Exercises

1	2	3	4	5	6	7	8	9

Surname, First name

1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0

Fundamentals of Artificial Intelligence Programme
Practice exam

Multiple choice

1p **1a** A number of clusters has to be specified to perform hierarchical clustering.

(a) False (b) True

1p **1b** The larger K is in K-means clustering, the better the model typically fits the training data.

(a) False (b) True

1p **1c** Chose the most efficient operation to select data from a pandas dataframe in the data engineering class:

(a) df.query("col1 > @value")
(b) df[df.col1 > value]
(c) df.loc[df.col1 > value]
(d) df[df['col1'] > value]

2p **1d** Which of the following statements are true?

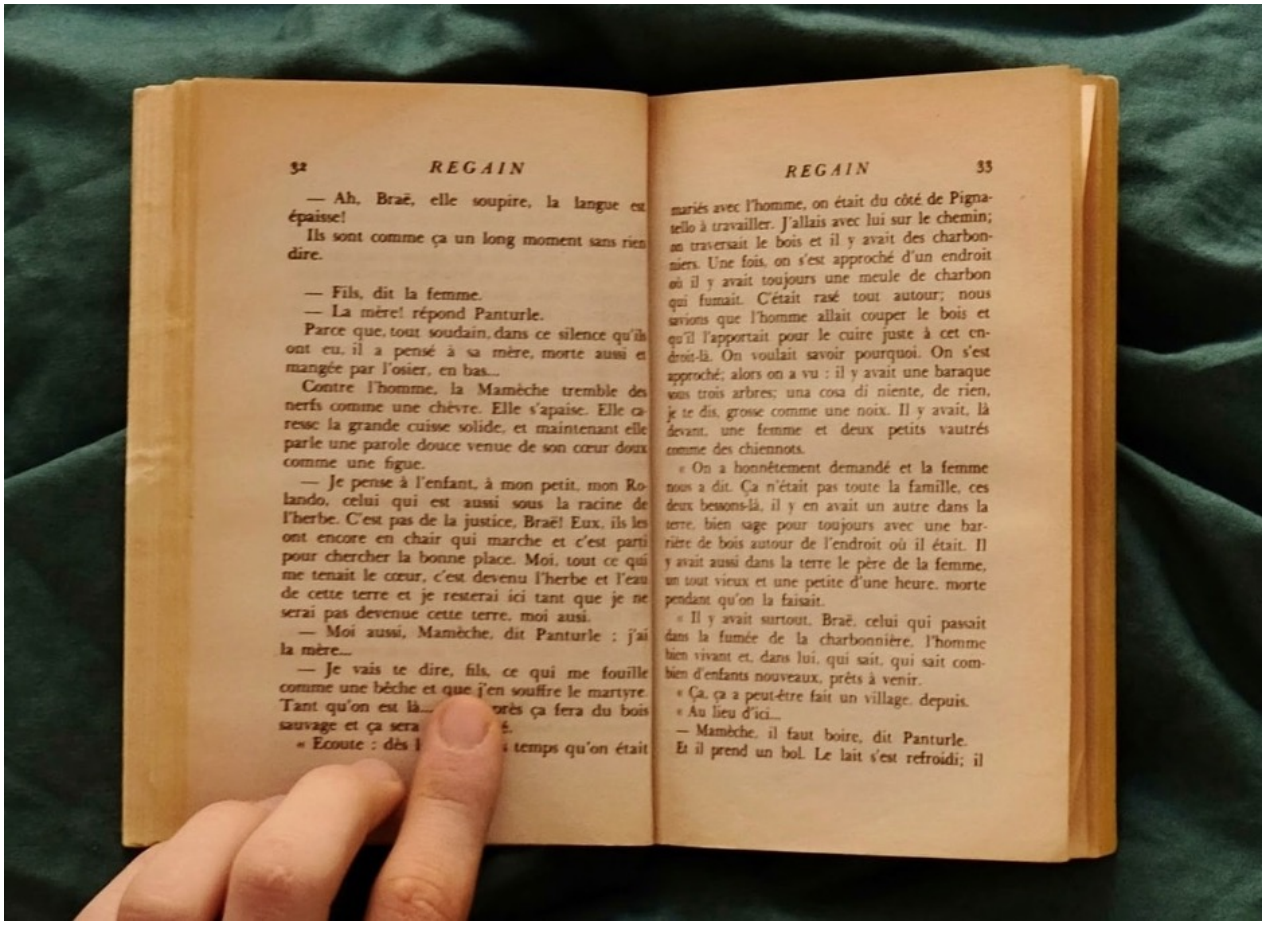
(a) In a conceptual data schema, the main usage restrictions and concepts are defined (like legal constraints, or descriptions on who can access the data, etc.)
(b) In a logical data schema describes the results of analyzing a dataset, e.g., the logical conclusions one can draw from it
(c) Data schemas define (among other thing) relevant entity types
(d) Different logical data schemas can be derived from the same conceptual data schema

Case study 1

2p **2a Describe two main issues in crowdsourced training data creation and techniques to address those issues. Answer briefly.**

1. **Quality Control**: Crowdsourced data often suffers from inconsistent or incorrect labeling due to varying skill levels and potential biases of contributors.
   - **Solution**: Implement quality control techniques such as redundancy (using multiple annotators per task), consensus-based voting, and using gold-standard questions to identify and filter unreliable annotators.
2. **Bias in Data**: The collected data may have inherent biases, such as demographic or cultural biases, that can lead to biased models.
   - **Solution**: Use diverse and representative crowdsourcing pools to minimize bias, and apply post-collection bias analysis and rebalancing techniques to ensure data is more representative.

3p **2b You want to build a ML model to help fully digitize books. Consider the following task design and description presented to crowd workers on Amazon Mechanical Turk. Identify 3 potential quality-related issues with this task design. How can they be improved?**



52 REGAIN

— Ah, Braë, elle soupire, la langue est épaisse!

Ils sont comme ça un long moment sans rien dire.

— Fils, dit la femme.

— La mère! répond Panturle.

Parce que, tout soudain, dans ce silence qu'ils ont eu, il a pensé à sa mère, morte aussi et mangée par l'osier, en bas...

Contre l'homme, la Mamèche tremble des nerfs comme une chèvre. Elle s'apaise. Elle caresse la grande cuisse solide, et maintenant elle parle une parole douce venue de son cœur doux comme une figue.

— Je pense à l'enfant, à mon petit, mon Rolando, celui qui est aussi sous la racine de l'herbe. C'est pas de la justice, Braë! Eux, ils les ont encore en chair qui marche et c'est parti pour chercher la bonne place. Moi, tout ce qui me tenait le cœur, c'est devenu l'herbe et l'eau de cette terre et je resterai ici tant que je ne serai pas devenue cette terre, moi aussi.

— Moi aussi, Mamèche, dit Panturle : j'ai la mère...

— Je vais te dire, fils, ce qui me fouille comme une bêche et que j'en souffre le martyre. Tant qu'on est là... près ça fera du bois sauvage et ça sera

« Ecoute : dès l temps qu'on était

REGAIN 53

mariés avec l'homme, on était du côté de Pignatello à travailler. J'allais avec lui sur le chemin; on traversait le bois et il y avait des charbonniers. Une fois, on s'est approché d'un endroit où il y avait toujours une meule de charbon qui fumait. C'était rasé tout autour; nous savions que l'homme allait couper le bois et qu'il l'apportait pour le cuire juste à cet endroit-là. On voulait savoir pourquoi. On s'est approché; alors on a vu : il y avait une baraque sous trois arbres; una cosa di niente, de rien, je te dis, grosse comme une noix. Il y avait, là devant, une femme et deux petits vautrés comme des chiennots.

« On a honnêtement demandé et la femme nous a dit. Ça n'était pas toute la famille, ces deux besons-là, il y en avait un autre dans la terre, bien sage pour toujours avec une barrière de bois autour de l'endroit où il était. Il y avait aussi dans la terre le père de la femme, un tout vieux et une petite d'une heure, morte pendant qu'on la faisait.

« Il y avait surtout, Braë, celui qui passait dans la fumée de la charbonnière, l'homme bien vivant et, dans lui, qui sait, qui sait combien d'enfants nouveaux, prêts à venir.

« Ça, ça a peut-être fait un village, depuis.

« Au lieu d'ici...

— Mamèche, il faut boire, dit Panturle.

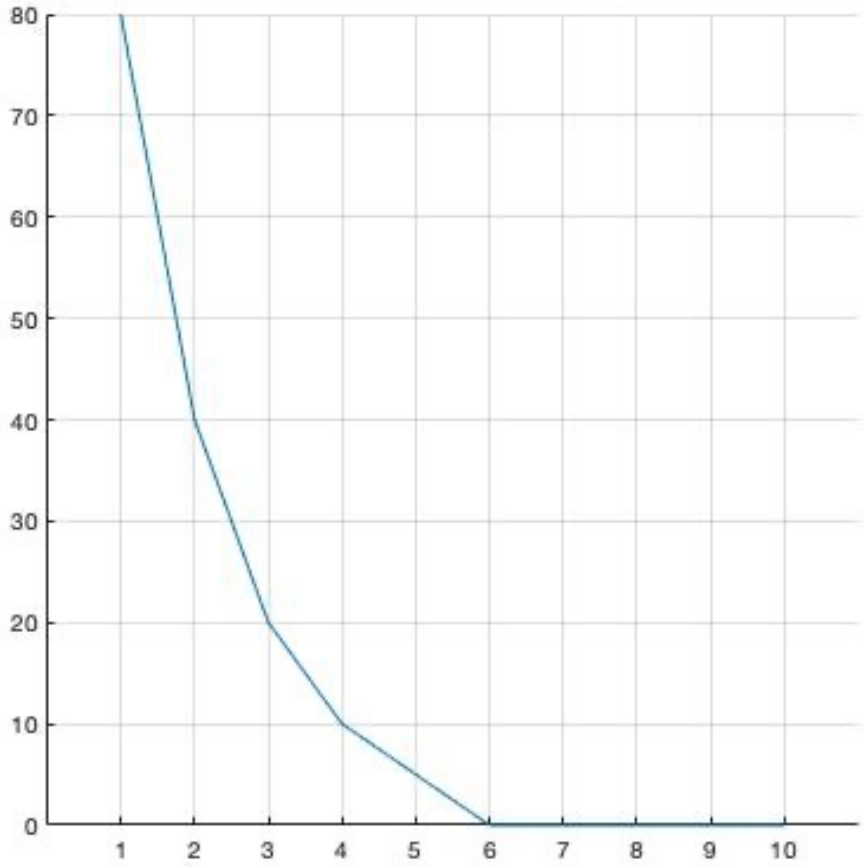
Et il prend un bol. Le lait s'est refroidi; il

In this task, you are required to transcribe the content of 10 photographs of pages in a book. Each photograph contains two pages, and you will be shown two pages at a time (as shown in the example image above). Each page contains approximately 100 words on average. You will be provided with two text-areas below each photograph, where you have to enter the transcription of each page. Required qualifications to take part in this task - submitted at least 100 tasks on Amazon Mechanical Turk. You will be paid an hourly wage of 15 USD.

1. **Image Quality and Legibility**: The quality of the photographs might be inconsistent, resulting in blurriness or unclear text (e.g., glare, poor focus, uneven lighting).
   - **Improvement**: Provide specific guidelines on how to capture high-quality images (e.g., good lighting, focus, stable surface) or review images before assigning transcription tasks to ensure legibility.
2. **Fatigue and Complexity of Task**: Transcribing the content of ten photographs containing two pages each might be too lengthy and could lead to fatigue, which in turn affects the quality of the work.
   - **Improvement**: Break down the task into smaller, more manageable units (e.g., fewer pages per task) to reduce fatigue and maintain quality. You could also implement a time limit to avoid prolonged work sessions.
3. **Ambiguity and Lack of Guidance for Difficult Text**: Some words in the book may be challenging to interpret (e.g., due to old fonts or worn text). Without guidance, transcribers may introduce errors when faced with unclear parts.
   - **Improvement**: Provide specific instructions on how to handle illegible text (e.g., marking unclear words with a placeholder or flagging them for review) and offer a few example transcriptions for guidance. Additionally, a review mechanism could be used to double-check transcriptions by multiple workers.

Case study 2

Given the figure below of eigenvalue spectrum, where x-axis is the principal components and y-axis is the variance in %:



1p **3a** What is the intrinsic dimensionality of the data?

2p **3b** Calculate the exact variance kept when choosing a dimensionality of 4 PCA features. And the percentage of kept variance after reduction.

2p **3c** What is the minimum number dimensions be after applying PCA dimensionality reduction if at least 70% of the variance is required to be kept. Explain your solution.

Case study 3

You are implementing a system in which you can model crowd behaviour. You implement different agents in a digital environment which move around in a virtual train station to see how what layout works best to make sure people aren't confused about where to go. Every agent has a set of rules which determine how it behaves, for instance determining their goal, how they react to other people, how they can understand the building, etc. You have a confusion score which keeps track of whether people feel lost or not. With this system, you hope to design better train stations for real people.

2p **4** Which of the four main definitions of AI suits the purpose of your agents best? Explain your answer

The most suitable definition of AI for your agents is **"Acting Rationally"**.

Explanation:

- The goal of your agents is to **move around in a virtual train station** and determine the best layout to ensure people are not confused or feel lost. Each agent follows a set of rules to achieve goals, react to other people, and understand the environment.
- This behavior aligns well with the **"acting rationally"** definition of AI, where agents are designed to take actions that maximize the likelihood of achieving a specific goal (in this case, reducing confusion).
- Rational agents use reasoning to make decisions that will lead to the optimal outcome, similar to how your agents navigate the environment and minimize the confusion score.

The other definitions (thinking humanly, thinking rationally, acting humanly) are less suitable because this system is more concerned with **effective actions** in an environment rather than simulating human thought processes or behaviors in detail. The primary focus is on **optimal navigation and behavior** in a given space.

Case study 4

Considering the standard k-means clustering algorithm using the Euclidean distance and the dataset shown in the table,

Data sample	Data value
1	-2
2	9
3	5
4	-4
5	10
6	6

2p **5a If the initialization of the centroids is at locations -14 and -8, how many iterations will the algorithm need to converge? What are the final centroids and clusters?**

To solve this k-means clustering problem, we will go through each iteration until the centroids converge.

Initial Centroids

Centroid 1: -14
Centroid 2: -8

Data Points

$-2, 9, 5, -4, 10, 6$

Iteration 1

Assign Data Points to Clusters:

1. Distance to centroid -14 and -8:
   - -2: Distance to -14 = 12, to -8 = 6 → Assigned to cluster 2
   - 9: Distance to -14 = 23, to -8 = 17 → Assigned to cluster 2
   - 5: Distance to -14 = 19, to -8 = 13 → Assigned to cluster 2
   - -4: Distance to -14 = 10, to -8 = 4 → Assigned to cluster 2
   - 10: Distance to -14 = 24, to -8 = 18 → Assigned to cluster 2
   - 6: Distance to -14 = 20, to -8 = 14 → Assigned to cluster 2

All data points are assigned to cluster 2.

New Centroids:

- Cluster 1 has no points → remains unchanged.
- Cluster 2 centroid = Average of $[-2, 9, 5, -4, 10, 6]$ = 4

Iteration 2

Now, centroids are: Centroid 1: -14
Centroid 2: 4

Assign Data Points to Clusters:

1. Distance to centroid -14 and 4:
   - -2: Distance to -14 = 12, to 4 = 6 → Assigned to cluster 2
   - 9: Distance to -14 = 23, to 4 = 5 → Assigned to cluster 2
   - 5: Distance to -14 = 19, to 4 = 1 → Assigned to cluster 2
   - -4: Distance to -14 = 10, to 4 = 8 → Assigned to cluster 2
   - 10: Distance to -14 = 24, to 4 = 6 → Assigned to cluster 2
   - 6: Distance to -14 = 20, to 4 = 2 → Assigned to cluster 2

Again, all data points are assigned to cluster 2.

New Centroids:

- Cluster 1 has no points.
- Cluster 2 centroid = Average of $[-2, 9, 5, -4, 10, 6]$ = 4

The centroid for cluster 2 remains 4, indicating convergence.

Result

- **Number of Iterations to Converge**: 2
- **Final Centroids**:
  - Centroid 1: -14 (remains unchanged)
  - Centroid 2: 4
- **Clusters**:
  - Cluster 1: No points
  - Cluster 2: All data points $[-2, 9, 5, -4, 10, 6]$

2p **5b The initialization of the centroids is at two different random data points x and y, x<y, define x and y such that the k-mean clustering provides two non-empty clusters. Also, what are the values of the final two centroids?**

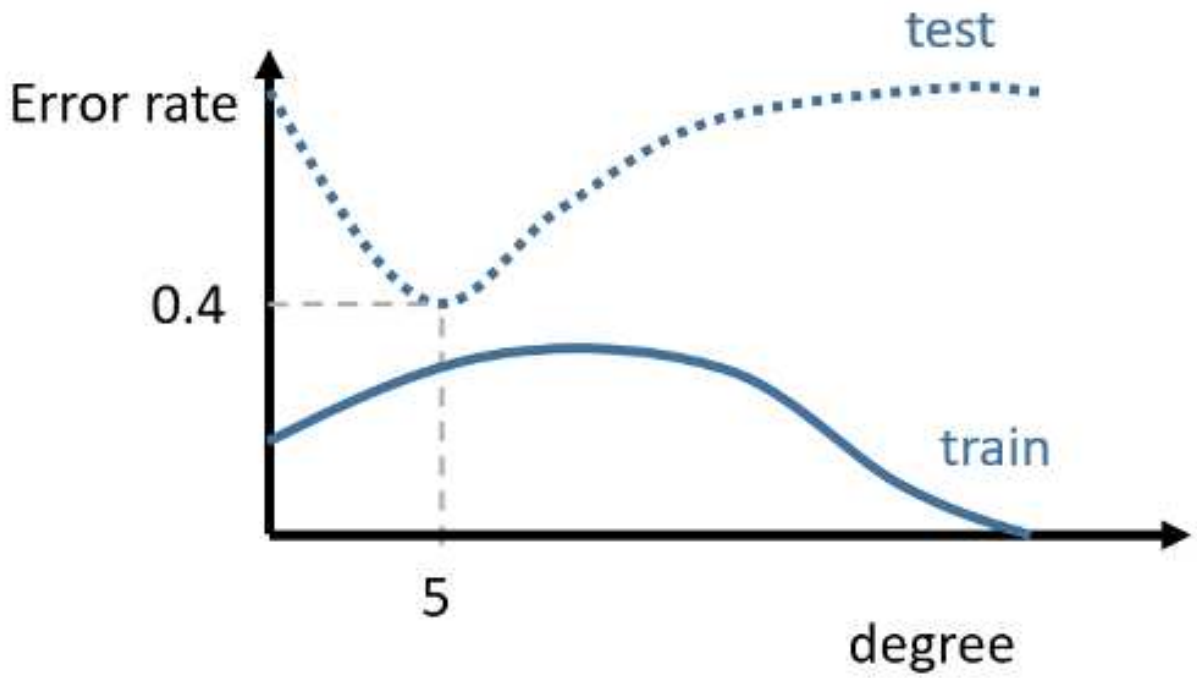
4p **5c Apply hierarchical clustering to the dataset using the single linkage criterion and the Euclidean distance (Note: Single linkage criterion assigns the distance between clusters as the minimum distance between members). Indicate the number of clusters that are obtained when the threshold is 2.5 and the data samples in the clusters. What is the range of threshold that will result in two clusters?**

Explanation:

- **At Threshold 2.5:**
  - Merges occur only at distances less than or equal to **2.5**.
  - Clusters formed are based on these merges, resulting in **3 clusters**.
- **For Thresholds Between 3 and 7:**
  - The merge at distance **3** reduces the number of clusters from **3** to **2**.
  - No further merges occur until the distance reaches **7**.
  - Thus, any threshold ≥ **3** and < **7** maintains **two clusters**.
- **At Threshold ≥ 7:**
  - All data points are merged into a single cluster.

Case study 5

Charles faces a classification task. It seems like the task is not linearly separable, which is why he extracts some polynomial features first, and afterwards trains a classifier on the extracted polynomial features. To figure out which degree should be used for the classification task, he generates the following complexity curve below using a single train and test set (holdout).



1p **6a** Explain why the polynomial degree should be tuned using test data and not using train data.

The polynomial degree should be tuned using test data (or validation data) and not train data to ensure that the model generalizes well to unseen data and avoids overfitting.

If we tune the polynomial degree using the training data, we might choose a high-degree polynomial that perfectly fits the training set, resulting in low training error but potentially very high error on unseen test data (as seen in the increasing test error for higher degrees) This phenomenon is called **overfitting**, where the model learns not only the underlying patterns but also the noise in the training data, leading to poor performance on new data.

By tuning with test data, we can find a degree that balances the model's complexity, providing a good fit for both the training and test datasets—leading to a model that performs well on unseen data. The goal is to achieve low error on both train and test sets, finding the optimal point of **generalization**.

1p **6b** In class we have discussed that the generalization error can be decomposed into: bias, variance, and the irreducible error.

True or False: When a larger degree is used, the model has a smaller bias. Explain your answer by explaining the meaning of bias, and explain why it is smaller or larger.

Bias refers to the error that is introduced by approximating a real-world problem (which m be extremely complex) by a much simpler model. It measures how far off the expected (average) predictions of the model are from the true underlying function. A model with hig bias pays little attention to the training data and oversimplifies the model, leading to systematic errors in the predictions—a phenomenon known as **underfitting**.

When using a **larger degree** in models like polynomial regression, you increase the model' complexity. A higher-degree polynomial can fit the training data more closely by capturing more of the underlying patterns and nuances in the data. This increased flexibility allows th model to approximate the true function more accurately, thereby **reducing the bias**.

Why Bias is Smaller with Larger Degree Models:

- **Increased Flexibility:** Higher-degree models have more parameters and can fit a wide variety of functions. This flexibility enables the model to closely follow the training data points.
- **Better Approximation:** With more degrees of freedom, the model can better approximate complex relationships in the data, aligning more closely with the true underlying function.
- **Reduced Systematic Error:** As the model becomes more capable of capturing the true patterns, the systematic error (bias) decreases because the difference between the expected predictions and the actual values diminishes.

However, it's important to note that while increasing the degree reduces bias, it may **increase variance**. High-degree models can become overly sensitive to the fluctuations in the training data, capturing noise as if it were a true signal—a phenomenon known as **overfitting**.

2p **6c** Charles is training a logistic regression classifier on top of the polynomial features. However, if he extracts too many polynomial features, sometimes his computer runs out of memory and throws an error message. He is using gradient descent to train his logistic regression classifier.

Explain which variants of gradient descent he should choose to resolve this issue, and explain why the issue is resolved.

To resolve the memory issue, Charles should use **Stochastic Gradient Descent (SGD)** or **Mini-batch Gradient Descent**.

Explanation:

- **Stochastic Gradient Descent (SGD)**: Instead of computing the gradient of the loss function using the entire dataset, SGD updates the model parameters for each training example one at a time. This significantly reduces memory usage since it only requires a small portion of the data to be loaded into memory at any given time. Therefore, SGD is particularly suitable when the dataset or the number of features is large, reducing the memory burden.
- **Mini-batch Gradient Descent**: This is a variant that falls between full-batch gradient descent and SGD. Mini-batch gradient descent computes the gradient using a small batch of data points (e.g., 32, 64, or 128 examples) instead of the entire dataset or a single example. This allows for a compromise between convergence stability (as compared to SGD) and memory efficiency. Mini-batch gradient descent still requires significantly less memory compared to full-batch gradient descent, as it only operates on smaller subsets of the data.

Why the Issue is Resolved: The primary memory issue occurs because computing the gradient over all the polynomial features requires a lot of memory, especially when the number of features becomes large. By using **SGD or Mini-batch Gradient Descent**, Charles can train his logistic regression model with much less data loaded into memory at any given time, thus avoiding memory overflow and making the training process feasible even with a large number of polynomial features. These methods reduce the **computational load per iteration** and make it easier to train models on resource-constrained systems.

2p **6d** Charles decides to use a polynomial degree of 5. His test set contains 10 samples. He estimates using the graph above that the generalization error of his SVM is 40%.

Explain why some machine learners may consider his estimate of 40% to be “cheating”, untrustworthy or overly optimistic. How should the generalization error be estimated instead?

Some machine learners may consider Charles's estimate of **40% generalization error** to be "cheating," untrustworthy, or overly optimistic because he is using a **single test set** containing only **10 samples**. This is problematic for a few reasons:

1. **Small Test Set Size**: A test set of only 10 samples is insufficient to provide a reliable estimate of the model's performance. The variance in the error estimate will be very high with such a small test set, which means that the results can be overly optimistic or pessimistic depending on the particular data points in the set. A small sample size does not adequately represent the full complexity and variability of the data distribution.
2. **Overfitting to the Test Set**: If the model evaluation is done on a single test set, and especially if there has been any tuning of hyperparameters (e.g., polynomial degree, regularization strength) based on this test set, it can lead to overfitting to the test data. This would make the performance estimate biased and give an **overly optimistic** view of the model's ability to generalize to unseen data.

Proper Way to Estimate Generalization Error

To get a more reliable estimate of the generalization error, Charles should consider using techniques such as:

1. **Cross-Validation**:
   - **K-Fold Cross-Validation**: In this approach, the dataset is split into **K** equally-sized folds. The model is trained **K** times, each time using a different fold as the validation set and the remaining **K-1** folds as the training set. The final performance is calculated by averaging the results from all **K** folds. This provides a more reliable estimate of the model's performance by ensuring that every data point is used for both training and validation, reducing the chance of overfitting to a single small test set.
2. **Repeated Train-Test Splits**:
   - Another approach is to **repeatedly split** the data into different training and test sets, train the model, and then average the results. This can also provide a better estimate of how well the model will generalize, as the results are less dependent on a particular split.
3. **Larger Test Set**:
   - If cross-validation is not feasible, Charles should at least increase the size of the test set to make sure it is representative of the entire dataset. A larger test set helps reduce the variance in the error estimate and provides a more trustworthy evaluation of the model's performance.

2p **6e** For Charles' classification problem, the priors of the classes are p(a) = 0.20 and p(b) = 0.80. Explain whether a generalization error of 40% is good or not and why.

For Charles' classification problem, where the **class priors** are:

- $p(a) = 0.20$
- $p(b) = 0.80$

This means that **class b** is much more frequent compared to **class a**, implying an **imbalanced dataset**.

If Charles' model has a **generalization error of 40%**, it means that the model is **incorrect** in 40% of the test samples, resulting in an accuracy of **60%**. However, to determine whether this is good or not, we need to compare it with the **baseline accuracy** derived from the class distribution.

Baseline Accuracy:

- Since $p(b) = 0.80$, a **naive classifier** that always predicts the majority class (class **b**) would have an accuracy of **80%**. This is known as the **baseline accuracy**.

Analysis:

- Charles' **generalization error of 40%** corresponds to an accuracy of **60%**, which is **worse than the baseline** accuracy of **80%**.
- In other words, if the model simply predicted **class b** for every instance, it would achieve **80% accuracy** due to the imbalanced class distribution. Charles' model, however, is only achieving **60% accuracy**, indicating that it is underperforming compared to even a very simple approach that ignores the features entirely and just predicts the majority class.

Conclusion:

- A **40% generalization error** is **not good** in this case because it is **worse than the baseline** that could be achieved by simply predicting the majority class.
- The model needs improvement to perform better than the naive strategy of predicting only the majority class. Charles could consider using techniques such as **balancing the dataset** (e.g., oversampling the minority class or undersampling the majority class), **using class weights**, or employing models that can handle class imbalance better.

Case study 6

You would like to use the linear model

$Y_i = \beta_0 + \beta_1 X_{i1} + \ldots.. + \beta_p X_{ip}, \quad$ where $i = 1, \ldots, n.$

2p **7a** In which type of situation would you use Ridge regression or Lasso rather than the least squares fit to estimate the coefficients $\beta_0, \beta_1 \ldots, \beta_p$? Give two examples and explain your reasoning.

Example 1: High Multicollinearity

- **Situation**: You have a dataset with many features that are highly correlated with each other.
- **Reasoning**: When features are highly correlated, **least squares estimates** can become unstable and have high variance, leading to unreliable coefficient estimates.
  - **Ridge regression** (L2 regularization) adds a penalty proportional to the **squared magnitude** of the coefficients. This helps **shrink the coefficients**, making them more stable and reducing the variance. Ridge is particularly useful when you want t **retain all features** while mitigating the impact of multicollinearity.

Example 2: Feature Selection with Many Features

- **Situation**: You have a large number of features, and you believe that only a subset of them is actually important for predicting the outcome.
- **Reasoning**: When you have many features, using least squares can lead to a model that overfits the training data, especially if many of the features are irrelevant.
  - **Lasso regression** (L1 regularization) is useful in this case because it adds a penalty proportional to the **absolute value** of the coefficients, leading to **sparse solutions** where some coefficients are driven to zero. This effectively performs **feature selection**, retaining only the most important features and producing a more interpretable model.

2p **7b** In which type of situation would you use Ridge regression rather than Lasso?

You would use **Ridge regression** instead of **Lasso** in the following type of situation:

Situation: Many Features with Small Effects, and Multicollinearity is Present

- **High-Dimensional Data with Many Small Effects**: If you have a dataset with a **large number of features** and you believe that **most or all features contribute** to the target variable, even if their contributions are relatively small, **Ridge regression** is a better choice.
  - **Reasoning**: Ridge regression (L2 regularization) **shrinks** the coefficients towards zero but does not force them exactly to zero. It retains all features by shrinking their effects proportionally. This is beneficial if you believe that **all features have some predictive power** and you do not want to exclude any feature completely. Ridge is also well-suited when you want a **stable model** in the presence of **multicollinearity**, as it helps reduce the variance of the coefficient estimates.

Comparison with Lasso:

- **Lasso** (L1 regularization) tends to **drive some coefficients to zero**, effectively performing **feature selection**. This is useful when you expect **only a few features** to be significant, while the rest are irrelevant.
- On the other hand, **Ridge** keeps all the features, which can be preferable if all features are thought to be important in the model to some extent.

Example:

- Consider a **genomic dataset** with thousands of gene expression levels, where you believe that many genes are correlated and all have **small effects** on predicting a disease outcome. Ridge regression is preferable here because it will keep all the gene features, ensuring that none are excluded, while reducing overfitting due to multicollinearity.

1p **7c** In which type of situation would you use Lasso rather than Ridge regression?

Situation: Feature Selection or Sparse Solutions

- **High-Dimensional Data with Only a Few Important Features**: If you have a dataset with a **large number of features** but you suspect that **only a subset** of these features is truly important for predicting the target variable, **Lasso regression** is more suitable.
  - **Reasoning**: Lasso regression (L1 regularization) adds a penalty proportional to the **absolute values** of the coefficients, which results in some coefficients being **exactly zero**. This effectively performs **feature selection**, retaining only the most important features and making the model more interpretable. Lasso is especially helpful when you want to identify which features are significant and exclude irrelevant or redundant ones.

Example:

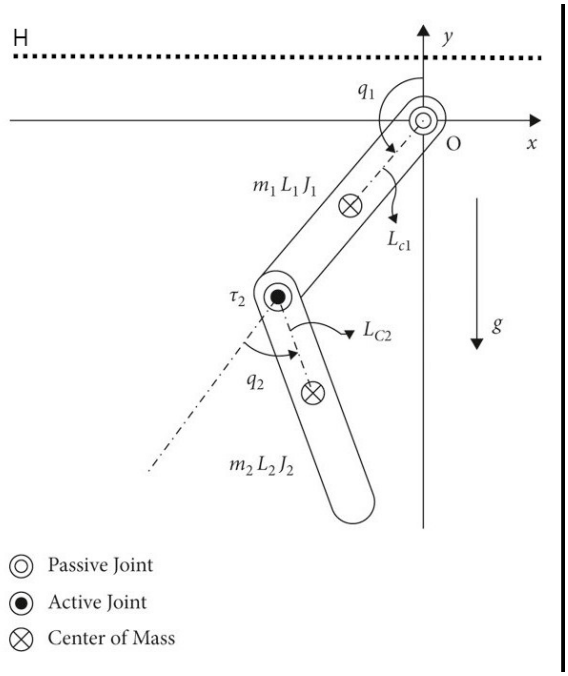
- **Marketing Campaign Analysis**: Suppose you have data with **hundreds of features** representing different customer characteristics (e.g., age, income, browsing habits, etc.) for predicting whether a customer will respond to a marketing campaign. You suspect that **only a few characteristics** actually influence the response. **Lasso regression** can be used to identify the most influential features by driving irrelevant coefficients to zero, leading to a **simpler** and **more interpretable** model with **better generalization**.

Comparison with Ridge Regression:

- **Ridge regression** (L2 regularization) shrinks coefficients but **does not eliminate** any of them, which means all features are retained with reduced importance. This is preferable when you believe **all features** have some predictive power.
- **Lasso**, on the other hand, is better suited for situations where you believe **only a subset** of the features are important, and you want to identify or **select** those important features while setting others to zero.

Case study 7

The picture below shows an "Acrobot" environment. A build-in motor exerts a torque (angular acceleration) on active joints, which can accelerate the joint's rotation in both directions, whereas passive joints swing freely. The task is to swing any part of the Acrobot over the dashed line at height *H*, while using as little torque and time as possible.



2p **8a** Define a low-dimensional continuous state space for the Acrobot

A low-dimensional continuous state space for the **Acrobot** environment can be defined by the following state variables:

1. **Angles of the Two Links** (q_1, q_2): The angles of the two links relative to a reference axis (e.g., the vertical axis). These angles capture the current configuration of the acrobot.
2. **Angular Velocities of the Two Links** ($\dot{q}_1, \dot{q}_2$): The angular velocities of the two links. These represent how fast the links are moving and in which direction.

Thus, the low-dimensional continuous state space can be represented as:

$S = [q_1, q_2, \dot{q}_1, \dot{q}_2]$

This state space is sufficient to describe the Acrobot's dynamics, as it captures both the configuration and the movement of the two links, which are necessary for planning the required torques and predicting future states.

1p **8b** Define a low-dimensional continuous action space the Acrobot could use.

A low-dimensional continuous action space for the **Acrobot** can be defined as the **torque** applied at the **active joint**. The action space could be represented by a single continuous variable:

- **Torque (τ)**: A continuous value representing the amount of torque applied by the motor at the active joint. This torque can be positive or negative, allowing for movement in either direction.

Thus, the action space can be defined as:

$A = [\tau]$

The torque can vary within a certain range, depending on the physical limits of the motor. By using a continuous action space, the Acrobot can finely adjust the torque applied to the active joint to achieve precise control for reaching the target height while minimizing energy use.

3p **8c** Define a reward function, which would incentivize a reinforcement learning algorithm to learn the above task. Explain why your reward solves the task.

Reward Function:

1. **Reaching the Target Height**:
   - Provide a **large positive reward** when the Acrobot's endpoint reaches or exceeds the target height H, e.g., reward = $+100$. This motivates the Acrobot to achieve the target height.
2. **Distance to Target (Before Reaching)**:
   - If the Acrobot has not yet reached height H, assign a **negative reward** proportional to the distance from the highest point of the Acrobot to the target height H, e.g., $R_{\text{distance}} = -|H - y_{\text{end}}|$, where y_{end} represents the height of the end of the Acrobot. This encourages the Acrobot to move closer to the target.
3. **Penalty for Torque**:
   - Penalize the **magnitude of the applied torque** to encourage minimal energy use, e.g., reward = $-\alpha|\tau|$, where α is a small positive constant to control the penalty strength.
4. **Time Penalty**:
   - Apply a **small negative reward** for each time step, e.g., reward = -1 per time step, to encourage the Acrobot to reach the target as quickly as possible.

Final Reward Function:

$$R = \begin{cases} +100 - \alpha|\tau| - 1 & \text{if } y_{\text{end}} \geq H \text{ (reached the target)} \\ -|H - y_{\text{end}}| - \alpha|\tau| - 1 & \text{if } y_{\text{end}} < H \text{ (not yet reached)} \end{cases}$$

↳

Deep Q-networks (DQN)

1p **8d** Which deep reinforcement learning algorithm from the lecture could be used to learn this task? You do not need to justify your answer.

Deep Q-networks (DQN)

Case study 8

Due to the opioid crisis afflicting the USA, machine-learning based Prescription Drugs Monitoring Program platforms (PDMPs) have been recently introduced, with the aim to predict patient's probability to develop an opioid addiction or misuse. Based on the risk scores generated by these algorithms, medical decisions are taken by physicians regarding which drugs to prescribe and to whom. The use of these systems is legally enforced upon physicians who are expected to make use of the risk scores provided. Taking medical decisions that deviate from the algorithmic risk scores could even cause them to lose their practicing license.

Even though they play a relevant role in crucial healthcare decisions, these systems are completely opaque to the end-users (patients and physicians alike). NarxCare's algorithms are proprietary and the company owning these systems is not ready to share information regarding, for example, the nature and weight of the proxies that inform the risk scores. This can lead to unjust treatment of patients: in fact, patients can be denied pain medication due to correlations established by the algorithm that have nothing to do with their eligibility to receive opioid medications.

This is what happened to Kathryn, a woman suffering from a pathology requiring pain medication. She was suddenly discharged from the hospital and her physician refused to continue treating her condition. She later realized that this happened because she was red-flagged by the machine learning system as being at a high risk of drug misuse (this means, she was not informed that a machine learning system was involved in the decision-making process in the first place). Even though she could later understand why she was denied medical care and why her risk score was too high even though she never misused drugs (she had sick pets requiring strong medications and this also informed her own risk score), she could do nothing to change the risk score that was unjustly attributed to her.

3p **9a** Which approaches towards transparency do you think would be needed to ameliorate unjust outcomes ensuing from the use of NarxCare algorithms?

3p **9b** In your opinion, to what extent are physicians to be considered morally responsible for the decisions taken in the case under scrutiny? Can you recognize any responsibility gaps?

3p **9c** Which forms of contestability do you think should be designed into these systems to empower patients that are attributed a risk score that does not mirror their actual drug consumption levels? Can any values conflicts emerge from allowing these forms of contestability?

3p **9d** Which kind of explanatory information would be needed, in your opinion, for physicians to understand and evaluate the suitability of the outputs produced by these systems

This page is left blank intentionally