Методические материалы по практическим работам

А.Е. Сулавко

Защищенное исполнение доверенного интеллекта

# Практическое занятие на тему: ОЦЕНКА ЭНТРОПИИ НПБК, ОБУЧЕННОГО ПО ГОСТ Р 52633.5. МОДЕЛИРОВАНИЕ АТАКИ ИЗВЛЕЧЕНИЯ ЗНАНИЙ

**Цель работы:** получение навыков обучения нейросетевого преобразователя биометрия-код (согласно ГОСТ Р 52633.5) и моделирования атаки извлечения знаний из него с помощью программного пакета AIC desktop; изучение процесса оценки энтропии обученного НПБК.

## ОПИСАНИЕ ПРОГРАММНОГО ПАКЕТА AIC DESKTOP

### Назначение и сферы применения AIC desktop

***AIC desktop* -** программный комплекс для проведения научно-исследовательских работ по машинному обучению (графический интерфейс AIC desktop представлен на рис.1.). Данный программный комплекс позволяет производить анализ данных с применением признанных методов математической статистики для выявления закономерностей. В пакете реализовано множество известных и новых методов машинного обучения (нейронные сети, Байесовские сети, иммунные сети и др.), разработанных научной группой AIC.

Визуальный редактор позволяет создавать сложные архитектуры многослойных искусственных нейронных сетей (сверточных, полносвязных, гибридных, рекуррентных). Сети можно объединять в ансамбли и комплексировать с другими (в том числе не сетевыми) моделями. Можно решать задачи классификации, регрессии и извлечения признаков.

Вероятности ошибок 1-го и 2-го рода, а также точность распознавания можно оценить разными способами, с учетом множества параметров, что позволяет получать результаты с высокой достоверностью. Поддерживается как прямой численный эксперимент, так и тестирование по ГОСТ Р 52633.3.

*AIC desktop* поддерживает разработку нечетких экстракторов, нейросетевых преобразователей биометрия-код (НПБК), обучаемых по ГОСТ Р 52633.5, а также гибридных НПБК с возможностью оценки стабильности выходных разрядов с помощью нескольких метрик.

Информацию об образах можно загружать в виде векторов признаков (xml, txt) и в виде сырых (неразмеченных, частично размеченных) данных (bmp, wav, edf, csv и др.), относящихся к любой предметной области (звук, изображения, термограммы, эхограммы, ЭКГ, ЭЭГ, биометрические и биомедицинские образы и др.). После загрузки данные преобразуются в унифицированный формат.

По окончанию серии вычислительных экспериментов может быть сгенерирован научный отчет в автоматическом режиме, содержащий графики, таблицы, описания использованных методов с отсылкой на соответствующие публикации.

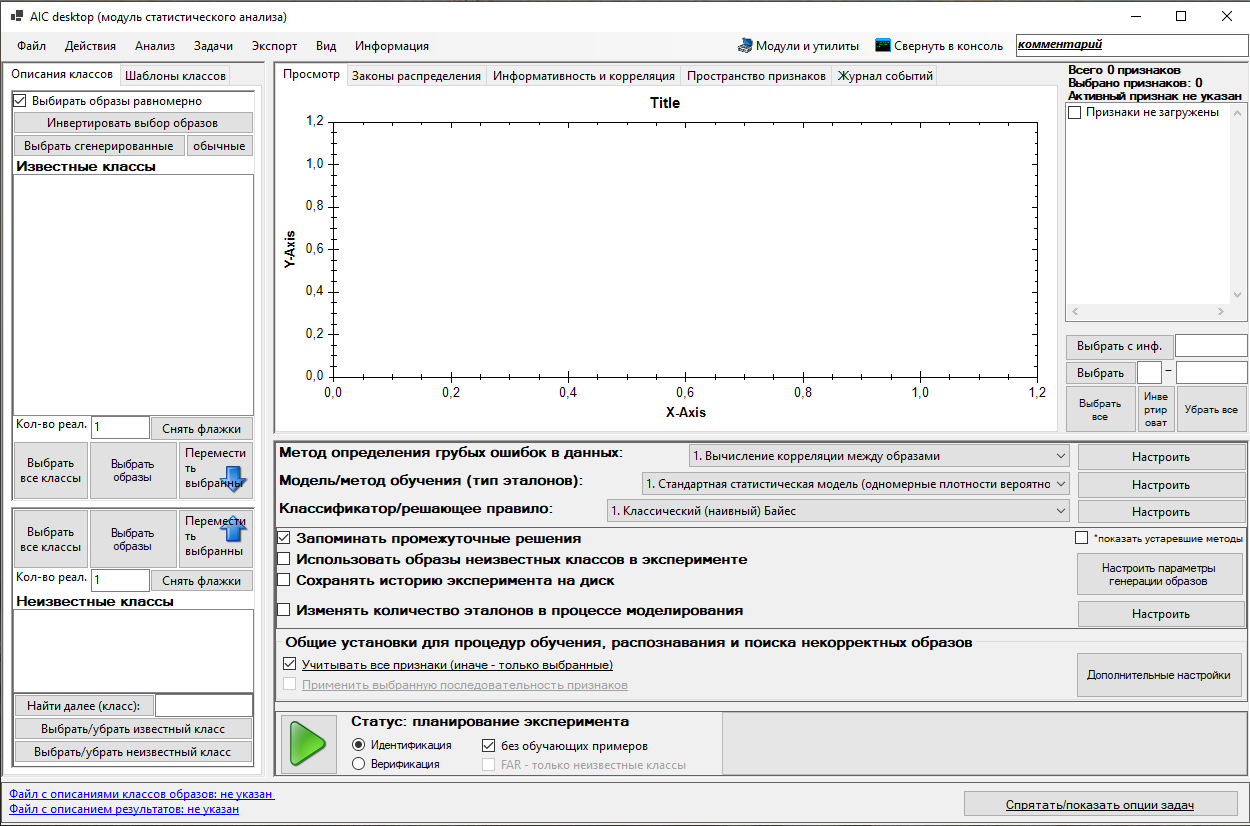


Рис.1. Графический интерфейс AIC desktop

### Установка AIC desktop

AIC desktop можно получить, перейдя на главную страницу [aiconstructor.ru](http://aiconstructor.ru/).

Так как AIC desktop является бесплатной, но не свободно распространяемой программой, а использование программы допускается только в рамках научных исследований и в целях ознакомления, то для получения программного продукта необходимо заполнить форму, позволяющую связаться с разработчиками программы. Указав некоторые данные, а также сферу научных интересов и исследований, для которых необходим AIC desktop, можно запросить программу для дальнейшего использования.

В случае проведения практических работ, программный продукт предоставляется преподавателем.

### Руководство по установке и запуску программы AIC desktop

На данный момент в AIC desktop реализована поддержка операционных систем Windows 7/8/10. Рекомендуемый объем оперативной памяти от 8 Гб и более (в зависимости от выполняемых задач и объемов анализируемых данных).

Для корректной работы AIC требуется следующее программное обеспечение:

1. Python 3.6. Требуется также установить 3 дополнительных библиотеки в следующем порядке:

1.1. Keras (version 2.2.5);

1.2. Tensorfow (version 1.6.0);

1.3. Protobuf (version 3.6.0);

2. NET Core 3.0 (и выше).

Для запуска AIC откройте исполняемый файл **nnview.exe** в папке с исходными файлами программы.

## ВЫПОЛНЕНИЕ РАБОТЫ

### Установка набора данных клавиатурного почерка

Для начала выполнения практической работы, помимо программного пакета AIC desktop, необходимо скачать на рабочий ПК набор данных образцов рукописного почерка. Датасет расположен по [ссылке](http://aiconstructor.ru/page14247028.html) и представлен в виде обезличенных (деперсонализированных) рукописных образов, воспроизведенных 130 подписантами возрастом 18-50 лет. Каждый испытуемый выбирал слово или набор символов по своему усмотрению, которые воспроизводил на графическом планшете фирмы Wacom с частотой опроса 200 точек в секунду и 1024 уровнями давления пера на планшет. Для данной практической работы необходим датасет, представляющий собой обработанные данные (не сырые) в xml-формате (в сокращенном формате), который содержит описания этих же образов, но представленных в виде векторов из 556 признаков.

### Загрузка и подготовка данных в AIC desktop

Для того чтобы загрузить данные рукописных образцов в AIC desktop, необходимо, находясь в главном окне приложения, перейти на вкладку **«Файл»** → **«Загрузить данные об образах»** → **«Сокращенный формат xml»** и выбрать данные, раннее скачанные на ПК. При удачной загрузке датасета, в левой части программы отобразятся классы образов и их реализации так, как показано на рисунке 2. Образы «Чужих» с помощью кнопки **«Переместить выбранные»** переносим вниз.

По ГОСТу НПБК обучается на данных «Свой» (не менее 11 примеров) и данных «Чужой» (не менее 64 примеров). Так как примеров в данном датасете всего 65, то каждый по отношению к каждому является «Чужим», в итоге при обучении каждый НПБК обучается на 64-ех «Чужих». Чтобы определиться с классами для дальнейшей работы, переводим образы «Чужие» в **«Неизвестные образы»**, так как они содержат в себе образы различных дикторов.

Дополнительно проверяем наличие галочки в поле **«Выбирать реализации равномерно».**

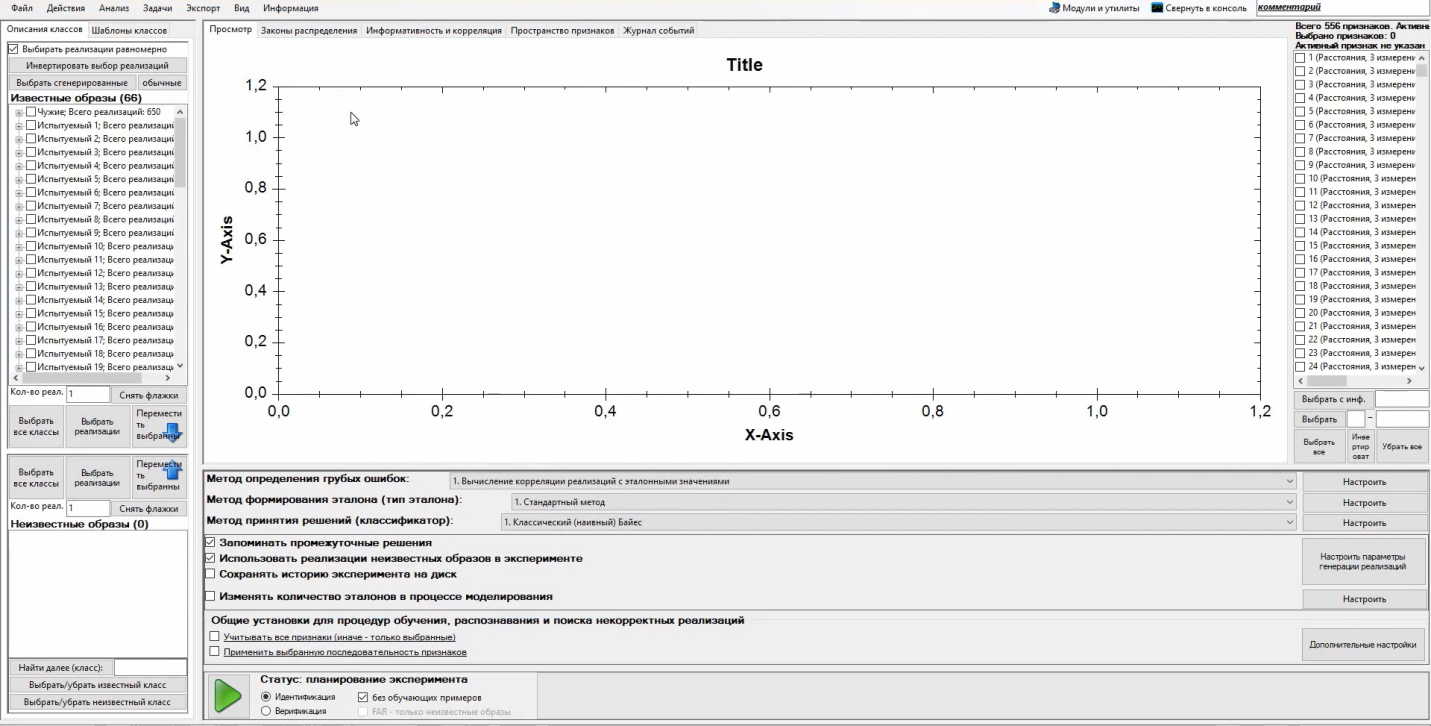


Рис.2. Подготовка данных в AIC desktop

### Формирование эталонов классов и

Для того чтобы создать эталонный образец каждого класса образов, необходимо проделать следующие шаги (рис. 3):

1. Выбрать все классы, забрав по 20 реализаций из каждого
2. Метод формирования эталона выбрать **«Персептрон (ЗНК, ГОСТ 52633.5, 1 слой, НПБК)»**

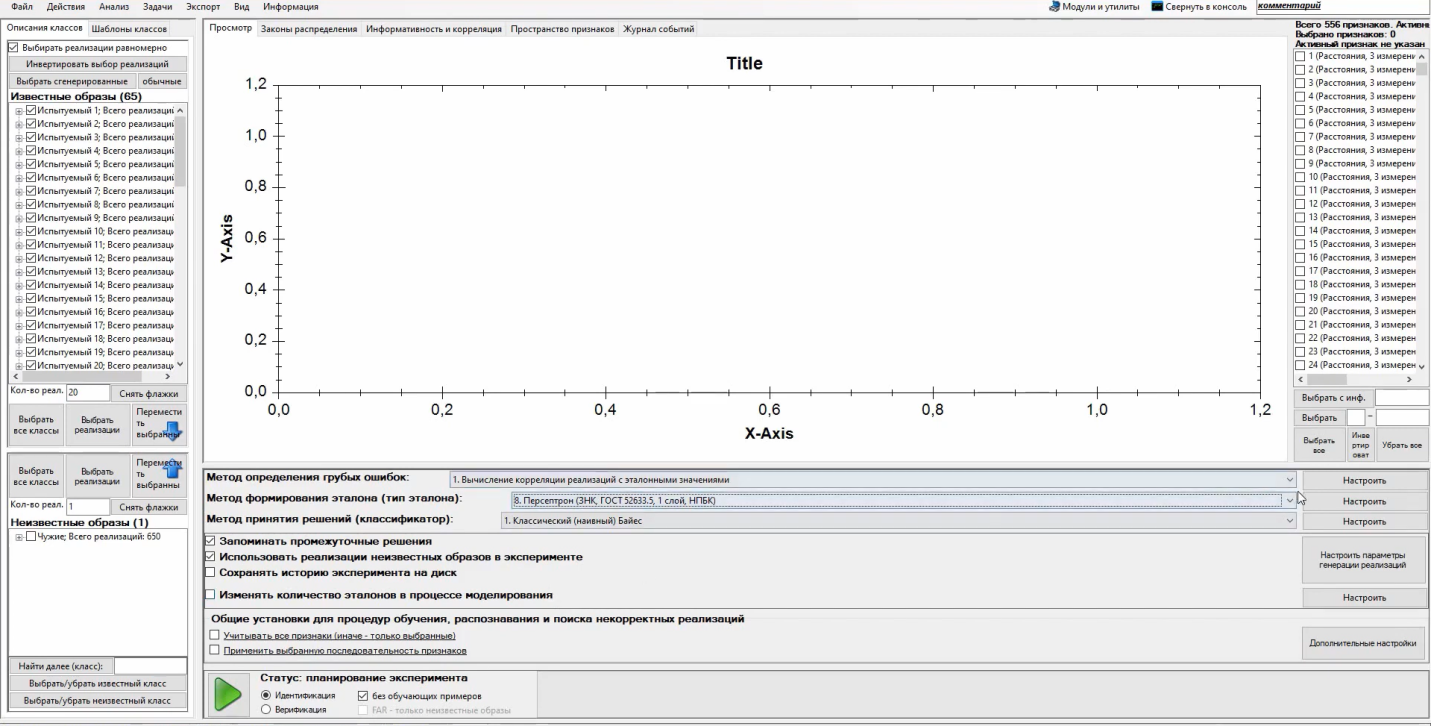


Рис.3. Выбор метода формирования эталонов

1. Далее необходимо настроить выбранный персептрон, нажав на кнопку **«Настроить»**, расположенную правее указанного метода (рис. 4): в открывшемся окне необходимо настроить первые два параметра. Число признаков указать равное 139, а число входов нейронов равное 4. Также необходимо снять галочки, стоящие напротив обоих параметров.

**Примечание:** количество признаков, в рассматриваемом датасете равно 556. Если 556 разделить на 4, получаем 139 (значения персептрона, настраиваемого в пункте 3). Во избежание атаки Маршалко, суть которой заключается в восстановлении идеального образа «Свой» (без тотального перебора) путем выявления общих связей нейронов (даже, если они зашифрованы) и общих весов, необходимо настроить НПБК таким образом, чтобы каждый признак входил только в один нейрон, что и было реализовано с помощью указанных выше параметров.

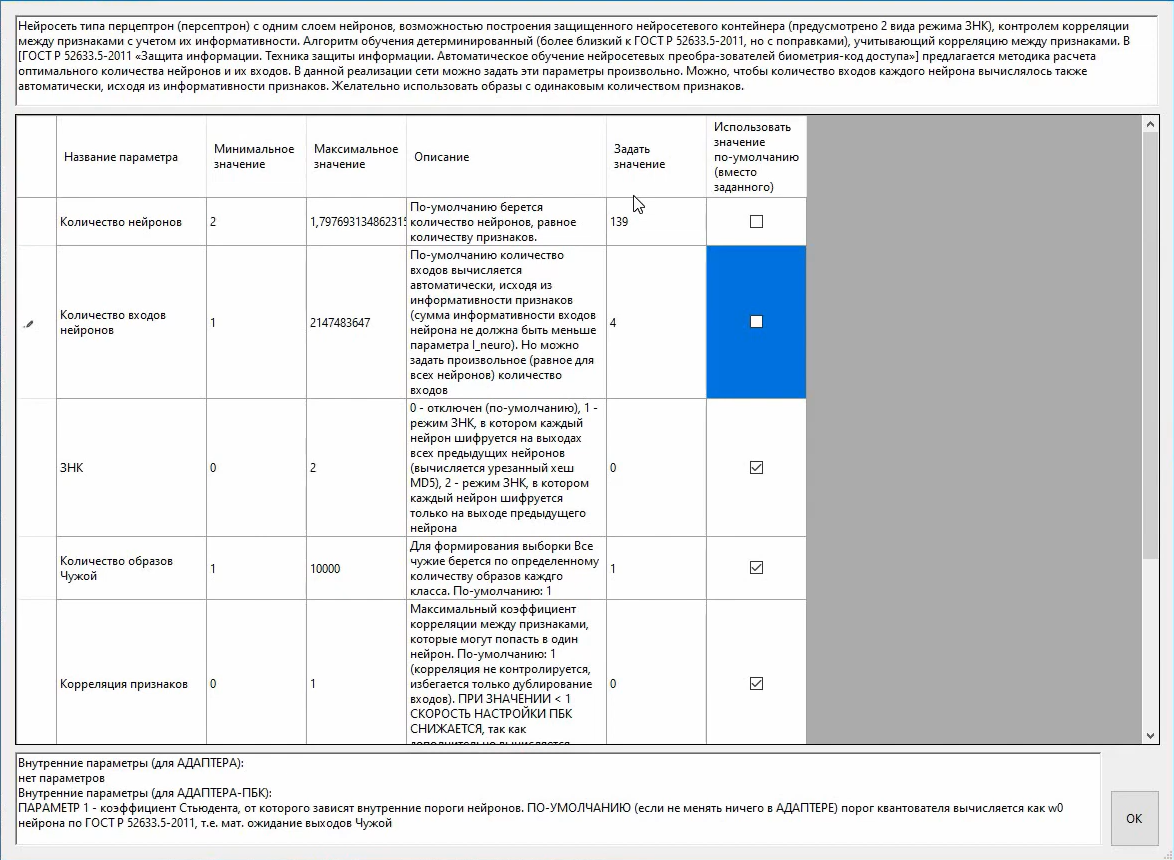
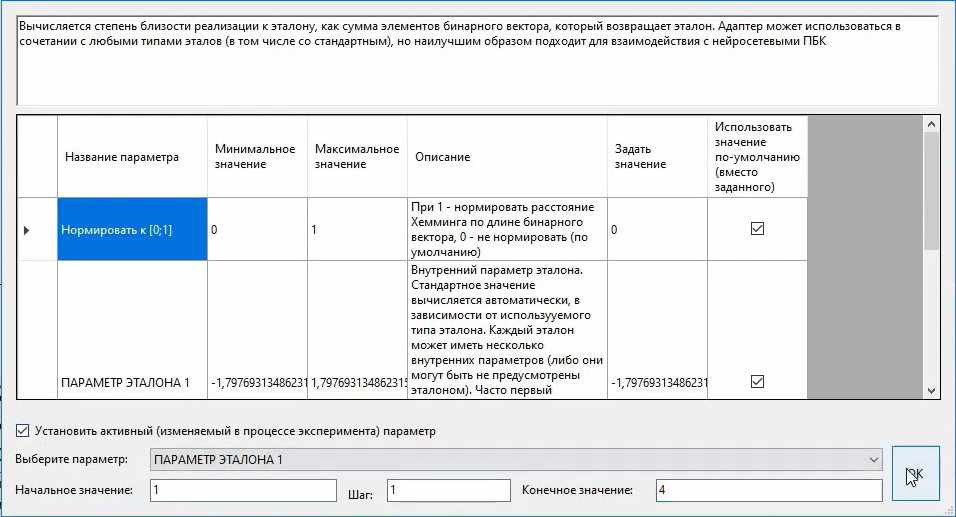


Рис. 4. Настройка персептрона

1. Метод принятия решений (классификатор) выбрать **АДАПТЕР-ПБК (мера Хэмминга, бинарные векторы)** и осуществить его настройку, нажав на кнопку **«Настроить»**, расположенную правее указанного метода (рис. 5): выбираем начальное равное 1, шаг 1, конечное значение равное 4. Также необходимо установить галочку в поле **«Установить активный (изменяемый в процессе эксперимента) параметр».**

Рис. 5. Настройка Адаптера-ПБК

1. В поле, расположенном ниже определения методов необходимо выбрать параметры:

* **«Запоминать промежуточные значения»**
* **«Использовать реализации неизвестных образов в эксперименте»**
* **«Сохранять историю эксперимента на диск»**

1. Выставить режим верификации без обучающих примеров в поле **«Статус»** (не выбирать **«FAR – только неизвестные образы»**)
2. Перейти на вкладку **«Задачи»** → **«Сформировать эталоны выбранных классов»** → **«Из выбранных реализаций»**
3. После формирования эталонов необходимо нажать кнопку пуск (зеленый треугольник в поле **«Статус»**) и запустить процесс работы Адаптера.

### Анализ ошибок 1-го и 2-го рода

В открывшемся после обучения окне (рис. 6) можно наблюдать графики ошибок первого и второго рода, пристраиваемые по разным параметрам. Построим график по порогу, задав значение нейронов равное 139, а значение активного параметра установим равное 4.

Дополнительно, помимо ошибок первого и второго рода, можно оценить значение EER.

**Примечание:** изменение активного параметра позволяет минимизировать ошибку второго рода. Это может быть необходимо, если при генерации ключей, не генерируется ключ на выходе.

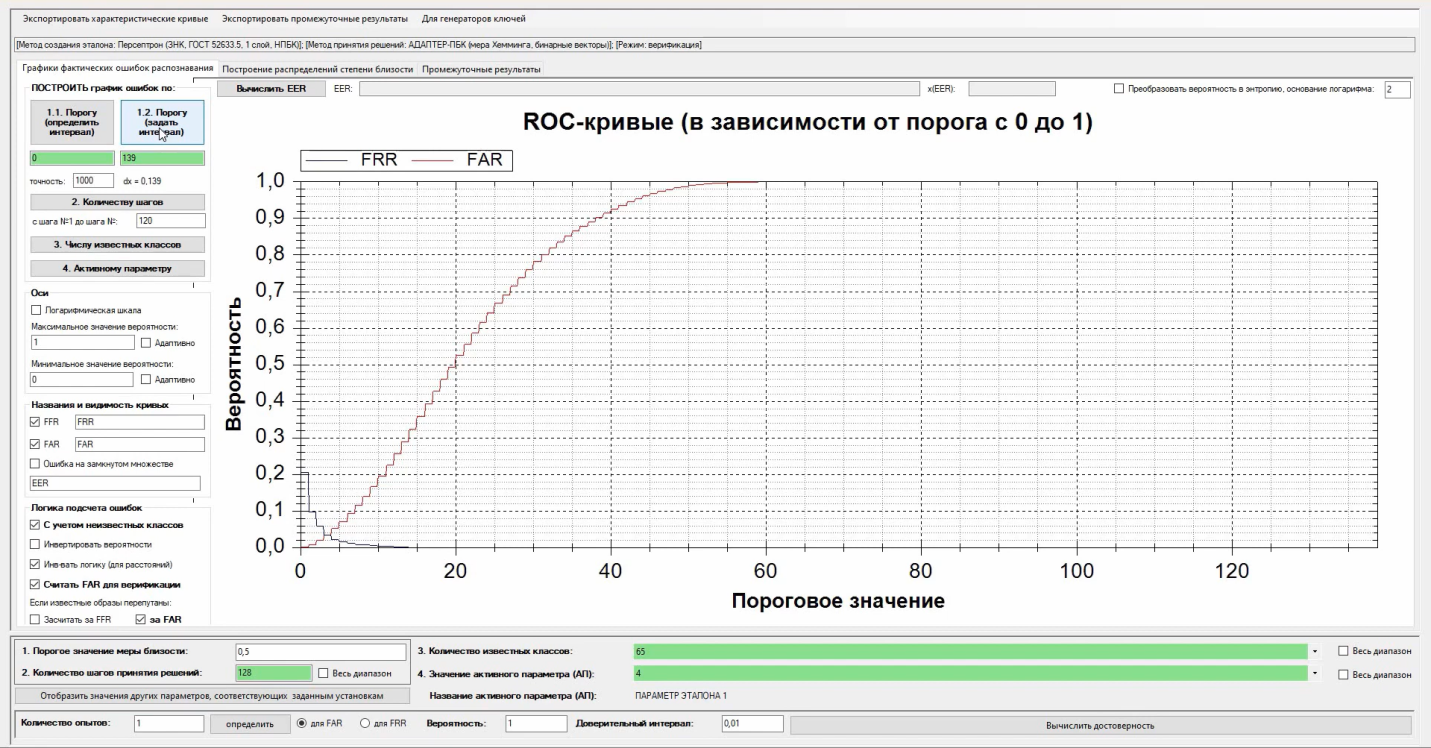
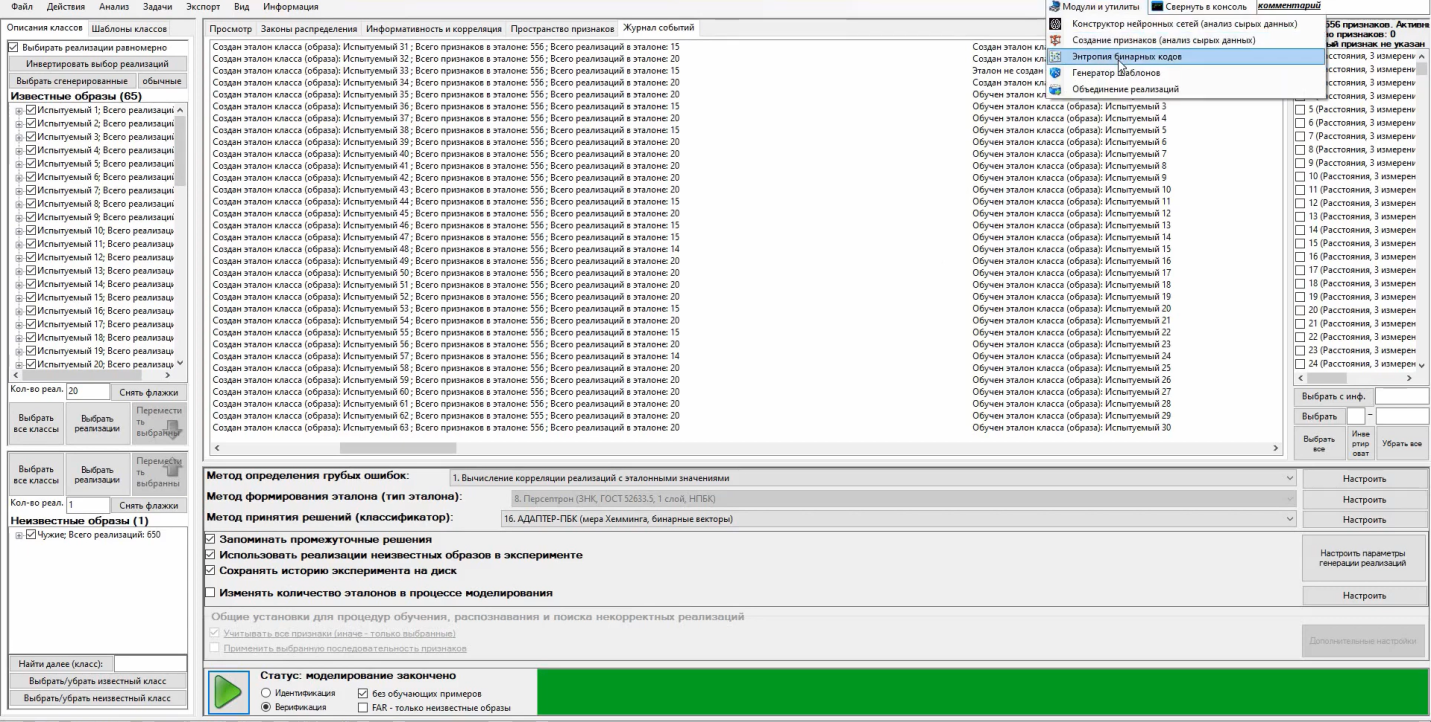


Рис. 6. Построение графика по порогу

### Энтропия бинарных кодов

Для того чтобы посмотреть, как выглядит энтропия выходных кодов, необходимо перейти на вкладку **«Энтропия бинарных кодов»** в верхней части меню, как показано на рисунке 7.

Рис. 7. Энтропия бинарных кодов

В открывшемся окне, в правом верхнем углу, необходимо нажать на кнопку **«Загрузить данные»** и выбрать вариантс кодом **\_0\_3 комментарий** в конце. После открытия кодов, выполняем переход во вкладку **«Графики»** и выбираем тип графика **Стабильность кодов Чужой-k (расстояние Хэмминга).** Дополнительно выбираем параметр **«показывать средние оценки»** и строим график **(**рис. 8).

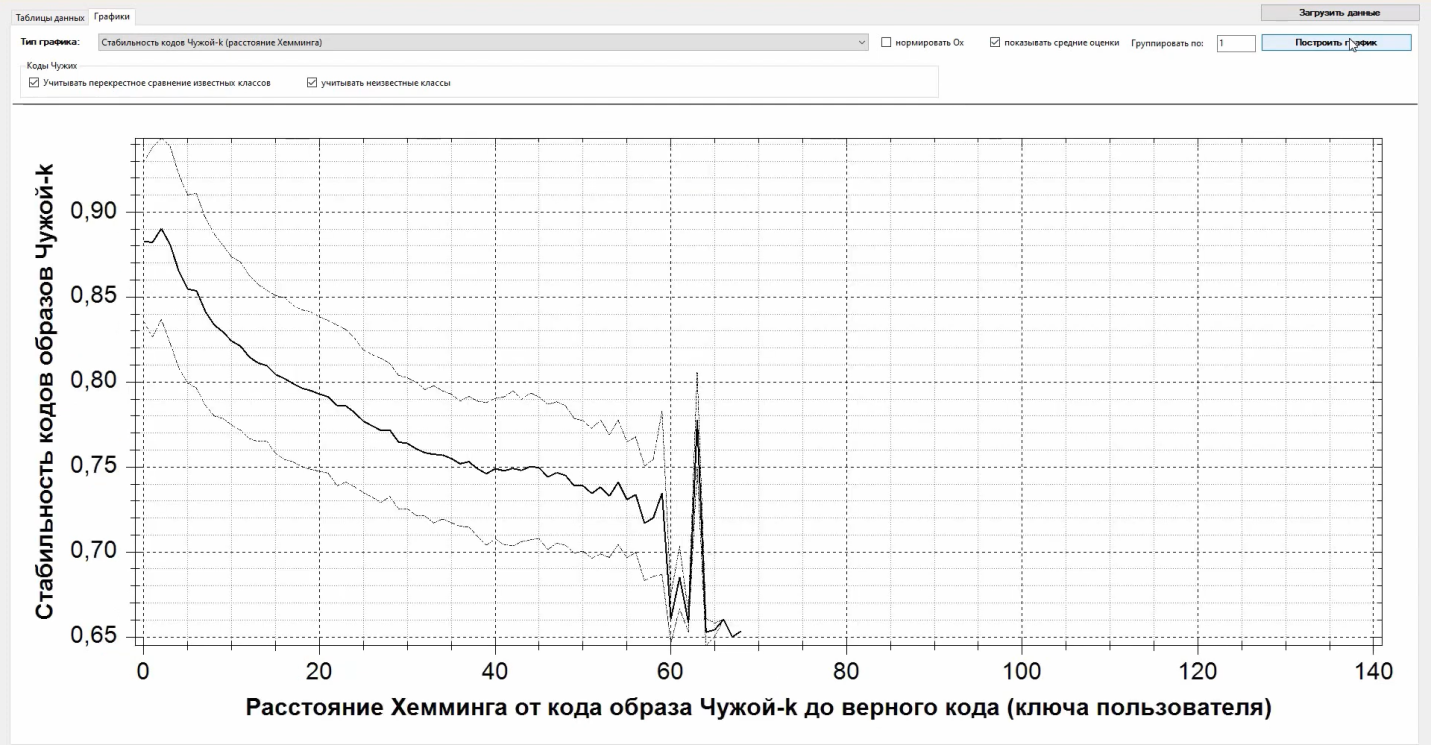


Рис. 8. Построение графика стабильности кодов «Чужих»

По оси X приведенного выше графика отложено расстояние Хэмминга отклика ПБК на образ класса под номером k, то есть на какие-то образы «Чужих» до верного кода пользователя «Свой» (подразумевается, что для образа «Свой» расстояние Хэмминга равно 0). Из графика видно, что стабильность кодов «Чужой» довольно высокая и при приближении к нулю повышается, это значит, что злоумышленник может вычислить стабильность.

## Самостоятельная работа

Повторите эксперимент в режиме *защищенного нейросетевого контейнера* (ЗНК):

1. Настройте персептрон (метод формирования эталона) таким образом, чтобы в поле для ЗНК стояла 1 и уберите галочку со значений по умолчанию;
2. Настройте АДАПТЕР-ПБК (метод принятия решения) таким образом, чтобы значение ПАРАМЕТРА ЭТАЛОНА 1 было равно 4. Также уберите галочки со значения по умолчанию и со значения установки активного параметра.
3. Сформируйте новые эталоны классов образов
4. Запустите обучение в режиме верификации
5. Постройте график ошибок первого и второго по порогу для 139 нейронов и проанализирует полученные графики. Почему на полученном графики произошли изменения в сравнении с графиком, представленном на рисунке 6?
6. Постройте и проанализируйте энтропию бинарных кодов полученного ПБК (необходимо загрузить последний проведенный эксперимент). Каким образом изменился график в сравнении с графиком, представленном на рисунке 8?

**КОНТРОЛЬНЫЕ ВОПРОСЫ**

**ДОПОЛНИТЕЛЬНЫЕ МАТЕРИАЛЫ И ИСТОЧНИКИ**