1. **Как можно решить недостаток функции активации ReLU?**

Недостаток ReLU (Rectified Linear Unit)заключается в том, что она может привести к «смерти» нейронов, то есть к тому, что некоторые из них перестанут работать и будут выдавать нулевые значения на выходе. Это может произойти, если входное значение меньше нуля. В таком случае ReLU будет возвращать ноль, и градиент (мера изменения функции) также будет равен нулю. Это означает, что нейрон не будет обучаться и останется в состоянии «смерти».

Чтобы решить эту проблему, можно использовать нормализацию данных. Нормализация входных данных может сделать их более подходящими для использования с ReLU, что поможет избежать ситуаций, когда входные значения становятся слишком большими или слишком маленькими. Также можно использовать функции Leaky ReLU или ELU.

1. **Что такое двойственная задача?**

В машинном обучении двойственная задача — это задача оптимизации, которая получается из исходной задачи путём преобразования её в задачу минимизации функции Лагранжа.

Функция Лагранжа представляет собой сумму целевой функции и ограничений, умноженных на соответствующие множители Лагранжа (неотрицательные коэффициенты). Двойственная задача заключается в максимизации этой функции по множителям Лагранжа при условии, что ограничения исходной задачи выполняются как равенства.

Двойственные задачи используются в машинном обучении для решения задач оптимизации с ограничениями, таких как линейное программирование, квадратичное программирование и другие. Они позволяют получить оптимальные значения параметров модели, учитывая ограничения на ресурсы или другие факторы.

**3. Как используется метод Лагранжа для SVM?**

В SVM метод Лагранжа применяется для нахождения оптимальных параметров модели, которые минимизируют ошибку классификации.

Для этого строится функция Лагранжа, которая включает в себя целевую функцию и ограничения на параметры модели. Затем с помощью метода множителей Лагранжа находятся оптимальные значения параметров, при которых функция Лагранжа достигает минимума (решается двойственная задача).

Таким образом, метод Лагранжа позволяет найти оптимальные параметры модели опорных векторов, что обеспечивает наилучшую классификацию данных.