



Доклад

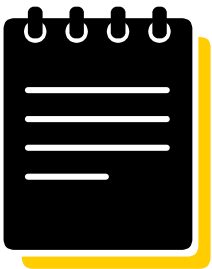
на тему: «Разработка системы поддержки принятия решений для прогнозирования оттока клиентов на базе АП Logiном »

к.ф.-м.н., доцент
Прокопенко Н.Ю.

Постановка задачи



Цель: разработать библиотеку компонентов анализа данных для прогнозирования оттока клиентов в телекоммуникационных компаниях с помощью платформы Loginom.



Задачи:

- Изучить методы анализа Data Mining
- Освоить возможности АП Loginom
- Разработать библиотеку компонентов
- Протестировать на реальных данных

Объект исследования

- Деятельность телекоммуникационных компаний, направленная на сохранение абонентской базы и привлечение новых абонентов.

Предмет исследования

- Возможности информационно-аналитической платформы Loginom, методы и модели Data Mining

DFD



Структура СППР

Подсистема сбора и хранения данных
Deductor, Loginom



Подсистема очистки и подготовки данных к
анализу и построению аналитических моделей
Loginom



Подсистема моделирования и прогнозирования
Loginom



Подсистема отображения данных
Loginom, дашборды Tableau

Подсистема сбора и хранения данных

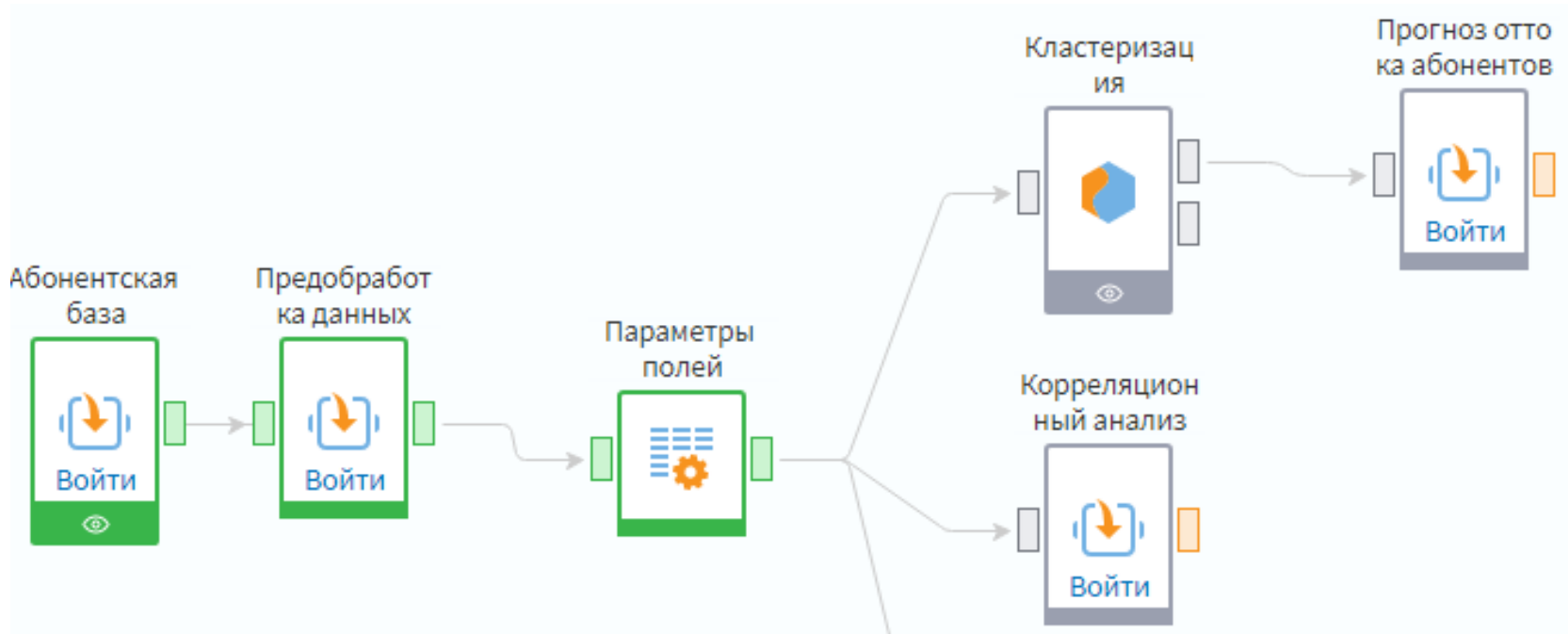
Структура ХД в Deductor



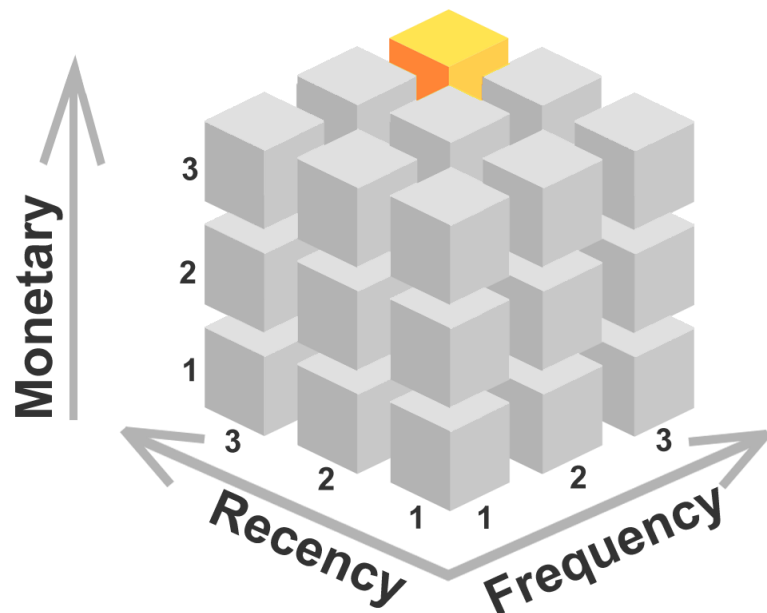
Преимущества ХД

- высокая скорость получения данных из хранилища;
- автоматическая поддержка непротиворечивости данных;
- возможность получения и сравнения срезов данных;
- наличие удобных средств для просмотра данных в хранилище;
- обеспечение целостности и достоверности хранящихся данных.

Визуальное проектирование компонент сценария

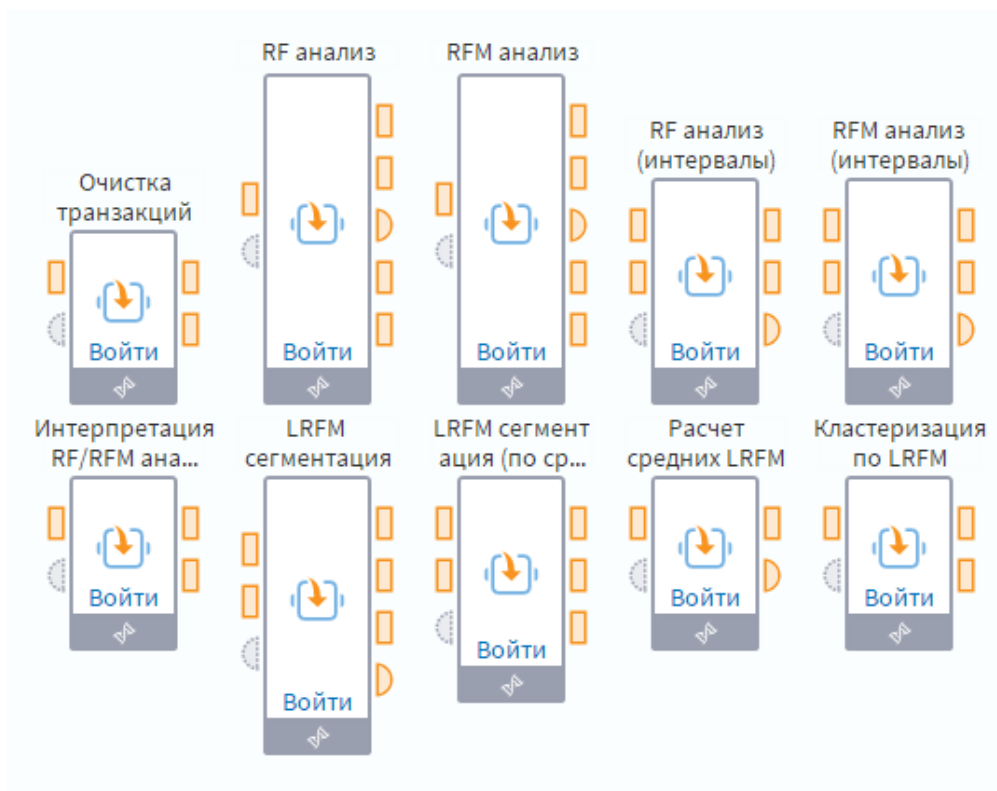


Методы и алгоритмы



- RF
- RFM
- LRFM
- Кластеризация по LRFM

Компоненты библиотеки



1. Очистка транзакций
2. RF
3. RFM
4. RF и 5. RFM (по интервалам)
6. Интерпретация RF/RFM
7. LRFM сегментация
8. LRFM (по средним)
9. Расчет средних LRFM
10. Кластеризация по LRFM

1. Входной набор данных

- Дата
- ID клиента
- Сумма

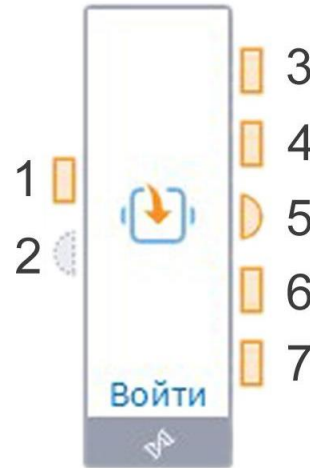
2. Управляющие переменных

- Дата актуальности
- Учёт клиентов с низким F
- Минимальное F
- Количество групп (R)
- Количество групп (F)

3. Сегментированные данные

- ID клиента
- RF код
- Группа Resency
- Группа Frequency

RF анализ



4. Выходной набор данных

- ID клиента
- Давность
- Количество

5. Переменные

- Дата актуальности
- Количество групп (R)
- Количество групп (F)
- Средняя давность
- Средняя частота

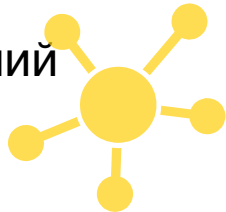
6. Границы квантования

- Идентификатор интервалов
- Номер интервала
- Нижняя граница
- Верхняя граница

Кластеризация по LRFM

1. Входной набор данных

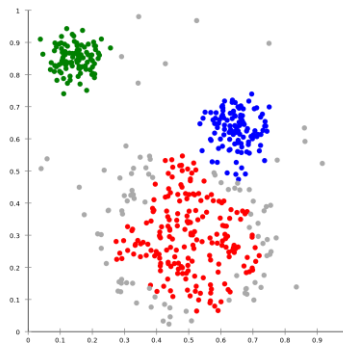
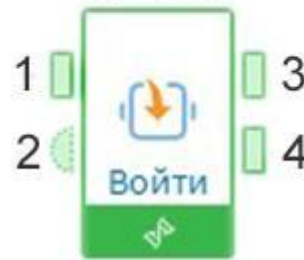
- ID клиента
- Длина отношений
- Давность
- Количество
- Сумма



2. Переменные для кластеризации

- Автоопределение числа кластеров
- Минимальное количество кластеров
- Максимальное количество кластеров
- Порог разделения кластеров
- Количество кластеров

Кластеризация по LRFM



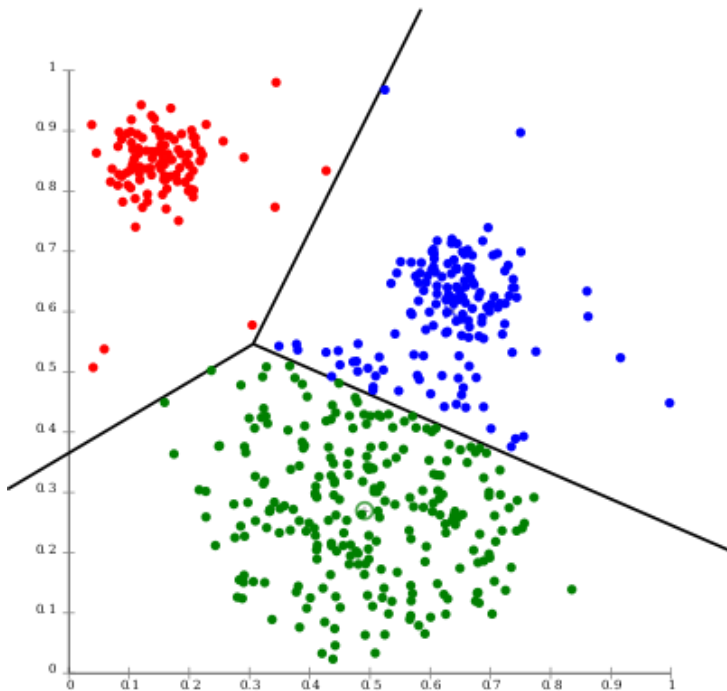
3. Выходной набор данных

- ID клиента
- Номер кластера
- Длина отношений
- Давность
- Количество
- Сумма

4. Центры кластеров

- Номер кластера
- Длина отношений
- Давность
- Количество
- Сумма

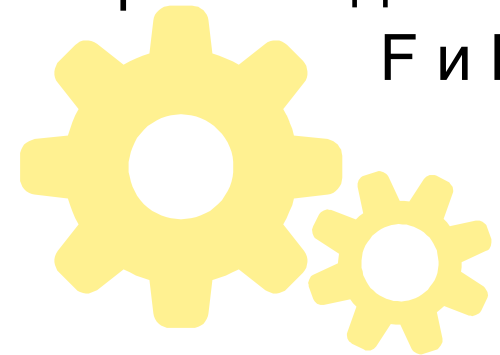
Кластерный анализ



- Объединение в устойчивые группы
- Использование алгоритмов k-means и g-means
- Сложности с автоматической интерпретацией

Функциональность библиотеки компонентов

1. Предварительная очистка транзакций
2. Произвольное число интервалов для параметров R, F и M
3. Экспертное задание границ интервалов
4. Просмотр и выгрузка отчётов
5. Просмотр результатов с помощью визуализаторов
6. Пользовательская интерпретация сегментов
7. Выбор количества сегментов
8. Автоматический расчет границ интервалов для R, F и M



Набор данных

компании Мегафон (база данных включала 2195000 записей об абонентах).



Основными полями, характеризующие абонентов были их средние (за определенный период) показатели, связанные с входящим и исходящим траффиком, также были сформированы производные переменные, отражающие доли потребления типов услуг.



Выводы

- Показана необходимость применения специализированного программного обеспечения и современных методов анализа данных для прогнозирования оттока клиентов в телекоммуникационных компаниях.
- Разработанная СППР в виде библиотеки компонентов на базе АП Loginot даст возможность повысить эффективность маркетинговых кампаний за счет точного сегментирования клиентских баз, прогнозирования оттока абонентов и их отклика на рассылку о новых тарифах.



Спасибо за внимание!