Anomaly Detection

  Romeo Kienzler

In this lecture we want to talk about anomaly detection using neural networks.

We want to do this in a unsupervised manner

Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

So remember, in supervised learning we have our data..

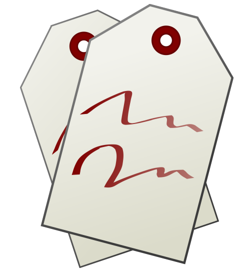


Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

..but each item in our data set needs to be assigned with a label, either a class or a continuous values.

Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

We call this target value which we want to predict.

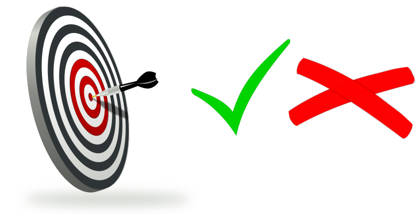


Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

In the case of anomaly detection this can be a binary target indicating an anomaly or not

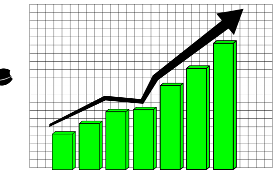


Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

Or a continuous value - so an anomaly score or R-U-L - value for example. R-U-L stands for remaining useful lifetime

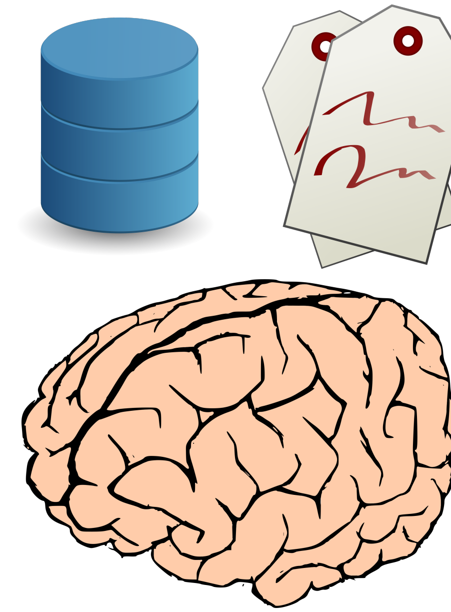
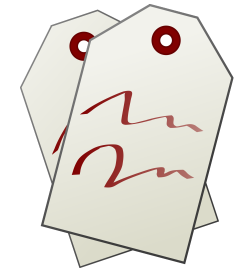


Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

Given that labeled data a model is trained



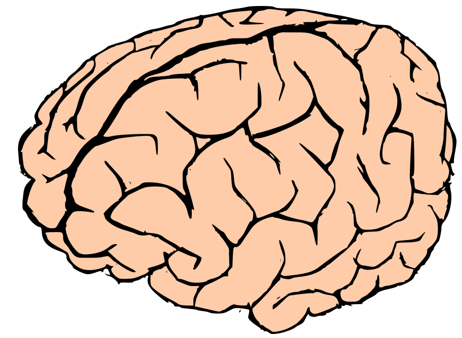


Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

- in this case - a supervised mode since we tell the model to replicate the hidden functions underlying the labeled dataset.

Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

since there is far more unlabelled data available

Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

than labeled……..

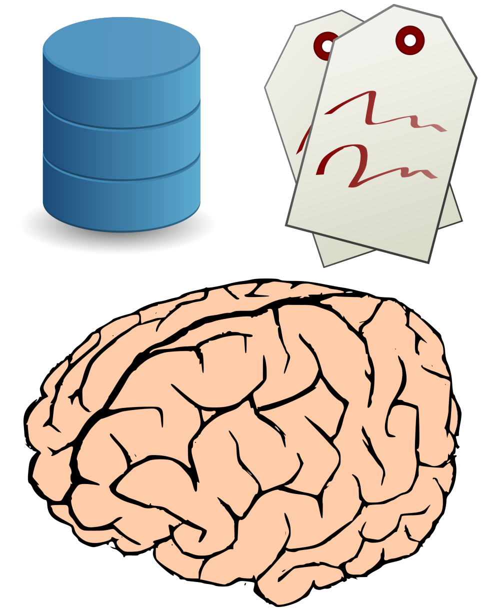
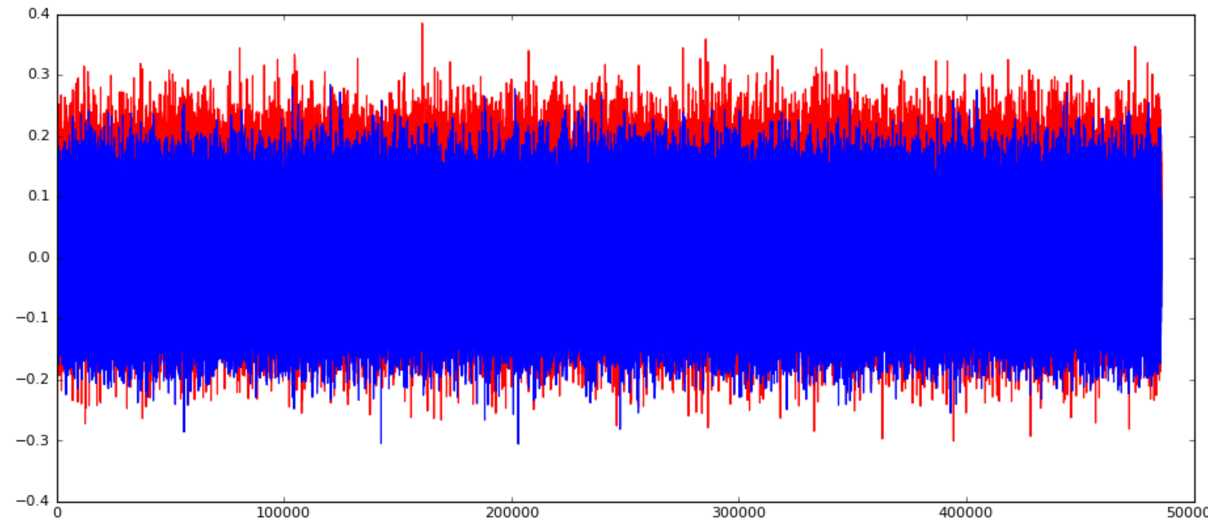


Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

IBM Watson IoT™ Romeo Kienzler

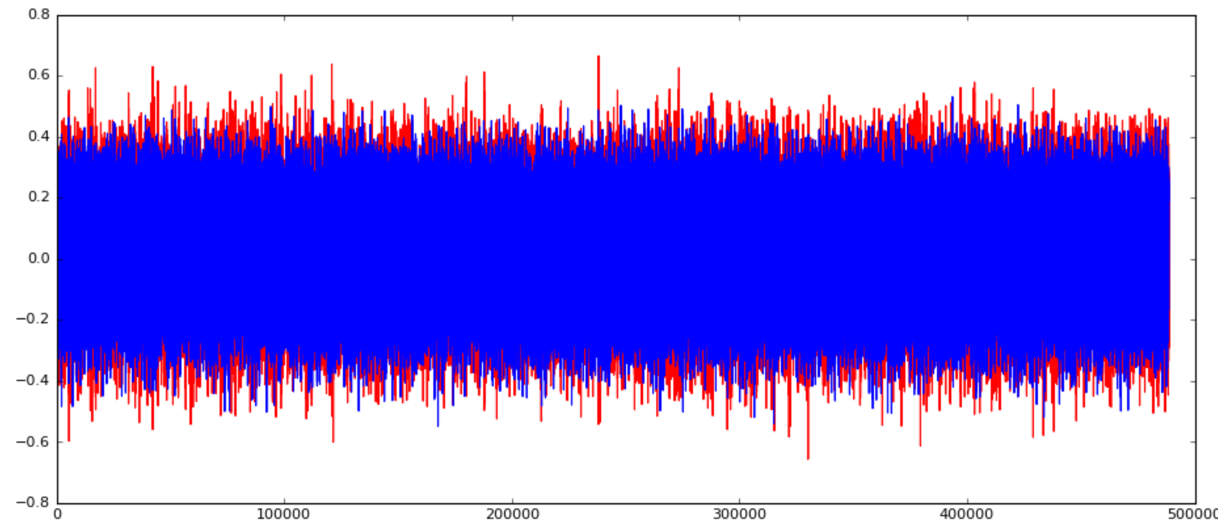
therefore we want to get rid of the boss



Data Credit:
http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website

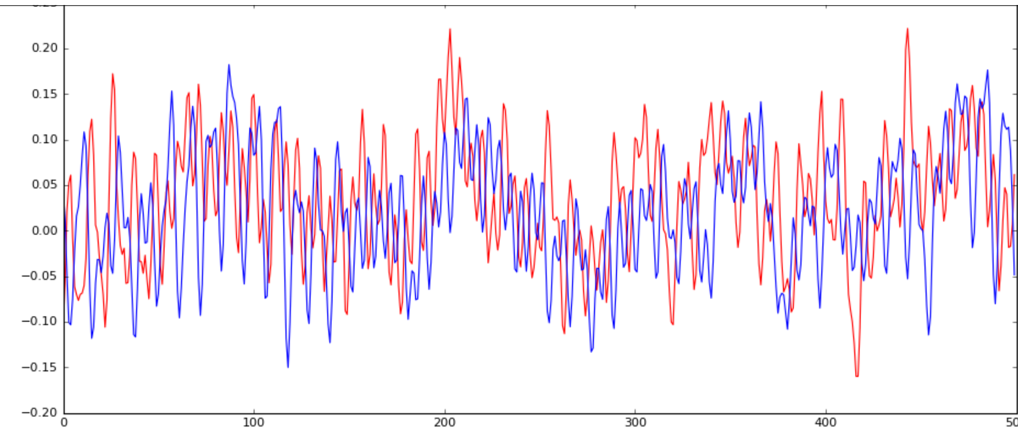
IBM Watson IoT™ Romeo Kienzler

So it turns out that this is not a trivial task. So look at this time series from two accelerometer based vibration sensors attached to a bearing. One vibration axis is shown in red, the other one in blue. So how can we predict if the bearing is in a healthy or faulty condition? In this case I’m telling you that the bearing is totally fine. So let’s have a look at data from a faulty bearing.



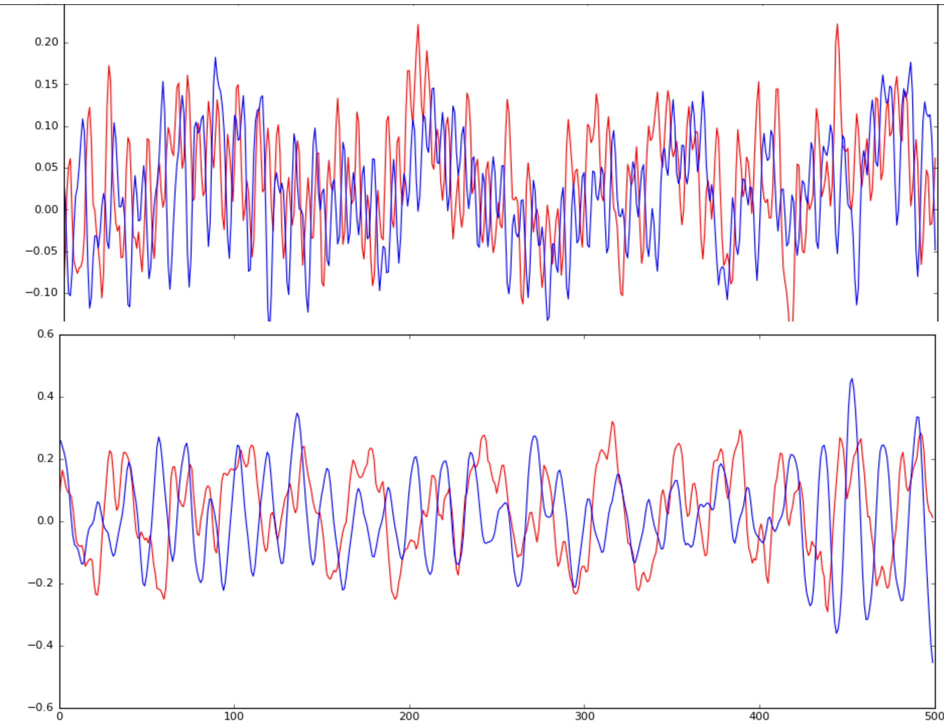
Data Credit:
http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website

Hard to tell, isn't it? …. :) … so let's zoom in a bit and check again….



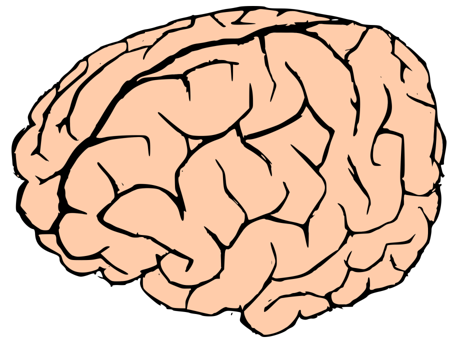
Data Credit:
http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website

So on the top we see healthy data…

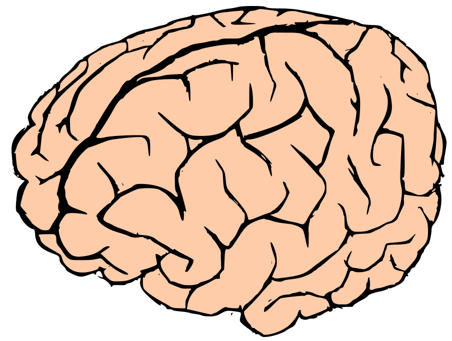


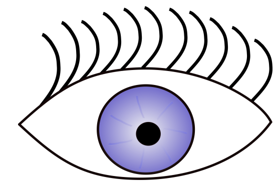
Data Credit:
http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website

…and on the bottom faulty one…ok…at least we can tell that they are different. We are using our biological neural network in order to see this. So what if we use an artificial neural network for doing the same?

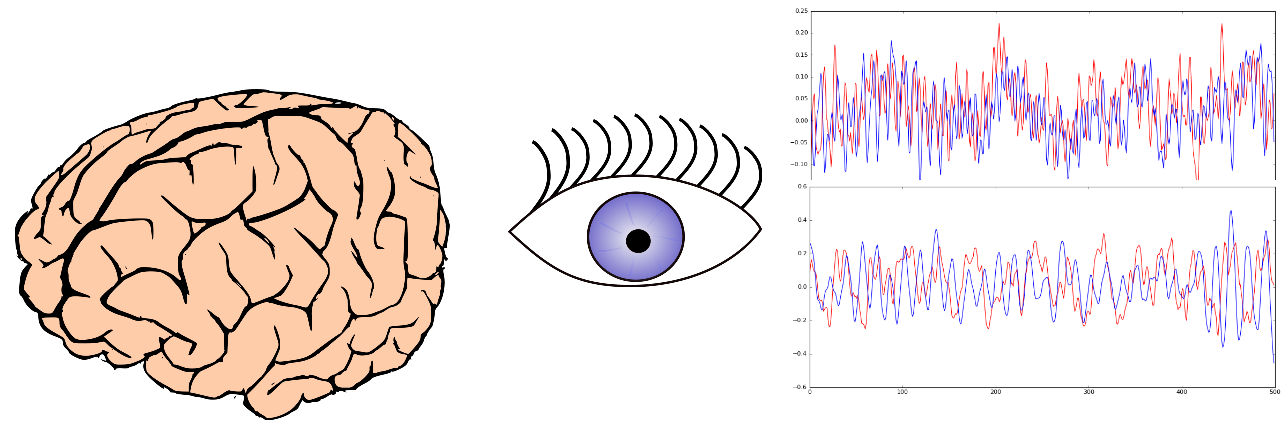


So if your brain…





looks



at two different time series plots it might be able to spot the difference. So in theory we would implement a neural network which does exactly the same. A convolution neural network for example can look at the two different figures and tell us weather if finds an anomaly or not. To be honest, I haven’t tried that, so I leave this up to you as an exercise :)

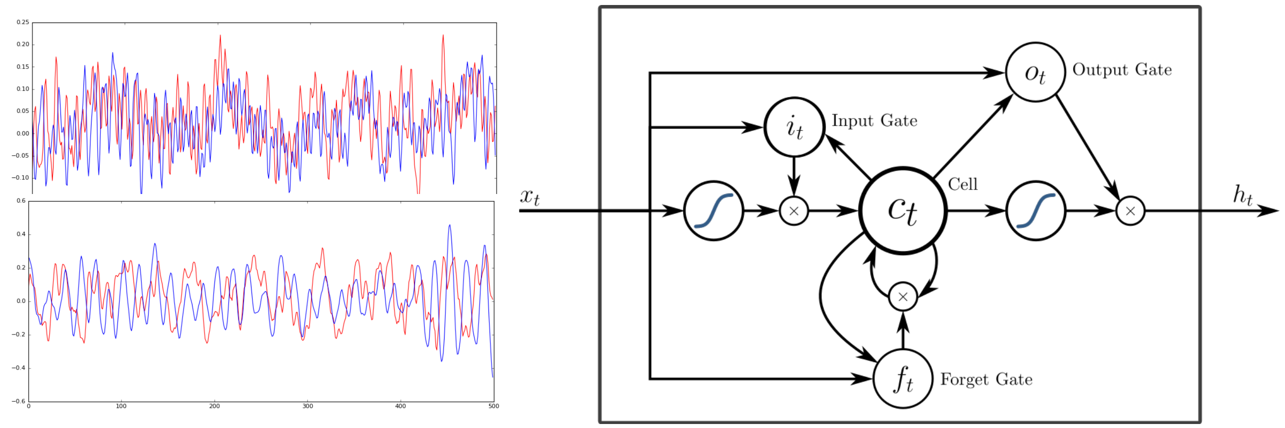
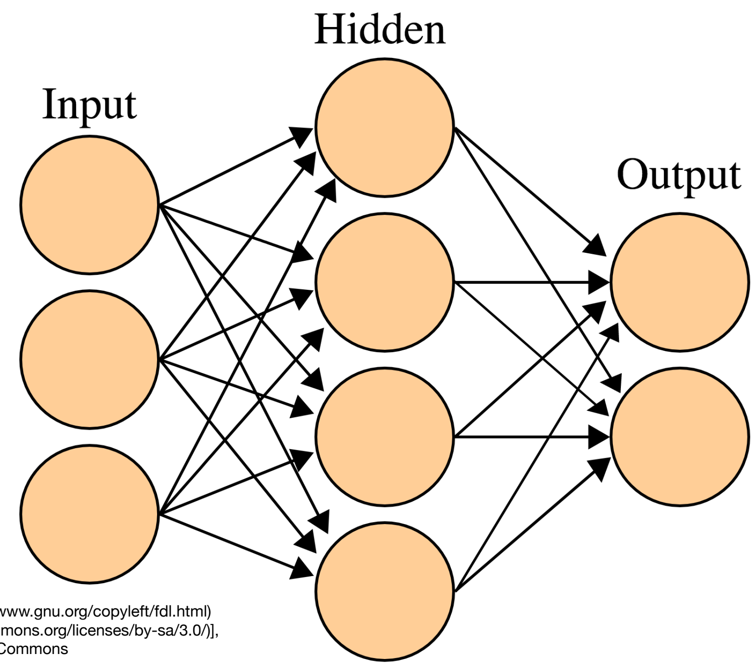


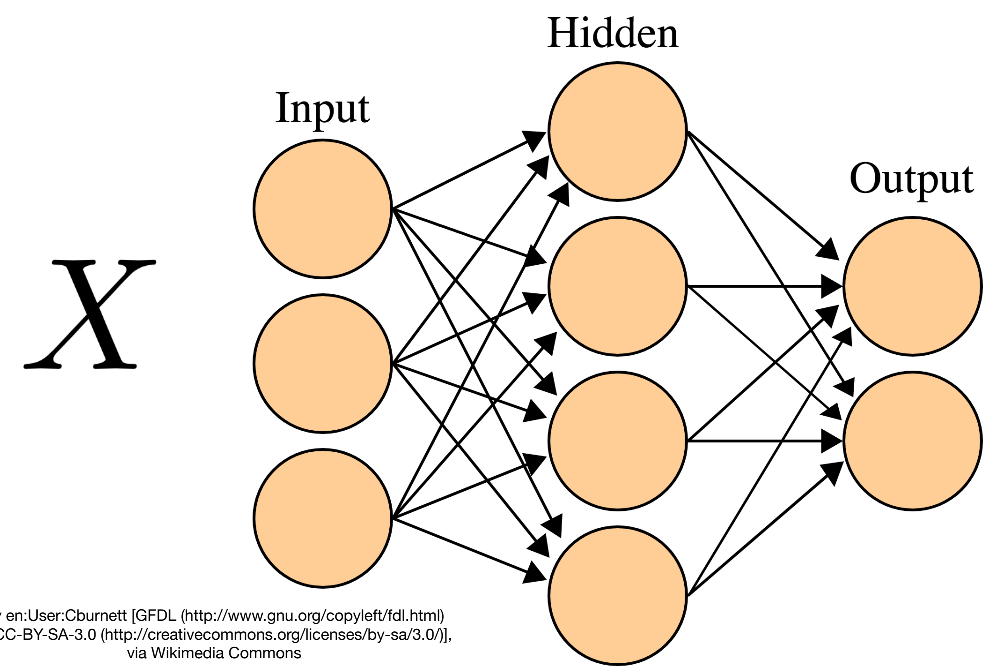
Image Credit:
By https://commons.wikimedia.org/wiki/User:BiObserve (Raster version previously uploaded to Wikimedia) Alex Graves, Abdel-rahman Mohamed, and Geoffrey Hinton (original) Eddie Antonio Santos (SVG version with TeX math) [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons
Data Credit:
http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website

But what we can do is using a LSTM - a Long Short Term Memory network for analysis. Remember, a LSTM is touring complete and ideally fit's time series and sequence data.



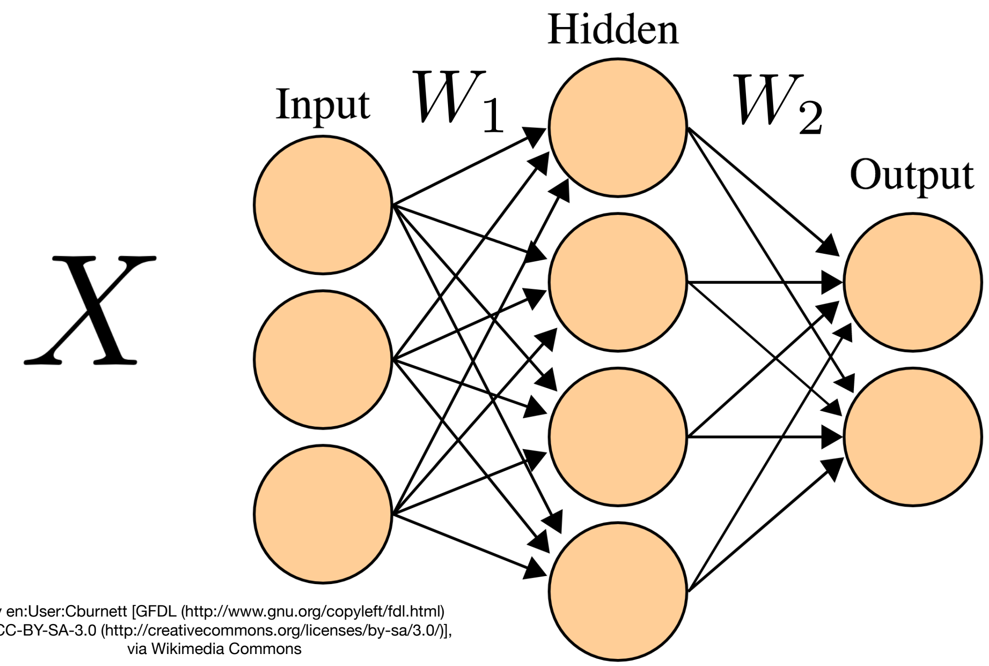
By en:User:Cburnett [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/)], via Wikimedia Commons

But isn’t there a problem? Since this is unsupervised machine learning we have no way in training this neural network because we need a label



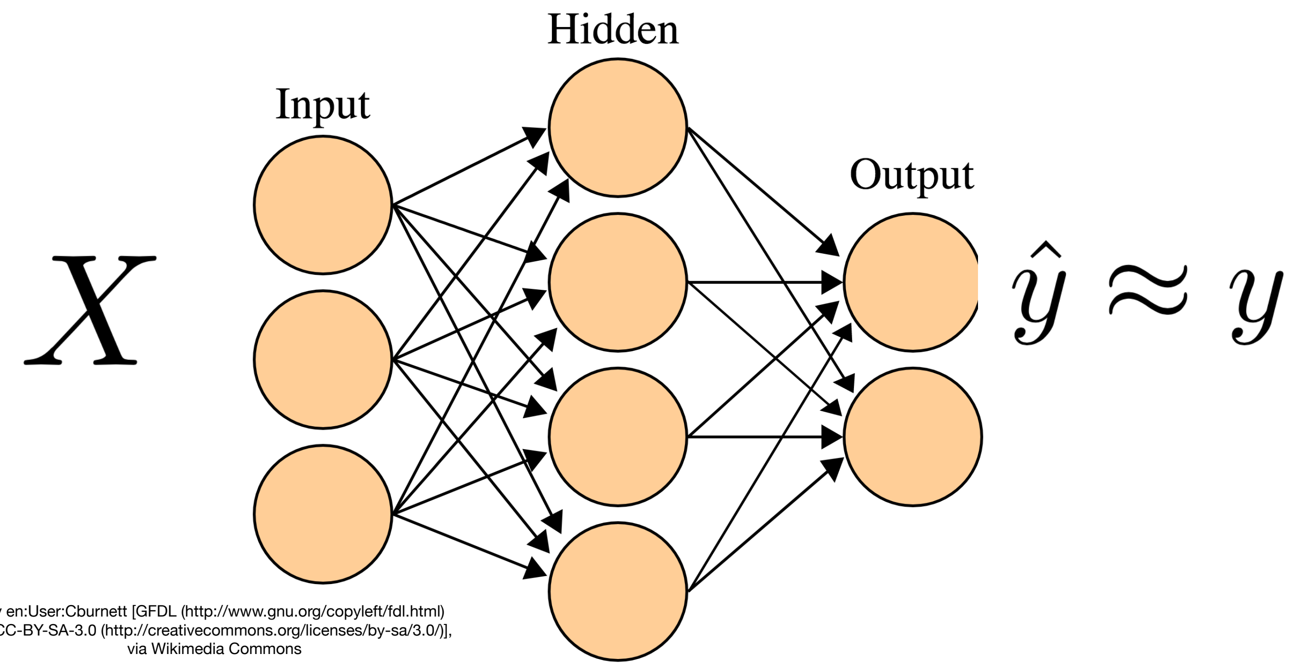
By en:User:Cburnett [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/)], via Wikimedia Commons

So the task of the neural network is to take X…

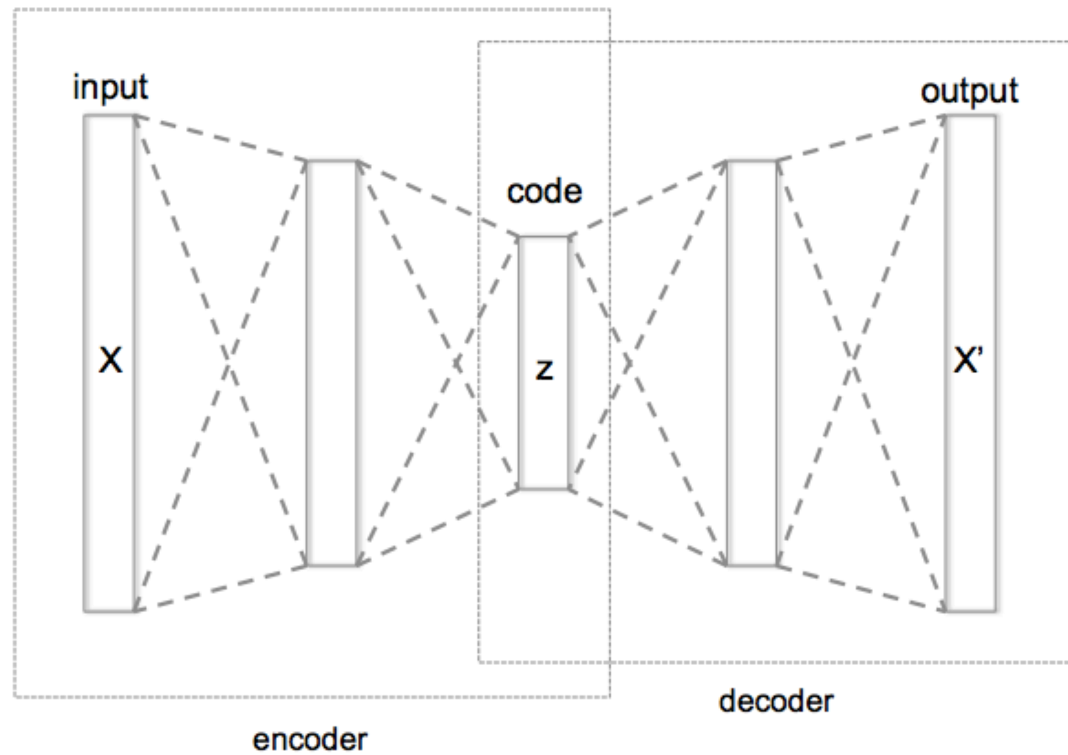


By en:User:Cburnett [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/)], via Wikimedia Commons

and adjust the weights..so that it



reconstructs y.



But hopefully you remember what an auto encoder does. So an auto encoder is trained by using X as input and as output as well. So an auto encoder will try to reconstruct it's input at the output side. And if we now train an LSTM auto encoder with healthy time series data it will learn to reconstruct healthy data but will have a hard time to reconstruct faulty data.

Image Credits openclipart.org:
by doodleguy

But how do we know if we are training the model with healthy or faulty data? So this is quite simple.

Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

Because in a real world scenario

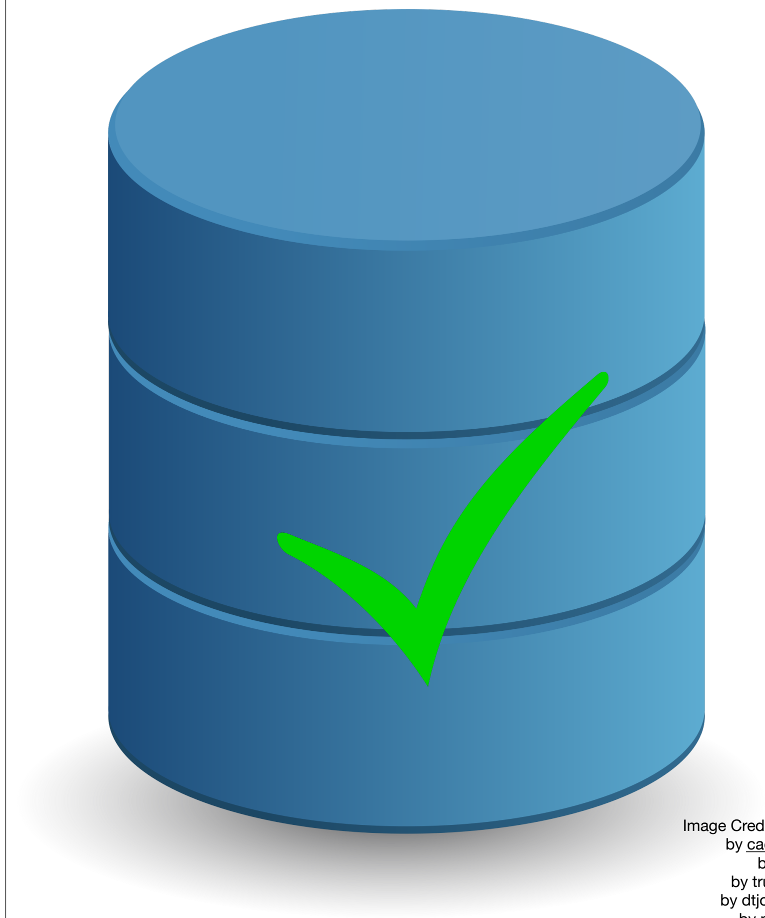
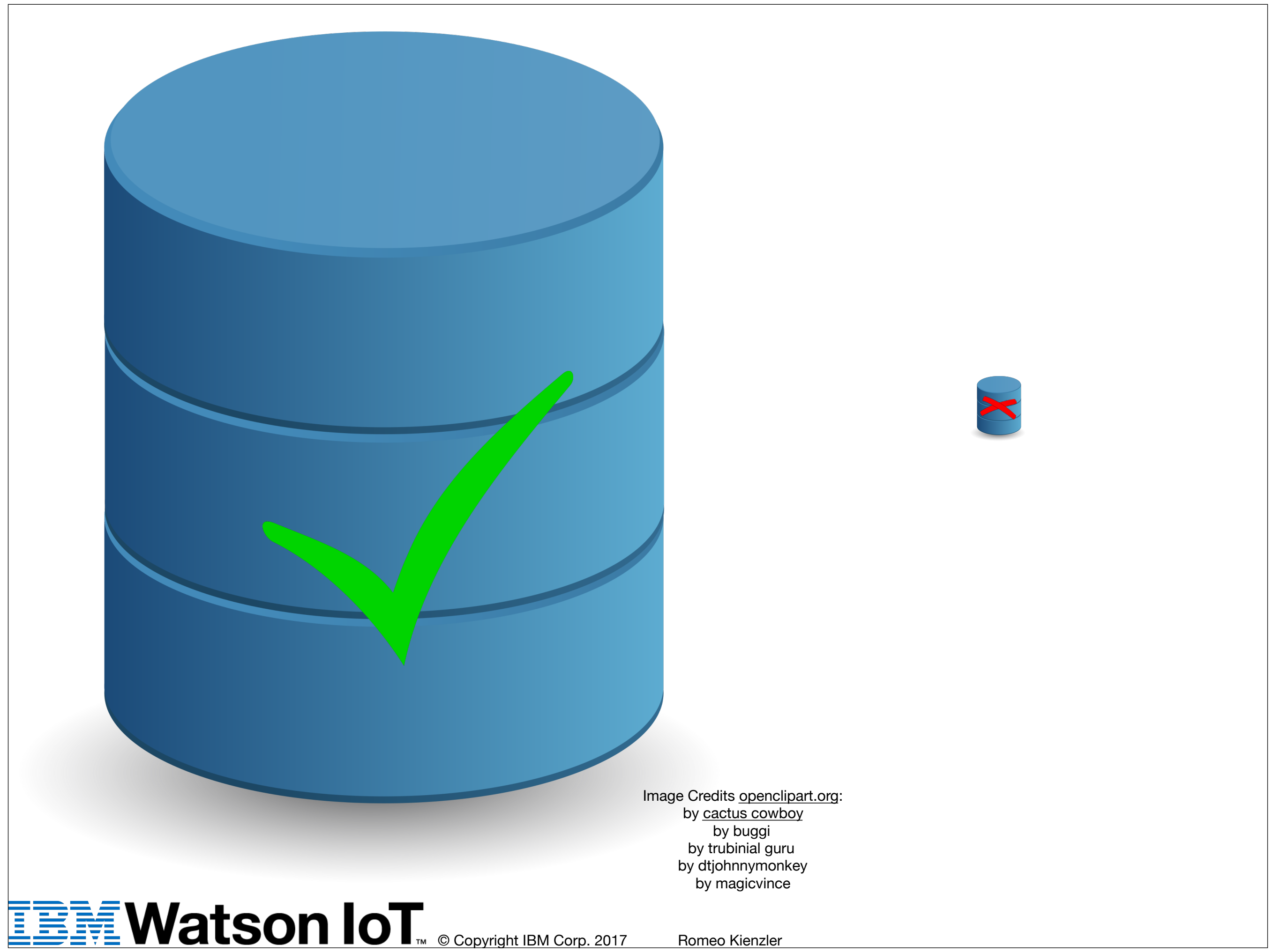


Image Credits openclipart.org:
by cactus cowboy
by buggi
by trubinial guru
by dtjohnnymonkey
by magicvince

there is much more healthy…



than broken data available for training. So if we randomly pick a sample chances are very low that we pick faulty data. There are much more strategies behind this with are often domain specific, therefore we won't go into more details here.

Implement an anomaly detector

In the next module we will implement an LSTM auto encoder based anomaly detector in Keras, so stay tuned :)