



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

**ESTIMACION DE REGLAS DE DECISIONES CONJUNTIVAS DE
CONSUMIDORES EN UN EXPERIMENTO DE ANÁLISIS CONJUNTO A
PARTIR DE DATOS DE EYE TRACKING**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

DELPHINE MARGUERITE MARIE-LOUISE ALICE FALLER

**PROFESOR GUÍA:
ANDRÈS MUSALEM SAID**

**MIEMBROS DE LA COMISIÓN:
RICARDO MONTOYA MOREIRA
DENIS SAURE VALENZUELA**

**SANTIAGO DE CHILE
ENERO 2015**

ESTIMACIÓN DE REGLAS DE DECISIONES CONJUNTIVAS DE CONSUMIDORES EN UN
EXPERIMENTO DE ANÁLISIS CONJUNTO A PARTIR DE DATOS DE *EYE TRACKING*

El estudio de la forma en la que el consumidor toma decisiones de compra es de vital importancia en marketing. Los equipos de marketing necesitan entender lo que valora el consumidor al momento de elegir un producto o servicio y por qué lo valora, para enriquecer el producto de acuerdo a sus gustos y mejorar la manera de exponer informaciones sobre el producto.

En las técnicas usuales de análisis conjunto, se analizan las decisiones del consumidor frente a elecciones diseñadas por el investigador para deducir sus preferencias y gustos en los distintos atributos de un producto e inferir sus elecciones futuras. Esta memoria busca predecir no sólo la elección sino también la regla de decisión, utilizando los movimientos de ojos del consumidor durante el proceso de elección. En particular, se estudia la regla conjuntiva, no-compensatoria, en la cual se definen niveles mínimos aceptables para cada atributo. Las alternativas cuyos niveles sobrepasan estos niveles mínimos para todos los atributos son satisfactorias. El estudio se desarrolla con datos de un experimento de análisis conjunto en el cuál se pidió a los participantes seguir distintas reglas conjuntivas, que varían por sus distintos niveles mínimos aceptables.

Luego de un análisis descriptivo del comportamiento visual de los participantes, se elaboran distintos modelos a partir de los datos de *eye tracking*. Estos modelos logran predecir la regla conjuntiva usada con una tasa de acierto de 88% y elecciones futuras con una *hit rate* de 75%, con 10 preguntas de entrenamiento. Se compara el desempeño de este enfoque con modelos que usan elecciones pasadas, que arrojan una tasa de 100% en la predicción de la regla usada con el mismo conjunto de entrenamiento. Finalmente, se aplican los modelos a un segundo experimento en el cual los participantes eligieron de manera libre, sin regla impuesta. Se logra predecir elecciones futuras con elecciones pasadas con una *hit rate* de 57% y con el movimiento ocular con una *hit rate* de 36%. Resulta que el hecho de agregar información de movimientos de ojos a los modelos que usan elecciones pasadas no mejora su desempeño. Se concluye que el comportamiento visual, si bien es consistente con las elecciones de un consumidor, no aporta más información acerca de su regla de decisión conjuntiva. Se propone, para futuras investigaciones, integrar el modelo con otros que suponen otras reglas de decisión, para abarcar la diversidad de reglas usadas en la realidad.

Agradecimientos

Quiero partir dándole las gracias a mis profesores Andrés Musalem y Ricardo Montoya, quienes me confiaron este trabajo, me guiaron con mucha dedicación y rigor, y me ayudaron a resolver las dudas que se me presentaron.

Gracias a mis amigos Marlene y Julio, y a Pierre-Baptiste, por su ayuda y apoyo en esta memoria.

Gracias a los funcionarios del departamento de Ingeniería Industrial por su disponibilidad y disposición a ayudar.

Tabla de contenidos

1. INTRODUCCIÓN	1
1.1. Objetivos	2
1.2. Alcances	2
1.3. Metodología	2
1.4. Resultados esperados	3
1.5. Estructura de la memoria	4
2. ANTECEDENTES GENERALES.....	5
2.1. Análisis conjunto	5
2.2. Reglas de decisión.....	5
2.3. <i>Eye tracking</i>	6
2.4. <i>Eye tracking</i> y decisiones.....	7
3. EXPERIMENTO	8
3.1. Descripción de las preguntas.....	8
3.2. Grabación del movimiento ocular	10
4. ANÁLISIS DESCRIPTIVO.....	12
4.1. Alternativa elegida por los participantes	12
4.2. Número de fijaciones	13
4.3. Transiciones.....	19
4.4. Predicción de la elección	21
4.5. Conclusiones del capítulo	26
5. INFERIR LA REGLA CONJUNTIVA CON ELECCIONES.....	27
5.1. Insumos y resultados del algoritmo	27
5.2. Evaluación de los algoritmos.....	28
5.3. Pasos de los algoritmos.....	29
5.4. Resultados.....	31
5.5. Conclusiones del capítulo	35
6. INFERIR LA REGLA CONJUNTIVA CON MOVIMIENTOS DE OJOS	36
6.1. Insumos y resultados del algoritmo	36
6.2. Metodología	37
6.3. Paso 1: predecir los dos atributos con umbral.	38

6.4.	Paso 2: predecir los 2 atributos con umbral y donde se sitúa el umbral.	39
6.5.	Paso3: Predecir el umbral para cada atributo.....	47
6.6.	Conclusiones del capítulo	55
7.	APLICACIÓN EN UNA BASE DE DATOS DE ELECCIONES LIBRES Y MODELO COMBINADO	56
7.1.	Descripción del experimento de las cafeteras	56
7.2.	Análisis descriptivo.....	58
7.3.	Algoritmos que usan elecciones	60
7.4.	Algoritmos que usan datos de movimiento ocular.....	63
7.5.	Algoritmos que combinan elecciones y datos de movimiento ocular.....	64
7.6.	Conclusiones del capítulo	68
8.	CONCLUSIONES	69
	GLOSARIO	71
	BIBLIOGRAFÍA.....	72
	ANEXOS	73
a.	Tablas de contingencia.....	73
b.	Resultados de los modelos combinando elecciones y datos de movimientos de ojos aplicados al experimento de las cafeteras.	74

ÍNDICE DE TABLAS

Tabla 1: Definición de los niveles de dificultad.....	9
Tabla 2: Promedio y Des. Estándar del número de AOI y AOI únicas miradas	15
Tabla 3: Regresiones lineales de los números de fijaciones por caso	16
Tabla 4: Matriz de transiciones entre productos.....	20
Tabla 5: Matriz de transiciones entre atributos.....	20
Tabla 6: Valores de las diagonales de las matrices de transiciones	21
Tabla 7: Porcentaje de casos en los cuales producto elegido= producto más mirado	21
Tabla 8: Predicción de los umbrales Enumeración- <i>Greedy</i>	31
Tabla 9: Tiempos computacionales de los algoritmos <i>Greedy</i> y de enumeración.....	34
Tabla 10: Proceso incremental de construcción de modelos.....	37
Tabla 11: Predicción de umbrales con movimientos de ojos: paso 2 (10 preguntas de entrenamiento).....	46
Tabla 12: Predicción de umbrales con movimientos de ojos: paso 3	54
Tabla 13: Atributos del experimento de cafeteras	56
Tabla 14: <i>Hit rate</i> para <i>Greedy</i> y Enumeración: experimento de las cafeteras	62
Tabla 15: <i>Hit rates</i> de referencia en el experimento de las cafeteras	62
Tabla 16: <i>Hit rate</i> para <i>Greedy</i> , Enumeración y modelo 6.2: experimento de las cafeteras	64
Tabla 17: Tabla de contingencia producto más mirado y producto elegido.....	73
Tabla 18: Tabla de contingencia último producto mirado y producto elegido	73

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Dispositivo de <i>eye tracking</i>	6
Ilustración 2: Matriz de alternativas y atributos	8
Ilustración 3: Numeración de los AOIs.....	10
Ilustración 4: Ejemplo de fijaciones en el experimento	11
Ilustración 5: ¿En qué pregunta se hicieron errores?	12
Ilustración 6: Histograma del número de fijaciones	14
Ilustración 7: Número promedio de AOI y AOI únicas fijadas y fijaciones repetidas por pregunta	15
Ilustración 8: Coeficientes de los niveles de dificultad en las regresiones lineales (1), (2) y (3).	18
Ilustración 9: Histograma del número promedio de fijaciones por persona.....	19
Ilustración 10: Transiciones posibles (fuente: [9]).....	19
Ilustración 11: Porcentajes de fijaciones en productos más mirados.....	22
Ilustración 12: Probabilidades condicionales de elección	24
Ilustración 13: Porcentaje de los casos en los cuales el producto elegido es el más mirado y/o el último mirado.....	25
Ilustración 14: Histograma del porcentaje de veces en las cuales producto elegido es el más mirado o el último mirado por persona	25
Ilustración 15: Input y Output del algoritmo con elecciones	27
Ilustración 16: Evaluación de los algoritmos	28
Ilustración 17: Evolución del ajuste de los modelos <i>Greedy</i> y de enumeración con el tamaño del conjunto de entrenamiento.....	32
Ilustración 18: <i>Hit rate Greedy</i> vs enumeración, 10 preguntas de entrenamiento	33
Ilustración 19: <i>Hit rate Greedy</i> vs enumeración, 6 preguntas de entrenamiento.....	34
Ilustración 20: Input y Output del algoritmo con elecciones.....	36
Ilustración 21: ¿Tienen umbrales los atributos más mirados?	38
Ilustración 22: Porcentaje de fijaciones en los atributos más mirados	39
Ilustración 23: Aspectos esperados de las probabilidades de mirar un nivel según el umbral	40
Ilustración 24: Número de fijaciones por aparición en los niveles de los atributos con o sin umbral.....	41
Ilustración 25: Predicción de elecciones con movimientos de ojos: paso 2.....	47
Ilustración 26: pvalores de la regresión lineal $nfixPerAp \sim 1$	49
Ilustración 27: Número de fijaciones en los atributos del participante 41	51
Ilustración 28: Variables X_n de las regresiones lineales	52
Ilustración 29: Modelo 6.2: R^2 de las regresiones lineales	53
Ilustración 30: Predicción de elecciones con movimientos de ojos: paso 3	54
Ilustración 31: Diseño del experimento de cafeteras.....	57
Ilustración 32: Evaluación de los algoritmos en el experimento de las cafeteras	58

Ilustración 33: Porcentaje de fijaciones por atributo en el experimento de las cafeteras	58
Ilustración 34: Porcentaje de fijaciones por atributo, atributos ordenados, en el experimento de las cafeteras	59
Ilustración 35: Fijaciones por aparición en los atributos más y menos fijados del experimento de las cafeteras	60
Ilustración 36: Predicción de elecciones con elecciones: experimento de las cafeteras...	61
Ilustración 37: Predicción de elecciones con movimientos de ojos: experimento de las cafeteras.....	63
Ilustración 38: Ejemplo de matriz de <i>hit rates</i> de un bucle de la heurística <i>Greedy</i>	65
Ilustración 39: Predicción de elecciones con elecciones y movimientos de ojos: experimento de las cafeteras	67
Ilustración 40: Predicción de elecciones, modelos combinados, experimento de las cafeteras.....	74
Ilustración 41: Variaciones del modelo 7.3 con gamma	75
Ilustración 42: Variaciones del modelo 7.4 con gamma.....	76

1. INTRODUCCIÓN

Los consumidores se ven enfrentados a un gran número de elecciones diarias, con cada vez más información sobre la oferta de productos. La forma como toman decisiones es un elemento de vital importancia para las empresas, porque les permite definir su oferta de productos y otras variables de marketing (tales como precio, distribución y promoción) de manera de ajustarles a la demanda.

Dentro de los métodos de investigación de mercado, la técnica de análisis conjunto se puede utilizar para entender y medir qué atributos de un producto son importantes para los consumidores. En esta técnica se analizan las elecciones de los consumidores dentro de un conjunto de alternativas de productos con distintos atributos (por ejemplo precio o calidad) para calcular la importancia relativa de estos atributos para los consumidores. De esta manera se diseñan productos que tienen fortalezas en los atributos más valiosos.

Los consumidores esperan de las empresas que personalicen sus productos según sus preferencias personales. Con este fin existe mucha literatura que intenta modelar el pensamiento del consumidor a nivel individual. En general se usan modelos que maximizan la utilidad para un consumidor de elegir tal o cual alternativa. El cálculo de esta utilidad resulta de las elecciones pasadas del consumidor y de varios supuestos sobre su regla de decisión. Por ejemplo Gilbride and Allenby (2004 [6]) construyen un modelo de elección que supone el uso de reglas conjuntivas, disjuntivas y compensatorias. Sin embargo, en este primer enfoque, solamente se observa el resultado, es decir las elecciones, y no el camino -el proceso de elección-. Como un mismo resultado puede surgir de distintos caminos, este enfoque queda así limitado.

Para entender cómo el consumidor llega a una decisión, se observa el proceso de elección. Varios métodos sirven para examinar el comportamiento de las personas cuando están tomando una decisión en un experimento: pedir a la persona describir oralmente su pensamiento, seguir los movimientos de su ratón de computadora o sus movimientos de ojos. Funaki, Jiang y Potters (Funaki et al., 2004, [6]) muestran que existe una estructura racional en el comportamiento visual, que los movimientos de ojos reflejan las decisiones de las personas y que si una persona sigue una regla de decisión determinada, se puede ver un patrón relacionado en sus movimientos de ojos. En esta memoria, se busca aprovechar las herramientas tecnológicas de *eye tracking* para analizar el comportamiento visual de los consumidores durante el proceso de elección y ver si se puede inferir el tipo de regla de decisión usada y predecir la decisión final con más precisión que con modelos que solamente usan las elecciones. Se pondrá énfasis en un tipo de reglas de decisión llamado reglas conjuntivas en que el participante elige la alternativa que sobrepasa un nivel mínimo para cada atributo.

1.1.Objetivos

Objetivo general:

Identificar las reglas de decisiones conjuntivas utilizadas por consumidores en aplicaciones de análisis conjunto con datos de *eye tracking*.

Objetivos específicos:

En esta memoria se busca:

- Determinar los patrones de fijación de los ojos cuando los participantes usan reglas conjuntivas.
- Inferir el tipo de regla conjuntiva usada a partir de los datos de movimientos de ojos del participante y con esto predecir sus decisiones futuras.
- Comparar el uso de datos de *eye tracking* con métodos de predicción basados solamente en las elecciones.
- Integrar variables de *eye tracking* en el modelo basado en elecciones para mejorar su rendimiento.

1.2. Alcances

Se presentan los alcances esperables y los límites del proyecto:

- Se trabaja con los datos de experimentos efectuados por Martin Meißner en 2013.
- Dentro de las reglas de decisión, se trabaja solamente con la regla conjuntiva.
- Solo se considera el comportamiento visual y las elecciones para estudiar la regla de decisión, no se incluyen variables socio-demográficas o de otra naturaleza, que se incorporan en la heterogeneidad no observada.
- Esta memoria tiene un fin de investigación, no se van a examinar aplicaciones en la industria.

1.3. Metodología

Se analizan los datos de movimientos de ojos (fijaciones y transiciones) y las elecciones de un experimento de análisis conjunto realizado con 60 personas y un total de 720 preguntas. Se pidió a los participantes elegir una alternativa dentro de 4 hoteles, usando una regla conjuntiva. Para una descripción del experimento ver el capítulo 3.

Primero se hace un análisis descriptivo para ver cómo se comportan visualmente las distintas personas: número de fijaciones y transiciones, relación con la elección y con la disposición espacial de las alternativas, dinámica de los movimientos. Se investiga si el comportamiento visual de los participantes refleja sus elecciones.

Luego, se busca ajustar modelos para predecir el tipo de regla conjuntiva. Una regla conjuntiva puede ser caracterizada por un conjunto de niveles mínimos aceptables, llamados umbrales, uno para cada atributo. Las alternativas que pasan el umbral para todos los atributos son aceptables. Los modelos buscan predecir los umbrales de cada persona. Se definirán dos tipos de variables explicativas:

- primero variables que dependen de elecciones pasadas,
- segundo variables con relación a los movimientos de ojos.

Para el modelo o algoritmo con elecciones se usa la heurística *Greedy*, elaborada en primera instancia para reglas lexicográficas, que se modifica para aplicarse a reglas conjuntivas. Se compara con un segundo modelo que comprende las variables de comportamiento ocular. Las variables de *eye tracking* se desarrollan a partir de los descubrimientos del análisis descriptivo y pueden por ejemplo involucrar variables como el número de fijaciones en un atributo con un nivel sobre el número de fijaciones totales.

Finalmente, se testean los modelos con los resultados de un segundo experimento de análisis conjunto en el cuál los participantes estuvieron libres en su método de decisión. Este segundo experimento se hizo sobre otra categoría de productos (cafeteras) con el mismo número de atributos (seis). Se ve cuál de los dos enfoques (elecciones o movimiento ocular) logra predecir mejor las decisiones. Además se realiza otro modelo que busca combinar los dos para mejorar la predicción.

1.4. Resultados esperados

- Descripción de la situación actual, análisis de la literatura relevante.
- Análisis descriptivo: estadísticas sobre el comportamiento visual de los participantes.
- Resultado de los modelos para inferir el tipo de regla conjuntiva utilizando las elecciones que existen en la literatura.
- Modelo para inferir el tipo de regla conjuntiva utilizando datos de movimiento ocular.
- Métricas de comparación de los dos modelos.
- Resultados de los modelos aplicados a otra base de datos, usando otro experimento en el cual el participante eligió de manera libre.

- Modelo para inferir el tipo de regla conjuntiva utilizando una combinación de elecciones y datos de movimiento ocular.

1.5. Estructura de la memoria

Capítulo 1: Introducción

Plantea el contexto del trabajo y la motivación, describe las características generales y los pasos principales de la investigación, objetivos y alcances.

Capítulo 2: Antecedentes generales

Se muestra una revisión de la literatura relativa a los temas de las reglas de decisión, del *eye tracking* y los modelos de elecciones de reglas conjuntivas.

Capítulo 3: Experimento

Se describen las características del experimento, la imposición de la regla conjuntiva y la captura de información visual.

Capítulo 4: Análisis descriptivo

Se explica cómo fue abordado el análisis exploratorio, su justificación y los resultados obtenidos en él.

Capítulo 5: Inferir la regla conjuntiva con elecciones

Se aplican dos modelos que usan elecciones pasadas para inferir la regla conjuntiva usada por el consumidor. Se comentan los resultados.

Capítulo 6: Inferir la regla conjuntiva con movimientos de ojos

Se crean varios modelos que usan el comportamiento ocular de un consumidor para inferir la regla conjuntiva usada por él. Se comparan con los modelos que usan elecciones.

Capítulo 7: Aplicación en una base de datos de elecciones libres y modelo combinado

Se aplican los modelos a otro experimento de elecciones libres. Se comparan los resultados y se crean modelos que combinan elecciones con movimientos de ojos para mejorar la predicción.

Capítulo 8: Conclusiones

Conclusiones finales del trabajo realizado y futuros caminos de investigación.

2. ANTECEDENTES GENERALES

2.1. Análisis conjunto

El análisis conjunto es una técnica de investigación de mercado que permite medir las preferencias y gustos de los consumidores en cierta categoría de productos. Es muy usada en marketing, en particular en la prueba de nuevos diseños de productos. Existen distintos tipos de análisis conjunto, pero en este trabajo se trabaja solamente con el análisis conjunto basado en elecciones (*Choice Based Conjoint*), que busca simular elecciones de la vida real. Consiste en una encuesta en la cual se pide al participante responder varias preguntas. Cada pregunta consiste en la elección de una alternativa dentro de un pequeño conjunto de productos. Los productos se caracterizan con sus niveles en distintos atributos como precio o diseño. Por ejemplo los niveles del atributo precio pueden ser \$99, \$199 y \$299. Se muestran varias combinaciones de niveles en las distintas preguntas para capturar todas las preferencias del consumidor, a nivel global o a nivel individual.

2.2. Reglas de decisión

Las reglas de decisión se refieren al conjunto de condiciones que permiten al consumidor aceptar o rechazar una alternativa dentro del conjunto de alternativas propuestas. Riedl, Brandstätter y Roithmayr (Riedl et al. 2008 [2]) describen trece reglas de decisiones, dentro de las cuales se destacan la conjuntiva, la disjuntiva y la lexicográfica (ver glosario para las definiciones).

Las reglas de decisión se reparten en dos grupos: el primer grupo abarca las reglas compensatorias en las cuales un nivel no satisfactorio de un atributo de una alternativa puede ser compensado por el buen nivel de la misma alternativa en otro atributo. El segundo grupo comprende las reglas no-compensatorias que no permiten esta compensación. En esta memoria se enfocará en la regla conjuntiva, que es no compensatoria. En esta regla, se definen niveles mínimos aceptables para cada atributo. Las alternativas cuyos niveles sobrepasan estos niveles mínimos para todos los atributos son satisfactorias. La regla no especifica cómo elegir en el caso de que dos o más alternativas sean satisfactorias, lo que se modela de dos maneras en la literatura:

- Elegir la primera alternativa satisfactoria examinada. Este método toma en cuenta que el costo de considerar una alternativa no es nulo. Supone entonces una evaluación secuencial de las alternativas y depende del orden de examinación de las alternativas. Stüttgen, Boatwright y Monroe (Stüttgen et al. 2012 [5])

consideran que el consumidor sigue este método. Riedl et al. (2008 [2]) usan este enfoque para modelar las reglas conjuntivas.

- Elegir la alternativa que entrega la mayor utilidad dentro de las alternativas satisfactorias. Gilbride and Allenby (2004 [6]), separan el proceso de elección en dos etapas. Su modelo discreto asume que el consumidor, en un primer paso, filtra las alternativas con una regla conjuntiva, para determinar un “choice set” y luego elige la alternativa de este conjunto que maximiza su utilidad con un modelo de elección regular.

2.3. *Eye tracking*

Las herramientas de *eye tracking*, o seguimiento de los movimientos de ojos, se desarrollaron a partir del siglo XX. Al inicio se estudiaron métodos de lectura, después en los años 80 se empezó a usar aquellas para resolver preguntas relacionadas con la interacción persona-ordenador. Luego se analizaron movimientos de ojos para ver cómo los usuarios de una página web reaccionan a los distintos diseños, saber en lo que se centran en una página web para ver por ejemplo dónde colocar una anuncio importante [1] o imágenes [8]. En el sector del retail se están usando herramientas de *eye tracking* para inferir las preferencias de los consumidores a partir de lo que miran, en vez de lo que dicen, que es más sesgado. En este ámbito el estudio de los movimientos de ojos de consumidores frente a tiendas virtuales podría bien reemplazar los *focus groups* [7]. En el futuro se espera un desarrollo mayor de los métodos de *eye tracking* con nuevos instrumentos tecnológicos, cada vez menos intrusivos, como los lentes Google Glass.



Ilustración 1: Dispositivo de *eye tracking*

Las herramientas de *eye tracking* usadas colectan dos tipos de datos:

- Las fijaciones: cuando el individuo posiciona su vista en un determinado lugar y la información adquirida puede ser procesada por el cerebro. Son los periodos en los cuales el ojo está inmóvil, y cuya duración suele estar entre los 200 y 500 milisegundos.
- Las transiciones: movimiento desde una fijación a otra, pero a diferencia de las fijaciones, la información vista en este movimiento no alcanza a ser procesada. La velocidad es alrededor de 20 - 40 milisegundos, dejándolo como el movimiento más rápido que hay en el cuerpo humano. Una persona, realiza en promedio 170 mil transiciones por día.

2.4. *Eye tracking* y decisiones

Varias investigaciones incluyen datos de movimientos de ojos para identificar reglas de decisiones. Riedl et al. (2008 [2]) identifican el uso de una regla de decisión (dentro de trece) con tres variables de seguimiento de ojos durante el proceso de elección y una variable relacionada al resultado. Su algoritmo de clasificación permite identificar la mayoría de las 13 reglas estudiadas pero no distingue las reglas conjuntivas y disyuntivas.

Gilbride and Allenby (2004 [6]) proponen un modelo de elección en dos etapas. Como la función de verosimilitud no es continua o diferenciable, no se pueden usar métodos de estimación usuales de gradientes como Newton-Raphson. Por lo tanto, se utilizan los métodos de aumentación de datos y de cadenas de Markov Monte Carlo para integrar sobre el parámetro. En este estudio no se consideran datos del proceso de elección, solamente los datos del experimento (valores de los atributos) y los resultados.

Stüttgen et al. (2012 [5]) consideran que el consumidor selecciona la primera alternativa que satisface los niveles mínimos de la regla conjuntiva. Se usa un modelo que combine datos de *eye tracking* y de elección y que supera un modelo de elección logit multinomial sobre el experimento realizado.

3. EXPERIMENTO

El presente estudio se desarrolla con datos que fueron capturados con herramientas de *eye tracking* durante un experimento de análisis conjunto basado en elecciones efectuado en el mes de septiembre 2013 en Alemania por Martin Meißner. El objetivo del experimento es descubrir patrones en el comportamiento visual de los encuestados cuando usan dos tipos de reglas de decisión no compensatorias, que son las reglas conjuntivas y lexicográficas. Se escogieron estas dos reglas porque corresponden a las reglas no-compensatorias más usadas.

3.1. Descripción de las preguntas

Un total de 60 personas respondieron a 24 preguntas de las cuales en 12 de ellas se les indicó seguir una regla conjuntiva. En esta memoria se estudian solamente las reglas conjuntivas. En cada pregunta, los participantes deben escoger entre cuatro alternativas de hoteles y una opción de no arriendo de un hotel. Pueden evaluar las distintas alternativas con respecto a seis atributos como la cercanía al centro de la ciudad, el precio o la calidad de la comida. Cada atributo puede tener un nivel bajo, medio o alto, siendo el nivel alto la mayor utilidad. Por ejemplo un nivel alto en “distancia al centro de la ciudad” representa un hotel cerca del centro, y un nivel alto en el atributo “precio” significa un precio bajo. En la Ilustración 2: Matriz de alternativas y atributos, se ve el ejemplo de una pregunta, con los seis atributos y la disposición espacial de ellos.

Atributo	Hotel 1	Hotel 2	Hotel 3	Hotel 4
Calidad de la comida	(-)	(-)	(-)	(o)
Recomendación de clientes	(-)	(+)	(o)	(-)
Distancia al centro de la ciudad	(+)	(o)	(+)	(o)
Vista al mar	(+)	(o)	(-)	(+)
Precio	(+)	(+)	(-)	(-)
Categoría de la habitación	(-)	(+)	(-)	(o)

Ilustración 2: Matriz de alternativas y atributos

Se pidió a los participantes elegir un hotel para un amigo hipotético, según indicaciones específicas, lo que permitió imponer la regla de decisión usada. Los participantes tenían a disposición una explicación del funcionamiento de la regla y los niveles mínimos para

cada atributo. Se prepararon respondiendo algunas tareas de entrenamiento. El experimento fue diseñado de manera de impedir los casos de empate: después de aplicar la regla queda una única alternativa.

En todos los casos del experimento, dos de los seis atributos tenían umbral, y para todos los casos el límite se situaba en el nivel medio. Para los cuatro atributos restantes, el nivel mínimo exigido era el nivel bajo, es decir que no había umbral para estos atributos. Por lo tanto, cada tipo de regla conjuntiva puede ser caracterizada por los dos atributos que tienen umbral.

La regla a seguir en las 12 preguntas de una persona era la misma: en las 12 preguntas, la alternativa elegida debía sobrepasar los mismos umbrales. Lo que cambiaba era la matriz de 6x4 niveles de las alternativas en cada atributo. Así se reproduce una situación en la cual el consumidor vuelve a elegir muchas veces un mismo producto (bienes o servicios de uso cotidiano por ejemplo).

Se diseñaron preguntas de 6 niveles de dificultad distintos, que cubren todas las configuraciones posibles. Cada uno de los participantes respondió dos preguntas de cada nivel de dificultad, en un orden aleatorio. Los niveles de dificultad se caracterizan con un conjunto $\{m; n\} \in \llbracket 1,4 \rrbracket^2$ donde m es el número de alternativas que sobrepasan el límite en uno de los atributos con umbral y n el número de alternativas que sobrepasa el límite en el otro atributo con umbral, con $m \geq n$. Se definen los seis niveles de dificultad en la Tabla 1. Cabe recordar que siempre hay una única alternativa que sobrepasa los dos límites.

Nivel	Número de alternativas que sobrepasan el límite en uno de los atributos con umbral	Número de alternativas que sobrepasan el límite en el otro atributo con umbral	Número de alternativas que sobrepasan el límite en los 2 atributos con umbral
1	1	1	1
2	2	1	1
3	2	2	1
4	3	1	1
5	3	2	1
6	4	1	1

Tabla 1: Definición de los niveles de dificultad

En los niveles 1, 2, 4 y 6, una sola alternativa sobrepasa el umbral en uno de los atributos relevantes, entonces en teoría basta mirar a un solo atributo para hacer la decisión. Pero éste supone que la persona pueda ir directamente a dicho atributo. En realidad la posición vertical de este atributo en comparación con el otro atributo con umbral podría

influir: si no aparece primero, lo más probable es que la persona vaya a mirar el otro atributo con umbral primero.

3.2. Grabación del movimiento ocular

Se siguieron los movimientos de ojos de los participantes usando lentes especiales que se pueden ver en la Ilustración 1: Dispositivo de *eye tracking*. La pantalla se divide en distintas áreas de interés (AOI o *Area of interest*), que corresponden a los distintos elementos visuales. Cada AOI es numerada, como se puede ver en la Ilustración 3. Esta numeración es compuesta de dos dígitos, el primero corresponde a la línea o al atributo (de 1 a 6) y el segundo a la columna o al producto (de 1 a 4).

Dein Freund hat festgelegt, dass für die folgenden zwei Eigenschaften der Pauschalreise ein Mindestwert erreicht werden soll: Qualität des Essens: sollte mittel (o) oder gut (+) sein
 Nähe zum Stadtzentrum: sollte 2km (o) oder 3km (+) sein
 Bitte wähle unter Berücksichtigung dieser Mindestwerte ein Angebot für Deinen Freund aus.

	Angebot A	Angebot B	Angebot C	Angebot D
Qualität des Essens	11 mittel (o)	12 mittel (+)	13 schlecht (-)	14 schlecht (-)
Weiterempfehlung in Prozent	21 90% (+)	22 70% (o)	23 70% (o)	24 90% (+)
Nähe zum Stadtzentrum	31 1km (+)	32 3km (-)	33 2km (o)	34 3km (-)
Meerblick	41 kein Meerblick (-)	42 kein Meerblick (-)	43 seitlicher Meerblick (o)	44 seitlicher Meerblick (o)
Preis	51 99 Euro (-)	52 99 Euro (+)	53 99 Euro (+)	54 99 Euro (o)
Zimmerkategorie	61 gehoben (o)	62 gehoben (o)	63 standard (-)	64 standard (-)

Ilustración 3: Numeración de los AOIs

Se tienen como resultado las fijaciones en cada AOI, el orden y la duración de aquellas. Cada círculo rojo de la Ilustración 4 corresponde a una fijación. Para ser analizada, esta información se dispone en una base de datos que contiene dos listas por pregunta, de igual tamaño. La primera lista agrupa las AOI fijadas en el orden de las fijaciones. Por ejemplo el participante miró primero la AOI 33, segundo la 34 (que corresponde al atributo 3 de la alternativa 4), luego la 64.

La segunda lista da la duración de cada fijación. En total se tiene 1440 listas, 2 por pregunta. Cabe señalar que por temas de tiempo se analizó solamente la lista de las fijaciones, dejando de lado las variaciones de duración.

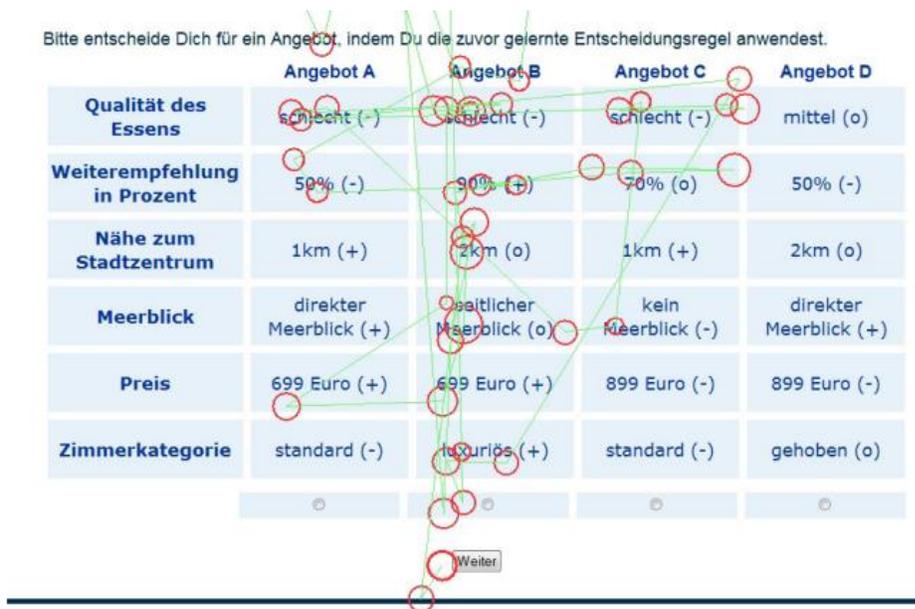


Ilustración 4: Ejemplo de fijaciones en el experimento

4. ANÁLISIS DESCRIPTIVO

En esta parte se hace una primera observación de los datos, con el objetivo de caracterizar el comportamiento ocular de los participantes. Este análisis se realiza sobre los 720 casos del experimento, que corresponden a las 12 preguntas de las 60 personas. Se busca identificar patrones del comportamiento ocular de los participantes, sus interpretaciones posibles, sus conclusiones en un ámbito de marketing. Este análisis plantea las intuiciones de los modelos para inferir la regla de decisión conjuntiva usada con movimientos de ojos que se desarrollan en el capítulo 6.

4.1. Alternativa elegida por los participantes

La primera pregunta de este análisis es la siguiente: ¿Se equivocaron los participantes? Permite verificar dos hipótesis: primero que el comportamiento visual refleja el seguimiento de una regla conjuntiva y segundo que siguieron la regla específica dada por las instrucciones.

Los participantes se equivocaron 9 veces sobre 720 casos, es decir en 1,25% de los casos, lo que se considera como bajo. Un 10% de los participantes se equivocaron al menos una vez (6 de 60). Estos números confirman que los participantes siguieron en su mayoría la regla conjuntiva dada por las instrucciones.

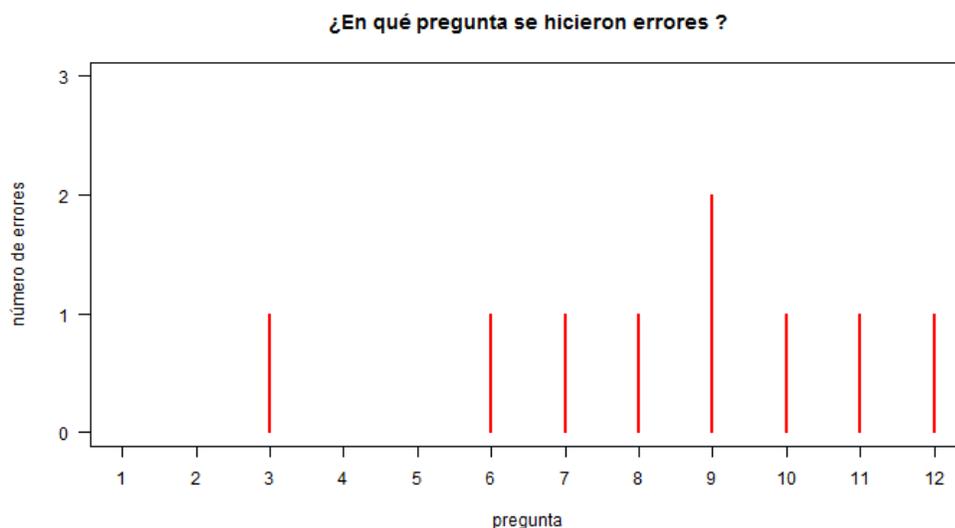


Ilustración 5: ¿En qué pregunta se hicieron errores?

En la Ilustración 5: ¿En qué pregunta se hicieron errores?, se nota que los errores se concentran en la segunda parte de las preguntas (preguntas 6 a 12), lo que podría provenir de un aburrimiento del participante. Otra razón podría ser que estas preguntas

en particular fueron particularmente difíciles, es decir que un nivel de dificultad en particular favorece los errores. Los niveles de dificultad de los casos con errores son los siguientes: {2; 5; 3; 2; 2; 5; 3; 2; 4}.

Para testear la hipótesis que el error es independiente del nivel de dificultad se hace un test de χ^2 .

$$\forall m \in \llbracket 1; 720 \rrbracket \text{ Err}_m = \begin{cases} 0 & \text{si la respuesta del caso } m \text{ es correcta} \\ 1 & \text{si no} \end{cases}$$

$$\forall m \in \llbracket 1; 720 \rrbracket, \text{ Dif}_m = \text{nivel de dificultad del caso } m$$

Hipótesis a testear: H_0 : *Err* y *Dif* son dos variables independientes

Se calcula un p-value = 0.1697, superior a 0.05. Entonces no se rechaza la hipótesis H_0 , es decir que el error es independiente del nivel de dificultad o en otras palabras los participantes que se equivocaron no lo hicieron por el nivel de dificultad.

¿Podría ser por aburrimiento? Para responder a esta pregunta se testea la independencia de las variables *Err* y de la variable que retorna el número de la pregunta con un test de χ^2 . La p-value es de 0.89, superior a 0.05, por lo que el error tampoco depende del número de la pregunta. También se prueba si el error es significativamente distinto en la primera mitad vs la segunda mitad de preguntas: se hace un test de χ^2 si el error y una variable *dummy* que indica si la pregunta está dentro de la primera o de la segunda mitad de las preguntas. La p-value es de 0.17, superior a 0.05, por consiguiente los errores no están estadísticamente situados en la segunda mitad de las preguntas. El número de preguntas, que queda relativamente bajo (12) en este experimento, no es fuente de error.

Un análisis de varianza ANOVA confirma estos resultados, entonces se concluye que los pocos errores se hicieron de manera aleatoria.

4.2. Número de fijaciones

Un primer paso de análisis del movimiento ocular consiste en observar cómo se comporta el número de fijaciones por caso. Cabe señalar que se toma en cuenta solamente las fijaciones en la matriz 6x4 de valores de atributos y alternativas, es decir los AOI 11 a 64 (ver Ilustración 3: Numeración de los AOIs), dejando de lado las fijaciones en las AOIs de instrucciones y en los botones de elección.

Se considera que el número de fijaciones en una parte de la pantalla (en una AOI) es un indicador de la atención y del interés que la persona puso en esta parte. Se pretende estudiar cuál es el nivel de atención que pone una persona para hacer una elección y su

variación con el tiempo y el nivel de dificultad de la pregunta. Se hace la hipótesis que la atención necesaria para responder una pregunta baja con el número de respuestas ya hechas y con el nivel de dificultad. Además, se verá si todos los participantes necesitan un mismo nivel de atención para responder a una pregunta. Por lo tanto, esta sección va a responder las siguientes preguntas:

- ¿Cuál es el número de fijaciones promedio?
- ¿Varía el número de fijaciones con la pregunta?
- ¿Varía el número de fijaciones con el nivel de dificultad?
- ¿Varía el número de fijaciones con la persona?

4.2.1. ¿Cuál es el número de fijaciones promedio?

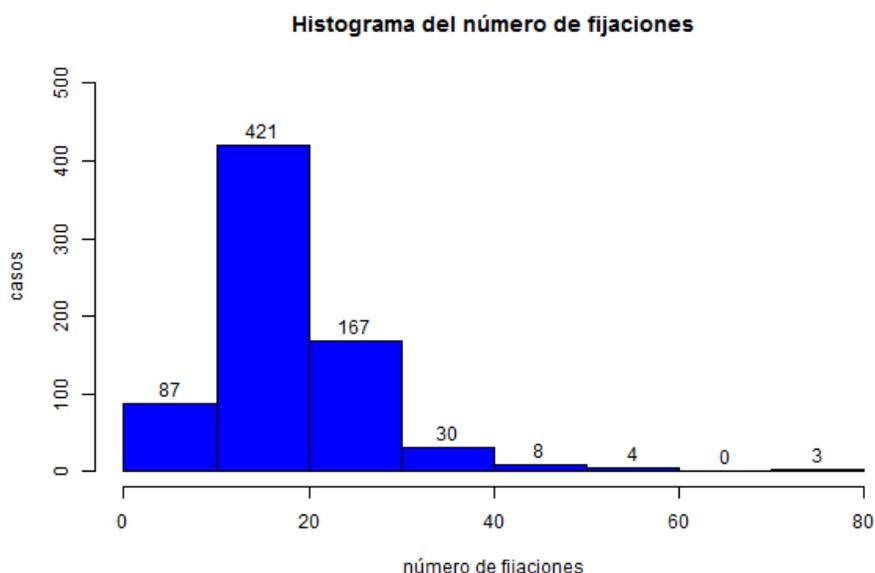


Ilustración 6: Histograma del número de fijaciones

Como se puede apreciar en la Ilustración 6: Histograma del número de fijaciones, en su mayoría los casos poseen entre 15 y 20 fijaciones. El número de fijaciones promedio es de 18.3 fijaciones por caso.

Existen dos situaciones particulares:

- Si el encuestado realizó pocas (entre 0 y 5) fijaciones, puede interpretarse de dos maneras:
 - como una eficiencia del participante (dependiendo del nivel de dificultad de la regla, 5 fijaciones pueden ser suficientes)
 - o como una falta de interés del participante por responder correctamente las preguntas.
- Si el encuestado realizó demasiadas fijaciones: Puede deberse a un problema de memoria, llamado de esta manera por la dificultad de recordar el valor de algunas

celdas de la matriz de atributos y alternativas. También, el encuestado puede volver al mismo AOI en una etapa de verificación.

Para considerar el efecto de memoria o de verificación, se introduce una segunda medida de las AOI fijadas, que mide el número promedio de AOI únicas miradas por caso. En una regla conjuntiva con dos atributos que tienen umbral, basta mirar máximo 8 AOI únicas (2 atributos por 4 alternativas). En la Tabla 2: Promedio y Des. Estándar del número de AOI y AOI únicas miradas, se nota que el número promedio de AOI únicas miradas es igual a 9.73, lo que indica que la gente tiende a mirar más AOI de lo que se necesita.

Variable	Promedio	Desviación estándar
Número de AOI miradas por caso	18.3	8.49
Número de AOI únicas miradas por caso	9.73	2.64

Tabla 2: Promedio y Des. Estándar del número de AOI y AOI únicas miradas

4.2.2. ¿Varía el número de fijaciones con la pregunta?

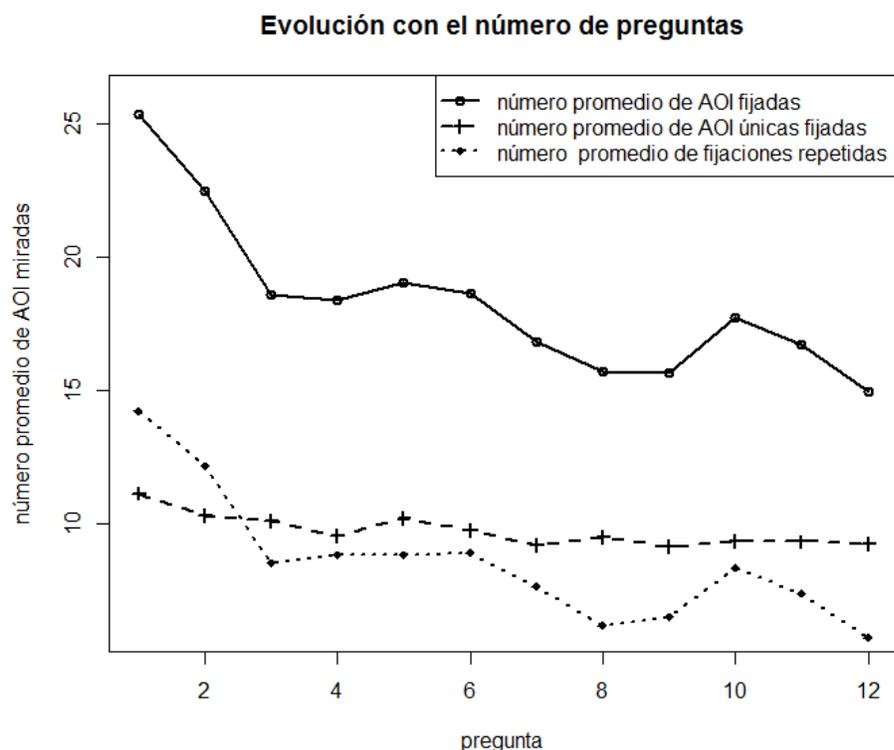


Ilustración 7: Número promedio de AOI y AOI únicas fijadas y fijaciones repetidas por pregunta

El gráfico de la Ilustración 7 permite deducir las siguientes hipótesis, que se testearan luego con tests estadísticos.

- (1) El número promedio de AOI miradas baja con las preguntas, lo que se puede atribuir a un aprendizaje del participante que se vuelve más eficiente con las preguntas.
- (2) El número promedio de AOI únicas miradas baja con el número de preguntas pero siempre queda superior a 8, lo que indica que los participantes en promedio miran más AOI de lo que necesitan.
- (3) El espacio entre las dos líneas se reduce, es decir el número de fijaciones repetidas (en una AOI ya mirada) baja con el número de preguntas, lo que señala que se memorizan y/o se verifican menos las informaciones.

Se hacen tres regresiones lineales: para las tres, las variables independientes son las mismas: el número de las preguntas como variable continua, el mismo al cuadrado para ver si existe un efecto no lineal y el nivel de dificultad como factor (5 variables dummy). Las variables dependientes son el número de fijaciones totales, el número de AOI únicas miradas y el número de fijaciones repetidas por caso.

Se definen las variables dummy :

$$\forall i \in [2; 6], \forall c \in [1; 720], \text{Niveldifi}_c = \begin{cases} 1 & \text{si el caso } c \text{ es del nivel de dificultad } i \\ 0 & \text{si no} \end{cases}$$

Las tres regresiones lineales siguen la siguiente ecuación:

$$Y = \alpha + \beta_{\text{pregunta}} * \text{Pregunta} + \beta_{\text{pregunta}^2} * \text{Pregunta}^2 + \sum_{i=2}^6 \beta_i * \text{Niveldifi}_i$$

Significancia y Coeficientes	Variable dependiente					
	(1) número de fijaciones totales		(2) número de AOI únicas miradas		(3) número de fijaciones repetidas	
Intercepto	***	23.76	***	10.13	***	13.63
Pregunta	***	-2.054	**	-0.3779	***	-1.677
Pregunta ²	***	0.1044	*	0.0183	***	0.086
Nivel dif 2		1.181	*	0.7217		0.459
Nivel dif 3	**	2.658	***	1.110	.	1.548
Nivel dif 4	***	4.768	***	1.779	***	2.999
Nivel dif 5	***	3.681	***	1.171	**	2.510
Nivel dif 6		1.372	**	1.595		-0.223
Códigos de significancia:		0.0001 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ''

Tabla 3: Regresiones lineales de los números de fijaciones por caso

La significancia y los coeficientes negativos de la variable “Pregunta” en las tres regresiones lineales de la

Significancia y Coeficientes		Variable dependiente					
		(1) número de fijaciones totales		(2) número de AOI únicas miradas		(3) número de fijaciones repetidas	
Intercepto		***	23.76	***	10.13	***	13.63
Pregunta		***	-2.054	**	-0.3779	***	-1.677
Pregunta ²		***	0.1044	*	0.0183	***	0.086
Nivel	dif 2		1.181	*	0.7217		0.459
Nivel	dif 3	**	2.658	***	1.110	.	1.548
Nivel	dif 4	***	4.768	***	1.779	***	2.999
Nivel	dif 5	***	3.681	***	1.171	**	2.510
Nivel	dif 6		1.372	**	1.595		-0.223
Códigos de significancia:		0.0001 ‘***’		0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘’

Tabla 3 confirman las hipótesis (1),(2) y (3). Se puede concluir que existe un aprendizaje de las personas, que necesitan menos fijaciones y menos verificación o memoria cuando avancen en las preguntas. Se destaca también la existencia de un efecto no lineal por la significancia de la variable “Pregunta²” y su coeficiente positivo, lo que denota un tope en el aprendizaje de las personas.

4.2.3.¿Varía el número de fijaciones con el nivel de dificultad?

Para estudiar el efecto del nivel de dificultad en los números de fijaciones (totales, únicas y repetidas) se grafican en la Ilustración 8 los coeficientes de las variables *Niveldifi* en las regresiones cuyos resultados están en la

Significancia y Coeficientes		Variable dependiente					
		(1) número de fijaciones totales		(2) número de AOI únicas miradas		(3) número de fijaciones repetidas	
Intercepto		***	23.76	***	10.13	***	13.63
Pregunta		***	-2.054	**	-0.3779	***	-1.677
Pregunta ²		***	0.1044	*	0.0183	***	0.086
Nivel	dif 2		1.181	*	0.7217		0.459
Nivel	dif 3	**	2.658	***	1.110	.	1.548
Nivel	dif 4	***	4.768	***	1.779	***	2.999
Nivel	dif 5	***	3.681	***	1.171	**	2.510
Nivel	dif 6		1.372	**	1.595		-0.223
Códigos de significancia:		0.0001 ‘***’		0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘’

Tabla 3. Cabe señalar que no todas las variables son significativas. El nivel de referencia usado en las tres regresiones es el nivel de dificultad 1, cuyo efecto se encuentra en el intercepto. Se ve que el nivel de dificultad influye en las fijaciones.

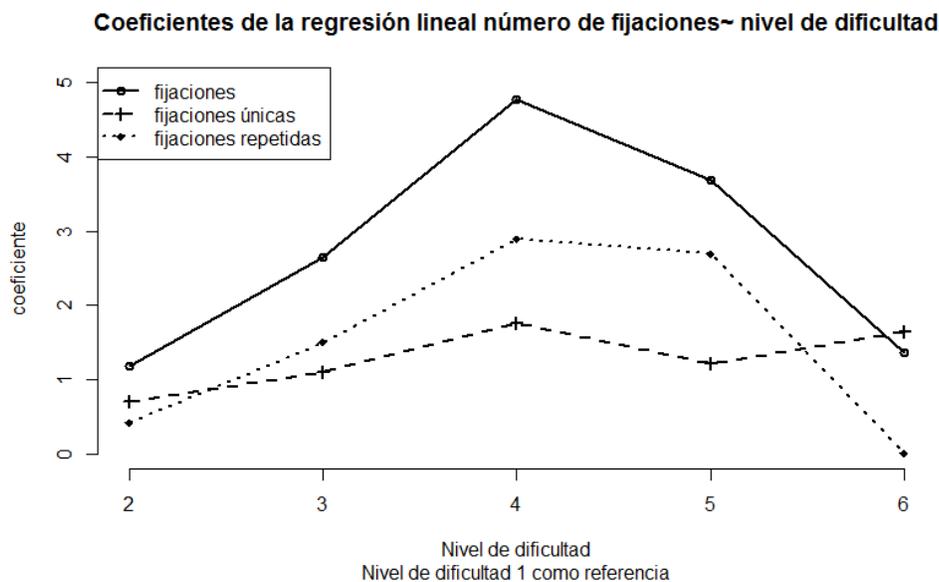


Ilustración 8: Coeficientes de los niveles de dificultad en las regresiones lineales (1), (2) y (3).

Se destacan las variables asociadas a los niveles 4 y 5: son significativas en las tres regresiones y tienen un coeficiente positivo. Los valores de los coeficientes indican que las preguntas con niveles de dificultad 4 y 5 requieren respectivamente alrededor de 5 y 4 fijaciones más que el nivel 1, lo que corresponde a un aumento importante dado que se hacen en promedio 18 fijaciones. Si se usa el nivel de dificultad 2,3 o 6 como referencia en vez del nivel 1, las variables de los niveles 4 y 5 siguen significativas y con coeficiente positivo en las regresiones con el número de fijaciones totales (1) y el número de fijaciones repetidas (3) pero no son significativas o tienen coeficientes menores en la regresión (2) con el número de AOIs únicas miradas. Se deduce que la razón por la cual los niveles 4 y 5 exigen más fijaciones totales que los demás niveles de dificultad es que requieren un mayor número de fijaciones repetidas y un igual número de fijaciones únicas. Es consistente con la definición de estos niveles (ver Tabla 1: Definición de los niveles de dificultad página 9): el participante tiene que acordarse de las alternativas que sobrepasan el primer atributo con umbral mientras está mirando el segundo, lo que implica más fijaciones de verificación o memoria.

4.2.4. ¿Varía el número de fijaciones con la persona?

Con el fin de descubrir cómo varía el número de fijaciones con la persona, se grafica el histograma del número promedio de fijaciones por caso de cada persona (ver Ilustración 9). No se ve una dispersión alta, es decir que el número de fijaciones no depende mucho de la persona. Se concluye por un lado que el número de fijaciones no da

mucha información acerca de cuál participante está respondiendo la pregunta, y de otro lado que las conclusiones del análisis del número de fijaciones son robustas a cambios de personas.

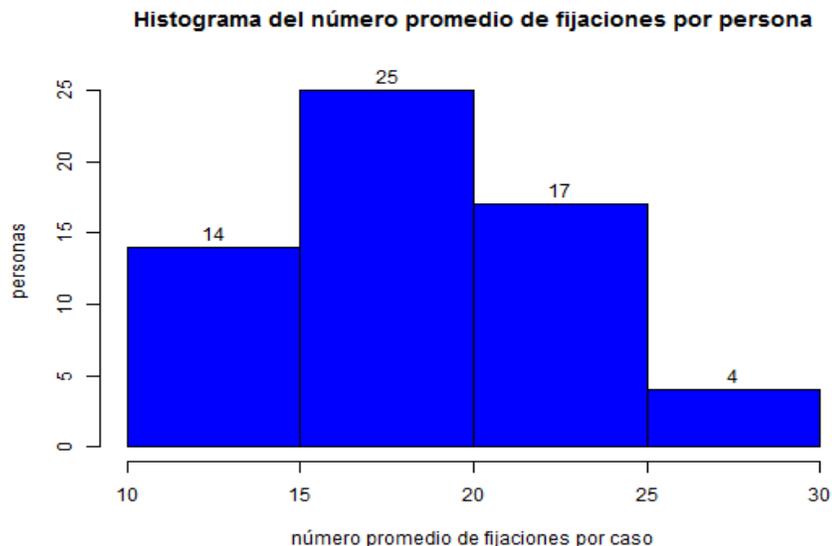


Ilustración 9: Histograma del número promedio de fijaciones por persona

4.3. Transiciones

Las transiciones entre las AOIs pueden dar información relevante sobre el razonamiento seguido por los participantes y las trayectorias de su atención.

Se clasifican las transiciones en tres grupos:

- Transiciones verticales: transiciones dentro de una misma alternativa.
- Transiciones horizontales: transiciones dentro de un mismo atributo.
- Transiciones diagonales: transiciones que no son verticales u horizontales.

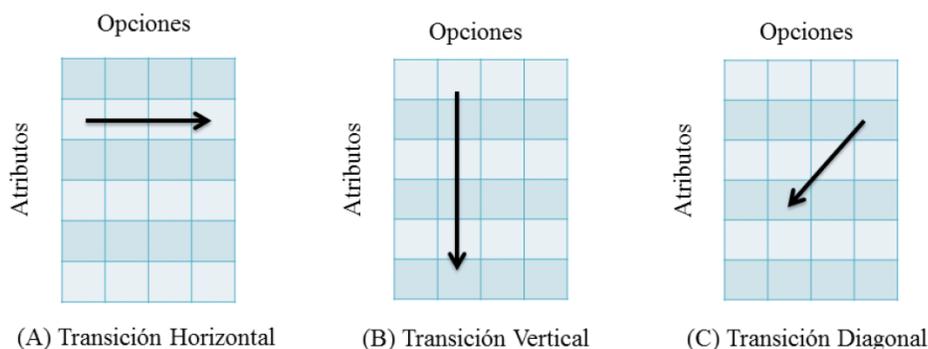


Ilustración 10: Transiciones posibles (fuente: [9])

Se calculan dos matrices de transiciones, una calculando las transiciones entre productos (ver Tabla 4) y la segunda entre atributos (ver Tabla 5). En la celda (i, j) de una de estas matrices se encuentra la probabilidad de pasar del objeto (producto o atributo) i al objeto j .

Productos	1	2	3	4
1	54%	34%	9%	3%
2	25%	45%	25%	4%
3	8%	19%	38%	36%
4	14%	19%	24%	44%

Tabla 4: Matriz de transiciones entre productos

Por ejemplo en la Tabla 4: Matriz de transiciones entre productos, la primera línea indica que, dentro de todas las transiciones que salen del producto 1, el 54% se quedan dentro del mismo producto, el 34% van a mirar el producto 2, el 9% el producto 3 y el 3% el producto 4. La segunda columna de esta matriz informa sobre la proveniencia de las transiciones que llegan al producto 2: el 34% llegan del producto 1, el 45% del producto 2, etc.

Atributos	1	2	3	4	5	6
1	65%	13%	5%	5%	5%	8%
2	21%	53%	12%	4%	4%	6%
3	7%	16%	55%	10%	5%	6%
4	4%	5%	11%	62%	12%	6%
5	4%	4%	4%	12%	62%	13%
6	8%	6%	5%	7%	6%	58%

Tabla 5: Matriz de transiciones entre atributos

Se observan varios aspectos comunes a las dos matrices. Primero, los valores de la diagonal son mayores a aquellas fuera de la diagonal, lo que refleja que los participantes tienen más probabilidad de seguir mirando el mismo atributo (transición horizontal) o el mismo producto (transición vertical) que de moverse a otro (transición diagonal). Si el participante no se queda mirando el mismo objeto, se movería en los objetos cercanos, lo que traduce una continuidad de la mirada.

Sin embargo, se diferencian las dos matrices. Se ve en la Tabla 6: Valores de las diagonales de las matrices de transiciones que el porcentaje promedio de transiciones dentro del mismo atributo es superior al porcentaje promedio de transiciones dentro del mismo producto, lo que indica que los participantes examinan la pantalla más atributo

por atributo que producto por producto. Este comportamiento es consistente con la regla conjuntiva impuesta, que obliga al participante a mirar dos atributos en particular (los dos atributos con umbral).

Valores de la diagonal	Promedio	Desviación Estándar
Productos	45	6,68
Atributos	59	4,62

Tabla 6: Valores de las diagonales de las matrices de transiciones

4.4. Predicción de la elección

A fin de hacer una primera predicción de elección con el comportamiento ocular, se busca ver en qué porcentaje de los casos el producto más mirado coincide con el producto elegido. Este porcentaje es de 53.1% (ver Tabla 7), a comparar con la probabilidad básica de un producto de ser elegido al azar dentro de cuatro alternativas que es de 25%.

Un comportamiento visual posible es de ir escaneando primero todas las alternativas disponibles, restringiendo poco a poco en las que sobreviven hasta fijarse en las últimas fijaciones en la alternativa que uno va a elegir. Con la intención de examinar si los participantes siguen este comportamiento, se separan las fijaciones en dos partes y se mira a la segunda parte de las fijaciones. Se calcula el producto más mirado en la segunda parte de las fijaciones, y también el último producto mirado para probar la hipótesis, que las personas vuelven a mirar el producto que van a elegir en los últimos pasos o al final de su razonamiento.

Se ve en la Tabla 7 que el último producto mirado, el producto más mirado y el producto más mirado en la segunda parte de las fijaciones son, en este orden, mejores indicadores del producto elegido. Se confirma la hipótesis según la cual las personas vuelven a mirar el producto que van a elegir al final de su razonamiento.

Variable	Valor
Porcentaje de casos en los cuales producto elegido= producto más mirado $Per_{ch=MF}$	53.1
Porcentaje de casos en los cuales producto elegido= producto más mirado en la segunda parte de las fijaciones $Per_{ch=MFSH}$	46.5
Porcentaje de casos en los cuales producto elegido= último producto mirado $Per_{ch=LF}$	61.1

Tabla 7: Porcentaje de casos en los cuales producto elegido= producto más mirado

Se denota que el porcentaje $Per_{ch=MF\&SH}$ queda inferior a $Per_{ch=MF}$ y a $Per_{ch=LF}$. En consecuencia, se deja de lado el cálculo del producto más mirado en la segunda parte de las fijaciones.

Para ver si la variable “producto más mirado” tiene sentido, se tiene que verificar que el producto más mirado se separa bien del resto en términos de fijaciones, y que hay pocos empates. Se grafica en la Ilustración 11 el porcentaje de fijaciones en el atributo más mirado, en el segundo más mirado, etc. Se ve que el producto más mirado abarca en promedio el 44% de las fijaciones y se demarca del resto, lo que confirma la validez de la variable.

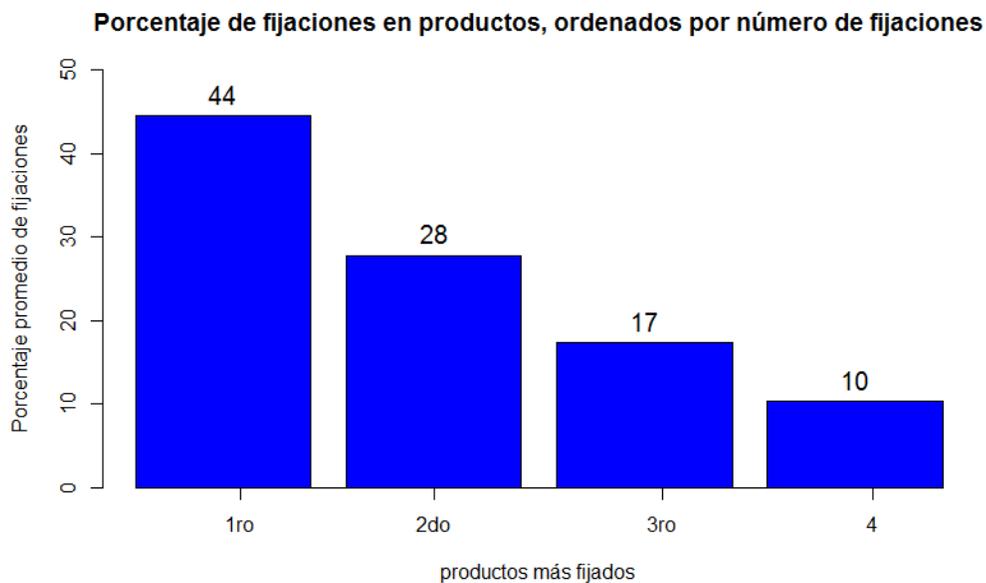


Ilustración 11: Porcentajes de fijaciones en productos más mirados

4.4.1. Efecto de la posición horizontal

Se construyen las tablas de contingencias del producto elegido con, respectivamente, el producto más fijado y el último producto fijado (ver Anexos a. Tablas de contingencia). Estas dos tablas permiten calcular para un producto en posición horizontal $i \in [1; 4]$ tres probabilidades condicionales que se grafican en la Ilustración 12: Probabilidades condicionales de elección:

- la probabilidad de ser elegido dado que es el más fijado, en azul en la Ilustración 12 :

$$\mathbb{P}(Ch|MF)_i$$

- la probabilidad de ser elegido dado que es el último producto fijado, en rojo en la Ilustración 12:

$$\mathbb{P}(Ch|LF)_i$$

- la probabilidad de ser elegido dado que es el último producto fijado y el más fijado, en verde en la Ilustración 12.

$$\mathbb{P}(Ch|MF\&LF)_i$$

La Ilustración 12: Probabilidades condicionales de elección permite ver la influencia de la posición horizontal de un producto sobre sus probabilidades condicionales de ser elegido. En general, las tres son altas (superiores a 0.5) y $\mathbb{P}(Ch|MF\&LF)$ para un cualquier producto es de 0.8.

Con respecto a la probabilidad $\mathbb{P}(Ch|MF)_i$, se destacan las posiciones 3 y 4 cuyos productos tienen más poder predictivo de ser el más fijado. Se puede explicar por el hecho que los participantes del experimento, de nacionalidad alemán, leen de izquierda a derecha. Entonces un comportamiento posible usando una regla conjuntiva es de mirar primero los productos a la izquierda (1 y 2) y de moverse a la derecha solo si estos productos no son satisfactorios. En este razonamiento, el producto 4 recibe atención solo si se han eliminado las alternativas a su izquierda, lo que implica que el producto 4 va a ser elegido dado que el diseño del experimento impone que una alternativa supere el umbral. Este resultado parece indicar que, en condiciones de elecciones reales, un producto situado muy a la derecha tendrá menos probabilidad de ser elegido dado que hay más probabilidad de encontrar un producto satisfactorio antes de llegar a su posición. Además sería interesante confirmar si este resultado “se da vuelta” para consumidores de idiomas que se leen de derecha a izquierda como el árabe o el hebreo.

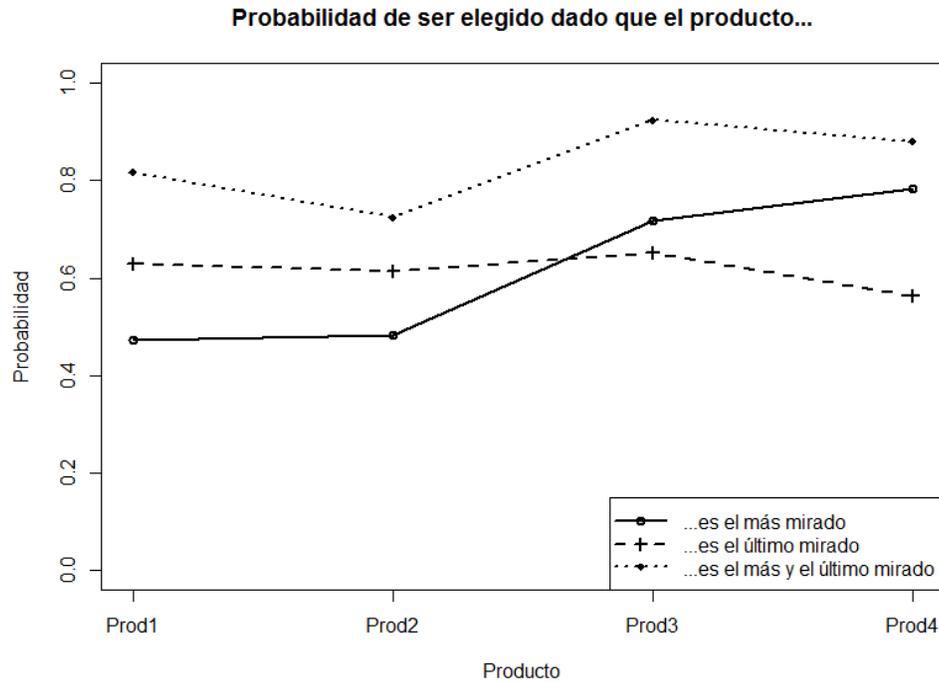


Ilustración 12: Probabilidades condicionales de elección

4.4.2. Efecto del número de preguntas y del nivel de dificultad

A continuación se estudia el efecto del número de preguntas y del nivel de dificultad en las probabilidades condicionales de elección. Se grafica en la Ilustración 13 el porcentaje por pregunta de los casos en los cuales:

- el producto más mirado corresponde al producto elegido.
- el último producto mirado corresponde al producto elegido.
- el último producto mirado corresponde al producto elegido y al producto elegido.

Para pasar de estos porcentajes a las probabilidades condicionales se debe multiplicar por una constante positiva. Las tres líneas se ven más o menos horizontal, lo que parece indicar que este efecto no depende de la pregunta. Tres regresiones logísticas confirman esta hipótesis con una variable “pregunta” no significativa. Se concluye que las probabilidades condicionales de elección no dependen del número de preguntas, es decir son indicadores de elección robustos a cambios en la cantidad de elecciones.

Además, el nivel de dificultad tampoco es significativo en las regresiones logísticas, por lo tanto son también robustos a cambio de dificultad de la elección.

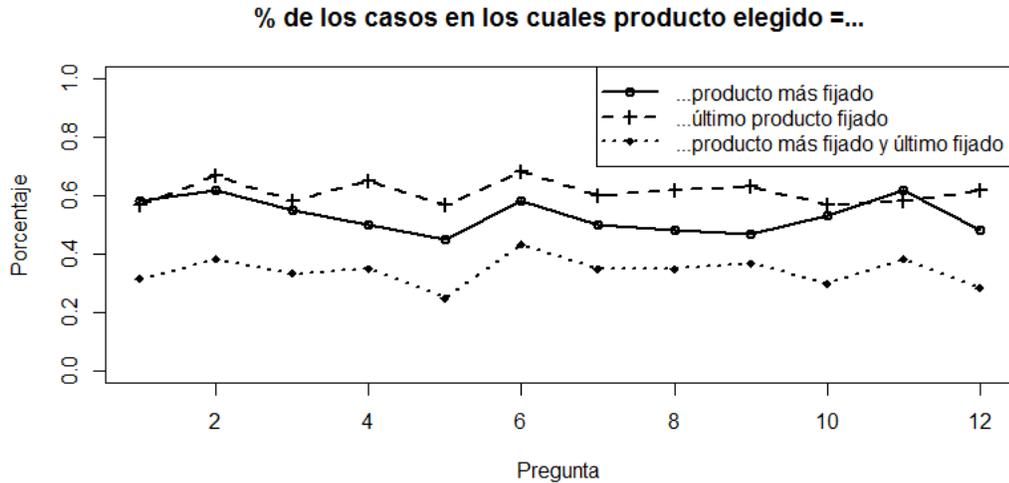


Ilustración 13: Porcentaje de los casos en los cuales el producto elegido es el más mirado y/o el último mirado

4.4.3. Efecto del participante

Con el objetivo de ver el efecto de la persona se grafica en la Ilustración 14, el histograma del promedio por persona del porcentaje de los casos en los cuales el producto más mirado corresponde al producto elegido. Estos histogramas parecen indicar que los porcentajes varían con la persona. El test de independencia de χ^2 da un p-valor inferior a $2 * 10^{-6}$ para los tres porcentajes, lo que indica que sí dependen de las personas. Por falta de más información acerca de los participantes no se puede investigar más la causa de esta heterogeneidad.

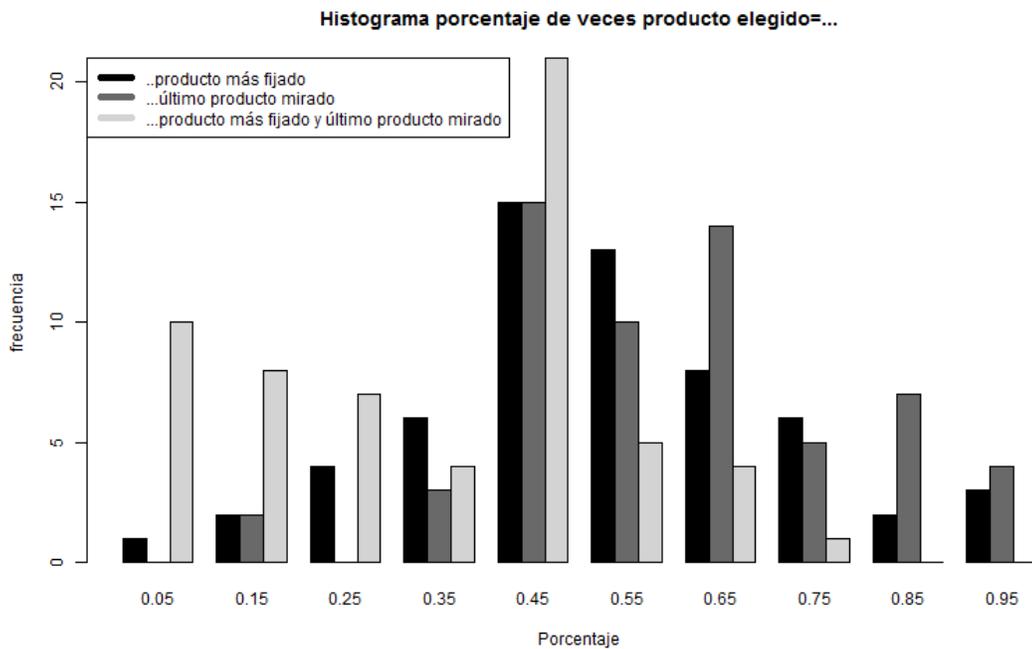


Ilustración 14: Histograma del porcentaje de veces en las cuales producto elegido es el más mirado o el último mirado por persona

4.5. Conclusiones del capítulo

En resumen, los participantes aprenden y son más eficientes haciendo menos fijaciones con el tiempo, pero todavía más que lo necesario. El nivel de dificultad también influye en el número de fijaciones: los niveles 4 y 5 requieren más fijaciones de verificación o de memoria. El número de fijaciones puede ser un indicador del número de preguntas respuestas anteriormente o del nivel de dificultad de una pregunta.

Por otro lado, no existe mucha heterogeneidad en el número de fijación promedio por participante, lo que indica que el número de fijaciones no da mucha información acerca de cuál participante está respondiendo la pregunta. Las conclusiones sobre el número de preguntas o el nivel de dificultad quedan las mismas independientemente de la persona.

Además, el hecho de ser el producto más mirado y/o el último producto mirado son buenos indicadores de elección, robustos a cambios en el número de preguntas y en el nivel de dificultad.

5. INFERIR LA REGLA CONJUNTIVA CON ELECCIONES

Para predecir las decisiones futuras de una persona se pueden observar sus elecciones pasadas. Este es el enfoque tradicional del análisis conjunto en marketing, en el cual se calculan las preferencias de los participantes por los distintos atributos y sus niveles. Este capítulo va más allá de solamente deducir preferencias: la idea es inferir la regla de decisión conjuntiva usada por cada participante, usando solamente sus elecciones hechas. Esta regla permite deducir el proceso mental de elección del participante y prever su respuesta en elecciones futuras. Para cada persona, se separan las 12 preguntas respondidas en un conjunto de preguntas de entrenamiento que representen las elecciones hechas y un conjunto de preguntas de prueba que corresponden a las elecciones futuras.

5.1. Insumos y resultados del algoritmo

En este capítulo se programa un algoritmo para inferir la regla conjuntiva usada por el participante a partir de sus elecciones en las i primeras preguntas (de las 12 preguntas respondidas, $i \in [1; 12]$), que representan el conjunto de entrenamiento. Se eligen las i primeras preguntas como conjunto de entrenamiento y las $12 - i$ últimas preguntas como conjunto de prueba (y no preguntas surtidas al azar) porque es la metodología que se usa en la realidad: se tienen las elecciones pasadas de un consumidor y se busca predecir su comportamiento futuro.

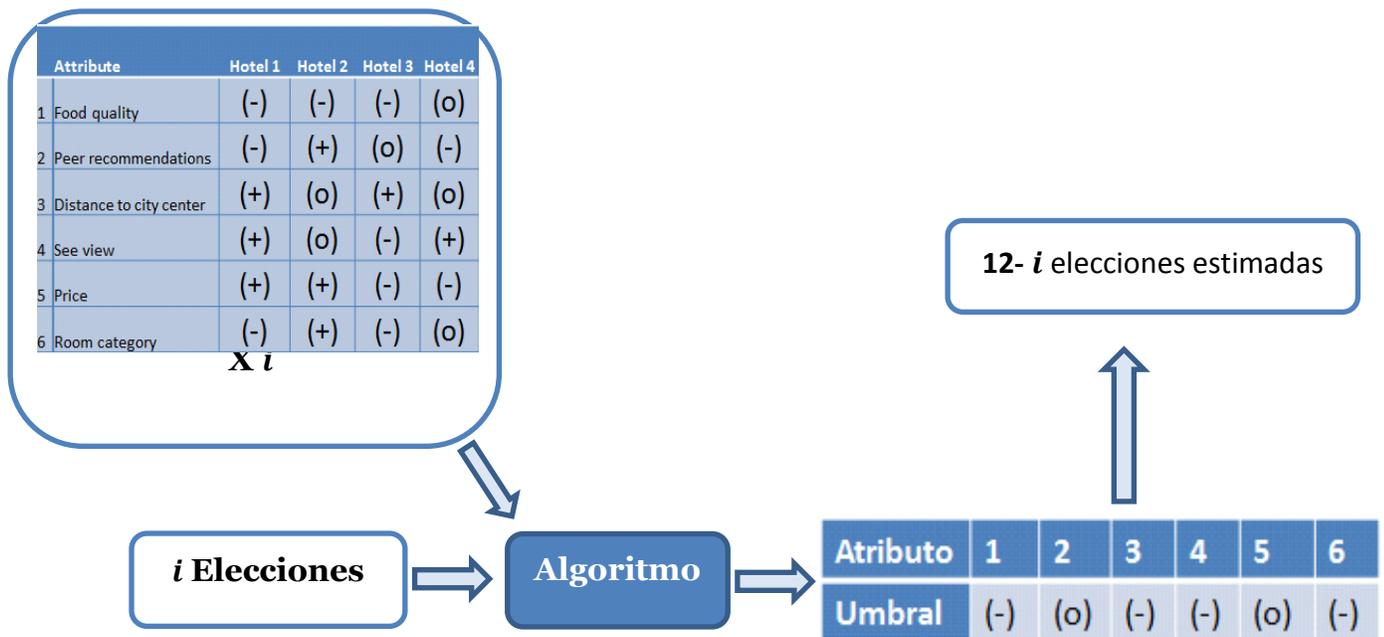


Ilustración 15: Input y Output del algoritmo con elecciones

Se busca determinar el tipo de regla conjuntiva usada por cada participante, es decir el umbral de cada atributo. El resultado del algoritmo corresponde entonces a un vector de 6 valores, que corresponden al nivel mínimo aceptado para cada atributo y que pueden tomar el valor 1 (umbral en el nivel bajo (-), o sin umbral), 2 (umbral en el nivel medio (o)) o 3 (umbral en el nivel alto (+)).

Se usan como insumos para cada pregunta en $\llbracket 1; i \rrbracket$ ($i < 12$) :

- La matriz 6x4 de valores de cada alternativa en cada atributo.
- La elección hecha por el participante (valor $\in \llbracket 1; 4 \rrbracket$).

Una vez que se tienen los umbrales estimados de cada participante, se puede reconstituir las elecciones que hubiera hecho el participante con esta regla en las $12 - i$ preguntas restantes (conjunto de prueba) y comparar estas elecciones estimadas con las elecciones reales.

5.2. Evaluación de los algoritmos

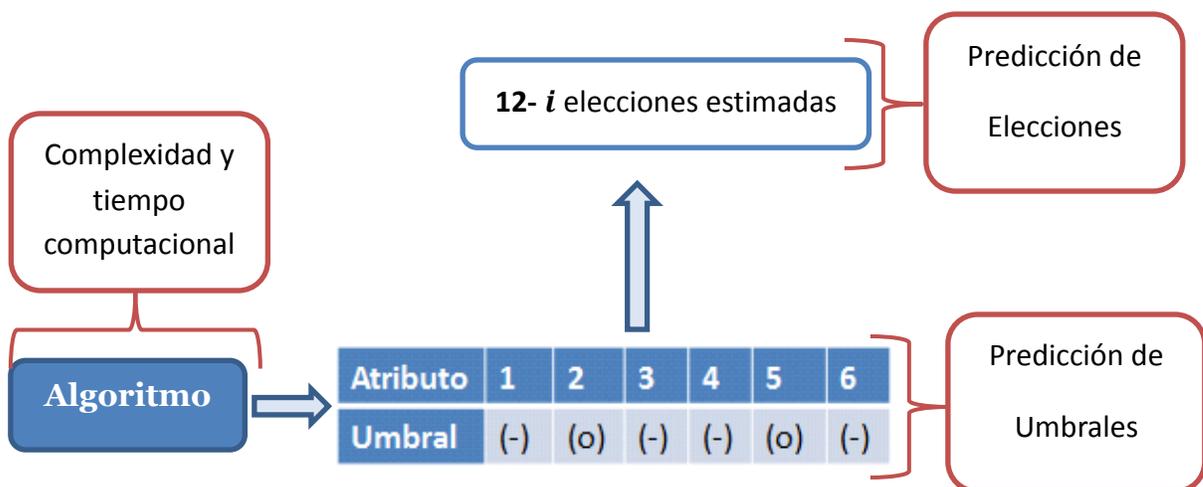


Ilustración 16: Evaluación de los algoritmos

Este trabajo busca encontrar algoritmos o modelos para inferir en poco tiempo reglas de decisiones de consumidores que reflejan su pensamiento y que pueden servir para predecir elecciones futuras. En consecuencia, los distintos algoritmos se evaluarán según tres ejes:

- Predicción de umbrales:

Para cada persona y para cada atributo, se compara el umbral predicho (un valor que puede ser 1, 2 o 3) con el umbral dado en las instrucciones (un valor que puede ser 1 o 2 dado que el experimento fue diseñado con umbrales solamente en el nivel medio). Se cuenta para cuantos atributos (de 0 a 6) coincidan los dos valores.

- Predicción de elecciones:

Se ve qué tan bien los umbrales (es decir la regla) predichos por el algoritmo con el conjunto de entrenamiento reflejan bien las elecciones hechas por el participante en el conjunto de prueba. Con este fin se crea una función *hit rate* que retorna para una combinación de umbrales y un conjunto de preguntas de una persona una medida acerca de “cuál porcentaje de elecciones hechas en el conjunto de preguntas haría el participante si tuviera esta combinación de umbrales”.

$$\text{hitrate}(\text{umbrales} \in \llbracket 1; 3 \rrbracket^6, \text{persona} \in \llbracket 1; 60 \rrbracket, \text{conjpreguntas} \subset \llbracket 1, 12 \rrbracket)$$

Para un participante dado y una pregunta dada, sean *Pass* el conjunto de alternativas que pasan el umbral y $n_{\text{pass}} = |\text{Pass}|$.

$$\left\{ \begin{array}{ll} \text{si } \left\{ \begin{array}{l} n_{\text{pass}} = 1 \\ \text{Pass} = \text{elección} \end{array} \right. & \text{entonces } \text{hit} = 1 \\ \text{si } \left\{ \begin{array}{l} n_{\text{pass}} > 1 \\ \text{elección} \in \text{Pass} \end{array} \right. & \text{entonces } \text{hit} = \frac{1}{n_{\text{pass}}} \\ \text{si no} & \text{hit} = 0 \end{array} \right.$$

Se calcula el hit para cada pregunta que pertenece a *conjpreguntas*. Al final se suman los hits obtenidos y la *hit rate* se calcula así :

$$\text{hitrate} = \frac{\sum_{\text{conjpreguntas}} \text{hit}}{|\text{conjpreguntas}|}$$

En conclusión la función *hit rate* permite calcular qué tan bien una combinación de umbrales refleja las elecciones del conjunto de preguntas *conjpreguntas*. Se usa en la evaluación de modelos para ver si los umbrales predichos por el modelo pueden predecir las elecciones hechas en el conjunto de prueba.

- Complejidad y tiempo computacional: se busca el algoritmo menos costoso.

5.3. Pasos de los algoritmos

En esta parte se va a presentar dos algoritmos: uno de enumeración completa y otro, llamado *Greedy*, menos costoso computacionalmente. Los dos usan la misma función *hit rate* que se usa en la evaluación de modelos pero esta vez la usan con preguntas del conjunto de entrenamiento (y no de prueba): testean varias combinaciones de umbrales para seleccionar la que tiene la mejor *hit rate* en el conjunto de entrenamiento. Difieren en el número de combinaciones distintas a probar: el algoritmo de enumeración testea

todas las combinaciones posibles mientras el *Greedy* destaca algunas combinaciones a probar.

5.3.1. Algoritmo de enumeración completa

Se testea un primer algoritmo que enumera todas las combinaciones posibles de umbrales. Dado que se tienen 6 atributos y que cada uno tiene 3 niveles, el universo comprende $3^6 = 729$ combinaciones.

El algoritmo aplica la función *hit rate* a todas las combinaciones, incrementando de (1,1,1,1,1,1) a (3,3,3,3,3,3), y a las preguntas del conjunto de entrenamiento. Retorna la combinación que tiene la mayor *hit rate* (puede ser 100%). Si varias combinaciones alcanzan la mayor *hit rate*, el algoritmo se queda con una combinación con los umbrales más bajos de las dos combinaciones. Por ejemplo, si las dos combinaciones $c1=(1,1,2,1,3,1)$ y $c2=(1,2,2,1,2,1)$ tienen una *hit rate* de 100%, el algoritmo guardará (1,1,2,1,2,1). De esta manera se retorna, dentro de las reglas con mayor *hit rate* en el conjunto de entrenamiento, la regla menos exigente, que corresponde a las preferencias más generales expresadas por la persona en el conjunto de entrenamiento.

5.3.2. Heurística *Greedy*

El algoritmo *Greedy* se inspira de un algoritmo del mismo nombre que ya existe para inferir reglas lexicográficas (ver [9] y [10]). Se modificó de manera de aplicarse a reglas conjuntivas.

Se hacen 6 bucles $b \in \llbracket 1; 6 \rrbracket$, cada bucle determina el valor del umbral de un atributo.

El primer bucle parte del umbral inicial (1,1,1,1,1,1), por el cual todas las alternativas son aceptables y de una *hit rate* en el conjunto de entrenamiento de 25% que corresponde a una selección al azar. Se calculan las *hit rates* en el conjunto de entrenamiento de las combinaciones (2,1,1,1,1,1) y (3,1,1,1,1,1) incrementando el atributo 1 por separado, luego las *hit rates* de las combinaciones (1,2,1,1,1,1) y (1,3,1,1,1,1) incrementando el atributo 2 por separado, etc... hasta incrementar el atributo 6. Se tiene al final una matriz de $6 * 2 = 12$ *hit rates*, de las cuales se calcula la máxima. Si esta *hit rate* máxima no es inferior a la máxima *hit rate* encontrada anteriormente, se actualiza la combinación inicial de umbrales modificando el atributo que tiene la máxima *hit rate* con el nivel en el cual la obtiene.

Se repiten estos pasos para el segundo bucle, probando los 5 atributos que no han sido modificados antes. Se para el algoritmo si se encuentra una *hit rate* en el conjunto de entrenamiento de 100%.

En el bucle $b \in \llbracket 1; 6 \rrbracket$ se testean $(7 - b) * 2$ combinaciones de umbrales, en consecuencia el máximo número de umbrales a probar por participante es de

$$\sum_{b=1}^6 (7 - b) * 2 = 42, \text{ a comparar con } 729 \text{ en el algoritmo enumerativo.}$$

5.4. Resultados

En esta sección se comparan los resultados de los dos algoritmos, en términos de predicción de umbrales, de predicción de elecciones y de complejidad.

5.4.1. Predicción de los umbrales

Para cada persona y para cada atributo, se compara el umbral predicho con un conjunto de entrenamiento de 10 preguntas con el umbral dado en las instrucciones. Se cuenta para cuantos atributos (de 0 a 6) coinciden los dos valores (ver Tabla 8). La columna i (de 0 a 6) de la tabla nos da el número de personas por las cuales el algoritmo hizo i errores en los umbrales.

Por ejemplo para la heurística *Greedy*, la primera columna indica que encontró los 6 límites correctos para 55 personas sobre 60. La tasa de acierto calcula el número total de umbrales correctamente predichos (acá $55 * 6 + 1 * 4 + 4 * 2 = 342$) sobre el número total de umbrales a predecir (6 por persona, es decir $6 * 60 = 360$). La tasa de acierto del *Greedy* es de 95% con 10 preguntas de entrenamiento, lo que se considera bueno. Se puede comparar con la tasa de acierto de una selección al azar que sería de 33% dado que cada atributo tiene 3 niveles.

Número de atributos que tienen un umbral predicho distinto al umbral verdadero	0	1	2	3	4	5	6	Tasa de acierto
Número de participantes- <i>Greedy</i>	60	0	0	0	0	0	0	100%
Número de participantes- Enumeración	60	0	0	0	0	0	0	100%

Tabla 8: Predicción de los umbrales Enumeración-Greedy

Las tasas de acierto del *Greedy* y del algoritmo de enumeración con 10 preguntas de entrenamiento por persona alcanzan un 100%. Cabe recordar que estas tasas resultan del hecho que todos los participantes usaron reglas conjuntivas en este experimento.

Con la intención de ver la evolución de la primera columna con tamaño del conjunto de entrenamiento (de 1 a 12 preguntas) se grafica en la Ilustración 17 el porcentaje de personas por las cuales el algoritmo de enumeración (respectivamente *Greedy*) encontró la combinación de 6 umbrales correcta, o sea la combinación dada por

las instrucciones que siguen los participantes. Primero se calcula sin tener en cuenta los errores (líneas roja y purpura) para ver el rendimiento de los algoritmos en el mejor caso. Además se grafican las líneas azul y verde que toman en cuenta los errores hechos por los participantes.

Evolución de la predicción de umbrales con tamaño del conjunto de entrenamiento

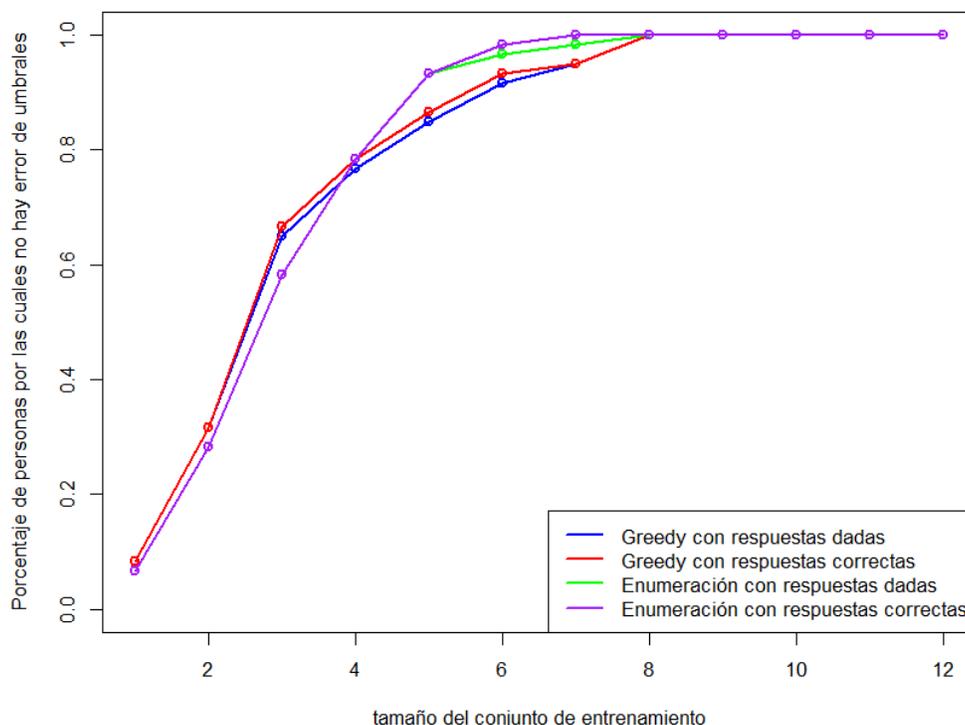


Ilustración 17: Evolución del ajuste de los modelos *Greedy* y de enumeración con el tamaño del conjunto de entrenamiento.

Se notan ajustes mejores al 85% con 5 preguntas de entrenamiento o más. Queda espacio por mejorar la predicción de umbrales en conjuntos de entrenamiento con menos de 5 preguntas.

5.4.2. Predicción de las elecciones

Se separan las preguntas de cada participante en dos conjuntos, uno de entrenamiento, con i preguntas, y el otro de test, con las $12 - i$ preguntas restantes. Se darán acá los resultados para $i = 10$ para ver el aspecto general de las curvas e $i = 6$ para hacer zoom sobre pequeños conjuntos de entrenamiento.

Para cada i , se calcula el umbral estimado con $j \leq i$ preguntas de entrenamiento. Luego, para ver si este umbral refleja bien las elecciones del participante, se calcula la *hit*

rate del umbral estimado en las $12 - i$ preguntas de test. Una *hit rate* es de 100% significa que el umbral estimado predijo todas las elecciones del conjunto de test y entonces revela bien la regla de decisión conjuntiva usada por el participante.

Caso $i = 10$:

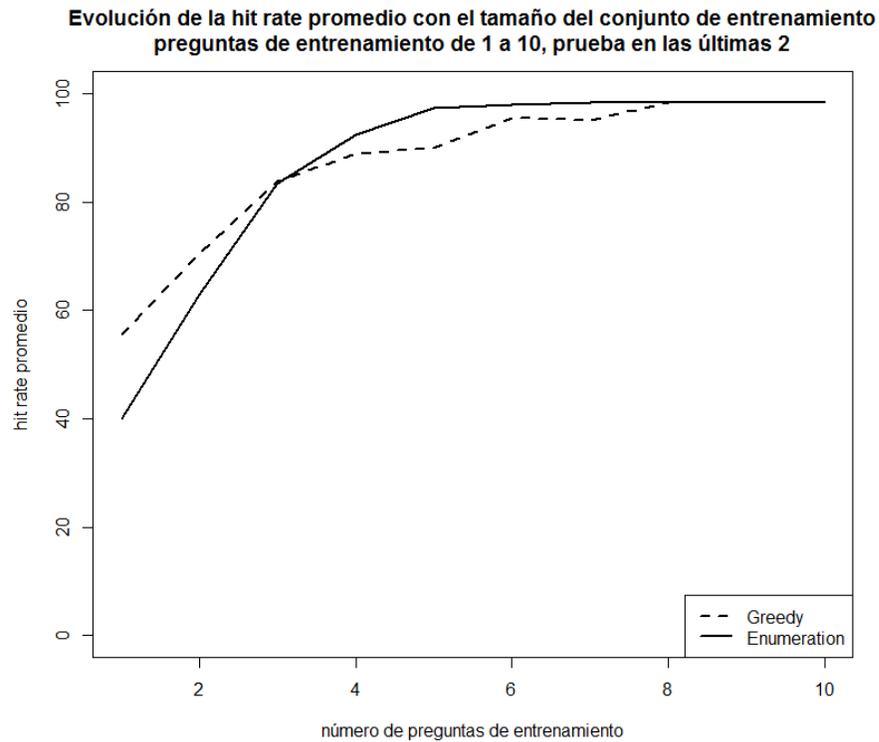


Ilustración 18: Hit rate Greedy vs enumeración, 10 preguntas de entrenamiento

En la Ilustración 18 se grafica la *hit rate* sobre las 2 preguntas de test de los umbrales estimados con $j \in \llbracket 1; 10 \rrbracket$ preguntas. Con 10 preguntas de entrenamiento, el algoritmo enumerativo llega al 98.3% de ajuste y el *Greedy* al 98%, y no al 100% por los errores de algunos participantes. La diferencia entre el *Greedy* y el enumerativo en las primeras preguntas se podría explicar por el hecho que el algoritmo enumerativo escoge la combinación de umbrales más general dentro de las que maximizan la *hit rate*.

Caso $i = 6$:

En la Ilustración 19 se grafica la *hit rate* sobre las 6 preguntas de test de los umbrales estimados con $j \in \llbracket 1; 6 \rrbracket$ preguntas. Se nota que se requieren al menos 3 preguntas para sobrepasar un 80% de ajuste.

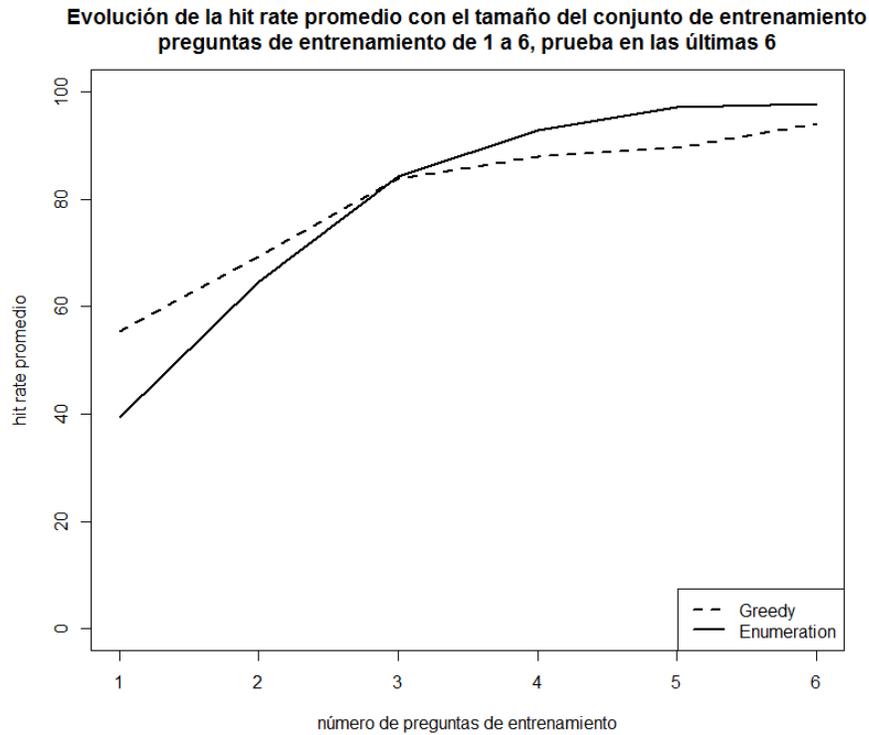


Ilustración 19: Hit rate Greedy vs enumeración, 6 preguntas de entrenamiento

5.4.3. Tiempo computacional y complejidad

Se anotaron los tiempos computacionales de los dos algoritmos. Se espera que el algoritmo de enumeración se demore más tiempo que el *Greedy*, dado que tiene 17 veces más combinaciones de umbrales a probar. Se comparan los tiempos necesarios para hacer la Ilustración 18 y la Ilustración 19 en la Tabla 9. Los cálculos se hacen con el programa R en un computador con un procesador Intel Xeon de RAM 8 GB y frecuencia 2.4 GHz.

Tiempos en s	Enumeración	Greedy	Factor
Ilustración 18(10-12)	2056	72	29
Ilustración 19 (6-6)	769	28	28

Tabla 9: Tiempos computacionales de los algoritmos Greedy y de enumeración

Esta diferencia en los tiempos computacionales se explica por la complejidad de los algoritmos. Si p es el número de personas, a el número de atributos y l el número de niveles para cada atributo, sin contar la complejidad de *hit rate*, *Greedy* tiene una complejidad en $o(p * a^2 * l)$ y el algoritmo de enumeración en $o(p * l^a)$.

5.5. Conclusiones del capítulo

Los dos algoritmos probados para inferir la regla de decisión con elecciones obtienen un desempeño muy alto: 8 preguntas de entrenamiento bastan para encontrar todas las reglas conjuntivas usadas por los participantes y con eso lograr un ajuste de 100% en la predicción de elecciones. Con solamente 3 preguntas de entrenamiento se encuentra la regla correcta para la mitad de los participantes y predicen el 80% de las elecciones futuras.

Los dos algoritmos obtienen resultados similares en términos de predicción de umbrales y de elecciones pero difieren mucho en la complejidad: el algoritmo de enumeración es mucho más costoso que el *Greedy*. En consecuencia, se considera al *Greedy* como el mejor de los dos.

Estos resultados muy buenos no dejan mucho espacio para mejorar agregando información del comportamiento visual, no obstante se deben al hecho que los participantes siguieron una regla conjuntiva como les fue indicado en las instrucciones. En la vida real se supone que las personas no necesariamente usan este tipo de reglas para elegir. En el capítulo 7 se aplicarán estos modelos a un experimento de elección libre para comprobar estos modelos a la realidad de las decisiones.

6. INFERIR LA REGLA CONJUNTIVA CON MOVIMIENTOS DE OJOS

Este capítulo tiene como meta inferir la regla de decisión usada por cada participante, pero esta vez usando solamente el comportamiento ocular. Este enfoque puede ser útil cuando se tiene acceso a los movimiento de ojos de una persona (con lentes de tipo Google Glass por ejemplo) pero no a sus elecciones (elección escondida o no concluida). Se quiere estudiar si los movimientos de ojos de una persona entregan menos, igual o más información que sus elecciones sobre su regla de decisión conjuntiva.

6.1. Insumos y resultados del algoritmo

Se prueban distintos algoritmos para inferir la regla conjuntiva usada por el participante a partir de sus fijaciones en i de las 12 preguntas respondidas ($i \in \llbracket 1; 12 \rrbracket$), que representan el conjunto de entrenamiento. En este capítulo se considerará solamente un $i = 10$.

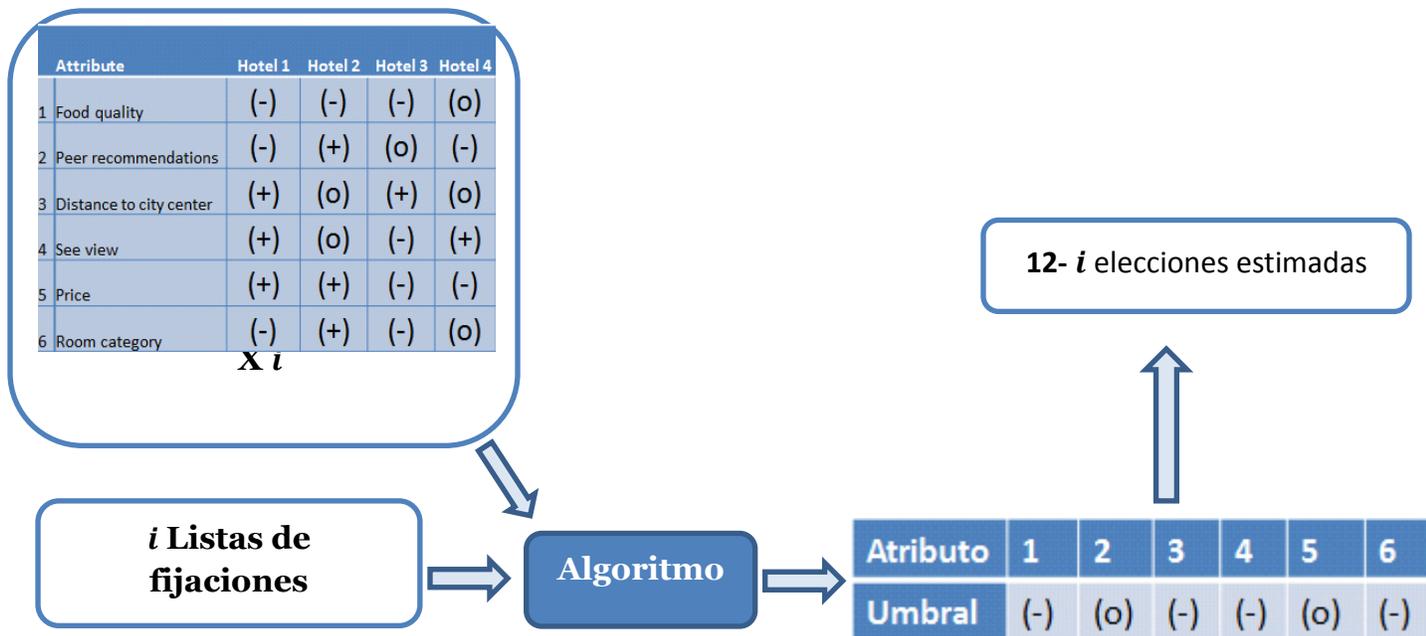


Ilustración 20: Input y Output del algoritmo con elecciones

De la misma manera que en el capítulo anterior, se busca determinar el tipo de regla conjuntiva usada por cada participante, es decir el umbral de cada atributo. El resultado del algoritmo corresponde entonces a un vector de 6 valores, que corresponden al nivel mínimo aceptable para cada atributo.

Se usan como insumos para cada pregunta en $[[1; 10]]$:

- La matriz 6x4 de valores de cada alternativa en cada atributo.
- Las fijaciones hechas por el participante (lista de AOI miradas en orden).

Una vez que se tienen los umbrales estimados de cada participante, se calcula la *hit rate* de esta regla en las 2 preguntas restantes (conjunto de prueba) para ver si los umbrales entregados por el modelo reflejan bien las elecciones del conjunto de prueba. La evaluación de los modelos se hace según los tres criterios presentados en 5.2 (predicción de umbrales, predicción de elecciones y complejidad), con $i = 10$. Cuando se habla de *hit rate* de un modelo, se refiere a la *hit rate* que calcula qué tan bien los umbrales entregados por el modelo reflejan bien las elecciones del conjunto de prueba.

6.2. Metodología

Para construir los modelos con variables de *eye tracking* se seguirá un proceso incremental descrito en la Tabla 10. Empieza con un modelo que simplemente busca cuáles son los dos atributos que tienen umbral (es decir un límite en un nivel mayor que 1) para terminar con un modelo que infiere la regla completa, sabiendo que el participante usó una regla conjuntiva. Cabe recordar que el experimento fue diseñado de tal manera que, en todas las preguntas usadas en el experimento, dos atributos bastaban para tomar la decisión. Esta metodología incremental permite construir un modelo paso por paso e identificar el desempeño de cada paso por separado.

Paso	Sabiendo que	Se busca predecir
1	<ul style="list-style-type: none"> • La regla es conjuntiva • 2 atributos tienen umbrales • El límite está en el nivel medio 	<ul style="list-style-type: none"> • los 2 atributos con umbral
2	<ul style="list-style-type: none"> • La regla es conjuntiva • 2 atributos tienen umbrales 	<ul style="list-style-type: none"> • los 2 atributos con umbral • el límite del umbral para cada uno
3	<ul style="list-style-type: none"> • La regla es conjuntiva 	<ul style="list-style-type: none"> • El umbral para cada atributo.

Tabla 10: Proceso incremental de construcción de modelos

En el paso 2 se usara el método del paso 1 para determinar cuáles son los 2 atributos con umbral, y luego, para estos 2 atributos solamente, se aplicarán modelos para determinar en qué nivel (2 o 3) se sitúa este umbral.

6.3. Paso 1: predecir los dos atributos con umbral.

En este paso se hacen tres supuestos: la regla es conjuntiva, 2 atributos tienen umbrales y el límite está en el nivel medio. Se busca predecir cuáles son los dos atributos que tienen este umbral en el nivel medio (los 4 atributos restantes no tienen umbral, es decir que todos los niveles son aceptables).

Con el fin de predecir los dos atributos que tienen umbral a partir de los datos de movimientos oculares, uno puede primero enfocarse en los dos atributos más mirados. Se calculan para cada persona los atributos más mirados en todas las fijaciones de las preguntas del conjunto de entrenamiento. En la Ilustración 21 se grafica el porcentaje de las personas por las cuales el atributo más mirado es uno de los atributos con umbral (primera columna), el segundo atributo más mirado es con umbral (segunda columna), el tercer atributo más mirado es con umbral (tercera columna), etc, considerando las 12 preguntas.

Se ve que el atributo más mirado de una misma persona (respectivamente el segundo más mirado) es, para el 97% de las personas (respectivamente el 87%), uno de los dos atributos con umbral. Se puede concluir que los dos atributos más mirados de una persona dan una buena indicación para inferir los atributos con umbral.

¿Son los dos atributos más mirados por persona los que tienen umbral?

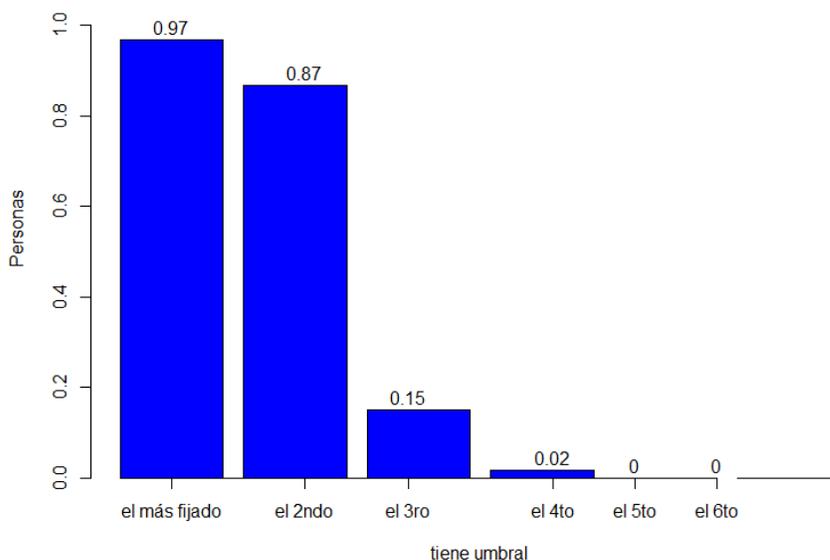


Ilustración 21: ¿Tienen umbrales los atributos más mirados?

Para ver si los dos atributos más mirados se separan bien del resto en términos de fijaciones, se grafica en la Ilustración 22 el porcentaje de fijaciones en el atributo más mirado, en el segundo más mirado, etc. El primero abarca el 50% de las fijaciones y se demarca bien del resto, el segundo con una separación menos radical pero todavía clara.

Se calcula también que en promedio el 30% de las fijaciones se hacen en AOIs irrelevantes, es decir AOIs que no pertenecen a los dos atributos con umbral. Un regresión lineal indica que esta cantidad baja con el número de preguntas (coeficiente de -0.5 fijaciones por pregunta adicional, significancia al 5%) lo que denota que los participantes al inicio miran todos los atributos pero avanzando en las preguntas se concentran en los atributos importantes. Este resultado puede servir en el mundo comercial para determinar si una persona está eligiendo por primera vez o si ya tiene experiencia en una elección dada.

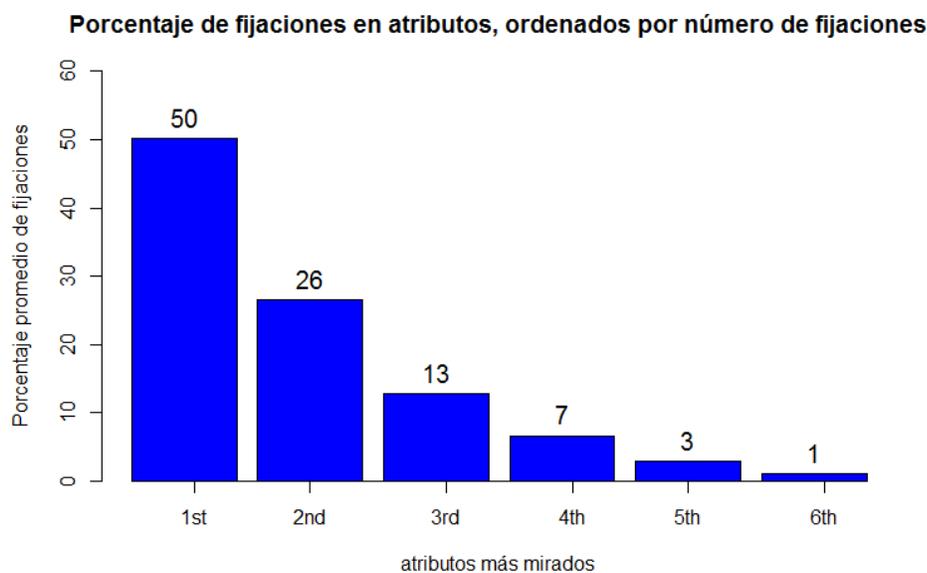


Ilustración 22: Porcentaje de fijaciones en los atributos más mirados

En suma, los participantes concentran en promedio el 76% de sus fijaciones en dos atributos, los cuales corresponden con probabilidad 0.85 a la pareja de los dos atributos con umbral. A medida que responden más preguntas, concentran sus fijaciones en los atributos con umbrales.

6.4. Paso 2: predecir los 2 atributos con umbral y donde se sitúa el umbral.

En esta parte se hacen dos supuestos: la regla es conjuntiva y 2 atributos tienen umbrales. Se busca predecir cuáles son los dos atributos que tienen umbral (usando el paso 1) y el nivel en el cual se sitúa este límite (puede ser los niveles 2 o 3). Cabe recordar que el experimento fue diseñado con umbrales solamente en el nivel medio, lo que no

permite estudiar el movimiento ocular de las personas cuando el límite está en el nivel alto (3) y restringe el análisis.

Con objeto de descubrir dónde se sitúa el umbral se hace la hipótesis que para un atributo dado, los participantes demuestran más interés por los niveles que son aceptables, es decir que sobrepasan el umbral. Según esta hipótesis, la probabilidad $p_{p,a,l}$ del participante p de mirar el nivel l del atributo a depende de si el nivel l pasa el umbral del atributo a o no. Se espera que el aspecto del gráfico de los $p_{p,a,1}$, $p_{p,a,2}$ y $p_{p,a,3}$ sean distintos si el umbral del atributo a está en 1, 2 o 3, como se puede ver en la Ilustración 23:

- Si el umbral se sitúa en el primer nivel (figura a la izquierda), se miran los tres niveles igualmente.
- Si el umbral se sitúa en el nivel medio (figura al medio), se miran más los niveles 2 y 3.
- Si el umbral se sitúa en el nivel alto (figura a la derecha), se mira más el nivel 3.



Ilustración 23: Aspectos esperados de las probabilidades de mirar un nivel según el umbral

A fin de comprobar si esta hipótesis coincide con la realidad, se aproxima la probabilidad $p_{p,a,l}$ del participante p de mirar el nivel l del atributo a al ratio $nFixPerAppear_{p,a,l}$ del número de fijaciones en este nivel dividido por el número de apariciones de este nivel. Es importante corregir por el número de apariciones para no castigar los niveles que, por diseño, aparecen poco y entonces abarcan pocas fijaciones.

En la Ilustración 24 se grafican los boxplots de $nFixPerAppear_{p,a,1}$, $nFixPerAppear_{p,a,2}$ y $nFixPerAppear_{p,a,3}$ para los atributos con umbral en el nivel 2 (tres boxplot de la izquierda) y para los atributos sin umbral (tres boxplot de la derecha). Afortunadamente no se puede graficar el caso de los atributos que tienen umbral en el nivel 3 porque no se probó esta configuración en la experimentación. Se nota bien la diferencia entre los atributos con umbral, por los cuales los niveles 2 y 3 son más fijados, y los atributos sin umbral, por los cuales cada nivel recibe una atención similar y baja. Este gráfico indica que, a nivel global, se confirma la hipótesis que para un atributo

dado, los participantes demuestran más interés por los niveles que son aceptables. El siguiente paso es probar esta hipótesis a nivel individual y usar esta diferencia de interés para determinar donde se sitúa el umbral.

Boxplot número de fijaciones por aparición de los niveles de los atributos con o sin umbral fijaciones de las 12 preguntas incluidas

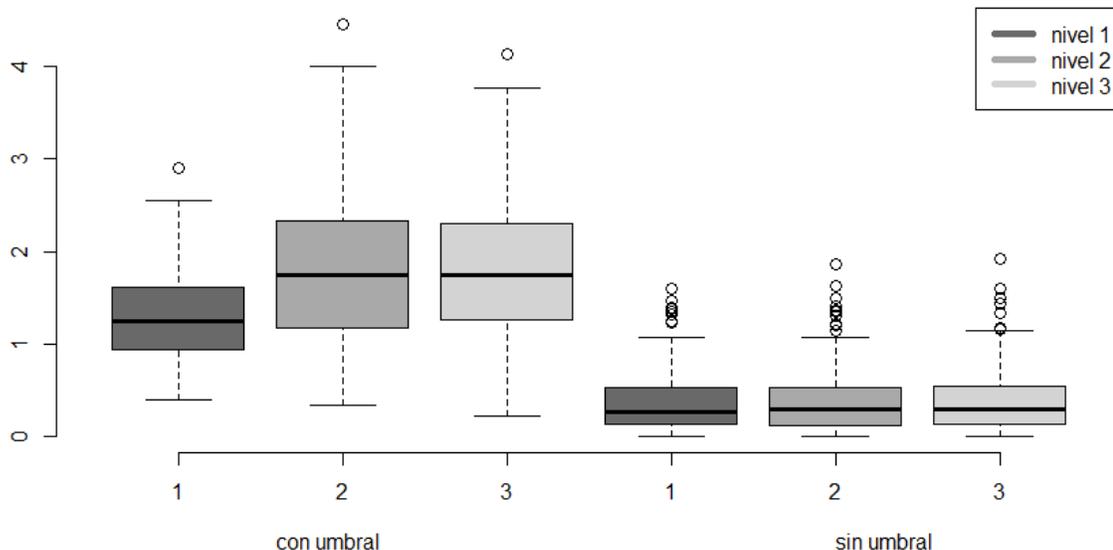


Ilustración 24: Número de fijaciones por aparición en los niveles de los atributos con o sin umbral

En base a este comportamiento observado, se crean 4 modelos:

- Modelos 1 y 2 que usan una estimación de verosimilitud
- Modelo 3 que elige el nivel que abarca más fijaciones por aparición.
- Modelo 4 que prueba la hipótesis que las medias de las fijaciones por aparición de los niveles 2 y 3 son iguales.

6.4.1. Modelos 1 y 2: usan una estimación de verosimilitud

Estos modelos buscan calcular para cada atributo susceptible de tener umbral tres verosimilitudes: una verosimilitud L_1 de tener el umbral en el nivel 1, una verosimilitud L_2 de tener el umbral en el nivel 2 y una verosimilitud L_3 de tener el umbral en el nivel 3. Se compararan las verosimilitudes para encontrar el nivel límite o umbral. Estas tres verosimilitudes tienen como parámetros las tres probabilidades $p_{p,a,1}$, $p_{p,a,2}$ y $p_{p,a,3}$ que corresponden al interés del participante por cada nivel. Las tres $p_{p,a,1}$, $p_{p,a,2}$ y $p_{p,a,3}$ están en el rango $[0; 1]$ y suman 1 pero son desconocidas. Se estiman maximizando las verosimilitudes.

Se definen y se calculan:

$f_{p,q,a,l}$ = Número de fijaciones de la persona p en el nivel l del atributo a de la pregunta q , obtenido a partir de la lista de fijaciones de la pregunta.

$nap_{p,q,a,l}$ = Número de apariciones del nivel l del atributo a de la pregunta q de la persona p , obtenido a partir de la matriz de los niveles de cada alternativa en la pregunta.

Entonces, corrigiendo por el número de apariciones de los niveles, la probabilidad por un participante p de fijar sus ojos $f_{p,q,a,l}$ veces en nivel l del atributo a en la pregunta q es:

$$\left(\frac{nap_{p,q,a,l} * p_{p,a,l}}{\sum_{m=1}^3 nap_{p,q,a,m} * p_{p,a,m}} \right)^{f_{p,q,a,l}}$$

Haciendo el producto sobre las preguntas y los niveles, tenemos que la verosimilitud de las fijaciones hechas por el participante p en el atributo a es:

$$\mathcal{L}(p_{p,a,1}; p_{p,a,2}; p_{p,a,3}) = \prod_{\substack{t \in \llbracket 1; 12 \rrbracket \\ l \in \llbracket 1; 3 \rrbracket}} \left(\frac{nap_{p,q,a,l} * p_{p,a,l}}{\sum_{m=1}^3 nap_{p,q,a,m} * p_{p,a,m}} \right)^{f_{p,q,a,l}}$$

Para probar donde se sitúa el umbral se hacen hipótesis sobre los parámetros $p_{p,a,1}$, $p_{p,a,2}$ y $p_{p,a,3}$. Se calculan para cada atributo que podría tener umbral:

- 1) la verosimilitud L_1 de tener el umbral en el nivel 1: supone que las tres probabilidades $p_{p,a,1}$; $p_{p,a,2}$; $p_{p,a,3}$ son iguales y entonces que cada una vale $1/3$.

$$\mathcal{L}_1 = \mathcal{L}\left(\frac{1}{3}; \frac{1}{3}; \frac{1}{3}\right)$$

- 2) la verosimilitud L_2 de tener el umbral en el nivel 2: supone que las probabilidades $p_{p,a,2}$ y $p_{p,a,3}$ son iguales.

$$\mathcal{L}_2 = \max_{\substack{p_{p,a,2}=p_{p,a,3} \\ p_{p,a,1}=1-2p_{p,a,2} \\ p_{p,a,l} \in [0;1]}} \mathcal{L}(p_{p,a,1}; p_{p,a,2}; p_{p,a,3})$$

- 3) la verosimilitud L_3 de tener el umbral en el nivel 3: supone que las probabilidades $p_{p,a,1}$ y $p_{p,a,2}$ son iguales.

$$\mathcal{L}_3 = \max_{\substack{p_{p,a,1}=p_{p,a,2} \\ p_{p,a,3}=1-2p_{p,a,2} \\ p_{p,a,l} \in [0;1]}} \mathcal{L}(p_{p,a,1}; p_{p,a,2}; p_{p,a,3})$$

Cabe destacar que el modelo con verosimilitud L_1 está anidado del modelo con verosimilitud L_2 (respectivamente L_3) imponiendo la restricción $p_{p,a,1} = p_{p,a,2}$ (respectivamente $p_{p,a,2} = p_{p,a,3}$). Por lo tanto, se puede aplicar un test de ratio de verosimilitud para determinar si el modelo con verosimilitud L_2 (respectivamente L_3) ajusta mejor a los datos que el modelo con verosimilitud L_1 .

Un primer algoritmo usa entonces 2 pruebas de ratio de verosimilitud (con un nivel de significancia de 5%) y sigue los siguientes pasos:

- si (L_2 ajusta mejor que L_1) y (L_3 no ajusta mejor que L_1) entonces el umbral está en 2.
- si (L_3 ajusta mejor que L_1) y (L_2 no ajusta mejor que L_1) entonces el umbral está en 3.
- si (L_2 no ajusta mejor que L_1) y (L_3 no ajusta mejor que L_1) entonces el umbral está en 1.
- si (L_2 ajusta mejor que L_1) y (L_3 ajusta mejor que L_1) entonces se comparan L_2 y L_3 para determinar el umbral.

Este algoritmo se aplica a los dos atributos más fijados, que según el paso 1 son los que tienen umbral. Su principal desventaja es que puede retornar un umbral igual a 1, lo que contradice el supuesto hecho que este atributo tiene umbral mayor a 1. Se compila un segundo algoritmo para evitar esta desventaja y siempre poner el límite en el nivel 2 o 3. Esta variante no usa test de ratio de verosimilitud, solamente compara los valores de L_2 y L_3 :

- si $L_2 < L_3$ entonces el umbral está en 3.
- Si no el umbral está en 2.

Los resultados en predicción de umbrales de estas dos variantes con 10 preguntas de entrenamiento se presentan en la Tabla 11. La variante 1 encuentra los 2 umbrales correctos para solamente 5 personas sobre 60, con una tasa de acierto total de 75%. La variante 2 mejora este resultado, encontrando la regla correcta para 20 personas sobre 60 con una tasa de acierto total de 84%.

Los resultados en términos de predicción de elecciones se grafican en la Ilustración 25. La variante 1 tiene una *hit rate* de solamente 43% con 10 preguntas de entrenamiento, lo que se considera como bajo. Para investigar donde se equivoca el algoritmo, se calcula el número de atributos declarados “con umbral” por los cuales el algoritmo retorna un umbral superior a 1: 47, lo que representa solamente el 37% de los 120 atributos. Además, sobre estos 120 atributos que, según el paso 1, tienen umbral, solamente 31 (26%) son atribuidos a un umbral de 2. La variante 2 no tiene este problema pero tampoco identifica todos los umbrales correctos. Logra una *hit rate* de 63% con 10 preguntas de entrenamiento, lo que es bajo considerando los supuestos hechos. Se

concluye que las diferencias observadas en el comportamiento visual al nivel general en la Ilustración 24 no logran ser capturadas por este modelamiento.

6.4.2. Modelo 3: elige el nivel que abarca más fijaciones per aparición.

Se vio anteriormente que los participantes ponen más interés en los niveles que superan el umbral de un atributo. Ahora se enuncia además la hipótesis que las personas ponen la máxima atención en el nivel donde se sitúa el umbral. Entonces se construye un algoritmo que, para los dos atributos detectados en el paso 1, atribuye el umbral al nivel que tiene el mayor número de fijaciones por aparición (considerando todas las preguntas de conjunto de entrenamiento).

Los resultados de este modelo en términos de predicción de umbrales se pueden ver en la Tabla 11, y en términos de predicción de elecciones en la Ilustración 25. El modelo obtiene una tasa de acierto de 76% en los umbrales, lo que no es satisfactorio. Esta tasa baja se refleja en la *hit rate* de la predicción de elecciones en el conjunto de prueba que llega a 53% con 10 preguntas de entrenamiento, lo que está por debajo del modelo 2. En consecuencia, no se guardará este modelo.

Un efecto extraño llama la atención en la curva de *hit rate* del modelo en la Ilustración 25. Sorprendentemente, la *hit rate* baja cuando el conjunto de entrenamiento aumenta entre 3 y 9 preguntas. Este hecho contra-intuitivo se debe probablemente a la mayor desventaja de este modelo: es muy sensible a pequeñas variaciones del número de fijaciones por aparición de los distintos niveles. La Ilustración 24 indica que al nivel general, el número de fijaciones por aparición de los niveles 2 y 3 son muy similares cuando el atributo tiene un umbral en el nivel 2. En este caso, el aumento de una o dos fijaciones puede cambiar el resultado. Además no considera el hecho que las personas hacen menos fijaciones a medida que van respondiendo más preguntas. Para evitar estos problemas, se construye un cuarto modelo que compara estadísticamente las medias de los números de fijaciones por aparición de los niveles 2 y 3.

6.4.3. Modelo 4: prueba la hipótesis que las medias de los números de fijaciones por aparición de los niveles 2 y 3 son iguales.

Se vio en la Ilustración 24 que a nivel global, las fijaciones por aparición del nivel 2, $nFixPerAppear_{p,a,2}$ igualan las del nivel 3 $nFixPerAppear_{p,a,3}$ cuando el atributo tiene umbral en el nivel 2. La hipótesis que los participantes demuestran más interés por los niveles que son aceptables indica que en el caso de un umbral en el nivel 3, $nFixPerAppear_{p,a,3}$ sería mayor a $nFixPerAppear_{p,a,2}$. Este modelo aprovecha esta diferencia para determinar si el umbral se sitúa en 2 o en 3.

Para cada persona, compara con un test estadístico la hipótesis siguiente para los atributos con umbral:

H₀: Las medias de los números de fijaciones por aparición de los niveles 2 y 3 son iguales.

Versus

H₁: la media del número de fijaciones por aparición del nivel 3 es mayor que aquella del nivel 2.

Si no se rechaza H₀, las medias son iguales, lo que significa que la persona pone tanta atención en el nivel 2 como en el nivel 3. Se concluirá que el umbral se sitúa en el nivel 2. Si se rechaza H₀, la persona pone más atención en el nivel 3, por lo tanto se pondrá el umbral en el nivel 3.

Se busca comparar la media de dos poblaciones de tamaño distinto y de varianza desconocida, por lo que se usa el test de Welch [11].

Se define $nap_{p,a,l}$, el número de apariciones del nivel l del atributo a en el conjunto de preguntas de entrenamiento de la persona p , es decir el número de AOI en que apareció el nivel l . A continuación se notará nap_l dado que la persona p y el atributo susceptible de tener umbral a no varían en el test. Para cada aparición se calcula el número de fijaciones en la AOI correspondiente.

Por consiguiente, se obtiene un vector $Y_l = (f_1^l, \dots, f_{nap_l}^l)$ donde f_i^l es el número de fijaciones en la AOI de la $i^{ésima}$ aparición del nivel l .

Entonces se tiene para el nivel 2 y el nivel 3 a dos vectores: $Y_2 = (f_1^2, \dots, f_{nap_2}^2)$ y $Y_3 = (f_1^3, \dots, f_{nap_3}^3)$.

El test de Welch prueba la hipótesis siguiente con un nivel de significancia de 5%:

$$H_0: \hat{f}_2 = \hat{f}_3 \quad vs \quad H_1: \hat{f}_2 < \hat{f}_3$$

$$\hat{f}_l = \frac{\sum_{i=1}^{nap_l} f_i^l}{nap_l}$$

En pequeños conjuntos de entrenamiento (entre 1 y 6 preguntas), puede ser que las observaciones en $Y_2 = (f_1^2, \dots, f_{nap_2}^2)$ y/o en $Y_3 = (f_1^3, \dots, f_{nap_3}^3)$ no son suficientes para conducir el test. En estos casos donde no se tiene bastante información para determinar dónde se sitúa el umbral en un atributo por el cual se sabe que tiene umbral mayor a 1, se retorna un umbral de 2 para guardar las preferencias más generales (es decir menos

exigentes) de la persona. Eso explica los buenos resultados del modelo 4 en conjuntos de entrenamiento pequeños (ver Ilustración 25: Predicción de elecciones con movimientos de ojos: paso 2). Cabe señalar que estos buenos resultados son consecuencia del diseño del experimento (que no tiene atributos con umbral en el nivel 3), y no reflejan un buen desempeño del modelo que se pueda traspasar a otro experimento.

En términos de predicción de umbrales (ver Tabla 11), el modelo 4 arroja muy buenos resultados con 10 preguntas de entrenamiento: alcanza una tasa de acierto de 93% y encuentra la regla correcta para 46 personas sobre 60.

En términos de predicción de elecciones (ver Ilustración 25), con 10 preguntas de entrenamiento obtiene una hit rate de 86% en el conjunto de prueba, y supera los demás modelos de este paso 2.

6.4.4. Resultados de los modelos del paso 2

Se presentan los resultados en términos de predicción de umbrales (ver Tabla 11) y de elecciones futuras (ver Ilustración 25).

De la misma manera que en 5.4.1 para los algoritmos *Greedy* y de enumeración, se compara en la Tabla 11, para cada persona y para cada atributo, el umbral predicho con un conjunto de entrenamiento de 10 preguntas con el umbral dado en las instrucciones. Se cuenta para cuántos atributos (de 0 a 6) coincidan los dos valores.

Número de atributos que tienen un umbral predicho distinto al umbral verdadero	0	1	2	3	4	5	6	Tasa de acierto
Modelo 1 tests de ratio de verosimilitud	5	21	34	0	0	0	0	75%
Modelo 2 comparación L_2 y L_3	20	23	14	2	0	0	0	84%
Modelo 3 mayor número de fijaciones por aparición	4	29	25	2	0	0	0	76%
Modelo 4 test de comparación de medias	46	5	7	1	1	0	0	93%

Tabla 11: Predicción de umbrales con movimientos de ojos: paso 2 (10 preguntas de entrenamiento)

En conclusión, el mejor modelo del paso 2 es el modelo 4 que compara las medias de los números de fijaciones por aparición de los niveles 2 y 3.

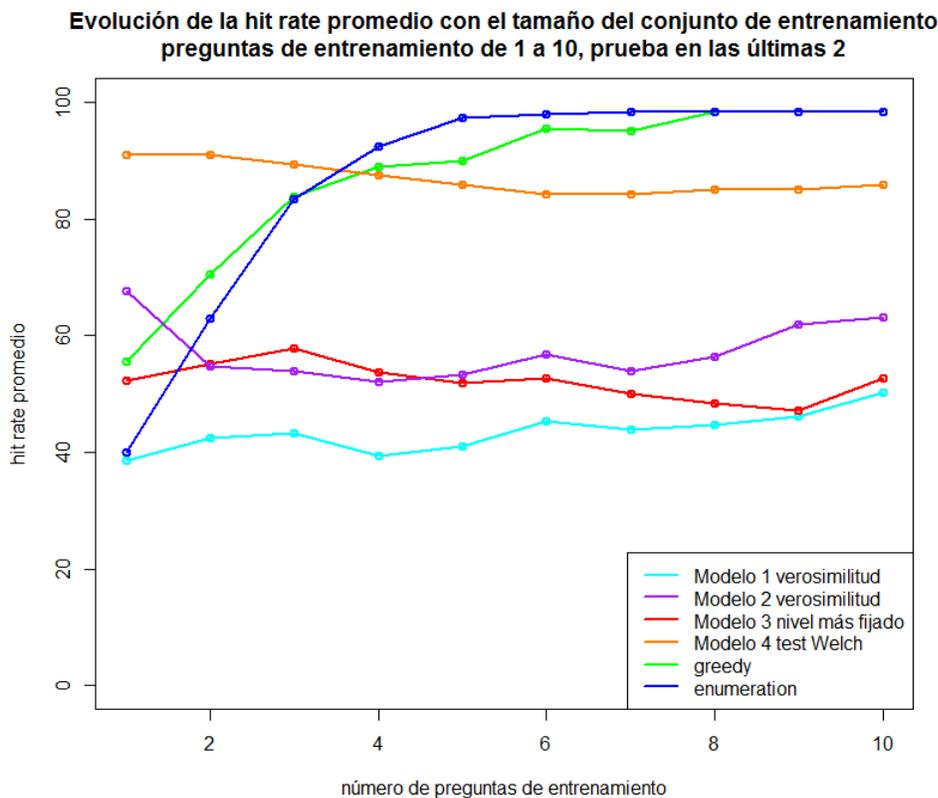


Ilustración 25: Predicción de elecciones con movimientos de ojos: paso 2

6.5. Paso3: Predecir el umbral para cada atributo.

En esta parte se hace el único supuesto que la regla es conjuntiva. Se busca predecir cuáles atributos tienen umbral (límite superior a 1) y el nivel en el cual se sitúa este límite. Se prueban dos enfoques distintos.

6.5.1. Modelo 5: tests de Welch para comparar las medias de los números de fijaciones por aparición de los niveles 1,2 y 3.

El primer enfoque se basa en la hipótesis que para un atributo dado, los participantes demuestran más interés por los niveles que son aceptables, es decir que sobrepasan el umbral. De esta hipótesis resultan los distintos aspectos de las probabilidades de mirar un nivel según el umbral como se vio en la Ilustración 23. En el paso 2 se usó un test de Welch para comparar \hat{f}_2 y \hat{f}_3 sabiendo que el umbral era mayor a 1. En este paso se generaliza el test de Welch para comparar \hat{f}_1 , \hat{f}_2 y \hat{f}_3 para cada uno de los 6 atributos y determinar así la atención puesta en cada nivel.

El test de Welch no puede comparar más de 2 poblaciones, por lo tanto se usan 3 tests de Welch:

$$W_{12}: H_0: \hat{f}_1 = \hat{f}_2 \quad vs \quad H_1: \hat{f}_1 < \hat{f}_2$$

$$W_{13}: H_0: \hat{f}_1 = \hat{f}_3 \quad vs \quad H_1: \hat{f}_1 < \hat{f}_3$$

$$W_{23}: H_0: \hat{f}_2 = \hat{f}_3 \quad vs \quad H_1: \hat{f}_2 < \hat{f}_3$$

Y el algoritmo hace los siguientes pasos:

Si $\hat{f}_1 = \hat{f}_2$ (W_{12})

entonces si $\hat{f}_2 = \hat{f}_3$ (W_{23}) entonces $\hat{f}_1 = \hat{f}_2 = \hat{f}_3$, umbral =1

si no $\hat{f}_2 < \hat{f}_3$ (W_{23}) entonces $\hat{f}_1 = \hat{f}_2 < \hat{f}_3$, umbral =3

si no: $\hat{f}_1 < \hat{f}_2$ (W_{12})

entonces si $\hat{f}_2 = \hat{f}_3$ (W_{23}) entonces $\hat{f}_1 < \hat{f}_2 = \hat{f}_3$, umbral =2

si no $\hat{f}_2 < \hat{f}_3$ (W_{23}) entonces $\hat{f}_1 < \hat{f}_2 < \hat{f}_3$, umbral =3

Se prueba el algoritmo con distintos niveles de significancia: $\alpha = 10\%$, $\alpha = 5\%$ y $\alpha = 1\%$. Este cambio no afecta mucho los resultados, por lo que se mostrará solamente la variante con $\alpha = 5\%$.

Los resultados de predicción de umbrales se encuentran en la Tabla 12 y son medios: el modelo encuentra 66% de los umbrales. Pero en términos de predicción de elecciones son poco satisfactorios (ver Ilustración 30): con 10 preguntas de entrenamiento tiene una *hit rate* de 32%, justo mejor que una selección al azar que tendrá una *hit rate* de 25%.

Se concluye que el test de Welch para comparar las medias no se logra extender cuando no se sabe si el atributo tiene umbral o no.

6.5.2. Modelo 6: Determinar el número de atributos con umbral para aplicar el modelo 4

El modelo 4 con test de Welch tiene buenos resultados cuando se sabe el número de atributos con umbral pero no se logra generalizar. Entonces con el modelo 6 se cambia el enfoque: se hará un paso inicial para determinar cuáles atributos tienen umbral y luego aplicar el modelo 4 en ellos para determinar donde se sitúa el límite.

Variante 1: p-values de la regresión lineal $nFixPerAppear_{p,a,l} \sim l$

Para determinar si un atributo dado tiene umbral o no, primero se vuelve a la hipótesis inicial que dice que los participantes demuestran más interés por los niveles que son aceptables y a los patrones vistos en la Ilustración 24: Número de fijaciones por aparición en los niveles de los atributos con o sin umbral. Se destaca que si el atributo no tiene umbral, el número de fijaciones por aparición de este no varía con el nivel. En cambio, si el atributo tiene umbral, el número de fijaciones por aparición será distinto para los niveles aceptables y no aceptables. Se usa esta observación en un atributo dado para determinar si tiene umbral o no.

Para un atributo y una persona dados, se usa una regresión lineal para determinar si el número de fijaciones por aparición $nFixPerAppear_{p,a,l}$ depende del nivel l .

La variable dependiente es $Y = (Y_1, Y_2, Y_3) = (f_1^1, \dots, f_{nap_1}^1, f_1^2, \dots, f_{nap_2}^2, f_1^3, \dots, f_{nap_3}^3)$ que contiene los números de fijaciones de las apariciones del nivel 1, después aquellos del nivel 2 y al final aquellos del nivel 3. En total se tienen $nap_{p,a,1} + nap_{p,a,2} + nap_{p,a,3}$ observaciones, lo que corresponde al número de apariciones del atributo en el conjunto de entrenamiento, es decir $4 * \text{número de preguntas}$.

La variable dependiente es un factor con los niveles de cada aparición. $X = (1, \dots, 1, 2, \dots, 2, 3, \dots, 3)$. También tiene $nap_{p,a,1} + nap_{p,a,2} + nap_{p,a,3}$ observaciones. Se construyen tres variables dummy a partir de X, una para cada nivel l , $D_l = \begin{cases} 1 & \text{si } X = l \\ 0 & \text{si no} \end{cases}$.

La ecuación de la regresión lineal es la siguiente: $Y = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3$.

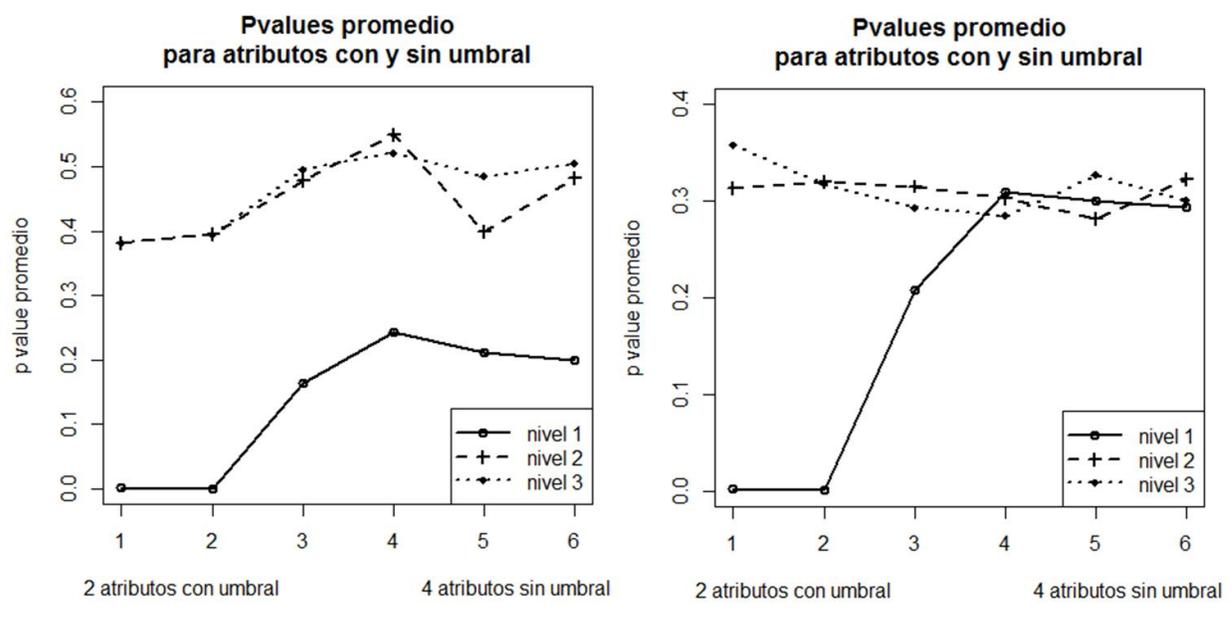


Ilustración 26: pvalues de la regresión lineal $nfixPerAp \sim l$

Para ver si el número de fijaciones por aparición depende del nivel se investiga si las variables D_l son significativas o no. En la Ilustración 26 se grafican los promedios (a la izquierda) y las desviaciones estándares (a la derecha) sobre las 60 personas de los p-values de las variables asociadas a los distintos niveles para los 6 atributos de una persona. Los atributos no están en el orden vertical usual pero separados según si tienen umbral para la persona o no: por ejemplo los 2 primeros valores de la línea roja del gráfico a la izquierda corresponden a los p-values promedios de la variable D_1 los 2 atributos con umbral de una persona. Se deben comparar con los p-values de los 4 atributos sin umbral, que son los 4 valores de la derecha. Aca se nota la diferencia en los p-values del nivel 1: son muy bajas para los atributos con umbral (es decir que la variable D_1 es muy significativa) pero suben para los atributos sin umbral (es decir que D_1 es menos significativa), lo que es esperable según la hipótesis hecha. Esta distinción no se hace con los p-values asociados a los niveles 2 y 3: no se ve una diferencia en los promedios y la desviación estándar es muy alta. Por todo lo anterior, se decide usar la significancia de la variable D_1 para decidir si el atributo tiene umbral o no.

Queda determinar qué nivel de significancia α es lo más apropiado para separar los atributos con y sin umbral. Los atributos con $pvalue(D_1) \leq \alpha$ serán etiquetados “con umbral” (variable D_1 significativa) mientras los atributos con $pvalue(D_1) > \alpha$ serán considerados “sin umbral” (variable D_1 no significativa). El gráfico de la izquierda de la Ilustración 26 da un orden de magnitud de los p-values asociados a D_1 son muy bajas para los atributos con umbral (el gráfico les aproxima a 0, tienen promedio de $7 * 10^{-4}$) y del orden de 0.15 en promedio para los atributos sin umbral. Esto indica que se necesita un α bajo para hacer la diferencia. Se prueba el modelo con distintos α : $\alpha = 5\%$, $\alpha = 1\%$ y $\alpha = 0.5\%$. Los mejores resultados se obtienen con $\alpha = 0.5\%$.

Entonces se determina con la regresión lineal si el atributo tiene umbral o no y se aplica el modelo 4 con test de Welch a los atributos con umbral. Este modelo obtiene mejores resultados que el modelo 5: se puede ver en la Tabla 12 que logra con 10 preguntas de entrenamiento una tasa de acierto de 86% en la predicción de umbrales (versus un 66% para el modelo 5) y una *hit rate* de 72% (versus 32% para el modelo 5) en la predicción de elecciones (ver Ilustración 30).

Variante 2: determinar el número de atributos con umbral $nthr$ y aplicar el modelo 4 en los $nthr$ atributos más mirados.

Un segundo enfoque es determinar el número de atributos con umbral de la persona, sea $nthr$, y después aplicar el modelo 4 en los $nthr$ atributos más mirados. Cabe recordar que en el paso 2 se aplicó el modelo 4 con $nthr = 2$. En otras palabras, busca encontrar los atributos más importantes para una persona (cuántos son y cuáles son) y en un segundo paso cuál es el límite aceptable para estos atributos.

Para determinar $nthr$, ésta variante busca usar información de las fijaciones a un nivel más agregado, es decir las fijaciones en los distintos atributos y no en los distintos

niveles de un atributo. La intuición de esta variante surge de la Ilustración 27 donde se grafica el número de fijaciones en los atributos ordenados del más fijado al menos fijado para el participante 41, que se eligió como ejemplo. Se ve una distinción clara entre los dos atributos más fijados a la izquierda y los 4 restantes, lo que parece indicar que son 2 atributos con umbral ($nthr = 2$). Esta variante usa varias regresiones lineales para realizar de manera automática la separación hecha “al ojo” para la persona 41 y determinar $nthr$.

Se considera que una persona puede tener de 1 a 5 atributos con umbral. Se hacen 5 regresiones lineales que corresponden a las 5 posibilidades para $nthr$. Para $n \in \llbracket 1; 5 \rrbracket$, la regresión lineal RL_n toma como variable dependiente el número ordenado de fijaciones por atributo y como variable explicativa una variable X_n que separa los n atributos con umbral de los $6 - n$ atributos sin umbral: $X_1 = (1,0,0,0,0)$, $X_2 = (1,1,0,0,0)$, $X_3 = (1,1,1,0,0)$, hasta X_5 .

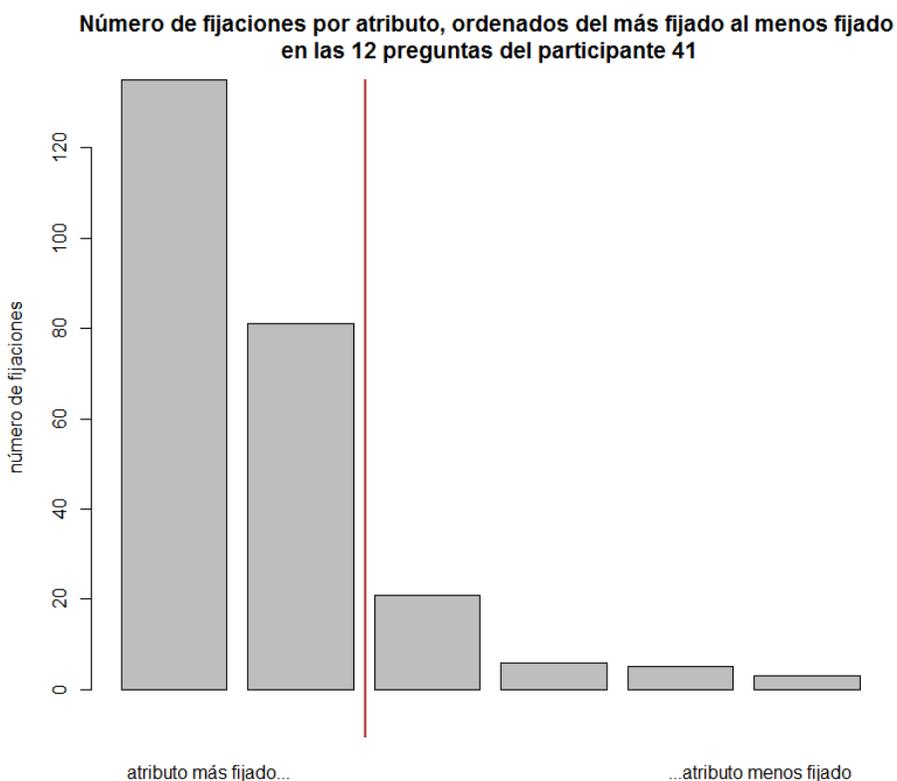


Ilustración 27: Número de fijaciones en los atributos del participante 41

Se calculan los R^2 asociados a cada regresión: R_n^2 mide qué tan bien el hecho de tener n atributos con umbral explica la varianza del número de fijaciones. Se guarda el $nthr$ que corresponde al mayor R^2 :

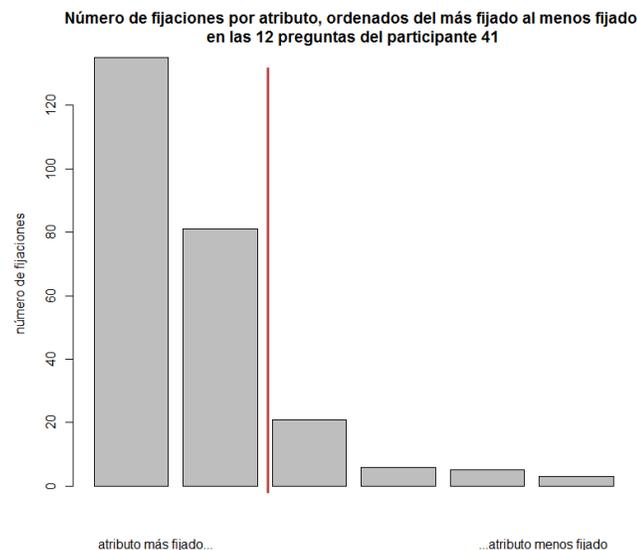
$$nthr = \operatorname{argmax}_{n \in \llbracket 1; 5 \rrbracket} R_n^2$$

En la Ilustración 28 se puede ver el ejemplo del participante 41, por el cual la variable X_2 debería tener un R^2 mayor a los demás R_n^2 .

A continuación se detallan las regresiones lineales. Para cada persona p , cada atributo a y cada pregunta t del conjunto de entrenamiento, se calcula el número de fijaciones $nfix_{p,a,t}$. Se tienen entonces $6 * t$ observaciones de la variable dependiente donde t es el número de preguntas del conjunto de entrenamiento. Para cada pregunta t , y cada persona p , se ordenan los 6 valores $nfix_{p,a,t}$ en orden creciente. Esa operación corresponde a ordenar los atributos del más mirado al menos mirado como se hizo en la Ilustración 27. La variable dependiente Y concatena estos $6 * t$ valores (cada grupo de 6 ordenado).

Como variable dependiente para la regresión n se usa:

- El número de la pregunta (para tomar en cuenta el hecho que las personas hacen más fijaciones en las primeras preguntas)
- la variable X_n que separa los n atributos con umbral de los $6 - n$ atributos sin umbral.



X1	1	0	0	0	0	0
X2	1	1	0	0	0	0
X3	1	1	1	0	0	0
X4	1	1	1	1	0	0
X5	1	1	1	1	1	0

Ilustración 28: Variables X_n de las regresiones lineales

La ecuación de la regresión lineal RL_n es la siguiente:

$$RL_n : \quad nfix_{p,a,t} \sim \beta_0 + \beta_1 X_n(a) + \beta_2 * t$$

Para tener un orden de magnitud de los R^2 se grafican en la Ilustración 29 los boxplot de los R^2 obtenidos para todas las personas. Varían entre 0.1 y 0.8, lo que es aceptable. Dado que los participantes del experimento tienen todos $nthr = 2$, los R^2 de la regresión RL_2 deberían ser mayores al resto. A nivel general se ve la diferencia con los R^2 de las regresiones RL_3 , RL_4 y RL_5 pero no se distingue bien de la regresión RL_1 . Los resultados del modelo van a indicar si esta distinción se hace bien a nivel individual.

La variante 2 del modelo 6 obtiene los mejores resultados dentro de los modelos con movimientos de ojos, tanto en predicción de umbrales (tasa de acierto de 88%, ver Tabla 12) como en predicción de elecciones (*hit rate* de 75% con 10 preguntas de entrenamiento, ver Ilustración 30).

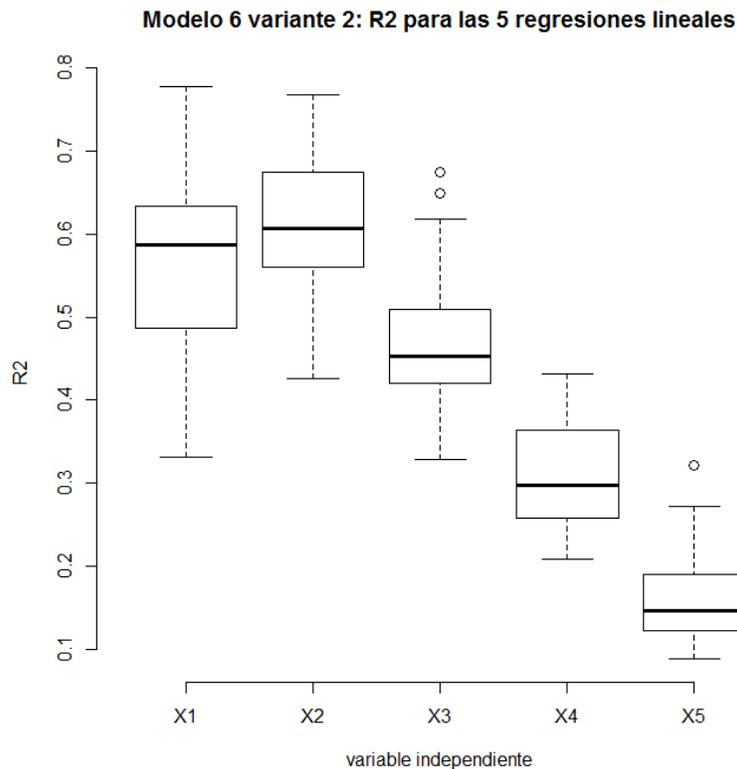


Ilustración 29: Modelo 6.2: R2 de las regresiones lineales

6.5.3. Resultados de los modelos del paso 3

Se presentan los resultados en términos de predicción de umbrales (ver Tabla 12) y de elecciones futuras (ver Ilustración 30). Como se mencionó precedentemente, el mejor modelo con movimientos de ojos es el modelo 6.2 que logra una tasa de acierto de

88% en predicción de umbrales con 10 preguntas de entrenamiento y una *hit rate* de 75% en predicción de elecciones.

En términos de complejidad, todos los modelos con movimientos de ojos se demoraron menos de 10 segundos en calcular los umbrales de 60 personas con 10 preguntas de entrenamiento, lo que se considera insignificante.

Número de atributos que tienen un umbral predicho distinto al umbral verdadero	0	1	2	3	4	5	6	Tasa de acierto
Modelo 5 tests de Welch $\alpha = 5\%$	2	12	29	16	1	0	0	66%
Modelo 6.1 $\alpha = 0.5\%$	29	16	10	5	0	0	0	86%
Modelo 6.2	27	26	6	0	1	0	0	88%

Tabla 12: Predicción de umbrales con movimientos de ojos: paso 3

Evolución de la hit rate promedio con el tamaño del conjunto de entrenamiento preguntas de entrenamiento de 1 a 10, prueba en las últimas 2

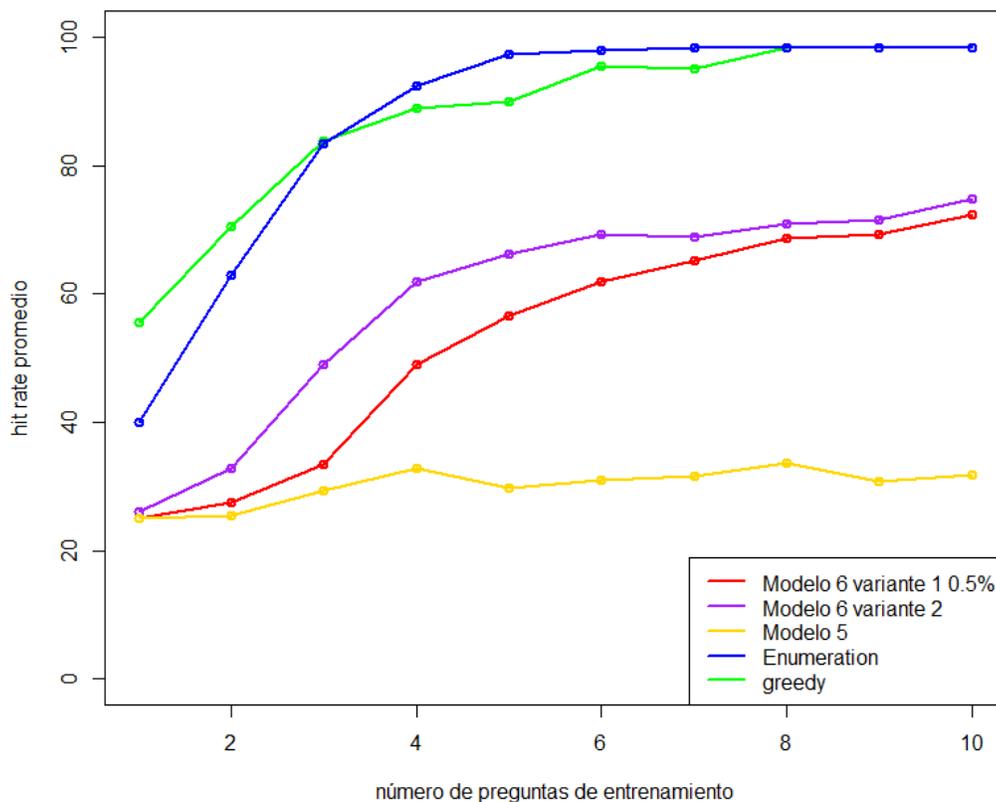


Ilustración 30: Predicción de elecciones con movimientos de ojos: paso 3

6.6. Conclusiones del capítulo

Se construyeron varios modelos de predicción de regla basados en movimientos de ojos, buscando aprovechar a nivel individual de los comportamientos vistos a nivel global. Alcanzan a predecir la regla conjuntiva usada con una tasa de acierto de 88% y elecciones futuras con una *hit rate* de 75% con 10 preguntas de entrenamiento lo que es bastante bueno. Sin embargo, el mejor modelo que se basa en el comportamiento ocular, el modelo 6.2, no logra superar la heurística *Greedy*, en términos de predicción de umbrales como en predicción de elecciones. Sin embargo, cabe destacar que la heurística *Greedy*, a la diferencia de los modelos con movimientos de ojos, utiliza información directamente relacionada con la métrica de comparación (*hit rate*). Se concluye que los movimientos de ojos, al menos en el marco de esta investigación, aportan menos información que las elecciones sobre el tipo de regla conjuntiva usada.

7. APLICACIÓN EN UNA BASE DE DATOS DE ELECCIONES LIBRES Y MODELO COMBINADO

Hasta ahora se ha trabajado con un experimento en el cual se impuso al participante la regla de decisión a usar. Sin embargo, en la vida real las personas eligen con libertad. Pueden seguir otra regla o una mezcla de distintas reglas. Este capítulo busca aplicar los modelos a otra base de datos en la cual los participantes eligieron de manera libre, comparar el desempeño de los modelos con elecciones con aquello de los modelos con movimientos de ojos en este experimento y finalmente combinar los dos para lograr un mayor rendimiento.

7.1. Descripción del experimento de las cafeteras

El experimento de análisis conjunto fue realizado en Alemania por Martin Meißner. Consiste en la elección de comprar una máquina cafetera, en donde el encuestado tiene cuatro alternativas diferentes: tres productos de cafeteras distintas entre sí, y la posibilidad de *No Comprar*. Las alternativas se despliegan horizontalmente y los atributos verticalmente, de la misma manera que en el experimento anterior y como se puede ver en la Ilustración 31. Se pidió a 62 personas responder 14 preguntas, de manera libre usando sus preferencias personales. Cada una de las alternativas fue creada al azar para cada persona y ocasión de respuesta, excepto la pregunta número 8, que se generó igual para todos los encuestados. El conjunto de entrenamiento de los modelos comprenderá las preguntas 1 a 10, dejando las últimas 4 para la prueba.

	Atributo	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Número de niveles	Ranking de los niveles
1	Marca	Braun	Krups	Philips	Severin	4	?
2	Material	Stainless steel	Plastic	Brushed aluminum		3	?
3	Sistema	Pad	Capsule			2	?
4	Diseño	Design A	Design B	Design C	Design D	4	?
5	Precio por taza	€ 0.12	€ 0.22	€ 0.32		3	(3,2,1)
6	Precio	€ 99.99	€ 129.99	€ 159.99	€ 189.99	4	(4,3,2,1)

Tabla 13: Atributos del experimento de cafeteras

Los productos se caracterizan con 6 atributos (*Marca, Material, Sistema, Diseño, Precio por Taza* y *Precio* de la maquina) que son expuestos en la Tabla 13. Se destacan dos

mayores diferencias con el experimento de los hoteles. Primero, todos los atributos no tienen el mismo número de niveles: los atributos *Marca*, *Diseño* y *Precio* tienen 4 niveles, *Material* y *Precio por taza* tienen 3 niveles y el atributo *Sistema* tiene solamente 2 niveles.

If you were in the market to buy a new single-cup coffee brewer and these were your only options, which would you choose?

	A	B	C	D
Brand 				None: I wouldn't choose any of these.
Material 	Plastic 	Brushed aluminium 	Stainless steel 	
System 	Pad 	Capsule 	Pad 	
Design 				
Price of a cup 	32 Cent 	22 Cent 	12 Cent 	
Price 	189. ⁹⁹ 	129. ⁹⁹ 	159. ⁹⁹ 	

Ilustración 31: Diseño del experimento de cafeteras

Además, en el experimento de los hoteles cada atributo tiene un ranking claro y universal de preferencia en los niveles: el nivel 1 entrega la menor utilidad al participante, el nivel 2 una utilidad media y el nivel 3 la mayor utilidad. Por consiguiente, el ranking creciente de preferencia de los niveles es (1,2,3), y eso para todos los participantes y todos los atributos. Para el experimento de las cafeteras, los atributos 5 y 6 (*Precio por taza* y *Precio*) también tienen un ranking genérico. En cambio, para los atributos 1 a 4 (*Marca*, *Material*, *Sistema* y *Diseño*) no existe ranking de preferencias compartido. Por ejemplo las preferencias por el diseño A, B, C o D cambian según la persona. Tampoco es claro que este ranking exista para cada persona, es decir que cada participante pueda ordenar los distintos niveles en orden de utilidad creciente. Eso implica que una regla conjuntiva ya no puede ser caracterizada por un vector de 6 umbrales pero más bien con un conjunto de niveles aceptables para cada atributo (6 conjuntos de niveles aceptables). Por lo tanto, se modifica la función *hit rate* usada en el experimento de los hoteles para calcular la predicción de elecciones: en vez

de tomar como input un vector de 6 umbrales, tomará 6 conjuntos de valores aceptables para poder aplicarse al experimento de las cafeteras.

En este experimento no se impuso la regla a usar, por lo tanto no se puede evaluar la predicción de umbrales de los algoritmos. Como se ve en la Ilustración 32, solamente se van a evaluar la predicción de elecciones en el conjunto de prueba (preguntas 11 a 14) y la complejidad.

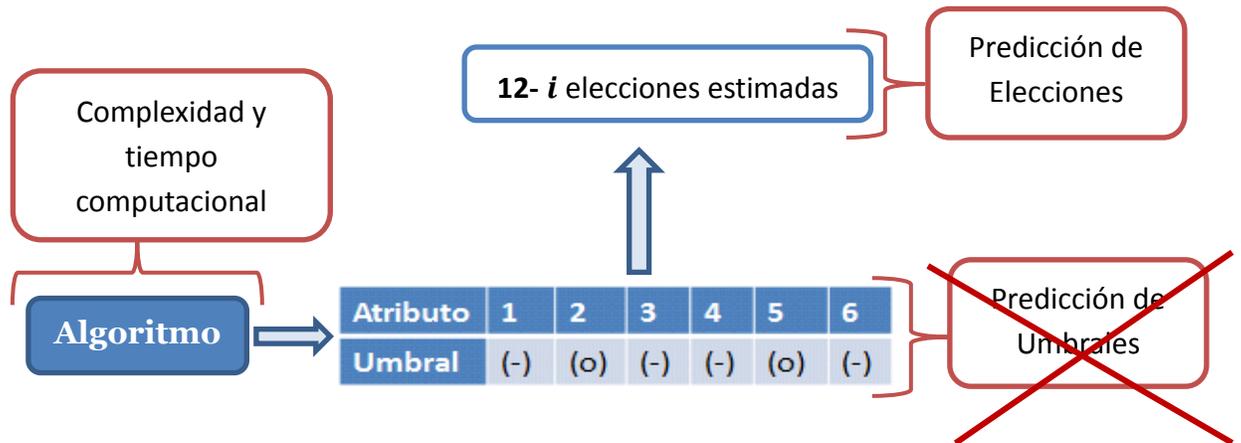


Ilustración 32: Evaluación de los algoritmos en el experimento de las cafeteras

7.2. Análisis descriptivo

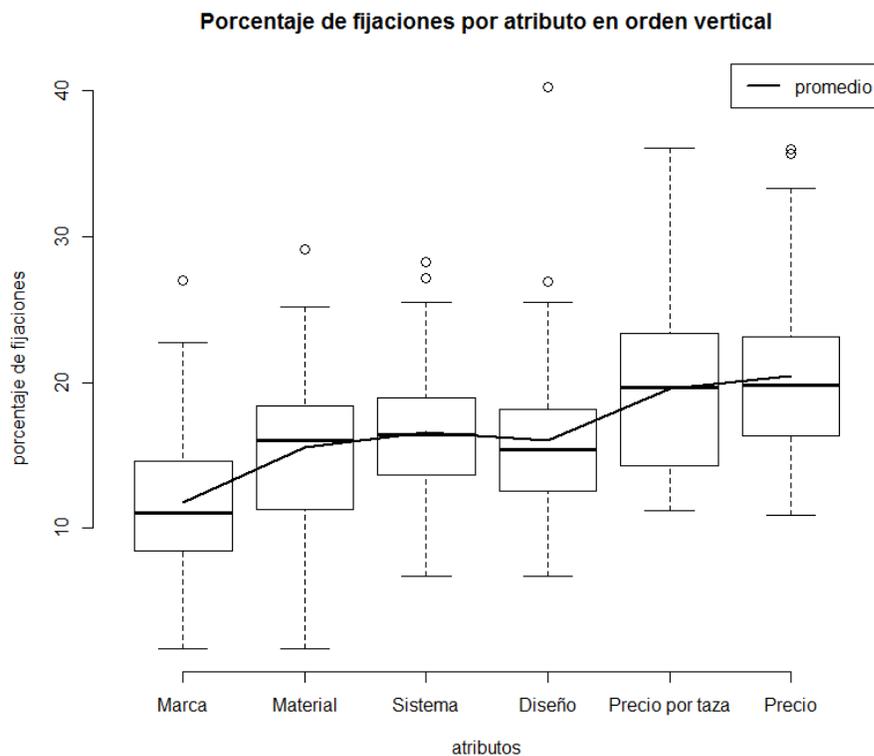


Ilustración 33: Porcentaje de fijaciones por atributo en el experimento de las cafeteras

A continuación se presenta un corto análisis descriptivo de los movimientos de ojos en el experimento de las cafeteras. En la Ilustración 33 se grafican los boxplots y el promedio de los porcentajes de fijaciones de los distintos atributos. Los atributos 5 y 6 (*Precio por taza* y *Precio*) son en promedio los más mirados, lo que indica que tienen mayor importancia a nivel general.

En la Ilustración 34 se grafican los boxplots y el promedio de los porcentajes de fijaciones de los distintos atributos pero esta vez con los atributos ordenados del más fijado al menos fijado para cada persona. Se observa una continuidad del primero al último atributo. Al contrario del experimento de las cafeteras, no se nota una separación entre atributos más importantes de un lado (eran los atributos con umbral en el experimento de los hoteles) y atributos menos importantes (sin umbral), a lo menos a nivel general.

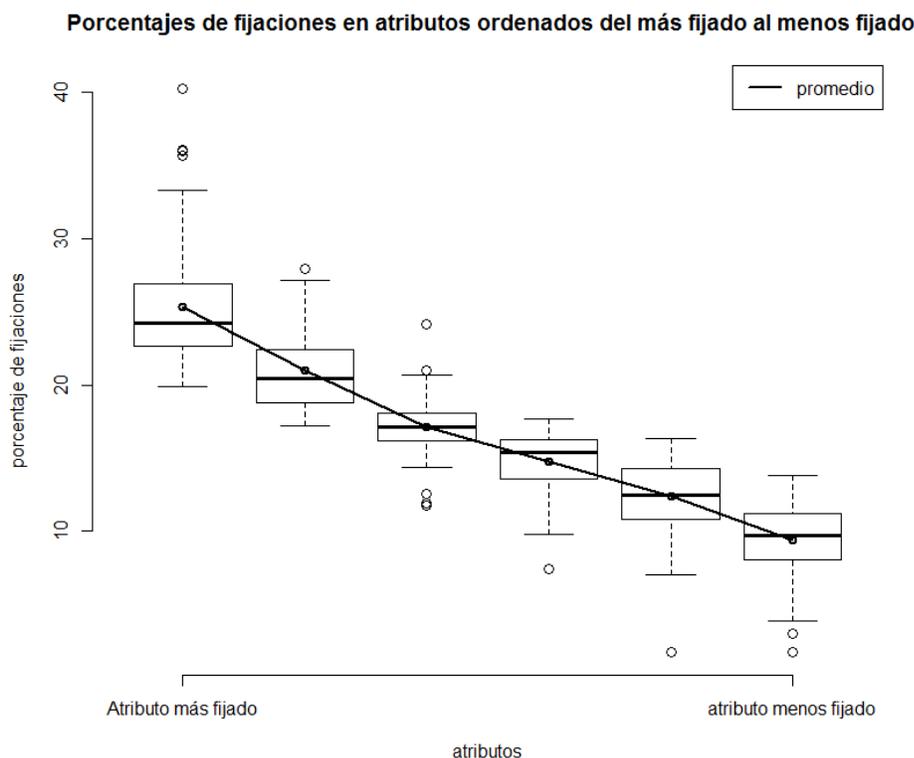


Ilustración 34: Porcentaje de fijaciones por atributo, atributos ordenados, en el experimento de las cafeteras

De otro lado se busca probar si la hipótesis que para un atributo dado, los participantes demuestran más interés por los niveles que son aceptables podría ser aplicable en este experimento y si se podrían encontrar los patrones vistos en la Ilustración 24. Con este fin se grafican en la Ilustración 35 las fijaciones por aparición de los distintos niveles para los dos atributos más fijados (a la izquierda, corresponden a atributos “con umbral” del experimento con hoteles) y para los dos atributos menos

fijados (a la derecha, corresponden a atributos “sin umbral”). Mientras que en el experimento de los hoteles se notaba una diferencia clara entre los patrones de números de fijaciones por aparición si el atributo era con o sin umbral (ver Ilustración 24), no se ven patrones distintos en la Ilustración 35. Se deduce que, a nivel global, en los movimientos de ojos no hay evidencia que los participantes siguieron una regla conjuntiva.

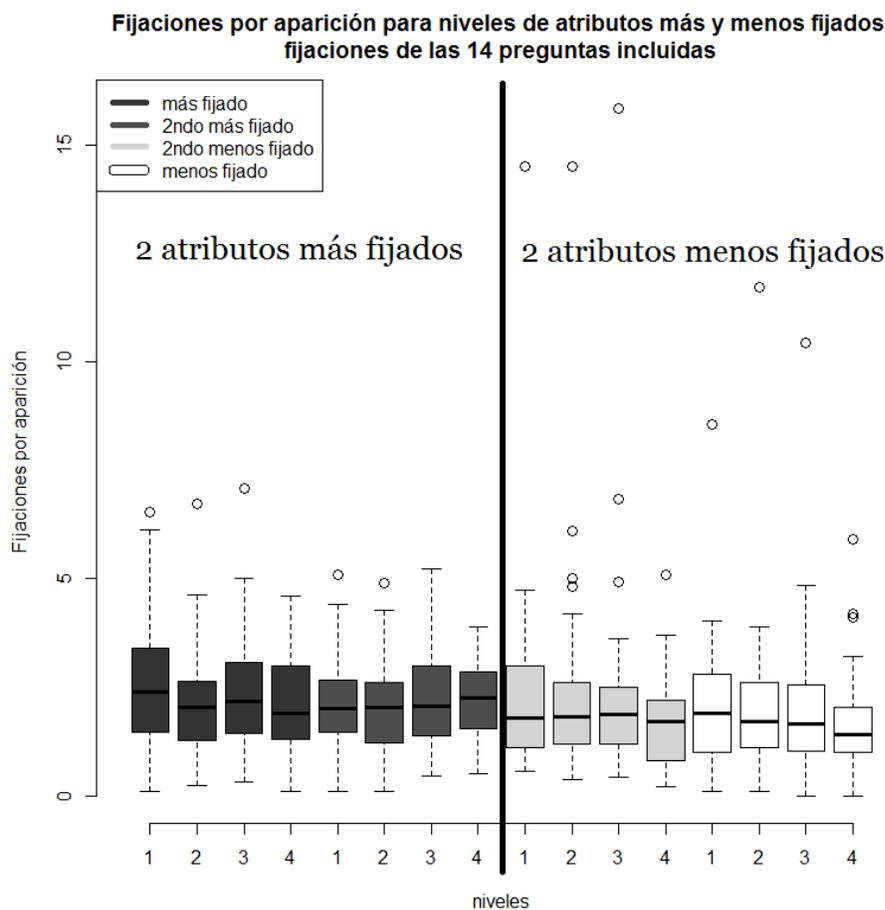


Ilustración 35: Fijaciones por aparición en los atributos más y menos fijados del experimento de las cafeteras

7.3. Algoritmos que usan elecciones

En esta parte se aplican el algoritmo de enumeración y la heurística *Greedy* al experimento de las cafeteras. La inexistencia de un ranking universal de preferencia de los niveles de los atributos 1 a 4 (ver 7.1) obliga a agregar un paso inicial a los algoritmos para calcular estos rankings para cada persona. Se hace entonces el supuesto que estos rankings existen para cada persona, es decir que cada participante puede ordenar los distintos niveles de un atributo en orden de utilidad creciente.

Para una persona p y un atributo $a \in \llbracket 1; 4 \rrbracket$ dado, el ranking de preferencias se construye a partir de las elecciones del conjunto de entrenamiento: se calcula para cada nivel el número de elecciones de alternativas que tenían este nivel. Además se calcula el número de apariciones de este nivel en las preguntas del conjunto de entrenamiento. Los niveles se ordenan según su número de elecciones por aparición $nChoicesPerAppear_{p,a,l}$ (del nivel l del atributo a de la persona p en el conjunto de entrenamiento).

Con estos rankings se aplican los pasos del algoritmo de enumeración y del *Greedy*, que entregan una combinación de umbrales que caracteriza la regla conjuntiva que ajusta mejor a las elecciones del participante. Esta combinación de umbrales se transforma en 6 conjuntos de niveles aceptables, todavía usando los mismos rankings. La *hit rate* del conjunto de prueba se calcula tomando como input estos conjuntos de valores aceptables como se vio en 7.1.

Cabe señalar que, al contrario del experimento de los hoteles, el algoritmo *Greedy* aplicado al experimento de las cafeteras no empieza con los umbrales en (1,1,1,1,1) pero en los niveles que aportan menor utilidad según los rankings calculados con las elecciones.

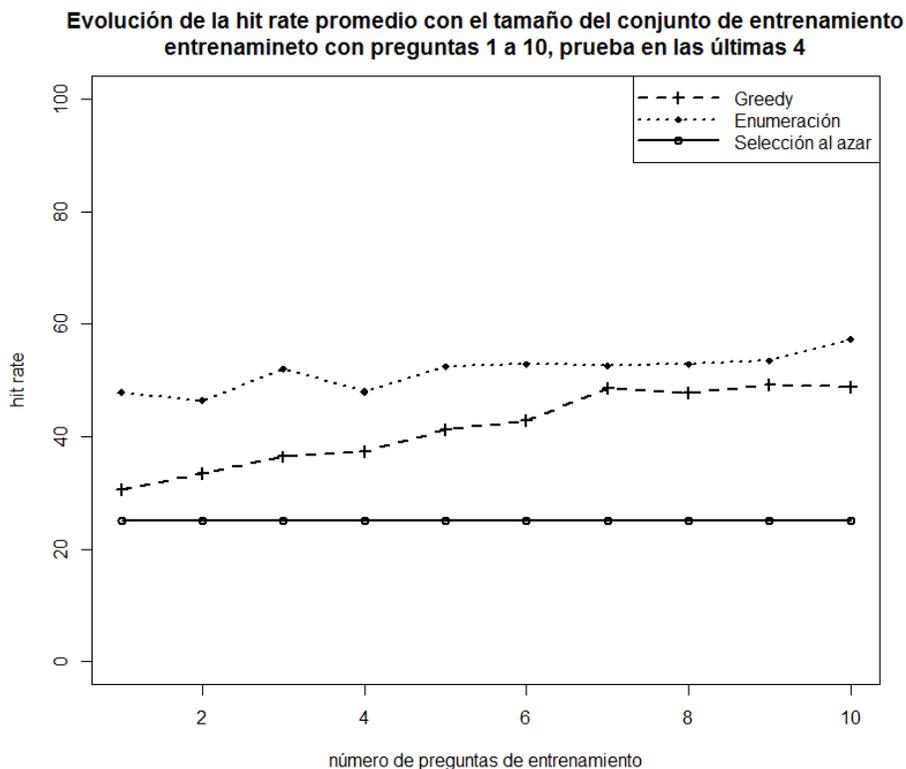


Ilustración 36: Predicción de elecciones con elecciones: experimento de las cafeteras

Los resultados obtenidos se muestran en la Ilustración 36 que grafica la *hit rate* en el conjunto de prueba para conjuntos de entrenamiento que van de 1 a 10 preguntas. La fuerte caída de las *hit rates* de este experimento con respecto al experimento de los hoteles se debe al hecho que los participantes eligen ahora de manera libre sin seguir

necesariamente una regla conjuntiva. Por otra parte, se puede ver que el algoritmo de enumeración supera el *Greedy* pero que esta diferencia se reduce con un mayor conjunto de entrenamiento. La Tabla 14 revela las *hit rates* en los conjuntos de entrenamiento y de prueba para 10 preguntas de entrenamiento. Primero se nota que los dos algoritmos ajustan mucho mejor en el conjunto de entrenamiento que en el conjunto de prueba, lo que se puede interpretar como un sobre-ajuste. Segundo, los dos métodos logran resultados al menos dos veces mejor que una selección al azar, lo que significa que la regla conjuntiva alcanza, en alguna medida, a representar elecciones reales hechas de manera libre.

<i>Hit rate</i>	Conjunto de entrenamiento (10 primeras preguntas)	Conjunto de prueba (4 ultimas preguntas)
<i>Greedy</i> conjuntivo	74%	49%
Enumeración	81%	57%

Tabla 14: *Hit rate* para *Greedy* y Enumeración: experimento de las cafeteras

Para tener un punto de referencia de las *hit rates* obtenidas, se muestran en la Tabla 15 las *hit rates* de dos otros modelos: un modelo que usa la heurística *Greedy* lexicográfica, hecho por Paula Urrutia [9] y de un modelo logit multivariado elaborado por Andrés Musalem. Cabe enfatizar que estas *hit rates* son solo un punto de referencia y no se pueden comparar con las *hit rates* de la Tabla 14 porque los modelos no usan la misma metodología. En efecto, difieren en:

- Número de personas: los modelos de la Tabla 15 usan 60 personas en vez de 62.
- Uso de la alternativa *No Comprar*: el modelo *Greedy* lexicográfico no incluye a las respuestas que indiquen *No Comprar*.
- Conjuntos de entrenamiento y de prueba: el modelo *Greedy* lexicográfico usa un 70% seleccionado al azar de las preguntas para entrenar y prueba con el 30% restante, mientras el Logit hace una validación “*leave one out*” usando 13 preguntas para predecir una respuesta y los modelos de la Tabla 14 usan otra metodología.

<i>Hit rate</i>	Conjunto de entrenamiento	Conjunto de prueba
<i>Greedy</i> lexicográfico	42%	30%
Logit Multivariado	89%	63%

Tabla 15: *Hit rates* de referencia en el experimento de las cafeteras

7.4. Algoritmos que usan datos de movimiento ocular

En esta parte se aplican modelos elaborados en 6.5 con movimientos oculares al experimento de las cafeteras. Del mismo modo que con los algoritmos que usan elecciones, la inexistencia de un ranking universal de preferencia de los niveles de los atributos 1 a 4 (ver 7.1) obliga a agregar un paso inicial a los algoritmos para calcular estos rankings para cada persona. Para una persona p y un atributo $a \in \llbracket 1; 4 \rrbracket$ dado, el ranking de preferencias se construye a partir de las fijaciones por aparición del conjunto de entrenamiento. Se calculan los $nFixPerAppear_{p,a,l}$ y se ordenan los niveles del atributo a según su número de fijaciones por aparición. Este paso se basa en la hipótesis que para un atributo dado, los participantes demuestran más interés por los niveles que son aceptables.

A continuación se mostrarán los resultados (Ilustración 37 y Tabla 16) del modelo 6.2 que obtuvo los mejores resultados.

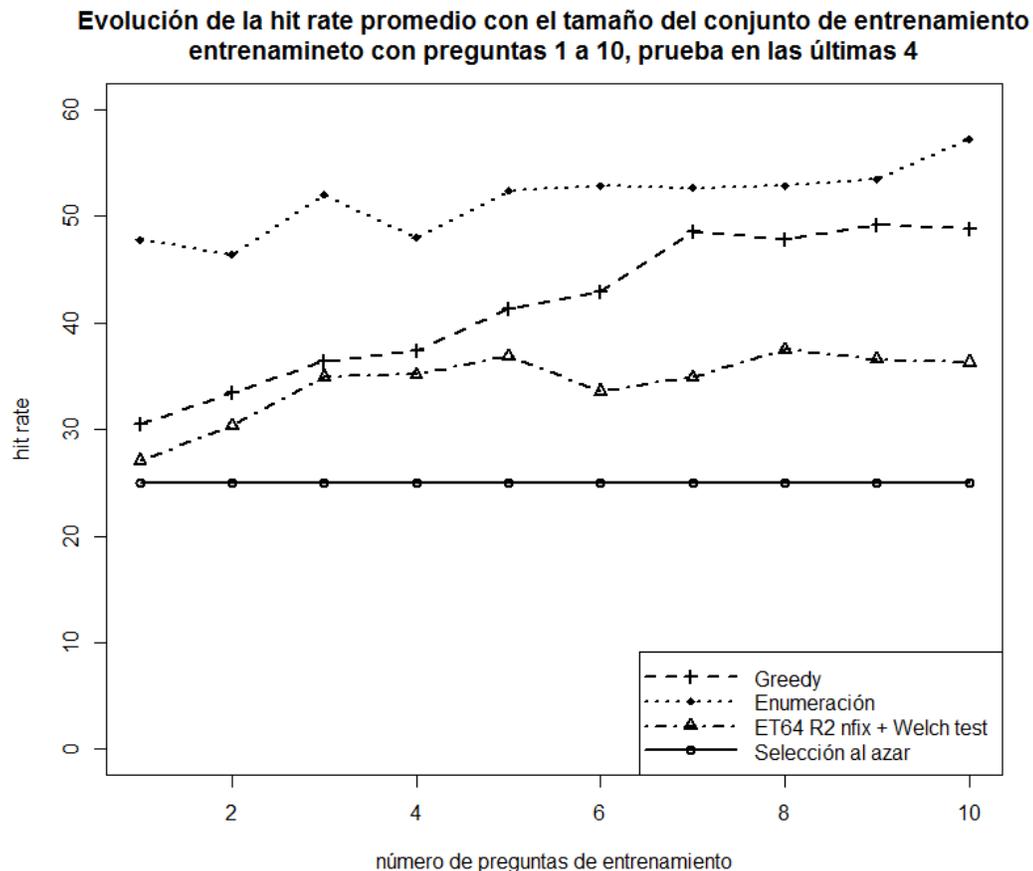


Ilustración 37: Predicción de elecciones con movimientos de ojos: experimento de las cafeteras

Los resultados obtenidos se muestran en la Ilustración 37 que grafica la *hit rate* en el conjunto de prueba para conjuntos de entrenamiento que van de 1 a 10 preguntas (se cambió el eje de la hit rate que va de 0 a 50% ahora) y en la Tabla 16 que revela las *hit*

rates en los conjuntos de entrenamiento y de prueba para 10 preguntas de entrenamiento. No son buenos: el modelo retorna una hit rate de 36% con 10 preguntas de entrenamiento, solamente 11 puntos más que la selección al azar.

Hit rate	Conjunto de entrenamiento (10 primeras preguntas)	Conjunto de prueba (4 ultimas preguntas)
<i>Greedy</i> conjuntivo	74%	49%
Enumeración	81%	57%
Modelo 6.2 con movimientos de ojos	36%	36%

Tabla 16: *Hit rate* para *Greedy*, Enumeración y modelo 6.2: experimento de las cafeteras

7.5. Algoritmos que combinan elecciones y datos de movimiento ocular

Esta parte busca combinar elecciones y datos de movimiento ocular para obtener un modelo que supere los resultados del *Greedy* o del algoritmo de enumeración.

De nuevo, la inexistencia de un ranking universal de preferencia de los niveles de los atributos 1 a 4 (ver 7.1) obliga a agregar un paso inicial a los algoritmos para calcular estos rankings para cada persona. Para una persona p y un atributo $a \in \llbracket 1; 4 \rrbracket$ dado, el ranking de preferencias se construye a partir no solo de las elecciones por aparición pero también de las fijaciones por aparición del conjunto de entrenamiento. Se calculan $nFixPerAppear_{p,a,l}$ y $nChoicesPerAppear_{p,a,l}$, se normalizan los dos y se crea una variable mixta:

$$\begin{aligned}
 nFix\&ChoicesPerAppear_{p,a,l} = \\
 &nFixPerAppear_{p,a,l}^{normalizado} \\
 &+ nFixPerAppear_{p,a,l}^{normalizado}
 \end{aligned}$$

Se ordenan los niveles del atributo a por valores crecientes de $nFix\&ChoicesPerAppear_{p,a,l}$.

Modelo 7 Variante 1: combinar elecciones y fijaciones para calcular los rankings

Se hace un primer modelo combinado, llamado 7.1, que usa estos rankings combinados y luego el *Greedy* normal con las fijaciones. Resulta que cambiar los rankings tiene poco impacto en los resultados (ver la predicción de elecciones en la Ilustración 40 en anexos). Se concluye que los rankings calculados usando elecciones, movimientos de ojos o los 2 combinados no difieren mucho a nivel general, lo que se confirma con un análisis descriptivo rápido. A continuación todos los modelos combinados usarán estos rankings obtenidos con fijaciones y elecciones.

Modelo 7 Variante 2: usar las fijaciones en caso de empate de hit rate en el Greedy

Luego se construye un segundo modelo, el modelo 7.2 que usa las fijaciones para determinar que atributo cambiar si hay un empate de *hit rate* entre varios atributos uno de los bucles del *Greedy*. En efecto, en cada bucle se calcula una matriz *H* de *hit rates* que entrega para cada nivel (columnas) de cada atributo (líneas) la *hit rate* que se obtiene al modificar solamente este atributo con este nivel. Se muestra un ejemplo en la Ilustración 38.

```
$hitrates$loop1
  level
att  2  3  4
  1 35 45 20
  2 30 30  0
  3 40  0  0
  4 45 50 20
  5 48 50  0
  6 35 38 30
```

Ilustración 38: Ejemplo de matriz de *hit rates* de un bucle de la heurística *Greedy*

En este bucle, la heurística *Greedy* modifica el vector de umbrales con el atributo y el nivel que tienen mayor *hit rate* en la matriz. Pero aparece un problema cuando hay varios atributos y/o niveles que tienen la máxima *hit rate*. Este caso surge en promedio en 1.5 bucles sobre los 6 máximos efectuados por la heurística. Se resuelve en el modelo con las fijaciones: dentro de los atributos con empate, se modifica el atributo que tiene mayor número de fijaciones y dentro de los niveles con empate, se modifica el nivel con menos utilidad (para más generalidad). Esta modificación tampoco tiene mucho impacto en los resultados (ver la predicción de elecciones en la Ilustración 40 en anexos).

Modelo 7 Variante 3: combinar elecciones y fijaciones en la modificación de atributos en el Greedy.

Con el fin de agregar más datos de movimientos de ojos a la heurística *Greedy* se diseña un tercer modelo 7.3 que incluye más datos de movimientos de ojos. Como se mencionó

anteriormente, en cada bucle se calcula en la heurística *Greedy* una matriz H de *hit rates* que entrega para cada nivel (columnas) de cada atributo (líneas) la *hit rate* que se obtiene al modificar solamente este atributo con este nivel. Además se crea una segunda matriz, la matriz F , del mismo tamaño, que entrega para cada nivel (columnas) de cada atributo (líneas) el número de fijaciones por aparición de este nivel de este atributo, normalizado para estar en el rango $[0; 100]$. De esta manera se pueden comparar las dos matrices H y F . La matriz H representa la información de las elecciones y la matriz F de los movimientos de ojos. Para combinar las 2 se introduce una tercera matriz M que es una combinación lineal de H y de F con un parámetro γ que corresponde al peso dado a las elecciones:

$$M = \gamma H + (1 - \gamma)F$$

Se prueba este modelo con distintos valores de γ ($\gamma = 0.25$; $\gamma = 0.5$; $\gamma = 0.75$) pero los resultados nunca superan los del *Greedy* (ver la predicción de elecciones en la Ilustración 41 en anexos). Los resultados empeoran cuando se baja el peso de las elecciones γ , lo que indica que las fijaciones no aportan información valiosa en este modelo. De la misma manera que el modelo 3 que elige el nivel que tiene el mayor número de fijaciones por aparición, este modelo 7.3 tiene la desventaja de no incluir un test estadístico para comparar los números de fijaciones por aparición.

Modelo 7 Variante 4: combinar elecciones y fijaciones en la modificación de atributos en el Greedy con tests de Welch

El modelo 7.4 es similar al modelo 7.3 en el sentido que modifica la heurística *Greedy* creando una matriz M que combina la matriz H de *hit rates*, basada en elecciones, con otra matriz basada en movimientos de ojos. Esta vez se reemplaza la matriz F por una matriz W de mismo tamaño tal que el valor del atributo a y del nivel l de la matriz W sea

$$W_{a,l} = 100 * (1 - pvalue(test\ Welch(l, level\ que\ entrega\ menor\ utilidad)))$$

El test de Welch hecho en $W_{a,l}$ compara las medias de los números de fijaciones por aparición del nivel l con los del nivel del atributo a que aporta menos utilidad según el ranking que usa fijaciones y elecciones. Por ejemplo para un atributo que tiene un ranking de (1,2,3,4), el nivel 1 es el menos preferido, por lo tanto el test de Welch hecho en $W_{a,l}$ con $l \in \llbracket 2; 4 \rrbracket$ tendrá la siguiente hipótesis nula:

$$H_0: \hat{f}_1 = \hat{f}_l \quad vs \quad H_1: \hat{f}_1 < \hat{f}_l$$

De esta manera se obtiene una matriz W cuyos valores están en el rango $[0; 100]$ y representan la probabilidad de que el número de fijaciones por aparición del nivel sea mayor a aquello del nivel que entrega menor utilidad.

La matriz final M usada para cambiar los atributos es una combinación lineal de H y de W con un parámetro γ que corresponde al peso dado a las elecciones:

$$M = \gamma H + (1 - \gamma)W$$

Se prueba este modelo con distintos valores de γ ($\gamma = 0.25$; $\gamma = 0.5$; $\gamma = 0.75$) pero los resultados nunca superan los del *Greedy* (ver la predicción de elecciones en la Ilustración 42 en anexos). Los resultados empeoran cuando se baja el peso de las elecciones γ , lo que indica que las fijaciones en este modelo tampoco agregan información valiosa.

En la Ilustración 39 se grafican las predicciones de elecciones (*hit rates* de las 4 últimas preguntas para conjuntos de entrenamiento de 1 a 10 preguntas) de los modelos 7.1 (que cambia únicamente los rankings), 7.2 (que usa las fijaciones en caso de empate de *hit rate*), 7.3 (que combina la matriz de *hit rates* con una matriz de fijaciones por apariciones en el *Greedy*) con $\gamma = 0.5$ y del modelo 7.4 (que combina la matriz de *hit rates* con una matriz de p-values de tests de Welch) con $\gamma = 0.5$.

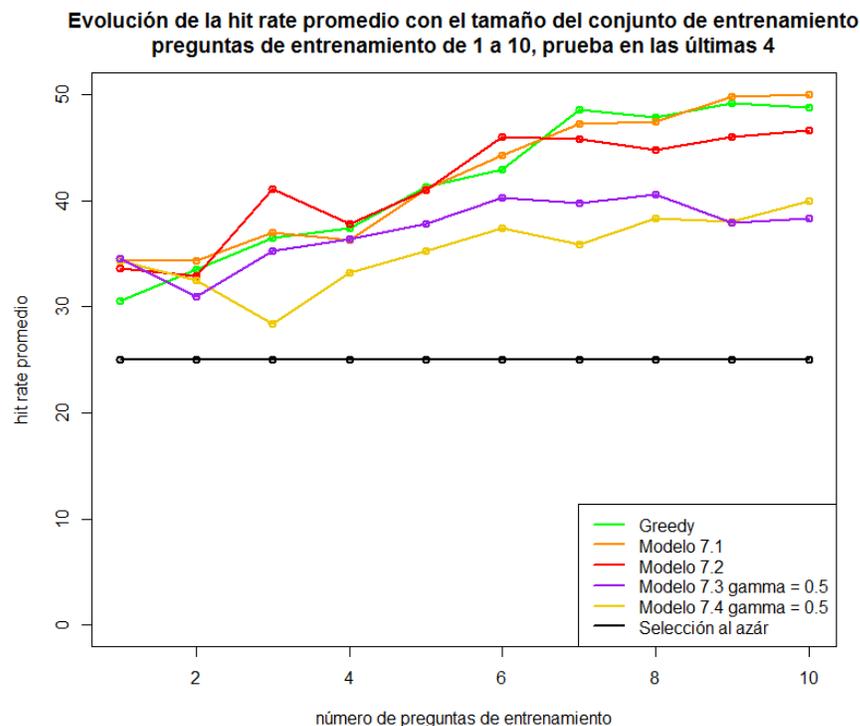


Ilustración 39: Predicción de elecciones con elecciones y movimientos de ojos: experimento de las cafeteras

Se destacan las siguientes conclusiones:

- El modelo 7.1 (en naranja) obtiene resultados muy similares a la heurística *Greedy*.
- El modelo 7.2 (en rojo) supera el *Greedy* puntualmente con un conjunto de entrenamiento de 3 y 6 preguntas pero al final con 10 preguntas de entrenamiento queda por debajo.
- El peak de *hit rate* del modelo 7.2 en la pregunta 3 se debe a un aumento del número de empates en las *hit rates* de la heurística *Greedy* cuando uno pasa de 2 a 3 preguntas de entrenamiento.
- Los modelos 7.3 y 7.4 (en violeta y amarillo) no superan al *Greedy* salvo por un conjunto de entrenamiento de 1 sola pregunta.
- La caída del modelo 7.4 para conjuntos de entrenamiento de 1 a 3 preguntas se explica con el hecho que para pequeños conjuntos de entrenamiento, las observaciones de fijaciones o de apariciones no son suficientes para conducir el test de Welch, entonces el algoritmo retorna una matriz W vacía.

7.6. Conclusiones del capítulo

Este capítulo prueba los modelos en un contexto más cerca de la realidad de las elecciones verdaderas, lo que permite generalizar la investigación. Primero, todos los modelos superan una selección al azar, lo que indica que la regla conjuntiva logra en alguna medida capturar el proceso de decisión de las personas, lo que es importante.

Los algoritmos que usan elecciones tienen una *hit rate* en el conjunto de prueba de 49% por el *Greedy* y de 57% por el algoritmo de enumeración, situándose en el orden de magnitud de otros modelos de elección usados generalmente como el *logit*. En cambio, los modelos que usan los movimientos de ojos de los participantes obtienen a lo más una *hit rate* de 36%, por lo tanto se concluye que la información visual, en conjunto con el supuesto de una regla conjuntiva, no es tan útil como las elecciones. Los modelos que combinan el algoritmo *Greedy* con datos de comportamiento ocular tampoco logran superar el *Greedy*.

Cabe señalar que otro modelo de movimientos de ojos, suponiendo una regla compensatoria y hecho por Paula Urrutia [9] arroja resultados mejores que una heurística *Greedy* con regla lexicográfica, que un *mixed logit* o que modelos de fijaciones suponiendo una regla no-compensatoria en el experimento de las cafeteras. Se concluye que la información ocular en algunos casos permite superar modelos que usan elecciones.

8. CONCLUSIONES

A la hora de estudiar las reglas de decisión con las cuales los consumidores eligen un producto, observar el proceso de elección permite capturar información que no está presente en las elecciones. En este trabajo se analizó el comportamiento ocular de participantes de un experimento de análisis conjunto. Se investigó si modelos que usan movimientos de ojos pueden inferir la regla de decisión conjuntiva usada y predecir la decisión final con más precisión que con modelos que solamente utilizan las elecciones.

El comportamiento visual resulta ser consistente con las elecciones, puesto que el producto más fijado o el último producto fijado corresponden en la mayoría de los casos al producto elegido. Además, dentro de los niveles de un atributo, se miran más los niveles que son aceptables para el consumidor. Se deduce que una persona pone más atención visual en los productos y atributos más importantes para ella, por lo tanto se puede inferir la probabilidad para un producto de ser elegido. Sin embargo, predecir la elección no es suficiente para entender completamente al consumidor: falta comprender las razones por las cuales eligió tal producto en vez de otro, es decir su regla de decisión.

Resulta que en el caso de reglas conjuntivas, los algoritmos que usan elecciones predicen el tipo de regla usada con una precisión sin parangón: 8 preguntas de entrenamiento bastan para encontrar todas las reglas conjuntivas usadas por los participantes y con eso lograr un ajuste de 100% en la predicción de elecciones. Con solamente 3 preguntas de entrenamiento se encuentra la regla correcta para la mitad de los participantes y predicen el 80% de las elecciones futuras. En comparación, los modelos que usan los movimientos de ojos alcanzan a predecir la regla conjuntiva usada con una tasa de acierto de 88% y elecciones futuras con una *hit rate* de 75% con 10 preguntas de entrenamiento. Estos resultados son bastante buenos, pero quedan por debajo de los algoritmos que usan elecciones. En definitiva, esta investigación concluye que, en el caso de reglas conjuntivas, el comportamiento ocular aporta menos informaciones que las elecciones sobre la regla usada.

A aplicar los modelos en una base de datos de elecciones libres, resulta que la regla de decisión conjuntiva alcanza en alguna medida modelar el proceso de decisión real de consumidores, dado que se predicen correctamente hasta el 57% de las elecciones hechas con modelos que usan elecciones. Los resultados de predicción obtenidos con modelos que usan fijaciones son menos satisfactorios y se limitan a una *hit rate* de 36%. Combinar datos de movimientos de ojos con la heurística *Greedy* tampoco logra mejorar los resultados. Falta ver si el hecho de agregar información ocular al algoritmo de enumeración arrojará mejores resultados.

Finalmente, la información visual coincide con las elecciones pero no aporta más información que ellas para predecir la regla conjuntiva usada. Además, integrar variables de *eye tracking* en el modelo basado en elecciones no mejora su rendimiento.

Trabajos futuros:

Este trabajo caracterizó el comportamiento ocular de participantes siguiendo una regla conjuntiva. Se vio que el patrón de fijaciones en los distintos niveles de un atributo depende del umbral de la persona en este atributo. Se observaron patrones distintos si el umbral era en el nivel 1 o en el nivel 2. Sin embargo, esta observación quedó limitada por el diseño del experimento, que no incluye a ninguna pregunta en la cual el umbral está en el nivel 3. Sería interesante analizar un experimento con un mayor número de niveles por atributo y umbrales en cada uno de ellos.

La información visual puede alcanzar un alto poder predictivo, dependiendo de los modelos de predicción y los supuestos de elección que se apliquen. Esta investigación supuso el uso de reglas conjuntivas, pero otros modelos suponen el uso de otras reglas de decisión, compensatorias como no compensatorias (ver trabajo de Paula Urrutia [9]). Estos modelos por separado alcanzan buenos resultados cuando se impone la regla a seguir, pero sus desempeño cae en bases de datos de elecciones libres porque modelan solamente una regla. Se propone combinarlos usando un modelo que primero determina la regla de decisión usada (ver [9]) y luego aplicar el algoritmo que conviene.

En esta investigación se buscaron patrones de movimientos de ojos a nivel general para luego aplicarles a nivel individual con modelos que usan el comportamiento visual de un participante para deducir su regla de decisión. Pero se pueden imaginar modelos que usan no solamente el comportamiento visual de un participante pero el de todos para inferir la regla del mismo participante. Se podrían usar algoritmos de *machine learning*, usando las fijaciones de todos los participantes y variables personales del participante para inferir su regla de decisión y así automatizar el trabajo de búsqueda de patrones.

GLOSARIO

- Regla de decisión Conjunto de condiciones que permiten al consumidor aceptar o rechazar una alternativa dentro del set de alternativas propuestas.
- Regla de decisión compensatoria Regla de decisión en la cual un nivel no satisfactorio de un atributo de una alternativa puede ser compensado por el buen nivel de la misma alternativa en otro atributo.
- Regla conjuntiva Regla de decisión no-compensatoria en la cual se definen niveles mínimos aceptables para cada atributo. Las alternativas cuyos niveles sobrepasan estos niveles mínimos para todos los atributos son satisfactorias.
- Regla disjuntiva Regla de decisión no-compensatoria en la cual se definen niveles mínimos aceptables para cada atributo. Las alternativas cuyos niveles sobrepasan estos niveles mínimos en cualquier de los atributos son satisfactorias.
- Regla lexicográfica Regla de decisión no-compensatoria que cuenta con un orden de importancia de los atributos. Se elige la alternativa que toma el mejor valor en el atributo más relevante. En el caso de empate, se procede a evaluar el siguiente atributo según la escala de importancia establecida. Se continúa de esta manera hasta quedar con una única opción.
- AOI Área de Interés: parte de la pantalla que contiene una única información.
- Segunda parte de las fijaciones Si n es el número de fijaciones de un caso, la segunda parte contiene las fijaciones $\frac{n}{2}$ a n si es n par o $\frac{n+1}{2}$ a n si es impar.
- Fijación repetida Fijación en una AOI que ya recibió una fijación en la misma pregunta.
- Umbral En una regla conjuntiva, el umbral de un atributo indica el límite aceptable para la regla: las alternativas que sobrepasan o igualan este nivel límite son aceptables, se rechazan las otras.

BIBLIOGRAFÍA

Artículos de blogs o de periódicos en línea:

[1] Gregory Ciotti, 7 *Marketing Lessons from Eye-Tracking Studies*, blog Kissmetrics. <<http://blog.kissmetrics.com/eye-tracking-studies/>> [consulta: 05 junio 2014]

[7] Emily Glazer, *The Eyes Have It: Marketers Now Track Shoppers' Retinas*, *The Wall Street Journal*. <<http://online.wsj.com/news/articles/SB10001424052702303644004577520760230459438>> [consulta: 06 julio 2014]

[8] Anne Aula y Kerry Rodden, *Eye-tracking studies: more than meets the eye*, blog oficial de Google. <<http://googleblog.blogspot.com/2009/02/eye-tracking-studies-more-than-meets.html>> [consulta: 06 julio 2014]

Artículos científicos:

[2] Riedl, R., Brandstätter, E., & Roithmayr, F. 2008. Identifying decision strategies: A process-and outcome-based classification method. *Behavior Research Methods*, 40(3): 795-807.

[3] Musalem, A., Meißner, M., & Huber, J. 2013. Do Motivated and Incidental Processing Distort Conjoint Choices? *Available at SSRN*.

[4] Funaki, Y., Jiang, T., & Potters, J. 2014. *Eye tracking Social Preferences*. Working Paper.

[5] Stüttgen, P., Boatwright, P., & Monroe, R. T. 2012. A satisficing choice model. *Marketing Science*, 31(6): 878-899.

[6] Gilbride, T. J., & Allenby, G. M. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, 23(3): 391-406.

[10] Kohli, Rajeev, and Kamel Jedidi. 2007. Representation and inference of lexicographic preference models and their variants. *Marketing Science* 26.3: 380-399.

Memoria:

[9] Urrutia Moreira, Paula Javiera, 2014, *Estudio del uso de reglas de decisión lexicográficas utilizando herramientas de eye tracking*, Memoria para optar al título de Ingeniera Civil Industrial, Santiago, Universidad de Chile.

Sitio web:

[11] “Welch's t test” de la encyclopedia Wikipedia. Disponible en <https://en.wikipedia.org/wiki/Welch%27s_t_test> [consulta: 26 septiembre 2014]

ANEXOS

a. Tablas de contingencia

		Producto elegido				
		1	2	3	4	Total
Producto más fijado	1	116	46	36	48	246
	2	45	156	53	70	324
	3	9	3	81	20	113
	4	2	3	3	29	37
	Total	172	208	173	167	720

Tabla 17: Tabla de contingencia producto más mirado y producto elegido

		Producto elegido				
		1	2	3	4	Total
Último producto fijado	1	98	29	15	14	156
	2	29	118	23	22	192
	3	20	23	107	14	164
	4	25	38	28	117	208
	Total	172	208	173	167	720

Tabla 18: Tabla de contingencia último producto mirado y producto elegido

b. Resultados de los modelos combinando elecciones y datos de movimientos de ojos aplicados al experimento de las cafeteras.

Resultados de los modelos 7.1 y 7.2:

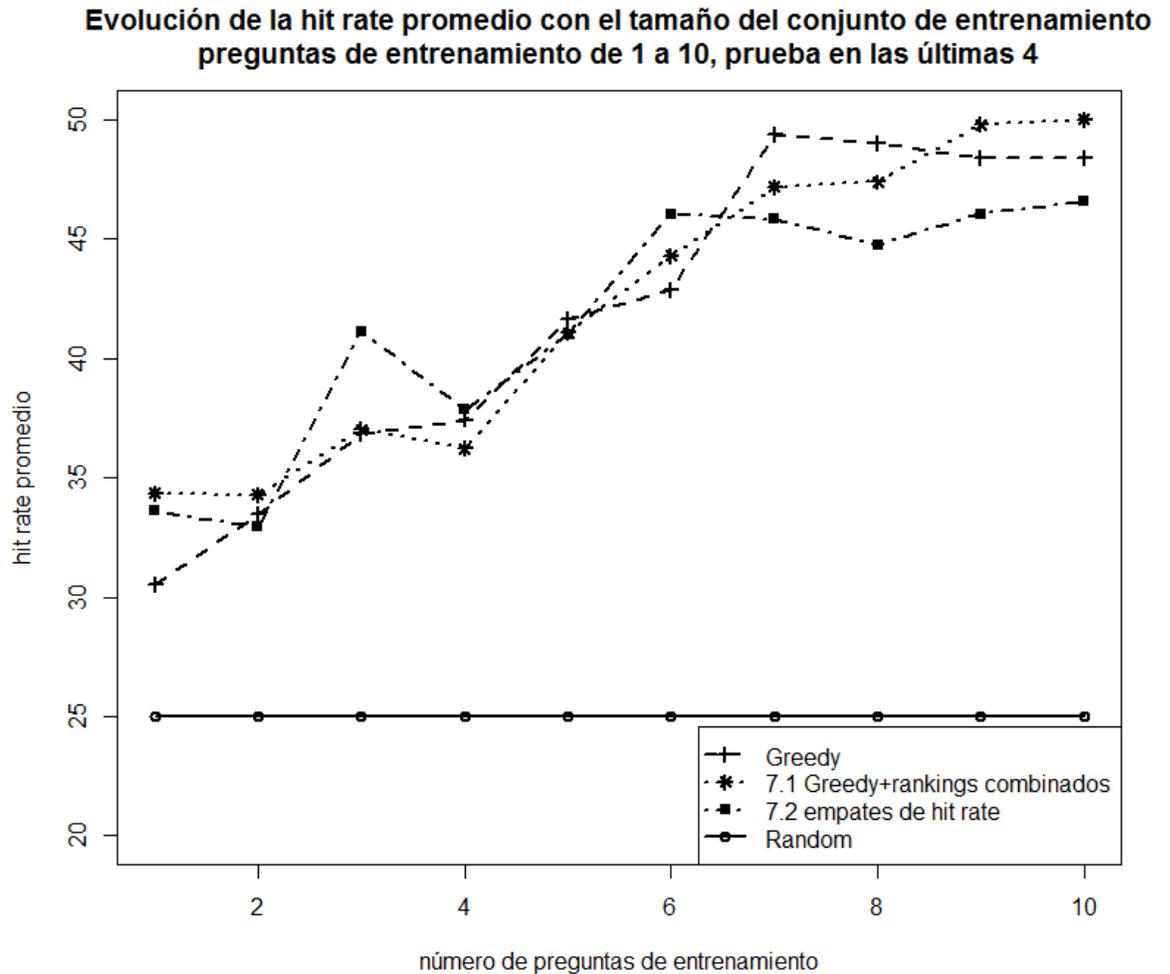


Ilustración 40: Predicción de elecciones, modelos combinados, experimento de las cafeteras

Variación del modelo 7.3 con $\gamma \in \{0.25; 0.5; 0.75\}$

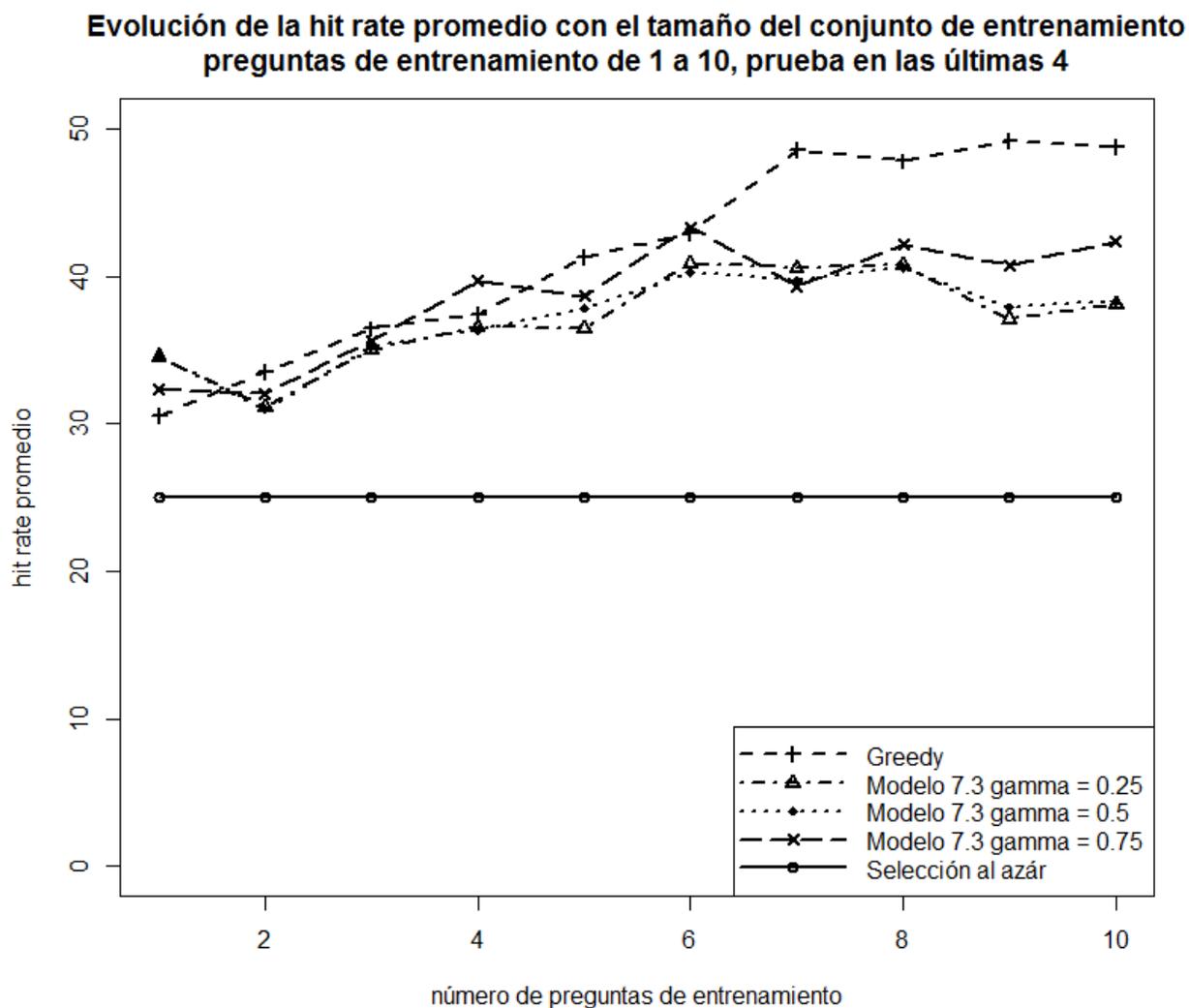


Ilustración 41: Variaciones del modelo 7.3 con gamma

Variación del modelo 7.4 con $\gamma \in \{0.25; 0.5; 0.75\}$

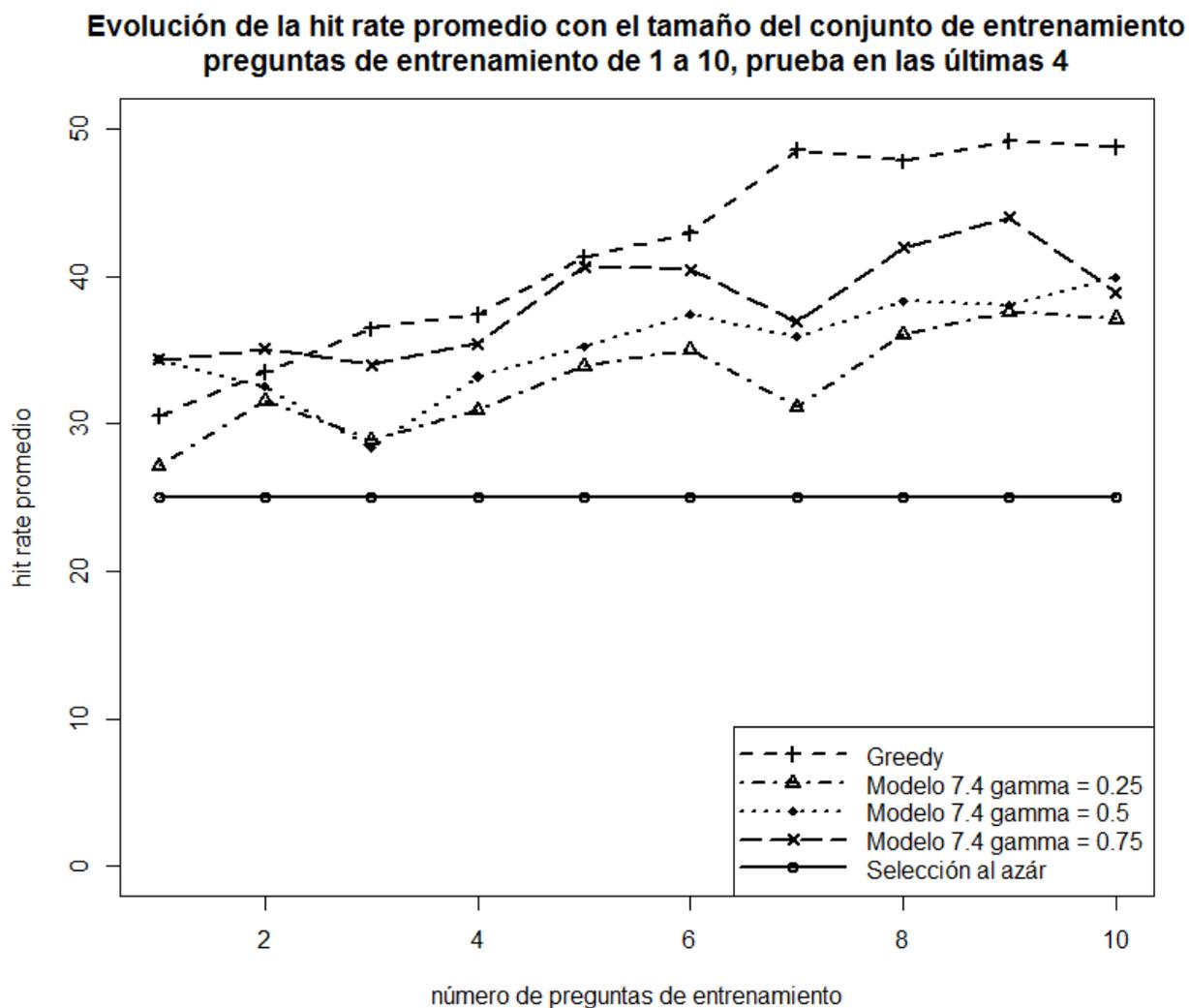


Ilustración 42: Variaciones del modelo 7.4 con gamma