

АНАЛІЗ МОЖЛИВОСТІ ПРОГНОЗУВАННЯ ЦІН

Хміль О. С.

Науковий керівник — проф. Ситніков Д.Е.

Харківський національний університет радіоелектроніки
(61166, Харків, просп. Науки, 14, каф. Системотехніки, тел. (057) 702-00-00)

This article discusses the situation of deep learning models in the stock market in our time, describes the problems that arise when using these models, the reasons for their occurrence, and ways to prevent them. It is justified that mainly should be used historical price information to build a predictive model rather than other information from various sources. There have been proposed algorithms for use and implementation. Mostly on the simplest algorithms due to their speed of analysis and, as a result, efficiency. And as a result, a conclusion is drawn on development trends.

Согласно данным различных новостных ресурсов, которые анализируют рынок, — примерно 80% объема торговли на фондовом рынке США генерируется алгоритмами. Если «машины» занимают доминирующее положение, то могут ли быть шансы получить прибыль у человека не использующего данные алгоритмы?

Прежде всего, большинство этих алгоритмов используют очень простые методы, которые являются не более чем цепочкой закодированных правил или обычной моделью линейной регрессии, а нацеленность этих алгоритмов — спрогнозировать на миллисекунды вперед.

Для более долгосрочного прогнозирования следует использовать модели глубокого обучения, которые способны отыскивать куда более сложные паттерны в данных. Но в данный момент возможность такого прогноза сомнительна и еще не подтверждена. Тем не менее, крупные финансовые компании содержат огромное количество инженеров и специалистов по машинному обучению занимающихся решением этой задачи. Из чего можно сделать вывод об тенденциях в инвестировании. Также стоит понимать что инвестиции крупных компаний сильно отличаются от индивидуального инвестирования, чем меньше объем приобретаемых акций тем меньше влияние оказывается на рынок.

Рассмотрим какие способы можно применить для анализа рынка. Самые простые это LSTM (Long short-term memory) или Temporal Convolutional Networks и их использование на истории цен за определенный период. Из сложных способов — применять модель Convolutional LSTM для анализа текста из различных источников, таких как веб-сайты финансовых новостей, социальные сети или тому подобное, чтоб решить пойдет ли конкретная акция вверх или вниз в ближайшем будущем.

Однако существует гипотеза эффективного рынка, которая гласит что цены на акции отражают всю имеющуюся информацию на данный момент. И вполне вероятно что любая новая информация, которая может повлиять на цены, уже включена в неё к тому времени, когда модель начнет анализировать её и если следовать этой гипотезе то использование в анализе любых данных помимо цен является излишним. Из чего можно сделать вывод что для анализа следует использовать только информацию об истории цен для построения прогностической модели.

Люди довольно давно пытаются спрогнозировать направление цен путем нахождения закономерностей в данных состояния рынка в прошлом, что по своей сути является формой технического анализа. Часто люди смотрят на диаграммы пытаясь увидеть некие повторяющиеся паттерны. Для этой задачи возможно использовать нейронные сети. Но если бы действительно паттерны существовали, то нейронные сети достаточно легко смогли бы найти их. Проблема в том, что даже если в прошлых данных были найдено некий паттерн — это еще не означает что эти паттерны смогут повториться в будущем. Модель может замечательно объяснять примеры из значений в прошлом, но не иметь никакой предсказательной силы — это называется переобучение. И в результате затраченных усилий будет получена модель, которая будет выполнять функцию справочной таблицы. Нейронная сеть даже может найти закономерности в случайно сгенерированных данных, но то чему научиться нейронная сеть будет бесполезным для прогнозирования. Как человек может увидеть лица или объекты в облаках — так и модель машинного обучения может находить паттерны в совершенно случайных данных из-за чего следует быть осторожным с техническим анализом.

Особое внимание следует обратить на то как логарифмическая модель степенного закона смогла спрогнозировать пузырь Биткойна в 2018 году за 8 дней до его взрыва, что не означает что она сможет предсказать следующий пузырь или сработать на другой криптовалюте. Тем не менее, это интересное наблюдение и его следует принять во внимание.

Резюмируя можно сделать вывод о том что в данный момент возможность создания модели позволяющей спрогнозировать тенденцию на достаточно долгий период — маловероятно, но в любом случае если она будет создана, то это неизменно окажет огромное влияние на рынок. Тем не менее это не значит что нельзя создать и обучить модель на выявление, например, сезонных тенденций или тому подобных паттернов и предсказывать цену с определенной погрешностью. Но а влияние простых алгоритмов, которые и так уже занимают доминирующее положение, будет со временем расти, что и происходит на протяжении нескольких лет и судя по прогнозам так будет происходить и дальше. Ведь в связи с их преимуществом на промежутке времени в несколько миллисекунд при достаточной мощности они не имеют конкурентов.