## توابع التشابه:

في مجال Q-Learning لأنظمة التوصية، تلعب وظيفة التشابه دورًا رئيسيًا في تعريف فضاء الحالة. تحدد هذه الوظيفة مدى تشابه الحالات المختلفة (مجموعات المستخدم والفيلم) مع بعضها البعض. يؤثر هذا التشابه على التوازن بين الاستكشاف خلال عملية التعلم.

فيما يلي بعض توابع التشابه الشائعة الاستخدام في Q-Learning للتوصية:

**1. توابع تشابه البنود (Item-based Similarity):**

* يركز هذا النهج على التشابه بين الأفلام بناءً على تقييمات المستخدمين المختلفين لها مثل **تشابه التمام (Cosine Similarity)**، **معامل ارتباط بيرسون (Pearson Correlation Coefficient)**

**2** **. تابع تشابه المستخدمين (User-based Similarity):**

* يركز هذا النهج على التشابه بين المستخدمين بناءً على تقييماتهم لأفلام مختلفة.
  + **تشابه جاكار (Jaccard Similarity):** يقيس هذا المعيار نسبة عدد الأفلام التي قام المستخدمان بتقييمها بشكل مشابه (أعجبهم أو كرهوها) إلى إجمالي عدد الأفلام التي قام بتقييمها على الأقل مستخدم واحد. يحدد هذا التشابه المستخدمين الذين لديهم تفضيلات متداخلة للأفلام.
  + **أقصى زيادة (MaxInc)** هو معيار يُستخدم لتقييم أداء أنظمة التوصية. يعتمد على فكرة تعظيم عدد العناصر الموصى بها بشكل صحيح مع تقليل عدد العناصر الموصى بها بشكل غير صحيح.
  + **معامل ديس:(Dicecorffeccent)** المعروف أيضًا باسم **معامل تشابه ديس**، هو معيار آخر يُستخدم لتقييم أداء أنظمة التوصية. يقيس التشابه بين مجموعتين من العناصر، مثل مجموعة من العناصر الموصى بها ومجموعة من العناصر ذات الصلة بالأرض الحقيقية.

**وفي هذا الجزء قمنا بتجريب توالع تشابه المستخدمين وسنتحدث عن كل تابع بالتفصيل**

### تشابه جاكار:

**معامل جاكار**، المعروف أيضًا باسم **معامل تشابه جاكار**، هو مقياس يُستخدم لقياس التشابه بين مجموعتين من العناصر. يُستخدم بشكل شائع في سياق استرجاع المعلومات وأنظمة التوصية لقياس التشابه بين مجموعات العناصر أو المستندات.

**حساب معامل جاكار:**

يتم حساب معامل جاكار باستخدام الصيغة التالية:

حيث:

* يمثل عدد العناصر المشتركة بين كل من المجموعة A والمجموعة B (التقاطع).
* يمثل العدد الإجمالي للعناصر في كل من المجموعة A والمجموعة B (الاتحاد).

**تفسير معامل جاكار:**

* **معامل جاكار أعلى:** يشير معامل جاكار أعلى إلى أن المجموعتين لهما درجة تداخل أو تشابه أعلى. بعبارة أخرى، يتشاركان نسبة أكبر من العناصر المشتركة.
* **معامل جاكار أدنى:** يشير معامل جاكار أدنى إلى أن المجموعتين لهما درجة تداخل أو تشابه أقل. بعبارة أخرى، يتشاركان نسبة أصغر من العناصر المشتركة.
* **القيم المتطرفة:**
  + **معامل جاكار = 1:** عندما يكون معامل جاكار 1، فهذا يعني أن المجموعتين متطابقتان. يتشاركان جميع العناصر نفسها ولا يحتويان على عناصر فريدة.
  + **معامل جاكار = 0:** عندما يكون معامل جاكار 0، فهذا يعني أن المجموعتين لا تحتويان على أي عناصر مشتركة. لا يتشاركان أي عناصر ويختلفان تمامًا.

كما كنا نعلم سابقا اننا نبدأ من قيمة عشوائية في شبكة qtable فنحن نتوقع هنا نتائج مختلفة للمكافئة مع اختلاف عملية تنفيذ البرنامج والصورتين الموضحتين في الشكل 1والشكل2. ولعل اهم ما استتنتجناه من الشكلين هو:

**اختلاف نهج المكافأة الأول (عينة عشوائية للحالة الابتدائية)**

كما توقعنا، يؤدي اختيار الحالة الابتدائية العشوائي في كل تشغيل لوظيفة تعلم Q إلى اختلاف في المكافآت لكل حلقة. يحدث هذا لأن تسلسل الحالات والإجراءات التي يواجهها العامل agent يختلف بناءً على النقطة التي يبدأ منها.

* **استكشاف مقابل استغلال:** تلعب الحالة الابتدائية دورًا في عملية الموازنة بين الاستكشاف والاستغلال. في البداية، قد يركز العامل على الاستكشاف (تجربة إجراءات جديدة) بدلاً من الاستغلال (اختيار الإجراء ذي أعلى مكافأة متوقعة بناءً على المعرفة الحالية). تساعد الحالات الابتدائية العشوائية في الاستكشاف عن طريق إجبار العامل على تجربة أجزاء مختلفة من البيئة.
* **تباين في مسار التعلم:** نظرًا لأن نقطة البداية عشوائية، فإن تسلسل الحالات والإجراءات التي يواجهها العامل ستختلف عبر التكرارات. يمكن أن يؤدي هذا التباين إلى اختلاف في قيم المكافأة لكل حلقة خلال عملية التعلم.

A graph with blue and orange lines

Description automatically generated

الشكل 1مكافأة تعتمد على معامل جاكار لكل حلقة خلال تعلم (تجربة اولى)

A graph with a line going up

Description automatically generated

الشكل 2 مكافأة تعتمد على معامل جاكار لكل حلقة خلال تعلم (تجربة ثانية)

في تعلم Q، يوضح رسم متوسط المكافأة لكل حلقة أداء العامل بمرور الوقت خلال عملية التدريب. تمثل كل حلقة تفاعلًا كاملاً بين العامل والبيئة، وتمثل المكافأة المكافأة التراكمية التي يتلقاها العامل لإجراءاته داخل تلك الحلقة.

**محاور الرسم البياني:**

* يمثل المحور X رقم الحلقة، ويمثل المحور Y متوسط المكافأة المحققة عبر جميع الحلقات حتى تلك النقطة.
* **سرعة التعلم:** يعكس معدل زيادة متوسط المكافأة سرعة اكتساب العامل للمعرفة حول البيئة.
* فيما يلي تفصيل لما قد يكشفه الرسم البياني:

**اتجاه صاعد (**الشكل 2**):**

* يشير الاتجاه التصاعدي العام في متوسط المكافأة إلى أن العامل يتعلم بشكل فعال. إنه يكتشف تدريجيًا إجراءات تؤدي إلى مكافآت أعلى على المدى الطويل.
* **استقرار الأداء:** يشير الرسم البياني ذو النمو الإيجابي المستمر إلى أن العامل يتعلم استراتيجية مستقرة. قد تشير التقلبات أو الانخفاضات الحادة إلى الحاجة إلى إجراء تعديلات في خوارزمية التعلم أو استراتيجية الاستكشاف.

**استقرارات أو تقلبات (**الشكل 1**) :**

* يظهر الرسم البياني فترات يركد فيها متوسط المكافأة أو يتقلب. يمكن أن يكون هذا بسبب مواجهة العامل لأجزاء صعبة من البيئة أو حاجته إلى مزيد من الاستكشاف للعثور على استراتيجيات مثالية.
* **التقارب:** عندما وصل متوسط المكافأة إلى ذروة واستقر، فقد يعني ذلك أن العامل قد تقارب على سياسة مثالية أو شبه مثالية للمهمة المحددة وهذا موضح في الشكل 2.

هناك جانب اخر لا يقل اهمية عن متوسط المكافئة قيم Q وتوزيعها في التعزيز الآلي:

يعد **جدول Q** أحد هياكل البيانات الأساسية في تعلم Q. يخزن هذا الجدول **قيم Q**، والتي تمثل المكافأة المستقبلية المتوقعة التي يمكن للوكيل توقعها من خلال اتخاذ إجراء معين في حالة معينة.

* **الحالة:** تمثل الحالة الوضع الحالي الذي يوجد فيه العامل داخل البيئة وهذه المكافئة تمثل **مكافأة على المدى الطويل أي** لا تأخذ قيمة Q فقط المكافأة الفورية التي يتم تلقيها لاتخاذ الإجراء، ولكن أيضًا المكافآت المستقبلية التي يتوقعها العامل من خلال اتباع أفضل مسار للعمل بعد ذلك.
* **السياسة المثلى:** تشير السياسة المثلى إلى تسلسل الإجراءات التي تؤدي إلى أعلى مكافأة تراكمية ممكنة بمرور الوقت.

**تفسير توزيع قيم Q:**

يمكن أن يوفر توزيع قيم Q في جدول Q رؤى حول تقدم تعلم العامل. فيما يلي تفصيل لبعض التفسيرات الشائعة:

* **توزيع موحد:** إذا تم توزيع جميع قيم Q بالتساوي عبر النطاق (على سبيل المثال، من -1 إلى +1)، فهذا يعني أن العامل لم يتعلم الكثير بعد، وأن قيم Q عشوائية في الغالب.
* **توزيع منحرف نحو القيم الإيجابية:** كما في الصورة التي قدمتها، يشير التوزيع الذي يميل نحو قيم Q الإيجابية إلى أن العامل تعلم أن اتخاذ إجراءات يؤدي بشكل عام إلى مكافآت إيجابية على المدى الطويل.
* **توزيع منحرف نحو القيم السلبية:** يشير هذا السيناريو إلى العكس، حيث تعلم العامل أن معظم الإجراءات تؤدي إلى عقوبات.
* **توزيع ذروة حول قيمة معينة:** قد يشير هذا إلى أن العامل قد تقارب على مجموعة من الإجراءات المتوقع أن تحقق مكافآت معتدلة، ولكنها قد لا تكون السياسة المثلى بعد وهذا ما نلاحظه في لشكل 3 حيث أن قيم المكافأت النستقبلية التراكمية تتجمع حول القيمتين 0.4 و -0.3ا

A graph of a distribution of q values

Description automatically generated

الشكل 3 توزيع قيم المكافأت النستقبلية التراكمية

### تشابه MaxInc

من المثير للاهتمام أن "MaxInc" ليس مصطلحًا دارجا في تعلم Q، ولكن من المحتمل أنه تابع تشابه مخصصة تستخدم في تطبيق معين.

**تفسير لـ MaxInc كتابع تشابه:**

* **التركيز على أقصى زيادة:** يشير جزء "MaxInc" إلى أن الدالة تعطي الأولوية لأكبر زيادة في قيم Q عند مقارنة حالتين.

**كيف يمكن أن تعمل:**

1. **حساب فرق قيمة Q:** بالنظر إلى حالتين (s1 و s2)، يحسب MaxInc الفرق المطلق بين قيم Q المقابلة لكل إجراء ممكن (a).
2. **تحديد الحد الأقصى للزيادة:** لكل إجراء، سيحدد بعد ذلك الحد الأقصى للزيادة في قيمة Q بين الحالتين.
3. **وهذه صيغته الرياضية:**

A graph with blue and orange lines

Description automatically generated

الشكل 4مكافأة تعتمد على معامل أقصى زيادة لكل حلقة خلال تعلم

A graph with blue squares

Description automatically generated

الشكل 5 توزيع قيم المكافأت النستقبلية التراكمية

من ملاحظة الشكلين أعلاه نلاحظ انحدار قيمة المكافأة بعد حلقة معينة وكانت النتيجة مشابهة عند اختلاف الجالات الابتدائية لذلك يمكننا القول ان مسار التعلم في حالة تابع جاكار اكثر استقرارا ولكن ليس بالضرورى أكثر دقة وهاذا ما سيتم توضيحه لاحقا.

### معامل ديس (معامل تشابه ديس):

معامل ديس، المعروف أيضًا باسم معامل تشابه ديس، هو مقياس إحصائي يُستخدم لتقييم أداء أنظمة التوصية. يقيس مدى تشابه مجموعتين من العناصر، مثل مجموعة العناصر الموصى بها ومجموعة العناصر ذات الصلة بالأرض الحقيقية.

لحساب معامل ديس، يتم اتباع الخطوات التالية:

1. **إنشاء جدول مطابقة:** يتم إنشاء جدول يوضح تواجد العناصر في كل من المجموعتين. يتكون الجدول من عمودين أحدهما للعناصر الموصى بها والآخر للعناصر ذات الصلة بالأرض الحقيقية. يتم وضع علامة "X" في الخلية إذا كان العنصر موجودًا في المجموعة المقابلة، وعلامة "-" إذا لم يكن موجودًا.
2. **حساب عدد العناصر المشتركة:** يتم حساب عدد العناصر الموجودة في كلتا المجموعتين.
3. **حساب عدد العناصر في كل مجموعة:** يتم حساب عدد العناصر في كل من المجموعتين بشكل منفصل.
4. **حساب معامل ديس:** يتم استخدام الصيغة التالية لحساب معامل ديس:

**تفسير معامل ديس:**

* **القيمة 1:** تشير إلى أن المجموعتين متماثلتان تمامًا، حيث تحتوي كلتا المجموعتين على نفس العناصر.
* **القيمة 0:** تشير إلى أن المجموعتين مختلفتان تمامًا، حيث لا توجد عناصر مشتركة بينهما.
* **القيم بين 0 و 1:** تشير إلى درجات متفاوتة من التشابه بين المجموعتين. كلما زادت القيمة، زاد التشابه..

**عيوب معامل ديس:**

* **عدم مراعاة أهمية العناصر:** لا يأخذ معامل ديس في الاعتبار أهمية العناصر الفردية. على سبيل المثال، قد يكون نظام التوصية الذي يوصي بعنصرين مهمين وعشرين عنصرًا غير مهم أفضل من نظام يوصي بعشرين عنصرًا غير مهم، ولكن معامل ديس سيكون متماثلًا لكلا النظامين.
* **التأثر بحجم المجموعات:** قد يتأثر معامل ديس بحجم المجموعات. على سبيل المثال، قد يكون نظام التوصية الذي يوصي بعنصرين من مجموعة صغيرة أفضل من نظام يوصي بعشرين عنصرًا من مجموعة كبيرة، ولكن معامل ديس سيكون أكبر للنظام الأخير.

A graph with blue and orange lines

Description automatically generated

الشكل 6 مكافأة تعتمد على معامل ديس لكل حلقة خلال تعلم

A graph of a number of values

Description automatically generated with medium confidence

الشكل 7 توزيع قيم المكافأت النستقبلية التراكمية

**النقاط الرئيسية وملاحظات حول أداء معامل ديس**

* **اتجاه زيادة المكافأة:** يبدو أن الاتجاه الأولي لزيادة المكافآت قد تلاشى ثم وصل القعر عند الحلقة 85. على عكس الزيادة المستقرة التي لوحظت في نابع جاكاردز
* **زيادة المكافأة النهائية:** على الرغم من الانحدار السلبي، هناك زيادة نهائية في المكافأة، حيث تصل إلى قيمة عالية.
* **عدم التناسق:** على الرغم من أن المكافأة النهائية إيجابية، إلا أن تناسق زيادة المكافأة بشكل عام أقل مقارنة بتابع جاكارد في معظم أنحاء الرسم البياني.
* **تركيز التوزيع:** يشير مخطط التوزيع التركيز على مستوى مكافأة مستقر ولا يتجمع بشكل واضح خول قيمة معينة وهذا غير جيد بالنسبة للعميل.
* قد يعلق العامل في "الاقصى المحلي" - وهي حالة يحصل فيها على مكافأة مناسبة ولكنه يواجه صعوبة في الاستكشاف أكثر واكتشاف مسار نحو مكافآت أعلى.
* قد لا يكون هيكل المكافأة أو خوارزمية التعلم مناسبة للمهمة، مما يؤدي إلى سلوك التعلم غير المتناسق هذا.

بشكل عام، يسلط التحليل الضوء على مشكلة محتملة في عملية تعلم العامل. يشير اتجاه المكافأة المستقر والتركيز على مستوى مكافأة معين إلى قيود في الاستكشاف أو هيكل المكافأة نفسه. قد تكون هناك حاجة إلى مزيد من التحقيق والتعديلات المحتملة على خوارزمية التعلم أو تصميم المكافأة لتحقيق مسار تعلم أكثر تناسقًا وأمًاثلية.