

Возможность применения метода нормализующих потоков для извлечения редких гамма событий в эксперименте TAIGA

Крюков А.П.^{1,*}, А.Ю. Разумов¹, Волчугов П.А.¹, Гресь Елизавета²,
Демичев А.П.¹, Дубенская Ю.Ю.¹, Журов Д.П.¹, С.П. Поляков³,
Е.Б. Постников¹

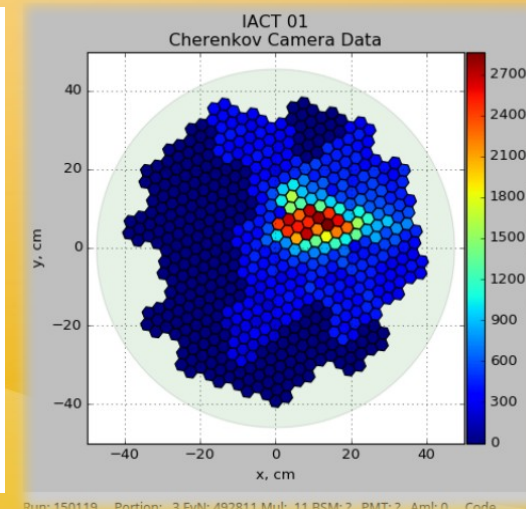
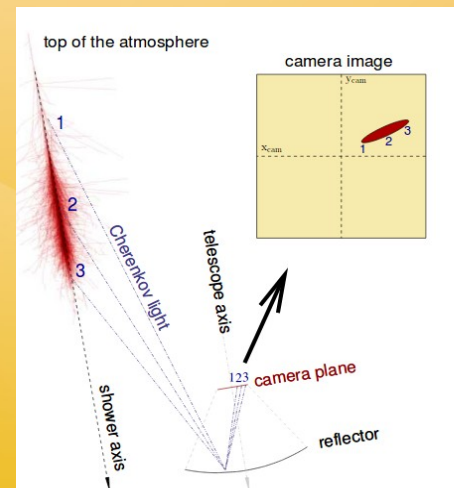
1) НИИЯФ МГУ; 2) НИИПФ ИГУ; 3) Институт проблем информаники и
автоматизации НАН РА

Работа выполнена при финансовой поддержке РФФб грант 24-11-00136

*) Email: kryukov@theory.sinp.msu.ru

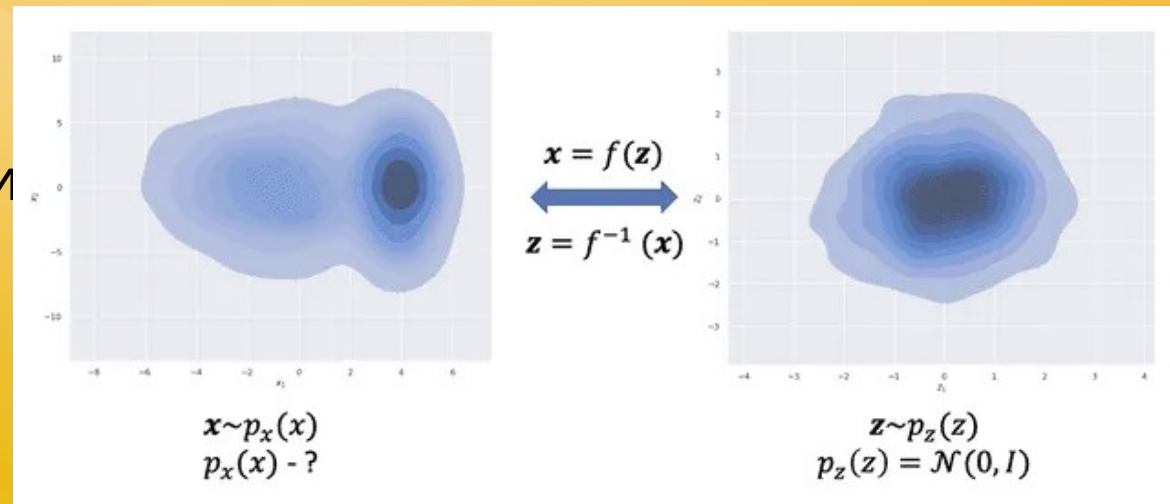
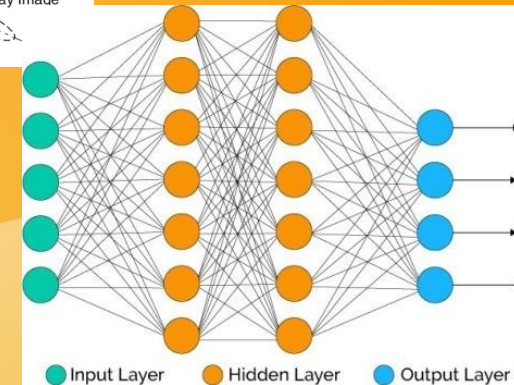
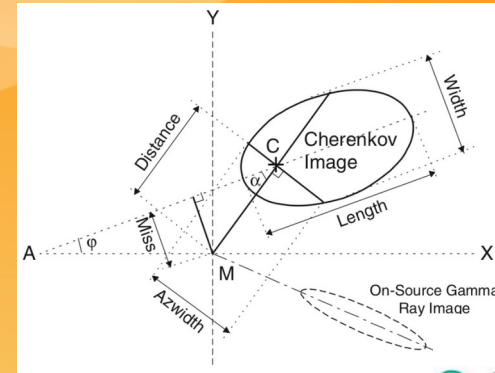
TAIGA-IACT

- Атмосферные черенковские телескопы находятся на астрофизическом полигоне НИИПФ ИГУ в Тункинской долине (в 50 км от оз. Байкал) в республике Бурятия на территории гамма-обсерватории TAIGA.
- АЧТ являются телескопами-рефлекторами с 4-метровым сегментированным сферическим зеркалом. В фокусе телескопов находится камера с 560-595 ФЭУ для регистрации черенковского излучения ШАЛ.



Методы выделения гамма событий

- Традиционный способ – обрезание по параметрам Хилласа
- Метод искусственных нейронных сетей
 - Полносвязанные сети
 - Сверточные сети
 - ...
 - Нормализующие потоки
 - ...



Особенности гамма астрономии

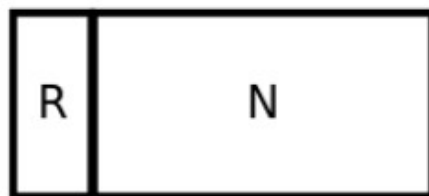
- Основной поток космических лучей – это зараженные частицы
 - Протоны, ядра гелия
 - Кислород, азот, ..., железо, ...
 - Гамма кванты всерхвысоких энергий составляет менее 0.1%
- Черенковский свет от ШАЛ, регистрируемый АЧТ, очень слабый (от нескольких сотен нескольких тысяч фотоэлектронов)
 - Большие шумы

Аномальные события против редких событий

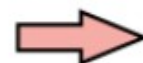
Редкие события



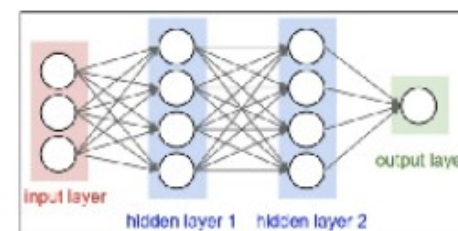
Обучающий набор данных: $R \sim N$



Тестовый набор данных: $R \ll N$

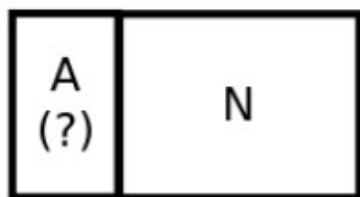


Two-class learning



Classifier

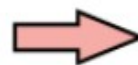
Аномальные события



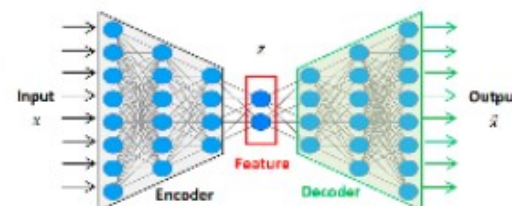
Обучающий набор данных: $A \ll N$



Тестовый набор данных: $A \sim N$



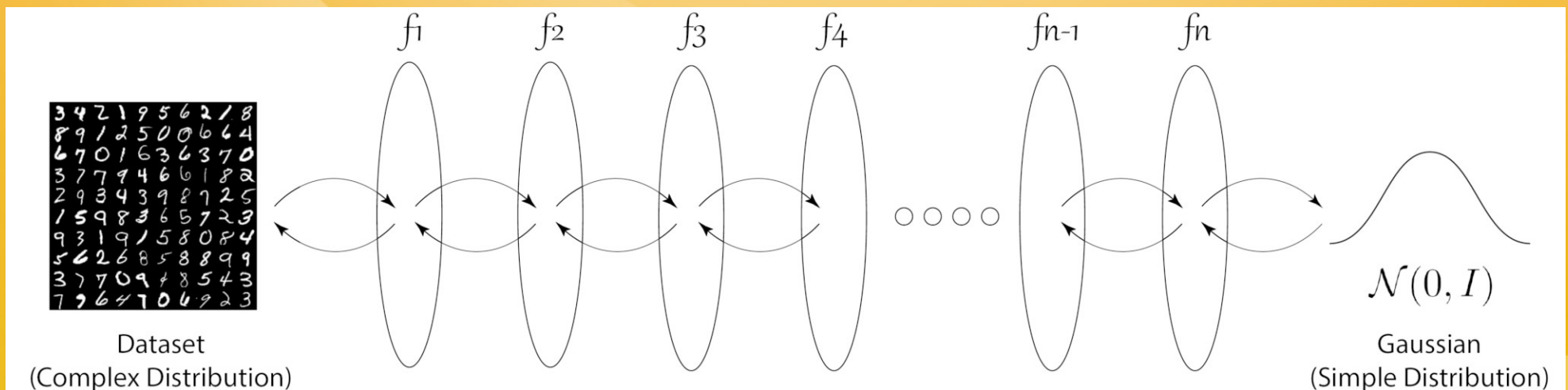
One-class learning



min "MSE"

Нормализующие потоки

- NF преобразует простое распределение случайного вектора в сложное.
- Это выполняется при помощи цепочки взаимнообратимых преобразований
 $f_1 \leftrightarrow f_2 \leftrightarrow \dots \leftrightarrow f_n$
- Алгоритм основан на теореме о замене переменных под интегралом.



Алгоритм выделения аномальных событий

- Отбор/выделение существенных признаков (понижение размерности) данных
- Преобразование распределения неаномальных данных в обучающей выборке в нормальное;
 - Любая непрерывная и биективная функция является монотонной
 - → сэмпл с большой вероятностью в исходном распределении перейдет в верхнюю часть (многомерного) нормального распределения (близко среднему)
- Поскольку отображение оптимизируется под выборку неаномальных данных, аномалии имеют низкую вероятность
 - в исходном пространстве со сложным распределением (к тому же имеем только выборку) сложно провести границу между нормальными и аномальными данными
 - для нормального распределения – легко, например, граница определяется 2σ (95%)

Цепочка преобразований

$$\begin{aligned} \mathbf{z}_{i-1} &\sim p_{i-1}(\mathbf{z}_{i-1}) \\ \mathbf{z}_i &= f_i(\mathbf{z}_{i-1}), \text{ thus } \mathbf{z}_{i-1} = f_i^{-1}(\mathbf{z}_i) \end{aligned} \tag{1}$$

$$\begin{aligned} p_i(\mathbf{z}_i) &= p_{i-1}(f_i^{-1}(\mathbf{z}_i)) \left| \det \frac{df_i^{-1}}{d\mathbf{z}_i} \right| \\ &= p_{i-1}(\mathbf{z}_{i-1}) \left| \det \left(\frac{df_i}{d\mathbf{z}_{i-1}} \right)^{-1} \right| && \Leftarrow \text{the inverse func theorem.} \\ &= p_{i-1}(\mathbf{z}_{i-1}) \left| \det \frac{df_i}{d\mathbf{z}_{i-1}} \right|^{-1} && \Leftarrow \text{Jacobians of invertible func.} \end{aligned}$$

$$\log p_i(\mathbf{z}_i) = \log p_{i-1}(\mathbf{z}_{i-1}) - \log \left| \det \frac{df_i}{d\mathbf{z}_{i-1}} \right| \tag{2}$$

Цепочка преобразований

- Необходимо построить цепочку преобразований такую, что
 - Легко вычислялся якобиан перехода $\rightarrow$ треугольная матрица
 - Легко вычислялись обратные функции $\rightarrow$ Экспонента.

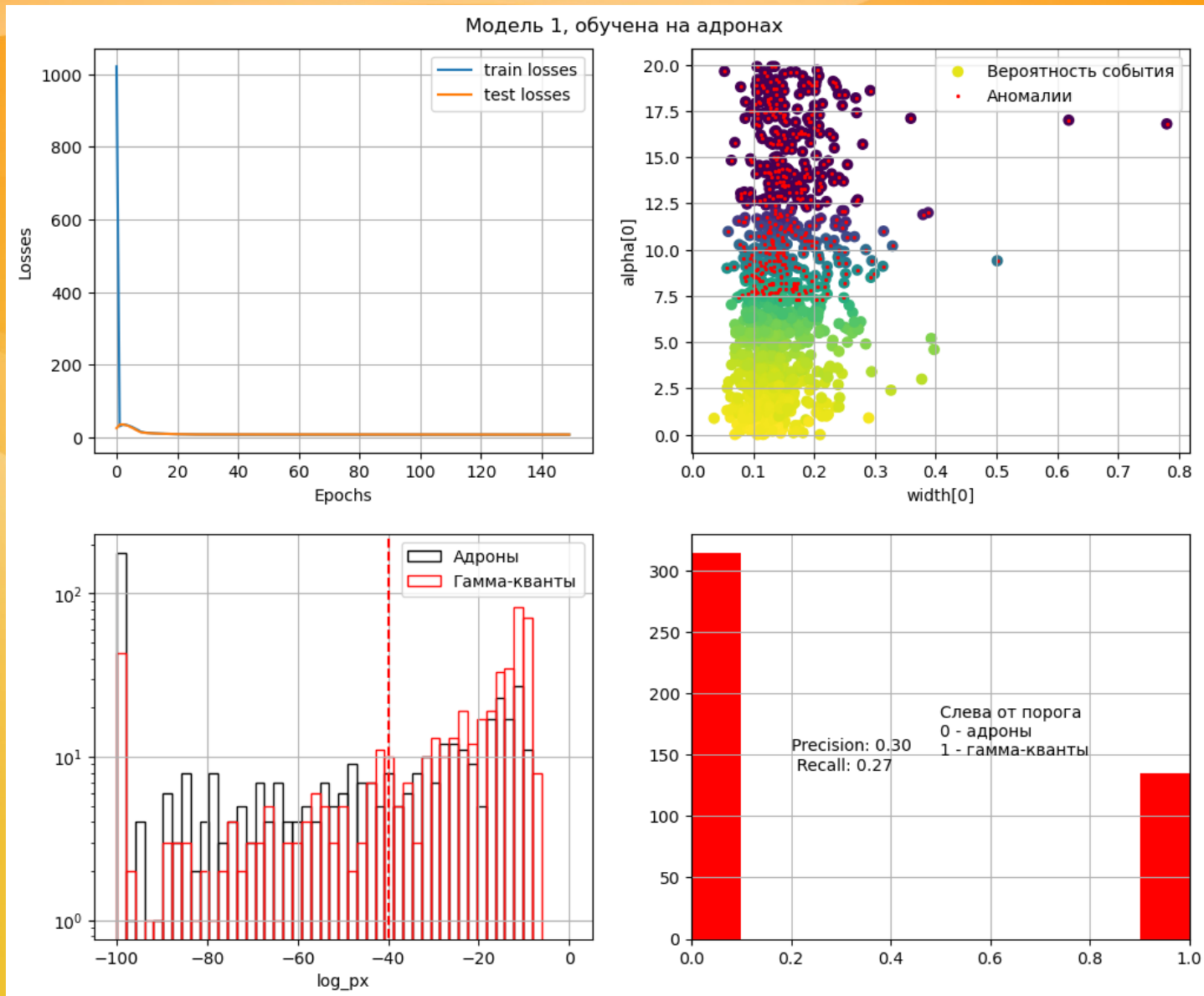
$$\begin{cases} \mathbf{y}_{1:d} &= \mathbf{x}_{1:d} \\ \mathbf{y}_{d+1:D} &= \mathbf{x}_{d+1:D} \odot \exp(\sigma(\mathbf{x}_{1:d})) + \mu(\mathbf{x}_{1:d}) \end{cases} \Leftrightarrow \quad (2)$$

$$\Leftrightarrow \begin{cases} \mathbf{x}_{1:d} &= \mathbf{y}_{1:d} \\ \mathbf{x}_{d+1:D} &= (\mathbf{y}_{d+1:D} - \mu(\mathbf{y}_{1:d})) \odot \exp(-\sigma(\mathbf{y}_{1:d})) \end{cases} \quad (3)$$

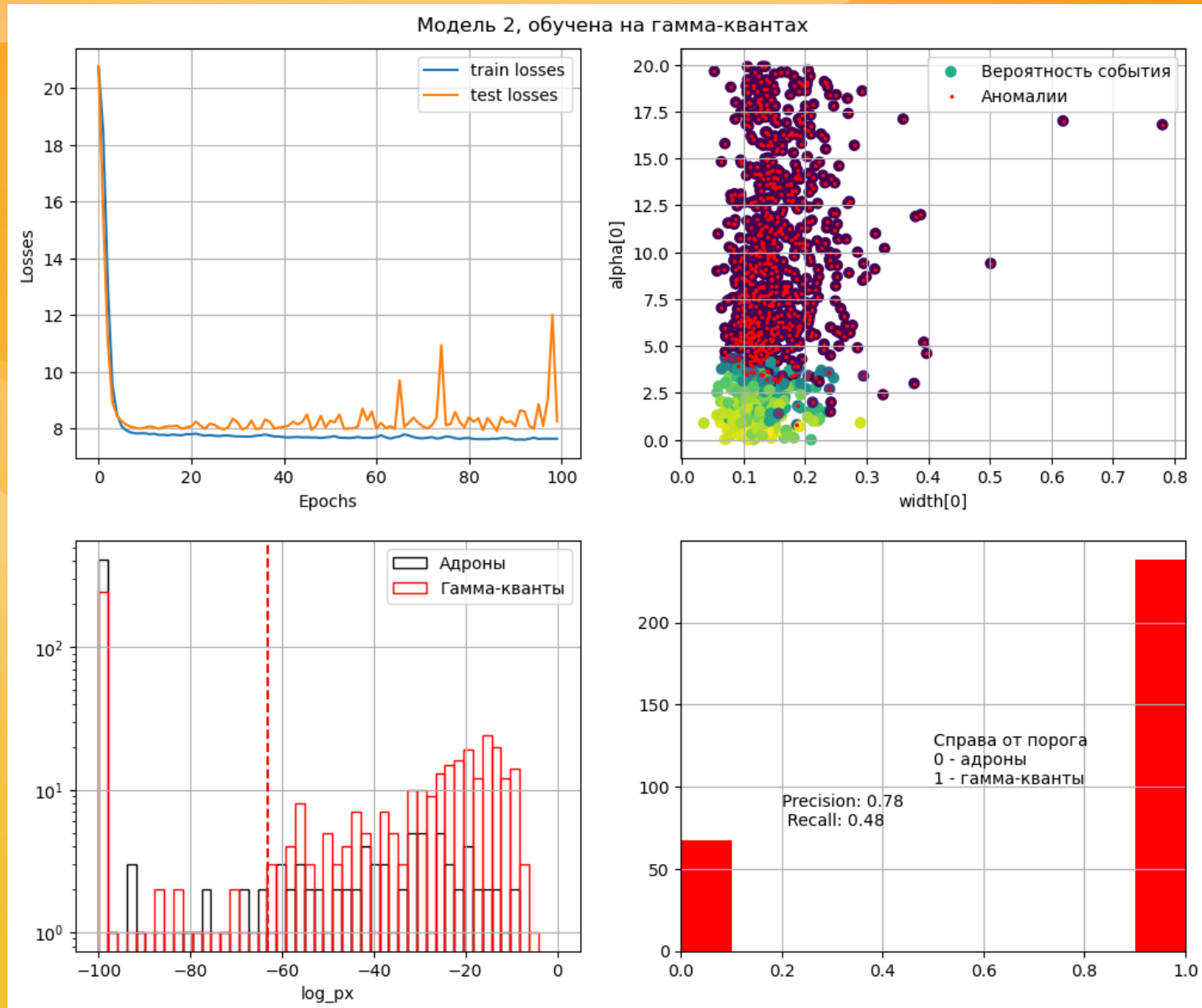
Параметры NF

- Параметры Хилласа, включенные в обучение NF
 - $\text{Size} > 100$
 - $0.02 < \text{Width} < 0.8$
 - $\text{Alpha} < 20$
 - $\text{Dist} < 3$
- Обучающая и валидационная выборка
 - $\text{Train_size} = 3000$
 - $\text{Test_size} = 500$
 - Блоки μ и σ содержат 2 скрытых полносвязанных слоя по 512 нейронов

Адроны – нормальные события



Гамма – нормальные события



Заключение

- Наиболее перспективным использованием NF является случай, когда в качестве аномальных событий рассматриваются протоны, а не гамма.
- На небольших нейросетевых моделях было продемонстрировано, что нормализующие потоки обладают потенциалом, позволяющий выделять редкие гамма события.
- Однако, надежность выделения гамма является недостаточным для практического использования в гамма астрономии.
 - Возможно, что причиной является малый объем обучающей выборки.
 - Возможно требуется использовать продвинутые нейросетевые модели NF с большим количеством блоков преобразования, слоев и нейронов в слоях.
- Предполагается продолжить изучение данного метода на более сложных моделях и существенно увеличенных по размеру обучающей выборки.

Спасибо за внимание.