**Tipos de Validación:**

* Interna: con todos los datos.
* Externa aleatoria: separando en datos de entrenamiento y de prueba
* Externa k-fold: es aquella donde hacemos validaciones con diferentes conjuntos de entrenamiento y prueba y sacamos los indicadores promedio de todos los resultados.
* Leave One Out: No es muy popular y es muy costosa pero se puede usar en muestras pequeñas de datos.
* Bootstrap

**¿Qué son las Neural Networks?**

* Neuronas y enlaces
* Aprendizaje automático
* Funciones de activación y pérdida
* Tipos de aprendizaje
* Elevado coste computacional
* Ventajas:
  + Adaptativas
  + Paralelizables
  + Tolerancia a fallos
  + Potencial predictivo
* Inconvenientes
  + Coste computacional
  + Caja negra
  + Menos explicativo

**XGboost**

Checar la documentación oficial para más detalles sobre este método.

**Text Mining:**

* Calidad de la información
* Encontrar texto en cualquier parte, cantidad de información
* Procesamiento numérico
* Escalabilidad del proceso, podemos analizar una conversación como una enciclopedia o tweets de algún tema.
* Proceso:
  + Obtención de la información
  + Bag of words, transformar el texto
  + Lematización – Método de machine learning
  + Análisis avanzados (traducción, detección de temáticas)
  + Modelización
* Paquete nltk en Python es útil para hacer text mining.