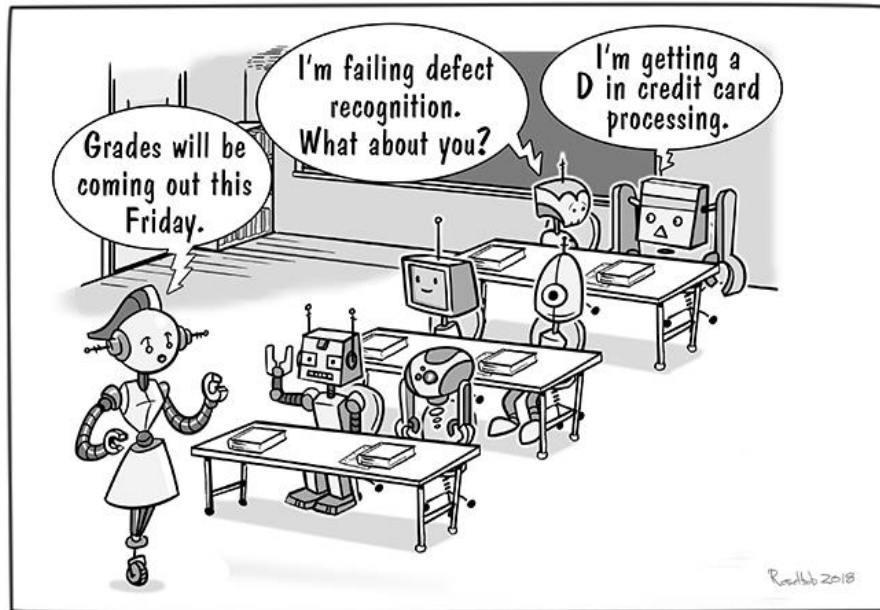


## Why can't DI/ML models learn?

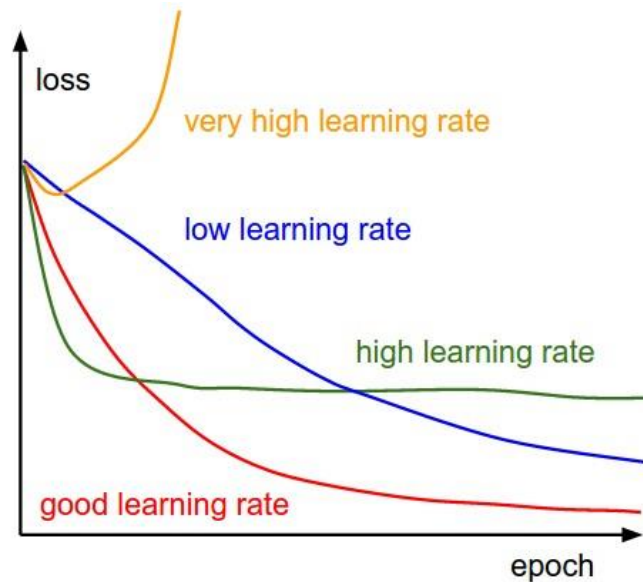
Unfortunately, when developing deep learning and machine learning models, we are faced with a problem such as the inability of models to learn. These problems sometimes have an easy solution, while some of them are problems that cannot be solved. Below I have compiled brief information about the examples of the most common of these problems in the form of articles, pleasant information 😊



**1)** A problem that we will never want to face (applies to all neural networks) is that the task assigned to the neural network is at a level that cannot be performed (perhaps partial accuracy can be achieved, but the desired success cannot be achieved). Unfortunately, solving this problem will not be easy at all. What needs to be done is to analyze the data used for the task and check whether we can solve it ourselves. We also need to realize that neither machine learning nor deep learning are magic wands and they cannot complete every task successfully.



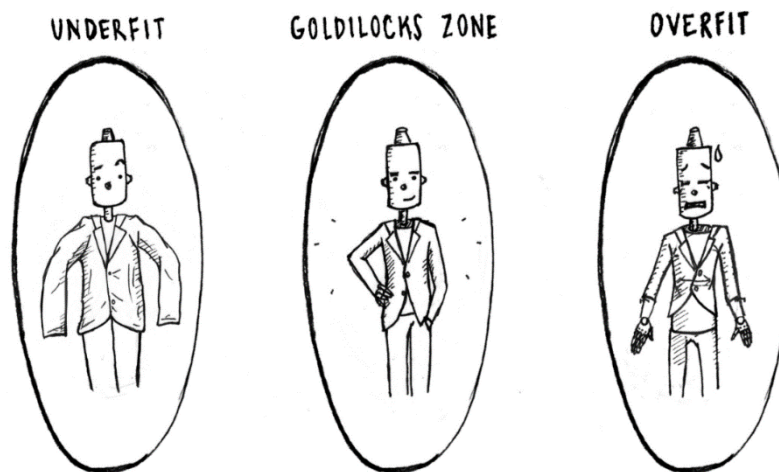
2) One of the errors that we can consider very simply is that the learning rate has either a very large or a very small value. In such cases, it is difficult to approach the most correct point when updating weights. The learning rate value used as a standard is 0.001.



3) Unbalanced data set is one of the reasons leading to this problem. For example, let there be a classification of data included in classes A and B. However, 95% of the data set belongs to class A, while only 5% belongs to class B. This situation is a major obstacle to the learning of the model. Although structures and algorithms are important when developing a model, a balanced data set is at the top of the order of importance.



4) The model structure we have established can be very large or very small. If it is small, we encounter problems with underfitting (inability to learn), if it is large, we encounter problems with overfitting (excessive learning status). In both cases, the generalization we want to achieve (the more generalized our model is, the more successful it is in validation and test data. I think it would not be wrong to think of generalization as a model that adapts to all conditions) which leads us to move away from the situation.



Because of these reasons, we need to set up a structure suitable for the dataset we have.

**Do not be afraid to try it!**