



Dawid Kaliszewski

Modelowanie glikemii w procesie insulinoterapii

Promotor

dr hab. inż. Zenon Gniazdowski

Cel pracy

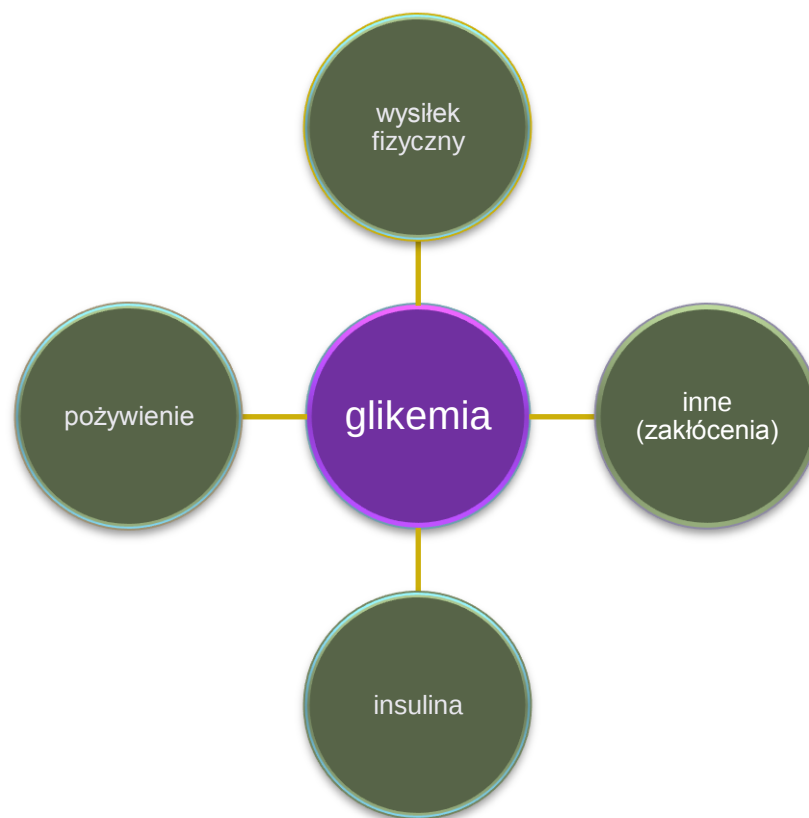
- Zbudowanie modelu predykcyjnego przyszłych wartości glikemii diabetyka leczonego za pomocą insulinoterapii.
- Model musi uwzględniać istotne w procesie czynniki, być odporny na zakłócenia i jak najbardziej trafny.
- Gdyby się udało, można by trafniej dobierać dawki insuliny.

Etapy pracy

1. Analiza zagadnienia i wybór metody modelowania;
2. Projekt i implementacja algorytmów przygotowania danych medycznych do modelowania;
3. Opracowanie słowników danych pomocniczych;
4. Implementacja algorytmów numerycznych;
5. Budowa różnych modeli i ocena ich trafności.

Modelowane zjawiska

- Na glikemię mają wpływ różne czynniki.
- Dynamika procesów – efekty następują po czasie
- Zjawiska ciągłe, ale dyskretne dane pomiarowe.



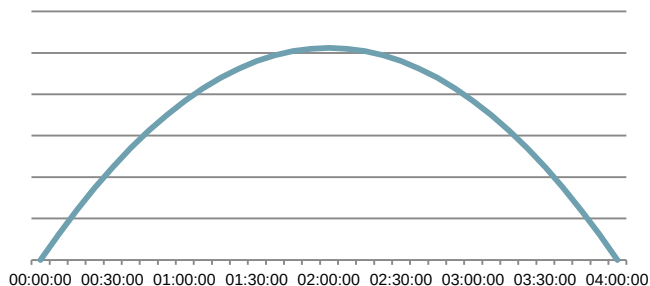
Istotne czynniki

Na podstawie analizy zagadnienia wybrano 17 potencjalnie istotnych czynników, należących do nich:

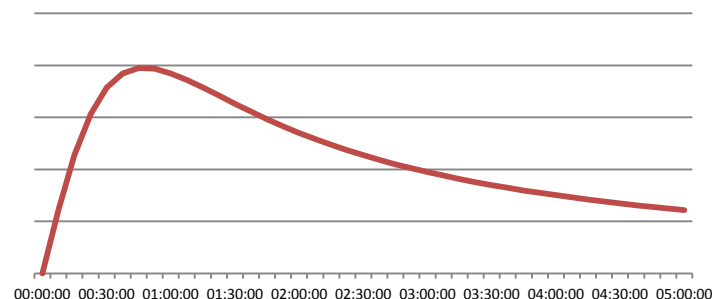
- związane z pożywieniem (masa, wartość energetyczna, masa poszczególnych składników pokarmu: białka, tłuszczu, węglowodanów);
- związane z wysiłkiem fizycznym – czas trwania i rodzaj;
- dawka insuliny – wielkość i rodzaj;
- poprzednie wartości glikemii;
- czas od wystąpienia czynnika do wynikowego pomiaru glikemii (Δt).

Wpływ czynników na zjawisko

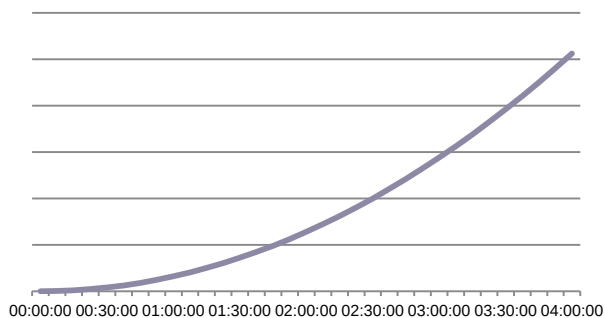
siła działania insuliny w czasie



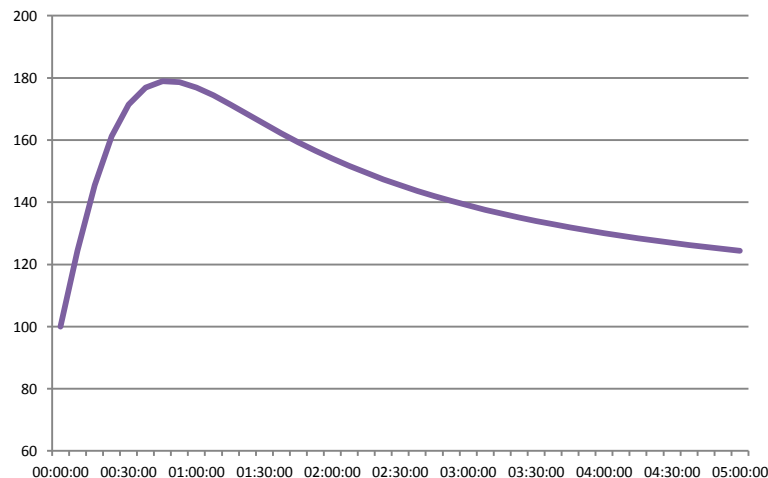
energia uwalniana z posiłku



zużycie energii podczas wysiłku fizycznego



glikemia



Dane pacjenta

Data	Godzina	Glikemia	Nazwa produktu	Ilość	Masa (g)	Rodzaj wysiłku	Czas trwania	Dawka
2011-06-01	01:10:00	95						
	07:45:00	115	ryż biały		112			10
			mleko		298			
	09:40:00	203						4
	11:05:00	119	serek wiejski		200			3
			wafle ryżowe	1 szt.				
	12:35:00	201						10
	13:00:00		serek waniliowy		150			
			napój XL		250			
			jabłko		200			
	15:55:00	177				rower 20km/h	115	
	16:40:00	70	batonik muesli	2 szt.				
	18:40:00	135	ryż biały		220			13
			kotlet schabowy		140			
			marchew z groszkiem		220			
2011-06-02	21:55:00	148	banan		220			6
	23:35:00	123						
	00:45:00	148						
	08:00:00	169	ryż biały		160			11
			mleko		180			
	09:30:00	90						
	10:00:00		Prince Polo XXL		50			5
			napój XL		250			
	10:30:00	75						
	11:30:00	208						3
	12:50:00	231						7
	13:20:00		serek wiejski		200			
			wafle ryżowe	1 szt.				
	14:25:00	197						12
	15:00:00		makaron		160			
			kotlet mielony		140			
			papryka czerwona		240			
	15:50:00	162				rower 20km/h	70	
	18:00:00		chleb jasny		80			12
			masło		16			
			ser żółty		56			
			kiełbasa zwyczajna		48			
	19:40:00	66	sok pomarańczowy		220			
	22:15:00	114						



Przygotowanie danych liczbowych

- Analizowane dane pacjenta mają postać liczbowo-opisową i zawierają błędy oraz braki. Nie da się ich analizować metodami matematycznymi w oryginalnej postaci.
- Konieczne jest ich przetworzenie do postaci liczbowej za pomocą szeregu procesów, np.:
 1. uzupełnienie brakujących atrybutów za pomocą danych słownikowych;
 2. zagregowanie wartości czynników występujących w jednej chwili;
 3. uzupełnienie brakujących wartości pomiarów glikemii;
 4. obliczenie Δt między czynnikami wejściowymi a pomiarem.
- Niezbędne opracowanie słownika danych o wartościach odżywczych produktów spożywczych.

Analiza regresji

- Poszukiwanie modelu danego ogólnym wzorem:

$$f(\mathbf{x}, \mathbf{b}) = b_0 + \sum_{i=1}^k b_i g(x_i) + \sum_{i=1, j=1, i \neq j}^k b_{ij} g(x_i, x_j)$$

- Po dobraniu struktury modelu należy na podstawie macierzy zmiennych objaśniających $\mathbf{X}$ oraz wektora zmiennej objaśnianej $\mathbf{y}$ wyznaczyć wektor parametrów $\mathbf{b}$ za pomocą metody najmniejszych kwadratów:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

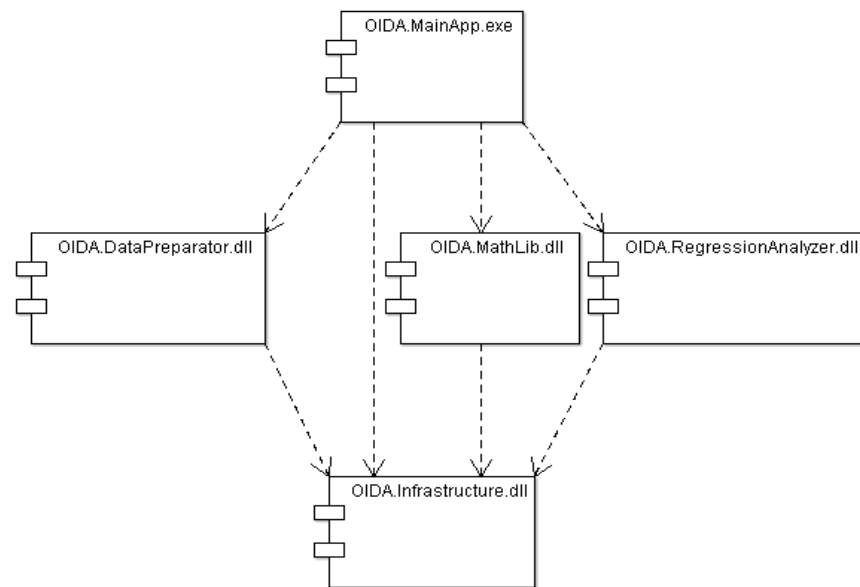
- Ocena istotności modelu oraz parametrów za pomocą testów statystycznych, a następnie odrzucenie najmniej istotnego parametru – regresja krokowa wsteczna.

Redukcja wymiaru

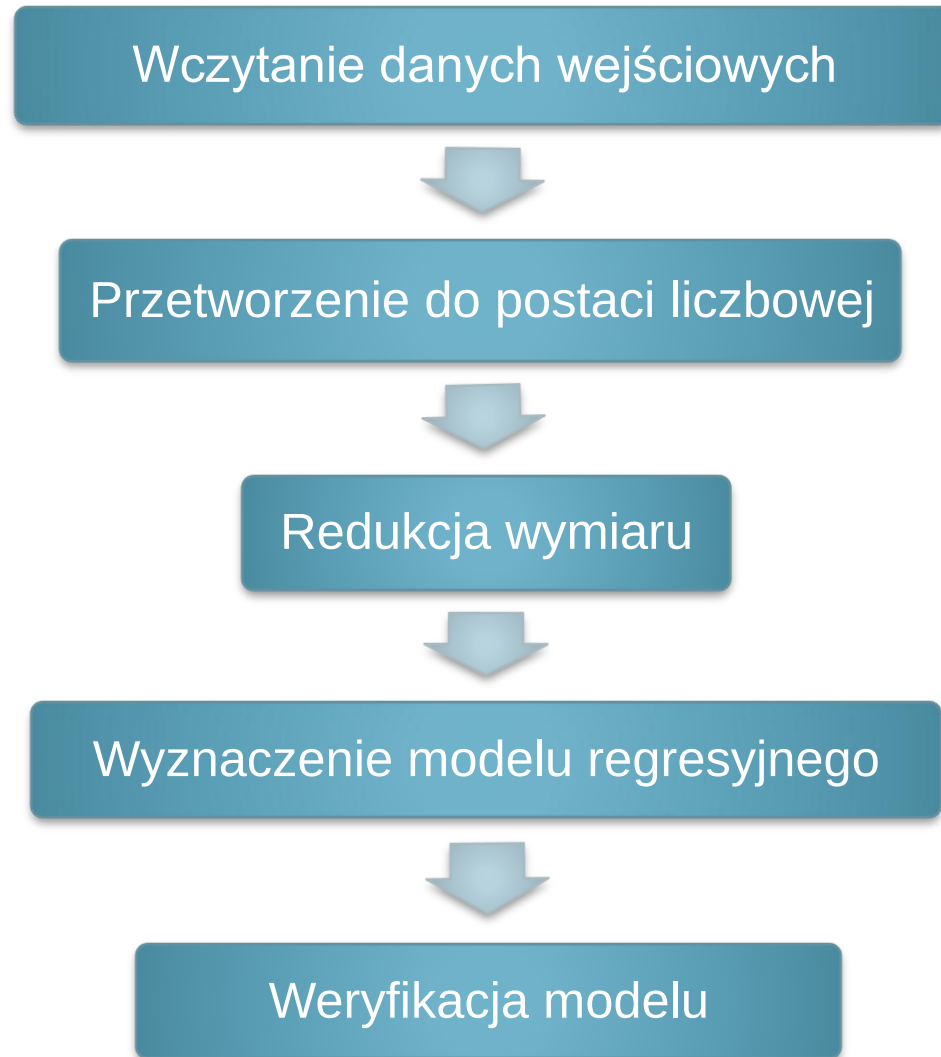
- Istnienie wielu zmiennych objaśniających może oznaczać skomplikowany i czasochłonny w wyznaczeniu model.
- Zakładając liniową zależność między zmiennymi, można ich liczbę zredukować za pomocą analizy głównych składowych. Dzięki temu można uprościć obliczenia (kosztem utraty dokładności).
- Do wyznaczenia układu własnego macierzy współczynników korelacji między zmiennymi objaśniającymi zastosowano metodę Jacobiego.

Implementacja aplikacji

- Każdy moduł spełnia inne zadanie.
- Możliwość oddzielnego wykorzystania każdego modułu lub efektu jego pracy.



Przebieg modelowania



Analizowane modele

- Liniowy – najprostsz;
- Kwadratowy – dodatkowo wzajemne zależności między zmiennymi objaśniającymi;
- Czasowy – szczególnie uwzględniający wpływ Δt między wystąpieniem czynnika, a wynikowym pomiarem;
- Ekspercki – bazujący na wiedzy dziedzinowej, nie zawierający wszystkich możliwych powiązań, a jedynie te teoretycznie istotne.

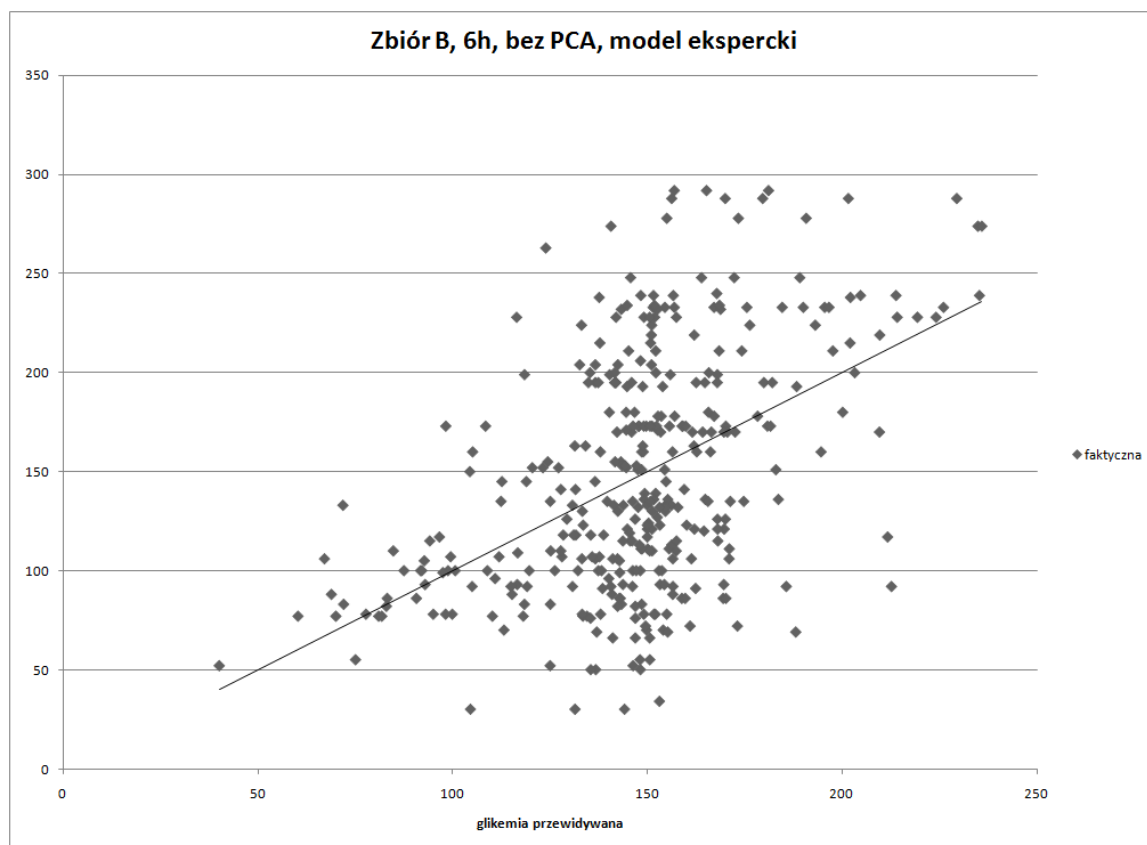
Przebieg eksperymentu

1. Dwa zbiory danych medycznych (10 dni i miesiąc);
2. Dwie wartości Δt_{max} (3h i 6h);
3. Bez PCA, albo dwa sposoby doboru zmiennych objętych PCA;
4. Cztery typy modeli.

W efekcie 48 różnych kombinacji – potencjalnych modeli do zweryfikowania.

Wyniki eksperymentu

- Najlepszy współczynnik determinacji modelu wyniósł zaledwie 0,29.
- Znaczny rozrzut między wartościami przewidywanymi i zmierzonymi.



Podsumowanie i wnioski

- Zostały wykonane moduły aplikacji:
 1. przetwarzająca dane do postaci liczbowej;
 2. wyznaczająca modele regresyjne.
- Niestety żaden model badanego zjawiska nie okazał się na tyle dobrze dopasowany, aby móc go zastosować w praktyce.
- Co dalej?
 1. bardziej skomplikowane modele (*ln*, *exp*, *sqrt*, itd.);
 2. zastosowanie metod z teorii układów dynamicznych;
 3. metody sztucznej inteligencji – SSN.



Dziękuję za uwagę