



Control 2

El siguiente control está pensado para una duración de 2:30 horas. Ponga de inmediato su nombre en las hojas de respuesta, si lo desea puede utilizar también las hojas de enunciado para responder.

Pregunta 1

1. (1 pto) Explicar en qué consiste la aplicación de neurociencia al estudio del comportamiento del usuario en la Web.
2. Suponga que tiene acceso a una base de datos de sesiones de navegación de un conjunto de usuarios junto a una base de datos de respuestas fisiológicas, entre las que se incluyen la posición ocular, la dilatación pupilar en el tiempo y el potencial eléctrico de la corteza cerebral medido con un electroencefalograma (EEG). Además, considere que tiene acceso a un mapeo de los objetos presentes en los sitios Web donde ocurren las navegaciones, es decir, las coordenadas y tamaños de aquellos objetos.

Si no existe mayor información que la mencionada, explique:

- a (1 pto) ¿Cómo confeccionaría el vector de características?
 - b (1 pto) Qué modelos utilizaría para generar posibles clasificaciones con respecto al comportamiento de los usuarios, detallando qué es necesario para cada modelo.
 - c (1 pto) ¿Qué tipo de conclusiones se podrían obtener a partir de este estudio?
3. (1 pto) Si además se entregan datos adicionales sobre el comportamiento del mouse, es decir, se conoce cuáles objetos específicamente fueron clickeados durante las sesiones. Para este caso, ¿qué información buscaría obtener y cuáles modelos utilizaría?
 4. (1 pto) ¿Cuáles complicaciones pueden surgir al utilizar este tipo de análisis en el estudio del comportamiento del usuario en la Web? ¿Cómo podría subsanar aquellas complicaciones?



INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

IN5526 - Web Intelligence

Profesores: Juan D. Velásquez
Pablo Tapia
Auxiliar: Gino Slanzi

Hoja Respuesta Pregunta 1

Nombre:

Hoja Respuesta Pregunta 1

Nombre:



Pregunta 2

Suponga que se desea realizar un ranking de popularidad de usuarios de Twitter, basado en los seguidores que tiene cada usuario. A modo de prueba se le entregan los siguientes datos sobre 5 usuarios de esa red social:

Usuario	Sigue a
1	4
1	2
2	1
2	3
2	4
3	4
4	1
4	2
4	5
5	1
5	3
5	4

1. (0,5 pts) Represente los datos de la tabla en forma de grafo dirigido.
2. (2,5 pts) En base al grafo anterior, estime la popularidad de cada usuario mediante PageRank para tres iteraciones. Use la tabla presente en la hoja de respuesta.
3. (1 pto) Proponga al menos un método alternativo para evaluar la popularidad de los usuarios de Twitter.
4. (1 pto) Pruebe su método propuesto con los datos entregados y compare esos resultados con lo obtenido mediante PageRank. Comentar.
5. (1 pto) Explique las diferencias entre los algoritmos PageRank y la versión clásica de HITS.

$$r_p^{(i+1)} = \frac{b}{N} + (1 - b) \sum_{\forall q, q \rightarrow p}^{I_p} \frac{r_q^{(i)}}{\text{outdegree}(q)} \quad (1)$$



Hoja Respuesta Pregunta 2

Nombre:

Nodos	PageRank por Iteración			
	0	1	2	3
1				
2				
3				
4				
5				



INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

IN5526 - Web Intelligence

Profesores: Juan D. Velásquez
Pablo Tapia
Auxiliar: Gino Slanzi

Hoja Respuesta Pregunta 2

Nombre:



Pregunta 3

1. (1 pto) A partir de lo expuesto en el curso: De qué forma agrega valor en una empresa el Web Mining? De un ejemplo.
2. (1 pto) En el lenguaje del marketing digital, ¿Qué se define como conversión? De tres ejemplos.
3. (1 pto) Qué es el Web Usage Mining y como se puede aplicar para mejorar la usabilidad de un sitio web? De un ejemplo
4. (2 pts) Escriba en forma de pseudo código el procedimiento para realizar una sesionización de un archivo de weblogs. Considere que el weblog a procesar está compuesto por los siguientes campos.
5. (1 pto) Qué problema del procedimiento de reconstrucción de sesiones está relacionado con el Path Completion Heuristic? En qué consiste esta técnica? De un ejemplo ocupando un diagrama (grafo).



INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

IN5526 - Web Intelligence

Profesores: Juan D. Velásquez
Pablo Tapia
Auxiliar: Gino Slanzi

Hoja Respuesta Pregunta 3

Nombre:



INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

IN5526 - Web Intelligence

Profesores: Juan D. Velásquez
Pablo Tapia
Auxiliar: Gino Slanzi

Hoja Respuesta Pregunta 3

Nombre: