



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Видове машинно обучение

В темата ще бъдат разгледани следните основни въпроси:

- Основните видове машинно обучение: с надзор, без надзор и с утвърждаване
- Модели за машинно обучение: с надзор, без надзор и с утвърждаване
- Сравнение между изкуствен интелект и машинно обучение

----- www.eufunds.bg -----

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.



1. Основни видове машинно обучение

Както бе споменато в първа глава машинното обучение може да се раздели на 3 основни групи: с надзор, без надзор и с утвърждаване.

1.1. Обучение с надзор

Обучението с надзор (Supervised Learning) е вид метод на машинно обучение, при който се предоставят примерни етикетирани данни на системата за машинно обучение, за да се обучи и на тази база да се прогнозира изхода.

Системата създава модел, използвайки етикетирани данни, за да разбере съвкупността от данни и да се обучи за всеки тип данни. След завършване на обучението и обработката се тества модела, като му се предоставят примерни данни, за да се провери дали той прогнозира точния резултат или не.

Целта на обучението с надзор е да съпостави входните данни с изходните данни. Обучението с надзор е подобно ученето на ученика под надзора на учителя. Пример за обучение с надзор е филтрирането на спам.

Обучението с надзор може да бъде групирано допълнително в две категории алгоритми: класификация (Classification) и регресия (Regression).

----- www.eufunds.bg -----



1.2. Обучение без надзор

Обучението без надзор е метод на обучение, при който машината се учи без никакъв надзор. Обучението се предоставя на машината с набор от данни, които не са етикетирани, класифицирани или категоризирани, и алгоритъмът трябва да действа върху тези данни без никакъв надзор. Целта на обучението без надзор е да преструктурира входните данни в нови характеристики или група обекти с подобни модели. При обучението без надзор няма предварително определен резултат. Машината се опитва да намери полезна информация от огромното количество данни. Обучението без надзор може допълнително да бъде класифицирано в две категории алгоритми: клъстеризация (Clustering) и асоциация (Association).

1.3. Обучение с утвърждаване

Обучението с утвърждаване е метод на обучение, базиран на обратна връзка, при който обучаващ агент получава награда за всяко правилно действие и получава наказание за всяко грешно действие. Агентът се учи автоматично с тези обратни връзки и подобрява ефективността си. При обучението с утвърждаване агентът взаимодейства със средата и я изследва. Целта на агента е да максимизира мярката за кумулативно възнаграждение и следователно така да подобрява своето представяне. Примери за обучение с утвърждаване са роботизираното куче, което автоматично научава

----- www.eufunds.bg -----



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

движението на ръцете си, автономните автомобили и игра срещу противник човек.

2. Модели за машинно обучение

Моделът на машинно обучение се определя като математическо представяне на изхода от процеса на обучение. Машинното обучение е изследване на различни алгоритми, които могат да се подобряват автоматично въз основа на стари данни, натрупване на опит и изграждане на модел. Моделът на машинно обучение е подобен на компютърен софтуер, предназначен да разпознава поведения въз основа на предишен опит или данни. Алгоритъмът за обучение открива закономерности в данните за обучение и извежда ML модел, който прави прогнози за нови данни.

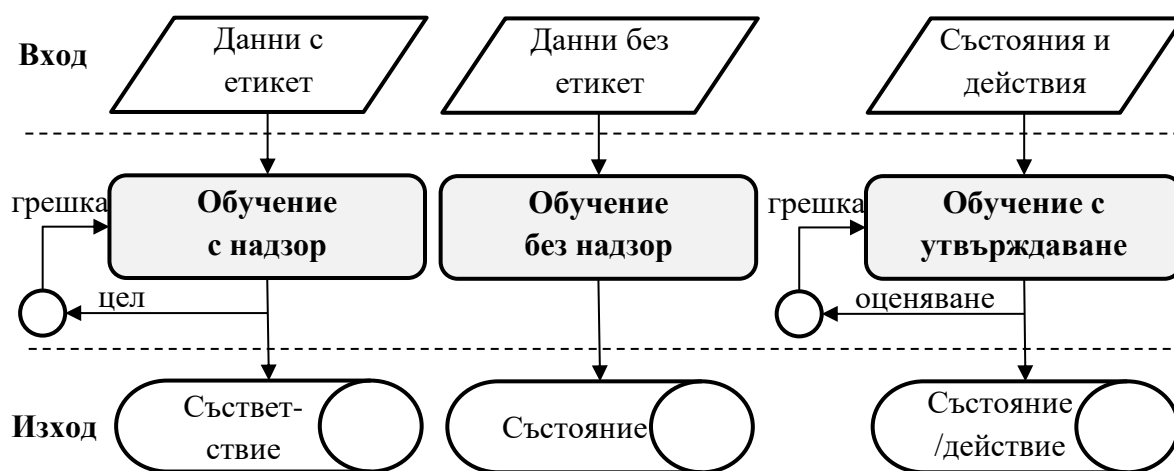
Моделите на машинно обучение могат да се разбират като програма, която е обучена да намира еталонни шаблони в нови данни и да прави прогнози. Тези модели са представени като математическа функция, която приема заявки под формата на входни данни, прави прогнози за входните данни и след това осигурява определен изход в отговор на заявката. Първо, тези модели се обучават върху множество от данни и след това те се предоставят на алгоритъм за осмисляне на данните, извличане на шаблон от подадените данни и учене от тези данни. След като тези модели се обучат, те могат да се използват за прогнозиране на необработени множества от данни.

----- www.eufunds.bg -----

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.



Трите основни модели за машинно обучение са представени схематично на фиг. 1.3.



Фиг. 1.3. Модели за машинно обучение

2.1. Модели за машинно обучение с надзор

Обучението с надзор е най-простият за разбиране модел на машинно обучение, в който има точно съответствие между входните данни, които се наричат данните за обучение с известен етикет и резултата като изход. Най-просто казано, моделът работи на принципа на входно-изходните двойки. Изисква се създаване на функция, която може да бъде обучена с помощта на множество от данни за обучение, след което се прилага към неизвестни данни и прави прогноза за тях. Обучението с надзор е базирано на задачи обучение и тествано върху множества от данни с етикети.

----- www.eufunds.bg -----



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Моделът на обучението с надзор може да се приложи върху прости проблеми от реалния живот. Например, при множество от данни, състоящо се от възраст и височина, след това може да се изгради модел на обучение с надзор, за да се предвиди височината на човека въз основа на възрастта му.

Моделите на обучение с надзор се класифицират допълнително в две категории: модели за регресия и класификация

При *модели за регресия* изходът е непрекъсната променлива. Някои често използвани модели за регресия са: линейна регресия (Linear Regression), дърво на решенията (Decision Tree), случайна гора (Random Forest) и невронни мрежи (Neural Networks).

Моделите за класификация са вторият тип модели за обучение с надзор, които се използват за генериране на категорични заключения от наблюдаваните стойности. Например, моделът за класификация може да идентифицира дали имейлът е спам или не, дали купувачът ще закупи продукта или не и т.н. Алгоритмите за класификация се използват за прогнозиране на два класа или за категоризиране на изхода в различни групи.

При моделите за класификация се проектира класификатор, който идентифицира множеството от данни в различни категории и на всяка категория се присвоява етикет.

В машинното обучение съществуват два вида класификации:

Двоична класификация: Ако проблемът има само два възможни класа, то моделът се нарича двоичен класификатор.

----- www.eufunds.bg -----

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.



Многокласова класификация: Ако проблемът има повече от два възможни класа, това е моделът е многокласов класификатор.

Някои често използвани модели за класификация са: логистична регресия (Logistic Regression), опорни вектори SVM (Support Vector Machine) и мрежи на Бейс (Naïve Bayes).

2.2. Модели за машинно обучение без надзор

Моделите за машинно обучение без надзор прилагат процес на обучение, противоположен на тези с надзор, което означава, че позволява на модела да се учи от множество от данни за обучение без етикети. Въз основа на тези неетикирани данни моделът прогнозира изхода. Използвайки обучение без надзор, моделът научава *скрити модели* от множеството от данни сам без надзор.

Моделите на обучение без надзор се използват главно за изпълнение на три основни задачи: клъстеризация (Clustering), правила за асоциация (Association Rule Learning) и намаляване на размерността (Dimensionality Reduction).

2.3. Модели за обучение с утвърждаване

При обучението с утвърждаване алгоритъмът научава действия за дадено множество от състояния, които водят до дадено целево изходно

----- www.eufunds.bg -----



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

състояние. Това е модел на обучение, базиран на обратна връзка, който приема сигнали за обратна връзка след всяко състояние или действие чрез взаимодействие с околната среда. Тази обратна връзка работи като награда (положителна за всяко добро действие и отрицателна за всяко лошо действие), а целта на агента е да увеличи максимално положителните награди, за да подобри своето представяне.

Поведението на модела при учене с утвърждаване е подобно на човешкото учене, тъй като хората научават неща чрез опит като обратна връзка и взаимодействие с околната среда.

Някои популярни алгоритми, които попадат в обучението с утвърждаване, са Q-обучение (Q-learning), основано на равенството на Белман (Bellman equation); състояние-действие-възнаграждение-състояние-действие SARSA (State-Action-Reward-State-Action), който е базиран на политика алгоритъм, вземащ решение на основата на Марков процес, и дълбока Q мрежа DQN (Deep Q Network или Deep Q Neural network), който е Q-обучение в невронни мрежа.

3. Сравнение между изкуствен интелект и машинно обучение

В заключение ще бъде разгледана основната разлика между AI и ML:

„Изкуственият интелект AI е по-голяма концепция за създаване на интелигентни машини, които могат да симулират човешко мислене и поведение, докато машинното обучение ML е приложение или подмножество

----- www.eufunds.bg -----

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.



от AI, което позволява на машините да се учат от данни, без да бъдат програмирани изрично.“ [https://www.javatpoint.com/difference-between-artificial-intelligence-and-machine-learning

AI обхваща много други дисциплини, за да симулира успешно мисловния процес. В допълнение към машинното обучение, AI обикновено включва:

1. **Обработка на естествен език:** Процесът на въвеждане на естествен език и преобразуването му във форма, която компютърът може да използва.
2. **Разбиране на естествен език:** Процесът на дешифриране на естествения език, за да се въздейства върху смисъла, който предоставя.
3. **Представяне на знания:** Способността да се съхранява информация във форма, която прави възможен бърз достъп.
4. **Планиране** (под формата на търсене на цел): Способността да се използва съхранена информация, за да се правят заключения в реално време (почти на момента, в който се случват действията, и с малко забавяне, понякога толкова кратко, че човек да не забележи).
5. **Роботика:** Способността да се действа по заявки от потребител в някаква физическа форма.

Таблица 1.1. обобщава основните разлики между изкуствения интелект и машинното обучение

----- www.eufunds.bg -----



Таблица 1.1. Основни разлики между изкуствения интелект и машинното обучение

Характеристика	Изкуствен интелект AI	Машинно обучение ML
Същност	Изкуственият интелект е технология, която позволява на машината да симулира човешко поведение.	Машинното обучение е подмножество от AI, което позволява на машината автоматично да се учи от минали данни, без да програмира изрично.
Цел	Да се създаде интелигентна компютърна система имитираща човека, която да решава сложни проблеми. Създават се интелигентни системи, които да изпълняват всяка задача като човек.	Да позволи на машините да се учат от данни, така че да могат да дават точен резултат. Машините се обучават с данни, за да изпълняват определена задача и да дават точен резултат.
Подгрупи	Машинно обучение и дълбоко обучение	Дълбоко обучение
Обхват	Много широк обхват - AI се	Ограничен обхват - ML се

----- www.eufunds.bg -----



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

	стреми да създаде интелигентна система, която може да изпълнява различни сложни задачи.	стреми да създаде машини, които могат да изпълняват само онези специфични задачи, за които са обучени.
Заинтересованост	за максимизирането на шансовете за успех.	главно за точността и моделите.
Основни приложения	поддръжка на клиенти с помощта на chatboats, експертна система, игра на онлайн игри, интелигентен хуманоиден робот и др.	онлайн система за препоръчване, алгоритми за търсене в Google, предложения за автоматично маркиране на приятели във Facebook и др.
Класификация спрямо възможностите	три типа, които са ограничен AI, общ AI и супер AI	три типа, които са обучение с надзор, обучение без надзор и обучение с утвърждаване.
Обхват	учене, разсъждение и самокорекция	учене и самокорекция при запознаване с нови данни
Данни	структурирани, полуструктурирани и неструктурирани данни	структурирани и полуструктурирани данни

----- www.eufunds.bg -----

Проект BG05M2OP001-2.016-0003 „Модернизация на Национален военен университет "В. Левски" - гр. Велико Търново и Софийски университет "Св. Климент Охридски" - гр. София, в професионално направление 5.3 Компютърна и комуникационна техника“, финансиран от Оперативна програма „Наука и образование за интелигентен растеж“, съфинансирана от Европейския съюз чрез Европейските структурни и инвестиционни фондове.