When we build the model produced by these three algorithms, we get some idea about what attributes might affect the death rate. We pick gender, age, max and min temperature, total snow and rain mm, and city as our features. Because gender and age will affect individual body fitness, temperature and mm also has a great influence on the body, city decide the medical level. I think it is reasonable.

But after we built the model (after undersampling), we find all the three algorithms have the similar result, low precision, a little high recall and low accuracy. We know in this model:

TP means the machine think the individual is fatal, and it is.
FP means the machine think the individual is fatal, but it is not.
TN means the machine think the individual is not fatal, but it is fatal.
FN means the machine think the individual is not fatal, and it is.

Precision =TP/(TP+FP). Low precision means our FP is too large. We know that we have 20643 not fatal cases and only 288 fatal cases. Therefore, it means in this model, we cannot predict if the cases if fatal based on the features we have under the samples we have. It might be we get some noisy or unnecessary features influence our prediction.

Recall = TP/(TP+FN). Our recall is 66%-80% which is a little high. Therefore, the FN should be lower than TP, almost 3 times. Like we pick 75%. TP/(TP+FN)=3/4= 3/(3+1). We know TP should not very large since we only have 288 fatal cases. So based on our feature, the model think in majority of cases, the individual is fatal. That is the same as what FP shown. That means the condition of fatal cases are more changeable than not fatal cases.

We can know after undersampling, the cases of fatal do not have a clear standard, at least for the features we pick, we cannot predict the death rate based on them. Although we know the death of Covid19 related to age, but with a small samples, this relationship is not that important, the death rate might be more complex than it seem like to be.