تحليل المشاعر القائم على الذكاء الاصطناعي كأداة لتقييم فعالية خدمات المعلومات (1)

د. ليلى سيد سميع

أستاذ مساعد بقسم المكتبات والوثائق وتقنية المعلومات
كلية الآداب – جامعة القاهرة
Laila_mtu@yahoo.com

تاريخ الاستلام: 25 يونيو 2022 **تاريخ القبول: 30 يونيو 2022**

المستخلص:

تناقش هذه الدراسة أسلوب تحليل المشاعر، كأحد تطبيقات الذكاء الاصطناعي واستكشاف فرص الاستفادة منه كمؤشر لتقييم خدمات المعلومات من خلال قياس رضاء المستفيدين عن طريق التحليل الآلي لما ينشر من تعليقات عبر وسائل التواصل الاجتماعي للصفحات الخاصة بمؤسسات المعلومات، مع استعراض نماذج للأدوات والتقنيات الرئيسية المستخدمة في عمليات التحليل الآلي للمشاعر، وأبرز المشكلات التي قد تواجه الاستثمار الفعال لذلك الأسلوب في بيئة المكتبات العربية، وتم الاستعانة بالمنهج الاستكشافي من خلال إجراء دراسة على عينة بلغت (665) تغريدة/تعليقا للمستفيدين عبر وسائل التواصل الاجتماعي (تويتر وفيس بوك) خلال الفترة من ديسمبر 2021 وحتى نهاية مايو 2022، بالبحث بعدد (16) توسيمًا وكلمة مفتاحية ذات علاقة بخدمات المكتبات والمعلومات، وتم إجراء التحليلات اللازمة باستخدام منصة Google Colab من خلال بناء نموذج للتعلم الآلي بتطبيق خوارزمية Naïve-Bays. وأشارت نتائج الدراسة إلى أهمية الاستفادة من البيانات الناتجة عن التحليلات التفصيلية التي يقدمها النموذج، مثل: تحديد قطبية آراء المستفيدين (إيجابية – محايدة – سلبية) نحو ما تنشره المكتبة من منشورات، في قياس رضاء المستفيدين بما يسهم في دعم عمليات اتخاذ القرار لتحسين مستوى تقديم الخدمات، ويمكن للمكتبة أن تعتمد عند التطبيق على أحد التطبيقات الجاهزة أو أن تقوم ببناء النموذج الخاص بها سواء من خلال التقنيات المدفوعة أو مفتوحة المصدر، وأشارت الدراسة إلى عدد من التحديات في هذا الصدد، من بينها ما يتعلق بطبيعة اللغة العربية ولهجاتها عند معالجة اللغة الطبيعية، وقلة ما يتوافر من أدوات أومجموعات بيانات جاهزة

(1) **تم عرضه في الندوة العلمية الأولي عن تطبيقات الذكاء الاصطناعي فى مجالات العلوم الاجتماعية والانسانيات :الخصائص . البرمجيات . آليات التنفيذ، كلية الآداب، جامعة القاهرة، في الفترة من 29 – 30 ديسمبر 2021.**

أومكتبات برمجية تتعامل مع النص العربي بفاعلية، وطبيعة محتوى التعليقات وغموض دلالة بعضها، وكذلك ضعف المشاركة الفعالة للمستفيدين في التعليقات، فضلًا عن التحديات ذات العلاقة بقضايا الخصوصية، وأوصت الدراسة بضرورة تبني دراسات بينية يتعاون فيها باحثون من تخصصات علوم المكتبات والمعلومات، والحاسبات، واللغويات لتصميم وبناء أدوات برمجية جاهزة وموارد معجمية تدعم إمكانات اللغة العربية، يمكن أن تستخدمها المكتبات وغيرها من مؤسسات المعلومات بصورة مباشرة دون الحاجة للتعقيدات البرمجية لعمليات تحليل الآراء بصورة مفصلة.

الكلمات المفتاحية: تحليل المشاعر(1)؛ معالجة اللغة الطبيعية(2)؛ رضاء المستفيدين(3)؛ خدمات المعلومات(4)؛ شبكات التواصل الاجتماعي(5).

أولًا: الإطار المنهجي للدراسة:

1/1- المقدمة:

شهدت السنوات الماضية نموًا متزايدًا في عدد مستخدمي مواقع التواصل الاجتماعي، فقد تحولت من مجرد وسائل للتواصل الشخصي أو الاستخدام لأغراض الترفيه والتسلية إلى قنوات أو منصات تستخدمها الهيئات والمنظمات بمختلف أنواعها بغرض التواصل الفعال مع جمهور المستفيدين من الخدمات التي تقدمها، وكان من بين تلك المنظمات والهيئات، المكتبات وغيرها من مؤسسات المعلومات بطبيعة الحال، وحرصت الكثير من المكتبات وغيرها من مؤسسات المعلومات على إنشاء الحسابات الرسمية الخاصة بها على منصات مثل: تويتر وفيس بوك وغيرها، لبث المعلومات ونشر الأخبار والمستجدات نظرًا لما تقدمه من إمكانات بث المعلومات مهما بلغ حجمها، ونشرها/ مشاركتها بأسرع وقت مع أكبر عدد من المستفيدين وبصورة يسيرة بالنسبة لهم (عصام عبيد، 2014)، فضلًا عما تقدمه من مؤشرات بديلة في إطار قياسات الويب استنادًا على عدد مرات الإعجاب والمشاركة وعدد المتابعين لصفحة ما (أمل وجيه حمدى، 2019). وقد أصبحت منصة فيس بوك على سبيل المثال أحد أهم قنوات التواصل التي تبث من خلالها المكتبات الروابط المهمة الخاصة بمصادر المعلومات والفهرس الإلكتروني وغيرها من الروابط الخاصة بالخدمات التي تقدمها المكتبة لمستفيديها (Lamba, M., & Madhusudhan, M., 2022).

وتلعب صفحات التواصل الاجتماعي الخاصة بالمكتبات وغيرها من مرافق المعلومات – إذا ما أحسن استثمار إمكاناتها – دورًا مهمًا في التسويق لخدمات المعلومات والتعرف على استفسارات وتعليقات المستفيدين، مما قد يساعد في عمليات تطوير وتحسين مستوى تقديم الخدمات بهذه المكتبات من خلال تلك الوسيلة المجانية (هبة المتبولي، 2019)، وقد عزز من أهمية شبكات التواصل الاجتماعي مع ظهور أدوات وتقنيات تساعد في تحليل الكميات الضخمة من البيانات التي تتاح عبرها واستخلاص المؤشرات التي تجعل المؤسسة أكثر فهمًا لمستفيديها بهدف تحسين مستوى تقديم الخدمة.

ويعد مجال تحليل المشاعر من المجالات التي تحظى بالاهتمام في مجال قياس ولاء المستخدمين من خلال تعليقاتهم ومراجعاتهم بشأن منتج ما، أو آرائهم بشأن قضية أو موضوع ما باستخدام أدوات وتقنيات تساعد في تحليل تلك البيانات واستخلاص المؤشرات التي تساعد المؤسسة في فهم مستفيديها بصورة أكبر بهدف تحسين مستوى تقديم الخدمة، وتم توظيف أسلوب تحليل المشاعر في بعض المجالات كالتسويق وفي مجال الدراسات الإعلامية (ريهام سامي، 2021)، ويمكن من خلال هذا الأسلوب تحليل التعليقات والتغريدات وغيرها من آراء المستفيدين بما في ذلك مراجعات المنتجات أيضًا من خلال الاستفادة من وجهات النظر المتعددة لجمهور

المستفيدين بشأن منتج واحد أو خدمة واحدة، فضلًا عن استخدامه أيضًا لتحليل مراجعات الأفلام وعروض الكتب، كما استخدم أسلوب تحليل المشاعر في تحليل التعليقات النصية والإجابات عن الأسئلة المفتوحة باستقصاءات الرأي الموجهة للمستفيدين (Drus & Khalid, 2019).

ويقصد بأسلوب تحليل المشاعر تحليل الآراء النصية وتصنيفها إلى آراء إيجابية وسلبية باستخدام تقنيات وأدوات معالجة اللغة الطبيعية، ويقع هذا المجال في علم الحاسوب كجزء من مجال اهتمام علوم البيانات مرتبط بمجال الذكاء الاصطناعي (تعلم الآلة Machine Learning) – ومجال التنقيب في النصوص، وهو أداة تصنيف النص الأكثر شيوعًا وتستخدمه الشركات للوقوف على تعليقات المستخدمين بخصوص منتجاتهم عبر وسائل التواصل الاجتماعي (Papachristopoulos et al, 2019).

وعند الحديث عن خدمات المعلومات، فلا يمكن إغفال مفهوم رضاء المستفيدين عن الخدمات، وقد تعاني المكتبات وغيرها من مؤسسات المعلومات في كثير من الأحيان في التعرف بدقة وفهم احتياجات المستفيدين، وقد يشوب الاستبيانات كأداة لاستطلاع آراء المستفيدين في هذا الصدد بعض المشكلات ولعل من بينها؛ عزوف المستفيدين عن الإجابة، وقد تم توظيف إمكانات الذكاء الاصطناعي في هذا السياق للتعامل مع البيانات الكيفية والتي تصدر بصورة تلقائية عن المستفيد نتيجة لتعامله واستخدامه للخدمة، ومن أمثلة تلك الإمكانات ما يتعلق بتطبيق أسلوب تحليل المشاعر Sentiment analysis – محلل النبرات Tone Analyzer – روبوتات المحادثة الآلية Chatbots.

وقد تم استخدام أسلوب تحليل المشاعر على النص الذي يتم استرجاعه من خلال تعليقات المستفيدين وتغريداتهم ومنشوراتهم أو من خلال تحليل التغذية الراجعة feedback من استجاباتهم أو تعليقاتهم على الأسئلة المفتوحة الواردة باستبيانات المكتبة، أو مراجعات الكتب، فضلًا عن ذلك ظهر اتجاه متزايد نحو استخدام مصطلح الرؤى insights وتطبيق تقنيات التعلم العميق Deep Learning لتعظيم الاستفادة منها في فهم طبيعة واحتياجات المستفيدين (Al-Otaibi, 2018) .

2/1- مشكلة الدراسة ومبررات اختيارها:

تتميز الفترة الزمنية الحالية بأنها حقبة البيانات الضخمة (حجمًا، ونوعًا، وتسارعًا)، وقد شكلت وسائل التواصل الاجتماعي أحد أهم روافد البيانات الضخمة التي من الممكن أن تمد المكتبات وغيرها من مؤسسات المعلومات بكم هائل من تلك البيانات من خلال ما ينشر على صفحاتها من تعليقات ومشاركات أو حتى استخدام للرسوم التعبيرية emojis، وتتاح هذه التعليقات والآراء للجميع بما يمكن أن يعكس صورة ذهنية ما عن المكتبة وخدماتها سواء بالسلب أو الإيجاب، نتيجة لذلك الكم من البيانات قد يصعب على المكتبة متابعته واستخراج المؤشرات اللازمة منه بصورة يدوية، وبخاصة وأن البيانات في هذه الحالة بيانات كيفية وليست كمية، وقد أدى ذلك لتنامي الحاجة إلى توظيف تقنيات الذكاء الاصطناعي في جلب وتحليل مشاعر أو آراء المستفيدين من خدمات المعلومات من خلال استثمار تلك البيانات الخام التي تعد ذات قيمة كبيرة إذا ما كان هناك تفاعل من جانب المستفيدين حيث تساعد في تحديد درجة الاستقطاب (إيجابي – سلبي – محايد) وغيرها من الإمكانات.

وتسعى الدراسة لاستكشاف فرص تطبيق أسلوب تحليل المشاعر في بيئة المكتبات العربية كمؤشر للدلالة على شعور وآراء المستفيدين فيما يقدم لهم من خدمات، مع محاولة رصد ما قد ينتج من مشكلات عند التطبيق ذلك الأسلوب.

3/1- تساؤلات الدراسة:

- ما المقصود بأسلوب تحليل المشاعر، وما أبرز الإمكانات التي يمكن أن يقدمها في مجال خدمات المعلومات؟
- ما أبرز التقنيات والأدوات التي تتطلبها عمليات التحليل الآلي للمشاعر؟
- هل يمكن الاستفادة من أنظمة تحليل المشاعر في الحصول على مؤشرات تتعلق بمدى رضاء المستفيدين عن خدمات المعلومات؟
- ما أبرز الصعوبات والتحديات التي قد تواجهها عمليات تحليل المشاعر في بيئة المكتبات العربية من واقع الدراسة الاستكشافية؟

4/1- أهمية الدراسة:

تكمن أهمية الدراسة في مساعدة المكتبات وغيرها من مؤسسات المعلومات في توظيف واستثمار تطبيقات الذكاء الاصطناعي فيما يتعلق بأسلوب تحليل المشاعر، وما يتضمنه من أدوات وتقنيات في تحليل الكميات الضخمة من البيانات المنشورة عبر صفحاتها الرسمية على مواقع التواصل الاجتماعي واستخلاص المؤشرات التي تجعلها أكثر فهمًا لمستفيديها بهدف تحسين مستوى تقديم الخدمة.

5/1- أهداف الدراسة:

يهدف البحث إلى دراسة إمكانية تطبيق أسلوب تحليل المشاعر المعتمد على الذكاء الاصطناعي في تقييم فعالية خدمات المعلومات عن طريق قياس مدى رضاء المستفيدين عن خدمات المعلومات من خلال ما ينشرونه من تعليقات عبر وسائل التواصل الاجتماعي لمؤسسات المعلومات عن طريق بناء نموذج لتحليل المشاعر كأداة للاستكشاف، وتسعى الدراسة إلى تحقيق الأهداف الفرعية التالية :

- الوقوف على أسلوب تحليل المشاعر وأبرز تقنيات الذكاء الاصطناعي لمعالجة اللغة الطبيعية المستخدمة في هذا الصدد.
- استكشاف إمكانات وفرص تطبيق أسلوب تحليل المشاعر لقياس مستوى رضاء المستفيدين عن خدمات المعلومات في بيئة المكتبات العربية.
- رصد التحديات التي تحول دون الاستفادة المثلى من أسلوب تحليل المشاعر كأداة يمكن استخدامها بصورة منتظمة لإمداد متخذ القرار في المكتبات وغيرها من مؤسسات المعلومات في عمليات تقييم خدمات المعلومات.

6/1- حدود الدراسة:

الحدود اللغوية:

- التعليقات والتغريدات المنشورة على الصفحات الرسمية للمكتبات مع التركيز على تجميع التعليقات باللغة العربية.

الحدود الجغرافية:

- أول عشر تغريدات/تعليقات تم نشرها نتيجة لعملية البحث بالكلمات المفتاحية والتوسيمات التي تم تحديدها بالدراسة على الصفحات الرسمية لمكتبات أو مؤسسات معلومات عربية بترتيب ظهورها على موقعي تويتر وفيس بوك.

الحدود الزمنية:

- التغريدات والتعليقات المنشورة على منصتي تويتر وفيس بوك والتي تم تجميعها خلال الفترة من أول ديسمبر 2021 وحتى نهاية مايو 2022 (بصرف النظر عن تواريخ نشر تلك التغريدات أو التعليقات، حيث تمثل تلك الفترة الزمنية تاريخ التجميع).

الحدود الموضوعية:

- فرص وإمكانات تطبيق أسلوب تحليل المشاعر باستخدام تقنيات الذكاء الاصطناعي لمعالجة اللغة الطبيعية في قياس رضاء المستفيدين كمؤشر لتقييم خدمات المعلومات بصورة عامة.

7/1- منهج الدراسة وإجراءاتها:

يعتمد البحث في الأساس على المنهج الاستكشافي الوصفي من خلال إجراء دراسة على عينة من تعليقات المستفيدين المنشورة عبر وسائل التواصل الاجتماعي (تويتر وفيس بوك) لصفحات المكتبات ومؤسسات المعلومات العربية (جامعية وعامة ومدرسية) خلال الفترة من شهر ديسمبر 2021 وحتى نهاية مايو 2022، بالاستعانة بواجهة برمجة التطبيقات لمنصة تويتر Twitter API للتجميع الآلي لتغريدات تويتر، والتجميع اليدوى لتعليقات المستفيدين عبر منصة فيس بوك، وتم تكوين مجموعة بيانات dataset بلغت (665) تعليقًا تم تجميعها من خلال البحث في المنصتين بعدد 16 توسيمًا وكلمة مفتاحية ذات علاقة بخدمات المكتبات والمعلومات (سوف يتم سردها في الجزء التطبيقي)، وتم إجراء التحليلات اللازمة بالاستفادة من خدمات منصة Google Colab باستخدام لغة البايثون واستيراد مجموعة مكتبات الأكواد ذات العلاقة من خلال مفكرة جوبيتر Jupyer Notebook، وتم بناء نموذج التعلم الآلي بتطبيق خوارزمية Naïve-Bays.

وقد تم القيام بدراسة استطلاعية قبل عملية التنفيذ النهائي للنموذج لعينة مبدئية من تعليقات وتغريدات المستفيدين عبر وسائل التواصل الاجتماعي خلال النصف الأول من شهر ديسمبر 2021 تم الحصول عليها من خلال البحث بتوسيمات المصطلحين (#خدمات_المكتبات – #خدمات_المعلومات) عبر منصتي تويتر وفيس بوك لتكوين مجموعة بيانات dataset استرشادية بلغت (50) تعليقًا، بالإضافة إلى الاستعانة بواجهة برمجة التطبيقات Twitter API للتجميع الآلي لعدد (100) تغريدة مع تجريب عدد من المنصات لإجراء التحليلات اللازمة (خدمات سحابة آي بي إم IBM Cloud مع ربطها بتطبيق Postman – منصة Google Colab، بالإضافة إلى عدد من التطبيقات الجاهزة)، وذلك للخروج بمؤشرات تتعلق بالتقنيات التي يقترح استخدامها بالدراسة وكذلك مجموعات البيانات[2].

(2) **تم عرض المؤشرات الأولية للدراسة الاستطلاعية في ورقة بعنوان : "تحليل المشاعر (الآراء) المعتمد على الذكاء الاصطناعي كأداة لتقييم فعالية خدمات المعلومات" في: أعمال الندوة العلمية الأولي: تطبيقات الذكاء الاصطناعي في العلوم الاجتماعية والإنسانيات: الخصائص – البرمجيات – آليات التنفيذ (القاهرة: 29 -30 ديسمبر 2021). والتي نظمت بكلية الآداب – جامعة القاهرة.**

8/1- الدراسات السابقة:

بالبحث في الإنتاج الفكرى من خلال بنك المعرفة المصري، والباحث العلمي من جوجل Google Scholar، تم رصد عدد من الدراسات التي تشير إلى تنامي الاهتمام بأسلوب تحليل المشاعر لقياس آراء المستفيدين بغرض تقييم فعالية ما يقدم لهم من خدمات أو منتجات بصورة كبيرة في مجال التسويق، وفي مراجعات الأفلام ومراجعات الكتب، ومن أمثلة ذلك دراسة شهد العشي وضحى البادي (2021)، كما تمت دراسات عدة في مجال الإعلام والاتصال الجماهيري (ريهام سامي، 2021)، والبحوث الإعلامية (طارق الخليفي، 2019)، ودراسة سبل توظيفها في تحليل مشاعر مستخدمى مواقع التواصل الاجتماعي في وقت أزمة كورونا (حمزة، 2021)، فضلًا عن الدراسات التي تمت بهدف مناقشة ورصد التقنيات والأدوات المستخدمة، والتحديات المتعلقة بالموضوع ومن بينها على سبيل المثال دراسة (Alessia et al, 2015) وغيرها الكثير من الدراسات في هذا الصدد، وفيما يلي عرض لنماذج من الدراسات التي تمت في مجال المكتبات، وتلك التي اهتمت بمعالجة مشكلات وقضايا تطبيق أسلوب تحليل المشاعر على المحتوى باللغة العربية :

الدراسات المتعلقة بتطبيق أسلوب تحليل المشاعر في مجال المكتبات:

وفيما يتعلق بمجال المكتبات، توجد دراسات أجنبية تناولت تطبيقه لقياس وتقييم آراء المستفيدين بشأن خدمات المكتبات، ولكنها ما زالت محدودة، فعلى سبيل المثال تم استخدام أسلوب تحليل المشاعر في تحليل مراجعات المستفيدين على تطبيق الخدمات المقدمة بالمكتبة الرقمية الوطنية للهند من خلال مجموعة بيانات بلغت 4560 تقييمًا للمستخدمين للتطبيق سواء من خلال نظام أندرويد أو نظام أي أو إس، وتم تحليل البيانات من خلال برنامج تحليلات .AppFollow. وأظهرت نتائج التحليل وجود 2130 تعليقًا تعطي مشاعر إيجابية و1808 تعليقات تعطي مشاعر سلبية تجاه التطبيق من خلال بيئة الأندرويد، بينما أظهرت وجود 6 مشاعر إيجابية و 22 مشاعر سلبية لنظام .iOS ، وتوصلت الدراسة من خلال مؤشرات التحليل بأن مستخدمي الأندرويد أكثر رضًا، مقارنة بمستخدمي iOS ورصدت عمليات التحليل أيضًا الشكاوى الأكثر شيوعًا التي ذكرها المستخدمون وهي الأخطاء الوظيفية، وتعطل التطبيق، والشكوى من الكتب التي يجب تنزيلها قبل القراءة بالإضافة لمشكلة وجود بعض ملفات PDF فارغة عند التنزيل، وأوصت الدراسة بمجموعة من التحسينات لمطورى الخدمة لجعل التطبيق أكثر سهولة للحصول على رضاء المستفيدين (Sharma, Gulati, Kaur, & Chakravarty, 2021)

وفي سياق تطبيق أسلوب تحليل المشاعر على استبيانات استقصاء الرأي، وتأتي دراسة مور (2017) التي هدفت إلى تطوير نموذج لتحليل الآراء ليتم تطبيقه لتحليل الأسئلة المفتوحة في أداة الاستبيان LibQual. واعتمدت الدراسة على مجموعة بيانات للتدريب مكونة من 514 تعليقًا ، تم اختيارها عشوائيًا من خمس استقصاءات للرأي من +LibQUAL وتمت مراجعة ردود الاستقصاء يدويًا وتم تصنيفها من حيث قطبية المشاعر إلى إيجابية أو سلبية، واستخدام منصة استخراج البيانات RapidMiner مفتوحة المصدر، للاستفادة من تلك التعليقات لإنشاء إطار عمل لإنشاء ملفات متجهات الكلمات الإيجابية والسلبية الخاصة بالمكتبة لتشغيل نموذج تحليل المشاعر يسهم في عمليات المقارنة التراكمية على مدار السنوات لمدى رضاء المستفيدين وفقًا للمحاور المتضمنة داخل الاستبيان، وأشارت النتائج كذلك إلى وجود توزيع متعادل فيما بين نسب التعليقات الإيجابية والسلبية فيما يتعلق بأسئلة القسم الخاص بخدمات المكتبات، وأظهرت النتائج توافقًا بين مؤشرات تحليل القطبية الخاصة بكل من الخدمات وإتاحة المعلومات والمبنى من جانب وبين نتائج التحليل الكمي لأسئلة الاستبيان المغلقة لذات المحاور.

كما صدرت دراسة أخرى لبحث تطبيق أسلوب تحليل المشاعر على تعليقات المستفيدين الواردة بالمسح السنوى الذي تجريه مكتبة جامعة هيلينك المفتوحة لتقييم الخدمات، وتقترح استخدام تحليل المشاعر على تعليقات المستخدمين حول المكتبة ومركز المعلومات، وقد سلطت نتائج التحليل الضوء على المعلومات الكامنة حول جوانب محددة من المكتبة كان لا يمكن التوصل إليها بطرق التحليل التقليدية (Apachristopoulos, 2019).

دراسة أخرى تمت في محاولة لتحليل المشاعر ومعدل تكرار الكلمات من واقع تغريدات المستفيدين لعدد 18 مكتبة هندية (5 مكتبات أكاديمية و 5 مكتبات حكومية و 5 مكتبات مدرسية و 3 مكتبات عامة)، وتم تجميع التغريدات المنشورة خلال شهري يونيو ويوليو 2019، وتحليل البيانات باستخدام برنامج "R" للتحليل الإحصائي، وتوصلت نتائج عملية تحليل المشاعر بأن المكتبات الهندية أقل نشاطًا على منصة تويتر،وبتحليل سحابة الكلمات التي تم توليدها آليًا من واقع الكلمات الأكثر تكرارا والواردة بالتغريدات، لوحظ وجود اختلاف في التغريدات وفقًا لنوع المكتبات، كما أظهر تحليل المشاعر للتغريدات أن معظم تغريدات المكتبات إيجابية بطبيعتها، وقد أوصت الدراسة بضرورة استخدام المكتبات الهندية لمنصة تويتر بصورة أكثر فعالية للترويج لمجموعاتها وخدماتها (Patra, 2019).

وفي سياق مجال خدمات المعلومات أيضًا، أجريت دراسة تجريبية لتطبيق أسلوب تحليل المشاعر لتقديم خدمة جديدة لمستخدمي المكتبات من خلال تحديد الأماكن المرتبطة بتغريدات ما، وتم استرداد البيانات من تويتر باستخدام 20 استفسارًا مختلفًا متعلقًا بالإنتاجية باستخدام منصة RapidMiner وتم إجراء تحليل المشاعر باستخدام برنامج AYLIEN لتحليل النص، واستخراج ما مجموعه 6416 تغريدة من تويتر لمدة 13 يومًا، علاوة على ذلك ، تم تحديد 676 علامة تصنيف بارزة، والعثور على 83 علامة تصنيف مرتبطة بموقع جغرافي في التغريدات،ولوحظ أن المملكة المتحدة كانت الدولة الأكثر شعبية والتي تم استخدامها كعلامة تصنيف على تويتر، تليها الهند والولايات المتحدة والصين ونيجيريا، وفيما يتعلق بالقطبية ، وجد أن معظم التغريدات ذات قطبية محايدة وأن معظم التغريدات الإيجابية كانت ذات قيمة منخفضة، يمكن إعادة تجميع المعلومات التي تم تحليلها باستخدام هذه الإستراتيجية كخدمة مدمجة قائمة على الوقت ويمكن تقديمها لمستخدمي المكتبات بتنسيقات مختلفة، لا تقدم هذه الدراسة طريقة حديثة لتلبية احتياجات المعلومات لمستخدمي اليوم فحسب ، بل تقترح أيضًا طريقة جديدة لإجراء التسويق في المكتبات باستخدام التنقيب عن وسائل التواصل الاجتماعي وتحليل المشاعر (Lamban & Madhusudhan, 2018).

الدراسات المتعلقة بمشكلات التعامل مع اللغة العربية:

تناولت بعض الدراسات مشكلات التعامل مع اللغة العربية وكيفية بناء موارد وتجميعات بيانات أو حلول معجمية تساعد في عمليات التحليل الآلي للنص العربي، ومن بين تلك الدراسات، الدراسة التي تناولت استخدام تحليل المشاعر والآراء لتقييم تغريدات المستخدمين باللغة العربية على القضايا المتعلقة بالأخبار والمنتجات والخدمات وما إلى ذلك، ورصدت الدراسة مجموعة من التحديات الناتجة عن استخدام اللغة العربية والعدد المحدود للمنشورات البحثية ومجموعات البيانات التي تم جمعها وتحليلها لهذا الغرض، وقام الباحثون في هذه الدراسة باتباع النهج المعتمد على البناء المعجمي، فقاموا ببناء معجم للمشاعر وأداة مبنية على هذا المعجم، وأشارت النتائج أن الأداة المقترحة تعمل بشكل جيد (Al-Ayyoub, Essa & Alsmadi, 2015).

كما قامت دراسة حبيبي والجيهاد (2020) ببناء موارد معجمية أساسية يمكن استخدامها كأداة لتحليل المشاعر في التعليقات الواردة باللغة العربية لتكون نواة تستخدم في بناء البرمجيات والتطبيقات الخاصة بتحليل والمشاعر، وذلك لما وجدوه من قصور في هذا

الصدد، وقام الباحثان ببناء مدونة عربية تشتمل على ما يزيد عن 100000 تعليق مكتوب باللغة العربية الفصحى تم تحميلها من مواقع مختلفة، وقام فريق العمل بوسم ما يقارب 25 ٪من التعليقات يدويًا.

وأوصت دراسة أخرى باتباع النهج الهجين الذي يجمع بين التوجيه الدلالي باستخدام طريقة بناء المعجم وتقنيات التعلم الآلي، ومن خلال هذا النهج ، قام المصنف المعجمي بتسمية بيانات التدريب، وهي مهمة تستغرق وقتًا طويلًا يتم إعدادها يدويًا في كثير من الأحيان، وتم استخدام ناتج المصنف المعجمي كبيانات تدريب لمصنف التعلم الآلي بتطبيق خوارزمية آلة التوجيه الداعمة SVM، وأظهرت نتائج التجربة أن النهج الهجين الذي تم اتباعه قد حسّن من مقياس F للمصنف المعجمي وكذلك معامل الدقة في النتائج (Aldayel, & Azmi, 2016).

يتضح من عرض الدراسات السابقة في الموضوع قلة الدراسات التي تمت باللغة العربية، مع التركيز على مجالات التسويق، والدراسات الإعلامية، والناحية التقنية من حيث مشكلات التعامل مع اللغة العربية وبناء الموارد التي تدعم اللغة العربية، ولم يتم رصد دراسات مخصصة تتناول أسلوب تحليل المشاعر في المكتبات باللغة العربية أثناء عملية البحث في الإنتاج الفكرى، وفيما يتعلق بالإنتاج الفكرى باللغة الإنجليزية كان هناك ثراء في الموضوع مقارنة بما نشر باللغة العربية، وتم رصد مجموعة من الدراسات التي تتناول سبل تطبيقه في مجال المكتبات.

ثانياً: الإطار النظرى للدراسة:

1/2- المقصود بتحليل المشاعر: Sentiment Analysis

يعد أسلوب تحليل المشاعر مجالًا فرعيًا من فروع التنقيب عن البيانات واللغويات الحاسوبية، ويعنى بالتقنيات الحاسوبية التي تقوم باستخراج، تصنيف، فهم ومقارنة الآراء التي يتم التعبير عنها في المصادر الإخبارية الإلكترونية، والتعليقات المنشورة عبر وسائل التواصل الاجتماعي وغيرها من المحتوى الذي يتم إنشاؤه بواسطة المستخدم (Zimbra, D, 2010 & Chen)، ويعرف كذلك بأنه العملية التي تتم آليًا لتحديد وتصنيف الآراء التي يتم التعبير عنها في جزء من النص، وبصفة خاصة لتحديد ما إذا كان موقف كاتب النص تجاه موضوع ما أو منتج معين أو ما إلى ذلك إيجابيًا أم سلبيًا أم محايدًا (Lexico, 2022). كما يعرف أيضًا من منظور مجال التسويق بأنه التصنيف المعتمد على الخوارزميات لتقييمات المستفيدين للعلامات التجارية إلى تقييمات إيجابية، وسلبية ومحايدة والتي ترد في تعليقاتهم ومنشوراتهم عبر وسائل التواصل الاجتماعي (Oxford Reference, 2022).

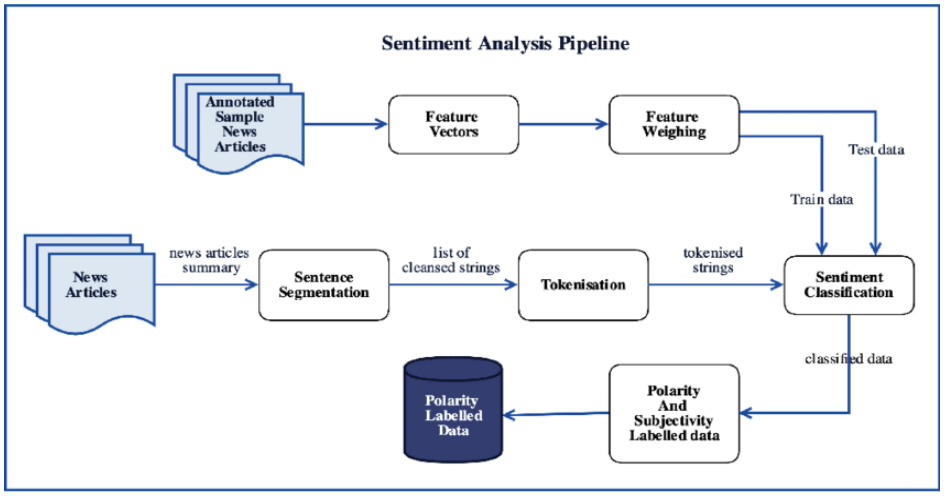
ويقدم الخليفى (2019) عرضًا مفصلًا للتعريفات المختلفة للمصطلح وعلاقته بالمصطلحات الأخرى ذات العلاقة ومن أهمها: مصطلح التنقيب عن البيانات، وبصفة عامة يشير مصطلح تحليل المشاعر القائم على الذكاء الاصطناعي إلى العملية التي يتم بمقتضاها تجميع الآراء النصية للتعرف على درجة قطبيتها من خلال تصنيفها إلى آراء إيجابية وسلبية ومحايدة، كما يمكن تصنيف الآراء الواردة بالنص لتعبر عن عواطف أو مشاعر أكثر تقدمًا مثل: الفرح و الحزن والغضب وذلك بالاستعانة بتقنيات وأدوات معالجة اللغة الطبيعية.

2/2- خطوات تحليل المشاعر:

تمر عملية تحليل المشاعر بسلسلة من الخطوات المتسلسلة والتي تعرف بـ sentiment analysis pipline وتبدأ بعملية تجميع البيانات والتي تعد أحد الخطوات المهمة في تلك العملية، وتوجد العديد من التقنيات والأدوات التي يمكن أن تستخدم لتجميع

البيانات، ويعتمد بعضها على طبيعة المصدر الذي يتم جلب أو تجميع البيانات منها، وتساعد تلك الأدوات في سرعة الحصول على البيانات بصورة أسرع وأدق وأشمل، مثل Twitter API على سبيل المثال.

وتأتي بعد ذلك مرحلة ما قبل المعالجة preprocessing أو ما يعرف بتجهيز البيانات للمعالجة data preparation حيث يمر النص بسلسلة من الخطوات لتنقيته وتنظيفه من أية تمثيلات أو كلمات أو علامات لا حاجة لها أثناء المعالجة والتحليل وذلك بهدف زيادة نسبة الدقة في أداء النموذج، ويشمل ذلك قائمة الكلمات الموقوفة Stop words. ثم تأتي عمليات تصنيف البيانات وتمثيلها باستخدام الخوارزميات المناسبة، ويلخص (Mandalapu,, et al, 2019) سلسلة خطوات تحليل المشاعر في الشكل التالي:



شكل رقم (1) خطوات تحليل المشاعر باستخدام نهج تعلم الآلة، نقلًا عن (Mandalapu, et al, 2019)

بالإشارة إلى الشكل رقم (1) والذي يلخص سلسلة الخطوات الرئيسية التي يتم من خلالها بناء نموذج تحليل المشاعر، يتضح أن هناك مجموعة من التقنيات والأدوات التي ينبغي توافرها للقيام بتلك العملية، والتي تمكن من استيراد مكتبات الأكواد، واستيراد تجميعة البيانات dataset، وتنظيف وتنقية البيانات data cleaning، وتحويل البيانات data transformation، ثم تقسيم تجميعة البيانات لقسمين splitting data (مجموعة تدريب training set ومجموعة اختبار testing set)، ثم تهيئة النموذج model fitting باستخدام أحد الخوارزميات مثل: خوارزمية نايف بايز على سبيل المثال، ولابد أن تسبق تلك الخطوات بفهم جيد وواضح للمشكلة والتي تدفع لبناء النموذج لتحليل المشاعر وتزويد المستخدم بالنتيجة(Ahmad, etal…, 2020) .

3/2- أبرز التقنيات والأدوات التي تتطلبها عمليات التحليل الآلي للمشاعر (اتجاهات التطبيق):

تعد منصات التواصل الاجتماعي والمدونات وغيرها من تطبيقات الجيل الثاني للويب (تلك التطبيقات التي تسمح للمستخدم أن ينشئ المحتوى، أو أن يشارك الرأي من خلال التعليقات والتقييمات وغيرها) من المصادر الأساسية للبيانات التي ينشئها المستخدم، وتحوى كما هائلا من البيانات الناتجة عن نشاط استخدام تلك المنصات، حيث يقوم المستخدمون من جميع أنحاء العالم بالتعبير عن آرائهم حول مجموعة متنوعة من الموضوعات، وتعد عمليات تحليل هذا الكم الهائل من البيانات الضخمة يدويًا أمرا مستحيلا، وبالتالي كان لابد من وجود تقنيات فعالة وذكية يمكنها التحليل الآلى وتحديد قطبية هذه البيانات النصية واستخراج الملامح والخصائص الرئيسية

للنص، وتوجد مجموعة متعددة من الأدوات والتقنيات المتاحة حاليًا لتصنيف البيانات التي تم إنشاؤها بواسطة المستخدم (Ahmad, Aftab, Muhammad, & Ahmad, 2017).

ويوجد ثلاث تقنيات مستخدمة لهذا الغرض:

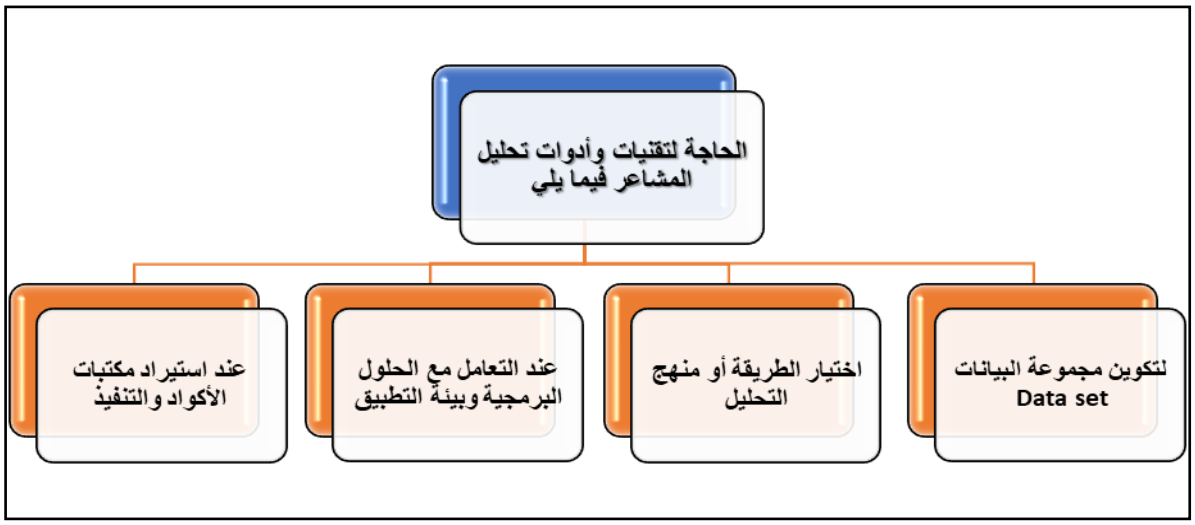
- التقنيات القائمة على المعجم Lexicon-based approach،
- التقنيات القائمة على تعلم الآلة Machine learning approach؛
- والتقنيات الهجينة (الأسلوب الذي يجمع بين المعجم القائم على التعلم الآلي) Hypered Approach.

وتحليل المشاعر يمكن تنفيذه باتباع نهج تعلم الآلة بنوعيها التعليم الموجه supervised و التعليم غير الموجه unsupervised learning، وكذلك شبه الموجه semi supervised أو بالطريقة المبنية على القواعد Lexicon - based والتي تعتمد على وضع قاموس بالكلمات تقوم الآلة بتصنيف النص بناء عليه، أو باتباع النهج المختلط، وذلك على النحو التالي (الخليفي، 2019):

- يعتمد الأسلوب الذي يتبع نهج تعلم الآلة على تكوين مجموعة من البيانات لأغراض تعليم الآلة training data حيث يتم عمل نموذج تدريب من خلالها للاستفادة منه في النموذج الذي يتم إدخال البيانات الجديدة فيه، ويمكن الاستفادة من أحد النماذج الجاهزة والمبنية وفقًا للخوارزميات المستخدمة في هذا الصدد مثل خوارزمية بايز Naïve bays – أو خوارزمية آلة المتجهات الداعمة SVM KNN، ويمكن تصنيف البيانات الكبيرة من خلال التعلم باستخدام البيانات الموسمة يدويًا عن طريق خبراء في المجال الموضوعى من البشر (supervised machine learning)، وبصفة عامة لكل نهج العديد من المتغيرات والتقنيات (Alessia et al, 2015).
- وفيما يتعلق بالنهج القائم على المعجم Lexicon-based فإنه يعتمد على واحدة من طريقتين: إما الاعتماد على القواميس التي تشتمل على مجموعة من المصطلحات التي تساعد في استنتاج المشاعر الموجودة داخل النص المراد تحليله، أو الاعتماد على موارد تشتمل على مجموعة من الكلمات المعدة سلفًا والتي تحتوى على بيان قطبية الكلمة وما إذا كانت ذات دلالة إيجابية (مثل كلمات: سعيد – جيد – عظيم) أو أن تكون ذات دلالة سلبية (مثل كلمات: حزين – سيئ – ضعيف) أو أن تكون محايدة (مثل الكلمات والجمل الاستفهامية التي يراد بها توجيه الأسئلة، أو الاستفسارات بصورة عامة).
- ويمزج النهج المختلط بين أساليب تعلم الآلة والنهج القائم على المعجم في محاولة للاستفادة من خصائص ومزايا كل أسلوب منهما .

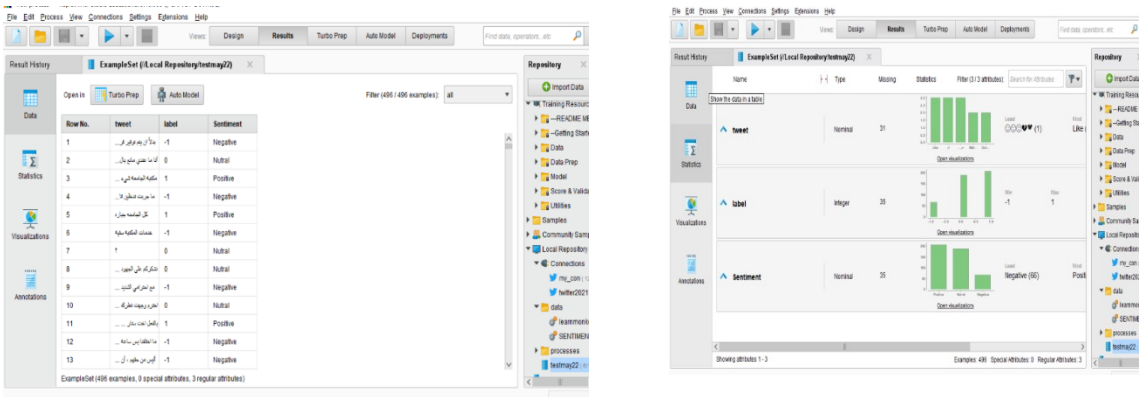
ويشير الشكل رقم (2) إلى المراحل المختلفة التي تتطلب استخدام تقنيات تحليل المشاعر، وفيما يلي عرض لأهم تلك التقنيات التي تم التعامل معها لأغراض استكشاف الطرق والوسائل التي يمكن تطبيقها في مجال قياس آراء المستفيدين من خدمات المعلومات:

- منصات العمل التي تعطي ميزة الحصول على بيانات الاعتمادية Credential ومنها IBM Cloud على سبيل المثال والتي تمكن من اختبار البيانات داخل بعض برمجيات التطبيق مثل Postman وغيرها.
- الحصول على حسابات مطورين يتم طلبها من إدارات مواقع التواصل الاجتماعي وأشهرها Twitter developer account في حالة الرغبة في جلب التغريدات أو التعليقات بصورة آلية (من خلال واجهة برمجة التطبيق API)



شكل رقم (2) الخطوات التي تتطلب استخدام تقنيات وأدوات تحليل المشاعر

- الاستفادة من المنصات التي تتيح الموارد وتجميعات البيانات datasets مثل GitHub على سبيل المثال (إلا أنه توجد صعوبة للحصول على تجميعة بيانات تتعلق بمستفيدي المكتبات باللغة العربية).
- الاستفادة من المنصات مفتوحة المصدر والتي تمكن من تنفيذ عمليات اختبار البيانات أو الحصول على تحليلات المشاعر (منها ما يعتمد على وجود معرفة متقدمة بالبرمجة – ومنها ما يقدم المساعدة لمن لديه معرفة مبدئية مثل **Google Colab** – ومنها ما يقدم التحليل للنص مباشرة إلا أنها تقدم معلومات مختصرة وتواجه إشكالات في تصنيف قطبية النص العربي) .
- كما يمكن استخدام برنامج orange أو SentiStrength أو Microsoft azure وRapidMiner Studio وغيرها، بالإضافة إلى التطبيقات المتاحة عبر الويب مثل Monkey learn و Kaggle، و Peltarion والذي يمكن تكامله مع تطبيق Bubble ، ويعرض الشكل رقم (3) لأحد تلك البرمجيات التي تم استخدامها لأغراض الدراسة الاستكشافية، وهو برنامج :RapidMiner

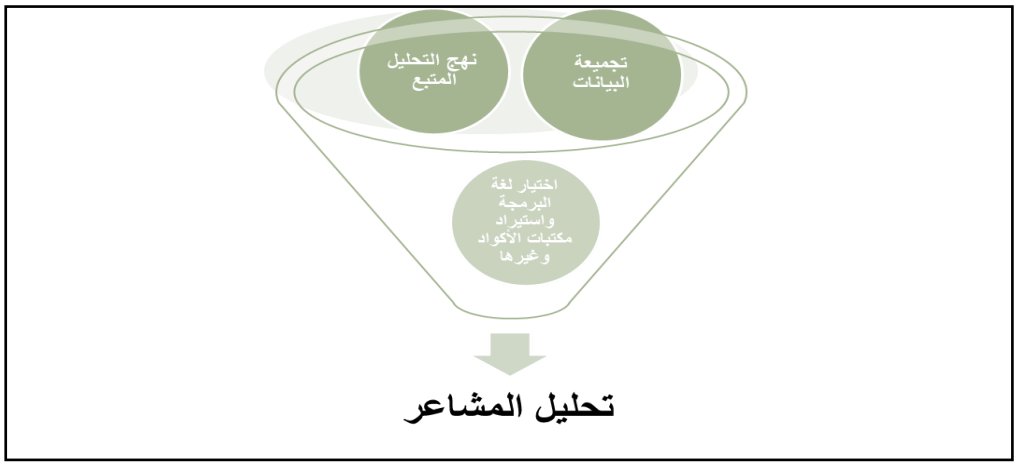


شكل رقم (3) نموذج لتحليل عينة البيانات باستخدام برنامج RapidMiner على سبيل المثال

- الاستفادة من مفكرات البرمجة مثل Jupyter notebook والمدعمة في كثير من منصات العمل مثل IBM cloud وGoogle colab وKaggle وغيرها، حيث تمكن من استيراد مكتبات الأكواد البرمجية وبخاصة بلغة البايثون مما يسهل الكثير من الخطوات ومن أمثلتها: tweepy التي تسمح بجلب تغريدات تويتر– nltk وهي مجموعة أدوات لمعالجة اللغة الطبيعية، و pandas التي تسمح بعمليات بمعالجة وتحليل البيانات و matplotlib و wordcloud التي تسمح بإنشاء سحابة الكلمات وغيرها من المكتبات.

ويلخص الشكل رقم (4) الأركان الرئيسية التي تتفاعل معًا لبناء نموذج تحليل المشاعر وهي:

- مجموعة البيانات dataset
- اختيار نهج التحليل.
- اختيار لغة البرمجة واستيراد مكتبات الأكواد ذات العلاقة.



شكل رقم (4) التفاعل ما بين العناصر الأساسية لأسلوب تحليل المشاعر

ثالثاً: الإطار التطبيقي:

اعتمدت الدراسة الاستكشافية في جانبها التطبيقي على نهج تعلم الآلة لما تم رصده في مراجعة الإنتاج الفكري من حيث فعاليته وموثوقيته للتنقيب عن الرأي وتصنيف المشاعر، وقد تم الحصول على حساب لإتاحة واجهة تطبيقات تويتر Twitter API من خلال التسجيل عبر منصة مطوري تويتر، ووفقًا لذلك تم الحصول على مفتاح واجهة التطبيق API Key ورمز الوصول Access Token وباقي بيانات التحقق من الهوية التي تمكن من جلب التغريدات لأغراض البحث العلمي، وللحفاظ على خصوصية البيانات سوف يتم عرض التعليقات والتغريدات دون الإشارة إلى اسم المكتبة أو اسم المستفيد وفي إطار نماذج توضيحية فقط:

وتضمن الإطار العام لبناء النموذج ثلاث خطوات رئيسية: تجميع وتحميل بيانات تدريب النموذج Data Collection – معالجة النص text preprocessing – بناء النموذج model. وتمثلت إجراءات الدراسة فيما يلي:

1/3- تجميع البيانات: Data Collection

- لتصميم النموذج لابد من وجود مجموعة بيانات dataset مصنفة إلى إيجابية وسلبية ومحايدة، ومن الممكن الرجوع إلى مجموعة من المصادر التي تتيح البيانات الجاهزة مثل GitHub و Kaggle وغيرها، أو أن يتم تجميعها، ولأغراض الدراسة تم تجميع التعليقات والتغريدات يدويًا نتيجة لعدم وجود مجموعات بيانات جاهزة في مجال المكتبات باللغة العربية، ولأنه من بين الأهداف الأساسية استكشاف طبيعة محتوى التعليقات من حيث دلالتها في سياق خدمات المكتبات والمعلومات.
- تم اختيار لغة البرمجة بايثون python وتهيئة واستخدام مفكرة جوبيتر Jupyter Notebook واستيراد مكتبات الأكواد اللازمة مع الأخذ في الاعتبار مكتبات الأكواد البرمجية التي تتعامل مع النص العربي (تم الاستقرار على اختيار منصة Google Colab في ضوء ما أسفرت عنه نتائج الدراسة الاستطلاعية نتيجة لمرونتها).
- تم تجميع البيانات من خلال البحث بالتوسيمات (الهاشتاج) والكلمات المفتاحية (16) لمجموعة من المصطلحات التي تظهر عبر منشورات مؤسسات المعلومات عبر منصتي تويتر وفيس بوك والتي تم ملاحظتها أثناء الدراسة الاستطلاعية، وتحديد تلك الكلمات المفتاحية في ضوء ما تم رصده خلال الدراسة الاستطلاعية لعينة مبدئية من التعليقات، ويقدم الجدول رقم (1) بيانًا مفصلًا بالكلمات المفتاحية التي تم استخدامها عبر منصتي تويتر وفيس بوك.

جدول رقم (1) بيان بالتوسيمات والكلمات المفتاحية التي تم استخدامها عبر منصتي تويتر وفيس بوك لتجميع التعليقات والتغريدات

الكلمات المفتاحية الأساسية	عدد التغريدات المسترجعة	فيس بوك	الإجمالي
بنك المعرفة المصري	28	73	101
فعاليات المكتبة	11	78	89
المكتبة الرقمية*	55	20	75
مركز مصادر التعلم	30	40	70
ورش عمل المكتبة (ورش العمل & المكتبة)	0	53	53
ساعات العمل & المكتبة	3	49	52
خدمات & المكتبة	17	30	47
#خدمات_المكتبة*	22	18	40
#خدمات_المعلومات & مكتبة	3	23	26
المصادر الإلكترونية*	3	20	23
الفهرس الإلكتروني للمكتبة	0	21	21
فهرس المكتبة*	3	16	19
المكتبة الرئيسية	10	5	15
#خدمات_المعلومات*	4	10	14
المكتبة المركزية & جامعة	6	8	14
خدمة الاستعارة	0	4	4
خدمة الإعارة	0	2	2
خدمات الإعارة	0	0	0
المجموع	195	470	665

* المصطلح عام يتداخل فيه تغريدات/منشورات خارج نطاق تخصص المكتبات والمعلومات، فضلًا عن ظهوره على صفحات لمجموعات اهتمام أو أفراد وليست لمكتبات أو مؤسسات معلومات (فتم استبعادها)

- تم تجميع البيانات باستخدام طريقتين: اعتمدت الطريقة الأولى على الاستدعاء الآلي للتغريدات من موقع تويتر وفقًا لخصائص معينة مثل: التوسيم أو البحث بالكلمة المفتاحية وذلك لتوافر تلك الإمكانية بمنصة تويتر وفقًا لحساب المطورين، بالإضافة إلى التجميع اليدوي للتعليقات من موقع فيس بوك وتم دمج البيانات معًا (665) تغريدة وتعليق في مجموعة واحدة dataset وحفظها داخل ملف من نوع csv وتحميلها لبرنامج التطبيق Google colab.
- تضمنت العينة مجموعة من الرسوم التعبيرية Emojis التي وردت كجزء من التغريدة في حالة وجودها، مع مراعاة التعبير عنها في صورة نص (smiling face – care – angry, ... etc.).
- تم تصنيف مجموعة البيانات التي تم تجميعها من حيث القطبية إلى إيجابية – سلبية – محايدة وذلك يدويًا لأن هناك من الدراسات ما تقترح إمكانية تصنيف بيانات التدريب يدويًا من خلال خبراء بشر وبخاصة في حالة مجموعات البيانات الصغيرة، وتم اعتبارها مجموعة التعلم training data مع تطبيق منهج تعليم الآلة machine learning، وتم تقسيم مجموعة البيانات إلى مجموعة تدريب وتمثل 80٪ ومجموعة اختبار وتمثل 20٪.

نتائج تحليل محتوى الغريدات والتعليقات:

هناك عدد من الإمكانات التي تقدمها منصة Google colab على سبيل المثال، والتي يمكن الاستفادة منها في التعرف على آراء المستفيدين، وطبيعة الموضوعات الواردة في التعليقات نتيجة عملية التفاعل مع منشورات المكتبة وغيرها من مؤسسات المعلومات، وفيما يلي عرض لأبرز تلك الإمكانات:

استعراض التغريدات:

1 to 5 of 5 entries Filter

index	id	tweet	label	Sentiment
0	1	RT @EKB_EG: إلحق متحمسون للعودة إلى ورش العمل وجهًا لوجه	1	Positive
1	2	47) RT @Tahanialdo: ♦ المكتبة الرقمية @EdTech_Books مصدر مجاني وأكثر من رائع لمجموعة مهمة من كتب التعليم الإلكتروني عن بعد و المدمج والبحث العلم...	1	Positive
2	3	للوصول إلى Reaxys من Elsevier من بنك المعرفة المصري https://t.co/2ltEaeLMSW https://t.co/UFK1oBenC8	0	Nutral
3	4	@KFNLGOV @Naseej المكتبة تقدم خدماتك رائعة ..شكرا لكل القائمين عليها @ibnm1980 @H1261980	1	Positive
4	5	#كتابة_خطة_البحث	0	Nutral

Show 25 per page

Like what you see? Visit the data table notebook to learn more about interactive tables.

شكل رقم (5) إمكانية عرض أول 5 تعليقات

تعد إمكانات عرض البيانات النصية من تعليقات وغيرها من أبرز مميزات منصات برمجيات تحليل المشاعر، ويوضح الشكل رقم (5) أحد الطرق، وهي عرض أول 5 تعليقات (سواء أكانت تم جلبها من الحساب الخاص بتويتر مباشرة، أو من خلال ملف بصيغة csv أو txt).

الحصول على بيانات إحصائية:

يمكن للمكتبة أن تحصل على مجموعة من البيانات التفصيلية كما هو موضح بالشكل رقم (6) مثل: إجمالي عدد التغريدات، وعدد التمثيلات الخاصة بكل تغريدة، مع إمكانية استعراض أطول تغريدة وأقصر تغريدة وهكذا:

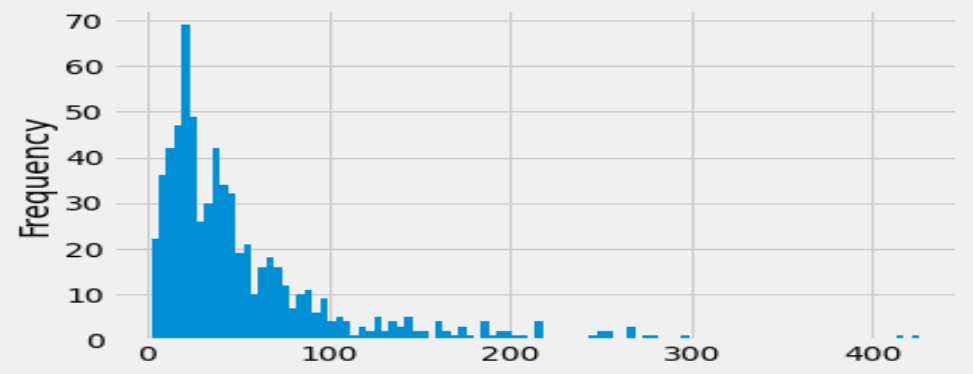
```
[276] # data type info
      df.info()

      <class 'pandas.core.frame.DataFrame'>
      RangeIndex: 665 entries, 0 to 664
      Data columns (total 4 columns):
       #   Column     Non-Null Count  Dtype
      ---  ------     --------------  -----
       0   id         665 non-null    int64
       1   tweet      665 non-null    object
       2   label      665 non-null    int64
       3   Sentiment  665 non-null    object
      dtypes: int64(2), object(2)
      memory usage: 20.9+ KB
```

	label	length
count	665.000000	665.000000
mean	0.284211	53.168421
std	0.737633	54.144067
min	-1.000000	2.000000
25%	0.000000	20.000000
50%	0.000000	37.000000
75%	1.000000	66.000000
max	1.000000	426.000000

شكل رقم (6) إمكانية عرض مجموعة من البيانات الإحصائية عن التعليقات التي تمت معالجتها بعينة الدراسة

ويوضح الشكل رقم (7) التوزيع البياني للتعليقات والتغريدات وفقًا لمعدلات تكرارها حسب طولها.



شكل رقم (7) إمكانية عرض مجموعة من البيانات الإحصائية عن التعليقات التي تمت معالجتها بعينة الدراسة

وتعد إمكانات استعراض التعليقات سواء السلبية أو الإيجابية من الأمور المهمة للمكتبة والتي تساعدها في الوصول إلى المزيد من الفهم لاحتياجات المستفيدين، وبخاصة فيما يتعلق بالتعليقات السلبية، يليها التعليقات المحايدة، بينما تظهر التعليقات الإيجابية كلمات للمدح والثناء أو توجيه الشكر بصورة عامة، وهو ما يوضحه الشكل رقم (8):

شكل رقم (8) إمكانية استعراض نصوص التعليقات وفقًا للمتوسطات الحسابية لأطوالها

كما يمكن للمكتبة أن تستعرض التغريدات كاملة، وعلى الرغم من وجود التغريدات أو التعليقات على الصفحة الرسمية للمكتبة، إلا أن التعامل معها بصور منفردة أمر غاية في الصعوبة والمشقة، بينما يتم عرض التعليقات من خلال تطبيقات تحليل المشاعر في صورة مجمعة وبأكثر من طريقة عرض، فضلًا عن إمكانات عرض التوسيمات سواء السلبية أو الإيجابية أو المحايدة أو دون تحديد، ويوضح الشكل رقم (9) أحد طرق العرض.

شكل رقم (9) إمكانية عرض مجموعة البيانات كاملة

إمكانية عرض التعليقات وفقًا لدرجة الاستقطاب (إيجابي أو سلبي):

يمكن للمكتبة كذلك استعراض التعليقات مصنفة وفقًا لمستوى قطبيتها: (التعليقات الإيجابية – التعليقات السلبية – التعليقات المحايدة)، مما يساعد المكتبات في الوقوف على نقاط الضعف، أو ربما تكون المشكلة في عدم معرفة المستفيد بتوافر خدمات معينة، أو أن المكتبة اتخذت إجراءات أو قرارات بحاجة لأن تقوم بالتواصل الفعال مع المستفيدين لتوضيح تلك الأمور بما يساعد من دعم المستفيد للمكتبة وزيادة مستوى رضائه عن خدماتها، ويعرض الشكل رقم (10) لنماذج لاستعراض التعليقات وفقًا لمستوى قطبيتها:

positive

	tweet	label	length
0	RT @EKB_EG: نحن منحسون للعوده إلى ورش العمل و...	1	56
1	RT @Tahanialdo: 📌المكتبة الرقمية @EdTech_B... 47)	1	144
3	@KFNLGOV @Naseej ..... المكتبة تقدم خدمات رائعة	1	94
8	...مكتبة الجامعة شيء جبار اعتقد هذي اقدم واكبر ملء	1	57
10	كل الجامعه جباره	1	16
...	...	...	...
646	...بالتوفيق لهذه الجهود الجبارة في خدمة العلم وال	1	53
648	اشتقنا لكم 😍💔	1	13
650	ما شاء الله عليكم 💐👍	1	20
660	يعطيكم العافية 👍👍	1	17
663	تم الرد من قبل الموظف وخدمتي وجزاكم الله الجنة	1	46

302 rows × 3 columns

negative

	tweet	label	Sentiment	length
5	... !!! ليش معلقة من ٥ لحد ٩ والطلبات كيف بدالكون	-1	Negative	77
6	...بدأ ان يتم توفير فرص وظيفية مالتمه لهذه الفئ	-1	Negative	95
9	...ما جربت تدخلين قاعات القراءة ما تقدرين تقرين له	-1	Negative	96
11	خدمات المكتبة سيئة	-1	Negative	18
14	... اذكركم ان ما يقلون Seedling مع احترامي الشديد	-1	Negative	101
...	...	...	...	...
656	... نواجه صعوبه في توصيل الكتب اللي يستعرق ٤ ايام	-1	Negative	79
657	...طلبه صارلي اسبوعين متى يوصل الطلب مافي احد يد	-1	Negative	66
658	...شوفوا انا هنّ مع الموقع اول شي ما يكمل عمليه ال	-1	Negative	99
661	في عطله العيد المكتبة معلّقة ؟	-1	Negative	29
662	...لا أعلم متى يكون الموظف المسؤول عن الرسائل الج	-1	Negative	121

113 rows × 4 columns

nutral

	tweet	label	length
2	...من بنك المعرفة ا Elsevier من Reaxys الوصول إلى	0	101
4	كتابة_حلقة_البحث#	0	17
7	...أنا ما عندي مانع بالمشاركة في هذه الفعالية شري	0	67
12	؟	0	5
13	... نشكركم على الجهود المبذولة بس نتمنى انها تفتح	0	79
...	...	...	...
636	هل المشاركة مفتوحة للجميع ام فقط لطلبة الجامعة؟	0	47
647	المكتبة مفتوحه هالفتره؟	0	24
655	...زين ضبط معاي الموقع دفعت حق الكتب ومادفع حق تو	0	53
659	شلون اطلب الكتب اونلاين	0	23
664	هل ممكن تمديدها لمدة أسبوع آخر؟	0	31

250 rows × 3 columns

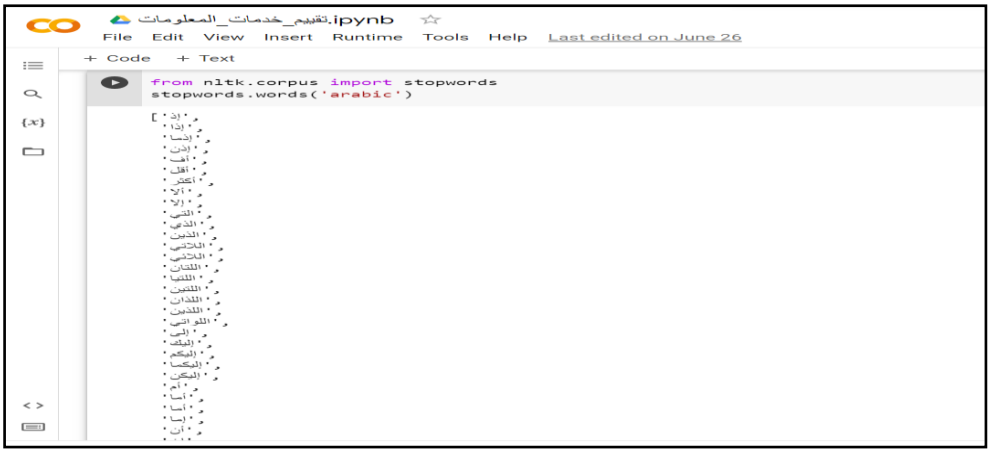
شكل رقم (10) إمكانية عرض التعليقات وفقًا لمستوى قطبيتها

2/3- تنظيف البيانات والمعالجة المسبقة: Data Preprocessing (cleaning)

تساعد عمليات تنظيف البيانات والمعالجة المسبقة في الحصول على مزيد من الدقة في بناء النموذج وقدرته على التنبؤ بقيم المدخلات الجديدة، ويعتبرالبعض تنظيف وتنقية البيانات في مشروعات معالجة اللغة الطبيعية نصف العمل، ويتم في هذه المرحلة استبعاد كلمات الوقف stop words والعلامات الخاصة، والقيام بعملية التجذيع stemming وغيرها من العمليات، وتم القيام بالعمليات التالية للمعالجة المسبقة لمجموعة بيانات الدراسة:

استبعاد كلمات الوقف :stop words

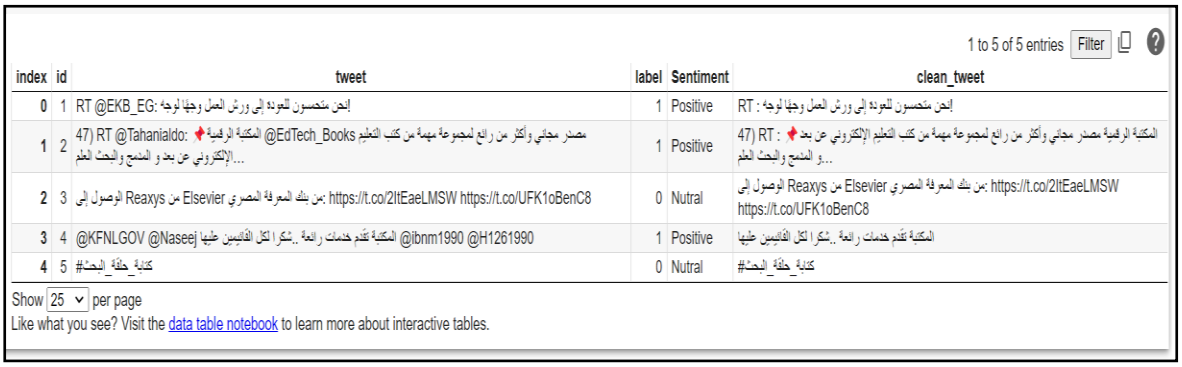
تم تنقية التعليقات باستبعاد كلمات الوقف من خلال مكتبة الأكواد البرمجية الخاصة بمعالجة اللغة العربية nltk من خلال استيراد stopwords كما هو موضح بالشكل رقم (10)



شكل رقم (11) تنقية البيانات باستبعاد كلمات الوقف

إزالة العلامة الخاصة (@) :

يشير الشكل رقم (12) إلى عملية تنقية البيانات بإزالة بعض العلامات الخاصة مثل: علامة البريد الإلكتروني @ كأحد الخطوات في عملية تنقية البيانات وذلك بهدف زيادة نسبة الدقة في النتائج.



1 to 5 of 5 entries Filter

index	id	tweet	label	Sentiment	clean_tweet
0	1	RT @EKB_EG: إنحن متحسون للعوده إلى ورش العمل وجهًا لوجه	1	Positive	RT : إنحن متحسون للعوده إلى ورش العمل وجهًا لوجه
1	2	47) RT @Tahanialdo: المكتبة الرقمية @EdTech_Books مصدر مجاني وأكثر من رائع لمجموعة مهمة من كتب التعليم الإلكتروني...	1	Positive	47) RT : المكتبة الرقمية مصدر مجاني وأكثر من رائع لمجموعة مهمة من كتب التعليم الإلكتروني عن بعد و المدمج والبحث العلم...
2	3	الوصول إلى Reaxys من Elsevier من بنك المعرفة المصري: https://t.co/2ItEaeLMSW https://t.co/UFK1oBenC8	0	Nutral	الوصول إلى Reaxys من Elsevier من بنك المعرفة المصري: https://t.co/2ItEaeLMSW https://t.co/UFK1oBenC8
3	4	@KFNLGOV @Naseej المكتبة تقدم خدمات رائعة ..شكرا لكل القائمين عليها @ibnm1990 @H1261990	1	Positive	المكتبة تقدم خدمات رائعة ..شكرا لكل القائمين عليها
4	5	#كتابة_حلقة_البحث	0	Nutral	#كتابة_حلقة_البحث

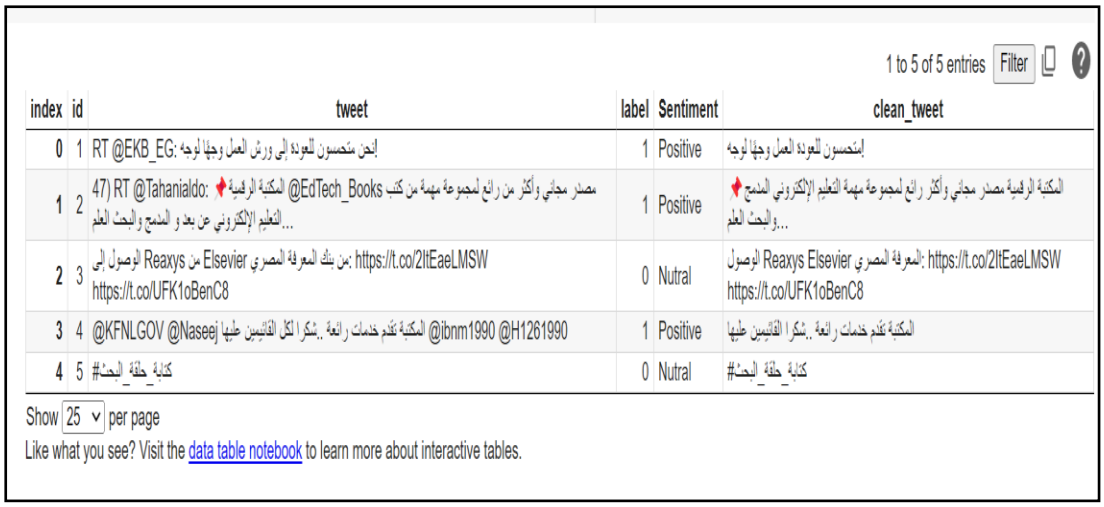
Show 25 per page

Like what you see? Visit the data table notebook to learn more about interactive tables.

شكل رقم (12) عرض التعليقات قبل وبعد عملية تنقية البيانات بإزالة بعض العلامات الخاصة بعينة الدراسة

إزالة إعادة التغريدات RT :

يوضح الشكل رقم (13) خطوة أخرى من خطوات تنقية التعليقات والتغريدات بالعينة، حيث تم إزالة



1 to 5 of 5 entries Filter

index	id	tweet	label	Sentiment	clean_tweet
0	1	RT @EKB_EG: إنحن متحسون للعوده إلى ورش العمل وجهًا لوجه	1	Positive	إنحن متحسون للعوده العمل وجهًا لوجه
1	2	47) RT @Tahanialdo: المكتبة الرقمية @EdTech_Books مصدر مجاني وأكثر من رائع لمجموعة مهمة من كتب التعليم الإلكتروني...	1	Positive	المكتبة الرقمية مصدر مجاني وأكثر رائع لمجموعة مهمة التعليم الإلكتروني المدمج والبحث العلم...
2	3	الوصول إلى Reaxys من Elsevier من بنك المعرفة المصري: https://t.co/2ItEaeLMSW https://t.co/UFK1oBenC8	0	Nutral	الوصول Reaxys Elsevier المعرفة المصري: https://t.co/2ItEaeLMSW https://t.co/UFK1oBenC8
3	4	@KFNLGOV @Naseej المكتبة تقدم خدمات رائعة ..شكرا لكل القائمين عليها @ibnm1990 @H1261990	1	Positive	المكتبة تقدم خدمات رائعة ..شكرا لكل القائمين عليها
4	5	#كتابة_حلقة_البحث	0	Nutral	#كتابة_حلقة_البحث

Show 25 per page

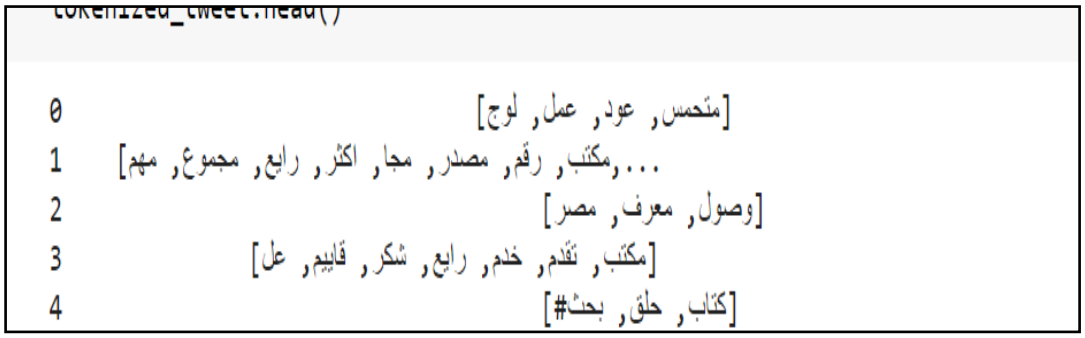
Like what you see? Visit the data table notebook to learn more about interactive tables.

شكل رقم (13) عرض التعليقات قبل وبعد عملية تنقية البيانات إزالة إعادة التغريدات RT بعينة الدراسة

إعادة التغريدات:

القيام بعمليات تجذيع الكلمات Stemming:

ويعد مجذع بورتر Porter Stemmer الأداة الأشهر في هذا الصدد ولكنها لا تتعامل مع اللغة العربية، ولذلك تم التعامل مع Snowball Stemmer، إلا أنه من الملاحظ كما هو موضح بالشكل رقم (14) أن تجذيع الكلمات عند القيام بتحليل المشاعر قد يفقد الكلمات دلالتها في التعبير عن القطبية، فضلًا عن أن فصل الجمل في صورة كلمات منفصلة قد لا يعبر بوضوح عن المعنى المقصود ومن الممكن أن يذهب بالمعنى من الإيجابية إلى السلبية والعكس صحيح.



شكل رقم (14) عرض التعليقات قبل وبعد عملية تجذيع الكلمات باستخدام Snowball Stemmer

إزالة العلامات الخاصة و URL و http:

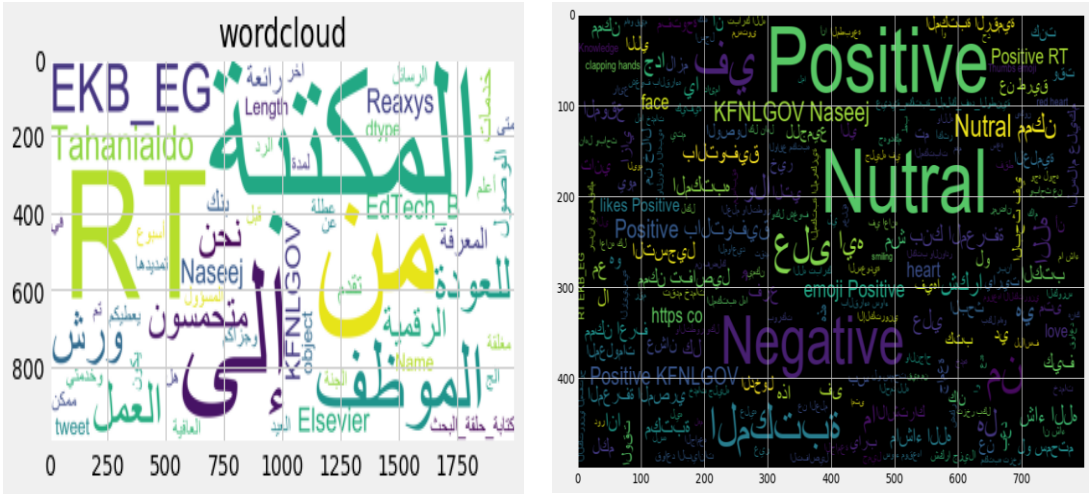
تم رصد مجموعة من العلامات الخاصة الأخرى مثل: URL وhttp وتم إجراء بعض العمليات لإزالتها، يوضح الشكل رقم (15) عينة التعليقات قبل وبعد إزالة هذه العلامات:

	id	tweet	label	Sentiment	clean_tweet
5	6	... !!! ليش مغلقة من ٥ لحد ٩ والطالبات كيف يذاكرن	-1	Negative	مغلقة والطالبات يذاكرن احنا الوقت مسموحلنا نطلع
6	7	...بدلاً ان يتم توفير فرص وظيفية ملائمه لهذه الفئ	-1	Negative	...بدلا توفير وظيفية ملائمه لهذه الفئه العزيزه ال
9	10	...ما جربت تدخلين قاعات القراءة ما تقدرين تقرين ك	-1	Negative	...جربت تدخلين قاعات القرا تقدرين تقرين كلمتين سو
11	12	خدمات المكتبة سئية	-1	Negative	خدمات المكتبة سئية
14	15	... اذكركم ان ما يقفون Seedling مع احترامي الشديد	-1	Negative	...احترامي الشديد اذكركم يقفون هناك أناس إنسانيته
17	18	...ما اختلفنا بس ساعة ما بتشتت العائلة كلن منا له	-1	Negative	اختلفنا ساعة بتشت العائلة أسره
18	19	...أليس من حقهم ، أن يكون لهم أوقاتهم التي يقضونه	-1	Negative	...أليس حقهم يكون أوقاتهم التي يقضونها أبنا وفلذا
24	25	...سؤال ظريف ... تعال واسال كل مستفيد يدخل المكتب	-1	Negative	... سؤال ظريف تعال واسال مستفيد يدخل المكتبه تقرا
25	26	ربما اجابتك أظرف Smiling face with open mouth	-1	Negative	ربما اجابتك أظرف
26	27	سيئة جدا	-1	Negative	سيئة

شكل رقم (15) عرض التعليقات قبل وبعد عملية إزالة العلامات الخاصة وURL وhttp بعينة الدراسة

استعراض سحابة الكلمات:

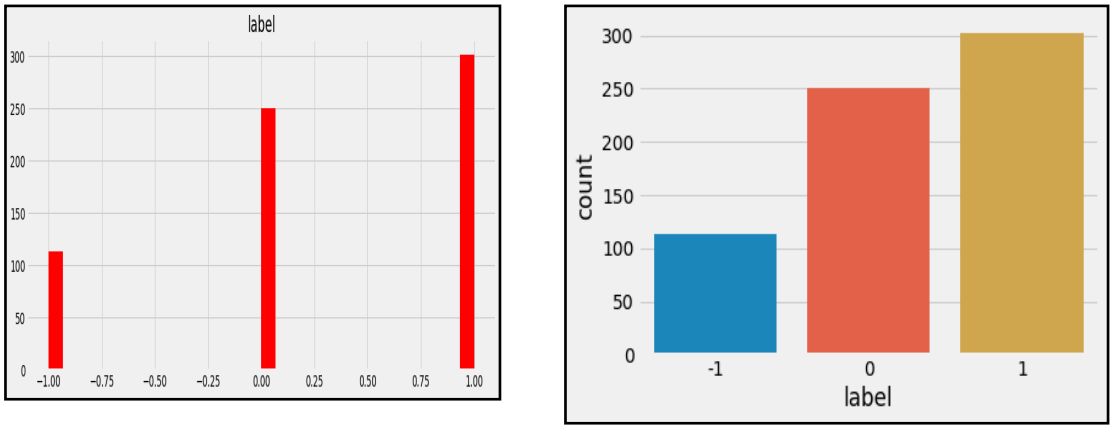
بعد انتهاء عمليات تنقية البيانات المحملة على ملف CSV يمكن استعراض سحابة الكلمات word cloud، وقد تقوم المكتبة بعرض سحابة الكلمات التي تعبر عن الشعور الإيجابي فقط، أو السلبي فقط، أو المحايد أو جميعها معًا، وتفيد سحابة الكلمات في التمثيل المرئي للموضوعات التي تم تناولها في التعليقات والتي تكرر ظهورها، ويحدد عدد مرات التكرار المعطيات التي يتم بناء سحابة الكلمات وفقًا لها، بالإضافة لحجم مجموعة البيانات، ولابد من تحميل مكتبات أكواد التعامل مع النص العربي حتى تتمكن سحابة الكلمات من إظهار الكلمات بشكل صحيح (مثل Arabic Reshaper وBidi-python)، ويوضح الشكل رقم (16) الفرق في عرض سحابة الكلمات قبل وبعد استخدام مكتبات الأكواد التي تدعم النص العربي مثل مكتبة Arabic Reshaper على سبيل المثال:



شكل رقم (16) الفرق في عرض سحابة الكلمات قبل وبعد استخدام مكتبات الأكواد التي تدعم النص العربي

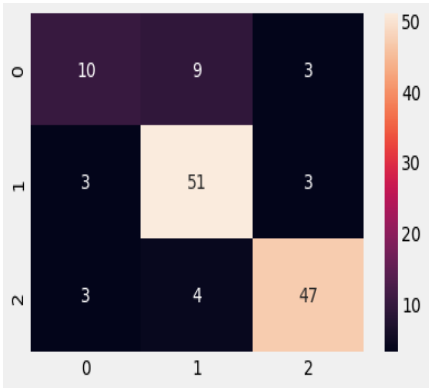
3/3- بناء النموذج:

- بعد دراسة خصائص التعليقات والتغريدات التي تم تجميعها بهدف التعرف على مدى دلالتها وإمكانيتها في دعم عمليات تقييم خدمات المعلومات، تم اختبار المجموعة التدريبية بتطبيق خوارزمية Naïve-Bays لتحليل مجموعة الاختبار test data والتصنيف القطبي لها إلى إيجابية، ومحايدة وسلبية، مع الأخذ في الاعتبار أن الهدف الأساسي ليس القيام باختبار نموذج لتحليل المشاعر، ولكن استكشاف إمكانات وفرص تطبيقه لاستخلاص مؤشرات تساعد في تقييم مستوى تقديم الخدمات وما يواجهها من مشكلات كأداة يمكن استخدامها بصورة منتظمة لإمداد متخذ القرار في المكتبات وغيرها من مؤسسات المعلومات.
- وتم الحصول على التوزيعات التكرارية لعدد التعليقات بالعينة وفقًا لمستوى القطبية : الإيجابية (1)، والسلبية (-1) والمحايدة (0) كما هو موضح بالشكل رقم (17):



شكل رقم (17) يوضح توزيع عدد التعليقات بالعينة وفقًا لمستوى القطبية : الإيجابية (1)، والسلبية (-1) والمحايدة (0)

- وكما يشير الشكل رقم (17)، تمثل التعليقات الإيجابية بالعينة نسبة 45٪ من إجمالي التعليقات، وتمثل التعليقات المحايدة نسبة 37.5% ، بينما تمثل التعليقات السلبية نسبة 17٪.
- وقد تمت عملية تدريب البيانات، واختبار النموذج باستخدام خوارزمية Naïve Bays، باستدعاء خاصية sklearn وبلغت قيمة F score 81٪ ، ويشير الشكل رقم (18) للبيانات التفصيلية لاختبار النموذج فيما يتعلق بالدقة ومعدلات الاستدعاء والتحقيق (ملاحظة: بلغت نسبة الدقة في عينة الدراسة الاستطلاعية 50٪ وذلك نتيجة لعدم إجراء عمليات تنقية وتنظيف البيانات عليها، فضلًا عن صغر حجمها نسبيًا مقارنة بحجم العينة الحالي).



	precision	recall	f1-score	support
-1	0.62	0.45	0.53	22
0	0.80	0.89	0.84	57
1	0.89	0.87	0.88	54
accuracy			0.81	133
macro avg	0.77	0.74	0.75	133
weighted avg	0.80	0.81	0.81	133

شكل رقم (17) يوضح توزيع عدد التعليقات بالعينة وفقًا لمستوى القطبية : الإيجابية (1)، والسلبية (-1) والمحايدة (0)

التحديات والصعوبات:

تم حصر مجموعة من التحديات، بعضها تحديات عامة والبعض الآخر تحديات ترتبط بمواقع المكتبات وذلك على النحو التالي:

- تحديات تتعلق بطبيعة اللغة العربية نفسها ولهجاتها عند معالجة اللغة الطبيعية.

- ما يتوافر من موارد أو تجميعات بيانات وأدوات ومكتبات برمجية تتعامل مع النص العربي بفاعلية وتوليد سحابة الكلمات.
- تحديات تتعلق بطبيعة التعليقات وغموض دلالتها، أو التعبير عن النص العربي بأبجدية لاتينية Arabizi or Franco-Arabic، ووجود بعض التعليقات التي تحمل السخرية.
- تحديات تتعلق بضعف المشاركة الفعالة للمستفيدين من خلال التعليقات حيث يغلب استخدام الرسوم التعبيرية، أو عدم وجود تعليقات، أو وجود تعليقات ليس لها علاقة بالمنشور.
- التحديات ذات العلاقة بقضايا الخصوصية، فعلى الرغم من وجود تعليقات المستفيدين وتغريداتهم بصورة مفتوحة على الصفحات العامة للمؤسسات، إلا أن الأفراد وحتى المكتبات وغيرها من المؤسسات، قد لا تفضل استخدام بياناتها وآرائها في عمليات التحليل هذه.
- ضعف المشاركة الإيجابية على تويتر مقارنة بفيس بوك في بعض الدول العربية مثل مصر على سبيل المثال.

هل يمكن الاستفادة من أنظمة تحليل المشاعر في عمليات تقييم فعالية خدمات المعلومات؟

من واقع الدراسة الاستكشافية، يمكن القول: إنه في الوقت الراهن وفي ظل ضعف المشاركة والتفاعل من جانب المستفيدين فيما يتعلق بما ينشر من منشورات على صفحات المكتبات، فالمادة الخام الأساسية ونقطة الانطلاق لها تتمثل في وجود مجموعة من البيانات Dataset؛ وفي ظل تقديم أغلبها تقييم مجملا على درجة استقطاب الرأي من (إيجابي – سلبي – محايد)؛ فضلًا عن مشكلات تطبيق اللغة العربية في كثير من البرمجيات مفتوحة المصدر التي يمكن استخدامها، مما يتطلب وجود متخصص للتعامل معها، الأمر الذي قد لا يتوافر في بعض المكتبات؛ فإن هذه العملية تحتاج لبذل جهد مضاعف حتى يتسنى الاستخدام الأمثل لها، حتى يمكن التغلب على هذه العقبات التي تتطلب اتخاذ التدابير التي تكفل تذليل تلك العقبات من خلال رفع الوعي بأهمية تشجيع المستفيدين على المشاركة الفعالة، والتنمية المهنية للعاملين في المكتبات للتدريب على بعض التقنيات التي لا تتطلب قدرًا كبيرًا من المعرفة البرمجية، إلا أن ذلك لا يمنع من أهمية استخدامها وفائدتها في قياس التقييم العام من جانب المستفيدين إذا ما توافر لدى المكتبة القوى البشرية المؤهلة لذلك.

نتائج الدراسة الاستكشافية:

أظهرت نتيجة تحليل العينة الاستكشافية من التعليقات ما يلي:

- هناك مجموعة من جوانب الإفادة التي تعود على المكتبات وغيرها من مؤسسات المعلومات عند تطبيق أسلوب تحليل المشاعر على تعليقات المستفيدين من خدماتها عبر صفحات التواصل الاجتماعي الخاصة بها.
- لا تقتصر جوانب الإفادة على عمليات تحليل درجة الاستقطاب والوقوف على المشاعر الإيجابية والسلبية، وإنما يمكن الاستفادة من البيانات الناتجة عن التحليلات التفصيلية التي يقدمها النموذج، مثل: حجم ما ينشر من تعليقات والموضوعات التي تتناولها، في قياس رضاء المستفيدين بما يسهم في دعم عمليات اتخاذ القرار لتحسين مستوى تقديم الخدمات.
- بلغت نسبة الدقة في النموذج 81٪، ومثلت التعليقات الإيجابية الغالبية من التعليقات بنسبة 45٪، يليها التعليقات المحايدة بنسبة 37.5٪، ثم التعليقات السلبية بنسبة 17٪، ولوحظ عند تحليل محتوى التعليقات المحايدة أنها تركز على توجيه تساؤلات واستفسارات للمكتبة وبصفة خاصة في المكتبات العامة، وكان أغلب تلك التساؤلات حول

ساعات فتح المكتبة وكيفية الاشتراك في الفعاليات والأنشطة، والمراحل السنية للأطفال للتسجيل ضمن أنشطة معينة، وتعد التعليقات السلبية أكثر دلالة من حيث المحتوى حيث تشتمل على بعض نقاط الضعف من وجهة نظر المستفيدين.

- هناك عدد من الخيارات للمكتبة عند التطبيق ما بين التطبيقات الجاهزة، أو أن تقوم ببناء النموذج الخاص بها سواء من خلال التقنيات المدفوعة أو مفتوحة المصدر، إلا أن البرمجيات الجاهزة مفتوحة المصدر تواجه عددا من التحديات، من بينها صعوبة التعامل مع النصوص باللغة العربية، وقلة ما يتوافر من أدوات أومجموعات بيانات جاهزة أومكتبات برمجية تتعامل مع النص العربي بفاعلية.
- لوحظ ضعف المشاركة الفعالة للمستفيدين في التعليقات أو الاكتفاء باستخدام الرسوم التعبيرية وتسجيلات الإعجاب أو المشاركة، بالإضافة إلى طبيعة محتوى التعليقات وغموض دلالة بعضها حيث غلب عليها تسجيلات الإعجاب أو المشاركة أو ذكر أسماء الأشخاص mention دون كتابة نص للتعليق في كثير من منشورات المكتبات.
- تواجه عمليات تحليل المشاعر ببعض التحديات ذات العلاقة بقضايا الخصوصية، ويجب على المكتبة أن تراعي خصوصية بيانات المستفيدين عند القيام بعمليات التحليل وأن تكون هناك سياسة معلنة في هذا الصدد.
- هناك عدد من لغات البرمجة المستخدمة في هذا الصدد مثل لغة البايثون و json ولغة R . مع تنوع أساليب تحليل المشاعر ما بين لغة الآلة أو التعلم العميق أو الأسلوب الهجين، وتؤدي أداة Jupyter Notebook كأداة مساعدة في بناء نموذج تحليل المشاعر دورًا كبيرًا في تسهيل التعامل مع بعض الأكواد البرمجية، كما تساعد في استيراد مكتبات الأكواد مثل pandas، wordCloud، numpy وغيرها مما يسهم في إنجاز العمل بصورة أفضل.
- على الرغم من الأهمية التي تقوم بها عملية تنقية البيانات data cleaning فقد لوحظ أن تجذيع الكلمات stemming على سبيل المثال- عند القيام بتحليل المشاعر- قد يفقد الكلمات دلالتها في التعبير عن القطبية، فضلًا عن فصل الجمل في صورة كلمات منفصلة، قد لا يعبر بوضوح عن المعنى المقصود ومن الممكن أن يذهب بالمعنى من الإيجابية إلى السلبية والعكس صحيح، بينما تفيد بعض عمليات تنقية النص الأخرى مثل: إزالة العلامات الخاصة وإعادة التغريد في زيادة الدقة في عمليات التحليل.
- يمكن استخدام هذا الأسلوب في تحليل استقصاءات الرأي التي تنطوى على أسئلة موضوعية (كيفية).
- في بعض البرمجيات أو في حالة تعامل من لديه خلفية برمجية في الموضوع يمكن الحصول على مزيد من المؤشرات والتحليلات مثل: أهم الكلمات المفتاحية التي ورد تصنيفها من حيث الخصائص.

التوصيات:

- توصي الدراسة بضرورة تبني دراسات بينية يتعاون فيها باحثون من تخصصات علوم المكتبات والمعلومات، والحاسبات، واللغويات لتصميم أدوات برمجية جاهزة تدعم إمكانات اللغة العربية ويمكن أن تستخدمها المكتبات وغيرها من مؤسسات المعلومات بصورة مباشرة دون الحاجة للتعقيدات البرمجية لعمليات تحليل الآراء بصورة مفصلة.
- ينبغي بذل مزيد من الجهد من جانب العاملين بالمكتبات وغيرها من مؤسسات المعلومات لرفع وعي المستفيدين بالمشاركة الفعالة ومتابعة صفحاتهم عبر وسائل التواصل الاجتماعي وغيرها.

- تصميم موارد معجمية تدعم إمكانات اللغة العربية، يمكن أن تستخدمها المكتبات وغيرها من مؤسسات المعلومات بصورة مباشرة لتصنيف التعليقات.
- الاهتمام بإعداد مجموعة من الدورات التدريبية بهدف إلقاء الضوء على أسلوب تحليل المشاعر وأبرز التقنيات المستخدمة به خاصة على مستوى التنمية المهنية للعاملين بالمكتبات وغيرها من مؤسسات المعلومات.
- ينبغي رفع الوعي لدى القائمين على التحليل الآلي للمشاعر بأخلاقيات الذكاء الاصطناعي (ويمكن الاسترشاد في هذا الصدد بتوصية وإرشادات اليونسكو في هذا الصدد).
- التشجيع على إعداد مزيد من الدراسات التي تتناول قضايا وجوانب تفصيلية تتعلق بالموضوع ومن بينها على سبيل المثال، دراسة تطبيقه في تحليل الأسئلة المفتوحة الواردة باستبيانات استقصاء الرأي للمكتبات، وبناء نماذج تطبيقية على أنواع معينة من المكتبات، ومدى إمكانية توظيفه في عمليات تنمية المقتنيات استنادًا على تقييمات الكتب وغيرها من المصادر سواء للكتب المطبوعة أو الإلكترونية، واستخدامه لتقييم تطبيقات الهواتف الذكية للمكتبات، وكيفية التعامل مع الرسوم التعبيرية والمشاركات وتسجيلات الإعجاب والإشارة إلى أسماء أشخاص mentions ودلالتها في تحليل الرأي.

قائمة المصادر والمراجع

أولا: المراجع العربية:

إبراهيم حبيبي، وعبد الحميد الجيهاد (2020)، بناء موارد أساسية لتحليل الآراء والمشاعر الواردة في التعليقات باللغة العربية. ***Journal of Engineering and Computer Science (JECS)*** **21(2), 18-35.**

أمل وجيه حمدى (2019)، دراسة تأثير وجود القياسات البديلة **Altmetrics** بمحركات البحث على استخدام الباحثين للمصادر الإلكترونية، بحوث في علم المكتبات والمعلومات22 ، (مارس)، 79-142.

حمزة السيد حمزة (2021)، توظيف تطبيقات الذكاء الاصطناعى لتحليل مشاعر مستخدمى مواقع التواصل الاجتماعي في الوقت الفعلي لأزمة جائحة فيروس كورونا، المجلة المصرية لبحوث الرأى العام**.202-149 ,20(2) ,**

ريهام سامى (2021)، اتجاهات مستخدمى مواقع التواصل الاجتماعى نحو جائحة كورونا (كوفيد-19): تحليل من المستوى الثانى لدراسات مدخل معالجة اللغة الطبيعية **Social media users' attitudes towards Corona Pandemic (Covid19): A secondary analysis of Natural language processing approach studies.** مجلة البحوث الإعلامية، 56(56-ج3)، 1301-1350.

شهد عبد الحليم العشى، وضحى هزاع بادى (2021)، تحليل البيانات الضخمة من الشبكات الاجتماعية كعامل تمكين لإدارة المعرفة المستدامة للاتجاهات التّسويقيّة. ***Journal of Information Studies & Technology (JIS&T)***, 2021(2), 8.

طارق الخليفى (2019)، تنقيب بيانات وسائل التواصل الاجتماعى واستخداماته فى البحوث الإعلامية، تحليل المشاعر نموذجًا، مجلة البحوث والدراسات الإعلامية**73-1 ,8(8)**

عصام محمد عبيد (2014)، واقع تويتر (Twitter) فى المكتبات الجامعية العربية كخدمة من خدمات الإحاطة الجارية : دراسة تحليلية تقويمية مع طرح معايير الاستخدام، مجلة التعليم عن بعد والتعليم المفتوح، مج 2، ع3، ص 111 – 181.

هبة أحمد محمد المتبولى (2019)، المكتبات الجامعية المصرية والعربية على شبكات التواصل الاجتماعي: دراسة وصفية تحليلية مقارنة، بحوث في علم المكتبات والمعلومات23 ,(سبتمبر), 199-256.

ثانيا: المراجع الأجنبية:

Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. (2017). Machine learning techniques for sentiment analysis: A review. *Int. J. Multidiscip. Sci. Eng*, 8(3), 27.

Ahmad, I. S., Bakar, A. A., Yaakub, M. R., & Darwich, M. (2020). Beyond sentiment classification: A novel approach for utilizing social media data for business intelligence. *International Journal of Advanced Computer Science and Applications*, 11(3).

Al-Ayyoub, M., Essa, S. B., & Alsmadi, I. (2015). Lexicon-based sentiment analysis of arabic tweets. *International Journal of Social Network Mining*, 2(2), 101-114.

Aldayel, H. K., & Azmi, A. M. (2016). Arabic tweets sentiment analysis–a hybrid scheme. *Journal of Information Science*, 42(6), 782-797.

Al-Otaibi, S., Alnassar, A., Alshahrani, A., Al-Mubarak, A., Albugami, S., Almutiri, N., & Albugami, A. (2018). Customer satisfaction measurement using sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 9(2), 106-117.

Alessia, D., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications*, 125(3).

Chen, H., Zimbra, D. (2010). AI and opinion mining. Intell. Syst. 3(25), 74–80.

Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *Procedia Computer Science*, 161, 707-714.

Lamba, M., & Madhusudhan, M. (2018). Application of sentiment analysis in libraries to provide temporal information service: a case study on various facets of productivity. *Social Network Analysis and Mining*, 8(1), 1-12.

Lamba, M., & Madhusudhan, M. (2022). Sentiment Analysis. In *Text Mining for Information Professionals* (pp. 191-211). Springer, Cham.

Mandalapu, A. C., Gunabalan, S., Sadineni, A., Cai, T., Al Hasan Haldar, N., & Li, J. (2019,

November). Correlate influential news article events to stock quote movement. In *International Conference on Advanced Data Mining and Applications* (pp. 331-342). Springer, Cham.

Moore, M. T. (2017). Constructing a sentiment analysis model for LibQUAL+ comments. *Performance Measurement and Metrics*. Available at: < https://www.emerald.com/insight/content/doi/10.1108/PMM-07-2016-0031/full/pdf?casa_token=RH4en7wyUbcAAAAA:U_6nq5GYHpawJVmoBQyE2uaCVXyky_36JY1HQdeH-X_tozPI2_gowVPJk-Tuthxf5jHq1A5QBqd5X11T4XZP6uQNZpXeH1_6xNYOlr8Z1h3BKyhq2wqg6w>

Oxford Reference (2022). Sentiment analysis.. Retrieved 28 June 2022, from https://www.oxfordreference.com/view/10.1093/acref/9780191803093.001.0001/acref-9780191803093-e-1306

Papachristopoulos, L., Ampatzoglou, P., Seferli, I., Zafeiropoulou, A., & Petasis, G. (2019). Introducing sentiment analysis for the evaluation of library's services effectiveness. *Qualitative and Quantitative Methods in Libraries*, 8(1), 99-110.

Patra, S. K. (2019). How Indian libraries tweet? Word frequency and sentiment analysis of library tweets. *Annals of Library and Information Studies* (*ALIS*), 66(4).

Sentiment Analysis: Meaning & Definition for UK English *in*: Lexico.com. (2022). Retrieved 28 June 2022, from https://www.lexico.com/definition/sentiment_analysis.

Sharma, R., Gulati, S., Kaur, A., & Chakravarty, R. (2021). Users' Sentiment Analysis toward National Digital Library of India: a Quantitative Approach for Understanding User perception.

●●●

AI-Based Sentiment Analysis

as a Tool to Measure the Information Services' Effectiveness

An Exploratory Study

Dr. Laila Sayed Samea

Assistant Professor – Dept. of Libraries, Archives, and Information Technology
Faculty of Arts – Cairo University
Laila_mtu@yahoo.com

The study explores the opportunities of using sentiment analysis technique, as one of the artificial intelligence applications, to measure the user satisfaction on the library and information services as an indicator for its effectiveness. The study reviews main tools and techniques used, the areas of application in Arabic libraries' environment through the analysis of users' comments published in its social media pages. The study discusses the challenges of implementing this method in the Arabic libraries' environment in terms of the significance of the comments, the polarity classification process to positive, negative, or neutral feelings about the information services provided which supports the Decision-making process to improve the library and information services provided. An exploratory study was conducted for a sample of users' comments on social media pages. A dataset of (665) tweets/comments was collected through December 2021 to end of April 2022, by searching the terms related to library services via Twitter and Facebook platforms. The NLP process was performed using the free tool Google Colab. The result of the analysis shows the capability of this method to provide a detailed data with required analysis in various forms which help identify the user opinions towards the library posts/tweets. There are some sentiment analysis applications that can be used to generate the sentiments automatically, but they provide limited capabilities especially with Arabic text. Some challenges are identified related to the nature of Arabic language, and its dialects in the NLP process, and generating the word cloud, the availability of software tools and code libraries to deal with Arabic text effectively. Some other challenges are related to the nature of the comments/tweets and their meaning ambiguity, using Franco-Arab to express the Arabic text, the lack of active participation of users in the terms of their comments, in addition to the challenges related to privacy issues. The study recommends the collaboration among the researchers from disciplines of library and information sciences, computer science, and linguistics to conduct interdisciplinary studies to design ready-made software tools that support the capabilities of the Arabic language and can be used by libraries and other information institutions directly without the need for the programming complexities to get detailed opinion analysis data.

***Keywords*: *Sentiment Analysis*(1), *Natural Language Processing*(2), *User Satisfaction*(3), *Information Services*(4)**