https://doi.org/10.15407/csc.2024.02.065
UDC 004.05

O.S. BYCHKOV, Doctor of Technical Sciences, Professor,
Taras Shevchenko National University of Kyiv
Bohdana Havrylyshyna str., 24, Kyiv, Ukraine, 02000,
ORCID: https://orcid.org/0000-0002-9378-9535,
oleksiibychkov@knu.ua

O.V. GEZERDAVA, Master,
Taras Shevchenko National University of Kyiv,
Bohdana Havrylyshyna str., 24, Kyiv, Ukraine, 02000,
ORCID: https://orcid.org/0009-0002-9494-5541,
oleksandrgezerdava@gmail.com

K.K. DUKHNOVSKA, Ph.D, Associate Professor,
Taras Shevchenko National University of Kyiv,
Bohdana Havrylyshyna str., 24, Kyiv, Ukraine, 02000,
ORCID: https://orcid.org/0000-0002-4539-159X,
kseniia.dukhnovska@knu.ua

O.I. KOVTUN, Ph.D, Associate Professor,
Taras Shevchenko National University of Kyiv,
Bohdana Havrylyshyna str., 24, Kyiv, Ukraine, 02000,
ORCID: https://orcid.org/0000-0003-0871-5097,
kovok@ukr.net

O.O. LESHCHENKO, PhD, Associate Professor,
Taras Shevchenko National University of Kyiv,
Bohdana Havrylyshyna str., 24, Kyiv, Ukraine, 02000,
ORCID: https://orcid.org/0000-0002-3997-2785,
lesolga@ukr.net

FITNESS TRACKER DATA ANALYTICS

The health status of patients is recorded in various sources, such as medical records, portable devices (smart watches, fitness trackers, etc.), forming a characteristic current health status of patients. The goal of the study was the development of medical card software for the analysis of data from fitness bracelets. This will provide an opportunity to collect data for further use of cluster analysis and improvement of the functionality and accuracy of medical monitoring.

The object of the study is the use of linear regression to analyze and predict heart rate based on data collected using fitness bracelets. In order to solve this problem, an information system was developed that uses linear regression to analyze the effect of parameters such as Very Active Distance, Fairly Active Minutes, and Calories on the heart rate (Value).

Cite: Bychkov O.S., Gezerdava O.V., Dukhnovska K.K., Kovtun O.I., Leshchenko O.O. Medical Card Information System for Data Analysis from Fitness Bracelets. *Control Systems and Computers*, 2024, 2, 65—76. https://doi.org/10.15407/csc.2024.01.065

Training and validation were performed on data from fitness bracelets. The results confirm the effectiveness of linear regression in predicting heart rate based on the parameters of fitness bracelets. The accuracy of the model was compared under the conditions of aggregation and without it, which allows us to draw conclusions about the optimal conditions for using linear regression for the analysis of fitness data.

The study proves the adequacy of the obtained results according to the Student's criterion. The calculated Student's t test is 1.31, with the critical test — 2.62. Which proves the adequacy of the developed model.

The results of the study confirm that the linear regression model is an effective tool for individual monitoring and optimization of physical activity based on data from fitness bracelets.

It is worth considering that the use of linear regression has its limitations and is not always the best choice for complex nonlinear dependencies. In such cases, other machine learning methods may need to be considered.

Keywords: *digital medical card, records, IoT, Fitbit tracker.*

Introduction

The development of artificial intelligence (AI) in medicine has stimulated interest in using its capabilities in various areas of society. Some of the most important areas are improving the accuracy of diagnostics, reducing healthcare costs, and preventing the spread of infectious diseases.

Access to quality data is crucial for the successful development and implementation of AI algorithms in medicine. However, existing data is often limited and does not take into account the patients' previous health conditions. Therefore, there is a need for "sensitive data" that is collected with the patient's consent and anonymized for machine learning.

This data not only helps improve individual care but also allows for the identification of epidemics and healthcare crises. AI algorithms can effectively analyze large amounts of data, which allows for rapid response to epidemics and the allocation of resources to prevent them. This approach strengthens public health surveillance and response systems.

This research is devoted to the development of an information technology that can analyze medical information and identify common patterns, trends, and templates that can be used for diagnostics and treatment of patients.

In the current context of rapid technological development and the impact of informatization on all spheres of life, the implementation of a medical information system (MIS) is an extremely important and necessary step in the development of medicine. The development of information systems for processing primary medical data (electronic medical record) is especially important.

Therefore, this research is devoted to the development of an electronic medical record, which is a key stage in the collection and systematization of medical data and its further processing using AI and other technologies. Scientific research on this topic is important because it can significantly improve the quality and availability of healthcare, as well as make it more efficient.

Literature Review and Problem Statement

Implementation of a modern medical information system that utilizes artificial intelligence for data collection, systematization, and processing can have a significant impact on various aspects of medicine and public health. Digital medical records of patients will serve as a source of information for this system. This, in turn, will provide physicians with more tools for disease diagnosis and allow them to predict disease progression and health risks for patients.

The study [1] presents the results of research using a convolutional neural network (CNN) on 144.170 electronic medical records, which included 63 types of childhood diseases. Based on the study results, an intelligent diagnostic aid tool for pediatricians was created.

In the work [2], electronic medical records were used to study the quality of care for people with Parkinson's disease. Specialized software was used to collect statistical data, and a sorting method using closed tabs was used to structure specific data.

In the study [3], an evaluation of various data stream processing methods from electronic medical records of patients suffering from chronic obstructive pulmonary disease was conducted. The main goal of this study was to establish similarity between patients, namely highlighting the most representative and clinically significant data from their electronic medical records. For this purpose, the following data were selected as clinical features: gender, body mass index, smoking status, and personal history of atopy. To quantitatively assess the contribution of each clinical feature to the overall similarity of patients, the authors developed a metric of relative variability, which compares the variability of the feature among N nearest neighbors of each observation with the variability of the feature in the entire cluster.

The aim of the work [4] was to develop an electronic medical record for patients with schizophrenia to improve the management of their clinical information. The study shows that such a record gives the medical team access to up-to-date and complete patient information, which, in turn, helps them make more informed and effective treatment decisions. The electronic medical record allows hospital staff to better organize and systematize information about patients, making it more accessible and convenient to use. This can significantly facilitate the treatment process for patients with schizophrenia, as the electronic medical record saves time and resources, and makes the process more clear and organized. Thanks to these advantages, the electronic medical record can help hospitals better manage patient information and, as a result, improve the overall provision of medical care.

The study [5] used data from the "Epic" health information management system. Using electronic clinical data, monitoring was conducted on the condition of patients with juvenile rheumatoid arthritis, revealing that correlates for the overall assessment of arthritis by a physician differ depending on the subtype of the disease.

Also, using data from the "Epic" health information management system, the study [6] analyzed a variety of real-time data streams from patients arriving at the emergency department of the hospital in order to forecast total hospitalization within a short period of time. Hospital bed planners closely collaborated with the research group to determine their requirements. They requested to send daily forecasts four times a day regarding the need for beds in the next four and eight hours for further comparison with their internal reports. As part of the study, a program was developed that automatically formats and sends emails to bed planners every four reporting hours. This program is based on a machine learning algorithm used to forecast bed availability in the hospital.

Besides medical information systems, more and more attention is being drawn to IoT in the field of medicine, which is capable of filling these systems with information in the future. An example of such research can be found in [7]. In this work, a device for conducting electroencephalograms of the brain was developed. The device records at least 16 channels of signals from the human brain and then transmits electronic clinical data to the patient's history, allowing for automated data analysis in the future.

Also, in study [8], an analysis of the accuracy of sleep registration using wearable devices, including Xiaomi Mi Smart Band 5, was conducted, comparing their data with polysomnography (PSG) data — the gold standard in sleep assessment. The study showed significant differences between sleep indicators obtained from Xiaomi Mi Smart Band 5 and PSG. Xiaomi Mi Smart Band 5 cannot completely replace PSG in determining sleep stages but can be useful as an initial means of assessing sleep quality. Wearable devices can be used to monitor sleep and serve as a basis for consultation with a doctor or adjustment of sleep habits.

In the study [9], the accuracy of sleep recording using wearable devices, including the Xiaomi Mi Smart Band 5, was analyzed by comparing their data with polysomnography (PSG) data — the gold standard in sleep assessment.

The study showed significant differences between the sleep parameters obtained from the Xiaomi Mi Smart Band 5 and PSG. The use of paired *t*-tests and Bland-Altman plots confirms a large difference between the means of different sleep parameters measured by both methods.

EBE analysis showed different levels of sensitivity, specificity, and agreement between Xiaomi and PSG in determining different sleep stages, such as light sleep, deep sleep, and REM sleep.

While the Xiaomi Mi Smart Band 5 cannot fully replace PSG in sleep stage determination, they can be useful for consumers as an initial tool for assessing sleep quality. These devices can be used to monitor sleep and serve as a basis for consultation with a doctor or for adjusting one's sleep habits.

Thus, the study confirms that wearable devices, including the Xiaomi Mi Smart Band 5, may be suitable for simplified sleep quality monitoring, but their accuracy does not allow them to fully replace professional PSG studies in sleep phase assessment.

In turn, the data format and their compatibility become a key element for further systematization and processing of data. Study [10] proposes a model for compatibility of medical data standards, privacy protection methods, and medical image measurements. This model applies Health Level Seven (HL7) and Digital Imaging and Communications in Medicine (DICOM) standards to medical image data standards. This approach ensures increased access to medical image data in accordance with privacy laws through de-identification methods. This study focuses on proposing a standard for the measurement values of standard materials, which eliminates the uncertainty in measurements that was not foreseen by previously existing standards for medical image analysis. The study found that medical image data standards are consistent with existing standards and also provide privacy protection for any medical images using de-identification methods.

In medical practice, it is important not only information technologies that help in the treatment of a patient or in the diagnosis of his disease. Information technologies that predict threats to human health are also important. Thus, in study [11], a model for predicting the thermal state of a person during physical activity in a hot environment is proposed. The model took into account the characteristics of the environment, namely air temperature, relative humidity and wind speed. The intensity of the load and the duration of its impact on the person are also taken into account here. As a result of the work of such a model, a dynamic of characteristics is obtained that reflects the physiological state of a person in hot conditions during physical exertion.

Overall, these studies confirm the potential of IoT and modern medical technologies to improve the collection, processing, and analysis of medical data. They demonstrate promising opportunities for the use of wearable devices to monitor patients and improve access to medical information. However, at the same time, they emphasize the importance of further research and improvement of technologies to ensure their reliability and accuracy in clinical use.

Information technology for disease analytics can combine elements of machine learning for data analysis and risk forecasting, data processing from wearable devices such as fitness trackers for real-time health monitoring, electronic medical records for accessing a patient's medical history and test results. This will provide advantages such as improving diagnostics through early detection and disease forecasting, personalizing treatment and disease prevention, increasing efficiency, and saving costs in the healthcare system.

The Purpose and Objectives of the Research

The aim of the research is to develop an information system for medical records that provides analysis of data obtained from fitness trackers. To achieve this goal, the following tasks were set:

- development and training of a mathematical model in the form of linear regression to predict heart rate;
- development of an electronic medical card with a module for predicting changes in heart rate based on data obtained from the fitness tracker.

The expected results of the research are the creation of an information system that will collect data from fitness trackers and use them to predict heart rate. As a result, there will be an improvement in the accuracy of medical monitoring through the use of data collected from fitness trackers. The information system will enable early detection of cardiovascular problems.

dailyActivity_merged.csv — Открыть в приложении «Microsoft Excel»

Id	ActivityDate	TotalSteps	TotalDistance	TrackerDistance	LoggedActivitiesDistance	VeryActiveDistance	ModeratelyActiveDistance	LightActiveDistance	SedentaryActiveDistance	VeryActiveMinutes	FairlyActiveMinutes	LightlyActiveM
1503960366	4/12/2016	13162	8.5	8.5	0	1.87999999523163	0.550000011920929	6.05999994277954	0	25	13	328
1503960366	4/13/2016	10735	6.96999979019165	6.96999979019165	0	1.57000005245209	0.689999997615814	4.71000003814697	0	21	19	217
1503960366	4/14/2016	10460	6.73999977111816	6.73999977111816	0	2.44000005722046	0.400000005960464	3.91000008583069	0	30	11	181
1503960366	4/15/2016	9762	6.28000020980835	6.28000020980835	0	2.14000010490417	1.25999999046326	2.82999992370605	0	29	34	209
1503960366	4/16/2016	12669	8.15999984741211	8.15999984741211	0	2.71000003814697	0.409999996423721	5.03999996185303	0	36	10	221
1503960366	4/17/2016	9705	6.48000001907349	6.48000001907349	0	3.19000005722046	0.779999971389771	2.50999999046326	0	38	20	164
1503960366	4/18/2016	13019	8.59000015258789	8.59000015258789	0	3.25	0.639999985694885	4.71000003814697	0	42	16	233
1503960366	4/19/2016	15506	9.88000011444092	9.88000011444092	0	3.52999997138977	1.32000005245209	5.03000020980835	0	50	31	264
1503960366	4/20/2016	10544	6.67999982833862	6.67999982833862	0	1.96000003814697	0.479999989271164	4.23999977111816	0	28	12	205
1503960366	4/21/2016	9819	6.34000015258789	6.34000015258789	0	1.3400000333786	0.349999994039536	4.65000009536743	0	19	8	211
1503960366	4/22/2016	12764	8.13000011444092	8.13000011444092	0	4.76000022888184	1.12000000476837	2.24000000953674	0	66	27	130
1503960366	4/23/2016	14371	9.03999996185303	9.03999996185303	0	2.80999994277954	0.870000004768372	5.3600001335144	0	41	21	262
1503960366	4/24/2016	10039	6.40999984741211	6.40999984741211	0	2.92000007629395	0.209999993443489	3.27999997138977	0	39	5	238
1503960366	4/25/2016	15355	9.80000019073486	9.80000019073486	0	5.28999996185303	0.569999992847443	3.94000005722046	0	73	14	216
1503960366	4/26/2016	13755	8.78999996185303	8.78999996185303	0	2.32999992370605	0.920000016689301	5.53999996185303	0	31	23	279
1503960366	4/27/2016	18134	12.210000038147	12.210000038147	0	6.40000009536743	0.409999996423721	5.40999984741211	0	78	11	243
1503960366	4/28/2016	13154	8.52999973297119	8.52999973297119	0	3.53999996185303	1.1599999666214	3.78999996185303	0	48	28	189
1503960366	4/29/2016	11181	7.15000009536743	7.15000009536743	0	1.05999994277954	0.5	5.57999992370605	0	16	12	243
1503960366	4/30/2016	14673	9.25	9.25	0	3.55999994277954	1.41999995708466	4.26999998092651	0	52	34	217
1503960366	5/1/2016	10602	6.80999994277954	6.80999994277954	0	2.28999996185303	1.60000002384186	2.92000007629395	0	33	35	246
1503960366	5/2/2016	14727	9.71000003814697	9.71000003814697	0	3.21000003814697	0.569999992847443	5.92000007629395	0	41	15	277
1503960366	5/3/2016	15103	9.65999984741211	9.65999984741211	0	3.73000001907349	1.04999995231628	4.88000011444092	0	50	24	254
1503960366	5/4/2016	11100	7.15000009536743	7.15000009536743	0	2.46000003814697	0.870000004768372	3.8199999332428	0	36	22	203
1503960366	5/5/2016	14070	8.89999961853027	8.89999961853027	0	2.92000007629395	1.08000004291534	4.88000011444092	0	45	24	250
1503960366	5/6/2016	12159	8.02999973297119	8.02999973297119	0	1.97000002861023	0.25	5.80999994277954	0	24	6	289
1503960366	5/7/2016	11992	7.71000003814697	7.71000003814697	0	2.46000003814697	2.11999988555908	3.13000011444092	0	37	46	175
1503960366	5/8/2016	10060	6.57999992370605	6.57999992370605	0	3.52999997138977	0.319999992847443	2.73000001907349	0	44	8	203
1503960366	5/9/2016	12022	7.71999979019165	7.71999979019165	0	3.45000004768372	0.529999971389771	3.74000000953674	0	46	11	206
1503960366	5/10/2016	12207	7.76999998092651	7.76999998092651	0	3.34999990463257	1.1599999666214	3.25999999046326	0	46	31	214

Fig. 1. Examples of datasets

```
Fitabase Data 4.12.16-5.12.16
    dailyActivity_merged.csv
    dailyCalories_merged.csv
    dailyIntensities_merged.csv
    dailySteps_merged.csv
    heartrate_seconds_merged.csv
    hourlyCalories_merged.csv
    hourlyIntensities_merged.csv
    hourlySteps_merged.csv
    minuteCaloriesNarrow_merged.csv
    minuteCaloriesWide_merged.csv
    minuteIntensitiesNarrow_merged.csv
    minuteIntensitiesWide_merged.csv
    minuteMETsNarrow_merged.csv
    minuteSleep_merged.csv
    minuteStepsNarrow_merged.csv
    minuteStepsWide_merged.csv
    sleepDay_merged.csv
    weightLogInfo_merged.csv
```

Fig. 2. Complete data set

```
In 42
def count_unique_ids(df):
    return df['Id'].nunique()
print("heartrate:", count_unique_ids(heartrate))
print("activity:", count_unique_ids(activity))
print("calories:", count_unique_ids(calories))
print("intensities:", count_unique_ids(intensities))
print("sleep:", count_unique_ids(sleep))
print("weight:", count_unique_ids(weight))
Executed at 2024.01.10 15:35:29 in 72ms

heartrate: 14
activity: 33
calories: 33
intensities: 33
sleep: 24
weight: 8
```

Fig. 3. Identified data

Research Materials and Methods

The dataset used in this study was taken from the official Kaggle website [12]. The data was collected through Amazon Mechanical Turk between December 3, 2016 and December 5, 2016. In total, the dataset contains data from Fitbit users who, in accordance with legal requirements, consented to the provision of personal data from the tracker, including physical activity, heart rate, and sleep monitoring at the minute level. Individual reports in the dataset can be analyzed by export session identifier (column A) or timestamp (column B). The difference in output means the use of different types of Fitbit trackers and individual tracking behaviors/preferences.

In Fig. 1 shows an example of data elements from dailyActivity_merged.csv:

The complete data set consists of 18 csv files (Fig. 2).

The analysis of unique identifiers turns out to be a key stage in understanding the properties of data and their distribution in the specified set (Fig. 3).

Identifying and comparing the number of unique IDs to the total number of records can help reveal potential anomalies, such as duplicates or data loss. This procedure allows for the detection of potential data quality issues, which is critical for the reliability and correctness of further analysis.

In addition, unique identifiers can be used as keys to join data from different sources or tables. This facilitates the identification and matching of information about the same objects or observations in different sources.

As can be seen from Fig. 3, not all activity indicators are present for the individuals who provided their data for analysis. This will be taken into account in the study.

The Pandas library in the Python programming language was used to process and transform the data contained in different sets (heartrate, intensities, calories, activity, sleep) (Fig. 4). This analysis is aimed at processing the date and time in text format for further analysis and use of this data in relevant research.

These actions are performed to standardize the representation of date and time into a special date/time object type, which simplifies further operations with time series analysis, sorting, filtering, and data visualization. This approach facilitates ease of processing and makes analysis more accessible for further scientific conclusions and use in research.

In addition, it is worth noting that for the convenience of this study, the data of all individuals was also summarized by each available parameter and presented in the form of graphs. First of all, the results of the study of the distribution of heart rate (HR) in the heartrate_selected_columns data-

```
# intensities
intensities['ActivityHour'] = pd.to_datetime(intensities['ActivityHour'], format="%m/%d/%Y
  %I:%M:%S %p")
intensities['time'] = intensities['ActivityHour'].dt.strftime("%H:%M:%S")
intensities['date'] = intensities['ActivityHour'].dt.strftime("%m/%d/%y")

# calories
calories['ActivityHour'] = pd.to_datetime(calories['ActivityHour'], format="%m/%d/%Y %I:%M:%S
  %p")
calories['time'] = calories['ActivityHour'].dt.strftime("%H:%M:%S")
calories['date'] = calories['ActivityHour'].dt.strftime("%m/%d/%y")

# activity
activity['ActivityDate'] = pd.to_datetime(activity['ActivityDate'], format="%m/%d/%Y")
activity['date'] = activity['ActivityDate'].dt.strftime("%m/%d/%y")

# sleep
sleep['SleepDay'] = pd.to_datetime(sleep['SleepDay'], format="%m/%d/%Y %I:%M:%S %p")
sleep['date'] = sleep['SleepDay'].dt.strftime("%m/%d/%y")
```

Fig. 4. Date and time format standardization for further analysis: Python code

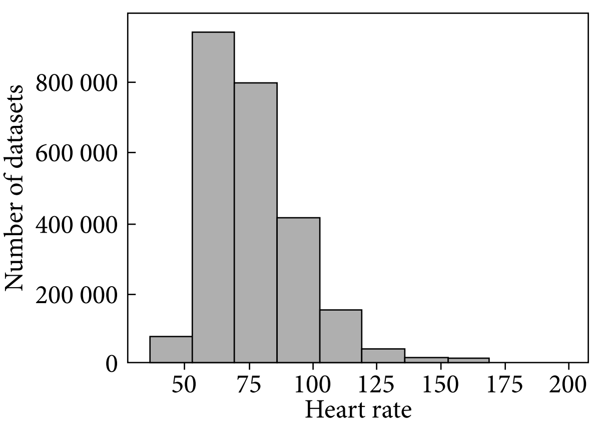

Fig. 5. Graph of heart rate distribution by activity states

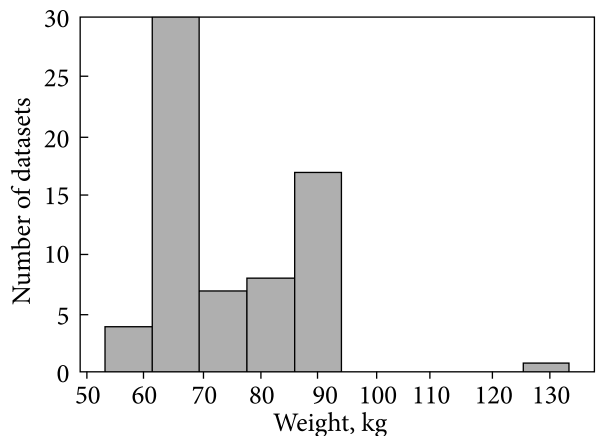

Fig. 6. Weight distribution chart

set were selected. This dataset contains HR values in beats per second, collected using a fitness bracelet to track health.

The graph (Fig. 5) of the distribution of values in the "Value" column in the heartrate_selected_columns dataset has several peaks. The first peak is in the region of 60 beats per second, and the second peak is in the region of 75 beats per second. This means that there are two main clusters of HR values in the dataset.

The first cluster, which is in the region of 60 beats per second, may represent a state of rest. This HR range coincides with the HR range that is typically observed in people at rest. The second cluster, which is in the region of 75 beats per second, may represent a state of activity. This HR range is higher than the HR range that is typically observed in people at rest. In addition, the graph has other peaks, which means that there are some HR values that are significantly higher or lower than the two main clusters. This may be due to the fact that some individuals lead a more active or less active lifestyle.

In turn, the graph (Fig. 6) shows the distribution of user weight in kilograms. Weight varies from 50 to 130 kg. A significant amount of data is concentrated in the 60¾70 kg range, and the largest number of values is in the 70¾80 kg range.

Another graph (Fig. 7) shows the distribution of calories consumed by the user. The number of

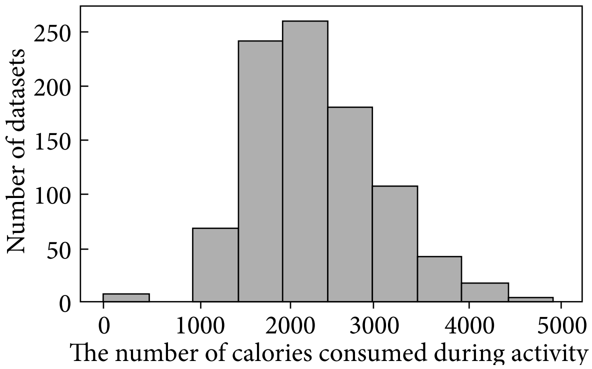

Fig. 7. Calorie distribution chart

calories varies from 0 to over 5000. The largest amount of data is concentrated in the range from 2000 to slightly over 3000 calories.

Thus, the data were prepared for model training by processing and standardizing information on physical activity, heart rate, and sleep monitoring collected using a Fitbit tracker.

Mathematical Model in the Form of Linear Regression for Predicting Heart Rate

This study uses a mathematical model of linear regression, which considers the dependence of the heart rate (labeled as "Value") on three main variables:

$$Value = \beta_0 + \beta_1 \times \textit{Very Active Distance} + \beta_2 \times \times \textit{Fairly Active Minutes} + \beta_3 \times \textit{Calories} + \varepsilon, \quad (1)$$

where:

Value — heart rate (dependent variable);

Very Active Distance — variable representing the distance traveled by the user during very physical activities such as running or vigorous aerobics. It is measured, for example, in kilometers or meters;

Fairly Active Minutes — the variable reflects the number of minutes during which the user spent some physical activity. This may include walking or light aerobics. Measured in minutes;

Calories — the variable represents the number of calories consumed by the user during the activity. It can take into account calories consumed during exercise and basal metabolism. It is measured in kilocalories (kcal) or joules (*J*);

β_0 — a constant (transition term);

β_1, β_2, β_3 — regression coefficients that determine the influence of the relevant variables;

ε — reflects random errors originating from unaccounted factors or random deviations from expectations.

This model allows you to quantify the effect of each variable on the heart rate and predict its value. Random error (ε) takes into account unknown or unpredictable factors that can affect the heart rate and introduce random deviations into predictions.

Development of an Electronic edical Card

The presented medical card system uses data from a fitness bracelet to monitor and diagnose the patient's health (Fig. 8). Fitness bracelets collect data such as the number of steps, heart rate, sleep quality and other parameters characterizing the state of health. The information system synchronizes data in real time.

In particular, one of the important directions is the use of data obtained from fitness bracelets for objective assessment of patients' condition. Thanks to the use of methods of machine learning and data analysis, it becomes possible not only to diagnose the current state of health, but also to predict its possible risks and trends. The presented information system uses a module for predicting a person's heart rate based on data obtained from a fitness bracelet. The module is built on the basis of the mathematical model (1).

To develop this module, popular Python libraries for working with data, modeling and visualization are used: pandas for data processing, scikit-learn (sklearn) for modeling linear regression, and matplotlib for visualization of results.

Real and predicted heart rate values are displayed on the graph (Fig. 9), which allows you to visually compare how well the model adapts to real data. This makes it easier to understand the effectiveness and scope of the linear regression model in this context.

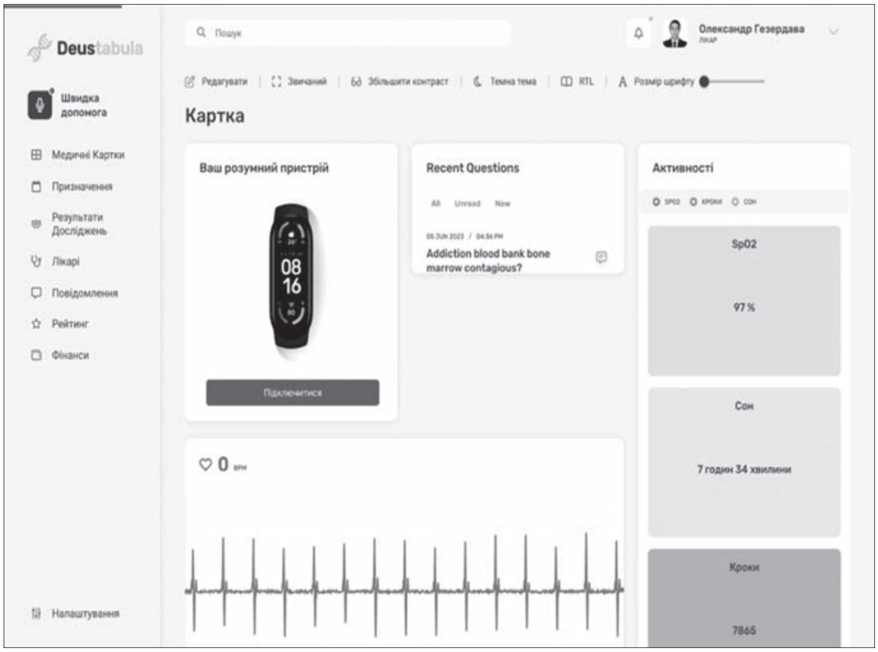

Fig. 8. Medical card interface

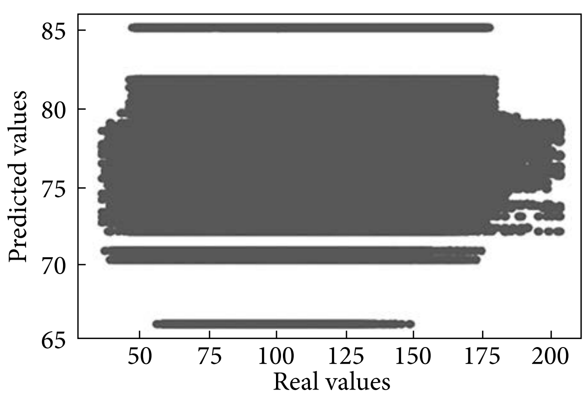

Fig. 9. Graphic presentation of forecasting results

Many points on the graph are closely clustered together, indicating a high degree of correlation between the predicted and actual values of heart rate. This suggests that the linear regression model effectively captures the relationship between the selected independent variables and heart rate, confirming its adequacy for these specific data.

However, attention should be paid to some outliers or deviations from the main cluster of points. These cases may indicate situations where the model encounters difficulties in accurately predicting heart rate. Perhaps they reflect the influence of other factors or anomalies in the data, which should be further examined and taken into account in the analysis of the results.

The graph also shows three horizontal lines, which may represent the mean or median values for different data groups. This may indicate the presence of subgroups or peculiarities in the data that could affect the model's results. Investigating these sections could lead to a deeper understanding of the dynamics of the interaction between the selected variables and heart rate.

Discussion of the Results of the Information System Using

The use of linear regression in the architecture of information systems for analyzing and predicting heart rate based on data from fitness trackers is an important research direction. Linear regression is a well-known statistical method widely used in solving numerous tasks of statistical data processing. It works well on large and sparse datasets without complex trends.

In building this model, we assumed an approximate linear relationship between heart rate and characteristics of physical activities. This has a simple explanation. During physical activities, a person's muscles require more oxygen. Oxygen is used to produce energy. During inhalation, oxygen enters the arterial blood, which is transported by the heart's work. During physical exertion, the heart works harder to supply oxygen to working muscles, which can affect heart rate. Typically, an increase in physical activity leads to an increase in heart rate. This explains the model's results.

The modeling showed that the expected heart rate will increase by 0.3996623 beats per minute with each meter walked. At the same time, since we use characteristics of a single physical activity, the expected heart rate will decrease by 0.11206182 with each second of this activity and by 0.00063471 with each calorie burned.

The model's adequacy has been proven, as the Student's criterion calculated is significantly smaller than the critical value.

From Fig. 2, it can be seen that the standard error, which shows how far the actual heart rate value can be from the expected value, is negligible.

Comparative analysis of the loss function helps to understand how different optimization methods affect the quality and speed of model training.

Forecasting heart rate has been studied for a long time and by various methods. However, unlike other models, the proposed model is the most accessible to users. You only need to have a fitness tracker and software for analyzing data from fitness trackers. The model predicts heart rate at any time during physical activity, unlike the model described in study [13], which predicts only for five seconds.

With the developed model, the user can plan physical workouts without harming their health. Since this model is aimed at predicting heart rate during exercise, which distinguishes it from the LSTM-BiLSTM-Att model [14], where the expected heart rate is calculated for a person at rest.

The study was conducted on data from young people of almost the same physical fitness level. How model (1) will behave for other age groups was not investigated.

Among the disadvantages of the model, it should be noted that linear regression typically uses independent variables. However, here, characteristics of a single event are considered, which is unlikely to prove that they are independent. However, the adequacy of the model has been proven.

Conclusions

As a result, user data from fitness trackers was systematized, and a mathematical model was created that can be used for individual health monitoring and optimization of physical activity to maintain cardiovascular health. The research results confirm that the linear regression model is an effective tool for individual monitoring and optimization of physical activity based on fitness tracker data. The study proves the adequacy of the obtained results by the Student's criterion. The calculated Student's t criterion is 1.31, with a critical value of 2.62, which proves the adequacy of the developed model.

Based on this model, an information system was developed, programmatically implemented, and described, which uses linear regression to study the impact of parameters such as Very Active Distance, Fairly Active Minutes, and Calories on heart rate (Value). The feature of the developed electronic medical card is that it can be used not only to record a person's health status but also for monitoring and forecasting it.

REFERENCES

1. Li, X., Wang, H., He, H., Du, J., Chen, J., & Wu, J. (2019). "Intelligent diagnosis with Chinese electronic medical records based on convolutional neural networks". *BMC bioinformatics*, 20, pp. 1—12. https://doi.org/10.1186/s12859-019-2617-8.
2. Connor, K. I., Siebens, H. C., Mittman, B. S., Ganz, D. A., Barry, F., Ernst, E. J., ... & Vickrey, B. G. (2020). "Quality and extent of implementation of a nurse-led care management intervention: care coordination for health promotion and activities in Parkinson's disease (CHAPS)". *BMC health services research*, 20, pp. 1—17. https://doaj.org/article/1a86bc727d50495e968f63d952da1982.

3. Pikoula, M., Kallis, C., Madjiheurem, S., Quint, J. K., Bafadhel, M., & Denaxas, S. (2023). "Evaluation of data processing pipelines on real-world electronic health records data for the purpose of measuring patient similarity". *Plos one*, 18 (6), e0287264. https://doi.org/10.1371/journal.pone.0287264.
4. Safdari, R., Hamidi, M., Aghaee, M., & Ghazi Saeedi, M. (2017). "Designing electronic card of health for schizophrenic patients". *Payavard Salamat*, 10 (6), pp. 479–487.
5. Miller, M.L., Ruprecht, J., Wang, D., Zhou, Y., Lales, G., McKenna, S., & Klein-Gitelman, M. (2011). "Physician assessment of disease activity in JIA subtypes. Analysis of data extracted from electronic medical records". *Pediatric Rheumatology,* 9, pp. 1–7. https://doi.org/10.1186/1546-0096-9-9.
6. King, Z., Farrington, J., Utley, M., Kung, E., Elkhodair, S., Harris, S., ... & Crowe, S. (2022). "Machine learning for real-time aggregated prediction of hospital admission for emergency patients". *NPJ Digital Medicine*, 5 (1), 104, pp. 1–12, https://doi.org/10.1038/s41746-022-00649-y.
7. Malakhov, K.S. (2023). "Insight into the digital health system of Ukraine (ehealth): Trends, definitions, standards, and legislative revisions". *International Journal of Telerehabilitation*, 15 (2). https://doi.org/10.5195/ijt.2023.6599.
8. Khiani, S., Iqbal, M.M., Dhakne, A., Thrinath, B.S., Gayathri, P.G., & Thiagarajan, R. (2022). "An effectual IOT coupled EEG analysing model for continuous patient monitoring measurement: sensors", 24, no 100597. https://doi.org/10.1016/j.measen.2022.100597.
9. Martínez-Martínez, F.J., Concheiro-Moscoso, P., Miranda-Duro, M.D. C., Boedo, F.D., Muiño, F.J. M., & Groba, B. (2020). "Validation of self-quantification Xiaomi band in a clinical sleep unit". In *Proceedings. MDPI.* Vol. 54, No. 1, p. 29. https://doi.org/10.3390/proceedings2020054029.
10. Kwon, O., & Yoo, S.K. (2021). "Interoperability Reference Models for Applications of Artificial Intelligence in Medical Imaging". *Applied Sciences*, 11 (6), p. 2704. https://doi.org/10.3390/app11062704.
11. Yermakova, I., Nikolaienko, A., Tadeieva, J., Bogatonkova, A., Solopchuk, Y., Gandhi, O.(2020). "Computer model for heat stress prediction during physical activity." 2020 IEEE 40th International scientific conference electronics and nanotechnology (ELNANO). Kyiv, Ukraine, 2020. April 22-24, pp. 569–573. https://doi.org/10.1109/ELNANO50318.2020.9088846.
12. FitBit Fitness Tracker Data. [online]. Available at: <https://www.kaggle.com/datasets/arashnic/fitbit> [Accessed 01 March, 2024].
13. Zhu, Z., Li, H., Xiao, J., Xu, W., & Huang, M. C. (2022). "A fitness training optimization system based on heart rate prediction under different activities". *Methods*, 205, pp. 89–96.
14. Lin, H., Zhang, S., Li, Q., Li, Y., Li, J., & Yang, Y. (2023). "A new method for heart rate prediction based on LSTM-BiLSTM-Att". *Measurement*, 207, 112384.

Received 30.04.2024

О.С. Бичков, д.т.н., професор,
Київський національний університет імені Тараса Шевченка,
02000, м. Київ, вул. Богдана Гаврилишина, 24, Україна,
ORCID: https://orcid.org/0000-0002-9378-9535,
oleksiibychkov@knu.ua

О.В. Гезердава, магістр,
Київський національний університет імені Тараса Шевченка,
02000, м. Київ, вул. Богдана Гаврилишина, 24, Україна,
ORCID: https://orcid.org/0009-0002-9494-5541,
oleksandrgezerdava@gmail.com

К.К. Духновська, к.т.н., доцент,
Київський національний університет імені Тараса Шевченка,
02000, м. Київ, вул. Богдана Гаврилишина, 24, Україна,
ORCID: https://orcid.org/0000-0002-4539-159X,
kseniia.dukhnovska@knu.ua

О.І. Ковтун, кандидат фіз.-мат. наук, доцент,
Київський національний університет імені Тараса Шевченка,
02000, м. Київ, вул. Богдана Гаврилишина, 24, Україна,
ORCID: https://orcid.org/0000-0003-0871-5097,
kovok@ukr.net

О.О. Лещенко, к.т.н., доцент,
Київський національний університет імені Тараса Шевченка, Україна,
02000, м. Київ, вул. Богдана Гаврилишина, 24, Україна,
ORCID: https://orcid.org/0000-0002-3997-2785,
lesolga@ukr.net

АНАЛІТИКА ДАНИХ З ФІТНЕС-БРАСЛЕТІВ

Вступ. Значний розвиток та досягнення в галузі штучного інтелекту (ШІ) спричинили інтерес до його впровадження в різних сферах суспільства. Не стала винятком і сфера охорони здоров'я та медицини. Серед напрямів використання ШІ в медицині можна виокремити багато аспектів. Водночас на найбільшу увагу заслуговує підвищення точності діагностики, зниження витрат на охорону здоров'я, профілактичне попередження епідеміологічних захворювань та географічне розширення отримання медичних послуг за допомогою телекомунікаційних технологій.

Мета статті. Метою дослідження є розробка інформаційної системи медичної картки для аналітики даних з фітнес браслетів. Це уможливить збирання даних для подальшого використання кластерного аналізу отриманих даних і вдосконалення функціоналу та точності медичного моніторингу.

Для досягнення цієї мети потрібно було виконати такі завдання:

- розробити та навчити математичну модель у вигляді лінійної регресії. Ця модель використовує параметри *Very Active Distance*, *Fairly Active Minutes* та *Calories* для прогнозування серцевого ритму (*Value*);
- розробити електронну медичну картку та в якості додаткового функціоналу впровадити в неї модуль прогнозування зміни у серцевому ритмі на основі даних, отриманих з фітнес-браслету.

Методи. Системний аналіз, лінійна рекурсія.

Результати. В результаті розроблено інформаційну систему медичної картки, яка може отримувати дані з фітнес-браслету і призначена для моніторингу та діагностування стану здоров'я пацієнта в реальному часі. Інформація з фітнес-браслета може включати кількість кроків, серцевий ритм, якість сну та інші параметри, які характеризують стан здоров'я.

Інформаційна система може синхронізувати дані в реальному часі та використовувати їх для покращення оцінки загального стану здоров'я пацієнта. Наприклад, зміни в серцевому ритмі або якості сну можуть вказувати на потенційні проблеми, які можуть вимагати подальшої уваги або медичного втручання.

Використання цих даних також може допомогти в лікуванні хронічних захворювань, таких як діабет або гіпертонія, завдяки моніторингу впливу різних факторів на стан здоров'я пацієнта. Процес також забезпечує лікарям можливість глибше зрозуміти стан здоров'я пацієнта, що дозволяє їм надавати більш персоналізовану та ефективну допомогу.

Висновки. У результаті було систематизовано дані користувачів фітнес браслетів та створено математичну модель, яка може використовуватися для індивідуального моніторингу здоров'я та оптимізації фізичної активності для підтримання здоров'я серцево-судинної системи. Результати дослідження підтверджують, що модель лінійної регресії є ефективним інструментом для індивідуального моніторингу та оптимізації фізичної активності на основі даних фітнес-браслетів. В дослідженні доводиться адекватність отриманих результатів за критерієм Стьюденса. Розрахований критерій Стьюденса t дорівнює 1.31, при критичному — 2.62, що і доводить адекватність розробленої моделі.

На основі цієї моделі було розроблено, програмно реалізовано та описано інформаційну систему, яка використовує лінійну регресію для вивчення впливу параметрів, таких як *Very Active Distance, Fairly Active Minutes* та *Calories*, на серцевий ритм (*Value*). Особливість розробленої електронної медичної картки полягає в тому, що її можливо використовувати не лише для фіксування стану здоров'я людини, але й для його моніторингу та прогнозування.

***Ключові слова**: цифрова медична картка, записи, IoT, трекер Fitbit.*