# LAB – Week 13

## 1/ e-Reader

Download the file online\_retailer\_dataset\_training.csv and online\_retailer\_dataset\_scoring.csv on the UV.

Objectives:

A marketing analyst would like to figure out which customers he could expect to buy the new eReader and on what time schedule, based on the company’s last release of a high-profile digital reader. He would like to segment customers in order to provide them the appropriate advertising.

Knowing that there are four possible values of eReader\_Adoption:

[1]Innovator: purchase within the 1st week

[2]Early Adopter: purchase within 2 or 3 weeks

[3]Early Majority: purchase > 3weeks but <= 2 months

[4]Late Majority: purchase after the first two month

The criteria and suggestions for segmenting customers are as follows:

[1] Those who are most likely to purchase immediately (predicted innovators) can be contacted and encouraged to go ahead and buy as soon as the new product comes out. They may even want the option to pre-order the new device. On the other hand, perhaps very little marketing is needed to the predicted innovators, since they are predicted to be the most likely to buy the eReader in the first place. [2] Those who are probably likely to purchase earlier (often are opinion leaders, predicted early adopter)

[3] Those who are less likely (predicted early majority) might need some persuasion, perhaps a free digital book or two with eReader purchase or a discount on digital music playable on the new eReader. [4] The least likely (predicted late majority), can be marketed to passively, or perhaps not at all if marketing budgets are tight and those dollars need to be spent incentivizing the most likely customers to buy

## 2/ Breast Cancer detection

Download the “Breast\_cancer.csv” file. A hospital is asking you to provide them with a binary classifier to help the doctors detect if a breast cancer is metastatic (M) or benign (B), based on many criteria that are analyzed on pictures (texture, smoothness, compactness, symmetry, etc.). If the classifier predicts a cancer as metastatic, then deeper analyses will be done. If the classifier predicts a cancer as benign, there will organize a check-up meeting 3 months later. It is worth noting that the hospital prefers to proceed to deeper analyzes if it is predicted as metastatic, even though it might be benign (wrong prediction), than the opposite wrong prediction which to predict the cancer as benign whereas it metastatic. Indeed, it could cost the life of the patient and also cost more to the hospital and patient if they wait too long to treat the cancer. Moreover, it appears that the evaluation of the texture is not very accurately guessed with image analyze. Therefore, they ask you to verify if it is an important factor to take into account while predicting a diagnosis.