



TALLERES COMPLEJIDAD EN SALUD Y ECOLOGÍA: RIESGOS Y DATOS

FACTORES





Enfermedades no transmisibles y Medicina Traslacional

13 de septiembre 2023



PROYECTO 42



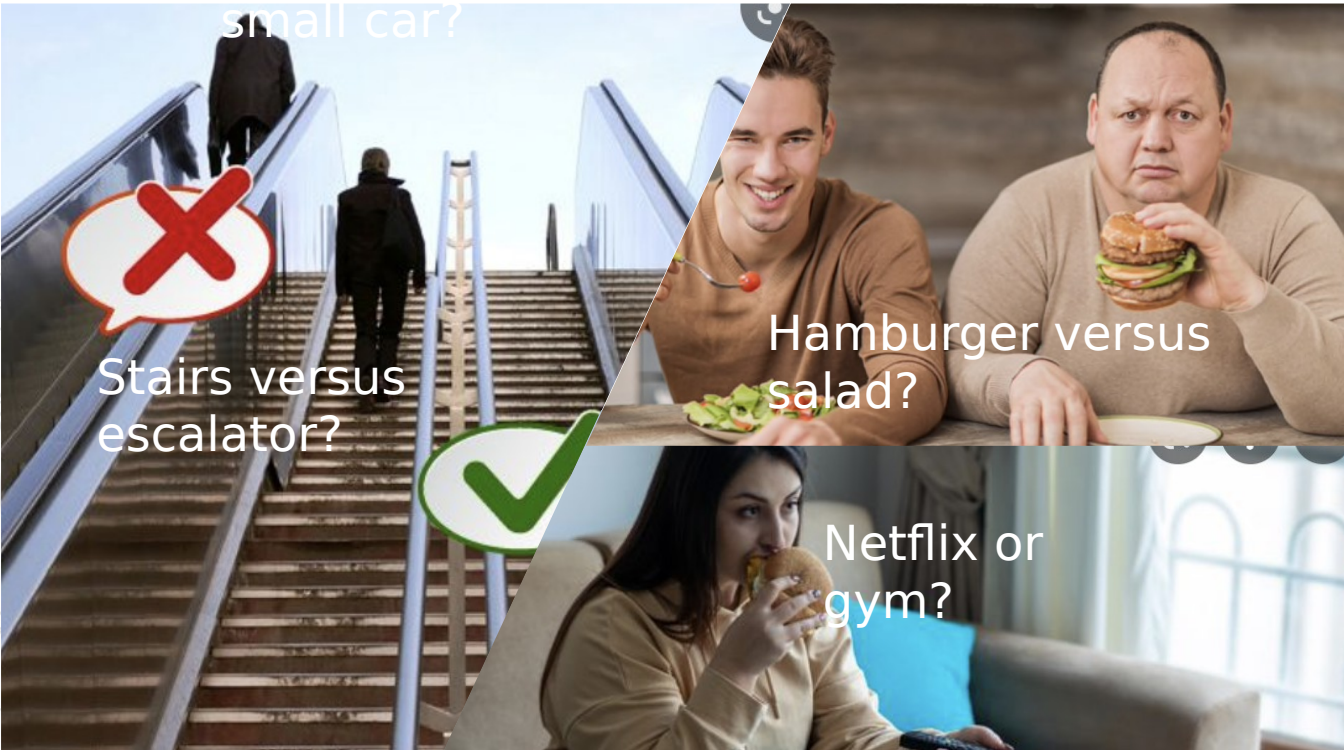
Epi-PUMA



Cree rápidamente su propio modelo predictivo centrado en una predicción vinculada a una decisión en un entorno amigable donde no se requieren conocimientos especializados.



El origen de las enfermedades no transmisibles ¡y transmisibles!



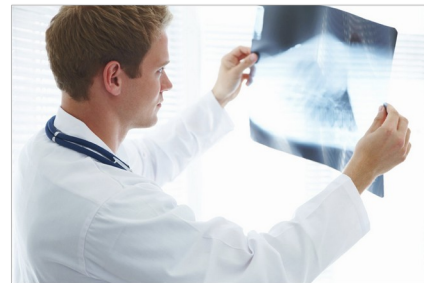
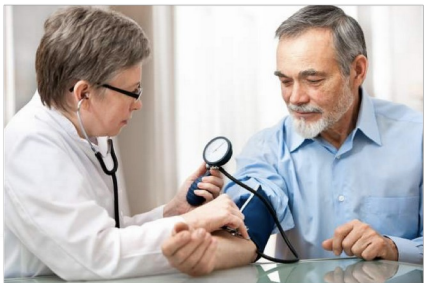
Son consecuencias de nuestras decisiones

Resist/don't resist?

Invade/don't invade?

Su mitigación depende de otras decisiones

Diagnostics



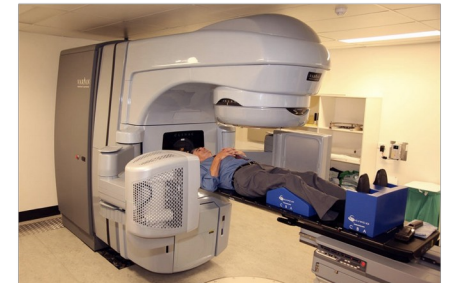
QBC Fluorescence Microscopy

- Upgrade your microscope to fluorescence
- Malaria diagnostics
- Tuberculosis diagnostics
- Other blood borne parasite diagnostics

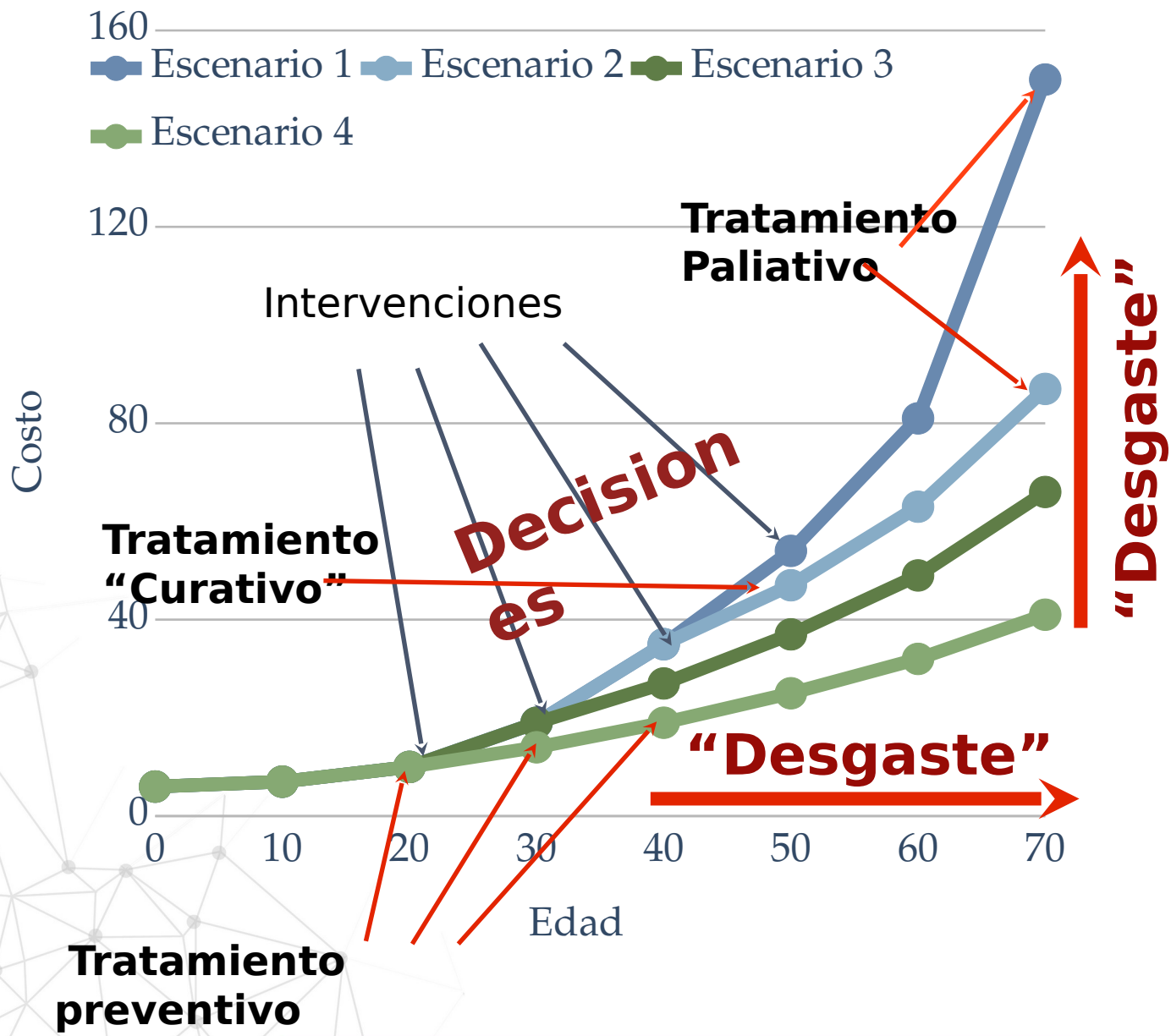


QBC Europe

Treatments



Y las consecuencias de estas decisiones...



Escenario 1: Onset of obesity at 20
 Continued obesity and onset of metabolic syndrome at 40
 Onset of diabetes at 50
 Onset of renal failure at 60
 Death at 70

Escenario 2: Onset of obesity at 20
 Continued obesity and onset of metabolic syndrome at 40
 Onset of diabetes at 50
 Adequate control and treatment of comorbidities at 60
 Bad health and high cost at 70

Escenario 3: Overweight at 20
 Obesity at 40
 Onset metabolic syndrome at 50
 Onset of diabetes at 60
 Continued diabetes but no serious comorbidities at 60
 Ill health and moderate cost at 70

Escenario 4: Overweight at 30
 Obesity at 50
 Onset of metabolic syndrome at 60
 Onset of diabetes at 70 but relative health

¿Cómo se predice que trayectoria alguien seguirá? Y ¿Por qué?

Sin embargo, a pesar de una inversión enorme...

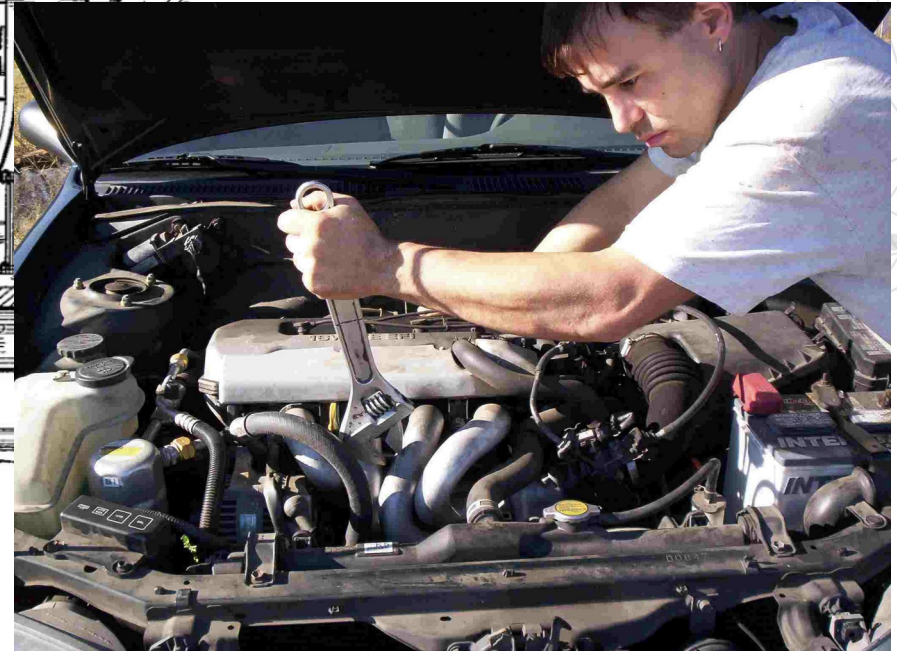
- Los gobiernos no han podido solucionarlos
- Las sociedades no han podido solucionarlos
- La ciencia no ha podido solucionarlos

¿Por qué?

La medicina considera que...

**El mundo es una máquina
¡Incluyendo gente!**

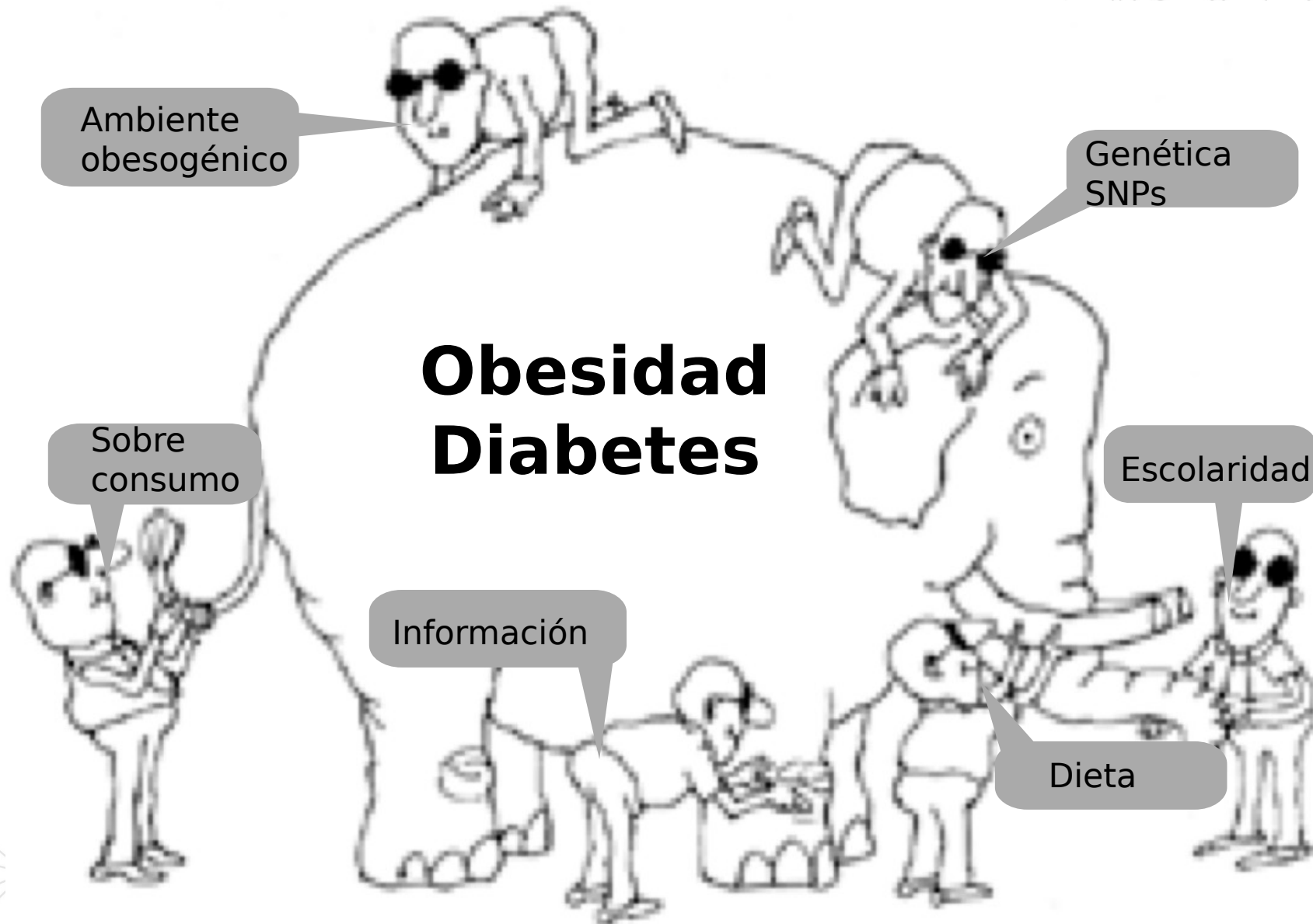
**Es el paradigma de
la medicina "curativa"**



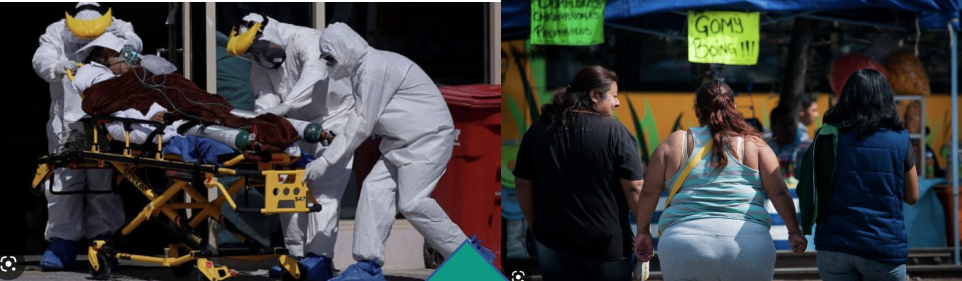
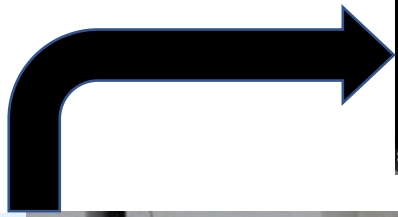
**Ha sido muy exitoso.
Pero...**

¡Somos “Sistemas Complejos Adaptativos”!

- Multifactorial
- Múltiples causas directas y indirectas
- Multiescala
- Multidisciplinaria



¿Cómo tomamos nuestras decisiones?



- ¿Dónde habrá más pérdida de biodiversidad debido al cambio climático?
- ¿Quién está en más riesgo de infectarse por T.cruzi?
- ¿Qué es más importante en determinar la distribución de un especie particular?
- ¿Cuándo hay más riesgo de un aumento en casos de SARS-Cov-2?



López-Gatell, AMLO say no to enforcing use of coronavirus face masks

MND Staff October 28, 2020



The effectiveness of face masks is 'overstated', says Mexico's coronavirus czar.

Consecuencias
Objetivos y metas estratégicos



**NO HAY
 DECISION SIN
 PREDICCIÓN**

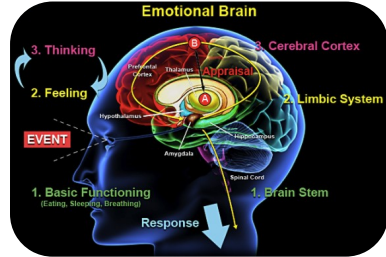
$$P(C|X)$$

INTELIGENCIA HÍBRIDA

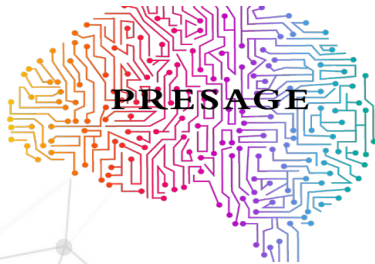
El Santo Grial para solucionar los problemas de los Sistemas Complejos Adaptativos

Potencia de procesamiento simbólico limitada ~ 30 bytes/seg

- No hablamos "base de



+



Las computadoras no "entienden" nada

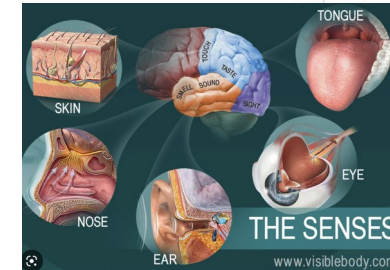
- Produce máquinas - especialistas

¿Qué quiero predecir?

¿Con qué lo quiero predecir?

$$P(C | X(t))$$

¿Qué modelo predictivo se usará?



+



LA PROBABILIDAD DE C DADO $\mathbf{X}(t) = (X_1, X_2, \dots, X_N)$

Nuestra predicción es una *inferencia estadística*

Tenemos un modelo de predicción
para cada una de estas situaciones.
¡Y muchas más!

$$P(C|X(t))$$

C = aguilas
X = mi moneda
X = todas nuestras monedas

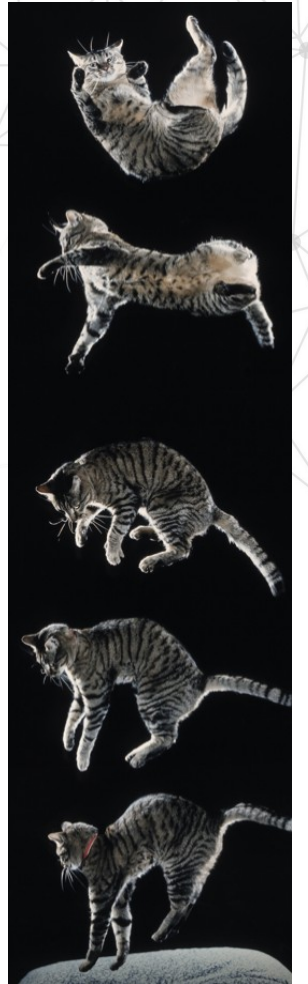
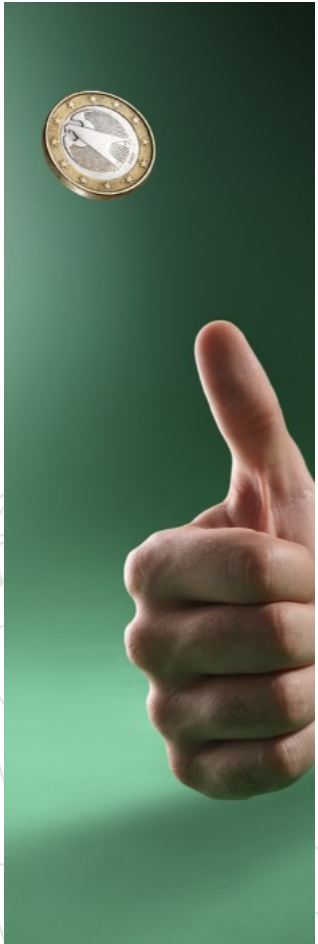
$$P(C) = 0.5$$

Bayesian priors

Nuestros prejuicios
Nuestras creencias

C = gato aterriza patas abajo
X = mi gato
X = gatos en general

$$P(C) \sim 1.0$$



¿Cómo se evalúan las predicciones?

- Se compone un “ensamble estadístico”
 - P.ej. se puede lanzar mi moneda N veces o N monedas una vez – representan diferentes comparaciones
 - $P(C)$ se convierte a una hipótesis nula. Se compara $P(\text{aguila} \mid \text{mi moneda})$, como hipótesis alternativa, con $P(\text{aguila}) = 0.5$ para ver como tan probable es las lanzas de la moneda representan monedas justas
 - Se calcula $P(\text{aguila} \mid \text{mi moneda})$ por lanzar la moneda N veces y ver con que frecuencia salio aguila. Se puede usar prueba de hipótesis para ver como tan probable era para observar una discrepancia entre $P(\text{aguila} \mid \text{mi moneda})$ y 0.5 por casualidad
 - Para el gato se puede formar un ensamble por dejar caer un gato N veces o N gatos una vez y comparar con la hipótesis nula que $P(\text{gato aterriza patas abajo}) = 1$, o podemos creerlo hasta vemos evidencia al contrario. Hay ensambles que son difíciles armar.
 - ¿Qué evidencia se necesita y cómo se incorpora en nuestro modelo?

¿Cómo ajustamos nuestras expectativas?

Aprendemos... esperamos

- $P(C) \square P(C|\mathbf{X}) = P(\mathbf{X}|C)P(C)/P(\mathbf{X}) \square P(C|\mathbf{X},\mathbf{X}') = P(\mathbf{X}'|C,\mathbf{X})P(C|\mathbf{X})/P(\mathbf{X}'|\mathbf{X}) \square \dots$
- “Corregimos” nuestro estimado inicial (prior) debido a la información \mathbf{X} a través de la Teorema de Bayes
 - $P(\mathbf{X}|C)$ es la verosimilitud (likelihood) que se verá los datos \mathbf{X} en los casos C y $P(\mathbf{X})$ es la evidencia
 - Llegamos a un nuevo estimado (el posterior) que toma en cuenta la nueva información que sirve como un nuevo prior
- “Corregimos” nuestro nuevo prior debido a la información \mathbf{X}' a través de la Teorema de Bayes de nuevo
 - $P(\mathbf{X}'|C,\mathbf{X})$ es el nuevo verosimilitud (likelihood) que se verá los nuevos datos \mathbf{X}' en los casos que tienen tanto C como \mathbf{X} y $P(\mathbf{X}'|\mathbf{X})$ es la evidencia acondicionado en la información inicial \mathbf{X}
 - Llegamos a un nuevo estimado (el posterior) que toma en cuenta la nueva información \mathbf{X}' que sirve como un nuevo prior etc etc

¿Qué X podemos/debemos incluir?

- ¿Un sitio web que dice que la teoría que los gatos aterrizan patas abajo es una conspiración neo-liberal/comunista/extraterrestre/...?
- ¿Un sitio web www.funanimalfacts.org que dice que los gatos no aterrizan patas abajo más que cualquier otro animal?
- Un estudio que apareció en Scientific American <https://www.scientificamerican.com/article/why-do-cats-land-on-their-feet-physics-explains/>
- Aunque esto no suena serio, es muy serio. La ciencia misma esta llena de prejuicios y sesgos tanto explícitos como implícitos. La disciplinariedad misma es un ejemplo de esto. Ni hablar de las autoridades y los individuos.

Ejemplos de cosas que se quiere predecir y que consideraremos en los casos de uso

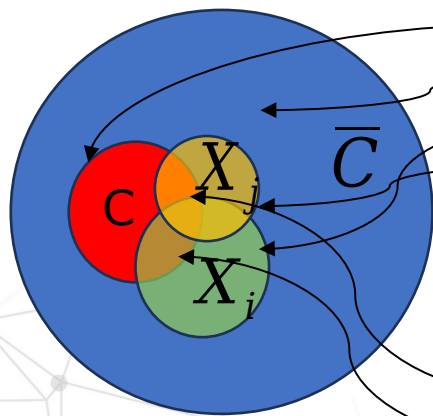
1. La obesidad C_1 ¿Cuánto pesa cada factor de riesgo en este padecimiento?
2. Diagnóstico de COVID19 a través de la sintomología C_2
3. Riesgo de fallecer de COVID19 en una trayectoria clínica C_3
4. ¿Es probable que esta persona fallecerá de COVID19? C_4
5. ¿Es probable que esta persona padece de hypoglucemia? C_5

La Obesidad: ¿Cuánto pesa cada factor de riesgo en este padecimiento? ? C_1

- Empezamos con $P(C_1 = \text{obeso})$ como hipótesis nula y $X_i = \text{no haces ejercicio actualmente}$.
- Se quiere saber si $P(C_1 = \text{obeso} | X_i = \text{no haces ejercicio actualmente}) > P(C_1 = \text{obeso})$ o $< P(C_1 = \text{obeso})$
 - Si una persona ya es obesa $P(C_1 = \text{obeso} | X_i)$ puede explicar parcialmente porque la persona es obesa y si una persona todavía no es obesa $P(C_1 = \text{obeso} | X_i)$ puede cuantificar el riesgo que será obesa.
 - ¿Se calcula $P(C_1 = \text{obeso} | X_i = \text{no haces ejercicio actualmente})$ cómo las “monedas” o como los “gatos”?
 - Gatos: Usando la Inteligencia Humana estimamos $P(C_1 | X_i)$ sin saber que otros factores \mathbf{X} son relevantes y sin saber que es $P(\cdot | \cdot)$ - nuestro cerebro es “caja negra”
 - “Experiencia” (p.ej. médico vs. físico) nos permite hacer una estimado de riesgo (sesgada hasta muy sesgada) de mayor o menor precisión
 - “Monedas”: Se forma un ensamble estadístico de N personas, cuentas $N(C_1 X_i)$ y $N(X_i)$ y calculas $P(C_1 = \text{obeso} | X_i) = N(C_1 X_i) / N(X_i)$
 - Suponemos que todos los miembros del ensamble son “idénticos”. Mi riesgo viene del riesgo calculado respecto al ensamble
- Para la medicina “personalizada” se quiere calcular $P(C_1 | \mathbf{X})$ donde el perfil $\mathbf{X} = (X_1, X_2, \dots, X_N)$ es único. ¿Cómo se valida la predicción? $N(C_1 \mathbf{X}) = 0, 1$
- Ya no hay posibilidad de un ensamble de personas \mathbf{X} idénticas. Se requiere modelos de Aprendizaje de máquina/Inteligencia Artificial

¿Cómo se calcula $P(C|\mathbf{X})$?

- $P(C|\mathbf{X}) = N(C\mathbf{X})/N(\mathbf{X})$ No se puede calcular directamente
= $P(\mathbf{X}|C)P(C)/P(\mathbf{X})$ Teorema de Bayes – se puede incluir Bayesian prior
usando la Inteligencia Humana
- $P(\mathbf{X}|C) \square$ Aproximación de Naïve Bayes
- $S(C|\mathbf{X}) =$ Score total $S(C|\mathbf{X}) >/<$ cero \square \mathbf{X} más/menos probable ser obeso
- Contribución de la variable al score total



- C = obesos
- \bar{C} = no obesos
- X_j = no hacen ejercicio
- X_i = comen más que lo recomendado
- $N(C)$ = número de obesos
- $N(\bar{C})$ = número de no obesos
- $N(X_j)$ = número que no hacen ejercicio
- $N(X_i)$ = número que comen más que lo recomendado
- $N(C \cap X_j)$ = número que son obesos y no hacen ejercicio
- $N(C \cap X_i)$ = número que son obesos y comen más que lo recomendado

Todo es puro conteo

¿Qué se quiere del modelo de $P(C|\mathbf{X})$?

- ¿Que es predictivo?
 - Se puede juzgar objetivamente como tan bueno es el modelo a través de muchas métricas estandares de desempeño
 - Se puede incluir y excluir variables X_i y juzgar si se aumentó/disminuyó el desempeño
- ¿Qué explica?
 - ¿Cuándo importa? Hay problemas donde los “por qué” no importan (p.ej. Traducción automática) versus otros donde los “por qué” si importan mucho (p.ej. diagnóstico médico). ChatGPT versus las plataformas de Chilam. Caja negra versus caja blanca.
- ¿Qué representa múltiples realidades?
 - La realidad actual es representado por cada elemento del ensamble c siendo representado por un estado $\mathbf{X}(c)$. Con el modelo $P(C|\mathbf{X})$ se puede crear cualquier escenario de “que-si” para calcular $P(C|\mathbf{X}'(c))$
- Predictabilidad es objetivo y explicabilidad es subjetivo (p.ej. muy dependiente de tu disciplina)
- Para optimizar la predictabilidad y crear nuevas realidades se requiere la Inteligencia Artificial y para optimizar la explicabilidad se requiere la Inteligencia Humana

¿Qué se quiere de la toma de decisión usando el modelo de $P(C|\mathbf{X})$?

- Si C representa algo bueno/malo se quiere aumentar/disminuir $P(C|\mathbf{X}) = P(C|X_1, X_2, \dots, X_N)$
- Se cambia $P(C|X_1, X_2, \dots, X_N)$ por cambiar uno o más X_i
- Estos cambios pueden ser debido a la dinámica actual del sistema – p.ej. Reducción en el número de susceptibles – o por una intervención humana – p.ej. Un programa de vacunación o educación.
- Hay demasiado posibles intervenciones para probarlas uno por uno y ver si ayudaron o no
- Se requiere saber el grado de causalidad entre C y X_i – X_i es directamente causal, indirectamente causal, pura correlación.
- Se requiere saber si X_i es “actionable” – es algo que tu intervención puede cambiar. P.ej. sexo no es actionable pero estilo de vida, potencialmente, si es
- $P(C|X_1, X_2, \dots, X_N)$ puede representar un enorme número de posibles mundos/realidades además de lo nuestro



Resumen

1. Cada problema de mayor grado que se enfrenta la humanidad se debe a una mala toma de decisión que tiene consecuencias adversas
2. Cada decisión requiere una predicción - predicción es el fundamento de la vida, la ciencia, la medicina,..., todo.
3. Únicamente los Sistemas Complejos Adaptativos toman decisiones
4. Para predecir SCA se requiere modelos de inferencia estadística $P(C|X)$ que: i) Son “complejos” - multifactoriales - y “adaptativos” - que las relaciones entre los C (“efectos”) y X_i (“causas”) pueden cambiar en tiempo.
5. El marco Bayesiano es lo más adecuado para modelar los SCA y calcular $P(C|X)$ - modela como los seres humanos (seres vivos) construimos modelos de predicción.
 - Incorpora el aprendizaje naturalmente.
 - Puede incorporar tanto “creencias” como datos

Resumen

6. La multi-factorialidad es clave. La disciplinarietà y el reduccionismo son los obstáculos principales porque la ciencia no ha avanzado más en solucionar o mitigar estos grandes problemas
7. La multifactorialidad implica la necesidad de BIG, DEEP data para representar en datos
8. Hoy en día estos datos son digitales y nosotros no hablamos base de datos
9. Así para calcular $P(C|\mathbf{X})$ se requiere algoritmos de ML/AI
10. Hay diferentes tipos de algoritmos, cada uno con sus pros y cons. Ninguno puede substituir la necesidad de la Inteligencia Humana para la interpretación de $P(C|\mathbf{X})$ y su uso en la toma de decisión.



¿Qué hacen nuestras plataformas?

- Se puede plantear y validar

Cierre

Discusión, retroalimentación y seguimiento

- Posibilidades de colaboración
 - Involucramiento en:
 - Colecta de datos
 - Curación y validación de datos
 - Análisis de datos
 - Uso de los datos para:
 - Investigación
 - Docencia
 - Toma de decisión
 - Uso de las plataformas para:
 - Investigación
 - Docencia
 - Toma de decisión