

## НИР

### Отчет

### Постановка

Дан массив интенсивностей отраженного излучения на различных длинах волн в фиксированном диапазоне  $\{S_j\}_{j=1}^M \in \mathbf{R}^N$ , где  $M$  - кол-во объектов,  $N$  - кол-во длин волн (признаков).

Пусть  $y_1, \dots, y_M$  - соответствующие каждому объекту физические характеристики (в нашем случае концентрации). Таким образом получаем обучающую выборку:  $\{S_j, y_j\}$ .

Предполагается, что изменение интенсивностей каждого объекта  $j$  связаны с изменением его характеристик.

Задача состоит в определении характеристик вещества  $y_k$  на неизвестных объектах на основе всех предыдущих измерений  $y_j$  заданном диапазоне.

### Природа данных

Предметом изучения являются рамановские спектры белков, которые зависят от следующих факторов:

- Амидные группы - различные структуры белков характеризуются ими (меняется форма и положение пиков)
- Дисульфидные связи - выражаются в виде пика
- Ароматические аминокислоты - позиции пиков

Внешние факторы так же оказывают влияние на экспериментальные данные:

- рН эффект - при различных условиях вносит шумы в сигнал
- Температура - меняет интенсивность пиков ( $I(T_1) > I(T_2), T_1 > T_2$ )
- Реакции с раствором - приводят к сдвигам или расширению пиков

Также сильным источником шума является флуоресценция. Существует достаточно методов для устранения ее влияния, но нами будет использовано вычитание выпуклой оболочки  $convS_j$  спектра  $S_j$

### Итоги

В рамках НИР были изучены актуальные подходы по обработке сигналов, а также найдены библиотеки рамановских спектров. Также был спланирован и поставлен физический эксперимент на основе которого будут проводиться дальнейшие исследования.