

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ВСЕРОССИЙСКАЯ ГРУППА ТЕОРИИ ИНФОРМАЦИИ ИЕЕЕ
АКАДЕМИЯ ИНФОРМАТИЗАЦИИ ОБРАЗОВАНИЯ
ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ООО «ОТКРЫТЫЕ РЕШЕНИЯ»
ОБЩЕСТВО «ЗНАНИЕ» РОССИИ
ПРИВОЛЖСКИЙ ДОМ ЗНАНИЙ

*XXII Международная
научно-техническая конференция*

**ПРОБЛЕМЫ ИНФОРМАТИКИ
В ОБРАЗОВАНИИ, УПРАВЛЕНИИ,
ЭКОНОМИКЕ И ТЕХНИКЕ**

Сборник статей

Декабрь 2022 г.

Пенза

УДК 004
ББК 32.81я43+74.263.2+65.050.2я43
П781

П781 **ПРОБЛЕМЫ ИНФОРМАТИКИ В ОБРАЗОВАНИИ,
УПРАВЛЕНИИ, ЭКОНОМИКЕ И ТЕХНИКЕ :**
сборник статей XXII Международной научно-технической
конференции. – Пенза: Приволжский Дом знаний, 2022. – 356 с.

ISBN 978-5-8356-1800-2
ISSN 2311-0406

Под редакцией *В.И. Горбаченко*, доктора технических наук,
профессора;
В.В. Дрождина, кандидата технических наук,
профессора

Информация об опубликованных статьях предоставлена в систему Рос-
сийского индекса научного цитирования (РИНЦ) по договору
№ 573-03/2014К от 18.03.2014.

ISBN 978-5-8356-1800-2
ISSN 2311-0406

© Пензенский государственный
университет, 2022
© АННМО «Приволжский Дом знаний», 2022

*XXII International
scientific and technical conference*

**PROBLEMS OF INFORMATICS
IN EDUCATION, MANAGEMENT,
ECONOMICS AND TECHNICS**

December, 2022

Penza

**«НЕРВ» – СИСТЕМА ИНТЕЛЛЕКТУАЛЬНОГО
АНАЛИЗА ПРОЦЕССОВ ОКАЗАНИЯ
ГОСУДАРСТВЕННЫХ УСЛУГ**

М.И. Кревский

**«NERVE» – THE SYSTEM OF INTELLECTUAL PROCESS
ANALYSIS FOR DIGITAL PUBLIC SERVICES**

M.I. Krevskiy

Аннотация. Разработан инструмент для анализа и мониторинга течения бизнес-процессов оказания государственных слуг. Применены технологии машинного обучения и анализа процессов. Система реализована на языке Python. Используется аналитиками портала mos.ru.

Ключевые слова: машинное обучение, анализ процессов, анализ графов, наука о данных.

Abstract. A process-mining tool for the workflow of digital public service delivery has been developed in Python. Machine learning and process analytics technologies were applied. The system is used by mos.ru portal analysts.

Key words: machine learning, process mining, graph analysis, data science.

Оказание государственных услуг преимущественно представляет собой сложный бизнес-процесс с большим количеством этапов. Целью разработки является создание системы интеллектуального анализа и мониторинга течения бизнес-процессов. Система реализована на языке Python. Используется аналитиками портала mos.ru.

На данный момент «Нерв» находится на стадии работающего прототипа и реализует следующий функционал:

автоматизированное создание запроса и выгрузка логов из базы данных государственных услуг;

расчет статистических показателей услуги и ее этапов;

визуализация суммарного графа оказания услуги за выбранный период

и графа оказания услуги в рамках единичного обращения пользователя;

отображение сравнительной диаграммы времени прохождения этапа в рамках оказания услуги разным пользователям;

проверка соответствия бизнес-процессов регламенту.

Алгоритм работы аналитика с инструментом «Нерв».

1. Выбор параметров набора данных для построения графа услуг:
 - a. интересующей услуги с вводом идентификатора или единого кода услуги;
 - b. временного отрезка для отображения;
 - c. параметров отображаемой на графе информации;
 - d. режима отображения только процессов оказания услуг, начавшихся в выбранный период времени.

2. На обобщенном графе оказания услуги (рис. 1), просуммированном по всем обращениям пользователей за период, с помощью фильтров, применения метрик и проверки соответствия бизнес-процессов регламенту аналитик ищет проблемные участки, нарушения регламента и «бутылочные горлышки», задерживающие процесс. После осмотра графа целиком следует выбрать проблемные ребра – переходы между этапами.

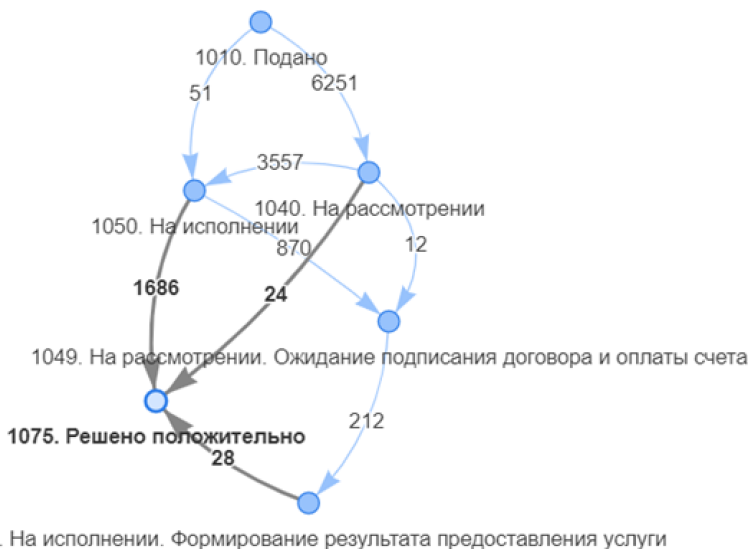


Рис. 1. Обобщенный граф оказания услуги

3. Переход между интересующими этапами выбирается в интерфейсе для построения диаграммы распределения времени переходов (рис. 2). На диаграмме аналитик видит упорядоченные по возрастанию временные отрезки выбранного перехода между этапами в рамках разных обращений пользователей. Каждая точка соответствует обращению пользователя, имеет

информацию об уникальном идентификаторе обращения и времени перехода.

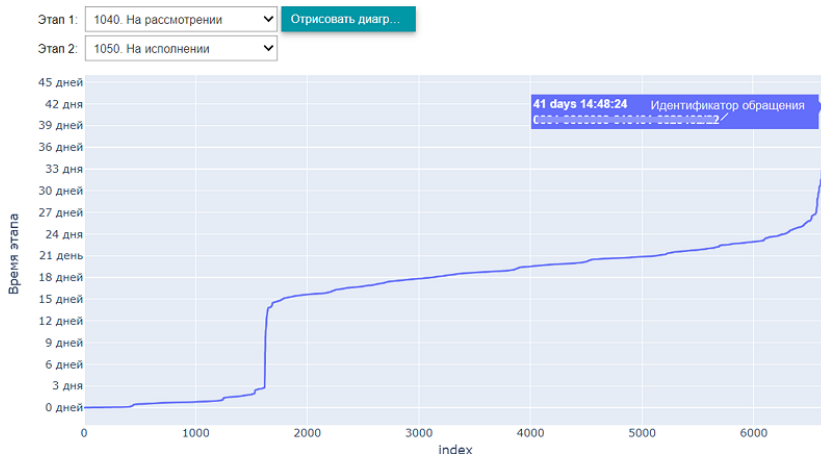


Рис. 2. Диаграмма распределения времени перехода между этапами

4. Ориентируясь на форму графика, аналитик выбирает несколько обращений так, чтобы они наглядно демонстрировали причину отклонений на обобщенном графе. Их идентификаторы вводятся в соответствующее поле интерфейса, после чего отдельно визуализируется граф оказания государственной услуги по обращению (рис. 3). На конкретных примерах пользователь системы может сделать вывод о природе проблемы в бизнес-процессе и направить информацию в орган исполнительной власти.

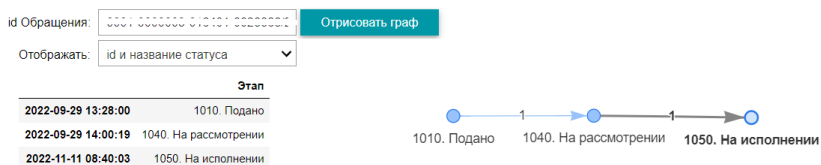


Рис. 3. Граф обращения пользователя

Реализована, но пока еще не встроена в общий проект, векторизация графа оказания услуги пользователю методом `Ttrace2vec` [1, 2] по аналогии с методом `Doc2vec` из сферы обработки естественных языков. Такие свойства данных логов, как последовательность с внутренней логикой

и правилами, конечное число разных видов этапов и объединение в группы (предложения), позволяют проследить аналогию с естественным языком. Этап соответствует слову, процесс оказания услуги в рамках одного обращения пользователя – предложению, а лог целиком – тексту. Полученные в результате векторизации эмбединги последовательностей этапов можно использовать при решении задач кластеризации и классификации. Качество эмбедингов позволяет достаточно эффективно разделять разные шаблоны оказания услуг и классифицировать, например, услуги, оказанные с нарушениями.

В разработке находится элемент «Нерва», позволяющий строить прогнозы дальнейшего прохождения бизнес-процесса оказания услуги. Также по аналогии с обработкой естественных языков предлагается применить рекуррентные нейронные сети для предсказания следующих этапов, возможных проблем и задержек при выполнении услуги и их вероятностей. Такой функционал позволит быстро и даже упреждающе реагировать на неблагоприятно складывающиеся обстоятельства оказания конкретных услуг, поможет эффективно распределить ресурсы и внимание органов исполнительной власти и добиться снижения процента услуг с нарушением регламента.

Несмотря на то, что на данный момент система «Нерв» дорабатывается, она активно используется отделом мониторинга и анализа государственных услуг ГКУ «Инфогород». Внедрение инструмента в работу отдела дало следующие преимущества:

позволило наглядно визуализировать поэтапную модель бизнес-процесса в виде ориентированного графа;

обеспечило быстрый доступ к статистике как по отдельным этапам, так и по услуге целиком;

позволило находить узкие места процесса и рассматривать их на примере конкретных случаев оказания услуг.

Библиографический список

1. Pieter De Koninck, SeppevandenBroucke& Jochen De Weerd. Act2vec, trace2vec, log2vec, and model2vec: Representation Learning for Business Processes. 16th International Conference on Business Process Management. Sydney, NSW, Australia, pp. 305–321, 2018. [Электронный ресурс]: SpringerLink, 2018. – Режим доступа: https://link.springer.com/chapter/10.1007/978-3-319-98648-7_18, свободный. – Загл. с экрана. – Яз. англ.

2. Кревский М.И. Построение продвинутых векторных моделей процессов: магистерская диссертация. – М.: МФТИ (НИУ), 2022. – 52 с.

УДК 519.23

**АНАЛИЗ АНСАМБЛЕВЫХ
МЕТОДОВ КЛАССИФИКАЦИИ
ДЛЯ ПРОГНОЗИРОВАНИЯ ПОСЛЕОПЕРАЦИОННЫХ
ОСЛОЖНЕНИЙ У БОЛЬНЫХ ЖЕЛЧНО-КАМЕННОЙ
БОЛЕЗНЬЮ**

Р. Н. Кузнецов, О. Ю. Кузнецова

**ANALYSIS OF ENSEMBLE CLASSIFICATION METHODS
TO PREDICT POSTOPERATIVE COMPLICATIONS
IN PATIENTS WITH CHOLELITHIASIS**

R. N. Kuznetsov, O. Yu. Kuznetsova

Аннотация. В этой работе описаны ансамблевые алгоритмы. Как правило, ансамблевые модели объединяют несколько базовых моделей для повышения производительности прогнозирования. Наиболее известным примером модели ансамбля является случайный лес, который, значительно упрощая логику алгоритма, объединяет несколько деревьев решений и агрегирует их прогнозы, используя большинство голосов.

Ключевые слова: прогнозирование послеоперационных осложнений, случайный лес, ROC-кривая, ансамбль голосования, деревья решений.

Abstract. In this paper, ensemble algorithms are described. As a rule, ensemble models combine several basic models to improve forecasting performance. The most famous example of an ensemble model is a Random Forest, which, greatly simplifying the logic of the algorithm, combines several decision trees and aggregates their predictions using a majority of votes.

Key words: prediction of postoperative complications, random forest, ROC curve, voting ensemble, decision trees.

Аналогично случайному лесу, ансамбль голосования оценивает несколько базовых моделей и использует голосование для объединения отдельных прогнозов, чтобы получить окончательный результат. Однако