# Clase 2:

## Modelos de Machine learning:

* Supervisados: Tienen etiqueta, es decir, puede ver si tiene algo es o no es viendo los datos de la bd. Tienen x|y siendo la y la etiqueta.
  + Clasificación: Se usan variables discretas
    - Regresion logística
    - Naive Bayes
    - Arboles de decision
    - Random forest
    - KNN
    - Super vector machine
  + Regresion: Se usan variables continuas.
    - Regresion Lineal
    - Arboles de decision
    - Random forest
    - KNN
    - Super vector machine
* No supervisados: No tiene etiqueta. Solo tiene la x en lo que es x|y
  + Clustering: Se ve cuánto se parecen grupos y en qué se diferencian también.
  + Deteccion de anomalias
  + Reduccion dimensionalidades
* Modelos de refuerzo : No lo veremos en el semestre.

## Metricas de clasificacion

* Metrica continua
* Recall
* Precision
* Accuracy
* F1 score

## Metodologías para llevar proyectos adelante

* SEMMA
* KDD
* CRISP - DM
  + Contexto de negocio
  + Analisis de datos
  + Preparacion de datos
  + Modelamiento
  + Evaluacion
  + Despliegue

# Clase 3

Los hiperparametros son los parámetros de detalles específicos del proceso de aprendizaje

En super vector machine **si o si** hay que normalizar los datos. Estandarizar o escalar también funciona.

Órdenes de magnitud son importantes.

Regresión lineal simple se usa R2

Si es ajustada es

Heterocedasticidad → Cuando los datos no tienen correlación.

Homocedasticidad → Cuando los datos tienen correlación.