

Стекинг и блендинг.

Презентация по курсу «основы машинного обучения»

Шабанов Н.А; 5030102/10101.

Основная идея концептов.

- Стекинг является подтипом так называемого ансамблевого обучения, блендинг - простейший подтип стекинга.
- Ансамблевое обучение - объединяет различные модели для решения одной задачи. Результаты этих моделей комбинируются для совершения финального предсказания.
- Ансамблевое обучение требует больших вычислительных ресурсов.
- При этом может значительно повышать точность предсказания в определенных случаях.
- Слабо подходит для простых задач, при этом может быть демонстрируема на их примере.

Стекинг.

Подтип ансамблевого обучения.

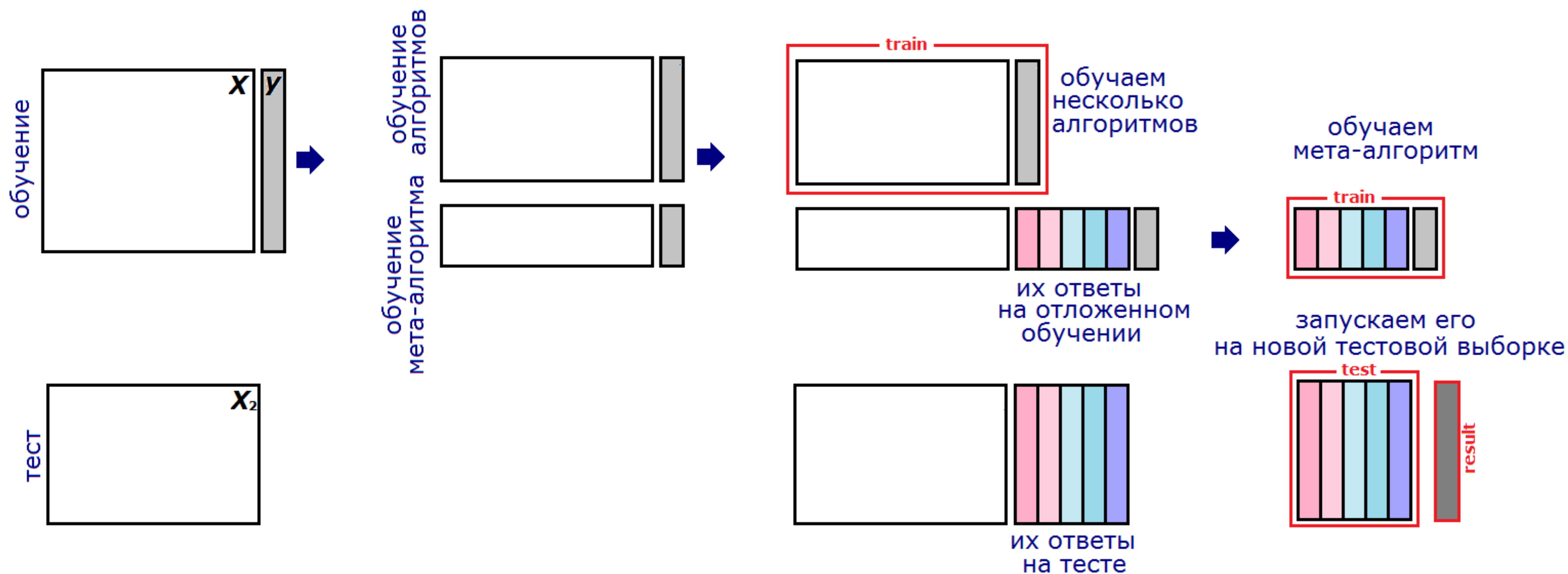
- Основная идея - для ансамбирования можно использовать один из обученных алгоритмов (метаалгоритмов) машинного обучения.
- В большинстве случаев голосование по большинству для классификационной задачи и поиск среднего для задачи регрессии превосходит по эффективности и качеству базовые алгоритмы.
- На практике используется для решения реальных бизнес-задач и соревнованиях по спортивному анализу данных.
- Был предложен Д. Волпертом в 1992 году, который с его помощью победил на Kaggle WISE 2014.
- Не всегда повышает точность предсказаний.

Блендинг.

Подтип ансамблевого обучения.

- Является простой формой стекинга.
- Основная идея - обучается несколько алгоритмов на обучающей выборке, на основе тестовой выборки обучается метапризнак, на основе которого определяется финальное предсказание.
- Главный недостаток - деление выборки, алгоритмы и метапризнак не используют весь объем данных.
- Для повышения качества нужно усреднять блендинг с разными разбиениями обучения.
- Подходит для простых задач.

Блендинг.

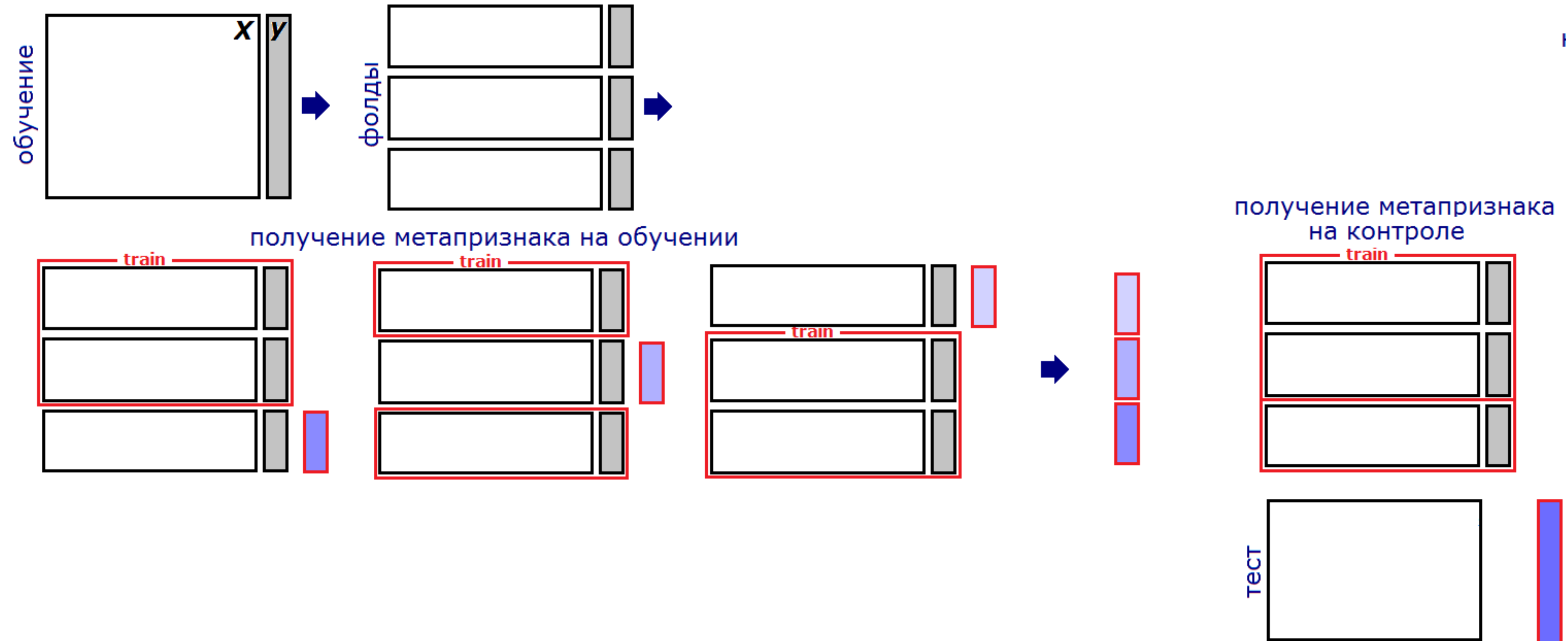


Классический стекинг.

Подтип ансамблевого обучения.

- Основная идея - выборка разбивается на части (фолды), на них обучаются базовые алгоритмы (каждый алгоритм на всех, кроме одного), на последнем фолде формируются метапризнаки, ответы базовых алгоритмов рассматриваются как значения признаков на соответствующем фолде.
- Главный недостаток - метапризнаки на обучении и на тестовых данных могут не совпадать. Это может вызывать переобучение и потерю качества предсказаний.
- Данная проблема решается добавлением нормального шума к метапризнакам.

Классический стекинг.



Используемые алгоритмы и методики.

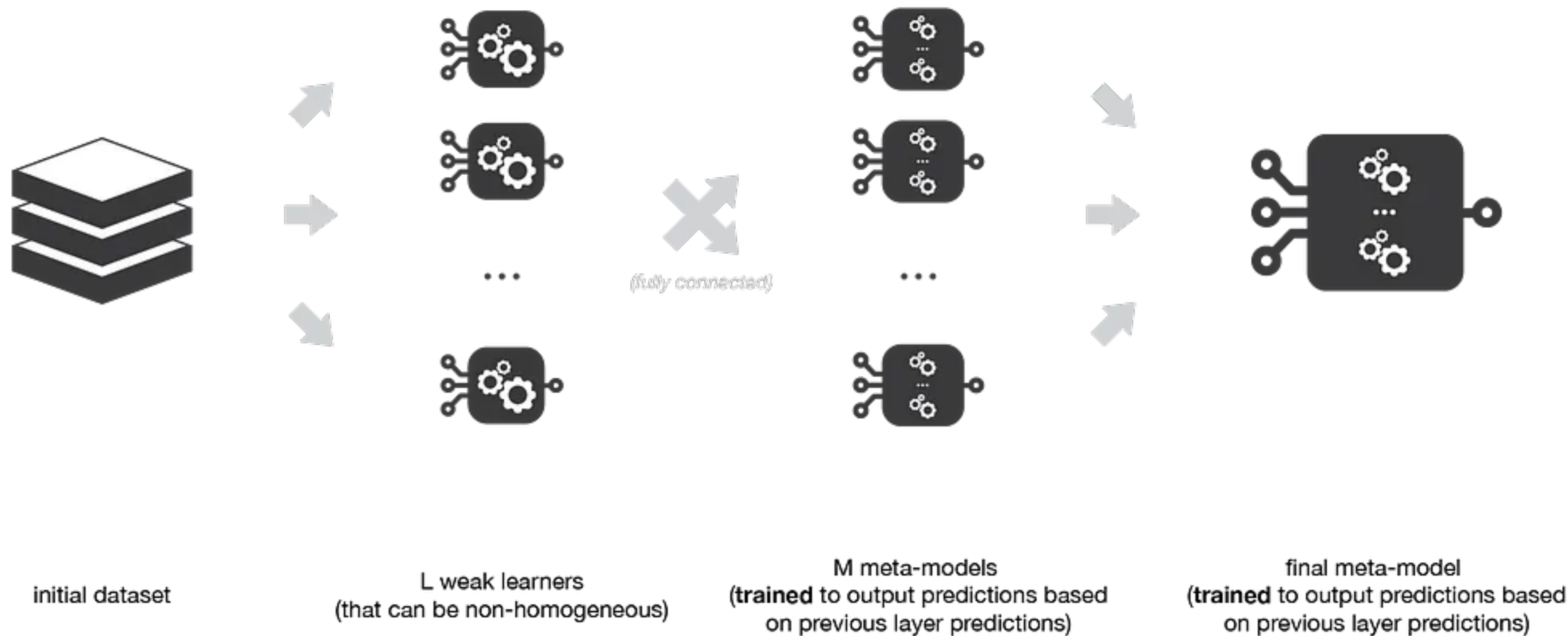
- Для стекинга используют различные алгоритмы, комбинируются разные подходы.
- Например, как базовые алгоритмы могут быть взята логистическая регрессия, метод опорных векторов, метод ближайших соседей KNN. Это - пример для классификационной задачи.
- От мета-алгоритма требуется оптимизировать заданный функционал качества. Для классификационной задачи как мета-модель можно использовать нейросеть.
- На практике для ускорения работы может использоваться параллелизация во время обучения - каждый базовый алгоритм будет обучаться на разных ядрах, параллельно.

Многоуровневый стекинг.

- Для определения понятия многоуровневого стекинга вводится понятие мета-мета-признака и мета-мета-алгоритма. Грубо говоря, к стекингу добавляется еще один 'уровень' мета-признаков перед последним мета-алгоритмом.
- Многоуровневый стекинг плохо подходит для бизнес-задач, но хорошо подходит для решения задач по спортивному анализу данных.
- Увеличение количества «уровней», как правило, повышает потребность в данных или во времени обучения.
- Для каждого уровня обучения мета-алгоритмов можно использовать отдельный алгоритм их обучения.

Многоуровневый стекинг.

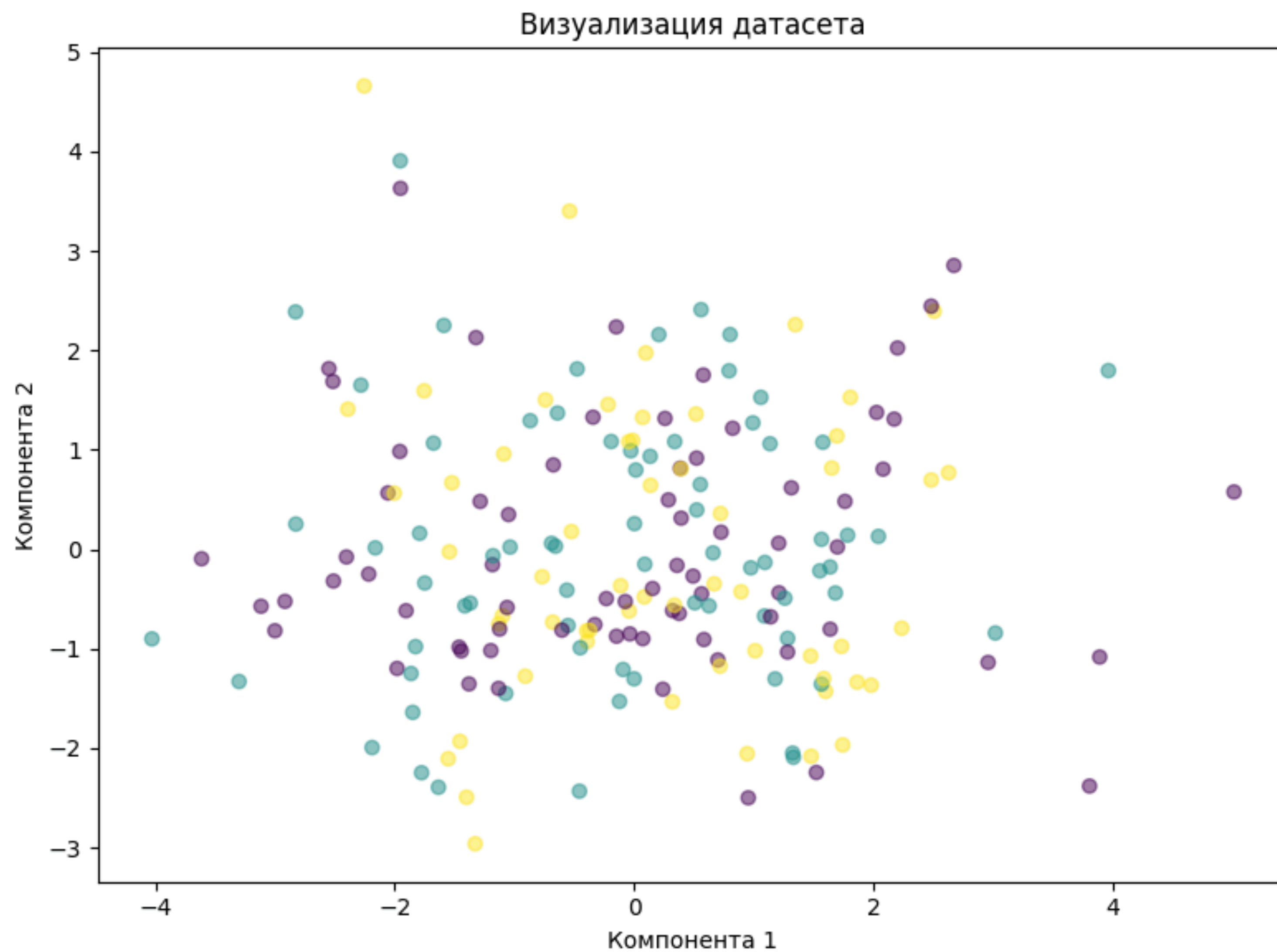
Графический пример реализации.



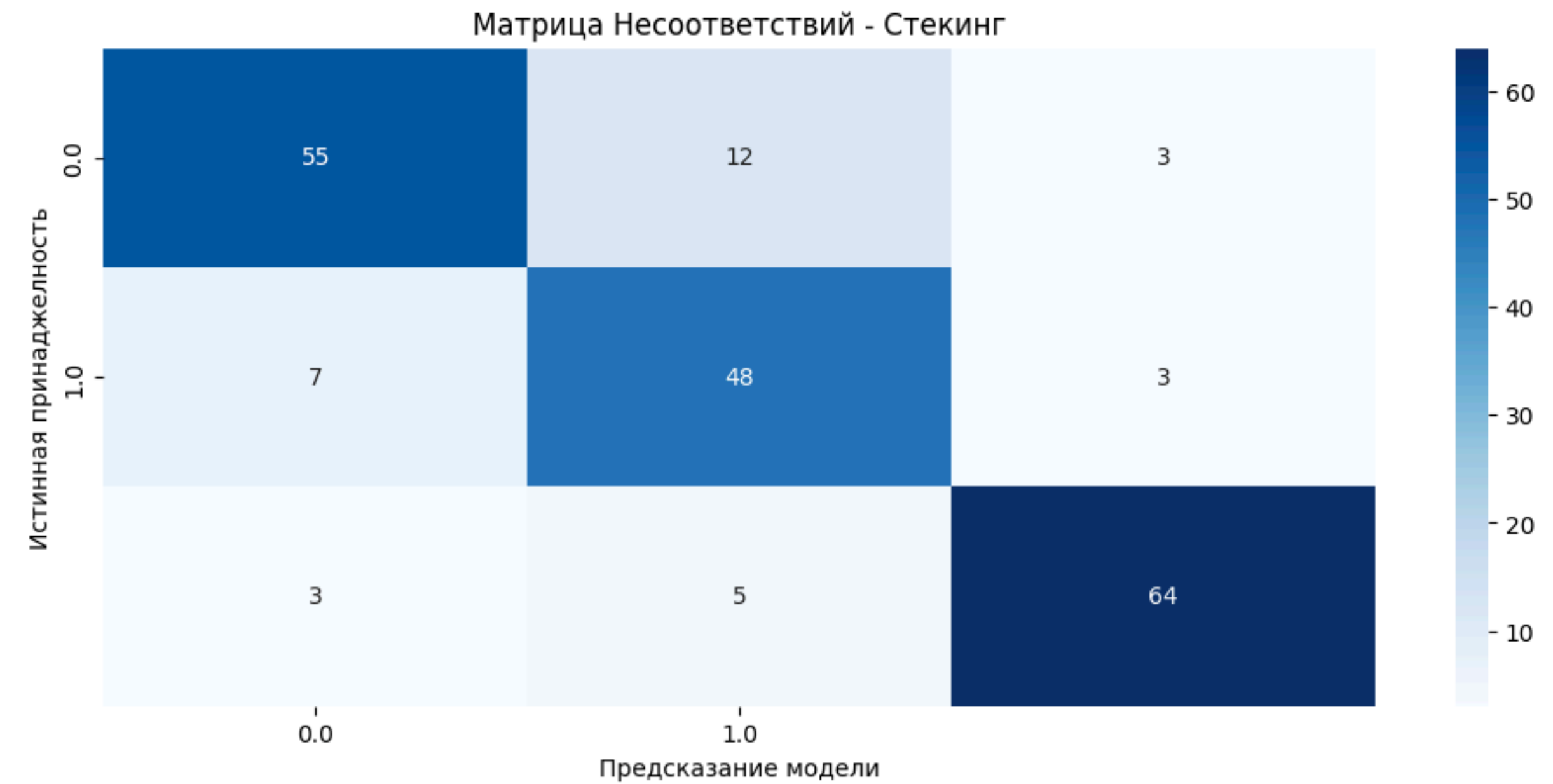
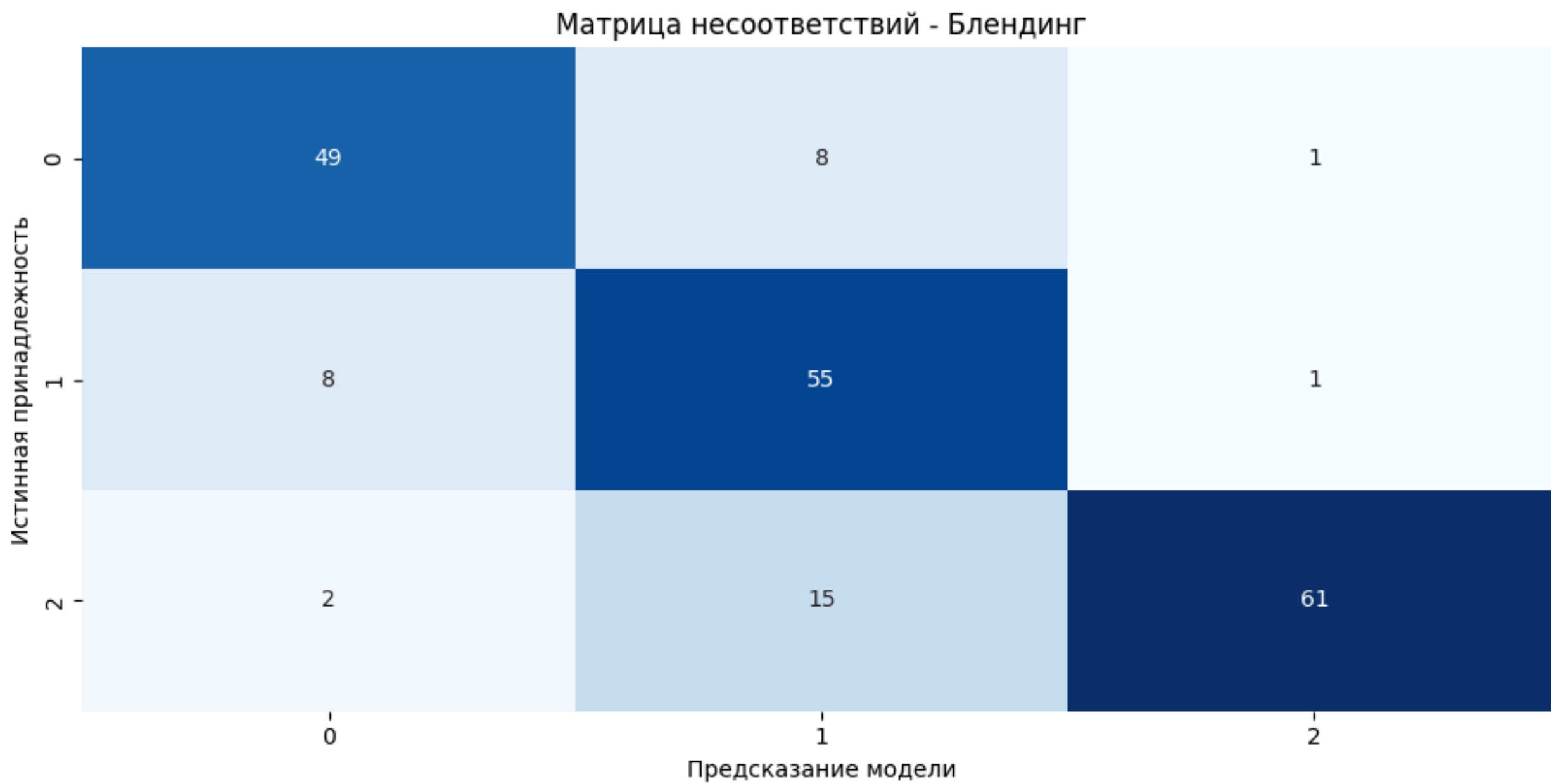
Применимость на практике.

- При решении простейших задач качество стекинга и блендинга обычно сравнимы с лучшим базовым алгоритмом.
- Стекинг требует больших выборок в большинстве случаев. Он может работать и на малых выборках, но нужно правильно подобрать базовые алгоритмы и метаалгоритм.
- В случае правильной подборки базовых алгоритмов и мета-алгоритма, можно решать сложные бизнес-задачи даже при наличии в выборке аномальных данных. Комбинация регрессионных алгоритмов (базовые алгоритмы) и деревьев решений (мета-алгоритм) могут давать намного более точные предсказания, чем давали бы эти подходы по отдельности.

Пример реализации.



Пример реализации.



Результаты.

- Рассмотрены методы для небольшой классификационной задачи на 1000 элементов, три класса.
- Стекинг точнее, примерно на 5%.
- Блендинг обучается 0.19 секунд, Стекинг 1.5238. Для маленькой задачи замедление почти в 10 раз!

Источники

- <https://alexanderdyakonov.wordpress.com/2017/03/10/стекинг-stacking-и-блендинг-blending/>
- Polikar, R. (2006). «Ensemble based systems in decision making.»
- <https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205>

Спасибо за внимание!