

# ИСПОЛЬЗОВАНИЕ РЕКУРРЕНТНЫХ НЕЙРОСЕТЕЙ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

## USING RECURRENT NEURAL NETWORKS FOR TIME SERIES FORECASTING

V. Perepelkin

*Summary:* Neural networks are widely used in various fields, including medicine, finance, manufacturing and science. They are already successfully used for solving problems of classification, computer vision, video image processing, natural language processing, data and image generation, and many others. In the work under study, the problematic aspects of using recurrent neural networks for forecasting time series are considered. Time series are ubiquitous in the world around us, industry and science. It represents collected historical data, such as the number of people living in a territory, production volumes, fuel consumption volumes, etc. The prediction of future values of time series is an important task in order to be able to prepare in advance for upcoming changes. The construction of a model for predicting the incidence of the COVID-19 virus based on recurrent neural networks is also considered as an example.

*Keywords:* forecasting, time series, neural networks, models.

Предлагаю рассмотреть систему прогнозирования в аспекте суждения о состоятельности будущего процесса или объекта. Система прогнозирования является важнейшим инструментом для принятия решений в областях технологий и науки.

Нейронная сеть — это компьютерная программа, которая работает по принципу человеческого мозга: она пропускает входные данные через систему «нейронов» — более простых программ, взаимодействующих между собой, после чего выдаёт некий результат вычислений на основе этого взаимодействия, учитывая опыт и ошибки прошлых запусков программы. Концепция нейронных сетей была разработана на основе изучения биологических нейронных сетей, которые присутствуют в мозге людей и животных. В процессе работы нейронной сети информация обрабатывается и передается через множество связей между нейронами. Это позволяет нейронной сети обучаться на основе опыта и принимать решения на основе полученных данных. Нейронные сети находят широкое применение в различных областях, включая медицину, финансы, производство и науку. Они используются для анализа и обработки больших объемов данных, определения трендов и прогнозирования будущих событий. Нейронные сети могут обучаться на основе большого количества данных и использовать эту информацию для классификации, распознавания образов, определения паттернов и других задач. Это по-

*Перепелкин Вадим Юрьевич*

*аспирант, Московский государственный университет технологий и управления (Первый казачий университет)  
vip@rambler.ru*

*Аннотация:* Нейронные сети находят широкое применение в различных областях, включая медицину, финансы, производство и науку. Они уже успешно используются в решении задач классификации, компьютерного зрения, работы с видео изображениями, обработки естественных языков, генерации данных и изображений и многих других. В исследуемой работе рассмотрены проблемные аспекты использования рекуррентных нейросетей для прогнозирования временных рядов. Временные ряды повсеместно встречаются в окружающем мире, промышленности и науке. Они представляют собой собранные статистические данные за прошлые периоды, например, количество проживающих людей на какой-либо территории, объемы производства продукции, объемы потребления топлива и т.д. Важной задачей является прогнозирование будущих значений временных рядов, чтобы иметь возможность заранее подготовиться к предстоящим изменениям. В качестве примера рассмотрено построение модели прогнозирования заболеваемости вирусом COVID-19 на базе рекуррентных нейросетей.

*Ключевые слова:* прогнозирование, временные ряды, нейронные сети, модели.

зволяет создавать более эффективные и точные системы, которые могут автоматизировать многие процессы. На текущий момент разработано огромное количество различных типов нейросетей и их архитектур. В зависимости от типа решаемой задачи выбирается та или иная архитектура нейросети. Некоторые типы нейросетей, как например сверточные нейросети, показали свою эффективность в решении задач компьютерного зрения, когда необходим анализ изображения с целью нахождения какого-либо объекта и последующего его распознавания, либо обработки потокового видео с детекцией и сегментацией. Другие типы используются для генерации новых данных, как например, генеративные нейросети, которые способны создавать новые изображения на основании описания, заданного пользователем.

Временной ряд представляет собой в совокупности собранный в разные моменты времени статистический материал о значении каких-либо параметров исследуемого процесса. Вот несколько примеров временных рядов: ежедневные значения температуры воздуха в некотором регионе, объемы потребления электроэнергии, количество произведенной сельскохозяйственной продукции, популяция населения и т.д.

Исторические данные, собранные за некоторый период времени, могут быть использованы для анализа, поиска скрытых закономерностей, корреляции с раз-

личными параметрами, которые могут влиять на измеряемые значения и прогнозирование будущих значений. Прогнозирование является важным процессом, который позволяет строить планы на будущее, готовиться к предстоящим изменениям, и принимать взвешенные решения. Важно отметить, что точно спрогнозировать значения интересующей величины невозможно. Даже банальный прогноз погоды на ближайший день с учетом накопленных исторических данных и сезонности процесса, а также использования специальных технических инструментов, которые позволяют собирать показатели и получать представления об изменениях, наблюдаемых в атмосфере, далеко не всегда бывает точным. Таким образом прогнозируемая величина всегда может отличаться от реально наблюдаемой, и в данном случае важно то, на сколько она будет отличаться, т.е. ошибка прогнозирования. Очевидно, чем больше период времени, на который дается прогноз, тем больше будет ошибка. Нужны требования к допустимому интервалу, в котором может находиться ошибка.

Хорошим примером может послужить компания, которая занимается производством и продажей электроники. Один из ключевых показателей для такой компании — это объем продаж. Для руководства компании было бы полезно прогнозировать объем продаж на основе исторических данных. Для начала нужно собрать данные по продажам за последние несколько лет. Далее нужен некоторый механизм или система, которая позволит на основании собранных данных строить прогнозы на будущие периоды. Например, предположим, что мы хотим прогнозировать объем продаж на следующий год. Мы можем подать исторические данные по продажам в нашу систему, и она выдаст прогноз на основе обнаруженных закономерностей и трендов. Это позволит компании планировать свои бизнес-операции, управлять запасами и ресурсами, и принимать более обоснованные решения.

На текущий момент существует несколько различных способов прогнозирования. Начиная от экспертных оценок, заканчивая математическими моделями, такими как GARCH и ARIMA, которые позволяют путем анализа исторических данных подбирать параметры и строить математические модели, которые могут использоваться для прогнозирования.

Одним из самых современных способов прогнозирования является использование нейросетей. Но использование нейросетей для прогнозирования временных рядов имеет ряд проблем. Классические полносвязные нейросети или сверточные имеют фиксированный размер входного слоя, что не соотносится с тем, временной ряд может быть произвольной длины. Кроме того, указанные выше типы нейросетей не обладают памятью, т.е. результат работы нейросети зависит только от входных

параметров и не учитывает обработанные ранее данные.

Для решения данных проблем обратимся к нейросетям, которые используются в задачах обработки естественного языка нейросетями. При обработке, например, напечатанного текста возникают аналогичные проблемы, т.к., во-первых, длина текста может варьироваться. Во-вторых, слова или предложения, подаваемые на вход, могут быть связаны друг с другом, иначе говоря, каждое слово, которое обрабатывает нейросеть должно быть обрабатывать не отдельно, а в рамках контекста, в котором оно используется, т.е. нейросеть должна обладать памятью.

Общий класс нейросетей, которые используются для обработки естественного языка, называется рекуррентные нейросети (RNN). Рассмотрим их принцип работы.

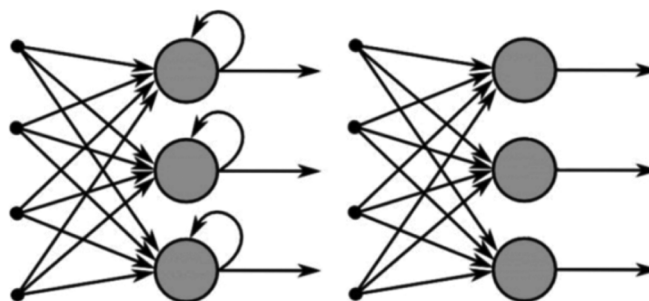


Рис. 1. Отличия полносвязного слоя от рекуррентного

На рисунке 1 показаны отличия рекуррентного слоя от обычного слоя. Если обычный слой получает на вход только значения от предыдущего слоя, то рекуррентный слой получает на вход помимо значений от предыдущего слоя еще и значения от самого себя, сгенерированные на предыдущем этапе. Таким образом, рекуррентный слой имеет внутреннее состояние, которое зависит от ранее обработанных данных, и которое учитывается при формировании выходных данных.

Существует большое количество различных типов рекуррентных сетей, но в общем виде логика их работы может быть выражена формулой:

$$y_t = \sigma_y(W_y h_t + b_t)$$

$$h_t = \sigma_h(W_h X_t + U_h h_{t-1} + b_h)$$

Где:  $X_t$  — вектор входного слоя;  $h_t$  — вектор скрытого слоя;  $y_t$  — вектор выходного слоя;  $W, U, b$  — матрица и вектор обучаемых параметров;  $\sigma_y, \sigma_h$  — функции активации.

Использование рекуррентных нейросетей в обработке естественного языка доказало свою эффектив-

ность. На вход RNN подаются слова из предложения, каждое слово кодируется таким образом, чтобы размер входных данных был фиксированным. При этом, при обработке каждого нового слова нейросеть учитывает уже обработанные ранее слова, т.е. понимает контекст, в котором используется слово.

Аналогичный подход может быть использован и для обработки временных рядов. Мы будем двигаться по временному ряду с окном заданной длины, которое будет подаваться на вход нейросети, следующее значение будет использоваться для обучения сети. Т.е., например, у нас есть ряд из 5000 значений, мы будем идти окном в 10 значений начиная с 0-го элемента, на каждой итерации сдвигая окно на 1 вправо до тех пор, пока не пройдем весь ряд до 5000 элемента. На первой итерации на вход сети будут поданы элементы с 0 по 9, а 10 элемент будет использован для проверки корректности предсказания сети и обратного распространения ошибки для обучения. На второй итерации на вход будут поданы элементы с 1 по 10, а элемент 11 использован для проверки, и т.д. пока не будет обработан весь ряд. При этом на каждой новой итерации нейросеть будет учитывать уже обработанные элементы и таким образом сможет находить закономерности во временном ряду.

Для построения модели будет использован язык программирования Python, который де-факто является стандартом для решения задач машинного обучения и построения моделей, основанных на нейросетях. Также будут использоваться библиотека Pandas, которая позволяет работать с данными, и библиотека PyTorch, которая позволяет создавать и обучать нейросети. Программа будет обрабатывать статистические данные о количестве заболевших вирусом COVID-19. Данные будут разделены на обучающую выборку и проверочную. Для оценки корректности прогнозируемых значений будет использоваться средняя квадратичная ошибка. Модель, обученная на обучающей выборке, будет прогнозировать будущие значения на несколько дней вперед.

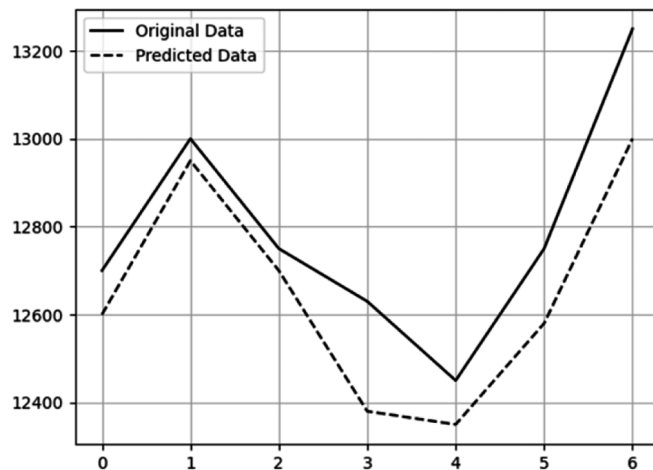


Рис. 2. График реальных и прогнозируемых значений

На рисунке 2 показаны результаты эксперимента применения RNN для прогнозирования количества заболевших вирусом COVID-19. Синяя линия — реальная статистика заболевших. Красная линия — прогнозируемые значения. Как можно убедиться, спрогнозированные значения довольно неплохо соотносятся с реальным, повторяют общий тренд, но безусловно отличаются от реальных.

Таким образом, использование рекуррентных нейронных сетей для прогнозирования временных рядов является очень перспективным направлением в области машинного обучения. Применение RNN позволяет достичь высокой точности прогнозирования и избежать проблем затухания градиента при обучении на длинных последовательностях. В научной литературе также можно найти информацию о применении RNN для решения аналогичных задач, в которых также подтверждается эффективность применения RNN сетей для решений задачи данного класса в сравнении с использованием других подходов.

#### ЛИТЕРАТУРА

1. Омеляненко, Я. Эволюционные нейросети на языке Python: практическое руководство / Я. Омеляненко; пер. с англ. В.С. Яценкова. — Москва: ДМК Пресс, 2020. — 310 с. — ISBN 978-5-97060-854-8. — Текст: электронный. — URL: <https://znanium.com/catalog/product/1210699> (дата обращения: 13.06.2023).
2. Альсова, О.К. Исследование временных рядов в среде R: учебное пособие / О.К. Альсова. — Новосибирск: Изд-во НГТУ, 2021. — 88 с. — ISBN 978-5-7782-4337-8. — Текст: электронный. — URL: <https://znanium.com/catalog/product/1866886> (дата обращения: 14.06.2023).
3. Воейко, О.А. Анализ временных рядов и прогнозирование: практикум / О.А. Воейко. — Москва; Берлин: Директ-Медиа, 2019. — 175 с. — ISBN 978-5-4499-0178-1. — Текст: электронный. — URL: <https://znanium.com/catalog/product/1873514> (дата обращения: 14.06.2023).
4. Ларионова, И.А. Статистика. Анализ временных рядов: учебное пособие / И.А. Ларионова. — Москва: ИД МИСиС, 2001. — 73 с. — Текст: электронный. — URL: <https://znanium.com/catalog/product/1281162> (дата обращения: 14.06.2023).
5. Виноградская, Н.А. Управление производством: методы экономического прогнозирования и планирования: практикум / Н.А. Виноградская, Е.Н. Елисева, О.О. Скрыбин. — Москва: Изд. Дом МИСиС, 2013. — 96 с. — ISBN 978-5-87623-687-6. — Текст: электронный. — URL: <https://znanium.com/catalog/product/1240023> (дата обращения: 14.06.2023).

© Перепелкин Вадим Юрьевич (vup@rambler.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»