

## ПРИМЕНЕНИЕ БАЙЕСОВСКИХ ГРАФОВЫХ МОДЕЛЕЙ ДЛЯ ПОИСКА ЗАШУМЛЕННЫХ ДЛИН ВОЛН В NIR-СПЕКТРОСКОПИИ.

Нетроголов Р.Ю.

(Университет ИТМО),

Научный руководитель – канд. физ.-мат. наук, Деева И.Ю.

(Университет ИТМО)

**Введение.** Многомерные данные, полученные с ИК-прибора, могут быть зашумлены. В силу очень большого количества взаимосвязей между частотами, нами была предложена и опробован анализ не фильтрами (которые ограничены своим окном), а байесовскими сетями, которые анализируют всю информацию сразу и являются вычислительно реализуемыми.

### Основная часть.

Преобразование многомерных данных при помощи PCA и последующая кластеризация уже используется в быстром качественном анализе[1]. Однако такие алгоритмы очень чувствительны к шумам – отношение сигнал/шум значительно ниже, чем у других спектроскопических методов. Как правило, задачи об удалении шумов решаются с помощью растяжения, сглаживания или коррекции базовой линии: почти все методы используют фильтры[2,3]. Однако у них есть четкое «рецептивное поле», то есть они ограничены, поэтому для обработки всего спектра сразу и общего вывода о шумах (что имеет больший физический смысл, так как, например, очень часто полосы поглощения смещаются. Исключение составляет лишь «область отпечатков пальцев») необходим метод, который позволит глобально и не слишком ресурсозатратно оценить взаимосвязи между длинами волн и убрать шумы. Более того, метод должен также быть интерпретируемым. Под эти критерии попадают байесовские сети.

С помощью байесовских сетей решается задача об очищении данных от шумов на основе ряда структурных критериев:

- 1) Весов между двумя вершинами
- 2) Размера марковского окружения

Оценка результата проводилась при помощи DBSCAN метриками кластеризации: AMI, Silhouette, гомогенность, полнота, V-мера, Calinski-Harabasz Index.

### Выводы.

1. С помощью байесовских сетей удалось повысить качество разделения данных очень существенно (с абсолютно неразделяемых до идеально разделенных)
2. Проведено сравнение предложенных способов, выявлен наиболее лучший по вышеперечисленным критериям.

### Список использованных источников:

1. Riu, J.; Gorla, G.; Chakif, D.; Boqué, R.; Giussani, B. Rapid Analysis of Milk Using Low-Cost Pocket-Size NIR Spectrometers and Multivariate Analysis. *Foods* 2020, 9, 1090. doi: 10.3390/foods9081090
2. Бёккер Ю. Мир химии. Спектроскопия. Перевод с немецкого Л.Н.Казанцевой. Москва - 2021. - 527с. // Применение ЭВМ в ИК-спектроскопии, 192-193 с.
3. Bian, X. (2022). Spectral Preprocessing Methods. In: *Chemometric Methods in Analytical Spectroscopy Technology*. Springer, Singapore. doi: 10.1007/978-981-19-1625-0\_4

Нетроголов Р.Ю. (автор)

Подпись

Деева И.Ю. (научный руководитель)

Подпись