1. Машинное обучение - это:  
     
   а) Метод для обучения компьютера человеческим навыкам (0%)  
   б) Тип искусственного интеллекта, использующий нейронные сети (0%)  
   в) Подраздел искусственного интеллекта, который позволяет компьютеру учиться на основе данных (100%)  
   г) Алгоритм для выполнения конкретной задачи, не требующий обучения (0%)  
   д) Только технология для обработки текста (0%)
2. Bias (смещение) в контексте машинного обучения:  
     
   а) Означает, что модель слишком сложная и подстроена под обучающие данные (0%)  
   б) Обозначает ошибку, которая происходит из-за предвзятости модели к определенным результатам (100%)  
   в) Является синонимом разброса модели (0%)  
   г) Представляет собой количество параметров в модели (0%)  
   д) Означает, что модель не способна адекватно обучаться на данных (0%)
3. Какие из следующих являются основными типами задач в машинном обучении? Выберите все верные утверждения:  
     
   а) Задачи классификации (25%)  
   б) Задачи регрессии (25%)  
   в) Задачи кластеризации (25%)  
   г) Задачи анимации (0%)  
   д) Задачи обучения с подкреплением (25%)  
   е) Задачи перевода языка (0%)
4. Линейная регрессия является:  
     
   а) Нелинейной моделью (0%)  
   б) Параметрической моделью (0%)  
   в) Моделью без обучения (0%)  
   г) Нелинейной моделью с неограниченным числом параметров (0%)  
   д) Параметрической моделью с линейной зависимостью между переменными (100%)
5. Функция потерь использующаяся в линейной регрессии:  
     
   а) Cross-Entropy (0%)  
   б) Hinge Loss (0%)  
   в) Mean Absolute Error (MAE) (0%)  
   г) Mean Squared Error (MSE) (100%)  
   д) Log Loss (0%)
6. Ridge и Lasso регрессии отличаются друг от друга:  
     
   а) Только типом данных, которые могут быть обработаны (0%)  
   б) Подходом к устранению мультиколлинеарности (0%)  
   в) Способом штрафования коэффициентов в модели (100%)  
   г) Количеством используемых итераций (0%)  
   д) Различными функциями потерь (0%)
7. Каковы преимущества метода градиентного спуска для обучения модели линейной регрессии?  
     
   а) Меньше требуется вычислительных ресурсов (0%)  
   б) Гарантированно сходится к глобальному оптимуму (0%)  
   в) Может застрять в локальных оптимумах (0%)  
   г) Эффективно работает на больших объемах данных (100%)  
   д) Требует меньше итераций для сходимости (0%)
8. При использовании полиномиальной регрессии с полиномиальной степенью 2:  
     
   а) Модель становится менее гибкой и склонной к переобучению (0%)  
   б) Увеличивается сложность модели, что может привести к переобучению (0%)  
   в) Уменьшается вероятность переобучения (0%)  
   г) Модель способна улавливать нелинейные зависимости между переменными (100%)  
   д) Происходит уменьшение числа параметров модели (0%)
9. Какие методы регуляризации могут быть использованы в линейной регрессии? Выберите все верные утверждения:  
     
   а) L1 регуляризация (Lasso) (33%)  
   б) L2 регуляризация (Ridge) (33%)  
   в) Elastic Net регуляризация (0%)  
   г) DropOut регуляризация (33%)  
   д) Bootstrap регуляризация (0%)  
   е) Случайный лес как метод регуляризации (0%)
10. Какие из перечисленных являются методами обработки выбросов при работе с регрессионными моделями? Выберите все верные утверждения:  
      
    а) Удаление выбросов (33%)  
    б) Использование медианы вместо среднего значения (33%)  
    в) Игнорирование выбросов при обучении модели (0%)  
    г) Использование методов робастной регрессии (например, RANSAC) (33%)  
    д) Применение алгоритма SMOTE (0%)
11. Метод k-ближайших соседей (k-NN) относится к:  
      
    а) Обучению с учителем (0%)  
    б) Обучению без учителя (0%)  
    в) Ленивому обучению (100%)  
    г) Обучению с подкреплением (0%)  
    д) Полностью надзорному обучению (0%)
12. Support Vector Machines (SVM) могут использоваться для:  
      
    а) Только для задачи регрессии (0%)  
    б) Только для задачи кластеризации (0%)  
    в) Классификации и регрессии (100%)  
    г) Решения только линейно разделимых задач (0%)  
    д) Обработки текстовых данных (0%)
13. Дерево решений в задаче классификации:  
      
    а) Всегда создает бинарное дерево (0%)  
    б) Гарантированно не подвержено переобучению (0%)  
    в) Не может обрабатывать категориальные признаки (0%)  
    г) Может быть глубоким и склонным к переобучению (100%)  
    д) Не поддерживает ансамблевые методы (0%)
14. Алгоритм классификации основанный на принципе "голосования" различных моделей:  
      
    а) Логистическая регрессия (0%)  
    б) Random Forest (0%)  
    в) Naive Bayes (0%)  
    г) Ансамбль (Ensemble) (100%)  
    д) k-NN (0%)