

СРАВНЕНИЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ РАМАНОВСКИХ СПЕКТРОВ СЫВОРОТКИ КРОВИ

Е.В. Сорокина, Ю.А. Христофорова
«Самарский национальный исследовательский университет имени
академика С.П. Королева», г. Самара

Ключевые слова: спектроскопия комбинационного рассеяния, сыворотка крови, хроническая сердечная недостаточность

Тенденция роста хронической сердечной недостаточности (ХСН) является серьезной проблемой здравоохранения во всем мире [1]. Сегодня для диагностики проводят физиологические тесты и анализ крови на натрийуретический пептид мозгового типа (BNP), уровень которого повышается при повышенной нагрузке на мышцы сердца. Существует потребность в альтернативном методе для диагностики ХСН. Для этого может подойти спектральный анализ рамановского рассеяния, которое представляет собой суперпозицию сигналов, обусловленных вкладом активных молекул, присутствующих в исследуемом образце. [2]

Цель данной работы – сравнение алгоритмов машинного обучения для классификации по особенностям спектральных характеристик рамановского рассеяния крови пациентов с ХСН и контрольной группы.

Для получения спектров рамановского рассеяния сыворотки крови использовался зеленый лазер с центральной длиной волны 532 нм. Анализ был проведен на 270 образцах крови, 190 из которых были получены от пациентов с подтвержденным диагнозом ХСН и 80 от волонтеров-доноров для контрольной группы.

Для выделения компонентного состава был применен метод многомерного разрешения кривых (MCR-ALS) и получено 4 спектральных профиля и относительные значения концентраций для них. По полученным значениям была построена бинарная модель классификации на основе метода опорных векторов (MCR-SVM).

Также был использован другой метод компонентного анализа – метод проекции на латентные структуры (PLS), по результатам которого была построена модель бинарного классификатора на основе дискриминантного анализа (PLS-DA) и метода опорных векторов (PLS-SVM).

Для количественной оценки производительности полученных моделей были рассчитаны значения площади под характеристической кривой ROC-AUC.

На рисунке 1 представлены ROC-кривые моделей бинарной классификации пациентов с ХСН и контрольной группы.

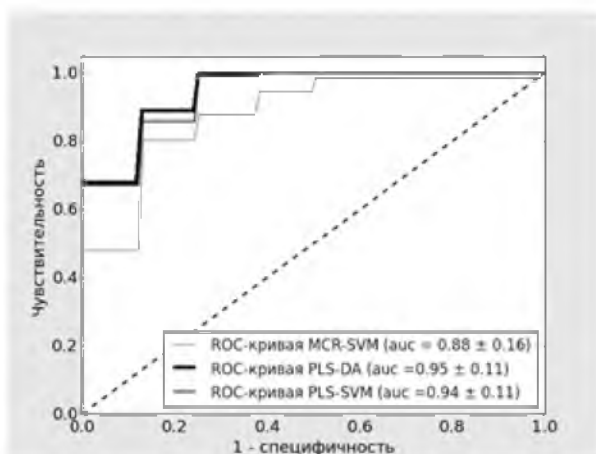


Рисунок 1. ROC-кривые моделей бинарной классификации.

При использовании для классификации относительных концентраций MCR с помощью SVM была достигнута ROC-AUC 0,88. При использовании для классификации методов PLS-DA и PLS-SVM были достигнуты значения 0,95 и 0,94 соответственно. Различия были оценены с помощью Z-теста и не являются статистически значимыми ($p\text{-value} > 0,05$) [4]. Учитывая особенности построения SVM моделей, возможно применения данных методов для выделения стадий заболевания.

Список использованных источников

1. Savarese, G. Global burden of heart failure: a comprehensive and updated review of epidemiology / G. Savarese, P. M. Becher, L. H. Lund, P. Seferovic, G. M. C. Rosano, A. J. S. Coats // *Cardiovascular research*. – 2022. – Vol. 118(17). – P. 3272-3287.
2. Parachalil, D. R. Potential of Raman spectroscopy for the analysis of plasma/serum in the liquid state: recent advances / D. R. Parachalil, J. McIntyre, H. J. Byrne // *Analytical and Bioanalytical Chemistry*. – 2020. – Vol. 412. – P. 1993-2007.
3. X.Feng, A. J Moy, H. T. M. Nguyen, Raman active components of skin cancer. // *Biomedical Optics Express*, 2017, v. 8. No. 6.
4. Hanley, J. A. A method of comparing the areas under receiver operating characteristic curves derived from the same cases / J. A. Hanley, B. J. McNeil // *Radiology*. – 1983. – Vol. 148(3). – P. 839-843.

Сорокина Елена Владимировна, студент гр. 6131-120404D, hellitio15@gmail.com
Христофорова Юлия Александровна, к.ф.-м.н., доцент каф. лазерных и биотехнических систем, khristoforova.yua@ssau.ru