СПбГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

А.Е. Сулавко

Машинное обучение в приложениях биометрии

Ансамблирование моделей классификации биометрических образов

Конспект лекции

СПбГЭТУ «ЛЭТИ», 2022 г.

АНСАМБЛИРОВАНИЕ МОДЕЛЕЙ КЛАССИФИКАЦИИ БИОМЕТРИЧЕСКИХ ОБРАЗОВ

Еще одним подходом в задачах классификации биометрических образов является ансамблирование моделей машинного обучения с целью повышения эффективности распознавания и классификации образов.

Тема ансамблирования достаточно широко изучена в машинном обучении и может быть представлена в виде самых разных модификаций, однако, в задачах классификации биометрических образов ансамблирование представляет собой объединение и обучение нескольких (базовых) алгоритмов классификации на одной и той же обучающей выборке с последующей комбинированием их предсказаний для новых тестируемых образов. Математическим обоснованием этой идеи служит теорема Кондорсе (XVIII в.), которая утверждает: если мнения экспертов независимы и вероятность правильного решения каждого из них больше 0,5, то с увеличением количества экспертов вероятность правильного решения комитета экспертов возрастает и стремится к единице. Причем чем выше вероятность верного решения для каждого эксперта в отдельности, тем выше вероятность верного решения комитета.

Помимо классического метода усреднения модели (в общем случае, это любое усреднение по Коши), существуют несколько методов объединения базовых алгоритмов в ансамбли с их последующим обучением, среди которых для задач классификации биометрических образов особым образом можно выделить следующие 3:

1. Бэггинг (bootstrap aggregating) – метаалгоритм композиционного обучения машин, основная идея которого заключается в обучении базовых (слабых) классификаторов на разных подмножествах обучающей выборки. Базовые классификаторы могут быть идентичными или иметь разные архитектуры. Бэггинг уменьшает дисперсию голосов базовых классификаторов и помогает избежать переобучения. Принцип работы бэггинга схож с принципами работы метода случайных классификаторов, а также методов накопления сигналов при их обнаружении и заключается в повышении отношения сигнал / шум.

2. Бустинг (boosting) – семейство алгоритмов машинного обучения, преобразующих слабые обучающие алгоритмы к сильным. Бустинг строит ансамбль путём тренировки каждого нового классификатора, уделяя больше внимания обучению на тех тренировочных примерах, которые предыдущие модели классифицировали ошибочно (например, путем присвоения весов обучающим примерам), и имеет тенденцию к переобучению.
3. Стекинг (stacked generalization) предполагает построение многослойных структур из ансамблей классификаторов, когда выходные данные ансамбля первого слоя воспринимаются ансамблем второго слоя как входные данные (метапризнаки). При использовании стекинга увеличивается необходимый для обучения объем выборки, так как для корректного обучения метамодели каждый слой требуется настраивать на разных тренировочных примерах.

Одной из реализаций ансамблирования (в том числе используемых в задачах классификации биометрических образов) являются комитеты нейронных сетей. В сущности, комитет представляет собой объединение двух или нескольких нейронных сетей (подсетей), каждая из которых решает свою отдельную задачу (ВИ и ПИ на примере иммунной сети ниже). В случае комитета сетей имеет смысл также говорить о комитетах мер близости, а также комитетах критериев согласия (также на примере иммунной сети).

Любая система, проектируемая с целью обеспечения информационной безопасности (в том числе системы биометрической идентификации/верификации), функционирует в потенциально враждебной внешней среде. Регулирование взаимодействий такой системы с внешней средой определяется принятой политикой безопасности. В случаях выявления угроз информационной безопасности такая система предпринимает определенные меры реагирования. Таким образом, значительная часть задач, решаемых системами информационной безопасности, по своей сути может быть сведена к проблеме нахождения отличий «своего» от потенциально опасного «чужого»[1].

Опираясь на вышеперечисленные особенности, по ряду параметров делающие системы информационной безопасности схожими с иммунной системой живых организмов (человека), одним из перспективных

направлений в области проектирования подобных систем становится подход, реализующий так называемые искусственные иммунные сети (ИИС). Данная технология известна человечеству с конца прошлого века, и, несмотря на то, что в отличие от искусственных нейронных сетей, эволюционных и генетических алгоритмов, ИИС пока не заняли прочного места в современных информационных технологиях[3], они по-прежнему представляют большой интерес в области информационной безопасности и претерпевают постоянные улучшения и модификации.

Важно отметить, что ИИС (как и ИНС) – крайне упрощенная конструкция, которая не подразумевает строгого соответствия своему биологическому прототипу. Искусственные иммунные системы (сети) в компьютерных науках принято рассматривать как семейство алгоритмов, основанных на соответствующих теориях о естественной иммунной системе (ЕИС): дендритных клеток, негативного отбора, клональной селекции (положительного отбора), сетевых алгоритмов (последние чаще всего называют иммунными сетями, а не системами, далее все типы иммунных моделей будем называть ИИС).

Иммунная система содержит множество клеток (макрофаги, дендритные клетки, лимфоциты), которые обладают способностью обнаруживать и удалять чужеродные организмы (антигены). Назовем все такие клетки детекторами – вычислительными элементами, способными анализировать распознаваемый образ либо его отдельные фрагменты и реагировать на него пропорционально тому, насколько этот образ соответствует антигену. Шкала реакций задана на интервале действительных чисел [0; 1], где 0 – полная уверенность в том, что клетка принадлежит организму (гипотеза «Свой»), а 1 – полная уверенность в обратном (гипотеза «Чужой»). Силу взаимодействия между клетками ИИС и антигеном также называют аффинностью. Каждый детектор следует рассматривать как бинарный классификатор, состоящий из нескольких функций, последовательно применяющихся к биометрическому образу. Образ представляет собой вектор признаков фиксированной длины $\bar{a} = \{a1, a2, \ldots, an\}$, где n – количество признаков, которое должно присутствовать в образе. В общем виде получение реакции i-го детектора на входной образ $\bar{a}$ можно описать формулой (1):

$$u_i = \varphi_\chi(y' = \phi(y = f_x(\bar{\alpha} = R(\bar{a}, \Psi_i), \breve{g}, \Theta_i), T_i))$$

Опишем функции детектора и их параметры:

1) $\bar{\alpha} = R(\bar{a}, \Psi_i)$ функция-рецептор, предоставляющая интерфейс взаимодействия для детектора и антигена. Данная функция извлекает ŋ из n признаков, содержащихся в ā, Ψi – множество номеров признаков из ā, которые должен анализировать i-й детектор. Вектор ᾱ = {a1, a2, …, aŋ} является подмножеством ā с собственной сквозной нумерацией элементов;

2) $y = f_x(\bar{\alpha}, \breve{g}, \Theta_i)$ функция-ядро детектора, параметрический функционал, который вычисляет близость вектора ᾱ к эталону класса образов «Свой»; x – тип функционала (для построения ядер детекторов могут применяться следующие функционалы: мера Минковского, разные вариации меры Байеса–Минковского, «наивный Байес» в дифференциальной и интегральной форме, параметрические критерии); ğ – вектор параметров функционала, которые влияют на характер вычислений; Θi = {μ1, μ2, …, μŋ, σ1, σ2, …, σŋ}, μj и σj – статистические оценки параметров распределения значений j-го признака из вектора ᾱ. Данные из множества Θi рассчитываются на основании нескольких случайных примеров из обучающей выборки (далее фолд). Разные функционалы образуют различные виды детекторов, которые дают слабо коррелированные решения относительно друг друга. Из любого функционала можно получить разные меры близости за счет изменения параметров ğ.

3) $y' = \phi(y, T_i)$ функция нормирования откликов y относительно порога Ti, который вычисляется в процессе настройки i-го детектора. Функция нормирования может иметь 2 реализации:

$$\phi(y, T_i) = y/T_i$$

если y – это расстояние от ᾱ до эталона класса «Свой» (чем меньше, тем ближе), тогда Ti – это максимальное значение функции-ядра i-го детектора, при поступлении на его вход обучающих образов «Свой», либо:

$$\phi(y, T_i) = T_i/y$$

если y – это вероятность того, что ᾱ принадлежит классу «Свой», тогда Ti – это минимальное значение функции-ядра i-го детектора, при поступлении на его вход обучающих образов «Свой». Физический смысл y (расстояние или вероятность) зависит от функции-ядра соответствующего детектора (например, для меры Евклида требуется использовать первый вариант, а для «наивного Байеса» – второй);

4) $u_i = \varphi_\chi(y')$ функция активации, дополнительный нелинейный элемент детектора, который определяет особенности реагирования на антиген. Функция активации также необходима, чтобы привести отклик детектора к области значений [0; 1]. В качестве функций активации имеет смысл использовать либо наиболее быструю из сигмоидальных, либо применять функции, которые дают как можно более отличающиеся результаты, чтобы создавать детекторы с низкой коррелированностью решений на базе однотипных мер близости (в целом x в большей степени влияет на характер преобразований детектора, чем x).

Детекторы можно разделить на две группы: врожденный и приобретенный иммунитет, и рассматривать их как два комитета (ансамбля) слабых классификаторов, обучаемых при помощи разных алгоритмов. Коллективное решение комитета из N детекторов может быть вычислено как среднее частных решений:

$$\ddot{u} = \Phi(\overline{D}^* = \{D_1^*, \dots D_n^*\}, \bar{a}) = \frac{1}{N}\sum_{i=1}^{N} \varphi(D_i^*, \bar{a}) = \frac{1}{N}\sum_{i=1}^{N} u_i$$

Врожденный иммунитет (ВИ) передается посредством генов, органы иммунной системы формируются еще при эмбриональном развитии. У эмбрионов В-лимфоциты образуются в печени и костном мозге. В предлагаемой модели костный мозг является местом пребывания иммунокомпетентных детекторов, параметры и состав которых определяются в ходе эмбрионального (и постэмбрионального) развития. Так, ВИ формируется в процессе итерационного обучения ИИС с использованием тренировочной и валидационной выборок. Последняя используется для промежуточной оценки надежности решений ИИС при смене поколения

детекторов (схожие практики применяются при обучении ИНС). Обе выборки являются непересекающимися подмножествами обучающей выборки.

Приобретённый иммунитет (ПИ) развивается с течением жизни и определяет способность организма обезвреживать специфические антигены, которые попадали в организм ранее. В формировании ПИ участвует тимус – орган, в котором происходит созревание и обучение Т-лимфоцитов. В предложенной модели тимус осуществляет настройку и отбор иммунокомпетентных детекторов, используя валидационную выборку. Адаптивный иммунный ответ приводит к появлению иммунологических клеток памяти (представленных в ИИС детекторами), которые долгое время пребывают в «спящем состоянии» до повторной встречи с тем же антигеном. В рассматриваемой модели ИИС приобретенный иммунитет формируется в процессе функционирования ИИС. Если решение об отнесении образа к категории «Свой» или «Чужой» является неоднозначным, могут генерироваться новые иммунокомпетентные детекторы.

Идея объединения классификаторов в комитет основана на теореме Кондорсе. Однако на практике решения классификаторов, играющих роль экспертов, в той или иной мере коррелированы. Чем ниже коррелированность решающих правил, тем более ощутим положительный эффект при их комплексировании. Таким образом, имеются следующие гиперпараметры, которые влияют на эффективность комитета детекторов:

- N – количество детекторов;
- RD – матрица коэффициентов корреляции Пирсона r(ūi, ūl) между решениями всех возможных пар детекторов, где ūi – вектор реакций i-го детектора на примеры образов «Чужих» из тренировочной или валидационной выборки;
- Δu – сила детекторов, их способность давать как можно более высокие показатели разницы средних уровней реакции на образы «Свой» и «Чужой».

При появлении в ИИС идентичных или слабых детекторов происходит апоптоз – процесс программируемой клеточной гибели для уничтожения дефектных клеток. Необходимо, чтобы решения всех детекторов врожденного и приобретенного иммунитета не являлись полностью коррелированными. Поэтому после генерации детектора должна

осуществляться проверка идентичности параметров нового детектора и уже существующих. При обнаружении «двойника» его следует удалить и сгенерировать детектор снова. При этом значения параметров ğ можно считать равными, когда они отличаются менее чем на 10–1. Конечно, решения детекторов будут в той или иной степени коррелированы (но не на 100 %). Чем сильней различаются параметры Di и Dl, тем менее коррелированы решения i-го и l-го детекторов.

При обучении рассмотренной модели ИИС с учителем используются бэггинг и бустинг. На каждой итерации происходит генерация новой популяции детекторов, которые настраиваются с учетом нескольких случайных тренировочных примеров (бэггинг), и выполняется промежуточная оценка их эффективности как на тренировочной, так и на валидационной выборке, слабые детекторы уничтожаются (апоптоз), в результате появляется новое поколение иммунокомпетентных (более эффективных) детекторов. Мерой эффективности (обученности) детекторов можно считать Δu.

Со временем в биометрических образах пользователей могут происходить изменения, что продиктовано множеством как внешних по отношению к человеку факторов, так и его внутренним состоянием. В таком случае можно говорить о так называемом концептуальном дрейфе модели (дрейфе данных). Под данным термином, в общем случае, понимается любое изменение отношений между входными и выходными данными модели машинного обучения с течением времени.

Для учета подобных изменений, то есть для возможности работы модели в динамическом режиме, рассматриваемая модель ИИС реализует алгоритм дообучения иммунной сети с подкреплением (формирование приобретенного иммунитета). Если обучающая выборка нерепрезентативна, эффективность детекторов ВИ может не соответствовать оценке Δu, при этом нет гарантий, что плохо настроенные детекторы в действительности преодалевают барьер Кондорсе. Обойти барьер Кондорсе можно, если дать возможность детекторам ПИ голосовать за коллективное решение детекторов ВИ.

Основными свойствами рассмотренной выше модели являются:

1. Распределенный характер вычислений и проявление эмерджентности (ИИС может повышать качество решений в процессе функционирования, в

отличие от базовых классификаторов, точность распознавания для ИИС выше, чем для каждого классификатора в отдельности).

2. Достаточно устойчивый процесс обучения (склонность к переобучению незначительна при формировании ВИ).
3. Способность ИИС к адаптации, обусловленная двойной пластичностью: структурной и параметрической (меняются параметры и состав детекторов).
4. Взаимодействие – врожденный иммунитет формирует параметры, которые влияют на механизм подкрепления детекторов приобретенного иммунитета.
5. Надежность решений ИИС зависит от объема и чувствительности популяции детекторов.
6. Формирование памяти при помощи механизма приобретенного иммунитета.