

А. И. Бредихин

### **ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ ЧИСЛЕННОСТИ НАСЕЛЕНИЯ РЕГИОНА**

*В данной статье описывается применение нейронной сети, а именно многослойного перцептрона в задаче прогнозирования численности населения региона на примере Ханты-Мансийского автономного округа – Югры. При решении данной задачи использованы данные о численности населения ХМАО-Югры за период с 1989 по 2018 год, на основании которых сделан прогноз численности населения ХМАО-Югры на 2019–2029 годы. По результатам прогнозирования сделаны выводы о применимости нейронных сетей в задаче прогнозирования.*

*Ключевые слова: искусственный интеллект, многослойный перцептрон, численность населения, прогнозирование, демография, регион.*

A. I. Bredikhin

### **THE USAGE OF A NEURAL NETWORKS FOR FORECASTING POPULATION A REGION**

*This article describes the use of a neural network, namely, a multi-layer perceptron in the problem of forecasting the population of the region on the example of the Khanty-Mansiysk Autonomous Okrug-Ugra. In solving this problem, we used data on the population of KHAMAO-Yugra for the period from 1989 to 2018, on the basis of which the forecast of the population of KHAMAO-Yugra for 2019–2029 was made. Based on the results of forecasting, conclusions are made about the applicability of neural networks in the forecasting problem.*

*Keywords: artificial intellect, multi-layer perceptron, population, forecasting, demography, region.*

---

### **Введение**

В настоящее время достаточно популярны методы прогнозирования временных рядов. Это объясняется потребностью людей «заглядывать в будущее» для эффективного управления процессами, ресурсами и т. п. Одним из примеров временного ряда является численность населения (страны, региона, города и т. д.).

Существуют различные методы прогнозирования временных рядов. Среди них выделяют методы регрессионного анализа, к которым относят метод наименьших квадратов [1], цель которого – поиск ближайшей к наблюдаемым точкам прямой, и метод максимального правдоподобия [2]. Сюда же относят и авторегрессионные модели, в которые заложено предположение о зависимости значения процесса от некоторого количества предыдущих значений [3]. Отдельно выделяются нейросетевые методы прогнозирования временных рядов [4].

Основным преимуществом нейросетевых методов прогнозирования временных рядов по сравнению с методами регрессионного анализа является их нелинейность [5], т. е. способность выявлять нелинейные зависимости между текущими и прогнозируемыми значениями процессов. Именно поэтому было принято решение рассмотреть в данной работе именно нейронные сети для решения задачи прогнозирования численности населения региона.

Прогнозирование численности населения позволит сделать определенные выводы по дальнейшему развитию инфраструктуры и экономики федерального округа, региона, города, что важно как для городов и регионов, так и для федеральных округов и страны в целом. Со-

гласно полученным выводам главы городов, регионов смогут принимать оптимальные стратегии развития, что при их реализации повысит уровень жизни населения.

В данной статье описано прогнозирование численности населения региона на примере ХМАО-Югры. Для этого использовался многослойный перцептрон и метод скользящего окна. Для обучения перцептрона использовались данные о численности населения ХМАО-Югры за период 1989–2018 гг. На их основании сделан прогноз численности населения на 2019–2029 гг.

### Полносвязный перцептрон

Полносвязный перцептрон – один из простых видов нейронных сетей. Тем не менее, при обучении он может научиться решать задачи, которые трудно поддаются формализации или не поддаются ей вообще. При этом все зависит от настроек сети.

Простейший перцептрон состоит из входного и выходного слоев нейронов. Нейроны соседних слоев связаны друг с другом по принципу «каждый с каждым». Архитектура простейшего перцептрона показана ниже.

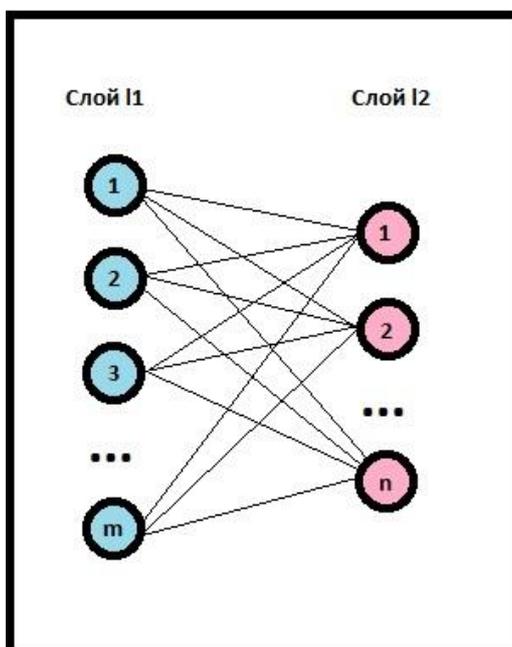


Рисунок 1 – Архитектура простейшего перцептрона

Каждая связь между двумя нейронами имеет собственный весовой коэффициент, а каждый нейрон имеет собственный порог (или смещение). Все весовые коэффициенты и пороги являются настраиваемыми в ходе обучения сети параметрами.

Количество слоев нейронов в перцептроне может быть любым. Тогда перцептрон будет иметь, помимо входного и выходного, скрытые слои нейронов.

Рассмотрим **функционирование перцептрона**. Пусть первый слой нейронов состоит из  $m$  нейронов, а второй – из  $n$  нейронов (см. рис. 1). Тогда вычисление значений нейронов второго слоя производится по следующим формулам [6]:

$$l_{2j} = F(S_{2j}), j = \overline{1..n}, \quad (1)$$

$$S_{2j} = \left( \sum_{i=1}^m l_{1i} * \omega_{ij} \right) - T_{2j}, \quad (2)$$

где  $F$  – значение функции активации от взвешенной суммы  $S_{2j}$ ;

$l_{1i}$  и  $l_{2j}$  – значения  $i$ -го нейрона первого слоя нейронов и  $j$ -го нейрона второго слоя соответственно;

$\omega_{ij}$  – значение связи между  $i$ -м нейроном первого слоя нейронов и  $j$ -м нейроном второго слоя соответственно;

$T_{2j}$  – значение порога (смещения)  $j$ -го нейрона второго слоя.

Для обучения многослойного перцептрона используется **алгоритм обратного распространения ошибки (АРО)**, который относится к методам обучения с учителем.

Рассмотрим **обучение многослойного перцептрона**. Ошибка формируется на последнем слое нейронов перцептрона и определяется как разность между выходной реакцией сети (значениями нейронов последнего слоя нейронов)  $y$  и эталоном  $t$  [6, с. 63]:

$$\gamma_j = y_j - t_j. \quad (3)$$

Далее происходит изменение значений весов и порогов по следующим формулам [6, с. 63]:

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \alpha \gamma_j F'(S_j) y_i, \quad (4)$$

$$T_j(t+1) = T_j(t) + \alpha \gamma_j F'(S_j), \quad (5)$$

где  $\alpha$  – скорость обучения сети;

$t$  и  $t+1$  – моменты времени до и после изменения весов и порогов соответственно;

индексы  $i$  и  $j$  обозначают нейроны первого и второго слоя нейронов соответственно.

Слои перцептрона обучаются по *процедуре обучения Розенблатта*, согласно которой значение скорости обучения постоянно в процессе всего времени обучения и принимает значения в промежутке  $(0;1]$  [7, с. 43].

### Обработка данных

Как уже было сказано ранее, для прогнозирования были взяты значения численности населения ХМАО-Югры за 1989 – 2018 годы. Они приведены в таблице 1.

Таблица 1 – Численность населения ХМАО за 1989–2018 годы [8]

|                             |             |             |             |             |             |             |
|-----------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| <b>Год</b>                  | <b>1989</b> | <b>1990</b> | <b>1991</b> | <b>1992</b> | <b>1993</b> | <b>1994</b> |
| Численность населения, чел. | 1268439     | 1267030     | 1280139     | 1271505     | 1267751     | 1279483     |
| Прирост                     | -           | -1409       | 13109       | -8634       | -3754       | 11732       |
| <b>Год</b>                  | <b>1995</b> | <b>1996</b> | <b>1997</b> | <b>1998</b> | <b>1999</b> | <b>2000</b> |
| Численность населения, чел. | 1292985     | 1303285     | 1316774     | 1342991     | 1359069     | 1359646     |
| Прирост                     | 13502       | 10300       | 13489       | 26217       | 16078       | 577         |
| <b>Год</b>                  | <b>2001</b> | <b>2002</b> | <b>2003</b> | <b>2004</b> | <b>2005</b> | <b>2006</b> |
| Численность населения, чел. | 1383449     | 1432817     | 1437729     | 1456509     | 1469011     | 1478178     |
| Прирост                     | 23803       | 49368       | 4912        | 18780       | 12502       | 9167        |
| <b>Год</b>                  | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011</b> | <b>2012</b> |
| Численность населения, чел. | 1488297     | 1505248     | 1519962     | 1532243     | 1537134     | 1561238     |
| Прирост                     | 10119       | 16951       | 14714       | 12281       | 4891        | 24104       |
| <b>Год</b>                  | <b>2013</b> | <b>2014</b> | <b>2015</b> | <b>2016</b> | <b>2017</b> | <b>2018</b> |
| Численность населения, чел. | 1584063     | 1597248     | 1612076     | 1626755     | 1646078     | 1655074     |
| Прирост                     | 22825       | 13185       | 14828       | 14679       | 19323       | 8996        |

Все значения численности населения были подвергнуты нормализации – они приведены в промежуток  $[0;1]$  путем деления.

Для прогнозирования данного временного ряда используется **метод скользящего окна**. Он характеризуется длиной окна  $p$ , равной числу элементов ряда, одновременно подаваемых на нейронную сеть [7, с. 56].

Обучающую выборку  $X$  и прогнозируемые значения  $Y$  в таком случае можно представить в следующем виде (длина временного ряда равна  $m$ ):

$$X = \begin{bmatrix} x(1) & x(2) & \dots & x(p) \\ x(2) & x(3) & \dots & x(p+1) \\ \dots & \dots & \dots & \dots \\ x(m-p) & x(m-p+1) & \dots & x(m-1) \end{bmatrix}, \quad Y = \begin{bmatrix} x(p+1) \\ x(p+2) \\ \dots \\ x(m) \end{bmatrix},$$

что эквивалентно перемещению окна по ряду  $x(t)$  с шагом, равным 1 [7, с. 56].

### Применение нейронной сети для прогнозирования

Для проектирования нейронной сети была определена длина скользящего окна исходя из длины прогнозируемого периода (2019–2029 гг.). Для данного периода требуется предсказать 11 значений численности населения (2029–2019+1). Следовательно, размер входа нейронной сети должен быть равным 11.

Затем был спроектирован многослойный перцептрон с одним скрытым слоем. Его архитектура описана в таблице 2.

Таблица 2 – Архитектура нейронной сети

| Номер слоя | Размер слоя (входа) | Размер матрицы весовых коэффициентов | Функция активации |
|------------|---------------------|--------------------------------------|-------------------|
| 1          | 11                  | 11*23                                | ReLU              |
| 2          | 23                  | 23*1                                 | Сигмоидная        |
| 3          | 1                   | -                                    | -                 |

После проектирования сеть была обучена на обучающей выборке. Количество эпох составило 2000, коэффициент скорости обучения – 0,25. Результаты обучения представлены ниже.

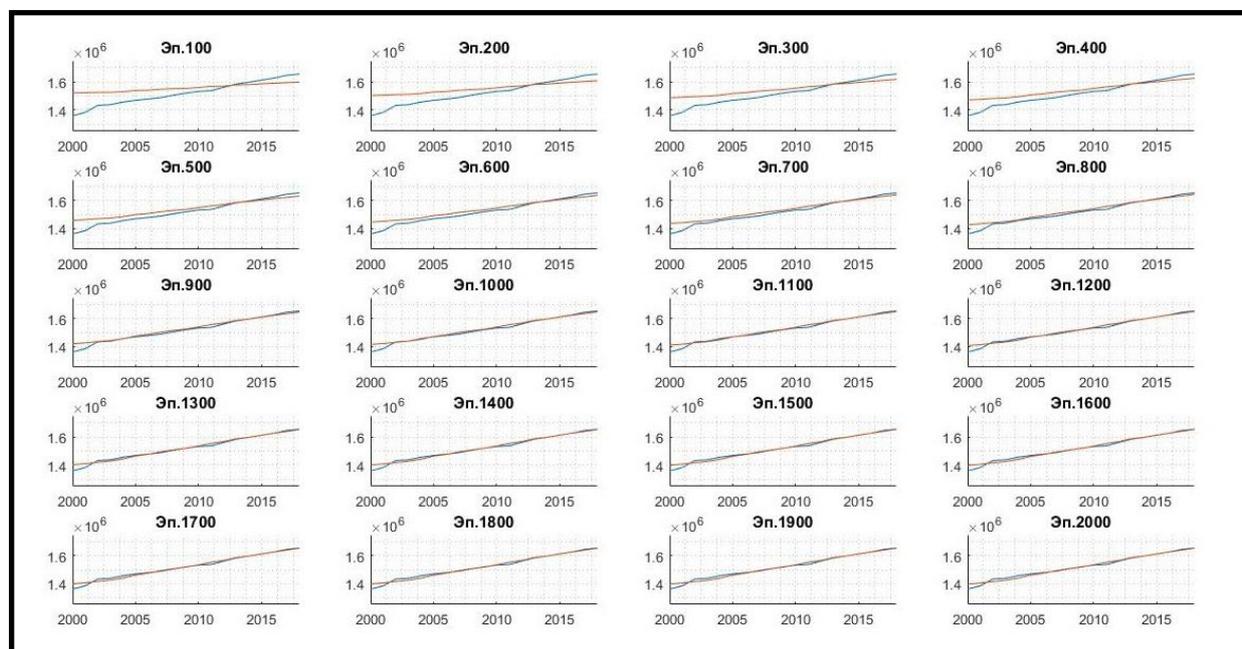


Рисунок 2 – Результаты обучения

**Замечание.** Синим цветом обозначен график динамики численности населения за 2000–2018 гг., а желто-оранжевым – график предсказанных нейронной сетью значений населения ХМАО за данные годы.

Как видно из рисунка 2, наблюдается постепенная сходимость реакции сети к динамике численности населения ХМАО, что свидетельствует об успешном процессе обучения сети.

Затем обучение нейронной сети было продолжено – количество эпох составило 2000 (с 2001-ю по 4000-ю), коэффициент скорости обучения был увеличен до 0,55. Результаты продолженного обучения представлены ниже.

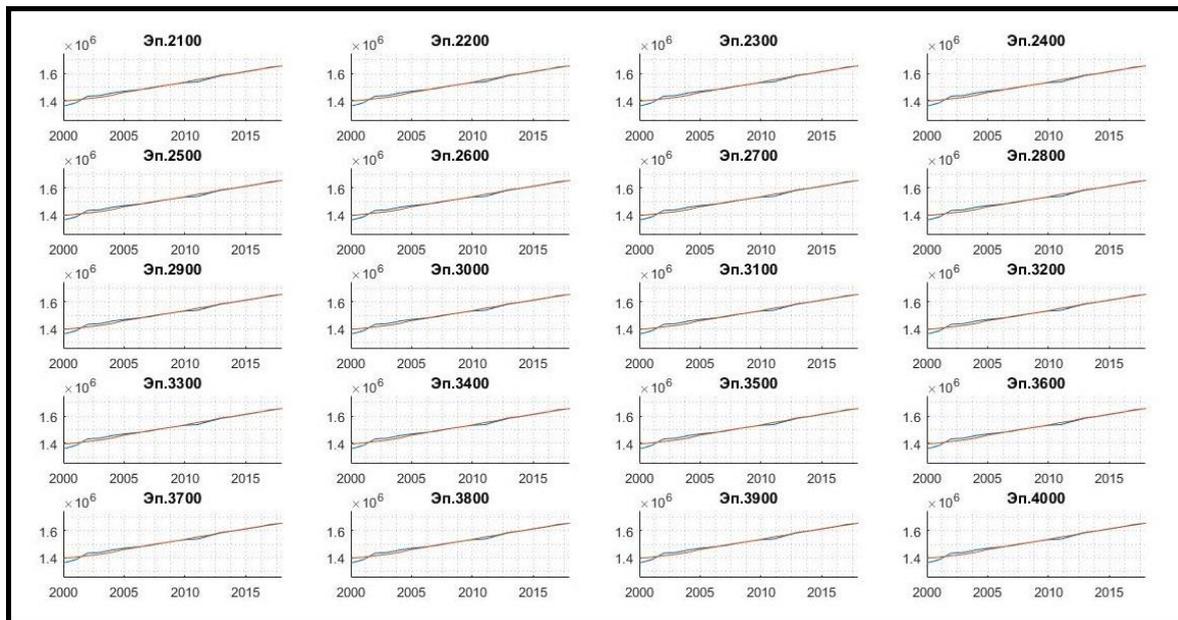


Рисунок 3 – Результаты продолженного обучения

Прогресса в обучении не наблюдается, следовательно, нейронную сеть можно считать обученной. При этом она не очень корректно подстраивается к резким перепадам численности населения.

### Результаты прогнозирования

На вход обученную нейронную сеть были поданы значения численности населения ХМАО за 2008–2018 годы, и таким образом получено спрогнозированное значение численности населения ХМАО за 2019 г. Данные действия были повторены для всех годов с 2020 по 2029. График предсказанных значений представлен ниже.

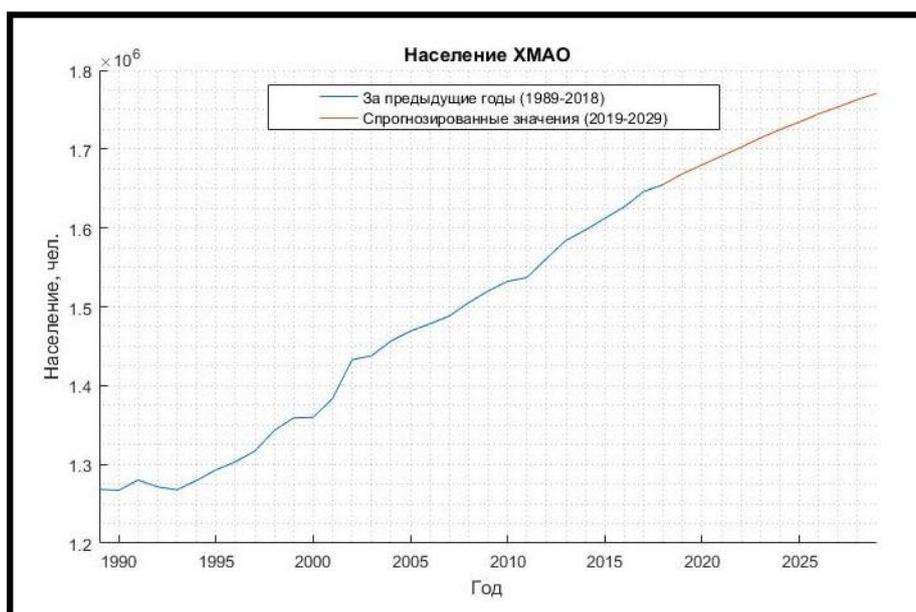


Рисунок 4 – Результаты прогнозирования численности населения ХМАО

Полученные с помощью прогнозирования значения численности населения ХМАО были сведены в таблицу 3.

Таблица 3 – Спрогнозированные значения численности населения ХМАО

| Год                         | 2019    | 2020    | 2021    | 2022    | 2023    | 2024    |
|-----------------------------|---------|---------|---------|---------|---------|---------|
| Численность населения, чел. | 1668701 | 1679951 | 1691125 | 1702113 | 1714102 | 1724625 |
| Прирост                     | 13627   | 11250   | 11174   | 10988   | 11989   | 10523   |
| Год                         | 2025    | 2026    | 2027    | 2028    | 2029    | –       |
| Численность населения, чел. | 1734281 | 1744638 | 1753508 | 1762900 | 1770805 | –       |
| Прирост                     | 9656    | 10357   | 8870    | 9392    | 7905    | –       |

**Замечание.** Спрогнозированные значения численности населения ХМАО-Югры в будущем будут близки к реальным при сохранении текущих тенденций развития округа.

Исходя из данных таблицы 3, население ХМАО-Югры будет продолжать расти. Однако темпы роста со временем будут снижаться. При сохранении текущих тенденций развития округа через несколько десятков лет население ХМАО-Югры может пойти на убыль, что может негативно сказаться на экономическом состоянии региона.

Одну из причин возможного в будущем падения темпов роста населения ХМАО-Югры автор статьи видит в слабо развитой инфраструктуре региона, а именно – большие расстояния между городами (порядка 200–300 км; исключение – города Нефтеюганск и Сургут, расстояние между которыми составляет около 60 км).

Для решения данной проблемы автор предлагает развитие инфраструктуры ХМАО-Югры с помощью усиленного развития городов с населением менее 100 тыс. человек и поселков.

### Заключение

В результате выполнения данной работы была спроектирована и обучена нейронная сеть, которая была применена для прогнозирования численности населения Ханты-Мансийского автономного округа – Югры. Сеть обучилась достаточно успешно, так что полученные с ее помощью данные о численности населения ХМАО-Югры на 2019–2029 годы можно считать в той или иной степени корректными. Из этого следует, что многослойный перцептрон вместе с применением метода скользящего окна вполне пригоден для прогнозирования временных рядов.

Что же касается анализа результатов прогнозирования, то согласно им в 2019–2029 гг. возможно падение темпов роста численности населения ХМАО-Югры и, следовательно, некоторое ухудшение экономики региона. Причиной вполне может быть слабо развитая инфраструктура региона, которая может быть решена путем усиленного развития малых городов и поселков ХМАО-Югры.

### Литература

1. Draper, N. Applied regression analysis / N. Draper, H. Smith. – New York : Wiley : In Press, 1981. – 693 p. – Text : direct.
2. Maximum likelihood estimation. – Text : direct // Wikipedia : The free encyclopedia. – URL: [https://en.wikipedia.org/wiki/Maximum\\_likelihood](https://en.wikipedia.org/wiki/Maximum_likelihood) (Date of circulation 15.09.2019).
3. Бокс, Дж. Анализ временных рядов, прогноз и управление / Дж. Бокс, Г. М. Дженкинс. – Москва : Мир, 1974. – 406 с. – Текст : непосредственный.
4. Mazengia, D. H. Forecasting Spot Electricity Market Prices Using Time Series Models : Thesis for the degree of Master of Science in Electric Power Engineering / D. H. Mazengia. – Gothenburg : Chalmers University of Technology, 2008. – 89 p. – Text : direct.
5. Хайкин, С. Нейронные сети : полный курс / С. Хайкин. – Москва : Вильямс, 2006. 1104 с. – ISBN 978-5-8459-0890-2. – Текст : непосредственный.

6. Татьянакин, В. М. Подход к формированию архитектуры нейронной сети для распознавания образов / В. М. Татьянакин. – Текст : непосредственный // Вестник Югорского государственного университета. – Ханты-Мансийск. – 2016. – № 2 (41). – С. 61–64.

7. Головкин, В. А. Нейронные сети: обучение, организация и применение : учебное пособие для вузов / В. А. Головкин ; общая редакция А. И. Галушкина. – Москва : ИПРЖР, 2001. – Кн. 4. – 256 с. – Текст : непосредственный

8. Население Ханты-Мансийского автономного округа – Югры. – Текст : электронный // Википедия : свободная энциклопедия. – URL: [https://ru.wikipedia.org/wiki/%D0%9D%D0%B0%D1%81%D0%B5%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5\\_%D0%A5%D0%B0%D0%BD%D1%82%D1%8B-%D0%9C%D0%B0%D0%BD%D1%81%D0%B8%D0%B9%D1%81%D0%BA%D0%BE%D0%B3%D0%BE\\_%D0%B0%D0%B2%D1%82%D0%BE%D0%BD%D0%BE%D0%BC%D0%BD%D0%BE%D0%B3%D0%BE\\_%D0%BE%D0%BA%D1%80%D1%83%D0%B3%D0%B0\\_%E2%80%94%D0%AE%D0%B3%D1%80%D1%8B](https://ru.wikipedia.org/wiki/%D0%9D%D0%B0%D1%81%D0%B5%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5_%D0%A5%D0%B0%D0%BD%D1%82%D1%8B-%D0%9C%D0%B0%D0%BD%D1%81%D0%B8%D0%B9%D1%81%D0%BA%D0%BE%D0%B3%D0%BE_%D0%B0%D0%B2%D1%82%D0%BE%D0%BD%D0%BE%D0%BC%D0%BD%D0%BE%D0%B3%D0%BE_%D0%BE%D0%BA%D1%80%D1%83%D0%B3%D0%B0_%E2%80%94%D0%AE%D0%B3%D1%80%D1%8B) (дата обращения 22.01.2019).