

УДК: 621.771.013: 621.314.222.6

10.33942/sit042223

## УЗАК МӨӨНӨТТҮҮ ЭС ТУТУМДУН НЕГИЗИНДЕ КАЙТАЛАНУУЧУ НЕЙРОНДУК ТАРМАКТЫ КОЛДОНУУ МЕНЕН ИНТЕЛЛЕКТУАЛДЫК ТАРМАКТАГЫ ЭЛЕКТР ЖҮКТӨМҮН ПРОГНОЗДОО

**Фролова Татьяна Арнольдовна** магистрант, «Электр менен жабдуу» кафедрасы, И. Раззаков атындагы Кыргыз мамлекеттик техникалык университети, Кыргыз Республикасы, Бишкек ш., Ч. Айтматов проспекти 66, e-mail: [goludo@mail.ru](mailto:goludo@mail.ru)

**Гасанова Дарья Александровна** магистрант, «Электр менен жабдуу» кафедрасы, И. Раззаков атындагы Кыргыз мамлекеттик техникалык университети, Кыргыз Республикасы, Бишкек ш., Ч. Айтматов проспекти 66, e-mail: [angel\\_777.23@mail.ru](mailto:angel_777.23@mail.ru)

Электрдик жүктү болжолдоо акылдуу тармакта маанилүү роль ойнойт. Кыска мөөнөттүү электрдик жүктү болжолдоо жүктүн сааттарын же жумаларын алдын ала айтат. Электрдик жүктөрдүн убакыт катарларынын сызыктуу эмес, стационардык эмес жана сезондук эмес мүнөзүнөн улам так болжолдоо кыйынга турат. Бул макалада бул көйгөйдү чечүү үчүн узак мөөнөттүү эс тутумга (LSTM) негизделген кайталануучу нейрон тармагы (RNN) изилденет. LSTM негизиндеги RNN так болжолдоолорду жасоо үчүн электрдик жүктөө убакыт сериясындагы узак мөөнөттүү көз карандылыктарды колдоно алат. LSTM негизиндеги RNN узак прогноздук горизонт менен татаал электрдик жүктөмдөрдүн убакыт серияларын так алдын ала айтууга жөндөмдүү экенин көрсөтүү үчүн эксперименттер жүргүзүлдү. Анын иштеши көптөгөн башка болжолдоо ыкмалары менен жакшыраак салыштырылат.

**Негизги сөздөр:** электрдик жүктү болжолдоо, бир өлчөмдүү убакыт сериясы, акылдуу тармак, кайталануучу нейрондук тармак (RNN), узак мөөнөттүү кыска мөөнөттүү эс тутум (LSTM).

## ПРОГНОЗИРОВАНИЕ ЭЛЕКТРИЧЕСКОЙ НАГРУЗКИ В ИНТЕЛЛЕКТУАЛЬНОЙ СЕТИ С ИСПОЛЬЗОВАНИЕМ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ НА ОСНОВЕ ДОЛГОВРЕМЕННОЙ ПАМЯТИ

**Фролова Татьяна Арнольдовна** магистрант кафедры «Электроснабжение» Кыргызского Государственного Технического Университета им. И. Раззакова, Кыргызская Республика, г. Бишкек, пр. Ч. Айтматова, 66, г. Бишкек, Кыргызстан, e-mail: [goludo@mail.ru](mailto:goludo@mail.ru)

**Гасанова Дарья Александровна** магистрант кафедры «Электроснабжение» Кыргызского Государственного Технического Университета им. И. Раззакова, Кыргызская Республика, г. Бишкек, пр. Ч. Айтматова, 66, г. Бишкек, Кыргызстан, e-mail: [angel\\_777.23@mail.ru](mailto:angel_777.23@mail.ru)

Прогнозирование электрической нагрузки играет жизненно важную роль в интеллектуальной сети. Краткосрочное прогнозирование электрической нагрузки прогнозирует нагрузку на несколько часов или несколько недель вперед. Из-за нелинейного, нестационарного и несезонного характера временных рядов электрических нагрузок точное прогнозирование затруднено. В этой статье исследуется рекуррентная нейронная сеть (RNN) на основе долговременной памяти (LSTM) для решения этой проблемы. RNN на основе LSTM может использовать долгосрочные зависимости во временных рядах электрических нагрузок для более точного прогнозирования. Проведены эксперименты, чтобы продемонстрировать, что RNN на основе LSTM способна точно прогнозировать временные ряды сложной электрической нагрузки с длинным горизонтом прогнозирования. Его производительность выгодно отличается от многих других методов прогнозирования.

**Ключевые слова:** прогнозирование электрической нагрузки, одномерные временные ряды, интеллектуальная сеть, рекуррентная нейронная сеть (RNN), долговременная кратковременная память (LSTM).

## ELECTRIC LOAD FORECASTING IN SMART GRID USING LONG- SHORT-TERM-MEMORY BASED RECURRENT NEURAL NETWORK

Frolova Tatjana Arnoldovna *master of the department "Power supply" of the Kyrgyz State Technical University I. Razzakov, Bishkek, Kyrgyzstan, e-mail: [goludo@mail.ru](mailto:goludo@mail.ru)*  
Gasanova Daria Aeksandrovna *master of the department "Power supply" of the Kyrgyz State Technical University I. Razzakov, Bishkek, Kyrgyzstan, e-mail: [angel.777.23@mail.ru](mailto:angel.777.23@mail.ru)*

*Electric load forecasting plays a vital role in smart grid. Short term electric load forecasting predicts the load that is several hours to several weeks ahead. Due to the nonlinear, nonstationary and nonseasonal nature of the electric load time series, accurate forecasting is challenging. This paper explores Long-Short-Term-Memory (LSTM) based Recurrent Neural Network (RNN) to deal with this challenge. LSTM-based RNN is able to exploit the long term dependencies in the electric load time series for more accurate forecasting. Experiments are conducted to demonstrate that LSTM-based RNN is capable of forecasting accurately the complex electric load time series with a long forecasting horizon. Its performance compares favorably to many other forecasting methods.*

**Key words:** *Electric load forecasting, univariate time series, smart grid, recurrent neural network (RNN), long-short-term memory (LSTM)*

## I. ВВЕДЕНИЕ

По сравнению с традиционной электросетью, интеллектуальная сеть способна обеспечить более интеллектуальную, эффективную, устойчивую и надежную электроэнергию за счет использования передовой инфраструктуры и информационных технологий. Прогнозирование электрической нагрузки играет все более важную роль в интеллектуальных сетях. Для поставщиков коммунальных услуг принципиально важно заранее моделировать и прогнозировать энергетические нагрузки, находить баланс между производством и спросом, снижать себестоимость продукции и внедрять различные схемы ценообразования для реагирования на спрос.

В зависимости от продолжительности горизонта прогнозирования прогнозирование электрической нагрузки подразделяется на четыре категории: долгосрочное прогнозирование нагрузки, среднесрочное прогнозирование нагрузки, краткосрочное прогнозирование нагрузки и сверхкраткосрочное прогнозирование нагрузки [1]. Краткосрочное прогнозирование нагрузки прогнозирует нагрузку на несколько часов или несколько недель вперед на основе наблюдаемых данных временного ряда нагрузки [2].

Поскольку электрическая нагрузка представляет собой в первую очередь одномерный временной ряд [3], для прогнозирования электрической нагрузки можно применять многие общие методы прогнозирования временных рядов. Общий класс методов составляют модели статистического прогнозирования, включающие модели авторегрессии (AR), скользящего среднего (MA), авторегрессионного интегрированного скользящего среднего (ARIMA) [4] [5] и ряд их вариантов [6]. В частности, ARIMA является одним из самых популярных и часто используемых методов прогнозирования временных рядов. Однако эти методы работают в предположении, что наблюдаемые временные ряды и будущие временные ряды связаны линейно, что делает их менее эффективными для временных рядов со значительными нелинейными характеристиками. Было проделано много работы по распространению их приложений на нелинейное прогнозирование, например модель авторегрессионной условной гетероскедастики (GARCH) [7]. Тем не менее, эти статистические модели имеют хорошие характеристики прогнозирования только для стационарных данных [8] [9], в то время как временные ряды электрической нагрузки могут быть нестационарными.

Другой класс методов прогнозирования основан на искусственных нейронных сетях (ИНС). В последнее десятилетие ИНС стали чрезвычайно популярными в прогнозировании электрических нагрузок. По сути, ИНС имитируют человеческий мозг, автоматически изучая закономерности и шаблоны из прошлого опыта и получая обобщенные результаты. В отличие от методов линейного прогнозирования на основе ARIMA, ИНС представляют собой набор нелинейных самоадаптирующихся методов, которые управляются данными, а это означает, что нет необходимости в каких-либо предварительных знаниях о взаимосвязи между моделями и переменными данными. Хорошо известно, что ИНС способны

аппроксимировать любую нелинейную функцию. ИНС обычно могут достигать разумных результатов, особенно для сложных моделей и временных рядов [10]. Существует обширная литература по использованию ИНС для прогнозирования электрических нагрузок, таких как многослойный персептрон с прямой связью (MLP) [11], нелинейные авторегрессионные модели с экзогенными входными данными (NARX), нейронная сеть [12], обобщенная регрессионная нейронная сеть (GRNN) [11]. 13], поддержка векторной регрессии (SVR) [14] и др.

Несмотря на то, что были проведены обширные исследования, точное прогнозирование электрической нагрузки остается проблемой для интеллектуальных сетей. Прогнозирование электрической нагрузки обычно представляет собой задачу прогнозирования одномерных временных рядов, которая является более сложной, чем соответствующая задача прогнозирования многомерных временных рядов. Потому что нет дополнительной информации из других источников данных, которую можно использовать для обучения [15]. Кроме того, по сравнению с линейными, стационарными и сезонными временными рядами, временные ряды электрических нагрузок являются нелинейными, нестационарными и несезонными, где несезонность означает отсутствие явной периодичности во времени. Трудно точно прогнозировать такие временные ряды в долгосрочной перспективе. Поэтому необходимы дополнительные усилия для разработки более эффективных методов прогнозирования.

В этой статье мы решаем эту проблему, используя Long Short-Term-Memory (LSTM) [16], которая представляет собой специальную архитектуру рекуррентной нейронной сети (RNN), которую можно использовать для более точного изучения временных последовательностей и долгосрочных зависимостей, чем Деер. Нейронные сети (DNN) и обычные RNN [17]. Мы разработаем новую схему прогнозирования электрической нагрузки на основе LSTM. Эта схема способна точно прогнозировать сложные нелинейные, нестационарные и несезонные одномерные временные ряды электрических нагрузок на длительный период прогнозирования. Оставшаяся часть этой статьи организована следующим образом. Задача прогнозирования одномерных временных рядов представлена в разделе II. В разделе III представлена схема прогнозирования электрической нагрузки с помощью RNN на основе LSTM. Эксперименты и выводы приведены в разделах IV и V соответственно.

## II. ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ НА НЕСКОЛЬКО ШАГОВ ВПЕРЕД

Рассматривая одномерный временной ряд электрической нагрузки с  $N$  наблюдениями  $\{X_{t1}, X_{t2}, \dots, X_{tN}\}$ , задача многошагового прогнозирования состоит в том, чтобы использовать эти  $N$  зарегистрированных точек данных для прогнозирования следующих  $N$  точек данных  $\{X_{tN+1}, X_{tN+2}, \dots, X_{tN+N}\}$  в будущем существующего временного ряда. Параметр  $N > 1$  является горизонтом прогнозирования. Даже при краткосрочном прогнозировании электрической нагрузки горизонт прогнозирования может быть очень большим, т. е.  $N \gg 1$ . Как правило, данные об электрической нагрузке интеллектуальной сети получают с помощью интеллектуальных счетчиков или PMU (блока измерения мощности). Если интеллектуальные счетчики имеют интервал выборки 15 минут, горизонт прогнозирования составляет  $N = 96$  для прогнозирования нагрузки на 24 часа (на сутки вперед). PMU имеет еще более высокую частоту дискретизации с интервалом дискретизации в доли секунды, а горизонт прогнозирования может быть очень длинным.

### *А. Рекурсивная стратегия*

Существуют три стратегии, которые обычно используются для прогнозирования временных рядов на несколько шагов вперед, т. е. рекурсивная стратегия, прямая стратегия и стратегия с несколькими входами и несколькими выходами (MIMO) [18]. Наиболее интуитивно понятной и традиционной стратегией прогнозирования является рекурсивная

стратегия [19], в которой применяется метод прогнозирования временных рядов на один шаг вперед с одной моделью прогнозирования  $f(\cdot)$ . Конкретно,

$$X_{t+1} = f(X_t, X_{t-1}, \dots, X_{t-d+1}) + \epsilon, \quad (1)$$

где  $t \in \{d, d + 1, \dots, N\}$ ,  $d$  размерность оценки, аддитивный шум, оценка  $f : \mathbf{R}^d \rightarrow \mathbf{R}$  и  $\mathbf{R}$  обозначает реальное поле. Чтобы спрогнозировать  $H$  шагов вперед, мы сначала спрогнозируем оценку  $X_{t+N+1}$  на один шаг вперед, используя (1). Затем, с прогнозируемым  $X_{t+N+1}$  как частью входного временного ряда, следующим шагом является оценка  $X_{t+N+2}$  с использованием той же модели прогнозирования на один шаг вперед (1). Эта процедура выполняется рекурсивно, пока мы не оценим  $X_{t+N+H}$ .

Хотя рекурсивная стратегия интуитивно понятна и проста в применении, она чувствительна к накоплению ошибок прогнозирования, особенно при большом горизонте прогнозирования. В рекурсивной стратегии ошибки прогнозирования на предыдущих этапах распространяются и накапливаются, что ухудшает точность последующего прогнозирования [18].

#### Б. Прямая стратегия

Другой стратегией прогнозирования на несколько шагов вперед является прямая стратегия [20]. В отличие от рекурсивной стратегии, прямая стратегия строит  $H$  различных моделей прогнозирования для каждого горизонта прогнозирования на основе наблюдаемых данных временного ряда. Конкретно,

$$X_{t+h} = f_h(X_t, X_{t-1}, \dots, X_{t-d+1}) + \epsilon h, \quad (2)$$

где  $h \in \{1, 2, \dots, H\}$ ,  $f_h$  —  $h$ -я модель прогнозирования, а  $\epsilon h$  — аддитивный шум, связанный с  $h$ -й моделью. Поскольку прямая стратегия не использует прогнозируемое значение в качестве входных данных для прогнозирования, она не склонна к накоплению ошибок. Однако модели прогнозирования  $H$  обучаются отдельно и независимо друг от друга, что может привести к условной независимости между прогнозируемыми значениями  $H$  [21]. Такой эффект независимости не позволит методам прогнозирования отражать статистическую зависимость между прогнозируемыми данными, что ухудшит качество прогнозирования.

#### В. Стратегия с несколькими входами и несколькими выходами (MIMO)

И рекурсивная стратегия, и прямая стратегия рассматриваются как стратегии с одним выходом, потому что они отображают несколько входов (вектор) в один выход (скаляр) [18]. Напротив, стратегия MIMO представляет собой стратегию прогнозирования, которая использует несколько входных данных для создания нескольких выходных данных [22]. При использовании стратегии MIMO результатом прогнозирования является временной ряд (вектор), а не скаляр. Все данные в этом выходном векторе генерируются одной и той же моделью, обученной с использованием одних и тех же наблюдаемых данных временных рядов. Конкретно,

$$\{X_{t+1}, X_{t+2}, \dots, X_{t+H}\} = F(X_t^t, X_{t-1}^t, \dots, X_{t-d+1}^t) + \epsilon, \quad (3)$$

где  $\epsilon$  — вектор шума и  $F: \mathbf{R}^d \rightarrow \mathbf{R}^H$ . По сравнению со стратегиями с одним выходом, стратегия MIMO способна смягчить проблему условной независимости. Его преимущество заключается в сохранении временной статистической зависимости в прогнозируемых временных рядах. Однако, поскольку стратегия MIMO прогнозирует все данные с помощью одной и той же модели прогнозирования, ее гибкость и изменчивость могут быть не такими сильными, как у других стратегий прогнозирования [23].

### III. RNN НА ОСНОВЕ LSTM ДЛЯ ПРОГНОЗИРОВАНИЯ ЭЛЕКТРИЧЕСКОЙ НАГРУЗКИ

#### А. Рекуррентные нейронные сети (RNN)

Многие нейронные сети с прямой связью, такие как MLP (многослойный перцептрон), DNN (глубокая нейронная сеть), CNN (сверточная нейронная сеть) и т. д., достигли самых современных результатов в различных контролируемых или неконтролируемых приложениях машинного обучения. Их успех сильно зависит от предположения о

независимости обучающих и тестовых данных [24]. Когда данные во временном ряду зависят друг от друга или предположение о независимости не выполняется, их эффективность обучения ухудшается из-за их недостаточной способности моделировать долгосрочные зависимости. Прогнозирование временных рядов — это типичный сценарий, в котором текущие точки данных связаны с предыдущими точками данных. Зависимость от длительного времени фактически является основой для прогнозирования временных рядов. Кроме того, нейронные сети с прямой связью ограничивают входные данные и целевые объекты векторами фиксированной длины [25], что также делает их неудобными для обучения последовательностей (например, временных рядов).

Напротив, RNN разработаны специально для работы с последовательными данными или временными рядами [26]. По сравнению с нейронными сетями с прямой связью, которые позволяют сигналам проходить только вперед от входа к выходу, RNN позволяют сигналам проходить как вперед, так и назад. Они создают петли в сети и позволяют устанавливать внутренние соединения между скрытыми устройствами. С помощью таких внутренних связей RNN больше подходят для использования информации из прошлых данных для прогнозирования будущих данных. В частности, RNN позволяют исследовать временные отношения между данными, которые находятся далеко друг от друга [27].

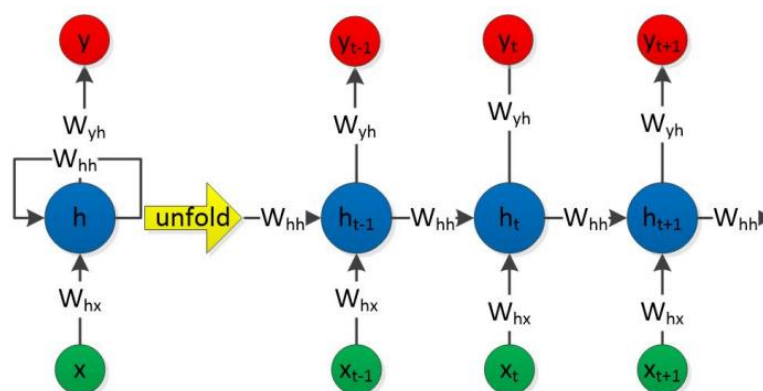


Рис. 1: Архитектура RNN.

На рис. 1 показана архитектура RNN. Учитывая входной временной ряд  $x = \{x_1, x_2, \dots, x_T\}$ , RNN вычисляет скрытую последовательность состояний  $h = \{h_1, h_2, \dots, h_T\}$ , а также выходную последовательность  $y = \{y_1, y_2, \dots, y_T\}$  итеративно, используя следующую систему уравнений

$$h_t = f(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \tag{4}$$

$$y_t = g(W_{yh}h_t + b_y). \tag{5}$$

В (4)-(5)  $W_{hx}$ ,  $W_{hh}$  и  $W_{yh}$  обозначают входную скрытую весовую матрицу, скрытую весовую матрицу и скрытую выходную весовую матрицу соответственно. Векторы  $b_h$  и  $b_y$  представляют смещение скрытого слоя и выходного слоя соответственно. Кроме того,  $f(\cdot)$  и  $g(\cdot)$  — функции активации для скрытого слоя и выходного слоя соответственно. RNN использует скрытое состояние  $h_t$  на временном шаге  $t$  для запоминания сети. Скрытое состояние захватывает всю информацию, включенную в предыдущие временные шаги.

Прогнозирование временных рядов на несколько шагов вперед показывает многоступенчатые зависимости, поскольку прогноз вневыборочных данных  $x_{t+h}$  зависит от входных данных, наблюдаемых в гораздо более ранний момент времени  $t_e$ , где  $t \gg t + h$ . Тем не менее, когда интервал зависимостей данных увеличивается, простые RNN, как правило, все больше страдают от проблемы исчезновения градиента [28]. Другими словами, влияние входных данных при  $t_e$  на прогнозируемые данные  $x_{t+h}$  быстро затухает с течением времени

$t+h-t_e$ . Следовательно, простые RNN могут быть не лучшим выбором для прогнозирования проблем с долгосрочными зависимостями.

*Б. Архитектура LSTM (долговременная кратковременная память)*

LSTM — это эффективная архитектура RNN, представленная Хохрайтером и Шмидхубером в 1997 году [16] и с тех пор усовершенствованная многими людьми [29]. LSTM был в основном мотивирован и разработан для преодоления проблемы исчезающих градиентов стандартной RNN при работе с долгосрочными зависимостями.

В стандартной RNN общая нейронная сеть представляет собой цепочку повторяющихся модулей, сформированных как ряд простых скрытых сетей, таких как один сигмовидный слой. В отличие от стандартной RNN, которая имеет серию повторяющихся модулей с относительно простой структурой, скрытые слои LSTM имеют более сложную структуру. В частности, LSTM вводит концепции вентиля и ячеек памяти в каждом скрытом слое. Блок памяти в основном состоит из четырех частей: входного вентиля  $i$ , вентиля забывания  $f$ , выходного вентиля  $o$  и самоподключенных ячеек памяти  $C$ . Входной вентиль контролирует ввод активаций в ячейку памяти. Выходной вентиль узнает, когда выводить активации в последовательную сеть. Ворота забывания помогают сети забыть прошлые входные данные и сбросить ячейки памяти. Кроме того, осторожно применяются мультипликативные вентили, чтобы ячейки памяти могли получать доступ и хранить информацию в течение длительного интервала времени. Такая структура может эффективно смягчить проблему исчезающего градиента [30]. Это делает LSTM архитектурой, подходящей для проблем с долгосрочными зависимостями.

Поскольку вентили не могут получить какую-либо информацию из вывода ячейки памяти, когда выходной вентиль закрыт, LSTM не знает, какой длины должна быть память для модели. Чтобы решить эту проблему, к ячейкам памяти LSTM можно добавить глазок. Работая в качестве непосредственного наблюдателя, соединения глазка позволяют всем гейтам проверять состояния ячеек [31]. На рис. 2 показана архитектура общего блока памяти LSTM с добавленными глазковыми соединениями.

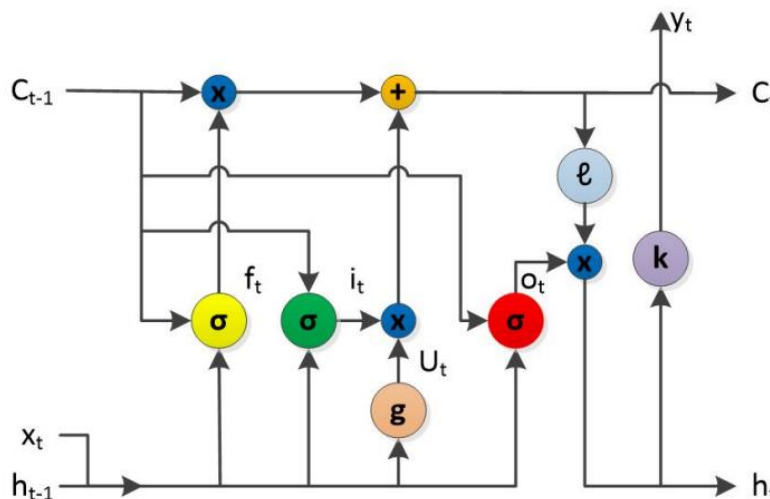


Рис. 2: Архитектура блока памяти LSTM с глазковыми соединениями.

*В. Схема прогнозирования RNN на основе LSTM*

Учитывая преимущества LSTM в прогнозировании временных рядов, мы используем схему RNN на основе LSTM в этой статье 4 для прогнозирования временных рядов электрической нагрузки. Схема применяет LSTM с глазком.

Учитывая входной временной ряд  $x = \{x_1, x_2, \dots, x_T\}$ , LSTM отображает входной временной ряд в две выходные временные последовательности  $h = \{h_1, h_2, \dots, h_T\}$  и  $y = \{y_1, y_2, \dots, y_T\}$  итеративно путем обновления состояний ячеек памяти с помощью следующей процедуры.

Во-первых, согласно рис. 2, ворота забывания применяются, чтобы помочь LSTM решить, какую информацию выбросить из состояния ячейки. Сигмовидная функция  $\sigma(\cdot)$  используется для вычисления активации ворот забывания как

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + W_{fc}C_{t-1} + b_f). \quad (6)$$

Выход  $f_t$  из (6) представляет собой значение от 0 до 1, соответствующее последнему состоянию ячейки  $C_{t-1}$ . Значение 0 означает полное забвение последнего состояния, а значение 1 означает полное сохранение последнего состояния.

Затем нам нужно сообщить LSTM, какая новая информация будет храниться в новом состоянии ячейки. Начнем с того, что LSTM использует сигмовидный слой, который называется входным слоем ворот, где

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + W_{ic}C_{t-1} + b_i), \quad (7)$$

решать, какую информацию обновлять. Слой сигмоиды  $g(\cdot)$  строит вектор  $U_t$  для хранения новых значений-кандидатов, которые должны быть добавлены к новому состоянию ячейки как

$$U_t = g(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (8)$$

Затем старое состояние ячейки  $C_{t-1}$  обновляется до нового состояния ячейки  $C_t$  с оцененными  $f_t$  и  $U_t$ . В частности, старое состояние ячейки умножается на  $f_t$ , чтобы забыть информацию из последнего состояния. Значения-кандидаты умножаются на уровень входного вентиля, чтобы решить, сколько новой информации нужно обновить до нового состояния ячейки, что дает

$$C_t = U_{it} + C_{t-1}f_t. \quad (9)$$

Другой сигмовидный слой  $\sigma(\cdot)$  затем используется в качестве выходного элемента для фильтрации и вывода состояния ячейки как  $o_t$ , где

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + W_{oc}C_{t-1} + b_o). \quad (10)$$

Кроме того, сигмовидная функция активации вывода ячейки  $l(\cdot)$  применяется к состоянию ячейки, которая затем умножается на вывод  $o_t$  для получения желаемой информации.

$$h_t = o_t l(C_t). \quad (11)$$

Что касается выхода блока памяти, то используется функция активации выхода  $k(\cdot)$ , т.е.

$$y_t = k(W_{yh}h_t + b_y). \quad (12)$$

В (6)-(12) матрицы  $W_{ix}$ ,  $W_{fx}$ ,  $W_{ox}$ ,  $W_{cx}$  — соответствующие матрицы входных весов,  $W_{ih}$ ,  $W_{fh}$ ,  $W_{oh}$ ,  $W_{ch}$  — рекуррентные матрицы весов,  $W_{yh}$  — скрытая матрица выходных весов,  $W_{ic}$ ,  $W_{fc}$ ,  $W_{oc}$  обозначают весовые матрицы глазковых соединений. Векторы  $b_i$ ,  $b_f$ ,  $b_o$ ,  $b_c$ ,  $b_y$  являются соответствующими векторами смещения.

#### IV. ЭКСПЕРИМЕНТАЛЬНЫЕ ОЦЕНКИ

##### A. Установка эксперимента

В этом разделе мы представляем наши эксперименты по применению схемы RNN на основе LSTM для прогнозирования электрической нагрузки. Мы сравнили предложенную схему RNN на основе LSTM со следующими методами: SARIMA, которая представляет

собой модель Seasonal Autoregressive Integrated Moving Average [32]; NARX — модель нелинейной авторегрессионной нейронной сети с экзогенными входными данными [12]; SVR (регрессия опорных векторов), которая является очень популярной моделью в прогнозировании финансовых временных рядов [14]; и NNETAR, которая представляет собой модель нейронной сети с прямой связью для одномерного прогнозирования временных рядов с одним скрытым слоем и запаздывающими входными данными. В качестве показателя эффективности использовались два критерия оценки: среднеквадратическая ошибка (RMSE) и средняя абсолютная процентная ошибка (MAPE) между реальными значениями и результатами прогнозирования.

Поскольку большинство сравниваемых методов разработаны для общего прогнозирования временных рядов, для объективного сравнения мы сравнили их эффективность с двумя наборами данных: набором данных об электрической нагрузке, который мы собрали, и набором данных о пассажирах авиакомпаний, который широко используется в качестве эталона для алгоритма. оценка. Набор данных о пассажирах авиакомпаний представляет собой временной ряд, описывающий месячное количество пассажиров международных авиакомпаний [33]. Этот набор данных включает 144 наблюдения всего за 12 лет. Как показано на рис. 3, наблюдается очевидный восходящий тренд и сильные сезонные колебания. В результате этот набор данных помогает нам изучить эффективность нашей схемы при прогнозировании коротких временных рядов с мультипликативными сезонными закономерностями.

Набор данных по электрической нагрузке представляет собой одномерный временной ряд, описывающий потребление электроэнергии в инженерном корпусе нашей школы. Этот набор данных содержит образцы энергопотребления здания, регистрируемые каждые 15 минут. Как показано на рис. 5, этот набор данных представляет собой сильный нестационарный и несезонный временной ряд, что создает большие трудности для долгосрочного прогнозирования временных рядов.

*Б. Результаты эксперимента с набором данных о потреблении электроэнергии*

В экспериментах мы использовали 904 выборки данных набора данных о потреблении электроэнергии для прогнозирования потребления электроэнергии.

ТАБЛИЦА I:

*Результаты RMSE и MAPE для набора данных о потреблении электроэнергии*

Методы прогнозирования	RMSE	MAPE
SVR	0.2044	0.1775
NNETAR	0.1952	0.1689
LSTM	0.0702	0.0535
NARX	0.1446	0.1192
SARIMA	0.2537	0.2001

$H = 96$  шагов вперед. Другими словами, мы использовали потребление электроэнергии за последние 10 дней, чтобы спрогнозировать потребление электроэнергии на следующий день. Из рис. 5 видно, что по сравнению с временными рядами пассажирских авиаперевозок этот временной ряд электрических нагрузок является более сложным, без какой-либо явной сезонной закономерности или тренда. Нестационарность и несезонность создают серьезную проблему для традиционных методов прогнозирования. Относительно более длинный горизонт прогнозирования, т. е.  $H = 96$ , делает точное прогнозирование еще более сложным.

Результаты эксперимента на рис. 5 показывают, что LSTM прогнозировал довольно хорошо по сравнению с исходным временным рядом. На рис. 6 более четко показано, что LSTM превзошел все остальные методы с лучшими прогнозируемыми временными рядами. Таблица II показывает, что LSTM превзошел все остальные методы с наименьшими ошибками прогнозирования. В этом сложном сценарии прогнозирования электрической



нагрузки эффективность остальных четырех методов была признана в целом неудовлетворительной. Хотя NARX уловил общую тенденцию реального временного ряда, результат прогнозирования был довольно ложным, что привело к неточному прогнозированию с большими RMSE и MAPE. SARIMA также не сработала в этом долгосрочном прогнозировании из-за нестационарности и несезонности.

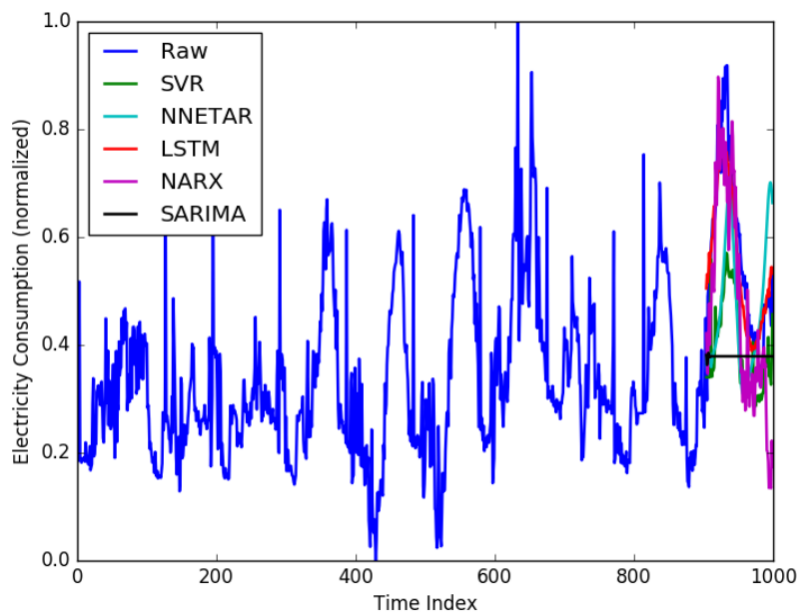


Рис. 3: Сравнение набора данных о потреблении электроэнергии внутри и вне выборки.

### V. ВЫВОДЫ

В этой статье мы предлагаем использовать рекуррентную нейронную сеть (RNN) на основе долговременной памяти (LSTM) для решения сложной задачи краткосрочного прогнозирования электрической нагрузки. Используя долгосрочные зависимости в

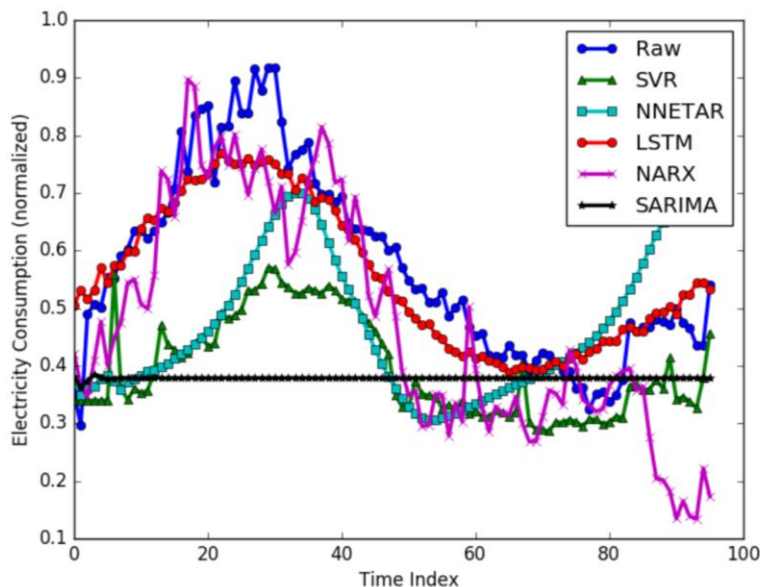


Рис. 4: Сравнение прогнозов для набора данных о потреблении электроэнергии.

временных рядов, LSTM способен прогнозировать сложные одномерные временные ряды электрической нагрузки с сильной нестационарностью и несезонностью. Эксперименты проводятся с коротким эталонным набором данных о пассажирах международных авиакомпаний и длинным набором данных о потреблении электроэнергии. Результаты эксперимента показывают, что метод прогнозирования на основе LSTM может

превзойти большинство традиционных методов прогнозирования в сложной задаче краткосрочного прогнозирования электрической нагрузки.

#### ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

- [1] C. Nataraja, M. Gorawar, G. Shilpa, and J. S. Harsha, "Short term load forecasting using time series analysis: A case study for karnataka, india," *International Journal of Engineering Science and Innovative Technology*, vol. 1, pp. 45–53, 2012.
- [2] L. Hernandez, C. Baladr ´ on, J. M. Aguiar, L. Calavia, B. Carro, ´ A. Sanchez-Esguevillas, F. P ´ erez, ´ A. Fern ´ andez, and J. Lloret, "Artificial ´ neural network for short-term load forecasting in distribution systems," *Energies*, vol. 7, no. 3, pp. 1576–1598, 2014.
- [3] E. Almeshaei and H. Soltan, "A methodology for electric power load forecasting," *Alexandria Engineering Journal*, vol. 50, no. 2, pp. 137– 144, 2011.
- [4] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [5] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and arima models for time series forecasting," *Applied Soft Computing*, vol. 11, no. 2, pp. 2664–2675, 2011.
- [6] K.-Y. Chen and C.-H. Wang, "A hybrid sarima and support vector machines in forecasting the production values of the machinery industry in taiwan," *Expert Systems with Applications*, vol. 32, no. 1, pp. 254– 264, 2007.
- [7] C.-L. Hor, S. J. Watson, and S. Majithia, "Daily load forecasting and maximum demand estimation using arima and garch," in *Probabilistic Methods Applied to Power Systems, 2006. PMAAPS 2006. International Conference on. IEEE, 2006*, pp. 1–6.
- [8] K. M. Kam, *Stationary and non-stationary time series prediction using state space model and pattern-based approach*. The University of Texas at Arlington, 2014.
- [9] M. Lineesh, K. Minu, and C. J. John, "Analysis of nonstationary nonlinear economic time series of gold price: A comparative study," in *International Mathematical Forum*, vol. 5, no. 34. Citeseer, 2010, pp. 1673–1683.
- [10] S. Panigrahi, Y. Karali, and H. Behera, "Time series forecasting using evolutionary neural network," *International Journal of Computer Applications*, vol. 75, no. 10, 2013.
- [11] C. Garc ´ ia-Ascanio and C. Mate, "Electric power demand forecasting ´ using interval time series: A comparison between var and impl," *Energy Policy*, vol. 38, no. 2, pp. 715–725, 2010.
- [12] M. Espinoza, J. A. Suykens, R. Belmans, and B. De Moor, "Electric load forecasting," *IEEE Control Systems*, vol. 27, no. 5, pp. 43–57, 2007.
- [13] Y. Li and J. Wang, "The load forecasting model based on bayes-grnn," *Journal of Software*, vol. 7, no. 6, pp. 1273–1280, 2012.
- [14] C.-N. Ko and C.-M. Lee, "Short-term load forecasting using svr (support vector regression)-based radial basis function neural network with dual extended kalman filter," *Energy*, vol. 49, pp. 413–422, 2013.
- [15] J. Du Preez and S. F. Witt, "Univariate versus multivariate time series forecasting: an application to international tourism demand," *International Journal of Forecasting*, vol. 19, no. 3, pp. 435–451, 2003.
- [16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [17] H. Sak, A. W. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling." in *INTERSPEECH, 2014*, pp. 338–342.
- [18] S. B. Taieb, G. Bontempi, A. F. Atiya, and A. Sorjamaa, "A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition," *Expert systems with applications*, vol. 39, no. 8, pp. 7067–7083, 2012.
- [19] C. Hamzac,ebi, D. Akay, and F. Kutay, "Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3839–3844, 2009.

- [20] D. R. Cox, "Prediction by exponentially weighted moving averages and related methods," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 414–422, 1961.
- [21] N. H. An and D. T. Anh, "Comparison of strategies for multi-step-ahead prediction of time series using neural network," in *2015 International Conference on Advanced Computing and Applications (ACOMP)*. IEEE, 2015, pp. 142–149.
- [22] G. Bontempi, "Long term time series prediction with multi-input multioutput local learning," *Proc. 2nd ESTSP*, pp. 145–154, 2008.
- [23] S. B. Taieb, A. Sorjamaa, and G. Bontempi, "Multiple-output modeling for multi-step-ahead time series forecasting," *Neurocomputing*, vol. 73, no. 10, pp. 1950–1957, 2010.
- [24] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv preprint arXiv:1506.00019*, 2015.
- [25] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in neural information processing systems*, 2014, pp. 3104–3112.
- [26] L. Medsker and L. Jain, "Recurrent neural networks," *Design and Applications*, 2001.
- [27] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks." *ICML (3)*, vol. 28, pp. 1310–1318, 2013.
- [28] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [29] A. Graves, N. Jaitly, and A.-r. Mohamed, "Hybrid speech recognition with deep bidirectional lstm," in *Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on*. IEEE, 2013, pp. 273–278.
- [30] A. Graves, "Neural networks," in *Supervised Sequence Labelling with Recurrent Neural Networks*. Springer, 2012, pp. 15–35.
- [31] F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, "Learning precise timing with lstm recurrent networks," *Journal of machine learning research*, vol. 3, no. Aug, pp. 115–143, 2002.
- [32] Y. Chakhchoukh, P. Panciatici, and L. Mili, "Electric load forecasting based on statistical robust methods," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 982–991, 2011.
- [33] J. Faraway and C. Chatfield, "Time series forecasting with neural networks: a comparative study using the airline data," *Applied statistics*, pp. 231–250, 1998.
- [34] R. Adhikari and R. Agrawal, "An introductory study on time series modeling and forecasting," *arXiv preprint arXiv:1302.6613*, 2013