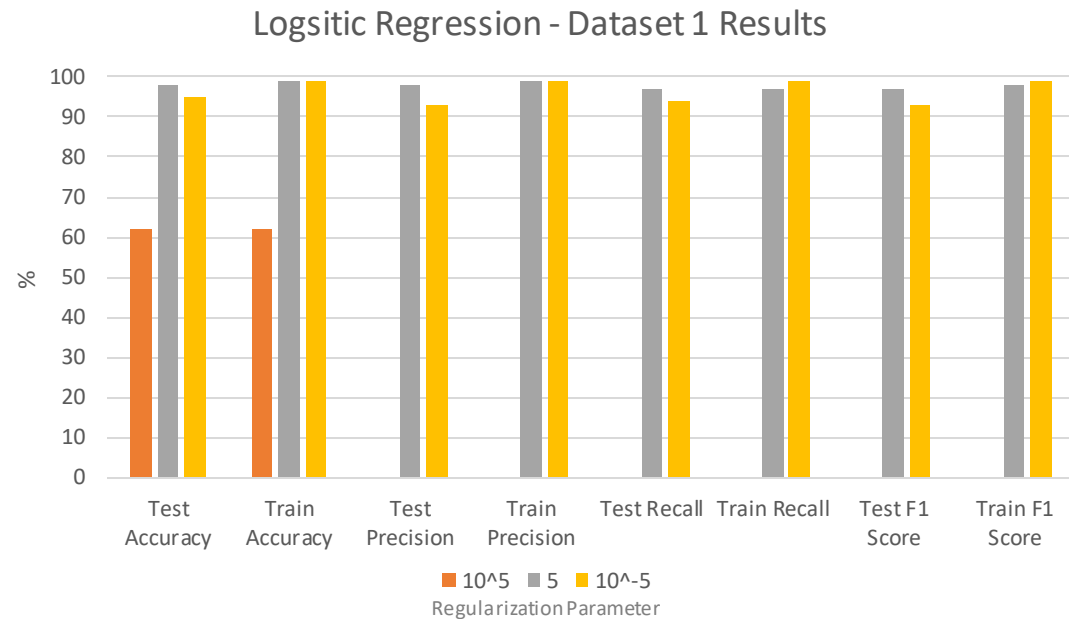
Keshara Weerasinghe, Jiho Lee, Rishabh Jain





Logistic Regression

- Results and Analysis

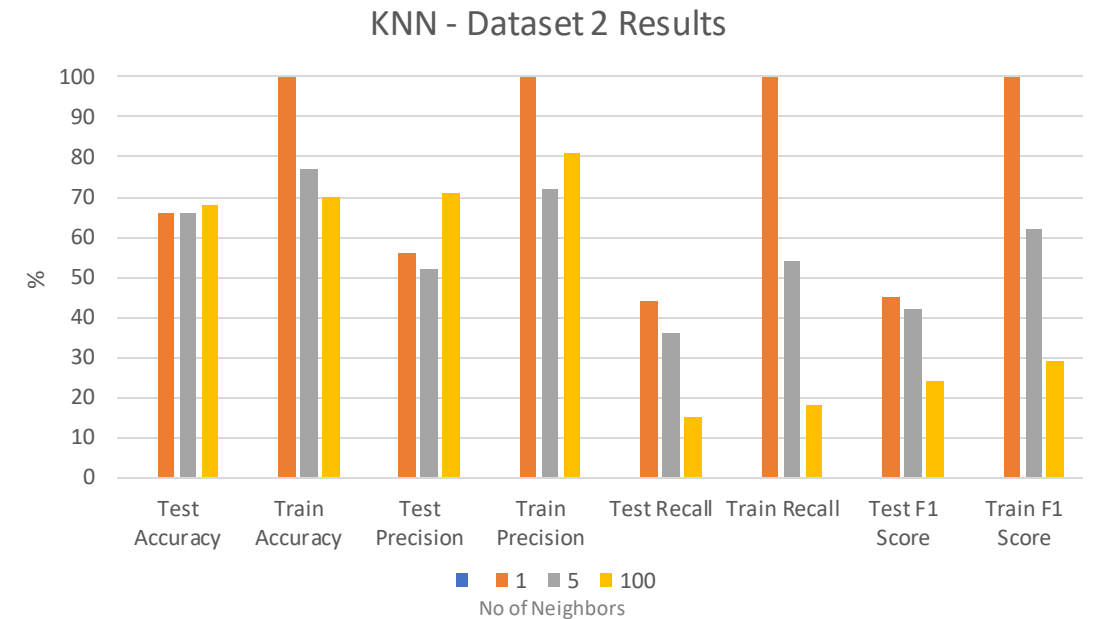
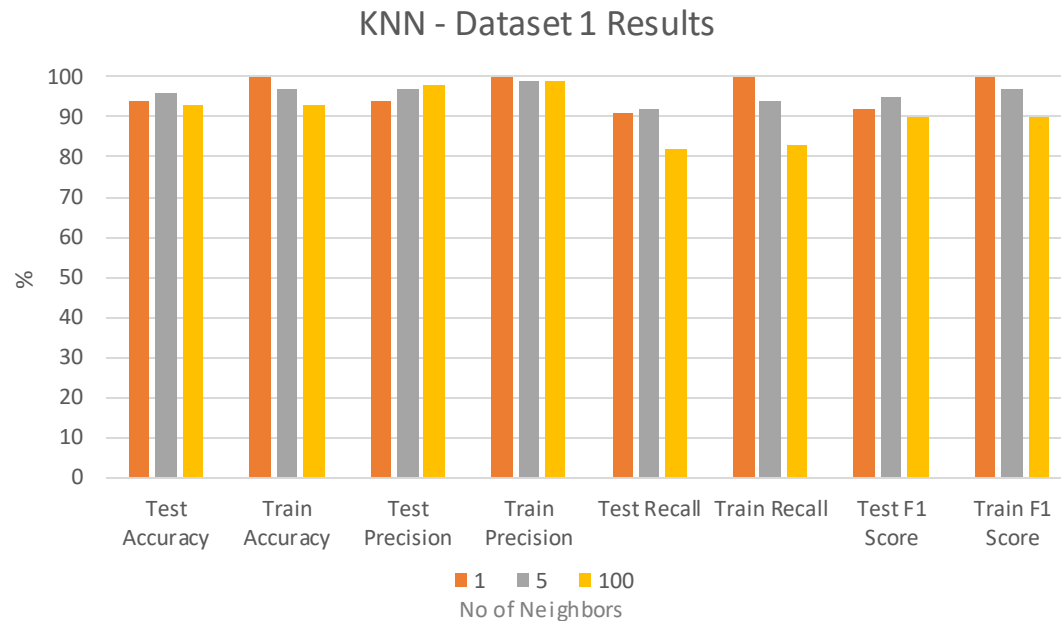


- * **Weak** regularization leads to **overfitting**
- * **Strong** regularization leads to **underfitting**



K Nearest Neighbor

- Results & Analysis

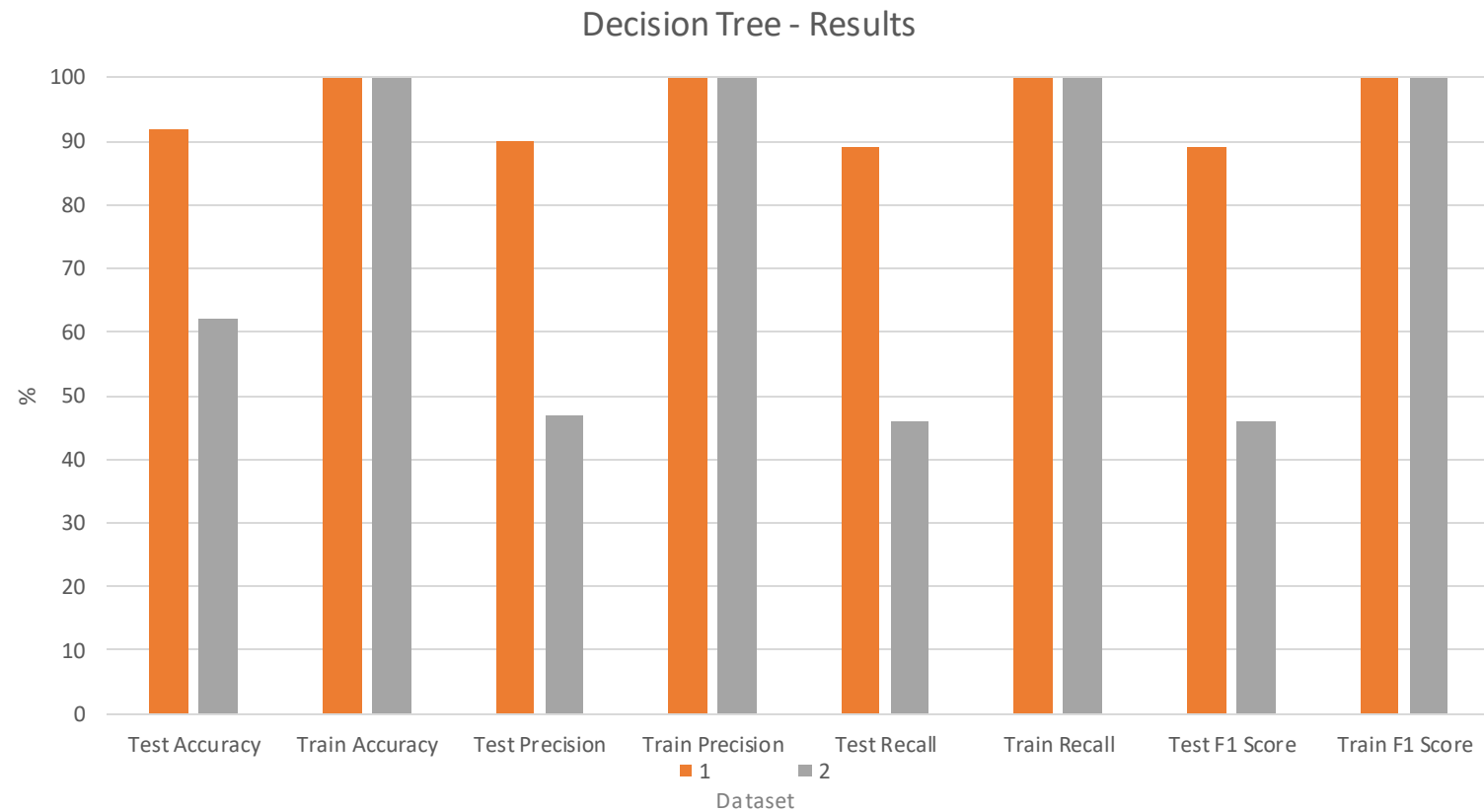


- 1 neighbor leads to **overfitting** – perfect accuracy on training
- Large number of neighbors leads to **underfitting**



Decision Tree

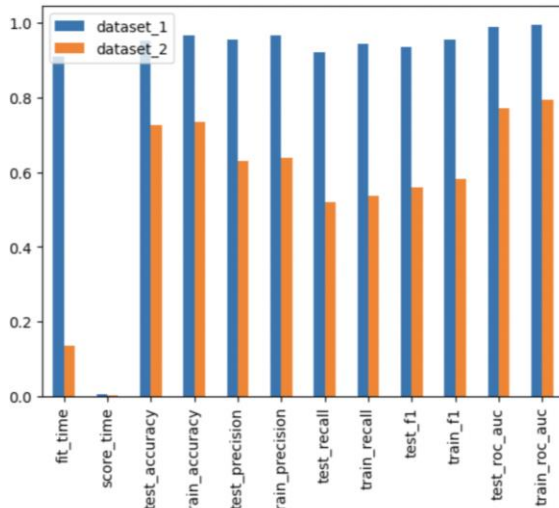
- Results & Analysis



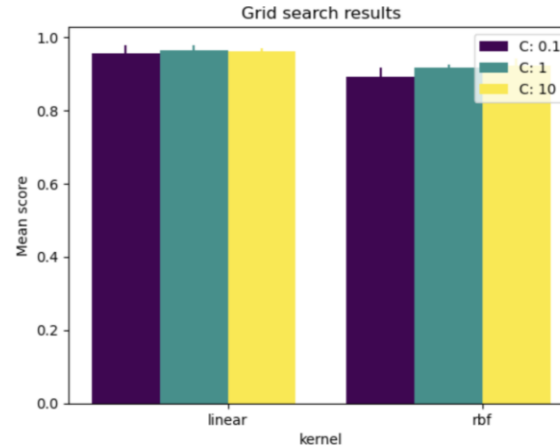


SVM – Support Vector Machines

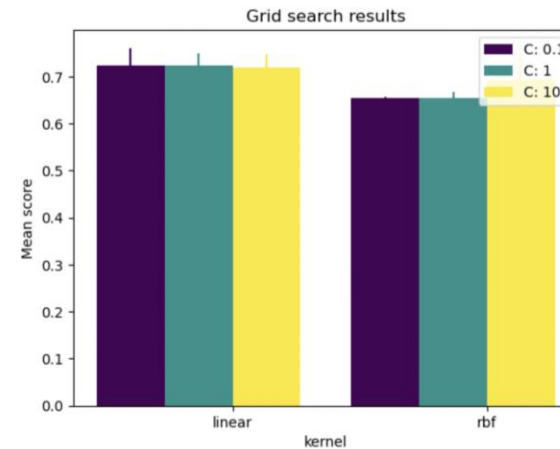
• Results & Analysis



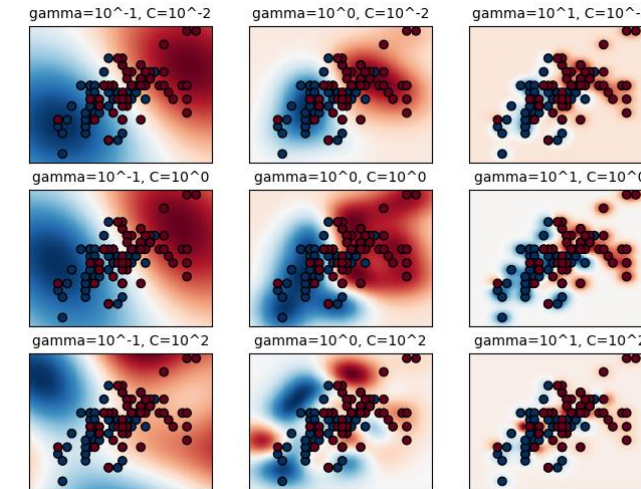
Accuracy Metrics



Dataset 1 Hyperparameter Tuning Results



Dataset 2 Hyperparameter Tuning Results



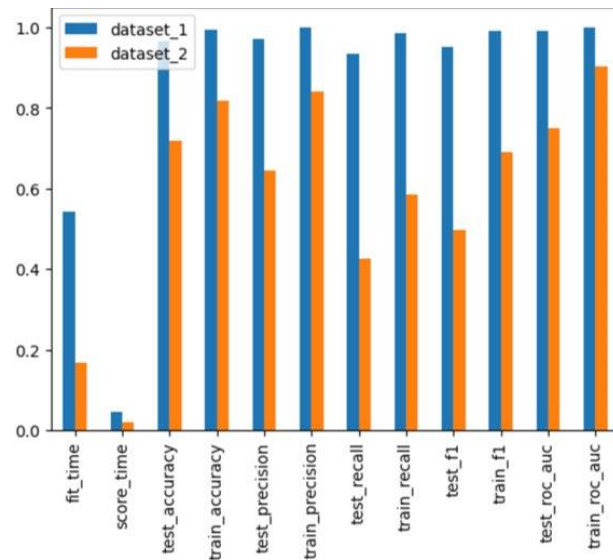
- Large value of C leads to **overfitting**
- Small value of C leads to **underfitting**

Image Credits: Scikit learn RBF SVM Parameter Tuning

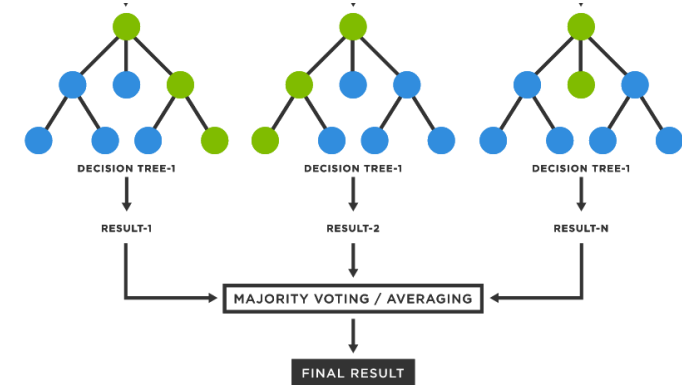
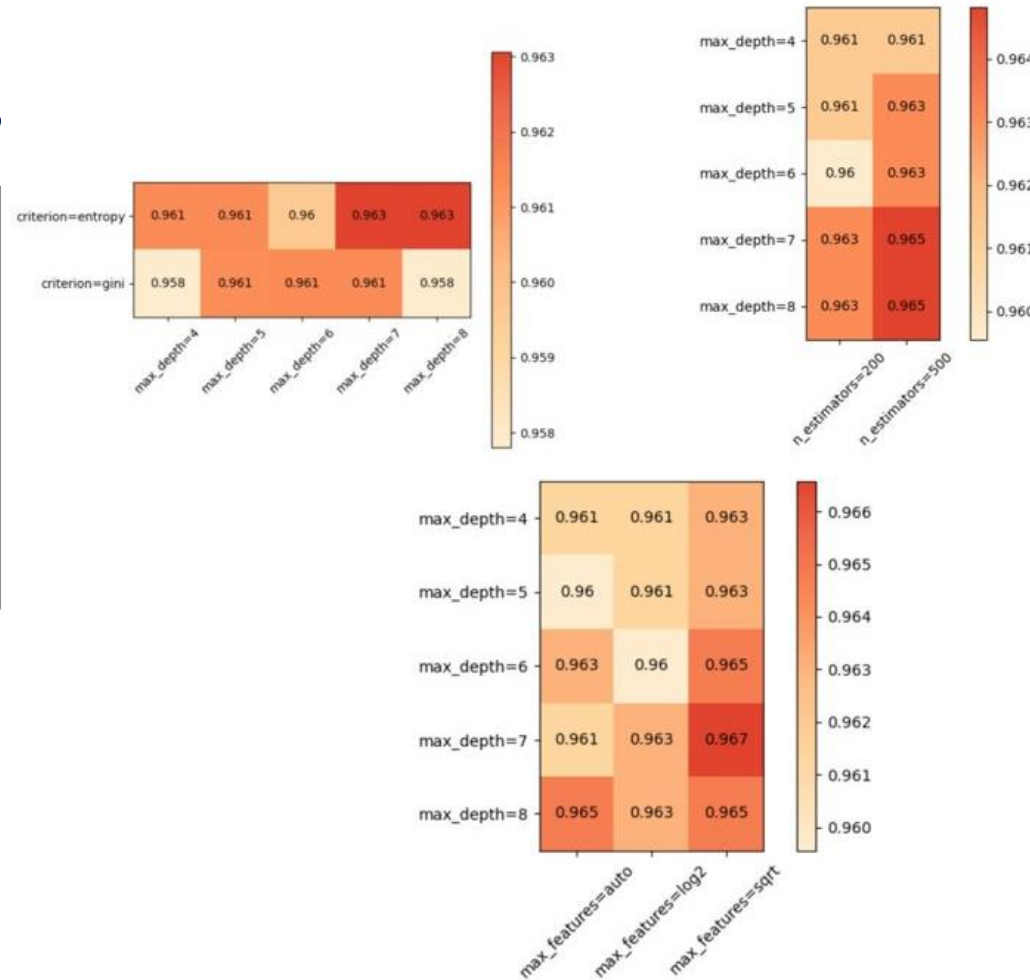


Random Forest

- Results & Analysis



Accuracy Metrics



- Large value of C leads to **overfitting**
- Small value of C leads to **underfitting**

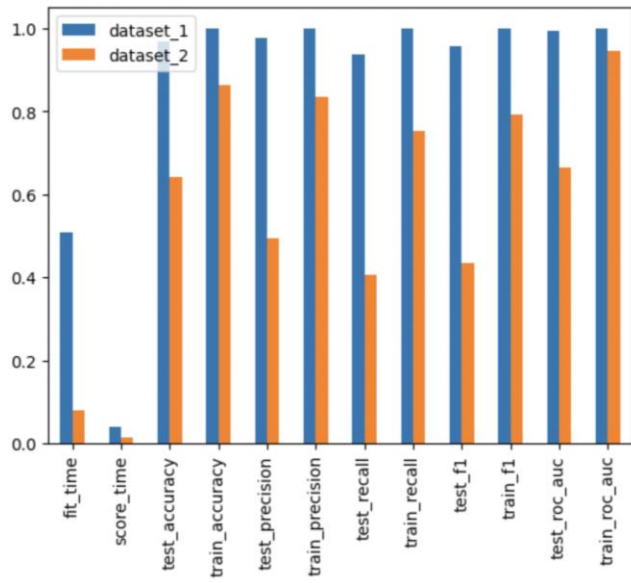
Image Credits: TIBCO Random Forest Glossary

Dataset 1 Hyperparameter Tuning Results

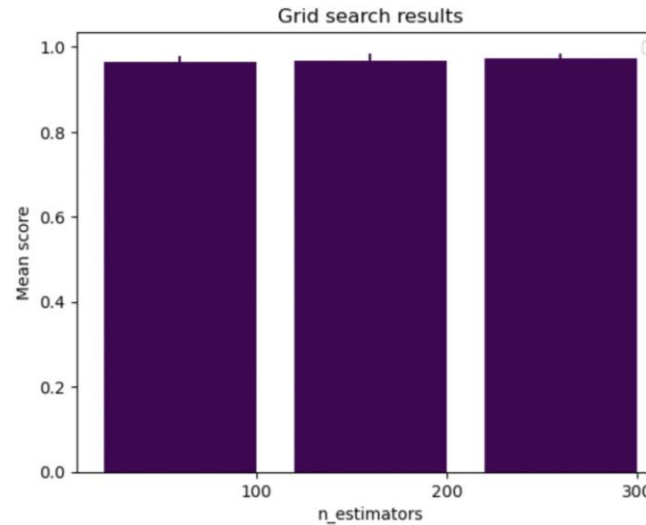


Boosting (AdaBoost)

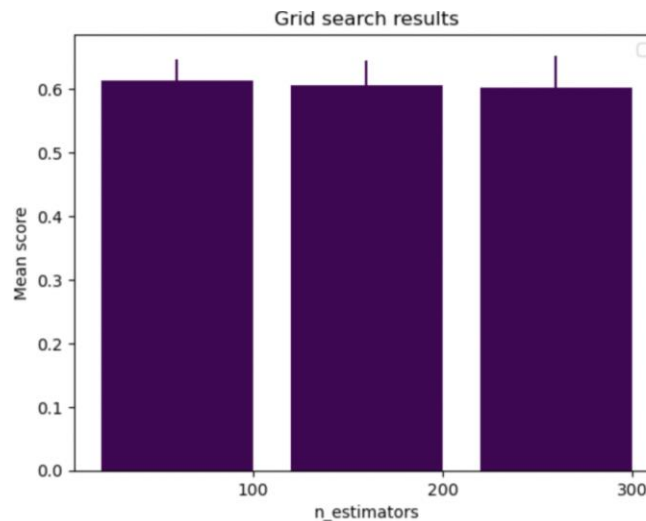
• Results & Analysis



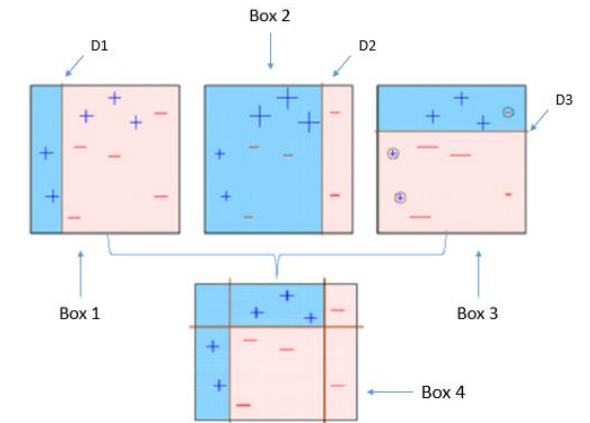
Accuracy Metrics



Dataset 1 Hyperparameter Tuning Results



Dataset 2 Hyperparameter Tuning Results



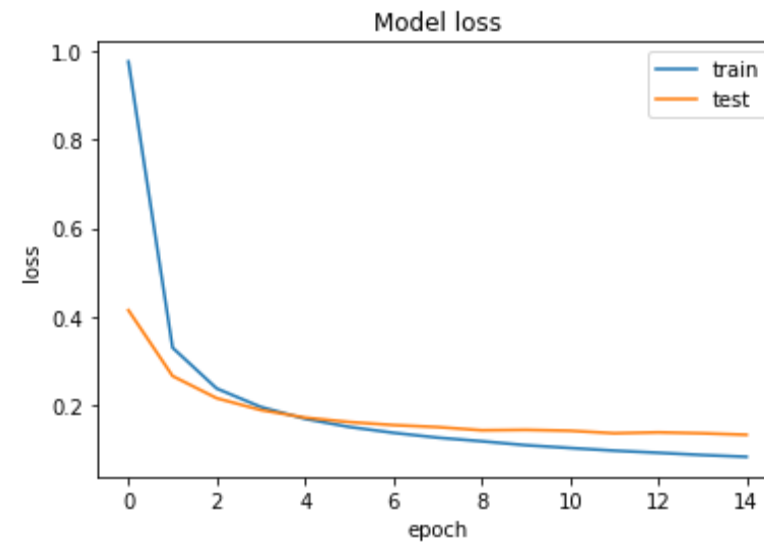
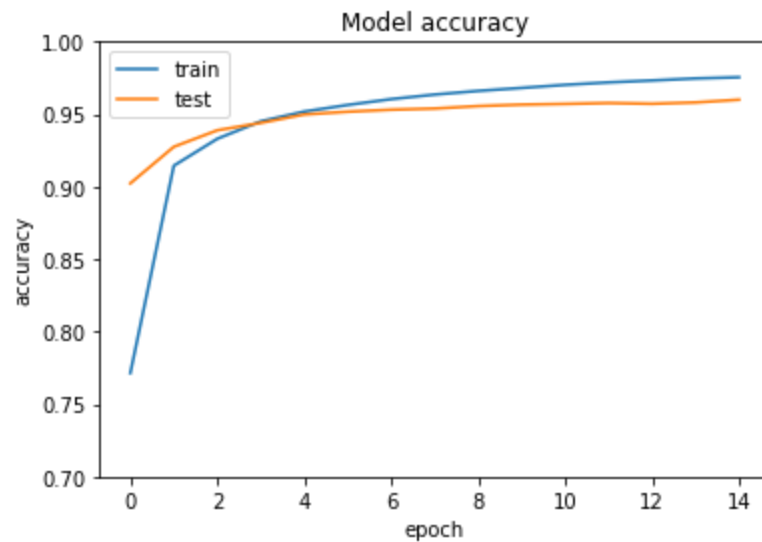
- **Large value of C leads to overfitting**
- **Small value of C leads to underfitting**

Image Credits: [Akash Desarda – Towards Data Science](#)



Neural Networks with MNIST dataset

# of hidden layers	Hidden units	Weight initialization	Bias initialization	Learning rate
2	(25, 25)	Xavier uniform	zeros	0.001



Final train accuracy: 97.5

Final test accuracy: 96.0



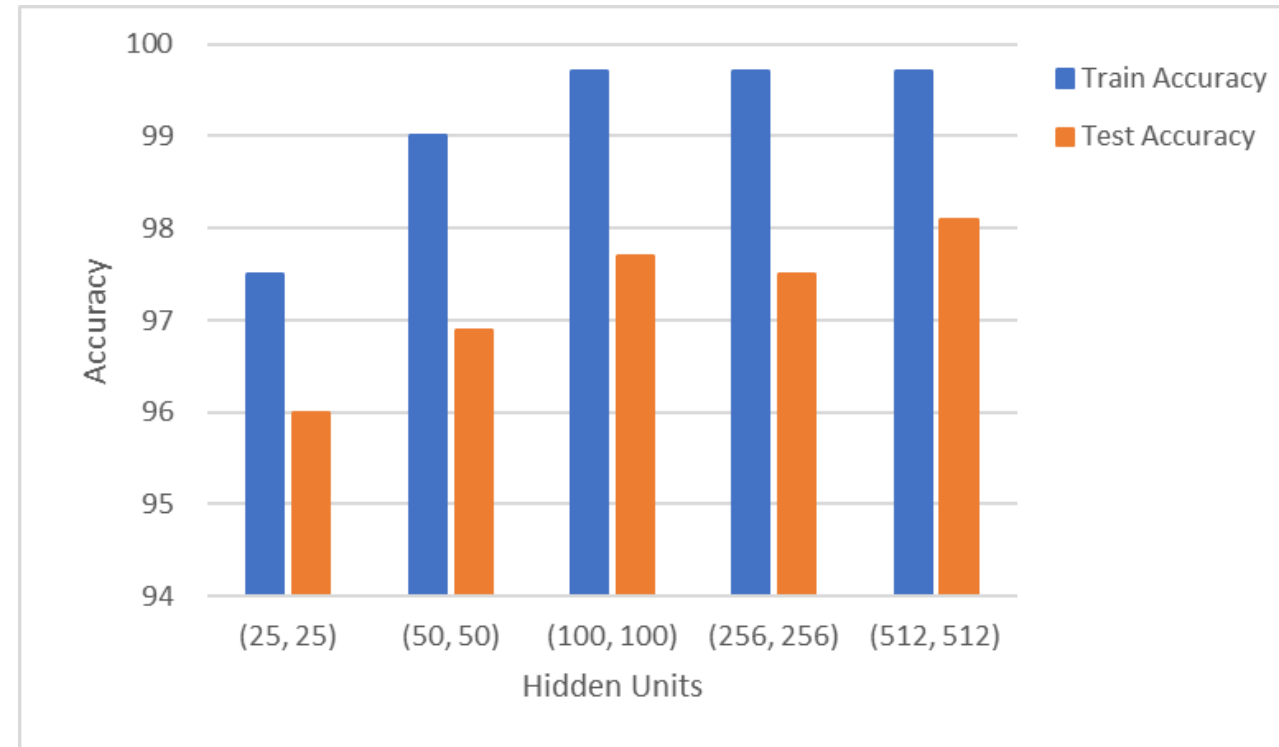
Neural Networks with MNIST dataset

- Different number of hidden units

# of hidden layers	Hidden units	Weight initialization	Bias initialization	Learning rate	Train Accuracy	Test Accuracy
2	(25, 25)	Xavier uniform	zeros	0.001	97.5	96.0
2	(50, 50)	Xavier uniform	zeros	0.001	99.0	96.9
2	(100, 100)	Xavier uniform	zeros	0.001	99.7	97.7
2	(256, 256)	Xavier uniform	zeros	0.001	99.7	97.5
2	(512, 512)	Xavier uniform	zeros	0.001	99.7	98.1



Different number of hidden units



--> As the number of hidden units increases, the accuracy of the model tends to increase.



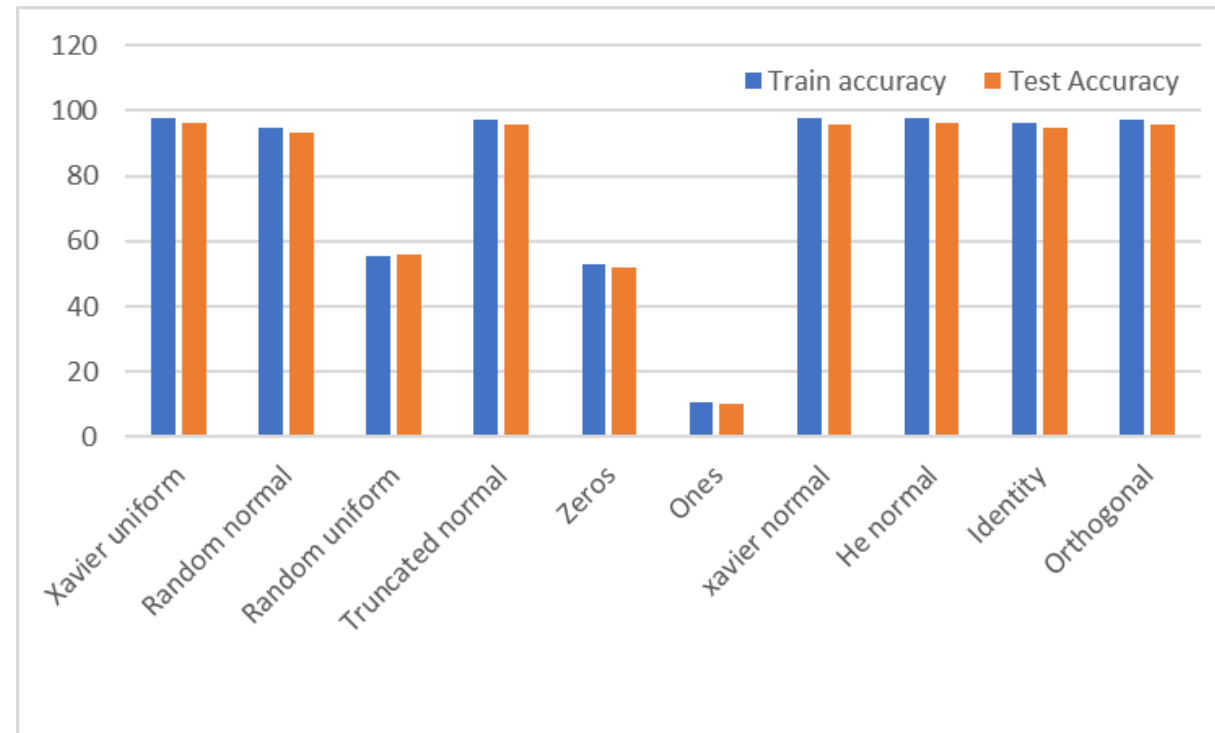
Neural Networks with MNIST dataset

- Different weight initialization

# of hidden layers	Hidden units	Weight initialization	Bias initialization	Learning rate	Train Accuracy	Test Accuracy
2	(25, 25)	Xavier uniform	zeros	0.001	97.5	96.0
2	(25, 25)	Random normal	zeros	0.001	94.6	93.4
2	(25, 25)	Random uniform	zeros	0.001	55.3	55.7
2	(25, 25)	Truncated normal	zeros	0.001	97.1	95.6
2	(25, 25)	zeros	zeros	0.001	52.9	51.9
2	(25, 25)	ones	zeros	0.001	10.8	10.1
2	(25, 25)	Xavier normal	zeros	0.001	97.6	95.8
2	(25, 25)	He normal	zeros	0.001	97.6	96.0
2	(25, 25)	Identity	zeros	0.001	96.0	94.8
2	(25, 25)	Orthogonal	zeros	0.001	97.3	95.8



Different weight initialization



--> The experiment shows that the weight initialization method has a great impact on learning the model.

Most effective: Xavier normal, Orthogonal

Least effective: Ones, Zeros



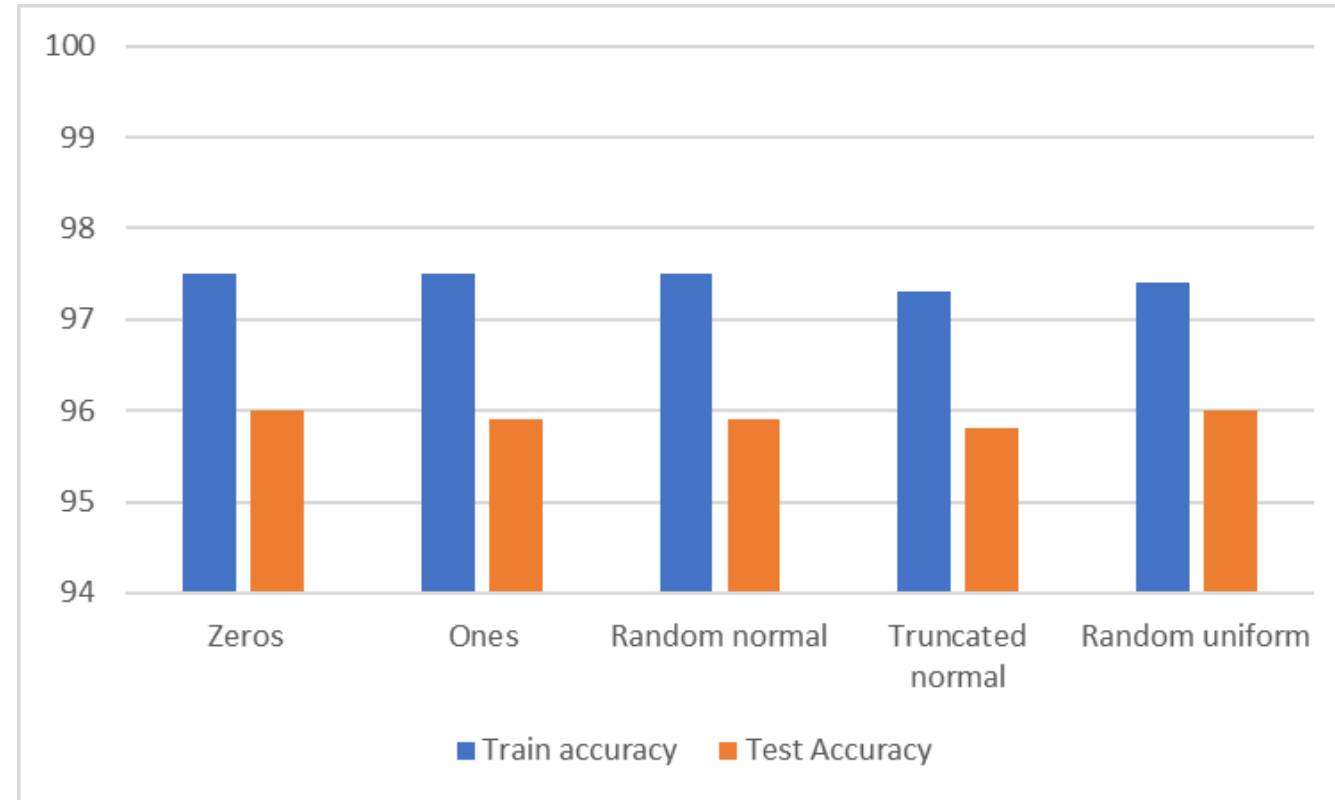
Neural Networks with MNIST dataset

- Different bias initialization

# of hidden layers	Hidden units	Weight initialization	Bias initialization	Learning rate	Train Accuracy	Test Accuracy
2	(25, 25)	Xavier uniform	zeros	0.001	97.5	96.0
2	(25, 25)	Xavier uniform	ones	0.001	97.5	95.9
2	(25, 25)	Xavier uniform	Random normal	0.001	97.5	95.9
2	(25, 25)	Xavier uniform	Truncated normal	0.001	97.3	95.8
2	(25, 25)	Xavier uniform	Random uniform	0.001	97.4	96.0



- **Different bias initialization**



--> It shows that the bias initialization method has little effect on learning accuracy compared to weight initialization



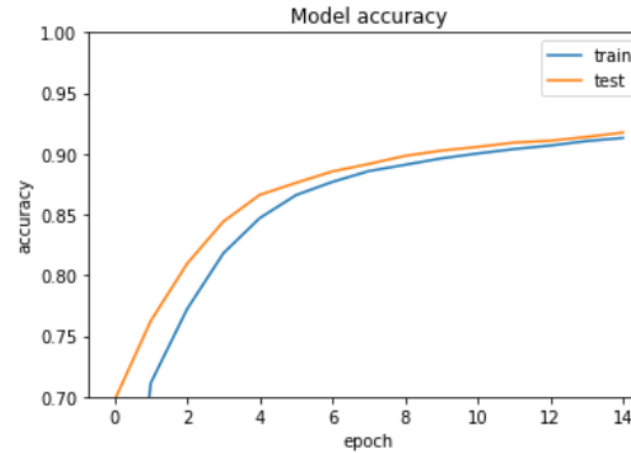
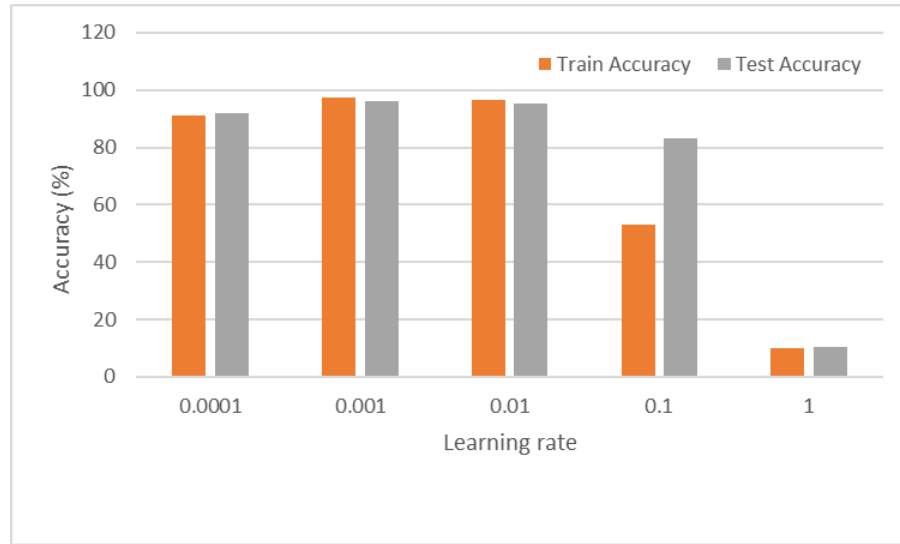
Neural Networks with MNIST dataset

- Different learning rate

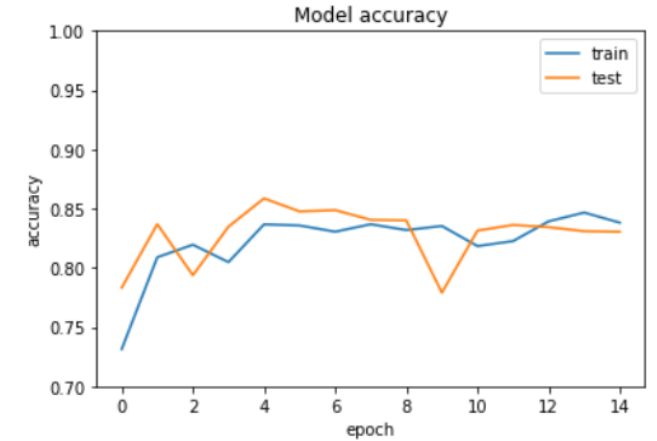
# of hidden layers	Hidden units	Weight initialization	Bias initialization	Learning rate	Train Accuracy	Test Accuracy
2	(25, 25)	Xavier uniform	zeros	0.001	97.5	96.0
2	(25, 25)	Xavier uniform	zeros	0.01	96.6	95.1
2	(25, 25)	Xavier uniform	zeros	0.1	52.9	83.0
2	(25, 25)	Xavier uniform	zeros	1	10.0	10.3
2	(25, 25)	Xavier uniform	zeros	0.0001	91.3	91.8



Different learning rate



(a) Learning rate = 0.0001



(b) Learning rate = 0.1

--> The accuracy of the learning model slowly increases compared to the default learning rate

In contrast, the model has problems with learning to converge the accuracy when the learning rate is too high.

Choosing appropriate learning rate is important.