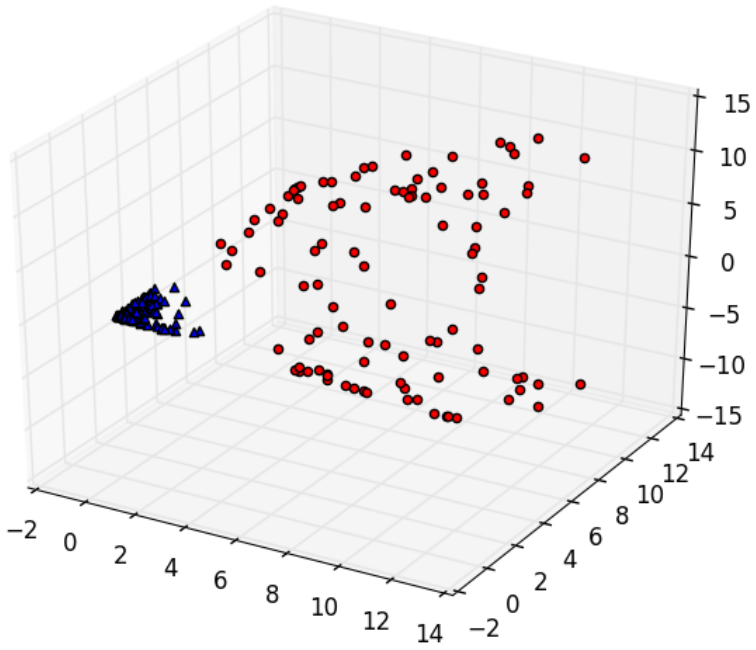
Assignment 2
Q1 Report

Kernel Trick:

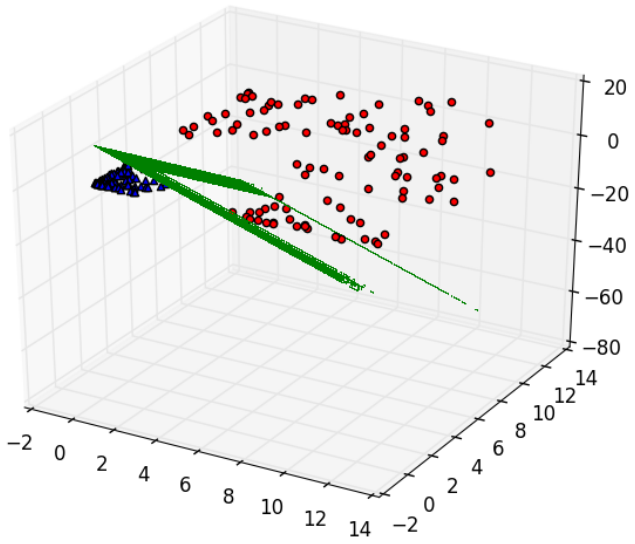
Input Data format: X=(x1,x2)

Tried the following kernels:

KernelX=(x1*x1, x2*x2, x1*x2)
--Works correctly and acheived 100% training accuracy.
Input data points:

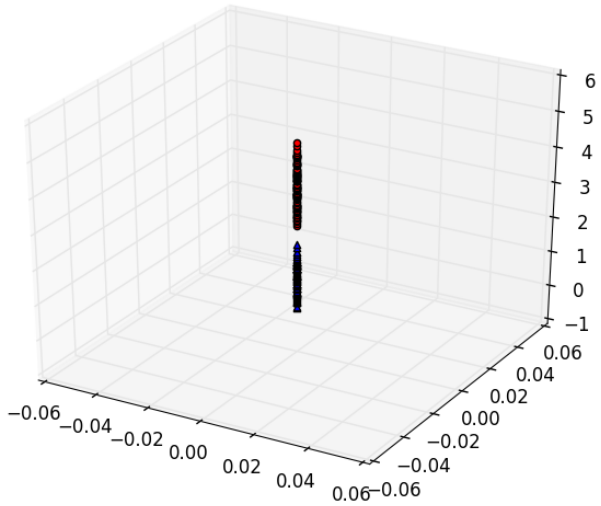


3_D decision surface:

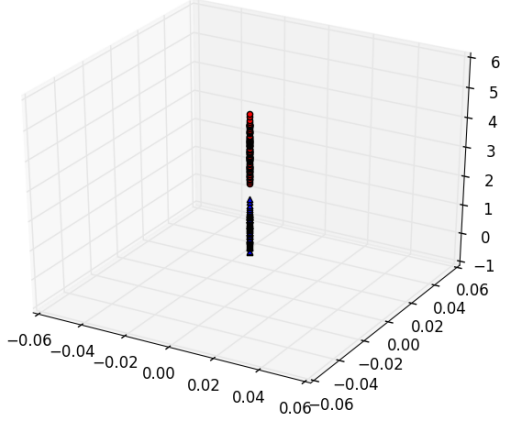


KernelX=(0, 0, sqrt(x1*x1+x2*x2))
--Works correctly and acheived 100% training accuracy.

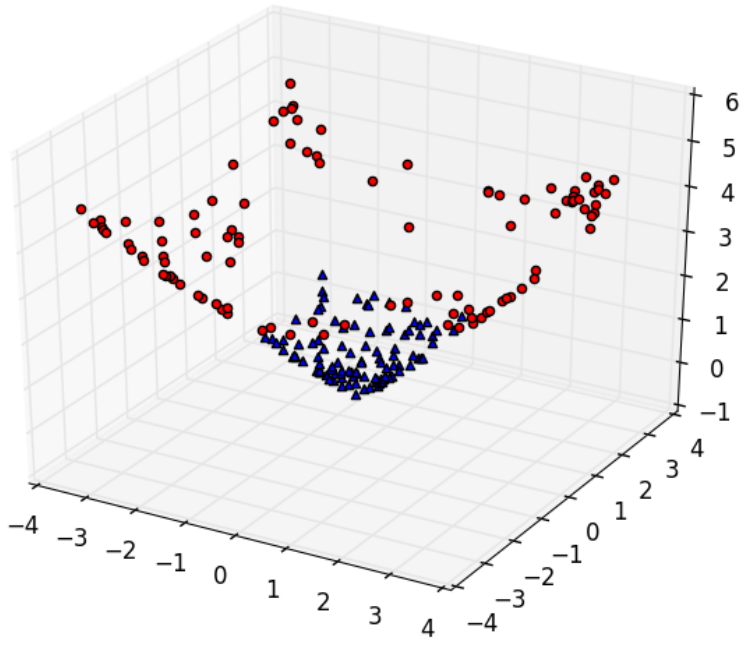
Input Data points:



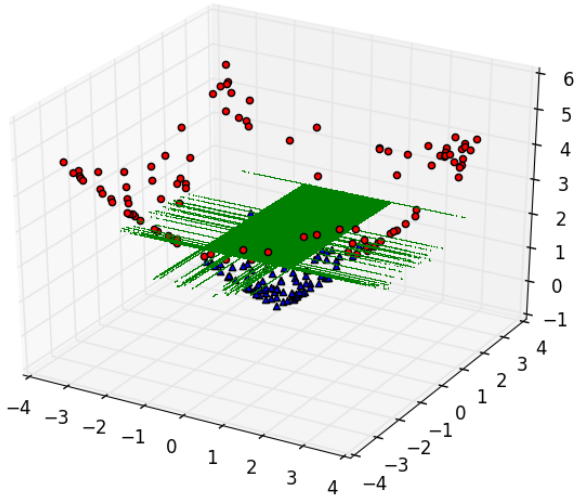
3-D surface:



KernelX=(x1, x2, sqrt(x1*x1+x2*x2))
--Works correctly and acheived 100% training accuracy.
Input data points graph:



3-D decision surface:



Note: Final kernel used: KernelX=(x1, x2, sqrt(x1*x1+x2*x2)), where InputX=(x1,x2)

Letter Classification:

Let us see some common kernels used with SVMs and experiment results on these kernels by varying the hyperparameters:

Polynomial kernel:

Equation is:

$$k(\mathbf{x_i}, \mathbf{x_j}) = (\mathbf{x_i} \cdot \mathbf{x_j} + 1)^d$$

where d is the degree of the polynomial.

Gaussian radial basis function (RBF)

It is general-purpose kernel; used when there is no prior knowledge about the data.

Equation is:

$$k(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma \|\mathbf{x_i} - \mathbf{x_j}\|^2)$$

for:

$$\gamma > 0$$

Sometimes parametrized using:

$$\gamma = 1/2\sigma^2$$

Sigmoid kernel

We can use it as proxy for neural networks. Equation is

$$k(x, y) = \tanh(\alpha x^T y + c)$$

Results of the model trained on data with 14000 training samples and tested on 2000 samples:

Kernel	C	Gamma	Accuracy	Precision	Recall	F1_score
RBF	10	0.1001	0.9721	0.9724209	0.9721	0.9720932
RBF	5	0.1001	0.9708	0.9712174	0.9708	0.9707886
RBF	0.9	0.2	0.9465	0.9488150	0.9465	0.944766
Poly	10	0.1001	0.9429	0.943796	0.9429	0.942819
Poly	5	0.1001	0.9431	0.9437564	0.9431	0.942993
Poly	0.9	0.2	0.9461	0.9467209	0.9461	0.946059

Linear	10	0.15	0.8537	0.855860	0.8537	0.853454
Linear	0.1	0.1	0.851	0.853145	0.851	0.850669

Observations:
--RBF kernel outperforms every other kernel.
--RBF kernel takes the maximum time for fitting and predicting the model.

The two discriminative features are :

1)xedge
2)yedge