

داده‌کاوی رفتار خریداران مواد غذایی در مشهد

ناصر شاهنوشی، علی فیروززارع، میترا ژاله‌رجبی^۱

تاریخ دریافت: ۱۳۹۹/۰۶/۱۶

تاریخ پذیرش: ۱۳۹۹/۰۸/۲۲

چکیده

شناخت عامل‌های اقتصادی- اجتماعی مؤثر بر رفتار خرید مواد غذایی شهروندان از جهات مختلف از جمله برنامه‌ریزی برای ارتقای رفاه خریداران، هدایت رفتار خرید آتی، هدایت راهبردها و برنامه‌های تولید و بازاریابی و به تبع آن کاهش هزینه‌های مازاد و غیرضروری در این مرحله‌های، جلوگیری از هدررفت منابع‌های ارزشمند محیط‌زیستی و منابع طبیعی، کاهش حجم پسماندهای مواد غذایی و کاهش هزینه‌های بلندمدت درمان، دارای اهمیت بسزایی است. در این راستا این پژوهش تلاش کرد با بهره‌گیری از دانش داده‌کاوی و الگوریتم C4.5 درخت تصمیم و استفاده از داده‌ها و اطلاعات ۲۲۰ خریدار شهر مشهد، الگوریتم‌های تصمیم‌سازی خریداران در زمینه انتخاب سلامت، قیمت و طعم مواد غذایی به عنوان اولویت نخست ایشان در هنگام خرید این محصولات را بررسی و ارزیابی کند. نتایج این پژوهش نشان داد که متغیر سن خریداران مهم‌ترین متغیری است که می‌تواند بر رفتار خریداران در انتخاب مواد غذایی مؤثر باشد. متغیرهای دیگری نیز در سطح‌های بعدی بر انتخاب سلامت، قیمت یا طعم ماده غذایی به عنوان اولویت نخست خریداران مواد غذایی مؤثرند که از آن جمله می‌توان به متغیرهای درآمد خانوار، جنسیت و تحصیلات خریدار و شمار اعضای خانوار اشاره کرد که در طراحی راهبردها و برنامه‌های بازاریابی محصول‌ها و فرآورده‌های غذایی و همچنین تدوین راهبردهای سلامت توسط متولیان امر، بایستی مورد توجه قرار گیرد.

طبقه‌بندی JEL: C38، M31، Q18.

واژگان کلیدی: درخت تصمیم، الگوریتم C4.5، داده‌کاوی، معیار انتخاب.

^۱ به ترتیب استاد، استادیار (نویسنده مسئول) گروه اقتصاد کشاورزی دانشگاه فردوسی مشهد و دکتری اقتصاد کشاورزی، سازمان صنعت، معدن و تجارت استان زنجان

مقدمه

درک و مدیریت رابطه پیچیده بین مصرف‌کنندگان و محصولات و فرآورده‌های غذایی از جمله بنیادی‌ترین وظایف بازاریابان حوزه مواد غذایی، است. بیشتر پژوهش‌ها در مورد مصرف مواد غذایی بر این باورند که طعم خوب و درجه بالایی از سلامتی با یکدیگر ناسازگار هستند. بررسی در مورد اهمیت نسبی معیارهای انتخاب غذا بیانگر این است که در بسیاری از کشورها، طعم و مزه بر سلامتی اولویت دارد (Luomala et al., 2015). به عنوان مثال، در روسیه مهم‌ترین معیار انتخاب مواد غذایی، طعم و مزه است، در حالی که سلامتی در جایگاه پنجم قرار دارد (Honkanen & Frewer, 2009). مصرف‌کنندگان نیز اغلب بر این باورند که طعم و مزه و سلامتی در تضاد هستند (Luomala et al., 2004). (Raghuathan et al., 2006) در نتایج بررسی‌های خود نشان دادند هنگامی که به مصرف‌کنندگان مواد غذایی نشانه‌هایی از ناسالم بودن غذا داده می‌شود، انتظار دارند طعم آن بهتر از هنگامی باشد که گفته می‌شود ماده غذایی سالم باشد که البته این تأثیرها به صورت ناخودآگاه بروز می‌یابد، همچنین در این زمینه می‌توان به نتایج بررسی‌های (Mai and Hoffmann, 2015) مراجعه کرد. اما نتایج پژوهش دیگری نشان داد که در فرانسه، غذای ناسالم خودبه‌خود با بدمزه بودن ارتباط دارد، در حالی که غذای سالم با خوشمزه بودن ارتباط دارد (Werle et al., 2013).

از سوی دیگر، یکی از معیارهای دیگری که به ویژه در شرایط نامناسب اقتصادی و به طور خاص در دهک‌های متوسط و پایین درآمدی، ممکن است نقش ویژه‌ای در انتخاب مواد غذایی ایفا کند، قیمت مواد غذایی است. در چنین شرایطی خریداران مواد غذایی بایستی بین مواد غذایی سالم و با طعم مناسب، با قیمت‌های مختلف، دست به انتخاب بزنند و چه بسا ممکن است این انتخاب در شرایط نامساعد اقتصادی اربیبی و گرایش زیادی به سمت معیار قیمت داشته باشد.

بنابراین می‌توان گفت خریدار در هنگام خرید مواد غذایی با سه معیار متفاوت قیمت، سلامت و طعم و مزه روبرو است، که چنانچه به اهمیت انتخاب مناسب در چنین شرایطی توجه نشود، ممکن است در بلندمدت آسیب‌های جبران‌ناپذیری بر بدنه جامعه وارد آید. بر این مبنای لزوم مداخله‌های بنیادی و گسترده برای آشنا کردن افراد جامعه با اصول تغذیه به عنوان فراسنجه (پارامتر) اصلی سبک درست زندگی و فراهم آوردن امکانات مناسب عملی در این راستا مشخص و بدیهی است. از سوی دیگر شناخت ناکافی از الگوهای انتخاب خریداران مواد غذایی می‌تواند

داده کاوی رفتار خریداران...۳

ضمن زبان‌های شدید به واحدهای تولیدی مواد غذایی، ضایعات گسترده محصولهای کشاورزی و هدررفت منابع طبیعی و محیط‌زیست را در سطح کلان به دنبال داشته باشد. بنابراین لزوم برنامه‌ریزی در این زمینه به شدت احساس می‌شود. به یقین لازمه برنامه‌ریزی‌ها و سیاست‌گذاری‌های درست و کارآمد در این رابطه، شناسایی عامل‌های مؤثر بر اولویت‌های ایشان در انتخاب مواد غذایی است. در کنار پژوهش‌های اشاره شده در بالا، می‌توان به نتایج پژوهش‌های انجام شده توسط (Steiner, 1979), (Ajzen & Fishbein, 1980), (Murcott, 1989), (Ajzen, 1991), (Shepherd, 1999), (Köster and Mojet, 2007) اشاره کرد که در پژوهش‌های خود به بررسی اختلاف الگوهای غذایی افرادی مختلف و شناسایی عامل‌های مؤثر بر آن پرداخته‌اند. محققان زیادی نیز به تحلیل عامل‌های مؤثر بر رفتار مصرف‌کنندگان در زمینه سلامت مواد غذایی پرداخته‌اند. (Dosman, et al., 2001) در بررسی‌های خود دریافتند که زنان، افراد سالمند و خریداران با درآمد بالاتر نسبت به سایر گروه‌ها اهمیت بیشتری برای سلامت مواد غذایی قائل می‌باشند. نتایج بررسی‌های صورت گرفته توسط (Flynn et al., 1994) گویای آن است که زنان نسبت به مردان توجه بیشتری به سلامت رژیم غذایی و مواد غذایی منتخب خود داشته‌اند. (Krewski et al., 1994) نیز در بررسی‌های خود دریافتند که افراد سالمند بیشتر از افراد جوان سلامت مواد غذایی را در رژیم غذایی منتخب خود لحاظ می‌دارند. پژوهش صورت گرفته توسط (Galdas et al., 2005) نشان داده است که مردان نسبت به زنان روش‌های نادرست‌تری در سبک زندگی و الگوی تغذیه داشته، نسبت به بیماری‌های خود کم اهمیت‌تر بوده، نشانه‌های هشداردهنده را نادیده گرفته و دیرتر به مرکزهای درمانی مراجعه می‌کنند. بر مبنای یافته‌های (Baker, 2003) زنان، افراد سالمند، افراد با تحصیلات بالاتر و خانوارهای دارای کودکان و نوجوانان بیش از دیگر گروه‌ها، در انتخاب‌های خود سلامت مواد غذایی را در نظر گرفته و نگران مخاطره‌های ناشی از رژیم غذایی نادرست می‌باشند.

بنا بر استانداردهای سازمان بهداشت جهانی هر فرد باید بیشینه ۵ گرم نمک در روز مصرف کند، این در حالی است که نمک مصرفی ایرانی‌ها نزدیک به دو برابر استاندارد و معادل ۹,۵۲ گرم نمک در روز است (Rezaei et al., 2018). نتایج بررسی‌ها نشان می‌دهد سرانه قند و شکر مصرفی ایرانی‌ها ۷۱,۸ گرم در روز است در حالی که هر فرد روزانه بیشینه باید ۴۰ گرم قند مصرف کند (FAO, 2013). این موارد بیانگر این است که الگوی مصرف ایرانیان، فاصله قابل توجهی با استانداردهای سازمان بهداشت جهانی دارد.

با توجه به موارد اشاره شده در بالا در زمینه نقش و جایگاه قیمت مواد غذایی در هنگام خرید، اهمیت سلامت افراد جامعه، فاصله قابل توجه با الگوی بهینه مصرف مواد غذایی، هزینه‌های بالای درمان در بلندمدت، هدررفت منابع طبیعی و ضایعات گسترده محصولات غذایی، اهمیت شناخت عوامل مؤثر بر انتخاب مواد غذایی توسط افراد جامعه در زمان خرید، بیش از پیش احساس می‌شود و بر این مبنا مطالعه حاضر با بهره‌گیری از تکنیک درخت تصمیم داده‌کاوی به دنبال بررسی و شناسایی عوامل اقتصادی-اجتماعی مؤثر بر رفتار خرید مواد غذایی و معیار اولویت نخست ایشان در زمان خرید و انتخاب مواد غذایی است، تا بر پایه عوامل‌های شناسایی شده از الگوریتم‌های تصمیم خریداران، پیشنهادهایی در راستای هدفمندسازی راهبردها و برنامه‌های عرضه‌کنندگان مواد غذایی، ارتقای سطح رفاه شهروندان در زمان خرید و کارایی بیشتر بازاریابی محصول‌های غذایی ارائه دهد.

مواد و روش‌ها

داده کاوی فرایند کشف دانش^۱ وابستگی و ارتباط، الگوهای درونی، تغییرها، ساختارها و ناهنجاری‌ها، در میان میزان زیادی داده ذخیره‌شده در پایگاه‌های داده یا انبارهای داده یا مخزن‌های اطلاعات دیگر است. به دلیل در دسترس بودن حجم عظیمی از داده‌ها به صورت الکترونیکی و نیاز به تبدیل چنین داده‌هایی به اطلاعات و دانش سودمند برای کاربردهای مختلف، این دانش در سال‌های اخیر بسیار استفاده شده است. این کاربردها در زمینه‌های مختلفی از جمله هوش مصنوعی، یادگیری ماشین، تجزیه و تحلیل بازار، آمار و سیستم‌های پایگاه داده، مدیریت تجارت و پشتیبانی تصمیم‌گیری به چشم می‌خورند. در واقع در این شیوه، فرایند جستجوی خبره و ماهری به منظور شناسایی الگوهای معتبر، جدید، بالقوه، سودمند و قابل درک در داده‌ها صورت می‌گیرد (Marban et al., 2008). اصلی‌ترین دلیلی که سبب شده است داده‌کاوی در کانون توجه‌ها قرار گیرد، مسئله در دسترس بودن حجم گسترده‌ای از داده‌ها و نیاز شدید به تحلیل این داده‌ها و استخراج دانش سودمند از آنهاست. دو هدف اصلی داده‌کاوی پیش‌بینی و توصیف است که در هر کدام، هدف ایجاد الگویی به منظور شناسایی و یا پیش‌بینی خروجی‌هاست (Kusiak & Smith, 2007). در الگوریتم‌های توصیفی، بر یافتن الگوهای توصیفی داده‌ها که می‌توانند به وسیله انسان تعبیر شوند، تمرکز می‌شود (Kantardzic, 2003).

¹ Process of discovering knowledge

² Anomalies

داده کاوی رفتار خریداران...۵

طبقه‌بندی یکی از روش‌های داده کاوی است که داده‌ها را در گروه‌ها یا کلاس‌های از پیش تعریف‌شده ترسیم می‌کند. این روش، یک روش یادگیری نظارت‌شده^۱ است که برای ایجاد قانون‌هایی برای طبقه‌بندی داده‌های آزمون در گروه‌ها یا کلاس‌های از پیش تعریف‌شده، به داده‌های آموزشی دارای برچسب^۲ نیاز دارد. این فرایند دارای دو مرحله است. مرحله اول مرحله یادگیری است، که در آن داده‌های آموزش تجزیه و تحلیل می‌شود و قانون‌های طبقه‌بندی تولید می‌شود. مرحله بعدی، مرحله طبقه‌بندی است، جایی که داده‌های آزمون با توجه به قانون‌های تولیدشده، در طبقه‌های مختلف، طبقه‌بندی می‌شوند. از آنجا که الگوریتم‌های طبقه‌بندی ایجاب می‌کنند کلاس‌ها بر مبنای ارزش‌های (مقادیر) ویژگی داده‌ها^۳ تعریف شوند، در این پژوهش برای هر خریدار یک ویژگی "کلاس" ایجاد شده است که می‌تواند یک مقدار از آن را که شامل سلامت، قیمت و طعم است، به خود اختصاص دهد.

درخت‌های تصمیم^۴ یکی از متداول‌ترین، قوی‌ترین و شهودی‌ترین ابزارهای طبقه‌بندی می‌باشند. الگویی که توسط یک درخت تصمیم ایجاد می‌شود، می‌تواند به چند قانون "اگر - آنگاه" تبدیل شود و رابطه‌ها میان ویژگی‌های خروجی و ویژگی‌های ورودی در آن نمایان می‌شود. بر خلاف درخت تصمیم که به تولید قانون می‌پردازد، در شبکه عصبی مصنوعی تنها پیش‌بینی نهایی بیان می‌شود و چگونگی آن در خود شبکه پنهان می‌ماند (Corporation, 1999). علت نام‌گذاری این روش به عنوان درخت تصمیم این است که این درخت فرایند تصمیم‌گیری برای تعیین طبقه پاسخ یک داده ورودی را نشان می‌دهد.

برتری‌های استفاده از درخت تصمیم نسبت به دیگر روش‌های داده‌کاوی شامل دقت بالای درخت تصمیم نسبت به دیگر الگوهای داده‌کاوی، قابلیت کاربرد برای انواع مختلف داده‌ها اعم از طبقه‌ای و پیوسته، قابلیت محاسبه سریع‌تر نسبت به دیگر الگوهای طبقه‌بندی و قابلیت درک راحت‌تر قانون‌های به‌دست آمده در آن نسبت به دیگر الگوهای طبقه‌بندی است (Menga, 2020).

همان‌طور که در شکل ۱ نشان داده شده است، یک درخت تصمیم شامل یک گره ریشه‌ای واحد، برخی گره‌های داخلی و چندین گره برگ است. گره ریشه آغاز درخت تصمیم است. گره‌های

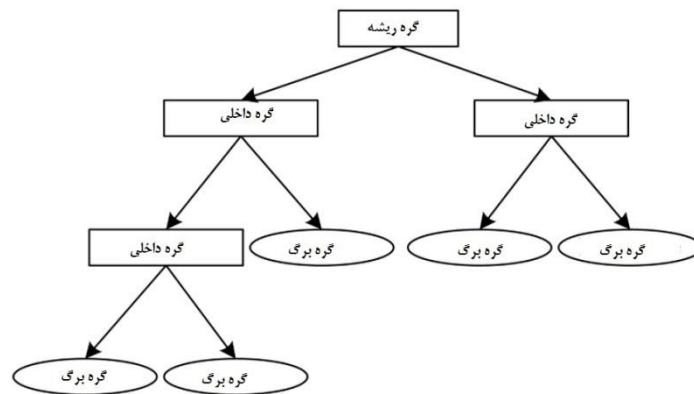
¹ Supervised learning method

² Labeled training data

³ Data attribute values

⁴ Decision Trees

درونی گره ریشه و گره‌های برگ را به هم متصل می‌کنند. هر گره برگ دارای یک برچسب کلاس است. مسیرهای از گره ریشه به گره‌های برگ، قانون‌های طبقه‌بندی را نشان می‌دهد.



شکل (۱) ساختار درخت تصمیم

Figure (1) Decision tree structure

در مقایسه با دیگر روش‌های طبقه‌بندی، درخت تصمیم یک روش یادگیری فعال و همچنین یک الگوی جعبه سفید^۱ است. پس از ساخت درخت تصمیم، می‌توان مجموعه‌ای از قانون‌های خرید مواد غذایی را شناسایی کرد. این قانون‌ها می‌تواند به عنوان راهنمای عملیاتی برای متولیان حوزه بازاریابی مواد غذایی برای نظارت و ارزیابی استفاده شود (Menga, 2020).

الگوریتم CART^۲، ID3^۳، CHAID^۴، C4.5^۵ از الگوریتم‌های پر استفاده درخت تصمیم هستند (Menga, 2020). در الگوریتم C4.5، درخت تصمیم از بالا به پایین ساخته می‌شود. برای تعیین ویژگی که برای خریداران مختلف، باید در ریشه درخت آموزش و شاخه‌های پایینی آن باشد، از معیاری به نام بهره اطلاعات^۶ استفاده می‌شود تا مشخص شود هر کدام از ویژگی‌ها شامل سن، جنسیت، درآمد خانوار، تحصیلات، بعد خانوار، وضعیت تأهل و دانش غذایی، تا چه حد قادر هستند داده‌های آموزشی خریداران را بر حسب طبقه آنان که بیانگر اولویت ایشان در خرید مواد غذایی است و سه گروه قیمت، سلامت و طعم را شامل می‌شود، جدا کنند. پس از انتخاب ویژگی

¹ White box model

² Classification And Regression Trees

³ Iterative Dichotomizer3

⁴ Chi-square automatic interaction detector

⁵ Classifier 4.5

⁶ Information Gain

داده کاوی رفتار خریداران...۷

از خریدار که باید در ریشه درخت قرار گیرد، برای هر یک (یا هر طیف) از مقادیر ممکن آن، یک شاخه ایجاد شده و داده‌های آموزشی بر مبنای ویژگی هر شاخه مرتب می‌شوند. سپس عملیات بیان شده در بالا برای داده‌های هر شاخه تکرار می‌شود، تا بهترین ویژگی برای گره بعدی انتخاب شود.

بهره اطلاعات یک ویژگی، عبارت است از مقدار کاهش آنروپی که به واسطه جداسازی داده‌ها از طریق این ویژگی به دست می‌آید. به عبارت دیگر، بهره اطلاعات Gain(S,A) برای یک ویژگی مانند A (به عنوان مثال تحصیلات یا درآمد) نسبت به مجموعه داده‌های S با استفاده از فرمول (۱) به صورت رابطه زیر تعریف می‌شود:

$$\text{Gain}(S, A) \equiv \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \quad (1)$$

که در آن Values(A) مجموعه همه مقدار ویژگی‌های A بوده و S_v زیرمجموعه‌ای از S است که برای آن A دارای مقدار V است. در تعریف بالا عبارت اول مقدار آنروپی داده‌ها و عبارت دوم مقدار آنروپی مورد انتظار بعد از جداسازی داده‌هاست. آنروپی نیز میزان خلوص یا بی‌نظمی مجموعه‌ای از داده‌های آموزشی را مشخص می‌کند. اگر مجموعه S شامل داده‌های مثبت و منفی از یک مفهوم هدف باشد، آنروپی S نسبت به این دسته‌بندی بولی به صورت زیر تعریف می‌شود (Sugumaran et al., 2008).

$$\text{Entropy}(S) \equiv -p_{\oplus} \log_2 p_{\oplus} - p_{\ominus} \log_2 p_{\ominus} \quad (2)$$

که p_⊕ نسبت داده‌های مثبت به کل داده‌ها و p_⊖ نسبت داده‌های منفی به کل داده‌هاست. همچنین log₂0=0 فرض می‌شود. اگر ویژگی هدف دارای C مقدار مختلف باشد آنروپی S نسبت به این دسته بندی C گانه به صورت زیر تعریف می‌شود:

$$\text{Entropy}(S) \equiv \sum_{i=1}^C -p_i \log_2 p_i \quad (3)$$

که در آن p_i نسبتی از S است که به دسته i تعلق دارند (Sugumaran et al., 2008).

تعیین اینکه کدام روش داده کاوی بهتر است به تفسیر مسئله توسط کاربران بستگی دارد. به طور معمول عملکرد الگوریتم‌ها با ارزیابی دقت نتیجه بررسی می‌شود. دقت طبقه‌بندی با تعیین درصد خریدارانی که در کلاس درست خود قرار می‌گیرند، محاسبه می‌شود.

عملکرد یک درخت تصمیم با استفاده از شاخص‌های نرخ درست مثبت (TPR)^۱، نرخ اشتباه مثبت (FPR)^۲، دقت (P)^۳، یادآوری (R)^۴ و مقدار F ارزیابی می‌شود (Geetha et al., 2010):

معیار TPR، که به آن میزان ضربه^۵ بیانگر نسبت خریدارانی است که در هر گروه به درستی طبقه‌بندی شده‌اند و به شکل ذیل محاسبه می‌شود:

$$TP \text{ Rate} \approx \frac{\text{Number of customers correctly classified}}{\text{Total number of customers}} \quad (۴)$$

معیار FPR نشان‌دهنده نسبت خریدارانی است که با استفاده از درخت تصمیم به درستی طبقه‌بندی نشده‌اند و چگونگی محاسبه آن به صورت زیر است:

$$FP \text{ Rate} \approx \frac{\text{Number of customers incorrectly classified}}{\text{Total number of customers}} \quad (۵)$$

معیار دقت و یادآوری نیز به صورت زیر محاسبه می‌شوند:

$$\begin{aligned} precision &= \frac{TP \text{ rate}}{TP \text{ rate} + FP \text{ rate}} \\ Recall &= TP \text{ rate} \end{aligned} \quad (۶)$$

معیاری که معیارهای FPR و TPR را ترکیب می‌کند، میانگین هارمونیک موزون دقت و یادآوری است و به صورت زیر محاسبه می‌شود:

$$F - \text{Measure} = \frac{2 \times precision \times Recall}{precision + Recall} \quad (۷)$$

^۱ TP (True Positive) rate

^۲ FP (False Positive) rate

^۳ Precision

^۴ Recall

^۵ Hit Rate

داده کاوی رفتار خریداران...۹

آماره کاپا^۱ نیز به منظور محاسبه دقت طبقه در مقایسه با دقت کل به کار می‌رود (Congalton, 1991) و بر پایه مؤلفه‌های ماتریس خطا به صورت زیر محاسبه می‌شود (Fitzgerald & Lees, 1994). اگر ماتریس خطا به صورت زیر باشد:

$$\begin{array}{cccc} p_{1,1} & \dots & p_{1,j} & p_{1,o} \\ \cdot & \dots & \cdot & \cdot \\ p_{i,1} & \dots & p_{i,j} & p_{i,o} \\ p_{o,1} & \dots & p_{o,j} & N \end{array}$$

نسبت توافق عمومی $p_0 = \sum_{i=1,N} p_{ii}$ و نسبت شانس مورد انتظار توافق $p_c = \sum_{i=1,N} p_{i0} p_{0i}$ است که برای محاسبه شاخص کاپا به صورت زیر محاسبه می‌شود:

$$\hat{K} = \frac{p_0 - p_c}{1 - p_c} \quad (A)$$

درخت تصمیمی مناسب‌تر است که ارزش این شاخص‌ها در آن بیشتر باشد (Fitzgerald & Lees, 1994).

در این پژوهش از الگوریتم J48 (معادل الگوریتم C4.5 در نرم‌افزار WEKA) به منظور یادگیری درخت تصمیم استفاده شده است.

در این پژوهش به منظور شناسایی عامل‌های مؤثر بر اولویت نخست افراد در انتخاب مواد غذایی از میان سلامت غذا، قیمت غذا و طعم آن پرسشنامه به عنوان ابزار گردآوری اطلاعات استفاده شد. برای طراحی پرسشنامه در آغاز با بررسی پیشینه پژوهش‌های انجام شده برخی از عامل‌های اقتصادی - اجتماعی مؤثر بر اولویت افراد در زمینه مواد غذایی تعیین شدند، آنگاه با برگزاری نشست‌هایی با کارشناسان مرتبط با موضوع مورد بحث - به منظور استفاده از دیدگاه کارشناسی - نسبت به اصلاح پرسشنامه طراحی شده اقدام شد تا از این طریق سعی شود تا حد امکان بیشتر متغیرهای مهم اقتصادی - اجتماعی تعیین‌کننده در تدوین پرسشنامه لحاظ شوند.

به منظور تعیین شمار نمونه و انجام آزمون‌های لازم به منظور اطمینان از اعتبار و قابلیت اعتماد آن یک پیش‌مطالعه^۲ انجام شد. در این راستا ۳۰ پرسشنامه تکمیل گردید و بر پایه نتایج بررسی‌ها

¹ Kappa Statistic

² Pilot study

نتایج و بحث

همان گونه که در میانی نظری نیز بیان شد، در این پژوهش در راستای بررسی فراسنجه‌ها و عامل‌های اقتصادی-اجتماعی مؤثر بر رفتار خریداران در اولویت نخست انتخاب مواد غذایی از الگوریتم C4.5 درخت تصمیم استفاده شده است.

به منظور آشنایی بیشتر با متغیرهای به کار گرفته شده در الگوی مورد بررسی، در این قسمت در آغاز، در جدول ۱ به توصیف این متغیرها پرداخته شده است.

جدول (۱) توصیف متغیرهای مورد استفاده در درخت تصمیم

Table (1) Description of the variables used in the decision tree

متغیر	شرح
اولویت اول	سلامت غذا=۰؛ قیمت غذا=۱؛ طعم غذا=۲
سن	کمتر از ۲۵ سال=۱، ۲۵-۳۵=۲، ۳۵-۴۵=۳، بالای ۴۵ سال=۴
جنسیت	مذکر=۱؛ مؤنث=۰
درآمد خانوار	درآمد پایین‌تر از کمینه دستمزد=۱، درآمد میانگین (بین کمینه دستمزد تا دو برابر آن)=۲، درآمد بالا (درآمد بالاتر از دو برابر حداقل دستمزد)=۳
تحصیلات	دیپلم و کمتر از آن=۱، کاردانی=۲، کارشناسی=۳، کارشناسی ارشد و دکترا=۴
بُعد خانوار	یک یا دو نفر=۱، سه نفر=۲، چهار و بیش از آن=۳
دانش مرتبط با مواد غذایی	خیر=۰، بله=۱

بنا بر اطلاعات جدول ۱، در این پژوهش متغیرهای اقتصادی-اجتماعی مؤثر بر اولویت نخست افراد در زمینه انتخاب مواد غذایی متغیرهای سن (age)، جنسیت (sex)، درآمد خانوار (inc)، تحصیلات (edu)، بعد خانوار (nofam) و دانش در زمینه مواد غذایی (knowledge) در نظر گرفته شده‌اند.^۱

به منظور اطمینان از نبود زمینه بروز خطای حفظ داده‌ها و نیز اطمینان از دقت طبقه بندی صورت گرفته توسط الگوی درخت تصمیم، داده‌ها به دو گروه آموزش و آزمون تقسیم می‌شود. به این منظور از ۲۲۰ مشاهده نمونه‌گیری شده، ۴۰ مشاهده به صورت تصادفی به عنوان داده‌های آزمون انتخاب شدند و در گروه آموزش و برازش الگو در نظر گرفته نشده‌اند.

نتایج الگوریتم C4.5 درخت تصمیم - همان گونه که پیشتر نیز تبیین شد، این پژوهش که درصدد بررسی و تحلیل عامل‌های اقتصادی-اجتماعی مؤثر بر الگوی تغذیه‌ای افراد و اولویت

^۱ این متغیرها پس از بررسی متغیرهای مختلف بسیار، با توجه به خروجی درخت تصمیم ارائه شده، در این جدول ارائه شده‌اند.

داده کاوی رفتار خریداران... ۱۱

نخست ایشان در زمان خرید و انتخاب مواد غذایی است، از درخت تصمیم که یکی از مناسب ترین روش های داده کاوی است و الگوریتم C4.5 به عنوان یکی از دقیق ترین الگوریتم های درخت تصمیم استفاده کرده است. در این راستا در آغاز به بررسی رفتار عمومی خریداران در انتخاب مواد غذایی پرداخته شده است و آنگاه رفتار گروه های مختلف خریدار به تفکیک ویژگی های تأثیرگذار به منظور تحلیل بیشتر در این زمینه بررسی شده است. پیش از تحلیل الگوریتم درخت تصمیم به دست آمده، لازم است دقت طبقه بندی انجام شده با استفاده از معیارهای مختلف بررسی نیکویی درخت برآزش شده، بررسی شود.

جدول (۲) دقت طبقه بندی خریداران بر مبنای اولویت اول خرید مواد غذایی

Table (2) precision of buyers' classification based on the first priority of food purchase

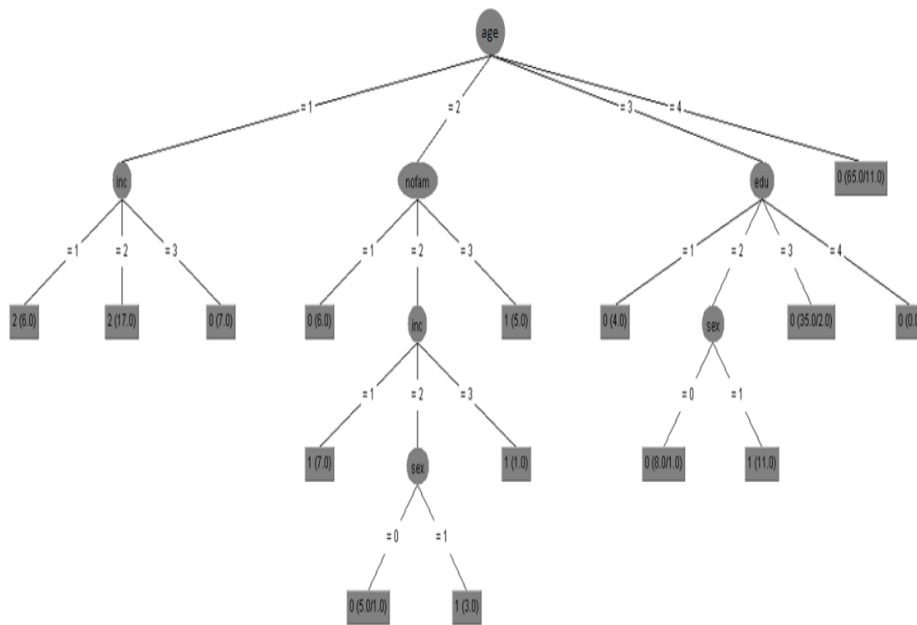
نتایج طبقه بندی الگوریتم C4.5 درخت تصمیم			شمار مشاهده ها	اولویت	
سلامت مواد غذایی	قیمت مواد غذایی	طعم و مزه مواد غذایی			
۱۱۵	۰	۰	۱۱۵	سلامت مواد غذایی	
٪۱۰۰	٪۰	٪۰			
۷	۲۷	۰	۳۴	قیمت مواد غذایی	
٪۲۰	٪۸۰	٪۰			
۸	۰	۲۳	۳۱	طعم و مزه مواد غذایی	
٪۲۶	٪۰	٪۷۴			
دقت طبقه بندی			۹۱،۶۶		
آماره کاپا			۰،۸۳		
F-Measure	Recall	Precision	FP rate	TP rate	آماره های دقت
۰،۹۴	۱	۰،۸۸	۰،۲۳	۱	(گروه ۱)
۰،۸۸	۰،۸۰	۱	۰	۰،۸۰	(گروه ۲)
۰،۸۵	۰،۷۴	۱	۰	۰،۷۴	(گروه ۳)
۰،۹۲	۰،۹۲	۰،۹۳	۰،۱۵	۰،۹۲	مجموع

Source: research findings

منبع: یافته های پژوهش

بنا بر اطلاعات جدول ۲، می توان گفت درخت برآزش شده در شکل ۱، دارای دقت طبقه بندی بالایی است و این درخت توانسته است ۱۰۰ درصد (۱۱۵ از ۱۱۵) خریداران گروه اول (خریدارانی که سلامت مواد غذایی برای آنان اولویت اول است)، ۸۰ درصد (۲۷ از ۳۴) خریداران گروه دوم

خریدارانی که قیمت مواد غذایی برای آنان اولویت اول است)، ۷۴ درصد (۲۳ از ۳۱) خریداران گروه سوم (خریدارانی که طعم و مزه مواد غذایی برای آنها اولویت اول است) و در مجموع در حدود ۹۲ درصد کل خریداران را به درستی طبقه‌بندی کند. بر این اساس آماره‌های دقت و یادآوری به ترتیب معادل ۰,۹۳ و ۰,۹۲ و بر مبنای آنها آماره F، معادل ۰,۹۲ محاسبه شده است، که بیانگر این است که درخت تصمیم به دست آمده دارای دقت بالایی می‌باشد. بر مبنای درخت تصمیم ارائه شده در شکل ۲، سن افراد به عنوان مهم‌ترین عامل تعیین‌کننده اولویت نخست افراد (سلامت غذا، قیمت و یا طعم) در انتخاب مواد غذایی تعیین شده است. سن خریداران در ۴ گروه زیر ۲۵ سال، ۲۵-۳۵، ۳۵-۴۵، و بالای ۴۵ طبقه‌بندی شده است.



شکل (۲) الگوریتم C4.5 درخت تصمیم داده‌های آزمون

Figure (2) Algorithm C4.5 of the test data decision tree

بنا بر نتایج به دست آمده، در گروه‌های سنی تعیین شده، درآمد، شمار اعضای خانوار و تحصیلات به عنوان عامل‌های مؤثر در قرارگیری خریداران در یکی از گروه‌های خریداران با اولویت نخست سلامت، قیمت و طعم مواد غذایی شناسایی شدند. تأثیرگذاری این عامل‌های به نحوی است که در خریداران زیر ۲۵ سال درآمد خریدار، در خریداران ۲۵-۳۵ سال شمار اعضای خانوار و برای خریداران ۳۵-۴۵ سال تحصیلات افراد عامل تعیین‌کننده اولویت نخست خریداران در انتخاب

داده کاوی رفتار خریداران...۱۳

مواد غذایی می‌باشند. نتایج گویای آن است که جنسیت در رتبه‌های پایین‌تر تأثیرگذاری و آن هم در گروه‌های سنی بین ۲۵ سال تا ۴۵ سال قرار گرفته است. همچنین بر مبنای درخت تصمیم به دست آمده متغیرهای وضعیت تاهل و دانش خریداران در زمینه مواد غذایی تأثیر معنی‌داری بر انتخاب خریداران نداشته و به عنوان عامل تعیین‌کننده در تصمیم‌گیری‌ها وارد نشده‌اند.

در درخت تصمیم با الگوریتم C4.5، عددهای اول هر گره برگ نشان‌دهنده اولویت نخست خریدار در انتخاب مواد غذایی است که عدد (۰) نشان‌دهنده سلامت مواد غذایی، عدد (۱) قیمت مواد غذایی و عدد (۲) طعم و مزه مواد غذایی به عنوان اولویت نخست فرد در هنگام انتخاب است. در هر گره برگ در درون پرانتز عدد اول نشان‌دهنده شمار مثال موجود در نمونه داده شده با ویژگی‌های مشخص شده و عدد دوم نشان‌دهنده شمار خطا در طبقه‌بندی گروه خریداران است. به عنوان مثال اولویت نخست زنان ۳۵ تا ۴۵ ساله با تحصیلات کاردانی، سلامت مواد غذایی بوده که ۸ نفر با این ویژگی در نمونه وجود داشته و از این شمار ۱ نفر به اشتباه در این گروه طبقه‌بندی شده است.

نتایج به دست آمده از طبقه‌بندی عمومی خریداران از یک سو بیانگر تأثیر منفی سن کم، شمار اعضای خانوار زیاد و درآمد کم در انتخاب سلامت مواد غذایی به عنوان اولویت اول خریداران است و از سوی دیگر نشان‌دهنده تأثیر مثبت همسویی افزایش سن، تحصیلات و درآمد در اهمیت به سلامت مواد غذایی از سوی خریداران است. به عبارت دیگر بنا بر نتایج به دست آمده و همسان با نتایج بررسی‌های Baker, (2003) تحصیلات دانشگاهی توانسته دیدگاه صحیحی در زمینه چگونگی انتخاب مواد غذایی به کار رفته در رژیم غذایی افراد جامعه ایجاد نماید و همسان با بررسی‌های (Dosman, et al., (2001، Krewski et al., (1994 و Baker, (2003 این دیدگاه با افزایش سن و مشاهده پیامدها و عوارض رژیم‌های نادرست و با افت توانایی جسمانی افراد تقویت می‌شود. نتایج گویای آن است که پایین بودن قدرت خرید (چه به دلیل پایین بودن درآمد خانوار و چه به دلیل شمار اعضای خانوار زیاد) عامل‌های دیگری غیر از سلامت به صورت قیمت و یا طعم مواد غذایی را در اولویت نخست خریداران قرار می‌دهد که با نتایج به دست آمده از مطالعات (Dosman, et al., (2001 هم‌راستا است. از دیگر نتایج قابل توجه در درخت تصمیم به دست آمده برای خریداران نتیجه به دست آمده برای مردان بالای ۲۵ تا ۳۵ سال در مقایسه با زنان با شرایط مشابه بوده است که همسان نتایج بررسی‌های (Flynn et al., (1994، Galdas et al.,

(2005) و Baker, (2003) مبنی بر توجه بیشتر زنان به سلامت مواد غذایی، نسبت به مردان است. بر این مبنا در شرایط همانند با وجود این که زنان سلامت مواد غذایی را اولویت نخست انتخاب مواد غذایی می‌دانند، قیمت مواد غذایی اولویت اصلی مردان بوده است. اگر چه این نتیجه را می‌توان به مسئولیت مردان در زمینه مدیریت و سرپرستی خانوار و نگرانی مادران در زمینه سلامت اعضای خانوار نسبت داد، بالا بودن آمار برخی بیماری‌ها در میان مردان نسبت به زنان در جامعه نیز خود مؤید این نتیجه‌گیری است.

بنا بر نتایج پژوهش‌های انجام شده توسط لون و همکاران (۲۰۰۹) و Luomala et al., (2015) به نظر می‌رسد ترسیم درخت تصمیم در گروه‌های جنسیتی خریداران تحلیل‌های روشن‌تری از تأثیر عامل‌های در اولویت نخست خریداران مواد غذایی ارائه دهد. از این رو در ادامه درخت‌های تصمیم با الگوریتم C4.5 برای خریداران در گروه‌های مختلف جنسیتی استخراج و تفسیر شده است.

بر پایه اطلاعات جدول ۳ که نشان‌دهنده دقت درخت تصمیم ارائه شده در مورد رفتار زنان و مردان خریدار است، درخت تصمیم رفتار زنان توانسته است در ۹۶٫۷ درصد موارد طبقه‌بندی صحیحی را در مورد رفتار خرید این گروه ارائه دهد و این در حالی است که این آمار در مورد مردان خریدار ۸۸٫۳ درصد است.

داده کاوی رفتار خریداران...۱۵

جدول (۳) طبقه‌بندی خریداران بر حسب جنسیت و بر مبنای اولویت خرید مواد غذایی
 Table (3) Classification of buyers by gender and by food purchase priority

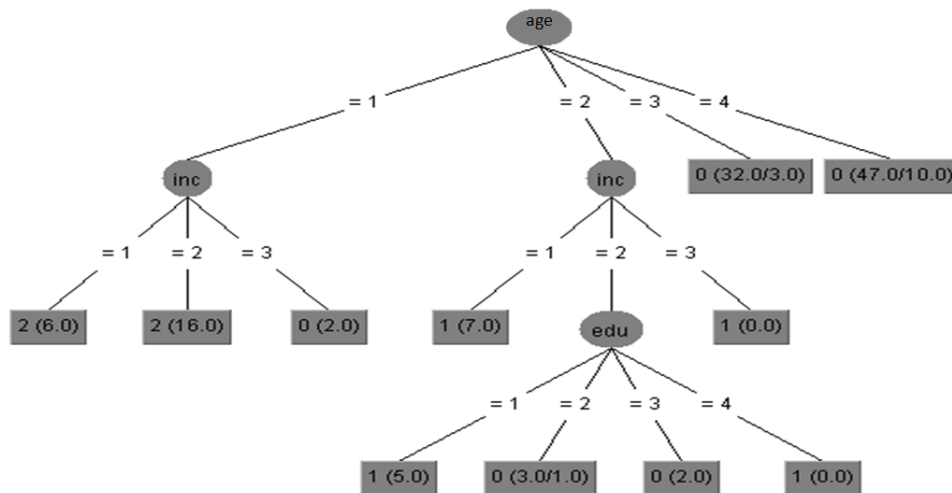
نتایج طبقه‌بندی الگوریتم C4.5 درخت تصمیم						شمار مشاهده‌ها				
طعم و مزه مواد غذایی		قیمت مواد غذایی		سلامت مواد غذایی		مرد	زن	اولویت		
مرد	زن	مرد	زن	مرد	زن					
۰	۰	۰	۰	۷۲	۴۳	۷۲	۴۳	سلامت مواد غذایی		
%۰	%۰	%۰	%۰	%۱۰۰	%۱۰۰					
۰	۰	۱۲	۱۵	۷	۰	۱۹	۱۵	قیمت مواد غذایی		
%۰	%۰	%۶۳	%۱۰۰	%۳۷	%۰					
۲۲	۰	۰	۰	۷	۲	۲۹	۲	طعم و مزه مواد غذایی		
%۷۶	%۰		%۰	%۲۴	%۱۰۰					
					زن					
					مرد					
					%۸۸,۳۳		دقت طبقه‌بندی			
					۰,۷۷		آماره کاپا			
F-Measure		Recall		Precision		FP rate		TP rate		آماره‌های دقت
مرد	زن	مرد	زن	مرد	زن	مرد	زن	مرد	زن	
۰,۹۱	۰,۹۸	۱	۱	۰,۸۳	۰,۹۶	۰,۲۹	۰,۱۲	۱	۱	(گروه ۱)
۰,۷۷	۱	۰,۶۳	۱	۱	۱	۰	۰	۰,۶۳	۱	(گروه ۲)
۰,۸۶	۰	۰,۷۶	۰	۱	۰	۰	۰	۰,۷۶	۰	(گروه ۳)
۰,۸۸	۰,۹۵	۰,۸۸	۰,۹۷	۰,۹۰	۰,۹۳	۰,۱۷	۰,۰۸	۰,۸۸	۰,۹۷	مجموع

Source: research findings

منبع: یافته‌های پژوهش

در مورد تحلیل رفتار مردان خریدار که نتایج آن در شکل ۲ ارائه شده‌است، مشاهده می‌شود پس از متغیر سن خریدار (به عنوان مهم‌ترین متغیر)، متغیر درآمد در سطح دوم و تحصیلات خریداران در سطح بعدی، متغیرهای تأثیرگذار بر رفتار خرید ایشان می‌باشند. تحلیل رفتار مردان با در نظر گرفتن الگوریتم C4.5 درخت تصمیم شکل ۳، نشان می‌دهد که برابر با انتظار و نتایج بررسی‌های پیشین مردان بالای ۳۵ سال سلامت را اولویت نخست خرید

مواد غذایی می‌دانند. بر مبنای درخت به دست آمده در حالی که سلامت مواد غذایی اولویت نخست مردان زیر ۲۵ سال و با درآمد بالا است، برای مردان زیر ۲۵ سال و با درآمد کم و متوسط طعم و مزه مواد غذایی جایگزین سلامت مواد غذایی می‌شود که این نتایج موید تأثیر مثبت افزایش درآمد بر انتخاب مواد غذایی بر مبنای سلامت آن است، که این نتیجه همانند نتیجه به دست آمده در بررسی انجام شده توسط Steenhuis et al., (2011) است. در میان مردان ۲۵ تا ۳۵ ساله اولویت نخست افراد در درجه اول تابع درآمد و در سطح بعدی تابع تحصیلات است. انتظار بر این است که با افزایش درآمد و تحصیلات توجه به سلامت غذایی افزایش یافته و به جای قیمت و طعم مواد غذایی، سلامت اولویت اصلی خریداران باشد. اما بر خلاف انتظار در گروه سنی مردان ۲۵ تا ۳۵ سال مشاهده می‌شود که مردان ۲۵ تا ۳۵ سال با درآمد بالا، قیمت را اولویت نخست در خرید مواد غذایی در نظر می‌گیرند، که این بیانگر کم توجهی ایشان به شاخص‌های سلامت بوده و ممکن است در سال‌های آینده سلامت جامعه را با خطر روبه رو سازد. با کنار هم قرار دادن نتایج به دست آمده برای خریداران مرد می‌توان چنین نتیجه‌گیری کرد که برابر با انتظار و نتایج بررسی‌های پیشین، افزایش سن (و به دنبال آن کاهش توان جسمی و افزایش ابتلا به بیماری و یا تشدید بیماری‌ها)، موجب افزایش توجه به سلامت مواد غذایی در خرید مواد غذایی می‌شود.



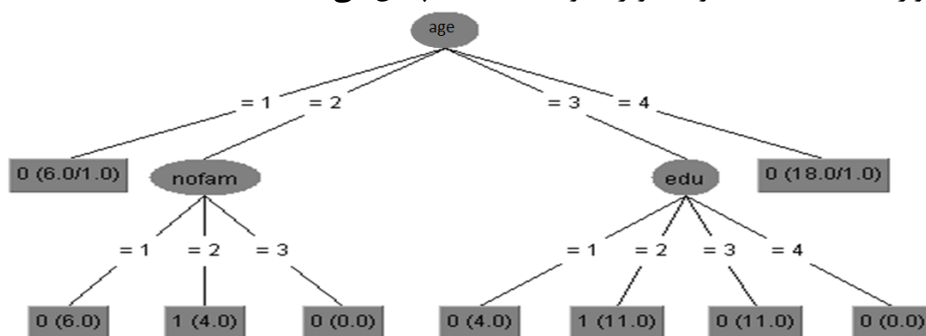
شکل (۳) درخت تصمیم رفتار مردان خریدار در انتخاب مواد غذایی

Figure (3) Decision tree of men buyers' Behavior in food choice

داده کاوی رفتار خریداران...۱۷

شکل ۴ که ارائه‌دهنده درخت تصمیم تحلیل رفتار زنان خریدار است، نشان می‌دهد که در این گروه پس از سن خریداران (به عنوان تعیین‌کننده‌ترین متغیر)، متغیرهای شمار اعضای خانوار و تحصیلات متغیرهایی هستند که رفتار خریداران را در انتخاب بین سلامت، قیمت و طعم مواد غذایی به عنوان اولویت اول در انتخاب مواد غذایی تحت تأثیر قرار می‌دهند. بر مبنای نتایج به دست آمده، زنان جوان با سن کمتر از ۲۵ سال با توجه به حساسیت این گروه افراد جامعه در زمینه حفظ سلامتی و جوانی و زنان سالمند بالای ۴۵ سال به دلیل ضعف قوای جسمانی و نگرانی در زمینه بروز و تشدید بیماری‌ها فارغ از دیگر ویژگی‌های اقتصادی-اجتماعی همگی سلامت مواد غذایی را مهم‌ترین اولویت در خرید و انتخاب، عنوان کرده‌اند. در مورد زنان ۲۵ تا ۳۵ سال اولویت‌ها در انتخاب مواد غذایی به شمار اعضای خانواری که فرد متعلق به آن است بستگی دارد. به طوری که برای زنان دارای خانواده ۱ یا ۲ نفره سلامت و برای زنان عضو خانوار پرجمعیت‌تر (۳ نفر) قیمت مواد غذایی در اولویت قرار دارد. برای زنان ۳۵ تا ۴۵ سال، انتخاب مواد غذایی تابع تحصیلات ایشان بوده و برابر با انتظار، با افزایش تحصیلات، سلامت مواد غذایی در اولویت اول انتخاب مواد غذایی ایشان است.

با جمع‌بندی نتایج به دست آمده برای خریداران زن می‌توان چنین نتیجه‌گیری کرد که برابر با انتظار، افزایش سن، تحصیلات و درآمد (با کاهش شمار اعضای خانوار) موجب افزایش توجه خریداران به سلامت مواد غذایی می‌شود و کاهش تحصیلات و افزایش شمار اعضای خانوار قیمت و طعم مواد غذایی را در کانون توجه ایشان قرار می‌دهد. از این‌رو یکی دیگر از گروه‌های هدف در زمینه برنامه‌ریزی و اصلاح نگرش در زمینه الگوی مصرف مواد غذایی زنان ۲۵ تا ۳۵ سال، با خانوارهای با جمعیت متوسط و زیاد و با تحصیلات پایین می‌باشند.



شکل (۴) درخت تصمیم رفتار زنان خریدار در انتخاب مواد غذایی
Figure (4) Decision tree of women buyers' Behavior in food choice

پیش‌بینی طبقه‌بندی خریداران در گروه‌های با اولویت سلامت، قیمت و طعم مواد غذایی داده‌های آموزش و آزمون - همان گونه که پیشتر نیز بیان شد، بنا بر نتایج جدول ۲ الگوی درخت تصمیم برآزش شده توانسته است در ۹۱٫۶ درصد موارد، طبقه‌بندی عضویت خریداران در طبقه‌های مختلف با اولویت نخست سلامت، قیمت و طعم و مزه مواد غذایی را به درستی انجام دهد. نتایج گویای آن است که ۱۰۰ درصد از خریداران با اولویت نخست سلامت مواد غذایی، ۸۰ درصد خریداران با اولویت اول قیمت و ۷۴ درصد خریداران با اولویت نخست طعم و مزه مواد غذایی به درستی طبقه بندی شده‌اند. اما با توجه به احتمال حفظ کردن روابط میان متغیرها از سوی الگو، به منظور ارزیابی توانایی قدرت پیش‌بینی‌کنندگی الگو، مشاهده‌ها حذف شده از نمونه، به صورت داده‌های آزمون (که در فرایند آموزش الگو وارد نشده‌اند)، به الگو معرفی و طبقه‌بندی برای این مثال‌ها بر مبنای درخت برآزش شده در فرایند آموزش صورت گرفت که نتایج آن در جدول ۴ ارائه شده است.

جدول (۴) طبقه‌بندی خریداران گروه آزمون بر مبنای اولویت خرید مواد غذایی

نتایج طبقه‌بندی الگوریتم C4.5 درخت تصمیم			شمار مشاهده‌ها	اولویت	
سلامت مواد غذایی	قیمت مواد غذایی	طعم و مزه مواد غذایی			
۱۴	۰	۰	۱۴	سلامت مواد غذایی	
٪۱۰۰	٪۰	٪۰	۱۲	قیمت مواد غذایی	
۱	۱۱	۰	۱۴	طعم و مزه مواد غذایی	
٪۸	٪۹۲	٪۰			
۵	۰	۹			
٪۳۶	٪۰	٪۶۴			
٪۸۵				دقت طبقه‌بندی	
۰٫۷۷				آماره کاپا	
Precision	Recall	F-measure	FP rate	TP rate	آماره‌های دقت
۰٫۷۰	۱	۰٫۸۲	۰٫۲۳	۱	(گروه ۱)
۱	۰٫۹۲	۰٫۹۶	۰	۰٫۹۲	(گروه ۲)
۱	۰٫۶۴	۰٫۷۸	۰	۰٫۶۴	(گروه ۳)
۰٫۸۹	۰٫۸۵	۰٫۸۵	۰٫۰۸	۰٫۸۵	مجموع

Source: research findings

منبع: یافته‌های پژوهش

داده کاوی رفتار خریداران...۱۹

نتایج این جدول نشان می‌دهد که ۸۵ درصد خریداران بخش آزمون، با بهره‌گیری از درخت تصمیم به دست آمده از بخش آموزش، به درستی در گروه خود پیش‌بینی شده‌اند. بر این مبنا ۱۰۰ درصد خریداران با اولویت نخست سلامت (۱۴ نفر از ۱۴ نفر)، ۹۲ درصد خریداران با اولویت نخست قیمت (۱۱ نفر از ۱۲ نفر) و ۶۴ درصد خریداران با اولویت نخست طعم و مزه مواد غذایی (۹ نفر از ۱۴ نفر) در خرید مواد غذایی، به درستی در گروه خود پیش‌بینی شده‌اند. بر این مبنا نسبت خریدارانی است که در هر گروه به درستی طبقه‌بندی شده‌اند.

نتیجه‌گیری و پیشنهادها

این پژوهش که به بررسی رفتار خرید مواد غذایی توسط شهروندان مشهودی، می‌پردازد، سه معیار تأثیرگذار در خرید مواد غذایی را که هر یک نقش ویژه‌ای بر رفتار انتخاب خریداران دارد، مورد واکاوی قرار داده است. شرایط اقتصادی جامعه، وضعیت سلامت و بیماری‌ها (از جمله بیماری‌های قلبی عروقی، دیابت، پوکی استخوان و برخی بیماری‌های بدخیم) و مرگ‌ومیرهای گسترده ناشی از الگوهای نادرست مصرف مواد غذایی و طبع و ذائقه انسان به سمت مواد غذایی خوش طعم و خوشمزه، سه معیاری هستند که البته نمی‌توان گفت همواره خریدار را در هنگام انتخاب، به سمت گزینه واحدی رهنمون سازند. لذا این انتخاب، همواره یکی از مهم‌ترین انتخاب‌ها در زمان خرید مواد غذایی است. این پژوهش از میان عامل‌های اقتصادی-اجتماعی مختلفی، که در رفتار انتخاب مصرف‌کننده می‌تواند نقش مؤثری ایفا کند، با بهره‌گیری از درخت تصمیم و الگوریتم C4.5 آن به بررسی نقش متغیرهای سن، جنسیت، درآمد خانوار، تحصیلات، بعد خانوار، وضعیت تأهل و دانش مرتبط با مواد غذایی، بر رفتار انتخاب خریداران پرداخته است.

بنا بر نتایج این پژوهش، افراد با سن بالای ۴۵ سال و افراد با سن بین ۳۵ تا ۴۵ سال، متناسب با میزان تحصیلات خود، توجه بیشتری به معیار سلامت مواد غذایی معطوف داشته‌اند. به عبارت دیگر، افزایش سن و به دنبال آن آغاز نشانه‌هایی از بروز برخی بیماری‌ها در بدن به خودی خود توانسته است، معیار سلامت را در اولویت ذهنی اول خریداران جای دهد، این در حالی است که خریداران ۳۵ تا ۴۵ ساله، متناسب با سطح تحصیلات خود، تا حدود اندکی رفتارهای متفاوتی در خرید مواد غذایی خواهند داشت. بنا بر بررسی‌های انجام شده، نتایج برخی بررسی‌های گویای این است که متغیرهای سن، جنسیت و دانش، متغیرهای مؤثری بر رفتار خرید مواد غذایی هستند (Azmawani, 2015; Hardin-Fanning and Gokun, 2014). بر این مبنا، سن و

جنسیت دو متغیری هستند که بر انتخاب غذای سالم، تأثیر معنی‌داری دارند (Hardin-Fanning & Gokun, 2014) همچنین بر مبنای نتایج این بررسی‌ها، دانش، در نگرش عمدی، تأثیر دارد (Azmawani, 2015; Hardin-Fanning and Gokun, 2014) و از این رو، افزایش دانش در توجه به موضوع، تأثیر خواهد گذاشت (Bang et al, 2000). بر این مبنای یافته‌ها نشان می‌دهد که بین دانش و تصمیم‌گیری‌های خرید رابطه وجود دارد (Hamdan et al, 2013). این در حالی است که همان‌گونه که در این پژوهش دیده می‌شود، چنانچه دانش را با توجه به تعریف ارائه شده توسط بریلیانا و مارسیتو به صورت آگاهی یا آشنایی حاصل از تجربه یا یادگیری تعریف کنیم (Briliana & Mursito, 2017)، می‌تواند در قالب تحصیلات یا دانش مرتبط بررسی شود، که همان‌گونه که در بالا بیان شد، تحصیلات متغیری تأثیرگذار بر رفتار خرید مواد غذایی توسط شهروندان مشهدی، دانسته شده است، در شرایطی که متغیر دانش مرتبط، در هیچ یک از درخت تصمیم‌های ارائه شده، متغیری تعیین‌کننده در تمایز رفتار، تشخیص داده نشده است، که البته این نتیجه در پژوهش انجام شده توسط هاردین-فانینگ و گوکان در سال ۲۰۱۴ نیز دیده می‌شود (Hardin-Fanning & Gokun, 2014). بر این مبنای به نظر می‌رسد ضمن نیاز به بررسی‌های بیشتر این موضوع در قالب پژوهش‌های آتی (با توجه به نتایج متناقض در پژوهش‌های مختلف)، لازم است متولیان امر تغذیه، توجه ویژه‌ای در قالب برنامه‌های آگاهی‌دهی و اطلاع‌رسانی به شهروندانی که از سطح تحصیلاتی پایین‌تری برخوردارند، معطوف نمایند. همچنین محققان بازار به طور کلی و برندهای تولید مواد غذایی به طور خاص، باید فعالیت‌های بازاریابی خود را با تمرکز بر ایجاد آگاهی در مورد هماهنگی آنها با محصولات سالم طراحی کنند. اگر مصرف‌کنندگان در برخی از بازارها میان‌سال و مسن هستند، محصولات جدید و چگونگی تبلیغ این محصولات باید برابر با معیارهای سلامتی و تأثیرهایی که از خرید مواد غذایی سالم در ذهن مصرف‌کنندگان تأیید می‌شود، تهیه و ارائه شود.

متغیر مؤثر دیگری که بر رفتار خریداران مواد غذایی در شهر مشهد، نقش ایفا می‌کند، متغیر سطح درآمد خریدار است که البته در پژوهش‌های مختلف نتایج متفاوتی از تأثیرگذاری آن ارائه شده است. به عنوان مثال این متغیر در پژوهش انجام شده توسط (Steenhuis et al., 2011)، نقش تعیین‌کننده‌ای در خرید مواد غذایی بر مبنای معیار سلامت ایفا کرده است (2011) در حالی که بر مبنای نتایج پژوهش انجام شده توسط Hardin-Fanning & Gokun در سال ۲۰۱۴، این متغیر نتوانسته است نقشی تأثیرگذار بر رفتار خرید مواد غذایی، داشته باشد. از این

داده کاوی رفتار خریداران... ۲۱

مهم این نتیجه قابل استنتاج است، که بستر اقتصادی جامعه مورد بررسی در پژوهش، می‌تواند نقشی قابل توجه بر تأثیرگذاری این متغیر اعمال کند. انتظار بر این است که جامعه‌ها با متوسط سطح رفاه اقتصادی بالاتر، بتوانند فارغ از سطح درآمد خود، انتخابی مناسب در خرید مواد غذایی داشته باشند، این در حالی است که در جامعه‌هایی که وضعیت اقتصادی با تنگنای بیشتری رو به رو است، این متغیر نقش مهم‌تری ایفا خواهد کرد. لذا بر این مبنا، به محققان بازار و فعالان عرصه بازاریابی مواد غذایی پیشنهاد می‌شود ضمن در نظر گرفتن، سطح درآمدی جامعه مخاطب خود، به تولید و عرضه مواد غذایی متناسب با نیاز و خواست ایشان اقدام کنند.

افزون بر موارد اشاره شده، نتایج این پژوهش همانند نتایج پژوهش انجام شده توسط Hardin-Fanning & Gokun, 2014. مبین این نکته است که زنان در انتخاب مواد غذایی، توجه بیشتری به معیار سلامتی در مقایسه با معیارهای قیمت و طعم و مزه، از خود نشان می‌دهند. بر این مبنا چنانچه هدف متولیان سلامت جامعه، ارتقای وضعیت سلامت الگوی مصرف مواد غذایی باشد، لازم است ضمن آموزش اهمیت شاخص‌های سلامت مواد غذایی به مردان - به عنوان گروه تعیین‌کننده خریدار مواد غذایی-، خرید خانوادگی و به همراه همسر را تشویق و ترویج کنند.

توجه به این برنامه‌ها می‌تواند، هزینه‌های سنگین درمان که در بلندمدت بر بدنه جامعه تحمیل می‌شود را کاهش داده و کیفیت زندگی بهتری را برای شهروندان به ارمغان آورد. البته در بلندمدت بایستی ارتقای وضعیت اقتصادی خانواده‌ها به منظور کاهش دغدغه قیمت برای ایشان در زمان خرید، مد نظر سیاست‌گذاران قرار گیرد تا از این طریق ناحیه ممکن خرید برای انتخاب خریداران بهبود یابد.

منبع‌ها

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50: 179-211.
- Ajzen, I., and Fishbein, M. (1980). Understanding attitudes and predicting social behavior. Englewood Cliffs, NJ: Prentice-Hall.
- Azmawani, A. R., Asrarhaghghi, E., Suhaimi, A. R. (2015). Consumers and halal cosmetic products: Knowledge, religiosity, attitude and intention. *Journal of Islamic Marketing*, 6(1): 148-163.
- Baker, G. A. (2003). Food safety and fear: factors affecting consumer response to food safety risk. *International Food and Agribusiness Management Review*, 6(1): 1-11.

- Bang, H. K., Ellinger, A. E., Hadjimarcou, J., and Traichal, P. A. (2000). Consumer concern, knowledge, belief, and attitude toward renewable energy: An application of the reasoned action theory. *Psychology and Marketing*, 17(6): 449-468.
- Briliana, V., and Mursito, N. (2017). Exploring antecedents and consequences of Indonesian Muslim youths' attitude towards halal cosmetic products: A case study in Jakarta. *Asia Pacific Management Review*, 22: 176-184.
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of the Environment*, 37: 35-46.
- Corporation, T. C. (1999). Introduction to Data Mining and Knowledge Discovery.
- Dosman, D. M., Adamowicz, W. L., Hruday, S. E. (2001). Socio-economic determinants of health- and food safety-related risk perceptions. *Risk Analysis*, 21(2): 307-317.
- Fitzgerald, R. W. Lees, B. G. (1994). Assessing the classification accuracy of multisource remote sensing data. *Remote Sensing of the Environment*, 47, 362-368.
- Flynn, J., Slovic, P., Mertz, C. K. (1994). Gender, race and perception of environmental health risks. *Risk Analysis*, 14(6), 1101-1108.
- Galdas, P.M., Cheater, F., and Marshall, P. (2005). Men and health help-seeking behavior: literature review. *Journal of Advanced Nursing*, 49(6), 616-623.
- Geetha, S., Ishwarya, N., Kamaraj, N. (2010). Evolving decision tree rule based system for audio stego anomalies detection based on Hausdorff distance statistics. *Information Sciences* 180, 2540-2559.
- Hamdan, H., Issa, Z. M., Abu, N., Jusoff, K. (2013). Purchasing decisions among Muslim Consumers of processed halal food products. *Journal of Food Products Marketing*, 19(1), 54-61.
- Hardin-Fanning, F., Gokun, Y. (2014). Gender and age are associated with healthy food purchases via grocery voucher redemption. *Rural Remote Health*. 14(3): 2830.
- <http://www.fao.org/faostat/en>
- Kantardzic, M., (2003). *Data Mining: Concepts, Models, Methods, and Algorithms*. John Wiley.
- Köster, E. P., Mojet, J. (2007). Theories of food choice development. In L. Frewer & H. van Trijp (Eds.), *Understanding consumers of food products*. 93-124. Abington, Cambridge, England: Woodhead Publishing Limited.
- Krewski, D., Slovic, P., Bartlett, S., Flynn, J., Mertz, C. (1994). Health risk perceptions in Canada (ERC 94-3, Environmental Risk Management Working Paper). Edmonton, Alberta, Canada: University of Alberta.
- Kusiak, A., Smith, M. (2007). Data mining in design of products and production systems, *Annual Reviews in Control*, 31, 147- 156.

داده کاوی رفتار خریداران...۲۳

- Luomala, H., Jokitalo, M., Karhu, H., Hietaranta-Luoma, H. L., Hopia, A., Hietamäki, S. (2015). Perceived health and taste ambivalence in food consumption. *Journal of Consumer Marketing*. 32 (4). 290 -301.
- Mai, R. and Hoffmann, S. (2015). How to combat the unhealthy tasty intuition: the influencing role of health consciousness, *Journal of Public Policy & Marketing*, 34 (1): 63-83.
- Marbán, O., Menasalvas, E., Fernández-Baizán, C. (2008). A cost model to estimate the effort of data mining projects (DMCoMo). *Information Systems*. 33. 133-150.
- Menga, X., Zhanga, P., Xub, Y., Xiea, H. (2020). Construction of decision tree based on C4.5 algorithm for online voltage stability assessment. *Electrical Power and Energy Systems*, 118.
- Murcott, A. (1989). Sociological and social anthropological approaches to food and eating. *World Review of Nutrition and Dietetics*, 55, 1–40.
- Rezaei, Sh., Mahmoudi, Z., Sheidaei, A., Aryan, Z., Mahmoudi, N., Gohari, K., Yoosefi, M., Hajipour, M., Dilmaghani-Marand, A., Soleimanzadehkhayat, M., Gholami, A., Mirab Samiee, S., Moradi, Gh., Larijani, B., Farzadfar, F. (2018). Salt intake among Iranian population: the first national report on salt intake in Iran. *Journal of Hypertension*. 36. 1. 10.1097
- Shepherd, R. (1999). Social determinants of food choice. *Proceedings of the Nutrition Society*, 58, 807–812.
- Steenhuis, I. H. M., Waterlander, W.E., de Mul, A. (2011). Consumer food choices: the role of price and pricing strategies. *Public Health Nutrition*: 14(12), 2220–2226.
- Steiner, J. E. (1979). Human facial expressions in response to taste and smell stimulation. In W. R. Hayne & P. L. Lewis (Eds.). *Advances in child development and behavior* (13, 257–295). New York: Academic Press.
- Sugumaran, V., Sabareesh, G. R., Ramachandran, K. I. (2008). Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine. *Expert Systems with Applications* 34, 3090–3098.
- Werle, C. O. C., Trendel, O. & Ardito, G. (2013). Unhealthy food is not tastier for everybody: the 'healthy tasty' French intuition, *Food Quality and Preference*, 28 (1): 116-121.
- Rezaei, Sh., Mahmoudi, Z., Sheidaei, A., Aryan, Z., Mahmoudi, N., Gohari, K., Yoosefi, M., Hajipour, M., Dilmaghani-Marand, A., Soleimanzadehkhayat, M., Gholami, A., Mirab Samiee, S., Moradi, Gh., Larijani, B., Farzadfar, F. (2018). Salt intake among Iranian population: the first national report on salt intake in Iran. *Journal of Hypertension*. 36. 1. 10.1097



Data mining of Food Buyers behavior in Mashhad

Naser Shahnoushi, Ali Firoozzare, Mitra Jalerajabi¹

Received: 06 Sep.2020

Accepted: 12 Nov.2020

Extended Abstract

Introduction

Recognizing the socio-economic factors affecting the food buying behavior of citizens in various aspects, including planning to improve customer welfare, conducting future buying behavior, guiding production and marketing strategies and programs, and consequently reducing unnecessary costs in these stages, preventing the loss of valuable environmental resources and natural resources, reducing the volume of food waste and reducing the long-term costs of treatment, is very important. In this regard, the present study tried to investigate the decision-making algorithms of buyers in the field of health, price and taste of food as their first priority when buying these products.

Materials and method

In order to determine the extent to which each of the characteristics, including age, gender, household income, education, household size, marital status, and food knowledge, can categorize buyers according to their first priority in purchasing food- Which includes three criteria of health, price and taste of food, in this study, C4.5 algorithm was used to learn the decision tree of 220 buyers in Mashhad. Decision trees are one of the most common, powerful, and intuitive classification tools. Compared to other classification methods, the decision tree is an active learning method as well as a white box model. Once the decision tree is built, a set of food purchase rules can be identified.

¹ Respectively: Professor, Assistant Professor, faculty staff of agricultural economics department of Ferdowsi university of Mashhad & Ph.D. Industry, Mine and Trade Organization of Zanjan
Email: firooz@ferdowsi.um.ac.ir

Results and discussion

The results of this study showed that the variable of buyers' age is the most important variable that can affect the behavior of buyers in food selection. Other variables in the next levels also affect the choice of health, price or taste of food as the first priority of food buyers, including the variables of household income, gender and education of the buyer and the number of household members, Which should be considered in the design of strategies and marketing plans for food products as well as the development of health strategies by those in charge.

Suggestion

Health officials should pay special attention to citizens with lower levels of education in the form of awareness programs. Also, market researchers in general and food brand managers in particular, should design their marketing activities with a focus on raising awareness about their compliance with healthy products. Food marketers are advised to supply food according to their target communities' needs and wants and their level of income.

JEL Classification: C38 ·M31 ·Q18.

Key Words: Decision Tree , C4.5 Algorithm, Data mining, choice criteria.