

УДК 343.8 : 311.2 (045)
DOI: 10.22213/2410-9304-2025-3-117-124

Программная реализация систем обработки статистических данных производственного сектора пенитенциарной системы

Д. С. Пономарев, кандидат технических наук, ИжГТУ имени М. Т. Калашникова;
Научно-исследовательский институт Федеральной службы исполнения наказаний

В статье рассмотрена программная реализация и итоги разработки информационной системы для работы со статистическими данными производственного пенитенциарного сектора. Актуальность работы обусловлена как нормативно-правовой документацией, так и масштабами организованных на территории Российской Федерации производств пенитенциарной системы: по официальным актуальным данным общий объем производства пенитенциарной составил 44,3 млрд руб.

Проведен анализ отчетных статистических форм пенитенциарной системы. Представлены результаты теоретико-множественного и теоретико-информационного анализа статистических данных производственного пенитенциарного сектора. Установлено, что для производственного сектора пенитенциарной системы исследование трудовых ресурсов является одной из самых важных составляющих (в свою очередь он может включать параметры, которые отражают результаты воспитательной работы, медико-санитарного обеспечения спецконтингента). Исходя из полученных результатов, были разработаны решения для формирования архитектур баз данных для проведения статистического анализа деятельности производственного сектора. Приведены методы, подходы и технические решения, которые были применены в разработке: принципы объектно-ориентированного подхода, функционального подхода, паттернов проектирования информационных систем, SOLID-принципы. Рассмотрены особенности применения в обработке данных таких библиотек, как Pandas, SciKit-Learn, TensorFlow, Keras, CatBoost и Plotly. Приведены подходы к разработке пользовательских интерфейсов.

Предложены подходы к развертыванию информационной системы при помощи веб-технологий. Рассмотрена разработка системы контроля управления доступом для пользователя информационной системы. Практическая значимость работы заключается во внедрении системы в подразделения ФСИН России, где она может быть использована для анализа производственных процессов, прогнозирования рисков и поддержки принятия решений. В дальнейшем планируется продолжить исследования в области масштабирования системы, обеспечения информационной безопасности и поддержания жизненных циклов программного обеспечения.

Ключевые слова: информационная система, производство, пенитенциарная система, системный анализ, статистика.

Введение

Согласно официальным данным, в структуру производственного сектора УИС входит 596 исправительных учреждений. В исправительных учреждениях на работах, связанных с приносящей доход деятельностью, привлечено к труду 127,9 тыс. осужденных, а общий годовой объем производства составил 44,3 млрд руб. (ФСИН России: официальный сайт. Москва. URL: <https://fsin.gov.ru/?%2F=>). Для контроля деятельности организованных в пенитенциарной системе производств в Федеральной службе исполнения наказаний были утверждены отчетные статистические формы (отчет о трудовой адаптации осужденных (далее – ОТАО)), а также разработано программное обеспечение, которое позволяет хранить статистические данные. Однако на сегодняшний день в пенитенциарной системе отсутствует какое-либо утвержденное или общепринятое программное обеспечение,

позволяющее проводить согласованную между подразделениями работу с собранными данными.

Исходя из рассмотренных исследований, было установлено, что актуальность разработки информационной системы (ИС) для производственно-трудового пенитенциарного сектора обусловлена несколькими факторами. Во-первых, это необходимость соответствия нормативно-правовым документам, которые регулируют развитие информационных технологий в УИС. Во-вторых, рост доли числа трудоустроенных осужденных и увеличение объемов производства в пенитенциарных учреждениях требуют внедрения современных систем управления и прогнозирования. В-третьих, специфика производственного сектора УИС, связанная с повышенными рисками в области безопасности и качества продукции, которая делает необходимым создание специализированных информационно-

аналитических комплексов. В-четвертых, отсутствие утвержденного программного обеспечения и общепринятых между подразделениями ФСИН России методик анализа, расчета и прогноза использования статистических учтенных данных в оценке и прогнозе производственно-трудовой деятельности в пенитенциарной системе.

Основной целью данной статьи являлось представить основные результаты разработки и применения информационной системы для обработки статистических данных производственно-трудового сектора пенитенциарной системы.

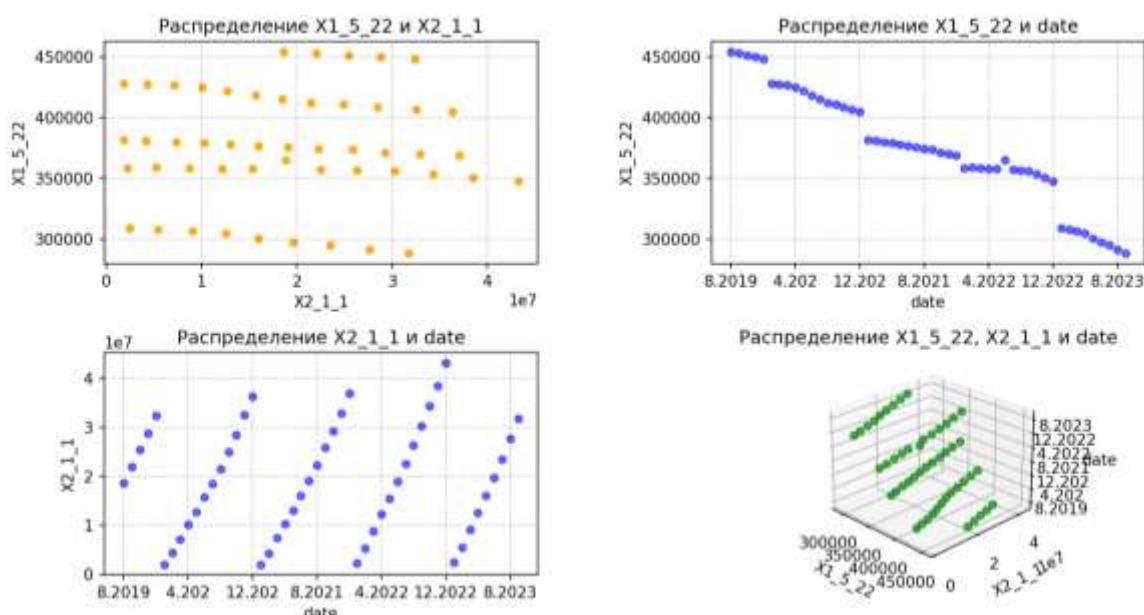
Методы и подходы

В работе использовались как теоретические, так и экспериментальные методы исследования. Теоретическая часть включала анализ нормативно-правовых документов, теоретико-множественный и теоретико-информационный анализы производственных процессов и разработку методологических основ для создания информационных систем. Экспериментальная часть базировалась непосредственно на самой разработке ИС, а также на прикладном применении разработанных инструментов.

В процессе разработки ИС были использованы принципы как объектно-ориентиро-

ванного подхода (ООП), так и функционального подхода (ФП). Применены общеизвестные паттерны проектирования и системного дизайна [1], *SOLID*-принципы [2, 3], предложены и применены системы индексации при помощи алгоритма *b*-дерева [4]. Рассмотрены и выделены связи «Один ко многим», «Многие ко многим». Разработана архитектура базы данных. Использованы результаты *OLAP*-анализа.

Пример визуализации *OLAP*-куба при помощи разработанных ранее программ для ЭВМ (Пономарев Д. С., Горохов М. М., Пономарев С. Б. Программа для графического анализа взаимных распределений числовых параметров из статистических отчетных форм ФСИН России. Свидетельство о регистрации программы для ЭВМ RU 2023685776, 29.11.2023. Заявка № 2023680076 от 29.09.2023; Пономарев Д. С., Горохов М. М. Программа для автоматизации создания отчетов разведочного анализа данных. Свидетельство о регистрации программы для ЭВМ RU 2024681798, 13.09.2024. Заявка № 2024680458 от 02.09.2024) представлен на рисунке.



Пример взаимного распределения между количеством трудоустроенного спецконтингента ($X_{m=1,p=5,o=22}$) и объемами производства ($X_{m=2,p=1,o=1}$)

Example of mutual distribution between the number of employed convicts ($X_{m=1, p=5, o=22}$) and production volumes ($X_{m=2, p=1, o=1}$)

Как видно из данного рисунка, учтенные данные по объемам производства являются

циклическим времененным рядом (где цикл равен одному году), в то время как учет трудо-

устроенного спецконтингента является трендовым. Более подробно методология сопоставления циклических и трендовых временных рядов для статистических данных пенитенциарной системы была раскрыта в опубликованной ранее работе (Пономарев Д. С. Разработка алгоритмов обработки временных рядов при работе с статистическими отчетными формами производственного сектора пенитенциарной системы // Инженерный вестник Дона. 2024. № 2 (110). С. 184–191). Проведенный же *OLAP*-анализ позволил более детально изучить и выявить закономерности и особенности взаимодействия между параметрами и показателями производственного сектора пенитенциарной системы, а также организовать срезы данных по заданным критериям. Стоит также отметить, что согласно утвержденным отчетным формам пенитенциарной системы, на сегодняшний день для производственного сектора ведется учет более 500 параметров и показателей, поэтому более детальное представление результатов *OLAP*-анализа (из-за объема данных) выходит за рамки поставленных задач данной статьи.

Теоретико-множественный анализ производственно-трудового сектора пенитенциарной системы

Для начала был проведен теоретико-множественный анализ производственно-трудового сектора пенитенциарной системы. В процессе работы осуществлен комплексный анализ базовых алгоритмов управления производственными процессами с оценкой перспектив внедрения системного подхода. Представлены инновационные решения по созданию систем упорядочивания статистических данных, соответствующих стандартам ведомственной отчетности и учитывающих требования иерархического распределения. Проведена детализация практического применения принципов реляционного моделирования. Установлено, что применение алгоритмов индексации на основе *b*-дерева может стать одним из подходящих решений исходя из имеющейся структуры данных, которая утверждена пенитенциарным сектором на сегодняшний день. Рассмотрена разработка алгоритмов и методов для работы с данными как с временными рядами исходя из существующих систем и правил учета данных за определенные периоды. Предложены и рассмотрены на примерах подходы к формированию ключевых параметров эффективности (в некоторой литературе *key performance indicators* [5, 6]). Рассмотрены основы для организации

срезов данных, а также примеры применения постулатов теории множеств и реляционной теории. Проведено исследование по формированию архитектур баз данных для ИС.

Было установлено, что систему учета данных можно представить в виде иерархической структуры. В качестве «основы» следует рассматривать класс «Заказы»: объем заказов на текущий год формируется в первый месяц года (то есть для параметров, которые характеризуют объемы заказов, максимальное значение будет в январе). К концу года все сформированные «Заказы» должны быть выполнены, соответственно, в декабре их значение будет равно нулю (согласно ведомственной отчетности). Объем сформированных заказов напрямую влияет на общий объем производства к концу года. Соответственно, максимальное значение объемов производства будет в декабре каждого года. Аналогичным образом дело обстоит с данными по расходам, остатками и браком продукции. Объем производства, расходы, остатки и брак продукции влияют на следующий уровень – доходы и прибыль, а также задолженности: при увеличении расходов и бракованной продукции доходы прибыль могут снижаться, а задолженности, соответственно, будут расти, и наоборот. В конечном итоге (следующий уровень иерархии) для всех перечисленных классов и соответствующих им параметров учет может вестись (согласно утвержденной системе отчетности) в рамках либо учреждения, где организовано производство, либо определенной отрасли производства.

Влияние криминальной субкультуры, изначальное отсутствие трудовой адаптации у спецконтингента (в некоторых случаях), наличие социально значимых заболеваний (таких как СПИД, ВИЧ, туберкулез), отсутствие возможности уволить работника и отсутствие возможности проведения кадрового поиска накладывают свой отпечаток на организацию производственных процессов в учреждениях пенитенциарной системы. Таким образом, проведение исследований для производств пенитенциарной системы представляет достаточно большой интерес, а с учетом масштабности организованного в Российской Федерации производственного пенитенциарного сектора актуальность решения представленных вопросов может иметь значение для экономической составляющей непосредственно всей страны в целом.

Технические решения в разработке

В ходе проведенного анализа компьютерных методов обработки информации для производ-

ственno-трудового сектора пенитенциарной системы были определены подходы, алгоритмы и методы разработки для создания инструментов проведения статистического анализа данных и обучения моделей.

Основываясь на проведенном теоретико-множественном и теоретико-информационном анализах, была установлено, что одним из ключевых моментов является создание инструментов для обработки многомерных данных.

Поэтому были разработаны фильтры данных, которые позволяют организовать не только срезы данных, но и провести сравнительный анализ между разными периодами (например, сравнить результаты объемов производства (или какого-либо другого важного для производства параметра, рассмотренного ранее) текущего года относительно предыдущего в заданные временные отрезки по заданным отраслям производства и заданным территориям). Рассмотрены практические примеры прикладной статистической обработки и тесты производительности локального сервера на примере выборки из 10 тыс. точек, записанных в базу данных, которая соответствует архитектуре, описанной ранее (данные были собраны за 5 лет). В результате проведенных тестов было установлено, что время получения отфильтрованных данных для дальнейшего использования не превышает 0,1 секунды (для тестирования использована библиотека *time*). Также были разработаны инструменты нормализации и стандартизации данных; инструменты для автоматизации разведочного анализа данных (в некоторых источниках *exploratory data analysis* [7, 8]); созданы инструменты графической визуализации данных.

Разработка инструментов машинного обучения

Разработка инструментов машинного обучения велась на примере методов глубокого обучения и методов градиентного бустинга (Брюс П., Брюс Э., Гедек П. Практическая статистика для специалистов Data Science: 50 важнейших понятий с использованием R и Python. 2-е изд., перераб. и доп. СПб.: БХВ-Петербург, 2021. С. 284; Протодьяконов А. В., Пылов П. А., Садовников В. Е. Алгоритмы Data Science и их практическая реализация на Python: учеб. пособие. Москва; Вологда: Инфра-Инженерия, 2022. С. 203–207) (примеры использования методов градиентного бустинга представлены в опубликованных ранее работах отечественных авторов [9, 10], в зарубежных публикациях для данного метода используется термин – *gradient boosting machine* [11–13] (далее – *GBM*), реализация

GBM, как правило, осуществляется при помощи таких библиотек (которые относятся к открытому ПО), как *XGBoost*, *Catboost* (является отечественной разработкой), *LightGBM*. В основе разработки инструментов глубокого обучения были использованы такие библиотеки, как *Pandas*, *SciKit-Learn*, *TensorFlow*, *Keras* (согласно исследованию [14] представленные библиотеки являются одними из самых популярных и применяемых при обработке данных с помощью *Python*). Рассмотрены возможные варианты использования оптимизаторов. В проводимой разработке выбор был остановлен на алгоритме *Adam*. Это было обосновано тем, что *Adam* сочетает в себе преимущества *Momentum* и *RMSProp*. Оптимизатор *Adam* динамически изменяет скорость обучения для каждого параметра модели на основе средних значений градиентов и их квадратов.

В разработке были применены паттерны и принципы проектирования ИС на основе ООП. Использован конструктор класса – *__init__*. Разработаны теоретические основы для создания конкретных классов по принципу «Разделение ответственности»:

класс *TrainingDeviceManager* – отвечает за выбор устройства для обучения (*GPU* или *CPU*) и настройку соответствующего окружения;

класс *DataHandler* – занимается загрузкой данных, их предобработкой и масштабированием;

класс *ModelBuilder* – отвечает за создание, обучение и оценку модели;

класс *AppInterface* управляет интерфейсом пользователя, обеспечивая интуитивное взаимодействие с программой.

Рассмотрены основные шаги по разработке, разработаны *UML*-диаграммы (классов, деятельности). Разработка приведенных классов основывалась на таких паттернах проектирования ИС, как *MVC* (*Model-View-Controller*); «Инкапсуляция»; «Композиция». Из *SOLID*-принципов были использованы: принцип единственной ответственности и принцип открытости-закрытости при разработке классов и функций, принцип разделения интерфейса при создании внешней пользовательской оболочки, принцип инверсии зависимостей при создании абстракций. В целом, технические решения базировались как на стандартных библиотеках: *io*; *tempfile*; *subprocess*; *time*, так и на внешних: *Streamlit*; *TensorFlow*; *Keras*; *Pandas*; *Matplotlib*; *Scikit-learn*; *Openpyxl*; *xlwt*.

Помимо разработанных инструментов глубокого обучения, были созданы инструменты для создания моделей на основе *GBM*. В основе дан-

ной разработки была применена библиотека *Catboost*, это было сделано по ряду причин: из-за имеющихся категориальных данных, работа с которыми имеет значение в расчетах; низких рисках переобучения моделей; использования в алгоритмах обучения более строгой *L2*-регуляризации; автоматизации обработки пропущенных данных; более упрощенной настройки гиперпараметров. Дополнительными преимуществами, которые повлияли на выбор данной библиотеки, явились: *Catboost* является отечественной разработкой; имеется возможность ее полной интеграции с другими известными библиотеками (например, такими как *SciKit-Learn*, *Pandas* и др.); кроме того, данная библиотека появилась среди открытого ПО сравнительно недавно, что повышает интерес к ее использованию.

В данной разработке был применен ФП. Были использованы такие паттерны проектирования ИС, как: *Factory* (Фабричный метод); *Strategy* (Стратегия); *TemplateMethod* (Шаблонный метод); *Adapter* (Адаптер); *Command* (Команда); *Observer* (Наблюдатель).

Также были применены и другие библиотеки. В частности, использование библиотеки *Pandas* – для подготовки и предобработки данных, *SciKit-Learn* – для разделения выборки на тестовую и обучающую, а также для проведения тестов на предсказательную способность обученной модели, *Streamlit* – для визуализации пользовательских интерфейсов управления разработанным инструментом (актуальность использования данной библиотеки для научных исследований рассмотрена в работах зарубежных авторов [15, 16]), *Plotly* – для мониторинга процесса обучения, *Catboost* – для проведения непосредственно самого обучения, а также последующего использования уже обученных моделей на новых данных, *OS* – для сохранения обученных моделей и сопроводительных данных по ним в файловой системе.

Были разработаны функции: *main()* – главная функция, которая отвечает за выбор режима (обучение новой модели или загрузка существующей) и отображение интерфейса с помощью *Streamlit*; *Load_data()* – функция, которая загружает данные из файла; *Train_model()* – данная функция обучает модель *CatBoost*, возвращает обученную модель, точность, метрики, и важность признаков; *Load_model()* – функция загружает сохраненную ранее модель *CatBoost* из файла; *CatBoostClassifier* – этот класс был использован непосредственно из библиотеки *CatBoost*, использовался для классификации.

Таким образом, были разработаны инструменты машинного обучения. Следующей задачей являлось провести развертывание ИС и организовать доступ пользователей. Рассмотрим реализацию данных шагов в следующем разделе.

Организация СКУД и подходы к развертыванию

Также была разработана система контроля управлением доступом (СКУД). Так как разработка ИС ориентирована на производства пенитенциарной системы, то создание аутентификации пользователей с заранее заданными логином и паролем будет более актуальной, чем системы авторизации. В нашем случае разработка СКУД была рассмотрена с возможностью администрирования.

Шифрование паролей базировалось на основе библиотеки *bcrypt* (более подробно использование которой было рассмотрено в опубликованных ранее работах [17, 18]), организация хранения зашифрованных паролей рассмотрена на примере файлов *JSON* (хотя использование отдельной БД для этих целей было также возможно). Учтена возможность добавления новых пользователей через администратора.

Помимо этого, были рассмотрены возможности развертывания ИС на локальном, облачном и собственном удаленном серверах. В качестве облачного сервера был рассмотрен сервис *streamlit.share.io*. Для развертывания на удаленном сервере рассмотрено применение контейнеров *Docker*. Разработаны инструкции для создания и запуска контейнеров.

Таким образом, в представленной работе были рассмотрены итоги разработки информационной системы для производственного пенитенциарного сектора. В качестве продолжения представленной разработки могут быть рассмотрены: поддержание жизненных циклов информационных систем и оркестрация контейнеров, масштабирование, дальнейшее обеспечение информационной безопасности веб-систем.

Заключение

В представленной работе осуществлена комплексная разработка и апробация программной реализации для информационной системы, ориентированной на статистическую обработку данных производственного сектора пенитенциарной системы. Проведенный теоретико-множественный и теоретико-информационный анализ позволил сформировать методологическую основу для создания эффективной архитектуры хранения и обработки статистических данных, соответствующей специфике ведомственной отчетности УИС.

Разработанная информационная система базируется на современных подходах программной инженерии, включая объектно-ориентированное и функциональное программирование, применение паттернов проектирования, *SOLID*-принципов, а также использование библиотеки инструментов для анализа и визуализации данных. Реализована возможность обработки многомерных временных рядов, анализа производственных показателей, обучения моделей машинного обучения и визуализации результатов.

Особое внимание было уделено вопросам внедрения информационной системы в условиях ограниченного доступа и повышенных требований к безопасности, характерных для учреждений ФСИН. Разработана система управления доступом пользователей, предусмотрены варианты развертывания на различных платформах, включая использование технологий контейнеризации.

Практическая значимость работы заключается в ее готовности к интеграции в реальные подразделения УИС для оптимизации производственно-го планирования, прогнозирования и поддержки управленических решений. В дальнейшем предполагается развитие системы в направлении масштабирования, обеспечения информационной безопасности и поддержки жизненного цикла программного обеспечения.

Библиографические ссылки

- Шмаков С. Э., Щенникова Е. В., Определенцева А. Е. Взаимодействие паттернов проектирования для эффективной web-разработки // Ученые записки УлГУ. Серия: Математика и информационные технологии. 2021. № 2. С. 90–96.
- Федорова О. Single responsibility principle // Системный администратор. 2024. № 1-2 (254-255). С. 62–63.
- Дехтиевский С. А., Бударный Г. С. Принципы SOLID и их применение в разработке программного обеспечения // Научный аспект. 2024. Т. 43, № 4. С. 5635–5641.
- Graefe G. More modern b-tree techniques // Foundations and Trends in Databases. 2024. Т. 13. № 3. С. 169–249.
- Uzule K., Zarina V., Shina I. Fostering performance management by identifying key performance indicators for human capital in integrated reporting // Business: Theory and Practice. 2024. Т. 25. № 2. С. 420–433.
- Sanosra A., Maimun A., Satoto E.B., Qomariah N. Role of key performance indicators and compensation in improving employee integrity and performance / International Journal of Economics and Management Studies. 2022. Т. 9. № 2. С. 35–43.
- GrihaTofik Isa I., Zulkarnaini Z., Novianti L., Elfadonna F., Agustri S. Exploratory Data Analysis (EDA) dalam Dataset Penerimaan Mahasiswa Baru Universitas XYZ Palembang / Smart Comp: Jurnalnya Orang Pintar Komputer. 2023. Т. 12. № 3. С. 600–609.
- Qian Zh., Sun K. Extensive analysis of rain in Australia by exploratory data analysis, feature engineering and modeling // Theoretical and Natural Science. 2023. Т. 7. № 1. С. 63–71.
- Королев Н. С., Сенько О. В. Метод повышения эффективности обучения градиентного бустинга, основанный на модифицированных функциях потерь // Автоматика и телемеханика. 2022. № 12. С. 78–88.
- Кирюшин Ю. Н., Мухеев М. Ю. Построение деревьев решений при помощи градиентного бустинга // Современные информационные технологии. 2023. № 37 (37). С. 40–43.
- Song Z., Xia J., Wang G., She D., Hu C., Hong S. Regionalization of hydrological model parameters using gradient boosting machine // Hydrology and Earth System Sciences. 2022. Т. 26. № 2. Р. 505–524.
- Sun Ya., Zhang F., Lin H., Xu Sh. A forest fire susceptibility modeling approach based on light gradient boosting machine algorithm // Remote Sensing. 2022. Т. 14, № 17. Р. 4362.
- Massalha E., Sharon A., Oren D., Fishman B., Segev A., Matetzky Sh., Maor E. A gradient boosting machine learning model enhances mortality prediction in non-st-segment elevation myocardial infarction // Journal of the American College of Cardiology. 2022. Т. 79. № 9. С. 1131.
- Samira Gholizadeh. Top Popular Python Libraries in Research // Journal of Robotics and Automation Research. 2022. 3(2). Р. 142–145.
- Chanin Nantasesamat, Avratana Biswas, J.M. Nápoles-Duarte, Mitchell I. Parker, Roland L. Dunbrack Jr. Building Bioinformatics Web Applications with Streamlit // Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development. 2023. №27. Р. 679-699.
- Yaganteeswarudu Akkem, Biswas Saroj Kumar, Aruna Varanasi. Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation // Indian Journal of Science and Technology. 2023. Vol.16(48), P. 4688-4702.
- Erna Nababan. Analysis Performance Bcrypt Algorithm to Improve Password Security from Brute Force // Journal of Physics Conference Series. 2021. 1811(1): 012129. doi:10.1088/1742-6596/1811/1/012129
- Toras Pangidoan Batubara, Syahril Efendi, Erna Budhiarti Nababan, Analysis Performance BCRYPT Algorithm to Improve Password Security from Brute Force // Journal of Physics: Conference Series. 2021. 1811 (2021): 012129. doi:10.1088/1742-6596/1811/1/012129

References

- Shmakov S.E., Shchennikova E.V., Opredelelenceva A.E. [Interaction of design patterns for effective web development]. *Scientific notes of Ulsu. Series: Mathematics and information technology*. 2021. No. 2. P. 90-96.

2. Fedorova O. [Single responsibility principle]. *System administrator*. 2024. No. 1-2 (254-255). P. 62-63.
3. Dekhtievsky S.A., Budarny G.S. [SOLID principles and their application in software development]. *Scientific aspect*. 2024. Vol. 43. No. 4. P. 5635-5641.
4. Graefe G. More modern b-tree techniques / Foundations and Trends in Databases. 2024. T. 13. No. 3. P. 169-249.
5. Uzule K., Zarina V., Shina I. Fostering performance management by identifying key performance indicators for human capital in integrated reporting / Business: Theory and Practice. 2024. T. 25. No. 2. P. 420-433.
6. Sanosra A., Maimun A., Satoto E.B., Qomariah N. Role of key performance indicators and compensation in improving employee integrity and performance / International Journal of Economics and Management Studies. 2022. T. 9. No. 2. P. 35-43.
7. GrihaTofik Isa I., Zulkarnaini Z., Novianti L., El-faladonna F., Agustri S. Exploratory Data Analysis (EDA) of the Dataset of the PenerimaanMahasiswaBaru Universitas XYZ Palembang / Smart Comp: Jurnalnya Orang PintarKomputer. 2023. Vol. 12. No. 3. Pp. 600-609.
8. Qian Zh., Sun K. Extensive analysis of rain in Australia by exploratory data analysis, feature engineering and modeling // Theoretical and Natural Science. 2023. Vol. 7. No. 1. Pp. 63-71.
9. Korolev N.S., Senko O.V. [Method for improving the efficiency of gradient boosting training based on modified loss functions]. *Automation and Telemechanics*. 2022. No. 12. P. 78-88.
10. KiryushinYu.N., MikheevM.Yu. [Construction of decision trees using gradient boosting]. *Modern information technologies*. 2023. No. 37 (37). P. 40-43.
11. Song Z., Xia J., Wang G., She D., Hu C., Hong S. Regionalization of hydrological model parameters using gradient boosting machine // *Hydrology and Earth System Sciences*. 2022. Vol. 26. No. 2. P. 505-524.
12. Sun Ya., Zhang F., Lin H., Xu Sh. A forest fire susceptibility modeling approach based on light gradient boosting machine algorithm // *Remote Sensing*. 2022. T. 14. No. 17. P. 4362.
13. Massalha E., Sharon A., Oren D., Fishman B., Segev A., Matetzky Sh., Maor E. A gradient boosting machine learning model enhances mortality prediction in non-st-segment elevation myocardial infarction // *Journal of the American College of Cardiology*. 2022. T. 79. No. 9. P. 1131.
14. Samira Gholizadeh. Top Popular Python Libraries in Research // *Journal of Robotics and Automation Research*. 2022. 3(2). P. 142-145.
15. ChaninNantasenamat, Avratanu Biswas, J.M. Napolis-Duarte, Mitchell I. Parker, Roland L. Dunbrack Jr. Building Bioinformatics Web Applications with Streamlit // *Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development*. 2023. No. 27. P. 679-699.
16. YaganteeswaruduAkkem, Biswas Saroj Kumar, Aruna Varanasi. Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation // *Indian Journal of Science and Technology*. 2023. Vol.16(48), pp. 4688-4702.
17. Erna Nababan. Analysis Performance Bcrypt Algorithm to Improve Password Security from Brute Force // *Journal of Physics Conference Series*. 2021. 1811(1): 012129. doi:10.1088/1742-6596/1811/1/012129
18. TorasPangidoan Batubara, Syahril Efendi, Erna BudhiartiNababan, Analysis Performance BCRYPT Algorithm to Improve Password Security from Brute Force // *Journal of Physics: Conference Series*. 2021. 1811 (2021): 012129. doi:10.1088/1742-6596/1811/1/012129

Software Implementation of Statistical Data Processing Systems for the Production Sector of the Penitentiary System

D.S. Ponomarev, PhD in Engineering, Izhevsk State Technical University named after M. T. Kalashnikov; Research Institute of the Federal Penitentiary Service

The article considers the results of the development of an information system for working with statistical data of the industrial penitentiary sector. The relevance of the work is due to both the regulatory and legal documentation and the scale of the penitentiary system production organized in the territory of the Russian Federation: according to official current data, the total volume of penitentiary production amounted to 44.3 billion rubles. The analysis of the reporting statistical forms of the penitentiary system was carried out. The results of the set-theoretical and information-theoretical analysis of statistical data of the industrial penitentiary sector are presented. It was established that for the industrial sector of the penitentiary system, the study of labor resources is one of the most important components (in turn, it can include parameters that reflect the results of educational work, medical and sanitary support of the special contingent). Based on the results obtained, solutions were developed for the formation of database architectures for statistical analysis of the activities of the industrial sector. The methods, approaches and technical solutions that were applied in the development are presented: principles of the object-oriented approach, the functional approach, information system design patterns, SOLID principles. The features of using such libraries as Pandas, SciKit-Learn, TensorFlow, Keras, CatBoost and Plotly in data processing are considered. Approaches to the development of user interfaces are given. Approaches to the deployment of an information system using web technologies are proposed. The development of an access control system for the user of the information system is considered. The practical significance of the work lies in the implementation of the system in the divisions of the Federal Penitentiary Service of Russia,

where it can be used to analyze production processes, predict risks and support decision-making. In the future, it is planned to continue research in the field of system scaling, ensuring information security and maintaining software life cycles.

Keywords: information system, production, penitentiary system, systems analysis, statisticsds.

Получено: 31.03.25

Образец цитирования

Пономарев Д. С. Программная реализация систем обработки статистических данных производственного сектора пенитенциарной системы // Интеллектуальные системы в производстве. 2025. Т. 23, № 3. С. 118–124. DOI: 10.22213/2410-9304-2025-3-118-124.

For Citation

Ponomarev D.S. [Software implementation of statistical data processing systems for the production sector of the penitentiary system]. *Intellektual'nye sistemy v proizvodstve*. 2025, vol. 23, no. 3, pp. 118-124. DOI: 10.22213/2410-9304-2025-3-118-124.