



الجمهورية العربية السورية

جامعة دمشق

كلية الهندسة المعلوماتية

قسم الذكاء الصناعي واللغات الطبيعية

## Intelligent Content Management System

مشروع أُعدَّ لنيل درجة الإجازة في الهندسة المعلوماتية

بإشراف:

د. ناديا البقاعي

إعداد الطلاب:

محمد سامي العش

محمد علاء خير الله

مهند الطباع

معاذ الخضري

آب – 2021



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ  
وَقُلِ اعْمَلُوا فَسَيَرَى اللَّهُ عَمَلَكُمْ وَرَسُولُهُ وَالْمُؤْمِنُونَ وَسَتُرَدُّونَ إِلَىٰ عِلْمِ الْغَيْبِ  
وَالشَّهَادَةِ فَيُنَبِّئُكُم بِمَا كُنتُمْ تَعْمَلُونَ

إلهي لا يطيب الليل إلا بشكرك ولا يطيب النهار إلا بطاعتك.. ولا تطيب اللحظات إلا بذكرك.. ولا تطيب الآخرة  
إلا بعفوك.. ولا تطيب الجنة إلا برؤيتك

ونحن نقف على عتبة التخرج نتقدم بأسمى آيات الشكر والامتنان والتقدير والمحبة.

إلى الذين يحملون أقدس رسالة في الحياة.

إلى الذين مهدوا لنا طريق العلم والمعرفة.

إلى جميع أساتذتنا الأفاضل.

ونخص بالشكر الجزيل الدكتورة **ناديا البقاعي** التي أشرفت على هذا المشروع وتابعت تطوراتها، لك منا كل

الاحترام والتقدير.

كما ونتوجه بالشكر الجزيل للمهندس **إياد الشامي** الذي لمسنا فيه أسمى تجليات الإنسانية عطاءً ونزاهةً وسمواً  
وأخلاقاً.

## أهدي هذا العمل

إلى التي أسكنتني قلبها فغمرتني بحبها ... إلى القمر الذي أضاء ظلام عقلي وأضاء طريقي ... إلى الشمس التي أذابت جمود قلبي وفجرت به ينابيع الحياة ... إلى التي ما فارقتني بعطفها وأمانها طوال السنين  
أمي الحبيبة

إلى من اشتعل رأسه شيباً وانحنى ظهره كي نستقيم نحن ... إلى اليد التي جرحت بأشواك دربي لتفرشه زهوراً ...  
إلى من استمد من ثباته قوتي، ومن اسمه فخري وعزتي  
أبي الحبيب

إلى من شدد الله به عضدي ... إلى من أرى فيه نفسي ... إلى شمسي وقمري ... إلى قلبي وملاكي  
أخي عمر

إلى التي من نظراتها استمد الأمل ... إلى الوردة الفواحة التي تعطر قلبي ... إلى زينة حياتي ومصدر سعادتي  
صغيرتي ماسة

إلى من بابتسامها أرى الحياة ... إلى المحبة التي لا تنضب والخير الذي لا ينتهي ... إلى نبض القلب  
أختي راما

إلى من تعثرت أبادينا معاً ... إلى من شاركتني طفولتي وربيعان شبابي، تفوقي وسعادتي ... توأم روحي ورفيقة  
دربي ... إلى صديقة قلبي وأوقاتي  
أختي رنيم

إلى أجمل الأقدار التي وهبني الله إياها ... إلى البعيدين عن عيني المقيمين في قلبي ... إلى أبطالتي وجيشي الأوحـد  
جدي الغالي – بيان

إلى من سار معي في هذا الدرب الطويل ... إلى من شاركني تعب المسير ... إلى من رسمت هذا الحلم معهم  
أعضاء الإبابا

إلى أصدقاء السنوات، من صنعت معهم أجمل الذكريات ... إلى القلوب التي أحبت وتآخت ولا تزال  
إبراهيم – موفق – ورد

مهند الطباع

أهدي هذا العمل

إلى مثلي الأعلى صاحب القلب الكبير ....  
إلى من علّمني حُبَّ الحياة وكان قُدوتي فيها ....  
إلى من حصد الأشواك عن دربي ليُمهّد لي طريق العلم والحياة ....  
إلى من علّمني المثابرة والسعي دوماً إلى الأفضل مهما كانت الظروف ومهما واجهتني الصعاب..  
إلى من حملت رايته بُكلِّ افتخار ستبقى كلمّاتك نُجوماً أهتدي بها في دربي طوال حياتي ....  
رحم الله روحك التي رحلت عن الدنيا وما زالت حيّة في قلوبنا ... رحم الله ضحكاته لا تنسى وملامحه لا تغيب  
وحديثاً اشتقنا لسماعه وجمعنا بك في الفردوس الأعلى.

والدي الغالي

إلى بلسم الشفاء وصاحبة القلب الرقيق ....  
إلى التي غمرتني بفيض حنانها ....  
إلى النور الذي أثار دربي والسراج الذي لا ينطفئ نوره أبداً ....  
إلى من زرعت لديّ روح المثابرة والاجتهاد ....  
إلى من بذلت جهد السنين من أجل أن اعتلي سلالم النجاح ....  
إلى من ربّنتني صغيراً ونصحتني كبيراً ... فُرّة عيني وفوادي أُمي الغالية أطل الله في عمرها.

والدتي الغالية

إلى من بهم يشتدّ عضدي وتعلو همّتي .... إلى من بهم أكبر وعليهم أعتمد .... إلى من شاركوني الحياة بخلوها ومُرّها  
.... إلى من وقفوا بجانبني بكل خطوةٍ خطيتها.

أختي المهندسة الغالية راما وأخي الغالي كنان

إلى من جمعني بهم منبر العلم والصداقة .... إلى من رافقوني طوال مشواري الدراسي .... إلى من أكنّ لهم  
أسمى عبارات المحبة والاحترام.

زملائي وأصدقائي الأعزاء

محمد سامي العش

أهدي هذا العمل المتواضع

إلى:

❖ والدين الكريمين حفظهما الله

❖ وإلى كل أفراد أسرتي

❖ إلى كل الأصدقاء، ومن كانوا برفقتي ومصاحبتني أثناء دراستي في  
الجامعة

❖ وإلى كل من لم يدخر جهداً في مساعدتي

❖ وإلى كل من ساهم في تلقيني ولو بحرف في حياتي الدراسية

محمد علاء خير الله

## أهدي هذا العمل

إلى من زرعت الحب والأمل في قلبي منذ سنين.. إلى خيط الأمل الذي ينير لي المستقبل.. إلى من غمرتني عطفاً وحناناً.. إلى الحزن الدافئ الذي ألجأ إليه دوماً في لحظات قوتي وضعفي.. إلى من تعطيني الحب والأمان دون مقابل.. إلى من كان دعاؤها دوماً سر نجاحي.. إلى من وجودها هو الحياة.. إلى نعمة سألهم الله عليها ما حييت ولن أوفيهما حقها.. إليك يا كل الحياة.

أمي ...

إلى العظيم بهيبته ووقاره.. إلى قوتي الأولى.. إلى من علمني القيم والمبادئ وزرع بذور الأخلاق وحب العلم في قلبي.. إلى من رفعت رأسي عالياً افتخاراً به.. إلى من سعى وشقى دوماً لأنعم بالراحة والهناء.. إلى من حصد الأشواك عن دربي ليمهد كل الطرق أمامي.. إلى من آمن بقدرتي على تجاوز كل الصعوبات.. إلى السند والقوة في الحياة.. إليك يا سندي.

أبي ...

إلى من مهد الطريق لخطواتي الأولى وأنار أولى الشمعات في دربي..

إلى من كانت نصائحه زاداً أعتمد عليه في المستقبل.

كل من علمني ...

إلى من هم امتداد روحي.. إلى من بهم أكبر وعليهم أعتمد.. إلى من كانوا إلى جانبي دائماً.. إلى ألوان حياتي ورياحين حديقتي.. إلى من هم أجمل ما في الحياة.. إلى السند والعون عند الحاجة.

إخوتي وعائلتي ...

إلى شركاء دربي.. إلى من أضافوا للحياة محبة وضحكات هونت صعوباتها ومشقاتها.. إلى ينباع الصدق الصافي.. إلى من معهم سعدت، وبرفقتهم في دروب الحياة الحلوة والحزينة سرت.. إلى من تعلمت منهم الكثير لكم مني كل المحبة والتقدير.

أصدقائي الأعزاء ...

معاذ الخضري

## الملخص التجريدي

يُعد اقتصاد صُنَاع المحتوى سوقاً ضخماً غير مُستكشف إلى حدٍّ كبير. حيث أصبح هناك الكثير من الوسائل الرقمية المساعدة للمؤثرين في إدارة أعمالهم. ويرى كبار المستثمرين في وادي السيليكون على نحو متزايد أن صناعة المحتوى الرقمي هي المسار المالي التالي الذي يجب استغلاله على الإنترنت.

وعند الحديث عن المحتوى العربي على الإنترنت بلغة الأرقام سنجد الأرقام والنسب متواضعة جداً، فعلى الرغم من أن اللغة العربية ذات تاريخ عريق، وواحدة من أكثر اللغات من حيث عدد المتكلمين، ولها يوم دولي يُحتفى بها وتعد إحدى اللغات الرسمية للأمم المتحدة، فإن حضورها الإلكتروني يكاد يكون ضئيلاً للغاية، وذلك على الرغم من أن عدد مستخدمي الإنترنت العرب يبلغ 96 مليون شخص.

وتشير معظم التقارير العالمية التي تهتم بالمحتوى الرقمي على الإنترنت أن نسبة تمثيل اللغة العربية على الشبكة لا يتناسب مع أهمية اللغة وعدد مستخدميها ومتصفحي الإنترنت من العرب، فقد أشار تقرير صادر عن الأمم المتحدة أن المحتوى العربي على الإنترنت لا يمثل سوى 3 بالمائة من إجمالي المحتوى العالمي.

ومن أهم العوامل التي تساهم في ضعف المحتوى العربي، وجود عدد كبير من الأدوات المتطورة التي صدرت بلغة إنجليزية ولا تدعم اللغة العربية ككثير من أدوات إنشاء المواقع أو إدارة المحتوى وغيرها.

لذلك انطلقنا للعمل على تسخير تقنيات الذكاء الصناعي لبناء نظام يحوي العديد من الأدوات المتطورة للمساهمة في دعم المحتوى العربي لأننا نؤمن بمهارات وقدرات شعبنا العربي، فلقد قمنا بتطوير نظام لإدارة المحتوى يحوي كافة الأدوات الأساسية اللازمة مع وجود إضافات ذكية ليصبح النظام النهائي يضاهي أنظمة إدارة المحتوى الأجنبية ويساهم في تخطي الصعوبات والمشاكل التي تواجه صناع المحتوى العرب ويساهم في نشر المحتوى العربي بأكبر قدر. حيث تم إضافة عدد من الخدمات وأهمها:

- استخراج الكلمات والعبارات المفتاحية للنصوص بشكل تلقائي: ساهمت هذه الخدمة في تسهيل صناعة المحتوى من حيث كلفة الوقت والجهد اللازم لصياغة الكلمات المفتاحية لكل محتوى يقوم المحرر بكتابته.
- توصيف الصور باللغة العربية: ساهمت هذه الخدمة في استخراج توصيف تلقائي للصور باللغة العربية وتضمن ذلك التوصيف ضمن صفحة الويب مما يساهم في رفع كفاءة محركات البحث الخارجية والداخلية (المُضمَّنة ضمن الموقع).
- التعرف على الأشخاص داخل الصور: نستطيع من خلال هذه الخدمة تحديد هوية الأشخاص داخل الصور وتضمن هوياتهم ضمن التوصيف مما يرفع كفاءة البحث أيضاً.



- المصحح التلقائي باللغة العربية: ساهمت هذه الخدمة في رفع كفاءة وجودة المنشورات باللغة العربية وخصوصاً عند صُناع المحتوى المبتدئين.

- البحث بعدة لغات: ساهمت هذه الخدمة في رفع كفاءة البحث من خلال السماح بتحديد لغة البحث سواء لغة عربية أو انجليزية للحصول على نتائج أكثر دقة.

وهكذا تكون هذه الخدمات أدّت دوراً مهماً في تطوير عمليات البحث وجعلها أكثر كفاءة للوصول للمحتوى من قبل محرّكات البحث الخارجية أو الداخلية المضمنة داخل الموقع نفسه، بالإضافة لمساعدة صُناع المحتوى على تحرير المحتوى وإدارته بشكل ذكي.

# جدول المحتويات

1	جدول المحتويات
5	جدول الرسوم التوضيحية
7	جدول الجداول
8	جدول المصطلحات
9	مقدمة
9	تمهيد
9	أهداف المشروع
9	تعريف المشروع
11	الفصل الأول: الدراسة النظرية
11	Page rank
13	Part of speech tags
14	Machine learning
15	Neural Networks
17	CNN
20	Recurrent Neural Networks
22	تدريب الشبكات العصبونية التكرارية
23	Long-Short Term Memory (LSTM)
26	Word2Vec
26	CBOW
27	Skip-gram model
28	الفصل الثاني: الدراسة المرجعية

28	1- مشاريع مشابهة
28	النظام WordPress
31	النظام Joomla
31	النظام Drupal
32	النظام Contentful
33	نظام Strapi
34	نظام Wagtail
34	نظام Apostrophe
35	2- استخراج العبارات المفتاحية
36	2.1- Unsupervised methods
36	2.1.1 Graph based Methods
43	2.1.2 Statistical Methods
48	2.2- Supervised methods
48	Automatic Keyphrase Extraction (KEA)
50	3- توصيف الصور
50	منهجيات من الأدنى إلى الأعلى
51	منهجيات من الأعلى إلى الأدنى
53	منهجيات هجينة
55	4- التعرف على الوجوه
55	dynamic way
58	Static way
59	الفصل الثالث: الدراسة التحليلية
59	المتطلبات الوظيفية
61	المتطلبات الغير وظيفية
61	فئات المستخدمين

62	مخطط حالات الاستخدام
62	حالات الاستخدام ذات التوصيف العام
63	حالات الاستخدام ذات التوصيف المفصل
69	توصيف واجهة الاستخدام
69	الأنساق المعمارية وأنساق التصميم المستخدمة
70	الأنظمة الجزئية المكونة
70	توزيع الأنظمة الجزئية على العتاد
70	تخزين المعطيات الدائم
71	الفصل الرابع: الدراسة التصميمية
71	الاستخراج التلقائي للعبارات المفتاحية
72	PV-DM
72	PV-DBOW
73	التعرف على الوجوه
74	توصيف الصور
77	التطبيق
77	قراءة النصوص OCR
78	المصحح التلقائي للأخطاء
78	التحقيق
78	اللغة المستخدمة
78	بيئة التطوير المستخدمة
79	أهم المكاتب المستعملة
80	الفصل الخامس: الاختبارات والنتائج
80	مقاييس التقييم
80	مقاييس تقييم توليد شرح تلقائي للصور
82	مقاييس تقييم استخراج الكلمات المفتاحية

83	التعرف على الوجوه.....
84	توصيف الصور.....
87	استخراج العبارات المفتاحية.....
90	خدمة قراءة النصوص.....
90	المصحح التلقائي.....
91	الخاتمة.....
92	المراجع.....

## جدول الرسوم التوضيحية

رسم توضيحي 1- حساب وزن العقدة e	11
رسم توضيحي 2- مثال عن ال POS tagging	13
رسم توضيحي 3- مخطط لشبكة عصبونية بسيطة	16
رسم توضيحي 4- ارتباط الصورة مع العصبونات في طبقة ال Convolution	18
رسم توضيحي 5- تطبيق مرشح على جزئين من الصورة stride = 1	18
رسم توضيحي 6- تطبيق مرشح على جزء من الصورة	18
رسم توضيحي 7- آلية تصغير الصورة في طبقة ال pooling	19
رسم توضيحي 8- آلية عمل ال max pooling	19
رسم توضيحي 9- شبكة عصبونية تلافيفية	20
رسم توضيحي 10- أنواع الشبكات العصبونية التكرارية	21
رسم توضيحي 11- مخطط يوضح بنية الشبكة العصبونية التكرارية	21
رسم توضيحي 12- مخطط يوضح الحالة الخفية في الشبكة التكرارية	21
رسم توضيحي 13- البنية العامة لشبكة LSTM	23
رسم توضيحي 14- موقع ذاكرة الشبكة State Cell	23
رسم توضيحي 15- بوابة النسيان في شبكة LSTM	23
رسم توضيحي 16- المرحلة الثانية من مراحل LSTM	24
رسم توضيحي 17- المرحلة الثالثة من مراحل LSTM	24
رسم توضيحي 18- المرحلة الرابعة من مراحل LSTM	25
رسم توضيحي 19- بنية نموذج CBOW	26
رسم توضيحي 20- بنية نموذج Skip-gram	27
رسم توضيحي 21- استخراج العبارات المفتاحية بطريقة Text Rank	35
رسم توضيحي 22- يوضح خطوات عمل خوارزمية ال Topic rank	38
رسم توضيحي 23- بيان مبني بطريقة ال position rank	42

رسم توضيحي 24- أفضل العبارات مع وبدون MMR.....	46
رسم توضيحي 25- آلية التدريب والاستخراج في KEA .....	48
رسم توضيحي 26- مثال يوضح نتائج التعرف على الأغراض .....	50
رسم توضيحي 27- البنية العامة لنموذج Show and tell.....	51
رسم توضيحي 28- البنية العامة لنموذج ATT.....	53
رسم توضيحي 29- image gradients to keypoint descriptor .....	55
رسم توضيحي 30- اتجاه سهم تحديد القتامة.....	55
رسم توضيحي 31- تحويل صورة إلى فضاء HOG .....	56
رسم توضيحي 32- مقارنة بين صورتين في فضاء HOG .....	56
رسم توضيحي 33- تعديل وضع الوجه بالاستفادة من face landmark estimation .....	57
رسم توضيحي 34- حالات الاستخدام ذات التوصيف العام.....	62
رسم توضيحي 35- بنية نموذج PV-DM .....	71
رسم توضيحي 36- بنية نموذج PV-DBOW .....	72
رسم توضيحي 37- بنية آلية الاهتمام، حيث H تعبر عن الحالة الخفية الناتجة عن فالك التميز في اللحظة السابقة، E(I) تعبر عن مصفوفة سمات أجزاء الصورة (ناتج تطبيق المرمز E على الصورة I)، وهي تتكرر ذاتها في كل لحظة. الطبقات FC1، FC2، FC3 لها أوزان يتم تعلمها أثناء التدريب.....	75
رسم توضيحي 38- بنية فالك التميز. حيث H تعبر عن الحالة الخفية، W تعبر عن الكلمة المولدة في كل لحظة، E(I) تعبر عن مصفوفة سمات أجزاء الصورة (ناتج تطبيق المرمز E على الصورة I) و ATT تُمثل بنية الاهتمام...75	75
رسم توضيحي 39- آلية عمل Teacher Forcing .....	76

## جدول الجداول

- جدول 1: يوضح أنواع مختلفة من المصطلحات المعجمية مع رموزها المختصرة وبعض الأمثلة لكل منها. .... 13
- جدول 2: خرج عملية تدريب Skip-gram ..... 27
- جدول 3: نتائج أبحاث مختلفة في مهمة وصف الصور ..... 54
- جدول 4: المتطلبات الوظيفية ..... 60
- جدول 5: حالة استخدام استخراج العبارات المفتاحية ..... 63
- جدول 6: حالة استخدام توصيف الصور ..... 64
- جدول 7: حالة استخدام تعديل صلاحيات المسؤولين ..... 65
- جدول 8: حالة استخدام التعرف على الوجوه ..... 66
- جدول 9: حالة استخدام إضافة وجه جديد ..... 67
- جدول 10: حالة استخدام المصحح التلقائي للخطأ ..... 68
- جدول 11: مقارنة بين النتائج التي تم تحقيقها ونتائج الأعمال السابقة في عملية التعرف على الوجوه ..... 83
- جدول 12: مقارنة بين النتائج التي تم تحقيقها ونتائج الأعمال السابقة ..... 84



## جدول المصطلحات

رمز الاختصار	الانكليزي	العربي
POS	Part of speech	
KEA	Automatic Keyphrase Extraction	
NN	Neural Networks	الشبكة العصبونية
TFIDF	Term Frequency — Inverse Document Frequency	
DNNs	Deep Neural Networks	الشبكات العصبونية العميقة
CNN	Convolutional Neural Network	الشبكات العصبونية التلافيفية
RELU	Rectified Linear Units Layer	
NER	Named entities recogniser	
MMR	Maximal Marginal Relevance	
API	application programming interface	واجهة التخاطب البرمجية
MySQL		نظام إدارة قواعد البيانات
Seq2Seq	Sequence to Sequence	
SOS	Start of sentence	
EOS	End of sentence	
LSTM	Long Short-Term Memory	
RNN	Recurrent Neural Networks	الشبكات العصبونية التكرارية
HOG	Histogram of Oriented	
CBOW	continuous bag of words	
PV-DM	Paragraph Vector: A distributed memory model	
PV-DBOW	Paragraph Vector without word Distributed bag of words ordering	
REST	Representational state transfer	
OCR	Optical character recognition	قراءة النصوص
PKE	python keyphrase extraction	

## مقدمة

### تمهيد

أدى تطور التكنولوجيا وسهولة وصول المستخدمين للإنترنت إلى استعمال الإنترنت في مجالات لم يكن ليستعمل فيها من قبل مثل كاستعماله للتسويق وعرض الأخبار وغيرها الكثير، وترافق هذا الأمر مع زيادة كبيرة في أعداد المستخدمين، وبالتالي زيادة كبيرة جداً في حجم البيانات التي يتم إضافتها وتداولها كل يوم، وبحسب بعض الدراسات، أنتجت 90% من البيانات في العالم في السنتين الماضيتين، ويُتوقع أن تزيد بمقدار 40% سنوياً.

ترافق الزيادة في حجم البيانات إلى زيادة الحاجة لتخزين وإدارة هذه البيانات، ونتيجة هذه الحاجة ظهرت أنظمة إدارة المحتوى.

مع تطور الذكاء الصناعي وإدخاله في مجالات حياتية كثيرة كان من ضمن المجالات التي تدخل فيها وحال تحسينها أنظمة إدارة المحتوى.

### أهداف المشروع

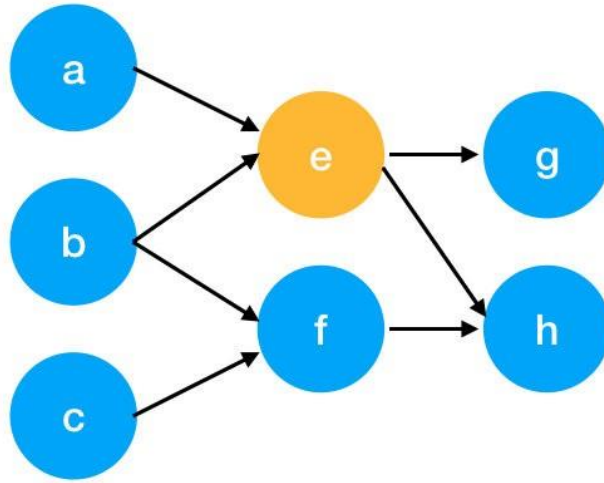
أكثر من 80% من البيانات التي نقوم بتوليدها بشكل يومي هي بيانات غير معنونة أو منظمة بطريقة يمكن الاستفادة منها للحصول على معلومات، مما يجعل عمليات التحليل والمعالجة للاستفادة منها أصعب بكثير. حيث نجد أن قطاعات العمل بحاجة ماسة لعملية مؤتمتة لتحليل وفهم تلك البيانات والاستفادة منها. وخاصة في البيانات الموجودة باللغة العربية حيث نجد أنها تفتقر بشكل كبير لأنظمة قادرة على تحليل تلك البيانات وعنونتها للاستفادة منها بمجالات واسعة.

يهدف المشروع لإيجاد النظام العربي الذكي الأول المتكامل لإدارة محتوى الويب وبدعم كامل للغة العربية ليكون بدوره المشروع الأول في تقديم الخدمات المتكاملة لإدارة محتويات الويب المختلفة وباللغة العربية.

### تعريف المشروع

نظام ذكي يقدم خدمات إدارة المحتوى وباللغة العربية، فهو يقدم مجموعة من الخدمات الأساسية كعملية التخزين للبيانات وعرض المنشورات السابقة مع تاريخ نشرها والشخص الذي قام بعملية النشر، وكذلك يسمح النظام بإضافة أكثر من مسؤول ويتم تحديد صلاحيات لكل شخص من هؤلاء الأشخاص المسؤولين، كما أنّ النظام تميز بعدد من الخدمات عن باقي أنظمة إدارة المحتوى كخدمة استخراج الكلمات المفتاحية من النصوص العربية. فقد لاحظنا افتقار الأنظمة المشابهة لدعم اللغة العربية. حيث يقوم النظام باستخراج الكلمات المفتاحية بشكل تلقائي

فور انتهاء المستخدم من كتابة المنشور ويعرض له قائمة بالكلمات المفتاحية المناسبة ليختار منها ما يناسبه. بالإضافة لخدمة توليد شرح مناسب للصور وباللغة العربية، والتي يستطيع المستخدم الاستفادة منها وإضافة هذا التوصيف لصفحة البرمجة الخاصة بصفحة الويب والتي ستمنح محركات البحث قدرة أكبر على تحديد محتوى الموقع، كما ويقدم النظام المقترح خدمة التعرف على وجوه الأشخاص الموجودين ضمن صورة وتعتبر هذه الخدمة مكملية لخدمة توليد شرح للصورة، ويتميز النظام بالقدرة على البحث عن محتوى معين باللغتين العربية والإنكليزية، كما أنه تم إضافة مجموعة من الخدمات التي تساعد المحرر في عملية الكتابة وتمنحه الدقة والسرعة فتم إضافة مصحح تلقائي لتصحيح الأخطاء الإملائية وإعطاء اقتراح لتصحيحها، كما ويملك المستخدم القدرة على إضافة صورة ويقوم النظام باستخراج الكتابة الموجودة ضمن الصورة أو ضمن الملفات التي لها صيغة من نمط Pdf مما يسهل على المستخدم عملية إعادة الكتابة.



رسم توضيحي 1- حساب وزن العقدة e

## الفصل الأول: الدراسة النظرية

### Page rank

هي خوارزمية تستخدم لحساب الأوزان لصفحات الويب، حيث تمثل جميع صفحات الويب ضمن بيان موجه، فتمثل كل صفحة ويب بعقدة ضمن هذا البيان، وإذا كانت صفحة الويب A تحتوي على رابط لصفحة الويب B، فيمكن تمثيلها كوصلة موجهة من A إلى B. [1]

$$S(V_i) = (1 - d) + d * \sum_{j \in In(v_i)} \frac{1}{|Out(V_j)|} S(V_j) \quad (1)$$

$S(V_i)$ : وزن أو أهمية الصفحة  $i$ .

$d$ : معدل التخامد (يكتب في بعض الأحيان  $\alpha$ ) وقيمته بين ال 0 وال 1 ويمثل احتمال الانتقال من عقدة إلى عقدة عشوائية.

$$\begin{aligned}
 In(v_e) &= \{a, b\}, j \in \{a, b\} \\
 \sum_{j \in \{a, b\}} \frac{1}{|Out(V_j)|} S(V_j) &= \frac{1}{|Out(V_a)|} S(V_a) + \frac{1}{|Out(V_b)|} S(V_b) \\
 &= \frac{1}{|\{e\}|} S(V_a) + \frac{1}{|\{e, f\}|} S(V_b) \\
 &= S(V_a) + \frac{1}{2} S(V_b)
 \end{aligned}$$

$In(V_i)$ : الوصلات الداخلة إلى  $i$ .

$Out(v_j)$ : الوصلات الخارجة من  $j$ .

فمثلا لحساب قيمة أهمية العقدة  $e$  من الرسم التوضيحي (1) نقوم بتبسيط المعادلة كما يلي:

$$S(V_i) = (1 - d) + d * \left( S(V_a) + \frac{1}{2} S(V_b) \right) \quad (2)$$

ثم نقوم بتنظيم جدول يعبر عن البيان والوصلات الداخلة والخارجة حيث نعوض 1 عندما يكون هناك وصلة بين العقدة في السطر  $i$  والعقدة في العمود  $j$ ، لينتج لدينا الجدول بالشكل التالي:

	a	b	e	f
a	0	0	0	0
b	0	0	0	0
e	1	1	0	0
f	0	1	0	0

ثم في مكان الوصلة بين  $i$  و  $j$  نعوض قيمة  $\frac{1}{|Out(V_i)|}$

	a	b	e	f
a	0	0	0	0
b	0	0	0	0
e	1	0.5	0	0
f	0	0.5	0	0

ويمكن تمثيل الجدول السابق بمصفوفة كما يلي:

weights of inbound nodes of e

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1.5 \\ 0.5 \end{bmatrix} \text{ weight of e}$$

inbound links of e

Why	not	tell	someone	?
adverb	adverb	verb	noun	punctuation mark, sentence closer

رسم توضيحي 2- مثال عن ال POS tagging

## Part of speech tags

هي عملية تعتمد على معالجة اللغات الطبيعية، حيث تقوم بتحديد نوع الكلمة ضمن الجملة (اسم - فعل - حرف ...الخ) وذلك اعتماداً على تعريف الكلمة وسياقها.

يحتوي POS على عدّة مستويات فالمستوى الأبسط يكتفي بتحديد نوع الكلمة (اسم - فعل - حرف)، أما المستويات الأكثر تعقيداً تعمل على تحديد نوع الاسم (مثنى - جمع - مفرد)، وكذلك نوع الفعل (ماض - مضارع - أمر)، كما أنه يوجد مستويات أخرى تستطيع إعطاء معلومات أكثر عن الكلمة.

يبين الرسم التوضيحي (2) أن لكل كلمة المصطلح المعجمي الخاص بها، ولكن لتبسيط عملية الكتابة وبدلاً من كتابة كامل المصطلح المعجمي لكل كلمة تم إنشاء رموز مختصرة لكل مصطلح معجمي "tags". [2]

Lexical Term	Tag	Example
Noun	NN	Paris, France, Someone, Kurtis
Verb	VB	work, train, learn, run, skip
Determiner	DT	the, a
...	...	

Why not tell someone ?  
WRB RB VB NN .

جدول 1: يوضح أنواع مختلفة من المصطلحات المعجمية مع رموزها المختصرة وبعض الأمثلة لكل منها.

## التعلم الآلي Machine learning

نشأت فكرة التعلم الآلي Machine Learning من أهمية الاستفادة من الكمية الكبيرة من البيانات في تصميم نماذج قادرة على اتخاذ القرار بدون تحديد مسبق للعلاقة التي تربط الدخل بالخرج حيث أن هذه العلاقة ليست دائماً معلومة لدينا وليس من السهل اكتشافها بالنسبة للإنسان.

ومن هذا المنطلق جاءت أهمية التعلم الآلي حيث يتم التعلم من البيانات المتوفرة لدينا لاكتشاف أفضل علاقة تربط بين الدخل والخرج، بحيث تمتلك القدرة على التعميم وتتمكن من توقع نتائج بيانات جديدة.

يعتمد التعلم التلقائي على:

### 1- البيانات Data:

يمكن أن تكون مصنفة labelled أو غير مصنفة unlabeled.

### 2- النموذج/ الخوارزمية Model:

هي العلاقة الرياضية التي ستستخدم لربط الدخل بالخرج.

### 3- تابع الخطأ loss function:

يعطي مقدار الخطأ الناتج عن توقع نموذج ما، تابع لمجموعة محددة من المتغيرات.

### 4- خوارزمية التحسين Optimization Algorithm:

وهي عملية يتم تنفيذها دورياً لإيجاد أفضل قيم لمتغيرات النموذج (الأوزان).

يقسم التعلم التلقائي حسب نوع البيانات إلى:

1. التعلم بالإشراف Supervised learning: حيث تكون البيانات مصنفة Labelled Data.
2. التعلم بدون إشراف Unsupervised learning: حيث تكون البيانات مصنفة Unlabelled Data.

مراحل التعلم الآلي:

### 1- تجميع البيانات ومعالجتها Data collection and pre-processing:

حيث يتم جمع البيانات من المصادر المتنوعة ومعالجتها كإزالة الضجيج، ثم يتم تقسيم البيانات إلى أقسام للتدريب والاختبار.

### 2- استخراج السمات Features engineering and extraction:

يتم تمثيل السمات المعبرة عن البيانات بشكل رياضي واختيار الأفضل منها.

### 3- اختيار النموذج model Selection:

يتم اختيار النموذج حسب نوع البيانات ودرجة تعقيدها كما وتلعب متطلبات المسألة دوراً في عملية الاختبار.

### 4- التدريب Train:

يتم استخدام قسم التدريب من البيانات وذلك لمحاولة إيجاد قيم الأوزان التي تحقق أقل قيمة ممكنة لتابع الخطأ.

#### 5- الاختبار والتقييم Test & Evaluate:

يتم تجريب النموذج الذي تمّ إيجاده مسبقاً على قسم الاختبار وهي مجموعة من بيانات جديدة لا يتم تعديل الأوزان بناءً عليها وذلك لقياس الدقة وتقييم النموذج.

## الشبكات العصبونية Neural Networks

هي سلسلة من الخوارزميات التي تسعى إلى التعرف على العلاقات الأساسية في مجموعة من البيانات من خلال عملية تسمح بتكيف الشبكات العصبونية مع المدخلات المتغيرة؛ لذا فإن الشبكة تحقق أفضل نتيجة ممكنة دون الحاجة إلى إعادة تصميم معايير الإخراج، ولذلك تعتبر الشبكات من أهم النماذج المستعملة لحل مسائل متنوعة وذلك بسبب قابليتها لحل طيف واسع من المسائل ذات درجات تعقيد مختلفة.

تعتمد الشبكة العصبونية الاصطناعية على مجموعة من الوحدات أو العقد المتصلة تسمى الخلايا العصبونية الاصطناعية، والتي تحاكي الطريقة التي تعمل بها الخلايا الدماغية في معالجة البيانات وتكوين المعارف.

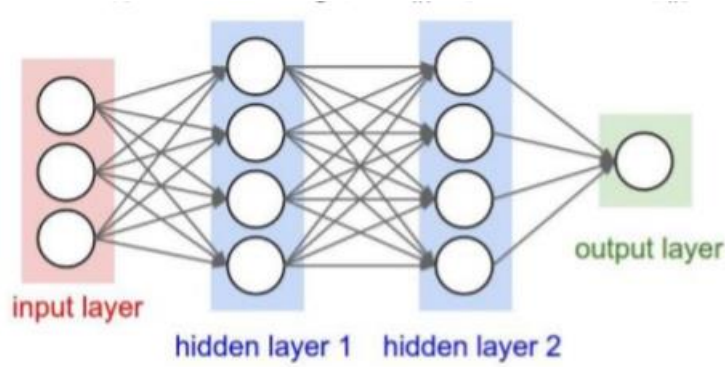
تتلقى الخلايا العصبونية إشارة وتقوم بمعالجتها ثم تنتقل الإشارة من خلية عصبونية إلى خلية عصبونية أخرى عن طريق وجود اتصال بين الخليتين يشبه هذا الاتصال نقاط الاشتباك العصبي في الدماغ البيولوجي.

تكون الإشارة عند الاتصال هي رقم حقيقي، ويتم حساب ناتج كل خلية عصبية بواسطة بعض الوظائف غير الخطية لمجموع مدخلاتها. تسمى الوصلات بالحواف. عادةً ما يكون للخلايا العصبونية والحواف ثقل يتكيف مع تقدم التعلم.

يزيد الوزن أو ينقص من قوة الإشارة عند الاتصال. قد يكون للخلايا العصبونية عتبة بحيث يتم إرسال الإشارة فقط إذا تجاوزت الإشارة المجمعة تلك العتبة. عادة، يتم تجميع الخلايا العصبونية في طبقات. قد تؤدي الطبقات المختلفة إلى تحولات مختلفة على مدخلاتها. تنتقل الإشارات من الطبقة الأولى (طبقة الإدخال) إلى الطبقة الأخيرة (طبقة الإخراج).

تتألف الشبكة العصبونية البسيطة من عدة طبقات تحوي كل طبقة مجموعة من خلايا (عصبونات)، حيث يتصل كل عصبون بجميع العصبونات في الطبقة التالية Fully connected ويكون خرج كل طبقة دخلاً للطبقة التي تليها.





رسم توضيحي 3- مخطط لشبكة عصبونية بسيطة

من الرسم التوضيحي (3) نلاحظ أنّ مكونات الشبكة العصبونية البسيطة ما يلي:

طبقة الدخل Input Layer: يكون دخل الشبكة مجموعة من السمات المعبرة عن عينة ما.

طبقة مخفية Hidden Layer: تحوي مجموعة من الخلايا neurons.

طبقة خرج Output Layer: تعطي توقع النموذج.

ترتبط كل طبقة مع الطبقة التي تليها بشعاع أوزان  $W_i$ .

خرج الطبقة هو نتيجة تطبيق تابع التفعيل على الجداء السلمي لشعاع الأوزن بشعاع الدخل لها.

$$f(x) = G(W_2 \cdot (S(W_1 \cdot x))) \quad (3)$$

$W_i$ : شعاع الأوزان بين الطبقة  $i - 1$  و الطبقة  $i$ .

$x$ : شعاع الدخل.

$G, S$ : توابع تفعيل activation function حيث يتم تطبيق تابع تفعيل على خرج كل طبقة لتحديد مجال الخرج.

تتعلم الشبكات العصبونية (أو يتم تدريبها) من خلال معالجة الأمثلة، كل منها يحتوي على "مدخلات" و "نتيجة" معروفة، مما يشكل ارتباطات مرجحة بين الاثنتين، والتي يتم تخزينها داخل بنية بيانات الشبكة نفسها. عادةً ما يتم إجراء تدريب الشبكة العصبونية من مثال معين عن طريق تحديد الفرق بين الخرج المعالج للشبكة (غالبًا ما يكون تنبؤًا) والخرج المستهدف. هذا هو الخطأ. تقوم الشبكة بعد ذلك بضبط ارتباطاتها الموزونة وفقًا لقاعدة التعلم وباستخدام قيمة الخطأ هذه. ستؤدي التعديلات المتتالية إلى إنتاج الشبكة العصبونية لمخرجات تشبه بشكل متزايد المخرجات المستهدفة.

يتم ذلك باستخدام خوارزمية gradient descent مع خوارزمية back propagation لإيجاد القيمة الصغرى لتابع الخطأ من خلال دراسة تأثير تغير الأوزان على تابع الخطأ وذلك باشتقاق تابع الخطأ بالنسبة لكل وزن الذي ينتج عنه شعاع الزيادة الأعظمية وتعديل الأوزان بالاتجاه المعاكس له.

$$\Delta w_{ij} = -\alpha \frac{\partial E}{\partial w_{ij}} \quad (4)$$

عادة ما يتم استخدام الشبكات العصبونية العميقة Deep Neural Networks وهي شبكة عصبونية تحوي على أكثر من طبقة مخفية layer hidden حيث كلما زاد عمق الشبكة استطاعت التعامل مع مسائل أكثر تعقيداً. [1]

## الشبكات العصبونية التلافيفية CNN

شبكة الـ CNN مشابهة جداً للشبكة العصبونية العادية تدرّب عادةً باستخدام supervised learning، تم اختراع هذه الشبكة لتسهيل التدريب وتقليل عدد الـ parameters مع نفس عدد Hidden units.

تختلف الشبكات التلافيفية عن الشبكات العصبونية أنه في الشبكات العصبونية العادية يتصل كل عصبون في كل طبقة مع جميع العصبونات في الطبقة التالية (الرسم التوضيحي (3)) مما يزيد عدد الأوزان بشكل كبير خاصة عندما نتعامل مع صور ذات حجم كبير، أما الشبكات العصبونية التلافيفية يتصل العصبون في طبقة ما فقط مع جزء من العصبونات في الطبقة التالية له مما يقلل عدد المتغيرات بشكل كبير ويوفر الكثير من العمليات الحسابية.

لذلك تعتبر الشبكات العصبونية التلافيفية واحدة من أهم النماذج المستعملة في عملية معالجة الصور وذلك لأنه إذا كان لدي أوزان كثيرة سأحتاج لحساب العديد من المتحولات مما يزيد احتمالية حدوث مشكلة overfitting وبما أن الشبكة التلافيفية تخفض عدد الوصلات فهي تقوم بالتخلص من هذه المشكلة.

تتألف الشبكة العصبونية من مجموعة من الطبقات التي تعمل على استخلاص السمات الأكثر تعبيراً عن الصورة لينتج عنها تمثيل شعاعي عن كل صورة وبهذا توفر الكثير من العمل المضني في استخراج السمات ومعالجة الصور قبل البدء بعملية التدريب وهذه الطبقات هي:

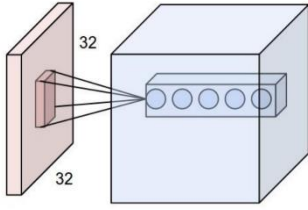
### **-1 Input layer:**

يكون عدد عصبونات في هذه الطبقة مساوياً لأبعاد الصورة مضروباً ب 3 في حال كانت الصورة ملونة،

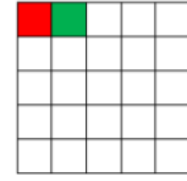
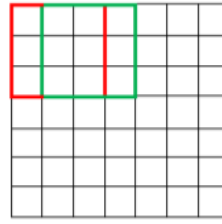
فلكل لون من ألوان البيكسل الواحد يكون له عصبون.

### **-2 Convolution layer:**

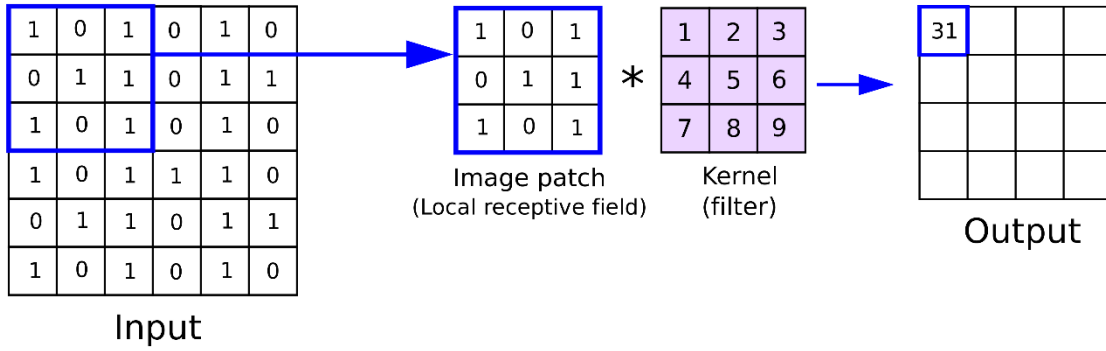
هذه الطبقة مهمتها استخراج السمات الخاصة بالصورة نفسها كالحواف، الزوايا، الخطوط، ويتم ذلك عن طريق تطبيق عدد من المرشحات على كل جزء من الصورة.



رسم توضيحي 4- ارتباط الصورة مع العصبونات في  
طبقة ال Convolution



رسم توضيحي 5- تطبيق مرشح على جزئين من  
الصورة stride = 1



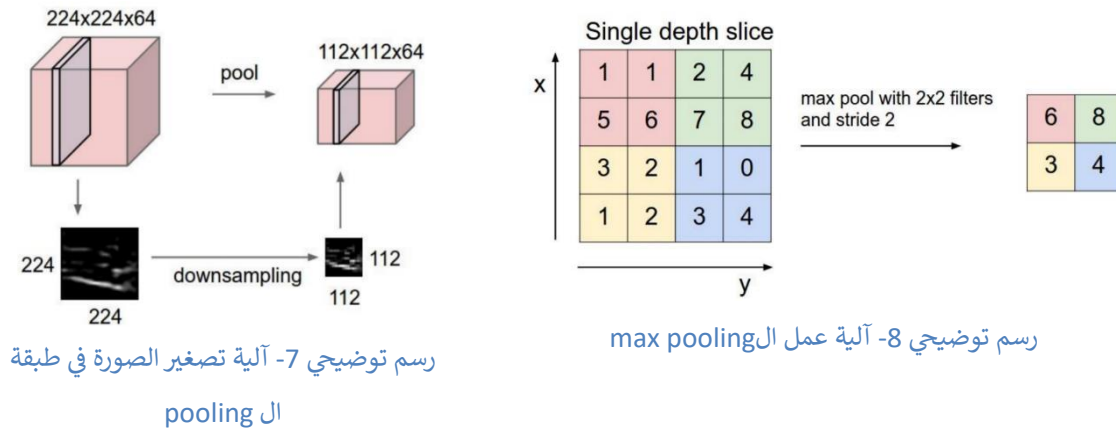
رسم توضيحي 6- تطبيق مرشح على جزء من الصورة

يتم تطبيق المرشح على كل جزء من أجزاء الصورة وذلك بتحريكه على الصورة بخطوة تدعى Stride يمكن التحكم بها حسب الحاجة وذلك كما هو مبين في الرسم التوضيحي (4).

ويكون كل عنصر من الصورة الجديدة هو مجموع جداء كل عنصر من المرشح مع البيكسل المقابل له من الصورة الأصلية وذلك كما هو مبين في الرسم التوضيحي (6).

ويمكن أن نضيف على حواف الصورة Padding للمحافظة على المعلومات على الحواف وعلى حجم الصورة.

كل عصبون من هذه الطبقة يوصل مع جزء من الصورة وبالتالي كل عصبون منها أصبح مسؤول عن الجزء الذي يأخذه حصراً وذلك كما هو مبين في الرسم التوضيحي (5).



### -3 Pooling layer:

يتم في هذه الطبقة القيام بتصغير لأبعاد الصورة مع المحافظة على المعلومات المهمة الموجودة في الصورة وهذا ما يوضحه الرسم التوضيحي (7)، وتمنع حدوث ال overfitting، ولهذه الطبقة عدّة أنواع:

- I. Average Pooling: حيث نأخذ متوسط القيم من كل جزء من الصورة.
- II. Sum Pooling: نأخذ مجموع قيم الجزء.
- III. Max pooling: تأخذ أعلى قيمة من قيم الأجزاء وهذا ما يوضحه الرسم التوضيحي (8).

### -4 Rectified Linear Units Layer:

كل عصبون له دخل وحيد وتعمل هذه الطبقة للقيام ب threshold، إضاءة...، ولها عدّة أنواع:

- $Max(0, x)$  for thresholding at zero
- Sigmoid:  $(1 + e^{-1})^{-1}$
- Tanh(x)

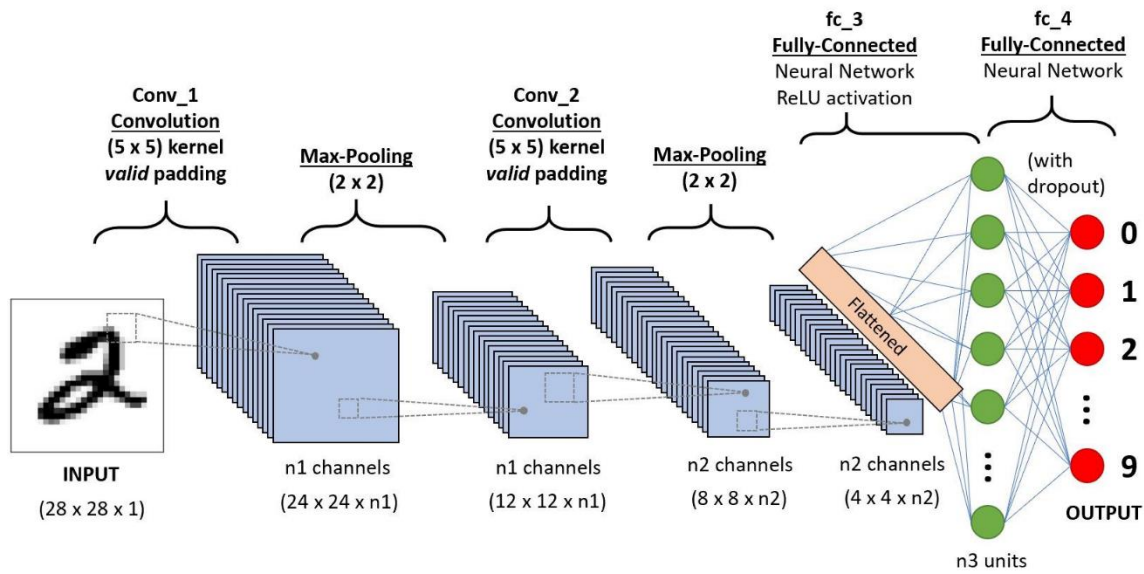
### -5 Fully Connected Layer:

شبكة عصبونية عادية، كل عصبون فيها يتصل مع كافة عصبونات الطبقة التي تسبقها.

### -6 Output Layer:

وهنا يمكن تطبيق عدّة توابع منها:

- MAX: يستخدم للتنبؤ بصف واحد من k صف.
- Sigmoid: يستخدم للتنبؤ بالاحتمالات المستقلة وقيمة تكون ضمن المجال [0,1].
- Euclidean: يستخدم للتراجع إلى labels ذات قيم  $[-\infty, +\infty]$ .



رسم توضيحي 9- شبكة عصبونية تلافيفية

بالنظر إلى الرسم التوضيحي (9) تظهر شبكة عصبونية تلافيفية مكونة من أنواع مختلفة من الطبقات.

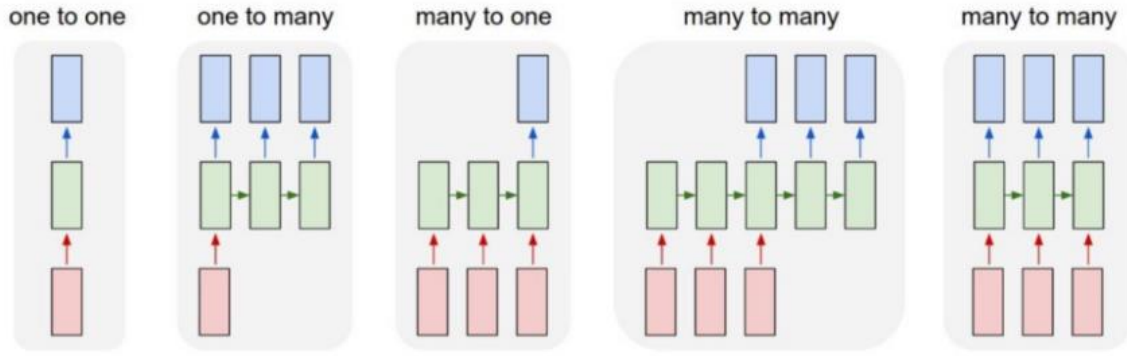
عادةً في مصنفات الصور تكون طبقات Convolution و Pooling لاستخراج السمات من الصور ويمكن تعلم الأوزان المستخدمة فيها، أما طبقات fully connected فتستعمل لتصنيف الصورة حسب السمات الناتجة ويتم تركيب هذه الطبقات حسب الحاجة.

## الشبكات العصبونية التكرارية Recurrent Neural Networks

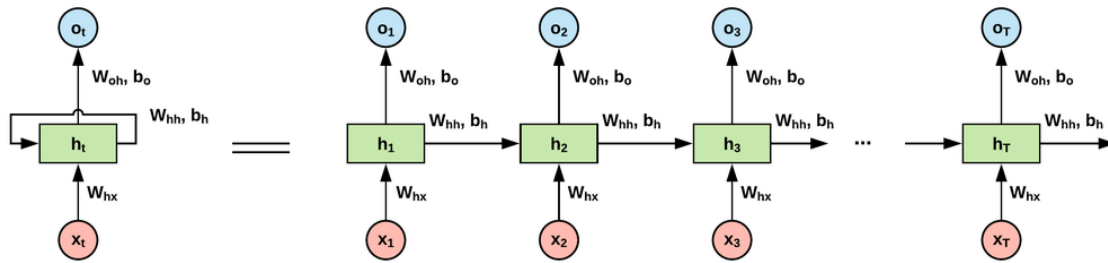
هي أحد أنواع الشبكات العصبونية التي تمتلك ذاكرة حيث أن الخرج لا يعتمد على الدخل فقط إنما أيضاً على الحالة السابقة. [3]

تكمّن أهمية هذا النوع من الشبكات العصبونية بأنه يستفيد من المعلومات المتسلسلة حيث أنّ الشبكات العصبونية التقليدية تتعامل مع كل دخل وخرج على أنّه حالة مستقلة كما أنّ الشبكات العصبونية التكرارية تستطيع التعامل مع أطوال مختلفة لأشعة الدخل والخرج.

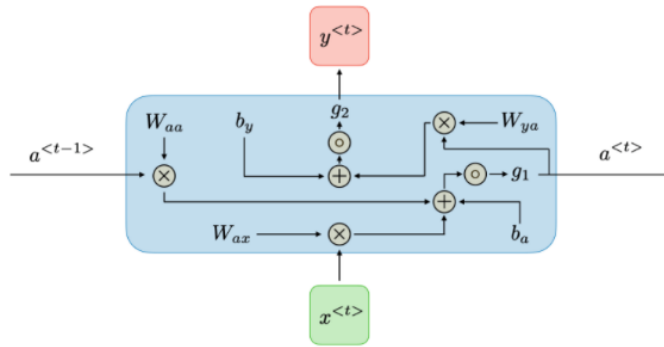
غالباً ما يستعمل هذا النوع من النماذج في مجالات معالجة اللغات الطبيعية والتعرف على الكلام وتوليد النصوص والترجمة.



رسم توضيحي 10- أنواع الشبكات العصبونية التكرارية



رسم توضيحي 11- مخطط يوضح بنية الشبكة العصبونية التكرارية



رسم توضيحي 12- مخطط يوضح الحالة الخفية في الشبكة التكرارية

لها عدّة أنواع حيث لكل نوع استعمالات مختلفة حسب المسألة وهي موضحة في الرسم التوضيحي (10)

- One to one: لها دخل واحد وخرج واحد وتعمل كشبكة عصبونية بسيطة.
- One to many: لها دخل واحد (في المرحلة الأولى) وأكثر من خرج.
- Many to one: لها أكثر من دخل وخرج وحيد.
- Many to many: لها أكثر من دخل وأكثر من خرج.

بالنظر إلى الرسم التوضيحي (11) يمكن فهم البنية العامة للشبكات التكرارية.

يمكن الاستفادة في فهم المعادلات التالية وذلك بالنظر إلى الرسم التوضيحي (12):

$$h_t = g_1 (W_{hh} h_{t-1} + W_{xh} x_t + b_a)$$

$$y_t = g_2(W_{hy} h_t + b_y)$$

$h_t$ : هي الحالة الخفية hidden state في الخطوة الزمنية  $t$  حيث تتعلق بدخل الشبكة في الخطوة الحالية  $t$  والحالة الخفية السابقة  $t-1$ .

$y_t$  هو خرج الشبكة في الخطوة الزمنية  $t$  وهو مرتبط بالحالة الخفية في اللحظة  $t$ .

$W_{hh}, W_{xh}, W_{hy}$ : مصفوفات الأوزان.

نلاحظ أن جميع الطبقات تتشارك مصفوفات الأوزان مما يقلل عدد المتغيرات.

### تدريب الشبكات العصبونية التكرارية

يتم تدريب الشبكات العصبونية التكرارية بخوارزمية مشابهة لخوارزمية backpropagation تدعى backpropagation through time.

كما في خوارزمية backpropagation التقليدية علينا حساب مشتق تابع الخطأ بالنسبة لجميع الأوزان:

حيث أن مشتق تابع الخطأ بالنسبة لـ  $W_{ya}$  يعتمد فقط على الحالة الحالية  $\langle y_3, a^{<t>}, \hat{y}_3 \rangle$ .

$$\frac{\partial E_3}{\partial W_{ya}} = \frac{\partial E_3}{\partial \hat{y}_3} \frac{\partial \hat{y}_3}{\partial W_{ya}} = \frac{\partial E_3}{\partial \hat{y}_3} \frac{\partial \hat{y}_3}{\partial z_3} \frac{\partial z_3}{\partial W_{ya}} = (\hat{y}_3 - y_3) \otimes s_3$$

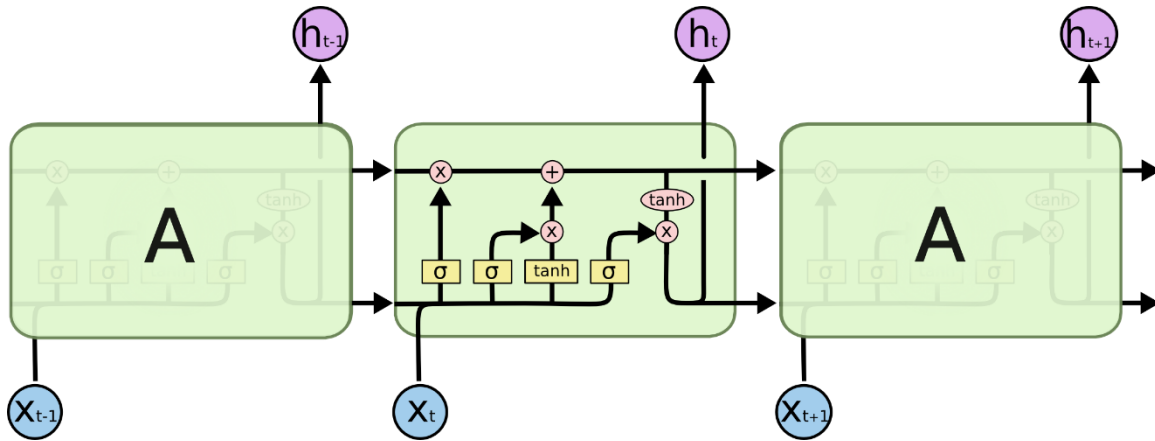
أما المشتق بالنسبة لبقية الأوزان  $W_{aa}, W_{ax}$  فهو مشتق تابع الخطأ بالنسبة لـ  $a_t$  وهو تابع للأوزان عبر المراحل السابقة ولذلك علينا أخذ المشتقات عبر الزمن بعين الاعتبار (مشتق تابع الخطأ بالنسبة للحالات الخفية عبر الزمن  $a_1, a_2, a_3$ ) فيكون تابع مشتق تابع الخطأ بالنسبة للأوزان هو مجموع مشتقات تابع الخطأ بالنسبة لهذه الأوزان عبر المراحل الزمنية.

$$\frac{\partial E_t}{\partial W} = \sum_{k=0}^t \frac{\partial E_t}{\partial \hat{y}_t} \frac{\partial \hat{y}_t}{\partial a_t} \frac{\partial a_t}{\partial a_k} \frac{\partial a_k}{\partial W}$$

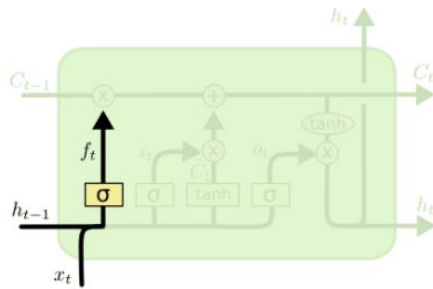
تكمّن مشكلة هذه الشبكات في أنها لا تستطيع التعامل مع الترابطات طويلة المدى Long Term Dependencies وذلك لأنه نتيجة ضرب المشتقات ببعضها يمكن للمشتق أن يصغر كثيراً حتى يختفي Vanishing Gradient أو أن يكبر كثيراً Exploding Gradient عبر الطبقات المتعددة حيث أن عدد الطبقات سيزداد بازدياد مدى الترابطات.

عادة يتم التعامل مع تضخم المشتق من خلال تحديد الحد الأعلى التي يمكن أن يأخذه Clipping Gradient.

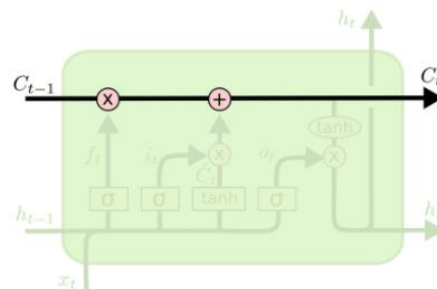
أما بالنسبة لاختفاء المشتق فهنا جاءت بنى جديدة تتعامل مع هذه المشكلة مثل LSTM.



رسم توضيحي 13- البنية العامة لشبكة LSTM



رسم توضيحي 15- بوابة النسيان في شبكة LSTM



رسم توضيحي 14- موقع ذاكرة الشبكة State Cell

## Long-Short Term Memory (LSTM)

تعد شبكة ال LSTM نوعاً خاصاً من الشبكات العصبونية التكرارية، وقد صممت هذه الشبكات للتعامل مع الترابطات طويلة المدى Long Term Dependencies والتي كانت تشكل مشكلة في الشبكات التكرارية RNN.

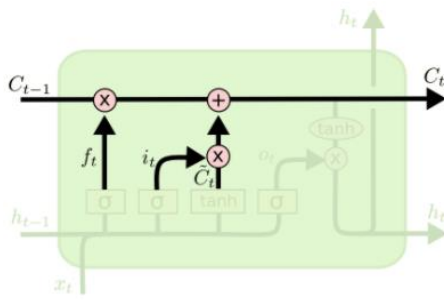
إن الشبكات العصبونية التكرارية تحوي على طبقة واحدة مكررة عدّة مرات أما شبكات ال LSTM فهي تحوي على 4 أنواع الطبقات تتفاعل فيما بينها. [4]

يمثل الخط العلوي في الرسم التوضيحي (14) ذاكرة الشبكة ويدعى Cell State حيث يتم تطبيق بعض العمليات عليه لاختيار المعلومات التي ستبقى والتي ستحذف وذلك من خلال عمل البوابات الأربعة السابق ذكرها بالطريقة الآتية:

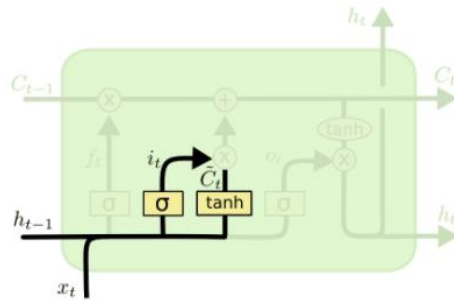
### 1- المرحلة الأولى:

يتم في هذه المرحلة تحديد المعلومات التي سيتم السماح لها بالمرور عبر ذاكرة الشبكة من خلال استعمال تابع التنشيط sigmoid، وتسمى هذه البوابة ببوابة النسيان، وهي موضحة في الرسم التوضيحي (15).





رسم توضيحي 17- المرحلة الثالثة من مراحل LSTM



رسم توضيحي 16- المرحلة الثانية من مراحل LSTM

دخل هذه البوابة هو الحالة الخفية السابقة  $h_{t-1}$  الدخل الحالي  $x_t$  يطبق عليهما تابع تفعيل من نوع sigmoid فيكون الخرج قيمة بين الصفر والواحد لكل قيمة مقابلة من الخلية الذاكرة  $C$  حيث 0 يعني نسيان القيمة المقابلة من الذاكرة تماماً و1 يعني إبقائها تماماً وأي قيمة بينهما هي توزيع احتمالي لإمكانية إبقاء القيمة المقابلة لها من الذاكرة .

وفي نهاية هذه المرحلة نكون قد حددنا المعلومات التي نريد نسيانها ومعلومات التي نريد الاحتفاظ بها.

فيكون شكل المعادلة كما يلي:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

## 2- المرحلة الثانية:

يتم في هذه المرحلة تحديد المعلومات التي سيتم الاحتفاظ بها من المعلومات الجديدة التي تم الحصول عليها وتظهر هذه المرحلة في الرسم التوضيحي (16).

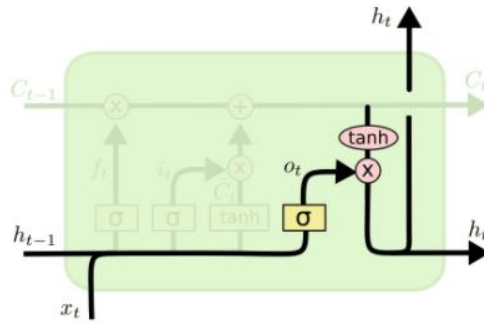
تتألف هذه المرحلة من بوابتين الأولى تدعى بوابة الدخل input gate يتم من خلالها تحديد أي القيم في الذاكرة سيتم تغييرها:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

والثانية بوابة الترشيح Candidate تعطينا شعاعاً بالقيم الجديدة المرشحة لتضاف إلى الذاكرة

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

## 3- المرحلة الثالثة:



رسم توضيحي 18- المرحلة الرابعة من مراحل LSTM

المراحل السابقة حددت لنا التغيرات التي ستطرأ على الذاكرة وكل ما تبقى هنا هو تطبيقها، حيث يتم في هذه المرحلة دمج خرج البوابات السابقة  $f_t, i_t, \tilde{C}_t$  فنحدد قيم الذاكرة التي ستبديل والقيم الجديدة التي ستحل بدلاً عنها وهذا ما هو موضح في الرسم التوضيحي (17).

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$C_t$ : هي الحالة الجديدة للذاكرة.

$C_{t-1}$ : هي الحالة القديمة للذاكرة يتم ضربها بخرج بوابة النسيان لننسى أو نحافظ على كل قيمة منها.

$t_i * \tilde{C}_t$ : هذا الحد لإضافة القيم الجديدة.

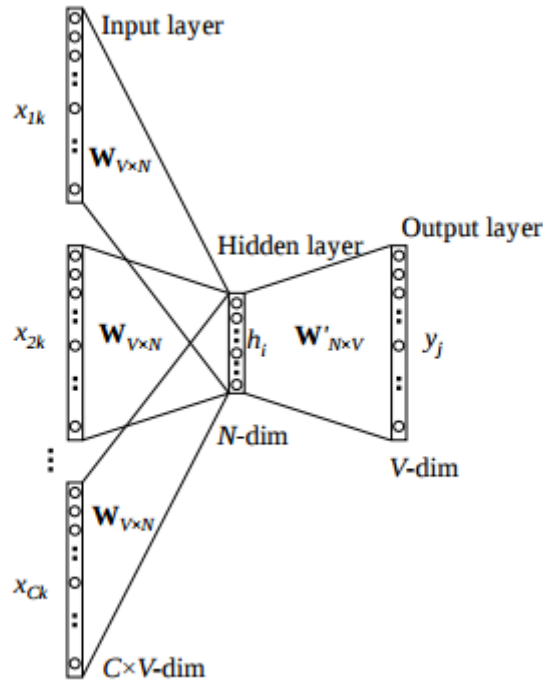
#### 4- المرحلة الرابعة:

في هذه المرحلة تدعى البوابة output gate وهي موضحة في الرسم التوضيحي (18) وتعتمد على الحالة الخفية السابقة والدخل وتعطي مجموعة القيم التي سنختار الخرج النهائي منها:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

أخيراً يتم تحديد الأجزاء التي نريدها من نتيجة بوابة الخرج باستخدام حالة الذاكرة بعد تطبيق تابع  $\tanh$  عليها وبذلك نخرج فقط الأجزاء التي تهتمنا حسب المعادلة:

$$h_t = o_t * \tanh(C_t)$$



رسم توضيحي 19- بنية نموذج CBOW

## Word2Vec

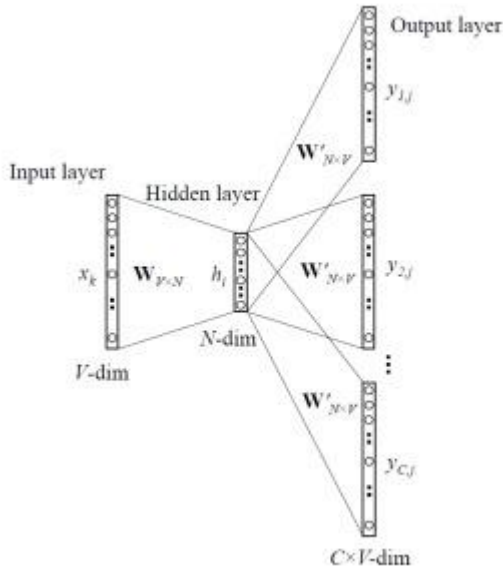
هي آلية من أجل تحويل أي كلمة إلى شعاع. وهي ليست خوارزمية واحدة بل عبارة عن دمج تقنيتين ( CBOW, skip-gram model) وكل من هاتين التقنيتين عبارة عن شبكة عصبونية ضحلة تقوم بتعلم أوزان الطبقة المخفية التي تمثل أشعة الكلمات. [5]

### CBOW

طريقة عمل CBOW تعتمد على التنبؤ بالكلمة التالية في الجملة بالاعتماد على السياق الذي يمكن أن يكون كلمة واحدة أو مجموعة كلمات.

خطوات عمل CBOW:

1. تحويل الكلمات إلى شكل one-hot-encoding وبحيث تكون أبعاد كل شعاع  $(V \times 1)$  حيث  $V$  هي عدد كلمات اللغة.
2. ضرب مصفوفة الدخل بمصفوفة الأوزان  $W_{v,n}$  في الطبقة المخفية والتي تكون أبعادها  $(V \times N)$  حيث  $N$  هو عدد العصبونات في هذه الطبقة ويعتبر الشعاع الناتج  $h$  هو التمثيل الشعاعي لكلمة الدخل.
3. ضرب الشعاع  $h$  بمصفوفة الأوزان  $W_{n,v}$  في الطبقة المخفية والتي تكون أبعادها  $(N \times V)$  ويعتبر الشعاع الناتج  $y$  هو احتمال من أجل كل كلمة في اللغة أن تكون هي الكلمة التالية ضمن سياق الدخل.



Input	Output(Context1)	Output(Context2)
Hey	this	<padding>
this	Hey	is
is	this	sample
sample	is	corpus
corpus	sample	corpus
using	corpus	only
only	using	one
one	only	context
context	one	word
word	context	<padding>

جدول 2: خرج عملية تدريب Skip-gram

رسم توضيحي 20- بنية نموذج Skip-gram

4. حساب الخطأ وتعديل الأوزان.
  5. يتم تكرار الخطوات السابقة حتى يتم التقارب أو الوصول إلى عدد تكرارات محدد.
- طريقة عمل CBOW تعتمد على التنبؤ بالكلمة التالية في الجملة بالاعتماد على السياق الذي يمكن أن يكون كلمة واحدة أو مجموعة كلمات، وتظهر بنية CBOW في الرسم التوضيحي (19).

## Skip-gram model

تتبع خوارزمية Skip-gram نفس خطوات CBOW لكنها تقوم بعكس بنية الشبكة. ففي هذه الخوارزمية نحاول التنبؤ بالسياق بالاعتماد على الكلمة الحالية ويمكن ان يكون السياق كلمة أو مجموعة كلمات فمثلا إذا كان لدينا الجملة التالية

" Hey, this is sample corpus using only one context word."

تكون بيانات التدريب كما في الجدول (2).

وبما أنه يوجد أكثر من خرج يتم تدريب الدخل مع كل خرج على حدا بإضافة الخطأ إلى الخطأ النهائي ومن ثم تعديل الأوزان.

تمثل الأوزان بين طبقة الدخل والطبقة المخفية أشعة الكلمات.

وتظهر بنية Skip-gram model في الرسم التوضيحي (20).

## الفصل الثاني: الدراسة المرجعية

### 1- مشاريع مشابهة

يوجد العديد من الأنظمة المخصصة لعملية إدارة المحتوى ومن أشهرها: WordPress , Drupal, Joomla , Contentful , Strapi, Wagtail, Apostrophe .

حيث تنقسم الأنظمة إلى نوعين:

1. Traditional CMS

2. Headless CMS

بعض هذه الأنظمة يعتمد على الويب بشكل مخصص في إدارة المحتوى (Traditional CMS) مثل WordPress و Joomla و Drupal وبعضها يقوم بفصل عملية إدارة المحتوى عن عملية العرض، حيث يدير المحتوى بشكل منفصل باستخدام API (Headless CMS) ومن ثم يقوم المستخدم بعرض المحتوى بالشكل الذي يريده (صفحة ويب، تطبيق للهاتف المحمول، تطبيق للحاسب.... الخ) مثل Strapi و Wagtail و Apostrophe و Contentful.

### النظام WordPress

هو نظام إدارة محتوى إلكتروني مفتوح المصدر، مبني بلغة php وقواعد بيانات MySQL، تتيح الوظائف التي يتوفر عليها WordPress إدارة أي موقع ويب خصوصاً المدونات. حيث تشير الإحصاءات إلى أن WordPress يُشغل حوالي 33% من مواقع الويب.

يعتمد هذا النظام على معمارية الإضافات Plugins، وهي عبارة عن برامج مساعدة تُقدم العديد من الوظائف المخصصة لتمكين احتياجات المستخدمين الخاصة وزيادتها، الاحتياجات الخاصة مثل: تقوية تواجد موقع المستخدم على محركات البحث لتحسينها (SEO) ومثل عناصر معدة مسبقاً لإظهار المعلومات الخاصة للعميل للمستخدمين المُسجلين في النظام Registered users لعرض المزيد من المحتويات مثل Widgets وشريط التنقل navigation bar. أيضاً هناك الإضافات التي تقدم خدمات وظيفية للموقع مثل خدمات الترجمة وشرح الصور وعنونة المقالات.

بالمقارنة مع الإضافات التي تقدم الخدمات نفسها التي يقدمها نظامنا المقترح:

### الإضافات المتعلقة بتوليد كلمات المفتاحية:

#### إضافة العنوان التلقائي للنصوص QwebmasterAutoTagger:

تقوم الإضافة بتوليد الكلمات المفتاحية بشكل تلقائي للمنشورات الخاصة بالمدونة Posts. تعمل هذه الإضافة عن طريق إرسال محتوى المنشور إلى مخدم ليتم معالجة النص باستخدام نماذج التعلم التلقائي لاستخراج

سلسلة من الكلمات المفتاحية المناسبة لهذا النص. حيث لا يتم تخزين أي معلومة متعلقة بالمدونة أو المنشور لضمان حماية سياسة الخصوصية.

ومن محاسن هذه الإضافة أنها توفر ميزة توليد الكلمات المفتاحية المناسبة لكل المنشورات الموجودة ضمن الموقع بشكل تلقائي دون تدخل المستخدم.

أما مساوئ هذه الإضافة أنها:

- 1- لا تدعم اللغة العربية أي لا يمكنها استخراج الكلمات المفتاحية من النصوص العربية.
- 2- تحتاج لإرسال المحتوى إلى مخدم خارجي من أجل معالجته.

#### إضافة الذكاء الاصطناعي لمنشئي المحتوى JANEY AI:

وهي إضافة تحتوي على عدة أدوات تعتمد على الذكاء الاصطناعي. فهي تساعد المدونين على تحديد الكلمات المفتاحية وجلب صور المرتبطة بمضمون المنشور والمراجع المناسبة للمحتوى. تعمل الإضافة من خلال تحليل محتوى المنشور واقتراح الكلمات المفتاحية المناسبة والصور المرتبطة بذاك المضمون.

ومن محاسن هذه الإضافة أنها:

توفر خدمات متعددة متعلقة بالذكاء الاصطناعي (توليد كلمات مفتاحية بشكل تلقائي مع عرض نسبة الارتباط بينها وبين المحتوى، اقتراح صور مناسبة لمضمون المحتوى مأخوذة من موقع Unsplash)

كما لها عدة مساوئ:

- لا تدعم اللغة العربية.
- يتم إرسال المحتوى إلى مخدم خارجي من أجل معالجته.
- الإضافة مدفوعة وغير مجانية.

#### الإضافات المتعلقة بترجمة للمحتوى:

##### إضافة دعم تعدد اللغات PolyLang:

تقوم هذه الإضافة بتوفير الموقع بعدة لغات يختارها المستخدم. حيث يقوم المدون عند كل منشور يريد إضافته بكتابة محتوى المنشور وكلماته المفتاحية وأقسامه بجميع اللغات التي يريد دعمها في الموقع. تدعم الإضافة اللغات التي تكتب من اليمين إلى اليسار وبالعكس أيضاً. تعتمد هذه الإضافة بشكل جوهري على الترجمة اليدوية لجميع أقسام الموقع من قبل صاحب المدونة لتوفيرها لاحقاً للمستخدم بحسب طلبه.

وهناك عدة محاسن لهذه الإضافة أهمها:

- نشر محتوى الموقع بالعديد من اللغات.
- توفر أنواع اشتراكات مختلفة (مجانية، شهرية، سنوية).
- سهولة الاستخدام، حيث يقوم المدون فقط بكتابة المحتوى كاملاً باللغة المطلوبة وإسناده لها وتتولى الإضافة العرض الصحيح لكل محتوى مع اللغة الموافقة له.

كما أنه يوجد سيئة لهذه الإضافة وهي أنها تعتمد على الترجمة اليدوية للمحتوى والتي تعتبر مهمة مجهدة لصاحب المدونة.

#### إضافة الترجمة التلقائية GTranslate:

تقوم الإضافة بالترجمة التلقائية للمدونة لأكثر من 103 لغة وذلك بالاعتماد على خدمة Google Translate للترجمة. تدعم اللغات التي تكتب من اليمين لليسار أيضاً. توفر اشتراك مجاني بمزايا محدودة واشتراك مدفوع بكامل المزايا. قابلة للتوافق مع أساسيات SEO لتحسين تصنيف المدونة ضمن محركات البحث. تعتمد الإضافة على تقنيات الحوسبة السحابية لتوفير خدمة الترجمة بالسرعة القصوى. تستخدم شبكات عصبونية ذات أداء عالي ومقارب للترجمة البشرية ضمن برنامج الاشتراك المدفوع.

وهناك عدة محاسن لهذه الإضافة أهمها:

- تدعم اللغة العربية.
  - توفير ترجمة المحتوى لأكثر من 100 لغة.
  - توفير ميزة الترجمة التلقائية دون الحاجة لتدخل المستخدم.
- كما أنه يوجد سيئة لهذه الإضافة وهي أنها إضافة مدفوعة وغير مجانية.

#### الإضافة ClassifAi:

تقدم هذه الإضافة العديد من الخدمات التي تعتمد على تقنيات التعلم التلقائي لتوفر خدمات مثل: استخراج الكلمات المفتاحية من النصوص – توليد شرح مبسط لمحتوى الصورة ضمن صفحات المدونة (يوضع ضمن الرمز Alt ضمن صفحات HTML) – اقتصاص الصور الذكي – توليد كلمات مفتاحية للصور. تعمل هذه الإضافة بالاعتماد على خدمات معالجة اللغات الطبيعية المقدمة من IBM Watson وخدمات معالجة الصور المقدمة من Microsoft Azure.

وهناك عدة محاسن لهذه الإضافة أهمها:

- تحوي نسخة مجانية ونسخ مدفوعة تحوي مزايا أكثر.
- توفر خدمة استخراج الكلمات المفتاحية تلقائياً للمحتوى (نصوص أو صور).
- توفر خدمة توليد شرح للصور وإضافته للصورة ضمن الموقع.
- توفر ميزة القص الذكي للصورة بحيث يتم إبقاء الأشياء المهمة في الصورة وقص الأجزاء الغير مهمة

أما بالنسبة لمساوئ هذه الإضافة فهي كالتالي:

- جميع الخدمات السابقة لا تدعم اللغة العربية.
- لا تتميز بسهولة الاستخدام: تحتاج لإعداد حساب على كل من IBM, Microsoft Azure, من أجل أن تعمل ClassifAi.

## النظام Joomla

هو إدارة محتوى إلكتروني مجاني ومفتوح المصدر، مبنيّ بلغة php ويعتمد على مبدأ البرمجة غرضية التوجه OOP ويعتمد على قواعد البيانات MySQL، يضم خدمات عديدة منها: ميزة النسخ المخبأة لسرعة تحميل المواقع و RSS-Feedback وميزة البحث ويدعم خدمات الترجمة المعتمدة على الموقع الجغرافي للمستخدم لتحديد لغته بشكل افتراضي.

يعد هذا النظام هو ثاني أشهر نظام عالمياً لإدارة المحتوى على الويب عالمياً. ويعتمد هذا النظام على معمارية الملحقات Extensions، وهي عبارة عن برامج مساعدة تُقدم العديد من الوظائف. لكن من الجدير بالذكر أنه حتى الآن لا يوجد الكثير من البرامج الملحقة التي تعتمد على الذكاء الاصطناعي لتوفير خدمات ذكية للمستخدمين سوى خدمة روبوت الدردشة كونها الميزة الأكثر طلباً لمواقع التجارة خصوصاً والتي تعتبر من المواقع الأكثر انتشاراً بالمقارنة مع باقي المواقع.

## النظام Drupal

هو نظام لإدارة المحتوى بمواقع شبكة الإنترنت، ويمكن أن يعمل أيضاً كنظام للتدوين أو كنظام لبناء تطبيقات الويب. يستخدم معمارية الوحدات Modules. بدأ تصميمه ليكون برنامج لإنشاء المنتديات، ولكن تطور إلى أبعد من هذا نظراً لتصميمه الجيد وبنيته المرنة التي يمكن التحكم بها وتطويرها بسهولة. يستخدم Drupal في مواقع شهيرة مثل Debian Planet و Spread Firefox. حيث يعد البرنامج الأكثر شهرة في مجال بناء المجتمعات القائمة على الويب. Drupal مكتوب بلغة برمجة php.

بداية من Drupal 6، أصبح دعم اللغة العربية (واللغات التي تكتب من اليمين لليسار عموماً) مدمجاً في Drupal، حيث يمكن للنظام عرض المواقع من اليمين لليسار ببساطة ودون أي تعديلات في السمات المستخدمة.

بالمقارنة مع الإضافات التي تقدم الخدمات نفسها التي يقدمها نظامنا المقترح:

### الوحدة ArabicTranslation:

وهي الوحدة الأساسية للترجمة في النظام. تحوي الترجمة الخاصة بالعناصر الأساسية ضمن أي موقع يتم إنشاؤه بواسطة النظام السابق. هذه الوحدة خرجت من إطار الدعم والتحديث وأصبحت قديمة.

محاسن هذه الوحدة:

ترجمة معدة مسبقاً لكامل عناصر الموقع باللغة العربية

مساوئ هذه الوحدة:

الترجمة محصورة فقط لواجهات الموقع وليس للمحتوى.



### الوحدة ImageTagger:

تمكن صاحب الموقع من إضافة نقاط علامة على الصور المرفقة ضمن المنشورات وترميز كل نقطة بكلمة أو عبارة معينة. تستخدم كثيراً في مواقع التسوق حيث يوضع pin عند كل غرض من أغراض الصورة وعند الضغط عليه يتم عرض الكلمة المفتاحية tag المناسب لهذا الموضع من الصورة.

ومن محاسن هذه الوحدة أنه يمكن إضافة كلمات مفتاحية تساعد المستخدم على التعرف على عناصر الصورة والاستفادة منها بإظهار نتائج مشابهة لما يريد.

أما مساوئ هذه الوحدة أنها:

- يتم إضافة الكلمات المفتاحية المعبرة عن عناصر الصورة بشكل يدوي.
- لا تدعم اللغة العربية في وضع الكلمات المفتاحية.

### الوحدة KeywordAnalysis:

تستخدم هذه الوحدة من أجل استخراج الكلمات المفتاحية التي تعبر عن نص المنشور بشكل تلقائي. تعمل بشكل محلي دون الحاجة للاعتماد على واجهة تخاطب برمجية خارجية API. حيث يتم استخراج الكلمات المفتاحية بالاعتماد على خوارزميات مفتوحة المصدر.

عند استخراج الكلمات المفتاحية يتم عرضها ضمن واجهة مخصصة مع نسبة مئوية تعبر عن مدى التأكد من كون تلك الكلمة مناسبة لمحتوى المنشور ليقوم المستخدم بالاختيار من القائمة، الكلمات الأكثر صلة بالمحتوى ومن ثم تأكيد خياره.

ومن محاسن هذه الوحدة أنها توفر ميزة الاستخراج التلقائي للكلمات المفتاحية التي تعبر عن محتوى المنشور واقتراحها على المدون ليختار منها.

وأما مساوئ هذه الوحدة أنها لا تدعم اللغة العربية.

## Contentful النظام

هو نظام إدارة محتوى متقدم غير مجاني، حيث أنه ينتمي للأنظمة التي لا تحوي على واجهات تفاعل Headless CMS ولكن من أهم ميزاته أنه عبارة عن منصة إدارة محتوى متكاملة يتم تقديمها من خلال API مما يسمح باستخدامه بالتوافق مع أي قناة اتصال رقمية، أي أن المستخدم ليس محصوراً بمواقع الويب فقط.

وهي بنفس الوقت تعتبر سيئة للمستخدمين الذين لا يجيدون صناعة تطبيقات الويب لعرض المحتوى الخاص بهم فهم المسؤولين عن بناء الموقع الخاص بهم وربطه مع المحتوى المقدم من مُخدمات Contentful عبر API.

يعتمد هذا النظام لتقديم الخدمات على مبدأ الوحدات Modules، فيمكن إضافة وحدات مخصصة يختارها المستخدم لمساعدته في صناعة المحتوى الخاص به، وبالمقارنة مع الوحدات التي تقدم الخدمات نفسها التي يقدمها نظامنا المقترح:

الوحدة Arabic Translation Module:

تقوم بالترجمة الآلية للمحتوى من اللغة الإنجليزية إلى اللغة العربية وبالعكس، ولكن هذه الوحدة أصبحت قديمة جداً ومهملة (Obsolete) وغير مدعومة.

الوحدة Image Tagger Module:

تساعد المستخدم على إضافة دبابيس على الصور (pins) يمكن من خلالها ربط كل دبوس مع رابط لمنتج أو منشور آخر، وبالتالي ساهمت في إعطاء الصورة معنى من خلال الربط اليدوي مع منتجات أو منشورات أخرى.

الوحدة Keyword Analysis:

وهي وحدة من أجل تنميط النصوص بشكل تلقائي (استخراج الكلمات المفتاحية)، وتعمل من دون الاعتماد على مكاتب خارجية أو API، حيث عند كتابة المحتوى تقوم باقتراح كلمات مفتاحية مناسبة للنص بشكل تلقائي ويقوم المحرر بالاختيار من خلال واجهة تفاعلية الكلمات المناسبة والتي يريد تثبيتها مع النص. ولكن يعد من أهم مساوئها أنها لا تدعم اللغة العربية.

## نظام Strapi

هو نظام إدارة محتوى مفتوح المصدر مبني بواسطة NodeJs مع واجهات تحكم بسيطة وعصرية.

ومن مميزاته:

- خيارات متعددة لأنواع المحتوى الممكن إضافتها للنظام (collectionTypes-singleType-component)
- مبني بواسطة NodeJs ويقدم سرعة وحماية عالية جداً
- يعتمد على نظام Plugins
- يسمح باستخدام العديد من أنواع ال database ويشمل Sql و NoSql
- يقدم Rest API بالإضافة إلى GraphQL API
- يدعم تعدد اللغات للمحتوى (من ضمنها اللغة العربية)

عيوبه:

- حجم المشروع كبير نوعاً ما والتعديل على البنية الداخلية له ليس بالأمر السهل.
- كل مستخدم سيقوم بتحديد ال database الخاصة به ومكان رفع المشروع وتخليده بالكامل.

الفرق بين مثل هذا النوع من ال Headless CMS وأنواع أخرى مثل Contentful هي أن Contentful يقدم خدماته على شكل software as a service أو saas حيث يكون تخديم المشروع بشكل كامل من قبل مقدم الخدمة وهو المتحكم بنوع ال database ومكان التخديم وكل شيء متعلق بعملية ال deployment.

فيمكننا القول إن من عيوب أي Open Source Headless CMS أنه self-hosted أي بحاجة لشخص لديه معرفة تقنية بكيفية ربط ال database وتخديم المشروع من خلال سيرفر خاص. وبالمقابل النوع الآخر يقدم لك CMS جاهز ومخدّم بشكل مباشر لكن بهذه الحالة فأنت تخاطر بتوقف عملك بحال حدوث أي مشكلة من قبل المخدم الخاص بمقدم الخدمة. كما أن مثل هذه الخدمات تكون مدفوعة ومقدمة من قبل شركات ولا تتواجد بشكل open source

## نظام Wagtail

هو نظام إدارة محتوى مفتوح المصدر مبني بواسطة Django مع واجهات تحكم بسيطة، من مميزاته:

- يمكن استخدامه وكأنه Headless CMS حيث يمكن استخدامه للوصول إلى البيانات على شكل API لكن بدعم أقل من Headless CMS الطبيعي، فهناك أنت تتحكم بأنواع البيانات بالكامل، أما هنا فأنت لديك صفحات فيها محتوى فأنت يمكنك الوصول لهذا المحتوى على شكل json فقط.
- يدعم تعدد اللغات
- يدعم العديد من SQL Databases

أما عيوبه فهي:

- واجهة التحكم غير سهلة الاستخدام للمستخدم الغير تقني.
- لا يوجد دعم NoSQL Databases.

## نظام Apostrophe

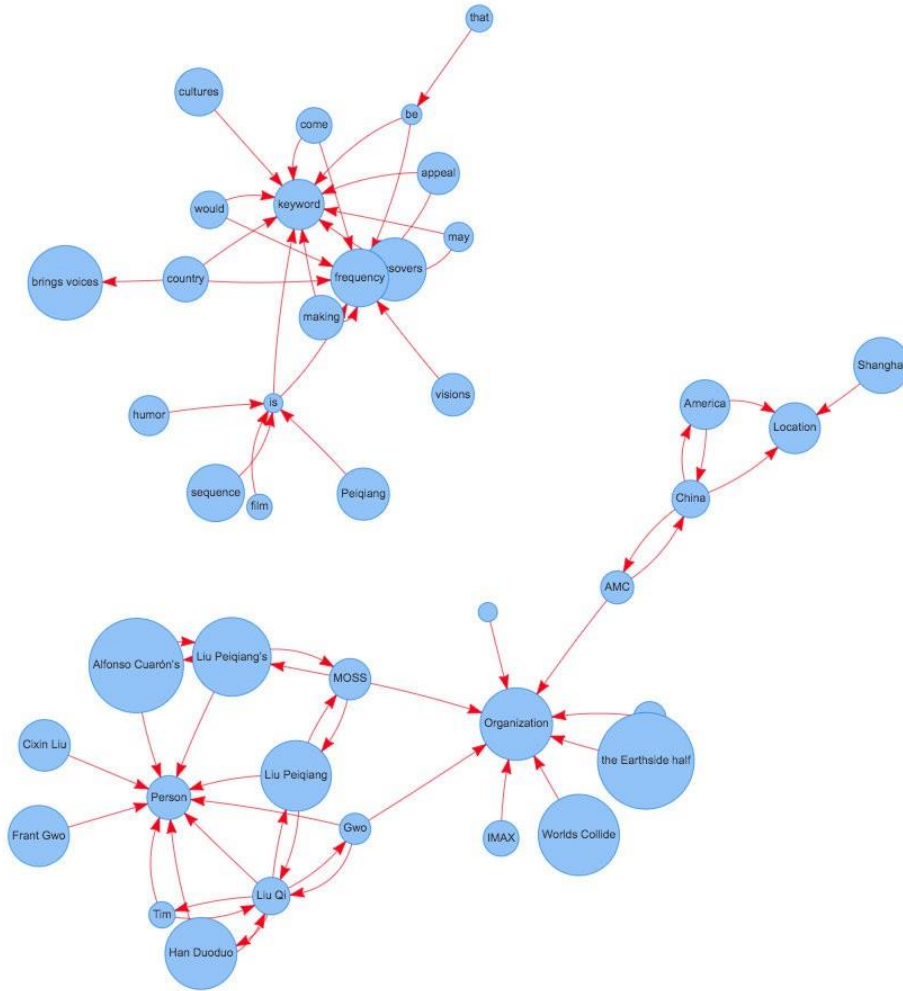
وهو نظام إدارة محتوى مفتوح المصدر مبني بواسطة NodeJs و MongoDB مع واجهات تحكم مميزة جداً، ومن

مميزاته:

- تحكم كامل بالشكل النهائي للصفحات (Front-End) الخاصة بالموقع من خلال Editor WYSIWYG، أي أن ما تراه أثناء بناء الصفحة هو ما سيظهر في الصفحة النهائية.
- يمكن استخدامه وكأنه Headless CMS حيث يمكن استخدامه للوصول إلى البيانات على شكل API لكن بدعم أقل من Headless CMS الطبيعي.
- يعتمد على Nunjucks من أجل بناء templates لاستخدامها لاحقاً في أرجاء الموقع مما يعطي تحكم أكبر وأبسط لعملية إضافة المحتوى.

عيوبه:

- لا يوجد دعم SQL Database حيث يتم استخدام Mongoddb بشكل تلقائي.
- تحتاج لتنصيب Imagemagick على الجهاز الذي تعمل عليه من أجل رفع الصور (وهي عبارة عن مجموعة برامج مجانية ومفتوحة المصدر تعمل عبر الأنظمة الأساسية لعرض الصور النقطية وإنشائها وتحويلها وتعديلها وتحريرها).
- يعمل بشكل أفضل مع Linux based system وبحال كنت تعمل على Windows فينصح باستخدام windows subsystem for Linux او WSL.



رسم توضيحي 21- استخراج العبارات المفتاحية بطريقة Text Rank

## 2- استخراج العبارات المفتاحية

تسمح العديد من المواقع لمستخدميها بالقيام بتحديد الكلمات المفتاحية للنصوص التي قاموا بإنشائها مثل موقع Tumblr، وكذلك تقوم العديد من المواقع الإخبارية والمدونات بكتابة الكلمات المفتاحية لمنشوراتها بشكل يدوي حيث يسهل وجود الكلمات المفتاحية عملية البحث عن موضوع معين، ويظهر ترابط المنشورات ببعضها، ويساعد في تحسين ترتيب المنشور على مواقع البحث.

على الرغم من أهمية كتابة الكلمات المفتاحية لكل منشور إلا أن القيام بهذه العملية بشكل يدوي يستغرق جهداً وزمناً كبيرين بالإضافة إلى أن هذه العملية مكررة ومملة، ولذلك ظهرت طرق عديدة لتنفيذ هذه المهمة.

تقسم الطرق التي تقوم بعملية الأتمتة إلى قسمين رئيسيين:

(supervised methods, Unsupervised methods)

## Unsupervised methods -2.1

### Graph based Methods 2.1.1

تقوم هذه الطريقة على بناء بيان للمستند ثم تقوم بتقييم الكلمات بناءً على موضع هذه الكلمات ومركزها بطريقة مشابهة لطريقة الـ page rank، ولها عدة أنواع تختلف عن بعضها في طريقة إنشاء البيان وبطريقة حساب الأوزان وسنشرح كل منها على حدة [6][7]:

#### 1. Text rank

تعتمد هذه الخوارزمية بشكل أساسي على الـ page rank، حيث يتم تشكيل بيان موجه، وتكون العقد ضمن هذا البيان هي عبارة عن كلمات كما في الرسم التوضيحي (21):  
يتم تمثيل خوارزمية page rank باستعمال بيان غير موزون، أما خوارزمية Text rank فتتمثل من خلال بيان موزون لأنه يمكن الاستفادة من ترابط الكلمات الموجود في اللغة.  
وبالتالي يتم تعديل المعادلة (1) لتصبح كما يلي:

$$WS(V_i) = (1 - d) + d * \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} WS(V_j) \quad (5)$$

حيث  $W_{ij}$  تشير إلى قوة العلاقة بين العقدة  $i$  والعقدة  $j$ .

لتشكيل البيان يجب أن يتم تقسيم النص إلى كلمات ومن ثم يجب إيجاد الـ POS لهذه الكلمات، ويتم حذف الكلمات الغير مرشحة لتكون كلمات مفتاحية وذلك بناءً على الـ POS الخاص بهذه الكلمات، فمثلاً حروف الجر في اللغة العربية غير مرشحة لتكون كلمات مفتاحية، أما الأسماء والصفات فغالباً ما تكون هي الكلمات المفتاحية.

عند الانتهاء من عملية ترشيح الكلمات، يتم تشكيل البيان بحيث يتم وضع كل كلمة مرشحة في عقدة، ويتم إنشاء وصلات بين العقد التي تكون كلماتها متتالية ضمن النص.

أصبح لدينا بيان غير موجه وغير موزون لأن جميع الوصلات حالياً لها نفس الوزن، يتم حساب الأوزان بين الوصلات من خلال تقسيم النص إلى مجالات، ودراسة المسافة بين الكلمات ضمن هذا المجال.  
يتم تشغيل الخوارزمية الموضحة في المعادلة (5) للحصول على أوزان الرؤوس.

عادة ما يتم اعتماد الثلث الأعلى وزناً من العقد كلمات مفتاحية، ويمكن اعتماد أعلى  $N$  رقم كلمات مفتاحية.

يمكن تحسين الخوارزمية السابقة من خلال حساب وزن عقدتين متتاليتين أو أكثر معاً من العقد المرشحة (العقد ذات الثلث الأعلى وزناً)، وقد تبين من خلال التجريب أن استعمال تسلسل من كلمتين يعد حلاً جيداً.

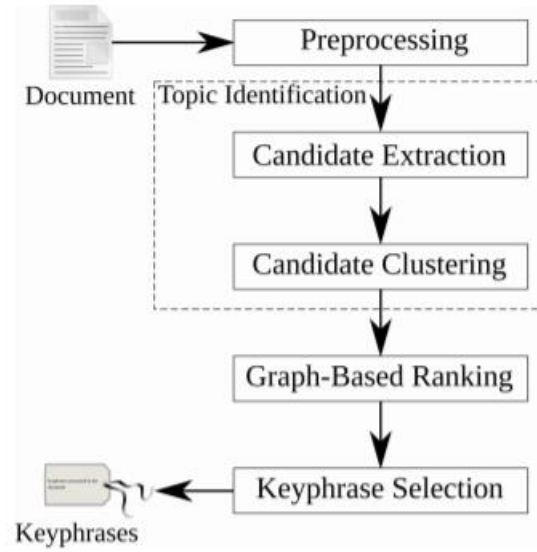
من أهم الانتقادات التي تم توجيهها لهذه الخوارزمية أنها تعتمد عند حساب ال bi-grams على العقد ذات الثلث الأعلى وزناً، في حين أنه من الممكن أن يكون لعقدتين Uni-grams أوزاناً قليلة كل على حدا، ولكن عند القيام بجمعها معاً من الممكن الحصول على وزن مرتفع.

وهناك انتقاد أخرى بأن عملية تحديد مجال التقسيم الذي يستعمل لحساب أوزان العقد يتم بشكل تجريبي. [8]

## 2. Single rank

تعتبر هذه الخوارزمية تنفيذاً لخوارزمية Text rank مع وجود اختلاف واحد بينهما، حيث يتم حساب ال N-grams لجميع العقد وليس فقط للعقد ذات الثلث الأعلى وزناً، لأنه من الممكن أن يكون لعقدتين Uni-grams أوزاناً قليلة كل على حدا، ولكن عند القيام بجمعها معاً من الممكن الحصول على وزن مرتفع.

ومن عيوب هذه الطريقة أنها لازالت تعتمد على تحديد مجال التقسيم الذي يستعمل لحساب أوزان العقد يتم بشكل تجريبي. [6]



رسم توضيحي 22- يوضح خطوات عمل خوارزمية ال Topic rank

### 3. Topic rank

تقوم هذه الخوارزمية على تصنيف النص إلى مجموعة من المواضيع، حيث يكون الموضوع هو مجموعة من الكلمات المرشحة المتشابهة، يتم بعد ذلك ترتيب هذه المواضيع وفقاً لأهميتها في المستند، ويتم اختيار عبارة رئيسية واحدة لكل موضوع، فينتج لدينا كلمات مفتاحية مرتبة بالأهمية.

تتكون Topic rank من ثلاثة مراحل موضحة في الرسم التوضيحي (22) وهي:

#### I. تعريف المواضيع:

يتم أولاً استخراج جميع الكلمات المرشحة لأن تكون كلمات مفتاحية وذلك عن طريق اختيار الكلمات الاسمية، ويتم اعتبار كل مجموعة متتالية من الكلمات الاسمية عبارة مرشحة واحدة.

يتم بعدها تصنيف الكلمات إلى مجموعة من المواضيع حسب درجة التقارب بين هذه الكلمات (Clustering) حيث تعتبر كلمتين متشابهتان إذا كان يوجد تقارب بينهما لا يقل عن 25%، ونتيجة لذلك من الممكن أن تنتمي نفس الكلمة لموضوعين مختلفين.

#### II. ترتيب المواضيع حسب الأهمية:

ويتم ذلك من خلال بناء بيان يقوم بترتيب المواضيع حسب درجة أهمية هذه المواضيع.

ويكون هذا البيان من النوع الموزون والكامل (أي يوجد وصلات بين أي عقدتين ضمن هذا البيان)، حيث تكون العقد هي عبارة عن المواضيع، ويتم حساب وزن الحفة بناءً على قوة العلاقة الدلالية بين هذين الموضوعين، ويتم تحديد قوة العلاقة الدلالية بناءً على مدى قرب الكلمات الرئيسية المرشحة للموضوعين من بعضهما في المستند ويتم تعريفها على النحو التالي:

$$w_{i,j} = \sum_{c_i \in t_i} \sum_{c_j \in t_j} \text{dist}(c_i, c_j) \quad (6)$$

$$\text{dist}(c_i, c_j) = \sum_{p_i \in \text{pos}(c_i)} \sum_{p_j \in \text{pos}(c_j)} \frac{1}{|p_i - p_j|} \quad (7)$$

$\text{dist}(c_i, c_j)$ : هي المسافة بين موضع كلمتين مفتاحيتين مرشحتين  $i, j$  في المستند.

$t_i, t_j$ : الموضوع  $i$  والموضوع  $j$ .

$c_i, c_j$ : الكلمات المفتاحية المرشحة  $i, j$ .

$\text{pos}(c_i)$ : جميع مواضع الكلمة المفتاحية  $i$ .

بطريقة حساب أوزان الوصلات بين المواضيع يكون قد تم التخلص من مشكلة تحديد مجال التقسيم الذي يستعمل لحساب أوزان العقد والذي كان من أحد عيوب كل من ال Single rank وال Text rank، وذلك لأن ال Topic rank لا يتطلب الحاجة إلى اختيار حجم النافذة لأنه بدلاً من ذلك يحسب الأوزان بناءً على المسافات بين المواضيع.

بعدها يتم تطبيق خوارزمية Text rank لترتيب المواضيع حسب الأهمية، مع الاستفادة من كون البيان كامل والذي يعطي رؤية أكثر شمولاً.

### III. اختيار الكلمات المفتاحية

بعد الانتهاء من عملية ترتيب المواضيع حسب أهميتها، يتم اختيار أعلى  $k$  موضوع ويتم اختيار عبارة مفتاحية واحدة لكل موضوع، فيكون الناتج النهائي هو  $k$  عبارة مفتاحية لكامل المستند.

يتم اختيار العبارة المفتاحية الأكثر أهمية بوحدة من الطرق التالية:

- 1- يتم اختيار العبارة المفتاحية التي ظهرت أولاً ضمن المستند، وتقوم هذه الطريقة على اعتبار أن الكلمات المفتاحية الأهم تكون في بداية المستند.
- 2- يتم اختيار العبارة المفتاحية الأكثر تكراراً ضمن المستند.



3- يتم اختيار العبارة التي تقع في منتصف المصنف، أي العبارة الأكثر تشابهاً مع باقي العبارات الموجودة ضمن نفس الموضوع.

من عيوب هذه الطريقة أنه من الممكن أن تكون ثاني عبارة مرشحة لموضوع ما أكثر أهمية من أعلى عبارة لموضوع آخر، ولكن هذه الطريقة ستأخذ أعلى عبارة مرشحة لكل موضوع.

كما أن هذه الخوارزمية تعتبر خوارزمية بطيئة التنفيذ وذلك لأنه يتم حساب انتماء كل كلمة لكافة المواضيع ضمن المستند وهذا يحتاج زمناً كبيراً. [Z]

#### 4. Position rank

تعمل هذه الخوارزمية في ثلاث خطوات رئيسية:

##### I. بناء الرسم البياني (Graph) على مستوى الكلمات

في هذه المرحلة يتم بناء بيان Graph غير موجه، بعد تطبيق POS Filter على كلمات النص، ثم يتم اختيار الكلمات التي تمثل أسماء (Nouns) أو صفات (adjectives)، ثم يتم اختيار هذه الكلمات لتمثل عقداً (Node) ضمن البيان.

ويتم إضافة وصلة (Edge) بين عقدتين بناءً على co-occurrence relation للكلمات التابعة لها وهي علاقة مفادها أنه يتم إضافة وصلة بين عقدتين في حال ورود الكلمتين التابعة لهما ضمن النص بحيث يكون الفرق بينهما أقل من قيمة عتبة معين  $W$ ، أو بصيغة أخرى ورود الكلمتين (الأولى تابعة للعقدة الأولى والثانية تابعة للعقدة الثانية) ضمن مجال طوله  $W$  ضمن النص (نفس الخطوة الأولى في خوارزمية الـ [Text rank](#)).

ويكون وزن هذه الوصلة مساوي لعدد مرات ظهور الكلمتين التابعتين لهذه الوصلة ضمن مسافة لا تتجاوز  $W$ ، أو بصيغة أخرى عدد المجالات من طول  $W$  والتي تحوي الكلمتين التابعة لعقد لهذه الوصلة.

يكون خرج هذه المرحلة هو مصفوفة الأوزان الخاصة بالنص  $M$  حيث تكون قيمة الخلية  $M_{ij}$  مساوية لوزن الوصلة بين العقدة  $u_i$  والعقدة  $v_j$  وبحال عدم وجود وصلة تكون القيمة مساوية للصفر.

##### II. تطبيق خوارزمية Page-Rank بالاستفادة من أماكن ورود كل كلمة في النص

في هذه المرحلة يتم حساب شعاع Position Page Rank لعقد البيان السابق، حيث يتم تهيئة شعاع الأوزان  $S$  بقيم ثابتة مساوية لـ  $\frac{1}{|V|}$ ، ثم يمكن حساب الشعاع في المرحلة التالية بتطبيق العلاقة المبدئية التالية:

$V$ : العقد ضمن البيان.

$$S(t + 1) = \tilde{M}.S(t) \quad (8)$$

$t$ : المرحلة الحالية.

$\tilde{M}$ : هي مصفوفة الأوزان بعد تطبيق العلاقة التالية عليها:

$$\widetilde{m}_{ij} = \begin{cases} m_{ij} / \sum_{j=1}^{|V|} m_{ij} & \text{if } \sum_{j=1}^{|V|} m_{ij} \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

ولكن لضمان عدم وقوع الخوارزمية السابقة في حلقات وضمن الانتقالات العشوائية يتم إضافة عامل  $\alpha$  لتصبح المعادلة بالشكل التالي:

$$S = \alpha . \tilde{M} . S + (1 - \alpha) . \tilde{p} \quad (10)$$

$\tilde{p}$ : شعاع يعبر عن امكانية الانتقال من العقدة الحالية لأي عقدة أخرى.

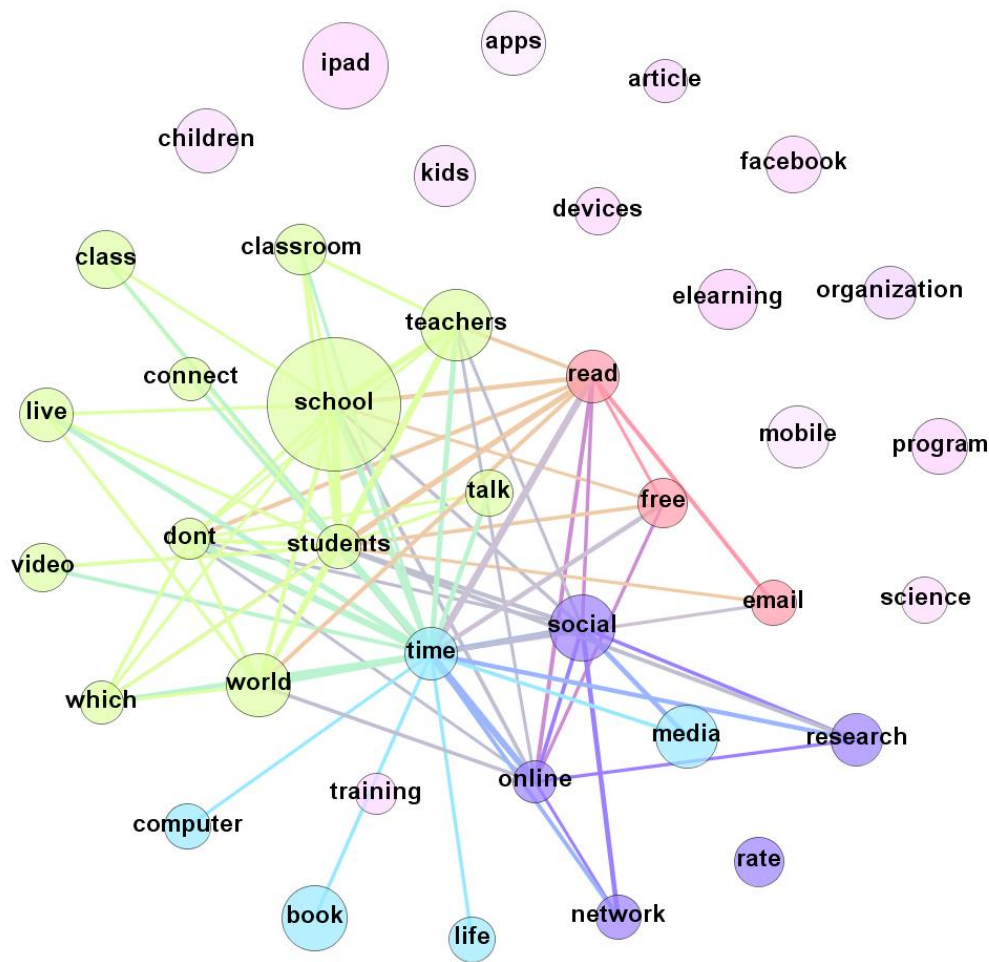
وهنا تأتي فائدة هذه الخوارزمية حيث أنها تقوم بحساب هذا الشعاع بناء على بعد الكلمة عن بداية النص وعدد مرات ظهور هذه الكلمة ضمن النص.

حيث يتم تعريف معامل  $p_i$  لكل كلمة على أنه:

مقلوب ترتيب أو بعد الكلمة عن بداية النص، وبحال ورود الكلمة أكثر من مرة فتكون قيمة  $p_i$  مساوية لمجموع قيم المقلوب لترتيب كل مرات ظهور هذه الكلمة ضمن النص

ثم يتم تعريف الشعاع  $\tilde{p}$  على الشكل التالي:

$$\tilde{p} = \left[ \frac{p_1}{p_1+p_2+\dots+p_{|V|}}, \frac{p_2}{p_1+p_2+\dots+p_{|V|}}, \dots, \frac{p_{|V|}}{p_1+p_2+\dots+p_{|V|}} \right] \quad (11)$$



رسم توضیحي 23- بیان مبني بطريقة ال position rank

لتصبح العلاقة النهائية لحساب وزن عقدة  $v_i$  بالشكل التالي:

$$S(v_i) = (1 - \alpha) \cdot \tilde{p}_i + \alpha \cdot \sum_{v_j \in Adj((v_i))} \frac{w_{ji}}{o(v_j)} S(v_j) \quad (12)$$

$\alpha$ : معدل الخامد وقيمته بين ال 0 وال 1 ويمثل احتمال الانتقال من عقدة إلى عقدة عشوائية.

يتم حساب هذه العلاقة بشكل تكراري حتى يصبح الفرق بين كل دورتين أصغر من عتبة معينة أو الوصول لعدد مرات تكرار محدد (100 دورة مثلا).

### III. اختيار الكلمات المرشحة ذات الوزن الأعلى

بعد حساب شعاع الأوزان لكل كلمة في المرحلة السابقة، يتم دمج الكلمات المرشحة (عقد البيان) والتي ظهرت بجانب بعضها في النص لتكوين جمل من 3 كلمات أو أقل كما هو موضح في الرسم التوضيحي (23)، ثم يتم حساب وزن كل جملة بناءً على مجموع أوزان الكلمات التابعة لها ضمن شعاع

الأوزان السابق، وفي النهاية يتم ترتيب الجمل بناء على هذه الأوزان ويكون الخرج هو الجمل ذات الوزن أو الأهمية الأعلى.

تتميز هذه الخوارزمية بأنها تأخذ 3 معايير أثناء حساب الأوزان:

- جميع مواضع ورود الكلمة ضمن النص.
- بعد كل كلمة عن بداية النص.
- عدد مرات ظهور الكلمة ضمن النص.

وقد أثبتت هذه الطريقة فعاليتها أكثر من الطرق التي تعتمد فقط على أول ظهور للكلمة أو عدد مرات ظهور الكلمة ضمن النص. [9]

إن جميع طرق ال Graph based Methods ما عدا ال Topic rank لها عيب مشترك وهو أنها تحسب لجميع العبارات المرشحة نفس الأهمية في حين يجب الأخذ بعين الاعتبار جميع المواضيع الموجودة ضمن النص، ونتيجة لهذا العيب من الممكن أن تكون العبارات المفتاحية الناتجة لا تستوفي كامل النص.

## Statistical Methods 2.1.2

يتم في هذه الطريقة تحديد العبارات المفتاحية بشكل احصائي يتعلق بعدد مرات تواجد هذه العبارة في النص وذلك وفق إحدى الطرق التالية: [7]

### 1. (TFIDF)

تعمل خوارزميات التعلم الآلي مع البيانات الرقمية ولا يمكننا استخدام البيانات النصية المتوفرة "كما هي".

لذلك يمكن تعريف TFIDF بأنها عملية إحصاء رقمي تعكس مدى أهمية كلمة لمستند أو مجموعة من المستندات، وتعد هذه الطريقة من أكثر الطرق المستعملة إلى يومنا هذا.

يتم حساب احتمال ال TFIDF لكل عبارة، ثم يتم اختيار العبارات الأعلى احتمالاً ذلك وفق مرحلتين:

### -1 Term Frequency (TF):

يمثل مقياساً لعدد مرات تكرار كلمة  $w$  في المستند  $d$ ، وبهذا يمكن تعريفه أنه عدد مرات تكرار كلمة  $w$  مقسوماً على عدد الكلمات الكلي في المستند.

$$TF(w, d) = \frac{\text{عدد مرات تكرار } w}{\text{العدد الكلي للكلمات في المستند } d} \quad (13)$$

## 2- Inverse Document Frequency (IDF):

يمثل مقياساً لأهمية الكلمة، حيث أنه يوجد كلمات تتكرر كثيراً لكن أهميتها تكون قليلة مثل (من، على، في،.... الخ).  
تحسب كما يلي:

$$IDF(w, D) = \ln \left( \frac{\text{عدد المستندات الكلي } D}{\text{عدد المستندات التي تحوي الكلمة } w} \right) \quad (14)$$

وعندها يكون القانون النهائي كما يلي:

$$TFIDF(w, d, D) = TF(w, d) * IDF(w, D) \quad (15)$$

تم استعمال اللوغاريتم في العلاقة (14) لأن قيم ال TF بين الصفر والواحد، وعدم استخدام اللوغاريتم يؤدي إلى قيم مرتفعة في IDF لبعض الكلمات، ومن أجل عدم السماح IDF بالسيطرة على الناتج الكلي يتم استخدام اللوغاريتم. من الواضح أن استعمال هذه الخوارزمية بحاجة إلى نصوص من نفس نوع النصوص التي سيتم معالجتها لكي يكون حساب ال IDF بشكل صحيح.

من عيوب هذه الطريقة أنها تقوم باعتبار الكلمات المترادفة (funny , humorous) كلمات منفصلة أي أنه في حال كان يوجد كلمتين لهما نفس المعنى فيجب أن يتم اعتبارهما كلمة واحدة مكررة مرتين، في حين أن هذه الطريقة تعتبرهما كلمتين مكررتين مرة واحدة، ويمكن حل هذه المشكلة باستعمال تقنية Word2vec وهو عبارة عن شبكة عصبونية مدربة على عدد كبير من النصوص تعمل على إنشاء ارتباطات بين الكلمات، مما يسمح لها باكتشاف الكلمات المتشابهة واقتراح كلمات إضافية لجملته جزئية.

كما أن لاستعمال ال [TFIDF](#) عيب آخر وهو أنها عادة ما تميل لاختيار العبارات التي لها أقل عدد من الكلمات، لأن العبارات التي لها عدد كلمات كبير يكون تكرارها أقل ضمن النص. [10]

## 2. KPMiner

تتحقق هذه الخوارزمية من خلال 3 مراحل:

### 1. تحديد العبارات المرشحة:

يتم اعتماد تقطيع النص إلى عبارات مرشحة وذلك من خلال تقطيع النص عند الكلمات الشائعة (Stopward) وعند علامات الترقيم، ولكن نتيجة لهذا التقطيع سينتج لدينا عدد كبير من العبارات المرشحة، ويوجد عدد منها ليس له أهمية أو معنى ولتخفيف عدد هذه

العبارات يتم اعتماد العبارات التي تكرر جذرها  $n$  مرة ضمن النص (في النصوص الانكليزية تكون غالباً  $n = 3$ ، ولكن يمكن أن تزداد أو تنقص عن هذا العدد وذلك تبعاً لحجم النص). كما أنه يتم إضافة شرط آخر لتقليل عدد العبارات المرشحة وهو يتعلق بموضع ظهور العبارة الأول ضمن النص حيث أنه لوحظ بالتجريب على النصوص الطويلة أن العبارات المفتاحية لا يكون ظهورها الأول بعد عتبة معينة (تم اعتماد القيمة 400 كقيمة مثلى للعتبة من خلال عملية التجريب).

## 2. حساب أهمية العبارات المرشحة:

يتم اعتماد [TFIDF](#) لحساب أهمية العبارات المرشحة مع إضافة معامل جديد يعطي أهمية (تعزيزاً) للعبارات ذات عدد الكلمات الأعلى وذلك للتخلص من المشكلة الموجودة في [TFIDF](#)، ويحسب هذا المعامل كما يلي:

$$B_d = \frac{|N_d|}{|P_d| * \alpha} \text{ and if } B_d > \sigma \text{ then } B_d = \sigma \quad (16)$$

$d$ : مستند ما.

$|N_d|$ : العدد الكلي للعبارات المرشحة في المستند  $d$ .

$|P_d|$ : عدد العبارات المرشحة التي تتكون من أكثر من كلمة ضمن المستند  $d$ .

$\sigma$ : ثابت قيمة 2.3 .

$\alpha$ : ثابت قيمته 3.

كما ويتم إضافة معامل آخر هو يتعلق بالموضع وتصبح عندها المعادلة النهائية لحساب الأهمية كما يلي:

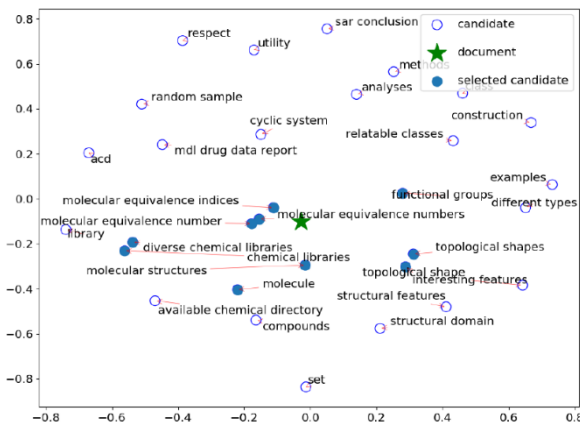
$$w_{ij} = tf_{ij} * idf * B_i * P_f \quad (17)$$

$P_f$ : عامل الموضع وتكون قيمته تساوي الواحد في حال لم يتم استعمال الموضع ضمن الحساب.

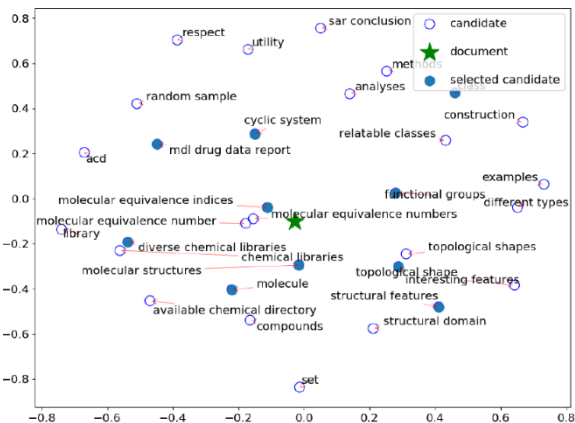
## 3. تحديد العبارات المرشحة النهائية:

يتم تحديد عدد  $N$  من قبل المستخدم وهو عدد العبارات التي يريد المستخدم لها الظهور، حيث يتم إظهار أعلى  $N$  عبارة من حيث الأهمية.

ما يميز هذه الخوارزمية أنها لا تضع حداً لعدد الكلمات ضمن العبارة المرشحة في حين تقوم معظم الخوارزميات الأخرى بتحديد عدد الكلمات ضمن العبارة المرشحة ب 3 على الأكثر. [11]



## أفضل العبارات بدون استخدام خوارزمية MMR



## أفضل العبارات مع استخدام خوارزمية MMR

رسم توضيحي 24- أفضل العبارات مع وبدون MMR

**Embed Rank .3**

وهي طريقة تستخدم لتحديد العبارات المفتاحية وتعتمد في أساسها على ال Embedding (تحويل الكلمات والجمل إلى أشعة في فضاء مستمر) وتتألف من ثلاث مراحل:

## 1- استخراج العبارات:

تحديد جمل أو أجزاء من النص يمكن أن تمثل عبارات مفتاحية، ويتم ذلك عن طريق تحديد العبارات التي تتطابق مع سلسلة POS محددة مثلا (اسم - اسم - صفة) أو عن طريق استخراج أسماء العلم من النص عن طريق NER ويمكن استخدام طرق أخرى أيضا.

## 2- تحويل النص والعبارات إلى أشعة:

في البداية نقوم بتنظيف النص والإبقاء فقط على الأسماء والصفات لأن ذلك يساعد على تخفيض الضجيج الناتج عن المعلومات غير المهمة عند تحويل النص إلى شعاع، ثم يتم تحويل العبارات التي تم استخراجها في المرحلة الأولى والنص إلى أشعة في فضاء  $Z_n$  ذات أبعاد محددة ( $n$ ) عن طريق Embedding model مدرب مسبقا.

### 3- حساب التشابه:

الفضاء  $Z_n$  ثم يتم اختيار العبارات الأكثر تشابهاً:

$$\text{Cos}_{\text{sim}} = \frac{\vec{C} \cdot \vec{\text{Doc}}}{\|\vec{C}\| * \|\vec{\text{Doc}}\|} = \sum_{i=1}^n \frac{C_i * \text{Doc}_i}{\sqrt{\sum_{i=1}^n C_i^2} * \sqrt{\sum_{i=1}^n \text{Doc}_i^2}} \quad (18)$$

لكن استخدام الزاوية بين  $\vec{Doc}$  و  $\vec{C}$  فقط كمقياس للتشابه يمكن أن يؤدي إلى ظهور عبارات مفتاحية متشابهة جداً كما هو موضح بالرسم التوضيحي (24). ومن أجل حل هذه المشكلة وإيجاد تنوع في العبارات المفتاحية يمكن استخدام خوارزمية الصلة الهامشية القصوى MMR:

$$MMR = \arg \max_{C_i \in R \setminus S} [\alpha \cdot \text{Cos}_{sim}(C_i, Doc) - (1 - \alpha) \max_{C_j \in S} \text{Cos}_{sim}(C_i, C_j)] \quad (19)$$

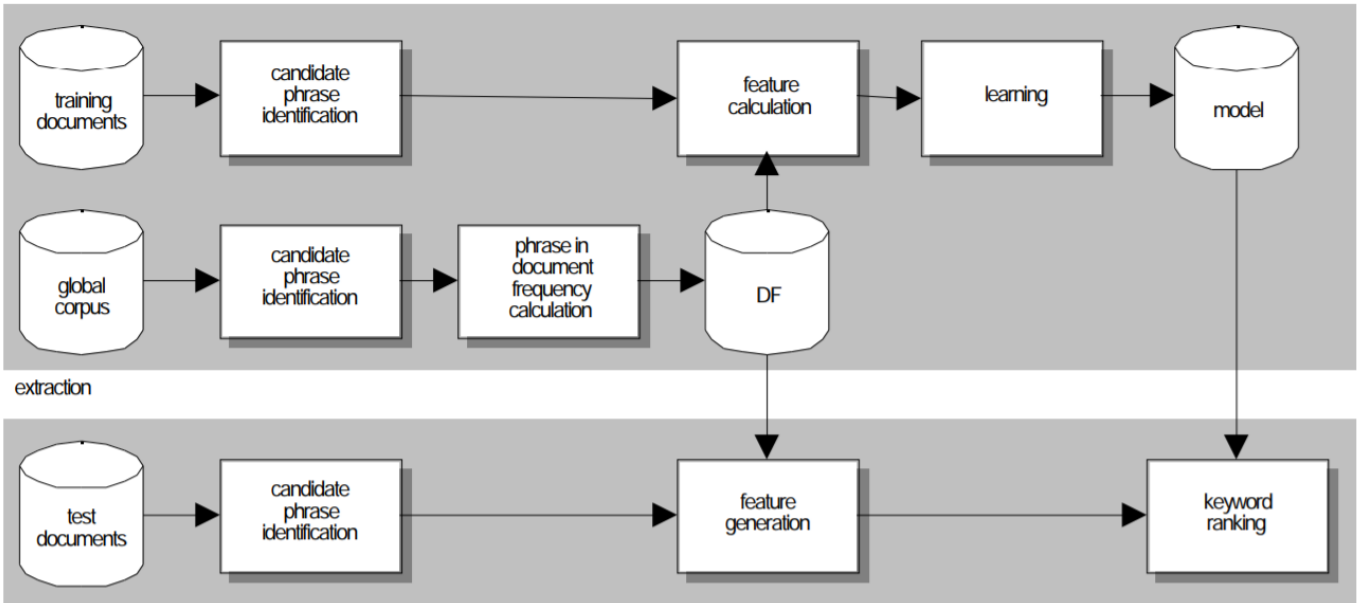
في البداية تقوم الخوارزمية بتهيئة مجموعة S تحوي شعاع العبارة الأكثر تشابهاً مع شعاع النص ومجموعة R تحتوي باقي اشعة العبارات في كل خطوة تقوم الخوارزمية بما يلي:

- 1- من أجل كل عنصر في R يتم حساب التشابه مع جميع عناصر S واختيار النتيجة الأكبر.
- 2- يتم تطبيق المعادلة (19) على المصفوفة السابقة ومصفوفة التشابه بين عناصر R و  $\vec{Doc}$ .
- 3- يتم اختيار الشعاع الموافق لأكبر نتيجة وإزالة هذا الشعاع من R وإضافته إلى S.

يتم تكرار الخطوات السابقة طالما أن R تحتوي عناصر [12].



#### training



رسم توضيحي 25- آلية التدريب والاستخراج في KEA

## Supervised methods -2.2

### Automatic Keyphrase Extraction (KEA)

تعتمد هذه الطريقة على المعاجم بشكل أساسي للقيام باستخراج العبارات المفتاحية.

تقسم عملية استخراج العبارات المفتاحية إلى مرحلتين:

- 1- عملية التدريب ويتم فيها تدريب نموذج من نوع نايف بايز على نصوص تحوي عبارات مفتاحية مستخرجة بشكل يدوي.
- 2- عملية استخراج العبارات المفتاحية من مستند جديد وذلك باستعمال النموذج المدرب.

يبين الرسم التوضيحي (25) كيف تتم عمليتي التدريب واستخراج العبارات المفتاحية، ويلاحظ وجود خطوات مشتركة سنبدأ بشرحها أولاً.

#### Candidate phrases ✓

تتم هذه العملية عبر ثلاث مراحل:

- **تنظيف النص:**  
ويتم فيها تقسيم النص إلى قطع (token)، ثم يتم حذف علامات الترقيم والأرقام وكل الرموز التي ليس لها دلالات مفيدة.
- **تحديد العبارات المرشحة:**  
يتم أولاً اعتماد طول أقصى للعبارات المرشحة (3 كلمات عادةً)، ويتم حذف العبارات التي تبدأ أو تنتهي بكلمات شائعة (stopward) وهي الكلمات تكون غير مرشحة لأن تكون كلمات مفتاحية مثل (عندما، أين، الخ..).
- **Case-folding and stemming**

يتم في هذه المرحلة حذف السوابق واللواحق من الكلمة وإيجاد جذر الكلمة، وتهدف هذه العملية بمعالجة العبارات التي لها نفس الحذر مع اختلاف في سوابق ولواحق كلماتها (proof net , proof nets).

#### ✓ Feature calculation :

من أجل كل عبارة مرشحة يتم حساب **TFIDF** وموضع الكلمة ضمن المستند والذي يعرف على أنه عدد الكلمات التي تسبق ظهور العبارة الأول مقسوماً على العدد الكلي للكلمات في المستند ويكون الناتج لها هو رقم بين الصفر والواحد.

بعد الانتهاء من توضيح الخطوات المشتركة لعملية التدريب واستخراج العبارات المفتاحية، سيتم توضيح عملية التعلم التي تتم في عملية التدريب فقط.

#### ✓ Learning :

تبدأ هذه المرحلة بحذف جميع الكلمات المكررة لمرة واحدة وذلك لتقليل حجم البيانات المستعملة في عملية التدريب.

تتم عملية التدريب من خلال الاستفادة من قيمتي Feature calculation التي تم حسابهما في المرحلة السابقة عن طريق نموذج نايف بايز الذي يتعلم تقسيم النتائج إلى قسمين.

القسم الأول هي العبارات المفتاحية، والقسم الثاني العبارات التي لا تصلح لأن تكون عبارات مفتاحية.

من المؤكد أن عملية التعلم تتم بالاستفادة من وجود نصوص تحوي عبارات مفتاحية تم استخراجها بشكل يدوي.

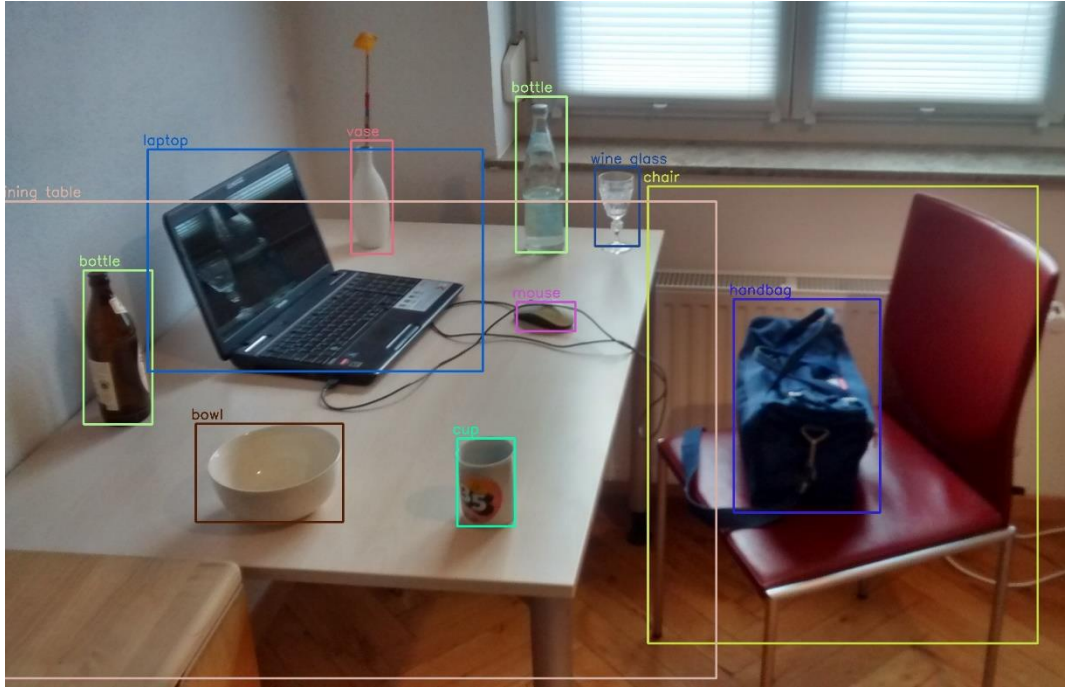
أما ما يميز عملية استخراج العبارات المفتاحية عن عملية التدريب هو الخطوة التالية:

#### ✓ Extraction of new keyphrases :

عند استخراج العبارات الأساسية من المستندات الجديدة، يتم حساب Feature calculation لكل عبارة مرشحة ويحسب احتمال أن تكون هذه العبارة هي عبارة مفتاحية، ثم يتم تحديد العبارات ذات أعلى احتمال في المجموعة النهائية من العبارات الرئيسية. كما ويمكن للمستخدم تحديد عدد العبارات الرئيسية التي يريد لها أن تظهر.

من عيوب هذه الخوارزمية أنها بحاجة إلى نصوص تم استخراج العبارات المفتاحية منها لكي تستطيع القيام بعملية التعلم، وهذا الأمر ليس من السهل القيام به في اللغة العربية خصوصاً مع عدم وجود بيانات جاهزة لعملية التعلم.

[13]



رسم توضيحي 26- مثال يوضح نتائج التعرف على الأغراض

### 3- توصيف الصور

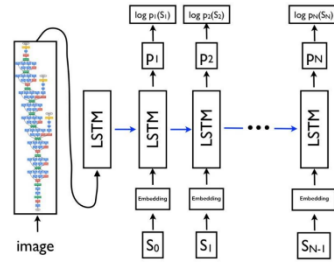
يعتبر توليد وصف لغوي للصور البصرية أحد أصعب المهام في مجال الرؤية الحاسوبية، لأنه يتطلب فهماً واسعاً للمحتوى البصري والعلاقات الدلالية بين الأغراض، كما يتطلب وجود نموذج لغوي قوي كي يولد جملاً مختلفة وصحيحة لغوياً.

كما إن لعملية توصيف الصور أهمية كبيرة من أجل عمليات البحث فمحركات البحث غير قادرة على البحث ضمن الصور وذلك لعدم فهمها محتوى الصور حيث أن معالجة الصور وفهم ما تحتويه أثناء البحث يحتاج إلى زمن أكبر بكثير مما تحتاجه عملية معالجة النصوص والبحث ضمنها.

ونتيجةً لذلك تنافس الباحثين وقاموا بإيجاد حلول مختلفة لهذه المشكلة حيث تم تقسيم الطرق إلى الأنواع التالية:

#### منهجيات من الأدنى إلى الأعلى

تعتمد هذه المنهجيات على تحليل الصورة لاكتشاف مجموعة كلمات تعبر عن أغراض أو علاقات ظاهرة في الصورة، ثم تستخدم هذه الكلمات لتوليد جملة متكاملة تصف الصورة عبر نموذج لغوي ما.



رسم توضيحي 27- البنية العامة لنموذج Show and tell

تعتمد معظم هذه المنهجيات تحويل الكلمات المستخرجة من الصورة إلى بيان وتحويل هذا البيان إلى جملة نصية عبر مجموعة من القواعد أو القوالب المبنية يدوياً، أو استخدام نماذج لغوية لتوليد الوصف بناءً على الكلمات المستنتجة مسبقاً.

بعض الطرق المستخدمة في هذه المنهجيات هي اكتشاف الأغراض من الصورة ثم استخدامها في قوالب جاهزة لتوليد الجمل الوصفية، كما فعل Farhadi et al [14] في عام 2010.

كانت هذه المنهجيات هي السائدة حتى عام 2015، حيث تفوقت عليها الطرق الأخرى، ومن سيئات هذه الطريقة قلة التنوع في الوصف الناتج وصعوبة تحديد قواعد كافية لأنواع المشاهد المختلفة.

## منهجيات من الأعلى إلى الأدنى

تعتبر هذه المنهجيات المشكلة مشابهة لمشكلة ترجمة لغوية، حيث تتم الترجمة من اللغة البصرية التي هي سمات مستخرجة من الصورة، عادةً عبر شبكات الطي العصبونية [Convolutional Neural Networks](#)، إلى لغة نصية وهي الوصف.

ظهرت بشكل أساسي في عام 2015 حيث طرح Vinyal et al [4] البحث المعروف بـ Show and Tell، يعتمد هذا النموذج على مبدأ المُرمز وفالك الترميز Encoder-Decoder، حيث يتم ترميز الصورة باستخدام مصنف صور GoogLeNet [15] مدرب على مجموعة بيانات ضخمة جداً مثل ImageNet [16]. تستخدم الطبقة قبل الأخيرة من هذا المصنف، مع إضافة طبقة Fully Connected يتم تدريبها، لاستخراج تمثيل شعاعي لملامح الصورة، حيث ساد استعمال الطبقات المبكرة من مصنفات الصور كأداة لاستخراج سمات للصورة في مسائل الرؤية الحاسوبية العديدة، ويمكن النظر إلى الرسم التوضيحي (27) لفهم بنية النموذج.

يتم توليد الوصف عبر نموذج لغوي هو عبارة عن شبكة عصبونية عودية [Recurrent Neural Network](#) من نوع [Long-Short Term Memory](#) [17] المعروفة بـ LSTM. من حسنات هذا النوع من الشبكات العودية أنه

يتجنب مشكلة الشبكات العصبونية العودية البسيطة المعروفة بـ Vanishing Gradient ، حيث يصعب تدريب ذلك النوع من الشبكات العصبونية بسبب ضعف عملية التعلم للأوزان نتيجة مشكلة التلاشي.

يتم إدخال سمات الصورة المستخرجة إلى النموذج اللغوي في اللحظة  $t=0$  ويقوم النموذج بتوليد كلمة ما، تدخل إليه مع الحالة الخفية Hidden State وحالة الخلية Cell State، التي ينتجها النموذج في اللحظة التالية  $t=1$ ، وهكذا حتى يصدر النموذج كلمة خاصة تعبر عن نهاية الوصف.

تم تمثيل الكلمات باستخدام نموذج تضمين الكلمات Word Embedding، الذي يحول شعاع الكلمة من شكل One-hot encoding حجمه بحجم قاموس الكلمات الكامل إلى شعاع بفضاء آخر من حجم أصغر، يساوي حجم فضاء شعاع ملامح الصورة. هذا النموذج يتم تعلمه خلال تعلم النظام الكامل. يلخص عمل النموذج بالعلاقات التالية:

$$X_{-1} = CNN(I)$$

$$X_t = W_e S_t : t \in \{0 \dots N - 1\}$$

$$p_{t+1} = LSTM(X_t) : t \in \{0 \dots N - 1\}$$

حيث:

I هي صورة الدخل

$CNN(I)$  هي سمات الصورة المستخرجة

$S_t$  هو تمثيل الكلمة الحقيقية في اللحظة  $t$  على شكل One-hot vector

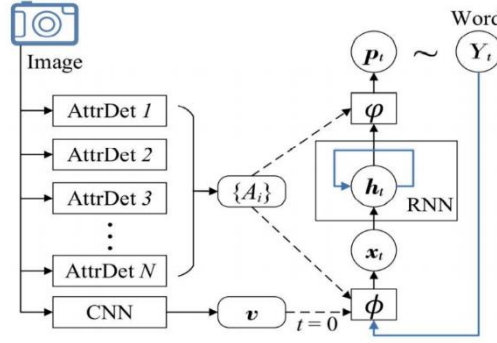
$W_e$  هو توزيع احتمالي للكلمة المتوقعة للحظة  $t+1$

فيما بعد، صدر بحث Show, Attend and Tell من Xu et al [4] الذي أضاف مبدأ الاهتمام الذي نشأ في مجال الترجمة من قبل Bahndau et al [18]، إلى مجال توليد وصف الصور. يعتمد هذا المبدأ على التركيز على أجزاء محددة من سمات الصورة في كل لحظة زمنية، كما يقوم الإنسان بالتركيز على مناطق محددة من الصورة أثناء وصفها.

يمكن لهذا النموذج استغلال معلومات بصرية أكثر لأنه يحافظ على سمات الصورة ويتم إدخالها (بعد معالجتها بنموذج الاهتمام) إلى نموذج اللغة في كل لحظة زمنية، مقارنة بالطريقة السابقة التي تستخدم شعاع سمات الصورة في اللحظة الزمنية الأولى فقط.

يختلف هذا النموذج عن النموذج السابق بعدة تفاصيل:

يستخدم هذا النموذج مصنف صور مختلف Oxford VGGnet [19] مدرب على ImageNet [16] بدون تعديل أوزان، ويستخرج السمات من طبقة الطي الرابعة في المصنف، ليحصل على مصفوفة سمات من الشكل



رسم توضيحي 28- البنية العامة لنموذج ATT

14x14x512 يقوم بتحويلها إلى الشكل 196 x 512، هي عبارة عن شعاع سمات طوله 512 لكل منطقة من 196 منطقة مختلفة في الصورة.

يتكون نموذج اللغة من LSTM [17] مشابه للنموذج السابق، مع تمرير شعاع سياق في كل لحظة، يعبر عن السمات بعد تطبيق آلية الاهتمام.

يتم حساب شعاع السياق من آلية الاهتمام التي قد تطبق بواحد من شكلين:

- Hard Attention: يحدد جزءاً واحداً من الصورة يجب الاهتمام بسماته في كل لحظة، مما يجعله غير مستمر وغير قابل للاشتقاق أثناء التعلم.
- Soft Attention: يحدد أهمية كل جزء من الصورة نسبياً مقارنة بباقي الأجزاء (بشكل توزيع احتمالي) فهو مستمر وقابل للاشتقاق وأسهل للتدريب.

يتم تعلم آلية الاهتمام، حيث تحدد الجزء أو الأجزاء التي تهتم بها في كل لحظة زمنية بناءً على الحالة الخفية Hidden State لنموذج اللغة في تلك اللحظة وسمات الصورة.

حقق هذا النموذج تفوقاً ملحوظاً على ما سبقه في المعايير المختلفة، بالرغم من أن رسم عمل آلية الاهتمام يوضح اهتمام النموذج بمناطق من الصورة أحياناً لا تعبر ولا ترتبط بالكلمات المولدة.

### منهجيات هجينة

تدمج هذه المنهجيات من النوعين السابقين، بحيث تستخرج مجموعة واصفات نصية من الصورة، وتقوم بعملية الترجمة اعتماداً على خليط بين هذه الوصفات وسمات مستخرجة أيضاً من الصورة لترجمتها إلى شكل نصي هو وصف الصورة.

أهم الأبحاث في هذه المنهجيات كان من قبل You et al [20] ، حيث تم تقديم نموذج يدعى ATT يشبه ما سبق في المنهجيات من الأعلى إلى الأدنى، لكنه يولد عدداً من الوصفات النصية عبر مصنف متعدد الأصناف Multilabel Classifier بالإضافة إلى استخراج سمات من الصورة، ويتم توليد الوصف اعتماداً على سمات الصورة فقط في اللحظة الأولى، ثم يستخدم آليات انتباه تحدد أهمية الوصفات النصية المستخرجة في كل

Name	BLEU 1	BLEU 2	BLEU 3	BLEU 4	METEOR
ATT	0.709	0.537	0.402	0.304	0.243
Show, Attend and Tell	0.718	0.504	0.357	0.250	0.230
Show and Tell	0.666	0.451	0.304	0.203	-

جدول 3: نتائج أبحاث مختلفة في مهمة وصف الصور

لحظة، التي تستخدم مع الحالة الخفية لنموذج اللغة لتوليد جملة الوصف، يمكن النظر للرسم التوضيحي (28) لفهم بنية ATT.

يوظف هذا النموذج آليتي اهتمام، بأوزان مختلفة:

- آلية اهتمام الدخل: تحدد أهم الوصفات المستخدمة كدخل لنموذج اللغة، اعتماداً على الكلمة السابقة والصفات المستخرجة. يتم استخدام نموذج تضمين الكلمات للحد من عدد المتحولات اللازمة بتصغير الأشعة المعبرة عن كل كلمة.
- آلية اهتمام الخرج: تستخدم لتعديل خرج نموذج اللغة بناءً على الحالة الخفية لنموذج اللغة في اللحظة الحالية والصفات المستخرجة، ونتيجتها هي توزيع احتمالي على جميع كلمات القاموس المستخدم، يحدد احتمال اختيار كل كلمة منه في هذه اللحظة.

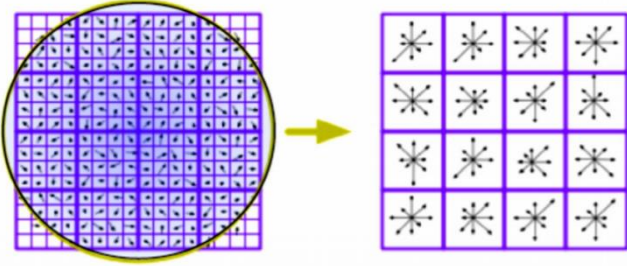
حيث: تعبر  $A_j$  عن قائمة الوصفات النصية المستخرجة، و  $v$  عن سمات الصورة الرقمية، و  $\phi$  عن آلية اهتمام الدخل، و  $\varphi$  عن آلية اهتمام الخرج، و  $P_t$  عن التوزيع الاحتمالي على قاموس الكلمات لخرج النموذج في اللحظة  $t$ ، و  $Y_t$  عن الكلمة المختارة كخرج في اللحظة  $t$ .

أثناء التعلم، يتم إضافة حدود إلى تابع الخطأ تشجع كلاً من آليتي الاهتمام على تعلم:

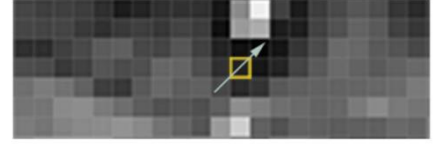
- الاهتمام بواصفات مختلفة من الصورة الواحدة عبر الزمن
- الاهتمام بعدد محدود من واصفات الصورة في اللحظة الزمنية الواحدة.

تفوق هذا النموذج على ما سبقه وحقق نتائج مبهره، ينسب هذا التفوق للقدرة على الدمج بين الوصفات العامة للصورة، مع التركيز على أجزاء مختلفة منها أثناء توليد الوصف، مع السمات المكانية للصورة المستخدمة في اللحظة الأولى فقط لإعطاء لمحة عامة عن سمات الصورة للنموذج اللغوي، بينما تواجه نماذج الأعلى للأدنى صعوبة في وصف التفاصيل الدقيقة في الصورة بسبب اعتمادها على سمات مكانية بدقة محددة.

نرى النتائج في الجدول (3) للنماذج السابقة على مجموعة بيانات MS-COCO [21]، بالاختبارات المتنوعة. تعد هذه الاختبارات معيارية في مجال الترجمة، حيث تقارن بين تشابه جملتين.



رسم توضيحي 29- image gradients to keypoint descriptor



رسم توضيحي 30- اتجاه سهم تحديد القتامة

#### 4- التعرف على الوجوه

##### dynamic way

يمكن مساعدة محركات البحث من خلال تحديد هوية الأشخاص الظاهرين في الصورة بدلاً من الاكتفاء بتوصيف وجود شخص ضمن الصورة من الأفضل أن يتم ذكر من هو هذا الشخص الظاهر ضمن الصورة خصوصاً في حال كان هذا الشخص شخصية مشهورة.

على الرغم أن الإنسان يستطيع بشكل بديهي تحديد هوية الأفراد من خلال التعرف على وجوه الأشخاص، لكن الأمر ليس بنفس السهولة بالنسبة للأنظمة الحاسوبية، فعملية التعرف على الوجوه تمر بعدة مراحل وهي كما يلي:

##### تحديد الوجوه الموجودة ضمن الصورة

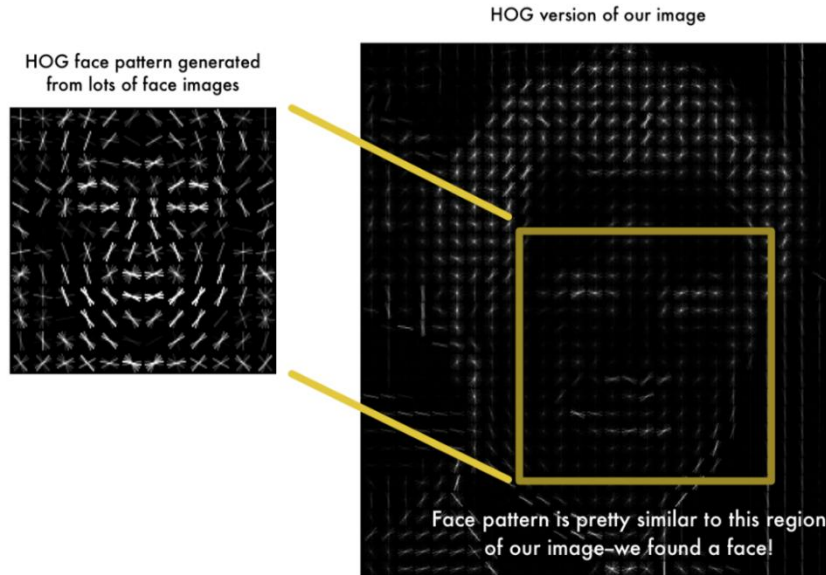
تعرف هذه المرحلة باسم face detection ويتم فيها إيجاد موقع كل الوجوه الموجودة ضمن الصورة. في البداية يتم تحويل الصورة من صورة ملونة إلى صورة رمادية، فلا يوجد حاجة للعمل على الصور الملونة ويمكن إيجاد الوجوه من خلال صورة رمادية.

يتم تحديد مدى ظلام النقطة (pixel) الحالية مقارنة بالنقط المحيطة بها مباشرة ثم يتم رسم سهم يوضح الاتجاه الذي تصبح فيه الصورة أكثر قتامة، وذلك من أجل كل نقطة ضمن الصورة وهذا ما هو موضح بالرسم التوضيحي (29). أي أن كل نقطة يتم استبدالها بسهم يمثل اتجاه القتامة وعندها ستكون الأسهم بنفس الاتجاه لصورتين تحويلان نفس الشخص إحداها قاتمة والأخرى فاتحة.





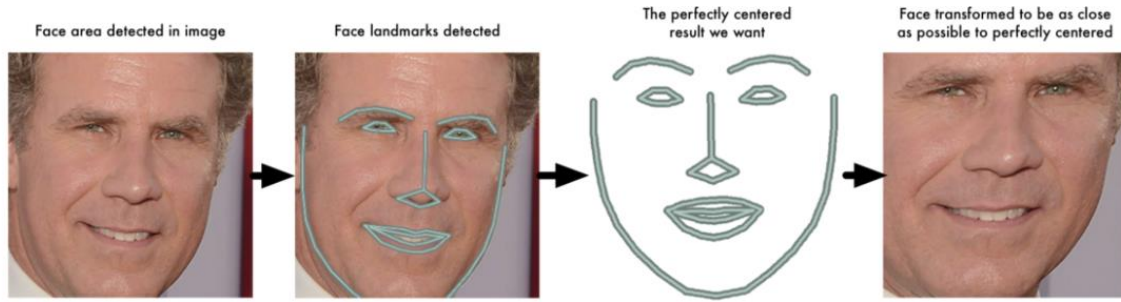
رسم توضيحي 31- تحويل صورة إلى فضاء HOG



رسم توضيحي 32- مقارنة بين صورتين في فضاء HOG

تظهر هنا مشكلة في أن عدد التفاصيل الموجود ضمن الصورة كبير وفائض عن الحاجة، فيتم تقسيم الصورة إلى  $16 * 16$  مربع ومن أجل كل مربع يتم إيجاد جميع الأسهم الناتجة عن النقاط الموجودة ضمنه ويتم استبدال كل مربع من هذه المربعات بسهم طوليته تتعلق بعدد تكرارات كل اتجاه للأسهم وهو ما يظهر في الرسم التوضيحي (31)، تعرف هذه العملية بـ HOG، ويكون النتيجة النهائية هي تمثيل بسيط من الصورة الأساسية وهو ما يظهر في الرسمين التوضيحين (32).

يتم تحديد مكان الوجه في صورة في فضاء HOG من خلال قياس التشابه بين أجزاء الصورة وصورة مرجعية معينة مسبقاً من نمط HOG تم استخراجها من خلال التدريب على مجموعة من الوجوه كما في الشكل التوضيحي (203). [22]



رسم توضيحي 33- تعديل وضع الوجه بالاستفادة من face landmark estimation

### تعديل وضع الوجه

يتم القيام بتعديل الصورة بحيث تكون العيون والذقن والشفاه دائماً في مكان معين من الصورة وذلك لتسهيل عملية المقارنة التي ستتم لاحقاً، وتتم هذه العملية من خلال خوارزمية تدعى face landmark estimation، حيث يتم استخراج المعالم الأساسية للوجه ويتم اجراء مجموعة من التحويلات على هذه الملامح (دوران - تغيير حجم - انزياح..). وهو ما يظهر في الرسم التوضيحي (33).

### ترميز الوجوه (استخراج السمات)

الطريقة الأسهل لإيجاد اسم الشخص الظاهر في الصورة هي مقارنة الصورة المراد إيجاد اسم الشخص فيها مع مجموعة من الصور التي تم حساب المرحلتين السابقتين لها والمعروف اسم الشخص الموجود ضمنها وفي حال كانت إحدى الصورتين متشابهتين عندها يمكن القول إن كلتا الصورتين لنفس الشخص، لكن في حال كانت كمية الصورة التي يتم المقارنة معها كبير عندها سيكون هنالك حاجة للكثير من الوقت لإجراء عملية المقارنة، لذلك سيتم استخراج السمات المهمة من الصورة وذلك اعتماداً على شبكة من نمط CNN وسيكون ناتج هذه العملية هو عمود بقياس 128.

### البحث عن اسم الشخص من الترميز

تعد هذه الخطوة الأسهل بين جميع الخطوات حيث تهدف لإيجاد الصورة الأكثر تشابهاً من بين مجموعة الصور المدخلة مسبقاً وبين الصورة التي يراد تحديد الشخص الموجود ضمنها، لذلك يتم تطبيق حساب المسافة الإقليدية لتحديد الصورة الأقرب.

تتميز هذه الطريقة بأنها سريعة جداً، ويمكن استعمالها في تطبيقات الزمن الحقيقي، كما أنه يمكن إضافة وجوه جديدة للتعرف عليها دون الحاجة للقيام بإعادة الترتيب.

## Static way

يتم تطبيق كل من تحديد الوجوه الموجودة ضمن الصورة وتعديل وضع الوجه بنفس الطريقة السابقة ويبدأ الاختلاف في مرحلة استخراج السمات فتكون وظيفة شبكة [CNN](#) ليس فقط استخراج السمات بل إيجاد احتمال ارتباط الشخص مع كل صف من صفوف البيانات التي تم التدريب عليها، وذلك باستخدام طبقة SoftMax في طبقة الخرج عدد العصبونات ضمنها بعدد الأشخاص الموجودين في البيانات، ولكن هذه الطريقة تعتبر غير فعالة لأنه ليس من المنطقي أن يتم إعادة التدريب على كامل البيانات عند كل إضافة لوجه جديد.

يوجد طريقة أخرى يتم فيها تطبيق كل من تحديد الوجوه الموجودة ضمن الصورة وتعديل وضع الوجه واستخراج السمات بنفس الطريقة السابقة، ويكون الاختلاف في البحث عن اسم الشخص من الترميز بدلاً من القيام بحساب المسافة الاقليدية يتم استعمال إحدى خوارزميات [التعلم التلقائي](#) المختصة بالقيام بعمليات التصنيف، فيتم تدريب الخوارزمية للقيام بفصل النقاط إلى مجموعات ولكن هذه الطريقة لتكون فعالة بحاجة إلى أكثر من وجه لكل شخص وذلك لمنع احتواء كل صف على وجه واحد.

## الفصل الثالث: الدراسة التحليلية

### المتطلبات الوظيفية

رقم المتطلب	اسم المتطلب	شرح المتطلب
01	استخراج العبارات المفتاحية	يتم معالجة نصوص باللغة العربية عن طريق إيجاد العبارات المفتاحية التي تعبر عن هذا النص وتساعد في عمليتي البحث والفهرسة عن طريق وضع العبارات المفتاحية ضمن واصفات صفحة ال HTML أو من خلال كتابتها بشكل علني ضمن الصفحة.
02	توصيف الصور	معظم محركات البحث لا تقوم بالبحث ضمن الصور، وإنما تكتفي بالبحث ضمن النصوص، ولذلك يتم استعمال توصيف للصور والتي توضح لمحركات البحث على محتوى الصورة، ويتم ذلك عن طريق إضافة توصيف أسفل الصورة، وإضافة هذا التوصيف ضمن واصفات ال HTML كنص بديل في حال فشل تحميل الصورة.
03	إنشاء حساب جديد	يجب على المستخدم امتلاك حساب لكي يقوم بالاطلاع على محتوى الموجود ضمن هذا الحساب.
04	تسجيل دخول	لا يمكن للمستخدم إدارة المحتوى دون القيام بعملية تسجيل الدخول.
05	كتابة منشور جديد	يمكن للمستخدم كتابة منشور جديد.
06	التحكم بوقت النشر	من الممكن أن يكون المنشور بحاجة إلى مدة طويلة للكتابة لذلك تم السماح للمستخدم بكتابة منشور ووضعها في المسودات ويستطيع إكمال الكتابة في الوقت الذي يريده.

07	تعديل منشور	في حال أراد المستخدم تعديل منشور في وضع الجدولة أو منشور تم نشره سابقاً يمكنه القيام بذلك.
08	تعديل صلاحيات المسؤولين	يمكن للمستخدم المالك إضافة محررين أو مالكين آخرين أو أعضاء بصلاحيات محددة.
09	التعرف على الوجوه	يمكن للمستخدم التعرف على وجوه الأشخاص الموجودين ضمن الصورة في حال كانت الشخصية مشهورة ومعروفة.
10	إضافة وجه جديد للتعرف عليه	يمكن للمستخدم في حال كان لديه شخص تتكرر صورته في الكثير من الصور أن يقوم بإضافة صورة لهذا الشخص واعتباره شخصاً معروفاً، ويصبح من الممكن التعرف عليه بشكل تلقائي.
11	خدمة قراءة النصوص	لتوفير الوقت على المستخدم في حال كان يقوم بعملية إعادة كتابة المنشور من صورة يمكن له الاستفادة من هذه الخدمة والتي تقوم باستخراج النص الموجود باللغة العربية ضمن صورة.
12	البحث باللغتين	يتم البحث عن الجمل المطابقة الموجودة في المنشورات السابقة وذلك سواء أراد المستخدم البحث عن جملة باللغة الإنكليزية أو عن جملة باللغة العربية.
13	المصحح التلقائي للأخطاء	تكثر الأخطاء الإملائية وأخطاء الترقيم عند المحررين لذلك كان لابد من وجود خدمة المصحح التلقائي للأخطاء والتي ستشير لمكان الخطأ وتمنح المستخدم اقتراح للتصويب.

جدول 4: المتطلبات الوظيفية

## المتطلبات الغير وظيفية

- زمن الاستجابة:

يجب أن تُنفذ كافة الخدمات خلال زمن مقبول، وذلك لتحقيق الغاية منها وهي توفير الوقت على المستخدمين.

- قابلية إعادة الاستخدام:

إن أي نظام يمكن أن يعاد استخدامه في مكان آخر بجهد قليل، هو نظام فعال يوفر جزءاً من الكلفة والزمن في حال استخدامه كجزء من نظام آخر. ويتم ذلك من خلال توفير الخدمات بشكل منفصل عن بعضها من خلال إتاحة استعمال أجزاء منها.

- إمكانية التوسيع Scalability:

إمكانية إضافة ميزات جديدة في المستقبل لتحسين النظام، وذلك بفضل اعتماد مبدأ ال micro services في جميع جزئيات النظام كافة، وتوسيع الخدمات الحالية بإضافة functionality جديد إليها، بحيث تكون أكثر تخصصية .

- قابلية التعميم Generalizability:

القدرة على التعامل مع أنواع النصوص المختلفة (أخبار- رياضة...)، وكذلك التعامل مع الصور بإضاءات مختلفة.

- الحماية Security:

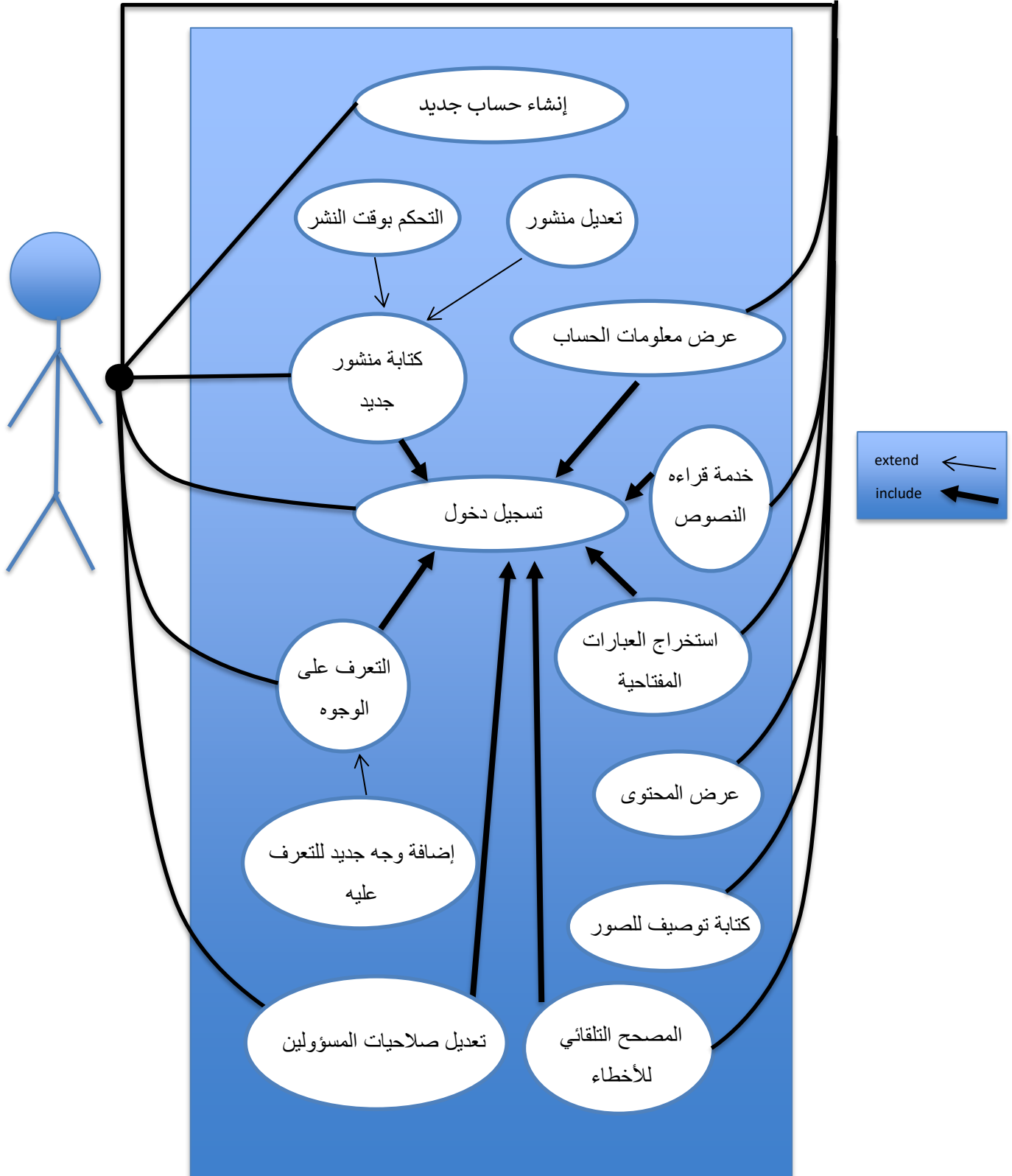
من أهم المتطلبات حماية البيانات بشكل جيد، والحفاظ على سرية المستخدمين، والتأكد من عدم وصولها إلا من قبل الأشخاص المعنيين.

## فئات المستخدمين

الشريحة المستهدفة هي صناع المحتوى باللغة العربية الذين يرغبون بإدارة محتوى المواقع بطريقة سهلة وعملية توفر لهم الكثير من الوقت والجهد.

## مخطط حالات الاستخدام

حالات الاستخدام ذات التوصيف العام



رسم توضيحي 34- حالات الاستخدام ذات التوصيف العام

## حالات الاستخدام ذات التوصيف المفصل

### استخراج العبارات المفتاحية

اسم حالة الاستخدام	استخراج العبارات المفتاحية
الرقم التعريفي للحالة	01
الممثل الرئيسي	مدير المحتوى
الممثل الثانوي	_____
الوصف	يتم إيجاد العبارات المفتاحية التي تعبر عن النص المراد معالجته وعرضها على المستخدم للاختيار بينها.
الشروط المسبقة	1- إدخال المستخدم للنص بشكل صحيح. 2- تحديد عدد العبارات المفتاحية التي يريد المستخدم ظهورها.
تدفق الأحداث	<ul style="list-style-type: none"> <li>يقوم المستخدم بإدخال النص المراد استخراج العبارات المفتاحية له.</li> <li>يتم القيام بعمليات تنظيف للنص.</li> <li>يتم استخراج العدد المطلوب من العبارات المفتاحية.</li> <li>يتم إظهار العبارات المفتاحية مرتبين حسب احتمال كل منها ومع امكانية التعديل عليها أو حذفها.</li> </ul>
السيناريوهات البديلة	_____

جدول 5: حالة استخدام استخراج العبارات المفتاحية



توصيف الصور

اسم حالة الاستخدام	توصيف الصور
الرقم التعريفي للحالة	02
الممثل الرئيسي	مدير المحتوى
الممثل الثانوي	_____
الوصف	يتم توليد جملة توصف الصورة المدخلة.
الشروط المسبقة	1- إدخال المستخدم للصورة بالشكل الصحيح. 2- اقتصاص الصورة أو التعديل عليها لتأخذ الحجم المطلوب.
تدفق الأحداث	<ul style="list-style-type: none"> <li>• يقوم المستخدم برفع الصورة.</li> <li>• يتم القيام بعمليات المعالجة المطلوبة للصورة.</li> <li>• استخراج توصيف مناسب للصورة.</li> <li>• إظهار التوصيف المناسب للصورة.</li> </ul>
السيناريوهات البديلة	_____

جدول 6: حالة استخدام توصيف الصور

تعديل صلاحيات المسؤولين

اسم حالة الاستخدام	تعديل صلاحيات المسؤولين
الرقم التعريفي للحالة	08
الممثل الرئيسي	مدير المحتوى
الممثل الثانوي	_____
الوصف	يتم إضافة مسؤولين جدد وإعطائهم صلاحيات محددة من قبل المالك.
الشروط المسبقة	<p>1- من يقوم بعملية التعديل يملك صلاحيات تعديل للمسؤولين.</p> <p>2- أن يملك الشخص المراد تعديل الصلاحية له حساباً على الموقع.</p>
تدفق الأحداث	<ul style="list-style-type: none"> <li>يقوم الشخص الذي يملك صلاحيات تعديل صلاحيات الأشخاص باختيار الأشخاص الذين يريد تعديل الصلاحيات لهم من خلال إدخال اسم أولئك الأشخاص بشكل مطابق لاسمهم في الموقع.</li> <li>يتم تحديد الصلاحيات المراد منحها لهؤلاء الأشخاص.</li> </ul>
السيناريوهات البديلة	<ul style="list-style-type: none"> <li>يقوم الشخص الذي يملك صلاحيات تعديل صلاحيات الأشخاص باختيار الأشخاص الذين يريد تعديل الصلاحيات لهم من خلال إدخال اسم أولئك الأشخاص بشكل مطابق لاسمهم في الموقع.</li> <li>في حال لم يتم العثور على الشخص الذي يراد منحه صلاحيات جديدة، عندها تظهر رسالة تخبر المسؤول عن أن الشخص غير موجود أو أن الاسم المدخل خاطئ.</li> </ul>

جدول 7: حالة استخدام تعديل صلاحيات المسؤولين

## التعرف على الوجوه

اسم حالة الاستخدام	التعرف على الوجوه
الرقم التعريفي للحالة	09
الممثل الرئيسي	مدير المحتوى
الممثل الثانوي	_____
الوصف	يمكن للمستخدم التعرف على وجوه الأشخاص الموجودين ضمن الصورة في حال كانت الشخصية مشهورة ومعروفة.
الشروط المسبقة	1- الشخص المراد التعرف عليه يكون شخصية مشهورة. 2- تحتوي الصورة على وجه واضح.
تدفق الأحداث	<ul style="list-style-type: none"> <li>• يقوم المستخدم برفع الصورة.</li> <li>• يتم القيام بعمليات المعالجة المطلوبة للصورة.</li> <li>• التعرف على كل الأشخاص الموجودين في الصورة.</li> <li>• إظهار النتيجة.</li> </ul>
السيناريوهات البديلة	<ul style="list-style-type: none"> <li>• يقوم المستخدم برفع الصورة.</li> <li>• يتم القيام بعمليات المعالجة المطلوبة للصورة.</li> <li>• لم يتم التعرف على أي شخص ضمن الصورة.</li> <li>• إظهار رسالة أن الشخص غير معروف.</li> </ul>

جدول 8: حالة استخدام التعرف على الوجوه

إضافة وجه جديد

اسم حالة الاستخدام	إضافة وجه جديد
الرقم التعريفي للحالة	10
الممثل الرئيسي	مدير المحتوى
الممثل الثانوي	_____
الوصف	يمكن للمستخدم في حال كان لديه شخص تتكرر صورته في الكثير من الصور أن يقوم بإضافة صورة لهذا الشخص واعتباره شخصاً معروفاً، ويصبح من الممكن التعرف عليه بشكل تلقائي.
الشروط المسبقة	وضوح الوجه المراد إضافته.
تدفق الأحداث	<ul style="list-style-type: none"> <li>يقوم المستخدم برفع الصورة.</li> <li>يتم إضافة اسم الشخص.</li> <li>يتم القيام بعمليات المعالجة المطلوبة للصورة.</li> <li>إضافة الوجه إلى البيانات المحفوظة مسبقاً.</li> </ul>
السيناريوهات البديلة	<ul style="list-style-type: none"> <li>يقوم المستخدم برفع الصورة.</li> <li>يتم القيام بعمليات المعالجة المطلوبة للصورة.</li> <li>يتم إضافة اسم الشخص.</li> <li>في حال كان الاسم موجد مسبقاً يتم إعلام المستخدم أن المستخدم موجود مسبقاً.</li> </ul>

جدول 9: حالة استخدام إضافة وجه جديد

المصحح التلقائي للأخطاء

المصحح التلقائي للأخطاء	اسم حالة الاستخدام
12	الرقم التعريفي للحالة
مدير المحتوى	الممثل الرئيسي
_____	الممثل الثانوي
يتم تحديد الأخطاء الموجودة في النص وعرض التصحيح المناسب لها.	الوصف
_____	الشروط المسبقة
<ul style="list-style-type: none"> <li>• يقوم المستخدم بإدخال النص.</li> <li>• يتم القيام بعملية تحويل النص إلى كلمات من خلال التلخيص من إشارات الترقيم.</li> <li>• يتم تحديد الكلمات التي تحوي على خطأ.</li> <li>• إظهار الكلمات الخاطئة مع التصحيح المناسب لها.</li> </ul>	تدفق الأحداث
_____	السيناريوهات البديلة

جدول 10: حالة استخدام المصحح التلقائي للخطأ

## توصيف واجهة الاستخدام

تم اعتماد واجهات الاستخدام المصممة من قبل Strapi والتي تم تصميمها لإدارة المحتوى، وتتمتع هذه الواجهات بسهولة التعديل في حال أراد المصمم إجراء تعديلات عليها، وكذلك تتميز الواجهات بالوضوح وسهولة الاستخدام، بالإضافة إلى السرعة في الأداء، ويمكن مشاهدة الواجهات من خلال زيارة الموقع الرسمي لـ <https://strapi.io>.

## الأنساق المعمارية وأنساق التصميم المستخدمة

ستكون server/client حيث أنه سيكون هنالك العديد من المستخدمين الذين يستفيدون من الخدمات في آن واحد.

سيتم تقسيم التطبيق الى أربع أقسام رئيسية:

### 1- Models :

يتم جمع ال models معاً في قسم واحد، والتي تتواصل مع المخدم وتقوم بتنفيذ المهمات التي يطلبها المخدم.

### 2- View :

الواجهات التي سيقوم المستخدم بالعمل عليها والتي تتواصل مع المخدم وتقوم بعرض النتائج.

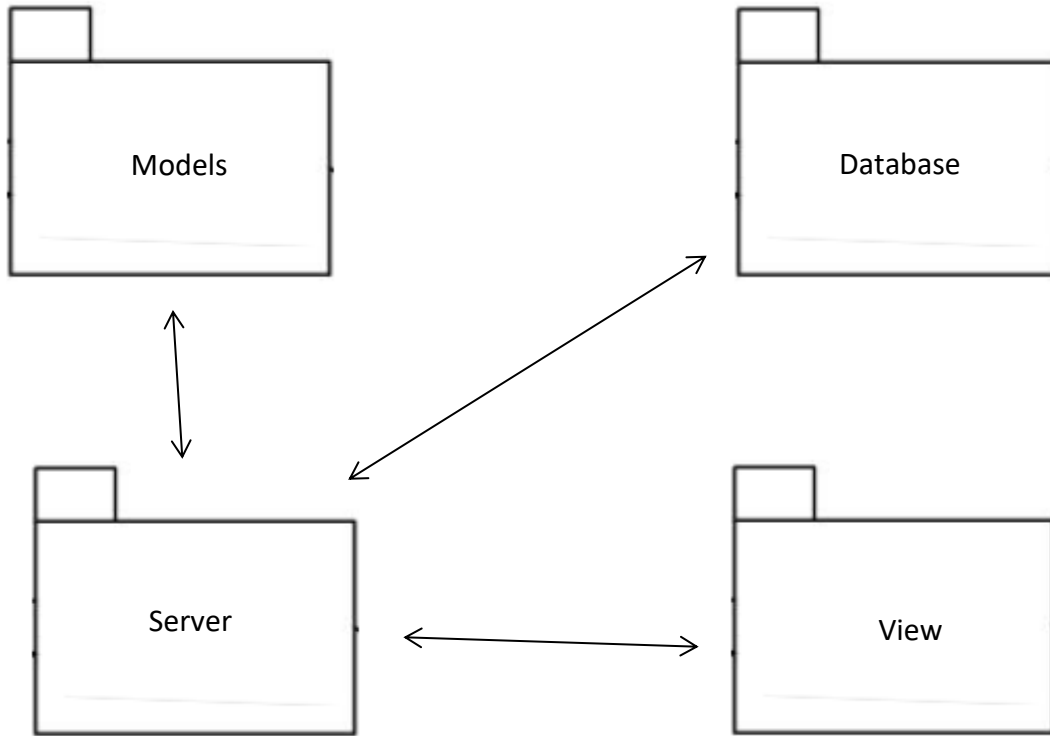
### 3- Database:

يجب أن تخزن البيانات في قسم منفصل عن باقي الأقسام لذلك تخزن في قاعدة بيانات مما يسمح بوصول سريع وعملي للبيانات.

### 4- Server:

يعد صلة الوصل بين جميع الأقسام وهو المسؤول عن تحديد الأوامر التي يجب أن تنفذ في باقي الأقسام.

## الأنظمة الجزئية المكونة



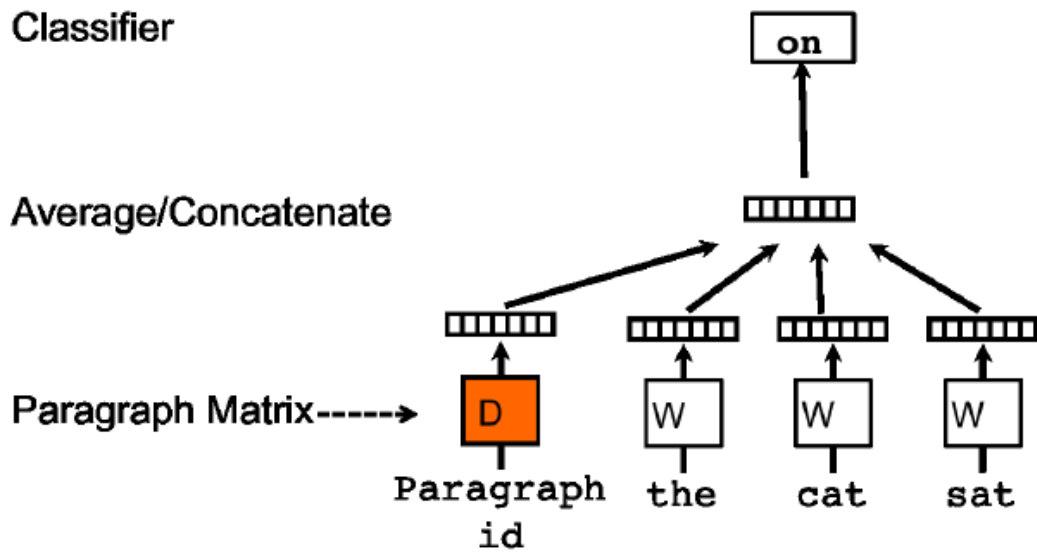
## توزيع الأنظمة الجزئية على العتاد

يتم عمل التطبيق على أكثر من جهاز، حيث يتم تجسيد التطبيق من خلال server يقوم بحفظ ومعالجة جميع الطلبات التي تأتي من مجموعة ال client وبالتالي يكون السيرفر هو عقدة الوصل بين هذه ال client من خلال .REST Api

## تخزين المعطيات الدائم

تحتاج أنظمة إدارة المحتوى لوجود محتوى يتم إدارته أي أنّ المحتوى يجب أن يكون مخزناً ضمن النظام لتتم إدارته، لذلك سيتم استعمال قواعد المعطيات للقيام بعملية التخزين وإجراء الاستعلام عليها، ولقد تم اعتماد SQLite والتي تتميز بما يلي:

- لا توجد مشاكل بالترميز لا سيما مع اللغة العربية.
- صغر الحجم.
- سهولة نقل البيانات من مزود إلى آخر.
- سهولة التركيب.
- لا تحتاج أية إعدادات تركيب خاصة على المزود.
- تدعم حجم قاعدة البيانات إلى 2 تيرابايت.



رسم توضيحي 35- بنية نموذج PV-DM

## الفصل الرابع: الدراسة التصميمية

### الاستخراج التلقائي للعبارات المفتاحية

يتم القيام بالتعرف على الكيانات الموجودة ضمن النص وذلك من خلال استخدام مكتبة [Stanza](#) والتي تتميز بالسرعة والدقة.

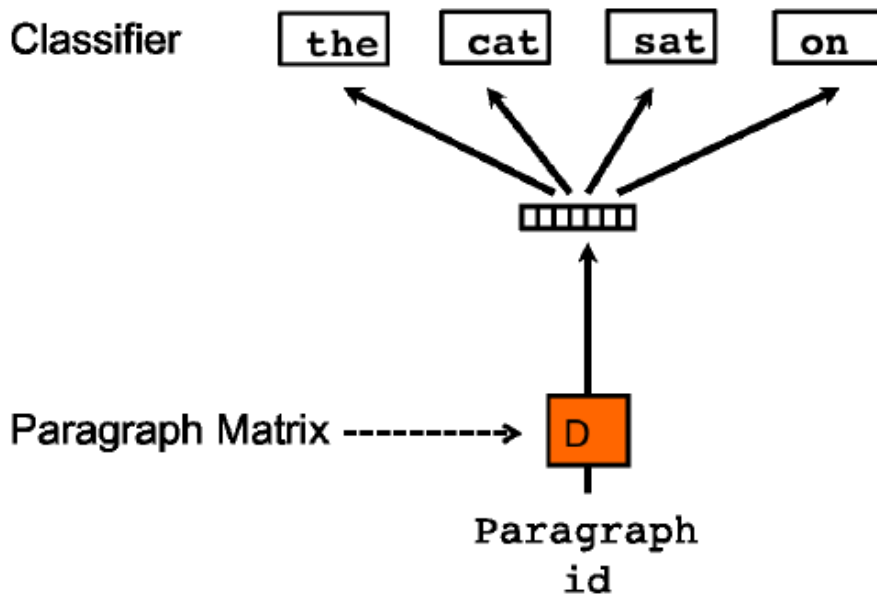
يتم في المرحلة الثانية القيام بعملية Embedding للكيانات التي تم استخراجها بالخطوة السابقة وذلك من أجل تحويل النص والعبارات المفتاحية إلى أشعة، ويمكن تطبيق عملية ال Embedding بطريقتين، وقد تم تنفيذ كل منهما:

#### 1. Doc2vec

هي خوارزمية تقوم بتحويل أي نص (أيًا كان طول النص) إلى شعاع في فضاء مستمر، ومن أجل شرح هذه الخوارزمية يجب الاطلاع على خوارزمية [Word2Vec](#) لأنها تعتمد بشكل كبير عليها. [23]

وبشكل مشابه جدا لخوارزمية [word2vec](#) يوجد تقنيتين في ال Doc2vec:





رسم توضيحي 36- بنية نموذج PV-DBOW

### PV-DM

في هذه التقنية يتم استخدام نموذج [CBOW](#) ويتم إضافة token من أجل كل نص في بيانات التدريب، ويتم إدخال ال token الخاص بكل نص للنموذج أثناء التدريب على كلمات هذا النص من أجل تعلم الشعاع d المعبر عن هذا النص، وهو ما يمكن ملاحظته في الرسم التوضيحي (35).

### PV-DBOW

هذا النموذج مشابه جدا لنموذج [skip-gram](#) والاختلاف الوحيد فقط في الدخول حيث يكون الدخول هنا token النص الذي يتم التدريب عليه، وهو ما يمكن ملاحظته في الرسم التوضيحي (36).

تم تدريب هذا النموذج على نصوص ويكيبيديا.

## 2. Bert model

يتم الترميز في Bert باستعمال عدّة طبقات، حيث أنّ الطبقة الأولى تقوم بتحويل كل كلمة في النص إلى شعاع بطول 768، أما الطبقة الثانية فهي مسؤولة عن التمييز بين نصين وقياس التشابه والاختلاف بينهما، في حين أن الطبقة الأخيرة تعطي ترميزاً للكلمة يتعلق بموضع الكلمة ضمن الجملة، وعندها يكون الخرج

الناتج بقياس  $512 \times 768$ ، ولكن للقيام بعلمية المقارنة بين كلمة ونص يجب تحويل النص إلى شعاع بطول 768 ويتم ذلك عن طريق أخذ المتوسط لل 512 قيمة فيكون الشعاع الناتج بطول 768، وهذا ما يسمح بقياس التشابه بين نص وكلمة.

بعد الانتهاء من تنفيذ عملية ال Embedding يتم قياس التشابه بين شعاع النص وشعاع العبارات المفتاحية واختيار الأشعة الأكثر تشابهاً وذلك بطريقة cosine similarity.

تم تطبيق طرق [Graph based Methods](#) لاستخراج العبارات المفتاحية وذلك من خلال استخدام مكتبة [PKE](#) والتي تعتمد في عملها على مكتبة spacy، لكن هذه المكتبة لا تدعم اللغة العربية لذلك قمنا بتعديل الشيفرة البرمجية الخاصة بهذه المكتبة فتم استبدال الدوال التي تعتمد على مكتبة spacy والتي لا تدعم اللغة العربية بدوال أخرى تدعم اللغة العربية بالاستعانة بمكتبة al camels tools المتخصصة للغة العربية وبهذا نكون قد حققنا المهمة المطلوبة من خلال إجراء عملية تكامل بين al camels tools و [PKE](#) لدعم اللغة العربية.

## التعرف على الوجوه

بناءً على متطلبات المسألة والحاجة إلى إعطاء صلاحية للمستخدم لإضافة شخص جديد للتعرف عليه، وكذلك الحاجة إلى الاكتفاء بصورة واحدة لكل شخص، تم اعتماد طريقة [dynamic](#) للقيام بهذه المهمة.

وقد تم استعمال أداة [Dlib](#) والتي وفرت الكثير من الجهد لتحديد الوجوه، تقوم المكتبة [بتحديد أماكن الوجوه الموجودة ضمن الصورة](#) باستعمال طريقة HOG.

وباستعمال نفس المكتبة يتم [تعديل وضع الوجه](#) في الصورة.

وعند الانتهاء من هذه المرحلة يجب الانتقال لمرحلة [ترميز الوجوه](#) ويتم القيام بهذه المهمة باستعمال شبكة من نمط [CNN](#) مدربة على استخراج السمات المهمة من الوجوه، وقد تم استعمال مكتبة [Dlib](#) التي توفر شبكة ال [CNN](#) المطلوبة والمدربة بشكل مسبق لذلك لا فائدة من القيام بإعادة تدريب هذه الشبكة، ويكون خرج هذه المرحلة هو شعاع بطول 128، ومن الجدير بالذكر أن مكتبة [Dlib](#) من أفضل الطرق المستعملة للتعرف على الوجوه فقد حققت أداءً ممتازاً بلغ 98.37%.

وبعد الانتهاء من [ترميز الوجوه](#) يتم الانتقال للخطوة الأخير وهي [البحث عن اسم الشخص من خلال الترميز](#) حيث تهدف لإيجاد الصورة الأكثر تشابهاً من بين مجموعة الصور المدخلة مسبقاً وبين الصورة التي يراد تحديد الشخص الموجود ضمنها، لذلك يتم تطبيق حساب الفرق بين الشعاعين واختيار الشعاع الأكثر تشابهاً من بين الأشعة الناتجة

عن صور تم إدخالها مسبقاً وبين الشعاع الناتج عن الصورة التي يتم البحث عن الشخص الموجود ضمنها، مع وجود عتبة لقبول أن الصورتين لنفس الشخص أو لشخصين مختلفين.

تتطلب المسألة إدخال مجموعة من الوجوه للقيام بعملية مقارنة معها، لذلك تم اختيار بيانات تحوي وجوه أهم الشخصيات في العالم والمقدمة من جامعة سيول في العام 2019 وتدعى باسم [Labeled Faces in the Wild](#) وتحوي هذه البيانات على صورة لـ 5749 شخصاً مختلفاً من أهم الشخصيات في العالم، ولكن وعلى الرغم من العدد الكبير للشخصيات المتواجدة إلا أن هذه الشخصيات أغلبها لأشخاص ليسوا من العرب لذلك تم إضافة حوالي 100 شخص بشكل يدوي.

كما وتم السماح للمستخدمين بالقيام بإضافة وجوه جديدة حيث سيتم استخراج شعاع السمات للصورة المدخلة وإضافتها للبيانات الموجودة مسبقاً مع التأكد من أنّ الشخص المضاف غير موجود ضمن الوجوه المعروفة مسبقاً.

## توصيف الصور

تم اعتماد بُنية مشابهة لبُنية منهجية الأعلى للأدنى ضمن بحث Show, Attend and Tell [4]، حيث يعتمد النموذج بنية encoder-decoder، فيتم ترميز سمات الصورة بشعاع أو مصفوفة في حالة الاهتمام باستخدام مُرمز ما، ويتم استخدام فاكك ترميز لتحويل هذه السمات إلى كلمات تصف محتويات الصورة.

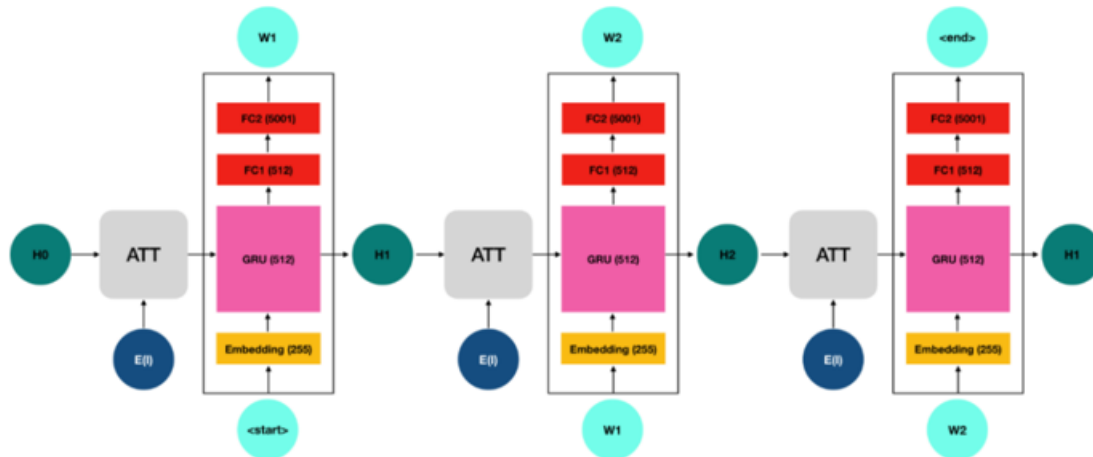
### المُرمز Encoder:

تم استخدام جزء من نموذج تصنيف الصور InceptionV3 مُدرَّب على مجموعة بيانات كبيرة وهي ImageNet [16]، لاستخراج السمات من الصورة، حيث تم استخدام جميع الطبقات عدا الأخيرة من نموذج التصنيف بأوزانها المتعلمة على مجموعة البيانات تلك، لأن الطبقة قبل الأخيرة تُعطي شعاع يمثل سمات كامل الصورة.

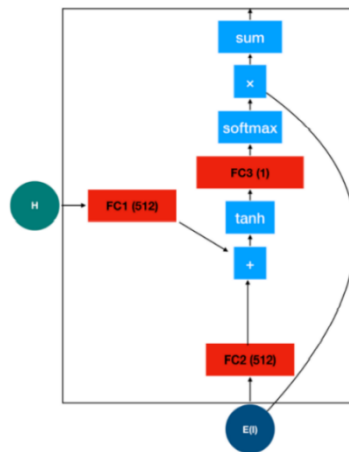
### فاكك الترميز Decoder:

دخل فاكك الترميز يكون عبارة عن الكلمة السابقة (عند بداية توليد الوصف تكون <start>) وسمات الأجزاء المختلفة من الصورة، والحالة الخفية السابقة (عند بداية توليد الوصف تكون الحالة الخفية السابقة شعاع من الأصفار بالحجم المناسب).

يتم في هذا النموذج أيضاً تمثيل جميع الكلمات على شكل رقم يدل على موقعها في قاموس الكلمات الكلي. يعتمد هذا النموذج على آلية الاهتمام، التي هي عبارة عن نموذج، دخله هو سمات الأجزاء المختلفة من الصورة،



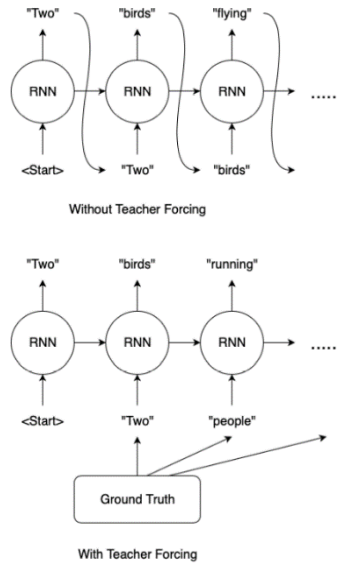
رسم توضيحي 38- بنية فالك الترميز. حيث  $H$  تعبر عن الحالة الخفية،  $W$  تعبر عن الكلمة المولدة في كل لحظة،  $E(i)$  تعبر عن مصفوفة سمات أجزاء الصورة (ناتج تطبيق المرمز  $E$  على الصورة  $I$ ) و  $ATT$  تمثل بنية الاهتمام.



رسم توضيحي 37- بنية آلية الاهتمام، حيث  $H$  تعبر عن الحالة الخفية الناتجة عن فالك الترميز في اللحظة السابقة،  $E(i)$  تعبر عن مصفوفة سمات أجزاء الصورة (ناتج تطبيق المرمز  $E$  على الصورة  $I$ )، وهي تتكرر ذاتها في كل لحظة. الطبقات  $FC1$ ،  $FC2$ ،  $FC3$  لها أوزان يتم تعلمها أثناء التدريب

والحالة الخفية السابقة، وخرجه عبارة عن شعاع سمات جديد يعبر عن سمات من الصورة مُثَقَّلَة بأهميتها في اللحظة الحالية، بناءً على سياق الجملة المولدة حتى الآن (المتمثل بالحالة الخفية السابقة) وسمات الأجزاء، يمكن النظر للرسم التوضيحي (37) لفهم أكبر لبنية فالك الترميز.

تتكون آلية الاهتمام من طبقة Fully Connected خاصة بسمات أجزاء الصورة، وطبقة Fully Connected خاصة بالحالة الخفية السابقة، يتم جمع خرج كل منهما كطريقة لدمج ناتجيهما، ويتم تطبيق طبقة Fully Connected أخرى على الناتج تتكون من عصبون واحد، حجم خرجها يكون شعاع بعده يساوي عدد أجزاء الصورة التي نملك لها سمات، تابع التفعيل لهذه الطبقة هو تابع SoftMax يعطي توزيعاً احتمالياً على أجزاء الصورة، يعبر هذا التوزيع عن احتمال أهمية كل جزء من هذه الأجزاء في هذه اللحظة، ويتم استخدامه لتثقيل



رسم توضيحي 39- آلية عمل Teacher Forcing

هذه الأجزاء، كل جزء حسب احتمال أهميته في اللحظة الحالية. يتم جمع أشعة سمات الأجزاء المختلفة بعد تثقيفها، لنحصل على شعاع واحد يكون الغالب فيه عادةً سمات عدد محدود جداً من أجزاء الصورة، تمثل الأجزاء التي تهتم فيها الآلية في هذه اللحظة، يمكن النظر للرسم التوضيحي (38) لفهم بنية آلية الاهتمام.

نلاحظ أن في هذا النموذج، سمات أجزاء الصورة تدخل لفاكك الترميز عند توليد كل كلمة، وذلك لأن آلية الاهتمام ستتخلص من السمات غير المفيدة منها بناءً على سياق الجملة (الحالة الخفية) بإعطائها أهمية منخفضة، مما يتخلص من مشاكل الضجيج وOverfitting.

يتم استخدام نموذج تضمين لتمثيل الكلمة السابقة في الفضاء الموجودة فيه السمات، ويتم استخدام GRU، ويمكن اعتبار آلية الاهتمام إضافة تُغيّر من شعاع الحالة الخفية، ونستخدم الكلمة السابقة مع خرج آلية الاهتمام كدخل لفاكك ترميز مطابق للمستخدم في الحالة السابقة.

تم تدريب النموذج باستخدام Adam Optimizer وتابع الخطأ SparseCategoricalCrossentropy.

أثناء التدريب، بدلاً من استخدام الكلمات المولدة من النموذج لتوقع الكلمات التالية يتم استخدام الكلمات الموجودة في جملة الوصف الحقيقية الموجودة في مجموعة البيانات المعتمدة. تسمى هذه التقنية Teacher Forcing وتساعد النموذج على التعلم بسرعة حيث تقلل احتمال المتابعة بجملة خاطئة كاملة بسبب توقع كلمة خاطئة واحدة. مثال وبالنظر للرسم التوضيحي (39). إذا كانت الجملة الواصفة

”Two people reading a book“ وقام النموذج بتوليد كلمة two ، ثم كلمة birds ، من المتوقع من النموذج أن يتابع بالخطأ وتوقع شيء مثل Flying باعتبار أن الكلمة التالية للطير من المتوقع أن تكون تطير في عدد كبير من الأمثلة. في حال استخدام الكلمة الثانية من الجملة الواصفة الحقيقية، يتم التخلص من هذا الخطأ وزيادة احتمال توليد كلمة تالية مناسبة للأشخاص، وليس للطير.

ومن المشاكل الملاحظة أثناء عملية التدريب والاختبار هي أن اللغة العربية تعتبر من اللغات الغنية بالكلمات وبالتالي عند التدريب والقيام بعملية Tokenization ، ومن أجل الأداء الأفضل، نقوم عادةً بالاحتفاظ بعدد محدود من الكلمات الأكثر تكراراً top\_k vocab، فعند تجربة الاحتفاظ بـ 5000 كلمة فقط، كانت النتائج التوصيف تحوي على <UNK>، وهي كلمة تظهر عندما يكون الخرج عبارة عن كلمة غير موجودة ضمن قاموس الكلمات المستخدم Vocab، لذلك قمنا برفع هذا العدد إلى 10000، وبهذا نكون قد قمنا بضم عدد أكبر بكثير من كلمات اللغة العربية الأكثر تكراراً.

## التطبيق

تم بناء النظام بالاعتماد على منصة Strapi مفتوحة المصدر نظراً لشهرته الواسعة من حيث المجتمع الإلكتروني الكبير له Community، وبسبب القدرة على تطوير الإضافات له بشكل سهل، حيث يتوفر توثيق متكامل له Documentation، يُمكن المطورين من تطوير الإضافات المختلفة، وبالإضافة إلى أنه من النوع Headless CMS أي يمكن المستخدم من إدارة محتواه الإلكتروني على جميع المنصات ولا يقتصر فقط على تطبيقات الويب.

## قراءه النصوص OCR

قد يحتاج مدير المحتوى لنسخ النص الموجود في صورة، وهذا الأمر يستغرق وقتاً طويلاً خصوصاً لو أنّ العملية مستمرة بشكل متكرر، لذلك وبهدف توفير الوقت المهدور على المستخدم وتوفير كافة احتياجاته ضمن النظام تم إضافة خدمة قراءه النصوص من صورة وتحويل الكتابة الموجودة ضمن الصورة إلى نص يمكن تعديله من قبل المستخدم، وقد تم دعم كل من اللغتين العربية والإنكليزية وذلك لتحقيق متانة أكبر ضمن النظام، ونظراً لأنّ هذه الخدمة من المواضيع البحثية التي لاقت اهتمام مطورين كثر حيث شهدت مؤخراً تحديثات وتحسينات كثيرة جعلتها تنتشر على نطاق واسع وباتت تتوافر في الكثير من الأنظمة بدقة عالية نسبياً مكنّتها من التعرف على أنواع مختلفة من النصوص، فضلاً عن وجود العديد من الأدوات المتاحة لتنفيذ هذه الخدمة، تم اختيار إحدى الأدوات الجاهزة لتنفيذ هذه المهمة وذلك لأنه ليس هناك فائدة وجدوى كبيرين من إعادة تنفيذها والعمل عليها طالما أنّها متوافرة بدقة وسرعة تفي بالغرض المرجو منها ضمن النظام، وقد تم اختيار مكتبة pytesseract وهي مكتبة مفتوحة المصدر بلغة بايثون تمثل غلاًفاً لـ Tesseract-OCR المطورة من قبل شركة غوغل، والتي تم تصميمها اعتماداً على [الشبكات العصبونية](#) من نوع LSTM، وتتميز هذه المكتبة بدعمها لأكثر من 100 لغة لكن تم الاكتفاء باللغتين العربية والإنكليزية.

## المصحح التلقائي للأخطاء

يقع الإنسان في الكثير من الأخطاء الإملائية وأخطاء التنسيق أثناء قيامه بالكتابة، لذلك يمكن أن يكون لوجود المصحح التلقائي فائدة كبيرة في التخلص من هذه الأخطاء من خلال تنبيه المستخدم لمكان وجود الخطأ، ويمكن تحقيق فائدة أكبر في حال تم اقتراح تصحيح لهذا الخطأ، لذلك تم العمل على توفير هذه الخدمة ضمن النظام لما لها من فائدة كبيرة في رفع دقة الكتابة والتخلص من الأخطاء الإملائية وأخطاء التنسيق.

تقسم عملية التصحيح إلى قسمين يتم في القسم الأول تصحيح أخطاء التنسيق، ويتم في القسم الثاني تصحيح الأخطاء الإملائية.

يتم تصحيح أخطاء التنسيق من خلال استخدام regular expression والتي من خلالها يتم تحديد مكان الخطأ ثم يتم اقتراح regular expression جديد كتصحيح للخطأ، وفي حال أراد المستخدم تصحيح الخطأ يتم استبدال التعبير القديم الخاطئ بالتعبير الجديد والذي هو تصحيح للتعبير السابق.

أما بالنسبة للأخطاء الإملائية فيتم أولاً تحويل النص إلى مجموعة من الكلمات من خلال التخلص من جميع علامات الترقيم وذلك من خلال استخدام regular expression، بعد عملية التحويل يتم التحقق من صحة كل كلمة من كلمات النص وذلك بالاستفادة من خدمة موجودة مسبقاً تقوم بتصويب الكلمة في حال كانت الكلمة تحوي خطأ إملائياً، هذه الخدمة متوفرة باسم Arabic Spell Checker وليس هناك فائدة وجدوى من إعادة تنفيذها والعمل عليها طالما أنها متوفرة بدقة وسرعة تفي بالغرض المرجو منها ضمن النظام.

## التحقيق

### اللغة المستخدمة

تم الاعتماد على اللغة البرمجية بايثون التي باتت الخيار المفضل للمطورين في مجال التعلم العميق نظراً لبساطتها ومرونتها وتوفيرها لمجموعة كبيرة من الموارد القادرة على تمكين المطورين من العمل بشكل أسرع، فضلاً عن توافقها مع العديد من المكاتب الخاصة بهذا المجال.

### بيئة التطوير المستخدمة

أجريت عمليات التدريب ضمن بيئة العمل المطورة من Google Colab المقدمة من شركة Google، وهي بمثابة منصة لتدريب وتنفيذ عمليات التعلم العميق Deep Learning ونماذج [التعلم التلقائي](#) Machine Learning التي عادةً ما تتطلب وتفرض قيوداً على عملية التطوير مثل امتلاك عتاد حاسوبي

قوي قادر على معالجة العمليات المطلوبة وذاكرة ضخمة قادرة على استيعاب الحجم الكبير لبيانات التدريب، حيث أنّ هذه المنصة تقدم هذه الموارد بشكل مجاني فنقوم بإتاحة GPU بمواصفات عالية كما تقدم ذاكرة RAM بحجم 12GB و Hard Disk بحجم 80GB، ولكن هذه الموارد لم تكن كافية لنا في عملية التدريب لذلك تم الاشتراك بخدمة colab pro والتي تعطي GPU أسرع من نوع nvidia-smi مع ذاكرة تخزين RAM بحجم أكبر ومساحة تخزين أكبر.

## أهم المكتاتب المستعملة

### 1- TensorFlow:

تعتبر واحدة من أطر التعلم الآلي Machine Learning Framework مفتوحة المصدر والمتوافقة مع لغة بايثون والتي تجعل [التعلم الآلي](#) أسرع من خلال تسهيل عملية الحصول على البيانات ونماذج التدريب وتقديم التنبؤات وتحسين النتائج المستقبلية. وعليه تم اعتماد هذه المكتبة لاحتوائها على أعداد كبيرة من نماذج وخوارزميات [التعلم الآلي](#) والتعلم العميق.

### 2- Python Imaging Library (PIL):

واحدة من أهم المكتاتب المستعملة للتعامل مع الصور عند استعمال لغة البرمجة بايثون، وهي مكتبة مفتوحة المصدر تمنح القدرة على التعامل مع الصور والتعديل عليها وإجراء تحويلات على الصور مثل الانزياح والدوران والخ...

### 3- Dlib:

عبارة عن مجموعة أدوات حديثة تحتوي على خوارزميات وأدوات [للتعلم الآلي](#). يتم استخدامها في مجموعة واسعة من المجالات بما في ذلك الروبوتات والأجهزة المضمنة والهواتف المحمولة. وهي مفتوحة المصدر.

### 4- Stanza:

عبارة عن مجموعة من الأدوات الدقيقة والفعالة لتحليل اللغوي للعديد من اللغات البشرية. بدءاً من النص الخام إلى التحليل النحوي والتعرف على الكيانات، تستخدم لتنفيذ عمليات معالجة اللغات الطبيعية.

### 5- PKE:

عبارة عن مجموعة أدوات استخراج عبارة مفاتيح مفتوحة المصدر تعتمد على لغة python.



## الفصل الخامس: الاختبارات والنتائج

### مقاييس التقييم

تعد عملية تقييم النماذج الإحصائية من أكثر العمليات أهمية. ففي حال تم تقييم النموذج بالاعتماد على مقاييس غير كافية، ومن ثم اتخاذ قرارات لتحديد النموذج الأفضل بناءً على تلك النتائج، سيؤول ذلك إلى عدم الوصول إلى النتيجة الأفضل.

لذلك يجب اختيار المقياس المناسب والكافي للعملية من أجل الحصول على نتائج صحيحة.

### مقاييس تقييم توليد شرح تلقائي للصور

#### • المقياس BELU:

هو أحد المقاييس الأولى التي تم استخدامها لقياس التشابه بين جملتين. تم اقتراحه في البداية للترجمة الآلية، يعمل عن طريق حساب مطابقة n-grams في الترجمة الناتجة إلى n-grams في النص المرجعي، حيث سيكون unigram كل كلمة ومقارنة bigram ستكون كل زوج من الكلمات. يتم إجراء المقارنة بغض النظر عن ترتيب الكلمات. يتم حساب عدد المطابقات وكلما زاد عددها كانت ترجمة النموذج أفضل. يمكن التعبير عن المقياس بالشكل الرياضي التالي:

$$p = \frac{m}{w_t}$$

حيث m: عدد الكلمات الموجودة في الجملة الناتجة عن النموذج ووجدت في الجمل المرجعية أيضاً.  
wt: هو عدد الكلمات الموجودة في الجملة الناتجة عن النموذج.

يتراوح مقياس BELU بين 0 و 1 حيث ستحصل جمل قليلة على درجة 1 في حال مطابقتها مع الجملة المرجعية أي النتيجة المثالية غير ممكنة في الممارسة العملية لأن الترجمة يجب أن تطابق المرجع تماماً وهذا غير ممكن حتى من قبل البشر. وينطبق هذه الأمر تماماً مع حالتنا في توصيف الصور وقد تكون معقدة بشكل أكبر.

#### • المقياس ROUGE:

تم اقتراحه في البداية لتقييم أنظمة التلخيص، ويتم هذا التقييم من خلال مقارنة تداخل n-grams وتسلسل الكلمات وأزواج الكلمات، ونظراً لأن مقياس ROUGE يعتمد بشكل كبير على الاسترجاع، فإنه يفضل الجمل الطويلة. كما أن إصدار ROUGE-L يقيس أساساً أطول التكرارات المشتركة بين زوج من الجمل.

#### • المقياس METEOR:

هو مقياس آخر للترجمة الآلية. يتم تعريفه على أنه المتوسط التوافقي للدقة precision و recall لتطابقات unigram بين الجمل. بالإضافة إلى ذلك، فإنه يستخدم المرادفات وإعادة الصياغة. يعالج المقياس العديد من

أوجه القصور في BELU مثل تقييم recall وعدم وجود مطابقة واضحة للكلمات. كما يمكن التعبير عن هذا المقياس بالشكل الرياضي:

$$p = \frac{m}{w_t}$$

$$R = \frac{m}{w_r}$$

$$f_{mean} = \frac{10 \times p \times R}{R + 9 \times p}$$

m: عدد الكلمات الموجودة في الجملة الناتجة عن النموذج ووجدت في الجمل المرجعية أيضاً  
 $w_t$ : هو عدد الكلمات الموجودة في الجملة الناتجة عن النموذج  
 $w_r$ : هو عدد الكلمات الموجودة في الجملة المرجعية  
 $f_{mean}$ : هو المتوسط التوافقي لكل من أجل precision و recall

تعمل المقاييس القائمة على gram-n بشكل جيد عندما يكون هناك تداخل كبير بين الجمل المرجعية والجمل الناتجة عن النماذج، لكنها تفشل في تحديد التشابه الدلالي عندما تكون الكلمات الشائعة نادرة، يعالج METEOR هذه المشكلة إلى حد باستخدام مطابقة المرادفات المستندة إلى [25] WordNet ولكن مجرد النظر إلى المرادفات قد يكون مقيداً للغاية لالتقاط التشابه الدلالي الكلي.

#### • المقياس CIDEr:

هو مقياس حديث تم اقتراحه لتقييم جودة توصيف الصور. يقيس التوافق بين وصف الصورة الناتج عن النموذج  $c_i$  والجمل المرجعية  $S_i$  حيث  $S_i = \{s_{i1}, \dots, s_{im}\}$  هي مجموعة الجمل التي تصف الصورة  $i$  مقدمة من قبل الإنسان، ومن أجل تطبيق هذا المقياس يتم توليد مجموعة 1-4 grams بعد ذلك، يتم حساب التكرارات المشتركة في الجمل المرجعية والجملة الناتجة. في CIDEr، على غرار tf-idf، يتم تقليل وزن n-grams الشائعة في جميع أوصاف الصور. أخيراً، يتم حساب التشابه الجيببي بين n-grams للمرشح والمراجع. تم تصميم CIDEr كمقياس متخصص لتقييم توصيف للصور، ومع ذلك، فهو يعمل بطريقة لغوية بحتة، ويوسع المقاييس الحالية فقط مع ترجيح tf-idf فوق n-grams، يؤدي هذا أحياناً إلى ترجيح تفاصيل غير مهمة للجملة بشكل أكبر، مما يؤدي إلى تقييم تسمية توضيحية غير فعال نسبياً.

## مقاييس تقييم استخراج الكلمات المفتاحية

- **المطابقة التامة:**

تعتمد هذه الطريقة على القيام بعملية مطابقة بين الكلمات المفتاحية المستخرجة من قبل الإنسان والتي تسمى Kgold مع الكلمات المفتاحية التي يقوم النموذج باستخراجها Ktext، وعندها يمكن أن تكون النتيجة صحيحة في حال حصول تطابق كامل بين كل من Kgold و Ktext، وتكون النتيجة خطأ في حال وجود أي اختلاف بينهما، وقد لا يكون مقياساً جيداً للجودة الإجمالية لمجموعة الكلمات الرئيسية المستخرجة، حيث توجد العديد من الحالات التي يفشل فيها هذا النوع من المطابقات، كالعبارات المعجمية التي لها نفس المعنى. [30]

- **part of and Includes**

وهنا يتم اعتبار العبارة الناتجة صحيحة في حال كانت العبارة الرئيسية المستخرجة هي جزء من العبارة الأساسية القياسية الذهبية أو أن تتضمن العبارة الرئيسية المستخرجة عبارة المفاتيح القياسية الذهبية، أو أن العبارتين التي ظهرت لهما نفس المعنى، ويتم العمل على هذه الطريقة بشكل يدوي من قبل الإنسان ولا يمكن تنفيذها بشكل آلي. [30]

الأداء	طريقة قياس الاختلاف	البيانات التي تم الاختبار عليها	التقنية
90.95%	MAP	LFW	Khoi et al. [26]
94.04%	Cosine similarity	LFW	Xi et al. [27]
88.5%	SVM	LFW	Arigbabu et al [25]
75.3%	Cosine similarity	LFW	Hussain et al [28]
88.10	Bayesian Classifier	LFW	Annalakshmi et al. [29]
90.42%	KNN	LFW	Our model

جدول 11: مقارنة بين النتائج التي تم تحقيقها ونتائج الأعمال السابقة في عملية التعرف على الوجوه

## التعرف على الوجوه

يوجد عدد من الطرق التي من الممكن أن تحقق نتائج أفضل لكنها بحاجة إلى أكثر من صورة لكل شخص، كما أنه هنالك طرق من الممكن أن تعطي نتائج أفضل لكنه لا يمكن إضافة وجوه جديدة على الوجوه التي تم التدريب عليها، ونتيجة متطلبات النظام كان لابد من الاعتماد على الطريقة التي تم شرحها سابقاً، وبالنظر إلى الجدول (11) يمكن الاطلاع على النتائج التي تم التوصل لها.

يجب الأخذ بعين الاعتبار عدد الوجوه التي يتم البحث ضمنها لأنها تؤثر على الأداء.

البيانات التي تم الاختبار عليها تحوي 5700 وجه مختلف.

['Neymar', 'Lionel Messi',  
'Cristiano Ronaldo']



['Sami Alesh', 'Mohamad Alaa  
Khairalla', 'Muhannad Al\_Tabba']

Name	BELU 1	BELU 2	BELU 3	BELU 4	METEOR
ATT (EN)	0.709	0.537	0.402	0.304	0.243
Show, Attend and Tell (EN)	0.718	0.504	0.357	0.25	0.230
Show and Tell (EN)	0.666	0.451	0.304	0.203	-
Obeida ElJundi encoder-decoder (end-to-end) (AR)	0.35	0.19	0.10	0.05	-
Obeida ElJundi encoder-decoder (translation) (AR)	0.26	0.11	0.04	0.02	-
Our model trained on MS-COCO with 5k vocab size (AR)	0.22	0.09	0.03	0.01	-
Our model trained on MS-COCO with 6k vocab size (AR)	0.24	0.10	0.04	0.01	-
Our model trained on Flickr8k with 10k vocab size (AR)	0.3	0.17	0.07	0.02	-
Our model trained on MS-COCO with 10k vocab size (AR)	0.42	0.2	0.12	0.05	-

جدول 12: مقارنة بين النتائج التي تم تحقيقها ونتائج الأعمال السابقة

## توصيف الصور

قمنا باختبار النموذج الأفضل (الذي يحوي على آلية اهتمام Attention) على مجموعتين من البيانات المترجمة للغة العربية MS-COCO و Flickr8k، تتكون المجموعة الأولى من 82000 صورة ولكل صورة 5 جمل من التوصيف، أما المجموعة الثانية تتكون من 8000 صورة ولكل صورة 5 جمل من التوصيف، وقمنا بالمقارنة مع النماذج المشابهة باللغة العربية، ولاحظنا أن النتائج الخاصة بنا أصبحت مقاربة إلى حد ما من هذه النماذج مع ملاحظة أن النموذج الذي تم العمل عليه يتم تدريبه بشكل أقل من النماذج المشابهة وأن النموذج الذي تم العمل عليه مخصص للغة العربية والتي تحوي عدد كلمات أكبر بكثير مما تحويه اللغة الإنكليزية والتي يتم المقارنة معها، حيث أثناء التدريب يكفي الاحتفاظ بأكثر 5000 كلمة مكررة ضمن مجموعة البيانات في حالة اللغة الإنكليزية، أما اللغة العربية فعند اختيار الرقم ذاته لاحظنا ضعف النتائج مقارنةً باللغة الإنكليزية وذلك لغنى اللغة العربية بالكلمات

## الفصل الخامس: الاختبارات والنتائج

بشكل أكبر بكثير، فقمنا بزيادة عدد الكلمات ل 10000 مما ساهم بشكل كبير في رفع مستوى النتائج، ويوضح الجدول (12) الفرق في النتائج بين النماذج السابقة والنموذج الخاص بنا.

صور بعض النتائج:

	
امراة تبسّم على هاتفها الخلوي	كلب صغير لطيف يقف على حقل أخضر على العشب
	
رجل يقف في الخارج يرتدي نظارات	غرفة معيشة
	
يوجد تلفزيون يجلس على طاولة قهوة بزهور	سيارة بيضاء متوقفة بين السيارات المتوقفة





تقاطع شارع المدينة في شارع على الجانب



غرفة سرير أنيق به الكثير من الجدران الحمراء



رجل يرقد على منحدر مع حقيبة بالقرب من نهر يمر تحت جسر



حوض المطبخ بجوار موقد



يتم تجميع مجموعة من الأشخاص في مؤتمر تنس



طفل في قميص البرتقالي في قميص بنفسجي ينظر إلى الكاميرا

## استخراج العبارات المفتاحية

تظهر الجداول التالية العبارات المفتاحية التي تم الحصول عليها للنص التالي:

"بلغ الرئيس اللبناني ميشال عون الموفد الأميركي الوسيط المفاوضات المباشرة لترسيم الحدود البحرية الجنوبية، ضرورة استمرار مفاوضات ترسيم الحدود الجانب الإسرائيلي وأكد عون أن لبنان المتمسك بسيادته أرضه ومياهه، يريد تنجح مفاوضات الترسيم وأضاف الصعوبات برزت يمكن تذليلها خلال بحث معمق يركز الحقوق الدولية ومواد قانون البحار جهته، تمنى السفير جون ديروشي خلال لقائه عون استمرار عملية التفاوض مصلحة الجميع وحضر اللقاء الجانب الأميركي السفارة الأميركية بيروت السفيرة دوروثي شيا والمستشار السياسي اودري كيرانين ، الجانب اللبناني الوزير السابق سليم جريصاتي والمدير العام لرئاسة الجمهورية الدكتور انطوان شقير والمستشاران العميد بولس مطر واسامة خشاب وكانت لبنان تبّلع رسمياً، الإثنين الماضي، الجانب الأميركي تأجيل جولة التفاوض المرتقبة الخامسة الأسبوع إسرائيل بشأن ترسيم الحدود البحرية، وفق أفاد مصدر عسكري لبناني يذكر أن موعد الجلسة الخامسة كان محدداً الثاني كانون الأول ديسمبر الجاري وقال المصدر العسكري اللبناني لوكالة فرانس برس ، تبّلعنا رسمياً تأجيل جلسة المفاوضات المباشرة، واستبدالها بجلسة خاصة الجانب اللبناني ، تحديد الأسباب ، لافتاً الجانب الأميركي طلب التأجيل وجاء الاعلان تأجيل الجلسة، اتهام وزير الطاقة الإسرائيلي يوفال شتاينتس تشرين الثاني نوفمبر الجاري، لبنان بأنّه غير موقفه بشأن حدوده البحرية إسرائيل سبع مرات ، محذراً احتمال تصل المحادثات طريق مسدود المقابل، نفت الرئاسة اللبنانية الاتهام الإسرائيلي، مؤكدة أن موقف بيروت ثابت مسألة ترسيم الحدود البحرية ويتمسك لبنان بحقوقه البحرية ويرفض التنازل عنها لصالح إسرائيل ، قدم طاولة المفاوضات مستندات ووثائق وخرائط تثبت حقه بحدود مياهه البحرية المفاوضات اللبنانيون تمسكوا بمطلب بزيادة كيلومتراً إضافية الـ كيلومتراً، يعني نصف حقل كاريش ملك للبنان المقابل، وصفت الحكومة الإسرائيلية المطالب بـ الاستفزازية وعقد لبنان و إسرائيل جولات المفاوضات المباشرة، شهر تشرين الأول أكتوبر الماضي، برعاية الولايات المتحدة، والأمم المتحدة"

Bert		
consine similarity	cosine similarity with Max Sum Similarity	cosine similarity with Maximal Marginal Relevance
'الحدود البحرية الجنوبية' (0.66592383)	عون	الحدود البحرية الجنوبية
(0.6598224, 'لبنان')	إسرائيل	بولس مطر
(0.65261567, 'كاريش')	بولس مطر	فرانس برس
(0.60083336, 'والأمم المتحدة')	أنطوان شقير	يوفال شتاينتس
(0.59737706, 'الولايات المتحدة')	يوفال شتاينتس	عون
(0.585487, 'اودري كيرانين')	أسامة خشاب	وزير الطاقة الإسرائيلي
'وزير الطاقة الإسرائيلي' (0.5780264)	فرانس برس	اللبنانيون
(0.5726024, 'الحكومة الإسرائيلية')	اللبنانيون	اودري كيرانين
(0.56952417, 'بيروت')	بيروت	سليم جريصاتي
(0.5604782, 'اللبنانيون')	الحكومة الإسرائيلية	لبنان
(0.5514766, 'فرانس برس')	وزير الطاقة الإسرائيلي	والأمم المتحدة
(0.54576415, 'جون ديروشي')	اودري كيرانين	أسامة خشاب
(0.533293, 'دوروثي شيا')	الولايات المتحدة	دوروثي شيا
(0.53264636, 'سليم جريصاتي')	والأمم المتحدة	كاريش
(0.51274294, 'ميشال عون')	لبنان	بيروت
(0.50562584, 'أسامة خشاب')		
(0.48531654, 'يوفال شتاينتس')		



Doc2Vec		
consine similarity	cosine similarity with Max Sum Similarity	cosine similarity with Maximal Marginal Relevance
(0.27111173, 'لبنان')	إسرائيل	لبنان
(0.25431168, 'بيروت')	كاريش	والأمم المتحدة
(0.24436428, 'اللبنانيون')	والأمم المتحدة	الحكومة الإسرائيلية
(0.22952476, 'عون')	الحكومة الإسرائيلية	كاريش
(0.2293303, 'الحدود البحرية الجنوبية')	اودري كيرانين	إسرائيل
(0.21968275, 'يوفال شتاينتس')	الولايات المتحدة	أسامة خشاب
(0.19578224, 'بولس مطر')	أسامة خشاب	دوروثي شيا
(0.17823738, 'وزير الطاقة الإسرائيلي')	دوروثي شيا	وزير الطاقة الإسرائيلي
(0.17452979, 'أنطوان شقير')	أنطوان شقير	يوفال شتاينتس
(0.16755138, 'دوروثي شيا')	وزير الطاقة الإسرائيلي	بولس مطر
(0.16422088, 'فرانس برس')	بولس مطر	الولايات المتحدة
(0.15956807, 'أسامة خشاب')	يوفال شتاينتس	الحدود البحرية الجنوبية
(0.14620665, 'سليم جريصاتي')	الحدود البحرية الجنوبية	أنطوان شقير
(0.14573568, 'ميشال عون')	اللبنانيون	عون

Graph Methods	
Topic Rank	Position Rank
(0.1387454851962053, 'لبنان')	(0.05933243894520201, 'الجانب اللبناني')
(0.1341860575852838, 'الصالح إسرائيل')	(0.05123485216823974, 'الرئيس اللبناني')
(0.09159419053878529, 'الصعوبات')	(0.0500116088060779, 'الجانب الأميركي')
(0.08637457694234399, 'التنازل')	(0.048043515841470416, 'لبنان بحقوقه البحرية')
(0.08275045397814093, 'مفاوضات الترسيم')	(0.040077890466995635, 'اللقاء الجانب الأميركي')
(0.07634668922719787, 'الاستفزازية')	(0.03674842125773139, 'المصدر العسكري اللبناني')
(0.06892838820202164, 'تذليلها')	(0.02796742790850995, 'نوفمبر الجاري، لبنان')
(0.057626691332985286, 'طاولة المفاوضات مستندات ووثائق وخرائط')	(0.027304241084396247, 'لبنان المتمسك')
(0.05384282262485679, 'موعد الجلسة الخامسة')	(0.026162853528865982, 'عون')
(0.05380937782125933, 'مصدر عسكري لبناني')	(0.02045983092775691, 'مفاوضات الترسيم')


TF-IDF	
(تبرسيم الحدود, 0.18481766596463556)	(تبل, 0.13120671363431946)
(المفاوضات المباشرة, 0.1728252614777919)	(اللبناني, 0.13045435798792446)
(الجانب, 0.17052553149682265)	(الحدود, 0.12323911757899589)
(الحدود البحرية, 0.16896456954822364)	(المفاوضات, 0.12062676947638896)
(ترسيم, 0.163143134688033)	(الإسرائيلي, 0.10884110224909814)
(لبنان, 0.15489087321285724)	(الجانب الأمريكي, 0.10408535184649952)
(إسرائيل, 0.14131177325121796)	(السفيرة, 0.1025451667151097)
(عون, 0.1394700707697926)	(أن, 0.10124659504660297)

نتائج عملية المقارنة بين الطرق السابقة لتحديد الطريقة الأفضل:

	المطابقة التامة	part of and Includes
Bert, Doc2Vec	0.4424	0.733
Topic Rank	0.0479	0.095
Position Rank	0.3351	0.0613
TF-IDF	-	0.319

## خدمة قراءه النصوص

ملف PDF	<p>العربية. حيث يقوم النظام باستخراج الكلمات المفتاحية بشكل تلقائي فور انتهاء المستخدم من كتابة المنشور ويعرض له قائمة بالكلمات المفتاحية المناسبة ليختار منها ما يناسبه. بالإضافة لخدمة توليد شرح مناسب للصور وباللغة العربية، والتي يستطيع المستخدم الاستفادة منها وبإضافة هذا التوصيف لصفحة البرمجة الخاصة بصفحة الويب والتي ستمنح محركات البحث قدرة أكبر على تحديد محتوى الموقع، كما ويقدم النظام المقترح خدمة التعرف على وجوه الأشخاص الموجودين ضمن صورة وتعتبر هذه الخدمة مكتملة لخدمة توليد شرح للصورة، ويتميز النظام بالقدرة على البحث عن جملة معينة باللغتين العربية والإنكليزية، كما أنه تم إضافة مجموعة من الخدمات التي تساعد المحرر في عملية الكتابة وتمنحه الدقة والسرعة فتم إضافة مصحح تلقائي لتصحيح الأخطاء الإملائية وإعطاء اقتراح لتصحيحها، كما ويملك المستخدم القدرة على إضافة صورة ويقوم النظام باستخراج الكتابة الموجودة ضمن الصورة أو ضمن الملفات التي لها صيغة من نمط Pdf مما يسهل على المستخدم عملية إعادة الكتابة.</p>
النتيجة	<p>العربية. حيث يقوم النظام باستخراج الكلمات المفتاحية بشكل تلقائي فور انتهاء المستخدم من كتابة المنشور ويعرض له قائمة بالكلمات المفتاحية المناسبة ليختار منها ما يناسبه. بالإضافة لخدمة توليد شرح مناسب للصور وباللغة العربية» والتي يستطيع المستخدم الاستفادة منها وبإضافة هذا التوصيف لصفحة البرمجة الخاصة بصفحة الويب والتي ستمنح محركات البحث قدرة أكبر على تحديد محتوى الموقع، كما ويقدم النظام المقترح خدمة التعرف على وجوه الأشخاص الموجودين ضمن صورة وتعتبر هذه الخدمة مكتملة لخدمة توليد شرح للصورة ععلق تبك عن 51م7 امينة تين العربية والإنكليزية « كما أنه تم إضافة مجموعة من الخدمات التي تساعد المحرر في عملية الكتابة وتمنحه الدقة والسرعة فتم إضافة مصحح تلقائي لتصحيح الأخطاء الإملائية وإعطاء اقتراح لتصحيحها كما ويملك المستخدم القدرة على إضافة صورة ويقوم النظام باستخراج الكتابة الموجودة ضمن الصورة أو ضمن الملفات التي لها صيغة من نمط 504 مما يسهل على المستخدم عملية إعادة الكتابة.</p>

صورة	
النتيجة	<p>How to turn on text to speech</p>

## المصحح التلقائي

النص قبل التصحيح	<p>ادخل نصاً هتا من اجل التدقيق، قم باضغط على زر ABC على شريط الاختصارات ( الأدوات ) . انقر بالزر هتا الأيمن على الكيمة المعلمة بالإخمر لترى الاقتراحات، المتوغة</p>
النص بعد تصحيح الأخطاء الاملائية	<p>ادخل نصاً هذا من اجل التدقيق، قم بالضغط على زر ABC على شريط الاختصارات ( الأدوات ) . انقر بالزر هذا الأيمن على الكبيرة المعلمة بالأحمر لترى الاقتراحات، المتأخرة</p>
النص بعد تصحيح أخطاء التنسيق	<p>ادخل نصاً هتا من اجل التدقيق، قم باضغط على زر ABC على شريط الاختصارات (الأدوات). انقر بالزر هتا الأيمن على الكيمة المعلمة بالإخمر لترى الاقتراحات، لمتوغة</p>
النص بعد تصحيح أخطاء التنسيق والأخطاء الاملائية	<p>ادخل نصاً هذا من اجل التدقيق، قم بالضغط على زر ABC على شريط الاختصارات (الأدوات). انقر بالزر هذا الأيمن على الكبيرة المعلمة بالأحمر لترى الاقتراحات، المتأخرة</p>

## الخاتمة

قدمنا في هذا البحث الإطار العام لنظام إدارة محتوى ذكي يدعم اللغة العربية من خلال عدّة خدمات ذكية لمساعدة صانعي المحتوى العربي، كان أبرزها خدمة تنميط النصوص العربية (استخراج الكلمات المفتاحية) التي ساهمت في رفع كفاءة المحتوى العربي وتحسين نتائج البحث، وساهمت أيضاً بتقليل كلفة إنشاء المحتوى من خلال اختصار الوقت اللازم لإنشاء الكلمات المفتاحية لكل مستند، أما من وجهة نظر المستخدمين المتابعين للمحتوى فساهمت بإعطائهم فكرة أكبر وأسرع لما يحتوي المنشور قبل القيام بقراءته، أيضاً تم إنجاز خدمة التوصيف التلقائي للمشاهد البصرية باللغة العربية، والتي اعتُبرت من المسائل التي تحمل تحديات كبيرة في الذكاء الصناعي كونها تربط بين مجال معالجة الصور ومجال معالجة اللغات الطبيعية، والتي ساهمت في استخراج توصيف تلقائي للصور وتضمين ذلك التوصيف ضمن صفحة الويب مما يساهم في رفع كفاءة محركات البحث الخارجية والداخلية (المضمنة ضمن الموقع) وبالتالي رفع تقييم الموقع (SEO)، وقد أظهر النظام المقترح قدرة على توليد وصف بجودة جيّدة جعلته يقترب من نتائج الأعمال المشابهة التي تفوقت في هذا المجال والتي تمت بكمية موارد أكبر بكثير مما يتوقّر لدينا، فضلاً عن خدمات أخرى كالتعرف على الأشخاص ضمن الصور والتي قمنا بمكاملتها مع الخدمة السابقة لوضع وصف للصور يحوي أسماء الأشخاص الموجودين في الصورة، وخدمات أخرى أيضاً مثل المصحح التلقائي الذي ساهم في رفع مستوى المحتوى العربي بشكل كبير وخصوصاً عند صانعي المحتوى المبتدئين، وخدمة البحث بعدّة لغات التي أدّت للحصول على نتائج ذات صلة أكبر من السابق، وكذلك خدمة استخراج النص الموجود ضمن الصورة باللغتين العربية والإنجليزية لمساعدة الكاتب على تحويل المحتوى المطبوع للشكل الرقمي، فكل هذه الخدمات السابقة عزّزت هدف المشروع لكونها من المهام التي كثيراً ما يعاني منها صناع المحتوى العربي، لنكون بذلك وصلنا إلى نظام يضاهي الأنظمة الأجنبية لمساعدة صانعي المحتوى بشكل فعال وسهل لإدارة المحتوى بشكل احترافي.

\*\*\*تم بعون الله\*\*\*

- [1] j. schmidhuber, Deep learning in neural networks: An overview, 2015.
- [2] K. Pykes, "Part Of Speech Tagging for Beginners," 26 Nov 2020. [Online]. Available: <https://towardsdatascience.com/part-of-speech-tagging-for-beginners-3a0754b2ebba>.
- [3] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation , (1997)
- [4] Oriol Vinyals, Alexander Toshev, Samy Bengio, Dumitru Erhan, "Show and Tell: A Neural Image Caption Generator", arxiv.org,20 Apr 2015.[online]. Available: <https://arxiv.org/pdf/1411.4555.pdf>
- [5] Mikolov, Tomas, et al. 'Efficient Estimation of Word Representations in Vector Space'. ArXiv:1301.3781 [Cs], Sept. 2013. arXiv.org, [Online]. Available: <http://arxiv.org/abs/1301.3781>.
- [6] LIANG Xu, "Understand TextRank for Keyword Extraction by Python," Medium, 18 Feb 2019. [Online]. Available: <https://towardsdatascience.com/textrank-for-keyword-extraction-by-python-c0bae21bcec0>.
- [7] F. B. D. Adrien Bougouin, "TopicRank: Graph-Based Topic Ranking for Keyphrase Extraction," Universite of Nantes, LINA, France, 2013.
- [8] P. T. Rada Mihalcea, "TextRank: Bringing Order into Texts," University of Michigan, Texas, September 2018.
- [9] C. C. Corina Florescu, "PositionRank: An Unsupervised Approach to Keyphrase Extraction," University of North Texas, USA, 2017.
- [10] V. Jayaswal, "Text Vectorization: Term Frequency — Inverse Document Frequency (TFIDF)," 4 Oct 2020. [Online]. Available: <https://towardsdatascience.com/text-vectorization-term-frequency-inverse-document-frequency-tfidf-5a3f9604da6d>.

- [11] A. R. Samhaa R. El-Beltagy, "KP-Miner: Participation in SemEval-2," The American University in Cairo, Cairo, 2010.
- [12] K. Bennani-Smires, "Simple Unsupervised Keyphrase Extraction Using Sentence Embeddings," arXiv.org, sept 2018. [Online]. Available: <http://arxiv.org/abs/1801.04470..>
- [13] G. W. P. G. F. Ian Witten, "KEA: Practical Automatic Keyphrase Extraction," Berkeley, CA, USA, 1999.
- [14] Farhadi, Ali, Mohsen Hejrati, Mohammad Amin Sadeghi, Peter Young, Cyrus Rashtchian, Julia Hockenmaier, and David Forsyth. "Every picture tells a story: Generating sentences from images." In European conference on computer vision, pp. 15-29. Springer, Berlin, Heidelberg, 2010
- [15] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1-9. 2015.
- [16] Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "Imagenet: A largescale hierarchical image database." In 2009 IEEE conference on computer vision and pattern recognition, pp. 248-255. Ieee, 2009.
- [17] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9, no. 8 (1997): 1735-1780.
- [18] Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." arXiv preprint arXiv:1409.0473 (2014).
- [19] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
- [20] You, Quanzeng, Hailin Jin, Zhaowen Wang, Chen Fang, and Jiebo Luo. "Image captioning with semantic attention." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4651-4659. 2016.

- [21] Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. "Microsoft coco: Common objects in context." In European conference on computer vision, pp. 740-755. Springer, Cham, 2014.
- [22] Geitgey, Adam. "Machine Learning Is Fun! Part 4: Modern Face Recognition with Deep Learning". Medium, 24 Sept. 2020. [Online]. Available: <https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78>.
- [23] Le, Quoc V., and Tomas Mikolov. 'Distributed Representations of Sentences and Documents'. ArXiv:1405.4053 [Cs], 2, May 2014. arXiv.org, [Online]. Available: <http://arxiv.org/abs/1405.4053>.
- [24] I. Shrivastava, "Exploring Different Keyword Extractors — Evaluation Metrics and Strategies," medium, 14 Jul 2020. [Online]. Available: <https://medium.com/gumgum-tech/exploring-different-keyword-extractors-evaluation-metrics-and-strategies-ef874d336773>.
- [25] Arigbabu, O.A.; Ahmad, S.M.S.; Adnan, W.A.W.; Yussof, S.; Mahmood, S. Soft biometrics: Gender recognition from unconstrained face images using local feature descriptor. arXiv 2017, arXiv:1702.02537.
- [26] Khoi, P.; Thien, L.H.; Viet, V.H. Face Retrieval Based on Local Binary Pattern and Its Variants: A Comprehensive Study. Int. J. Adv. Comput. Sci. Appl. 2016, 7, 249–258
- [27] Xi, M.; Chen, L.; Polajnar, D.; Tong, W. Local binary pattern network: A deep learning approach for face recognition. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 3224–3228.
- [28] Arashloo, S.R.; Kittler, J. Efficient processing of MRFs for unconstrained-pose face recognition. In Proceedings of the 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), Rlington, VA, USA, 29 September–2 October 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 1–8.

- [29] Annalakshmi, M.; Roomi, S.M.M.; Naveedh, A.S. A hybrid technique for gender classification with SLBP and HOG features. *Clust. Comput.* 2019, 22, 11–20.
- [30] Torsten Zesch, Iryna Gurevych; “Approximate Matching for Evaluating Keyphrase Extraction”. [Online]. Available: <https://aclanthology.org/R09-1086.pdf>