Wybrane metody matematyczne w zastosowaniach na przykładzie projektu „Domestic – domowy asystent osób starszych i chorych”

dr hab. inż. Jerzy Wtorek, prof. PG

dr inż. Artur Poliński

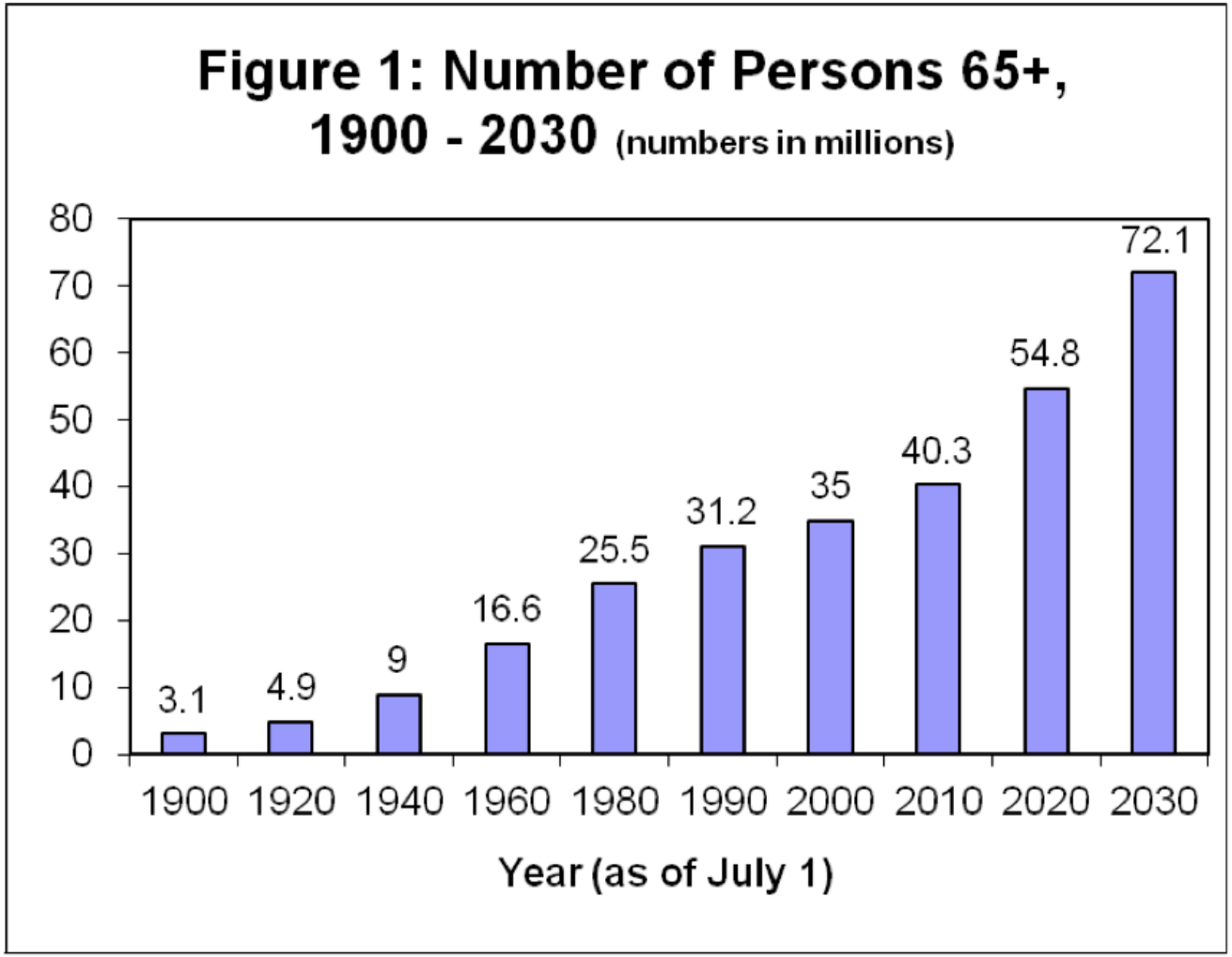
Katedra Inżynierii Biomedycznej, Politechnika Gdańska

Uzasadnienie

Przewiduje się, że zarówno w USA jak i Unii Europejskiej, pomijając problemy możliwe problemy ekonomiczne, zabraknie osób opiekujących się starszymi, chorymi i inwalidami.

Uzasadnienie

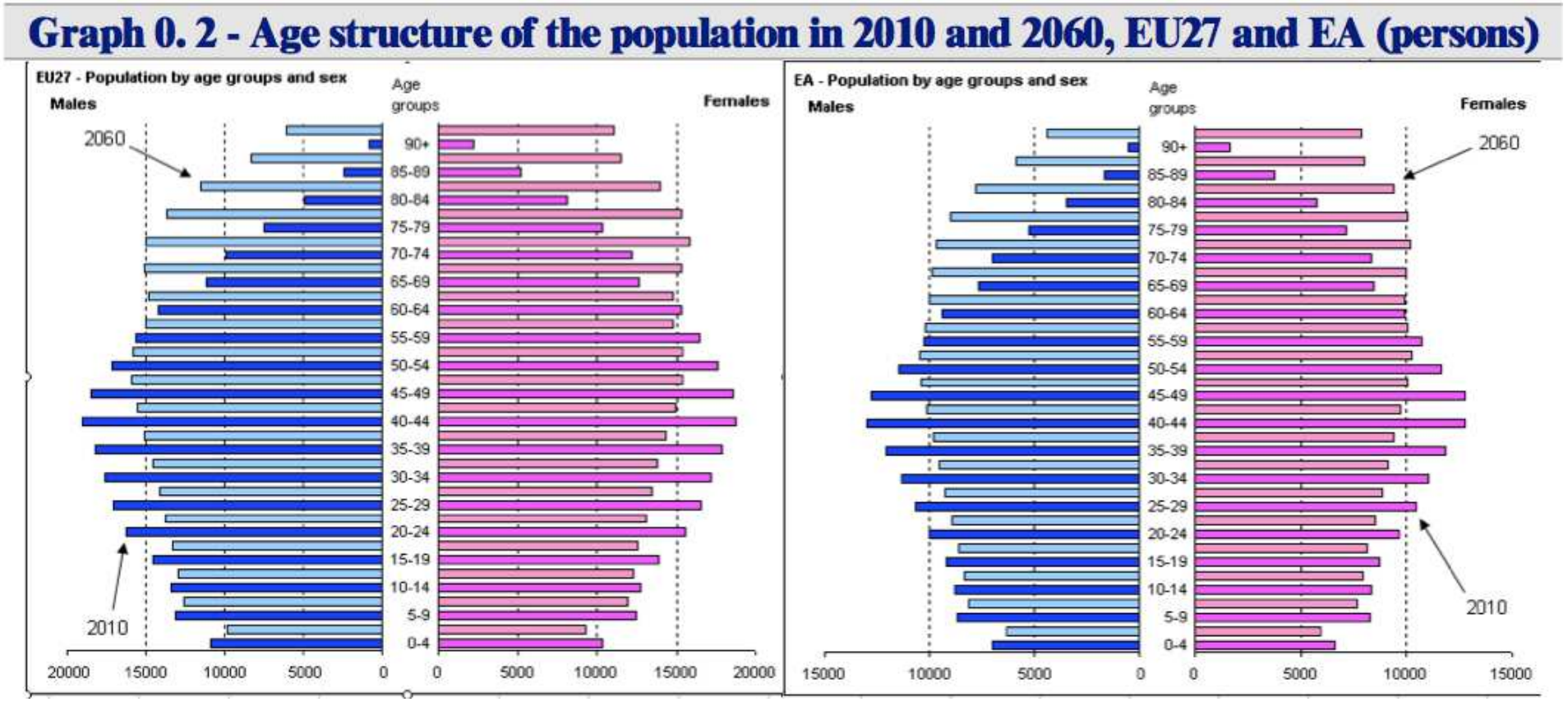
USA



Źródło: Administration on Aging, A Profile of Older Americans: 2011. Administration on Aging, U.S. Department of Health & Human Services, (2011).

Uzasadnienie

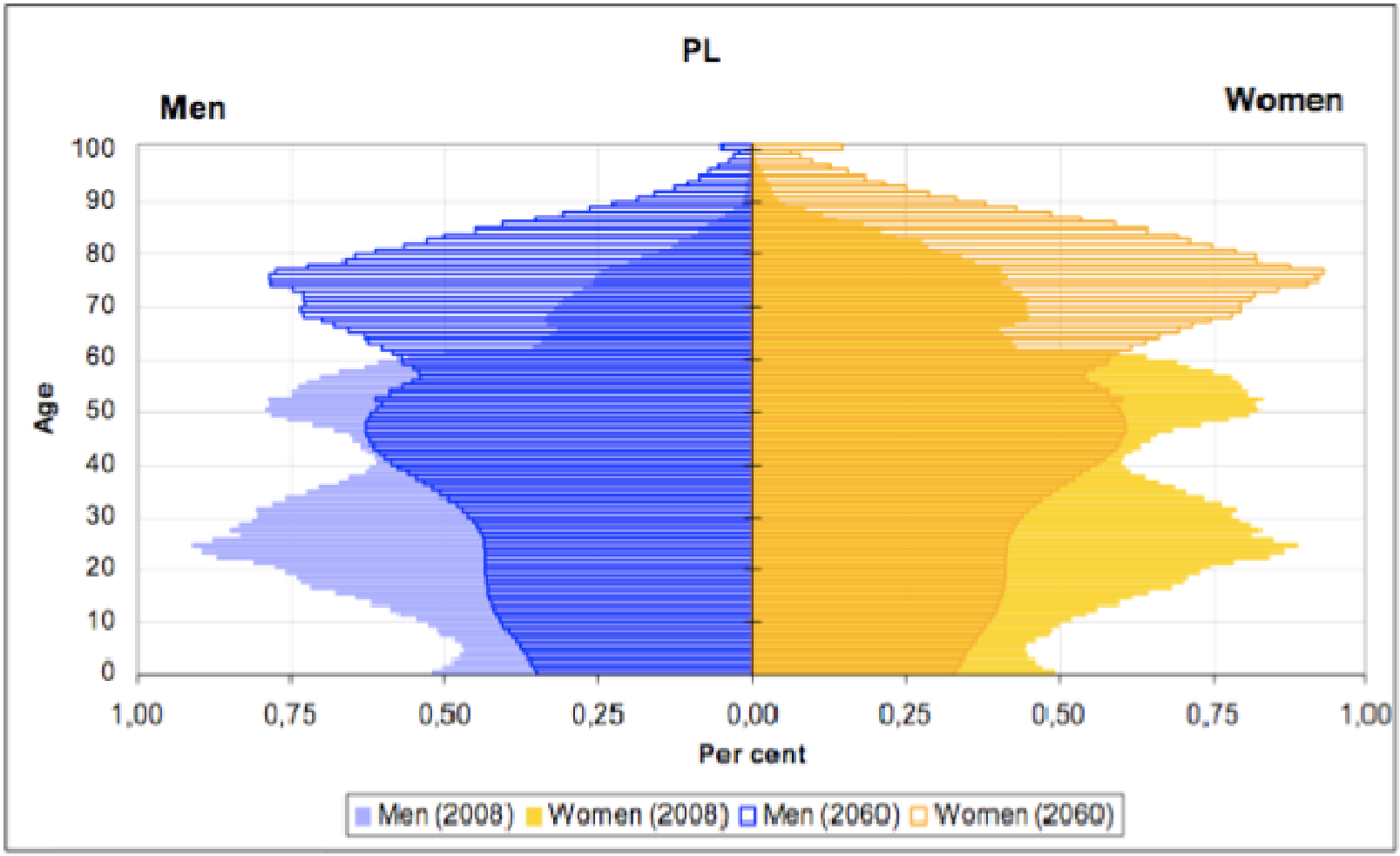
EU



Źródło: European Union, The 2012 Ageing Report: Underlying Assumptions and Projection Methodologies. EUROPEAN ECONOMY 4, European Union, (2011).

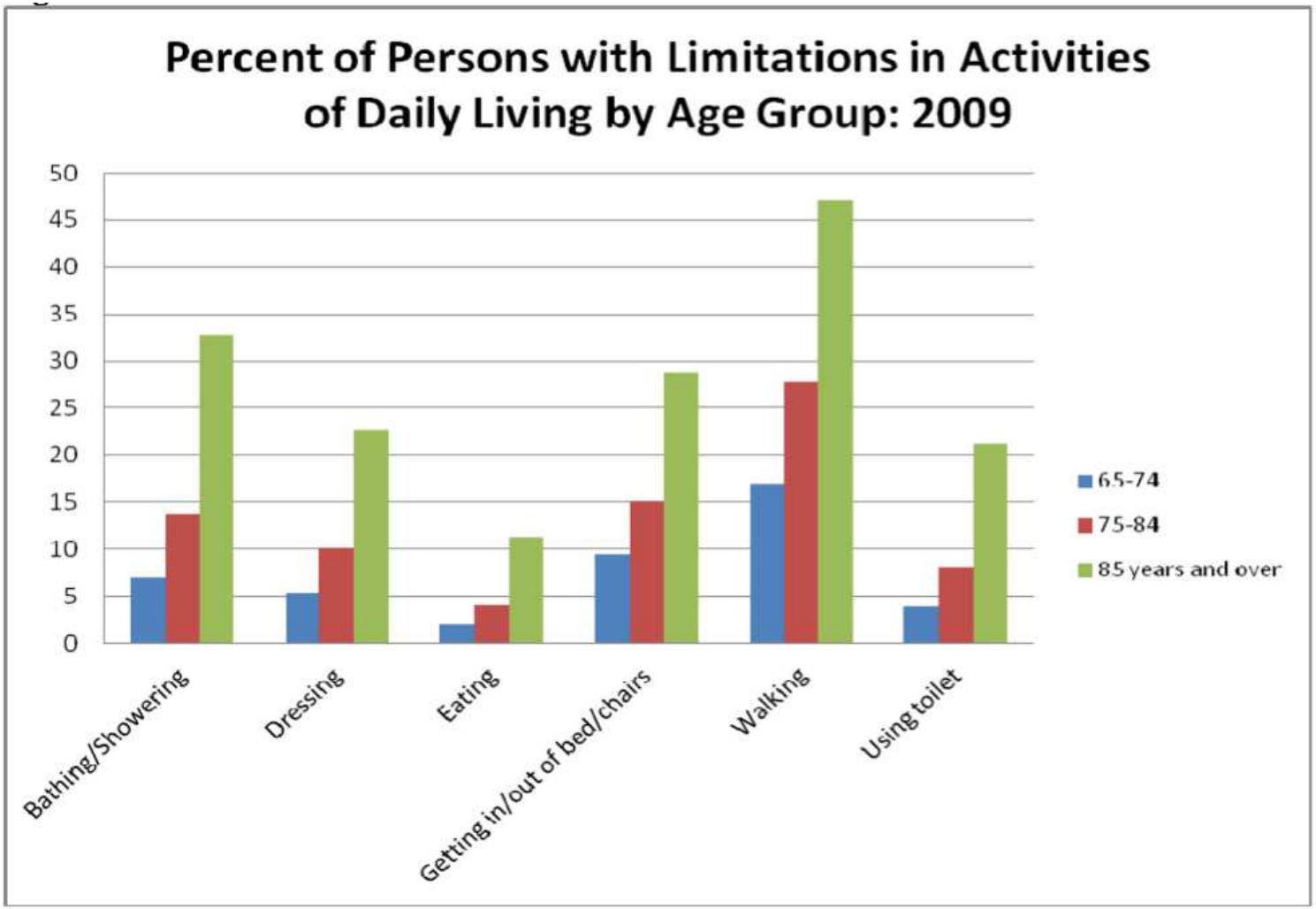
Uzasadnienie

PL



Źródło: Eurostat, EUROPOP2008, 2008

Uzasadnienie



Źródło: Administration on Aging, A Profile of Older Americans: 2011. Administration on Aging, U.S. Department of Health & Human Services, (2011).

Uzasadnienie

Znacznie dłuższy czas życia kreuje nowe problemy, a przynajmniej zmienia się skala niektórych, np. przewlekłe choroby, ograniczenie niektórych funkcji, itp.

Jednocześnie wzrastają koszty leczenia i zapotrzebowanie na profesjonalnych opiekunów.

Wstępne wnioski:

- populacja 65+: około 30% całej populacji w EU/USA w 2060
- Populacja 50+: istotny wzrost wydatków na opiekę zdrowotną
- Problemy ekonomiczne (koszty opieki, koszty opieki zdrowotnej)
- Zasoby ludzkie (ograniczona liczba w wieku produkcyjnym, ograniczona liczba opiekunów formalnych i nieformalnych)

Każde rozwiązanie redukujące konieczność udziału opiekuna i poprawiające jakość życia jest

OCZEKIWANE

Szczególnie oczekiwane są rozwiązania

INTELIGENTNE

Technologie wspomagające

Technologie wspomagające to te, które pomogą:

- Tobie i mi,
- Osobom z upośledzeniami,
- Osobom chorym,
- Osobom starszym

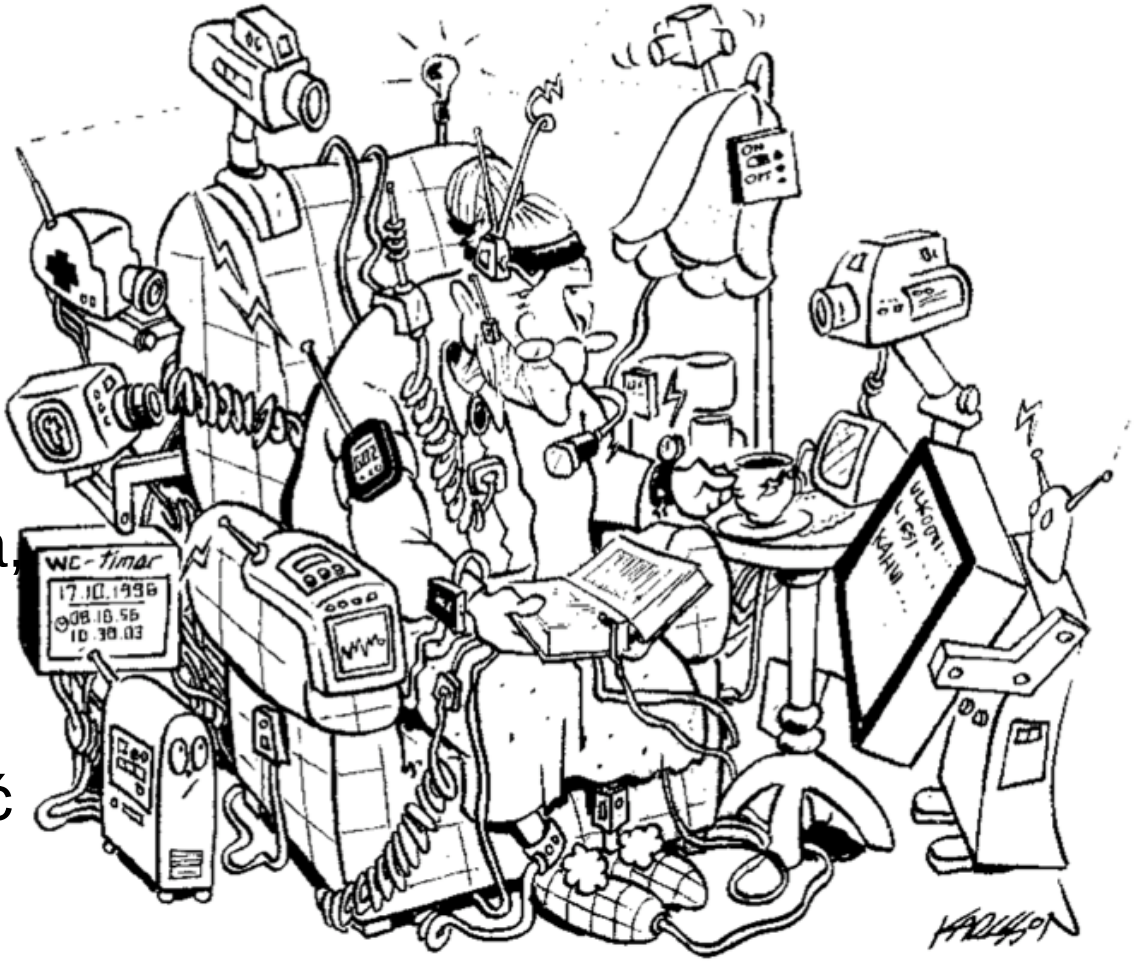
wykonywać czynności, które są trudne lub niemożliwe do wykonania

Źródło: IEEE Spectrum, September, 2012

Technologie wspomagające

Wymagania:

- Przyjazne w użyciu,
- Łatwe do zrozumienia,
- Ukryte,
- Tanie,
- O niskich kosztach użycia,
- Niezawodne,
- Bezobsługowe,
- Zapewniające prywatność
- ……



Źródło: R.Karlsoon, Printed in Helsingin Sanomat, October 18, 1996.

Spersonalizowana opieka medyczna

Rozwiązania ICT wspomagające opiekę zdrowotną:

- Typowe przyrządy medyczne (e.g. pulsoksymetry, ciśnieniomierze, glukometry, itd.)

- Dobre samopoczucie/stan fizyczny (np. wagi, mierniki aktywności, przyrządy ćwiczeń, inne)

- Specjalizowane przyrządy (np. zdalne rejestratory EKG)

- Aplikacje do zarządzania Indywidualnymi Kartami Pacjenta (Personal Health Records - PHR)

- Aplikacje do ćwiczeń mentalnych (demencja, Alzheimer, asysta w przemieszczaniu, układy przypominające, itp.

-Zarządzanie procesem leczenia,

- inne.

POLITECHNIKA GDAŃSKA

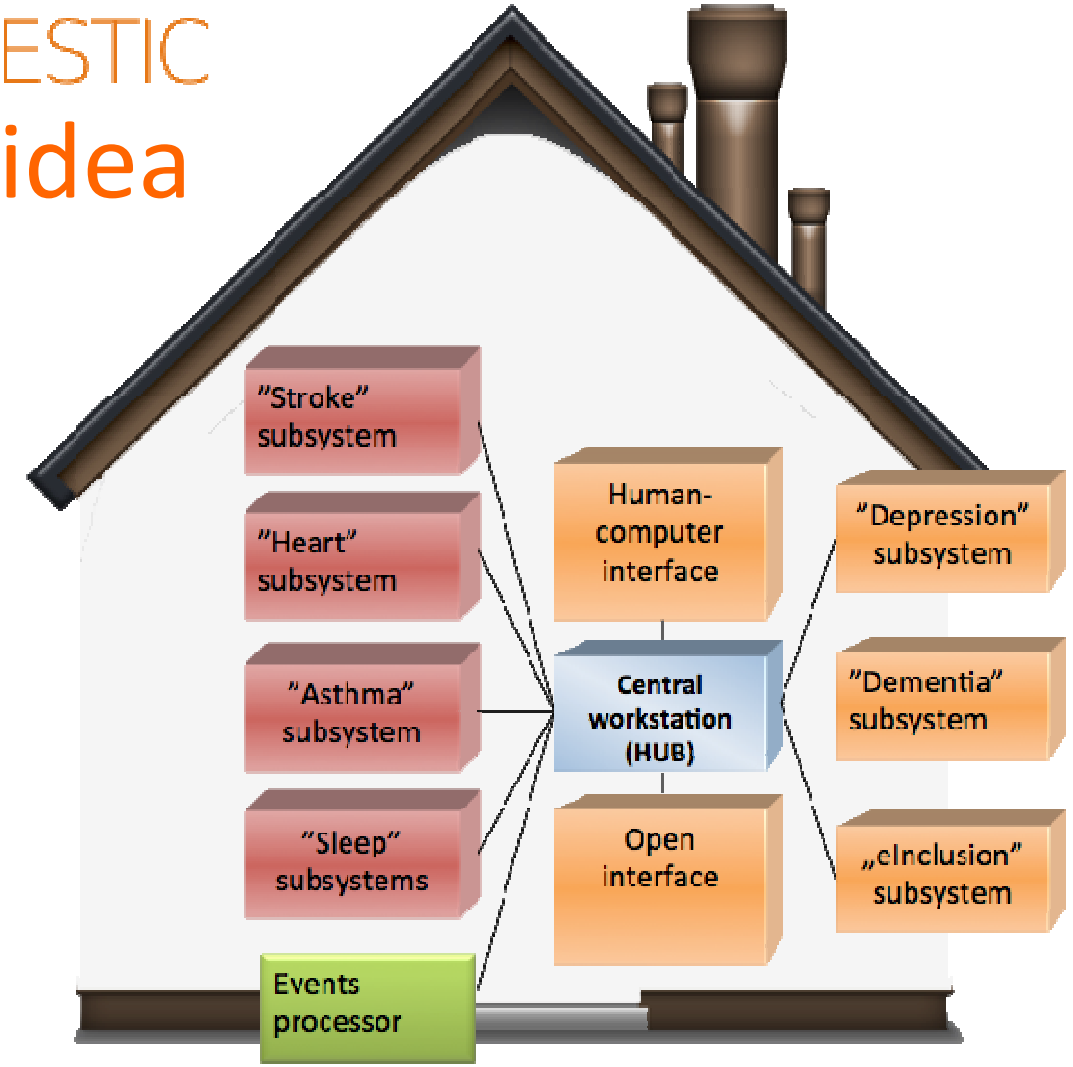
Zaprojektowane w KIB, PG

DOMESTIC

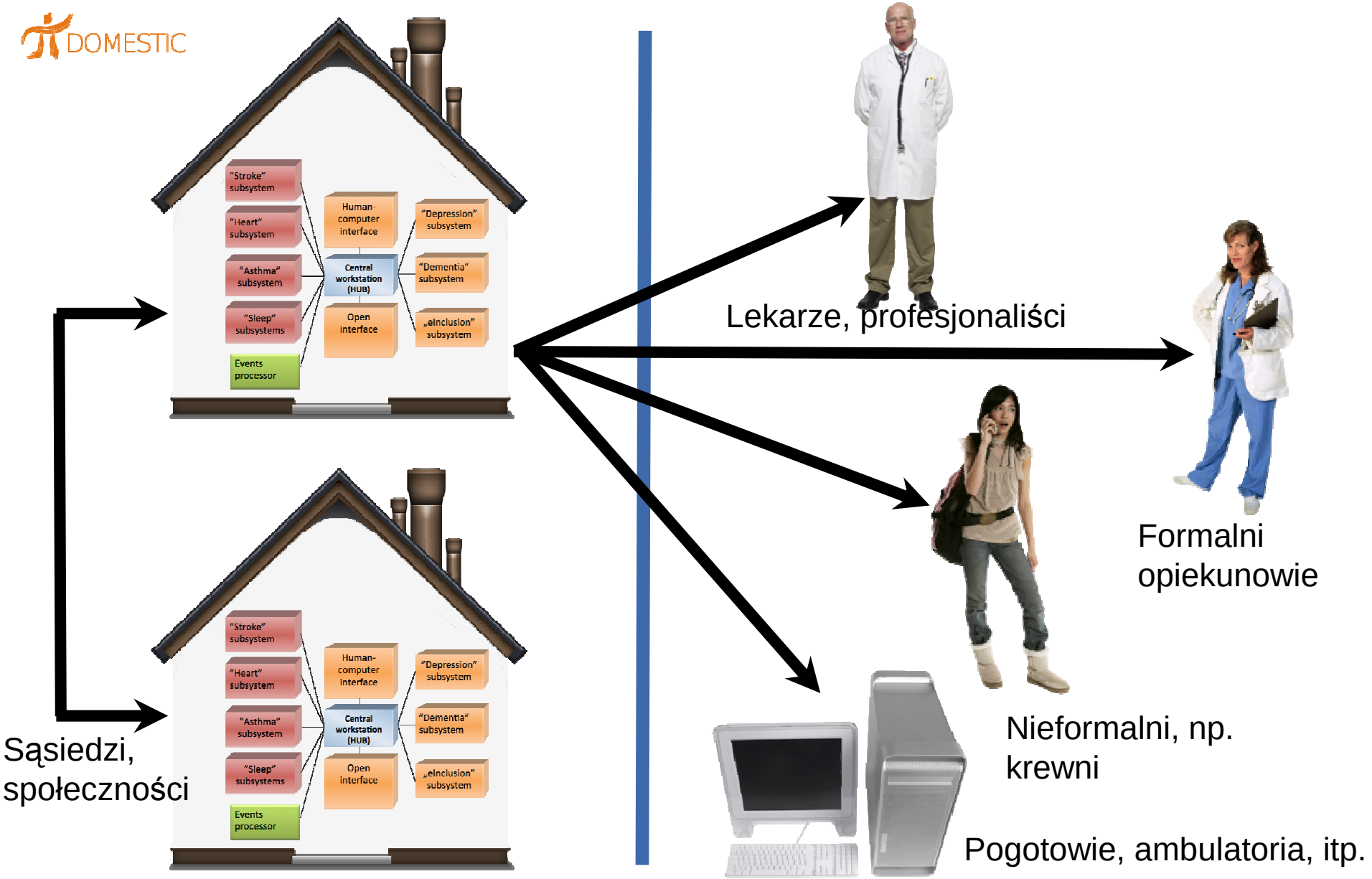
Praca współfinansowana przez Europejski Funduszu Rozwoju Regionalnego, projekt: UDA-POIG.01.03.01-22-139/09-00 -“Domowy asystent osób starszych i chorych – DOMESTIC”, Innowacyjna Gospodarka 2007-2013, Narodowa Strategia Spójności

Spersonalizowana opieka medyczna

DOMESTIC idea



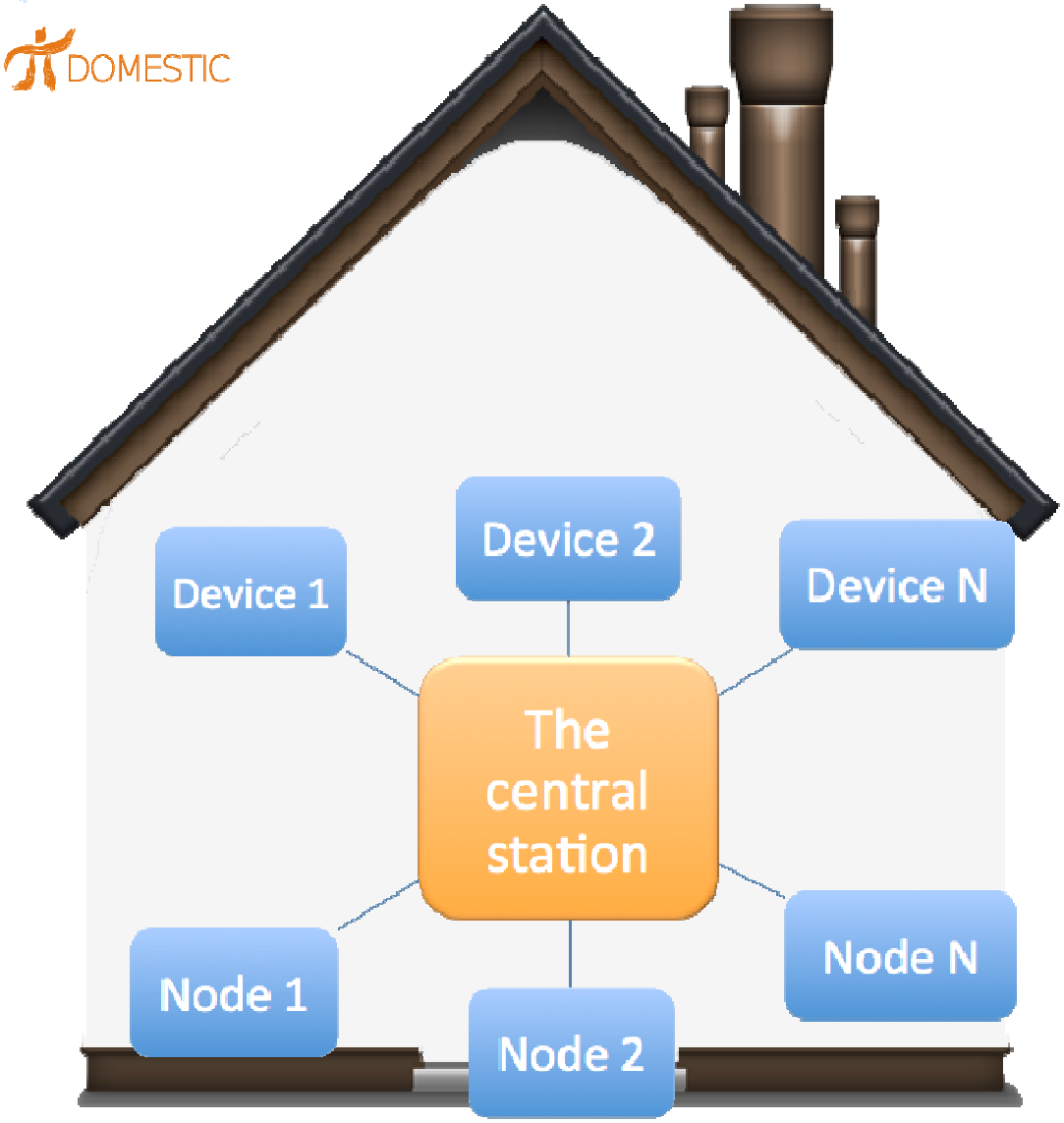
Spersonalizowana opieka medyczna

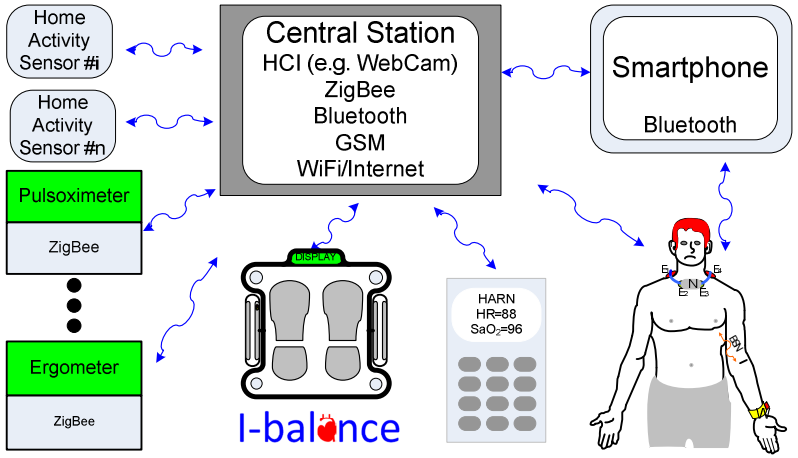


Spersonalizowana opieka medyczna -> Stacja centralna

DOMESTIC

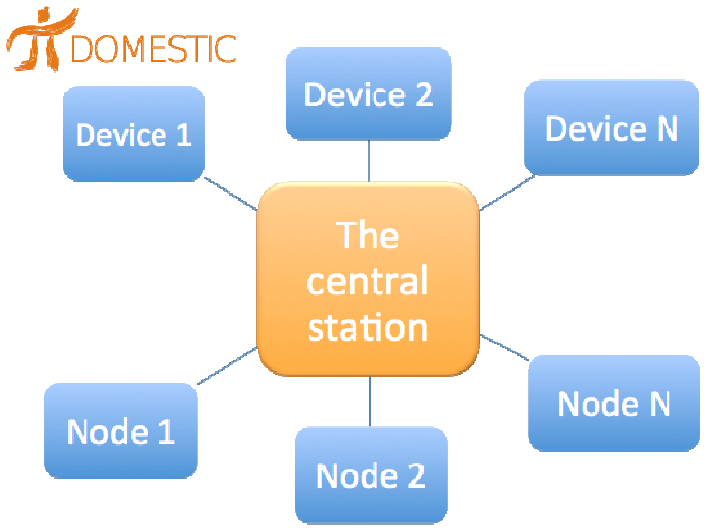
POLITECHNIKA GDAŃSKA





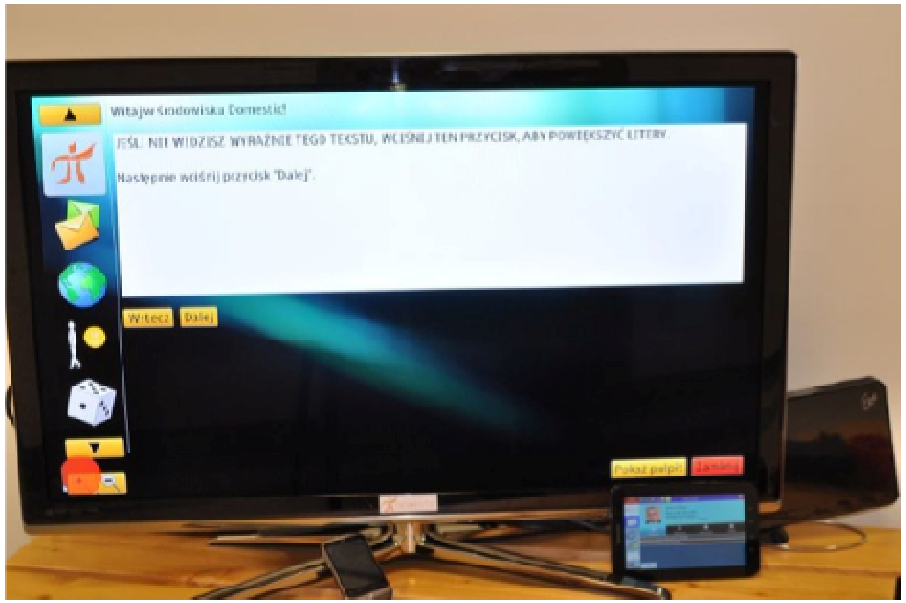
Krótkookresowe i długookresowe zbierania danych i analiza: dane medyczne, aktywność, bezpieczeństwo, zdarzenia, itp.

Spersonalizowana opieka medyczna -> KIB PG



Interfejsy:
Bluetooth, Zigbee, WiFi, GSM, inne

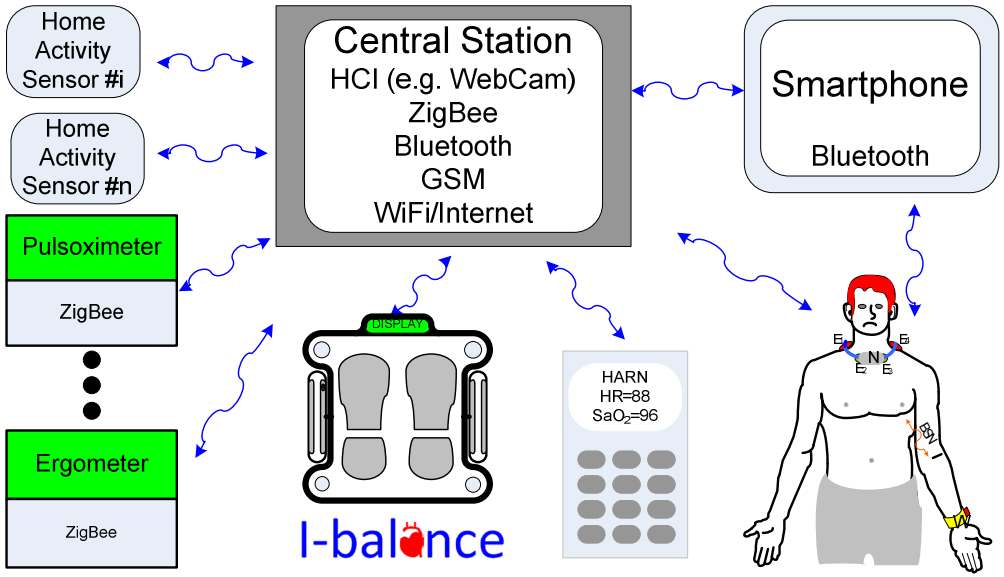
Funkcje:
Dane: zbieranie, przetwarzanie, odkrywanie wiedzy, wizualizacja, informacja (alerty), itp.



Spersonalizowana opieka medyczna -> Stacja centralna

DOMESTIC

- Stacja centralna
- Waga
- Krzesło (narzuta)
- Pilot
- Wanna
- SleAP
- PathMon
- Dmuchawka
- Systemy wykorzystujące Kinect
- Inne (inteligentne gniazdka, czujniki aktywności, …)



I-waga

Urządzenie składa się z układu pomiaru masy za pomocą czujników tensometrycznych oraz rozbudowane jest o układ umożliwiający dodatkowe pomiary: Body Mass Index i sześciu odprowadzeń elektrokardiograficznych optyczny detektor dla pomiaru saturacji krwi SaO_2 oraz o układ transmisji bezprzewodowej do komputera-laptopa

E-narzuta

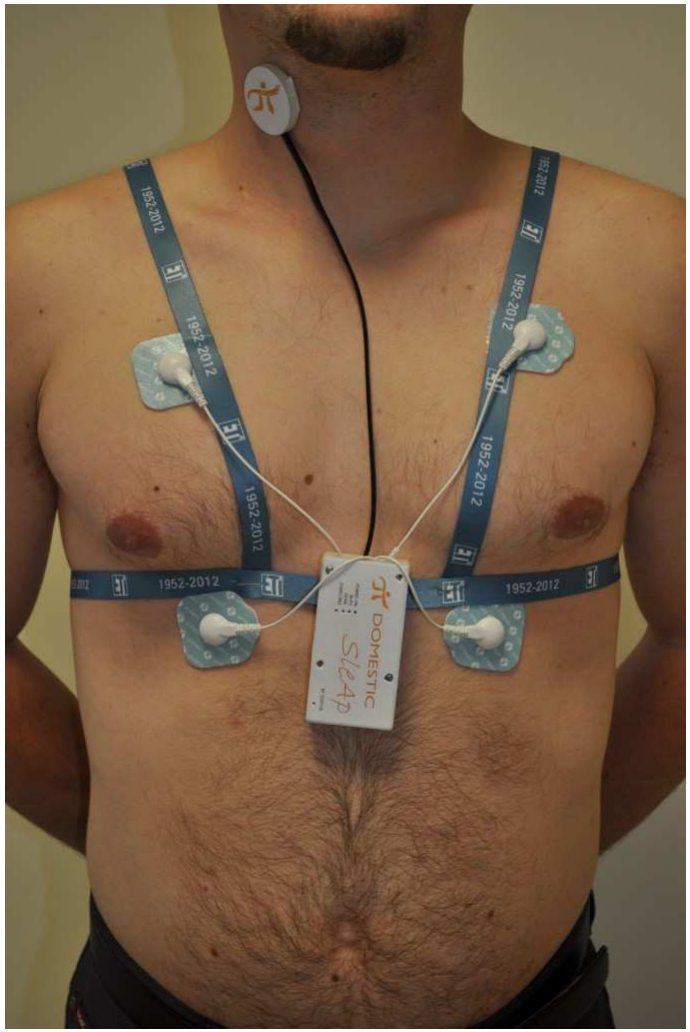
Poza przetwarzaniem EKG narzuta wyposażona jest w dwusektorowy, ciśnieniowy czujnik położenia osoby siedzącej. Wartości pomiarowe wskazują na aktywność fizyczną osoby siedzącej i są raportowane do stacji nadzorującej.

HARN

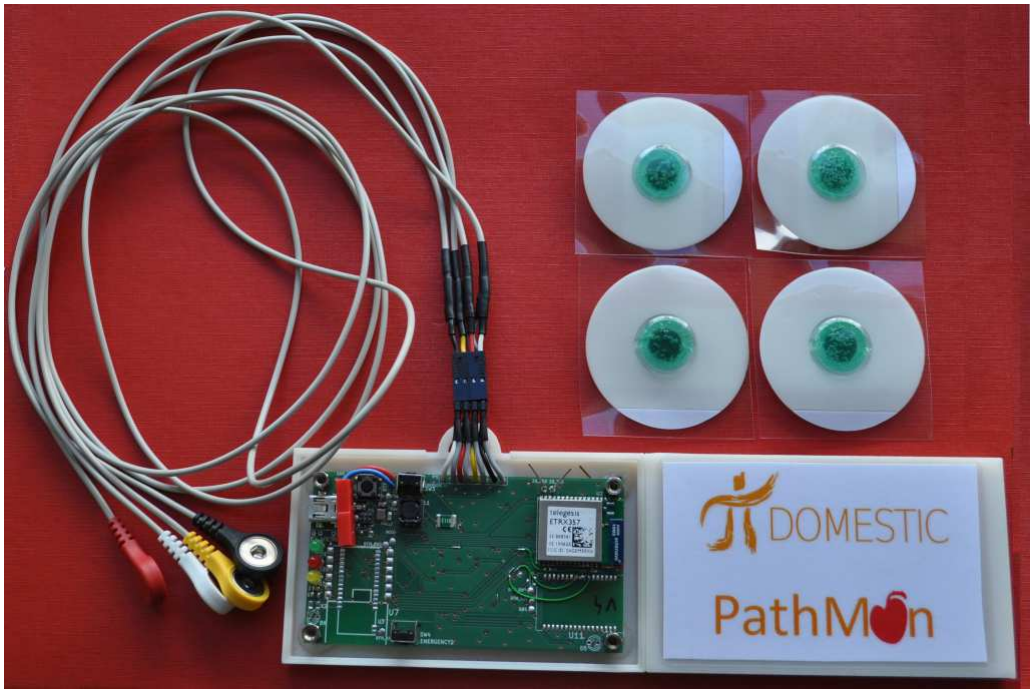
Mobilny monitor pracy serca, aktywności oddechowej, ruchowej, temperatury i postawy.

E-wanna

- **Kontroluje proces napełniania wanny – stabilizując temperaturę wody i jej poziom,**
- **Wykryje osobę w wannie oraz w przybliżeniu jej pozycję,**
- **Zna aktualny poziom i temperaturę wody w wannie,**
- **Zasugeruje moment opuszczenia kąpieli,**
- **Sprawdzi czy osoba w wannie aktywnie się porusza,**
- **Zmierzy częstość oddechu osoby w wannie,**
- **Zmierzy częstość uderzeń serca,**
- **Wykryje moment wchodzenia i wychodzenia z wanny,**
- **Poinformuje opiekuna o wykrytym stanie zagrożenia.**



Detekcja bezdechu za pomocą pomiaru EKG i impedancji klatki piersiowej



Monitoring sygnału EKG oraz mechanicznej pracy serca (za pomocą pomiaru zmian impedancji klatki piersiowej)

POLITECHNIKA GDAŃSKA

Predykcja sygnałów

Klasyfikacja

Analiza EKG

Detekcja bezdechu

Modelowanie pomiarów impedancyjnych

Analiza Fouriera

W przypadku określenia liczby próbek branych do predykcji pożyteczna może być znajomość występowania trendów okresowych zmian ciśnienia, jak i innych sygnałów.

W celu przebadania, czy takie trendy występują przeprowadzono analizę widmową Fouriera.

Przekształcenie Fouriera

Przyjmijmy, że funkcja jednej zmiennej $f(x)$ jest transformowalna (nie wnikając, jakie założenia muszą być spełnione, aby tak było).

Transformata Fouriera (prosta)

$$F(\omega) = \int_{-\infty}^{+\infty} f(x) e^{-i\omega x} dx$$

Transformata odwrotna

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{i\omega x} d\omega$$

Dyskretne Przekształcenie Fouriera

Niech będzie dany sygnał y_m określony w równoodległych punktach t_m na przedziale [0,2π]

Transformata prosta

$$Y_k = \frac{1}{M}\sum_{m=0}^{M-1} y_m e^{-ikt_m} = \frac{1}{M}\sum_{m=0}^{M-1} y_m e^{-ik\frac{2\pi}{M}m} \qquad k = -M/2,\ldots,M/2$$

$$t_m = 2\pi m/M, \qquad m = 0,\ldots,M-1$$

Dla nierównomiernych

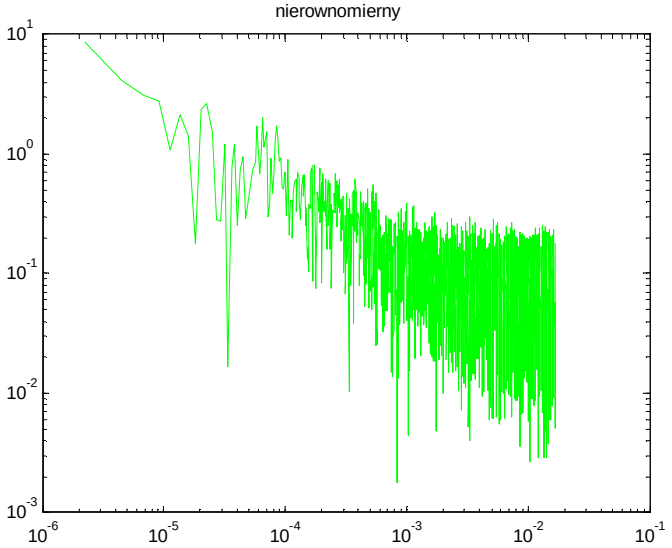
$$Y_k = \frac{1}{N}\sum_{n=0}^{N-1} y_n e^{-ikt_n} \qquad k = -N/2,\ldots,N/2$$

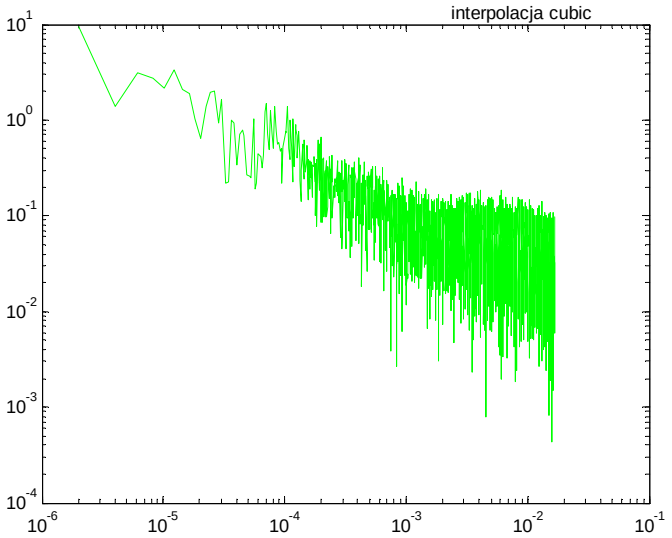
$$0 \le t_n \le 2\pi, \qquad n = 0,\ldots,N-1$$

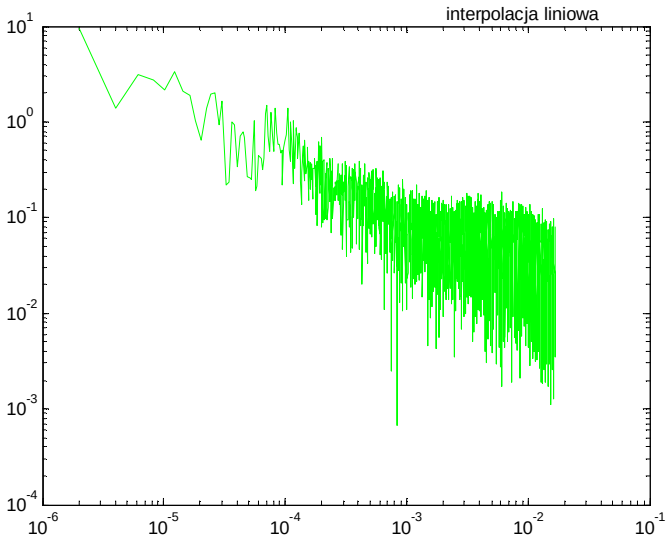
Przykład – widmo sygnału ciśnienia

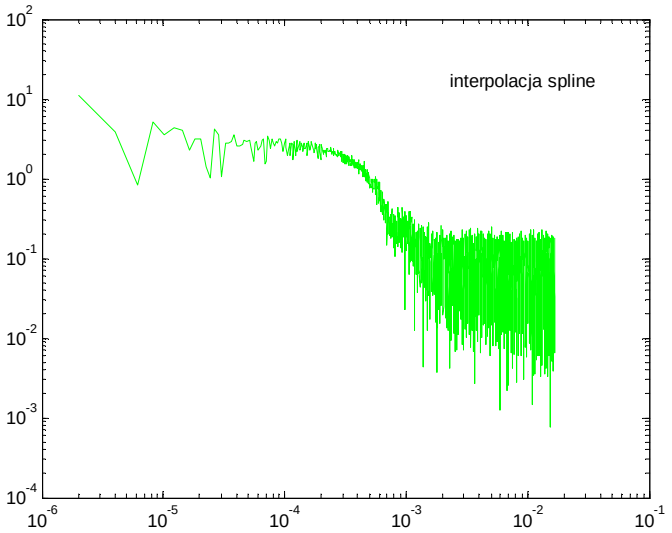
- analiza sygnału dla próbkowania nierównomiernego
- interpolacja i następnie analiza
  - liniowa,
  - spline – funkcje sklejane 3 stopnia,
  - cubic – interpolacja Hermita wielomianem stopnia 3.

Analiza Fouriera









Przykładowe wyniki dla ciśnienia skurczowego mierzonego inwazyjnie. Na osi poziomej częstotliwość w Hz.

Predykcja

- Wykorzystanie filtru adaptacyjnego
- Wykorzystanie rozwinięcia w szereg Taylora
- Ekstrapolacja
- Metody łączone
- Łańcuchy Markowa

Filtr adaptacyjny

- **Model**

$$p(n \mid P_{n-1}) = \sum_{k=1}^{N} a_k \, p(n-k)$$

- **Minimalizujemy**

$$E[(p(n) - p(n \mid P_{n-1}))^2]$$

Filtr adaptacyjny

$$E[p(n-1)p(n)] = \begin{bmatrix} E[p(n-1)p(n)] \\ E[p(n-2)p(n)] \\ \dots \\ E[p(n-N-1)p(n)] \\ E[p(n-N)p(n)] \end{bmatrix} = \begin{bmatrix} R_{pp}(1) \\ R_{pp}(2) \\ \dots \\ R_{pp}(N-1) \\ R_{pp}(N) \end{bmatrix} = r$$

$$R = E[p(n-1)p^T(n-1)] = \begin{bmatrix} R_{pp}(0) & R_{pp}(1) & \dots & R_{pp}(N-2) & R_{pp}(N-1) \\ R_{pp}(1) & R_{pp}(0) & \dots & R_{pp}(N-3) & R_{pp}(N-2) \\ \dots & \dots & \dots & \dots & \dots \\ R_{pp}(N-2) & R_{pp}(N-3) & \dots & R_{pp}(0) & R_{pp}(1) \\ R_{pp}(N-1) & R_{pp}(N-2) & \dots & R_{pp}(1) & R_{pp}(0) \end{bmatrix}$$

$$Ra = r$$

Filtr adaptacyjny

$$R_{pp}(j-k) = R_{pp}(k-j) = E[p(n-k+1)\,p(n-j+1)]$$

Filtr adaptacyjny

- **Wada**
  - **próbkowanie równomierne**

Filtr adaptacyjny

- **Problem**
  - **długość filtru**
  - **estymacja funkcji autokorelacji**

Estymacja funkcji autokorelacji

– **Możemy wybrać estymator obciążony**

$$\hat{R}_{pp}(k) = \frac{1}{M}\sum_{n=k+1}^{M} p(n)\,p(n-k), \quad k = 0,1,\ldots., N$$

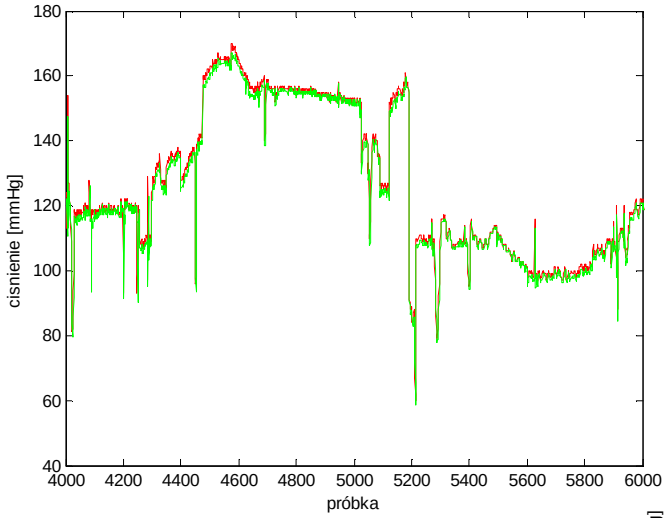
– **Analogicznie możemy wybrać estymator nieobciążony postaci**

$$\hat{R}_{pp}(k) = \frac{1}{M-k}\sum_{n=k+1}^{M} p(n)\,p(n-k), \quad k = 0,1,\ldots., N$$

Zapominanie danych historycznych

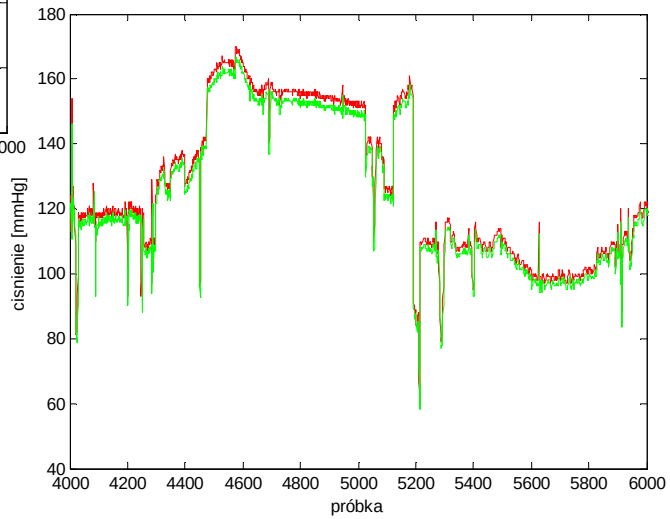
$$\hat{R}_{pp}(k) = \frac{1}{M} \sum_{n=k+1}^{M} \lambda^{M-n} p(n)\, p(n-k), \quad k = 0,1,\ldots., N$$

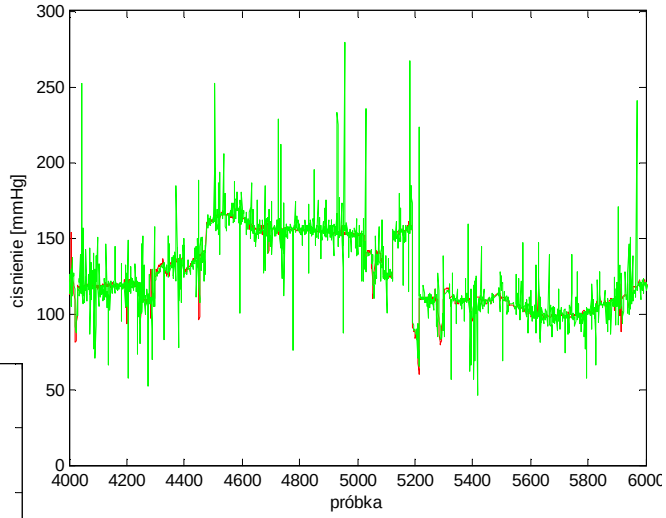
Filtr adaptacyjny



a)

b)

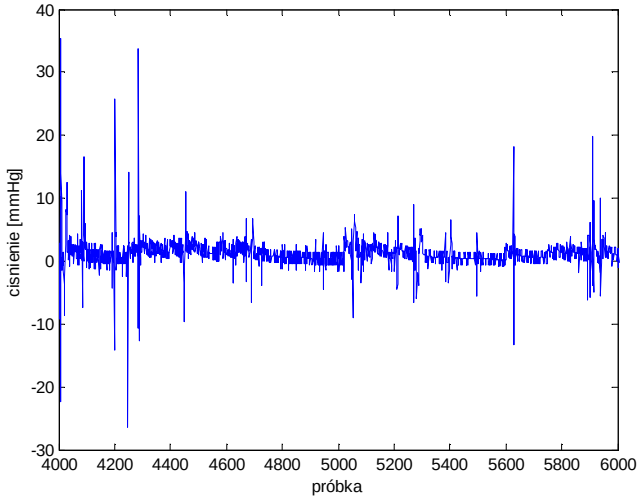




c)

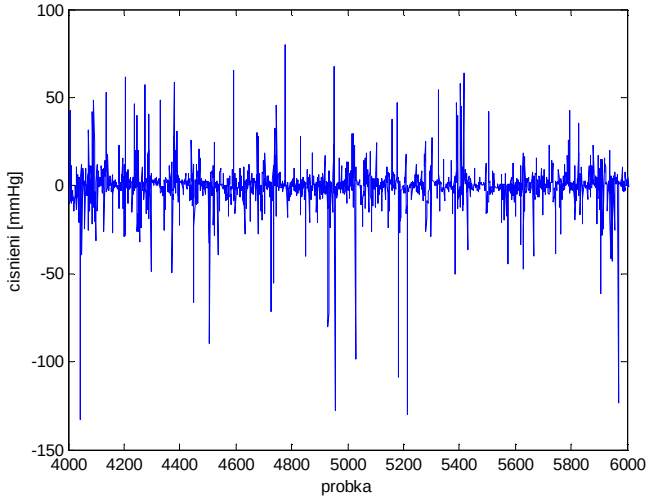
Wyniki dla predykcji ciśnienia tętniczego dla filtru długości 12 próbek. Rysunek a) pełna informacja, b) model z obcięciem o długości 5 krotność długości filtru, c) model z zapominaniem, współczynnik zapominania równy 0.9. Kolorem czerwonym oznaczono pomiary, a zielonym przewidywane wartości.

Filtr adaptacyjny

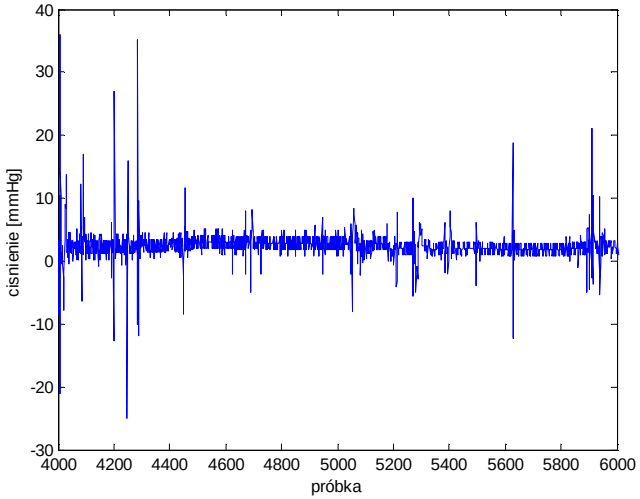


a)

b)



c)



Błędy dla predykcji ciśnienia tętniczego dla filtru długości 12 próbek. Rysunek a) pełna informacja, b) model z obcięciem o długości 5 krotność długości filtru, c) model z zapominaniem, współczynnik zapominania równy 0.9.

Rozwinięcie w szereg Taylora

Prognozujemy $p(t+a)$ wykorzystując dane historyczne $p(t-b_k)$

$$p(t+a) = \sum_{k=0}^{n} \alpha_k \, p(t-b_k)$$

Rozwinięcie w szereg Taylora

Rozwijamy $p(t+a)$ i $p(t-b_k)$ w szereg Taylora

$$p(t+a) = \sum_{i=0}^{n} \frac{p^{(i)}(t)}{i!} a^i$$

$$p(t-b_k) = \sum_{i=0}^{n} \frac{p^{(i)}(t)}{i!} (-1)^i b_k^i$$

Rozwinięcie w szereg Taylora

Stąd

$$\begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ -b_1 & -b_2 & \dots & -b_{n-1} & -b_n \\ \dots & \dots & \dots & \dots & \dots \\ (-b_1)^{n-1} & (-b_2)^{n-1} & \dots & (-b_{n-1})^{n-1} & (-b_n)^{n-1} \\ (-b_1)^{n} & (-b_2)^{n} & \dots & (-b_{n-1})^{n} & (-b_n)^{n} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_{n-1} \\ \alpha_n \end{bmatrix} = \begin{bmatrix} a^0 \\ a^1 \\ \dots \\ a^{n-1} \\ a^n \end{bmatrix}$$

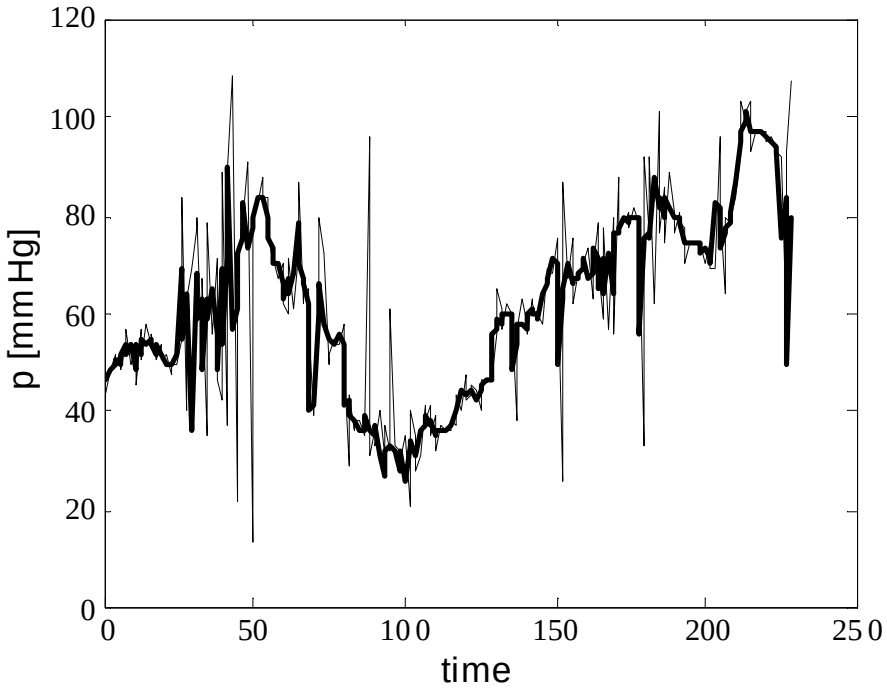
Rozwinięcie w szereg Taylora

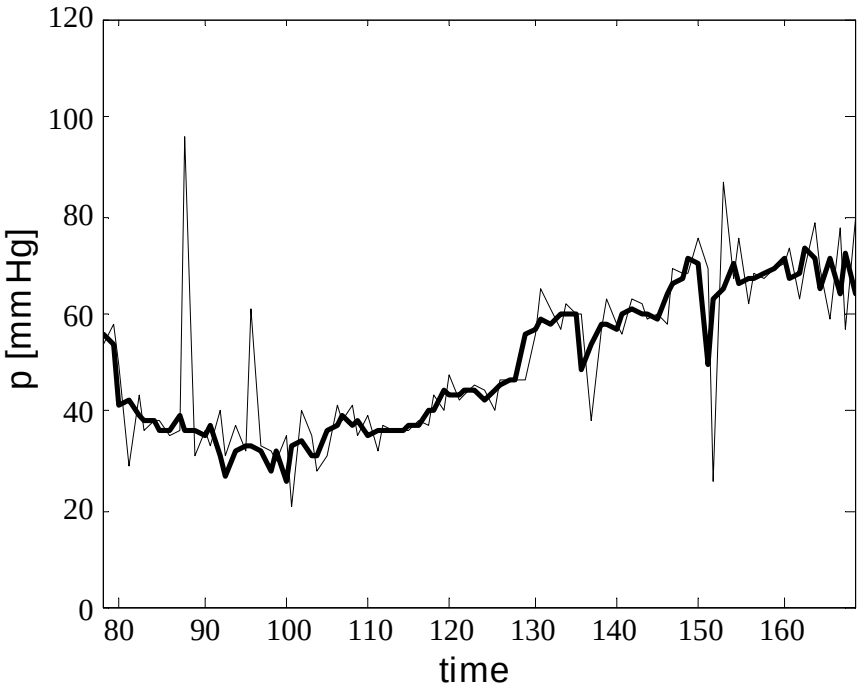
- **Zaleta**
  - **nie zależy od momentu czasu**
  - **próbkowanie nierównomierne**

- **Wady**
  - **dobre dla sygnałów aproksymowanych wielomianem (możliwie niskiego stopnia)**
  - **wrażliwe na szumy**

Rozwinięcie w szereg Taylora

Przykład predykcji





Ekstrapolacja

Dane

- f **– dany sygnał o skończonej energii, o widmie zawartym w obszarze** Ω
- g**- znany fragment sygnału** f

Zadanie

odtworzyć f **znając** g

Metoda iteracyjna

$$y_0(x_j) = 0$$

$$y_{m+1}(x_j) = y_m(x_j) + \sum_{x_k \in B} \sin c_\Omega(x_j - x_k)\left(g(x_k) - y_m(x_k)\right)$$

Sanz i Huang "Discrete and continuous band-limited Signac extrapolation" IEEE Trans. Acoustics, Speech, and signal processing vol. assp-31, no.5, 1983, pp 1276-1285

Ekstrapolacja

- **Stabilność dla funkcji *sinc***
- **Inne jądra h i regularyzacja λ**
- **Próbkowanie równomierne i nierównomierne**

Ekstrapolacja

Podobne podejście jest przedstawione w pracy

T. Strohmer, On discrete band-limited signal extrapolation, Contemp. Math. 190:323-337, 1995.

Zamiast iterować równanie splotowe rozwiązuje się metodą sprzężonych gradientów problem

$$p_0 = \lambda S p,$$

$$p_{m+1} = p_m + \lambda S(p - p_m)$$

Ekstrapolacja

S jest operatorem, D jądrem Dirichleta, a r liczbą próbek sygnału ($0 \le t_1 < t_2 < \ldots t_r < 1$)

$2M+1 \le r$

$$Sp(t) = \sum_{j=1}^{r} p(t_j) D_M(t - t_j)$$

$$D_M(t) = \sum_{k=-M}^{M} e^{2\pi i k t} = \frac{\sin(2M+1)\pi t}{\sin \pi t}$$

Ekstrapolacja

Inne podejście jest w pracy

H.G. Feichtinger, K.Gröchenig, and T.Strohmer, Efficient Numerical Methods in Non-uniform Sampling Theory, Numerische Mathematik, 69:423-440, 1995.

Ekstrapolacja

Aproksymują poszukiwaną funkcję wzorem

$$p(t) = \sum_{k=-M}^{M} a_k e^{2\pi i k t}$$

gdzie a_k są to nieznane współczynniki aproksymacji (trzeba je wyznaczyć)

Ekstrapolacja

wtedy

$$Sp(t) = \sum_{j=1}^{r} p(t_j) D_M(t - t_j) = \sum_{j=1}^{r} \sum_{k=-M}^{M} a_k e^{2\pi i k t_j} \sum_{l=-M}^{M} e^{2\pi i l t} e^{-2\pi i l t_j}$$

Przestawiając sumy dostajemy

$$= \sum_{l=-M}^{M} \left[\sum_{k=-M}^{M} \left(\sum_{j=1}^{r} e^{-2\pi i (l-k) t_j} \right) a_k \right] e^{2\pi i l t}$$

Ekstrapolacja

Zdefiniujmy macierz

$$T_{lk} = \sum_{j=1}^{r} e^{-2\pi i(l-k)t_j} \qquad |l|, |k| \le M$$

Korzystając z tego mamy

$$Sp(t) = \sum_{l=-M}^{M} \left[\sum_{k=-M}^{M} T_{lk} a_k \right] e^{2\pi i l t}$$

Ekstrapolacja

Z drugiej strony

$$Sp(t) = \sum_{j=1}^{r} p(t_j) D_M(t - t_j) = \sum_{k=-M}^{M} \left(\sum_{j=1}^{r} p(t_j) e^{-2\pi i k t_j} \right) e^{2\pi i k t}$$

Definiując

$$b_k = \sum_{j=1}^{r} p(t_j) e^{-2\pi i k t_j} \qquad |k| \le M$$

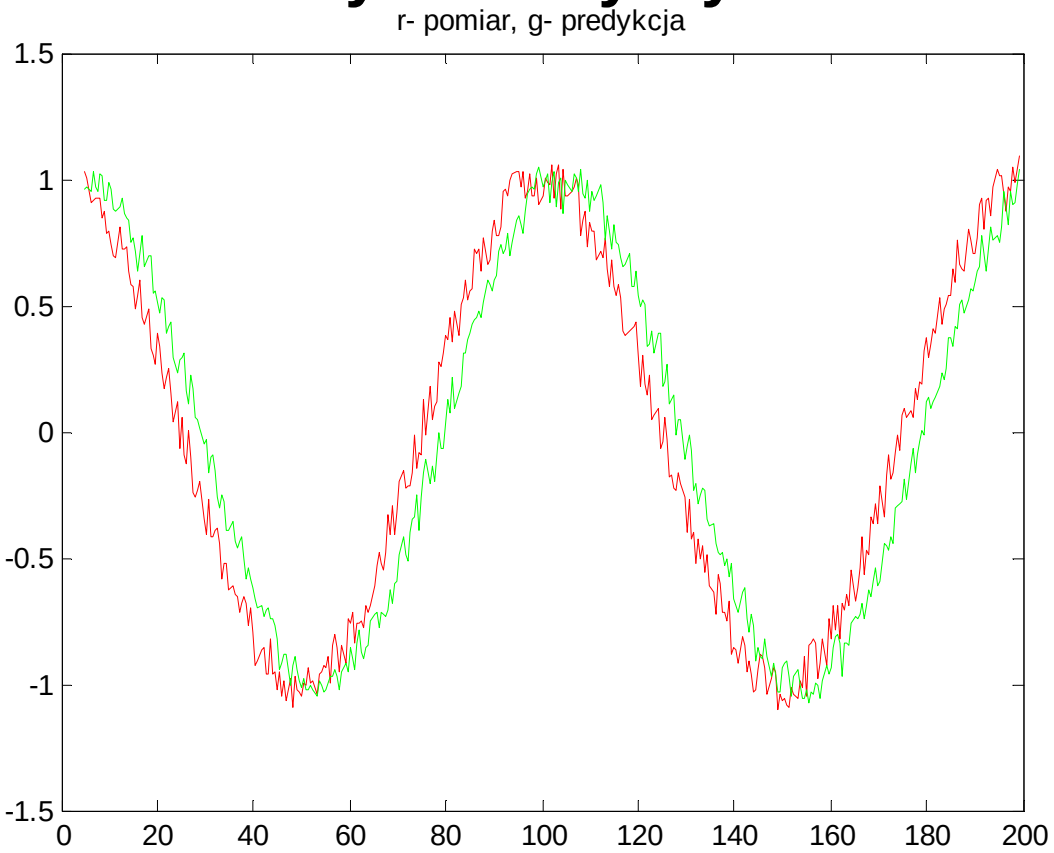
Ekstrapolacja

Dostajemy układ $Ta=b$, z którego wyznaczamy a, a to daje aproksymację p.

To jest prostsza wersja, którą można rozszerzyć na próbkowanie nierównomierne (algorytm jest podany w cytowanej pracy)

Ekstrapolacja

Przykładowy wyniki



czerwony - oryginalny sygnał, zielony – predykcja

Model - porównanie

Rozważamy dwa modele

$$s(t+a)=\sum_{k=0}^{n}\alpha_k s(t-b_k)$$

$$s(t+a)=\sum_{k=0}^{n}\alpha_k s(t-b_k)+\alpha_{n+1}$$

$$a,\ b_k>0,\ b_i\neq b_j \quad \text{dla } i\neq j$$

Rozwinięcie w szereg Taylora - modyfikacja

$$\begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ -b_0 & -b_1 & \dots & -b_{n-1} & -b_n \\ \dots & \dots & \dots & \dots & \dots \\ (-b_0)^{m-1} & (-b_1)^{m-1} & \dots & (-b_{n-1})^{m-1} & (-b_n)^{m-1} \\ (-b_0)^{m} & (-b_1)^{m} & \dots & (-b_{n-1})^{m} & (-b_n)^{m} \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \dots \\ \alpha_{n-1} \\ \alpha_n \end{bmatrix} \approx \begin{bmatrix} a^0 \\ a^1 \\ \dots \\ a^{m-1} \\ a^m \end{bmatrix}$$

Optymalizacja

$$\left\| s(t+\Delta t) - \sum_{k=0}^{n} \alpha_k s(t-k\Delta t) \right\|_2^2 =$$

$$\sum_{i=0}^{m} \left(s(t_i+\Delta t) - \sum_{k=0}^{n} \alpha_k s(t_i-k\Delta t) \right)^2$$

$$\left\| s(t+\Delta t) - \sum_{k=0}^{n} \alpha_k s(t-k\Delta t) - \alpha_{n+1} \right\|_2^2 =$$

$$\sum_{i=0}^{m} \left(s(t_i+\Delta t) - \sum_{k=0}^{n} \alpha_k s(t_i-k\Delta t) - \alpha_{n+1} \right)^2$$

POLITECHNIKA GDAŃSKA

Problem - jak wybrać m i n?

- Zmiana parametrów pacjenta w czasie – jak szybka?
- Dokładność pomiarów

Przyjmujemy

- $n=1, 2, ..., 4$
- $m=n, 2(n+1)-1, 3(n+1)-1, ..., 5(n+1)-1$

Algorytm konkurencyjny

$$\hat{s}(t) = \begin{cases} \hat{s}_1(t) & dla \quad \left|\hat{s}_1(t-\Delta t) - s(t-\Delta t)\right| \le \left|\hat{s}_2(t-\Delta t) - s(t-\Delta t)\right| \\ \hat{s}_2(t) & dla \quad \left|\hat{s}_2(t-\Delta t) - s(t-\Delta t)\right| < \left|\hat{s}_1(t-\Delta t) - s(t-\Delta t)\right| \end{cases}$$

Można wykorzystać więcej metod

Metody łączne

Kombinacja liniowa predyktorów

$$\hat{s}(t) = w_1(t)\hat{s}_1(t) + w_2(t)\hat{s}_2(t)$$

$$w_1(t) = \frac{Var((\hat{s}_2(t) - s(t))/s(t))}{Var((\hat{s}_1(t) - s(t))/s(t)) + Var((\hat{s}_2(t) - s(t))/s(t))}$$

$$w_2(t) = \frac{Var((\hat{s}_1(t) - s(t))/s(t))}{Var((\hat{s}_1(t) - s(t))/s(t)) + Var((\hat{s}_2(t) - s(t))/s(t))}$$

$$w_2(t) + w_1(t) = 1$$

Można wykorzystać więcej metod

Miara błędu

Jak porównać wyniki dla danych różnej długości?

$$q = \sqrt{\frac{\sum_{i=1}^{N} (\hat{s}(t_i) - s(t_i))^2}{N}}$$

Testy

Dane pomiarowe:

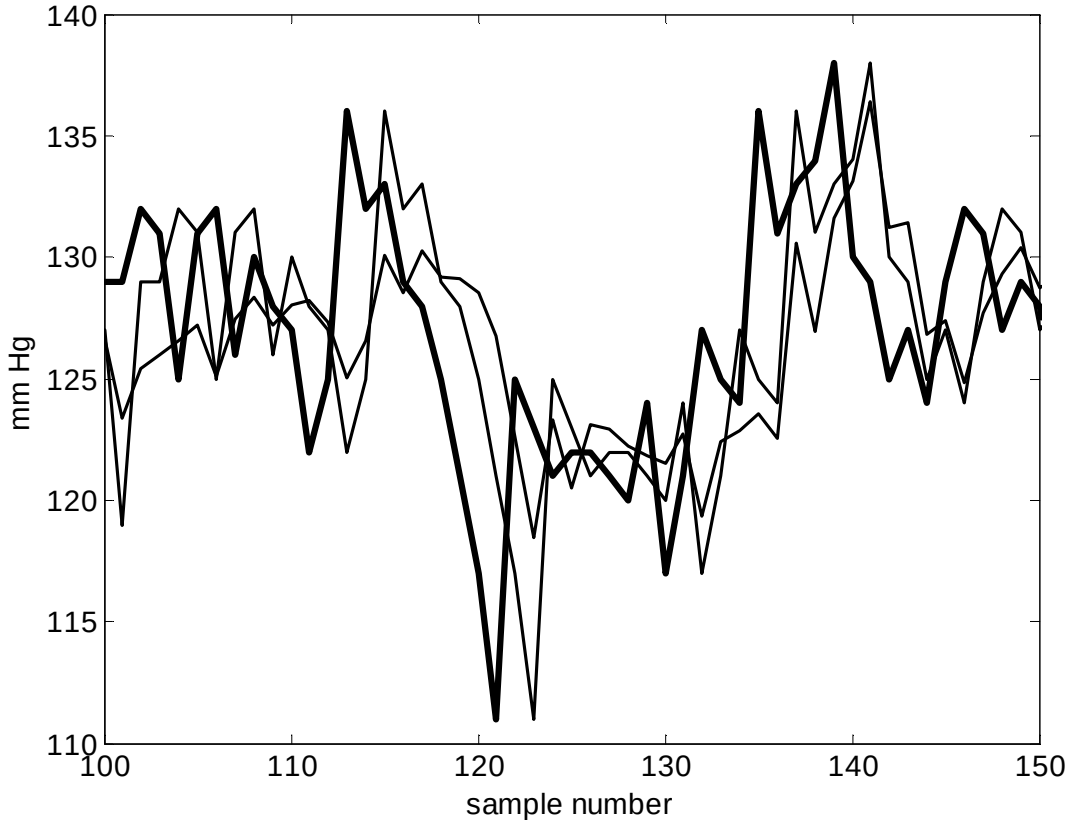
Trzy kobiety 47, 52 i 70 lat oraz jedne mężczyzna w wieku 86 lat.

Ciśnienie skurczowe i rozkurczowe z tętnicy promieniowej.

Wybrane po 2 najdłuższe nieprzerwanie mierzone wartości ciśnienia (od 257 do 1177 próbek co 30 sekund)

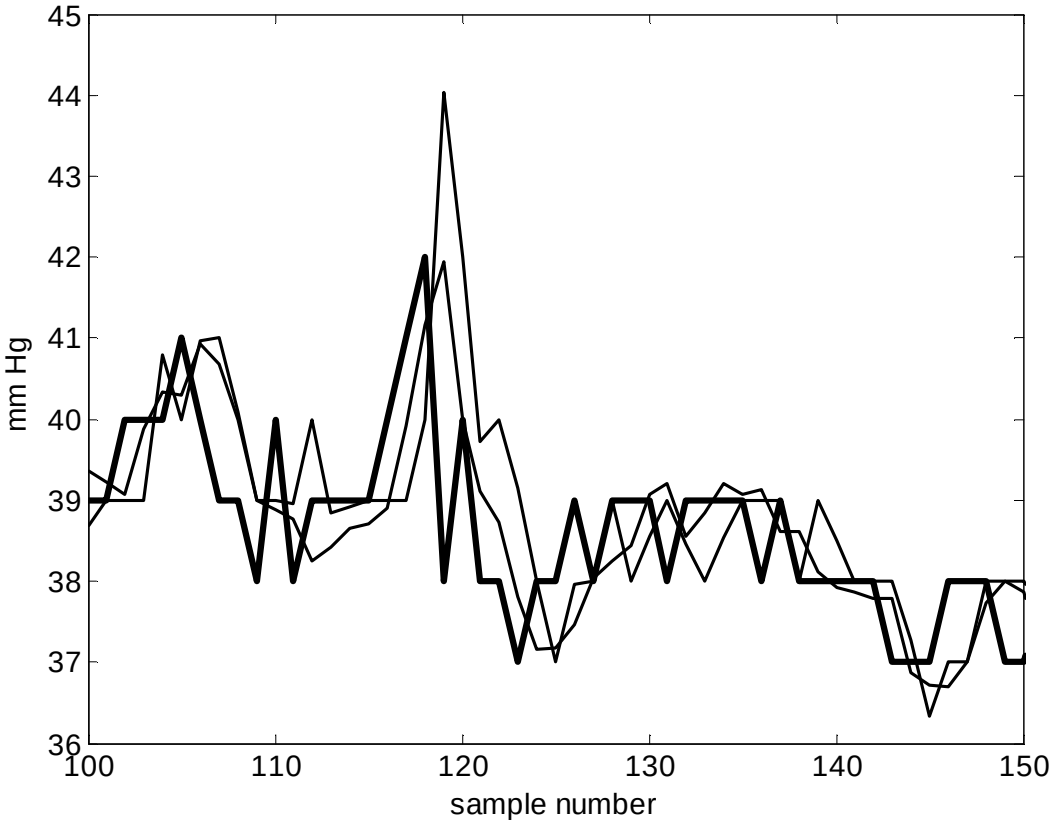
Wyniki testów

Przykład słabej predykcji – najlepsza i najgorsza

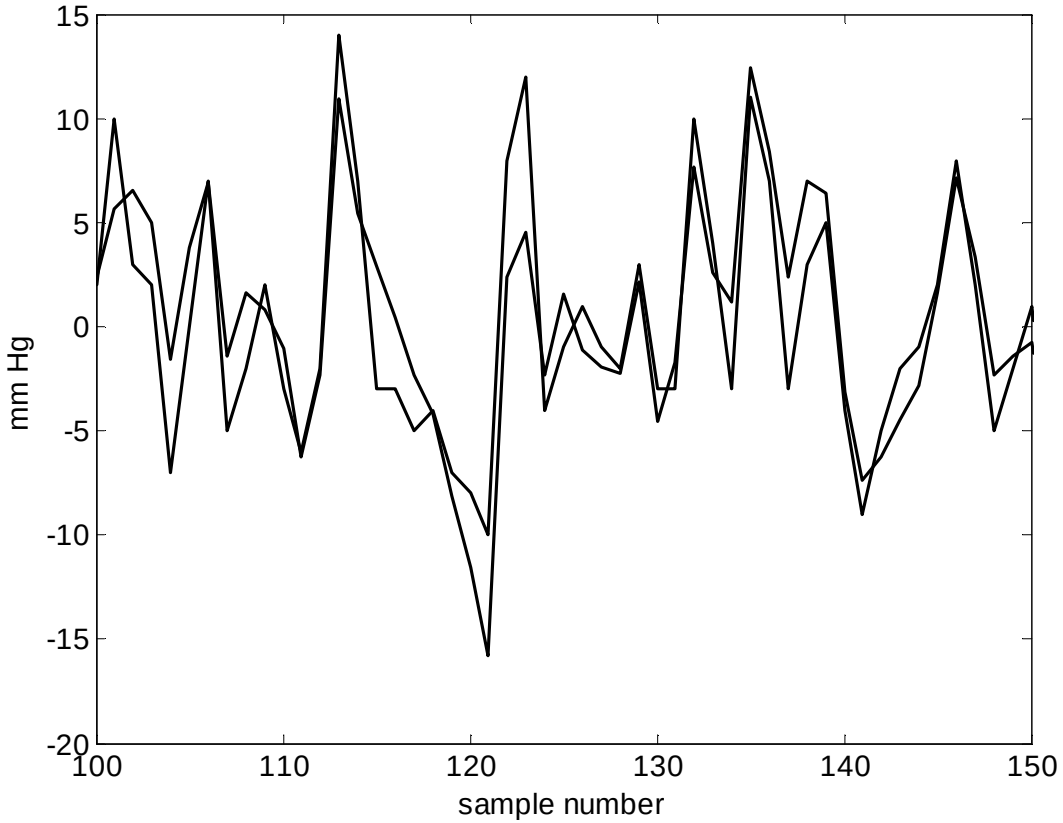


Wyniki testów

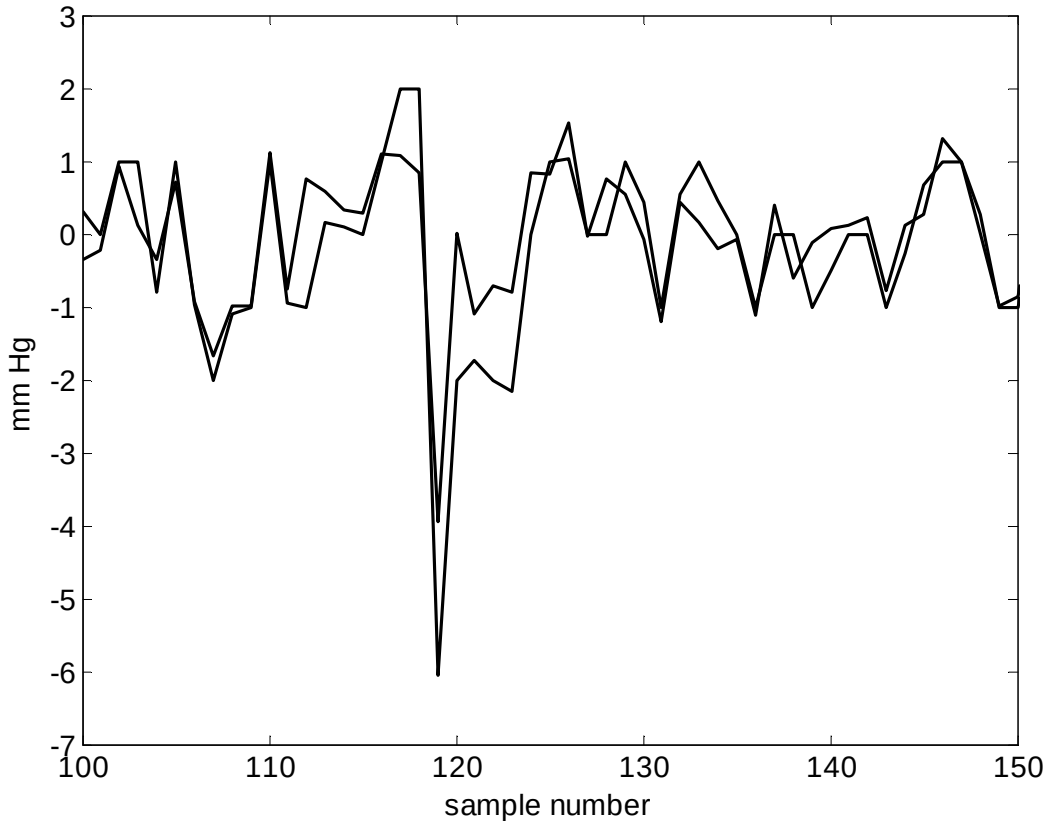
Przykład dobrej predykcji – najlepsza i najgorsza



Przykład błędu dla słabej predykcji – najlepsza i najgorsza

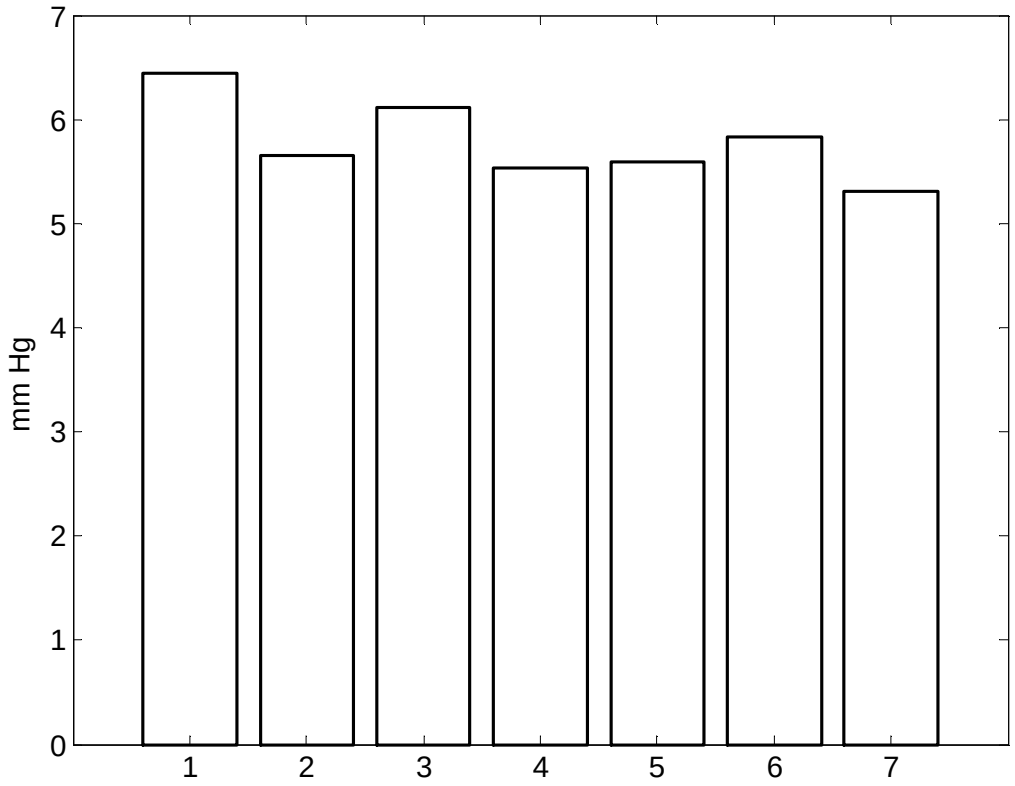


Przykład błędu dla dobrej predykcji – najlepsza i najgorsza



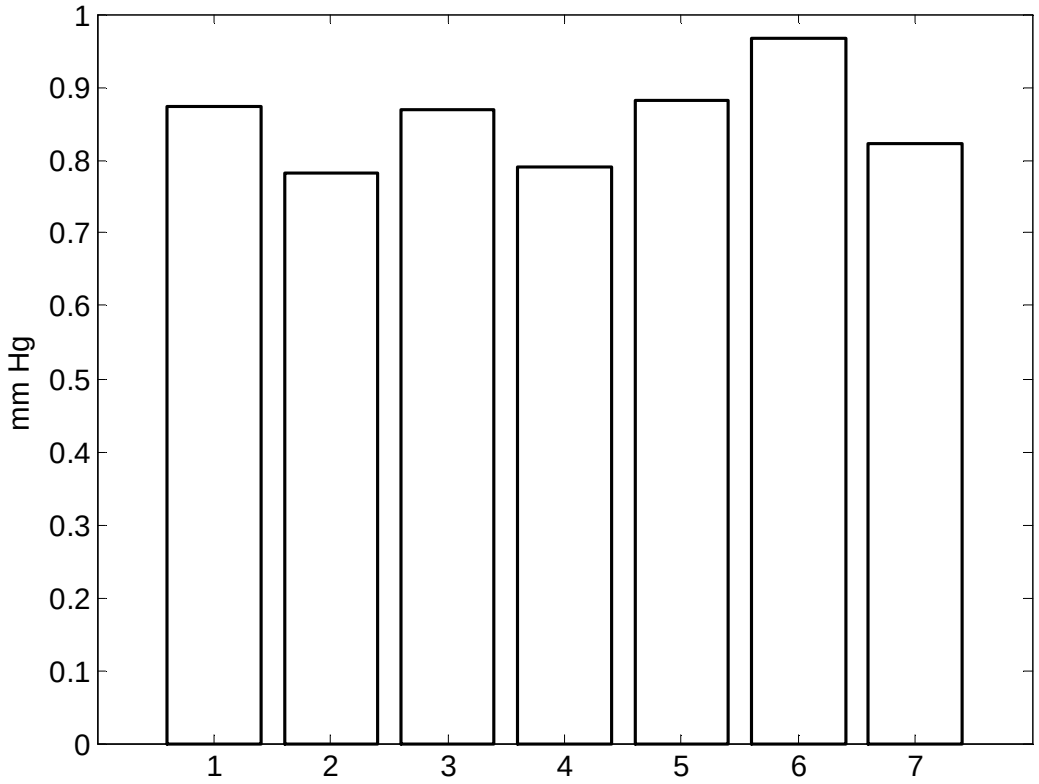
Wyniki testów

Błąd predykcji dla różnych metod dla słabej predykcji



Wyniki testów

Błąd predykcji dla różnych metod dla dobrej predykcji



Problemy praktyczne

- Wiele czynników wpływających na mierzone parametry (np. ciśnienie)
- Czynniki zmienne w czasie – także zależność od leków
- Zależność wyników od stabilności pacjenta
- Nie wszystkie parametry wpływające na dany sygnał fizjologiczny są możliwe do monitorowania
- Błędy pomiarowe

Wnioski

- **Brak najlepszej metody**
- **Zadowalająca dokładność predykcji w większości przypadków**

Wykorzystanie kilku parametrów do predykcji

POLITECHNIKA GDAŃSKA

Cel

Nieinwazyjny ciągły monitoring ciśnienia

Czy można wykorzystać np. częstość akcji serca?

Czy istnieje zależność pomiędzy parametrami

Korelacja

- surowe dane
- uśrednione dane (MA(3),MA(4),MA(5)) - odszumianie
- współczynnik nachylenia prostej ($y=ax+b$) - trendy

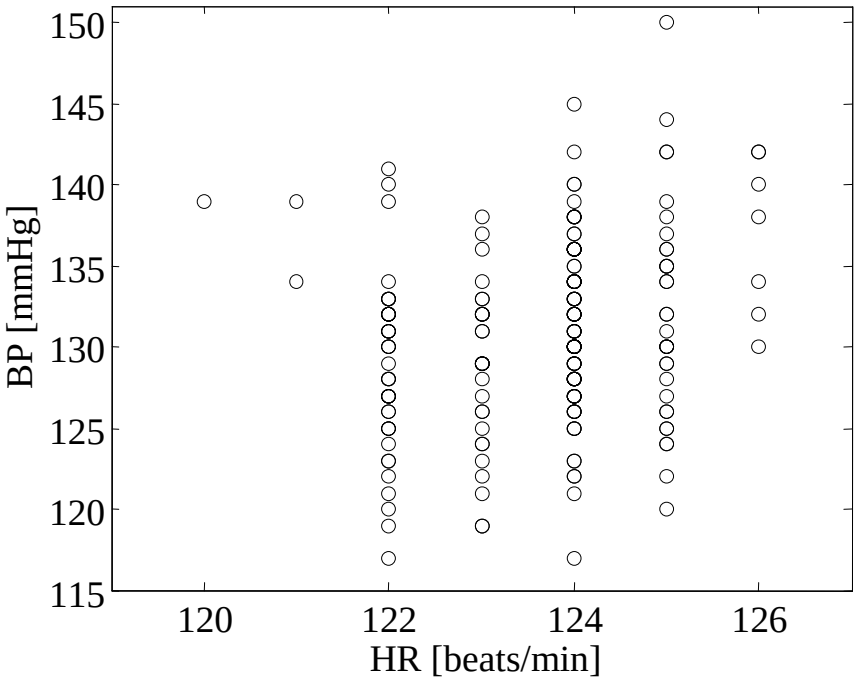
POLITECHNIKA GDAŃSKA

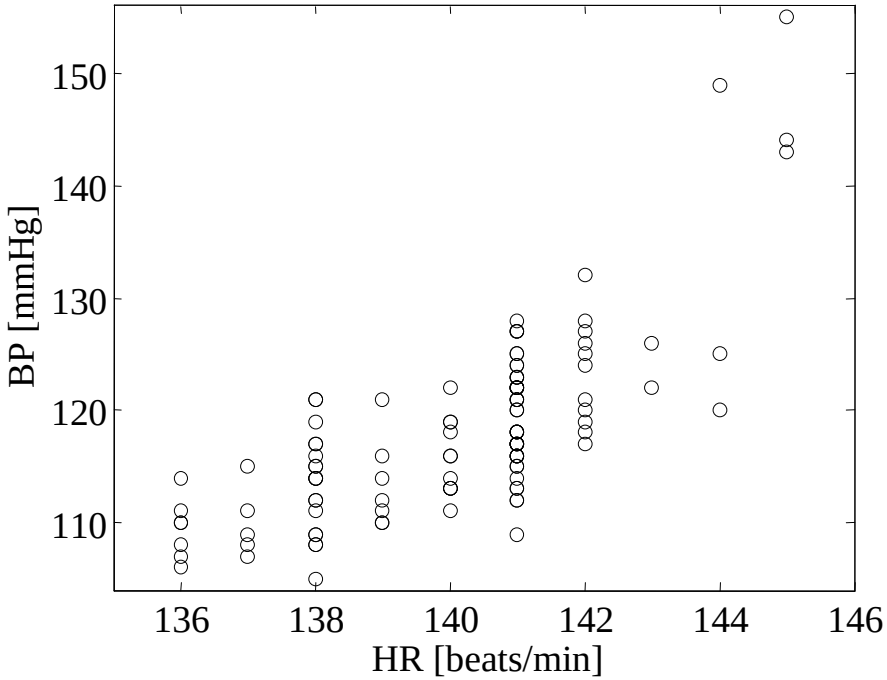
Analiza ciśnienia i częstości akcji serca ciągi min długości 100 próbek

Przykładowe wyniki

SBP vs HR – ten sam pacjent, różne momenty czasu. Współczynnik korelacji 0.229 (p<0.001) i 0.734 (p<0.001).

Surowe dane

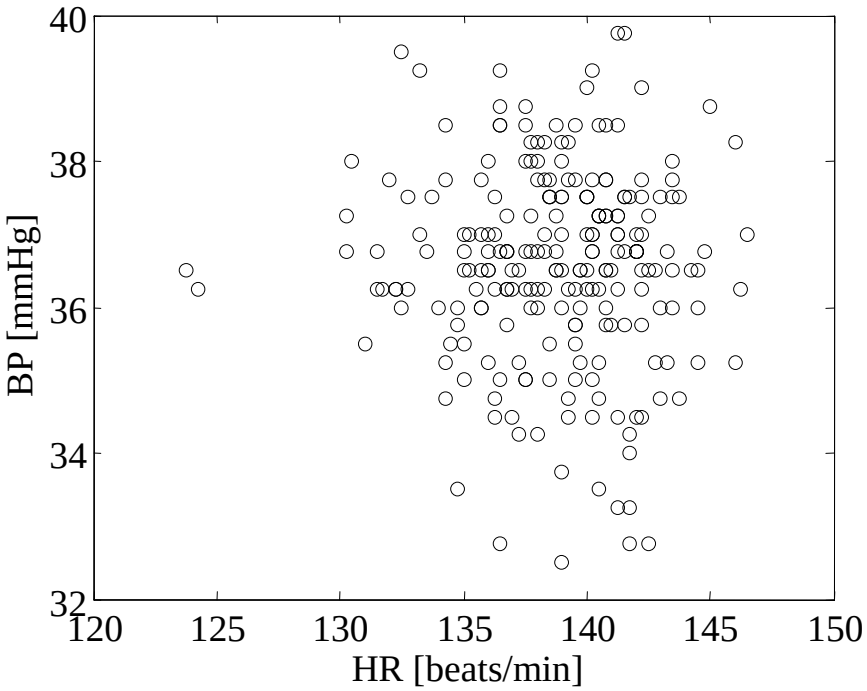


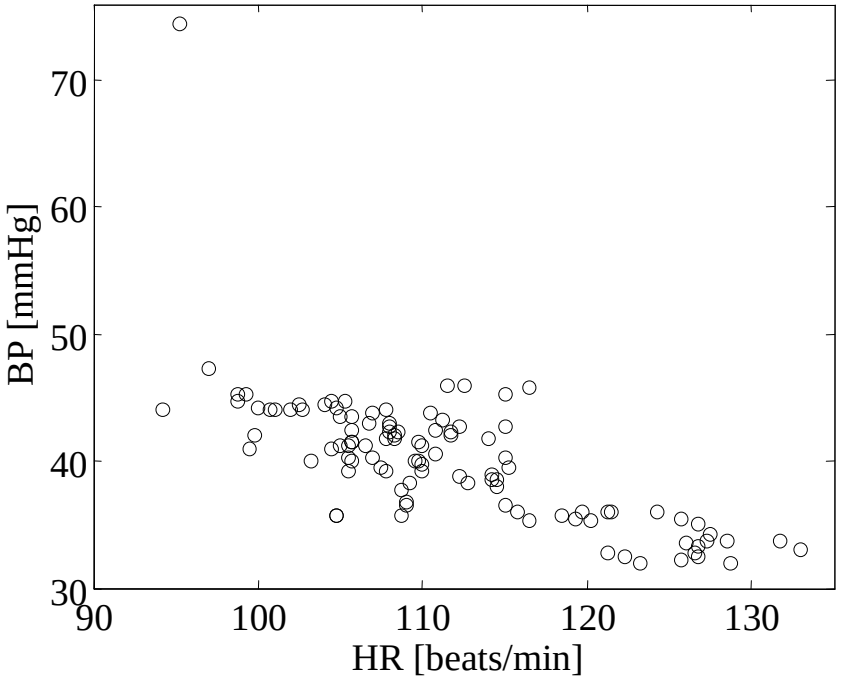


Przykładowe wyniki

DBP vs HR – ten sam pacjent, różne momenty czasu. Współczynnik korelacji -0.029 (p>0.05) i –0.708 (p<0.001).

Dane po filtracji

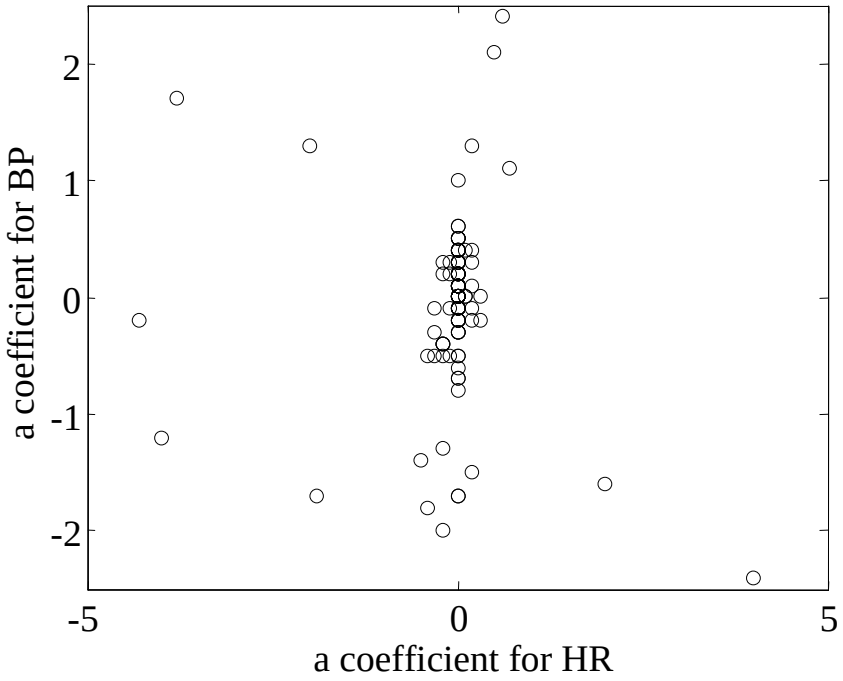


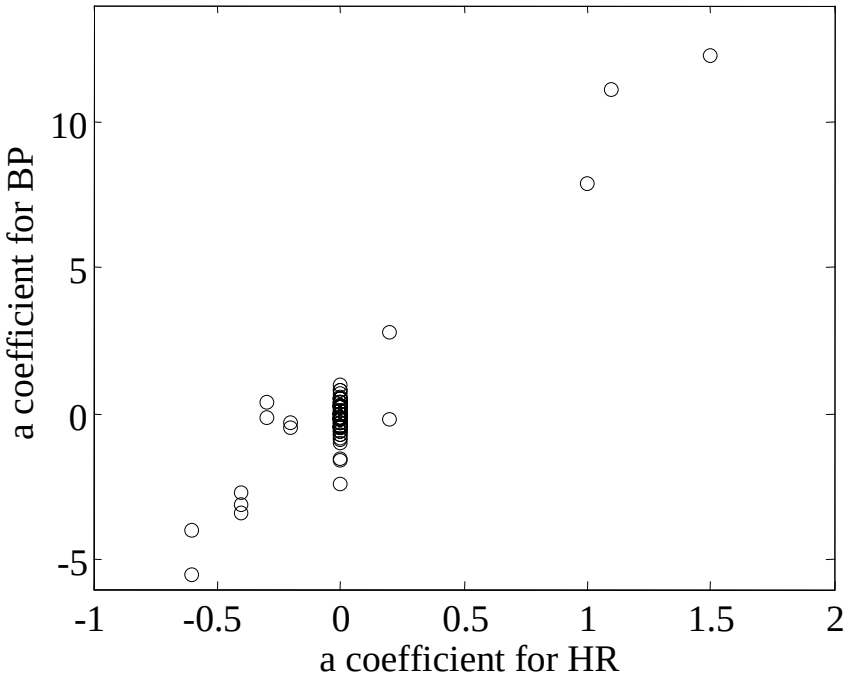


Przykładowe wyniki

SBP vs HR – ten sam pacjent, różne momenty czasu. Współczynnik korelacji –0.161 (p>0.05) i 0.939 (p<0.001).

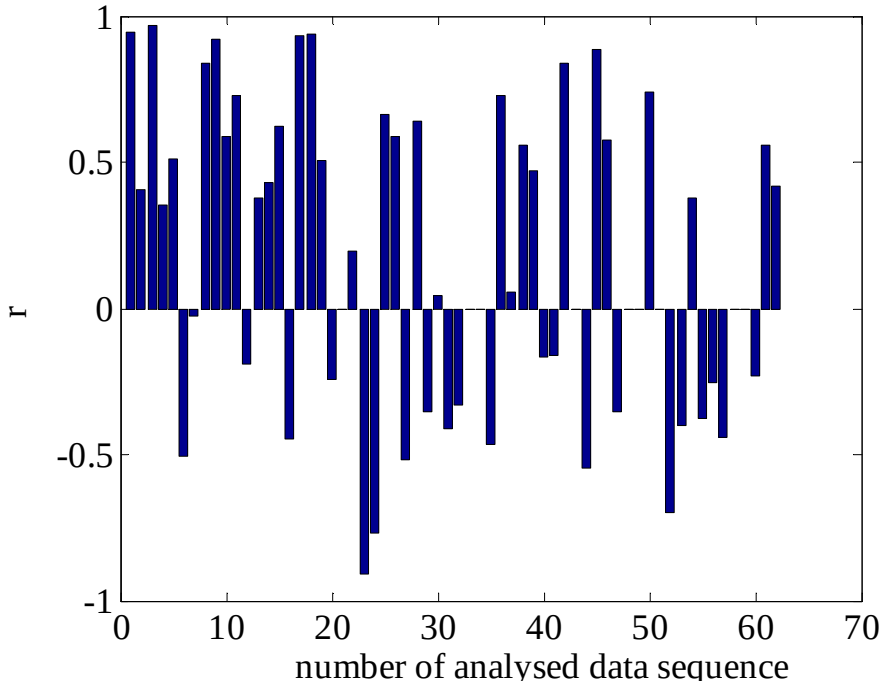
Trendy

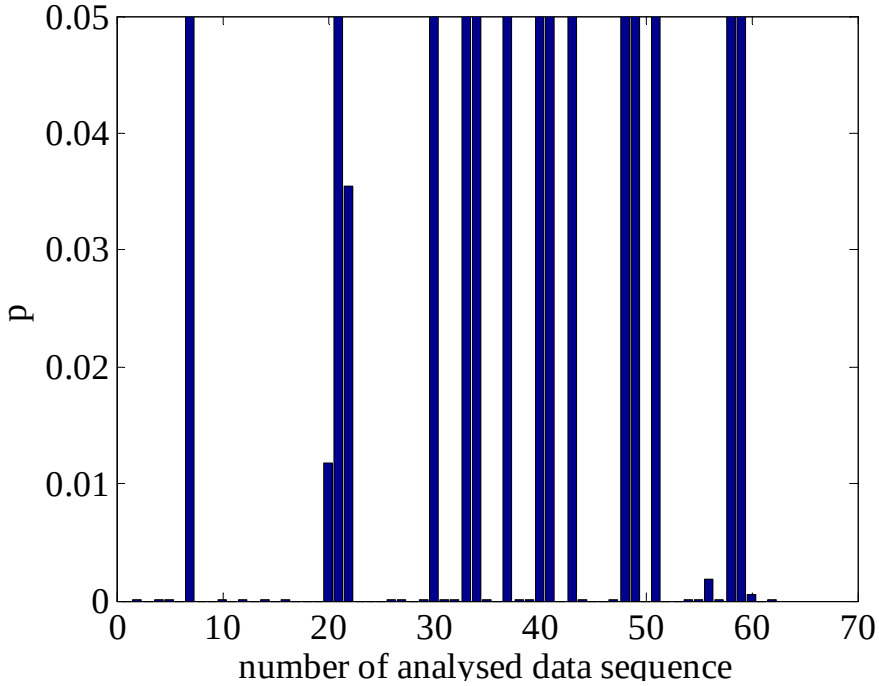




Przykładowe wyniki

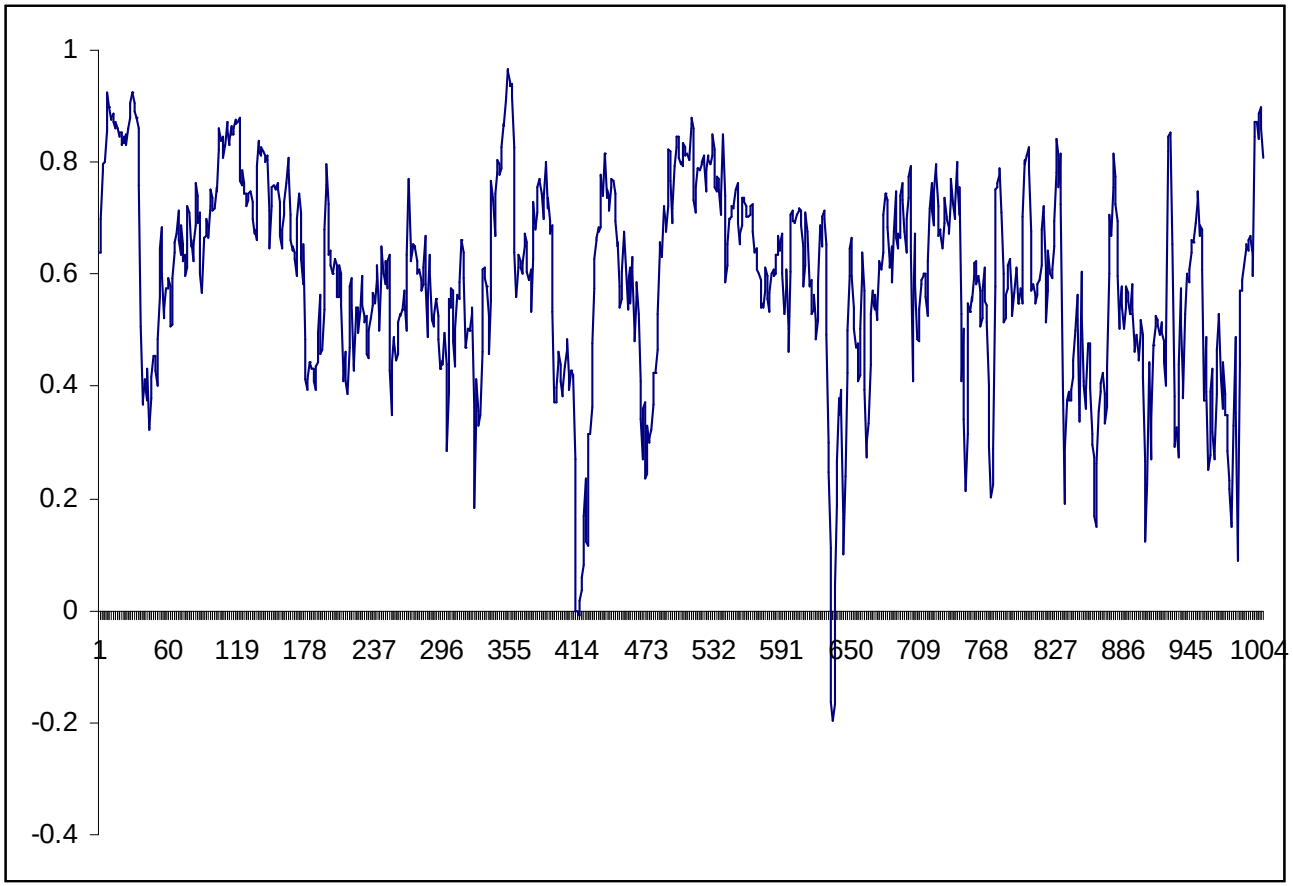
Przykładowe wyniki zmian współczynnika korelacji pomiędzy SBP i HR w czasie (sygnały filtrowane)





Korelacja ciśnienia i HR

Krótkookresowa korelacja



Analiza korelacji - wnioski

Duża zmienność

Zależność liniowa lub jej brak

Predykcja – łańcuchy Markowa

Idea zaczerpnięta z pracy

Baier V., Baumert M., Caminal P., Vallverdú M., Faber R., Voss A.,

Hidden Markov Models Based on Symbolic Dynamics for Statistical Modeling of Cardiovascular Control in Hypertensive Pregnancy Disorders,

IEEE Transactions on biomedical engineering: vol. 53, nr 1 2006.

Łańcuchy Markowa

$S = \{s_0, s_1, s_2, ..., s_r\}$ – zbiór stanów

Jeżeli łańcuch jest aktualnie w stanie s_i i przechodzi do stanu s_j z prawdopodobieństwem p_{ij} oraz to prawdopodobieństwo nie zależy od stanów w jakich łańcuch znajdował się poprzednio, to mówimy, że proces jest łańcuchemMarkowa.

Łańcuchy Markowa

Ciąg zmiennych losowych $(X_n)_{n=0}^{\infty}$ o wartościach w przeliczalnym zbiorze S (przestrzeni stanów) nazywamy łańcuchem Markowa wtedy i tylko wtedy, gdy dla każdego $n \in N$ i każdego ciągu s_0, s_1, s_2, ..., $s_n \in S$ mamy

$$P(X_n=s_n|X_{n-1}=s_{n-1}, \ldots, X_1=s_1, X_0=s_0) = P(X_n=s_n|X_{n-1}=s_{n-1}),$$

jeżeli

$$P(X_{n-1}=s_{n-1}, \ldots, X_1=s_1, X_0=s_0) > 0.$$

Łańcuchy Markowa

Macierz $P = (p_{ij})_{si,sj \in S}$ nazywamy macierzą przejścia na S, gdy wszystkie jej wyrazy są nieujemne, a ponadto suma każdego wiersza wynosi 1, czyli dla każdego $s_i \in S$

$$\sum_{s_j \in S} p_{ij} = 1$$

Łańcuchy Markowa

Jeżeli $(X_n)_{n=0}^{\infty}$ jest łańcuchem Markowa, to rozkład zmiennej losowej X_0 nazywamy rozkładem początkowym. Łańcuch Markowa nazywamy jednorodnym (jednorodnym w czasie), gdy istnieje macierz

$$\hat{P} = (\hat{p}_{ij})_{s_i, s_j \in S}$$

będąca dla każdego n jego macierzą przejścia w n-tym kroku.

Łańcuchy Markowa - predykcja

Wektorem stanu w chwili t nazywamy wektor postaci

$$y(t) = (y_1(t), y_2(t), ..., y_r(t)),$$

gdzie $y_i(t)$ dla $i \in \{1, 2, ...r\}$ jest prawdopodobieństwem znalezienia się w stanie $s_i \in S$ w chwili t.

Łańcuchy Markowa - predykcja

Prognoza

$$y_p(T+1) = y(T)\hat{P}$$

$$y_p(T+1) = y(T)\hat{P}(T+1)$$

Histogram

Wybór szerokości klasy *h*

$$f_n(x) = \frac{N \leftarrow x_i \in x}{nh}$$

$$N \leftarrow x_i \in x$$

liczba wartości x_i należących do x,

n – liczba przedziałów

Ocena dopasowania estymatora

$$R(f_n) = E\int_{-\infty}^{\infty}[f_n(x) - f(x)]^2\,dx$$

Łańcuchy Markowa - histogram

Jakie h jest optymalne?

$$h_0 = \frac{c}{\sqrt[3]{n}}$$

Dla dużych wartości n

$$c = \sqrt[3]{6 \Big/ \int [f'(x)]^2 dx}$$

Łańcuchy Markowa - histogram

Jeżeli rozkład jest normalny to

$$c = 2\sqrt[3]{3\sqrt[6]{\pi}}\sigma \approx 3.486\sigma$$

$$S_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}$$

POLITECHNIKA GDAŃSKA

Inny dobór

David Freedman and Persi Diaconis,

On the Histogram as a Density Estimator: L_2 Theory,

Z. Wahrscheinlichkeitstheorie verw. Gebiete 57, 453-476 (1981)

Łańcuchy Markowa - histogram

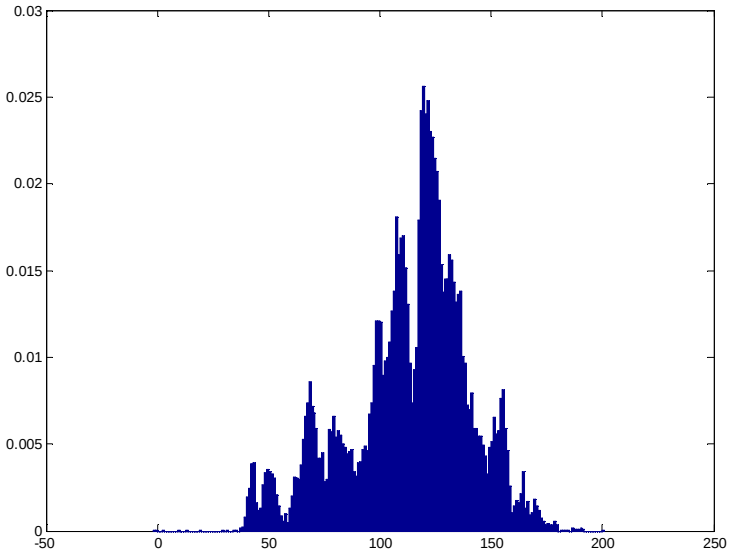
Brak znajomości c, a więc

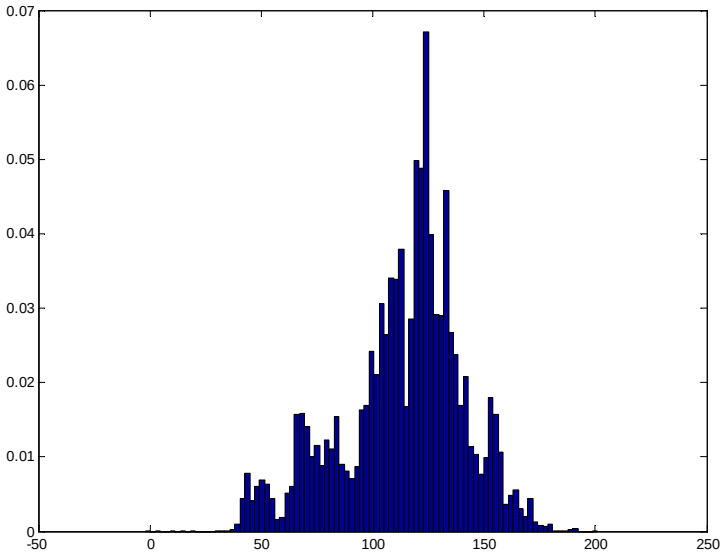
Choose the cell width as twice the interquartile range of the data,

divided by the cube root of the sample size.

Łańcuchy Markowa - histogram

Histogramy – te same dane, różna ilość podprzedziałów

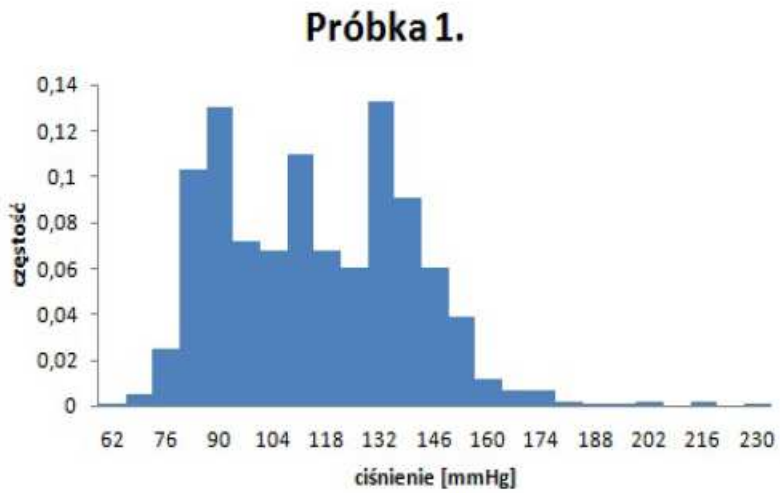




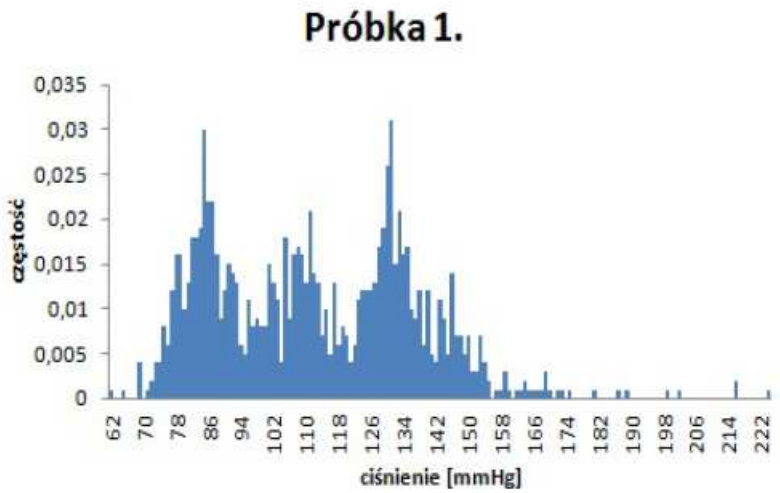
Łańcuchy Markowa - histogram

Przykładowe wyniki

$h_* = 8,7 \approx 9$



$h = 1,3 \approx 1$



* oznacza h wyznaczone z założeniem rozkładu normalnego, bez * - estymacja z wykorzystaniem przybliżenia pochodnych

Łańcuchy Markowa - predykcja

POLITECHNIKA GDAŃSKA

Przykładowa estymowana macierz przejścia

	105-111	112-118	119-125	126-132	133-139	140-146	147-153	154-160	161-167
105 - 111	0,33	0,50	0,17	0,00	0,00	0,00	0,00	0,00	0,00
112 - 118	0,03	0,50	0,44	0,03	0,01	0,00	0,00	0,00	0,00
119 - 125	0,00	0,10	0,84	0,05	0,00	0,01	0,00	0,00	0,00
126 - 132	0,00	0,03	0,56	0,32	0,06	0,00	0,00	0,03	0,00
133 - 139	0,00	0,00	0,17	0,06	0,72	0,06	0,00	0,00	0,00
140 - 146	0,20	0,20	0,00	0,00	0,20	0,20	0,00	0,20	0,00
147 - 153	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
154 - 160	0,00	0,00	0,00	0,67	0,00	0,33	0,00	0,00	0,00
161 - 167	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00

POLITECHNIKA GDAŃSKA

Problemy

- wyznaczeniem rozkładu
- weryfikacją jednorodności łańcuchów Markowa
  - – testy wykazały jej brak, nawet przy analizie uwzględniającej rytm dobowy

Elementami zbioru treningowego X_T są instancje (x_i,k_i) będące listą wartości atrybutów i wybranego

atrybutu decyzyjnego - klasy

$$X_T = \{(x_1, k_1), (x_2, k_2), \ldots, (x_M, k_M)\},\ x_i \in X \subset R^N,\ k_i \in \{k_1, k_2, \ldots k_K\}$$

Przykładowy wektor

$$x_i = (a_{i1}, a_{i2}, \ldots, a_{iN})$$

Klasyfikacja

Klasyfikator przypisuje nieznanym obserwacjom określoną klasę. Przypisanie to dokonywane jest na podstawie funkcji $f(x)$, za której zbudowanie odpowiada klasyfikator.

Klasyfikacja

Najprostszy przykład – klasyfikator binarny

Zakładamy dwie klasy – stan pacjenta stabilny oraz stan pacjenta niestabilny

Jeśli instancja jest przypisana do klasy niestabilnej i klasyfikator przewiduje dla niej taką klasę, to zalicza się ją do *true positive* (TP). Jeśli przypisuje ją do klasy stabilnej – to liczy się ją jako *false negative* (FN). Jeśli instancja w rzeczywistości posiada klasę stabilną i zostaje sklasyfikowana tak samo, to zalicza się ją do *true negative* (TN). W przeciwnym przypadku liczy się ją jako *false postive* (FP)

Klasyfikator bayesowski

Zakładamy, że mamy dane prawdopodobieństwa przynależności do klas $p(k_1), p(k_2), \ldots, p(k_N)$ oraz prawdopodobieństwa warunkowe wystąpienia wektora obserwacji x w poszczególnych klasach $p(x|k_1), p(x|k_2), \ldots, p(x|k_N)$.

Korzystając z twierdzenia Bayesa można obliczyć prawdopodobieństwo wystąpienia klasy k pod warunkiem zaobserwowania wektora x:

$$p(k_i \mid x) = \frac{p(x \mid k_i)\, p(k_i)}{p(x)}$$

$$p(x) = \sum_{i=1}^{N} p(x \mid k_i)$$

Klasyfikator bayesowski

Wyszukując maksymalne prawdopodobieństwo warunkowe możemy określić do jakiej klasy należą zaobserwowane dane. Problemem jest znajomość rozkładu prawdopodobieństwa.

$$f(x) = \arg\max_{k_i} p(k_i \mid x)$$

Klasyfikator bayesowski - modyfikacja

Przy konstrukcji klasyfikatora można wziąć pod uwagę także różne koszty klasyfikacji (poprawne i błędne, różne dla różnych klas). Można to zrobić dodając macierz kosztów

Ferdinand van der Heijden, Image Based measurement Systems – object recognition and parameter estimation, John Wiley & Sons, 1994

Klasyfikator bayesowski - modyfikacja

POLITECHNIKA GDAŃSKA

Szukamy minimum funkcji kosztu

$$f(x) = \arg\min_{\substack{k_i \\ 1 \le i \le N}} \left\{ \sum_{j=1}^{N} C(k_i, k_j)\, p(k_j \mid x) \right\}$$

$$= \arg\min_{\substack{k_i \\ 1 \le i \le N}} \left\{ \sum_{j=1}^{N} C(k_i, k_j)\, \frac{p(x \mid k_j)\, p(k_j)}{p(x)} \right\}$$

$$= \arg\min_{\substack{k_i \\ 1 \le i \le N}} \left\{ \sum_{j=1}^{N} C(k_i, k_j)\, p(x \mid k_j)\, p(k_j) \right\}$$

Klasyfikator bayesowski - modyfikacja

W przypadku macierzy kosztów na diagonali możemy mieć 0 (poprawne przypisanie do klasy), a poza nią wartości dodatnie.

W przypadku, gdy

$$C(k_i, k_j) = 1 - \delta(i, j)$$

(δ - delta Kroneckera) to nasz problem sprowadza się do klasycznego przypadku znalezienia maksimum funkcji

$$f(x) = \arg\max_{k_i} p(k_i \mid x)$$

Klasyfikator bayesowski - modyfikacja

W przypadku zastosowania tzw. „rejection option” dodajemy klasę k_0

$$k = k_0 \qquad C_{rej} < e(x)$$

$$k = \arg\max_{\substack{k_i \\ 1 \le i \le N}} p(k_i \mid x) \qquad C_{rej} \ge e(x)$$

gdzie

$$e(x) = 1 - \max_{\substack{k_i \\ 1 \le i \le N}} \{ p(k_i \mid x) \}$$

C_{rej} jest to koszt odrzucenia

Klasyfikator bayesowski - modyfikacja

Klasyfikator naiwny – zakładamy wzajemną niezależność atrybutów

$$p(x_i \mid K_j) = \prod_{k=1}^{N} p(a_{ik} \mid k_j) = p(a_{i1} \mid k_j) \cdot p(a_{i2} \mid k_j) \cdot \ldots \cdot p(a_{iN} \mid k_j)$$

Zaleta – redukcja wymiaru

Klasyfikator bayesowski - modyfikacja

Przekleństwo wymiarowości (ang. *curse of dimensionality*) - im więcej wymiarów tym bardziej unikalne stają się posiadane dane.

Można stosować klasyfikatory bayesowskie łączące dane zależne (pośrednie rozwiązanie)

POLITECHNIKA GDAŃSKA

Estymacja rozkładu

- **Histogram**
- **Estymator jądrowy**

Estymator jądrowy

Niech $(y_1, \ldots y_L)$ będzie próbką pobraną z rozkładu o nieznanej gęstości f, gdzie L to wielkość próby prostej pobranej z rozkładu. Wartość funkcji gęstości prawdopodobieństwa estymowanej funkcją gaussowską dla punktu wynosi

$$g(y,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

Klasyfikator bayesowski – estymacja rozkładu

Typowo w naiwnym klasyfikatorze bayesowskim zakłada się, ze każda klasa ma rozkład normalny

Odchylenie standardowe i wartość oczekiwaną estymuje się z próbki

$$\hat{f}(y) = \frac{1}{L}\sum_{i=1}^{L} g(y, y_i, \sigma_k)$$

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Klasyfikator NS poszukuje najbliższego wektora ze zbioru wektorów uczących.

Wynik zależy od sposobu liczenia odległości.

Modyfikacja - kNS - wybiera się odległości do k najbliższych sąsiadów.

Podobnie działa klasyfikator najbliższego centrum (NC).

Klasyfikator najbliższego sąsiada (NS)

Przykładowe miary odległości

- **odległość Czebyszewa**
- **odległość miejska**
- **odległość euklidesowa**
- **odległość Minkowskiego**
- **odległość Hamminga**
- **odległość Mahalanobisa**

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Odległość Czebyszewa (maksimum)

$$d = \max_{1 \le i \le n} |x_i - y_i|$$

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Odległość miejska

(taksówkowa, Manhattan distance, city-block)

$$d = \sum_{i=1}^{n} \left| x_i - y_i \right|^2$$

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Odległość euklidesowa

$$d = \sqrt{\sum_{i=1}^{n} \left| x_i - y_i \right|^2}$$

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Odległość Minkowskiego

$$d = \left(\sum_{i=1}^{n} |x_i - y_i|^q \right)^{1/q}$$

q=1, q=2, q=∞ – szczególne przypadki

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Odległość Hamminga

Określa na ilu pozycjach różnią się dwa wektory

Przykładowo

d(0011,1100) = 4, d(122, 220) = 3.

Można także stosować wektory literowe itp.

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Odległość Mahalanobisa

$$d = \sqrt{(x-y)^T S^{-1} (x-y)}$$

gdzie S jest macierzą kowariancji

Klasyfikator najbliższego sąsiada (NS)

POLITECHNIKA GDAŃSKA

Można warzyć współrzędne

$$d = \left(\sum_{i=1}^{n} w_i \left| x_i - y_i \right|^q \right)^{1/q}$$

Można skalować współrzędne

$$a' = (a - \min_A(a)) / (\max_A(a) - \min_A(a))$$

Klasyfikacja – wykrywanie anomalii

POLITECHNIKA GDAŃSKA

Detekcja anomalii czy też nowości (ang. *outlier detection*, *novelty detection*) pozwala wykryć dane w znacznym stopniu odmienne, wyjątkowe i niespójne z większością danych wejściowych.

Klasyfikacja – wykrywanie anomalii

POLITECHNIKA GDAŃSKA

Metody

- Statystyczne
- Odległościowe
- Na podstawie gęstości
- inne

POLITECHNIKA GDAŃSKA

Statystyczne wykrywanie anomalii

Bazują na rozkładzie prawdopodobieństwa konstruowanym w sposób parametryczny (zakładamy rozkład) bądź nieparametryczny

Trening – budowa modelu

Przykładem metody nieparametrycznej jest wykorzystanie histogramu do określenia prawdopodobieństwa wystąpienia danych

Problem przy większej liczbie wymiarów

POLITECHNIKA GDAŃSKA

Statystyczne wykrywanie anomalii

Dla danych składających się z wielu atrybutów można wyznaczyć *próg anomalii* (ang. *outlier score*).

Niskie prawdopodobieństwo występowania danej instancji oznacza wysoki *próg anomalii*.

$$próg = \frac{1}{\#F} \sum_{f \in F} w_f (1 - p_f)$$

gdzie w_f oznacza wagę nadaną każdemu atrybutowi , p_f oznacza prawdopodobieństwo a priori wystąpienia wartości atrybutu , a F zbiór atrybutów

Odległościowe wykrywanie anomalii

Można wykorzystać ideę najbliższego sąsiedztwa badanej instancji.

Brak założeń dotyczących rozkładu danych, na których operujemy.

Problem – dobór miary odległości.

POLITECHNIKA GDAŃSKA

Odległościowe wykrywanie anomalii

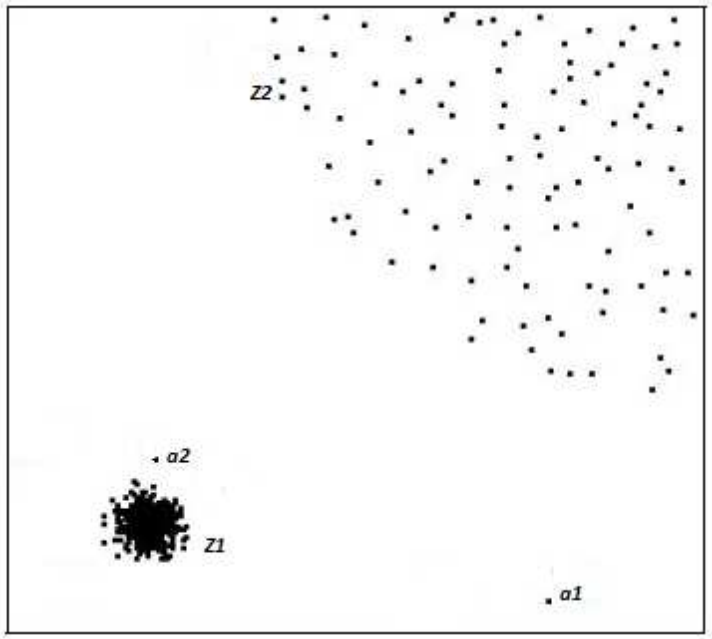
Przykładowa metoda – DB(k, λ) – Outlier

Instancję x ze zbioru testowego zalicza się jako nietypową jeśli nie więcej niż k instancji treningowych jest w odległości λ lub mniejszej od x.

Modyfikacja – metoda DB(pct, d_{min})

Instancje zalicza jako nietypową, jeśli co najmniej pct% instancji w zbiorze treningowym jest w odległości większej niż d_{min} od tej instancji.

Klasyfikacja – wykrywanie anomalii



Z1 **- 400 obiektów**
Z2 **– 100 obiektów**

a1,a2 **- anomalie**

Problem - tylko *a1* jest anomalią albo cały *Z2* to anomalie

Nie da się dobrać współczynników tak, aby tylko *a1* i *a2* były anomaliami

POLITECHNIKA GDAŃSKA

Wykrywanie anomalii na podstawie gęstości

Przykład - metoda LOF (ang. *local outlier factor*)

Nadaje każdej instancji stopień, w jakim można ją uważać za anomalię. Metoda ta sprawdza jak bardzo odizolowana jest instancja od otaczającego ją sąsiedztwa.

Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, Jörg Sander, LOF: Identifying Density-Based Local Outliers,

Proc. ACM SIGMOD 2000 Int. Conf. On Management of Data, Dalles, TX, 2000

LOF – local outlier factor

$$LOF_k(p) = \frac{\#N_k(p)}{\sum_{a \in N_k(p)} \frac{lrd_k(a)}{lrd_k(p)}}$$

gdzie:

- $lrd_k(a)$ - **local reachability density punktu p - określa zagęszczenie w k-sąsiedztwie obiektu a**
- $lrd_k(p)$ - **local reachability density punktu p - określa zagęszczenie w k-sąsiedztwie obiektu p**
- $\#N_k(p)$ - **liczba elementów w k-sąsiedztwie punktu p**

POLITECHNIKA GDAŃSKA

LOF – local outlier factor

$$lrd_k(p) = \frac{\#N_k(p)}{\sum_{o \in N_k(p)} reach - dist_k(p,o)}$$

$$reach - dist_k(p,o) = \max\{k - distance(o), dist(p,o)\}$$

POLITECHNIKA GDAŃSKA

LOF – local outlier factor

- metoda lokalna
- brak kategoryzacji
- jeżeli nowy pomiar jest "bliski" pomiarom treningowym, to LOF~1
- ocena "jak mocno" dany obiekt jest "odbiegający"

Klasyfikacja – przykłady wyników

POLITECHNIKA GDAŃSKA

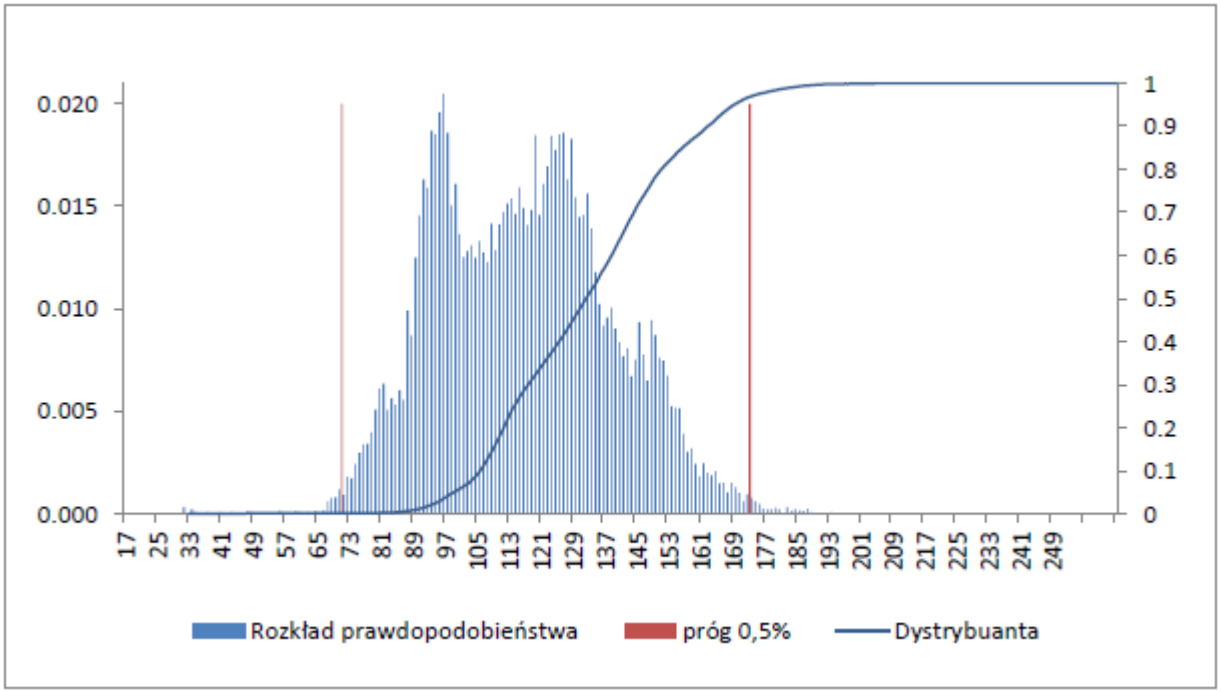
Przykładowe wyniki

Dane typowe – $[param \cdot \sigma - \mu, param \cdot \sigma + \mu]$,

$param = 1,2,3$

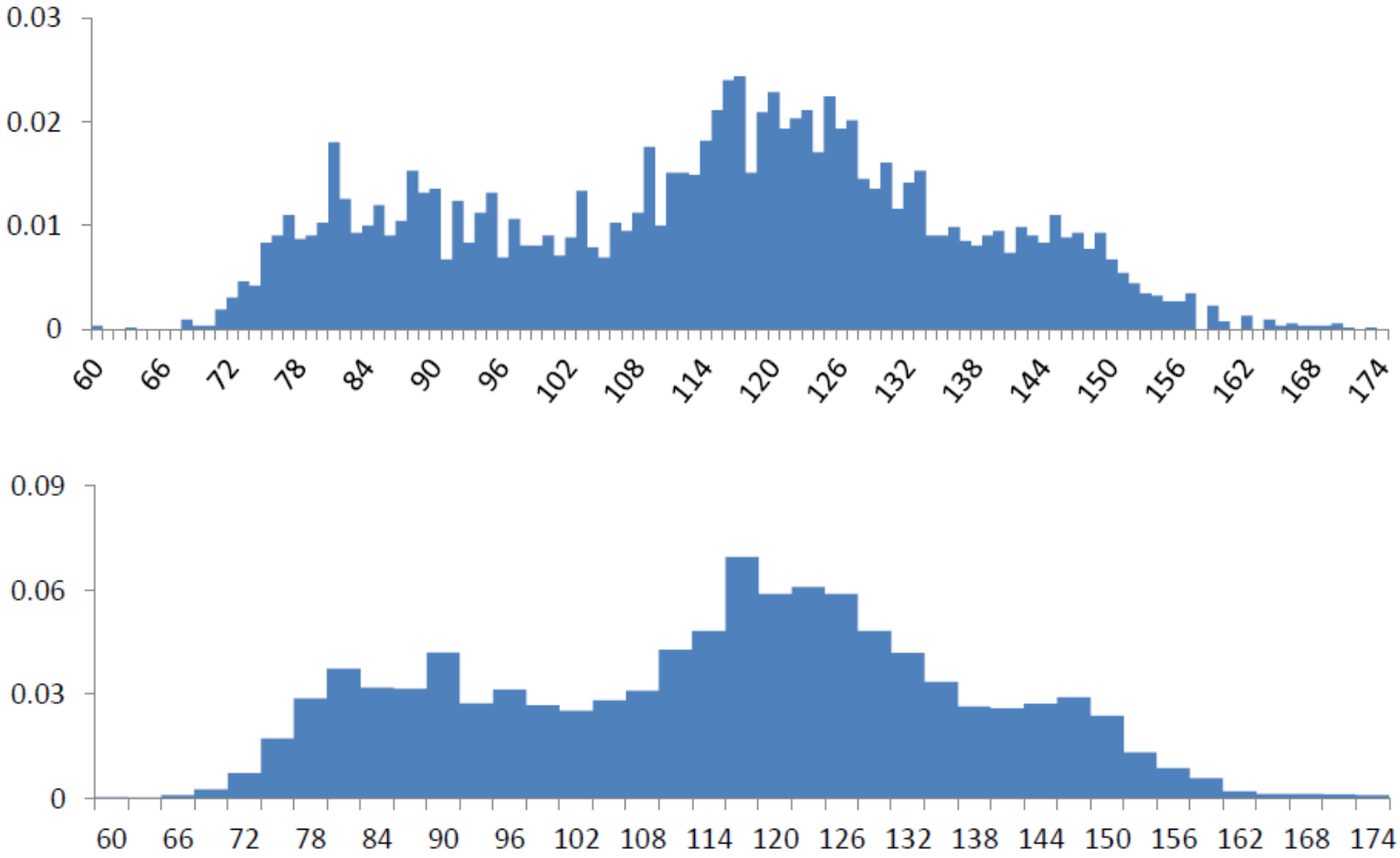
param	instancje stabilne (% ogółu)	instancje niestabilne (% ogółu)
1	21185 (82,3%)	4554 (17,7%)
2	25529 (99,2%)	211 (0,8%)
3	25724 (99,9%)	16 (0,1%)

Klasyfikacja – przykłady wyników



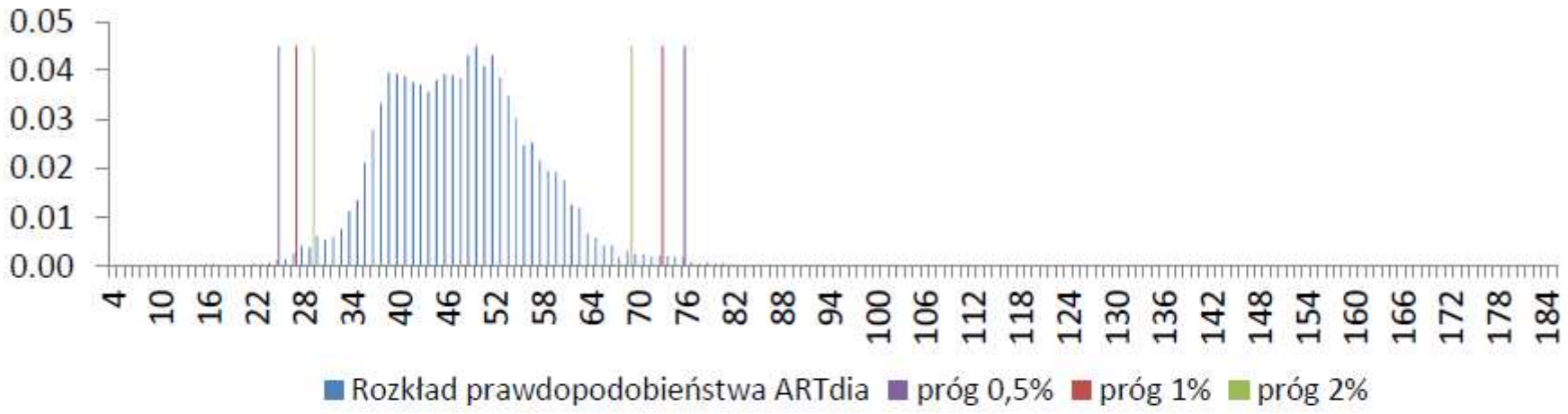
Rozkład prawdopodobieństwa tętniczego ciśnienia skurczowego (skala na lewej osi) wraz z dystrybuantą (skala na prawej osi) oraz progiem wyznaczonym na jej podstawie.

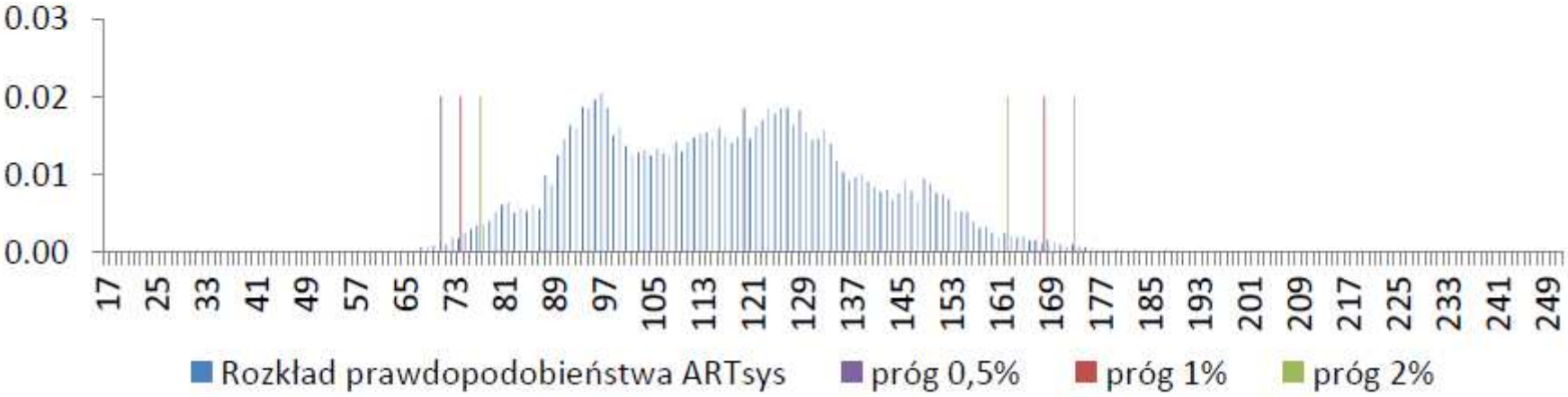
Klasyfikacja – przykłady wyników



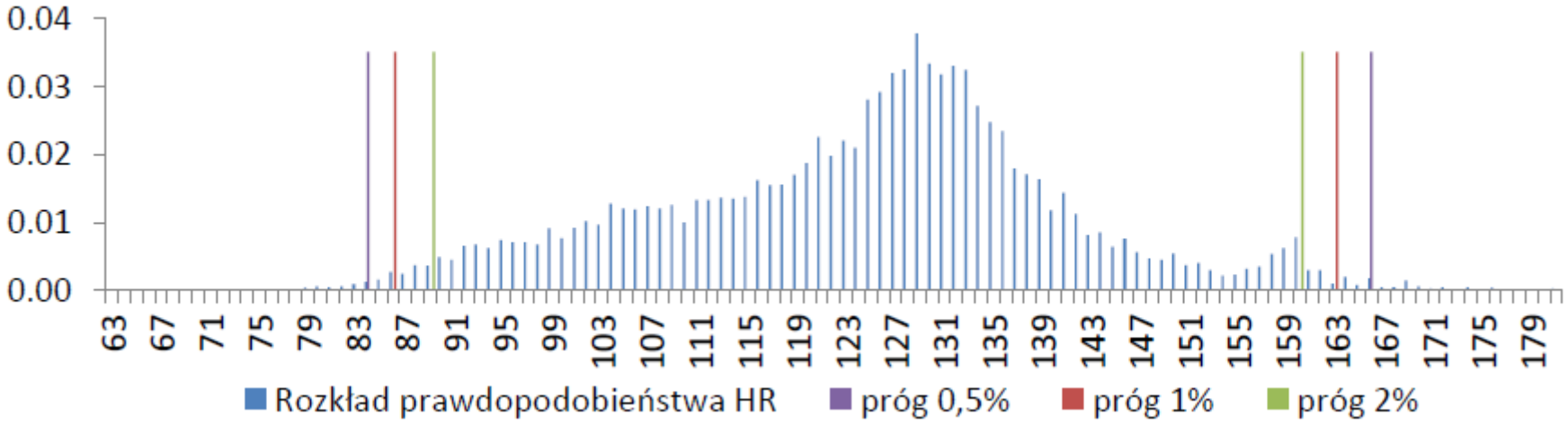
Przykład wpływu szerokości przedziału na estymowany rozkład prawdopodobieństwa dla tętniczego ciśnienie skurczowego

Klasyfikacja – przykłady wyników





Klasyfikacja – przykłady wyników



Klasyfikacja – przykłady wyników

POLITECHNIKA GDAŃSKA

Przykładowe miary jakości klasyfikacji

$$\text{TP rate} = TP/(TP+FN)$$

$$\text{TN rate} = TN/(TN+FP)$$

Klasyfikacja – przykłady wyników

POLITECHNIKA GDAŃSKA

próg [%]	param	Dokładność klasyfikacji [%]	TN rate	TP rate
0.5	1	87%	90%	71%
	2	92%	92%	78%
	3	94%	94%	56%
		91%	**92%**	**68%**
1	1	87%	90%	73%
	2	91%	91%	75%
	3	93%	93%	38%
		90%	**91%**	**62%**
2	1	87%	90%	74%
	2	90%	90%	70%
	3	93%	93%	36%
		90%	**91%**	**60%**

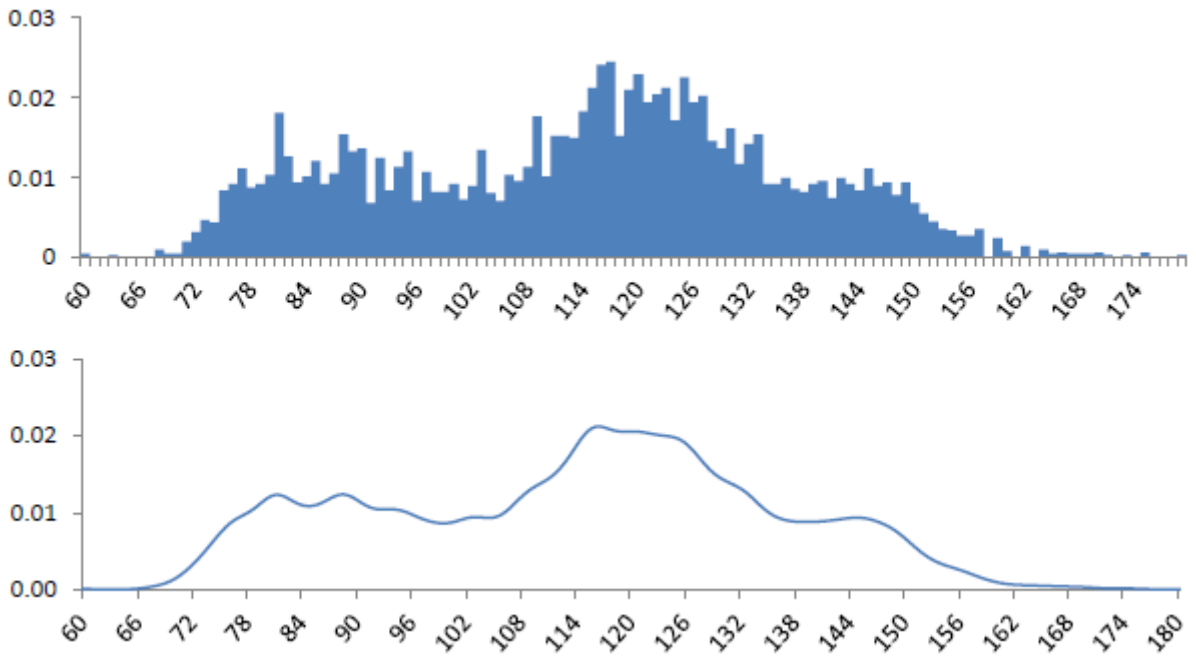
Wyniki działania naiwnego klasyfikatora bayesowskiego dla rozkładów prawdopodobieństwa z różnym umiejscowieniem skrajnych progów cięcia dla trzech zbiorów instancji

Klasyfikacja – przykłady wyników

Nazwa zestawu obowiązująca w programie	param					
	1		2		3	
	TP rate	TN rate	TP rate	TN rate	TP rate	TN rate
$h_{ARTdia} = 2, h_{ARTsys} = 2, h_{EtCO2} = 0.04, h_{HR} = 2, h_{RR} = 1$	72%	90%	85%	90%	46%	91%
"*Optymalne szerokości*"	72%	90%	85%	90%	43%	91%
$h_{ARTdia} = 2, h_{ARTsys} = 4, h_{EtCO2} = 0.05, h_{HR} = 3, h_{RR} = 1$	72%	90%	86%	91%	49%	91%
$h_{ARTdia} = 3, h_{ARTsys} = 5, h_{EtCO2} = 0.06, h_{HR} = 4, h_{RR} = 1$	72%	90%	86%	91%	49%	91%
$h_{ARTdia} = 4, h_{ARTsys} = 6, h_{EtCO2} = 0.07, h_{HR} = 5, h_{RR} = 1$	71%	90%	86%	92%	55%	91%
$h_{ARTdia} = 5, h_{ARTsys} = 7, h_{EtCO2} = 0.08, h_{HR} = 6, h_{RR} = 1$	72%	90%	85%	92%	47%	91%

Wpływ estymacji rozkładu na wyniki klasyfikacji naiwnym klasyfikatorem bayesowskim

Klasyfikacja – przykłady wyników



Rozkład prawdopodobieństwa dla tętniczego ciśnienia skurczowego oraz gęstość prawdopodobieństwa tego rozkładu estymowana z wykorzystaniem estymatora jądrowego.

Klasyfikacja – przykłady wyników

param	Dokładność klasyfikacji	TN rate	TP rate
1	76%	75%	84%
2	87%	87%	56%
3	100%	100%	0%

Wyniki klasyfikacji naiwnym klasyfikatorem bayesowskim z wykorzystaniem rozkładów prawdopodobieństwa otrzymanych metodą estymacji jądrowej.

Klasyfikacja – przykłady wyników

Klasyfikator	Nazwa zestawu obowiązująca w programie	param	Dokładność	TN rate	TP rate
NBC	*"Optymalne szerokości"*	1	87.0%	90.2%	71.8%
		2	89.8%	89.8%	84.6%
		3	91.0%	91.2%	43.2%
NBC z połączonymi ARTdia i ARTsys	*"Optymalne szerokości"*	1	81.6%	89.2%	50.0%
		2	94.8%	95.0%	74.6%
		3	95.8%	95.8%	45.2%
	$"h_{ARTdia} = 5, h_{ARTsys} = 7, h_{EtCO2} = 0.08, h_{HR} = 6, h_{RR} = 1"$	1	84.2%	90.0%	58.2%
		2	94.2%	94.2%	79.6%
		3	94.8%	94.8%	42.8%

Wyniki klasyfikacji dla naiwnego klasyfikatora bayesowskiego oraz naiwnego klasyfikatora bayesowskiego wykorzystującego łączny rozkład prawdopodobieństwa dla ciśnienia rozkurczowego oraz skurczowego.

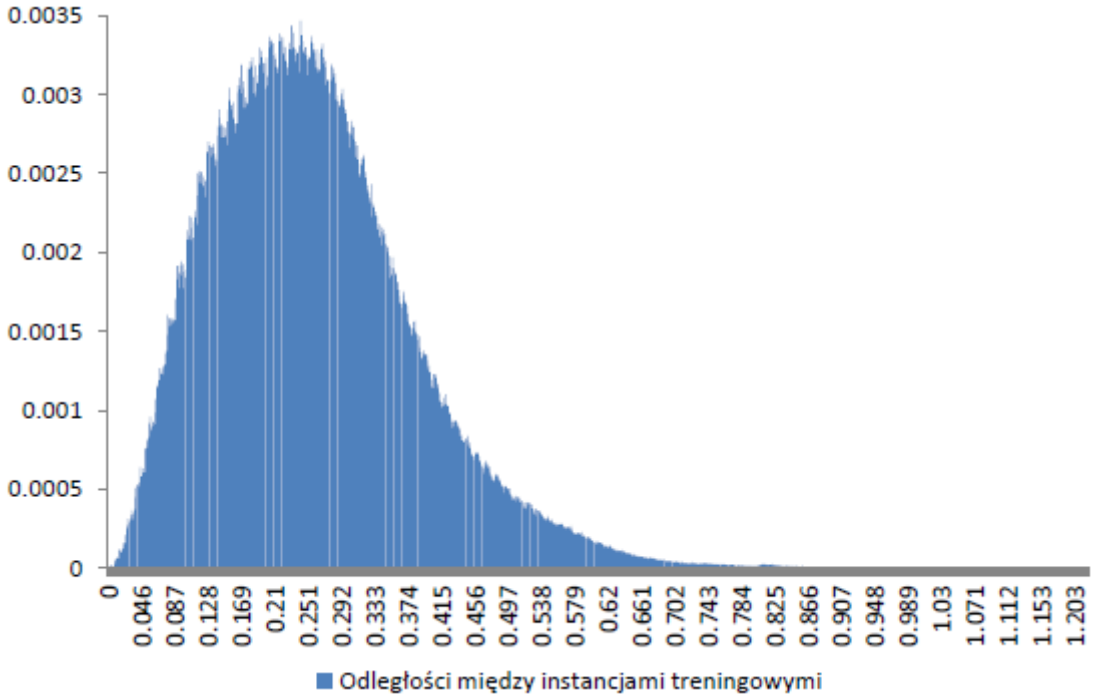
Klasyfikacja – przykłady wyników

Liczba najbliższych sąsiadów k	param	Dokładność klasyfikacji	TN rate	TP rate
1	1	86.4%	97.4%	37.0%
	2	99.2%	100.0%	9.8%
	3	100.0%	100.0%	4.4%
Średnia dla $k = 1$		**95.2%**	**99.1%**	**17.1%**
3	1	86.4%	98.4%	29.2%
	2	99.2%	100.0%	4.2%
	3	100.0%	100.0%	1.6%
Średnia dla $k = 3$		**95.2%**	**99.5%**	**11.7%**
5	1	86.2%	98.6%	28.0%
	2	99.2%	100.0%	3.8%
	3	100.0%	100.0%	0.0%
Średnia dla $k = 5$		**95.1%**	**99.5%**	**10.6%**

Dokładność klasyfikacji algorytmem k-NN dla różnej ilości najbliższych sąsiadów k.

Problem – liczba instancji niestabilnych ze wzrostem k

Klasyfikacja – anomalie - przykłady



Częstość występowania euklidesowych odległości między instancjami w zbiorze treningowym.

Klasyfikacja – anomalie - przykłady

POLITECHNIKA GDAŃSKA

Próg Anomalii	Procent wykrytych anomalii w zbiorze testowym
0.98	7.93%
0.99	1.84%
0.995	0.52%

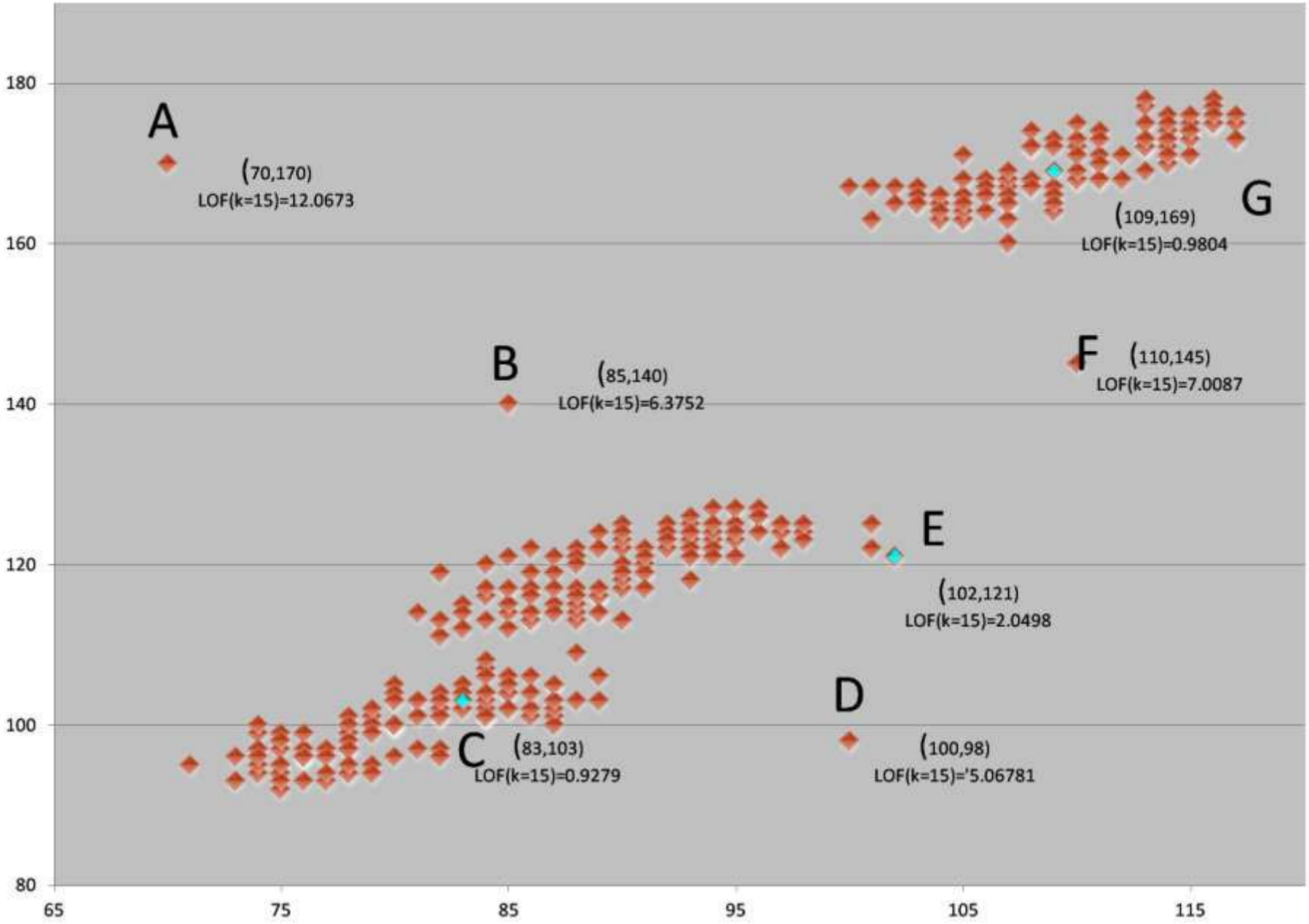
Procent anomalii wykrytych w zbiorze instancji testowych dla różnych wartości *progu anomalii*

Klasyfikacja – anomalie - przykłady

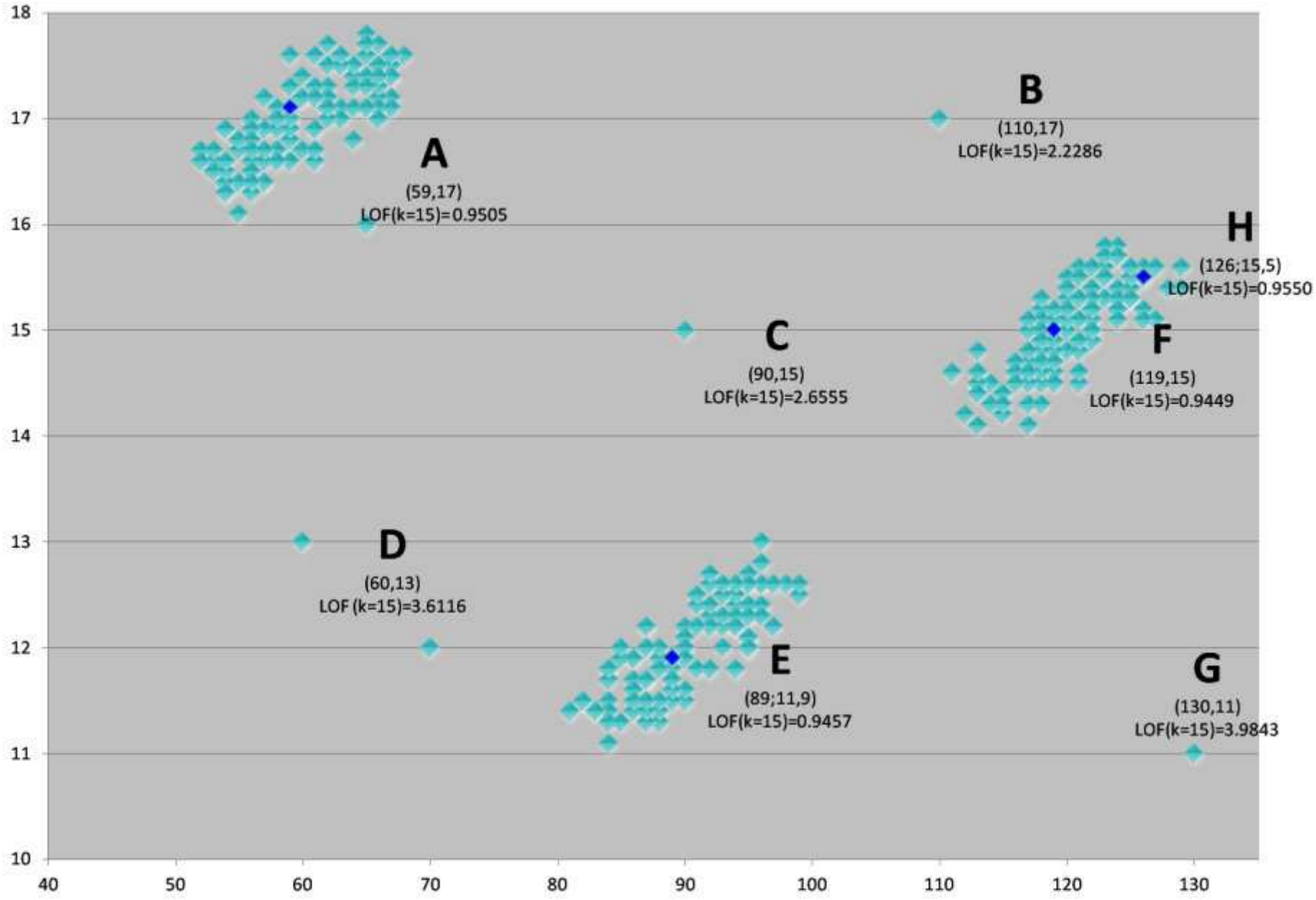
d_{Min}	pct	Procent wykrytych anomalii
0.8	98.0%	14.67%
	99.0%	21.90%
	99.5%	30.27%
1	98.0%	2.40%
	99.0%	9.65%
	99.5%	15.18%

Procent anomalii wykrytych w zbiorze instancji testowych dla różnych wartości d_{min} oraz pct

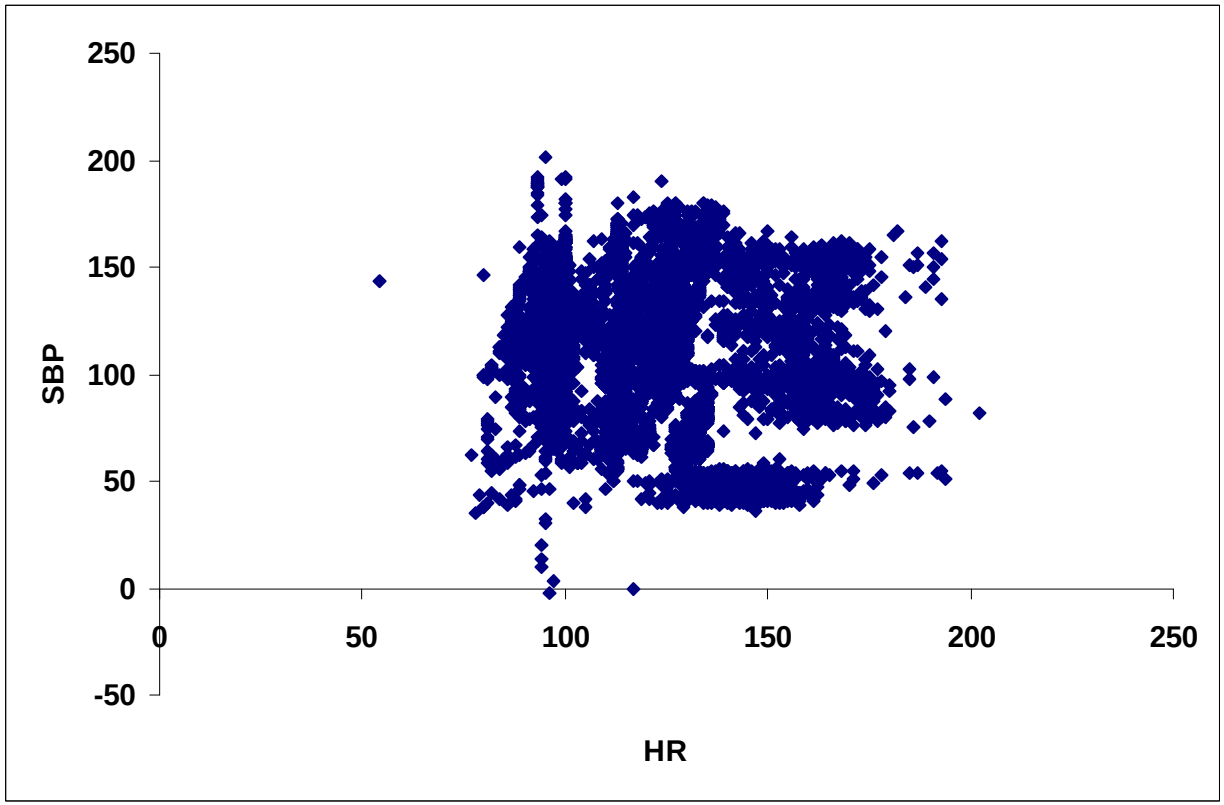
Klasyfikacja – anomalie - przykłady



Klasyfikacja – anomalie - przykłady



Klasyfikacja – anomalie - przykłady



POLITECHNIKA GDAŃSKA

Wnioski z symulacji LOF

- jeżeli pomiar jest „w normie”, to klasyfikator LOF~1

-duże wartości LOF (>3) wskazują na mocno odbiegający pomiar

- LOF w przedziale (2,3) – reguły decyzyjne

Klasyfikacja

POLITECHNIKA GDAŃSKA

Eksperymenty dotyczące klasyfikatorów bayesowskich pokazały duże znaczenie sposobu estymacji rozkładu prawdopodobieństwa.

Problemem może też być liczba parametrów wykorzystywanych w klasyfikatorze.

Dla modelowanych danych dla klasyfikatora k-NN instancji o klasie niestabilnej było zbyt mało, aby skutecznie wykorzystywać go do identyfikacji stanów zagrożonych.

Algorytmy wykrywające anomalie nie muszą mieć sklasyfikowanych instancji, ale pojawia się problem ustawienia odpowiednich progów detekcji stanów nietypowych.

Klasyfikacja – problemy praktyczne

POLITECHNIKA GDAŃSKA

- **nieregularne/niedokładne pomiary**
- **wpływ rytmu dobowego/tygodniowego/...**
- **nieregularne przyjmowanie leków**
- **zróżnicowana aktywność dzienna**
- **możliwa zmiana stanu pacjenta w czasie (np. zmiana leków)**

POLITECHNIKA GDAŃSKA

Dziękujemy za uwagę