SYSTEM AUTOMATYCZNEJ ANALIZY WZORCÓW ZMIENNOŚCI W RYTMIE SERCA PŁODU

Paweł Łabaj
Instytut Informatyki, Wydział Automatyki, Elektroniki i Informatyki, Politechnika Śląska w Gliwicach

1 WPROWADZENIE

W obecnych czasach komputery mają coraz większy wpływ na nasze życie. Znalazły zastosowanie już w prawie wszystkich dziedzinach, odciążając człowieka w jego pracy. Medycyna nie jest tutaj wyjątkiem. Początkowo komputery służyły głównie do lepszej organizacji pracy czy też usprawnienia przechowywania danych medycznych. Jednak wraz ze wzrostem wydajności komputerów poszerzał się wachlarz możliwych do wykonania przez nie zadań. Obecnie najważniejszym kierunkiem rozwoju jest komputerowo wspomagana diagnostyka medyczna oraz integracja szpitala w jeden system informatyczny.

Już w odległych czasach świadectwem tego, że płód żyje, były jego ruchy odczuwane przez matkę [1]. Od tamtego okresu bardzo wiele się zmieniło, obecne położnictwo ma do swojej dyspozycji dużo lepsze metody oceny stanu płodu. Od kilkudziesięciu lat największe znaczenie ma ocena czynności serca płodu. Prawidłowe działanie układu krążenia jest najprostszym wyznacznikiem prawidłowego rozwoju płodu. Spośród wielu metod uzyskiwania sygnału częstości serca płodu, obecnie najczęściej stosowana jest ultradźwiękowa metoda dopplerowska [6]. Oprócz pomiaru sygnału częstości uderzeń serca dokonuje się również rejestracji aktywności skurczowej macicy oraz ruchów płodu. Proces ten nazywany jest badaniem kardiotokograficznym (KTG) i stanowi podstawową biofizyczną metodę oceny stanu płodu.

Klasyczna analiza zapisu kardiotokograficznego jest dokonywana przez lekarza w sposób wzrokowy i sprowadza się do stwierdzenia, czy zapis jest prawidłowy czy nieprawidłowy. Taki sposób oceny zapisu jest niezwykle trudny ze względu na złożony kształt krzywych. Innym problem związanym z oceną wzrokową jest rozbieżność pomiędzy różnymi ekspertami klinicznymi oceniającymi ten sam zapis. Występują sytuacje [5], w których ten sam lekarz po pewnym czasie odmiennie interpretuje ten sam zapis. Powodem tego jest fakt, że jakość oceny zapisu przez lekarza bardzo zależy nie tylko od jego doświadczenia, ale również od stanu emocjonalnego w momencie dokonywania oceny. Takie rozbieżności w ocenie zapisów KTG są niedopuszczalne, szczególnie w sytuacji, gdy mamy do czynienia z ciążą zagrożoną, gdzie konieczna jest odpowiednio wczesna, ale jednocześnie poprawna i obiektywna diagnoza. Taką obiektywną i powtarzalną diagnozę mogłaby zapewnić kardiotokografia komputerowa, gdzie system komputerowy przeprowadza dokładną analizę ilościową zapisu kardiotokograficznego.

W chwili obecnej kierunkiem rozwoju komputerowych systemów kardiotokograficznych jest wyposażenie ich w diagnostykę stanu płodu na podstawie ilościowych parametrów charakteryzujących zapis kardiotokograficzny. Temat i zakres mojej pracy został tak dobrany, aby wpisywał się w ten trend. Złożone podłoże większości aspektów związanych ze stanem płodu wymogło stosowanie zaawansowanych technik opartych głównie na sieciach neuronowych czy klasyfikatorach statystycznych [8, 12].

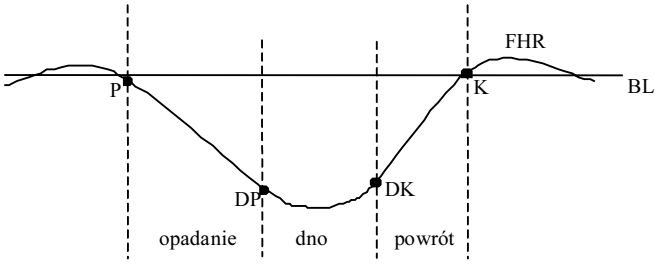
W prezentowanej pracy celem było użycie parametrów ilościowej oceny zapisów KTG do prognozowania wystąpienia konkretnej oznaki nieprawidłowego rozwoju płodu. Projektowany system ma na celu analizę wzorców zmienności

w rytmie pracy serca płodu. Do tych wzorców zaliczane są akceleracje i deceleracje, czyli odpowiednio chwilowe przyspieszenia i spowolnienia częstości uderzeń serca płodu. Jednak tylko deceleracje będą brane pod uwagę w opracowanym systemie. Wynika to z tego, że pojawianie się konkretnych typów deceleracji w zapisie KTG może wskazywać na stan zagrożenia płodu, stąd prawidłowe rozpoznanie typu zjawiska jest niezmiernie istotne. Natomiast akceleracje nie niosą ze sobą negatywnej prognostycznie informacji odnośnie stanu płodu. Klasyfikacja zjawisk deceleracji przeprowadzona będzie zarówno pod względem ich typu, jak i mocy. Parametry ilościowej oceny deceleracji zostały dobrane tak, aby w pełni odpowiadały najnowszym wytycznym odnośnie klasyfikacji deceleracji opracowanym przez *National Insitute of Child Health and Human Development* (NICHD) [7]. Opracowany system do automatycznej analizy wzorców zmienności na pewnym etapie klasyfikacji deceleracji będzie zawierał inteligentny system decyzyjny. Zostanie on zbudowany w oparciu o analizę przydatności następujących typów klasyfikatorów: systemy liniowe oraz sieci neuronowe, zarówno w postaci klasycznej (perceptrony wielowarstwowe oraz sieci o radialnych funkcjach bazowych), jak i w postaci probabilistycznych sieci neuronowych. Zostanie przeprowadzony szereg eksperymentów mających na celu wyłonienie najlepszego klasyfikatora zarówno ze względu na typ, jak i dobór ustawień.

2 DECELERACJE

Od początku badań nad charakterystyką zapisu kardiotokograficznego szczególną uwagę poświęca się deceleracjom, czyli chwilowym obniżeniom krzywej FHR od linii podstawowej. Zgodnie z wytycznymi NICHD [7] deceleracja definiowana jest jako zwolnienie czynności serca płodu o więcej niż 15 bpm i czasie trwania nie krótszym niż 15 s. W ramach deceleracji rozróżnia się trzy fazy: opadanie, dno i powrót (rys. 1).

Określenie granic dla każdej z faz jest bardzo ważnym, a zarazem złożonym zadaniem. W przeciwieństwie do ogólnie przyjętych definicji, w pracy przyjęto, że dno jest szerszym pojęciem niż najniższy punkt deceleracji. Dno reprezentuje grupę próbek FHR położonych w „dolinie" deceleracji. Takie podejście wynika z faktu, że w wielu przypadkach w wyniku zakłóceń próbka znajdująca się najniżej wcale nie musi być dnem. Może to być błędna wartość. Przyjęcie za dno większego fragmentu eliminuje wpływ tego typu zakłóceń na ocenę deceleracji. Dodatkowo, badanie charakteru dna ma wpływ na klasyfikację deceleracji.



Rys. 1. Przykładowa deceleracja z zaznaczonymi fazami.

Jak już wspomniano, w pracy skupiono się na klasyfikacji zjawisk deceleracji. Klasyfikacja będzie się opierać na najnowszych wytycznych NICHD [7]. Jednak jej podstawy są prawie niezmienne od konferencji w Amsterdamie z 1972 r., gdzie zdecydowano się przyjąć podział zaproponowany przez Hona [4]. Zaproponował on podział na deceleracje wczesne, późne i zmienne. Deceleracja wczesna i późna zawdzięczają swoją nazwę położeniu względem skurczu macicy. Natomiast deceleracja zmienna (epizodyczna) może występować w różnych fazach skurczu.

Wytyczne NICHD [7] kładą zwiększony nacisk na to, czy zjawiska deceleracji są powiązane z występującym periodycznie skurczem macicy czy nie. Wynikiem tego jest określenie deceleracji wczesnych i późnych, które wykazują powiązanie ze skurczami macicy, jako periodycznych. Ponadto określenie deceleracji zmiennych jako epizodycznych, gdyż nie wykazują powiązania ze skurczami macicy.

Podział deceleracji tylko względem typu nie jest wystarczający do pełnego określenia stanu płodu. Uważa się, że do kompleksowej oceny predykcyjnego znaczenia deceleracji konieczne jest wprowadzenie gradacji ze względu na moc zjawiska. Najczęściej stosowany jest podział zaproponowany przez Hona [9], który zostanie też zaimplementowany w projektowanym systemie.

Dopiero znając stopień deceleracji ze względu na moc, można w pełni określić znaczenie

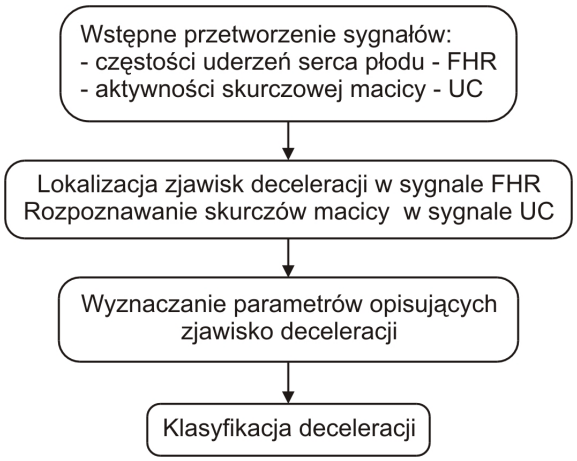
obserwowanej deceleracji w odniesieniu do predykcji stanu płodu.

3 METODYKA

Celem pracy jest stworzenie systemu do automatycznej analizy deceleracji. Jak wspomniano, decelerację są klasyfikowane zarówno pod względem typu, jak i mocy. Podział ze względu na moc jest zadaniem prostym. Podział względem typu jest trudniejszy i wymaga zastosowania specjalnie dobranego klasyfikatora. W pracy chciano sprawdzić przydatność do tego celu kilku typów klasyfikatorów. Są to: klasyfikatory liniowe oraz sieci neuronowe, zarówno w podejściu klasycznym jak i w postaci probabilistycznych sieci neuronowych [2, 3, 10, 11].

Jako materiał badawczy zastosowano zapisy archiwalne systemu MONAKO z trzech ośrodków klinicznych. Materiał badawczy stanowiło 101 zapisów KTG o łącznym czasie trwania 285 godzin. System komputerowego wspomagania monitorowania płodu MONAKO zidentyfikował 2384 zjawisk jako deceleracje. W trakcie klasyfikacji dokonanej przez ekspertów klinicznych do dalszej analizy zostały wybrane jedynie 383 zjawiska.

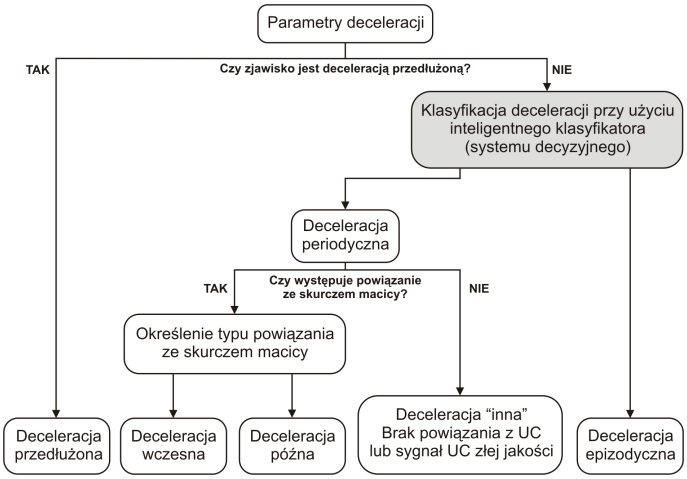
Zapisy archiwalne systemu MONAKO obejmowały pełne spróbkowane sygnały rejestrowane w czasie badania kardiotokograficznego. Do potrzeb pracy posłużyły dwa pliki stowarzyszone z badaniem. W jednym znajdował się sygnał częstości uderzeń serca płodu (FHR), natomiast w drugim pliku znajdował się sygnał aktywności skurczowej macicy (UC).



Rys. 2. Kolejne etapy działania systemu.

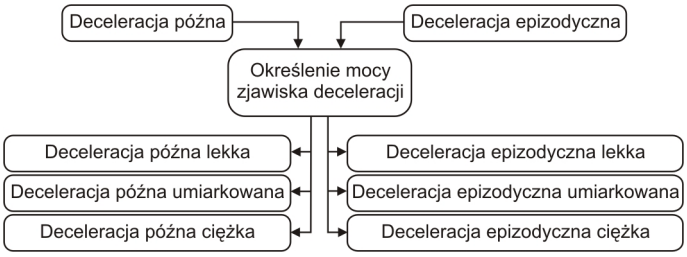
W dotychczas stosowanym rozwiązaniu (system MONAKO) deceleracja była opisywana tylko za pomocą: maksymalnej amplitudy, czasu trwania i pola zjawiska. Jednak jest to zestaw parametrów niewystarczający dla klasyfikacji deceleracji. W prezentowanym podejściu tych parametrów będzie znacznie więcej. Będą one starały się opisać charakter poszczególnych typów deceleracji.

Wybór najlepszego typu klasyfikatora zostanie dokonany z użyciem pakietu *STATISTICA Neural Networks*. Poszczególne moduły systemu zostały zrealizowane w środowisku LabVIEW.



Rys. 3. Schemat klasyfikacji deceleracji pod względem typu.

Głównym kryterium zastosowania tych programów była ich dostępność w Instytucie Techniki i Aparatury Medycznej (ITAM) w Zabrzu. Nie umniejsza to faktu, że jest to oprogramowanie najwyższej klasy, mające szereg zalet, które w wielu testach porównawczych było uznawane za najlepsze w swojej klasie[10].



Rys. 4. Schemat klasyfikacji deceleracji pod względem mocy.

Jednym z ważniejszym aspektów pracy jest dobranie odpowiednich parametrów, które będą opisywać charakter deceleracji. W oparciu o wcześniej dokonywane próby [12], jednocześnie mając na uwadze zalecenia najnowszych

wytycznych [7] oraz zalecenia uznanych autorytetów z dziedziny położnictwa, zdecydowano się na zdefiniowanie własnego zbioru parametrów opisujących charakter deceleracji i to zarówno zjawiska jako całości, jak i poszczególnych jego faz.

Ze względu na to, że jakość oceny eksperta zależy od wielu czynników, takich jak: doświadczenie, zmęczenie, nastrój i inne, zdecydowano się skorzystać z kilku ekspertów. Przed dokonaniem klasyfikacji przez ekspertów poproszono ich o dokładne zapoznanie się z wytycznymi NICHD [7]. Do analizy wybrano te deceleracje, które zostały zgodnie przez wszystkich ekspertów zaklasyfikowane jako periodyczne bądź epizodyczne. Zastosowanie części wspólnej z decyzji kilku ekspertów jest najbezpieczniejszym podejściem, jednak powoduje znaczne ograniczenie ilości materiału badawczego. Spośród wybranych 383 zjawisk 55 to deceleracje periodyczne, a 328 to deceleracje epizodyczne. Po sklasyfikowaniu deceleracji przez ekspertów klinicznych oraz mając wyliczone parametry opisujące te deceleracje, można było przystąpić do konstruowania systemu decyzyjnego.

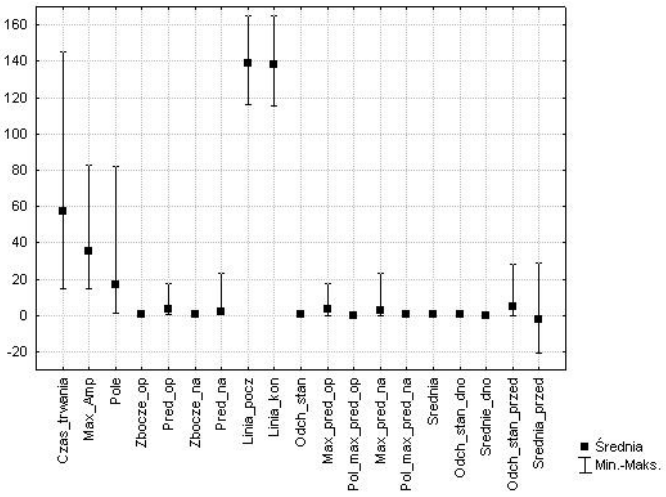
Wynikiem działania systemu decyzyjnego jest podział deceleracji na periodyczne i epizodyczne. Do pełnej klasyfikacji pod względem typu w wypadku deceleracji periodycznej konieczny jest jeszcze podział na wczesne i późne. Z wytycznych NICHD [7] wynika, że takiego podziału można dokonać jedynie poprzez określenie położenia deceleracji względem skurczu macicy. Z tego względu zdecydowano się oprzeć na zależności pomiędzy dnem deceleracji i szczytem skurczu macicy przy określaniu, czy daną decelerację należy zaklasyfikować jako wczesną czy późną. Dotychczas opisane etapy pozwalają na obliczenie podstawowych i szczegółowych parametrów deceleracji oraz dokonanie podziału względem typu na deceleracje (rys. 3):

1 epizodyczne
2 periodyczne:
– wczesne
– późne
3 przedłużone

Jednak jak wspomniano, do pełnej oceny istotności deceleracji konieczne jest jeszcze określenie mocy deceleracji epizodycznych i późnych (rys. 4).

4 KLASYFIKATOR NEURONOWY

System decyzyjny do klasyfikacji deceleracji na periodyczne i epizodyczne został zrealizowany przy wykorzystaniu programu *STATISTICA*, a w szczególności pakietu *Neural Networks* w oparciu o parametry opisu deceleracji oraz decyzję ekspertów klinicznych odnośnie klasyfikacji tych deceleracji. Otrzymano zestaw 19 parametrów ciągłych opisujących zjawisko deceleracji (parametry wejściowe) oraz jeden parametr określający klasę deceleracji (parametr wyjściowy). Ze względu na to, że wartości parametrów wejściowych reprezentowały różne zakresy (rys. 5), zdecydowano się dokonać ich normalizacji.



Rys. 5. Rozrzut wartości parametrów wejściowych.

Po normalizacji wartości parametrów wejściowych, drugim problemem dotyczącym przygotowania danych jest podział na zbiory: uczący, walidacyjny i testowy. Oczywiście można wykorzystać w tym celu wbudowane w program *STATISTICA Neural Networks* mechanizmy. Jednak należy zauważyć, że przy każdym eksperymencie dane te będą losowane do zbiorów inaczej. Przy bardzo dużej liczbie przypadków nie ma to znaczenia, jednak w rozpatrywanej sytuacji może mieć to znaczenie. Z tego względu, w celu zapewnienia możliwości porównywania wyników z różnych eksperymentów, dokonałem jednokrotnego losowego przydziału przypadków do trzech ww. zbiorów.

Jeśli chodzi o liczebność poszczególnych zbiorów, to najczęściej stosowany jest podział: 2-1-1; uczący (U) – walidacyjny (W) – testowy (T). Jednak [2, 3] zaleca się zwiększenie

liczebności zbioru uczącego kosztem zbioru testowego. Uwzględniając powyższe, ustalono następujące liczebności zbiorów: 205-107-71, co w przybliżeniu wyraża stosunek 6-3-2.

Zdecydowano się sprawdzić trzy typy klasyfikatorów: klasyfikatory liniowe, sieci neuronowe (MLP i RBF) oraz probabilistyczne sieci neuronowe. Celem przeprowadzonych doświadczeń było wybranie optymalnego typu klasyfikatora jako systemu decyzyjnego oraz dobranie optymalnych parametrów ustawień dla danego systemu decyzyjnego. W eksperymentach posługiwano się *Asystentem Projektanta Sieci*, który wchodzi w skład *STATISTICA Neural Networks*.

4.1 *Klasyfikatory liniowe*

Problemem przy zadaniach klasyfikacyjnych jest określenie *a priori* złożoności badanego zjawiska. Z tego względu dobrym rozwiązaniem jest sprawdzenie, jak najprostszy klasyfikator liniowy radzi sobie z danym zagadnieniem. Takie same przesłanki skłoniły mnie do przeprowadzenia eksperymentu, w którym jako system decyzyjny zostanie przyjęty klasyfikator liniowy. Wyniki będą swego rodzaju punktem odniesienia dla wyników uzyskiwanych przez inne klasyfikatory. Być może okaże się, że problem podziału deceleracji na epizodyczne i periodyczne jest liniowo separowalny.

4.2 *Sztuczne sieci neuronowe – MLP*

W drugim eksperymencie zostanie przetestowana przydatność typowych sieci neuronowych. Wielowarstwowe sieci perceptronowe są najbardziej charakterystycznym przedstawicielem sieci neuronowych. Najprawdopodobniej rozpatrywany problem nie jest liniowo separowalny, więc wyniki uzyskane przez sieci MLP powinny być znacznie lepsze niż wyniki dla klasyfikatora liniowego.

4.3 *Sztuczne sieci neuronowe – RBF*

Sieci o radialnych funkcjach bazowych prezentują odmienne podejście do podziału przestrzeni wejściowej. O ile dla sieci MLP był to podział z użyciem hiperpłaszczyzn, o tyle dla sieci RBF jest to interpretacja skupieniowa. Być może omawiany problem klasyfikacji deceleracji na periodyczne i epizodyczne jest problemem skupieniowym. Eksperyment numer 3 powinien rozwiać te wątpliwości.

4.4 *Probabilistyczne sieci neuronowe (PNN)*

W eksperymencie tym zostaną przetestowane, pod kątem przydatności do omawianej klasyfikacji, probabilistyczne sieci neuronowe. Sieci te stosują zupełnie odmienne podejście do klasyfikacji.

4.5 *Sieci PNN – wpływ współczynnika wygładzania*

W eksperymencie tym zostanie sprawdzony wpływ współczynnika wygładzania na jakość sieci PNN. W eksperymencie zostanie użyta najlepsza sieć z poprzedniego eksperymentu. Współczynnik wygładzania określa zdolność generalizacji sieci PNN.

4.6 *Sieci PNN – wpływ prawdopodobieństw a-piori*

Istnieje możliwość ustalenia prawdopodobieństwa *a priori* dla sytuacji, gdy któraś z klas jest liczniej reprezentowana od innej. W rozpatrywanym przypadku ma miejsce taka właśnie sytuacja, ponieważ deceleracje epizodyczne są znacznie liczniejsze od deceleracji periodycznych. W eksperymencie tym zostanie zbadane, jak zdefiniowanie rzeczywistych prawdopodobieństw klas wpłynie na jakość sieci PNN. Do badań w tym eksperymencie została wybrana sieć eksperymentu czwartego.

4.7 *Sieci PNN – wpływ macierzy kosztów*

W ostatnim eksperymencie zostanie sprawdzona druga z możliwych modyfikacji sieci PNN, a mianowicie macierz kosztów. Realizuje ona ideę, że wszystkie błędne klasyfikacje są złe, ale niektóre z nich są ekstremalnie złe. W naszym wypadku ze względu na to, że deceleracji periodycznych jest mniej, bardziej szkodliwe jest klasyfikowanie deceleracji periodycznej jako epizodyczną. Doświadczenie to zostanie przeprowadzone przy użyciu sieci z eksperymentu czwartego.

5 WYNIKI

W pierwszym eksperymencie sprawdzono przydatność klasyfikatorów liniowych jako systemu decyzyjnego do klasyfikacji deceleracji. W pierwszym kroku eksperymentu został znaleziony najlepszy klasyfikator analizujący wszystkie 19 parametrów wejściowych. W tabeli 1 zaprezentowano wyniki osiągnięte przez znaleziony najlepszy klasyfikator. Klasyfikator prawidłowo rozpoznał wszystkie (55) deceleracje periodyczne oraz tylko 243 z 328 deceleracji epizodycznych. Wyniki te pokazują podstawową wadę klasyfikatorów liniowych. Znalezione rozwiązanie idealnie klasyfikuje deceleracje periodyczne, niestety jednocześnie błędnie klasyfikując ponad 1/4 deceleracji epizodycznych.

Tabela 1. Jakości i błędy klasyfikatora liniowego analizującego wszystkie 19 parametrów.

Jakość klasyfikatora dla zbioru			Błąd klasyfikatora dla zbioru		
U	W	T	U	W	T
75,12%	82,24%	78,87%	0,244	0,242	0,279

W drugim kroku pozwolono, aby *Asystent Projektanta Sieci* programu *STATISTICA Neural Networks* z wykorzystaniem dostępnych algorytmów wybrał tylko optymalne parametry wejściowe. W ten sposób zostało wybranych 15 parametrów wejściowych spośród 19, a osiągnięte wyniki prezentuje tabela 2.

Tabela 2. Jakości i błędy dla klasyfikatora liniowego wykorzystującego ograniczoną liczbę parametrów

Jakość klasyfikatora dla zbioru			Błąd klasyfikatora dla zbioru		
U	W	T	U	W	T
80,49%	85,98%	84,51%	0,244	0,243	0,281

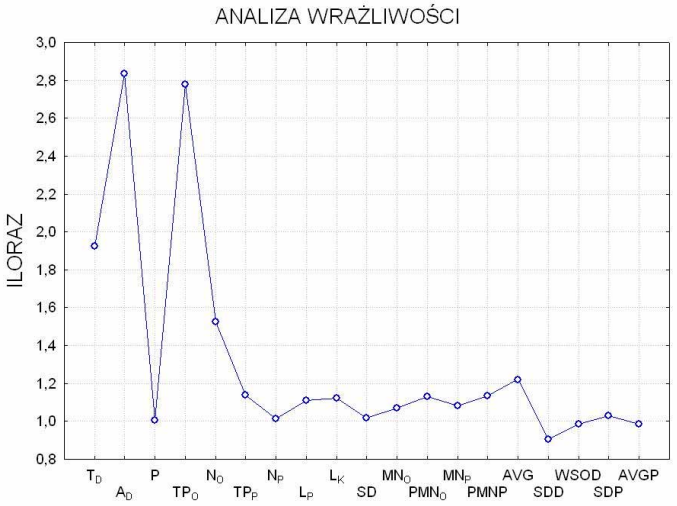
Uzyskane wskaźniki jakości są lepsze niż uzyskane dla klasyfikatora z 19 parametrami wejściowymi. Wpłynęła na to lepsza identyfikacja deceleracji epizodycznych (262 z 328) bez pogarszania wykrywalności deceleracji periodycznych (55). Usunięcie 4 parametrów wejściowych (PMN_O, MN_P, SDD i SDP) umożliwiło znalezienie lepszej hiperpłaszczyzny. Uzyskane klasyfikatory liniowe bardzo dobrze radzą sobie z identyfikacją deceleracji periodycznych, niestety odbywa się to kosztem deceleracji epizodycznych. Najwyraźniej zagadnienie klasyfikacji deceleracji jest zbyt złożone dla klasyfikatorów liniowych.

W drugim eksperymencie sprawdzono przydatność sieci neuronowych o architekturze MLP. W pierwszym kroku eksperymentu została znaleziona najlepsza sieć typu MLP wykorzystująca wszystkie 19 parametrów wejściowych. W tabeli 3 prezentowane są wyniki osiągnięte przez znalezioną najlepszą sieć o strukturze MLP 19:19-30-1:1. Uzyskane wskaźniki jakości sieci są zdecydowanie lepsze od wskaźników jakości klasyfikatora liniowego. Poprawnie rozpoznano 53 z 55 deceleracji periodycznych i 322 z 328 deceleracji epizodycznych.

Tabela 3. Jakości i błędy dla sieci o strukturze MLP 19:19-30-1:1.

Jakość sieci dla zbioru			Błąd sieci dla zbioru		
U	W	T	U	W	T
98,05%	99,07%	95,77%	0,204	0,132	0,309

Na rys. 6 przedstawiono wyniki analizy wrażliwości jakości sieci na usunięcie poszczególnych parametrów wejściowych. Jedne z najważniejszych parametrów to TP_O i N_O, które w bezpośredni sposób opisują fazę opadania deceleracji. Można powiedzieć, że najprawdopodobniej wynik sieci można poprawić, usuwając niektóre parametry (np. SDD).



Rys. 6. Wartości ilorazów parametrów wejściowych dla analizy wrażliwości.

Przypuszczenie to zostało sprawdzone w drugim kroku eksperymentu, gdzie *Asystent Projektanta Sieci* dobrał optymalny zestaw parametrów wejściowych. W tabeli 4 zostały zap-

rezentowane wyniki najlepszej sieci o strukturze MLP 10:10-14-1:1.

Tabela 4. Jakości i błędy sieci o strukturze MLP 10:10-14-1:1.

Jakość sieci dla zbioru			Błąd sieci dla zbioru		
U	W	T	U	W	T
99,02%	98,13%	94,37%	0,134	0,204	0,311

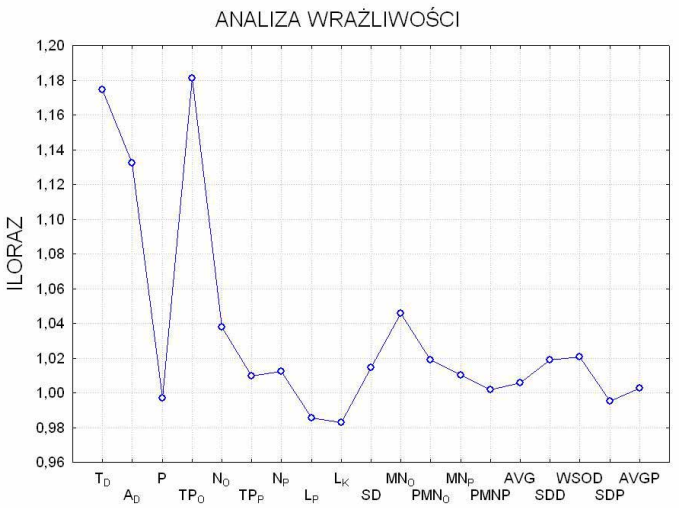
Sieć ta uzyskała porównywalne rezultaty w stosunku do sieci wykorzystującej 19 parametrów wejściowych. Poprawnie sklasyfikowała 53 z 55 deceleracji periodycznych i 322 z 328 deceleracji epizodycznych. Zmianie uległo jedynie rozmieszczenie poprawnie sklasyfikowanych przypadków w zbiorach (uczącym, walidacyjnym i testowym). Można powiedzieć jednak, że sieć ta jest lepsza, gdyż przy takiej samej jakości ma znacznie prostszą strukturę. W sieci tej w procesie klasyfikacji biorą udział jedynie następujące parametry wejściowe: T_D, A_D, TP_O, TP_P, L_K, MN_O, PMN_O, MN_P, SDP, AVGP.

W trzecim eksperymencie sprawdzono przydatność sieci neuronowych o architekturze RBF jako systemu decyzyjnego przeznaczonego do klasyfikacji deceleracji. W pierwszym kroku eksperymentu została znaleziona najlepsza sieć typu RBF wykorzystująca wszystkie 19 parametrów wejściowych. W tabeli 5 zaprezentowano wyniki osiągnięte przez znalezioną najlepszą sieć o strukturze RBF 19:19-29-1:1. Uzyskane przez sieć o architekturze RBF jakości są, podobnie jak w wypadku sieci MLP, bardzo dobre. Sieć poprawnie sklasyfikowała 51 z 55 deceleracji periodycznych i 312 z 328 deceleracji epizodycznych.

Tabela 5. Jakości i błędy dla sieci o strukturze RBF 19:19-29-1:1

Jakość sieci dla zbioru			Błąd sieci dla zbioru		
U	W	T	U	W	T
93,17%	97,20%	95,77%	0,197	0,203	0,247

Na rys. 7 przedstawiono wykres analizy wrażliwości jakości sieci na usunięcie poszczególnych parametrów wejściowych, dla sieci typu RBF zakres wartości ilorazów jest mniejszy. Najwyższe wartości ilorazu osiągnęły parametry T_D, A_D, TP_O, N_O i MN_O. Jest to prawie ten sam zestaw, co dla sieci MLP. Oznacza to, że zestaw ten w największym stopniu wpływa na klasyfikację, niezależnie od typu sieci neuronowej. Ponadto jakość obu sieci świadczy o tym, że parametry te bardzo dobrze oddają charakter obu typów deceleracji, dzięki czemu jakości klasyfikacji są tak dobre. Godny zauważenia jest również fakt, że dla uzyskanej sieci parametry L_K i L_P pogarszają wynik, przeciwnie niż dla klasyfikatorów liniowych. Jakość uzyskana przez sieć RBF potwierdza, że klasyfikacja dokonywana przez klasyfikator liniowy opiera się najprawdopodobniej na niewłaściwych parametrach.



Rys. 7. Wartości ilorazów parametrów wejściowych dla analizy wrażliwości.

W drugim kroku *Asystent Projektanta Sieci* stworzył sieć o strukturze RBF 5:5-98-1:1. Sieć ta ma optymalny zestaw parametrów wejściowych. Jej wyniki zostały przedstawione w tabeli 6. Zaprojektowana sieć uzyskała lepsze rezultaty niż wszystkie wcześniejsze klasyfikatory, jednocześnie wykorzystuje jedynie 5 z 19 parametrów wejściowych. Optymalny zestaw parametrów wejściowych tworzą: T_D, A_D, TP_O, N_O, MN_O i jest to zestaw, o którym była mowa przy okazji analizy wrażliwości sieci RBF z wariantu pierwszego. Sieć poprawnie sklasyfikowała 53 z 55 deceleracji periodycznych i 323 z 328 deceleracji epizodycznych. Nastąpiła poprawa w stosunku do sieci z eksperymentu 2 jedynie o jedną deceleracje, jednak w klasyfikacji bierze udział jedynie pięć parametrów wejściowych.

Tabela 6. Jakości i błędy dla sieci o strukturze RBF 5:5-98-1:1.

Jakość sieci dla zbioru			Błąd sieci dla zbioru		
U	W	T	U	W	T
99,02%	97,20%	97,18%	0,112	0,158	0,179

W czwartym eksperymencie sprawdzono przydatność ostatniego typu klasyfikatora, a mianowicie probabilistycznych sieci neuronowych (PNN) jako systemu decyzyjnego przeznaczonego do klasyfikacji deceleracji. W pierwszym kroku eksperymentu została znaleziona najlepsza sieć typu PNN wykorzystująca wszystkie 19 parametrów wejściowych. W tabeli 7 zaprezentowano wyniki osiągnięte przez znalezioną najlepszą sieć o strukturze PNN 19:19-205-2-2:1. Wyniki uzyskane przez sieć typu PNN wydają się być całkiem dobre. Niestety stosunkowo wysoka jakość wynika z faktu, że sieć nauczyła się rozpoznawać tylko decelerację epizodyczne, które stanowią ponad 85% wszystkich deceleracji. Sieć poprawnie sklasyfikowała wszystkie decelerację epizodyczne i żadnej deceleracji periodycznej. Odnośnie analizy wrażliwości, wartości ilorazów zawierają się w przedziale od 0,9988 do 1,0000.

Wyniki uzyskane w drugim kroku (tabela 8), gdzie *Asystent Projektanta Sieci* dobrał optymalny zestaw parametrów, są lepsze jedynie w zbiorze uczącym. Wybrana sieć ma strukturę PNN 14:14-205-2-2:1. Z danych wynika, że sieć typu PNN w swojej podstawowej postaci nie nadaje się do klasyfikacji deceleracji, gdyż prawdopodobieństwo przynależności do deceleracji epizodycznej jest znacznie większe niż prawdopodobieństwo przynależności do deceleracji periodycznej, niezależnie od wartości parametrów wejściowych. Z tego względu w kolejnych eksperymentach sprawdzony zostanie wpływ modyfikacji na przydatność sieci typu PNN do klasyfikacji deceleracji. Jako sieć bazowa zostanie użyta sieć PNN 14:14-205-2-2:1 otrzymana w drugim wariancie eksperymentu z podrozdziału 4.4.

Tabela 7. Jakości i błędy dla sieci o strukturze PNN 19:19-205-2-2:1.

Jakość sieci dla zbioru			Błąd sieci dla zbioru		
U	W	T	U	W	T
85,37%	85,98%	85,92%	0,791	0,795	0,791

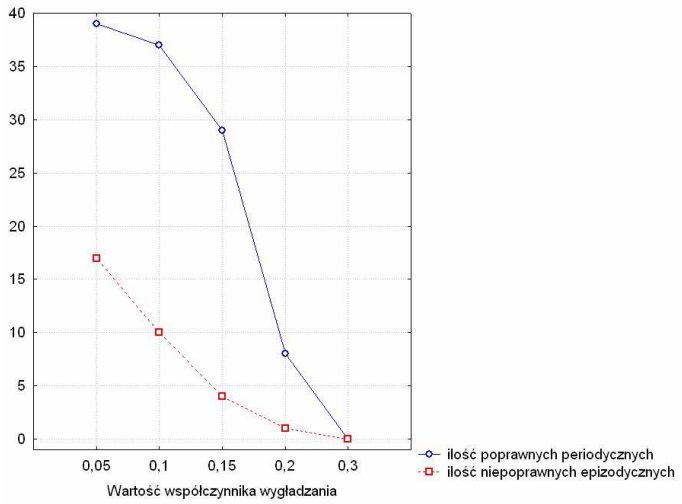
Tabela 8. Jakości i błędy dla sieci o strukturze PNN 14:14-205-2-2:1.

Jakość sieci dla zbioru			Błąd sieci dla zbioru		
U	W	T	U	W	T
89,27%	85,98%	84,51%	0,842	0,803	0,791

W piątym eksperymencie sprawdzono wpływ współczynnika wygładzania na jakość wyników sieci PNN 14:14-205-2-2:1. Współczynnik wygładzania decyduje o zasięgu „strefy wpływów" danego neuronu w przestrzeni parametrów wejściowych. W programie *STATISTICA Neural Networks* domyślnie jest on ustawiony na wartość 0,2. Eksperyment polegał na modyfikacji współczynnika wygładzania dla istniejącej sieci PNN 14:14-205-2-2:1 i ponownym jej uczeniu. Dla sieci typu PNN ponowne uczenie zawsze jest równoznaczne z uczeniem od nowa. W tabeli 9 zaprezentowano wyniki uzyskane dla różnych ustawień współczynnika. Analizując tabelę 9, moglibyśmy odnieść wrażenie, że dla wartości współczynnika wygładzania równej 0,3 sieć jest najlepsza, ponieważ jakości w różnych zbiorach są zbliżone. Niestety, jeśli przyjrzeć się dokładniej liczbie poprawnie sklasyfikowanych deceleracji każdego typu, wniosek będzie zupełnie inny.

Tabela 9. Jakości i błędy dla sieci o strukturze PNN 14:14-205-2-2:1 przy różnych wartościach współczynnika wygładzania.

Wartość współczynnika wygładzania	Jakość sieci dla zbioru			Błąd sieci dla zbioru		
	U	W	T	U	W	T
0,3	85,37%	85,98%	85,92%	0,789	0,780	0,769
0,2	89,27%	85,98%	84,51%	0,842	0,803	0,791
0,15	99,02%	85,98%	81,69%	0,912	0,828	0,816
0,1	100,00%	85,05%	83,10%	0,986	0,861	0,855
0,05	100,00%	82,24%	80,28%	0,999	0,893	0,890



Rys. 8. Zależność liczby wykrytych deceleracji periodycznych i liczby niewykrytych deceleracji epizodycznych w zależności od wartości współczynnika wygładzania.

Na wykresie z rys. 8 przedstawiono zależność liczby wykrytych deceleracji periodycznych i liczby niewykrytych deceleracji epizodycznych w zależności od wartości współczynnika wygładzania. Z wykresu wynika, że najlepsza sytuacja, pod kątem wykrytych deceleracji periodycznych, jest dla wartości współczynnika wygładzania równej 0,05.

Im mniejsza wartość współczynnika wygładzania, tym lepsze dopasowanie do zbioru uczącego. Ponieważ 30 z 55 deceleracji periodycznych znajduje się w zbiorze uczącym, to im mniejsza wartość współczynnika wygładzania, tym więcej poprawnie wykrytych deceleracji tego typu. Sieć jest dopasowana do zbioru uczącego, jednak nie ma właściwości generalizującej. Dla współczynnika o wartości 0,3 sieć ma właściwość generalizacji. Jednak ponieważ 85% próbek stanowią deceleracje epizodyczne, nawet deceleracje periodyczne są traktowane jako epizodyczne. Podsumowując, można powiedzieć, że zmiana współczynnika wygładzania nie poprawia zdolności poprawnej klasyfikacji bez utraty przez sieć zdolności do generalizowania na nowe nieznane przypadki.

Dotychczasowym problemem dla sieci PNN w analizowanym zagadnieniu był fakt znacznej przewagi liczebnej jednej klasy w stosunku do drugiej. W takiej sytuacji pomocna może się okazać modyfikacja, której celem jest dodanie do sieci informacji o prawdopodobieństwie występowania przypadków obu klas. W rozpatrywanej sytuacji prawdopodobieństwo występowania deceleracji periodycznych wynosi w przybliżeniu 15%, a epizodycznych 85%. W eksperymencie szóstym, bazując na modelu z eksperymentu 4, dokonujemy ponownego uczenia sieci PNN 14:14-205-2-2:1 z ustawionymi powyższymi prawdopodobieństwami. Uzyskane wyniki prezentuje tabela 10.

Tabela 10. Jakości i błędy dla sieci o strukturze PNN 14:14-205-2-2:1 z modyfikacją prawdopodobieństwa *a priori* i bez modyfikacji

Typ sieci	Jakość sieci dla zbioru			Błąd sieci dla zbioru		
	U	W	T	U	W	T
Bez modyfikacji	89,27%	85,98%	84,51%	0,842	0,803	0,790
Z modyfikacją	89,27%	85,98%	84,51%	0,839	0,800	0,788

Jak widać, w analizowanym problemie modyfikacja nie dała żadnych pozytywnych rezultatów. Być może ustawienie innych prawdopodobieństw poprawiłoby wynik. Jednak inne prawdopodobieństwa byłyby niezgodne z wartościami rzeczywistymi otrzymanymi z materiału badawczego.

Podstawy teoretyczne sieci PNN dopuszczają jeszcze jeden typ modyfikacji struktury sieci. Jest nim wprowadzenie macierzy kosztów. W macierzy tej zawarte są informacje, jak należy traktować poszczególne decyzje sieci, które błędy powinny być bardziej kosztowne, a które mniej. W wypadku rozpatrywanego problemu logiczne wydaje się, że bardziej kosztowne powinno być błędne zaklasyfikowanie deceleracji periodycznej jako epizodycznej. Z tego powodu w programie *STATISTICA Neural Networks* koszt błędnej klasyfikacji deceleracji epizodycznej jako periodycznej nie był zmieniany, zwiększaniu natomiast podlegał koszt drugiego typu błędu.

Jak widać w tabelach 11 i 12, macierz kosztów ma wpływ na jakość sieci PNN i jej przydatność do klasyfikacji deceleracji. W danym zestawieniu należało by się zastanowić, co jest ważniejsze: czy ogólna jakość sieci czy jakość klasyfikacji deceleracji periodycznych (występujących znacznie rzadziej). Wybierając wersję sieci PNN z wagą błędnej klasyfikacji deceleracji periodycznej jako epizodycznej równą trzy, można powiedzieć, że jest ona nieco lepsza od klasyfikatora liniowego, jednak jednocześnie dużo gorsza od obu klasycznych typów sieci neuronowych. Być może umiejętna kombinacja wszystkich trzech typów modyfikacji sieci PNN dałaby lepsze rezultaty.

Otrzymane rezultaty dowodzą, że spośród sprawdzonych klasyfikatorów najlepsze okazały się sieci neuronowe o architekturach MLP i RBF.

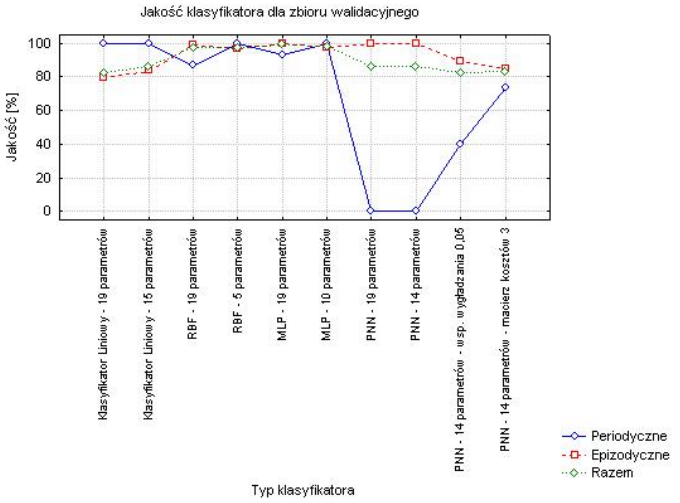
Tabela 11. Jakości i błędy dla sieci o strukturze PNN 14:14-205-2-2:1 dla różnych konfiguracji macierzy kosztów.

Waga błędnej klasyfikacji deceleracji periodycznej jako epizodycznej	Jakość sieci dla zbioru			Błąd sieci dla zbioru		
	U	W	T	U	W	T
1 (domyślnie)	89,27%	85,98%	84,51%	0,841	0,802	0,790
2	96,59%	88,79%	87,32%	0,796	0,733	0,720
3	89,27%	83,18%	74,65%	0,765	0,687	0,673

Tabela 12. Liczba poprawnie sklasyfikowanych deceleracji w sieci PNN 14:14-205-2-2:1 dla różnych konfiguracji macierzy kosztów.

Waga błędnej klasyfikacji deceleracji periodycznej jako epizodycznej	Liczba poprawnie sklasyfikowanych deceleracji periodycznych	Liczba poprawnie sklasyfikowanych deceleracji epizodycznych
1 (domyślnie)	8 (14,5%)	327 (99,7%)
2	38 (69,1%)	317 (96,6%)
3	48 (87,3%)	277 (84,5%)

Często podkreśla się, aby przy projektowaniu i uczeniu klasyfikatorów największą uwagę zwracać na jakość dla zbioru walidacyjnego. To ona jest uznawana za najważniejszy wyznacznik jakości klasyfikatora. Na rys. 9 przedstawiono zestawienie jakości różnych klasyfikatorów dla zbioru walidacyjnego. Jak można zauważyć, najwyższą jakość, całościową jak i dla poszczególnych typów deceleracji, w zbiorze walidacyjnym osiągnęła sieć neuronowa typu RBF wykorzystująca pięć parametrów wejściowych. Równie dobrze prezentuje się sieć typu MLP. Podsumowując, można powiedzieć, że jako system decyzyjny do klasyfikacji deceleracji najlepiej nadają się sieci neuronowe RBF i MLP z ograniczoną liczbą parametrów. Niestety zarówno klasyfikatory liniowe, jak i sieci typu PNN nie radzą sobie z zadaniem klasyfikacji deceleracji w zbadanych konfiguracjach. Być może odpowiednie bardzo specyficzne ustawienie parametrów sieci typu PNN umożliwi uzyskanie wyników zbliżonych do sieci typu RBF i MLP.



Rys. 9. Zestawienie jakości różnych klasyfikatorów dla zbioru walidacyjnego dla deceleracji periodycznych, epizodycznych i łącznie.

Godne uwagi jest to, że dla obu typów najlepszych sieci optymalne zestawy parametrów wejściowych pokrywają się. Zestaw optymalny dla sieci RBF (T_D, A_D, TP_O, N_O i MN_O) oddaje podstawowe informacje odnośnie kształtu, na które został położony nacisk w wytycznych NICHD. Oznacza to, że przyjęte definicje tych parametrów są poprawne oraz przeprowadzony podział deceleracji na trzy fazy jest poprawny.

6 PODSUMOWANIE

W ramach pracy nie powstał wprawdzie gotowy do użycia system do automatycznej analizy wszystkich wzorców zmienności częstości serca płodu, jednak wykonana praca dostarczyła wielu cennych informacji, które zostaną wykorzystane przy stworzeniu takiego systemu. Ze względu na złożoność problemu potrzebne są dalsze prace w tym kierunku. Po analizie wyników przeprowadzonych eksperymentów konieczna będzie ostateczna decyzja odnośnie struktury systemu decyzyjnego do klasyfikacji deceleracji na periodyczne i epizodyczne. Porównywalne bardzo dobre wyniki sieci typu MLP i RBF powodują, że należałoby się zastanowić nad stworzeniem systemu składającego się z sieci obu typów. Takie połączenie powinno zwiększyć uniwersalność rozwiązania i zapewnić większą stabilność działania. Dalszych prac wymaga również algorytm wykrywania dna. Wprawdzie wyniki sieci mogą sugerować, że wszystko zostało wyznaczone prawidłowo, jednak w niektórych sytuacjach dno nie jest wyznaczane w sposób optymalny. System ten nie będzie samodzielnym programem, lecz zostanie włączony jako moduł do systemu monitorowania zagrożeń płodu wspomaganego komputerem – MONAKO, który od wielu lat jest rozwijany w Instytucie Techniki i Aparatury Medycznej (ITAM) w Zabrzu.

BIBLIOGRAFIA

1) Bręborowicz G.H., Gadzinowski J.: Biofizyczna diagnostyka płodu i noworodka, Ośrodek Wydawnictw Naukowych, 1998.
2) Chen C.H.: Fuzzy logic and neural network handbook, IEEE Press, 1996.
3) Cyran K.A.: Wykłady z przedmiotu „Sztuczne sieci neuronowe”, 2005.
4) Hon E.H.: An atlas of fetal heart rate patterns, New Haven: Harty Press, 1968.

5) Jeżewski J.: Wpływ metody wyznaczania linii podstawowej częstości uderzeń serca na parametry zmienności długoterminowej, 1997, Praca doktorska, Akademia Medyczna w Poznaniu.
6) Jeżewski J., Wróbel J., Horoba K.: Comparison of Doppler ultrasound and direct electrocardiography acquisition techniques for quantification of fetal heart rate variability, IEEE Transactions on Biomedical Engineering, 2006, Vol. 53, 855-864.
7) National Institute of Child Health and Human Development Research Planning Workshop, Electronic fetal heart rate monitoring: Research guidelines for interpretations, American Journal of Obstetrics and Gynaecology, December 1997, Vol. 177, 1385-1390.
8) Sabbatini R., M., E.: Neural Networks for Classification and Pattern Recognition of Biological Signals, Proc. of 15th International IEEE/EMBS, 1993, San Diego, 265-266.
9) Słomko Z.: Biofizyczne monitorowanie w medycynie perinatalnej, PWN, 1991..
10) StatSoft: *STATISTICA 7 Sieci Neuronowe*, Dokumentacja programu, 2005.
11) Tadeusiewicz R.: Sieci Neuronowe, Akademicka Oficyna Wydawnicza RM, 1993.
12) Warrick P., Emily Hamilton E., Macieszczak M.: Neural Networks Based Detection of Fetal Heart Rate Patterns, Proc. of Inter. Joint Conference of Neural Networks, Montreal, 2005, 23-28.