Performing K-fold Cross Validation on Experiment- 1:

K-Value	Avg of Accuracy	Var of Accuracy	Avg of Precision	Var of Precision	Avg of Recall values	Var of Recall Values	Avg of F1 scores	Var of F1 scores
2	0.945	0.0002	0.959	2.78×10^{-6}	0.93	0.0009	0.94	0.0002
3	0.965	4.804×10^{-5}	0.979	0.0002	0.951	0.0001	0.96	4.81×10^{-5}
4	0.965	0.0004	0.98	0.0003	0.948	0.001	0.96	0.0004
5	0.97	0.001	0.99	0.0002	0.955	0.001	0.97	0.0008
6	0.98	0.0004	1.0	0.0	0.959	0.0017	0.97	0.0004
7	0.965	0.002	0.983	0.0007	0.953	0.003	0.96	0.001
8	0.97	0.0027	0.99	0.0006	0.947	0.005	0.96	0.002
9	0.985	0.0004	1.0	0.0	0.966	0.002	0.98	0.006
10	0.985	0.0005	0.983	0.002	0.98	0.0016	0.98	0.001

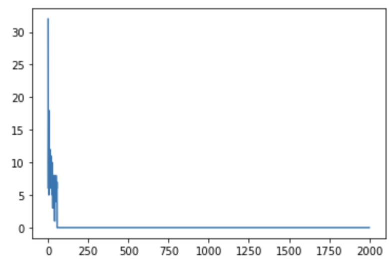
Table - 1: Containing values of the performance metrics and their variances for different values of K for 1st dataset

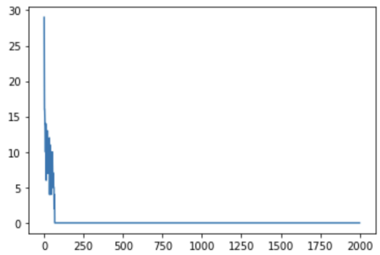
Performing k-fold cross validation on the first dataset we can observe the performance metrics table for 10 different k values and for k = 9/ 10 it gives the best result, which is expected as we increase the k the accuracy of the dataset increases but the variance also increases. So, choosing 5 would be optimal in this case as the accuracy is good and the variance is also less.

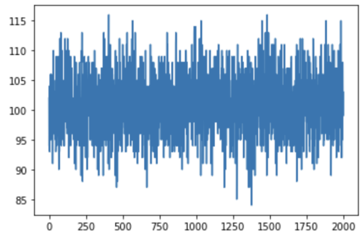
Comparing the three experiments:

Dataset. no	Accuracy	Precision	Recall	F1_score
1	0.975	1.0	0.9504	0.974
2	0.955	0.958	0.948	0.953
3	0.475	0.517	0.541	0.529

Table - 2: Training and test accuracy for experiment 1, 2 and 3







By the Performance metrics we can clearly see that the values of accuracy of 1 and 2 is higher than the 3rd dataset.

In the Perceptron Learning Algorithm we assume that the given data is linearly separable and if the assumption fails the model is not able to provide accurate output. Looking at the accuracy of the 3rd dataset we can say that it is not linearly separable whereas the first two are linearly separable. We can see in the graph, the number of misclassification v/s the iteration doesn't converge to zero in the 3rd dataset.