

РАЗРАБОТКА МОДУЛЯ КОНФИГУРАЦИИ КРИОХРАНИЛИЩА ДЛЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ «КРИОБАНК» С ВОЗМОЖНОСТЬЮ ПРОГНОЗИРОВАНИЯ ЗАПОЛНЯЕМОСТИ КРИОХОЛОДИЛЬНИКОВ

Солодун Александр Евгеньевич
магистрант,

МИРЭА – Российский технологический университет.
Россия, г. Москва

Научный руководитель:

Старичкова Юлия Викторовна

кандидат технических наук, профессор.

МИРЭА – Российский технологический университет
Россия, г. Москва

АННОТАЦИЯ

Информационная система «Криобанк» является узкоспециализированным программным обеспечением. Оно разработано для ФГБУ «НМИЦ ДГОИ им. Дмитрия Рогачева» Минздрава России и является частью комплексной информационной системы медицинского центра. Рассматриваемый сервис информационная система «Криобанк» занимается сбором и хранением информации о конфигурации криохранилища, а также о сданном биоматериала донорами. Благодаря этой системе пользователи видят актуальную информацию о конфигурации криохолодильников, а также о количестве и видах хранящегося в них биоматериала. Однако в этой медицинской системе отсутствует функционал отвечающий за прогнозирование сданного в будущем количества биоматериала. Для этих целей был разработан внешний модуль, который бы позволили сотрудникам ФГБУ «НМИЦ ДГОИ им. Дмитрия Рогачева» иметь удобный инструментарий, благодаря которому можно будет делать прогнозы, что избавит медицинский центр от неожиданной нехватке биоматериала, которая бы повлекла за собой остановку трансплантационных отделений.

ANNOTATION

The Cryobank information system is a highly specialized software. It was developed for FSBI "NMIC DGOI im. Dmitry Rogachev" of the Ministry of Health of Russia and is part of the integrated information system of the medical center. The service in question, the Cryobank information system, collects and stores information about the configuration of the cryopreservation facility, as well as about the donated biomaterial by donors. Thanks to this system, users see up-to-date information about the configuration of cryocholoders, as well as about the amount and types of biomaterial stored in them. However, this medical system lacks the functionality responsible for predicting the amount of biomaterial delivered in the future. For these purposes, an external module was developed that would allow employees of the FSBI "NMIC DGOI im. Dmitry Rogachev " to have convenient tools, thanks to which it will be possible to make predictions, which will save the medical center from an unexpected shortage of biomaterial, which would entail the shutdown of transplant departments.

Ключевые слова: Информационная система; экстраполяция; машинное обучение.

Keywords: Information system; extrapolation; machine learning.

В ходе статьи будут рассмотрены алгоритмы для прогнозирования в контексте разрабатываемого модуля, после чего будет выбран лучший из приведенных, который лег в основу разработанного модуля конфигурации информационной системы «Криобанк».

Как выбрать правильный алгоритм для решения поставленной задачи. Существует четыре основных типа задач машинного обучения:

1. Обучение с учителем.
2. Обучение без учителя.
3. Обучение с частичным привлечением учителя.
4. Обучение с подкреплением.

Обучение происходит в три этапа:

- Анализ данных.
- Нахождение шаблонов.
- Предсказание на основе шаблонов.

Обучение с учителем строится на основе имеющихся данных. Изначально существует тренировочные данные и путем подгонки результатов находим оптимальную модель для прогнозирования. Такой подход наиболее распространенный, так как легче всего обучать программу с выбранными данными. Существует два типа задач:

- Задачи регрессии. При прогнозировании непрерывных значений, таких как цена валюты на бирже, или же как в рамках данной статьи: прогнозирование заполняемости криоохладильников, у таких задач нет конкретного ограничения значений, результатом могут быть любые числа. Визуальный пример регрессии на рисунке 1.

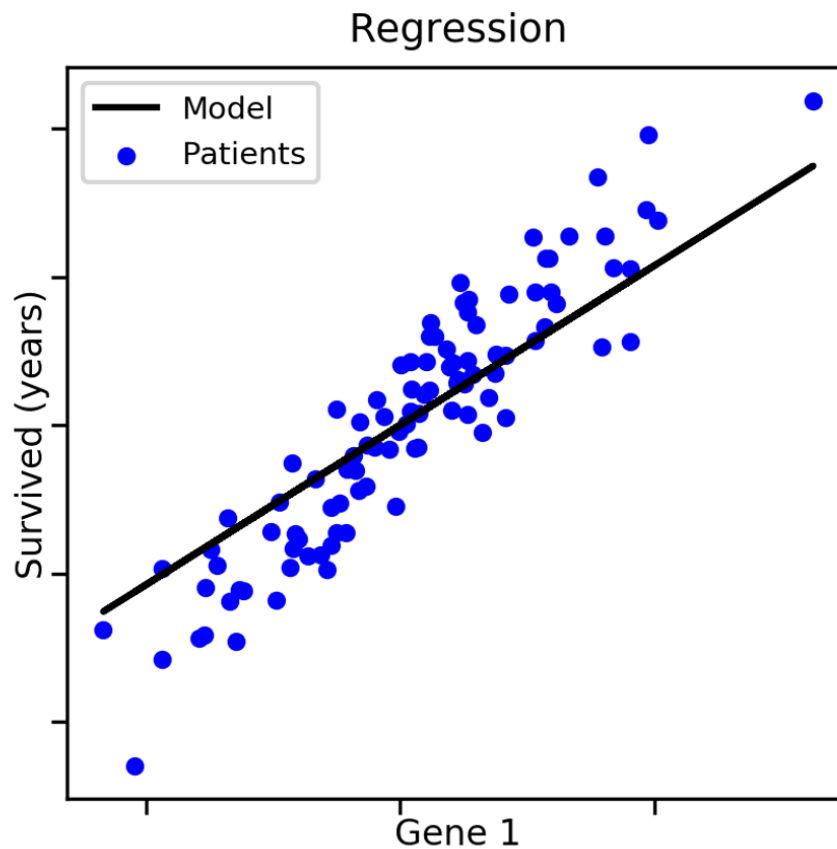


Рисунок 1 – визуализация регрессии

- Задачи классификации. Задачи с конечным числом ответов. На вопрос «Забьет ли команда гол» существует два ответа: да или нет, такой тип задач и называется задачами классификации. Визуальный пример классификации на рисунке 2.

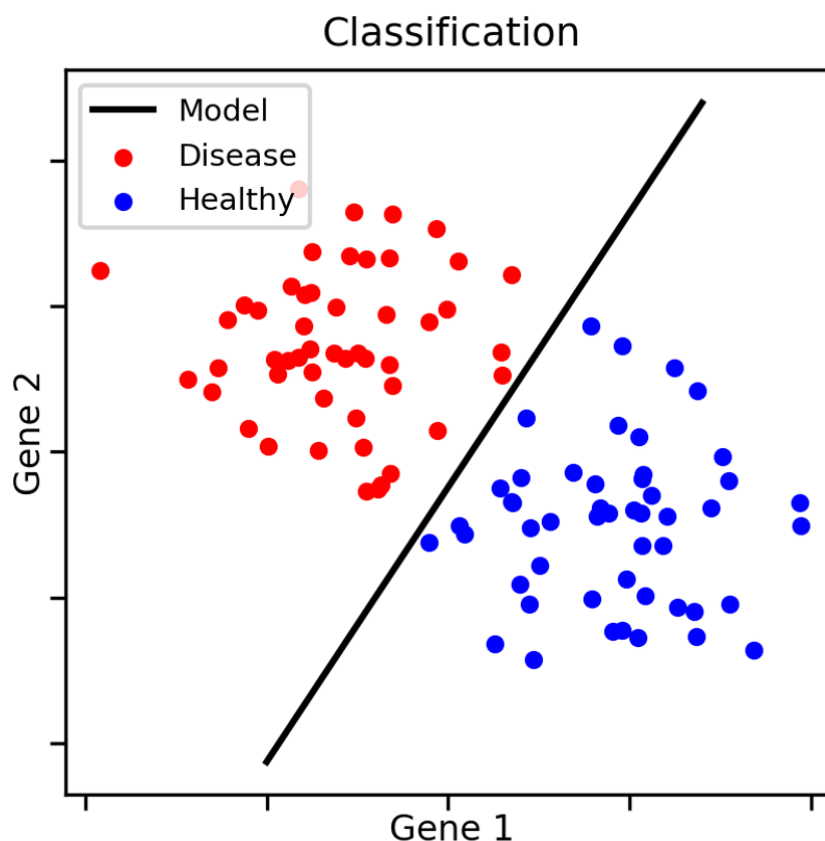


рисунок 2 – визуализация классификации

Обучение без учителя. При таком подходе к обучению машинного интеллекта, ключевым является нахождением всевозможных корреляций, а также группировка данных, так как изначально отмеченных и сгруппированных данных нет, как в случае в обучении с учителем. Пример такой постановки задачи: нужно сгруппировать животных в зоопарке исходя из их запахов. В таких задачах есть три способа достичь нужный результат, разберем на примере задачи выше про животных:

- Путем кластеризации – группировка животных в кластеры исходя из попарных сходств, рисунок 3.

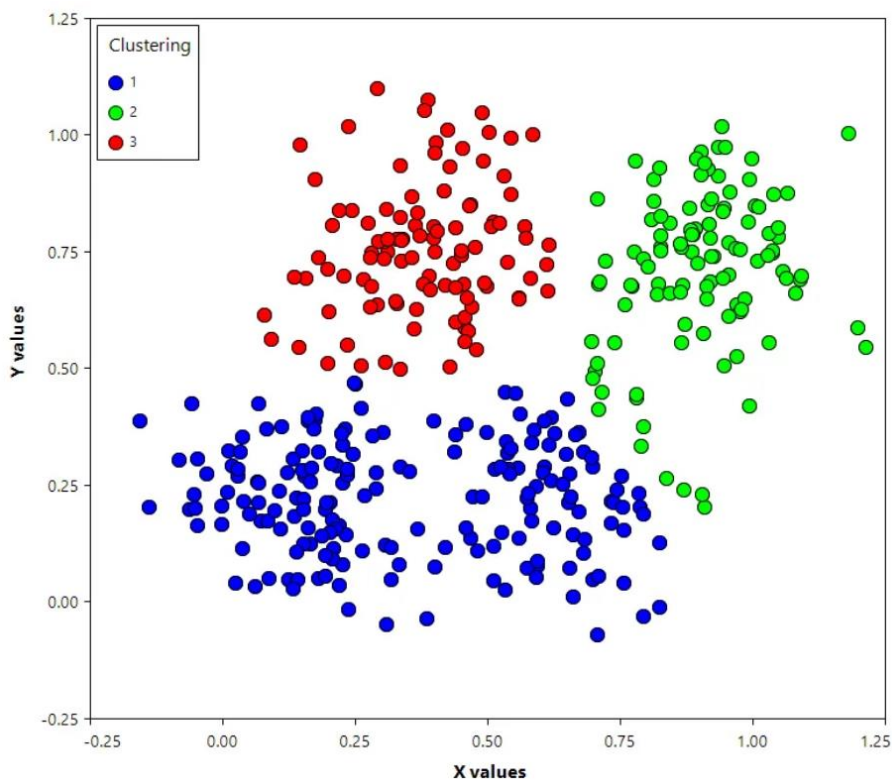


Рисунок 3 – визуализация кластеризации

- Путем поиска ассоциативных правил. Происходит поиск животных с одинаковыми характеристиками - признаками.
- Путем снижения размерности. Сутью данного подхода является с помощью математических преобразований отказаться от части признаков, не потеряв при этом нужную информацию.

Обучение с частичным привлечением учителя. В таком подходе в задаче существуют два типа данных: размеченные и неразмеченные. У неразмеченных данных происходит поиск общих признаков, у размеченных в добавок к поиску общих признаков присутствует известный целевой признак. Такого типа задачи считаются чем-то средним между теми двумя типам задач описанные ранее.

Обучение с подкреплением. Такой подход к обучению искусственного интеллекта считается самым сложным, но самым подающим надежды. Суть такого подхода в обучении на практике, в отличие от выше описанных подходов, в которых обучение происходит в пассивном режиме – передача данных и обнаружение корреляций. В

данном подходе необходимо научить машину взаимодействовать с внешней средой, благодаря чему поиск данных и выделение в них нужных признаков будет происходить без участия человека.

В рамках разрабатываемого модуля, задача, которую приходится решать – является задачей обучения с учителем типа регрессии.

В задачах регрессии ключевым является выбрать нужную функцию из всех, которая бы минимизировала потери. У регрессии есть родственные понятия: экстраполяция и интерполяция. Интерполяция – способ выбора функции из множества функций, которая бы проходила через заданные точки, в свою очередь экстраполяция – это предсказание выбранной функции проходящий через заданные точки, только уже вне интервала этих точек, изображено на рисунке 4.



Рисунок 4 – визуализация экстраполяции

Даны две точки x^* . Ближайшими к экстраполируемой точке (X_{k-1}, Y_{k-1}) , равны и (X_k, Y_k) , экстраполяция дает функцию:

$$Y(x^*) = Y_{k-1} + \frac{X^* - X_{k-1}}{X_k - X_{k-1}} (Y_k - Y_{k-1}),$$

является идентичной интерполяции, если $X_{k-1} < X^* < X_k$

В рамках задачи разрабатываемого модуля с функцией предсказания заполняемости криохолодильников как нельзя лучше подходит экстраполяция. На основе данных о количестве ранее доноров и сданного ими биоматериале можно прогнозировать в будущем количество сданного биоматериала, взяв за основу алгоритм экстраполяции.

Итак, в рамках данной статьи был выбран алгоритм, который лег за основу программного, отвечающего за предсказания заполнения криохолодильников в криохранилище в НМИЦ ДГОИ им. Дмитрия Рогачева.

Список литературы:

1. НМИЦ ДГОИ им. Дмитрия Рогачева, официальный сайт [Электронный ресурс]. – Режим доступа: <https://fnkc.ru/> (дата обращения: 25.05.2022)
2. Визуализация данных: основные правила, полезные приемы и инструменты [Электронный ресурс]. – Режим доступа: <https://www.owox.ru/blog/articles/data-visualization/> (дата обращения: 27.04.2022)
3. Старичкова Ю. В., Воронин К. А., Борисова Н. В., Масчан М. А., Румянцев А. Г. Подходы к интеграции комплекса программных средств управления процессами и сложноструктурированными медицинскими данными с медицинскими и лабораторными информационными системами в учреждениях здравоохранения. 2020
4. Хасти Тревор. - Основы статистического обучения: интеллектуальный анализ данных, логический вывод и прогнозирование. – 2020