

**О.А. ЖУКОВСЬКА**

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна, e-mail: zhukovskaya71@gmail.com.

**Л.С. ФАЙНЗІЛЬБЕРГ**

Міжнародний науково-навчальний центр інформаційних технологій і систем НАН та МОН України; Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна, e-mail: fainzilberg@gmail.com.

## **ОЦІНЮВАННЯ КОРИСНОСТІ БІНАРНОГО КЛАСИФІКАТОРА НА ОСНОВІ УДОСКОНАЛЕНОГО МЕТОДУ ROC-АНАЛІЗУ**

**Анотація.** Сформульовано означення корисності бінарного класифікатора в сенсі зменшення апіорного ризику помилкової класифікації. Запропоновано достатні умови, що гарантують корисність діагностичного тесту згідно з цим означенням. Отримані умови надали змогу вдосконалити традиційний ROC-аналіз завдяки обмеженню відповідного фрагмента ROC-кривої. Показано, що пряма, яка обмежує фрагмент ROC-кривої гарантовано корисного тесту, збігається з відомою ізолінією ефективності, що відповідає рівню апіорного ризику. Визначено допустимі межі співвідношення втрат від помилки пропуску цілі та хибної тривоги, за яких тест із відповідними операційними характеристиками залишається корисним для скринінгу захворювання з відомим преваленсом. На основі отриманих результатів обґрунтовано ефективність нового методу аналізу та інтерпретації електрокардіограм, який ґрунтується на визначенні оригінальної діагностичної ознаки у фазовому просторі та надає змогу виявляти осіб з високим ризиком ішемічної хвороби серця на ранніх стадіях захворювання.

**Ключові слова:** бінарний класифікатор, ROC-крива, діагностична ознака, аналіз та інтерпретація ЕКГ.

### **ВСТУП**

Бінарні класифікатори широко застосовують у різних галузях. Так, наприклад, для профілактики захворювань потрібно проводити масові обстеження населення (скринінг) для виявлення пацієнтів з високим ризиком потенційно небезпечних хвороб [1]. У банківській сфері активно використовують системи скорингу [2], що забезпечують виявлення неблагонадійних позичальників. Перелік таких прикладів може бути продовжений.

Відомі різні підходи до оцінювання ефективності бінарних класифікаторів [3]. Зручним засобом є метод, який базується на аналізі так званої операційної характеристичної кривої (Receive Operating Characteristic curve, ROC) [4, 5]. Традиційний ROC-аналіз ґрунтується на оцінюванні двох характеристик діагностичного тесту: чутливості та специфічності [6–15], які, по суті, оцінюють імовірності помилок пропуску цілі та хибної тривоги, прийнятих у теорії статистичних рішень. Для інтегральної оцінки ефективності бінарного класифікатора, що реалізує діагностичне вирішувальне правило, найчастіше визначають площу під ROC-кривою [16].

У роботах [17–21] запропоновано розвинення ROC-аналізу для випадку, коли множина класів, що розпізнаються, містить більше двох діагнозів. Удосконалення методу ROC-аналізу для порівняння бінарних класифікаторів щодо очікуваних середніх втрат від неправильних рішень запропоновано в [22–24].

Мета цієї статті — подальше вдосконалення методу ROC-аналізу для забезпечення вибору діагностичного тесту, який гарантовано зменшує апіорний ризик помилкової діагностики.