

4/2022,  
iyul-  
avgust  
(№ 00060)



## ВНЕДРЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СТАТИСТИЧЕСКИЕ И ЭКОНОМИЧЕСКИЕ ИССЛЕДОВАНИЯ

**Родионов Андрей Александрович,**

*Докторант 1 курса Института повышения квалификации кадров и статистических исследований. Ташкент, Узбекистан. [andre-rodionov@mail.ru](mailto:andre-rodionov@mail.ru)*

*DOI: [https://doi.org/10.55439/EIT/vol10\\_iss4/a25](https://doi.org/10.55439/EIT/vol10_iss4/a25)*

### **Аннотация**

Данная статья раскрывает важность применения технологий искусственного интеллекта в статистике. Раскрывается потенциал использования искусственного интеллекта при прогнозировании статистических показателей. Показывается практическое применение данных систем в статистических исследованиях. Также, выделяется роль больших данных в контексте использования и применения их совместно с технологиями, работающими на искусственном интеллекте.

**Ключевые слова:** искусственный интеллект, инженерия искусственного интеллекта, интеллектуальные системы, статистические исследования, прогнозирование, разработка программ, анализ данных, большие данные, нейросети, программирование.

### **Аннотация**

Ушбу мақола статистикада сунъий интеллект технологияларидан фойдаланиш муҳимлигини очиб беради. Статистик кўрсаткичларни башорат қилишда сунъий интеллектдан фойдаланиш имкониятлари аниқланди. Ушбу тизимларнинг статистик тадқиқотларда амалий қўлланилиши кўрсатилган. Шунингдек, катта маълумотларнинг роли улардан фойдаланиш ва сунъий интеллектга асосланган технологиялар билан биргаликда қўлланилиши контекстида таъкидланган.

**Калит сўзлар:** сунъий интеллект, сунъий интеллект муҳандислиги, интеллектуал тизимлар, статистик тадқиқотлар, прогнозлаш, дастурий таъминотни ишлаб чиқиш, маълумотларни таҳлил қилиш, катта маълумотлар, нейрон тармоқлар, дастурлаш.

### **Abstract**

This article reveals the importance of using artificial intelligence technologies in statistics. The potential of using artificial intelligence in predicting statistical indicators is revealed. The practical application of these systems in statistical research is shown. Also, the role of big data is highlighted in the context of their use and application in conjunction with technologies based on artificial intelligence.

**Keywords:** artificial intelligence, artificial intelligence engineering, intelligent systems, statistical research, forecasting, software development, data analysis, big data, neural networks, programming.

## **Введение**

Технологии развиваются все быстрее год за годом. Появилось такое понятие, как искусственный интеллект (ИИ), которое выполняет задачи, изначально закрепленные за человечеством. С целью того, чтобы облегчить учебный процесс, а также в целях получения актуальных знаний, в программы образовательного процесса год за годом вводят изучение все новых языков программирования [1]. Как отметил Ш.М.Мирзиеев, выступая на открытии музыкального фестиваля «Шарк Тароналари» в 2019 году: «В наши дни идет активная полемика о генеративной музыке, искусственном интеллекте. Выдвигаются предположения, что в будущем – искусство, в том числе и музыка будут создаваться без участия человека, только техникой».

Отличие технологий искусственного интеллекта от обычного программирования заключается в том, что при обычном программировании выполняется алгоритмизация задач, а при технологиях искусственного интеллекта - происходит скрытое обучение моделей. Такой метод эффективно подойдет для изучения экономических дисциплин, так как экономическое моделирование и прогнозирование, построенное при помощи искусственного интеллекта - поможет студентам лучше понять принцип работы нейронных сетей и машинного обучения [2].

Также, экономические дисциплины крайне хорошо подходят в качестве объекта для внедрения искусственного интеллекта. То есть, программы и языки программирования, осуществляющие поддержку нейросетей и искусственного интеллекта могут внедряться в рамки образовательного процесса, тем самым помогая изучать студентам экономические дисциплины. В этой статье мы поговорим об инженерии искусственного интеллекта, знаниях, которыми нужно обладать для создания искусственного интеллекта, а также расскажем о программах, с помощью которых могут быть реализованы нейросети в качестве основного инструмента при работе с экономическими и статистическими задачами. Эта структура поможет исследовать тему внедрения изучения искусственного интеллекта в рамках образовательной программы для высших учебных заведений (ВУЗов).

## **Анализ тематической литературы**

На сегодняшний день существует большое количество накопленных знаний в сфере использования искусственного интеллекта в статистическом анализе.

Например, Н.Нильсон [1], в своем труде «Принципы искусственного интеллекта», определил, что роль искусственного интеллекта при проведении статистических исследований различной сложности – по-настоящему велика, и ее трудно переоценить. Кроме этого, похожим образом высказывались о том же такие российские ученые, как А. Винокуров [5], Л. Гагарина [6], Е. Боровская [10], которые также подчеркнули, что «у искусственного интеллекта есть большой потенциал в статистических исследованиях, так как данные технологии могут использоваться в статистике и экономике для того, чтобы производить сложные и комплексные анализы, направленные на достижение высокоточных и высокоэффективных результатов.

Особое место, по мнению И. Михайловой [13], занимает прогнозирование с применением искусственного интеллекта. Она считает, что нейросети и машинное обучение позволяют «достигать высокоточных результатов в прогнозировании, достигая отметку точности в 90-99%». Это говорит о высокой эффективности и актуальности

применения искусственного интеллекта в статистических и экономических исследованиях.

Также, можно отметить таких ученых из стран СНГ, как В. Новикову [14], и А. Исхакову [12], которые рассматривали различные возможности и способы применения искусственного интеллекта в рамках статистического и экономического исследований.

Помимо трудов отечественных авторов (указанные выше авторы представляют регион СНГ), хочется отметить немалый вклад ученых из зарубежных стран. Например, Д. Слэйгл [2], описал системы искусственного интеллекта с точки зрения его эффективного и актуального применения в экономических и статистических системах. В своих трудах, Э. Пап [4], описал системы искусственного интеллекта с практической и теоретической точек зрения, и его работы были высоко оценены западным и мировым научным сообществом. Ученый А. Тей [7] описал логический подход к искусственному интеллекту, а автор С. Рассел [9] представил учебник «Искусственный интеллект: современный подход», в котором отразил значимую роль ИИ в статистике и экономике.

Крайне важным и полезным литературным трудом является труд ученого М. Джонса [15], который описал, как можно применять системы искусственного интеллекта в различных статистических и экономических приложениях.

Хочется отметить, что существует огромное количество трудов, в которых содержится полезная информация об искусственном интеллекте и его применении в статистике. В данной работе представлены наиболее интересные и современные литературные труды.

### **Методология исследования**

К используемым методам исследований можно отнести – метод наблюдения, который представлен в данной работе. В последние десятилетия для прогнозирования широко используются нейронные сети [3]. Рассмотрим особенности применения нейронных сетей, которые показывают их преимущества по сравнению с другими существующими методами при выборе модели прогноза.

1. Результативность при решении неформализованных или плохо формализованных задач. Из общеизвестных преимуществ методов на основе нейронных сетей следует выделить одно самое привлекательное — отсутствие необходимости в строгой математической спецификации модели, что особенно ценно при прогнозировании плохо формализуемых процессов. Известно, что большинство финансовых, бизнес и других подобных задач плохо формализуется [4].

2. Устойчивость к частым изменениям среды. Достоинства нейронных сетей становятся заметными, когда часто изменяются «правила игры»: среда, в которой существует прогнозируемый процесс, а также характер воздействия влияющих факторов [5].

3. Результативность при работе с большим объемом противоречивой информации. Нейронные сети будут предпочтительнее там, где имеется большое количество анализируемых данных, в которых скрыты закономерности. В этом случае автоматически учитываются также различные нелинейные взаимодействия между влияющими факторами [6].

4. Результативность при работе с неполной информацией. Целесообразно использование нейронных сетей в задачах с неполной или «зашумленной» информацией, а также в задачах, для которых характерны интуитивные решения.

Отбор значащих факторов. На первом этапе выделяется максимальное число из значащих, влияющих на прогноз факторов. Такие дополнительные факторы, влияющие на поведение прогнозируемой величины, называют экзогенными (внешними) или артефактами [7].

Здесь же выбирается интервал наблюдения, т. е. выясняется, по какому количеству предшествующих значений временного ряда осуществляется прогноз.

Предобработка данных. На втором этапе устраняются несущественные, по мнению эксперта, и не влияющие на прогноз данные [8]. При необходимости также восстанавливается пропущенная информация, устраняются аномальные выбросы, убираются высокочастотные шумы. Умело проведенная предобработка данных позволяет значительно улучшить качество прогноза.

Построение модели. На следующем этапе для данного анализируемого процесса выбираются наиболее подходящая парадигма и структура нейронной сети, а также алгоритм и параметры ее обучения [9].

Прогнозирование (получение результата). Эксперименты осуществляются по схеме, аналогичной той, при которой производилось обучение. При этом необходимо иметь в виду, что существуют определенные ограничения и недостатки, связанные с использованием нейронных сетей для прогнозирования. Во-первых, для эффективного прогнозирования, как правило, требуется некоторый минимум наблюдений. Однако существует много задач, когда такое количество статистических данных недоступно [10].

Следует отметить, что удовлетворительная модель прогноза с использованием нейронной сети все же может быть построена даже в условиях нехватки данных. При этом модель будет уточняться при поступлении в нее свежих данных. Во-вторых, нельзя забывать о значительных временных затратах для достижения удовлетворительного результата. Эта проблема не столь существенна, если исследуется небольшое число временных последовательностей, однако обычно прогнозирующая система в области управления производством включает от нескольких сотен до нескольких тысяч временных последовательностей [11].

В-третьих, обучить и эксплуатировать нейронную сеть для решения многих задач, как правило, может и не специалист, но надежно интерпретировать результаты, а также численно оценивать значимость получаемых прогнозов способны только специалисты, имеющие навыки в моделировании нейронных сетей [12].

### **Анализ и результаты**

Рассмотрим простой пример, когда потребуется прогнозирование при помощи нейросетей. Например, есть данные по доходам и расходам школ США по годам. Данные не структурированы по названиям штатов, доходу и расходу (рис.1). Требуется создать экспертную модель, и продумать, какое прогнозирование применить. Допускается прогнозирование по задаваемым параметрам [13]. Например, человек будет вводить прогнозный год, название штата, и расход школ в этот год. Нейросеть будет давать значение прибыли. Прибыль необходимо получить самостоятельно в данном случае. То есть от расхода требуется отнять доход.

STATE	YEAR	ENROLL	TOTAL_REVENUE	FEDERAL_REVENUE	STATE_REVENUE	LOCAL_REVENUE	TOTAL_EXPENDIT		
Alabama	1992	2678885	304177	1659028	715680	2653798	1481703	735036	174053
Alaska	1992	1049591	106780	720711	222100	972488	498362	350902	37451
Arizona	1992	3258079	297888	1369815	1590376	3401580	1435908	1007732	609114
Arkansas	1992	1711959	178571	958785	574603	1743022	964323	483488	145212
California	1992	26260025	2072470	16546514	7641041	27138832	14358922	8520926	2044688
Colorado	1992	3185173	163253	1307986	1713934	3264826	1642466	1035970	364760
Connecticut	1992	3834302	143542	1342539	2348221	3721338	2148041	1142600	48542
Delaware	1992	645233	45945	420942	178346	638784	372722	194915	30595

**Рис.1. Пример данных из задачи**

Использование табличного редактора Microsoft Excel как инструмента при создании прототипа простой Экспертной Системы. О системе PROLOG. Ознакомившись с задачей, следует заняться созданием Экспертной Системы. Существует система PROLOG - инструмент логического программирования. На нем можно создать экспертную модель. Однако, для большего понимания, хочется показать прототип экспертной системы на Microsoft Excel. Однако, PROLOG настолько интересен, что нельзя не показать, как бы строилась ЭС на нем. К примеру, создадим небольшую базу знаний на PROLOG. Сначала следует создать текстовый файл. В нем следует указать параметры, например, за 3 года - для штата Алабама, указывая при этом только прибыль (рис.2).

```
alabama(1992, 2678885). % Данные дохода по штату Алабама за 1992 год  
alabama(1993, 2827391).% Данные дохода по штату Алабама за 1993 год  
alabama(1994, 2970223). % Данные дохода по штату Алабама за 1994 год
```

**Рис.2. Данные в текстовом документе.**

После чего, требуется зайти в PROLOG, и экспортировать данный файл, выбрав пункт «Consult» в разделе «File». После экспорта появится окошко с вводом данных. Именно сюда и следует вводить запросы (рис 3).



**Рис.3 Поле ввода в PROLOG**

Затем, следует сделать какой-либо запрос, который интересует пользователя. Например, мы можем запросить все данные по штату Алабама (в нашем случае, это будет год и статья дохода).

```
?- alabama(X,Y).|
```

**Рис.4. Запрос в PROLOG**

Как видно на рисунке 4, мы сделали запрос по штату Алабама. Указали название штата - «alabama», а также ввели 2 параметра. Синтаксис системы PROLOG подразумевает написание параметров с большой буквы. То есть в данном случае это X и Y.

```
?- alabama(X,Y).  
X = 1992,  
Y = 2678885 ;  
X = 1993,  
Y = 2827391 ;  
X = 1994,  
Y = 2970223.
```

**Рис. 5. Получение данных из БЗ**

Как показано на рисунке 5, мы получили данные по 3 годам, где X- год, а Y- прибыль школ из штата Алабама. Причем данные с 1992 по 1994 год (из данных задачи). При этом, стоит отметить, что для встраивания всех данных с задачи, данный способ будет не рационален и труден. Легче показать принцип ЭС на базе программы Microsoft Excel. При этом, будут учтены требования ЭС, то есть будут выводиться структурированные данные после обработки их значений. В новых версиях программы Microsoft Office Excel, присутствует функция ЕСЛИМН, которая является частью, отвечающей за составления программы для задачи. То есть при ее помощи можно задать условия, и получить в зависимости от них, результат.

Начнем с того, что что мы имеем следующие данные:

- Название штата;
- Год;
- Общий доход;
- Федеральный, городской и местный доход (что в сумме дают общий доход);
- Общий расход;
- Федеральный, городской и местный расход (что в сумме дают общий расход)

Вполне можно убрать данные о федеральных и пр. бюджетах, оставляя только общие доходы и расходы. Год останется также. Однако, плохо определяем будет пункт - «Название штата». Для него и применим функцию ЕСЛИМН, которая поможет распределить штаты по нумерации. Например, зададим, что вместо штата Алабамы, будем получать значение «1», для Аляски - «2», и так далее. Вначале, сделаем сортировку по алфавиту в столбике названий штата. Помимо этого, следует добавить ячейку прибыли, которая будет получена в результате вычитания Расхода от Дохода (рис. 6).

Теперь, помимо дохода и расхода, имеем показатель прибыли. Который сможет сказать о том, что если показатель больше 0, то это профицит, а если меньше, то дефицит. Затем, перейдем непосредственно к созданию прототипа ЭС. В новом поле (после «Прибыли»), укажем следующую формулу: =ЕСЛИМН(A2="Alabama";1;A2="Arizona";3;A2="Alaska";2). Она говорит о том, что если в значении названия штата есть Алабама, то новое поле станет равно 1, для Аляски - 2, и так далее. Для примера и дальнейшей работы, остановимся на 3 штатах.

24	Alabama	2014	7357267	7587737	-230470	1
25	Alabama	2015	7360222	7501799	-141577	1
26	Alabama	2016	7498567	7708845	-210278	1
27	Alaska	1992	1049591	972488	77103	2
28	Alaska	1993	1191398	1126398	65000	2
29	Alaska	1994	1052282	1116160	-63878	2
30	Alaska	1995	1117068	1203817	-86749	2
31	Alaska	1996	1120346	1240031	-119685	2
32	Alaska	1997	1131950	1237103	-105153	2
33	Alaska	1998	1206195	1222893	-16698	2

**Рис.6. Данные из задачи про доход школ.**

Как видно на рисунке 6, появился новый столбец, со значениями от 1 до 3, отвечающий за номер штата. Теперь, появилась возможность для формулировки задачи. У нас есть данные по годам - расход, доход, и само значение года. Также, есть номер штата (от 1 до 3). Далее, потребуется использовать программу с встроенной функцией «нейросетей» для составления прогноза за указанный пользователем год. Воспользуемся программой STATISTICA. Она содержит нужный компонент, то есть функцию прогнозирования на основе искусственного интеллекта - «нейросетей». Реализация задачи прогнозирования с использованием Нейросетей (при построении модели прогнозирования) при помощи программы STATISTICA. Вначале, следует открыть программу, и экспортировать таблицу. Для этого заходим в пункт «Файл», и нажимаем «Открыть». Выбираем нужный файл, и экспортируем его. Далее, следует зайти в пункт - «Добыча данных», и нажать на кнопку «Нейронные сети» (рис. 7).

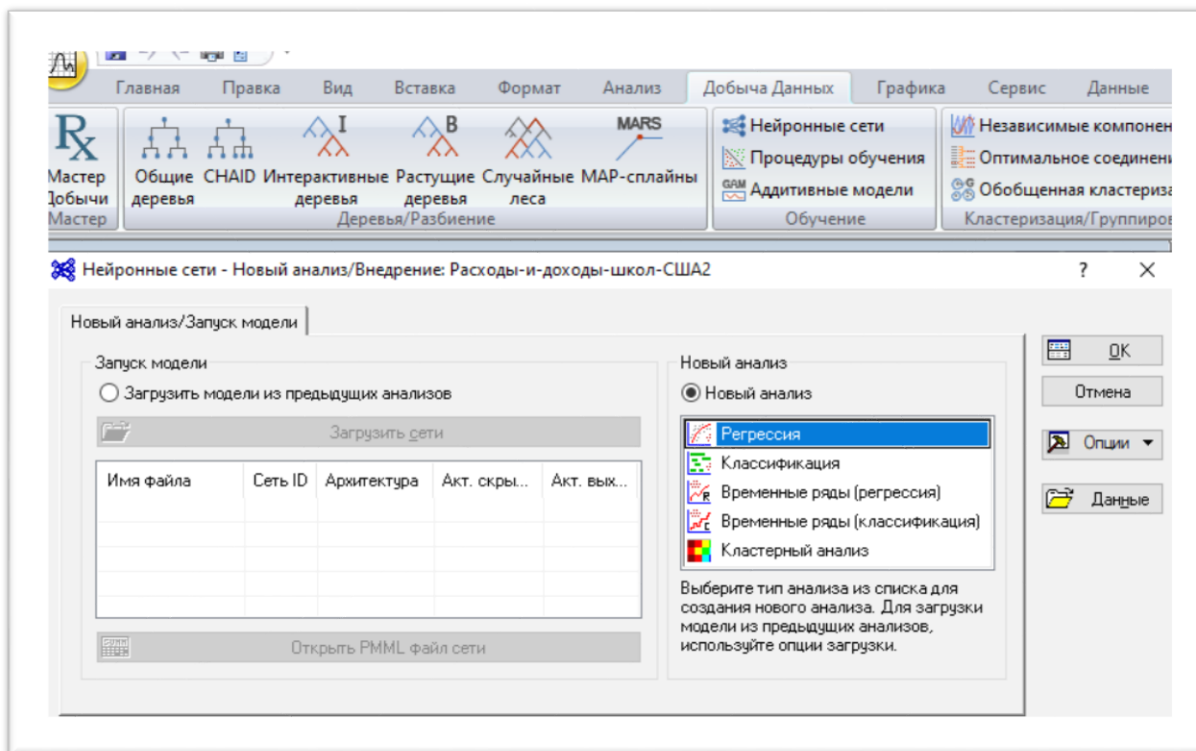


Рис.7. Нейронные сети в программе STATISTICA.

После чего, выбираем пункт регрессия, и выбираем переменные. Далее, нужно выбрать переменные, и указать их в нужных ячейках. То есть, мы укажем год, расход и номер штата, а программа выдаст нам ожидаемое значение прибыли. Сейчас нужно нажать «ОК», а после «Обучить». Можем посмотреть, с какой точностью программа будет выполнять прогноз. Нажимаем на «Итоги моделей», где получим 5 поколений обученной модели.

Итоги моделей (Расходы-и-доходы-школ-США2)							
N	Архитектура	Производительность обуч.	Контр. производительность	Тест. производительность	Ошибка обучения	Контрольная ошибка	Тестовые ошибки
1	MLP 3-4-1	0,883846	0,941414	0,356417	5,800700E+09	2,086286E+09	1,278589
2	MLP 3-4-1	0,759716	0,708359	0,227730	1,131640E+10	7,779776E+09	2,096731
3	MLP 3-3-1	0,843089	0,985493	0,388144	7,655419E+09	7,298262E+08	9,181541
4	MLP 3-9-1	0,892187	0,893915	0,350882	5,426283E+09	2,931479E+09	1,469525
5	MLP 3-8-1	0,893171	0,888719	0,507205	5,351160E+09	3,060876E+09	1,816342

Рис.8. Точность прогноза.

Как показывает рисунок 8 - в 5 поколении точность прогноза составляет около 90%.

Результаты прогноза

Прибыль	Год	Перб	Расход
-449834	2022	1	200000
-324049	2023	2	300000
563588	2023	3	100000
-336576	2023	2	10000

Например, в таблице 1, вводим значения года, расхода, и номера штата (1-3). Тут, планируется профицит для штата 3 при значении расхода в 100 тыс. - 563.5 тысячи, и дефицит для штата 2 при значении расхода в 300 тыс - (минус 324 тысячи). Эти данные получены на основе данных с задачи.

Таким образом, мы решили задачу, сделав возможным создавать прогнозные значения, создали прототип ЭС на Excel, ознакомились с системой PROLOG. Также, создали модель с регрессионным анализом на основе нейронных сетей в программе STATISTICA.

**Заключение**

Одна из самых распространенных ошибок в статистике — непонимание роли искусственного интеллекта. Термины ИИ и машинное обучение используются так, как если бы они были синонимами, но эти два понятия имеют очень разные значения [14]. Искусственный интеллект включает в себя создание сложных систем, которые используют алгоритмы для получения выводов, в то время как машинное обучение — это когда компьютеры получают информацию и учатся действовать без участия человека [15]. В этой статье были представлены основы искусственного интеллекта и машинного обучения, описаны их соответствующие приложения и представлены некоторые преимущества, которые они предоставляют. Разработчики создают программное обеспечение, которое может учиться на собственном опыте, а не следовать заранее определенному пути. Оно способно анализировать огромные объемы данных и распознавать закономерности способами, которые были бы невозможны для людей. Системы искусственного интеллекта используются всеми видами бизнеса, от крупных игроков, таких как Google и Яндекс, до небольших, таких как косметическая компания, которая находит уникальные оттенки для своей основы. Также, в статье было показано, как можно применять технологии искусственного интеллекта. Выделена роль искусственного интеллекта в следующих аспектах: например, в улучшении ведения бизнеса. Один из способов сделать это — использовать чат-боты с искусственным интеллектом на своем веб-сайте. Чат-бот будет давать ответы клиентам, которые ищут информацию о продуктах и услугах. Когда клиенты связываются с компаниями, они, как правило, спешат и не имеют времени ждать и слушать автоматическое сообщение о том, что агент скоро будет с ними. Поэтому, если они получают немедленный ответ, они могут быть удовлетворены тем, что их проблема была решена быстро, вместо того чтобы чувствовать разочарование, когда им приходилось ждать. Компании всегда ищут способы

сохранить лояльность клиентов, поэтому они готовы тратить на них больше ресурсов. Системы искусственного интеллекта могут предоставлять информацию о клиентах и предлагать целевые сделки на основе этой информации. Например, банк может обратиться к каждому клиенту с предложением, специально адаптированным к его потребностям. Это увеличивает шансы клиентов покупать в этом банке, поскольку очевидно, что они заботятся о них, а не просто берут их деньги, ничего не предлагая взамен.

#### **Список литературы**

1. Нильсон, Н. Принципы искусственного интеллекта / Н. Нильсон. - М.: Радио и связь, 2019. – с. 27.
2. Слэйгл, Дж. Искусственный интеллект / Дж. Слэйгл. - М.: Мир, 2018. - с. 44.
3. Люгер, Дж.О. Искусственный интеллект: стратегии и методы решения сложных проблем / Дж.О. Люгер. - М.: Диалектика, 2016. – с. 101.
4. Э.Пап. «Искусственный интеллект: теория и приложения». - Издание «Artificial Intelligence S», 2021 год. – с. 133.
5. Винокуров А.Ю. Информационные технологии: учебно-методический комплекс. - Ульяновск: Ульян, гос. техн, ун-т, 2017. – с. 144.
6. Гагарина Л.Г. Автоматизированные информационные системы: учеб, пособие. - М.: МИЭТ, 2018. – с. 97.
7. Тей, А. Логический подход к искусственному интеллекту / А. Тей, П. Грибомон, и др. - М.: Мир, 2018. – с. 54-55.
8. К.Аггарвал. «Искусственный интеллект: учебник». - Издание - «Charu», 2021 год. – с. 211.
9. Рассел, С. Искусственный интеллект: современный подход / С. Рассел, П. Норвиг. - М.: Вильямс, 2018. – с. 107.
10. Е.В. Боровская, Н.А.Давыдова. - М.Лаборатория знаний, 2020. – с. 77.
11. Дмитриева М. А. Влияние искусственного интеллекта на сферу маркетинга / М. А. Дмитриева. - Интернет-маркетинг, 2018. – с. 82-88.
12. Исхакова А. Ф. Применение искусственного интеллекта / А. Ф. Исхакова. - Вестник современных исследований, 2018.- с. 261-262.
13. Михайлова И.С. Перспективы использования искусственного интеллекта в сфере образования. - Современные научные исследования и разработки, 2018. - с. 475-476.
14. Новикова В. С. К вопросу о специфике развития искусственного интеллекта. - Дневник науки, 2018. - с. 8.
15. Джонс М. Т. Программирование искусственного интеллекта в приложениях. - М.: ДМК Пресс, 2018. – с. 311.