* با وجود حجم عظیم داده ها ،تمرکز هوش مصنوعی به سمت توانمندی های یادگیری ماشین و یادگیری عمیق معطوف شده است ،جایی که ماشین ها می‌توانند به طور مستقل یاد بگیرند .تفاوت هوش مصنوعی **(AI)** و یادگیری ماشین در این است که هوش مصنوعی هوشمند است اما تعامل مستقیمی با محیط ندارد و نمی‌تواند از آن یاد بگیرد.

*عامل **(Agent)** سیستمی است که می‌تواند اطلاعاتی از محیط جمع آوری کرده و براساس آن عمل کند و یک خروجی داشته باشد. عامل می‌تواند یک نرم افزار یا الگوریتم باشد.خودمختاری **(Autonomy)** به میزان یادگیری عامل از تجربیات خود در محیط اشاره دارد و به عامل امکان می‌دهد به صورت مستقل تر و خودمختار تر عمل کند .عملکرد **(Performance)**معیاری برای ارزیابی رفتار عامل است که تعیین می‌کند آیا عامل وظایف مورد انتظار خود را در محیط به درستی انجام می‌دهد یا خیر.اهداف**(Goals)** مقاصدی هستند که عامل برای رسیدن به آن‌ها تلاش می‌کند .در یادگیری ماشین ،هدف تنها یادگیری نیست بلکه رسیدن به یک مأموریت کلی است .سودمندی **(Utility)** به ارزیابی عملکرد داخلی عامل اشاره دارد که به آن اجازه می‌دهد در صورت وجود چندین مسیر برای رسیدن به هدف ،بهترین مسیر را با تصمیم گیری منطقی انتخاب کند .این معیار با ارزیابی عملکرد مدل متفاوت است و بیشتر به ارزیابی درونی خود مدل توجه دارد . محیط **(Environment)** جهانی است که اطراف عامل را توصیف می‌کند و ویژگی هایی مانند دسترس پذیری،قطعی یا غیرقطعی بودن ،پویایی ،و اپیزودیک یا زمان واقعی بودن را شامل می‌شود .همچنین سرعت و نوع داده هایی که وارد محیط می‌شوند بر عملکرد آن تأثیر می‌گذارد .ویژگی **(Feature)** به خصوصیات داده ها و مقادیر آن‌ها اشاره دارد .به عنوان مثال ((رنگ پوست قهوه ای است)) یک ویژگی است که در آن رنگ پوست <<صفت >>و قهوه ای <<مقدار >> آن است .در واقع ،ویژگی ها در آموزش مدل ها مشابه متغیرها **(Variable)** در آمار هستند ،اما می‌توانند یک متغیر یا ترکیبی از چند متغیر باشند .اگر ویژگی ها در یک بُعد قابل تعریف نباشند ،می‌توان از روش هایی برای افزودن بُعد ماتریکس ها به مدل استفاده کرد تا این ویژگی ها قابل فهم و تحلیل شوند .انتخاب ویژگی **(Feature Selection)** فرآیندی است که ویژگی های مهم و موردنیاز برای پیش بینی خروجی های یک مدل آماری را از میان ویژگی های مختلف انتخاب می‌کند ،چه ویژگی های اصلی چه آن‌هایی که طی مهندسی ویژگی **(Feature Engineering)** ساخته می شوند.

* یادگیری ماشین به چهار دسته تقسیم می‌شود:

۱) یادگیری نظارت شده ۲) یادگیری بدون نظارت ۳) یادگیری نیمه نظارتی ۴) یادگیری تقویتی

* یادگیری نظارت شده(Supervised Learning) : در این نوع یادگیری ،الگوریتم ها با داده های آموزشی شامل ورودی‌ها و خروجی های مشخص آموزش می‌بینند تا قوانینی عمومی را یاد بگیرند و مدلی بسازند که بتواند پیش‌بینی کند . در واقع ،الگوریتم با استفاده از داده های برچسب‌دار می‌تواند رویدادهای آینده را پیش‌بینی کند و خروجی را با خروجی درست مقایسه کرده و مدل را بر این اساس اصلاح کند .به این دلیل به آن نظارت شده می‌گویند که نیاز به یک ناظر انسانی برای برچسب گذاری یا لیبل زدن داده ها دارد .آموزش تا زمانی ادامه می‌یابد که مدل به سطح دقت مطلوب برسد.

این یادگیری شامل انواع مختلفی است:

-طبقه بندی (Classification) : پیش‌بینی دسته بندی ها براساس داده های آموزشی ،مانند تشخیص اینکه یک فرد بیمارهست یا خیر .تکنیک های طبقه بندی مرزهای تصمیم گیری را تعریف می‌کنند و شامل روش هایی مانند ماشین بردار پشتیبان(Support Vector Machine) ، بیز ساده(Naive Bayes) ، بیز گاوسی(Gaussian Bayes) ، نزدیک‌ترین همسایه (k-Nearest Neighbors) و رگرسیون لجستیک (Logistic Regression) هستند.

-رگرسیون(Regression) : پیش‌بینی مقادیر عددی در داده های پیوسته .از مدل‌های معروف رگرسیون می‌توان به رگرسیون خطی(Linear Regression) ، رگرسیون چندجمله‌ای(Polynomial Regression) ، ماشین بردار پشتیبان(SVR) ، تجمیع‌ها(Ensembles) ، درخت‌های تصمیم (Decision Trees) و شبکه‌های عصبی (Neural Networks)اشاره کرد.

-پیش بینی (Forecasting) : استفاده از داده های سری زمانی برای پیش بینی براساس داده های گذشته و حال است.

-تجمیع (Ensemble) : این روش چندین مدل یادگیری ماشین مختلف را ترکیب می‌کند تا نتیجه بهتری روی یک نمونه جدید ارائه دهد ،به طوری که مدل ها به عنوان ویژگی ها عمل می‌کنند.

* هیچ یک از این مدل‌ها به صورت مطلق برتری خاصی ندارند و انتخاب مدل به هدف تحقیق بستگی دارد .در بسیاری از موارد ،چند مدل آموزش داده می‌شوند و بهترین مدل براساس معیارهای ارزیابی عملکرد انتخاب می‌شود.