# Machine Learning

Unidad # 3 - Aprendizaje Supervisado Avanzado y Aprendizaje No Supervisado

CC57 – 2019-1

Profesor
Andrés Melgar

# Competencias a adquirir en la sesión

- Al finalizar la sesión el alumno comprenderá el funcionamiento del **aprendizaje inductivo**.
- Al finalizar la sesión el alumno implementará **modelos algorítmos de regresión** usando conjuntos de datos.
- Al finalizar la sesión el alumno **entenderá** el algoritmo de **regresión logística**.
- Al finalizar la sesión el alumno **aplicará** el algoritmo de **regresión logística** para obtener modelos algorítmicos.

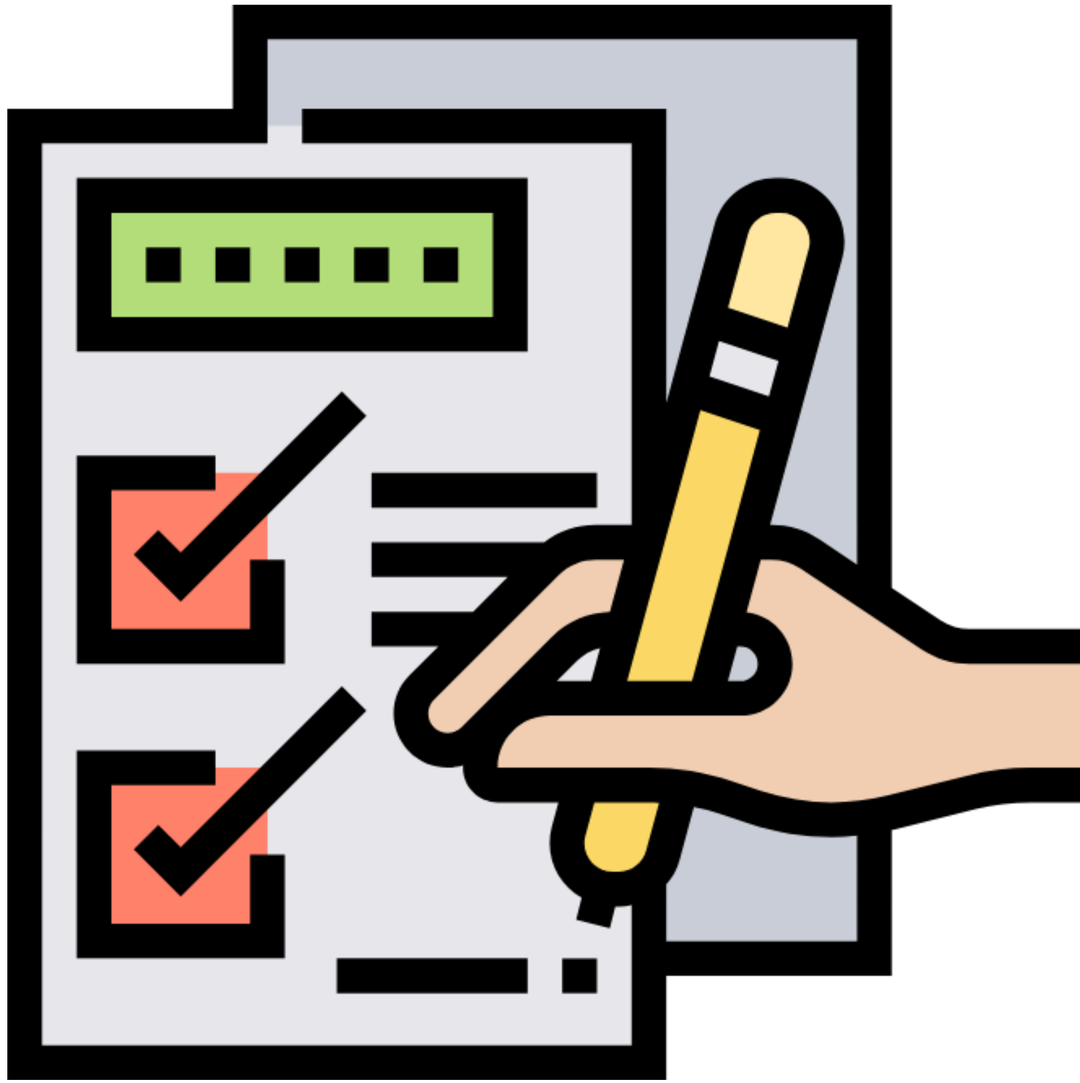


Imagen obtenida a través de la URL https://www.flaticon.com/free-icon/checklist_1467445

# Revisión de la sesión anterior

- ¿Qué **problema** resuelven los modelos lineales?
- ¿Cómo se **representan** los modelos lineales?
- ¿En qué se **fundamente** la regresión lineal?
    - ¿Qué se debe tener en cuenta debido a esto en la fase de entrenamiento?

# Métricas de Evaluación

## Texto guía

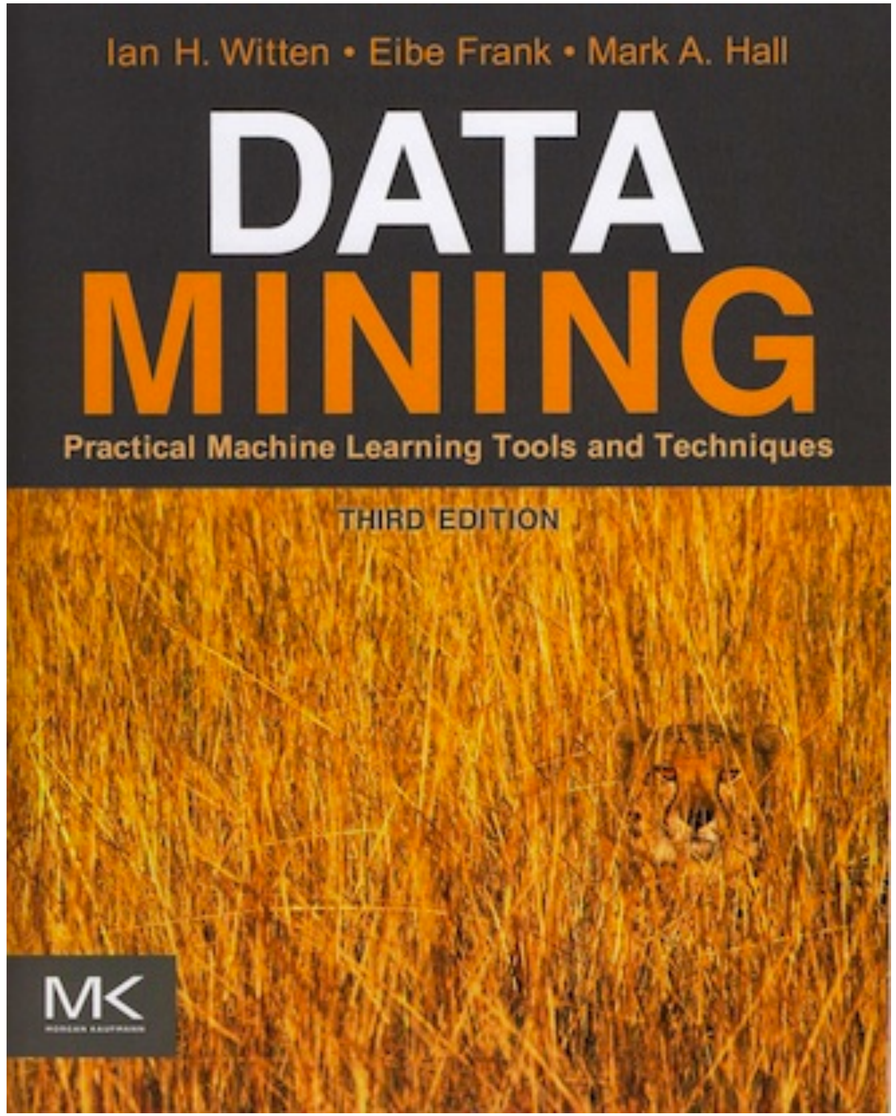
Witten, Ian H., Frank, Eibe, and Hall, Mark A.. 2011. ***Data Mining : Practical Machine Learning Tools and Techniques with Java Implementations.*** San Francisco: Elsevier Science & Technology.

CHAPTER

Algorithms: The Basic Methods

4

4.6 LINEAR MODELS



# Regresión lineal multi-respuesta

- La **regresión lineal** típicamente se ha utilizado para dar solución a **problemas de predicción** con atributos numéricos.
  - La **regresión lineal** se puede utilizar también para **problemas de clasificación** con atributos numéricos.
  - De hecho, se puede utilizar cualquier **técnica de regresión**, ya sea lineal o no lineal.
- Usar la regresión lineal para separar **dos clases** de valores es simple
  - Se usa una **única recta** para separa a las dos clases.
  - Le recta sirve como **frontera de decisión.**

# Regresión lineal multi-respuesta

- ¿Cómo separar **más de dos valores**?
  - Se lleva a cabo **una regresión para cada posible valor** de la clase, estableciendo la salida igual a 1 para las instancias que pertenecen a la clase y 0 para las que no.
  - El resultado es una **expresión lineal** para cada clase.
  - Entonces, dado una instancia desconocida, **se calcula el valor de cada expresión lineal** y se elige el de mayor valor.
  - Este esquema se llama a veces la **regresión lineal multi-respuesta.**

# Regresión lineal multi-respuesta

- Una forma de ver la regresión lineal multi-respuesta es imaginar que esta se aproxima a una **función de pertenencia numérica** para cada clase
  - La función retornará **1** en los casos que la instancia pertenezca a dicha clase
  - La función retornará **0** en los casos que la instancia no pertenezca a dicha clase
  - Dada una nueva instancia, se calcula el valor de la **función de pertenencia** para cada clase
  - Se selecciona la clase cuya función numérica retorna **mayor valor**.

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Regresión lineal multi-respuesta

- La regresión lineal multi-respuesta a menudo da buenos resultados en la práctica. Sin embargo, presenta dos **inconvenientes**:
  - En primer lugar, los valores que produce la función de membresía **no son probabilidades** ya que pueden caer fuera del rango *[0,1]*.
  - En segundo lugar, la **regresión por los mínimos cuadrados** asume que:
    - Los **errores** son estadísticamente independientes.
    - Se **distribuyen** con la misma desviación estándar.
  - Esta última suposición **no se cumple**.

# Regresión lineal multi-respuesta

- La regresión lineal multi-respuesta a menudo da buenos resultados en la práctica. Sin embargo, presenta dos **inconvenientes**:
  - En primer lugar, los valores que produce la función de membresía **no son probabilidades** ya que pueden caer fuera del rango *[0,1]*.
  - En segundo lugar, la **regresión por los mínimos cuadrados** asume que:
    - Los **errores** son estadísticamente independientes.
    - Se **distribuyen** con la misma desviación estándar.
  - Esta última suposición **no se cumple**.

# Regresión lineal multi-respuesta

- La técnica estadística llamada de **regresión logística** no sufre de estos problemas.
  - En lugar de la aproximación a los valores 0 y 1, se construye un modelo lineal basado en una variable transformada.

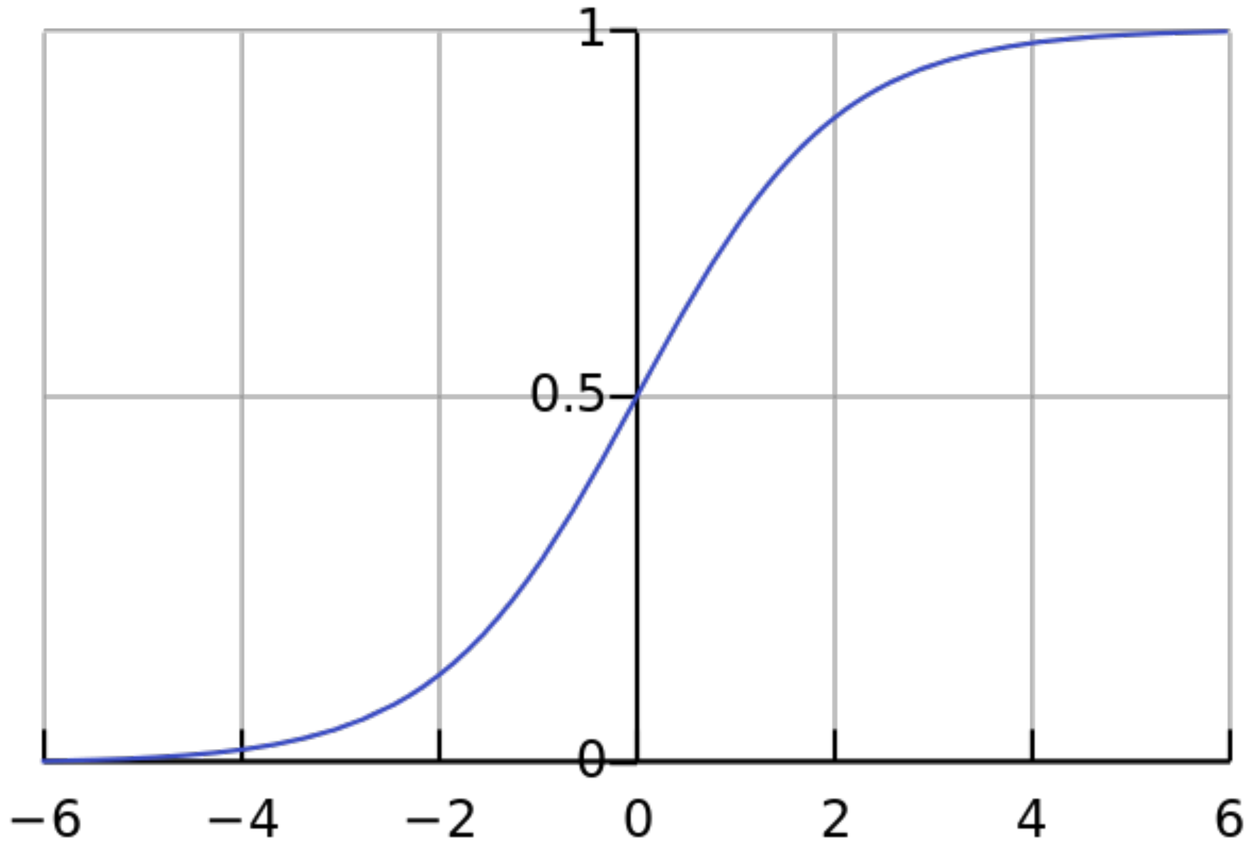



Imagen tomada de https://es.wikipedia.org/wiki/Regresi%C3%B3n_log%C3%ADstica

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Transformación Logística

- Si ***p*** es la probalidad de éxito de un evento, se define el ***ratio odds*** como el proporción de éxito en relación al fracaso: $\frac{p}{1-p}$

- Si la probabilidad de éxito posee el valor de 0.5
  - El ratio odds es 0.5/(1-0.5) = 0.5/0.5 = 1
  - Se lee un éxito por un fracaso

- Si la probabilidad de éxito posee el valor de 0.75
  - El ratio odds es 0.75/(1-0.75) = 0.75/0.25 = 3
  - Se lee tres éxitos por un fracaso

- Si la probabilidad de éxito posee el valor de 0.25
  - El ratio odds es 0.2/(1-0.2) = 0.2/0.8 = 0.25
  - Se lee 0.25 éxitos por un fracaso
  - O un éxito por cada 4 fracasos

# Transformación Logística

- Mientras mayor sea la probabilidad de éxito mayor será el *ratio odds*.

| p | ratio odds |
|---|---|
| 0.10 | 0.11 |
| 0.30 | 0.43 |
| 0.50 | 1.00 |
| 0.70 | 2.30 |
| 0.90 | 9.00 |
| 0.99 | 99.00 |
| 0.999 | 999.00 |

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Transformación Logística

- La función ***logit*** se define como el logaritmo natural del *ratio odds*.

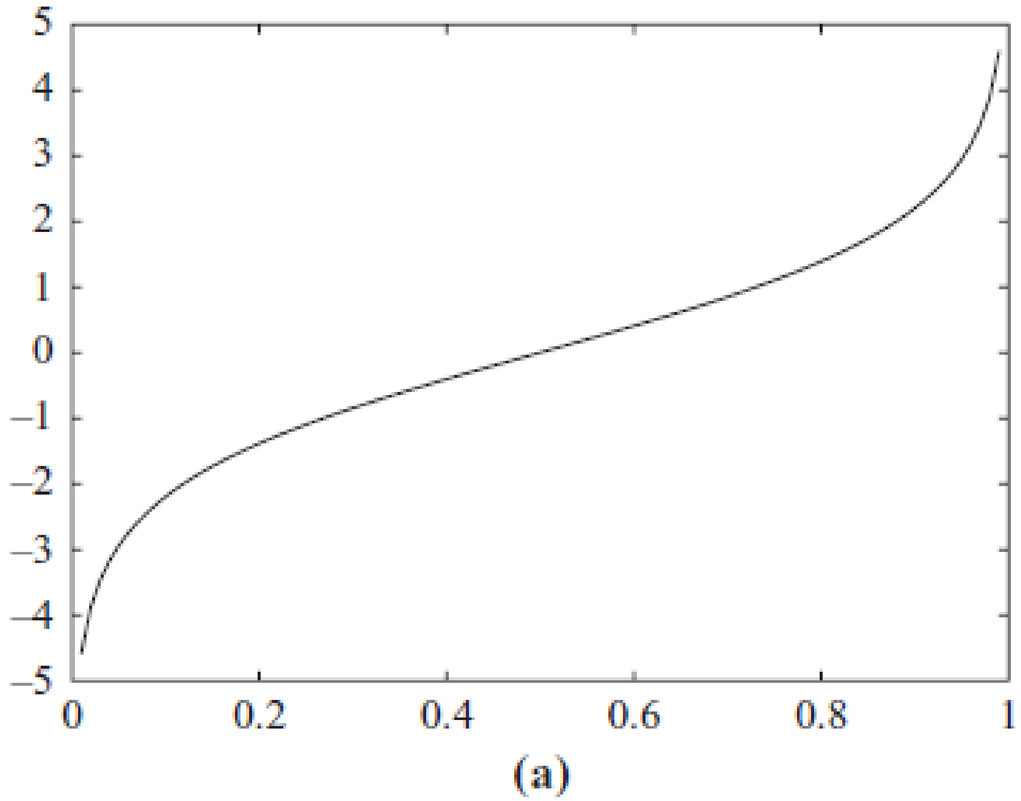
$$logit(p) = log\left(\frac{p}{1-p}\right) = log(p) - log(1-p)$$

| p | ratio odds | logit |
|---|---|---|
| 0.10 | 0.11 | -0.95 |
| 0.30 | 0.43 | -0.37 |
| 0.50 | 1.00 | 0.00 |
| 0.70 | 2.30 | 0.37 |
| 0.90 | 9.00 | 0.95 |
| 0.99 | 99.00 | 1.99 |
| 0.999 | 999.00 | 2.99 |

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Transformación Logística

**Función logit**




| p | ratio odds | logit |
|---|---|---|
| 0.10 | 0.11 | -0.95 |
| 0.30 | 0.43 | -0.37 |
| 0.50 | 1.00 | 0.00 |
| 0.70 | 2.30 | 0.37 |
| 0.90 | 9.00 | 0.95 |
| 0.99 | 99.00 | 1.99 |
| 0.999 | 999.00 | 2.99 |

Los valores de la función logit van desde [-∞,+∞]

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Transformación Logística

- La función **logística** se define como: $P(t) = \frac{1}{1 + e^{-t}}$
- Está función se utiliza para estimar la probabilidad de éxito de un evento determinado.

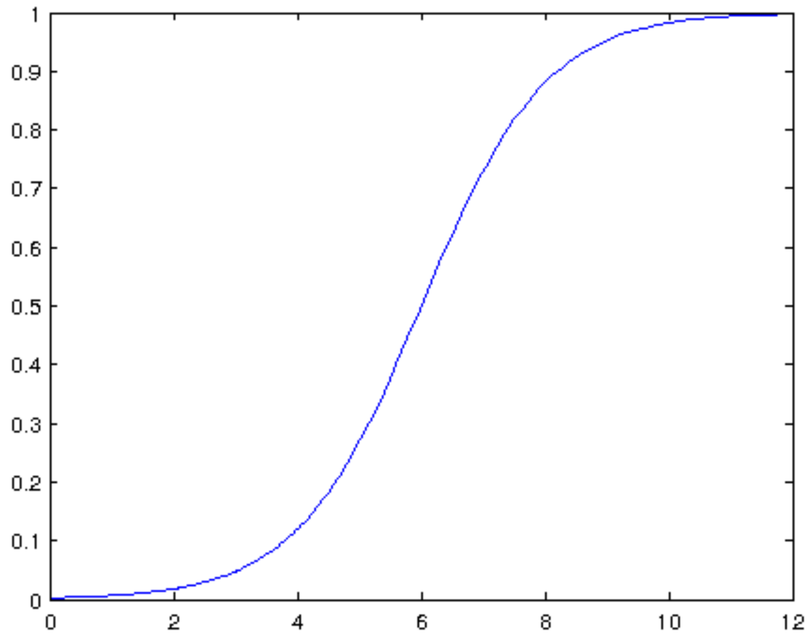



Imagen tomada de http://stackoverflow.com/questions/8252069/matlab-plotting-the-shifted-logistic-function

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Regresión Logística

- En la **regresión lineal**, la salida en sí misma es tomada como el resultado esperado.
  - Sirve para **predecir** valores.
- En **problemas de clasificación**, la salida sirve como frontera de decisión.
  - Sirve para determinar la **pertenencia** a determinada clase.
- La **regresión logística** permite determinar la **probabilidad** de que una instancia pertenezca a una clase.

# Regresión Logística

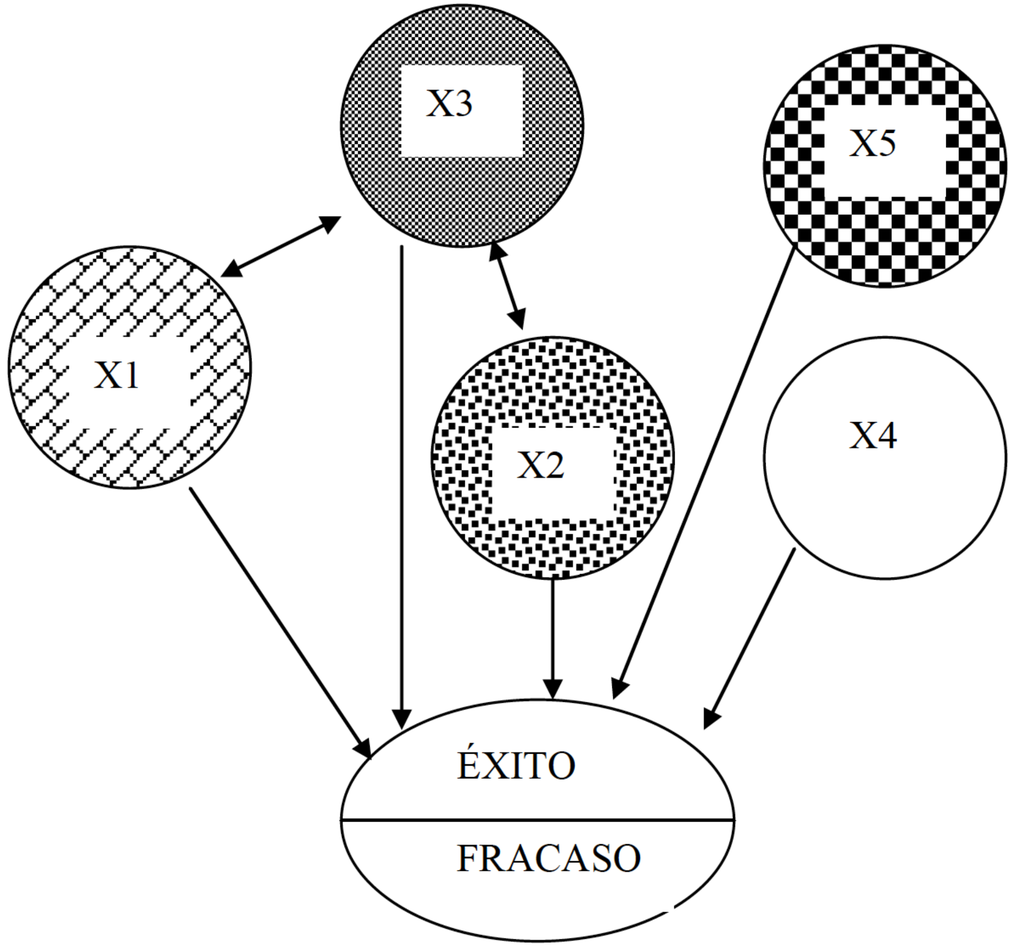



Imagen tomada de Imagen tomada de http://www.uru.edu/fondoeditorial/libros/pdf/manualdestatistix/cap10.pdf

# Transformación Logística

- Supongamos que se tienen solamente **dos clases** para clasificar.
- La **probabilidad** de que una instancia pertenezca a determinada clase se puede describir como:

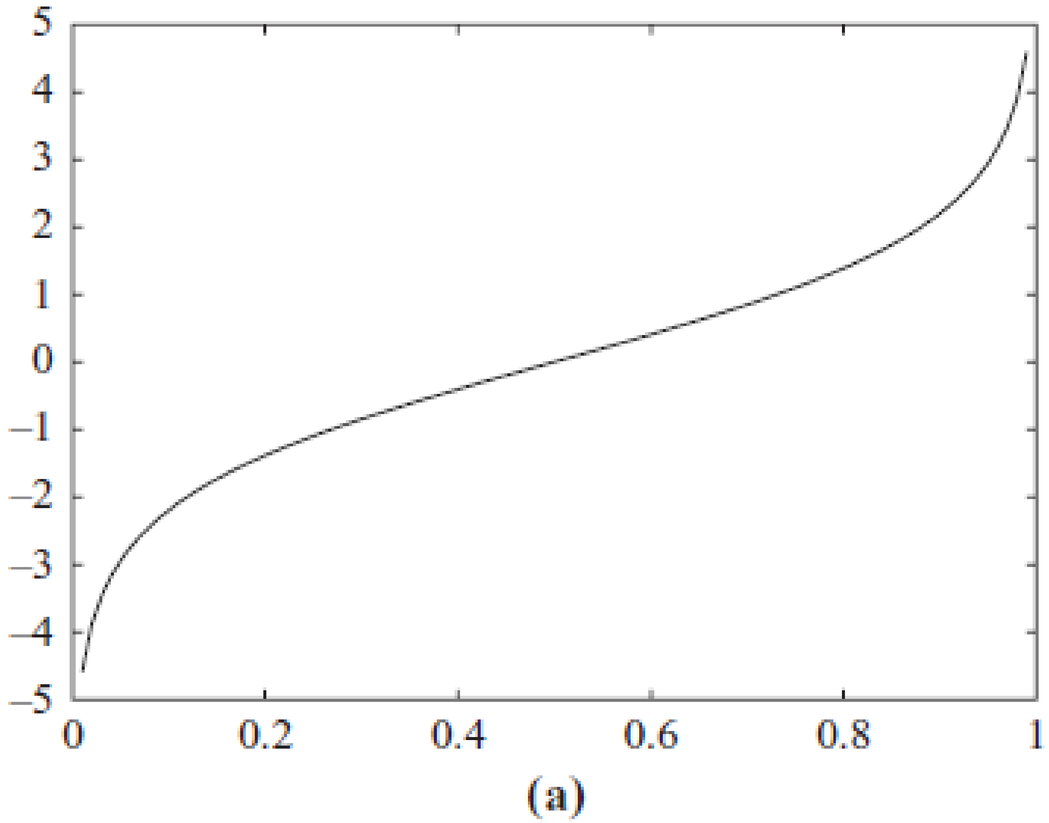
$$Pr[1 \mid a_1, a_2, a_3, \ldots, a_k]$$

- La regresión logística reemplaza la variable **objetivo original** usando la función lineal siguiente:

$$log\left(\frac{Pr[1 \mid a_1, a_2, a_3, \ldots, a_k]}{1 - Pr[1 \mid a_1, a_2, a_3, \ldots, a_k]}\right)$$

# Transformación Logística

- Los valores resultantes ya no están restringidas al intervalo de 0 a 1, puede estar en cualquier lugar entre menos infinito y el más infinito.

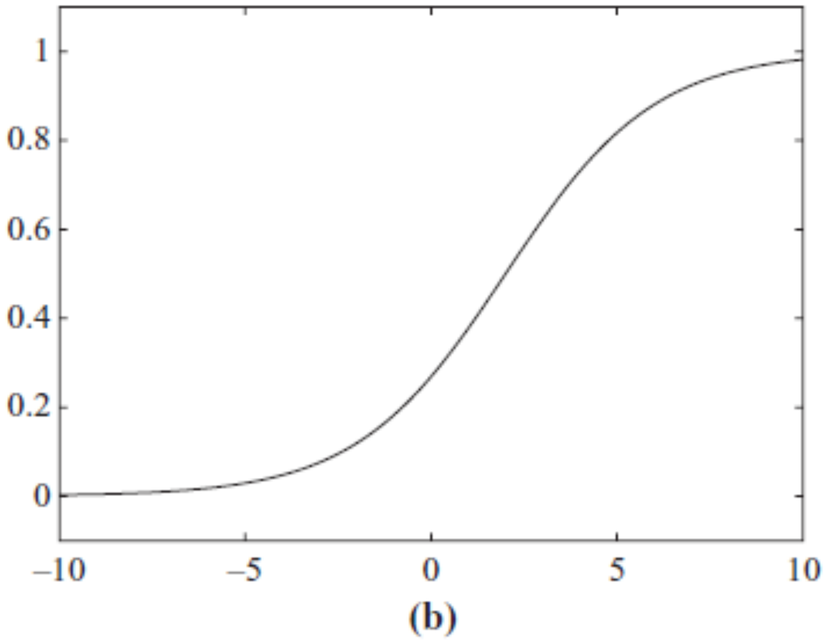



**Transformación logit**

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Transformación Logística

- Para calcular la probabilidad del éxito se usa la función logística.
- De esta manera: $$Pr[1 \mid a_1, a_2, a_3, \ldots, a_k] = \frac{1}{1 + exp(-w_0 - w_1 a_1 - \ldots - w_k a_k)}$$




**Función de regresión logística**

WITTEN, I. H.; FRANK, E.; HALL, M. A. ***Data Mining: Practical machine learning tools and techniques.*** Third Edition. Morgan Kaufmann, 2011. ISBN 978-0-12-374856-0.

# Transformación Logística

- De forma similar a la regresión lineal, los pesos deben ser encontrados a partir de los datos de entrenamiento
  - En la regresión lineal el ajuste se hace minimizando el cuadrado de las diferencias del valor real y el valor predicho
  - En la regresión logística se usa el logaritmo de la probabilidad en su lugar.
  - De esto se obtiene:

$$\sum_{i=1}^{n} (1-x)log(1-Pr[1\,|\,a_1,a_2,a_3,\ldots,a_k]) + xlog(Pr[1\,|\,a_1,a_2,a_3,\ldots,a_k])$$

    - Donde x puede ser tanto 1 como 0.
    - Se busca maximizar esta sumatoria.

# Competencias a adquirir en la sesión

- Al finalizar la sesión el alumno comprenderá el funcionamiento del **aprendizaje inductivo**.
- Al finalizar la sesión el alumno implementará **modelos algorítmos de regresión** usando conjuntos de datos.
- Al finalizar la sesión el alumno **entenderá** el algoritmo de **regresión logística**.
- Al finalizar la sesión el alumno **aplicará** el algoritmo de **regresión logística** para obtener modelos algorítmicos.

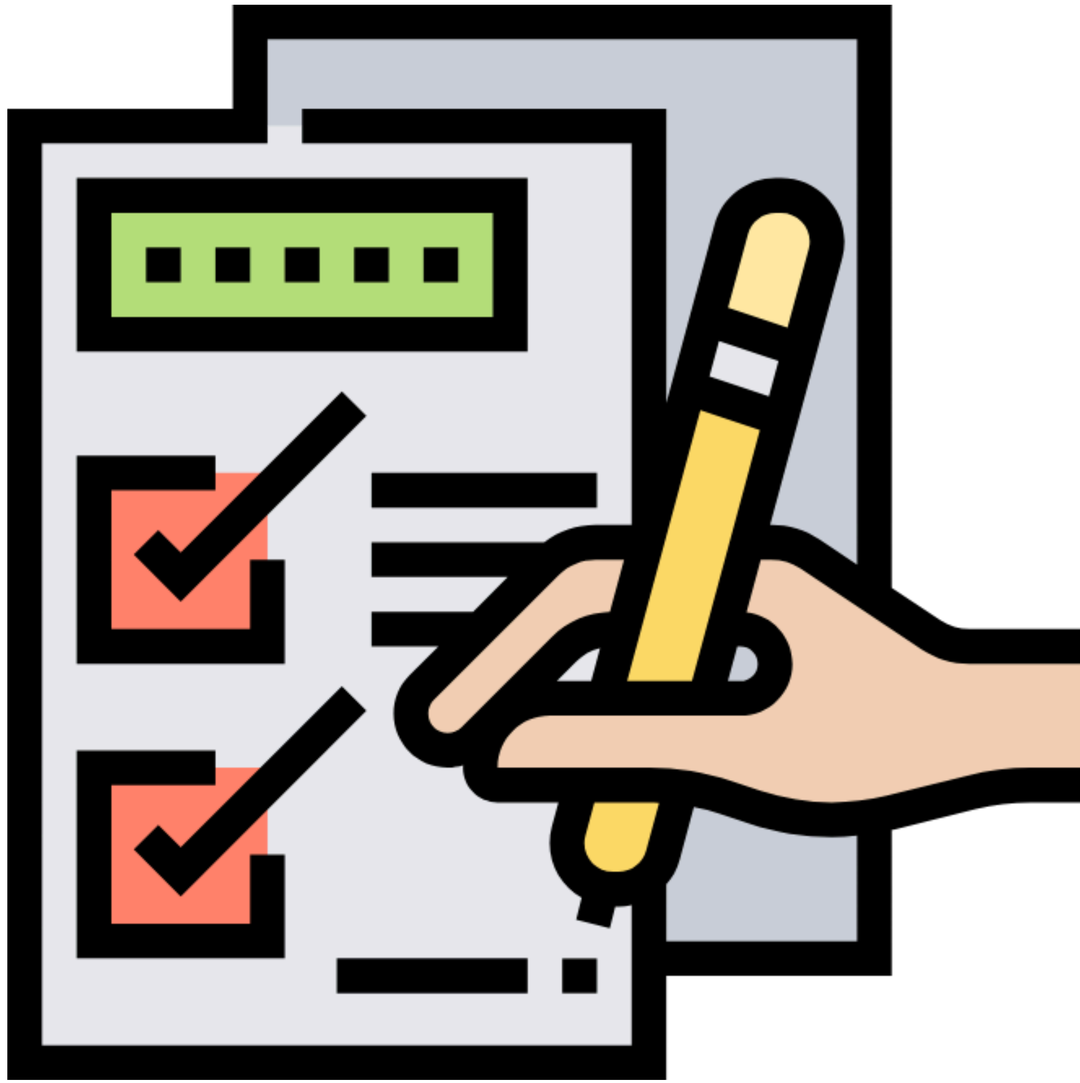


Imagen obtenida a través de la URL https://www.flaticon.com/free-icon/checklist_1467445