

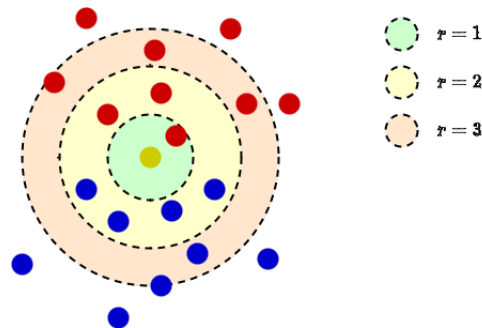
## Inteligência Artificial I - Prática 02

Nesta aula prática trabalharemos em um sistema de segmentação de clientes como um desafio de introdução ao Aprendizado de Máquina. Para esta finalidade, podemos seguir uma estratégia que é base de um dos algoritmos mais famosos da Inteligência Artificial e buscar uma forma de resolver o problema considerando a **distância entre pontos**, mais especificamente a **Distância Euclidiana**.

Neste conceito, a distância entre dois pontos  $p$  e  $q$  é medida através de suas coordenadas em um Plano Cartesiano, através da equação:

$$d_{p,q}^2 = (x_q - x_p)^2 + (y_q - y_p)^2$$

Suponha que você está trabalhando em uma empresa de *streaming* e você foi convidado(a) para uma reunião. Nesta reunião, a equipe de análise de perfis de clientes te apresentou um conjunto de dados de 20 novos clientes da empresa, o formato da figura abaixo:



Nesta figura, temos uma representação gráfica do problema de análise de padrões em um sistema bidimensional. Nele, um novo usuário, o padrão em amarelo, acessou a plataforma e buscou informações a respeito de alguns títulos de filmes e o interesse é verificar se este novo usuário pode ser atribuído ao grupo de pessoas que têm **preferencia** por comédia (padrões em azul), ou das pessoas que **preferem** drama (padrões em vermelho). Para que este objetivo seja alcançado é sugerido estabelecer que o padrão amarelo seja o centro de um círculo a ser traçado no seu entorno. Este círculo define sua região de interesse. Portanto, para uma análise mais consistente, é necessário também definir um valor para o raio deste círculo e, posteriormente, verificar **quantos padrões estão dentro da área de interesse junto ao padrão amarelo**. Dessa maneira, o critério de tomada de decisão será: **se houverem mais padrões azuis do que vermelhos dentro da área de interesse, então o padrão amarelo deve ser atribuído ao grupo de pessoas que gostam de comédia, sendo que, caso contrário, ele deve ser atribuído ao grupo de pessoas que gostam de drama**.

Partindo do código base disponível no CANVAS nas extensões `.ipynb` (*Notebook*) e `.py` (*Python File*), resolva o que se pede:

**Questão 1)** Estabeleça o valor de  $r = 0.3$  e verifique se o padrão amarelo pertence ao grupo azul ou ao grupo vermelho, através do critério da Distância Euclidiana.

**Questão 2)** Retrabalhe o seu código e crie uma função para o cálculo da Distância Euclidiana, onde o valor de  $r$  será um parâmetro a ser escolhido pela pessoa que utilizará o seu código na empresa. Ou seja, a pessoa pode escolher qualquer valor para  $r$  e o código retornar como resposta a qual grupo o padrão amarelo pertence.

**Questão 3)** Repita os itens a) e b) considerando agora que os analistas ficaram Satisfeitos com os resultados que você obteve nas tarefas anteriores e decidiram enviar a você um problema mais complexo. Agora, são analisados os perfis de 50 clientes que se dividem entre gostar dos gêneros Ficção Científica e Ação. Assim como no caso anterior, existem 3 novos clientes na plataforma que ainda não tiveram seu perfil definido como fã de Ficção ou de Ação e **seu objetivo agora é estabelecer uma forma de fazer com que eles tenham o perfil atribuído a um dos gêneros em questão.**