User-Centric Democratization towards Social Value Aligned Medical AI Services

Zhaonian Zhang , **Richard Jiang***
LIRA Center, Lancaster University, Lancaster LA1 4YW, England
z.zhang47@lancaster.ac.uk, r.jiang2@lancaster.ac.uk

Abstract

Democratic AI, aiming at developing AI systems aligned with human values, holds promise for making AI services accessible to people. However, concerns have been raised regarding the participation of non-technical individuals, potentially undermining the carefully designed values of AI systems by experts. In this paper, we investigate Democratic AI, define it mathematically, and propose a user-centric evolutionary democratic AI (u-DemAI) framework. This framework maximizes the social values of cloud-based AI services by incorporating user feedback and emulating human behavior in a community via a user-in-the-loop iteration. We apply our framework to a medical AI service for brain age estimation and demonstrate that non-expert users can consistently contribute to improving AI systems through a natural democratic process. The u-DemAI framework presents a mathematical interpretation of Democracy for AI, conceptualizing it as a natural computing process. Our experiments successfully show that involving non-tech individuals can help improve performance and simultaneously mitigate bias in AI models developed by AI experts, showcasing the potential for Democratic AI to benefit end users and regain control over AI services that shape various aspects of our lives, including our health.

1 Introduction

The primary goal of AI is to develop technologies that improve people's quality of life and address significant societal challenges [Hodges, 2006; Taddeo and Floridi, 2018]. Its origins lie in simulating human intelligence [Hodges, 2006]. Over the past five years, AI has been increasingly applied in various areas, such as discovering new drugs and vaccines [Jumper *et al.*, 2021], addressing environmental issues [Gomes *et al.*, 2019], predicting humanitarian crises [Tomašev *et al.*, 2019], and influencing policymaking [Lee *et al.*, 2019]. With the growing prominence of AI, ethical considerations have gained greater attention [Conitzer *et al.*, 2017], as the public seeks increased decision-making autonomy in matters relevant to their daily lives [Olson Jr, 1971].

To date, AI technology has primarily been developed by a handful of technology giants, including Microsoft, Google, and Facebook. [Shen, 2017] highlighted that approximately 10,000 individuals in only 7 countries were responsible for coding all AI worldwide. This concentrated control poses a significant challenge as it hampers the progress and potential of AI applications. In the worst case, AI systems can exhibit biases and lack generalizability to the broader population, hindering the realization of their intended societal benefits.

Toward fulfilling the social values of AI, Democratic AI (DemAI) has recently gained traction as a process to guarantee that the public can be well-informed in the decision-making process. By aligning AI processes with modern democratic values, the field can contribute to "Social Good" instead of relying solely on professional codes of ethics or legislative restrictions on technologies and industries.

Democratization of AI is the idea of providing everyone with equal access to resources, opportunities and benefits related to AI [Strouse *et al.*, 2021]. This concept of Democratic AI has yet to be fully defined [Garvey, 2018]. However, it can be understood to mean the ability for everyone to participate in the process of building AI, with full freedom of choice, regardless of their knowledge of AI.

Democratizing AI offers advantages in reducing AI monopolies and oligopolies [Ahmed *et al.*, 2020]. It lowers entry barriers, enabling individuals with no prior AI experience to access AI. Open sharing of data and algorithms, along with cloud infrastructures, allows for broader AI access regardless of financial resources. Increased availability of open datasets and algorithms enhances solutions to complex AI problems while reducing costs. Open-source frameworks like PyTorch and TensorFlow contribute significantly to deep learning progress and talent development, expediting AI skill growth. These factors collectively accelerate AI advancement, enhance accessibility, and generate societal value.

Despite its potential benefits, democratic AI faces significant challenges. Concerns have been raised about the involvement of non-technical individuals, which can undermine the carefully crafted values of AI systems established

*Correspondence Author

by experts [Rao, 2020]. Bias represents a major risk in democratic AI systems, even in those developed by highly qualified engineers [Zou and Schiebinger, 2018]. Inexperienced contributors to an AI system can generate unfair results, leading to misjudgment and serious consequences. Identifying the origin of bias is challenging, and rectifying it can be costly for companies, with uncertain success. Furthermore, the lack of a clear definition or mathematical description of democratic AI leaves it in a realm of imagination. Despite its perceived merits, the development of democratic AI has been hindered by these issues, with no widely accepted model to address the aforementioned challenges.

In this paper, we aim to address the challenges associated with Democratic AI by exploring its potential value in addressing societal and technological issues. We propose a user-centric Democratic AI (u-DemAI) framework that maximizes benefits for individuals through a natural user-in-the-loop iteration process. To demonstrate the efficacy of our framework, we conduct a case study focusing on brain age estimation in medical services. Our experimental results demonstrate the significant benefits of our u-DemAI framework, which involves users in the decision-making process and enables freedom of choice within a community for AI services. Importantly, the success of our experiments serves as a practical proof-of-concept implementation of Democratic AI, highlighting its advantages over conventional expert-ruled AI services by engaging non-experts.

2 Preliminary

2.1 What is Democratic AI

In politics, democracy is a system wherein the people have the power and liberty to decide on the governing system. With the advent of Artificial Intelligence (AI), new topics have emerged in the social sciences, such as AI governance and the pros and cons of introducing AI into democracy. Conversely, when democracy is incorporated into AI, a new branch of AI known as Democratic AI emerges. This is also referred to as AI democratization, which is a process of bringing democracy to AI. When it comes to AI services, users are expected to gain more authority in the governance of AI. To gain a better understanding of AI democratization, it is essential to explore some basic concepts that can serve as fundamentals for constructing our theoretical framework.

As the need for democratic AI grows, some researchers have attempted to describe their ideal ones. [Nguyen *et al.*, 2022] think democratic AI is a machine learning system underlying principle for building large-scale distributed. It relies on the hierarchical self-organization of well-connected distributed learning agents who have limited and highly personalized data and can adjust themselves based on the underlying duality of specialized and generalized processes.

[Shashi *et al.*, 2022] think federated analytics (FA) is a good basic structure of democratic AI, but the current realization of federated learning (FL) adopts single server-multiple client architecture with limited scope for FA, which often results in learning models with poor generalization, so they built a democratic AI based on FL, which can empower generalization capability of models over clouds.

As AI is currently dominated by an oligopoly of centralized mega-corporations, that focus on the interests of their stakeholders, [Montes and Goertzel, 2019] want to change that situation and propose a democratic AI as a distributed, decentralized, and democratized market for AI services to run on distributed ledger technology.

In summary of these existing works, we can come to give a generalized definition of democratic AI as below,
"Democratic AI means an AI implementation that involves relevant people in the optimization of AI services that promotes the social values to benefit the communities of people."

Inevitably, people need to be included in the loop of the optimization toward social values, accessible to everyone concerned with full rights to select the best AI service, according to Montes's expectations of AI [Montes and Goertzel, 2019], while it breaks the monopoly of big companies on AI. Putting users' requests in the loop of AI service optimization is an inevitable solution to improve the generalization and personalization abilities of AI services, which matches the expectation of [Nguyen *et al.*, 2022]. Based on the above definitions, we will build a math description of democratic AI called u-DemAI in the below sections and examine it in our case study.

2.2 Overview of Our u-DemAI Framework

The u-DemAI framework proposed in this paper is a comprehensive, multi-function, public system that offers various optimization strategies, models and datasets to its users. It encourages people to share their trained models based on their local data, where AI services are made available to users via clouds. For the uploaded models, the system tests and records their performances, creating a list for users to choose from.

When users want to make a prediction, u-DemAI employs multiple cloud-based AI services and engages users in the optimization loop. Users can select their desired models, preferences, and optimization measures. The models are the uploaded trained models, with their performance list as a reference. Target preferences include various indicators such as fairness over gender, age groups, race, etc. The optimization measures may include options for the highest accuracy, the fairest results, or a balanced prediction. The prediction is based on the results from several single cloud services, leveraging the concept of evolutionary optimization to find the best weights for each cloud service. The weights of single services can be adjusted according to the users' feedbacks of different communities, such as those of different age groups, towards optimizing social values such as the fairness of prediction. In summary, our u-DemAI enables open AI Services by providing a democratic process through which users can select the desired models. Figure 1 illustrates the architecture of our proposed u-DemAI framework.

2.3 Case Study: Medical Brain Age Estimation

In this paper, we focus on a specific medical task as our case study, given its relevance to democratic AI and potential benefits to patients. We select brain age estimation, a recent medical challenge, as our case study, which requires both high predictive accuracy and age-wise fairness [Cole *et al.*, 2017; Luders *et al.*, 2016].

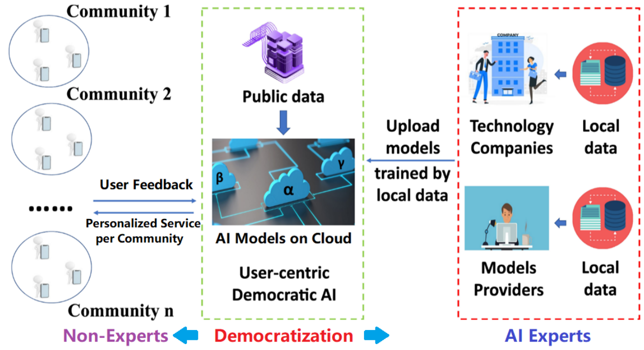


Figure 1. The community-adaptive democratic process of the proposed u-DemAI framework over cloud-based AI services.

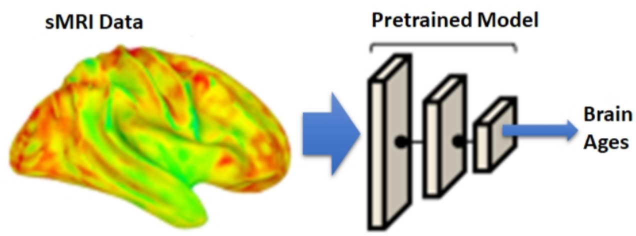


Figure 2. Case study: Brain age estimation using AI models.

The human brain can be modeled as a Turing's Type-B machine with randomly interconnected neurons [Jiang and Crookes, 2019], and it is closely linked to mental health [Jiang *et al.*, 2022]. Brain age estimation is a common classification or regression task that provides a health indicator of the human brain [Peng *et al.*, 2021]. Previous studies have successfully employed various AI models to predict individuals' brain age using neuroimaging data, demonstrating strong performance [Cole *et al.*, 2017; Luders *et al.*, 2016]. By training deep learning models on healthy samples, we can obtain AI models capable of estimating the biological age of a healthy person based on brain neuroimaging data [Cole *et al.*, 2017]. The process is illustrated in Figure 2.

3 Modelling Democratic AI

3.1 Democratic AI beyond Clouds

Unlike its literal interpretation, democratic AI is deeply intertwined with cloud computing technology. The reason is that when democratic AI aims to make AI accessible to the general public, these AI services necessitate internet access from any location, making cloud computing an indispensable component [Nguyen *et al.*, 2022; Shashi *et al.*, 2022]. Leading AI companies like Microsoft, Google, and IBM have all expressed their intentions to democratize AI, with cloud computing occupying a central position in this endeavor.

These prominent AI providers offer AI services accessible through the cloud for a wide range of users, including edge users, smart home users, hospitals, and industries. However, a crucial question remains: How can we select the most suitable AI service? This is precisely where our u-DemAI algorithm comes into play. In this paper, we present our new algorithm as a means to involve users and maximize the benefits and social values derived from AI services.

3.2 Community Based User-Centric DemAI

Standing on the side of users, every user in a community of an AI service would like to evaluate the AI service using their own experience, rather than fully depending on the told specifications by the service provider. Considering we have a set of different measures fm_kg using user preference, to simplify the measures and enable an easier choice by users, we can define a combined measure F over k measures,

$$F = \sum_k \ {}_k m_k \tag{1}$$

Here ${}_k$ are the preference degrees set up by users. For example, if users care more about accuracy, the ${}_k$ for accuracy will be set bigger than others; when users in another community come to care more about fairness, the ${}_k$ for fairness will be set bigger. Consequently, based on F, users can easily make the choice over a Service i. In this case study, we set m_1 as the prediction accuracy, and m_2 as the Pearson coefficient (PC) to indicate the degree of ageism of predicted results. In our case study, we assume all users have ${}_1$= ${}_2$=0.5 to balance accuracy and fairness.

Considering a user can access various AI services, these AI services can then be combined together by their weights, that impact the performance F,

$$c_D^j = \sum_i w_i c_i^j \tag{2}$$

Here, c_i^j is the prediction on the data of the j-th user from the i-th service, and w_i is the weight for that service. c_D^j denotes the combined prediction via the democratic process.

It is noted that the process in Eq.(2) is influenced by users since fw_ig are defined by users' personalized requests and c_D^j is the combined outcome from user feedback. In the optimization process of our user-centric DemAI framework, users are included in the loop, to choose their preference towards social values by setting up $f \ {}_k g$. The democratic process can then be described as an optimization process via users' choice on various AI services,

$$w_i = \arg\max F(w_i; \ {}_k; c_i^j) \tag{3}$$

Consequently, users are the final decision-maker with full control on how to use AI services over the clouds since the selection of single AI services, the weights of each service and the optimization outcome from the DemAI are all dependent on users' requests.

The philosophy behind this is that a user may always try to give those good services more weights. In a broader sense, such a democratic process is a natural computing process, similar to bird flocking, DNA computing, swarm intelligence, artificial life, etc. It is a natural consequence of individual behavior or actions that are included in the loop of the AI service optimization for a community of users. Inevitably, we will face two fundamental math challenges on democratic AI,

1. How to include users in the loop of the democratic process?

2. How to guarantee the democratic process converges on the best optimum?

To answer the above two fundamental questions of democratic AI, we will propose a natural democratic computing process for our u-DemAI framework.

3.3 Evolutionary Democratic Process

In principle, Democratic AI can be considered a natural computing process via human behavior [Koster *et al.*, 2022] from a community of users. In order to find the best way to integrate AI services over the cloud using people's feedback in the loop of optimization towards social values, we borrow the concept of particle swarm optimization (PSO) to emulate an evolutionary democratic model.

Here select *n* initial particles to appear at random positions with random initial velocities. Each initial particle is an n-dimension weights data of single services as $(\ _1, \ _2, \ldots, \ _n)$. Now we record the current position of the *i*-th particle as $x_i = (\ _{i1}, \ _{i2}, \ldots, \ _{in})$, the current velocity of the particle as v_i, the current best position of the particle as y_i, and the number of iterations is t, and the symbol g will be used to denote the loss function. We can update the next best position via the equation below (in the loop of people's feedback from the community of AI services),

$$y_i(t+1) = \begin{cases} y_i(t); & g(x_i(t+1)) \geq g(y_i(t)) \\ x_i(t+1); & g(x_i(t+1)) < g(y_i(t)) \end{cases} \quad (4)$$

We define $\hat{y}$ as the global optimal position among all particles in Eq.(5),

$$\begin{aligned} \hat{y}(t) &\in \{y_0(t), y_1(t), \ldots, y_s(t)\} \mid g(\hat{y}(t)) \\ &= \min\{g(y_0(t), g(y_1(t), \ldots, g(y_s(t)\} \end{aligned} \quad (5)$$

The step of velocity update is specified separately for each dimension *j* so that V_{ij} denotes the *j*-th dimension of the velocity vector associated with the *i*-th particle. And the velocity update equation is shown below:

$$\begin{aligned} V_{ij}(t+1) = \omega V_{ij}(t) &+ \varphi_1 rand(0,1)(y_{ij}(t) - x_{ij}(t)) \\ &+ \varphi_2 rand(0,1)(\hat{y}_j(t) - x_{ij}(t)) \end{aligned} \quad (6)$$

Among them, ω, φ_1 and φ_2 are constants between (0, 1). The update of the current position of *i*-th particle is,

$$x_i(t+1) = x_i(t) + V_i(t+1) \quad (7)$$

We assume x(t) is the position of a particle at time *t*, $a = \varphi_2 / (\varphi_1 + \varphi_2)$, according to the proof from [Van Den Bergh and others, 2007], we can get equation (8):

$$\lim_{t \to +\infty} x_t = (1 - a)y + a\hat{y} \quad (8)$$

Here x(t) is the position of a particle at time t. Eq.(8) implies that the particles converge to a value derived from the line connecting the personal best to the global best.

In the subsequent experiments of our case study, we set $\omega = 0.2$, $\varphi_1 = \varphi_2 = 0.2$, the number of iterations is 100, and the number of initial particles is 300.

The mathematical interpretation of the evolutionary democratic process provides rigorous proof of the convergence of Democratic AI toward optimal social values. In parallel, the democratic process represents a natural computing phenomenon intertwined with the collective behavior of individuals within a community.

3.4 AI Services for Brain Age Estimation

Brain age estimation has become a key measure for healthy aging [Cole and Franke, 2017]. In our case study, we assume we use four online AI services over cloud, and each offers their brain age estimation using an AI model that has remarkable performance in estimating brain ages,

1. Convolutional Neural Networks [Zhang *et al.*, 2022]: The CNN model was implemented using Keras & Tensorflow.
2. GoogLeNet (Inception V1) [Couvy-Duchesne *et al.*, 2020]: This model is based on Inception V1. Compared to Google's Inception V1, it changes the softmax layer to a fully connected layer as the final layer so that this task becomes a regression task instead of a classification task.
3. ResNet [Peng *et al.*, 2021]: This model is based on ResNet that consists of 5 residual blocks, each followed by a max pooling layer of kernel size 3×3×3 and stride 2×2×2, and one fully connected block.
4. SVR [Cole *et al.*, 2017]: This model is based on SVM but for the regression purpose for brain age estimation.

By accessing the above cloud AI services, we will test our u-DemAI framework with different communities of users and see if it can automatically optimize for various groups of people in different conditions.

Here, we consider three different communities of users: young group (16–30 years old), middle age group (31–60) and elder group (61–100). We will test our u-DemAI framework with each community and see if it will achieve better performance for each different group.

3.5 Datasets

The dataset we used here is based on [Cole *et al.*, 2017], including 2641 healthy individuals' brain sMRI with extra information such as their age and gender. The sample age ranged from 16 to 90 years old, the average age of samples is 35.8 years old, and the standard deviation of age is 16.2 years. Of the participants, 53% are females and 47% males. The sample details can be found in [Cole *et al.*, 2017].

In our project, we use two different kinds of data as input for the services. One is Gray Matter and White Matter Maps, the other is Surface-Based Processing of Gray Matter. The Gray Matter and White Matter Maps were distributed by the PAC organization. This kind of data is used for the input of self-defined CNN, ResNet and GoogLenet in our project. As for Surface-Based Processing of Gray Matter, we extract the vertex-wise measurements of cortical thickness and surface area based on the sMRIs by using FreeSurfer 6.0 [Fischl, 2012]. This kind of data is used for the input of SVR.

4 Experimental Results

4.1 Experimental Setup

The dataset we use here includes 2641 healthy individuals' brain sMRIs and information from samples such as their age and gender. We use 75% of them to train the services and 25% to test the performance of services.

In this work, Mean Absolute Error (MAE) between the sample's chronological age and predicted age represents the estimation performance of services, which is frequently used in brain age prediction papers [Couvy-Duchesne *et al.*, 2020; Peng *et al.*, 2021; Cole *et al.*, 2017], with the least MAE representing the highest accuracy in age prediction.

We evaluate the fairness of services and judge the degree of ageism in predicted results by 3 criteria. One is the Pearson coefficient between the brain age gap (chronological age minus predicted age) and chronological age, low coefficient denotes that the performance of service is not influenced by the true age significantly, which means the service is good at resisting ageism. The equation is shown below:

$$\rho_{X;Y} = cov(X;Y)/\sigma_X \sigma_Y \tag{9}$$

where *X* is brain age gap, *Y* is true age, cov means the covariance, σ_X is the standard deviation of *X*, and σ_Y is the standard deviation of *Y*.

The slope rate of the brain age gap with increasing chronological age [Couvy-Duchesne *et al.*, 2020] is a good evaluation criterion. The low rate represents that the predicted age and the performance of the service are less influenced by the chronological age, which means the service is a fair AI with less of ageism. We can know their equation is shown below:

$$\lim_{\Delta a \to 0} \left| \frac{G(a + \Delta a) - G(a)}{\Delta a} \right| \tag{10}$$

Here *a* represents age, *G* means brain age gap. We analyze them from the slope of the brain age gap-chronological age line.

We also use the standard deviation of absolute error between chronological age and predicted age to evaluate the degree of ageism in predicted results, lower standard deviation represents a stronger ability against ageism and unfairness. The equation is shown below:

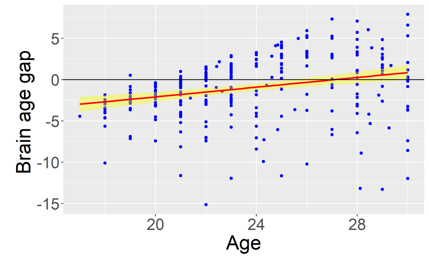
$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1}} \tag{11}$$

where *n* is the number of samples, x_i is the absolute value of the brain age gap of *i*-th sample, $\bar{x}$ is the average of the absolute value of brain age gap of all samples.

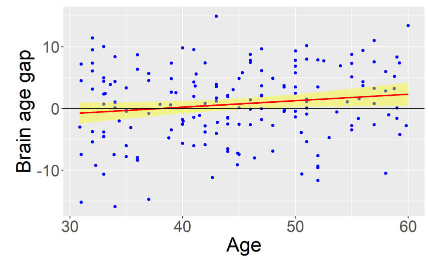
4.2 Ageism in Single Models

We test four popular single services, including self-designed CNN [Zhang *et al.*, 2022], ResNet [Peng *et al.*, 2021], GoogLeNet [Couvy-Duchesne *et al.*, 2020] and SVR [Cole *et al.*, 2017] to check if ageism could be method-dependent, it also shows the degree of fairness.

Figure 3 represents GoogLeNet's brain age gap plotted as a function of the chronological age in different age groups. This figure shows that ageism indeed exists in GoogLeNet's brain age estimation because the slope of the brain age gap-chronological age line varies at different age groups, the predicted age of the young is generally higher than their actual age, while the predicted age of the elderly will be underestimated. In 17-30 age group, the red fit line is almost under the black horizontal line, which implies young people's predicted age is always overestimated. In 30-60 age group, the red line through the black line, the brain age gap begins to change from negative to positive. In 60-90 age group, the red line stays beyond the black line, implying old people's predicted age is always underestimated.



(a) Young



(b) Middle Aged

(c) Elder

Figure 3. Unfairness and ageism in brain age estimation for single service. Here, age ranges for Young, Middle and Elder groups are: 16-30, 31-60 and 61-100. It shows the brain age gap as a function of the chronological age using the GoogLeNet model in different groups. The best fit of line regression (red) in each plot with the 95% prediction interval (yellow area) denotes the degree of bias.

Table 1 shows the details of 4 single services' slopes of fitted lines for age gaps in different age groups. It says 4 services have different degrees of ageism in their predicted results and the fairest algorithm is always different between different age groups, in other words, some services are suitable for predicting young samples', and some services are suitable for elder samples. GoogLeNet is the fairest service for young people, so as CNN for middle-aged people and ResNet for old people. Besides, SVM always has the worst fairness compared to other methods.

4.3 Evaluation on Democratic Process

In this part, we will explore the performance of our novel u-DemAI in age estimation towards alleviating ageism. The estimation results of u-DemAI evolved from the results of

Services/ Indicators	Young	Middle	Elder
Single service 1: GoogLeNet	0.29	0.10	0.20
Single service 2: ResNet	0.39	-0.08	0.16
Single service 3: CNN	0.30	0.01	0.19
Single service 4: SVM	0.42	0.21	0.20

Table 1. The details of single services' slopes of fitted lines for age gap in different age groups.

single services using CNN, ResNet, GoogLeNet and SVR, respectively. In order to deal with ageism in brain age estimation, we set three age groups (young, middle, and elder) as optional personality requests. For each age group, we build a separate u-DemAI service upon all 4 single model services. When we predict the brain age, the u-DemAI learned from each user group makes predictions using the appropriate weights that are optimized in the user-inclusive loops per group. Its prediction results fully consider the brain characteristics of different age groups by adjusting the weights of single services in it.

The variation of individual services' weights and the optimization target function in the u-DemAI models at different age groups during training are shown in Figure 4. The blue lines, brown lines, green lines and purple lines represent the weights of GoogLeNet, ResNet, self-defined CNN, and SVM in the u-DemAI models for different age groups respectively, the red line shows the variation of the loss function in the u-DemAI model per age group. As shown in Figure 4, for the young group, both weights of individual services and the loss function converge to a definite value roughly in 45 iterations. As for the middle group, they both converge in about 36 iterations, which is similar to the elder group.

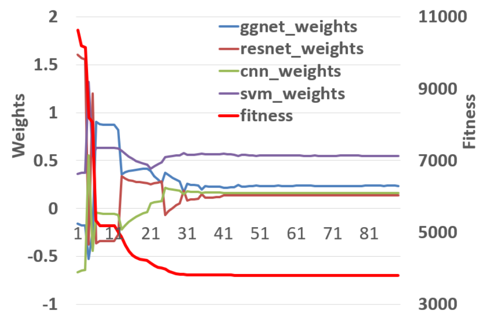
Figure 5 represents the brain age gap of five different services as functions of the chronological age. The brain age gap plot shows u-DemAI is the best service to deal with ageism during the brain estimation, because its brain age gap varies minimally with aging, in contrast, SVM shows the biggest ageism compare to other methods and the fairest single service is self-designed CNN. Single services always underestimate old people's age and overestimate young, but u-DemAI always predicts higher than chronological age.

The details of the test results are shown in Table 2. We use three criteria to judge fairness and ageism in predicted results. The slopes of fitted lines for the age gap show the fairest service is u-DemAI, followed by self-designed CNN. The results of ResNet and GoogLeNet are similar with bad fairness, and SVM has the biggest ageism in its predicted results. In terms of the standard deviation of absolute error, u-DemAI has the lowest value at 2.67. In terms of the Pearson coefficient between the brain age gap and true age, our u-DemAI achieves the best fairness, amazingly with a Pearson coefficient as low as 0.01, and SVM is the worst at 0.55.

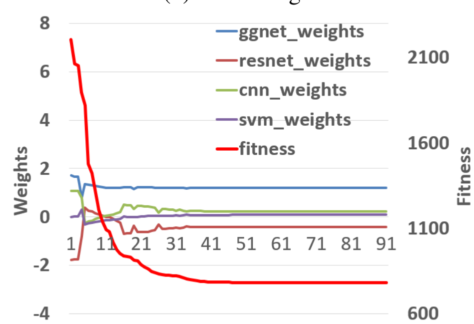
In terms of estimation accuracy, u-DemAI achieved the best prediction accuracy, whose MAE is 2.99. GoogLeNet's performance is the best among single services with 3.7, and SVM is still the worst among them with 5.19.

We also compare u-DemAI with the expert-guided model combination via ensemble methods based on the same dataset

(a) Young



(b) Middle Aged



(c) Elder

Figure 4. The user-in-the-loop iterative process in the u-DemAI framework for different age communities. We can see both the service weights and the cost function converges successively along with user feedback.

from PAC 2019. [Couvy-Duchesne *et al.*, 2020] build an ensemble model combined by seven different algorithms, the Pearson coefficient of its predicted results is 0.21, and MAE is 3.33. [Da Costa *et al.*, 2020] used shallow machine learning methods to build an ensemble model with an MAE of 3.76, and [Zhang *et al.*, 2022] donate a nonlinear age-adaptive ensemble learning whose MAE is 3.19. Our u-DemAI which combines services together via non-expert users shows the best performance on accuracy and fairness of all the above expert-designed ensemble methods.

Overall, u-DemAI demonstrated exceptional results across all measures when incorporating a user-in-the-loop process involving non-expert users. These findings underscore the immense potential of democratic AI in providing social values for individuals and communities.

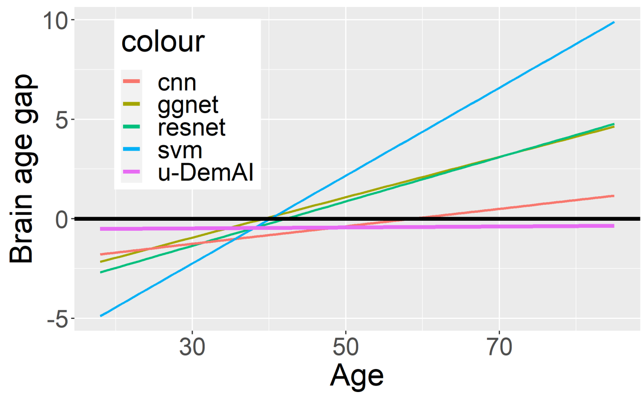


Figure 5. Test performance of AI models/services. It represents the brain age gap in 5 different services as functions of the whole chronological age. Here, ggnet means GoogLeNet, resnet means ResNet, cnn means a self-defined CNN, and svm refers to SVM. We can see our u-DemAI reduced the bias apparently.

Services	Slopes	SDAE	PC	MAE
Single service 1: GoogLeNet	1.13	3.09	0.34	3.70
Single service 2: ResNet	0.96	3.87	0.33	3.92
Single service 3: CNN	0.38	3.49	0.13	4.34
Single service 4: SVM	1.39	3.86	0.55	5.19
Ensemble Service A [Couvy-Duchesne *et al.*, 2020]			0.21	3.33
Ensemble Service B [Da Costa *et al.*, 2020]				3.76
Ensemble Service C [Zhang *et al.*, 2022]				3.19
Our u-DemAI	**0.04**	**2.67**	**0.01**	**2.99**

Table 2. The details of tested services' performance. Here, the 1st column is the slopes of fitted lines for the age gap, the 2nd column is the standard deviation of absolute error (SDAE), the 3rd column is the Pearson coefficient (PC) between the brain age gap and true age, and the last column is MAE. Our u-DemAI has consistently achieved the best among all measures.

5 Discussion

AI plays a crucial role in driving technological transformation, and concerns regarding its fairness have escalated in recent years [Posner *et al.*, 2020]. The discussion primarily revolves around two key aspects: the diversity of AI users and the presence of biased outcomes in AI predictions. [Posner *et al.*, 2020] contend that the AI field is currently facing a diversity crisis, emphasizing the significance of this juncture.

To address this, the concept of democratic AI can be a good solution to lower the entry barriers for people to use AI. However, democratic AI has several drawbacks, such as the lack of a clear definition and mathematical description. Also, democratizing AI can make the problem of bias in prediction results more obvious. Further, it is also worried that Inexperienced AI developers can lead to results that are far from the truth, which can have serious consequences.

Our work clarified the democratic AI from these worries in several aspects. First of all, we present a clear definition of democratic AI that engages people in its optimization process toward social values. Secondly, we establish a mathematical description of the democratic process, where DemAI can be interpreted as a natural computing process based on the human behavior of various communities. Furthermore, we leveraged the evolutionary algorithm for DemAI and proposed our u-DemAI framework with guaranteed convergence in its optimization process. Finally, through our case study, we demonstrated that by including non-expert users in the optimization loop, democratic AI can help achieve the best performance (in terms of accuracy and fairness) beyond both single and combined AI services designed by experts.

It is worth adding that bringing democracy into AI is a cross-disciplinary task across social sciences, math and computer science. However, in its implementation, inevitably democratic AI will involve AI services over the cloud, while users and people are involved via cloud-based platforms. Therefore, the DemAI topic is deeply rooted in cloud technology, where its security needs to be studied in our future work.

Our u-DemAI approach has several novel features. We provide a wide-covered definition of democratic AI that emphasizes its social values and formulate a mathematical interpretation of the democratic AI process as a natural computational process upon human behavior. We also present mathematical proof of the convergence of such a natural democratic process towards the optimization of social values by involving non-expert people in the loop. In the end, we demonstrated our novel u-DemAI method in the case study of a medical service, and successfully validated the great value of democratic AI as a promising future trend to implement AI for the benefit of people.

6 Conclusion

In conclusion, for the first time, we developed a novel democratic AI system, termed User-Centric Evolutionary Democratic AI (u-DemAI), which is mathematically explainable and can be customized to different groups of people. We applied u-DemAI on a medical service task of brain age estimation and demonstrated that the proposed DemAI framework can safely converge towards better social values for users and communities by engaging non-AI-expert people in the optimization loops, showing better performance beyond expert-designed single model services or combined services, in term of both accuracy and fairness in the brain age estimation. Our contribution thus provides a useful example to demonstrate the merits of AI democratization for both theoretical development and practical demonstration.

Ethical Statement

There are no ethical issues.

Acknowledgments

This work was supported in part by the U.K. EPSRC under Grant EP/P009727/1 and Grant EP/T518037/1, and in part by the Leverhulme Trust under Grant RF-2019-492.

References

[Ahmed *et al.*, 2020] Shakkeel Ahmed, Ravi S Mula, and Soma S Dhavala. A framework for democratizing ai. *arXiv preprint arXiv:2001.00818*, 2020.

[Cole and Franke, 2017] James H Cole and Katja Franke. Predicting age using neuroimaging: innovative brain ageing biomarkers. *Trends in neurosciences*, 40(12):681–690, 2017.

[Cole *et al.*, 2017] James H Cole, Rudra PK Poudel, Dimosthenis Tsagkrasoulis, Matthan WA Caan, Claire Steves, Tim D Spector, and Giovanni Montana. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163:115–124, 2017.

[Conitzer *et al.*, 2017] Vincent Conitzer, Walter Sinnott-Armstrong, Jana Schaich Borg, Yuan Deng, and Max Kramer. Moral decision making frameworks for artificial intelligence. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

[Couvy-Duchesne *et al.*, 2020] Baptiste Couvy-Duchesne, Johann Faouzi, Benoît Martin, Elina Thibeau-Sutre, Adam Wild, Manon Ansart, Stanley Durrleman, Didier Dormont, Ninon Burgos, and Olivier Colliot. Ensemble learning of convolutional neural network, support vector machine, and best linear unbiased predictor for brain age prediction: Aramis contribution to the predictive analytics competition 2019 challenge. *Frontiers in Psychiatry*, 11:593336, 2020.

[Da Costa *et al.*, 2020] Pedro F Da Costa, Jessica Dafflon, and Walter HL Pinaya. Brain-age prediction using shallow machine learning: predictive analytics competition 2019. *Frontiers in psychiatry*, 11:604478, 2020.

[Fischl, 2012] Bruce Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.

[Garvey, 2018] Colin Garvey. A framework for evaluating barriers to the democratization of artificial intelligence. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[Gomes *et al.*, 2019] Carla Gomes, Thomas Dietterich, Christopher Barrett, Jon Conrad, Bistra Dilkina, Stefano Ermon, Fei Fang, Andrew Farnsworth, Alan Fern, Xiaoli Fern, et al. Computational sustainability: Computing for a better world and a sustainable future. *Communications of the ACM*, 62(9):56–65, 2019.

[Hodges, 2006] Andrew Hodges. B. jack copeland (ed.), the essential turing: The ideas that gave birth to the computer age. oxford: Clarendon press, 2004. pp. viii+ 613. isbn 0-19-825079-7.£ 50.00 (hardback). isbn 0-19-825080-0.£ 14.99 (paperback). *The British Journal for the History of Science*, 39(3):470–471, 2006.

[Jiang and Crookes, 2019] Richard Jiang and Danny Crookes. Shallow unorganized neural networks using smart neuron model for visual perception. *IEEE Access*, 7:152701–152714, 2019.

[Jiang *et al.*, 2022] Richard Jiang, Paul Chazot, Nicola Pavese, Danny Crookes, Ahmed Bouridane, and M Emre Celebi. Private facial prediagnosis as an edge service for parkinson's dbs treatment valuation. *IEEE Journal of Biomedical and Health Informatics*, 26(6):2703–2713, 2022.

[Jumper *et al.*, 2021] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[Koster *et al.*, 2022] Raphael Koster, Jan Balaguer, Andrea Tacchetti, Ari Weinstein, Tina Zhu, Oliver Hauser, Duncan Williams, Lucy Campbell-Gillingham, Phoebe Thacker, Matthew Botvinick, et al. Human-centred mechanism design with democratic ai. *Nature Human Behaviour*, 6(10):1398–1407, 2022.

[Lee *et al.*, 2019] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, et al. Webuildai: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–35, 2019.

[Luders *et al.*, 2016] Eileen Luders, Nicolas Cherbuin, and Christian Gaser. Estimating brain age using high-resolution pattern recognition: Younger brains in long-term meditation practitioners. *Neuroimage*, 134:508–513, 2016.

[Montes and Goertzel, 2019] Gabriel Axel Montes and Ben Goertzel. Distributed, decentralized, and democratized artificial intelligence. *Technological Forecasting and Social Change*, 141:354–358, 2019.

[Nguyen *et al.*, 2022] Minh NH Nguyen, Shashi Raj Pandey, Tri Nguyen Dang, Eui-Nam Huh, Nguyen H Tran, Walid Saad, and Choong Seon Hong. Self-organizing democratized learning: Toward large-scale distributed learning systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[Olson Jr, 1971] Mancur Olson Jr. *The Logic of Collective Action: Public Goods and the Theory of Groups, with a new preface and appendix*, volume 124. Harvard University Press, 1971.

[Peng *et al.*, 2021] Han Peng, Weikang Gong, Christian F Beckmann, Andrea Vedaldi, and Stephen M Smith. Accurate brain age prediction with lightweight deep neural networks. *Medical image analysis*, 68:101871, 2021.

[Posner *et al.*, 2020] Tess Posner, Li Fei-Fei, et al. Ai will change the world, so it's time to change ai. *Nature*, 588(7837):S118–S118, 2020.

[Rao, 2020] Anand S. Rao. Democratization of ai: A double-edged sword. August 16 2020.

[Shashi *et al.*, 2022] Raj Pandey Shashi, Huu Nhat Minh Nguyen, Dang Tri Nguyen, Thar Kyi, H Tran Nguyen, Zhu Han, and Choong Seon Hong. Edge-assisted democratized learning towards federated analytics. 2022.

[Shen, 2017] Lucinda Shen. Former u.s. cto: The 'robot apocalypse' could happen. here's how you stop it. November 15 2017.

[Strouse *et al.*, 2021] DJ Strouse, Kevin McKee, Matt Botvinick, Edward Hughes, and Richard Everett. Collaborating with humans without human data. *Advances in Neural Information Processing Systems*, 34:14502–14515, 2021.

[Taddeo and Floridi, 2018] Mariarosaria Taddeo and Luciano Floridi. How ai can be a force for good. *Science*, 361(6404):751–752, 2018.

[Tomašev *et al.*, 2019] Nenad Tomašev, Xavier Glorot, Jack W Rae, Michal Zielinski, Harry Askham, Andre Saraiva, Anne Mottram, Clemens Meyer, Suman Ravuri, Ivan Protsyuk, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*, 572(7767):116–119, 2019.

[Van Den Bergh and others, 2007] Frans Van Den Bergh et al. *An analysis of particle swarm optimizers*. PhD thesis, University of Pretoria, 2007.

[Zhang *et al.*, 2022] Zhaonian Zhang, Richard Jiang, Ce Zhang, Bryan Williams, Ziping Jiang, Chang-Tsun Li, Paul Chazot, Nicola Pavese, Ahmed Bouridane, and Azeddine Beghdadi. Robust brain age estimation based on smri via nonlinear age-adaptive ensemble learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:2146–2156, 2022.

[Zou and Schiebinger, 2018] James Zou and Londa Schiebinger. Ai can be sexist and racist—it's time to make it fair, 2018.