

**ФУНКЦИОНАЛЬНЫЕ ХАРАКТЕРИСТИКИ
СИСТЕМЫ БИЗНЕС-АНАЛИТИКИ С КОНСТРУКТОРОМ
НЕЙРОСЕТЕЙ**

Содержание

1	Введение	3
2	Основные понятия системы.....	3
3	Назначение системы	5
4	Ключевые функциональные блоки.....	5
4.1	Приложения	5
4.2	Конструктор ETL	6
4.3	Администрирование	7
4.4	Безопасность.....	7
4.5	Предиктивная аналитика	10
5	Протоколирование	10
6	Заключение	12

1 Введение

Система бизнес-аналитики с конструктором нейросетей (далее – Система аналитики) позволяет объединять данные из различных источников, визуализировать полученные данные, а блок нейронных сетей позволяет строить и обучать нейронные сети для с целью выявления тенденций и прогнозирования развития событий, что дает дополнительную информацию для принятия управленческих решений.

Архитектура Системы аналитики спроектирована с учетом необходимости гарантированной доставки и быстрой обработки больших объемов данных, прозрачности ETL-потоков, возможности масштабирования решения и оперативного добавления новых типов источников данных.

Ключевые преимущества продукта:

- Многоступенчатый сценарий процесса ETL с возможностью использования кастомизированных математических моделей преобразования входных данных, ввода обучающих выборок для предиктивной аналитики (обучения цепочек нейросетей);
- Гарантированная и согласованная доставка преобразованной информации в аналитическое хранилище;
- Раздел безопасности, позволяющий настраивать роли пользователей и отслеживать их действия в системе;
- Интерактивные аналитические приложения.

2 Основные понятия системы

Таблица 1. Основные понятия Системы аналитик

Термин	Описание
ETL-процесс	Процесс извлечение, преобразования и отправки данных в аналитическое хранилище
ETL-хранилище (PostgreSQL)	Хранилище на postgresql, в которое выгружаются данные из внешних источников, для дальнейшего преобразования и отправки в аналитическое хранилище

DWH-хранилище (Apache Druid)	СУБД, обеспечивающая колоночное хранение большого объема данных, выступает как Аналитическое хранилище
Виджет	Это график, построенный на основе выбранного шаблона визуализации и данных
Источники данных	Это внешние системы, к которым с помощью коннекторов может подключиться Инженер данных и получить их структуру данных
Глобальные фильтры	Работают в рамках приложения. Глобальные фильтры позволяют устанавливать фильтрацию по разным значениям полей
Дашборд	Информационная панель, которая состоит из совокупности виджетов, имеет признак интерактивности
Домен	Автономная область, в том числе безопасности, которая характеризуется следующими признаками: уникальными пользователями, ETL-хранилищами, аналитическим хранилищем, приложениями, дашбордами, данными
Конструктор	Раздел в системе для конструирования ETL-процесса, управления безопасностью, управлению настройками визуализации, и администрирования системы
Кросс-фильтрация	Функционал на дашборде, при котором пользователь выбирает на виджете значение(ия) и другие виджеты фильтруются согласно этому запросу. Кросс-фильтрация работает только в рамках одного дашборда
Приложение	Обособленная предметная область, которая состоит из совокупности дашбордов, и обладает собственной безопасностью
Служебная база данных	База postgresql, предназначенная для хранения объектов: описание источников данных, описание ETL-хранилищ, описание аналитических хранилищ, описание приложений (дашбордов, виджетов и правил по работе с ними), описание ETL-задач, информацию по подсистеме прав доступа (пользователи, группы, роли), служебные таблицы, фильтры
Шаблон (шаблон визуализации)	Тип визуального представления информации. Например: кольцевая диаграмма, таблица, карточка с текстом, столбчатый график

Шина данных (Apache Kafka)	Распределенный брокер сообщений, используется для доставки данных из внутреннего ETL-хранилища до аналитического хранилища
Предиктивная аналитика	Раздел для создания и настройки своих собственных нейронных сетей и алгоритмов машинного обучения.

3 Назначение системы

Основной целью Системы аналитики является автоматизация информационно–аналитической деятельности организации в бизнес–процессах с целью получения полного и наглядного доступа к консолидированной информации, отражающей деятельность организации. Данная информация может быть выражена в виде совокупности бизнес-показателей и визуальных представлений, в следствии чего появится возможность выработки и принятия качественно новых управленческих решений в оптимальный срок.

4 Ключевые функциональные блоки

Система состоит из следующих блоков:

- Администрирование
- Безопасность
- Приложения
- Конструктор ETL
- Предиктивная аналитика

4.1 Приложения

Раздел Приложения предназначен для создания аналитических приложений и интерактивных дашбордов. Приложение — это обособленная предметная область, которая состоит из совокупности дашбордов, и обладает собственной безопасностью. Дашборд — это информационная панель, которая состоит из совокупности виджетов, имеет признак интерактивности. Пользователь может создавать неограниченное число приложений, в каждом приложении может быть неограниченное количество дашбордов, каждый

дашборд может содержать неограниченное количество виджетов благодаря вертикальному скроллу.

С приложениями могут работать Дизайнеры приложений (создание, редактирование, удаление, публикация, настройка глобальных фильтров) и Пользователи (просмотр, использование глобальной фильтрации).

К каждому приложению можно назначить доступ в рамках настроек безопасности через группы доступа. В том числе каждое приложение имеет возможность быть опубликованным (то есть стать видимым) для роли Пользователь.

На каждое приложение могут быть назначены глобальные фильтры, которые позволяют отфильтровать информацию на всех дашбордах в рамках приложения согласно настроенным и выбранным значениям.

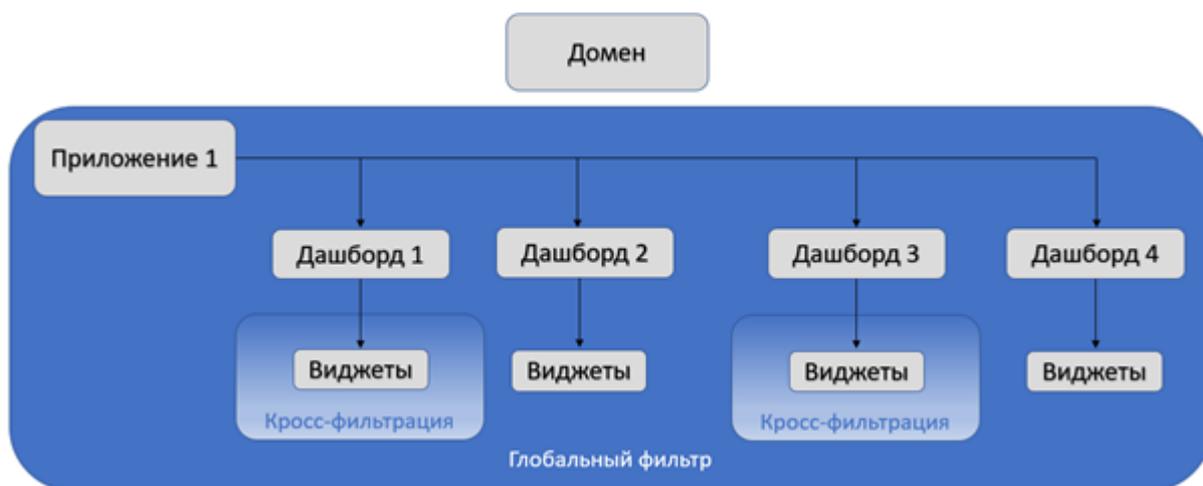


Рисунок 2. Фильтрация

На каждом дашборде пользователь может настраивать кросс-фильтрацию, которая позволяет отфильтровать интересующую информацию в рамках конкретного дашборда для проведения сценарного анализа.

4.2 Конструктор ETL

Конструктор ETL – это основной раздел Дата-инженера. В конструкторе реализован следующий функционал:

- Подключение к источникам данных для получения их структуры, в том числе по каждой таблице отдельно

- Работа с ETL-хранилищем и DWH-хранилищем (аналитическим) в части просмотра структуры, ввода бизнес-наименований, заведение описаний для таблиц и полей и т.д.

- Создание ETL-задач на извлечение, трансформацию, загрузку в аналитическое хранилище, подключение ML моделей

- Мониторинг статуса ETL-задач и их шагов

4.3 Администрирование

Администрирование – это основной раздел Мастер администратора и Администратора домена. В данном блоке системы реализован следующий функционал:

- Создание и настройка доменов (назначение серверов ETL-хранилища, Аналитического хранилища, Шины данных количество неограниченно в рамках одного домена)

4.4 Безопасность

Безопасность – раздел меню, отвечающий за пользователей, их права и действия в системе. С помощью данного раздела, Мастер администратор (подробнее про его функционал в руководстве мастера администратора) заводит новых пользователей, добавляет их в домены и просматривает их активные сессии. Администратор домена видит подготовленный список пользователей для своего домена, назначает им Бизнес-роли и распределяет в различные группы доступа.

Функциональные возможности работы пользователей с системой определяются такой сущностью, как «Роль». У одного пользователя может быть несколько ролей в рамках одного домена. В разных доменах у одного и того же пользователя может быть разная роль.

Роль состоит из набора привилегий. Привилегия – это право на выполнение атомарного бизнес-действия.

При создании домена, в системе есть следующие предустановленные роли:

Таблица 2. Роли в Системе

Роль	Возможности	Доступы
<p>Мастер Администратор (данная роль не назначается, она появляется автоматически при развертывании системы)</p>	<ul style="list-style-type: none"> - Создать домен; - Назначить Администратора домена; - Редактировать/блокировать домены; - Создать пользователей и назначить им домен (или несколько); - Просматривать и редактировать список пользователей и доменов. 	<p>Имеет доступ к информации о серверах, доменах, пользователях по всем доменам</p>
<p>Администратор домена (данная роль назначается Мастером администратором)</p>	<ul style="list-style-type: none"> - Просматривать список пользователей по домену; - Назначить пользователям роли; - Создать/редактировать/удалить группы доступа; - Назначить пользователям группу доступа; - Просматривать пользователей и список групп доступа с информацией о пользователях и к чему дан доступ: приложения, таблица данных, модель данных, источники данных, схемы в ETL-хранилище, глобальные фильтры; - Просматривать все сервисы, которые назначены домену Мастер администратором; - Создавать схемы в ETL-хранилище. 	<p>Имеет доступ к информации о серверах своего домена, Пользователях, Ролях, Группях доступа по конкретному домену, администратором которого он назначен.</p>
<p>Инженер данных (данная роль назначается Администратором)</p>	<ul style="list-style-type: none"> - Подключение к внешним источникам данных - Настройка процесса извлечения, преобразования 	

<p>домена, роль появляется автоматически при создании нового домена)</p>	<p>данных и их загрузки в аналитическое хранилище, настройка расписания процесса (ETL – процесс); - Работа с моделями данных, хранилищами, создание дополнительных структур данных.</p>	
<p>Дизайнер приложений (данная роль назначается Администратором домена, роль появляется автоматически при создании нового домена)</p>	<p>- Конструирование аналитических приложений и дашбордов, предназначенных для работы остальных пользователей, на основе информации в ETL или аналитическом хранилище; - Создание виджетов, добавление виджетов на дашборд; - Настройка глобальных фильтров на приложении; - Использование кросс- фильтров на дашборде</p>	
<p>Пользователь (данная роль назначается Администратором домена, роль появляется автоматически при создании нового домена)</p>	<p>- Работа с аналитическими приложениями без возможности создания собственных дополнительных приложений, дашбордов, виджетов; - Использование кросс- фильтров на дашборде - Использование глобальных фильтров в приложении</p>	
<p>Дизайнер отчетов (данная роль назначается Администратором домена, роль появляется</p>	<p>- Конструирует отчеты в Генераторе отчуждаемой отчетности и настраивает рассылку отчетов.</p>	<p>Имеет доступ к генератору отчетности.</p>

<i>автоматически при создании нового домена)</i>		
<i>Аналитик ML (данная роль назначается Администратором домена, роль появляется автоматически при создании нового домена)</i>	- Работа с конструктором нейронных сетей.	Имеет доступ к разделу «Конструктор нейронных сетей».

4.5 Предиктивная аналитика

Предиктивная аналитика — это уникальный функциональный блок, который позволяет создавать и настраивать свои собственные нейронные сети и алгоритмы машинного обучения. С помощью интуитивно понятных инструментов и готовых шаблонов пользователь может создать нейросеть/алгоритм для различных задач: от предсказания стоимости продукта до прогнозирования поведения клиентов. Этот компонент системы входит в вариант расширенной поставки, встроенный в общую систему архитектуры, включая сервисы протоколирования, безопасности и администрирования.

Основная цель блока – моделирование закономерностей в данных. Достижение цели означает нахождение закономерности с достаточным показателем качества. Как только цель моделирования достигнута, полученная модель сохраняется для последующего использования в промышленном режиме.

5 Протоколирование

Логирование представляет из себя процесс создания журналов, который включает в себя сбор и организацию информации о функционировании системы в отдельных файлах с возможностью легкого доступа к ним при необходимости. Эти файлы содержат отчет о том, какие действия были

выполнены конкретными пользователями, когда они были выполнены, как система отреагировала на различные события.

Специальное программное обеспечение отвечает за запись всех важных событий для системы. В процессе создания журнала используются настройки, определяющие уровень детализации записей. Помимо создания самого журнала, важным аспектом настройки логирования является организация надежного хранения журналов. Это обеспечивает быстрый и легкий доступ к журналам. Журналы могут храниться на локальном компьютере, удаленном жестком диске, отдельном сервере, облачном хранилище или любом другом надежном носителе, ограниченном для доступа посторонних.

6 Заключение

Система аналитики является полностековой системой поддержки принятия решений, которая предоставляет организациям все необходимые инструменты для эффективного сбора, анализа и интерпретации данных. Она помогает принимать обоснованные решения на основе фактов и трендов, повышает эффективность бизнес-процессов, улучшает прогнозирование и планирование, а также оптимизирует затраты и ресурсы.