

DOI: 10.51790/2712-9942-2024-5-4-05

## ПРОГРАММНЫЙ МОДУЛЬ ПРОГНОЗИРОВАНИЯ ПЛАНОВОЙ ПОТРЕБНОСТИ В МЕДИЦИНСКОЙ ПРОДУКЦИИ

**В. Д. Крыжановский**

Воронежский государственный университет, г. Воронеж, Российская Федерация  
ORCID: <http://orcid.org/0009-0003-6145-1649>, [✉ vlad0kryzh@gmail.com](mailto:vlad0kryzh@gmail.com)

*Аннотация:* данная статья посвящена оценке достаточности имеющихся данных для факторного анализа с целью создания формулы плановой потребности в рамках годовой заявки с целью снижения трудовых затрат персонала. Для реализации использовался язык программирования C# и средства Microsoft Excel. В наборе данных рассматриваются такие данные, как численность льготополучателей региона, страдающих сахарным диабетом, среднегодовое изменение такой численности, количество получаемого препарата, половой состав пациентов, страдающих диабетом, их возрастной состав и количество пациентов, разделенных на три группы по количеству льготных категорий: с одной категорией, двумя категориями и тремя и более категориями. В работе используются упрощения, не оказывающие существенного влияния на получаемые значения в силу специфики предметной области. Полученные результаты указывают на перспективность реализации данного метода, планируется создание и обучение нейронной сети для формирования более точных формул расчета плановой потребности. Также сделан вывод о необходимости изучить вопрос автоматизированного создания модели для факторного анализа с целью проверки качества текущей модели. Для позиций, имеющих достаточный объем данных для обучения нейронной сети, эффективно предсказывающей плановую потребность, имеет смысл использовать ее, а не рассматриваемое решение.

*Ключевые слова:* факторный анализ, плановая потребность, нейронные сети, здравоохранение.

*Для цитирования:* Крыжановский В. Д. Программный модуль прогнозирования плановой потребности в медицинской продукции. *Успехи кибернетики*. 2024;5(4):40–44. DOI: 10.51790/2712-9942-2024-5-4-05.

*Поступила в редакцию:* 09.10.2024.

*В окончательном варианте:* 11.11.2024.

## DEMAND FORECASTING SOFTWARE FOR MEDICAL PRODUCTS

**V. D. Kryzhanovsky**

Voronezh State University, Voronezh, Russian Federation  
ORCID: <http://orcid.org/0009-0003-6145-1649>, [✉ vlad0kryzh@gmail.com](mailto:vlad0kryzh@gmail.com)

*Abstract:* this study evaluates the sufficiency of available data for factor analysis to develop an equation for calculating planned annual requirements, aiming to reduce labor costs. The implementation uses C# and Microsoft Excel. The dataset includes such variables as the number of patients with diabetes mellitus in the region, the average annual change in their numbers, the number of drugs received, the gender and age of the patients, and the distribution of patients into three categories of government benefits they are entitled to. Simplifications are applied where they do not significantly impact the results. The findings highlight the potential of this approach. We are going to develop and train a neural network to obtain more precise equations. The study also emphasizes the need to explore automated model creation for factor analysis to assess the current model's quality. For cases where sufficient data is available to train a neural network that reliably predicts planned requirements, adopting neural network solutions is recommended over the current approach.

*Keywords:* factor analysis, demand, neural networks, healthcare.

*Cite this article:* Kryzhanovsky V. D. Demand Forecasting Software for Medical Products. *Russian Journal of Cybernetics*. 2024;5(4):40–44. DOI: 10.51790/2712-9942-2024-5-4-05.

*Original article submitted:* 09.10.2024.

*Revision submitted:* 11.11.2024.

### Введение

Сбор плановой потребности в медицинской продукции — ответственный процесс, позволяющий сотрудникам центрального органа здравоохранения региона произвести государственную закупку

препаратов, необходимых льготным пациентам региона. В рамках этого процесса производится анализ плановой потребности, заключающийся в указании каждой медицинской организацией региона, выписывающей льготные рецепты, количества медицинской продукции для каждой из возможных для заявления позиций единого справочника-каталога лекарственных препаратов, медицинских изделий, лечебного питания и иной медицинской продукции на тот период, на который собирается потребность в рамках лимитов, установленных центральным органом здравоохранения региона Российской Федерации.

В рамках данной статьи будут рассматриваться препараты, содержащие в себе одно действующее вещество — «метформин» (гидрохлорид метформина) — в дозировке 850 мг. В силу специфики задачи, решаемой в данной статье, персонифицированная потребность рассматриваться не будет.

Данный препарат входит в перечень «жизненно необходимых и важнейших лекарственных препаратов». С точки зрения фармакодинамики, препарат снижает гипергликемию, но не приводит к развитию гипокликемии. Препарат используется для профилактики сахарного диабета у пациентов, страдающих предиабетом, а также назначается при сахарном диабете второго типа [1].

Метформин — один из самых назначаемых противодиабетических препаратов с многокомпонентным органным и системным действием, эффективно и долгосрочно контролирующий уровень глюкозы в крови и предоставляющий дополнительные преимущества пациентам с сахарным диабетом второго типа в отношении контроля некоторых факторов риска сердечно-сосудистых заболеваний (избыточная масса тела, дислипидемия, инсулинорезистентность и состояния, ассоциированные с ней), снижения сердечно-сосудистой смертности, риска некоторых форм рака.

Возможности применения метформина при СПКЯ (синдром поликистозных яичников) у взрослых и подростков, при НАЖБП (неалкогольная жировая болезнь печени) у детей закреплены в актуальных клинических рекомендациях. В остальном же значительный потенциал плейотропных эффектов препарата в области онкологии, неврологии, офтальмологии, антивозрастной медицины требует дальнейшего изучения [2].

Ранее была реализована рекуррентная нейронная сеть с использованием ДКП («долгая краткосрочная память», англ. «Long Short-Term Memory», или LSTM), позволяющая прогнозировать персонифицированную потребность в данном препарате более точно, чем у «наивного» алгоритма [1].

В силу технических особенностей используемого программного решения точность предсказания может изменяться даже при обучении на одних и тех же данных, что может затруднить промышленную эксплуатацию данного метода прогнозирования.

Также важным моментом использования нейронных сетей для предсказания плановой потребности является потребность в большом количестве данных о назначении препарата.

В частности, для обучения упомянутых выше нейронных сетей были использованы имеющиеся на тот момент данные о льготном лекарственном обеспечении населения за период 01.01.2021–16.11.2023 — два набора по 1050 дней. Для обучения используются данные о 960 днях, для тестирования — о последних 90 днях.

В регионе 1 фактическая потребность составила 136260 единиц препарата, а расчетная потребность на тот же период — 149520 единиц, что означает превышение фактического расхода на 9,73% [1].

В регионе 2 фактическая потребность составила 282150 при расчетной в 294540, что означает превышение на 4,39%.

Кроме того, следует отметить, что реализованная ранее нейронная сеть корректно предсказала специфические колебания предъявленной потребности в препарате, что, однако, не позволило ей корректно предсказать всплеск потребности в препарате, вызванный увеличением количества обратившихся за препаратом пациентов. Это также может быть критично для промышленной эксплуатации системы.

Таким образом, данная статья посвящена проверке возможности эффективно использовать имеющиеся статистические данные для прогнозирования плановой потребности с использованием факторного анализа.

### **Реализация**

В качестве выборки, на которой производится выведение формулы назначения препарата, используются обезличенные данные реальной предъявленной потребности населения двух субъектов Российской Федерации за период 01.01.2021–31.08.2024.

В качестве периода, на котором производится расчет, выберем данные за 01.01.2021–31.12.2023. В качестве данных, которые будут сверяться с прогнозируемым количеством, — период 01.01.2024–31.08.2024. То есть был произведен расчет потребности на год, его домножение на 2/3 и сравнение полученного числа с фактически предъявленной и плановой потребностью. Использование домножения может казаться сомнительным решением, однако в силу специфики потребности в препарате может рассматриваться как не ведущее к существенным ошибкам.

Альтернативным решением является отказ от использования промежутка с 01.01.2024 года и использование для проверки эффективности данные за 2023 год, однако это сократило бы тот набор данных, который используется для вычисления коэффициентов. При этом отбрасывание трети имеющихся данных рассматривается как путь, оказывающий более существенное влияние на результат.

В качестве параметров формулы потребности региона в препарате рассматриваются следующие параметры:

- 1) численность льготополучателей региона, страдающих сахарным диабетом;
- 2) среднегодовое изменение численности пациентов, страдающих сахарным диабетом;
- 3) половой состав пациентов, страдающих сахарным диабетом (женщины страдают им чаще);
- 4) возраст пациентов, страдающих сахарным диабетом;
- 5) количество имеющих у пациента льготных категорий и диагнозов;
- 6) количество получаемого препарата.

Основная идея факторного анализа заключается в разложении некоторой результирующей переменной — в нашем случае, количества заявленного препарата — на множество компонентов, каждый из которых характеризует влияние того или иного фактора.

Основными предпосылками для применения данного метода являются:

- 1) определение модели путем логического анализа;
- 2) объединение влияющих на результат показателей, если это возможно;
- 3) применение метода производится к такому отрезку времени, за который не происходит существенного изменения свойств объекта — в нашем случае, это множество пациентов и их заболеваний.

Сначала был произведен факторный анализ количества назначений в разрезе количества имеющихся льготных категорий и влияния их количества на вероятность назначения препарата (когда одна из категорий связана с диабетом). После этого был произведен факторный анализ количества назначенного препарата в единицах измерения препарата в зависимости от возраста, среднего количества назначаемого препарата, полового состава и среднего возраста пациентов.

### Полученные результаты

Регион 1:

$$G = (p_1 \times 0,95 + p_2 \times 1 + p_3 \times 1,2) \times 60 \times cW \div cM \div |m - 60|.$$

Регион 2:

$$G = (p_1 \times 1 + p_2 \times 0,95 + p_3 \times 1,15) \times 60 \times cW \div cM \div |m - 66|.$$

$G$  — плановая потребность в препарате,  $p_1$  — количество пациентов с одной категорией с учетом ожидаемого изменения численности пациентов на основе данных за предыдущий период;  $p_2$  — количество пациентов с двумя категориями с учетом ожидаемого изменения численности пациентов на основе данных за предыдущий период;  $p_3$  — количество пациентов с тремя и более категориями с учетом ожидаемого изменения численности пациентов на основе данных за предыдущий период;  $m$  — средний возраст пациентов, обратившихся за льготным лекарственным обслуживанием в прошлом году;  $cW$  и  $cM$  соответственно — количество женщин и мужчин среди обратившихся пациентов. Множитель 60 обозначает 60 таблеток препарата в данной дозировке, а вычитаемое из  $m$  значение под знаком модуля — средний возраст льготополучателей, округленный до количества лет.

В рамках примерного сравнения это не имеет решающего значения, но в рамках промышленной эксплуатации необходимо предусмотреть некоторое решение для случая обращения значения под знаком модуля в ноль.

Введем допущение о том, что в 2024 году не появится новых льготополучателей. В сравнении с классическим методом сбора плановой потребности это не является отклонением от формальной

процедуры: потребность планируется на уже имеющихся в регионе льготополучателей. В рамках промышленной эксплуатации имеет смысл добавление параметров, указывающих на ожидаемое изменение количества пациентов с различным количеством льготных категорий и вызванного этим колебания среднего возраста.

Далее домножим полученное значение на  $2/3$  и сравним результат с фактически предъявленной потребностью.

При применении формулы расчета, полученной на данных региона 1, для предсказания потребности в регионе 1, полученная точность прогноза составила 81%. Для региона 2 — 74%.

При применении формулы расчета, полученной на данных региона 2, для предсказания потребности в регионе 1, полученная точность прогноза составила 71%. Для региона 2 — 79%.

Полученные результаты являются недостаточными для промышленной эксплуатации полученных формул, однако демонстрируют, что имеющийся набор данных охватывает факторы, достаточные для предсказания фактически предъявленной потребности.

Следует отметить, что реализованная ранее нейронная сеть на аналогичном наборе данных аналогичного состава, без учета 2024 года, демонстрировала превышение прогнозного фактического расхода на 9,73%. Для региона 2 аналогичное превышение составило 4,39%. Данное превышение может быть критично для дорогостоящих, наркосодержащих и имеющих малый срок годности препаратов.

### **Причины получения частично удовлетворительного результата**

Имеющиеся данные касаются исключительно препаратов, полученных в рамках льготного лекарственного обслуживания, то есть влияние препаратов, закупленных за счет пациента, не учитывается.

Отсутствие более подробных сведений о состоянии здоровья пациентов, данных их анализов, исполнения и неисполнения рекомендаций лечащих врачей также сказывается на эффективности прогнозирования потребности в медицинской продукции.

Также имеющаяся выборка содержит преимущественно заболевания, которые «дают» региональные и федеральные льготы части льготного лекарственного обеспечения, могут не содержать фактически имеющиеся заболевания и не содержат назначения, сделанные пациентам в стационарах.

Кроме того, использование множителя  $2/3$  к предсказываемой потребности для сравнения может быть некорректно, хотя данный препарат не имеет ярко выраженной сезонности применения. Также следует отметить, что возможной является недостаточная качественность самой модели.

Для позиций с большим количеством данных о льготном лекарственном обеспечении, вероятно, будет эффективнее применять иные решения, а реализация сбора персонифицированной потребности на текущем составе набора данных с помощью факторного анализа не представляется возможной.

Тем не менее, результат рассматривается как достаточно перспективный для дальнейшего совершенствования конкретной методологии формирования формулы плановой персонифицированной потребности в медицинской продукции.

В отличие от реализованного ранее метода формирования заявки, факторный анализ не требует наличия большого количества данных о произведенных назначениях, что позволяет применять предложенную методологию для более широкого списка позиций, доступных для заявки в рамках различных заявочных кампаний и с меньшей задержкой на накопление первичной информации.

В дальнейшем планируется произвести обучение нейронной сети для получения в качестве выходных параметров данной сети коэффициентов, необходимых для построения формулы составления плановой потребности, а также изучить вопрос автоматизированного создания модели для факторного анализа для проверки качества текущей модели.

## **ЛИТЕРАТУРА**

1. Крыжановский В. Д. Использование рекуррентных нейронных сетей для прогнозирования плановой потребности региона. *Актуальные проблемы прикладной математики, информатики и механики*. Воронеж: Научно-исследовательские публикации; 2023. С. 1411–1415.
2. Друк И. В., Ряполова Е. А. Метформин: обновленные рекомендации и плейотропный потенциал. *Терапия*. 2016;4(8):44–51.
3. *Архив Корнеллского университета*. Режим доступа: <https://arxiv.org/pdf/1909.00590v2.pdf>.

4. Tarkov M., Kozhushko O. Regression Analysis of Text Ranking Algorithms by Neural Networks. *Bulletin of the Novosibirsk Computing Center. Series: Computer Science*. 2016;39:63–70. DOI: 10.31144/bncc.cs.2542-1972.2016.n39.p63-70.
5. Терехов С. А. *Введение в байесовы сети*. М.: Московский инженерно-физический институт; 2003. 186 с.
6. Платформа TensorFlow: официальный сайт. Режим доступа: <https://www.tensorflow.org/>.
7. Шовин В. А., Гольдяпин В. В. Факторное моделирование с помощью нейронной сети. *Математическое моделирование и численные методы*. 2016;2:85–103. DOI: 10.18698/2309-3684-2016-2-85103.
8. Лиля В. Б., Пучков Е. В. Методология обучения рекуррентной искусственной нейронной сети с динамической стековой памятью. *Программные продукты и системы*. 2014;4:132–135. DOI: 10.15827/0236-235X.108.132-135.
9. Тетерин Д. А., Хабибулин Р. Ш., Гудин С. В. Обзор применения искусственных нейронных сетей в управлении социальными и экономическими системами. *Научные ведомости. Серия Экономика. Информатика*. 2018;45(3):574–583. DOI: 10.18413/2411-3808-2018-45-3-574-583.
10. Иберла К. *Факторный анализ*. М.: Финансы и статистика; 1980. 398 с.