

DOI 10.58351/2949-2041.2024.11.6.021

УДК 004.83

Выжигин Александр Юрьевич, кандидат технических наук, доцент, заведующий кафедрой, Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации; кандидат технических наук, доцент, Российский технологический университет МИРЭА, г. Москва
Vyzhigin Alexander Yurievich, Candidate of Technical Sciences, Associate Professor, Head of Department, Russian Academy of National Economy and Public Administration under the President of the Russian Federation; candidate of technical sciences, associate professor, Russian Technological University MIREA, Moscow
SPIN-code:5086-0171
ResearcherID:T-9882-2018

Москалев Илья Сергеевич, студент
МИРЭА – Российский технологический университет, г. Москва
Ilya Sergeevich Moskalev, student,
MIREA – Russian Technological University, Moscow

Сидоренко Никита Павлович, студент,
МИРЭА – Российский технологический университет, г. Москва
Nikita Pavlovich Sidorenko, student,
MIREA – Russian Technological University, Moscow

Трубиенко Олег Владимирович,
кандидат технических наук, доцент, заведующий кафедрой,
МИРЭА – Российский технологический университет, г. Москва
Oleg Vladimirovich Trubienko, Candidate of Technical Sciences,
Associate Professor, Head of the Department,
MIREA – Russian Technological University, Moscow

МОДЕЛИРОВАНИЕ АНАЛИТИЧЕСКОГО ПРОЦЕССА РАСЧЕТА ЗАВИСИМОГО ПОКАЗАТЕЛЯ В МАССИВАХ ДАННЫХ MODELING OF THE ANALYTICAL PROCESS OF CALCULATION OF THE DEPENDENT INDICATOR IN DATA ARRAYS

Аннотация: Вся наша жизнь состоит из моделирования процессов обеспечения трудовой деятельности. Не то что день, час нельзя представить без моделирования какого-либо аналитического процесса, состоящего по большей степени из когнитивного анализа данных. Когнитивный анализ – процесс информационно-аналитической деятельности, главная суть которого состоит в анализе и представлении субъективной информации. В основном, так все и происходит, если речь не идет о специалистах по работе с большими данными.

Современные специалисты в области аналитики прибегают к моделированию аналитических процессов. Моделирование аналитических процессов – один из методов предиктивной аналитики, смысл которого заключается в создании формализованной модели описания расчета искомого показателя.

Научная новизна исследования заключается в следующем:

1. расширены возможности моделирования аналитических процессов за счет применения методов вычислительной математики;
2. разработана интеграция концепции машинного обучения для моделирования аналитических процессов;
3. разработан метод расчета истинных значений показателей на основе численного решения дифференциальных уравнений методом Рунге-Кутты второго порядка;



Цель работы – повышение эффективности моделирования аналитических процессов за счет применения методов вычислительной математики.

Метод или методология проведения работы: в статье использовались такие методы как: Рунге-Кутта второго порядка (один из методов приближенного решения дифференциального уравнения, относящийся к классу методов Рунге-Кутта); предиктивная аналитика (процесс прогнозирования информации на основе полученных данных); вычислительная математика (комплекс методов для проведения приближенных расчетов).

Результаты: разработана модель аналитического мониторинга смертности граждан Российской Федерации.

Область применения результатов: полученные результаты могут использоваться работниками всех государственных учреждений Российской Федерации.

Abstract: Our whole life consists of modeling of processes of work activity support. Not just a day, an hour cannot be imagined without modeling any analytical process, which consists mostly of cognitive data analysis. Cognitive analysis is a process of information-analytical activity, the main essence of which is to analyze and present subjective information. That's basically how it works, unless we're talking about big data practitioners.

Modern analytics professionals resort to analytical process modeling. Analytical process modeling is one of the methods of predictive analytics, the meaning of which is to create a formalized model to describe the calculation of the desired indicator.

Scientific novelty of the research consists in the following:

- 1) the possibilities of modeling analytical processes are extended by applying the methods of computational mathematics;
- 2) integration of machine learning concept for modeling analytical processes is developed;
- 3) a method for calculating the true values of indicators based on the numerical solution of differential equations by the Runge-Kutta method of the second order has been developed;

Work purpose

increasing the efficiency of modeling analytical processes through the use of computational mathematics methods.

Methodology

The methods used in the article were: Runge-Kutta second order (one of the methods for approximate solution of a differential equation, belonging to the class of Runge-Kutta methods); predictive analytics (the process of predicting information based on the obtained data); computational mathematics (a set of methods to perform approximate calculations).

Results: a model of analytical monitoring of mortality of citizens of the Russian Federation has been developed.

Practical implications

The obtained results can be used by employees of all state institutions of the Russian Federation.

Ключевые слова: экспертная оценка, предиктивная аналитика, модель, моделирование, моделирование аналитических процессов, аналитика, язык программирования Python, интерполяция функции, вычислительная математика.

Keywords: peer review, predictive analytics, model, modeling, analytical process modeling, analytics, Python programming language, function interpolation, computational mathematics.

Введение

Современный мир сложно представить без моделирования процессов решения задач профессиональной деятельности. Это вполне логично, ведь не проводить предиктивный анализ событий – провал в работе, общении, а также возможность нанесения наибольшего количества вреда злоумышленниками. Несмотря на имеющиеся системы поддержки принятия решений (например, СППР «Игла»), в мире ведутся научные работы в области моделирования аналитических процессов.



Для решения поставленной проблемы было проведено исследование, цель которого – повышение эффективности моделирования аналитических процессов за счет применения методов вычислительной математики.

Для достижения поставленной цели были решены следующие задачи:

- проведен обзор и анализ существующих моделей аналитических процессов;
- проведена интеграция концепции машинного обучения для моделирования аналитических процессов;
- описана интеграция методов вычислительной математики для моделирования аналитических процессов;
- обоснована пригодность метода Рунге-Кутты второго порядка для решения аналитических задач;
- описана формализованная модель итогового показателя;
- обоснована корректность формализованной модели;
- проведено описание работы модели аналитических процессов;
- проведено тестирование работоспособности модели;

Анализ существующих моделей аналитических процессов

Для того, чтобы приступить к исследованию следует изучить уже имеющиеся решения для подобного рода задач (табл. 1).

Модель Портера "Пяти сил" – аналитическая модель, разработанная Майклом Портером, которая позволяет анализировать конкурентные силы в отрасли и определить уровень привлекательности отрасли для инвестирования или конкуренции.

SWOT-анализ – аналитическая модель, которая позволяет оценить сильные и слабые стороны компании, а также возможности и угрозы, с которыми компания сталкивается на рынке.

Модель Бостонской консалтинговой группы (BCG) – аналитическая модель, которая помогает определить стратегическую позицию продуктов компании на рынке и принять решение о дальнейших инвестициях.

Анализ цепочки создания стоимости (Value chain analysis) – аналитическая модель, которая позволяет разбить деятельность компании на отдельные звенья и определить, какие этапы создают большую добавленную стоимость.

Модель прогнозирования продаж – аналитическая модель, использующая статистические методы для прогнозирования объемов продаж в будущем на основе исторических данных, маркетинговых и экономических показателей.

Таблица 1

Анализ существующих моделей аналитических процессов

| Модели аналитических процессов | + | - |
|--------------------------------|--|--|
| Пять сил | 1. Позволяет оценить конкурентную среду предприятия. 2. Помогает определить стратегические преимущества и угрозы на рынке. 3. Позволяет выявить потенциальные риски и возможности для бизнеса. | 1. Может не учитывать индивидуальные особенности отрасли или региона. 2. Не всегда учитывает изменения внешней среды и быстрые технологические изменения. 3. Требуется дополнительного анализа для принятия конкретных стратегических решений. |
| SWOT-анализ | 1. Помогает выявить внутренние и внешние факторы, влияющие на бизнес. 2. Позволяет определить | 1. Может быть слишком общим и не давать конкретных рекомендаций для развития предприятия. 2. Требуется дополнительной оценки |



| | | |
|-------------------------------|---|--|
| | <p>сильные и слабые стороны предприятия.</p> <p>3. Помогает выявить возможности и угрозы для развития бизнеса.</p> | <p>и анализа для принятия стратегических решений.</p> <p>3. Исключает возможность учета изменений внешней среды и быстрых технологических изменений.</p> |
| Цепочка создания стоимости | <p>1. Позволяет выявить основные этапы процесса создания стоимости в компании.</p> <p>2. Помогает оптимизировать процессы и улучшить эффективность бизнеса.</p> <p>3. Позволяет выявить возможности для увеличения прибыли и сокращения издержек.</p> | <p>1. Может быть сложным и требовать больших затрат времени и ресурсов.</p> <p>2. Не всегда учитывает индивидуальные особенности компании или отрасли.</p> <p>3. Требуется дополнительного анализа для принятия конкретных стратегических решений.</p> |
| Модель прогнозирования продаж | <p>1. Помогает предсказать спрос на продукцию или услуги компании.</p> <p>2. Позволяет более точно планировать производство и закупки материалов.</p> <p>3. Помогает оптимизировать маркетинговые стратегии и увеличить прибыль.</p> | <p>1. Могут быть ошибки в прогнозе из-за изменений внешней среды или конкуренции.</p> <p>2. Требуется постоянного обновления данных и анализа для достижения точных прогнозов.</p> <p>3. Не всегда учитывает поведение потребителей и их предпочтения.</p> |

По данным из табл.1 делаем вывод, что лучшей аналитической моделью является модель SWOT-анализ. Несмотря на равное количество плюсов и минусов, данная модель выигрывает по содержанию своих преимуществ.

Интеграция машинного обучения в моделирование аналитических процессов

Машинное обучение в данном исследовании – один из ключевых моментов в рассматриваемом научном исследовании. Машинное обучение – анализ данных, суть которого заключается в выделении в таблице факторов, зависящих от итогового показателя, а также выявлении функциональной зависимости значащих столбцов для проведения дальнейших расчетов. Функциональная зависимость, как правило, рассчитывается при помощи созданной формализованной модели. Формализованная модель – математический объект, описывающий взаимодействие зависимых переменных между собой. То есть, по-простому говоря, формула. Любая формула (например, сила тяжести) является формализованной моделью. Для их создания ничего сверхъестественного делать не нужно. Нужно всего лишь уметь рассуждать. Обратимся к датасету (таблице) для объяснения всего вышесказанного на примере (рис. 1).

| | Родившиеся | Умершие | в том числе дети до 1 года | Естественный прирост, убыль (-) | Браки | Разводы |
|---|------------|---------|----------------------------|---------------------------------|--------|---------|
| 1 | 110674 | 103916 | 382 | 6758 | 84013 | 44807 |
| 2 | 113318 | 115684 | 392 | -2366 | 96167 | 44158 |
| 3 | -2644 | -11768 | -10 | 0 | -12154 | 649 |
| 4 | 9.2 | 8.7 | 3.4 | 0.5 | 7 | 3.7 |
| 5 | 9.5 | 9.7 | 3.3 | -0.2 | 8.1 | 3.7 |
| 6 | 96.8 | 89.7 | 103 | 0 | 86.4 | 100 |

Рис. 1 Таблица по демографии населения Российской Федерации за январь-июнь 2023 года



На рис. 1 изображена таблица с данными по смертности населения с января по июнь 2023 года. Задача моделирования аналитического процесса по данной таблице заключается в определении нужного нам показателя для создания практико-теоретических расчетов, а также влияющих на него факторов. Наш показатель (Y) – умершие. То есть сумма умерших людей. На данный показатель влияют два фактора (X): Естественный прирост, убыль (-) и в том числе дети до 1 года. Данный показатель учитывает значения таким образом, что учет ведется тогда, когда хотя бы есть какие-то значения в одном из двух факторов. Следовательно,

$$Y = X1 + X2$$

$$Y = \text{Умершие}$$

$$X1 - \text{Естественный прирост, убыль(-)}$$

$$X2 - \text{в том числе дети до 1 года}$$

Интеграция методов вычислительной математики

С машинным обучением разобрались. Но, чтобы создать моделирование процесса, необходимо еще грамотно интегрировать математический аппарат в саму модель. С этой задачей легко справится вычислительная математика. Вычислительная математика – раздел прикладной математики, отвечающий за приближенный подсчет значений в фактических данных. Из данной дисциплины был взят математический аппарат под названием численное решение дифференциального уравнения методом Рунге-Кутты второго порядка. Помимо своего основного предназначения (решение задачи Коши в дифференциальных уравнениях), данный метод используется в моделировании процессов, решении задач оптимизации, предиктивной аналитике и т.д.

Из рис. 1 мы выделили нашу выборку данных, необходимую для исследования. Теперь необходимо составить сами уравнения (рис. 2).

$$\frac{dk_1}{dx} = 382x^5 + 392x^4 - 10x^3 + 3.4x^2 + 3.3x + 103$$

а)

$$\frac{dz_1}{dx} = 6758x^5 - 2366x^4 + 0.5x^2 - 0.2x$$

б)

Рис. 2 а, б дифференциальные уравнения

Зависимая переменная (X) – значение месяца (1,...,6). Уравнения составлены. Остается их решить методом Рунге-Кутты второго порядка (рис. 3).

$$y_{i+1} = y_i + \frac{h}{2} \cdot (K_1 + K_2) \quad i = 0, 1, 2, \dots$$

где $K_1 = f(x_i, y_i),$
 $K_2 = f(x_i + h, y_i + h \cdot K_1)$

Рис. 3. Метод Рунге-Кутты второго порядка

После запрограммированных расчетов получаем новые данные (рис. 4).



| | Умершие | в том числе дети до 1 года | Естественный прирост, убыль (-) |
|---|-------------|----------------------------|---------------------------------|
| 1 | 2450 | 353 | 2097 |
| 2 | 67289 | 6143 | 61146 |
| 3 | 484868 | 37767 | 447101 |
| 4 | 1.98884e+06 | 142299 | 1.84654e+06 |
| 5 | 5.9704e+06 | 404978 | 5.56542e+06 |
| 6 | 1.46925e+07 | 960669 | 1.37318e+07 |

Рис. 4. Результат итоговых вычислений

Отсюда мы делаем вывод, что методы приближительных вычислений играют огромную роль в прикладных науках, связанных с анализом данных. Вычислительная математика потому называется вычислительной, что она несет в себе прогнозирование того или иного показателя. По данным из рис. 4 мы видим, что с марта по июнь 2023 года прогнозировалось по методу Рунге-Кутты второго порядка колоссальное количество смертей. Более чем в 100 раз отличаются прогнозируемые результаты от реальных. Следовательно, можно сказать, что ситуация была благосклонной для уменьшения количества смертности в связи с достойными социально-экономическими мерами поддержки государства. Оценка реального и практического результатов помогают составить переходный процесс от одного периода к другому, что являет собой неотъемлемую часть работы в области науки о данных.

Заключение

Проведенные эксперименты доказали эффективность методов вычислительной математики при моделировании аналитических процессов. Цель работы была достигнута. Проведена интеграция концепции машинного обучения для моделирования аналитических процессов. Описана интеграция методов вычислительной математики для моделирования аналитических процессов. Обоснована пригодность метода Рунге-Кутты второго порядка для решения аналитических задач. Описана формализованная модель итогового показателя. Обоснована корректность формализованной модели. Проведено описание работы модели аналитических процессов. Проведено тестирование работоспособности модели.

Список литературы:

1. Актуальные проблемы развития бизнес-анализа в условиях ориентации на инновационный путь развития: монография / под ред. В.И. Бариленко. – М.: Финансовый университет, 2012.
2. Баканов М.И., Мельник М.В., Шеремет А.Д. Теория экономического анализа: учеб. / под ред. М.И. Баканова. – 5-е изд., перераб. и доп. – М.: Финансы и статистика, 2008.
3. Бебнева Е.В. Бухгалтерский учет: учеб, пособие. – М.: Эксмо, 2008.
4. Трегуб И.В. Имитационное моделирование: учеб, пособие. – М.: Финакадемия, 2007.
5. Ченг Ф. Ли, Финнерти Дж. И. Финансы корпораций: теория, методы и практика / пер. с англ. – М.: ИНФРА-М, 2000.
6. Шеремет А.Д., Негашев Е.В. Методика финансового анализа деятельности коммерческих организаций. – М.: ИНФРА-М, 2003

References:

1. Actual problems of business analysis development in the conditions of orientation to the innovative way of development: monograph / edited by V.I. Barilenko. – Moscow: Financial University, 2012.
2. Bakanov M.I., Melnik M.V., Sheremet A.D. Theory of economic analysis: textbook / edited by M.I. Bakanov. – 5th ed., revision and supplement. – M.: Finance and Statistics, 2008.
3. Bebneva E.V. Accounting: textbook. – Moscow: Eksmo, 2008.
4. Tregub I.V. Simulation modeling: textbook. – Moscow: Finakademiya, 2007.
5. Cheng F. Lee, Finnerty J. I. Finance of corporations: theory, methods and practice / per. from Engl. – M.: INFRA-M, 2000.
6. Sheremet A.D., Negashev E.V. Methodology of financial analysis of commercial organizations. – M.: INFRA-M, 2003

