

Trabajo grupal

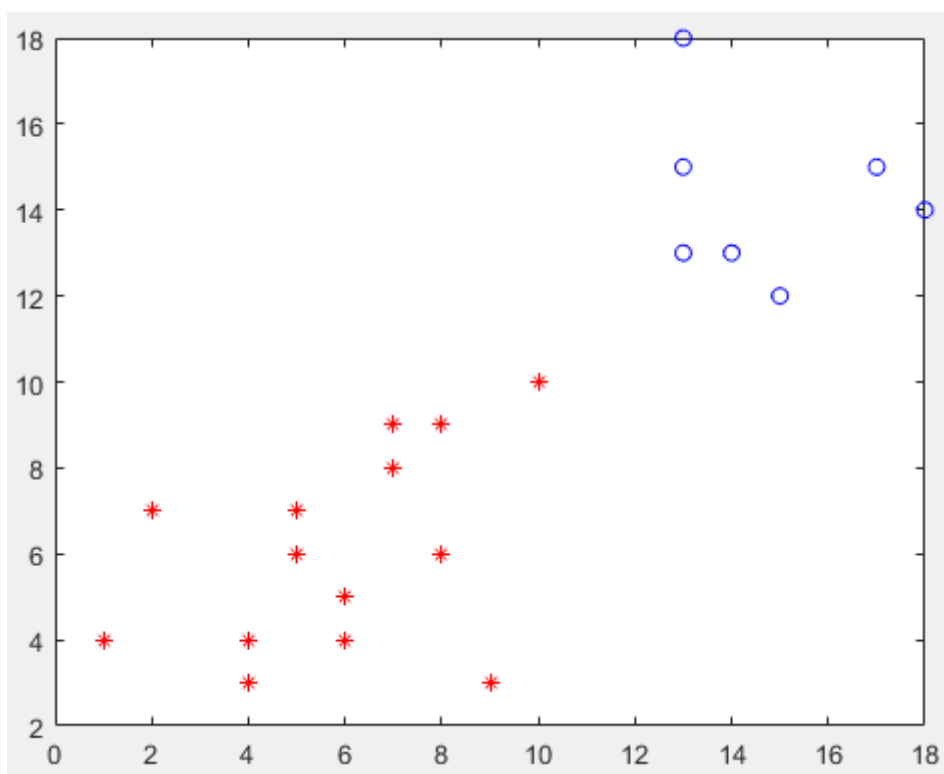
En este trabajo debéis resolver un problema de agrupamiento mediante algoritmos voraces. En un problema de agrupamiento partimos de un conjunto de datos que queremos dividir en varios subconjuntos, de tal forma que los datos que sean más cercanos deben pertenecer al mismo subconjunto, y los datos que estén más alejados deben pertenecer a subconjuntos diferentes.

Recordad que la técnica de algoritmos voraces es una de las más sencillas a la hora de ejecutarse, por lo que es bastante rápida. Esto es así porque una vez que se toma una decisión (en este problema la asignación de un dato a un subgrupo concreto), esta decisión ya no cambia nunca más. Debido a esta simplicidad y rapidez, los algoritmos voraces no siempre obtienen la solución óptima.

Deberéis pensar qué estrategia es la mejor para este problema: en qué orden se estudian los datos, cómo se elige el grupo al que pertenece el dato... para obtener la mejor solución posible en diferentes conjuntos de datos.

Ejemplo

Tenemos los datos que se muestran en la siguiente figura y queremos dividirlos en dos subgrupos. El algoritmo voraz debe determinar que los datos marcados como asteriscos rojos pertenecen a un subgrupo y los datos marcados como círculos azules pertenecen a otro subgrupo.



Parámetros del programa

El programa debe leer un fichero de texto de sigue el siguiente formato:

- En la primera línea siempre se almacenan 2 números enteros. El primero de ellos indica el número de datos que hay en total. El segundo número indica el número de subgrupos que se quieren crear.

- A partir de la segunda línea se almacenan los datos del conjunto de datos. Habrá una línea por cada dato, y cada línea está compuesta dos números enteros que indican la posición de dicho dato: el valor en el eje x y el valor en el eje y.

El programa debe asignar un subgrupo a cada uno de los datos de entrada, y almacenar dicha solución en otro fichero de texto. Este fichero contendrá una línea por cada dato a agrupar, y dicha línea solo contendrá un número entero, que indique el subgrupo asignado. Por tanto, el fichero de salida tendrá una línea menos que el fichero de entrada.

Para el ejemplo anterior, el fichero de entrada será el siguiente:

```
23  2
7   8
2   7
4   4
7   9
8   6
1   4
10  10
8   9
5   6
5   7
5   6
4   3
6   4
6   5
9   3
17  15
13  13
13  18
13  18
13  15
15  12
14  13
18  14
```

Y el fichero de salida será el siguiente:

```
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
2
2
```

2
2
2
2
2
2

Evaluación del algoritmo

Para evaluar cuánto de bueno es el algoritmo que habéis desarrollado, debéis implementar otro algoritmo que compare el porcentaje de aciertos entre el subgrupo asignado a cada dato por vuestro algoritmo y el subgrupo real al que pertenecen.

Desarrollo del trabajo

Dentro del grupo debéis probar varias estrategias y, de entre todas ellas, seleccionar la que obtenga mayores porcentajes de acierto sobre los conjuntos de datos que han sido proporcionados.

Evaluación dentro de la asignatura

Este trabajo tiene un peso de un 20% dentro de la nota de la asignatura. Se debe realizar en grupos de 4 alumnos y la calificación obtenida será igual para todos los miembros del grupo. Esta calificación se obtiene a través de dos aspectos:

- 50% - competición entre compañeros. Se entregarán 3 nuevos conjuntos de datos de los que no se entrega la solución ideal. Con el algoritmo voraz definitivo, debéis asignar los subgrupos a cada dato y entregar los tres ficheros de salida. Yo evaluaré el porcentaje de acierto para cada uno de ellos. Con el porcentaje de acierto de todos los alumnos, se hará un ranking que determinará la nota de esta parte.
- 50% - presentación del algoritmo implementado. Se debe realizar una presentación de 10 minutos en la que se expliquen la versión final del algoritmo y al menos otras dos estrategias diferentes que se hayan probado. Esta presentación se realizará por un solo miembro del grupo, que será elegido por sorteo en el momento de presentar.