



El Conductoma: Un Nuevo Paradigma

Christopher Stephens

C3 – Centro de Ciencias de la Complejidad e
Instituto de Ciencias Nucleares, UNAM
XXVIII Congreso Mexicano de Psicología
I Congreso Latinoamericano de Psicología



¿Qué es El Conductoma?

Un intento para predecir y entender la conducta humana holísticamente usando...

1. Un marco conceptual y teórico novedoso
2. Datos grandes y profundos (Big, deep data) - multi-escala → multi-disciplinario
Porque nuestra conducta depende de un sinfin de factores.
3. Modelos de predicción de Aprendizaje de Máquinas que son "precisos", explicables y útiles
Porque no hay un concepto de decisión sin una predicción correspondiente.

La perspectiva "omica" es para indicar que se intentará caracterizar la totalidad de los factores que entran en la descripción y predicción de la conducta humana



¿Por qué un Conductoma?

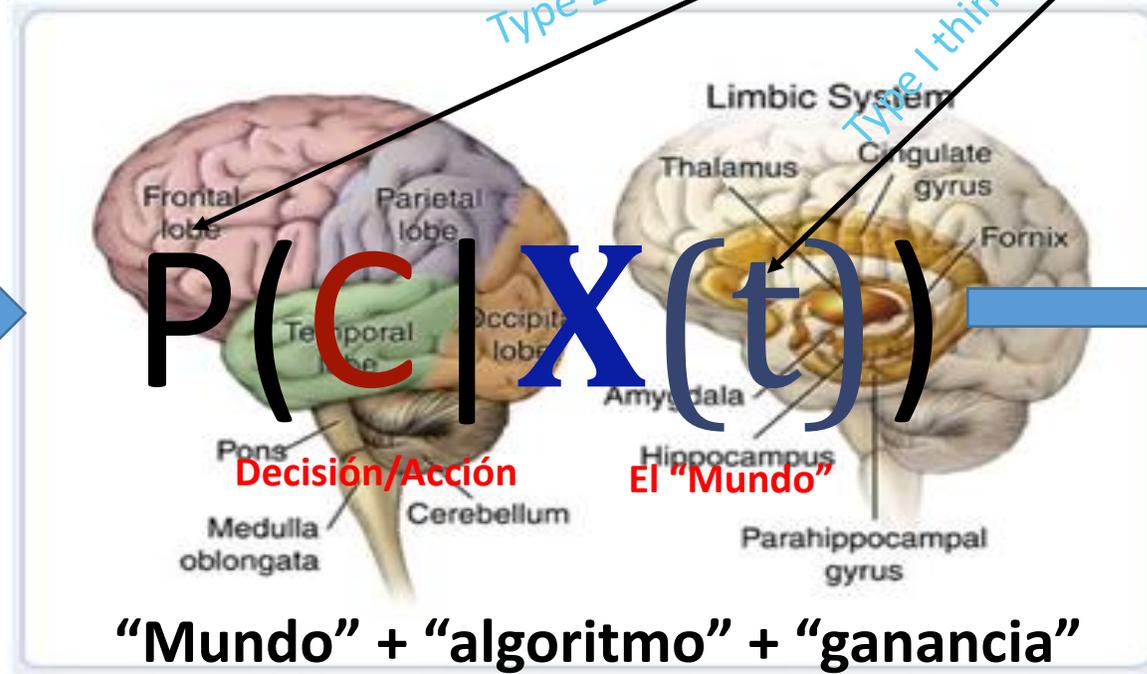
1) El Marco Conceptual y Teórico

¿Qué es El Conductoma?



Datos + Significando + Valor →
Predicción + Decisión + Acción

Datos +
Significado
+ Valor



Type 2 thinking

Type 1 thinking

Predicción
+ Decisión
+ Acción



Hay muchas alternativas para cualquier decisión y muchas posibles acciones

¿Funcionó?

DSV → PDA se evalúa ser bueno o malo según un criterio de desempeño - ganancia

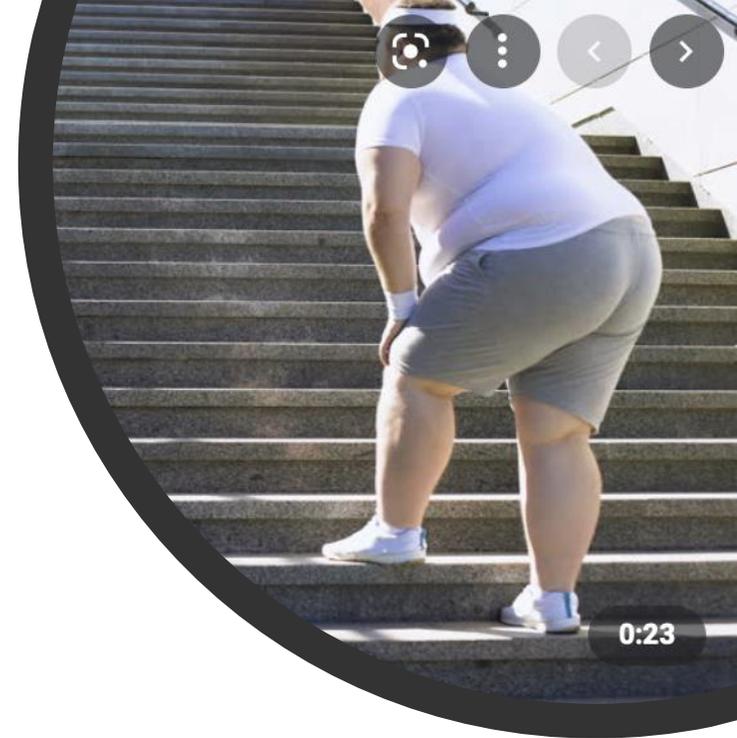


$P(C | X(t))$ representa nuestro modelo de la realidad y su percepción.

- Es una **inferencia estadística** que es basada en un **conjunto** de “**eventos**” (reales – ensamble “**externo**” o imaginados - ensamble “**interno**”)
- El “**externo**” es objetivo y empírico – se cuenta eventos y relaciones entre efectos y sus potenciales causas (estadística frecuentista)
- “**Interno**” es subjetivo y asociado con nuestro **modelo mental** del mundo (estadística Bayesiana).

Elevators vs Stairs

How Long Does it Take to Reach the Top of These Famous Structures?



Burn Calories
Not Electricity



Take the stairs!

Skip the elevator and escalator. Walking up stairs just 2 minutes per day helps prevent weight gain. And it helps the environment by saving electricity.

¿Por qué tomamos “malas” versus “buenas” decisiones?



Los retos de los “conjuntos” internos y externos

El conjunto externo (fenomenológico)

- $\mathbf{X}(t)$ tiene muchas dimensiones – muchos datos/escalas/disciplinas
- $P(\mathbf{C}|\mathbf{X}(t)) = N(\mathbf{C} \mathbf{X}(t))/N(\mathbf{X}(t))$
- Si $\mathbf{X}(t)$ tiene muchas dimensiones $\rightarrow N(\mathbf{X}(t)) = 0, 1$
- Por eso hay que usar modelos de aprendizaje de máquinas
- ¿Cómo se mide la ganancia?

El conjunto interno (teórico)

- “ $P(\mathbf{C}|\mathbf{X}(t))$ ” viene del modelo mental de la persona
- ¿Cómo se mide/describe el “mundo”/“algoritmo”/“ganancia”?

(Casi) cualquier decisión es multi-objetiva

- En la “Teoría de la Elección Racional” somos agentes que maximamos una función única de utilidad – hay nada más un objetivo
- En el experiment del elevador hay multiples objetivos y utilidades
 - DT = la diferencia en tiempo entre tomar la escalera versus el elevador;
 - DE = la diferencia en energia gastado en uno versus el otro;
 - DH = el beneficio de salud percibido;
 - DS = el beneficio social;
 - Y hay otros.
- No son mutuamente satisfactorios. Es un sistema “frustrado”. P.ej. diferencia de tiempo y de energia.
 - El valor relativo de uno versus el otro es una función de los estados del “sujeto”, “objeto” y el “ambiente”.
 - P.ej., si uno esta cansado (efecto del sujeto) será más probable tomar el elevador. Si el elevador es muy pequeño o si la escalera es muy empanada (efecto del objeto) puede afectar la decision. Si hay información disponible al lado de los dos que habla del beneficio de salud de uno versus el otro (efecto del ambiente), también puede afectar la decision.

¿Cómo se toma una decisión?

¿Cuánto te vale tu decisión?

- Existen N_v funciones de valor/utilidades: $\mathbf{V} = (v_1, v_2, \dots, v_{N_v})$ – gasto energético, tiempo etc.
- Una acción A_k causa cambios en estas funciones $D\mathbf{V}(A_k) = (Dv_1(A_k), Dv_2(A_k), \dots, Dv_{N_v}(A_k))$.
- $D\mathbf{V}(A_k)$ puede tomar diferentes valores dependiendo de si se mide antes o después de la acción
 - Pos-acción, $D\mathbf{V}(A_k)$ representa las ganancias reales o percibidas - resultados. En lo general son subjetivos, pero a veces pueden ser comparado con la realidad.
 - Pre-acción, $\langle Dv_1(A_k) \rangle$ representa una predicción de la ganancia debida a la acción. $\langle \dots \rangle$ no necesariamente significa que es un valor esperado asociado con un conjunto externo, aunque podría ser, pero, más bien significa que uno tiene un modelo interno de predicción para estimar el cambio.
- Se hipotetiza que la probabilidad para la acción A_k es
- $P(A_k | \langle D\mathbf{V}(A_k) \rangle) = P(A_k | \langle Dv_1(A_k) \rangle, \langle Dv_2(A_k) \rangle, \dots, \langle Dv_{N_v}(A_k) \rangle)$
- Así, la probabilidad para la acción A_k se acondiciona en las ganancias predichas de la acción asociadas con las funciones de ganancia.
- Una decisión es “buena” o “mala” respecto a una función de valor v_i si $Dv_i(A_k) > 0$ (“buena” decisión) versus $Dv_i(A_k) < 0$ (“mala” decisión). Así, decisiones pueden ser frustradas – buenas para algunas funciones de valor y malas para otras.

¿Cómo se evalúa una utilidad?

- Si se supone que se toma una decisión y una acción correspondiente según los cambios predichos en un conjunto de utilidades hay que preguntar ¿Cómo se hace la predicciones?
- $P(Dv_i(A_k) \mid \mathbf{X}(\text{Sujeto}), \mathbf{X}(\text{Objeto}), \mathbf{X}(\text{Ambiente}))$
- $\langle Dv_i(A_k) \rangle = F(P(Dv_i(A_k) \mid \mathbf{X}(\text{Sujeto}), \mathbf{X}(\text{Objeto}), \mathbf{X}(\text{Ambiente})))$
- i.e., dentro de nuestro modelo mental del mundo existe un modelo de predicción que estima la probabilidad de una cierta ganancia dado que se implementa la acción en un estado dado del mundo (sujeto, objeto, ambiente) y una ganancia esperada.
 - P.ej., si el sujeto esta en el estado “muy cansado”, la probabilidad para un aumento percibido grande de esfuerzo, DE, por tomar la escalera sería más alto que en el estado “no-cansado” con la consecuencia que $P(\text{escalera} \mid \langle DE(\text{escalera}) \rangle_1) < P(\text{escalera} \mid \langle DE(\text{escalera}) \rangle_0)$, donde $\langle DE(\text{escalera}) \rangle_1$ es el esfuerzo predicho por tomar la escalera dado que el sujeto esta en el estado 1 = “muy cansado” y $\langle DE(\text{escalera}) \rangle_0$ es el esfuerzo predicho por tomar la escalera dado que el sujeto esta en el estado 0 = “no cansado”

THE ULTIMATE ANSWER
TO LIFE, THE UNIVERSE
AND EVERYTHING IS

42

2) Datos Grandes y Profundos Proyecto 42

Project 42 y el Conductoma

(CONACyT Fronteras, CONACyT Redes, PAPIIT, SECTEI y Microsoft)

Fase I: (03-05/2014) 1,076 académicos y no-académicos de 12 dependencias de la UNAM (ICN, IFC, FC, IB, II, IG, IF, IM, IIMAS) 2,524 variables - Genetic, epidemiological, physiological,...

Epidemiological: Personal (81), Personal history (130), Family History (548), Self-health evaluation (226), Nutrition (220), Lifestyle (390), Health knowledge (293); **Genetic** (772); **Anthropometric and physiological** (49).

Fase II: (2017-2018) 282 estudiantes de carrera de la FM; (06/17) 400 trabajadores y profesores de la FM. Agregando variables psicológicas.

Fase III: (2019) Seguimiento de los 1,076 de Fase I. Otro análisis de sangre y agregando variables psicológicas, datos detallados de “un día en tu vida”, más variables fisiológicos en tiempo,

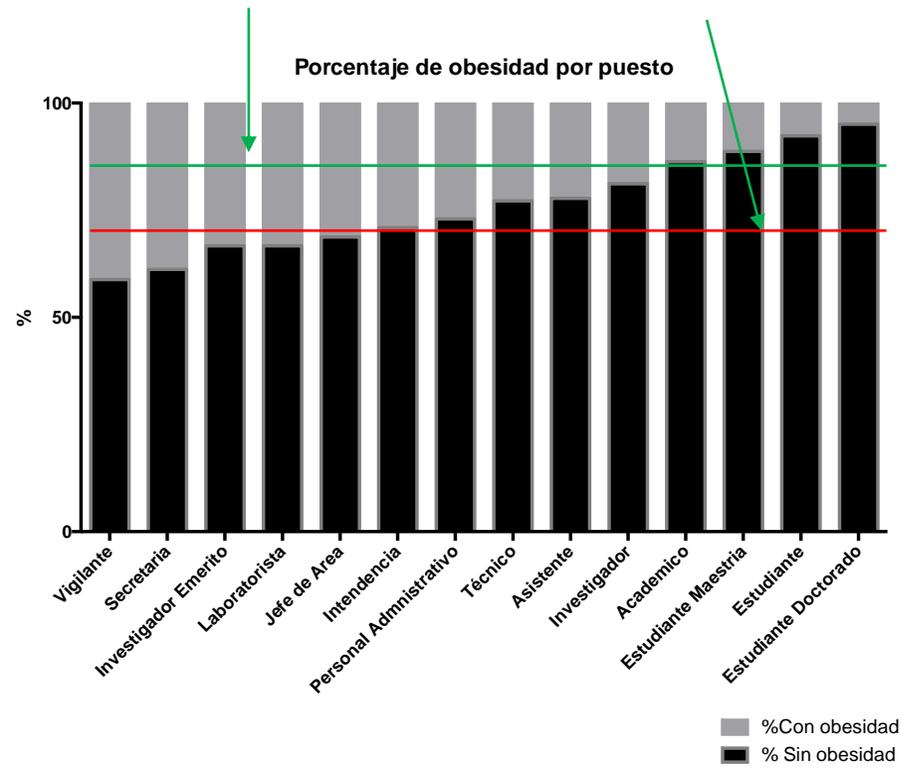
Fase IV: (2020-21) 600 estudiantes de carrera de la FM, Fac. Psic., Ibero, FES Zaragoza
Múltiples instrumentos psicológicos. Impacto de COVID 19. Vínculo con programa de Salud en tu Vida del SECTEI.

En total tenemos datos profundos de más de 3000 sujetos. Varios de estos datos ya están públicos.

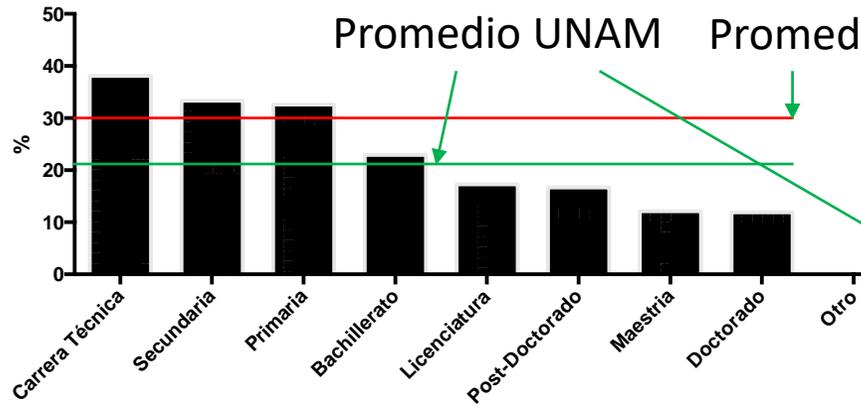
Promedio UNAM

Promedio nacional

Porcentaje de obesidad por puesto



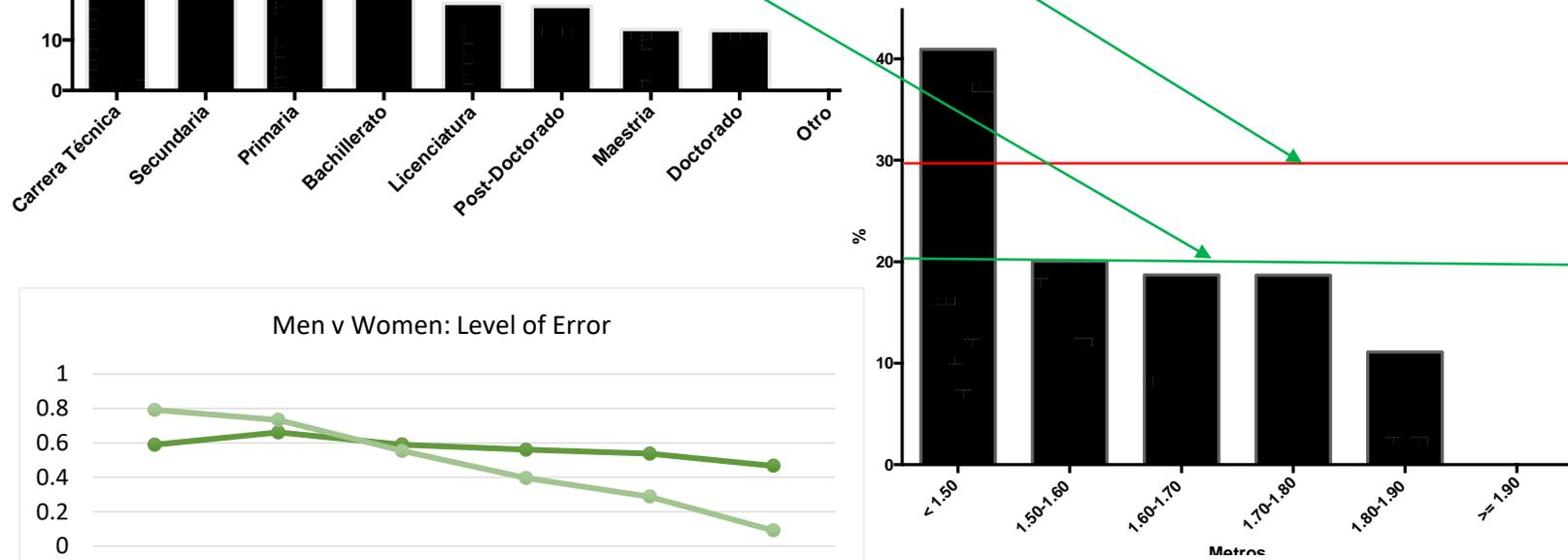
Proporción de obesos por escolaridad



Promedio UNAM

Promedio nacional

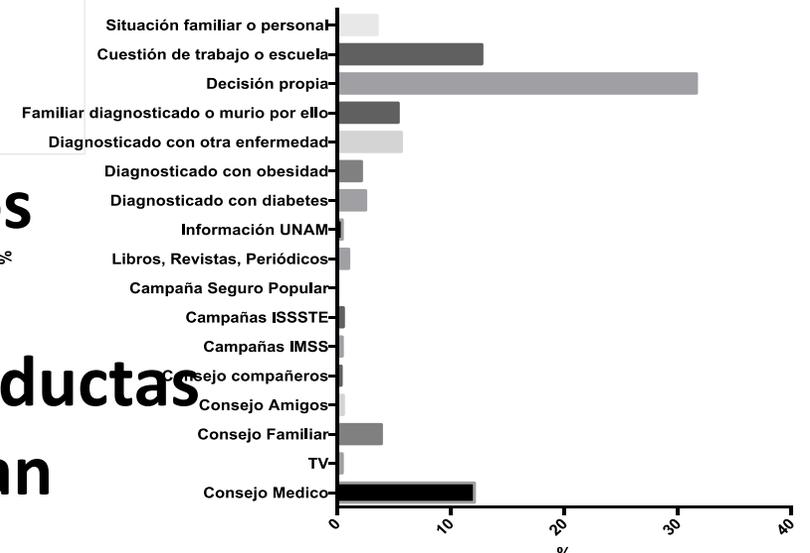
Proporción de obesos (Estatura)



Men v Women: Level of Error

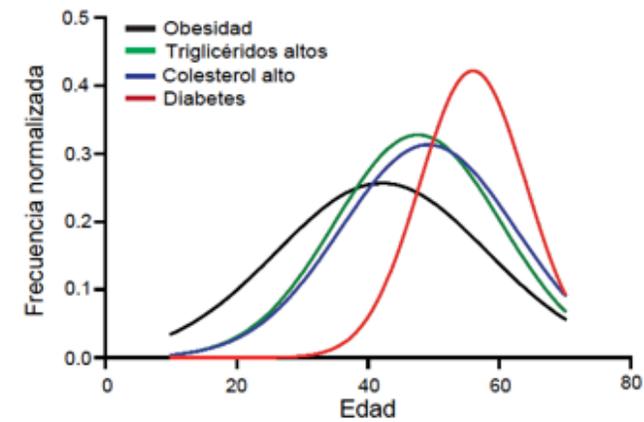


Si has cambiado tus hábitos (nutrición, ejercicio, estilo de vida etc.) ¿Porqué los cambiaste?



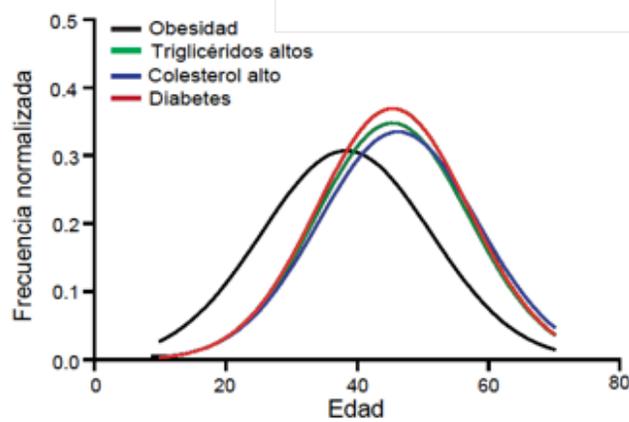
A

Académicos



B

No académicos



Atrás de estos resultados están las conductas que los causan

3) Modelos de predicción de Aprendizaje de Máquinas

Un Conductoma explicita: “Haces ejercicio entre semana?”

- $P(C = \text{Haces ejercicio entre semana} = \text{NO} \mid \text{factores extrínsecos e intrínsecos } \mathbf{X})$

- $S(C|\mathbf{X}) = \ln\left(\frac{P(C|\mathbf{X})}{P(\bar{C}|\mathbf{X})}\right) = \sum_{i=1}^N s_i + \ln(P(C)/P(\bar{C}))$

donde $s_i = \ln\left(\frac{P(X_i|C)}{P(X_i|\bar{C})}\right)$ es el score/peso del factor de riesgo X_i

- Si el conjunto de factores \mathbf{X} del Conductoma es tal que $S(C|\mathbf{X}) > 0$ así estos factores indicant mayor propensidad a ser sedentario, en que la persona es más propensa no hacer ejercicio entre semana y al contrario, si $S(C|\mathbf{X}) < 0$.

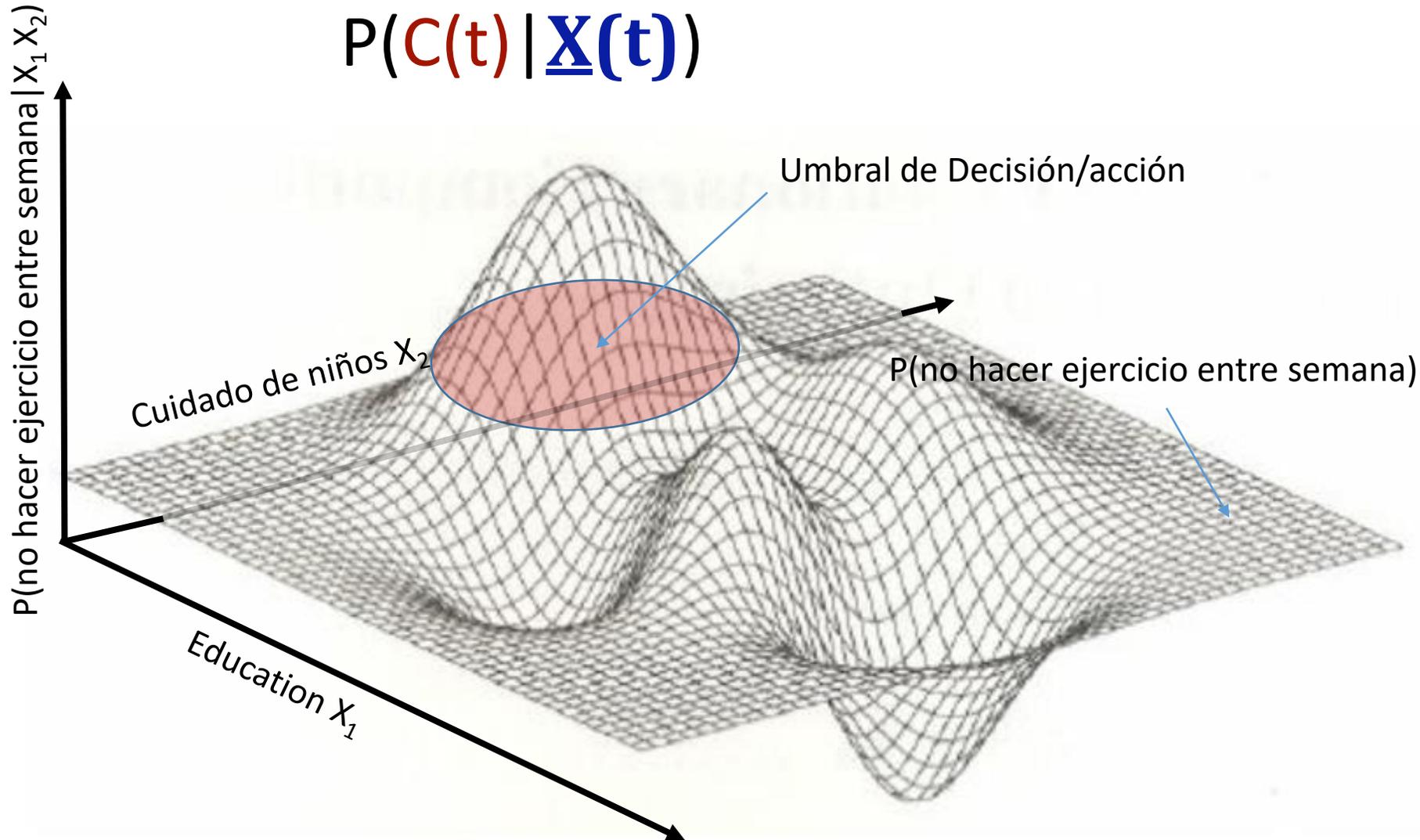
Los factores más importantes en el Conductoma para la conducta – No haces ejercicio entre semana

Pregunta	Respuesta	Número de personas con X	Número de personas que no hacen ejercicio y X	N	Nc	% que no hacen ejercicio	% que no hacen ejercicio y X	Predictive model weight (score)	Statistical reliability (Epsilon)	Causa o consecuencia
¿Qué quehaceres realiza?: Cuidado de niños	Sí	41	29	292	120	41.10%	70.73%	0.38	3.86	Causa
¿Qué tan regular es su horario para ir a dormir?	1 - 2 hrs	17	14	292	120	41.10%	82.35%	0.67	3.46	Ambos
¿Realiza ejercicio en fin de semana?	No	182	97	292	120	41.10%	53.30%	0.06	3.35	Ambos
¿A qué hora se transporta a su casa?	15:00	32	22	292	120	41.10%	68.75%	0.34	3.18	Causa
¿Cuántas horas duerme entre semana?	4-5 horas	65	39	292	120	41.10%	60.00%	0.18	3.1	Consecuencia
¿Aproximadamente cuantas horas libres tiene al día entre semana?: No sé	Sí	24	17	292	120	41.10%	70.83%	0.39	2.96	Causa
¿Dónde come entre semana?: Posición 2	En puestos de la calle	9	8	292	120	41.10%	88.89%	0.9	2.91	Ambos
¿Qué quehaceres realiza?: Lavar el baño	Sí	172	89	292	120	41.10%	51.74%	0.03	2.84	Causa
¿Qué quehaceres realiza?: Sacudir	Sí	158	82	292	120	41.10%	51.90%	0.03	2.76	Causa
¿En qué tipo de vehículo se transporta de su casa al trabajo? y ¿Cuánto dura cada uno aproximadamente EN MINUTOS?: Metro: Valor	60 min	11	9	292	120	41.10%	81.82%	0.65	2.75	Causa
¿Cómo consigue sus colaciones?: La compro en un puesto	Sí	50	30	292	120	41.10%	60.00%	0.18	2.72	Ambos
¿Qué quehaceres realiza el fin de semana?: Cuidado de niños	Sí	60	35	292	120	41.10%	58.33%	0.15	2.71	Causa
¿Dónde desayuna? Seleccione por orden de frecuencia.: Posición 1	En la cocina del trabajo	27	18	292	120	41.10%	66.67%	0.3	2.7	Ambos

Los factores más importantes en el Conductoma para la conducta – Haces ejercicio entre semana

Pregunta	Valor	Respuesta	Número de personas con X	Número de personas que no hacen ejercicio y X	% que no hacen ejercicio	% que no hacen ejercicio y X	Predictive model weight (score)	Statistical reliability (Epsilon)
¿En qué tipo de vehículo se transporta de su casa al trabajo? y ¿Cuánto dura cada uno aproximadamente EN MINUTOS?: Auto propio	Y	Sí	164	55	41.10%	33.54%	-0.3	-1.97
Cintura	(8.199, 76.28]		30	7	41.10%	23.33%	-0.52	-1.98
¿Aproximadamente cuantas horas libres tiene al día entre semana?: Tarde (En Horas)	Y	Sí	111	35	41.10%	31.53%	-0.34	-2.05
¿Aproximadamente cuantas horas libres tiene al día en fin de semana?: Noche (En Horas): Valor	2		40	10	41.10%	25.00%	-0.48	-2.07
¿Dónde realiza la mayoría de su ejercicio?	A2	Calle	18	3	41.10%	16.67%	-0.7	-2.11
¿Cómo considera que es su comida?	A4	Ligero	25	5	41.10%	20.00%	-0.6	-2.14
¿Cómo realiza su jornada laboral? y ¿Cuánto tiempo (en HORAS) aproximadamente? En movimiento: Valor	6		16	2	41.10%	12.50%	-0.85	-2.32
¿Aproximadamente cuantas horas libres tiene al día entre semana?: Mañana (En Horas): Valor	1		16	2	41.10%	12.50%	-0.85	-2.32
¿Aproximadamente cuantas horas libres tiene al día entre semana?: Tarde (En Horas): Valor	2		30	6	41.10%	20.00%	-0.6	-2.35
¿Aproximadamente cuantas horas libres tiene al día entre semana?: Mañana (En Horas)	Y	Sí	40	8	41.10%	20.00%	-0.6	-2.71
¿En qué tipo de vehículo se transporta de su casa al trabajo? y ¿Cuánto dura cada uno aproximadamente EN MINUTOS?: Auto propio: Valor	30		27	4	41.10%	14.81%	-0.76	-2.78
Grado de estudios	Doctorado		47	9	41.10%	19.15%	-0.63	-3.06
Puesto	Investigador		24	2	41.10%	8.33%	-1.04	-3.26
¿Realiza ejercicio en fin de semana?	Y	Sí	107	23	41.10%	21.50%	-0.56	-4.12

$$P(C(t) | \underline{X}(t))$$



Esta construcción es puramente fenomenológica – se pone todo (\underline{X}) dentro de la máquina $P(C|X)$ y vemos que sale. Esta es la parte de Inteligencia Artificial. El papel de la Inteligencia Humana es interpretar lo que significa. La parte más vital es la búsqueda de la causalidad.

El paisaje tiene un sinfín de dimensiones – genéticas, epigenéticas, fisiológicas, psicológicos, sociológicos, ambientales, económicos, políticos,...

- **El Proyecto del Conductoma y la construcción de un Conductoma específico es factible** – se puede calcular $P(C|X)$ de los datos de Project 42 usando algoritmos Bayesianos de Aprendizaje de Máquinas en una manera que exhibe predictabilidad, explicabilidad y utilidad.
- Hay muchas conductas por modelar.
- Hay un sinfin de factores que contribuyen a un Conductoma particular- ¿Cómo se supera la disciplinarietà asociado con concentrar en un subconjunto y ignorar los demás?
- ¿Cómo se pasa de un acercamiento fenomenológico $P(C|X)$ a un entendimiento más “mecánico” y causal de nuestros modelos mentales? $P(C|X)$ es una ventana indirecta pero ponderosa. ¿Cómo vamos más allá que eso?

Actividad física
- actigrafía.
Estefa Espítia

El Conductoma de la OBESIDAD

Factores culturales
Rolando Diaz-Loving

Locus de control
Leonardo Solis

Efecto socio-economico
Dagmara Wrzecionkowska

Alimentación
infantil
Gabriela Guevarra

Modelos
mentales
Miriam Barajas

