

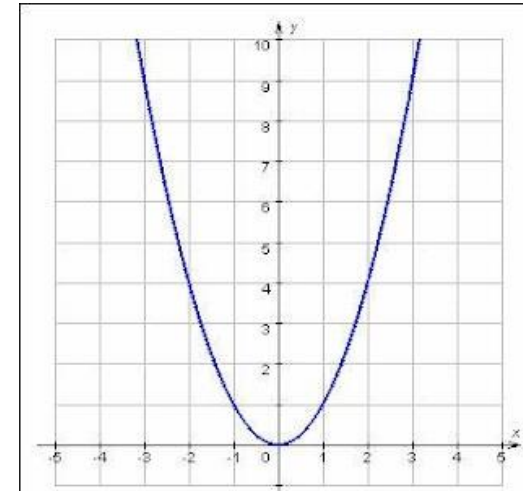
Семинар 3

Оптимизаторы и метрики качества модели

Почему алгоритму тяжело сойтись?

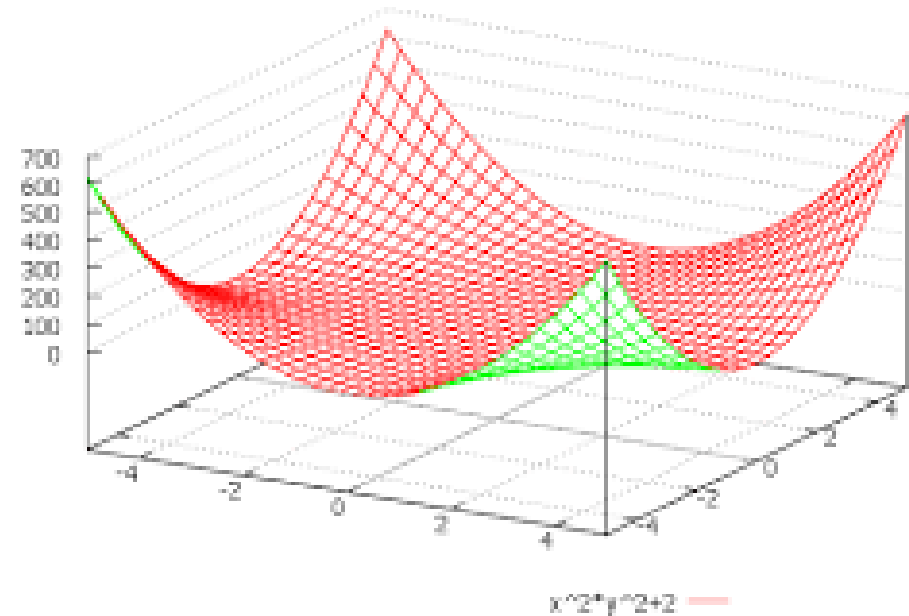
1. $y = w_1 x$

1 параметр + функция потерь



2. $y = w_1 x + w_0$

2 параметра + функция потерь

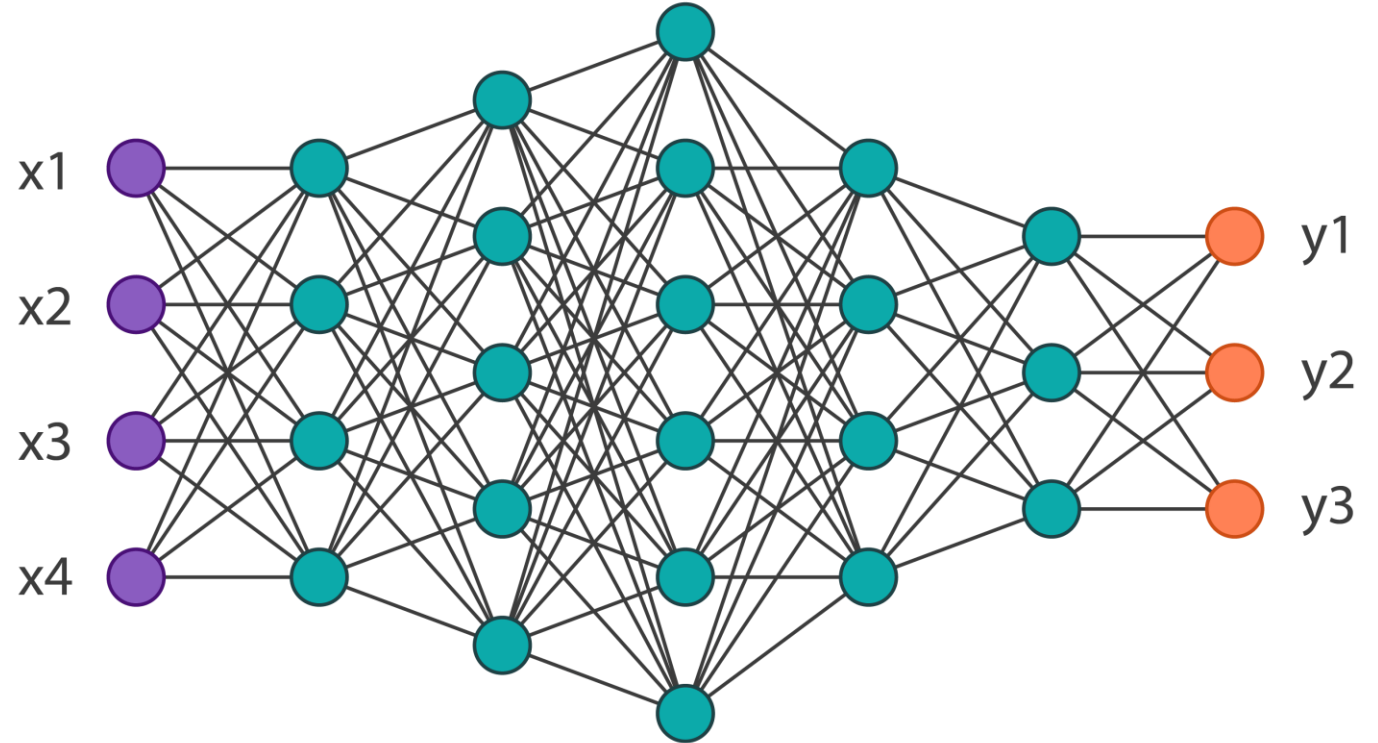


Почему алгоритму тяжело сойтись?

У этой нейросети
параметров:

$4 \cdot 4 + 4 \cdot 5 + 5 \cdot 6 + 6 \cdot 4 +$
 $4 \cdot 3 + 3 \cdot 3 = 16 + 20 +$
 $30 + 24 + 12 + 9 = 111$
параметров + функция
оптимизации

То есть стоит задача поиска
минимума в **111-мерном**
пространстве



Почему алгоритму тяжело сойтись?

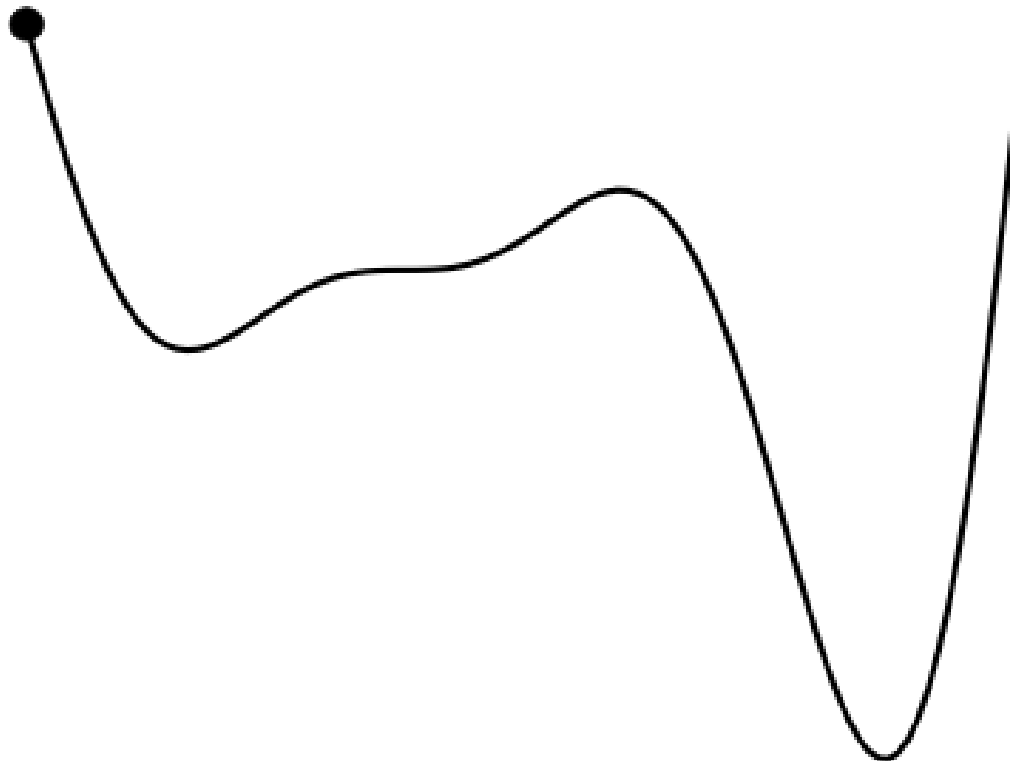
Нейросеть Stable Diffusion имеет **890 миллионов** параметров.

Нейросеть DALL-E-2 имеет **4.5 миллиарда** параметров.



Оптимизаторы

Простой оптимизатор работает так



Оптимизаторы

Формулы:

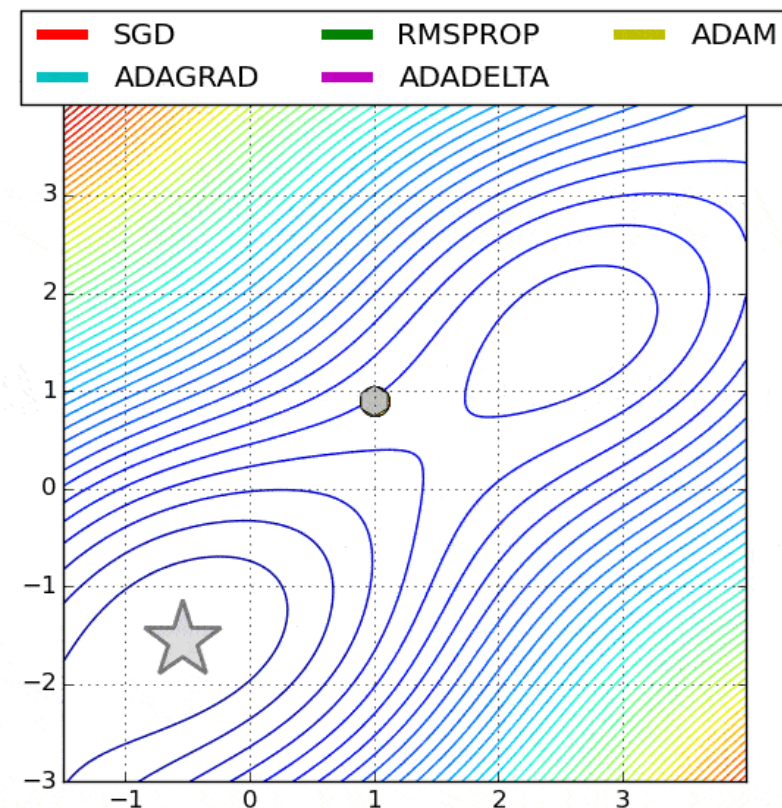
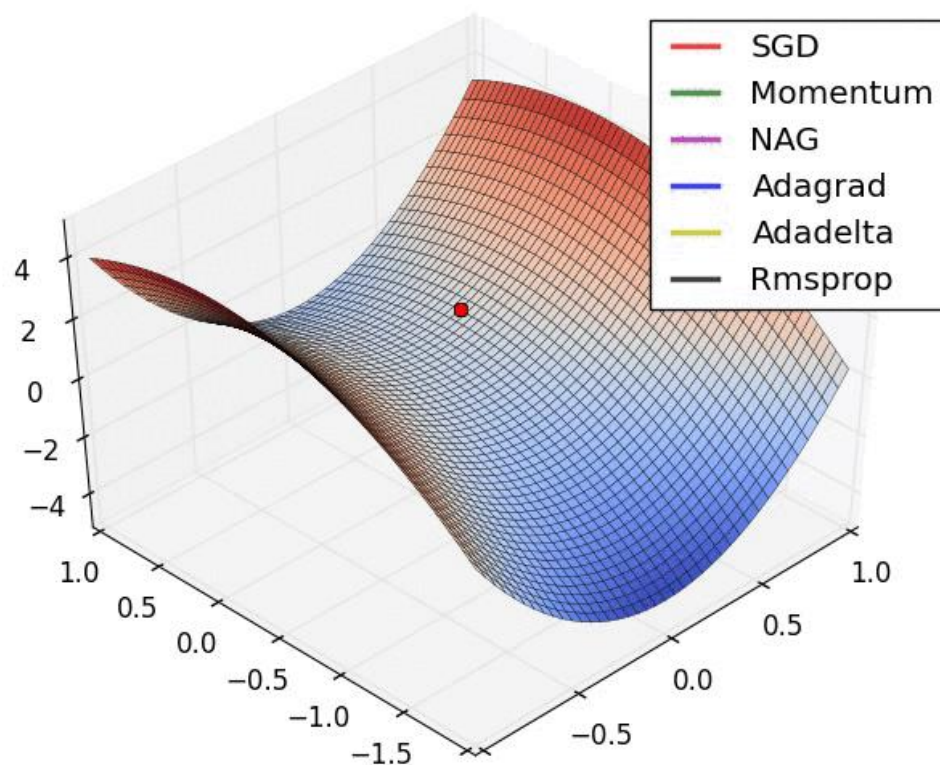
$$\theta_j = \theta_j - \eta \cdot \overbrace{\frac{\partial C}{\partial \theta_j}}^{\text{Backprop}}$$

$$\theta = \theta - \eta \nabla J(\theta) + \gamma v_t$$

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{\epsilon + \sum_{\tau=1}^t (\nabla J(\theta_{\tau,i}))^2}} \nabla J(\theta_{t,i})$$

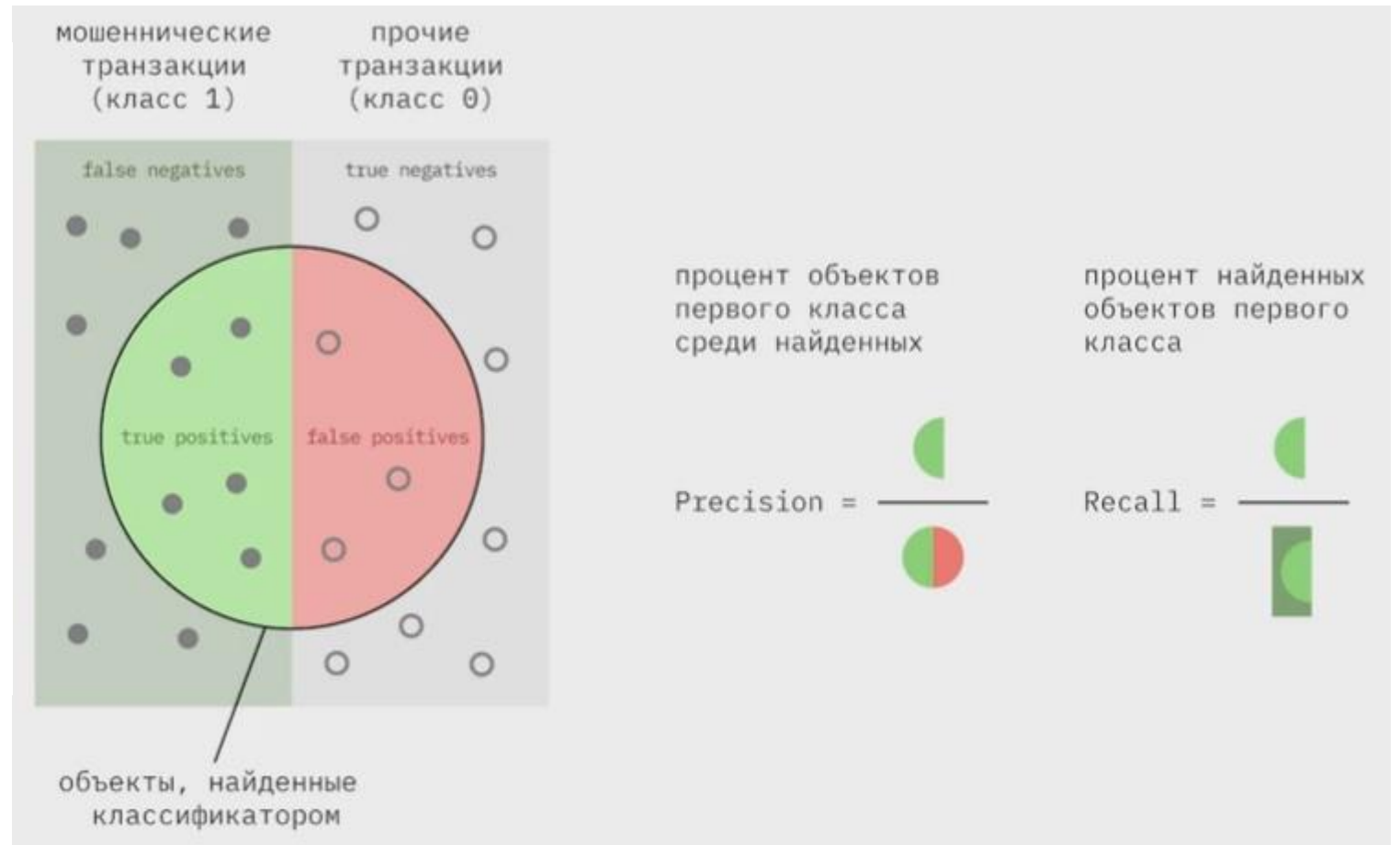
Оптимизаторы

Посложнее оптимизаторы работают так



Метрики качества: классификация

		Predicted class	
		P	N
Actual Class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)



Метрики качества: регрессия

1. **Квадратичная** (Q — MSE)

$$L(a, x, y) = (a(x) - y)^2$$

2. **Абсолютная** (Q — MAE)

$$L(a, x, y) = |a(x) - y|$$

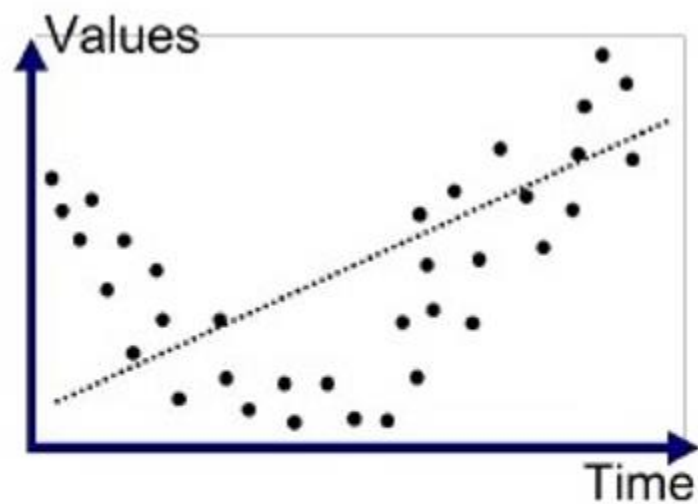
3. **Логарифмическая** (Q — MSLE)

$$L(a, x, y) = (\log(a(x) + 1) - \log(y + 1))^2$$

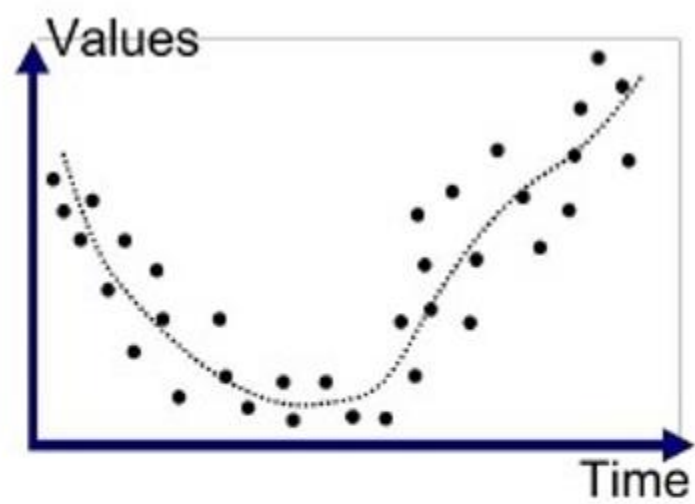
4. **Абсолютная-процентная** (Q — MAPE)

$$L(a, x, y) = \frac{|a(x) - y|}{y}$$

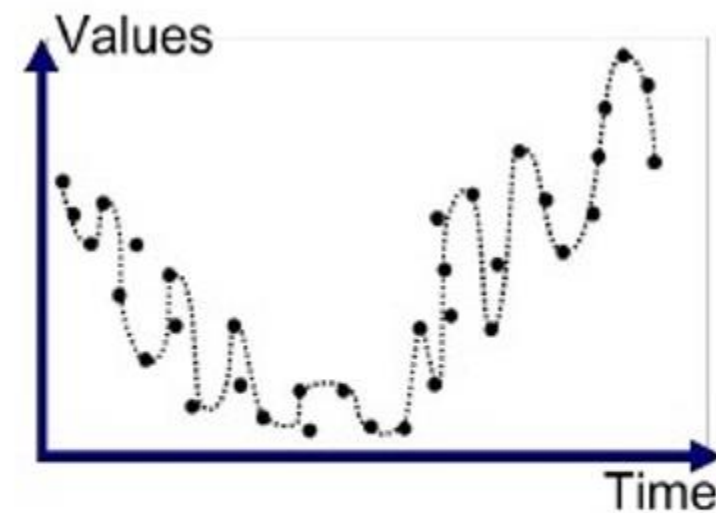
Переобучение



Underfitted



Good Fit/Robust



Overfitted

Переобучение



Практика

Пишем код