NETFLIX DATA SCIENCE INTERVIEW QUESTIONS

WHAT IS DIFFERENCE BETWEEN BATCH AND ONLINE GRADIENT DESCENT

In batch gradient descent, the model looks at the entire dataset at once to calculate the gradient (the direction to minimize error) and update parameters. This means it calculates the average gradient across all data points and then makes one update.
Pros: More stable and accurate, since it uses all data at each step.
Cons: Can be slow and memory-intensive, especially with large datasets, since it needs to process all data at once.
Online Gradient Descent

In online gradient descent, the model updates parameters one data point at a time. Each new data point provides a quick update without waiting for all data to be processed. This approach is also called stochastic gradient descent (SGD) because each update uses a random single point, adding some randomness to the updates.
Pros: Faster and requires less memory, as it only looks at one data point at a time. Works well for very large datasets or data that's continuously updating (e.g., real-time applications).
Cons: Less stable because updates are noisier (one point at a time can vary a lot), so it may "zigzag" toward the minimum rather than taking a direct path.

@karunt

WHAT MAKE RELU AN EFFECTIVE ACTIVATION FUNCTION?

The Rectified Linear Unit, or ReLU works so well because -

Simplicity: ReLU is easy to compute. It simply takes any negative value and turns it into zero, while keeping positive values as they are. This makes it fast and efficient.

- Formula: ReLU(x)=max(0,x)
- ReLU(x)=max(0,x)\text{ReLU}(x) = \max(0, x)

Avoids the Vanishing Gradient Problem: Many activation functions (like sigmoid or tanh) squish values to be between -1 and 1, causing gradients to shrink and slowing down learning. ReLU avoids this by not limiting the positive side, allowing gradients to stay larger, which helps in learning.

Sparse Activation: ReLU turns off (outputs zero) for any negative input. This makes the network "sparse" by reducing unnecessary signals, which improves efficiency and reduces the chances of overfitting.

In short, ReLU is popular because it's fast to compute, helps with efficient learning, and prevents certain problems other functions have.

EXPLAIN ANOVA TEST? (FOLLOW-UP: EXPLAIN MEANING OF P-VALUES)

The ANOVA test is used to compare the means of three or more groups to determine if there is a significant difference among them. How ANOVA Works -

1. Null Hypothesis (H_0): All group means are equal (i.e., any observed differences are due to random variation).
2. Alternative Hypothesis (H_1): At least one group mean is significantly different from the others.

A p-value represents the probability of observing the test results, or something more extreme, under the assumption that the null hypothesis is true.

- Low p-value (< 0.05): Indicates that the observed data is unlikely under the null hypothesis, so we have evidence to reject it in favor of the alternative hypothesis. This suggests a statistically significant effect.
- High p-value (≥ 0.05): Suggests that the observed data is plausible under the null hypothesis, so we fail to reject it. There isn't strong evidence for a significant difference.

@karunt

WHAT ARE THE KEY METRICS YOU WOULD CONSIDER WHEN EVALUATING THE PERFORMANCE OF A RECOMMENDATION ALGORITHM?

Precision@K and Recall@K: These measure the relevance of the top K recommended items. Precision@K is the proportion of relevant items in the top K recommendations, while Recall@K measures the proportion of all relevant items that appear in the top K.

Hit Rate: Measures how often the recommended list contains at least one item that the user interacts with or rates highly, indicating that the model is generating some relevant suggestions.

Normalized Discounted Cumulative Gain (NDCG): Measures the quality of ranked lists by considering the position of relevant items in the recommendations, with higher rewards for higher-ranking relevant items. This is especially useful for ordered lists, like search results or top recommendations.

Diversity: Evaluates how different the recommended items are from each other. High diversity ensures users are not just shown similar items repeatedly, which improves engagement.

Click-Through Rate (CTR): The ratio of users who click on recommended items to those who view them. A higher CTR suggests that the recommendations are capturing user interest.

HOW WOULD YOU BUILD AND TEST A METRIC TO COMPARE TWO USERS' RANKED LISTS OF MOVIE/TV SHOW PREFERENCES?

A few metrics to consider are -

Kendall's Tau: Measures how similarly two lists are ranked by counting the number of pairwise swaps needed to convert one list into the other. It's a good choice when you want to assess the order of preferences rather than exact placement.

Spearman's Rank Correlation: Measures correlation based on rank, ignoring exact scores but comparing the relative order of items. It's helpful if you only have the ranks of items and want a measure that handles ties.

Normalized Discounted Cumulative Gain (NDCG): This is a relevance-based metric that measures how well a ranked list aligns with a ground truth list, useful if some items in the list are considered more "relevant" than others.

A few other things to consider are - making sure the list lengths are the same, and conducting testing by swapping order to see how much the above metrics change etc.

WAS THIS HELPFUL?

Be sure to save it so you can come back to it later!

@karunt