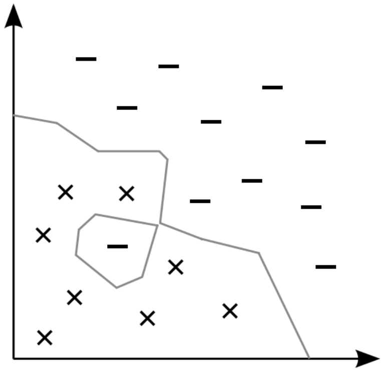
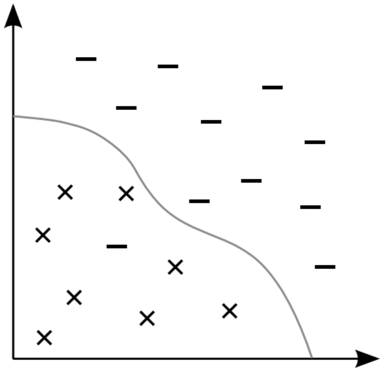
Q1-1: Figure shows the decision boundaries (DB) for two nearest-neighbour classifiers - 1NN and 3NN. Consider following statements and choose the correct option (True/False for all the statements A/B/C/D).

A. *DB1 belongs to 3NN while DB2 belongs to 1NN.*
B. *DB2 belongs to 3NN while DB1 belongs to 1NN.*
C. *DB1 gives zero test error.*
D. *DB1 gives zero training error.*



Decision Boundary 1

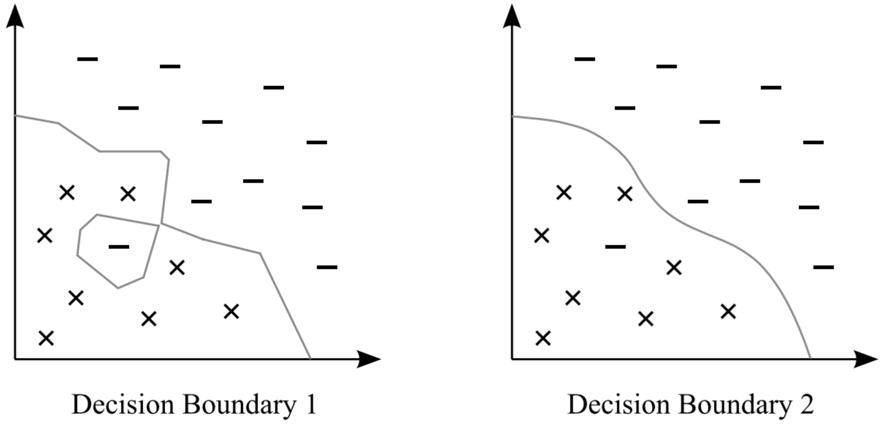


Decision Boundary 2

1. A: True, B: False, C: True, D: False
2. A: False, B: True, C: False, D: True
3. A: False, B: True, C: True, D: False
4. A: True, B: False, C: False, D: True

Q1-1: Figure shows the decision boundaries (DB) for two nearest-neighbour classifiers - 1NN and 3NN. Consider following statements and choose the correct option (True/False for all the statements A/B/C/D).

A. *DB1 belongs to 3NN while DB2 belongs to 1NN.*
B. *DB1 belongs to 1NN while DB2 belongs to 3NN.*
C. *DB1 gives zero test error.*
D. *DB1 gives zero training error.*



1. A: True, B: False, C: True, D: False
2. A: False, B: True, C: False, D: True
3. A: False, B: True, C: True, D: False
4. A: True, B: False, C: False, D: True

Q1-2: Table shows all the training points in 2D space and their labels. Assume 3NN classifier and euclidean distance. What should be the labels of the points A: (1, 1) and B(2, 1)?

1. A: +, B: -
2. A: -, B: +
3. A: -, B: -
4. A: +, B: +

x	y	label
0	0	+
1	0	+
2	0	+
2	2	+
0	1	-
0	2	-
1	2	-
3	1	-

Q1-2: Table shows all the training points in 2D space and their labels. Assume 3NN classifier and euclidean distance. What should be the labels of the points A: (1, 1) and B(2, 1)?

1. A: +, B: -
2. A: -, B: +
3. A: -, B: -
4. A: +, B: +

x	**y**	**label**
0	0	+
1	0	+
2	0	+
2	2	+
0	1	-
0	2	-
1	2	-
3	1	-

3 nearest neighbors to point A are (0, 1) [-], (1, 0) [+], (1, 2) [-]. Hence, the label should be **[-]**

3 nearest neighbors to point B are (2, 0) [+], (2, 2) [+], (3, 1) [-]. Hence, the label should be **[+]**

Q2-1: In a distance-weighted nearest neighbor, which of the following weight is **NOT** appropriate? Let p be the test data point and x_i {i = 1: N} be training data points.

1. $w_i = d(p, x_i)^{½}$
2. $w_i = d(p, x_i)^{-2}$
3. $w_i = \exp(-d(p, x_i))$
4. $w_i = 1$

Q2-1: In a distance-weighted nearest neighbor, which of the following weight is **NOT** appropriate? Let p be the test data point and x_i {i = 1: N} be training data points.

1. $w_i = d(p, x_i)^{½}$
2. $w_i = d(p, x_i)^{-2}$
3. $w_i = \exp(-d(p, x_i))$
4. $w_i = 1$

The intuition behind weighted kNN, is to give more weight to the points which are nearby and less weight to the points which are farther away. Any function whose value decreases as the distance increases can be used as a function for the weighted knn classifier. w = 1 is also apt as it reduces to our traditional nearest-neighbor algorithm.

Q2-2: Which of the following statement is true?

1. kNN is a “lazy” learning algorithm as it does virtually nothing at testing time.
2. kNN is a non-parametric technique.
3. In incremental growth technique, we start with all training instances and remove non-useful training instances.
4. Standardizing numeric features always helps in getting better accuracy.

Q2-2: Which of the following statement is true?

1. kNN is a “lazy” learning algorithm as it does virtually nothing at testing time.
2. kNN is a non-parametric technique.
3. In incremental growth technique, we start with all training instances and remove non-useful training instances.
4. Standardizing numeric features always helps in getting better accuracy.

    1. kNN is a “lazy” learning algorithm as it does virtually nothing at **training** time.
    2. kNN is a non-parametric technique. This is correct.
    3. In incremental **deletion** technique, we start with all training instances and remove non-useful training instances.
    4. Standardizing numeric features **DOESN’T ALWAYS** helps in getting better accuracy. Consider the case where a feature important for the label gets a much smaller range after standarization.

Q3-1: Select the correct option.

A. *Instance based learning is sensitive to range of feature values.*

B. *Training is very efficient.*

C. *Occam's razor is an example of hypothesis space bias.*

1. Statement A is true. Statement B, C are false.
2. Statement A, B are true. Statement C is false.
3. Statement B, C are true. Statement A is false.
4. All Statements are true.

Q3-1: Select the correct option.

A. *Instance based learning is sensitive to range of feature values.*

B. *Training is very efficient.*

C. *Occam's razor is an example of hypothesis space bias.*

1. Statement A is true. Statement B, C are false.
2. Statement A, B are true. Statement C is false.
3. Statement B, C are true. Statement A is false.
4. All Statements are true

Occam's razor is an example of **preference bias,** i.e – Prefer one hypothesis over another even though they have similar training accuracy. For example, we prefer smaller trees in the hypothesis space of decision trees

Q3-2: Select the correct option.

A. *K-d trees can reduce the classification time.*
B. *Without an Inductive Bias we have no rationale to choose one hypothesis over another.*
C. *In a KNN using Euclidean distance, the distance between neighbors may be dominated by the large number of irrelevant attributes. And hence, it might fail.*

1. Statement A, C are true. Statement B is false.
2. Statement A, B are true. Statement C is false.
3. Statement B, C are true. Statement A is false.
4. All Statements are true.

Q3-2: Select the correct option.

A. *K-d trees can reduce the classification time.*

B. *Without an Inductive Bias we have no rationale to choose one hypothesis over another.*

C. *In a KNN using Euclidean distance, the distance between neighbors may be dominated by the large number of irrelevant attributes. And hence, it might fail.*

1. K-d tree can speed up the computation of nearest neighbors.
2. Without inductive bias means we don't know any property of the data
3. If we have a lot of irrelevant attributes/features, their impact on the distance can overwhelm that of the relevant features, and then lead to poor predictions.

1. Statement A, C a
2. Statement A, B are true. Statement C is false.
3. Statement B, C are true. Statement A is false.
4. All Statements are true.