МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ

РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное

образовательное учреждение высшего образования

«Самарский национальный исследовательский университет

имени академика С.П. Королева»

Институт информатики и кибернетики

Факультет информатики

Кафедра программных систем

ОТЧЁТ

по лабораторной работе № 2

«Прогнозирование временных рядов»

по курсу «Нейронные сети»

Выполнил:

Каликин М.А.

гр. 6132-020402D

Проверила:

Жданова А.Н.

Самара 2022

СОДЕРЖАНИЕ

[1 Математическая постановка задачи 3](#_Toc121425674)

[2 Архитектура сети, параметры, методы обучения 4](#_Toc121425675)

[3 Вычислительные эксперименты 20](#_Toc121425676)

[4 Результаты работы 41](#_Toc121425677)

[ЗАКЛЮЧЕНИЕ 44](#_Toc121425678)

[ССЫЛКИ 45](#_Toc121425679)

# **1 Математическая постановка задачи**

Набор данных содержит в себе 3526 строк данных за период с 2007-01-01 по 2021-01-01. Набор данных использовался и был скачен с сайта yahoo finance, за пример взяты акции компании Apple.

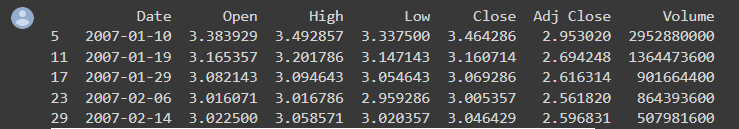
В наборе данных 7 классов:

Date, Open, High, Low, Close, Adj Close, Volume

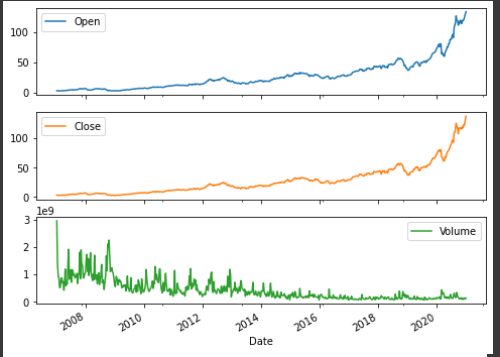
# **2 Архитектура сети, параметры, методы обучения**

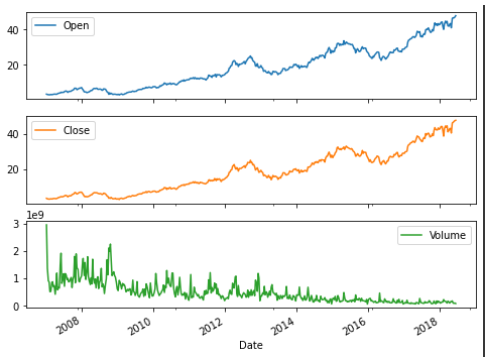
Для выполнения работы использовался фреймворк TensorFlow, и вспомогательные библиотеки Python. С помощью TensorFlow были построены модели для прогнозирования временных рядов.

Загрузим этот набор данных в colab.research.google.com и прочитаем их при помощи библиотеки pandas (файл формата \*csv).



Выполним вывод графиков некоторых функций с течением времени.





**Выполним разделение на выборки.**

Используем разделение на 70%, 20%, 10% для обучающих, проверочных и тестовых наборов данных. Данные не перемешиваются случайным образом перед разделением. Происходит это по двум причинам:

1. Гарантия того, что разделение данных на т.н. окна последовательных выборок по-прежнему возможно.
2. Гарантия того, что результаты проверки/тестирования будут более реальными, потому что они оцениваются на основе данных, собранных после обучения модели

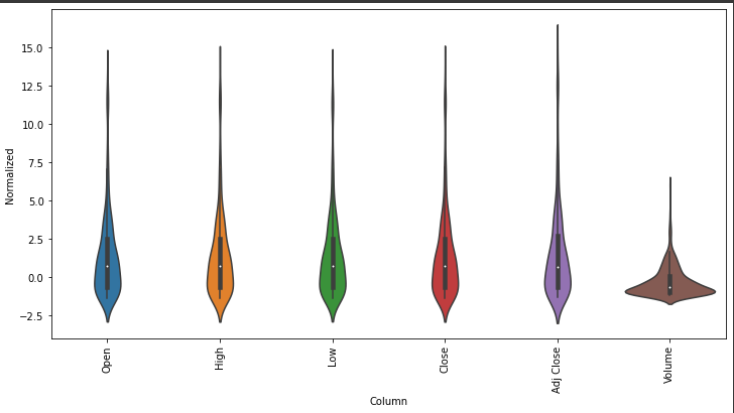
**Нормализуем данные**

Перед тем как начать обучение нейронной сети, нужно масштабировать функции. Нормализация – это распространенный способ масштабирования: вычесть среднее значение и разделить на стандартное отклонение каждого признака.

Среднее значение и стандартное отклонение следует вычислять только с использованием обучающих данных, чтобы модели не имели доступа к значениям в проверочных и тестовых наборах.

Модель не должна иметь доступа к будущим значениям в тренировочном наборе данных во время обучения и нормализация должна выполняться с использованием скользящих средних.

Посмотрим на распределение функций.



**Окно данных.**

Модели будут делать набор прогнозов на основе окна последовательных выборок из данных.

Основные особенности окон ввода:

* Ширина (количество временных шагов) окон ввода и метки.
* Смещение времени между ними.
* Какие функции используются в качестве входных данных, меток или того и другого.

Создаются различные модели (включая линейные модели, модели DNN, CNN и RNN) и используются они для обеих целей:

* Прогнозы с одним выходом и несколькими выходами.
* Прогнозы с одним и несколькими временными шагами.

В этом разделе основное внимание уделяется реализации окна данных, чтобы его можно было повторно использовать для всех этих моделей.

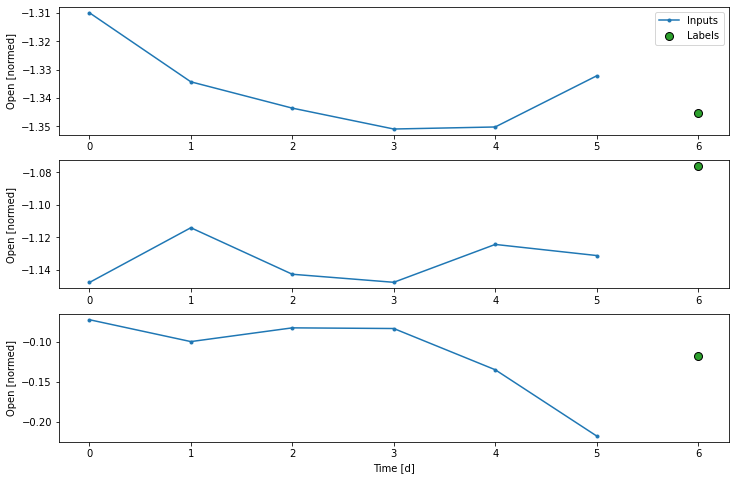
В зависимости от задачи и типа модели может потребоваться создание различных окон данных.

Начнем с создания класса WindowGenerator. Метод \_\_init\_\_ включает всю необходимую логику для индексов ввода и меток.

Он также принимает обучающие, оценочные и тестовые кадры данных в качестве входных данных. Позже они будут преобразованы в tf.data.Dataset окон.

Метод split\_window преобразует их в окно входных данных и окно меток.

Данный график выравнивает входные данные, метки и (более поздние) прогнозы на основе времени, к которому относится элемент:

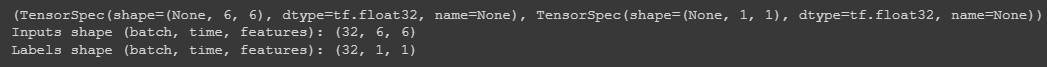


**Создадим датасет.**

Метод make\_dataset возьмет временной ряд DataFrame и преобразует его в tf.data.Dataset из пар (input\_window, label\_window) с помощью функции tf.keras.utils.timeseries\_dataset\_from\_array

Объект WindowGenerator содержит обучающие, проверочные и тестовые данные.

Свойство Dataset.element\_spec сообщает о структуре, типе данных и форм элементов набора данных. Итерация по Dataset дает конкретные т.н. партии.



**Одноступенчатые модели(Single step).**

Самая простая модель, которую вы можете построить на такого рода данных, – это модель, которая предсказывает значение одной функции – 1 временной шаг (один день) в будущее, основываясь только на текущих условиях.

Итак, начнем с построения моделей для прогнозирования значения Open (цены открытия) на один день вперед.

Настроим объект WindowGenerator для создания этих одношаговых пар (input, label). Объект window создает tf.data.Dataset из обучающих, проверочных и тестовых наборов, что позволяет легко перебирать пакеты данных.

**Базовый уровень(Baseline).**

Перед созданием обучаемой модели было бы неплохо иметь базовый уровень производительности в качестве точки для сравнения с более поздними более сложными моделями.

Эта первая задача состоит в том, чтобы предсказать цену открытия на один день вперед, учитывая текущее значение всех признаков. Текущие значения включают текущую цену.

Итак, начнем с модели, которая просто возвращает текущую цену в качестве прогноза, прогнозируя «без изменений». Это разумная базовая линия, поскольку цена изменяется медленно.

В качестве потерь используем tf.losses.MeanSquaredError. Вычисляет среднее квадратическое значение ошибок между метками и прогнозами.

В качестве метрик tf.metrics.MeanAbsoluteError. Вычисляет среднюю абсолютную ошибку между метками и прогнозами.

Модель делает прогнозы на один день вперед на основе одного входного временного шага.

Метки, сдвинуты вправо на один день.

**Линейная модель(Linear).**

Самая простая обучаемая модель, которую вы можно применить к этой задаче, – это вставить линейное преобразование между входом и выходом. В этом случае результат временного шага зависит только от этого шага:

Слой tf.keras.layers.Dense без набора activation является линейной моделью. Слой преобразует только последнюю ось данных из (batch, time, inputs) в (batch, time, units); он применяется независимо к каждому элементу по осям batch и time.

В данной модели используется лишь 1 слой Dense.

Так как мы будем обучать множество моделей, то упакуем процедуру обучения в функцию def.

В качестве потерь используем tf.losses.MeanSquaredError. Вычисляет среднее квадратическое значение ошибок между метками и прогнозами.

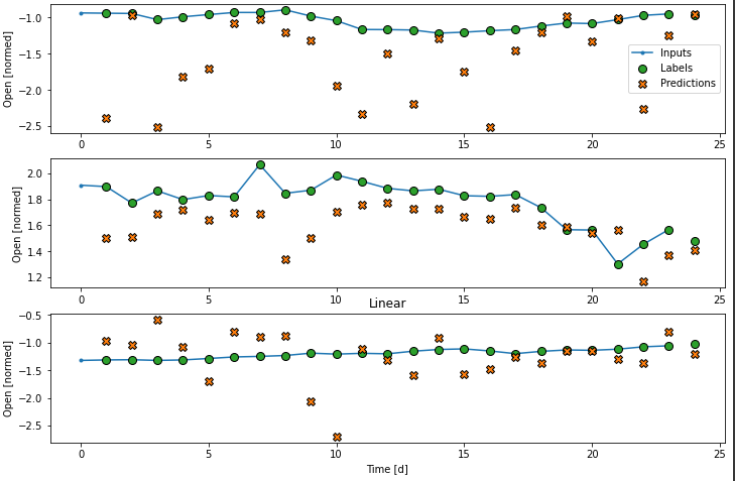
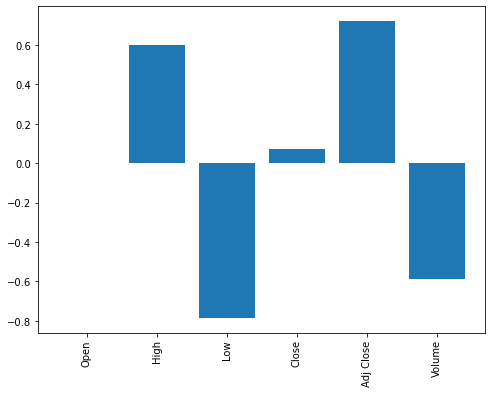
В качестве метрик tf.optimizers.Adam. Вычисляет среднюю абсолютную ошибку между метками и прогнозами.

В функции активации используем tf.losses.MeanSquaredError. Вычисляет среднее квадратическое значение ошибок между метками и прогнозами.

Также в качестве ранней остановки используем функцию tf.keras.callbacks.EarlyStopping, чтобы прекратите обучение, когда отслеживаемый показатель перестал улучшаться.

Количество эпох составляет 20.

Одним из преимуществ линейных моделей является то, что их относительно просто интерпретировать. Вы можете вытащить веса слоя и визуализировать вес, назначенный каждому входу:

**Dense модель**

В Dense модели используются 3 полносвязных слоя Dense. Первые 2 из которых состоят из 64 нейронов. Используется функции активации – нелинейная функция и (Relu), а также выходной слой Dense как в линейной функции.

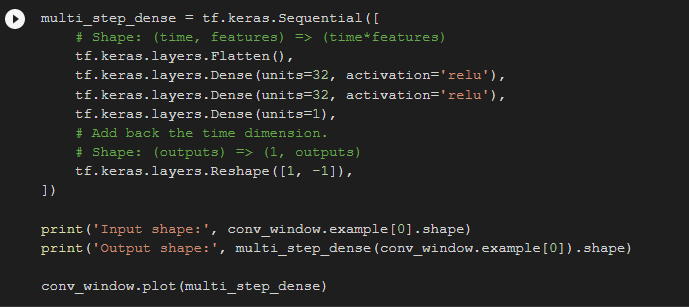
Обучим модель при помощи, ранее упакованной в функцию def процедуры обучения.

**Multi-step dense модель**

Baseline, linear и dense модели обрабатывали каждый временной шаг независимо. Здесь модель будет принимать несколько временных шагов в качестве входных данных для получения одного вывода.

Flatten в данной модели, используется в качестве первого слоя и выравнивает входные данные модели затем используются 3 слоя полносвязных слоёв Dense. Первые 2 слоя состоят из 32 нейронов. В качестве функции активации используется нелинейная функция и (Relu). А также выходной слой Dense как в линейной функции. На выходе используется слой Reshape который преобразует данные в заданную форму.

Обучим модель при помощи, ранее упакованной в функцию def процедуры обучения.



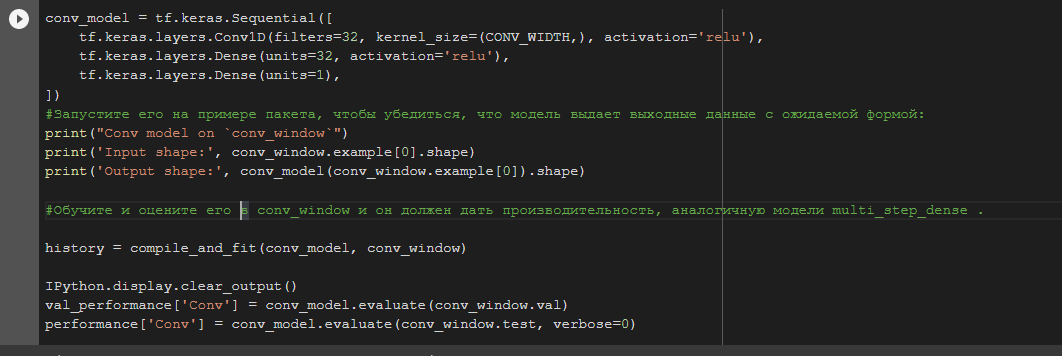
**Сверточная модель**

Слой свертки ( tf.keras.layers.Conv1D ) также использует несколько временных шагов в качестве входных данных для каждого прогноза.

Представлена та же модель, что и в multi\_step\_dense, переписанная с помощью свертки.

tf.keras.layers.Flatten и первый tf.keras.layers.Dense заменяются tf.keras.layers.Conv1D. В нем задается размер свертки равный 3.

tf.keras.layers.Reshape больше не нужен, так как свертка сохраняет ось времени в своих выходных данных.

Обучим модель при помощи, ранее упакованной в функцию def процедуры обучения.

**Реккурентная модель**

Рекуррентная нейронная сеть (RNN) — это тип нейронной сети, хорошо подходящий для данных временных рядов. RNN обрабатывают временной ряд шаг за шагом, сохраняя внутреннее состояние от шага к шагу.

Важным аргументом конструктора для всех слоев Keras RNN, таких как tf.keras.layers.LSTM , является аргумент return\_sequences . Этот параметр может настроить слой.

В данной модели LSTM используется в качестве первого слоя и запоминает данные. Он состоит из 32 нейронов, а также параметр return\_sequences, который находится в состоянии True, слой возвращает результат для каждого входа. А также выходной слой Dense как в линейной функции.

С return\_sequences=True модель можно обучать на данных за 24 дня за раз.

**Модели с несколькими выходами**

До сих пор все модели предсказывали одну выходную характеристику T (degC) для одного временного шага.

Все эти модели можно преобразовать для прогнозирования нескольких функций, просто изменив количество единиц в выходном слое и настроив окна обучения, чтобы включить все функции в labels ( example\_labels ).

Базовая

Используем ту же базовую модель ( Baseline ), но на этот раз с повторением всех функций вместо выбора конкретного label\_index.

Dense

Используем ту же Dense модель ( Dense ), но на этот раз с повторением всех функций вместо выбора конкретного label\_index.

RNN

Используем ту же RNN модель ( LSTM ), но на этот раз с повторением всех функций вместо выбора конкретного label\_index.

Остаточная RNN

Baseline модель, использованная ранее, использовала тот факт, что последовательность не сильно меняется от одного временного шага к другому. Каждая модель, обученная в этом руководстве до сих пор, была случайным образом инициализирована, а затем должна была узнать, что результат представляет собой небольшое изменение по сравнению с предыдущим временным шагом.

В анализе временных рядов обычно строят модели, которые вместо предсказания следующего значения предсказывают, как значение изменится на следующем временном шаге. Точно так же остаточные сети — или ResNets — в глубоком обучении относятся к архитектурам, в которых каждый уровень добавляется к накапливающемуся результату модели.

tf.initializers.zeros , чтобы гарантировать, что начальные прогнозируемые изменения будут небольшими и не пересилят остаточное соединение. Здесь нет проблем с нарушением симметрии для градиентов, поскольку zeros используются только на последнем слое.

**Multi-step модели**

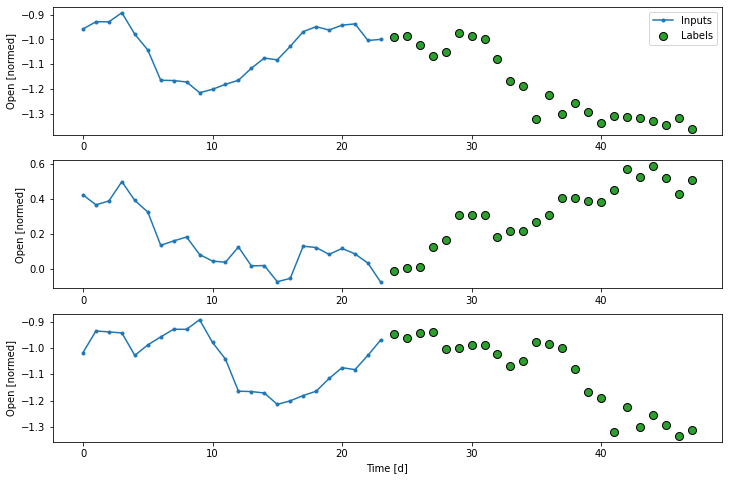
Как модели с одним выходом, так и модели с несколькими выходами в предыдущих разделах делали прогнозы с одним временным шагом, на один день вперед.

В многоэтапном прогнозировании модель должна научиться прогнозировать диапазон будущих значений. Таким образом, в отличие от одноступенчатой модели, в которой предсказывается только одна точка будущего, многоступенчатая модель предсказывает последовательность будущих значений.

Есть два грубых подхода к этому:

1. Прогнозы одиночного выстрела, когда весь временной ряд прогнозируется сразу.
2. Прогнозы авторегрессии, при которых модель делает только одношаговые прогнозы, а ее выходные данные возвращаются в качестве входных данных.

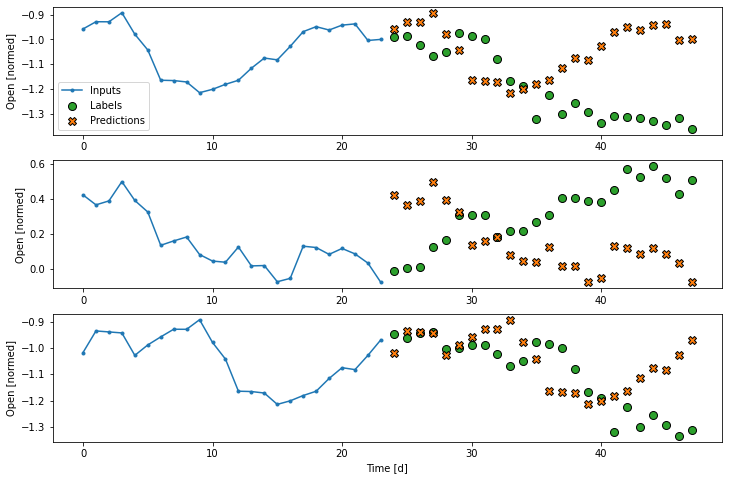
Все модели будут предсказывать все функции на всех выходных временных шагах.

Для многоступенчатой модели обучающие данные снова состоят из ежедневных выборок. Однако здесь модели научатся предсказывать будущее на 24 дня, учитывая 24 дня прошлого.

**Baselines**

Простая базовая линия для этой задачи – повторить последний временной шаг ввода для необходимого количества выходных временных шагов.

Базовый(повторение)

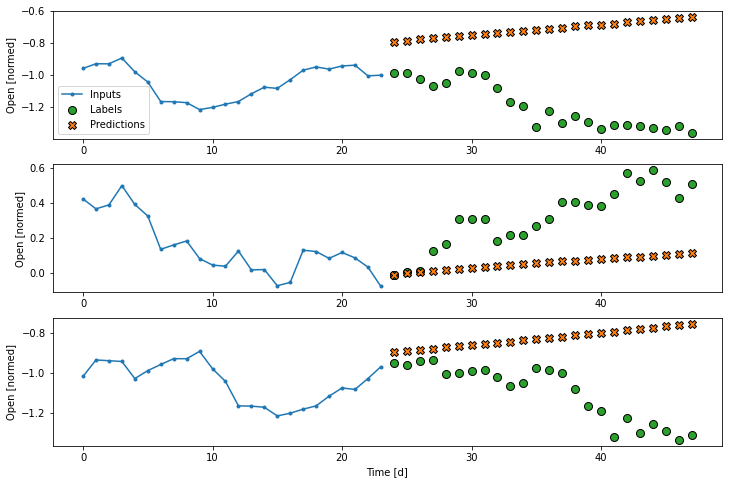
Поскольку эта задача состоит в том, чтобы предсказать 24 дня в будущем, учитывая 24 дня в прошлом, другой простой подход состоит в том, чтобы повторить предыдущий день, предполагая, что завтра будет похоже.

**Single-shot модели**

Одним из высокоуровневых подходов к этой проблеме является использование «однократной» модели, в которой модель делает прогноз всей последовательности за один шаг.

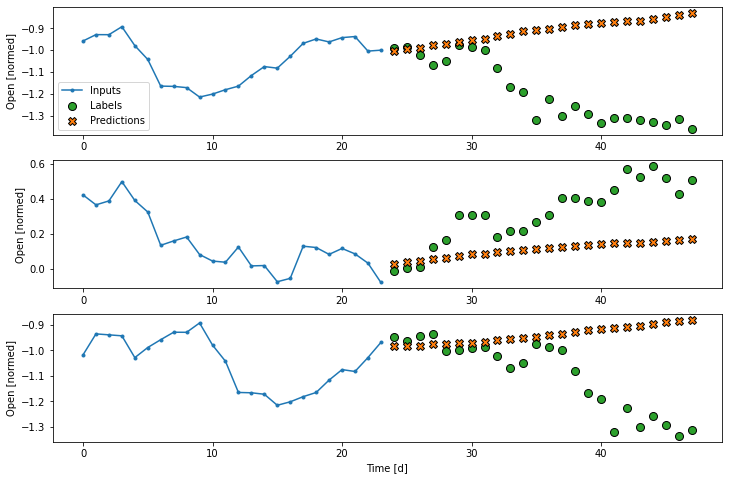
Это может быть эффективно реализовано как tf.keras.layers.Dense с выходными единицами OUT\_STEPS\*features . Модель просто должна преобразовать этот вывод в требуемый (OUTPUT\_STEPS, features).

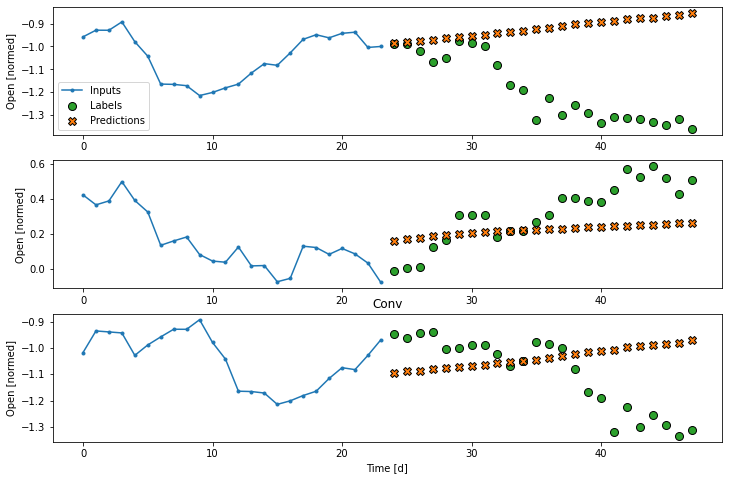
Linear модель

Простая линейная модель, основанная на последнем входном временном шаге, работает лучше любой базовой линии, но у нее недостаточно мощности. Модель должна прогнозировать временные шаги OUTPUT\_STEPS на основе одного входного временного шага с линейной проекцией. Он может зафиксировать только низкоразмерный фрагмент поведения, вероятно, основанный в основном на времени суток и времени года.

Dense модель

Добавление tf.keras.layers.Dense между входом и выходом делает линейную модель более мощной, но по-прежнему основано только на одном временном шаге ввода. В модели меняется количество нейронов.

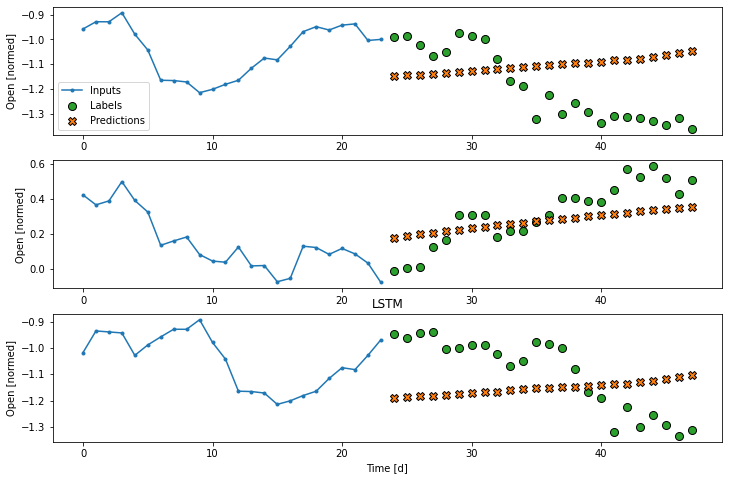
Сверточная (CNN)

Сверточная модель делает прогнозы на основе истории фиксированной ширины, что может привести к более высокой производительности, чем плотная модель, поскольку она может видеть, как все меняется с течением времени. В модели меняется количество нейронов.

Реккурентная(RNN)

Рекуррентная модель может научиться использовать длинную историю входных данных, если она имеет отношение к прогнозам, которые делает модель. Здесь модель будет накапливать внутреннее состояние за 24 дня, прежде чем сделать один прогноз на следующие 24 дня.

В этом однократном формате LSTM нужно выводить только на последнем временном шаге, поэтому установим return\_sequences=False в tf.keras.layers.LSTM.



**Авторегрессионная модель(Autoregressive model)**

Все вышеперечисленные модели предсказывают всю выходную последовательность за один шаг.

В некоторых случаях для модели может быть полезно разложить этот прогноз на отдельные временные шаги. Затем выходные данные каждой модели могут быть возвращены в себя на каждом шаге, и прогнозы могут быть сделаны в зависимости от предыдущего, как в классическом Генерация последовательностей с рекуррентными нейронными сетями. Одним из явных преимуществ модели этого типа является то, что ее можно настроить для получения выходных данных различной длины.

Модель будет иметь ту же базовую форму, что и одноэтапные модели LSTM из более ранних версий: слой tf.keras.layers.LSTM , за которым следует слой tf.keras.layers.Dense , который преобразует выходные данные слоя LSTM в прогнозы модели.

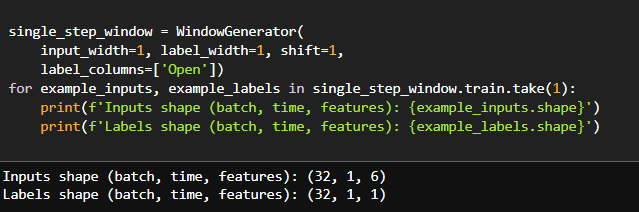
tf.keras.layers.LSTM — это tf.keras.layers.LSTMCell , обернутый в tf.keras.layers.RNN более высокого уровня, который управляет для вас результатами состояния и последовательности.

В этом случае модель должна вручную управлять входными данными для каждого шага, поэтому она использует tf.keras.layers.LSTMCell непосредственно для интерфейса нижнего уровня с одним временным шагом.

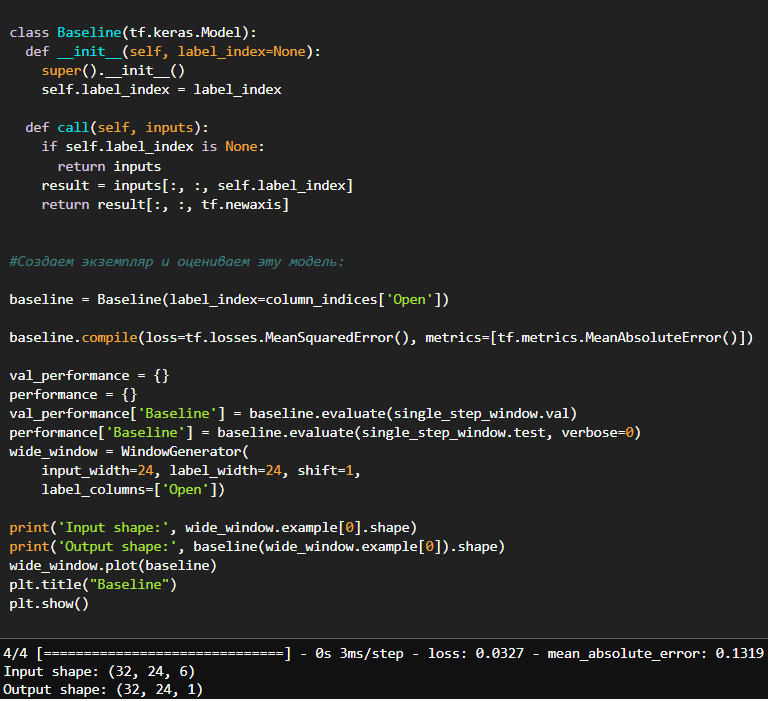
Первый метод, в котором нуждается эта модель, — это метод warmup для инициализации ее внутреннего состояния на основе входных данных. После обучения это состояние будет захватывать соответствующие части истории ввода. Это эквивалентно одноэтапной модели LSTM из более раннего.

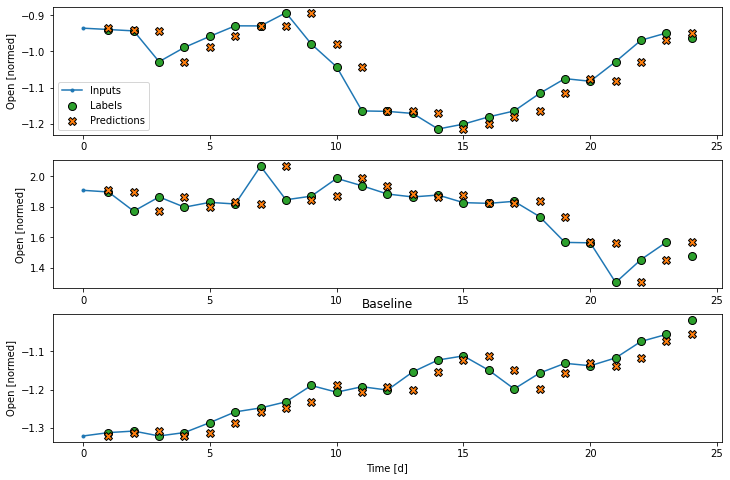
# **3 Вычислительные эксперименты**

**Single-step модели**



**Baseline**



****

На приведенных выше графиках трех примеров одноэтапная модель работает в течение 24 дней.

Синяя линия Inputs показывает входную температуру на каждом временном шаге. Модель получает все функции, этот график показывает только цену открытия.

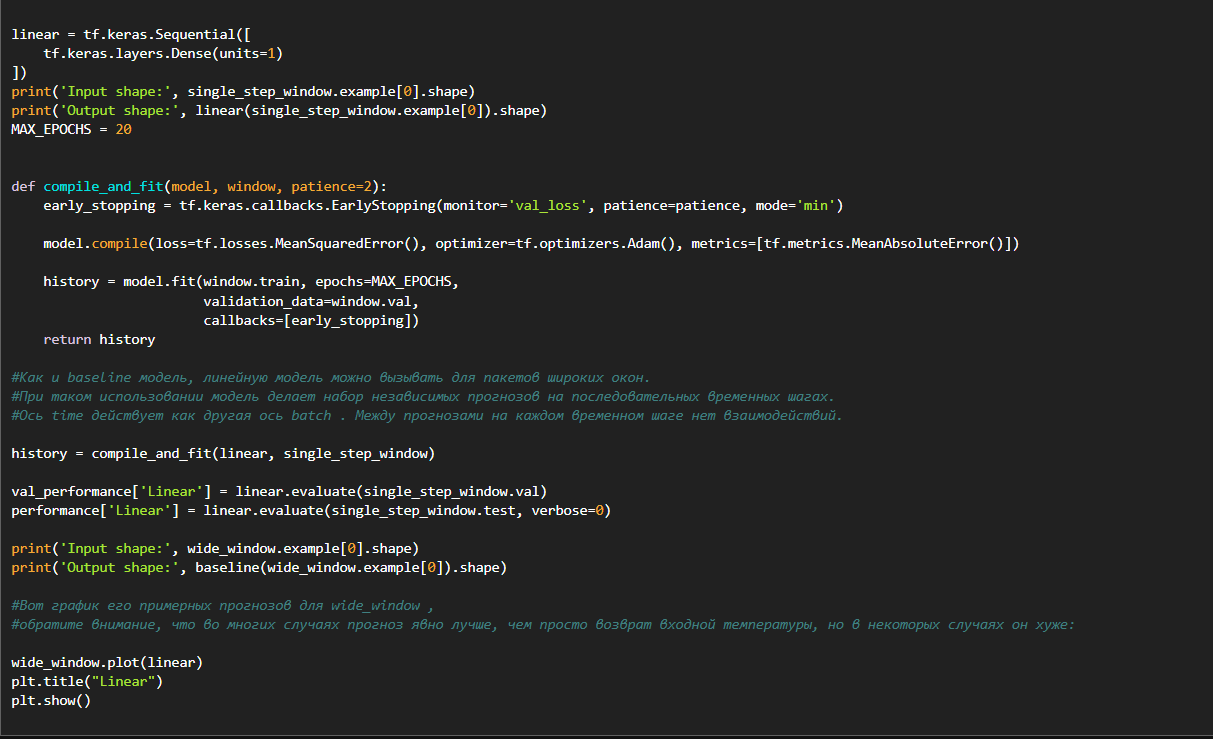
Зеленые линии - Labels показывают целевое значение прогноза. Эти точки отображаются во время прогнозирования, а не во время ввода. Поэтому диапазон меток смещен на 1 шаг относительно входов.

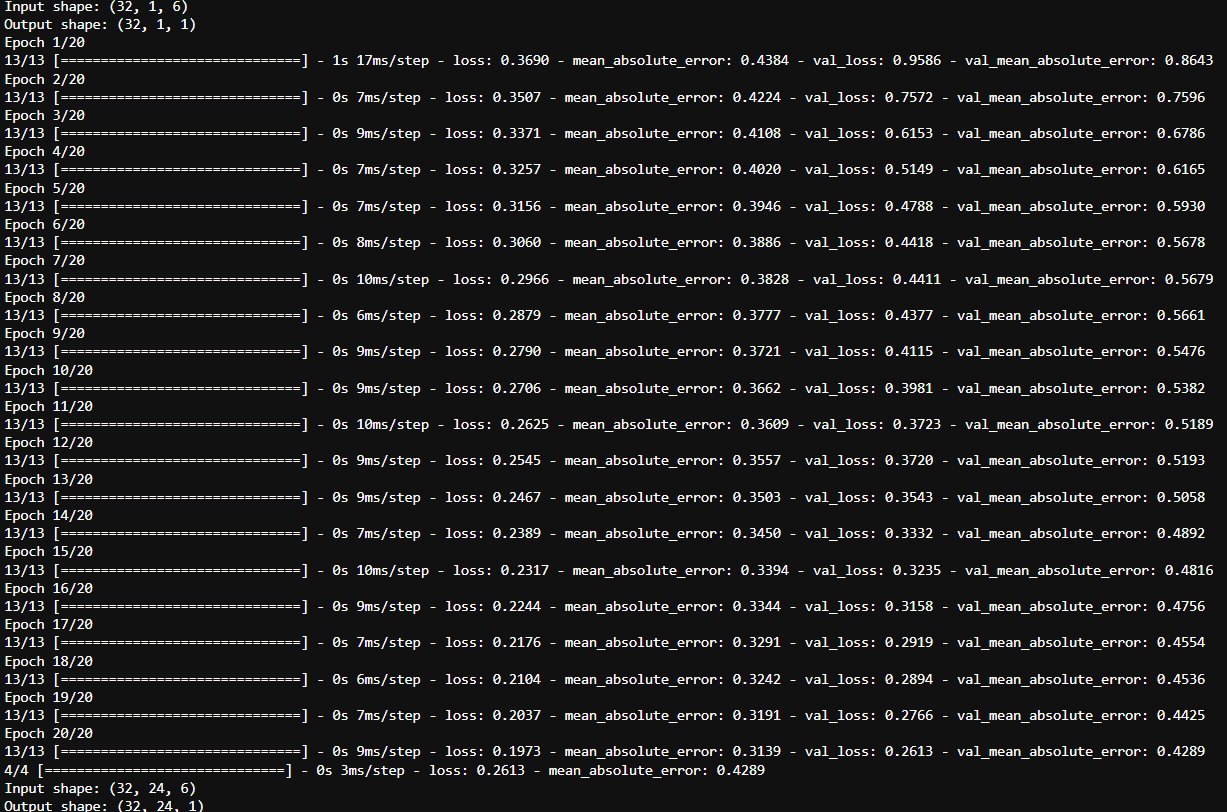
Оранжевые кресты Predictions — это прогнозы модели для каждого выходного временного шага. Если бы модель предсказывала идеально, прогнозы попадали бы прямо в Labels.

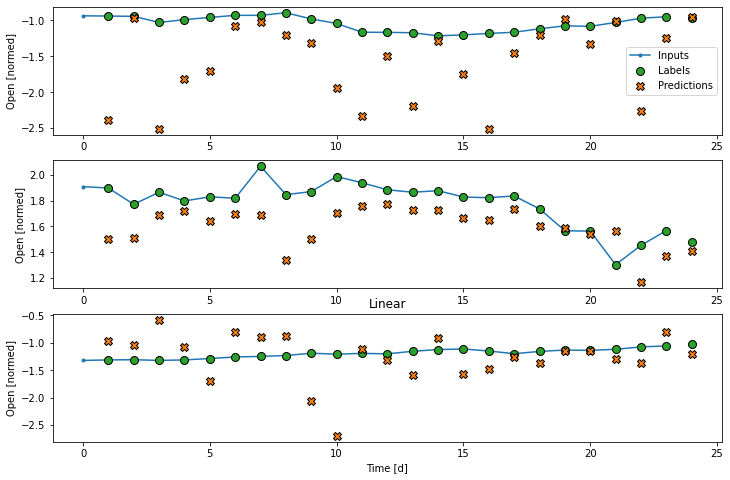
**Линейчатая модель**

Одним из преимуществ линейных моделей является то, что их относительно просто интерпретировать. Вы можете вытащить веса слоя и визуализировать вес, назначенный каждому входу.

Как и baseline модель, линейную модель можно вызывать для пакетов широких окон. При таком использовании модель делает набор независимых прогнозов на последовательных временных шагах. Ось time действует как другая ось batch. Между прогнозами на каждом временном шаге нет взаимодействий.



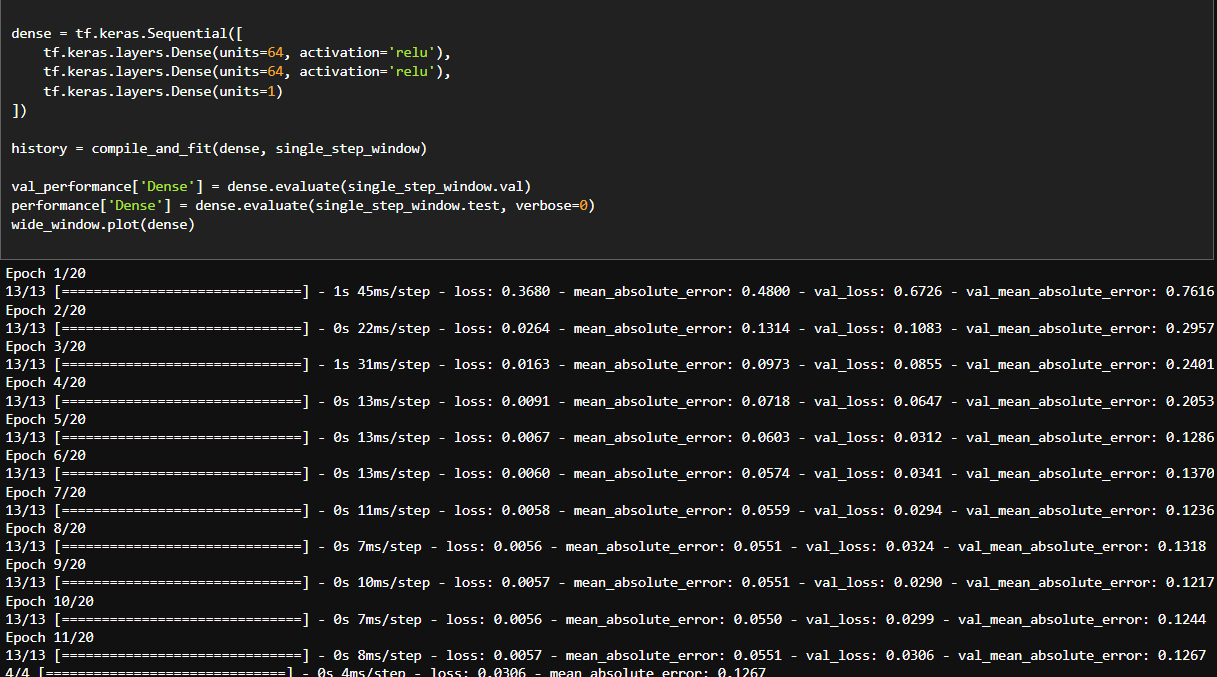


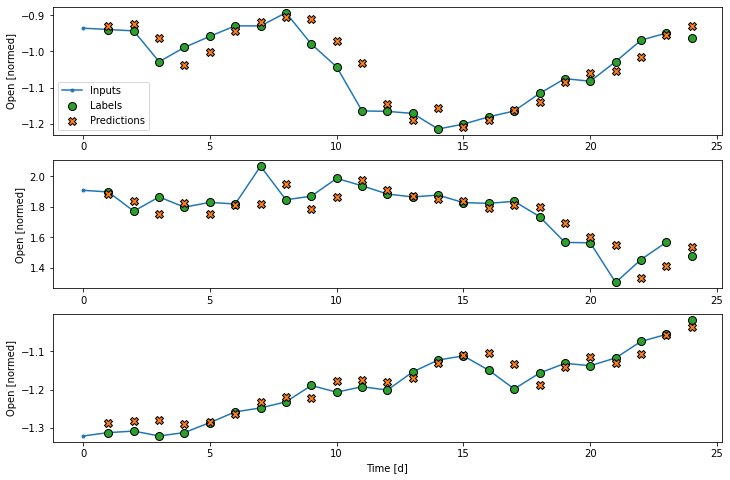
****

**Dense модель**

Прежде чем применять модели, которые фактически работают с несколькими временными шагами, стоит проверить производительность более глубоких и мощных моделей с одним входным шагом.

Вот модель, похожая на linear модель, за исключением того, что между входом и выходом укладывается несколько Dense слоев:

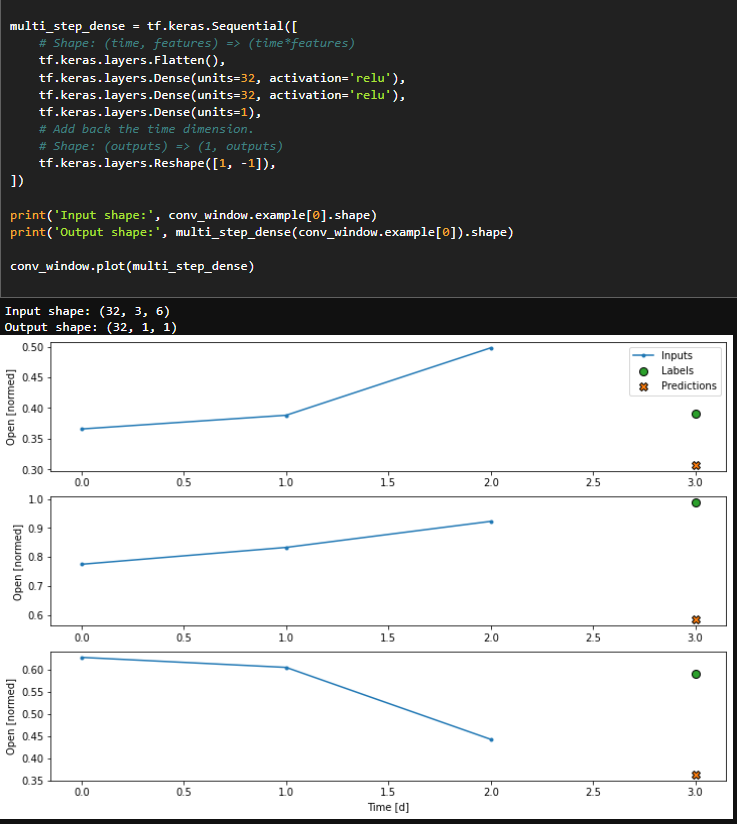


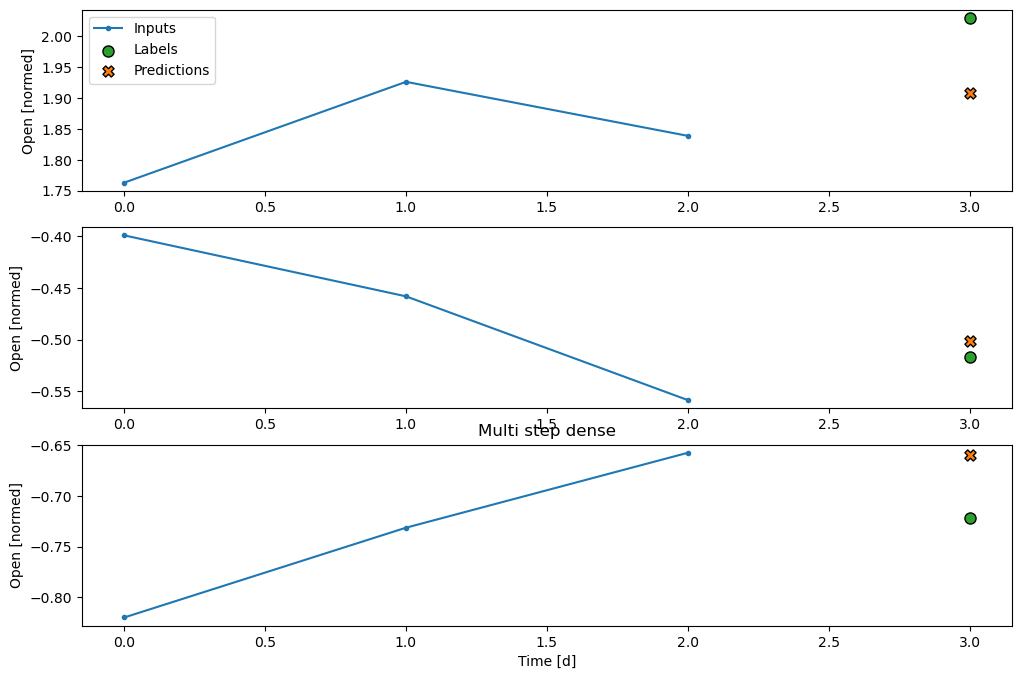


**Multi-step dense модель**

Одношаговая модель не имеет контекста для текущих значений входных данных. Он не может видеть, как входные объекты меняются с течением времени. Чтобы решить эту проблему, модели требуется доступ к нескольким временным шагам при прогнозировании.

Основным недостатком этого подхода является то, что результирующая модель может выполняться только на входных окнах именно такой формы.



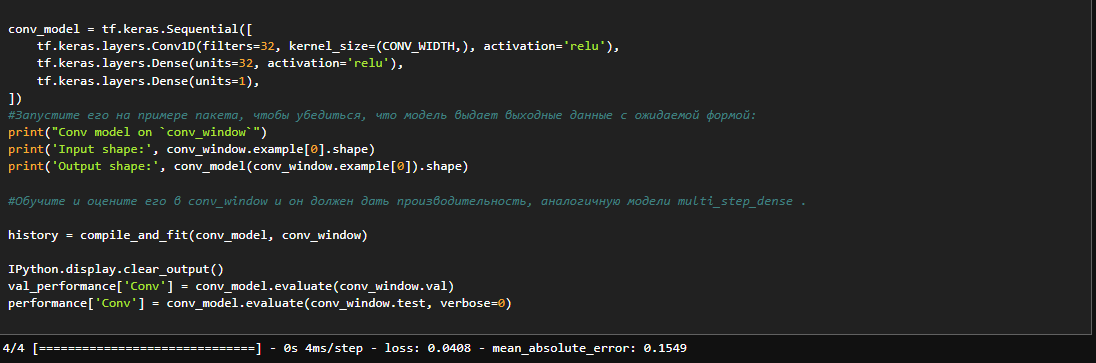


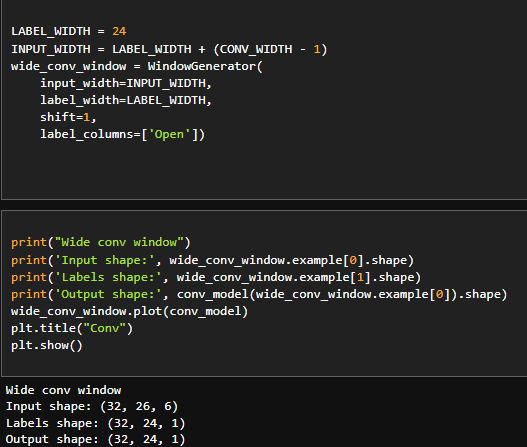
**Сверточная модель(CNN)**

Обучим и оценим его в conv\_window и он должен дать производительность, аналогичную модели multi\_step\_dense.

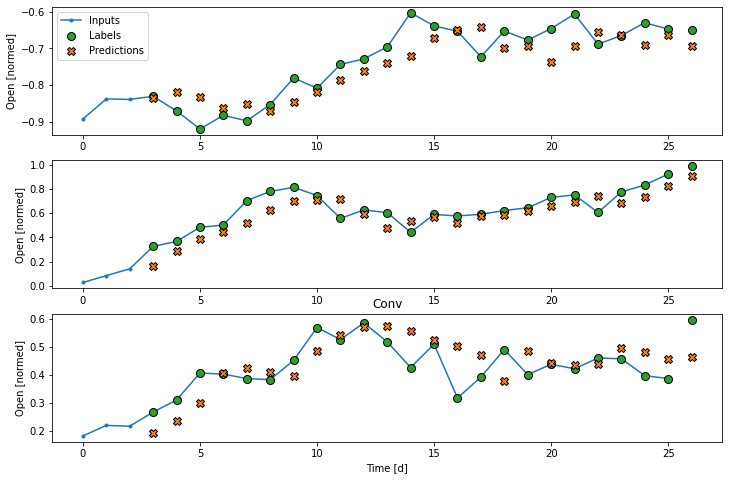
Разница между этой conv\_model и моделью multi\_step\_dense заключается в том, что conv\_model можно запускать на входах любой длины. Сверточный слой применяется к скользящему окну входных данных.

Чтобы обучение или построение графика работали, вам нужно, чтобы метки и прогноз имели одинаковую длину. Поэтому создадим WindowGenerator для создания широких окон с несколькими дополнительными временными шагами ввода, чтобы длина метки и предсказания совпадала.

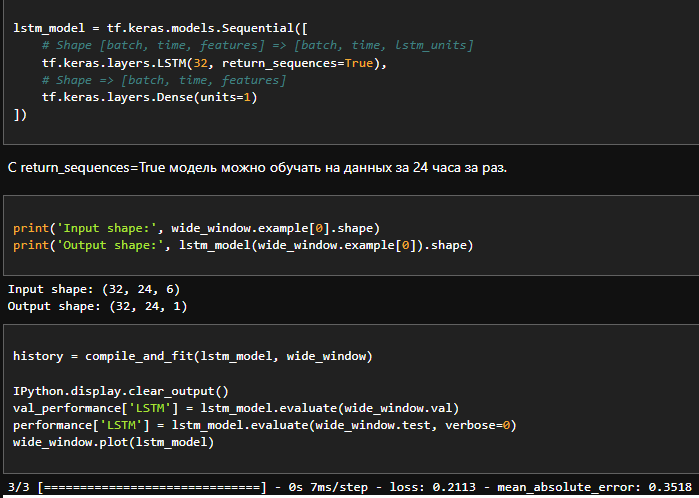


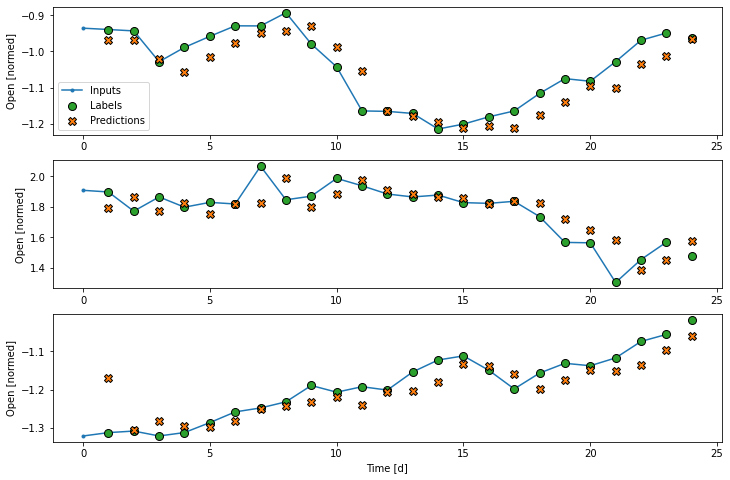


Теперь мы можем построить прогнозы модели в более широком окне. Обратим внимание на 3 шага входного времени перед первым прогнозом. Каждый прогноз здесь основан на 3 предыдущих временных шагах:



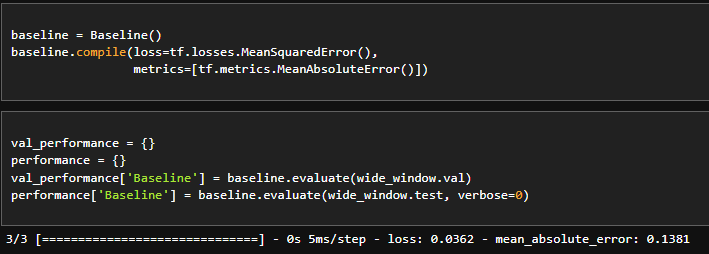
**Реккурентная нейронная сеть(RNN)**



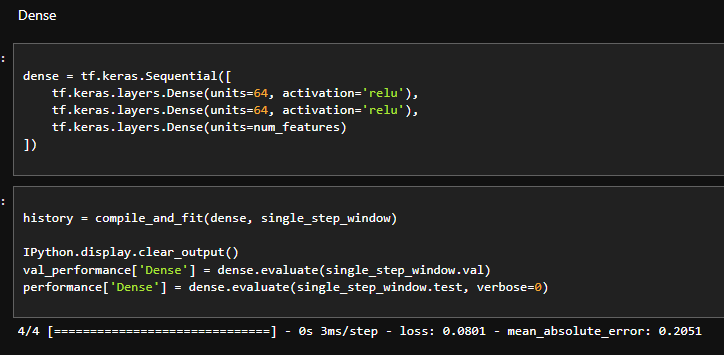


**Модели с несколькими выходами**

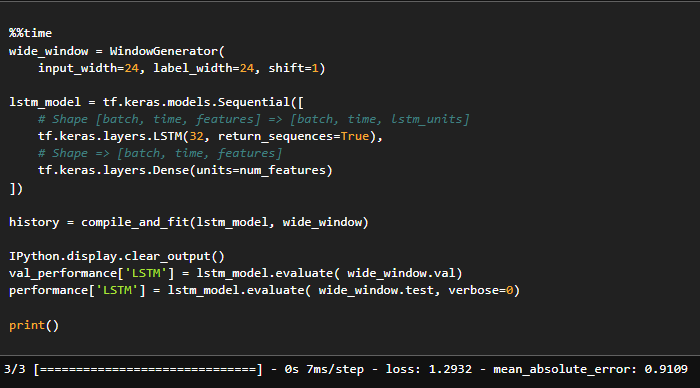
Baseline



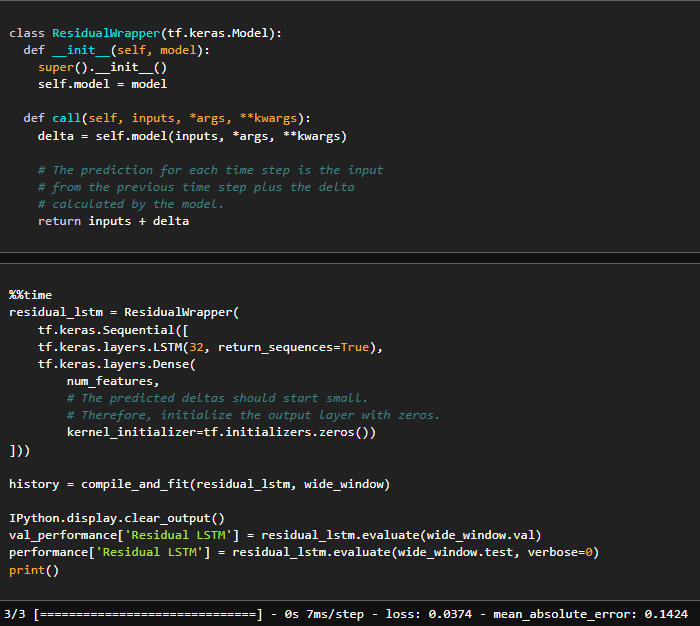
Dense



LSTM



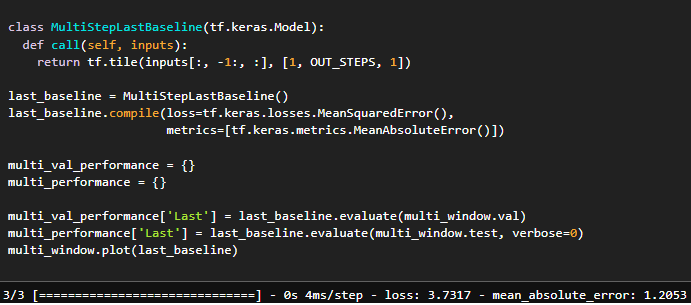
Residual LSTM

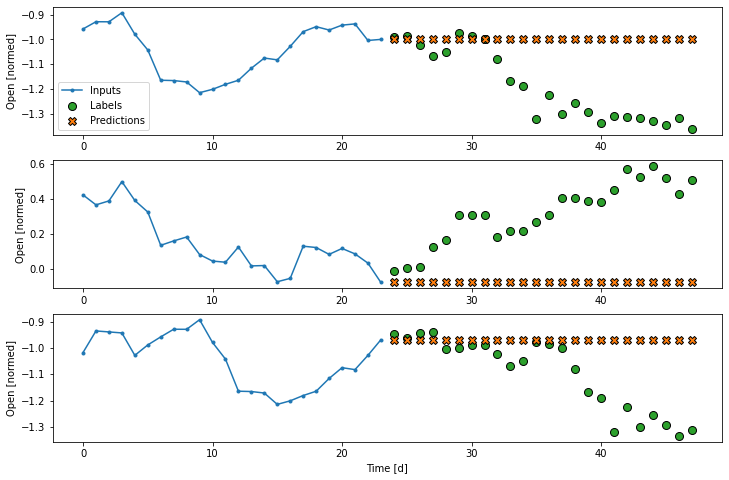


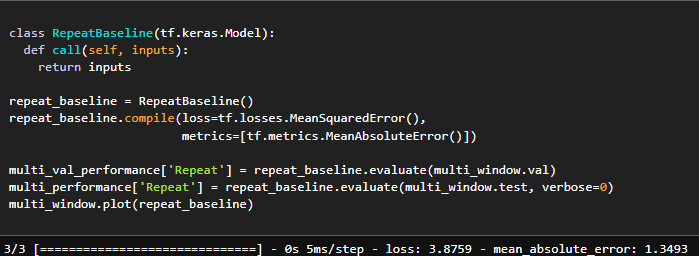
**Многошаговые модели**

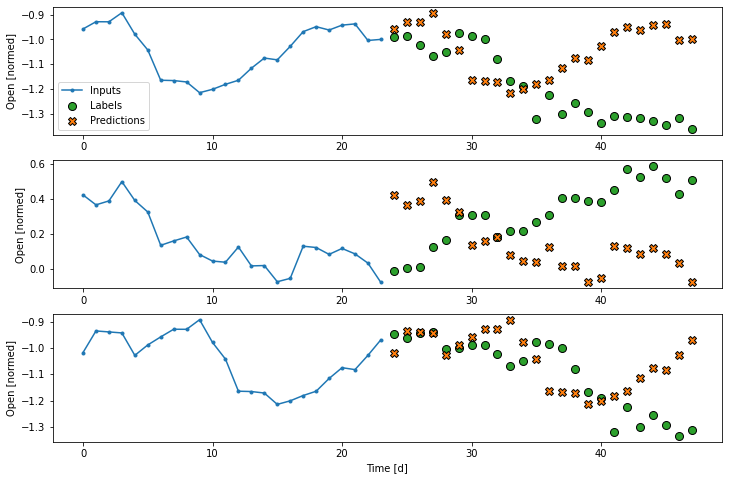


**Baseline**

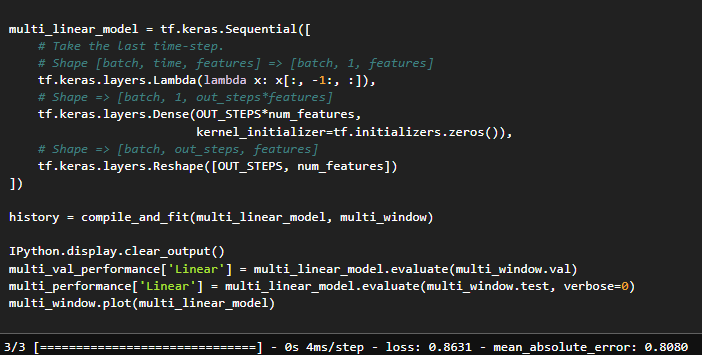


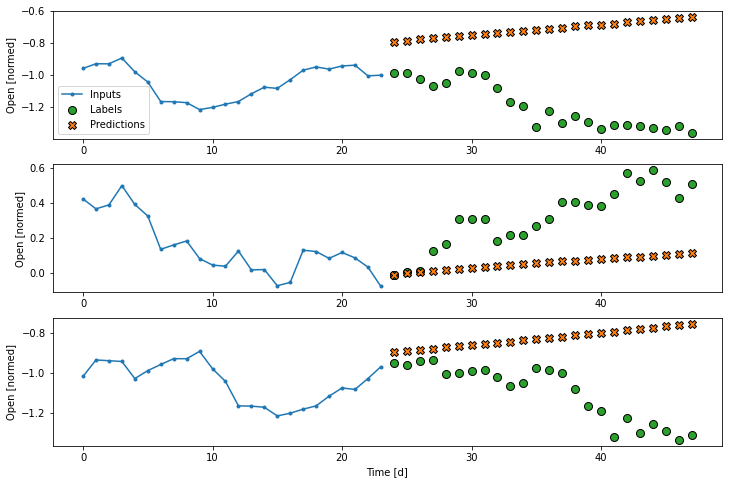


**RepeatBaseline** 

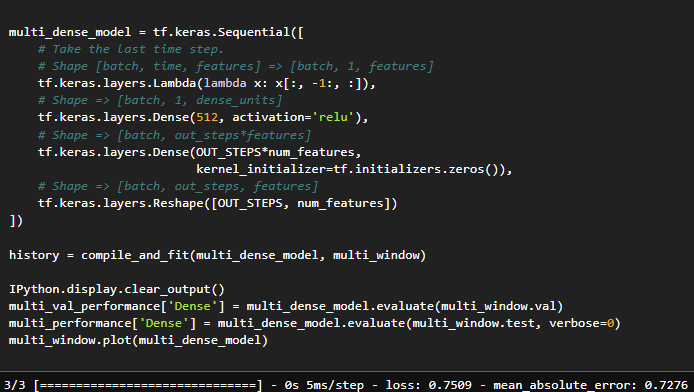


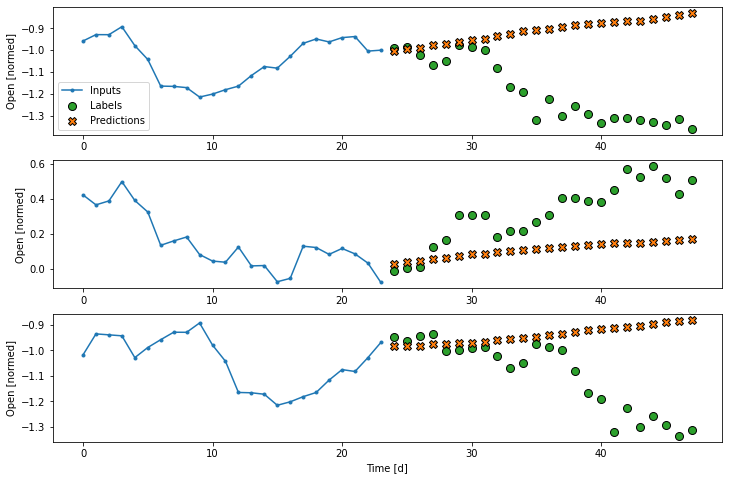
**Linear**



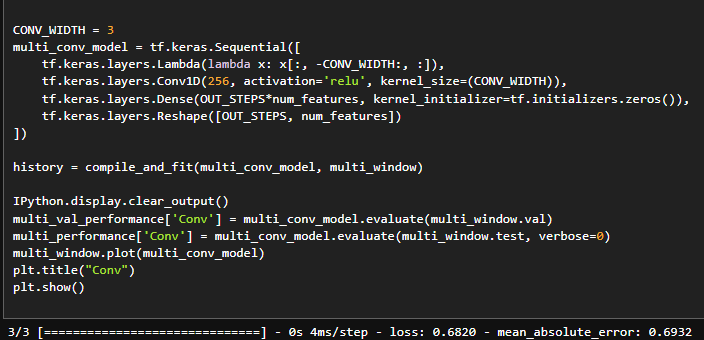


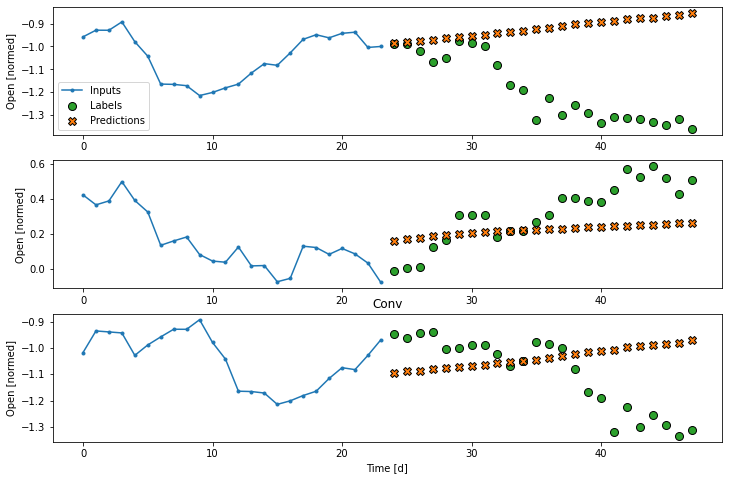
**Dense**



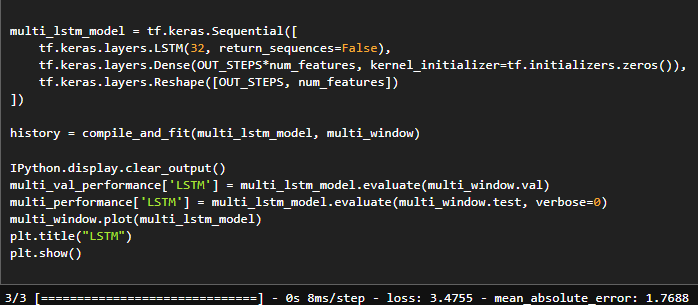


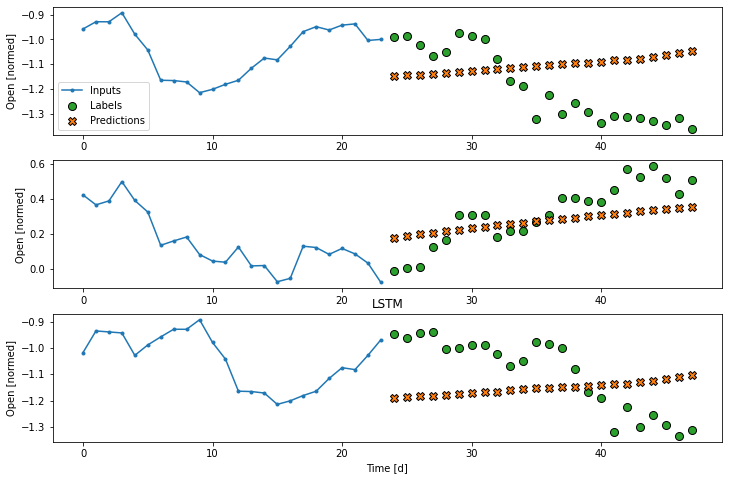
**Сверточная(CNN)**



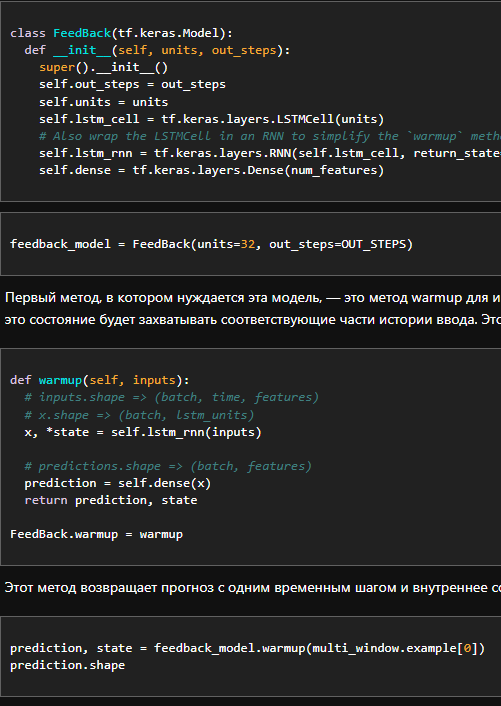
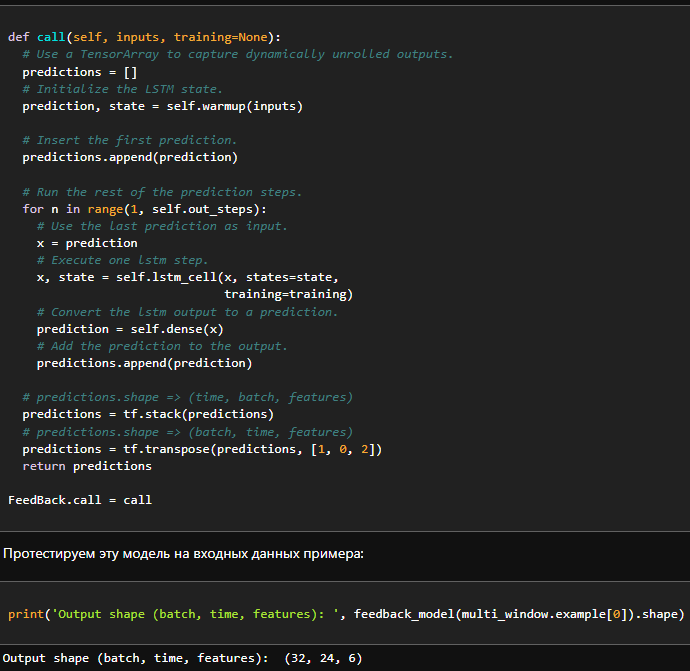
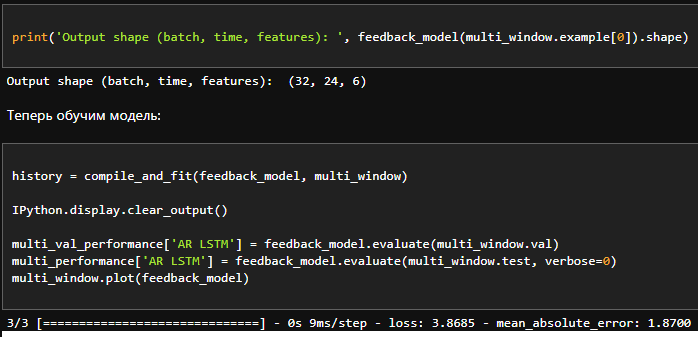


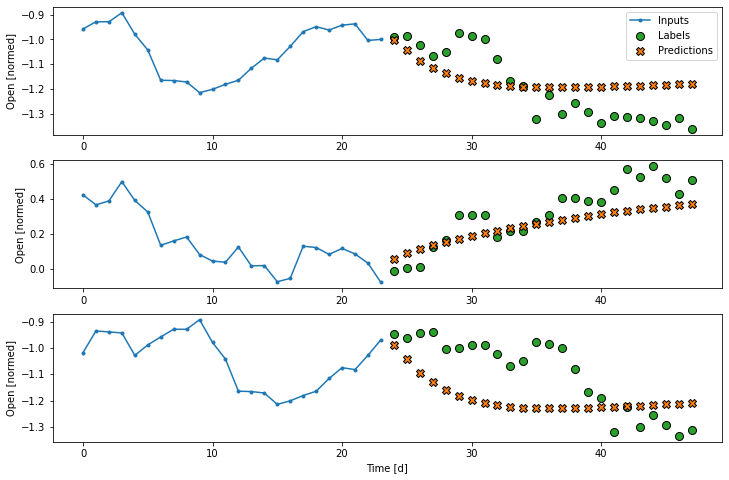
**Реккурентная(RNN)**





**Авторегрессионная модель(Autoregressive model)**

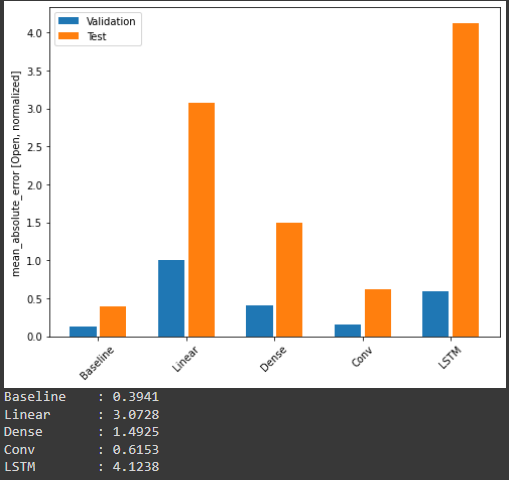


# **4 Результаты работы**

Были разработаны модели для одношаговых и многошаговых моделей нейронных сетей.

**Для одношаговых моделей**.

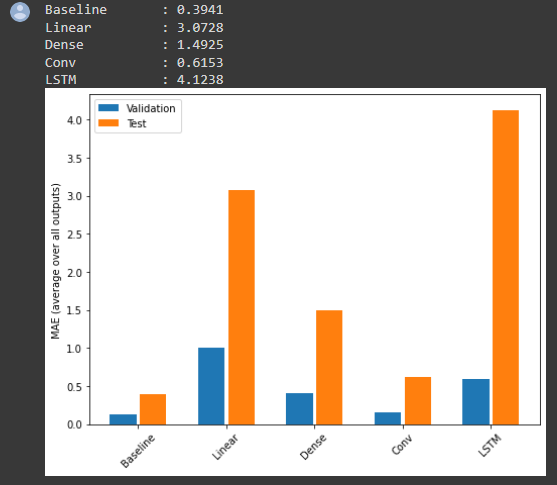
Представим величину ошибки всех моделей на графике и проанализируем их



Самой выигрышной является базовая модель. Выигрыш, достигнутый при переходе от плотной модели к многоплотной и сверточной, составляет всего несколько процентов (если вообще есть), а LSTM модель показала себя явно хуже, как и линейная модель. Таким образом, эти более сложные подходы могут быть бесполезными для решения этой проблемы, но невозможно было узнать, не попробовав, и эти модели могут быть полезны для решения проблемы.

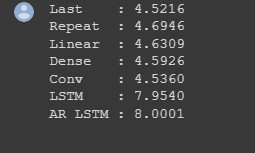
Из графика видно, что LSTM показала себя лучше на проверочной выборке, но очень плохо на тестовом сете. В остальных случаях

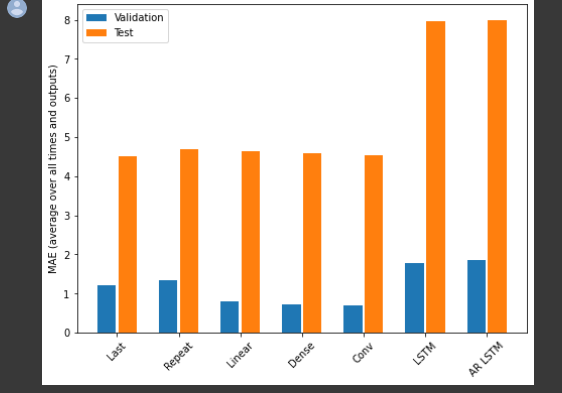
Для моделей с несколькими выходами величина ошибки и график выглядя следующим образом.



Самой выигрышной является базовая модель. Выигрыш, достигнутый при переходе от базовой модели к остаточной реккурентной, составляет около половины процента, а LSTM модель показала себя явно хуже, как Dense модель.

**Для многошаговых.**





Все модели не проявили себя.

# **ЗАКЛЮЧЕНИЕ**

В ходе выполнения лабораторной работы были изучены основные принципы работы с обучением нейронных сетей. Обучены различные нейронные сети на наборе из 3526 данных.

Проведены эксперименты по сравнению различных моделей нейронных сетей.

Явно наблюдается увеличение отдачи в зависимости от сложности модели.

Метрики для моделей с несколькими выходами в первой половине этого руководства показывают среднюю производительность по всем выходным функциям. Эти характеристики аналогичны, но также усреднены по временным шагам вывода.

В результате было выявлено, что одношаговые модели ведут себя гораздо лучше, чем многошаговые модели на данном наборе данных.

# **ССЫЛКИ**

Набор данных:

<https://finance.yahoo.com/quote/AAPL/history?period1=1167609600&period2=1609459200&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true>

GitHub: <https://github.com/MikhailKalikin/CNN_RNN_LAB2>