

# **ПРИМЕНЕНИЕ АЛГОРИТМОВ НА МНОЖЕСТВЕ НАБОРОВ ДАННЫХ В РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЕ ДЛЯ СОКРАЩЕНИЯ ВРЕМЕНИ ОБУЧЕНИЯ СИСТЕМЫ.**

Мальцева М.Г., магистр кафедры компьютерных технологий  
«Университета ИТМО», [mariymlv@gmail.com](mailto:mariymlv@gmail.com)

## **Аннотация**

Обучение с подкреплением— это один из способов машинного обучения, в ходе которого испытуемая система (агент) обучается, взаимодействуя с некоторой средой. Главное достоинство этого метода — его сравнительная простота для понимания, но его сложно реализовать: наблюдаются действия обучаемого объекта и в зависимости от результата поощряется данный объект, либо наказывается, что не позволяет обучаемому объекту извлечь полную информацию о том, как следует действовать. Цель работы — обучить систему так, чтобы для каждого набора данных был найден лучший алгоритм за максимально короткое время и наименьшее число итераций.

## **Введение**

Существует большое число различных алгоритмов выбора признаков, но ни один из них не является универсальным и одинаково эффективным для решения большинства задач машинного обучения. Более того, универсального алгоритма, превосходящего все остальные на всех наборах данных, не существует. Поэтому в настоящее время проблема выбора подходящего алгоритма для решения конкретной задачи обычно относится к экспертной — подбором алгоритма занимается специалист, ее решающий.

## **Мета-обучение**

Распространенным решением подобного рода задач является применение всех доступных алгоритмов к рассматриваемой задаче и

выбор одного — с лучшей производительностью. Однако это решение требует много времени и весьма трудоемко. Хорошо зарекомендовавшим себя подходом для преодоления подобного рода проблем выбора оптимального алгоритма является мета-обучение. Основная идея мета-обучения заключается в том, чтобы описать задачи (представленные набором данных) через мета-данные, создавая новую задачу, в которой объектами являются задачи, а целевым значением — алгоритм, лучше всего подходящий для решения данной задачи. На основе этой информации для нового набора данных рекомендуется алгоритм, который показал лучшую эффективность на похожих наборах данных.

Вместе с тем следует признать, что такое решение перекладывает временные затраты с этапа применения на этап обучения: для  $M$  алгоритмов полный перебор на новом наборе данных потребует  $O(M)$  запусков алгоритмов, а в мета-обучении —  $O(1)$ . Однако для применения системы мета-обучения необходимо ее обучить на  $N$  наборах данных, что потребует  $O(MN)$  запусков алгоритмов. Эта проблема усугубляется тем, что при росте  $M$  необходимо увеличивать число наборов данных, чтобы избежать переобучения. Указанная проблема препятствует масштабируемости систем мета-обучения, поскольку время их обучения при большом числе алгоритмов может измеряться годами.

Для решения описанной проблемы в данной работе предлагается разрешить применение системы до ее полного обучения.

## **Обучение с подкреплением**

Система была реализована таким образом, что на каждом шаге заполнения базы данных выбирается следующий шаг так, чтобы максимизировать математическое ожидание точности прогноза. Для этого необходимо учитывать, с какой вероятностью новый набор данных окажется близким к тем или иным наборам данных. Проблема заключалась в том, чтобы найти выгодные действия, принимая эмпирически лучшие шаги как можно чаще. Распространенным способом решения этой дилеммы является использование алгоритма многоруких бандитов.

Для решения задачи был использован алгоритм UCB1. Он построен по принципу оптимизма в условиях неопределенности. Было

реализованы несколько алгоритмов обучения системы: алгоритм обучения с подкреплением по наборам данных, по алгоритмам, смешанный (по наборам данных и алгоритмам), алгоритм случайного выбора.

В первом случае мы считаем, что «ручками» однорукого бандита выступают наборы данных, и к ним необходимо применять алгоритмы, чтобы максимизировать выигрыш. При этом выигрыш на конкретном наборе данных должен умножаться на вероятность того, что новый набор данных окажется ближайшим именно к этому набору. Во втором случае считается, что «ручками» однорукого бандита выступают алгоритмы, и необходимо максимизировать выигрыш применения каждого из алгоритмов. В смешанном алгоритме ручками выступают и алгоритмы и наборы данных.

### **Предложенный алгоритм и описание экспериментов**

Была создана матрица качества работы алгоритмов на датасетах обучения. Эта матрица заполняется в ходе обучения системы, и именно этот процесс оптимизировался. Заранее была заполнена матрица качества работы алгоритмов на датасетах тестирования. Случайным образом были перемешаны алгоритмы, затем последовательно выбирались алгоритмы и запускались на тех датасетах, на которых еще не было ни одного запуска алгоритма. Постепенно заполнялась матрица. Для обучения системы было построено четыре системы, для каждой из которых для выбора следующей пары алгоритм - датасет используется один из четырех алгоритмов.

Затем, была заполнена матрица качества работы на датасетах обучения. На каждом шаге после добавления в матрицу нового элемента системе подавались датасеты тестирования. Для каждого такого датасета сравнивалось, насколько эффективен тот алгоритм, который система возвращает. В итоге заполнялся массив, в котором для каждой системы и для каждого датасета, а также для каждого шага хранилась оценка качества. Затем для каждого набора данных выбирался лучший алгоритм по оценке качества.

### **ЗАКЛЮЧЕНИЕ**

Были проведены эксперименты, в результате которых можно сделать вывод, что смешанный алгоритм обучает систему лучше всего, но работает дольше остальных. Планируется реализовать несколько

алгоритмов активного обучения. После реализации алгоритмов полученную программу можно будет применять для решения сложных практических задач. В дальнейшем также планируется предоставить результаты сравнения разных подходов обучения системы и сделать выводы о качестве работы рекомендательной системы.

## Литература

1. Finite-time Analysis of the Multiarmed Bandit Problem / P. Auer, N. Cesa-Bianchi, P. Fischer // *Machine Learning*. – 2002. Vol. 47. – P. 235–256.
2. A Feature Subset Selection Algorithm Automatic Recommendation Method / G. Wang, Q. Song, H. Sun, X. Zhang, B. Xu, Y. Zhou. – Technical Report. Department of Computer Science & Technology. – 2013. P. 34.
3. Sample mean based index policies with  $O(\log n)$  regret for the multi-armed bandit problem / R. Agrawal // *Advances in Applied Probability*. – 1995. – Vol. 27. – P. 1054–1078.
4. Introduction to reinforcement learning / R.S. Sutton, A.G. Barto. – Cambridge, MA : MIT Press, 1998. – P. 322.
5. No Free Lunch Theorems for Search / David H., Wolpert D., William W., G. Macready // The Santa Fe Institute 1399 Hyde Park Road Santa Fe, NM, 87501 February 23. – 1996.
6. Dimension Reduction / P. Cunningham // Technical Report UCI-CSI-2007-7, August 8th, 2007. University College Dublin. – 2007.
7. Online Black-Box Algorithm Portfolios for Continuous Optimization / P. Baudis and P. Posaik // Czech Technical University in Prague Faculty of Electrical Engineering, Department of Cybernetics Technická 2, 166 27 Prague 6, Czech Republic. – 2014
8. Automatic (Offline) Configuration of Algorithms / T. Stutzle and M. Lopez-Ibanez. – 2015.
9. Automatic Design of Evolutionary Algorithms for Multi-Objective Combinatorial Optimization / Leonardo C. T. Bezerra, M. Lopez-Ibanez, and T. Stutzle / IRIDIA, Université Libre de Bruxelles (ULB), Brussels, Belgium. – 2014.
10. A modern Bayesian look at the multi-armed bandit / S. L. Scott. // *Applied Stochastic Models in Business and Industry*. – 2010. P. 639–658.