أدوات كاملKamel tools لمعالجة النصوص العربية عبر مواقع التواصل الاجتماعي دراسة تحليلية في إطار البيانات الضخمة

إعداد

د. طاهر محمد محمد أمين

موجه عام الاعلام التربوي بالمنيا

مجلة البحوث في مجالات التربية النوعية

معرف البحث الرقمي DOI: 10.21608/JEDU.2024.312826.2101

المجلد العاشر. العدد54 . سبتمبر 2024

الترقيم الدولي

P-ISSN: 1687-3424 E- ISSN: 2735-3346

موقع المجلة عبر بنك المعرفة المصري https://jedu.journals.ekb.eg/

موقع المجلة http://jrfse.minia.edu.eg/Hom

العنوان: كلية التربية النوعية . جامعة المنيا . جمهورية مصر العربية

أدوات كاملkamel tools لمعالجة النصوص العربية عبر مواقع التواصل الاجتماعي دراسة تحليلية في إطار البيانات الضخمة

إعداد

طاهر محمد محمد أمين

دكتوراه الآداب اعلام جامعة المنيا

Tahrt.m.amin@gmail.com

مستخلص البحث:

يتناول هذا البحث تحليل المشاعر باستخدام أدوات "كامل" المتخصصة في معالجة النصوص العربية، مع التركيز على تطبيقات هذا التحليل في فهم والتعامل مع بعض القضايا المعاصرة. يستعرض البحث كيفية توظيف تقنيات الذكاء الاصطناعي لتحليل النصوص العربية، مع تسليط الضوء على التحديات التي تواجه معالجة اللغة العربية وأهمية تحسين دقة النتائج. كما يركز البحث على أمثلة واقعية تُظهر كيف يمكن لتحليل المشاعر أن يسهم في تقديم رؤى أعمق حول القضايا الاجتماعية والسياسية والاقتصادية، مما يساعد في تحسين عمليات صنع القرار وتطوير استراتيجيات فعالة للتعامل مع هذه القضايا.

يسعى البحث إلى تقديم رؤى معمقة حول كيفية تعبير النصوص عن المشاعر المتعلقة بهذه القضايا من خلال أدوات "كامل"، مما يساهم في فهم أفضل لتأثير الأحداث المعاصرة على العواطف العامة وتطوير استراتيجيات فعالة للتعامل مع القضايا الاجتماعية والإعلامية.

بناءً على التحليل، أظهرت نتائج تحليل المشاعر باستخدام أدوات "كامل" للنصوص العربية تباينًا واضحًا في ردود فعل الجمهور تجاه مختلف الأحداث. بالنسبة للحدث الأول المتعلق بطرد رجل وأسرته من المنزل، عبّرت غالبية التعليقات (66%) عن مشاعر إيجابية، حيث أظهر الناس تعاطفًا كبيرًا مع الأسرة ودعوات لتقديم المساعدة. في المقابل، أبدت 22% من التعليقات مشاعر سلبية، منتقدة الظروف الاقتصادية التي أدت إلى هذا الوضع، فيما كانت 12% من التعليقات محايدة.

وفيما يتعلق بالحدث الثاني، الذي يخص ابنة قتيل الإسماعيلية، أظهرت 70% من التعليقات مشاعر إيجابية، حيث تعاطف الجمهور مع العائلة وأعرب عن دعمه لتحقيق العدالة. بينما 24% من التعليقات كانت سلبية، حيث انتقدت العنف والجريمة، وكانت نسبة ضئيلة جدًا من التعليقات (1%) محايدة.

بالنسبة للحدث الثالث، الذي تمحور حول مواجهة أحمد العوضي وحسام بوجي، كانت أغلب التعليقات إيجابية (81%)، حيث دعم الجمهور موقف العوضي وتعاطف مع حسام بوجي في وجه التنمر. ومع ذلك، ظهرت 14% من التعليقات سلبية، حيث انتقدت بعض الجوانب المتعلقة بالحادث، فيما كانت 5% محايدة.

أما الحدث الرابع الذي يتعلق بإنهاء الثأر بين قبيلتين من خلال وليمة حضرها الآلاف، فقد أظهرت 67% من التعليقات مشاعر إيجابية تجاه هذه المصالحة، حيث رحّب الجمهور بروح التسامح وإنهاء النزاع. ومع ذلك، كانت 27% من التعليقات سلبية، منتقدة الخصومة الطويلة وتأثيرها السلبي على العائلات، بينما كانت 6% من التعليقات محايدة.

وأخيرًا، في الحدث الخامس الذي يتناول رسالة بالدموع من معلمة المنصورة، كانت المشاعر السلبية هي الغالبة (53%)، حيث انتقد الجمهور الموقف بشدة. ومع ذلك، عبّرت 24% من التعليقات عن مشاعر إيجابية، مع إظهار تعاطف مع المعلمة، فيما كانت 23% من التعليقات محايدة.

تعكس هذه النتائج تفاوتًا كبيرًا في كيفية استجابة الجمهور للأحداث المختلفة، حيث برزت المشاعر الإيجابية بشكل ملحوظ في حالات التعاطف والمصالحة، بينما كانت المشاعر السلبية أكثر وضوحًا في القضايا المثيرة للجدل أو المرتبطة بالعنف.

الكلمات الرئيسة:

معالجة اللغات الطبيعية– تحليل المشاعر– البينات الضخمة– أدوات كامل– البيانات الضخمة– مواقع التواصل الاجتماعي

Abstract

This research focuses on sentiment analysis using the "Kamel" tools specialized in Arabic text processing, emphasizing the applications of this analysis in understanding and addressing contemporary issues. The study reviews how artificial intelligence techniques can be employed to analyze Arabic texts, highlighting the challenges faced in processing the Arabic language and the importance of improving result accuracy. It also focuses on real-life examples showing how sentiment analysis can provide deeper insights into social, political, and economic issues, thereby helping to improve decision-making processes and develop effective strategies for addressing these issues.

The research aims to provide in-depth insights into how texts express sentiments related to these issues through "Kamel" tools, contributing to a better understanding of the impact of contemporary events on public emotions and the development of effective strategies for addressing social and media-related issues.

Based on the analysis, the results of sentiment analysis using "Kamel" tools for Arabic texts showed a clear variation in public reactions to different events. For the first event, related to the eviction of a man and his family from their home, the majority of comments (66%) expressed positive sentiments, showing significant empathy towards the family and calls for assistance. In contrast, 22% of comments showed negative sentiments, criticizing the economic conditions that led to this situation, while 12% of comments were neutral.

Regarding the second event, which concerns the daughter of the Ismailia murder victim, 70% of comments showed positive sentiments, as the public empathized with the family and expressed support for justice. Meanwhile, 24% of comments were

negative, criticizing violence and crime, with a very small percentage (1%) being neutral.

For the third event, which revolved around the confrontation between Ahmed Elawady and Hossam Bougy, most comments were positive (81%), as the public supported Elawady's position and empathized with Bougy in the face of bullying. However, 14% of comments were negative, criticizing some aspects related to the incident, while 5% were neutral.

The fourth event, which involved ending a feud between two tribes through a feast attended by thousands, showed that 67% of comments had positive sentiments towards this reconciliation, with the public welcoming the spirit of forgiveness and ending the conflict. Nevertheless, 27% of comments were negative, criticizing the long-standing feud and its negative impact on families, while 6% were neutral.

Finally, in the fifth event, which deals with a tearful letter from a teacher in Mansoura, negative sentiments predominated (53%), as the public strongly criticized the situation. However, 24% of comments expressed positive sentiments, showing empathy towards the teacher, while 23% of comments were neutral.

These results reflect a significant variation in public responses to different events, with positive sentiments notably emerging in cases of empathy and reconciliation, while negative sentiments were more apparent in contentious or violence-related issues

Key words

Natural Language Processing – NLP- Sentiment Analysis- Big Data- Kamel Tools- Social Media Platforms

مقدمة

تتسارع التطورات في مجال الذكاء الاصطناعي ومعالجة اللغات الطبيعية، مما يعزز قدرتنا على تحليل وفهم النصوص بشكل غير مسبوق. في هذا السياق، تبرز أهمية أدوات "كامل" لمعالجة النصوص العربية كأدوات رئيسية في فحص وتحليل المشاعر والتفاعلات البشرية عبر النصوص المكتوبة. توفر هذه الأدوات إمكانيات متقدمة في تحليل النصوص العربية، التي تتسم بتنوعها وثرائها اللغوي، مما يتطلب تقنيات متطورة للتعامل مع تعقيداتها.

تتمثل أهمية هذه الدراسة في تطبيق أدوات "كامل" لتحليل المشاعر على قضايا معاصرة ذات طابع اجتماعي، سياسي، واقتصادي، بهدف تقديم رؤى دقيقة حول كيفية تأثير هذه القضايا على الجمهور. من خلال تحليل التعليقات والردود المرتبطة بأحداث محددة، نسعى إلى فهم أعمق لكيفية استجابة المشاعر وتوجيه الاستراتيجيات الفعالة في التعامل مع القضايا الحساسة. سيسهم هذا البحث في تحسين تقنيات تحليل النصوص العربية وتقديم أدوات أكثر دقة وفعالية لفهم المشاعر البشرية في السياقات المختلفة.

استخدمت الدراسة أسلوب تحليل المشاعر (Sentiment Analysis) لتحليل التعليقات التي تم نشرها على مواقع التواصل الاجتماعي، وذلك لتحديد قطبية المشاعر إما إيجابية أو سلبيه والتي عبر عنها الجمهور تجاه الأحداث المختلفة التي تم تغطيتها، وتم ذلك عن طريق استخدام تقنيات التعلم الآلي والذكاء الاصطناعي لتحليل نصوص التعليقات وتصنيفها على مستوى الوثيقة المراد تحليلها.

في عصر التكنولوجيا الحديثة والذكاء الاصطناعي، أصبحت أدوات تحليل المشاعر تلعب دورًا محوريًا في فهم تفاعلات الجمهور وتوجهاتهم تجاه القضايا المختلفة. يأتي

هذا البحث ليلقي الضوء على تحليل المشاعر باستخدام أدوات "كامل" المتخصصة في معالجة النصوص العربية، مُركزًا على كيفية تطبيق هذه الأدوات في فهم وتناول القضايا المعاصرة ذات الأبعاد الاجتماعية والسياسية والاقتصادية.

تتمثل أهمية هذا البحث في تقديم رؤية شاملة لكيفية توظيف تقنيات الذكاء الاصطناعي لتحليل النصوص العربية، خاصةً في ظل التحديات التي تواجه معالجة اللغة العربية. يهدف البحث إلى تسليط الضوء على كيفية تحسين دقة النتائج التي توفرها هذه الأدوات، وكيفية استفادة الباحثين وصانعي السياسات من هذه التحليلات في تطوير استراتيجيات فعالة للتعامل مع القضايا المعاصرة.

من خلال تحليل المشاعر عبر أمثلة واقعية، يعرض البحث كيفية استخدام أدوات "كامل" للحصول على رؤى أعمق حول كيفية تعبير النصوص عن المشاعر المرتبطة بالأحداث المختلفة. كما يعرض البحث تباين ردود فعل الجمهور تجاه مجموعة من الأحداث، بدءًا من القضايا الاجتماعية مثل طرد أسرة من منزلها، وصولًا إلى المصالحة بين قبيلتين، وكيفية تأثير هذه الأحداث على المشاعر العامة.

تُظهر نتائج التحليل التباين الواضح في ردود فعل الجمهور، حيث تتفاوت المشاعر بين الإيجابية والسلبية بناءً على نوع الحدث وتأثيره. بينما كانت المشاعر الإيجابية بارزة في حالات التعاطف والمصالحة، كانت المشاعر السلبية أكثر وضوحًا في القضايا المثيرة للجدل أو المرتبطة بالعنف. هذا التباين يعكس أهمية أدوات "كامل" في تقديم بيانات دقيقة ومفيدة تساعد في فهم الديناميات الاجتماعية والتفاعل معها بطرق أكثر فعالية.

من خلال هذه الدراسة، نطمح إلى تعزيز الفهم لكيفية تأثير الأحداث المعاصرة على العواطف العامة، وتقديم رؤى تسهم في تحسين عمليات صنع القرار وتطوير استراتيجيات مناسبة للتعامل مع القضايا الاجتماعية والإعلامية.

تكمن أهمية هذا البحث أيضًا في تسليط الضوء على الأساليب المبتكرة التي يمكن استخدامها لتحليل المشاعر في النصوص العربية، وهي لغة تتميز بتعقيداتها الفريدة وتنوعها. أدوات "كامل" توفر إطارًا متقدمًا لمعالجة هذه التحديات، مما يساهم في تحسين دقة نتائج التحليل وتوسيع نطاق تطبيقاته في مجالات متعددة. إن معالجة النصوص العربية تتطلب فهماً عميقاً للغة وأساليب تحليل متطورة تتناسب مع خصوصياتها، وهذا البحث يسعى لتقديم حلول عملية لهذه التحديات.

تأتي هذه الدراسة في سياق متزايد من الاهتمام العالمي بتحليل المشاعر كأداة فعالة لفهم ردود فعل الجمهور والتفاعل مع القضايا المعاصرة. بالنظر إلى تنوع الأحداث التي تم تحليلها، يتضح كيف يمكن أن تلعب أدوات تحليل المشاعر دورًا رئيسيًا في تقديم رؤى متكاملة حول ردود فعل الجمهور وتعزيز فهمنا لكيفية تأثير الأحداث على العواطف والمواقف.

أهمية الدراسة تتمثل أهمية هذه الدراسة في تعزيز دقة وكفاءة تحليل المشاعر عبر النصوص العربية باستخدام أدوات "كامل". من خلال تحليل التعليقات والردود المرتبطة بقضايا معاصرة، توفر الدراسة رؤى قيمة حول كيفية استجابة الجمهور للأحداث الاجتماعية والسياسية والاقتصادية. هذا الفهم العميق يعزز القدرة على تطوير استراتيجيات تواصل وإعلام فعالة، ويدعم تحسين تقنيات الذكاء الاصطناعي في معالجة النصوص. كما تساهم الدراسة في تقديم معلومات جديدة حول الاتجاهات

الاجتماعية والتغيرات الثقافية، مما يساعد في توجيه السياسات واتخاذ القرارات بطريقة مستنيرة.

أولا الأهمية النظرية: –

تتمثل الأهمية النظرية لهذه الدراسة في تعميق الفهم الأكاديمي لعملية تحليل المشاعر في النصوص العربية من خلال أدوات "كامل". تسهم الدراسة في توسيع نطاق الأبحاث في مجال معالجة اللغة الطبيعية والذكاء الاصطناعي بتطبيق تقنيات متقدمة على اللغة العربية، مما يعزز المعرفة حول كيفية تعبير النصوص عن المشاعر وكيفية تصنيفها بدقة. كما توفر الدراسة إطاراً نظرياً لفهم التحديات والفرص المرتبطة بتحليل النصوص في سياقات ثقافية ولغوية محددة، مما يدعم تطوير نماذج وتحليلات أكثر تعقيداً وملاءمة للغة العربية. هذه الإسهامات النظرية تعزز من قاعدة المعرفة الأكاديمية في مجال تحليل النصوص وتفتح آفاقاً جديدة للبحث والتطبيق في معالجة اللغات الطبيعية.

ثانيا الأهمية التطبيقية

تقديم حلول عملية لتحسين تحليل المشاعر في النصوص العربية باستخدام أدوات "كامل"، مما يعزز من كفاءة ودقة النتائج في مجالات متعددة. تسهم الدراسة في تطوير أدوات تحليل النصوص لتكون قادرة على معالجة قضايا معقدة وذات طابع حساس في المجتمع العربي، مثل القضايا الاجتماعية والسياسية والاقتصادية. هذه التطبيقات تساعد المؤسسات الإعلامية، الشركات، والجهات الحكومية على فهم ردود أفعال الجمهور بشكل أفضل واتخاذ قرارات مبنية على بيانات دقيقة.

بالإضافة إلى ذلك، تتيح هذه الدراسة الفرصة لتطوير استراتيجيات تواصل أكثر فعالية تستند إلى تحليل دقيق للمشاعر العامة، مما يسهم في تحسين حملات التوعية الاجتماعية، وإدارة الأزمات، والتعامل مع القضايا الحساسة. كما يمكن تطبيق النتائج

في تصميم برامج دعم اجتماعي وسياسات حكومية تراعي مشاعر الجمهور وتلبي احتياجاتهم بشكل أفضل.

اهداف الدراسة

- تحليل المشاعر في النصوص العربية تطبيق أدوات "كامل" لتحليل المشاعر وتحديد كيفية تعبير النصوص العربية عن المشاعر الإيجابية، السلبية، والمحايدة.
- فحص فعالية الأدوات: تقييم مدى فعالية أدوات "كامل" في معالجة وتحليل النصوص العربية، وتقديم توصيات لتحسين أدائها في هذا السياق.
- تطبيق التحليل على قضايا معاصرة: استخدام تقنيات تحليل المشاعر لدراسة تأثير القضايا الاجتماعية والسياسية والاقتصادية المعاصرة على الجمهور، واستكشاف كيفية استجابة النصوص لهذه القضايا.
- تطوير استراتيجيات التواصل : تقديم رؤى تساعد في تحسين استراتيجيات التواصل والإعلام بناءً على نتائج تحليل المشاعر، مما يعزز من فعالية الرسائل والاستراتيجيات المتبعة.
- إثراء المعرفة الأكاديمية: تقديم إسهامات نظرية في مجال معالجة النصوص العربية والذكاء الاصطناعي، من خلال تعزيز فهم كيفية معالجة وتفسير النصوص في سياقات ثقافية ولغوية محددة.
- استكشاف التحديات والفرص: تحليل التحديات المرتبطة بتحليل النصوص العربية وتقديم حلول عملية لتحسين دقة التحليل وتوسيع نطاق استخدام الأدوات في أبحاث متعددة المجالات.

الدراسات السابقة

راعي الباحث ترتيب هذه الدراسات من الأحدث إلى الأقدم كما يلي: -

1- تهدف دراسة حسام فايز عبد الحي(2024)[1] إلى تحليل تفاعل الجمهور العربي مع مضامين الحرب على غزة 2023 عبر الصفحات الإخبارية على مواقع التواصل الاجتماعي، باستخدام تقنيات تحليل المشاعر ونمذجة الموضوعات. تم تحليل تعليقات الجمهور على صفحات الجزيرة مصر وArabic News BBC على فيسبوك من 7 أكتوبر 2023 إلى 23 نوفمبر 2023، باستخدام تقنيات معالجة اللغات الطبيعية ولغة بايثون. كشفت النتائج عن مشاعر إيجابية تضامنية تجاه فلسطين وغزة، مع تصدر الإعجابات كأكثر أشكال التفاعل. كما أظهرت نمذجة الموضوعات هيمنة ست موضوعات رئيسية، تشمل التضامن مع غزة والاعتداءات الإسرائيلية على المدنيين.

2- **دراسة (Singh, Maliha) 2023 [2] ركزت الدراسة على تحليل المشاعر** العامة والرأي فيما يتعلق بفيروس كرونا وقانون الإغاثة والأمن الاقتصادي، استخدم الباحث نموذج تخصيص **Dirichlet (LDA)** الكامن لنمذجة الموضوع، أظهرت نتائج تحليل المشاعر أن المشاعر العامة تحسنت بشكل كبير بعد إعلان قانون **Cares**، ولكن بعد ذلك

[1] عبد الرحيم، حسام فايز عبد الحي. (2024). تفاعل الجمهور مع المضامين المتعلقة بالحرب على غزة 2023 عبر الصفحات الإخبارية على مواقع التواصل الاجتماعي: دراسة في إطار البيانات الضخمة وفق أسلوبي تحليل المشاعر ونمذجة الموضوعات .*مجلة البحوث الإعلامية، ع69، ج3* ، 1445

[2] Singh, M. (2023). Public sentiment and opinion regarding the CARES act. Business Economics, 58(1), 24-33. doi:https://doi.org/10.1057/s11369-022-00299-8

تلاشى بعد أسبوع قبل التحسن مرة أخرى كتاريخ لتلقي فحص التحفيز الأول. يحدد الباحث بعض البرامج التي تلقت الدعم والانتقاد لقانون **Cares** من خلال نمذجة الموضوع، تقدم الدراسة نظرة ثاقبة على استجابة الجمهور لقانون **Cares** وتأثيرها على المشاعر العامة. هذه الدراسة مفيدة لصانعي السياسات والباحثين المهتمين بفهم رد فعل الجمهور على السياسات الحكومية وحزم التحفيز.

3- دراسة (Furqan Rustam) وآخرون 2023 [3] تدرس آراء الناس حول الوجبات السريعة وتحليلها باستخدام نمذجة **LDA** على مطاعم الوجبات السريعة باستخدام التغريدات غير المهيكلة. اعتمد البحث على استخدام تحليل البيانات لفهم الآراء والاتجاهات حول الوجبات السريعة. فمع ظهور منصات التواصل الاجتماعي، أصبحت مشاركة التقييمات معيارًا اجتماعيًا في مجتمع اليوم الحديث. يتحقق الأشخاص من آراء العملاء على مواقع الشبكات الاجتماعية حول مطاعم الوجبات السريعة والمواد الغذائية المختلفة قبل زيارة المطاعم وطلب الطعام. يمكن للمطاعم التنافس لتحسين جودة العناصر أو الخدمات المعروضة من خلال تحليل التعليقات المقدمة من العملاء بعناية. يميل الناس إلى زيارة المطاعم التي لديها عدد أكبر من التقييمات الإيجابية. وفقًا لذلك، يعد جمع التعليقات يدويًا من العملاء لكل منتج عملية تتطلب عمالة مكثفة؛ وينطبق الشيء نفسه على تحليل المشاعر. للتغلب على هذا، استخدمت الدراسة تحليل المشاعر، الذي يستخرج تلقائيًا معلومات مفيدة من البيانات. تركز الدراسات الحالية على نماذج التعلم الآلي. ونتيجة لذلك،

[3] Mujahid, M., Rustam, F., Alasim, F., Siddique, M., & Ashraf, I. (2023). What people think about fast food: Opinions analysis and LDA modeling on fast food restaurants using unstructured tweets. PeerJ Computer Science, doi:https://doi.org/10.7717/peerj-cs.1193

يتم إهمال تحليل أداء نماذج التعلم العميق في المقام الأول ونماذج المجموعات العميقة بشكل خاص. تحقيقًا لهذه الغاية، تتبنى هذه الدراسة العديد من نماذج المجموعات العميقة بما في ذلك الذاكرة ثنائية المدى طويلة المدى والوحدة المتكررة المسورة(**BilSTM+GRU**) ، و**LSTM+GRU**، و+**GRU**الشبكة العصبية المتكررة (**GRU+RNN**)، ونماذج **BilSTM+RNN** باستخدام التغريدات غير المهيكلة التي تم جمعها ذاتيًا. تتم مقارنة أداء الأساليب القائمة على المعاجم بنماذج المجموعات العميقة لتصنيف المشاعر. بالإضافة إلى ذلك، تستخدم الدراسة نمذجة تخصيص **Dirichlet** الكامن (**LDA**)لتحليل الموضوع. للتجارب، يتم جمع التغريدات الخاصة بأكبر خمس شركات تقدم الوجبات السريعة والتي تشمل كنتاكي فرايد تشيكن وبيتزا هت وماكدونالدز وبرجر كينج وصواي. تكشف النتائج التجريبية أن نماذج المجموعات العميقة تحقق نتائج أفضل من النهج القائم على المعاجم وأن **BILSTM+GRU** تحصل على أعلى دقة تبلغ 95.31٪ لثلاث مشاكل فئوية. تشير نمذجة الموضوع إلى أن أكبر عدد من المشاعر السلبية يتم تمثيله في مطاعم Subway بكلمات سلبية عالية الكثافة. يظل غالبية الناس (49٪) محايدين فيما يتعلق باختيار الوجبات السريعة، ويبدو أن 31٪ يحبون الوجبات السريعة بينما لا يحب البقية (20٪) الوجبات السريعة.

4- **دراسة حمزة خليل [4] 2021 توظيف تطبيقات الذكاء الاصطناعي لتحليل مشاعر مستخدمي مواقع التواصل الاجتماعي** في الوقت الفعلي

[4] خليل، حمزة السيد حمزة. (2021). توظيف تطبيقات الذكاء الاصطناعي لتحليل مشاعر مستخدمي مواقع التواصل الاجتماعي في الوقت الفعلي لأزمة جائحة فيروس كورونا، المجلة المصرية لبحوث الرأي العام.20(2)

لأزمة جائحة فيروس كورونا لقد أدت جائحة كورونا إلى نقاش عام ومهم على وسائل التواصل الاجتماعي، وبالتالي يمكن أن يساعد فهم هذه المناقشات المسئولين والأفراد علي تجاوز هذه الجائحة، ولذلك هدفت الدراسة إلي تحليل المشاعر العامة تجاه تفشي فيروس كورونا، وتحديد الموضوعات السائدة في المناقشات المتعلقة بالفيروس علي موقع تويتر. وطبقت الدراسة منهجية التعلم الآلي " Machine learning" في مجال الذكاء الاصطناعي "Artificial Intelligence" لتحليل البيانات التي تم جمعها من موقع تويتر الفترة من 1 مارس إلى 30 مايو 2020، وشملت عينة الدراسة 109154 تغريدة.

وقد أسفرت الدراسة عن ثلاثة جوانب رئيسية حول المشاعر العامة بشأن جائحة كورونا، أولاً- تم تقسيم اتجاهات المواطنين نحو انتشار وأعراض فيروس كورونا إلى ثلاث مراحل؛ ثانيًا- أظهرت نتائج تحليل المشاعر أن معظم المواطنين (الرقميين) لديهم مشاعر سلبية تجاه فيروس كورونا علي موقع تويتر، وجاءت مشاعر الخوف في المرتبة الأولي؛ ثالثًا- استنادًا إلى نمذجة الموضوعات تم تقسيم الموضوعات المتعلقة بـ "فيروس كورونا" علي تويتر -التي عبرت عن مخاوف المستخدمين- إلى خمس فئات رئيسية على النحو التالي:1) بيئة الرعاية الصحية؛ 2) والدعم النفسي والعاطفي،3) واقتصاد الأعمال؛ 4) والتغير الاجتماعي؛ 5) التوتر والإجهاد النفسي. وفي ضوء ذلك يمكن استخلاص الآتي: (أ) يمكن استخدام موقع تويتر كمدخل واعد ليعكس معرفة المواطن؛ (ب) يمكن استخدام التغريدات لتحديد تقلبات حدة الأزمات بمرور الوقت؛ (ج) تحليل المشاعر للتغريدات يؤكد صحة تطبيق رسائل التغريدات المحددة جغرافيا لتحديد المناطق شديدة التأثر بالأزمة.

5- دراسة خديجه عبدالله (2019)[5] بعنوان البيانات الضخمة والتحديات، وهدفت الدراسة الى التعرف عن قرب عن البيانات الضخمة وخصائصها وسماتها، بالإضافة الى التعرف على التحديات التي تواجه البيانات الضخمة في ظل التطور الحاصل في تكنولوجيا المعلومات والانفجار الرقمي، واعتمدت الدراسة على المنهج الوصفي، ومن اهم النتائج التي توصلت لها الدراسة أن البيانات الضخمة علم حديث ومتجدد ويحتاج الى موارد بشريه تستفيد منه ولكن قبل ذلك لابد من إيجاد المؤسسات التعليمية المتخصصة في مثل هذه العلوم فالبيانات الضخمة تحتاج الى مهارات في تكنولوجيا المعلومات والاتصالات والتحليل وغيرها من العلوم الأخرى، وحققت الشركات نجاحا خلال استغلال تحليلات البيانات الضخمة في تسويق منتجاتها وتوسيع نطاق مشترياتها، ومن أهم توصيات الدراسة التطرق اكثر الى البينات الضخمة في الدول العربية والتوجه نحو هذا الموضوع في الأبحاث والدراسات لتذكير المؤسسات والشركات بأهمية البيانات الضخمة، على المؤسسات التعليمية وضع خطط واستراتيجيات نحو إدخال هذا العلم ضمن تخصصها والتركيز عليه.

---

[5] اليحيائية، خديجة بنت عبد الله (2018، مارس6-8). البيانات الضخمة: التحديات] بحث مقدم] . المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربي البيانات الضخمة وآفاق استثمارها: الطريق نحو التكامل المعرفي. مسقط، المغرب.

6-دراسة (Stryk 2015)[6] تناولت هذه الدراسة موضوع البيانات الضخمة من جهة التعريف بأفضل ممارسات المسئولين عن معالجه البيانات الضخمة التي تم الاطلاع عليها وآليات تحليلها واستخراج المنافع منها، واعتمدت الدراسة على منهج دلفاي وهو منهج تفاعلي يعتمد على التواصل مع مجموعه من الخبراء من اجل استطلاع آرائهم حول موضوع معين، والتعرف على كيفية قيام المنظمات بإعداد البيانات وتنظيمها ليتم استخدامها بشكل مفيد في إدارة المنظمة من خلال مساعدة لمنظمات عل استخراج معلومات أكثر فعالية للحصول على قيمة مضافه لها، ومن أهم نتائج الدراسة أن السبب الرئيسي الذى دفع المنظمات للقيام بمعالجه البيانات الضخمة وتحليلها والاستثمار في ذلك لأجل تقليل المخاطر التنظيمية والتقنية وتحقيق أرباح على من هذه البيانات

مشكلة الدراسة

تتمثل مشكلة الدراسة في الحاجة إلى تحسين فهم كيفية تفاعل الجمهور العربي مع الأحداث والقضايا المعاصرة عبر مواقع التواصل الاجتماعي، وخاصة من خلال تحليل المشاعر المرتبطة بتلك التفاعلات. تتسم النصوص العربية بتنوعها وتعقيدها، مما يمثل تحدياً في تحليلها بدقة باستخدام أدوات معالجة النصوص. كما أن الدراسات السابقة تفتقر إلى تطبيقات متكاملة لتحليل المشاعر ونمذجة الموضوعات في السياقات الثقافية العربية، مما يجعل من الصعب تقديم رؤى دقيقة حول كيفية استجابة الجمهور للقضايا الكبيرة مثل النزاعات والحوادث الكبرى.

[6] Stryk, B. (2015). How do organizations prepare and clean Big Data to achieve better data governance? A Delphi Study.

تركز المشكلة على عدم وضوح كيفية تعبير الجمهور العربي عن مشاعره في سياقات معينة وكيفية تأثير هذه المشاعر على التفاعل مع المحتوى المتعلق بتلك القضايا. بالإضافة إلى ذلك، هناك نقص في الدراسات التي تستخدم تقنيات متقدمة مثل نمذجة الموضوعات وتحليل المشاعر بشكل منهجي على النصوص العربية لتقديم رؤى حول المشاعر العامة وتحليل التفاعلات في إطار البيانات الضخمة.

بالتالي، تسعى الدراسة إلى سد هذه الفجوة من خلال استخدام أدوات "كامل" لتحليل النصوص العربية، وتقديم نتائج دقيقة حول تفاعل المشاعر والتفاعل مع القضايا المعاصرة، مما يعزز فهم كيفية تأثير هذه القضايا على الجمهور ويوفر بيانات قيمة لصناع القرار والباحثين.

مجتمع الدراسة وعينتها:

أولا مجتمع الدراسة التحليلية: يتمثل مجتمع الدراسة التحليلية في صفحات الصحف المصرية على مواقع التواصل الاجتماعي، حيث تمتلك العديد من الصحف والمؤسسات الإخبارية في مصر صفحاتها الرسمية الخاصة على شبكات التواصل الاجتماعي، والتي يستخدمونها لمشاركة المقالات الإخبارية والتحديثات مع قرائها. تعليقات القراء على قضايا الأحداث الجارية بهدف تحليل المشاعر مع تلك القضايا للمساعدة في اتخاذ القرار المناسب لمساعدة صانع القرار فيما يراه مناسبا.

عينة الاحداث: تم اختيار خمسة احداث لتحليلها نتيجة لتفاعل الجمهور معها خلال الفترة من نوفمبر 2021 وحتى مارس 2022 لتصبح أشهر الأحداث متابعة وهي كالتالي:

- جريمة مذبحة الإسماعيلية: ". وشملت 21255 تعليق.

- طرد عائلة فقيرة لم تتمكن من دفع الايجار بسبب الضغوط الاقتصادية". وشملت 24552 تعليق.
- انتهاك حق الخصوصية: تأثير التصوير في الأماكن العامة وشملت 24958 تعليق.
- انتهاء خصومة عائلية دامت 17 عامًا في أسوان: بسبب الثأر وشملت 6736 تعليق. حادث التنمر الالكتروني على مواقع التواصل الاجتماعي: وشملت **24326** تعليق.

مبررات اختيار الاحداث الخمسة سالفة الذكر لتحليلها:

شهدت هذه الأحداث اعلى تفاعلات من جانب الجمهور الرقمي، واستمرار الحدث لفترة زمنية ليست قصيرة بالإضافة الى:

1. جريمة مذبحة الإسماعيلية:

حادث جنائي ذات طابع خطير لها تأثير كبير على الأمان العام والاستقرار في المجتمع، حيث شهدت الواقعة تعدى شاب صغير على رجل طاعن في السن وقام بذبح عنقه أمام الناس في وضح النهار مما الذعر بين الناس لبشاعة الحادث المخالف للتقاليد المصرية الطيبة.

2. مأساة طرد عائلة فقيرة لم تتمكن من دفع الايجار بسبب الضغوط الاقتصادية: يسلط الحدث الضوء على قضية اجتماعية حساسة تتعلق بالفقر وحقوق الإنسان. وشهد الحدث تفاعلا من جانب الجمهور ورغبتهم في مساعدة هذا الرجل

3. انتهاك حق الخصوصية وتصوير معلمة دون إذنها: تناول الحدث حقوق الفرد والأخلاقيات المهنية في سياق التكنولوجيا والتصوير في الأماكن العامة. وشهد الحدث

ردود فعل مختلفة ما بين قبول فكرة التصوير ورفض فكرة تصوير الأشخاص أثناء رحلاتهم.

4. انتهاء خصومة ثأرية في أسوان

- يمثل الحدث قصة نجاح في حل نزاع عائلي دام لفترة طويلة. وحل مشكلات الثأر

5. حادث التنمر الالكتروني على مواقع التواصل الاجتماعي: يعكس التحديات الاجتماعية والنفسية للتنمر في المجتمع.

التعريفات الإجرائية

البيانات الضخمة:Big Data –

هي تلك البيانات التي لا يمكن الحصول عليها وتخزينها وأدارتها بواسطة برمجيات قواعد البيانات التقليدية وان حجم هذه البيانات تنمو بمرور الوقت أو بالتقدم التكنولوجي وتظهر في تطبيقات مختلفة عن بعضها البعض[7]. والبيانات الضخمة هي مصادر البيانات ذات الحجم والسرعة الكبيرة وهي تمثل مجموعة متنوعة من البيانات التي تتطلب أدوات وطرق جديدة لاستخراجها وتنظيمها وأدارتها ومعالجتها بطريقة فعالة، ويعرفها الاتحاد الدولي للاتصالات بأنها نموذج لتمكين وجمع وتخزين وإدارة وعرض مجموعات واسعة من البيانات ذات الخصائص غير المتجانسة[8].

تحليل البيانات الضخمة Big Data Analysis: –

[7] Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, *19*(2), 171-209. https://doi.org/10.1007/s11036-013-0489-0
[8] Salem, F. (2017). The Arab World Online 2017: Digital Transformations and Societal Trends in The Age of the 4th Industrial Revolution.

مجموعة كبيرة جدا من البيانات التي يتم إنتاجها من قبل الأشخاص الذين يستخدمون الإنترنت، والتي لا يمكن تخزينها وفهمها واستخدامها إلا بمساعدة أدوات وأساليب خاصه[9]، كما أن هناك أدوات متعددة لتحليل البيانات الضخمة وما يقصده الباحث هنا أدوات تحليل مواقع التواصل الاجتماعي وسوف نذكر منها البعض

- **Nodxel**
- **IBM I2**
- **Orange**
- **weka**
- استخدام لغة برمجة بايثون**Python** وتحليلها باستخدام **Cloud Meaning**
- **Processing Language Natura**

ويري الباحث تحليل البيانات الضخمة هو التحليل الذي يتم من خلاله الحصول على المعلومات المهمة من البيانات الضخمة التي يصعب التعامل معها باستخدام الأدوات التقليدية ويستخدم في هذا التحليل الأدوات الذكية مثل الذكاء الصناعي والتعليم الآلي لتصحيح الأخطاء الإنسانية وتحليل التغيرات الزمنية.

[9] Cambridge university "big data" Cambridge online dictionary. https://dictionary.cambridge.org/dictionary/english/big-data?q=BIG+DATA

تحليل المشاعر (Sentiment Analysis):

تحليل المشاعر والمعروف أيضا باسم استخراج الرأي Opinion miningهو المهمة الأساسية لتصنيف نص أساسي آلياً أو تلقائياً، وتحديد ما إذا كان اتجاهه، أو ميله إيجابيا، أو سلبياً، أو محايداً[10]

تحليل المشاعر، المعروف أيضًا باسم التنقيب عن الرأي، هو استخدام معالجة اللغة الطبيعية، وتحليل النص، واللغويات الحسابية لتحديد واستخراج المعلومات الشخصية من المواد المصدر. غالبًا ما يكون ناتج تحليل المشاعر عبارة عن درجة أو قيمة ثنائية تشير إلى شعور النص (على سبيل المثال، إيجابي، سلبي، محايد). يستخدم تحليل المشاعر على نطاق واسع في معالجة اللغة الطبيعية واللغويات الحسابية لاستخراج المعلومات الذاتية من النص، ويستخدم على نطاق واسع في تطبيقات مثل مراقبة وسائل التواصل الاجتماعي، وأبحاث السوق، والتنقيب عن الرأي. يمكن أن تستند تقنيات تحليل المشاعر إلى مجموعة متنوعة من الأساليب، مثل استخدام التعلم الآلي أو الأساليب المستندة إلى القواعد أو الأساليب القائمة على المعجم.

[10] Oliveira, D. J., Bermejo, P.H, & Santos, P.A. (2016). Can social media Reveal the Preferences of Voters? A Comparison Between Sentiment Analysis and Traditional Opinion Polls. Journal Of Information Technology & Politics, 14(1), 34-45.

أدوات كامل

أدوات كامل CAMeL Tools هي مجموعة من الأدوات مفتوحة المصدر بلغة بايثون لمعالجة اللغة العربية آليا. توفر CAMeL Tools حاليًا عدة من التقنيات المُمَكِنة للمعالجة المسبقة والنمذجة الصرفية (التحليل والتوليد وإعادة التصريف) وتحديد اللهجة والتعرف على الكيانات المسماة وتحليل المشاعر. في هذا العرض التقديمي، نقدم بعض التحديات التي تواجه العمل مع اللغة العربية في نطاق الذكاء الاصطناعي وتحليل اللغات الطبيعية، ثم نَصِف فلسفة التصميم التي اتبعناها في بناء CAMeL Toolsوالوظائف التي توفرها.

مواقع التواصل الاجتماعي (Social Networking Sites): -

هي تطبيقات قائمة على الإنترنت تسهل إنشاء شبكات منظمة من الأفراد أو الكيانات أو الأشياء الافتراضية المرتبطة على أساس الخصائص الاجتماعية، وهي تسهل التفاعلات المتعددة بما في ذلك إنشاء المحتور ومشاركته من قبل المستخدمين[11]، وهي منظومة من الشبكات الإلكترونيّة التي تسمح للمشترك فيها بإنشاء موقع خاص به، ومن ثم ربطه عن طريق نظام اجتماعي إلكتروني مع أعضاء آخرين لديهم الاهتمامات والهوايات نفسها.

الإطار المعرفي

خلفية تاريخية عن البيانات الضخمة

ظهر مصطلح البيانات الضخمة في بدايات عام 2000م، وزادت استخداماته – حديثًا– حيث انتشر المصطلح ضمن أوائل الاتجاهات التقنية المقبلة، إذ توقعت

[11] Salem, F. (2017). The Arab World Online 2017: Digital Transformations and Societal Trends in the Age of the 4th Industrial Revolution.

أهميتها وانتشارها – مراكز الأبحاث التقنية مثل: وماكينزي، وشركة IBM ولم يخلُ الأمر من اهتمام دوائر سياسية كبيرة بهذا الموضوع التقني، مثل: إدارة الرئيس الأمريكي السابق، والمفوضية الأوربية التي عدت البيانات الضخمة رصيدًا أساسيا للاقتصاد، كما ركزت دوائر علمية أبحاثها في هذا المجال مثل: مؤسسة العلوم الوطنية الأمريكية، ومجلس الأبحاث الهندسية العلوم الطبيعية الكندي، ومعهد مهندسي الكهرباء والإلكترونيات الأمريكي، وبرنامج الأبحاث والابتكار الأوروبي، ومجلة الطبيعة، ومجلة العلوم، وقطاع الأعمال والاقتصاد مثل: منتدى "دافوس الاقتصادي"، كما انتشر الحديث عن البيانات الضخمة في وسائل الإعلام مثل : نيويورك تايمز، وول ستريت جورنال، والإيكونيميست، ومجلة الفيصل العلمية.[12]

ظهر الاهتمام بالبيانات الضخمة في الدول المتقدمة مثل الولايات المتحدة الأمريكية وشمال أوروبا.. ونظرا لتوافر أدبيات التخصص يتم تناول تعداد 1880م كأول مشكلات البيانات الضخمة بأمريكا الشمالية، أي ما يقرب من 40 عاما، تم جمع هذا التعداد باستخدام الوسائل والتكنولوجيا التي وضعها Herman Holerithالخبير الإحصائي، ومن هنا أثبتت تلك التكنولوجيا قدرتها على تقليل الوقت اللازم لجدولة التعداد من أكثر من ثماني سنوات (كما هو الحال لتعداد 1880م إلى سنة تعداد 1890م)، وأعلن عن مجموع السكان 62947711 بعد ستة أسابيع من التجهيز[13].

ثم ظهر قانون الضمان الاجتماعي 1935 في عهد في عهد الرئيس الأمريكي franklin d.roosevelts كواحدة من أكثر الجهود الطموحة في جمع البيانات على الاطلاق. وكانت IBMهي الشركة الفائزة المسئولة عن جمع وحفظ سجلات العمل

[12] محمد، الحسن شعبان احمد. (2021). *تحليلات البيانات الضخمة في المكتبات العامة المصرية: دراسة تحليلية للواقع والتخطيط للمستقبل*] أطروحة دكتوراه غير منشورة] ، جامعة المنيا، ص 43

[13] القصاص، علا نبيل (2020). مستقبل تكنولوجيا البيانات الضخمة، الهيئة المصرية العامة للكتاب، ص 17.

الخاصة ب 26 مليون أمريكى عامل و3 مليون موظف، حيث سجل على كل بطاقة اسم العامل وغيرها من المعلومات الهامة ومن خلال البطاقات، تم إنشاء كشاف الموظفين المؤهلين، وتم إنشاء حساب دفتر حساب لكل موظف مؤهل، ومن خلال تلك البطاقات وردت عدة تقارير، وكان الأجر مسجلا لكل عامل[14].

بدأ في الظهور البرمجيات المساعدة في عملية تخزين تلك البيانات الضخمة وكيفية التعامل معها، فظهر ما يسمى بالعملاق Colossus 1943ويعتبر أول آلة رقمية قابلة للبرمجة تم إنشاؤها من قبل المهندس توماس Thomas، فتكنولوجيا العملاق أثرت بشكل كبير على تطوير أجهزة الحاسوب كما ساعدت في تأكيد عملية السرعة الهائلة لأجهزة الكمبيوتر الإلكترونية.

أهمية البيانات الضخمة: –

تكتسب البيانات الضخمة أهميتها من كونها تُستخدم على نطاق أوسع في القطاع المصرفي؛ لفهم العملاء، وفي القطاع التعليمي؛ لتحديد الطلاب المعرضين للخطر، وتمكين الطلاب من تحقيق تقدم دراسي، وفي العمل الحكومي، مثل إدارة المرافق، والوكالات، والتعامل مع الازدحام المروري، أو منع جريمة، وفي القطاع الصحى: مثل إدارة سجلات المرضى، وإعداد الخطط العلاجية والوصفات الطبية، وفي قطاع التصنيع؛ لتحسين الانتاج والجودة[15].

[14] Zarate Santovena, A. (2013). Big data: evolution, components, challenges and opportunities [Master, Massachusetts Institute of Technology]. P1 http://hdl.handle.net/1721.1/80667

[15] Shaw, J. (2014). Why “Big Data” is a big deal. *Harvard Magazine*, *3*, 30-35.

شكل رقم (1) أهمية البيانات الضخمة [16]

مفهوم البيانات الضخمة: –

1) البيانات **(Data)**: تمثل الحقائق، وهي الأرقام أو التعليمات المقدمة في شكل مناسب؛ لفهمها وتفسيرها بما يتناسب مع عمليات الاتصال، أو المعالجة عن طريق الإنسان أو الحاسوب17
2) البيانات الضخمة: هي كميات من البيانات كبيرة الحجم، عالية السرعة، متنوعة الصيغ، بحيث تحتاج تكلفة باهظة، ومجهودًا ضخمًا؛ لإمكانية معالجتها باستخدام تقنيات قواعد البيانات، والبرمجيات التقليدية.18

خصائص البيانات الضخمة: –

وضع معهد SAS **(Statistical Analysis System)** خمس خصائص للبيانات الضخمة وهذا المعهد بمثابة مطور أمريكي متعدد الجنسيات مرتكز في شمال كارولينا يطور مجموعة برمجيات تحليلية لتسهيل الإتاحة **Access** الإدارة **Manage**

16 https://imarticus.org/5-ways-to-understand-the-importance-of-big-data/
17 Reitz, J. M. (2002). ODLIS: Online Dictionary of Library and Information Science. WCSU Libraries.
18 Gartner Glossary. Big Data. https://www.gartner.com/en/information-technology/glossary/big-data

التحليل **Analyze**والأخبار والتقرير **Report**عن البيانات المساعدة في اتخاذ القرارات وهي كالتالي:

1) **الحجم** (إن المنظمات تجمع البيانات من مصادر متنوعة بما في ذلك المعاملات التجارية ووسائط الإعلام أو أجهزة الاستشعار).
2) **السرعه** (إن تيارات البيانات في سرعة لم يسبق لها ويجب التعامل معها بطريقة حسنة التوقيت).
3) **التنوع** (إن البيانات الضخمة تأتى في جميع الأنواع والتنسيقات سواء وثائق نصية مهيكلة أو غير مهيكلة والبريد الإلكتروني والفيديو والصوت)
4) **التباين Variability** (تدفقات البيانات الضخمة غير متسقة بدرجة كبيرة)
5) **التعقيد Complexity** (تأتى البيانات الضخمة من مصادر متعددة، مما يجعل من الصعب الربط والنقل عبر عدة نظم)[19]

مجالات الاستفادة من البيانات الضخمة: –

1– القطاع الحكومي والبيانات الضخمة (Big Data): –

عندما يتعلق الأمر بإدارة البيانات، أغلب المنظمات الحكومية تواجه مشكلة وجود كميات هائلة من البيانات في أنظمة الحاسوب، ومعظم هذه البيانات غير منظمة أو مُهيكلة (unstructured data) وهذا يعني أنها لا تناسب أي نموذج بيانات معرّف مسبقاً. لفهم الأنماط الموجودة في هذه البيانات يجب أن تطبق المنظمات الحكومية نماذج إحصائية تسعى لالتقاط ومعالجة كميات هائلة من البيانات غير المهيكلة وتسمى هذه العملية بالبيانات الضخمة.

[19] Sas Institute. "Big Data What it Is and Why It Matters" Sas https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

أغلب المنظمات الحكومية لا تملك عدد كافي من الموظفين أو لا تملك القدرة الحسابية اللازمة، لإدارة وتحليل جميع البيانات الخاصة بهم، ومع طبيعة البيانات المتغيرة وارتفاع حجمها أصبحت الاستعانة بأدوات البيانات الضخمة من خلال الحوسبة السحابية **(Cloud Computing)** أمراً ضرورياً. فأصبح بإمكان المختصين بتطوير الخدمات الحكومية رصد مدى رضا المواطنين عن الخدمات المقدَّمة لهم. وعلى ضوء النتائج المحللة يمكن استنتاج ما يلزم عمله للتطوير والتحسين، حيث أصبح مسح آراء الجمهور عن طريق الاستبيانات التقليدية مكلفاً وغير مجدٍ في كثير من الأحيان، وذلك نظراً لتنوع البيانات الديموغرافية وثقافات المتعاملين. إن من أكبر المصادر لتلك البيانات الضخمة هي البيانات المسجلة من خلال عمليات التعداد السكاني والتسجيل في قواعد البيانات الحكومية، حيث يمكن أن تستنتج الحكومات معلومات ثمينة جداً من خلال تحليل تلك البيانات المخزنة.

2- تحليل البيانات الضخمة (Big Data) وتحسين التعليم: -

أدى استخدام أدوات التعلم عبر الإنترنت والبرامج القائمة على التفاعل بصورة متزايدة في مجال التعليم إلى زيادة حجم البيانات، واختلاف نوعية البيانات الكبيرة التي يُمكن جمعها من بيئات التعلم، فهنا نجد بيانات كبيرة عن المتعلمين، وخبرات التعلم لدى المتعلمين، كما نجد بيانات متعمقة داخل بيئات التعلم، وبيانات حول التفاعلات الاجتماعية في بيئات التعلم، وبيانات مُفصلة عن أنشطة التعلم من نصوص ووسائط ومقاطع فيديو وغيرها، كما تختلف هذه البيانات في نوعيتها وعمقها بنسب متفاوتة.

يمكن الاستفادة من تحليل هذه الأنواع من البيانات الضخمة في التعليم، لتوفير مجموعة مُتنوعة من الفرص والخيارات بهدف تحسين تعلم الطلاب من خلال التعلم التكيفي أو التعليم القائم على الكفاءة، مما ينتج عنه تعلم أفضل نتيجة لتشخيص أسرع

وأكثر تعمقاً في بيانات حقيقية تراكمية لاحتياجات التعلم أو المتاعب التي تواجهه أثناء عملية التعلم، بما في ذلك تقييم المهارات مثل التفكير المنظم، والتعاون، وحل المشاكل في سياق عميق، وتقييم أصيل لمجال وموضوع المعرفة، بالإضافة لتحديد التدخلات المستهدفة لتحسين نجاح الطلاب وخفض التكاليف الإجمالية للطلاب والمؤسسات، واستخدام البيئات القائمة والمعلومات المعقدة في صنع القرارات وتحديد السياسات، ويمكن أن توفر هذه البيانات أدوات حديثة وفعالة لقياس أداء الطلاب للمهام التعليمية، ويمكن أن تُساعد كذلك في تصميم بيئات تعلم تصميماً مُخصصاً وفق احتياجات مُحددة للطلاب، ويمكن أن تعطي تحليلاً واضحاً لردود الفعل الفردية والجماعية لمجموعة من القضايا التعليمية وغيرها من المميزات.

الدراسات الإعلامية والبيانات الضخمة

وعلي مستوى الدراسات الإعلامية يمكن أن تستفيد من البيانات الضخمة بعدة طرق، تتمثل إحدى الطرق في استخدام البيانات الضخمة لتحليل الأنماط والاتجاهات في استهلاك الوسائط. يمكن أن يساعد هذا الباحثين على فهم كيفية استهلاك مختلف الديموغرافيات للوسائط، وكيف تتغير أنماط استهلاك الوسائط بمرور الوقت. بالإضافة إلى ذلك، يمكن استخدام البيانات الضخمة لتحليل فعالية أنواع مختلفة من الإعلانات، والتي يمكن أن تساعد الشركات الإعلامية على إنشاء حملات تسويقية أكثر فاعلية. علاوة على ذلك، يمكن استخدام البيانات الضخمة لتتبع كيفية انتشار المعلومات على وسائل التواصل الاجتماعي، والتي يمكن أن تساعد الباحثين على فهم كيفية مشاركة الأفكار والرسائل ونشرها عبر الإنترنت. أخيرًا، يمكن استخدام البيانات الضخمة لتحليل تأثير الأخبار والمعلومات على الرأي العام، والتي يمكن أن تساعد الباحثين على فهم كيفية تشكيل وسائل الإعلام للتصورات والمواقف العامة.

هناك طريقة أخرى يمكن أن تستفيد بها الدراسات الإعلامية من البيانات الضخمة وهي استخدامها لتحليل الشعور بالمحتوى عبر الإنترنت ومنشورات وسائل التواصل الاجتماعي. يمكن أن يساعد ذلك الباحثين على فهم شعور الأشخاص تجاه مواضيع أو مشكلات أو منتجات معينة، ويمكن أيضًا استخدامه لتتبع التغييرات في المشاعر بمرور الوقت. يمكن أيضًا استخدام البيانات الضخمة لتحليل أداء محتوى الوسائط، مثل مقاطع الفيديو والمقالات والبودكاست، من خلال تتبع المقاييس مثل المشاهدات والمشاركات والمشاركة. يمكن أن يساعد هذا الشركات الإعلامية في تحديد أنواع المحتوى الأكثر شيوعًا وفعالية. بالإضافة إلى ذلك، يمكن استخدام البيانات الضخمة لتحليل تأثير العوامل المختلفة على صناعة الإعلام، مثل دور الخوارزميات في توصيات المحتوى، وتأثيرات الأخبار المزيفة، وظهور التسويق المؤثر. بشكل عام، توفر البيانات الضخمة ثروة من المعلومات التي يمكن استخدامها لاكتساب رؤى حول صناعة الوسائط والطرق التي يتم بها استهلاك الوسائط وإنشائها ومشاركتها.

هناك طريقة أخرى يمكن أن تستفيد بها الدراسات الإعلامية من البيانات الضخمة وهي استخدامها لتتبع أداء وسائل الإعلام الفردية والصحفيين ومنشئي المحتوى. يمكن أن يساعد ذلك الباحثين على فهم المنافذ الإعلامية والصحفيين ومنشئي المحتوى الأكثر تأثيرًا وفعالية. يمكن أيضًا استخدام البيانات الضخمة لتحليل انتشار المعلومات المضللة والأخبار المزيفة والمعلومات المضللة، والتي يمكن أن تساعد الباحثين على فهم كيفية انتشار هذه الظواهر وكيف يمكن مكافحتها.

يمكن أيضًا استخدام البيانات الضخمة لدراسة تأثير وسائل الإعلام على المجتمع. على سبيل المثال، يمكن للباحثين استخدام البيانات المتعلقة

باستهلاك الوسائط لدراسة العلاقة بين التعرض لوسائل الإعلام والنتائج المختلفة مثل الاستقطاب السياسي وتكوين الرأي العام والمشاركة المدنية.

علاوة على ذلك، يمكن استخدام البيانات الضخمة لتحليل تأثير التقنيات الجديدة مثل الواقع الافتراضي، والواقع المعزز، والذكاء الاصطناعي على صناعة الإعلام، وكيف يمكن استخدامها لإنشاء أنواع جديدة من التجارب الإعلامية.

بشكل عام، توفر البيانات الضخمة كمية غير مسبوقة من المعلومات التي يمكن استخدامها لاكتساب رؤى حول صناعة الوسائط والطرق التي يتم بها استهلاك الوسائط وإنشائها ومشاركتها. يمكن أن يساعد ذلك الباحثين والشركات الإعلامية وصناع السياسات على فهم المشهد الإعلامي بشكل أفضل واتخاذ قرارات أكثر استنارة.

مستقبل سوق البيانات الضخمة

سوق عمل البيانات الضخمة يشهد نموًا كبيرًا ومتسارعًا في السنوات الأخيرة، ومن المتوقع أن يستمر هذا النمو في المستقبل القريب. تزايد حجم البيانات وأهميتها في مختلف القطاعات يعني زيادة الطلب على المهارات والخبرات المتعلقة بتحليل البيانات الضخمة واستخدامها.

يتطلب سوق عمل البيانات الضخمة وجود محترفين متخصصين قادرين على التعامل مع حجم البيانات الكبير وتحليلها وتفسيرها بطرق فعالة. يتضمن ذلك المهارات الفنية مثل التعلم الآلي، وتحليل البيانات، والبرمجة، والإحصاء، وأيضًا المهارات الناعمة مثل قدرة التفكير التحليلي وحل المشكلات.

وتشمل فرص العمل في سوق البيانات الضخمة العديد من الوظائف مثل علماء البيانات (Data Scientists) ومهندسي البيانات (Data Engineers)ومحللي البيانات (Data Analysts) ومختصي تحليل السلوك

(Behavioral Analysts)وخبراء تعلم الآلة (Machine Learning)
(Experts)ومهندسي الذكاء الاصطناعي (Artificial Intelligence)
(Engineers)وغيرهم.

بالإضافة إلى ذلك، يمكن أن يتطلب سوق عمل البيانات الضخمة أيضًا احتياجات للمديرين والمستشارين الذين يمتلكون فهمًا استراتيجيًا للبيانات الضخمة ويمكنهم توجيه الأعمال واتخاذ القرارات الاستراتيجية بناءً على التحليلات.

من المتوقع أن يستمر الطلب على الخبرات المتعلقة بالبيانات الضخمة في التزايد في المستقبل، حيث تواصل الشركات والمؤسسات اعتماد تحليل البيانات للحصول على مزايا تنافسية واتخاذ قرارات أفضل. يتطلب ذلك توفر متخصصين مؤهلين ومتعلمين باستمرار لمواكبة التطورات التكنولوجية والاحتياجات المتغيرة في سوق العمل.

ثانيا المبحث الثاني:

التنقيب في الآراء وتحليل المشاعر

Researching Opinions and Analyzing Feelings

تحليل المشاعر وأفضل الطرق لتحليل النصوص العربية

يعد البحث عن الآراء وتحليل المشاعر إحدى الطرق التي يمكن من خلالها استخدام البيانات الضخمة لصالح الدراسات الإعلامية. كما ذكرنا سابقًا، يمكن استخدام البيانات الضخمة لتحليل الشعور بالمحتوى عبر الإنترنت ومنشورات الوسائط الاجتماعية. يمكن أن يساعد ذلك الباحثين على فهم شعور الأشخاص تجاه مواضيع أو مشكلات أو منتجات معينة، ويمكن أيضًا استخدامه لتتبع التغييرات في المشاعر بمرور الوقت.

يعد تحليل المشاعر أسلوب يستخدم معالجة اللغة الطبيعية والتعلم الآلي لتحديد واستخراج المعلومات الشخصية من النص، طريقة شائعة تستخدم لتحليل المشاعر المتعلقة بالمحتوى عبر الإنترنت، يمكن استخدامه لتصنيف النص على أنه إيجابي أو سلبي أو محايد، ويمكنه أيضًا تقديم رؤى حول المشاعر والآراء المعبر عنها في النص.

نـتـائـج الـدراسة

نموذج لتحليل المشاعر باعتبارها أحد روافد البيانات الضخمة

شكل رقم (1) يوضح إزالة الروابط والوسوم والأحرف الزائدة

أ- **التعلم الآلي**: تم استخدام تقنيات التعلم الآلي والذكاء الاصطناعي لتحليل النصوص (كل تعليق بمفرده) وتصنيفها إلى فئات مختلفة، مثل **إيجابي – وسلبي– ومحايد**، كما هو موضح بالشكل التالي

شكل رقم (98) يوضح تصنيف التعليقات الى (سلبي، إيجابي، محايد)

الخطوات التالية للتوضيح

قام الباحث باستخدام لغة بايثون واستخدام مكتبات متخصصة:

الخطوة الأولي: استخدم الباحث أداة **Tokenizers Word**

أداة **Tokenizer**هي أداة لتحليل النصوص وتقطيع النص الطويل إلى وحدات صغيرة، مثل الكلمات أو الجمل أو الفقرات.

تحسين جودة البيانات المدخلة في **tokenizer** يساهم في زيادة دقة وفعالية نتائج التحليل النصي الذي يتم بعدها، كما هو موضح نلاحظ تقطيع الجمل (التعليق الواحد) الى كلمات متقطعة

شكل رقم (1) يوضح تقطيع الجمل الى كلمات باستخدام أداة tokenizer

كما أنها أداة تستخدم في معالجة اللغة الطبيعية **(NLP)** لتحويل النص الحر إلى مجموعة من الكلمات المفردة التي يمكن معالجتها بشكل منفصل، وتم استخدام هذه الأداة في النصوص التي تم جمعها وهي من تطبيقات **NLP** مثل تصنيف النصوص واستخراج المعلومات والترجمة الآلية.

يمكن استخدام أداة **Tokenizers** لتحليل نصوص مختلفة مثل النصوص الخاصة بالأحداث الجارية المتنوعة سواء التعليمية والإخبارية والاجتماعية، والأحداث الخاصة بالجريمة، وتساعد هذه الأدوات في فهم النصوص واستخلاص المعلومات المهمة منها.

لاحظ الباحث من الممكن أن تكون أدوات **Tokenizers** متخصصة في عمليات معينة، مثل:

أ- **توكنايزر الكلمات:** وهو أداة تقوم بتقسيم النص الحر إلى مجموعة من الكلمات المفردة كما هو موضح(ناس،دي،لو ،كان،تجار ،مخدر ،ما ،كان،سيب،حكوم،تجيبل،لهم،حق،ناس،دي، لو ،كان،تجار ،مخدر ،كان،زمان،فيه،دم،سايح،ناس،كتير ،من ،أهل،قاتل،ميت،ده،كان،زمان،هربانيين،خايف،من،ثأر)

ب-توكنايزر الجمل: وهو أداة تقوم بتقسيم النص الحر إلى مجموعة من الجمل المفردة.

ت-توكنايزر المفردات: وهو أداة تقوم بتحديد الكلمات التي تشكل المفردات المستخدمة في اللغة، ويتم استخدام هذه الأداة عادةً في عمليات تصحيح الإملاء.

ث-توكنايزر الأرقام: وهو أداة تقوم بتحويل الأرقام في النص الحر إلى صيغة محددة.

يمكن استخدام التحليل اللغوي للنصوص **(NLP)** وتقنيات التعلم الآلي لتحليل المشاعر والمعاني المختلفة في اللغة العامية المصرية. يمكن بدء التحليل باستخدام

مكتبات **NLP** المتاحة مثل **NLTK** وإعداد مجموعة بيانات تحتوي على عينات من اللغة العامية المصرية.

لكن يجب ملاحظة أن اللغة العامية المصرية تحتوي على تعابير وألفاظ شائعة ومتداولة وغير رسمية قد تكون صعبة التحليل، لذلك يجب إجراء دراسات وتحليلات أكثر تفصيلًا لتحديد السياق الثقافي والاجتماعي واللغوي لهذه اللغة، واتضح من البحث ان اللغة الإنجليزية في التحليل لها مواردها وسهوله استخدامها

بعض أدوات **Tokenizers** الشهيرة في مجال معالجة اللغة الطبيعية والتي تم استخدامها تشمل:

أ- أداة **NLTK Tokenizer**: وهي أداة تستخدم في Python لتحويل النص الحر إلى مجموعة من الكلمات المفردة. تدعم هذه الأداة عدة أنواع من توكنايزر الكلمات، بما في ذلك توكنايزر الفراغات وتوكنايزر الكلمات المفتاحية وتوكنايزر الأنماط.

ب-توجد أيضًا أدوات **Tokenizers** أخرى مثل أداة **Moses Tokenizer** وأداة **Gensim Tokenizer** وغيرها. ويتم استخدام هذه الأدوات في مختلف التطبيقات في مجال معالجة اللغة الطبيعية.

الخطوة الثانية استخدم أداة إزالة الغموض (Disambiguation)

إزالة الغموض **(Disambiguation)** عملية تحديد المعنى الصحيح لكلمة أو عبارة في سياق معين، خاصةً في اللغات التي تحتوي على كلمات متعددة الدلالات.

يمكن تحقيق إزالة الغموض باستخدام الأساليب التالية:

أ- **استخدام السياق**: يمكن استخدام السياق الذي توجد فيه الكلمة الملتبسة لتحديد المعنى الصحيح. فعلى سبيل المثال، إذا كان النص يتحدث عن المجال البرمجي، يكون المقصود بكلمة "رمز" هو الرمز البرمجي.

ب-**الاستنتاج النحوي:** يتم استخدام قواعد اللغة النحوية لتحديد المعنى الصحيح للكلمة. فعلى سبيل المثال، يمكن استخدام التحليل النحوي لتحديد ما إذا كانت الكلمة تستخدم كفعل أو اسم، ومن ثم تحديد المعنى المناسب.

ت-**استخدام القواميس والمعاجم:** يمكن استخدام القواميس والمعاجم لتحديد المعنى الصحيح للكلمة، حيث يتم البحث في المصادر المعتمدة للحصول على معاني الكلمات، وبالفعل قام الباحث بإعداد قاموس يحتوي على أكثر من 5000 كلمة.

في هذه الخطوة اعتمد الباحث على

كلاً من (Maximum Likelihood Estimation) **MLE** و **BERT** في مهمة إزالة الغموض (Disambiguation) في معالجة اللغة الطبيعية.

طريقة **MLE** تقوم على تحديد الاحتمالات الأكثر توافقاً مع النص، حيث تقوم بحساب احتمالات الأحداث بناءً على توزيع الاحتمالات في النص، وبناءً على تاريخ الأحداث في مصدر النص، وتعتمد هذه الطريقة على تصميم نموذج إحصائي لحساب هذه الاحتمالات.

تستخدم تقنية **BERT** شبكات عميقة لتوليد تمثيلات للنصوص، حيث تعتمد على المدخلات المزدوجة والتعلم الذاتي لتوليد تمثيلات دقيقة للنصوص، ويمكن استخدامها في إزالة الغموض بناءً على تحليل السياق المحيط بالكلمة الملتبسة.

على الرغم من أن **BERT** تمتلك أداءً ممتازاً في مختلف مهام معالجة اللغة الطبيعية، فإن استخدام **MLE** قد يكون أكثر فعالية في مهمة إزالة الغموض، خاصةً في اللغات التي تتضمن تعدد المعاني وتحديداً اللغة العربية، يجب اختيار الأسلوب الأنسب للمهمة وفقًا لمتطلباتها وطبيعة اللغة المستخدمة.

بالنسبة لتطبيقات **MLE** و **BERT** في معالجة اللغة الطبيعية، يمكن استخدام **MLE**لحل مشكلات التصنيف والتنبؤ في المعالجة اللغوية، بينما يمكن استخدام **BERT**في مهام تحليل النص وتحديد العلاقات بين الجمل والمفردات.

يمكن استخدام **MLE** في تطبيقات مثل الترجمة الآلية، حيث يمكن استخدام النماذج الإحصائية لتحديد الاحتمالات الأكثر توافقًا مع النص، وبناء ذلك يمكن توليد الترجمة المناسبة. يمكن أيضًا استخدام **MLE** في تحليل النصوص، حيث يمكن تدريب نماذج إحصائية على الاحتمالات الأكثر توافقًا مع المعاني المختلفة للكلمات.

من ناحية أخرى، يمكن استخدام **BERT** في تطبيقات مثل تحليل المشاعر، حيث يمكن استخدامها لتحديد العلاقات بين الجمل والمفردات وتحليل السياقات اللغوية المختلفة. يمكن أيضًا استخدام **BERT** في مهام تحليل النصوص الطويلة وتحليل النصوص المكتوبة بلغات مختلفة.

بشكل عام، يمكن استخدام كلاً من **MLE** و **BERT** في تطبيقات معالجة اللغة الطبيعية المختلفة، ويمكن تحديد الأسلوب الأنسب للمهمة وفقًا للمتطلبات والظروف المحددة للمشروع.

الخطوة الثالثة بعد إجراء عملية الـ Tokenizer على النص، يمكن الاستمرار في معالجة اللغة الطبيعية باستخدام أدوات لتحليل المورفولوجيا

والتي تهتم بدراسة هيكل الكلمات وأشكالها المختلفة والعلاقات بينها، ويمكن تقسيم أدوات تحليل المورفولوجيا إلى عدة أنواع، مثل:

أ. المعالجة المورفولوجية القائمة على القواعد (rule-based morphology): وتستخدم هذه الأدوات مجموعة من القواعد اللغوية والتي

تم إنشاؤها يدويًا لتحليل الأشكال المختلفة للكلمات وتحديد الجذر واللواحق والبادئات وغيرها من الأشكال المختلفة للكلمات.

ب. المعالجة المورفولوجية القائمة على القاموس (dictionary-based morphology): وتستخدم هذه الأدوات قوائم الكلمات الموجودة في القاموس والتي تحتوي على معلومات المورفولوجيا الخاصة بالكلمات مثل الأشكال المختلفة للكلمات والجذور واللواحق والبادئات وغيرها، وبالتالي تمكن الأداة من تحديد المعلومات المورفولوجية للكلمات الموجودة في النص.

ت. التحليل المورفولوجي الاحتمالي (probabilistic morphology analysis): وتستخدم هذه الأدوات النماذج الإحصائية والذكاء الاصطناعي لتحليل الأشكال المختلفة للكلمات وتحديد المعلومات المورفولوجية الخاصة بها.

بعد تحليل المورفولوجيا، يمكن الاستمرار في معالجة اللغة الطبيعية بواسطة أدوات لتحليل النحو، والتي تهتم بدراسة العلاقات بين الكلمات وتحديد دور كل كلمة في الجملة.

الخطوة الرابعة استخدام Morphtokenegy

عبارة عن محلل نحوي ومورفولوجيا مفتوح المصدر للغة العربية، يتم استخدامه لتحليل النصوص وفهم بنية الجمل والكلمات وعلاقاتها في النص، يستخدم **MorphTokenEGy** قواعد لغوية متعددة ومتنوعة لتحليل النصوص، ويستند على خوارزمية معالجة اللغة الطبيعية والتعلم الآلي.

يستخدم **MorphTokenEGy** لتحليل النصوص في مجالات متعددة مثل التعليم والترجمة الآلية ومعالجة اللغة الطبيعية وغيرها، كما يمكن استخدامه في تطوير التطبيقات اللغوية ومحركات البحث والذكاء الاصطناعي، يتميز **Morphtokenegy**

بسهولة الاستخدام والتكامل مع التطبيقات الأخرى، كما يمكن تخصيصه وتعديله لتلبية احتياجات محددة لمشروع معين.

جدول رقم (1) يوضح وصف عام للأحداث

منشور رقم	1	2	3	4	5
وصف	قاعد في عز البرد هو وبناته على الرصيف يبكي بعد طرده من المنزل	ابنة قتيل الإسماعيلية تكشف ما حدث وعلاقة القاتل بوالدها	أحمد العوضي وجهاً لوجه مع حسام بوجي	وليمة لـ 10000 شخص تنهى سلسال الدم بين قبيلتين	رسالة بالدموع من معلمة المنصورة
أعجبني	37k	61k	261k	35k	69k
أحببت	6.1k	8.9k	164k	3.9k	12k
أحزنني	25k	34k	1.1k	0.03k	9.2k
ادعمه	0.8k	0.6k	18k	0.5k	1.9k
أضحكني	0.07k	0.4k	1.1k	0.6k	2.3k
أكره	1k	1.5k	0.4k	0.05k	2k
مشاهدات	10m	18m	19m	4m	9.3m

لاحظ الباحث

كما هو موضح من الشكل تنوعت الأحداث التي تفاعل معها الشعب المصري ما بين قضايا اجتماعية وقضايا وجرائم، وهذه الأحداث متنوعة التفاعل على حسب سياق الحدث، فنسب المشاهدات مختلفة ونسبة المتابعة لكل حدث مختلف، وتختلف الطريقة التي يتفاعل بها المصريون مع هذه الأحداث حسب سياق الحدث. على سبيل المثال، قد يكون الناس أكثر عرضة للتفاعل مع حدث يحدث في مجتمعهم أو يؤثر عليهم شخصيًا، قد يكونون أيضًا أكثر عرضة للتفاعل مع حدث تتم مناقشته في وسائل الإعلام أو يتم التحدث عنه من قبل أصدقائهم وعائلاتهم.

ولابد من ملاحظة أنه لن يتفاعل جميع المصريين مع هذه الأحداث بنفس الطريقة، قد يهتم بعض الأشخاص بأحداث معينة أكثر من غيرهم، قد يكون بعض الأشخاص أكثر نشاطًا في محاولة تغيير الأشياء، بينما قد يكون البعض الآخر أكثر سلبية.

الطريقة التي يتفاعل بها الشعب المصري مع الأحداث معقدة ومتنوعة. يتأثر بعدد من العوامل، بما في ذلك تجارب الفرد الشخصية ومعتقداته وقيمه.

يعرض الجدول خمس منشورات متنوعة تغطي مجموعة من المواضيع الاجتماعية والإنسانية. يبرز المنشور الثالث، الذي يتعلق بالممثل أحمد العوضي، كأكثر المنشورات شعبية من حيث المشاهدات والتفاعل. حقق هذا المنشور 19 مليون مشاهدة، مع 261 ألف إعجاب و164 ألف تعليق، مما يشير إلى الاهتمام الكبير بالشخصيات العامة والمشاهير في وسائل التواصل الاجتماعي.

القضايا الاجتماعية والتعاطف العام

المنشور الثاني، الذي يتناول قصة ابنة قتيل الإسماعيلية، أثار تعاطفًا كبيرًا، حيث حصل على أعلى نسبة مشاركات بلغت 34 ألف. هذا يدل على أن القضايا الاجتماعية الحساسة تحرك مشاعر الجمهور وتدفعهم للمشاركة والتفاعل. بالمثل، المنشور الأول عن شخص مشرد يبكي على الرصيف حقق تفاعلاً ملحوظًا، مما يؤكد على تأثير القصص الإنسانية المؤثرة.

التباين في الاهتمامات والتفاعلات

يظهر تباين واضح في الاهتمام بين المنشورات. فبينما حظي منشور أحمد العوضي باهتمام كبير، نجد أن المنشور الرابع عن وليمة لـ 10000 شخص حصل على أقل عدد من المشاهدات والتفاعلات. هذا قد يشير إلى أن الأحداث الاجتماعية العامة، رغم

أهميتها، قد لا تثير نفس مستوى الاهتمام مقارنة بالقضايا الشخصية أو الإنسانية المؤثرة.

العاطفة والتأثير

المنشور الخامس، وهو رسالة عاطفية من المنصورة، حقق نسبة عالية من الإعجابات مقارنة بعدد المشاهدات. هذا يشير إلى أن المحتوى العاطفي يمكن أن يحدث تأثيرًا قويًا على المتابعين، حتى لو لم يصل إلى جمهور واسع كالمنشورات الأخرى.

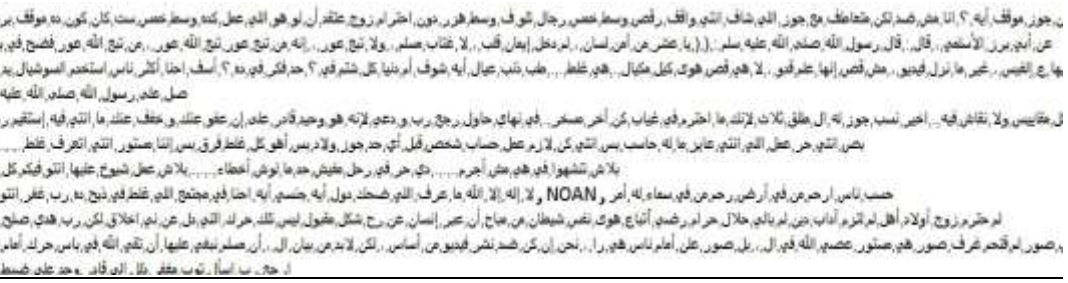
من خلال هذا التحليل، نستنتج أن الجمهور يميل للتفاعل بشكل أكبر مع المحتوى الذي يثير العواطف أو يتعلق بشخصيات مشهورة. القضايا الاجتماعية والإنسانية تحظى باهتمام كبير وتشجع على المشاركة. في حين أن الأحداث العامة قد تجذب الانتباه، إلا أنها لا تحقق نفس مستوى التفاعل مقارنة بالقصص الشخصية أو القضايا المؤثرة عاطفيًا.

جدول رقم (2) يوضح نتائج تحليل المشاعر بناء على الخطوات السابق ذكرها [20]

منشور رقم	1	2	3	4	5
وصف أو مضمون الحدث	قاعد في عز البرد هو وبناته على الرصيف	ابنة قتيل الإسماعيلية تكشف ما حدث وعلاقة القاتل بوالدها	أحمد العوضي وجهاً لوجه مع حسام بوجي	وليمة لـ 10000 شخص تنهى سلسال الدم	رسالة بالدموع من معلمة المنصورة

[20] الاحداث الخمس
https://www.facebook.com/watch/live/?ref=watch_permalink&v=636190177377345
https://www.facebook.com/Youm7/videos/1061927104598529/
https://www.facebook.com/133880733349264/posts/8183254755078448/
https://www.facebook.com/AlmasryAlyoum/videos/704124204015956/
https://www.facebook.com/watch/live/?ref=watch_permalink&v=229418499226333

		يبكي بعد طرده من المنزل			بين قبيلتين	
تعليقات	إجمالي	855	2900	2900	290	290
	إيجابي	%66	%70	%81	%67	%24
	سلبي	%22	%24	%14	%27	%53
	محايد	%12	%1	%5	%6	%23
إجمالي	النسبة المئوية	%100	%100	%100	%100	%100



شكل رقم يوضح نموذج من التعليقات الإيجابية والسلبية

تعقيب الباحث

الحدث الأول ابنة قتيل الإسماعيلية وبنفس الطريقة على الأحداث الأخرى.

يمكن رصد بعض العواطف والمشاعر في النص الخاص بابنة قتيل الإسماعيلية. ومن بين العواطف التي يمكن تحديدها في هذا النص: (الحزن، الصدمة، الغضب، القلق، الرغبة في العدالة، الإحباط، الحسرة، والرجاء للرحمة والصبر)

هنا بعض التوضيحات حول تحليل المشاعر في النص:

1. **الحزن:** يتجلى في عبارات مثل "بكاء شديد" و"صدمة ياناس" و"صبركم" و"رحمة" و"سلوان" و"راجعون"، حيث يعبر المعلقين عن الحزن والألم الذي يشعروا به بسبب وفاة أبيها ويعزي أهلها وذويها.
2. **الصدمة:** يظهر في العبارات مثل "ذهول وصدمة" و"مشاهدة الفيديو"، مما يشير إلى استغراب الجمهور وصدمته عندما سمع خبر الوفاة وشاهد مقطع الفيديو المرتبط به.
3. **الغضب:** يتجلى في عبارات مثل "مشتمينا" و"القاتل" و"اتقوا الله شوية" و"حسبي الله ونعم الوكيل"، حيث يعبر الكاتب عن الغضب والاستياء من القاتل والتجاوزات التي يراها في المجتمع.
4. **القلق:** يعبر الكاتب عن القلق بشأن أهل المتوفى وحقهم في العدالة، ويدعو إلى إجراء تحقيق شرعي والاهتمام بشكل ملائم بالأسرة المتضررة.
5. الرجاء للعدالة: يعبر الكاتب عن رغبته في تحقيق العدالة والقصاص للمتسببين في الوفاة، ويطالب بإجراء التحقيقات اللازمة لتحقيق العدالة لأفراد الأسرة المتضررة.
6. **الإحباط والحسرة:** يظهر الإحباط والحسرة في بعض العبارات مثل "مش متحمسين لاخد حقهم" و"متضايقاني قد ناس" و"دي مش قضيتكم بس"، حيث يشعر الكاتب بالإحباط من عدم اهتمام البعض بالموضوع وعدم حصول العدالة المطلوب
7. **الحزن العميق:** يتجلى الحزن العميق في عبارات مثل "قلبي محطم" و"لم أستطع التعبير عن ألمي" و"الدموع لا تتوقف"، وهذا يعكس مدى الألم والحزن الذي يشعر به الكاتب بسبب فقدان أبيه. إنه يعبر عن مشاعر الحزن والتجاهل للفراغ الذي تركه أبوه في حياته.

8. **الحنين والشوق:** يتضح من النص أن الكاتب يشعر بالحنين والشوق لأبيه، حيث يذكر "الله يرحمك يا بابا" و"أشتاقك"، وهذا يظهر رغبته في وجود أبيه مرة أخرى وتبادل اللحظات الجميلة معه.
9. **التفكير في الماضي:** يوحي النص بأن الكاتب يفكر في الأوقات الماضية التي قضاها مع أبيه، ويرغب في إحضار ذكريات تلك اللحظات واسترجاعها، وذلك من خلال العبارات مثل "أتذكر يا بابا" و"ذكريات جميلة"، مما يعكس رغبته في الاحتفاظ بالذكريات الجميلة والتواصل معها.

ومن الممكن أيضا تصنيف المشاعر وفقا لنموذج روبرت بلوتشيك (Robert Plutchik)[21] هو نموذج لتصنيف المشاعر يعتمد على ثمانية مشاعر أساسية، حيث يعتبر الشعور بالمشاعر كمركب من هذه المشاعر الأساسية. إليك قائمة بالمشاعر الأساسية في نموذج بلوتشيك:

- السعادة (Joy): شعور السرور والرضا والسعادة.
- الحزن (Sadness): شعور الحزن والألم والتألم.
- الغضب (Anger): شعور الغضب والاستياء والغضب.
- الخوف (Fear): شعور الخوف والقلق والهلع.
- الاشمئزاز (Disgust): شعور الاشمئزاز والاستنكار والاشمئزاز.
- الثقة (Trust): شعور الثقة والأمان والاعتماد.
- المفاجأة (Surprise): شعور الدهشة والمفاجأة والاستغراب.
- الاستنزاف (Anticipation): شعور التوقع والترقب والتحضير.

وفقًا لنموذج بلوتشيك، يُمكن تجسيد العديد من الشعور بالمشاعر المختلفة عن طريق تركيب هذه المشاعر الأساسية.

[21] استخدم حمزة خليل هذا النموذج في رصد المشاعر تجاه جائحة كورونا

وجدير بالذكر تم إنشاء سحب للكلمات باستخدام كلمات المشاعر الأكثر تكراراً والتي تم تصنيفها باستخدام كلمات المشاعر الأكثر استخداما على سبيل المثال:

شكل رقم يوضح سحابة الكلمات الخاصة بأحد الاحداث

يوضح سحابة الكلمات الخاصة بأحد الاحداث

يوجد عدد من التوصيات الخاصة بهذه الدراسة: –

- **تحسين أدوات التحليل**

– تطوير أدوات تحليل المشاعر العربية لتحسين الدقة.

– تحسين معالجات اللغة الطبيعية للغة العربية بتوسيع قاعدة البيانات اللغوية

- **توسيع نطاق الدراسة والتطبيقات**

- دراسة تأثير الاختلافات الثقافية والجغرافية في تحليل المشاعر.
- تطبيق التحليل على مزيد من القضايا الاجتماعية والسياسية المعاصرة.

- **تعزيز التعاون بين التخصصات**

تشجيع التعاون بين اللغويين ومهندسي الحاسوب لتحسين الأدوات.

تعزيز البحث بين فرق العمل متعددة التخصصات.

- **تعزيز التطبيقات العملية والتدريب**

اقتراح استخدام التحليل في المؤسسات الإعلامية والمجال التجاري.

توفير ورش عمل وبرامج تدريبية على أحدث الأدوات والتقنيات.

المراجع

أولا المراجع العربية

1. عبد الرحيم، حسام فايز عبد الحي. (2024). تفاعل الجمهور مع المضامين المتعلقة بالحرب على غزة 2023 عبر الصفحات الإخبارية على مواقع التواصل الاجتماعي: دراسة في إطار البيانات الضخمة وفق أسلوبي تحليل المشاعر ونمذجة الموضوعات. مجلة البحوث الإعلامية.
2. خليل، حمزة السيد حمزة. (2021). توظيف تطبيقات الذكاء الاصطناعي لتحليل مشاعر مستخدمي مواقع التواصل الاجتماعي في الوقت الفعلي لأزمة جائحة فيروس كورونا، المجلة المصرية لبحوث الرأي العام.20(2)
3. عزب، عمرو محمد. (2022). العوامل المؤثرة في بناء أجندة القضايا الاجتماعي في الصحافة المصرية المطبوعة والإلكترونية،] رسالة دكتوراه غير منشورة] ، جامعة القاهرة.

4. راشد، رائد صلاح الدين. (2021). دور صحافة الموبايل في تشكيل اتجاهات الجمهور المصري نحو الأحداث الجارية،] أطروحة ماجستير غير منشورة] ، جامعة المنصورة.

5. احمد، مي عمر سعيد. (2022) التأثيرات المتبادلة بين أجندة الصحافة الدولية والنقاشات على تويتر لانتخابات الرئاسة الأمريكية،] أطروحة ماجستير غير منشورة] ، جامعة المنيا.

6. اليحيائية، خديجة بنت عبد الله (2018، مارس6-8). البيانات الضخمة: التحديات] بحث مقدم] . المؤتمر الرابع والعشرون لجمعية المكتبات المتخصصة فرع الخليج العربي البيانات الضخمة وآفاق استثمارها: الطريق نحو التكامل المعرفي. مسقط، المغرب.

7. محمد، الحسن شعبان احمد. (2021). *تحليلات البيانات الضخمة في المكتبات العامة المصرية: دراسة تحليلية للواقع والتخطيط للمستقبل*] أطروحة دكتوراه غير منشورة] ، جامعة المنيا، ص 4

8. 1 القصاص، علا نبيل (2020). مستقبل تكنولوجيا البيانات الضخمة، الهيئة المصرية العامة للكتاب.

9. 1- فؤاد، أحمد (2013). مبيعات أمازون. 426 سلعة بالثانية. متاح من خلال سكاى نيوز http://cutt.us/Yzx7y

10. عبد اللطيف، محمود رمضان أحمد. (2021). تبني المؤسسات الصحفية المصرية تكنولوجيا الذكاء الاصطناعي في إنتاج وتحرير الأخبار والموضوعات الصحفية: في ضوء تجارب بعض الصحف الأجنبية .المجلة المصرية لبحوث الرأي العام، مج20، ع3 ، 1 .68 -

11. عبد الحميد، عمرو محمد محمود. (2020). توظيف تطبيقات الذكاء الاصطناعي في إنتاج المحتوى الإعلامي وعلاقتها بمصداقيته لدى الجمهور المصري مجلة البحوث الإعلامية، ع55، ج5.

12. الشريف، سلوى أحمد محمد أبو العلا. (2022). تحليل مشاعر تغريدات توتير أثناء الانتخابات الرئاسية الأمريكية 2020 باستخدام إطار البيانات الضخمة .المجلة المصرية لبحوث الرأي العام، مج21، ع2.

ثانيا المراجع الاجنبية

1. Singh, M. (2023). Public sentiment and opinion regarding the CARES act. Business Economics, 58(1), 24-33. doi:https://doi.org/10.1057/s11369-022-00299-8
2. Mujahid, M., Rustam, F., Alasim, F., Siddique, M., & Ashraf, I. (2023). What people think about fast food: Opinions analysis and LDA modeling on fast food restaurants using unstructured tweets. PeerJ Computer Science, doi:https://doi.org/10.7717/peerj-cs.1193
3. Faliang, H., Changan, Y. Yingzhou, B., Lu, J., Liqiong, L., & Wang, X. (2022). Multi-granular document-level sentiment topic analysis for online reviews. Applied Intelligence, 52(7), 7723-7733. doi:https://doi.org/10.1007/s10489-021-02817-1 Stryk, B. (2015). How do organizations prepare and clean Big Data to achieve better data governance? A Delphi Study.
4. Zarate Santovena, A. (2013). Big data: evolution, components, challenges and opportunities [Master, Massachusetts Institute of Technology]. P1 http://hdl.handle.net/1721.1/80667
5. Zarate Santovena, A. (2013). Big data: evolution, components, challenges and opportunities [Master, Massachusetts Institute of Technology]. P15 http://hdl.handle.net/1721.1/80667
6. Shaw, J. (2014). Why "Big Data" is a big deal. *Harvard Magazine*, *3*, 30-35.
7. https://imarticus.org/5-ways-to-understand-the-importance-of-big-data/
8. Mutula, S. (2016). Big Data Industry: Implication for the Library and Information Sciences. African Journal of Library, Archives & Information Science, 26(2), 93.

9. Bertot, J. C., Butler, B. S., & Travis, D. M. (2014). Local big data: The role of libraries in building community data infrastructures. Proceedings of the 15th annual international conference on digital government research.
10. Wang, C., Xu, S., Chen, L., & Chen, X. (2016). Exposing library data with big data technology: A review. 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS),
11. Reitz, J. M. (2002). ODLIS: Online Dictionary of Library and Information Science. WCSU Libraries.
12. Gartner Glossary. Big Data. https://www.gartner.com/en/information-technology/glossary/big-data
13. Dictionaries,O.(2017).Englishoxford‹living‹dictionaries.: https://en.oxforddictionaries.com/definition/big_data
14. Merriam-Webster.(2017).bigdata.From https://www.merriamwebster.com/dictionary/big%20data
15. Sas Institute. "Big Data What it Is and Why It Matters" Sas https://www.sas.com/en_us/insights/big-data/what-is-big-data.html
16. Demchenko, Y., Ngo, C., de Laat, C., Membrey, P., & Gordijenko, D. (2013). Big security for big data: Addressing security challenges for the big data infrastructure. Workshop on secure data management, - UNECE, H. (2013). What Does 'Big Data'Mean for Official Statistics? *Translated by High-level Group on the Modernization of Statistical Production and Services*.
17. Solon, O. (2014). A simple guide to Care. data. *Wired*. https://www.wired.co.uk/article/a-simple-guide-to-care-data
18. Banjo, S. (2014). Wal-Mart Notches Web Win Against Rival Amazon. *Wall Street Journal*. https://www.wsj.com/articles/SB10001424052702303417104579544080308837964

19. zeforia. (2018). *The Top 15 Valuable Facebook Statistics – Year in Review*. https://yoursuperinspector.com/wpcontent/uploads/2018/09/Top_20_stats_for_facebook.pdf

20. Infragistics Adds Predictive Analytics, Machine Learning and More to Reveal Embedded Business Intelligence Tool: Updates allow users to see predictions based on historical trends and access big data sources including Azure SQL, Azure Synapse, Google Big Query, and Salesforce. (2020, Apr 03). *NASDAQ OMX's News Release Distribution Channel* https://www.proquest.com/wire-feeds/infragistics-adds-predictive-analytics-machine/docview/2385727227/se-2

21. Leppänen, Leo, Hanna Tuulonen, and Stefanie Sirén-Heikel, (2020). "Automated Journalism as a Source of and a Diagnostic Device for Bias in Reporting."**Media and Communication** Volume 8, Issue 3P p 39–49

22. 1 Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. Journal of Computational Science, 36, 101003. doi:

23. https://doi.org/10.1016/j.jocs.2019.05.009

24. Farhadloo M, Rolland E (2016) Fundamentals of sentiment analysis and its applications. Springer, Cham, pp 1–24

25. Haider S, Tanvir Afzal M, Asif M, Maurer H, Ahmad A, Abuarqoub A (2021) Impact analysis of adverbs for sentiment classifcation on twitter product reviews. Concurr Comput Pract Exp. https://onlinelibrary.wiley.com/doi/10.1002/cpe.4956

26. 1 Liu B (2012) Sentiment analysis and opinion mining. Synth Lect Hum Lang Technol 5(1):1–167 Liu B, Zhang L (2012) A survey of opinion mining and sentiment analysis. Springer, Boston, pp 415– 463. https://doi.org/10.1007/978-1-4614-3223-4_13

27.1 Saad S, Saberi B (2017) Sentiment analysis or opinion mining: a review. Int J Adv Sci Eng Inf Technol –1660:(5)7 1666https://doi.org/10.18517/ijaseit.7.5.2137

28. D'Aniello, G., Gaeta, M., & La Rocca, I. (2022). KnowMIS-ABSA: An overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis. The Artificial Intelligence Review, 55(7), 5543-5574.

29. Jing, X. (2017). Concept-level sentiment analysis of online hotel. Available from ProQuest Dissertations & Theses Global Retrieved from https://www.proquest.com/dissertations-theses/concept-level-sentiment-analysis-online-hotel/docview/2008665967/se-2

30. Han, H., Zhang, Y., Zhang, J., Yang, J., & Zou, X. (2018). Improving the performance of lexicon-based review sentiment analysis method by reducing additional introduced sentiment bias. PLoS One, 13(8)

1.31 للاطلاع على القاموس الخاص بالكلمات الإيجابية والسلبية: -

32. https://drive.google.com/drive/folders/193s-AZ_BzLxxKF97xPtC-V9mG74VyiG4?usp=share_link

33. Osmani, A., & Jamshid, B. M. (2022). Weighted joint sentiment-topic model for sentiment analysis compared to ALGA: Adaptive lexicon learning using genetic algorithm. Computational Intelligence and Neuroscience, pp 50

34. Kolchyna, O., Souza, T. T. P., Treleaven, P., & Aste, T. (2015). Twitter sentiment analysis: Lexicon method, machine learning method and their combination. Ithaca: Cornell University Library, arXiv.org. Retrieved from https://www.proquest.com/working-papers/twitter-sentiment-analysis-lexicon-method-machine/docview/2076258448/se-2

35. Li, Q., Li, X., Du, Y., Fan, Y., & Chen, X. (2022). A new sentiment-enhanced word embedding method for sentiment analysis. *Applied Sciences*, *12*(20), 10236.

36. Ibrahim, R. M. (2018). Sentiment analysis of arabic tweets – implicit semantic-based approach . Available from ProQuest Dissertations & Theses Global. Retrieved from https://www.proquest.com/dissertations-theses/sentiment-analysis-arabic-tweets-implicit/docview/2597835208/se-2

37. Alanazi, N. (2021). Opinion mining challenges and case study: Using twitter for sentiment analysis towards Pfizer/BioNTech, moderna, AstraZeneca/Oxford, and sputnik COVID-19 vaccines (Order No. 28867246). Available from ProQuest Dissertations & Theses Global. (2617309929). Retrieved from https://www.proquest.com/dissertations-theses/opinion-mining-challenges-case-study-using/docview/2617309929/se-2

38. Zhilin, O. (2021). Challenges in opinion mining: An analysis using vaccine hesitancy (Order No. 29274154). Available from ProQuest Dissertations & Theses Global. (2700375258). Retrieved from https://www.proquest.com/dissertations-theses/challenges-opinion-mining-analysis-using-vaccine/docview/2700375258/se-2

39. Hossain, A., Karimuzzaman, M., Hossain, M. M., & Rahman, A. (2021). Text mining and sentiment analysis of newspaper headlines. Information, 12(10), 414. doi: https://doi.org/10.3390/info12100414

40. Cooper, D. (2020). Evolving Synthetic Data Generating Programs for Text Recognition Tasks: Evolving a Diverse Population of SyntheticData Generators Using Genetic Programming with Novelty Search for Robust Text Recognition Models Available https://www.proquest.com/dissertations-theses/b-evolving-

synthetic-data-generating-programs/docview/2674874946/se-2

41. Ahmed, S., Pasquier, M., & Qadah, G. (2013, March). Key issues in conducting sentiment analysis on Arabic social media text. In 2013 9th International Conference on Innovations in Information Technology (IIT) (pp. 72-77). IEEE.

42. Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.