Exercise: Conditional mutual information and naive Bayes classifiers

Consider a naive Bayes classifier with two features, X_1 and X_2. What is the conditional mutual information between the features conditioned on Y, defined below.

$$I(X_1; X_2|Y) = KL(p(X_1, X_2|Y)||p(X_1|Y)p(X_2|Y)) = H(X_1|Y) - H(X_1|X_2, Y) \qquad (1)$$

State any assumptions you need to make to solve this problem.