MALIS
Group Exercise

November 29 2022

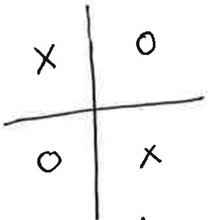
Group Name:	
Group Members:	

1 Bias Variance Decomposition

[3 points] Suppose you collected a sufficiently large dataset generated by a polynomial of degree 4. Characterize the bias-variance of the estimates of the following models on the data with respect to the true model by choosing the appropriate entry.

Model	Bias	Variance
Linear regression	low/high	low/high
Polynomial regression with degree 4	low/high	low/high
Polynomial regression with degree 10	low/high	low/high

2 Support Vector Machines and Kernels



Plot 1

$\phi(x)$	Y
1 1 1 1	1
1 1 -1 -1	-1
1 -1 1 -1	-1
1 -1 -1 1	1

Consider a training set consisting of points in the 2D space $x = \{(1,1), (1,-1), (-1,1), (-1,-1)\}$ with labels $y = \{1, -1, -1, 1\}$.

(a) [½ point] Is the dataset linearly separable in the original space? Justify your answer.

No. See plot 1

(b) [1½ points] Consider the following feature transformation: $\phi(\mathbf{x}) = [1, x_1, x_2, x_1 x_2]^T$, where x_1 and x_2 are the first and second coordinates of x. Your prediction function in this feature space is $\hat{y}(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x})$. Give the coefficients, $\mathbf{w}$, of a maximum-margin decision surface separating the positive from the negative examples.
Hint: You should be able to do this by inspection, without the need of any significant computation.

$W = [0\ 0\ 0\ 1]^T$

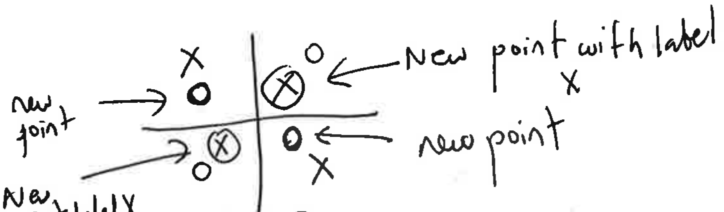
(c) [1½ points] Plot the training set in the original space. Add one training sample to the plot so that the five samples can no longer be linearly separated in the feature space $\phi(\mathbf{x})$ using the coefficients $\mathbf{w}$ you estimated in the previous question.

See plot 2

(d) [1½ points] What kernel $K(\mathbf{u}, \mathbf{v})$ does this feature transformation ϕ correspond to?

$K(u,v) = 1 + u_1 v_1 + u_2 v_2 + u_1 u_2 v_1 v_2$

Note: These 2 questions were part of the final exam in 2019.



Plot 2