

3. Canini KR, Suh B, and Pirolli PL. 2011. Finding credible information sources in social networks based on content and social structure. 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing: IEEE. p 1-8.
4. A. Jain, A. Shakya, H. Khatter and A. K. Gupta, "A smart System for Fake News Detection Using Machine Learning," 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad, India, 2019, pp. 1-4, doi: 10.1109/ICICT46931.2019.8977659.
5. Sprenkamp, K., Jones, D. G., & Zavolokina, L. (2023). Large Language Models for Propaganda Detection. *arXiv preprint arXiv:2310.06422*.

### ПАРАМЕТРИЗАЦІЯ АДАПТИВНОЇ МОДЕЛІ ІНДИКАТОРУ СТАНУ ОРГАНІЗАЦІЙНОЇ СИСТЕМИ УПРАВЛІННЯ

Льїна О.П., Скибик С.Я. / O.P.Ilyina, S.Skybyk

Інститут програмних систем НАН України / Institute of Software Systems NAS of Ukraine,

Проспект Академіка Глушкова, 40 корп.5, Київ, 03187, тел. (044) 526-55-07,

E-mail: [ilyina.elenal@ukr.net](mailto:ilyina.elenal@ukr.net), [sskybyk@gmail.com](mailto:sskybyk@gmail.com).

The study develops weight coefficients for a weighted voting model assessing organizational system states using an ensemble of classification models. It identifies key input data requirements and supports adaptive indicators and retraining diagnostics.

Виклики створення індикаторів стану систем організаційного управління за умов глибокої невизначеності зумовили побудову моделі індикатору за допомогою гетерогенного ансамблю для класифікаційного машинного навчання, запропоновану в [1]. Інтеграція результатів  $R_i$  класифікаторів  $CL_i$  здійснюється в [1] зваженим голосуванням, при якому ваговий коефіцієнт для  $R_i$  має вигляд

$$P_i = 1/4 \sum_{j=1}^4 ((1 - KD_{ij}) \cdot KI_{ij}) \cdot Q_i$$

де  $KD_{ij}$  – рівень критичності  $j$ -ї вимоги від  $i$ -го класифікатора до властивостей вхідних даних;  $KI_{ij}$  (0 або 1) – індекс наявності порушення  $j$ -ї вимоги у навчальній вибірці;  $Q_i$  – коефіцієнт якості класифікації.

В даній роботі запропоновано аргументовані експертні оцінки значень  $KD_{ij}$  для ансамблю 5 класифікаторів з [1].

На основі аналізу широкого спектру публікацій, зокрема [2-4], кожен із методів був охарактеризований наявністю та рівнем критичності вимог, які він виставляє до використовуваних даних, згідно до моделі

$$MA = \{D_i, O, \{A_{ij}\}_{j=1, \dots, 4}\}_{i=1, \dots, 4}$$

де  $D_i$  – вид вимоги до характеристик (1 – вид розподілу, 2 – незалежність, 3 – збалансованість спостережень, 4 – шкала значень);

$O$  – оцінка рівня критичності вимоги,  $O \in \{0, 0.5, 1\}$ , де 0 – відсутність вимог (обмежень), 0.5 – припустимість порушень, 1 – критичність;

$A_{ij}$  –  $j$ -й аспект аргументації (1 – наявність первинного обмеження, 2 – методи досягнення робастності; 3 – доступність програмного продукту, що реалізує такі методи; 4 – ризики використання методів).

Аргументовані оцінки, наведені далі в табл.1, надають можливість використання ансамблевої моделі [1] в якості адаптивної щодо стану предметної області як при формуванні індикаторів за умов глибокої невизначеності, так і для діагностики потреб у повторному навчанні моделі використовуваного індикатору (на основі суттєвої відмінності між поточними та навчальними значеннями вагових коефіцієнтів для елементів ансамблю класифікаторів).

Табл. 1 – Оцінки рівня критичності вимог до навчальних даних з аргументацією

	Вид розподілу ознак $A_{1j}$	Незалежність ознак $A_{2j}$	Збалансованість спостережень $A_{3j}$	Тип значень $A_{4j}$
<b>Байсівський класифікатор</b>				

<i>O</i>	<b>0</b>	<b>0</b>	<b>0.5</b>	<b>0</b>
<i>D<sub>1</sub></i>	Необхідність точної специфікації ймовірнісної моделі	Припущення про незалежність ознак	Незбалансованість даних може знизити якість класифікації	Адаптація моделі до типів даних
<i>D<sub>2</sub></i>	Використання моделей змішування з процесом Діріхле, що дозволяють даним визначати форму розподілу	Моделювання залежностей через багатовимірні розподіли або умовні залежності	Використання апріорних розподілів, регуляризації, масштабування даних.	Вибір відповідних ймовірнісних розподілів для різних типів ознак
<i>D<sub>3</sub></i>	Пакети в R, такі як <i>brms</i> і <i>DirichletProcess</i>	Пакет <i>brms</i> у R	Пакети R <i>brms</i> , <i>caret</i>	Пакет <i>brms</i> у R
<i>D<sub>4</sub></i>	Мінімальні	Відсутні; врахування залежностей підвищує точність моделі	Ризик втрати інформативності через масштабування	Відсутні
<b>Метод опорних векторів SVM</b>				
<i>O</i>	<b>0</b>	<b>0</b>	<b>0.5</b>	<b>0.5</b>
<i>D<sub>1</sub></i>	Відсутне; класичні лінійні класифікатори не можуть ефективно розділити класи	Відсутне; необхідність оптимізації при сильній кореляції	Упередженість до більш численного класу	Чутливий до масштабу ознак
<i>D<sub>2</sub></i>	Використання ядрових функцій для моделювання складних нелінійних залежностей	Проектування даних у простори вищої розмірності за допомогою ядрових функцій	Використання вагових коефіцієнтів, методів ресемплінгу, SMOTE	Масштабування даних (нормалізація, стандартизація)
<i>D<sub>3</sub></i>	Пакети R <i>e1071</i> , <i>kernelab</i> , <i>caret</i> реалізують SVM з різними ядрами	Пакети R <i>e1071</i> , <i>kernelab</i> , <i>caret</i>	Пакети R <i>DMwR</i> , <i>ROSE</i> , <i>caret</i>	Пакет R <i>caret</i> надає інструменти для попереднього оброблення даних
<i>D<sub>4</sub></i>	Мінімальні; метод стійкий до різних розподілів	Незначні; вплив лише на обчислювальну складність	Можлива втрата інформативності	Некритичні обмеження
<b>Випадковий ліс</b>				
<i>O</i>	<b>0</b>	<b>0</b>	<b>0.5</b>	<b>0</b>
<i>D<sub>1</sub></i>	Відсутне	Відсутне	Упередженість до більш численного класу	Відсутне
<i>D<sub>2</sub></i>	Непотрібні	Випадковий вибір ознак при розщепленні знижує вплив кореляції	Зміна ваг класів, ресемплінг	Непотрібні
<i>D<sub>3</sub></i>	Пакети R <i>randomForest</i> , <i>ranger</i> , <i>randomForestSRC</i> , <i>imbalanced</i>			
<i>D<sub>4</sub></i>	Відсутні	Відсутні	Можлива втрата інформативності	Відсутні
<b>Метод найближчих сусідів kNN</b>				
<i>O</i>	<b>0</b>	<b>0.5</b>	<b>0.5</b>	<b>0.5</b>
<i>D<sub>1</sub></i>	Відсутне	Залежності ознак спотворюють результати	Упередженість до більш численного класу	Чутливість до масштабу ознак
<i>D<sub>2</sub></i>	Непотрібні	Зменшення розмірності (PCA, t-SNE)	Використання вагових коефіцієнтів, методів ресемплінгу	Нормалізація, стандартизація ознак
<i>D<sub>3</sub></i>	Пакети R <i>caret</i> , <i>ROSE</i> , <i>DMwR</i> , <i>kknn</i>			
<i>D<sub>4</sub></i>	Мінімальні	Можлива втрата інформативності	Втрата інформативності	Невірне масштабування зміщує результати
<b>Логістична регресія</b>				
<i>O</i>	<b>0</b>	<b>0.5</b>	<b>0.5</b>	<b>0.5</b>
<i>D<sub>1</sub></i>	Відсутне; ненормальність розподілу впливає на надійність	Мультиколінеарність між ознаками	Упередженість до більш численного класу	Потребує коректного кодування категоріальних змінних
<i>D<sub>2</sub></i>	Методи трансформації	Виключення кореляцій,	Використання вагових	One-hot encoding, effect

	(логарифми, корені) за наявності викидів	PCA, L1/L2 регуляризація	коефіцієнтів, методи ресемплінгу	coding, ordinal encoding
<i>D<sub>3</sub></i>	Пакет <i>stats</i> з функцією <i>glm()</i> у R, пакет <i>caret</i>			
<i>D<sub>4</sub></i>	Мінімальні	Можлива втрата інформативності	Можлива втрата інформативності	Неправильне кодування може знизити точність

#### Список використаної літератури

1. Ільїна О.П., Скибик С.Я. Моделювання ресурсного індикатору безпеки інтересів розподіленої системи організаційного управління з використанням класифікаційних методів машинного навчання // *Інформаційні технології і автоматизація – 2024*: матеріали XVII міжнародної науково-практичної конференції (Одеса, 31 жовтня – 1 листопада 2024 р.). – Одеса: Видавництво ОНТУ, 2024. – С. 69–72. – 847 с.
2. Bishop, C.M. (2006). *Pattern Recognition and Machine Learning*. New York: Springer.
3. Aggarwal, C.C. (2015). *Data Mining: The Textbook*. Springer.
4. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*. Springer.

### OPTIMIZING SKIN IMAGE SEGMENTATION WITH FOURIER AND GRAPH-BASED METHODS

**Кіншаков Е.В., Парфененко Ю.В.** / Kinshakov E.V., Parfenenko Yu., V.  
Сумський державний університет / Sumy State University  
40007, Суми, вул. Харківська, 116,  
E-mail: edikkinshakov@gmail.com, yuliya\_p@cs.sumdu.edu.ua

This paper introduces advanced methods for skin disease image segmentation using the Dermnet dataset, one of the largest resources in dermatology. Traditional approaches like Watershed and thresholding often fail due to the complex textures, color variations, and noise present in skin images. To address these challenges, novel techniques were proposed. First, the Fourier transform reduces high-frequency noise, preparing images for segmentation. Then, min-cut/max-flow graph algorithms minimize energy functions, enabling precise separation of pathological and healthy areas. Additionally, a piecewise smooth approximation improves boundary detection, refining segmentation results. Experiments demonstrated a 15% accuracy improvement over traditional methods. Processing time was also significantly reduced, enhancing the reliability and efficiency of automated diagnostic systems.

**Keywords:** segmentation, machine learning, image processing, skin diseases, Fourier transform, graph algorithms, computational optimization, piecewise approximation.

**Introduction and Problem Statement.** Image segmentation for medical diagnostics, particularly for skin diseases, poses challenges due to the complex textures, uneven boundaries, and noise in affected areas. The Dermnet dataset, a rich resource for research, highlights the limitations of traditional methods like thresholding and Watershed, which struggle with accuracy. Key challenges include distinguishing healthy and pathological areas, uneven illumination, and texture noise.

To address these issues, advanced methods have been proposed: Fourier transforms to reduce noise, graph algorithms for precise boundary segmentation, and piecewise smooth approximation for improved detection of pathological areas. This integrated approach enhances accuracy and minimizes false positives, offering a significant improvement over traditional algorithms.

**Problem solution and Results.** The first step in segmentation is preprocessing to remove noise and artifacts affecting accuracy. High-frequency noise from uneven illumination, glare, and skin texture is addressed using a Fourier transform, which filters out unwanted components in frequency space. Formally, this can be written as: