

ПРОЕКТИРОВАНИЕ СПЕЦИАЛИЗИРОВАННОЙ КОМПЬЮТЕРНОЙ СИСТЕМЫ ДИАГНОСТИКИ СОСУДОВ ГОЛОВНОГО МОЗГА НА ПРИМЕРЕ ИШЕМИЧЕСКОГО ИНСУЛЬТА

Лобачева М.В., группа КСД-06м

Руководитель проф. Скобцов Ю.А.

Ишемический инсульт – наиболее частая форма острых нарушений мозгового кровообращения. Он представляет собой зону некроза, образовавшуюся вследствие грубых, стойких нарушений метаболизма нейрональных структур, возникающих в результате недостаточности кровоснабжения головного мозга. Ежегодно инсульт поражает около 6 млн. человек в мире [1]. Поэтому проблема своевременной диагностики инсультов в настоящее время является актуальной и представляет собой интерес для врачей различных специальностей.

Целью работы является разработка методов прогнозирования риска ишемического инсульта с выводом правил на основе генетического программирования.

Для прогнозирования ишемического инсульта используются следующие факторы риска:

1. Возраст. На каждые 10 лет жизни риск развития инсульта увеличивается в 3 раза
2. Пол. Как правило, страдают мужчины старше 55 лет, женщины - старше 65 лет.
3. Повышенное артериальное давление. Риск инсульта у больных с АД более 160/95 мм. рт.ст. возрастает в 4 раза в сравнении с пациентами, которые имеют нормальное давление, а при АД более 200/115 мм. рт.ст. – в 10 раз.
4. Атеросклероз.
5. Сахарный диабет.

6. Курение. Увеличивает риск развития инсульта вдвое. Ускоряет развитие атеросклероза сонных и коронарных артерий. Прекращение курения приводит через 2-4 года к снижению риска инсульта.

7. Ожирение.

8. Злоупотребление алкоголем.

9. Гиподинамия.

10. Транзиторные ишемические атаки.

11. Стеноз сонных артерий.

12. Нарушения липидного обмена.

13. Факторы образа жизни (отсутствие физич. активности, нарушение питания, фактор стресса).

Для решения поставленной задачи используется аппарат генетического программирования.

Генетическое программирование (ГП) использует процедуры поиска, основанные на механизмах естественного отбора и наследования. В них используется эволюционный принцип выживания наиболее приспособленных особей – потенциальных решений [2].

В отличие от генетических алгоритмов каждое потенциальное решение представляется деревом, состоящим из функциональных вершин, которые являются внутренними узлами деревьев, и терминалов, которые формируют листья деревьев.

Таким образом, для решения задачи с помощью ГП необходимо выполнить следующее: определить терминальное множество, определить функциональное множество, определить фитнес-функцию, определить значения параметров (размер популяции, максимальная глубина особей в начальной популяции, максимальная глубина особей в развивающейся популяции, метод генерации, вероятность функционального узла, вероятность терминального узла, вероятность кроссинговера, вероятность репродукции, вероятность мутации, вероятность выбора функционального узла для кроссинговера, метод селекции и т.п.)

После этого можно разрабатывать непосредственно сам алгоритм решения конкретной задачи.

В качестве обучающего множества в работе используются данные обследования пациентов, имеющих факторы риска и развившийся из-за них ишемический инсульт и группа пациентов, у которых ишемического инсульта нет.

Входное обучающее множество представим в виде булевых переменных. Для этого исходные данные преобразуем следующим образом:

1. Возраст > 30.
2. Возраст > 40.
3. Возраст > 50.
4. Возраст > 60.
5. Возраст > 70.
6. Пол (1- мужской, 0 - женский).
7. АД > 120/80 мм. рт.ст.
8. АД > 160/95 мм. рт.ст.
9. АД > 200/115 мм. рт.ст.
10. Атеросклероз (1 – есть, 0 - нет).
11. Сахарный диабет (1 – есть, 0 - нет).
12. Курение (1 – да, 0 - нет).
13. Ожирение (1 – да, 0 - нет).
14. Злоупотребление алкоголем (1 –да, 0 - нет).
15. Гиподинамия (1 – есть, 0 - нет).
16. Отсутствие физической активности (1 – да, 0 - нет).
17. Транзиторные ишемические атаки (1 – да, 0 - нет).
18. Стеноз сонных артерий (1 – да, 0 - нет).
19. Нарушения липидного обмена (1 – да, 0 - нет).

Набор терминалов для данной задачи состоит из параметров, перечисленных выше.

Набор функции состоит из логических операторов: И, ИЛИ, И-НЕ, ИЛИ-

НЕ. Поэтому решением является логическое выражение, использующее признаки пациента в качестве переменных.

При использовании ГП мы работаем с популяцией, которая состоит из закодированных соответствующим образом особей (деревьев), подвергающихся воздействию генетических операторов, для нахождения оптимального решения, которым считается дерево, наилучшим образом решающее поставленную задачу [3].

Начальная популяция представляет собой сгенерированные рекурсивно случайным образом деревья. Сначала генерируется функциональный узел и его аргументы. Для каждого следующего узла определяется, будет ли он функциональным или терминальным, в зависимости от чего выбирается значение из соответствующего множества. Чтобы предотвратить появление слишком больших деревьев, устанавливается параметр, регулирующий глубину особей.

При создании деревьев используются следующие методы:

1. Полный. При генерации случайного дерева каждая ветвь имеет одинаковую (максимальную) глубину.
2. Растущий. При генерации случайного дерева каждая ветвь может иметь различную глубину, но не более чем максимальная.
3. Комбинированный. Половина деревьев всей популяции генерируется полным методом, вторая половина – растущим.

Деревья оцениваются относительно определенной специальной функцией приспособленности. Переменная диагноза принимает значения 1 (высокая степень риска развития ишемического инсульта) или 0 (низкая степень риска развития ишемического инсульта). Для каждого дерева вычисляется доля пациентов с правильно поставленным диагнозом. Это и будет являться фитнес-функцией.

Производство последующих поколений достигается применением генетических операторов репродукции, кроссинговера, мутации.

Алгоритм решения поставленной задачи состоит из следующих

ОСНОВНЫХ ЭТАПОВ:

- а) инициализация или выбор исходной популяции хромосом;
- б) оценка приспособленности хромосом в популяции;
- в) проверка условия останова алгоритма;
- г) селекция хромосом;
- д) применение генетических операторов;
- е) формирование новой популяции;
- ж) выбор наилучшей хромосомы.

В качестве критерия останова служит выполнение определенного числа итераций или указание определенного числа повторения лучшего результата.

В настоящее время ГП в области диагностики и прогнозирования конкурирует с нейронными сетями и обладает некоторыми преимуществами. Данная работа расширяет область применения ГП на примере актуальной проблемы прогнозирования риска ишемического инсульта.

Перечень ссылок

1. Валикова Т.А., Алифирова В.М. Инсульт, этиология, патогенез, классификация, клинические формы, лечение и профилактика / Электронный ресурс. Способ доступа: URL: <http://medbookaide.ru/books/fold1002/book1117/p1.php>
2. Рутковская Д., Пилинский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер.с польск. И.Д. Рудинского. - М.:Горячая линия – Телеком, 2006.– 452 с.
3. Simon Kent - Diagnosis of Oral Cancer using Genetic Programming.- Technical Report CSTR-96-14 CNES-96-04 - July 1996.