

Relatório de Implementação - Vida Extraterrestre -

Grupo de trabalho: 76

Membros:

Tiago Robalo nº 51628

Alexandre Vidal nº44366

Implementação:

Este projeto foi desenvolvido em conjunto pelos dois alunos em igual parte (50% / 50%). A implementação encontra-se organizada por ficheiros sendo que os ficheiros correspondem a determinadas classes com respetivos métodos. Todas os métodos seguem a implementação por contrato.

Os Sistemas Operativos utilizados durante a realização deste projeto foram o Windows 10 Pro Versão 1803.

1) Módulo 'constants.py', que guarda as constantes que foram consideradas relevantes à realização deste projeto.

2) Módulo 'inputFile.py', com a seguinte Classe:

Classe inputFile, que é inicializada passando o nome do ficheiro de entrada. Cria uma lista em que cada elemento da lista corresponde a uma linha do ficheiro. Efetua outras verificações relevantes referentes à existência de Centroides Iniciais.

Utilizamos os métodos 'setExemplars' e 'setSamples' para criar 2 listas distintas. Uma, em que cada elemento da lista corresponde à linha inteira de cada um dos Exemplares/ Samples. E outra que contém os Exemplares/ Samples como Objetos do tipo Alien devidamente criados para serem utilizados pela Classe kmeansclustering. São usadas vários 'getters' para fazer return da informação necessária.

3) Módulo 'Alien.py', com a seguinte classe:

Classe Alien, que é inicializada com o Identificador (ID) de cada elemento e o seu respetivo feature vector. Implementa métodos que permitem calcular a distância entre vários feature vectors, calcular a sua dimensão etc. São Objetos desta Classe que são passados como argumentos para métodos da Classe kmeansclustering.

4) Módulo 'Cluster.py', com a seguinte classe:

Classe Cluster, que é inicializada passando como argumentos, a lista dos Objetos e o tipo de Objetos passados. Implementa métodos que permitem calcular a distância entre vários feature vectors da mesma lista, métodos que calculam os Centroides dos Clusters e fazem update se necessário, devolvem a variância etc. Objetos desta Classe são utilizados em métodos da Classe kmeansclustering.

5) Módulo 'kmeans.py', com a seguinte classe:

Classe kmeansclustering. Implementa métodos algorítmicos inicializados com, uma lista com Objetos Alien como Samples, o tipo de Obetos, o número de Clusters a formar, Objetos Alien como Centroides Iniciais (se aplicável) e a condição verbose (se aplicável). Permite 'Clusterizar' isto é, criar grupos de Samples que são semelhantes entre si calculando Centroides Artificiais. Tem também métodos que devolvem a Dissimilaridade entre Clusters. Devolve uma lista com os Clusters formados.

6) Módulo 'outputfile.py', com a seguinte classe:

Classe outputFile, inicializada com o nome inicial do ficheiro, a lista com os Clusters formados resultantes do kmeansclustering, e a lista resultante de inicializar e Classe InputFile. Implementa métodos que escrevem o resultado para um ficheiro de saída e constroem um gráfico representativo dos Clusters e os seus elementos constituintes.

7) Módulo 'clusteringAliens', com a seguinte função:

Função 'main', que consolida todos os módulos e classes associados à resolução deste projeto. Faz a asserção de erros de input.

Execução do software

O software é executado através da linha de comandos do Windows usando o seguinte comando:

py clusteringAliens.py k t filename.txt

Em que, clusteringAliens é o nome do ficheiro a chamar para correr o programa, k é o número de Clusters a formar, t o número de Trials que deve ser corrido o kmeansclustering e filename.txt o nome do ficheiro de entrada em formato .txt.