По поводу статьи с профилями масс-спектров.

Нам нужно будет обсудить, что вставлять в результаты.

У меня следующие предложения:

Использовать 2 типа автоэнкодеров, denoise и ванильный автоэнкодер.

(1) denoise автоэнкодер

Для обучения denoise заранее не генерить зашумленные профили, как мы это делали для ванильного. На вход подавать исходные профили.

Denoise сам будет их шумить на входе, а на выходе использовать исходный профиль.

Для классификации тестовых профилей, нужно будет отменить первичное зашумление в энкодере. Сделать так, чтобы латентный слой рассчитывался по исходному начальному профилю без зашумления.

Если есть такая готовая модель автоэнкодера, то хорошо. Если ее нет, то можно самим имитировать denoise с помощью ванильного. При обучении на ванильный подавать зашумленный профиль, а на выход подавать исходный профиль без шума и получим denoise автоэнкодер.

Denoise автоэнкодеру можно разрешить шумить довольно сильно при обучении.

(2) ванильный автоэнкодер.

Его обучать нужно с умеренным шумом. Это нужно, чтобы мы не ушли из области значений для видов Bacillus.

В denoise автоэнкодере эта проблема не такая острая, поскольку выходной профиль берется без шума.

3) Сравнение автоэнкодеров со старым методом главных компонент будем проводить как обсуждали ранее.

Будем классифицировать несколько зашумленных и незашумленных исходных профилей.

Введение к статье я уже написал. Теперь осталось описать методы (библиотеки, которые использовались) и результаты.