

Inteligencia Artificial (IS) 2019/20

ANEXO - Trabajo

Selección de características

Juan Galán Páez

1. Introducción

Tal y como se ha informado (mediante el foro de avisos específico de la plataforma), atendiendo a las instrucciones de las diferentes instancias académicas responsables, se ha realizado una reprogramación de los contenidos y de la evaluación de la asignatura.

Dado que el porcentaje de valoración de la parte práctica pasa a ser el 60% sobre el total de contenidos, se elabora el presente Anexo a la propuesta de trabajo que se indica en el título conteniendo las indicaciones de ampliación de dicha propuesta.

2. Un nuevo objetivo a cubrir

Con objeto de no alterar la numeración de los objetivos de la propuesta inicial añadiremos un Objetivo 0 independiente del resto. Con esto se pretende que esta ampliación no afecte a los avances que ya algunos alumnos hayan realizado hasta el momento.

El **Objetivo 0** de este trabajo consiste en ampliar el estudio teórico sobre selección de características que se incluirá en el documento del trabajo. La temática de selección de características es amplia y diversa. La nueva documentación ampliada debe presentar en líneas generales el problema de la selección de características y los diferentes tipos de métodos existentes. En la medida de lo posible, ampliaremos información sobre el tema (no es difícil encontrar información en inglés en la red) ya sea a nivel general o tratando en detalle un temas concreto a elegir por los alumnos. Algunos ejemplos (aunque existen muchas mas posibilidades) de temas en los que se puede profundizar son:

- Análisis detallado de las 3 grandes familias de métodos de selección de características (filtrado, envoltura e integrados), comparación, ventajas e inconvenientes de cada tipo, cuando usar uno y otro, etc.
- Estudio y comparación de diferentes algoritmos de búsqueda (que son de tipo wrapper o de envoltura).
- Importancia de características por permutación (también de tipo wrapper o de envoltura)
- Detallar el funcionamiento de algún método concreto de tipo integrado (embedded methods) como pueden ser:
 - Importancia de características (feature importance) en árboles de decisión.
- Detallar el funcionamiento de algún método concreto de tipo filtrado (filter methods).
- Es fácil encontrar artículos que describen un algoritmo o aplicación concreta. En este caso, tras comprender el caso, procederíamos a resumir en nuestra memoria el caso de uso o algoritmo concreto.

En general se deja libertad para elegir el tema a tratar. Lo importante es mostrar que se ha entendido la problemática en general y su importancia.

Algunas consideraciones:

- Los alumnos tendrán libertad de buscar documentación y referencias en Internet, siempre que estas sean citadas y su contenido no sea copiado literalmente. Este debe ser entendido y sintetizado.
- Aunque existe material en castellano, se recomienda consultar referencias en inglés ya que son más ricas y fiables (recordemos que prácticamente toda la documentación científica existente está escrita en inglés).

Algunas referencias interesantes (aunque existen muchas más):

- Serie de 6 artículos que introducen conceptos generales. [Enlace](#)
- En los capítulos 10, 11 y 12 del [siguiente libro](#) se introducen algunos conceptos y algoritmos tanto de búsqueda como de filtrado.
- El [siguiente artículo](#) presenta una revisión de métodos de búsqueda (de tipo wrapper) similares a los algoritmos objetivo de este trabajo.
- El [siguiente artículo](#) presenta de forma resumida la descripción o especificación de las técnicas y algoritmos más populares.
- Importancia de características (feature importance) basada en permutación. [Enlace](#)
- El [siguiente enlace](#) proporciona nociones básicas sobre los fundamentos del cálculo de la importancia de características (feature importance) en árboles de decisión.

3. Su inclusión en la memoria

Ahora este trabajo consta de tres objetivos cuya consecución se plasmará en la memoria como se indica:

- La consecución del Objetivo 0 constituirá el contenido de un nuevo apartado denominado “Marco teórico” que cada grupo de trabajo puede añadir al principio o bien al final, pero antes de la bibliografía. En caso de ir al final, el título más adecuado sería “Anexo. Marco teórico”. La bibliografía y los enlaces consultados (con fecha de último acceso) utilizados en este nuevo apartado deberán ir igualmente referenciados en el texto e incluidos en el apartado Bibliografía común, al final de la memoria.
- Respecto de la memoria, el resto de objetivos se mantienen igual.

4. En relación con la defensa del trabajo

El nuevo objetivo y su apartado en el trabajo no formarán parte de la exposición inicial o presentación del alumno en el acto de defensa. Sí podrá ser objeto de alguna cuestión por parte del tutor en el tiempo dedicado a las preguntas.

5. Criterios de evaluación

El apartado dedicado a los criterios de evaluación de la propuesta inicial se ve afectado en los detalles que se subrayan a continuación, manteniéndose igual el texto restante.

El apartado dedicado al documento, queda redactado como sigue.

- El documento – artículo en formato PDF. Deberá tener una extensión mínima de 6 páginas a las que el Anexo sumará entre 2 y 4 páginas más. (El resto del párrafo se mantiene igual)

En cuanto a los criterios de valoración, los cambios se subrayan a continuación:

Para la evaluación se tendrá en cuenta el siguiente criterio de valoración, considerando una nota máxima de 6 puntos para el trabajo:

- El código fuente (hasta 2.5 puntos): (el resto del párrafo se mantiene.)
- El documento – artículo científico (hasta 2 puntos):
Se valorará la claridad de las explicaciones, el razonamiento de las decisiones, el análisis y presentación de resultados, y el uso del lenguaje. Se valorará el uso de una plantilla de las que se facilitan. Igualmente, no se evaluará el trabajo si se detecta cualquier copia textual de otras fuentes: la redacción debe ser personal, en particular en el caso del contenido del Anexo a que este documento se refiere.
- La presentación y defensa (hasta 1.5 puntos): (el resto del párrafo se mantiene.).
- Mejoras: Se valorarán hasta con 1 punto extra sin superar el máximo de 6 puntos totales del trabajo.

6. Resto de apartados

El resto del documento queda redactado en los mismos términos iniciales.

7. Comentarios finales

Se insiste a los alumnos en que deben suscribirse al Foro específico dedicado a las cuestiones sobre este trabajo. El uso del mismo va orientado a compartir/despejar dudas generales que puedan surgir en la realización de los trabajos. También pueden iniciarse hilos con un mensaje que plantee una duda genérica y elegir “Responder en privado” a la hora de compartir código o detalles sobre el diseño más concretos de la solución que el alumno esté desarrollando y sobre la que realiza la consulta. Esto no anula la posibilidad de contactar con la tutora por correo.

La suscripción al foro permite que el alumno reciba notificaciones por correo de los nuevos mensajes que se incorporen al mismo y que pueden ser de su interés. La dirección de correo de cada alumno es modificable en cualquier momento por el interesado.