

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»  
НАВЧАЛЬНО-НАУКОВИЙ КОМПЛЕКС «ІНСТИТУТ ПРИКЛАДНОГО  
СИСТЕМНОГО АНАЛІЗУ»  
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

Бакалаврська дипломна робота на тему:  
«Моделі і оцінювання прогнозів індексу людського розвитку»

ВИКОНАЛА:

---

СТУДЕНТКА ІV КУРСУ ГРУПИ КА-24

КОЛОМІЄЦЬ КАТЕРИНА

НАУКОВИЙ КЕРІВНИК:

Д.Т.Н., ПРОФ. БІДЮК П. І.

# Актуальність теми

---

- Дослідження тенденцій розвитку людського капіталу нації
- Дослідження способів покращення якості життя
- Дослідження локальних проблем людського розвитку
- Прийняття рішень стосовно впровадження державних реформ у сфері охорони здоров'я, освіти.

# Об'єкт, предмет і мета роботи

---

**Об'єкт** – часові ряди індексів розвитку людини України та його складових, методи їх аналізу, побудови моделей для прогнозування

**Предмет** – математичні моделі і методи опису нестаціонарних процесів, оцінювання та аналізу якості прогнозів

**Мета** роботи – побудувати адекватні моделі для прогнозування індексу людського розвитку декількома методами, порівняти результати.

**Методи** дослідження – теорія моделювання і прогнозування часових рядів, регресійний аналіз, статистичні методи аналізу якості прогнозів та адекватності моделей

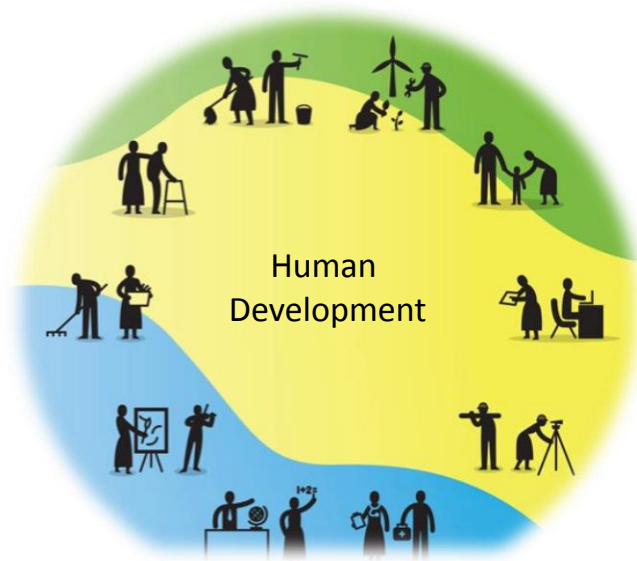
# Постановка задачі

---

1. Дослідити особливості процесів, пов'язаних із формуванням індексу людського розвитку
2. Виконати огляд методів для моделювання нестационарних процесів
3. Знайти дані для проведення аналізу
4. Побудувати моделі на знайдених вхідних даних у системі Eviews
5. Розробити та реалізувати програму для моделювання і прогнозування нестационарних процесів кількома способами
6. Провести якісний аналіз побудованих моделей

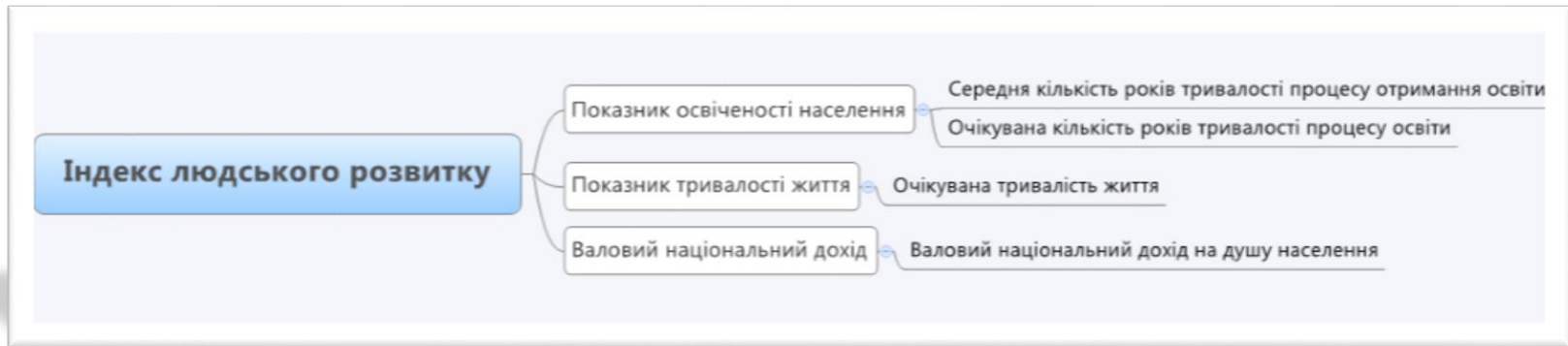
# Що таке людський розвиток?

---



# Індекс людського розвитку. Формування

---



$$ІЛР = \sqrt[3]{I_{здоров'я} \cdot I_{освіченість} \cdot I_{економіка}}$$

# Складові ІЛР

---

$$I_{\text{здоров'я}} = \frac{LE - 20}{LE_{\max} - 20}$$

$$I_{\text{освіченість}} = \frac{MYSI + EYSI}{2}$$

$$EYSI = \frac{EYS}{18}$$

$$MYSI = \frac{MYS}{15}$$

$$I_{\text{економіка}} = \frac{\ln(GNIpc) - \ln(GNIpc_{\min})}{\ln(GNIpc_{\max}) - \ln(GNIpc_{\min})}$$

# Опис вхідних даних

The screenshot shows the World Bank Data website. At the top, there is a navigation bar with the World Bank logo, the tagline 'Working for a World Free of Poverty', and language options: English, Español, Français, العربية, Русский, 中文. A search bar is also present. Below the navigation bar, there is a 'Data' section with a red header. Underneath, there are tabs for 'By Country', 'By Topic', 'Indicators', 'Data Catalog', 'Microdata', 'Initiatives', 'What's New', 'Support', and 'Products'. A sub-navigation bar shows 'This page in: English, Español, Français, العربية, 中文'. The main content area features a line graph on the left, a 'Find an indicator' search box, and a 'The World at a Glance' box. Below the graph, there is a section titled 'World Development Indicators' with a red underline. To the left of this section, there are two article teasers: 'Troubled waters: en: sustainable fishing a oceans' and 'Protecting life on lan'. To the right, there is a description of the World Development Indicators: 'The primary World Bank collection of development indicators, compiled from officially-recognized international sources. It presents the most current and accurate global development data available, and includes national, regional and global estimates.'

- GNI per capita, PPP (constant 2011 international \$)
- Life expectancy at birth, total (years)
- School life expectancy, both sexes (years)
- Mean years of schooling of the population age 25+. Total(years)



# Нестационарні часові ряди

---

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=0}^q b_j \varepsilon(k-j),$$

$$b_0 = 1, \sum_{j=1}^q b_j = 1$$

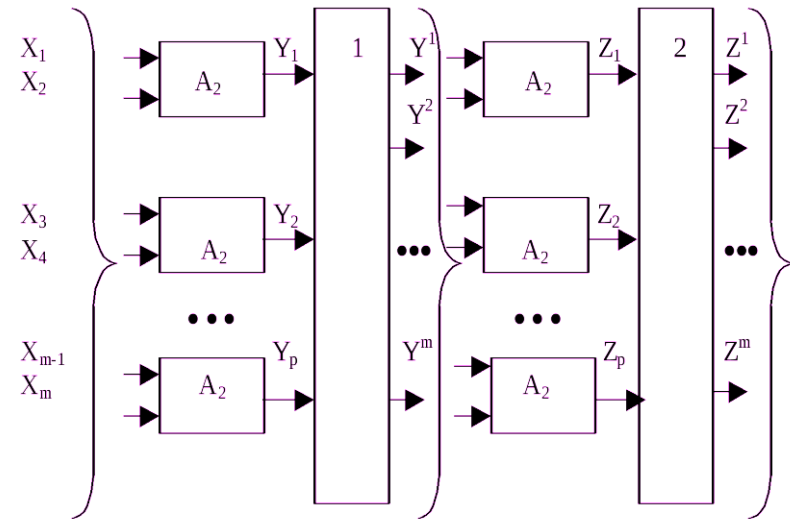
де корені характеристичних рівнянь лежать в одиничному колі,

$\{\varepsilon(k)\}$  - випадковий процес.

# Використані методи побудови моделей

$$y_t = C + \sum_{i=1}^R AR_i y_{t-i} + \sum_{j=1}^M MA_j \varepsilon_{t-j} + \varepsilon_t$$

APIKC(p,d,q)



Метод групового урахування аргументів

# Критерії для вибору моделі

---

- Інформаційний критерій Акаїке:

$$AIC = n \cdot \ln\left(\sum_{k=1}^n (y(k) - \hat{y}(k))^2\right) + 2p, p - \text{порядок регресії}$$

Критерії якості оцінок прогнозів:

- Середньоквадратична похибка

$$СКП = \sqrt{\frac{1}{n} \sum_{k=1}^n (y(k) - \hat{y}(k))^2}$$

- Середня абсолютна похибка у відсотках:

$$САПП = \frac{1}{n} \sum_{k=1}^n \frac{|y(k) - \hat{y}(k)|}{y(k)} \cdot 100\%$$

- Коефіцієнт Тейла

$$U = \frac{\sqrt{\frac{\sum_{k=1}^n (y(k) - \hat{y}(k))^2}{n}}}{\sqrt{\frac{\sum_{k=1}^n y^2(k)}{n}} + \sqrt{\frac{\sum_{k=1}^n \hat{y}^2(k)}{n}}}$$

# Критерії для вибору моделі. Адекватність моделі

---

Коефіцієнт детермінації:

$$R^2 = 1 - \frac{\hat{\varepsilon}'\hat{\varepsilon}}{(y - \mu)'(y - \mu)}$$

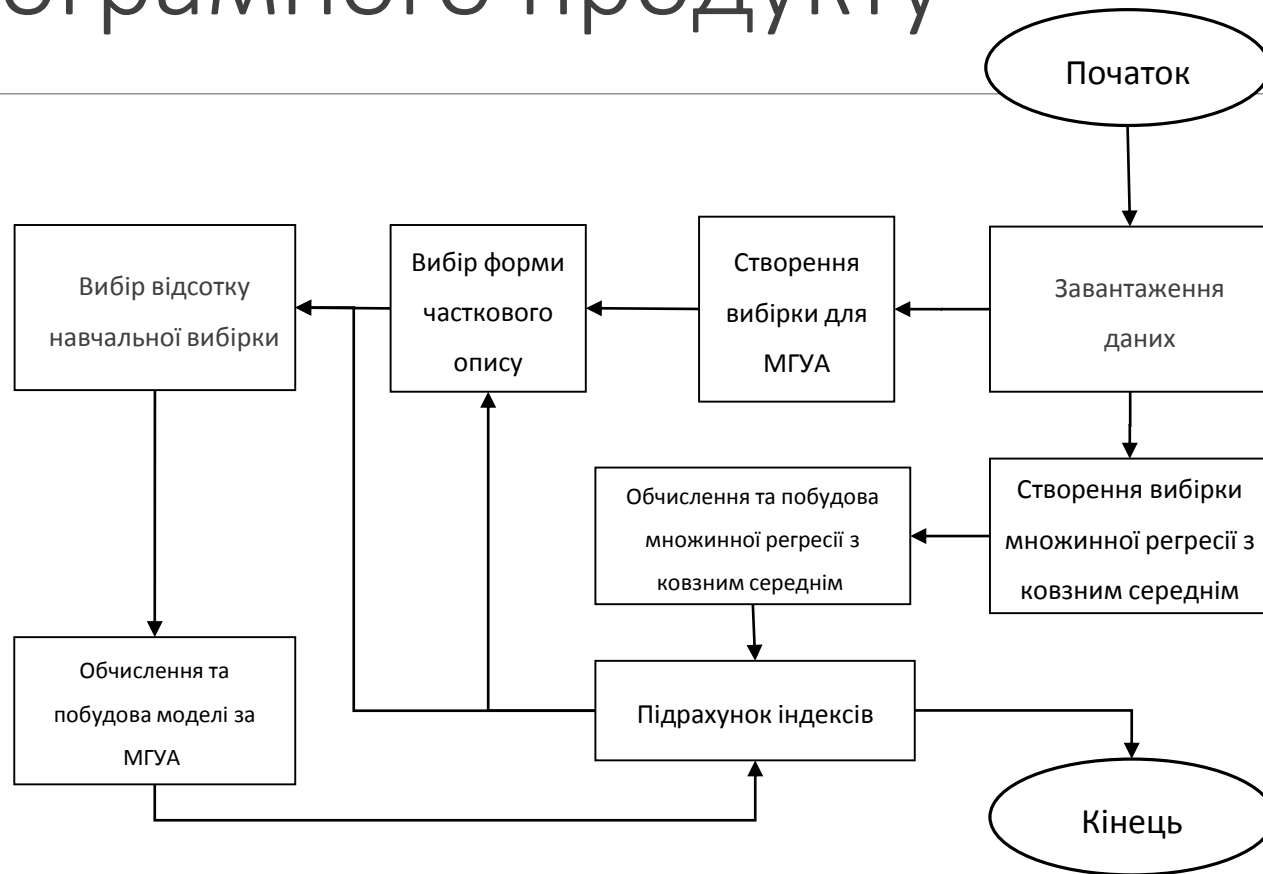
Сума квадратів похибок:

$$SSE = \sum_{k=1}^n e^2(k) = \sum_{k=1}^n (y(k) - \hat{y}(k))^2$$

Статистика Дарбіна-Уотсона:

$$DW = \frac{\sum_{t=2}^N (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^N \varepsilon_t^2}$$

# Блок-схема розробленого програмного продукту



# Розроблений програмний продукт

**Forecast analysis**

Info

	LEB	MYS	GNI	EYS	HDI
1	68.213	10.227	4373....	13.029	0.65
2	67.86	10.232	4663.81	13.182	0.65
3	68.287	10.237	5221.11	13.674	0.67
4	68.276	10.238	5566....	14.032	0.67
5	68.211	11.238	6160....	14.319	0.69
6	68.185	11.24	6951....	14.233	0.70
7	67.957	11.267	7196....	14.412	0.71
8	68.078	11.27	7734....	14.524	0.71
9	68.222	11.262	8399....	14.6	0.72
10	68.251	11.262	8697....	14.635	0.72
11	69.19	11.268	7350....	14.637	0.72
12	70.265	11.297	7738....	14.649	0.73
13	70.809	11.298	8089....	14.761	0.74

**HDI**

Life expectancy at birth

Expected Years of Schooling

Mean Years of Schooling

Gross National Income per capita

1990    1990

GMDH. 1-year forecast

$HDI_t = f(GNI_{t-1}, LEB_{t-1}, EYS_{t-1}, MYS_{t-1})$

$a_0 + a_1 * x_1 + a_2 * x_2$

Training Sample, %    Calculate

Open the window with the result

open file

open file

open file

open file

open file

Mixed regression    Upload Data

ARMA(3)

Mixed Regression with Moving average

of the sequence of 3 values:

$HDI_t = a_0 GNI_{t-1} + a_1 LEB_{t-1} + a_2 EYS_{t-1} + a_3 MYS_{t-1} + a_4 MA_3$

$HDI(t) = -0.0 * GNI(t-1) + 0.00088 * LEB(t-1) + 0.02047 * EYS(t-1) + 0.00134 * MYS(t-1) + 0.51544 * MA(3)$

Forecast	Akaike	MAPE	DW	Model	Theil	MSE	R-sq.	RSS
0.753	-106.976	0.365%	2.055	ARMA(3)	7e-06	1e-05	0.954	0.000124
0.749	-103.518	0.658%	1.841	GMDH(2,90%)	2.7e-05	3.9e-05	0.953	0.000591

# Результати роботи

Forecast	Akaike	MAPE	DW	Model	Theil	MSE	R-sq.	RSS
0.753	-106.976	0.365%	2.055	ARMA(3)	7e-06	1e-05	0.954	0.000124
0.754	-109.953	0.555%	1.669	GMDH(2,50%)	1.8e-05	2.6e-05	1.142	0.000385
0.75	-107.724	0.556%	1.768	GMDH(2,60%)	2.1e-05	3e-05	1.072	0.000446
0.753	-113.003	0.491%	2.019	GMDH(2,70%)	1.5e-05	2.1e-05	1.055	0.000314
0.749	-116.833	0.495%	2.205	GMDH(2,80%)	1.1e-05	1.6e-05	0.972	0.000243
0.748	-116.474	0.463%	2.273	GMDH(2,90%)	1.2e-05	1.7e-05	0.985	0.000249
Forecast	Akaike	MAPE	DW	Model	Theil	MSE	R-sq.	RSS
0.753	-106.976	0.365%	2.055	ARMA(3)	7e-06	1e-05	0.954	0.000124
0.749	-114.735	0.514%	2.159	GMDH(3,90%)	1.3e-05	1.9e-05	0.925	0.00028
0.748	-115.981	0.499%	2.555	GMDH(3,80%)	1.2e-05	1.7e-05	0.975	0.000257
0.75	-114.047	0.486%	2.496	GMDH(3,70%)	1.4e-05	2e-05	1.047	0.000293
0.748	-108.244	0.527%	1.403	GMDH(3,60%)	2e-05	2.9e-05	0.925	0.000431
0.748	-65.512	1.816%	0.356	GMDH(3,50%)	0.000348	0.000496	2.662	0.007441

# Моделі, побудовані програмою

---

$$\text{HDI}(t) = 5e-06 * \text{GNI}(t-1) + 0.00088 * \text{LEB}(t-1) + 0.02047 * \text{EYS}(t-1) + 0.00134 * \text{MYS}(t-1) + 0.51544 * \text{MA}(3)$$

$$\text{МГУА}(\varphi(x, y) = a_0 + a_1x + a_2y, 80\%):$$

$$\begin{aligned} \text{HDI} = & -0.0 + (0.113) * (-0.011 + (0.331) * (0.091 + (0.0) * (\text{LEB}) + (0.007) * (\text{MYS})) + (0.684) * (0.044 + (0.0) * (\text{LEB}) + (0.047) * (\text{GNI}))) + \\ & + (0.887) * (0.005 + (0.164) * (0.091 + (0.0) * (\text{LEB}) + (0.007) * (\text{MYS})) + (0.843) * (0.132 + (0.003) * (\text{MYS}) + (0.043) * (\text{GNI}))) \end{aligned}$$

$$\text{МГУА}(\varphi(x, y) = a_0 + a_1x + a_2y + a_3xy, 80\%):$$

$$\begin{aligned} \text{HDI} = & 0.028 + ((0.677) * (-0.043 + ((0.002) * (\text{MYS})) + ((0.033) * (\text{GNI})) + (0.0) * (\text{MYS} * \text{GNI}))) + ((0.242) * (3.558 + ((-0.201) * (\text{GNI})) + ((-0.331) * (\text{EYS})) + (0.023) * (\text{GNI} * \text{EYS}))) + (0.059) * (-0.043 + ((0.002) * (\text{MYS})) + ((0.033) * (\text{GNI})) + (0.0) * (\text{MYS} * \text{GNI})) * 3.558 + ((-0.201) * (\text{GNI})) + ((-0.331) * (\text{EYS})) + (0.023) * (\text{GNI} * \text{EYS})) \end{aligned}$$



# Висновки

---

- У даній дипломній роботі виконано аналіз та моделювання індексу людського розвитку як нестационарного процесу.
- Побудовано моделі для короткострокового прогнозування значення індексу людського розвитку.
- Ковзне середнє третього порядку у регресійній моделі може говорити про те, що реформування у сферах охорони здоров'я та освіти дають свої результати через 3 і більше років.
- За результатами дослідження, для покращення людського розвитку в Україні необхідно популяризувати максимально тривалу освіту.
- Обидва методи для побудови моделі дали дуже хороші результати.

---

Дякую за увагу