

ПОСТРОЕНИЕ И АНАЛИЗ МАКРОЭКОНОМИЧЕСКИХ СУБИНДЕКСОВ ЦЕН НА ОСНОВЕ ДАННЫХ ИЗ ИНТЕРНЕТ

А. В. Якубович

Белорусский государственный университет, г. Минск;

hanna.yakubovi@gmail.com;

науч. рук. – В. И. Малюгин, канд. физ.-мат. наук, доц.

В статье представляются инструментальные средства и полученные с их применением результаты оценивания и анализа динамики изменения стоимости продовольственных товарных групп из потребительской корзины Республики Беларусь на основе соответствующих субиндексов. Для оценивания стоимости потребительских корзин используется информация с сайта самой крупной торговой сети Республики Беларусь.

Ключевые слова: потребительская корзина; анализ интернет-данных; краткосрочное прогнозирование; машинное обучение; субиндексы потребительских цен.

ВВЕДЕНИЕ

Традиционно для решения разнообразных задач анализа текущего состояния и прогнозирования экономических процессов используется статистика, предоставляемая государственными статистическими органами. Многие макроэкономические показатели, например, ВВП, индекс потребительских цен и др., рассчитываются со значительной временной задержкой и допускают последующие уточнения. Это ограничивает возможности их применения не только для построения своевременных прогнозов, но и для оценки экономической ситуации в текущий момент времени. В этих условиях в мировой практике все большее внимание уделяется построению экономических индикаторов на основе «микро-данных» (*firm-level data*), поступающих непосредственно от различных категорий участников экономических отношений [1, 2]. В последнее время все более актуальным становится новый подход, известный как «наукастинг» (*nowcasting*), целью которого является анализ и краткосрочное прогнозирование экономических индикаторов в режиме реального времени [3].

Исследования, представляемые в данном докладе, можно отнести к данному подходу. Целью работы является разработка алгоритмического и программного инструментария, предназначенного для оценки и краткосрочного прогнозирования в режиме онлайн индексов цен на продовольственные товары из потребительской корзины и отдельных товарных групп на основе ежедневных данных публикуемых в Интернет самой крупной торговой сетью потребительских товаров на территории

Республики Беларусь. Основным результатом исследования является разработка программы, позволяющей решать следующие задачи: извлечение данных с сайта торговой сети, очищение и нормировка данных; морфологический анализ данных, классификация и расчет стоимости потребительской корзины и отдельных товарных групп.

ФОРМИРОВАНИЕ ТОВАРНЫХ ГРУПП НА ОСНОВЕ ДАННЫХ ИЗ СЕТИ ИНТЕРНЕТ

Продуктовая часть потребительская корзина включает в состав несколько видов продовольственных товаров со своим удельным весом. Таким образом для ее подсчета сначала оценивается удельная стоимость отдельных товарных групп. Все данные в сети Интернет хранятся на веб-страницах. Формирование товарных групп на их основе представляет собой процесс, состоящий из следующих этапов: загрузка веб-страницы и извлечение информации из нее (web scrubbing), предобработка данных (очистка и приведение к консистентному виду), классификация полученных данных по товарным группам. На основе полученных данных оценивается удельная стоимость всех товарных групп m_i . Тогда стоимости потребительской корзины оценивается по формуле:

$$C = \sum_{i=1}^n w_i m_i, \quad (1)$$

где w_i – удельный вес i -ой товарной группы.

ОЦЕНИВАНИЕ СТОИМОСТИ ПОТРЕБИТЕЛЬСКОЙ КОРЗИНЫ РЕСПУБЛИКИ БЕЛАРУСЬ

Как известно, *потребительская корзина* – представляет собой перечень товаров на месячный или годовой период, который необходим для обеспечения нормального уровня жизни человека (семьи). В рамках данного исследования оценивается продуктовая составляющая потребительской корзины для трех социально-демографических групп населения: пенсионной, детской и трудоспособной. Согласно рис. 1 официально установленные корзины для различных групп населения различаются по структуре. Считается, что трудоспособному населению необходимо больше мяса, овощей и хлеба, детям необходимо употреблять больше молока и фруктов, пенсионеры, как правило, употребляют меньше рассматриваемых продуктов, чем трудоспособное население и даже дети.

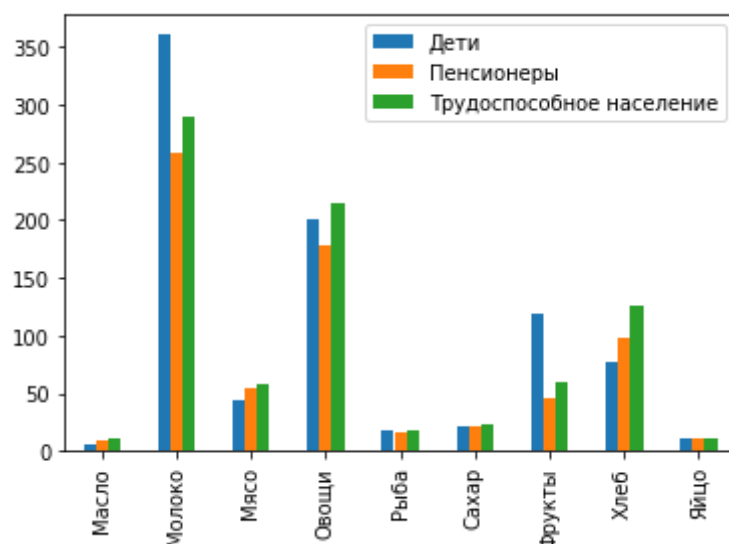


Рис. 1 Сравнение корзин трудоспособного населения, пенсионной и детской

Официальные стоимости этих корзин последний раз были утверждены в Постановлении Министерства труда и социальной защиты Республики Беларусь в 2018 г. В соответствии с этим размеры минимальных потребительских бюджетов для разных социально-демографических групп определяются следующим образом: трудоспособное население – 446 белорусских рублей 2 копейки; пенсионеры – 349 белорусских рублей 35 копеек; дети – 379 белорусских рублей 23 копейки. Из них продукты питания составляют 48.6 %. В табл. 1 приводятся официальные стоимости и статистические оценки стоимости продуктовой составляющей потребительской корзины на ноябрь 2018 г. для различных групп населения.

Таблица 1

Сравнение стоимостей продуктовой составляющей потребительских корзин

Группа	Официальная стоимость, руб.	Статистическая оценка, руб.
Дети	2211.00	2105.50
Пенсионеры	2037.41	2040.50
Трудоспособное население	2601.19	2476.71

Из табл. 1 следует, что оценка стоимости потребительской корзины пенсионеров практически совпадает с официально установленным значением. Для трудоспособного населения и детей есть расхождения (около 5 %).

Причина расхождений может быть выявлена с помощью построенных субиндексов потребительских цен, которые вычисляются на основе дан-

ных торговой сети, публикуемых в Интернет. В работе используются цепные индексы в виде темпов роста оценки стоимости некоторой потребительской корзины (или товарной группы) вида (1) по отношению к предыдущему периоду (моменту) времени:

$$s(t) = m(t) / m(t-1), \quad (2)$$

где $m(t)$ – оценка удельной стоимости товарной группы в момент времени t .

На рис. 2 приводятся значения субиндексов для двух товарных групп за шесть месяцев наблюдений с ноября по май 2019 г..

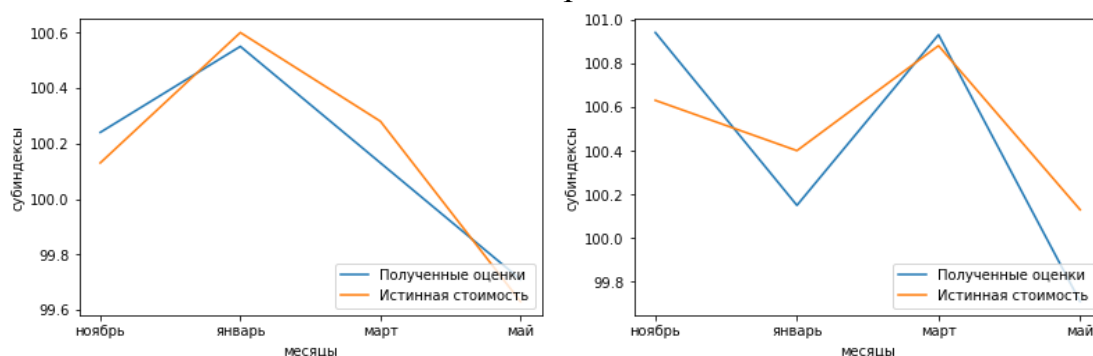


Рис. 2 Месячные изменения стоимости корзин:
слева – товарная группа «яйца», справа – товарная группа «хлеб»

Из рис. 2 видно, что значение субиндекса для товарной группы «хлеб» в ноябре меньше, чем официально установленная стоимость, что может быть следствием некоторых краткосрочных факторов. Также стоит заметить, что для товарной группы «яйца» расхождений почти нет, что объясняется постоянством цен на продукты данной группы.

Библиографические ссылки

1. Малюгин В. И., Новопольцев А. Ю., Гринь Н. В. Статистический анализ кредитоспособности реального сектора белорусской экономики на основе микроданных // Банковский вестник. 2018. № 14. С. 4–98.
2. Малюгин В. И., Крук Д. Э., Милевский П. С. Индекс экономических настроений белорусской экономики: методические, модельные и инструментальные средства построения и применения // Банковский вестник. 2019. № 16. С. 4–29.
3. Якубович А. В. О краткосрочном прогнозировании потребительских цен на основе интернет-данных торговых сетей в режиме онлайн // 10-я Междунар. науч.-практ. конф. студентов и аспирантов ВШЭ : сб. трудов. М., 2019. С. 202.