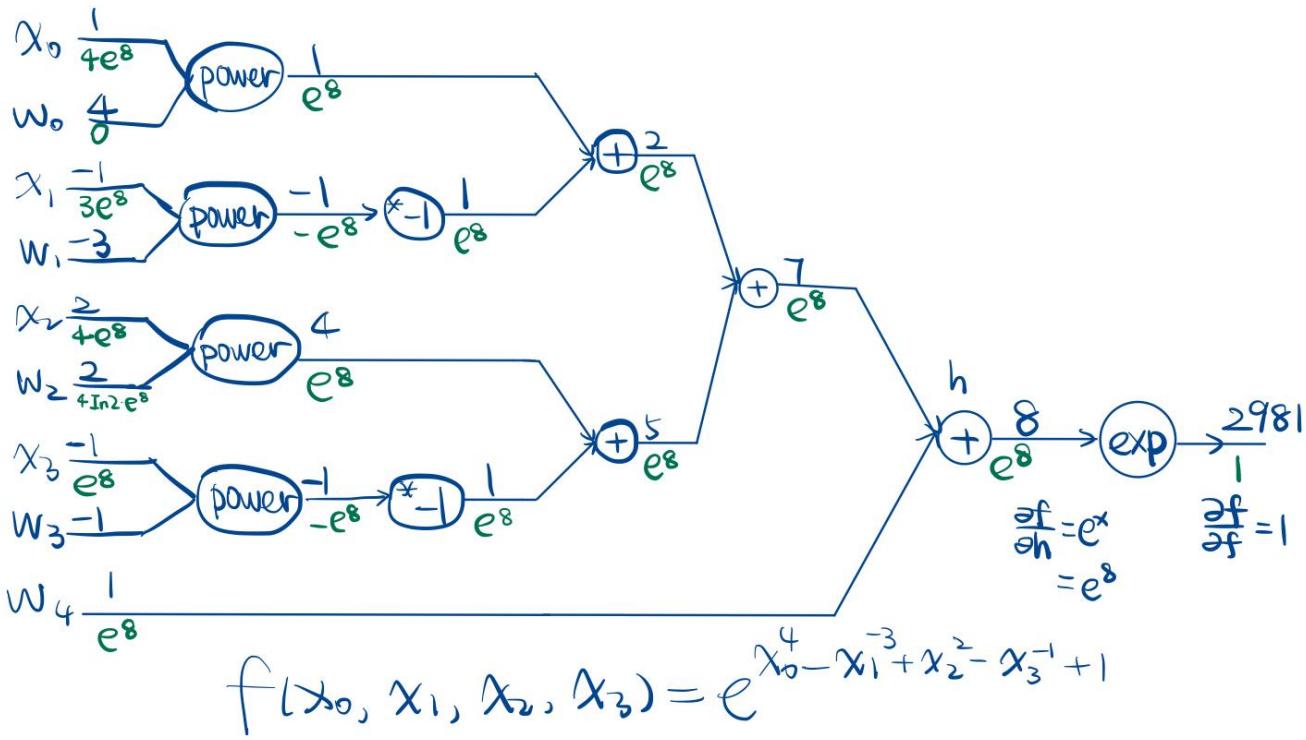
Question 1:



Question 2:

When I conducted the procedure with different numbers of hidden units, I noticed some interesting patterns. Generally, increasing the number of hidden units led to improved performance and accuracy. However, there was a notable exception when the number of hidden units was set to 10.

With a smaller number of hidden units (2 or 10), the neural network seemed to struggle in capturing complex patterns. The model's capacity appeared to be limited, resulting in lower accuracy and difficulty in learning the underlying patterns in the dataset.

On the other hand, when using a moderate number of hidden units (50 or 100), the network's capacity improved significantly. It became more adept at capturing complex features and patterns, leading to better overall performance and increased accuracy.

However, as the number of hidden units increased substantially (e.g., 1000 or 10000), the network's performance showed diminishing returns. While the model had an even greater

capacity to learn intricate relationships, it became more prone to overfitting. Overfitting occurs when the model becomes too specialized in the training data and fails to generalize well to new, unseen examples.