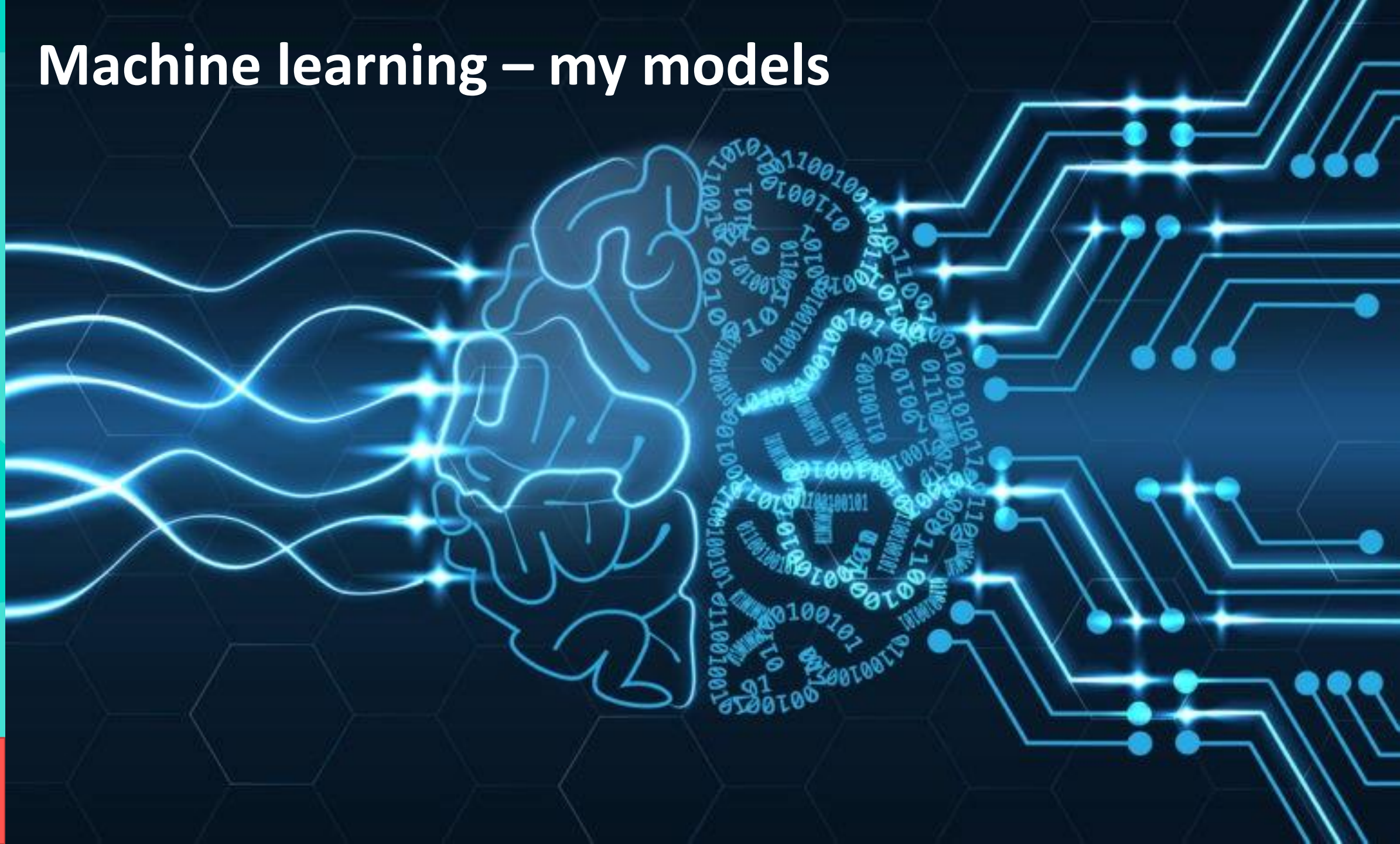


Machine learning – my models



Introduction

Here you will find three machine learning models I developed while learning the models.

I tried to use practical dataset and context to train the models and purpose was to use machine learning modes to solve 3 real business problems.

If you carefully see the models, you must understand 3 basic machine learning models and its real business use.

Linear regression

The first model is linear regression.

Here I wanted to predict the expiry return based on the sale to respective distributors. I have used 36 months of historical data including distributor and product wise sale and expiry return. Then developed the linear regression model to understand the relationship between sales and respective expiry return.

After the training with the historical data, the model was able to predict expected expiry return for specific sale to specific distributor.

Linear regression algorithm

1 of 3

Event

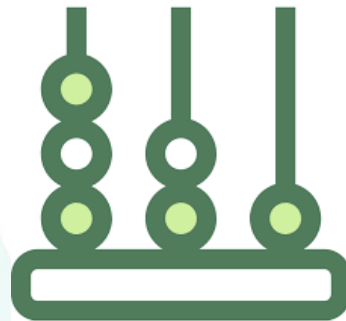


Expiry
10 pcs



Sold
1000 pcs

Data



123

Database

38 months of data

By distributor

By SKU

With volume

With amount

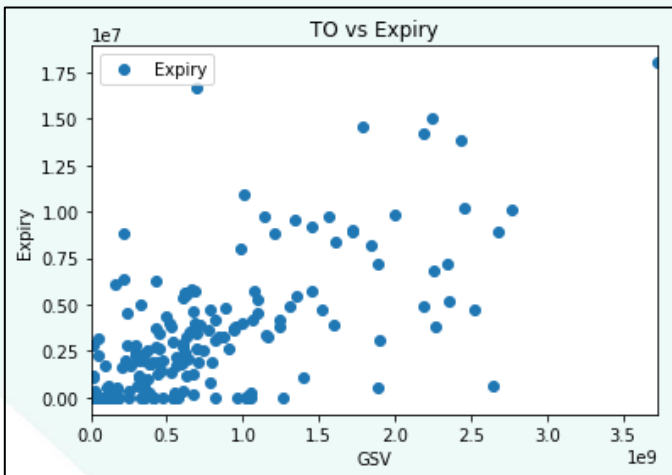
GSV

Expiry

1,000,000 + rows

Linear regression algorithm

Linear relationship



Expectation

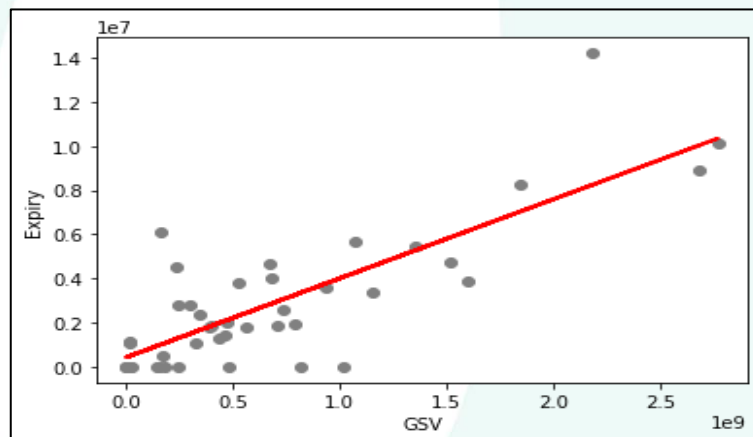
Expiry considering sale to customer

Expiry considering sale by SKUs

Model



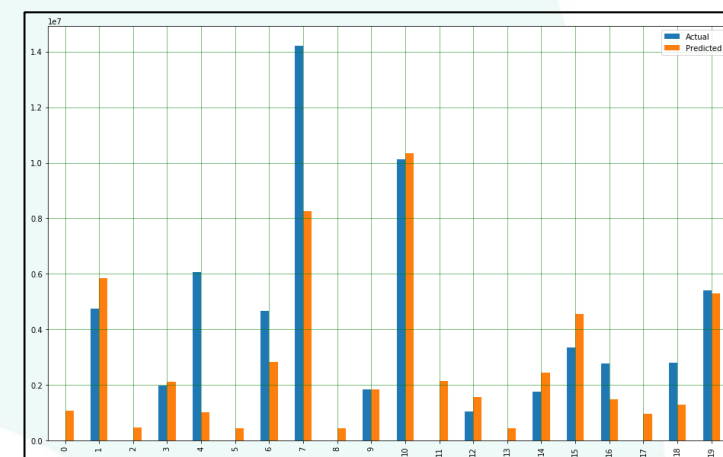
Linear regression



Prediction



Accuracy 67%



Logistic regression

The logistic regression model was used to predict whether a savings project will be successful or not.

All the projects we tried for were not successful, few projects were unsuccessful. I used the historical data of last three years projects including who was the owner, what was the deadline, what was the value, what type of project was that one etc.

With the logistic regression model, I trained the dataset, then the model was able to understand with some special set of characteristics some projects were unsuccessful.

Now the model can predict at the beginning of any project, that the project will be successful or not. Based on that prediction we can take appropriate measures to make the project successful.

Logistic regression algorithm

2 of 3

Event



200 to 250
VIP projects



200 to 200
Projects dropped



30 to 50
Projects dropped

Data

Owner

Deadline

Value

Risk factor

Cash/non-cash

Network

Calculation

SRS number

RM/PM

Buying/R&D

Category

Status

Database

3 years of savings data

12 parameters

Result

643 projects

Logistic regression algorithm

2 of 3

Question/ expectation

Predict which project will
be implemented

Predict which project will
dropped

Machine learning model



Logistic regression



Prediction



Accuracy 77%

K-means clustering

The K-means clustering model helped me to predict efficient transport route to deliver products from factory to depot.

Several routes were used to transport goods from different factories to different depots and customers. I wanted to classify which route were efficient and which were not considering cost per ton, load-fill by weight and load-fill by volume.

I trained the model with four months of transport data. Then identified some routes were costly and inefficient.

Based on the result of this model, I remapped few routes to get efficiency.

K-means clustering algorithm

3 of 3

Event

Logistic lane performance

Some lanes high CPT

Some lanes low CPT

Some lanes high load-fill

Some lanes low load-fill

Data

CPT

Load-fill weight

Load-fill volume

Database

**4 months lane
performance**

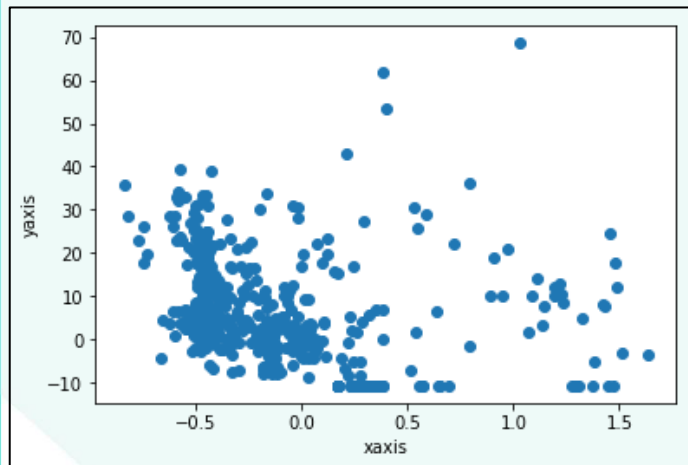
713 lanes

12 columns

3 parameters

K-means clustering algorithm

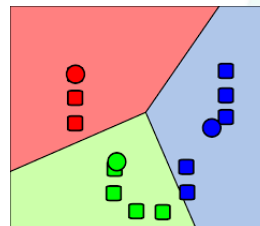
Data shape



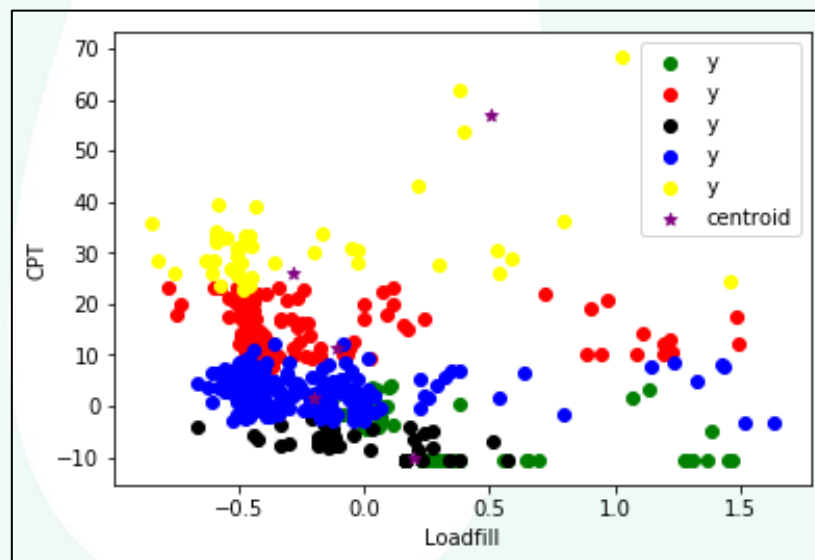
Expectation

Classify lanes according to performance

Machine learning model



K-means clustering



Prediction



Accuracy 77%

