**Logo

Description automatically generated**

**San Francisco Bay University**

**CS483 - Fundamentals of Artificial Intelligence**

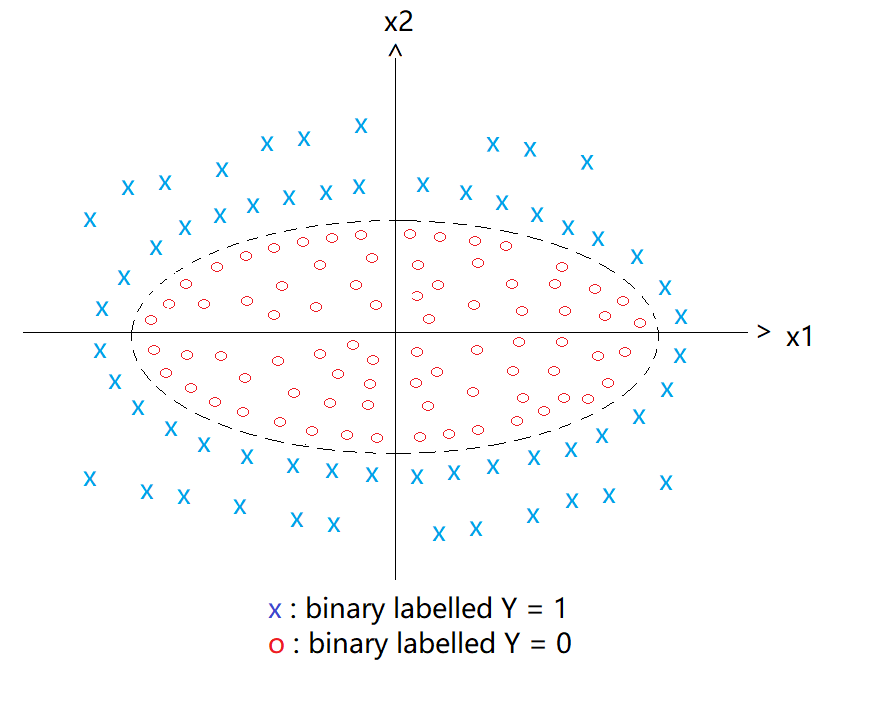
**2022 Summer Quiz #1**

**Student Name: Student ID:**

**Instruction:**

1. **Put your answer right after each question in the answer sheet**

1. Assuming two-features (and ) binary labelled Y dataset and the boundary line of classification shown as follows, try to figure out the hypothesis function, loss function and cost function for the training model and explain why



2. Given 1,000 samples dataset with more features and labelled Y’s values, 80% training set and 20% validation set will be taken as splitting in data preprocessing. If the training model takes KNN algorithm, what is K’s value to create the overfitting? Explain why, and how to avoid it. Based on the rule of thumb, what is the appropriate K’s value to make error rate minimum?