AI in Social Determinants of Health Analysis and Intervention: A New Frontier in Health Equity

Rasit Dinc

Rasit Dinc Digital Health & AI Research

Published: June 15, 2025 | AI Cardiology

DOI: 10.5281/zenodo.17996664

Abstract

AI in Social Determinants of Health Analysis and Intervention: A New Frontier in Health Equity The Social Determinants of Health (SDoH)—the conditions i...

AI in Social Determinants of Health Analysis and Intervention: A New Frontier in Health Equity

The **Social Determinants of Health (SDoH)**—the conditions in which people are born, grow, live, work, and age—are now widely recognized as the primary drivers of health outcomes, accounting for up to 80% of modifiable health factors [1]. Addressing SDoH is critical for achieving health equity, yet the sheer volume, complexity, and unstructured nature of the data involved have historically posed a significant challenge to effective analysis and intervention. Enter **Artificial Intelligence (AI)**, a transformative technology that is rapidly emerging as a powerful tool to unlock the potential of SDoH data and reshape public health strategies.

The Data Challenge and AI’s Analytical Power

SDoH data is inherently diverse, spanning clinical records, census data, geographical information systems (GIS), social media, and community-level surveys. Much of this crucial information, such as housing instability, food insecurity, or transportation barriers, is often buried within the unstructured text of electronic health records (EHRs), including clinical notes and social work assessments.

AI, particularly **Natural Language Processing (NLP)** and **Machine Learning (ML)**, is proving instrumental in overcoming this data fragmentation [2].

Natural Language Processing (NLP) for Data Extraction

NLP, a subfield of AI, is revolutionizing the extraction of SDoH from clinical narratives. Traditional methods of manual chart review are time-consuming

and prone to human error. NLP models, however, can automatically scan vast amounts of unstructured text to identify, classify, and extract relevant SDoH concepts.

| NLP Technique | Application in SDoH Analysis | | :--- | :--- | | **Rule-Based NLP** | Identifying pre-defined keywords and phrases related to SDoH (e.g., "homeless," "unemployed"). | | **Machine Learning NLP** | Training models (e.g., deep learning) to recognize complex SDoH contexts and sentiment in text. | | **Word Embeddings (e.g., Word2Vec)** | Creating vector representations of words to understand semantic relationships, allowing for the identification of related SDoH concepts even if the exact phrase is not used. |

Studies have demonstrated that NLP models can achieve high specificity (nearing 95.9%) in identifying SDoH factors, significantly augmenting the retrieval of information hidden in clinical notes [2]. This high-throughput extraction is essential for creating comprehensive patient profiles that extend beyond purely clinical data.

Machine Learning for Predictive Modeling

Once SDoH data is extracted and structured, ML algorithms are deployed for predictive modeling and risk stratification. Algorithms such as **Random Forest**, **Support Vector Machines**, and **Neural Networks** are commonly used to:

1. **Risk Prediction:** Predict adverse health outcomes, such as hospital readmissions, emergency department (ED) revisits, or in-hospital mortality, by incorporating SDoH variables [3]. For instance, ML models incorporating SDoH have been shown to improve risk prediction for in-hospital mortality after hospitalization for heart failure, particularly in vulnerable populations [4]. 2. **Classification and Imputation:** Classify patients into risk groups based on their SDoH profile and impute missing SDoH data, which is a common challenge in real-world datasets. 3. **Targeted Intervention:** Identify specific SDoH factors that are most strongly associated with poor outcomes, allowing health systems to allocate resources more effectively and design targeted interventions.

From Analysis to Intervention: The Path to Health Equity

The true value of AI in SDoH lies not just in analysis but in its application to intervention strategies. By providing a granular, real-time understanding of a patient's social context, AI facilitates a shift from reactive care to proactive, personalized intervention.

Personalized and Proactive Care

AI-driven insights enable healthcare providers to move beyond a one-size-fits-all approach. For a patient identified by an AI model as high-risk due to a combination of chronic disease and transportation barriers, the intervention can be immediately tailored to address the social factor—for example, by automatically generating a referral to a community-based transportation service rather than simply adjusting medication.

Ethical Considerations and Explainable AI (XAI)

As AI systems become more integrated into SDoH analysis, ethical considerations surrounding bias and fairness are paramount. If AI models are trained on biased data, they risk perpetuating or even amplifying existing health disparities. This is why the concept of **Explainable AI (XAI)** is gaining traction in this domain [5].

XAI aims to make the decision-making process of AI models transparent and understandable. By using XAI techniques, researchers can audit models to ensure they are not unfairly penalizing specific demographic groups based on SDoH factors. The goal is to build **equitable AI** systems that promote fairness and trust among diverse patient populations.

Conclusion

AI is not merely an incremental improvement in digital health; it represents a fundamental shift in how we understand and address the root causes of health inequity. By leveraging the power of NLP to extract hidden SDoH data and ML to build predictive, risk-stratifying models, healthcare systems can transition to a more holistic, personalized, and proactive model of care. The continued development and ethical deployment of AI in SDoH analysis and intervention promise a future where health outcomes are less determined by social circumstances and more by effective, data-driven action.

References

[1] Magnan, S. (2017). *Social Determinants of Health 101: Key Concepts and Ways to Address*. Health Affairs*. [2] Abbott, E. E., Apakama, D., Richardson, L. D., Chan, L., & Nadkarni, G. N. (2024).* ***Leveraging Artificial Intelligence and Data Science for Integration of Social Determinants of Health in Emergency Medicine: Scoping Review****.* JMIR Medical Informatics*, 12(1), e57124. [https://medinform.jmir.org/2024/1/e57124] (https://medinform.jmir.org/2024/1/e57124) [3] Ong, J. C. L. (2024).* ***Artificial intelligence, ChatGPT, and other large language models in social determinants of health research****.* npj Digital Medicine*, 7(1), 1-3. [https://pmc.ncbi.nlm.nih.gov/articles/PMC10829781/] (https://pmc.ncbi.nlm.nih.gov/articles/PMC10829781/) [4] Segar, M. W., et al. (2022).* ***Machine Learning–Based Models Incorporating Social Determinants of Health Improve Risk Prediction of In-Hospital Mortality After Hospitalization for Heart Failure****.* Circulation: Cardiovascular Quality and Outcomes*, 15(7), e008905. [https://pmc.ncbi.nlm.nih.gov/articles/PMC9260645/] (https://pmc.ncbi.nlm.nih.gov/articles/PMC9260645/) [5] Yan, J. (2025).* ***Revisiting the social determinants of health with explainable AI****.* American Journal of Epidemiology*, 194(3), 365-367. [https://academic.oup.com/aje/advance-article/doi/10.1093/aje/kwaf205/8256413] (https://academic.oup.com/aje/advance-article/doi/10.1093/aje/kwaf205/8256413)