

вычисленных разностей. Из этого следует, что сеть Кохонена позволяет разделить множество входных образов на несколько групп в соответствии с характерными признаками.

Преимущество - позволяет разделить множество входных образов на кластеры, обладающие некоторыми общими свойствами.

3. Многослойный персептрон:

Многослойный персептрон является одной из базовых и самых распространенных моделей нейронных сетей, позволяющих решать широкий спектр задач, одна из которых - задача распознавания образов. Многослойный персептрон относится к сетям, обучающимся на основе алгоритма обратного распространения ошибки, весовые коэффициенты подстраиваются итерационно.

Преимущество - может формировать на выходе произвольную многомерную функцию при соответствующем выборе количества слоев, диапазона изменения сигналов и параметров нейронов [1].

Недостатки:

- требуется обучающая выборка значительного объема, содержащая множество образцов каждого типа;

- для каждого из обучающих примеров в процессе обучения многократно происходит пересчет значений всех нейронов сети и связей между ними.

Выбор конкретной архитектуры ИНС для распознавания временных рядов должен быть основан на имеющемся наборе обучающих данных, требованиях к времени обучения, а также ограничений объема памяти необходимого для работы и обучения ИНС. Исходя из вышеизложенного, целесообразной является разработка такой архитектуры ИНС, которая удовлетворит требованиям задачи распознавания и использует преимущества существующих ИНС.

Список использованных источников:

1. Серебряная, Л. В. Методы и алгоритмы принятия решений: учебно-методическое пособие / Л. В. Серебряная, Ф. И. Третьяков. – Минск: БГУИР, 2016. – 64 с.
2. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-3/W3, 2013 CMRT13 - City Models, Roads and Traffic 2013, 12 – 13 November 2013, Antalya, Turkey, 13 - 16 с.
3. Dictionary of Philosophy of Mind [Электронныйресурс] / McCulloch, Warren Sturgis - May 2004. - Режимдоступа: <https://sites.google.com/site/minddict/mcculloch-warren-sturgis>. - Датадоступа: 17.03.2018
4. Artificial Neural Network [Электронныйресурс] / Wikipedia. – Режимдоступа: https://en.wikipedia.org/wiki/Artificial_neural_network. – Дата доступа: 18.03.2018.
5. Найман, Э. Малая энциклопедия трейдера / Э. Найман. – М.: Альпина Паблишер, 2003. – 378 с.
6. Хайкин, С. Нейронные сети: полный курс / С. Хайкин. – М.: Вильямс, 2006. – 104 с.

АНАЛИЗ ПОБОЧНЫХ ДАННЫХ СОСТОЯНИЯ СИСТЕМЫ И ДАННЫХ РЕЗУЛЬТИРУЮЩИХ ТОЧЕК СИСТЕМЫ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Мялик И.С.

Куликов С.С. – к.т.н., доцент

Уже сейчас большинство программных системы обладают теми или иными возможностями мониторинга собственной работоспособности. Одним из видов такого мониторинга является сохранение данных состояния системы или данных результата выполнения каких-либо ключевых функций, т.е. протоколирование определенных, чаще всего, скрытых от пользователя, сведений о системе. Как правило, протокол работы сохраняется в файлы на сервере или в базы данных с определенной моделью для будущей обработки. В данный момент не до конца решенным остается вопрос обработки большого объема протоколов работы системы в реальном времени.

Допустим, перед нами стоит задача обработки большого объема протоколов работы системы размером в несколько ТБ (например, системы показа рекламы).

Одна из возможных стратегий может быть описана следующим образом:

1. Простейшей эвристикой определяются «нужные пользователи», например, по числу заходов на сайт за неделю.
2. Теперь для каждого сайта считается его посещаемость «нужными пользователями» и всеми пользователями, определяется доля «правильных пользователей» от числа всех посетителей сайта.
3. Далее сортируются сайты по убыванию рассчитанного показателя, выбираются, например, 300 сайтов, начиная с начала, на которых находится целевая аудитория.

Вероятнее всего результатом решения этой задачи будут веб-сайты содержащие один и тот же тип информации. Можно предположить, что результатом будут веб-сайты по аренде и продаже жилья. Таким образом, определяется новый этап задачи, а именно: каким образом разместить рекламные баннеры. На каких сайтах следует разместить аренду домов в Минске, а на каких продажу квартир в Ингушетии. Для выполнения этой задачи, необходимо работать с инструментами Big Data и иметь доступ к данным для их анализа.

Инструмент Big Data

1. Необходимо выбрать место для хранения протоколов работы системы, им может быть файловая система или база данных.
2. Далее следует эти данные обрабатывать, сортировать по разным параметрам, находить веб-сайты в верхних 100\300\1000 строчках рейтинга, определять долю целевых пользователей в общем трафике. Далее будет использована парадигма MapReduce и искусственные алгоритмы вида:
 - фильтрации;
 - оптимизации: соединения «map-join» с использованием распределенного кэша;
 - оптимизации: применение «combiner», как одно из требований реализации «reducer»;
 - соединения «reduce-join», к примеру, для поиска рейтинговых сайтов;
3. Хотя данное заключение верно не только для Big Data, но также и для большого объема данных становится крайне критично, для этого необходимо построить некий процесс предобработки данных, с помощью машинного обучения.

Машинное обучение

Первый этап работы – информационная очистка и предобработка данных:

- работа с аномальными значениями;
- удалить без информативные данные, такие как: технические посещения и т.д.;
- нормализовать URL-адреса;
- работа с пропущенными, при трассировке, значениями.

Второй этап – превращение данных, в понятные для модели векторы.

На этом этапе существует множество подходов, но нет однозначной методики их применения. Было выделено несколько общепринятых подходов для решения этой задачи:

- проанализировать домен и сгруппировать веб-сайты по нему;
- скачать описательную часть страницы и проанализировать ее;
- скачать весь код страницы и постараться выявить тематику страницы;
- связать посещения пользователем сайтов в цепочку и определить логику в последовательности переходов.

Последний этап – применение машинного обучения.

Для применения машинного обучения нету определенного подхода, это наиболее творческий вариант, требующий применения здравого смысла и умения итерационно улучшать выбранное решение.

Список использованных источников:

Real-WorldMachineLearning. Henrik Brink, Joseph W. Richards.

КОНТРОЛЬ КАЧЕСТВА РАЗРАБОТКИ ПРОГРАММНЫХ СРЕДСТВ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Новицкий А.А.

Бахтизин В.В. – профессор каф. ПОИТ, к.т.н., доцент

Работы по контролю качества разработки востребованы в течение всего жизненного цикла программных средств (ПС). Выбор оптимальных для проекта работ по контролю качества (КК) отвечает целям повышения надежности итогового продукта, уменьшения издержек, повышения рентабельности процесса разработки и является актуальной темой. Повсеместное внедрение программных продуктов и рост сложности используемых технологий и ПС становится причиной роста цены ошибки в продукте. Поэтому не теряет актуальности вопрос дальнейшей оптимизации процессов контроля качества (КК) разработки программных средств.

В данном докладе рассматривается классификация тестирования в зависимости от необходимости выполнения программного кода. Это позволяет четко выделить как минимум одну пригодную для анализа группу работ. По данному критерию методы тестирования делятся на статические, не требующие запуска программного кода, и динамические, непосредственно взаимодействующие с функционирующим приложением или его моделью.

Статические виды тестирования включают работы по верификации требований и анализу программного кода, вследствие чего могут обеспечивать наиболее раннее выявление различных проблем, потому обязательны к включению в план разработки [1].

Динамические методы тестирования в основном обеспечивают обнаружение дефектов функциональности в ходе проверки тестовых сценариев или же отвечают задачам проверки устойчивости системы к превышению пределов нормального функционирования, закрытости от влияния извне. Подходы к динамическому тестированию наиболее разнообразны, что создает дополнительные трудности в выборе работ для полного избыточного покрытия функций разрабатываемого ПС.

Независимо от наличия автоматизации процесса динамического тестирования общим вариантом критерия оценки покрытия является покрытие, основанное на спецификации или требованиях. Главное требование состоит в покрытии некоторого утвержденного минимума – набора требований. Разница в механизме доступа обуславливает специфику проведения тестирования, предпочтительного для каждого из методов тестирования. Таким образом, рассматривать следует критерии эффективности ручного и