

## Инструкция к использованию библиотеки brain-interface

### в целях прогнозирования паттерна ЭЭГ

#### v 0.04b

Данная библиотека предназначена для работы с ЭЭГ данными, получаемыми от нейрогарнитуры Brain expert с целью их анализа посредством машинного обучения. Выполняет следующие процессы:

- Прогнозирование (предикция) значений амплитуд по отведениям
- Классификация ЭЭГ паттерна

Данная инструкция описывает процесс ПРОГНОЗИРОВАНИЯ

*\* справочная информация:*

Под прогнозированием в данном руководстве понимается: вычисление ожидаемого значения амплитуд по одному или нескольким отведениям отсроченными от последних полученных реальных данных ЭЭГ. Глубина прогноза (интервал дальности предикции) зависит от программных установок вычислений (см ниже раздел обучение модели), а также от временного шага записи исходных данных: так при дискретной частоте регистрации ЭЭГ записанной с частотой 1 Hz (дискретная частота регистрации установленная на нейрогарнитуру или вашем энцефалографе), каждая последующая запись данных ЭЭГ, будет соответствовать 1 с временного шага, поэтому прогноз выполненный на один шаг будет отсрочен также на 1 с, а при частоте 100 Hz, соответственно один шаг будет 0,01 с, этот показатель следует учитывать при интерпретации выбранных данных и расчете выбираемой глубины прогноза.

Прогнозирование возможно по следующим направлениям:

- вычисление ожидаемого однократного значения амплитуд в период сразу после регистрации
- вычисление ожидаемого однократного значения амплитуд в отсроченный период после регистрации на заданный интервал
- вычисление ожидаемого тренда амплитуд в отсроченный период после регистрации на заданный интервал

Точность прогноза зависит от выбранной глубины прогнозирования, исходных данных и параметров модели.

*\* использование библиотеки:*

Для прогнозирования в библиотеке используется технология анализа временных рядов с использованием рекуррентных нейронных сетей глубокого машинного обучения keras. Библиотека в целях удобства использования реализована с поддержкой графического интерфейса, позволяющего вносить и получать переменные как в цифровом виде, так и с помощью диаграмм.

Практическое применение возможно для интерпретации регистрируемой ЭЭГ в интересах разработки систем поддержки принятия решений, экспериментальном поиске закономерностей, создании прикладных нейросетевых моделей. Библиотека предоставляется по принципу AS IS (как есть), лицензия GNU GPL, вся ответственность за использование библиотеки и возможный ущерб лежит на пользователе. Рекомендуется проверять обновления на сайте. ООО ЭМРАТЕХ предоставляет также обученные модели на коммерческой основе для различных практических задач (<https://emra.tech/software>), использование которых возможно с помощью данной библиотеки при дистанционном подключении к серверам компании.

## 1. Формат исходных данных

Исходные данные должны соответствовать нижеследующему формату:

при использовании архивных записей - файл csv с разделителями значений ";"

значения колонок данных (см рис.1)

t - временной ряд, отражающий значения времени в секундах, записанные при регистрации ЭЭГ отведения F1,F2,F3,F4,T3,T4,C3,C4,P3,P4,O1,O2 - отражающий значения амплитуд, полученных при регистрации

light - логическое значение (0,1) активности в данный временной момент модуля световой стимуляции. 0- False, 1 -True

sound - логическое значение (0,1) активности в данный временной момент модуля звуковой стимуляции. 0- False, 1 -True

trigger - логическое значение (0,1) наличия или отсутствия фактора воздействия, связанного по мнению исследователя с ожидаемыми изменениями ЭЭГ. 0- False, 1 -True.

Рис.1

t	F1	F2	F3	F4	T3	T4	C3	C4	P3	P4	O1	O2	light	sound	trigger
0.3	0.9384	6.978	-1.188	-1.423	0.5779	0.6957	1.778	3.097	-1.7	-1.423	-0.02443	1.378	0	0	0
0.302	0.4319	6.517	-1.702	-1.581	0.616	0.3752	1.573	2.916	-1.241	-1.063	0.5178	1.617	0	0	0
0.304	-0.1661	5.876	-2.284	-1.886	0.5398	0.1329	1.378	2.535	-0.9819	-0.8403	0.938	1.737	0	0	0
0.306	-0.764	5.157	-3.023	-2.343	0.235	0.01563	1.28	1.998	-0.8842	-0.7035	1.216	1.778	0	0	0
0.308	-1.221	4.494	-3.687	-2.782	-0.04446	0.05471	1.212	1.397	-0.9038	-0.6605	1.319	1.799	0	0	0
0.31	-1.503	4	-4.202	-3.201	-0.3048	0.1329	0.977	0.8793	-0.9233	-0.6409	1.339	1.737	0	0	0
0.312	-1.727	3.517	-4.626	-3.601	-0.5017	0.1563	0.6155	0.4983	-0.9819	-0.68	1.28	1.597	0	0	0
0.314	-1.927	3.112	-4.925	-3.944	-0.6605	0.05471	0.1759	0.2589	-1.124	-0.7426	1.158	1.44	0	0	0
0.316	-2.026	2.933	-5.066	-4.122	-0.7621	-0.1407	-0.1856	0.2198	-1.221	-0.7426	1.08	1.279	0	0	0
0.318	-2.001	3.056	-5.041	-4.103	-0.743	-0.3674	-0.2833	0.3762	-1.221	-0.68	1.119	1.197	0	0	0
0.32	-1.786	3.416	-4.842	-3.906	-0.6605	-0.6019	-0.1075	0.6986	-1.104	-0.5628	1.299	1.197	0	0	0
0.322	-1.487	3.91	-4.626	-3.645	-0.6414	-0.8832	0.1563	1.08	-0.9233	-0.5237	1.539	1.299	0	0	0
0.324	-1.105	4.393	-4.36	-3.34	-0.6033	-1.204	0.4104	1.436	-0.7621	-0.5432	1.798	1.419	0	0	0
0.326	-0.6478	5	-3.945	-2.94	-0.4255	-1.446	0.6546	1.778	-0.5813	-0.6409	2.037	1.518	0	0	0
0.328	0.01661	5.775	-3.322	-2.362	-0.0254	-1.641	0.9575	2.179	-0.381	-0.8012	2.316	1.58	0	0	0
0.33	0.7973	6.73	-2.541	-1.683	0.4954	-1.727	1.456	2.618	-0.1417	-0.9809	2.638	1.679	0	0	0
0.332	1.694	7.73	-1.661	-0.9463	0.9971	-1.602	2.14	3.058	0.1172	-1.223	2.936	1.799	0	0	0
0.334	2.616	8.73	-0.764	-0.1842	1.48	-1.36	2.775	3.439	0.298	-1.462	3.136	1.898	0	0	0
0.336	3.438	9.573	0.07474	0.4954	1.835	-1.047	3.195	3.698	0.3371	-1.723	3.239	1.98	0	0	0
0.338	4.078	10.13	0.7557	0.959	2.096	-0.7035	3.439	3.815	0.2785	-1.942	3.278	2.018	0	0	0
0.34	4.493	10.44	1.196	1.156	2.178	-0.3674	3.42	3.796	0.2198	-2.142	3.239	2.038	0	0	0
0.342	4.659	10.42	1.379	1.08	2.02	-0.1016	3.136	3.62	0.01954	-2.321	3.039	2.018	0	0	0

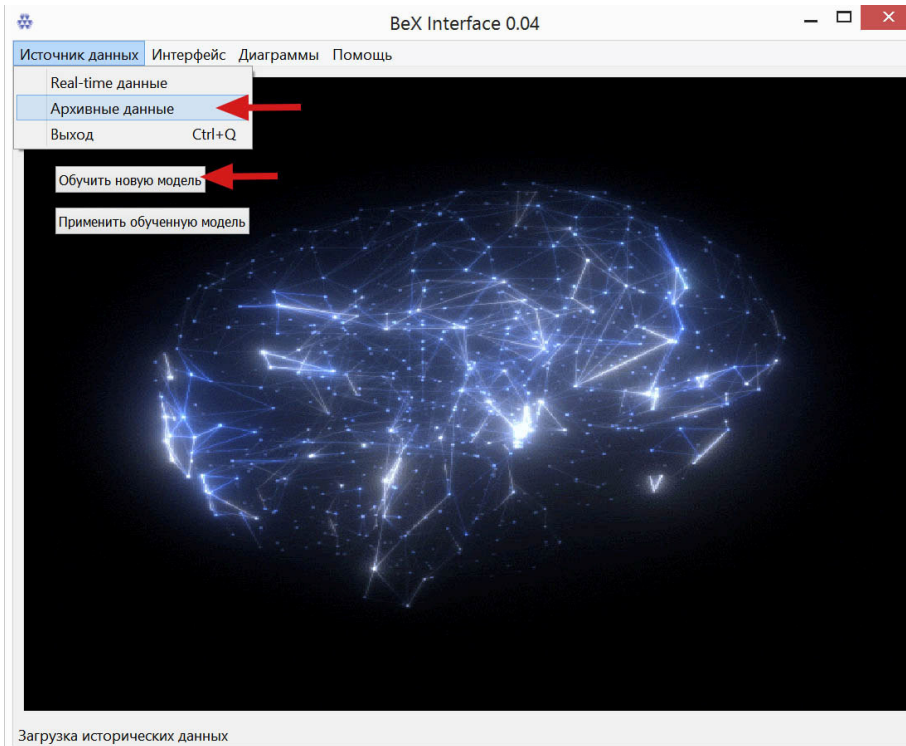
при использовании текущих данных (при подключенной нейрогarnитуре) - переменные формируются приложением brainexpert.exe аналогично архивным непосредственно отправляясь в цикл вычислений. Для использования с нейрогarnитурой необходимо также скачать в каталог с данным приложением библиотеку brain\_socket ([https://emra.tech/media/bex\\_socket.zip](https://emra.tech/media/bex_socket.zip))

## 2. Обучение модели

### 2.1. Подготовка dataset

Для загрузки архивных данных выберите либо кнопку в основном окне "Обучить новую модель", или в панели верхнего меню "Источник данных/ Архивные данные" (рис.2)

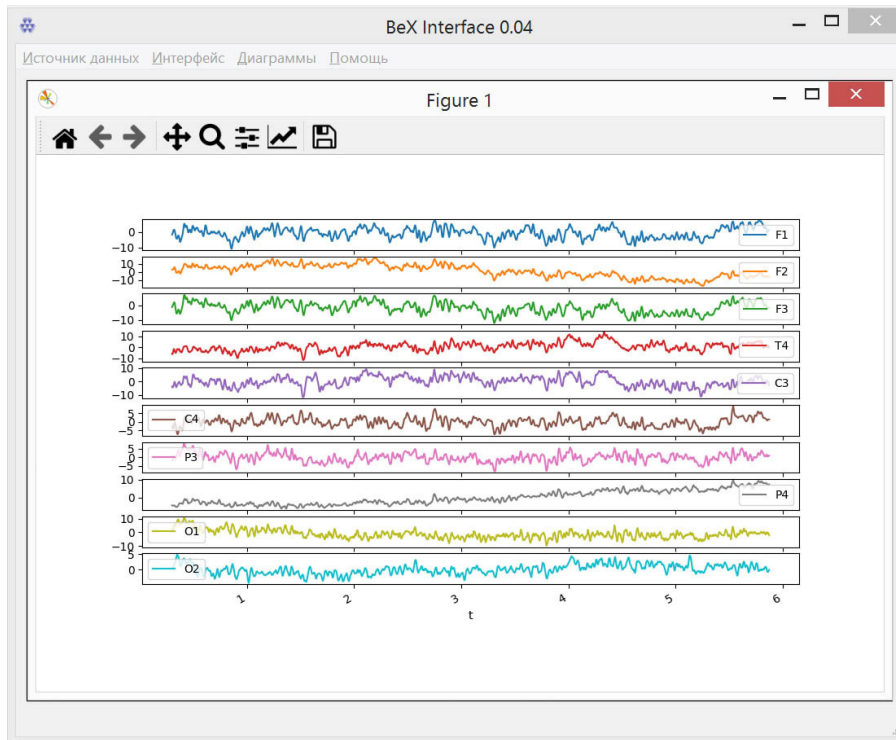
Рис.2



Для использования данных сразу с нейрогарнитуры ВЕХ (текущие данные) выберите в панели верхнего меню "Источник данных/Real-time данные".

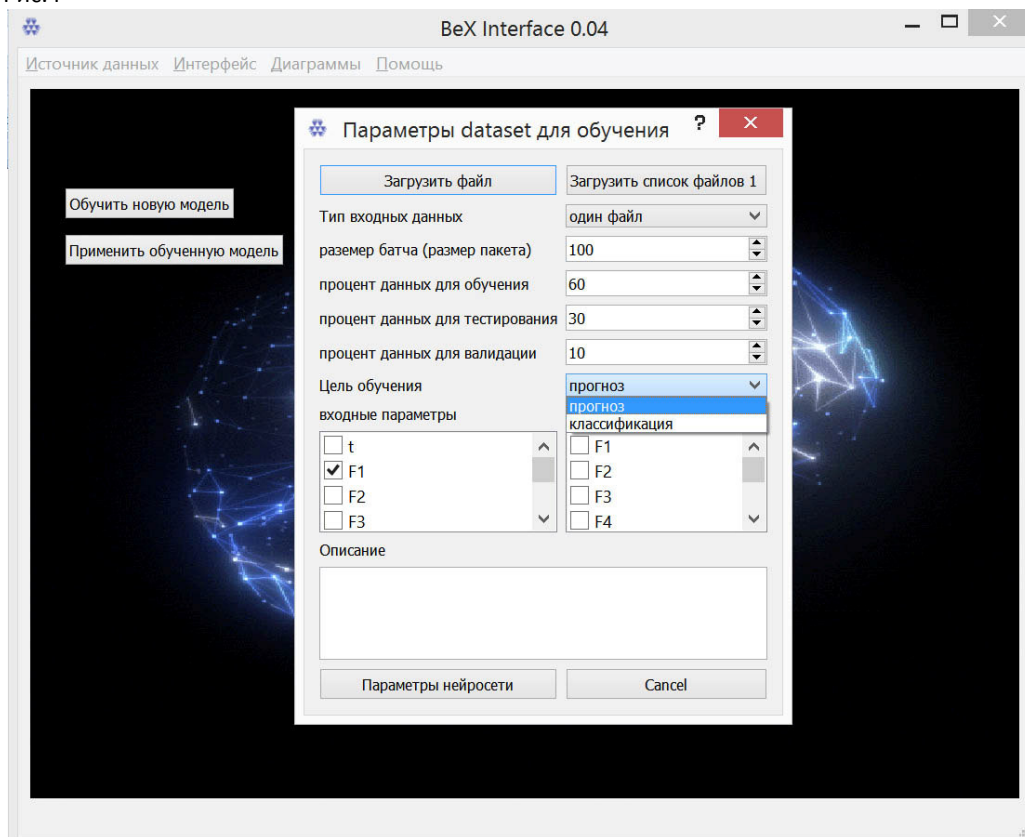
Для наглядности данные можно просмотреть на графике в виде диаграммы (рис.2 ) нажав на кнопку "Диаграммы" и установить флажок ("Показывать диаграммы").

Рис.3



После выбора источника данных откроется окно настроек dataset . В открывшемся окне выберите опцию "Цель обучения/ прогноз" (рис.4)

Рис.4



Выполните загрузку файла нажав "Загрузить файл". Для примера возьмите любой файл из архива примеров теста [https://emra.tech/media/test\\_data\\_RH.zip](https://emra.tech/media/test_data_RH.zip) (данные записей ЭЭГ с триггером: мысленное сжатие правой руки в кулак) либо из [https://emra.tech/media/test\\_data\\_PR.zip](https://emra.tech/media/test_data_PR.zip), где хранятся записи без триггерного воздействия.

\* поддержка текущих данных для обучения нейросетевой модели не применяется

После загрузки файла в поле "файл входных данных" появится наименование файла и пути к нему, а также информация о количестве строк данных для ориентировки (рис.5).

**Тип входных файлов** оставьте "один файл", т.к. текущая версия библиотеки поддерживает обучение модели прогнозирования только по одному файлу данных.

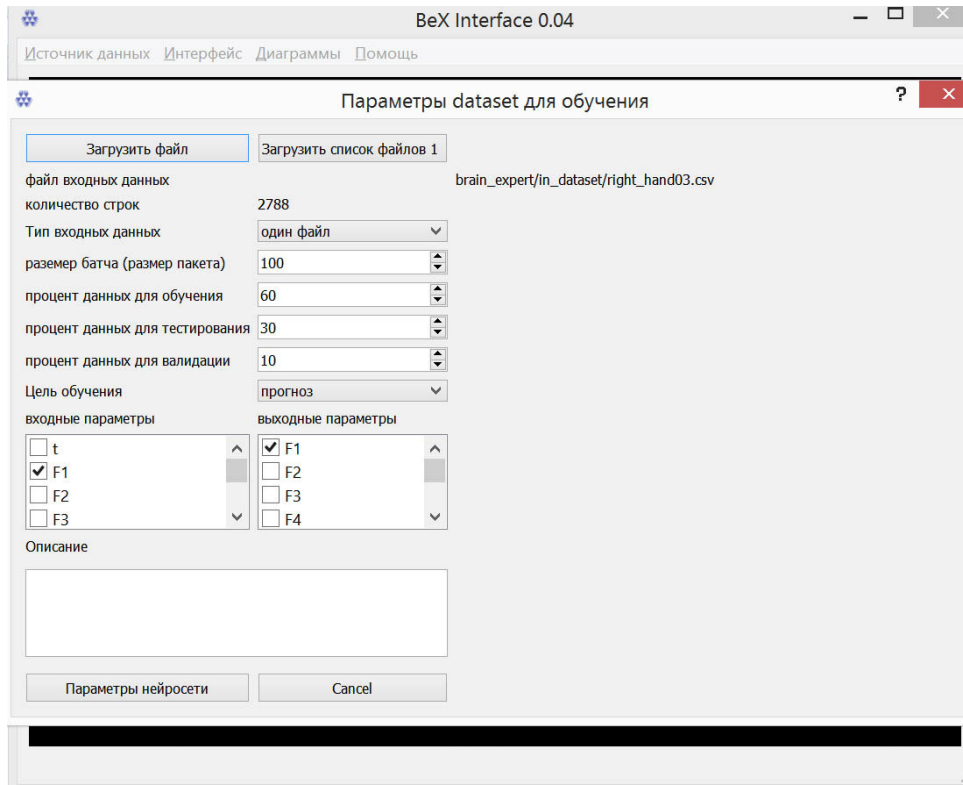
**Размер батча(размер пакета)** будет делить исходный файл данных на сегменты, в рамках которого осуществляется обучение модели. Выбор размера пакета имеет важное значение в плане как обучения так и применимости обученной модели, т.к. захватывает паттерн ЭЭГ за определенный этим размером временной интервал. Например, если вас интересует модель для оценки участка записи за 0.4 секунды, в ваших исходных данных частота дискретизации (см выше справочную информацию) равна 500 Hz, значит запись идет с интервалом 0.002 с или 50 значений за 0.1 с ( $50 \cdot 0.002$ ), соответственно для целевого паттерна записи в 0.4 секунды нужен батч 200 ( $0.4 / 0.1 \cdot 50$ ). Важно что при применении уже обученной модели, следует в анализируемых данных указать аналогичный размер пакета для интерпретируемых данных.

В примере ниже установите батч 200.

### Доли исходных данных для обучения, тестирования и валидации

Установите процент записи для обучения, тестирования и валидации данных по своему усмотрению. Для примера установите значения 60, 30,10 соответственно (рис.5).

Рис.5



Также следует указать по каким колонкам исходных данных планируется осуществлять обучение и прогноз, для этого выберите в разделе входные параметры интересующие колонки для источника данных. А в выходных параметрах установите целевые данные для прогнозирования.

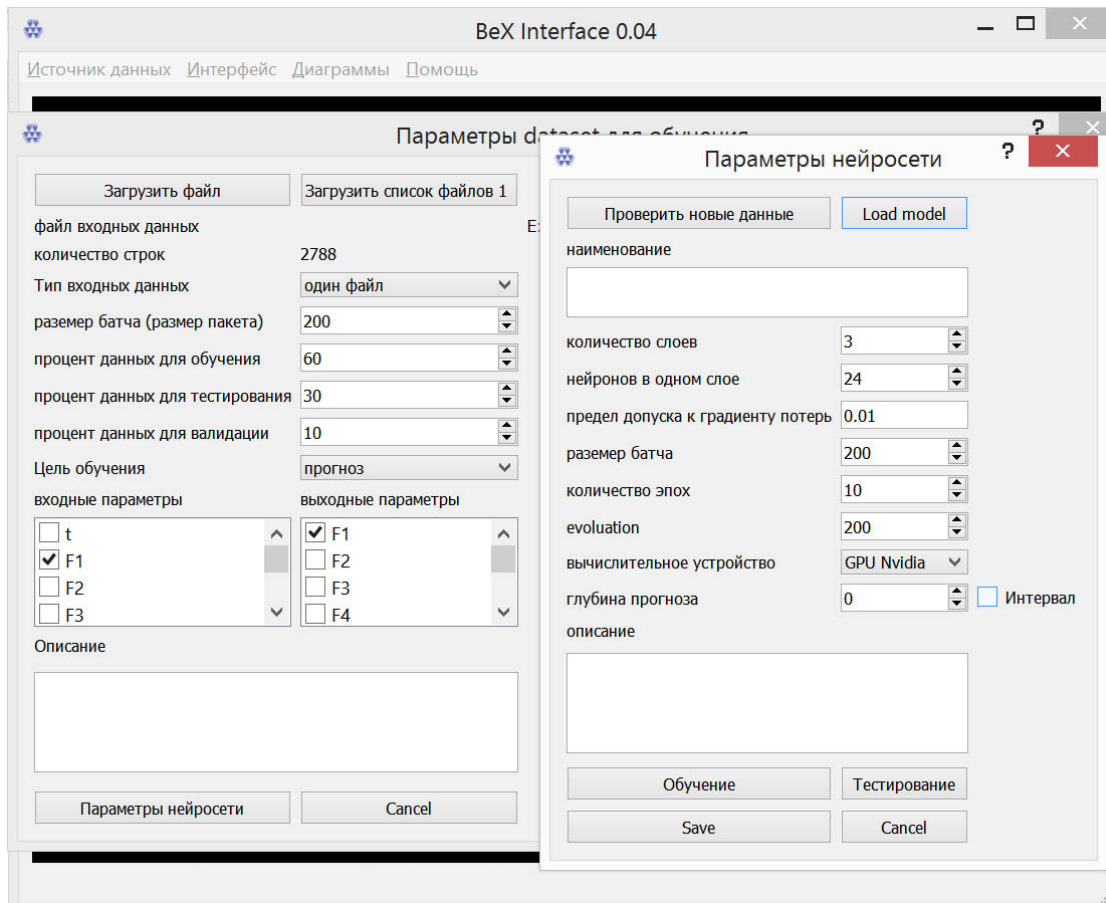
\*Важно! При обучении выходной искомым параметр обязательно должен быть отмечен и входных данных. При использовании, этого уже не требуется.

На этом подготовка датасет завершена. Нажмите "Параметры нейросети" и дождитесь появления окна настроек параметров нейросети (рис.6)

## 2.2. Обучение модели

после активации окна "Параметры нейросети" и дождитесь появления окна настроек параметров нейросети (рис.6)

Рис.6



**проверить данные** -запускает процесс анализа новых исходных данных по уже обученной модели (см раздел 3)

**загрузить модель** - загружает конфигурацию обученной модели из указанного места (см раздел 3) при необходимости дообучения. . Вы можете разрабатывать свою модель или приобрести готовые модели обученные под различные задачи в ООО ЭМРАТЕХ (<https://emra.tech/>).

**наименование** - введите наименование, если планируете сохранять модель, и при этом не используете PostgreSQL и серверную обработку данных, поля "описание" не будут сохраняться, поэтому рекомендуем формировать наименование, помогающее интерпретировать сохраненную модель. В примерах обученных моделей predict\_1\_f2\_eeg\_rh05\_200\_500 означает что модель обучена по прогнозированию одного интервального шага по отведению F2 на исходных данных файла rh05 с размером пакета 200 при дискретной частоте 500

**количество слоев и нейронов** - имеет значение для оптимизации сети при прогнозировании амплитуд по двум и более отведениям, если вы изучаете только одно отведение, то рекомендуется использовать 1-3 слоя и не более 30 нейронов, этого вполне достаточно для прогноза высокой точности до 1-2 с.

**предел допуска** - устанавливает целевой градиент потерь при обратном распространении ошибки. После его достижения модель будет считаться обученной. Чем меньше градиент там точнее расчет, но дольше обучение. Рекомендуется установить 0.01

**размер батча** - переносится из датасета, и можно менять в интересах оптимизации модели без перенастройки dataset (см описание в разделе 2.1)

**количество эпох** - количество проходов по датасету при обучении, при большем значении будет занимать больше время

**evolution** - дополнительный ограничитель внутри датасета, если ваши исходные данные значительны, а локальные вычислительные ресурсы ограничены, то рекомендуется ввести ограничение. В идеальных условиях этот показатель должен быть равен количеству строк (см. рис.5) исходного датасета. Влияет на точность прогноза. Для загруженного примера используйте значение от 200 до 1000.

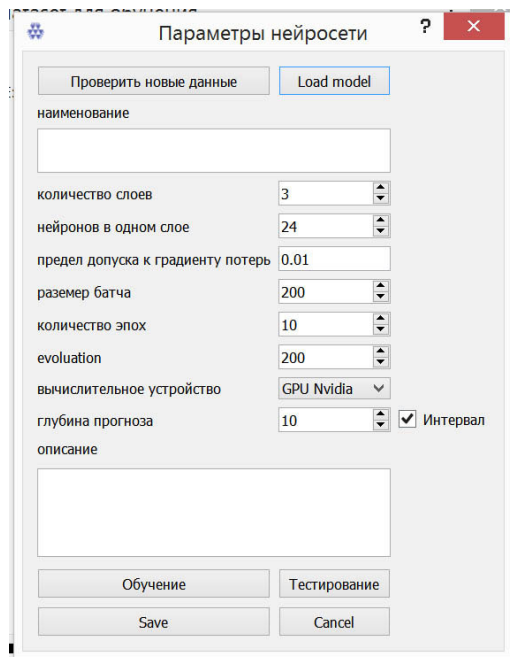
**вычислительное устройство** - определяется по умолчанию как наиболее подходящее на вашем устройстве можете указать принудительно.

**глубина прогноза** - выставляется в шагах временного ряда исходных данных,

если равно "0" то означает однократное расчетное значение, ближайший шаг по временному интервалу, больше "0" устанавливает количество шагов для интервала. Глубина прогноза (интервал дальности предикции) зависит от дискретной частоты регистрации ЭЭГ. Например при записи с частотой 1 Hz (дискретная частота регистрации установленная на нейрогарнитуре или вашем энцефалографе), каждая последующая запись данных ЭЭГ, будет соответствовать 1 с временного шага, поэтому прогноз выполненный на один шаг будет отсрочен также на 1 с, а при частоте 100 Hz, соответственно один шаг будет 0,01 с. В приведенном примере шаг равен 0.002 с.

для получения однократного значения на указанную глубину прогноза оставьте поле "Интервал" не отмеченным, для расчета серии интервальных значений поставьте флажок в поле "Интервал" рис.7

Рис.7



**описание** - используется только при работе PostgreSQL или серверной обработке данных

**После подготовки параметров нажмите "Обучение"**

В соответствии с установленными параметрами начнется тренировка модели, прогресс можно увидеть в окне терминала вашего интерпретатора или IDE (рис. 8)

Рис.8

```
Console [E:\python\python.exe]
bex_interface.py [E:\python\python.exe]
165/200 [=====>.....] - ETA: 1s - loss: 0.2613
167/200 [=====>.....] - ETA: 0s - loss: 0.2600
169/200 [=====>.....] - ETA: 0s - loss: 0.2589
171/200 [=====>.....] - ETA: 0s - loss: 0.2578
173/200 [=====>.....] - ETA: 0s - loss: 0.2566
175/200 [=====>....] - ETA: 0s - loss: 0.2554
177/200 [=====>....] - ETA: 0s - loss: 0.2545
179/200 [=====>....] - ETA: 0s - loss: 0.2533
181/200 [=====>... ] - ETA: 0s - loss: 0.2522
183/200 [=====>... ] - ETA: 0s - loss: 0.2512
185/200 [=====>... ] - ETA: 0s - loss: 0.2500
187/200 [=====>.. ] - ETA: 0s - loss: 0.2489
189/200 [=====>.. ] - ETA: 0s - loss: 0.2476
191/200 [=====>.. ] - ETA: 0s - loss: 0.2465
193/200 [=====>.. ] - ETA: 0s - loss: 0.2456
195/200 [=====>.. ] - ETA: 0s - loss: 0.2446
197/200 [=====>.. ] - ETA: 0s - loss: 0.2434
199/200 [=====>.. ] - ETA: 0s - loss: 0.2423
```

где последовательно будет отражаться процесс обучения и полученный коэффициент потерь. Длительность обучения может быть довольно значительной в зависимости от вышеприведенных параметров. Если вы не использовали ограничитель `evolution`, то на обычных компьютерах может занять от нескольких часов до неск. дней. Для загруженного примера используйте значение 200.

Итоговое значение по результатам обучения и тестирования прогноза будет высвечиваться в окне терминала (рис.9) в виде значений прогнозного (`predict`), реального (`true future`) и процентного уровня отклонения (отклонение). Чем меньше отклонение тем точнее предсказывает модель. на рисунке 9 точность в процентах достигнута 100 - 0.65 (99.35%).

Рис.9

```
Console [E:\python\python.exe]
bex_interface.py [E:\python\python.exe]

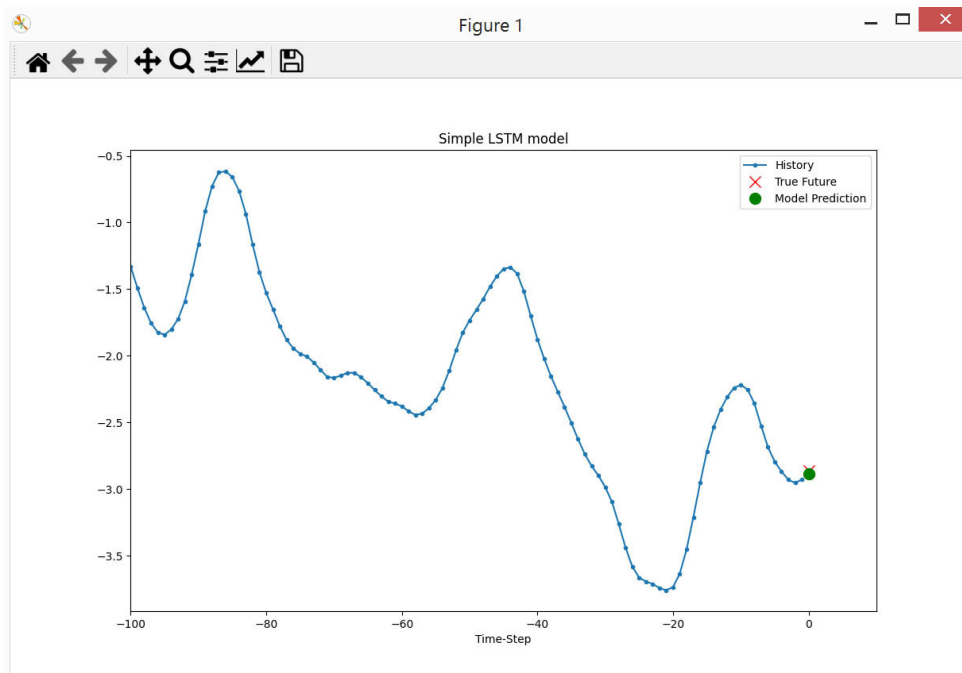
1/4 [=====>.....] - ETA: 0s
4/4 [=====>.....] - 0s 8ms/step

1/4 [=====>.....] - ETA: 0s
4/4 [=====>.....] - 0s 8ms/step
true future: [-2.86460225]
predict [-2.8831432]
отклонение % [0.64724713]
```



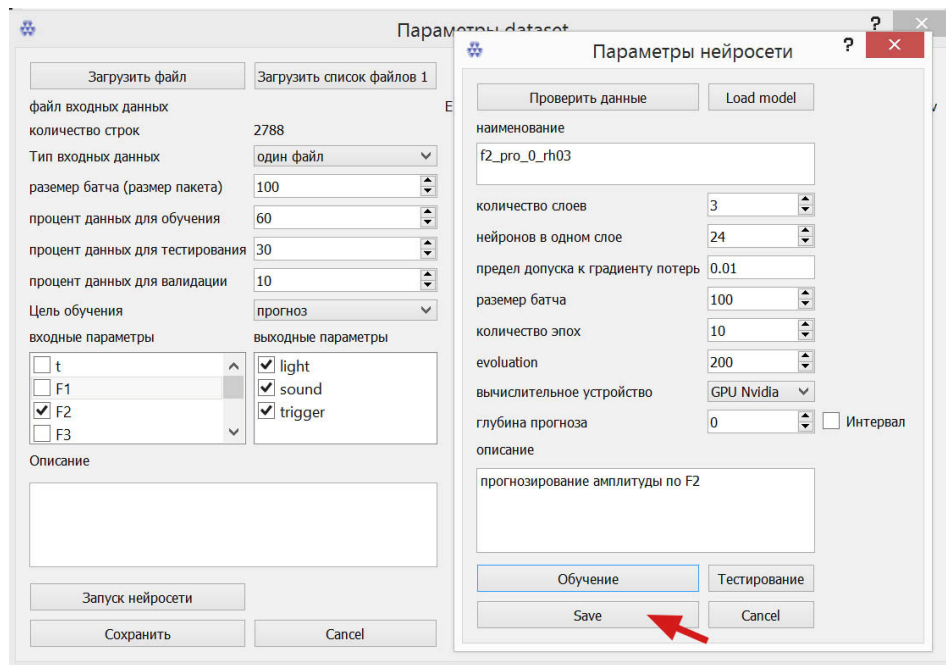
А также в виде графика для наглядности (рис.10)

Рис. 10



Если вас устраивает результат обучения модели, то нажмите кнопку "Save" в окне параметров нейросети (рис. 11)

Рис.11

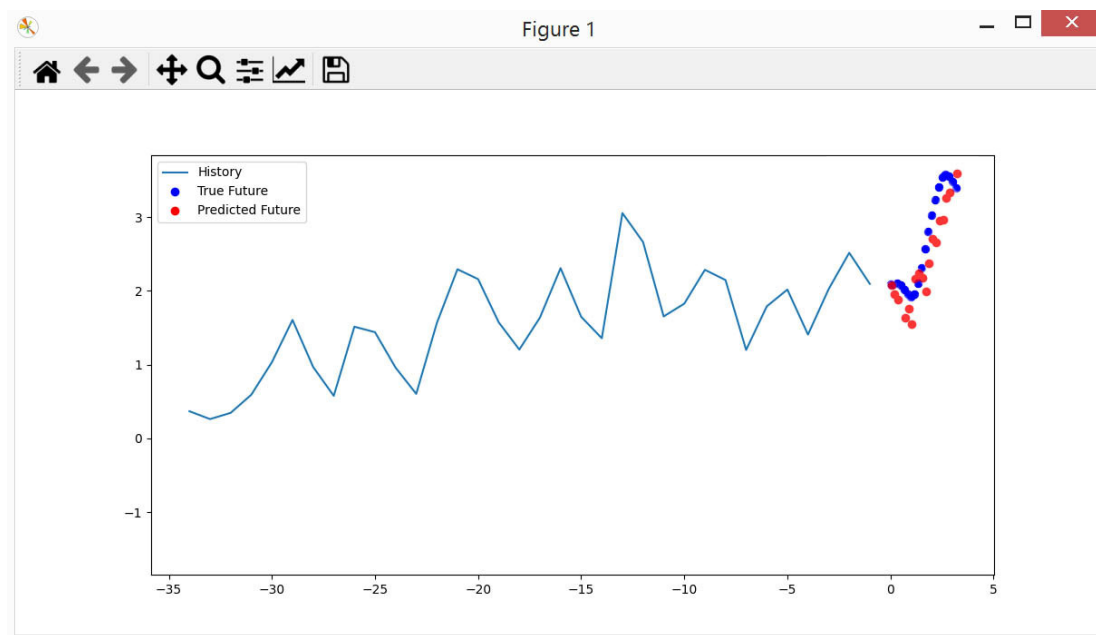


Данные всех настроек датасета, модели, и полученных расчетов сохраняются в каталоге `n_model` под указанным вами наименованием модели.

При использовании нескольких отведений для прогнозирования, а также отсроченное прогнозирование потребует значительных вычислительных ресурсов для обучения модели. Поэтому вы можете переложить вычислительную задачу на сервер `emra.tech` (). Для этого необходимо получить файл для доступа по `ssh` протоколу к вычислительному серверу компании (обратитесь в ООО ЭМРАТЕХ ). Разместите полученный файл в каталоге библиотеки и установите переключатель вычислительное устройство в режим "сервер EMRA", затем сохраните модель

Обучение в интервальном диапазоне также отображается в виде последовательности полученных значений и в виде графика (рис.12). \*Важно! При расчете интервальных значений для достижения хороших результатов рекомендуется указывать не менее 3 слоев и не менее 20 нейронов на слой.

Рис.12

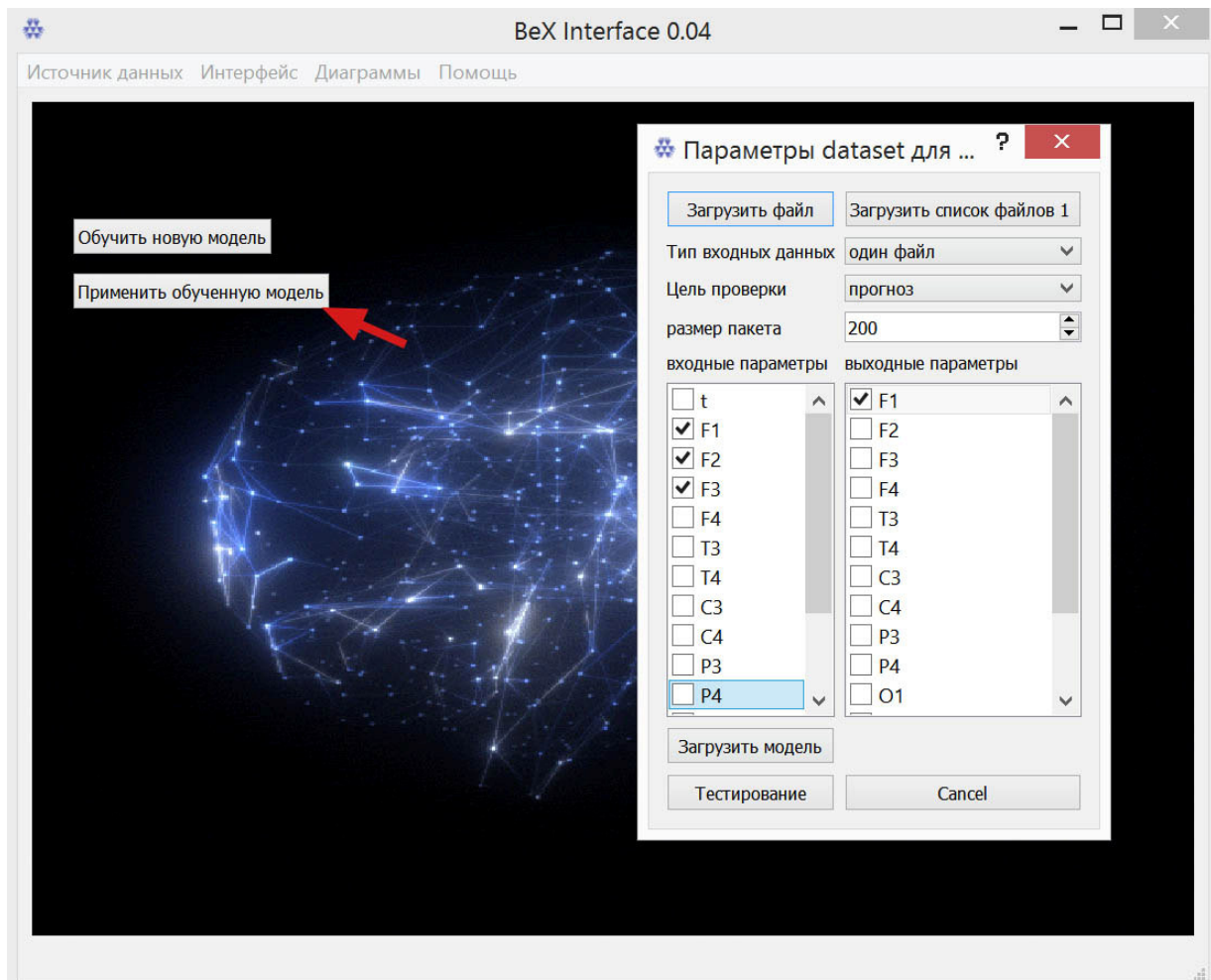


### 3. Применение обученной модели

Сохраненную ранее обученную модель вы можете применить к новым данным для прогноза:

- если вы используете данные с нейрогарнитуры выберите в панели верхнего меню "Источник данных/Real-time данные" (рис.2) и нажмите "Применить обученную модель" (рис. 13).
- если вы используете данные из записи выберите в основном окне "Применить обученную модель" (рис. 13), в открывшемся окне загрузите файл исходных данных

Рис.13



**Загрузить файл** - открывает файл для загрузки архивной записи. Если вы получаете данные непосредственно с нейрогарнитуры, то не загружайте данные из файла. При загрузке архивных данных появятся краткие сведения, включающие наименование файла и количества строк данных для ориентировки (рис. 14).

**Загрузить модель** - выберите наименование папки с моделью, которую вы обучили на предыдущем этапе, или распакуйте в каталог `n_model` библиотеки модели из архива готовых примеров ([https://emra.tech/media/test\\_cnn.zip](https://emra.tech/media/test_cnn.zip)). Если модель в указанном каталоге присутствует, после сообщения "модель загружена" появится наименование и путь к модели для ориентировки (рис. 14).

**Тип входных данных** - для прогнозирования оставьте один файл

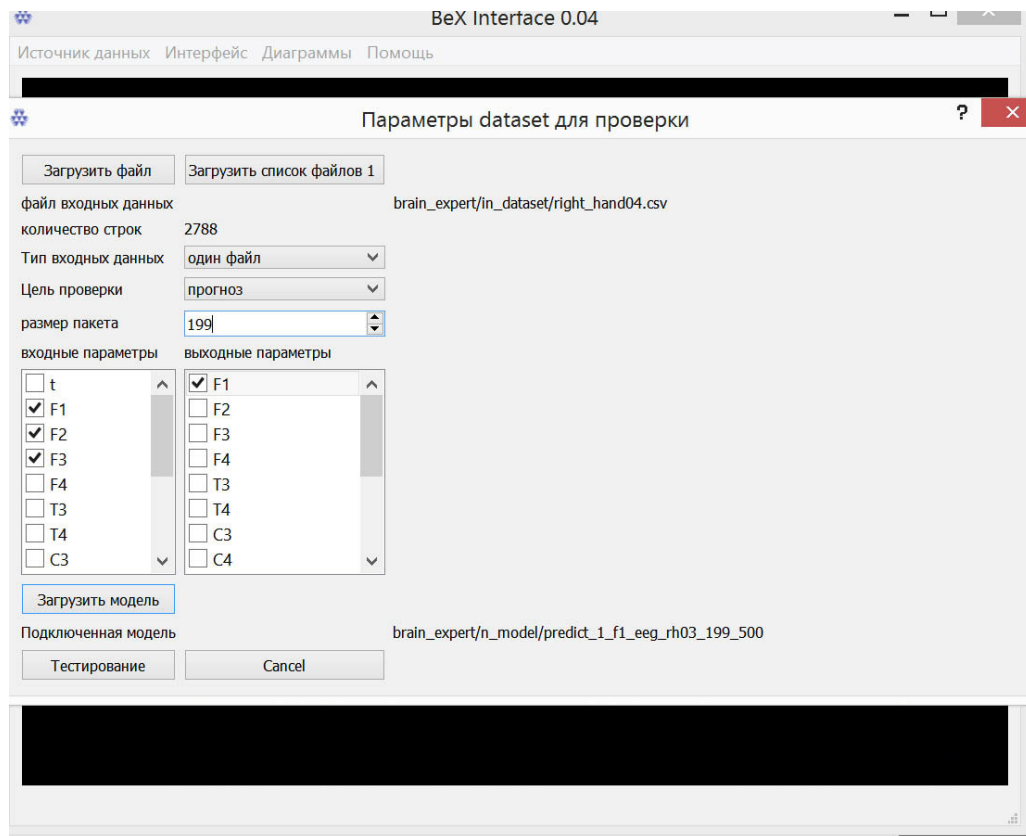
**Цель проверки** - выберите прогноз

**Размер пакета** - рекомендуется устанавливать размер пакета аналогичный размеру в обученной модели или меньше (см описание в разделе 2.1). При этом будут задействованы последние записи исходных данных эквивалентные размеру пакета, или текущие регистрируемые данные от нейрогарнитуры с указанным размером пакета. При использовании из примера возьмите значение из наименования готовой модели предпоследнюю группу цифр (см рекомендацию по наименованию из раздела 2.2)

**Входные параметры** - используйте аналогичные тем, под которые обучена модель

**Выходные параметры** - используйте аналогичные тем, под которые обучена модель

Рис. 14



Нажмите "Тестирование" и дождитесь получения результата, который также будет выведен в виде цифрового значения в окне терминала и в виде графика уже только с прогнозными значениями.

