ФАКТОРЫ ФОРМИРОВАНИЯ ЦЕНЫ НА ПРЕДМЕТЫ ИСКУССТВА С ПРИМЕНЕНИЕМ ТЕКСТОВОГО АНАЛИЗА НОВОСТЕЙ В ТВИТТЕРЕ[1]

Е.А. Федорова, Д.В. Зарипова, И.С. Демин

DOI: 10.33293/1609-1442-2020-2(89)-114-131

В данной работе были подтверждены гипотезы о влиянии индекса настроений в сети Твиттер на ценообразование предметов искусства и разницу между предварительной оценкой экспертов и итоговой ценой аукциона. Гипотезы были протестированы с помощью выборки из 83 картин, выбранных на основе рейтингов интернет-ресурса ARTNET[2] о самых дорогих когда-либо проданных произведениях искусства за последние 10–15 лет. Выборка состояла из 25 художников, по каждому из них был составлен индекс настроений в сети Твиттер. Данный индекс был создан путем проведения сентимент-анализа каждого твита[3] о художнике по хэштегу за период от двух до четырех месяцев между анонсами продажи в открытых источниках и непосредственной продажей работы по двум словарям AFINN и NRC.

Ключевые слова: текстовый анализ, ценообразование, арт-объект, инвестиционный объект, Твиттер.

JEL: C53, Z11.

ВВЕДЕНИЕ

В 2017 г. на известном во всем мире аукционе Sotheby's была продана картина Леонардо да Винчи под названием «Спаситель мира». Эта сделка стала рекордной в сфере искусства, ее стоимость достигла 450 млн долл. с учетом премии агенту[4]. Примечательно, что эксперты оценивали ее стоимость до покупки предыдущим владельцем не более 40 млн долл., что составляет только 11-ю часть ее итоговой цены на аукционе сегодня. Однако еще до финальной сделки о покупке картины даже сам факт ее присутствия на торгах вызвал огромный ажиотаж в средствах массовой информации и социальных сетях. В первую очередь это было связано с тем, что картина – заказ французского короля – долгое время считалась утерянной и была найдена случайно. К тому же важным предметом обсуждений стало предположение, что ее подлинность не была достоверно доказана экспертами. Огромное число публикаций в СМИ и социальных сетях, рекордная цена продажи, но при этом низкая ожидаемая цена экспертов наталкивают на вопрос, существует ли взаимосвязь между итоговой стоимостью предмета искусства и освещением возможной сделки в социальных сетях.

© Федорова Е.А., Зарипова Д.В., Демин И.С., 2020 г.

Федорова Елена Анатольевна, д.э.н., профессор, Департамент финансов, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия; ecolena@mail.ru

Зарипова Диана Вилевна, бакалавр, факультет экономических наук, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия

Демин Игорь Сергеевич, д.э.н., профессор, Департамента анализ данных и принятия решений Финансовый университет при Правительстве РФ, Москва, Россия

[1] Статья подготовлена по результатам исследований, выполненных за счет бюджетных средств по государственному заданию Финуниверситета 2019 г., Москва, 2019.

[2] Интернет-проект ArtNet [Электронный ресурс]. URL: http://www.artnet.com/

[3] Словарь языка интернета.ru / под ред. М.А. Кронгауза. М.: АСТ-Пресс, 2016. С. 220. URL: https://ru.wiktionary.org/wiki/твит.

[4] Принадлежавшая Д. Рыболовлеву картина да Винчи ушла за рекордную сумму. 16.11.2017 [Электронный ресурс]. URL: https://www.rbc.ru/society/16/11/2017/5a0cd4409a79476c3877792.

Целью исследования является оценка влияния индекса настроений в Твиттер, на процесс ценообразования в искусстве. Данное исследование обладает научной новизной, так как: 1) был разработан индекс настроений в Твиттере, который предлагается рассматривать как фактор ценообразования предметов искусства; 2) результаты исследования доказывают то, что предметы искусства относятся к инвестиционным активам, поскольку новости, опубликованные в социальных сетях, влияют на их ценообразование.

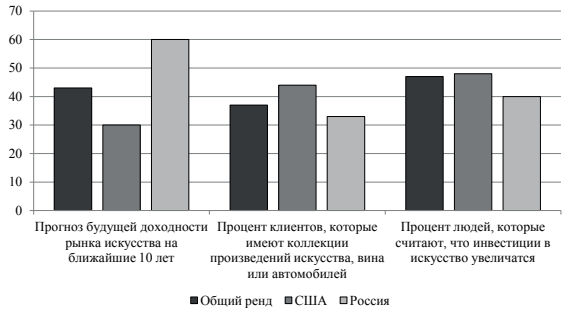

Рис. 1. Прогноз развития рынка на ближайшие 10 лет

Источник: Delloite: Art&Finance report 2017 [Электронный ресурс]. URL: https://www2.deloitte.com/lu/en/pages/art-finance/articles/art-finance-report.html

1. ОБЗОР ЛИТЕРАТУРЫ И ГИПОТЕЗЫ ИССЛЕДОВАНИЯ

Арт-рынок (т.е. рынок произведений искусств) – интересная сфера, которой начинает интересоваться все большее число людей. После экономических кризисов 2008 и 2011 гг. инвесторы начали больше вкладываться в рынок произведений искусства. Его особенности сегодня обусловлены тем, что общая прибыльность объекта во многом зависит не от мастерства художника, а от ряда определенных факторов, которые можно просчитать и предугадать. Этот процесс расчета и предсказаний будет представлен в данной работе.

В России рынок покупки и продажи работ в сфере искусства нельзя назвать высокоразвитой областью. Однако даже сейчас процесс развития арт-сферы набирает все более серьезные обороты, что можно заметить на примере растущего числа выставок как современного, так и более раннего искусства. Государственные учреждения, Третьяковская галерея, проводят выставки художников Куинджи, Бэнкси, Айвазовского и Мунка и собирают огромные очереди с момента открытия и до последнего дня работы экспозиций[5].

Из данных на рис. 1, представленных выше и опубликованных в ежегодном отчете *Deloitte&Touch* о финансах и искусстве, можно заключить, что Россия обладает большим потенциалом развития на рынке искусства. Так, например, предполагаемый рост будущего потенциала на рынке искусства, оцениваемый в 60%, уступает только азиатскому рынку, где число потенциальных покупателей в разы выше. В данный момент в Москве функционируют и офисы таких известных аукционных домов, как Sotheby's и Christie's. Одно из направлений их работы – поддержка отечественных производителей и выявление талантливых художников для дальнейшей работы с ними на мировом уровне. В Великобритании, США и Франции аукционный бизнес особенно развит, известные компании формируют каталоги и проводят открытые и закрытые аукционы для продажи полотен и скульптур. В Москве подобные события собирают большое число непосредственных участников и просто любителей и ценителей искусства. Например, ежегодная и единственная в России арт-ярмарка Cosmoscow нацелена на ознакомление коллекционеров нашей

[5] В. Мединский, Министр культуры, рассказал о появлении интернет-мемов про очереди в музей. 21.03.2018 [Электронный ресурс]. URL: https://www.rbc.ru/society/21/03/2018/5ab22e3e9a7947d8a494bace#ws.

страны и стран СНГ с новыми современными веяниями арт-рынка[6].

Ценообразование на этом рынке, процесс формирования цен в искусстве далеко неоднозначены. На сегодняшний день арт-сфера является популярным выбором для огромного числа инвесторов. Поведение людей на рынке искусства сегодня показывает, что покупатели воспринимают в основном картины в качестве инвестиционных объектов. Большого объема информации о динамике цен в прошлом, художественных индексах и аналитических отчетах о состоянии рынка недостаточно для того, чтобы сделать достаточно достоверный прогноз будущих цен. Для построения успешных стратегий на арт-рынке важно понимать механизмы ценообразования, особенности как глобального, так и регионального рынков, современные тенденции, а также другие факторы, влияющие на изменения цен. Сегодня технологический процесс и постоянно возрастающая скорость жизни в современном мире сделали социальные сети неотъемлемой и даже рутинной частью повседневности. В связи с этим многие исследователи сталкиваются с необходимостью анализа текстов, опубликованных в социальных сетях, для получения более точных оценок. Эмоциональный анализ публикаций в Твиттере и других социальных сетях уже несколько лет успешно применяется, например, для прогнозирования результатов политических выборов (Wang, Kazemzadeh et al., 2012) в других отраслях. В данной работе исследуется влияние настроений пользователей в социальной сети Твиттер в качестве одного из основных факторов ценообразования на произведения искусства. В работе выдвинут ряд следующих гипотез.

Гипотеза 1. Индекс настроений (позитивная окраска) в публикациях в сети Твиттер положительно влияет на аукционную стоимость произведений искусства.

На основе анализа литературы можно прийти к выводу о том, что новости влияют на ценообразование финансовых активов (Agarwal et al., 2016). Исследования, описанные выше, получили положительный результат – обнаружена зависимость между сентиментной, эмоциональной окраской новостей и продажей финансовых активов на бирже (Bollen, Mao, Zeng, 2011). Можно также отнести произведения искусства к финансовому активу, так как многие коллекционеры и бизнесмены, участвующие в торгах, совершают покупки в целях инвестирования.

Данная гипотеза поможет оценить влияние индекса настроений в сети на финальную стоимость арт-объекта. Она была сформирована ввиду отсутствия эффективных предсказательных моделей для определения цены на произведения искусства. Однако все чаще можно заметить, что художники, которые считаются популярными среди публики, т.е. собирают большое число лайков и откликов в социальных сетях, также становятся все более востребованными на этом рынке. Например, популярность, которую достиг известный уличный художник Бэнкси через социальные сети, позволяет ему продавать картины на аукционах и получать за каждую более миллиона долларов. Таким образом, в данной работе предполагается, что для успешного формирования стоимости произведения искусства, помимо стандартных факторов состояния и параметров работы, данных о художнике и истории картины, необходимо также применять индекс настроений в сети.

Следующая гипотеза, которая тестируется в данном исследовании, также затрагивает индекс настроений в сети Твиттер, но отражает влияние на разницу между предварительной и финальной ценой.

Гипотеза 2. Разница между предварительной оценкой и итоговой ценой арт-объекта находится в линейной зависимости от позитивной

[6] Официальный сайт международной ярмарки современного искусства Cosmoscow [Электронный ресурс]. URL: https://www.cosmoscow.com/ru/visitors/.

и негативной эмоциональной окраски публикаций в Твиттере*:* чем позитивнее твиты о художнике за период до реальных продаж*, тем большей* будет разница между аукционной и ожидаемой ценой.

Данная гипотеза может объяснить существенные различия между предварительной оценкой экспертов и реальной ценой произведения искусства. Исследования, проведенные ранее, доказывали, что существует реальное влияние эмоционально окрашенных публикаций на изменение цен на финансовые активы. Автор выдвигал также гипотезу о манипулировании ценами с помощью публикаций генеральных директоров компаний в Твиттере (Chan, 2003). Как правило, стоимость работы художника после торгов часто попадает в диапазон, который определяют эксперты-оценщики. Однако данное правило не распространяется на самые популярные полотна, часто их цена во много раз превышает ожидания. Данная гипотеза основана на предположении о том, что на финальную стоимость подобных работ влияет мнение людей. Поэтому и разница между оценкой, основанной только на характеристиках самой картины, и финальной суммой сделки оказывается существенной. В основу данного предположения входят результаты продажи самых дорогих полотен, которые были на слуху у большинства людей до продажи, представляли огромную историческую и художественную ценность и широко обсуждались в Интернете.

В случае положительного результата интересно предположить, что существует возможность косвенно управлять ценой на арт-объекты путем создания ажиотажа в социальных сетях и средствах массовой информации.

Подтверждение первых друг гипотез может свидетельствовать о поведении покупателей на рынке искусства. Таким образом, можно понять, являются ли они больше ценителями или инвесторами, когда совершают такие покупки.

2. МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

2.1. Текстовый анализ

Существуют два вида анализа текстов, которые помогают человеку анализировать и интерпретировать данные: количественный и качественный. Количественный анализ вполне понятен с точки зрения экономики: он заключается в построении диаграмм, отчетов и таблиц на основе текста. Качественный анализ имеет более сложную структуру, так как он основан на выявлении информации непосредственно внутри самого текста. Это направление активно развивается на основе методов машинного обучения, так как огромный объем информации не позволяет анализировать данные вручную.

Текстовый анализ – автоматизированный процесс, который позволяет выделять и классифицировать информацию из таких текстовых источников, как электронные письма, новости, публикации и твиты. Данный анализ пользуется особой популярностью в сфере бизнеса, так как помогает выделить специальную для ее информацию, например о новом продукте или мероприятии, и в дальнейшем разделить ее на позитивную и негативную. Возможности текстового анализа позволяют работать компаниям все более эффективно. Объем текстов, которые должны быть подвергнуты машинному обучению, различается в зависимости от заданных факторов и результатов, которые необходимо достичь. Так, например, для обучения компьютерной системы задаче определения темы сообщения требуется примерно по 250 примеров из каждой области, а для проведения качественного сентимент-анализа, о котором будет говориться далее, нужно не менее 500 примеров на каждую отдельную эмоцию.

На сегодняшний момент вследствие широкого распространения текстового анализа существует большое число техник для его проведения. Некоторые из этих методов анализа встречаются и используются ежедневно.

К наиболее распространенным и продвинутым методам относится классификация текста. Классификация текста – процесс присвоения тексту определенных тегов и категорий. Это один из наиболее полезных инструментов машинного обучения, так как он позволяет структурировать абсолютно любой заданный текст.

Подробнее стоит остановиться на одном методе – именно на текстовом анализе (сентимент-анализе), так как он будет использован в дальнейшем в данной работе. Эмоциональная окраска текста зачастую характеризует предмет изложения больше, чем его основное содержание: так, очевидно, что одно и то же событие люди с разными взглядами будут описывать с разными эмоциями. Выявление эмоций широко используется в бизнесе, например, чтобы выявить отношение покупателей к новому продукту через отзывы в Интернете. Существует ряд методов, которые позволяют распознавать эмоции и группировать тексты по их тональности. Сентимент-анализ – автоматизированный процесс выявления признаков эмоций из письменных текстов или устного языка. При этом, как правило, используются словари-классификаторы, в которых уже имеется оценка эмоциональной окраски слов (Демин, Рогов, Федорова, 2019). Классификатор настроений часто включают в оценку репутации бренда, исследования рынка, а также в определение настроений масс для проведений связанных исследований, о чем будет более подробно рассказано далее.

Особенностью предсказательных моделей, которые используются в данной работе, является наличие в них индекса настроения Твиттера. Анализ литературы по теме показывает, что чаще всего в качестве данного значения рассчитывается отношение позитивных твитов к негативным, деленным на их общее число. Однако словари, которые использовались для сентимент-анализа в данной работе, представляют собой более сложную систему и позволяют определить более глубокий спектр эмоций.

Текстовый анализ твитов в данном исследовании был проведен на языке *R* с использованием двух бесплатных словарей тональности – NRC и AFINN. Второй словарь был специально создан для работы именно с социальной сетью Твиттер.

Выборка текстовых данных, взятая из социальной сети Твиттер, была проанализирована по определенным параметрам. Каждый твит представлял собой набор слов, из которого были удалены дополнительные хэштеги, упоминания других профилей, вебссылки, а также стоп-слова. Так как все анализируемые твиты были исключительно на английском языке, то к стоп-словам относятся артикли, цифры и предлоги. Таким образом, каждая публикация представляется в виде набора слов. Далее с помощью словарей определения тональности оценке присваивается некоторая эмоциональная окраска.

Словари, которые используются для определения тональности текста, различаются между собой, поэтому необходимо определять, какие из них больше всего подходят для проведения конкретного исследования. Например, они могут различаться по специфике работы: одни содержат набор слов, распределенный по эмоциям на основе экспертных мнений, другие берут за основу методы машинного обучения, которые с помощью предварительных тестовых выборок учатся определять тональность текста. Большое значение имеют объем словаря, число признаков тональности, а также его направленность: есть словари, предназначенные для анализа бизнес-информации и интернет-публикаций.

Так, словарь AFINN, который был разработан профессором Университета Дании Финном Орупом Нельсеном, был создан экспертами вручную и содержит 2476 слов, поделенных на позитивные и негативные значения. Специфическая особенность словаря AFINN состоит в том, что он не только делит слова на положительные и негативные, но и указывает степень/уровень их положительности или негативности: для каждого термина указана степень/уровень позитивности

по шкале от –5 (термины с наиболее отрицательным значением) до 5 (наиболее позитивно окрашенные слова). Например, слово *curious* («любопытный», «забавный») имеет позитивность 1, а слово *terrific* («потрясающий») – позитивность 4. Одной из ключевых причин использования этого словаря в работе является наличие в нем жаргонных слов, ругательств и сильно эмоционально окрашенных слов, которые встречаются именно в постах социальных сетей или микроблогах, но не в формальных источниках.

Другой словарь для определения эмоциональной тональности теста – NRC. Он состоит из 6468 терминов, причем каждый имеет от 1 до 10 тональностей. В них включаются:

- позитив (positive);
- негатив (negative);
- злость (anger);
- ожидание (anticipation);
- отвращение (disgust);
- страх (fear);
- радость (joy);
- печаль (sadness);
- удивление/сюрприз (surprise);
- доверие (trust).

Данный словарь выделяется по числу слов, а также по числу хорошо интерпретируемых эмоций, что позволяет углубить текстовый анализ.

2.2. Ценообразование и факторы оценки произведений

Существует много работ об эффективной и адекватной оценке произведений искусства, которые, например, выделяют общие категории для оценки картин, выполненных в технике живопись: размеры холста; сохранность; имя художника; сюжет; материал; провенанс (история картины от момента создания до продажи: предыдущие владельцы, участие работы на выставках); инвестиционная привлекательность работы; качество работы; композиция; историко-документальное значение; мода на рынке. В данной работе будут исследованы только некоторые из них, чтобы добиться максимально эффективной оценки при использовании минимального числа параметров, а также введение нового, ранее неиспользованного в работах фактора – сентимент-анализа.

Принципы ценообразования в арт-индустрии существенно отличаются от других рынков. С одной стороны, цена картины не может быть определена как точка равновесия между спросом и предложением. С другой стороны, производственные затраты, статус и провенанс также не позволяют свободно установить цену на произведение искусства. Некоторые исследователи (Goetzmann, Spiegel, 1995) утверждают, что художественный объект стоит ровно столько, сколько устраивает покупателя и продавца. В 1966 г. Университет Ланкастера начал использовать в качестве основы методов ценообразования объектов искусства «classical balance de peintres» («классический баланс художников»), описанный в XVIII в. французским художником и художественным критиком Роджером де Пьелем (de Piles, 1708).

Идентификация характеристик картины, т.е. определение композиции, материала и цвета, была попыткой построить модель эффективного ценообразования арт-объектов. По результатам серии эмпирических исследований определения стоимости произведений (Sagot-Duvauroux et al., 1992; Velthuis, 2005) установлено, что на цену картины больше всего влияют три группы факторов: характеристики самой работы, характеристики художника и внешние факторы. Многие исследования при определении стоимости прежде всего сосредоточены на особенностях самой работы.

Одной из наиболее популярных характеристик, положительно влияющих на цену, является размер холста. Тем не менее существует критическое значение размера работы (Sagot-Duvauroux, 2003). Слишком большие картины требуют больше места для хранения и характеризуются чрезвычайно сложной транспортировкой. Ввиду этого будут наблю-

даться снижение спроса на их размещение на выставках и, соответственно, меньший провенансом, что напрямую влияет на стоимость этого художественного актива. Более того, картины, выполненные маслом, обычно дороже выполненных в любой другой технике. В одной из работ автор объясняет эту тенденцию долговечностью масляных картин и мастерством художника, который их использует. В исследовании (Sproule, Valsan, 2006) было доказано, что наиболее дорогие работы написаны на холстах с использованием масляных красок. Именно это сочетание обладает наиболее значимым эффектом в определении цены в данной работе. Некоторые исследователи также считают, что основная ценность картины – ее сюжет. Отмечается, что сюжет не может быть определяющим фактором цены, так как он также зависит от вкуса, тенденций моды, политической ситуации и мнений экспертов (Sproule, Valsan, 2006). В другой работе этот вывод подтверждается доказательством гипотезы о том, что в современном мире общий эффект сюжета уже не является настолько значительным (Sagot-Duvauroux et al., 1992). Кроме того, наличие подписи художника или даты на произведении положительно влияет на его цену. Наличие этих двух характеристик является особым доказательством подлинности произведения. Произведения искусства с более высокой степенью достоверности ценятся намного выше других. Наиболее спорной переменной являются имя художника или его принадлежность к определенной школе (Rengers, Velthuis, 2002).

Отдельным фактором для анализа является «эффект смерти художника». Он связан с резким ростом цен на работы художника сразу после его смерти, так как большинство полотен ныне живущих художников ценятся значительно ниже, чем уже умерших. Это можно объяснить логически; более того, (Agnello, Pierce, 1996) выявили нелинейную корреляцию между ценностью картины и возрастом художника, в котором картина была написана. Покупатели готовы платить больше за картины старых и более опытных художников. Для сравнения: в 2019 г. был побит новый рекорд стоимости живописного произведения, написанного ныне живущим художником. Картина «Портрет художника» Дэвида Хокни была продана за 90,3 млн долл. на аукционе Christie's в Нью-Йорке. Ранее данный рекорд находился на отметке менее 60 млн долл., отданных за скульптуру Джеффа Кунса[7]. При этом картина Хокни все равно занимает лишь 5-е место среди самых дорогих из когда-либо проданных полотен, и это единственная позиция ныне живущего художника среди 200 строчек рейтинга.

В еще одном исследовании о формировании цен на картины бельгийских художников, проданных на аукционах в последних трех декадах XX в., авторы доказали важность всех указанных выше параметров (Renneboog, Van Hootte, 2013). Анализ более 10 тыс. аукционных цен привел авторов к выводу о том, что существует несколько значимых факторов, которые существенно влияют на стоимость произведения. Так, эти авторы выявили, что конечная стоимость картин, написанных маслом, оказывается намного выше стоимости картин, написанных акварелью, пастелью или другими красками. Еще один фактор, который также оказался значимым, помимо наличия подписи художника и размеров работы, – продажа картины на определенном аукционе. Картины, проданные такими общеизвестными аукционными домами, как Christie's и Sotheby's, имеют гораздо большую стоимость, чем те, что были выставлены в порядке частной сделки, или на менее знаковых для общественности аукционах. Также в работе было проанализировано художественное направление, в котором выполнена картина, т.е. в качестве категориальной переменной авторы учитывали разные стили искусства. На основе данной переменной, включенной в их регрессию, они сделали вывод о том, что работы, выполненные в жанре сюрреализма

[7] Koons J. And his balloon dogs. 21.01.2018. [Электронный ресурс]. URL: http://www.dailyartmagazine.com/jeff-koons-balloon-dog/.

и постимпрессионизма, были наиболее продаваемыми, а экспрессионизма и символизма – наиболее дорогими (Renneboog, Van Houtte, 2013).

Таким образом, к исследуемым в данной работе факторам мы относим: наличие подписи, масляная техника, площадь работы (кв. см), факты продаж на аукционах и индекс настроений в сети Твиттер, с помощью которого можно выявить уникальность, моду и ведущие тенденции сегодня.

Причина выбора именно этих параметров, описанных в табл. 1, обусловлена их наибольшей значимостью исходя из результатов предыдущих исследований.

Основной задачей исследования является построение эффективной предсказательной модели ценообразования на предметы искусства. Однако для подтверждения сформулированных выше гипотез также необходимо сравнить несколько типов моделей.

Модель 1 поможет тестировать гипотезу 1 о повышении точности предсказания цены на произведение искусства при использовании в качестве одного из факторов индекса настроений в сети Твиттер:

$$\ln price = \propto_0 + \beta_0 \cdot living + \beta_1 \cdot square + \\ + \beta_2 \cdot oil + \beta_3 \cdot sign + \beta_4 \cdot auction + \\ + \beta_5 \cdot sentiment.$$

Модель показывает зависимость линейного логарифма цены произведения и включает такие факторы, как смерть художника на момент продажи (*living*), площадь картины (*square*), подпись художника (*sign*), масляная

Таблица 1

Описание и спецификация факторов модели

Цена после торгов	*price*	Зависимая переменная 1 (*Y*)
Разница между предварительной оценкой экспертов и ценой торгов	*dif*	Зависимая переменная 2 (del *Y*)
Размер	*square*	Независимая количественная переменная. Отражает зависимость между площадью картины (длина × высота) и ее стоимостью
Фактор смерти художника на момент продажи	*living*	Независимая дамми-переменная. 1 – на момент продажи работы художник уже умер, 0 – художник жив на момент продажи произведения. Определяет зависимость между фактом смерти художника и итоговой ценой картины
Подпись	*sign*	Независимая дамми-переменная. 1 – при наличии подписи художника на арт-объекте, 0 – при ее отсутствии. Определяет зависимость между наличием подписи и итоговой ценой картины
Масло	*oil*	Независимая дамми-переменная. 1 – работа выполнена масляными красками, 0 – в любом ином случае. Определяет значение масляной техники в ценообразовании работы
Аукцион	*auction*	Независимая дамми-переменная. 1 – продажа картины была осуществлена на открытых аукционных торгах, 0 – в любом ином случае, т.е. при сделке, проведенной в частном порядке. Данный фактор отражает зависимость между продажей именно на аукционе и ее стоимостью
Индекс настроений в сети Твиттер	*sentiment*	Независимая количественная переменная. Отражает связь между индексом настроений сети и стоимостью работы. Подробнее о нем см. далее

И с т о ч н и к: расчеты авторов.

техника исполнения (*oil*) и продажа непосредственно на аукционе (*auction*), в нее также добавлен некоторый фактор сентимент-анализа (*sentiment*). β выражает коэффициент при каждом параметре. По результатам данной модели можно выявить, наблюдается ли увеличение прогнозной силы модели в случае добавления индекса настроений в сети.

Модель 2 может быть использована для тестирования второй сформулированной в данной работе гипотезы. Ее главное отличие от предыдущих моделей заключается в том, что в качестве зависимой переменной выступает различие между итоговой ценой и предварительной оценкой экспертов:

$$\ln dif = \propto_0 + \beta_0 \cdot living + \beta_1 \cdot square + \\ + \beta_2 \cdot oil + \beta_3 \cdot sign + \beta_4 \cdot auction + \\ + \beta_5 \cdot sentiment.$$

Модель показывает зависимость линейного логарифма цены произведения и включает такие факторы, как смерть художника на момент продажи (*living*), площадь картины (*square*), подпись художника (*sign*), масляная техника исполнения (*oil*) и продажа непосредственно на аукционе (*auction*), в нее также добавлен некоторый фактор сентимент-анализа (*sentiment*).

3. ОПИСАНИЕ ДАННЫХ

Каждый художник был выбран на основе рейтинга самых дорогих из когда-либо проданных произведений искусства при условии, что оно выставлялось на торги аукционных домов или в частном порядке в течение последних 20 лет. Также для чистоты эксперимента были добавлены самые известные современные художники, которые также выставляли свои полотна на торги в течение заданного периода. Таким образом, для отбора твитов использовались данные о самой работе, задавался поиск по хэштегам в сети Твиттер с указанием имени, имени и фамилии вместе и только фамилии художника внутри периода двух–четырех месяцев до продажи работы. Дата продажи и предварительная оценка экспертов находятся в открытом доступе на сайтах аукционных домов Sotheby's и Christie's. Выбор данного промежутка времени обусловлен предварительной публикацией аукциона или частного продавца о продаже полотна в открытых источниках и возможным дальнейшим обсуждением данной информации в сети. Так, например, после реставрации картины Леонардо да Винчи «Спаситель мира», которая считалась утраченной, и выставления ее на торги на аукционе появилось более 2 тыс. сообщений об этом событии в средствах массовой информации.

Таким образом, выборка твитов на каждого отдельного художника содержит от 20 до 102 публикаций только на английском языке по теме за период в два–четыре месяца, что показано в табл. 2. Помимо этого, была собрана информация о числе репостов на каждый отдельный твит и получена их сумма для каждого отдельного художника. Результаты также отображены в табл. 2.

Помимо самого текста публикации и хэштега, по которому осуществлялся запрос, выборка содержит данные о числе репостов, лайков и комментариев для каждого твита. Затем на основе данных твитов была оценена их эмоциональная окраска как положительная, нейтральная или негативная и соответственно проведен сентимент-анализ на позитивность предоставленных твитов.

В выборку включено 25 художников и 90 наблюдений, т.е. произведений искусства. Эта разница связана с тем, что в один аукционный день на торги могло быть выставлено несколько картин одного художника, что позволяет увеличить общую выборку и число наблюдений. Выборка данных состоит из цены объекта, разницы между предварительной оценкой экспертов и реальной конечной ценой, размерами картины: площади, наличия подписи, техникой живописи, аукционной

Таблица 2

Число твитов и репостов о художнике на английском языке в заданном периоде

Художник	Число опубликованных твитов на английском языке за период	Число репостов всех использованных твитов
Леонардо да Винчи	66	786
Поль Сезанн	38	1633
Клод Моне	32	856
Бэнкси	71	6944
Пабло Пикассо	102	4122
Сальвадор Дали	31	1651
Фрэнсис Бэкон	43	140
Жан-Мишель Баския	26	143
Дэмьен Херст	29	537
Дэвид Хокни	32	2277
Поль Гоген	41	1041
Ганс Гольбейн (Младший)	20	54
Кит Харинг	20	48
Виллем де Кунинг	37	823
Густав Климт	34	1018
Джефф Кунс	33	1135
Казимир Малевич	25	269
Амедео Модильяни	51	1430
Такаси Мураками	23	348
Эдвард Мунк	29	889
Рене Магритт	20	18
Джексон Поллок	35	901
Энди Уорхол	36	423
Марк Ротко	37	972
Яёи Кусава	23	1891

И с т о ч н и к: расчеты авторов.

или частной продажей и результатов сентимент-анализа.

Предварительная оценка стоимости картины отражает мнение экспертов и их примерные оценки стоимости произведения на будущих торгах. За основу исследования был взят рейтинг самых дорогих художественных работ, выполненных в технике живопись, и отдельный рейтинг самых дорогих работ современных художников из ныне живущих или родившихся в XX в. Данные рейтинги были взяты из интернет-портала ArtNet[8]. Каждая строка отражает отдельное произведение и его главные характеристики. Цена работы, а также ее оценка была взята из открытых источников – сайтов Sotheby's и Christie's.

В нашем исследовании были использованы данные наиболее дорогих и даже рекордных сделок в сфере искусства, данные об их итоговой стоимости также были взяты из интернет-источников: ArtNet и ARTInvestment.ru. Аналогично на официальных сайтах Sotheby's и Christie's указаны наличие подписи, размеры работы и тип краски. В данной работе исследовалось только то, написана ли работа масляными красками или иной техникой. Недостающие данные, отсутствующие по причине давности проведения торгов или в связи с продажей не на аукционах именно двух этих компаний, были получены с помощью двух сервисов: ArtNet и ARTInvestment.ru.

Результаты сентимент-анализа и данные индекса для каждого художника описаны на рис. 2. Общий балл тональности посчитан по всем твитам и делен на число твитов по каждому художнику. Например, наиболее

[8] Интернет-проект ArtNet [Электронный ресурс]. URL: http://www.artnet.com/.

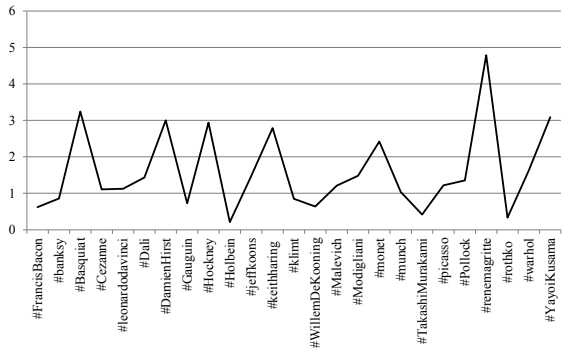

Рис. 2. Распределение индекса позитивности в сети Твиттер по художникам по словарю AFINN

И с т о ч н и к: расчеты авторов.

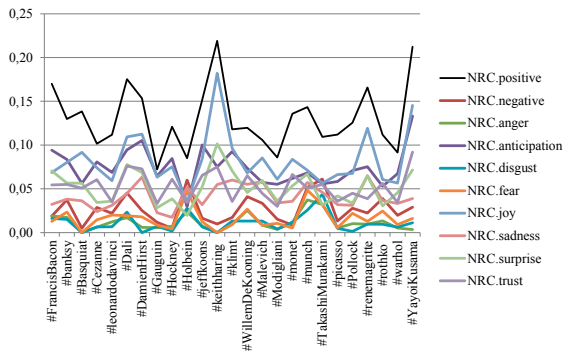

Рис. 3. Распределение индекса позитивности в сети Твиттер по художникам (по словарю NRC)

И с т о ч н и к: расчеты авторов.

высоким, а значит, позитивным индексом настроений в сети Твиттер по словарю AFINN обладают художники Рене Магритт, Жан-Мишель Баския, Дэмьен Херст и Яёи Кусама.

Результаты распределения эмоций, определенные словарем NRC, представлены на рис. 3 ниже. Таким образом, видно сильное преобладание позитивных эмоций над любыми другими, ввиду чего именно они будут использоваться для исследования. Можно заметить, что наиболее позитивными и радостными эмоциями обладают такие художники, как Рене Магритт, Кит Харинг, Яёи Кусама и Дали, что немного отличается от результатов словаря AFINN.

4. РЕЗУЛЬТАТЫ И ДИСКУССИЯ

Для данной модели будет использован фактор сентимент-анализа, рассчитанный с использованием словаря AFINN. С помощью упомянутого словаря была определена эмоциональная окраска твитов о каждом художнике с точки зрения позитивности. Результаты построения всех гипотез представлены в табл. 3.

Результаты оценки свидетельствуют о том, что указанные факторы объясняют модель на 36,8%. При этом при использовании

Таблица 3

Результаты оценки моделей

Название моделей	Модель 1	Модель 2a	Модель 2b
c	14,6*** (1,12)	14,1*** (1,32)	14,82*** (1,11)
auction (Аукцион)	–0,61 (0,72)	–1,2 (0,85)	–1,76** (0,77)
sign (Подпись)	0,05 (0,53)	0,09 (0,62)	0,085 (0,62)
oil (Масло)	1,34*** (0,47)	1,48*** (0,54)	1,47*** (0,54)
sentiment (Сентимент-анализ)	0,42** (0,19)	0,43* (0,23)	6,23** (3,037)
square (Площадь)	0,00004*** (0,0000116)	0,00004*** (0,000013)	0,00004*** (0,000013)
living (Фактор смерти)	–1,12 (0,68)	–0,64 (0,78)	–0,78 (0,77)
Число наблюдений	83	75	75
R^2	0, 3680	0, 3683	0, 3738

Пояснение: *, **, *** значимость на 10-, 5- и 1%-м уровне соответственно. В скобках приведены стандартные ошибки оценок коэффициентов модели.

разумного коэффициента значимости, равного 5%, техника масло, индекс настроений в сети и размеры картины оказывают наибольшее влияние на стоимость. Фактор смерти художника обычно повышает стоимость картины, т.е. картины художников, которые уже умерли на момент продажи, ценятся гораздо выше картин живущих. Данные выводы логичны, так как только одна работа ныне живущего художника входит в рейтинг первых 100 картин по стоимости. Однако данный параметр подтверждается только при 10%-м уровне значимости. Можно заметить, что работы маслом и размер картины (выраженный в квадратных сантиметрах) не отвергается и при 1%-й значимости.

Таким образом, подтверждается основная гипотеза данного исследования: индекс настроений в сети Твиттер положительно влияет на аукционную стоимость произведения искусства. Это означает, что публикации в социальных сетях между моментом анонса продажи и фактическими торгами серьезно влияют на конечную стоимость картины. Воздействие данного фактора не удивительно, так как роль социальных сетей в формировании цены разных активов и отношения масс к какой-либо компании или личностям были широко исследованы и доказана. Таким образом, согласно построенной регрессии, при нарастании позитивности твитов на единицу цена картины увеличивается на 0,41.

Модель 2 отличается от первой тем, что в качестве зависимой переменной рассматривается разница между итоговой аукционной ценой и предварительной оценкой. Однако зависимая переменная имеет распределение χ^2-квадрат, нежелательное для исследования, ввиду этого в данной модели также будет взят логарифм разницы цен.

Таким образом, наиболее значимыми параметрами по-прежнему являются индекс настроений в сети, масляная техника и размер картины. Как было показано ранее, наиболее интересные результаты в данной модели должны быть представлены факторами сентимент-анализа и фактом продажи на открытых торгах. Однако индекс настроений в сети Твиттер, хотя и близок 5%, но оказался значимым только на уровне 10%, что тоже является очень хорошим результатом на такой небольшой выборке. Фактор аукциона не существенно повлиял на цену, однако это также легко объяснить тем, что, как правило, на экспертную оценку никак не влияют условия продажи картины, а наиболее интересные лоты продают в рамках как закрытого, так и открытого аукциона. Часто бывает такое, что даже если картина должна была выставляться на открытых торгах, авторитетный музей, компания или коллекционер могут купить картину, предварительно договорившись с продавцом или агентом. Например, сейчас таким образом пополняет и расширяют свою коллекцию музеи Катара и Абу-Даби. В таком случае цена или предварительные оценки экспертов часто не имеют значения, так как произведение искусства во многом ценится как статусное приобретение для поддержания имиджа.

Так как в данных моделях в переменной «индекс настроений в сети Твиттер» был использован только словарь AFINN, то для более точного результата можно использовать данный индекс, рассчитанный с помощью и другого словаря – NRC. В данной модели также будет рассмотрена позитивность. Мы провели исследование с помощью этого словаря (см. рис. 2) и с его помощью разделили все слова в публикациях на слова с разной эмоциональной окраской. Число слов с любой другой эмоциональной окраской, кроме позитивной, оказалось небольшим, поэтому мы берем только слова с позитивной окраской, а остальные отбрасываем. Можно заметить, что в отличие от той же модели с использованием другого словаря индекса настроений в сети AFINN, описанной ранее, фактор продажи на аукционе является значимым на уровне 5%, как и индекс настроений в сети.

Данные регрессии подтверждают: действительно существует влияние позитивных новостей и публикаций в Твиттере на конеч-

ную стоимость картины, как и размер картины, и работа масляными красками. Однако эти последние два фактора не стоит учитывать именно в данной модели, так как они не меняются с момента публикации анонса о продаже произведения до его продажи. При этом легко объяснить причину того, почему наблюдается настолько высокая значимость этих двух параметров. Как указано в описании предыдущей модели, и размер, и техника арт-объекта прямо и существенно влияют на его стоимость, а оценка экспертов наиболее дорогих картин не достигает таких высоких значений, поэтому и наиболее высокая разница наблюдается у полотен, проданных за рекордные суммы. В какой-то мере высокие цены были достигнуты в результате не только понимания высокой ценности картины, но и трансформации искусства в бизнес и использования факта продажи или покупки произведения искусства в целях пиара. Исследуем, например, полотна и инсталляции Дэмьена Хёрста. Известный художник в интервью не раз говорил о том, что не соглашается продавать свои работы до тех пор, пока не будет предложена рекордная и даже абсурдная цена, так как «его интересуют не деньги, а влияние в мире искусства». Исходя из того факта что в данной регрессии было подтверждено влияние индекса позитивного настроения в сети Твиттер, можно сделать вывод о доказательстве второй гипотезы: позитивная эмоциональная окраска публикаций в Твиттере имеет линейную зависимость с разницей между предварительной оценкой и итоговой ценой арт-объекта. Это свидетельствует о том, что при большом числе позитивных публикаций по теме итоговая цена будет линейно расти. Интересно предположить, что в таком случае возможна намеренная фальсификация мнений в социальных сетях и, соответственно, рост обсуждения картины, который в конечном итоге может привести к росту цены.

Модель 3 отображает зависимость между таким параметром, как число репостов всех постов в сети Твиттер за заданный период и отличие аукционной и оценочной стоимости работы. Анализ результатов показал, что параметр, отвечающий за число репостов, не является значимым. Однако этот факт легко объяснить тем, что наиболее обсуждаемыми стали художники не с самыми дорогими работами, т.е. разница между оценкой и аукционной ценой не такая высокая. Так, например, наиболее популярным художником, если судить по числу репостов в Твиттере, является Бэнкси (6944 репоста), но исследуемая в данной работе картина был продана только за 1,4 млн долл. при предварительной оценке в 800 тыс. долл. На основе этого можно сделать вывод о том, что гипотеза 3 не подтверждается, а значит, число репостов в социальных сетях не влияет на разницу между итоговой ценой и предварительной оценкой экспертов.

Гипотезы 1 и 2, которые были поставлены в исследовании, подтвердились. В целом это доказывает, что настроения людей и их восприятие искусства во многом зависят от заданных тенденций, моды и мнения большинства. Таким образом, нельзя отрицать, что общий ажиотаж вокруг какой-то работы и отношение к ней аналитиков, заинтересованных покупателей, продавцов и просто зрителей в современном мире диктуют ее значимость. Этот фактор еще раз заставляет исследователей, аналитиков и экспертов учитывать его для построения модели.

ЗАКЛЮЧЕНИЕ

Данная работа является новаторской и вносит вклад в научное сообщество, так как в ней описаны новые методы ценообразования произведений искусства. Нами были построены успешные модели, которые доказали значимость некоторых параметров. Таким образом, индекс настроений в сети Твиттер оказался значимым на уровне 5% для всех протестированных моделей, а техника работы маслом и размеры работы подтверждены

на 1%-м уровне значимости. Так, в ней впервые используется индекс настроений в сети Твиттер как фактор ценообразования на предметы искусства. При этом по результатам нашего исследования данный фактор является значимым и существенно влияет на итоговую стоимость картины. Этот вывод аналогичен ранним исследованиям того наблюдения, что текстовый анализ существенно влияет на инвестиционную стоимость актива (Tetlock, 2007).

Также можно сделать вывод о том, что искусство сегодня не является благом, стоимость которого нельзя оценить. Наоборот, протестированные модели показывают, насколько цену данного фактора можно предсказать. Сентимент-анализ новостных источников влияет на стоимость финансовых активов (Agarwal et al., 2016), поэтому интересно отметить, что сегодня предметы искусства должны иметь аналогичную систему ценообразования, как и любой финансовый и инвестиционный актив. В дополнение было объяснено поведение покупателей на рынке искусства. Доказательство гипотез о значимости индекса настроений в сети приводит к выводу о том, что потребители рынка искусства являются больше не ценителями, которые покупают работы, исходя исключительно из собственного желания, а инвесторами, которые рассматривают предметы искусства как долгосрочные финансовые активы.

Практическая значимость данного исследования заключается в том, что полученные нами результаты доказывают, что для построения более эффективной оценки произведений искусства эксперты должны учитывать фактор настроений людей и их реакций на события в данной области. Сентимент-анализ новостных источников и социальных сетей позволяет повысить результативность оценок. Таким образом, в вопросах ценообразования предметов искусства могут быть использованы те же методы оценки, что и для инвестиционных активов.

Список литературы / References

Демин И.С., Рогов О.Ю., Федорова Е.А. (2019). Применение словарей тональности для текстового анализа // Прикладная информатика. Т. 79. № 1. [Demin I., Rogov O., Fedorova E. (2019). Application of the tonality dictionaries for the sentimental analysis. *Journal of the Applied Informatics*, vol. 79, no. 1 (in Russian).]

Agarwal A., Sharma V., Sikka G., Dhir R. (2016). Opinion mining of news headlines using SentiWordNet. *Colossal Data Analysis and Networking (CDAN)*. IEEE, March, pp. 1–5.

Agnello R.J., Pierce R.K. (1996). Financial returns, price determinants, and genre effects. *Journal of Cultural Economics*, no. 20, pp. 359–383.

Bollen J., Huina Mao, Xiaojun Zeng (2011). Twitter mood predicts the stock market. *Journal of Computational Science,* no. 2.1, pp. 1–8.

Chan W.S. (2003). Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics*, vol. 70, no. 2, pp. 223–260.

Goetzmann W.N., Spiegel M. (1995). Non-temporal components of residential real estate appreciation. *Review of Economics and Statistics*, no. 77 (1), pp. 199–206.

Piles de R. (1989). Cours de peinture par principes. Ed. by J. Thuillier. Paris: Gallimard.

Mao H., Counts S., Bollen J. (2011). Predicting financial markets: Comparing survey, news, twitter and search engine data. Preprint.

Rengers M., Velthuis O. (2002). Determinants of prices for contemporary art in Dutch Galleries, 1992–1998. *Journal of Cultural Economics*, no. 26 (1), pp. 1–28.

Renneboog L., Van Houtte T. (2002). The monetary appreciation of paintings: From realism to Magritte. *Cambridge Journal of Economics*, no. 26 (3), pp. 331–357.

Sagot-Duvauroux D. (2003). Art prices. In: Towse R. (ed.). *A Handbook of Cultural Economics*. Cheltenham: Edward Elgar.

Sagot-Duvauroux D., Pflieger S., Rouget B. (1992). Factors affecting price on the contemporary art market. In: Towse R., Khakee A. (eds). *Cultural*

Economics. Berlin, Heidelberg: Springer-Verlag, pp. 91–102.

Sproule R., Valsan C. (2006). Hedonic models and pre-auction estimates: Abstract art revisited. *Economics Bulletin*, no. 26 (5), pp. 1–10.

Tetlock P.C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, vol. 62, no. 3, pp. 1139–1168.

Velthuis O. (2005). Talking prices: Symbolic meanings of prices on the market for contemporary art. Princeton, Oxford: Princeton University Press.

Wang H., Can D., Kazemzadeh A., Bar F., Narayanan S. (2012). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. *Proceedings of the ACL*, July. System Demonstrations Association for Computational Linguistics, pp. 115–120.

ПРИЛОЖЕНИЕ

Результаты оценки сентимент-анализа работ некоторых художников по твитам

Таблица П1

Первые 50 твитов выборки по художнику Бэнкси с указанием числа репостов, лайков, комментариев и сентимент-анализа для каждого

doc_id	text	hashtag	comment	repost	like	AFINN.po	AFINN.po	NRC.positi	NRC.nega	NRC.anger	NRC.antic	NRC.disgu	NRC.fear	NRC.joy	NRC.sadne	NRC.surpri	NRC.trust
1	banksy	#banksy	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
2	Banksy Bankitalia DE	#banksy	1	60	139	0	0	0	0	0	0	0	0	0	0	0	0
3	If Banksy was Italian s	#banksy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	We are shredding the	#banksy	2	17	189	0	0	0	0	0	0	0	0	0	0	0	0
5	Banksy ou le paradox	#banksy	10	50	101	0	0	0,25	0	0	0	0	0	0	0	0	0
6	Banksy G i r l W i t h	#banksy	0	13	305	0	0	0,142857	0	0	0,285714	0	0	0,142857	0,142857	0,142857	0
7	McDonalds recent tre	#banksy	4	22	68	2	0,08	0,16	0	0	0	0	0	0,04	0	0,04	0,04
8	belgian joke following	#banksy	8	73	162	2	0,4	0	0,4	0	0	0	0	0	0	0,2	0
9	Ci sono dischi di certi a	#banksy	1	5	34	2	0,133333	0	0	0	0	0	0	0	0	0	0
10	My Banksy tweet has	#banksy	3	8	55	4	0,266667	0,2	0,066667	0,133333	0,2	0	0	0,2	0,266667	0,2	0
11	My dear RobinRoberts	#banksy	4	15	106	2	0,2	0,1	0	0	0	0	0	0	0	0	0
12	Banksy is kinda badass	#banksy	15	14	299	-3	-1	0	0	0	0	0	0	0	0	0	0
13	A video has appeared	#banksy	14	54	114	0	0	0,142857	0	0	0	0	0	0	0	0	0,142857
14	The democrats have p	#banksy	1	14	51	0	0	0	0	0	0	0	0	0	0	0	0
15	Banksy is a provocate	#banksy	5	6	29	0	0	0,058824	0,058824	0	0	0	0,058824	0	0	0	0
16	The team over Bunny	#banksy	7	13	105	2	0,285714	0,142857	0	0	0,142857	0	0	0	0	0	0,285714
17	Banksy just gave the f	#banksy	9	7	51	3	0,230769	0,307692	0	0	0,230769	0,076923	0	0,230769	0,153846	0,230769	0,153846
18	Fucking genius Banksy	#banksy	1	18	102	-4	-0,66667	0,166667	0	0	0	0	0	0	0	0	0
19	This latest Banksy eve	#banksy	9	24	127	1	0,076923	0,153846	0,153846	0,076923	0,076923	0,076923	0	0,076923	0	0	0,076923
20	This is so so bloody go	#banksy	7	120	261	0	0	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
21	banksy	#banksy	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
22	In you missed what this	#banksy	5	10	64	-2	-0,66667	0,333333	0	0	0	0	0	0	0	0	0
23	Did you hear about Ba	#banksy	3	11	38	-1	-0,0625	0,0625	0,1875	0,0625	0,0625	0,0625	0,0625	0,0625	0,1875	0,125	0
24	Banksy Putin paid way	#banksy	10	176	346	0	0	0	0	0	0	0	0	0	0	0	0
25	Haven t been this excit	#banksy	0	3	17	3	0,428571	0,428571	0	0	0,571429	0	0	0,285714	0,142857	0,285714	0,285714
26	Is who knew Banksy w	#banksy	10	100	400	0	0	0	0	0	0	0	0	0	0	0	0
27	I am so proud of Bank	#banksy	6	8	37	4	0,363636	0,363636	0,181818	0	0,272727	0	0,181818	0,363636	0,090909	0,272727	0,090909
28	Auctioneers had just s	#banksy	5	15	33	1	0,04	0,12	0,16	0	0,08	0	0,12	0,08	0,12	0,08	0,04
29	Soooo today someone	#banksy	6	5	113	0	0	0	0	0	0	0	0	0	0	0	0
30	La toile de Banksy aur	#banksy	12	41	258	0	0	0,153846	0	0	0,076923	0	0	0,076923	0,076923	0,076923	0
31	Big up to Banksy for fl	#banksy	6	43	144	1	0,04	0,16	0	0	0,04	0	0	0,08	0,04	0,04	0,04
32	The best prank in histo	#banksy	2	27	42	0	0	0,2	0,2	0	0,2	0	0	0,2	0,2	0,4	0
33	The best thing I ve see	#banksy	6	82	323	0	0	0	0	0	0	0	0	0	0	0	0
34	C est Banksy lui m me	#banksy	9	139	207	-5	-0,16129	0,032258	0,032258	0,064516	0,032258	0	0	0,032258	0,032258	0,064516	0
36	Quand ce que tu viens	#banksy	6	104	175	0	0	0	0	0	0	0	0	0	0	0	0
37	This may be the greate	#banksy	3	11	105	3	0,3	0	0	0	0	0	0	0	0	0	0
38	Banksy	#banksy	0	0	0	0	0	0,5	0	0	0,5	0	0	0,5	0	0	0,5
39	Dat moment This is ju	#banksy	0	26	78	4	1	0,25	0	0	0,25	0	0	0,25	0	0	0,25
40	Banksy has shared vi	#banksy	71	1400	2300	-2	-0,16667	0	0	0	0	0	0	0	0	0	0
41	Ok I m ready to drink t	#banksy	3	5	45	0	0	0	0	0	0,25	0	0	0	0	0	0
43	Have admired Banksy	#banksy	21	352	624	3	0,1875	0,125	0,0625	0,0625	0	0,0625	0,0625	0,0625	0,0625	0,125	0,0625
44	Banksy n a rien invent	#banksy	3	128	211	0	0	0	0	0	0	0	0	0	0	0	0
45	What a great move Ri	#banksy	1	19	39	3	0,2	0	0,066667	0	0,066667	0	0	0	0,066667	0	0
46	A Banksy Painting Sel	#banksy	16	41	117	0	0	0,125	0	0	0	0	0	0	0	0	0,125
47	Exit through the shredc	#banksy	1	20	34	0	0	0	0	0	0	0	0	0	0	0	0
48	How an artist with a st	#banksy	0	0	0	1	0,1	0,1	0	0	0,1	0	0	0,1	0	0	0,2
49	What do Banksy Lore	#banksy	0	0	0	2	0,133333	0,133333	0	0	0,133333	0	0	0,133333	0,066667	0,066667	0,066667
50	Think of it as Recyclin	#banksy	0	0	0	2	0,125	0,1875	0,0625	0	0,1875	0	0,0625	0,0625	0	0	0,1875

И с т о ч н и к: расчеты авторов.

Таблица П2

Первые 30 твитов выборки по художнику Сальвадор Дали с указанием числа репостов, лайков, комментариев и сентимент-анализа для каждого

doc_id	text	hashtag	comment	repost	like	AFINN.po	AFINN.po	NRC.positi	NRC.nega	NRC.anger	NRC.antic	NRC.disgu	NRC.fear	NRC.joy	NRC.sadne	NRC.surpri	NRC.trust
1	Loving the	#Dali	0	1	0	2	0,666667	0,333333	0	0	0	0	0	0,333333	0	0	0,333333
2	It s not oft	#Dali	1	105	10	3	0,375	0,375	0,25	0,125	0,125	0,125	0,125	0,25	0,25	0,25	0
3	Just editin	#Dali	1	8	5	-3	-0,21429	0,071429	0,142857	0,071429	0,214286	0,071429	0,071429	0,071429	0,071429	0,071429	0
4	I am not a	#Dali	1	7	5	1	0,045455	0,181818	0,045455	0	0	0,045455	0,045455	0,045455	0	0,045455	0,136364
5	Dal was	#Dali	0	7	0	3	0,157895	0,157895	0,052632	0	0,157895	0	0	0,052632	0,052632	0,105263	0,052632
6	Salvador[	#Dali	0	0	0	0	0	0,117647	0,058824	0	0	0,058824	0	0	0,058824	0	0
7	I ve perso	#Dali	0	5	4	0	0	0,1	0	0	0,1	0	0	0,1	0,1	0,1	0
8	This is the	#Dali	0	24	6	6	0,428571	0,142857	0	0	0,071429	0	0	0,142857	0	0,142857	0,071429
9	But the on	#Dali	9	99	34	4	0,444444	0,111111	0,222222	0,111111	0,111111	0,111111	0,111111	0,111111	0,222222	0,111111	0
10	Dal was	#Dali	0	3	1	3	0,157895	0,157895	0,052632	0	0,157895	0	0	0,052632	0,052632	0,105263	0,052632
11	Dal had a	#Dali	1	24	6	3	0,176471	0,176471	0	0	0	0	0	0	0	0	0,058824
12	A filming d	#Dali	0	20	5	4	0,235294	0,176471	0	0	0	0	0	0	0	0	0
13	A visit to M	#Dali	0	56	4	4	0,444444	0,111111	0	0	0	0	0	0	0	0	0
14	Every sprii	#Dali	0	11	4	0	0	0	0,083333	0	0	0	0	0	0,083333	0	0
15	And he wa	#Dali	3	40	19	0	0	0	0	0	0	0	0	0	0	0	0
16	Inquiring a	#Dali	0	30	7	-2	-0,14286	0,142857	0,071429	0	0,071429	0	0	0,071429	0,071429	0,142857	0
17	The clown	#Dali	2	338	166	-2	-0,15385	0,230769	0,076923	0,076923	0,153846	0,076923	0,076923	0,076923	0,076923	0,153846	0
18	f she coulc	#Dali	3	58	16	-4	-0,16667	0,041667	0,125	0,041667	0	0,125	0,041667	0	0	0	0
19	Flashback	#Dali	7	52	16	5	0,217391	0,086957	0	0	0	0	0	0,043478	0	0,086957	0,086957
20	Salvador[	#Dali	0	17	6	5	0,416667	0,166667	0	0	0,083333	0	0	0,166667	0	0,083333	0
21	In honor o	#Dali	0	31	19	2	0,2	0,1	0	0	0	0	0	0	0	0	0,1
22	In St Peter	#Dali	1	24	3	2	0,285714	0,285714	0	0	0,285714	0	0	0,285714	0	0,285714	0,285714
23	Salvador D	#Dali	0	36	12	0	0	0,111111	0	0	0,111111	0	0	0,111111	0,111111	0,111111	0
24	Losing the	#Dali	3	79	30	-5	-0,2	0,16	0,04	0,04	0,08	0	0	0,04	0,08	0,04	0
25	If she coul	#Dali	2	44	7	-4	-0,17391	0,043478	0,086957	0,043478	0	0,086957	0	0	0	0	0
26	If you have	#Dali	18	133	23	0	0	0,444444	0	0	0,222222	0	0	0,222222	0,111111	0,111111	0,333333
27	Born onth	#Dali	0	39	16	2	0,166667	0,083333	0	0	0	0	0	0	0	0	0
28	Happy birtl	#Dali	5	256	156	3	1	0,666667	0	0	0,666667	0	0	0,666667	0	0,333333	0,333333
29	Intelligenc	#Dali	0	39	17	3	0,333333	0,333333	0	0	0,222222	0	0,111111	0,333333	0	0	0,333333
30	Salvador D	#Dali	2	65	48	8	0,4	0,15	0,05	0	0	0	0	0,1	0	0,05	0,1

И с т о ч н и к: расчеты авторов.

Таблица П3

Первые 50 твитов выборки по художнику Леонардо да Винчи с указанием числа репостов, лайков, комментариев и сентимент-анализа для каждого

1	Salvator M	#leonardo	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
2	Rare paint	#leonardo	0	14	8	0	0	0	0	0	0	0	0	0	0	0	0
3	See touch	#leonardo	0	15	39	4	0,444444	0	0	0	0	0	0	0	0	0	0
4	Faces w	#leonardo	0	20	61	4	0,333333	0,166667	0	0	0	0	0	0,166667	0	0	0,083333
5	TY so muc	#leonardo	3	11	29	0	0	0	0	0	0	0	0	0	0	0	0
6	The last da	#leonardo	3	25	49	0	0	0	0	0	0	0	0	0	0	0	0
7	WalterIsa	#leonardo	2	33	64	0	0	0	0	0	0	0	0	0	0	0	0
8	WalterIsa	#leonardo	6	15	54	1	0,142857	0,428571	0	0	0,142857	0	0	0,285714	0	0,142857	0,142857
9	Leonardo	#leonardo	0	33	40	0	0	0	0	0	0	0	0	0	0	0	0
10	Why did V	#leonardo	1	20	37	2	0,181818	0,181818	0	0	0,090909	0,090909	0	0,090909	0,181818	0,090909	0
11	What Get	#leonardo	2	4	10	-2	-0,11765	0,235294	0	0	0,058824	0	0,058824	0	0	0,058824	0,117647
12	Detail The	#leonardo	0	3	16	0	0	0	0,066667	0,066667	0	0	0,066667	0	0,066667	0	0
13	I love ever	#leonardo	7	62	179	4	0,222222	0,333333	0	0	0,166667	0	0	0,277778	0,111111	0,111111	0,055556
14	Leonardo	#leonardo	0	4	14	0	0	0,25	0	0	0	0	0	0	0	0	0
15	Here are 5	#leonardo	0	5	9	0	0	0,2	0	0	0	0	0	0	0	0	0,2
16	The great	#leonardo	3	3	17	-2	-0,25	0	0,25	0	0	0	0	0	0	0	0
17	The greate	#leonardo	0	44	71	3	0,272727	0,090909	0	0	0	0	0	0	0	0	0
18	Sometime	#leonardo	1	7	32	3	0,333333	0,222222	0	0	0,111111	0	0	0,111111	0	0	0,111111
19	TONIGHT	#leonardo	1	39	59	5	0,625	0,25	0	0	0	0	0	0,125	0	0	0
20	WalterIsa	#leonardo	0	8	9	2	0,333333	0,166667	0	0	0	0	0	0	0	0	0
21	wanna rea	#leonardo	0	5	23	4	0,5	0,125	0	0	0,125	0	0	0,125	0	0	0,125
22	PIAZZA D	#leonardo	1	5	18	0	0	0	0	0	0	0	0	0	0	0	0
23	Happy put	#leonardo	1	6	13	3	0,333333	0,222222	0	0	0,111111	0	0	0,111111	0	0	0,222222
24	Likewise i	#leonardo	1	12	25	0	0	0	0	0	0	0	0	0	0	0	0
25	Leonardo	#leonardo	5	13	37	0	0	0,1	0	0	0	0	0	0	0	0	0
27	Christie s	#leonardo	0	0	0	0	0	0,2	0	0	0,2	0	0	0,2	0,2	0,2	0
28	Did somet	#leonardo	0	4	20	0	0	0	0	0	0	0	0	0	0	0	0,083333

Окончание табл. П3

29	Here s a si	#leonardo	2	28	54	0	0	0	0,142857	0,142857	0	0	0,142857	0	0	0,142857	0
30	Technolog	#leonardo	1	16	21	2	0,333333	0,333333	0,166667	0,166667	0,166667	0,166667	0,166667	0,166667	0,166667	0	0,166667
31	Drawn for	#leonardo	4	22	142	0	0	0	0	0	0	0	0	0	0	0	0
32	Walterlsa	#leonardo	0	12	18	0	0	0,222222	0	0	0,222222	0	0,111111	0,111111	0,111111	0,111111	0
33	The rivalry	#leonardo	0	19	26	0	0	0,125	0,125	0,125	0	0	0	0	0	0	0
34	Art Lovers	#leonardo	4	22	44	-5	-0,35714	0,357143	0,071429	0,071429	0,285714	0,071429	0,071429	0,214286	0,285714	0,214286	0
35	Simplicity	#leonardo	0	11	25	0	0	0	0	0	0,142857	0	0	0	0,142857	0	0
36	Paramoun	#leonardo	0	54	71	2	0,222222	0,111111	0,111111	0	0	0	0,111111	0	0	0	0,111111
37		#leonardo	0	0	1	3	0,6	0,2	0	0	0,2	0	0	0,2	0	0,2	0,2
38	Drawings	#leonardo	0	2	5	0	0	0	0	0	0	0	0	0	0	0	0
39	Watch cor	#leonardo	0	5	5	0	0	0	0	0	0,222222	0	0,111111	0	0	0	0
40	The noble	#leonardo	0	2	5	6	0,75	0,25	0	0	0	0	0	0,125	0	0	0,125
41	Sample pa	#leonardo	2	10	27	3	0,428571	0	0	0	0	0	0	0	0	0	0
42	Behold W	#leonardo	0	6	3	3	0,5	0,166667	0	0	0	0	0	0	0	0	0
43	DaVinci s	#salvatorn	6	9	26	-3	-0,16667	0	0	0	0	0	0	0	0	0	0
44	Wearing n	#salvatorn	0	1	15	2	0,181818	0,272727	0	0	0	0	0	0	0	0	0
45	Leonardo	#salvatorn	0	2	7	9	0,45	0,2	0,05	0	0,2	0	0,05	0,15	0,05	0,1	0
46	The world	#salvatorn	9	100	178	0	0	0	0	0	0,125	0	0	0	0	0	0
47	On Wedne	#salvatorn	0	5	7	1	0,083333	0	0	0	0,083333	0	0	0	0	0	0
48	More brilli	#salvatorn	1	1	3	5	0,555556	0,222222	0	0	0,222222	0	0	0,222222	0	0	0,222222
49	Christiesl	#salvatorn	7	4	12	0	0	0	0	0	0,111111	0	0	0	0	0	0
50	This da Vi	#salvatorn	3	8	13	-3	-0,375	0	0,125	0,125	0,125	0	0,125	0	0,125	0	0

И с т о ч н и к: расчеты авторов.

Таблица П4

28 твитов выборки по художнику Дэмьен Хёрст с указанием числа репостов, лайков, комментариев и сентимент-анализа для каждого

doc_id	text	hashtag	comment	repost	like	AFINN.po	AFINN.po	NRC.positi	NRC.nega	NRC.anger	NRC.antic	NRC.disgu	NRC.fear	NRC.joy	NRC.sadne	NRC.surpri	NRC.trust
1	If this is ar	#DamienH	0	0	2	0	0	0,625	0	0	0,625	0	0	0,625	0,625	0,625	0
2	The giant	#DamienH	1	7	2	0	0	0	0	0	0	0	0,1	0	0	0	0,1
3	Another de	#DamienH	0	0	0	6	0,352941	0,176471	0	0	0,117647	0	0	0,117647	0,058824	0,117647	0,058824
4	Signed An	#DamienH	0	2	1	5	0,294118	0,176471	0,176471	0	0,058824	0	0,058824	0,176471	0,117647	0	0,058824
5	artsy pres	#DamienH	0	0	0	-2	-0,08696	0,086957	0,043478	0	0,130435	0	0,086957	0,086957	0,043478	0,086957	0
6	If only hirs	#DamienH	0	1	0	0	0	0,142857	0	0	0,142857	0	0	0,142857	0	0,142857	0,285714
7	Just pick c	#DamienH	0	0	0	0	0	0,2	0	0	0,1	0	0	0	0	0	0,1
8	Do you kn	#DamienH	12	61	13	0	0	0,214286	0,071429	0	0,142857	0	0	0,142857	0,071429	0,142857	0,071429
9	Beautiful w	#DamienH	1	9	2	3	0,5	0,166667	0	0	0	0	0	0,166667	0	0	0
10	Perfect da	#DamienH	1	2	0	5	0,357143	0,214286	0	0	0,071429	0	0	0,071429	0	0	0,071429
11	This magn	#DamienH	0	0	0	5	0,294118	0,117647	0	0	0,117647	0	0	0,117647	0,058824	0,117647	0,117647
12	I went dov	#DamienH	0	0	0	9	0,6	0,066667	0	0	0	0	0	0,066667	0	0	0
13	I had so m	#DamienH	0	2	0	4	0,25	0,1875	0	0	0,125	0	0	0,125	0,0625	0,0625	0
14	WHAT a ¯	#DamienH	0	0	0	0	0	0	0,166667	0,166667	0,166667	0	0	0	0,166667	0	0
15	Let your ne	#DamienH	0	0	0	2	0,133333	0,133333	0,066667	0	0,133333	0	0	0,133333	0,133333	0,066667	0,066667
16	Cannot un	#DamienH	1	0	0	0	0	0	0,142857	0	0	0	0,142857	0	0,142857	0	0
17	Beautiful V	#DamienH	0	0	0	3	0,272727	0,090909	0	0	0	0	0	0,090909	0	0	0
18	Beautiful p	#DamienH	0	0	0	6	0,375	0,125	0	0	0,0625	0	0	0,125	0,0625	0,0625	0
19	Wow anot	#DamienH	0	0	0	11	0,647059	0,117647	0	0	0,058824	0	0	0,117647	0,058824	0,117647	0,058824
20	Yep it s all	#DamienH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	Excited to	#DamienH	0	44	15	7	0,269231	0,038462	0,038462	0	0,038462	0	0	0,038462	0,038462	0,038462	0,076923
22	7 major sc	#DamienH	0	35	15	0	0	0,05	0	0	0,1	0	0	0	0	0	0
23	Just annou	#DamienH	2	80	25	2	0,1	0,1	0	0	0,05	0	0	0,05	0,05	0,05	0
24	Best wishe	#DamienH	2	39	8	4	0,25	0,0625	0	0	0,0625	0	0	0,0625	0	0,0625	0,0625
25	GeorgeMi	#DamienH	1	11	3	0	0	0,25	0	0	0	0	0	0	0	0	0,25
26	It is so bea	#DamienH	0	27	6	5	0,333333	0,466667	0	0	0,333333	0	0,066667	0,333333	0	0,133333	0,4
27	The more	#DamienH	5	39	4	5	0,416667	0,25	0	0	0,166667	0	0	0,166667	0,083333	0	0,166667
28	Our miracl	#DamienH	14	178	38	4	0,190476	0,238095	0	0	0,142857	0	0,047619	0,190476	0	0,095238	0,095238

И с т о ч н и к: расчеты авторов.

Рукопись поступила в редакцию 09.09.2019 г.

FACTORS OF PRICE FORMATION FOR ART OBJECTS WITH THE APPLICATION OF TEXT ANALYSIS OF TWITTER NEWS

E.A. Fedorova, D.V. Zaripova, I.S. Demin

DOI: 10.33293/1609-1442-2020-2(89)-114-131

Elena A. Fedorova, National Research University Higher School of Economics, Moscow, Russia; ecolena@mail.ru
Diana V. Zaripova, National Research University Higher School of Economics, Moscow, Russia
Igor S. Demin, Financial University under the Government of the Russian Federation, Moscow, Russia

Acknowledgement. The article is based on the results of the research carried out at the expense of government funds of the Financial University. Moscow, 2019.

This work confirmed the hypotheses about the influence of the mood index on Twitter on the pricing of art objects and the difference between the experts' estimations and the final price of the auction. The hypotheses were tested with the use of a sample of 83 paintings selected on the basis of ratings of ARTNET's online resource about the most expensive works of art ever sold in the last 10–15 years. The sample consisted of 25 artists, for each of them was made an index of moods on Twitter. This index was created by a sentimental analysis of each tweet about the artist on the hashtag for a period of 2 to 4 months between the announcements of sales in the open sources and the direct sale of the work with the use of the two dictionaries AFINN and NRC.
Keywords: text analysis, pricing, art object, investment object, Twitter.
JEL: C53, Z11.

Manuscript received 09.09.2019