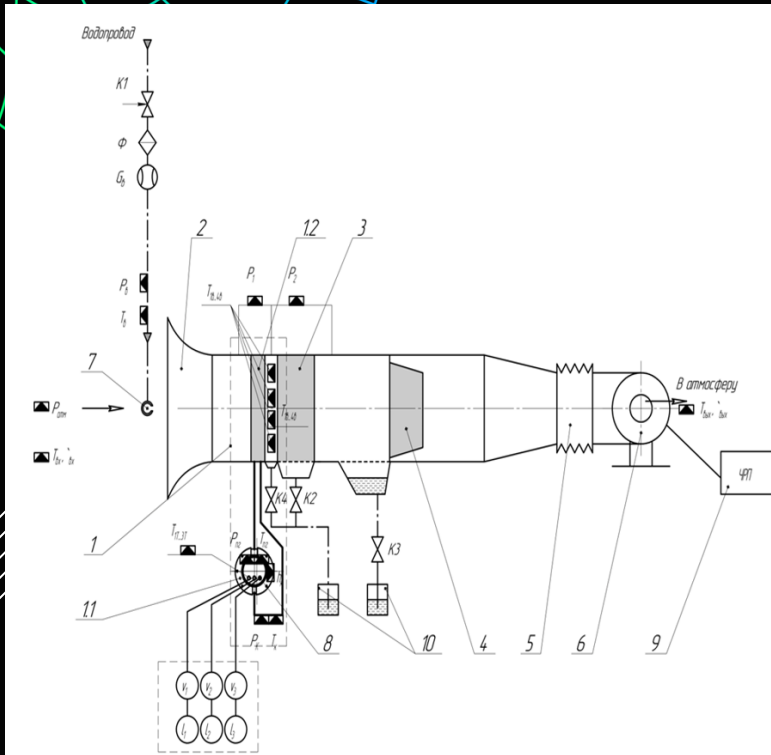


# Разработка модели нейросетей для прогнозирования поведения тепловой установки

---

Выполнил: Лашкевич Д.М.  
Науч.рук.: Корлякова М.О.

# Объект



Вентилятор стенда осуществляет проток атмосферного воздуха через теплообменную поверхность. При наличии расхода воздуха осуществляется подача напряжения на трубчатый электронагреватель. ТЭН, прогреваясь передают тепло фреону и доводят его до кипения. Пар фреона, поднимаясь к трубному пространству теплообменника сообщает тепло потоку воздуха и конденсируется. Конденсат фреона возвращается в парогенератор самотеком.

# Постановка задач



Сформировать модель предсказания теплоэнергетической установки на заданном горизонте прогнозирования

01

## Анализ данных

Провести обзор и предварительный анализ доступных данных о поведении тепловой установки

02

## Выбор модели ML

Проведём сравнительный анализ моделей LSTM и BI-LSTM

03

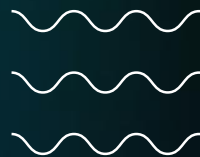
## Качество

Оценить качество модели на отложенных данных и провести анализ точности прогнозов

04

## Анализ результатов

Проанализировать полученные результаты и сделать выводы

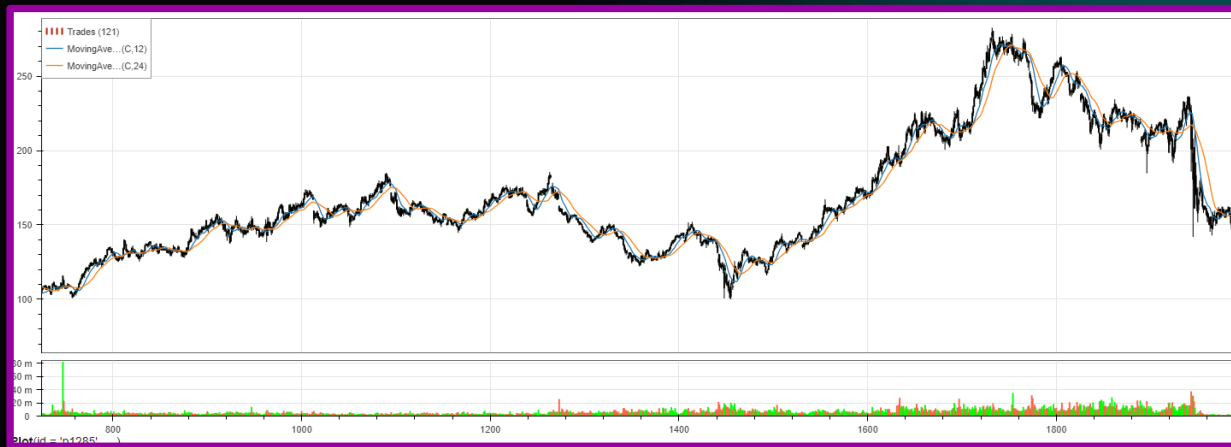
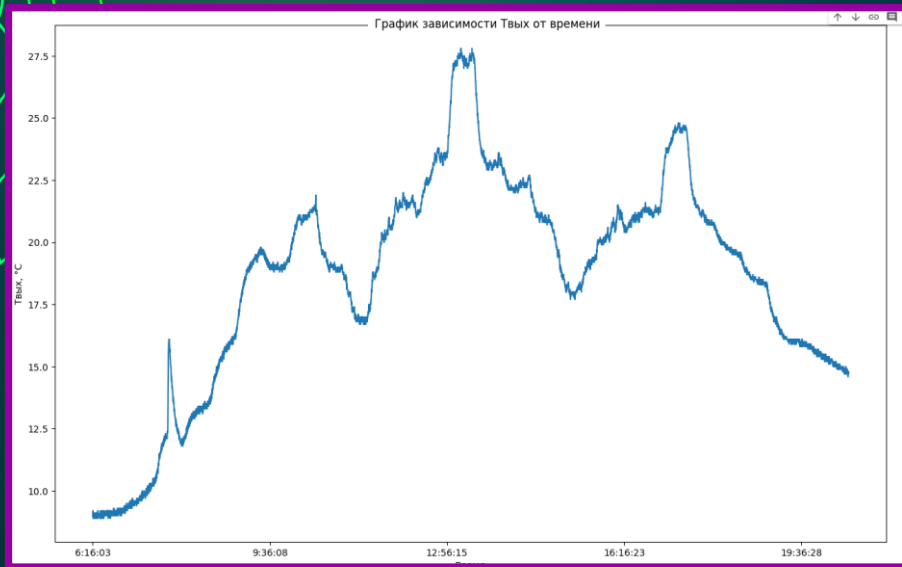


# Временные ряды

$$y = \langle x(t_0), x(t_1), \dots, x(t_n) \rangle$$

$$dt = t_i - t_{i+1}, i = \overline{0, n}$$

$$|x(t_i) - x(t_{i+1})| < e$$



# Методы

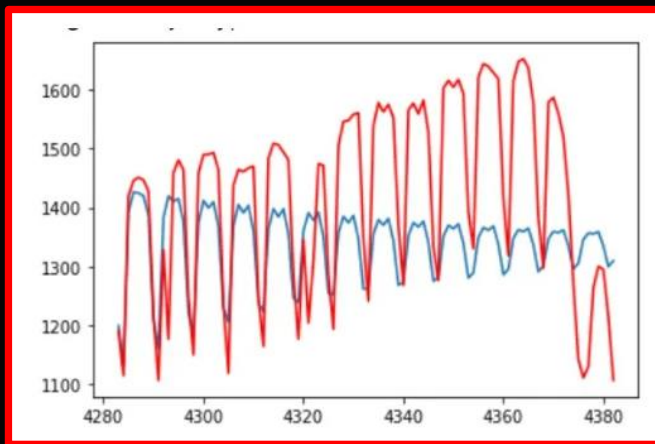
## Метод авторегрессии (AR)

$$X_t = \varphi_0 + \varphi_1 X_{t-1} + \dots + \varphi_i X_{t-i} + \varepsilon_t$$

$X_t$  - текущее значение временного ряда

$\varphi_0 \dots \varphi_i$  - коэффициенты авторегрессии

$\varepsilon_t$  - случайная ошибка



потребность жилых домов в электроэнергии

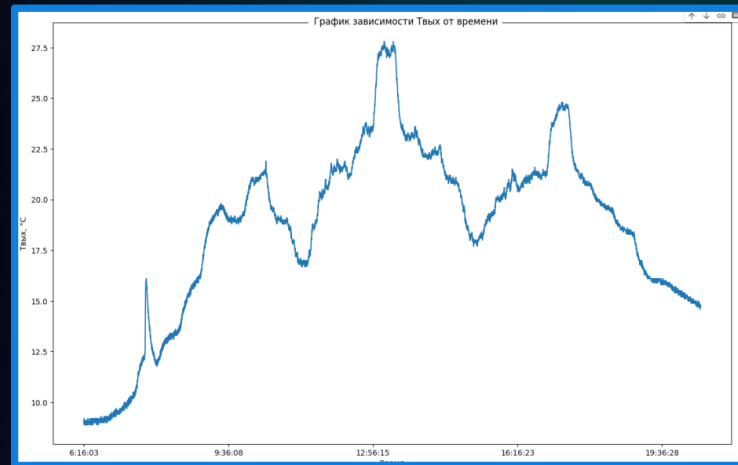
## Метод не авторегрессии (NAR)

$$X_t = f(Z_1, Z_2, \dots, Z_i, \varepsilon_t)$$

$X_t$  - текущее значение временного ряда

$Z_1 \dots Z_i$  - экзогенные переменные или значения других временных рядов

$\varepsilon_t$  - случайная ошибка



# Метрики



## MAPE (Mean Absolute Percentage Error)

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| * 100\%$$

n - количество наблюдателей

$y_i$  - фактическое значение для i-го наблюдения

$\hat{y}_i$  - прогнозируемое значение для i-го наблюдения

$\left| \frac{y_i - \hat{y}_i}{y_i} \right| * 100\%$  - абсолютное

отклонение в процентном

выражении между фактическим и прогнозируемым значением

## $R^2$ (коэффициент детерминации)

$$R^2 = 1 - \frac{RSS}{TSS}$$

RSS - сумма квадратов остатков (сумма квадратов разностей между фактическими и прогнозируемыми значениями)

TSS - общая сумма квадратов (сумма квадратов разностей между фактическими значениями и их средним)

# Подзадачи:



## Сбор данных

Необходимо собрать данные о работе тепловой установки



## Предобработка данных

Полученные данные требуется предварительно обработать, включая очистку от выбросов, заполнение пропущенных значений, нормализацию и т.д.



## Выбор модели ML

Выбрать архитектуру нейронной сети, которая лучше всего подходит для прогнозирования временных рядов



## Обучение модели

Обучить и оценить качество модели на отложенных данных и провести анализ точности прогнозов.

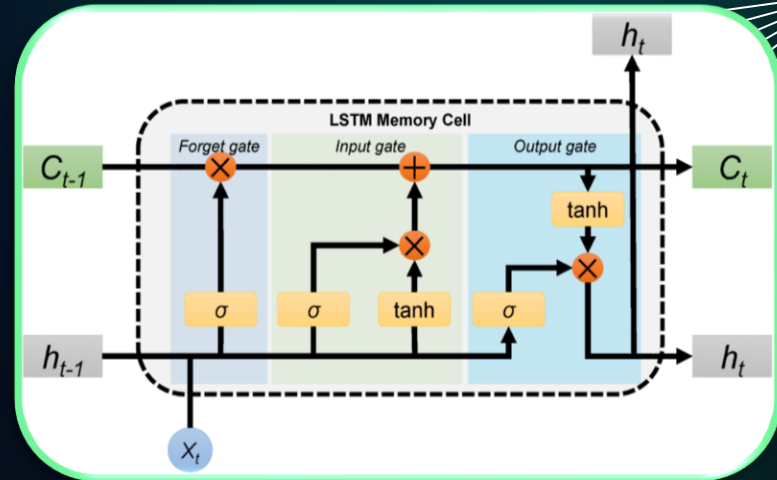


# LSTM

Схема LSTM состоит из нескольких основных компонентов:

- ❖ входного ворота (input gate);
- ❖ забывания (forget gate);
- ❖ выходного ворота (output gate);
- ❖ ячейки памяти (memory cell);

Идея работы LSTM заключается в том, что каждый из этих компонентов позволяет сети принимать решения о том, какую информацию нужно сохранить в памяти и какую информацию передать дальше.





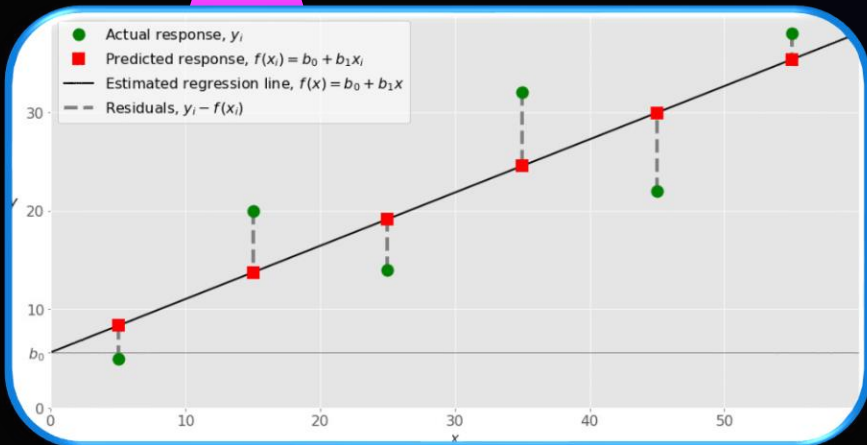
# Linear Regression

−

1. Ограниченность в моделировании сложных зависимостей
2. Чувствительность к выбросам

+

1. Простота и интерпретируемость
2. Эффективность на простых моделях
3. Малое количество параметров



$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n + \varepsilon$$

$Y$  - зависимая переменная

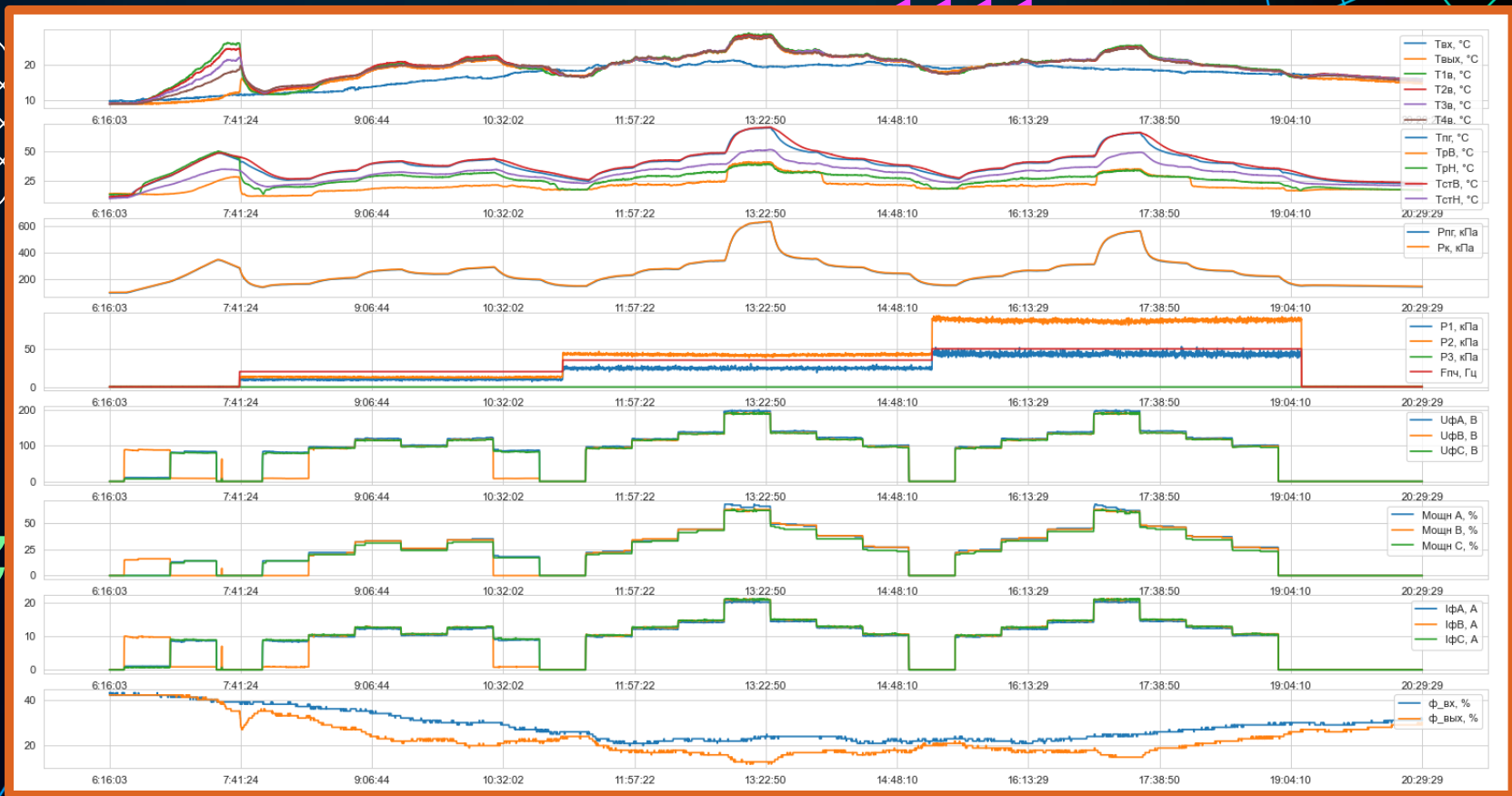
$X_0 \dots X_n$  - независимые переменные

$\beta_0 \dots \beta_n$  - коэффициенты регрессии

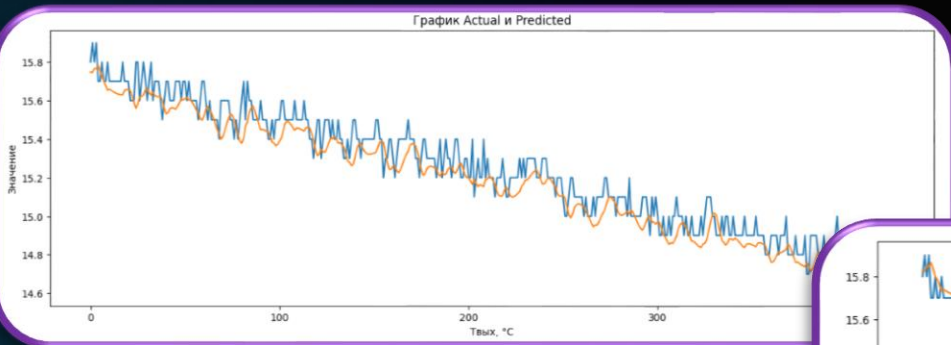
$\varepsilon$  - ошибка



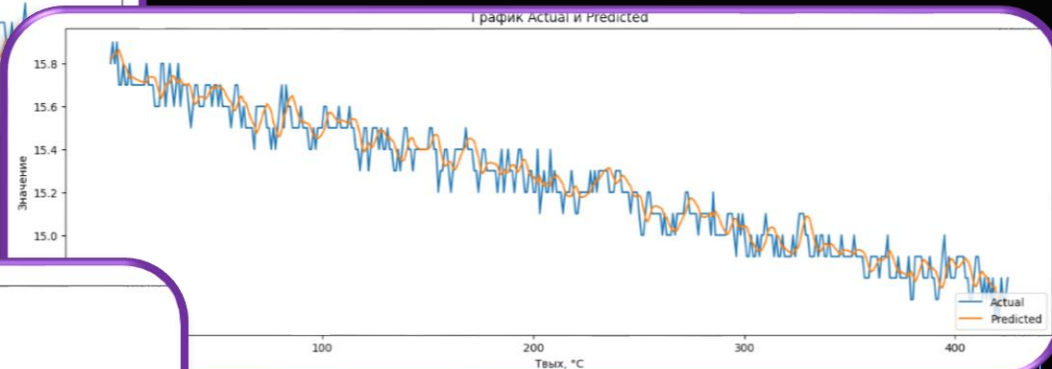
# Данные исследуемого объекта



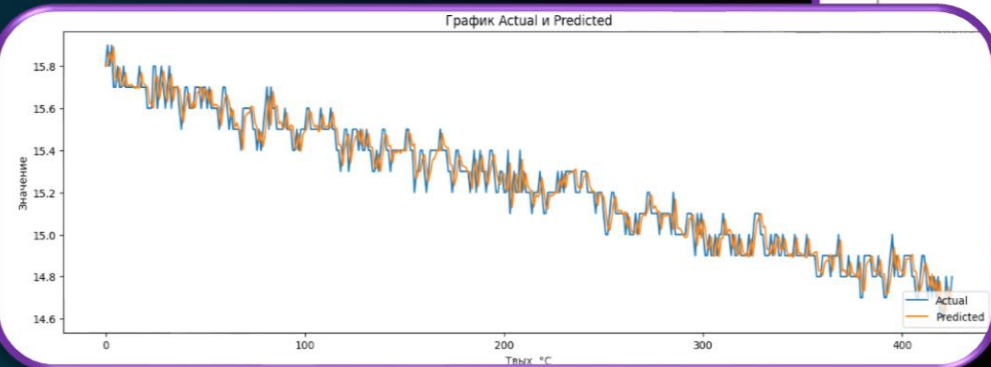
# Прогнозирование данных тепловой установки



LSTM



Bi-LSTM

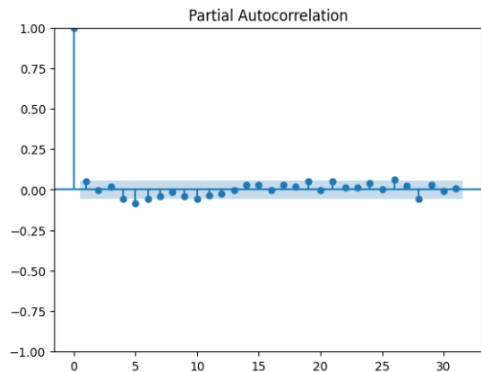


LinearRegression

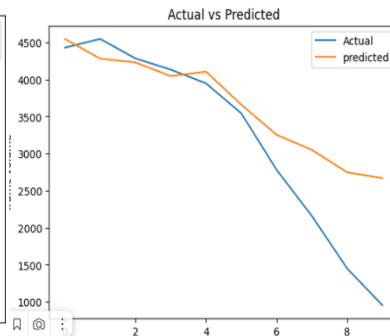
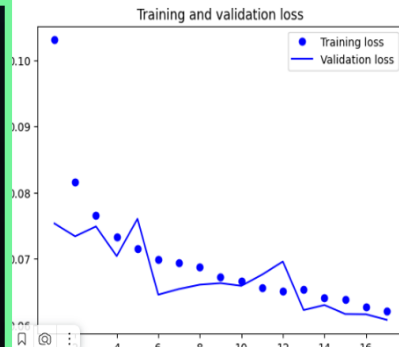


# Сравнительный анализ моделей LSTM и BI-LSTM

## Частичная автокорреляция



BI-LSTM



MSE is : 577626.247288984

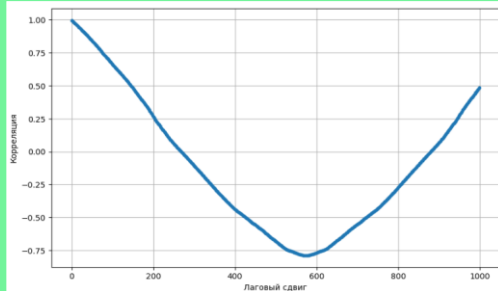
MAE is : 517.5730712890625

RMSE is : 760.0172677571109

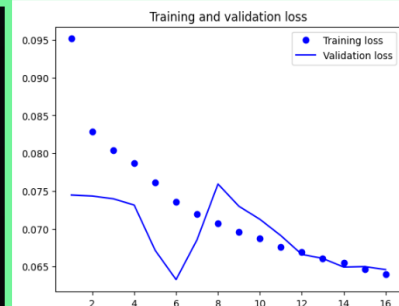
MAPE is : 34.67115162028535

R2 is : 0.6249963458677712

## Корр. на лаговых сдвигах



LSTM



MSE is : 1116790.5618079305

MAE is : 788.28173828125

RMSE is : 1056.7831195699193

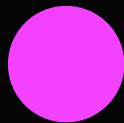
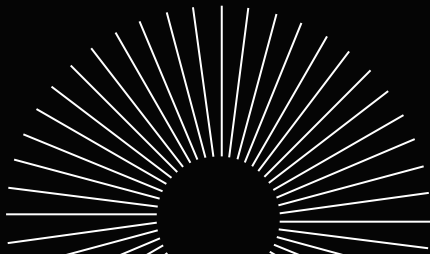
MAPE is : 48.40023892319808

R2 is : 0.2749627574163289

# Заключение



1. Bi-LSTM способна учитывать информацию с обоих направлений, что делает её более эффективной для анализа временных рядов с комплексными зависимостями.
2. LSTM может быть более вычислительно эффективной и простой в реализации для некоторых задач.
3. Линейная регрессия может быть эффективна на небольших наборах данных с простыми зависимостями, но с увеличением объема данных и сложности зависимостей, нейронные сети могут оказаться более эффективными.



Дата сет	Архитектур а сети	Тест: MAPE, R2
Данные за 13.04.23	LinReg	0.409, 0.933
	LSTM	0.518, 0.893
	Bi-LSTM	0.436, 0.929
Данные за 04.04.23, 07.04.23, 13.04.23, 14.04.23	LinReg	0.722, 0.997
	LSTM	0.772, 0.997
	Bi-LSTM	0.752, 0.997

