

УДК 004.8

СПОСОБЫ МАШИННОГО ОБУЧЕНИЯ

М.А. ШЕВЦОВ

(Представлено: канд. физ.-мат. наук, доц. Ю.Ф. ПАСТУХОВ)

Рассматриваются способы машинного обучения, отдельные из них поддерживают несколько видов обучения, некоторые могут быть реализованы только с определенным алгоритмом обучения. В данной работе рассмотрены основные из рассмотренных способов.

1. Обучение с учителем (англ. **supervised learning**) – один из способов машинного обучения, в ходе которого испытуемая система принудительно обучается с помощью примеров «стимул-реакция».

Между входами и эталонными выходами (стимул-реакция) может существовать некоторая зависимость, но она неизвестна. Известна только конечная совокупность пар «стимул – реакция», называемая обучающей выборкой. На основе этих данных требуется восстановить зависимость (построить модель отношений стимул – реакция, пригодных для прогнозирования), то есть построить алгоритм, способный для любого объекта выдать достаточно точный ответ.

Под учителем понимается либо сама обучающая выборка, либо тот, кто указал на заданных объектах правильные ответы.

Постановка данного эксперимента предполагает наличие экспериментальной системы, метода обучения и метода испытания системы или измерения характеристик.

Основные алгоритмы:

- линейная и логистическая регрессия;
- машина опорных векторов;
- наивный Байес;
- нейронная сеть;
- повышение градиента;
- классификационные деревья и случайный лес.

2. Обучение без учителя (англ. **unsupervised learning**) – способ машинного обучения, при котором испытуемая система спонтанно обучается выполнять поставленную задачу без вмешательства со стороны экспериментатора. Метод пригоден для задач, в которых известны описания множества объектов (обучающей выборки) и требуется обнаружить внутренние взаимосвязи, зависимости, закономерности, существующие между объектами.

Обучение без учителя часто противопоставляется обучению с учителем, когда для каждого обучающего объекта принудительно задается «правильный ответ», и требуется найти зависимость между стимулами и реакциями системы.

Экспериментальная схема обучения без учителя часто используется в теории распознавания образов, машинном обучении. При этом в зависимости от подхода формализуется в ту или иную математическую концепцию.

Основные алгоритмы:

- алгоритм кластеризации, такой как K-mean, иерархическая кластеризация или модели смеси;
- алгоритмы сокращения размерности, такие как PCA, ICA или autoencoder;
- обнаружения аномалий, чтобы найти выбросы в данных, т.е. наблюдения, которые не соответствуют шаблонам набора данных.

3. Обучение с подкреплением (англ. **reinforcement learning**) – один из способов машинного обучения, в ходе которого испытуемая система (агент) обучается, взаимодействуя с некоторой средой. Откликом среды (а не специальной системы управления подкреплением, как это происходит в обучении с учителем) на принятые решения являются сигналы подкрепления, поэтому такое обучение является частным случаем обучения с учителем, но учителем является среда или ее модель. Также нужно иметь в виду, что некоторые правила подкрепления базируются на неявных учителях.

Учитывая его состояние и состояние окружающей среды, агент выберет действие, которое максимизирует его вознаграждение или будет исследовать новую возможность. Эти действия изменяют состояние среды и агента. Они также будут интерпретироваться, чтобы дать вознаграждение агенту. Выполняя этот цикл много раз, агенты улучшат свое поведение.

Агент воздействует на среду, а среда воздействует на агента. О такой системе говорят, что она имеет обратную связь. Такую систему нужно рассматривать как единое целое, поэтому линия раздела между средой и агентом достаточно условна.

Основные алгоритмы:

- Q-Learning.

Заключение

Выбор алгоритма зависит от типа задач, которые предстоит решить, а также от доступных данных и алгоритма, по которому будет построена модель. Некоторые виды задач могут поддерживать различные способы машинного обучения, а некоторые только один.

ЛИТЕРАТУРА

1. Айвазян, С.А. Прикладная статистика: основы моделирования и первичная обработка данных / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М. : Финансы и статистика, 1983.
2. Айвазян, С.А. Прикладная статистика: исследование зависимостей / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М. : Финансы и статистика, 1985.
3. Прикладная статистика: классификация и снижение размерности / С.А. Айвазян [и др.] – М. : Финансы и статистика, 1989.
4. Вапник, В.Н. Восстановление зависимостей по эмпирическим данным / В.Н. Вапник. – М. : Наука, 1979.