

УДК 330.46:519.2

К. Ю. Кононова

доктор економічних наук, доцент
Харківський національний університет імені В.Н. Каразіна
kateryna.kononova@gmail.com

А. С. Усачова

студент кафедри економічної кібернетики та прикладної економіки
Харківський національний університет імені В.Н. Каразіна
anastasiausachova@gmail.com

ПРОГНОЗУВАННЯ КУРСУ АКЦІЙ З УРАХУВАННЯМ ПОКАЗНИКІВ ПОПУЛЯРНІСТІ ПОШУКОВИХ ЗАПИТІВ

У статті запропонована модель прогнозування курсу акцій з урахуванням поведінкових факторів, в якості яких використані показники популярності пошукових запитів заданої тематики, отримані за допомогою інструменту Google Trends. Для дослідження було обрано дві компанії з різною історією розвитку: Apple та Royal Dutch Shell. На основі досліджень, присвячених поведінковим фінансів, сформований словник з 67 слів економічного, соціального і політичного характеру, що відображають настрої трейдерів, та виявлено ті пошукові запити, показник популярності яких найбільшою мірою впливає на динаміку курсу акцій. З використанням нейронних мереж, на вхід яких, крім іншого, подавалися показники популярності пошукових запитів, проведено прогнозування курсу для двох компаній з різних секторів економіки. Також було здійснено прогнозування з різними комбінаціями цих пошукових запитів, які були об'єднані за смисловим навантаженням. Для побудови нейронних мереж використовували додаток Neural Network Toolbox програмного пакету MathLab. Проведені експериментальні дослідження з різною архітектурою нейронних мереж, були використані тришарові та двошарові нейронні мережі. В ході проведення експерименту виявлено, що для компанії інформаційного сектора якість моделі істотно зросла за рахунок включення поведінкових факторів, в той час як для компанії переробного сектора значного поліпшення прогнозу не спостерігалось. Що, втім, природним чином впливає зі специфіки дослідження. Отримані результати свідчать про адекватність побудованих моделей і підтверджують доцільність використання показників популярності пошукових запитів для прогнозування курсу акцій.

Ключові слова: прогнозування курсу акцій, поведінкові фактори, пошукові запити, нейронна мережа.

JEL Classification: C45, C51, E47, G17.

K. Yu. Kononova

D. Sc. (Economics), Professor
V.N. Karazin Kharkiv National University
kateryna.kononova@gmail.com

A. S. Usachova

Student
V.N. Karazin Kharkiv National University
anastasiausachova@gmail.com

STOCK PRICE FORECASTING TAKING INTO ACCOUNT THE POPULARITY OF SEARCH QUERIES

The article proposes to study the impact of behavioral factors on forecasting the stock price, and presents the model of the above-mentioned forecast. The factors that are supposed to be taken into account are used as indicators of the popularity of search queries within the certain topic. Two companies (Apple and Royal Dutch Shell) were selected for research, because they have different development histories. The Glossary based on papers on this subject has been created; it consists of 67 terms of economic, social and political meaning reflecting sentiment-oriented behaviors of traders, and the criteria for search of high popularity queries have been determined. We assume that the queries characterized by high popularity index make a great impact on the dynamics of the stock price. Besides, the article addresses the issue of neural networks as upon researching we observe that they might serve as indicators of the popularity of search queries. Furthermore, we introduced the forecast for two companies from different sectors of the economy. In addition, forecasting was made with various

combinations of these search queries, which were combined according to the semantic load. During the experiments it was revealed that for the information-sector company the quality of the model increased significantly due to the inclusion of behavioral factors, while for the processing-sector company the data did not significantly improve the forecast. This follows from the specificity of the study. The results can indicate the adequacy of the constructed models and confirm the feasibility of using the popularity indicators of search queries for forecasting the stock price.

Key words: stock price forecasting, behavioral factors, search queries, neural network.

JEL Classification: C45, C51, E47, G17.

Е. Ю. Кононова

доктор економічних наук, доцент

Харківський національний університет ім. В.Н. Каразіна

kateryna.kononova@gmail.com

А. С. Усачёва

студент кафедри економічної кібернетики і прикладної економіки

Харківський національний університет ім. В.Н. Каразіна

anastasiausachova@gmail.com

ПРОГНОЗИРОВАНИЕ КУРСА АКЦИЙ С УЧЕТОМ ПОКАЗАТЕЛЕЙ ПОПУЛЯРНОСТИ ПОИСКОВЫХ ЗАПРОСОВ

В статье предложена модель прогнозирования курса акций с учетом поведенческих факторов, в качестве которых использованы показатели популярности поисковых запросов заданной тематики, полученные с помощью инструмента Google Trends. Для исследования были выбраны две компании с разной историей развития: Apple и Royal Dutch Shell. На основе исследований, посвященных поведенческим финансам, сформирован словарь из 67 слов экономического, социального и политического характера, отражающих настроения трейдеров, и выявлены те поисковые запросы, показатель популярности которых в наибольшей степени влияет на динамику курса акций. С использованием нейронных сетей, на вход которых, кроме прочего, подавались показатели популярности поисковых запросов, проведено прогнозирование курса для двух компаний из разных секторов экономики. Также было осуществлено прогнозирование с различными комбинациями этих поисковых запросов, которые были объединены по смысловой нагрузке. Для построения нейронных сетей использовали приложение Neural Network Toolbox программного пакета MathLab. Проведенные экспериментальные исследования с разной архитектурой нейронных сетей, были использованы трехслойные и двухслойные нейронные сети. В ходе экспериментирования выявлено, что для компании информационного сектора качество модели существенно возросло за счет включения поведенческих факторов, в то время как для компании перерабатывающего сектора значительного улучшения прогноза не наблюдалось. Что, впрочем, естественным образом следует из специфики исследования. Полученные результаты свидетельствуют об адекватности построенных моделей и подтверждают целесообразность использования показателей популярности поисковых запросов для прогнозирования курса акций.

Ключевые слова: прогнозирование курса акций, поведенческие факторы, поисковые запросы, нейронная сеть.

JEL Classification: C45, C51, E47, G17.

Постановка проблеми. Кожної години акції тисячі компаній змінюють свою вартість. Прогнозування змін курсу акцій є актуальним та важливим завданням для власників компаній, професійних трейдерів та багатьох інших учасників фінансових ринків. В другій половині ХХ століття з'явилися дослідження, які наголошують на нерациональності вибору людини під час купівлі-продажу товарів та послуг, яка притаманна також й інвесторам. На думку дослідників, здійснить людина купівлю-продаж акцій чи ні, доволі сильно залежить від її емоційного стану в конкретний момент часу. Різні поведінкові фактори можуть вплинути на прийняття нерационального рішення. Це дослідження присвячено перевірці гіпотез поведінкової економіки щодо доцільності використання показників популярності пошукових запитів для поліпшення якості моделей прогнозування курсів акцій.

Аналіз останніх досліджень і публікацій. У кінці 70-х років ХХ століття виник новий напрям в науці – поведінкові фінанси, який поєднує економіку та психологію поведінки людини. Початок цьому напрямку поклали експерименти Данієла Канемана та Амоса Тверські, які виявили та описали ефекти, що впливають на вибір людини – ефект якоря, ефект доступності, ефект відображення та інші. В сукупності ці ефекти сформували теорію можливостей (Канеман, 2013). У 2017 році Нобелівську премію з економіки отримав професор Чиказького

університету Річард Талер за внесок у вивчення поведінкової економіки. У своїй останній книзі він досліджує такі особливості людської поведінки як обмежена раціональність, соціальні переваги та брак самоконтролю. Особливу увагу Талер приділяв поведінці інвесторів, він стверджує, що рішення про купівлю-продаж акцій може прийматися в залежності від того сита або голодна людина, виграла чи програла його улюблена команда (Талер, 2017).

У рамках концепції поведінкових фінансів Johan Bollen, Huina Mao та Xiao-Jun Zeng запропонували модель прогнозування індексу Доу-Джонса на основі колективних настроїв. Ними аналізувався текстовий зміст щоденних постів в Twitter, що з використанням сервісу Google Profile of Mood States (GPOMS) оцінювався за 6 параметрами: спокійний, насторожений, впевнений, енергійний, добрий і щасливий. Отримані часові ряди суспільного настрою вчені співвідносили зі значенням індексу Доу-Джонса. На основі самоорганізованої розмитої нейронної мережі, на вхід якої подавалися значенні настроїв, ними було отримано прогноз, точність якого складала близько 87,6 % (Bollen, 2011).

Науковці бізнес-школи Уорика Tobias Preis, Helen Susannah Moat та H. Eugene Stanley запропонували модель прогнозування індексу Доу-Джонса на основі аналізу пошукових запитів в Google. За допомогою сервісу Google Trends ними була зібрана інформація щодо популярності пошукових запитів (в основному це терміни, пов'язані з фондовим ринком) з 2004 по 2011 рік. Для виявлення залежності між пошуковими запитами та величиною індексу був використаний симулятор інвестиційних торгів (Tobias, Moat, Stanley, 2013). Damien Challet та Ahmed Bel Hadj Ayed для прогнозування часових рядів SPY, які відображають індекс «Standard and Poors 500», використовували не тільки слова пов'язані з фінансами та економікою, але й із захворюваннями, класичними марками автомобілів, комп'ютерними іграми (Challet, Bel Hadj Ayed A., 2014). Ці дослідження показали доцільність використання даних Google Trends для прогнозування динаміки фондового ринку.

Мета та завдання роботи. Мета статті полягає в дослідженні впливу поведінкових факторів на курс акцій компаній, що належать до різних секторів економіки – традиційних (переробного) та нових (інформаційного). Для її досягнення в роботі були поставлені наступні завдання: сформувати пошукові словники, провести збір і підготовку даних, оцінити ступінь взаємозв'язку динаміки популярності пошукових запитів і курсу акцій та сформувати вектори входів нейронної мережі, розробити пул моделей багатосарових нейронних мереж для прогнозу вартості акцій, провести серію експериментів з мережами різної архітектури, проаналізувати отримані результати та обґрунтувати доцільність використання поведінкових факторів при прогнозуванні курсів акцій.

Перший етап – базова модель. Для дослідження було обрано дві компанії, що належать до різних секторів економіки – переробного (Royal Dutch Shell) та інформаційного (Apple).

Дані про вартість акцій компаній Apple та Shell було зібрано на момент закриття роботи біржі (Сервіс Google Finance, 2017) за період з 1 січня 2010 року по 21 вересня 2017 року. Курс акцій Shell не мав сильних коливань за останні роки, курс Apple змінювався доволі часто та коливався у великих діапазонах (рис. 1). Обидва ряди є нестационарними, тому апробація нетрадиційних підходів до їх прогнозування, заснованих на методах поведінкових фінансів, є актуальною задачею.

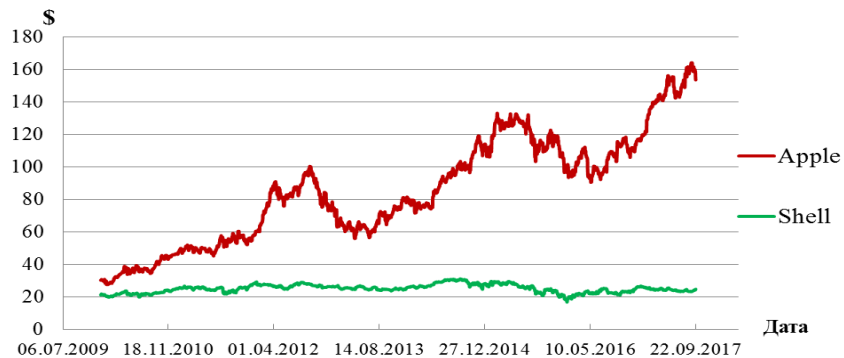


Рис. 1 – Динаміка зміни курсу акцій компаній Apple та Shell

Джерело: побудовано авторами

На першому етапі експериментування для обох компаній було розроблено моделі нейронних мереж, на вхід яких подавалися лише попередні значення часового ряду. Було визначено, що для динаміки акцій компанії Shell значущими є значення курсу за 2 минулі дні, для Apple – за 4 попередні дні.

Для побудови нейронних мереж було використано додаток Neural Network Toolbox програмного пакету MathLab. В результаті серії експериментів по критерію мінімуму помилки тестування було виявлено найкращі архітектури мереж для обох компаній. Для Apple це двошарова мережа з 5 нейронами в першому скритому шарі, помилка тестування на неї склала $MSE = 1,3894$. Для Shell це двошарова мережа з 3 нейронами в першому скритому шарі, помилка тестування – $MSE = 0,2812$. З цими контрольними значеннями будемо порівнювати результати подальших експериментів.

Другий етап – модель з урахуванням поведінкових факторів. Для виявлення настроїв, що можуть впливати на курс акцій Apple та Shell, на основі описаних вище досліджень та дослідження (Peter, 2015) було складено словник, до якого крім слів економічного, соціального та політичного характеру, що відображають настрої трейдерів, увійшли самі назви компаній Apple та Shell. Складений таким чином словник містить 67 слів.

Для того, щоб проаналізувати як часто люди цікавились цими словами в період з 1 січня 2010 року по 21 вересня 2017 року було використано інструмент Google Trends (Сервіс Google Trends, 2017). В результаті отримано ряд з 2821 значень для 67 пошукових запитів.

Для уточнення прогнозу курсу акцій Apple, отриманого на основі динаміки часового ряду, було відібрано 14 слів, для яких кореляція показника популярності пошукового запиту та курсу акцій була найбільша, а саме: gas, chance, forex, freedom, labor, markets, money, oil, sell, Shell, trader, travel, war, world. Також було здійснено прогнозування з різними комбінаціями цих слів, які були об'єднані за смисловим навантаженням. Такі експерименти були проведені в двох варіантах: в першому кількість вхідних векторів збільшено на кількість слів, а в другому – розраховано загальну суму показників популярності пошукових запитів за кожен день, яка подавалась на вхід мережі одним значенням. На рис. 2 зображена архітектура нейронної мережі, яка використовувалась при прогнозуванні вартості акцій компанії Apple.

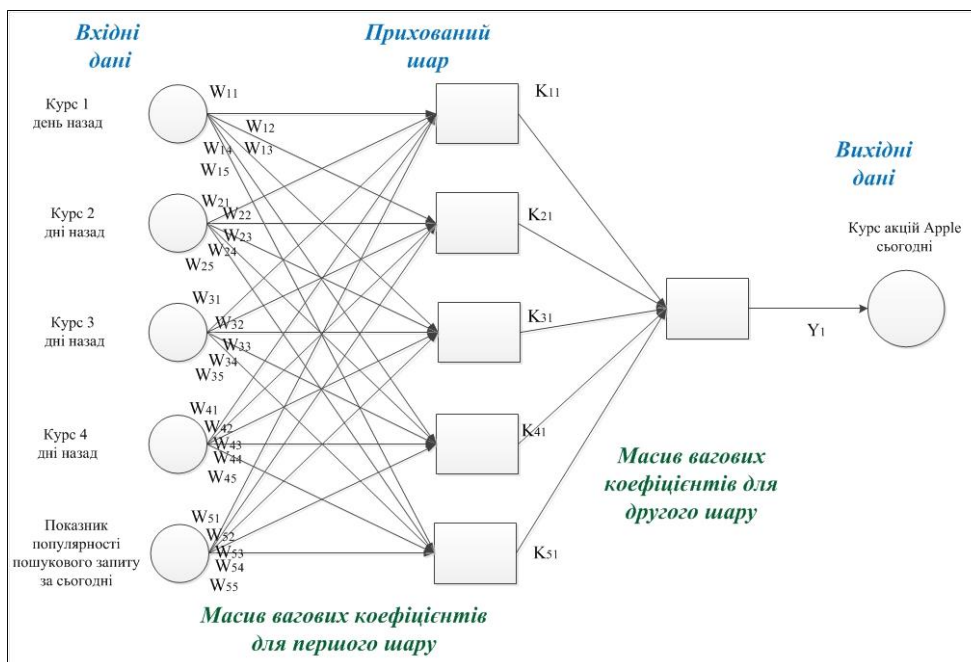


Рис. 2 – Архітектура двошарової нейронної мережі для компанії Apple

Джерело: створено авторами

Загалом з цими словами було проведено 137 експериментів, з них 32 (21% від усіх експериментів) показали кращі результати, ніж при прогнозуванні без пошукових запитів. Для кожного слова підбиралась найкраща архітектура нейронної мережі. В табл. 1 представлено результати 10 експериментів для курсу акцій компанії Apple, які дали найменшу похибку та експеримент, результати якого були використані в якості контрольного значення.

Таблиця 1

Результати прогнозування курсу акцій компанії Apple

№	Пошуковий запит	Кількість шарів	Кількість нейронів			Дані по результатам навчання	Дані по результатам тестування
			1 шар	2 шар	3 шар	Похибка навчання, *103)	MSE, *103
1	—	2	5	1		0.0867	1.3894
2	sell	3	5	4	1	0.0844	0.3652
3	Shell	3	4	4	1	0.0858	0.4279
4	forex	2	6	1		0.0851	0.4561
5	markets, sell, money	2	5	1		0.0650	0.4963
6	forex, trader	2	5	1		0.0830	0.5002
7	markets	2	5	1		0.0882	0.5634
8	sell	3	4	4	1	0.0846	0.5798
9	travel	3	5	4	1	0.0799	0.6158
10	travel	2	4	1		0.0879	0.6755
11	gas, oil	2	5	1		0.0863	0.7069

Джерело: створено авторами

Найкращі результати було отримано на трьох-шаровій нейронній мережі [5, 4, 1], на вхід якої подавався пошуковий запит «sell». З ним отримано мінімальне середнє значення суми квадратів відхилень на тестовій вибірці – 0,3652, що практично в чотири рази краще за базовий варіант.

На рис. 3 представлено порівняння прогнозних даних, отриманих в експериментах 1 та 2 (табл. 1). Бачимо, що модель з пошуковим запитом «sell» дуже точно прогнозує реальні дані, особливо це помітно на тій частині ряду, де в контрольному варіанті спостерігалось значне відхилення.

Для компанії Shell за результатами кореляційного аналізу було обрано 12 пошукових запитів: chance, crisis, financial market, food, freedom, markets, marriage, oil, ring, society, trader, war. Також окрім саме слів було використано їх комбінації у двох варіантах. Усього для прогнозування вартості акцій компанії Shell проведено 102 експерименти. В 21 експерименті (21% від усіх варіантів) середнє значення суми квадратів відхилень було менше контрольного варіанту.

У табл. 2 наведені результати 10 найбільш успішних експериментів для компанії Shell та результати найкращого експерименту серед тих, які не враховували пошукові запити.

Найкращі результати було отримано на трьох-шарових мережах [4, 3, 1], на вхід яких подавалися пошукові запити society та crisis, MSE=0,2797 (експерименти 2 та 3, табл. 2). Але на відміну від попередньої серії експериментів, в цьому випадку не було значного покращення якості прогнозу – результати майже не відрізняються один від одного. Якщо при прогнозуванні курсу акцій Apple похибка зменшилася майже в 4 рази, то в даному випадку –лише на 0,5%.

На рис. 4 зображено фрагмент прогнозних даних, отриманих в експериментах 1 та 2 (табл. 2), вони майже не відрізняються один від одного.

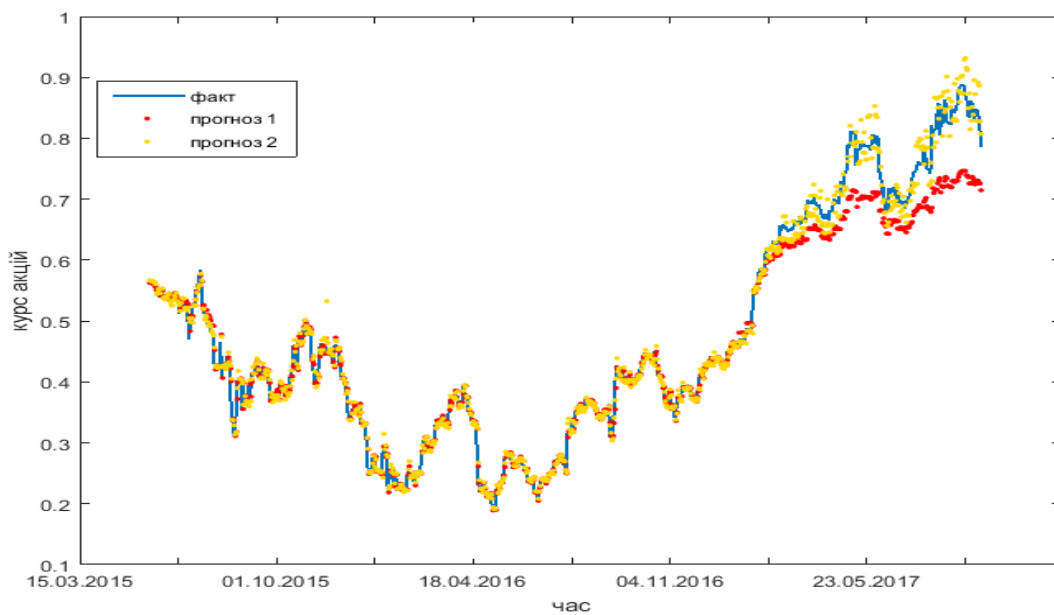


Рис. 3 – Прогнозні дані, отримані в експериментах 1 та 2 для компанії Apple
Джерело: створено авторами

Таблиця 2

Результати прогнозування курсу акцій компанії Shell

№	Пошуковий запит	Кількість шарів	Кількість нейронів			Дані по результатам навчання	Дані по результатам тестування
			1 шар	2 шар	3 шар	Похибка навчання, *103	MSE, *103
1	—	2	3	1		0.3870	0.2812
2	society	3	4	3	1	0.3840	0.2797
3	crisis	3	4	3	1	0.3900	0.2797
4	financial market	2	4	1		0.3910	0.2798
5	society	2	3	1		0.3790	0.2798
6	marriage+ring	3	3	3	1	0.3920	0.2798
7	marriage+ring	2	3	1		0.3910	0.2798
8	Oil	2	3	1		0.3920	0.2799
9	trader	2	3	1		0.3900	0.2799
10	freedom	2	3	1		0.3840	0.2800
11	marriage	3	4	4	1	0.3870	0.2801

Джерело: створено авторами

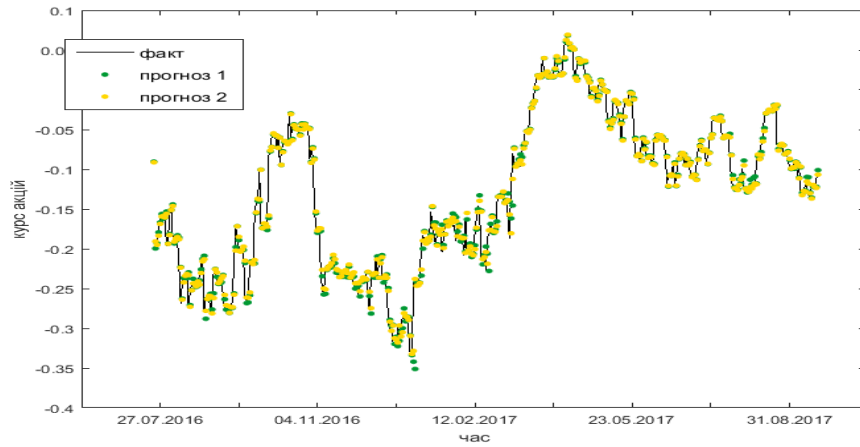


Рис. 4 – Прогнозні дані, отримані в експериментах 1 та 2 для компанії Shell

Джерело: створено авторами

Висновки

Результати експериментування свідчать, що для компаній інформаційного сектора моделі прогнозування з використанням поведінкових факторів дозволяють суттєво поліпшити якість прогнозу. Можна сказати, що вартість акцій таких компаній доволі сильно залежить від людських настроїв. В той же час на вартість акцій компаній традиційних секторів економіки настрої, скоріш за все, впливають в меншій мірі.

Ми мали 6 пошукових запитів, які використовувались в експериментах для Apple та в експериментах для Shell, але пошукові запити, які дають найкращі результати прогнозування, для обох компаній різні. Тільки запит *trader* зустрічається в обох таблицях: в першій в комбінації з іншими пошуковими запитами, в другій – як самостійне слово. Цікаво, що при прогнозуванні курсу акцій компанії Apple найбільш впливовими стали слова, пов'язані з економікою: *sell*, *markets*, *forex*, *money*, *trader*. А на вартість акцій компанії Shell більш впливають запити, пов'язані з соціальними аспектами: *society*, *freedom*, *crisis*.

Порівнюючи отримані результати з результатами вчених бізнес-школи Уоріка, бачимо, що в їх дослідженнях найбільш впливовими виявились слова: *debt*, *inflation*, *economics*, *metals*, *stocks*. В нашому випадку ці слова не є значущими. Та навпаки, ті слова, які в роботі дослідників бізнес-школи Уоріка впливали найменшим чином на індекс Доу-Джонса, доволі непогано покращують якість прогнозування в нашому випадку: *ring*, *labor*, *world*, *freedom*, *financial markets*.

Література

1. Канеман Д. Думай медлено, решай быстро. [Текст] / Д. Канеман; пер. з англ. Н. Парфенова. – М.: АСТ, 2013. – 710 с.
2. Талер Р. Х. Новая поведенческая экономика. [Текст] / Р.Х. Талер; пер. з англ. А. Е. Прохорова. – М.: Эксмо, 2017. – 368 с.
3. Bollen J. Twitter mood predicts the stock market / J. Bollen, M. Huina, Z. Xiao-Jun // Journal of Computational Science. – 2011. – Vol. 2(1).
4. Tobias P. Quantifying trading behavior in financial markets using Google Trends / P. Tobias, H.-S. Moat, E. Stanley // Scientific reports. – 2013. – Vol. 3 (1684).
5. Challet D., Bel Hadj Ayed A. Predicting financial markets with Google Trends and not so random keywords [Electronic resource]. – Accessed mode : <https://arxiv.org/pdf/1307.4643v3.pdf>.
6. Сервіс Google Finance [Електронний ресурс]. – Режим доступу : <https://www.google.com/finance>.
7. Eckley Peter. Measuring economic uncertainty using news-media textual data [Electronic resource]. – Accessed mode : https://mpira.ub.uni-muenchen.de/64874/1/MPRA_paper_64874.pdf.
8. Сервіс Google Trends [Електронний ресурс]. – Режим доступу : <https://trends.google.ru/trends/>

References

1. Kahneman, D. (2013) Thinking, Fast and Slow. Moscow: AST (in Russian).
2. Thaler, Richard H. (2017). New Behavioral Economics. Moscow : Eksmo (in Russian).
3. Bollen, J., Huina, M., Xiao-Jun, Z. (2011). Twitter mood predicts the stock market. Journal of Computational Science.
4. Tobias, P., Moat, H.-S., Stanley, E. (2013). Quantifying trading behavior in financial markets using Google Trends. Scientific reports.
5. Challet D., Bel Hadj Ayed A. (2014). Predicting financial markets with Google Trends and not so random keywords. Retrieved from <https://arxiv.org/pdf/1307.4643v3.pdf>.
6. Google Finance. Retrieved from <https://www.google.com/finance>.
7. Peter, E. (2015). Measuring economic uncertainty using news-media textual data. Retrieved from https://mpira.ub.uni-muenchen.de/64874/1/MPRA_paper_64874.pdf.
8. Google Trends. Retrieved from <https://trends.google.ru/trends/>.