

Neuronowy system informatyczny wspomagający proces prognozowania masy mleka uzyskanego podczas doju

Streszczenie

W pracy podjęto badania mające na celu określenie możliwości wykorzystania sztucznych sieci neuronowych w procesie prognozowania poziomu wydajności mleczności krów. Celem było zaprojektowanie, wytworzenie i przetestowanie systemu informatycznego, opartego na technologii sztucznych sieci neuronowych, służącego do predykcji masy wydojonego mleka. Proponowana aplikacja „Milko” jest systemem informatycznym, którego zadaniem jest prognoza masy uzyskanego w czasie doju mleka. Jej jądrem jest sztuczna sieć neuronowa, nauczona z wykorzystaniem pozyskanych danych empirycznych. Program uzupełniony jest w system pomocy zawierający ogólne wiadomości o produkcji mleka przez krowy, informacje o zasadach korzystania z aplikacji, będące wsparciem dla potencjalnego użytkownika oraz elementarne wiadomości o sieciach neuronowych.

Wprowadzenie

Ostatnia dekada to dynamiczny rozwój nauk informatycznych oraz technologii komputerowych. Ten bezsporny fakt stworzył m.in. możliwości rozwoju szeregu dyscyplin naukowych, w tym również szeroko rozumianej inżynierii rolniczej. Coraz częściej proponowane jest efektywne wykorzystanie w rolnictwie nowoczesnych metod informatycznych związanych ze sztuczną inteligencją, a w szczególności aplikacji bazujących na metodach sztucznych sieci neuronowych. Główną zaletą modeli neuronowych jest ich zdolność do uogólniania nabytej, podczas uczenia, wiedzy. Obszary, w których modele neuronowe dobrze się sprawdzają to, m.in. prognozowanie, rozpoznawanie, klasyfikowanie oraz diagnozowanie.

Celem pracy było zaprojektowanie, wytworzenie i przetestowanie systemu informatycznego, opartego na technologii sztucznych sieci neuronowych, służącego do predykcji masy wydojonego mleka.

Podstawą do podjęcia tematu była obserwacja rosnącej dynamiki zmian zachodzących w mleczarstwie, a zwłaszcza intensyfikacja produkcji wyrobów, których półproduktem jest mleko. Wzrastająca konkurencja na rynku mleka i stale rosnące wymagania co do jakości surowca, zmuszają producentów do zwiększania wydajności krów oraz zmniejszania strat. W związku z tym, w celu przewidywania poziomu produkcji mleka, przy wystąpieniu określonych warunków w czasie doju, zaistniała potrzeba opracowania narzędzia do prognozowania tego zjawiska w warunkach polskich. Opracowany system może być również wykorzystany przez rolników indywidualnych, jako instrument wspomagający procesy decyzyjne w produkcji mleka.

Projektowanie i implementacja sztucznej sieci neuronowej

Zgodnie z przyjętymi założeniami jądrem całego systemu informatycznego winna być, uprzednio wygenerowana, sztuczna sieć neuronowa. Dlatego najważniejszym etapem pracy nad powstaniem systemu było wytworzenie adekwatnego modelu neuronowego o jak najlepszej jakości [2]. Zostało to osiągnięte w kilku fazach:

1. W oparciu o studia literaturowe i badania własne wybrano następujące czynniki mające znaczący wpływ na ilość wydojonego mleka:
 - wiek krowy,

- zawartość krwi rasy holsztyńsko-fryzyjskiej,
- dzień laktacji, w jakim znajduje się krowa,
- sposób doju,
- pora doju.

Wielkości te uznano za reprezentatywne cechy charakteryzujące proces pozyskiwania mleka. W konsekwencji przyjęto je jako pięć zmiennych wejściowych dla tworzonej sieci neuronowej. Jako pierwszą zmienną wyjściową przyjęto masę wydojonego mleka [1].

2. Budowa adekwatnego zbioru uczącego w oparciu o posiadane dane empiryczne. Materiał badawczy do pracy pozyskano w oparciu o dane pochodzące z prac magistrów wykonanych w Instytucie Inżynierii Rolniczej. Kolejnym etapem było uporządkowanie danych i zbudowanie zbioru uczącego, który posłużył następnie do zaprojektowania i uczenia sieci neuronowej (rys. 1). Zbiór zawierał 154 przypadki uczące i podzielony został standardowo na podzbiory w stosunku odpowiednio 2 : 1 : 1 (uczący : walidacyjny : testowy).
3. W trakcie testów zauważono, że wyniki błędów dla zbiorów walidacyjnego i testowego są do siebie zbliżone. Sugerowało to, że sieć dobrze generalizuje nabytą wiedzę. Zrezygnowano więc ze zbioru testowego i przeznaczono jego przypadki na rzecz powiększenia zbiorów uczącego i walidacyjnego w celu uzyskania jeszcze lepszych wyników.
4. Przetestowanie wybranych algorytmów uczących i wyszukanie odpowiedniej topologii sztucznej sieci neuronowej. Najlepszy neuronowy model predykcji masy wydojonego mleka, na podstawie posiadanego zbioru danych, uzyskano przy zastosowaniu jednokierunkowych sieci wielowarstwowych typu perceptron wielowarstwowy *MLP (MultiLayer Perceptron)*, uczonych z wykorzystaniem algorytmu wstecznej propagacji błędów *BP (Back Propagation)* i douczanych przy użyciu algorytmu *gradientów sprzężonych*. Na rys. 2. przedstawiono schemat wygenerowanego modelu neuronowego [4].

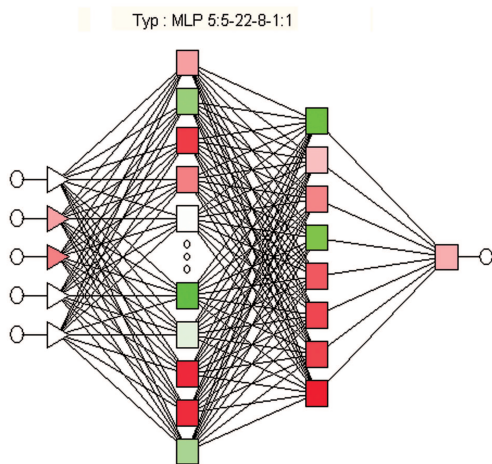
Technologie użyte w trakcie projektowania systemu informatycznego „Milko”

Według modelu kaskadowego pierwszą fazą tworzenia aplikacji jest określenie wymagań. Efektem jest sporządzenie dokumentu zawierającego konkretne cele, wymagania, zakres i

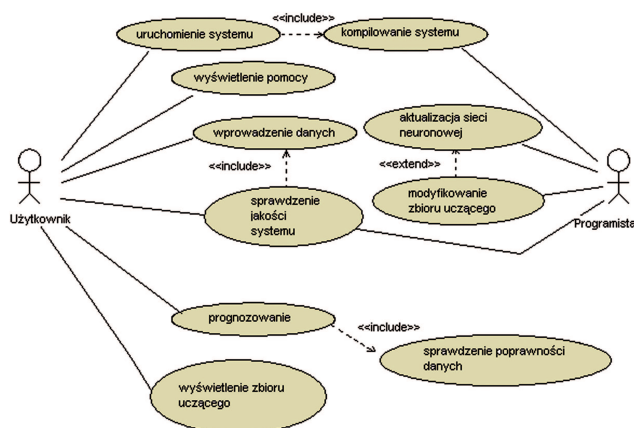
	A	B	C	D	E	F
	HF [%]	Wstaw funkcję k krowy	Pora doju	Dzień laktacji	Sposób doju	Masa wydojonego mleka [kg]
1	62,5	13	rano		przezienny	13,5
2	62,5	13	rano		równoczesny ze stymulacją wymienia	10,6
3	62,5	13	wieczór		równoczesny ze stymulacją wymienia	7,1
4	62,5	13	wieczór		przezienny	7,7
5	62,5	13	rano		równoczesny	10,0
6	62,5	13	rano		przezienny ze stymulacją wymienia	11,6
7	62,5	13	wieczór		przezienny ze stymulacją wymienia	7,5
8	62,5	13	wieczór		równoczesny	8,0
9	50	5	rano	154	równoczesny ze stymulacją wymienia	7,4
10	50	5	rano	155	równoczesny ze stymulacją wymienia	7,6
11	50	5	wieczór	150	równoczesny	5,2
12	50	5	rano	148	równoczesny ze stymulacją wymienia	6,8
13	50	5	wieczór	154	równoczesny ze stymulacją wymienia	5,1
14	50	5	rano	149	równoczesny ze stymulacją wymienia	7,0
15	50	5	rano	147	równoczesny ze stymulacją wymienia	6,1
16	50	5	rano	152	równoczesny	7,3
17	50	5	rano	154	równoczesny ze stymulacją wymienia	7,4
18	50	5	rano	151	równoczesny	7,3
19	50	5	rano	151	równoczesny	7,3

Rys. 1. Fragment zbioru uczącego

ogólny opis projektowanej aplikacji. Kolejny etap, zwany fazą projektowania, polega na wykonaniu szczegółowego opisu implementacji systemu. W fazie tej niezwykle ważna jest znajomość środowiska programistycznego, gdyż jego możliwości i ograniczenia determinują funkcjonalność przyszłego systemu [6]. Prócz założeń projektowych, rezultatem postępowania powinny być diagramy modelujące obiektowe podejście do wytwarzanego systemu (rys. 3). Do wykonania diagramu posłużono się programem Dia 0.96.



Rys. 2. Topologia wytworzonej sieci neuronowej typu MLP



Rys. 3. Diagram przypadków użycia

Zgodnie z założeniem jądrem aplikacji ma być wytrenowana sztuczna sieć neuronowa. Konieczny więc stał się wybór narzędzia do budowy i uczenia sieci neuronowych. Do tego celu wykorzystano moduł „Sieci neuronowe” zaimplementowany w programie *Statistica v.7.0*. Wygenerowana sieć neuronowa wyeksportowana została następnie do generatora kodu, umożliwiającego eksport kodu sieci neuronowej do tworzonej aplikacji „*Milko*”.

Obecnie dostępnych jest wiele, zróżnicowanych pod względem efektywności, narzędzi programistycznych. Przy projektowaniu budowanego systemu informatycznego „*Milko*” wykorzystano środowisko *Borland C++ Builder 6.0*. Uznano, że platforma ta, poprzez swoją wszechstronność, bogaty zestaw narzędzi i nowatorskie rozwiązania, będzie najlepszym środowiskiem do wytworzenia zamierzonego oprogramowania [7]. Aplikacja pisana była w języku *C++*, więc wygenerowany przez program *Statistica v.7.0* kod *C* musiał ulec pewnym modyfikacjom wynikającym z ewolucji rodziny języków *C*.

Do aplikacji dołączony został system pomocy, którego celem jest szybkie nauczenie użytkownika pracy z programem i pomoc w odpowiednim doborze zakresu danych. System pomocy wykonany został za pomocą programu *Help&Manual 4*.

Testowanie programu przeprowadzane było podczas implementacji programu, a także po wytworzeniu wszystkich modułów systemu informatycznego. Celem testowania była przede wszystkim detekcja i korekta ewentualnych błędów w aplikacji, jak również ocena jej niezawodności. Odbywało się ono na dwa sposoby: statyczny, sprowadzający się do analizy kolejnych instrukcji kodu, i dynamiczny, polegający na porównaniu rezultatów otrzymanych po wykonaniu kolejnych fragmentów aplikacji z wynikami oczekiwanymi. Bardzo ważnym i wysoce zaawansowanym narzędziem, umożliwiającym wykrycie wszystkich błędów w aplikacji, nawet tych najbardziej ukrytych jest zawarty w środowisku *C++ Builder*) *debugger*. Umożliwia on obserwację przebiegu poszczególnych funkcji aplikacji, znajdowanie błędnych fragmentów kodu, pracę krokową czy punkty wstrzymania, śledzenie zmiennych, a także wybór poziomu ostrzegania.

Ostatni etap (według modelu kaskadowego) stanowiła faza konserwacji oprogramowania, polegająca na przekazaniu użytkownikowi wytworzonej i przetestowanej aplikacji. Zadaniem użytkownika jest zgłaszanie wykrytych błędów i propozycji udoskonaleń do twórców oprogramowania.

Testowanie aplikacji „Milko”

Testowanie i ocena jakości wytworzonej aplikacji ma duże znaczenie dla poprawności jej działania. Proces ten przekłada się na zweryfikowanie czy wytworzona aplikacja spełnia ogólne normy przyjęte dla systemów IT oraz czy spełnia postawione przed nią zadania w fazie określania wymagań klienta.

Weryfikacja

Weryfikacja polegała na sprawdzeniu czy wytworzony system spełnia postawione mu uprzednio wymagania zdefiniowane w poprzednich fazach cyklu życia oprogramowania.

Proces weryfikacji przeprowadzony został bezpośrednio po zakończeniu pracy przy projektowaniu i implementacji systemu „Milko”. Odbył się on w dwóch etapach:

- statycznym, sprowadzającym się do analizy kolejnych instrukcji kodu,
- dynamicznym, polegającym na porównaniu rezultatów otrzymanych podczas działania aplikacji z wynikami oczekiwanymi, otrzymanymi podczas doświadczeń polowych.

Wynikiem weryfikacji była pełna zgodność funkcjonowania aplikacji z wymaganiami postawionymi w fazie określania wymagań oraz poprawność przeprowadzania procesu prognozowania. Uznano, że porównanie wyników empirycznych i wyników otrzymanych podczas korzystania z aplikacji będzie najlepszym sposobem oceny jakości systemu. Po przeprowadzonym testowaniu uznano, że błąd mieści się w dopuszczalnych granicach i umożliwia przeprowadzenie wiarygodnej prognozy w przyjętym zakresie danych.

Walidacja

Zadaniem walidacji jest wykazanie, że wytworzony system nadaje się do praktycznego wykorzystania. Program „Milko” przetestowany został przez grupę potencjalnych użytkowników, czego efektem było określenie jego cech funkcjonalnych:

- przyjazny interfejs użytkownika: wszystkie funkcje programu umieszczone są w rozwijanym menu głównym programu i nazwane są w taki sposób, aby użytkownik nie miał wątpliwości, jak otworzyć interesujący go moduł,
- szybkość: obsługa, wyświetlanie modułów, jak i sam proces prognozowania odbywa się bardzo szybko, co nie naraża użytkownika na oczekiwania i straty czasu,
- łatwość obsługi: poruszanie się po programie jest bardzo łatwe i intuicyjne.

Dodatkowo program „Milko” przetestowany został pod kątem zgodności działania na różnych systemach operacyjnych z rodziny „Windows” i na różnych platformach sprzętowych, co umożliwiło określenie jego wymagań systemowych i sprzętowych.

Program „Milko” nie wymaga instalowania dodatkowych bibliotek zewnętrznych, gdyż wszystkie wymagane składniki oprogramowania zostały dołączone do programu na etapie kompilacji, poprzez użycie odpowiednich opcji kompilatora.

Minimalne wymagania sprzętowe i systemowe:

- komputer klasy PC z procesorem Pentium,
- pamięć operacyjna RAM 64 MB (zalecane 128 MB),
- 20 MB wolnego miejsca na dysku HDD,
- system operacyjny Windows 98/2000/XP.

Opis systemu informatycznego „Milko”

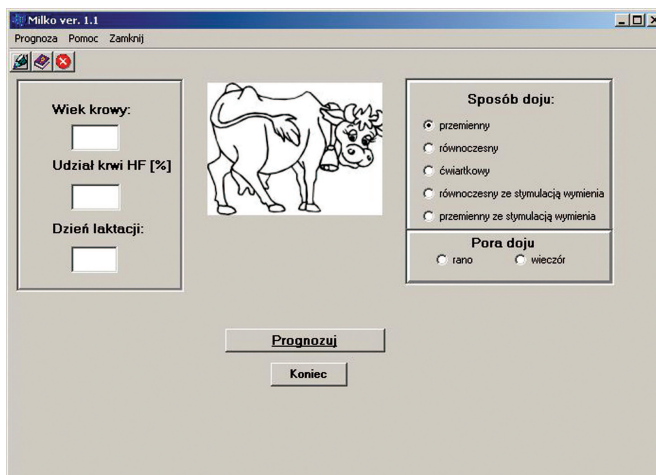
Program został skompilowany w taki sposób, aby nie była konieczna procedura instalacyjna. Program „Milko.exe” można uruchomić bezpośrednio z dowolnego folderu, w

którym się znajduje zbiór wykonywalny. W tym samym folderze powinien znajdować się zbiór pomocy „Milko.hlp” oraz podfolder „Dane”, w którym zawarty jest zbiór uczący „DaneKompl.txt.” Dzięki temu program można uruchomić z dowolnego nośnika danych, np. pendrive'a, dyskietki itp. Plikem uruchamiającym aplikację jest „Milko.exe”. Po uruchomieniu pojawia się okno informacyjne (rys. 4), które mówi użytkownikowi o zasadzie działania programu.

Po kliknięciu przycisku start wyświetli się główne okno prognozy (rys. 5).



Rys. 4. „Milko” okno informacyjne



Rys. 5. „Milko” - główne okno prognozy

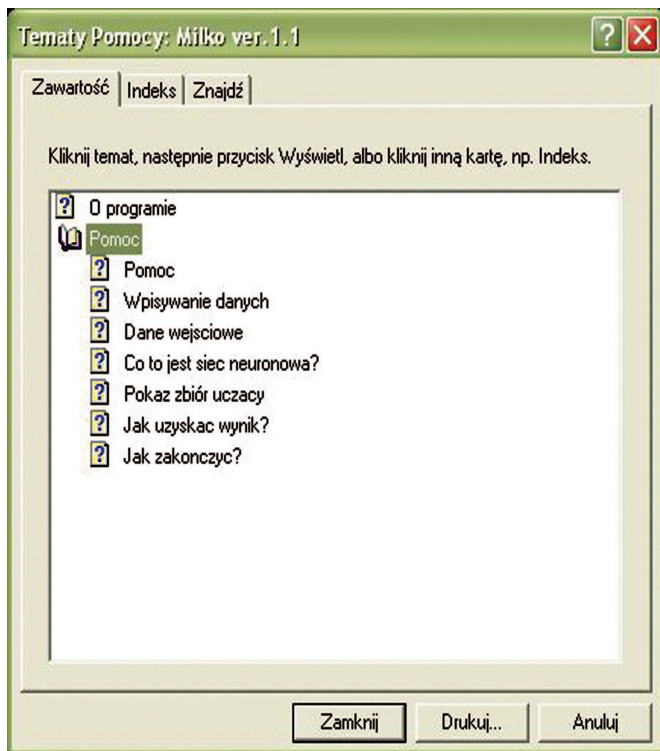
Na główne okno prognozy składają się:

- Kontrolki służące do wprowadzania danych wejściowych. Trzy kontrolki *MaskEdit* do wprowadzania danych numerycznych o określonej masce i siedem kontrollek *RadioButton*, zgrupowane w osobnych panelach umożliwiające wybór sposobu i pory doju.
- Przycisk „Prognozuj” uruchamiający mechanizm prognozowania.
- Kontrolka „Label”, za pośrednictwem której wyświetlany zostaje wynik prognozy.
- Obraz graficzny.
- Menu główne.

W menu głównym okna prognozowania dostępne są następujące opcje:

- Prognozuj masę mleka - uruchamia mechanizm prognozowania.
- Pokaż zbiór uczący - otwiera okno tekstowe z całym zbiorem uczącym, który posłużył do trenowania sieci. Daje to możliwość użytkownikowi wykorzystania danych w symulatorze sieci i uczenia własnych modeli neuronowych.

- Pomoc - uruchamia system pomocy, dający użytkownikowi obszerniejsze informacje na temat zmiennych wejściowych. Ważną informacją na tym panelu są zakresy danych, jakie przyjmuje program do prognozowania.
- O programie wyświetla ogólne informacje na temat programu.
- Zamknij aplikację - zamyka okno i powraca do okna informacyjnego.



Rys. 6. Okno pomocy aplikacji „Milko”

Wprowadzanie danych wejściowych obarczone jest ścisłą kontrolą błędów, umożliwiającą podanie ich tylko z góry określonego zakresu. W momencie, kiedy użytkownik wprowadzi na wejście sieci daną wykraczającą poza ustalony zakres, poinformowany zostanie o tym odpowiednim komunikatem i proces prognozowania nie zostanie przeprowadzony. O ustalonym zakresie użytkownik jest informowany na bieżąco po najechnaniu na wybrane pole, w postaci dymka informacyjnego. Inną możliwością uzyskania informacji o dopuszczalnych zakresach jest skorzystanie z pomocy w menu „Pomoc” (rys. 6).

Uwagi końcowe

1. Wytworzony system komputerowy „Milko” pozwala prognozować wielkość produkcji mleka. Złożoność modelowanego zjawiska pozwala przyjąć, że średni błąd predykcji popełniany przez aplikację jest akceptowalny.
2. Proponowany system informatyczny spełnia założenia funkcjonalne postawione w fazie określania wymagań oraz wszelkie wymogi inżynierii oprogramowania pod względem jakości i funkcjonalności.
3. Na podstawie przeprowadzonych testów stwierdzono, iż techniki neuronowe są właściwym narzędziem generalizacyjnym, umożliwiającym wytworzenie modelu do prognozy produkcji mleka.
4. Najlepszy neuronowy model do predykcji masy wydojonego mleka, na podstawie posiadanego zbioru danych, uzyskano przy zastosowaniu jednokierunkowych sieci wielowarstwowych typu *MLP* uczonych z wykorzystaniem algorytmu wstecznej propagacji błędów i douczanych przy użyciu algorytmu gradientów sprzężonych.
5. Zaprojektowana aplikacja jest narzędziem umożliwiającym oszacowanie ilości wydojonego mleka na podstawie danych agrotechnicznych. Może być również z powodzeniem wykorzystana przez rolników w celu doboru odpowiednich zwierząt przeznaczonych do użytkowania.

Literatura

- [1] Rutkowska D., Piliński M., Rutkowski L.: Sieci neuronowe, algorytmy genetyczne i systemy rozmyte. Wydawnictwo Naukowe PWN, Warszawa-Łódź 1997
- [2] Osowski S.: Sieci neuronowe do przetwarzania informacji. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2000
- [3] Hertz J., Krogh A., Palmer R. G.: Wstęp do teorii obliczeń neuronowych. WNT, Warszawa 1993
- [4] Boniecki P.: Sieci neuronowe typu MLP oraz RGB jako komplementarne modele aproksymacyjne w procesie predykcji plonu pszenżyta. Journal of Research and Applications in Agricultural Engineering, Poznań 2004, Vol. 49(1), str. 28-33
- [5] Boniecki P., Weres J.: Wykorzystanie technik neuronowych do predykcji wielkości zbiorów wybranych płodów rolnych. Journal of Research and Applications in Agricultural Engineering, Poznań 2003, Vol. 48, str. 56-60
- [6] Jaszkiwicz A.: Inżynieria oprogramowania. Wydawnictwo Helion S.A., Gliwice 1997
- [7] Daniluk A.: C++Builder. Kompendium programisty.

Neural information system helping in the process of predicting the mass of obtained milk during milking

Summary

Neural techniques constitute an alternative to describing and analyzing empirical systems in agricultural engineering. Artificial neural networks, referred to as neural networks for short, are a rapidly growing field of knowledge with applications reaching into many areas of science. The properties of such networks are ideal for many practical applications. Neural networks constitute a universal approximation system for representing multi-dimensional data sets. They have an ability to learn and adapt to changing environments and generalize on the acquired knowledge, which makes them a system of artificial intelligence. At the heart of the network's operation are learning algorithms which allow for designing a proper network structure and selecting its parameters to best suit the problem at hand. The thesis describes utilisation of neural network model type MLP (MultiLayer Perceptron) to build information system “Milko” for prediction the mass of milk during milking.