**Customer Sentiment Analysis**

**المتقدم : حمزة عبدالعظيم حاج اصطيفي**

**Task 1**

في البداية سوف أقوم بشرح ووصف ما فهمته من ال task المقدم لي :

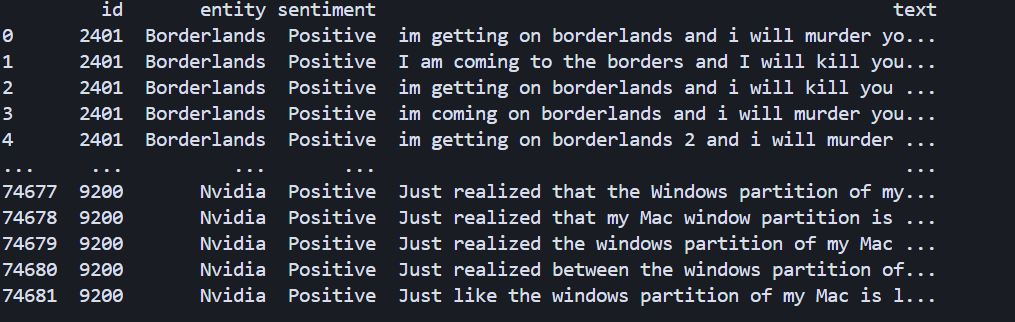
حسب الشرح انه يجب ان أقوم بتنفيذ تطبيق يقوم بتصنيف التعليقات التي تكتب   
في بالتعليقات إلى negative و positive و neutral حيث من خلالها نقيم مدى أداء المنتجات التي يجري عليها التقييم و ما سوف نقوم بعد هذه الخطوة .

في البداية لدي القاعدة التالية  
  
**حيث تحتوي على التالي :**

قرابة ال 74682 سجل حيث كل سجل يحوي الصفات التالية   
1- ID وهو عنوان للرسالة او التعليق المكتوب قيم مستمرة   
2- لدينا ال sentiment الذي يحمل قيمة التصنيف وهوclass label .

3- لدينا entity بالكاتب التعليق .

4- بالإضافة لل text المكتوب من قبله

وهذه صورة عنها :

من القاعدة التالية يتوضح سير عمل التالي :

* حيث سوف نقوم بأخذ كل من text , sentiment للقيام بتحليل ومعرفة التصنيف الصحيح ولقيام ذلك نقوم بالخطوات الأساسية :

الآن نقوم بجعل البيانات التي لدينا نظيفة لتناسب التحليل وتسهل العمل .  
حيث نقوم بالتالي حسب ما وارد بالكود :

حيث اولاً لتسهيل العمل وبما انه قمنا بجعل التنفيذ على خوارزميتين الأولى (naïve bayes) و الثانية (svm) حيث بإنشاء class يتم وضع فيه جميع ما يلزم كخطوة أولى للنموذج من تنظيف واعداد متغيرات الكلاس وغيرها من الأمور تقسيم . كما هو في الكود من السطر (30-56)

حيث استخدمت دوال خاصة بمكتبة ال (sklearn) منها :

train\_test\_split تقوم بتقسيم البيانات التي لدي

TfidfVectorizer تساعد في عملية تصنيف اسهل حيث تقوم بتحويل النص لسلة من الأرقام وقوم بتعلم وتحويلها الي ارقام واخذ مجموعة من الكلمات التي تكررت اكثر حسب ما يدخل في max\_feature انا خدت الأنسب الي 5000 كلمة .

accuracy\_score من شان معرفة دقة النموذج ورح نعرف من خلال الواجهة المشغولة .

confusion\_matrix مصفوفة الادراك تم عملها من خلال heatmap

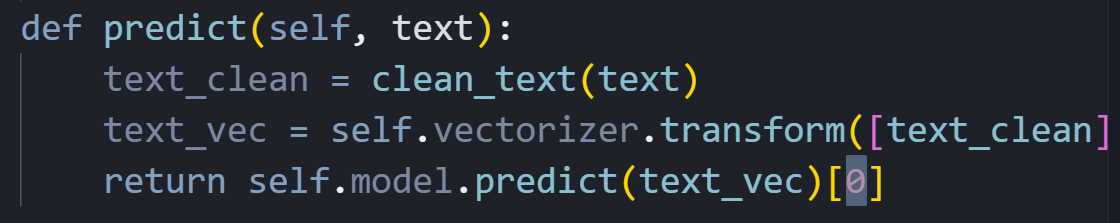
ولبناء النموذجين السابقين توفر لنا المكتبة نعمل object منهم كل من الخوارزميتين :

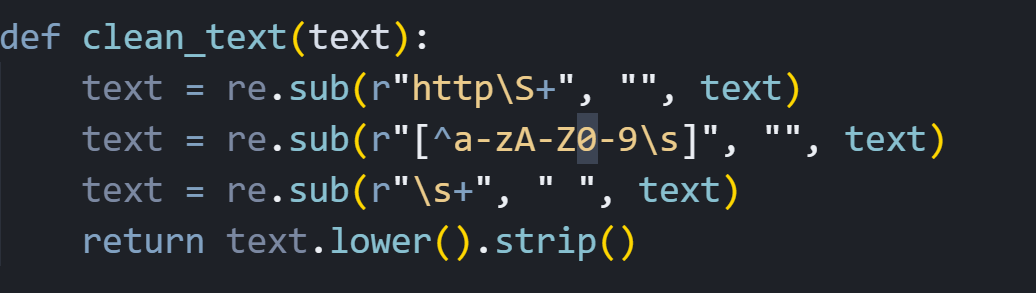
LinearSVC .

MultinomialNB .

تم بناء دالتين تساعد في تنظيف النص و دالة للتوقع .

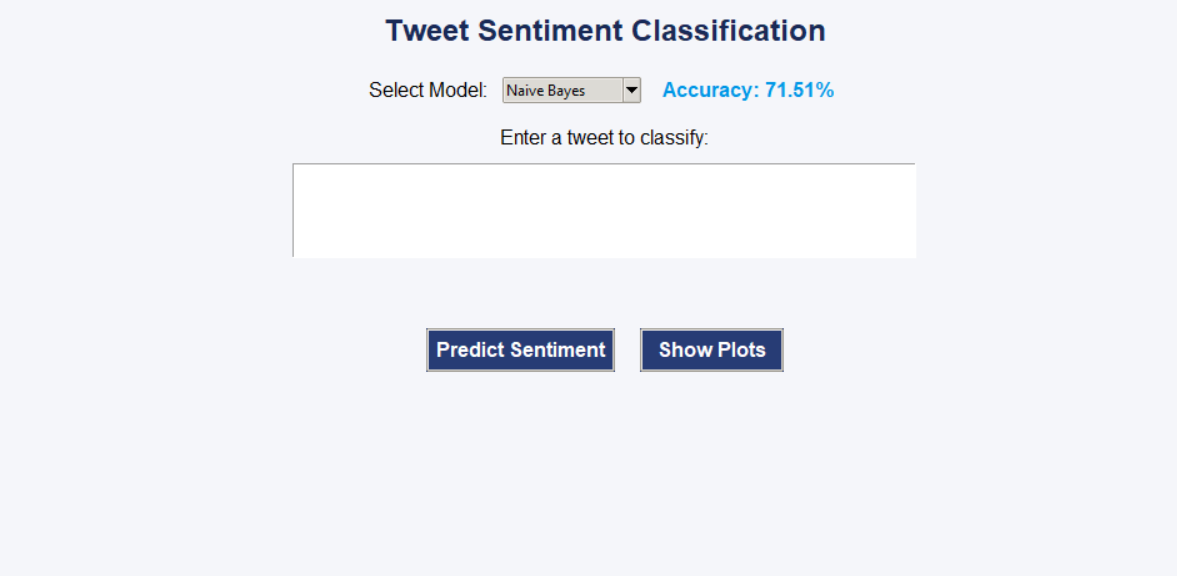
موجودات في الكود



الدالة الثانية : 

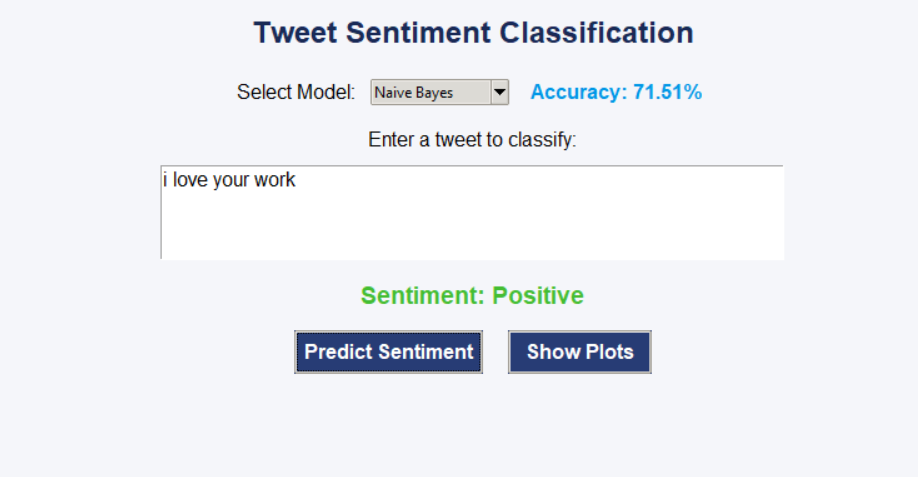
قمت ببناء class يقوم باظهار واجه تفاعلية للقيام بإدخال نص والتوقع بتصنيفه حسب ما تم تدريب به النموذج المختار .

- حيث يتم اختيار الخوارزمية المتبعه   
 - ندخل التعليق ونضغط على زر predict Sentiment

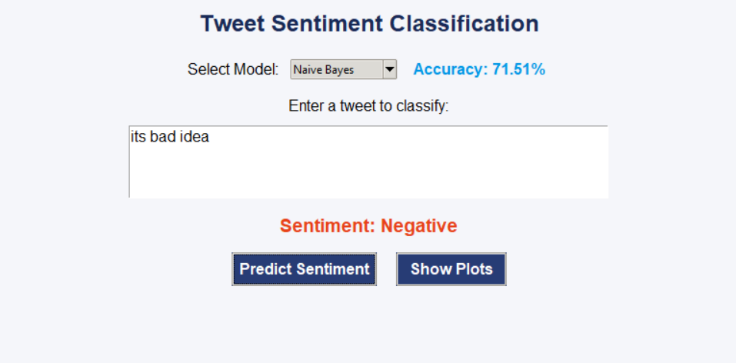
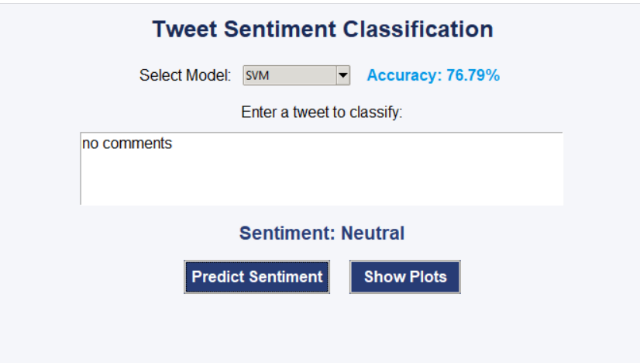
الواجه :

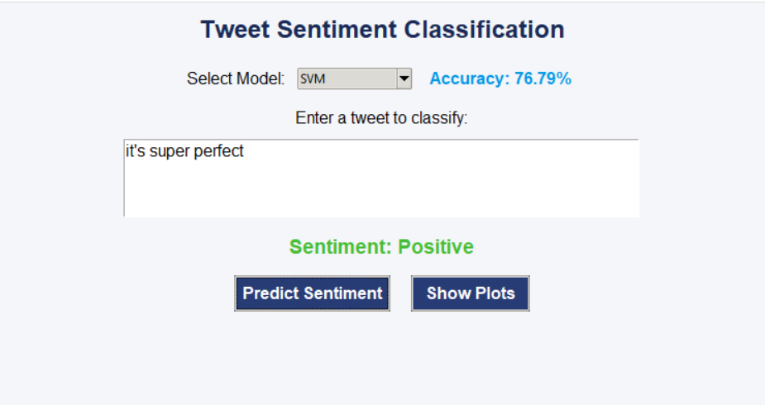
اذا قمنا بإدخال النص التالي :   
 I love your work

نجد النتيجة انها ...هي positive كما في الشكل التالي :



وكل شيء موضح بالواجه ,بعض التجارب مع تعليقات مختلفة في كلا الخوارزميتين :





**المكاتب المستخدمة :**

لدينا المكاتب التالية :  
tkinter : مكتبة رسومية مدمجة في بايثون تُستخدم لبناء واجهات المستخدم (GUI) مثل النوافذ والأزرار.

Sklearn (scikit-learn) :  
مكتبة قوية للتعلم الآلي توفر أدوات جاهزة لبناء نماذج تصنيف، انحدار، تجميع، وتقسيم البيانات.

Matplotlib : مكتبة لرسم المخططات والرسوم البيانية ثنائية الأبعاد (2D) بطريقة مرنة ومخصصة.

Seaborn : مكتبة مبنية على matplotlib لتسهيل رسم الرسوم الإحصائية الجميلة والمعقدة بسهولة.

Pandas : مكتبة لتحليل البيانات ومعالجتها، وتوفر هياكل بيانات قوية مثل DataFrame للتعامل مع الجداول.

Re : مكتبة مدمجة في بايثون تُستخدم لمعالجة النصوص باستخدام التعابير النمطية

(Regular Expressions) .

Os : مكتبة مدمجة تتيح التعامل مع نظام التشغيل، مثل الوصول للملفات والمجلدات وتعديل المسارات.