



Diplomado Inteligencia Artificial para la Transformación de Negocios


Clase 01 - El valor estratégico de los datos y fundamentos para entender la IA

Sebastián Egaña Santibáñez 

Enlaces del profesor

 <https://segana.netlify.app>

 <https://github.com/sebaegana>

 <https://www.linkedin.com/in/sebastian-egana-santibanez/>

Sobre el docente

Experiencia, educación, etc.

Enlaces

-  sebastian.egana@udd.cl
-  <https://segana.netlify.app>
-  <https://www.linkedin.com/in/sebastian-egana-santibanez/>

Clase 01

- Contexto actual: datos como activo estratégico.
- Rol de la IA en la toma de decisiones.

Estrategia basada en datos (45 min)

Diferencia entre decisiones intuitivas vs. basadas en datos.

Por ejemplo, en el ámbito personal, ¿cómo decidir entre una oferta de trabajo que implica un mayor salario pero requiere re localizarse o una oferta con menor salario pero un mejor equilibrio entre la vida laboral y personal?

La dualidad acá está entre tomar una decisión en base a nuestro instinto o en base a un sistematización de puntos a favor o puntos en contra. Esto independiente del **método** o **metodología**.

Para esto, debemos desarrollar un framework orientado a la toma de decisiones; en la actualidad se habla de **decision science**, como el campo multidisciplinario orientado a desarrollar un enfoque sistemático y orientado en datos (data-driven) para resolver problemas y optimizar desempeño y resultados.

Como áreas relacionadas se encuentran las matemáticas, estadística, psicología, economía, ciencias de la computación, etc.

El punto principal se basa en tomar decisiones informadas; una herramienta para esto corresponde al método científico.

Método científico

Corresponde a la metodología aplicada en ciencias orientada a entender la realidad y generar nuevos conocimientos en base a la observación sistemática, medición y experimentación en base al análisis y formulación de hipótesis. Se puede descomponer en las siguiente etapas:

1. Observación
2. Planteamiento de hipótesis
3. Experimentación / Medición
4. Análisis de resultados
5. Conclusión

En este sentido, **decision science** busca tomar los aspectos positivos del conocimiento científico en el ámbito de la toma de decisiones para organizaciones.

Podemos afirmar lo siguiente: “Cada decisión en salud, finanzas o negocios se basa en una predicción — explícita o implícita. El valor/desafío está en hacerla con evidencia”.

- Decisiones: independiente del tipo de institución, es necesario tomar decisiones de gestión
- Predicción: visión del futuro en base a la información del pasado
- Evidencia: relación con el método científico y **decision science**

Framework de trabajo

El enfoque de **decision science** plantea el siguiente framework:

1. Definir el problema
2. Reunir y analizar datos
3. Desarrollar y evaluar alternativas
4. Seleccionar e implementar soluciones

Esto orientado a tener claridad en la toma de decisiones (no desarrollar/invertir en nada que no tenga sustento en datos), identificar tendencias y generar predicciones para reducir/interiorizar la incertidumbre.

- Por ejemplo: generar una gestión correcta de camas dentro de un recinto hospitalario para aumentar la rentabilidad/bienestar de maximizar la ocupación del recinto.

Ejemplos de uso de datos en distintos sectores (finanzas, salud, retail).

Ecuación de Mincer en economía

$$\ln w_i = \beta_0 + \rho S_i + \beta_1 X_i + \beta_2 X_i^2 + \gamma^\top Z_i + \varepsilon_i \quad (1)$$

- w_i : salario (normalmente por hora, en términos reales).
- S_i : años de escolaridad
- X_i : experiencia potencial (se utiliza también un coeficiente al cuadrado)
- Z_i : puede incluir variables de género, región, etc.



Por lo general este modelo se utiliza para poder explicar factores que influyen dentro de la renta de las personas.

Tiempo promedio de hospitalización:

En un caso simple, podríamos pensar en el promedio de días que pasa cada paciente por alguna patología en particular.

Tenamos los siguientes datos de altas pacientes con neumonía y queremos predecir el tiempo promedio de hospitalización en días:

```
tiempo <- c(2,3,1,5,7,8,9,3,4,6,2,10,11,7,5,4,12,8,9,6)
alta    <- c(1,1,1,1,1,0,0,1,1,1,1,0,0,1,1,1,0,1,1,1)
```

Recordar que acá la censura corresponde a los pacientes que no han sido dados de alta (alta = 0).

Table 1: Conteos: total, eventos (DELTA=1), censuras (DELTA=0)

Total	Eventos_DELTA1	Censuras_DELTA0	Proporcion_evento
20	15	5	0.75

Table 2: Resumen global de TCC (días)

n	min	mean	mean_obs	sd	median	max
20	1	6.1	4.8	3.161	6	12

Veamos el análisis de Weibull para estos datos:

Table 3: Parámetros Weibull 2P (ajuste con censura)

Caso	shape (k)	scale ()
Con censura	1.5848	7.8569

Table 4: Mediana, media y percentiles t_p ($p=0.1, 0.5, 0.9$) con censura

Caso	Mediana	Media	$p=0.1$	$p=0.5$	$p=0.9$
Con censura	6.2347	7.0506	1.8993	6.2347	13.2985

- Concepto de Data-Driven Organization.

Fundamentos estadísticos esenciales (60 min)

- Conceptos clave: promedio, varianza, correlación, causalidad.
- Incertidumbre y confianza en la interpretación de resultados.
- Ejercicios de interpretación (ejemplos sin cálculo formal).

Calidad y gobernanza de datos (45 min)

- Dimensiones de calidad de datos: completitud, exactitud, consistencia.
- Riesgos de decisiones con datos de mala calidad.
- Gobernanza y ética de datos: responsabilidad y compliance.

Conceptos iniciales

Conceptos básicos de series de tiempo

Se habla de series de tiempo cuando nos referimos a una variable sobre la cual observamos su comportamiento a través del tiempo. Un ejemplo conocido, puede ser el principal indicador de productividad de un país, denominado Producto interno bruto (PIB). Veamos la siguiente imagen:



Figure 1: Obtenida desde: <https://revistapaideia.cl/ciclos-economicos-que-son-tipos-y-fases>

Pasemos ahora a hablar de ciertos componentes de las series de tiempo.

Tendencia, ciclo, estacionalidad y ruido

- Tendencia: corresponde al comportamiento general o a la dirección de largo plazo de la serie. Se debe tener cuidado que en algunas variables esto no siempre se cumple.

Veamos el caso para el precio de la acción de LATAM Airlines :

Referencias

- Vecillas Martin, D., Berruezo Fernández, C., & Gento Municio, A. M. (2025). Systematic Review of Discrete Event Simulation in Healthcare and Statistics Distributions. *Applied Sciences*, 15(4), 1861. [Enlace aquí](#)
- Heins, J., Schoenfelder, J., Heider, S., Heller, A. R., & Brunner, J. O. (2022). A scalable forecasting framework to predict COVID-19 hospital bed occupancy. *INFORMS Journal on Applied Analytics*, 52(6), 508-523. [Enlace aquí](#)
- Harper, P. R., & Shahani, A. K. (2002). Modelling for the planning and management of bed capacities in hospitals. *Journal of the Operational research Society*, 53(1), 11-18.