



INSTYTUT SYSTEMÓW INFORMATYCZNYCH
ul. S. Kaliskiego 2, 00-908 Warszawa, tel. (+4822) 68 39 50 4, fax (+4822) 68 37 8 5 8

WYDZIAŁ CYBERNETYKI

WOJSKOWA AKADEMIA TECHNICZNA

<http://www.isi.wat.edu.pl>

Komputerowa diagnoza medyczna – tworzenie i interpretowanie

prof. dr hab. inż. Andrzej Walczak



<http://www.isi.wat.edu.pl>



Agenda

1. Po co budujemy komputerowe wspomaganie diagnostyki medycznej?
2. Wymagania na IT wdrażane w medycynie
3. Przypisanie przypadku diagnozowanego do choroby – klasyfikacja
4. Choroby skóry wg IDC i ograniczenia klasyfikacji przypadku w medycynie
5. Przykładowe modele klasyfikacji
6. Wyniki i ich kliniczna weryfikacja
7. Wnioski



Cele IT w diagnostyce medycznej

1. **Wspomaganie interpretacji danych** (analiza obrazów medycznych, badania genotypu, itp.)
2. Rozwój aparatury medycznej i jej obsługi – **przetwarzanie danych pomiarowych** (pomiar FFR w kardiologii, zdalny pomiar EKG, itp., TELEMEDYCYNA)
3. A może „kiosk medyczny” w przyszłości?

Powstaje pytanie: jaki cel mają aplikacje wspomagające diagnozowanie w codziennej praktyce lekarskiej w przychodniach?

Odpowiedź: trudno nazwać ten cel badawczym. Raczej ma on **charakter czysto utylitarny**.

Wymagania na IT wdrażane w medycynie

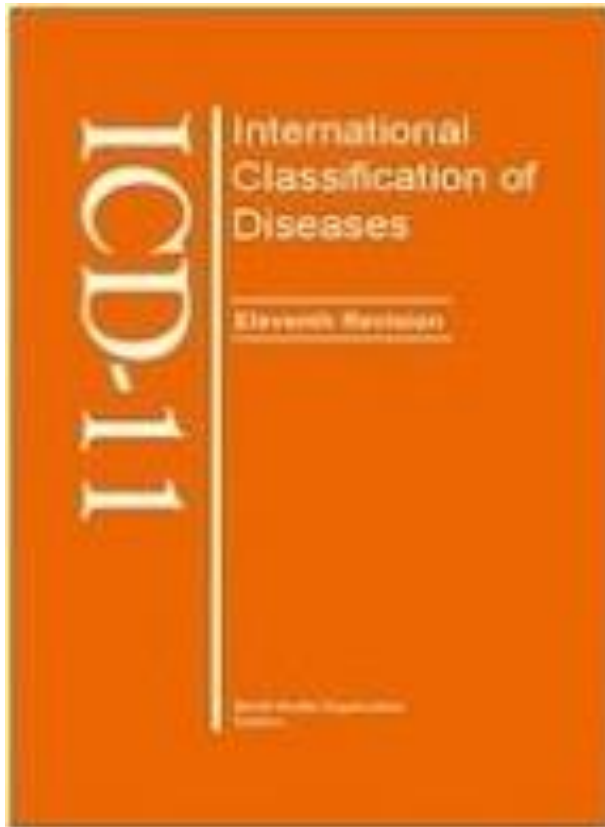
Dokumentacja techniczna do certyfikowania wyrobu medycznego musi zawierać:

- wyniki analizy ryzyka
- wyniki oceny klinicznej

Wobec tego aplikacja IT wspomagająca diagnozowanie musi posiadać rodzaj „tabliczki znamionowej”, która poinformuje lekarza jakie jest ryzyko błędnej diagnozy przy jej stosowaniu.



Wymagania na IT wdrażane w medycynie



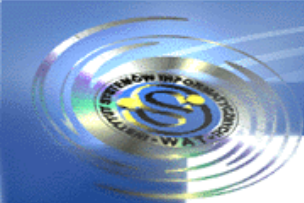


Wymagania na IT wdrażane w medycynie

Definitions

All ICD entities will have definitions:

- that give key descriptions and guidance on what the meaning of the entity/category is
- in human readable terms - to guide users
- All ICD entities will have a concise definition (100 words) in the print version
- More detailed definitions can be found ONLINE
- **the Content Model**
- **Diagnostic Criteria**
- across the whole classification and the versions

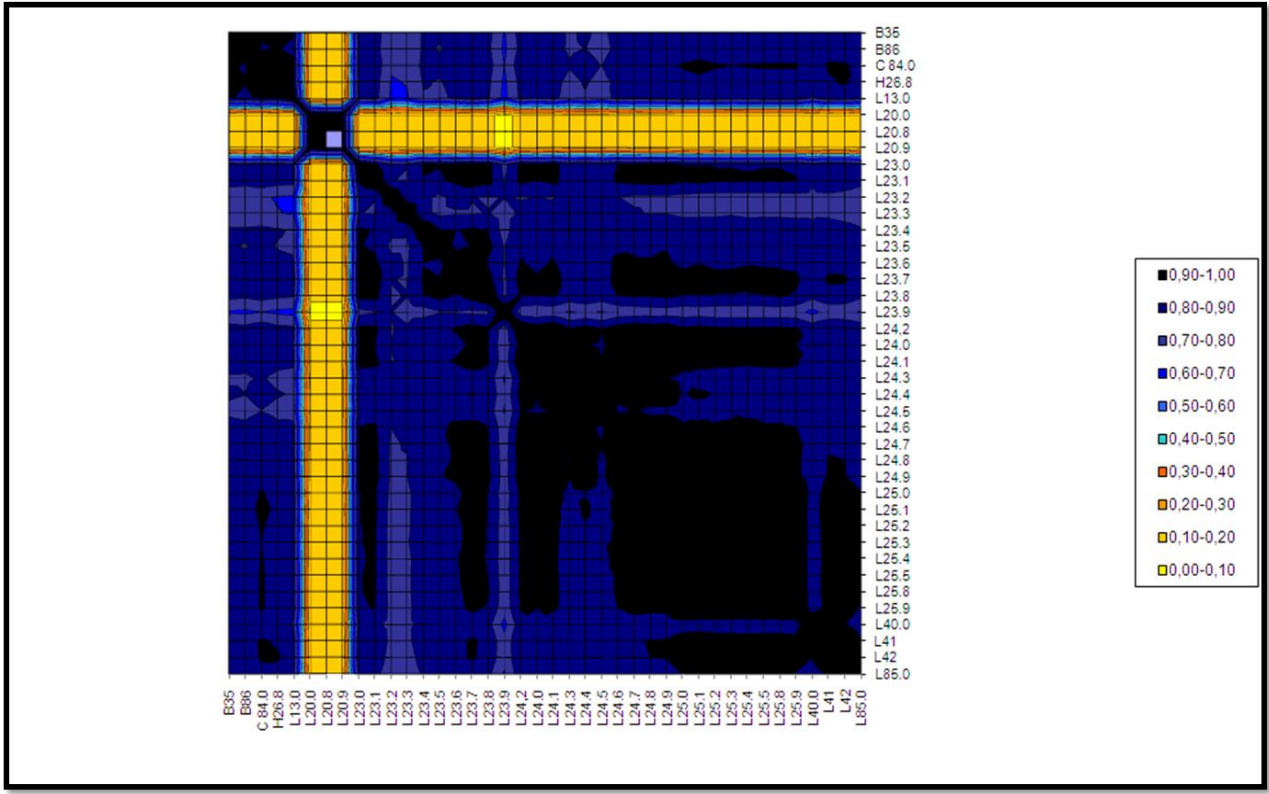


Podobieństwo chorób zapisanych zgodnie z IDC (alergie i choroby skóry)

	B35	B86	C 84.0	H26.8	L13.0	L20.0	L20.8	L20.9	L23.0
B35	1,00	0,96	0,96	0,96	0,96	0,25	0,25	0,25	0,92
B86	0,96	1,00	0,94	0,95	0,94	0,26	0,26	0,26	0,91
C 84.0	0,96	0,94	1,00	0,97	0,96	0,26	0,27	0,27	0,92
H26.8	0,96	0,95	0,97	1,00	0,95	0,25	0,24	0,24	0,91
L13.0	0,96	0,94	0,96	0,95	1,00	0,26	0,26	0,26	0,90
L20.0	0,25	0,26	0,26	0,25	0,26	1,00	0,99	0,99	0,25
L20.8	0,25	0,26	0,27	0,24	0,26	0,99	1,00	1,00	0,26
L20.9	0,25	0,26	0,27	0,24	0,26	0,99	1,00	1,00	0,26
L23.0	0,92	0,91	0,92	0,91	0,90	0,25	0,26	0,26	1,00
L23.1	0,93	0,92	0,93	0,92	0,91	0,24	0,25	0,25	0,96
L23.2	0,83	0,83	0,84	0,83	0,82	0,22	0,22	0,22	0,89
L23.3	0,84	0,84	0,84	0,83	0,82	0,25	0,25	0,25	0,89
L23.4	0,90	0,90	0,91	0,90	0,89	0,25	0,25	0,25	0,95
L23.5	0,90	0,89	0,90	0,89	0,89	0,23	0,24	0,24	0,95
L23.6	0,91	0,91	0,92	0,91	0,90	0,24	0,25	0,25	0,96
L23.7	0,93	0,93	0,94	0,93	0,92	0,25	0,25	0,25	0,96
L23.8	0,91	0,91	0,92	0,91	0,90	0,25	0,25	0,25	0,96

Podobieństwo chorób zapisanych zgodnie z IDC (alergie i choroby skóry)

<http://www.isi.wat.edu.pl>





Przypisanie przypadku diagnozowanego do choroby – klasyfikacja

Załóżmy, że mamy k klas decyzyjnych dec_1, \dots, dec_k . Wtedy funkcją rozróżnialności nazywamy funkcję:

$$F: D \rightarrow R^k$$

która dla danego obiektu przypisuje k wartości — współczynniki wsparcia dla poszczególnych klas. Im większy współczynnik wsparcia tym bardziej prawdopodobne, że obiekt należy do danej klasy.

Lidmila I. Kunchewa, *Combining Pattern Classifiers, Methods and Algorithms*, John Wiley & Sons, Inc., Hoboken, New Jersey 2004



Przypisanie przypadku diagnozowanego do choroby – klasyfikacja

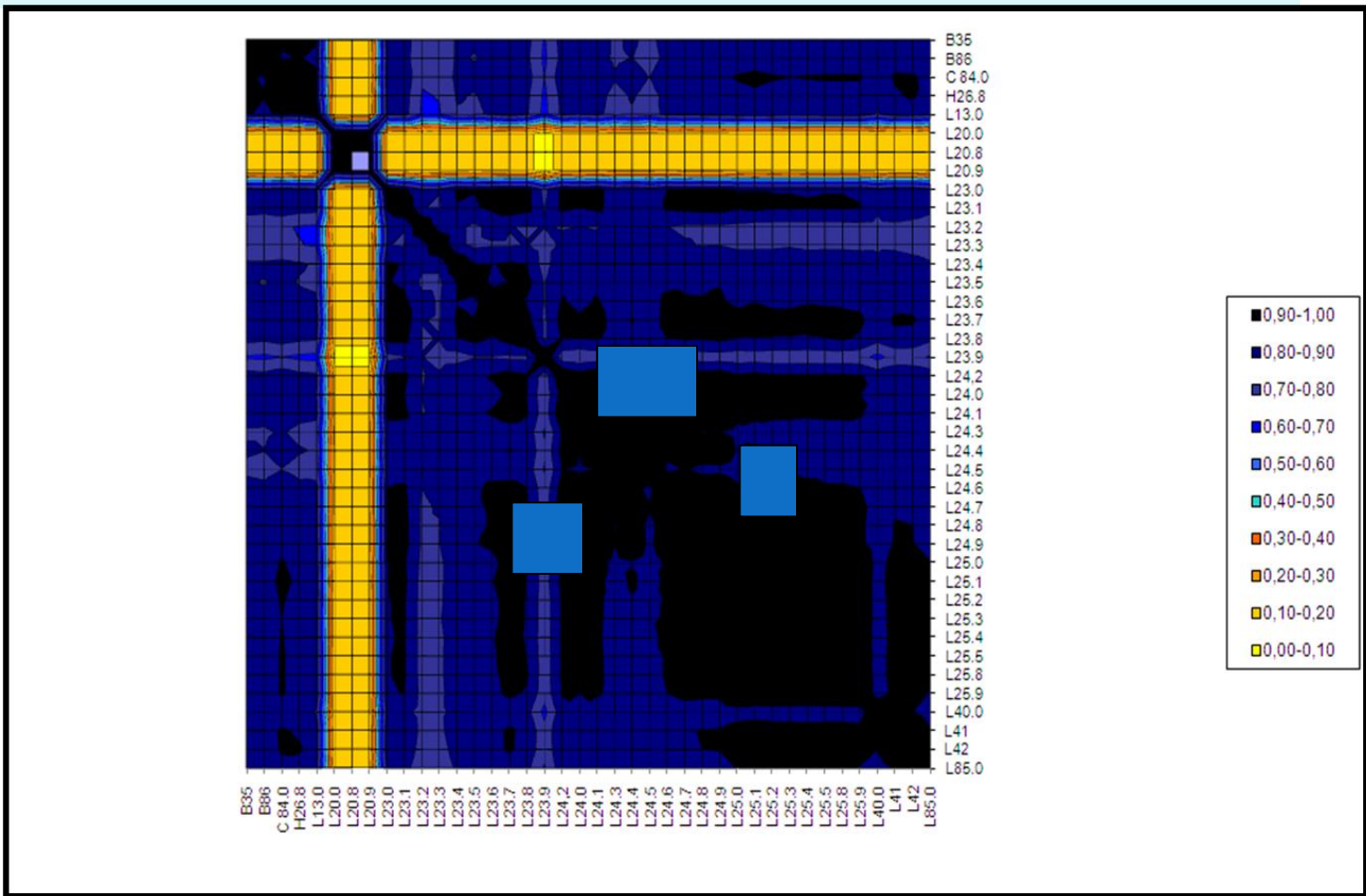
Niech $F = (f_1, \dots, f_k)^T$. Obiekt x jest przypisywany do klasy o najwyższym współczynniku wsparcia, tzn. jeśli:

$$M = \max_i \{f_i(x)\}$$

Lidmila I. Kunchewa, *Combining Pattern Classifiers, Methods and Algorithms*, John Wiley & Sons, Inc., Hoboken, New Jersey 2004

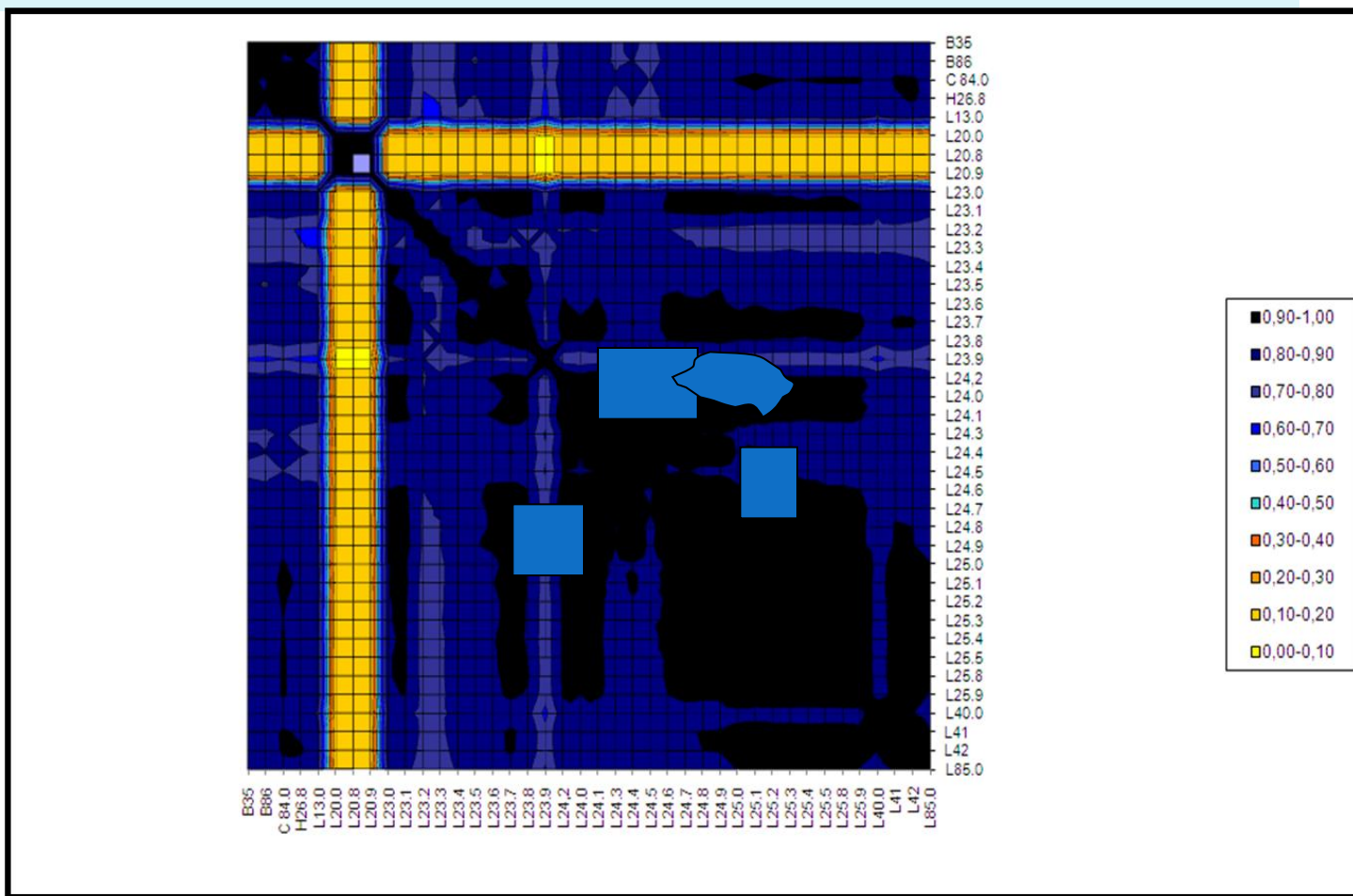
Podobieństwo chorób zapisanych zgodnie z IDC (alergie i choroby skóry)

<http://www.isi.wat.edu.pl>





Podobieństwo chorób zapisanych zgodnie z IDC (alergie i choroby skóry)

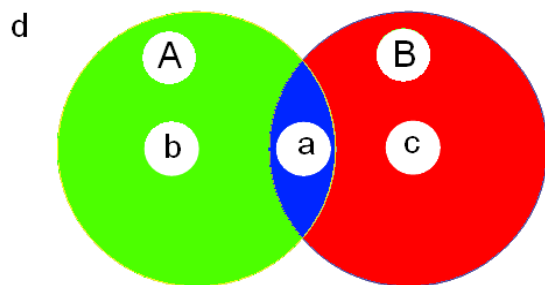


Przykładowe modele klasyfikacji

Metody tworzenia funkcji wsparcia mogą być różne. Najczęściej spotykane w zastosowaniach medycznych to:

- Sieci neuronowe (rozmyte sieci neuronowe)
- Ukryte modele Markowa
- Sieci Bayesa
- Metoda k-NN sąsiadów
- Drzewa decyzyjne
- Klasyfikacje oparte o logikę rozmytą

Przykładowe modele klasyfikacji



A – zbiór cech choroby (z bazy danych medycznych)

B – zbiór symptomów zaobserwowanych u pacjenta



Przykładowe modele klasyfikacji

1. Klasyfikatory oparte na podobieństwie zbioru objawów choroby u pacjenta do zbioru objawów chorób zapisanych w bazie wiedzy medycznej systemu. Współczynniki wsparcia mają w tym przypadku postać stopnia podobieństwa zbiorów wyrażonego liczbą rzeczywistą dodatnią.
2. **Sieci Bayesa.** Współczynnik wsparcia jest w tym przypadku prawdopodobieństwem występowania choroby. .



Cechy rezultatów

1. Przy bazie wiedzy medycznej wielokrotnie zweryfikowanej (niezależnie) przez ekspertów medycznych wszystkie klasyfikatory wskazywały bardzo podobny zbiór chorób przy zróżnicowanych oczywiście współczynnikach wsparcia.
2. Wystarczyły niewielkie ilościowo braki (rzędu 5% objawów) w bazie wiedzy medycznej aby wskazania klasyfikatorów znacznie się różniły.

Cechy rezultatów

lp	Nazwa choroby	Wskaźnik sieci Bayesa	Wskaźnik wsp. Jaccarda	Wskaźnik wsp. przekrycia	Wskaźnik wsp. unormowany	Udział wazony unormowany
1	Łuszczyca	0,49	0,07	0,61	0,216	
2	Grzybica skóry	0,41	0,05	0,42	0,163	
3	Nieokreślone kontaktowe zapalenie skóry wywołane kosmetykami	0,26	0,02	0,21	0,091	
4	Świerzb	0,21	0,03	0,24	0,098	
5	Kontaktowe zapalenie skóry z podrażnienia wywołane przez nieokreślone czynniki	0,19	0,02	0,18	0,072	
6	Nieokreślone kontaktowe zapalenie skóry wywołane przez nieokreślone czynniki	0,19	0,02	0,21	0,078	
7	Nieokreślone kontaktowe zapalenie skóry wywołane barwnikami	0,15	0,02	0,21	0,070	
8	Nieokreślone zapalenie skóry wywołane innymi substancjami chemicznymi	0,14	0,02	0,21	0,068	
9	Łupież różowy Gilberta	0,12	0,04	0,33	0,091	
10	Alergiczne zapalenie skóry wywołane przez przylepiec	0,12	0,02	0,15	0,054	

Weryfikowanie wyników w testach klinicznych

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

TP – true positive

FN- false nagative

Weryfikowanie wyników w testach klinicznych

Diagnozowana JCH	Liczba pacjentów	Odsetek pacjentów w próbie	Czułość sieci Bayesa	Czułość współczynnik a przekrycia
AZS				
Pozycja rankingu 1/2	31/45	69%	0.82	0.85
Pozycja rankingu 1/10	44/45	96%	0.99	0.99



Wnioski

1. Nie należy oczekiwać od klasyfikatora wskazywania jednej choroby BO TO BĘDZIE ZALEŻEĆ OD CECH BAZY WIEDZY MEDYCZNEJ I JEST SPECYFICZNE DLA DANEJ DZIEDZINY MEDYCYNY
2. Nawet relatywnie prosta postać klasyfikatora zapewnia uzyskanie wstępnych rezultatów o czułości najprawdopodobniej akceptowalnej dla środowiska medycznego ale NA PEWNO CHARAKTERYSTYCZNEJ DLA KONKRETNEJ CHOROBY LUB GRUPY CHORÓB