

Hespress Stories Sentiment Analysis Classification TF-IDF and SVC

Table of Contents

- [1. Introduction](#)
- [2. Importing Libraries](#)
- [3. Loading Data](#)
- [4. Data Exploration](#)
- [5. Data Preprocessing](#)
- [6. Exploratory Data Analysis](#)
- [7. Modeling](#)
- [8. Final Results](#)

Hespress

Moroccan News/Stories and Comments from Hespress

In this Project we gonna use the 11 csv files of the Stories only

About Dataset > Kaggle Link : [Hespress](#)

- The Arabic NLP scene is still developing, this dataset will help with that!
- I've scraped over 10 000 stories from Hespress with details such as the author, the publishing date, the topic.

- I've also scraped all the comments associated with them which resulted in over 300K comments, this is very interesting for sentimental analysis, understanding the general opinion, and maybe even predicting election results, since with each comment a score by the readers is associated.

::: Importing Libraries :::

[Go to Table of Contents](#)

```
In [42]: import os
import pandas as pd
import matplotlib.pyplot as plt
import re
import string
from collections import Counter

import nltk

# nltk.download('punkt') # make sure to uncomment this line
# nltk.download('stopwords') # make sure to uncomment this line

from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.util import ngrams
from pyarabic.araby import strip_tashkeel, normalize_ligature, strip_tatweel

from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.svm import LinearSVC
from sklearn.metrics import classification_report, accuracy_score
from sklearn.model_selection import RandomizedSearchCV
from sklearn.pipeline import Pipeline

import warnings
warnings.filterwarnings('ignore')
pd.set_option('display.max_colwidth', None)
```

```
In [2]: %%html
<style>
.output_wrapper .output {
    overflow-y: visible;
    height: fit-content;
}
</style>
```

```
from IPython.display import display, HTML
display(HTML("<style>.container { width:100% !important; }</style>"))
display(HTML("<style>.output_result { max-width:100% !important; }</style>"))
display(HTML("<style>.prompt { display:none !important; }</style>"))
```

```
from IPython.display import display, HTML display(HTML("")) display(HTML("")) display(HTML(""))
```

::: Loading Data :::

[Go to Table of Contents](#)

```
In [3]: # create an empty list to hold all the dataframes
dfs = []

# get a list of all CSV files in the data directory
data_dir = "data"
csv_files = [f for f in os.listdir(data_dir) if f.endswith('.csv')]

# Loop through the CSV files and append each one to the list of dataframes
for file in csv_files:
    file_path = os.path.join(data_dir, file)
    df = pd.read_csv(file_path)
    dfs.append(df)

# concatenate all the dataframes together into one dataframe
combined_df = pd.concat(dfs)
```

```
In [4]: df = combined_df[['story', 'topic']].copy()
```

::: Data Exploration :::

[Go to Table of Contents](#)

```
In [5]: # print a sample of the combined dataframe
df.sample(2)
```

في خطوة تهدف إلى تجاوز التداعيات السلبية التي أثارها قراره القاضي بأن ينوب عضو أو عضوان عن كل فريق في التصويت على مشاريع القوانين في ظل الحجر الصحي، الذي جر عليه اتهامات بخرق الدستور، اعتمد مجلس المستشارين التصويت عن بعد.وشهدت فترة الطوارئ الصحية جدلا كبيرا في البرلمان المغربي بسبب القرارات التي اعتمدها مكتبها مجلسي النواب والمستشارين القاضية بحضور عدد قليل من النواب والمستشارين إلى قبة البرلمان، وذلك ضمن الإجراءات الصحية للتخفيف من خطر انتشار الفيروس في المؤسسة التشريعية.وانطلق الجدل خلال تصويت لجنة المالية والتنمية الاقتصادية بمجلس النواب على مشروع قانون يجيز للحكومة تجاوز سقف التمويلات الخارجية من الاقتراض، حيث نبه البرلمان عن فيدرالية اليسار عمر بلافريج إلى أن عدد المصوتين كان تسعة في مقابل معارضة برلماني واحد، لكن رئيس اللجنة قال إن عدد المصوتين بلغ 43 مقابل معارضة برلماني واحد.وفي سابقة من نوعها، صوت مجلس المستشارين أمس الاثنين على مشروع قانون رقم 30.20 بسن أحكام خاصة تتعلق بعقود الأسفار والمقامات السياحية وعقود النقل الجوي للمسافرين، كما تم التوصل به من مجلس النواب، معتمدا التصويت عن بعد في الجلسة العامة.وباستعمال شاشتين كبيرتين وسط القاعة الكبرى للجلسات، صوت 54 برلمانيا لصالح المشروع الحكومي الذي قدمته وزيرة السياحة نادية فتاح العلوي، في حين لقي مشروع القانون معارضة 6 برلمانيين، دون تسجيل أي امتناع.وسبق لمكتب مجلس المستشارين أن قرر في آخر اجتماع له اعتماد التصويت الإلكتروني عن بعد على النصوص التشريعية الجاهزة، مشددا على أن الهدف هو ضمان مشاركة أوسع للمستشارين في عملية التصويت.ولتعميم التصويت الإلكتروني عن بعد وتأمين إعماله في أحسن الظروف مستقبلا، أكد مكتب المجلس وجوب اتخاذ كل التدابير والإجراءات الإدارية والتقنية اللازمة لتأهيل منظومة الاتصال السمعي البصري والمعلوماتي على مستوى قاعة الجلسات وقاعات اجتماعات اللجان، مشددا على أهمية تطوير تطبيق معلوماتي متقدم يتيح إمكانية التغلب على كل الصعوبات التقنية ذات الصلة.وضمن إجراءات مواجهة جائحة كورونا، اعتمد مجلسا البرلمان في وقت سابق حضور رؤساء الفرق والمجموعات النيابية، أو من ينوب عنهم، بالإضافة إلى عضوين عن كل فريق ومجموعة نيابية.

politique

جمعت الحاجة إلى السير "نحو وعي بيئي جديد" مجموعة من الأكاديميين والباحثين المغاربة، في كتاب جماعي جديد صدر عن "مركز فاطمة الفهرية للأبحاث والدراسات".وعُنونَ هذا الكتاب الجماعي الصادر رقميا، بفعل الجائحة بـ (التربية البيئية وسؤال التنمية والأخلاق "نحو وعي بيئي جديد")، ونسّق أبحاثه الأكاديمي أحمد الفراك.ويتضمّن هذا الكتاب، على سبيل المثال لا الحصر، أبحاثا في فلسفة الأخلاق البيئية، وأهمّ اتجاهات ونظريات علم اجتماع البيئة، والبعد الأخلاقي في بناء الوعي الحضاري بالبيئة في الفكر الإسلامي، وأخرى تتّخذ من التربية البيئية مدخلا للوعي بالمشترك الإنساني، وأثر التغيّرات المناخية على التنوّع البيولوجي.وشارك في هذا المؤلّف باحثون هم: جمال الدين ناسك، وربيع الحمداوي، ومحمد شرايمي، ويونس المرابط، وعبد الله الجباري، وأحمد الفراك، ورشيد أمشنوك، ومحمد الفتات، ولبنى السحايمي، وأسماء غيلان، ويونس الخليلشي، ولحسن بلالي ومريم رشيد، وسميرة عيسو، ومصطفى العادل، وأحمد أهلال.ويقول الكتاب في تقديمه إنّّه رغم ما للاهتمامات الفكرية القديمة من قيمة عظمت في تركيز النّظر على المعطى الكوني في تكامل عناصره وتفاعلها، وتأثير ذلك على حياة الإنسان وتفكيره؛ إلا أنّ نظرية "أخلاق الفضيلة" لأرسطو طاليس، و"أخلاق الواجب" لكانط، و"أخلاق المنفعة" لوليام جيمس، لا تنفع في استيعاب مشكلات "البيئة والإنسان" المستحدثة في عالمنا المعاصر، ولا في تنمية "الوعي الكافي" لإيقاف نزيف هذا الكون جراء استفحال جشع الإنسان وتسلّطه على الطبيعة، وبالتاليّ على نفسه والأحياء والحياة من حوله.ويذكر الكتاب في تقديمه أنّه "دون تربية بيئية جذرية ومستدامة، لا يكتفي الوعي الجديد الذي تحمله نداءات فلسفة البيئة المعاصرة والأخلاق التطبيقية بوصف الواقع البيئس وإنشاء خطابات تتحدّث عن الطّموح والآمال والأحلام الجميلة، بلغة متأسّفة وعاجزة عن الإقناع والتأثير، بل يتجاوز هذا إلى بذل وتنظيم الجهود الجماعية لتحقيق الوعي العالميّ المشترك بالأزمة التي يستشعرها الإنسان طوعا وكرها".ثمّ يعدّ جهد تحقيق الوعي العالمي المشترك بالأزمة، يضيف الكتاب، يُشرك الجميع دولا ومجتمعات، وأفراد وهيئات، في الشّمال والجنوب، والشّرق والغرب، قصّد "إنجاز وعي بيئي جديد وإعداد برنامج الحلّ المشترك وتنفيذه".ليذكر بعد ذلك تقديم الكتاب الذي خطّه أحمد الفراك، أستاذ الفلسفة بجامعة عبد المالك السعدي في تطوان، أنّ هذا المولود البحثي "محاولة جادة سابقة ولاحقة في سبيل الإسهام العلمي والفكري في تجديد الوعي البيئي والتربية على حفظ المعطى الإيكولوجي الذي نحيا في وسطه".يُذكر أنّ الطبعة الأولى من كتاب "التربية البيئية وسؤال التنمية والأخلاق - نحو وعي بيئي جديد"، صدرت عن "مفاد" مركز فاطمة الفهرية للأبحاث والدراسات، وهي ثالث المؤلّفات الصادرة عن سلسلة "كتب جماعية محكمة"، ويمكن قراءتها وتحميلها مجّانا على شبكة الإنترنت.

art-et-culture

In [6]: df.shape

Out[6]: (11000, 2)

In [7]: df.isna().sum()

Out[7]: story 0
topic 0
dtype: int64

```
In [8]: df.duplicated().sum()
```

```
Out[8]: 4
```

```
In [9]: # identify duplicated rows  
duplicated_rows = df[df.duplicated()]  
duplicated_rows
```

جدّد عبد اللطيف الجواهري، والي بنك المغرب، الدعوة إلى اعتماد إجراءات تحفيزية لدعم مشروع الأداء عبر الهاتف النقال، الذي أصبح منظومة جاهزة منذ مدة، لكن إطلاقه على نطاق واسع لم يتم بعد. وأشار الجواهري، في ندوة صحافية رقمية عقدها الثلاثاء عقب اجتماع مجلس البنك، أن الأداء عبر الهاتف يعتبر من الحلول الأساسية لمحاربة تداول النقد (الكاش) الذي ارتفع منذ بدء الجائحة. وأكد والي بنك المغرب أن اعتماد الأداء عبر الهاتف سيمكن من خفض وقت معالجة العمليات وتكلفتها، كما شدد خاصة للأداء عبر الهاتف، وزاد: (m-wallet) على أنه مهم من أجل تحقيق الشمول المالي الذي يسعى إليه المغرب. وذكر الوالي أن المغرب حالياً يتوفر على 1.5 محفظة "على المستويين التقني والمؤسساتي كل شيء أصبح جاهزاً، لكن على البرلمان أن يُقر إجراءات تحفيزية لتشجيع التجار على الانخراط في هذه المنظومة، لأننا ضيعنا وقتاً طويلاً". واقترح المتحدث اعتماد إجراءات تحفيزية على المستوى الضريبي لتشجيع الانخراط في الأداء عبر الهاتف، وإطلاق عملية تجريبية في إحدى الجهات من أجل الوقوف عليها نتائج العملية وفوائدها. وأكد الجواهري أنه "سيحرص شخصياً على ضمان البدء بالعمل بنظام الدفع عبر الهاتف على نطاق واسع في أقرب وقت ممكن"، مشدداً على أنه "سيطرق جميع الأبواب من أجل ذلك". وبخصوص ارتفاع الديون المتعثرة، قال الجواهري إنها ارتفعت في نهاية يوليو بـ 7 مليارات درهم، حيث انتقلت من 70 مليار درهم إلى 77 مليار درهم، وهي بذلك تمثل 8 في المائة من مجموعة القروض. وحسب الوالي، يجب التمييز بين الديون المتعثرة الناتجة عن تأجيل الأقساط الذي أقرته البنوك والديون الحقيقية المتعثرة، واعتبر أن ارتفاع هذه الأخيرة ليست حكرًا على المغرب، بل يهم جميع الدول بسبب تداعيات الجائحة والحجر الصحي. وأوضح الجواهري أن "القروض التي تضمنها الدولة بالنسبة للبنوك تمثل خطراً محدوداً"، لكنه أثار الانتباه إلى رصد ارتفاع حاد في حالات التخلف عن السداد، خصوصاً لدى شركات قروض الاستهلاك. ومن المرتقب أن يعقد الجواهري، حسب ما أعلنه في الندوة الصحافية، اجتماعاً مُرتقباً مع المجموعة المهنية لبنوك المغرب في شهر نونبر، لتقييم تأثير القروض المتعثرة على نتائج المنظومة البنكية، وتحديد مصادر هذه التعثر ما بين المقاولات والأسر.

economie

مع تفاقم الأوضاع الصحية والاقتصادية والاجتماعية وحتى النفسية... بسبب تمّدّد وباء "كورونا" الناتج عن فيروس "كوفيد 19"، في عدد من دول العالم، أصبحت الأنظار تتّجه نحو المختبرات العلمية التي وصلت فيها الأبحاث إلى مراحل متقدّمة، حيث بدأت الكثير منها في إصدار تقارير تبشّر بتقدم كبير على هذا المستوى. وهكذا، أكّدت منظمة الصحة العالمية أن هناك أكثر من 200 لقاح قيد التجربة في هذا الخصوص، 26 منها وصلت إلى مرحلة التجارب السريرية على البشر.. وأمام هذه التطورات، طرحت العديد من الأسئلة، فأمام خطر شلّ الحياة الاقتصادية عالمياً، وأصبح يهدّد الإنسانية جمعاء، هل يمثل تعدّد التجارب أمراً صحياً، يعزّز التنافس، لتحقيق سبق علمي، سيسجّله التاريخ؟ أم أن في ذلك هدر للإمكانات وللوقت، أمام وباء متسارع يقتضي التنسيق والتضامن؟ وهل ستتاح التلقيحات في حال توافرها بالنسبة للدول الفقيرة، بأثمنة في المتناول، أم أن هاجس الربح سيطغى على حساب المعاناة الإنسانية؟ وهل يمكن القبول بتحقيق ربح رغم أن الأمر يتعلق بجائحة خطيرة عابرة للحدود، علماً أن الكثير من الدول خصّصت تمويلات ومنح كبيرة، لإجراء البحوث المتعلقة بهذا الشأن، مصدرها الضرائب التي يؤدّيها المواطنون؟ تتضارب وتتباين الأجوبة الواردة في هذا الخصوص، ففي الوقت الذي يؤكّد فيه البعض على ضرورة تنسيق الجهود العلمية والمبادرات الدولية، وتجنيد الإمكانيات المالية والتقنية، على طريق تعزيز الأبحاث، في أفق طرح لقاح غير مكلف، يتيح وقف تمّدّد الوباء وعودة الحياة إلى طبيعتها، يعتبر البعض الآخر، أن تنويع اللّقاحات بين مختلف المختبرات والشركات، هو مدخل سليم سيدعم التعاطي مع الموضوع من زوايا متعددة، من شأنه أن يغني حقل التجارب، ويوسّع من هوامشها نجاحها، في إطار من التنافس، الذي سيسمح بتسريع وتيرة الأبحاث، ويدعم الجهود الرامية إلى انحسار الوباء ووقف خطره.. في شهر أبريل الماضي، وفي عرّ تطوّر تداعيات الوباء، أعلن عدد من العلماء والأطباء والممّولين والمصنّعين من مختلف البلدان (الولايات المتحدة الأمريكية، وإيطاليا، وبريطانيا، والصين، وأستراليا وكندا، والأرجنتين وإسبانيا، وهولندا، وفرنسا، واليابان والنرويج، وكوريا الجنوبية، وسويسرا، وجنوب إفريقيا، والسويد، وألمانيا، وسنغافورة)، تكثيف الجهود، في إطار تعاون دولي تشرف عليه منظمة الصحة العالمية، من أجل التوصل بشكل سريع إلى طرح لقاح فعّال وآمن ضد الوباء، يكون متاحاً للجميع. ورغم القصور الملحوظ الذي طبع تعاطي منظمة الصحة العالمية مع الجائحة في بداياتها، فقد أطلقت قبل بضعة أسابيع تجربة "التضامن" السريرية، كسبيل لدعم الجهود الدولية في إيجاد علاج ناجع لمرض كوفيد-19، وتتمحور هذه المبادرة حول اعتماد مقارنة بين عدد من الخيارات العلاجية بعدد من البلدان، لتقييم الأنجع منها، على مستوى القضاء على الفيروس، والمساعدة في التوصل السريع إلى طرح أدوية فاعلة في هذا الخصوص. كما طالبت دول العالم بتوفير حوالي 100 مليار دولار كدعم يساعد على إيجاد لقاح للوباء، والحدّ من تفشّيه. وفي نفس السياق، دعا الملياردير "بيل غيتس" إلى إنتاج اللّقاح داخل البلدان النامية، محدّراً من أنّ هذه الأخيرة، لن تتمكن من الحصول على اللّقاح في الوقت المطلوب، في حال خضوعه لمنطق السوق التجارية، ما سيعمّق جروح الأوضاع الصحية والاجتماعية أكثر.. فيما اعتبر الأمين العام الأممي "أنطونيو غوتريس" أن اللقاح ضد فيروس "كوفيد 19" يمثل "منفعة عامة عالمية"، ومن جانبها عبّرت بعض دول الاتحاد الأوروبي، ومؤسسة "بيل غيتس" عن توجّهها نحو جمع أكثر من 8 مليارات يورو لصالح منظمة الصحة العالمية، كسبيل لإحداث وكالة دولية، تعمل على توفير اللقاحات والأدوية لدول العالم قاطبة. وضعت الجائحة مجموعة من القيم والمبادئ المتصلة بالتعاون والتضامن الدوليين، على محكّ الواقع، ففي مقابل هذه الجهود البناءة، فضّلت الكثير من الدول الكبرى التعاطي مع الأزمة من منظور المصالح الضيّقة، والانكفاء على الذات، ووصل الأمر إلى حدّ منع تصدير الآليات الطبيّة نحو الخارج.. ورغم أن الأمر يتعلّق بوباء عالمي عابر للحدود، فقد أكّدت الكثير من الدول، كما هو الشأن بالنسبة للولايات المتحدة، وعدد من البلدان الأوربية على ضرورة إعطاء الأولوية للحاجيات الداخلية من اللّقاح أولاً، بعد إنتاجه.. وفيما يعتبر الكثير من الخبراء والمهتمين أن خطورة الأمر تقتضي تخليّ الشركات المنتجة عن هامش الربح، في سبيل إيصال الأدوية واللّقاحات إلى مختلف الفئات المجتمعية، وإلى كل الدول الفقيرة، حيث تتضاعف المعاناة الصحية، بوجود إشكالات اقتصادية واجتماعية وبيئية أكثر قساوة، تبرز الكثير من شركات الأدوية العالمية توجّهها إلى تحقيق الربح من وراء هذه اللقاحات، بنيّتها في توظيف ذلك لتطوير البحوث العلمية.. بين نظرة متشائمة تنخرط فيها حتى منظمة الصحة العالمية، ترى بأنه لا مجال للقاح فعال إلا بعد حلول عام 2021، وبخاصة وأن تطوير لقاحات سابقة ضدّ عدد من الفيروسات الخطرة، تطلّب مدداً تجاوزت العقدين من الزمن أحياناً، حيث يؤكّد الكثير من الخبراء والمختصين أن تطوير اللقاحات لا يخلو من صعوبات وتعقيدات جمة، تنضاف إلى كلفته العالية مادياً، ونظرة متفائلة ترى في تقدّم الأبحاث الجارية، وبلوغها إلى مستويات جيّدة، يمثل مؤشراً على اقتراب طرح لقاح قد ينهي الصفحة القائمة التي فتحتها الجائحة على العالم.. شكّل إعلان البدء في إنتاج اللقاح الروسي، الذي تم إطلاق اسم "سبوتنيك 5" عليه، بارقة أمل في هذا الخصوص.. رغم الانتقادات الغربية التي شكّكت في نجاعته

orbites

122

وشروط سلامته.. وقد سارعت العديد من البلدان نحو إبرام عقود وصفقات للحصول على اللقاح قبل خروجه إلى حيز الوجود.. وبغض النظر عن الوقت الذي سيكون فيه اللقاح جاهزاً، وعن الدولة التي سيكون لها السبق في هذا الخصوص، فالواضح أن هذه المرحلة ستمثل محطة حاسمة لقياس مستوى الرقي الإنساني لدى الكثير من الدول المتقدمة التي طالما تبجحت بالدفاع عن حقوق الإنسان

ترأس الملك محمد السادس، مساء اليوم، جلسة عمل خصصت لمتابعة تدبير انتشار وباء فيروس "كورونا" ببلادنا ومواصلة اتخاذ مزيد من الإجراءات لمواجهة أي تطور، وفق ما ذكره بلاغ للديوان الملكي. في ما يلي بلاغ للديوان الملكي: "ترأس صاحب الجلالة الملك محمد السادس، نصره الله، اليوم الثلاثاء بالقصر الملكي بالدار البيضاء، جلسة عمل خصصت لمتابعة تدبير انتشار وباء فيروس كورونا ببلادنا، ومواصلة اتخاذ مزيد من الإجراءات لمواجهة أي تطور. ويندرج هذا الاجتماع في إطار المتابعة المستمرة من طرف جلالة الملك، حفظه الله، لتطورات هذا الوباء، منذ بداية انتشاره على الصعيد العالمي، ومباشرة بعد ظهور الحالات الأولى على التراب الوطني. وفي بداية هذا الاجتماع، استفسر جلالة الملك وزير الصحة عن آخر تطورات الوضعية الصحية ببلادنا، والطاقة الاستيعابية للمستشفيات والوحدات الصحية، بمختلف جهات المملكة، وكذا توفير جميع مستلزمات السلامة الصحية، بما في ذلك مواد التعقيم والأدوية. وبناء على المعطيات التي قدمها السيد وزير الصحة، وحرصاً من جلالة الملك، حفظه الله، على الرفع من قدرات المنظومة الصحية الوطنية في مواجهة هذا الوباء، أصدر جلالة الملك، حفظه الله، القائد الأعلى ورئيس أركان الحرب العامة للقوات المسلحة الملكية، تعليماته السامية للمفتش العام للقوات المسلحة الملكية، قصد وضع المراكز الطبية المجهزة، التي سبق لجلالته أن أمر بإحداثها لهذا الغرض، بمختلف جهات المملكة، رهن إشارة المنظومة الصحية بكل مكوناتها، إن اقتضى الحال وعند الحاجة. وتندرج هذه المبادرة الملكية السامية في إطار المقاربة الاستباقية، التي أمر جلالة الملك باعتمادها منذ بداية ظهور هذا الوباء، وتعزيز الإجراءات الوقائية والاحترازية غير المسبوقة، التي اتخذتها القطاعات والمؤسسات المعنية، بهدف الحد من انتشاره، ومواجهة تداعياته الصحية والاقتصادية والاجتماعية. وبهذه المناسبة، تم استعراض مدى تنفيذ الإجراءات التي تم اتخاذها، بتوجيهات سامية من جلالته، والتي تهم إغلاق المجال الجوي والبحري المغربي أمام المسافرين، وإلغاء التجمعات والتظاهرات الرياضية والثقافية والفنية، وإحداث صندوق خاص بتدبير جائحة فيروس كورونا "كوفيد 19"، وتوقيف الدراسة بالمدارس والجامعات، والإغلاق المؤقت للمساجد، وتعليق الجلسات بمختلف محاكم المملكة، إضافة إلى مجموعة من الإجراءات التي بادرت لاتخاذها السلطات المختصة، في مجالات النقل العمومي، وإغلاق المحلات العمومية غير الضرورية. كما وجه جلالة الملك، نصره الله، السلطات المختصة للسهر على حسن تطبيق التدابير الناجعة المتخذة في مجال ضمان تزويد الأسواق عبر التراب الوطني، بجميع المواد الغذائية والاستهلاكية، وبمواد التطهير والتعقيم، بصفة منتظمة ومتواصلة، ومحاربة مختلف أشكال الاحتكار والزيادة في الأسعار. وفي نفس السياق، أصدر جلالة الملك تعليماته السامية لرئيس الحكومة ولجميع القطاعات المعنية، باتخاذ جميع الإجراءات اللازمة، والإعداد لمرحلة جديدة، إن اقتضى الحال ذلك. حضر هذا الاجتماع رئيس الحكومة السيد سعد الدين العثماني، ووزير الداخلية عبد الوافي لفتيت، ووزير الصحة خالد آيت الطالب، والجنرال دو كوردارمي عبد الفتاح الوراق المفتش العام للقوات المسلحة الملكية، والجنرال دو كوردارمي محمد حرمو قائد الدرك الملكي، والمدير العام للأمن الوطني المدير العام لمراقبة التراب الوطني السيد عبد اللطيف الحموشي.

politique

قررت المديرية الإقليمية لوزارة التربية الوطنية بالتأطير إغلاق مؤسسة تعليمية خصوصية بعد تسجيل 3 إصابات مؤكدة بفيروس كورونا المستجد في صفوف تلاميذها. وتقرر تعليق الدراسة بالمؤسسة ذاتها واعتماد نمط "التعليم عن بعد" ابتداء من يوم الثلاثاء على سبيل الاحتياط، إلى حين صدور بلاغ جديد يُعلن فيه عن تاريخ استئناف نمط التعليم الحضوري. وجاء اتخاذ هذا القرار بتنسيق مع السلطات الترابية والصحية والأكاديمية الجهوية للتربية والتكوين بجهة الشرقية وجمعية آباء وأولياء التلاميذ بالمؤسسة ذاتها، تفعيلاً للمذكرتين الوزاريين المتعلقتين بتدبير الموسم الدراسي في ظلّ جائحة فيروس كورونا

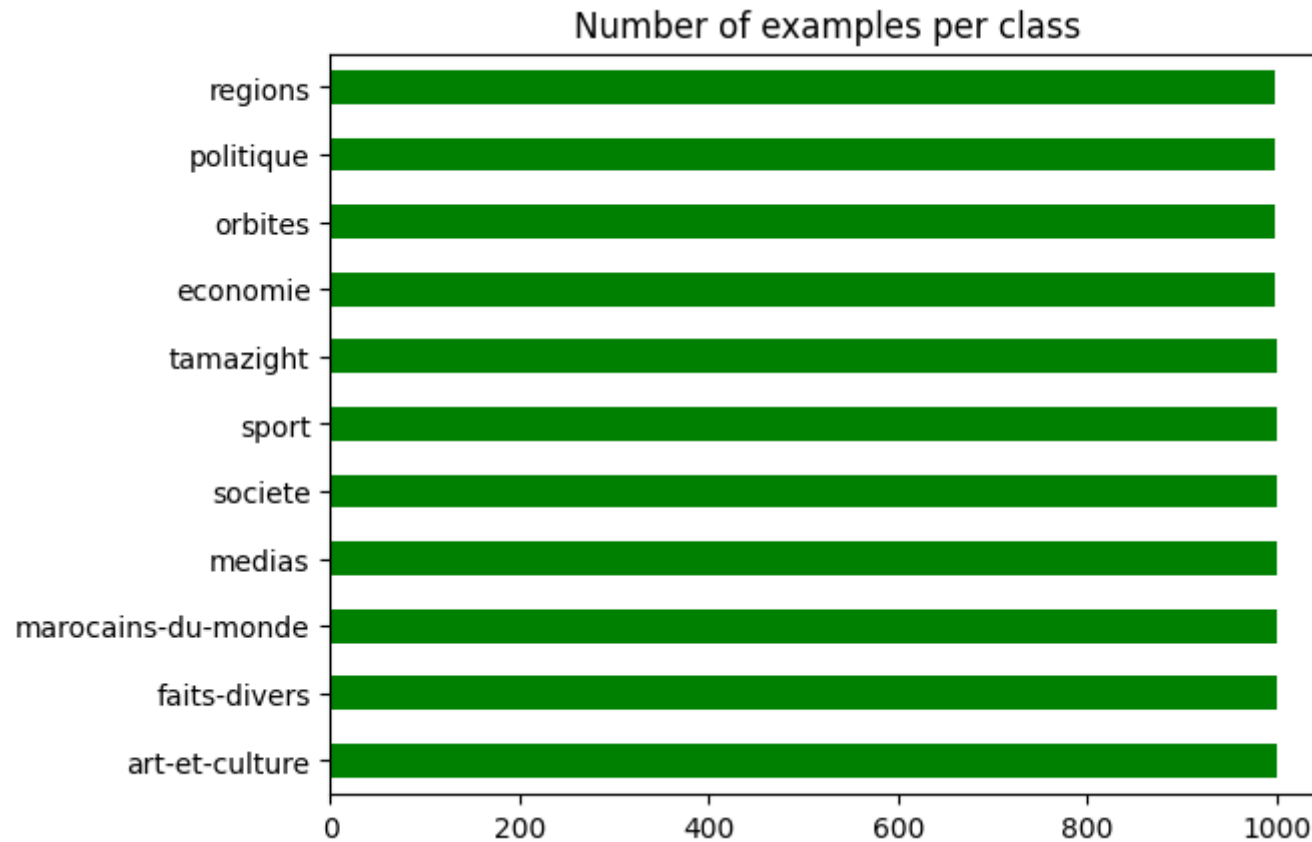
regions

```
In [10]: df.drop_duplicates(inplace=True)
```

```
In [11]: print(df['topic'].value_counts())
df['topic'].value_counts().plot(kind='barh',color='green')
plt.title('Number of examples per class')
plt.show()
```

art-et-culture	1000
faits-divers	1000
marocains-du-monde	1000
medias	1000
societe	1000
sport	1000
tamazight	1000
economie	999
orbites	999
politique	999
regions	999

Name: topic, dtype: int64



```
In [12]: ## create a regular expression pattern that matches any non-Arabic letter or non-space character
## non_arabic_pattern = re.compile('[^â-_\s]+')

## replace any non-Arabic characters in the "text" column with a space
## df['story'] = df['story'].str.replace(non_arabic_pattern, ' ')
```


1.Viewing Sample with idx-2 before Data Preprocessing

```
In [13]: df.iloc[2]
```

```
Out[13]: story    تائهة في (Alice)، تشير مشاهدة فيلم قصير ضمن الثلاثية الأخيرة للمخرج المغربي هشام العسري إلى جملة من المرجعيات الثقافية والأدبية والأسطورية العالمية التي وظفها. كان يا مكان، امرأة بطة  
    متربص ينتظر فرصة للانقضاض على ضحاياه أو ضحية Bayard وهو يحاول، بجشع، بيع ما يظن أنه يصنعه، الكلب The mad hatter "بلاد العجائب". عقارب الساعة وهي تدور إلى الخلف، صانع القبعات  
    4. إنها المرأة البطلة التي تمشي بحثاً عن المجهول وكأنها تأخذ كلبها في نزهة، أو أنها في حداد على موته، أو لربما هي ببساطة غاضبة أو مجنونة أو أنها أعقل من العقلاء. في الفيلم القصير - بدون عنف  
    مشاهدة تبرز قوة المرأة (Cruelty Free)، The last Arab Movie، Android and Zombies: ضمن ثلاثية أفلام قصيرة هي على التوالي)، من تأليف وسيناريو وإخراج هشام العسري -  
    ٥، (شخصية البطلة "ندى" في الفيلم)، وكيف أنها تستطيع مواصلة السير رغم الألم والابتسام رغم الأحزان. تلتقط الكاميرا مشية أنثوية مُختالة وأنيقة بخطى ثابتة، فالحذاء ذو الكعب العالي يؤلم القدمين لكن المرأة تحا  
    وفجأة... يصبح الجمال "Keep your standards, heels and head always high": ذاتها قائلة Coco Chanel فظ على مقاييسها وعلى رأسها مرفوعاً ككعب حذاءها، كما نظرت لذلك كوكو شانيل  
    جريمة؟ من السهل اكتشاف الشعور بالجنون. هل من السهل السير في طريق الانتقام؟ معانٍ سخيصة وسريالية؛ لكن ما خفي في الفيلم أعظم، فمن يشتري الزهور؟ ومن يفك معاني ضاعت دلالاتها مع الأيام وما تحمله من  
    P ثقل عنف وعادات وتقاليد وأحكام؟ قد يكون الأهم من كل ذلك هو واقع المجتمع المغربي وكيف يريد الكل... نريد جميعاً... فقط، أن نشاهد، أن نتفرج، أن ننظر، أن نرمق، أن نحقق وأن نسرق نظرات ونقوم ب  
    على ما يُسمح وما لا يُسمح بمشاهدته. نحب ببساطة - حد التبسيطية - أن "ننشغل" بقصص الآخرين، مسيطرين على قصصهم. لهذا "الاهتمام" دلالاته العميقة والمأساوية، فقد تكون تلك القصص lan séquence  
    ص، هي كل ما يمكن للمغربي أن "يسيطر" عليه، من خلال تحديقه. نتوقع وننتظر وننوي دائماً الأسوأ في الناس ومنهم. نظن أننا نمتلك أو نمسك بالحقيقة بين أيدينا الصفر من أية حقيقة وبكل ما يحوم حولها لكن،  
    من يمسك بها حقاً؟ لا أحد يعلم. في الفيلم لمسة جيدة وإبداعية، لمسة عميقة بأسلوب الفنان المغربي المقيم في الخارج حسن الحجاج: المغرب الحقيقي، (العميق كما يسميه البعض)، والهاتف العمومي وصناديق الكوكاك  
    ولا القديمة والبلاستيكية الحمراء، حتى فرن المايكرويف الذي يمثل تابوتاً لجثة الكلب، ينتمي لنفس الأسلوب البصري المبدع. إنه فنٌ يعرض ويبرز واقعاً مؤلماً: خضر وفواكه مُلقاة في كل مكان، لكن الناس ما زالوا  
    والتي هي بعيدة المنال بالنسبة لهم. الألوان والواقعية المؤلمة في الفيلم، كعادة كل أفلام هشام العسري، نقد اجتماعي واضح لواقع بني، Maserati: جانحين، يصفون بها المرأة البطلة الجميلة والمنتمية لمستوى سيارتها ال  
    س. يتجلى ذلك في فيلم "بدون عنف" الذي نحن بصده، وذلك من خلال إظهار هذا الواقع عارياً من أي جمالية. أليس هذا الأسلوب مظلماً وظالماً لمجتمع بأكمله بثقافته وتاريخه وذكاء وإبداعات مواطنيه وإلا: أين ا  
    حيث يُشد "The Portrait of a Lady" للمساء الصيفي في إحدى القرى الإنجليزية في روايته Henry James لأمل؟ يشبه وصف هشام العسري للواقع المغربي في فيلمه القصير هذا وصف الروائي الأمريكي  
    ويختزلها في صورة سوداء وعبثية. لعل هشام العسري يضع هنا بصمة مؤثرة للجدل على واقع مؤلم، هو في الغالب واقعٌ مسكوت عنه، لكن، هل لا أم The afternoon tea وهُ الروائي معالم طقوس شاي المساء  
    ل. في إصلاح مجتمعنا، إذا كنا نعتقد حقاً أن الأمر يستحق الإصلاح
```

::: Data Preprocessing :::

[Go to Table of Contents](#)

```
In [14]: # create a regular expression pattern that matches any non-Arabic Letter
non_arabic_pattern = re.compile('[^\u0600-\u06FF\s]+')

# replace any non-Arabic letters in the "text" column with an empty string
df['story'] = df['story'].str.replace(non_arabic_pattern, '')
```

```

In [15]: # define a translation table that maps each Arabic punctuation mark to a space
arabic_punctuations = '``÷×-“”!|+|~{}',.‘”:/«»[%^&*()_<>‘’'
translation_table = str.maketrans(arabic_punctuations, ' ' * len(arabic_punctuations))

# apply the translation table to the "text" column
df['story'] = df['story'].str.translate(translation_table)

In [16]: # Replace 2 or more spaces with 1 space
df['story'] = df['story'].replace('\s+', ' ', regex=True)

In [17]: # define function to normalize arabic text
def normalize_arabic_text(text):
    """
    Normalizes Arabic text by removing diacritics, replacing ligatures with their component letters,
    and removing tatweel characters.
    """
    normalized_text = strip_tashkeel(text)
    normalized_text = normalize_ligature(normalized_text)
    normalized_text = strip_tatweel(normalized_text)

    return normalized_text

# define function to remove arabic stop words
def remove_stopwords(text):
    stop_words = set(stopwords.words('arabic'))
    words = word_tokenize(text)
    return ' '.join([word for word in words if word not in stop_words])

# define function to normalize specific letters
def normalize_letters(text):
    text = re.sub(' [ى]', ' ي', text)
    text = re.sub(' [إئإ]', ' ا', text)
    return text

# def remove_newlines_and_tabs(text):
#     cleaned_text = text.replace('\n', '').replace('\t', '')
#     return cleaned_text

# apply functions to 'story' column in our dataframe
df['story'] = df['story'].apply(normalize_arabic_text)
df['story'] = df['story'].apply(remove_stopwords)
df['story'] = df['story'].apply(normalize_letters)

```

2.Viewing the same Sample with idx-2 after Data Preprocessing

```
In [18]: df.iloc[2]
```

```
Out[18]: story    تشير مشاهدة فيلم قصير ضمن الثلاثية الاخيرة للمخرج المغربي هشام العسري جملة المرجعيات الثقافية والادبية والاسطورية العالمية وظفها مكان امراة بطلة تائهة بلاد العجائب عقارب الساعة وهي تدور  
الخلف صانع القبعات يحاول بجشع بيع يظن انه يصنعه الكلب متربص ينتظر فرصة للانقضاض ضحاياه ضحيته انها المرأة البطلة تمشي بحثا المجهول وكانها تاخذ كليها نزهة انها حداد موته لربما ببساطة غاضبة مجذ  
ونة انها اعقل العقلاء الفيلم القصير بدون عنف تأليف وسيناريو واخراج هشام العسري ضمن ثلاثية افلام قصيرة التوالي مشاهد تبرز قوة المرأة شخصية البطلة ندي الفيلم وكيف انها تستطيع مواصلة السير رغم الالم و  
الابتسام رغم الاحزان تلتقط الكاميرا مشية انثوية مختالة وانيقة بخطي ثابتة فالحذاء الكعب العالي يؤلم القدمين المرأة تحافظ مقاييسها وعلي راسها مرفوعا ككعب حذاءها نظرت لذلك كوكو شانيل ذاتها قائلة وفجأة يصبح  
الجمال جريمة السهل اكتشاف الشعور بالجنون السهل السير طريق الانتقام معان سخيصة وسريالية خفي الفيلم اعظم يشتري الزهور يفكك معاني ضاعت دلالاتها الايام تحمله ثقل عنف وعادات وتقاليد واحكام يكون الاهم  
واقع المجتمع المغربي وكيف يريد الكل نريد جميعا فقط نشاهد نتفرج ننظر نرمق نحدق وان نسرق نظرات ونقوم يسمح يسمح بمشاهدته نحب ببساطة حد التبسيطة ننشغل بقصص الآخرين مسيطرين قصصهم لهذا الاهد  
مام دلالاته العميقة والمساوية فقد تكون القصص يمكن للمغربي المقيم الخارج حسن الحجاج المغرب الحقيقي العميق يسميه البعض والهاتف العمومي وصناديق الكوكا كولا القديمة والبلاستيكية الحمراء فرن المايكرويف يمثل  
تابوتا لجثة الكلب ينتمي لنفس الاسلوب البصري المبدع فن يعرض ويبرز واقعا مؤلما خضر وفواكه ملقاة مكان الناس زالوا جانعين يصفون المرأة البطلة الجميلة والمنتمية لمستوي سيارتها ال والتي بعيدة المنال بالنسبة  
الالوان والواقعية المؤلمة في الفيلم كعادة افلام هشام العسري نقد اجتماعي واضح لواقع بنيس يتجلى فيلم بدون عنف بصده وذلك خلال اظهار الواقع عاريا جمالية اليس الاسلوب مظلما وظالما لمجتمع باكملة بثقافته وتاريخه  
وذكاء وابداعات مواطنيه والا الامل يشبه وصف هشام العسري للواقع المغربي فيلمه القصير وصف الروائي الامريكي للمساء الصيفي القري الانجليزية روايته يشوه الروائي معالم طقوس شاي المساء ويختزلها صور  
ة سوداء وعبثية هشام العسري يضع بصمة مثيرة للجدل واقع مؤلم الغالب واقع مسكوت عنه امل اصلاح مجتمعنا كنا نعتقد الامر يستحق الاصلاح  
topic  
art-et-culture  
Name: 2, dtype: object
```

::: Exploratory Data Analysis :::

[Go to Table of Contents](#)

```
In [19]: df.topic.unique().tolist()
```

```
Out[19]: ['art-et-culture',  
'economie',  
'faits-divers',  
'marocains-du-monde',  
'medias',  
'orbites',  
'politique',  
'regions',  
'societe',  
'sport',  
'tamazight']
```

```
In [20]: # topic encoding  
from sklearn.preprocessing import LabelEncoder  
  
# initialize LabelEncoder  
encoder = LabelEncoder()
```

```
# fit and transform 'topic' column, and add 1 to the encoded labels
df['topic_encoded'] = encoder.fit_transform(df['topic']) + 1
df[['topic', 'topic_encoded']].sample(10)
```

Out[20]:

	topic	topic_encoded
816	sport	10
92	orbites	6
513	sport	10
785	economie	2
355	sport	10
511	societe	9
109	economie	2
490	societe	9
769	regions	8
18	societe	9

In [21]:

```
df.head(3)
```

وجه بيت الشعر المغرب وزير الثقافة والشباب والرياضة رسالة موسومة لماذا تحولت احلام بيضاء كوابيس سوداء اشار خلالها بيت الشعر تاسس سنة بهدف تحقيق جملة الاهداف تروم جميعها تعزيز مكانة الشعر المجتمع والحياة وترسيخ مكانته الناس كحامل لقيم الحلم والخيال الاهداف توطين الشعر المغربي المقررات الدراسية وتشجيع التلاميذ والطلبة قراءته وتذوق جمالياته اللحظة ينتصر لكل مدهش وانساني وجاء الرسالة توصلت هسبريس بنسخة بيت الشعر نجح المغرب خلال مسيرته الطويلة تمتد مدي ربع قرن اقناع المنظمة العالمية للتربية والثقافة والعلوم اليونيسكو باحداث يوم عالمي للشعر طريق المقترح تقدم تبنته الحكومة المغربية عهد الراحل عبد الرحمان اليوسفي نجح ضمان مكانة عالمية لجائزته الشعرية المعروفة الراكنة شعراء العالم يتطلعون للفوز علاوة انتظام منشوراته الشعرية ومجلته الرصينة البيت واستدامة برامجه الشعرية بالتعاون عدد الشركاء امنوا بجدية مشروع مؤسستنا وجوده مستواه فنيا وجماليا واشارت الرسالة ذاتها انه سعي بيت الشعر المغرب تقديم خدمة جديدة للشعر المغربي يربطه بالفضاء الابداعي العام وخاصة الموسيقى والغنائي واستعادة لحظات طيبة الذكر تعانق شعرنا المغربي عبد الرفيق جواهري وادريس الجاي والخمار الكونني وحسن المفتي واحمد الطيب لعلي الحداني اوتار والجان عبد السلام عامر وعبد النبي الجبراري وحسن القديري تقدم بمشروع فني عنوان احلام بيضاء واوضح بيت الشعر المشروع عبارة مقطوعات غنائية جعلت الشعر المغربي متنا بهدف استعادة ماض جميل اثري خلاله الشعراء المغاربة السجل الشعري للاغنية المغربية فقد منحوها قصائد تحولت خلال اوتار ملحنين مقتدرين اغان تتردد السنة وشفاها الناس المناسبات والافراح والاعراس بعضها سيحظي باهتمام مطربين مشاركة اعداوا ادائها وتسجيلها جديد باصواتهم وورد الرسالة ذاتها انه نجاح يتحق الشعراء كتبوا قصائدهم داخل تجربتهم الانسانية وفي افق الرؤية امتلكوها اللغة والمجتمع والكون وليس الطلب ارغامات سوق الغناء وحسب الرسالة ذاتها فانه لاول مرة تاريخ برنامج الوزارة لدعم الاغنية المغربية تتقدم حياة متخصصة كبيت الشعر المغرب بمشروع رؤية مندمجة تربط الشعر المغربي بافقه الفني يحظي بمساحة اوسع للانتشار عبر الاغنية يصير بمقدور الجمهور المغربي يتعرف شعرنا المغاربة خلال دواوينهم الشعرية عبر امسياتهم الثقافية خلال الاغنية كالحظة فنية تتجمع وتنصهر عدة ابعاد شعرية ولحنية موسيقية وطربية غنائية وقد سبق للفنانة زيداني جعلت الشعر المغربي افقا لتفكيرها واشتغالها تضيف الرسالة ذاتها عندما قدمت بعضا نصوصه الجميلة للشعراء بوجمعة العوفي وعبد الهادي السعيد خاضت تجربة فنية الشاعر المغربي الكبير عبد الله زريقة بمعية ثلة الشعراء والموسيقيين الاجانب يجعل وجودها ضمن المشروع تثمينا لهذه الارادة تلتقي برغبة بيت الشعر المغرب يكون شعرنا المغربي حاضرا مختلف الحوامل تتيح الذبوع والانتشار مسرح تشكيل اغنية واكد بيت الشعر تشبيك الفنون بينها وتجسير الصلات مكوناتها اهم مخرجات المشروع يجمع شعراء مرموقين بملحنين مقتدرين وموزعين اكفاء وعازفين ماهرين علاوة الحضور الواعي والباذخ للفنانة زيداني انخرطوا المشروع نعيد للكلمة الشعرية بريقها لحظة تعالقاها للحن والموسيقي واختتم بيت الشعر رسالته بالتساؤل لماذا تحولت الاحلام البيضاء كوابيس سوداء ولماذا خيبت اللجنة شكلتموها دعم الاغنية المغربية والارتقاء بمستواها ظنكم وظن الشعب المغربي يمول ماله العام سياسة ثقافية فاشلة وماذا ستفعلون لانصاف بيت الشعر المغرب وتحقيق المساواة وتكافؤ الفرص ينص عليها دستور المملكة وكيف ستردون خلال اجراءات عملية وتدابير ادارية الضجة الفضيحة خلفتها نتائج السنة مجال دعم الاغنية مضيئا اننا نطعن نتائج الدورة برنامج دعم الموسيقى والاغنية واننا ننتظر ستقومون

art-et-culture

1

ظل استمرار حالة الطوارئ الصحية المرتبطة بجائحة كورونا اعلن مهرجان الرباط الدولي لسينما المؤلف تنظيم دورته الخامسة والعشرين الممتدة وتكرم الدورة سيدة المسرح واول وزيرة فنانة تاريخ المغرب المعاصر المرحومة الفنانة ثريا جبران غادرتنا دار البقاء شهر غشت الماضي والتي تركت وراءها تجربة فنية مجال التشخيص تجاوزت حدود المسرح تخوم السينما والتلفزيون وقال المهرجان بيان يجدد اللقاء بمحبي الفن السابع دورته الخامسة والعشرين تصادف السنة اليوبيل الفضي للمهرجان مسار فني حافل بصمت جمعية مهرجان الرباط الدولي للثقافة والفنون موعد قار تنفتح عاصمة المملكة محيطها العربي والافريقي والدولي لتقاسم التجارب الابداعية والتعريف بقيم التسامح والانفتاح والتعايش لطالما ميزت الشعب المغربي وذكر المهرجان انه ظل الظروف الصحية الحالية وعملا بمتطلبات حالة الطوارئ والاجراءات الاحترازية اعلنت عنها بلادنا محاصرة جائحة كورونا تقرر الابقاء الدورة لتكون دورة يتطلبه رصد للامكانيات التقنية والرقمية شانها ضمان تتبع الجمهور الوفي لمختلف انشطة المهرجان بنفس الزخم المعهود الدورات العادية المرتقب تشرف اختيار جوائز المسابقة الدولية للفيلم القصير لجنة يراسها المخرج المغربي عهد بنسودة جانب مبدعين يمثلون ثقافات مختلفة وحساسيات فنية متنوعة تحل كندا ضيفة شرف الدورة سيكتشف عشاق الفن السابع سينما بصمتها صناعة وابداعا تعكس خصوصيات البلد المميز وبالموازاة المسابقة الرسمية للفلام سجل المهرجان بيانه عودة ورشة كتابة السيناريو يؤطرها المخرج الامريكي جيروم غاري مدي ايام موجهة للمهنيين والطلبة والشغوفين بكتابة السيناريو والحاملين لمشاريع افلام روائية طويلة وتماشيا الوضع الصحي يشهده العالم جراء جائحة كورونا المرتقب ينظم مهرجان الرباط الدولي لسينما المؤلف ندوات تقارب الاول مستقبل الصناعة السينمائية ضوء الوضعية الوبائية العالمية وتستشرف الثانية مستقبل الفرحة السينمائية قاعات العرض وسينما البيت بينما ستركز الثالثة حقوق المؤلف الوسائط الجديدة انطلاقا حالة المؤلفات السينمائية واورد المهرجان تكيف ادارته مستجدات الوضعية فرضها كوفيد واصرارها المضي قدما تنظيم الدورة الرقمية نابع بالاساس قناعتها بان السينما اسمي تعبير انشغالات الشعوب وتطلعاتها واحلامها مجملا بان اختيار مواضيع الندوات المقامة هامش المهرجان دليل الفن المواطن يسعي للمساهمة موقعه تجاوز الازمات نوعها

art-et-culture

1

تشير مشاهدة فيلم قصير ضمن الثلاثية الاخيرة للمخرج المغربي هشام العسري جملة المرجعيات الثقافية والادبية والاسطورية العالمية وظفها مكان امراة بطلة تائهة بلاد العجائب عقارب الساعة وهي تدور الخلف صانع القبعات يحاول بجشع بيع يظن انه يصنعه الكلب متربص ينتظر فرصة للانقضاض ضحاياه ضحيته انها المرأة البطلة تمشي بحثا المجهول وكانها تاخذ كلبها نزهة انها حداد موته لربما ببساطة غاضبة مجنونة انها اعقل العقلاء الفيلم القصير بدون

art-et-culture

1

عنف تاليف وسيناريو واخراج هشام العسري ضمن ثلاثية افلام قصيرة التوالي مشاهد تبرز قوة المرأة شخصية البطلة ندي الفيلم وكيف انها تستطيع مواصلة السير رغم الالم والابتسام رغم الاحزان تلتقط الكاميرا مشية انثوية مختالة وانيقة بخطي ثابتة فالحذاء الكعب العالي يؤلم القدمين المرأة تحافظ مقاييسها وعلي راسها مرفوعا ككعب حذائها نظرت لذلك كوكو شانيل ذاتها قائلة وفجأة يصبح الجمال جريمة السهل اكتشاف الشعور بالجنون السهل السير طريق الانتقام معان سخيقة وسريالية خفي الفيلم اعظم يشتري الزهور يفكك معاني ضاعت دلالاتها الايام تحمله ثقل عنف وعادات وتقاليد واحكام يكون الاهم واقع المجتمع المغربي وكيف يريد الكل نريد جميعا فقط نشاهد نتفرج ننظر نرمق نحدق وان نسرق نظرات ونقوم يسمح يسمح بمشاهدته نحب ببساطة حد التبسيطية ننشغل بقصص الآخرين مسيطرين قصصهم لهذا الاهتمام دلالاته العميقة والماساوية فقد تكون القصص يمكن للمغربي يسيطر خلال تحديقه تتوقع وننتظر وننوي دائما الاسوا الناس ومنهم نظن اننا نمتلك نمسك بالحقيقة ايدينا الصفر اية حقيقة وبكل يحوم حولها يمسك يعلم الفيلم لمسة جيدة وابداعية لمسة عميقة بأسلوب الفنان المغربي المقيم الخارج حسن الحجاج المغرب الحقيقي العميق يسميه البعض والهاتف العمومي وصناديق الكوكا كولا القديمة والبلاستيكية الحمراء فرن المايكرويف يمثل تابوتا لجثة الكلب ينتمي لنفس الاسلوب البصري المبدع فن يعرض ويبرز واقعا مؤلما خضر وفواكه ملقاة مكان الناس زالوا جائعين يصفون المرأة البطلة الجميلة والمنتمية لمستوي سيارتها ال والتي بعيدة المنال بالنسبة الالوان والواقعية المؤلمة في الفيلم كعادة افلام هشام العسري نقد اجتماعي واضح لواقع بنيس يتجلى فيلم بدون عنف بصده ذلك خلال اظهار الواقع عاريا جمالية اليس الاسلوب مظلما وظالما لمجتمع باكملة بثقافته وتاريخه وذكاء وابداعات مواطنيه والا الامل يشبه وصف هشام العسري للواقع المغربي فيلمه القصير وصف الروائي الامريكي للمساء الصيفي القرى الانجليزية روايته يشوه الروائي معالم طقوس شاي المساء ويختزلها صورة سوداء وعيثة هشام العسري يضع بصمة مثيرة للجدل واقع مؤلم الغالب واقع مسكوت عنه امل اصلاح مجتمعنا كنا نعتقد الامر يستحق الاصلاح

```
In [22]: shuffled_df = df.sample(frac=1)
```

```
In [23]: shuffled_df.head()
```

اعددنا نعد استحقاق مغرب كورونا حول السؤال تمحورت وجهة نظر قدمتها مؤسسة محمد عابد الجابري للفكر والثقافة حول رؤيتها لمستقبل المغرب تجاوز ازمة جائحة كورونا خلطت اوراق الدول وارغمتها مراجعة حساباتها ووضع مخططات مستقبلية لوضع جديد سيكون شك مغاير للوضع الحالي ومهدت مؤسسة عابد الجابري لتصورها حول مرحلة كورونا بالاشارة الوضع العام بالمغرب وصول الفيروس المملكة متسما بالتخبط تجلي متاعب اقتصادية حقيقية واحتقان اجتماعي وتراجع مستوي الحقوق والحريات العامة وتخبط سياسي وضياح المستوي الثقافي اضافة اشارت المؤسسة التخبط السياسي ينذر بمخاطر ناجمة تفكيك الحياة السياسية وتمييعها علاوة تتابع الحركات المنذرة بانفجار يباغت الجميع كانت الافاق تنسد يوما اخر جاء الوباء وخلق جوا جديدا نجمت عنه دينامية جديدة بمنطقة اخرى واولويات اخرى وتساءلت المؤسسة تحمل اسم ابرز رموز الفكر المغرب المعاصر تدخل جائحة كورونا باب رب ضارة نافعة متوقعة يكون الجواب السؤال بالايجاب لوجود مؤشرات دالة برزت الجائحة اجواء التراجعات والاحتقان ميز قبيل كورونا تراجعت عموما لفائدة ذهنية وتقديرات وسلوكات ومشاعر ايجابية اخرى مقدمتها سلوكات التضامن الوطني ومشاعر المشاركة الوجدانية المناخ الجديد حتمه الوباء العالمي ساعد تقدم السلوكات والمشاعر الايجابية جهة وتراجع اجواء التوتر جهة ثانية تري مؤسسة عابد الجابري لافقة الاجراءات اتخذتها مؤسسات الدولة كانت نقائصها وتعامل السلطات العمومية جائحة كورونا تنم الشعور بالمسؤولية وعن حد الذكاء واليقظة وإدراك اهمية العمل الاستباقي مواجهة الملمات الجو المشوب بالتفاؤل يستبطن سؤالين طرحتهما مؤسسة عابد الجابري الاول يتعلق بمدي الاستعداد لتمثل اهمية العمل الجماعي وفائدة مشاركة القوي الوطنية مواجهة التحديات والثاني يرتبط بالقدرة الانتباه عمق الازمة بلغها المغرب قبيل كورونا وإدراك اننا فرصة تبعدنا الخطر بعيد كورونا الطرف المعني بالاجابة الاستفهامين المذكورين المقام الاول الدولة ومؤسساتها المقام الاول فانهما مطروحان القوي الوطنية وفق تصور مؤسسة عابد الجابري مؤكدة الجميع معني بمصير البلد وبمستقبله وان الجميع معني بموقعه مؤسسات ونخبا واحزابا سياسية ومنظمات نقابية وجمعيات المجتمع المدني وعموم الشعب وللانتقال الوضعية المغرب تزحف جائحة كورونا العالم وضعية افضل نوهت مؤسسة عابد الجابري يجري يؤكد صحة نظرية كون النموذج التنموي البديل ينبغي يركز بالضرورة بناء كتلة تاريخية تتصدى لمهام اعادة بناء الدولة والمجتمع مبرزة الكتلة التاريخية ينبغي تكون نقطة بداية التأسيس للمرحلة الجديدة وذهبت المؤسسة القول الكتلة التاريخية تستطيع تشكل جوابا استثنائيا وضعية استثنائية تعيشها الدولة والمجتمع داعية تتم بلورة جواب الوضعية الاستثنائية الحالية بمقاربة تشاركية تضمن مشاركة الجميع لوضع معالم ومشروع مرحلة كورونا متوقعة يحالف النجاح المسعي صدقت النيات وتوحدت الارادات بناء ديموقراطية حقيقية تجعل الدولة خدمة المجتمع تحقيق الهدف بحسب تصور مؤسسة عابد الجابري ينبغي يبدأ خطوة يمكن تكون عربونا لارادة سياسية جادة فتح الافق وهي اقدام الدولة تنقية الاجواء بالاقدام اطلاق سراح المعتقلين السياسيين ومعتقلي الحركات الشعبية لتمتين الثقة وتعزيز التضامن الوطني تبلور جائحة كورونا

orbites

6

اقترب نادي الاهلي المصري حسم صفقة بدر بانون مدافع الرجاء الرياضي فترة الانتقالات المقبلة قطعت المفاوضات الاهلي واللاعب جهة ومع ادارة الفريق الاخضر جهة ثانية اشواط كبيرة خلال الايام القليلة الماضية وعلمت هسبورت مصادر مطلعة ادارتي الفريقين المصري والمغربي اتفقتا التفاصيل المالية الخاصة بالصفقة لتصير اكتمالها ابرز الانتقالات الدوريين المغربي والمصري سواء ناحية القيمة الرياضية للاعب ناحية المبلغ المالي المنتظر ينعش بشكل كبير خزينة الرجاء تعيش خصاصا فترة طويلة وشجع حسم صفقة مروان الهدودي المنتظر يلتحق بالفريق الاخضر عقب نهاية الموسم الحالي الدفاع الحسني الجديدي النسر السماح برحيل اعمدتها مرشحا للمغادرة فترة البحث افاق جديدة وتحسين دخله السنوي واكدت المصادر ذاتها المفاوضات الطرفين تجري فترة سرية تامة الفريقين معا يخوضان اخر مباريات الدوريين المغربي والمصري ويحضران لمباريات حارقة لحساب نصف نهائي دوري ابطال افريقيا شان صفقة انتقال بانون الاهلي المصري حال اكتمالها رسميا خلق مزيد الجدل الوسط الرياضي المغربي صارت البطولة الوطنية سوقا للاندية المصرية رحيل محمد اونايم واشرف بنشرفي وحמיד احدات وقت سابق لتقوية الاندية المحلية حساب الاندية المغربية خلال المواجهات تجمع الطرفين المستوي القاري ويعد بانون ابرز اللاعبين انجبتهم البطولة الوطنية السنوات الاخيرة قاد الرجاء التتويج باكثر لقب حمل شارة قيادة المنتخب المغربي للاعبين المحليين النسخة الاخيرة مسابقة الشان توج الفريق الوطني بالدار البيضاء

sport

10

متوقعا سلفا اختارت غالبية الاسر المغربية ارسال ابنائها حجرات المدارس خلال الموسم الدراسي المقبل واعتماد التعليم الحضوري الرغم خطر انتشار فيروس كورونا داخل قاعات الدراسة يقابله تطمينات الوزارة بضمان صحة التلاميذ ويواجه القرار الحكومي مجموعة التحديات ستعيق تنزله بشكل يلائم امكانيات المغرب جهة ويواكب الحالة الوبائية الصحية المعقدة جهة اخرى فمعروف غالبية الاسر المغربية ستضطر ارسال ابنائها المدارس بحكم غياب الامكانيات تتيج متابعة الدروس تبدي العديد الاسر المغربية حماسا واضحا بشأن استمرار اعتماد التعليم البعد ولذلك تضطر الاخرى ارسال ابنائها الاقسام تطمينات الوزير امزازي اكد صحة التلاميذ اولي الاولويات ويقف مهنيون تعانيه غالبية المدارس مستوي الاطوار التعليمية يعثر التلاميذ مرحاض مرفق صحي مناسب الماء الشرب وبالحاري يخصص للنظافة والتعقيم متسائلين كانت المؤسسات ستخضع لعملية التنظيف والتعقيم المستمر طيلة اليوم وسجل حسن الحداد نائب رئيس رابطة الدفاع حقوق المستهلكين وزير التعليم تشبث بالدخول المدرسي رغم المؤشرات تنذر بانتشار وباء كورونا وبزيادة عدد الاصابات والوفيات خلال المرحلة المقبلة وابدي الفاعل المدني تخوفه يتحول الدخول المدرسي بؤرة للفيروس حصل عيد الاضحى تحول ماساة ومناسبة لاثارة الالام والالام فاصبحنا نشكو بؤر التجمعات السكنية وبؤر المخالطين وبؤر المعامل والضيعات الفلاحية وبؤر المؤسسات الادارية ووصل الوباء العالم القروي وشدد الحداد تصريحه لهسبريس الوزير امزازي تشبث بالدخول المدرسي رغم مناشدة الاطراف والجهات بتأجيله وليس الغائه يغار مصلحة التلاميذ حقيقة فهم اولياؤهم فهم احوج يكون عودة فلذات اكبادهم مقاعد الدراسة يستريحوا مناوشاتهم وشغبهم وزاد الفاعل المدني الامر

societe

9

481

71

331

يتعلق بطبقة هشة تتعلق بالاطفال وصغار السن والاحداث والشباب يمكن التحكم حركيتهم ونشاطهم الطبيعي كانت الحواجز يمكن وضعها امامهم يثير الاحساس بالخوف انتقال العدوي اوساط التلاميذ والطلبة بكل فئاتهم وكذلك الاطر التعليمية واعتبر المتحدث المطلب الرئيس بالنسبة للرأي العام تاجيل الدخول المدرسي لفترة محددة الاعداد الجيد لهذه المرحلة وتوفير الوسائل والامكانيات شأنها انجاح خطة الوزارة وفي الوقت نفسه حفظ حياة التلاميذ والاطر التعليمية الخطر وتساءل الحداد الصدد سيتم التعامل اولياء التلاميذ بالمؤسسات التعليمية يظنون يرافقون ابناءهم الصغار خصوصا اوقات الدخول والخروج فهل سيخضعون بدورهم لاجراء الفحص وهل سيكون سالمين جمعها غبار حالتهم الصحية يخص البنات المدرسية وفضاءات الاستقبال داخل المؤسسات يضيف المتحدث الملمح العام اقتتار معظم المؤسسات التابعة للتعليم العمومي للتجهيزات الكافية وخصوصا الجانب المتعلق بالمرافق الصحية الكافية والمؤهلة وكذلك الفضاءات الخاصة بالاستراحة

اعرب الدولي المغربي حكيم زياش نجم اياكس امستردام الهولندي لكرة القدم رغبته يتعاقد ناديه الجديد تشيلسي الانجليزي زميله الحالي الحارس الكاميروني اندري اوانا المرشح بقوة لمغادرة الاراضي المنخفضة الصيف المقبل ووفقا لكشفته مصادر ديلي ميل البريطانية المايسترو المغربي يرغب يلحق الحارس اوانا النجم الكاميروني البالغ العمر عاما يريد الرحيل صوب الدوري الانجليزي سواء عبر بوابة تشيلسي توتنهام وازافت المصادر ذاتها انه حال نجح البلوز اللندني التعاقد الحارس اندري اوانا صفقة ضمه ستبلغ مليون استرليني انه يتبقى عقده عامان الفريق الهولندي وتمكن اوانا الحفاظ نظافة شبابه مباراة المسابقات الموسم وقت تحوم العديد الشكوك حول مستقبل حارس تشيلسي الحالي الاسباني كيا اريزابالاغا يرتقب يغادر ناديه الصيف المقبل يذكر النجم المغربي حكيم زياش البالغ العمر عاما سينتقل صفوف تشيلسي الصيف المقبل ينهي الموسم الحالي فريقه الحالي اياكس بعدما انهي البلوز التعاقد معه شهر الماضي مقابل ملايين استرليني

فندت مصالح وزارة التضامن والتنمية الاجتماعية والمساواة والاسرة جاء صفحة مشبوهة موقع التواصل الاجتماعي فيسبوك تدعي صرف الوزارة منحة مالية لمواجهة تداعيات جائحة كوفيد مستعملة بشكل تدليسي الاسم القديم لوزارة التضامن والتنمية الاجتماعية والمساواة والاسرة وبجانبه هويتها البصرية لوغو وكذبت الوزارة توضيح صادر عنها جملة وتفصيلا روجته الصفحة مغالطات لاغراض مبيتة مستغلة ظروف الحجر الصحي فرضته الجائحة موضحة للرأي العام انها باشرت الاجراءات القانونية المعمول الحالات وذلك باخطار مصالح النيابة العامة بالموضوع تتخذ حق مروجي الاخبار تراه مناسبة درءا لكل يفسد المواطنين وعلي مؤسسات الدولة جو التلاحم والتضامن طبع التعبئة الوطنية الشاملة لمواجهة تداعيات فيروس كورونا المستجد ودعت وزارة التضامن والتنمية الاجتماعية والمساواة والاسرة كافة المواطنين والمواطنات اعتماد مصادر اخبار موثوقة والرجوع الموقع الرسمي للوزارة وحساباته مواقع التواصل الاجتماعي للتأكد الاخبار تهم القطاع

```
In [24]: # Saving The Processed Data as Checkpoint
shuffled_df.to_csv('hespress_processed.csv', index=False)
```

```
In [25]: # Define a function to get Words count and Letters count per example and add 2 new columns
def get_word_and_letter_counts(df, column_name):
    # get word and letter counts
    df['word_count'] = df[column_name].apply(lambda x: len(x.split()))
    df['letter_count'] = df[column_name].apply(lambda x: len(x.replace(' ', '')))

    return df
```

```
In [26]: df = get_word_and_letter_counts(shuffled_df, 'story')
```

```
In [27]: df.sample(3)
```


	story	topic	topic_encoded	word_count	letter_count
640	<p>استدعت الجزائر سفيرها بجمهورية كوت ديفوار للتشاور عقب فتح البلد الافريقي قنصلية مدينة العيون المغربية وقالت وزارة الخارجية الجزائرية اليوم الخميس استدعاء السفير الجزائري يأتي ضد تصريحات مبطنة لوزير الشؤون الخارجية الايفواري واعتبرت الخارجية الجزائرية قنصلية كوت ديفوار بالصحراء المغربية تعد ضربا للالتزام الجماعي للبلدان الافريقية المؤسسة للاتحاد الافريقي القاضي بالتمسك بمبادئ المنظمة والعمل لتحقيق الاهداف المكرسة العقد التاسيسي تعلق بضرورة الوحدة والتضامن الدول والشعوب والدفاع السيادة والوحدة الترابية واستقلال الدول الاعضاء وكان علي كوليبالي وزير الاندماج الافريقي وايفواريي الخارج وجه رسائل مسبوقة مدينة العيون اعداء الوحدة الترابية للمملكة اكد كوت ديفوار تتمتع بحرية القرارات وترفض الاملاءات المتعلقة بتوجهها العلاقات الدولية مضيفا فتح قنصلية بالعيون يعتبر قرارا سياديا يتوازي قنوات كوت ديفوار الدبلوماسية المتجذرة التاريخ وقال الوزير الايفواري الندوة الصحافية جمعته بنظيره ناصر بوريطه وزير الشؤون الخارجية والتعاون الافريقي والمغاربة المقيمين بالخارج الاعراف الدبلوماسية لبلادنا نقبل تقديم الدروس والمواظ السياسة نقبل تدخل اطراف اخري قراراتنا السيادة دول حليفة كالمغرب</p>	politique	7	149	930
231	<p>اخرت المحكمة الابتدائية بمدينة اسفي يوم الثلاثاء ملف المتورطين اخفاء شاحنة كانت متنها اطنان الحشيش بمنطقة البدوزة الشهر الجاري وكانت عناصر المركز القضائي باسفي احالت يوم الاحد الماضي النيابة العامة المختصة بعاصمة عبدة اشخاص ضبطوا متلبسين باخفاء اطنان المخدرات كيلومترا مدينة اسفي باتجاه الوالدية كانت الشبكة تستعد لتهربها متن قارب مطاطي وتمكنت عناصر الدرك الملكي بسرية اسفي اجهاض عملية تهريب الكمية المخدرات عبر شواطئ اقليم اسفي توصلها بمعلومات دقيقة حول وجود شاحنة مخبة بالقرب الفنادق بمنطقة الشنينات التابعة لجماعة البدوزة وعلي متنها اطنان المخدرات وعلي اثر كمين نصيبته لمداهمة موقع العملية يوم الجمعة الماضي حجزت عناصر الدرك الكمية الكبيرة المخدرات وقوارب ومعدات اخري واوقفت منفذي عملية التهريب المزمع القيام جنح الليل واوضحت مصادر هسبريس المركز القضائي بالقيادة الجوية للدرك الملكي وسع دائرة ابحاثه ستبقي مفتوحة للوصول متورط القضية وضعت هواتف المشتبه فيهم الخبرة مضيفة الملف ينتظر يطيح برؤوس عديدة اوضحت شبكات تهريب المخدرات تعتمد ضواحي المدن الشاطئية للقيام بعملية تهريب الممنوعات متن قوارب عرض المحيط الاطلسي لتتمكن بعدها البواخر نقلها وجهات متعددة يشار عناصر الدرك الملكي الصعيد الوطني تمكنت خلال الشهر الماضي اجهاض عمليات شبيهة بالتي عرفها اقليم اسفي</p>	faits-divers	3	179	982
836	<p>قالت سفارة جمهورية مصر العربية بالمملكة المغربية السفارة تتلق اية تعليمات جديدة بشأن تعديل قواعد منح التأشيرات ولم يتم اجراء تعديلات القواعد المنظمة لمنح التأشيرات لرعايا المملكة لدخول مصر وذلك تعقيا تداولته وسائل الاعلام بشأن فرض قواعد جديدة لمنح تأشيرات الدخول لمواطني المملكة المغربية الراغبين زيارة مصر اوضحت السفارة بيان صادر عنها توصلت هسبريس مواطني المملكة المغربية مرحب دائما بلدهم الثاني مصر وشروط اعفاء الافواج السياحية الموافقة الامنية يطرا عليها تغيير تسمح القواعد المنظمة لمنح التأشيرات باعفاء الافواج السياحية المغربية بحد ادني افراد شرط الحصول الموافقة الامنية وبرزت سفارة جمهورية مصر العربية القواعد المنظمة لمنح تأشيرات الدخول للمجموعات السياحية تنص تتضمن المجموعة افراد بحد ادني يسافرون برنامج سياحي موحد وعلي رحلة طيران واحدة يتطلب احضار المجموعة يفيد حجز تذاكر الطيران وبرنامج الرحلة وهذه القواعد معمول فترة طويلة ولم يطرا عليها تغيير اكد بيان السفارة التسهيلات الممنوحة للمجموعات السياحية المغربية والافراد الراغبين زيارة مصر وجاذبية المنتج السياحي وتنوعه ساهم تحقيق طفرة التدفق السياحي البلدين خلال الفترة الاخيرة ولذلك تمت زيادة عدد رحلات مصر للطيران مطاري القاهرة ومحمد الخامس بمدينة الدار البيضاء وختمت السفارة ببيانها بدعوة المغاربة المتوجهين مصر الالتحاق مباشرة بالقسم القنصلي بالسفارة لانهاء اجراءات الحصول تأشيرة الدخول خلال اسبوعين تاريخ التقدم ودون اية معوقات طلبات اضافية وباسعار محددة ومعلنة وهي درهما مغربيا للتأشيرة السياحية درهما مغربيا لتأشيرة الزيارة ووفقا للضوابط المعلنة داخل القسم القنصلي وعلي صفحة السفارة موقع التواصل الاجتماعي فيسبوك وعدم التعامل وسطاء وذلك لمنع المتاجرة التأشيرات جانب السماسرة يتحايلون الراغبين السفر ويفرضون عليهم مبالغ مالية كبيرة وجه حق</p>	marocains-du-monde	4	240	1364

```
In [28]: # 1. Total words in the dataframe
total_words = df['word_count'].sum()
print("Words count in the dataframe:", total_words)

# 2. Total Letters in the dataframe
total_letters = df['letter_count'].sum()
print("Letters count in the dataframe:", total_letters)

# 3. Total words by each class
words_by_class = df.groupby('topic')['word_count'].sum()
print("\nWords count per class:")
print(words_by_class)

# 4. Total Letters by each class
letters_by_class = df.groupby('topic')['letter_count'].sum()
print("\nLetters count per class:")
print(letters_by_class)
```

Words count in the dataframe: 3193442
Letters count in the dataframe: 18176260

Words count per class:

topic	
art-et-culture	335340
economie	265836
faits-divers	121037
marocains-du-monde	293131
medias	430330
orbites	496053
politique	267358
regions	179464
societe	259165
sport	180461
tamazight	365267

Name: word_count, dtype: int64

Letters count per class:

topic	
art-et-culture	1868525
economie	1533354
faits-divers	673237
marocains-du-monde	1686830
medias	2437992
orbites	2805323
politique	1548699
regions	1025435
societe	1488042
sport	993393
tamazight	2115430

Name: letter_count, dtype: int64

In [29]: *# Define the function to get the most frequent n-gram for a given row*

```
def get_most_frequent_ngram(row, n):  
    # Tokenize the text in the 'story' column  
    tokens = nltk.word_tokenize(row['story'])  
    # Generate n-grams  
    n_grams = list(ngrams(tokens, n))  
    # Count the frequency of each n-gram  
    n_gram_counts = Counter([tuple(gram) for gram in n_grams])  
    # Get the most frequent n-gram  
    most_frequent_ngram = ' '.join(n_gram_counts.most_common(1)[0][0])  
    # Return the most frequent n-gram as a string  
    return most_frequent_ngram
```

```

# Create a new column with the most frequent 2-gram for each row
df['most_frequent_2gram'] = df.apply(lambda row: get_most_frequent_ngram(row, 2), axis=1)

# Create a new column with the most frequent 3-gram for each row
df['most_frequent_3gram'] = df.apply(lambda row: get_most_frequent_ngram(row, 3), axis=1)

```

```

In [30]: # Get the most frequent n-gram generally
# Concatenate the most_frequent_2gram and most_frequent_3gram columns into a single column
df['most_frequent_ngram'] = df['most_frequent_2gram'] + ', ' + df['most_frequent_3gram']
# Split the n-gram into a list
ngrams_list = df['most_frequent_ngram'].str.split(', ')
# Count the frequency of each n-gram
ngram_counts = Counter([tuple(ngram) for ngram in ngrams_list])
# Get the most frequent n-gram
most_frequent_ngram = ', '.join(ngram_counts.most_common(1)[0][0])
# Print the most frequent n-gram
print("Most frequent (2,3)n-gram:", most_frequent_ngram)

# Get the most frequent n-gram by class
# Group the dataframe by class
grouped = df.groupby('topic')
# Get the most frequent 2-gram for each class
most_frequent_2gram_by_class = grouped.apply(lambda x: get_most_frequent_ngram(x.iloc[0], 2))
# Get the most frequent 3-gram for each class
most_frequent_3gram_by_class = grouped.apply(lambda x: get_most_frequent_ngram(x.iloc[0], 3))
# Combine the most frequent 2-gram and 3-gram for each class into a single column
most_frequent_ngram_by_class = most_frequent_2gram_by_class + ', ' + most_frequent_3gram_by_class
# Print the most frequent n-gram for each class
print("\nMost frequent (2,3)n-gram by class:")
for c, ngram in most_frequent_ngram_by_class.items():
    print(f"Class {c}: {ngram}")

```

Most frequent (2,3)n-gram: وزارة الصحة , كشفت وزارة الصحة

Most frequent (2,3)n-gram by class:

Class art-et-culture: الفنان المغربي , وفق نمط موسيقي

Class economie: فيروس كورونا , الناتج المحلي الاجمالي

Class faits-divers: قسم العناية , قسم العناية المركزة

Class marocains-du-monde: سفارة المملكة , سفارة المملكة المغربية

Class medias: النيابة العامة , المحامي عبد المولي

Class orbites: عابد الجابري , مؤسسة عابد الجابري

Class politique: افراد الجالية , افراد الجالية المغربية

Class regions: حالة باقليم , المديرية الجهوية للصحة

Class societe: الاسر المغربية , غالبية الاسر المغربية

Class sport: الاهلي المصري , الدوريين المغربي والمصري

Class tamazight: النظام الجزائري , عبر التنسيق الوطني

```
In [31]: custom_stop_words =['من',
```

```
'في',  
'على',  
'و',  
'في',  
'يا',  
'عن',  
'مع',  
'ان',  
'هو',  
'علي',  
'ما',  
'اللي',  
'كل',  
'بعد',  
'ده',  
'اليوم',  
'أن',  
'يوم',  
'انا',  
'إلى',  
'كان',  
'ايه',  
'اللى',  
'الى',  
'دي',  
'بين',  
'انت',  
'أنا',  
'حتى',  
'لما',  
'فيه',  
'هذا',  
'واحد',  
'احنا',  
'اي',  
'كده',  
'إن',  
'او',  
'أو',  
'عليه',  
'ف',  
'دى',  
'مين',  
'الي',
```

, 'كانت'
, 'أمام'
, 'زي'
, 'يكون'
, 'خلال'
, 'ع'
, 'كنت'
, 'هي'
, 'فيها'
, 'عند'
, 'التي'
, 'الذي'
, 'قال'
, 'هذه'
, 'قد'
, 'انه'
, 'ريتويت'
, 'بعض'
, 'أول'
, 'ايه'
, 'الان'
, 'أي'
, 'منذ'
, 'عليها'
, 'له'
, 'ال'
, 'تم'
, 'ب'
, 'دة'
, 'عليك'
, 'اي'
, 'كلها'
, 'اللتى'
, 'هى'
, 'دا'
, 'انك'
, 'وهو'
, 'ومن'
, 'منك'
, 'نحن'
, 'زى'
, 'أنت'
, 'انهم'
, 'معانا'
, 'حتي'
, 'وانا'

```

'عنه',
'إلي',
'ونحن',
'وانت',
'منكم',
'وان',
'معاهم',
'معايا',
'وأنا',
'عنها',
'إنه',
'اني',
'معك',
'اننا',
'فيهم',
'د',
'انتا',
'عذك',
'وهي',
'معا',
'آن',
'انتتي',
'وانت',
'وان',
'ومع',
'وعن',
'معاكم',
'معاكو',
'معاها',
'وعليه',
'وانتم',
'واننتي',
'غ',
'|']

```

```

In [32]: words = nltk.word_tokenize(' '.join(df['story'].tolist()))
words = [word for word in words if word.lower() not in custom_stop_words]
word_freq = Counter(words)
most_common_words = word_freq.most_common(30)
# print(most_common_words)

top10_common_words = pd.DataFrame(most_common_words)
top10_common_words = top10_common_words.iloc[1:,:]
top10_common_words.columns = ['most_common_words', 'count']
top10_common_words.style.background_gradient(cmap='Greens')

```

Out[32]:

	most_common_words	count
1	المغربية	10401
2	كورونا	7740
3	الامازيغية	7559
4	المغربي	6545
5	حالة	6371
6	سنة	5943
7	وفي	5770
8	الحكومة	5661
9	محمد	5369
10	الوطنية	5248
11	الوطني	5152
12	عدد	4824
13	المغاربة	4689
14	حول	4653
15	رئيس	4645
16	العام	4643
17	بشكل	4466
18	عبر	4239
19	العامة	4081
20	عيد	3998
21	ذاته	3875
22	العالم	3872
23	الصحية	3736
24	بسبب	3689
25	يمكن	3571

	most_common_words	count
26	العمل	3535
27	القانون	3516
28	المائة	3454
29	الله	3359

```
In [33]: df.sample(2)
```

يرتقب يخوض الدولي المغربي
جواد الياميق لاعب جنوة
الايطالي لكرة القدم تجربة
جديدة بالدوري الاسباني نهاية
فترة اعارته ايام نادي سرقسطة
استقدمه خلال الميركاتو الماضي
وحسب مصادر اسبانية المدافع
المغربي مطلوباً اندية عقب
العروض الكبيرة بصم عليها
بقميص سرقسطة ويتعلق الامر
بهويسكا وقادش صعودهما
الموسم القسم الممتاز الليغا
يدخل البالغ العمر سنة ضمن
خطط نادي غرناطة للموسم
المقبل يريد الفريق الاندلسي
التعاقد معه لتقوية خطه
الخلفي والتوصل اتفاق نهائي
ناديه جنوة يملك عقده نهاية
موسم واكدت المصادر نفسها
جدية المفاوضات نادي هويسكا
بطل السيجوندا ووكيل اعمال
اللاعب الياميق بغية التعاقد
المدافع السابق لفريقي الرجاء
الرياضي واولمبيك خريبكة وكان
اللاعب سالف الذكر قريباً تحقيق
ابرز اهداف الموسم رفقة
سرقسطة ويتجلي العودة
القسم الممتاز الهزيمة الاخيرة
ضد نادي التشي سيلعب مباراة
فاصلة نادي خيرونا لحسم مركز
الصاعد الثالث الليغا اجلت امال
المجموعة الموسم المقبل

205

sport

10

132

715

القسم الممتاز

يرتقب يخوض الدولي

القسم الممتاز، يرتقب
يخوض الدولي

20

orbites

6

279

1657

معبر الكركات

عرقلة حركة السير

معبر الكركات، عرقلة
حركة السير

صفعة اممية جديدة وجهها
الامين العام للأمم المتحدة
انطونيو غوتيريس جبهة
البوليساريو الانفصالية بتحذيرها
خطورة الافعال تقوم مستوي
معبر الكركات الحدودي
وتواصل جبهة البوليساريو
تحضيرها يسمي الاعتصام
الشعبي المعبر الحدودي توقف
عشرات الصحراويين القادمين
مخيمات الرايوني الناحية

	topic	topic_encoded	word_count	letter_count	most_frequent_2gram	most_frequent_3gram	most_frequent_ngram
العسكرية الاولى منطقة تسمي الدكوج ويبدو زيارة الجنرال دوكوردارمي عبد الفتاح الوراق المفتش العام للقوات المسلحة الملكية المنطقة الجنوبية وبالضبط الجدار الامني للوقوف كثب تطورات الوضع اريكت ميليشيات البوليساريو كشفت مصادر متطابقة تفجر انقسام حاد قيادات الجبهة بشأن مخطط الحركة العفوية لاجللاق معبر الكركرات وتدفع قيادات الجبهة بضرورة التراجع الخطوة المرتقبة يمكن تحمل نتائج محسوبة العواقب مستوي الادانة خصوصا قرب صدور تقرير مجلس الامن الدولي نهاية الشهر المقبل تدعو قيادات اخرى بينها المسمي وزير الدفاع محمد لمين البوهالي بضرورة التصعيد لاعادة ملف الصحراء الواجهة الدولية وعبرت مصادر موريتانية امتعاضها الاستفزات المتكررة للبوليساريو مستوي المعبر الحدودي يعتبر منفذا اقتصاديا رئيسيا والي نواكشوط وذكر منبر اعلامي موريتاني التحرك الجديد للبوليساريو موجها الظاهر ضد المغرب موريتانيا تدفع ثمنه بالدرجة الاولي نظرا لاعتماد سوقها المحلي التموين مادة الخضار القادمة عبر الشاحنات الجارة المغرب واوضح المصدر ذاته تعطيل الحركة داخل معبر الكركرات شأنه يضاعف اسعار الخضراوات الاسواق الموريتانية النهاية عقاب مباشر للمستهلك الموريتاني وتحد سافر لقيادتها واستهتار واضح بمصالح البلد وسيادته وتلاعب خطير بامنه الغذائي ودعت الامم المتحدة البوليساريو عدم عرقلة حركة السير المدنية والتجارة							

story	topic	topic_encoded	word_count	letter_count	most_frequent_2gram	most_frequent_3gram	most_frequent_ngram
المنتظمة والامتناع اجراء بشكل تغييرا الوضع القائم المنطقة العازلة بالكركرات وذلك ردا التظاهرات المزمع القيام بههدف عرقلة حركة السير المدنية والتجارية وقال نائب المتحدث باسم الامين العام للامم المتحدة فرحان حق تصريح للسحافة منظمة الامم المتحدة بدعوات انفصاليي البوليساريو واذناها ذوي السوابق الاجرامية للقيام بتظاهرات جديدة بالكركرات بهدف عرقلة حركة السير المدنية والتجارية جدد دعوة الامم المتحدة التحلي باقصي درجات ضبط النفس ونزع فتيل توتر بالكركرات مضيفا بعثة المينورسو تراقب الوضع بالمنطقة							

::: Modeling :::

[Go to Table of Contents](#)

```
In [34]: # Splitting the Dataframe
X = df['story'] # Text data
y = df['topic_encoded'] # Target variable

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=101, shuffle=True) # Shuffling again

In [35]: print('X_train shape',X_train.shape)
print('y_train shape',y_train.shape)

X_train shape (8796,)
y_train shape (8796,)

In [36]: tfidf_vectorizer = TfidfVectorizer()

X_train_tfidf = tfidf_vectorizer.fit_transform(X_train)
X_test_tfidf = tfidf_vectorizer.transform(X_test)
```

```
In [37]: svc_classifier = LinearSVC(multi_class="ovr", random_state=101)
svc_classifier.fit(X_train_tfidf, y_train)

y_pred_svc = svc_classifier.predict(X_test_tfidf)

# Classification report for each class and for the whole test data
print(classification_report(y_test, y_pred_svc))
```

	precision	recall	f1-score	support
1	0.83	0.89	0.86	203
2	0.87	0.89	0.88	200
3	0.95	0.93	0.94	205
4	0.92	0.91	0.91	200
5	0.93	0.90	0.92	200
6	0.72	0.72	0.72	204
7	0.80	0.86	0.83	191
8	0.76	0.80	0.78	178
9	0.75	0.66	0.70	216
10	0.99	0.99	0.99	207
11	0.96	0.96	0.96	196
accuracy				0.86 2200
macro avg	0.86	0.86	0.86	2200
weighted avg	0.86	0.86	0.86	2200

```
In [38]: # Compute accuracy for each class
accuracy_by_class = {}
for i in range(len(df['topic_encoded'].unique())):
    class_name = df['topic'][df['topic_encoded']==i+1].values[0]
    y_true_class = (y_test == i)
    y_pred_class = (y_pred_svc == i)
    accuracy_by_class[class_name] = accuracy_score(y_true_class, y_pred_class)

# Compute overall accuracy
accuracy = accuracy_score(y_test, y_pred_svc)

# Print accuracy for each class and overall accuracy
print("LinearSVC - Accuracy per class:")
for class_name, acc in accuracy_by_class.items():
    print(f"\t{class_name}: {acc:.3f}")
print(f"\nLinearSVC - Overall accuracy: {accuracy:.3f}")
```

LinearSVC - Accuracy per class:

- art-et-culture: 1.000
- economie: 0.973
- faits-divers: 0.977
- marocains-du-monde: 0.989
- medias: 0.985
- orbites: 0.985
- politique: 0.949
- regions: 0.969
- societe: 0.964
- sport: 0.945
- tamazight: 0.998

LinearSVC - Overall accuracy: 0.864

```
In [39]: pipeline = Pipeline([
    ('tfidf', TfidfVectorizer()),
    ('clf', LinearSVC(multi_class="ovr", random_state=101))
])

parameters = {
    'tfidf__ngram_range': ((1, 2), (1, 3)),
    'tfidf__sublinear_tf': (True, False),
    'clf__C': [0.1, 1, 10],
    'clf__penalty': ['l1', 'l2'],
    'clf__loss': ['hinge', 'squared_hinge'],
    'clf__dual': [True, False],
    'clf__class_weight': [None, 'balanced'],
    'clf__max_iter': [1000, 2000],
    'clf__tol': [1e-3, 1e-4]
}

random_search = RandomizedSearchCV(pipeline, parameters, cv=5)
random_search.fit(X_train, y_train)

print("Best parameters:", random_search.best_params_)
```

Best parameters: {'tfidf__sublinear_tf': True, 'tfidf__ngram_range': (1, 2), 'clf__tol': 0.001, 'clf__penalty': 'l2', 'clf__max_iter': 1000, 'clf__loss': 'hinge', 'clf__dual': True, 'clf__class_weight': None, 'clf__C': 10}

```
In [40]: best_model = random_search.best_estimator_

# fit the best model on the entire training set
best_model.fit(X_train, y_train)

# predict on the test set
y_pred_grid = best_model.predict(X_test)
```

```
# evaluate the model
```

```
print(classification_report(y_test, y_pred_grid))
```

	precision	recall	f1-score	support
1	0.85	0.90	0.87	203
2	0.86	0.89	0.87	200
3	0.95	0.93	0.94	205
4	0.92	0.92	0.92	200
5	0.93	0.92	0.92	200
6	0.73	0.72	0.73	204
7	0.80	0.87	0.83	191
8	0.78	0.81	0.79	178
9	0.79	0.68	0.73	216
10	0.99	0.99	0.99	207
11	0.96	0.96	0.96	196
accuracy			0.87	2200
macro avg	0.87	0.87	0.87	2200
weighted avg	0.87	0.87	0.87	2200

```
In [41]: # Compute accuracy for each class
```

```
accuracy_by_class = {}
```

```
for i in range(len(df['topic_encoded'].unique())):
```

```
    class_name = df['topic'][df['topic_encoded']==i+1].values[0]
```

```
    y_true_class = (y_test == i)
```

```
    y_pred_class = (y_pred_grid == i)
```

```
    accuracy_by_class[class_name] = accuracy_score(y_true_class, y_pred_class)
```

```
# Compute overall accuracy
```

```
accuracy = accuracy_score(y_test, y_pred_grid)
```

```
# Print accuracy for each class and overall accuracy
```

```
print("GridSearchCV/LinearSVC - Accuracy per class:")
```

```
for class_name, acc in accuracy_by_class.items():
```

```
    print(f"\t{class_name}: {acc:.3f}")
```

```
print(f"\nGridSearchCV/LinearSVC - Overall accuracy: {accuracy:.3f}")
```

```
GridSearchCV/LinearSVC - Accuracy per class:  
  art-et-culture: 1.000  
  economie: 0.975  
  faits-divers: 0.976  
  marocains-du-monde: 0.988  
  medias: 0.985  
  orbites: 0.986  
  politique: 0.950  
  regions: 0.970  
  societe: 0.966  
  sport: 0.951  
  tamazight: 0.998
```

GridSearchCV/LinearSVC - Overall accuracy: 0.870

::: Final Results :::

[Go to Table of Contents](#)

Since the tuned LinearSVC performed better , lets interpret it's results

accuracy: The accuracy is the proportion of correctly classified samples out of the total number of samples in the test set.

precision: The precision is the ratio of true positives (TP) to the sum of true positives and false positives (FP). It measures the proportion of positive predictions that are correct. A high precision means that the model is making few false positive predictions.

recall: The recall is the ratio of true positives to the sum of true positives and false negatives (FN). It measures the proportion of actual positives that are correctly identified by the model. A high recall means that the model is correctly identifying many of the positive cases.

f1-score: The f1-score is the harmonic mean of precision and recall, and it provides a balanced measure of the model's accuracy. It ranges between 0 and 1, with a higher value indicating better performance.

macro avg: The macro average is the average of the precision, recall, and f1-score across all classes, weighted equally. It provides an overall summary of the model's performance across all classes, giving equal weight to each class regardless of its size.

weighted avg: The weighted average is the weighted average of the precision, recall, and f1-score across all classes, weighted by the number of samples in each class. It provides a summary of the model's performance that takes into account class imbalance.

The weighted average gives more weight to classes with more samples, while the macro average treats all classes equally. Which method to use depends on the specific problem and the goals of the analysis. If all classes are equally important, macro average may be more appropriate. If some classes are more important than others, weighted average may be more appropriate.

- **art-et-culture > Class 1:** Precision=0.85, Recall=0.9, F1-score=0.87
- **economie > Class 2:** Precision=0.86, Recall=0.89, F1-score=0.87
- **faits-divers > Class 3:** Precision=0.95, Recall=0.93, F1-score=0.94
- **marocains-du-monde > Class 4:** Precision=0.92, Recall=0.92, F1-score=0.92
- **medias > Class 5:** Precision=0.93, Recall=0.92, F1-score=0.92
- **orbites > Class 6:** Precision=0.73, Recall=0.72, F1-score=0.73
- **politique > Class 7:** Precision=0.80, Recall=0.87, F1-score=0.83
- **regions > Class 8:** Precision=0.78, Recall=0.81, F1-score=0.79
- **societe > Class 9:** Precision=0.79, Recall=0.68, F1-score=0.73
- **sport > Class 10:** Precision=0.99, Recall=0.99, F1-score=0.99
- **tamazight > Class 11:** Precision=0.96, Recall=0.96, F1-score=0.96

Accuracy: The overall accuracy of the classifier is 0.87, which means that 87% of the samples in the test set were correctly classified.

Precision: A precision of 0.85 for class 1 means that 85% of the samples predicted as class 1 were correctly classified as class 1. The precision is high (above 0.8) for most classes, indicating that the classifier has relatively few false positives.

Recall: A recall of 0.9 for class 1 means that 90% of the actual samples of class 1 were correctly classified as class 1. The recall is high (above 0.8) for most classes, indicating that the classifier has relatively few false negatives.

F1-score: The F1-score is high (above 0.8) for most classes, indicating that the classifier has good overall performance.

Overall, the classifier performs well on most classes, with high precision, recall, and F1-score. Some classes (e.g. 6 and 9) have lower precision and recall, indicating that the classifier has more difficulty distinguishing between these classes. The overall accuracy of 0.87 is also relatively high, indicating that the classifier is able to classify most samples correctly.

Possible enhancements that could improve the performance of the classifier:

1. Performing more in-depth analysis of misclassified instances can potentially improve the classifier's accuracy and generalization ability.
2. Asking Domain Experts for An Custom Stopword List.
3. Trying different text preprocessing techniques, such as stemming or lemmatization.
4. Using more advanced vectorization techniques, such as word embeddings or character-level embeddings.

5.Data augmentation: The current classifier uses only the available labeled data for training and evaluation. However, it may be possible to generate additional labeled data through techniques such as data synthesis or data annotation, which could improve the performance of the classifier.

6.Ensemble methods: The current classifier uses a single model to make predictions. However, it may be possible to improve performance by combining the predictions of multiple models using ensemble methods such as bagging or boosting.

If you find this notebook useful or interesting, please, support with an upvote :)

- If you have read this far, I want to thank you for taking the time to engage with my work.
- Whether you found it insightful or not, or if you noticed areas where I could improve or mistakes i have done, I will be grateful for your feedback.
- I believe that your feedback is essential for my personal and professional growth, and I would welcome any comments or suggestions that you may have.

If you enjoyed reading this kernal and are curious about my other work, I invite you to check out my other projects:

1. [Complete Hyperparameter Tuning 10 Classification Machine Learning Algorithms to Predict Heart Disease](#)
2. [Time Series - Sales Forecasting - Hyperparameter Tuning 4 evaluated Models](#)
3. [Market Segmentation Clustering - K-Means - Mean Shift - Hierarchical - DBSCAN - Spectral - Affinity Propagation](#)
4. [A/B Testing : Which version of the website should you use?](#)
5. [E-Commerce Sales Exploratory Data Analysis - Time Series Forecasting - Part 2](#)
6. [E-Commerce Price Predictive Modeling using 12 Regression Machine Learning Algorithms Full Explained with Complete Hyperparameter Tuning - Part 3](#)
7. [Customer Churn Prediction - Feature Selection and Hyperparameter Tuning - 4 Tree Based Algorithms](#)
8. [Complete 9 Dimensionality Reduction Techniques Comparison using the Kannada MNIST Data](#)

Follow me on [LinkedIn](#) and [kaggle](#) for upcoming tutorials and projects

Finally, I will appreciate your upvote and thank you again for your support.

In []: