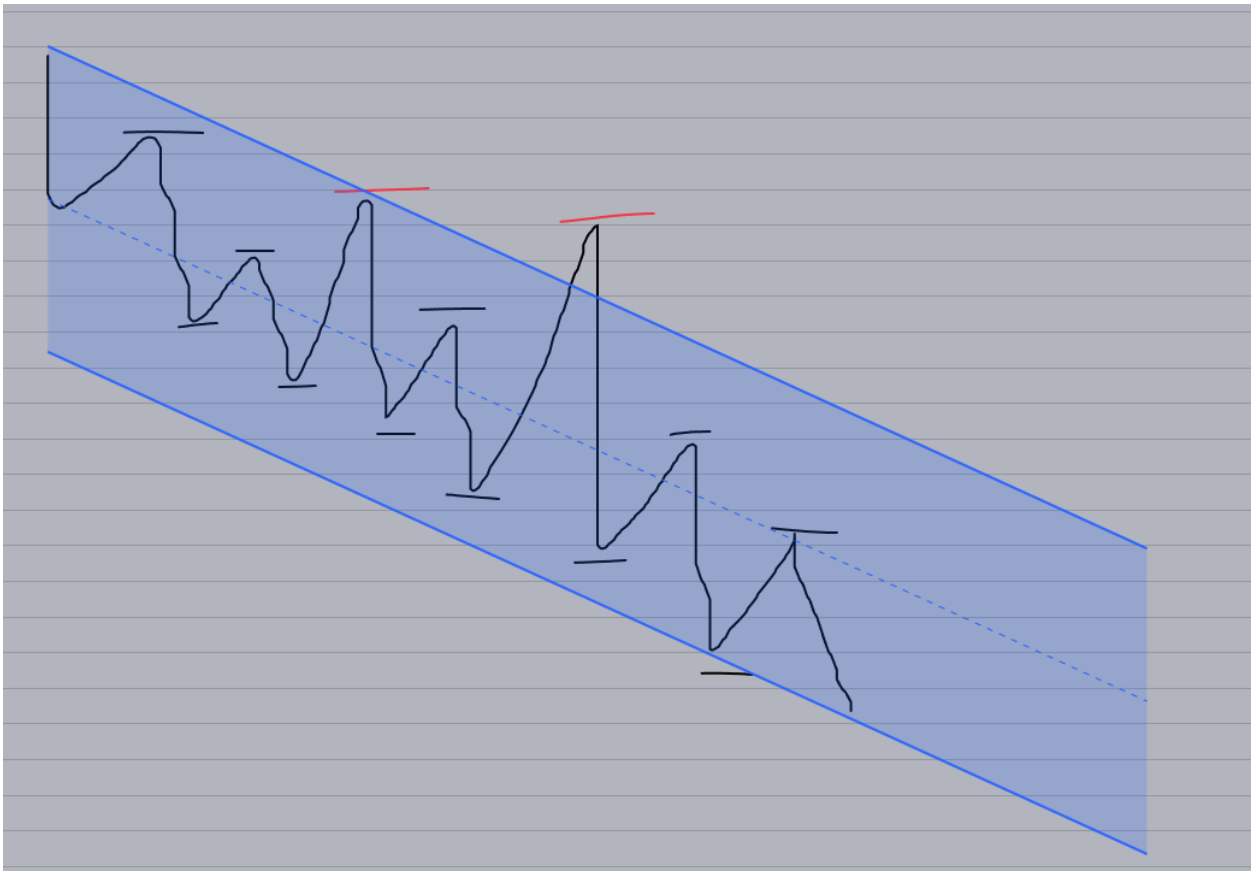
1. No, stopping mini-batch gradient descent immediately when the testing error increases is not a smart idea. This is because it is usual in machine learning for the testing error to vary briefly throughout the training process. The error may grow slightly before dropping again. It is preferable to watch the general trend in mistake over time and make judgements based on that trend rather than a single rise in error.



This picture shows how the mini batch gradient descent algorithm works some times there are little hiccups that make it seem like the error is going up however, over all the trend is still going down. If we were to stop it early we could miss out on finding out the best (smallest) error.

2.