

**ПРИЛОЖЕНИЕ НА NEURO – FUZZY КЛАСИФИКАТОР ЗА
РАЗПОЗНАВАНЕ НАЛИЧИЕТО НА БАКТЕРИИ В СУРОВО МЛЯКО ПО
СПЕКТРАЛНИТЕ ИМ ХАРАКТЕРИСТИКИ В БЛИЗКАТА
ИНФРАЧЕРВЕНА ОБЛАСТ**

**ADAPTIVE NEURO – FUZZY CLASSIFIER APPLIED TO NEAR
INFRARED SPECTRA OF RAW MILK FOR RECOGNITION OF
BACTERIAL PRESENCE**

Петя Велева - Донева
*Технически Университет Габрово,
pveleva@uni-sz.bg*

Цветелина Драганова
*Русенски Университет “Ангел Кънчев”,
cgeorgieva@uni-ruse.bg*

Стефка Атанасова
*Тракийски Университет Стара Загора,
atanassova@uni-sz.bg*

Валентина Кукенска
*Технически Университет Габрово,
vally@tugab.bg*

Abstract

Adaptive neuro – fuzzy classifier applied to near infrared spectra of raw milk for recognition of bacterial presence is presented in the paper. Healthy milk samples and milk samples contaminated with Staphylococcus aureus and Streptococcus agalactiae were used. Principal component analysis was used for spectral data processing to reduce the recognition features. First three components were used as a input data for the classifier. The output is one – sample class – contaminated or non-contaminated. The obtained average percent error for class non-contaminated samples is 3,88% and for class contaminated is 10,06%.

Keywords: milk, bacteria contamination, NIR spectroscopy, Neuro - Fuzzy, PCA

ВЪВЕДЕНИЕ

Един от основните проблеми за световната млечна промишленост, водещ до значителни икономически загуби е заболяването мастит при животните [4]. Той е причина за значително намаляване на млечния добив и промени в състава на млякото, което от своя страна води до рязко влошаване на

качеството на произведените млечни продукти и намаляване на срока за съхранение. Допълнителните загуби произтичат и от разходите за лечение, извънреден труд, както и от преждевременно унищожаване на животните.

Bennett, R. et. al. [2] изразяват своето становище, че във Великобритания маститът е икономически най-скъпото заболяване на

едрия рогат добитък. Yalcin, C. [17] изчислява, че разходите за лечение и превенция са между £ 69 - £ 228 на крава годишно, в зависимост от разпространението на заболяването или около £ 42 000 000 годишно [11]. Според Jones, G. M. et. al. [10] разходите в САЩ са около 1,7 до 2 милиарда долара всяка година или 11% от цялата млечкопреработваща индустрия.

Ето защо качеството на произвежданото мляко и неговата годност е въпрос, чиято актуалност в световен мащаб придобива все по-голямо значение. В тази връзка от съществено значение за млечкопреработващата индустрия и млечодобиващите ферми е използването на автоматизирани системи за контрол на производствения процес и качеството на произвежданата продукция, която да отговаря адекватно на изискванията на пазара.

В последните години се изследват възможностите на спектралният анализ в близката инфрачервена област (Near Infrared Spectroscopy - NIRS) за определяне на различни компоненти от състава на млякото като алтернатива на класическите химически и микробиологични методи за анализ [9,15]. Методът позволява бърз и точен анализ, без нужда от предварителна химична обработка на пробите. Той е недеструктивен, което позволява едновременно да се определят няколко компонента в анализираната проба. Основно негово предимство е, че анализът може да се извършва не само в специализирани лаборатории, а на място, на различни производствени линии и др. Освен за определяне на състава, публикувани са и опити за използването на NIR спектроскопия за идентификация на различни бактериални видове [5, 6, 14].

Сложният, нелинеен характер на спектралните характеристики на анализираните биологични обекти изисква използването на интелигентни подходи за обработка на данните. Ето защо напоследък ефективно се прилага NIRS, комбинирана с невронни мрежи и размита логика [7,16].

Един от най-използваните класификатори, базиран на размита логика и невронни мрежи е Adaptive Neuro – Fuzzy Inference System (ANFIS) [1,8].

Целта на това изследване е да се проучи възможността за откриване на патогенни

микроорганизми в проби от краве мляко чрез използване на спектроскопия в близката инфрачервена област и невронен класификатор, базиран на Fuzzy логика.

ИЗЛОЖЕНИЕ

Едни от основните патогени, най-често причиняващи мастит при животните са *Staphylococcus aureus* и *Streptococcus agalactiae* [3,12]. Този вид бактерии са обект и на нашето изследване.

В експеримента са използвани два вида проби – четвъртинкови проби от краве мляко и контаминирани със *Staphylococcus aureus* и *Streptococcus agalactiae*.

Четвъртинковите проби са събрани от 37 крави от порода Holstein (Awaji Agricultural Technology Institute, Japan) два пъти месечно, за девет месечен период преди сутрешното доене. Бактериологичният анализ е извършен съгласно приетите стандарти [13]. Наличието на *Staphylococcus aureus*, *Streptococcus agalactiae* и други бактерии от вида *Streptococcus* са открити в някои от пробите чрез класическите микробиологични методи.

Експерименталната група проби е приготвена чрез контаминиране на 9.9 ml мляко с 0.1 ml култура съдържаща от 10^2 CFU/ml до 10^{10} CFU/ml концентрация на бактерии от вида *Staphylococcus aureus* или *Streptococcus agalactiae*. 36 млечни проби са контраминирани с бактерии от вида *Staphylococcus aureus* и съответно 34 проби са контраминирани със *Streptococcus agalactiae*. Всички проби са поставени в инкубатор за 24 часа при температура 5°C преди спектралното измерване. Непосредствено преди спектралния анализ всяка проба е загрята до температура 40°C на водна баня, която се поддържа постоянна по време на измерванията.

Спектралните характеристики са получени със спектрофотометър NIRSystem 6500 (FOSS NIRSystems, Silver Spring, MD, USA), чрез измерване на абсорбцията на слой мляко с дебелина 1 mm. Те се представят като $[\log(1/T)]$, където T представлява коефициент на пропускане през пробата. В изследването са използвани спектрите в диапазона съответно: 600 ÷ 1064 nm и 1136 ÷ 1880 nm с интервал 2 nm. Прекъсването в спектралния диапазон е направено с цел избягване на

смущенията, породени от смяната на детекторите на използвания спектрофотометър.

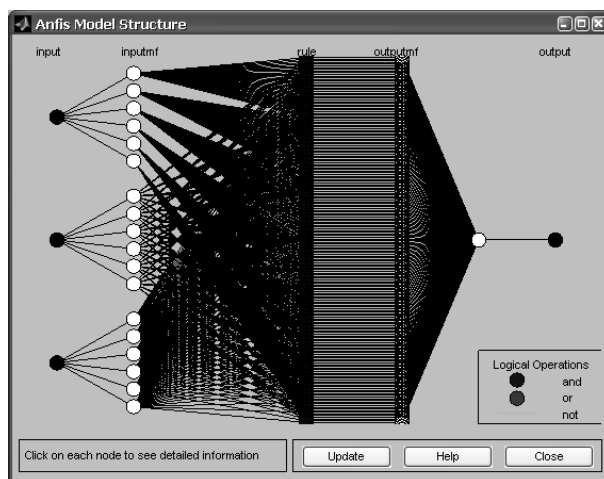
За всяка проба е определен клас както следва: на пробите инфектирани със *Staphylococcus aureus* или *Streptococcus agalactiae* е присвоен клас *Contaminated*, а на пробите, в които не е установено наличие на патогенни бактерии - клас *Non-contaminated*.

В изследването са използвани общо 325 проби от краве мляко. В 174 от тях не е установено наличие на патогенни бактерии. 60 млечни проби са контаминирани с бактерии от вида *Staphylococcus aureus* и съответно 91 проби са контаминирани със *Streptococcus agalactiae*. Измерените спектри са разделени в съотношение 2:1. Обучаващата извадка е формирана от 117 проби на здраво мляко и 104 проби, контаминирани с бактерии от вида *Staphylococcus aureus* и *Streptococcus agalactiae*. Тестовата извадка включва по 57 и 47 проби за съответните класове.

Приложен е Neuro-Fuzzy класификатор за разпознаване наличието на бактерии в сурово мляко по спектралните им характеристики в близката инфрачервена област. Класификаторът се базира на логически принципи за класификация. Спектралните характеристики от обучаващата извадка се използват за създаване на класификатора. Данните се трансформират чрез пресмятане на главните компоненти (PC). PCA намира линейни комбинации на оригиналните спектри, които описват максимална част от вариациите. Едновременно с това се намалява размерността на данните, без да се губи съществена информация.

Първите три главни компонента се използват като входни данни за класификатора, а изходът е само един – съответния зададен клас. Класификаторът има три скрити слоя. Структурата му е представена на фиг. 1.

Всяка входна променлива е представена чрез функция на принадлежност – Гаусов сигнал. ANFIS използва смесен алгоритъм за обучение от вид Sugeno и fuzzy inference systems (FIS), за да установи параметрите на функцията на принадлежност за всеки изход.



Фиг. 1. Структура на класификатора

За обучаващата FIS функция на принадлежност се използва комбинация от метод на най-малките квадрати и обратно разпространение на грешката за дадените входно / изходни данни.

При определянето на най – добрата топология на класификатора се използва средния процент грешка:

$$APE = \frac{1}{P} \sum_{i=1}^p \frac{|T_i - O_i|}{|T_i|} \cdot 100\% \quad (1) \quad (6)$$

където P е броят на пробите от обучаващата извадка;

T_i е i -ят желан изход;

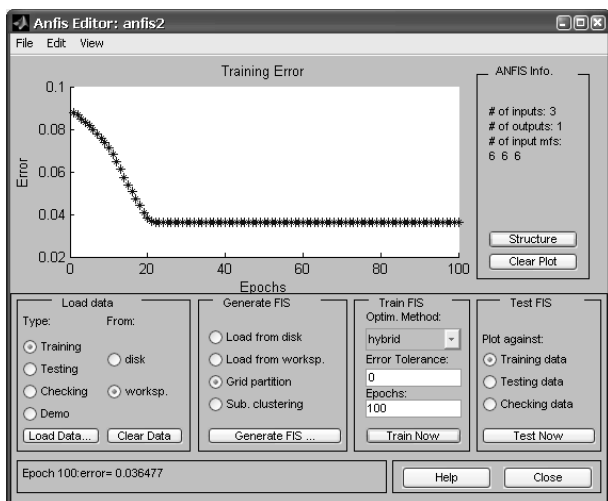
O_i е полученият i -ти изход от класификацията.

За реализирането на невронния класификатор се използва програмния продукт MathLab 7.9.0 (The MathWorks Inc., USA).

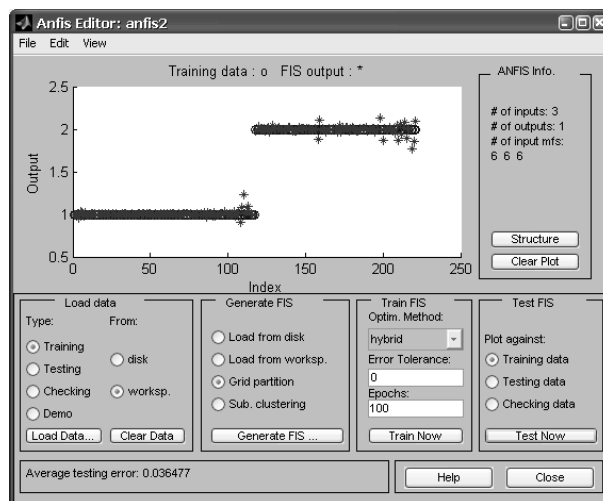
Грешката от обучението на невронно – размития класификатор е представена на фиг. 2.

Тестването на данните от обучаващата извадка дава минимална средна грешка – 0,036 (фиг. 3).

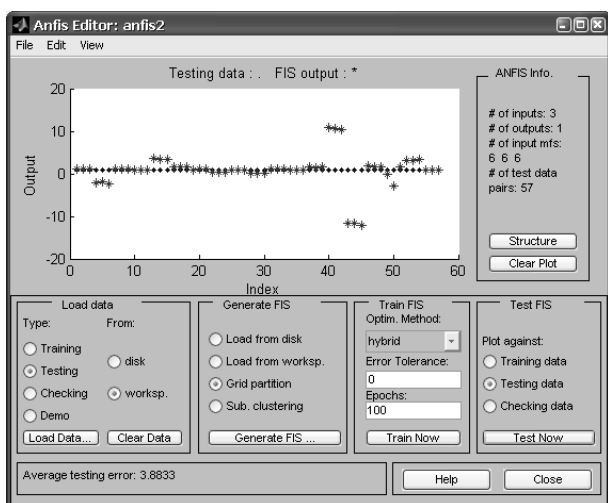
Резултатите, които се получават от класификацията на пробите от тестовата извадка за клас здрави е 3,88% средна грешка. За клас заразени проби грешката е 10,07%. Това вероятно се дължи на факта, че са обединени двата вида бактерии при проведеното изследване в един клас – клас *Contaminated*.



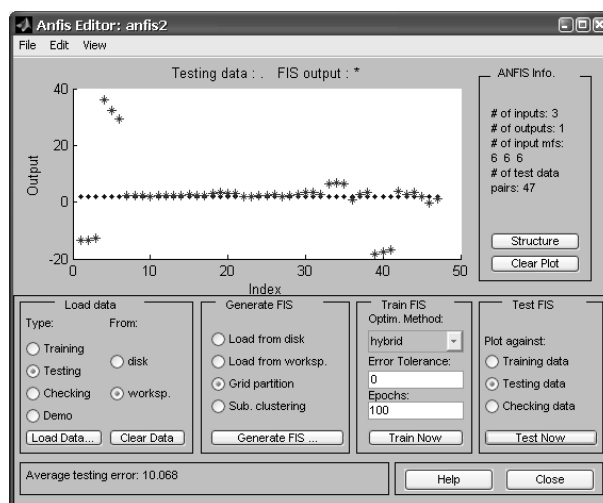
Фиг. 2. Грешка от обучението на невронно – размития класификатор



Фиг. 3. Тестване на класификатора с данните от обучаващата извадка



Фиг. 4. Тестване на класификатора с данните от тестовата извадка – здрави проби



Фиг. 5. Тестване на класификатора с данните от тестовата извадка – заразени проби

ЗАКЛЮЧЕНИЕ

Резултатите от анализа показват, че в спектрите на млечните проби в близката инфра-червена област се съдържа информация свързана с наличието и развитието на бактерии от изследваните видове.

Получените резултати дават основание да се твърди, че класификатор с невронно – размита структура може да се използва при разпознаване на здрави и заразени проби мляко. За клас *Non-Contaminated* средната грешка е задоволителна. Но за клас *Contaminated* ще се проведат допълнителни изследвания за двата класа поотделно с цел подобряване на грешката от разпознаване. Ще се изследва и влиянието на параметрите на невронно – размития класификатор върху точността на разпознаване на пробите.

БЛАГОДАРНОСТИ

Изследванията са подкрепени по договор № BG051PO001-3.3.04/28, „Подкрепа за развитие на научните кадри в областта на инженерните научни изследвания и иновациите”. Проектът се осъществява с финансовата подкрепа на Оперативна програма „Развитие на човешките ресурси” 2007-2013, съфинансирана от Европейския социален фонд на Европейския съюз“.

ЛИТЕРАТУРА

- [1] Babazadeh, Y., Mousavi, S., Akbarzadeh, M. (2008). Multidimensional Dynamic Modeling of Milk Ultrafiltration Using Neuro-Fuzzy Method and a Hybrid Physical Model, *Iranian Journal of Chemical Engineering*, Vol.5, No.2 (Spring), IChE
- [2] Bennett, R., Jpelaar, J. (2005). Updated estimates

- of the costs associated with thirty four endemic livestock diseases in Great Britain: a note, *Journal of Agricultural Economics*, 56:1, 135-144
- [3] Bradley, A.J., Green, M.J., Huxley, J.N. (2002). Making better use of milk samples: monitoring and investigating herd mastitis, *Cattle Practice*, 10:2, 105-112
- [4] Bradley, A., Leach, K., Breen, J., Green, L. and Green, M. (2007). Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales, *Veterinary Record*, 160, 253-258
- [5] Cena, H., He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality, *Trends in Food Science & Technology*, Vol.18, Issue 2, 72-83
- [6] Fagan, C., O'Donnell, C., Rudzik, L. and Wust, E. (2009). Milk and Dairy Products. In: *Infrared Spectroscopy for Food Quality and Control*. Sun, D.W. (ed.), Elsevier, Amsterdam, 241-273
- [7] Goodarzi, M., Olivieri, A., Freitas, M. (2009). Principal component analysis-adaptive neuro-fuzzy inference systems (ANFISs) for the simultaneous spectrophotometric determination of three metals in water samples, *J. Molecular and Biomolecular Spectroscopy*, 73:4, 608-614
- [8] Jang, S. R. (1993). ANFIS: Adaptive network based fuzzy inference system. *EEE Trans. Syst.*, Vol. 23(3), 665–685, Man, Cybern
- [9] Jankovska, R. and Šustova, K. (2003). Analysis of Cow Milk by Near-infrared Spectroscopy, *Czech J. Food Sci.*, 21, 123–128
- [10] Jones, G. M., Bailey, T. L. (2009). Understanding the Basics of Mastitis, ID 404-233
- [11] Kossaibati, M. A. and Esslemont, R. J. (2000). The costs of production diseases in dairy herds in England, *The Veterinary Journal*, 154: 41-51
- [12] Milne, M. H., Barrett, D.C., Fitzpatrick, J.L., Biggs, A.M. (2002). Prevalence and aetiology of clinical mastitis on dairy farms in Devon, *Veterinary Record*, 151:8, 241-243
- [13] Quinn, P.J., Carter, M.E., Markey, B.K., Carter, G.R. (1999). Clinical Veterinary Microbiology, Mastitis, 327-344
- [14] Saranwong, S., Kawano, S (2008). Interpretation of near infrared calibration structure for determining the total aerobic bacteria count of raw milk: interaction between bacterial metabolites and water absorptions, *J. Near Infrared Spectroscopic*, 16, 497-504
- [15] Tsenkova, R., Atanassova, S., Morita, H., Ikura, K., Toyoda, K., Iordanova, I., and Hakogi, E. (2006). Near infrared spectra of cows' milk for milk quality evaluation: disease diagnosis and pathogen identification, *J. Near Infrared Spectroscopic*, 14, 363 – 370
- [16] Wang, L., Zhang, C., Feng, Y. (2007). Neural Networks' Backward Forecasting Ability for Spectroscopy, *Proceedings of the International Conference University of Electronic Science and Technology of China, Information Computing and Automation*, 328-330
- [17] Yalcin, C. (2000). Cost of mastitis in Scottish dairy herds with low and high subclinical mastitis problems. *Turk Veterinerlik ve Hayvanclik Dergisi*, 24:5, 465-472