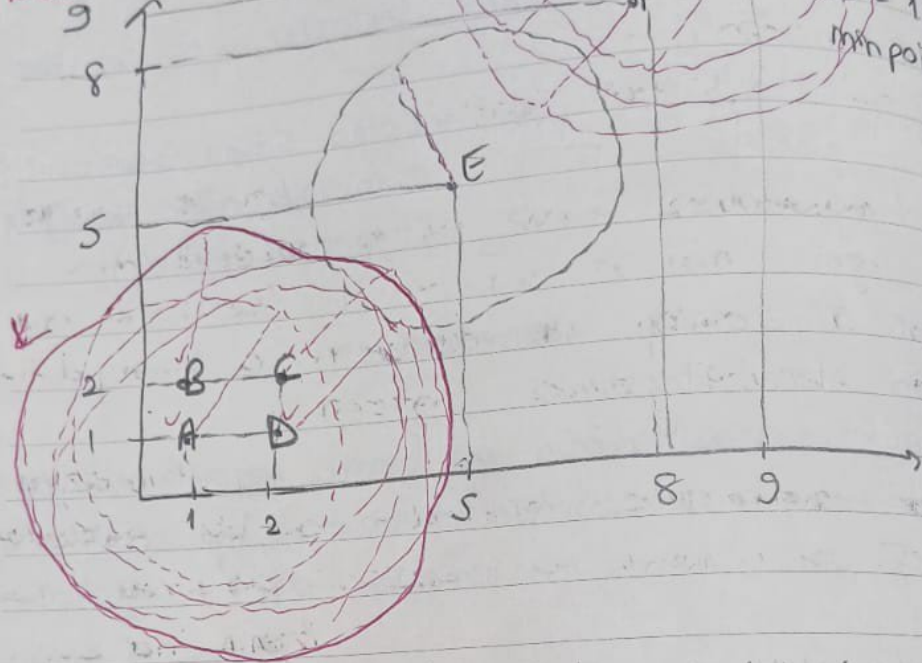


DBSCAN örneği



- Bu pembe kmenin içindeki bütün elemanların kendi etrafındaki comberden sıkları sıkladılar ve kome oldular
- E li sıklını sıklamadı ve kome olmadı veri setinde dışlandı.
- F, G ve H' de min pointi sıklayamadı ve veri setinde dışlandı. Bu veri setinde sadece k sıkladı

Reinforcement Öğrenme (Reinforcement Learning)

- Bir AI modelinin, bir ortamla etkileşime geçerek belirli bir görevi gerçekleştirtilmeyi öğrenmesini sağlayan bir öğrenme türüdür (veri yok deneme - yanılma)
- oyun yapay zekası, robotik, otomatik koordinatlar, alarmlar kullanılmaktadır.

Environ

State, Reward

Reinforcement

→ model

olayları

model

bu teknik

bir ser

→ model

ve olay

öğrenir

→ Depo

ian k

etme

→ Po

bir e

dur

ger

ed

→ Pe

→ ag

b

→ E

b

→

e

→

Environment → Eylemi gerçekleştirebilen olan ortam diye bilirik.

State / Reward → AI (Ajan) → Action (Gerekenlerde yaptığı eylemler)

peziştirmeli Öğrenme Türleri

→ **Model tabanlı peziştirmeli öğrenme** : Ajan gerçeği ve olayları modeller için bir iç model kullanır. Model ajanın mevcut durumunu ve gelecekteki durumunu tahmin etmek için kullanılır. Çevrenin doğru bir şekilde modellenmediği gerektirir.

→ **Model tabanlı olmayan peziştirmeli öğrenme** : Ajan gerçeği ve olayları modellere dayanmadan doğrudan gözlemlerle öğrenir. Gerçek dünyada daha etkilidir. Öğrenmesi zordur.

→ **Değer tabanlı peziştirmeli öğrenme** : Ajanın her durumu için bir değer belirleyen bir değer fonksiyonu tahmin etmesi gerekir.

→ **Politika tabanlı peziştirmeli öğrenme** : Ajan doğrudan bir eylem seçmek için bir politika kullanır. Mevcut durumda en iyi sonuç için hangi eylemleri seçmesi gerektiğini belirtir (politika). Politikayı optimize eder.

★ **Peziştirmeli Öğrenme terimleri**

→ **Agent (Ajan)** : Çevreyi algılayan, keşfeden hareket eden bir varlıktır. Eylemler yaparak öğrenir (AI)

→ **Environment** : Ajanın ya diğer varlıkların içinde bulunduğu durum veya ortamdır.

→ **Action** : Bir ajan ortam içinde gerçekleştirdiği eylemlerdir. aksiyondur.

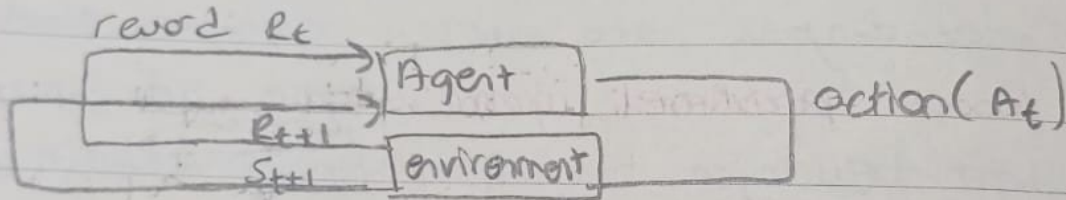
→ **State** : Ajanın yaptığı her şey sonucu ortamdan döndürülen sonuç, yonit, durumdur.

Reward (Ödül): İlla pozitif olmaz zorunda değil. Ajanın eylemini değerlendirmek için ajanın aldığı geri bildirimlerdir.

Policy: Ajanın mevcut duruma dayanarak bir sonraki eylemi için uyguladığı stratejidir.

Value: Bir eylemin beklenen uzun vadeli ödünün indirim faktörüyle hesaplanır. Kısa vadeli ödünün tersidir.

Q-Value: Genellikle normal değerdir ancak mevcut eylemi bir öz parametre olarak kullanılır. Birim yaptığımız eylemler sonucunda toplamış olduğumuz değerleri Q-Value olarak adlandırabiliriz.



State S_t

$R_{t+1} \rightarrow$ agent'in t+1'te yaptığı hareket sonucunda t+1'de gerçekleştiği duruma göre ifade eder.

Pekiştirmeli Öğrenme Algoritmaları

\rightarrow **Q-Learning** (Quality) \rightarrow Bir ortamda bir ajan en iyi eylemi öğrenmek için deneme-yanılma yöntemini kullanılır.

\rightarrow **Deep Q-Learning**: Sıhır öpücüğü söz konusudur.

En iyi eylemi sıhır öpücüğü ile bulmaya adanmış. Daha karmaşık ve daha karmaşık ortamlarda kullanılır.

\rightarrow Q-Learning de bir durumda olduğumuz action sonucunda value değeri bellidir.

Uygulama Alanları

\rightarrow Robotik, Kontrol, Oyun oynama, işletme, Verim

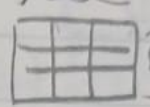
Q-Learning

• Bir ajan bir ortamda belirli eylemler gerçekleştirir ve eylemlerin sonuçlarını gözlemleyerek en uygun eylemleri öğrenir. Ajan her adımda ödeneceği bilgisi olmaya gerek yoktur.

Q öğrenme ödevi şeması:

Initialize Q-table

Action



state

0 ile başlayabilir.

↓

Choose an Action (Eğer table doluysa en yüksek değeri seç)

↓

Perform an Action (seçilen action davranışını uygula)

↓

measure Reward (Eylem uygulandıktan sonra ortamdan geri bildirim alıyoruz)

↓

Update Q-table (Q öğrenme işlemini gerçekleştirir)

↓

$$New\ Q(s,a) = Q(s,a) + \alpha \cdot [R(s,a) + \gamma \max_{a'} Q'(s',a') - Q(s,a)]$$

güncellenecek yeni değer

indirim oranı
öğrenme hızı
mevcut değer

s state'indeyken a action'a geçecek önceki değeri

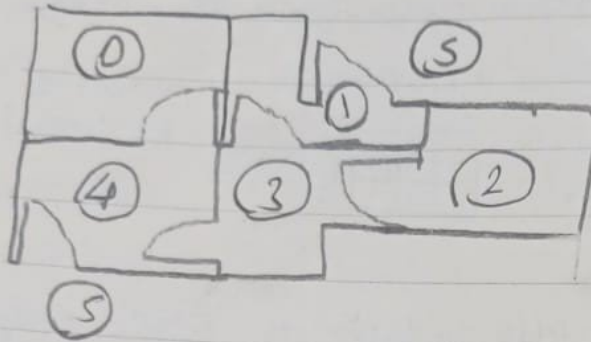
γ : kontrol stratejisi. Gelecekteki ödülleri bugünkü değeri temsil eder. $[0, 1]$ 'ıai şimdi elimizde olan ödülle mi ödülleri bekliyoruz. gelecekte gelecekteki ödülle mi ödülleri bekliyoruz

$\max_{a'} Q'(s',a')$ = Bir sonraki adıma gidecek olan adan sonra karşılaşılabilecek en yüksek değeri belirler

Formül gerektirir iken sodelestirelim $\alpha = 1$ olarak

new $Q(s,a) = R(s,a) + \gamma \max_{a'} Q'(s',a')$ all actionlardan en büyükü

Burum Calismasi



state = {0, 1, 2, 3, 4, 5}
action = {0, 1, 2, 3, 4, 5}

ör: 0 numaradan 4 numaraya geceriz iken 4 numarada action kulluyor.

Reward table

R \ A	0	1	2	3	4	5
0	-1	-1	-1	-1	0	-1
1	-1	-1	-1	0	-1	100
2	-1	-1	-1	0	-1	-1
3	-1	0	0	-1	0	-1
4	0	-1	-1	0	0	100
5	-1	0	-1	-1	0	100

-1 : Bu action secilmesin diye verilen negatif ödül
0 : 1 den daha iyidir. Pozitif reward.

Ara 1 numaralı 0 daha beslesin.

Q table

	0	1	2	3	4	5
0					80	
1				64		100
2				64		
3		80	51		80	
4	64			64		100
5		80			80	100

ilk besta biter
hepsi 0 kullu etm

$Q(1,5) = R(1,5) + 0.8 \cdot \max_{a'} Q(1,5) = 100$

state action 100 seçili

Agen 3 numaradan başlasın.

Geciklim o kadar fazla
gidebileceğim yerler.

$$Q(3,1) = R(3,1) + 0.8 \max \begin{bmatrix} Q(1,5) \\ Q(1,3) \end{bmatrix}$$

3. O kadar 1. o kadar

geçersen

$$Q(3,1) = 0 + 0.8 \cdot 100 = 80$$

Agen 2 den başlasın

$$Q(2,3) = \underbrace{R(2,3)}_0 + 0.8 \max \begin{bmatrix} Q(3,4), Q(3,2) \\ Q(3,1) \\ Q(3,3) \end{bmatrix} = \underline{\underline{64}}$$

80