**Assignment-5**

1. صح أو خطأ مع شرح الإجابة.
2. الافتراض الأساسي للنموذج (Naïve Bayes) هو أن كل ميزة مستقلة.

صحيح. يفترض (Naïve Bayes) أن الميزات المستخدمة لوصف المراقبة هي مستقلة بناءً على الفئة. على الرغم من هذا الافتراض يؤدي نموذج (Naïve Bayes) غالبًا إلى أداء جيد في الممارسة.

1. نظرية (Naïve Bayes) توفر وسيلة لحساب الاحتمال الشرطي P(A|B) من P(B) و P(A) وP(B|A).

صحيح. نظرية (Naïve Bayes) هي نظرية أساسية في الاحتمالات تصف احتمال حدوث حدث، بناءً على المعرفة السابقة بالشروط التي قد تكون متعلقة بالحدث. الصيغة هي P(A|B) = [ (P(B|A) \* P(A)) \ P(B) ] ،

حيث:

P(A|B) هو الاحتمال الشرطي، P(B|A) هو الاحتمالية،

P(A) هو الاحتمال السابق) احتمال وقوع الحدث قبل رؤية الدليل)، وP(B) هو الدليل.

1. لديك P(B|A) هل تستطيع حساب P(A|B).

خطأ. لحساب P(A|B)، تحتاج إلى معرفة P(A) (الاحتمال السابق لـ A) وP(B) (الدليل).

وجود P(B|A) وحده لا يكفي.

1. إذا كان لديك نموذج ذو دقة عالية (High Accuracy) على مجموعة البيانات التدريبية، يمكنك التأكد بنسبة 100٪ أن نموذجك جيد للغاية.

خطأ. الدقة العالية على مجموعة البيانات التدريبية لا تضمن أن يؤدي النموذج بشكل جيد على بيانات جديدة غير مرئية (Generalization). يُعتبر Overfitting هو مشكلة شائعة حيث يتعلم النموذج البيانات التدريبية بشكل جيد لكنه يفشل في التعميم إلى بيانات جديدة. يمكنك استعمال Precision and recall ومن ثم F1-score.

1. كيف يعمل (Naïve Bayes)؟

وتعتبر خوارزمية تعلم تستخدم لمهام التصنيف، مثل تصنيف النصوص. تُستخدم لحساب الاحتمالات الشرطية.

الاحتمال الشرطي هو قياس لاحتمال حدوث حدث ما عندما يحدث حدث آخر.

الفرضية الساذجة: تفترض أن كل ميزة تجعل تأثيرًا مستقلًا من الآخرين وتعطى نفس التأثير.

يتم اختيار فئة التصنيف بناءً على أعلى احتمال شرطي للفئة.

النظرية وراء خوارزميات البايز الساذج

يعتمد البايز الساذج على نظرية بايز الاحتمالية.

حساب الاحتمالات:

حساب احتمالات الفئات السابقة والاحتمالات الشرطية للميزات.

المرحلة التصنيفية:

تنبؤ الفئة:

لكل فئة، حساب احتمالية شرطية جديدة باستخدام المعلومات السابقة.

اختيار الفئة بأعلى احتمال شرطي كفئة تصنيف.

1. إذا كانت لدينا مجموعة البيانات التالية، فقم بتطبيق Naive Bayes للتنبؤ بالميزات الجديدة

X={color=orange, Calories=2, Tall=long, Tasty=Bad}

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fruit | Tasty | Tall | Calories | Color |
| Y | Good | Short | 3 | Yellow |
| Y | Bad | Long | 2 | Orange |
| Y | Good | Short | 3 | Orange |
| Y | Good | Short | 3 | Yellow |
| N | Bad | Short | 2 | Orange |
| N | Bad | Long | 2 | Yellow |
| N | Bad | Long | 2 | Yellow |
| N | Good | Short | 2 | Yellow |

|  |  |  |
| --- | --- | --- |
| N | Y | Color |
| 3/5 | 2/5 | Yellow |
| 1/3 | 2/3 | Orange |

الحل:

Y = 4/8 = 0.5

N = 4/8 = 0.5

|  |  |  |
| --- | --- | --- |
| N | Y | Calories |
| 4/5 | 1/5 | 2 |
| 0/3 | 3/3 | 3 |

|  |  |  |
| --- | --- | --- |
| N | Y | Tasty |
| 3/4 | 1/4 | Bad |
| 1/4 | 3/4 | Good |

|  |  |  |
| --- | --- | --- |
| N | Y | Tall |
| 2/5 | 3/5 | Short |
| 2/3 | 1/3 | Long |

لحساب X = {color=orange, Calories=2, Tall=long, Tasty=Bad}

P (Fruit | X) = [ P(Y) \* P(color=orange) \* P(Calories=2) \* P(Tall=long) \* P(Tasty=Bad)]

P (Fruit | X) = [ 0.5 \* 2/3 \* 1/5 \* 1/3 \* 1/4 ]

P (Fruit | X) = 0.0055

P (NotFruit | X) = [ P(Y) \* P(color=orange) \* P(Calories=2) \* P(Tall=long) \* P(Tasty=Bad)]

P (NotFruit | X) = [ 0.5 \* 1/3 \* 4/5 \* 2/3 \* 3/4 ]

P (NotFruit | X) = 0.0066

0.0066 > 0.0055

NotFruit > Fruit

إذا الإحتمال الأكبر هوNotFruit

1. ما هي الفكرة الأساسية وراء SVM؟ ما هو SVM؟

الفكرة الأساسية وراء (Support Vector Machines - SVM) هي البحث عن حدود فاصلة (margin) بين فئتين مختلفتين في البيانات.

الهامش هو المسافة بين حدود القرار وأقرب نقطة من كل فئة في (Dataset).

SVM يحاول تحقيق أكبر هامش (Maximum margin) ممكن.

SVM: هو خوارزمية في مجال تعلم الآلي تحت الاشراف تُستخدم لمهام التصنيف والتحليل التنبؤي. تم تطوير SVM للبحث عن حدود فاصلة بين مجموعات من البيانات، تستعمل اكتر في المهام ذات الأبعاد العالية.

الخصائص الرئيسية SVM:

1. يستخدم للتصنيف حيث يقوم بفصل البيانات إلى فئتين أو أكثر بناءً على الخصائص المستخدمة.
2. العثور على حدود قرار (decision boundary) بين مجموعات البيانات يكون لديها أكبر هامش (margin) ممكن.
3. الهامش هو المسافة بين حدود القرار وأقرب نقطة من كل فئة. يتم تحسين SVM لتحقيق أكبر هامش ممكن.
4. النواقل هي النقاط التي تقع على حدود القرار أو قريبة منها، وتلعب دورًا حاسمًا في تحديد حدود القرار.
5. التعامل مع البيانات التي لا يمكن تصنيفها بشكل خطي، بفضل استخدام وسائط (kernels) لتحويل البيانات إلى فضاء ذو أبعاد أعلى.
6. تظهر فعالية في التعامل مع مجموعات البيانات التي تحتوي على العديد من (Features) أو الأبعاد.
7. اشرح المعلمات التي تحتاج إلى تغييرها او تكييفها فيSVM ؟

بعض parameters التي تحتاج الى تغييرها في خوارزمية SVM:

1. معلمةC :

* تُمثل معلمة C قيمة تكلفة التصحيح. إذا كانت قيمة C عالية، فإن النموذج سيحاول تصحيح كل الأخطاء في التصنيف، مما قد يؤدي إلى نموذج معقد. ومن الجيد استخدام قيمة C عالية عندما تكون البيانات معقدة.

1. نوع الكيرنل (Kernel Type):

* يُستخدم الكيرنل لتحويل البيانات إلى فضاء ذو أبعاد أعلى. يمكنك اختيار نوع الكيرنل الذي يناسب نوع البيانات الخاص بك، مثل الكيرنل **linear**، وكيرنل **RBF**، وكيرنل **Polynomial**.

1. معلمة Gamma:

* تحكم معلمة Gamma في شكل الكيرنل. قيم عالية لـ Gamma تعني أن النقاط البعيدة يتم منحها أوزان أقل، مما يؤثر أكثر على حدود القرار. قيم منخفضة لـ Gamma تعني أن النقاط البعيدة يتم منحها أوزان أكبر.

ملاحظة: تعديل هذه المعلمات يعتمد على البيانات وطبيعة المشكلة. يُفضل استخدام تقنيات التحقق (Cross-Validation) لاختيار القيم المناسبة لهذه المعلمات او عن طريق (Tuning the hyper-parameters of an estimator).

1. افترض أن نموذج LR الخاص بنا يتدرب على مجموعتين من بيانات الميزات (x1، x2) ويخرج بالمعادلة التالية:f(x)=z= ax1+bx2+c، حيث a=3، وb=15 وc =-15، يتنبأ النموذج بـ y=1 إذا كانت z>=0 ويتنبأ بخلاف ذلك بـ 0.

أ. ارسم بعض البيانات وارسم المنطقة التي يتنبأ فيها النموذج بالرقمين 1 و9.

ب. ما هو التنبؤ بنموذج x1=-2، x2=2.

1. افترض أن لدينا مجموعة بيانات وقمنا برسم/تصور مجموعة البيانات كما هو موضح أدناه. اشرح ما هي الخوارزميات (LR، و/أو LogisticRegression، و/أو SVM، و/أو Naive Bayes) التي تقترحها لاستخدام مجموعة البيانات واشرح إجابتك؟ اشرح أيضًا أي الخوارزميات ستعمل بشكل أفضل على مجموعة البيانات