UDC 004.89

ESTIMATING BIOLOGICAL AGE USING KOLMOGOROV-ARNOLD NETWORKS ON SMALL DATA

V. Slipchenko, L. Poliahushko, V. Rudyk, V. Shatylo

Abstract. This article explores the issue of the application of Kolmogorov-Arnold Networks (KAN) for biological age estimation using a dataset of 344 male patients. The dataset includes biomarkers related to bone health and body composition. To enhance model performance, data preprocessing techniques such as polynomial interpolation for missing values and standardization were applied. Pearson and Spearman correlation analyses identified the most relevant biomarkers. Machine learning models were evaluated, along with neural networks and KANs. Experimental results demonstrate that KANs outperform traditional machine learning models and classical neural networks on small datasets. The optimal KAN architecture achieved a correlation coefficient of 0.93, a mean squared error of 18.81, and a mean absolute error of 2.8, surpassing the best-performing conventional models. These findings highlight the potential of KANs as a robust alternative for biological age estimation in resource-limited settings.

Keywords: Kolmogorov-Arnold Networks, biological age, bone mass density, machine learning, deep learning.

УДК 004.89

ОЦІНКА БІОЛОГІЧНОГО ВІКУ З ВИКОРИСТАННЯМ МЕРЕЖ КОЛМОГОРОВА-АРНОЛЬДА НА МАЛИХ ДАНИХ

В.Г. Сліпченко, Л.Г. Полягушко, В.І. Рудик, В.В. Шатило

Анотація. У статті досліджено можливості застосування мереж Колмогорова-Арнольда (KAN) для оцінки біологічного віку на основі аналізу вибірки, що складалася з 344 чоловіків. Набір даних містить біомаркери, які характеризують стан кісткової тканини та склад тіла. Для підвищення точності прогнозування використано методи попередньої обробки даних, зокрема поліноміальну інтерполяцію пропущених значень і стандартизацію показників. Кореляційний аналіз за Пірсоном і Спірменом дозволив визначити найбільш значущі біомаркери для побудови моделі. Було здійснено порівняльне тестування традиційних моделей машинного навчання, класичних нейронних мереж і мереж KAN. Результати експерименту показали, що архітектура KAN перевершує інші підходи при роботі з невеликими вибірками. Зокрема, оптимальна конфігурація KAN продемонструвала коефіцієнт кореляції 0,93, середньоквадратичну помилку 18,81 та середню абсолютну помилку 2,8 року, що є суттєво кращими показниками, ніж у конвенційних моделей. Отримані

результати підтверджують перспективність використання KAN для оцінки біологічного віку, особливо в умовах обмежених наборів медичних даних.

Ключові слова: мережі Колмогорова-Арнольда (KAN), біологічний вік, мінеральна щільність кісткової тканини, машинне навчання, глибинне навчання.

INTRODUCTION

Biological age (BA) is a comprehensive indicator of the functional state of the human body. Unlike chronological age, which counts years from birth, biological age assesses the physiological state of different body systems, reflecting the actual rate of aging processes [1, 2, 3]. Typically, the assessment of biological age relies on specific health markers, known as biomarkers.

One effective method for estimating biological age involves analyzing bone tissue conditions. Age-related changes in bone density and structural characteristics serve as reliable biomarkers, indicating an individual's aging process and risks for conditions such as osteoporosis and bone fractures [4]. Osteoporosis has become a growing global health concern. Prevalence statistics for osteoporosis vary significantly across populations: approximately 10 million individuals aged over 50 in the United States; 50% of women and 20% of men over 50 in the United Kingdom; 25–62% of postmenopausal women in India; and nearly 50% of individuals aged over 60 in Ukraine [5]. These data highlight the critical need for accurate and timely diagnostic and preventive measures.

Conventionally, statistical methods such as regression analysis and principal component analysis have been used to estimate biological age [6]. However, machine learning, in particular neural network-based models, has recently emerged as a more efficient approach for managing complex and multidimensional medical data. For instance, neural network models such as ST-ResNet achieved a mean absolute error (MAE) of only 0.455 years, while BoNet+ showed a MAE of 0.76 years, even when dealing with low-quality X-ray images [7, 8]. Nevertheless, these advanced models generally require large training datasets.

However, a significant challenge in this field remains the accurate prediction of biological age when small datasets are available. This situation is common in medical research conducted in resource-limited countries such as Ukraine. Under such conditions, classical machine learning methods often show errors of more than six years [9, 10]. One potential solution to this issue is the application of Kolmogorov-Arnold Networks (KAN) [11], which have recently demonstrated high performance even with limited data samples [12, 13].

The aim of this study is to evaluate the effectiveness of Kolmogorov-Arnold Networks (KAN) in estimating the biological age of men by analyzing their bone tissue conditions, particularly in situations of limited data availability.

To this aim, the following objectives have been defined:

– analyze and preprocess medical data to identify correlations between bone health indicators and biological age;

– compare the predictive accuracy of KAN with traditional machine learning methods and other neural network models;

– determine the optimal configuration and parameters of the KAN model for the most accurate biological age estimation.

The results of this study will help improve methods for early detection of bone aging and risk assessment for osteoporosis. The results can support preventive healthcare globally and help advance biological age research, especially in countries such as Ukraine, where collecting large amounts of medical data is difficult.

MATERIALS AND METHODS

Dataset. In this study, a dataset provided by the D.F. Chebotarev Institute of Gerontology of the National Academy of Medical Sciences of Ukraine [14] was used. The dataset includes the following biomarker indicators: body mass index (BMI), the fracture risk assessment tool score (FRAX) for all bones and the hip only, data on bone mass density of the lumbar spine, right and left femoral neck, proximal right and left femur, radial bone, and trabecular bone score, obtained from 345 male patients aged 40 to 92 years (mean age 59 years). More detailed information about the dataset is provided in Table 1.

Table 1. Detailed information about the dataset

Biomarker name	Unit	Mean	SE
Body mass index (BMI)	kg/m^2	27.182	0.243
10-year probability of major osteoporotic fractures (FRAX-all)	%	3.174	0.085
10-year probability of hip fractures (FRAX-hip)	%	1.771	0.044
Lumbar spine bone mineral density (LS-BMD)	g/cm^2	1.023	0.011
Femoral right neck bone mineral density (FN-BMD-R)	g/cm^2	0.765	0.008
Femoral left neck bone mineral density (FN-BMD-L)	g/cm^2	0.767	0.008
Proximal right femur (hip) bone mineral density (Hip-BMD-R)	g/cm^2	0.955	0.009
Proximal left femur (hip) bone mineral density (Hip-BMD-L)	g/cm^2	0.964	0.009
Bone mineral density ultra-distal radius of the forearm (UDR-BMD)	g/cm^2	0.744	0.005
The trabecular bone scores (TBS)	units	1.32	0.006

Pre-processing. To restore missing values in the dataset, a second-order polynomial interpolation method was used, specifically the interpolate method with the order = 2 parameter from the pandas library [15]. This method approximates missing data by constructing a polynomial of the specified order. This approach allows for the retention of rows where only a few indicators have missing values, thereby increasing the accuracy of subsequent analysis.

For data normalization, the StandardScaler method from the sklearn library [16] was employed, which scales the values of all indicators to a range from -1 to 1, with the mean value being centered around zero. The transformation is implemented through

the following formulas:

$$z = \frac{x - \mu}{\sigma},$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n}(x_i),$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2},$$

where z is the scale value; x is the current value; μ is the mean value; n is the sample size; and σ is the standard deviation.

This transformation ensures that all features contribute equally to the learning process, preventing any single variable from dominating due to differences in scale.

Correlations. To evaluate the correlation of indicators with the target variable, namely the chronological age of participants, Pearson and Spearman correlation coefficients were used [17]. The use of both correlation measures ensures that important dependencies in the dataset are not overlooked, capturing both linear and monotonic relationships.

The Pearson correlation coefficient measures the strength and direction of the linear relationship between two continuous variables and is calculated by the formula:

$$r_p = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2(y_i - \bar{y})^2}},$$

where x_i and y_i are data points; $\bar{x}$ is the mean of the x-values; $\bar{y}$ is the mean of the y-values; and n is the sample size.

Values range from -1 to 1, where 0 indicates no correlation, negative values indicate an inverse relationship and positive values indicate a direct relationship. This method is particularly effective for capturing linear dependencies.

The Spearman correlation coefficient assesses the monotonic relationship between variables, making it useful for detecting both linear and nonlinear trends. It is calculated by the formula:

$$r_S = 1 - \frac{6\sum_i^n d_i^2}{n(n^2 - 1)},$$

where d_i is the square of the difference in the ranks of the two coordinates for each point (x_i, y_i); and n is the sample size.

It works similarly to Pearson's coefficient; however, it is based on rank ordering and is more robust against outliers and non-normally distributed data.

P-value was used to determine the statistical significance of the observed correlations [18]. It represents the probability of obtaining the given correlation coefficient, or a more extreme one, assuming that the null hypothesis (i.e., no correlation exists) is true. A lower P-value indicates stronger evidence against the null hypothesis. Typically, a threshold of $\alpha = 0.05$ is used, meaning that correlations with $p < 0.05$ are considered statistically significant.

To obtain the p-value, the value of t-statistics is first calculated. It quantifies how much the observed correlation deviates from zero in terms of standard error [19]. Larger absolute values of t-statistics indicate stronger evidence of a significant

correlation. It is calculated as:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}},$$

where r is the correlation coefficient (Pearson or Spearman); and n is the sample size.

After obtaining this value, the p-value is derived from the t-distribution, considering that the degrees of freedom two less than sample size.

Accuracy metrics. To evaluate the effectiveness and accuracy of the predictions, several metrics were used, including the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), as well as a specific metric for assessing biological age – the Pearson's correlation coefficient between chronological and biological age (r).

The coefficient of determination measures how well a model explains the variance in the target variable. It ranges from 0 to 1, where 1 indicates a perfect fit, and 0 means the model explains no variance. A higher R^2 suggests better predictive accuracy. It is calculated as:

$$R^2 = \frac{\sum_i^n (\hat{y}_i - \bar{y})^2}{\sum_i^n (y_i - \bar{y})^2},$$

where $\hat{y}_i$ are predicted value; $\bar{y}$ are mean value; y_i are actual values; and n is the sample size.

Mean Absolute Error measures the average absolute error between predictions and actual values, giving equal weight to all errors. It is less sensitive to outliers and is in the same units as the original data. It is calculated as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|,$$

where $\hat{y}_i$ are predicted value; y_i are actual values; and n is the sample size.

Mean Squared Error, on the other hand, squares the errors, making it more sensitive to larger discrepancies, thus penalizing big errors more heavily. It is expressed in squared units of the original data. It is calculated as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2,$$

where $\hat{y}_i$ are predicted value; y_i are actual values; and n is the sample size.

Machine learning. For machine learning tasks, the sklearn library was utilized, along with compatible libraries such as xgboost [20], lightgbm [21], and catboost [22], which provide enhanced and specialized models. These libraries offer a comprehensive range of algorithms, from classical methods like linear regression, decision trees, and K-nearest neighbors to advanced techniques like boosting, bagging, and ensemble models. The inclusion of xgboost and lightgbm brings efficient gradient boosting algorithms that are highly effective in handling large datasets, while catboost offers robust solutions for categorical data. Together, these libraries enable the use of both traditional and state-of-the-art machine learning models.

To find the optimal parameters for testing each of the models, the GridSearchCV method from the sklearn library was used. It performs cross-validation for each combination of parameters, allowing the selection of the best-performing set of hyperparameters based on model performance metrics. This process helps ensure that

the model achieves its maximum efficiency by tuning parameters such as learning rate, number of estimators, or maximum depth, depending on the specific algorithm used.

For deep learning, the Keras library was employed, built on top of TensorFlow [23], providing a flexible and efficient framework for neural network development. The optimizers module from TensorFlow was used to fine-tune model training, offering various optimization algorithms such as Adagrad, Adam, and RMSprop to enhance convergence speed and accuracy. Additionally, the metrics module was used to track model performance during training and validation, ensuring effective evaluation of loss and accuracy. To improve network expressiveness, the Parametric Rectified Linear Unit (PReLU) activation layer was incorporated [24], allowing adaptive learning of activation parameters, which helps mitigate issues like vanishing gradients and enhances model generalization. These components collectively contributed to the effective implementation and optimization of neural network models in this study.

In 2024, Kolmogorov-Arnold Networks (KAN) were introduced as a promising alternative to Multi-Layer Perceptron (MLP) [25]. Unlike MLPs, which have fixed activation functions on nodes (neurons), KANs have learnable activation functions on edges (weights). KANs do not have linear weights – each weight parameter is replaced by a univariate function parameterized as a spline. To work with the software implementation of these networks, the pykan library [26] developed by the authors of the original article was used.

RESULTS

To preserve information in the already small dataset, the data was first sorted by chronological age. Then, polynomial interpolation was used to replace all zero or missing values with approximated values based on the neighboring patients' ages. This approach allowed retaining rows with missing values for some markers, instead of discarding them and filling in the missing data with values close to the actual ones.

As a result of the initial processing, the dataset contains 344 rows, each consisting of 10 biomarkers and the patients' chronological ages (Fig. 1).

	Age	BMI	FRAX-all	FRAX-hip	LS-BMD	FN-BMD-R	FN-BMD-L	Hip-BMD-R	Hip-BMD-L	UDR-BMD	TBS
0	71	26.149	6.4	2.6	0.818	0.678	0.764	0.874	0.924	0.660	1.245
1	61	25.690	3.3	0.6	0.953	0.857	0.832	1.037	1.003	0.696	1.276
2	77	29.668	2.3	0.9	1.466	0.899	0.959	1.250	1.222	0.824	1.452
3	47	28.905	1.9	0.1	1.060	0.852	0.925	1.021	1.096	0.894	1.493
4	60	26.122	3.3	0.8	1.309	0.748	0.915	1.000	1.137	0.809	1.442
...	...	...	...	...	...	...	...	...	...	...	...
339	57	26.534	3.2	0.5	0.942	0.808	0.726	0.964	0.996	0.775	1.074
340	69	27.251	2.2	0.6	0.903	0.568	0.651	0.778	0.837	0.679	1.286
341	65	32.872	1.7	0.3	0.931	0.776	0.654	0.947	0.645	0.793	1.251
342	44	23.991	2.1	0.1	1.106	0.930	0.864	1.067	1.026	0.857	1.536
343	64	24.212	2.2	0.5	0.764	0.743	0.756	0.968	0.885	0.758	1.351

Fig. 1. Initial data set

Using Pearson and Spearman methods, the corresponding correlation coefficients, p-values, and t-statistics were obtained, as presented in Table 2.

Given the low correlation, the variables with an absolute correlation of less than 0.15, namely BMI, LS-BMD, Hip-BMD-R, and Hip-BMD-L, were excluded from the dataset. After this, all data were standardized using the StandardScaler method and split into two parts: the training set (75%, 258 rows) and the test set (25%, 86 rows).

Table 2. Correlation of biomarkers with chronological age

Biomarker name	Pearson`s correlation	Pearson`s p-value	Pearson`s t-statistics	Spearman`s correlation	Spearman`s p-value	Spearman`s t-statistics
BMI	0.012	8.30618×10^{-1}	0.2141	0.008	8.84701×10^{-1}	0.1451
FRAX-all	0.176	1.05890×10^{-3}	3.3027	0.231	1.53804×10^{-5}	4.3862
FRAX-hip	0.636	2.55892×10^{-40}	15.2229	0.726	1.17256×10^{-57}	19.5464
LS-BMD	0.129	1.68264×10^{-2}	2.4023	0.115	3.32615×10^{-2}	2.1376
FN-BMD-R	-0.192	3.45605×10^{-4}	-3.6149	-0.199	1.98422×10^{-4}	-3.7618
FN-BMD-L	-0.180	8.04475×10^{-4}	-3.3815	-0.202	1.63985×10^{-4}	-3.8112
Hip-BMD-R	-0.132	1.40984×10^{-2}	-2.4674	-0.128	1.71852×10^{-2}	-2.3944
Hip-BMD-L	-0.123	2.24059×10^{-2}	-2.2938	-0.132	1.45014×10^{-2}	-2.4571
UDR-BMD	-0.357	9.13084×10^{-12}	-7.0636	-0.354	1.35219×10^{-11}	-7.0009
TBS	-0.192	3.49754×10^{-4}	-3.6117	-0.224	2.82140×10^{-5}	-4.2448

For the automatic selection of the best parameters for training the machine learning models, the GridSearchCV method was used with the accuracy parameter set to neg_mean_squared_error. The obtained results are presented in Table 3.

Table 3. Best parameters for training models

Model`s name	Best parameters
KNeighborsRegressor	'n_neighbors': 5
GradientBoostingRegressor	'max_depth': 4, 'n_estimators': 200
RandomForestRegressor	'max_depth': 10, 'n_estimators': 200
XGBRegressor	'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 100
LGBMRegressor	'learning_rate': 0.1, 'max_depth': 4, 'n_estimators': 100
CatBoostRegressor	'depth': 5, 'iterations': 200, 'learning_rate': 0.1
AdaBoostRegressor	'estimator': XGBRegressor(), 'learning_rate': 0.5, 'n_estimators': 50
BaggingRegressor	'max_samples': 1.0, 'n_estimators': 500
BayesianRidge	'alpha_1': 1e-06, 'alpha_2': 1e-06, 'lambda_1': 1e-08, 'lambda_2': 1e-06
ElasticNet	'alpha': 0.01, 'l1_ratio': 0.1, 'max_iter': 500
PLSRegression	'n_components': 4
DecisionTreeRegressor	'max_depth': 30, 'min_samples_split': 5
Lasso	'alpha': 0.1
MLPRegressor	'activation': 'relu', 'alpha': 0.01, 'hidden_layer_sizes': (50, 50), 'max_iter': 3000, 'solver': 'adam'

After determining the optimal parameters, the models were trained and tested, with the results shown in Table 4. The highest accuracy metrics were achieved by the CatBoostRegressor model, with a correlation coefficient between the predicted biological and chronological age of 0.92, MSE of 21.71, MAE of 3.12, and R^2 of 0.84.

Table 4. Machine learning model testing results

Model`s name	r	MSE	MAE	R^2	Time, s
KNeighborsRegressor	0.7051	73.0693	6.7744	0.4693	0.0046
GradientBoostingRegressor	0.9055	25.5564	3.2612	0.8144	0.1401
RandomForestRegressor	0.9122	23.8812	3.1461	0.8265	0.2318
XGBRegressor	0.9146	23.1514	3.0412	0.8319	0.0264
LGBMRegressor	0.9065	25.0477	3.1219	0.8181	0.0137
CatBoostRegressor	**0.9231**	**21.7106**	**3.1161**	**0.8423**	**0.1668**
AdaBoostRegressor	0.9162	23.1531	3.1663	0.8318	2.1076
BaggingRegressor	0.9147	23.2832	3.0378	0.8309	0.8463
BayesianRidge	0.7788	59.1709	6.4304	0.5703	0.0034
ElasticNet	0.7789	59.3665	6.4434	0.5688	0.0019
PLSRegression	0.7750	59.2691	6.4271	0.5695	0.0031
DecisionTreeRegressor	0.8882	30.5093	3.7819	0.7784	0.0025
Lasso	0.7795	59.5585	6.4594	0.5674	0.0019
MLPRegressor	0.9123	24.1630	3.3032	0.8245	2.9501

In addition to the selected machine learning methods, several neural network models with different architectures were tested (the number of neurons in the corresponding layers is indicated in the name, e.g., MLP_64_32_1 – first layer with 64 neurons, second layer with 32 neurons, third layer with 1 neuron). The following fixed parameters were chosen for each network: Adagrad optimizer with a learning rate of 0.01, mean absolute error loss function and batch_size equal to 25.

The best accuracy was achieved by the model with the following architecture: an input layer with 6 neurons (corresponding to the number of input parameters), the first hidden layer with 128 neurons, the second hidden layer with 64 neurons, the third hidden layer with 16 neurons, and an output layer with 1 neuron. After testing, it demonstrated the following results: a correlation coefficient between the predicted biological and chronological age of 0.91, MSE of 22.83, MAE of 3.13, and R^2 of 0.83. A comparison with other models is presented in Table 5.

Table 5. Results of testing classical neural networks

Model`s name	r	MSE	MAE	R^2	Time, s
MLP_256_128_64_16_1	0.8823	35.7992	4.3178	0.7401	1.2639
MLP_128_64_32_16_1	0.9098	27.2812	3.8839	0.8018	1.4921
MLP_128_32_16_8_1	0.8993	26.3451	3.5427	0.8086	1.9737
MLP_64_32_16_8_1	0.9056	24.8910	3.7605	0.8192	2.9644
MLP_32_16_8_1	0.8487	39.5009	4.7729	0.7131	4.4202
MLP_16_64_32_16_1	0.9039	25.4172	3.4627	0.8154	3.0542
MLP_64_32_1	0.8585	43.9039	4.3985	0.6990	5.1929
MLP_128_64_16_1	**0.9146**	**22.8338**	**3.1285**	**0.8341**	**2.1848**

Several variants of the Kolmogorov-Arnold network were also tested, as shown in Table 6. The corresponding architectural parameters are specified in the model's name: KAN_6_1_1_grid3_k3 – input layer dimension 6, number of hidden neurons 1, output layer dimension 1, grid intervals 3, cubic spline 3.

The optimal architecture was found to be the KAN model with 6 input parameters, 3 hidden layers, a grid size of 3, and a spline of order 5. The correlation coefficient reached 0.93, MSE was 18.81, MAE was 2.8, and R^2 was 0.86.

Table 6. Results of testing Kolmogorov-Arnold networks

Model`s name	r	MSE	MAE	R^2	Time, s
KAN_6_1_1_grid3_k3	0.8998	26.7811	3.4477	0.8055	3.5169
KAN_6_3_1_grid3_k3	0.9146	22.8338	3.1285	0.8342	4.3357
KAN_6_6_1_grid3_k3	0.9237	20.9441	2.9502	0.8479	4.8969
KAN_6_12_1_grid3_k3	0.9250	20.0287	2.8131	0.8545	6.0963
KAN_6_1_1_grid6_k3	0.8998	26.7811	3.4476	0.8055	3.5388
KAN_6_3_1_grid6_k3	0.9323	18.8108	2.7960	0.8634	4.3032
KAN_6_6_1_grid6_k3	0.9237	20.9441	2.9502	0.8479	4.8697
KAN_6_12_1_grid6_k3	0.9250	20.0287	2.8131	0.8545	6.1817
KAN_6_1_1_grid3_k5	0.8998	26.7811	3.4476	0.8055	3.5465

KAN_6_3_1_grid3_k5	**0.9323**	**18.8108**	**2.7960**	**0.8634**	**4.2122**
KAN_6_6_1_grid3_k5	0.9237	20.9441	2.9502	0.8479	4.8501

For each of the best models, the corresponding correlation plots between the predicted biological and chronological age were constructed (Fig. 2).

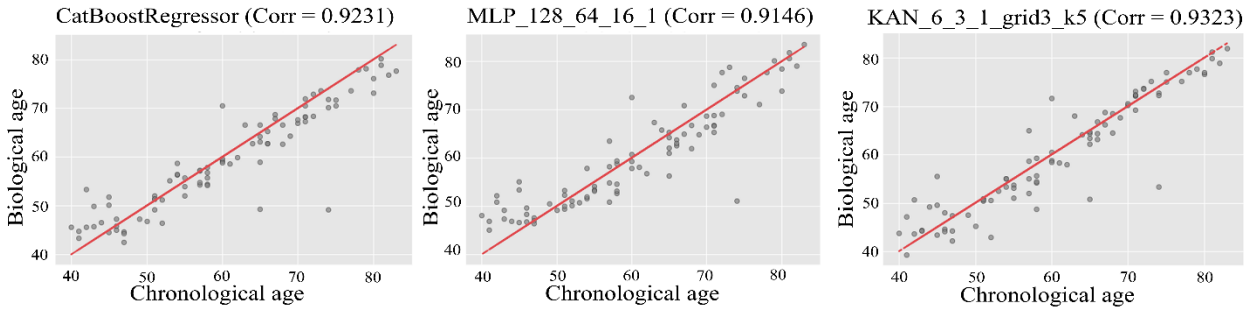

Fig. 2. Correlation graphs of models with the best accuracy rates

Figure 3 visualizes the architecture of the network that demonstrated the highest accuracy, namely KAN_6_3_1_grid3_k5.

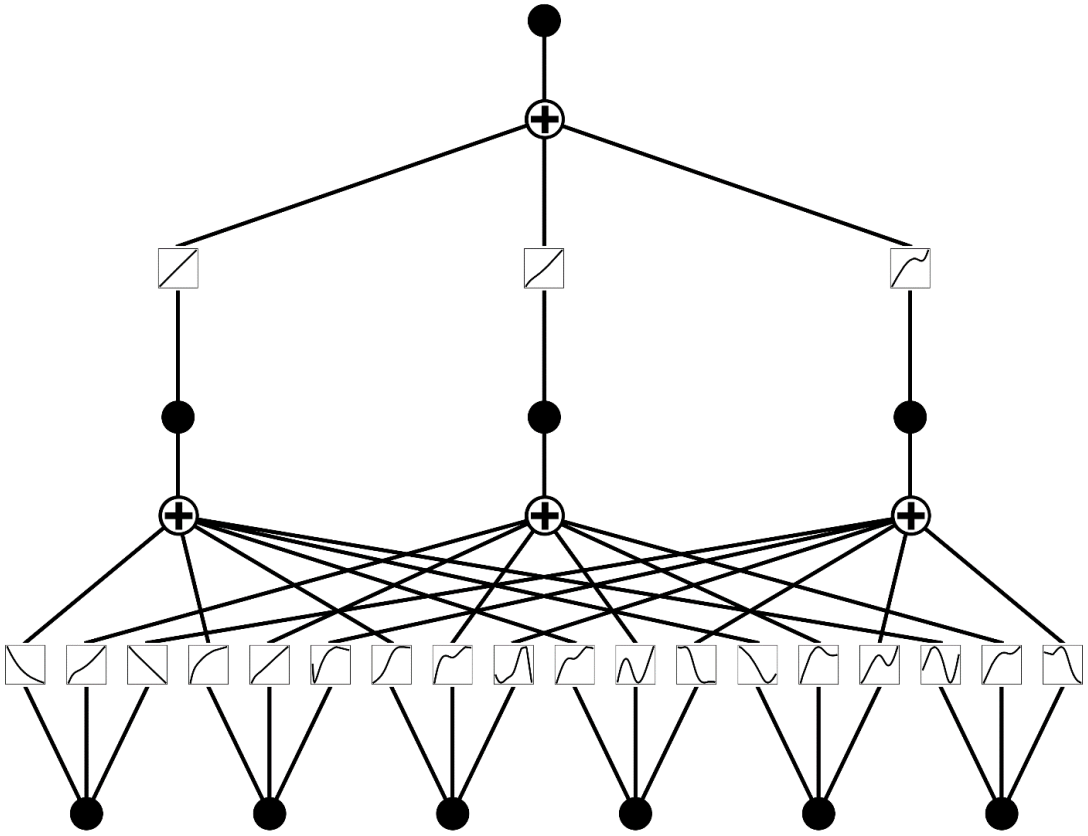

Fig. 3. Network architecture KAN_6_3_1_grid3_k5

DISCUSSION

This study is the first to apply the Kolmogorov-Arnold Networks (KAN) method for determining the biological age of men based on bone mineral density (BMD) parameters and several other key biomarkers. The obtained results demonstrate high estimation accuracy (correlation coefficient r = 0.93, MAE = 2.8 years), surpassing

classical approaches, including conventional and advanced machine learning models as well as deep learning techniques.

In comparison with previous studies, one notable work is that of Grygorieva et al. [27], which analyzed a mixed sample of 121 individuals (77 women and 44 men) and employed multiple regression methods to estimate the biological age of bone tissue. That study achieved a correlation coefficient r = 0.615 and MAE = 8.16 years. In our research, despite utilizing a larger dataset (344 participants, all male), the accuracy significantly improved due to the implementation of KAN. This underscores the ability of KAN to more effectively model the nonlinear relationships between biomarkers and biological age.

Another relevant example is a model developed for estimating the biological age of women, which analyzed a dataset with over 3200 individuals and achieved high accuracy (r = 0.93-0.94 and MAE = 2.1-2.2 years) [5]. However, this model was trained on a substantially larger dataset. In contrast, our method achieves comparable results (MAE = 2.8 years) even with a smaller sample size (344 men). This suggests that KAN can be an effective solution for tasks with limited data availability, particularly when focusing on the male population's specific characteristics.

A distinctive feature of this study is its exclusive focus on a male cohort. Many previous studies either primarily included women or utilized mixed samples, which could introduce biases due to sex-based differences in bone structure and density. By concentrating solely on men, we developed a more stable and homogeneous model that better accounts for the specific age-related and metabolic characteristics of the male body.

Another advantage of the proposed approach is its reliance on numerical biomarkers (such as DXA and FRAX) rather than imaging data, making the model less dependent on variations in diagnostic equipment settings and types. This universality facilitates the system's implementation in different clinical settings and could enhance the accuracy of osteoporosis risk screening and other age-related bone conditions.

Nevertheless, certain limitations should be noted. First, despite achieving higher accuracy compared to traditional methods, the overall sample size (344 individuals) remains relatively small for broad population generalizations. Second, the study was conducted exclusively on Ukrainian men, which may affect the generalizability of the results to other ethnic groups. Future research aims to expand the dataset, including representatives from different age groups, and further analyze the practical aspects of implementing this method in clinical practice.

Thus, the application of the Kolmogorov-Arnold Networks method for determining the biological age of men has demonstrated high accuracy and superiority over classical approaches. The obtained results have potential applications in the early prediction of age-related changes in the skeletal system and the timely diagnosis of osteoporosis. Additionally, they could serve as a foundation for developing personalized recommendations in preventive and sports medicine.

CONCLUSIONS

As a result of this study, Kolmogorov-Arnold Networks (KAN) were applied for the

first time to estimate biological age, leading to the following conclusions:

– the indicators of bone mineral density in men show a sufficient correlation with chronological age, making them viable for biological age estimation

– when estimating the biological age of men based on bone mineral density data, machine learning methods based on boosting, particularly the CatBoostRegressor, yield slightly better results ($r = 0.92$, $MSE = 21.71$, $MAE = 3.12$, $R^2 = 0.84$) compared to neural network models ($r = 0.91$, $MSE = 22.83$, $MAE = 3.13$, $R^2 = 0.83$);

– in an identical task, Kolmogorov-Arnold Networks demonstrated a slightly higher correlation between chronological and biological age ($r = 0.93$), but significantly lower errors ($MSE = 18.81$, $MAE = 2.8$, $R^2 = 0.86$), indicating superior performance;

– given the limited dataset, the results suggest that KAN models outperform neural networks and other machine learning methods, particularly in small data scenarios.

ACKNOWLEDGMENTS

The authors would like to thank DSc, Prof. N. Grygorieva and PhD, Sr. Research Fellow A. Musiienko from the D.F. Chebotarev Institute of Gerontology of the National Academy of Medical Sciences of Ukraine for providing the data used in this study.

REFERENCES

1. J. Bortz et al., “Biological age estimation using circulating blood biomarkers,” Communications Biology, vol. 6, no. 1, pp. 1–10, Oct. 2023, doi: https://doi.org/10.1038/s42003-023-05456-z.

2. L. Chen et al., “Modeling biological age using blood biomarkers and physical measurements in Chinese adults,” EBioMedicine, vol. 89, pp. 104458–104458, Feb. 2023, doi: https://doi.org/10.1016/j.ebiom.2023.104458.

3. S. Li, C. Wen, X. Bai, and D. Yang, “Association between biological aging and periodontitis using NHANES 2009-2014 and mendelian randomization,” Scientific Reports, vol. 14, no. 1, May 2024, doi: https://doi.org/10.1038/s41598-024-61002-9.

4. L. W. M. Diebel and K. Rockwood, “Determination of biological age: geriatric assessment vs biological biomarkers,” Current Oncol. Rep., vol. 23, no. 9, Jul. 2021, doi: https://doi.org/10.1007/s11912-021-01097-9.

5. V. Slipchenko, N. Grygorieva, L. Poliahushko, A. Musiienko, V. Rudyk, and V. Shatylo, “Estimation of the bone biological age using machine learning,” EUREKA Physics and Engineering, no. 1, pp. 175–186, Jan. 2025, doi: https://doi.org/10.21303/2461-4262.2025.003656.

6. S. E. C. Bafei and C. Shen, “Biomarkers selection and mathematical modeling in biological age estimation,” NPJ Aging, vol. 9, no. 1, Jul. 2023, doi: https://doi.org/10.1038/s41514-023-00110-8.

7. X. Chen, J. Li, Y. Zhang, Y. Lu, and S. Liu, “Automatic feature extraction in X-ray image based on deep learning approach for determination of bone age,” Future Gener. Comput. Syst., vol. 110, pp. 795–801, Sep. 2020, doi: https://doi.org/10.1016/j.future.2019.10.032.

8. J. Guo, J. Zhu, H. Du, and B. Qiu, “A bone age assessment system for real-world X-ray images based on convolutional neural networks”, Comput. & Elect. Eng., vol. 81, p. 106529, Jan. 2020, doi: https://doi.org/10.1016/j.compeleceng.2019.106529.

9. A.V. Pisaruk et al., “Human Biological age: Regression and Neural Network Models,” Fìzìologìčnij žurnal, vol. 69, no. 2, pp. 3–10, Mar. 2023, doi: https://doi.org/10.15407/fz69.02.003.

10. A. V. Pysaruk, V. P. Chyzhova, and V. B. Shatylo, “Biological Age of a Woman with Metabolic Syndrome,” Endokrynologia, vol. 28, no. 3, pp. 207–213, Sep. 2023, doi: https://doi.org/10.31793/1680-1466.2023.28-3.207.

11. G. Abramov, I. Gushchin, and T. Sirenka, “On the evolution of recurrent neural systems,” System research and information technologies, no. 4, pp. 77–85, Dec. 2024, doi: https://doi.org/10.20535/srit.2308-8893.2024.4.06.

12. S. Zinage, S. Mondal, and S. Sarkar, “DKL-KAN: Scalable Deep Kernel Learning using Kolmogorov-Arnold Networks,” arXiv (Cornell University), Jul. 2024, doi: https://doi.org/10.48550/arxiv.2407.21176.

13. H. Cui, S. Ning, S. Wang, W. Zhang, and Y. Peng, “ECG Signal Classification Using Interpretable KAN: Towards Predictive Diagnosis of Arrhythmias,” Algorithms, vol. 18, no. 2, pp. 90–90, Feb. 2025, doi: https://doi.org/10.3390/a18020090.

14. “D. F. Chebotarev State Institute of Gerontology” Kiev.ua, 2023. https://www.geront.kiev.ua/ (accessed Mar. 01, 2025).

15. Pandas, “Python Data Analysis Library — pandas: Python Data Analysis Library,” Pydata.org, 2019. https://pandas.pydata.org (accessed Mar. 01, 2025).

16. “scikit-learn: machine learning in Python — scikit-learn 0.22.2 documentation,” scikit-learn.org. https://scikit-learn.org/stable (accessed Mar. 01, 2025).

17. R. R. Wilcox, “Robust Correlation Coefficients That Deal with Bad Leverage Points,” Methodology, vol. 19, no. 4, pp. 348–364, Dec. 2023, doi: https://doi.org/10.5964/meth.11045.

18. H. Zhang and Z. Wu, “The Generalized Fisher’s Combination and Accurate P-value Calculation under Dependence,” Methodology, vol. 79, no. 2, pp. 1159–1172, Mar. 2022, doi: https://doi.org/10.1111/biom.13634.

19. E. Johannesson, J. Ohlson, and S. Zhai, “The Explanatory Power of Explanatory Variables,” Rev Account Stud, vol. 29, pp. 3053–3083, Jul. 2023, doi: https://doi.org/10.1007/s11142-023-09781-w.

20. “XGBoost Documentation — xgboost 2.1.3 documentation,” Readthedocs.io, 2022. https://xgboost.readthedocs.io/en/stable (accessed Mar. 01, 2025).

21. “Welcome to LightGBM’s documentation! — LightGBM 4.6.0 documentation,” Readthedocs.io, 2025. https://lightgbm.readthedocs.io/en/stable

(accessed Mar. 01, 2025).

22. “catboost/catboost,” GitHub, Aug. 22, 2022. https://github.com/catboost/catboost (accessed Mar. 01, 2025).

23. “TensorFlow,” TensorFlow, 2019. https://www.tensorflow.org (accessed Mar. 01, 2025).

24. Y. Bodyanskiy and S. Kostiuk, “Adaptive hybrid activation function for deep neural networks,” System research and information technologies, no. 1, pp. 87–96, Apr. 2022, doi: https://doi.org/10.20535/srit.2308-8893.2022.1.07.

25. Z. Liu et al., “KAN: Kolmogorov-Arnold Networks,” arXiv (Cornell University), p. 50, Apr. 2024, doi: https://doi.org/10.48550/arxiv.2404.19756.

26. “Welcome to Kolmogorov Arnold Network (KAN) documentation! — Kolmogorov Arnold Network documentation,” Github.io, 2024. https://kindxiaoming.github.io/pykan (accessed Mar. 01, 2025).

27. N. V. Grygorieva, H. S. Dubetska, N. M. Koshel, A. V. Pisaruk, and I. A. Antoniuk-Shcheglova, “Mathematical model of the bone biological age based on the bone mineral density and quality indicex and Ukrainian FRAX model,” Bol′, sustavy, pozvonočnik, vol. 12, no. 1, pp. 16–22, Mar. 2022, doi: https://doi.org/10.22141/pjs.12.1.2022.324.