

Студ. К.А. Ахраменок  
Науч. рук. доц., к.т.н. Н.А. Жилияк  
(кафедра информационных технологий, БГТУ)

## ПРИМЕНЕНИЕ СКОРИНГА В РАБОТЕ С BIGDATA

BigData — термин, появившийся совсем недавно. Большими данными принято называть огромные массивы информации со сложной неоднородной и\или неопределенной структурой [3]. Отсюда возникает проблема обработки, структурирования и извлечения полезной информации из больших наборов данных.

Цель нашей работы: проанализировать литературы по вопросу применения скоринга в работе с BigData, выявить виды скоринга, которые могут быть взяты из банковского опыта и применены к обработке больших объемов данных.

Сегодня термин BigData, как правило, используется для обозначения не только самих массивов данных, но также инструментов для их обработки. Главные характеристики, отличающие BigData от другого рода данных – три V: volume, velocity, variety [3]. Первая – большие объемы, вторая – необходимость в быстрой обработке и высокая скорость накопления этих данных, третье – разнообразие. Исходя из характеристик BigData, было принято решение проанализировать возможность использования скоринга в обработке больших объемов данных [3].

Скоринг (от англ. Scoring – подсчет очков в игре) – это модель классификации клиентской базы на различные группы, если неизвестна характеристика, разделяющая эти группы, но известны другие факторы, связанные с интересующей нас характеристикой [1]. В основе скоринговой система лежит предположение, что люди со схожими социальными показателями ведут себя одинаково. Таким образом, мы можем попытаться применить это предположение в работе с большими данными информации.

Посредством различных видов скоринга можно как классифицировать данные так и давать вероятностные прогнозы на происхождение того или иного события [2].

После анализа литературы было выяснено, что из всех существующих видов скоринга, для работы с большими объемами данных подходят только два: Application-scoring и Behavioral-scoring. Их описание применения в банковской сфере и возможность применения в работе с BigData приведены в таблице [1].

Таблица 1

Видскоринга	Применение в банковской сфере	Возможное применение в работе с BigData
application-scoring	Это самый распространенный и известный клиентам вид скоринга. В его основе лежат первичный сбор анкетных данных заемщика, их обработка компьютером и вывод результата: предоставлять заем или нет [2].	Каждому свойству исследуемого объекта присваивается вес. После этого вес всех свойств подсчитывается и на его основе выставляется оценка. Исходя из этой оценки, отнести объект к какой-либо группе.
behavioral-scoring	«скоринг поведения» - оценка наиболее вероятных финансовых действий заемщика. Такая система дает возможность прогнозировать изменение платежеспособности заемщика, корректировать установленные для него лимиты [3].	Может быть использован для решения задач прогнозирования на основе имеющихся данных за период времени. Данные будут проходить обработку посредством скоринговой модели, а на выходе мы будем получать предполагаемое поведение объекта.

На основе выше сказанного были выделены следующие этапы обработки больших объемов данных посредством скоринга:

1. определить интересующую характеристику;
2. провести сбор второстепенных сведений о объекте исследования;
3. разработать скоринговую модель на основе имеющихся данных;
4. провести обработку данных посредством скоринговой модели

Актуальность работы заключается в увеличении объемов информации, требующей обработки, недостатке качественных методов обработки больших объемов информации.

Научная новизна заключается в том, что до текущего момента не было предложено вариантов использования скоринга в работе с большими объемами данных.

В результате работы была проанализирована литература по теме исследования, были выявлены методы скоринга, которые можно применить в работе с BigData, были выделены этапы обработки больших объемов данных посредством скоринга.

#### ЛИТЕРАТУРА

1. The credit scoring toolkit: Theory and Practice for retail Credit Risk Management and Decision Automation/ Anderson R. [et al] // England: Oxford University Press. – 2007. – P. 8–13, 16–17.
2. Data Smart: Using Data Science to Transform Information into Insight / John W. Foreman [et al] // Wiley, UK. – 2013. – P. 2–5.
3. Bigdata простым языком / А.П. Благирев; Издательство АСТ – Москва, 2019. №1:27–32 с.