Foundations of Machine Learning Assignment -2

Qillis SVM:

doubeno bromobell 21 1. In the expression Wixtb, wand bare unknowns on some or or of the

For Ex. (10x+100 = 1000 ,)(+1) = 1

both are same lines. Since Wand bare unknowns fixed right side to +1 & - I, rescale wakw, b->kb according to that

2. Making +1&-1 just for our convenience. It malcies calculations egsier. Lamon de

In this way, if margin boundaries replaced by any arbitrary constants, solution is unchanged.

Kernels.

 $k(x,2) = k_1(x,2) + k_2(x,2)$

=> k(x,2) = xk,(x,2) + Bk2(x,2) . for x, B≥0

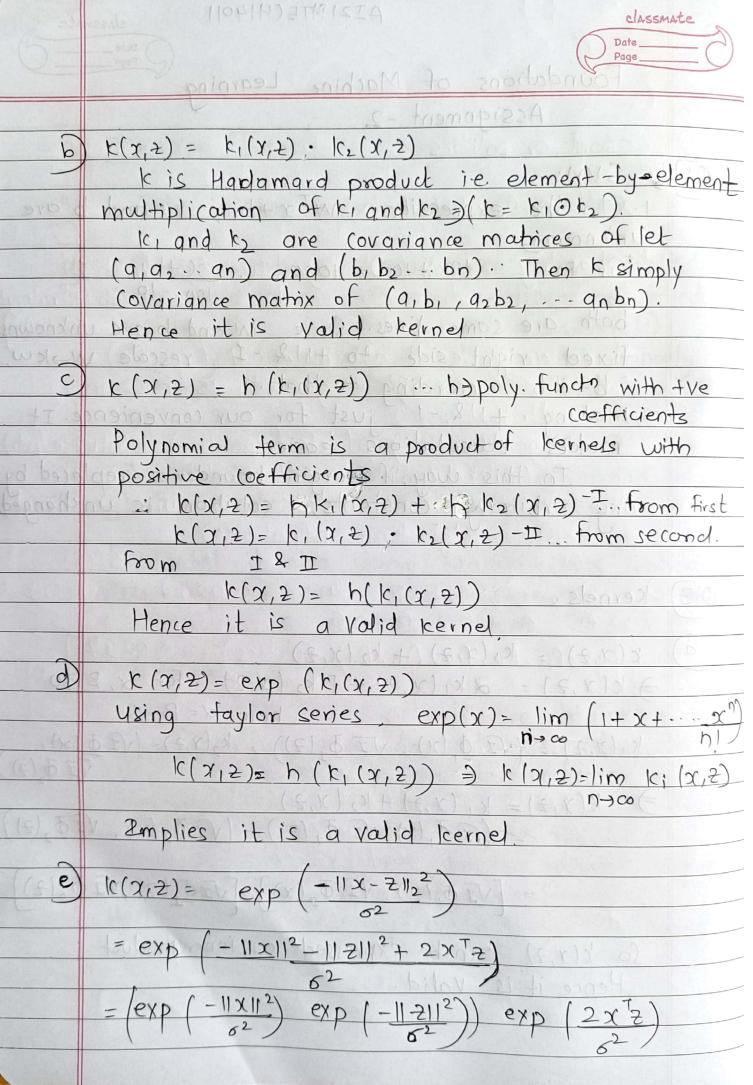
 $k_1(x_1, z) = (\sqrt{2} \Phi_1(x)), \sqrt{2} \Phi_1(z)), k_2(x_1, z) = (\sqrt{2} \Phi_1(x)), \sqrt{2} \Phi_2(z))$

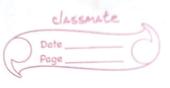
 $\exists |c(x, z) = |c_1(x, z) + |c_2(x, z)|$

 $= (\mathcal{I} \overline{\Phi}, (x), \sqrt{I} \overline{\Phi}, (z)) + (\mathcal{I} \overline{\Phi}_{2}(x), \sqrt{I} \overline{\Phi}_{3}(z))$

= $\left[\sqrt{J}\Phi_{1}(x), \sqrt{J}\Phi_{2}(x)\right], \left[\sqrt{J}\Phi_{1}(z), \sqrt{J}\Phi_{2}(z)\right]$

So, k(x, 2) can be expressed as inner product. Hence it is valid.





	= fbe) f(z) exp(k,(x,2))
	f(x) f(t) is one boine ovalu (x)) is one
	f(x) f(t) is one kernel exp(k,(x,t)) is one kernel and hence product of two kernels
	is a valid kernel.
	12 y vary perner.
0.2	Q 1 C.T 1 Sali
9	$S = \frac{1}{11011}$, $S \cdot T$, $1 = \sum x_i$
	2 02
	111112
	1 - 11.111 ² T
	$\frac{1}{ \omega ^2}$ $\frac{1}{ g ^2} = \omega ^2 - I$
	From SVM hard-margin duality, $\bar{w} = \frac{2}{3}\bar{\chi}_i \underline{y}_i \underline{x}_i - \underline{I}$ Here
	Here i=1
	Lagrangian, L(w,x)= 11w112+ = xi(1-y; xiTw).
	$\frac{2qq^{1}q^{1}q^{1}}{1=1}, \frac{1}{1}\omega_{1}^{2} + \frac{1}{2}\omega_{1}^{2} + \frac{1}{2}\omega_{1}^{2}$
1000	Primal maximum margin was
	P(w) = cup. L(w,x)= cup. [w ^2 + \frac{2}{3} x: (1-4) x. [w]
	Primal maximum margin was. P(w) = sup. L(w,x)= sup. [w _2 + \frac{1}{2} \pi; (1-y_1x_1, w)] x>0
	Durd is given by
	$D(\alpha) = \min_{w \in \mathbb{R}^d} L(w, \alpha) = \sum_{i=1}^d \alpha_i - \sum_{i=1}^d \alpha_i y_i x_i ^2$
	WERD i=1 /i=1 /2
	For max. margin SVM, D(x)=0
	For max. margin SVM, $D(x) = 0$ $\vdots \exists x \in \mathbb{Z}$ $\exists x \in \mathbb{Z}$ $\exists x \in \mathbb{Z}$ $\exists x \in \mathbb{Z}$ $\exists x \in \mathbb{Z}$
	i=1 112
) ZX = W 2 From I
	III THE REAL PROPERTY OF THE P
	$\exists \alpha := 1 - Foom I$
	Hence, 1 - Edi thus proved.
	g^2 $i=1$

Programming Questions

Q.4)

Ans. In this question, initially training and testing data are loaded, then they are filtered according to one-versus-one classifier for the digits '1' (+1) and '5' (-1). Values are stored as features and labels. Data is trained and tested using different kernel implementation techniques.

- a) Model is train and tested using Linear Kernel method. Overall accuracy calculated over entire test set is 0.97877 or 97.877%. Total no. of support vectors calculated are 28.
- b) Model is train and tested using Linear Kernel method for first given points. The given number of samples, their respective test accuracy and their respective support vectors are given below.

Samples	Test Accuracy	Total Support Vectors
50	0.98	2
100	0.99	4
200	0.99	8
800	0.981	14

c) Model is train and tested using Polynomial Kernel method.

True or False.

i. False.

At C=0.0001, For Q=2, training error is 0.02242 and for Q=5, it is 0.00641. Hence, this statement is False.

ii. True.

At C=0.001, For Q=2, total number of support vectors are 152 and for Q=5, they are 28. Hence, this statement is True.

iii. False.

At C=0.01, For Q=2, training error is 0.00448 and for Q=5, it is also 0.00448. Hence, this statement is False.

iv. True.

At C=1, For Q=2, test error is 0.01887 and for Q=5, it is 0.01651. Hence, this statement is True.

d) Model is train and tested using RBF Kernel method.

Training and Testing error values for given values of C are listed below.

C-Value	Training Error	Testing Error
0.01	0.00384	0.02123
1	0.00448	0.02123
100	0.0032	0.01887
10000	0.00256	0.01887
1000000	0.00128	0.02123

Lowest training error results for C=10⁶, which is 0.00128. Lowest testing error results for C=100 and C=10000, which is 0.01887.

Ans. In this question, initially training and validation data are loaded. Values are stored as features and labels, converted into arrays. Data is trained and tested using different kernel implementation techniques.

- a) Model is train and tested using Linear Kernel method. Training error calculated as 0, Testing error as 0.024 or 2.4%. Total number of support vectors calculated are 1084.
- b) In the first part, Model is train and tested using RBF Kernel method. Training error calculated as 0, Testing error as 0.5 or 50%. Total number of support vectors calculated are 6000.

In the second part, Model is train and tested using Polynomial Kernel method. Training error calculated as 0, Testing error as 0.021 or 2.1%. Total number of support vectors calculated are 1755.

All Kernels have lowest training error i.e., 0.