

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Комсомольский-на-Амуре государственный университет»

На правах рукописи

Попова Ольга Владимировна

**Интеллектуальная система управления процессами поддержания
параметров технических систем**

Направление 09.04.03
«Прикладная информатика»

**АВТОРЕФЕРАТ
МАГИСТЕРСКОЙ ДИССЕРТАЦИИ**

2022

Работа выполнена в ФГБОУ ВО «Комсомольский-на-Амуре государственный университет»

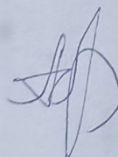
Научный руководитель: Григорьев Ян Юрьевич,
кафедра прикладной математики,
кандидат физико-математических наук,
доцент, КнАГУ, проректор по УР.

Рецензент: Анисимов Антон Николаевич,
кафедра информационной
безопасности, информационных
систем и физики, доцент,
кандидат физико-математических наук,
АМПУ.

Защита состоится 23 июня 2022 года в 9:50 на заседании государственной экзаменационной комиссии по направлению 09.04.02 «Прикладная информатика» в Комсомольском на Амуре государственном университете по адресу: 681000, г. Комсомольске на амуре, пр. Ленина, 27, аудитории 204/5.

Автореферат разослан 15.06.2022 г.

Секретарь ГЭК



Е. В. Абрамсон

ОБЩАЯ ХАРАКТЕРИСТИКА ДИСЕРТАЦИОННОЙ РАБОТЫ

Актуальность темы заключается в формировании подхода, основанного на интеллектуальных технологиях, обеспечивающего повышение эффективности технологического процесса на основе текущей информации о нем. Оборудование, используемое в современном производстве, позволяет получать информацию о технологическом процессе, контролировать значения производственных параметров в реальном времени. Формируемый набор имеет большую ценность с точки зрения анализа и выбора стратегии по управлению производством. Системы автоматического контроля параметров системы производства могут способствовать повышению эффективности производства. Особый успех можно достигнуть в таких технологиях, как переработка природного сырья в извлекаемый из него готовый продукт. Это обусловлено постоянно изменяемыми исходными характеристиками поступающих ресурсов, а также отклонениями в работе промышленного оборудования, человеческими факторами. Своевременное изменение характеристик технологического процесса позволит повысить в режиме реального времени эффективность переработки сырья и повысить прибыль предприятия.

Целью данной работы заключается разработка интеллектуальной системы управления процессами поддержания параметров технических систем.

Задачи:

- Рассмотреть процесс автоклавного окисления;
- Исследовать применимость «классических» методов обработки данных и методов машинного обучения к набору данных, сформированному на основе информации о рассматриваемом процессе;
- Разработать автоматизированную систему;
- Провести оценку полученных результатов.

Объектом исследования является производственная система химической и механической обработки сырья на предприятии.

Предметом исследования является инструментарий методов машинного обучения, включающего в себя применение нейронных сетей и метода XGBoost.

Методы:

- анализ;
- моделирование;
- аналитическое исследование;
- методы машинного обучения;
- методы регрессионного моделирования.

Научная новизна исследования:

- На основе непрерывного анализа данных о действующем производстве формируется управляющая интеллектуальная системы, определяющая оптимальные параметры процесса.

- Разрабатывается математическая модель с применением подходов машинного обучения для решения исследуемых производственных задач.

Достоверность и обоснованность результатов исследования.

Основные положения и выводы работы обоснованы, аргументированы применением аппарата математического моделирования интеллектуальных систем и методов машинного обучения, результаты оценены экспериментально на актуализированном наборе данных.

Практическая значимость заключается в направленности использования разработанной системы в реальных условиях производства автоклавного окисления с целью повышения ее эффективности.

Личный вклад автора заключается в выполнении основного объема теоретических и практических исследований, изложенных в диссертационной работе, включая разработку моделей, проведение анализа и оформления результата в виде публикаций.

Основные положения работы, выносимые на защиту:

- Рассмотрение процесса автоклавного окисления.
- Модели «классических» методов обработки данных и методов машинного обучения к набору данных, сформированному на основе информации о рассматриваемом процессе.
- Программная реализация.
- Проведение оценки полученных результатов.

Апробация результатов.

Конференции:

- Актуальные проблемы информационно-телекоммуникационных технологий и математического моделирования в современной науке и промышленности, Комсомольск-на-Амуре, 20–25 марта 2021 г.

- Наука, инновации и технологии: от идей к внедрению, Комсомольск-на-Амуре, 07–11 февраля 2022 г.

- 2020 international multi-conference on industrial engineering and modern technologies, fareastcon 2020, Vladivostok, 06–09 октября 2020 г.

Публикации:

1 Применение методов машинного обучения в задачах оценки технологических процессов. Попова О.В., Григорьев Я.Ю., Жарикова Е.П., Григорьева А.Л. Ученые записки Комсомольского-на-Амуре государственного технического университета. 2021. № 7 (55). С. 68-72.

2 Оценка технологических процессов методами машинного обучения. Попова О.В., Григорьев Я.Ю., Жарикова Е.П. В сборнике: Наука, Инновации и технологии: от идей к внедрению. Материалы Международной научно-практической конференции. Комсомольск-на-Амуре, 2022. С. 240-244.

3 Моделирование оценки параметров в технологических задачах. Попова О.В., Кириллов А.В. В сборнике: Наука, инновации и технологии от идей к внедрению.. Материалы Международной научно-практической конференции. Комсомольск-на-Амуре, 2022. С. 244-246.

4 Модели умного освещения. Попова О.В., Григорьев Я.Ю., Жарикова Е.П. В сборнике: Актуальные проблемы информационно-телекоммуникационных технологий и математического моделирования в современной науке и промышленности. Материалы I Международной научно-практической конференции молодых учёных. Комсомольск-на-Амуре, 2021. С. 98-100.

5 Моделирование системы контроля освещённости. Попова О.В., Григорьев Я.Ю., Жарикова Е.П., Григорьева А.Л. Ученые записки Комсомольского-на-Амуре государственного технического университета. 2021. № 1 (49). С. 10-18.

6 Lighting control system modeling. Popova O.V., Grigoryev J.U., Zharikova E.P. В сборнике: 2020 International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon 2020. 2020. С. 9271577.

Структура и объем.

Магистерская диссертация состоит из введения, четырех глав, заключения, списка использованных источников. Объем работы – 68 страниц, в том числе 39 рисунков, 4 таблиц и 1 приложения.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении раскрывается актуальность темы, определяются цели и задачи исследования.

В первой главе приводится описание исследуемых производственных процессов. Производится описание автоклавного окисления (РОХ) и цианирования, обеспечивающих извлечение золота из упорных золотых концентратов с месторождений Албазино и Майское. Блок-схема работы производства представлена на рисунке 1.

Блок-схема Амурского ГМК

(Автоклавное окисление/цианирование)



Рисунок 1 - Блок-схема работы производства

Цианирование – способ извлечения, выщелачивания, металлов, главным образом золота и серебра, из сравнительно бедных, тонковкрапленных руд, хвостов и других продуктов обогащения, основанный на селективном растворении металлов в слабых растворах цианидов и последующем их осаждении. Процесс автоклавного окисления (РОХ) – это процесс, применяемый для извлечения золота, меди, цинка, молибдена и урана из упорной руды. В технологии РОХ применяются повышенные температуры, высокое давление, и кислород для освобождения минералов из упорных руд.

Описаны главные достоинства использования технологии РОХ: быстрое отделение золота, меди и других минералов от упорных руд, достижение высоких характеристик показателей извлечений и снижение воздействия на окружающую среду.

Во второй главе описывается модель аналитической системы управления параметрами, позволяющими оптимизировать технологический

процесс. Моделирование производится на основе информации о содержании элементов и свойствах вещества, полученной с датчиков, контролирующих работу автоклавной установки. В основу закладываются методы машинного обучения и классические методы. К «классическим» методам относится регрессионное моделирование, производимое с помощью разных видов зависимостей.

Рассматривается предполагаемая схема работы системы. При работающей автоклавной установке, прогоняющей вещество через себя, через датчики поступают данные о его содержимом и других используемых для его переработки компонентах. Полученная информация позволит с помощью разрабатываемого алгоритма оценить значение выходной характеристики. Значимость выходного значения определяется оператором, который может корректировать характеристики работающей системы с целью изменения результирующей характеристики.

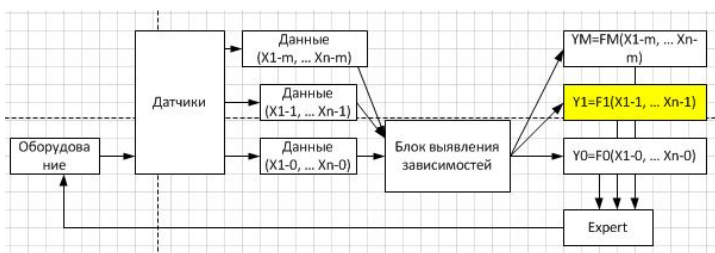


Рисунок 2 – Схема работы системы при «классическом» методе

Где, к примеру, F_0 - это линейная функция, F_1 - кубическая и т.д., x_1 - железо, x_2 – мышьяк и т.д. Y_0 – искомое значение параметра на основе данных x_1-0 и т.д.

Описывается анализ данных и вывод из выявленных функций, к примеру:

$$y = a + bx^{0.5} + cx + dx^{1.5} + ex^2 + fx^{2.5} + gx^3 + hx^{3.5} + ix^4 + jx^{4.5} + kx^5 \quad (1)$$

($a = 0,1367$, $b = 18409404$, $c = -4.5384$, $d = 4,42326$, $e = -1.91109$, $f = 5,146586$, $g = 3,3305$, $h = -1,631326$, $i = 3.76998$, $j = -4.4827$, $k = 2.20989$), c коэффициентом детерминации равным 0,3. Так же показан рисунок 3 графика функции.

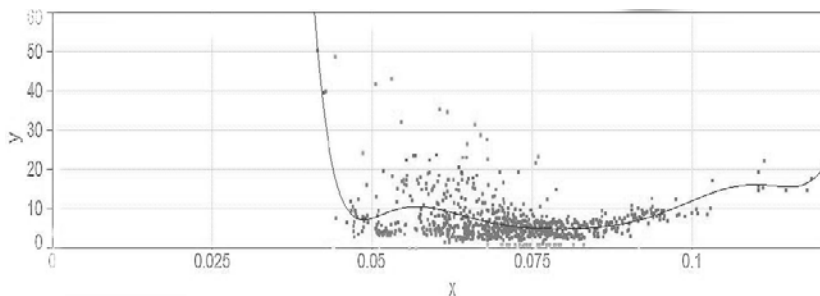


Рисунок 3 – График функции 1

$$y = a + bx^{0.5} + cx + dx^{1.5} + ex^2 + fx^{2.5} + gx^3 + hx^{3.5} + ix^4 \quad (2)$$

$$(a = 1,27, b = -3,83, c = 5,02, d = -3,75, e = 1,74, f = -5,16, g = 9,52, h = -$$

9,99, i = 4,56), с коэффициентом детерминации равным 0,38.

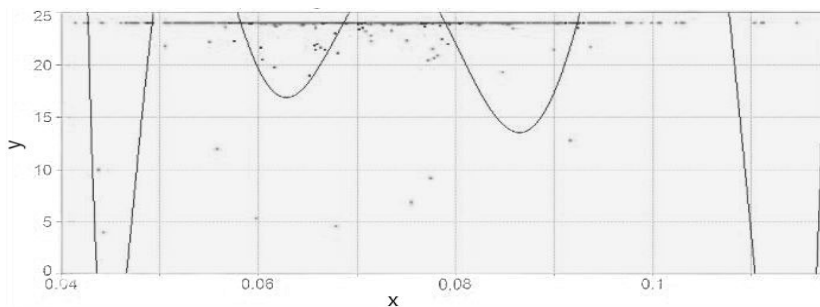


Рисунок 4 – График функции 2

Производится анализ большого количества достаточно сложных зависимостей, среди которых отсутствуют функции, обеспечивающие достаточно тесные связи. Коэффициенты детерминации некоторых представленных уравнений указывают на умеренную связь, но надежность результатов полученных с помощью этих зависимостей низкая. Приведенные функции имеют большие по сравнению с аналогами значения коэффициента детерминации, но не обеспечивают надежных связей. Аналитических зависимостей имеющих высокую или очень высокую связь по шкале Чеддока нет. Анализ множественных связей не позволяет найти аналитическую зависимость, обеспечивающую требуемую точность. Требуется другие

подходы к выявлению зависимостей, предлагается рассмотреть методы машинного обучения для решения поставленной задачи.

В третьей главе описываются методы машинного обучения, к которым относятся: нейронные сети, деревья решений, «случайный лес», кластеризация. Предлагаемые методы позволяют получить надежное решение в заданном диапазоне.

Приводится схема работы системы при реализации метода машинного обучения. Поступающие данные о технологическом процессе сохраняются в файл, который является исходным для реализации модели (например, нейронной сети). На выходе системы формируется информации о прогнозном значении результирующего параметра. Оценка полученных данных производится оператором.

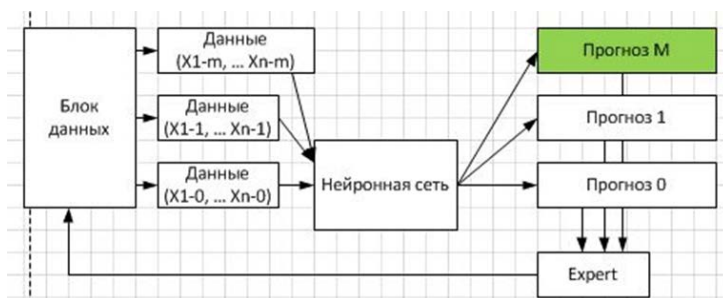


Рисунок 5 – Схема работы системы при машинном обучении

Показан рисунок работы схемы системы при выбранном методе.

Таблица обозначений данных представлена ниже.

Таблица 1 – Обозначения данных схемы работы при машинном обучении.

Обозначение данных	Название
x_0^1	Железо
x_1^1	Мышьяк
x_2^1	Хлориды
x_m^n

В качестве основных моделей для оценки решений поставленной задачи в работе рассматриваются «Нейронные сети» и «Деревья решений».

Описывается алгоритм XGBoost, который относится к деревьям решений с градиентным усилением, разработанным для увеличения скорости и производительности системы. XGBoost представляет собой ансамбль из деревьев, в котором обучается множество копий более слабой модели, не позволяющей достичь высокой точности на обучающей выборке данных. На каждом шаге слабая модель обеспечивает корректировку ошибок, допущенных предыдущими моделями. Предсказания всех моделей суммируются с определяемыми весами. Для обучения модели реализуется 1000 итераций с шагом 0.1 секунды и глубиной узлов деревьев 4. На рисунке 6 приводится схема алгоритма XGBoost.

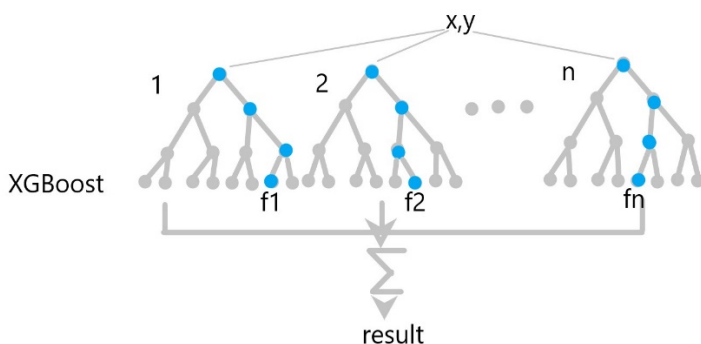


Рисунок 6 - Схема алгоритма XGBoost

Описывается архитектура нейронной сети. Прогнозы делаются на выходном или видимом слое, содержащем один параметр. На входном слое задаются несколько характеристик. Для обучения модели необходимо 100 эпох, со скоростью 0.01 секунды, с 7 скрытыми слоями. Эпоха – количество итераций в процессе обучения. На рисунке 7 изображается схема нейронной сети.

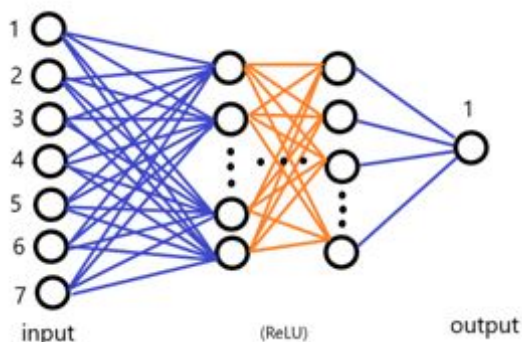


Рисунок 7– Архитектура нейронной сети MLP

В четвертой главе описывается построение модели и ее реализация для исследуемой задачи. На рисунке 8 представлена работа нейронной сети.

```

main x
Epoch 1198: | Train Loss: 0.00451 | Val Loss: 0.00868
torch.Size([182, 15])
torch.Size([94, 15])
Epoch 1199: | Train Loss: 0.00451 | Val Loss: 0.00868
torch.Size([182, 15])
torch.Size([94, 15])
Epoch 1200: | Train Loss: 0.00451 | Val Loss: 0.00868
Check Point

```

Рисунок 8 - Результат работы нейронной сети

Производится оценка характеристик нейронной сети с помощью функции потерь. Функция потерь представляет собой разность между фактическим значением и прогнозируемым. Для задач прогнозирования используется характеристика MSE (Mean Squared Error) равная средней арифметической квадрата разности между реальным и прогнозным значением:

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2, \text{ где } e_t = original_t - predict_t, n - \text{количество измерений.}$$

Нейронная сеть показывает высокие результаты, которых недостаточно для оценки адекватности модели. Приводятся метрики для метода

деревьев решений MSE и MAE. MAE (Mean Absolute Error) определяется как среднее арифметическое абсолютной разности между реальными и прогнозными значениями:

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t|, \text{ где } e_t = original_t - predict_t.$$

Результаты метрики обученной модели занесены в таблицу 2.

Остальные метрики рассчитаны для дополнительного сравнения.

Таблица 2 - Результаты метрик обученной модели XGBoost

Метрика	Значение
MSE	0.015
MAE	0.015
r2	0.92
EVS	0.92
ME	0
Median AE	0
MAPE	0

Полученные значения указывают на высокие показатели оценки модели.

На рисунке 9 представлена матрица корреляции для параметров модели, значения которой рассчитываются по критерию Пирсона.

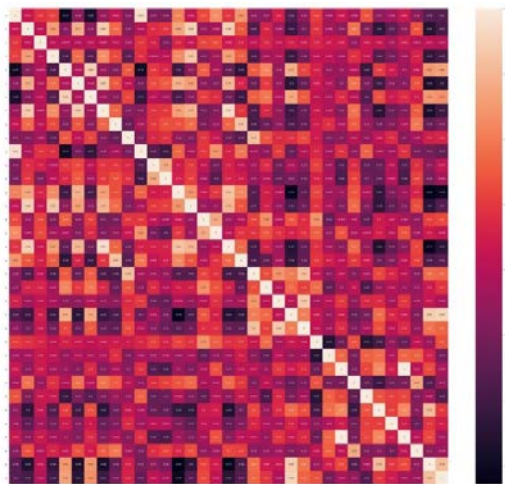


Рисунок 9 – Матрица корреляции

Цветом показываються значения в диапазоне от 0 до 1, чем ближе к 1, тем ячейка светлее, чем ниже, тем ячейка темнее. Чем ближе значение к 0, тем меньше вероятность связи. Метод позволяет оценить статистическую значимость различий двух или нескольких относительных показателей. Значение равное единице показывает абсолютную функциональную связь, поэтому в матрице игнорируется значения по диагонали.

На рисунке 10 показывается график Шепли, позволяющий оценить адекватность модели.

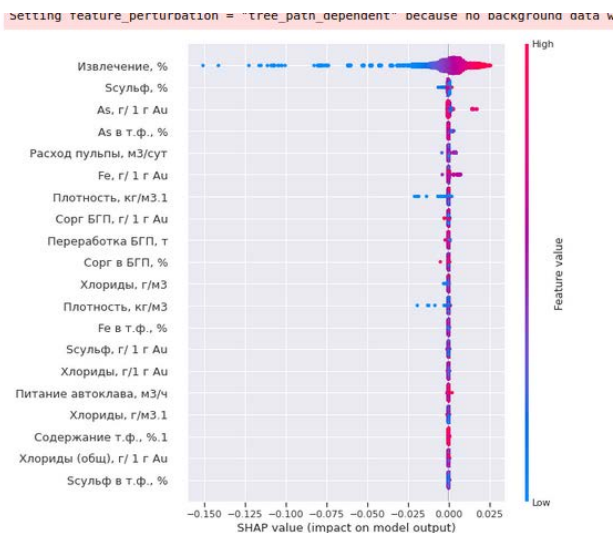


Рисунок 10 – График Шепли

По вертикали расположены параметры, по горизонтали вероятности влияния параметров на результирующий признак. Значения слева от центральной вертикальной линии — это отрицательный класс - 0, справа — положительный - 1. Толщина линии прямо пропорциональна количеству точек наблюдения. Чем выше параметр, тем он имеет большее влияние на систему. Чем краснее точки, тем большее значение имеет признак в этой точке.

Полученные от предприятия данные не позволяют обеспечить нормальное обучение модели, посредством интерполяции набор данных расширяется в 10 раз.

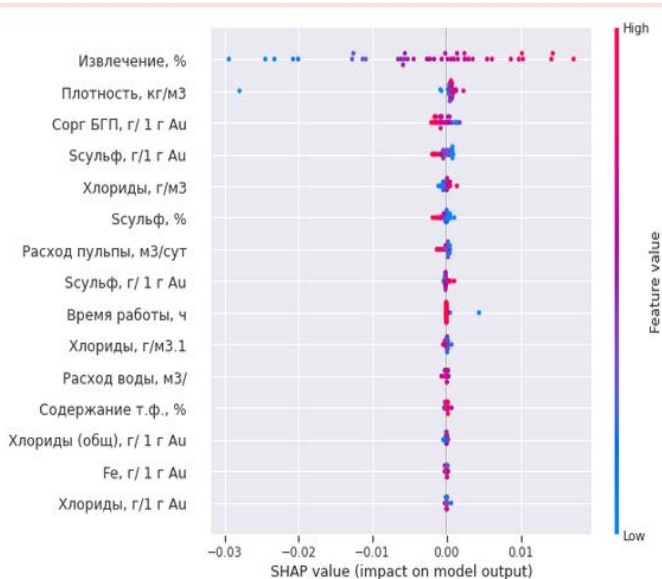


Рисунок 11 – График значений увеличенных данных

Увеличение данных, как и другие способы изменения не дают видимых результатов, что говорит о необходимости формирования новых требований к сбору и подготовке данных предприятием.

В заключении представляются выводы исследования о несостоятельности классических методов для описания рассматриваемой технологии. Указываются преимущества машинных методов, обеспечивающих на исследуемом диапазоне данных более точные результаты. Обозначается актуальность проблемы увеличения данных и их интерпретации, что требует формирование нового набора данных для повышения качества получаемых результатов.

Список использованных источников

- 1 Жарикова, Е. П. Методы контроля нештатных ситуаций в процессах транспортировки и переработки нефтепродуктов / Е. П. Жарикова, Р. А. Лосев, Я.Ю. Григорьев // В сборнике: актуальные проблемы информационно-телекоммуникационных технологий и математического моделирования в современной науке и промышленности. Материалы I Международной научно-практической конференции молодых учёных. Комсомольск-на-Амуре. 2021. С. 177 – 180.
- 2 Жарикова, Е. П. Применение искусственного интеллекта в задачах анализа состояния акваторий / Е. П. Жарикова, Я. Ю. Григорьев, А. Л. Григорьева // Морские интеллектуальные технологии. 2021. Т. 2. № 2 (52). С. 129 – 133.
- 3 Челухин, В. А. Применение математических моделей в информационных технологиях / В. А. Челухин, В. А. Абрамсон // Материалы I Международной научно-практической конференции молодых учёных. Комсомольск-на-Амуре: ФГБОУ ВО «КнАГУ». 2021. С. 117 – 120.
- 4 Амосов, О. С. Оценивание случайных последовательностей с использованием нечётких систем и кластеризации / О. С. Амосов, Д. С. Магола, Е. А. Малашевская // Информатика и системы управления. 2012. № 1 (31). С. 146 – 155.
- 5 Khanghah, E. R. Study of the relationship between overinvestment, leverage, and liquidity in companies listed in tehran stock exchange / E. R. Khanghah, B. Partovi // Revista Genero & Direito. 2020. № 9 (2). P. 558 – 572.
- 6 Pivnicka, K. Fish species-diversity in small czech and moravian streams / K. Pivnicka, J. Poupe; M. Svatora, 1995. № 40 (4). P. 177 – 180.
- 7 Андрейчиков, А. З. Интеллектуальные информационные системы / А. З. Андрейчиков, О.Н. Андрейчикова // М.: Финансы и статистика.–2004. 424с.
- 8 Бардин, А.К. Особенности оценки эффективности производственных процессов при их дифференциации / А. К. Бардин, В. Е. Сайкинов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2016. № 124. С. 1493 – 1504.

9 РД ФГБОУ ВО «КНАГУ» 013-2016 Текстовые студенческие работы. Правила оформления. - Введ. 2016-03-10. - Комсомольск-на-Амуре: ФГБОУ ВО «КНАГУ», 2016 – 55 с.

10 Ковалев, С. М. Интеллектуальные информационные технологии на производстве и в промышленности: аналитический обзор. / С. М. Ковалев, А. Н. Гуда, А. В. Суханов // Вестник РГУПС. – 2016. № 2. С. 54 – 71.

11 Ковалев, С. М. Информационные технологии, автоматика и телекоммуникации. / С. М. Ковалев, В. Снашел, А. Н. Гуда, А. В. Суханов, П. Кромер // Вестник РГУПС. – 2017. № 4. С. 65 – 88.

12 Воробьев, А. Е. Цифровизация нефтяной промышленности: базовые подходы и обоснование «интеллектуальных» технологий. / Тчаро Хоноре, А. Е. Воробьев, К. А. Воробьев // Вестник Евразийской науки, 2018. - №2. С. 1 – 17.

13 Голубенко, П. В. Анализ внедрения IT-технологий и интеллектуальных систем в структуру управления предприятий атомной промышленности. / Современные научные исследования и разработки. Томский политехнический университет, 2018. - № 11. С. 183 – 185.

14 Кульмаметов, Р. Основы конструирования и технологии создания интеллектуальных систем для прецизионного контроля параметров водно-химических режимов теплоносителей энергосистем и сточных вод в нефтегазовой промышленности. / Р. Кульмаметов, Я. Ф. Карамова // 14 Туполевские чтения. 2019. С. 715 –718.

15 Абдрахманов, А. Д. Цифровизация промышленности как новый способ развития бизнеса / А. Д. Абдрахманов, А. Е. Кунакбаева, А. О. Данковский// Актуальные проблемы развития управленческой и сервисной деятельности в цифровой среде. 2020 С. 124 – 127.

16 Ван Юань, Применение цифровых и интеллектуальных технологий в промышленности Китая / Юань Ван // Журнал Белорусского государственного университета. Экономика. – 2020. С. 100 – 105.

17 Эусебек Д. Э. Применение методов интеллектуального анализа данных в фармацевтической промышленности. / Д. Эусебек, Р. Б. Абдрахманов // Colloquium-journal. – 2020. № 31-1 (83). С. 31– 36.