

Part Two: Problems

Question One

1. Is any of the following problems an example of a supervised learning problem? (You can choose multiple choices). (10 pts)

- A. Given a set of patient data, including various features such as breast images, family history, cell nuclei measurements (for example, radius, texture, perimeter, area, smoothness of the cells), and the biopsy results of benign and malignant, the goal is to classify whether a tumor is benign (non-cancerous) or malignant (cancerous).
- B. Given images of crop (wheat or rice) leaves and biologist's labels in terms of whether the crop is healthy or suffering from a specific disease (for example, powdery mildew, leaf rust, or bacterial blight), the goal is to classify whether any new plant image is healthy.
- C. Given a large dataset of customer behaviors, such as purchase history, browsing patterns, demographics, and engagement metrics, the goal is to group customers into distinct segments based on similarities in their behavior and preferences. There are no predefined labels for the customer segments.
- D. Given historical data of a satellite's position, velocity, and various environmental factors (such as gravitational forces, solar radiation pressure, and atmospheric drag), the goal is to predict the satellite's future position in space at a given time.

2. Is any of the following problems an example of an unsupervised learning problem? (You can choose multiple choices). (10 pts)

- A. Given a large dataset of images of faces, each image is represented as a high-dimensional vector (e.g., each pixel in an image is a feature). The goal is to reduce the dimensionality of the data while preserving the most important information, so that the faces can still be recognized or compared with minimal computational cost.
- B. In a manufacturing plant, machines produce data through sensors, such as temperature, vibration, and sound levels. The goal is to detect unusual patterns or behaviors (anomalies) that could indicate machine faults or maintenance needs. The challenge is that there are no pre-observed labels of normal or faulty operation for every possible scenario.
- C. An autonomous vehicle must learn how to drive through complex environments, such as urban roads, highways, and intersections, while interacting with other vehicles, pedestrians, traffic signals, and obstacles. The goal is to enable the vehicle to make safe and efficient driving decisions, such as when to accelerate, brake, change lanes, or stop at traffic lights.
- D. In an online learning platform, some students may frequently revisit course materials and may benefit from more practice problems, while some other students need to be challenged with more advanced content. An algorithm might analyze student data (e.g., study habits, online learning activity, quiz completion times, etc.) and groups students with similar learning patterns.

3. Is any of the following methods capable of solving a supervised learning problem? (You can choose multiple choices). (10 pts)

- A. Linear regression
- B. Random forests
- C. Support vector machines
- D. K-nearest neighbors

4. For a quadratic form as follows (Eq. 1, please compute the gradient of this function $f(x)$ where x is a vector in R^n , A is a $n \times n$ matrix, B is a vector in R^n , and C is a scalar. (20 pts) (Write your derivation. Do not just copy-paste what GPT provides if you use GPT.)

$$f(x) = x^T A x + B^T x + C \quad (1)$$

5. The following function (Eq. 2) is a quadratic form, can you write out its gradient, and compare with the above gradient formula you derived in Problem 4. (10 pts)

$$f(x) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c \quad (2)$$

6. For a sigmoid function $\sigma(s) = \frac{1}{1+e^{-s}}$, where s is a scalar, prove that its derivative satisfies $\sigma'(s) = \sigma(s) \cdot (1 - \sigma(s))$. (20pts)