

Machine Learning: Theory & Algorithms I

Mathematics of Machine Learning

Математика машинного навчання

Література

1. S. Ghahseri-Schwartz, S. Ben-David. Understanding Machine Learning.
2. M. Mohri, A. Rostamizadeh, A. Talwaker. Foundations of Machine Learning.
3. D.J. MacKay. Information Theory, Inference, and Learning Algorithms.

Машинне навчання ^(ML) — область інформатики (більш точно, н.зв. штучного інтелекту), що досліджує статистичні алгоритми, які здатні модифікувати себе з використанням відомих даних ("навчатися") і робити змістовні висновки з раніше невідомих даних ("застосовувати

надбати знання").

Застосування: у загальному, де побудова явного алгоритма на протилежну статистичному, ("традиційний" підхід штучного інтелекту - імітація людської мислення) дуже складна (як правило, практично неможлива).

Зокрема,

- задачі, з якими добре вправляються люди або тварини: розпізнавання образів, мови, інші задачі класифікації (виявлення стема), переклад, спілкування природними мовами, водична транспортних засобів;

- задачі, навпаки, за межами людських можливостей: аналіз великих об'ємів даних у науці, медицині, бізнесі, знаходження прихованих закономірностей.

З іншого боку, відмінність від статистики - фокус на алгоритмах, відсутність наперед визначеної гіпотези про розподіл даних.

Наші задачі у цьому курсі:

1. Побудувати загальні математичні моделі машинного навчання та дослідити їхні властивості.

2. Застосувати цей загальний підхід до конкретних алгоритмів ML.

Більш точно, нас цікавить перше за все кероване статистичне пакетне пасивне навчання. Значення цих слів:

Кероване (з учителем supervised) - гради дані, на яких алгоритм навчається, ^(training data) мають суцільну інформацію, що відсутня у дані, до яких він застосовується і яку потрібно відновити (мітки класифікації, розмірочки, переклади, продовження фраз...) на відміну від некерованого (без учителя, unsupervised), де навчають на тестові (робочі) дані не відносяться (наприклад, задачі кластеризації)

Статистичне у цьому контексті означає, що навчальні дані генеруються деяким випадковим процесом, на fitting big ситуації "виперед", що допомагає чи забавляє.

Пакетне (batch learning) — алгоритм застосовується до тестових даних лише після обробки суттєвої кількості навчальних, на fitting big онлайн-навчання, де ці процеси можуть відбуватися одночасно.

Пасивне — алгоритм лише сприймає інформацію, що йому надходить, на fitting big активно, коли він може робити запити до оточення під час навчання.