مجله علوم تغذیه و صنایع غذایی ایران
سال نوزدهم، شماره ۲، تابستان ۱۴۰۳، صفحات ۱۰۵-۸۷

رویکردهای مبتنی بر هوش مصنوعی در ارزیابی الگوهای غذایی: یک مرور نظام مند

هانیه مالمیر[۱]، سمیه حسین پور نیازی[۲]، پروین میرمیران[۱]

۱- مرکز تحقیقات تغذیه در بیماری‌های غدد درون ریز، پژوهشکده علوم غدد درون ریز و متابولیسم، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران
۲- گروه تغذیه بالینی و رژیم درمانی، دانشکده علوم تغذیه و صنایع غذایی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران. پست الکترونیکی: s.hossainpour@yahoo.com

تاریخ دریافت: ۱۴۰۲/۵/۱۷ تاریخ پذیرش: ۱۴۰۲/۹/۴

چکیده

سابقه و هدف: توسعه هوش‌مصنوعی فرصت‌های جدیدی را برای تحقیق در زمینه علوم تغذیه فراهم کرده است. این مقاله با هدف مرور و بررسی جامع مطالعات مربوط به حوزه رژیم‌غذایی و الگوهای غذایی که از تکنیک‌های هوش‌مصنوعی و الگوریتم‌های یادگیری ماشین استفاده کرده‌اند، انجام شد.

مواد و روش‌ها: تمامی مطالعات به چاپ رسیده تا نوامبر ۲۰۲۳ با استفاده از پایگاه‌های اطلاعاتی Cochrane, PubMed و SCOPUS و موتور جستجوی Google Scholar و با کلیدواژه‌های مرتبط مورد جست‌وجو قرارگرفت.

یافته‌ها: بعد از بررسی کامل مقالات ۳۱ مقاله مرتبط انتخاب شدند که با هدف مطالعه حاضر هم‌خوانی داشتند. روش‌های مختلف یادگیری ماشین در پیش‌بینی الگوهای غذایی دقت متفاوتی دارند. به عنوان مثال شبکه عصبی هوشمند دقت بالاتری در پیش‌بینی پنجک‌های شاخص غذایی سالم دارد، در حالی که در مورد وعده‌های غذایی درخت تصمیم‌گیری دقت بالاتری دارد. کاربرد دیگر یادگیری ماشین، استخراج الگوهای غذایی و بررسی ارتباط آنها با بیماری‌های مختلف مانند چاقی، بیماری‌های قلبی، سکته مغزی، خطر مرگ ناشی از بیماری قلبی عروقی و سرطان می‌باشد. همچنین برخی از روش‌های یادگیری ماشین مانند درخت تصمیم قادر به ارائه مدل‌هایی برای پیش‌بینی میزان پیروی از رژیم غذایی مختلف مانند رژیم‌غذایی مدیترانه‌ای است.

نتیجه‌گیری: روش‌های مختلف هوش مصنوعی می‌توانند به شناخت بیشتر الگوهای غذایی مرتبط با بیماری‌های مزمن کمک کنند. مهمترین الگوریتم‌های مطرح در بررسی الگوهای غذایی درخت تصمیم گیری، جنگل تصادفی، میانگین کا، نزدیک ترین همسایه کا، روش‌های رگرسیونی، ماشین بردار پشتیبان و شبکه عصبی هوشمند هستند. این روش‌ها می‌توانند با دسته بندی و یافتن ارتباط پنهان بین گروه‌ها و مواد غذایی، به درک بهتر الگوهای غذایی مرتبط با بیماری‌های مزمن کمک کنند. برای درک بهتر این ارتباطات مطالعات بیشتری در این حوزه لازم است.

واژگان کلیدی: هوش مصنوعی، یادگیری ماشین، تغذیه، الگوی غذایی

پیام‌های اصلی

- هوش‌مصنوعی به شناخت بیشتر الگوهای غذایی مرتبط با بیماری‌های مزمن کمک می‌کند.
- الگوریتم‌های درخت تصمیم گیری، جنگل تصادفی، میانگین کا، نزدیک ترین همسایه کا، روش‌های رگرسیونی، ماشین بردار پشتیبان و شبکه عصبی هوشمند بیشترین کاربرد را در بررسی ارتباط الگوهای غذایی و بروز بیماری‌های مزمن دارند.
- بیماری‌های چاقی، بیماری‌های قلبی-عروقی، سکته مغزی، مرگ ناشی از بیماری قلبی عروقی و سرطان‌ها بیشتر مورد توجه بودند.

• مقدمه

داده‌ها در عصر امروزی یعنی عصر اطلاعات، عمده ترین دارایی نهادهای دخیل در سلامت هستند و موفقیت این سازمان‌ها در گروی جمع آوری، ذخیره و تحلیل آنها است. امروزه با توجه به ظهور نظام های اطلاعات یکپارچه و رشد فناوری اطلاعات، این مهم بیش از پیش نمایان شده است (۱). هوش مصنوعی شاخه‌ای از علوم کامپیوتر است که هدف آن

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

تقلید از فرآیندهای تفکر، توانایی‌های یادگیری و مدیریت دانش است. واژه "هوش مصنوعی (Artificial Intelligence)" اولین بار در سال ۱۹۵۵ توسط جان مک‌کارتی (John McCarthy)، محقق علوم کامپیوتر، در طرح پیشنهادی یک پروژه تحقیقاتی, پیشنهاد شد(۲). شاخه‌های گوناگونی از هوش مصنوعی در دانش‌های رایانه‌ای مورد استفاده قرار می‌گیرند، برخی از این شاخه‌ها شامل یادگیری ماشین (Machine Learning)، شبکه عصبی (Neural Networks)، بینایی ماشین (Machine Vision)، سامانه‌های متخصص (Expert System)، پردازش زبان طبیعی (Natural Language Processing)، الگوریتم ژنتیک (Genetic Algorithm)، و مفاهیم مرتبط با روباتیک (Robotic) می‌باشند (۳). در دهه‌های اخیر کاربرد هوش مصنوعی در پزشکی و پیراپزشکی گسترش‌یافته است. استفاده از هوش مصنوعی برای تشخیص‌های پزشکی و پیش‌بینی خطر بیماری‌های مزمن به سرعت در حال افزایش است. با استفاده از ابزارهای هوش مصنوعی معیارهای تشخیصی و درمانی مهمی برای بیماری‌های قلب و عروق، غدد درون ریز و سندروم متابولیک به دست آمده اند (۶-۴).

توسعه هوش مصنوعی فرصت‌های جدیدی را برای تحقیق در زمینه علوم تغذیه فراهم کرده است. بررسی و تفسیر رژیم‌های غذایی، طراحی نرم افزارهای کنترل کالری و دریافت مواد غذایی، طراحی نرم افزارهای تشخیص تصویر مواد غذایی، تنظیم و شخصی سازی رژیم‌های غذایی، پیش بینی خطر بیماری‌های مزمن و حتی ایجاد راهنماهای غذایی از جمله کاربردهای هوش مصنوعی در تحقیقات تغذیه می‌باشد (۱۵-۷). بیش از همه موارد ذکر شده، مطالعات قبلی از روش‌های مختلف هوش مصنوعی برای پیش بینی رژیم‌های غذایی مؤثر در ابتلا یا پیش گیری از بیماری‌های مزمن استفاده کرده اند. مطالعات مروری قبلی در حوزه هوش مصنوعی بیش از همه به بررسی کارآمدی این روش‌ها در ابزارهای سنجش دریافت غذایی، شخصی سازی رژیم غذایی و تعیین میکروبیوم روده پرداخته‌اند (۱۹-۱۶). الگوهای غذایی و مدل سازی در این زمینه کمتر مورد توجه قرار گرفته است. لذا این مقاله با هدف مرور و بررسی جامع مطالعات مربوط به حوزه رژیم غذایی و الگوهای غذایی که از تکنیک‌های هوش مصنوعی و الگوریتم‌های یادگیری ماشین استفاده کرده اند، انجام شد.

• مواد و روش‌ها

استراتژی جست و جوی مطالعات پیشین

به منظور بررسی مطالعات انجام شده در زمینه ی رژیم غذایی و هوش مصنوعی، تمامی مطالعات به چاپ رسیده تا نوامبر ۲۰۲۳ با استفاده از پایگاه‌های اطلاعاتی PubMed ، Cochrane و SCOPUS و موتور جستجوی Google Scholar مورد جست و جو قرار گرفت و در این جستجو از کلید واژه های"diet" یا "dietary pattern" یا "nutrition" یا " food group" برای بررسی رژیم غذایی، و از کلید واژه‌های " data mining" یا "artificial intelligence" یا "machine learning" یا "unsupervised learning" یا "supervised learning" یا "deep learning" یا "support vector machines" یا "reinforcement learning"یا "K-nearest neighbor " یا " decision tree" یا "random forest" یا " artificial neural network" برای یافتن مقالات در حوزه هوش مصنوعی و یادگیری ماشین استفاده شد. رفرانس‌های مطالعات قبلی نیز جهت از دست ندادن هیچ مطالعه ای در این زمینه مورد جستجو قرار گرفت. هیچ محدودیت زمانی در حین جستجو اعمال نشد. مراحل جستجوی مطالعات در شکل ۱ نشان داده شده است.

معیارهای ورود مطالعات

جستجوی مطالعات توسط دو محقق به صورت جداگانه انجام شد و تمامی مطالعاتی که به نحوی الگوی غذایی یا رژیم غذایی را با تکنیک‌های هوش مصنوعی و یادگیری ماشین بررسی کرده بودند، وارد مرور سیستماتیک شدند. مطالعات مورد بررسی باید از اطلاعات تغذیه ای افراد برای ساختن مدل استفاده کرده باشند و همچنین در مدل‌سازی از روش‌ها، تکنیک‌ها و الگوریتم‌های مطرح در حوزه ی هوش مصنوعی و یادگیری ماشین استفاده کرده باشند.

معیارهای خروج مطالعات

مطالعاتی که دارای یکی از شرایط زیر بودند وارد بررسی نشدند: ۱- مطالعات مروری و حیوانی و سلولی؛ ۲- مطالعاتی که الگوی غذایی یا دریافت یک ماده غذایی را بررسی کرده بودند ولی یافته‌ای منتشر نکرده و یا در مدل نهایی پیش بینی بیماری‌ها این معیار را وارد نکرده بودند؛ ۳- مطالعاتی که فقط به بررسی الگوهای مرتبط با بیماری‌ها پرداخته بودند و پارامترهای تغذیه ای و رژیم غذایی را بررسی نکرده بودند. ۴- مطالعاتی که به بررسی پرسشنامه‌های تغذیه ای و روش‌های رایج در ارزیابی تغذیه ای پرداخته بودند. ۵- مطالعاتی که به طراحی نرم‌افزارهای تغذیه‌ای پرداخته بودند. ۶- مطالعاتی که در زمینه شخصی سازی رژیم غذایی انجام شده‌اند. ۷- مطالعاتی که در زمینه میکروبیوم بودند و یا فاکتورهای ژنتیکی را بررسی کرده بودند.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

استخراج داده ها

از هر مطالعه اطلاعات مورد نیاز در مورد نام نویسنده ی اول، سال انتشار، طراحی مطالعه، کشور، سن افراد، جنسیت، حجم نمونه، طول مدت مطالعه، متغیر مواجهه و نحوه ی ارزیابی آن، متغیر پیامد و نحوه ی ارزیابی آن، روش‌های آنالیز (تکنیک‌ها و الگوریتم‌های یادگیری ماشین و نرم افزار مورد استفاده) و یافته‌ها استخراج گردید.

• یافته‌ها و بحث

در جست و جوی اولیه ۵۱۶۸ مقاله استخراج گردید. بعد از کنار گذاشتن مقالات تکراری، مطالعات مروری، حیوانی و آزمایشگاهی ۱۱۹۸ مقاله مورد بررسی قرار گرفتند. ۱۱۶۷ مقاله با هدف مطالعه حاضر مرتبط نبودند و لذا کنار گذاشته شدند. . با بررسی عنوان و چکیده ۳۱ مقاله جهت بررسی کامل و ورود به مطالعه انتخاب شدند (شکل ۱).

توصیف مطالعات

تمامی مطالعات در سال‌های ۲۰۰۸ تا ۲۰۲۳ به چاپ رسیده اند.از ۳۱ مطالعه بررسی شده، ۹ مطالعه دارای طراحی مقطعی، ۱۵ مطالعه کوهورت و ۷ مطالعه مورد-شاهدی بودند (جدول ۱). ۱۰ مطالعه در آسیا، ۱۰ مطالعه در آمریکا، ۶ مطالعه در اروپا، ۲ مطالعه در چندین کشور از قاره‌های متفاوت و ۳ مطالعه در آفریقا انجام شده بودند. در ۲۸ مطالعه گروه هدف بزرگسالان بودند و سه مطالعه بر روی کودک و نوجوان انجام شده بودند. حجم نمونه از ۲۱۲ نفر در مطالعه مورد شاهدی تا ۱۰۰ هزار نفر در مطالعه کوهورت متغیر بود.

روش‌ها و نرم افزارهای هوش مصنوعی

مبنای اصلی هوش مصنوعی یادگیری ماشین است. یادگیری ماشین در واقع مجموعه ای از الگوریتم‌ها و روش هایی است که برای طبقه بندی، خوشه بندی و پیش بینی استفاده می‌شود. روش‌های مورد استفاده در یادگیری ماشین را می‌توان به چند گروه اصلی تقسیم کرد: یادگیری نظارت شده (Supervised Learning)، یادگیری بدون نظارت (Unsupervised Learning) و یادگیری تقویتی (Reinforcement Learning) (۲۰). در جدول ۲ تکنیک‌ها و الگوریتم‌های مطرح در یادگیری ماشین که در حوزه علوم تغذیه کاربرد دارند توضیح داده شده است.

یادگیری نظارت شده مدلی را بر روی داده‌های ورودی و خروجی شناخته شده آموزش می‌دهد تا بتواند خروجی‌های آینده را پیش‌بینی کند و از تکنیک‌های طبقه بندی (Classification) و رگرسیون (Regression) استفاده می‌کند. از این روش هنگامی استفاده می‌شوند که دادگان ما دارای ویژگی (Tag) مشخص هستند. الگوریتم‌های (Algorithm) مورد استفاده شامل رگرسیون خطی (Linear Regression) و چند متغیره (Multivariate Regression)، رگرسیون لجستیک (Logistic Regression)، بیز ساده (Naïve Bayes)، درخت تصمیم گیری (Decision Tree)، نزدیکترین همسایه کا (K-nearest neighbor)، ماشین بردار پشتیبان (Support Vector Machine)، جنگل تصادفی (Random Forest)، و شبکه عصبی هوشمند (Artificial Neural Network) می‌باشد. یادگیری عمیق (Deep Learning) بخشی از یادگیری ماشین است که می‌تواند نظارت شده یا بدون نظارت باشد (۲۴-۲۱).

یادگیری بدون نظارت برای خوشه بندی (Clustering) و کاهش بعد دادگان (Dimensionality Reduction) استفاده می‌شوند. این روش اغلب برای استخراج ویژگی‌ها (Feature Extraction) مورد استفاده قرار می‌گیرند. الگوریتم‌های خوشه بندی میانگین کا (K-means clustering) و تحلیل مولفه اصلی (Principal Component Analysis) از جمله الگوریتم‌های مورد استفاده در این روش‌ها هستند (۲۶، ۲۵).

یادگیری تقویتی درواقع حوزه ای از یادگیری ماشین است که به ماشین ها اجازه می‌دهد تا برای رسیدن به اهداف خود با محیط پویای اطراف خود در تعامل باشند. با این کار، ماشین‌ها و عامل‌های نرم افزاری می‌توانند رفتار ایده آل را در یک زمینه خاص ارزیابی کنند(۲۷).

الگوهای غذایی

مطالعات در بررسی رژیم‌ها و الگوهای غذایی به چند شیوه عمل کرده اند: اغلب براساس شیوه‌ها و مدل‌های پیش بینی رژیم غذایی مرتبط با یک بیماری را گزارش کرده اند. برخی تنها دریافت‌های غذایی را بررسی کردند و برخی دیگر، دریافت‌های غذایی را به همراه سایر پارامترهای شیوه زندگی مانند میزان فعالیت بدنی، مصرف سیگار و نوشیدنی‌های الکلی، میزان خواب و سابقه بیماری‌های مزمن در مدل آورده اند. چند مطالعه نیز به بررسی عوامل مؤثر بر پیروی از رژیم‌های غذایی پرداخته بودند.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

جدول ۱. مشخصات مقالات در حوزه الگوی غذایی و هوش مصنوعی

نویسنده (سال انتشار)	طراحی مطالعه (نام مطالعه)	کشور	محدوده سنی	جنسیت	تعداد افراد	متغیر مواجهه	ارزیابی متغیر مواجهه	متغیر پیامد	ارزیابی متغیر پیامد	روش‌های آنالیز	یافته ها
هارتی[1] (۲۰۰۸)	مقطعی (North-South Ireland Food Consumption Survey 1997–1999)	ایرلند	۶۴-۱۸ سال	هر دو جنسیت	۱۳۷۹	الگوی غذایی	ثبت خوراک ۷ روزه	مقایسه دو سیستم کد گذاری براساس مواد غذایی و وعده غذایی	HEI Q1 vs. Q5	شبکه عصبی هوشمند درخت تصمیم با الگوریتم C5	ANN بر اساس سیستم کدگذاری مواد غذایی، دقت بالاتری از درخت تصمیم در پیش‌بینی پنجک‌های HEIداشت با این حال، درخت تصمیم دقت بالاتری از ANN بر اساس سیستم کدگذاری وعده غذایی داشت.
لازارو[2] (۲۰۱۲)	مقطعی (CYKIDS)	روسیه	۱۳-۹ سال	هر دو جنسیت	۱۱۴۰	الگوی غذایی	۱۵۴ FFQ آیتمی پرسشنامه کوتاه عادات غذایی	چاقی و اضافه وزن	نمودار وزن برای سن	درخت تصمیم با الگوریتم C5 برنامه WEKA	مصرف مکرر (۳ تا ۵ بار در هفته) از نوشیدنی ها، گوشت لذیذ، شیرینی ها، سرخ کردنی‌ها و غذاهای ناسالم خطر ابتلا به چاقی را تا ۷۵ درصد افزایش می‌دهد، در حالی که در کودکانی که الگوی غذایی مشابهی دارند، اما بیش از ۲ بار در هفته ماهی و غذای دریایی می‌خورند، ۳۳ درصد خطر چاقی کمتری دارند.
کاستورینی[3] (۲۰۱۳)	مورد شاهدی	روسیه	۸۶-۴۸	هر دو جنسیت	۵۰۰ بیمار ۵۰۰ شاهد	الگوی غذایی پیشین (رژیم مدیترانه ای) و الگوی غذایی پسین	FFQ	سندروم حاد کرونری سکته مغزی ایسکمیک	تشخیص پزشکی	رگرسیون لجستیک چندگانه، بیز ساده، درخت تصمیم ، شبکه عصبی هوشمند و ماشین‌های بردار پشتیبانی	هر دو رویکرد الگوی غذایی در اکثر الگوریتم‌های طبقه‌بندی به دقت طبقه‌بندی برابری دست یافتند. بنابراین، انتخاب به برنامه مورد نظر بستگی دارد
بیسبروک [4] (۲۰۱۵)	کورهوت cohort EPIC-NL	هلند	۷۰-۴۹	هر دو جنسیت	۳۶۶۴۴	الگوی غذایی	FFQ	بیماری قلبی و سکته مغزی	ثبت بیمارستانی	جنگل تصادفی میانگین کا PCA RRR	با استفاده از تمامی روش‌ها سه الگوی غذایی غربی،پیشگیرانه و سنتی شناسایی شدند. گنجاندن عوامل خطر در روش RRR منجر

[1] Hearty
[2] Lazarou
[3] kastorini
[4] Biesbroek

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

نویسنده (سال انتشار)	طراحی مطالعه (نام مطالعه)	کشور	محدوده سنی	جنسیت	تعداد افراد	متغیر مواجهه	ارزیابی متغیر مواجهه	متغیر پیامد	ارزیابی متغیر پیامد	روش‌های آنالیز	یافته ها
										نرم افزار SAS	به تغییرات کوچکی در الگوهای غذایی میشود که ارتباط بیشتری با بیماری قلبی و سکته مغزی دارند.
لی[5] (۲۰۱۷)	مورد-شاهدی بیمارستانی	چین	۲۷/۲±۴/۰ سال	زنان باردار	۱۱۹ بیمار ۲۳۹ شاهد	الگوی غذایی و سایرریسک فاکتورها	مصاحبه رو در رو	بیماری قلبی عروقی		شبکه عصبی هوشمند	۱۵ پیش بینی کننده بیماری قلبی عروقی شامل عوامل زمینه ای و تغذیه ای شناسایی شدند. دریافت میوه‌ها و سبزی ها، دریافت ماهی/ میگو/ گوشت/ تخم مرغ و دریافت شیر و شیر سویا موارد تغذیه ای مدل شبکه عصبی پس انتشار استاندارد بودند.
پانارتوس[6] (۲۰۱۸)	کوهورت ۱۰ ساله	روسیه	۴۵ سال	هر دو جنسیت	۲۰۲۰	الگوی غذایی (غذا و ماده مغذی)	FFQ نیمه کمی	ریسک ابتلا به بیماری‌های قلبی عروقی		رگرسیون خطی و آنالیز فاکتور الگوریتم نزدیکترین همسایه کا و جنگل تصادفی نرم افزار R	آنالیز فاکتور ۵ و ۳ عامل برای غذا و ماده مغذی نشان دادند که درصد قابل توجهی از واریانس را پوشش می‌داد. الگوریتم‌های نزدیک ترین همسایه کا و جگل تصادفی دقت بالاتری از رگرسیون و فاکتور آنالیز داشتند و هر دو یکسان عمل کردند.
استون[7] (۲۰۱۸)	مورد شاهدی	مکزیک				الگوی غذایی		چاقی دیابت نوع دو		مدل بیز ساده	افراد مبتلا به چاقی در همه گروه‌های غذایی به جز فست فودها دریافت بالاتری دارند. در افراد مبتلا به دیابت نوع دو و چاقی نیز دریافت‌های غذایی در همه گروه‌های غذایی به جز فست فودها و گوشت و ماهی بالاتر بود. افراد ابتلا به دیابت نوع دو دریافت کمتری از گروه‌های غذایی گوشت و ماهی، شکر و چربی، و فست فودها دارند.
آرکئوویلاس[8] (۲۰۲۰)	مقطعی	اسپانیا	بالای ۴۰ سال	هر دو جنسیت	۵۰۳	عوامل تغذیه ای و سایر ریسک فاکتورها	پرسشنامه و مصاحبه	پیروی از الگوی غذایی مدیترانه ای	مصاحبه	جنگل تصادفی ماشین بردار پشتیبان	۱۶ پیش بینی کننده شناسایی و وارد مدل شدند. عوامل تغذیه ای در مدل باقی نماندند

[5] Li
[6] Panaretos
[7] Easton
[8] Arceo-vilas

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

نویسنده (سال انتشار)	طراحی مطالعه (نام مطالعه)	کشور	محدوده سنی	جنسیت	تعداد افراد	متغیر مواجهه	ارزیابی متغیر مواجهه	متغیر پیامد	ارزیابی متغیر پیامد	روش‌های آنالیز	یافته ها
										همسایه نزدیک کا نرم افزار R	
بندار[9] (۲۰۲۰)	کوهورت	آمریکا		زنان باردار	۷۵۷۲	الگوی غذایی	FFQ	تولد زودهنگام جنین، وزن کم جنین، دیابت بارداری و پره اکلامپسی	ثبت	رگرسیون لجستیک چند متغیره، مدل‌های فرا یاددهنده نرم افزار R	مصرف بالای میوه و سبزی با خطر کمتر پره اکلامپسی، تولد زودهنگام و وزن کم جنین همراه بود.
فو [10] (۲۰۲۰)	کوهورت	چین	۴-۷ سال	۲۱۲۵ نوزاد-والد	هر دو جنسیت	متغیرهای مربوط به دوران جنینی و نوزادی	مصاحبه	چاقی / اضافه وزن	نمودار نمایه توده بدنی برای سن	یادگیری ماشین/ الگوریتم درخت تصمیم نرم افزار STATA	الگوی تغییرات نمودار BMI برای سن در سال اول تولد مهمترین پیش بینی کننده چاقی کودکان زمان شروع غذای کمکی در نوزادن نارس در این تغییرات بسیار اهمیت دارد (۶ ماه تمام و بالاتر)
هی [13] (۲۰۲۰)	کوهورت ۲۳ ساله	سازمان جهانی بهداشت/ چندین کشور	۱۸ سال	۱۰۰۰۰۰	هر دو جنسیت	دریافت پنیر، میوه و سبزی، الکل و سیگار، فشار خون	مصاحبه و پرسشنامه	مرگ و میر و بیماری‌های ناشی از مشکلات قلبی	ثبت	GPLVM[11] و PCA GTM[12] نرم افزار MATLAB	افزایش مصرف میوه و سبزی و غلات با خطر کمتر پرفشاری خون همراه است
زو [14] (۲۰۲۰)	مورد-شاهدی	آمریکا	۱۸-۲۵ سال	۱۰۶ بیمار ۱۰۶ شاهد	هر دو جنسیت	شاخص کیفیت رژیم غذایی (DQI)	پرسشنامه FFQ و تاریخچه غذایی	امتیاز خطرات بیماری‌های قلبی عروقی	مصاحبه و اندازه گیری	مدل رگرسیون و مدل تحلیل عاملی نرم افزار STATA	در این مدل دریافت سبزیجات اثر محافظتی و دریافت شیرینی ها، چربی اشباع و کربوهیدرات اثر افزاینده بر خطر بیماری‌های کاردیومتابولیک داشتند

[9] Bondar
[10] Fu
[11] Gaussian Process Latent Variable Model
[12] Generative topographic map
[13] He
[14] Xu

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

نویسنده (سال انتشار)	طراحی مطالعه (نام مطالعه)	کشور	محدوده سنی	جنسیت	تعداد افراد	متغیر مواجهه	ارزیابی متغیر مواجهه	متغیر پیامد	ارزیابی متغیر پیامد	روش‌های آنالیز	یافته ها
یو [15] (۲۰۲۰)	۱۸ مطالعه مورد-شاهدی و یک مطالعه کوهورت لانه گزینی		۵۹ سال	۳۱۵۵۱	هر دو جنسیت	الگوی غذایی	FFQ	سرطان مثانه	ثبت	درخت تصمیم الگوریتم C5 و C4.5 نرم افزار R	گوشت و فرآورده‌های گوشتی، چربی ها، روغن‌ها و فرآورده‌های آنها، غلات و محصولات غلات، سبزیجات و محصولات گیاهی، نوشیدنی‌ها (غیر شیر)) با خطر سرطان مثانه در مردان و زنان مرتبط هستند.
بوتو [16] (۲۰۲۱)	مقطعی	پرتقال	۱۵-۱۹ سال	۳۲۵	هر دو جنسیت	شاخص KIDMED	پرسشنامه	میزان پیروی از رژیم مدیرانه ای		درخت تصمیم SPSS	با در نظر گرفتن دسته بندی شاخص KIDMED عادات غذایی تنها عامل پیش بینی میزان پیروی از رژیم مدیترانه ای بودند. مصرف روزانه دومین سروینگ از میوه‌ها یک عامل مهم افتراق افرادی است که پایبندی بالایی به رژیم مدیترانه ای دارند. مصرف ۲ تا ۳ بار آجیل در هفته و مصرف سبزیجات پخته و خام در افراد پایبند بالاتر است. در افرادی که دو سروینگ میوه در روز مصرف نمی کردند، تنها کسانی که به طور همزمان از میوه تازه در صبحانه و سوپ سبزیجات در ناهار استفاده می‌کردند پایبندی بالایی به رژیم مدیترانه ای داشتند
سفلایی [17] (۲۰۲۱)	مورد شاهدی	ایران	۵۶ سال	۲۷۳ بیمار ۷۱۶ شاهد	هر دو جنسیت	الگوی مواد مغذی دریافتی	یادآمد ۲۴ ساعته	بیماری عروق قلبی	آنژیوگرافی	داده کاوی با C5 و QUEST درخت تصمیم C&R	از ۲۳ متغیر تنها پروتئین، منیزیوم، بیوتین، روی و کلسترول در مدل باقی ماندند
ژائو [18] (۲۰۲۱)	کوهورت مبتنی بر جمعیت (CARDIA)	آمریکا	۱۸-۳۰ سال	۵۱۱۵	هر دو جنسیت	الگوی غذایی	پرسشنامه FFQ	خطر ۱۰ ساله بیماری قلبی عروقی	ثبت	رگرسیون ماشین هسته بیزی نرم افزار R	در زنان مصرف گوشت قرمز و در مردان میوه‌ها بیشترین اثر را داشتند

[15] Yu
[16] Boto
[17] Soflaei
[18] Zhao

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

نویسنده (سال انتشار)	طراحی مطالعه (نام مطالعه)	کشور	محدوده سنی	جنسیت	تعداد افراد	متغیر مواجهه	ارزیابی متغیر مواجهه	متغیر پیامد	ارزیابی متغیر پیامد	روش‌های آنالیز	یافته ها
چویی[19] (۲۰۲۲)	دو مطالعه مقطعی فاز ۴ و ۷ (KNHANES)	کره جنوبی	۴۵ سال	۱۶۱۸۷	هر دو جنسیت	الگوی غذایی		الگوی غذایی	یادآمد ۲۴ ساعته	هوش مصنوعی	در سال ۲۰۰۷ الگوی غذایی شامل برنج سوپ سبزیجات مختلف و ماهی بود (سنتی) در سال ۲۰۱۶ الگوی غذایی غربی (گوشت و الکل) و الگوی غذایی شامل غذاهای سرخ شده، فراوری شده و نوشیدنی‌های شیرین شده بود
مورجنسترن[20] (۲۰۲۲)	کوهورت مبتنی بر جمعیت (CCHS 2.2)	کانادا	۵۰ سال	۱۲۱۳۰	هر دو جنسیت	الگوی ماده مغذی	یادآمد ۲۴ ساعته	بیماری قلبی عروقی	ثبت	جنگل تصادفی نرم افزار R	استفاده از مکمل، کافئین، الکل مهمترین متغیرهای پیش بینی بیماری قلبی هستند
موسوی[21] (۲۰۲۲)	مقطعی	ایران		۱۵۲۸	هر دو جنسیت			میزان پیروی از رژیم غذایی	پرسشنامه ۲۶ آیتمی	شبکه عصبی هوشمند و الگوریتم ژنتیکی نرم افزار MATLAB و SPSS	فاکتورهای پیش بینی مدل شامل مدت زمان ازدواج، دلیل مراجعه به کلینیک، وزن، نمایه توده بدنی، رضایت از وزن، زمان صرف ناهار و شام، و زمان خواب بودند
تاتولی[22] (۲۰۲۲)	مقطعی	ایتالیا	۷۵ سال	۱۳۹۹	هر دو جنسیت	الگوی غذایی	پرسشنامه FFQ	دیابت نوع دو	گزارشات پزشکی	PCA نرم افزار R	افراد دیابتی مصرف بالای محصولات لبنی، تخم مرغ، سبزیجات، میوه تازه، آجیل و روغن زیتون داشتند و مصرف شکر و شیرینی پایین تری داشتند افراد بدون دیابت مصرف بالای گوشت قرمز، گوشت فراوری شده، غذای آماده و الکل بیشتر و مصرف سبزیجات کمتر داشتند.
سیلوا[23] (۲۰۲۲)	کوهورت (ELSA- Brasil)	برزیل	۳۵ -۷۴ سال	۱۲۶۶۷	هر دو جنسیت			الگوی غذایی	پرسشنامه FFQ	الگوریتم خوشه بندی میانگین کا ماشین بردار پشتیبان	دو الگوی غذایی براساس خوشه بندی به دست آمدند و تمامی تکنیک‌های یادگیری ماشین دقت بالایی برای پیش بینی این دو الگو داشتند

[19] Choi
[20] Morgenstern
[21] Mousavi
[22] Tatoli
[23] Silva

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

نویسنده (سال انتشار)	طراحی مطالعه (نام مطالعه)	کشور	محدوده سنی	جنسیت	تعداد افراد	متغیر مواجهه	ارزیابی متغیر مواجهه	متغیر پیامد	ارزیابی متغیر پیامد	روش‌های آنالیز	یافته ها
										نزدیکترین همسایه کا درخت تصمیم جنگل تصادفی آنالیز بیز نرم افزار R	
یانگ[24] (۲۰۲۲)	کوهورت	چین	۶۵-۷۸	هر دو جنسیت	۳۹۹۵	گروه‌های غذایی	پرسشنامه FFQ	مرگ و میر بیماری قلبی، سرطان و هر علتی	ثبت	QGC نرم افزار R	غلات کامل، حبوبات، قارچ، سویا و محصولات سویا اصلی ترین گروه‌های غذایی شناسایی شده با خطر کمتر مرگ و میر ناشی از همه علل و CVD بودند، علاوه بر این، میوه‌ها و سویا و محصولات سویا مهم ترین گروه‌های غذایی در ارتباط با خطر کمتر سرطان بودند.
لیمکرت [25] (۲۰۲۲)	کوهورت	ایالات متحده آمریکا	۲۴-۵۰	هر دو جنسیت	۶۹۱	الگوی غذایی	FFQ	سندروم روده تحریک پذیر	شواهد بالینی / گزارش فردی	خوشه بندی میانگین کا نرم افزار پایتون	رژیم غذایی با مصرف زیاد میوه‌ها و سبزیجات و کاهش مصرف گوشت و شیرینی به طور همزمان، با ترجیح آب برای هیدراتاسیون، با کاهش خطر علائم فعال همراه بود.
کارمیش[26] (۲۰۲۳)	مورد-شاهدی Moroccan case–control study	مراکش	۱۸-۷۵	هر دو جنسیت	۱۴۸۳ بیمار ۱۴۸۳ شاهد	الگوی غذایی	پرسشنامه FFQ	سرطان کولورکتال	شواهد بالینی آزمایشات تشخیصی	خوشه بندی میانگین کا تجزیه و تحیل اجزا اصلی	دو الگوی غذایی پیشگیرانه و پر خطر با خوشه بندی میانگین کا شناسایی شدند که ارتباط قوی با شانس ابتلا به سرطان کولورکتال داشتند
طاهری[27] (۲۰۲۳)	مقطعی	ایران		زنان	۲۲۳	الگوی مواد مغذی	پرسشنامه FFQ	سندروم یش از قاعدگی		نزدیک ترین همسایه کا	ارائه مدلی برای پیش بینی ابتلا به سندروم قاعدگی دردناک که شامل قاعدگی دردناک (دیسمنوره)، ضخامت چین‌های پوستی، قاعدگی نامنظم، کل کالری دریافتی، کل فیبر مصرفی، اسیدهای چرب ترانس، مصرف

[24] Yang
[25] Limketkai
[26] Qarmiche
[27] Taheri

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

نویسنده (سال انتشار)	طراحی مطالعه (نام مطالعه)	کشور	محدوده سنی	جنسیت	تعداد افراد	متغیر مواجهه	ارزیابی متغیر مواجهه	متغیر پیامد	ارزیابی متغیر پیامد	روش‌های آنالیز	یافته ها
											سدیم، کل قند دریافتی، کل چربی دریافتی و بیوتین بود که دقت ۸۰ درصد داشت.
هوانگ[28] (۲۰۲۳)	کوهورت NHANES 2017–2020	آمریکا	>۱۸	هر دو جنس	۷۹۲۹	الگوی ریز مغذی‌های دریافتی	پرسشنامه FFQ	افسردگی	Patient Health Questionnaire (PHQ- 9)	ایکس جی بوستینگ	ارائه مدل ابتلا به افسردگی براساس فاکتورهای تغذیه ای و ریزمغذی‌های دریافتی
مارتین مولارس[29] (۲۰۲۳)	کوهورت NHANES 2003-2012	آمریکا	<۱۸	هر دو جنس	۹۷۰۶ بیمار ۲۱۶	سه مدل فاکتورهای سلامتی، تغذیه ای، ترکیبی	پرسشنامه FFQ	مرگ و میر ناشی از بیماری‌های قلبی عروقی	سیست مثبت	جنگل تصادفی SVM XG Boost	ارائه سه مدل براساس فاکتورهای سلامتی، گروهها و ریزمغذی‌های دریافتی و ترکیبی برای مرگ و میر که مدل ترکیبی کارایی و حساسیت و ویژگی بالاتری دارد
لامپیگنانو [30] (۲۰۲۳)	کوهورت	ایتالیا	۷۴	هر دو جنس	۱۹۲۹	الگوی غذایی	پرسشنامه FFQ	استئاتوز و فیبروز کبدی	Fatty Liver Index (FLI)و FIB-4 score	یادگیری ماشین	دو الگوی غذایی مشابه با الگوی غذایی غربی شناسایی شد که به مصرف بالای غذاهای فراوری شده، غلات تصفیه شده، الکل، نوشیدنی‌های شیرین، شکر و چربی‌های اشباع اشاره دارند.
کرتی[31] (۲۰۲۳)	مقطعی	هند	۱۰-۱۹	هر دو جنس	۱۲۳۱۸	الگوی ریزمغذی	پرسشنامه FFQ	چاقی و سندروم متابولیک		خوشخ بندی میانگین کا	۵ خوشه مرتبط با چاقی و سندروم متابولیک شناسایی شد. کمبود روی، فولات، آهن،ویتامین D ، B12 در خوشه‌ها بیشتر دیده شدند
ژائو[32] (۲۰۲۳)	کوهورت NHANES	آمریکا	<۱۸	هر دو جنس	۹۸۲۲	فاکتورهای تغذیه ای و شیوه زندگی	پرسشنامه	دیابت نوع دو		چنگل تصادفی XG Boost	دریافت مس رژیمی، میزان کادمکیوم و جیوه ادراری مهمترین پارامترهای پیش بینی دیابت هستند.
لی [33] (۲۰۲۳)	کوهورت NHANES	آمریکا	<۱۸	هر دو جنس	۵۷۲۵	فاکتورهای تغذیه ای و شیوه زندگی	پرسشنامه	سرطان پروستات		یادگیری ماشین	شاخص کیفیت رژیمی و فاکتورهایی مثل سن، کلسترول و RBC مهمترین عوامل در پیش بینی سرطان پروستات هستند.

[28] Huang
[29] Martin-Morales
[30] Lampignano
[31] Kirti
[32] Zhao
[33] Li

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

جدول ۲. تعاریف کلمات و اصطلاحات رایج در هوش مصنوعی

کلمات علمی	تعاریف
Artificial Intelligence	شاخه ای از علوم رایانه است که هدف اصلی آن تولید ماشین‌های هوشمندی است که توانایی انجام وظایفی که نیازمند به هوش انسانی (فرآیندهای تفکر, توانایی‌های یادگیری و مدیریت دانش) است را داشته باشد.
Machine Learning	از زیرمجموعه‌های هوش مصنوعی است که این امکان را برای سیستم‌ها فراهم می‌کند تا به صورت خودکار یادگیری و پیشرفت داشته باشند. تمرکز این سیستم بر توسعه برنامه‌های رایانه‌ای است تا بتوانند به داده‌ها دسترسی پیدا کرده و از آن‌ها برای یادگیری خود استفاده کنند
Supervised Learning	مدلی را بر روی داده های ورودی و خروجی شناخته شده آموزش می دهد تا بتواند خروجی های آینده را پیش بینی کند
Unsupervised Learning	اغلب برای استخراج ویژگی ها مورد استفاده قرار می گیرند
Reinforcement Learning	به ماشین ها اجازه می دهد تا برای رسیدن به اهداف خود با محیط پویای اطراف خود در تعامل باشند.
Algorithm	یک دنباله محدود از دستورالعمل‌های کاملاً تعریف شده‌است که معمولاً برای حل یک کلاس از مسائل خاص یا انجام یک محاسبات استفاده می‌شود
Classification	بر اساس داده‌های قبلی که دارای برچسب هستند، مدلی برای پیش بینی برچسب داده‌های جدید می‌سازد
Regression	پیش بینی و بیان تغییرات یک متغیر بر اساس اطلاعات متغیر دیگر
Linear Regression	رگرسیون خطی نوعی تابع پیش‌بینی‌کننده خطی است که در آن متغیر وابسته به صورت ترکیبی خطی از متغیرهای مستقل پیش‌بینی می‌شود، بدین معنی که هر کدام از متغیرهای مستقل در ضریبی که در فرایند تخمین برای آن متغیر به‌دست آمده ضرب می‌شود؛ جواب نهائی مجموع حاصل‌ضرب‌ها به علاوه یک مقدار ثابت خواهد بود که آن هم در فرایند تخمین به‌دست آمده‌است.
Multivariate Regression	رگرسیون چند متغیره تکنیکی است که برای اندازه گیری میزان ارتباط خطی متغیرهای مستقل مختلف و متغیرهای وابسته مختلف با یکدیگر استفاده می شود. این رابطه به دلیل همبستگی بین متغیرها خطی است.
Logistic Regression	تکنیکی آماری است برای نشان دادن تاثیر متغیرهای کمی یا کیفی بر متغیر وابسته دو وجهی (دو طبقه ای).
Decision Tree	درخت تصمیم گیری یک الگوریتم یادگیری ماشینی است که برای مسائل دسته بندی و پیش بینی استفاده می شود. درخت تصمیم گیری به صورت یک درخت با شاخه هایی که هر شاخه شامل یک سوال و با یک تصمیم است، ساخته می شود. هر شاخه از درخت یک مجموعه از داده ها را به دو یا چند شاخه دیگر تقسیم می کند به طوری که داده های هر شاخه بهترین پاسخ دسته بندی را در برابر سوال یا تصمیم مربوطه به دست می دهند.
Random Forest	ترکیبی از چندین درخت تصمیم گیری می‌باشد، که بر روی زمان آموزش و خروجی کلاس‌ها (کلاس‌بندی) یا برای پیش‌بینی‌های هر درخت به شکل مجزا، کار می‌کنند
Naïve Bayes	روشی برای دسته‌بندی پدیده‌ها بر پایه احتمالات است که با فرض استقلال متغیرهای تصادفی و براساس قضیه بیز ساخته می‌شوند. و براساس احتمال وقوع یا عدم وقوع یک پدیده انجام می شود.
K-nearest neighbor	یک متد آمار ناپارامتری است که برای طبقه‌بندی آماری و رگرسیون استفاده می شود. این الگوریتم برای طبقه بندی یک نقطه داده براساس نحوه طبقه بندی همسایگان آن استفاده می شود. مقدار K به تعداد نزدیک ترین نقاط داده همسایه برای انتخاب مشابه ترین ها اشاره دارد.
Generative topographic map	نقشه توپوگرافی تولیدی (GTM) یک روش یادگیری ماشینی است .در این مدل فرض می‌شود که داده‌ها با انتخاب احتمالی یک نقطه در فضای کم‌بعد قرار می گیرند و نقاط به فضای ورودی با ابعاد بالا (از طریق یک تابع صاف) و با اضافه کردن نویز در آن فضا به دست می‌آیند. GTM در سال ۱۹۹۶ در مقاله ای توسط کریستوفر بیشاپ، مارکوس اسونسن و کریستوفر کی آی ویلیامز معرفی شد.
Gaussian Process Latent Variable Model	مدل‌های متغیر پنهان فرآیند گاوسی (GPLVM) روش‌های کاهش ابعاد احتمالی هستند که از فرآیندهای گاوسی (GPs) برای یافتن جاسازی غیرخطی ابعاد پایین‌تر داده‌های با ابعاد بالا استفاده می‌کنند. آنها توسعه فرمول احتمالی PCA هستند. مدل به صورت احتمالی تعریف می‌شود و متغیرهای نهفته در حاشیه قرار می‌گیرند و پارامترها با حداکثر کردن احتمال به دست می‌آیند. مانند PCA هسته، آنها از یک تابع هسته برای تشکیل یک نقشه برداری غیر خطی (به شکل یک فرآیند گاوسی) استفاده می کنند. با این حال، در GPLVM نقشه برداری از فضای جاسازی شده (مخفی) به فضای داده (مانند شبکه های چگالی و GTM) است در حالی که در PCA هسته در جهت مخالف است.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

کلمات علمی	تعاریف
Support Vector Machine	مبنای کاری آن دسته‌بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌کنیم ابرصفحه ای را انتخاب کنیم که حاشیه اطمینان بیشتری داشته باشد. حل معادله پیدا کردن خط بهینه برای داده‌ها به وسیله روش‌های برنامه سازی غیرخطی که روش‌های شناخته شده‌ای در حل مسائل محدودیت‌دار هستند صورت می‌گیرد. قبل از تقسیمِ خطی برای اینکه ماشین بتواند داده‌های با پیچیدگی بالا را دسته‌بندی کند داده‌ها را به وسیله تابع phi به فضای با ابعاد خیلی بالاتر می‌بریم. برای اینکه بتوانیم مسئله ابعاد خیلی بالا را با استفاده از این روش‌ها حل کنیم از قضیه دوگانی لاگرانژ برای تبدیلِ مسئله مینیمم‌سازی مورد نظر به فرم دوگانی آن که در آن به جای تابع پیچیده phi که ما را به فضایی با ابعاد بالا می‌برد، تابعِ ساده‌تری به نامِ تابع هسته (کرنل) که ضرب برداری تابع phi است ظاهر می‌شود استفاده می‌کنیم. از توابع هسته مختلفی از جمله هسته‌های نمایی، چندجمله‌ای و سیگموید می‌توان استفاده نمود. یکی از معروفترین خودآموزها مربوط به است.
Artificial Neural Network	سیستم‌ها و روش‌های محاسباتی نوین برای یادگیری ماشینی، نمایش دانش و در انتها اعمال دانش به دست آمده در جهت بیش‌بینی پاسخ‌های خروجی از سامانه‌های پیچیده هستند. ایده اصلی این گونه شبکه‌ها تا حدودی الهام‌گرفته ازعملکرد یک سلول عصبی می باشد. یک شبکه عصبی مصنوعی، از سه لایه ورودی، خروجی و پردازش تشکیل می‌شود. هر لایه شامل گروهی از سلول‌های عصبی (نورون) است که عموماً با کلیه نورون‌های لایه‌های دیگر در ارتباط هستند، مگر این که کاربر ارتباط بین نورون‌ها را محدود کند؛ ولی نورون‌های هر لایه با سایر نورون‌های همان لایه، ارتباطی ندارند. نورون کوچک‌ترین واحد پردازشگر اطلاعات است که اساس عملکرد شبکه‌های عصبی را تشکیل می‌دهد.
Deep Learning	بخشی از یادگیری ماشین مبتنی بر شبکه عصبی هوشمند با یادگیری بازنمایی است.
Clustering	الگوریتم *خوشه‌بندی* اطلاعاتی را که ویژگی‌های نزدیک به هم و مشابه دارند را در دسته‌های جداگانه که به آن *خوشه* گفته می‌شود قرار می‌دهد.
Dimensionality Reduction	انتقال داده از فضای با بعد بیش تر به فضایی با بعد پایین تر به شکل تحت نظر است، به گونه ای که داده در فضای با بعد کمتر بتواند به طور معنی داری داده اصلی را نمایندگی کند و ویژگی های آن را در خود داشته باشد.
K-means clustering	خوشه‌بندی با هدف تجزیه تعداد مشاهدات به کا خوشه است که در آن هر یک از مشاهدات متعلق به خوشهای با نزدیکترین میانگین آن است، این میانگین به عنوان پیش‌نمونه استفاده می‌شود.
Principal Component Analysis	یک تکنیک آماری برای کاهش ابعاد یک مجموعه داده است و ویژگی‌هایی را که ارزش بیشتری فراهم می‌کنند برای ما استخراج می‌کند
C5	یکی از الگوریتم های مطرح در روش درخت تصمیم گیری که ضعف های الگوریتم C4.5 را پوشش داده است و به طور خودکار ویژگی های مهم را شناسایی می کند.
C4.5	یکی از الگوریتم های مطرح در روش درخت تصمیم گیری که قوانین خاصی در تقسیم بندی و هرس دادگان دارد.
QUEST	یکی از الگوریتم های مطرح در روش درخت تصمیم گیری است و برای یافتن یک ماتریس چرخشی بین دو سیستم مختصات از دو مجموعه مشاهدات طراحی شده است.
C&R	یکی از الگوریتم های مطرح در روش درخت تصمیم گیری است.
Super learner models	الگوریتمی است که از اعتبار سنجی متقابل برای تخمین عملکرد جندین مدل استفاده می کند. سپس با استفاده از عملکرد داده های آزمون میانگین وزنی بهینه از مدلها ایجاد می کند.
Reduced rank regression	رگرسیون رتبه کاهش یافته یک مدل رگرسیون خطی چند متغیره است که با هدف کاهش ابعاد توسعه یافته است.
Quantile G computation	محاسبه پنجکی جی تخمین هایی را از اثر افزایش همه مواجهه ها به اندازه یک پنجک به طور همزمان به دست می آورد که یک اثر ترکیبی است .
XG Boosting	یک الگوریتم نظارت شده که نسخه ای کارآمد و بهینه از الگوریتم تقویت گرادیان است. با موازی کردن کل فرایند تقویت زمان تمرین را بسیار بهبود می بخشد.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]



شکل ۱. فلوچارت مراحل ورود مطالعات به مرور سیستماتیک

دریافت‌های غذایی اغلب با پرسشنامه بسامد خوراک و یادآمد غذایی ۲۴ ساعته بررسی شده اند. الگوهای غذایی اغلب بدین ترتیب به دست آمده اند: در ابتدا آیتم‌های زیادی وارد مدل می‌شوند و ارتباط هر گروه یا آیتم غذایی با بیماری مورد نظر بررسی شده و در نهایت مواردی که ارتباط معناداری داشتند و اثرگذار بودند در مدل باقی می‌مانند. در ادامه یافته‌های مطالعاتی که پارامترهای رژیم غذایی را وارد مدل کرده بودند به تفضیل بیان خواهد شد:

در مطالعه Hearty و همکاران (۲۰۰۸) کیفیت رژیم غذایی که با دو روش کدگذاری براساس وعده غذایی و ماده غذایی به دست آمده بودند با هم مقایسه شدند و بدین منظور از دو روش شبکه عصبی هوشمند و درخت تصمیم گیری با الگوریتم C5 استفاده شده بود. یافته‌ها نشان دادند که در سیستم کدگذاری براساس ماده غذایی شبکه عصبی هوشمند دقت بالاتری در پیش بینی پنجک‌های شاخص غذایی سالم دارد. در حالی که در مورد وعده‌های غذایی درخت تصمیم گری دقت بالاتری دارد (۲۸). در مطالعه Lazarou و همکاران (۲۰۱۲) با استفاده از پرسشنامه‌های عادات غذایی و بسامد خوراک ۱۵۴ آیتمی، ۵ گروه غذایی انتخاب شدند و به همراه جنسیت به عنوان فاکتورهای مؤثر وارد مدل شدند. با کمک روش درخت تصمیم گیری و الگوریتم C5 با استفاده از نرم افزارWEKA ، الگوی غذایی ابتلا به چاقی و اضافه وزن شناسایی شد. بدین منظور فاکتورهای مد نظر چندین بار در سیستم اجرا (Run) شدند و قوانین (Rule) متعددی به دست آمد. در مرحله بعدی قوانین مشابه و بی معنی حذف شدند و در نهایت قوانین براساس فراوانی نسبت کودکان دارای اضافه وزن و چاقی مرتب شدند و الگوها به دست آمدند. یافته‌ها نشان دادند که رژیم غذایی حاوی نوشیدنی‌های شیرین، گوشت‌های فراوری شده، غذاهای سرخ کرده و ناسالم خطر چاقی را تا ۷۵ درصد بیشتر بود. در حالی که با رژیم غذایی مشابه با مصرف بیشتر از ۲ بار در هفته ماهی و غذای دریایی خطر چاقی ۳۳ درصد کمتر بود (۲۹). در مطالعه kastorini و همکاران (۲۰۱۳) الگوهای غذایی پیشین و پسین در ارتباط با شانس بروز سندروم حاد کرونری و سکته مغزی مقایسه شدند. الگوی غذایی مدیترانه ای با الگوهای غذایی به دست آمده از روش رگرسیون لجستیک چندگانه، بیز ساده، درخت تصمیم گیری، شبکه عصبی هوشمند و ماشین‌های بردار پشتیبانی دقت تقریبا برابری داشتند. به نظر می‌رسد انتخاب روش تعیین الگوی غذایی وابسته به هدف و دیتای در دسترس باشد (۳۰). در مطالعه Biesbroek و همکاران (۲۰۱۵) با استفاده از روش‌های جنگل تصادفی، میانگین کا، تجزیه و

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]
[DOI: 10.61186/nsft.19.2.87]

تحلیل مولفه اصلی و RRR و با کمک نرم افزار SAS الگوهای غذایی مرتبط با بیماری‌های قلبی و سکته مغزی بررس شدند. یافته‌ها نشان داد که سه الگوی غذایی رایج شامل الگوی غربی، پیشگیرانه و سنتی بیان کننده الگوهای غذایی رایج در بیماران مبتلا به بیماری‌های قلبی و سکته مغزی هستند و اضافه کردن عوامل خطر به RRR منجر به افزایش دقت شناسایی الگوهای غذایی می‌شود (۳۱).

در مطالعه Li و همکاران (۲۰۱۷) با استفاده از روش شبکه عصبی هوشمند و نرم افزار MATLAB، مدل پیش بینی خطرات بیماری قلبی مادرزادی در زنان باردار به دست آمد. پانزده پیش بینی کننده شامل عوامل زمینه‌ای و دریافت گروه‌های میوه و سبزی، گروه ماهی و میگو و گوشت و تخم مرغ و گروه شیر و شیرسویا در یک شبکه عصبی پس انتشار (Neural Network Backpropagation) با کاهش خطر ابتلا به بیماری قلبی مرتبط بودند (۹). در مطالعه Panaretos و همکاران (۲۰۱۸) ریسک ابتلا به بیماری‌های قلبی عروقی با سنجش الگوی غذایی و الگوی مواد مغذی دریافتی با کمک روش‌های نزدیکترین همسایه کا و جنگل تصادفی با نرم افزار R در یک مطالعه کوهورت ده ساله بررسی شد. یافته‌ها نشان دادند که هر دو روش دقت بالایی در انتساب افراد به گروه‌های کم خطر و پر خطر داشتند و یکسان عمل کردند (۳۲). در مطالعه Easton و همکاران (۲۰۱۹) با استفاده از روش آنالیز بیز مدل پیش بینی ابتلا به چاقی و دیابت نوع دو مورد بررسی قرار گرفت. بدین منظور دریافت‌های غذایی افراد در ۷ گروه غذایی تقسیم بندی شدند و به همراه عادات غذایی و اندازه‌های آنتروپومتریک افراد وارد مدل شدند. یافته‌ها نشان دادند که افراد مبتلا به چاقی در همه گروه‌های غذایی به جز فست فودها دریافت بالاتری دارند. در افراد مبتلا به دیابت نوع دو و چاقی نیز دریافت‌های غذایی در همه گروه‌های غذایی به جز فست فودها و گوشت و ماهی بالاتر بود. افراد ابتلا به دیابت نوع دو دریافت کمتری از گروه‌های غذایی گوشت و ماهی، شکر و چربی، و فست فودها دارند. مدل پیش بینی دیابت نوع دو با دقت بالاتری عمل می‌کرد درحالی که در مورد چاقی دقت مدل بالا نبود (۳۳). در مطالعه Arceo-vilas و همکاران (۲۰۲۰) میزان پیروی از رژیم غذایی مدیرانه‌ای با استفاده از چندین روش یادگیری ماشین (جنگل تصادفی، ماشین بردار پشتیبان و نزدیکترین همسایه کا) با نرم افزار R محاسبه شد. شانزده پیش بینی کننده برای میزان پیروی از رژیم غذایی مدیترانه ای در مدل باقی ماند که شامل عوامل آنتروپومتریک و سوابق پزشکی می‌شدند. عوامل تغذیه ای در مدل باقی نماندند (۳۴). در مطالعه Bondar و همکاران (۲۰۲۰) مدل پیش بینی کننده ابتلا به پره اکلامپسی و دیابت بارداری و زایمان زودرس و وزن کم جنین هنگام تولد با استفاده از روش رگرسیون لجستیک چند متغیره و مدل‌های فرا یاددهنده (Super Learner) و نرم افزار R به دست آمد. در این مدل‌ها دریافت بالای میوه و سبزی را دو عامل مؤثر در مدل پیش بینی کننده پره اکلامپسی و زایمان زودرس و وزن کم هنگام تولد معرفی می‌کنند و این عوامل نقش محافظتی دارند. در حالی که دریافت بالای سبزیجات با خطر بالاتری برای دیابت بارداری مرتبط بود (۷).

Fu و همکاران (۲۰۲۰) با تکنیک‌های یادگیری ماشین و نرم افزار STATA مدلی برای ابتلا به چاقی و اضافه وزن در کودکان طراحی کردند. در این مدل تغییرات نمایه توده بدنی در سال اول تولد و زمان شروع غذای کمکی مؤثرترین عوامل در پیش بینی چاقی کودکان هستند (۳۵). مطالعه He و همکاران (۲۰۲۰) با استفاده از مدل‌های مختلف نزدیک ترین همسایه کا و تجزیه و تحلیل اجزا اصلی در نرم افزار MATLAB مدلی برای پیش بینی میزان فشار خون سیستولیک و خطر مرگ ناشی از بیماری قلبی عروقی طراحی کردند. دو پارامتر اصلی در این مدل دریافت میوه‌ها و سبزی‌ها بودند (۳۶). مطالعه Xu و همکاران (۲۰۲۰) با کمک مدل‌های رگرسیونی و تحلیل عاملی با نرم افزار STATA مدل پیش بینی عوامل خطر بیماری‌های قلبی عروقی را طراحی کرند. در این مدل دریافت سبزیجات اثر محافظتی و دریافت شیرینی ها، چربی اشباع و کربوهیدرات اثر افزاینده بر خطر بیماری‌های کاردیومتابولیک داشتند (۳۶). مطالعه Yu و همکاران (۲۰۲۰) مدل پیش بینی کننده سرطان مثانه را با استفاده از دریافت‌های غذایی از گروه گوشت، چربی‌ها و روغن‌ها، غلات، سبزیجات و نوشیدنی‌های شیرین شده طراحی کردند. از الگوریتم‌های C5 و C4.5 و نرم افزار R در این مطالعه استفاده شد (۳۷). در مطالعه Boto و همکارن (۲۰۲۱) مدل پیش بینی میزان پیروی از رژیم غذایی مدیترانه ای در نوجوانان با کمک روش درخت تصمیم طراحی شد. یافته‌های مطالعه نشان داد که با در نظر گرفتن دسته بندی شاخص KIDMED عادات غذایی تنها عامل پیش بینی میزان پیروی از رژیم مدیترانه ای بودند. مصرف روزانه دو نوبت از میوه‌ها یک عامل مهم افتراق افرادی است که پایبندی بالایی به رژیم مدیترانه‌ای دارند. مصرف ۲ تا ۳ بار آجیل در هفته و مصرف سبزیجات پخته و خام در افراد پایبند بالاتر است. در افرادی که دو نوبت میوه در روز مصرف نمی‌کردند، تنها کسانی که به طور همزمان از میوه تازه در صبحانه و سوپ سبزیجات در ناهار استفاده می‌کردند پایبندی بالایی به رژیم مدیترانه‌ای داشتند (۳۸). در مطالعه Soflaei و همکاران (۲۰۲۱) با استفاده از روش‌های یادگیری ماشین و الگوریتم‌های QUEST، C5 و

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

C&R مدلی طراحی کردند که خطر ابتلا به بیماری قلبی عروقی را براساس دریافت پروتئین، منیزیوم، بیوتین، روی و کلسترول پیش بینی می‌کند(۳۹). Zhao و همکاران (۲۰۲۱) عوامل خطر ۱۰ ساله بیماری قلبی عروقی را با استفاده از رگرسیون ماشین هسته بیزی (Bayesian kernel machine regression) و نرم‌افزار R بررسی کردند و در مدل نهایی دریافت گوشت قرمز در زنان و دریافت کم میوه‌ها در مردان بیشترین اثرات را داشتند (۴۰). در مطالعه Choi و همکاران (۲۰۲۲) با استفاده از روش‌های هوش مصنوعی الگوهای غذایی بزرگسالان در دو فاز مطالعه KNHANES شناسایی شدند. در سال ۲۰۰۷ الگوی غذای مرسوم الگوی سنتی شامل دریافت برنج، سوپ، سبزیجات مختلف و ماهی بود. در سال ۲۰۱۶ الگوی غذایی غربی (مبتنی بر گوشت و الکل) و الگوی غذایی شامل مواد غذایی سرخ شده و فراوری شده و نوشیدنی‌های شیرین شده مشاهده شد (۴۱). در مطالعه Morgenstern و همکارن (۲۰۲۲) با روش جنگل تصادفی مدلی برای پیش بینی بیماری‌های قلبی عروقی طراحی کردند. دریافت مکمل، کافئین و الکل را مهمترین پارامترهای این مدل بودند. برای طراحی مدل از نرم افزار R استفاده شد(۱۲). در مطالعه موسوی و همکاران با استفاده از روش‌های شبکه عصبی هوشمند و الگوریتم ژنتیکی و نرم افزار MATLAB مدل پیش بینی کننده پیروی از رژیم غذایی به دست آمد. در این مدل فاکتورهای پیش بینی کننده شامل مدت زمان ازدواج، دلیل مراجعه به کلینیک، وزن، نمایه توده بدنی، رضایت از وزن، زمان صرف ناهار و شام، و زمان خواب بودند (۴۲). در مطالعه Tatoli و همکاران (۲۰۲۲) رژیم غذایی افراد دیابتی در مدل‌های یادگیری بدون نظارت منعکس کننده دریافت بالاتر محصولات لبنی، تخم مرغ، سبزیجات، میو ها، آجیل، روغن زیتون و دریافت کمتر شیرینی‌ها و مواد قندی بود. افراد بدون دیابت دریافت بالای گوشت قرمز، گوشت فراوری شده، غذای آماده و الکل بیشتر و مصرف سبزیجات کمتر داشتند. به نظر می‌رسد افراد دیابتی به دلیل شرایط ویژه ای که دارند بیشتر از توصیه‌های تغذیه‌ای پیروی می‌کنند. از نرم افزار R در این بررسی استفاده شده است (۴۳). در مطالعه Silva و همکاران (۲۰۲۲) با کمک الگوریتم‌های خوشه بندی دو الگوی غذایی در بزرگسالان برزیلی یافت شد و با کمک چندین روش یادگیری ماشین با دقت بالایی این دو الگو در افراد پیش بینی شدند(۱۴). در مطالعه Yang و همکاران (۲۰۲۲) با استفاده از روش پیشرفته محاسبه پنجک جی در حدود ۴۰۰۰ نفر از سالمندان چینی گروه‌های غذایی مرتبط با مرگ و میر به علت بیماری قلبی و سرطان و هر علتی شناسایی شدند. غلات کامل، حبوبات، قارچ، سویا و محصولات سویا اصلی ترین گروه‌های غذایی شناسایی شده با خطر کمتر مرگ و میر ناشی از همه علل و CVD بودند، علاوه بر این، میوه‌ها و سویا و محصولات سویا مهم ترین گروه‌های غذایی در ارتباط با خطر کمتر سرطان بودند (۴۴). Limketkai و همکاران (۲۰۲۲) ۵ الگوی غذایی اصلی را که توسط بیماران در حین بهبودی از سندروم روده تحریک‌پذیر اجرا می‌شد، شناسایی کردند. دو الگوی غذایی شبیه رژیم غذایی سنتی غربی بود که با مصرف زیاد چربی‌های اشباع شده، گوشت قرمز، قندهای افزوده شده و مصرف کم غذاهای غنی از فیبر مانند میوه‌ها و سبزیجات شناسایی شد، یک الگوی غذایی شبیه به یک رژیم غذایی متعادل بود، و دو الگوی غذایی مشابه رژیم غذایی گیاهی بودند که مصرف زیاد میوه ها، سبزیجات، حبوبات، آجیل و دانه‌ها و غلات کامل در حالی که پروتئین‌های حیوانی را محدود می‌کردند، داشتند. در این مطالعه با استفاده از روش خوشه بندی میانگین کت به این نتیجه رسیدند که رژیم غذایی با مصرف زیاد میوه‌ها و سبزیجات و کاهش مصرف گوشت و شیرینی به طور همزمان، با ترجیح مصرف آب برای هیدراتاسیون، با کاهش خطر علائم فعال کولیت اولسراتیو و کرون همراه است (۴۵). مطالعه Qarmiche و همکاران (۲۰۲۳) با کمک الگوریتم‌های خوشه‌بندی میانگین کا و تجزیه و تحلیل اجزا اصلی دو الگوی غذایی پیشگیرانه و پر خطر در ارتباط با سرطان کولورکتال شناسایی شد. دریافت متوسط از همه گروه‌های غذایی با میوه و سبزی بیشتر شانس کمتری برای ابتلا به سرطان کولورکتال و دریافت بالای روغن‌های گیاهی، کیک، شکلات، گوشت قرمز، پنیر، شکر و کره شانس بیشتری برای ابتلا داشتند (۴۶). در مطالعه طاهری و همکاران (۲۰۲۳) با کمک الگوریتم نزدیک ترین همسایه کا مدل پیش بینی ابتلا به سندروم پیش از قاعدگی ارائه کردند. این مدل که شامل فاکتورهای قاعدگی دردناک (دیسمنوره)، ضخامت چین‌های پوستی، قاعدگی نامنظم، کل کالری دریافتی، کل فیبر مصرفی، اسیدهای چرب ترانس، مصرف سدیم، کل قند دریافتی، کل چربی دریافتی و بیوتین بود دقت ۸۰ درصدی در پیش بینی داشت(۴۷). در مطالعه Huang و همکاران (۲۰۲۳) با استفاده از روش ایکس جی بوستینگ (XG Boosting) که یک روش نظارت شده در یادگیری ماشین است پارامترها و فاکتورهای تغذیه ای مرتبط با افسردگی شناسایی شدند. از بین تمامی موارد ۲۴ مورد از ریزمغذی‌ها و پارامترهای تغذیه‌ای مرتبط در مدل باقی ماندند. بیشترین میزان تأثیر از دریافت پتاسیم (۶/۸ درصد)، دریافت ویتامین E (۵/۷ درصد)، تعداد آیتم‌های غذایی و نوشیدنی‌های دریافت‌شده (۵/۷ درصد)، و دریافت ویتامین K (۵/۶ درصد) گزارش شد. این مدل با AUC ۰/۶۰۳و حساسیت ۰/۹۴۳ و ویژگی ۰/۱۶۳ مدل مناسبی

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

برای شناسایی افسردگی می‌باشد (۴۸). در مطالعه Martin-Morales (۲۰۲۳) در سه مدل با پارامترهای سلامتی، فاکتورهای تغذیه ای و ترکیب این عوامل با کمک روش‌های یادگیری ماشین مانند الگوریتم جنگل تصادفی، ایکس جی بوستینگ و غیره عوامل مؤثر بر مرگ و میر ناشی از بیماری‌های قلبی عروقی شناسایی شدند. در مدل پارامترهای سلامتی اندازه‌های تن سنجی، فشار خون و نشانگرهای خونی وارد مدل شدند. در مدل تغذیه‌ای دریافت ریزمغذی‌ها و گروه‌های غذایی وارد مدل شدند. یافته‌ها نشان دادند که توجه به الگوی غذایی و پارامترهای تغذیه ای به اندازه فاکتورهای تن‌سنجی، نشانگرهای خونی و فشارخون می‌توانند در تشخیص بیماری‌های قلبی عروقی کمک کننده باشند و ترکیب فاکتورهای سلامتی با اطلاعات تغذیه‌ای منجر به ایجاد مدل با کارایی بهتر می‌شود (۴۹). در مطالعه Lampignano و همکارن (۲۰۲۳) با استفاده از نرم افزار R و الگوریتم‌های یادگیری ماشین مدل مرتبط با بروز استئاتوز و فیبروز کبدی در سالمندان ایتالیایی بررسی شد. در هر دو حالت الگوی غذایی مشابه با الگوی غذایی غربی شناسایی شد که به مصرف بالای غذاهای فراوری شده، غلات تصفیه شده، الکل، نوشیدنی‌های شیرین، شکر و چربی‌های اشباع اشاره دارند (۵۰). Kirti و همکارن (۲۰۲۳) در مطالعه مقطعی در بین نوجوانان ۱۰ تا ۱۹ ساله هندی الگوهای غذایی مرتبط با چاقی و سندروم متابولیک را شناسایی کردند. در این مطالعه از الگوریتم‌های خوشه بندی و میانگین کا استفاده شد. پنج خوشه با توجه به شرایط بالینی افراد مورد مطالعه شناسایی شدند و پارامترهای تغذیه ای مرتبط برای چاقی، سندروم متابولیک، دیابت، پرفشاری خون و دیس لیپیدمی تعریف شدند. کمبود روی، فولات، آهن، ویتامین D، B12 در خوشه‌ها بیشتر دیده شدند (۵۱). در مطالعه Zhao و همکاران (۲۰۲۳) با کمک روش‌های جنگل تصادفی و ایکس جی بوستینگ مدل‌های سنجش مواجهه با فلزات سنگین و خطر ابتلا به دیابت نوع دو تبیین شد. دریافت مس رژیمی، میزان کادمیوم ادراری و میزان جیوه ادراری مهمترین پارامترهای پیش بینی کننده دیابت هستند (۵۲). Li و همکاران (۲۰۲۳) با کمک شاخص تغذیه ای و فاکتورهای مطرح در سرطان پروستات مدل پیش بینی ابتلا به سرطان پروستات تبیین شد. بدین منظور از سه الگوریتم یادگیری ماشین استفاده شد (۵۳).

روش‌های مختلف هوش مصنوعی می‌توانند به شناخت بیشتر الگوهای غذایی مرتبط با بیماری‌های مزمن کمک کنند. مهمترین الگوریتم‌های مطرح در بررسی الگوهای غذایی درخت تصمیم گیری، جنگل تصادفی، میانگین کا، نزدیک ترین همسایه کا، روش‌های رگرسیونی، ماشین بردار پشتیبان و شبکه عصبی هوشمند هستند. این روش‌ها می‌توانند با دسته بندی و یافتن ارتباطات پنهان بین گروه‌ها و مواد غذایی، به درک بهتر الگوهای غذایی مرتبط با بیماری‌های مزمن کمک کنند. با توجه به محدودیت مطالعات انجام شده در زمینه ارتباط بین روش‌های هوش مصنوعی با عوامل تغذیه‌ای، مطالعات بیشتری جهت تعیین روش‌های یادگیری ماشین دقیق در ارائه مدل‌ها برای پیش‌بینی میزان پیروی از رژیم غذایی و نیز استخراج الگوهای غذایی و ارتباط آن با بیماری‌های مزمن توصیه می‌شود.

تضاد منافع

نویسندگان اعلام می‌دارند که هیچ گونه تضاد منافعی در پژوهش حاضر وجود ندارد.

References

1. Bali RK. Clinical Knowledge Management: Opportunities and Challenges. . Hershey: Idea Group Inc (IGI); 2005.
2. SIYAH B. The document in which the words "Artificial Intelligence" were written for the first time 2021. Available from: https://www.kaggle.com/general/246669.
3. libraries N. Artificial intelligence: Wikipedia; 2023. Available from: https://en.wikipedia.org/wiki/Artificial_intelligence.
4. Gholamhosseini L, Damroodi M. Evaluation of Data Mining Applications in the Health System. ajaums-jps. 2015;10(1):39-48.
5. Jahanbakhsh HMAHFAM. Application of Data Mining in Health. Health Information Management. 2011;9(2):297-304.
6. Pashaei; E, Fard MK, Chale A. An overview of data mining methods in the field of health. The second international conference and the third national conference on the application of new technologies in engineering sciences2015.
7. Bodnar LM, Cartus AR, Kirkpatrick SI, Himes KP, Kennedy EH, Simhan HN, et al. Machine learning as a strategy to account for dietary synergy: an illustration based on dietary intake and adverse pregnancy outcomes. Am J Clin Nutr. 2020;111(6):1235-43.
8. Bodnar LM, Kirkpatrick SI, Naimi AI. Machine learning can improve the development of evidence-based dietary guidelines. Public Health Nutr. 2022;25(9):2566-9.
9. Li H, Luo M, Zheng J, Luo J, Zeng R, Feng N, et al. An artificial neural network prediction model of congenital heart disease based on risk factors: A hospital-based case-control study. Medicine (Baltimore). 2017;96(6):e6090.
10. Mezgec S, Koroušić Seljak B. Deep Neural Networks for Image-Based Dietary Assessment. J Vis Exp. 2021(169).
11. Monlezun DJ, Carr C, Niu T, Nordio F, DeValle N, Sarris L, et al. Meta-analysis and machine learning-augmented mixed effects cohort analysis of improved diets among 5847 medical trainees, providers and patients. Public Health Nutr. 2022;25(2):281-9.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

12. Morgenstern JD, Rosella LC, Costa AP, Anderson LN. Development of machine learning prediction models to explore nutrients predictive of cardiovascular disease using Canadian linked population-based data. Appl Physiol Nutr Metab. 2022;47(5):529-46.

13. Rein D, Ternes P, Demin R, Gierke J, Helgason T, Schön C. Artificial intelligence identified peptides modulate inflammation in healthy adults. Food & Function. 2019;10(9):6030-41.

14. Silva VC, Gorgulho B, Marchioni DM, Araujo TA, Santos IS, Lotufo PA, et al. Clustering analysis and machine learning algorithms in the prediction of dietary patterns: Cross-sectional results of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). J Hum Nutr Diet. 2022;35(5):883-94.

15. Tao D, Yang P, Feng H. Utilization of text mining as a big data analysis tool for food science and nutrition. Comprehensive Reviews in Food Science and Food Safety. 2020;19(2):875-94.

16. Kirk D, Catal C, Tekinerdogan B. Precision nutrition: A systematic literature review. Computers in biology and medicine. 2021;133:104365.

17. Matusheski NV, Caffrey A. Diets, nutrients, genes and the microbiome: recent advances in personalised nutrition. 2021;126(10):1489-97.

18. Miyazawa T, Hiratsuka Y, Toda M, Hatakeyama N, Ozawa H. Artificial intelligence in food science and nutrition: a narrative review. 2022;80(12):2288-300.

19. Oh YJ, Zhang J, Fang ML, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. The international journal of behavioral nutrition and physical activity. 2021;18(1):160.

20. Mukhamediev RI, Popova Y, Kuchin Y, Zaitseva E, Kalimoldayev A, Symagulov A, et al. Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges. Mathematics. 2022;10(15).

21. Santosh KC, Das N, Ghosh S. Chapter 2 - Deep learning: a review. In: Santosh KC, Das N, Ghosh S, editors. Deep Learning Models for Medical Imaging: Academic Press; 2022. p. 29-63.

22. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science. 2021;2(3):160.

23. Robinson KG, Akins RE. Chapter 24 - Machine learning in epigenetic diseases. In: Tollefsbol TO, editor. Medical Epigenetics (Second Edition). 29: Academic Press; 2021. p. 513-25.

24. Lugo Reyes SO. Chapter 21 - Artificial intelligence in precision health: Systems in practice. In: Barh D, editor. Artificial Intelligence in Precision Health: Academic Press; 2020. p. 499-519.

25. El Bouchefry K, de Souza RS. Chapter 12 - Learning in Big Data: Introduction to Machine Learning. In: Škoda P, Adam F, editors. Knowledge Discovery in Big Data from Astronomy and Earth Observation: Elsevier; 2020. p. 225-49.

26. Schneider P, Xhafa F. Chapter 8 - Machine learning: ML for eHealth systems. In: Schneider P, Xhafa F, editors. Anomaly Detection and Complex Event Processing over IoT Data Streams: Academic Press; 2022. p. 149-91.

27. Kaelbling LP, Littman, M.L., and Moore, A.W. Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research. 1996;4:237-85.

28. Hearty AP, Gibney MJ. Analysis of meal patterns with the use of supervised data mining techniques--artificial neural networks and decision trees. Am J Clin Nutr. 2008;88(6):1632-42.

29. Lazarou C, Karaolis M, Matalas AL, Panagiotakos DB. Dietary patterns analysis using data mining method. An application to data from the CYKIDS study. Comput Methods Programs Biomed. 2012;108(2):706-14.

30. Kastorini CM, Papadakis G, Milionis HJ, Kalantzi K, Puddu PE, Nikolaou V, et al. Comparative analysis of a-priori and a-posteriori dietary patterns using state-of-the-art classification algorithms: a case/case-control study. Artif Intell Med. 2013;59(3):175-83.

31. Biesbroek S, van der A DL, Brosens MCC, Beulens JWJ, Verschuren WMM, van der Schouw YT, et al. Identifying cardiovascular risk factor–related dietary patterns with reduced rank regression and random forest in the EPIC-NL cohort12. The American Journal of Clinical Nutrition. 2015;102(1):146-54.

32. Panaretos D, Koloverou E, Dimopoulos AC, Kouli GM, Vamvakari M, Tzavelas G, et al. A comparison of statistical and machine-learning techniques in evaluating the association between dietary patterns and 10-year cardiometabolic risk (2002-2012): the ATTICA study. Br J Nutr. 2018;120(3):326-34.

33. Easton JF, Román Sicilia H, Stephens CR. Classification of diagnostic subcategories for obesity and diabetes based on eating patterns. Nutr Diet. 2019;76(1):104-9.

34. Arceo-Vilas A, Fernandez-Lozano C, Pita S, Pértega-Díaz S, Pazos A. Identification of predictive factors of the degree of adherence to the Mediterranean diet through machine-learning techniques. PeerJ Comput Sci. 2020;6:e287.

35. Fu Y, Gou W, Hu W, Mao Y, Tian Y, Liang X, et al. Integration of an interpretable machine learning algorithm to identify early life risk factors of childhood obesity among preterm infants: a prospective birth cohort. BMC Medicine. 2020;18(1):184.

36. He X, Matam BR, Bellary S, Ghosh G, Chattopadhyay AK. CHD Risk Minimization through Lifestyle Control: Machine Learning Gateway. Scientific Reports. 2020;10(1):4090.

37. Yu EYW, Wesselius A, Sinhart C, Wolk A, Stern MC, Jiang X, et al. A data mining approach to investigate food groups related to incidence of bladder cancer in the BLadder cancer Epidemiology and Nutritional Determinants International Study. Br J Nutr. 2020;124(6):611-9.

38. Bôto JM, Marreiros A, Diogo P, Pinto E, Mateus MP. Health behaviours as predictors of the Mediterranean diet adherence: a decision tree approach. Public Health Nutr. 2021:1-13.

39. Soflaei SS, Shamsara E, Sahranavard T, Esmaily H, Moohebati M, Shabani N, et al. Dietary protein is the strong predictor of coronary artery disease; a data mining approach. Clin Nutr ESPEN. 2021;43:442-7.

40. Zhao Y, Naumova EN, Bobb JF, Claus Henn B, Singh GM. Joint Associations of Multiple Dietary Components With Cardiovascular Disease Risk: A Machine-Learning

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

Approach. American Journal of Epidemiology. 2021;190(7):1353-65.

41. Choi I, Kim J, Kim WC. Dietary Pattern Extraction Using Natural Language Processing Techniques. Frontiers in Nutrition. 2022;9.
42. Mousavi H, Karandish M, Jamshidnezhad A, Hadianfard AM. Determining the effective factors in predicting diet adherence using an intelligent model. Scientific Reports. 2022;12(1):12340.
43. Tatoli R, Lampignano L, Bortone I, Donghia R, Castellana F, Zupo R, et al. Dietary Patterns Associated with Diabetes in an Older Population from Southern Italy Using an Unsupervised Learning Approach. Sensors (Basel). 2022;22(6).
44. Yang J, Yang A, Yeung S, Woo J, Lo K. Joint Associations of Food Groups with All-Cause and Cause-Specific Mortality in the Mr. OS and Ms. OS Study: A Prospective Cohort. Nutrients. 2022;14(19).
45. Limketkai BN, Hamideh M, Shah R, Sauk JS, Jaffe N. Dietary Patterns and Their Association With Symptoms Activity in Inflammatory Bowel Diseases. Inflamm Bowel Dis. 2022;28(11):1627-36.
46. Qarmiche N, El Kinany K, Otmani N, El Rhazi K, Chaoui NEH. Cluster analysis of dietary patterns associated with colorectal cancer derived from a Moroccan case-control study. BMJ Health Care Inform. 2023;30(1).
47. Taheri R, ZareMehrjardi F, Heidarzadeh-Esfahani N, Hughes JA, Reid RER, Borghei M, et al. Dietary intake of micronutrients are predictor of premenstrual syndrome, a machine learning method. Clin Nutr ESPEN. 2023;55:136-43.
48. Huang AA, Huang SY. Exploring Depression and Nutritional Covariates Amongst US Adults using Shapely Additive Explanations. Health science reports. 2023;6(10):e1635.
49. Martin-Morales A, Yamamoto M, Inoue M, Vu T, Dawadi R, Araki M. Predicting Cardiovascular Disease Mortality: Leveraging Machine Learning for Comprehensive Assessment of Health and Nutrition Variables. Nutrients. 2023;15(18).
50. Lampignano L, Tatoli R, Donghia R, Bortone I, Castellana F, Zupo R, et al. Nutritional patterns as machine learning predictors of liver health in a population of elderly subjects. Nutrition, metabolism, and cardiovascular diseases : NMCD. 2023;33(11):2233-41.
51. Kirti K, Singh SK. Obesogenic diet and metabolic syndrome among adolescents in India: data-driven cluster analysis. BMC cardiovascular disorders. 2023;23(1):393.
52. Zhao M, Wan J, Qin W, Huang X, Chen G, Zhao X. A machine learning-based diagnosis modelling of type 2 diabetes mellitus with environmental metal exposure. Computer methods and programs in biomedicine. 2023;235:107537.
53. Li S, Hu X. Assessing the Risk of Prostate Cancer with Nutritional and Environmental Factors: A Cross-Sectional Study from National Health and Nutrition Examination Survey 2001-2010. Nutrition and cancer. 2023;75(5):1361-72.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]

Artificial Intelligence-based Approaches to Assess Dietary Patterns: A Systematic Review

Malmir H[1], Hosseinpour-Niazi S[1], Mirmiran P[1]

1- Nutrition Research Center in Endocrine Diseases, Research School of Endocrinology and Metabolism, Shahid Beheshti University of Medical Sciences, Tehran, Iran

*2- *Corresponding author: Nutrition Research Center in Endocrine Diseases, Research School of Endocrinology and Metabolism, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

Received 8 Aug, 2023 **Accepted 25 Nov, 2023**

Background and Objectives: Development of artificial intelligence has provided novel opportunities for research in the field of nutrition sciences. This study was carried out with the aim of reviewing and comprehensively assessing studies linked to the field of diet and food patterns, using artificial intelligence and machine learning algorithms.

Materials & Methods: All studies published until June 2023 were searched using PubMed Cochrane, EMBASE and SCOPUS databases with associated keywords. No time and language restrictions were used.

Results: After a complete review of the articles, 31 relevant articles were selected that were consistent with the purpose of the present study. Various machine-learning methods have various accuracy in predicting food patterns. For example, intelligent neural network is further accurate in predicting healthy food index quintiles, while it is further accurate in decision tree meals. Another use of machine learning includes extracting food patterns and investigating their relationships with various diseases such as obesity, heart diseases, stroke, risks of death from cardiovascular diseases and cancers. Machine learning methods such as decision trees are able to provide models for predicting adherence to various diets such as the Mediterranean diet.

Conclusion: Various artificial intelligence methods can help better understand food patterns linked to chronic diseases. The most important algorithms in the study of food patterns are decision tree, random forest, K-means, K-nearest neighbor, regression methods, support vector machine and intelligent neural network. These methods can help better understand dietary patterns associated with chronic diseases by categorizing and finding hidden associations between the groups and foods. Further studies are needed to better understand these connections.

Keywords: Artificial intelligence, Machine learning, Nutrition, Food patterns

[Downloaded from nsft.sbmu.ac.ir on 2026-11-10]

[DOI: 10.61186/nsft.19.2.87]