

Exploratory Data Analysis

07/2023



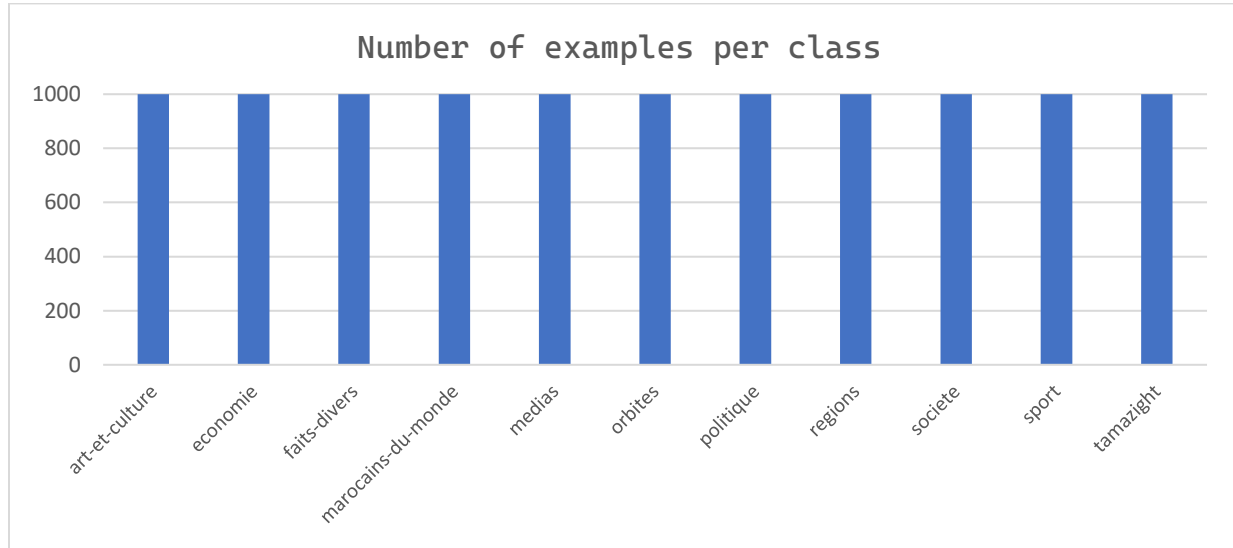
WideBot
وايد بوت

Prepared By:

Abdullah Mohamed

EDA Report

- Number of examples per class



- Top 10 frequent 3-grams generally

الجملة	م
الملك محمد السادس	1
مشيرا إلى أن	2
في المائة من	3
لجريدة هسبريس الإلكترونية	4
فيروس كورونا المستجد	5
في تصريح لهسبريس	6
تصريح لجريدة هسبريس	7
في تصريح لجريدة	8
سعد الدين العثماني	9
حالة الطوارئ الصحية	10

- Top 10 frequent 3-grams for art-et-culture class

الجملة	م
الثقافة والشباب والرياضة	1
المهرجان الدولي للفيلم	2
وزارة الثقافة والشباب	3
لجريدة هسبريس الإلكترونية	4
مشيرا إلى أن	5
المركز السينمائي المغربي	6
تصريح لجريدة هسبريس	7
في تصريح لهسبريس	8
في تصريح لجريدة	9
الإشارة إلى أن	10

- Top 10 frequent 3-grams for economie class

الجملة	م
في المائة من	1
فيروس كورونا المستجد	2
الاقتصاد والمالية وإصلاح	3
والمالية وإصلاح الإدارة	4
في المائة في	5
من السنة الجارية	6
العام لمقاولات المغرب	7
في المائة خلال	8
أزمة فيروس كورونا	9
حالة الطوارئ الصحية	10

- Top 10 frequent 3-grams for faits-divers class

الجملة	م
النيابة العامة المختصة	1
تدبير الحراسة النظرية	2
إشراف النيابة العامة	3
تحت إشراف النيابة	4
تحت تدبير الحراسة	5
وهن إشارة البحث	6
الحراسة النظرية وهن	7
النظرية وهن إشارة	8
العامة للأمن الوطني	9
مناصر الدرك الملكي	10

- Top 10 frequent 3-grams for marocains-du-monde class

الجملة	م
المغاربة العالقين في	1
الجالية المغربية المقيمة	2
أفراد الجالية المغربية	3
بالمغاربة المقيمين بالخارج	4
المكلفة بالمغاربة المقيمين	5
الشؤون الخارجية والتعاون	6
الملك محمد السادس	7
مشيرا إلى أن	8
بالخارج وشؤون الهجرة	9
المقيمين بالخارج وشؤون	10

- Top 10 frequent 3-grams for medias class

الجملة	م
التي نشرت أن	1
المنبر ذاته فإن	2
من المساء التي	3
الملك محمد السادس	4
مشيرة إلى أن	5
الأحداث المغربية التي	6
الجريدة ذاتها أن	7
أخبار اليوم أن	8
أخبار اليوم التي	9
ورد بها أن	10

- Top 10 frequent 3-grams for orbites class

الجملة	م
الملك محمد السادس	1
على الله عليه	2
الله عليه وسلم	3
فيروس كورونا المستجد	4
الولايات المتحدة الأمريكية	5
النبي على الله	6
لجريدة هسبريس الإلكترونية	7
بها في ذلك	8
مشيرا إلى أن	9
لا يمكن أن	10

- Top 10 frequent 3-grams for politique class

الجملة	م
الملك محمد السادس	1
مشيرا إلى أن	2
سعد الدين العثماني	3
حزب العدالة والتنمية	4
التجمع الوطني للأحرار	5
الأمين العام لحزب	6
الشؤون الخارجية والتعاون	7
الخارجية والتعاون الإفريقي	8
وزير الشؤون الخارجية	9
والتعاون الإفريقي والمغاربة	10

- Top 10 frequent 3-grams for regions class

الجملة	م
بفيروس كورونا المستجد	1
المديرية الجهوية للصحة	2
خلال الـ 24 ساعة	3
جديدة بـ فيروس كورونا	4
بني ملال خنيفرة	5
في كل من	6
فيروس كورونا المستجد	7
إصابة جديدة بـ فيروس	8
الـ 24 ساعة الماضية	9
الفقيه بن صالح	10

- Top 10 frequent 3-grams for societe class

الجملة	م
وزارة التربية الوطنية	1
لجريدة هسبريس الإلكترونية	2
تصريح لجريدة هسبريس	3
بفيروس كورونا المستجد	4
في تصريح لجريدة	5
التعليم عن بعد	6
في كل من	7
في المائة من	8
خلال الـ 24 ساعة	9
هسبريس الإلكترونية أن	10

- Top 10 frequent 3-grams for sport class

الجملة	م
البالغ من العمر	1
الملكية المغربية لكرة	2
المغربية لكرة القدم	3
الجامعة الملكية المغربية	4
في وقت سابق	5
الدفاع الحسني الجديدي	6
فيروس كورونا المستجد	7
تفشي فيروس كورونا	8
خلال الانتقالات الصيفية	9
في التعاقد مع	10

- Top 10 frequent 3-grams for tamazight class

الجملة	م
الملكي للثقافة الأمازيغية	1
الطابع الرسمي للأمازيغية	2
المعهد الملكي للثقافة	3
تفعيل الطابع الرسمي	4
مشروع القانون التنظيمي	5
رأس السنة الأمازيغية	6
لتفعيل الطابع الرسمي	7
اللغة الأمازيغية في	8
بتفعيل الطابع الرسمي	9
الوطني للغات والثقافة	10

- Length of examples in words and letters

Number of words	Number of letters
4,140,471	25,191,160