

Título o Nombre de las notas

Cristo Daniel Alvarado

11 de marzo de 2025

Índice general

Libros	2
Lenguajes de Programación	2
Páginas y Cursos en Línea	2
Cursos en Coursera	2
Otros Cursos y Recursos	2
Referencias en Teoría de la Decisión	3
Libros Clásicos y Rigurosos	3
Libros con Aplicaciones en Machine Learning y Data Science	3
Temas Clave en Teoría de la Decisión	3
1. Introducción	5
Aprendizaje Estadístico	5
Spam de Email	6
2. Vista Panorámica del Aprendizaje Supervisado	8
Introducción	8
Mínimos Cuadrados y Vecinos Cercanos	8

Información y Recursos

Página del libro (incluye los datasets) de donde estoy sacando algo de contenido: [The Elements of Statistical Learning](#).

§0.1 LIBROS

Escrito por ChatGPT: Si buscas un enfoque teórico y práctico para aprender Machine Learning, estos libros son altamente recomendados:

- **"Pattern Recognition and Machine Learning"** - Christopher M. Bishop Matemáticamente riguroso, ideal para alguien con formación en matemáticas.
- **"The Elements of Statistical Learning"** - Hastie, Tibshirani y Friedman Recomendado para obtener fundamentos teóricos sólidos.
- **"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow"** - Aurélien Géron Enfoque práctico con Python, excelente para comenzar con código.

§0.2 LENGUAJES DE PROGRAMACIÓN

- **Python** (Principal) - Librerías clave: NumPy, Pandas, Scikit-Learn, TensorFlow, PyTorch.
- **R** - Útil para estadística avanzada y visualización de datos.

§0.3 PÁGINAS Y CURSOS EN LÍNEA

§0.3.1 CURSOS EN COURSERA

- Curso de Andrew Ng: [Machine Learning](#) Matemáticas y conceptos clave explicados de manera clara.
- [Deep Learning Specialization](#) Curso avanzado enfocado en redes neuronales.

§0.3.2 OTROS CURSOS Y RECURSOS

- [Fast.ai: Practical Deep Learning for Coders](#) Curso gratuito con enfoque práctico en código.

- [Kaggle Learn](#) Mini cursos interactivos y competencias con datos reales.
- [CS229 de Stanford](#) Notas de clase de Andrew Ng en Stanford.
- **StatQuest (YouTube)** Explicaciones claras de teoría estadística aplicada a ML.

Lo siguiente es sacado de ChatGPT (habla de estadística y de donde encontrar cosas sobre teoría de la decisión y cosas así):

§0.4 REFERENCIAS EN TEORÍA DE LA DECISIÓN

Si quieres profundizar en **Teoría de la Decisión**, aquí tienes algunos libros y recursos avanzados que se alinean con una formación matemática rigurosa.

§0.4.1 LIBROS CLÁSICOS Y RIGUROSOS

- **Statistical Decision Theory and Bayesian Analysis** – James O. Berger
Un enfoque matemático sólido de la teoría de la decisión estadística, con énfasis en métodos bayesianos.
- **Theory of Point Estimation** – Erich L. Lehmann y George Casella
Cubre teoría de la estimación bajo diferentes criterios de optimalidad y conecta con la teoría de la decisión.
- **Decision Theory: Principles and Approaches** – Giovanni Parmigiani y Lurdes Inoue
Presenta teoría de la decisión clásica y bayesiana con aplicaciones en estadística y machine learning.
- **Introduction to Statistical Decision Theory** – John Pratt, Howard Raiffa, Robert Schlaifer
Un libro clásico que introduce la teoría de la decisión en un contexto estadístico.
- **Foundations of Decision Theory** – Itzhak Gilboa
Más filosófico y matemático, ideal si te interesan los fundamentos axiomáticos de la teoría de la decisión.

§0.4.2 LIBROS CON APLICACIONES EN MACHINE LEARNING Y DATA SCIENCE

- **Bayesian Decision Theory and Machine Learning** – David Barber (Parte del libro "Bayesian Reasoning and Machine Learning")
Relaciona la teoría de la decisión con aprendizaje automático.
- **Pattern Recognition and Machine Learning** – Christopher M. Bishop
Enfatiza la teoría de la decisión bayesiana en clasificación y reconocimiento de patrones.

§0.4.3 TEMAS CLAVE EN TEORÍA DE LA DECISIÓN

Si ya tienes nociones de probabilidad y optimización, te pueden interesar:

- **Decisiones Bayesianas** (Bayesian Decision Theory)

- **Teoría de Juegos y Decisión** (Game Theory & Decision Making)
- **Riesgo y utilidad** (Risk & Utility Theory)
- **Reglas de decisión óptimas** (Optimal Decision Rules)
- **Criterios de Minimax y Bayes**

Si buscas algo más aplicado o enfocado en optimización estadística, hay muchos otros recursos disponibles.

CAPÍTULO 1

INTRODUCCIÓN

El objetivo de estas notas es el de dar un panorama general sobre el Machine Learning.

§1.1 APRENDIZAJE ESTADÍSTICO

Definición 1.1.1 (Aprendizaje Estadístico)

El **aprendizaje estadístico** consiste en

El aprendizaje estadístico juega un rol clave en muchas áreas de la ciencia, finanzas y la industria. Tales ejemplos pueden verse en los siguientes problemas:

- Predecir cuando un paciente hospitalizado debido a un ataque cardiaco tendrá un segundo ataque cardiaco. Esta predicción debe basarse en la demografía del paciente.
- Predecir precios de acciones a 6 meses a futuro.
- Identificar números escritos a mano en códigos postales.
- Identificar factores de riesgo para cáncer de próstata, basado en variables clínicas y demográficas.

Definición 1.1.2 (Demografía)

La **demografía** es la ciencia que estudia las poblaciones humanas, su dimensión, estructura, evolución y características generales, así como los procesos que determinan su formación, conservación y desaparición.

El objetivo es tomar información, interpretarla y decir algo acerca de ella, más precisamente, se pretende dar una *predicción*.

Usando la información se construirá un **aprendedor** (o **learner**) que será la pieza clave para predecir la salida de nueva fuente de información. Queremos un *good learner*.

Hay dos tipos de aprendizaje:

- **Supervisado:** En este tipo de aprendizaje hay la presencia de una variable de salida que nos permite guiar el proceso de aprendizaje.
- **No Supervisado:** Observamos solo las características y no tenemos medida de que tan buena es la salida del proceso.

Hablaremos de algunos ejemplos de aprendizaje supervisado:

§1.2 SPAM DE EMAIL

Ejercicio 1.2.1 (Spam de Email)

Se tienen 4601 mensajes en email hacia una persona y pretendemos determinar si cada uno de ellos era un email basura o *spam*. Diseñar un detector de spam automático que pueda filtrar el spam antes de que las bandejas de entrada de los usuarios se llenen.

Solución:

Para cada uno de los 4601 mensajes, podemos asignar dos estados de salida:

email ó spam

Este es un tipo de problema dentro del área de aprendizaje supervisado y es llamado **problema de clasificación**.

	george	you	your	hp	free	hpl	!	our	re	edu	remove
spam	0.00	2.26	1.38	0.02	0.52	0.01	0.51	0.51	0.13	0.01	0.28
email	1.27	1.27	0.44	0.90	0.07	0.43	0.11	0.18	0.42	0.29	0.01

Figura 1. Tabla donde se muestran las probabilidades de que cierta palabra aparezca en un correo del tipo spam y email.

Nuestro método de aprendizaje decide cuáles de características usar y como: Por ejemplo puede usar una regla como la siguiente:

```
1 if george < 0.6 & you > 1.5 then spam, else email.
```

o también:

```
1 if 0.2*you - 0.3*george > 0 then spam, else email.
```

En este problema no todos los errores son iguales, queremos evitar filtrar email que realmente corresponda a email. Una forma de atacar este problema se discute más adelante. ☐

Observación 1.2.1

No existe un lenguaje de programación usado en este texto a menos que se indique lo contrario. Si no parece que es algún lenguaje conocido, como el caso del ejercicio anterior, asumiremos que este es pseudocódigo.

Ejercicio 1.2.2 (Reconocimiento de Números en Manuscrita)

Se tienen miles de cartas escritas a mano con códigos postales en las mismas. Cada carta es escaneada y almacenada. De cada una de las cartas se extrae el área que contiene el código postal y se almacena. El objetivo es determinar el número que está escrito en el código postal.

Solución:

Lo que se hace es primero, separar cada uno de los números en la imagen para después, colocar ese número en un grid de 16×16 bits en escala de grises con cada uno de éstos variando de intensidad desde 0 hasta 255. Se pretende con esto determinar cada uno de los números en cada una de las imágenes.

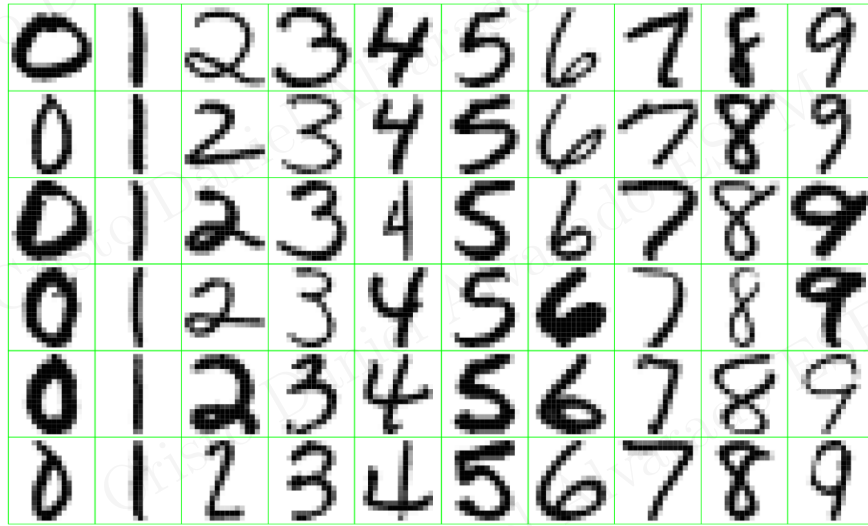


Figura 2. Diversos números sacados de cartas en códigos postales.

Este es un problema de clasificación cuyo error debe mantenerse bajo para poder dar un resultado lo mejor posible. ☐

Otro problema interesante es el de *regresión*, en el que básicamente se da un input de información varia y se requiere dar un output de cierto parámetro que puede o no ser deducido a partir de esos valores introducidos inicialmente.

CAPÍTULO 2

VISTA PANORÁMICA DEL APRENDIZAJE SUPERVISADO

§2.1 INTRODUCCIÓN

El objetivo es que dada entrada de información demos una salida de la misma que podemos medir (ya que en aprendizaje supervisado eso es lo que se hace). Pueden ser muchas entradas de información de diferentes tipos y varias salidas igualmente de diferentes tipos.

Definición 2.1.1 (Terminología Machine Learning)

Los *inputs* son a menudo llamados **predictors** o **independent variables** (**variables independientes** en español), los *outputs* son llamados **responses** o **variables dependientes** (**respuestas** en español).

Se adoptará mucha terminología matemática vista en cursos de álgebra lineal y temas un poco más avanzados, por lo que no se ahondará mucho en ello. Sin embargo, es relevante mencionar el tipo de información con el que puede que estemos trabajando a menudo:

§2.2 MÍNIMOS CUADRADOS Y VECINOS CERCANOS
