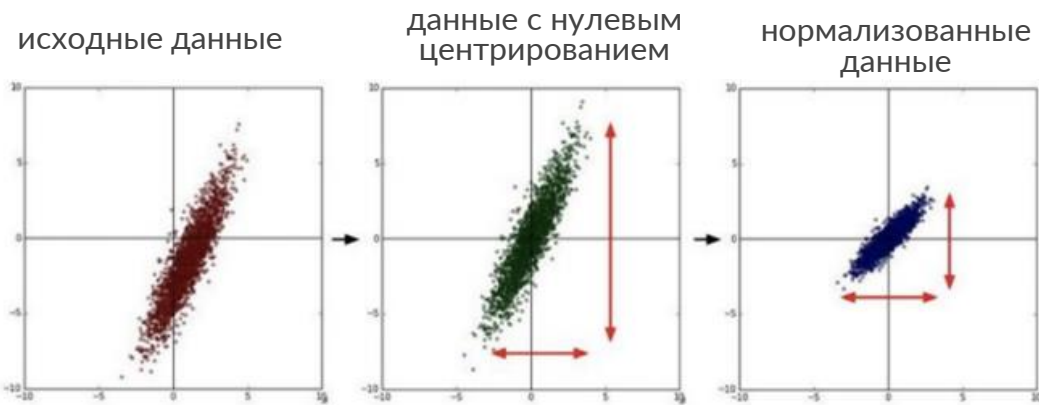
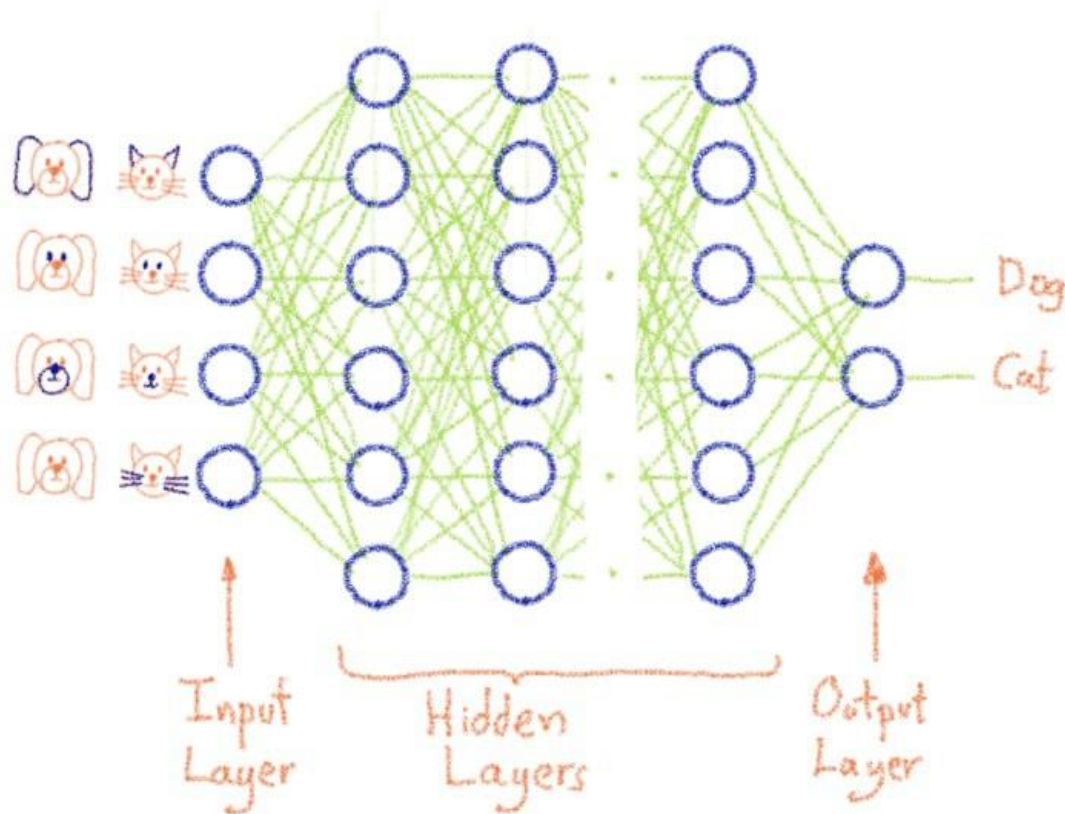




# Методы регуляризации нейронных сетей

**Тетерин Михаил Александрович**





“Обычно хорошей идеей является нормализация входов линейной модели”

//среднее по мини батчу

$$\mu_{\beta} \leftarrow \frac{1}{m} \sum_{i=1}^m x_i$$

//отклонение по мини батчу

$$\sigma_{\beta}^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\beta})^2$$

//нормализация

$$\hat{x}_i \leftarrow \frac{x_i - \mu_{\beta}}{\sqrt{\sigma_{\beta}^2 + \epsilon}}$$

//масштабирование и сдвиг

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma, \beta}(x_i)$$



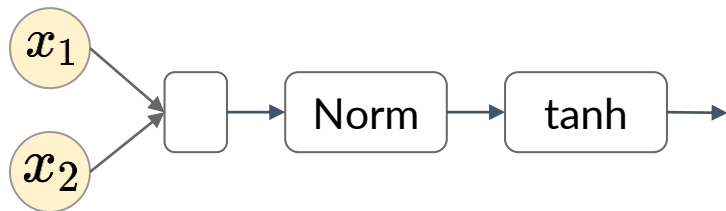
Ввод: значение  $x$  по мини батчу:

$$\beta = \{x_{1..m}\}$$

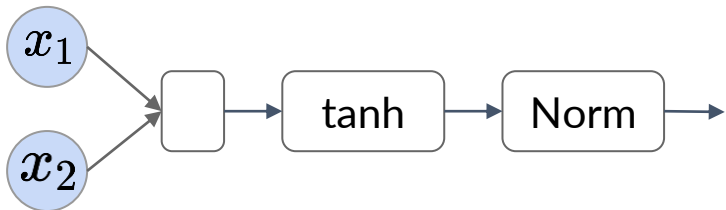
параметры обучения:  $\gamma, \beta$

Вывод:  $\{y_i = BN_{\gamma, \beta}(x_i)\}$

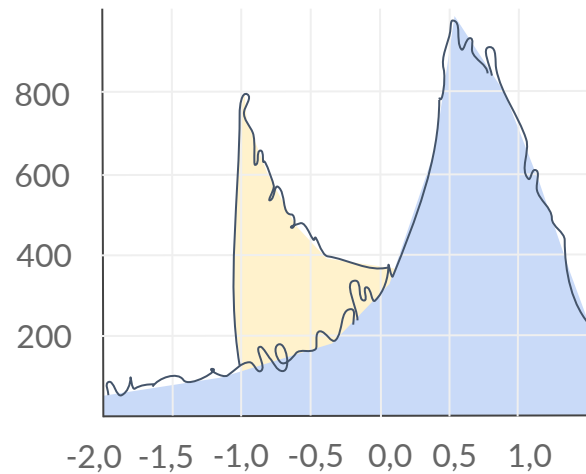
## A Нормализация до нелинейности

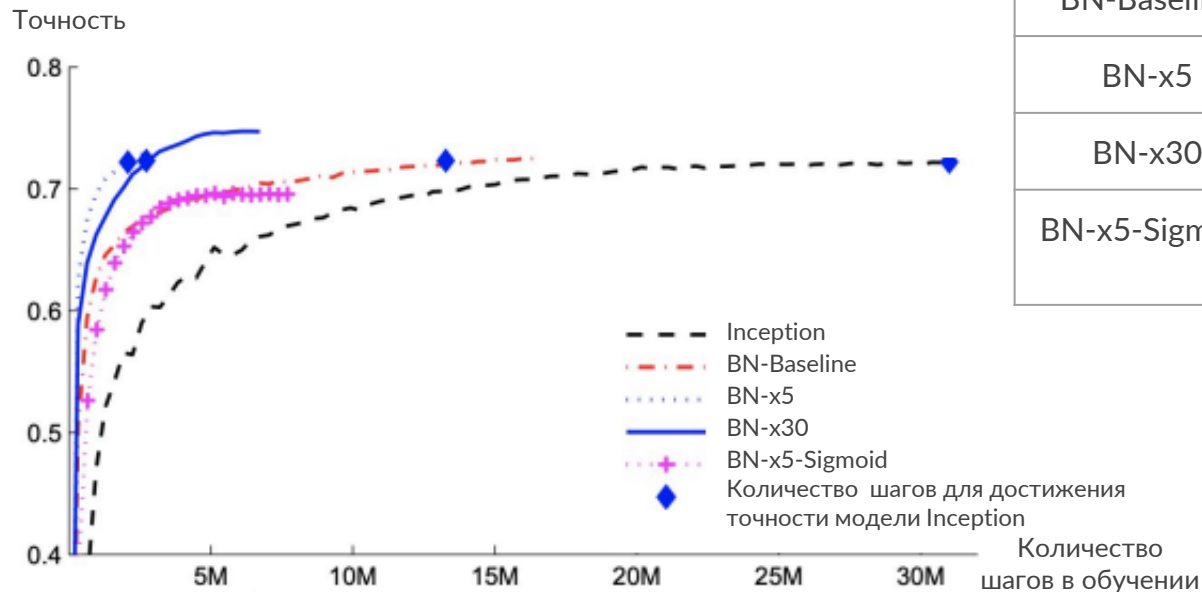


## B Нормализация после нелинейности



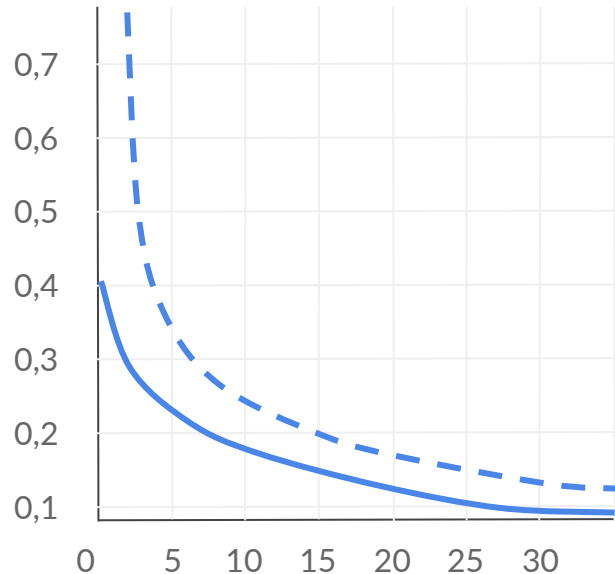
## Реализация сэмплирования



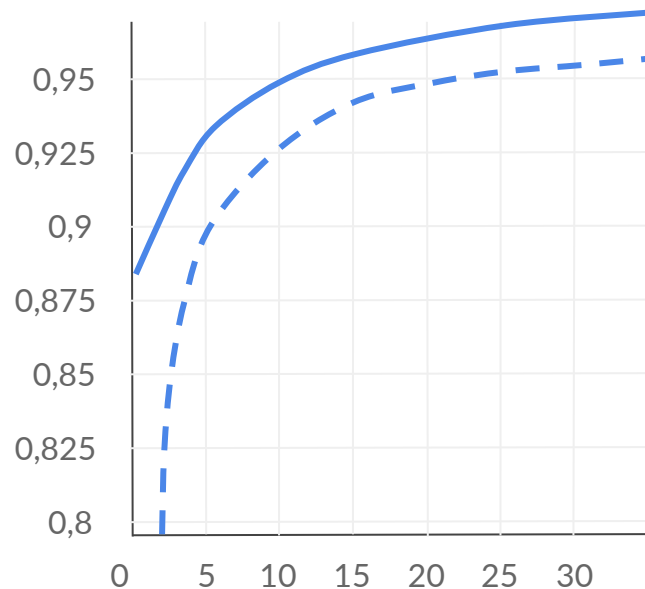


Модель	Шаги к 72.2%	Максимальная точность
Inception	$31.0 \cdot 10^6$	72.2%
BN-Baseline	$13.3 \cdot 10^6$	72.7%
BN-x5	$2.1 \cdot 10^6$	73.0%
BN-x30	$2.7 \cdot 10^6$	74.8%
BN-x5-Sigmoid		69.8%

Функция потерь



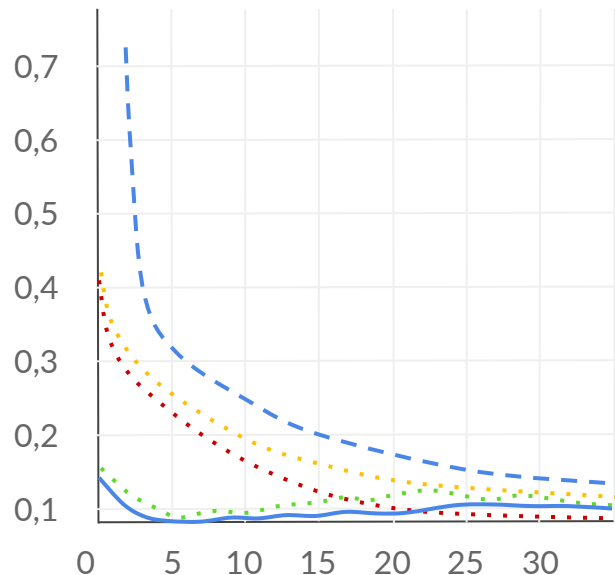
Точность



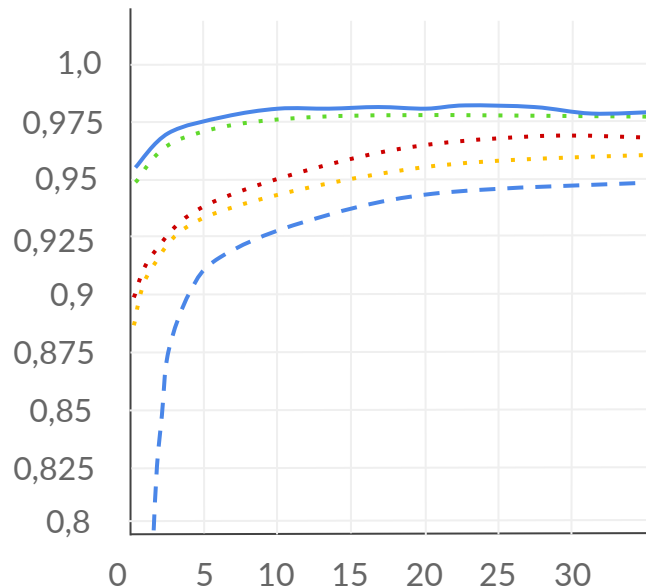
- Без нормализации
- С нормализацией

# Сравнение разных методов для улучшения обучения нейронных сетей

Функция потерь

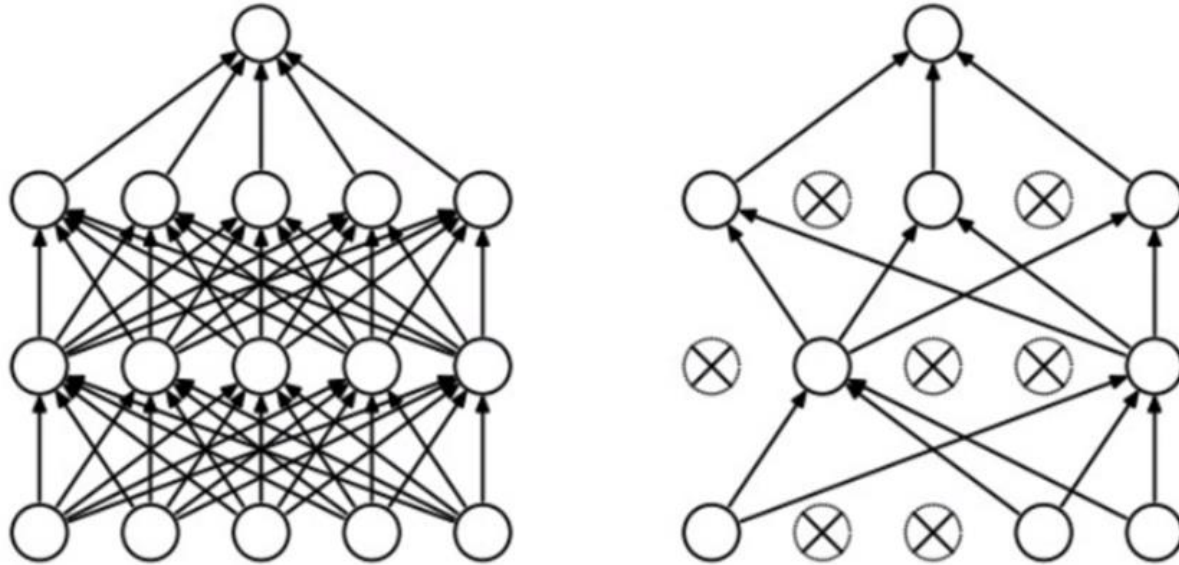


Точность



- Начальная модель
- Adam + Xavier + BN
- ... Xavier
- ... BN
- ... Adam





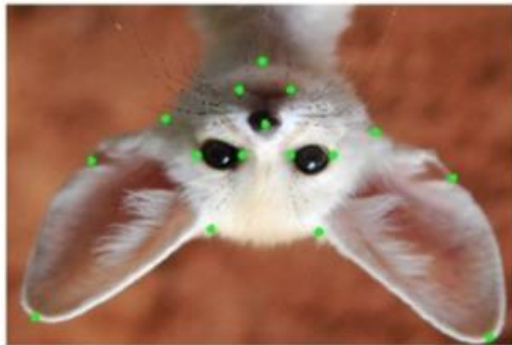
Оригинал



Горизонтальное отображение

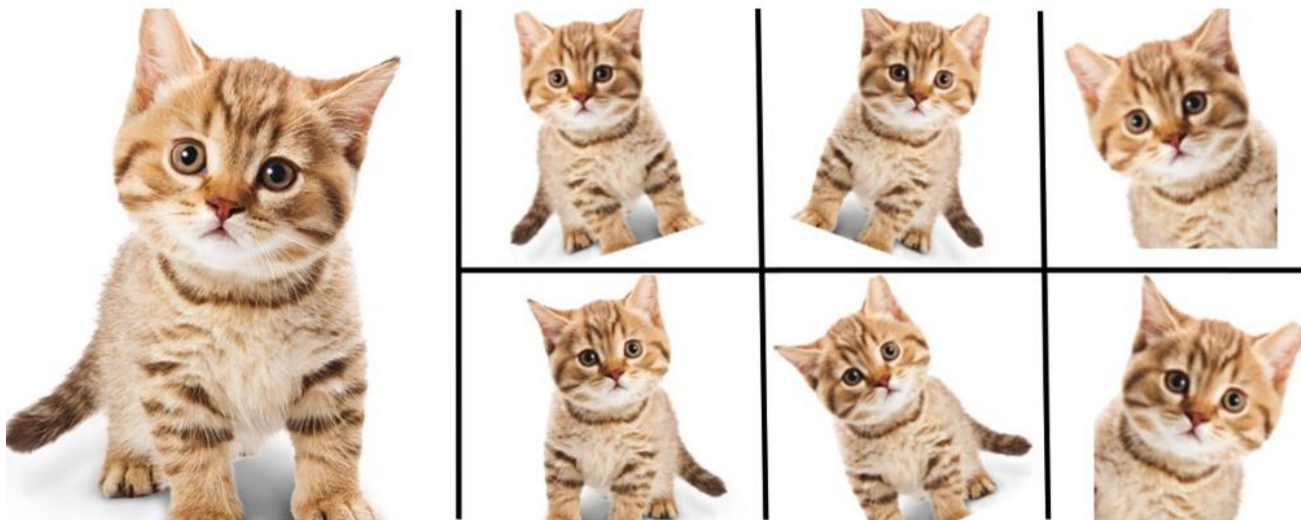


Вертикальное отображение



Поворот





# Спасибо за внимание

**Тетерин Михаил Александрович**

ОБРАЗОВАТЕЛЬНЫЕ ПРОГРАММЫ В ОБЛАСТИ  
ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА