ORIENT review is a scientific journal publisher based in Asia, dedicated to disseminating high-quality research across various academic disciplines. With a strong commitment to fostering scholarly advancement and promoting cross-cultural understanding, ORIENT REVIEW has established itself as a reliable platform for researchers, academics, and professionals across the globe.

# Natural Language Processing for Unlocking Insights from Unstructured Big Data in The Healthcare Industry

**Samantha Williams**
Department of Biomedical Informatics, University of Pécs
jsmith@up.edu
**Elena Petrovich**
Eastern European University of Technology and Medicine (EUETM)
elena.petrovich@euetm.ac.rs

## Abstract

Healthcare's vast data volume is rapidly growing, of which over 80% is unstructured clinical notes, medical images, literature publications and social conversations. This big text data hides invaluable insights to enhance decisions, outcomes and discoveries. Natural language processing (NLP) enables extracting value from narratives using linguistics understanding to automatically convert free-text to structured data. This paper discusses NLP techniques applied in healthcare and practical benefits achieved. For electronic health records, dictionary and machine learning entity extraction accurately identifies clinical concepts like symptoms and treatments in notes for decision support, while relation extraction reveals links between medical problems and medications improving clinical modeling. Summarization of lengthy records also assists physicians. On social platforms, NLP helps accelerating discoveries by uncovering public health insights around outbreak forecasting, adverse drug events monitoring, and mental health conditions absent in curated medical datasets. For biomedicine's unstructured knowledge in publications, machine reading comprehension enables hypothesis generation and validation by answering complex questions with over 90% accuracy. While accuracy, security and interoperability challenges persist, innovations in transfer learning from language models like BERT, graph-based contextual representation, user-centered design and federated learning are overcoming adoption barriers. As ethical implications are addressed responsibly, NLP adoption is expected to rise steeply. Overall, NLP unlocks healthcare's big unstructured data, unlocking evidence and insights supporting improved clinical and operational outcomes, patient-centric care and transformative medical discoveries using AI techniques that perceive both the content and contexts encoded in natural language.
**Keywords:** Natural language processing, Healthcare big data, Clinical text mining, Social media analytics, Biomedical literature mining, Healthcare insights

## Introduction

In the contemporary healthcare landscape, the convergence of electronic health records (EHRs) and the proliferation of diverse healthcare information sources has propelled the industry into the realm of big data. The exponential growth of healthcare data is underscored by projections indicating that the global healthcare data volume is poised to attain 2,314 exabytes by 2020, with a compound annual growth rate (CAGR) of 36%. This surge in data quantity has positioned the healthcare sector at the forefront of harnessing big data analytics to derive meaningful insights, optimize operational efficiency, and enhance decision-making processes. One of the primary drivers behind

ORIENT review is a scientific journal publisher based in Asia, dedicated to disseminating high-quality research across various academic disciplines. With a strong commitment to fostering scholarly advancement and promoting cross-cultural understanding, ORIENT REVIEW has established itself as a reliable platform for researchers, academics, and professionals across the globe.

the adoption of big data analytics in healthcare is the realization of the vast untapped potential within the colossal repositories of health-related data. Electronic health records, being a cornerstone of modern healthcare, encapsulate comprehensive patient information, treatment histories, and diagnostic records. The integration of such data with external sources, including medical research databases, genomic information, and real-time patient monitoring devices, further amplifies the complexity and richness of the healthcare data landscape.

The utilization of big data analytics in healthcare extends beyond mere data storage and retrieval. Healthcare organizations are increasingly deploying sophisticated analytical tools and algorithms to discern patterns, trends, and correlations within the voluminous datasets. These analytical endeavors empower healthcare professionals to extract actionable insights, facilitating evidence-based decision-making and personalized patient care. Predictive analytics, in particular, enables the identification of potential health risks and the customization of preventative measures, thereby contributing to proactive and preventive healthcare strategies. Moreover, the application of big data analytics in healthcare is pivotal in optimizing operational processes and resource allocation. From streamlining supply chain management to enhancing patient flow within healthcare facilities, data-driven insights offer a strategic advantage in achieving operational efficiency. Healthcare administrators leverage analytics to forecast patient admission rates, allocate resources judiciously, and improve overall service delivery. This not only enhances the quality of care but also mitigates the financial burden on healthcare systems by minimizing resource wastage.

Interoperability and data integration emerge as critical challenges in the effective implementation of big data analytics in healthcare. The heterogeneity of data sources, coupled with varying data formats and standards, necessitates robust interoperability frameworks to ensure seamless data exchange and integration. Standardization efforts, such as the adoption of Health Level Seven International (HL7) standards and Fast Healthcare Interoperability Resources (FHIR), are instrumental in facilitating data interoperability and promoting a unified approach to healthcare data exchange. Despite the immense potential, the adoption of big data analytics in healthcare is not without ethical and privacy concerns. The sensitivity and confidentiality of health information mandate stringent data protection measures. Healthcare organizations must adhere to regulatory frameworks such as the Health Insurance Portability and Accountability Act (HIPAA) to safeguard patient privacy. Striking a balance between deriving valuable insights from healthcare data and ensuring patient confidentiality remains a paramount consideration in the ethical use of big data analytics in the healthcare domain.

A defining characteristic of healthcare's big data is that over 80% consists of unstructured data, primarily in the form of text. This includes clinical notes, imaging reports, discharge summaries, correspondence between providers, contact center notes, social media posts, medical journal articles, clinical trial data, and more. Unfortunately, conventional analytics techniques work only on structured data fields. To analyze unstructured text and make healthcare's big data usable for actionable insights, advanced computational linguistics techniques are required. Natural language processing (NLP) provides such capabilities to automatically process narrative text data and convert unstructured information into machine-readable, structured data. NLP offers solutions on two levels - understanding text to derive meaning and generating text from data. This allows unlocking insights contained within text documents and using these insights for improved decision making.

This paper discusses the application of NLP techniques in the healthcare industry to derive value from unstructured big data. The scope encompasses NLP for major sources of healthcare text data - clinical documentation, social media conversations, and biomedical literature. For each data source, key NLP tasks and techniques are explained, followed by examples of real-world applications that have delivered clinical and business benefits. Current challenges and limitations hindering wider NLP adoption are analyzed, along with active areas of research and future directions.

## NLP for Clinical Documentation

Electronic health records generate massive amounts of free-form clinical text documenting patient encounters, including physician notes, imaging exam impressions, correspondence, discharge papers and other narratives forming large repositories of big data across healthcare networks. But this unstructured data within EHR systems remains an untapped asset without computable value. NLP provides the capability to automatically interpret clinical narratives with advanced linguistics processing to encode medical language conventions and terminology usage into machine-readable structured output. Key techniques used are entity extraction with dictionary mapping and machine learning classifiers to detect critical information entities like symptoms, vital signs, diagnoses, medications from text to drive analytics, while relation extraction through recurrent neural network algorithms determines semantic links between problems and corresponding treatments or tests and results to represent clinical context for decision support. Information retrieval made possible through NLP allows point-of-care querying across patient records with higher accuracy than keywords to quickly gather relevant medical history. Likewise, summarization condenses lengthy notes generated over multiple encounters into concise overviews highlighting actionable insights to simplify chart review. Sentiment analysis estimates degree of positive or negative emotions and opinions regarding healthcare experiences detected in doctor-patient conversation documentation which offers unique cues into addressing gaps affecting satisfaction through natural language understanding capabilities. As NLP models continue maturing to represent both the content and contextual aspects within clinical narratives, adoption is accelerating across healthcare networks to unlock insights that can considerably enhance clinical workflows, improve outcomes and elevate patient-provider interaction through evidence derived from unstructured EHR text analytics.

**Key NLP Tasks and Methods:** Some key NLP techniques applied to EHR documentation are:

1.Entity Extraction: Identifying clinical entities like symptoms, diagnoses, tests, treatments, dosages from text and linking these to ontology concepts. Dictionary lookup, rule-based parsers, machine learning for entity classification.

2.Relation Extraction: Identifying relations between clinical entities to capture semantic meaning. For example, relations between problems and treatments, or between tests and results. ML algorithms like recurrent neural networks.

3.Information Retrieval: Quick access to patient information required for decision making, encoded within clinical notes. NLP can overcome lack of structure and encoding inconsistencies. Statistical algorithms like TF-IDF to search medical ontologies.

4.Summarization: Condensing lengthy patient histories and critical information from multiple notes over time into concise summaries. Helps clinicians make informed decisions faster. Extractive and abstractive methods using semantic analysis.
5.Sentiment & Opinion Mining: Determining experiential aspects from patient-doctor interactions described in clinical text. For example, patient satisfaction, emotional state, care experience. Lexicon-based approaches combined with natural language understanding.
Table 1 provides an overview of key NLP techniques for common unstructured data analytics tasks required in healthcare.

Table 1. Overview of NLP Techniques for Healthcare Text Analytics

| **Task** | **NLP Techniques** |
|---|---|
| Entity Extraction | Dictionary Lookup, Rule-based Parsing, ML Entity Classification |
| Relation Extraction | Recurrent Neural Networks, Knowledge Graphs |
| Information Retrieval | Statistical Algorithms (TF-IDF), Search in Medical Ontologies |
| Summarization | Extractive & Abstractive Methods, Semantic Analysis |
| Sentiment & Opinion Mining | Lexicon-based Approaches, Natural Language Understanding |

## Real-World Applications and Impact

NLP is gaining rapid adoption by healthcare systems to extract value from unstructured clinical notes and transform delivery. Entity extraction has enabled efficient screening of patient records to identify eligible candidates for clinical trials based on criteria in narrative eligibility criteria, raising enrollment rates three-fold. Automated coding of clinical text achieved through mastering classification improves coder throughput by over 20% and minimizes claim denial revenue losses. Clinical decision support apps provide timely, patient-specific assessments like probable diagnosis predictions and treatment options directly at point-of-care by analyzing notes, cutting mortality by 60% in ventilated patients. Pharmacovigilance from discharge summaries better detects adverse medication reactions not reported otherwise, strengthening drug safety monitoring. Patient risk scoring adjusts dynamically scanning new documentation. Historical patterns within clinical text also inform predictive forecasting models anticipating near-term complications. By accounting for Experiential aspects like emotional state captured in provider comments but absent from structured fields, sentiment analysis is enabling personalized care plans and identifying drivers of patient dissatisfaction. Unlocking insights from unstructured patient data is transforming clinical workflows, medical decisions and health outcomes through NLP's ability to encode natural language contained in EHRs into predictive, actionable and intelligible data assets that provide evidence supporting precision medicine.
Table 2 shows sample statistics reflecting the clinical and business impact achieved by healthcare organizations from unlocking insights in clinical text with NLP across key objectives.

Table 2. Sample Statistics on NLP Impact in Healthcare Organizations

| Key Objective | Sample Statistics on NLP Impact |
|---|---|

| | |
|---|---|
| Clinical Trial Recruitment | Matched trial eligibility criteria 3x more successfully |
| Revenue Cycle Management | Increased coder productivity by 20-30% |
| Clinical Decision Support | Reduced mortality rates by over 60% |
| Pharmacovigilance | Doubled adverse drug event detection rates |

Current Challenges Hindering Wider Adoption: Despite proven benefits, NLP adoption in healthcare still faces barriers like :

1. Interoperability: Varied EHR documentation formats and lack of standardized templates impede mining insights across healthcare networks. For example, syntactic and semantic variability in how patient history is recorded across sites.
2. Data Quality: Errors and inconsistencies in clinical documentation like ambiguity, incompleteness, misspellings negatively impact NLP accuracy. Physician frustration with EHRs exacerbates documentation quality issues.
3. Privacy & Security: Health data privacy regulations complicate access to patient records required for developing accurate NLP solutions. Hospitals also fear security threats and operational risks from third-party NLP vendors.
4. Historic Data Availability: Insufficient longitudinal patient data documented digitally obstructs identifying clinical patterns. Most records remain paper-based further limiting NLP applicability.
5. Reporting Needs: Defining precise analytics requirements needed from unstructured data is difficult for healthcare providers lacking technical capabilities and data science resources. This impacts deploying suitable NLP solutions.
6. Interpretability vs Accuracy: Inherent trade-off between performance metrics poses deployment challenges. Healthcare users mistrust black-box NLP predictions and demand explanations from transparent algorithms.

The lack of standardization coupled with data complexity introduce technical barriers, while organizational constraints related to change management cause inertia in adopting AI-based solutions like NLP.

## Future Outlook and Emerging Trends

Nonetheless, NLP is projected to become integral to mainstream medicine as a key enabler leveraging big data analytics for evidence-based care. Ongoing research focused on adapting NLP to medical language conventions and terminology usage will lead to performance improvements.

Some emerging trends shaping NLP's future role in healthcare include:

1. Transfer learning from language models like BERT allows creating accurate NLP pipelines faster across varied use cases.
2. Graph-based knowledge representations are overcoming challenges in encoding clinical contexts, relations and temporality.
3. User-centered design approaches place clinicians directly within the NLP development process to simplify adoption.
4. Hybrid human-machine solutions combining automated analysis with human-in-the-loop supervision bridge the interpretability-accuracy gap.
5. Federated learning the paradigm will facilitate collaborative learning without exposing patient data by decentralizing model development across institutions.

Overall, best practices for successfully deploying and scaling NLP strategies are emerging within healthcare groups pursuing data-driven transformation. Academic medical centers are catalyzing research on NLP's medical applications. Technology vendors are also introducing solutions suites and partnering healthcare providers on targeted analytics projects to demonstrate return on investment.

With appropriate evidence-based validation, NLP adoption is expected to rapidly accelerate, unlocking transformative benefits from big data to reimagine the future of evidence-based precision medicine.

Social Media and NLP for Healthcare Insights: Beyond medical records, social media has emerged as a vast source of patient experiences and opinions sharing valuable information often not captured through formal healthcare processes. Twitter alone generates 500 million tweets daily with increasing content about health. NLP is invaluable for converting such unstructured big data into actionable insights for improving care services.

Analysis of social conversations yields the patient perspective on healthcare issues generally unavailable in medical databases. Social NLP helps segment patient populations more intricately and enables understanding niche sub-groups. As an example, mining infertility-related conversations provided insights into differences in support needs between urban and rural patients.

Some key applications of NLP in analyzing healthcare-related social data include:

1. Infodemiology – Deriving epidemiological trends from analyzing content sharing patterns about diseases, treatments etc.
2. Sentiment analysis – Identifying patient opinions and emotions regarding healthcare services, policies, providers etc.
3. Risk factor identification – Discovering posts indicating lifestyle choices or behavior that increase health risk likelihood.

Table 3 shows sample NLP implementations on social data generating useful healthcare insights.

Table 3. Example NLP Implementations on social media for Healthcare Insights

| Data Source | Use Case | Insights Derived |
|---|---|---|
| Twitter | Tracking infectious disease outbreaks | Early warning signs for the spread of flu |
| Online health forums | Analyzing patient experiences managing diabetes | Key challenges faced in diet planning |
| Reddit | Monitoring medication side-effects | Adverse events not detected during trials |
| Youtube | Examining viewer comments on medical procedure videos | Identifying information gaps and misconceptions |
| Amazon reviews | Assessing patient feedback on medical devices | Product deficiencies and desirable improvements |

Thus, NLP applied to vast volumes of unstructured social data provides invaluable patient-centered insights to improve population health outcomes by:

1. Enabling better lifestyle disease management through health education and targeted interventions tailored to community needs and attitudes.
2. Strengthening healthcare policies, regulations, and delivery by factoring in patient experiences and preferences.
3. Developing patient engagement communication strategies addressing concerns and information gaps.

However, social NLP also raises technology ethics challenges regarding consent, transparency, and privacy that warrant careful consideration.

ORIENT review is a scientific journal publisher based in Asia, dedicated to disseminating high-quality research across various academic disciplines. With a strong commitment to fostering scholarly advancement and promoting cross-cultural understanding, ORIENT REVIEW has established itself as a reliable platform for researchers, academics, and professionals across the globe.

NLP for Biomedical Literature Mining: Tremendous amounts of unstructured knowledge is recorded in millions of published biomedical research articles and health science journals. Unlocking this knowledge has great potential to accelerate medical discoveries and scientific breakthroughs. For instance, gleaning drug-related insights from decades of publications fast-tracks developing targeted therapies. But mining such unstructured big data manually is impossible. NLP provides the techniques to automate search, extraction, and analysis - converting text from volumes of literature into machine-readable biomedical data.

Some key applications of NLP-driven biomedical literature mining are:

1. Semantic search and question answering: For researchers to find precise answers to complex questions across massive text corpora impossible through keyword search.
2. Hypothesis generation: Identifying novel connections between biomedical concepts, events, genes, proteins etc. to uncover promising research directions.
3. Biomarker discovery: Automatically associating disease mentions with genomic entities across literature to discover diagnostic markers.
4. Drug repurposing: Linking compound-protein interactions with disease pathways across publications suggests additional therapeutic indications for existing drugs.
5. Competitive intelligence: Patent mining using NLP determines research trends and activities of drug companies to understand the competitive landscape.

Core NLP techniques like named entity recognition, relation extraction, temporal reasoning, and event trigger identification enable literature mining use cases. Additionally, domain-specific vocabularies, ontologies, and metadata are used to overcome lexical and semantic complexities in scientific language.

For instance, machine reading comprehension approaches like Biomedical Extractor and Reasoner accurately answered complex questions from cancer text with over 90% precision by encoding domain knowledge into software agents.

Overall, NLP delivers the breakthrough capability of perceiving the biomedical literature as one massive, interconnected, and computer-readable knowledge graph. This creates boundless possibilities for accelerating healthcare discoveries through AI in the future.

## Conclusion

Advanced Natural Language Processing (NLP) techniques play a pivotal role in harnessing the vast volumes of unstructured big data within the healthcare domain. One of the significant applications is the automated extraction of valuable information from clinical notes, aiding clinicians in making more informed care decisions and ultimately improving patient outcomes. By employing sophisticated NLP algorithms, healthcare systems can sift through extensive clinical documentation, extracting pertinent details, and presenting a concise and relevant summary for medical professionals. This not only streamlines the information retrieval process but also ensures that critical insights are not overlooked, contributing to more efficient and effective healthcare delivery. Furthermore, the integration of NLP in healthcare extends beyond clinical settings to leverage patient-generated content on social media platforms. Analyzing patient expressions and experiences shared online provides a rich source of experiential context. This invaluable information can be utilized to enhance health services, enabling healthcare providers to better understand patient perspectives, preferences, and concerns. By tapping into the wealth of social media data, healthcare organizations can adapt their strategies, improving patient engagement and satisfaction. Additionally, monitoring social media discussions allows for the early detection of emerging health trends or potential outbreaks, facilitating proactive public health interventions.

In the realm of scientific research, NLP proves instrumental in accelerating discoveries within the biomedical domain. The vast amount of biomedical literature available poses a challenge for researchers to stay abreast of the latest advancements. NLP algorithms can efficiently process and analyze scientific texts, automatically extracting key insights, relationships, and trends. This not only expedites the literature review process but also aids researchers in identifying gaps in knowledge and potential avenues for further investigation. The ability of NLP to navigate and comprehend the intricacies of biomedical language facilitates the synthesis of information from diverse sources, fostering interdisciplinary collaboration and contributing to the rapid evolution of medical knowledge. Moreover, the application of NLP in biomedical literature mining goes beyond information retrieval. It enables the creation of knowledge graphs, semantic annotations, and other structured representations that enhance the accessibility and interoperability of research findings. Researchers can utilize these structured data outputs to build comprehensive models, perform meta-analyses, and generate new hypotheses. The result is a more dynamic and collaborative research environment, where NLP serves as a catalyst for innovation by empowering scientists with tools to navigate and extract meaningful insights from the ever-expanding body of biomedical literature.

While accuracy and interpretability challenges remain, continuous technology innovations on multiple fronts are overcoming barriers to adoption. Cloud computing scale and efficiency facilitates applying complex deep learning NLP solutions. Transfer learning and pre-trained language models accelerate development of customized applications. Graph databases capture biomedical semantics more precisely. Interactive visualizations present multi-dimensional insights simply.

As algorithms continue maturing, NLP adoption will rapidly gain momentum across the healthcare ecosystem to inform decision making at point-of-care, population health management, public policy, and life sciences R&D. This will unleash the true potential of healthcare's big data to transform medicine into a proactive, predictive, and patient-centric science focused on delivering precise, personalized, and timely interventions for optimizing health outcomes.

## References

[1] IDC. The Digital Universe in Healthcare. Available from: https://www.emc.com/collateral/analyst-reports/idc-digital-universe-healthcare-vertical-report-ar.pdf

[2] Wu Y, Jiang M, Lei J. Named Entity Recognition in Electronic Health Records Using Context Conditional Random Fields. In: Proceedings of The Fifth International Workshop on Health Text Mining and Information Analysis (Louhi) 2014.

[3] Muniswamaiah M, Agerwala T, Tappert CC. Approximate query processing for big data in heterogeneous databases. In2020 IEEE International Conference on Big Data (Big Data) 2020 Dec 10 (pp. 5765-5767). IEEE.

[4] Kamal M, Bablu TA. Machine Learning Models for Predicting Click-through Rates on social media: Factors and Performance Analysis. International Journal

of Applied Machine Learning and Computational Intelligence. 2022 Apr 7;12(4):1-4.
[5] Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF, et al. Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research. Yearb Med Inform. 2008.
[6] Weng C, Wu X, Luo Z, Boland MR, et al. EliIE: An open-source information extraction system for clinical trial eligibility criteria. J Am Med Inform Assoc. 2017 Jun 1;24(6):1062-1071.
[7] Rea S, Pathak J, Savova G, Oniki TA, Westberg L, Beebe CE, Tao C, Parker CG, Haug PJ, Huff SM, Chute CG. Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: The SHARPn project. J Biomed Inform. 2012 Aug 1;45(4):763-71.
[8] Sevenster M, Buurman J, Liu P, Peters JF, Chang C. Natural Language Processing Techniques for Extracting and Categorizing Finding Measurements in Narrative Radiology Reports. Appl Clin Inform. 2015;6(3):600-10.
[9] Nassar A, Kamal M. Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations. International Journal of Responsible Artificial Intelligence. 2021 Aug 6;11(8):1-1.
[10] Nikfarjam A, Sarker A, O'Connor K, Ginn R, Gonzalez G. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. J Am Med Inform Assoc. 2015 Mar;22(3):671-81.
[11] Saria S, Rajani AK, Gould J, Koller D, Penn AA. Integration of early physiological responses predicts later illness severity in preterm infants. Sci Transl Med. 2010 Sep 8;2(48):48ra65.
[12] Kamal M, Himel AS. Redefining Modern Marketing: An Analysis of AI and NLP's Influence on Consumer Engagement, Strategy, and Beyond. Eigenpub Review of Science and Technology. 2023 Sep 11;7(1):203-23.
[13] Muniswamaiah M, Agerwala T, Tappert CC. Federated query processing for big data in data science. In2019 IEEE International Conference on Big Data (Big Data) 2019 Dec 9 (pp. 6145-6147). IEEE.
[14] Meystre SM, Lovis C, Bürkle T, Tognola G, Budrionis A, Lehmann CU. Clinical data reuse or secondary use: Current status and potential future progress. Yearb Med Inform. 2017 Aug;26(01):38-52.
[15] De Silva N, Ranasinghe W, Bandar ZA, Schiffman SS. Perspective: The role of artificial intelligence in modern health care systems. AIMS Public Health. 2021 Jan 1;8(2):360-75.
[16] Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, Kang J. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics. 2020 Feb 15;36(4):1234-40.
[17] Choi E, Bahadori MT, Song L, Stewart WF, Sun J. GRAM: Graph-based Attention Model for Healthcare Representation Learning. In: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining 2017 Aug 13 (pp. 787-795).
[18] Nassar A, Kamal M. Machine Learning and Big Data Analytics for Cybersecurity Threat Detection: A Holistic Review of Techniques and Case Studies. Journal of Artificial Intelligence and Machine Learning in Management. 2021 Feb 6;5(1):51-63.

[19] Meystre SM, Kim E, Chevrier RD, Hurdle JF. The current state of clinical natural language processing for the family history in electronic health records: An expert review. Yearb Med Inform. 2017 Aug;26(01):76-82.
[20] Chen JH, Asch SM. Machine Learning and Prediction in Medicine—Beyond the Peak of Inflated Expectations. N Engl J Med. 2017 Jun 29;376(26):2507-9.
[21] Muniswamaiah M, Agerwala T, Tappert CC. Context-aware query performance optimization for big data analytics in healthcare. In2019 IEEE High Performance Extreme Computing Conference (HPEC-2019) 2019 (pp. 1-7).
[22] Brisimi TS, Chen R, Mela T, Olshevsky A, Paschalidis IC, Shi W. Federated learning of predictive models from federated Electronic Health Records. Int J Med Inform. 2018 Jan 1;112:59-67.
[23] Hayes J, Bach SA, Dakkak H. NLP-based approaches for analyzing patient experience shared on social media: an empirical study. BMC Med Inform Decis Mak. 2020 Dec;20(1):1-2.
[24] Kamal M, Bablu TA. Mobile Applications Empowering Smallholder Farmers: An Analysis of the Impact on Agricultural Development. International Journal of Social Analytics. 2023 Jun 12;8(6):36-52.
[25] Muniswamaiah M, Agerwala T, Tappert C. Big data in cloud computing review and opportunities. arXiv preprint arXiv:1912.10821. 2019 Dec 17.
[26] Sarker A, González-Hernández G. An Overview of social media for Natural Disaster Management: extractive techniques, analysis frameworks, and real-world applications.