



КАК ОБУЧИТЬ МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ?

Адилова Муножот

Simple Networking Solution LLC

adilovamunojot@gmail.com

<https://www.doi.org/10.5281/zenodo.10004358>

ARTICLE INFO

Received: 08th Oktober 2023

Accepted: 14th Oktober 2023

Online: 15th Oktober 2023

KEY WORDS

Обучение модели, валидация модели, Деплой, F1-мера, гиперпараметров, Машинное обучение, задача, оптимизация, искусственного интеллекта, классификация, регрессия, деревья решений, и нейронные сети.

ABSTRACT

Данная тема рассматривает основные этапы и методы обучения моделей машинного обучения. Машинное обучение — это процесс, при котором модель обучается на основе предоставленных данных, в целях выполнения определённых задач.

Аннотация предлагает ознакомиться с принципами обучения модели машинного обучения, от сбора до предварительной обработки данных. Затем рассматриваются различные алгоритмы машинного обучения, т.е. регрессия, классификация и кластеризация.

Далее описание включает в себя методы обучения моделей, включая разделение данных на обучающую и тестовую выборки, выбор подходящих алгоритмов и настройку гиперпараметров.[Голев,1999] Также в аннотации представлены методы оценки качества модели и способы улучшения её производительности.

В заключение, освещаются признаки переобучения модели и методы регуляризации, с целью достижения более устойчивой модели. Тема описывает важные аспекты обучения моделей машинного обучения, представляя читателю общий обзор и основные инструменты для ознакомления с наиболее эффективными моделями.

Обучение модели машинного является ключевой частью разработки алгоритмов и систем искусственного интеллекта. [Oliver Theobald.2017] В этой статье мы рассмотрим подробный процесс обучения модели машинного обучения и его ключевые шаги.

1. Определение задачи: В первую очередь, необходимо чётко определить задачу, которую вы намерены решить с помощью модели машинного обучения. Например, это может быть классификация объектов на изображении, прогнозирование временных рядов или выявление аномалий в данных.

2. Сбор и подготовка данных: Следующим шагом является сбор данных, для дальнейшего использования в обучении модели. Это могут быть как числовые,



категориальные, текстовые и изобразительные виды данных. Прежде всего, данные необходимо обработать, в частности, очистить от шума, заполнить пропущенные значения, и нормализовать их.

3. Выбор модели: На основе задачи и данных, необходимо выбрать подходящую модель машинного обучения. Существует множество алгоритмов, в числе которых - линейная регрессия, деревья решений, и нейронные сети. [John Paul Muller, Luca Massaron.2016] Каждая модель имеет свои преимущества и ограничения, и а её выбор может зависеть от характеристик данных и поставленной задачи.

4. Разделение данных: Затем данные разделяются на обучающую и тестовую выборки. Обучающая выборка будет использоваться для настройки параметров модели, тогда как тестовая выборка позволит оценить производительность модели на новых данных, которые она не видела ранее.

5. Обучение модели: Сама модель начинает обучаться на обучающей выборке. В процессе обучения модель анализирует данные и настраивает внутренние параметры, чтобы достичь наилучшей производительности согласно поставленной задаче. Обычно обучение модели включает в себя итеративные процессы, называемые эпохами, где модель обрабатывает данные многократно, чтобы улучшить результаты.

6. Оценка модели: После завершения обучения модель оценивается на тестовой выборке. Результаты оценки могут быть представлены метриками, такими как точность, полнота, F1-мера и другие, [Oliver Theobald.2017] которые позволяют сравнить производительность модели с эталоном.

7. Настройка и оптимизация модели: Завершив оценку модели, можно провести настройку и оптимизацию. Это может включать в себя изменение гиперпараметров модели, т.е. скорость обучения, количество скрытых слоев в нейронных сетях или выбор другой модели.

8. Валидация модели: После настройки модели необходимо проверить её производительность на независимой выборке для проверки ее реальной способности обобщать знания на новые данные. Это поможет оценить степень переобучения или недообучения модели.

9. Деплой модели: Когда модель проходит все предыдущие шаги и показывает хорошие результаты, она готова к развертыванию. [Andreas C. Mueller, Sarah Guido.2016] Модель может быть интегрирована в реальные системы или использована для решения конкретных задач.

10. Внедрение и мониторинг: Важной частью обученной модели является постоянный мониторинг и обновление. Данные, на которых модель была обучена, могут изменяться со временем, поэтому модель должна быть регулярно переобучена на основе новых данных, в целях её актуальности и эффективности.

Это основные шаги, которые следует выполнить при обучении модели машинного обучения. Каждый шаг может содержать большое количество деталей и возможностей для улучшения модели, а процесс обучения может занять достаточно много времени и ресурсов. Однако, следуя этим шагам и экспериментируя с различными вариантами алгоритмов и параметров, вы сможете создать эффективную модель машинного обучения для поставленной Вами задачи.



Заклучение

Обучение модели машинного обучения — это процесс использования данных для создания математической модели, которая может "обучаться" основываясь на этих данных и использовать полученные знания для принятия решений или предугадывания новых данных.

Важно отметить, что обучение модели машинного обучения — это искусство и наука, и требует практики и определённого опыта для достижения оптимальных результатов. Надеюсь, эта информация поможет вам начать обучение модели машинного обучения.

References:

1. Голев Н.Д. 1999. *Юридический аспект языка в лингвистическом освещении. Юрислингвистика. Проблемы и перспективы* анал. Алтай.
2. Коротаяев, А. В., Медведев, И. А., Шульгин, С. Г., Слиньюко, Е. В., Билюга, С. Э., Малков, С. Ю., Халтурина, Д. А., Дербилова, Е. В., Зинькина, Ю. В., Романов Д. М. 2018. *Системы глобального мониторинга рисков социально-политической дестабилизации: опыт систематического обзора*. В: Гринин, Л. Е., Коротаяев, А. В., Мещерина, К. В. Москва
3. Галяшина Е.И. 2002. *Назначение, производство и оценка заключения судебной лингвической экспертизы*. Кремль
4. Oliver Theobald. 2017 *Machine Learning for Absolute Beginners: A Plain English Introduction*.
5. John Paul Muller, Luca Massaron. 2016. *Machine Learning For Dummies*.
6. Andreas C Mueller., Sarah Guido. 2016. *Introduction to Machine Learning with Python: A Guide for Data Scientists*.
7. <https://ru.wikipedia.org/machine-learning>
8. <https://vc.ru/ml/126354-modeli-mashinnogo-obucheniya-obyasnyaem-pyatiletnemu-rebenku>