Naïve Bayes

التعلم باشراف – تصنيف

تستند الى بايز الإحصائية, يفترض أن وجود ميزة معينة في صنف معين غير مرتبط بوجود أي ميزة أخرى.

من السهل بناء نموذج و مفيد بشكل خاص لمجموعات البيانات الكبيرة جدا. الى جانب البساطة, من المعروف أنها تتفوق حتى على أساليب التصنيف المتطورة للغاية.

يطلق عليها هذا الاسم لانه يتم تبسيط حساب الاحتمالات الخاصة بكل فرضية لجعل حسابها قابلا للتتبع

ويعمل بشكل رئيسي على مبدأ الاحتمالات

**ومن اهم مزاياها :**

أنها تتطلب كمية صغيرة من بيانات التدريب لتقدير المعلمات اللازمة.

أشهر أستخدامها تصنيف المستندات وتصفية البريد العشوائي

و من ايجابياتها : انه سهل الفهم والبناء نسبيا و سهلة التدريب, حتى مع مجموعة بيانات صغيرة وسريعة.

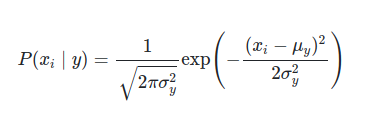
وسلبياتها: يفترض ان كل ميزة مستقلة وليست الحالة دائما في البيانات ,

يعمل على تصنيف وغير قوي في التوقع, و يرتبط ببعض الفرضيات.

**أنواع مصنف بايز**

1 - gaussian :

بتم توزيع القيم المستمرة المرتبطة بكل القيم وفقا للتوزيع الغوسية ويتم تقديرها باستخدام أقصى احتمال.

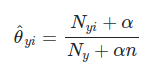


2 - Multinomial Naive Bayes :

تستخدم مع البيانات المنفصلة و متعددة الخرج

هنا لدينا أشعة vector الميزات هي عبارة عن " تكرارات" أو (fraquencies) للأحداث التي تتولد باستخدام التوزيع ويسمى متعدد الحدود.

يستخدم هذا النوع عادة" في تصنيف الوثائق.



3 - Bernouli Naive Bayes :

خوارزمية التصنيف الثنائي

تتطلب هذه الفئة أن يتم تمثيل العينات كمتجهات مميزة ذات قيمة ثنائية, قد يقوم بتكوين ثنائي المدخلات وذلك اعتماد على المعلمات

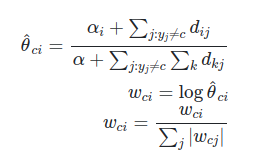


Complement Naive Bayes :

تعمل خوارزمية CNB على تلائم (تكييف) خوارزمية ال MNB متعددى الحدود بشكل خاص لمجموعة بيانات غير متوازنة , يستخدم CNB إحصائيات من مكمل كل فئة لحساب أوزان النموذج, وتبين ان تقدير المعلمات ل CNB أكثر استقرار من لMNBويتغلب عليها ايضا في مهام التنصيف النصي

وصمم لتصحيح الافتراضات "المعلمات" التي وضعها المصنف القياسي MNB

وهي مناسبة بشكل خاص لمجموعة البيانات غير المتوازنة



Categorical Naive Bayes

تتعامل مع الخرج الفئوي وشبيهى بعمل ال MNB اذ انها تعمل بشكل جيد مع التصنيف الثنائي



وهنا المصدر يوضح آلية عمل القوانين وتوضيح كل متحول بها

<https://scikit-learn.org/stable/modules/naive_bayes.html>

<https://www.kaggle.com/blackblitz/gaussian-naive-bayes>

<https://github.com/ritu-19/Machine-Learning-Projects/tree/ab61174747dcd06bef2497201ff8251582ad26e3/Spam_Classifier>