

Министерство образования Республики Беларусь
Учреждение образования
"Белорусский Государственный университет информатики
и радиоэлектроники"

Лабораторная работа №8
“Рекуррентные нейронные сети для анализа временных рядов”
по учебной дисциплине “Машинное обучение”

Выполнил:

Студент гр. 956241 Дубовик Н.О.

Минск 2020

Данные: Набор данных для прогнозирования временных рядов, который состоит из среднемесячного числа пятен на солнце, наблюдаемых с января 1749 по август 2017. Данные в виде csv-файла можно скачать на сайте Kaggle -> <https://www.kaggle.com/robervalt/sunspots/>

Результат выполнения заданий опишите в отчете.

Задание 1.

Загрузите данные. Изобразите ряд в виде графика. Вычислите основные характеристики временного ряда (сезонность, тренд, автокорреляцию).

На рисунке 1 изображен график ряда, на рисунке 2 был построен и изображен тренд, на рисунке 3 была найдена сезонность, на рисунке 4 была вычислена автокорреляция.

```
plt.figure(figsize=(12,6))
plt.plot(dataset[DATE_COLUMN_NAME], dataset[VALUE_COLUMN_NAME], "-")
plt.xlabel("Year")
plt.ylabel("Value")
plt.grid(True)
```

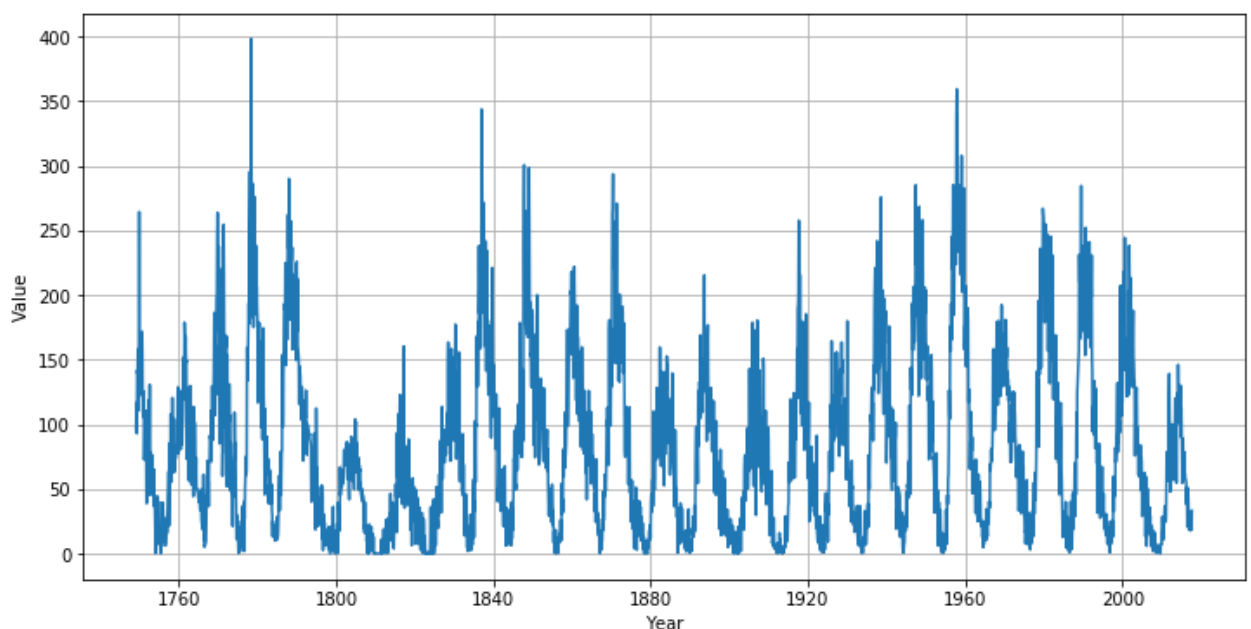


Рисунок 1 – График ряда

```
sunspot_number=dataset[VALUE_COLUMN_NAME]
trend=sunspot_number.rolling(12).mean()
plt.figure(figsize=(12,6))
```

```
plt.plot(dataset[DATE_COLUMN_NAME], trend, "-")
plt.xlabel("Year")
plt.ylabel("Value")
plt.grid(True)
```

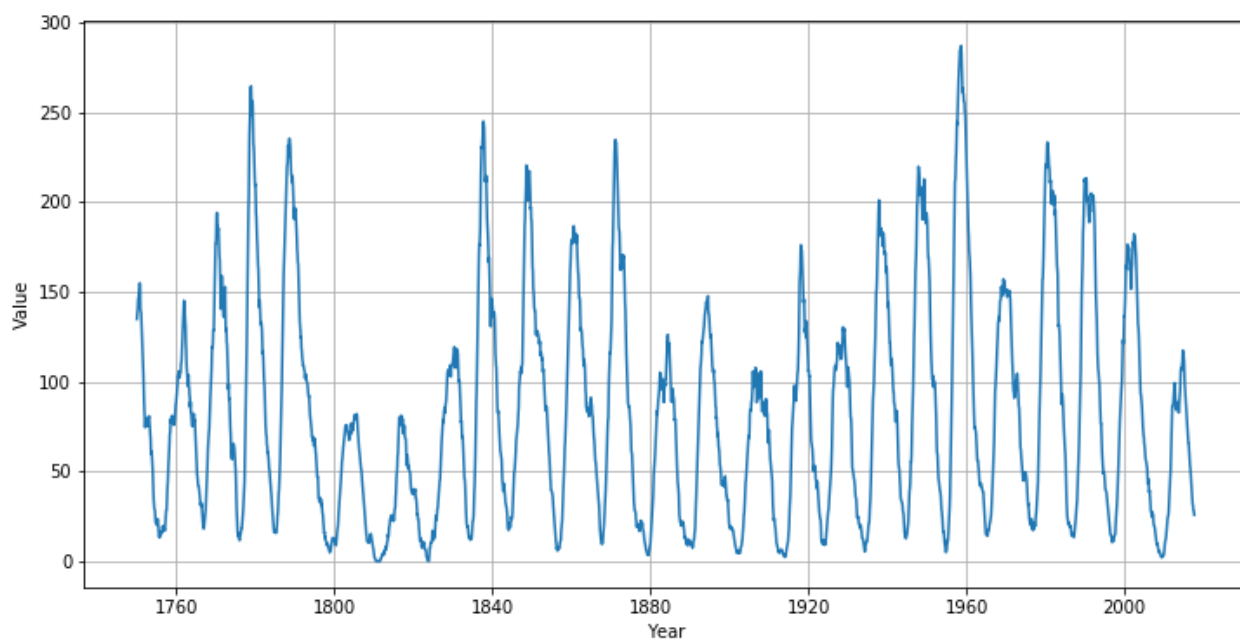


Рисунок 2 – Тренд

```
import statsmodels.tsa.seasonal as seasonal
dataset.index = dataset[DATE_COLUMN_NAME]
decomposition = seasonal.seasonal_decompose(dataset[VALUE_COLUMN_NAME], model='additive')
```

```
plt.figure(figsize=(20,4))
plt.plot(decomposition.seasonal)
plt.ylabel("Value")
plt.xlabel("Year")
```

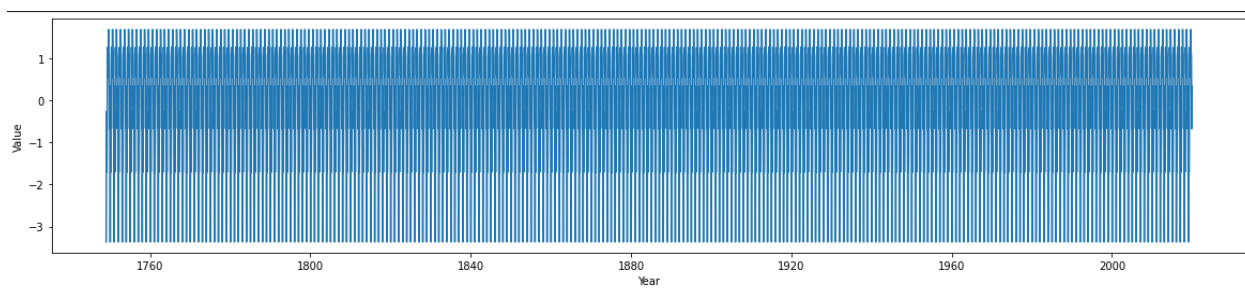


Рисунок 3 – Сезонность

```
plt.figure(figsize=(16,8))
pd.plotting.autocorrelation_plot(dataset[VALUE_COLUMN_NAME])
```

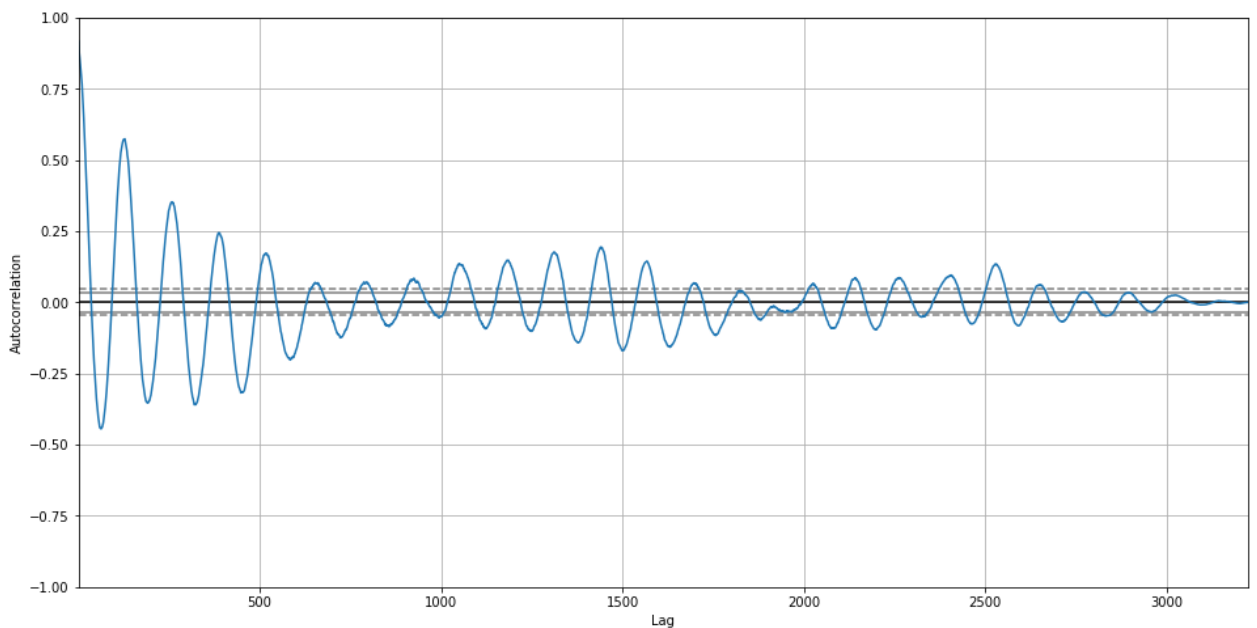


Рисунок 4 – Автокорреляция

Задание 2.

Для прогнозирования разделите временной ряд на обучающую, валидационную и контрольную выборки.

Данные были разделены в отношении 60/20/20.

Задание 3.

Примените модель ARIMA для прогнозирования значений данного временного ряда.

Были применены следующие гиперпараметры при работе с ARIMA, рисунок 5:

Были использованы параметры $p=5$, $d=1$, $q=0$, где

p : Порядок авторегрессии тренда

d : Порядок изменения тренда

q : Тренд скользящей средней

```
model = ARIMA(train_dataset[VALUE_COLUMN_NAME], (9,0,1), dates
=train_dataset[DATE_COLUMN_NAME])
```

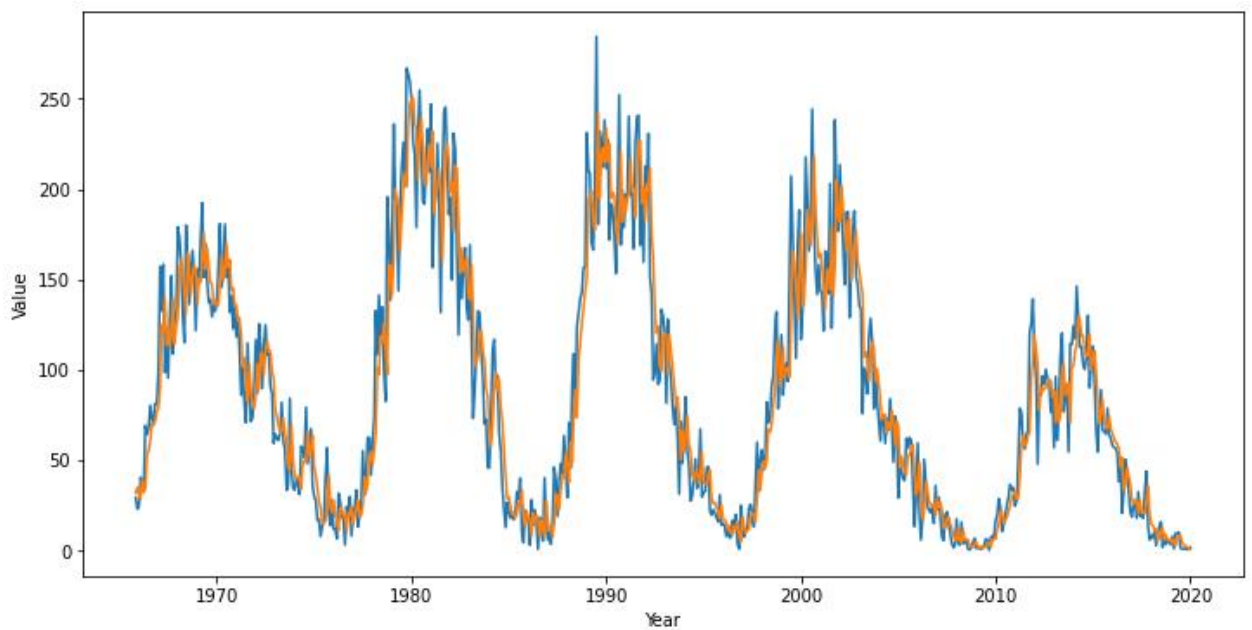


Рисунок 5 – График ряда и прогноза ARIMA

Задание 4.

Повторите эксперимент по прогнозированию, реализовав рекуррентную нейронную сеть (с как минимум 2 рекуррентными слоями).

Была создана сеть, состоящая из 2 слоев LSTM. В выходном слое один нейрон. Также были применены оптимизатор Adam и функция потерь “mae”.

Реализация:

```
model = tf.keras.Sequential([
    tf.keras.layers.LSTM(8, input_shape=x_train.shape[-
2:], return_sequences=True),
    tf.keras.layers.LSTM(8),
    tf.keras.layers.Dense(1)
])
model.compile(optimizer='adam', loss='mae')
```

Результат обучения на рисунках 6 и 7.

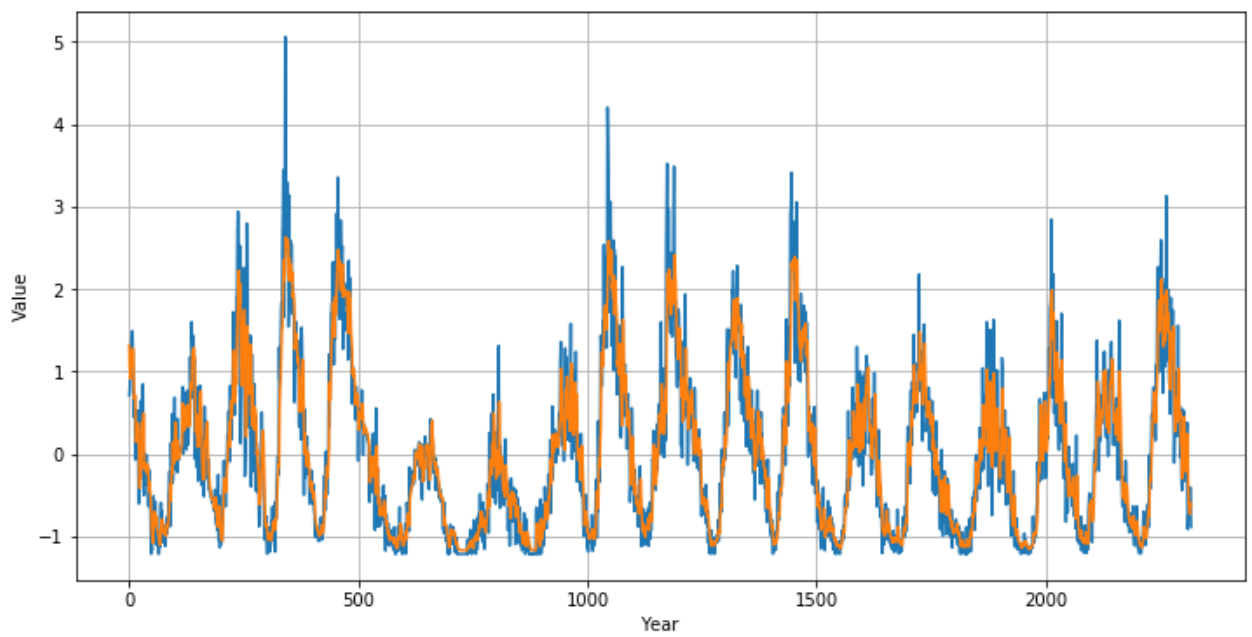


Рисунок 6 – График валидационного ряда и прогноза нейронной сети

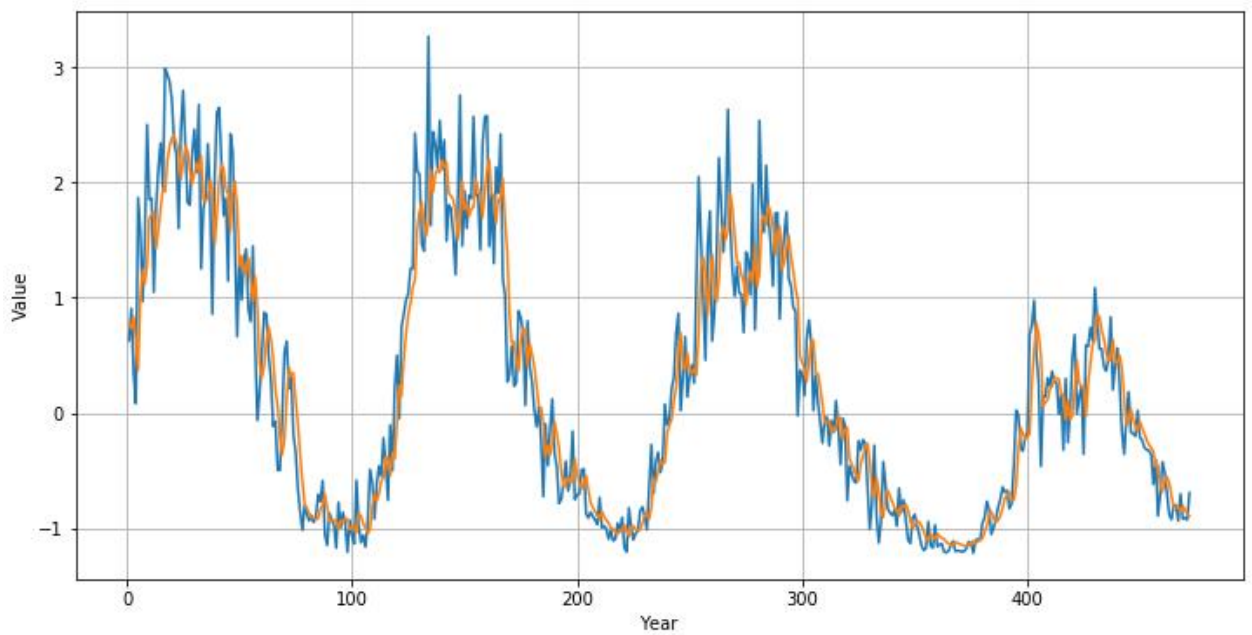


Рисунок 7 – График тестового ряда и прогноза нейронной сети

Задание 5.

Сравните качество прогноза моделей.

Наилучший результат удалось получить на модели ARIMA.