

**1. Что такое линейная регрессия и для каких задач она применяется?**

**В.** Это метод, который моделирует зависимость целевой числовой переменной от входных признаков в виде линейной функции.

**2. Какой принцип лежит в основе метода наименьших квадратов для оценки коэффициентов линейной регрессии?**

**В.** Минимизация суммы квадратов разностей между истинными и предсказанными значениями.

**3. Какая из следующих метрик НЕ применяется для оценки качества моделей регрессии?**

**С.** Accuracy (Точность).

**4. Чем отличается аналитический метод оценки коэффициентов (МНК) от численного метода (например, градиентного спуска) при обучении линейной регрессии?**

**А.** Аналитический метод гарантирует точное решение, а численный – приближённое; численный метод может использоваться при больших количествах признаков и больших объёмах данных.

**5. Что показывает коэффициент детерминации  $R^2$  в модели регрессии?**

**В.** Долю дисперсии целевой переменной, объяснённую моделью (от 0 до 1, где значения ближе к 1 свидетельствуют о лучшем качестве).

**6. Почему важно масштабировать (нормализовать или стандартизировать) данные при использовании градиентного спуска?**

**В.** Приведение признаков к единому масштабу помогает обеспечить более стабильную и быструю сходимость алгоритма, поскольку градиенты для признаков с разными масштабами будут сравнимы.

**7. Чем L1-регуляризация (Lasso) отличается от L2-регуляризации (Ridge) в задаче линейной регрессии?**

**В.** L1-регуляризация имеет свойство обнулять некоторые коэффициенты (что ведёт к разреженной модели), а L2-регуляризация снижает значения коэффициентов, но обычно не обнуляет их.

**8. Что подразумевается под полиномиальными признаками, и зачем они используются в модели регрессии?**

**А.** Это признаки, полученные путём возведения исходных признаков в степень или вычисления их взаимодействий; их применение позволяет моделировать нелинейные зависимости с помощью линейных моделей.

**9. Что такое переобучение (overfitting) модели, и каким образом его можно обнаружить?**

**В.** Ситуация, когда модель слишком точно подстраивается под обучающие данные (низкая ошибка на обучающем наборе, но высокая ошибка на тестовом), что можно обнаружить сравнением показателей (например,  $R^2$ , MAE) между тренировочной и тестовой выборками.