Ramezani *et al. BioData Mining* (2025) 18:14
https://doi.org/10.1186/s13040-025-00428-x

BioData Mining

RESEARCH

Open Access

Agenda setting for health equity assessment through the lenses of social determinants of health using machine learning approach: a framework and preliminary pilot study

Maryam Ramezani[1,2†], Mohammadreza Mobinizadeh[3†], Ahad Bakhtiari[1,2], Hamid R. Rabiee[4], Maryam Ramezani[4], Hakimeh Mostafavi[2], Alireza Olyaeemanesh[3], Ali Akbar Fazaeli[1,2], Alireza Atashi[5], Saharnaz Sazgarnejad[6,7], Efat Mohamadi[2] and Amirhossein Takian[1,2,8*]

†Maryam Ramezani and Mohammadreza Mobinizadeh first author-equal contribution.

*Correspondence:
Amirhossein Takian
takian@tums.ac.ir

Full list of author information is available at the end of the article

Abstract

Introduction The integration of Artificial Intelligence (AI) and Machine Learning (ML) is transforming public health by enhancing the assessment and mitigation of health inequities. As the use of AI tools, especially ML techniques, rises, they play a pivotal role in informing policies that promote a more equitable society. This study aims to develop a framework utilizing ML to analyze health system data and set agendas for health equity interventions, focusing on social determinants of health (SDH).

Method This study utilized the CRISP-ML(Q) model to introduce a platform for health equity assessment, facilitating its design and implementation in health systems. Initially, a conceptual model was developed through a comprehensive literature review and document analysis. A pilot implementation was conducted to test the feasibility and effectiveness of using ML algorithms in assessing health equity. Life expectancy was chosen as the health outcome for this pilot; data from 2000 to 2020 with 140 features was cleaned, transformed, and prepared for modeling. Multiple ML models were developed and evaluated using SPSS Modeler software version 18.0.

Results ML algorithms effectively identified key SDH influencing life expectancy. Among algorithms, the Linear Discriminant algorithm as classification model was selected as the best model due to its high accuracy in both testing and training phases, its strong performance in identifying key features, and its good generalizability to new data. Additionally, CHAID in numeric models was the best for predicting the actual value of life expectancy based on various features. These models highlighted the importance of features like current health expenditure, domestic general government health expenditure, and GDP in predicting life expectancy.

Conclusion The findings underscore the significance of employing innovative methods like CRISP-ML(Q) and ML algorithms to enhance health equity. Integrating this platform into health systems can help countries better prioritize and address health inequities. The pilot implementation demonstrated these methods' practical applicability and effectiveness, aiding policymakers in making informed decisions to improve health equity.

Keywords Equity, Social determinant of health (SDH), Machine learning

Introduction

Health equity and Social Determinants of Health (SDH) play an overriding role in public health [1–3]. Health inequities arising from social injustices are preventable, underscoring the importance of addressing upstream SDH to improve population health and equity. Evaluating health equity ensures improved life expectancy and healthier lives for everyone, regardless of socioeconomic status. Understanding the connections between various SDH and health outcomes is necessary to develop helpful interventions and programs. This approach fosters the development of strategic policies to tackle the fundamental causes of health inequities and facilitate sustainable improvements in population health [1, 2].

AI tools are being extensively developed and can be utilized to adopt policies aimed at creating a more equitable world, as well as for controlling and managing various problems [4–6]. Moreover, AI tools play a significant role in achieving health-related Sustainable Development Goals (SDG), Universal Health Coverage (UHC) [7], reducing social and economic disparities, and improving health outcomes [8] particularly within high-income countries [7]. The importance of advances in AI tools to reduce poverty and improve public services has been gaining attention in recent international meetings and assemblies [7].

Within AI, ML is a powerful tool that allows systems to plan and make decisions based on data and past experiences without the need for explicit programming [9]. ML, as a branch of AI, can be utilized to improve health systems' function through hypothesis generation and testing and uncovering data trends [6]. The term "ML" refers to a diverse array of models and techniques focused on algorithmic modeling [10]. It can be categorized into three main groups: supervised learning (using known patterns in training data), unsupervised learning (finding new patterns in data), and reinforcement learning (using rewards and punishments in a dynamic environment) [7]. The application of ML to health data modeling is becoming increasingly common, offering significant advancements in our understanding of health and potential interventions [10]. These methods surpass traditional capabilities by integrating mathematics, statistics, and computer science, driving the development of advanced computational systems [11]. Developing curated data and the capacity to manage strategic changes is a fundamental step to leveraging AI and ML. ML has the potential to inform more equitable health interventions by analyzing datasets without prior assumptions and uncovering patterns. Thus, it can help identify and address disparities in health outcomes [6, 7].

Studies have shown that AI tools can identify gender disparities in finance, healthcare, human development, psychology, security, equity, and socio-cultural aspects [12]. Creative applications of AI can be deployed to track equity among marginalized populations, such as people with disabilities, ultimately supporting advocacy for these groups [13]. AI technologies can also identify racial subgroups suffering from inequity [14]. Various health indicators can be assessed and monitored using these methods, and inequalities can be measured for them. One of the most important of these indicators is life expectancy, which is considered a key metric for evaluating health systems.

Life expectancy, the average number of years a person is expected to live from birth to death, is a key indicator for assessing human health. It also comprehensively measures economic development, education, and health systems. Most existing studies focus on qualitative analyses of one or a few factors, needing more quantitative analyses of multiple factors. This gap makes it difficult to identify the predominant factor influencing life expectancy precisely. Given the various conditions and complications current in the society today, many factors must be considered to predict life expectancy accurately. Consequently, various ML models have been developed to address this need [3]. Given the importance of utilizing cutting-edge sciences and ML algorithms to predict and interpret accessible data to overcome challenges and complex problems, the vital role of AI and ML has become evident.

By employing advanced ML algorithms, this research aims to uncover deeper insights into the factors contributing to health inequities. These insights are essential for informing policy recommendations that effectively target interventions and promote equitable health outcomes. Specifically, the study focuses on the applicability of various ML algorithms in assessing health equity related to SDH. To guide this exploration, we pose two critical research questions: How can ML algorithms be utilized to set the agenda for health equity based on SDH? Which ML algorithms, as a pilot, accurately identify the SDH that influences life expectancy and assist in prioritizing health equity interventions?

Through addressing these questions, this study seeks to contribute to the growing body of knowledge on the intersection of ML and public health, ultimately paving the way for more informed and equitable health policies.

Method

Accordingly, this study used the CRISP-ML(Q) (Cross-Industry Standard Process for Machine Learning with Quality Assurance) method. This iterative approach allows for backtracking to earlier phases at each level of the investigation. This method encompasses six phases, from scope definition to deployment application maintenance, and integrates quality assurance methodologies to address ML development challenges [15].

Although ML is widely used in various fields, a standard process model must be used to enhance the success and efficiency of ML applications. Meanwhile, the outcomes of ML initiatives can be evaluated at three levels: business success, ML success, and economic success.

CRISP-ML(Q) method phases

The six steps of CRISP-ML(Q) were applied in this study. Monitoring and maintenance were not performed due to the study's limited scope. Future work can include these aspects to ensure long-term performance [15].

Business and data understanding

A conceptual model was developed based on a desk review and document analysis, validated by the research team to define business goals and understand the available data. A variety of key documents were reviewed in this study. Including laws, development plans, and health-related frameworks from the Islamic Republic of Iran (Supplementary

File 1). Additionally, relevant data from 2000 to 2020 was collected from the Global Health Observatory, World Bank, and WHO EMRO.

Data preparation

Clean, transform, and prepare data for modeling. The analysis was conducted using 20 years of data, with each year comprising 140 different features collected from a variety of sources. Prior to model development, a thorough data preprocessing procedure was implemented. This phase involved meticulously filling in any missing data points and then employing feature engineering methods to significantly reduce the dimensionality of the dataset. Life expectancy was selected as the health outcome in this study. Preprocessing included deleting variables with inadequate data, removing missing data using the Missing Data Handling option in SPSS Statistics under the Transform menu. Missing values in each feature were imputed using a linear trend interpolation method, accounting for the temporal sequence of the time-series data across multiple years. This method uses a straightforward regression model where the variable containing missing values functions as the dependent variable and the case sequence number (representing years) serves as the independent variable.

A feature selection process, using SPSS Modeler, was conducted to identify a subset of 140 features, selecting only the most pertinent and non-redundant variables. Employing statistical analysis, the study examined correlations and other univariate measures related to the dependent variable of life expectancy. Variables exhibiting strong, significant correlations with life expectancy were kept; those with weak or redundant associations were excluded. The importance of the selected features was assessed to ensure that the preserved variables provided unique and valuable contributions, minimizing redundancy. Feature engineering techniques decreased the number of features from 140 to 20, resolving high dimensionality relative to the sample size. This reduction enhanced computational efficiency, mitigated overfitting, and improved the generalizability of machine learning models.

Modeling

Because of the continuous distribution of life expectancy, we investigated the issue from two different viewpoints. Employing regression, our "numeric models" approach treated life expectancy as a continuous variable for direct estimation. The second method involved discretizing the continuous life expectancy data, converting the solution of problem into "classification models". The life expectancy values were categorized into three distinct groups: the first category includes values greater than or equal to 69 and less than 72, the second category includes values greater than or equal to 72 and less than 75, and the third category includes values greater than or equal to 75 and less than 78. The dual perspective facilitated the precise numerical estimation of life expectancy and its stratification into meaningful categories to enhance decision-making. SPSS AutoML uses an automated procedure to systematically assess a diverse array of algorithms, using key performance indicators, such as accuracy, which is evaluated via cross-validation. Employing AutoML in SPSS Modeler, regression, CHAID, and generalized linear models proved to be the optimal numeric predictors of life expectancy. Within the classification domain, automated machine learning (AutoML) selected Discriminator, Logistic Regression, and Bayesian Network algorithms as the highest-performing models.

Evaluation metrics

To ensure the accuracy of this model, the Analyze node was used. The performance of the predictive models is evaluated using various metrics depending on the type of models. Two primary categories of metrics are.

*From a numerical modeling standpoint, we use the following metrics:

(1) Minimum Error and (2) Maximum Error that represent the smallest and largest discrepancies between predicted and observed values. This analysis reveals the optimal and least optimal model prediction outcomes in relation to the true value.

(3) Mean Error calculates the average difference between predicted and actual values. While it reflects overall model accuracy, it may be biased by positive or negative errors in the predictions.

(4) Mean Absolute Error calculates the average of the absolute differences between the predicted and the actual values. In contrast to mean error, mean absolute error disregards the sign of errors, providing a more direct assessment of predictive accuracy.

(5) Standard Deviation measures the variability of the prediction errors. High standard deviation signifies inconsistent model predictions, whereas low standard deviation denotes stable performance.

(6) Linear Correlation assesses the strength and direction of the linear relationship between predicted and actual values. A correlation close to + 1 means a strong positive relationship, while values close to 0 suggest no linear relationship.

For classification models, which predict the categorical outcome for life expectancy the Coincidence Matrix (also known as the Confusion Matrix) is used to evaluate the performance. The Coincidence Matrix provides detailed information on the true positives, true negatives, false positives, and false negatives, enabling a more thorough analysis of the model's classification accuracy. Number of correctly predicted (Correct) and incorrectly predicted (Wrong) values are the other utilized metrics. These metrics reflect the number of samples where the model's predicted value for life expectancy category exactly matched the true value (Correct) or deviated from it (Wrong).

Reporting the 95% Confidence Interval (CI) for training and testing results ensures the robustness and reliability of the model's performance metrics. CI provides a range within which the true value of a metric is likely to fall, accounting for variability in the data. To enhance statistical rigor, the models were run 10 times with different data splits, and the results were aggregated to calculate the CIs. This process demonstrates that the reported performance metrics are not artifacts of random fluctuations and are representative of the model's generalizability.

Deployment

Deploy the model into a production environment. Since this study focused on developing and evaluating the pilot model, the deployment phase was not included. Future research can consider deployment and operationalization [15].

Monitoring and maintenance

Continuously track the model's performance and make necessary updates. Continuously track the model's performance and make necessary updates. Future research can consider deployment and operationalization. These two phases should be prioritized as part of the final platform model's development agenda. Using machine learning to reduce health inequities requires robust infrastructure development to support these phases. By focusing on deployment and maintenance, future efforts can ensure the long-term effectiveness and sustainability of the machine learning models in addressing health disparities.

Results

The results of this study are presented in three main sections: (1) Equity indicators in Upstream Documents and review, (2) ML-based framework for inequity assessment, and (3) Pilot implementation of the framework.

Equity indicators in upstream documents and review

After extracting the components through the upstream documents and desk review analysis, they were categorized and reported according to Fig. 1 (a complete list of indicators is provided in Appendix 1).

This framework integrates the six building blocks of the health system and Pickett's model of equity. The categorized indicators provide a comprehensive view of equity across health system dimensions.

ML-based framework for inequity ASSESSMENT

The proposed model (Supervised type) suggests the necessary data repositories for ML-based equity calculations. Based on the framework for ML analysis, there is a need for databases related to health outcomes and SDH, including data related to equity (Fig. 2).

This table outlines the necessary components for an inequity assessment framework using machine learning (ML):

Health Outcomes Repository: Includes data on life expectancy and various mortality rates (e.g., mental disorders, chronic respiratory diseases, cancer, infant mortality, cardiovascular diseases, traffic accidents, pregnant women, children under five, communicable and non-communicable diseases, diabetes).

SDH and Healthcare Data Repository: Contains data on healthcare provision and social determinants of health.

Equity Data Repository: Includes data on various demographic and social groups (e.g., racial and ethnic groups, gender, indigenous populations, age, color, wealth, disability, gender identity, sexual orientation, religion, nationality, identity, social class).

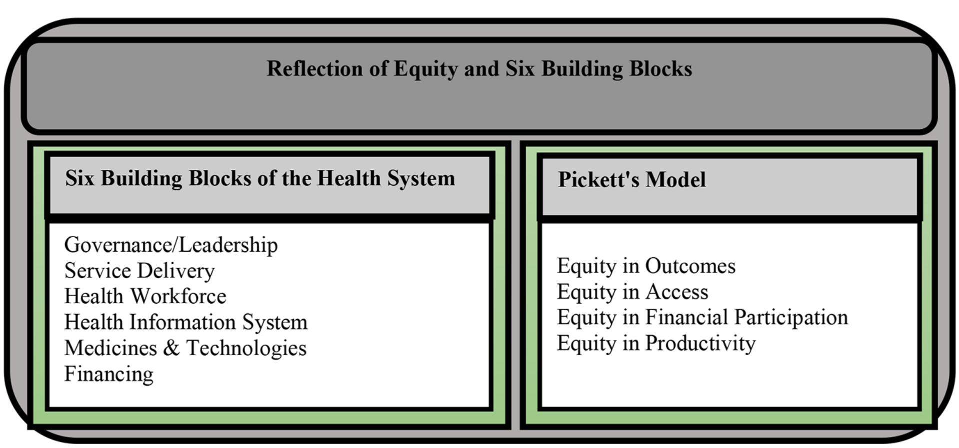

Fig. 1 Equity indicators framework based on Upstream Documents and review

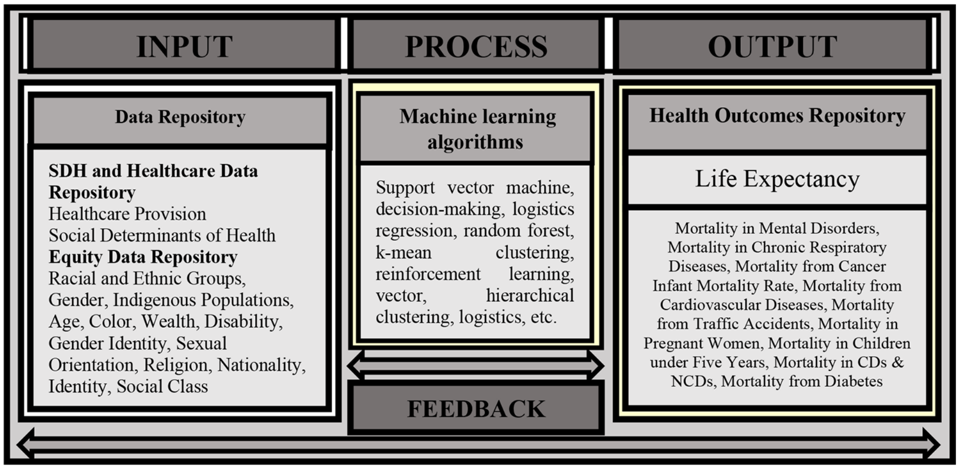

Fig. 2 Inequity assessment framework using ML

Pilot implementation of the framework

A pilot implementation of the framework was conducted to test the feasibility and effectiveness of using ML algorithms in assessing health equity. Life expectancy was chosen as the health outcome (as the target). This section is divided into two main parts: classification and numeric. The classification part includes algorithms such as Regression, CHAID, and Generalized Linear, while the numeric part includes algorithms such as Linear Discriminant, Logistic Regression, and Bayesian Network.

Feature selection analysis

This analysis helps identify key features that should be considered in the modeling process to improve the accuracy of predictions. The important features identified can be used as primary inputs in machine learning models to provide more accurate predictions of life expectancy as shown in Table 1.

These include risk of catastrophic expenditure for surgical care, social insurance contributions, domestic general government health expenditure per capita, and other health-related expenditures. These features directly impact life expectancy and are used in the modeling process.

Numeric models

The numeric models were used to predict value of data based on the selected features. The following algorithms were then applied: Regression, Generalized Linear Models and CHAID. The performance of these numeric algorithms is evaluated, demonstrating their ability to accurately predict the value of life expectancy.

Regression algorithm analysis

The Regression algorithm was used to analyze the data based on the selected features. The performance of the Regression algorithm was evaluated using various metrics, as Table 2 shows that despite good performance on training data we have MAE on test dataset is 3.48, with 95% Confidence Interval including CI for train (0.00,0.00) and CI for test (0.78,6.18), indicating that the true MAE value is likely to fall within this range.

Table 1 Feature selection analysis

Field
Risk of catastrophic expenditure for surgical care of people at risk
Social insurance contributions
Domestic general government health expenditure per capita PPP current
Risk of impoverishing expenditure for surgical care of people at risk
Current Health Expenditure CHE
Current health expenditure of GDP
Domestic general government health expenditure of GDP
Out of pocket expenditure of current health expenditure
Social Health Insurance SHI as of Current Health Expenditure CHE
Gross Domestic Product GDP
Current health expenditure per capita PPP current international
General government expenditure GGE
Domestic Private Health Expenditure PVTD
Domestic Private Health Expenditure PVTD as Current Health Expenditure CH
Domestic General Government Health Expenditure GGHED as Current Health Ex
Household out of pocket payment
General Government Expenditure GGE as Gross Domestic Product GDP
Domestic private health expenditure per capita PPP current internation
Out of pocket expenditure per capita PPP current international
GDP per capita PPP current international

Table 2 Performance of regression algorithm model

Performance Metrics	Testing	Training
Minimum Error	-3.875	-0.0
Maximum Error	7.093	0.0
Mean Error	-0.134	0.0
Mean Absolute Error	3.48	0.0
Standard Deviation	4.359	0.0
Linear Correlation	-0.132	1.0
Minimum Error	-3.875	-0.0

Table 3 Performance of generalized linear algorithm model

Performance Metrics	Testing	Training
Minimum Error	-3.375	-0.0
Maximum Error	14.242	0.0
Mean Error	0.876	-0.0
Mean Absolute Error	4.089	0.0
Standard Deviation	6.709	0.0
Linear Correlation	-0.559	1.0
Minimum Error	-3.375	-0.0

Generalized linear algorithm

This algorithm assumes that the target value (life expectancy) is generated from normal distribution. The hyper-parameters for out experiments are: Scale parameter method: Maximum likelihood estimate, Maximum iterations: 100, Singularity tolerance: 1E-007, Link function: Identity. Results in Table 3 demonstrates that its assumptions are unsuitable for the present dataset. The 95% Confidence Interval for train is (0.00,0.00) and for test is (– 0.07,8.25).

Table 4 Performance of CHAID algorithm model

Performance Metrics	Testing	Training
Minimum Error	-0.5	0.0
Maximum Error	2.9	0.0
Mean Error	0.633	0.0
Mean Absolute Error	0.8	0.0
Standard Deviation	1.282	0.0
Linear Correlation	0.935	1.0
Minimum Error	-0.5	0.0

Table 5 Linear discriminant accuracy

	Testing	Accuracy	Training	Accuracy
Correct	5	100%	16	100%
Wrong	0	0%	0	0%
Total	5	100%	16	100%

CHAID algorithm analysis

The CHAID (Chi-squared Automatic Interaction Detector) algorithm was used to analyze the data based on the selected features. The most important feature identified by the CHAID algorithm is domestic general government health expenditure as a percentage of current health expenditure, followed by current health expenditure per capita and domestic private health expenditure per capita. These features play a significant role in predicting life expectancy. The performance of the CHAID Algorithm was evaluated using various metrics, as shown in Table 4. The 95% Confidence Interval for train is (0.00,0.00) and for test is (0.01,1.59).

The CHAID algorithm demonstrated exceptional performance in both the testing and training phases. The linear correlation values of 1.0 and 0.935 in train and test partitions indicate a perfect correlation between the predicted and actual values. The mean absolute error values of 0.8 suggest that the model's predictions are extremely close to the actual values.

Overall, in numeric view, the CHAID algorithm demonstrated outstanding performance in predicting the value of life expectancy based on the selected features. The perfect linear correlation and minimal error values indicate that the model is highly effective in capturing the relationship between the features and the target variable.

Classification models

Linear discriminant algorithm analysis

The Discriminator algorithm was used to classify the data based on the selected features. The most important feature identified by the Discriminator algorithm is the current health expenditure, followed by domestic general government health expenditure as a percentage of current health expenditure, and domestic private health expenditure as a percentage of current health expenditure. These features play a significant role in predicting life expectancy.

The performance of the discriminator algorithm was evaluated using various metrics, as shown in Tables 5 and 6. The train 95% Confidence Interval is (100%,100%) and the test CI is (100%,100%).

This table shows the accuracy of the Linear Discriminant model on both the testing and training datasets.

Table 6 Linear discriminant confusion matrix

Testing	**L1**	**L2**	**L3**
L1	2	0	0
L2	0	1	0
L3	0	0	2
Training	**L1**	**L2**	**L3**
L1	7	0	0
L2	0	5	0
L3	0	0	4

Table 7 Logistic regression accuracy

	Testing	**Accuracy**	**Training**	**Accuracy**
Correct	4	80%	14	87.5%
Wrong	1	20%	2	12.5%
Total	5	100%	16	100%

Table 8 Logistic regression confusion matrix

Testing	**L1**	**L2**	**L3**
L1	1	0	1
L2	0	1	0
L3	0	0	2
Training	**L1**	**L2**	**L3**
L1	5	0	2
L2	0	5	0
L3	0	0	4

This table presents the confusion matrix of the Linear Discriminant model's classification results for both the testing and training datasets.

Overall, the Discriminator algorithm demonstrated strong performance in classifying the data and identifying key features that impact life expectancy.

Logistic regression algorithm analysis

The Logistic Regression algorithm was used to classify the data based on the selected features. The most important feature identified by the Logistic Regression algorithm is GDP per capita growth annual, followed by GDP per capita PPP current international, and current health expenditure. These features play a significant role in predicting life expectancy. The train 95% Confidence Interval is (87.45%,87.55%) and the test CI is (79.86%,80.14%).

The performance of the logistic regression algorithm was evaluated using various metrics, as shown in Tables 7 and 8.

This table shows the accuracy of the Linear Discriminant model on both the testing and training datasets.

This table presents the confusion matrix of the Linear Discriminant model's classification results for both the testing and training datasets.

Bayesian network algorithm analysis

The Bayesian Network algorithm was used to classify the data based on the selected features. The most important feature identified by the Bayesian Network algorithm is current health expenditure, followed by domestic general government health expenditure

Table 9 Bayesian network accuracy

	Testing	Accuracy	Training	Accuracy
Correct	3	60%	14	87.5%
Wrong	2	40%	2	12.5%
Total	5	100%	16	100%

Table 10 Bayesian network confusion

Testing	L1	L2	L3
L1	1	0	1
L2	0	2	1
L3	0	0	0
Training	**L1**	**L2**	**L3**
L1	5	0	0
L2	0	5	2
L3	0	0	4

as a percentage of current health expenditure, and domestic private health expenditure as a percentage of current health expenditure. These features play a significant role in predicting life expectancy. The performance of the Bayesian network algorithm was evaluated using various metrics, as shown in Tables 9 and 10. The train 95% Confidence Interval is (87.45%,87.55%) and the test CI is (59.79%,60.21%).

This table shows the accuracy of the Linear Discriminant model on both the testing and training datasets.

This table presents the confusion matrix of the Linear Discriminant model's classification results for both the testing and training datasets.

Linear Discriminant surpassed alternative methodologies, including Logistic Regression and Bayesian Networks, because of its inherent simplicity and efficacy with limited datasets characterized by many features. The data is assumed to be normally distributed, with all classes exhibiting equal covariance matrices. Using these assumptions serves to decrease model complexity and prevent overfitting, a prevalent issue when the number of features approaches the sample size. In addition, Linear Discriminant Analysis's dimensionality reduction, a consequence of projecting data onto a space that optimizes class separability, is advantageous when data is scarce.

Discussion

Health equity can be defined as providing an equal opportunity for all individuals to have a healthy life. In this regard, AI has emerged as a potential tool to create the necessary conditions for equitable access to health [16–18]. Therefore, health equity considerations should be incorporated into AI tools modeling to ensure equal opportunity for everyone in terms of having a healthy life. In this study, we proposed a model capable of establishing an appropriate data structure for examining the impact of SDH on equity. Among the various models tested, the Linear Discriminant algorithm was identified as the best performing model due to its high accuracy, strong feature identification, and good generalizability.

The effective implementation of equity strategies requires various measures [19]. For instance, further resampling or stratification may be necessary to achieve algorithmic fairness in patients of multiple vulnerable groups. Moreover, to ensure accurate predictions across diverse populations, the necessity of considering SDH and equity in ML applications has been revealed [18]. CRISP-ML(Q) is a promising tool for addressing current challenges by covering all development phases, from project idea formulation to maintaining and monitoring an existing ML application. Subsequently, CRISP-ML(Q) includes tasks to ensure quality throughout the ML development process, making it stable and general enough to support various knowledge discovery scenarios [15].

To evaluate health equity with a focus on social determinants of health, a life course perspective is essential. This perspective helps us understand and interpret racial and ethnic patterns in neuropsychological test performance. Contextual factors shape the environmental conditions encountered by racial and ethnic minorities, including geographic segregation, migration patterns, socioeconomic position, discrimination, and group resources [20]. Employing ML techniques to forecast life expectancy offers significant insights into the economic and health advancements of countries [11]. This approach ensures the development of ML tools in a method that promotes health equity [21]. Thus, life expectancy was selected by the panel of experts, comprising the research team and several specialists in the field of equity.

The Health Equity Across the AI Lifecycle (HEAAL) framework was developed to address the gaps in regulatory frameworks, accountability measures, and governance standards for AI in healthcare. HEAAL assesses five equity domains—accountability, fairness, fitness for purpose, reliability and validity, and transparency—across eight key decision points in the AI adoption lifecycle. It provides procedures for evaluating existing AI solutions and procedures for new ones, guiding healthcare organizations in mitigating the risk of AI exacerbating health inequities [22].

Biases in AI can perpetuate and exacerbate racial and ethnic inequities. Ensuring equity in algorithms should be a priority, but the lack of diversity in the AI field is concerning. There is a need for regulation and rigorous testing of algorithms to ensure their accuracy and fairness. Ethical standards for AI in healthcare are essential, and promoting transparency and accountability is crucial. To maximize the benefits of AI in healthcare, it must be approached with an equity lens during all phases of development. This approach can help reduce health disparities and ensure that the advantages of AI are distributed fairly across all population groups [23]. Additionally, regulatory strategies for uprooting bias in healthcare AI have been proposed to ensure ethical integration into health systems. These strategies emphasize the need for consensus around the regulation of algorithmic bias at the policy level, highlighting three overarching principles in bias mitigation that map to each phase of the algorithm lifecycle [24].

The HEAL framework, designed to quantitatively assess the performance equity of health AI technologies, is an example of such an approach. This framework evaluates whether AI models prioritize performance for subpopulations with worse health outcomes. For instance, in a case study applying the HEAL framework to a dermatology AI model, the HEAL metric showed that the model performed better for racial/ethnic subpopulations with poorer health outcomes [1]. Another study utilized the Behavioral Risk Factor Surveillance System (BRFSS) data to identify social determinants of health associated with inequities in self-rated health. This study highlighted the importance of using weighted logistic regression to predict health outcomes accurately, emphasizing the need for representative data to avoid biases [25]. MLOps (Machine Learning Operations) in healthcare is another critical area that addresses the challenges of deploying AI/ML tools in clinical workflows. Adherence to MLOps best practices ensures that AI/

ML models are generalizable, integrated, and robust, which is essential for maintaining health equity. This includes continuous monitoring and updating of models to detect and mitigate fairness drift [26]. Furthermore, a framework for identifying and mitigating biases in ML models has been proposed to ensure fair and equitable outcomes in public health. This framework provides guidance on incorporating fairness into different stages of the ML pipeline, from data processing to model evaluation, to prevent systematic biases and promote health equity [27].

ML has demonstrated significant potential in enhancing health equity by analyzing large datasets to identify patterns and factors that contribute to health disparities and address inequalities in health outcomes [21]. Based on the analyzed studies, ML can be beneficial in different ways; for example, ML can be used to analyze various data derived from social media content and the Google search engine. Although the extra data collection and analysis accelerate this effort, experts who generate the data must comprehend how their data are applied in developing ML algorithms [21]. ML algorithms can identify at-risk populations and predict health outcomes by analyzing data from various sources, such as electronic health records, social media, and public health surveys. According to the findings, designing the targeted interventions is one of the primary applications of ML in promoting health equity. Therefore, AI technologies enable healthcare providers and policymakers to design and implement interventions based on the specific needs of disadvantaged communities. ML can ensure equitable resource allocation by predicting which areas or populations are most qualified for using the medical resources [21].

This study indicated that SDHs are predictors of health outcomes. This aligns with the well-established understanding that sociodemographic and socioeconomic factors are crucial determinants of population health [28]. This study applied a case study from a part of the framework to life expectancy as an important indicator, demonstrating its potential to encourage explicit health equity assessment with ML. These predictive models are instrumental in pinpointing specific factors that affect average life expectancy, thereby assisting in developing strategies for enhancement. By integrating additional variables into the models through an expanded dataset, the accuracy of predictions can be further improved. This comprehensive analysis can serve as a valuable resource for NGOs, corporate entities, and governmental bodies in shaping future healthcare policies and initiatives [11]. ML can be utilized to analyze various datasets, such as Google Trends data. By examining the popularity and frequency of specific search terms over time, ML can provide insights into public interest and information needs during different phases of a pandemic. This information is crucial for informing the design and development of effective pandemic awareness systems [29].

Despite all the potential benefits of AI in health equity, it is crucial to understand the challenges associated with using ML in health equity. The lack of guidance, poor data, and ineffective execution are challenges that should be addressed when using ML in developing projects [21]. Another challenge is the potential bias within ML algorithms. If the training data is biased, the resulting predictions and interventions may also be biased, perpetuating existing health disparities. Therefore, it is important to continuously apply diverse and precise data and monitor ML models to decrease bias [19]. Ethical considerations regarding data privacy and ownership must also be addressed to ensure the ethical aspects of AI applications [21]. Moreover, previous studies indicate that one of the main obstacles that middle- and low-income countries struggle with is the design and development of AI toolsets and benchmarking [30].

This study makes several significant contributions to the field of health equity assessment using machine learning (ML) techniques, particularly focusing on social determinants of health (SDH). The key contributions of this study include the development of a comprehensive framework utilizing the CRISP-ML(Q) model to systematically assess health equity. This framework covers all phases of machine learning (ML) development, from project formulation to monitoring and maintenance, ensuring quality and stability throughout the process. Our study identified crucial social determinants of health (SDH) that significantly influence life expectancy. By employing various ML algorithms, including Linear Discriminant, Logistic Regression, and Bayesian Network for classification and CHAID, Regression, and Generalized Linear, we demonstrated the effectiveness of these models in predicting health outcomes with high accuracy. A pilot implementation of the proposed framework was conducted, focusing on life expectancy as a health outcome. The results from this pilot study validated the feasibility and effectiveness of using ML algorithms in health equity assessment, providing a practical example for future applications. We highlighted the importance of algorithmic fairness in ML applications for health equity, discussing the necessity of resampling or stratification to achieve fair predictions across diverse populations and addressing potential biases in the data and models. The findings of this study offer valuable insights for policymakers and healthcare providers. By integrating ML tools into health systems, countries can better prioritize and address health inequities, ultimately promoting a more equitable society. We provided strategic recommendations for developing the necessary infrastructure and regulatory frameworks to support the responsible use of AI in health equity assessments.

Conclusion

This study proposed a framework for assessing health equity using machine learning (ML) techniques, emphasizing the social determinants of health (SDH). The pilot study has demonstrated the potential of ML models, particularly the Linear Discriminant algorithm, in predicting life expectancy based on SDH. The Linear Discriminant and CHAID algorithms were selected as the best models due to their high accuracy and low error, strong performance in identifying key features, and good generalizability.

Agenda-setting to address health equity requires robust infrastructure and strategic prioritization. AI techniques and ML analysis offer promising solutions to address health inequities and reduce inequality, particularly in developing countries. The results suggest that AI systems can be designed to track equity and help the health system achieve its objectives. Overall, AI has the potential to analyze health equity and further global health goals.

The findings of this study underscore the necessity of developing the conditions for the effective use of ML tools. Therefore, comprehensive data sources, ensuring data quality, and establishing regulatory frameworks should be considered to support the responsible use of AI in health equity assessments. The presented pilot model of this study is a preliminary example of the applications of ML in addressing health inequities.

This theoretical insight calls for a paradigm shift in approaching health equity, promoting advanced technologies to inform and enhance public health strategies. The development and deployment of ML models should be prioritized as part of the agenda for reducing health inequities. This includes creating the necessary infrastructure and regulatory frameworks to support the effective use of AI in health equity assessments.

This study has several limitations that should be addressed in future research. Firstly, it was conducted within a specific period and utilized a limited number of data sources, which may not fully capture the variability and complexity of social determinants of health (SDH) over time. Future research should consider more comprehensive and diverse datasets, including longitudinal data, to provide more generalizable results. Secondly, although there are many machine learning (ML) models available, this study focused on a limited number of algorithms. Future studies should incorporate a broader range of ML models, including ensemble methods and deep learning techniques, to explore their potential in health equity assessment.

Additionally, the potential biases in the data and ML models were not extensively discussed. Biases in data collection and algorithmic processing can lead to skewed results and reinforce existing health inequities. Future research should deploy strategies to identify and mitigate these biases, such as using bias detection tools, applying fairness-aware ML algorithms, and ensuring diverse and representative datasets. As this study was designed as a pilot, the findings may not fully reflect real-world complexities. Future research should focus on developing and validating models in real-world settings, involving continuous monitoring and iterative improvements to ensure the accuracy and effectiveness of the models over time.

Ethical considerations, including data privacy, algorithmic transparency, and informed consent, are crucial for the responsible application of ML in health equity assessments. Future studies should establish clear ethical guidelines and frameworks to protect individuals' rights and ensure the ethical use of AI technologies. Finally, the implementation of ML models for health equity requires robust infrastructure and supportive policies. Future research should explore the development of necessary infrastructure, such as data-sharing platforms and regulatory frameworks, to facilitate the effective use of ML in health equity assessments.

Abbreviations

AI	Artificial intelligence
ML	Machine learning
WHO	World Health Organization
GDP	Gross Domestic Product

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13040-025-00428-x.

Supplementary Material 1

Acknowledgements

The Tehran University of Medical Sciences supported this research.

Author contributions

AT and MR conceived the study. AT supervised all evaluation phases and critically revised the manuscript. MM collaborated in all phases. MR and AB wrote the main manuscript text. AT, MM, SS, HR, HM, EM, and AB edited the manuscript. MM, MR, and MR collected and analyzed the data. AO, HR, MM, AAF, AB, AA, MR, and HM provided feedback on the result and manuscript. All authors reviewed the manuscript. AT is the guarantor.

Funding

Tehran University of Medical Sciences funded this research.

Data availability

The dataset presented in the study is available on request from the corresponding author during submission or after its publication.

Declarations

Ethics approval and consent to participate
This study received ethical approval from the Ethical Committee of the Tehran University of Medical Sciences (Approval ID: 56138); all methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication
Not applicable.

Adherence to national and international regulations
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Health Management, Policy and Economics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[2]Health Equity Research Centre (HERC), Tehran University of Medical Sciences, Tehran, Iran
[3]National Institute for Health Research, Tehran University of Medical Sciences, Tehran, Iran
[4]Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
[5]E-Health Department, Virtual School, Tehran University of Medical Science, Tehran, Iran
[6]School of Medicine, Tehran University of Medical Sciences, Tehran, Iran
[7]School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[8]Centre of Excellence for Global Health (CEGH), Department of Global Health and Public Policy, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Received: 29 July 2024 / Accepted: 28 January 2025
Published online: 10 February 2025

References
1. Schaekermann M et al. Health equity assessment of machine learning performance (HEAL): a framework and dermatology AI model case study. eClinicalMedicine 70:102479. https://doi.org/10.1016/j.eclinm.2024.102479
2. Dover DC, Belon AP. The health equity measurement framework: a comprehensive model to measure social inequities in health. Int J Equity Health. 2019;18:1–12.
3. Amos BK, Smirnov I. Determinants factors in predicting life expectancy using machine learning. Adv Eng Res (Rostov-on-Don). 2022;22(4):373–83.
4. Baqui P, et al. Comparing COVID-19 risk factors in Brazil using machine learning: the importance of socioeconomic, demographic and structural factors. Sci Rep. 2021;11(1):1–10.
5. Zhang Y, et al. Identifying urban built environment factors in pregnancy care and maternal mental health outcomes. BMC Pregnancy Childbirth. 2021;21(1):1–11.
6. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. J Global Health, 2018. 8(2).
7. Wahl B, et al. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ Global Health. 2018;3(4):e000798.
8. Murphy K, et al. Artificial intelligence for good health: a scoping review of the ethics literature. BMC Med Ethics. 2021;22(1):1–17.
9. El Naqa I, Murphy MJ. What is machine learning? Machine learning in radiation oncology. Springer; 2015. pp. 3–11.
10. Wiemken TL, Kelley RR. Machine learning in epidemiology and health outcomes research. Annu Rev Public Health. 2020;41(1):21–36.
11. Meshram SS. Comparative analysis of life expectancy between developed and developing countries using machine learning. In 2020 IEEE Bombay Section Signature Conference (IBSSC); 2020. IEEE.
12. Goulart CM, et al. Tools for measuring gender equality and women's empowerment (GEWE) indicators in humanitarian settings. Confl Health. 2021;15(1):1–16.
13. El Morr C, et al. A virtual community for disability advocacy: development of a Searchable Artificial intelligence–supported platform. JMIR Formative Res. 2021;5(11):e33335.
14. Thompson HM, et al. Bias and fairness assessment of a natural language processing opioid misuse classifier: detection and mitigation of electronic health record data disadvantages across racial subgroups. J Am Med Inform Assoc. 2021;28(11):2393–403.
15. Studer S, et al. Towards CRISP-ML (Q): a machine learning process model with quality assurance methodology. Mach Learn Knowl Extr. 2021;3(2):392–413.
16. Sunarti S, et al. Artificial intelligence in healthcare: opportunities and risk for future. Gac Sanit. 2021;35:S67–70.
17. Rattermann MJ, et al. Advancing health equity by addressing social determinants of health: using health data to improve educational outcomes. PLoS ONE. 2021;16(3):e0247909.
18. Reeves M, Bhat HS, Goldman-Mellor S. Resampling to address inequities in predictive modeling of suicide deaths. Volume 29. BMJ health & care informatics; 2022. 1.
19. Clark CR, et al. Health care equity in the use of advanced analytics and artificial intelligence technologies in primary care. J Gen Intern Med. 2021;36:3188–93.
20. Glymour MM, Manly JJ. Lifecourse social conditions and racial and ethnic patterns of cognitive aging. Neuropsychol Rev. 2008;18:223–54.
21. Intelligence NM. Striving for health equity with machine learning. Nat Mach Intell. 2021;3:653.

22. Kim JY, et al. Development and preliminary testing of Health Equity across the AI lifecycle (HEAAL): a framework for healthcare delivery organizations to mitigate the risk of AI solutions worsening health inequities. PLOS Digit Health. 2024;3(5):e0000390.
23. Murphy A et al. Bridging Health disparities in the Data-Driven World of Artificial Intelligence: a narrative review. J Racial Ethnic Health Disparities, 2024, pp. 1–13. https://doi.org/10.1007/s40615-024-02057-2
24. Thomasian NM, Eickhoff C, Adashi EY. Advancing health equity with artificial intelligence. J Public Health Policy. 2021;42(4):602.
25. Tumin D. Health Equity insights from Machine Learning models. J Gen Intern Med. 2021;36(8):2475–2475.
26. Ng MY, et al. Scaling equitable artificial intelligence in healthcare with machine learning operations. BMJ Health Care Inf. 2024;31(1):e101101.
27. Raza S. Connecting fairness in machine learning with public health equity. In 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI); 2023. IEEE.
28. Feng C, Jiao J. Predicting and mapping neighborhood-scale health outcomes: a machine learning approach. Comput Environ Urban Syst. 2021;85:101562.
29. Shakeri E, Far BH. Exploring the requirements of pandemic awareness systems: A case study of covid-19 using social media data. In Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering; 2020.
30. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. Lancet. 2020;395(10236):1579–86.

Publisher's note