Remember that in lecture, we talked about the statistical learning framework, where training, validation and test data are all independent samples drawn from the same underlying data distribution. State whether the statistical learning framework assumption applies in the following cases. In each case, if your answer is yes, explain why. If your answer is no, explain what is different between training and test data – \( \mu \) (the marginal over \( x \)), \( \eta \) (the conditional distribution of \( y|x \)), or something else – and justify your answer.

(1) (5 points) Alice wants to build a classifier that detects if a patient’s chest X-ray shows tuberculosis. She gathers a large training data set by collecting all patient chest X-rays from the UCSD hospital, and getting UCSD doctors to label them; this data is then used to build a classifier. The classifier is tested in the Mayo Clinic in New York.

(2) (5 Points) Bob wants to build a classifier that can predict whether a candidate would be a good engineer in his company. For this, he collects historical data of all applicants for engineers, along with who was hired; this is used to build a classifier, which is then tested on future job applications.