

## مروری کوتاه بر الگوریتم‌های یادگیری ماشین

علیرضا شبانی

گروه مهندسی صنایع، دانشکده‌ی مهندسی، دانشکدگان فارابی، دانشگاه تهران، تهران، ایران

### چکیده

یادگیری ماشین، یک زمینه برجسته در هوش مصنوعی، توجه زیادی را در حوزه دیجیتال به عنوان یک عنصر حیاتی از راه‌حل‌های هوشمند به خود جلب کرده است. هدف این مقاله ارائه‌ی یک نمای کلی و مختصر از الگوریتم‌های یادگیری ماشینی است که به طور معمول مورد استفاده قرار می‌گیرند. بررسی پرکاربردترین و محبوب‌ترین الگوریتم‌های یادگیری ماشین، ارزیابی نقاط قوت و ضعف آن‌ها از منظر کاربردی و بینش‌های به‌دست‌آمده از این بررسی‌ها به تصمیم‌گیرندگان در انتخاب مناسب‌ترین الگوریتم یادگیری که به بهترین وجه با نیازهای خاص برنامه‌هایشان مطابقت دارد، کمک می‌کند. تجزیه و تحلیل این پژوهش عواملی مانند دقت، کارایی محاسباتی، پایداری و سازگاری با انواع داده‌ها و چالش‌ها را در نظر می‌گیرد. درک محاسن و معایب این الگوریتم‌ها، تصمیم‌گیری آگاهانه، تضمین عملکرد بهینه و اجرای موفقیت‌آمیز در سناریوهای دنیای واقعی را تسهیل خواهد کرد.

واژگان کلیدی: الگوریتم‌های یادگیری ماشین، هوش مصنوعی، دیدگاه کاربردی، تصمیم‌گیری

## مقدمه

نقطه شروع مناسب در این مقاله، شروع با مفهوم اساسی یادگیری ماشین است. در یادگیری ماشین، یک برنامه رایانه‌ای با یادگیری از تجربه‌های خود، عملکرد خود را بهبود می‌دهد. این امر از طریق تصمیم‌گیری و پیش‌بینی بر اساس داده‌های گذشته حاصل می‌شود. به عنوان مثال، یک مدل را در نظر بگیرید که یاد می‌گیرد با تجزیه و تحلیل گزارش‌های پزشکی، بیماری سرطان را تشخیص دهد و یا پیش‌بینی کند. از آنجا که این مدل با بررسی داده‌های بیشتری از بیماران، تجربه بیشتری کسب می‌کند، عملکرد آن در پیش‌بینی صحیح و شناسایی موارد درست سرطان افزایش می‌یابد. این عملکرد با مقایسه نتایج مدل در مقابل نتایج تأیید شده توسط یک متخصص انکولوژیست با تجربه ارزیابی می‌شود [1-3].

یادگیری ماشینی کاربردهایی را در زمینه‌های مختلف مانند بازارهای مالی [4]، مراقبت‌های پزشکی و سلامت [5]، سیستم‌های توصیه‌گر [6]، مدیریت زنجیره‌ی تامین [7] و ... ارائه می‌دهد.

یادگیری ماشین بطور کلی شامل چهار دسته‌ی یادگیری تحت‌نظارت، یادگیری بدون‌نظارت، یادگیری نیمه‌نظارتی و یادگیری تقویتی است. انتخاب الگوریتم یادگیری ماشین مناسب به در دسترس بودن و ماهیت داده‌های آموزشی بستگی دارد [1-3]. در بخش‌های بعدی مروری بر برخی از الگوریتم‌های یادگیری ماشین که به طور گسترده استفاده می‌شوند، خواهد شد.

## الگوریتم‌های یادگیری ماشین

### ۱. گرادیان کاهشی (Gradient Descent)

گرادیان کاهشی یک روش بهینه‌سازی تکرار شونده است که برای به حداقل رساندن یک تابع از جنس هزینه استفاده می‌شود. این روش متکی بر محاسبه‌ی مشتق جزئی است که نشان دهنده‌ی شیب یا گرادیان تابع می‌باشد. در هر تکرار، ضرایب با هدف نزدیک شدن به حداقل‌های محلی پس از چند تکرار، با تفریق مشتق ضرب در نرخ یادگیری (اندازه گام) به‌روزرسانی می‌شوند. همگرایی زمانی حاصل می‌شود که الگوریتم به حداقل مقدار تابع هزینه برسد و در نتیجه کاهش بیشتری حاصل نشود [8]. سه حالت مختلف از روش گرادیان کاهشی وجود دارد که عبارتند از: گرادیان کاهشی تصادفی (Stochastic Gradient Descent)، گرادیان کاهشی دسته‌ای (Batch Gradient Descent) و گرادیان کاهشی دسته‌ای کوچک (Mini-Batch Gradient Descent) [8, 9].

در گرادیان کاهشی دسته‌ای، خطا برای هر نمونه در مجموعه داده آموزشی محاسبه می‌شود، اما به‌روز رسانی مدل تنها پس از ارزیابی همه نمونه‌های آموزشی رخ می‌دهد. مزیت اصلی الگوریتم BGD کارایی محاسباتی است. این الگوریتم یک گرادیان خطای پایدار و یک همگرایی پایدار تولید می‌کند. با این حال الگوریتم این اشکال را دارد که گرادیان خطای پایدار گاهی اوقات می‌تواند منجر به حالت همگرایی شود که بهترین حالتی نیست که مدل می‌تواند به دست آورد. علاوه بر این، BGD نیاز دارد که کل مجموعه داده آموزشی در حافظه موجود باشد [8-10].

در گرادیان کاهشی تصادفی خطا برای هر نمونه آموزشی در مجموعه داده محاسبه می‌شود و پارامترها برای هر نمونه آموزشی به روز خواهند شد. SGD می‌تواند برای مسائل خاص سریعتر از BGD باشد. SGD این مزیت را دارد که به‌روز رسانی‌های مکرر منجر به بهبود نرخ دقت می‌شود. با این حال، به‌روز رسانی‌های مکرر هزینه‌های محاسباتی را افزایش خواهد داد [8-10].

رویکرد گرادیان کاهشی دسته‌ای کوچک با ترکیب مفاهیم SGD و BGD به دست می‌آید. در این رویکرد مجموعه داده آموزشی به دسته‌های کوچک تقسیم می‌شود و برای هر یک از این دسته‌ها یک به‌روز رسانی انجام خواهد شد. بنابراین تعادلی بین پایداری SGD و کارایی BGD ایجاد می‌کند. از این الگوریتم می‌توان برای آموزش شبکه عصبی استفاده کرد و بنابراین بیشتر استفاده‌ی این الگوریتم در یادگیری عمیق است. رویکرد بهینه‌سازی گرادیان نزول در الگوریتم پس انتشار استفاده شده است که در آن از تابع گرادیان برای تنظیم وزن نوروها کمک گرفته می‌شود [8-10].

الگوریتم گرادیان کاهشی یک اشکال بالقوه دارد: اگر نرخ یادگیری بیش از حد بالا باشد، ممکن است به خاطر سرعت بهینه‌سازی، مقدار حداقل محلی واقعی را نادیده بگیرد، در حالی که سرعت یادگیری بسیار پایین ممکن است از همگرایی جلوگیری کند. انتخاب نرخ یادگیری بر رسیدن به مقدار حداقل و سرعت همگرایی تأثیر می‌گذارد. توصیه می‌شود از یک نرخ یادگیری در حال تغییر استفاده کنید که با کاهش خطا، سرعت آن کاهش می‌یابد [8-10].

## ۲.۲. رگرسیون خطی (Linear Regression)

رگرسیون یک رویکرد یادگیری تحت نظارت است. می‌توان از آن برای مدل سازی متغیرهای پیوسته و انجام پیش‌بینی استفاده کرد. نمونه‌هایی از کاربرد الگوریتم رگرسیون خطی عبارتند از: پیش‌بینی قیمت ملک، پیش‌بینی فروش، پیش‌بینی نمرات امتحانات دانش‌آموزان و پیش‌بینی حرکت قیمت سهام در بورس [11].

در رگرسیون مجموعه داده‌های برجسب‌دار و خروجی را داریم که خروجی توسط مقادیر متغیر ورودی تعیین می‌شود. بنابراین رویکرد یادگیری تحت نظارت است [11].

رگرسیون خطی چندین مزیت از جمله سهولت درک آن و همچنین مدیریت ساده‌ی بیش‌برازش مدل از طریق تکنیک‌های منظم‌سازی را ارائه می‌دهد. از طرفی می‌توان از SGD برای به‌روزرسانی مدل‌های خطی با داده‌های جدید استفاده کرد [11].

اگر رابطه‌ی بین متغیر کمکی و متغیر خروجی خطی باشد، رگرسیون خطی مناسب است. رگرسیون خطی تمرکز را از مدل‌سازی آماری به تجزیه و تحلیل داده‌ها تغییر می‌دهد. همچنین این روش برای یادگیری درمورد فرآیند تجزیه و تحلیل داده‌ها مناسب است. البته این الگوریتم علی‌رغم محاسن آن به دلیل ساده‌انگاری بیش از حد مسائل پیچیده‌ی دنیای واقعی، برای کاربردهای عملی توصیه نمی‌شود. رگرسیون خطی در مواجهه با مسائل غیرخطی نیز عملکرد مناسبی ندارد و مدیریت الگوهای پیچیده برای این الگوریتم دشوار است [11].

## ۳. رگرسیون چندمتغیره (Multivariate Regression)

یک مدل رگرسیون خطی ساده دارای یک متغیر وابسته تحت تأثیر یک متغیر مستقل است که به اندازه‌ی کافی پیچیدگی مسائل دنیای واقعی را پوشش نمی‌دهد. درواقع یک متغیر وابسته از عوامل متعددی تأثیر می‌پذیرد. به عنوان مثال، قیمت یک خانه به عواملی همچون موقعیت مکانی، اندازه، تعداد اتاق‌ها، امکانات رفاهی، نزدیکی به مراکز حمل و نقل و دسترسی به مراکز خرید بستگی دارد [12, 13]. در رگرسیون خطی ساده، رابطه‌ای یک به یک بین متغیر ورودی و متغیر خروجی برقرار است در صورتیکه در رگرسیون خطی چندگانه، رابطه‌ای چند به یک بین تعدادی از متغیرهای ورودی و یک متغیر خروجی وجود دارد [12, 13].

افزودن متغیرهای ورودی بیشتر به این معنی نیست که رگرسیون بهتر خواهد بود یا پیش‌بینی‌های بهتری ارائه می‌دهد. رگرسیون خطی چندگانه و ساده موارد کاربرد متفاوتی دارند و یکی از دیگری برتر نیست. در برخی موارد افزودن متغیرهای ورودی بیشتر می‌تواند اوضاع را بدتر کند زیرا منجر به بیش‌برازش خواهد شد. علاوه بر این، با معرفی متغیرهای ورودی اضافی، روابط متقابل بین آن‌ها می‌تواند پدیدار شود که منجر به چند خطی شدن خواهد شد. سناریوی ایده آل مستلزم آن است که هر متغیر ورودی با متغیر خروجی همبستگی نشان دهد و در عین حال استقلال خود را از یکدیگر حفظ کند [12, 13].

تکنیک چندمتغیره دارای مزیت‌های از جمله ارائه‌ی درکی جامع از روابط بین مجموعه متغیرهای مستقل و متغیرهای وابسته است. همچنین رگرسیون چندمتغیره مسائل دنیای واقعی را به شیوه‌ای واقع‌بینانه و عملی‌تر مدل می‌کند [12, 13].

## ۴. رگرسیون لجستیک (Logistic Regression)

از رگرسیون لجستیک برای مسائل کلاس‌بندی استفاده می‌شود و متغیر خروجی احتمال وقوع یک رویداد است. به عنوان مثال، پیش‌بینی خوش‌خیم یا بدخیم بودن یک تومور و یا کلاس‌بندی یک ایمیل به عنوان هرزنامه یا غیر هرزنامه. بنابراین رگرسیون لجستیک

با پیش‌بینی متغیر هدف سروکار دارد درحالی‌که رگرسیون خطی برای پیش‌بینی مقادیر متغیرهای پیوسته مانند پیش‌بینی قیمت املاک طراحی شده است [14].

رگرسیون لجستیک چندین مزیت از جمله سهولت اجرا، کارایی محاسباتی، کارایی در آموزش و منظم‌سازی ساده را ارائه می‌دهد. برخلاف برخی دیگر از الگوریتم‌ها، رگرسیون لجستیک نیازی به مقیاس‌بندی ویژگی ندارد. این الگوریتم عمدتاً برای حل مسائل در مقیاس صنعتی استفاده می‌شود. با توجه به اینکه خروجی رگرسیون لجستیک یک احتمال (عددی بین ۰ و ۱) است، نیاز به تعیین حد آستانه دارد تا طبقه‌بندی را تسهیل کند. رگرسیون لجستیک در برابر نویزهای کوچک در داده‌ها و چندخطی‌ها مقاوم است [14].

با این وجود این الگوریتم دارای معایبی مثل ناتوانی در حل مسائل غیرخطی به دلیل سطح تصمیم‌گیری ذاتا خطی می‌باشد. همچنین رگرسیون لجستیک مستعد بیش‌برازش است و برای عملکرد موثر به شناسایی همه‌ی متغیرهای مستقل نیاز دارد [14]. کاربردهای عملی رگرسیون لجستیک فراوان است که شامل پیش‌بینی خطر بیماری‌ها، تشخیص سرطان، پیش‌بینی مرگ‌ومیر بیماران و کاربردهای مهندسی مانند ارزیابی احتمال شکست در فرآیندها، سیستم‌ها یا محصولات می‌شود [14].

## ۵. درخت تصمیم (Decision Tree)

الگوریتم درخت تصمیم یک تکنیک یادگیری ماشینی نظارت‌شده است که در مسائل طبقه‌بندی و رگرسیون به کار می‌رود. این روش با تقسیم مکرر مجموعه داده بر اساس معیارهای خاص عمل می‌کند. در درخت تصمیم، برگ‌ها به عنوان تصمیمات نشان داده می‌شوند و پارتیشن‌بندی داده‌ها در گره‌ها نشان داده می‌شوند. در درخت تصمیم کلاس‌بندی متغیر تصمیم بصورت بله و خیر خواهد بود و در درخت تصمیم رگرسیون متغیر تصمیم پیوسته است [3, 15, 16].

درختان تصمیم برای مسائل کلاس‌بندی و رگرسیون کاربرد دارند و از درجه‌ی بالای تفسیرپذیری بهره می‌برند. همچنین می‌توانند بطور موثر انواع ویژگی‌ها را مدیریت کرده و در مواردی که مقدار یک ویژگی گم‌شده است، محتمل‌ترین مقدار نسبت به مقادیر دیگر را اختصاص دهند و عملکردی کارا داشته باشند. با همه‌ی این تفاسیر اما درخت تصمیم مستعد بیش‌برازش است که البته الگوریتم جنگل تصادفی که از رویکرد مدل‌سازی گروهی استفاده می‌کند، راه‌حل این مشکل است.

این الگوریتم می‌تواند ناپایدار و غیرقابل کنترل در اندازه‌ی درخت باشد. همچنین ممکن است در نمونه‌گیری دچار خطا شود و درنهایت یک بهینه‌ی محلی را ارائه دهد [15, 16].

علی‌رغم این محدودیت‌ها، درخت تصمیم در حوزه‌های مختلف مانند پیش‌بینی استفاده از یک کتاب در کتابخانه و یا پیش‌بینی ایجاد یک تومور کاربرد عملی دارد [15, 16].

## ۶. ماشین بردار پشتیبان (Support Vector Machine)

ماشین‌های بردار پشتیبان به عنوان ابزار همه‌کاره‌ای عمل می‌کنند که توانایی حل مسائل کلاس‌بندی و رگرسیون را دارند. در این روش باید یک ابرصفحه تعریف شود که در اصل مرز تصمیم است. هنگام مواجهه با مجموعه‌ای از نمونه‌های متعلق به کلاس‌های مختلف به این ابرصفحه برای جدا کردن این نمونه‌ها از یکدیگر نیاز خواهد بود. نمونه‌ها ممکن است بصورت خطی قابل تفکیک باشند یا نباشند که در اینصورت توابع ریاضی پیچیده به نام کرنل‌ها برای جداسازی نمونه‌هایی که اعضای کلاس‌های مختلف هستند مورد نیاز می‌باشند. هدف این روش کلاس‌بندی صحیح نمونه‌ها بر اساس الگوهای مشاهده شده در مجموعه داده‌ی آموزشی است [3, 17].

ماشین بردار پشتیبان می‌تواند داده‌های ساختار یافته و نیمه ساختار یافته را مدیریت کرده و در صورتیکه بتوان کرنل‌های مناسب را بدست آورد توانایی مدیریت و حل مسائل پیچیده را دارد. احتمال بیش‌برازش در این روش کمتر است و داده‌های با ابعاد بالا را پشتیبانی می‌کند. همچنین ماشین بردار پشتیبان در بهینه‌ی محلی گیر نخواهد کرد [17].

عملکرد این الگوریتم با مجموعه داده‌های بزرگ به دلیل افزایش زمان آموزش کاهش می‌یابد. علاوه بر این هنگامی که مجموعه داده دارای نویز باشد، SVM خوب کار نمی‌کند. این روش قابلیت تخمین احتمالی را ندارد و درک و تفسیرپذیری مدل نهایی آن ممکن است کمی دشوار باشد [17].

ماشین بردار پشتیبان در تشخیص سرطان، تشخیص تقلب در کارت‌های اعتباری، تشخیص دست‌خط، تشخیص چهره و طبقه‌بندی متن و ... کاربرد دارد [17].

هنگام انتخاب از بین رگرسیون لجستیک، درخت تصمیم و ماشین بردار پشتیبان، توصیه می‌شود که رگرسیون لجستیک استفاده شده سپس با بکارگیری درخت تصمیم، بهبود مدل ارزیابی شود. زمانیکه که مجموعه داده دارای تعداد زیادی نمونه و ویژگی است، SVM به یک گزینه قابل اعتماد تبدیل می‌شود.

## ۷. یادگیری بیزی (Bayesian Learning)

در یادگیری بیزی، فرآیند با انتخاب یک توزیع احتمال قبلی آغاز می‌شود و سپس با استفاده از داده‌های موجود به‌روزرسانی شده تا توزیع پسین به دست بیاید. با بدست آمدن مشاهدات جدید می‌توان توزیع خلفی قبلی را به عنوان توزیع احتمال قبلی جدید مورد استفاده قرار داد [18].

شبکه‌های بیزی ابزاری برای مدیریت مجموعه داده‌های ناقص ارائه می‌دهند و می‌توانند خطر بیش‌برازش را بدون نیاز به حذف تناقضات داده‌ها کاهش دهند [18].

البته یادگیری بیزی دارای معایب خاصی است. انتخاب یک توزیع احتمال قبلی می‌تواند یک کار چالش برانگیز باشد و توزیع پسین بطور قابل توجهی تحت تاثیر انتخاب توزیع احتمال قبلی است. اگر این انتخاب درست نباشد منجر به پیش‌بینی‌های اشتباه خواهد شد [18]. یادگیری بیزی را می‌توان در مسائل تشخیص پزشکی و ... استفاده کرد [18].

## ۸. بیز ساده (Naïve Bayes)

الگوریتم بی ساده یک روش ساده بر اساس احتمال شرطی است. در این رویکرد یک جدول احتمال وجود دارد که مدل است و از طریق داده‌های آموزشی به روز می‌شود. این "جدول احتمال" بر اساس مقادیر ویژگی ساخته شده که امکان جستجوی احتمالات کلاس را در هنگام پیش‌بینی نمونه‌های جدید فراهم می‌کند. فرض کلیدی و زیربنای بیز ساده، شرط مستقل بودن متغیرهای ورودی است. لازم به ذکر است که در دنیای واقعی، به ندرت اتفاق می‌افتد که همه‌ی متغیرهای ورودی مستقل از یکدیگر باشند [3, 19].

پیاده‌سازی روش NB آسان است و عملکرد قابل قبولی ارائه می‌دهد. این روش با داده‌های آموزشی کمتر نیز کار می‌کند و داده‌ها را پیوسته و گسسته را پوشش می‌دهد. الگوریتم بیز ساده می‌تواند مسائل کلاس‌بندی دوکلاسه و چندکلاسه را مدیریت کرده و همچنین پیش‌بینی‌های احتمالی نیز ارائه دهد. بیز ساده پایداری خوبی در برابر ویژگی‌های نامربوط دارد [3, 19].

NB دارای محدودیت‌های از جمله نیاز به حافظه زمان اجرا بیشتر نسبت به الگوریتم‌های دیگر مثل SVM و LR است. همچنین در مدل‌هایی که متغیرهای زیادی دارند، محاسبات بسیار زیادی دربر خواهد داشت. بیز ساده در مقابل سایر مدل‌هایی که به درستی آموزش داده شده و تنظیم شده‌اند، معمولاً دقت کمتری دارد [3, 19].

از این الگوریتم برای کاربردهایی مثل سیستم‌های توصیه‌گر و پیش‌بینی پیشرفت سرطان پس از پرتودرمانی استفاده می‌شود [19].

## ۹. نزدیک‌ترین همسایه (K-Nearest Neighbor)

الگوریتم KNN یک تکنیک کلاس‌بندی است. این تکنیک از پایگاه‌داده‌ای استفاده می‌کند که دارای نقاط داده‌ای است که در دو یا چند کلاس، طبقه‌بندی شده‌اند. KNN سعی می‌کند نمونه داده شده به آن را به عنوان یک مسئله کلاس‌بندی، به یک کلاس اختصاص دهد. الگوریتم نزدیک‌ترین همسایه هیچ فرض اساسی درمورد توزیع داده‌ها انجام نمی‌دهد که بنابراین به آن ناپارامتریک می‌گویند [20, 21].

KNN چندین مزیت از جمله سادگی و سهولت در اجرا، مقرون به صرفه بودن در ساخت مدل و انعطاف پذیری بالا در طبقه‌بندی مجموعه داده‌های چندکلاسه ارائه می‌دهد. این روش مخصوصاً برای مجموعه داده‌های چندکلاسه بسیار مناسب است [20, 21]. این الگوریتم از نظر محاسباتی با اجرا در تکرار بالا بسیار فشرده خواهد شد. نزدیک‌ترین همسایه در برابر ویژگی‌های نامربوط پایدار نیست و منجر به کاهش دقت می‌شود. کلاس‌بندی داده‌های جدید با این الگوریتم از نظر محاسباتی پرهزینه است. همچنین در مجموع داده‌های با ابعاد بالا دقت این روش کاهش پیدا می‌کند. قابل توجه است که این الگوریتم در مقایسه با سایر الگوریتم‌ها، کند در نظر گرفته می‌شود [20, 21].

با وجود این محدودیت‌ها، KNN دارای طیف گسترده‌ای از کاربردها، از جمله سیستم‌های توصیه‌گر، تشخیص پزشکی برای بیماری‌هایی با علائم مشابه، تشخیص دست‌خط، تجزیه و تحلیل ریسک مؤسسات مالی برای تأییدیه‌های وام، پیش‌بینی الگوهای رأی‌گیری برای احزاب مختلف سیاسی و تشخیص تصویر می‌باشد [20, 21].

#### ۱۰. خوشه‌بندی کی-میانگین (K-Means Clustering)

الگوریتم K-Means یک روش برای حل مسائل خوشه‌بندی است که به عنوان یک تکنیک بدون نظارت در نظر گرفته می‌شود. این روش از نظر محاسباتی بسیار کارآمدتر از خوشه‌بندی سلسله‌مراتبی است به خصوص در زمانی که مجموعه داده بزرگ باشد. K-Means در مقایسه با سایر روش‌های خوشه‌بندی تمایل به ساخت خوشه‌های کوچک‌تر دارد که در نتیجه دقت الگوریتم را افزایش می‌دهد. علاوه بر کارایی بالای این روش، سهولت اجرا و تفسیرپذیری نتایج خوشه‌بندی از جذابیت‌های این روش می‌باشد [3, 22, 23]. از سوی دیگر، تعیین تعداد بهینه‌ی خوشه‌ها در این روش دشوار است. همچنین عملکرد آن ممکن است در هنگام مواجهه با خوشه‌های کروی کاهش پیدا کند. علاوه بر این از آنجایی که پارتیشن‌های اولیه مختلف منجر به خوشه‌های نهایی متفاوت می‌شوند، عملکرد می‌تواند دچار اختلال شود. K-Means ممکن است در دام بهینه محلی بیفتد و هیچ راه‌حل منحصر به فردی برای این مسئله وجود ندارد، که باعث می‌شود گاهی برای انتخاب بهترین نتیجه، نیاز به اجرای چندباره باشد. این روش را می‌توان برای خوشه‌بندی اسناد، خوشه‌بندی مشتریان، خوشه‌بندی خودکار هشدارهای فناوری اطلاعات، تشخیص تقلب‌های بیمه‌ای و ... استفاده کرد [22, 23].

#### بحث و نتیجه‌گیری

در این مقاله سعی شده است تا متداول‌ترین الگوریتم‌های یادگیری ماشین برای حل مسائل کلاس‌بندی، رگرسیون و خوشه‌بندی بررسی شود. مزایا و معایب این الگوریتم‌ها به همراه مقایسه الگوریتم‌های مختلف (در صورت امکان) از نظر عملکرد، سرعت یادگیری و ... نیز مورد بحث قرار گرفته است. در کنار آن، نمونه‌هایی از کاربردهای عملی این الگوریتم‌ها بررسی شد. انواع تکنیک‌های یادگیری ماشین یعنی یادگیری تحت‌نظارت، یادگیری بدون نظارت و یادگیری نیمه‌نظارت‌شده نیز معرفی شدند. انتظار می‌رود که این پژوهش درک و بینشی درست به خوانندگان بدهد تا تصمیمی آگاهانه در شناسایی گزینه‌های موجود در الگوریتم‌های یادگیری ماشین بگیرند و سپس الگوریتم یادگیری ماشین مناسب را در زمینه حل مسئله خاص انتخاب کنند.

## منابع

- [1] Zhou, Z. H. (2021). Machine learning. Springer Nature.
- [2] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- [3] Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9(1), 381-386.
- [4] Asadi, S. (2019). Evolutionary fuzzification of RIPPER for regression: Case study of stock prediction. *Neurocomputing*, 331, 121-137.
- [5] Fotouhi, S., Asadi, S., & Kattan, M. W. (2019). A comprehensive data level analysis for cancer diagnosis on imbalanced data. *Journal of biomedical informatics*, 90, 103089
- [6] Asadi, S., Jafari, S., & Shokrollahi, Z. (2019). Developing a course recommender by combining clustering and fuzzy association rules. *Journal of AI and Data mining*, 7(2), 249-262.
- [7] Fu, W., & Chien, C. F. (2019). UNISON data-driven intermittent demand forecast framework to empower supply chain resilience and an empirical study in electronics distribution. *Computers & Industrial Engineering*, 135, 940-949.
- [8] Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.
- [9] Haji, S. H., & Abdulazeez, A. M. (2021). Comparison of optimization techniques based on gradient descent algorithm: A review. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 18(4), 2715-2743.
- [10] Khirirat, S., Feyzmahdavian, H. R., & Johansson, M. (2017, December). Mini-batch gradient descent: Faster convergence under data sparsity. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)* (pp. 2880-2887). IEEE.
- [11] Weisberg, S. (2005). *Applied linear regression* (Vol. 528). John Wiley & Sons.
- [12] Alexopoulos, E. C. (2010). Introduction to multivariate regression analysis. *Hippokratia*, 14(Suppl 1), 23.
- [13] Stockemer, D., & Stockemer, D. (2019). Multivariate regression analysis. *Quantitative methods for the social sciences: a practical introduction with examples in SPSS and stata*, 163-174.
- [14] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.
- [15] Song, Y. Y., & Ying, L. U. (2015). Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry*, 27(2), 130.
- [16] Kotsiantis, S. B. (2013). Decision trees: a recent overview. *Artificial Intelligence Review*, 39, 261-283.
- [17] Suthaharan, S., & Suthaharan, S. (2016). Support vector machine. *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, 207-235.
- [18] Ch, R., & MAP, M. (1997). Bayesian learning. book: *Machine Learning*. McGraw-Hill Science/Engineering/Math, 154-200.
- [19] Webb, G. I., Keogh, E., & Miikkulainen, R. (2010). Naïve Bayes. *Encyclopedia of machine learning*, 15(1), 713-714.
- [20] Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, 4(2), 1883.
- [21] Kramer, O., & Kramer, O. (2013). K-nearest neighbors. *Dimensionality reduction with unsupervised nearest neighbors*, 13-23.

- [22] Yadav, J., & Sharma, M. (2013). A Review of K-mean Algorithm. *Int. J. Eng. Trends Technol*, 4(7), 2972-2976.
- [23] Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, 9(8), 1295.



## A brief review of machine learning algorithms

**Alireza Shabani**

Department of Industrial Engineering, Faculty of Engineering, College of Farabi, Universty of Tehran, Tehran, Iran

### **Abstract**

Machine learning, a prominent discipline within the realm of artificial intelligence, has garnered significant attention in the digital sphere as a pivotal component of intelligent solutions. The primary objective of this scholarly article is to present a concise overview of commonly utilized machine learning algorithms. By examining the prevalent and well-regarded machine learning algorithms, assessing their strengths and limitations from an application-oriented standpoint, and deriving insights from these evaluations, stakeholders can make informed decisions regarding the selection of the most suitable learning algorithm that aligns with the specific requirements of their applications. The examination conducted in this study takes into account factors such as accuracy, computational efficiency, stability, and adaptability to diverse data types and challenges. A comprehensive understanding of the advantages and disadvantages of these algorithms will facilitate judicious decision-making, thereby ensuring optimal performance and successful deployment in practical scenarios.

**Keywords:** Machine learning algorithms, Artificial intelligence, applied perspective, Decision making