Analysis of our confusion-matrix and the implication in our customer retention model

There are 483 churners **Churning is declared as a positive condition**.

3333 total	Predicted negative	Predicted positive
True condition negative	TN	FP
True condition positive	FN	TP

By definition a confusion matrix C is such that Cij is equal to the number of observations known to be in group i but predicted to be in group j. Thus in binary classification, the count of true negatives is C[0,0] false negatives is C[1,0], true positives is C[1,1] and false positives is C[0,1].

- TP: are well predicted churners so as they are going to leave investing on them is necessary
- FP: are customers badly predicted as churners so expending will be unnecessary and an overcost
- TN: are customers that will keep loyal to our company
- FN: are non predicted churners we need to spend money on them as they are future churners

Results of our classifier compared to sklearn SVM implementation trained, as our classifier, without sklearn cross validation.

Sigma 0.1 epochs 100 - positive is churning

SGD: 0.344 accuracy	1263 TN	1587 FP	SKLearn	2675	175
	143 FN	340 TP		322	161

Sigma 0.1 epochs 1000 - positive is churning

SGD: 0.796 accuracy	2844 TN	6 FP	SKLearn	2806	44

	464 FN	19 TP		413	70

Sigma 0.1 epochs 5000 - positive is churning

SGD: 0.796 accuracy	2785 TN	65 FP	SKLearn	2849	1
	423 FN	60 TP		413	70

The goal of our classifier is to, given some data of our customers, to predict future leaving and modeling a tool that can help marketing department to design accurate retention campaigns. It's important to clarify that our classifier is not exactly precise but we measure two important measurements. With sigma 0.1 and 1000 epochs

- **Accuracy**: how often is the classifier correct?
    - (TP+TN)/total = (19+2844)/3333 = 0.85
- **Error Rate**: Overall, how often is it wrong?
    - (FP+FN)/total = (6+464)/3333= 0.14 (1-accuracy)
- **Recall**: When it's actually churn, how often does it predict churn?
    - TP/actual yes = TP/TP+FN = 0.03
- **Precision**: How good are we predicting churners?
    - TP/TP+FP = 19/25 = 0.76
- **F1 Score**: 2 * (precision * recall) / (precision + recall)
    - 2*(0.76*0.03)/(0.76+0.03) = 0.05

Despite this results, it can gives you an intuition about how much you should expense in successful retention campaigns. We have all our future churners in TP + FN and we will propose the following cost prediction function:

$$-\alpha * (TP + FP) + 100 * \beta * TP > 0$$

α *is the cost expended for retaining a customer*

β *is fraction of churn recovered*

TP *are precisely churn predicted customers*

FN *are bad churn predicted customers*

$$-\alpha * (19 + 6) + 100 * \beta * 19 > 0$$

$$-25\alpha + 100 * \beta * 19$$
$$\alpha < (100 * 19/25) * \beta$$
$$\alpha < 76 * \beta$$