

Предсказание курса на валютном рынке с помощью машинного обучения

С.В. Костин¹, А.В. Гайдель^{1,2}

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

²Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация. В работе исследуется эффективность двух алгоритмов машинного обучения для предсказания курса валют. Первый алгоритм основан на линейной регрессии, а второй – на рекуррентной нейронной сети с длинной кратковременной памятью. Главным преимуществом данных алгоритмов является то, что они хорошо подходят для прогнозирования временных рядов. Для обучения и построения модели регрессии применяется метод скользящего окна, позволяющий использовать предыдущие временные шаги для прогноза следующего шага. В качестве оценки полученных моделей используются коэффициент детерминации, среднеквадратическая ошибка потерь регрессии и средняя абсолютная ошибка потерь регрессии. Оба алгоритма показывают сравнимые результаты.

1. Введение

В современном мире с большой остротой проявляется интерес к качественному прогнозированию финансовых рынков. Это связано с быстрым развитием высоких технологий и, соответственно, с появлением новых инструментов анализа данных.

Проблемы прогнозирования временных рядов являются сложным типом задачи прогнозного моделирования. В отличие от прогнозного регрессионного моделирования, временной ряд также добавляет сложность зависимости последовательности между входными переменными [1]. В последние годы, у финансовых аналитиков стали вызывать большой интерес, так называемые искусственные нейронные сети – это математические модели, а также их программные или аппаратные реализации, построенные по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма. Впоследствии эти модели стали использовать в практических целях, как правило, в задачах прогнозирования. Актуальностью данной работы является сравнение традиционного алгоритма линейной регрессии с алгоритмом рекуррентной нейронной сети. Возможность обучения – одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Способности нейронной сети к прогнозированию напрямую следуют из ее способности к обобщению и выделению скрытых зависимостей между входными и выходными данными [1]. После обучения сеть способна предсказать будущее значение некой последовательности на основе нескольких предыдущих

значений или каких-то существующих в настоящий момент факторов. Следует отметить, что прогнозирование возможно только тогда, когда предыдущие изменения действительно в какой-то степени предопределяют будущие. Например, прогнозирование котировок акций на основе котировок за прошлую неделю может оказаться успешным, тогда как прогнозирование результатов завтрашней лотереи на основе данных за последние несколько лет почти наверняка не даст никаких результатов.

2. Регрессионная модель

Пусть объекты наблюдения характеризуются некоторыми признаками X_j , и интерес будет представлять другой признак этих объектов Y . Тогда требуется научиться по набору наблюдаемых признаков X_j , определять значение признака Y . Предполагается, что признак Y каким-то образом зависит от признаков X_j . В качестве набора признаков X будем использовать стоимость курса за предыдущий временной период. Если считать все эти признаки случайными величинами, то можно определить регрессию Y на X , как условное математическое ожидание:

$$y(x) = M[Y | X = x].$$

График функции $y(x_1, x_2, \dots, x_m)$ будем называть линией регрессии, признаки X_j – независимыми переменными (также регрессорами или предикторами), а признак Y – зависимой или выходной переменной.

Задача восстановления линии регрессии заключается в построении некоторой функции $y(x_1, x_2, \dots, x_m)$, которая наилучшим образом описывает данные из некоторой обучающей выборки U , в которой каждому вектору предикторов X ставится в соответствие зависимая переменная Y . Понятно, что эту функцию следует искать в некотором конкретном классе функций.

Наиболее популярным подходом является предположение, что данные отвечают линейной регрессионной модели вида:

$$y(x_1, x_2, \dots, x_m) = a^T x + b + \varepsilon.$$

где $a \in R^m$ и $b \in R$ – это неизвестные коэффициенты в уравнении регрессии, ε – гауссовский белый шум с нулевым средним и некоторой дисперсией σ^2 .

Для заданной выборки U , состоящей из пар (x, y) , оптимальные коэффициенты a и b определяются из условия минимума суммы квадратов отклонений:

$$J(a, b) = \sum_{(x,y) \in U} (y - \hat{y}(x_1, x_2, \dots, x_m))^2 \rightarrow \min.$$

Здесь подразумевается: $\hat{y}(x) = a^T x + b$. Подобный метод приближения данных с помощью кривых называется методом наименьших квадратов.

Для оценки качества приближения данных с помощью заданной регрессионной модели используется коэффициент детерминации, определяемый как:

$$R^2 = 1 - \frac{J(a,b)}{S_0},$$

где $J(a, b)$ – это сумма квадратов отклонений от линии регрессии, S_0 – это общая сумма квадратов отклонений от выборочного среднего. Чем больше значение коэффициента детерминации R^2 , тем лучше данные описываются заданной моделью. В случае, когда $R^2 = 1$, все точки из выборки в точности лежат на линии регрессии.

В одномерном случае x – это просто один числовой признак, так что модель имеет вид:

$$y(x) = ax + b + \varepsilon.$$

3. Рекуррентная модель нейронной сети

Рекуррентная нейронная сеть, на фундаментальном уровне, является типом плотно связанной нейронной сети. Однако ключевым отличием от обычных сетей прямой связи является введение времени - в частности, выход скрытого слоя в рекуррентной нейронной сети возвращается обратно в себя.

На рисунке выше мы видим простую рекуррентную нейронную сеть, входные узлы которой подаются в скрытый слой с сигмовидными активациями, как в любой нормальной плотно связанной нейронной сети. Далее вывод скрытого слоя возвращается в тот же скрытый слой.

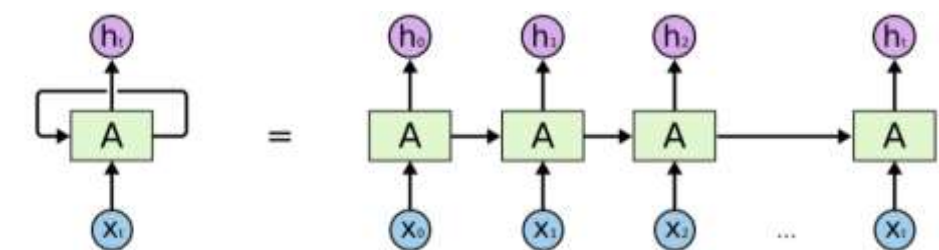


Рисунок 1. Рекуррентная нейронная сеть.

Выходы скрытого слоя пропускаются через блок концептуальной задержки, чтобы разрешить ввод в скрытый слой для возможности моделирования данных, которые зависят от времени или последовательности. Для рекуррентных нейронных сетей мы хотели бы иметь длинную память, чтобы сеть могла соединять отношения данных на значительные расстояния во времени. Такая сеть может добиться реального прогресса в понимании того, как соотносятся события на валютном рынке. Однако, чем больше временных шагов, тем больше вероятность того, что градиенты обратного распространения либо накапливаются, либо исчезают до нуля. Чтобы уменьшить исчезающую проблему градиента и, следовательно, позволить более глубоким сетям и рекуррентным нейронным сетям хорошо работать в практических условиях, должен быть способ уменьшить умножение градиентов, которые меньше нуля. Ячейка LSTM это специально разработанная логическая единица, которая поможет в достаточной мере уменьшить проблему исчезающего градиента, чтобы сделать рекуррентные нейронные сети более полезными для задач долговременной памяти, то есть для предсказания курса на валютном рынке. Это происходит путем создания состояния внутренней памяти, которое просто добавляется к обработанному входу, что значительно уменьшает мультипликативный эффект малых градиентов [2].

Сеть с длинной кратковременной памятью, или сеть LSTM, представляет собой рекуррентную нейронную сеть, которая обучается с использованием обратного распространения через время и преодолевает проблему исчезающего градиента. Сеть LSTM могут использоваться для создания больших рекуррентных сетей для решения сложных проблем последовательности в машинном обучении. Вместо нейронов в сетях LSTM есть блоки памяти, которые связаны через слои. Блок содержит ворота, которые управляют состоянием и выходом блока. Блок работает с входной последовательностью, и каждый логический элемент в блоке использует блоки активации сигмоида для управления, что делает изменение состояния и добавление информации, протекающей через блок, условным [3].

4. Экспериментальные исследования

В качестве набора данных возьмем стоимость американского доллара и японской йены, предоставленные Центробанком за период с 20.04.2018 по 20.04.2019. Задача состоит в том, чтобы на основе предоставленной информации сделать прогноз на некоторое количество дней. На вход нашим моделям будем подавать вектор признаков, представляющий собой стоимость курса в рублях за известный временной период, размер вектора будем называть окном, и на основе него будем делать прогноз стоимости курса на следующий временной шаг. Например, учитывая стоимость валюты в текущее время (t), которое мы хотим предсказать в следующий раз в последовательности ($t + 1$), мы можем использовать значение в текущее время (t), а также в два предыдущих временных шага ($t - 1$) и ($t - 2$) в качестве входных переменных. Если сформулировать как проблему регрессии, входными переменными являются ($t - 2$), ($t - 1$), t , а выходной переменной - ($t + 1$). Исходный набор данных о стоимости американского доллара представлен в виде графика на рисунке 2.

Для обучения и оценки моделей разобьем набор данных на тренировочный и тестовый в соотношении 7 к 3, т.е. 70 процентов данных для обучения моделей и 30 процентов данных для оценки. Будем изменять размеры окна на значения 1,3,5,10. В качестве оценки полученных результатов вычислим коэффициент детерминации, среднеквадратическая ошибка потерь

регрессии и средняя абсолютная ошибка потерь регрессии. Полученные результаты изображены в таблицах 1,2,3,4.

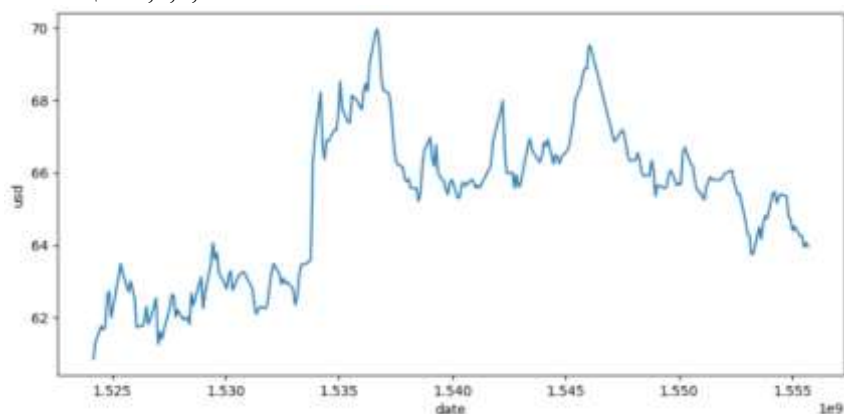


Рисунок 2. График стоимости доллара за период с 20.04.2018 по 20.04.2019.

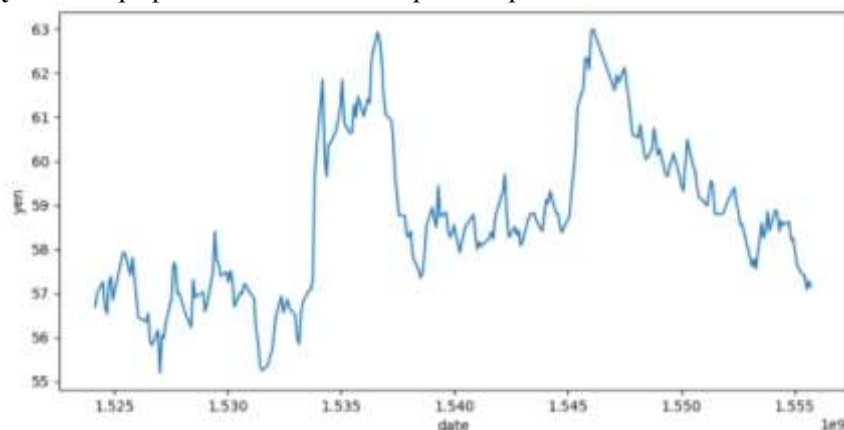


Рисунок 3. График стоимости йены за период с 20.04.2018 по 20.04.2019.

Таблица 1. Оценка модели линейной регрессии для американского доллара.

Н - ширина окна	Н = 1	Н = 3	Н = 5	Н = 10
R2 score	0.894	0.894	0.891	0.891
Absolute error	0.267	0.263	0.262	0.262
Squared error	0.164	0.165	0.165	0.165

Таблица 2. Оценка рекуррентной модели нейронной сети для американского доллара.

Н - ширина окна	Н = 1	Н = 3	Н = 5	Н = 10
R2 score	0.895	0.890	0.891	0.891
Absolute error	0.274	0.285	0.280	0.262
Squared error	0.163	0.172	0.165	0.165

Таблица 3. Оценка модели линейной регрессии для японской йены.

Н - ширина окна	Н = 1	Н = 3	Н = 5	Н = 10
R2 score	0.919	0.922	0.929	0.920
Absolute error	0.310	0.305	0.292	0.278
Squared error	0.160	0.156	0.144	0.129

В приведенных выше таблицах параметр Н - является шириной окна, т.е. стоимость курса за количество предыдущих дней, R2 score – коэффициент детерминации, Absolute error - средняя абсолютная ошибка потерь регрессии, Squared error - среднеквадратическая ошибка потерь регрессии.

Таблица 4. Оценка рекуррентной модели нейронной сети для японской йены.

Н - ширина окна	Н = 1	Н = 3	Н = 5	Н = 10
R2 score	0.921	0.924	0.927	0.904
Absolute error	0.315	0.307	0.292	0.300
Squared error	0.156	0.153	0.147	0.155

5. Заключение

Таким образом, в данной работе были построены и исследованы модели для предсказания курса валют, с использованием линейной регрессии и рекуррентной нейронной сети. Для формирования обучающей и тестовой выборки данных был использован метод окна со сдвигом на несколько временных шагов. Для оценки полученных моделей были вычислены коэффициент детерминации, средняя абсолютная ошибка потерь регрессии, среднеквадратичная ошибка потерь регрессии, для каждого временного окна. Исходя из этих оценок, можно сделать вывод, что полученные модели обладают высокой точностью и пригодны для использования на практике. Так же можно сделать вывод, что обе модели дают сравнимые результаты, и для предсказания курса на валютном рынке достаточно построить линейную регрессию.

6. Благодарности

Работа выполнена при поддержке грантов РФФИ № 19-29-01235 мк и № 19-29-01135 мк, экспериментальные исследования – в рамках госзадания ИСОИ РАН – филиала ФНИЦ «Кристаллография и Фотоника» РАН (соглашение № 007-ГЗ/Ч3363/26).

7. Литература

- [1] Мюллер, А. Введение в машинное обучение с помощью Python / А. Мюллер, С. Гвидо – М.: Питер, 2017. – 480 с.
- [2] Плас, Дж.В. Python для сложных задач: наука о данных и машинное обучение / Дж.В. Плас – М.: Питер, 2018. – 576 с.
- [3] Рашка, С. Python и машинное обучение / С. Рашка – М.: ДМК-Пресс, 2017. – 418 с.
- [4] Малов, С.В. Регрессионный анализ. Теоретические основы и практические рекомендации / С.В. Малов – С.-Пб.: СПбГУ, 2013. – 276 с.
- [5] Бринк, Х. Машинное обучение / Х. Бринк, Д. Ричардс, М. Феверолф – М.: Питер, 2017. – 336 с.
- [6] Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах – Москва: ДМК-Пресс, 2017. – 567 с.

Predicting exchange rate dynamics in the forex market using machine learning

S.V. Kostin¹, A.V. Gaidel^{1,2}

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

Abstract. The paper investigates the effectiveness of two machine learning algorithms for predicting the exchange rate. The first algorithm is based on linear regression, and the second is based on a recurrent neural network with long short-term memory. The main advantage of these algorithms is that they are well suited for predicting time series. To train and build a regression model, the sliding window method is used, which allows you to use the previous time steps to predict the next step. As an assessment of the obtained models, the coefficient of determination, the standard error of the regression losses, and the average absolute error of the regression losses are used. Both algorithms show comparable results.