پیش‌بینی سفر با هدف خرید (برخط و برون‌خط) با بهره‌گیری از رویکرد تلفیقی گرگ خاکستری و یادگیری عمیق (نمونه موردی: شهر تهران)

مقاله علمی – پژوهشی

محمدحنیف دسومی، دانشجوی دکترا، دانشکده عمران، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
علی نادران *، استادیار، دانشکده عمران، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
توفیق الهویرنلو، استاد، دانشکده علوم و فناوری‌های همگرا، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
*پست الکترونیکی نویسنده مسئول: naderan@srbiau.ac.ir

دریافت: ۱۴۰۲/۰۹/۲۰– پذیرش: ۱۴۰۳/۰۱/۲۵
صفحه ۹۲-۷۷

چکیده

یکی از مسائل مهم در زمینه تجارت الکترونیک، سیستم حمل و نقل، کنترل آلایندگی و مباحث پایداری، ارزیابی نوع سفر به منظور خرید برخط و برون‌خط است. بر این اساس، در این پژوهش تلاش شد تا با جمع‌آوری اطلاعات خرید ۱۰۰۰ نفر از ساکنین مناطق ۲ و ۵ تهران در سال ۱۳۹۹، از یک رویکرد تلفیقی مبتنی بر یادگیری ماشین به منظور تخمین نوع سفر با هدف خرید استفاده شود. رویکرد پیشنهادی شامل الگوریتم تلفیقی گرگ خاکستری و شبکه کانولوشنال عمیق می‌باشد. کاربرد الگوریتم بهینه‌ساز گرگ خاکستری از آن جهت حائز اهمیت است که می‌تواند علاوه بر تنظیم هایپرپارامترهای شبکه کانولوشن، به صورت همزمان عملیات انتخاب ویژگی را نیز انجام دهد. نتایج نشان داد که روش پیشنهادی با انتخاب ۷ ویژگی از ۱۰ ویژگی اولیه توانسته به دقت تخمین برابر با ۹۷/۸۱ درصد با معیار MSE برابر با ۰/۳۲۵ برسد. همچنین مقایسه عملکرد روش پیشنهادی با روش‌های دیگر نشان داد که پس از الگوریتم پیشنهادی با دقت ۹۷/۸۱ درصد، دقت مدل‌های یادگیری عمیق تکی، شبکه عصبی MLP، درخت تصمیم و KNN به ترتیب برابر با ۹۵/۶۳ ، ۹۰/۱۲ ، ۸۶/۴۹ و ۸۰/۱۶ درصد می‌باشد. نتایج این پژوهش دید وسیع و شفافی به برنامه‌ریزان حمل‌ونقل، مدیران کسب‌وکارهای برخط و مدیران قانون‌گذار در توسعه حمل‌ونقل شهری خواهد داد. این نتایج باعث طراحی مؤثر استراتژی‌ها، کاهش هزینه‌های حمل‌ونقل، کاهش آلایندگی هوا و رضایت شهروندان را در برخواهد داشت که همگی منتج به توسعه پایدار خواهد شد.

واژه‌های کلیدی: الگوریتم گرگ خاکستری، توسعه پایدار، سفرهای با هدف خرید برخط، سفرهای با هدف خرید برون‌خط، یادگیری عمیق

۱- مقدمه

آمار بسیاری نشان می‌دهد که پس از سفرهای با هدف کار، سفرهای با هدف خرید بیشترین تقاضا را در بین افراد دارد. از این رو کاهش، جایگزینی یا اصلاح این گونه سفرها با استفاده از سیاست‌ها و برنامه‌ریزی‌های مناسب می‌تواند بر سیستم حمل و نقل تأثیر مثبتی داشته باشد (Li et all. 2021). موضوع تقاضا و تولید سفر و همچنین تأثیر آن بر سیستم حمل‌ونقل شهری از اهمیت ویژه‌ای در برنامه‌ریزی حمل‌ونقل برخوردار است به ویژه در کشوری مانند ایران که در بسیاری از شهرها و استان‌هایش از تراکم ترافیک رنج می‌برد. بررسی‌ها نشان می‌دهد که پیشرفت و گسترش زیرساخت‌های برخط، مسیر زندگی انسان را به کلی دگرگون ساخته و نفوذ گسترده‌ای در جامعه جهانی از منظر مسائل کاری و تجاری تا نحوه خرید و تفریح انسان‌ها داشته است. فرآیند خرید برخط می‌تواند به صورت جستجو و خرید کالاها و خدمات از طریق اینترنت، تعریف شود. از طرف دیگر، این سیستم می‌تواند تأثیر بالایی بر نحوه سفرهای با هدف خرید برای شهروندان به همراه داشته باشد (Candra et all., 2022). از آنجایی که خریداران محصولات برخط اغلب به زمان تحویل سریع نیاز دارند،

سفارشات آن ها باید بلافاصله به محض ثبت، پردازش شوند. تمام عملیات مورد نیاز برای تحویل گرفتن سفارشات به مشتریان در یک بازه زمانی کوتاه فشرده می‌شود (پردازش، برداشت سفارش و تحویل). بنابراین، آنچه معمولاً اتفاق می‌افتد این است که سفارش‌ها زمانی وارد سیستم توزیع می‌شوند که سایر سفارش‌ها (قبلی) از قبل فرآیند توزیع را شروع کرده باشند (Shao, Derudder and Witlox, 2022). از آنجا که در سفارشات برخط، مجموعه سفارش‌های گسترده‌ای بر اساس موقعیت مکانی سفارش دهندگان با استفاده از سیستم لجستیکی انجام می‌شود، مهمترین مسأله، تحویل به دورترین مکان مشتری می‌باشد که بعضاً با اعتراض مشتریان به دلیل تأخیرهای احتمالی یا عدم رسیدن محصول در زمان مشخص شده همراه می‌باشد. این مسأله نشان دهنده چالش‌های بسیار زیاد در زمینه انتخاب خرید از طریق برخط یا برون خط است. در یک دسته بندی کلی می‌توان مزایای استفاده از خرید برخط از دید مشتریان را شامل صرفه جویی در زمان، صرفه جویی در هزینه رفت و آمد، استفاده از تخفیف‌های محصولات، خریداری در هر ساعت از شبانه روز، عدم قرارگیری در صف، جلوگیری از ازدحام جمعیت در فروشگاه‌ها، دسترسی آسان به لیست محصولات مدنظر همراه با مشخصات آن، و خرید آگاهانه اشاره کرد. از طرفی، خریدهای برخط، معایبی نظیر تأخیر در ارسال، عدم لمس فیزیکی محصول، هزینه‌های ارسال، و فرآیند طولانی عودت محصول را نیز شامل می‌شوند که این مسائل سبب شده بسیاری از خریداران تمایل به خرید برون‌خط از نزدیک‌ترین فروشگاه محل زندگی خود را داشته باشند. طبیعتاً، هر کدام از این دو نوع خرید (برخط و برون‌خط)، نیازمند ارزیابی سفرهای با هدف خرید می‌باشند. در جدول ۱، تعداد سفرهای با هدف خرید برخط و برون‌خط در یک روز عادی در سال ۱۳۹۹ در شهران تهران ارائه شده است (Dasoomi, Naderan and Allahviranloo, 2023; Annual Report of Urban Traffic and Transportation Organization in Tehran City, 2020). بسته به نوع خرید برخط یا برون خط، نوع سفر با هدف خرید از اهمیت بالایی برخوردار است. سیستم لجستیک که شامل انتقال، جابجایی و دسترسی به اطلاعات موقعیت مکانی شهروندان خواهد بود و مهمترین بخش در زنجیره تأمین شرکت‌های عرضه کننده محصولات برخط می‌باشد. از طرفی در خریدهای برون‌خط، نوع سفر و استفاده از سیستم‌های حمل‌ونقل توسط شهروندان، علاوه بر مسائل اقتصادی، مسائل زیست محیطی را نیز در بر میگیرد. بر این اساس، تخمین نوع سفر می‌تواند به عنوان موضوع اصلی این پژوهش اتخاذ شود. بدیهی است، بدلیل حجم بالای داده‌های مربوط به سفرهای با هدف خرید که در هر دو نوع خرید برخط و برون‌خط وجود دارد، نیازمند اتخاذ روش‌های نوین مبتنی بر فناوری‌های هوش مصنوعی و محاسبات می‌باشد (Bansal and Srivastava, 2021). مسأله تخمین سفرهای با هدف خرید برخط و برون خط از مهمترین مسائلی است که در این پژوهش با بهره‌گیری از روش های مبتنی بر یادگیری ماشین به دنبال تحقق آن می‌باشیم. الگوریتم‌های یادگیری عمیق در بخش‌های متنوع و مختلفی از حمل‌ونقل شهری قابل استفاده هستند و این الگوریتم‌ها در مانیتور کردن ترافیک، پیش‌بینی و بررسی تصادفات و بهینه‌سازی سیستم‌های حمل‌ونقل مورد توجه کارشناسان بخش حمل‌ونقل شهری قرار گرفته‌اند (Moshayedi and et al., 2023). بر این اساس لازم است که با سنجش عوامل مؤثر بر نوع انتخاب برخط یا برون‌خط، نوع سفر پیش‌بینی شود که این مسأله علاوه بر اهمیت انجام این پژوهش، ضرورت استفاده از روش های مبتنی بر هوش مصنوعی و یادگیری ماشینی را نشان می‌دهد. از این رو، با توجه به نقش تجارت الکترونیک در سفرهای با هدف خرید، در این پژوهش تلاش می‌شود تا با بهره‌گیری از هوش مصنوعی و روش‌های مبتنی بر آن، بتوان مدلی در راستای تخمین سفرهای با هدف خرید برخط و برون‌خط ارائه داد. سؤالات پژوهش که می‌تواند در ادامه تحقیقات به صورت قابل توجه مورد بررسی قرار گیرد طبق موارد زیر خواهد بود:

-عوامل اصلی تأثیرگذار بر انتخاب نوع خرید (برخط یا برون‌خط)

-بررسی دقت پیش‌بینی نوع سفر با هدف خرید بر اساس مدل اجرا شده با مدل‌های مشابه و اعتبارسنجی آن

در جدول ۲، تعداد سفرهای با هدف خرید برون‌خط شهر تهران بر اساس مطالعات جامع حمل و نقل سال ۱۳۹۹، در مناطق ۲۲ گانه ارائه شده است در این جدول مبدأ، محل سکونت شهروندان و مقصد محل خرید است. در جدول ۳ نیز بر حسب اطلاعات جمع‌آوری شده از پلتفرم‌های برخط، سفرهای با هدف خرید برخط در مناطق ۲۲ گانه تهران ارئه شده است (Dasoomi, Naderan and Allahviranloo, 2023).

جدول ۱. تعداد سفر با هدف خرید بر خط و برون خط در یک روز (شهر تهران)

عنوان	تعداد	توزیع (درصد)
سفر با هدف خرید بر خط	۵۲۶۲۱	٤/۹۱
سفر با هدف خرید برون‌خط	۱۰۱۸۶۵۲	۹۵/۰۹
کل	۱۰۷۱۲۷۳	۱۰۰/۰۰

جدول ۲. ماتریس مبدأ – مقصد سفرهای با هدف خرید برون‌خط (شهر تهران)

مقصد/مبدأ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	جمع کل
1	61,213	2,643	6,744	4,005	358	2,505	1,108	341	27	27	195	744	78	31	26	26	10	15	9	23	37	12	80,178
2	2,714	74,796	6,250	782	10,549	6,875	1,019	128	1,070	788	1,016	1,143	42	20	28	105	114	182	60	55	432	160	108,325
3	7,944	2,394	30,826	4,005	337	3,641	1,568	248	21	24	168	553	48	18	15	16	7	10	6	12	24	9	51,896
4	8,785	2,115	7,811	78,428	440	4,842	3,482	5,763	44	41	411	2,080	704	211	145	86	16	29	22	88	64	18	115,628
5	2,434	19,451	5,048	906	78,854	5,016	1,004	144	1,523	375	809	1,299	48	24	40	108	132	463	92	83	1,899	1,243	120,994
6	706	3,397	4,679	464	431	20,454	1,605	88	51	112	1,122	1,349	43	15	11	28	16	16	11	12	35	11	34,655
7	1,595	936	3,379	3,276	181	6,265	13,566	1,440	20	31	512	2,428	770	159	76	48	10	12	9	41	22	7	34,785
8	2,480	965	2,910	10,961	211	3,571	4,040	18,979	25	29	389	1,971	1,496	253	137	73	11	18	17	73	35	9	48,654
9	290	3,890	608	134	1,854	1,420	226	25	5,395	878	664	766	19	13	30	117	541	1,382	173	66	500	44	19,036
10	468	6,802	1,112	232	1,451	3,817	474	47	1,532	7,936	3,013	1,902	42	28	57	337	812	383	173	102	276	36	31,030
11	461	2,305	1,148	279	416	5,717	718	64	254	1,408	10,952	5,885	72	56	121	758	350	139	151	126	91	12	31,482
12	501	482	825	575	105	2,323	1,418	196	39	82	1,175	10,418	811	695	632	592	52	37	49	192	29	4	21,232
13	1,248	603	1,479	4,231	130	2,444	2,253	3,587	25	33	441	3,068	6,620	1,006	254	108	14	23	23	103	29	6	27,727
14	1,649	1,004	2,092	2,887	225	3,859	2,402	1,217	63	92	1,089	11,156	3,720	11,265	2,389	672	55	82	100	555	71	11	46,653
15	2,003	1,442	2,381	3,156	372	3,895	1,829	924	127	157	1,272	9,243	904	1,917	22,113	2,651	139	256	338	2,409	207	31	57,769
16	323	694	567	280	182	1,247	292	76	83	128	894	2,301	73	110	722	11,611	192	210	958	1,716	95	15	22,769
17	439	2,134	896	220	958	1,958	359	43	863	729	1,459	1,929	38	38	113	684	7,255	1,756	988	230	313	31	23,433
18	573	2,960	1,067	282	1,661	2,009	385	52	2,097	489	894	1,574	38	37	120	431	1,530	18,552	1,512	320	1,810	156	38,548
19	404	1,150	795	254	460	1,460	285	59	258	187	828	1,666	49	59	255	3,376	700	848	8,381	951	236	37	22,697
20	595	747	809	604	255	1,448	507	162	96	84	566	2,817	148	205	1,256	2,598	112	261	480	21,070	168	26	35,012
21	266	1,539	511	120	2,018	776	154	21	372	73	178	318	11	6	16	37	40	687	35	36	13,837	387	21,438
22	810	3,589	1,577	345	5,349	1,621	359	54	289	88	259	512	19	12	31	69	42	362	59	69	2,908	6,286	24,711
جمع کل	97,901	136,036	83,512	116,427	106,798	87,165	39,053	33,661	14,272	13,792	28,307	65,122	15,790	16,177	28,585	24,532	12,150	25,726	13,646	28,330	23,119	8,551	1,018,652

جدول ۳. ماتریس مبدأ – مقصد سفرهای با هدف خرید برخط (شهر تهران)

مقصد/مبدأ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	21	22	جمع کل
1	2,120	582	307	616	451	91	59	29	27	26	7	1	3	17	45	20	1	12	16	11	24	4,463
2	1,000	1,727	472	799	1,689	368	261	112	85	75	48	3	11	53	83	35	1	30	27	17	52	6,948
3	836	799	386	588	823	148	82	32	26	24	11	1	4	15	44	15	1	11	16	9	35	3,904
4	610	744	384	1,291	1,089	170	149	97	40	31	10	3	6	23	53	22	1	14	19	14	47	4,817
5	707	1,272	420	830	2,405	284	195	103	71	52	17	3	11	40	78	31	2	40	24	17	61	6,661
6	306	931	224	482	787	391	284	76	60	73	46	3	8	39	49	26	1	17	15	12	31	3,862
7	229	606	178	418	580	239	327	66	37	34	10	2	7	43	35	19	1	11	10	9	20	2,882
8	212	424	150	486	496	146	182	82	23	20	6	2	5	24	28	13	0	8	9	7	15	2,340
9	39	148	24	47	158	49	36	8	57	55	10	1	1	20	26	10	1	10	5	4	4	710
10	118	425	68	140	400	150	134	24	100	169	39	4	4	61	58	35	2	14	13	8	12	1,977
11	117	411	68	139	385	168	134	23	86	137	48	3	5	58	65	43	1	15	17	9	15	1,947
12	64	206	39	84	199	91	113	13	40	60	15	2	3	35	47	27	1	9	11	6	6	1,071
13	79	262	60	154	227	104	115	26	29	32	5	1	4	31	22	9	1	4	5	6	8	1,180
14	96	292	60	171	305	126	166	26	69	100	14	5	4	70	101	33	2	14	15	6	10	1,682
15	82	204	41	124	233	72	110	18	61	83	19	5	2	39	202	58	2	23	44	6	7	1,434
16	39	95	18	45	109	34	33	7	30	44	11	1	1	15	83	38	1	18	33	2	3	660
17	36	103	17	43	119	36	29	6	38	63	7	1	1	15	54	21	1	22	13	2	4	627
18	66	184	31	73	258	59	55	11	84	84	12	2	2	30	106	32	2	48	22	6	7	1,172
19	35	96	14	40	96	29	28	7	26	38	8	1	1	15	88	37	1	32	40	3	3	635
20	47	124	26	67	126	38	44	9	25	40	12	2	2	28	205	53	2	33	48	3	5	939
21	62	219	42	75	279	85	58	13	29	30	5	1	3	24	19	8	0	13	6	7	13	990
22	173	291	127	269	577	72	49	27	17	12	5	1	2	9	18	9	0	5	7	5	46	1,721
جمع کل	7,072	10,144	3,155	6,980	11,790	2,948	2,640	812	1,059	1,280	364	45	91	704	1,510	595	23	402	413	169	426	**52,621**

ساختار کلی این مقاله در ادامه مطرح می‌شود. بخش دوم، مطالعات پیشین در زمینه تخمین و پیش‌بینی نوع سفر با هدف خرید را ارائه میکند. بخش سوم، تعریف مسئله و اهداف تحقیق خواهد بود. بخش چهارم، روش تحقیق و چگونگی استفاده از مدل را ارائه خواهد کرد. در بخش پنجم پیاده‌سازی، تجزیه و تحلیل نتایج به همراه مقایسه این مدل با مدل‌های مشابه و اعتبار سنجی را شاهد خواهیم بود و در بخش ششم، نتیجه‌گیری این پژوهش ارائه خواهد شد.

۲- پیشینه تحقیق

مدلسازی سفر با هدف خرید برخط و برون‌خط با استفاده از الگوریتم‌های ابتکاری و فراابتکاری موضوعی جدید به همراه پیچیدگی‌های فراوان است، سعی شده است در این بخش از مطالعاتی که به طور مستقیم در زمینه حمل‌ونقل فعالیتی انجام داده‌اند اشاره شود و البته در زمینه‌های دیگر که تأثیر غیرمستقیمی بر ایجاد سفر با هدف خرید برخط و برون‌خط شود نیز به طور خلاصه بررسی شود. یادگیری ماشین و استخراج یک مدل کلی از روی داده‌ها است (Dasoomi, Naderan and Allahviranloo, 2023). یادگیری ماشین که نمونه‌ای پیشرفته از روش‌های هوش مصنوعی می‌باشند، معمولاً سیستم‌هایی را با توانایی یادگیری و ارتقاء تجربه به طور خودکار بدون برنامه‌ریزی خاص ارائه می‌دهند و به عنوان محبوب‌ترین فناوری‌های جدید در انقلاب صنعتی چهارم شناخته می‌شوند. الگوریتم‌های یادگیری ماشین معمولاً داده‌ها را برای یادگیری الگوهای مرتبط در مورد افراد، فرآیندهای تجاری، تراکنش‌ها، رویدادها و غیره مصرف و پردازش می‌کنند (Sarker, 2021).

۲-۱- شاو و همکاران

شاو و همکاران در مقاله‌ای به ارزیابی دسترسی فیزیکی و مجازی مبتنی بر موقعیت جغرافیایی شهروندان در سفرهای با هدف خرید پرداختند. این مقاله یک مطالعه ترکیبی از هر دو موضوع را با تمرکز بر مورد چین ارائه می‌کند. با استفاده از شاخص خرید برخط علی‌بابا، از یک مدل فضایی عمومی برای بررسی نقش دسترسی مجازی و واقعی در ۲۷۶ شهر در چین استفاده شده است. نتایج نشان می‌دهد که هم دسترسی واقعی (تعداد نسبی مراکز خرید و تعداد نسبی اتوبوس‌ها) و هم دسترسی مجازی (بر اساس درصد مشترکان پهنای باند و تعداد نسبی نقاط تحویل) باعث افزایش خرید برخط می‌شود (Shao, Derudder and Witlox, 2022).

۲-۲- دونگ و همکاران

دونگ و همکاران در مقاله خود با استفاده از رویکردهای یادگیری ماشینی یکپارچه به پیش‌بینی رفتار شهروندان در سفر با هدف خرید پرداختند. در این مقاله، مجموعه داده‌ای حاوی داده‌های رفتاری اکتبر و نوامبر ۲۰۱۹ از یک فروشگاه برخط بزرگ چند منظوره و همچنین الگوریتم‌های متنوع یادگیری ماشین در پایتون برای پیش‌بینی دقیق رفتار شهروندان استفاده شده است. این مقاله با استخراج ۵ مجموعه داده شامل ۱۰۰۰۰ مشاهده از یک میلیارد مشاهده و به کارگیری مفاهیم Label Encoder، توانست مدل‌ها را بسازد و داده‌های این مقاله را تحلیل کند. در نتیجه، این مقاله نشان داد که الگوریتم‌های پایپلاین و جنگل تصادفی بهترین عملکرد را دارند که هر دوی آنها دقت پیش‌بینی ۹۶٪ را انجام می‌دهند که به طور قابل توجهی بیشتر از سایر الگوریتم‌ها است (Dong, Tang and Zhang, 2022).

۲-۳- ژیونگ

ژیونگ در مقاله خود به ارزیابی تأثیر هوش مصنوعی و رفتار سفر با هدف خرید برخط شهروندان پرداخت. در این مقاله، به منظور تقویت بهتر سفرها از هوش مصنوعی استفاده می‌شود. بر این اساس با استفاده از پرسشنامه مشاهده شد که تمامی گروه‌های سنی در معرض سفر با هدف خرید برخط بوده‌اند که بیشتر آنها را جوانان تشکیل می‌دهند و همچنان تعداد سفر با هدف خرید در ماه بین ٤ تا ١١ برابر بوده است. همچنین این مطالعه میزان خرید برخط چین را بررسی کرد و نشان داد که تا دسامبر سال ۲۰۲۱، ۸۲۰ میلیون شهروند خریدبرخط در چین وجود دارد که پیش‌بینی می‌شود تا سال ۲۰۲۲ به ۹۱۰ میلیون نفر برسد (Xiong, 2022).

۲-٤- ژیاهو و هارادا

ژیاهو و هارادا در مقاله خود به پیش‌بینی عدم تمایل به خرید برخط بر اساس مدل‌های K-Means و SVM پرداختند. در این مقاله، با توجه به ویژگی‌های زمانی و متغیرهای داده

چندبعدی رفتارهای خرید شهروندان، یک مدل پیش‌بینی زیان بر اساس ترکیب تقسیم‌بندی مشتری k-means و ماشین بردار پشتیبانی (SVM) پیشنهاد شده است. این روش شهروندان را به سه دسته تقسیم می‌کند و گروه‌های اصلی مشتریان را تعیین می‌کند. ماشین بردار پشتیبان و رگرسیون برای پیش‌بینی ریزش مشتری مقایسه شدند. نتایج نشان می‌دهد که هر شاخص پیش‌بینی پس از تقسیم‌بندی مشتری به طور قابل‌توجهی بهبود یافته است، که ثابت می‌کند که تقسیم‌بندی خوشه‌بندی k-means ضروری است. دقت پیش‌بینی SVM بالاتر از پیش‌بینی رگرسیون بوده است (Xiahou and Harada, 2022).

۲-۵- ژیانگ و همکاران

ژیانگ و همکاران در مقاله خود به پیش‌بینی موقعیت مکانی شهروندان و تحلیل رفتار مبتنی بر یادگیری ماشین برای سرویس‌های مکانی بر اساس هزینه ارسال پرداختند. در این مقاله، یک روش جدید برای پیش‌بینی دقیق مغازه‌های خاصی که شهروندان در آن‌ها در مراکز خرید قرار دارند، پیشنهاد شده است. با توجه به نتایج پیش‌بینی مکان‌های مشتریان، راهکارهایی برای ارائه خدمات دقیق‌تر به مشتریان ارائه شده است. با استفاده از استخراج ویژگی و غربالگری از سوابق برخی از مشتریان واقعی در مراکز خرید ساخته شده است. برای اثبات اعتبار مدل پیشنهادی، انواع الگوریتم‌های یادگیری ماشین نیز بررسی شده است. نتایج نشان می‌دهد که روش پیشنهادی بهترین مبادله سرعت-دقت را به دست می‌آورد و می‌تواند فروشگاه‌هایی را که مشتریان در آن‌ها در مراکز خرید قرار دارند، به‌درستی مکان‌یابی کند. لازم به ذکر است که در مقایسه با سایر الگوریتم‌ها، مدل پیشنهادی دقیق‌تر است (Jiang et all., 2021).

۲-۶- لی و همکاران

لی و همکاران در مقاله خود به مقایسه و تفسیر الگوریتم‌های مختلف یادگیری ماشین در راستای پیش‌بینی تبدیل خرید برخط پرداختند. در این مقاله به بررسی سه سؤال اصلی پرداخته شده است. اول، مدل یادگیری ماشینی مناسب برای پیش‌بینی رفتار شهروندان برخط چیست؟ دوم، روش نمونه‌گیری داده مناسب برای پیش‌بینی رفتار شهروندان برخط چیست؟ سوم، آیا می‌توان نتایج پیش‌بینی رفتار شهروندان برخط یادگیری ماشین را تفسیر نمود؟ بر این اساس، ۳۷۴۷۴۹ داده رفتار مصرف‌کننده برخط از فروشگاه‌های گوگل یک مرکز خرید برخط ، تجزیه و تحلیل شده است. نتایج نشان می‌دهد که عملکرد مدل مجموعه‌ای مدل گرادیان افزایشی[1] برای پیش بینی تبدیل خرید مصرف‌کنندگان برخط مناسب‌ترین روش است و نمونه برداری بیش از حد بهترین روش برای کاهش عدم تعادل داده است. این مطالعه از لحاظ نظری با بررسی و پاسخ به مسائلی که هنگام استفاده از مدل‌های یادگیری ماشین برای پیش‌بینی تبدیل مصرف‌کننده برخط به وجود می‌آیند، به ادبیات بازاریابی و یادگیری ماشین کمک می‌کند. همچنین با بررسی ویژگی‌های مصرف‌کننده که برای هدف‌گیری مجدد تبلیغات مؤثر هستند، به ادبیات تبلیغات برخط کمک می‌کند (Li et al, 2021).

۲-۷- زبیدی همکاران

زبیدی و همکاران در مقاله‌ای به بررسی پیش بینی تقاضای سفر با استفاده از روش سیستم استنتاج عصبی-فازی تطبیقی پرداختند. این مقاله بر توسعه و ارائه بازاریابی از متغیرهای زبانی بر اساس ارزیابی تقاضا متمرکز شده است. در این مقاله از دو روش شبکه عصبی مصنوعی (ANN)، یک شبکه عصبی فازی (ENN) جهت تخمین استفاده شده است. با مقایسه نتایج پیش بینی مشخص شد که استفاده از شبکه با ساختارهای عصبی فازی با متغیرهای زبانی برای سیستم پیش بینی تقاضا بسیار مفید می‌باشد و دارای دقت بالایی است (Zubaidi et al, 2020).

با توجه به بررسی دقیق مطالعات پیشین که بخشی از آنها در این بخش ارائه گردید، میتوان بیان نمود که ایده اصلی این پژوهش، مبتنی بر بکارگیری الگوریتم یادگیری عمیق، به طور خاص شبکه عمیق کانولوشنال در طبقه‌بندی ویژگی‌ها به منظور تشخیص سفرهای با هدف خرید است. یکی از مهمترین راهکارهای حل این مسأله، بهبود و توسعه الگوریتم های یادگیری عمیق است که میتوان با تنظیم هایپرپارامترها و استفاده از الگوریتم‌های فراابتکاری در مرحله انتخاب ویژگی، از پیچیدگی محاسبات یادگیری عمیق کاست. مهمترین گام در راستای افزایش عملکرد طبقه‌بندی و کاهش هزینه‌های عملیاتی، انتخاب ویژگی کاهش ابعاد داده‌ها و انتخاب تنها

حیاتی‌ترین ویژگی‌ها است. پیدا کردن بهترین ویژگی، به خصوص در مورد فضای جستجوی وسیع به عنوان مسأله‌ای چالش برانگیز خواهد بود. یافتن کمترین و حیاتی‌ترین ویژگی‌هایی که جزئیات کافی برای توصیف سیستم را ارائه می‌دهند، هدف اصلی انتخاب ویژگی است. در میان الگوریتم‌های فراابتکاری، الگوریتم گرگ خاکستری به دلیل جستوجوی گسترده در فضای محاسباتی از اهمیت بالایی برخوردار است. در فرآیند شکار گرگ‌های خاکستری، سه دسته گرگ آلفا، بتا و دلتا فرآیند شکار (تشخیص ویژگی‌های مؤثر) را به طور همزمان انجام می‌دهند. گرگ آلفا به عنوان متقاضی بهینه برای حل بتا و دلتا در نظر گرفته می‌شود که انتظار می‌رود از موقعیت احتمالی طعمه آگاه باشد. بنابراین، سه بهترین راه‌حل که تا یک تکرار مشخص پیدا شده‌اند، حفظ می‌شوند و گرگ‌های دیگر (مثلاً امگا) را مجبور می‌کنند تا موقعیت‌های خود را در فضای تصمیم‌گیری مطابق با بهترین مکان تغییر دهند. این عمل سبب می‌شود تا انتخاب ویژگی‌های بهینه در فضای جستوجو به طور مداوم انجام شود. مکانیسم بروزرسانی گرگ‌ها تابعی از موقعیت سه بردار مختلف است که هر گرگ را به سه راه‌حل اول برتر ارتقا می‌دهد که به موجب آن، بهترین ویژگی‌ها از میان بهترین مقادیر بدست‌آمده انتخاب می‌شوند. همین مسأله نشان از کارایی بالای این الگوریتم در انتخاب ویژگی است.

۳- تعریف مسأله و اهداف تحقیق

انفجار خدمات فناوری اطلاعات با ایجاد صنایع و زمینه‌های تولید و خدمات جدید که حتی ۳۰ سال پیش قابل تصور نبود، عمیقاً اقتصاد و در نتیجه آن تقاضای سفر را دگرگون ساخته است. به طور خاص رفتار و نیازهای شهروندان، شرکت‌ها را متحول کرده است (Archetti and Bertazzi, 2021; Ardiansah, Chariri and Januarti, 2019). مشتریان در تجارت الکترونیک به عنوان نبض اصلی پویایی تجارت در بستر اینترنتی هستند. کسب و کارها نسبت به سال‌های گذشته با افزایش قابل توجهی همراه بوده و همچنان برخی از کاربران نسبت به خرید برخط اعتماد لازم را ندارند و ترجیح می‌دهند به صورت برون‌خط اقدام به خرید محصولات مورد نیاز خود کنند (Kharis, 2020). طبق گزارش اخیر لیپسمن، فروش محصولات برخط که به همراه خود سفر با هدف خرید برخط را شامل خواهد شد در سراسر جهان در سال ۲۰۲۳ به ۶/۵ تریلیون دلار خواهد رسید. پیش‌بینی می‌شود سهم بازار نسبت به کل بازار خرده‌فروشی از ۱۰/۴۰ درصد در سال ۲۰۱۷ (با مقدار فروش تقریباً ۲,۴ تریلیون دلار) تا ۲۲ درصد در سال ۲۰۲۳ رشد نماید. این مساله تأثیر بسیار زیادی بر مدیریت کسب و کار خرده فروشی دارد و با تمرکز بر جنبه مدیریت تقاضای حمل و نقل، بر نحوه تحویل به مشتریان تأثیر می‌گذارد. این مسأله ناشی از توسعه تکنولوژی که در سال‌های اخیر، روند رو به رشدی را در میان عموم مردم به همراه داشته است (Lipsman, 2019; Moon, Choe and Song, 2021). همچنان چالش اصلی، به نوع انتخاب خرید برخط[۲] یا برون خط[۳] توسط شهروندان می‌باشد. برای مثال، در کشور کره، شرکت‌های خرده‌فروشی بزرگ مانند گروه Shinsegae (یعنی E-mart، فروشگاه بزرگ Shinsegae) و Lotte Group (یعنی Lotte Mart، Lotte Department Store) که بر صنعت خرده‌فروشی در کره تسلط دارند، رقابت بالایی در زمینه خرید برخط و برون‌خط ایجاد کرده‌اند. به عنوان مثال، اگر چه Lotte Mart (فروشگاه برون‌خط) و Lotte Mart Mall (خرده فروشی برخط) هر دو از یک شرکت هستند، اما رقابت بالایی در زمینه فروش محصولات با یکدیگر دارند (Stocchi et al., 2018). هوش مصنوعی با سرعت بالایی در حال پیشرفت می‌باشد. مطالب علمی اغلب هوش مصنوعی را به عنوان ربات‌هایی با ویژگی‌های انسانی نشان می‌دهند اما حقیقت این است که هوش مصنوعی می‌تواند شامل هر چیزی شود (Jawandhiya, 2018). یادگیری ماشین(ML[۴]) نوعی از کاربردهای پیشرفته هوش مصنوعی است که توانایی بالایی در اعمال خودکار محاسبات ریاضی پیچیده در داده‌های بزرگ را دارد و می‌تواند برای پیش‌بینی سفرها با هدف خرید به کار آید (Jordan and Mitchell, 2015). مسئله یادگیری ماشین به مسئله یادگیری از طریق تجربه با توجه به ارزیابی فعالیت‌ها و عملکردها بر می‌گردد (Tsai et al., 2015). معمولاً مسائل محاسباتی علمی و مهندسی، دارای پیچیدگی‌های ذاتی و منابع فشرده هستند (Jeffers and Reinders, 2015). تعیین مجموعه بهینه‌ای از پارامترها در یک زمینه اجرایی مشخص یک کار پیچیده است، بنابراین برای حل این مسئله رویکردهای مختلفی مبتنی بر یادگیری ماشینی ارائه شده است (Memeti et al., 2019). از این رو در پژوهش حاضر

تلاش می‌شود به ارائه مدلی به منظور پیش‌بینی سفرها با هدف خرید برخط و برون‌خط با استفاده از داده‌های سری زمانی و بهره‌گیری از رویکرد یادگیری عمیق پرداخته شود.

مسأله تخمین میزان سفرهای با هدف خرید برخط و برون‌خط از مهمترین مسائلی است که در این پژوهش با بهره‌گیری از الگوریتم گرگ خاکستری و روش‌های مبتنی بر یادگیری ماشین به دنبال تحقق آن می‌باشیم. در این راستا، اهداف دیگری نیز مدنظر می‌باشد که در ادامه به معرفی آن‌ها پرداخته شده است:

-بررسی مقایسه‌ای ویژگی‌های خریدهای برخط و برون خط توسط مشتریان

-شناسایی عوامل مؤثر بر انتخاب خریدهای برخط و برون خط توسط مشتریان

-مقایسه رویکرد پیشنهادی با الگوریتم‌های پرکاربرد یادگیری ماشین نظیر شبکه عصبی MLP، درخت تصمیم، ماشین بردار پشتیبان و نزدیکترین همسایه

٤- روش تحقیق

جامعه آماری این پژوهش را ۱۰۰۰ کاربر فعال در حوزه تجارت الکترونیک ساکن تهران در مناطق ۲ و ۵ که در سال ۱۳۹۹ سفارش موفق در سرویس‌های برخط و برون‌خط داشته اند تشکیل داده‌اند و نمونه‌گیری مورد استفاده نیز به صورت هدفمند بوده است. در تمامی این پرسشنامه‌ها، از شاخص‌های سن، جنسیت، وضعیت تأهل، مالکیت خودرو، هزینه ارسال، زمان ارسال، قیمت محصول، درآمد خانوار، وضعیت اشتغال و میزان تحصیلات به عنوان شاخص های اثرگذار بر نوع سفر برای خرید استفاده شده است. بر این اساس، در گام نخست، با توجه به فراوانی داده‌ها در مناطق ۲ و ۵ کلانشهر تهران و محاسبات براساس روش کوکران در فرمول ۱ و برآورد خطای ۵ درصد، دقت ۹۵ درصدی و نسبت صفت معین ۰/۵، ۵۰۰ پرسشنامه از سفرهای برخط از طریق ارسال پیامک و ۵۰۰ پرسشنامه از سفرهای برون‌خط به صورت استقرار در مراکز خرید این دو مناطق از شهروندان اخذ شد. در شکل ۱، فلوچارت روش و گام‌های حل مساله ارائه شده است.

$$n = \frac{\frac{Z^2 pq}{d^2}}{1 + \frac{1}{N}\left[\frac{Z^2 pq}{d^2} - 1\right]} \qquad (۱)$$

n: حجم نمونه

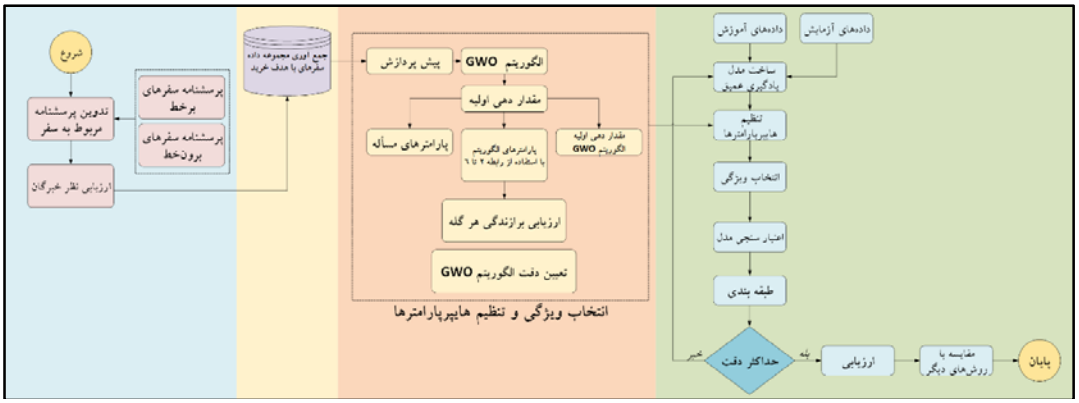
N: حجم جمعیت آماری

Z: مقدار متغیر نرمال واحد استاندارد

p: نسبتی از جمعیت داری صفت معین

q: نسبتی از جمعت فاقد صفت معین (1-p)

d: مقدار اشتباه مجاز یا درصد خطا



شکل ۱. فلوچارت روش پیشنهادی

بعد از مرحله خواندن مجموعه داده‌ها و عملیات پیش‌پردازش بر روی آنها، عملیات تنظیم هایپرپارامترها و انتخاب ویژگی‌های بهینه با استفاده از الگوریتم گرگ خاکستری انجام می‌گیرد. الگوریتم گرگ خاکستری بهینه‌سازی مراحل شکار گرگ‌ها را شبیه‌سازی می‌کنند. ساختار شکار شامل سه قسمت تعقیب و محاصره طعمه، آزار و اذیت طعمه تا توقف آن و نهایتا حمله به طعمه است. هر گرگ i به عنوان یک راه‌حل برای مسئله در فضای جستجو دارای بردار موقعیت

$W_i=<w_{i1},w_{i2},...,w_{in}>$ است که n ابعاد مسئله را نشان می‌دهد. برای ارزیابی موقعیت گرگ‌ها از تابع برازش (متناسب با تعریف مسئله) استفاده می‌شود. با توجه به مقادیر تابع برازش اولین بهترین گرگ با آلفا (α)، دومین بهترین گرگ با بتا (β) و سومین بهترین گرگ با (δ) نشان داده می‌شود. در طول فرآیند شکار (بهینه سازی)، گرگ‌ها موقعیت خود را با توجه به موقعیت سه گرگ آلفا، بتا و دلتا به روز می‌کنند (Kamel, Yaghoubzadeh and Kheirabadi, 2019; Mirjalil, Mirjalili and Lewis, 2014).

روند اجرای الگوریتم گرگ خاکستری بهینه‌ساز به شرح زیر می باشد:

ورودی:

N: تعداد گرگ‌ها (اعضای جمعیت)

T: تعداد اجرای الگوریتم

F: تابع برازش (مطابق با تعریف مسئله)

- خروجی

گرگ آلفا(α): بهترین راه حل بدست آمده

۱- جمعیتی از گرگ‌ها را به صورت تصادفی ایجاد خواهد کرد.

۲- بردارهای ضرائب A، C و a با توجه به روابط (۲) و (۳)، مقداردهی اولیه خواهد شد. بردار A دارای مقادیر تصادفی بین [a,a] می‌باشد که واگرایی[۵] را مدل می‌کند. زمانی که $|A|>1$ باشد، عامل‌های جستجو (گرگ‌ها) به دور شدن از طعمه و زمانی که $|A|<1$ برای حمله ملزوم می‌شوند. بردار C شامل مقادیر تصادفی در بازه $[۰,۲]$ است که به عامل‌ها کمک می‌کند تا از دام بهینگی[۶] محلی دوری کنند. مؤلفه های بردار a در طول تکرار[۷] اجرای الگوریتم به صورت خطی از ۲ به ۰ کاهش می‌یابند.

۳- مقدار برازش هر عامل جستجو(گرگ) محاسبه خواهد شد.

٤- در بین جمعیت، اولین بهترین راه حل به عنوان گرگ آلفا، دومین بهترین راه حل به عنوان گرگ بتا و سومین بهترین راه حل به عنوان گرگ دلتا انتخاب خواهد شد.

۵- اگر تعداد تکرارها به پایان رسیده است، گرگ آلفا مجدداً بازخواهد گشت.

٦- موقعیت گرگ‌ها را با توجه به موقعیت سه گرگ آلفا، بتا و دلتا با استفاده از روابط (۲) تا (٦) بروزرسانی خواهد شد.

۷- مقادیر برداری ضرائب A، C و a را بروز می‌شود و به مرحله ۳ بازخواهد گشت.

$$A = 2a.\overline{r_1}\ a \qquad (۲)$$

$$c = 2.\overline{r_2} \qquad (۳)$$

$$\overline{D}_a=|\overline{C}_1.\overline{X}_a\overline{X}|,\ \overline{D}_\beta=|\overline{C}_2.\overline{X}_\beta\overline{X}|\ ,\ \overline{D}_\delta=|\overline{C}_3.\overline{X}_\delta\overline{X}| \qquad (٤)$$

$$\overline{X}_1=X_aA_1.(\overline{D}_a),\overline{X}_2=X_\beta A_2.(\overline{D}_\beta),\overline{X}_3=X_\delta A_3.(\overline{D}_\delta) \qquad (۵)$$

$$\overline{X(t+1)}=\frac{\overline{X}_1+\overline{X}_2+\overline{X}_3}{3} \qquad (٦)$$

بر این اساس، در این الگوریتم، N تعداد گله گرگ و D تعداد متغیرهای تصمیم یا ابعاد مسئله بهینه‌سازی[۸] می‌باشد. لذا گله گرگ‌های خاکستری توسط یک ماتریس N×D شبیه سازی می‌شوند. هر سطر مربوط به یک راه حل ممکن از مسئله بهینه‌سازی است. در مدل پیشنهادی N تعداد رکوردهای مجموعه داده و D تعداد ویژگی‌ها می‌باشد. جمعیت گله که شامل تعداد زیادی گرگ است طبق معادله (۷) تعریف می‌شود. در مدل پیشنهادی، روش کار برای مجموعه داده بدین صورت است که الگوریتم گرگ خاکستری از N گله و هر گله نیز از ویژگی تشکیل شده است. هر گله توسط این D ویژگی تعریف می شود.

$$Population\ of\ GWO=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{12} \\ x_{22} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \qquad (۷)$$

در مجموعه x_i هر $x_i=(x_{i1},x_{i2},...,x_{id}),\ i=1,2,....,n$ نشان دهنده یک راه حل ممکن در فضای راه حل‌ها است. هر گله شامل یک گروه از گرگ‌های مهاجم است که به عنوان عناصر یک راه حل در نظر گرفته می‌شوند. تمام گرگ‌های مهاجم در یک گله به عنوان یک واحد کلی در نظر گرفته می‌شوند که به سمت یک مکان مناسب با منابع فراوان حرکت می‌کنند. اگر گله به یک موقعیت ایده‌آل برسد، راه حل بهینه‌ای بدست خواهد آمد. ارزیابی هر گله بر مبنای تابع هدف طبق معادله (۸) محاسبه می‌شود.

$$fit_i=1-\frac{Obj_i-worst(Obj)}{best(Obj)-worst(Obj)} \qquad (۸)$$

در معادله (۸)، fit_i برازندگی گله i ام است. پارامتر Obj_i مقدار تابع هدف برای گله i ام است. هر گله بر مبنای معیار فاصله محاسبه می‌شود. پارامترهای Worst و best بدترین و بهترین گله نسبت به طعمه هستند. در مدل پیشنهادی باید الگوریتم گرگ‌های خاکستری از حالت پیوسته به گسسته تبدیل شود. برای تبدیل اعداد به باینری طبق معادله (۹) استفاده

خواهد شد. در الگوریتم پیشنهادی، تابع مماس هذلولی از توابع V شکل مطابق با معادلات (۹) و (۱۰) تطبیق داده شده است. بنابراین الگوریتم پیشنهادی برای تطبیق جستجوی فضای ویژگی و ارائه بهترین ترکیب ویژگی استفاده می‌شوند (Hussien, Hussien and Hassanien, 2017).

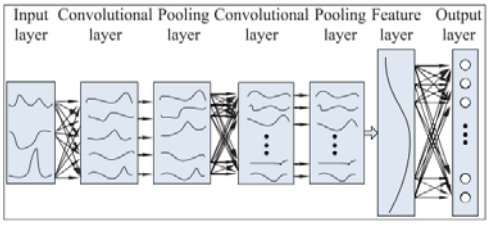
$$y^k = |tanhx^k| \quad (۹)$$

$$x_{ij}=\begin{cases}0, & \text{if rand}<y^k\\ 1, & \text{otherwise}\end{cases} \quad (۱۰)$$

در مدل پیشنهادی با استفاده از الگوریتم گرگ‌های خاکستری زیرمجموعه‌ای از ویژگی‌ها که به حالت بهینه منجر می‌شوند، انتخاب می‌گردد. تابع برازندگی برای انتخاب ویژگی از هر گله طبق معادله (۱۱) تعریف می‌شود. در معادله (۱۱)، |n| تعداد کل ویژگی‌ها و |S| تعداد ویژگی‌های انتخاب شده است. پارامتر accuracy درصد صحت و مقدار پارامترهای δ و ρ ثابت هستند و مقدار آن‌ها به ترتیب برابر با ۹۹ و ۱ می‌باشد.

$$Fitness = \delta.Accuracy + \rho.\frac{|n|-|S|}{|n|} \quad (۱۱)$$

CNN معمولا از دو بخش تشکیل شده است. در بخش اول، عملیات convolution و pooling به طور متناوب برای تولید ویژگی‌های عمیق داده‌های خام استفاده می‌شود. و در قسمت دوم، ویژگی‌ها برای طبقه‌بندی به یک طبقه بند متصل می‌شوند. برای درک بهتر، یک معماری معمولی CNN برای طبقه بندی با دو لایه همزمان و دو لایه ادغام در شکل ۲ نشان داده شده است (Zhao et al., 2017).



شکل ۲. معماری معمول یک شبکه CNN

در این پژوهش برای اعتبارسنجی پرسشنامه از روش کرونباخ طبق فرمول (۱۲) استفاده شده است (Kenedy, 2022). ۱۰۰ پرسشنامه به صورت تصادفی توسط آلفای کرونباخ در نرم افزار SPSS 22 مورد ارزیابی و اعتبار سنجی قرار گرفتند.

$$\alpha = \left(\frac{K}{K-1}\right)^2 [1-\frac{\sum S_i^2}{S_x^2}] \quad (۱۲)$$

K تعداد سؤالات مطرح شده خواهد بود و S_i واریانس نمرات برای سؤال i و S_x واریانس جمع نمرات برای تمامی سؤالات خواهد بود. آلفای کرونباخ یک شاخص برای سنجش همگونی سؤالات پرسشنامه محسوب خواهد شد. این شاخص نشان می‌دهد که سؤالات مطرح شده در پرسشنامه تا چه میزان با یکدیگر همبسته و چقدر مستقل از هم اندازه‌گیری خواهند شد. آلفای کرونباخ بین صفر و یک خواهد بود و پژوهش‌ها نشان داده است که معمولاً اگر این شاخص بین ۰/۶ تا ۰/۸ باشد، قابل اطمینان خواهد بود و اگر بالاتر از ۰/۷ و نزدیک به ۱ باشد دارای اعتبار سنجی با درجه عالی خواهد بود (Collins, 2007). نتایج آلفای کرونباخ در جدول ٤ ارائه شده است.

جدول ٤. اعتبارسنجی پرسشنامه

شاخص	آلفای کرونباخ
سن	۰/۷۱
جنسیت	۰/۷۳
وضعیت تأهل	۰/۷۲
مالکیت وسیله نقلیه	۰/۷۹
درآمد	۰/۷۸
شغل	۰/۷۹
تحصیلات	۰/۷۸
هزینه ارسال	۰/۷۳
زمان ارسال	۰/۷۶
هزینه محصول	۰/۷۷

جدول ٤ نشان داد که تمامی شاخص‌های مطرح شده در پرسشنامه دارای آلفای کرونباخ بالای ۰/۷ هستند. ضمناً پرسشنامه ارائه شده نیز در شکل ۳ قابل مشاهد است.

۱. این پرسشنامه برای کدام خرید ثبت شده است؟	خرید برخط	○
	خرید برون‌خط	○
۲. سن را انتخاب کنید.	بین ۱۸ تا ۲۵	○
	بین ۲۵ تا ۳۰	○
	بین ۳۰ تا ۳۵	○
	بالاتر از ۳۵	○
۳. جنسیت را انتخاب کنید.	زن	○
	مرد	○
٤. وضعیت تأهل را انتخاب کنید.	مجرد	○
	متأهل	○
۵. مالکیت وسیله نقلیه را انتخاب کنید.	بله	○
	خیر	○
٦. درآمد شهروند را انتخاب کنید.	کمتر از ۵ میلیون تومان	○
	بین ۵ تا ۱۰ میلیون تومان	○
	بین ۱۰ تا ۱۵ میلیون تومان	○
	بیش از ۱۵ میلیون تومان	○
۷. شغل شهروند را انتخاب کنید.	آزاد	○
	دانشجو	○
	کارمند تمام وقت	○
	کارمند پاره وقت	○
	بازنشسته	○
	فاقد شغل	○
	غیره	○
۸. تحصیلات شهروند را انتخاب کنید.	دیپلم و زیر دیپلم	○
	فوق دیپلم	○
	کارشناسی	○
	کارشناسی ارشد	○
	بالاتر از کارشناسی ارشد (پزشک و دکتری و بالاتر)	○

۹. فاصله تقریبی شهروند از نزدیک‌ترین سوپرمارکت با محل سکونت را انتخاب کنید.

کمتر از ۵۰۰ صد متر	○
بین ۵۰۰ تا ۱۵۰۰ متر	○
بین ۱۵۰۰ تا ۳۰۰۰ متر	○
بیش از ۳۰۰۰ متر	○

۱۰. شهروند گرامی، سه پارامتر روبرو را اولویت بندی کنید.

پارامتر / اولویت	زیاد	متوسط	کم
زمان ارسال	○	○	○
هزینه ارسال	○	○	○
هزینه محصول	○	○	○

در ادامه با ٦ سؤال مواجه خواهید شد، که بین دو خرید برخط و برون‌خط با شرایط مختلف می‌بایستی یک گزینه را انتخاب کنید.

۱۱. از بین دو حالت زیر کدام را انتخاب می‌کنید؟

	زمان ارسال	هزینه ارسال	مبلغ محصول	
خرید برخط	کمتر از ۱ ساعت	هر کیلومتر ۷ هزار تومان	بدون تخفیف	○
خرید برون‌خط			بدون تخفیف	○

۱۲. از بین دو حالت زیر کدام را انتخاب می‌کنید؟

	زمان ارسال	هزینه ارسال	مبلغ محصول	
خرید برخط	کمتر از ۱ ساعت	رایگان	بدون تخفیف	○
خرید برون‌خط			بدون تخفیف	○

۱۳. از بین دو حالت زیر کدام را انتخاب می‌کنید؟

	زمان ارسال	هزینه ارسال	مبلغ محصول	
خرید برخط	۱ روز بعد (زمان‌بندی شده)	رایگان	بدون تخفیف	○
خرید برون‌خط			بدون تخفیف	○

۱٤. از بین دو حالت زیر کدام را انتخاب می‌کنید؟

	زمان ارسال	هزینه ارسال	مبلغ محصول	
خرید برخط	۱ روز بعد (زمان‌بندی شده)	هر کیلومتر ۷ هزار تومان	۱۰ درصد تخفیف	○
خرید برون‌خط			بدون تخفیف	○

۱۵. از بین دو حالت زیر کدام را انتخاب می‌کنید؟

	زمان ارسال	هزینه ارسال	مبلغ محصول	
خرید برخط	کمتر از ۱ ساعت	رایگان	۱۰ درصد تخفیف با شرط حداقل خرید ۲۰۰ هزار تومانی	○
خرید برون‌خط			بدون تخفیف	○

۱٦. از بین دو حالت زیر کدام را انتخاب می‌کنید؟

	زمان ارسال	هزینه ارسال	مبلغ محصول	
خرید برخط	کمتر از ۱ ساعت	رایگان	بدون تخفیف	○
خرید برون‌خط			۱۰ درصد تخفیف	○

شکل ۳. نمونه سؤالات پرسشنامه

۵- پیاده‌سازی، تجزیه و تحلیل یافته‌ها

در گام نخست لازم است نتیجه آمار توصیفی جامعه آماری مشخص شود. این موضوع در جدول ۵ ارائه شده است. سپس داده‌های مورد مطالعه که شامل رفتار ۱۰۰۰ فرد در قالب ۱۰ ویژگی مدنظر می‌باشند با استفاده از الگوریتم گرگ خاکستری مورد ارزیابی قرار می‌گیرند. در جدول ٦ نتیجه حاصل از اعمال الگوریتم گرگ خاکستری نشان داده شده است.

جدول ۵. نتایج آماری توصیفی داده‌ها

متغیر		تعداد	درصد
سن	بین ۱۸ تا ۲۵	۳۵۰	۳۵/۰۰
	بین ۲۵ تا ۳۰	۳۲۱	۳۲/۱۰
	بین ۳۰ تا ۳۵	۱۹۸	۱۹/۸۰
	بالاتر از ۳۵	۱۳۱	۱۳/۱۰
جنسیت	زن	٤٦٨	٤٦/٨٠
	مرد	٥٣٢	٥٣/٢٠
وضعیت تأهل	مجرد	٥٨٥	٥٨/٥٠
	متأهل	٤١٥	٤١/٥٠
مالیکت وسیله نقلیه	بله	٤٨٦	٤٨/٦٠
	خیر	٥١٤	٥١/٤٠
درآمد	کمتر از ۵ میلیون تومان	٥٦	٥/٦٠
	بین ۵ تا ۱۰ میلیون تومان	٢٢٥	٢٢/٥٠
	بین ۱۰ تا ۱۵ میلیون تومان	٤١٣	٤١/٣٠
	بیش از ۱۵ میلیون تومان	٣٠٦	٣٠/٦٠
شغل	آزاد	١٨٢	١٨/٢٠
	دانشجو	٢٤٦	٢٤/٦٠
	کارمند تمام وقت	٢٩٢	٢٩/٢٠
	کارمند پاره وقت	٢٢٩	٢٢/٩٠
	بازنشسته	٣٦	٣/٦٠
	فاقد شغل	١٠	١/٠٠
	غیره	٥	٠/٥٠
تحصیلات	دیپلم و زیر دیپلم	٨٩	٨/٩٠
	فوق دیپلم	١٨٦	١٨/٦٠
	کارشناسی	٣٩٤	٣٩/٤٠
	کارشناسی ارشد	٢٩٨	٢٩/٨٠
	بالاتر از کارشناسی ارشد (پزشک و دکتری و بالاتر)	٣٣	٣/٣٠

.

جدول ٦. نتایج اعمال الگوریتم گرگ خاکستری در راستای انتخاب ویژگی

تعداد ویژگی اصلی	تعداد ویژگی انتخاب شده	دقت (٪)	زمان حل (ثانیه)
۱۰	۷	۹۵,٦۸	۸,٦۹

در ادامه پس از تنظیم هایپرپارامترها، به بررسی تخمین نوع سفر با هدف خرید، به کمک رویکرد پیشنهادی پرداخته می‌شود. در جدول ۷ نیز معماری بکاررفته برای الگوریتم CNN پس از تنظیم هایپرپارامترها ارائه شده است.

جدول ۷. تنظیمات اولیه مدل CNN

پارامتر	مقدار درنظر گرفته شده
ضریب یادگیری اولیه	۰/۰۰۵
عامل کاهش نرخ یادگیری	۰/۲
تعداد تکرار	٤۰
حلگر	Adam
معماری شبکه	Alexnet
تابع آموزش دهنده	Trainlm

در گام بعد به منظور مقایسه نتایج بدست آمده با سایر الگوریتم‌ها مانند K نزدیکترین همسایه (KNN)[۹]، درخت تصمیم(DT)[۱۰] و شبکه عصبی پرسپترون چندلایه (MLP)[۱۱] از سه شاخص: دقت، خطای میانگین مربعات (MSE)[۱۲] و ریشه خطای میانگین مربعات (RMSE)[۱۳] استفاده خواهد شد.

خطای میانگین مربعات میزان تفاوت میانگین مقادیر پیش‌بینی شده و مقادیر واقعی را نشان می‌دهد و میزان توانایی پیش‌بینی مدل را ارائه خواهد کرد. هر چه این شاخص کمتر باشد، مدل توانایی بهتری برای پیش‌بینی دارد.

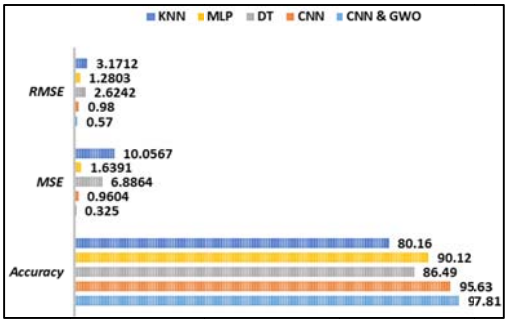
ریشه خطای میانگین مربعات، انحراف متوسط بین مقادیر پیش‌بینی شده و واقعی را نمایش می‌دهد و هرچه این شاخص کمتر باشد، نمایانگر عملکرد مناسب الگوریتم خواهد بود و از طریق فرمول‌های ۱۳ و ۱٤ محاسبه خواهد شد (Dou et al., 2023). نتایج این مقایسه در جدول ۸ و شکل ٤ نمایش داده شده است.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}|Y_i - Y_i^*|^2 \qquad (۱۳)$$

$$RMSE = \sqrt{MSE} \qquad (۱٤)$$

جدول ۸. مقایسه عملکرد الگوریتم‌ها

معیار	مدل تلفیقی	CNN	DT	MLP	KNN
دقت	۹۷/۸۱	۹۵/٦۳	۸٦/٤۹	۹۰/۱۲	۸۰/۱٦
MSE	۰/۳۲۵	۰/۹٦	٦/۸۸	۱/٦۳	۱۰/۰۵
RMSE	۰/۵۷۰	۰/۹۸	۲/٦۲	۱/۲۸	۳/۱۷



شکل ٤. مقایسه عملکرد الگوریتم‌ها

با توجه به جدول ۸ و شکل ٤، عملکرد الگوریتم پیشنهادی با سایر الگوریتم‌ها مقایسه شده است. نتایج نشان داد که الگوریتم پیشنهادی از لحاظ کارایی، برتری قابل توجهی نسبت به سایر الگوریتم‌ها دارد و بهبود نتایج با اضافه شدن الگوریتم بهینه‌ساز گرگ خاکستری در مقایسه با الگوریتم CNN به وضوح قابل مشاهده است.

٦- نتیجه‌گیری

امروزه تعداد زیادی از شهروندان خرید خود را به صورت برخط انجام می‌دهند. محبوبیت روزافزون گوشی‌های هوشمند منجر به افزایش زمان جستجوی شهروندان برای خرید برخط شده است. با این حال، بسته به نوع خرید برخط یا برون‌خط، نوع سفر با هدف خرید از اهمیت بالایی برخوردار است. سیستم حمل‌ونقل مهمترین بخش در زنجیره تأمین شرکت‌های عرضه کننده محصولات برخط می‌باشد. از طرفی در خریدهای برون خط، نوع سفر و استفاده از سیستم‌های حمل‌ونقل توسط شهروندان، علاوه بر مسائل اقتصادی، مسائل زیست محیطی را نیز دربر میگیرد. لذا می‌توان اذعان داشت که هدف اصلی پژوهش حاضر ارائه مدلی مبتنی بر یادگیری ماشین به منظور پیش‌بینی سفرهای با هدف خرید برخط و برون‌خط با استفاده

از داده‌های سری زمانی و بهره‌گیری از رویکرد تلفیقی گرگ خاکستری و یادگیری عمیق بوده است. در این پژوهش، با توجه به فراوانی داده‌ها در مناطق ۲ و ۵ کلانشهر تهران و محاسبات بر اساس فرمول کوکران، ۵۰۰ پرسشنامه از سفرهای برخط از طریق ارسال پیامک و ۵۰۰ پرسشنامه از سفرهای برون‌خط به صورت استقرار در مراکز خرید این دو مناطق از کاربرانی که در سال ۱۳۹۹ سفارش موفق در سرویس‌های برخط و برون‌خط داشته اند، جمع‌آوری شد. سپس با کمی‌سازی داده‌های کیفی و برچسب‌گذاری آن‌ها، به عنوان ورودی مدل استفاده شدند. لازم به ذکر است که از شاخص‌های سن، جنسیت، وضعیت تأهل، مالکیت وسیله نقلیه، هزینه ارسال، زمان ارسال، قیمت محصول، درآمد خانوار، وضعیت اشتغال و میزان تحصیلات به عنوان شاخص‌های اثرگذار بر نوع سفر با هدف خرید استفاده شده است. در گام بعد، با اعمال الگوریتم گرگ خاکستری به منظور تنظیم مقادیر هایپرپارامترها، مدل CNN آماده پردازش و طبقه‌بندی به منظور تخمین نوع سفر با هدف خرید شد. نتایج نشان داد که مدل پیشنهادی با انتخاب ۷ ویژگی وضعیت تأهل، مالکیت وسیله نقلیه، هزینه ارسال، زمان ارسال، قیمت محصول، درآمد خانوار، و وضعیت اشتغال به عنوان ویژگی‌های نهایی، توانسته نوع سفر را با دقت ۹۷/۸۱ درصد تخمین بزند. لازم به ذکر است که عملکرد روش پیشنهادی با مدل های کانولوشن CNN تکی، K–نزدیکترین همسایه (KNN)، درخت تصمیم (DT) و شبکه عصبی (MLP) نیز مقایسه شد. نتیجه این مقایسه نشان داد، پس از الگوریتم پیشنهادی با دقت ۹۷/۸۱ درصد، مدل یادگیری عمیق تکی با دقت ۹۵/۶۳ درصد، شبکه عصبی با دقت ۹۰/۱۲ درصد، درخت تصمیم با دقت ۸۶/۴۹ درصد و مدل KNN با ۸۰/۱۶ درصد قرار دارند. . نتایج این پژوهش دید وسیع و شفافی به برنامه‌ریزان حمل‌ونقل، مدیران کسب‌وکارهای برخط و مدیران قانون‌گذار در توسعه حمل‌ونقل شهری خواهد داد. این نتایج باعث طراحی مؤثر استراتژی‌ها، کاهش هزینه‌های حمل‌ونقل، کاهش آلایندگی هوا و رضایت شهروندان را در برخواهد داشت که همگی منتج به توسعه پایدار خواهد شد.

۷- پی نوشت‌ها

1. Extreme Gradient
2. Online or E-shopping
3. Offline
4. Machine Learning
5. Divergence
6. Local optima
7. Iteration
8. Optimum
9. K-Nearest Neighbors
10. Decision Tree
11. Multi-Layer Perceptron
12. Mean Squared Error
13. Root Mean Square Error

۸- مراجع

- Annual Report of Urban Traffic and Transportation Organization in Tehran City. (2020). Available online: https://www.ictte.ir/data/cnf1668321135/uploads/amar/1399.pdf (in Persian)
- Archetti, C., & Bertazzi, L. (2021). Recent challenges in Routing and Inventory Routing: E-commerce and last-mile delivery. *Networks*, 77(2), 255-268. **dx.doi.org/10.1002/net.21995**
- Ardiansah, M. N., Chariri, A., & Januarti, I. (2019). Empirical study on customer perception of e-commerce: Mediating effect of electronic payment security. *Jurnal Dinamika Akuntansi*, 11(2), 122-131. **doi.org/10.15294/jda.v11i2.20147**
- Bansal, A., & Srivastava, P. (2021). Factors affecting consumer buying behavior of online travel agencies. *Elementary Education Online*, 2021 May 10;20(1):2958
- Candra, S., Nita, S., Loang, OK., Basmantra, IN., & Wong, NK. (2022). Consumer Buying Behavior in Online Travel Agent: A Preliminary Finding. *In 2022 International Conference on Information Management and Technology (ICIMTech),* 2022 Aug 11 24-27. IEEE. https://10.0.4.85/ICIMTech55957.2022.9915179
- Collins, LM. (2007). Research Design and Methods. *In Encyclopedia of Gerontology*; Elsevier: Amsterdam, the Netherlands, Vol. 2, 419–429. **doi.org/10.1016/B0-12-370870-2/00162-1**
- Dasoomi, MH., Naderan, A., & Allahviranloo, T. (2023). Predicting the Choice of Online or Offline Shopping Trips Using a Deep Neural Network Model and Time Series Data: A Case Study of Tehran, Iran. *Sustainability*, 15(20), 14764. **doi.org/10.3390/su152014764**
- Dong, Y., Tang, J., & Zhang, Z. (2022). Integrated Machine Learning Approaches for E-commerce Customer Behavior Prediction. *In 2022 7th International Conference on Financial Innovation and Economic*

doi.org/10.1016/j.annals.2020.103083
- Lipsman, June 27, (2019). Global ecommerce 2019. https://www.emarketer.com/content/global-ecommerce.
- Memeti, S., Pllana, S., Binotto, A., Kołodziej, J., & Brandic, I. (2019). Using meta-heuristics and machine learning for software optimization of parallel computing systems: a systematic literature review. *Computing*, 101(8), 893-936. **doi.org/10.1007/s00607-018-0614-9**
- Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in engineering software*, 69, 46-61. **dx.doi.org/10.1016/j.advengsoft.2013.12.007**
- Moon, J., Choe, Y., & Song, H. (2021). Determinants of consumers' online/offline shopping behaviors during the COVID-19 pandemic. **doi.org/10.3390/ijerph18041593**
- Moshayedi, A.J., Roy, A.S., Taravet, A.;,Liao, L., Wu, J., Gheisari, M. (2023). A secure traffic police remote sensing approach via a deep learning-based low-altitude vehicle speed detector through uavs in smart cites: Algorithm, implementation and evaluation. *Future Transp*, 3, 189–209.
- Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), 1-21. **doi.org/10.1007/s42979-021-00592-x**
- Shao, R., Derudder, B., & Witlox, F. (2022). The geography of e-shopping in China: On the role of physical and virtual accessibility. Journal of Retailing and Consumer Services, 64, 102753. **doi.org/10.1016/j.jretconser.2021.102753**
- Stocchi, L., Michaelidou, N., Pourazad, N., & Micevski, M. (2018). The rules of engagement: How to motivate consumers to engage with branded mobile apps. *Journal of Marketing Management*, 34(13-14), 1196-1226. **doi.org/10.1080/0267257X.2018.1544167**
- Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: a survey. *Journal of Big data*, 2(1), 1-32. **doi.org/10.1186/s40537-015-0030-3**
- Xiahou, X., & Harada, Y. (2022). B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(2), 458-475. **doi.org/10.3390/jtaer17020024**

Development (ICFIED 2022), 1008-1015. Atlantis Press. **doi.org/10.2991/aebmr.k.220307.166**
- Dou, Z., Sun, Y., Zhu, J., & Zhou, Z. (2023). The Evaluation Prediction System for Urban Advanced Manufacturing Development, *Systems*, 11(8), 392. **doi.org/10.3390/systems11080392**
- Hussien, AG., Houssein, EH., & Hassanien, AE. (2017). A binary whale optimization algorithm with hyperbolic tangent fitness function for feature selection. *In 2017 Eighth international conference on intelligent computing and information systems (ICICIS)*, 2017 Dec 5, IEEE. 166-172. **doi.org/10.1109/INTELCIS.2017.8260031**
- Jawandhiya, P. (2018). Hardware design for machine learning. *International Journal of Artificial Intelligence and Applications (IJAIA)*, 9(1), 63-84. **dx.doi.org/10.5121/ijaia.2018.9105**
- Jeffers, J., & Reinders, J. (2015). High performance parallelism pearls volume two: multicore and many-core programming approaches. Morgan Kaufmann.
- Jiang, H., He, M., Xi, Y., & Zeng, J. (2021). Machine-Learning-Based User Position Prediction and Behavior Analysis for Location Services. *Information*, 12(5), 180. **doi.org/10.3390/info12050180**
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260. **doi.org/10.1126/science.aaa8415**
- Kamel, SR., YaghoubZadeh, R., & Kheirabadi, M. (2019). Improving the performance of support-vector machine by selecting the best features by Gray Wolf algorithm to increase the accuracy of diagnosis of breast cancer. *Journal of Big Data*, 6(1):90.
- Kennedy, I. (2022). Sample Size Determination in Test-Retest and Cronbach Alpha Reliability Estimates. Br. J. Contemp. Educ. 2022, 2, 17–29. **doi.org/10.52589/BJCE-FY266HK9**
- Khrais, LT. (2020). Role of Artificial Intelligence in Shopping Consumer Demand in E-Commerce. *Future Internet*, 12(12), 226. **doi.org/10.3390/fi12120226**
- Li, C., Wang, Y., Lv, X., & Li, H. (2021). To buy or not to buy? The effect of time scarcity and travel experience on tourists' impulse buying. *Annals of Tourism Research*, 86, 103083.

Predict the Shopping Trips (Online and Offline) using a combination of a Gray Wolf Optimization Algorithm (GWO) and a Deep Convolutional Neural Network (Case Study: Tehran)

MohammadHanif Dasoomi, Ph.D., Candidate, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
Ali Naderan, Assistant Professor, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
Tofigh Allahviranloo, Professor, Department of Mathematical Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran.

E-mail: naderan@srbiau.ac.ir

Received: February 2024- Accepted: June 2024

ABSTRACT
Online and offline shopping trips affect various aspects of urban life, such as e-commerce, transportation systems, and sustainability. To evaluate the factors that influence shoppers' choices, we propose a hybrid machine learning model that combines a gray wolf optimization algorithm and a deep convolutional neural network. We apply this model to estimate shopping trips based on a survey of 1000 active e-commerce users in districts 2 and 5 of Tehran, who made successful orders in both online and offline services in 2020. The gray wolf optimization algorithm performs feature selection and hyperparameters tuning for the deep convolutional neural network, which is a powerful deep learning model for image recognition and classification. Our model achieves an accuracy of 97.81% with an MSE of 0.325 by selecting seven out of ten features. The most important features are delivery cost, delivery time, product price, and car ownership. In addition, comparing the performance of the proposed method with other methods showed that the proposed algorithm with an accuracy of 97.81%, the accuracies of the single deep learning model, MLP neural network, decision tree, and KNN models were 95.63%, 90.12%, 86.49%, and 80.16%, respectively. This study offers valuable insights for transportation planners, e-commerce managers, and policymakers. It aims to help them design effective strategies to reduce transportation costs, lower pollutant emissions, alleviate urban traffic congestion, and enhance user satisfaction all while promoting sustainable development.

Keywords: Online Shopping Trips, Offline Shopping Trips, Gray Wolf Optimization, Deep Neural Network Model, Sustainable Development