



ЭКСПЕРИМЕНТАЛЬНАЯ ЭКОНОМИКА И ФИНАНСЫ

№1 2023

СОДЕРЖАНИЕ

МАТЕМАТИЧЕСКИЕ, СТАТИСТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ В ЭКОНОМИКЕ

Смирнов С.Д., Булгаков А.Л., Гончарова А.В.
Создание унифицированного представления
данных с сайта ООН для повышения точности
алгоритмов машинного обучения 3

Уринцов А.И., Алешина А.В., Мустафин А.М.
Нормализация нормативно-справочной
информации на примере справочника
номенклатур 17

ОТРАСЛЕВАЯ И РЕГИОНАЛЬНАЯ ЭКОНОМИКА

Дадов Т.А.
Пути повышения конкурентоспособности
авиатранспортной отрасли Российской
Федерации в условиях санкционных
ограничений 28

ФИНАНСЫ

Ермошин С.А., Булгаков А.Л., Шапова К.Д.
Создание HTML файла из RSS-ленты:
автоматизация обработки торговых лотов
с использованием Python 34

МЕНЕДЖМЕНТ

Воронкова С.Н., Алешина А.В., Ильин И.И.,
Булгакова Т.Н., Овсян С.А.
Актуальные проблемы банкротства
физических лиц 39

НАУЧНЫЙ ЖУРНАЛ

Основан в июне 2021 года

Выходит 4 раза в год

Зарегистрирован Федеральной службой по надзору в
сфере связи, информационных технологий и массовых
коммуникаций (РОСКОМНАДЗОР).

Рег. ПИ № ФС77-81413 от 30 июня 2021 года

ISSN 2782-3644

Учредители:

ООО «Издательство «КноРус»

ООО «Институт экспериментальной экономики и финансов
МГУ имени М.В. Ломоносова»

Адрес редакции:

Россия, 117218, Москва,
ул. Кедрова, д. 14, корп. 2
Многоканальный телефон/факс: +7 (495) 741-46-28

Сайт: www.eeaf.ru

Почта: welcome@eeaf.ru

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

А.Л. Булгаков

Главный редактор:

А.В. Алешина

Отпечатано в типографии ООО «Русайнс», 117218, Мо-
сква, ул. Кедрова, д. 14, корп. 2

Тираж 300 экз. Формат А4. Подписано в печать: 30.03.2023
Цена свободная

Все материалы, публикуемые в журнале, подлежат вну-
треннему и внешнему рецензированию

Издание не подлежит маркировке согласно п. 2 ст. 1
Федерального закона от 29.12.2010 № 436-ФЗ «О защите
детей от информации, причиняющей вред их здоровью и
развитию»

К СВЕДЕНИЮ АВТОРОВ

Уважаемые коллеги! Обращаем ваше внимание на то, что матери-
алы статей проходят обязательную экспертизу. После экспертизы
статьи поступают в Редакцию журнала, где проходят редактор-
скую и корректорскую правку. Редакция оставляет за собой право
сокращать объем статей и редактировать их в соответствии с тре-
бованиями научного журнала. Рукописи статей не возвращаются;
с авторами в переписку Редакция не вступает; гонорар авторам
не выплачивается.

ПЕРЕЧЕНЬ ДОКУМЕНТОВ И ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ СТАТЕЙ И СОПУТСТВУЮЩИХ МАТЕРИАЛОВ, НЕОБХОДИМЫХ ДЛЯ ПУБЛИКАЦИИ

Текст статьи, при оформлении которого необходимо соблюсти
следующие требования: объем статьи - до 60 тыс. знаков (1,5 авт.
листа); в статье должна быть следующая информация: ФИО авто-
ра(ов) полностью, место работы (учебы), контактная информация
(телефон, E-mail); аннотация и ключевые слова к статье, список
литературы (на русском и английском языках).

TABLE OF CONTENTS

MATHEMATICAL, STATISTICAL AND INSTRUMENTAL METHODS IN ECONOMICS

Smirnov S.D., Bulgakov A.L., Goncharova A.V.
Creation of a unified representation of data
from the UN website to improve the accuracy
of machine learning algorithms 3

Urintzov A.I., Aleshina A.V., Mustafin A.M.
Normalization of normative and reference
information using the example of a reference
book of nomenclatures 17

INDUSTRY AND REGIONAL ECONOMICS

Dadov T.A.
Ways to increase the competitiveness of the air
transport industry of the Russian Federation
under sanctional restrictions 28

FINANCE

Ermoshin S.A., Bulgakov A.L., Shchapova K.D.
Creating HTML file from RSS Feed: automating
trading lot processing with Python 34

MANAGEMENT

*Voronkova S.N., Aleshina A.V., Ilyin I.I.,
Bulgakova T.N., Ovsoyan S.A.*
Current issues of bankruptcy of individuals. 39

SCIENTIFIC JOURNAL

*Founded in June 2021
Published 4 times a year*

*Registered by the Federal Service for Supervision of
Communications, Information Technology and Mass
Media (ROSKOMNADZOR).*

Reg. PI No. FS77-81413 dated June 30, 2021

ISSN 2782-3644

Founders:

*Knorus Publishing House LLC, Institute of
Experimental Economics and Finance of the
Lomonosov Moscow State University LLC*

EDITORIAL OFFICE:

Russia, 117218, Moscow, Kedrova St., 14, bldg. 2
Multi-channel phone/fax: +7 (495) 741-46-28

Website: www.eeaf.ru

Mail: welcome@eeaf.ru

CHIEF EDITOR

Anna Valentinovna Aleshina

EDITORIAL TEAM

Andrey Leonidovich Bulgakov

*Printed at the printing house LLC Rusyns,
117218, Moscow, st. Kedrova, d.14, building 2*

*Circulation 300 copies. A4 format. Signed to print:
30.03.2023*

Free price

*All materials published in the journal are subject to
internal and external review.*

*The publication is not subject to labeling in
accordance with paragraph 2 of Art. 1 of the Federal
Law of December 29, 2010 No. 436-FZ "On the
Protection of Children from Information Harmful to
Their Health and Development"*

МАТЕМАТИЧЕСКИЕ, СТАТИСТИЧЕСКИЕ
И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ В ЭКОНОМИКЕ**Создание унифицированного представления
данных с сайта ООН для повышения
точности алгоритмов машинного обучения****Смирнов Сергей Дмитриевич**аспирант кафедры, экономический факультет МГУ
имени М. В. Ломоносова,
E-mail: smirbox@gmail.com**Булгаков Андрей Леонидович**к.э.н., профессор МИСАО,
доцент РЭУ имени Г. В. Плеханова,
с.н.с. МГУ имени М.В.Ломоносова
E-mail: bulgakoval@mymsu.ru**Гончарова Анастасия Валерьевна**РЭУ имени Г. В. Плеханова
E-mail: AnastasiaVG2002@yandex.ru

В данной статье рассматриваются современные подходы к повышению точности алгоритмов машинного обучения. Актуальность темы обусловлена постоянно растущими требованиями к эффективности моделей искусственного интеллекта в условиях увеличения сложности решаемых задач и объемов обрабатываемых данных. Автор анализирует ряд ключевых направлений, способствующих улучшению производительности алгоритмов. Особое внимание уделяется методам оптимизации параметров моделей, техникам объединения моделей, применению глубокого обучения, использованию предобученных моделей, а также работе с данными и методам регрессии и регуляризации. В статье приводятся мнения ведущих исследователей в области машинного обучения, обсуждаются преимущества и ограничения каждого из рассмотренных подходов. Автор подчеркивает, что повышение точности алгоритмов машинного обучения является комплексной задачей, требующей применения различных техник в зависимости от конкретной области применения и характера обрабатываемых данных. Статья представляет интерес для специалистов в области искусственного интеллекта, разработчиков систем машинного обучения, а также исследователей, работающих над улучшением производительности алгоритмов в различных прикладных областях.

Ключевые слова: машинное обучение, нейросети, параметрическое пространство, ансамбли нейросетей, глубокое обучение.

Введение

В современном мире искусственного интеллекта постоянно растут требования к эффективности и точности алгоритмов машинного обучения. По мере того, как усложняются решаемые задачи и увеличивается объемы обрабатываемых данных, исследователи разрабатывают все новые и новые методы и подходы для улучшения производительности моделей. В данной статье предлагаются к рассмотрению ряд перспективных направлений, которые способствуют повышению точности алгоритмов машинного обучения.

Оптимизация параметров моделей

Одним из ключевых направлений является оптимизация параметров моделей. Бергстра и соавторы отмечают: «Автоматический выбор параметров может значительно улучшить производительность алгоритмов машинного обучения, особенно в случае сложных моделей с большим количеством настроек» [1]. Исследователи предлагают использовать алгоритмы случайного поиска и байесовской оптимизации для эффективного подбора параметров (hyperparameters).

Автоматизация процесса настройки моделей выступает как ключевое достоинство этого метода. Данное преимущество приобретает особую значимость при взаимодействии с масштабными и комплексными архитектурными решениями. Тем не менее, важно принимать во внимание,

что подобная оптимизация может быть ресурсоемкой с точки зрения вычислительных мощностей, в частности, когда речь идет о моделях с обширным набором регулируемых параметров.

Методы объединения моделей

Еще один перспективный подход — применение методов объединения моделей (ensemble methods). Чжоу в своем исследовании анализирует различные техники построения ансамблей и приходит к выводу: «Комбинирование нескольких моделей часто позволяет достичь более высокой точности по сравнению с отдельными алгоритмами» [2]. Автор рассматривает такие методы, как бэггинг (bagging), бустинг (boosting) и стекинг (stacking), отмечая их преимущества и ограничения.

Способы комбинирования моделей способствуют уменьшению дисперсии и смещения, что зачастую ведет к совершенствованию обобщающих свойств. Тем не менее, необходимо принимать во внимание, что применение ансамблевых методов может повлечь за собой рост вычислительной нагрузки и продолжительности тренировки моделей.

Глубокое обучение

Существенную лепту в повышение точности алгоритмов машинного обучения вносит глубокое обучение (deep learning). В своем исследовании Гудфеллоу с соавторами акцентируют внимание на следующем: «Глубокие нейронные сети способны автоматически извлекать сложные признаки из данных, что часто приводит к значительному улучшению качества моделей» [3]. Вместе с тем, исследователи указывают на ряд трудностей, сопряженных с обучением глубоких сетей, среди которых выделяются проблемы затухания градиентов и переобучения.

Ключевое достоинство глубоких нейронных сетей кроется в их потенциале обрабатывать массивные наборы данных и справляться с комплексными задачами в разнообразных сферах, охватывающих компьютерное зрение

и анализ естественного языка. Однако следует учитывать, что процесс обучения глубоких моделей сопряжен с потребностью в существенных вычислительных мощностях и обширных массивах информации.

Использование предобученных моделей:

Ван и Ли [4] в своем исследовании предлагают любопытный метод повышения точности, основанный на применении предобученных моделей (pre-trained models). Они постулируют: «Использование предобученных моделей и адаптация их к новым задачам позволяет значительно улучшить результаты, особенно в условиях ограниченных данных» [4]. Исследователи иллюстрируют результативность данного подхода на примерах из области компьютерного зрения и обработки естественного языка.

Предобученные модели демонстрируют особую ценность в контекстах, где для решения специфической задачи имеется лишь ограниченный объем аннотированных данных. Тем не менее, важно принимать во внимание, что эффективность данной стратегии в значительной степени обусловлена степенью сходства между изначальной проблематикой и целевой задачей.

Работа с данными:

Улучшение качества данных и применение эффективных методов предобработки могут значительно повысить точность моделей. Однако стоит помнить, что некоторые методы увеличения объема данных могут быть специфичными для конкретных задач и не всегда применимы в общем случае.

Методы регрессии и регуляризации:

Наконец, нельзя не упомянуть о роли методов регрессии и регуляризации в повышении точности алгоритмов. Тибширани [6] в своей работе показывает: «Применение различных форм регуляризации позволяет снизить риск

переобучения и улучшить обобщающую способность моделей» [6]. Автор анализирует такие методы, как L1 и L2 регуляризация, dropout и ранняя остановка (*early stopping*).

Регуляризация помогает моделям лучше обучаться на новых данных, предотвращая переобучение. Однако выбор оптимального метода и степени регуляризации может быть нетривиальной задачей и часто требует экспериментального подбора.

Повышение точности алгоритмов машинного обучения является комплексной задачей, требующей применения различных подходов и техник. Комбинирование методов оптимизации параметров, объединения моделей, глубокого обучения и других рассмотренных подходов позволяет достичь значительных улучшений в производительности моделей.

Каждый из рассмотренных методов имеет свои преимущества и ограничения. Оптимизация параметров и методы объединения моделей могут быть особенно эффективны для улучшения классических алгоритмов машинного обучения. Глубокое обучение и использование предобученных моделей показывают выдающиеся результаты в задачах обработки сложных данных, таких как изображения и тексты. Работа с данными и методы регуляризации являются универсальными подходами, применимыми к широкому спектру задач и моделей.

Исследуемая ситуация

Имеем международные статистические данные ООН по Валовой Добавленной Стоимости (ВДС) в зависимости от вида экономической деятельности. Валовая добавленная стоимость (ВДС) представляет собой первичные доходы единиц-резидентов, участвующих в производстве товаров и услуг. ВДС исчисляется на уровне отраслей как разность между выпуском товаров и услуг и промежуточным потреблением. Рассчитывается в основных ценах. Необходимо обработать данные и представить их в удобном для работы формате, а затем применить технологии машинного обучения для прогнозирования ВДС в зависимости от страны.

Проблема

В настоящее время на сайте ООН в открытом доступе имеется большой объем информации, интересующей самые разные категории пользователей. Однако данные представлены в неструктурированном виде, при котором их визуальное восприятие, анализ и прогнозирование динамики изменения ВДС с использованием машинного обучения затруднительны без осуществления обработки *BigData*.

Цель исследования

Разработать универсальную, удобную и понятную логическую структуру для обработки различных баз данных с сайта ООН и продемонстрировать применение данной структуры для прогнозирования динамики ВДС в стране с помощью технологий машинного обучения.

Задачи исследования

1. Импорт базы данных
2. Предобработка данных
3. Анализ данных
4. Составление логической структуры
5. Подготовка к машинному обучению
6. Создание моделей машинного обучения

1. Импорт базы данных

Импортируем необходимые библиотеки

```
#структуры и анализ данных
#pandas — https://pandas.pydata.org
import pandas as pd

# операции с регулярными выражениями
#re — https://docs.python.org/3/library/re.html
import re

# предварительная обработка данных
# sklearn.preprocessing — https://scikit-learn.org/stable/modules/preprocessing.html
from sklearn.preprocessing import
```

LabelEncoder

```
# построение графиков
# plotly.express — https://plotly.com/python/
plotly-express/
import plotly.express as px

# чтение и запись файлов CSV
# csv — https://docs.python.org/3/library/csv.
html
```

import csv

Скачаем файл с данными о ВДС по видам экономической деятельности непосредственно с сайта ООН <https://data.un.org/>

```
data = pd.read_csv('https://data.un.org/_Docs/
SYB/CSV/SYB65_153_202209_Gross%20Value%20
Added%20by%20Economic%20Activity.csv', header = 1)
data.head()
```

	Region/Country/Area	Unnamed: 1	Year	Series	Value	Footnotes	Source
0	4	Afghanistan	1995	Agriculture, hunting, forestry and fishing (% ...	66.3	Data classified according to ISIC Rev. 4.	United Nations Statistics Division, New York, ...
1	4	Afghanistan	2005	Agriculture, hunting, forestry and fishing (% ...	37.3	Data classified according to ISIC Rev. 4.	United Nations Statistics Division, New York, ...
2	4	Afghanistan	2010	Agriculture, hunting, forestry and fishing (% ...	33.2	Data classified according to ISIC Rev. 4.	United Nations Statistics Division, New York, ...
3	4	Afghanistan	2015	Agriculture, hunting, forestry and fishing (% ...	27.3	Data classified according to ISIC Rev. 4.	United Nations Statistics Division, New York, ...
4	4	Afghanistan	2018	Agriculture, hunting, forestry and fishing (% ...	23.2	Data classified according to ISIC Rev. 4.;Excl...	United Nations Statistics Division, New York, ...

Рис. 1. Полученная база данных с сайта ООН по видам экономической деятельности (данные собраны авторами) [7].

Опишем основные столбцы полученной базы данных:

- **Unnamed: 1** — страна
- **Year** — год фиксации статистического значения
- **Series** — сфера экономической деятель-

ности

- **Value** — % от валовой добавленной стоимости (ВДС)

Посмотрим, как выглядит таблица с данными непосредственно на сайте ООН

14 Valeur ajoutée brute par type d'activité économique

Percentage distribution, current prices
Répartition en pourcentage, aux prix courants

Country or area ⁶	1995	2005	2010	2015	2018	2019	2020	Pays ou zone ⁶
Afghanistan ¹								Afghanistan ¹
Agriculture	66.3	37.3	33.2	27.3	23.2 ²	27.0 ²	28.3 ²	Agriculture
Industry	10.6	21.3	13.0	10.8	14.1 ³	14.8 ³	13.1 ³	Industrie
Services	23.0	41.4	53.8	61.9	62.7 ⁴	58.2 ⁴	58.7 ⁴	Services
Albania ¹								Albanie ¹
Agriculture ²	54.3	21.5	20.7	22.5	21.1	21.0	21.9	Agriculture ²
Industry ³	26.5	28.7	28.7	24.8	24.3	23.6	22.9	Industrie ³
Services ⁴	19.2	49.8	50.7	52.7	54.6	55.4	55.2	Services ⁴
Algeria ¹								Algérie ¹
Agriculture	10.4	8.0	8.6	12.1	12.3	12.9	14.6	Agriculture
Industry	47.9	59.7	51.4	37.3	40.6	39.0	33.5	Industrie
Services	41.7	32.3	40.0	50.6	47.1	48.1	51.9	Services
Andorra ¹								Andorre ¹
Agriculture ²	0.5	0.4	0.5	0.6	0.6	0.6	0.6	Agriculture ²
Industry ³	18.8	17.7	14.5	11.2	12.4	13.0	12.4	Industrie ³
Services ⁴	80.7	81.9	85.0	88.2	87.0	86.4	87.0	Services ⁴
Angola ¹								Angola ¹
Agriculture	7.4	5.0	6.2	9.1	8.6	7.0	9.9	Agriculture
Industry	67.4	59.7	52.1	42.1	48.1	52.3	46.8	Industrie
Services	25.2	35.3	41.7	48.8	43.3	40.7	43.3	Services
Anguilla ¹								Anguilla ¹
Agriculture	3.3	2.7	2.0	2.3	1.5	2.8	1.0	Agriculture
Industry	14.5	19.8	15.8	15.3	21.3	11.8	15.1	Industrie
Services	82.2	77.4	82.2	82.4	77.2	85.3	83.9	Services
Antigua and Barbuda ¹								Antigua-et-Barbuda ¹
Agriculture	1.9	2.0	1.8	1.8	2.0	1.9	2.5	Agriculture
Industry	14.4	16.5	18.3	18.7	23.8	23.6	23.8	Industrie
Services	83.7	81.5	79.9	79.5	74.2	74.5	73.7	Services
Argentina								Argentine
Agriculture	5.3	9.3	8.5	6.1	6.3	7.3	8.2	Agriculture
Industry	26.0	33.7	30.1	27.5	27.4	27.9	26.9	Industrie
Services	68.7	57.0	61.4	66.3	66.4	64.8	65.0	Services

Рис. 2. Данные на сайте ООН [7].

Заметим, что данные в формате csv (переменная *data*) не соответствуют исходной таблице с сайта ООН, поэтому их необходимо преобразовать.

2. Предобработка данных

Оставим только информативные столбцы

```
data = data[['Unnamed: 1':  
1', 'Year', 'Series', 'Value']]
```

Проверим базу данных на отсутствие пропусков и соответствие типов

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 4466 entries, 0 to 4465  
Data columns (total 4 columns):  
#   Column      Non-Null Count  Dtype  
---  ---  
0   Unnamed: 1  4466 non-null   object  
1   Year        4466 non-null   int64  
2   Series      4466 non-null   object  
3   Value       4466 non-null   float64  
dtypes: float64(1), int64(1), object(2)  
memory usage: 139.7+ KB
```

Рис. 3. Описание параметров датасета

Заметим, что пропусков в данных нет, и все столбцы имеют тип данных в соответствии со значениями в них.

Столбец *value* может содержать большие числовые значения, разделенные запятой на разряды, поэтому для того, чтобы сделать данный код универсальным для различных наборов данных с сайта ООН преобразуем столбец *value* к числовому формату

```
data['Value'] = pd.to_numeric(data['Value'])  
data['Value'].dtype  
dtype('float64')
```

Переименуем столбцы для удобства работы с таблицей

```
data = data.rename(columns={'Unnamed: 1':  
'country', 'Year': 'year', 'Series': 'field', 'Value': 'value'})
```

	country	year	field	value
0	Afghanistan	1995	Agriculture, hunting, forestry and fishing (% ...	66.3
1	Afghanistan	2005	Agriculture, hunting, forestry and fishing (% ...	37.3
2	Afghanistan	2010	Agriculture, hunting, forestry and fishing (% ...	33.2
3	Afghanistan	2015	Agriculture, hunting, forestry and fishing (% ...	27.3
4	Afghanistan	2018	Agriculture, hunting, forestry and fishing (% ...	23.2
...	...	...	...	...
4461	Zimbabwe	2010	Services (% of gross value added)	65.6
4462	Zimbabwe	2015	Services (% of gross value added)	65.7
4463	Zimbabwe	2018	Services (% of gross value added)	67.9
4464	Zimbabwe	2019	Services (% of gross value added)	67.3
4465	Zimbabwe	2020	Services (% of gross value added)	67.4

4466 rows x 4 columns

Рис. 4. Пример модифицированной базы данных

Таким образом наша база данных имеет следующие столбцы:

- **country** — страна
- **year** — год фиксации статистического значения
- **field** — сфера экономической деятельности
- **value** — % от валовой добавленной стоимости(ВДС)

Заметим, что в столбце **field** строки имеют неинформативную запись (*% of gross value added*), указывающая на то, что данные столбца **value** представляют собой процентное значение. Создадим универсальный код для удаления лишней информации в скобках столбца **field**. Будем спрашивать пользователя о необходимости удаления такой информации, так как в некоторых базах данных в зависимости от значения поля *field* значения поля *value* измеряются по-разному и данная информация будет необходима для лучшего понимания полученных данных из логического файла.

```
print('Убрать скобки с информацией  
из строк столбца field?')  
print('Да — 1, Нет — 0')  
flag = int(input())
```

```
if flag == 1:
    data['field'] = data['field'].apply(lambda x:
re.sub(r'([.]*?)', '', x))
```

Убрать скобки с информацией из строк столбца field?
Да - 1, Нет - 0
1

	country	year	field	value
0	Afghanistan	1995	Agriculture, hunting, forestry and fishing	66.3
1	Afghanistan	2005	Agriculture, hunting, forestry and fishing	37.3
2	Afghanistan	2010	Agriculture, hunting, forestry and fishing	33.2
3	Afghanistan	2015	Agriculture, hunting, forestry and fishing	27.3
4	Afghanistan	2018	Agriculture, hunting, forestry and fishing	23.2
...	...	...	...	...
4461	Zimbabwe	2010	Services	65.6
4462	Zimbabwe	2015	Services	65.7
4463	Zimbabwe	2018	Services	67.9
4464	Zimbabwe	2019	Services	67.3
4465	Zimbabwe	2020	Services	67.4

4466 rows x 4 columns

Рис. 5. Пример модифицированной базы данных датасета 2.

Проведем кодирование категорий для удобства составления логической структуры.

```
# Порядковое кодирование
le=LabelEncoder()
data['field']=le.fit_transform(data['field'].
values)
data['field'].unique()
```

array([0, 1, 2])

Таким образом, различные виды экономической деятельности получили соответствующие им номера:

- 0 — Agriculture, hunting, forestry and fishing
- 1 — Industry
- 2 — Services

3. Анализ данных

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4466 entries, 0 to 4465
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   country     4466 non-null   object
1   year        4466 non-null   int64
2   field       4466 non-null   int64
3   value       4466 non-null   float64
dtypes: float64(1), int64(2), object(1)
memory usage: 139.7+ KB
```

Рис. 6. Структура данных модифицированного датасета.

```
data['country'].nunique()
```

214

```
data['year'].unique()
```

```
array([1995, 2005, 2010, 2015, 2018, 2019, 2020])
```

```
data.duplicated().sum()
```

0

Следующим параметром проанализируем цену и воспользуемся для этого «Ящиком с усами» (Box plot). Ящик с усами — простой и удобный график, показывающий одномерное распределение вероятностей, или, проще говоря, концентрацию данных. Отрисовывает медиану (линия в центре), верхний и нижний квартили (стороны ящика), края статистически значимой выборки («усы») и выбросы (точки за «усами»). График позволяет быстро оценить где располагается большая часть данных (50% находятся внутри ящика), их симметричность (смещение медианы к одной из сторон ящика и/или длина «усов») и степень разброса — дисперсию (размеры ящика, размеры усов и количество точек-выбросов).

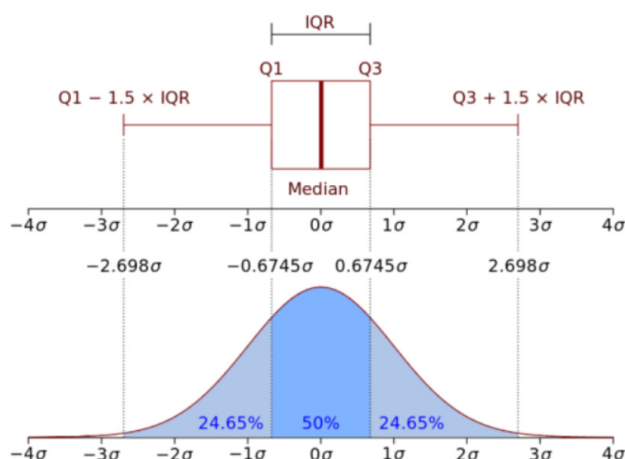


Рис. 7. Пример распределения выбросов в базе данных.

```
fig = px.box(data, x =»value«)
fig.show()
```

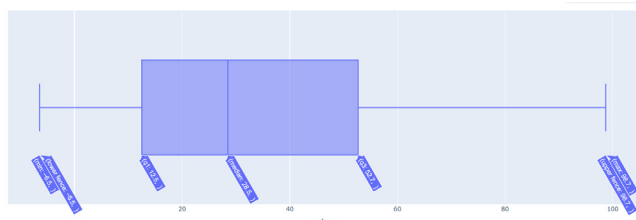


Рис. 8. Основные характеристики распределения.

```
fig = px.box(data, x =»field«, y =»value«)
fig.show()

fig = px.histogram(data, x =»field«, nbins = 6)
fig.show()

import seaborn as sns
sns.heatmap(data.corr(), annot = True)
```



Рис. 9. Корреляционная матрица ВДС и сферой экономической деятельности

Сделаем краткие выводы об имеющихся данных:

- В базе данных содержится 4466 записей о 214 странах
- Пропусков и дубликатов нет
- Данные распределены равномерно по столбцу field
- Наблюдается корреляция между значением ВДС и сферой экономической деятельности

4. Составление логической структуры

Для создания логического формата определим основные метки(параметры) для более удобной работы с данными: Индикатор.Объект.Период значение

- Индикатор характеризует измеритель процесса или явления – **(столбец field)**
- Объект характеризует конкретное место измерения индикатора – **Страна (столбец country)**
- Период характеризует время, в которое производились измерения – **Год (столбец year)**
- Значение – **ВДС (столбец value)**

В зависимости от наших потребностей будем фиксировать определенный параметр. Для более удобной фиксации в качестве подсказки пользователю будем выводить список уникальных значений выбранного столбца.

#Введем номер параметра, который хотим зафиксировать

```
print(«Enter the number»)
print(“”
0 — Фиксируем индикатор
1 — Фиксируем объект
2 — Фиксируем период”)
print()
n = int(input()) # 0, 1, 2
res = “
```

```
if n == 0:
    print(data[‘field’].unique()) # Фиксируем
экономическую сферу — индикатор
    print(“”
0 — Agriculture, hunting, forestry and fishing
1 — Industry
2 — Services”)
    print(«\nEnter field»)
    res = ‘field’
    tmp = input()
    print()
    print(‘Получим новый dataframe с нужной
структурой’)
    print()
    print(data.query(f’\nfield == {tmp} ’))
    new_data = data.query(f’\nfield == {tmp} ’)
    elif n == 1:
        print(data[‘country’].unique()) # Фиксируем
страну — объект
        print(«\nEnter country»)
        res = ‘country’
        tmp = input()
        print()
        print(‘Получим новый dataframe с нужной
структурой’)
        print()
        print(data.query(f’\ncountry == «{tmp}» ’))
        new_data = data.query(f’\ncountry ==
«{tmp}» ’)
        elif n == 2:
            print(data[‘year’].unique())
            print(«\nEnter year») # Фиксируем год — пе-
риод и экономическую сферу — индикатор
            tmp1 = input()
            print(data[‘field’].unique())
```

```
print(«\nEnter field»)
tmp2 = input()
print()
print(‘Получим новый dataframe с нужной
структурой’)
print()
print(data.query(f’\nyear == {tmp1} & field
== {tmp2} ’))
new_data = data.query(f’\nyear == {tmp1} &
field == {tmp2} ’)

Enter the number

0 — Фиксируем индикатор
1 — Фиксируем объект
2 — Фиксируем период

1
[‘Afghanistan’ ‘Albania’ ‘Algeria’ ‘Andorra’
‘Angola’ ‘Anguilla’
‘Antigua and Barbuda’ ‘Argentina’ ‘Armenia’
‘Aruba’ ‘Australia’ ‘Austria’
‘Azerbaijan’ ‘Bahamas’ ‘Bahrain’
‘Bangladesh’ ‘Barbados’ ‘Belarus’
‘Belgium’ ‘Belize’ ‘Benin’ ‘Bermuda’ ‘Bhutan’
‘Bolivia (Plurin. State of)’ ‘Bosnia and
Herzegovina’ ‘Botswana’ ‘Brazil’
‘British Virgin Islands’ ‘Brunei Darussalam’
‘Bulgaria’ ‘Burkina Faso’
‘Burundi’ ‘Cabo Verde’ ‘Cambodia’
‘Cameroon’ ‘Canada’ ‘Cayman Islands’
‘Central African Republic’ ‘Chad’ ‘Chile’
‘China’ ‘China, Hong Kong SAR’
‘China, Macao SAR’ ‘Colombia’ ‘Comoros’
‘Congo’ ‘Cook Islands’
‘Costa Rica’ ‘Côte d’Ivoire’ ‘Croatia’ ‘Cuba’
‘Curaçao’ ‘Cyprus’
‘Czechia’ ‘Dem. People’s Rep. Korea’ ‘Dem.
Rep. of the Congo’ ‘Denmark’
‘Djibouti’ ‘Dominica’ ‘Dominican Republic’
‘Ecuador’ ‘Egypt’
‘El Salvador’ ‘Equatorial Guinea’ ‘Eritrea’
‘Estonia’ ‘Eswatini’
‘Ethiopia’ ‘Fiji’ ‘Finland’ ‘France’ ‘French
Polynesia’ ‘Gabon’ ‘Gambia’
‘Georgia’ ‘Germany’ ‘Ghana’ ‘Greece’
‘Greenland’ ‘Grenada’ ‘Guatemala’
‘Guinea’ ‘Guinea-Bissau’ ‘Guyana’ ‘Haiti’
```

'Honduras' 'Hungary' 'Iceland'
 'India' 'Indonesia' 'Iran (Islamic Republic of)'
 'Iraq' 'Ireland'
 'Israel' 'Italy' 'Jamaica' 'Japan' 'Jordan'
 'Kazakhstan' 'Kenya'
 'Kiribati' 'Kosovo' 'Kuwait' 'Kyrgyzstan'
 «Lao People's Dem. Rep.»
 'Latvia' 'Lebanon' 'Lesotho' 'Liberia' 'Libya'
 'Liechtenstein'
 'Lithuania' 'Luxembourg' 'Madagascar'
 'Malawi' 'Malaysia' 'Maldives'
 'Mali' 'Malta' 'Marshall Islands' 'Mauritania'
 'Mauritius' 'Mexico'
 'Micronesia (Fed. States of)' 'Monaco'
 'Mongolia' 'Montenegro'
 'Montserrat' 'Morocco' 'Mozambique'
 'Myanmar' 'Namibia' 'Nauru' 'Nepal'
 'Netherlands' 'Netherlands Antilles [former]'
 'New Caledonia'
 'New Zealand' 'Nicaragua' 'Niger' 'Nigeria'
 'North Macedonia' 'Norway'
 'Oman' 'Pakistan' 'Palau' 'Panama' 'Papua
 New Guinea' 'Paraguay' 'Peru'
 'Philippines' 'Poland' 'Portugal' 'Puerto Rico'
 'Qatar'
 'Republic of Korea' 'Republic of Moldova'
 'Romania' 'Russian Federation'
 'Rwanda' 'Saint Kitts and Nevis' 'Saint Lucia'
 'Saint Vincent & Grenadines' 'Samoa' 'San
 Marino' 'Sao Tome and Principe'
 'Saudi Arabia' 'Senegal' 'Serbia' 'Seychelles'
 'Sierra Leone' 'Singapore'
 'Sint Maarten (Dutch part)' 'Slovakia'
 'Slovenia' 'Solomon Islands'
 'Somalia' 'South Africa' 'South Sudan' 'Spain'
 'Sri Lanka'
 'State of Palestine' 'Sudan' 'Sudan [former]'
 'Suriname' 'Sweden'
 'Switzerland' 'Syrian Arab Republic'
 'Tajikistan' 'Thailand'
 'Timor-Leste' 'Togo' 'Tonga' 'Trinidad and
 Tobago' 'Tunisia' 'Türkiye'
 'Turkmenistan' 'Turks and Caicos Islands'
 'Tuvalu' 'Uganda' 'Ukraine'
 'United Arab Emirates' 'United Kingdom'
 'United Rep. of Tanzania'
 'United States of America' 'Uruguay'
 'Uzbekistan' 'Vanuatu'

'Venezuela (Boliv. Rep. of)' 'Viet Nam'
 'Yemen' 'Zambia' 'Zanzibar'
 'Zimbabwe']

Enter country

Spain

Получим новый dataframe с нужной структурой

country year field value

3736 Spain 1995 0 4.2

3737 Spain 2005 0 3.1

3738 Spain 2010 0 2.6

3739 Spain 2015 0 3.0

3740 Spain 2018 0 3.0

3741 Spain 2019 0 2.9

3742 Spain 2020 0 3.4

3743 Spain 1995 1 30.6

3744 Spain 2005 1 30.3

3745 Spain 2010 1 25.2

3746 Spain 2015 1 22.1

3747 Spain 2018 1 22.0

3748 Spain 2019 1 22.2

3749 Spain 2020 1 22.3

3750 Spain 1995 2 65.2

3751 Spain 2005 2 66.6

3752 Spain 2010 2 72.1

3753 Spain 2015 2 74.9

3754 Spain 2018 2 75.0

3755 Spain 2019 2 74.9

3756 Spain 2020 2 74.2

Для возможности дальнейшей работы с преобразованными данными будем записывать получившийся dataframe в новый csv-файл.

new_data

new_data.to_csv('newf.csv')

	country	year	field	value
3736	Spain	1995	0	4.2
3737	Spain	2005	0	3.1
3738	Spain	2010	0	2.6
3739	Spain	2015	0	3.0
3740	Spain	2018	0	3.0
3741	Spain	2019	0	2.9
3742	Spain	2020	0	3.4
3743	Spain	1995	1	30.6
3744	Spain	2005	1	30.3
3745	Spain	2010	1	25.2
3746	Spain	2015	1	22.1
3747	Spain	2018	1	22.0
3748	Spain	2019	1	22.2
3749	Spain	2020	1	22.3
3750	Spain	1995	2	65.2
3751	Spain	2005	2	66.6
3752	Spain	2010	2	72.1
3753	Spain	2015	2	74.9

Рис. 10. Пример распределения ВДС по годам в Испании с 1995 года по 2015 год.

5. Подготовка к машинному обучению

Импортируем необходимые библиотеки

```
#Библиотека для машинного обучения
#scikit-learn — https://scikit-learn.org

import numpy as np
from sklearn.model_selection import train_
test_split
from sklearn import metrics
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.linear_model import
LinearRegression
```

Импортируем созданный нами новый набор данных

```
data1 = pd.read_csv(«/content/newf.csv»)
data1
```

	Unnamed: 0	country	year	field	value
0	3736	Spain	1995	0	4.2
1	3737	Spain	2005	0	3.1
2	3738	Spain	2010	0	2.6
3	3739	Spain	2015	0	3.0
4	3740	Spain	2018	0	3.0
5	3741	Spain	2019	0	2.9
6	3742	Spain	2020	0	3.4
7	3743	Spain	1995	1	30.6
8	3744	Spain	2005	1	30.3
9	3745	Spain	2010	1	25.2
10	3746	Spain	2015	1	22.1
11	3747	Spain	2018	1	22.0
12	3748	Spain	2019	1	22.2
13	3749	Spain	2020	1	22.3
14	3750	Spain	1995	2	65.2
15	3751	Spain	2005	2	66.6
16	3752	Spain	2010	2	72.1
17	3753	Spain	2015	2	74.9
18	3754	Spain	2018	2	75.0
19	3755	Spain	2019	2	74.9
20	3756	Spain	2020	2	74.2

Рис. 11. Пример датасета до подготовки к машинному обучению.

Удаляем колонки без названия

```
data1 = data1.drop(columns = ['Unnamed: 0'],
axis = 1)
data1
```

Проверим данные на null

```
data1.isnull().sum()

country    0
year       0
field      0
value     0
dtype: int64
```

Проверим данные на дублирующиеся записи

```
data1.duplicated().sum()
```

0

Проанализируем типы данных

```
data1.dtypes
```

```
country    object
year       int64
field      int64
value      float64
dtype: object
```

```
data1.describe()
```

	year	field	value
count	21.000000	21.000000	21.000000
mean	2011.714286	1.000000	33.323810
std	8.655304	0.83666	29.529374
min	1995.000000	0.00000	2.600000
25%	2005.000000	0.00000	3.400000
50%	2015.000000	1.00000	22.300000
75%	2019.000000	2.00000	66.600000
max	2020.000000	2.00000	75.000000

Рис. 12. Пример датасета после подготовки к машинному обучению

Создадим диаграммы, которые покажут нам баланс данных. Первая диаграмма *countplot* покажет нам, как часто встречается то или иное значение **value**. Вторая круговая диаграмма *pie* покажет нам, как часто встречается то или иное значение **value** в процентном соотношении.

```
plt.figure(figsize=(20,5))
plt.subplot(1,2,1)
sns.countplot(x='value', data=data1)
plt.title(«Показатель»)
plt.show()
```

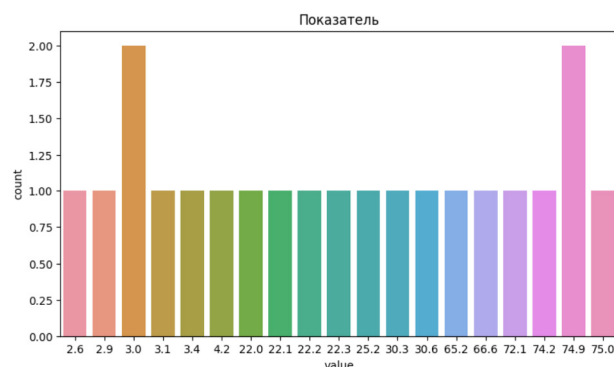


Рис. 13. Анализ выбросов в базе данных.

```
plt.figure(figsize=(15,7))
plt.subplot(1,2,1)
sns.countplot(x='field', data=data1)
plt.title(«Показатель»)
plt.subplot(1,2,2)
labels = data1['field'].value_counts(sort = True).index
sizes = data1['field'].value_counts(sort = True)
plt.pie(sizes, labels=labels,
autopct='%1.1f%%', startangle=270,)
plt.title('Процентное соотношение', size = 12)
plt.show()
```

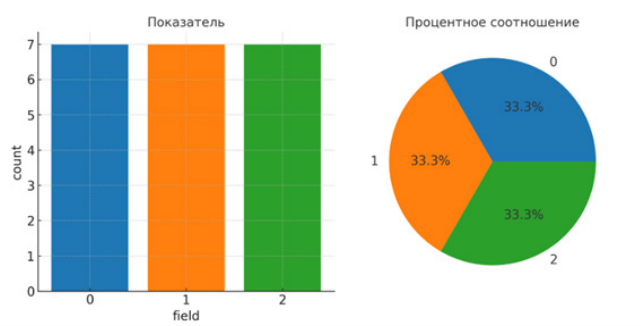


Рис. 14. Баланс данных в базе данных распределения ВДС.

Проверим, есть ли значения **value** меньше нуля

```
len(data1[data1['value'] < 0]) / len(data1)
```

Если такие есть, то удалим их для дальнейшей работы

```
data1.drop(data1[(data1['value']<0)].index,
inplace=True, axis=0)
```

Проведём кодирование столбца «Страны» (**country**)

```
le=LabelEncoder()
data1['country']=le.fit_
transform(data1['country'].values)
data1
```

Разобьём данные на выборки *train* и *test* для машинного обучения

```
X = data1[['country', 'year', 'field']]
Y = data1['value']
```

```
X_train, X_test, y_train, y_test = train_test_
split(X, Y, test_size=0.3, random_state=0)
```

6. Создание и оценка моделей машинного обучения

Создадим модель линейной регрессии и обучим её на наших данных. С помощью нашей модели попробуем предсказать, какие значения **value** мы получим на основе выборки *X_test*

```
model = LinearRegression()

model.fit(X_train, y_train)
```

Получаем результаты

```
from sklearn.metrics import mean_absolute_
error, mean_squared_error, r2_score
y = model.predict(X_test)
print ('MAE:', round (mean_absolute_error(y_
test, y),3)) #Метрики
print ('√MSE:', round (mean_squared_
error(y_test, y)**(1/2),3))
print ('R2_score:', round (r2_score(y_test,
y),3))
```

```
MAE: 7.875
√MSE: 9.433
R2_score: 0.844
```

```
model.score(X_train, y_train)

0.9691671890268119
```

```
from sklearn.tree import
DecisionTreeRegressor
TR = DecisionTreeRegressor(max_depth=2)
#Создаем модель
TR.fit(X_train, y_train) #Обучаем модель
Y_TR=TR.predict(X_test) #Предсказываем
значения для выборки
print ('MAE:', round (mean_absolute_error(y_
test, Y_TR),3)) #Метрики
print ('√MSE:', round (mean_squared_
error(y_test, Y_TR)**(1/2),3))
print ('R2_score:', round (r2_score(y_test,
Y_TR),3))
```

```
MAE: 2.487
√MSE: 3.056
R2_score: 0.984
```

```
TR.score(X_train, y_train)
```

```
0.9967412808835854
```

```
import matplotlib.pyplot as plt
```

```
dep, score=[],[]
for i in range(3,16):
TR = DecisionTreeRegressor(max_depth=i)
TR.fit(X_train, y_train)
Y_TR=TR.predict(X_test)
dep.append(i)
score.append(mean_squared_error(y_test,
Y_TR)**(1/2))
plt.rcParams['figure.figsize']=6,3
plt.plot(dep, score)
```

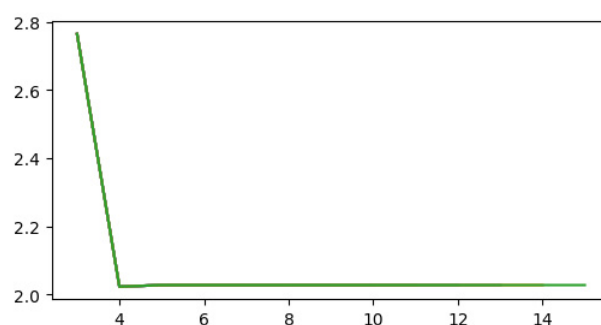


Рис. 15. Результаты предсказания значения ВДС с помощью метода линейной регрессии.

Заключение

Весь код, представленный ниже, может быть также применен к другим базам данных ООН с поправками определенных комментариев, написанных для лучшего понимания конкретного вопроса.

Список ссылок на некоторые файлы, для которых адаптирован проект (для удобства поделим данные на разделы):

Population and National accounts

GDP and GDP per capita: https://data.un.org/_Docs/SYB/CSV/SYB65_230_202209_GDP%20and%20GDP%20Per%20Capita.csv

Population, surface area and density: https://data.un.org/_Docs/SYB/CSV/SYB65_1_202209_Population,%20Surface%20Area%20and%20Density.csv

Population growth, fertility, life expectancy and mortality: https://data.un.org/_Docs/SYB/CSV/SYB65_246_202209_Population%20Growth,%20Fertility%20and%20Mortality%20Indicators.csv

Education

Education at the primary, secondary and tertiary levels: https://data.un.org/_Docs/SYB/CSV/SYB65_309_202209_Education.csv

Price and production indices

Consumer price indices: https://data.un.org/_Docs/SYB/CSV/SYB65_128_202209_Consumer%20Price%20Index.csv

Agricultural production indices: https://data.un.org/_Docs/SYB/CSV/SYB65_12_202209_Agricultural%20Index.csv

International merchandise trade

Total imports, exports and balance of trade: https://data.un.org/_Docs/SYB/CSV/SYB65_123_202209_Total%20Imports%20Exports%20and%20Balance%20of%20Trade.csv

Major Trading Partners: https://data.un.org/_Docs/SYB/CSV/SYB65_330_202209_Major%20Trading%20Partners.csv

Полный список файлов можно посмотреть на сайте <https://data.un.org/>

В данной статье была рассмотрена конкретная база данных. Была преобразована ее структура, создана своя логическая структура для более удобного анализа и обработки. Был реализован универсальный код для обработки

различных данных с сайта ООН. Были учтены различия в изначальных статистических данных и подготовлена модель линейной регрессии для машинного обучения.

Литература

1. Bergstra J., Yamins D., Cox D. D. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures // Proceedings of the 30th International Conference on Machine Learning. 2013. Vol. 28. P. 115–123.
2. Zhou Z. H. Ensemble Methods: Foundations and Algorithms. Chapman and Hall/CRC, 2012. 236 p.
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016. 800 p.
4. Wang J., Li M. Pre-trained Models for Natural Language Processing: A Survey // AI Open. 2021. Vol. 2. P. 225–236.
5. Smith J. W. The Effect of Data Quality on Machine Learning Algorithm Performance // Big Data. 2020. Vol. 8(1). P. 26–37.
6. Tibshirani R. Regression Shrinkage and Selection via the Lasso // Journal of the Royal Statistical Society: Series B (Methodological). 1996. Vol. 58(1). P. 267–288.
7. Статистические данные ООН по Валовой добавленной стоимости в разных странах по видам экономической деятельности // Портал ООН // URL: <https://data.un.org/> (дата обращения 03.02.2023)

CREATING A UNIFIED VIEW OF DATA FROM THE UN WEBSITE TO IMPROVE THE ACCURACY OF MACHINE LEARNING ALGORITHMS

Smirnov S. D., Bulgakov A. L., Goncharova A. V.

Lomonosov Moscow State University, Plekhanov Russian University of Economics, MISAO

This article discusses modern approaches to improving the accuracy of machine learning algorithms. The relevance of the topic is due to the ever-growing demands on the efficiency of artificial intelligence models in the context of increasing

complexity of problems and volumes of processed data. The author analyzes a number of key areas that contribute to improving the performance of algorithms. Particular attention is paid to methods for optimizing model parameters, model fusion techniques, the use of deep learning, the use of pre-trained models, as well as working with data and regression and regularization methods. The article presents the opinions of leading researchers in the field of machine learning, discusses the advantages and limitations of each of the approaches considered. The author emphasizes that improving the accuracy of machine learning algorithms is a complex task that requires the use of various techniques depending on the specific area of application and the nature of the data being processed. The article is of interest to specialists in the field of artificial intelligence, developers of machine learning systems, as well as researchers working on improving the performance of algorithms in various application areas.

Keywords: machine learning, neural networks, parametric space, neural network ensembles, deep learning.

References

1. Bergstra J., Yamins D., Cox D. D. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures // *Proceedings of the 30th International Conference on Machine Learning*. 2013. Vol. 28. P. 115–123.
2. Zhou Z. H. *Ensemble Methods: Foundations and Algorithms*. Chapman and Hall/CRC, 2012. 236 p.
3. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press, 2016. 800 p.
4. Wang J., Li M. Pre-trained Models for Natural Language Processing: A Survey // *AI Open*. 2021. Vol. 2. P. 225–236.
5. Smith J. W. The Effect of Data Quality on Machine Learning Algorithm Performance // *Big Data*. 2020. Vol. 8(1). P. 26–37.
6. Tibshirani R. Regression Shrinkage and Selection via the Lasso // *Journal of the Royal Statistical Society: Series B (Methodological)*. 1996. Vol. 58(1). P. 267–288.
7. UN Statistics on Gross Value Added in Different Countries by Type of Economic Activity // UN Portal // URL: <https://data.un.org/> (date accessed 03.02.2023)

МАТЕМАТИЧЕСКИЕ, СТАТИСТИЧЕСКИЕ
И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ В ЭКОНОМИКЕНормализация нормативно-справочной
информации на примере справочника
номенклатур**Уринцов Аркадий Ильич**

профессор, д. э. н., уч. звание «профессор»

E-mail: urintsov.ai@rea.ru

Алешина Анна Валентиновна

к. э. н., доцент экономического факультет МГУ

имени М. В. Ломоносова

E-mail: annaaleshina@mail.ru

Мустафин Артур Маратович

магистр РЭУ имени Г. В. Плеханова.

E-mail: mustafin_artur97@mail.ru

Качественно выстроенные справочники нормативно-справочной информации являются залогом успешных и качественных бизнес-процессов. Качественные (нормализованные) справочники позволяют бизнесу видеть точную информацию о Компании, принимать эффективные управленческие решения. Нормализация информации представляет собой инструмент, позволяющий получить качественные справочники, эффективно управлять данными НСИ, принимать эффективные управленческие решения. Общепринятого и эффективного способа проведения нормализацию на текущий момент не существует. Различные Компании и эксперты в указанной области используют различные инструменты и решения. Нормализацию осуществляют с применением минимум одного инструмента — MS Excel, путем поиска и актуализации информации в Интернет, нормативно-технической документации и документации поставщиков или производителей номенклатуры (если речь касается справочника Номенклатуры). Указанный способ не самый эффективный и предполагает высокие риски ошибки вследствие человеческого фактора. Для повышения эффективности нормализации данных предлагается использовать парсинг и машинное обучение.

Ключевые слова: нормализация данных, номенклатура, справочники, парсинг.

В эпоху цифровых технологий эффективное управление огромным объемом информации имеет решающее значение для принятия обдуманных и значимых решений, облегчения совместной работы и стимулирования инноваций. Разрозненный и неполный характер данных часто затрудняет их использование для анализа или системной интеграции. Нормализация служит мощным инструментом для наведения порядка и преобразования неструктурированных данных в формат, который является более управляемым, эффективным и надежным.

Обычно используются два основных типа нормализации: 1. Когда основное внимание уделяется организации данных в базах данных, позволяющие повысить целостность данных и сократить объем хранилища; 2. Нормализация номенклатуры — касается стандартизации имен, меток и идентификаторов. Она направлена на обеспечение согласованности имен файлов, структур папок, имен переменных и других соглашений об именовании, повышение ясности, удобочитаемости и совместимости в различных системах и инструментах. Используют следующие методы нормализации номенклатуры, например, преобразование всех имен в нижний или верхний регистр, замена пробелов символами подчеркивания, дефисами или другими разделителями, устранение знаков препинания, ударений или других специальных символов, добавление префиксов или суффиксов к именам для категоризации или указания свойств. Благода-

ря нормализации ускоряется обработка и поиск данных, а также уменьшается количество несоответствий и ошибок, связанных с дублированием. Основным преимуществом является то, что нескольким пользователям проще

работать с данными и понимать их.

История развития идеи информационного управления и машинного обучения насчитывает много десятилетий (см. рис. 1)

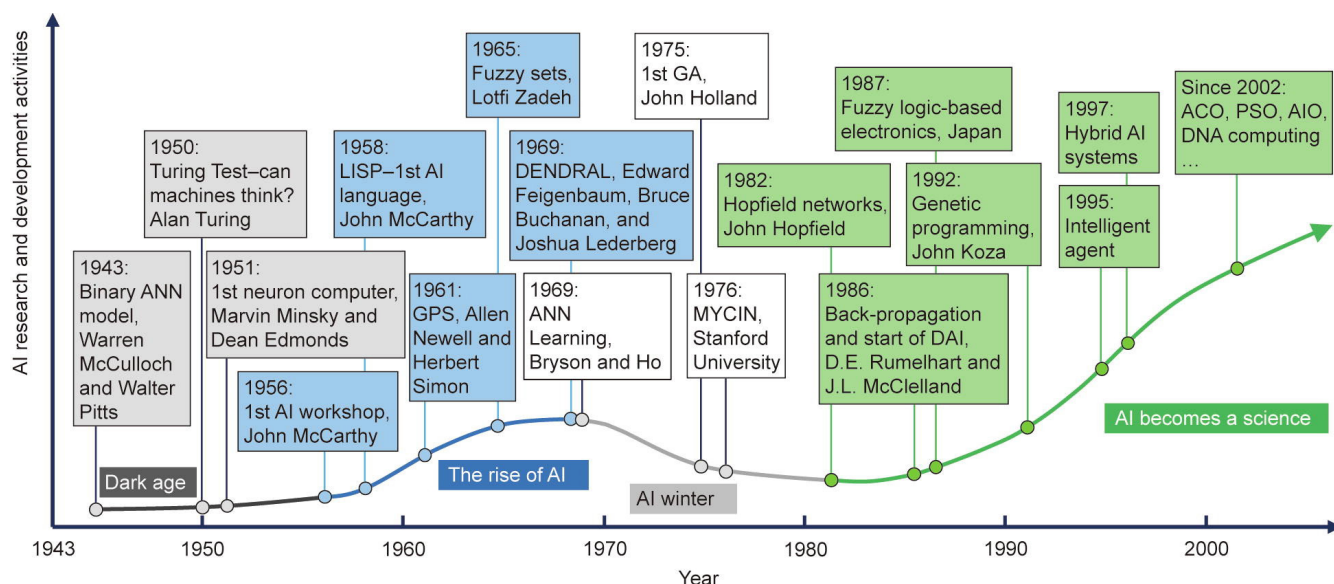


Рис. 1. Основные научные публикации по тематике искусственного интеллекта (1943–2022 годы) [1].

Бурное развитие алгоритмов, средств связи и вычислительных технологий заложило прочную основу для создания «умной» индустрии. В частности, появление алгоритмов искус-

ственного интеллекта (ИИ) способно сделать управление производством и нормативно-справочной информацией более интеллектуальным и интуитивно понятным [1].

Таблица 1. Типичные модели машинного обучения.

Модели машинного обучения	Контролируемый / неконтролируемый / полу-контролируемый	Дискриминационный / генеративный	Глубокое обучение / non-deep learning
К — означает кластеризацию	Не контролируемая	Генеративный	Не-глубокое обучение
К-ближайшие соседи	Контролируемая	Дискриминационный	Не-глубокое обучение
Машина опорных векторов	Контролируемая	Дискриминационный	Не-глубокое обучение
Скрытая марковская модель	Контролируемая	Дискриминационный	Не-глубокое обучение
Случайный лес	Контролируемая	Дискриминационный	Не-глубокое обучение
XGBoost	Контролируемая	Дискриминационный	Не-глубокое обучение
Ансамблевые методы	Контролируемая	Дискриминационный	Не-глубокое обучение
Сверточная нейронная сеть	Контролируемая	Дискриминационный	Глубокое обучение

Модели машинного обучения	Контролируемый / неконтролируемый / полу-контролируемый	Дискриминационный / генеративный	Глубокое обучение / non-deep learning
Рекуррентная нейронная сеть	Контролируемая	Дискриминационный	Глубокое обучение
Долговременная кратковременная память	Контролируемая	Дискриминационный	Глубокое обучение
Наивный Байес	Контролируемая	Генеративный	Не-глубокое обучение
Модель гауссовой смеси	Контролируемая	Генеративный	Не-глубокое обучение
Генеративные состязательные сети	Под частичным наблюдением	Генеративный	Глубокое обучение

Источник: составлено на основе статьи L.Wang. *From intelligence science to intelligent manufacturing Engineering* [1]

Благодаря Интернету разрозненные производственные процессы и системы могут быть объединены в сеть и скоординированы между собой [2]. В свою очередь, стандартизация является важным требованием для интеграции различных систем и процессов. Стандартизация имеет решающее значение для максимизации результатов бизнеса при использовании различных технологий и решений в интеллектуальном производстве и не только. Отметим, что эффективное администрирование сетевой инфраструктуры имеет важное значение для предприятий любого размера в современном быстро меняющемся технологическом мире. Традиционные методы управления информацией часто оказываются трудоемкими, приводящими к ошибкам и сложными в масштабировании [3].

Нормативно-справочная информация (НСИ) действует как стандартизированная структура для управления данными, относящимися к операциям различных участников в рамках конкретных процессов. Эта структура включает реестры, справочники, классификаторы и модификаторы, которые призваны унифицировать информацию в различных областях деятельности. Устанавливая единые стандарты обработки данных, НСИ обеспечивает общее понимание и способствует эффективной коммуникации и сотрудничеству. Стоит отметить, что упорядочивая нормативную и справочную информацию в единую систему,

повышается надежность аналитических данных, появляется возможность создания отчетов с высокой степенью детализации и оптимизация процедур обработки информации [4].

Машинное обучение революционизирует методы администрирования большими объемами нормативной информации, в особенности в следующих областях: обнаружение ошибок, очистка данных, интеграция и выполнение запросов. Исследования последних лет показали, что использование алгоритмов ИИ может частично автоматизировать сопоставление, удаление дубликатов и стандартизацию данных. После первоначальной ручной настройки модели машинного обучения можно обучать на новых массивах данных, чтобы постоянно улучшать ее способности. Отмечается, что чем больше данных обрабатывает модель машинного обучения, тем эффективнее она становится. В рамках решения задачи автоматизации систем нормативно-справочной информации отдельные компоненты последней могут функционировать как независимые подсистемы, каждая из которых требует собственного управления [5]. При этом иерархический вид группирования данных является наиболее востребованным при составлении справочников НСИ, поскольку дает уникальные возможности весьма стремительно обнаруживать необходимые фрагменты (подгруппы), и, соответственно, играет ключевую роль в формировании этих систем.

Такой подход, как обработка естественного языка (NLP) на сегодняшний день является одним из самых востребованных и используется в голосовых помощниках, автоматических переводах текста и фильтрации текста. В области искусственного интеллекта (ИИ) методы машинного обучения (МО), глубокого обучения (ГО) и обработки естественного языка (NLP), разработанные для обработки и понимания текстовых документов, помогают организовывать синтаксические отношения и структуру языка для автоматической обработки и изучения. Однако, эти методы, основанные на искусственном интеллекте, требуют валидации и тщательного тестирования надежности в различных областях применения [6].

Чтобы автоматически суммировать большое количество нормативно-справочной информации, каждый документ должен быть декомпозирован для получения встроенных в него функций с использованием алгоритмов мягких вычислений. Алгоритмы генерируют краткое описание для данного единичного документа или набора патентных документов. Подход к автоматическому обобщению может быть экстрактивным или абстрагирующим от данного документа. Экстрактивный метод объединяет наиболее значимые абзацы, предложения или фразы, взятые из документа(ов). Абстрактный метод генерирует резюме, используя предложения и словосочетания, которые могут быть полностью извлечены из исходного документа, а могут и не быть полностью использованы [7]. Цель автоматического обобщения состоит в том, чтобы информация, содержащаяся в исходном документе, автоматически извлекалась без существенной потери смысла [8].

Эффективным методом нормализации справочников номенклатур является парсинг. Парсинг позволяет автоматически извлекать и извлекать и структурировать информацию из различных источников, что облегчает ее использование и обновление. Парсинг также помогает уменьшить ошибки и несовместимости, возникающие при ручном вводе данных. Это особенно важно для справочников, которые требуют точной и актуальной информа-

ции, чтобы обеспечить эффективное управление и принятие решений.

Парсинг данных как эффективный инструмент обработки информации обеспечивает автоматизацию извлечения и структурирования информации, повышение точности и актуальности данных, повышение эффективности управления и принятия решений, снижение трудозатрат:

Таким образом, нормализация НСИ — это важный процесс для эффективного управления и манипулирования данными. Это обеспечивает упорядоченность и согласованность данных, повышая их качество, эффективность и надежность. Используя соответствующие методы нормализации, мы можем полностью раскрыть потенциал нормативно-справочной информации, что позволяет улучшить их анализ, принятие решений и интеграцию с другими системами.

Задача нормализации номенклатуры

Основной целью нормализации является приведение Записей Справочника к стандартизованному виду, поиск и выделение дубликатов, неполных, некорректных, противоречивых и неактуальных Записей с последующим принятием решения: приведению к стандартизованному виду, удалению или блокировке Записи [11, 12, 13, 14, 15].

Задачами Нормализации являются:

- унификация наименований номенклатуры;
- устранение ошибок, избыточности и дублирования справочных данных;
- обеспечение единообразия за счет использования единых стандартов классификации.

Нормализация Записей Справочника «Номенклатура» включает следующие шаги.

- выявление неполных, некорректных, неактуальных Записей, с последующей блокировкой или удалением;
- нормализация актуальных и востребованных в деятельности НКЗ Записей в соответствии с требованиями и под-

ходами, представленными в данной методике;

- поиск и выделение дублирующих записей после проведения нормализации;
- устранение дублей путем блокировки или удаления Записей;
- загрузка результатов Нормализации

в ИС (SAP ERP и SRM);

- проверка и тестирование корректности и полноты внесенных изменений в Записи Справочника.

Рассмотрим пример Справочника Номенклатуры теплоэнергетической компании (см. рис 2.).

Материал	Краткое наименование	Полное наименование
481353	АВТОБУС ПАЗ-4234-04	АВТОБУС ПАЗ-4234-04
109420	АВТОГЕРМЕТИК 80Г ТУ 2384-001-27-858	АВТОГЕРМЕТИК 80Г ТУ 2384-001-27-858
484542	АВТОГИДРОПОДЪЕМНИК АГП18.01 НА ШАССИ ЗИЛ	АВТОГИДРОПОДЪЕМНИК АГП18.01 НА ШАССИ ЗИЛ
425967	АВТОГИДРОПОДЪЕМНИК ВИПО-18-01	АВТОГИДРОПОДЪЕМНИК ВИПО-18-01
443528	АВТОГИДРОПОДЪЕМНИК НА ШАССИ ГАЗ-3309	АВТОГИДРОПОДЪЕМНИК НА ШАССИ ГАЗ-3309
443691	АВТОГРЕЙДЕР ГС-10.01	АВТОГРЕЙДЕР ГС-10.01
185755	АВТОДОВОДЧИК Н-М №5S 605-160	АВТОДОВОДЧИК Н-М №5S 605-160
425281	АВТОЗАРЯДКА В ПРИКУРИВАТЕЛЬ MINI USB	АВТОЗАРЯДКА В ПРИКУРИВАТЕЛЬ MINI USB
96741	АВТОКРАН ИВАНОВЕЦ Г/П 25Т	АВТОКРАН ИВАНОВЕЦ Г/П 25Т
443544	АВТОКРАН УРАЛ 5557	АВТОКРАН УРАЛ 5557
443545	АВТОКРАН УРАЛ КС 35714-1	АВТОКРАН УРАЛ КС 35714-1
515859	АВТОМАГНИТОЛА ACV AVS-930BR	АВТОМАГНИТОЛА ACV AVS-930BR
415400	АВТОМАТ ELCS 253 С 63 3-ПОЛЮС. 63А	АВТОМАТ ELCS 253 С 63 3-ПОЛЮС. 63А
252898	АВТОМАТ БЕЗОПАСНОСТИ БТ-169643-2 ТУРБИНЫ	АВТОМАТ БЕЗОПАСНОСТИ БТ-169643-2 ТУРБИНЫ
175212	АВТОМАТ ГАЗВОДЫ АВ-2М-900	АВТОМАТ ГАЗВОДЫ АВ-2М-900
175176	АВТОМАТ ГАЗВОДЫ АСБН-4	АВТОМАТ ГАЗВОДЫ АСБН-4
443692	АВТОМАТ ГАЗВОДЫ САТУРАТОР АВ-3	АВТОМАТ ГАЗВОДЫ САТУРАТОР АВ-3
415740	АВТОМАТ ГАШЕНИЯ ПОЛЯ АГП-30-42	АВТОМАТ ГАШЕНИЯ ПОЛЯ АГП-30-42
457836	АВТОМАТ ГОРЕНИЯ SIMENS LFL 1.333	АВТОМАТ ГОРЕНИЯ SIMENS LFL 1.333
420632	АВТОМАТ ГОРЕНИЯ SIMENS LME 22.331C2	АВТОМАТ ГОРЕНИЯ SIMENS LME 22.331C2
381911	АВТОМАТ ЗАЩИТЫ ЭЛ.ДВИГАТЕЛЯ MMS-63S 63А	АВТОМАТ ЗАЩИТЫ ЭЛ.ДВИГАТЕЛЯ MMS-63S 63А
343125	АВТОМОБИЛЬ GENESIS G80	АВТОМОБИЛЬ GENESIS G80
458964	АВТОШИНА 165/70 R13 ВА3-2108/09	АВТОШИНА 165/70 R13 ВА3-2108/09
493171	АВТОШИНА 17.5-25 Ф-120 БЕЛШИНА	АВТОШИНА 17.5-25 Ф-120 БЕЛШИНА
120781	АВТОШИНА 175 R16С БИ-522	АВТОШИНА 175 R16С БИ-522
457925	АВТОШИНА 175/65 R13 ВА3 ROSAVA	АВТОШИНА 175/65 R13 ВА3 ROSAVA
457926	АВТОШИНА 175/65 R14 АМТЕЛ ШИП.	АВТОШИНА 175/65 R14 АМТЕЛ ШИП.
266162	АВТОШИНА 175/65 R14 CONTINENTAL	АВТОШИНА 175/65 R14 CONTINENTAL

Рис. 2. Ненормализованный справочник Номенклатуры [9,16]

В результате нормализации данных приходим к следующему (см. рис. 3):

МТР	Краткое наименование материала (До нормализации)	Полное наименование материала (До нормализации)	Краткое наименование материала (После нормализации)	Полное наименование материала (После нормализации)
425840	МОНИТОР 31.5 LG 32QN600-B	МОНИТОР 31.5 LG 32QN600-B	МОНИТОР LG 32QN600-B 31,5"	МОНИТОР LG 32QN600-B; ДИАГОНАЛЬ ЭКРАНА 31,5ДЮЙМ; РАЗРЕШЕНИЕ МАКСИМАЛЬНОЕ 2560X1440; ЧАСТОТА КАДРОВ ДО 75ГЦ; ИНТЕРФЕЙСЫ DP,HDMI
436896	МОНИТОР 23.8 MSI PRO MP241	МОНИТОР 23.8 MSI PRO MP241	МОНИТОР MSI PRO MP241 23,8"	МОНИТОР MSI PRO MP241; ДИАГОНАЛЬ ЭКРАНА 23,8ДЮЙМ; РАЗРЕШЕНИЕ МАКСИМАЛЬНОЕ 1920X1080; ЧАСТОТА КАДРОВ ДО 60ГЦ; ИНТЕРФЕЙСЫ DP,HDMI,VGA
437417	МОНИТОР 23.8 ACER V247YBIP	МОНИТОР 23.8 ACER V247YBIP	МОНИТОР ACER V247YBIP 23,8"	МОНИТОР ACER V247YBIP; ДИАГОНАЛЬ ЭКРАНА 23,8ДЮЙМ; РАЗРЕШЕНИЕ МАКСИМАЛЬНОЕ 1920X1080; ЧАСТОТА КАДРОВ ДО 75ГЦ; ИНТЕРФЕЙСЫ DP,HDMI,VGA
438733	МОНИТОР 23.8 PHILIPS 241E2FD	МОНИТОР 23.8 PHILIPS 241E2FD	МОНИТОР PHILIPS 241E2FD 23,8"	МОНИТОР PHILIPS 241E2FD; ДИАГОНАЛЬ ЭКРАНА 23,8ДЮЙМ; РАЗРЕШЕНИЕ МАКСИМАЛЬНОЕ 1920X1080; ЧАСТОТА КАДРОВ ДО 75ГЦ; ИНТЕРФЕЙСЫ HDMI,DVI-D,VGA
439029	МОНИТОР 23.8 AOC 24B2XD	МОНИТОР 23.8 AOC 24B2XD	МОНИТОР AOC 24B2XD 23,8"	МОНИТОР AOC 24B2XD; ДИАГОНАЛЬ ЭКРАНА 23,8ДЮЙМ; РАЗРЕШЕНИЕ МАКСИМАЛЬНОЕ 1920X1080; ЧАСТОТА КАДРОВ ДО 75ГЦ; ИНТЕРФЕЙСЫ DVI-D,VGA

Рис. 3. Сравнение Ненормализованных и нормализованных данных [9,16]

Чтобы прийти к нормализованным записям справочника требуется найти признаки каждой записи (характеристики номенклатуры). Вручную данный процесс занимает в среднем 5–7 минут на 1 запись. Записей номенклатуры у различных Компаний отличается, начинается от 20 тыс. до нескольких миллионов записей.

Про нормализацию НСИ написано множество статей, исследований и презентаций.

Виноградов К. А., Никитина М. И., Жучков Д. В. в своей статье приводят описание качественных справочников, пример их структуры и подчеркивают важность единой комплексной модели справочника для целей достижения наиболее полно и системно предоставляемой нормативно-справочной информации. [17]

Наугольнова И. А. указывает на необходимость применения параметрических рядов. По мнению автора, это позволяет уменьшить себестоимость эксплуатации, сокращается номенклатура запасных частей и комплектующих изделий, повышается коэффициент загрузки и коэффициент полезного действия. [18]

Е.Антропов [10] отмечает, что качество данных напрямую зависит от качества управления нормативно-справочной информацией. Отмечает ряд типовых проблем в управлении НСИ, таких как:

- Организационные:
- отсутствие регламентов ведения справочников;
- несогласованность и конфликты интересов в ведении справочников;
- отсутствие центра компетенций или ответственного Подразделение за качество НСИ.
- Методологические:
- отсутствие методологии ведения справочников;
- несвоевременная актуализация справочников;
- отсутствие методики нормализации, как основополагающего инструмента документа, по актуализации справочников.
- Технологические:

- отсутствие возможности централизации справочников в единой информационной системе (справочники ведутся разрозненно в нескольких информационных системах);
- отсутствие интеграционных потоков между децентрализованными источниками (базами данных/хранения) Справочников.

Указанный перечень типовых проблем не является исчерпывающий.

Одним из инструментов и способов качественного управления справочниками является внедрение MDM-систем [5].

Рассмотрим применение парсинга для автоматизации поиска данных для целей нормализации с сайта [Электронный ресурс: <https://www.dns-shop.ru>].

В связи с тем, что сайт является динамическим, для начала требуется сформировать список всех товаров <https://www.dns-shop.ru/sitemap.xml>. Далее используем указанный код парсинга:

```
import pickle
import sys
from openpyxl import Workbook
from openpyxl.styles import Alignment,
Border, Font, Side
from tqdm import tqdm
from random import randint
from datetime import datetime
from time import sleep as pause
from bs4 import BeautifulSoup
import undetected_chromedriver as uc
def parse_characteristics_page(driver, url):
    «««» Парсит страницу товара по ссылке.»»»
    driver.get(url)
    pause(randint(7, 11))
    soup = BeautifulSoup(driver.page_source,
'html')
    #print(soup.prettify())
    name = soup.find('div', class_='»product-card-
description__title»)
    price = soup.find('div', class_='»product-
buy__price»)
    desc = soup.find('div', class_='»product-card-
description-text»)
    avail = soup.find('a', class_='»order-avail-
```

```

wrap__link ui-link ui-link_blue»)
    chars = soup.find_all('div', class_='»product-
characteristics__spec-title»)
    cvalue = soup.find_all('div', class_='»product-
characteristics__spec-value»)
    main_picture = soup.find('img',
class_='»product-images-slider__main-img»)
    pictures_soup = soup.find_all('img',
class_='»product-images-slider__img loaded tns-
complete»)
    pictures_list = []
    for i in pictures_soup:
        _ = pictures_list.append(i.get('data-src'))
    if _ is not None:
        pictures_list.append(_)
    span_tags = soup.find_all('span')
    for i in span_tags:
        if bool(str(i).find('data-go-back-
catalog')!=1):
            category = i
            tech_spec = {}
            for f1, f2 in zip(chars, cvalue):
                tech_spec[f1.text.rstrip().lstrip()] = f2.text.
rstrip().lstrip()
            notebook = {}
            notebook[«Категория»] = category.text.
lstrip(': ')
            notebook[«Наименование»] = name.text[15:]
            notebook[«Цена»] = int(price.text.replace(' ',
'')[:-1])
            notebook[«Доступность»] = avail.text if avail
is not None else 'Товара нет в наличии'
            notebook[«Ссылка на товар»] = url
            notebook[«Описание»] = desc.text
            notebook[«Главное изображение»] = main_
picture.get('src') if main_picture is not None else 'У
товара нет картинок'
            notebook[«Лист с картинками»] = pictures_
list
            notebook[«Характеристики»] = list(tech_
spec.items())
            # for i, j in notebook.items():
            # print(i, j)
            return notebook
    def get_all_category_page_urls(driver, url_to_
parse):
        «»» Получаем URL категории и парсим
ссылки с неё.»»»

```

```

page = 1
url = url_to_parse.format(page=page)
driver.get(url=url)
pause(10)
soup = BeautifulSoup(driver.page_source,
'html')
span_tags = soup.find_all('span')
for i in span_tags:
    if bool(str(i).find('data-role=»items-
count»')!=1):
        number_of_pages = [int(x) for x in str(i) if
x.isdigit()]
        res = int(''.join(map(str, number_of_pages)))
        pages_total = ((res // 18) + 1)
        print(f'Всего в категории {pages_total} стра-
ницы')
        urls = []
        while True:
            page_urls = get_urls_from_page(driver)
            urls += page_urls
            if page >= pages_total:
                break
            page += 1
            url = url_to_parse.format(page=page)
            driver.get(url)
            pause(randint(6, 9))

        return urls
    def get_urls_from_page(driver):
        «»» Собирает все ссылки на текущей стра-
нице.»»»
        soup = BeautifulSoup(driver.page_source,
'html')
        elements = soup.find_all('a', class_='»catalog-
product__name ui-link ui-link_black»)
        return list(map(
            lambda element: 'https://www.dns-shop.ru' +
element.get('href') + 'characteristics/',
            elements
        ))
    def to_excel(data, file_name=»table»):
        workbook = Workbook()
        sheet = workbook.active
        print('=' * 20)
        print('Начался экспорт в Excel Таблицу')
        column_names = [
            «Категория»,
            «Наименование»,

```

```

        «Цена»,
        «Доступность»,
        «Ссылка на товар»,
        «Описание»,
        «Главное изображение»,
        «Лист с картинками»,
        «Характеристики»,
    ]
    side = Side(border_style='thin')
    border = Border(
        left=side,
        right=side,
        top=side,
        bottom=side
    )
    alignment = Alignment(
        horizontal='center',
        vertical='center'
    )
    for column, name in enumerate(column_
names, 1):
        cell = sheet.cell(
            column=column,
            row=1,
            value=name
        )
        cell.font = Font(bold=True)
        cell.border = border
        cell.alignment = alignment
        counter = 1
        for index, value in enumerate(data, 2):
            for i in value.values():
                cell = sheet.cell(
                    column=counter,
                    row=index,
                    value=str(i) if type(i) == list else i
                )
                cell.alignment = Alignment(horizontal='left')
                counter += 1
            counter = 1
            for i in 'ABCDEFGHI':
                sheet.column_dimensions[i].width = 30
                workbook.save(f» {file_name} {datetime.
now().strftime('%d.%m.%y %H-%M-%S')}.xlsx»)
    def main():
        driver = uc.Chrome()
        urls_to_parse = [
            'https://www.dns-shop.ru/catalog/recipe/

```

```

e585499db2f27251/demontaz/?p={page}',
            'https://www.dns-shop.ru/
catalog/17a89bb916404e77/platy-
rasshireniya/?p={page}',
            'https://www.dns-shop.ru/catalog/
c8a984d0ba7f4e77/radiosistemy/?p={page}',
            'https://www.dns-shop.ru/
catalog/2c0f47131ade2231/aksessuary-dlya-
materinskix-plat/?p={page}',
            'https://www.dns-shop.ru/
catalog/17a89b8416404e77/karty-
videozaxvata/?p={page}',
        ]
        urls = []
        for index, url in enumerate(urls_to_parse):
            print(f'Получение списка всех ссылок
из {index+1} категории:')
            parsed_url = get_all_category_page_
urls(driver, url)
            urls.append(parsed_url)
            print(«Запись всех ссылок в файл url.txt:»)
            with open('urls.txt', 'w') as file:
                for url in urls:
                    for link in url:
                        file.write(link + «\n»)
                    with open('urls.txt', 'r') as file:
                        urls = list(map(lambda line: line.strip(), file.
readlines()))
            print(urls)
            info_dump = []
            for url in tqdm(urls, ncols=70, unit='товаров',
colour='blue', file=sys.stdout):
                info_dump.append(parse_characteristics_
page(driver, url))
                with open('dump_list_pickle.txt', 'wb+') as
file:
                    pickle.dump(info_dump, file)
                    with open('dump_list_pickle.txt', 'rb') as file:
                        info_dump = pickle.load(file)
                    to_excel(info_dump, file_name=»info_dump»)
            if __name__ == '__main__':
                main()
            print('=' * 20)
            print('Готово!')

```

В результате получим следующую информацию о номенклатуре:

- Категорию

- Наименование
- Цена
- доступен или нет к продаже
- Ссылка страницы с товаром
- Ссылку на главное изображение
- Ссылки на все изображения
- Характеристики
- Описание

Использование указанного инструмента позволяет снизить вероятность ошибки, вызванной человеческим фактором и сократить длительность процесса Нормализации данных.

Вывод

В настоящее время работа с информацией имеет большое значение для повышения эффективности деятельности промышленной компании. Хорошо выстроенная система обработки информации позволяет получать снижение затрат на хранение избыточных ресурсов на предприятии, более быстрое удовлетворение спроса покупателей, повышение выручки предприятия. Грамотное структурирование обрабатываемой информации позволяет повысить качество анализа данных на предприятии и как следствие, результативность работы предприятия.

Литература

1. Lihui Wang. From intelligence science to intelligent manufacturing. // Engineering, 5 (2019), pp. 615-618, <https://doi.org/10.1016/j.eng.2019.04.011>
2. Lu Y., Xu X., Wang L. Smart manufacturing process and system automation—a critical review of the standards and envisioned scenarios // Journal of Manufacturing Systems. — 2020. — Т. 56. — С. 312–325. <https://doi.org/10.1016/j.jmsy.2020.06.010>
3. Madushan Jayasekara, Chamoth, FranPyCisco 2022: Network Automation & Abstraction Solutions To Simplify Configuration Complexity (January 1, 2022). Available at SSRN: <https://ssrn.com/abstract=4176096> or <http://dx.doi.org/10.2139/ssrn.4176096>
4. Телишев А. М. Централизация единого информационного пространства инженерных справочников с применением средств автоматизации в объединениях производственных предприятий (холдингах, концернах) // Новые технологии, материалы и оборудование российской авиакосмической отрасли — АКТО-2016 // Сборник докладов Всероссийской научно-практической конференции с международным участием: в 2-х томах. Том 2. 2016 // Издательство: Академия наук Республики Татарстан (Казань)
5. Вечканова Ю. С., Федосин С. А. Алгоритмы машинного обучения в управлении нормативно-справочной информацией (НСИ). Издательство: Национальный исследовательский Мордовский государственный университет им. Н. П. Огарёва (Саранск) — 2021.
6. J. Pooley, V. Huang Multi-national patent litigation: management of discovery and settlement issues and the role of the judiciary Fordham Intellectual Property Media Entertain. Law J., 22 (2011), pp. 45–68, <https://heinonline.org/HOL/LandingPage?handle=hein.journals/frdipm22&div=5&id=&page=> (дата обращения 30.02.2023).
7. A.J.C. Trappey, C. V. Trappey An R&D knowledge management method for patent document summarization Ind. Manage. Data Syst., 108 (2) (2008), pp. 245–257
8. Trappey A. J. C. et al. Intelligent compilation of patent summaries using machine learning and natural language processing techniques //Advanced Engineering Informatics. — 2020. — Т. 43. — С. 101027. <https://doi.org/10.1016/j.aei.2019.101027> (дата обращения 30.02.2023).
9. Информация Министерства энергетики РФ от 17 апреля 2020 г. «Сформирован перечень системообразующих организаций российской экономики в части ТЭК» [Электронный ресурс] Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/73819140/>, (дата обращения 30.02.2023).
10. Евгений Антропов. Анализ и нормализация справочников IBS. // Электронный

- ресурс (2021) // URL: <https://ibs-business.ru/upload/iblock/120/xс733x0b3rtjrq4bqeiapkt5nvzsss1/Analiz-i-Normalizatsiya-short.pdf?ysclid=lrk6d3m95y866346718>
11. Про нормализацию НСИ простыми словами / Алексей Голиков. // Электронный ресурс // URL: <https://vc.ru/services/163097-pro-normalizaciyu-nsi-prostymi-slovami?ysclid=lrk6b1xzq6557572091> (дата обращения 30.02.2023)
 12. Чипура П., Шафорост Ф. Номенклатура как решение по нормализации НСИ. // Доклад. — 2018. — Режим доступа: https://static.1c.ru/bf/2018/p/02/16_45_%201С%20Номенклатура%20как%20решение%20по%20нормализации%20НСИ%20%28Чипура%20П,%20Шафорост%20Ф%29.pdf, свободный (дата обращения 30.02.2023).
 13. Гостев, С. Анализ и нормализация справочников IBS // Habr. — 2019. — Режим доступа: <https://habr.com/ru/articles/421047/>, свободный (дата обращения 30.02.2023).
 14. Применение аналогов в производстве в 1С ERP // Электронный ресурс Devgram. — 2020. // Режим доступа: <https://devgram.ru/blog/rabotav1s/primenenie-analogov-v-proizvodstve-v-1s-erp/?ysclid=lm0l87lnq7229557873>, свободный (дата обращения 30.02.2023).
 15. Ильин, А. В. Как устроен и как работает сервер в 1С // Электронный ресурс — 2019 // Режим доступа: <https://habr.com/ru/articles/439866/>, свободный (дата обращения 30.02.2023).
 16. Справочники номенклатур теплоэнергетической компании // Портал cntd-reglament.ru // Электронный ресурс // URL: https://cntd-reglament.ru/Nti_exp_teploenergetika_all.php (дата обращения 30.02.2023).
 17. Виноградов К. А., Никитина М. И., Жучков Д. В. Система ведения базы данных нормативно-справочной информации // Комплексная компьютеризация ЛПУ. 2004. № 3. С. 161–169.
 18. Наугольнова И. А. Параметрическая стандартизация для обоснования номенклатуры производства при процессном подходе к управлению затратами на про-

мышленном предприятии [Электронный ресурс] // И. А. Наугольнова. — DOI 10.24412/cl-36008-2021-2-156-161. — Текст: электронный // Новые импульсы развития: вопросы научных исследований: X Международная научно-практическая конференция. — 2021. — № 2. — С. 156–161. — URL: <https://cyberleninka.ru/article/n/parametricheskaya-standartizatsiya-dlya-obosnovaniya-nomenklatury-proizvodstva-pri-protsessnom-podhode-k-upravleniyu-zatratami-na>.

NORMALIZATION OF NORMATIVE AND REFERENCE INFORMATION USING THE EXAMPLE OF A REFERENCE BOOK OF NOMENCLATURES

Urintsov A. I., Aleshina A. V., Mustafin A. M.

Lomonosov Moscow State University, Plekhanov Russian University of Economics.

High-quality reference books of regulatory and reference information are the key to success and high-quality business processes. High-quality (normalized) directories allow businesses to see accurate information about companies and make effective management decisions. Normalization is a tool that allows you to obtain high-quality reference books, effectively manage master data, and make effective management decisions. There is currently no generally accepted and effective way to carry out normalization. Different manufacturers and experts in the said field use different tools and solutions. Normalization is carried out using at least one tool — MS Excel, by searching and updating information on the Internet, regulatory and technical documentation and documentation of suppliers or manufacturers' nomenclatures (if we are talking about the Nomenclature directory). This method is not the most effective and involves a high risk of errors caused by human factors. To improve the efficiency of normalization, the following tools should be used: parsing and machine learning.

Keywords: data normalization, nomenclature, reference books, parsing.

References

1. Lihui Wang. From intelligence science to intelligent manufacturing. // Engineering, 5 (2019), pp. 615-618, <https://doi.org/10.1016/j.eng.2019.04.011>
2. Lu Y., Xu X., Wang L. Smart manufacturing process and system automation—a critical review of the standards and envisioned scenarios // Journal of Manufacturing Systems. — 2020. — Vol. 56. — pp. 312–325. <https://doi.org/10.1016/j.jmsy.2020.06.010>
3. Madushan Jayasekara, Chamoth, FranPyCisco 2022: Network Automation & Abstraction Solutions To Simplify Configuration Complexity (January 1, 2022). Available at SSRN: <https://ssrn.com/abstract=4176096> or <http://>

- dx.doi.org/10.2139/ssrn.4176096
4. Telishev A. M. Centralization of a single information space of engineering reference books using automation tools in associations of manufacturing enterprises (holdings, concerns) // New technologies, materials and equipment of the Russian aerospace industry – ACTO-2016 // Collection of reports of the All-Russian scientific and practical conference with international participation: in 2 volumes. Volume 2. 2016 // Publisher: Academy of Sciences of the Republic of Tatarstan (Kazan).
 5. Vechkanova Yu. S., Fedosin S. A. Machine learning algorithms in the management of normative and reference information (NSI). Publisher: National Research Mordovian State University named after N. P. Ogarev (Saransk) – 2021.
 6. J. Pooley, V. Huang Multi-national patent litigation: management of discovery and settlement issues and the role of the judiciary Fordham Intellectual Property Media Entertain. Law J., 22 (2011), pp. 45–68, <https://heinonline.org/HOL/LandingPage?handle=hein.journals/frdipm22&div=5&id=&page=> (accessed 30.02.2023).
 7. A.J.C. Trappey, C.V. Trappey An R&D knowledge management method for patent document summarization Ind. Manage. Data Syst., 108 (2) (2008), pp. 245–257
 8. Trappey A. J. C. et al. Intelligent compilation of patent summaries using machine learning and natural language processing techniques //Advanced Engineering Informatics. – 2020. – Vol. 43. – p. 101027. <https://doi.org/10.1016/j.aei.2019.101027> (date of appeal 30.02.2023).
 9. Information from the Ministry of Energy of the Russian Federation dated April 17, 2020. «A list of system-forming organizations of the Russian economy in the fuel and energy sector has been formed» [Electronic resource] Access mode: <https://www.garant.ru/products/ipo/prime/doc/73819140/>, (accessed 30.02.2023).
 10. Evgeny Antropov. Analysis and normalization of IBS reference books. // Electronic resource (2021) // URL: <https://ibs-business.ru/upload/iblock/120/x0b3rtjjrq4bqeiapkt5nvzsss1/Analiz-i-Normalizatsiya-short.pdf?ysclid=lrk6d3m95y866346718> (accessed 30.02.2023).
 11. Golikov Alexey. About the normalization of NSI in simple words // Electronic resource (01.10.2020) // URL: <https://vc.ru/services/163097-pro-normalizaciyu-nsi-prostymi-slovami?ysclid=lrk6b1xzq6557572091> (accessed 30.02.2023).
 12. Chipura P., Shaforost F. Nomenclature as a solution for the normalization of NSI. // Report. – 2018. – Access mode: https://static.1c.ru/bf/2018/p/02/16_45_%201C%20Номенклатура%20как%20решение%20по%20нормализации%20НСИ%20%28Чипура%20П,%20Шафорост%20Ф%29.pdf (accessed 30.02.2023).
 13. Gostev, S. Analysis and normalization of IBS reference books // Habr. – 2019. – Access mode: <https://habr.com/ru/articles/421047> (accessed 30.03.2023).
 14. The use of analogues in production in 1C ERP // Electronic resource Devgram. – 2020. // Access mode: <https://devgram.ru/blog/rabotav1s/primenenie-analogov-v-proizvodstve-v-1s-erp/?ysclid=lm0l87lnq7229557873> (accessed 30.02.2023).
 15. Ilyin, A. V. How the server works in 1C // Electronic resource – 2019 // Access mode: <https://habr.com/ru/articles/439866/>, free (accessed 30.02.2023).
 16. Reference books of the nomenclature of the thermal power company // Portal cntd-reglament.ru // Electronic resource // URL: https://cntd-reglament.ru/Nti_exp_teploenergetika_all.php
 17. Vinogradov K. A., Nikitina M. I., Zhuchkov D. V. The system of maintaining a database of normative and reference information // Complex computerization of medical institutions. 2004. No. 3. pp. 161–169.
 18. Nagunova I. A. Parametric standardization for substantiating the nomenclature of production with a process approach to cost management at an industrial enterprise [Electronic resource] // I. A. Nagunova. – DOI 10.24412/cl-36008–2021–2–156–161. – Text: electronic // New impulses of development: issues of scientific research: X International scientific and practical conference. – 2021. – No. 2. – pp. 156–161. – URL: <https://cyberleninka.ru/article/n/parametriceskaya-standardizatsiya-dlya-obosnovaniya-nomenklatury-proizvodstva-pri-protsessnom-podhode-k-upravleniyu-zatratami-na>.

Пути повышения конкурентоспособности авиатранспортной отрасли Российской Федерации в условиях санкционных ограничений

Дадов Тамерлан Азаматович

Экономический факультет МГУ имени М. В. Ломоносова,
магистр

специалист отдела пассажирских услуг,

Международный аэропорт Нальчик

E-mail: dadovtamerlan@mail.ru

В настоящее время высокую политическую и экономическую значимость приобретает повышение конкурентоспособности авиатранспортной отрасли России с учетом сложившихся санкционных ограничений. Авиатранспортная отрасль имеет высокую значимость в сфере организации транспортной доступности регионов, перевозке пассажиров и грузов. В статье рассматриваются основные вызовы и проблемы развития авиатранспортной отрасли Российской Федерации в условиях санкционных ограничений и соответствующие меры повышения уровня конкурентоспособности российских авиакомпаний на основе активной государственной поддержки. Освещены утвержденные на государственном уровне планы в области развития импортозамещения зарубежных воздушных судов, оптимизации маршрутной сети и развития альтернативных рынков пассажирских и грузовых авиаперевозок.

Ключевые слова: авиатранспортная отрасль, пассажирские авиаперевозки, грузовые авиаперевозки, импортозамещение, государственная поддержка авиатранспортной отрасли.

Введение

В 2022 году российская авиатранспортная отрасль столкнулась с беспрецедентным санкционным давлением в виде целого ряда различных ограничений, которые создали множество новых отраслевых вызовов, некоторые из которых имеют критическое значение для всего транспортного комплекса Российской Федерации. По данным Федерального агентства воздушного транспорта (Росавиации) в 2022 году сразу девять российских авиакомпаний были вынуждены остановить свою деятельность ввиду существенного снижения уровня пассажирских и грузовых перевозок [2]. У ряда российских региональных авиакомпаний были приостановлены сертификаты эксплуатантов воздушных судов — основного разрешительного документа для выполнения коммерческих воздушных авиаперевозок. У одной из девяти авиакомпаний (резидента ОЭЗ в Калининградской области компании авиакомпании «СКОЛ») сертификат был полностью аннулирован, что привело к уходу компании с рынка. Стоит отметить, что в предыдущие 5 лет Росавиация приостанавливала и аннулировала сертификаты 5 компаний, что в целом свидетельствует о наличии негативных тенденций в российской авиационной отрасли и том, что введенные в отношении отрасли санкции оказывают непосредственное негативное влияние, в первую очередь, на уровень конкурентоспособности средних и малых российских авиационных компаний.

Внешние санкционные ограничения и закрытие ряда популярных авиационных маршрутов существенно сказались на объ-

еме перевозок через российские аэропорты (рисунок 1).

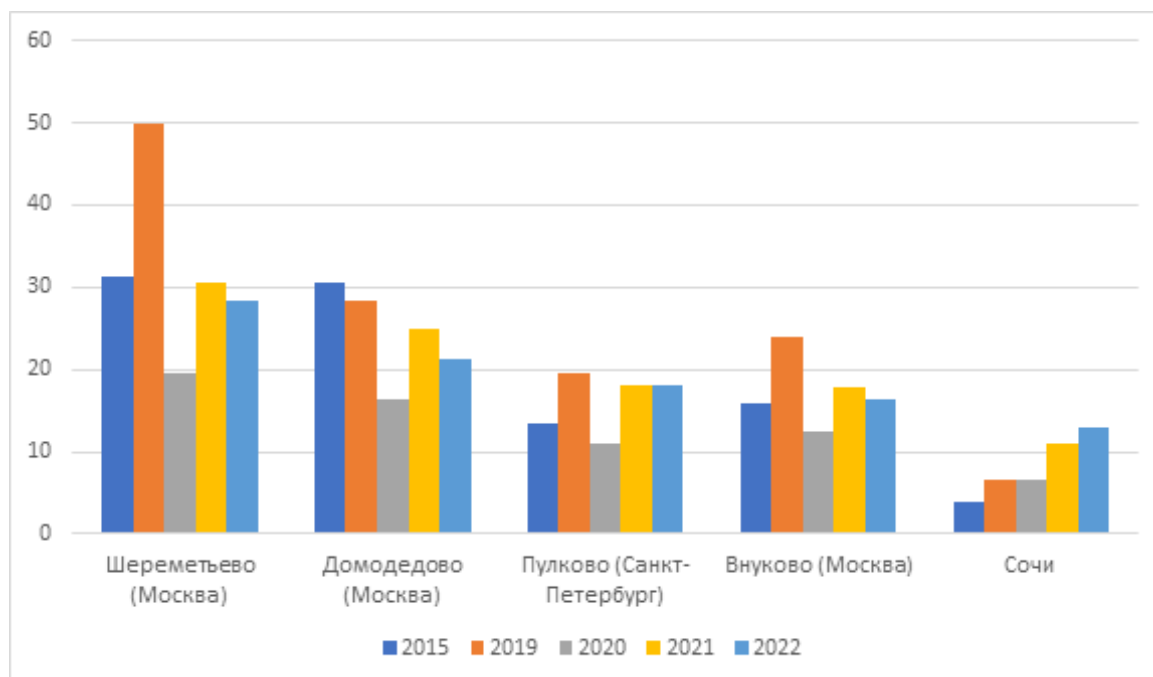


Рис. 1. Динамика пассажиропотока крупнейших аэропортов России за 2015–2022 гг. [2]

Как видно из графика (рис. 1), с 2015 года вплоть до пандемии 2020 г. наблюдался стремительный рост пассажиропотока – некоторые аэропорты демонстрировали двукратный рост числа обслуживаемых пассажиров, что в целом свидетельствовало о стабильных темпах развития авиатранспортной отрасли Российской Федерации как по внутренним, так и по международным авиаперевозкам. С февраля 2022 г. Росавиация перестала публиковать консолидированные данные об объемах обслуживания пассажиров в целях обеспечения национальной безопасности, но на основании данных информационных порталов было выявлено, что для большинства ключевых российских авиационных узлов динамика в 2022 г. была отрицательной.

Стратегия развития авиатранспортной отрасли Российской Федерации

В принятой в 2022 году распоряжением Правительства Российской Федерации № 1693-р Стратегии развития авиатранспортной отрасли

на период до 2030 года [1] отмечается, что российский авиатранспортный рынок характеризуется высоким уровнем концентрации авиационных перевозок, где ведущую роль занимают крупные рыночные игроки – группа компаний «Аэрофлот», S7 (АО «Авиакомпания «Сибирь»), «Уральские Авиалинии», «ЮТэйр»), общая доля авиаперевозок которых превышает 70%, а лидирующую позицию на протяжении последних лет неизменно занимает группа компаний «Аэрофлот», средневзвешенная доля рынка которой в последние годы составляет на уровне 43%.

В Стратегии подчеркивается, что в сложившихся условиях для российской авиатранспортной отрасли сформировались следующие вызовы и барьеры развития:

- рост потерь парка воздушных судов в связи с отсутствием возможности технического обслуживания самолетов зарубежного производства (в первую очередь среднемагистральных и дальнемагистральных пассажирских самолетов компаний Boeing и Airbus);

- недостаточно высокие темпы импортозамещения зарубежных воздушных судов техникой отечественного производства;
- рост финансовых рисков для авиакомпаний ввиду сокращения международного сообщения, ведущий к повышению вероятности банкротства, потери квалифицированных кадров;
- снижение уровня безопасности авиаперевозок;
- деградация маршрутной сети.

Эти и смежные риски могут стать причиной ухудшения динамики развития российской транспортной отрасли, снизить уровень транспортной связности регионов Российской Федерации, негативно отразиться на уровне мобильности населения, что, в свою очередь, может привести к снижению общих темпов социально-экономического развития государства.

Повышение конкурентоспособности российской авиатранспортной отрасли представляется в настоящий момент комплексной стратегической задачей, решение которой предполагает непосредственное участие государства в следующих областях:

- создание всех необходимых условий для интенсификации производства и последующего ввода в эксплуатацию среднемагистральных и дальнемагистральных воздушных судов отечественного производства;
- обеспечение необходимых темпов обновления парка воздушных судов российских авиакомпаний, имеющих статус системообразующих организаций транспортного комплекса Российской Федерации;
- – обеспечение стабильной финансовой поддержки ведущих российских авиакомпаний ввиду роста их финансовой нагрузки в кризисный период;
- – оптимизация сети воздушных перевозок с учетом геополитической ситуации;
- развитие альтернативных рынков пассажирских и грузовых авиаперевозок;

- развитие сотрудничества с дружественными государствами в авиатранспортной сфере.

В этой связи в сложившихся условиях наиболее очевидным стратегическим решением со стороны органов государственной власти является разработка и последующая поэтапная реализация комплексных мер поддержки ведущих российских авиакомпаний и предприятий авиационной отрасли, занимающихся разработкой и производством отечественных воздушных судов, комплектующих и компонентов, программного обеспечения, включая аэронавигационные базы данных и автоматизированные системы управления.

Меры государственной поддержки авиатранспортной отрасли Российской Федерации

В течение 2022 года был реализован комплекс мер государственной поддержки российской авиационной отрасли с целью обеспечения стабильности и устойчивости отечественной гражданской авиации. Главным инструментом поддержки являлось государственное субсидирование российских авиакомпаний в части возмещения операционных затрат на выполнение внутренних и международных авиаперевозок. В полном объеме были возмещены финансовые потери, связанные с отменой ряда международных маршрутов в связи с закрытием воздушного пространства зарубежных стран для российских авиакомпаний. Инструмент субсидирования так же использовался для поддержки наземной инфраструктуры (аэропортов) в большинстве регионов Российской Федерации и поддержки организации воздушного движения. Общая сумма государственной поддержки авиационной отрасли в 2022 году составила более 172 млрд рублей [4], что позволило сохранить необходимый уровень доступности авиационных перевозок, минимизировать потери ведущих российских авиакомпаний и обеспечить необходимый уровень транспортной связности субъектов Российской Федерации.

Стоит отметить, что внешние санкционные ограничения имеют определенный положительный эффект — они выступили катализатором роста темпов комплексного реформирования авиатранспортной отрасли Российской Федерации, обнажив ключевые ее проблемы, к которым можно отнести высокий уровень зависимости российских авиакомпаний от зарубежных воздушных судов. На протяжении 2022 года велась активная работа по переводу регистрации использующихся воздушных судов зарубежного производства в Государственный реестр гражданских судов Российской Федерации с целью поддержания необходимого уровня авиаперевозок по внутренним и доступным внешним маршрутам. В дополнение к этому активизировалась работа в области обеспечения внутрироссийского технического сопровождения самолетов (в т. ч. самолетов зарубежного производства), что позволило создать необходимые условия для поддержания летной годности всех судов российских авиакомпаний, прошедших перерегистрацию. Ярким примером реализации стратегических инициатив стало получение расширенного сертификата для осуществления периодического и оперативного технического сопровождения воздушных судов российской компанией «А-Техникс» (входит в группу компаний «Аэрофлот»), которая на протяжении последних лет ведет работу по обеспечению летной годности воздушных судов Boeing и Airbus.

Параллельно с этим велась активная работа по обеспечению необходимых мер поддержки российской авиастроительной отрасли с целью реализации мероприятий по интенсификации ввода в эксплуатацию воздушных судов отечественного производства. В конце 2022 года ПАО «Корпорация «Иркут» получила ряд технических одобрений по внесению необходимых изменений в типовую конструкцию среднемагистрального пассажирского самолета нового поколения MC-21, который в будущем станет основным среднемагистральным воздушным судном российских авиакомпаний и не уступает по своим характеристикам признанным лидерам рынка.

В феврале 2023 года Правительство Российской Федерации закончило формирование паспортов инвестиционных проектов для обновления парка воздушных судов за счет средств Фонда национального благосостояния (ФНБ) [6]. Этот шаг ориентирован на финансовую поддержку реализации контрактов с российскими предприятиями авиационной промышленности с целью обновления парка воздушных судов, а главным выгодоприобретателем станет флагман российской авиационной отрасли — группа компаний «Аэрофлот», которая на льготных условиях получит в свое распоряжение порядка 60 магистральных самолетов отечественного производства (MC-21, Superjet, TU-214). Общая сумма финансовой поддержки отечественной авиационной отрасли в 2023 году составит порядка 237 млрд рублей (в 2024–2025 г. на эти проекты будет выделены дополнительные 32 млрд рублей) [3].

Активная работа ведется и в области развития альтернативных рынков международных авиаперевозок. С октября 2022 года ряд российских авиакомпаний возобновил прямые рейсы в Таиланд. Прямые рейсы из Российской Федерации доступны на начало 2023 г. в Турцию, Алжир, Индию, Египет, Иран, Катар, Китай, Кубу, ОАЭ, Оман, Тунис и ряд других стран. В настоящий момент российские авиакомпании осуществляют полеты в 22 государства, но уже в ближайшие несколько месяцев этот перечень может быть расширен за счет возобновления авиасообщения с рядом государств Юго-Восточной Азии (Вьетнам, Лаос) [7]. Вместе с тем ситуация с развитием альтернативных направлений международных авиаперевозок остается нестабильной, так как со стороны западных государств ведется активная работа по сдерживанию российских авиакомпаний на международном рынке. В частности, Министерство торговли США на официальном уровне озвучивало возможные санкции для Турции в случае продолжения поддержки российских авиакомпаний, осуществляющих прямые рейсы на самолетах американского производства (Boeing). Но при этом российские туристы занимают значительную долю в общем объеме туристского потока в Турции, для которой данная отрасль является системно значимой.

Выводы

По результатам проведенного исследования обеспечение необходимых темпов восстановления основных показателей развития авиатранспортной отрасли и создание базовых основ для повышения ее конкурентоспособности в долгосрочном периоде может быть обеспечено за счет следующих направлений:

1) интенсификации усилий в области реализации необходимых инициатив по замещению зарубежных воздушных судов;

2) дальнейшего развития технологической базы и укрепления уровня технологического суверенитета для создания прорывных технологий, используемых в авиатранспортной отрасли;

3) активного развития сотрудничества Российской Федерации в дружественными зарубежными государствами в области международных пассажирских и грузовых авиаперевозок на взаимовыгодных условиях.

Литература

1. Распоряжение Правительства Российской Федерации от 25.06.2022 г. № 1693-р
2. Данные портала Федерального агентства воздушного транспорта Российской Федерации (Росавиации) [Электронный ресурс] // URL: <https://favt.gov.ru/dejatelnost-ajeroporty-i-ajerodromy-stat-dannye/> (дата обращения 15.02.2023)
3. Интервью А. В. Нерадько о результатах работы авиационной отрасли в 2022 году и планах на 2023 год [Электронный ресурс]. // Материалы всероссийской транспортной еженедельной информационно-аналитической газеты «Транспорт России» от 08.02.2023 // URL: <https://transportrussia.ru/razdely/vozdushnyj-transport/9706-iz-zony-sanktsionnoj-turbulentnosti-bez-poter.html> (дата обращения 15.02.2023)
4. Стенограмма заседания Правительства Российской Федерации от 16.02.2023 [Электронный ресурс] // Портал Правительства Российской Федерации // URL:

<http://government.ru/news/47804/> (дата обращения 16.02.2023)

5. Абдулина А. Минус девять за бортом. Авиарынок недосчитался игроков в 2022 году [Электронный ресурс] // Портал Коммерсант.ру от 20.01.2023 // URL: <https://www.kommersant.ru/doc/5774914> (дата обращения 15.02.2023)
6. Правительство сформировало паспорт инвестпроекта по лизингу водного и авиатранспорта // Портал Финмаркет.ру от 16.02.2023 [Электронный ресурс]. URL: <http://www.finmarket.ru/news/5898831> (дата обращения 16.02.2023)
7. Росавиация работает над возобновлением прямых рейсов во Вьетнам [Электронный ресурс]. // Портал Турдом.ру от 14.02.2023 URL: <https://www.tourdom.ru/news/rosaviatsiya-rabotaet-nad-vozobnovleniem-priamykh-reysov-vo-vietnam-.html> (дата обращения 15.02.2023)

WAYS TO INCREASE THE COMPETITIVENESS OF THE AIR TRANSPORT INDUSTRY OF THE RUSSIAN FEDERATION UNDER SANCTIONAL RESTRICTIONS

Dadov T. A.

Lomonosov Moscow state University

Currently, increasing the competitiveness of the Russian air transport industry is gaining high political and economic importance, taking into account the prevailing sanctions restrictions. The air transport industry is of high importance in the field of organizing transport accessibility of regions, transportation of passengers and cargo. The article discusses the main challenges and problems of the development of the air transport industry of the Russian Federation in the context of sanctions restrictions and appropriate measures to increase the competitiveness of Russian airlines based on active state support. The plans approved at the state level in the field of development of import substitution of foreign aircraft, optimization of the route network and development of alternative markets for passenger and cargo air transportation are highlighted.

Keywords: air transport industry, passenger air transportation, air cargo transportation, import substitution, state support for the air transport industry.

References

1. Decree of the Government of the Russian Federation dated June 25, 2022 No. 1693-r
2. Data from the portal of the Federal Air Transport Agency of the Russian Federation (Rosaviatsia) [Electronic resource] // URL: <https://fvt.gov.ru/deyatelnost-ajeroporty-i-ajerodromy-stat-dannye/> (accessed 15.02.2023)
3. Interview with A. V. Neradko on the results of the aviation industry in 2022 and plans for 2023 [Electronic resource]. // Materials of the all-Russian transport weekly information and analytical newspaper «Transport of Russia» dated February 8, 2023 // URL: <https://transportrussia.ru/razdely/vozdushnyj-transport/9706-iz-zony-sanktsionnoj-turbulentnosti-bez-poter.html> (accessed 02/15/2023)
4. Transcript of the meeting of the Government of the Russian Federation dated February 16, 2023 [Electronic resource] // Portal of the Government of the Russian Federation // URL: <http://government.ru/news/47804/> (accessed February 16, 2023)
5. Abdulina A. Minus nine overboard. The aviation market missed the players in 2022 [Electronic resource] // Kommersant.ru portal of 01/20/2023 // URL: <https://www.kommersant.ru/doc/5774914> (accessed 02/15/2023)
6. The government has formed a passport for an investment project for the leasing of water and air transport // Portal Finmarket.ru dated February 16, 2023 [Electronic resource]. URL: <http://www.finmarket.ru/news/5898831> (accessed 02.16.2023)
7. The Federal Air Transport Agency is working on the resumption of direct flights to Vietnam [Electronic resource]. // Portal Turdom.ru dated February 14, 2023 URL: <https://www.tourdom.ru/news/rosaviatsiya-rabotaet-nad-vozobnovleniem-priamykh-reysov-vo-vietnam-.html> (accessed February 15, 2023)

Создание HTML файла из RSS-ленты: автоматизация обработки торговых лотов с использованием Python

Ермошин Сергей Алексеевич

магистр, РЭУ имени Г. В. Плеханова

E-mail: serega-arr@yandex.ru

Булгаков Андрей Леонидович

к. э. н., профессор МИСАО,

доцент РЭУ имени Г. В. Плеханова,

с.н.с. МГУ имени М.В.Ломоносова

E-mail: z3900207@mail.ru

Щапова Ксения Дмитриевна

студент, экономический факультет МГУ

имени М. В. Ломоносова

E-mail: ksenya_shapova@mail.ru

В настоящее время информация играет все большее значение в мире. Скорость доступа к информации, скорость ее обработки, обработка в корректной форме, чтобы некорректно введенные данные не искажали результаты исследования, существенным образом влияет на принимаемые экономические решения. Быстрый доступ к актуальным данным может быть решающим фактором в принятии управленческих и инвестиционных решений. Обработка больших данных позволяет повысить эффективность принимаемых управленческих решений. При обработке больших данных возникает ряд вопросов: нахождение источников данных, корректная их обработка, аналитика данных. Особое значение для принятия торговых и инвестиционных решений при совершении финансовых операций представляет обработка биржевой информации. Сервисы с использованием технологии RSS предоставляют отличный способ отслеживать обновления на различных платформах, таких как торговые лоты. В данной статье будет рассмотрен пример кода на Python, который автоматизирует получение данных из RSS-ленты и создает HTML-файл для удобного просмотра.

Ключевые слова: Python, автоматизация обработки данных, RSS-ленты, искусственный интеллект, анализ больших данных.

В современном мире финансовых технологий автоматизация обработки данных в режиме реального времени играет ключевую роль в принятии оперативных решений и управлении рисками. Особую актуальность эта тема приобретает в контексте обработки биржевой информации, где скорость и точность анализа данных напрямую влияют на эффективность торговых стратегий и финансовые результаты.

Исследователи в области финансовых технологий активно разрабатывают и совершенствуют методы автоматизации обработки финансовой информации. В статье «Real-time Financial Data Processing Using Cloud Computing» Johnson и Smith предлагают использовать облачные технологии для масштабирования вычислительных мощностей при обработке больших объемов биржевых данных. Авторы утверждают: «Cloud-based solutions provide unprecedented flexibility and cost-efficiency for real-time financial data processing, allowing for rapid scaling of computational resources as market conditions change» [1]. Этот подход особенно актуален для организаций, сталкивающихся с пиковыми нагрузками во время высокой рыночной волатильности.

Значимость обработки финансовой информации подчеркивается многими авторами [2]. Получаемая информация с сайта биржи показывает, как текущие настройки на рынке, ожидания инвесторов, состояние макроэкономической среды. Чтобы совершать взвешенные инвестиционные или торговые решения необходимо иметь доступ к биржевой информации и в оперативном режиме, обрабатывать эту информацию.

Возможности получения финансовой информации из новостной RSS-ленты

Одним из наиболее доступных форматов получения новостей являются RSS — ленты. Ежегодно их количество преумножается, как и количество содержащихся в них файлов. Необходимо оптимизировать время на фильтрацию новостей.

Из официальной документации библиотеки «Feedparser» [3]: наиболее часто используемыми элементами в RSS-каналах (независимо от версии) являются заголовок, ссылка, описание, дата публикации и идентификатор записи. Дата публикации происходит из элемента pubDate, а идентификатор записи — из элемента guid.

Universal Feed Parser (универсальный анализатор каналов) прост в использовании; у него есть одна основная общедоступная функция — синтаксический анализ. parse принимает несколько аргументов, но требуется только один, и это может быть URL-адрес, имя локального файла или необработанная строка, содержащая данные канала в любом формате.

Из официальной документации библиотеки «Request» [4]: Python Requests — это библиотека, которая создана для быстрой и Requests работы с запросами. Стандартные HTTP-библиотеки Python, например та же Urllib3, часто требуют значительно больше кода для выполнения одного и того же действия, а это затрудняет работу.

В статье «Технологические аспекты построения системы сбора и предобработки корпусов новостных текстов для создания моделей языка» [5] авторы предлагают реализовать обработку новостных статей с помощью программной системы сбора. В дальнейшем предполагается создание языковых моделей. Источниками информации выступают RSS-ленты и новостные сайты. Ход работы включает: сбор ссылок на актуальную информацию, скачивание новостей из интернета, хранение, управление и выделения текстов из html-документов.

Созданная система сбора и обработки информации помогает эффективно собирать основы новостных статей. Также система даёт возможность структурировать информацию по дате, теме, источнику и мету происхождения, для создания более точных языковых моделей с опорой на актуальные новостные данные.

По мнению авторов для новостных моделей больше подходит метод, построенный на собранных корпусах, так как является наиболее точным.

В статье «Исследование правильности и эффективности средств парсинга информации на веб-ресурсах» [6] авторы выбрали двадцать инструментов для парсинга и описали их характеристики, в том числе методы, разработанные в рамках исследования: на основе регулярных выражений и строковых функций языка.

Сравнительная оценка происходила в рамках стандарта ГОСТ, и включала в себя следующие параметры:

1. Функциональные возможности;
2. Эффективность (время работы и занимаемая память);
3. Адаптируемость (возможность переноса на другое окружение).

Иерархальные средства парсинга тратят меньше времени на обработку документа, что удобно для многократного поиска по документу.

Потоковые средства парсинга используют последовательный перебор, поэтому пригодны для работы, где не требуется повторное обращение. Подходят для индексации и преобразовании формата данных.

Метод регулярных выражений подходит, когда необходимо достать конкретный элемент документа, структура которого известна.

Шаги обработки

Для проведения сбора анализа данных из новостной ленты RSS необходимо совершить следующую последовательность действий.

1. Установка библиотек

Для этого проекта используются библиотеки 'feedparser' и 'requests'. Первая позволяет легко разбирать данные из RSS-лент, а вторая — выполнять HTTP-запросы.

```
"""python
import feedparser
import requests
"""
```

2. Получение RSS-ленты

Начнем с написания функции 'parse_rss', которая принимает URL RSS-ленты и параметры запроса. Эта функция выполняет запрос, получает RSS-ленту и возвращает объект 'FeedParserDict'.

```
"""python
def parse_rss(url, params):
    response = requests.get(url, params=params,
headers={'User-Agent': 'Mozilla/5.0'}, verify=True)
    if response.status_code == 200:
        feed = feedparser.parse(response.text)
        return feed
    else:
        print(f'Не удалось получить RSS-ленту. Код
состояния: {response.status_code}')
        return None
"""
```

3. Генерация HTML

Следующая функция, 'generate_html', принимает объект RSS-ленты и имя файла HTML. Она создает HTML-файл, представляя данные из RSS-ленты в виде таблицы.

```
"""python
def generate_html(feed, html_filename):
    if feed is not None:
        with open(html_filename, 'w',
encoding='utf-8') as htmlfile:
            htmlfile.write(«<!DOCTYPE html>\n<html>\n
<head>\n<title>RSS Feed Results</title>\n</head>\n
```

```
n<body>\n»)
            htmlfile.write(f«<p>Всего элементов:
{len(feed.entries)}</p>»)
            htmlfile.write(«<table border='1'
cellpadding='5' style='border-collapse: collapse;'>\n
<tr>»)
            for field in ['title', 'published', 'link',
«summary»]:
                htmlfile.write(f«<th>{field.capitalize()}</
th>»)
            htmlfile.write(«</tr>»)
            for entry in feed.entries:
                htmlfile.write(«<tr>»)
                for field in ['title', 'published»]:
                    value = entry.get(field, «»)
                    htmlfile.write(f«<td>{value}</td>»)

                link_value = entry.get(«link», «»)
                htmlfile.write(f«<td><a href=' {link_value}'
target='_blank'>{link_value}</a></td>»)
                summary_value = entry.get(«summary», «»)
                htmlfile.write(f«<td>{summary_value}</td>»)
                htmlfile.write(«</tr>»)
            htmlfile.write(«</table>\n</body>\n</html>»)
            print(f'HTML-результаты сохранены в файл
{html_filename}')
        else:
            print(«Ошибка при парсинге RSS-ленты.»)
    """
```

4. Запуск программы

В блоке 'if __name__ == «__main__»:' определены параметры запроса для RSS-ленты, URL и имя файла HTML. Производится вызов функций 'parse_rss' и 'generate_html' для создания HTML-файла.

```
"""python
if __name__ == «__main__»:
    rss_url = «https://torgi.gov.ru/new/api/public/
lotcards/rss»

    params = {
        'dynSubjRF': '78',
        'biddType': «»,
        'biddForm': «»,
        'currCode': «»,
```

```

        'lotStatus': 'PUBLISHED, APPLICATIONS_
SUBMISSION',
        'biddEndFrom': "",
        'biddEndTo': "",
        'pubFrom': "",
        'pubTo': "",
        'aucStartFrom': "",
        'aucStartTo': "",
        'etpCode': "",
        'catCode': '100',
        'text': "",
        'noticeStatus': "",
        'amoOrgCode': "",
        'resourceTypeUse': "",
        'npa': "",
        'byFirstVersion': 'true'
    }
    html_filename = «feed_results.html»
    feed = parse_rss(rss_url, params)
    if feed:
        generate_html(feed, html_filename)
    ""

```

Выводы

Полученные результаты показывают возможность использования средств автоматизации для обработки финансовой информации. Полученный код демонстрирует, как использовать Python для автоматизации получения данных из RSS-ленты и их представления в виде HTML-файла. Можно адаптировать этот пример для своих нужд, изменяя параметры запроса и структуру HTML-файла.

Благодаря этому примеру можно создавать автоматизированные отчеты по обработке биржевой информации, отслеживать изменения в данных и интегрировать этот процесс в свои рабочие потоки при принятии инвестиционных решений.

Литература

1. Johnson, M., & Smith, K. Real-time Financial Data Processing Using Cloud Computing. // Journal of Financial Technology. — 2022 -

15(3), 287–301.

2. Carone, T. Future Automation: Changes to Lives and to Businesses by // World Scientific Publishing Co, Inc., 2018 // URL: https://doi.org/10.1142/9789813142343_0008
3. Documentation // Портал feedparser 6.0.11. // Электронный ресурс // URL: <https://feedparser.readthedocs.io/en/latest/>
4. Requests: HTTP for Humans // Портал requests.readthedocs.io // Электронный ресурс // URL: <https://requests.readthedocs.io/en/latest/>
5. Белозеров Андрей Александрович, Вахлаков Дмитрий Владимирович, Мельников Сергей Юрьевич, Пересыпкин Владимир Анатольевич, Сидоров Евгений Сергеевич. Технологические аспекты построения системы сбора и предобработки корпусов новостных текстов для создания моделей языка // Известия ЮФУ. Технические науки. № 12 (185). // Электронный ресурс (2016) // URL: <https://cyberleninka.ru/article/n/tehnologicheskie-aspekty-postroeniya-sistemy-sbora-i-predobrabotki-korpusov-novostnyh-tekstov-dlya-sozdaniya-modeley-yazyka>
6. Лыгина Н. И., Пудич А. С. Исследование правильности и эффективности средств парсинга информации на веб-ресурсах // Инновационная наука № 3–1. // Электронный ресурс (2017) // URL: <https://cyberleninka.ru/article/n/issledovanie-pravilnosti-i-effektivnosti-sredstv-parsinga-informatsii-na-veb-resursah>

CREATING HTML FROM RSS FEED: AUTOMATING TRADING LOT PROCESSING WITH PYTHON

Ermoshin S. A., Bulgakov A. L., Shchapova K. D.

MISAO, Plekhanov Russian University of Economics, Lomonosov MSU

Information is becoming increasingly important in the world today. The speed of access to information, the speed of its processing, processing it in the correct form so that incorrectly entered data does not distort the results of the study, significantly affecting the decisions made. Quick access to relevant data can be a decisive factor in management and management decisions. Big data processing Allows you to increase the efficiency of management decisions. When processing big data, a number of issues arise: finding data sources, their correct processing, data analysis. Of particular importance for making trading decisions and

making decisions when performing financial transactions is the processing of cryptocurrency information. Services using RSS technology provide an excellent way to track updates on various platforms, such as trading lots. This article will consider an example of Python code that automates the receipt of data from an RSS feed and creates an HTML file for easy viewing.

Keywords: Python, data processing automation, RSS feeds, artificial intelligence, big data analysis.

References

1. Johnson, M., & Smith, K. Real-time Financial Data Processing Using Cloud Computing. // Journal of Financial Technology. – 2022 -. 15(3), 287–301.
2. Carone, T. Future Automation: Changes to Lives and to Businesses by (World Scientific Publishing Co, Inc., 2018) // URL: https://doi.org/10.1142/9789813142343_0008
3. Documentation // Portal feedparser 6.0.11. // Electronic resource // URL: <https://feedparser.readthedocs.io/en/latest/>
4. Requests: HTTP for Humans // Portal requests.readthedocs.io // Electronic resource // URL: <https://requests.readthedocs.io/en/latest/>
5. Belozerv Andrey Aleksandrovich, Vakhlakov Dmitry Vladimirovich, Melnikov Sergey Yuryevich, Peresypkin Vladimir Anatolyevich, Sidorov Evgeny Sergeevich. Technological aspects of building a system for collecting and preprocessing news text corpora to create language models // Bulletin of SFedU. Technical sciences. No. 12 (185). // Electronic resource (2016) // URL: <https://cyberleninka.ru/article/n/tehnologicheskie-aspekty-postroeniya-sistemy-sbora-i-predobrabotki-korpusov-novostnyh-tekstov-dlya-sozdaniya-modeley-yazyka>
6. Lygina N. I., Pudich A. S. Study of the correctness and efficiency of information parsing tools on web resources // Innovative science No. 3–1. // Electronic resource (2017) // URL: <https://cyberleninka.ru/article/n/issledovanie-pravilnosti-i-effektivnosti-sredstv-parsinga-informatsii-na-veb-resursah/viewer>

Актуальные проблемы банкротства физических лиц

Воронкова Светлана Николаевна

Аспирант экономического факультета Воронежского государственного лесотехнического университета
E-mail: 9651021720@mail.ru

Алешина Анна Валентиновна

к. э. н., доцент экономический факультет МГУ имени М. В. Ломоносова
E-mail: annaaleshina@mail.ru

Ильин Илья Иванович

Инженер, ООО «Институт экспериментальной экономики и финансов МГУ имени М.В.Ломоносова»
E-mail: ilia_ilin1991@mail.ru

Булгакова Татьяна Николаевна

Аналитик ООО «Институт экспериментальной экономики и финансов МГУ имени М.В.Ломоносова»
E-mail: tate-888@yandex.ru

Овсян Степан Андраникович

Студент экономического факультета МГУ имени М.В.Ломоносова
E-mail: stepan-stepan-2014@bk.ru

Статья посвящена выявленным на практике и в научных работах проблемам банкротства физических лиц. В настоящее время данный вопрос очень актуален в силу нестабильности экономики и увеличения числа неплатежеспособных граждан. Большое внимание исследователей и практиков приковано к критериям определения неплатежеспособности. Анализ судебной практики позволил сделать вывод о том, что в разных делах наличие или отсутствие тех или иных критериев, свидетельствующих о неплатежеспособности, приводили к кардинально разным решениям. Кроме того, актуальной проблемой является определение стоимости имущества должника финансовым управляющим — зачастую у него нет достаточной квалификации для справедливой и качественной оценки, а также достаточной доступных сведений для определения всего актива, имеющегося у должника. В работе приведен ряд критических позиций и путей решения имеющихся проблем.

Ключевые слова: банкротство, неплатежеспособность, банкротство физического лица, Федресурс.

Нестабильность мировой и отечественной экономической ситуации, вызванной множеством внутренних и внешнеполитических процессов, стала причиной неплатежеспособности населения страны. В связи с невозможностью граждан выполнить свои долговые обязательства перед кредиторами, существует механизм защиты граждан и организаций, участвующих в экономическом обороте — это институт банкротства.

По своей сути несостоятельность (банкротство) является процедурой, в которой лицо не имеет возможности исполнить свои финансовые обязательства перед другими лицами. Согласно ст. 2 Федерального закона от 26.10.2002 № 127-ФЗ «О несостоятельности (банкротстве)» (далее — ФЗ № 127-ФЗ), под несостоятельностью (банкротством) (далее — банкротство) понимается признанная арбитражным судом или наступившая в результате завершения процедуры внесудебного банкротства гражданина неспособность должника в полном объеме удовлетворить требования кредиторов по денежным обязательствам, о выплате выходных пособий и (или) об оплате труда лиц, работающих или работавших по трудовому договору, и (или) исполнить обязанность по уплате обязательных платежей.

Институт банкротства появился в России относительно недавно, с введения в действие ФЗ № 127-ФЗ, вступившего в силу с 01.10.2015 года. Государство вкладывало во вновь введенную процедуру следующий смысл — «банкротство должно стать инструментом финансового оздоровления, возмож-

ность освобождения от долгового бремени и восстановления кредитоспособности населения» [2]. Однако на сегодняшний день правоприменительная практика и научные позиции указывают на недостаточную отлаженность механизма банкротства физических лиц, при этом количество неплатежеспособных граждан стремительно растет. Так, по статистике, к 1 апреля 2022 года число граждан, признанных несостоятельными составило

более 537 тыс. человек. За первые 3 месяца 2022 года 54 190 граждан были признаны банкротами, что на 33,6% больше, чем за тот же период 2021 года [3].

Согласно статистике государственного портала по банкротству Федресурс [11] количество граждан-банкротов в 2022 году выросло на 44,2% по сравнению с 2021 годом до 278.137 человек, при этом в 2021 году рост был на 62% до 192.833 человек (см. таблицу 1).

Таблица 1. Количество сообщений о введении судами процедур в отношении физических лиц и индивидуальных предпринимателей

	Окт- дек 2015	2016	2017	2018	2019	2020	2021	2022	Окт 2015-дек 2022	2022 к 2021, прирост
О признании гражданина банкротом и введении реализации имущества гражданина	870	19574	29827	43984	68980	119045	192833	278137	753250	44,4%
О признании обоснованным заявления о признании гражданина банкротом и введении реструктуризации его долгов	338	7730	7998	9749	15768	24292	33860	42415	142150	25,3%
Об утверждении плана реструктуризации долгов гражданина	5	165	188	179	204	209	329	461	1740	40,1%

Источник: [11, стр. 4].

За период существования процедуры потребительского банкротства, с октября 2015 года по декабрь 2022 года, несостоятельными стали уже 753250 граждан.

В большинстве случаев процедур банкротств граждан, граждане сами инициируют собственное банкротство: в 2022 году практически 96,7% случаев (94,9% в 2021). Доля конкурсных кредиторов в качестве заявителей составила 2,2% (была 3,9%), ФНС России – 1,0% (была 1,1%) [11]

Очевиден рост почти по всем показателям судебного банкротства. Примерно 66% дел возбуждено при долгах от 500 000 до 3 000 000 рублей. Меньшие суммы задолженно-

сти заявлены только в четверти обращений. В 2022 году рассмотрено 53 дела, где размер требований кредиторов составит свыше 1 млрд руб.

Хочется отметить, что не все судебные дела завершились освобождением от долгов. Исходя из статистики почти 2 000 человек получили отказ в списании задолженности [11].

В настоящий момент, вопрос существующих в отечественной практике проблем, связанных с банкротством физических лиц, стоит довольно остро. Осветим в данной работе некоторые из них.

Согласно ст. ст. 213.3, 223.2 ФЗ № 127-ФЗ, должники, претендующие на несостоятель-

ность, делятся на 2 категории в силу существования судебного (стандартного — через арбитражный суд) и внесудебного порядка (упрощенного — через МФЦ) признания гражданина банкротом. Критерием прохождения упрощенной процедуры банкротства является подача заявления гражданином без финансового управляющего в МФЦ, при условии, что сумма долга составляет от 25000 рублей до 1000000 рублей, в соответствии с положениями статьи 223.2 с указанием ряда оснований.

В одном из обзоров практики по делам о банкротстве отмечено, что внесудебная процедура имеет целью восстановление платежеспособности лица в случае возникновения у него сложной ситуации (потеря работы, доход ниже прожиточного минимума на каждого члена семьи) [4].

Несмотря на лояльный характер положений о судебной и внесудебной процедуре банкротства граждан в научной доктрине существуют позиции, что таким образом граждане ставятся в заведомо неравные условия. Более того, на практике наблюдается множество связанных с этими коллизиями. В основном коллизии сводятся к тому, что в судебной практике сложилась тенденция к признанию презумпции недобросовестности должника.

Довольно активно на практике и в науке обсуждаются критерии, по которым гражданин признается либо не признается банкротом — во-первых, неплатежеспособность, во-вторых недостаточность имущества. Согласно закону, необходимо использовать один из двух критериев или оба в совокупности. При этом нормы содержат в себе набор требований, которые должны присутствовать для признания несостоятельности, но конкретизирует методы разграничения критериев и условий их применения. Отечественный законодатель отошел от классической модели «неплатежеспособность-неоплатность» и пришел к более гибким методам [5].

Однако, на наш взгляд, переходя от законодательного регулирования к практике, можно увидеть множество правоприменительных

проблем, связанных с этими критериями. Так, в судебной практике отсутствует единообразный подход к применению критериев. В одном из решений суда гражданин был признан банкротом в силу наличия единственного критерия в виде «факта доказанности неспособности удовлетворения требований кредиторов» [6]. При этом множество судов берет за основу «критерии презумпции неплатежеспособности гражданина и исключения банкротства с учетом планируемых доходов гражданина» [7]. Приведенные примеры демонстрируют широкие границы судебного усмотрения в данной категории дел. Более того, в практике встречаются решения, где суды не приняли во внимание наличие у банкрота имущества, стоимостью выше объема денежных обязательств (по причине непринятия должником мер по его реализации), наличие стабильного соразмерного дохода [8], а также судебные решения с полностью противоположными решениями при тех же обстоятельствах.

Перейдем к другим актуальным проблемам, связанным с непосредственной процедурой банкротства. Одним из остро стоящих вопросов, вызывающих многочисленные споры, является вопрос оценки имущества. Согласно ФЗ № 127-ФЗ, ее осуществляет финансовый управляющий по своему письменному решению. Помимо этого, собрание кредиторов может привлечь к оценке реализуемого имущества профессионального оценщика. В случае его отсутствия в процессе оценки в решении финансового управляющего должен содержаться отчет, соответствующий правилам об оценочной деятельности и включающий в себя основные его элементы.

В большинстве случаев суды не признают определение стоимости верным при отсутствии в отчете непосредственно исследовательской части, содержащей методы оценки, данные о состоянии имущества. Полагаем, что при всей справедливости данных требований, нацеленных на справедливую и соразмерную оценку имущества, финансовый управляющий не является профессиональным оценщиком и соответственно не имеет достаточной для выполнения таких требований квалификации.

В силу отсутствия соответствующих знаний высокого уровня в процедуре оценки закон допускает возможность привлечения других лиц для оценки. Это возможно на основании определения суда, в случае если в ходатайстве будут приведены доказательства обоснованности привлечения и стоимости их услуг, а также при наличии согласия гражданина. Но хоть закон и предоставляет такую возможность, практика показывает, что такие ходатайства о привлечении профессионального оценщика нередко отклоняются судом по причине отнесения обязанности по определению стоимости имущества к компетенции финансового управляющего и достаточности открытых сведений, позволяющих произвести оценку.

Кроме того, даже в случае использования финансовым управляющим федеральных стандартов оценки к моменту реализации имущества на торгах, цена имущества может измениться. На наш взгляд, это повышает риски возникновения споров непосредственно на проводимых торгах. Именно поэтому, мы согласны с позицией многих ученых, отмечающих необходимость единообразного стандартного подхода к оценке.

Ряд проблем возникает и в вопросах получения полных и достоверных сведений об активах и обязательствах должника. Финансовый управляющий имеет право получать информацию об имуществе физического лица, заявившего о банкротстве, который в свою очередь обязан предоставлять все сведения о своем имуществе финансовому управляющему. При этом, в положениях закона есть ряд пробелов, связанных с этой обязанностью. Так, например, законодательно закреплена обязанность предоставлять сведения о сделках на сумму свыше 300 тыс. рублей, совершенных за последние 3 года, не охватывает безвозмездные сделки по передаче имущества (дарение), сделки по отчуждению долей по их номинальной стоимости и другие.

Также, ряд проблем связан с системами автоматизированного получения сведений. В практике имеется множество дел, где данные об имуществе должника, взятые финансовым

управляющим из Единого государственного реестра недвижимости (ЕГРН) были неактуальны ввиду отмены регистрации недвижимости и автотранспортных средств на должника. При этом о соответствующем судебном решении финансовый управляющий знать не мог. В свою очередь, поиски необходимых судебных решений на сайтах судов не дают результатов, так как не все они выкладываются своевременно. Представляется, что такое положение дел усложняет и делает громоздким весь процесс оценки «имущества должника и подтверждает необходимость создания единой картотеки сведений по судам общей юрисдикции с возможностью поиска процессов по личным данным конкретного гражданина, на что указано, в том числе, и в научной доктрине» [9].

Кроме того, существуют сложности при поиске полной информации об участии должника-банкрота в уставном капитале компаний или наличии у него в собственности акций юридических лиц. В Едином государственном реестре сведений о юридических лицах (ЕГРЮЛ) отсутствует возможность поиска сведений об участии конкретных физических лиц в капитале организаций. Только если существует информация, полученная из иных источников, о том, что должник-банкрот является учредителем ООО, то при проверке конкретного ООО в ЕГРЮЛ можно увидеть – является ли данное лицо учредителем ООО. Однако в ЕГРЮЛ невозможно такой поиск осуществить по фамилии должника-банкрота. Можно только вручную проверить конкретную компанию ООО, так как в сведениях в ЕГРЮЛ по ООО присутствует перечень всех участников ООО. Однако в информации в ЕГРЮЛ по акционерному обществу такая информация об учредителях отсутствует.

Проверить в каких компаниях должник-банкрот является учредителем или акционером можно только на платных информационно-аналитических порталах, таких как СПАРК.

Подводя итог, следует сказать, что несмотря на важную социальную функцию отечественного института банкротства физических лиц сегодня существует множество

проблем, связанных с данным вопросом. Основная их часть сводится к отсутствию в практике единообразного справедливого подхода к критериям определения несостоятельности, оценки имущества, а также действиям финансового управляющего. Представляется, что решение раскрытых в работе проблем кроется в конкретизации пределов судебного усмотрения в делах о банкротстве граждан, а также в более подробной регламентации статуса и полномочий финансового управляющего.

Литература

1. Илстинова М. Х., Одинцов С. В. Особенности оспаривания решений собрания кредиторов, «Имущественные отношения в Российской Федерации» // 2017, № 9, РФ, Москва, Издательство НОУ ВПО Международная академия оценки и консалтинга, стр. 34–52.
2. Ефременко А. В. Банкротство физических лиц в Российской Федерации: актуальные проблемы судебно-арбитражной практики // Вестник магистратуры. 2022. № 12–2 (135).
3. В России число граждан-банкротов превысило полмиллиона человек // [Электронный ресурс] Banki.ru от 08.04.2022 // URL: <https://www.banki.ru/news/lenta/?id=10964338>.
4. Федеральный закон от 26.10.2002 N 127-ФЗ (с изм. и доп.) «О несостоятельности (банкротстве)» // СПС «Консультант.Плюс».
5. Обзор практики по делам о банкротстве за июнь 2020 года (Филиал партнерства с ограниченной ответственностью Брайан Кейв Лейтон Пейзнер (Раша) ЛЛП в г. Москве, Практика по разрешению споров) // СПС «Консультант.Плюс».
6. Карелина С. А., Фролов И. В. Институт банкротства граждан по законодательству РФ (лекция в рамках учебного курса магистерской программы «Правовое регулирование несостоятельности (банкротства)» // Предпринимательское право. Приложение «Право и бизнес». 2017. N 4.
7. Постановление Девятого арбитражного апелляционного суда от 21.10.2016 N 09АП-47821/2016 по делу N А40–127048/16 // СПС «Консультант.Плюс».
8. Постановление Арбитражного суда Поволжского округа от 02.06.2020 N Ф06–59902/2020 по делу N А55–27676/2019; Постановление Арбитражного суда Центрального округа от 02.09.2020 N Ф10–2395/2020 по делу N А14–3450/2019 // СПС «Консультант.Плюс».
9. Постановление Арбитражного суда Московского округа от 29.09.2020 N Ф05–14447/2020 по делу N А40–308202/2019 // СПС «Консультант.Плюс».
10. Кудинова М. С. Актуальные аспекты процедуры банкротства граждан // Имущественные отношения в Российской Федерации. 2019. N 1. С. 66–78.
11. Банкротства в России: итоги 2022 года. Статистический релиз Федресурса // Портал по банкротству Федресурс // Электронный ресурс // URL: <https://download.fedresurs.ru/news/Банкротства%20статрелиз%202022.pdf>

CURRENT ISSUES OF BANKRUPTCY OF INDIVIDUALS

Voronkova S. N., Aleshina A. V., Ilyin I. I., Bulgakova T. N., Ovsoyan S. A.

Voronezh State Forestry University, Lomonosov Moscow State University, LLC "Institute of Experimental Economics and Finance of Moscow State University named after M. V. Lomonosov", MISAO, Plekhanov Russian University of Economics

The article is devoted to the problems of bankruptcy of individuals identified in practice and in scientific works. At present, this issue is very relevant due to the instability of the economy and the increase in the number of insolvent citizens. Much attention of researchers and practitioners is focused on the criteria for determining insolvency. Analysis of judicial practice allowed us to conclude that in different cases the presence or absence of certain criteria indicating insolvency led to radically different decisions. In addition, a pressing problem is the determination of the value of the debtor's property by the financial manager – often he does not have sufficient qualifications for a fair and high-quality assessment, as well as sufficient available information to determine all the assets of the debtor. The work provides a number of critical positions and solutions to existing problems.

Keywords: bankruptcy, insolvency, bankruptcy of an individual, Fedresurs.

References

1. Ilistinova M.Kh., Odintsov S. V. Features of challenging decisions of the meeting of creditors, «Property relations in the Russian Federation» // 2017, No. 9, RF, Moscow, Publishing House of the NOU VPO International Academy of Appraisal and Consulting, pp. 34–52.
2. Efremenko A. V. Bankruptcy of individuals in the Russian Federation: current problems of judicial and arbitration practice» // Bulletin of the Magistracy. 2022. No. 12–2 (135).
3. In Russia, the number of bankrupt citizens has exceeded half a million people // [Electronic resource] Banki.ru dated 04/08/2022 // URL: <https://www.banki.ru/news/lenta/?id=10964338>.
4. Federal Law of 26.10.2002 N 127-FZ (as amended and supplemented) «On Insolvency (Bankruptcy)» // SPS «Consultant.Plus».
5. Review of practice in bankruptcy cases for June 2020 (Branch of the limited liability partnership Brian Cave Leighton Paisner (Russia) LLP in Moscow, Dispute Resolution Practice) // SPS «ConsultantPlus».
6. Karelina S. A., Frolov I. V. Institute of bankruptcy of citizens under the legislation of the Russian Federation (lecture as part of the course of the master's program «Legal regulation of insolvency (bankruptcy)») // Entrepreneurial law. Appendix «Law and Business». 2017. N 4.
7. Resolution of the Ninth Arbitration Court of Appeal dated 21.10.2016 N 09AP-47821/2016 in case N A40–127048/16 // SPS «ConsultantPlus».
8. Resolution of the Arbitration Court of the Volga District dated 02.06.2020 N F06–59902/2020 in case N A55–27676/2019; Resolution of the Arbitration Court of the Central District dated 02.09.2020 N F10–2395/2020 in case N A14–3450/2019 // SPS «ConsultantPlus».
9. Resolution of the Arbitration Court of the Moscow District dated September 29, 2020 No. F05–14447/2020 in case No. A40–308202/2019 // SPS «ConsultantPlus».
10. Kudinova M. S. Current aspects of the bankruptcy procedure of citizens // Property relations in the Russian Federation. 2019. No. 1. Pp. 66–78.
11. Bankruptcies in Russia: results of 2022. Statistical release of Fedresurs // Bankruptcy portal Fedresurs // Electronic resource // URL: <https://download.fedresurs.ru/news/Банкротства%20статресли%202022.pdf>