Good Morning Guys,

In this dissertation, I’m gonna talk about how to get started with Machine learning from scratch to do that I am going to work with an open source library called scikit-learn.

Scikit-learn it’s a free software library in python which is used to code machine learning, data analytics, we will see scikit-learn in a minute

Tensor-flow it’s open source library created by Google brain group it’s used for everything for example automaticatly identiftying and labeling items from youtube’s videos, improve speech recognition, improve google translate

In this point let me do a reflection about what’s happening in the software industries, before it was unthinkable that large Companies make open its tools, nowadays in this big-data world large companies can make open its tools/libraries but not its data obviously google not share with us datasets.

Before we get started to write some lines of codes let me show what’s ML and why is important.

You can think of machine learning as a subfield of Artificial Intelligent, machine learning it’s the study of algorithm that learn from examples and experiences instead of relying on hard code rules so that’s the state of the art.

Nowadays machine learning can be used for everything for all industries – the limit is our mind – please let me list some applications ….

I’ll show you a business case that is very easy but it’s impossible to solve it without machine learning.

Imagine that we have been asked to work on an application which has to identify who is the best BI member to close out one particular ticket.

We know the ticket is created based on an email sent by one particular customer.

You know when we get the ticket in our queue => the ticket has many details for example category / type of incident / severity / Impacted devices / etc etc but all those information was filling by Service desk group.

But in the beginning the ticket only has a Description, subject and the sender.

Training dataset => the last column is called label, it identifies who is the BI member which close the ticket.

The whole table is our training data => think of this as all the examples we want the classifier to learn from.

Now let’s write down our traiming data in code.

Fit => Find patterns in data.

so let's take it for a spin and use it to classify a new Ticket request.