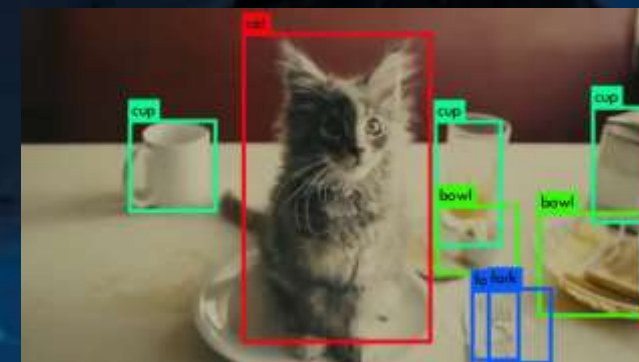
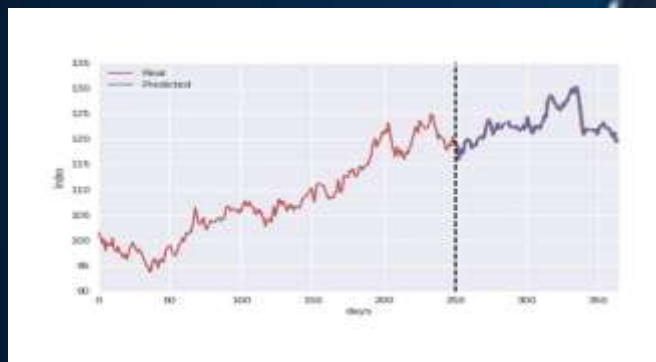
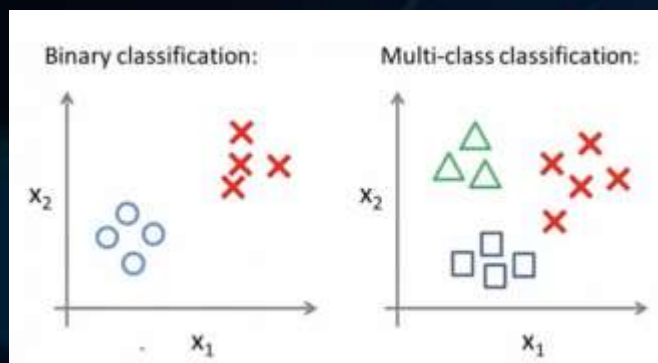


ПРЕДСКАЗАНИЕ ВРЕМЕННЫХ РЯДОВ ПРИ ПОМОЩИ НЕЙРОННЫХ СЕТЕЙ

Выполнил студент 211 группы
Попов Иван Васильевич под
руководством д.ф-м.н.
Голубцова Петра Викторовича.

Области применения нейронных сетей

- Классификация
- Прогнозирование
- Распознавание

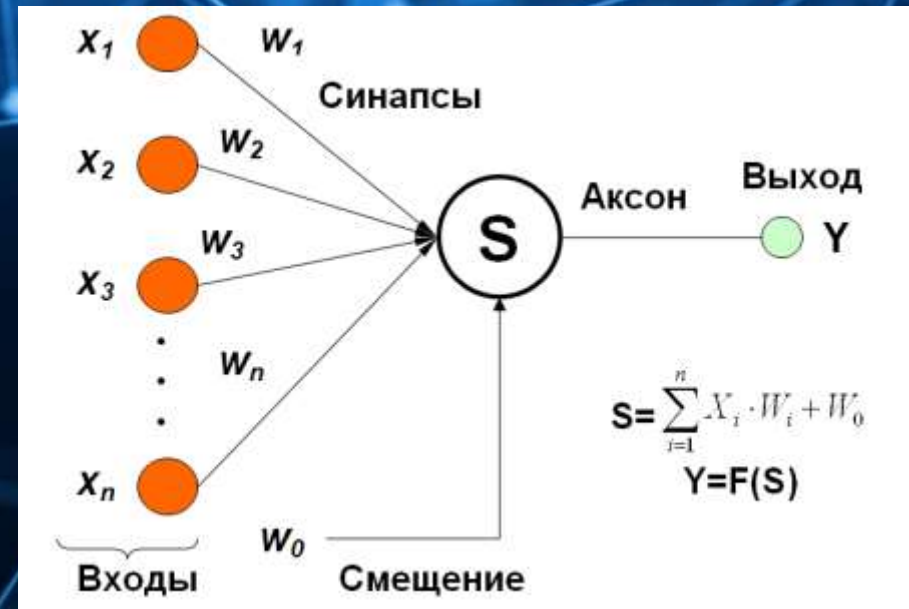
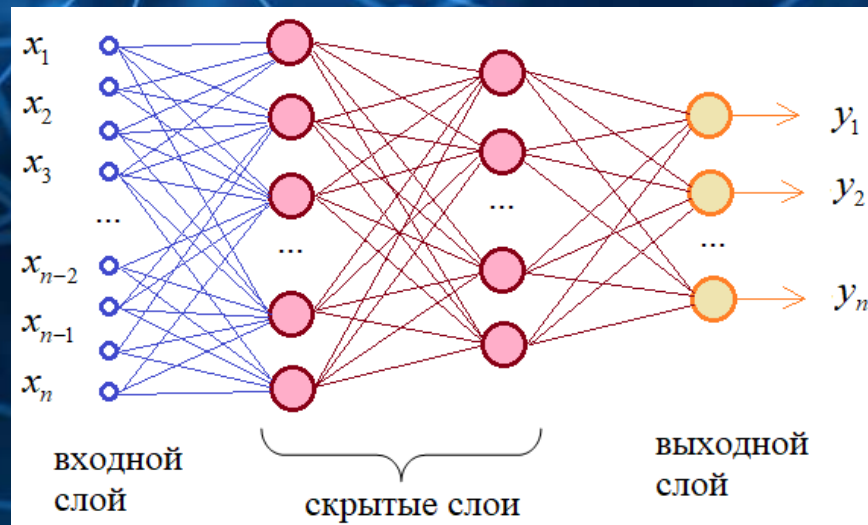


Математическая модель нейрона.

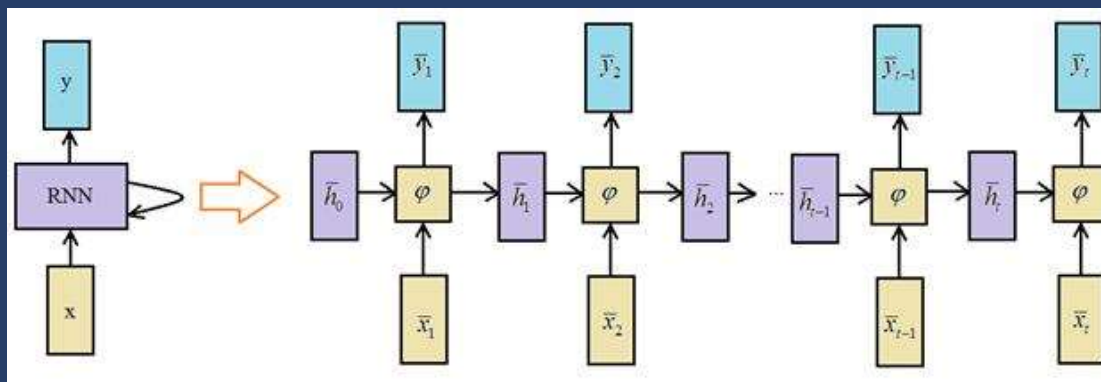
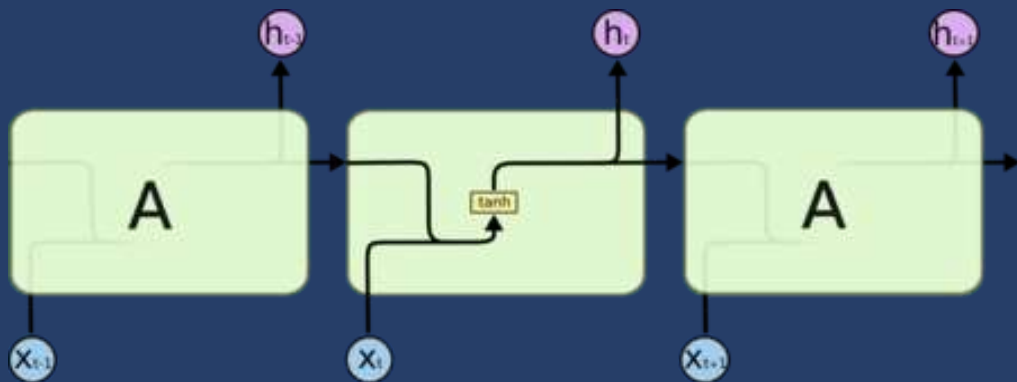
Полносвязные нейронные сети



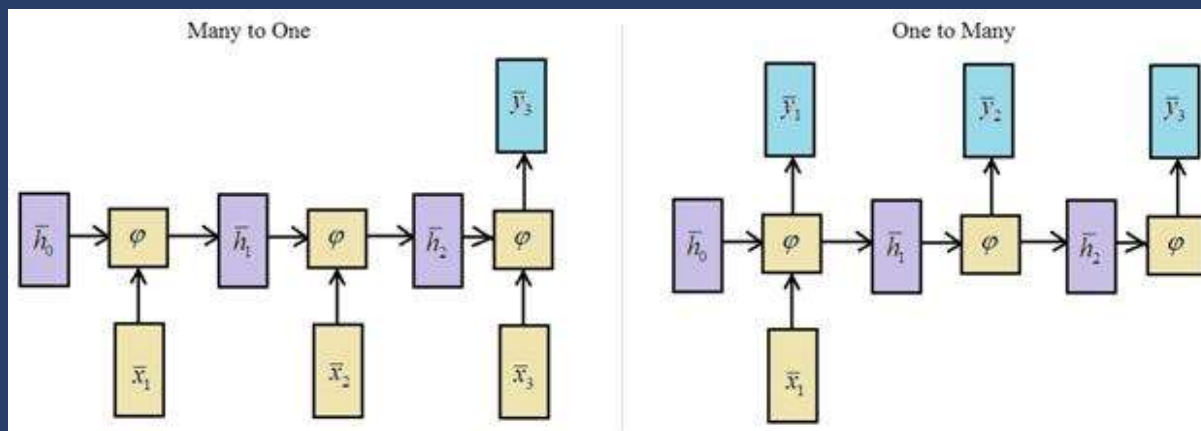
На картинке слева представлена биологическая модель нейрона. Именно на ее идеях основана математическая модель, приведенная на рисунке ниже.



Рекуррентные нейронные сети

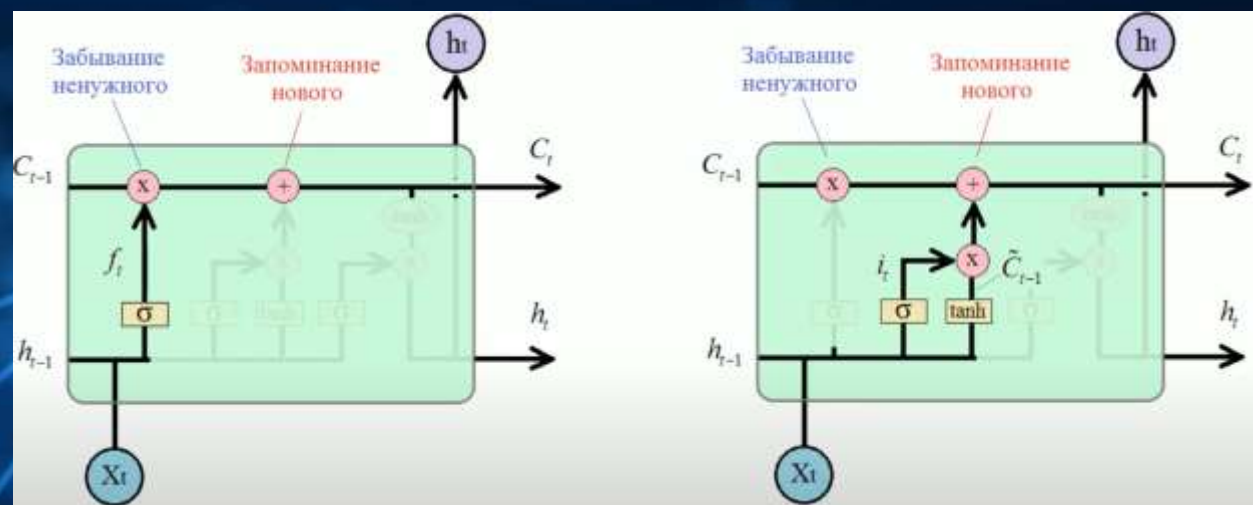
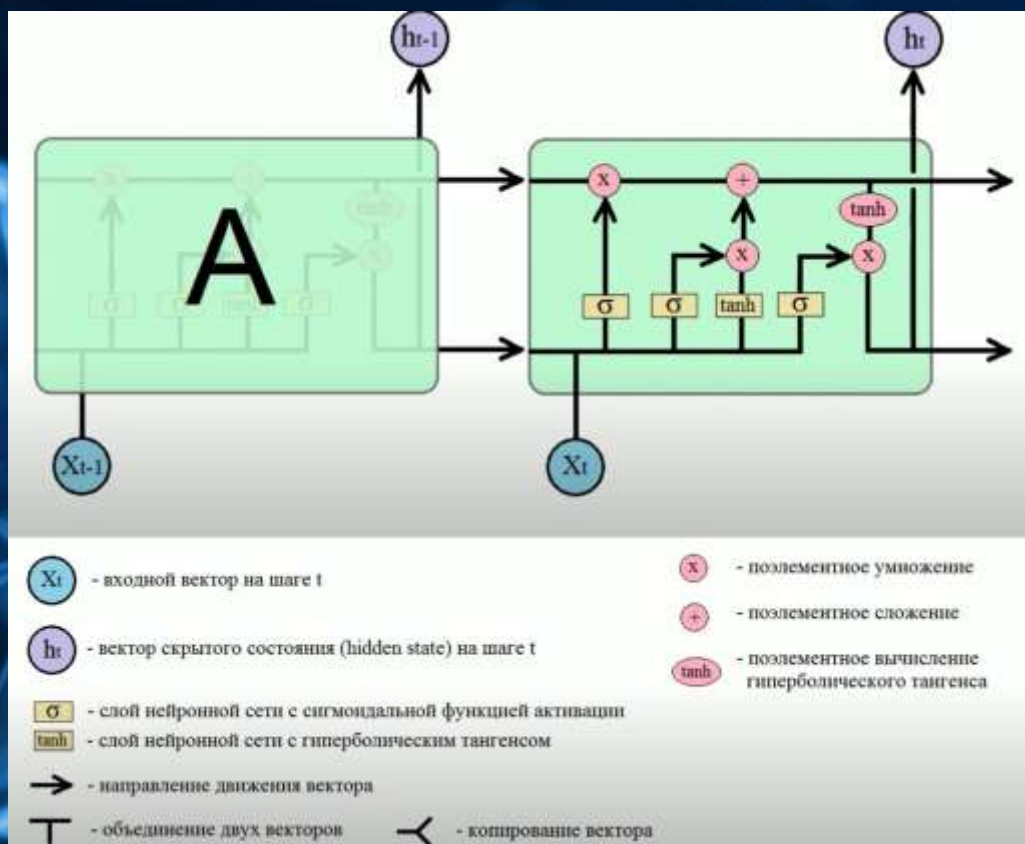


Это сети с памятью,
хранящие
информацию о том,
что в ней
происходило в
предыдущие
итерации обучения.



Достигается это
благодаря тому,
что каждый
нейрон в такой
сети имеет связь
самим с собой.

Архитектура LSTM(1997 год)



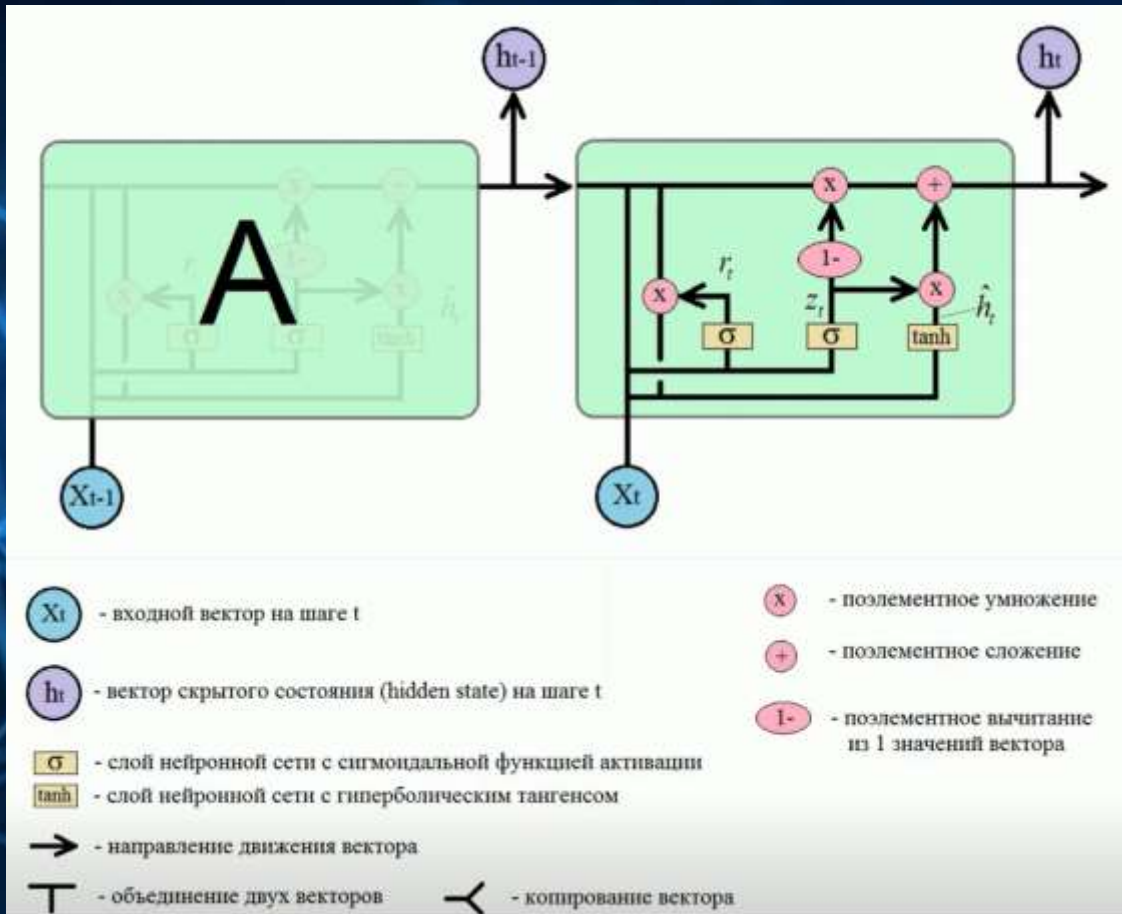
$$f_t = \sigma(W_t * [h_{t-1}, X_t] + b_f), f_t \in [0; 1].$$

$$i_t = \sigma(W_p * [h_{t-1}, X_t] + b_i)$$

$$\tilde{C}_{t-1} = \tanh(W_C * [h_{t-1}, X_t] + b_C)$$

Главным отличием данной архитектуры от простой рекуррентной сети является наличие верхнего канала, способного сохранять важный долгосрочный контекст!

Архитектура GRU(2014 год)



$$z_t = \sigma(W_z * [h_{t-1}, X_t] + b_z), z_t \in [0; 1]$$

$$r_t = \sigma(W_r * [h_{t-1}, X_t] + b_r)$$

$$\hat{h}_t = \tanh(W_{xh} * x_t + W_{hh} * (r_t \otimes h_{t-1}) + b_h)$$

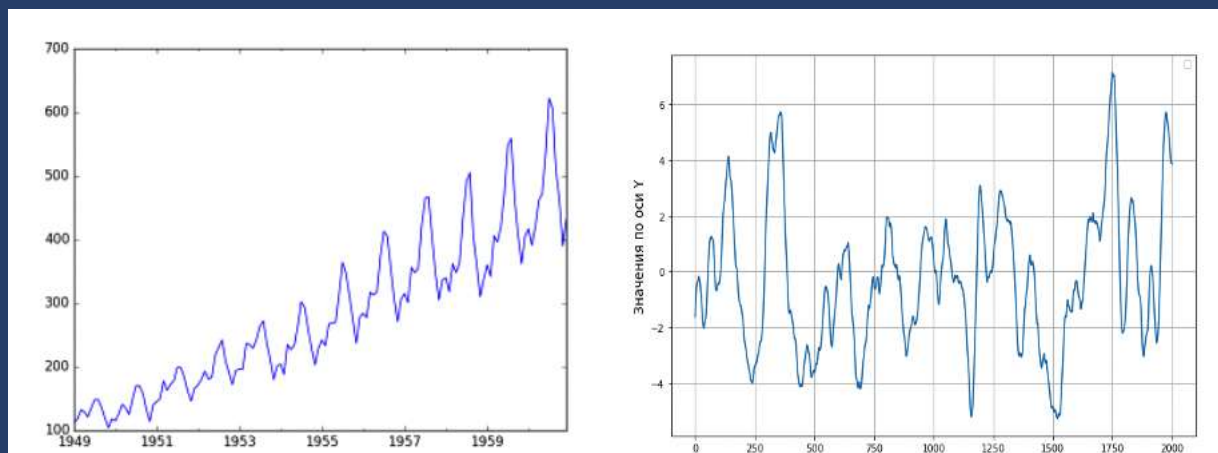
$$h_t = h_{t-1} \otimes (1 - z_t) + \hat{h}_t \otimes z_t$$

С развитием глубокого обучения появилась необходимость в ускорении работы нейронных сетей с сохранением их эффективности. GRU работает быстрее LSTM за счет меньшего числа настраиваемых в ходе обучения параметров, но иногда проигрывает в точности предсказаний.

Временные ряды



Многомерный временной ряд



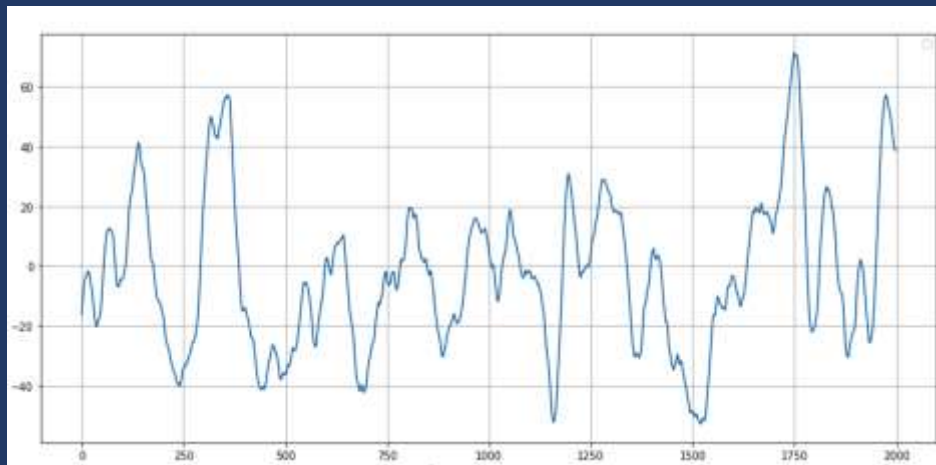
На картинке сверху приведены нестационарный и **стационарный** ряды.

Классические примеры временных рядов:

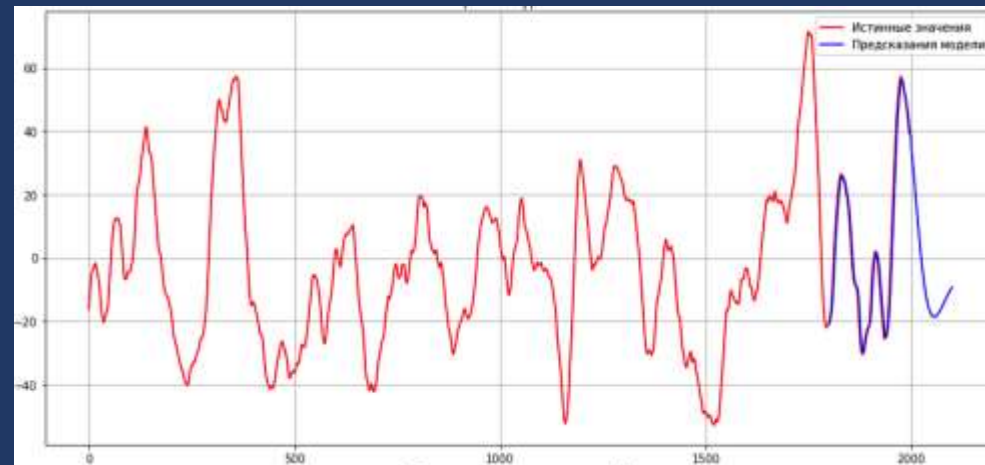
- Курс акций
- Прогноз погоды
- Объем продаж

В данной работе нейронные сети будут применяться для предсказания одномерных стационарных рядов.

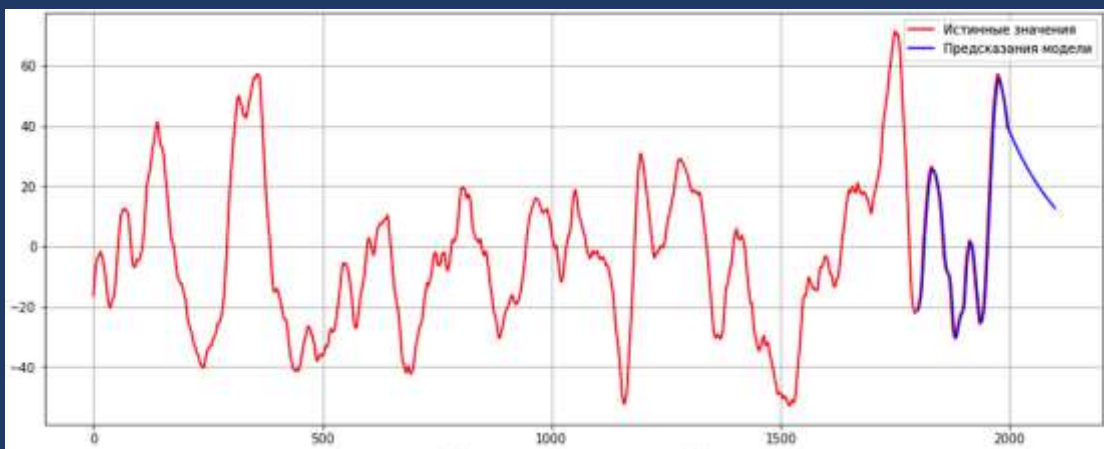
Результаты применения LSTM и GRU моделей



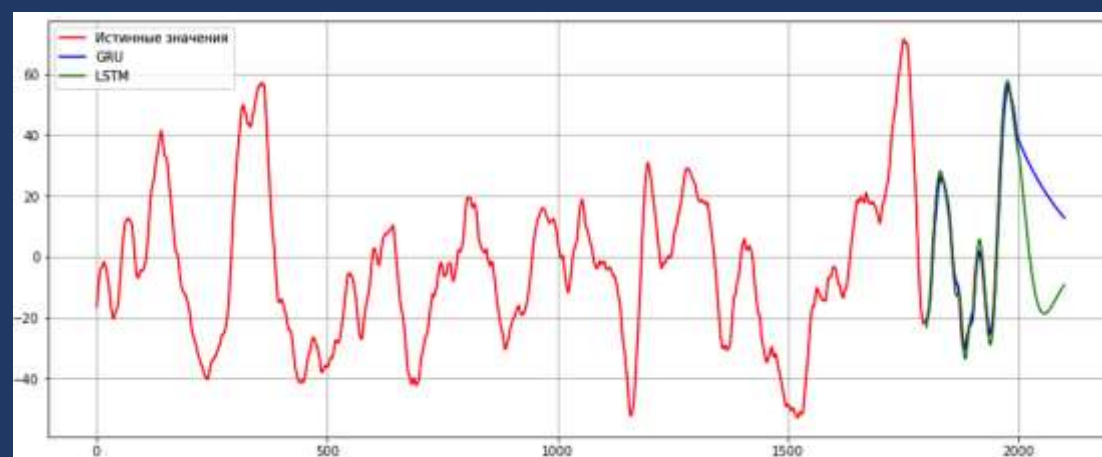
Первый вариант синтетических данных



LSTM модель

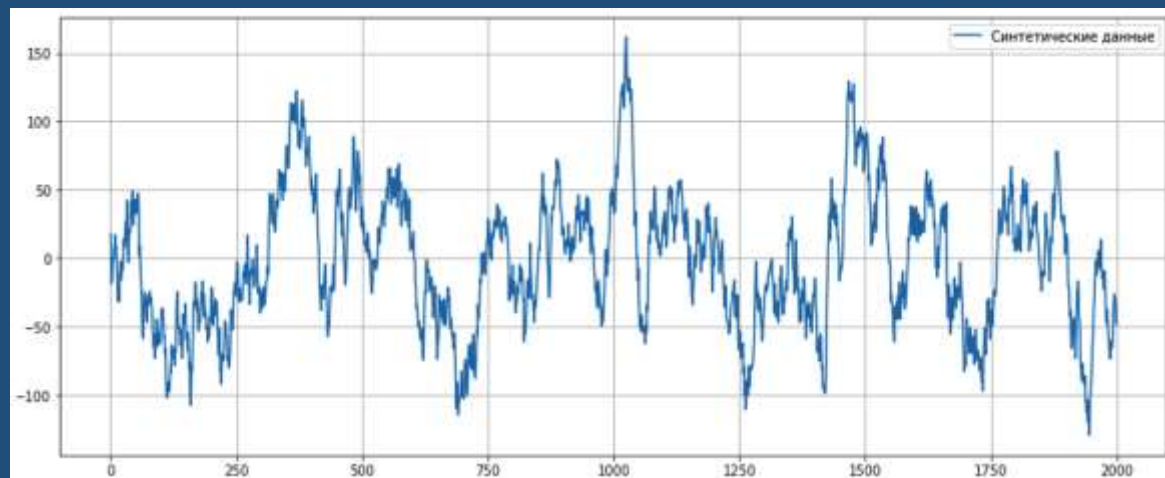


GRU модель

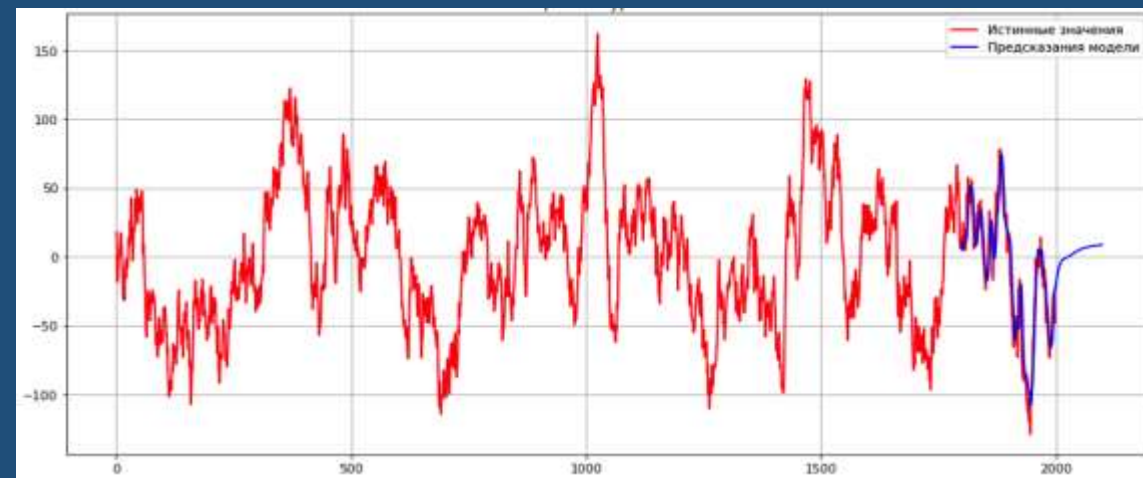


Сравнение LSTM и GRU

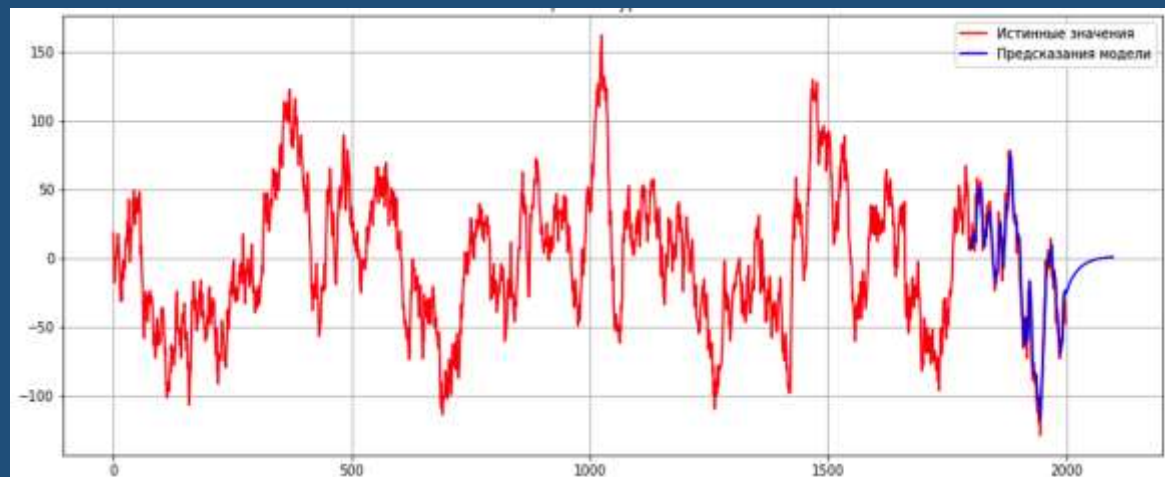
Результаты применения LSTM и GRU моделей



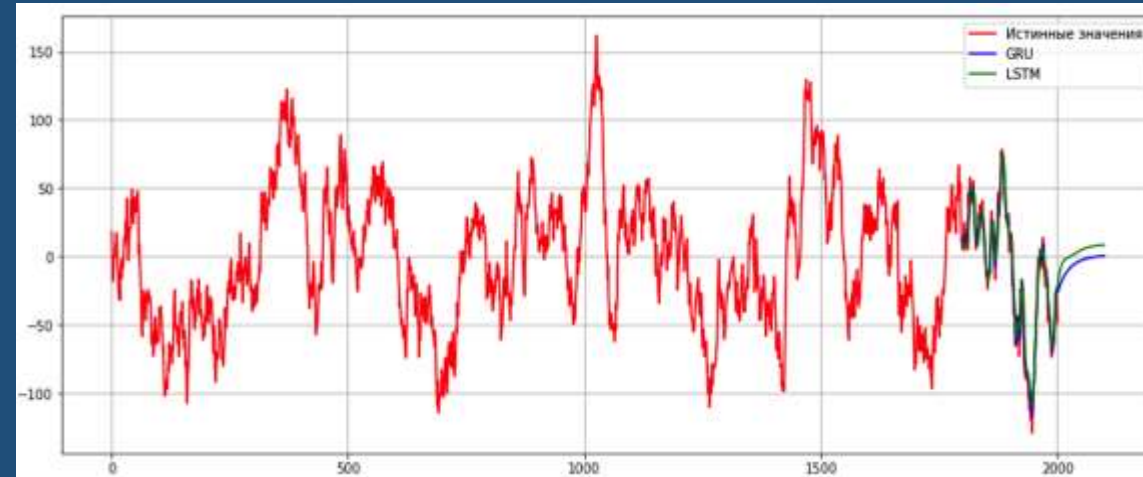
Второй вариант синтетических данных



LSTM модель



GRU модель



Сравнение LSTM и GRU

Результаты применения двух моделей

Первый временной ряд			Второй временной ряд		
Модель	MSE	Time	Модель	MSE	Time
LSTM	13.4	26min55sec	LSTM	237	26min32sec
GRU	15.6	25min50sec	GRU	145	25min53sec

Заключение

Действительно, архитектура GRU имеет меньшее число настраиваемых параметров и, соответственно, быстрее обучается по сравнению с LSTM моделью. Точность предсказания зависит от вида временного ряда. Для более плавных графиков лучше подходит LSTM модель.