CHAPTER 17

Various Uses Of Natural Language Processing

Yasin ER

Necmettin Erbakan University, Türkiye

To Cite This Chapter:
Er, Y. (2023). Various Uses Of Natural Language Processing. In H. Arikan, Y. Uzun & I. Cetin (Eds.), *Current Studies in Basic Sciences, Engineering and Technology 2023* (pp. 173–184). ISRES Publishing.

Introduction

The goal of the artificial intelligence (AI) branch of natural language processing (NLP) is to create models and algorithms that let computers comprehend, interpret, and produce human language. The first computer scientists started investigating the idea of utilizing machines to comprehend and interpret natural language in the 1950s, which is when natural language processing (NLP) first emerged (Katar et al., 2023).
In recent years, machine learning, a type of artificial intelligence, has emerged as a crucial NLP method. Large datasets are used to train algorithms so they can recognize patterns and correlations in the data and provide outputs or predictions.
NLP made tremendous strides in the 2000s thanks to the advent of neural networks, especially deep learning models. These models can perform a broad range of tasks, such as text categorization, sentiment analysis, and machine translation, since they can handle enormous volumes of text data and understand intricate linguistic patterns. In general, the history of natural language processing (NLP) and machine learning is marked by gradual advancements and innovations in computational and algorithmic methods. These days, a wide range of sectors and applications, from chatbots and virtual assistants to healthcare and finance, use NLP and machine learning.

Suicide ranks as the 13th most common cause of death worldwide, contributing to 5–6% of all fatalities (Lozano et al., 2013). Suicidal conduct is complex, and no one explanation applies to all situations. Information retrieval, text categorization, document summarization, text clustering, and topic modelling are some of the text mining techniques now in use. Conversely, these methods focus on extracting valuable information from text texts using various techniques as sentiment analysis, topic modelling, keyword extraction, and categorization (Munot & Govilkar, 2014). Unlike natural language processing (NLP), these methods have a narrower emphasis and don't usually centre on understanding textual meaning. NLP, on the other hand, is able to understand the mood and emotion that underlie texts, phrases, and sentences in addition to the meaning and context of individual words. This gives it an advantage over traditional text mining techniques in that it can comprehend complex texts more efficiently and extract more pertinent insights (Khurana et al., 2023).

Cancer is one of the major public health issues. The World Health Organisation (WHO) reported in 2019 that the primary cause of mortality globally is this illness. By 2030, it is estimated that there would be about 13 million cancer-related deaths worldwide (Allemani et al., 2018). Researchers employ a variety of techniques to battle cancer

since the illness is becoming more prevalent. One technique that has been utilised to diagnose cancer (Acs et al., 2020; Afolayan et al., 2022), as well as forecast its risk (Izci et al., 2020), likelihood of recurrence (Yang et al., 2020), and symptoms (Kourou et al., 2015; Yang et al., 2020), is artificial intelligence (AI). AI can offer a quick, easy, and safe method of treating many illnesses.

The Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) (Ely et al., 2001) and the Intensive Care Delirium Screening Checklist (ICDSC) (Inouye et al., 1990), two commonly used screening instruments for cognitive and behavioural dysfunction in critically ill patients, only provide a limited amount of guidance for the identification of behavioural phenotypes.

Despite the aforementioned challenges, a number of studies have found that there are delirium phenotypes and that they may be linked to varying rates of morbidity and death utilising CAMICU and ICDSC (Krewulak et al., 2018, 2020).

To detect terms and phrases like “combative” or “confused,” which may also be symptomatic of various phenotypes of NLP-diagnosed behaviour disturbance (NLP-Dx-BD), NLP may be used to examine the clinical notes of carers (Young, Holmes, Robbins, et al., 2023).

Many domains, such as big data (Jutla et al., 2013; Osinga et al., 2022), software effort estimation (Nassif et al., 2011), formal specification (Muhamad et al., 2019), augmented reality (Puspasari et al., 2020), e-commerce (Pujadi et al., 2020), health care (Vasilakis et al., 2008), safety-critical systems, and blockchain (Diaz et al., 2020), are among those where use cases are applicable. Its broader application serves as our source of inspiration. Use cases must be created using the free text criteria, which takes time, money, and human labour.

The majority of methods are semiautomated, require limited natural language, or both. Consequently, it is preferable to create use cases and recognise actors and their connections automatically, independent of formalism, constrained natural language, or extraneous information. Natural Language Processing (NLP) and Network Science are used to demonstrate a collection of algorithms that automatically construct use cases, actors, and their interactions from natural language needs.

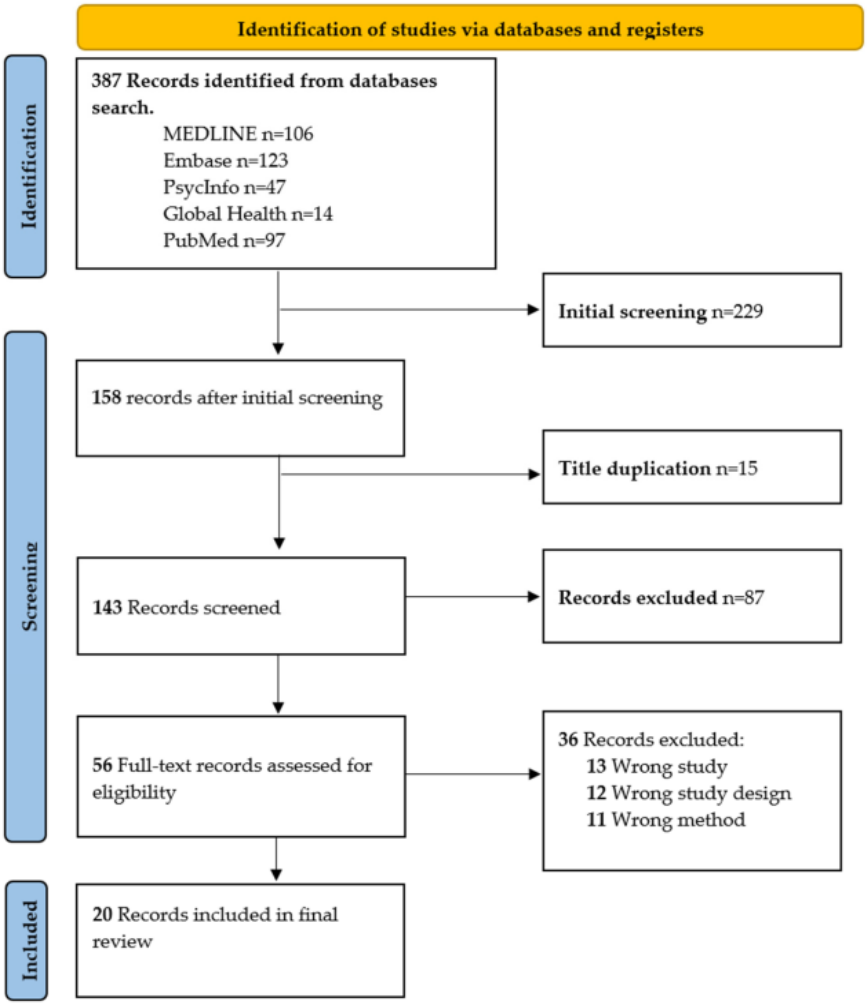
Materials and Methods

Preventing Suicide Ideation

387 hits were found overall in the preliminary search, which used the OVID platform to scan five different databases. Following the deduplication and preliminary filtering steps, 158 records in all were generated (Arowosegbe & Oyelade, 2023).

The 158 data were uploaded into Covidence, a systematic review management system, where the screening process was finished. After the entire text was reviewed, twenty (20) papers were evaluated and selected for publication. The (Preferred Reporting Items for Systematic Reviews and Meta-analyses) PRISMA (Page et al., 2021) flow diagram below is used to illustrate the search process, which is seen in Figure 1.

Figure 1.
PRISMA Diagram



The study tested the effectiveness of a suicide risk prediction model using Natural Language Processing (NLP/ML) in two South-eastern US emergency departments. The model (J. Pestian et al., 2016; J. P. Pestian et al., 2016), which was evaluated through interviews with 37 suicidal and 33 non-suicidal patients, demonstrated the feasibility of integrating technology into emergency department workflows. A prospective clinical trial also used NLP and semi-supervised machine learning techniques to differentiate between suicidal and non-suicidal individuals based on their conversations.

Neonatal fatalities have decreased globally by 51%, but not in low-income and middle-income countries (Hug et al., 2019; Mboya et al., 2020). Prompt delivery of high-quality healthcare services and early identification of pregnant women at risk for unfavourable outcomes during the prenatal period can enhance mother and neonatal survival (Kuhle et al., 2018). Machine learning and artificial intelligence models can be crucial in assessing risk factors for perinatal mortality and triaging pregnant women at high risk of severe postpartum depression, suicide ideation, and death. An NLP tool was developed to detect perinatal self-harm in electronic health records with sufficient recall and accuracy. (Zhong et al., 2018)created algorithms using NLP to detect pregnant women exhibiting suicidal behavior in electronic medical records. They used both structured and unstructured clinical notes to extract diagnostic information, demonstrating that using structured data and NLP to mine unstructured clinical notes significantly enhances the ability to identify suicidal behavior in pregnant women (Zhong et al., 2019). This approach led to an 11-fold increase in the number of pregnant women whose suicide behaviours were

identified. Using NLP significantly increases the sensitivity of screening for suicidal behaviour in electronic health records, but the proportion of verified suicidal behaviour was lower among women who did not have diagnostic codes for suicidal behaviour.

Mobile health applications (MHA) have the potential to improve access to evidence-based care for those with suicidal thoughts by addressing limitations in traditional mental health therapy. The proliferation of smartphones has enabled MHA to provide timely, convenient, discrete, and cheap assistance during severe crises. An automated algorithm for analysing and estimating the risk of suicide based on social media data was developed by (Coppersmith et al., 2018), which allows for scalable screening for suicide risk before engaging with a healthcare system.

The Boamente program, developed by (Diniz et al., 2022), uses textual data from users' smartphones to detect suicidal ideation. It uses an Android virtual keyboard to passively gather user messages and transfer them to a web service using Natural Language Processing (NLP) and Deep Learning.

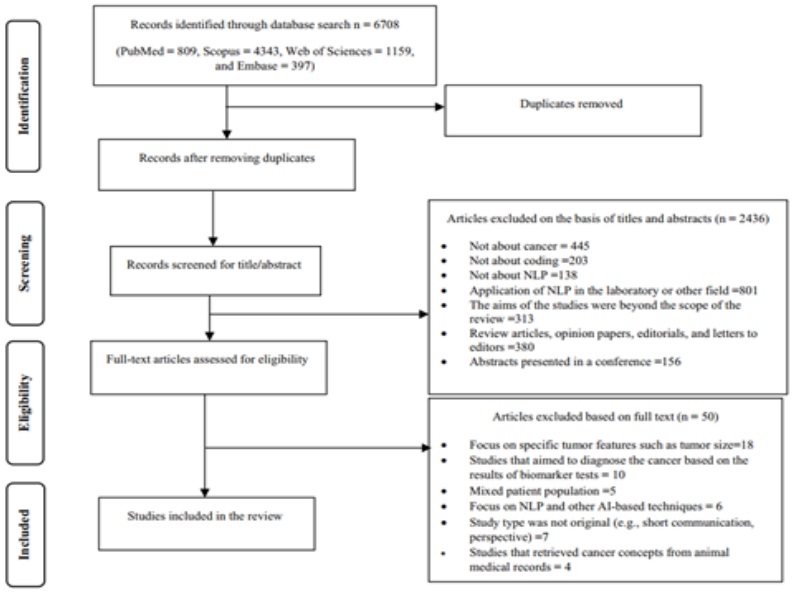
(Cook et al., 2016) used NLP and machine learning to predict suicide ideation and elevated mental symptoms among adults recently released from psychiatric inpatient or emergency hospital settings in Spain. They also created a domain Knowledge Aware Risk Assessment (KARA) model to enhance suicide identification in online counselling systems. The model can inform counsellors of potential risks in a user's information while engaging with the user.

Extracting Cancer Concepts

This work used natural language processing (NLP) to extract cancer concepts from papers in a systematic review manner. 6708 documents were initially retrieved in total. There were 2503 articles left for additional evaluation after duplicates were eliminated. The remaining papers' titles and abstracts were then filtered, and inclusion and exclusion criteria were used (Gholipour et al., 2023).

67 studies were deemed relevant after 2436 papers were eliminated due to the application of exclusion criteria. After going over the whole contents of these papers, 17 were chosen, and data was taken out of them (Figure 2).

Figure 2.
The PRISMA diagram of study selection



The study reveals that researchers have utilized various methods and algorithms to retrieve concepts from electronic texts in the field of cancer. The rule-based algorithm was the most commonly used, but deep learning has been more commonly used in healthcare (Casey et al., 2021; Shickel et al., 2017). Both rule-based and machine learning algorithms can demonstrate comparable performance when evaluated using the same dataset (Chapman et al., 2001; Solti et al., 2009). Despite the widespread adoption of deep learning methods, both rule-based and traditional algorithms remain popular due to their simplicity, ease of implementation, and interpretability (Rajula et al., 2020). Deep learning can be challenging to interpret due to unclear steps in the analytical output. Rule-based and traditional algorithms are more suitable for smaller datasets with few features, as they do not require massive amounts of data (Rajula et al., 2020; Vasudevan et al., 2021). However, ML techniques can lead to overfitting, which limits the generalizability of the model to different data sets and makes inaccurate predictions. Despite this, traditional algorithms may not always be a better approach than deep learning, as some situations may require more flexible and complex techniques (Vasudevan et al., 2021).

The application of Natural Language Processing (NLP) methods for cancer idea extraction is assessed in this systematic review, which is a novel area in the biological, clinical, and medical domains. According to the report, the most often used NLP terminology for extracting cancer concepts are SNOMED-CT and UMLS. In this discipline, rule-based procedures were the most employed approaches. The study advises academics to concentrate on combining the findings of these investigations and using ML and deep learning algorithms for information extraction. The outcomes of NLP software used to extract illness ideas from clinical documents such as radiology or laboratory reports should be compared in future studies. In order to assist developers in selecting the appropriate terminological system with appropriate coverage, future research should also focus on the content coverage of applicable terminological systems.

The study analysed the evaluation methods used in evaluating NLP algorithms, focusing on the recall (R), f1-score (F1), and precision (P) metrics. The research included articles using pre-developed software or software developed by researchers to interpret text and extract cancer concepts. Similar to previous studies, this review extracted concepts identified by included studies, the NLP methodology and tools used, and their application purpose and performance results.

The most commonly used terminologies in the articles were UMLS and SNOMEDCT, with UMLS being more frequently used. A 2020 study found that 42% of UMLS users were researchers, while 28% were programmers and software developers. The study divided articles concerning UMLS into six categories, with 78% falling under the NLP category.

The use of SNOMED-CT terminology in implementations has increased in recent years, while its use in theoretical discussions has recently been reduced. The results of this study indicated the practical use of this terminology to retrieve concepts from medical texts or documents.

A 2013 review paper categorized articles into five groups: unknown, theoretical, development and design, implementation, and evaluation. However, a 2013 study by Change et al. addressed 124 articles out of 622, highlighting the importance of this field and its attention in recent years. Most articles focused on the classification or coding of free-text clinical notes/narratives and radiology reports.

Despite the importance of content coverage as a metric in the evaluation of terminological systems, most articles in the review did not include this information in their results. This may be due to the focus on extracting concepts from narratives and identifying the best algorithms rather than evaluating applied terminological systems.

Behavioural Disturbance Phenotypes

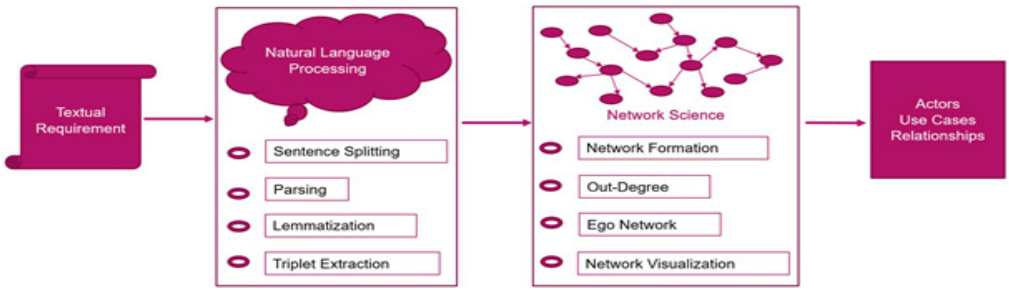
A retrospective, non-interventional research was conducted on a cohort of critically sick adult patients (≥18 years old) who were admitted to the three intensive care units (ICUs) of a hospital connected with a university in Melbourne, Australia, between May 1, 2019, and December 31, 2020 (Young, Holmes, Kishore, et al., 2023). The Austin Hospital Human Research Ethics Committee (LNR/19/Austin/38) accepted the project and disregarded the need for informed consent. Only the initial admission was included for patients who had multiple admissions. There were no further exclusion standards used. All patients received treatment during the trial period that was intended to lower the likelihood of developing delirium. This care included visits from family members, subdued lighting at night, less contact to promote sleep at night, and the use of visual and auditory aids wh0haracterization of NLP-Dx-BD (Young, Holmes, Robbins, et al., 2023) characteristics in patients are made possible using natural language processing in electronic carer notes. Furthermore, we discovered that the combination phenotype predominated and that critically sick research participants often switched between NLP-Dx-BD (Young, Holmes, Robbins, et al., 2023) behavioural phenotypes throughout the first 48 hours. Crucially, upon adjustment, antipsychotic medication was substantially more likely to be prescribed to patients who fit the agitated phenotype, either alone or in combination with other phenotypes. These findings have significant ramifications for trial design in this area as well as for our comprehension of the epidemiology of phenotypes of disordered behaviour in critical patients.

Extraction Of Use Case Diagram Elements

From natural language requirements, the method creates use cases, actors, and connections. We employ natural language processing (NLP) and network science methods. Sentence splitting, parsing, lemmatization, and triplet extraction are examples of NLP approaches. The method goes through many stages to identify the players from the textual requirement. Every stage enhances the performers' identification. The upper-level diagram of our suggested method is displayed in Figure 3.

Figure 3.

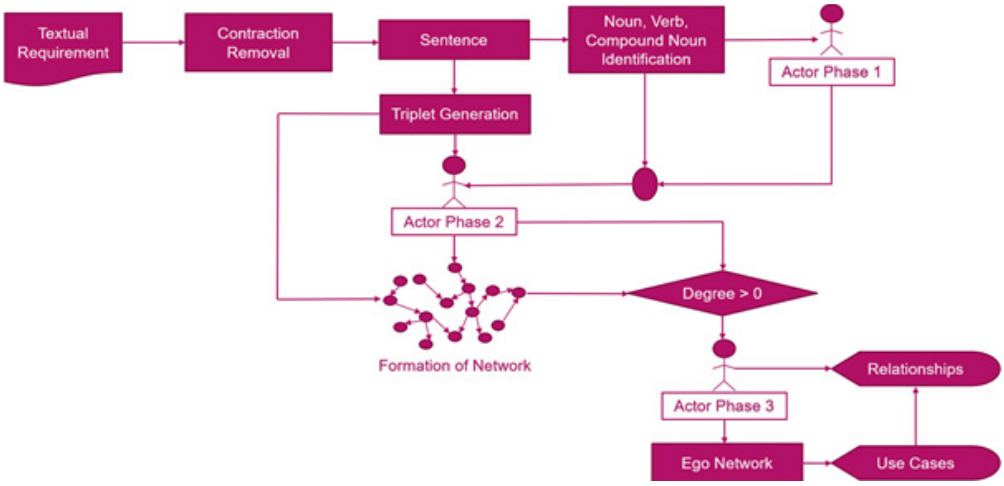
Use Case Diagram Approach



The method uses textual requirements as input to determine actors, use cases, and their connections. In Algorithm 1 (Imtiaz Malik et al., 2023), every textual requirement sentence is divided into a separate sentence. In Algorithm 2 (Imtiaz Malik et al., 2023), we have determined the noun, verb, and compound noun from every phrase. In the following algorithms, these nouns, verbs, and compound nouns make it easier to identify actors and use cases. Algorithm 3 (Imtiaz Malik et al., 2023) uses the noun list of each phrase as an input to determine who the actors in phase 1 are. The triplets in each sentence are recognised. Algorithm 4 (Imtiaz Malik et al., 2023) uses these triplets as input to determine who the actors in step 2 are. The graph's nodes and linkages may also be found using this approach. For the purpose of creating the network, identifying use cases, and identifying the players in the last stage, these actors and linkages are supplied in Algorithm 1. Figure 4 depicts the workflow of our suggested method.

Figure 4.

Process Flow Of Proposed Approach



For every case study, the outcomes are analysed in two categories. Using the manual actors and use cases found in (Al-Hroob et al., 2018), the outcomes are assessed in the first category.

The manual results in the second category are refined and found in (Al-Hroob et al., 2018). The majority of the manually selected use cases from the case studies include only one word, a verb, according to observations (Al-Hroob et al., 2018). The noun and verb need both be present in a use case, though. For an Assembly System Case Study (Imtiaz Malik et al., 2023), a few more actors and use cases have been found.

The findings show that either the triplets are missing, or the triplet does not have an actor in the topic, which accounts for the missing use instances. Our method distinguishes the actor from the subject in step 2 regardless of whether the phrase is written in the active or passive voice. For passive voice phrases, object-predicate-subject has been changed to subject-predicate-object.

An automated approach is proposed for generating actors, use cases, and their relationships without relying on formalism or restricted natural language. This approach uses Natural Language Processing (NLP) techniques and network science to generate both primary and external actors. The approach generates actors in three phases: natural language processing, network science, and use cases identification. The node values are set using NLP techniques. The approach comprises five algorithms: first, which discusses the extraction of actors and links through triplets of sentences; second, which identifies nouns, verbs, and compound nouns; third, which generates actors; fourth, which refines actors; and last, which generates the final actor, use cases, and relationships. The proposed approach outperforms the existing approach in five case studies, achieving 74.5% precision, 85% recall, 74% F-Measure, 25.45% False Discovery Rate, and 14.97% Miss Rate for the first category and 70% precision, 81% recall, 71.5% F-measure, 29.73% False Discovery Rate, and 19% Miss Rate for the second category (Imtiaz Malik et al., 2023).

References

Acs, B., Rantalainen, M., & Hartman, J. (2020). Artificial intelligence as the next step towards precision pathology. In *Journal of internal medicine* (Vol. 288, Issue 1, pp. 62–81). Blackwell Scientific Publications. https://doi.org/10.1111/joim.13030

Afolayan, J. O., Adebiyi, M. O., Arowolo, M. O., Chakraborty, C., & Adebiyi, A. A.

(2022). Breast Cancer Detection Using Particle Swarm Optimization and Decision Tree Machine Learning Technique. In *Intelligent Healthcare ISBN: 9789811681493*. Springer Nature Singapore. https://research.ebsco.com/linkprocessor/plink?id=6c82d6e6-4c41-3556-9415-13d2ff490849

Al-Hroob, A., Imam, A. T., & Al-Heisa, R. (2018). The use of artificial neural networks for extracting actions and actors from requirements document. *Information and Software Technology*, *101*, 1–15. https://doi.org/10.1016/j.infsof.2018.04.010

Allemani, C., Di Carlo, V., Harewood, R., Matz, M., Nikšić, M., Bonaventure, A., Coleman, M. P., Matsuda, T., Valkov, M., Johnson, C. J., Estève, J., Ogunbiyi, O. J., Azevedo e Silva, G., Chen, W.-Q., Eser, S., Engholm, G., Stiller, C. A., Monnereau, A., Woods, R. R., … Lewis, C. (2018). Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *The Lancet*, *391*(10125), 1023–1075. https://doi.org/10.1016/S0140-6736(17)33326-3

Arowosegbe, A., & Oyelade, T. (2023). Application of Natural Language Processing (NLP) in Detecting and Preventing Suicide Ideation: A Systematic Review. In *International Journal of Environmental Research and Public Health* (Vol. 20, Issue 2). MDPI. https://doi.org/10.3390/ijerph20021514

Casey, A., Davidson, E., Poon, M., Dong, H., Duma, D., Grivas, A., Grover, C., Suárez-Paniagua, V., Tobin, R., Whiteley, W., Wu, H., & Alex, B. (2021). A Systematic Review of Natural Language Processing Applied to Radiology Reports. In *BMC Medical Informatics and Decision Making 2021*. https://doi.org/10.1186/s12911-021-01533-7

Chapman, W. W., Chapman, B. E., Fizman, M., & Haug, P. J. (2001). A comparison of classification algorithms to automatically identify chest X-ray reports that support pneumonia. *Journal of Biomedical Informatics*, *34*(1), 4–14. https://doi.org/10.1006/jbin.2001.1000

Cook, B. L., Progovac, A. M., Mullin, B., Hou, S., Chen, P., & Baca-Garcia, E. (2016). Novel Use of Natural Language Processing (NLP) to Predict Suicidal Ideation and Psychiatric Symptoms in a Text-Based Mental Health Intervention in Madrid. *Computational and Mathematical Methods in Medicine*, *2016*. https://doi.org/10.1155/2016/8708434

Coppersmith, G., Fine, A. B., Crutchley, P., & Leary, R. (2018). Natural Language Processing of Social Media as Screening for Suicide Risk. In *Biomedical Informatics Insights* (Vol. 10). SAGE Publishing. https://research.ebsco.com/linkprocessor/plink?id=7063684e-7fb0-321f-bdcf-304c3b7cd1f6

Diaz, M., Soler, E., Llopis, L., & Trillo, J. (2020). Integrating Blockchain in Safety-Critical Systems: An Application to the Nuclear Industry. In *IEEE Access, Access, IEEE* (Vol. 8, pp. 190605–190619). IEEE. https://doi.org/10.1109/ACCESS.2020.3032322

Diniz, E. J. S., Teles, A. S., Fontenele, J. E., de Oliveira, A. C., Bastos, V. H., Teixeira, S., Dos Santos, R. M., Rabêlo, R. L., de Oliveira, A. K., & Calçada, D. B. (2022). Boamente: A Natural Language Processing-Based Digital Phenotyping Tool for Smart Monitoring of Suicidal Ideation. *Healthcare (Switzerland)*, *10*(4). https://doi.org/10.3390/healthcare10040698

Ely, E. W., Bernard, G. R., Speroff, T., Gautam, S., Dittus, R., May, L., Truman, B., Ely, E. W., Gordon, S., Margolin, R., Inouye, S. K., Francis, J., & Hart, R. P. (2001). Delirium in mechanically ventilated patients: Validity and reliability

of the Confusion Assessment Method for the intensive care unit (CAM-ICU). *Journal of the American Medical Association*, *286*(21), 2703–2710. https://doi.org/10.1001/jama.286.21.2703

Gholipour, M., Khajouei, R., Amiri, P., Hajesmaeel Gohari, S., & Ahmadian, L. (2023). Extracting cancer concepts from clinical notes using natural language processing: a systematic review. *BMC Bioinformatics*, *24*(1). https://doi.org/10.1186/s12859-023-05480-0

Hug, L., Alexander, M., You, D., & Alkema, L. (2019). National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. *The Lancet Global Health*, *7*(6), e710–e720. https://doi.org/10.1016/S2214-109X(19)30163-9

Imtiaz Malik, M., Azam Sindhu, M., & Ayaz Abbasi, R. (2023). Extraction of use case diagram elements using natural language processing and network science. *PloS One*, *18*(6), e0287502. https://doi.org/10.1371/journal.pone.0287502

Inouye, S. K., H, van D. C., Alessi, C. A., Balkin, S., Siegal, A. P., & Horwitz, R. I. (1990). Clarifying confusion: the confusion assessment method. A new method for detection of delirium. In *Annals of internal medicine* (Vol. 113, Issue 12, pp. 941–948). American College of Physicians--American Society of Internal Medicine. https://doi.org/10.7326/0003-4819-113-12-941

Izci, H., Tambuyzer, T., Tuand, K., Depoorter, V., Laenen, A., Wildiers, H., Vergote, I., L, V. E., H, D. S., Verdoodt, F., & Neven, P. (2020). A Systematic Review of Estimating Breast Cancer Recurrence at the Population Level With Administrative Data. In *Journal of the National Cancer Institute* (Vol. 112, Issue 10, pp. 979–988). Oxford University Press. https://doi.org/10.1093/jnci/djaa050

Jutla, D. N., Bodorik, P., & Ali, S. (2013). Engineering Privacy for Big Data Apps with the Unified Modeling Language. In *2013 IEEE International Congress on Big Data, Big Data (BigData Congress), 2013 IEEE International Congress on, bigdatacongress* (pp. 38–45). IEEE. https://doi.org/10.1109/BigData.Congress.2013.15

Katar, O., ÖZKAN, D., YILDIRIM, Ö., & Acharya, U. R. (2023). Evaluation of GPT-3 AI Language Model in Research Paper Writing. *Turkish Journal of Science & Technology*, *18*(2), 311–318. https://doi.org/10.55525/tjst.1272369

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications: An International Journal*, *82*(3), 3713–3744. https://doi.org/10.1007/s11042-022-13428-4

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V, & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. In *Computational and Structural Biotechnology Journal* (Vol. 13, Issue C, pp. 8–17). Elsevier. https://doi.org/10.1016/j.csbj.2014.11.005

Krewulak, K. D., Stelfox, H. T., Ely, E. W., & Fiest, K. M. (2020). Risk factors and outcomes among delirium subtypes in adult ICUs: A systematic review. *Journal of Critical Care*, *56*, 257–264. https://doi.org/10.1016/j.jcrc.2020.01.017

Krewulak, K. D., Stelfox, H. T., Leigh, J. P., Ely, E. W., & Fiest, K. M. (2018). Incidence and Prevalence of Delirium Subtypes in an Adult ICU: A Systematic Review and Meta-Analysis. In *Critical Care Medicine*. https://doi.org/10.1097/CCM.0000000000003402

Kuhle, S., Maguire, B., Zhang, H., Hamilton, D., Allen, A. C., Joseph, K. S., & Allen,

V. M. (2018). Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: a retrospective cohort study. In *BMC Pregnancy and Childbirth* (Vol. 18, Issue 1, pp. 1–9). BMC. https://doi.org/10.1186/s12884-018-1971-2

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., Abraham, J., Adair, T., Aggarwal, R., Ahn, S. Y., Alvarado, M., Anderson, H. R., Anderson, L. M., Andrews, K. G., Atkinson, C., Baddour, L. M., Barker-Collo, S., Bartels, D. H., Bell, M. L., … Lipshultz, S. E. (2013). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, *380*(9859), 2095–2128.

Mboya, I. B., Mohammed, M., Mwambi, H. G., Mahande, M. J., & Obure, J. (2020). Prediction of perinatal death using machine learning models: A birth registry-based cohort study in northern Tanzania. *BMJ Open*, *10*(10). https://doi.org/10.1136/bmjopen-2020-040132

Muhamad, Z. H., Abdulmonim, D. A., & Alathari, B. (2019). An integration of uml use case diagram and activity diagram with Z language for formalization of library management system. *International Journal of Electrical and Computer Engineering*, *9*(4), 3069–3076. https://doi.org/10.11591/ijece.v9i4.pp3069-3076

Munot, N., & Govilkar, S. (2014). Comparative Study of Text Summarization Methods. In *International Journal of Computer Applications* (Vol. 102, pp. 33–37). Foundation of Computer Science. https://research.ebsco.com/linkprocessor/plink?id=dc33e1ab-a8ba-3606-a4c6-71788655b1ed

Nassif, A., Capretz, L., & Ho, D. (2011). *A Regression Model with Mamdani Fuzzy Inference System for Early Software Effort Estimation Based on Use Case Diagrams*.

Osinga, S. A., Paudel, D., Mouzakitis, S. A., & Athanasiadis, I. N. (2022). Big data in agriculture: Between opportunity and solution. *Agricultural Systems*, *195*, 103298. https://doi.org/https://doi.org/10.1016/j.agsy.2021.103298

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology*, *134*, 178–189. https://doi.org/10.1016/j.jclinepi.2021.03.001

Pestian, J. P., Grupp-Phelan, J., Bretonnel Cohen, K., Meyers, G., Richey, L. A., Sorter, M. T., & Matykiewicz, P. (2016). A Controlled Trial Using Natural Language Processing to Examine the Language of Suicidal Adolescents in the Emergency Department. *Suicide and Life-Threatening Behavior*, *46*(2), 154–159. https://doi.org/10.1111/sltb.12180

Pestian, J., Sorter, M. T., Morency, L.-P., Rohlfs, L., McCullumsmith, C. B., Scherer, S., Gee, J. T., Connolly, B., & Cohen, K. B. (2016). A Machine Learning Approach to Identifying the Thought Markers of Suicidal Subjects: A Prospective Multicenter Trial. In *Suicide and Life-Threatening Behavior* (Vol. 47, pp. 112–121). Wiley. https://research.ebsco.com/linkprocessor/plink?id=2e7c0a65-8602-3c4d-b01c-bc67af4a7561

Pujadi, T., Simamora, B. H., Kumar, V., Hulu, Y., Tumar, & Wihendro. (2020). Modeling of E-Commerce Supply Chains Mobile Application. In *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), Cybernetics and Intelligent System (ICORIS), 2020 2nd International Conference on* (pp. 1–5).

IEEE. https://doi.org/10.1109/ICORIS50180.2020.9320757

Puspasari, S., Suhandi, N., & Iman, J. N. (2020). Augmented Reality Development for Supporting Cultural Education Role in SMB II Museum During Covid-19 Pandemic. In *2020 Fifth International Conference on Informatics and Computing (ICIC), Informatics and Computing (ICIC), 2020 Fifth International Conference on* (pp. 1–6). IEEE. https://doi.org/10.1109/ICIC50835.2020.9288619

Rajula, H. S. R., Fanos, V., Verlato, G., Antonucci, N., & Manchia, M. (2020). Comparison of conventional statistical methods with machine learning in medicine: Diagnosis, drug development, and treatment. *Medicina (Lithuania)*, *56*(9), 1–10. https://doi.org/10.3390/medicina56090455

Shickel, B., Tighe, P., Bihorac, A., & Rashidi, P. (2017). *Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis*. https://doi.org/10.1109/JBHI.2017.2767063

Solti, I., Cooke, C. R., Xia, F., & Wurfel, M. M. (2009). Automated classification of radiology reports for acute lung injury: Comparison of keyword and machine learning based natural language processing approaches. In *2009 IEEE International Conference on Bioinformatics and Biomedicine Workshop, Bioinformatics and Biomedicine Workshop, 2009. BIBMW 2009. IEEE International Conference on* (pp. 314–319). IEEE. https://doi.org/10.1109/BIBMW.2009.5332081

Vasilakis, C., Lecnzarowicz, D., & Lee, C. (2008). Application of Unified Modelling Language (UML) to the Modelling of Health Care Systems: An Introduction and Literature Survey. *IJHISI*, *3*, 39–52.

Vasudevan, A., Langenberg, D. R. van, & Con, D. (2021). Deep learning vs conventional learning algorithms for clinical prediction in Crohn's disease: A proof-of-concept study. In *World Journal of Gastroenterology* (Vol. 27, pp. 6476–6488). Baishideng Publishing Group Inc. https://research.ebsco.com/linkprocessor/plink?id=1f4c6807-61d3-3c73-9031-bdc67cc333d1

Yang, L., Fu, B., Li, Y., Liu, Y., Huang, W., Feng, S., Xiao, L., Sun, L., Deng, L., Zheng, X., Ye, F., & Bu, H. (2020). Prediction model of the response to neoadjuvant chemotherapy in breast cancers by a Naive Bayes algorithm. *Computer Methods and Programs in Biomedicine*, *192*. https://doi.org/10.1016/j.cmpb.2020.105458

Young, M., Holmes, N. E., Kishore, K., Amjad, S., Gaca, M., Serpa Neto, A., Reade, M. C., & Bellomo, R. (2023). Natural language processing diagnosed behavioural disturbance phenotypes in the intensive care unit: characteristics, prevalence, trajectory, treatment, and outcomes. *Critical Care*, *27*(1). https://doi.org/10.1186/s13054-023-04695-0

Young, M., Holmes, N., Robbins, R., Marhoon, N., Amjad, S., Neto, A. S., & Bellomo, R. (2023). Natural language processing to assess the epidemiology of delirium-suggestive behavioural disturbances in critically ill patients. In *Critical care and resuscitation : journal of the Australasian Academy of Critical Care Medicine* (Vol. 23, Issue 2, pp. 144–153). Elsevier B.V. on behalf of College of Intensive Care Medicine of Australia and New Zealand. https://doi.org/10.51893/2021.2.oa1

Zhong, Q.-Y., Karlson, E. W., Gelaye, B., Finan, S., Avillach, P., Smoller, J. W., Cai, T., & Williams, M. A. (2018). Screening pregnant women for suicidal behavior in electronic medical records: diagnostic codes vs. clinical notes processed by natural language processing. In *BMC Medical Informatics and Decision Making* (Vol. 18, Issue 1, pp. 1–11). BMC. https://doi.org/10.1186/s12911-018-0617-7

Zhong, Q.-Y., Mittal, L. P., Nathan, M. D., Brown, K. M., González, D. K., Cai, T.,

Finan, S., Gelaye, B., Avillach, P., Smoller, J. W., Karlson, E. W., Cai, T., & Williams, M. A. (2019). Use of natural language processing in electronic medical records to identify pregnant women with suicidal behavior : towards a solution to the complex classification problem. *European Journal of Epidemiology*, *34*(2), 153–162. https://research.ebsco.com/linkprocessor/plink?id=d66a6065-0eb8-3346-9a67-02f1d59ebb5e

About The Authors

Yasin ER, he works as a computer engineer at the Turkish State Meteorological Service. He has a master's degree from Hoca Ahmet Yesevi University. His studies are meteorological communication, enterprise-specific software development, and Windows/Linux-based system management. He is a member of various expert groups and working groups at the World Meteorological Organization, currently, He is studying for a doctorate at the Institute of Pure and Applied Science of Necmettin Erbakan University.

E-mail: yasineryasiner@gmail.com, ORCID: 0000-0001-6729-5278

Similarity Index:
The similarity index obtained from the plagiarism software for this book chapter is 12%.