مؤسسة النهضة العلمية

دراسة شاملة لتحليل المشاعر البلاغية في اللغة العربية

نوح صبري المتولى[1]، احمد السياط[2]، سعد العنزى[3]

[1]كلية علوم الحاسب والمعلومات، قسم علوم الحاسب، جامعة الجوف

كلية الحاسبات والذكاء الاصطناعي، قسم علوم الحاسب، جامعة القاهرة

[2] كلية علوم الحاسب والمعلومات، قسم علوم الحاسب، جامعة الجوف

[3] كلية علوم الحاسب والمعلومات، قسم علوم الحاسب، جامعة الجوف

[1] nselmitwally@ju.edu.sa, nouh.sabri@fci-cu.edu.eg

[2] asayat@ju.edu.sa

[3] sanazi@ju.edu.sa

الخلاصة:

تقوم معالجة اللغة العربية على مستويات متعددة؛ تتكامل هذه المستويات التكميلية في كثير من الأحيان، بشكل جيد مع بعضها البعض. تتألف مستويات تحليل الآراء والمشاعر للغة العربية ابتداء من المستوى الصوتي، النحوي، الصرفي والمعجمي، والدلالي، بالإضافة الى المستوى البلاغي. ويُعد تحليل الآراء والمشاعر موضع اهتمام كبير لأصحاب اللغات الإنجليزية والهندية الأوروبية، مع انحسار كبير على اللغة العربية، وهي لغة مليئة بالخصائص الخطابية والمعاني الضمنية التي لها دلالات ومعاني إيجابية وسلبية عبر المستويات اللغوية الستة – المذكورة آنفا. تقدم هذه الورقة اقتراحًا شاملاً وكاملاً لاستراتيجية معالجة اللغة العربية لتحليل الآراء والمشاعر على مستوى الخطاب النصي المكتوب والمقروء. يحتل هذا البحث أهميته بالأساس من كونه يعنى باللغة العربية التي تعاني نقصا في الاهتمام من حيث تحليل الآراء والمشاعر للنصوص البلاغية العربية. وقد توصلت النتائج الى أهمية تفوق بعض خوارزميات تعلم الآلة وبمعدل دقة عالي في تصنيف النصوص اللغوية في المكنز اللغوي من حيث أساليب التشبيه وصيغ المبالغة

الكلمات الجوهرية: تحليل الآراء والمشاعر، معالجة اللغات الطبيعة، تعلم الآلة، تنقيب البيانات.

1. المقدمة

لقد ساهمت علوم الذكاء الاصطناعي وتعلم الآلة والتعلم العميق في تطبيقات عديدة وشتى؛ ولاسيما في مجال معالجة اللغات الطبيعية على مستوى كبير من الاهتمام والتطبيق للمساهمة في خدمة البشرية في التعامل مع وسائل تكنولوجيا المعلومات بيسر وسهولة وفاعلية. ولا يخفى علينا أن اللغة العربية – لغة الوحي المنزل من السماء (القرآن الكريم) – لم يكن لها هذا الاهتمام الكبير في مجال هذه العلوم التقنية الحديثة. وأصبح من الواجب علينا كباحثين في هذا المجال العلمي المتخصص فتح المجال للتعرف على ما تم إنجازه وما هي مشاكل ومعوقات تطبيقات علوم الذكاء الاصطناعي للغة العربية عامة وللنص القرآني والنبوي على وجه الخصوص من خلال الأهداف والآليات المذكورة في هذا المقترح البحثي إن شاء الله تعالى.

لقد أصبحت وسائل التواصل الاجتماعية مثل تويتر والفيسبوك مهمة جدا في معرفة توجه الناس حول موضوع معين، ولذا تم استخدام واستغلال البيانات التي تنتجها وسائل التواصل الاجتماعي من قبل الشركات التجارية والحكومات لقياس الرأي

حول مواضيع أو منتجات معينة. ويوجد العديد من الدراسات العلمية التي تمت على البيانات الصادرة من وسائل التواصل الاجتماعي لتحليل مستخدمي وسائل التواصل الاجتماعي.

الهدف من هذا المشروع البحثي هو دراسة شاملة لتطبيقات وتحديات هذا العلم في لغتنا الحبيبة – اللغة العربية – ومدى معرفة إمكانية الاستفادة من هذه التطبيقات في مجالات كثيرة تخدم علومنا وثقافتنا ولغتنا الجميلة.

بداية سوف يتم عمل مسح أدبي لمعرفة الطرق المستخدمة في هذا النوع من الدراسات والتي تعتمد على دراسة الخوارزميات المطبقة في اللغة الإنجليزية والعربية على حد سواء. ومن ثم تحديد أوجه التقدم والقصور والتحديات في تفعيل مثل هذه العلوم الذكية الحديثة للاستفادة منها على مستوى كبير وفعال. ولاسيما أن ما تم إنجازه وتحقيقه في مثل هذه التطبيقات على اللغة العربية يعد نسبة ضئيلة جدا مقارنة باللغات الطبيعة الأخرى؛ حيث أن أغلب الدراسات تمت على اللغة الإنجليزية وعليه سوف نحاول أن ندرس إمكانية تكييف تلك الخوارزميات للعمل على معالجة النصوص المكتوبة باللغة العربية.

ويحتل هذا البحث أهمية كبرى من حيث كونه معنى باللغة العربية التي تعاني نقصا في الاهتمام من حيث تطبيقات الذكاء الاصطناعي – وعلم تعلم الآلة والتعلم العميق – التي تهتم بمعالجة البيانات النصية على خلاف كثير من اللغات الاخرى. فحسب علمنا وكما هو موضح في المسح الادبي أنه لا يوجد بحث علمي حتى الآن لرصد ومعرفة ما توصلت إليه تطبيقات اللغة العربية في هذا المجال على مستوى كبير وشامل. وعليه فهذه الدراسة العلمية سوف تساعد الباحثين وتضئ لهم الطريق في هذا المجال البحثي المهم.

أ- الأهمية والمبررات:

القيام بدراسة علمية شاملة فاحصة للجهود التي تمت في مجال تحليل الآراء القائم على أدوات وعلوم الذكاء الاصطناعي والتعلم العميق وتعلم الآلة. كما يحتل هذا البحث أهميته بالأساس من كونه يعنى باللغة العربية التي تعاني نقصا في الاهتمام من حيث تطبيقات الذكاء الاصطناعي التي تهتم بمعالجة البيانات النصية على خلاف كثير من اللغات الاخرى. فحسب علمنا وكما هو موضح في المسح الادبي أنه لا يوجد بحث علمي حتى الآن للوقوف عل ما توصلت إليه تطبيقات اللغة العربية في هذا المجال على مستوى كبير وشامل، إيجابا أو سلبا بشكل آلي دون مراجعة أو قراءة جميع ما كتب في في هذا المجال.

ب- أهداف البحث:

1- بناء مجموعة بيانات: استخلاص جميع البحوث الممكنة والتي تتحدث عن تطبيقات "تحليل الآراء" باللغة العربية

2- استخلاص خصائص التعبير عن البيانات اللغوية: تحليل البيانات المجمعة في مجموعة البيانات، لمعرفة اللغة المستخدمة في وسائل التواصل الاجتماعي وغيرها باللغة العربية وذلك من خلال خوارزميات تساعد بتحقيق الهدف، وتتبع مختلف التعابير المستعملة في هذا النطاق.

3- فحص شامل للتحديات الحالية في هذا المجال، ودراسة كيفية حلولها وماهي المعوقات والمشاكل التي تواجه الباحثين في هذا المجال؟

ج- كيفية تحقيقه:

1. الدراسة الفاحصة الشاملة بأسلوب المسح الأدبي في البحث العلمي
2. الوقوف على أهم المصادر والمراجع في هذا المجال
3. دراسة التطبيقات المستقبلية والحالية بالنسبة للغة العربية
4. الوقوف على أحدث التقنيات والخوارزميات التي ساعدت في حل مشاكل كثير من هذه التطبيقات
5. المشاركة في تأسيس ما يجب أن يقوم على البحث العلمي وأسسه ووسائله في تطبيقات تحليل الآراء في اللغة العربية

د- تفاصيل المقترح البحثي:

1- المسح الادبي العلمي: الاطلاع على البحوث ذات العلاقة المنشورة في المجلات والدوريات المحكمة
2- بناء مجموعة بيانات: استخلاص البيانات ومصادرها التي تتحدث اللغة العربية من منصات وسائل التواصل الاجتماعي ومكانز البيانات الأخرى المتخصصة في هذه التطبيقات
3- استخلاص خصائص التعبير عن لغة "تحليل الآراء": تحليل البيانات المجمعة في مجموعة البيانات، لمعرفة اللغة المستخدمة في الحديث عن طرق التعبير المستعملة في اللغة العربية وذلك من خلال خوارزميات تساعد بتحقيق الهدف.
4- العمل على تأسيس وفهم هذا النظام والقائم على تصنيف الآراء إيجابية أو سلبية أو غير ذلك: بناء على نتائج التحليل في المرحلة السابقة سوف يتم اختيار أبرز التقنيات والخوارزميات والطرق العلمية التي ساعدت وساهمت في فهم وتأسيس هذا العلم
5- التنفيذ: بعرض شيق ومستفيض ومرتب لأبرز هذه التقنيات والوسائل والأدوات.

هـ- طريقة التنفيذ:

المسح الأدبي للدراسات السابقة (Literature review)

القيام بمسح علمي نظري للطرق الآلية المستخدمة في تحليل البيانات المجموعة بمختلف المجالات، واستقراء وحصر ما تم إنجازه فيما يخص معالجة النصوص العربية آليا في مجال تحليل الآراء.

تحليل وتجميع البيانات (Data Collection & Analysis)

في هذه المرحلة يتم جمع البيانات (من مصادر مختلفة)، ثم بعد ذلك يتم تحليلها بعد معالجتها من المشاكل المتعلقة بها.

مرحلة التنفيذ (Design & implementation)

استخلاص خصائص التعبير عن لغة "تحليل الآراء": تحليل البيانات المجمعة في مجموعة البيانات، لمعرفة اللغة المستخدمة في الحديث عن طرق التعبير المستعملة في اللغة العربية وذلك من خلال خوارزميات تساعد بتحقيق الهدف.

مرحلة التحقق والاختبار (Test & evaluation)

القيام باختبار النماذج المتاحة من خوارزميات وتطبيقات، والتي تم التوصل إليها في المرحلة السابقة من خلال بيانات (Training data) والبيانات التي لم تستخدم في مرحلة التطوير (Testing data).

الخاتمة والدراسات المستقبلية (Conclusions and future perspectives)

استعراض نتائج البحث العلمي والتوصيات المستقبلية في المشاريع البحثية القادمة

2. المسح الأدبي

تركزت الدراسات المتخصصة في تحليل الآراء والمشاعر على مستويات عديدة، منها ما هو على مستوى الكلمة والجملة والنص المتعدد الجمل [1]. وكانت التصنيفات التي تحلل هذه النصوص واقعة بين ما هو إيجابي او سلبى أو محايد (طبيعي). وذلك بعد التعليق على النصوص من أصحاب اللغة والمتخصصين في معرفة المعاني التي تؤول اليها الكلمات والجمل من حيث القبول والرفض او الايجاب والسلب، الخ [2].

وتم التركيز على هذا الاتجاه وبكثرة في نصوص اللغة الإنجليزية [3] وغيرها من اللغات الأخرى مثل الصينية [4] والإيطالية [5] وغيرها. وقد حازت اللغة العربية على نصيب لا بأس به في هذا المستوى من التحليل والتصنيف مقارنة باللغات الأخرى. وقد تمت المعالجة والتصنيف على أسس علمية وخوارزميات متخصصة من تعلم الآلة والتعلم العميق.

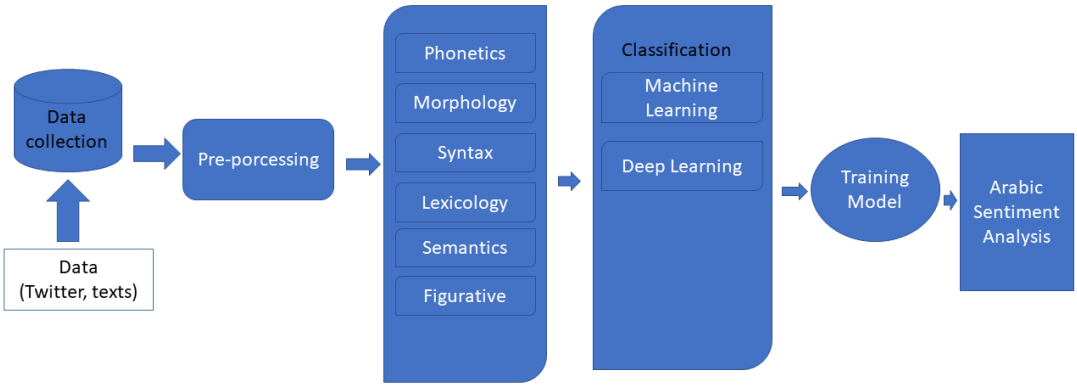
وأصبحت اللغة العربية محاكية وبصورة تقليدية لهذا التوجه والذي نعتبره تقليدا مباشرا للغات الأخرى ينافي طبيعة لغتنا الحبيبة والتي تنبثق معانيها وأساليبها من لغة النص القرآني المجيد.

وحين نستعرض مستويات معالجة اللغة العربية الستة، نجد أن كلا منها ينشأ عنه جرسا وحسا لغويا يقع أثره في النفس مباشرة من حيث القبول والرفض أو الايجاب والسلب.
فالمستوى الصوتي؛ والذي يعد المستوى الأول في مستويات معالجة اللغات الطبيعية له طابع خاص في اللغة العربية، من حيث الصفات والمخارج.. فالحروف التي يكون مخرجها من الحلق تختلف من حيث الصفة عن مخارج حروف اللسان (اوله – وسطه – نهايته) وذلك من حيث الشدة والرخاوة – التفخيم والترقيق.. الخ
وهذا مبحث كبير جدا في علوم القرآن والقراءات، والتي تولى هذا المستوى أهمية خاصة في التحليل والتوجيه. (راجع دراسات صوتية في تجويد الآيات القرآنية – الأستاذ الدكتور صبري المتولي – جامعة القاهرة) [6]

ونقيس على هذا المستويات الأخرى (المستوى النحوي والصرف والدلالي والمعجمي) كل من هذه المستويات يخرج لنا معنى وجرسا لغويا تنعقد به المعاني والدلالات المختلفة على الوجه الإيجابي او السلبي.

حتى يطوف بنا المقام ونصل في التوجيه الى المستوى الأعلى وهو المستوى البلاغي، وهو من أخص المستويات والذي منه يتم تصنيف الكلام والحوار والجمل الى معاني ذات طبيعة حسية وجمالية؛ إيجابية كانت او سلبية او محايدة. وسوف نستعرض امثلة كثيرة على هذا المستوى – موضوع البحث – في الفصل الثاني من هذه الورقة العلمية.

ويوضح الشكل 1 النموذج المقترح لمعالجة مستويات اللغة العربية من حيث مفهوم تحليل الآراء والمشاعر (Sentiment Analysis)، الذي يوضح التي تتم معالجة النصوص من خلالها بعد تجميعها وتنقيحها واستخلاص الشوائب منها، ثم معالجتها على المستوى (الصوتي – النحوي – الصرفي – المعجمي – والدلالي – البلاغي).



الشكل 1: نموذج تحليل الآراء والمشاعر للغة العربية [1]

3. بناء المكنز البلاغى اللغوى

تنقسم العلوم البلاغية الى ثلاثة اقسام: (1) علم البيان (2) علم المعاني (3) علم البديع.

وينبغي أن يُعْلَم أن مصطلح [الحقيقة والمجاز] خاص بعلم البيان ويعرف بأنه "علم يُبْحَثُ فيه عن بلوغ الأثر المطلوب، عن طريق بيان المعنى المطلوب، بالفن الأنسب من فنون المجاز. (راجع أسرار البلاغة في علم البيان– وكتاب دلائل الإعجاز– عبد القاهر الجرجاني) [7]

وإذا نظرنا الى الجناس – على سبيل المثال – والذي يعرف على أنه اتحاد الكلمتين في الحروف والاختلاف في المعنى، وانواعه:

1- الجناس التام: وهو اتحاد الكلمتين في عدد الحروف وترتيبها، وضبطها مع الاختلاف في المعنى. مثال: (هوى، هوى – بمعنى – أحب، سقط)
2- اختلاف الكلمتين في عدد الحروف أو ضبطها أو ترتيبها. مثال: (الشفيق، الشقيق – بمعنى – الصاحب، الأخ). وهذا كله يحدث نغما موسيقيا يثير النفس وتطرب له الأذن. ومثال ذلك: (ذهب محمد الى السوق واشترى ذهبا) هذا جناس تام لاتحاد الحروف واختلاف المعنى.
3- الجناس غير التام: هو اختلاف الكلمتين في ترتيب الحروف. مثال: (البدر – البرد)

الحقيقة والمجاز:

تعرف الحقيقة على انه استعمال اللفظ فيما وضع له.

والمجاز: هو اللفظ المستعمل في غير ما وُضِعَ له، ومثال ذلك: رأيت أسدا يقاتل في سبيل الله – حيث تم تشبيه الجندي بالأسد، مع حذف المشبه والتصريح بلفظ المشبه به.

ونستعرض فيما يلي بعض الأمثلة الشائعة والتي يعمل بها في كثير من المصادر لتوضيح مفاهيم الكناية والاستعارة والتشبيه والمجاز في اللغة العربية. والهدف من هذا هو التأكيد على أن المعاني التي تحتويها هذه الجمل والعبارات هي بالأحرى أن تكون معان واضحة من حيث الأثر الواقع على النفس من سماعها ومعرفتها (إيجابا أو سلبا) وهو المعمول به في دراسة تحليل الآراء والمشاعر – سواء كان اللغة المستعملة لغة تعليقات تجارية وتغريدات – أو لغات عامية دارجة – أو لغة عربية فصحى. ونستعرض في الجدول 1 تصميم المكنز اللغوي البلاغي والذي قامت على أساسه الدراسة العلمية لتحليل الآراء والمشاعر.

1- طويل النجاد كناية عن معنى شجاع (فهي كناية عن صفة) لأن طول النجاد يستلزم طول السيف وطول السيف يستلزم طول القامة وطول القامة يستلزم الشجاعة
2- رفيع العماد كناية عن معنى سيد وزعيم (فهي كناية عن صفة) لأن رفعة العماد تستلزم ارتفاع الخيمة وارتفاع الخيمة يستلزم أن صاحبها سيد وزعيم في قومه.
3- كثير الرماد كناية عن معنى كريم (فهي كناية عن صفة) لأن كثرة الرماد تستلزم كثرة الطبخ وكثرة الطبخ تستلزم كثرة الضيوف وكثرة الضيوف تستلزم كثرة الكرم.
4- بسطهمُ حرير كناية عن معنى سعداء وأعزاء (فهي كناية عن صفة)
5- بسطهمُ تراب كناية عن معنى محتاجين وأذلاء ومهانين (فهي كناية عن صفة)
6- نؤوم الضحى .كناية عن معنى كسول (فهي كناية عن صفة)
7- ألقى عصاه كناية عن معنى مستقر وثابت في مكان معين (فهي كناية عن صفة)
8- ناعم الكفين كناية عن عدم العمل وممارسة أي نشاط (فهي كناية عن صفة)
9- موطن أسراره كناية عن القلب (فهي كناية عن موصوف)
10- رغوة الشباب كناية عن الشيب (فهي كناية عن موصوف)
11- اليمن يتبع ظله كناية عن نسبة (نسب اليمن وهو البركة إلى الظل بدل الممدوح)
12- والمجد يمشي في ركابه عن نسبة (نسب المجد إلى الركاب بدل الممدوح)
13- قرب حاجبا من حاجب كناية عن معنى سيء العشرة (فهي كناية عن صفة)
14- لبست له جلد النمر كناية عن معنى إظهار العداوة وكشفها (فهي كناية عن صفة)
15- لا تقدم رجلا وتؤخر أخرى كناية عن معنى متردد (فهي كناية عن صفة)
16- أبى نظيف اليد
17- تعلمت لغة الضاد
18- يشار اليه بالبنان
19- وحملناه على ذات ألواح ودسر
20- الخيل معقود بنواصيها الخير
21- ويوم يعض الظالم على يديه
22- وقف مرفوع الرأس
23- ولا تجعل يدك مغلولة إلى عنقك ولا تبسطها كل البسط
24- الفصاحة في بيانه والبلاغة في لسانه
25- وأسمعت كلماتي من به صمم

26-رأيت أسدًا يحكي بطولاته
27-فلان يده طويلة
28-المسلم من سلم المسلمون من لسانه ويده
29-خالد مكتنز اللحم
30-هو غليظ الكبد
31-المجد بين ثوبيه، والكرم تحت ردائه
32-وَكُلُوا وَاشْرَبُوا حَتَّى يَتَبَيَّنَ لَكُمُ الْخَيْطُ الْأَبْيَضُ مِنَ الْخَيْطِ الْأَسْوَدِ مِنَ الْفَجْرِ

الجدول ا: تصميم المكنز البلاغي اللغوي

مسلسل	الجملة	التصنيف					الانطباع عن الجملة		
		صيغة مبالغة	استعارة	تشبيه	كناية	ايجابى	سلبى	طبيعى
1	كثير الرماد				√	√		
2	رأيت اسدا يقاتل في سبيل الله		√			√		
3	الرحمن الرحيم	√				√		

وقد تم بناء المكنز والذي احتوى حجمه على 2000 من النصوص المتفرقة من النصوص القرآنية والأحاديث النبوية الى غير ذلك من التغريدات المقتبسة من موقع تويتر. وتمت التعليقات على جميع النصوص مستعينين بذلك بثلاثة من أهل اللغة في هذا المضمار وتم استخلاص هذه النتائج المتعلقة بمصداقية التعليقات بمعيار (kappa) على النحو التالي:

$$k = \frac{P(A) - P(E)}{1 - P(E)} \quad (1)$$

الجدول 2: معيار الاتفاقية بين المعلقين

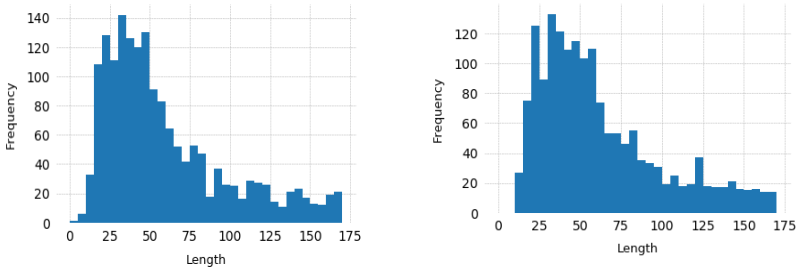
Metrics	**inter-annotator agreement (IAA)**
kappa	0.51
fleiss	0.51
alpha	0.51
scotts	0.49

4. نتائج الدراسة

نستعرض في هذا المقام بعض النتائج التي خلصت من إليها الدراسة – وحتى لا يطول التفصيل في هذه الورقة العلمية المطروحة – فسوف نكتفي بعرض بعض النتائج المقتبسة من كل مرحلة على النحو التالي:

أ. مرحلة معالجة النصوص

يوضح الشكل 2 حجم البيانات المستخدمة بعد معالجتها واستخراج منها الكلمات التي تشكل تكرارا في النصوص (مثل: ان – ما – كلا – لا.. الخ) وهذه تسمى (Arabic stop words)



الشكل 2: حجم المكنز اللغوي (ا) بعد المعالجة وقبل المعالجة (ب)

ب. مرحلة استخلاص الخصائص من النصوص

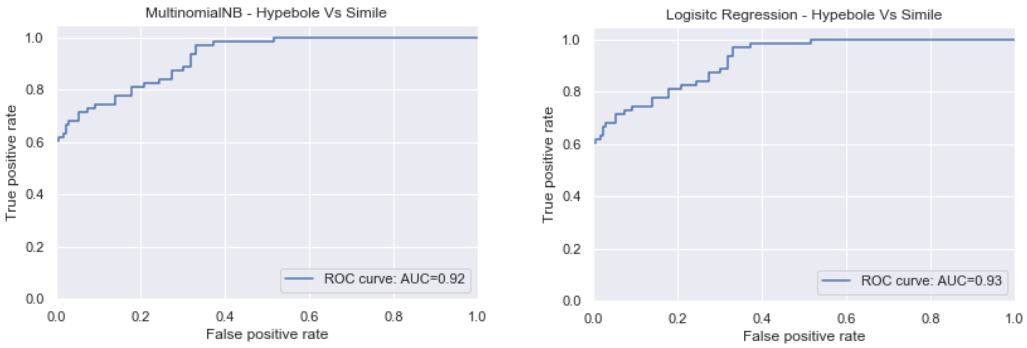
وقد تم استخلاص بعض الخصائص من النصوص مثل أسلوب صيغة المبالغة والتشبيه والاستعارة، وتصنيفها الى معانى إيجابية او سلبية (كما يوضح الشكل 3: الشكل السحابى لأكثر الكلمات تكرارا في كل من صيغة المبالغة والتشبيه)

الشكل 3: الشكل السحابى لتكرار الكلمات من النوع (ا) التشبيه – (ب) صيغة المبالغة

ت. مرحلة التدريب للبيانات باستخدام خوارزميات تعلم الآلة

تم استخدام خوارزميات متعددة للتصنيف والتمييز ونذكر منها: MultinomialNB, LogisticRegression, BernoulliNB, GaussianNB, KNeighbors, RandomForest, SVC and SGDClassifier

ويوضح الشكل 4 بعض نتائج وتقييمات هذه الخوارزميات والتي تم اختبارها على البيانات المجمعة في المكنز اللغوي (Arabic Figurative Corpus)



الشكل 4: (ا) نتيجة الخوارزمية LogisticRegression ب نتيجة الخوارزمية MultinomialNB

ث. مرحلة التقييم واستخراج النتائج

اثبتت النتائج الموضحة بالجدول 3 تفوق بعض خوارزميات تعلم الآلة على بعضها من حيث الدقة والقدرة على التصنيف والتمييز والتعرف والتعلم على البيانات المراد تصنيفها الى (صيغة مبالغة – تشبيه – سخرية) على النحو التالي:

الجدول 3 : نتائج الدراسة من حيث الدقة والقدرة على التصنيف لبعض خوارزميات تعلم الآلة

Machine Learning	Accuracy (%)	Recall (%)	F1-Score
MultinomialNB	87%	88%	87%
LogisticRegression	90%	89%	89%
SGDClassifier	89%	89%	89%

5. الخاتمة

تبرز أهمية اللغة العربية من كونها لغة الخطاب القرآني المعجز، ولغة النص النبوي المكتوب، واللغة التي يتكلم بها ملايين البشر في شتى بقاع المعمورة. وقد قامت كل أمة بخدمة لغتها والقيام على معالجتها آليا بفضل وقوة تكنولوجيا الذكاء الاصطناعي وعلم تعلم الآلة والتعلم العميق وغيرها من العلوم التطبيقية في هذا المجال. ونسأل الله التوفيق والسداد في عرض ما توصلت إليه التكنولوجيا المعاصرة من خدمة اللغة العربية ومعالجتها آليا في صورها المكتوبة والمنطوقة والمسموعة؛ حتى يتسنى للباحث المتخصص وطالب العلم الوقوف على هذه التطورات والإسهامات في هذا المجال الهام والمفيد لأمتنا ولغتنا الحبيبة.

وقد قامت الدراسة على علمية شاملة فاحصة للجهود التي تمت في مجال معالجة اللغات الطبيعية القائم على أدوات وعلوم الذكاء الاصطناعي والتعلم العميق وتعلم الآلة. كما يحتل هذا البحث أهميته بالأساس من كونه معني باللغة العربية التي تعاني نقصا في الاهتمام من حيث تطبيقات الذكاء الاصطناعي التي تهتم بمعالجة البيانات النصية على خلاف كثير من اللغات الاخرى. فحسب علمنا وكما هو موضح في المسح الادبي أنه لا يوجد بحث علمي حتى الآن للوقوف على ما توصلت إليه تطبيقات اللغة العربية في هذا المجال (تحليل الآراء للنصوص اللغوية) على مستوى كبير وشامل، إيجابا أو سلبا بشكل آلي وباستخدام خوارزميات تعلم الآلة.

6. المراجع

[1]

A. Alsayat and N. Elmitwally, “A comprehensive study for Arabic Sentiment Analysis (Challenges and Applications),” *Egypt. Inform. J.*, p. S1110866519300945, Jun. 2019.

[2]

L. Almuqren, L. Almuqren, and A. I. Cristea, “Framework for Sentiment Analysis of Arabic Text,” p. 4, 2016.

[3]

D. I. Hernández Farías, V. Patti, and P. Rosso, “ValenTO at SemEval-2018 Task 3: Exploring the Role of Affective Content for Detecting Irony in English Tweets,” in *Proceedings of The 12th International Workshop on Semantic Evaluation*, New Orleans, Louisiana, 2018, pp. 643–648.

[4]

D. Zhang, H. Lin, P. Zheng, L. Yang, and S. Zhang, “The Identification of the Emotionality of Metaphorical Expressions Based on a Manually Annotated Chinese Corpus,” *IEEE Access*, vol. 6, pp. 71241–71248, 2018.

[5]

A. T. Cignarella, C. Bosco, V. Patti, and M. Lai, “Application and Analysis of a Multi-layered Scheme for Irony on the Italian Twitter Corpus TWITTIRO`,” p. 8.

[6]

دراسات صوتية فى تجويد الآيات القرآنية، صبرى المتولى، 1439/2018 - مكتبة زهراء الشرق، جمهورية مصر العربية

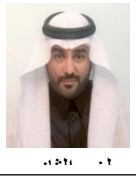
[7]

http://books-library.online/files/download-pdf-ebooks.org-kupd-3355.pdf
http://alqasimy.com/Files/books/e3jaz_alkor2an/dlail_ali3jaz_lljorjane/dlail_ali3jaz_lljorjane.pdf

7. جدول الألفاظ

English	عربي
Natural Language Processing	معالجة اللغات الطبيعة
Algorithms	الخوارزميات
Sentiment Analysis	تحليل الآراء والمشاعر
Machine learning	تعلم الآلة
Deep Learning	التعلم العميق

السيرة الذاتية للباحثين

	نوح صبري المتولي أستاذ مساعد – جامعة الجوف – قسم علوم الحاسب – كلية علوم الحاسب والمعلومات أستاذ مساعد – جامعة القاهرة – قسم علوم الحاسب – كلية الحاسبات والذكاء الاصطناعي
صورة المؤ 	أحمد السياط أستاذ مساعد – جامعة الجوف – قسم علوم الحاسب – كلية علوم الحاسب والمعلومات
صورة 	سعد العنزي أستاذ مساعد – جامعة الجوف – قسم علوم الحاسب – كلية علوم الحاسب والمعلومات

الخلاصة باللغة الانجليزية

A comprehensive study for Arabic Figurative Sentiment Analysis

Nouh Sabri Elmitwally[1], Ahmed Alsayat[2] and Saad Alanazi[3]

[1]Department of Computer Science, College of Computer and Information Sciences, Jouf University Aljouf, Saudi Arabia
[1]Department of Computer Science, Faculty of Computers and Artificial Intelligence, Cairo University, Egypt
[2]Department of Computer Science, College of Computer and Information Sciences, Jouf University Aljouf, Saudi Arabia
[3]Department of Computer Science, College of Computer and Information Sciences, Jouf University Aljouf, Saudi Arabia

[1] nselmitwally@ju.edu.sa, nouh.sabri@fci-cu.edu.eg

[2] asayat@ju.edu.sa

[3] sanazi@ju.edu.sa

Abstract

Arabic language processing works on multiple levels; less often, these complementary levels synergize well with each other. Arabic Language Sentiment Analysis (ALSA) levels consist of phonetics, morphology, syntax, lexicology, semantics, and figurative nature. The analysis of opinions and feelings is of interest in English and Indo-European languages, with little emphasis in the Arabic language, which is a language full of rhetorical characteristics and implicit meanings that have positive and negative connotations and meanings across the six linguistic levels. This paper presents a comprehensive and full proposal of a strategy for ALSA. The ALSA framework analyzes the opinions and feelings at all levels of language, in addition to the importance of building an annotated corpus, which helps to understand an Arabic sentence from the level of phonetics to the rhetorical and metonymy levels.

Keywords: *semantics; Arabic Sentiment Analysis; social sciences; machine learning; ALSA six-level framework*