

Recomendaciones: Todo es un Ensemble



Netflix prize was won by an ensemble of 107 models (2007):

- Matrix factorization
- Restricted Boltzman Machines
- k-NN
- Regression models

The BellKor solution to the Netflix Prize

Robert M. Bell, Yehuda Koren and Chris Volinsky
AT&T Labs – Research
BellKor@research.att.com

Our final solution ($\text{RMSE}=0.8712$) consists of blending 107 individual results. Since many of these results are close variants, we first describe the main approaches behind them. Then, we will move to describing each individual result.

The core components of the solution are published in our ICDM'2007 paper [1] (or, KDD-Cup'2007 paper [2]), and also in the earlier KDD'2007 paper [3]. We assume that the reader is familiar with these works and our terminology there.

Problemas con las Recomendaciones

The screenshot shows the TripAdvisor website for Santiago de Compostela, Spain. The top navigation bar includes links for Opinion, Unirse, Iniciar Sesión, and currency conversion. The main search bar shows "Encuentra: Qué hacer" and "Cerca de: Santiago de Compostel ...". Below the search bar, the breadcrumb navigation shows: Europa > España > Galicia > Provincia de A Coruña > Santiago de Compostela > Qué hacer en Santiago de Compostela.

Cosas que hacer en Santiago de Compostela

Ordenar por: Clasificación Reservar en línea

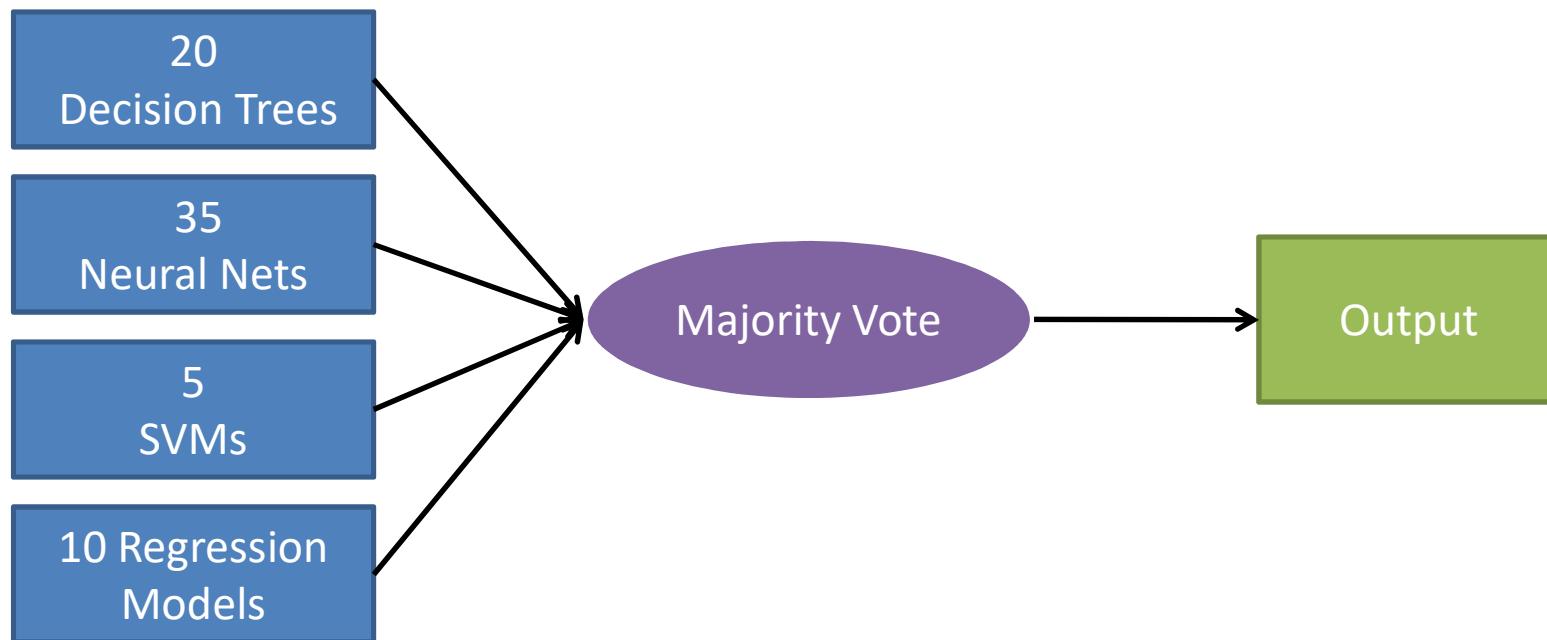
Actividad	Imagen	Detalles
Catedral de Santiago de Compostela		1 de 84 cosas que hacer en Santiago de Compostela CCCCCC 6.250 opiniones "Imponente" 26/10/2016 "Es imponente este templo. Por su e..." 25/10/2016 Edificios con valor arquitectónico desde EUR 42,09 €* Detalles
Botafumeiro		2 de 84 cosas que hacer en Santiago de Compostela CCCCCC 1.620 opiniones "Impresiona" 26/10/2016 "Expectacular" 25/10/2016 Puntos emblemáticos y de interés 1 visita guiada disponible desde EUR 560,00 €* Detalles
Casco histórico de Santiago de Compostela		3 de 84 cosas que hacer en Santiago de Compostela CCCCCC 1.153 opiniones "Muy eclesiástico" 25/10/2016 "Con mucho encanto!" 23/10/2016 Puntos emblemáticos y de interés 1 visita guiada disponible desde EUR 380,00 €* Detalles

TIPO DE ATRACCIÓN:

- Monumentos y puntos de interés (39)
- Visitas guiadas (26)
- Museos (10)
- Naturaleza y parques (7)
- Compras (12)
- Vida nocturna (54)
- Conciertos y espectáculos (1)

“Users do not trust explicit recommendations” (2016)

Problemas con las Recomendaciones



¡¡Cómo lo explicamos!!

Fundamentos teóricos

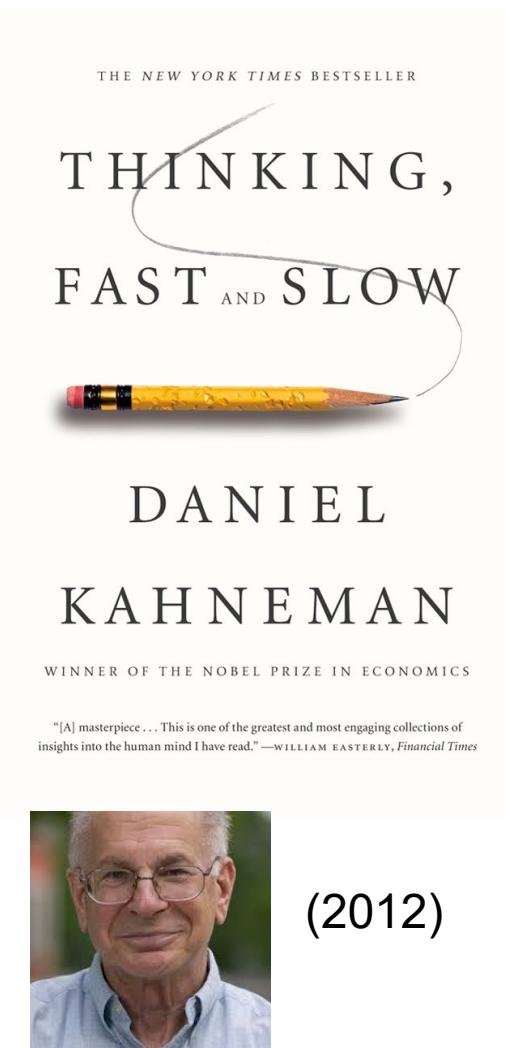
Más fácil con una teoría

Nuestra mente tiene dos sistemas:

- Sistema I: Rápido, automático, poco esfuerzo, sin control consciente.
- Sistema II: Lento, racional, costoso, con control consciente.

Psicólogos ante la teoría económica:

- Econos Vs. Humanos.
- El poder de la multidisciplinariedad



Tomando decisiones



Referencia: Brain Games

Economist.com	SUBSCRIPTIONS
OPINION	
WORLD	
BUSINESS	
FINANCE & ECONOMICS	
SCIENCE & TECHNOLOGY	
PEOPLE	
BOOKS & ARTS	
MARKETS & DATA	
DIVERSIONS	

Welcome to
The Economist Subscription Centre

Pick the type of subscription you want to buy or renew.

Economist.com subscription - US \$59.00
One-year subscription to Economist.com.
Includes online access to all articles from *The Economist* since 1997.

Print subscription - US \$125.00
One-year subscription to the print edition of *The Economist*.

Print & web subscription - US \$125.00
One-year subscription to the print edition of *The Economist* and online access to all articles from *The Economist* since 1997.

Referencia: (Ariely, 2010)

¿Qué es la toma de decisiones?

Definición 1:

The thought process of selecting a logical choice from the available options.

When trying to make a good decision, a person must weight the positives and negatives of each option, and consider all the alternatives. For effective decision making, a person must be able to forecast the outcome of each option as well, and based on all these items, determine which option is the best for that particular situation

Fuente: <http://www.businessdictionary.com/>

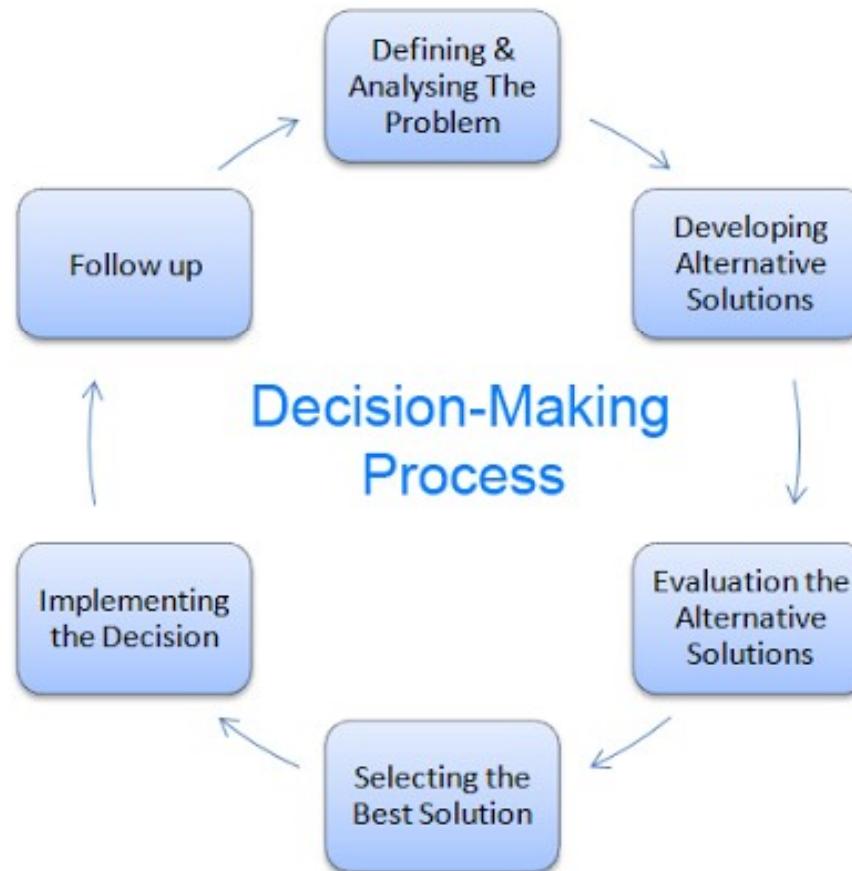
Definición 2:

Decision-making can be regarded as a problem-solving activity terminated by a solution deemed to be optimal, or at least satisfactory. It is therefore a process which can be more or less rational or irrational and can be based on explicit or tacit knowledge and beliefs.

Fuente: <https://en.wikipedia.org/wiki/Decision-making>

Decisiones racionales

Teoría de elección racional: el proceso



Fase I: Definiendo el problema



Tipo de problema: Informal

Tipo de problema: Serio



Ejemplo: ¿Tipo de problema?



2 euros

3.25 euros

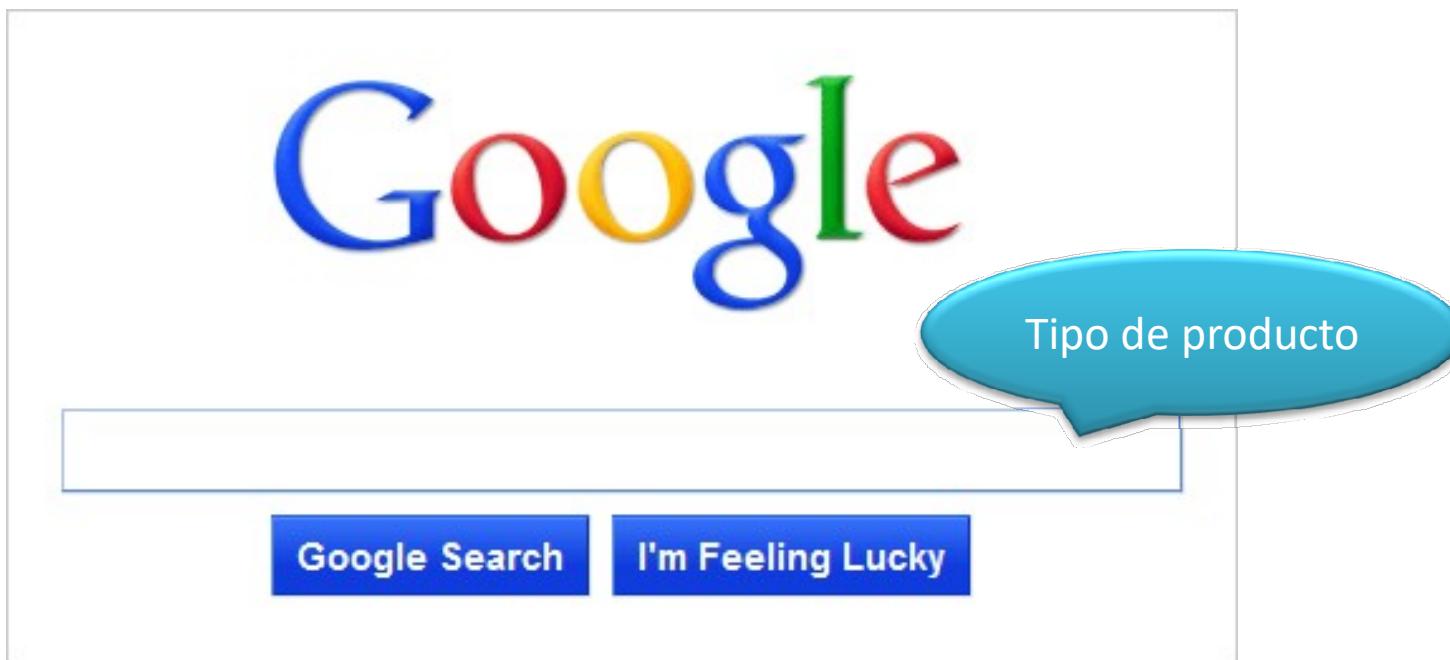
3.30 euros

Teoría de elección racional: Fase II

Fuentes de información:
Mi memoria



Fuentes de información: Buscadores



Fuentes de información: Tiendas online o tiendas físicas

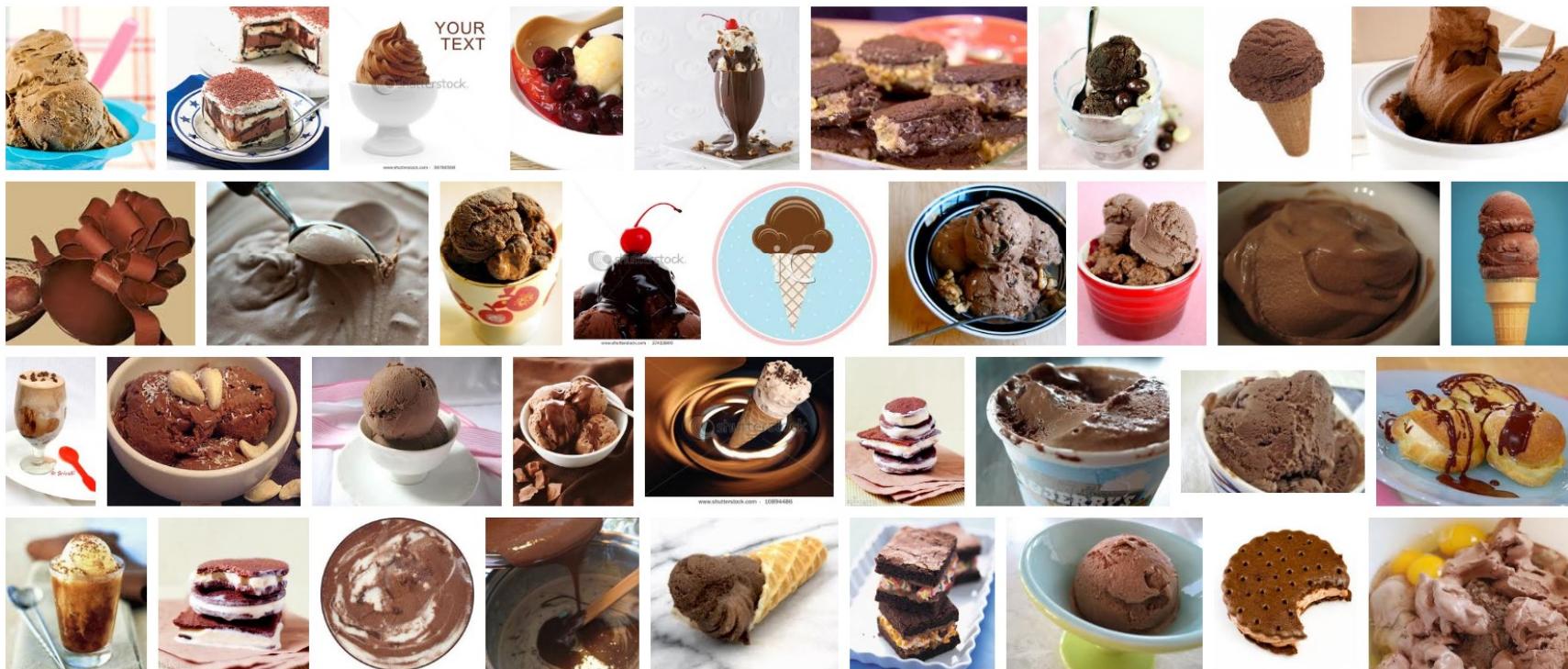


Fuentes de información: Mi red social

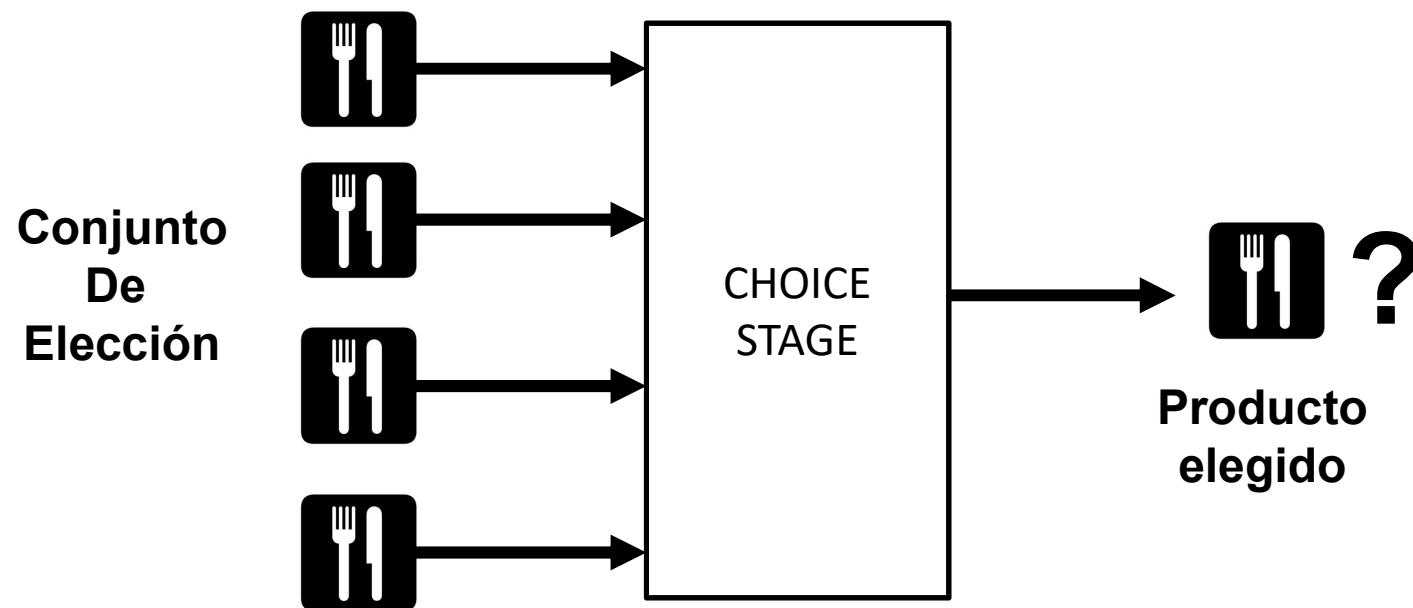


Teoría de elección racional: Fase III

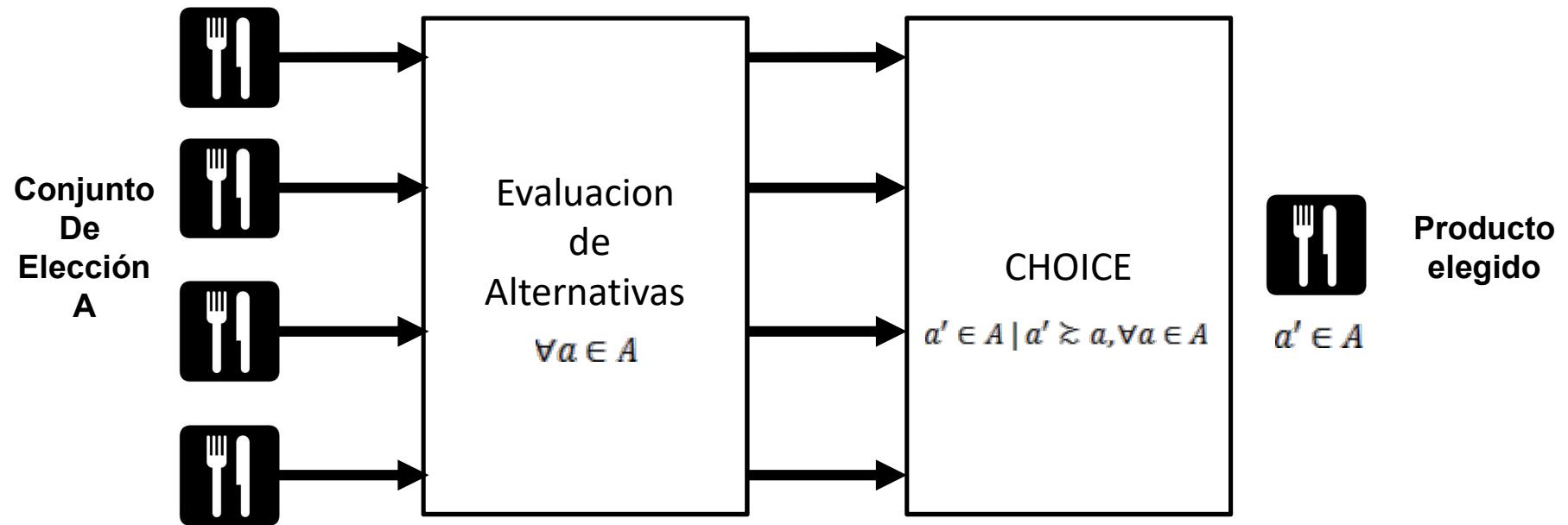
El Conjunto de Elección



Fase III y IV: El problema de la Elección



Fase III y IV: El problema de la Elección



Regla de Elección con Preferencias

$$RE(A, \geq) = \{a' \in A \mid a' \geq a, \forall a \in A\}$$

Regla de Elección con Preferencias

$$RE(A, \succsim) = \{a' \in A \mid a' \succsim a, \forall a \in A\}$$

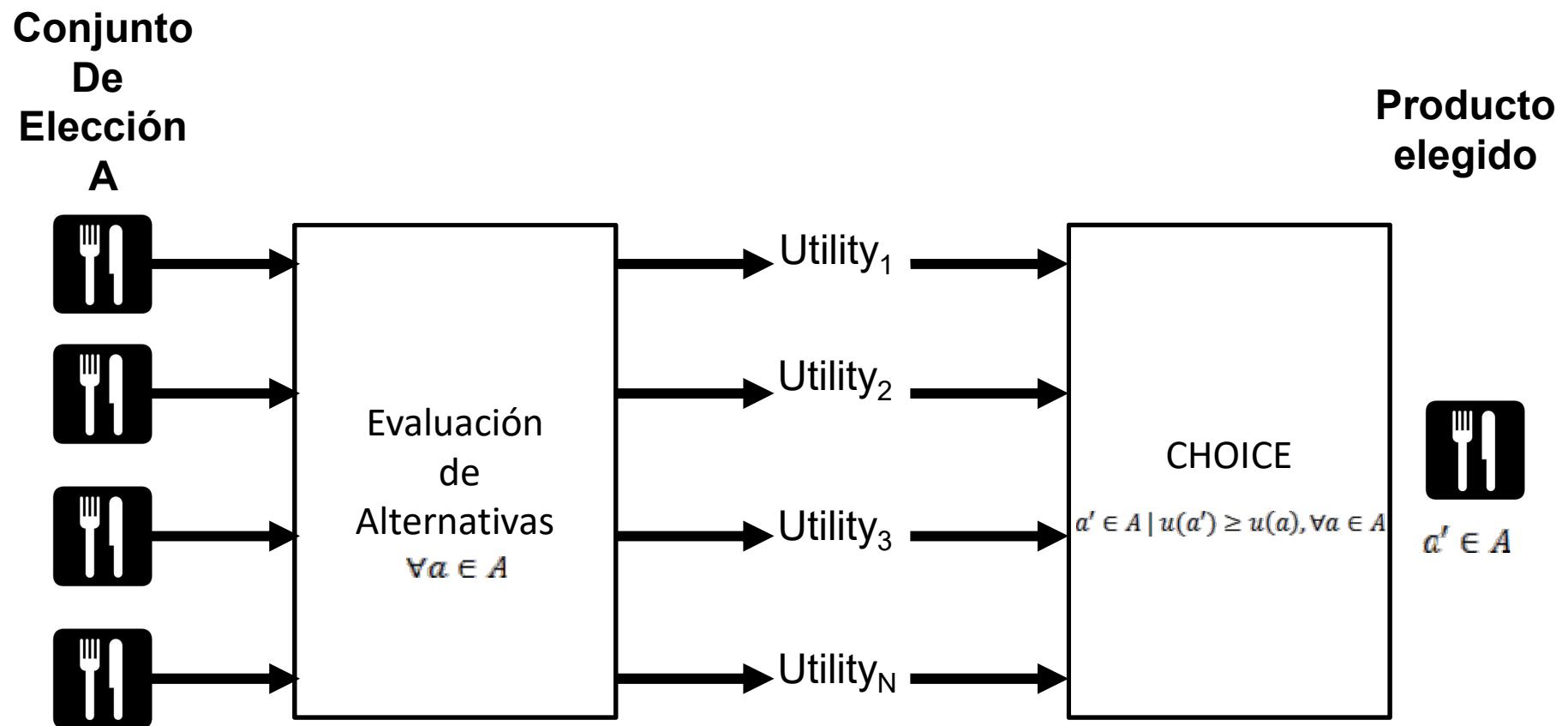
Axioma de la Teoría de la Utilidad

$$a \succsim b \Leftrightarrow u(a) \geq u(b)$$

Regla de Elección con Utilidades

$$RE(A, \geq) = \{a' \in A \mid u(a') \geq u(a), \forall a \in A\}$$

Teoría de la Utilidad



- ▶ ¿Qué es la Utilidad?:
 1. Perspectiva de Marshall:
 - Funcionalidad de una alternativa.
 - Se estima comparando los atributos de una alternativa con las preferencias del sujeto.
 - Concepto de Utilidad estándar.
 2. Perspectiva de Bentham:
 - Satisfacción que se espera obtener de una alternativa.
 - Concepto original de Utilidad.

- ▶ ¿Cómo construimos la función de utilidad?:
 1. Identificar los atributos relevantes de las alternativas
 2. Identificar los valores de los atributos anteriores
 3. Estimar las preferencias del individuo sobre los valores de los atributos.
 4. Matemáticamente: Función lineal sobre los valores de los atributos. Dado individuo c y alternativa a:

$$u(c,a) = \sum_k \beta_{c,k} x_{a,k}$$

Con K indicando el cto de valores de todos los atributos de a.

Teoría de la Utilidad

Ejercicio: ¿Cómo comparas entre estas alternativas?



Centro ciudad
20 mins de playa
100 m²
3 habitaciones
210.000 euros

10 min coche del centro
15 mins de playa
160 m²
4 habitaciones
250.000 euros

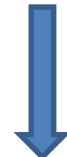
20 min coche del centro
Pie de playa
90 m²
3 habitaciones
150.000 euros

Teoría de la Utilidad

Estimate utilities

Alternatives	Short Distance	Medium Distance	Long Distance	Small Size	Big Size	Low Price	High Price
$A_1 - X(A_1)$	1	0	0	1	0	0	1
$A_2 - X(A_2)$	0	1	0	0	1	0	1
$A_3 - X(A_3)$	0	0	1	1	0	1	0

Preferences	Short Distance	Medium Distance	Long Distance	Small Size	Big Size	Low Price	High Price
My prefs - B	6	3	1	2	8	6	4



$$u(c, a) = \sum_k \beta_{c,k} x_{a,k}$$

Teoría de la Utilidad

Building a table

Alternatives	Utility	MU	Chosen Alternative
A ₁	12		
A ₂	15	15	A ₂
A ₃	9		

Conexión con estrategias de recomendación

- ▶ En la teoría de la elección racional, la utilidad es una función lineal sobre los valores de los atributos. Dado individuo c y un ítem a:

$$u(c, a) = \sum_k \beta_{c,k} x_{a,k}$$

Con x indicando los valores del atributo k del ítem a

Con b indicando la preferencia del individuo c sobre x

Con K indicando el cto de todos los valores de los atributos de a

- ▶ En la estrategia basada en contenido, la utilidad se puede calcular a través de la similaridad, y ésta a través de la medida del coseno:

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

- ▶ Por tanto:

utilidad(estrategia_basada_contenido) = utilidad(eleccion_racional) normalizada

- ▶ ¿Qué significa esta conexión?

La estrategia de recomendación basada en contenido asume que la toma de decisiones de los humanos es de tipo racional

Teoría de elección racional: Fase V

Fase V: Experiencia con el producto



Teoría de elección racional: Fase V

Fase V: Experiencia con el producto

ACME Restaurant Customer Survey

Thank you in advance for taking the ACME Restaurant survey. Please cross the box which is most relevant to your experience.

How satisfied are you with your experience at the ACME restaurant today?

	Very dissatisfied 1	-	-	-	-	Very satisfied 10				
The cleanliness of the restaurant	1	2	3	4	5	6	7	8	9	10
Ease of booking a table	1	2	3	4	5	6	7	8	9	10
The decor of the restaurant	1	2	3	4	5	6	7	8	9	10
The breadth of the food menu	1	2	3	4	5	6	7	8	9	10
The breadth of the wine menu	1	2	3	4	5	6	7	8	9	10
Catering for special diets	1	2	3	4	5	6	7	8	9	10
Quality of your eating experience	1	2	3	4	5	6	7	8	9	10
Quality of service from staff	1	2	3	4	5	6	7	8	9	10
The speed of service	1	2	3	4	5	6	7	8	9	10

How important or unimportant are the following requirements for you when visiting the ACME Restaurant?

	Totally unimportant 1	-	-	-	-	Very important 10				
The cleanliness of the restaurant	1	2	3	4	5	6	7	8	9	10
Ease of booking a table	1	2	3	4	5	6	7	8	9	10
The decor of the restaurant	1	2	3	4	5	6	7	8	9	10
The breadth of the food menu	1	2	3	4	5	6	7	8	9	10
The breadth of the wine menu	1	2	3	4	5	6	7	8	9	10
Catering for special diets	1	2	3	4	5	6	7	8	9	10
Quality of your eating experience	1	2	3	4	5	6	7	8	9	10
Quality of service from staff	1	2	3	4	5	6	7	8	9	10
The speed of service	1	2	3	4	5	6	7	8	9	10

Predicción: el problema del científico

Ejercicio: ¿Con qué atributos representarías estas alternativas?
¿Cómo harías la predicción considerando TUS preferencias?



El problema de la incertidumbre

Problema: ¿Hay completa seguridad de predecir el resultado/utilidad una vez tomada una decisión? ¿Es posible que no siempre se obtenga el mismo resultado/utilidad?

Ejemplo de decisión con incertidumbre: ¿Qué alternativa prefieres?

Opción A:

70% probabilidad de ganar 1.000 euros
30% probabilidad de no ganar nada

Opción B:

50% probabilidad de ganar 500 euros
50% probabilidad de ganar 200 euros.

Opción A:

33% probabilidad de ganar 2.500 euros
67% probabilidad de no ganar nada

Opción B:

34% probabilidad de ganar 2.400 euros
66% probabilidad de no ganar nada.

Teoría de la Utilidad Esperada

Problema: ¿Hay completa seguridad de predecir el resultado/utilidad una vez tomada una decisión? ¿Es posible que no siempre se obtenga el mismo resultado/utilidad?

Solución: *Teoría de la Utilidad Esperada*

Principio MEU (Maximum Expected Utility):

$$RE(A, \geq) = \{a' \in A \mid \hat{u}(a') \geq \hat{u}(a), \forall a \in A\} \Leftrightarrow a' = \arg \max_{a \in A} \hat{u}(c, a)$$

Con c indicando el usuario, a una alternativa, y \hat{u} la utilidad esperada (expected Utility).

La utilidad esperada \hat{u} se calcula en base a los posibles resultados de una Decision. Cada resultado r tiene una probabilidad de ocurrencia y una utilidad asociada. Por tanto, la utilidad esperada \hat{u} asociada a la elección de la alternativa a se calcula:

$$\hat{u}(c, a) = \sum_r P(Result = r) \hat{u}(c, r)$$

Teoría de la Utilidad Esperada

Ejercicio: ¿Qué alternativa prefieres?

Opción A:

$$0.7 \times 1000$$

70% probabilidad de ganar 1.000 euros
30% probabilidad de no ganar nada

$$0.5 \times 500 + 0.5 \times 200$$

Opción B:

50% probabilidad de ganar 500 euros
50% probabilidad de ganar 200 euros.

Opción A:

33% probabilidad de ganar 2.500 euros
67% probabilidad de no ganar nada

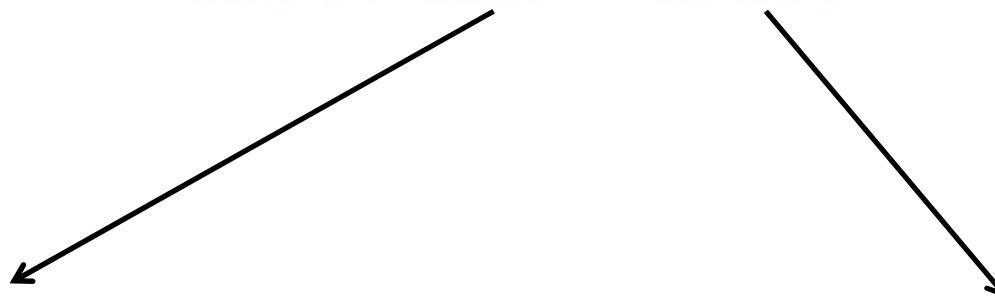
Opción B:

34% probabilidad de ganar 2.400 euros
66% probabilidad de no ganar nada.

Predicción: el problema del científico

Problema: En muchas situaciones, el científico no conoce directamente los valores de $P(\text{Result}=r)$ y/o de utilidad esperada \hat{u} :

$$\hat{u}(c, a) = \sum_r P(\text{Result} = r) \hat{u}(c, r)$$



¿Y si esta probabilidad no es constante, sino que puede variar si tenemos en cuenta evidencias disponibles del problema?

¿Y si esta utilidad depende de factores del usuario que el científico desconoce?

Estimando probabilidades con datos



Problema: ¿Y qué pasa si la probabilidad $P(\text{Result}=r)$ está condicionada por evidencias o factores externos? Es decir, tenemos que estimar $P(\text{Result}=r | \text{Evidencias})$.

Ejercicio:

Un taxi es el responsable de atropellar a un peatón de noche, dándose a la fuga tras el incidente. Dos compañías de taxi, la Verde y la Azul, son las que operan en la ciudad.

Te dan los siguientes datos:

1. El 85% de los taxis en la ciudad son Verdes y el 15% de ellos son Azules.
2. Un testigo identificó el taxi como un taxi Azul. En la investigación del suceso se ha comprobado la precisión del testigo en circunstancias similares a la de aquella noche, y se concluyó que el testigo identifica correctamente cada uno de los dos colores el 80% de los casos, y falla el 20% restante.

Cual es la probabilidad de que el taxi responsable del accidente sea Verde o Azul?

Problema: ¿Y qué pasa si la probabilidad $P(\text{Result}=r)$ está condicionada por el contexto o por otros factores externos? Es decir, tenemos que estimar $P(\text{Result}=r | \text{Evidencias})$.

Solución: *Teorema de Bayes e inferencia bayesiana*

Teorema de Bayes:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

Teorema de Bayes

¿De dónde viene el Teorema de Bayes?:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A|B) P(B) = P(A \cap B) = P(B|A) P(A)$$

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Interpretación del Teorema de Bayes con Hipótesis y Evidencias:

$$P(H/E) = \frac{P(E/H) \cdot P(H)}{P(E)}$$

where $P(H)$ = the previous or *a priori* probability that the hypothesis is true

$P(E)$ = the probability that an event will occur

$P(E/H)$ = the probability that the event will occur given that the hypothesis is true

Y el denominador es un factor de normalización que se calcula:

$$P(E) = \sum_i P(E|H_i) \cdot P(H_i)$$

Metodología: (1) Identificar las hipótesis, (2) Identificar las evidencias, (3) Identificar las probabilidades a priori, (4) Identificar las verosimilitudes, (5) Normalizar la ecuación

Ejercicio:

Un taxi es el responsable de atropellar a un peatón de noche, dándose a la fuga tras el incidente. Dos compañías de taxi, la Verde y la Azul, son las que operan en la ciudad.

Te dan los siguientes datos:

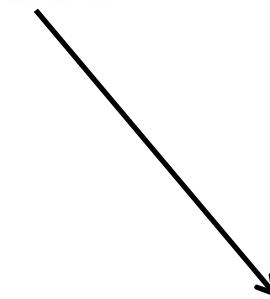
1. El 85% de los taxis en la ciudad son Verdes y el 15% de ellos son Azules.
2. Un testigo identificó el taxi como un taxi Azul. En la investigación del suceso se ha comprobado la precisión del testigo en circunstancias similares a la de aquella noche, y se concluyó que el testigo identifica correctamente cada uno de los dos colores el 80% de los casos, y falla el 20% restante.

Cual es la probabilidad de que el taxi responsable del accidente sea Verde o Azul?

Utilidades con variables no observadas

Problema: ¿Y qué pasa si la utilidad esperada depende de factores que el científico desconoce?

$$\hat{u}(c, a) = \sum_r P(\text{Result} = r) \hat{u}(c, r)$$



¿Y si esta utilidad depende de factores que el científico desconoce?

Utilidades con variables no observadas

Problema: ¿Y qué pasa si la utilidad esperada depende de factores que el científico desconoce?

Solución: *Modelos de utilidad aleatorios (MUA)*

$$u(c, a) = u_r(c, a) + \varepsilon_{c,a}$$

- $u_r(c, a)$ = utilidad determinista. Variable que puede calcular el científico utilizando los datos disponibles.
 - ε = utilidad no determinista. Variable aleatoria que depende de factores no observables.

¡¡¡La variable $u(c, a)$ se convierte entonces en una variable aleatoria!!!

El Principio de Decisión o Regla de Elección, queda:

$$RE(A, \geq) = \{a' \in A \mid P(\text{eleccion}(c) = a') \geq P(\text{eleccion}(c) = a), \forall a \in A\}$$

Y las probabilidades se calculan:

$$\begin{aligned} P(\text{eleccion}(c) = a') &= P(u(c, a') > u(c, a), \forall a \in A \text{ y } a \neq a') \\ &= P(u_r(c, a') + \varepsilon_{c,a'} > u_r(c, a) + \varepsilon_{c,a}, \forall a \in A \text{ y } a \neq a') \\ &= P(\varepsilon_{c,a} - \varepsilon_{c,a'} < u_r(c, a') - u_r(c, a), \forall a \in A \text{ y } a \neq a') \end{aligned}$$

Modelos de utilidad aleatorios

Y la probabilidad se calcula integrando sobre ε :

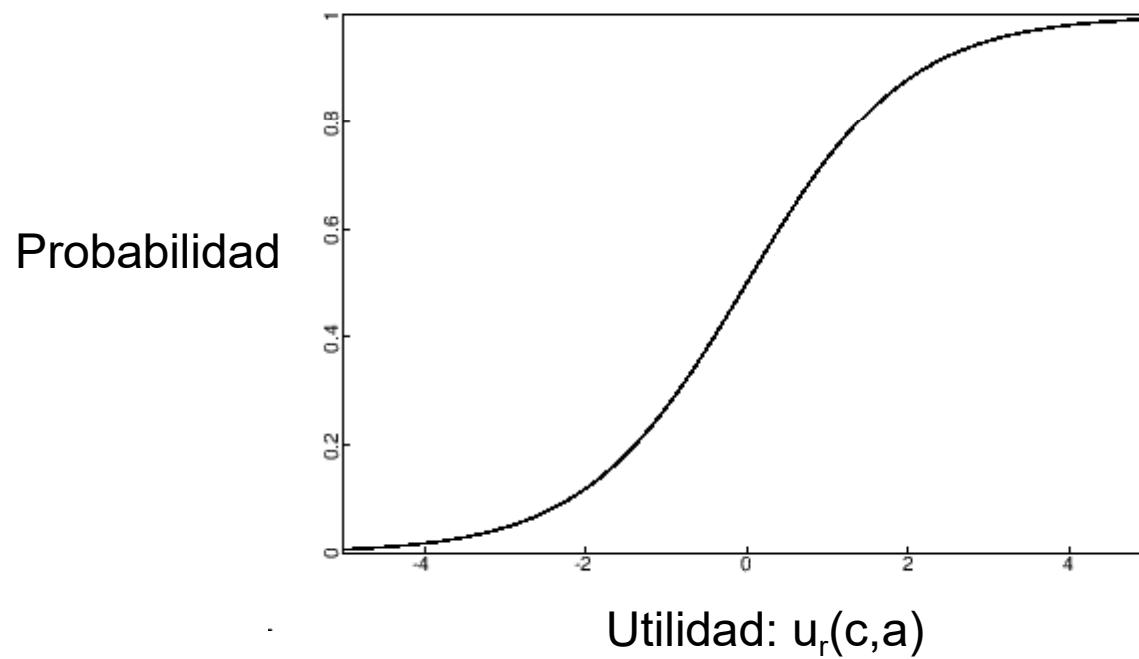
$$P(\text{elección}(c) = a') = \int_{\varepsilon_c = -\infty}^{\infty} I(\varepsilon_{c,a} - \varepsilon_{c,a'} < u_r(c, a') - u_r(c, a), \forall a \neq a') f(\varepsilon_c) d\varepsilon_c$$

Resolvemos la integral asumiendo una cierta distribución de probabilidad para ε , y obtenemos una función logística:

$$P(\text{elección}(c) = a') = \frac{e^{u_r(c, a')}}{\sum_j e^{u_r(c, a_j)}}$$

Modelos de utilidad aleatorios

Interpretación de la función logística:



Resultados



Rational choice-based models:

- Easy to explain
- Superior performance compared with basic baseline algorithms

Choice-based recommender systems

Paula Saavedra
CITIUS
University of Santiago de Compostela
Santiago de Compostela,
Spain
paula.saavedra@usc.es

Rosa Crujeiras
School of Mathematics
University of Santiago de Compostela
Santiago de Compostela,
Spain
rosa.crujeiras@usc.es

Pablo Barreiro
CITIUS
University of Santiago de Compostela
Santiago de Compostela,
Spain
pablobv70@gmail.com

Maria Loureiro
School of Business
University of Santiago de Compostela
Santiago de Compostela,
Spain
maria.loureiro@usc.es

Roi Durán
CITIUS
University of Santiago de Compostela
Santiago de Compostela,
Spain
roiduram@gmail.com

Eduardo Sánchez Vila
CITIUS
University of Santiago de Compostela
Santiago de Compostela,
Spain
eduardo.sanchez.vila@usc.es

ABSTRACT

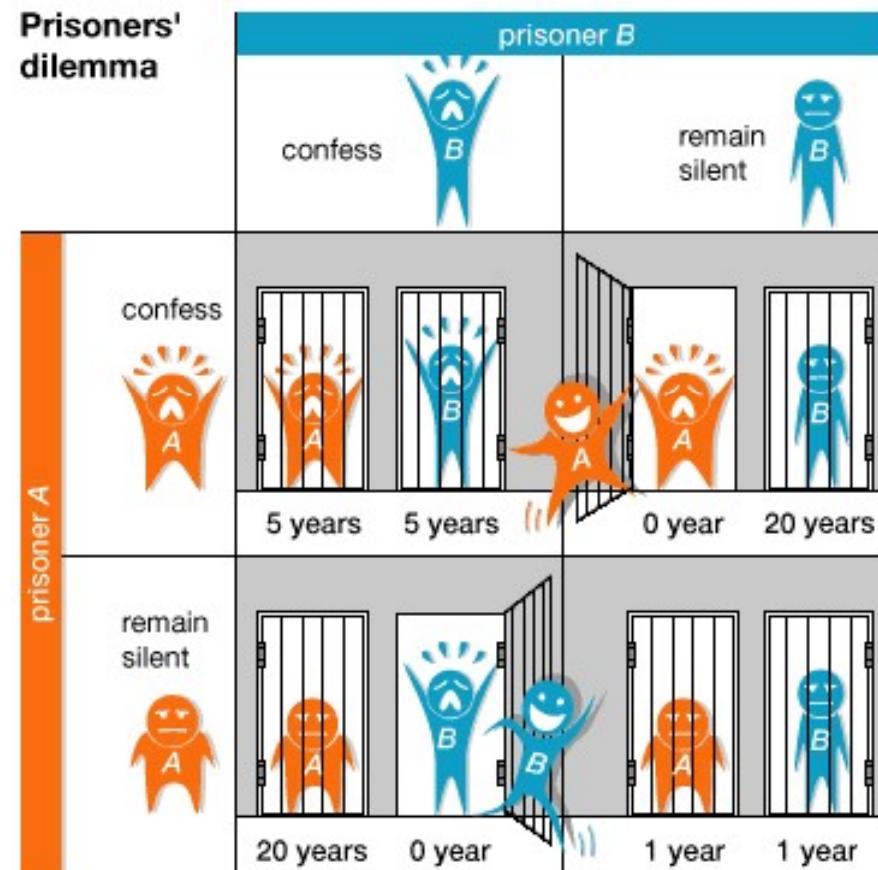
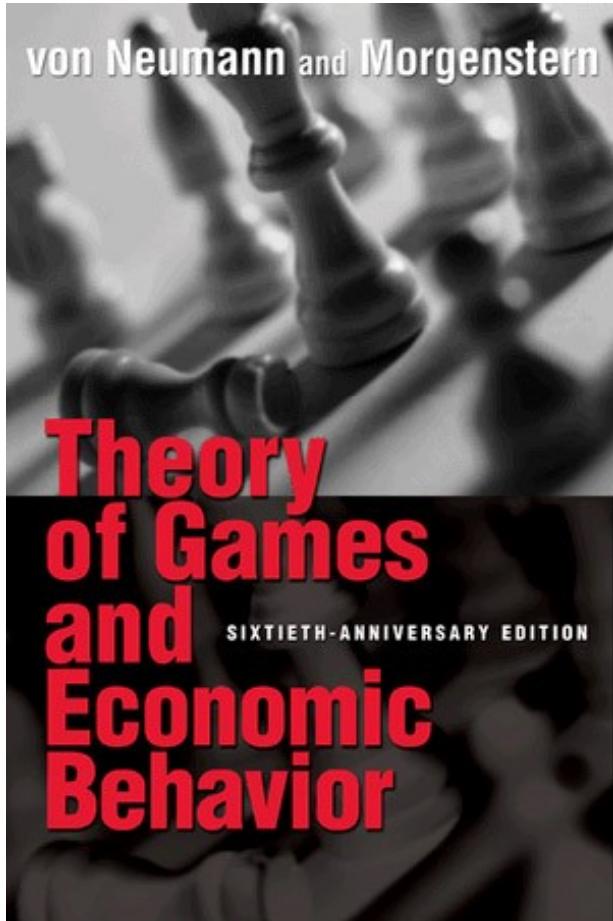
Choice-based models are proposed to overcome some of the limitations found in traditional rating-based strategies. The new approach is grounded on decision-making paradigms, such as choice and utility theories. Specifically, random utility models were applied in a recommendation problem. Prediction accuracy was compared with state-of-art rating-based algorithms in a gastronomy dataset. The results show the superior performance of choice-based models, which may suggest that real choices could bring more predictive power than ratings.

item's attributes [2]. These preferences can be used to predict the utility of any given item by comparing them with the values of item's attributes. Collaborative recommenders, on the other hand, take advantage of previous ratings provided by the available decision-makers to predict the utility of any given user-item pair [6]. This approach has been widely adopted as it removes the burden of knowing and managing item attributes as well as their corresponding values.

Many algorithms and models have been proposed under the collaborative paradigm. Among them, two families have gained major attraction: neighborhood algorithms and latent factor models. The neighborhood approach was the

(RECSYS, 2016)

Decisiones que dependen de otras decisiones



1944

Juego de Emociones

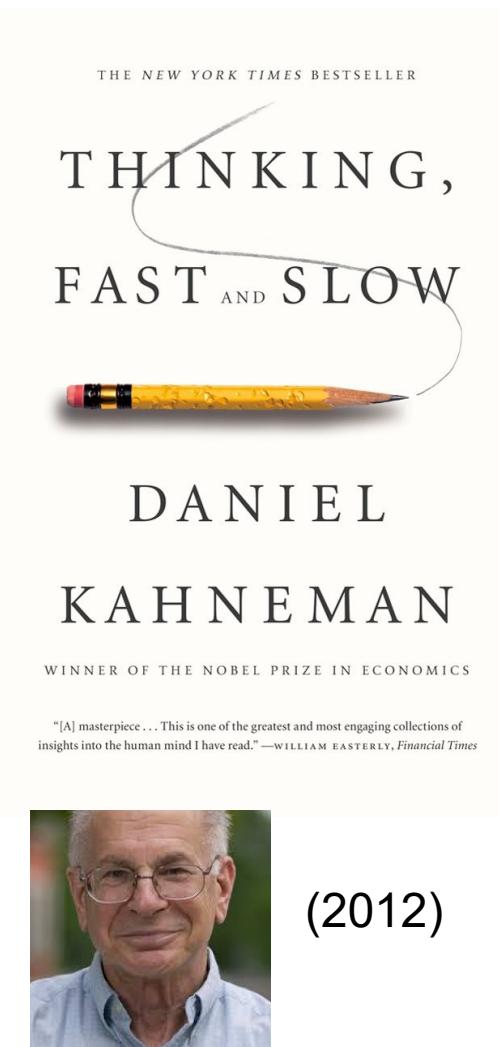
Estudiando el Sistema I

Nuestra mente tiene dos sistemas:

- Sistema I: Rápido, automático, poco esfuerzo, sin control consciente.
- Sistema II: Lento, racional, costoso, con control consciente.

Psicólogos ante la teoría económica:

- Econos Vs. Humanos.
- El poder de la multidisciplinariedad



Sistema I: Hipótesis del marcador somático

Emotions and decision-making:

- Emotions play a key role in making decisions
- Impairments/defects in emotional system reduces performance on decision-making tasks

The somatic marker hypothesis and the possible functions of the prefrontal cortex

ANTONIO R. DAMASIO*

Department of Neurology, Division of Behavioral Neurology and Cognitive Neuroscience, University of Iowa College of Medicine, Iowa City, Iowa, U.S.A.

SUMMARY

In this article I discuss a hypothesis, known as the somatic marker hypothesis, which I believe is relevant to the understanding of processes of human reasoning and decision making. The ventromedial sector of the prefrontal cortices is critical to the operations postulated here, but the hypothesis does not necessarily apply to prefrontal cortex as a whole and should not be seen as an attempt to unify frontal lobe functions under a single mechanism.

The key idea in the hypothesis is that 'marker' signals influence the processes of response to stimuli, at multiple levels of operation, some of which occur overtly (consciously, 'in mind') and some of which occur covertly (non-consciously, in a non-minded manner). The marker signals arise in bioregulatory processes, including those which express themselves in emotions and feelings, but are not necessarily confined to those alone. This is the reason why the markers are termed somatic: they relate to body-state structure and regulation even when they do not arise in the body proper but rather in the brain's representation of the body.

Examples of the covert action of 'marker' signals are the undeliberated inhibition of a response learned previously; the introduction of a bias in the selection of an aversive or appetitive mode of behaviour, or in the otherwise deliberate evaluation of varied option-outcome scenarios. Examples of overt action include the conscious 'qualifying' of certain option-outcome scenarios as dangerous or advantageous.

The hypothesis rejects attempts to limit human reasoning and decision making to mechanisms relying, in an exclusive and unrelated manner, on either conditioning alone or cognition alone.



(1996)

- Hypothesis 1: System I models (Emotional, Attentional) will show better performance than System II models (Rational).

- Hypothesis 2: Emotional models will ouperform Attentional models.

Métodos



HOME JOBS FINANCE TABLES AND CHARTS EVENTS PAIDCONTENT 50

Sep 30, 2010 - 12:12AM

ADVERTISEMENT

Facebook's Sandberg: In The Future, All Media Will Be Personalized

BY David Kaplan

5 Comments +1

There will always be a place for mass marketing, but in the next three- to five years, a website that isn't tailored to a specific user's in...



There will always be a place for mass marketing, but in the next three- to five years, a website that isn't tailored to a specific user's interest will be an anachronism, Facebook COO Sheryl Sandberg told Arianna Huffington at the latter's Advertising Week event. "People don't want something targeted to the whole world — they want something that reflects what they want to see and know," she said. Sandberg and Huffington also discussed Facebook's privacy issues and a certain movie being released this weekend.

RELATED

France calms fears over Facebook Timeline scare

A French tabloid set off a temporary worldwide panic that Facebook had published the private messages of...

Newsprint joins the internet of things
Print and digital media are so disconnected, they often appear a lifetime apart. But new technology promises...

No... an internet you won't leave journalism



The Premium Experience:
Neurological Engagement
on Premium Websites

Experimental Setup

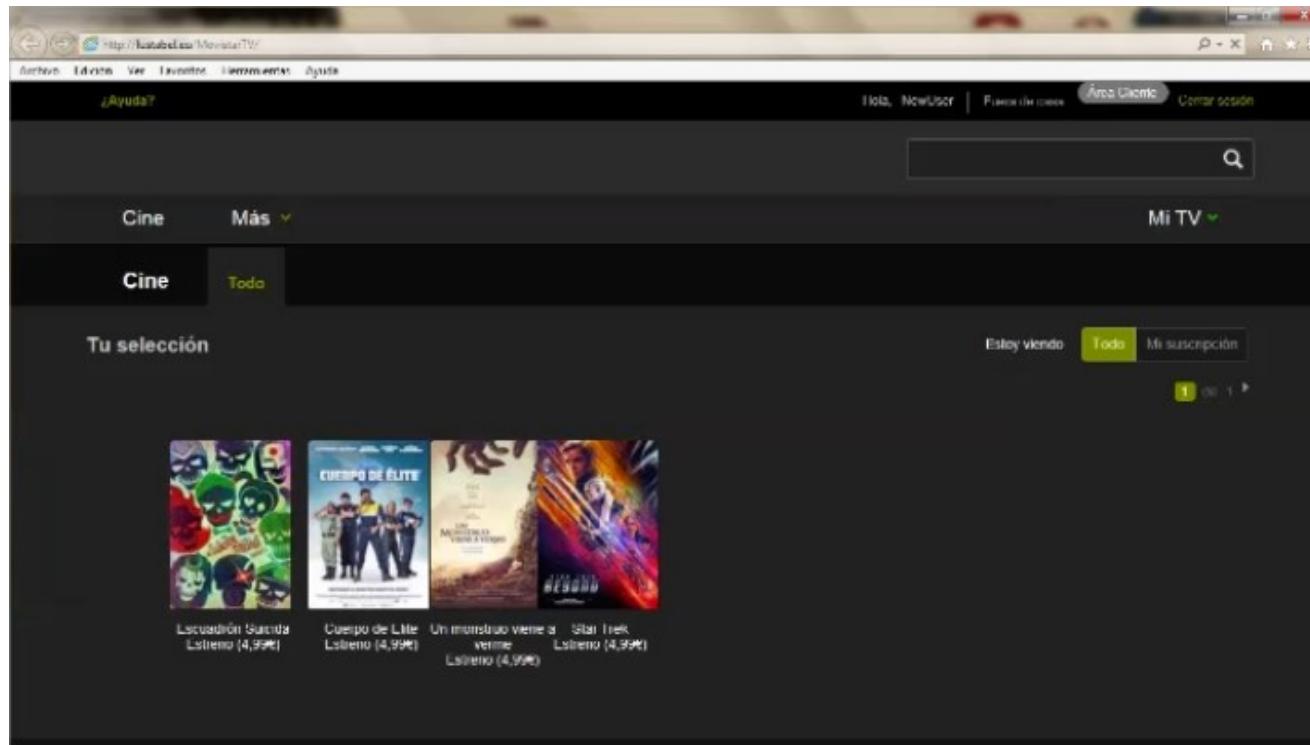


Experimental Setup

Table 1. Devices and recorded variables.

Devices	Models	Variables	Variable Type
EEG	Emotiv EPOC Headset	Frustration, Excitement and Engagement	Emotional
Facial Coding (FACEIT)	Logitech HD Pro webcam C920	Joy, Anger, Fear, Surprise, Contempt, Disgust, Sadness, Neutral, Positive and Negative	Emotional
Eye-Tracking	TOBii X2 @30Hz	TimeSpend and Fixations	Attentional
iMotions (Software)	Version 6.2		

Experimental Design



1 Trial per user:

- 4 Movies per situation
- 20 choice situations
- 80 Movies per Trial

Experimental Design

Table 2. Characterization of movies: attributes and values.

Feature	Values
Genre	Action, Comedy, Science Fiction, Drama
Novelty	Release, Catalog
Price	4,99 euros (Release), 0 euros (Catalog)

Factorial Design:

- 8 Profiles
- 10 Movies per Profile
- 80 Movies per Trial

Experiments



Choice Rule in probabilistic form:

$$CR(A, \geq) = \{a_i \in A \mid \mathbb{P}_i \geq \mathbb{P}_j, \forall a_j \in A\}$$

Form of probabilities: Standard Logit Function

$$\mathbb{P}_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}. \quad \text{with } V_{nj} = \beta_{nj} \times x_j$$

V_{nj} , the utility of alternative a_j for decision maker c_n

β_{nj} , the vector of preferences of decision maker c_n

x_j , the observed features of alternative a_j

Observed features

Table 3. Observed features.

Models	Features
Rational	Action, Comedy, Science Fiction, Drama, Release and Catalog
Emotional	Frustration, Excitement, Engagement, Joy, Anger, Fear, Surprise, Contempt, Disgust, Sadness, Neutral, Positive and Negative
Attentional	TimeSpend and Fixations

Specific Models

Table 4. Choice models used in this work: Rational (R), Attentional (A), and Emotional (E).

Model	Features
R	Action, Comedy, Fiction, Premiere
A1	TimeSpend
A2	FixationTime
E1	Action, Comedy, Fiction, Premiere, Anger, Frustration, Negative, Fear
E2	Action, Comedy, Fiction, Premiere, Excitement, Joy, Engagement
E3	Action, Comedy, Fiction, Premiere, Frustration, Excitement
E4	Action, Comedy, Fiction, Premiere, Frustration, Excitement, Engagement
E5	Action, Comedy, Fiction, Premiere, Anger, Fear
E6	Action, Comedy, Fiction, Premiere, Fear, Contempt, Disgust, Sadness
E7	Action, Comedy, Fiction, Premiere, Joy, Anger, Fear, Surprise
E8	Action, Comedy, Fiction, Premiere, Frustration, Excitement, Engagement, Joy, Anger
E9	Action, Comedy, Fiction, Premiere, Joy, Anger, Fear, Surprise, Contempt, Disgust, Sadness
E10	Action, Comedy, Fiction, Premiere, Sadness, Neutral, Positive, Negative
E11	Action, Comedy, Fiction, Premiere, Frustration, Excitement, Engagement, Neutral, Positive, Negative
E12	Action, Comedy, Fiction, Premiere, Anger, Fear, Surprise, Contempt, Disgust, Sadness, Neutral, Joy, Negative
E13	Action, Comedy, Fiction, Premiere, Frustration, Excitement, Engagement, Joy, Anger, Fear, Surprise, Contempt, Disgust, Sadness, Neutral, Negative

Evaluation

$$Accuracy = \frac{T_{CorrectPrediction}}{T_{AllPrediction}}$$

Resultados

Best Model for Single Subject

Table 5. Performance of choice-based models for each user. Accuracy results for top five models.

Users	Number of Choices	Top Five Models	Accuracy
J02	12	E8	0.787879
		E11	0.787879
		E4	0.757576
		E2	0.666667
		E6	0.636364
J06	14	A2	0.681818
		A1	0.621212
		E4	0.560606
		E9	0.560606
		R	0.545455
V09	20	R	0.522727
		E1	0.477273
		E0	0.454545
		E5	0.431818
		E6	0.431818

Best Model for Single Subject

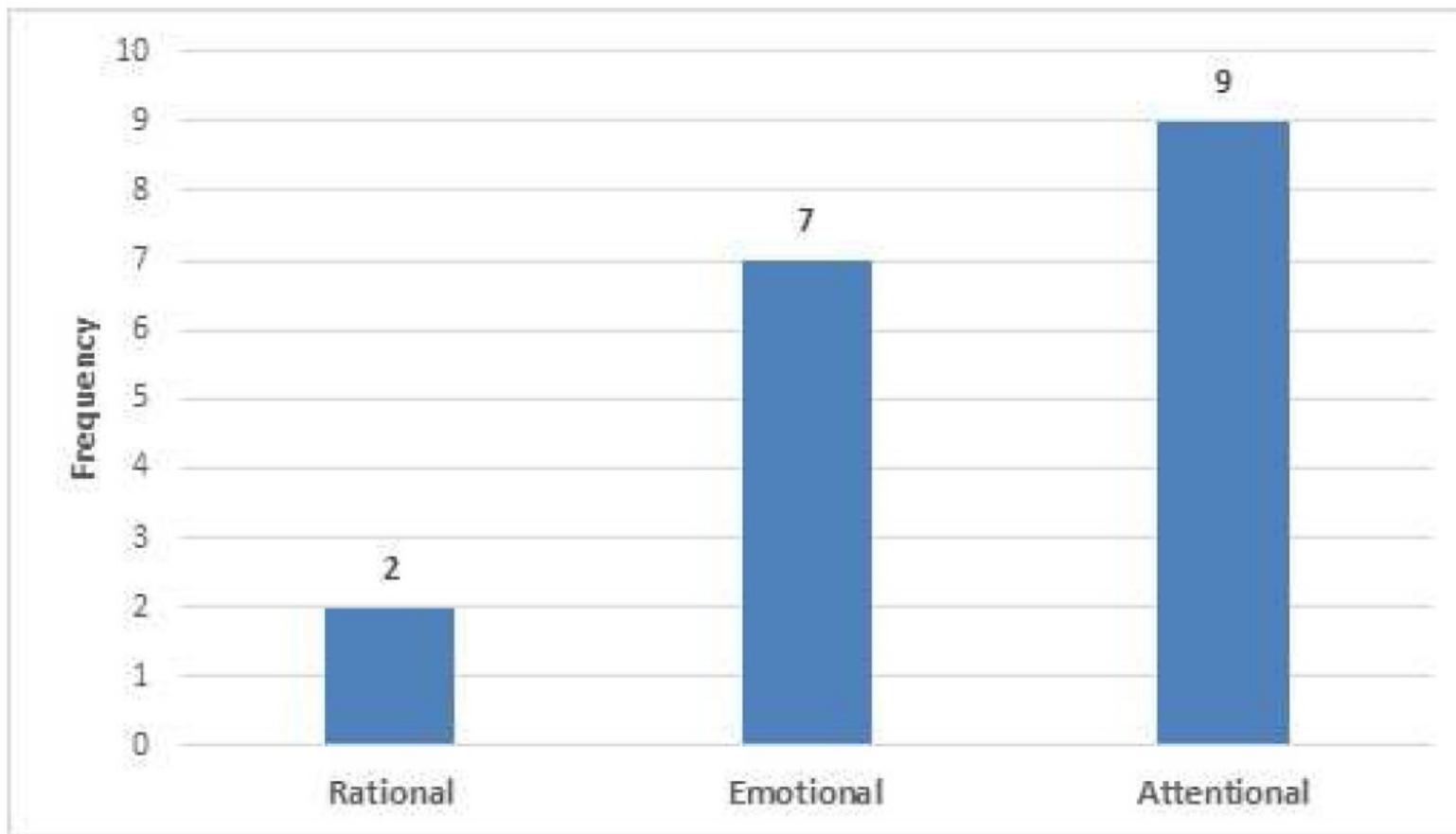


Fig. 3. Frequency of the Rational, Emotional, and Attentional models as best model.

Best Model for All Subjects

Table 6. Performance of choice-based models for average model for all users. Accuracy results for top five models.

Models	Average of Accuracy
Fixation	0.528282828
TimeSpend	0.522306397
Rational	0.520911496
Emotional6	0.464850890
Emotional4	0.448328523

Best Model for All Subjects

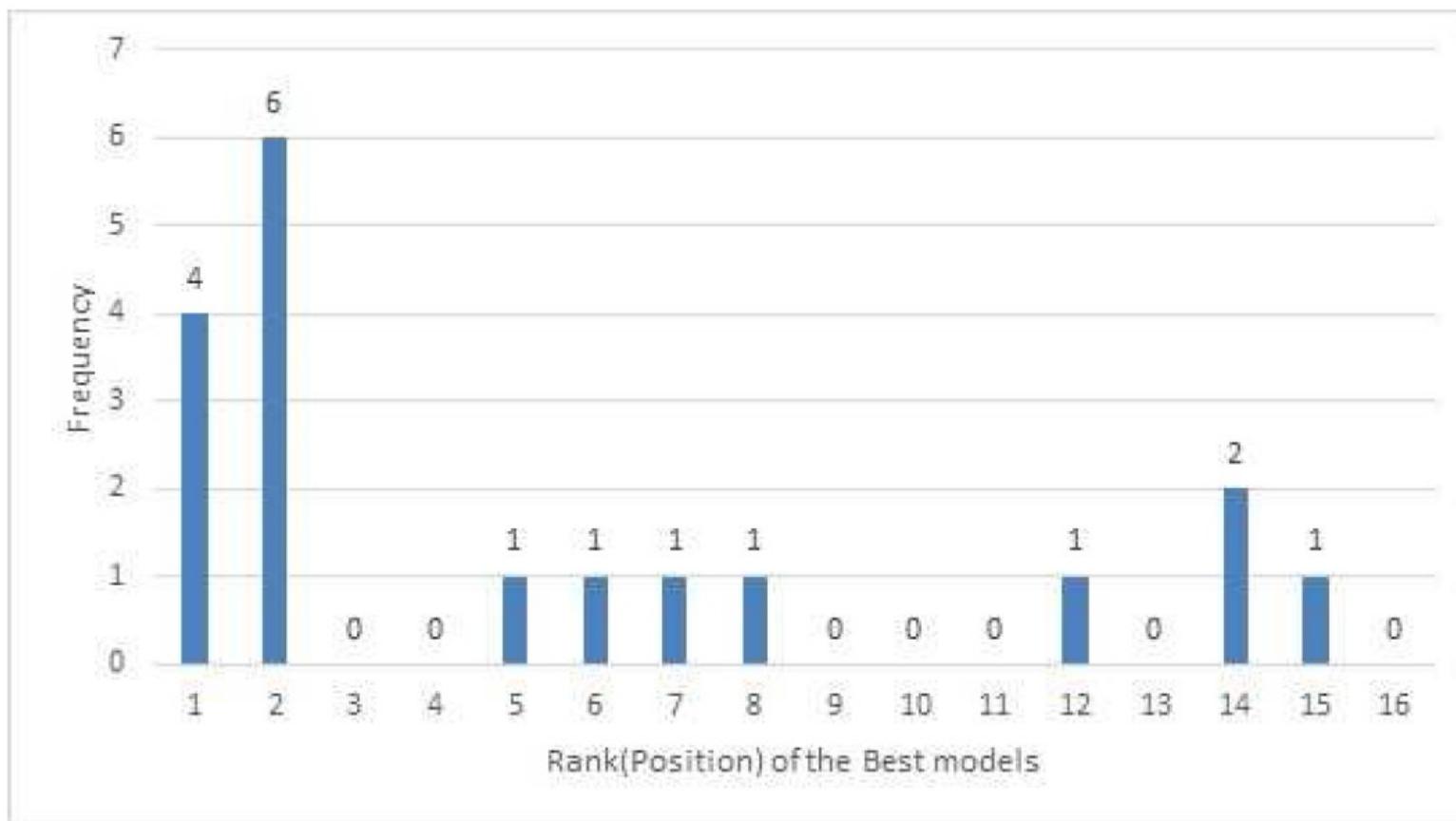


Fig. 4. Frequency of the rank of the best average model(Fixation model).

Discusión

Conclusiones principales

- Hypothesis I: System I models (Emotional, Attentional) will show better performance than System II models (Rational).
 - **RESULT: YES!**
- Hypothesis 2: Emotional models will ouperform Attentional models.
 - **RESULT: NO!**
- FINDING: No model is the best for every single subject!
- FUTURE WORK: Factors underlying the attention process?

Conexión con el Thick Data

Modelos Toma Decisiones => Thick Data!



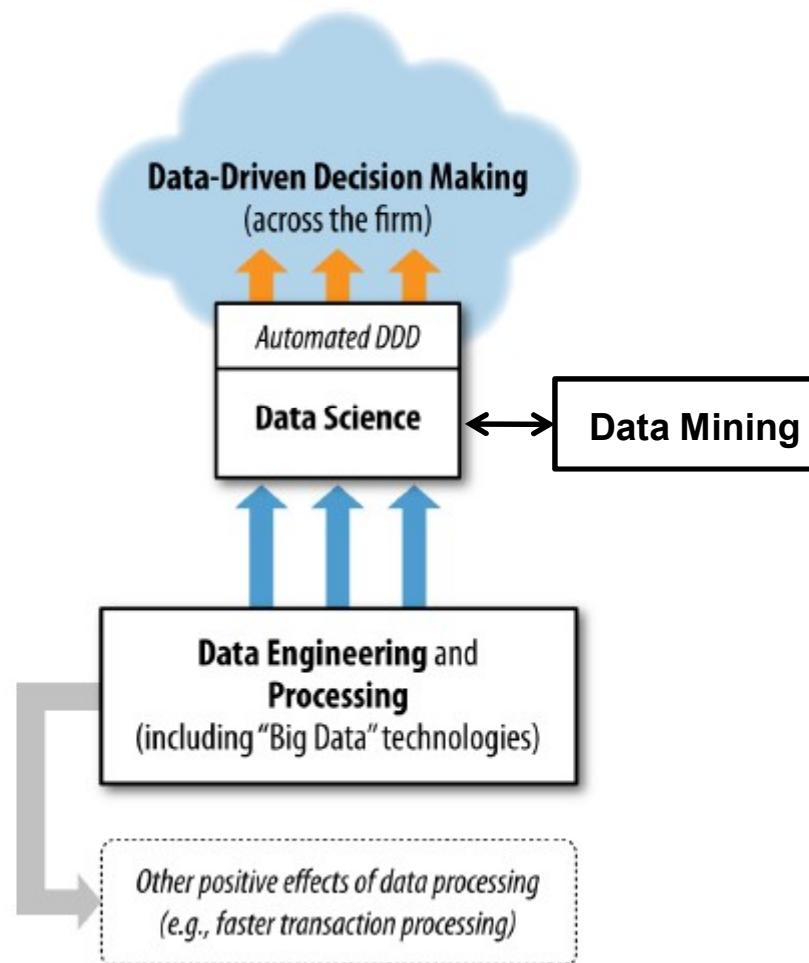
Modelos Toma Decisiones + Big Data => pensar analíticamente!

1. HIPÓTESIS

2. MODELOS CAUSALES / PREDICTIVOS

3. DATOS

Modelos Toma Decisiones + Big Data



Modelos de Toma de decisiones



Grupo de Sistemas *inteligentes*



Eduardo M. Sánchez Vila
eduardo.sanchez.vila@usc.es

Grupo de Sistemas Inteligentes
Universidad de Santiago de Compostela