Feature Selection

1. Consider a binary classification problem in a 2D feature space. What is the shape of the boundary separating the 2 classes in an ideal setting?

   - <mark>Linear</mark>
   - Parabola
   - Sigmoid
   - Perceptron

   This is the simplest functional form of a boundary.

2. Feature selection is characterized by: (check all that apply).

   - <mark>Identify features that best represent the relationship between two or more variables.</mark>

     Feature selection identifies features with predictive power.

   - Ensuring numerical features follow the same numerical range
   - Accounting for data changes over time (drift, seasonality, etc).
   - <mark>Remove features that don't influence the outcome.</mark>
   - Ensuring that the serving dataset is representative of future inference requests.

3. What is the definition of backward elimination?

   - <mark>In this method we start by selecting all the features. We then remove the least significant feature based on model performance. We repeat this step until no improvement is observed in model performance.</mark>
   - We start by selecting all features in the feature set and calculating their feature importances. We then prune features from the current feature set to select a subset of the features based on the feature importances, We recursively prune features on the new subset until no model performance improvement is observed.
   - We first start with no features. In each iteration we keep adding features which will increase the model performance until no performance improvement is observed.

4. Embedded methods combine the best of both worlds, filter and wrapper methods. Embedded methods are: (Check all that apply)

- Faster than filter methods

- More efficient than wrapper methods

- Faster than wrapper methods

  Wrapper methods are based on the greedy algorithm and thus solutions are slow to compute.

- More efficient than filter methods

  Filter methods suffer from inefficiencies as they need to look at all the possible feature subsets.