MLP-калькулятор

Руководство пользователя

Подробная инструкция по работе

Русский

Вольфганг Бёме

wolfgang.boehme@economic-modeling.de

www.economic-modeling.de

2 ноября 2026 г.

Документация на этом сайте разработана при поддержке ИИ Claude (Anthropic).

Содержание

1 Быстрый старт — первые шаги **2**
1.1 Обзор вкладок . . . 2
1.2 Минимальный порядок работы (2 минуты) . . . 2

2 Загрузка данных **3**
2.1 Примеры наборов данных . . . 3
2.2 Собственный CSV-файл . . . 3
2.2.1 Формат . . . 3
2.2.2 Загрузка . . . 4
2.3 Подготовка данных . . . 4

3 Настройка сети **5**
3.1 Параметры на вкладке «Конфигурация» . . . 5
3.2 Блок рекомендаций . . . 5

4 Обучение **6**
4.1 Чтение кривой Loss . . . 6

5 Бэктестинг **6**
5.1 Метрики . . . 6
5.2 Эталоны сравнения . . . 6
5.3 Экстраполяция . . . 7
5.4 Графики и таблица . . . 7

6 Прогнозирование **7**
6.1 Ручной ввод . . . 7
6.2 Пакетный прогноз . . . 7

7 Визуализация, анализ и экспорт **8**
7.1 Визуализация . . . 8
7.2 Экспорт . . . 8

8 Частые проблемы и их решение **8**

9 Полный пример работы **8**
9.1 Кросс-курсы (UZS) — хорошо прогнозируются . . . 8
9.2 Валютные данные (UZS) — распознать пределы . . . 9

10 Заключение **9**
10.1 Порядок работы кратко . . . 9
10.2 Самое важное . . . 10

1 Быстрый старт — первые шаги

1.1 Обзор вкладок

MLP-калькулятор (версия 20) — это один HTML-файл, который полностью работает в браузере. Он разделён на восемь вкладок:

Вкладка	Назначение
Конфигурация	структура сети, активации, скорость обучения, эпохи, размер батча, seed
Данные	загрузка примеров или своего CSV, нормализация, разбиение Train/Test
Обучение	запуск/остановка обучения, журнал, кривая Loss
Бэктестинг	качество прогноза на обучающих и тестовых данных, эталоны сравнения
Визуализация	граф сети, активные нейроны, распределение весов, активации
Прогноз	ручной ввод и пакетный прогноз в исходных значениях
Анализ	статистика сети и сведения о слоях
Экспорт	экспорт CSV и PNG

1.2 Минимальный порядок работы (2 минуты)

1. **Вкладка «Данные»** → нажмите пример «Кросс-курсы (UZS)». Рекомендуемые настройки применяются, сеть создаётся.
2. **Вкладка «Обучение»** → «Начать обучение» (около 15 секунд).
3. **Вкладка «Бэктестинг»** → посмотрите метрики и сравнение с простыми правилами.
4. **Вкладка «Прогноз»** → введите значения (заполнены последней строкой данных) → «Вычислить прогноз».

✓ Успех

Готово! Для собственных данных шаг 1 заменяется загрузкой CSV-файла; калькулятор подстраивает сеть под число входов и показывает рекомендуемые диапазоны на вкладке «Конфигурация».

2 Загрузка данных

2.1 Примеры наборов данных

На вкладке «Данные» доступны семь примеров. При загрузке калькулятор применяет настройки из документа «Рекомендации по параметрам обучения» и создаёт сеть.

Пример	Данные	Назначение
Валютные данные (UZS)	251 строка, USD, CNY, GBP, CHF → EUR	учебный пример: слабая связь, бэктестинг показывает пределы
Кросс-курсы (UZS)	251 торговый день, USD/UZS, EUR/USD, GBP/UZS, CNY/UZS → EUR/UZS	хорошо прогнозируется (EUR/UZS ≈ EUR/USD × USD/UZS)
XOR, AND, OR	4 строки, 2 входа	бинарная классификация, без тестовой выборки
Классификация по кругу	200 точек, 2 входа	нелинейная граница решения
Простая регрессия	50 точек, 1 вход	синусоида с шумом

При загрузке «Кросс-курсов (UZS)» примечание объясняет отличия от «Валютных данных (UZS)»; для валютного примера калькулятор указывает, что связь там почти отсутствует.

2.2 Собственный CSV-файл

2.2.1 Формат

```
Datum,USD_UZS,EUR_USD,GBP_UZS,CNY_UZS,EUR_UZS
2025-11-03,12350.00,1.16000,16294.60,1736.22,14328.04
2025-11-04,12377.15,1.16156,16444.69,1739.41,14362.95
```

Листинг 1: Пример CSV-файла (запятая)

```
Datum;USD_UZS;EUR_USD;GBP_UZS;CNY_UZS;EUR_UZS
03.11.2025;12350,00;1,16000;16294,60;1736,22;14328,04
```

Листинг 2: Те же данные в формате русского Excel (точка с запятой)

(i) Информация

- **Расположение:** файл должен храниться локально на вашем компьютере (или подключённом к нему диске); онлайн-файлы сначала скачайте. Файл обрабатывается только в браузере.
- **Строка заголовков** с именами столбцов рекомендуется; имена затем появляются во всех таблицах и полях ввода.
- **Разделитель:** запятая или точка с запятой. При точке с запятой допускается десятичная запятая.
- **Целевой столбец:** по умолчанию последний; после загрузки его можно изменить в списке «Целевой столбец (выход)». Все остальные числовые столбцы — входы (не более 10).
- **Текстовые столбцы и даты** игнорируются автоматически, строки с пропусками пропускаются — калькулятор сообщает об этом.
- **Значения:** исходные, не нормализованные.
- **Порядок:** временные ряды по возрастанию даты (сначала самые старые).

2.2.2 Загрузка

1. Вкладка «Данные» → «Выбрать файл» → выберите файл.
2. Строка состояния показывает число строк, входы, целевой столбец и игнорируемые столбцы.
3. Если число входов не совпадает с сетью, она пересоздаётся автоматически. Вкладка «Конфигурация» затем показывает рекомендацию для этих данных (см. раздел 3).

2.3 Подготовка данных

Все опции действуют **сразу** на загруженные данные — повторная загрузка не нужна. Затем обучите сеть заново.

Нормализовать данные в [0, 1]: рекомендуется при больших или разномасштабных значениях (напр. валютные курсы). Параметры (минимум, максимум) определяются *только по обучающим данным*. Ввод и вывод всегда остаются в исходных значениях.

Значение цели за предыдущий день как дополнительный вход: только для временных рядов. Значение целевого столбца из предыдущей строки добавляется как вход «*цель* (пред. день)»; первая строка отбрасывается. Тогда сеть может выучить как минимум правило «как вчера».

Разбиение Train/Test: доля обучающих данных (по умолчанию 80 %); остаток служит для бэктестинга. Два способа:

- **Хронологически** — последние строки являются тестом. Правильно для временных рядов, соответствует реальному прогнозу.
- **Равномерно** — каждая n-я строка является тестом. Подходит для данных без временного порядка (напр. классификация по кругу).

✓ **Успех**

Пример (кросс-курсы): всего 251 запись — обучение 200 записей, тест 51 запись (последние строки = тест, как при реальном прогнозе).

3 Настройка сети

3.1 Параметры на вкладке «Конфигурация»

- **Входные нейроны:** берутся из данных.
- **Число скрытых слоёв (2–5)** и для каждого слоя **нейроны (1–20)** и **функция активации** (ReLU, Sigmoid, Tanh, Linear).
- **Активация выхода:** Linear для регрессии, Sigmoid для бинарной классификации (цель только 0/1).
- **Функция потерь:** MSE (регрессия) или BCE (классификация).
- **Скорость обучения, эпохи, размер батча.**
- **Начальное значение (seed)** (по умолчанию 42): тот же seed и те же настройки дают точно тот же результат.
- **Перемешивать порядок обучающих данных в каждой эпохе** (рекомендуется): сеть не видит последними всегда одни и те же строки — важно для хронологических данных.

После изменений нажмите «Создать сеть».

3.2 Блок рекомендаций

Если данные загружены, вкладка показывает таблицу «Рекомендация для загруженного набора» со столбцами *Рекомендуется* и *Текущее*. Настройки вне рекомендации отмечены знаком △ — они допустимы, но могут ухудшить обучение. «Применить рекомендуемые настройки» восстанавливает значения и пересоздаёт сеть.

Для собственных данных действуют эмпирические правила:

- Скрытый слой 1 = 2–4 × число входов, скрытый слой 2 ≈ вдвое меньше.
- Цель непрерывная → выход Linear, потери MSE; цель только 0/1 → выход Sigmoid, потери BCE.
- Большие значения → включить нормализацию.

(i) **Информация**

Пример кросс-курсов (по умолчанию): 4 → 12 (ReLU) → 6 (ReLU) → 1 (Linear), MSE, скорость обучения 0,003, 2000 эпох, размер батча 16, seed 42 — 145 параметров.

4 Обучение

1. Вкладка «Обучение» → «Начать обучение».
2. Показываются ход выполнения, текущий и лучший Loss, время и живая кривая Loss; «Стоп» прерывает обучение.
3. После обучения бэктестинг выполняется автоматически; журнал показывает главные тестовые метрики.

4.1 Чтение кривой Loss

- **Хорошо:** равномерно падает и выходит на плато.
- **Сильно скачет:** уменьшите скорость обучения.
- **Остаётся высокой:** больше нейронов, больше эпох, проверьте нормализацию.

! Предупреждение

Loss рассчитывается на *нормализованных* обучающих данных и ничего не говорит о качестве прогноза на новых данных. Очень малый Loss может означать и переобучение. Решающей является вкладка «Бэктестинг».

5 Бэктестинг

Вкладка «Бэктестинг» применяет обученную сеть к обучающим и тестовым данным. Все значения — в исходной шкале.

5.1 Метрики

Метрика	Смысл
MAE	средняя абсолютная ошибка в единицах цели
RMSE	корень из средней квадратичной ошибки; сильнее штрафует большие ошибки
MAPE	средняя абсолютная ошибка в процентах фактического значения
R²	$1 - \sum(\text{ошибка}^2)/\sum(\text{факт} - \text{среднее})^2$; $R^2 < 0$: хуже постоянного среднего
Макс. ошибка	наибольшая абсолютная ошибка
Точность	только классификация, порог 0,5

Значимы величины в столбце *Тест (невиденные)*. Если ошибка на тесте более чем вдвое больше ошибки обучения, калькулятор предупреждает о возможном переобучении.

5.2 Эталоны сравнения

Малый MAPE сам по себе не признак качества — валютные курсы часто меняются за день менее чем на 0,2 %. Поэтому калькулятор сравнивает сеть в тестовом периоде с тремя простыми правилами:

- среднее обучающих данных (константа),
- последнее значение перед тестовым периодом (константа),

- значение предыдущей строки («как вчера»).

Вывод: «лучше всех простых правил», «примерно так же» (RMSE в пределах ±5 %) или «хуже». Для классификации эталоном служит самый частый класс обучения.

5.3 Экстраполяция

Для каждого входа калькулятор определяет обучающий диапазон (от минимума до максимума обучающих данных). Тестовые строки, где хотя бы один вход лежит вне него, отмечаются знаком △, затронутое значение выделяется цветом (▼ ниже, ▲ выше; отклонение — при наведении мыши). Таблица показывает по каждому входу число строк вне диапазона и наибольшее отклонение, абсолютно и в процентах ширины диапазона; раскрывающаяся подсказка объясняет смысл. Кроме того, средняя ошибка отмеченных строк сравнивается с ошибкой остальных.

5.4 Графики и таблица

«Факт и прогноз» показывает все строки (тестовые — на сером фоне), диаграмма рассеяния — прогноз против факта с диагональю. Таблица «Тестовые строки подробно» содержит входы, прогноз, факт и ошибку. «Бэктестинг в CSV» экспортирует все строки (точка с запятой, десятичная запятая).

6 Прогнозирование

6.1 Ручной ввод

- Поля подписаны именами столбцов, заполнены последней строкой данных и показывают обучающий диапазон.
- Ввод **всегда в исходных значениях** (десятичная запятая или точка). Нормализация выполняется внутри по параметрам обучающих данных.
- «Вычислить прогноз» показывает результат в исходных значениях, при классификации — также класс.
- Если значение лежит вне обучающего диапазона, предупреждение называет вход и отклонение, напр. при GBP_UZS = 17 000 в примере кросс-курсов: «GBP_UZS: выше максимума на 240,70 (28,5 % ширины диапазона)».

6.2 Пакетный прогноз

- **Загрузить CSV-файл** с теми же входными столбцами, что при обучении. Столбцы сопоставляются по заголовку; если файл содержит и целевой столбец, рассчитываются ошибки и метрики.
- **Пакет из тестовых данных:** «Скачать как пакетный файл (CSV)» создаёт файл с тестовыми строками (напр. `CrossRates_Batch_TestPeriod.csv`); «Сразу выполнить пакетный прогноз» рассчитывает их немедленно.
- Результат: метрики (MAE, RMSE, MAPE, R^2, средняя ошибка/смещение, макс. ошибка), график факт/прогноз, обзор экстраполяции и таблица по строкам; экспорт — «Результаты в CSV».

7 Визуализация, анализ и экспорт

7.1 Визуализация

- **Архитектура сети:** все нейроны и веса (красный — положительные, синий — отрицательные, толщина по модулю).
- **Активные нейроны (статус активации):** для текущего ввода на вкладке «Прогноз» (иначе последней строки данных) показывает только нейроны со статусом *нормальный* ($0{,}2 \le |a| < 0{,}8$); *слабые* ($|a| < 0{,}2$) — бледно, *сильные* ($|a| \ge 0{,}8$) включаются флажками. Неактивные нейроны — только серая точка. Толщина линии соответствует вкладу вес × активация.
- **Распределение весов** (гистограмма) и **значения активаций** (таблица со статусом).

7.2 Экспорт

- **Данные модели:** веса, архитектура, значения активаций (CSV).
- **Обучение и прогнозы:** история обучения, пакетные прогнозы, бэктестинг (CSV; бэктестинг и пакет — с точкой с запятой и десятичной запятой для русского Excel).
- **Визуализации:** граф сети, кривая Loss, распределение весов, бэктестинг факт/прогноз и диаграмма рассеяния (PNG).

8 Частые проблемы и их решение

Проблема	Решение
«Корректные строки данных не найдены»	Проверьте разделитель; не менее двух числовых столбцов; при точке с запятой — десятичная запятая, при запятой — точка.
Сеть и данные не согласуются (число входов)	Пересоздайте сеть на вкладке «Конфигурация» — при загрузке это происходит автоматически.
Loss не снижается	Увеличьте скорость обучения (напр. 0,003 → 0,005), больше эпох, включите нормализацию.
Loss сильно скачет / NaN	Уменьшите скорость обучения, сбросьте сеть, обучите заново.
Ошибка теста $> 2 \times$ ошибки обучения	Переобучение: сеть поменьше, меньше эпох, больше данных.
Сеть не лучше «как вчера»	Входы почти не объясняют цель; для временных рядов включите значение предыдущего дня как вход.
Много строк с экстраполяцией	Проверьте столбец ошибок отмеченных строк; удлините обучающий период.
Результат не воспроизводится	Используйте тот же seed и настройки; пересоздайте сеть перед обучением.

9 Полный пример работы

9.1 Кросс-курсы (UZS) — хорошо прогнозируются

1. Вкладка «Данные» → «Кросс-курсы (UZS)». Столбец даты игнорируется; обучение 200, тест 51 строка (хронологически).

2. Вкладка «Конфигурация»: рекомендация применена (4 → 12 → 6 → 1, все строки ✓).
3. Вкладка «Обучение» → «Начать обучение» (2000 эпох, около 15 секунд).
4. Вкладка «Бэктестинг»:

✓ Успех

Метрика	Обучение	Тест
MAE	4,68	5,34
RMSE	5,89	6,44
MAPE	0,033 %	0,037 %
R^2	0,9988	0,9959

Сравнение в тестовом периоде: правило «как вчера» MAPE 0,354 %, R^2 0,621. Вывод: «Сеть лучше всех простых правил (RMSE на 90 % ниже лучшего правила)».

5. Экстраполяция: в 27 из 51 тестовой строки один вход вне обучающего диапазона (USD/UZS до 14,8 % ниже, GBP/UZS до 24,7 % выше). Ошибка этих строк не больше (MAE 5,03 против 5,68) — GBP/UZS не влияет на цель.
6. Вкладка «Прогноз»: заполнено строкой 251 (USD_UZS 12 326,81; EUR_USD 1,1715; GBP_UZS 16 582,05; CNY_UZS 1 733,29) → прогноз EUR_UZS = 14 439,66 при фактическом значении 14 437,19.
7. «Сразу выполнить пакетный прогноз» повторяет оценку теста как пакет; «Результаты в CSV» экспортирует её.
8. Вкладка «Визуализация» → «Активные нейроны»: для этого ввода 10 нейронов нормальные, 1 сильный, 1 слабый и 11 неактивных.

9.2 Валютные данные (UZS) — распознать пределы

1. Вкладка «Данные» → «Валютные данные (UZS)» (4 → 16 → 8 → 1, 3000 эпох) и обучить.
2. Бэктестинг: MAPE на обучении 0,255 %, на тесте 0,557 %, R^2 на тесте −4,24. Калькулятор сообщает о возможном переобучении и о том, что сеть хуже простого правила «как вчера» (MAPE 0,117 %).
3. Включить опцию «Использовать значение цели за предыдущий день как дополнительный вход» и обучить заново: MAPE на тесте 0,120 %, R^2 0,638 — «примерно так же, как правило «как вчера»».

! Предупреждение

Вывод: MAPE ниже 1 % выглядит хорошо, но для валютных курсов не является признаком качества. В этом наборе дневные изменения EUR и входных курсов практически не связаны; лучше, чем «курс как вчера», получить нельзя. Только сравнение с простыми правилами в бэктестинге делает это видимым.

10 Заключение

10.1 Порядок работы кратко

1. Данные → 2. Проверка конфигурации → 3. Обучение → 4. Бэктестинг → 5. Прогноз → 6. Экспорт

10.2 Самое важное

- Всегда оставляйте разбиение Train/Test включённым; для временных рядов — хронологически.
- Включайте нормализацию при больших значениях — ввод и вывод остаются в исходных значениях.
- Оценивайте всегда по **тестовым метрикам** и **сравнению с простыми правилами**, а не по Loss.
- Учитывайте предупреждения об экстраполяции: вне обучающего диапазона прогноз менее обоснован.
- Тот же seed = тот же результат; при сравнении настроек фиксируйте seed.

Успехов в работе с нейронными сетями!