

УДК 336.774.3

ОСОБЕННОСТИ АНАЛИЗА И КЛАССИФИКАЦИИ ПРОЕКТНЫХ ДАННЫХ В ФИНАНСОВО-КРЕДИТНЫХ ОРГАНИЗАЦИЯХ



Д.М. Рагель

*к.э.н., доцент кафедры экономики
Белорусского государственного
университета информатики и
радиоэлектроники*

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
e-mail: ragel@bsuir.by*

Д.М. Рагель

В 2000 г. окончил экономический факультет Белорусского государственного университета информатики и радиоэлектроники по специальности «Экономическая информатика». В 2016 г. окончил аспирантуру Академии управления при Президенте Республики Беларусь. В 2018 г. защитил кандидатскую диссертацию. Научные интересы: майнинг данных в маркетинге, моделирование процессов, анализ данных, статистическое прогнозирование, макроэкономика.

Аннотация. В статье описывается подход к систематизации больших данных по профинансированным проектам на основании алгоритма минимизации среднеквадратичной ошибки разбиения массива. В публикации приводятся основные этапы реализации алгоритма, а также критерии по определению кластеров для классификации анализируемого массива данных.

Ключевые слова: анализ данных, проектное управление, большие данные, массив данных, среднеквадратическая ошибка, кластеризация массива, финансово-кредитные организации, кредит, данные, проектные показатели, управление портфелем проектов, эффективность.

Эффективность работы финансово-кредитных организаций – это качество кредитного портфеля организации на основании чего, особую важность приобретает анализ портфеля, отслеживания динамики реализации входящих в портфель проектов. Своевременная реакция на возникающие отклонения в эффективности реализации отдельных проектов непосредственно влияет на финансовое положение кредитной организации и ее рыночные возможности. Успешные финансово-кредитные организации в настоящее время формируют большие портфели проектов, по которым собираются данные различного качества в целях мониторинга их состояния. С учетом этого, правильно выстроенная система мониторинга данных является одним из ключевых факторов успешного развития деятельности финансово-кредитных организаций.

Анализ проектной деятельности компаний предполагает, помимо реализации базовых функций управления, последовательный сбор и систематизацию данных по всем этапам жизненного цикла. Если говорить, непосредственно, о последовательной реализации долгосрочного проекта с переходом от начального этапа, на котором осуществляется инвестирование, к финальному этапу полной окупаемости и выхода на запланированные производственные мощности, можно выделить четыре основных этапа анализа хода

инвестиционных проектов. В силу своей специфики, мониторинг каждого из этапов требует отслеживания своих специфических групп показателей, которые позволят наиболее полно отобразить особенности протекания того или иного этапа.

Предпроектный этап – это этап планирования, разработки различных планов и объединение их в единый бизнес-план проекта, в соответствии с которым будет осуществляться финансирование его реализации. На данном этапе составляются графики, утверждаются значения показателей, на основании которых в дальнейшем будет осуществляться мониторинг и проводиться корректировка хода реализации проекта.

Инвестиционный этап реализации проекта характеризуется началом финансирования капитальных затрат по проекту. На данном этапе контролируется непосредственно степень реализации графиков, связанных с освоением капитальных затрат и вводом объектов в эксплуатацию. Основные группы показателей, которые могут быть рассчитаны, связаны с вводом в эксплуатацию оборудования, а также основных фондов. Расчет и контроль по данному этапу может вестись на основании сопоставления фактических значений с плановыми по итогам контрольного периода (месяц, квартал, полугодие):

$$K_{\text{кп}} = \frac{K_{\text{факт.}}}{K_{\text{план}}}, \quad (1)$$

где $K_{\text{кп}}$ – коэффициент характеризующий выполнение запланированного значения показателя проекта;

$K_{\text{план}}$ – запланированное значение показателя освоения по группе капитальных затрат;

$K_{\text{факт}}$ – фактическое значение показателя, полученное в рассматриваемом периоде времени.

Следующим этапом является эксплуатационная стадия, которая начинается после подписания документов о вводе объектов в эксплуатацию. Данный этап, как правило, характеризуется началом выполнения основных видов деятельности и выходом на запланированные производственные мощности. Основное внимание на этой стадии сосредоточено на темпах использования, окупаемости, а также финансовых показателях, которые характеризуют общую эффективность и позволяют сделать выводы о соблюдении сроков выхода на запланированные производственные мощности. В данном случае основными показателями, которым следует уделить внимание и на основании которых можно выстроить систему контроля могут быть как натуральные (производство, реализация), так и финансовые, к примеру, чистая прибыль, выручка, рентабельность продаж, цены, а также показатели, характеризующие задолженность и возможности ее погашения.

С учетом особенностей механизма мониторинга, его периодичности, а также количества стадий и предусмотренного количества показателей по стадиям, в конце отчетных периодов накапливается достаточно большой массив данных, который возрастает на протяжении проектного жизненного цикла.

Итоговым и наиболее важным, с точки зрения аналитики, этапом является упорядочивание полученного массива данных, его классификация, систематизация, формулировка выводов о качестве портфеля профинансированных проектов. С учетом этого, на данном этапе целесообразно использование алгоритма, на основании которого может быть проведена кластеризация полученного массива данных. В случае больших данных кластеризация позволит в течение достаточно короткого промежутка времени сделать выводы об основных группах, полученных данных и за счет этого сформировать видение динамики качества профинансированного портфеля проектов.

Обобщение полученных данных может быть проведено на основании использования алгоритма минимизации среднеквадратичной ошибки разбиения массива. Данный тип кластеризации может быть применен для оптимального разбиения на группы. Основной задачей процесса разбиения данных на кластеры в данном случае является минимизация ошибок на основании соответствия средним значениям установленных для кластеров. Итог проведенной кластеризации – массив полученных данных соотнесенный с интервалами соответствия плановым значениям, на основании этого можно сделать выводы об оптимальности формирования портфеля проектов. С учетом этого, распределение имеющегося массива данных по группам может быть проведено следующим образом:

$$e^2 = \sum_{j=1}^k \sum_{i=1}^{n_j} \|x_i^{(j)} - c_j\|^2; \quad (2)$$

где e^2 – среднеквадратическая ошибка разбиения массива;
 j – порядковый номер кластера;
 i – порядковый номер переменной в рассматриваемом кластере;
 k – количество кластеров;
 n_j – количество значений в рассматриваемом кластере;
 $x_i^{(j)}$ – классифицируемая i -тая переменная кластера j ;
 c_j – точка со средним значением характеристики данных в рассматриваемом кластере j .

В зависимости от целей анализа и для удобства дальнейшей классификации алгоритм может состоять из следующих обязательных этапов:

1. Выбор изначальных k точек с наиболее типичными характеристиками проектов для определения изначальных центров кластеров;
2. Соотнесение всех имеющихся объектов с определенными на предыдущем этапе центрами кластеров;
3. Актуализация центров кластеров в соответствии с полученной фактической классификацией данных;
4. Дальнейшее соотнесение с центрами кластеров имеющихся данных, если изначальный критерий классификации данных не выполнен (переход ко второму этапу).

В целях оптимизации распределения проектных данных по кластерам для дальнейшего анализа может быть использована следующая интервальная шкала:

0 – 50% соответствия фактических значений плановым – неудовлетворительное качество портфеля проектов;

50 – 70% соответствия фактических значений плановым – портфель, который требует дополнительных мер по его нормализации;

70 – 90% соответствия фактических значений плановым – портфель удовлетворительного качества;

более 90% – оптимально профинансированный портфель проектов.

Предложенный подход позволяет оптимальным образом классифицировать массив данных по портфелю проектов и быстро реагировать на возникшие в течение отчетного периода проблемы с эффективностью их реализации. Алгоритм апробирован в процессе мониторинга портфеля проектов некоторых субъектов финансово-кредитной деятельности Республики Беларусь.

Список использованных источников

[1] Введение в анализ данных : учебник и практикум / Б.Г.Миркин. – Москва : Издательство Юрайт, 2020. – 174 с.

[2] Волнин В.А. Аналитическая фабрика: Как настроить финансовую аналитику под задачи бизнеса. – Альпина Паблишер, 2021. – 548 с.

[3] Мыльников Л.А. Статистические методы интеллектуального анализа данных. – БХВ-Петербург, 2021. – 119 с.

FEATURES OF ANALYSIS AND CLASSIFICATION OF PROJECT DATA IN FINANCIAL AND CREDIT ORGANIZATIONS

Dz.M. Rahel

*PhD, Associate Professor of the Department of
Economics, Belarusian State University of
Informatics and Radioelectronics*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus
e-mail: ragel@bsuir.by*

Abstract. The article describes an approach to systematization of big data on funded projects based on the algorithm for minimizing the mean square error of array partitioning. The publication provides the main stages of the implementation of the algorithm, as well as criteria for determining clusters for classifying the analyzed data array.

Keywords: data analysis, project management, big data, data array, standard error, array clustering, financial and credit organizations, credit, data, project indicators, project portfolio management, efficiency.