

Machine Learning

Guías de Estudio Semanales

Comentario general sobre los contenidos de la segunda unidad

Esta unidad se presenta un algoritmo de clasificación que se basa en un principio muy sencillo, asignar el grupo de un nuevo individuo a la clase del k vecino más próximo.

Recordemos los contenidos de la unidad:

2.1. Comprender la clasificación con los vecinos más cercanos.

2.2 El algoritmo kNN.

2.3 Distancias entre datos.

2.4 Elección de una k adecuada.

2.5. Preparando los datos

2.6. Ejemplo de aplicación.

El primer bloque (2.1-2.2) introduce el proceso de clasificación y muestra sus características principales.

El segundo bloque (2.3-2.4) muestra por un lado, la importancia que tiene escoger la distancia, en este caso solo se muestra la distancia euclídea y por otro, como escoger el número k de vecinos más próximos.

El tercer bloque (2.5-2.6) presenta algunas técnicas de preparación de los datos y a continuación, se realiza un ejemplo guía.

Materiales de estudio

Podéis basaros en el material que os resulte más conveniente para vuestra forma de estudio.

A continuación disponéis de una tabla en donde podéis ver, para cada apartado del tema, que material de consulta utilizar.

2.1. Comprender la clasificación con los vecinos más cercanos.	Cap. 3 (hasta "Understanding classification using nearest neighbors")
2.2 El algoritmo kNN.	Cap. 3 ("The kNN algorithm")
2.3 Distancias entre datos.	Cap. 3 ("Calculating distance")
2.4 Elección de una k adecuada.	Cap. 3 ("Choosing an appropriate k")
2.5. Preparando los datos	Cap 3 ("Preparing data for use with kNN")
2.6. Ejemplo de aplicación.	Cap 3 (de "Why is the kNN algorithm lazy" hasta "Summary")

Actividades complementarias

Las actividades prácticas que podéis realizar para reforzar el aprendizaje son:

- El paquete *animation* de R tiene la función *knn.ani()* que muestra visualmente el algoritmo. Podéis ejecutar el ejemplo de la función y después hacer modificaciones e incluso un pequeño informe.
- Reproducir el ejemplo de aplicación haciendo un informe dinámico.
- Para los más inquietos podéis hacer el informe dinámico de tal manera que admita otra fuente de datos y al hacer el informe lo tenga en cuenta sin modificar el código.

Si encuentras materiales complementarios de interés o que desees discutir puedes añadirlos a la actividad "**Upload Material añadido por el estudiante para compartir**" en Moodle 2. La idea es muy simple, al entrar en la actividad aparecen diferentes pestañas. Si se tiene un pdf, video, url, ... que quieres subir hay que ir a la pestaña añadir entrada y completar la ficha. Una vez completa hacer clic al botón inferior. Si se quieren ver las aportaciones actuales hay que ir Ver lista.