

## شناسایی مهم ترین پارامترهای مؤثر در تفکیک سفر با استفاده از روش نظریه مبنا

مهیار عربانی، علی منصور خاکی و بابک امانی

**چکیده:** فاکتورها و ویژگیهای متنوع و عدیده‌ای بر نوع سیستم انتخابی فرد در طول سفر اثر می‌گذرانند. از اینرو می‌توان با اطمینان گفت که شناسایی کوتاهترین الگوریتم‌های تصمیم‌گیری ممکن از میان تعداد زیاد پارامترهای اثرگذار، همواره یک گام اساسی در مدلسازیهای تفکیک سفر است. یکی از جدیدترین و قویترین ابزارهای ریاضی برای تعیین کوتاهترین الگوریتم تصمیم‌گیری، نظریه مبنا (Rough-Set) است که توسط پاولاک در سال ۱۹۹۱ ارائه شده است. این پژوهش به ارزیابی کارایی این نظریه در مدلسازیهای تفکیک سفر می‌پردازد و هدف آن شناسایی مهم‌ترین پارامترهای تصمیم‌گیرنده در نوع انتخاب افراد و استخراج قواعد و قوانین تصمیم‌گیری است تا با مشخص شدن آنها از میزان حجم وسیع اطلاعات لازم که در روشهای متداول مدلسازی تفکیک سفر مورد نیاز است، کاسته شود. در این تحقیق تلاش شده تا از طریق روش برازش خطی گام به گام (Stepwise) که یکی از متداول‌ترین روشها در تعیین پارامترهای مؤثر در تفکیک سفر است، نیز اقدام به شناسایی بهترین الگوریتم ممکن از صفات تعریف شده از ویژگیهای رفتاری فرد و سفر او شود تا مقایسه بهتر بتوان عملکرد روش نظریه مبنا را تحت بررسی قرار داد. نتایج حاصل از این بررسی‌ها عملکرد بهتر و قویتر نظریه مبنا در شناسایی بهترین الگوریتم‌های تصمیم‌گیری را نشان می‌دهد. همچنین با مطالعه موردی در مورد شهر همدان مشاهده می‌شود که روش نظریه مبنا دارای عملکرد مطلوبی در تعیین پارامترهای مؤثر در تفکیک سفر است. نتایج این پژوهش نشان می‌دهد که منظور تعیین پارامترهای مذکور، نظریه مبنا ابزاری قوی تر در مقایسه با روش برازش خطی گام به گام است.

**واژه های کلیدی:** نظریه مبنا، تفکیک سفر، الگوریتم تصمیم‌گیری

### ۱. مقدمه

مرحله تفکیک سفر از مهمترین مراحل پیش‌بینی در مدل‌های کلاسیک برنامه‌ریزی حمل و نقل است. اهمیت این مرحله به دلیل نقش حیاتی سیستم حمل و نقل عمومی در راهبرد و تعیین سیاست‌گذاری‌ها و تصمیم‌گیری‌های حمل و نقلی است. بدون استثناء می‌توان ادعا کرد که کلیه سیستم‌های حمل و نقل عمومی به شکل بهتری از فضا (سطح زمین) بهره می‌برند. همچنین امروزه با حضور سیستم‌های جابه‌جایی زیرزمینی و سایر تجهیزات ریلی،

مقاله در تاریخ ۱۳۸۳/۲/۱۳ دریافت شده و در تاریخ ۱۳۸۳/۱۲/۲۲ به تصویب نهایی رسیده است.

مهیار عربانی، دانشکده فنی و مهندسی، دانشگاه گیلان.  
علی منصور خاکی، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران.

بابک امانی، دانشکده تحصیلات تکمیلی، دانشگاه آزاد واحد تهران جنوب.

سطح بهره‌وری سیستم‌های حمل و نقلی افزایش یافته است و این سیستمها دیگر باعث بروز مسائل ترافیکی و مانند آن نمی‌گردند. بنابراین اگر اشخاص به جای استفاده از وسایل نقلیه شخصی به استفاده از سیستم حمل و نقل عمومی و ریلی تمایل پیدا کنند، سطح سرویس در سیستم‌های جاده‌ای نیز افزایش می‌یابد. نتایج تفکیک سفر بر کل بازدهی و عملکرد جابجایی یک شهر تأثیر گذار بوده و تعیین‌کننده کل فضاهای مورد نیاز یک شهر برای تخصیص به فعالیت‌های حمل و نقلی و مانند آن می‌باشند. بنابراین ساخت مدل‌های صحیح و مطمئن جهت برآورد نوع سیستم حمل و نقل انتخابی افراد یک گام اساسی و مهم است [1]. در این گام، یکی از مهم‌ترین موارد، شناسایی پارامترهای اصلی اثرگذار در نوع انتخاب افراد است. در این پژوهش با استفاده از نظریه اولیه (یا مبنا) پارامترهای اصلی مؤثر در نوع سیستم انتخابی افراد شناسایی و ارزیابی شده‌اند.

## ۲. مدل‌سازی تفکیک سفر

هر فردی در طول سفر خود بین حالات ممکن نحوه انجام یک سفر شیوه ای را بر می‌گزیند. به عنوان مثال یک فرد برای رفتن به مقصد می‌تواند یکی از موارد ذیل را انتخاب کند: خودش شخصاً رانندگی کند، شخص دیگری دنبالش بیاید و او را همراه خود ببرد، با اتوبوس به مقصد برود، پیاده روی کند یا سوار موتورسیکلت شود. باید توجه داشت که انتخاب هر یک از این حالات تا حد زیادی به ویژگیهای رفتاری افراد بستگی دارد و نحوه انتخاب در افراد مختلف، با توجه به عوامل مختلفی مانند نوع و هدف از سفر، هزینه های سفر و سطوح سرویس دهی متفاوت است.

ویژگیهای سفر از جمله عواملی هستند که تأثیر بسزایی در نوع انتخاب افراد دارند. به نظر می‌رسد که معمولاً در اکثر موارد افراد برای سفرهای کاری یا آموزشی خود از سیستمهای حمل و نقل عمومی استفاده می‌کنند. در حالی که در سفرهای اجتماعی، تفریحی و مانند آنها (در صورت داشتن وسیله نقلیه شخصی) تمایل به استفاده از وسیله نقلیه شخصی خود دارند.

علاوه بر مشخصات سیستم های گوناگون حمل و نقل و ویژگیهای سفرها، وضعیت‌ها و شرایط مختلف اجتماعی - اقتصادی افراد نیز در نحوه انتخاب آنها موثر است. این شرایط را می‌توان با توجه به سطوح درآمد، سن و غیره رده بندی کرد.

در این میان، آنچه که باید مورد توجه قرارگیرد، آن است که درصد قابل توجهی از افراد جامعه به دلایل گوناگون امکان استفاده از وسایل نقلیه شخصی را ندارند. بنابراین جابه‌جایی آنها منحصراً به استفاده از وسایل حمل و نقل عمومی وابسته است. از این گروه می‌توان افراد مسن، کم درآمد، با سن کم و افرادی را که در منزل یک وسیله نقلیه شخصی دارند ولی آن وسیله بیشتر در اختیار سایر افراد خانواده قرار داشته و آنها در مرتبه دوم استفاده قرار می‌گیرند را نام برد [2].

بدین ترتیب پارامترهای مؤثر در تفکیک سفر به دو گروه عمده تقسیم بندی می‌شوند. دسته اول صفاتی هستند که تابع ویژگیهای مبدأ و مولد سفر مانند هدف از سفر و امکان دسترسی می‌باشند و دسته دوم نیز وابسته به ویژگیهای مقصد و مسیرهای عبوری مانند هزینه و زمان سفر هستند.

با توجه به دسته بندی فوق، مدل‌های تفکیک سفر به دو گروه مجزا تقسیم می‌شوند. گروهی تابع دسته اول از پارامترها هستند (یعنی در صورتی که مرحله توزیع سفر درست بعد از مرحله تولید سفر مورد ارزیابی قرار گیرد) و در آنها مشخصات سفرها لحاظ نمی‌گردند. به این مدل‌ها مدل‌های قبل از توزیع گویند. گروهی دیگر تابع دسته دوم از پارامترها هستند (یعنی هنگامی که مدل درست بعد از مرحله توزیع سفر مورد ارزیابی قرار گیرد) و در آنها ویژگیهای مولد سفر نادیده گرفته می‌شوند، در این حالت این مدل‌ها، مدل‌های بعد از توزیع (Post-Distribution or Trip)

(Interchange) نامیده می‌شوند. امروزه محققان در مطالعات حمل و نقل تلاش می‌کنند تا در ساخت مدل‌های تفکیک سفر هر دو گروه از پارامترها را مورد ارزیابی قرار داده و از ترکیب مدل‌های قبل و بعد از توزیع استفاده کنند [2, 3].

علاوه بر تقسیم بندی فوق، مدل‌های تفکیک سفر را به دو دسته کلی تر مدل‌های متراکم و غیر متراکم (Disaggregate) دسته بندی می‌نمایند. این دسته بندی با توجه به سطوح تعریف شده برای این مدل‌ها صورت می‌گیرد. هرگاه در یک مدل برای ارزیابی و پیش بینی یک انتخاب نیاز به بررسی ویژگیهای رفتاری گروهی از افراد جامعه باشد (مثلاً یک منطقه)، آنگاه آن مدل از نوع مدل‌های متراکم است و در صورتی که یک مدل بر مبنای بررسی ویژگیهای رفتاری یک فرد بنا شده باشد از نوع مدل‌های غیر متراکم در نظر گرفته می‌شود.

از انواع مدل‌های متراکم می‌توان مدل‌های قبل از توزیع، مدل‌های بعد از توزیع و مدل‌های ترکیبی (Synthetic Models) را نام برد. از این مدل‌ها (مدل‌های متراکم) در سالهای متمادی (تا اوایل دهه ۱۹۸۰) استفاده شده است [1]. هر چند که محققانی چون وارنر (Warner, 1962)، اوآی و شالدینر (Oi and Shuldiner, 1962) مدعی بودند که این مدل‌ها دارای نقاط ضعف فراوانی هستند [4, 5]. لذا با پیدایش مدل‌های نسل دوم تفکیک سفر (غیر متراکم) و روند روبه افزایش کاربرد آنها به جای مدل‌های اولیه، به مرور مدل‌های متراکم کنار گذاشته شدند و جای خود را به مدل‌های تفکیک سفر غیر متراکم (مبنا فردی) دادند (See Williams, 1981) [6].

از مزایای عمده مدل‌های تفکیک سفر غیر متراکم (نسل دوم مدل‌ها) (مبنا فردی) می‌توان به موارد زیر اشاره نمود (Spear, 1977):

۱- در مدل‌های تفکیک سفر غیر متراکم اساس و مبنای مدل بر اساس تئوری رفتارهای فردی بنا نهاده شده است. بنابراین از مزایای مهم این مدل‌ها قابلیت برآورد و انتقال پذیری بهتر آنها در شرایط زمانی و مکانی مختلف می‌باشد.

۲- مدل‌های تفکیک سفر غیر متراکم (DM) در کاربرد اطلاعات از مدل‌های متعارف پیشین تفکیک سفر کارایی بهتری دارند. در این سیستم، اطلاعات کمتری جهت پیش بینی نوع انتخاب یک فرد نیاز است در حالی که در مدل‌های متراکم برآورد یک انتخاب گاهی میانگین انتخاب صدها نفر می‌باشد.

۳- مدل‌های تفکیک سفر غیر متراکم مدل‌هایی احتمالی هستند که احتمال انتخاب هر یک از سیستم‌های حمل و نقلی را در مقایسه با باقی موارد بیان می‌کنند و به طور دقیق مشخص نمی‌کنند که کدام یک از حالات انتخاب می‌گردند. لذا جهت پیش بینی و کاربرد آنها می‌توان از مفاهیم پایه ای احتمالات سود جست.

۴- در مدل‌های تفکیک سفر غیر متراکم امکان کاربرد هر تعداد پارامتر و با هر مشخصاتی برای پیش بینی پارامتر وابسته وجود دارد.

- هزینه‌های اعمالی سفر (Relative Monetary Costs)
- میزان دسترسی به پارکینگ و هزینه پارک خودرو
- سطح دسترسی به سیستم های حمل و نقل
- ب - پارامترهای کیفی:
- راحتی
- منظم بودن و قابل اعتماد بودن سیستم
- ایمنی
- عادت

#### ۴. تئوری Rough-Set

تئوری Rough-Set یا نظریه مبنا ابزار ریاضی قدرتمندی است که در تجزیه و تحلیل اطلاعات مبهم، نامطمئن و غیردقیق (یعنی مواردی که ابهام و عدم قطعیت در آنها وجود دارد) کارایی دارد. این نظریه اولین بار توسط پاولاک (Pawlak) در سال ۱۹۹۱ ارائه گردیده است [8].

تئوری Rough-Set در استخراج قوانین و روابط موجود از جداول تصمیم گیری (Skowron, 1995)، پردازش داده‌های کلینیکی (Tsumoto, 1998)، یادگیری و آموزش طراحی روابط ادراکی (Arciszewski et al., 1993)، استخراج قوانین پیش بینی حجم آب مورد نیاز (An et al., 1996) و پردازش و مدیریت اطلاعات روسازی‌ها (Attoh-Okine, 1997) کارایی‌های چشمگیری داشته است [9-12]. همچنین اخیراً یک چهار چوب کلی برای نظریه مبنا توسط Kryszkiewicz (1998, 1999) ارائه شده است. این چهارچوب قابلیت تحلیل اطلاعات ناقص را نیز به نظریه مبنا می‌دهد [13, 14].

روش Rough-Set یک روش آماری غیر پارامتری است که به منظور بررسی مجموعه فاکتورهای گوناگون و غیر محسوس کارایی دارد. این روش ابزاری جهت تبدیل مجموعه‌ای از اطلاعات مثل مجموعه موارد مشاهده شده در گذشته و یا تجارب اندوخته شده، در قالب یک ساختار اطلاعاتی و طبقه بندی موضوعات (موارد) مختلف با توجه به صفات و ویژگیهای مشخص هر یک از آنها است. اطلاعات ناقص باعث می‌شوند که نتوان موضوعات (Objects) مختلف را با توجه به مقادیر ارائه شده در صفات و ویژگیهای تعریف شده آنها از یکدیگر متمایز ساخت. بنابراین در این میان مجموعه‌های تقریبی بالا و پایین از مجموعه موضوعات، تنها مجموعه‌هایی هستند که دقیقاً با توجه به مقادیر کران‌های صفات موجود قابل تشخیص و شناسایی می‌باشند [15]. در ادامه بیشتر با مفاهیم پایه ای این تئوری آشنا خواهیم شد. (برای جزئیات بیشتر به مراجع Van Slownski and Stefanowski 1994, den Bergh et al 1997, Greco et al 1995 و Pawlak 1991 رجوع شود) [16-19].

دانش و اطلاعات بشری بر مبنای تجربه های بشر از پدیده ها و یافته های او شکل گرفته‌اند. این اطلاعات در یک سیستم کلی که

درحالی که در مدل‌های تفکیک سفر پیشین کاربران می توانستند تنها از تعداد محدودی پارامتر استفاده کنند. بنابراین در این مدلها (DM) امکان کاربرد پارامترهای اضافی که معرف بهتر سیاستها و تصمیم گیریهای حمل و نقلی باشد، به راحتی امکان پذیر است [1, 7].

به طور کلی با توجه به مباحث فوق و مقایسه مزایا و معایب مدل‌های مختلف می توان به دو نتیجه مهم دست یافت:

۱- در صورتی که اطلاعات بر مبنای ویژگیها و مشخصات فردی بنا شده باشند بهتر می توانند مفسر و معرف نوع سیستم حمل و نقلی انتخابی افراد باشند.

۲- ویژگیهای رفتاری مؤثر فرد در شیوه سفر انتخابی او را می توان توسط سه گروه از پارامترها و صفات، مورد ارزیابی قرار داد که عبارتند از:

ویژگیهای اجتماعی - اقتصادی سفر کننده  
مشخصات سفر

مشخصات انواع سیستمهای حمل و نقلی موجود

#### ۳. معرفی پارامترها و صفات مؤثر در تفکیک سفر

بهترین ساختار بررسی پارامترها و اطلاعات برای پیش بینی نوع سیستم انتخابی افراد، ارزیابی همزمان ویژگیها و صفات رفتاری فرد، مشخصات سفر و نوع سیستم انتخابی او است. به طور کلی این ویژگیها عبارتند از:

۱- خصوصیات سفر کننده:

- تعداد وسیله نقلیه شخصی خانوار

- داشتن گواهینامه رانندگی

- سن

- جنسیت

- ساختار خانوار

- درآمد

- فاصله از مرکز خرید شهر

- تراکم منطقه مسکونی

- و در نهایت سایر عواملی که میتوانند در نوع انتخاب فرد مؤثر باشند.

۲- ویژگیهای سفر (Characteristics of the Journey)

- هدف از سفر

- زمان روز

- مسافت سفر

- راستای سفر

۳- ویژگیهای تجهیزات حمل و نقلی:

الف - پارامترهای کمی:

- زمان سفر نسبی (Relative Travel Time)

- زمان اضافی سفر (Excess Travel Time)

همان سیستم اطلاعات تعریف شده در فوق است. همچنین می‌توان برای هر  $x$ ، توصیفی از آن در قالب زیر مجموعه‌هایی از صفات  $Q$  ( $P \subseteq Q$ ) ارائه نمود.

یک مفهوم اساسی در تئوری Rough-Set معرفی یک ارتباط دوتایی غیر قابل تمایز (تشخیص) است که آن را با نماد  $Ip$  نمایش می‌دهند. دو نمونه  $x$  و  $y$  را ( $x, y \in U$ ) با توجه به صفات  $P$  غیر قابل تمایز ( $P$ -indiscernible) گویند، هرگاه برای مجموعه صفات  $P$  ( $P \subseteq Q$ ) هر دو نمونه از توصیف یکسانی برخوردار باشند. لذا زیر مجموعه‌هایی از  $U$  که هر یک شامل گروهی از نمونه‌های با توصیف یکسان با توجه به مجموعه صفات  $P$  باشند را مجموعه‌های ابتدایی ( $P$ -elementary sets) می‌نامند. در صورتی که این تقسیم بندی برای کل مجموعه صفات  $Q$  صورت گیرد، به مجموعه‌های ابتدایی بدست آمده  $Q$ ، اتم‌ها (Atoms) نامیده می‌شوند. مجموعه‌های ابتدایی  $P$  ( $P \subseteq Q$ ) بین مجموعه اعضای (نمونه‌های)  $U$  طوری تقسیم بندی می‌شود که هر  $x$  ( $x \in U$ ) تنها به یکی از مجموعه‌های ابتدایی  $P$  تعلق داشته باشد.

معرفی دو مفهوم کلیدی دیگر برای توضیح نظریه Rough-Set ضروری است. فرض کنید  $P$  یک زیر مجموعه از  $Q$  ( $P \subseteq Q$ ) و  $X$  یک زیر مجموعه از  $U$  باشد. تقریب پایین  $P$  از  $X$  ( $P$ -lower approximation of  $X$ ) که با نماد  $P_L X$  معرفی می‌شود، زیرمجموعه‌ای از  $U$  است که اعضای آن تمامی موضوعاتی (نمونه‌هایی) هستند که دقیقاً به مجموعه‌های ابتدایی  $P$  تعلق دارند که شامل اعضای مجموعه  $X$  می‌باشند. به بیان دیگر اعضای  $P_L X$  تمامی نمونه‌هایی از  $U$  هستند که (فقط) به تمامی گروه‌های ایجاد شده با توجه به ارتباط غیر قابل تمایز  $Ip$  تعلق داشته و شامل اعضای  $X$  می‌باشند.

همچنین تقریب بالای  $P$  ( $P$ -upper approximation) از  $X$  که با نماد  $P_U X$  معرفی می‌شود، زیر مجموعه‌ای از  $U$  است که اعضای آن تمامی نمونه‌هایی هستند که به مجموعه‌های ابتدایی  $P$  تعلق داشته و دارای حداقل یک عضو از مجموعه  $X$  می‌باشند. به بیان دیگر اعضای  $P_U X$  تمامی نمونه‌هایی از  $U$  هستند که به تمامی گروه‌های ایجاد شده با توجه به ارتباط غیر قابل تمایز  $Ip$  تعلق داشته و حداقل شامل یک عضو از مجموعه  $X$  می‌باشند.

تفاوت بین دو مجموعه فوق را تحت عنوان کران  $P$  از  $X$  ( $P$ -boundary of  $X$ ) معرفی می‌کنند و به صورت  $Bnp(X) = P_U X - P_L X$  قابل ارائه است. بنابراین می‌توان استنتاج کرد که رابطه  $P_L X \subseteq X \subseteq P_U X$  برقرار بوده و نتیجتاً اگر عضوی ( $x$ ) متعلق به  $P_L X$  باشد آن  $x$  متعلق به  $X$  نیز خواهد بود. اگر عضو  $x$  متعلق به  $P_U X$  باشد آن عضو ممکن است به  $X$  تعلق داشته باشد. در نتیجه  $Bnp(X)$  معرف یک منطقه مشکوک خواهد بود که نمی‌توان در خصوص اعضای آن با اطمینان

اصطلاحاً سیستم اطلاعات نامیده می‌شود، ذخیره سازی می‌گردد. سیستم اطلاعات شامل داده‌هایی درباره موضوعات خاص مورد بحث و نظر (موضوعات، مشاهدات، نمونه‌ها، یافته‌ها، اتفاقات و...) و همچنین عوامل و ویژگی‌های موثر بر آنها (صفات، مشخصات، متغیرها، نشانه‌ها) است. این مجموعه صفات به دو دسته تفکیک می‌شوند. دسته اول که از نتایج آزمایشات و اندازه‌گیریها یا اطلاعات مشاهداتی برای ما قابل برآورد هستند را صفات وضعیت یا مستقل نامیده می‌شود. دسته دوم که مربوط به تصمیم‌گیری‌های کارشناسان و تشخیص و طبقه‌بندی و ارزیابی نتایج با توجه به آن صفات می‌گردند را صفات وابسته (Decision Attributes) گفته می‌شود [20]. پس برای هر اتفاق یا هر موضوع و پدیده‌ای دو دسته از صفات را می‌توان تعریف کرد. یک دسته از صفات جزء ویژگیها و مشخصات پدیده‌ها هستند و دسته دیگر در واقع صفات ارزیابی کننده یا وابسته می‌باشند. لذا با توجه به دسته دوم از این صفات و مشاهده تفاوت‌ها و طبقه‌بندیهای متفاوت بین آنها در موضوعات (نمونه‌های) مختلف، می‌توان به بررسی بر روی صفات مشخصات (وضعیت یا وابسته) نمونه‌ها پرداخت.

فرض بر این است که بتوان با رجوع به یک مجموعه مشخص از نمونه‌ها یا موضوعات  $U$ ، علل اختلاف و تمایز بین آنها را با مشاهده اطلاعات مربوط به هر یک از آنها شناسایی کرد. یک مجموعه محدود از صفات  $Q$  را می‌توان معرفی کرد که کل موضوعات با توجه به آن صفات قابل تعریف، شناسایی و ارزیابی باشند. از آنجا که هدف تئوری Rough-Set طبقه بندی و تمایز اطلاعات بر مبنای تفاوت‌های مشاهده شده در مقادیر صفات تعریف شده برای هر موضوع در مقایسه با موضوعات دیگر می‌باشد لذا هر صفت  $q$  ( $q \in Q$ ) باید از مقادیر متفاوتی در کران تعریف شده خود  $Uq$  برخوردار باشد. بنابراین هر صفت برای آنکه قابل تشخیص و ارزیابی باشد، نیاز به حداقل دو مقدار (سطح) دارد. در صورتی که صفت کمی باشد کران تعریف شده آن صفت، در عمل قابل تقسیم بندی به تعداد مناسبی بازه مجزا و کوچکتر است به قسمی که معرف خوبی برای پدیده مورد بررسی باشند. لذا باید از تعریف سطوح اضافی که در آنالیز نیز کارایی و کاربردی ندارند، اجتناب کرد. در این میان نحوه انتخاب بازه‌ها در داخل یک کران و در چگونگی مراحل تبدیل مقادیر صفات کمی به صورت سطوح کیفی (کد بندی شده) بسیار حائز اهمیت است. این تبدیل باید به صورتی باشد که تا حد ممکن اطلاعات زیادی از بین نروند.

هر  $x$  ( $x \in U$ ) را می‌توان به صورت برداری ارائه کرد که اجزای آن معیارهای ارزیابی شده مجزا برای نمونه  $x$  با توجه به هر یک از صفات تعریف شده  $Q$  هستند و به آن توصیف  $x$  (Description of  $x$ ) در قالب مقادیر مجموعه صفات  $Q$  می‌گویند. به جدولی که در آن توصیف هر یک از  $x$  ها ( $x \in U$ ) به وسیله صفات مجموعه  $Q$  ارائه شده باشد، جدول اطلاعات گفته می‌شود. این جدول در واقع

reducts) را هسته  $Y$  ( $Y$ -core) از صفات  $P$  تعریف می‌کنند و به صورت  $CORE_Y(P) = \cap RED_Y(P)$  قابل تعریف است. بدیهی است که صفات موجود در هسته مهم‌ترین صفات تعریف شده در جدول اطلاعات از مجموعه صفات موجود  $P$  هستند (و به عبارتی مناسب‌ترین صفات موجود برای انجام یک دسته بندی صحیح از نمونه‌ها یا موضوعات  $U$  می‌باشند).

به بیان دیگر جهت تحلیل و آنالیز یک جدول اطلاعات، کاربرد یکی از ترکیبات ساده شده صفات  $R$  ( $R \subseteq Q$ ) کفایت می‌کند. لذا دسته بندی  $Y$  در نمونه‌های  $U$  تنها با کمک مجموعه صفات  $R$  بدون آنکه هیچ اطلاعاتی حذف شود، امکان‌پذیر است و نیاز به سایر صفات تعریف شده در سیستم اطلاعات ( $Q - R$ ) ندارد. از سوی دیگر این امکان وجود دارد که با چشم پوشی و حذف هر یک از صفات تعریف شده (که به هسته تعلق ندارند) تغییری در سطح کیفیت طبقه بندی مشاهده نگردد اما اگر تنها یکی از صفات موجود در هسته حذف شود آنگاه دیگر امکان دستیابی به بالاترین سطح کیفیت تقریب با باقی صفات موجود امکان پذیر نیست.

در نتیجه با توجه به مباحث مطرح شده، نظریه مینا یک روش طبقه بندی جهت تفسیر اطلاعات است که مهم‌ترین نتایج حاصل از کاربرد و تحلیل آن موارد زیر هستند:

- ۱- ارزیابی وابستگی و ارتباط بین صفات وضعیت.
- ۲- دست یابی به کوچکترین یا کوتاهترین زیر مجموعه ممکن از متغیرها (صفات وضعیت) به قسمی که از سطح کیفیتی معادل با کل صفات برخوردار باشند (شناسایی کوتاهترین الگوریتم‌های تصمیم گیری).
- ۳- تعیین هسته. فصل مشترک الگوریتم‌های مختلف تصمیم گیری (یا مجموعه‌های مختلف ساده شده از صفات)، مجموعه صفاتی را تحت عنوان هسته (Core) تشکیل می‌دهند که در نبود هر یک از آنها دست یابی به بالاترین سطح کیفیت در تفسیر و ارزیابی اطلاعات و صفات امکان پذیر نیست.
- ۴- شناسایی و حذف صفات اضافی و غیرمؤثر موجود [15].

## ۵. کاربرد نظریه مینا در شناسایی پارامترهای تفکیک

### سفر

اولین گام جهت تحلیل اطلاعات به کمک تئوری Rough-Set و شناخت مهم‌ترین پارامترهای مؤثر در تفکیک سفر، تهیه جدول اطلاعات ورودی نظریه مینا (سیستم اطلاعات) است.

### ۵-۱. جدول اطلاعات ورودی نظریه مینا یا سیستم

#### اطلاعات

سیستم اطلاعات به صورت یک جدول اطلاعاتی است. هر یک از ردیف‌های این جدول معرف یکی از نمونه‌های مورد بررسی است و هر یک از ستونهای آن معرف یکی از صفات یا مشخصات ارزیابی

نظر داد که آیا به  $X$  تعلق دارند یا خیر. بدین ترتیب گروه‌های (دسته بندی‌های) متمایز حاصل از ارتباط  $I_p$ ، به عنوان یک ابزار اصلی در تئوری Rough-Set جهت شناخت و ارزیابی بهتر اطلاعات کاربرد دارند.

یک  $P$ -rough set شامل تمامی زیر مجموعه‌هایی از  $U$  است که تقریب‌های بالا و پایین  $P$  یکسانی دارند. به منظور رسیدن به هدف، تقریب یک مجموعه  $X$  ( $X \subseteq U$ ) با کمک دو گروه مجموعه وابسته انجام می‌گیرد. این گروهها تقریب پایین  $X$  ( $P_L X$ ) و تقریب بالای  $X$  ( $P_U X$ ) هستند. به این صورت می‌توان آنها را به عنوان یک حالت مشخص از مجموعه حد فاصل در نظر گرفت. بدین ترتیب تنها در حالتی که رابطه  $P_U X = P_L X$  برقرار باشد، با اطمینان می‌توان گفت که هر عضو  $x$  ( $x \in U$ ) با توجه به مجموعه صفات  $P$  به  $X$  ( $X \subseteq U$ ) تعلق دارد، در این حالت  $X$  را با توجه به صفات  $P$  قابل تعریف ( $P$ -definable) اطلاق می‌کنند. علاوه بر این با استفاده از معادله زیر می‌توان دقت تقریب  $X$  را بدست آورد (که حداکثر مقدار آن یک می‌باشد):

$$\frac{\text{card}(P_L X)}{\text{card}(P_U X)} \quad (1)$$

فرض کنید که  $Y = (Y_1, Y_2, \dots, Y_n)$  یک طبقه بندی مشخص از  $U$  باشد. لذا با توجه به طبقه بندی  $Y$ ، تقریب بالا و پایین  $P$  در  $Y$  مجموعه‌هایی هستند که به ترتیب اعضای آنها تقریب‌های بالا و پایین هر یک از گروه‌های تعریف شده ( $Y_i$ ) بوده و به صورت  $P_L Y = (P_L Y_1, P_L Y_2, \dots, P_L Y_n)$  و  $P_U Y = (P_U Y_1, P_U Y_2, \dots, P_U Y_n)$  قابل ارائه می‌باشند. کیفیت تقریب دسته بندی  $Y$  با توجه به مجموعه صفات  $P$  را با نماد  $\gamma_P(Y)$  معرفی می‌کنند. از نسبت کل تعداد نمونه‌های درست دسته بندی شده  $P$  (که متعلق به تقریبهای پایین  $Y_i$  ها،  $i=1, 2, \dots, n$  هستند) به کل تعداد اعضای در نظر گرفته شده بدست می‌آید. به این نسبت کیفیت طبقه بندی گفته می‌شود و زمانی به حداکثر مقدار خود می‌رسد که هر گروه  $Y_i$  در  $Y$  با صفات  $P$  قابل تعریف باشد.

مفهوم اساسی دیگر در این تئوری کاهش یا ساده سازی (Reduction) صفات است (در این روند برای یک دسته بندی  $Y$  از نمونه‌ها یا موضوعات  $U$ ، هدف یافتن کوتاهترین یا کوچکترین مجموعه ممکن ( $R$ ) از صفات مستقل بوده به طوری که از کیفیتی معادل با کیفیت طبقه بندی صفات اولیه و اصلی  $P$  برخوردار باشد). به کوچکترین زیر مجموعه ممکن  $R$  ( $R \subseteq P \subseteq Q$ ) که رابطه  $\gamma_R(Y) = \gamma_P(Y)$  برای آن برقرار باشد، ساده شده  $Y$  از صفات  $P$  ( $Y$ -reduct of  $P$ ) گویند و با نماد  $RED_Y(P)$  نمایش می‌دهند. (باید توجه داشت که یک جدول ساده اطلاعات ممکن است بیش از یک ترکیب از صفات کاهش یافته را دارا باشد). فصل مشترک تمامی ترکیبات کاهش یافته از صفات  $Y$  ( $Y$ -

با توجه به این طبقه بندی ها و سطوح یک جدول تصمیم گیری (یا سیستم اطلاعات) بدست می آید. در این جدول تمامی نمونه‌ها در قالب گروه های طبقه بندی شده مجزا (Distinct Categories) با توجه به صفت وابسته ارائه شده اند [15]. چگونگی تعریف بازه‌ها برای هر یک از صفات مورد بررسی در این پژوهش در جداول (۱-۱) و (۱-۲) مشاهده می‌شود. همچنین قسمتی از سیستم اطلاعات ورودی در جدول (۲) ارائه شده است.

## ۲-۵. نتایج تحلیل نظریه مبنا

پس از تهیه سیستم اطلاعات ورودی، این اطلاعات توسط نظریه مبنا مورد ارزیابی و تحلیل قرار گرفتند. نتایج حاصل از این تحلیل شامل چهار بخش عمده زیر است:

### الف) مجموعه های ساده شده (Reducts)

این مجموعه‌ها کوتاهترین ترکیبات ممکن از صفات مستقل هستند که دقیقاً می‌توانند تغییرات به وقوع پیوسته در پارامتر وابسته یا تصمیم گیری را بدون نیاز به سایر پارامترهای تعریف شده در سیستم ارزیابی کنند [15]. پارامتر وابسته در مبحث تفکیک سفر، نوع سیستم انتخابی فرد در طول سفر برای رسیدن به مقصد است. بنابراین الگوریتم های بدست آمده از این تحلیل از میزان دقتی معادل با کل صفات تعریف شده در سیستم برخوردار هستند. به بیان دیگر با کمک آنها میتوان تغییرات مشاهده شده در نوع سیستم انتخابی افراد را تفسیر و ارزیابی کرد و نیاز به در نظر گرفتن سایر پارامترها و صفات مستقل تعریف شده در سیستم اطلاعات نیست.

الگوریتم‌های تصمیم گیری حاصل از تحلیل نظریه مبنا در جدول (۳) ارائه شده اند. با توجه به نتایج تحلیل نظریه مبنا، پنج الگوریتم مختلف از ترکیب صفات برای ارزیابی نوع سیستم انتخابی فرد (یا تحلیل تابع سود او) وجود دارند. این الگوریتم‌ها کوتاهترین الگوریتم های حاصل از تحلیل نظریه مبنا می باشند. در این جدول اولین الگوریتم مؤید این مطلب است که تفاوت‌های موجود در سیستم انتخابی افراد را می‌توان با ترکیب ویژگی‌های درآمدی، زمان اضافی سفر، داشتن گواهینامه رانندگی، هزینه سفر، مسافت سفر، جنسیت فرد و نحوه دسترسی او به پارکینگ تحلیل نمود. همچنین دومین الگوریتم برای ارزیابی و تحلیل علل اختلاف در سیستم های انتخابی افراد، ترکیب ویژگی‌های درآمدی، زمان اضافی سفر، داشتن گواهینامه، هزینه سفر، مسافت، جنسیت و راستای سفر را پیشنهاد می‌کند. نتیجتاً می‌توان از هر یک از پنج الگوریتم بدست آمده برای ساخت مدل‌های تفکیک سفر استفاده کرد و این الگوریتم ها از لحاظ سطح دقت و کیفیت تقریب مزیتی بر یکدیگر ندارد.

شده هستند. در ستون آخر این جدول نیز پارامتر وابسته قرار می‌گیرد. بدین ترتیب هر ردیف در بر گیرنده مشخصات یکی از موضوعات مورد تحلیل می‌باشد [20]. بنابراین برای تهیه این جدول باید مراحل زیر را انجام داد:

- ۱- تهیه اطلاعات در خصوص موضوع مورد بررسی از طریق مشاهدات قبلی، جمع آوری نمونه ها و آمارگیری.
- ۲- طبقه بندی اطلاعات بدست آمده از مرحله اول [15].

## ۱-۱-۵. تهیه اطلاعات (Information Survey)

در این پژوهش، برای جمع آوری اطلاعات از روش آمارگیری در سطح شهر رشت استفاده شده است. سپس اطلاعات جمع آوری شده به صورت یک جدول در آمدند. این جدول در قالب ساختار جدول اطلاعات ورودی نظریه مبنا و به شکل طبقه بندی نشده می‌باشد. در هر یک از ردیف‌های این جدول مشخصات یک فرد، سفر او و سیستم انتخابی او درج شده است. ستونها نیز هر یک معرف یکی از انواع مشخصات و صفات مورد تحقیق و ستون آخر بیانگر نوع وسیله نقلیه یا سیستم حمل و نقل انتخابی فرد در طول سفر است.

## ۲-۱-۵. طبقه بندی اطلاعات

از نظریه مبنا به راحتی می‌توان جهت تحلیل داده های کمی بهره برد به شرط اینکه این اطلاعات ابتدا به صورت طبقه بندی شده (کد بندی شده) ارائه گردند. برای این منظور عموماً مقادیر پارامترها (صفات) را با تعریف تعدادی بازه (کران) به چندین سطح مختلف تقسیم می‌کنند. سپس به هر یک از این سطوح تعریف شده یک کد عددی اختصاص می‌دهند که معرف گروه یا سطح شناسایی شده از نمونه مورد بررسی با توجه به آن مشخصه تعریف شده برای نمونه است (این موضوع در خصوص صفات کیفی نیز صادق است). به بیان دیگر این کد مؤید این مطلب است که آن مشخصه تعریف شده برای فرد یا سفر او در کدام گروه طبقه بندی شده آن مشخصه قرار می‌گیرد. از دید محققان تعریف چنین بازه هایی نیازمند درک دقیق تر و جامع تر آنان از پدیده مورد پژوهش است. هر چند باید توجه داشت که تعریف این طبقه بندی ها برای صفات و پارامترها سبب پیدایش دو عیب عمده می‌شود:

- ۱- با تعریف این بازه ها گاه مقداری از اطلاعات مفید موجود حذف می‌شوند.
  - ۲- این بازه ها در اکثر مواقع به صورت ذهنی انتخاب می‌شوند و مبانی دقیقی برای تعریف آنها وجود ندارند.
- به طور کلی می‌توان بررسی های دقیق تری را برای تعریف بازه ها انجام داد، به صورتی که معرف بهتری در ارزیابی مشخصه مربوطه باشند. برای این منظور می‌توان مقادیر با عملکرد مشابه و یکسان را در یک بازه معین تعریف کرد.

جدول ۱ - ۱. طبقه بندی‌های تعریف شده برای صفات وابسته و مستقل

صفت تصمیم‌گیری یا وابسته (Decision Attribute)		
نحوه سفر یا سیستم انتخابی فرد (Travel Mode)		
۱	پیاده	
۲	سرویس (مثل سرویس اداره یا مدرسه)	
۳	اتوبوس (خط واحد)	
۴	تاکسی	
۵	تاکسی تلفنی (آژانس)	
۶	خودروی شخصی	
صفات وضعیت یا مستقل (Condition Attributes)		
۱	مالکیت خودرو	۶ هزینه سفر (بر حسب تومان)
۱	۰ (هیچ)	۱
۲	۱	۲ $1 < \leq 60$
۳	۲ یا بیشتر	۳ $60 < \leq 150$
		۴ $150 < \leq 300$
		۵ $300 < \leq 600$
		۶ $600 < -$
۲	درآمد	۷ داشتن گواهینامه رانندگی
۱	$\leq 60$	۱ بلی
۲	$60 < \leq 90$	۲ خیر
۳	$90 < \leq 180$	
۴	$180 < \leq 300$	
۵	$300 < \leq 450$	
۶	$450 < \leq 700$	
۷	$700 < \leq 1000$	
۸	$1000 < -$	
۳	زمان سفر	۸ جنسیت فرد
۱	$\leq 5$	۱ مرد
۲	$5 < \leq 15$	۲ زن
۳	$15 < \leq 25$	
۴	$25 < \leq 40$	
۵	$40 < -$	
۴	زمان اضافی سفر (بر حسب دقیقه)	۹ ساختار منزل
۱	$\leq 5$	افراد شاغل در منزل
۲	$5 < \leq 10$	سایر افراد غیر شاغل
۳	$10 < \leq 15$	۱
۴	$15 < \leq 20$	۲ یا بیشتر
۵	$20 < -$	۱ یا کمتر
		۲ یا بیشتر
		۱ یا کمتر
		۲ یا بیشتر
۵	مسافت سفر (بر حسب کیلومتر)	۱۰ هدف از سفر
۱	$\leq 0.5$	۱ کاری، آموزشی
۲	$0.5 < \leq 2$	۲ خرید، اجتماعی یا تفریح
۳	$2 < \leq 5$	۳ غیره
۴	$5 < \leq 10$	
۵	$10 < \leq 20$	
۶	$20 < -$	

جدول ۱ - ۲. طبقه بندیهای تعریف شده برای صفات مستقل

صفات وضعیت (مستقل)	
۱۱ تراکم منطقه مسکونی	۱۴ زمان روز
۱ در شهر (تراکم بالا)	۱ ساعات شلوغ روز
۲ در حومه (تراکم متوسط)	۲ ساعات غیر شلوغ
۳ خارج از حومه (تراکم کم)	
۱۲ راستای سفر	۱۵ فاصله تا مرکز خرید شهر
۱ به سمت مرکز خرید شهر	۱ $1 \leq$
۲ غیره	۲ $1 < 2 \leq$
	۳ $2 < 5 \leq$
	۴ $5 < 10 \leq$
۱۳ نحوه دسترسی به پارکینگ	
۱ پارک در محلی بدون هزینه یا عدم نیاز به محل پارک	
۲ پارک خودرو در محلی با پرداخت هزینه	

جدول ۲. اطلاعات طبقه بندی شده (Coded Table for Rough Set Analysis)

نمونه	صفات وضعیت (Condition Attributes)															صفت وابسته
	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	سیستم انتخابی
۱	۲	۵	۳	۱	۵	۴	۱	۲	۳	۱	۱	۲	۱	۱	۱	۶
۲	۳	۶	۴	۴	۵	۱	۱	۲	۴	۱	۱	۲	۱	۱	۳	۲
۳	۱	۲	۳	۱	۵	۱	۱	۱	۱	۱	۱	۲	۱	۱	۱	۲
۴	۳	۶	۴	۴	۴	۳	۱	۱	۴	۱	۱	۱	۱	۱	۳	۴
۵	۲	۵	۲	۲	۲	۲	۱	۱	۴	۱	۱	۲	۱	۱	۳	۴
۶	۲	۵	۳	۱	۵	۱	۱	۱	۴	۱	۱	۲	۱	۱	۳	۲
۷	۱	۶	۲	۱	۲	۲	۱	۲	۴	۱	۱	۲	۱	۱	۳	۴
۸	۱	۶	۳	۳	۴	۳	۲	۲	۴	۲	۱	۱	۱	۱	۳	۴
۹	۱	۵	۴	۳	۳	۱	۲	۲	۶	۱	۱	۲	۱	۱	۱	۲
۱۰	۱	۵	۳	۲	۳	۲	۱	۱	۶	۱	۱	۱	۱	۱	۱	۴
۱۱	۱	۵	۳	۲	۳	۲	۲	۱	۶	۱	۱	۲	۱	۱	۱	۴
۱۲	۱	۵	۳	۱	۲	۳	۲	۲	۶	۲	۱	۲	۱	۱	۱	۴
۱۳	۱	۵	۵	۳	۳	۳	۲	۲	۶	۱	۱	۱	۱	۱	۲	۴
۱۴	۱	۵	۳	۲	۲	۲	۲	۲	۶	۲	۱	۱	۱	۱	۲	۳
۱۵	۱	۴	۲	۲	۲	۲	۱	۱	۲	۱	۲	۲	۱	۱	۳	۴
۱۶	۱	۴	۴	۲	۵	۱	۱	۱	۲	۱	۲	۲	۱	۱	۳	۲
۱۷	۱	۴	۲	۲	۱	۱	۱	۱	۲	۲	۲	۲	۱	۲	۳	۱
۱۸	۱	۴	۲	۲	۱	۱	۱	۱	۲	۲	۲	۲	۱	۱	۳	۱
۱۹	۱	۱	۲	۲	۱	۱	۱	۱	۱	۲	۲	۲	۱	۱	۳	۱
۲۰	۱	۱	۲	۲	۱	۱	۱	۱	۱	۲	۲	۲	۱	۱	۳	۱
۲۱	۱	۴	۲	۱	۲	۵	۲	۲	۲	۲	۱	۲	۱	۱	۲	۵
۲۲	۱	۴	۳	۱	۲	۵	۲	۲	۲	۳	۱	۱	۱	۱	۲	۵
۲۳	۲	۷	۳	۱	۶	۴	۱	۲	۱	۱	۱	۲	۱	۱	۳	۶
۲۴	۱	۳	۲	۱	۳	۶	۲	۲	۲	۳	۱	۱	۱	۱	۱	۵
۲۵	۱	۱	۲	۲	۲	۲	۱	۱	۱	۲	۱	۱	۱	۱	۲	۳
۲۶	۲	۳	۳	۱	۳	۳	۱	۲	۲	۳	۱	۲	۱	۲	۳	۶



جدول ۳. مجموعه های ساده شده و هسته (Reducts and Cores)

مجموعه های ساده شده (Reducts)	
{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، مسافت، جنسیت، هزینه پارکینگ }	الگوریتم شماره ۱
{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، مسافت، جنسیت، راستای سفر }	الگوریتم شماره ۲
{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، مسافت، جنسیت، هدف از سفر }	الگوریتم شماره ۳
{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، زمان سفر، جنسیت، هدف از سفر }	الگوریتم شماره ۴
{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، مسافت، زمان روز، هدف از سفر }	الگوریتم شماره ۵
هسته (Core)	
{ درآمد، زمان اضافی سفر، داشتن گواهینامه رانندگی، هزینه سفر }	

جدول ۴. دقت و کیفیت طبقه بندی ها با توجه به وسیله انتخابی افراد

دقت	گروه وسیله نقلیه یا سیستم انتخابی فرد
۱	۱
۱	۲
۱	۳
۰.۹۶۷	۴
۱	۵
۰.۹۵۷	۶
دقت تقریب (Accuracy of Approximation) : ۰.۹۷۶	
کیفیت تقریب (Quality of Approximation) : ۰.۹۸۷	

(ب) هسته (Core)

صفات هسته، مجموعه صفاتی هستند که در تمامی الگوریتم های ساده شده از پارامترهای مستقل حضور دارند و در تمامی تئوریها مطرح می باشند [15]. لذا بدون در نظر گرفتن این پارامترها، امکان ارزیابی دقیقتر پارامتر وابسته وجود ندارد. به عبارتی این پارامترها شدیداً بر نوع انتخاب فرد تأثیر گذارند. در این پژوهش پارامترهای درآمد، داشتن گواهینامه رانندگی، زمان اضافی سفر و هزینه سفر صفات هسته را تشکیل می دهند.

(ج) دقت سطوح (Accuracy of Classes)

میزان دقت هر یک از سطوح صفت وابسته را می توان از تقسیم تقریب پایین بر تقریب بالای آن طبقه یا سطح تعریف شده بدست آورد [15]. همچنین می توان دقت و کیفیت طبقه بندی ها و صفات تعریف شده در سیستم را مورد ارزیابی قرار داد که میزان آن تا حدی بستگی به نحوه تعریف بازه ها و گروه بندیهای تعریف شده آنها دارد. نتایج بدست آمده در خصوص دقت سطوح پارامتر تصمیم گیری (نوع انتخاب) در جدول (۴) ارائه شده است. همانطور که مشاهده می شود در این جدول میزان دقت و کیفیت طبقه بندیها از مقادیر بالایی برخوردار می باشند و این بدین معنا است که با توجه به صفات و مشخصات تعریف شده در سیستم، با دقت بالایی

میتوان تغییرات مشاهده شده در نوع سیستم حمل و نقل انتخابی افراد را تفسیر و تحلیل نمود.

(د) قوانین یا قواعد (Rules)

قوانین در واقع روابط قطعی و دقیق (Exact) یا گاه نامطمئن (Approximate) بین صفات توصیفی با صفت یا پارامتر وابسته را بیان می کنند. قواعد عموماً به صورت عبارت « اگر... آنگاه... » بیان می شوند. یک قاعده قطعی همواره مؤید این مطلب است که هر یک از سطوح مؤلفه تصمیم گیری (وابسته)، همواره با سطوح یکسانی از پارامترهای مستقل همراه است (وضعیت ها یا شرایط یکسان تصمیمات یکسانی را به همراه خواهند داشت). در حالی که در یک قاعده غیر قطعی یا نامطمئن، بیش از یک مقدار (سطح) از صفت تصمیم گیری (وابسته) به ازای مقادیر (سطوح) یکسانی از صفات مستقل وجود دارد (وضعیت ها یا شرایط یکسان تصمیمات متفاوتی را به همراه خواهند داشت). بنابراین تنها در صورت استفاده از قوانین قطعی می توان یقین داشت که نتیجه تصمیم گیری (یا ارزیابی) در یک گروه مشخص قرار دارد [15]. قوانین بدست آمده از تحلیل نظریه اولیه یا مبنا در جدول (۵) ارائه شده اند.

کوتاهترین و مناسبترین ترکیبات ممکن از صفات و یا به عبارت دیگر شناسایی مهم ترین صفات تعریف شده در سیستم اطلاعات که مفسر تغییرات مشاهده شده در مقادیر پارامتر وابسته باشند، از روش گام به گام (Stepwise) استفاده می کنند.

در این روش پارامترهای مختلف جهت دست یابی به بهترین ترکیب خطی در ساخت مدل مورد آزمایش قرار می گیرند به صورتی که معادله خطی بدست آمده از آنها بیشترین میزان  $R^2$  را داشته باشد.

در این روند ابتدا میزان همبستگی (با  $R^2$ ) تک تک پارامترهای مستقل در ارزیابی پارامتر وابسته مورد بررسی قرار می گیرد تا تعیین شود که کدام پارامتر مستقل بیشترین میزان همبستگی را با پارامتر وابسته دارد. سپس در مرحله بعدی این فرآیند را با افزایش هر یک از پارامترهای مستقل دیگر به پارامتر اولیه، در قالب یک معادله برازش خطی سه متغیره ( یعنی با دو پارامتر مستقل ) ادامه می دهند و در هر مرحله میزان  $R^2$  بدست آمده را محاسبه می کنند.

این روند تا زمان تعیین بهترین پارامتر ثانویه از صفات مستقل ادامه خواهد داشت، به صورتی که معادله برازش خطی حاصل از آن با پارامتر اولیه در مقایسه با باقی ترکیبات سه متغیره، بیشترین میزان  $R^2$  را داشته باشد. این فرآیند تا زمانی ادامه می یابد که دیگر با افزایش یک پارامتر مستقل دیگر به مدل، تغییرات مشاهده شده در میزان  $R^2$  اندک بوده و قابل چشم پوشی باشد. در نتیجه پارامترهای موجود در معادله خطی بدست آمده از این روش به ترتیب ( از اولین

چرخه ) مهم ترین پارامترهای تعریف شده در سیستم هستند. به عبارت دیگر از دیدگاه روش رگرسیون، این پارامترها به بهترین وجهی می توانند تغییرات مشاهده شده در سطوح پارامتر وابسته را تحلیل کرده و در ساخت مدل مورد استفاده قرار گیرند.

## ۶. تحلیل اطلاعات و شناسایی پارامترها با استفاده از

### روش برازش خطی

روش برازش خطی یک روش ریاضی جهت بیان ارتباط خطی موجود بین پارامترهای مستقل و پارامتر وابسته است. به عبارتی از این طریق می توان تغییرات مشاهده شده در پارامتر وابسته را مورد ارزیابی قرار داد. در این روش از معادلات همبستگی ریاضی استفاده می شود که انتخاب بهترین فرم این معادلات و تعیین پارامترهای آنها نیاز به تجربه و مطالعات فراوانی در خصوص موضوع مورد بررسی دارد. شکل معمول استفاده از مدل همبستگی، استفاده از مدل خطی رگرسیون ( برازش خطی ) است که به فرم زیر می باشد:

$$Y = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (2)$$

که در آن  $Y$  متغیر وابسته،  $X$  ها متغیرهای مستقل و  $a$  ها پارامترهای مدل هستند که قبل از هر چیز در مدل تعیین می شوند.

از آنجا که هدف ما در این پژوهش، شناخت مهم ترین صفات اثرگذار در نوع سیستم انتخابی افراد می باشد، پارامتر وابسته یا باید معرف نوع سیستم انتخابی بوده و یا معرف تابع سودی برای فرد باشد که بیشترین مقدار را دارا است. پارامترهای مستقل نیز باید معرف ویژگیهای رفتاری فرد، مشخصات نوع سیستم انتخابی او و سفر فرد باشند.

با این وجود یک معضل عمده در ساخت و شکل گیری معادلات رگرسیون استفاده از پارامترهایی است که به بهترین وجهی بیانگر تغییرات مشاهده شده در سیستم انتخابی افراد باشند زیرا که استفاده از تمامی پارامترها جهت ارزیابی و برآورد آن پارامتر کاری دشوار، پرحجم و غیر عملی است. همچنین در این صورت نیاز به حجم وسیعی از اطلاعات ورودی می باشد که جمع آوری آنها کاری دشوار و حتی غیر ممکن است. بنابراین جهت تعیین

## جدول ۵. قواعد (Rules Generated by the Rough Set Analysis)

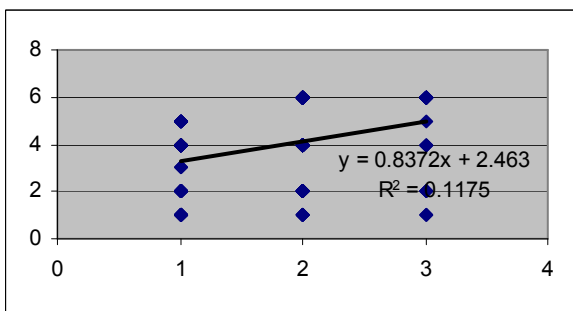
قواعد قطعی (Deterministic Rules) :
( ۶ = انتخاب ) => ( ۲ = زمان سفر )، ( ۴ = مسافت سفر )
( ۴ = انتخاب ) => ( ۳ = هزینه سفر )، ( ۲ = زمان اضافی سفر )
( ۱ = انتخاب ) => ( ۱ = هزینه )، ( ۱ = مسافت سفر )
( ۲ = انتخاب ) => ( ۱ = هزینه )، ( ۴ = زمان سفر )
( ۲ = انتخاب ) => ( ۵ = مسافت سفر )، ( ۱ = هزینه )
( ۴ = انتخاب ) => ( ۱ = راستای سفر )، ( ۳ = هزینه سفر )
( ۴ = انتخاب ) => ( ۲ = هدف )، ( ۶ = درآمد )
( ۵ = انتخاب ) => ( ۵ = هزینه )، ( ۸ = درآمد )
( ۴ = انتخاب ) => ( ۵ = زمان سفر )، ( ۴ = هزینه )
قواعد غیر قطعی (Non-Deterministic Rules) :
( ۴ = انتخاب ) یا ( ۲ = انتخاب ) => ( ۱ = داشتن گواهینامه رانندگی )، ( ۴ = ساختار )، ( ۲ = مالکیت خودرو )، ( ۷ = درآمد )

پارامترها در ارزیابی پارامتر وابسته نشان داده شده اند. همانطوریکه ملاحظه می شود به مرور از اولین چرخه میزان همبستگی هر یک از پارامترها با پارامتر وابسته کاهش می یابد. پس از پایان روش گام به گام این نتیجه حاصل گردید که پارامترهای هزینه پارکینگ، زمان روز، زمان سفر و جنسیت فرد بعد از سه پارامتر بالا، به ترتیب مهم ترین پارامترهای مستقل تعریف شده برای تحلیل نوع سیستم انتخابی فرد هستند. به عبارت دیگر از دیدگاه روش برازش خطی گام به گام، این مجموعه از صفات (هفت پارامتر فوق) که حاصل تحلیل در این روش است، بهترین الگوریتم برای ساخت مدل‌های تفکیک سفر معرفی می گردد.

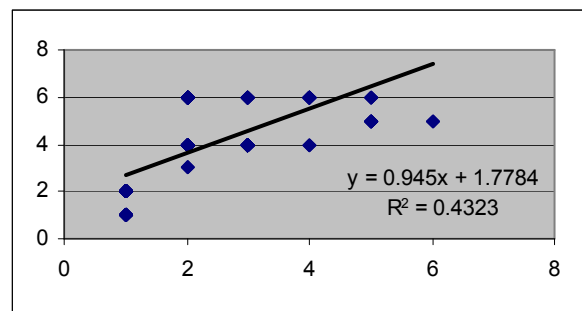
در این پژوهش با کاربرد روش گام به گام (Stepwise) به شکل فوق، تحلیل اطلاعات صورت پذیرفت و این نتیجه به دست آمد که پارامتر هزینه سفر به عنوان اولین و مهم ترین پارامتر از صفات مستقل، بیشترین میزان همبستگی را با پارامتر وابسته دارد. همچنین مشاهده شد که پارامترهای سهم مالکیت خودرو و داشتن گواهینامه رانندگی به ترتیب دومین و سومین پارامترهای مهم در سیستم اطلاعات هستند که به ترتیب معادلات سه متغیره و چهار متغیره بدست آمده از برازش خطی آنها بیشترین میزان همبستگی را (در مقایسه با باقی معادلات مشابه) در ارزیابی پارامتر وابسته دارند. جدول (۶) معادلات برازش خطی بدست آمده را نشان می دهد. همچنین در اشکال (۱)، (۲)، (۳) میزان  $R^2$  هر یک از

جدول ۶. معادلات برازش خطی گام به گام

گام	معادلات	R2	پارامترها
۱	$Y = 1.78 + 0.945X_1$	٪۴۳.۲	$X_1 =$ هزینه سفر
۲	$Y = 0.402 + 0.936X_1 + 0.808X_2$	٪۵۴.۲	$X_2 =$ تعداد خودرو
۳	$Y = 1.68 + 0.94X_1 + 0.671X_2 - 0.794X_3$	٪۵۸.۶۹	$X_3 =$ داشتن گواهینامه رانندگی
۴	$Y = -0.082 + 0.946X_1 + 0.637X_2 - 0.761X_3 + 1.73X_4$	٪۶۰.۵۳	$X_4 =$ هزینه پارکینگ
۵	$Y = -1.17 + 0.95X_1 + 0.628X_2 - 0.722X_3 + 1.8X_4 + 0.933X_5$	٪۶۱.۹۲	$X_5 =$ زمان روز
۶	$Y = -0.438 + 0.997X_1 + 0.581X_2 - 0.751X_3 + 1.71X_4 + 0.879X_5 - 0.198X_6$	٪۶۳.۱۲	$X_6 =$ زمان سفر
۷	$Y = 0.068 + 1.01X_1 + 0.576X_2 - 0.741X_3 + 1.81X_4 + 0.817X_5 - 0.204X_6 - 0.384X_7$	٪۶۴.۳۶	$X_7 =$ جنسیت



شکل ۲. نوع سیستم انتخابی بر حسب سطح مالکیت خودرو



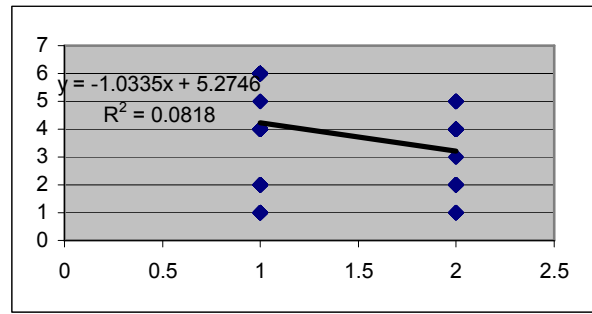
شکل ۱. نمودار نوع سیستم انتخابی بر حسب سطح هزینه سفر

هستند. بنابراین می توان نتیجه گرفت که الگوریتم های ساده شده نظریه مبنا در مقایسه با سایر الگوریتم های بررسی شده، دارای میزان همبستگی مطلوب تری در برآورد نوع انتخاب افراد هستند و از دیدگاه روش رگرسیون به شکلی مناسب تغییرات مشاهده شده در نوع سیستم انتخابی افراد را تحلیل می کنند.

در گام دوم بررسیهای انجام شده، میزان دقت و کیفیت تقریب به عنوان ارزیابی نتایج تحلیل روشهای مختلف استفاده شده است ( هر چند باید توجه داشت که میزان کیفیت تقریب، معیار اصلی مورد بررسی در تئوری Rough-Set برای شناسایی بهترین و کوتاهترین الگوریتمهای تصمیم گیری بوده و از اهمیت بیشتری برخوردار است).

این مقادیر بیانگر این مطلب است که پارامترهای مستقل تعریف شده در الگوریتم با چه میزان دقت و کیفیتی پارامتر وابسته را پیش بینی می کنند. به عبارتی معیارهای فوق، دقت و کیفیت الگوریتم بررسی شده در تحلیل ویژگیهای رفتاری افراد، مشخصات سفر و سیستم حمل و نقل انتخابی فرد را مورد ارزیابی قرار می دهند. به همین علت در این پژوهش دقت و کیفیت تقریب در الگوریتم های مختلف تصمیم گیری بررسی شده است. نتایج بررسی ها نشان می دهد که سایر الگوریتم های بررسی شده از صفات از میزان دقت و کیفیت تقریب کمتری نسبت به الگوریتم های حاصل از تحلیل نظریه مبنا برخوردار هستند. به عبارتی سایر الگوریتم های بررسی شده از صفات، ضعیف تر از الگوریتم های ساده شده نظریه مبنا در تحلیل و پردازش اطلاعات و ارزیابی نوع انتخاب فرد عمل می کنند. به عنوان مثال ملاحظه شد که الگوریتم حاصل از روش برازش خطی گام به گام و یا الگوریتم بدست آمده از صفات اصلی تعریف شده در مدل های تفکیک سفر دارای میزان دقت و کیفیت تقریب کمتری می باشند. به عنوان نمونه، میزان تفاوت مربوط به میزان دقت و کیفیت تقریب در الگوریتم های نظریه مبنا و الگوریتم برازش خطی گام به گام به ترتیب معادل ۲۸،۴٪ و ۱۶،۹٪ می باشند که مقادیر قابل توجهی هستند.

بنابراین با توجه به بررسیهای فوق می توان چنین نتیجه گرفت که هر دو مشخصه مهم مورد بررسی در این پژوهش برای الگوریتم های تصمیم گیری بدست آمده از تحلیل نظریه مبنا صادق هستند. اول اینکه معادلات برازش خطی چند متغیره حاصل از آنها دارای میزان همبستگی مطلوبی در برآورد نوع سیستم انتخابی فرد و تحلیل علل اختلاف در انتخابهای افراد ( در مقایسه با سایر الگوریتمها) هستند و دوم اینکه این الگوریتمها از بیشترین میزان دقت و کیفیت تقریب ( در مقایسه با سایر الگوریتمها) در پردازش و تفسیر اطلاعات و ارزیابی نوع انتخاب فرد برخوردار هستند. بنابراین می توان ادعا کرد که الگوریتمهای ساده شده حاصل از تحلیل نظریه مبنا بهتر از سایر الگوریتمها در تحلیل علل اختلاف در سیستم حمل و نقل انتخابی



شکل ۳. نمودار نوع سیستم انتخابی بر حسب وضعیت داشتن گواهینامه رانندگی

#### ۷. مقایسه نتایج روش نظریه مبنا و روش برازش خطی

در این پژوهش، در اولین گام برای بررسی و ارزیابی نتایج روش های نظریه مبنا و برازش خطی از معادلات رگرسیون استفاده گردید. برای این منظور میزان همبستگی معادلات رگرسیون چند متغیره حاصل از الگوریتم های مختلف تصمیم گیری در برآورد نوع سیستم انتخابی افراد محاسبه گردید و این میزان به عنوان اولین معیار در نظر گرفته شد. نتایج حاصل از این بررسیها که با استفاده از نرم افزارهای آماری روش رگرسیون انجام گردیده، در جدول (۷) درج شده است.

همانطوریکه ذکر شد، الگوریتم های بررسی شده در این پژوهش، الگوریتم های تصمیم گیری حاصل از تحلیل نظریه مبنا و الگوریتم بدست آمده از تحلیل برازش خطی گام به گام هستند. همچنین در این تحقیقات میزان همبستگی تعدادی از معادلات خطی الگوریتمهای مختلف به کار رفته در مدل های تفکیک سفر نیز مورد ارزیابی قرار گرفته اند. این الگوریتمها در واقع ترکیباتی از مهمترین صفات تعریف شده در مدل های تفکیک سفر ( به فرم های مختلفی ) می باشند. در این میان می توان به الگوریتم درآمد، هزینه سفر، تعداد خودرو، زمان سفر، زمان اضافی سفر، مسافت سفر و هدف از سفر اشاره کرد. این الگوریتمها از مهمترین الگوریتمهای بررسی شده بوده و در برگیرنده پارامترهای اصلی تعریف شده در مدل های مختلف تفکیک سفر هستند.

همانطوریکه در جدول (۷) مشاهده می شود، میزان تفاوت در  $R^2$  بدست آمده از الگوریتم های تصمیم گیری نظریه مبنا در برآورد نوع انتخاب فرد، با مقدار مشابه آن در الگوریتم حاصل از تحلیل برازش خطی گام به گام به طور متوسط ۴،۱۸٪ است که مقدار آن ناچیز بوده و به راحتی قابل چشم پوشی است. همچنین میزان  $R^2$  بدست آمده از الگوریتم های نظریه مبنا از مقادیر  $R^2$  حاصل از برای الگوریتمهای مختلف بررسی شده از صفات بیشتر است. به عنوان مثال می توان به الگوریتم درآمد، تعداد خودرو، زمان سفر، زمان اضافی سفر، هزینه سفر، مسافت سفر و هدف از سفر اشاره کرد که اکثراً شامل پارامترهای اصلی تعریف شده در مدل های مختلف

جدول ۷. مقایسه ضریب همبستگی، دقت و کیفیت الگوریتم‌های مختلف

کیفیت تقریب	دقت تقریب	R <sup>2</sup>	الگوریتم
۰.۹۸۷	۰.۹۷۶	٪۶۰.۴	* {L, H, G, F, E, D, B}
۰.۹۸۷	۰.۹۷۶	٪۶۰.۲	* {K, H, G, F, E, D, B}
۰.۹۸۷	۰.۹۷۶	٪۶۰	* {I, H, G, F, E, D, B}
۰.۹۸۷	۰.۹۷۶	٪۶۰.۲	* {I, H, G, F, D, C, B}
۰.۹۸۷	۰.۹۷۶	٪۶۰.۱	* {M, I, G, F, E, D, B}
۰.۸۱۸	۰.۶۹۲	٪۶۴.۳۶	** {M, L, H, G, F, C, A}
۰.۷۸۱۸	۰.۶۱۱۳	٪۵۱.۸	{F, D, C, B}
۰.۸۳۰۳	۰.۷۰۹۸	٪۵۵.۹	{F, D, C, B, A}
۰.۸۸۴۸	۰.۷۹۳۴	٪۵۶.۶	{F, E, D, C, B, A}
۰.۹۴۵۴	۰.۸۹۶۵	٪۵۷.۱	{I, F, E, D, C, B, A}
۰.۸۱۸	۰.۶۶۵	٪۵۶.۷	{M, L, F, D, C, B}
۰.۹۰۹	۰.۸۱	٪۵۸.۷	{M, L, F, E, D, C, B}
۰.۹۵۱۵	۰.۹۰۷۵	٪۵۹.۴	{M, L, I, F, E, D, C, B}
۰.۸۹۰۹	۰.۸۰۳۲	٪۵۹.۷	{O, M, L, F, D, C, B, A}
۰.۹۵۷	۰.۹۱۸	٪۵۹.۵	{O, M, L, I, F, E, D, C, B}
۰.۹۶۹	۰.۹۴۱	٪۶۱.۲	*** {O, M, L, I, F, E, D, C, B, A}

- \* کوتاهترین الگوریتم‌های حاصل از تحلیل نظریه مینا
- \*\* الگوریتم بدست آمده از تحلیل برازش خطی گام به گام
- \*\*\* الگوریتم مقایسه شده از ترکیب تمامی صفات متعارف بکار رفته در مدل‌های تفکیک سفر جهت ارزیابی دقیق تر نتایج
- A = مالکیت خودرو، B = درآمد، C = زمان سفر، D = زمان اضافی سفر، E = مسافت سفر، F = هزینه سفر
- G = داشتن گواهینامه رانندگی، H = جنسیت، I = هدف از سفر، J = تراکم منطقه مسکونی، K = راستای سفر،
- L = نحوه دسترسی به پارکینگ، M = زمان روز، N = فاصله منزل تا مرکز خرید شهر، O = ساختار خانواده
- میزان R<sup>2</sup> (Coefficient of Determination) بدست آمده از رگرسیون خطی تمامی پارامترها (هر ۱۵ صفت) ٪۶۷ است.

افراد عمل می‌کنند و مطمئن‌ترین الگوریتم‌ها در پردازش اطلاعات و در نتیجه مدل‌سازی های تفکیک سفر می‌باشند.

### ۸. ارزیابی نظریه مینا (مطالعه موردی شهر همدان)

برای حصول اطمینان بیشتر در استفاده از نظریه مجموعه مینا برای شهرهای ایران، کارآرایی این نظریه برای شهر همدان ارزیابی شده است. بر این اساس از جدول طبقه‌بندی تعریف شده (جدول شماره ۱-۱ و ۲-۱) برای صفات وابسته و مستقل استفاده شده است. این جداول همان جداولی هستند که برای طبقه‌بندی صفات وابسته و مستقل در شهر رشت استفاده شده‌است. در ادامه مشابه با روشی که در جدول شماره ۲ نشان داده شده است، اطلاعات موجود در رابطه با پارامترهای تفکیک سفر شهر همدان [21] طبقه بندی شده و سپس با استفاده از تئوری نظریه مینا ارزیابی شده است. جدول ۸ نتایج ارزیابی پارامترهای مؤثر در تفکیک سفر در مورد شهر همدان را در قالب الگوریتم‌های ساده شده و هسته نشان می‌دهد. با مقایسه نتایج این جدول و نتایج مربوط به ارزیابی پارامترهای مؤثر در

تفکیک سفر با نظریه مینا (جدول ۳) در شهر رشت ملاحظه می‌شود که مهمترین عوامل مؤثر در تفکیک سفر در هر دو شهر (رشت و همدان) بسیار نزدیک به هم بوده و هسته بدست آمده در این دو شهر کاملاً یکسان است. همچنین در جدول شماره ۹ مقادیر ضرایب همبستگی (R<sup>2</sup>) بدست آمده از الگوریتم‌های نظریه مینا و برازش خطی گام به گام در برآورد نوع انتخاب فرد برای شهر همدان و همچنین دقت تقریب و کیفیت تقریب آنها دیده می‌شود. همانطوری که در این جدول ملاحظه می‌شود میزان R<sup>2</sup> بدست آمده از الگوریتم نظریه مینا برای شهر همدان (مانند شهر رشت) نزدیک به مقدار R<sup>2</sup> مربوط به روش برازش خطی گام به گام است، اما از نظر دقت و کیفیت تقریب، نتایج بدست آمده توسط نظریه مجموعه مینا از دقت و کیفیت تقریب بهتری نسبت به نتایج حاصل از روش برازش خطی در تشخیص پارامترهای مؤثر در تفکیک سفر شهر همدان دارند. این نتیجه در مورد شهر رشت نیز در بخش قبل حاصل شده است و بنابراین موید کارایی بهتر روش Rough-set در مقایسه با روش برازش خطی گام به گام که یک روش متداول در تعیین پارامترهای مؤثر در تفکیک سفر است، می‌باشد.

در واقع این صفات همان مجموعه صفات مشخص شده در هسته تحلیل بوده و در پیش بینی نوع سیستم انتخابی فرد نقش عمده‌ای را ایفا می‌کنند. با این وجود پارامترهای تعریف شده در هسته، به تنهایی برای برآورد نوع انتخاب فرد کافی نیستند و باید از صفات اضافی دیگری برای پیش بینی دقیق تر آن استفاده کرد. به عنوان نمونه، اولین الگوریتم بدست آمده از تحلیل نظریه مبنا (جدول ۳) علاوه بر چهار مشخصه فوق، کاربرد سه پارامتر مسافت سفر، جنسیت فرد و هزینه پارکینگ را برای تجزیه و تحلیل یک مدل تفکیک سفر ضروری می‌داند.

با توجه به نتایج تحلیل تئوری Rough-Set، پنج الگوریتم مختلف از ترکیب صفات وجود دارند که می‌توان از هر یک از آنها در تهیه و تحلیل مدل‌های تفکیک سفر استفاده کرد و تفاوتی از لحاظ سطح دقت و کیفیت تقریب بین آنها وجود ندارد. به عنوان مثال تئوری Rough-Set علاوه بر الگوریتم فوق ترکیب ویژگیهای درآمدی، زمان اضافی سفر، هزینه سفر، داشتن گواهینامه، مسافت سفر، جنسیت

بر اساس موارد فوق نتیجه می‌شود که روش نظریه مبنا با توجه به سرعت و استفاده از حجم اطلاعات کمتر و دقت و کیفیت تقریب بهتر، می‌تواند بعنوان یک روش جایگزین برای روش برازش خطی گام به گام استفاده شود و سبب کاهش هزینه و افزایش سرعت و دقت مدل‌سازی گردد.

### ۹. نتیجه گیری

همواره محققان به دنبال روشی برای شناسایی مهم ترین صفات لازم جهت تعریف در توابع یا مدل‌های تفکیک سفر بوده‌اند. در این پژوهش از نظریه مبنا برای شناسایی کوتاهترین الگوریتم‌های تصمیم گیری استفاده شده است. برای این منظور، ۱۵ مشخصه مهم از ویژگیهای رفتاری افراد، مشخصات سفر و نوع سیستم انتخابی آنها مورد ارزیابی قرار گرفته‌اند. پس از تحلیل اطلاعات به کمک تئوری Rough-Set مشاهده شد که چهار صفت درآمد، داشتن گواهینامه، زمان اضافی سفر و هزینه سفر مهم‌ترین مشخصات تعریف شده از ویژگیهای رفتاری فرد و سفر او هستند.

### جدول ۸. مجموعه های ساده شده و هسته (Reducts and Cores) برای شهر همدان

مجموعه های ساده شده (Reducts)	
الگوریتم شماره ۱	{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، مسافت، زمان روز، هزینه پارکینگ }
الگوریتم شماره ۲	{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، مسافت، زمان روز، راستای سفر }
الگوریتم شماره ۳	{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، مسافت، جنسیت، هدف از سفر }
الگوریتم شماره ۴	{ درآمد، زمان اضافی سفر، گواهینامه رانندگی، هزینه سفر، زمان سفر، جنسیت، مالکیت خودرو }
هسته (Core)	
{ درآمد، زمان اضافی سفر، داشتن گواهینامه رانندگی، هزینه سفر }	

### جدول ۹. مقایسه ضریب همبستگی، دقت و کیفیت الگوریتم های مختلف برای شهر همدان

کیفیت تقریب	دقت تقریب	$R^2$	الگوریتم‌های حاصل از نظریه مبنا
۰.۹۸۲	۰.۹۷۱	٪۶۱.۳	{ M, L, G, F, E, D, B }
۰.۹۸۲	۰.۹۷۱	٪۶۲.۲	{ M, K, G, F, E, D, B }
۰.۹۸۲	۰.۹۷۱	٪۶۰.۴	{ I, H, G, F, E, D, B }
۰.۹۸۲	۰.۹۷۱	٪۶۰.۳	{ H, G, F, D, C, B, A }
			الگوریتم‌های حاصل از روش برازش خطی
۰.۸۳	۰.۷۱	٪۶۴.۵	{ M, L, H, G, F, C, A }
			الگوریتم‌های حاصل از کلیه صفات متعارف
۰.۹۵۸	۰.۹۳۵	٪۶۱.۸	{ O, M, L, I, F, E, D, C, B, A }
<ul style="list-style-type: none"> <li>• <math>A</math> = مالکیت خودرو، <math>B</math> = درآمد، <math>C</math> = زمان سفر، <math>D</math> = زمان اضافی سفر، <math>E</math> = مسافت سفر، <math>F</math> = هزینه سفر، <math>G</math> = داشتن گواهینامه رانندگی، <math>H</math> = جنسیت، <math>I</math> = هدف از سفر، <math>J</math> = تراکم منطقه مسکونی، <math>K</math> = راستای سفر، <math>L</math> = دسترسی به پارکینگ، <math>M</math> = زمان روز، <math>N</math> = فاصله منزل تا مرکز خرید شهر، <math>O</math> = ساختار خانواده</li> <li>• میزان <math>R^2</math> (Coefficient of Determination) بدست آمده از رگرسیون خطی تمامی پارامترها ( هر ۱۵ صفت ) ٪۶۸ است.</li> </ul>			

بررسی های بسیار وسیع و دقیق توسط یک تیم چند نفره از کارشناسان خبره و مسلط بر موضوع مورد پژوهش است.

### مراجع

- [1] Juan de Dios Ortuzar and Luis G. Willumsen, *Modeling Transport-3<sup>rd</sup> ed.*, John Wiley & Sons, 2001.
- [2] C. S. Papacostas and P.D.Prevedouros, "*Transportation Engineering and Planning-3<sup>rd</sup> ed.*", Prentice-Hall International, Inc., 2001, pp. 345-358.
- [3] Robert Lane, Timothy J. Powell and Paul Prestwood Smith, "*Analytical Transport Planning*", Gerald Duckworth, 1971. pp. 116-128.
- [4] Warner, S.L., "*Strategic Choice of Mode in Urban Travel: A study of Binary Choice*", Northwestern University Press, Evanston, 1962.
- [5] Oi, K.I.Y. and Shuldiner, P.W., *An Analysis of Urban Travel Demands*, Northwestern University Press, Evanston, 1962.
- [6] Williams, H.C.W.L., *Travel Demand Forecasting: an overview of theoretical developments* In D.J., Banister and P.G. Hall (eds), *Transport and Public Policy Planning*. Mansell, London, 1981.
- [7] Spear, B.D., "*Applications of new travel demand forecasting techniques to transportation: a study of individual choice models*", Final Report to the Office of Highway Planning, Federal Highway Administration, US Department of Transportation, Washington, D.C., 1997.
- [8] Attoh-Okine, N. "*Combining Use of Rough Set and Artificial Neural Networks in Doweled-Pavement-Performance Modeling-A Hybrid Approach*" *Journal of Transportation Engineering*, May/June, 2002.
- [9] Skowron, A., "*Extracting Laws from Decision Tables*" *Compute. Intelligence*, Vol.11, No. 2, 1995, pp 371-388.
- [10] Tsumoto, S., "*Automated Extraction of Medical Expert Systems Rules from Clinical Databases based on Rough-Set Theory*" *Information Science*, Vol. 112, 1998, pp 67-84.
- [11] An, A.J., et al., "*Discovering Rules for Water Demand Prediction: An enhanced rough set approach*" *Eng. Applic. Artif. Intell.*, Vol.9, No. 6, 1996, pp 645-653.
- [12] Attoh-Okine, N.O., "*Rough-Set Application to Data Mining Principles in Pavement Management Database*" *J. Compute. Civ. Eng.*, Vol.11, No. 4, 1997, pp 231-237.
- [13] Kryszkiewicz, M., "*Chapter 21: Properties of Incomplete Information Systems in the Framework of Rough-Sets*" *Rough set and Knowledge Discovery*, L. Polliowski and A. Skowron, eds., 1998, pp 422-450.
- [14] Kryszkiewicz, M., "*Rules in Incomplete Information Systems*" *Information Science*, 1999, pp 113, 271-292.

فرد و راستای سفر را برای برآورد پارامتر وابسته معرفی می کند. همچنین سومین الگوریتم بدست آمده از تحلیل نظریه مبنا مجموعه مشخصات درآمدی، زمان اضافی سفر، داشتن گواهینامه، هزینه سفر، مسافت سفر، جنسیت و هدف از سفر را برای ساخت مدل‌های تفکیک سفر پیشنهاد می کند. به عبارتی هر یک از پنج الگوریتم ارائه شده با ساختاری متفاوت بازگو کننده علل اختلاف در نوع سیستم حمل و نقل انتخابی افراد می باشند.

در این پژوهش نتایج تحلیل نظریه مبنا به منظور ارزیابی بهتر در تعیین مهم ترین صفات اثر گذار در سیستم انتخابی افراد با نتایج روش گام به گام مقایسه شده است. در عین حال میزان همبستگی هر یک از الگوریتم‌های تصمیم گیری نظریه مبنا در برآورد نوع سیستم انتخابی فرد مورد ارزیابی قرار گرفته است. نتایج بررسی‌ها نشان می‌دهد که تفاوت موجود در میزان  $R^2$  بدست آمده از الگوریتم های تصمیم گیری نظریه مبنا و الگوریتم حاصل از روش برازش خطی گام به گام اندک است. بنابراین می توان نتیجه گرفت که الگوریتم های حاصل از تحلیل نظریه مبنا همزمان با برخورداری از میزان دقت و کیفیت تقریب بالایی در پردازش اطلاعات و تحلیل ویژگیهای رفتاری افراد، دارای میزان همبستگی مناسبی در برآورد پارامتر وابسته هستند. در حالی که الگوریتم روش برازش خطی گام به گام از میزان دقت و کیفیت تقریب کم تری در تحلیل و پردازش اطلاعات برخوردار می‌باشد. همچنین ملاحظه شد که الگوریتم درآمد، تعداد خودرو، زمان سفر، زمان اضافی سفر، مسافت سفر، هزینه سفر و هدف از سفر که اکثراً شامل پارامترهای اصلی تعریف شده در مدل‌های مختلف تفکیک سفر هستند، علی رغم برخورداری از میزان همبستگی مناسب در ارزیابی نوع انتخاب فرد، از سطح کیفیت تقریب پایین تری در تحلیل اطلاعات برخوردار است. بدین ترتیب میتوان ادعا کرد که الگوریتم‌های ساده شده حاصل از تحلیل نظریه مبنا مطمئن ترین الگوریتم ها در پردازش و برآورد اطلاعات و در نتیجه مدل‌سازیهای تفکیک سفر می باشند. بنابراین با استفاده از الگوریتم های نظریه مبنا در مدل‌سازیهای تفکیک سفر، از حجم وسیع اطلاعات لازم که در روشهای قبلی مدل‌سازی مورد نیاز بودند، بمیزان زیادی کاسته می شود. همچنین سرعت و بازدهی پردازش اطلاعات و در عین حال دقت تصمیم گیری در حل مسئله مورد نظر به میزان قابل توجهی افزایش می یابد.

به طور کلی نظریه مبنا دارای قابلیت خارق العاده ای در شناسایی کوتاهترین الگوریتم های تصمیم گیری ممکن از کل مجموعه صفات تعریف شده در سیستم می باشد. لذا این تئوری می تواند به عنوان ابزاری قدرتمند در حل و ساده سازی انواع مسائل تصمیم گیری مورد استفاده قرار گیرد. هر چند باید توجه داشت که شناسایی جامع و دقیق الگوریتم های تصمیم گیری نیازمند انجام

ICS Research Report 38/95. Warsaw University of Technology, Warsaw, 1995.

[19] Pawlak, Z., "Rough sets: Theoretical aspects of reasoning about data" Kluwer Academic, Boston, 1991.

[20] Predki, B., Wilk, S., Rough Set Based Data Exploration Using ROSE System. In: Z.W.Ras, A.Skowron, eds. Foundations of Intelligent Systems, Lecture Notes in Artificial Intelligence, vol. 1609. Springer-Verlag, Berlin, 1999.

[21] Amani, B., "Evaluation of the use of Rough-Set theory in solution of transportation problems", M. Sc. Thesis, Faculty of Higher Education, Azad University of South Tehran, Iran, 2004.

[15] Nijkamp, P. and Pepping G., "Meta-Analysis for Explaining the Variance in Public Transport Demand Elasticities in Europe", Journal of Transportation and Statistics, January 1998, vol-1, no.-1.

[16] Van den Bergh, J.C.J.M., K. Button, P. Nijkamp, and G. Pepping. "Meta-Analysis of Environmental Policies" Dordrecht, Netherlands: Kluwer, 1997.

[17] Slowinski, R. and J. Stefanowski., "Handling Various Types of Uncertainty in the Rough Set Approach" Rough Sets, Fuzzy Sets and Knowledge Discovery. Edited by W.P. Ziarko. Berlin, Germany: Springer, 1994.

[18] Greco, S., B. Matarazzo, and R. Slowinski, "Rough Set Approach to Multi-Attribute Choice and Ranking Problems",