

Методи машинного навчання для оцінки віку серця за даними R-R інтервалів

Бакалаврська робота
Фрідман Рон
КА-23 ІПСА

Науковий керівник:

доцент, к.т.н.

Тимошенко Юрій Олександрович

Об'єкт предмет і мета дослідження:

Мета дослідження – розробка системи для оцінки віку серця людини за даними R-R інтервалів, які отримано після анотації та фільтрації шумів електрокардіограми.

Об'єктом дослідження є персональне здоров'я серця людини і моніторинг його стану.

Предметом дослідження - методи машинного навчання для прогнозування віку серця людини.

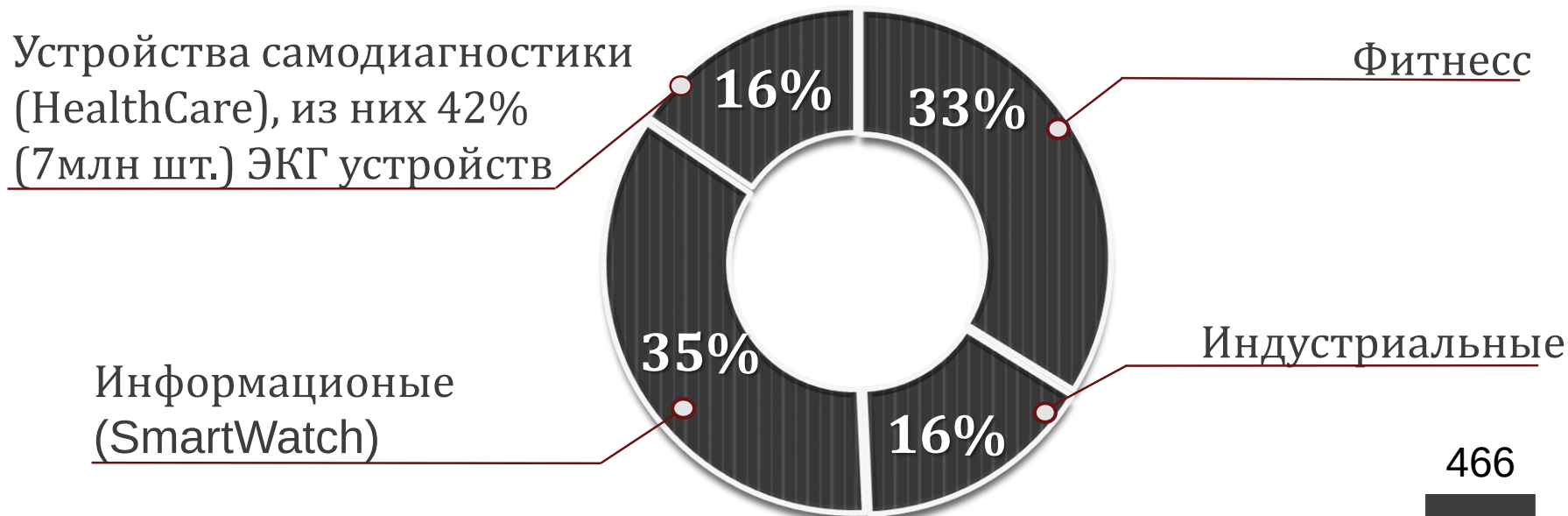
Завдання :

Завданнями дипломної роботи є:

- Необхідно з даних кардіоінтервалограми виділити ознаки за якими будуть побудовані моделі прогнозування біологічного віку людини.
- Необхідно побудувати моделі прогнозування віку серця на основі отриманих ознак, порівняти між собою та проаналізувати їх точність в залеженості від кількості біомаркерів.
- Надати змогу користувачеві за різними типами файлів ЕКГ сигналу на основі запропонованих алгоритмів оцінити біологічний вік.

Ринок smart clothes та wearable devices має експоненційне зростання

Носимые устройства



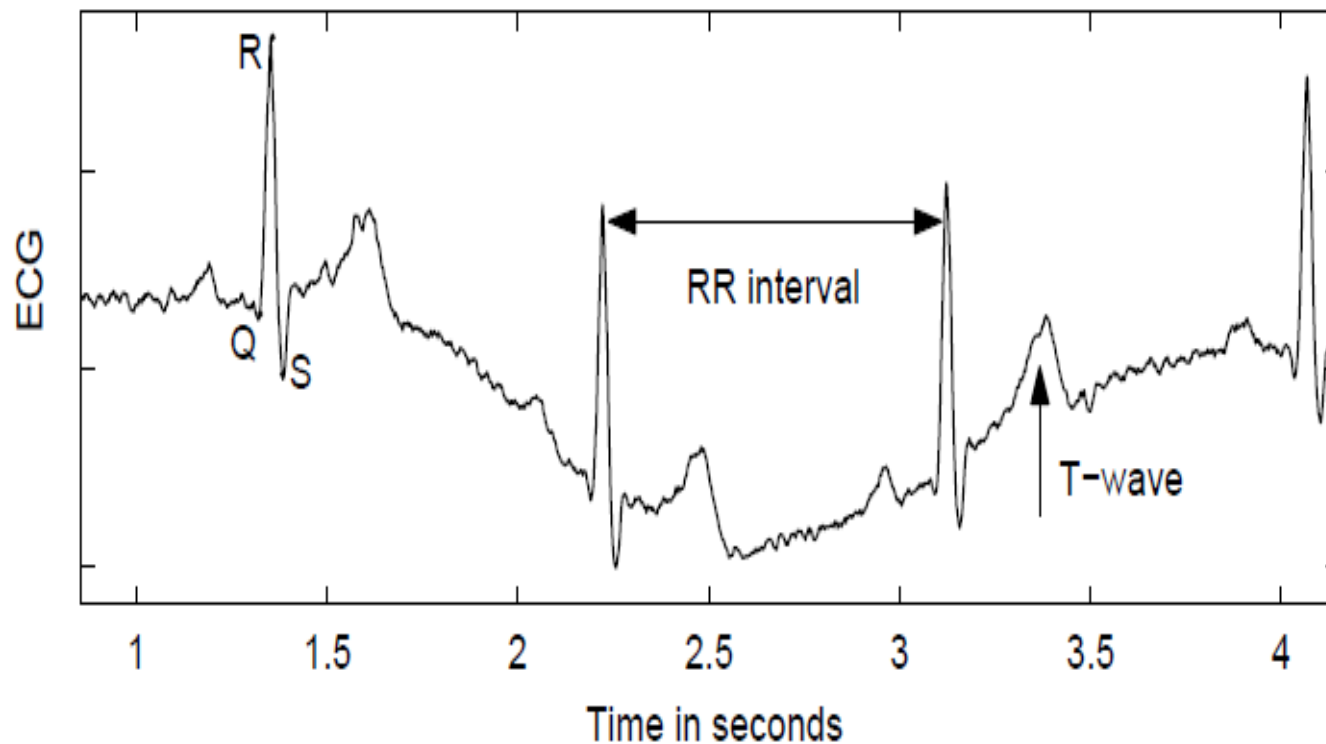
Сейчас - **100млн**
носимых устройств,
2018 – **466 млн**



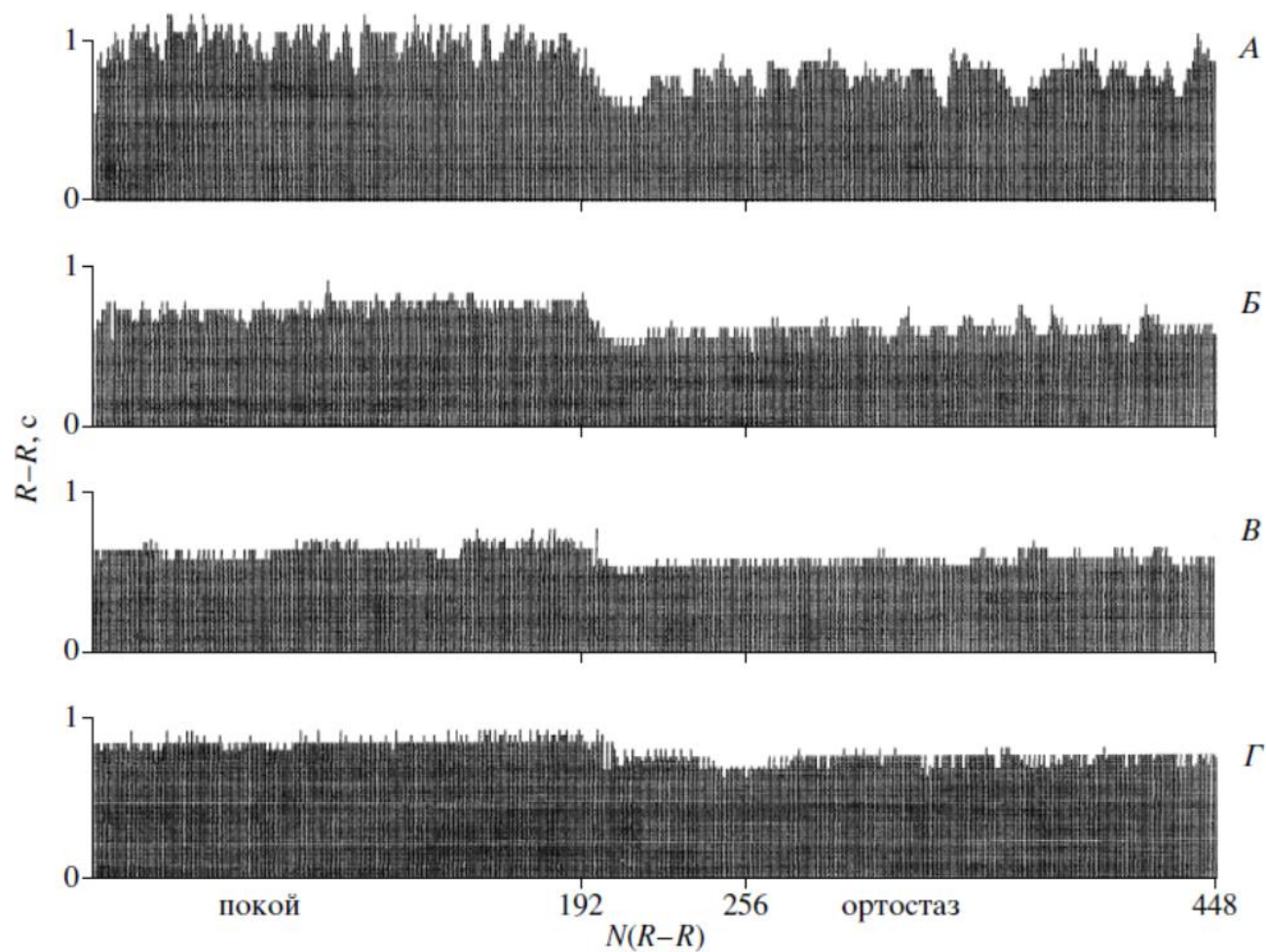
Що ж таке біологічний вік людини?

- Біологічний вік (БВ) - це показник рівня зносу структури та функції певного елемента організму, групи елементів і організму в цілому, виражений в одиницях часу шляхом співвіднесення значень заміряних індивідуальних біомаркерів з еталонними середнепопуляційними кривими залежностей змін цих біомаркерів від календарного віку.

Кардіограма і сьогодні залишається важливим фактором при діагностуванні хвороб ССС



Кардіогігнал –це часовий ряд, до якого в певному сенсі можливе застосування методів машинного навчання

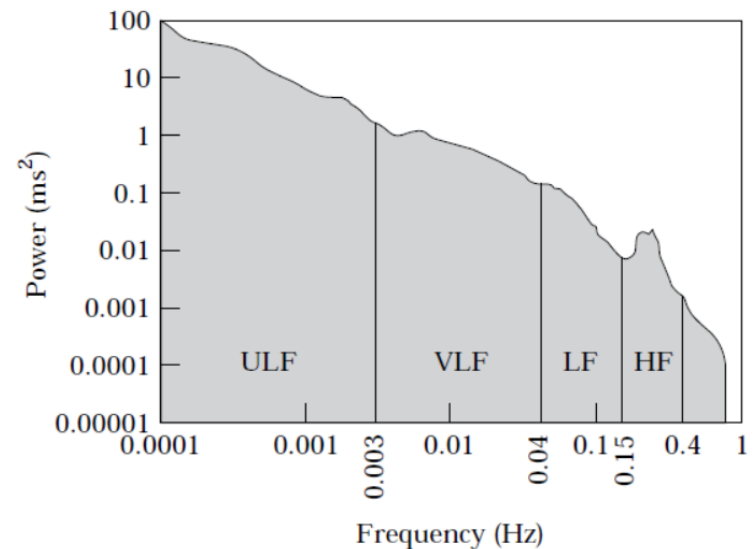
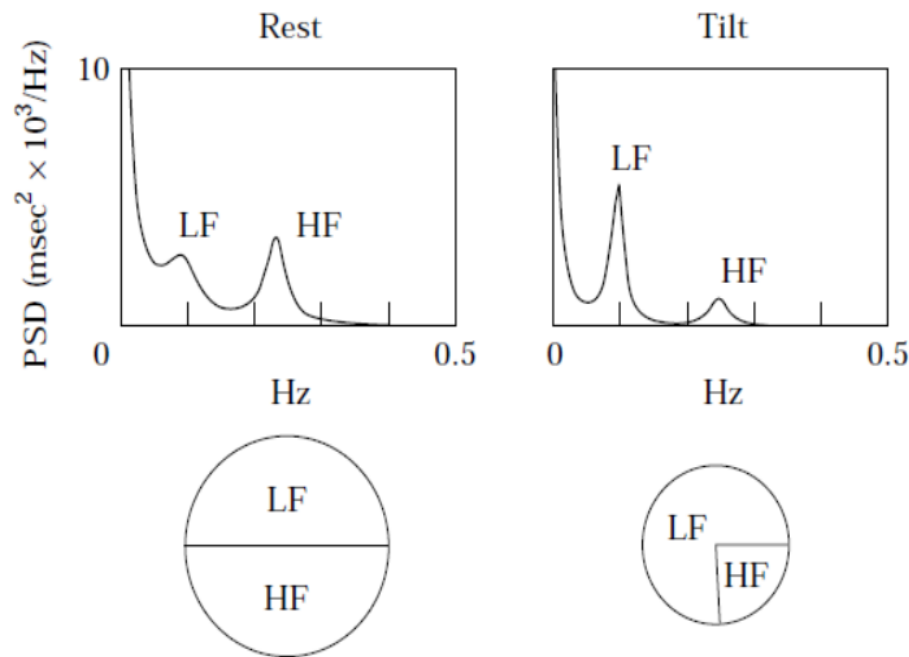


Кардіоінтервалографія – наступний крок в розвитку пульсової діагностики

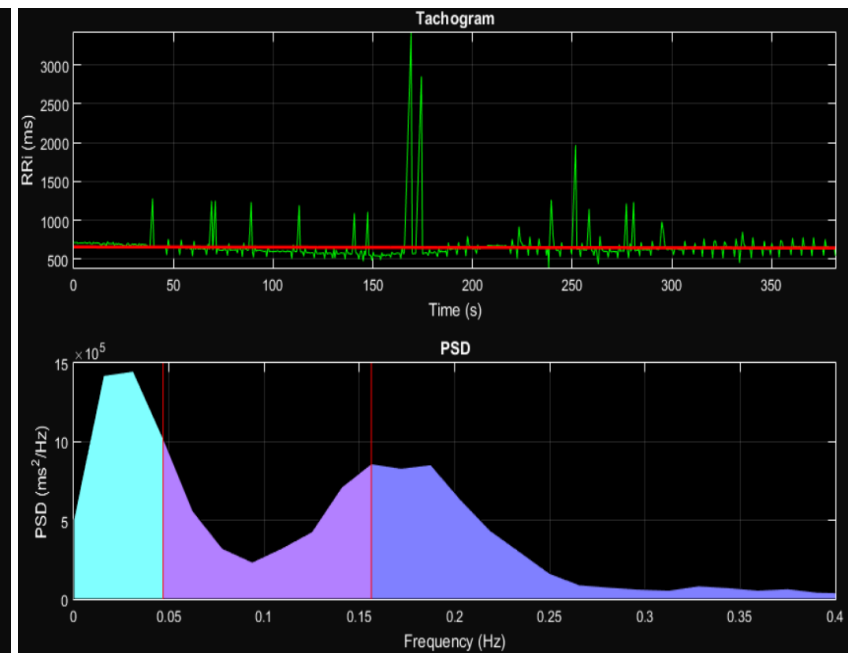
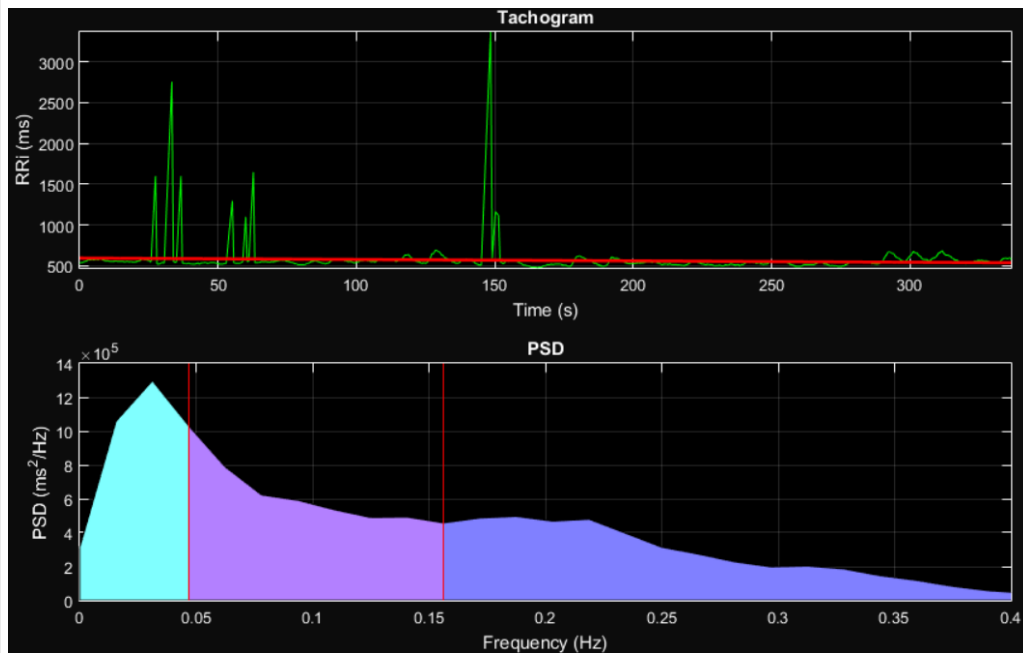
Table 3 Approximate correspondence of time domain and frequency domain methods applied to 24-h ECG recordings

Time domain variable	Approximate frequency domain correlate
SDNN	Total power
HRV triangular index	Total power
TINN	Total power
SDANN	ULF
SDNN index	Mean of 5 min total power
RMSSD	HF
SDSD	HF
NN50 count	HF
pNN50	HF
Differential index	HF
Logarithmic index	HF

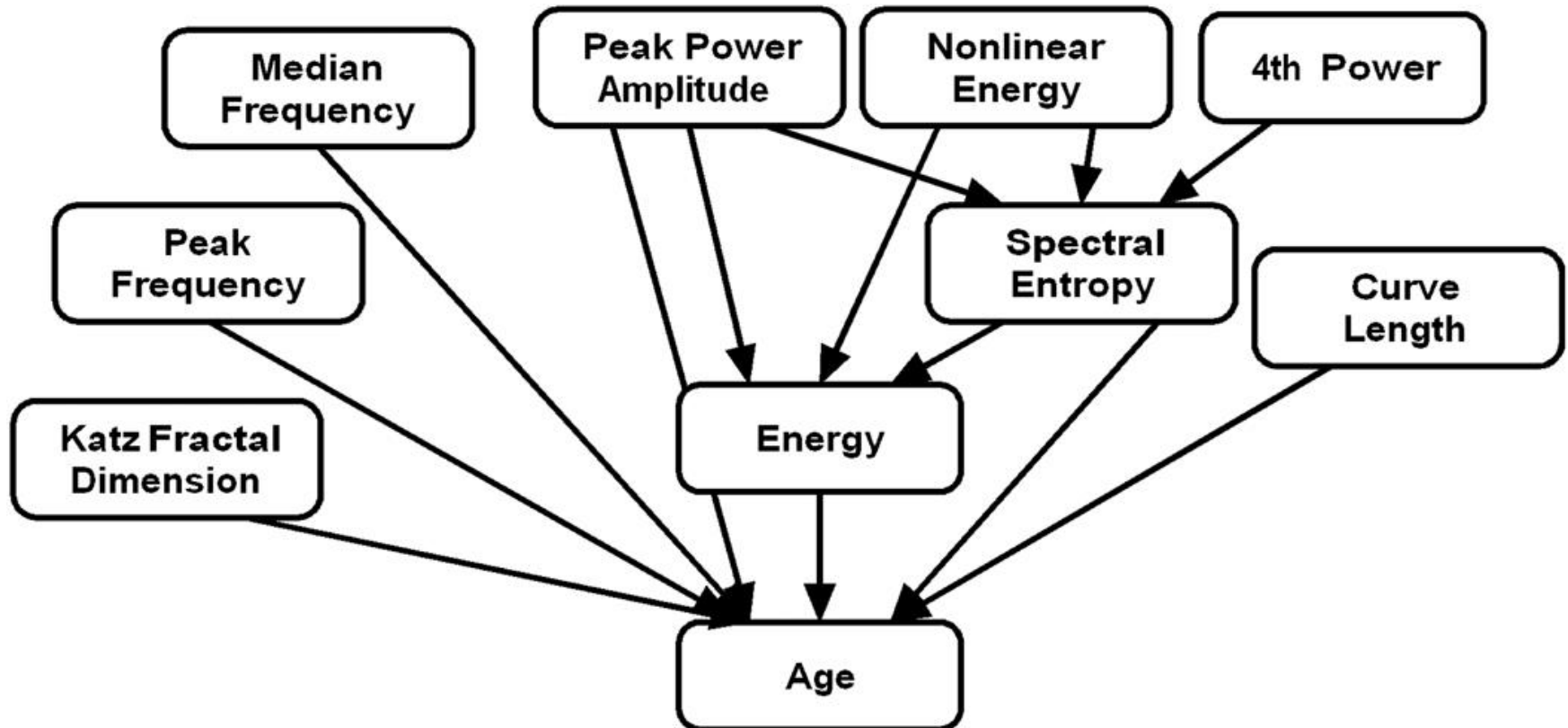
Новаторський підхід полягає в застосуванні спектральних методів



Порівняння показників HF LF на реальних прикладах



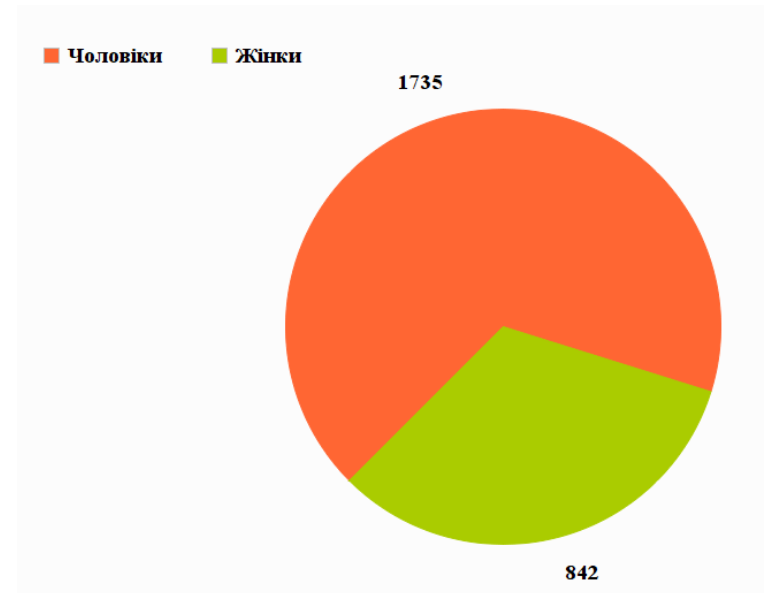
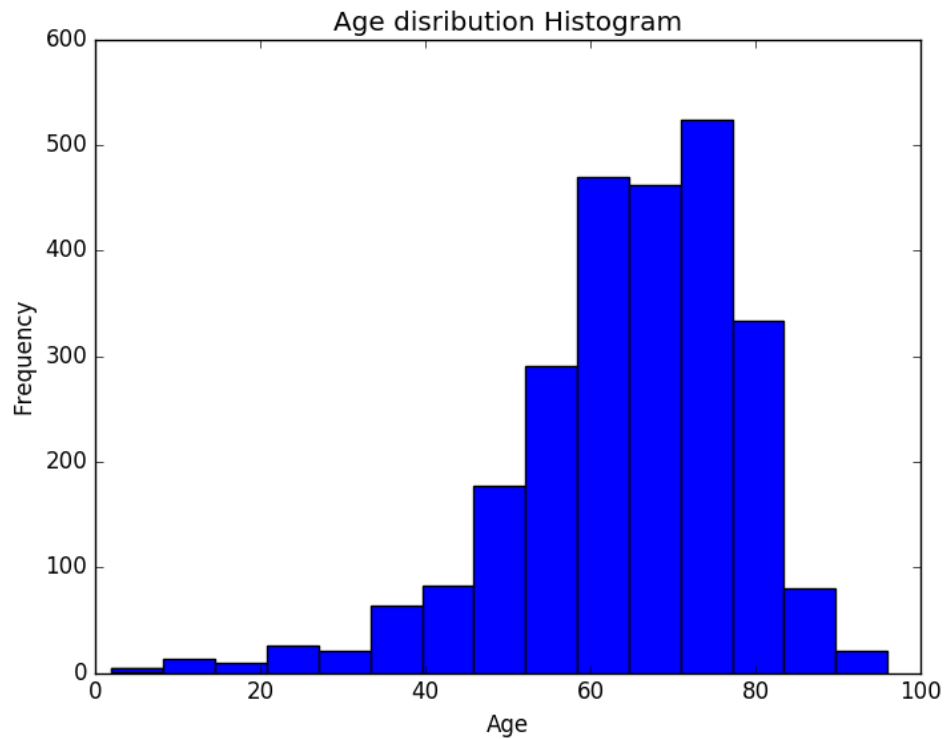
Побудова features matrix на основі отриманих індексів



Послідовність застосування методів Data mining



Опис даних



id90012as1 * x											
	A	B	C	D	E	F	G	H	I	J	K
1	RMSSD	SDNN	TP	SI	LF	HF	VLF	HF_LF	pNN50	Mean	AGE
2	34,09	43,63	1488,7	180	474,11	627,37	387,22	0,76	12,21	877	19

Критерії за якими порівнювались моделі

]

$$R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

де $\hat{y}_t$ – пораховані значення,

y_i – фактичні значення,

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$\text{САП} = \frac{1}{n} \sum_{t=0}^n (\hat{y}_t - y_t)$$

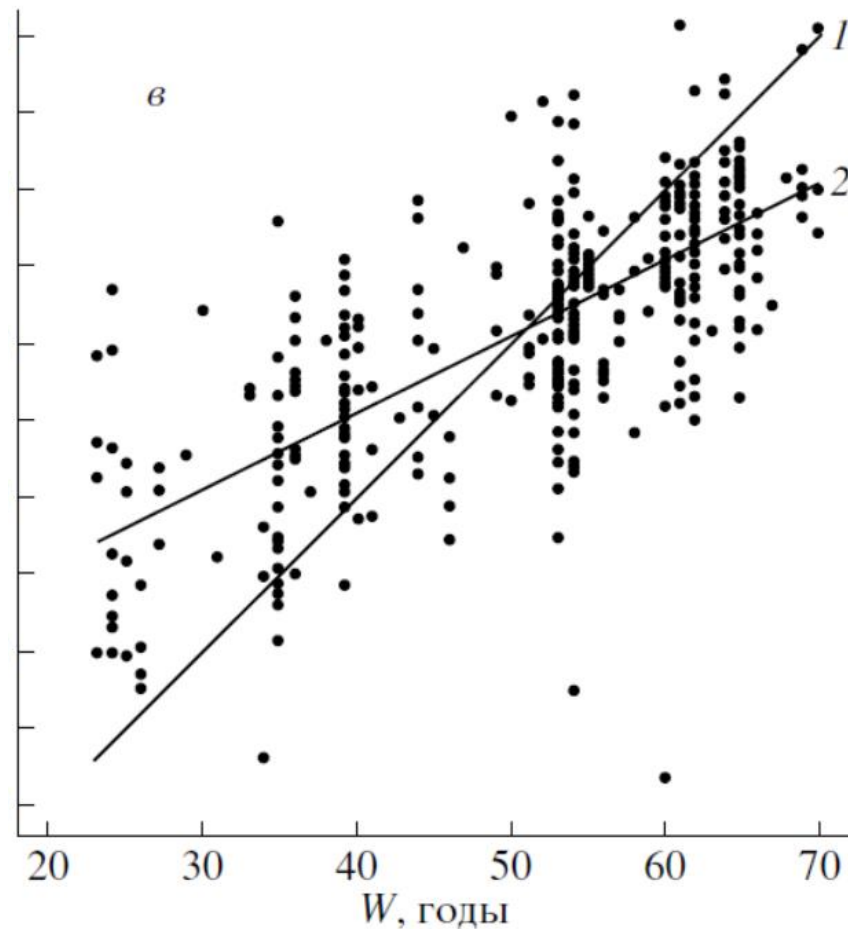
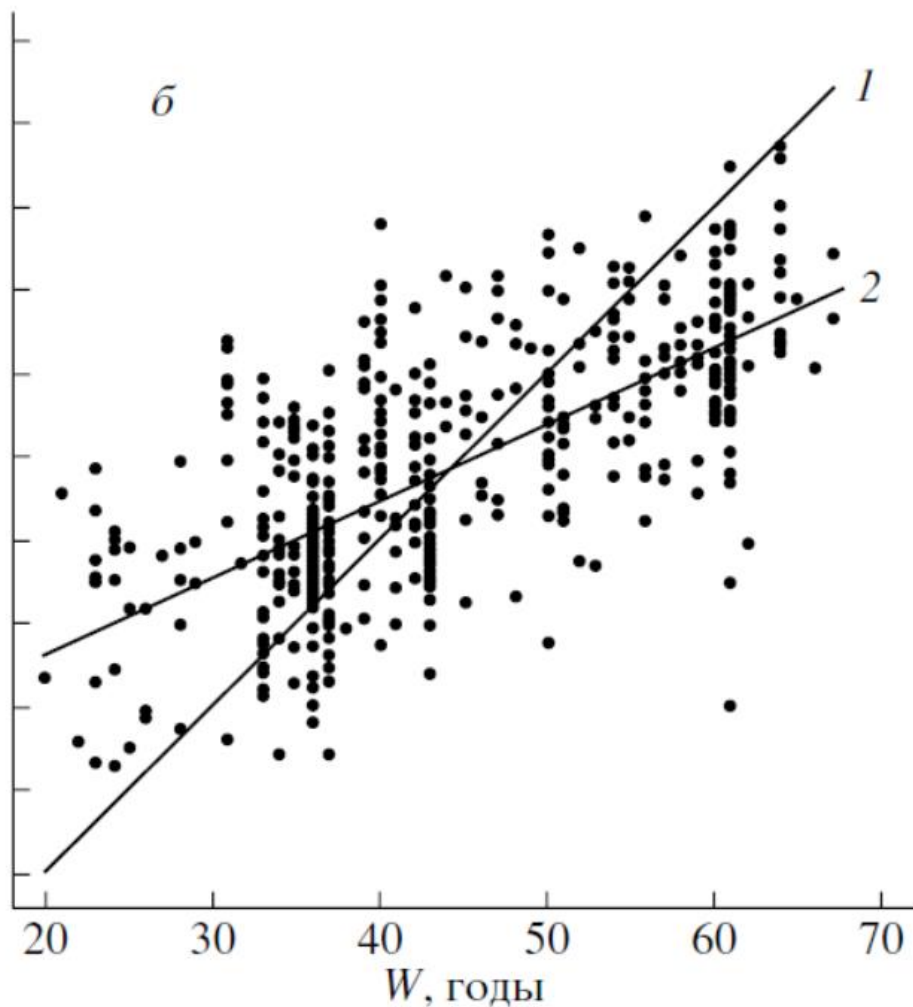
$$\text{САПП} = \frac{1}{n} \sum_{t=0}^n \frac{(\hat{y}_t - y_t)}{y_t} * 100\%$$

Регресійна модель

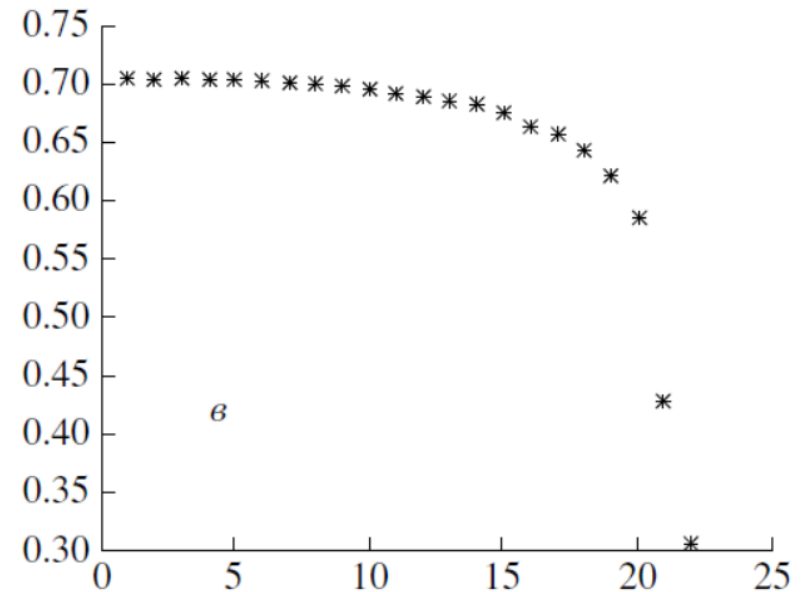
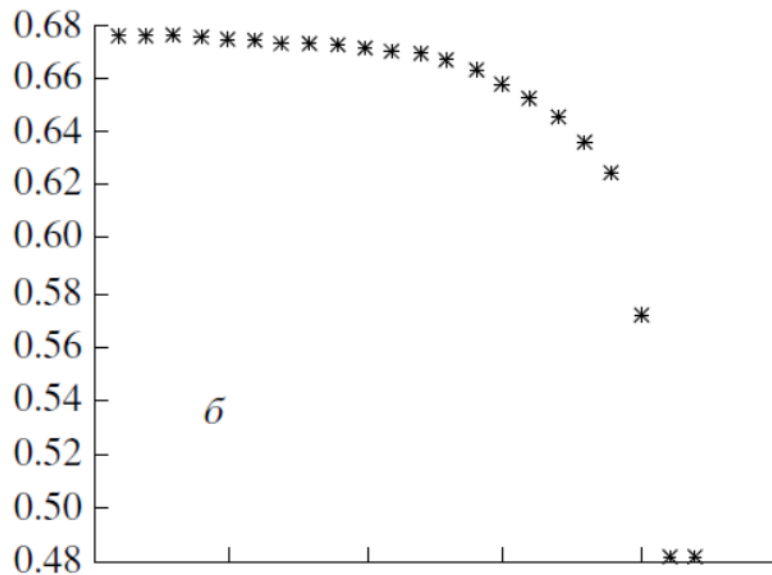
№	Назва параметрів	Константи лінійної моделі V_j			Середньоквадратичне відхилення σ_j параметрів
		загальна вибірка	жіноча вибірка	чоловіча вибірка	
1	ЧСС (с)	-6.10	-5.14	-5.83	13.85
2	СКО (л)	-1.58	-1.49	-1.93	0.02
3	P3(л)	-0.55	-1.09	-0.28	79.10
4	P3(с)	-1.81	-1.52	-2.07	47.86
5	P4(с)	42491,00	0.52	29252,00	87.12
6	РА	19360,00	18295,00	0.69	43831,00
7	ФС	24504,00	14611,00	20149,00	0.82
8	ЭС	41275,00	0.27	42492,00	0.35
9	K2	17564,00	24108,00	33635,00	0.08
10	A23	47.68	53.00	39.00	-

Значення констант лінійної моделі з десятима параметрами (біологічними маркерами), відібраними методом виключення

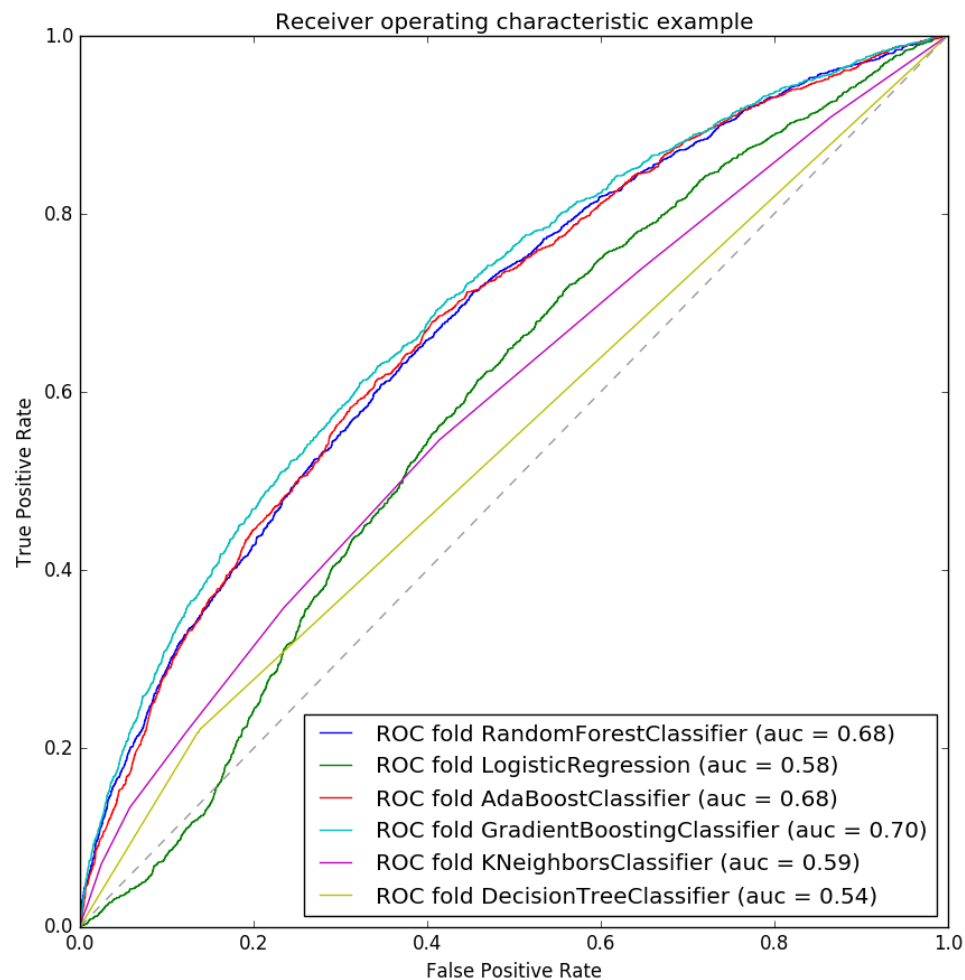
Результати регресійної моделі



Результати оцінки віку серця з застосуванням регресії



Результати роботи методів машинного навчання



Висновки

- В результаті виконання дипломної роботи розроблено програму для аналізу інтервалів та підсистеми, основаної на методах машинного навчання для якісного показника стану здоров'я серця, а саме – біологічного віку людини. Було порівняно різні підходи, як із застосуванням лінійної регресії так і з застосуванням сучасних модифікацій алгоритму градієнтного бустингу та випадкового лісу.
- Також на основі еталонної вибірки біологічних маркерів було отримано константи моделі біологічного віку для регресійної моделі. Адекватність моделі зазвичай визначається якістю апроксимації біологічного віку. Для кількісної оцінки використовується коефіцієнт кореляції між біологічним і календарним віком.

Подальші дослідження

- Щодо подальших досліджень, є сенс розвиватися в напрямку відбору найбільш інформативних ознак для побудови більш точної моделі оцінки віку серця людини за даними R-R інтервалів.

Дана робота була представлена на міжнародних конференціях

Статті

- Саган В.Ю., Фридман Р. Применение сетей байеса для анализа ЭКГ// Open Semantic Technologies for Intelligent Systems, Minsk, Belarus: Матеріали VI міжнародной научно-технической конференции – 2016. – с.579 - 582.
- Fridman R. Application of machine learning methods for heart pathologies detection using ECGs // SCCS-2015, Erfurt, Germany. – 2015. – с.36 - 41.

Участь у конференціях

- Fridman R., Sagan V. Approach to automated ECG analysis // Системний аналіз та інформаційні технології: Матеріали XII міжнародної науково-технічної конференції «Системний аналіз та інформаційні технології» (22–25 червня 2015 р. м.Київ). – К.:НТУУ «КПІ», 2015.
- Fridman R., Sagan V. Approach to ischemia detection using artificial neural network// Обчислювальний інтелект (результати, проблеми, перспективи): Матеріали III-ої міжнародної науково-практичної конференції 12-15 травня 2015 року, Україна, Київ-Черкаси. – 2015. – с. 155-157.
-

Дякую за увагу!