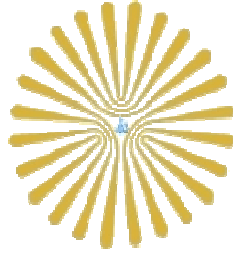


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه سям نور

دانشکده فنی و مهندسی

گروه مهندسی کامپیوتر و فناوری اطلاعات

پایان نامه

برای دریافت درجه کارشناسی ارشد

رشته مهندسی کامپیوتر گرایش نرم افزار

# ارائه مدلی برای سیستم های پیشنهاددهنده وب بر اساس کاربرد کاوی وب

نگارش:

فروغ یوسف زاده شوشتری

استاد راهنما:

دکتر صمد مومن بالله

استاد مشاور:

دکتر سید علی رضوی

بهمن ۱۳۹۰

خداوند ما

توفیق تلاش در شگفت،

صبر در نومی،

فداکاری در سکوت،

خوبی بی نمود،

مناعت بی غرور،

عشق بی هوس،

تنهایی در انبوه جمعیت و

دوست داشتن بی آنکه دوست بدانند را، عنایت فرما.

تشکر ویژه دارم از استادان عزیز آقایان دکتر صمد مومن بالله و دکتر سید علی رضوی که در تمام مراحل ارائه این پایان نامه مرا یاری کردند و از در اختیار گذاشتن هر زمانی برای یاری من دریغ نفرمودند. در نهایت جا دارد از تمامی اساتیدم در طول دوران تحصیل که راه را به من آموختند تشکر و قدردانی به عمل آورم.

یعنی دارم روزی می آید

که عمر آدمی را بر اساس سالهای عاشق بودنش می سنجند، و انسان بدون عشق یعنی انسان مرده.

ساده بگویم، خلاصه و بی پروا:

تقدیم به بهترین همسر دنیا

به پاس صداقت بیکرانیش، دستان پر مهرش، و بهمراهی حُسنکی نامذیرش.

تقدیم به پدر بزرگوار و مادر مهربانم

آن دو فرشته ای که از خواسته هایشان گذشتند، سختی ها را به جان خریدند و خود را سپر پلهای مشکلات و ناملایمات کردند تا من به جایگاهی که اکنون در آن ایستاده ام برسم. پروردگارا، پس توفیقم ده که هر لحظه سگم کز ارشان باشم و ثانیه های عمرم را در عصای دست بودنشان بگذرانم.

تقدیم به خواهرم، که وجودش شادی، نخس و صفایش مایه آرامش من است.

تقدیم به برادرم، که همواره در طول تحصیل متحمل زحمت بوده و وجودش مایه دلگرمی من می باشد.

## چکیده

بی شک یکی از مهمترین فناوری‌هایی که در عصر حاضر زندگی ما را تحت تاثیر قرار داده، اینترنت است. با این وجود، افزایش قابل توجه وبسایت‌ها، گسترش روز افزون اینترنت، رشد نمایی کاربران و محتویات وب انجام سازماندهی محتوای وب جهت تامین نیاز کاربران را به عملی چالش برانگیز تبدیل ساخته و تحلیل و بررسی اطلاعات وب را از اهمیت ویژه‌ای برخوردار کرده است. یک راه غلبه بر این مشکلات استفاده از برنامه‌هایی موسوم به سیستم‌پیشنهاددهنده است.

در این پایان نامه مدلی جدید برای سیستم پیشنهاددهنده ارائه شده که بر پایه کاربردکاوی کار می‌کند اما ساختار وبسایت را نیز برای ارائه پیشنهادات بهتر در نظر می‌گیرد، زیرا کاربر در صورتی از یک صفحه به صفحه دیگر می‌رود که بین آنها یک پیوند برقرار باشد. در این مدل، تابع علاقه کاربر به صفحه را با استفاده از زمان سپری شده کاربر در آن صفحه تخمین می‌زنیم و از این طریق داده‌های پرت زیادی را از فضای نمونه حذف می‌کنیم. در این مدل از درخت با حداکثر عمق ثابت استفاده شده که جستجو را بسیار راحت کرده است، همچنین با استفاده از لیست‌های مرتب زمان جستجوی درخت بسیار بهبود بخشیده شده است. این ترکیب باعث شده تا این سیستم بتواند میزان اثرگذاری پیشنهادات را به میزان قابل توجهی بهبود بخشد.

کلمات کلیدی: سیستم پیشنهاد دهنده وب، کاربردکاوی وب، ساختارکاوی وب، تابع علاقه، درخت پیشنهاد

## فهرست مطالب

۱۴	فصل ۱. مقدمه
۱۵	۱-۱. مقدمه
۱۷	۲-۱. تعریف مساله و سوالات اصلی تحقیق
۱۷	۳-۱. اهداف تحقیق
۱۸	۴-۱. کاربردهای متصور از تحقیق
۱۸	۵-۱. مراجع استفاده کننده از نتایج پایان نامه
۱۸	۶-۱. روش انجام تحقیق
۱۹	۷-۱. جامعه آماری و روش نمونه گیری
۲۰	۸-۱. ساختار پایان نامه
۲۱	فصل ۲. مروری بر منابع مطالعاتی
۲۲	۱-۲. مقدمه
۲۲	۲-۲. تکنیک‌های داده‌کاوی وب
۲۲	۱-۲-۲. محتواکاوی وب
۲۳	۲-۲-۲. ساختارکاوی وب
۲۳	۳-۲-۲. کاربردکاوی وب
۲۴	۴-۲-۲. متن کاوی وب
۲۴	۳-۲. شخصی‌سازی و سیستم‌های پیشنهاد دهنده
۲۵	۱-۳-۲. تعریف
۲۶	۲-۳-۲. طبقه بندی
۲۷	۳-۳-۲. نقش کاربردکاوی وب در سیستم‌های پیشنهاد دهنده
۲۷	۴-۲. مروری بر کارهای انجام شده
۳۱	فصل ۳. فرآیند کاربردکاوی وب
۳۲	۱-۳. مقدمه
۳۳	۲-۳. جمع آوری داده‌ها و پیش پردازش
۳۴	۱-۲-۳. منابع و انواع داده

.....	آماده‌سازی و پیش پردازش داده‌ها	۲-۲-۳	۴۲
.....	کشف الگوهای حرکتی کاربر از داده‌های کاربرد وب	۳-۳	۵۲
.....	کاوش قواعد انجمنی	۱-۳-۳	۵۳
.....	کشف الگوهای ترتیبی	۲-۳-۳	۵۴
.....	خوشه‌بندی	۳-۳-۳	۵۹
.....	استفاده از الگوهای کشف شده جهت سیستم پیشنهاد دهنده	۴-۳	۶۲
.....	<b>فصل ۴. معماری سیستم پیشنهادی</b>		۶۴
.....	مقدمه	۱-۴	۶۵
.....	جمع آوری داده‌ها	۲-۴	۶۷
.....	داده‌های استاندارد دانشگاه CTI Depaul	۱-۲-۴	۶۷
.....	جزئیات فایل داده‌های استاندارد CTI	۲-۲-۴	۶۷
.....	پاکسازی و آماده‌سازی داده‌ها	۳-۴	۷۰
.....	پاکسازی داده‌ها	۱-۳-۴	۷۰
.....	محاسبه تابع علاقه کاربر به صفحه وب	۲-۳-۴	۷۰
.....	کشف و استخراج الگوهای حرکتی کاربران	۴-۴	۷۶
.....	ساخت درخت جلسات	۱-۴-۴	۷۶
.....	ساخت درخت پیشنهاد	۲-۴-۴	۷۷
.....	پاکسازی درخت پیشنهاد	۳-۴-۴	۷۸
.....	ساخت لیست مرتب	۴-۴-۴	۷۸
.....	فاز آنلاین	۵-۴	۷۹
.....	<b>فصل ۵. پیاده‌سازی سیستم پیشنهادی</b>		۸۱
.....	مقدمه	۱-۵	۸۲
.....	فاز آفلاین	۲-۵	۸۲
.....	آماده‌سازی داده‌ها	۱-۲-۵	۸۲
.....	پاکسازی داده‌ها و محاسبه تابع علاقه	۲-۲-۵	۸۸
.....	ساخت درخت جلسات و درخت پیشنهاد	۳-۲-۵	۹۸
.....	فاز آنلاین	۳-۵	۱۰۰
.....	<b>فصل ۶. ارزیابی، نتایج و کارهای آتی</b>		۱۰۱
.....	مقدمه	۱-۶	۱۰۲
.....	ارزیابی	۲-۶	۱۰۲

۱۰۳.....	نتایج ارزیابی.....	۳-۶
۱۰۸.....	بحث.....	۴-۶
۱۰۹.....	نتیجه‌گیری و کارهای آتی.....	۵-۶
۱۱۰.....	<b>منابع</b> .....	
۱۱۳.....	<b>واژه‌نامه</b> .....	



## فهرست جدول‌ها

- جدول (۱-۲) - مقایسه کلی کارهای گذشته ..... ۳۰
- جدول (۱-۴) - ارزیابی تجربی برای انتخاب بهترین ضرایب ..... ۷۴
- جدول (۱-۶) - مجموعه داده‌ای استفاده شده در آزمایشات ..... ۱۰۲
- جدول (۲-۶) - مجموعه داده‌ای پس از پیش آماده‌سازی ..... ۱۰۲

## فهرست نمودارها

- نمودار (۱-۴) - نمودار هیستوگرام توزیع زمان سپری شده یک صفحه وب ..... ۷۲
- نمودار (۲-۴) - هیستوگرام توزیع لگاریتم زمان سپری شده ..... ۷۲
- نمودار (۳-۴) - نمودار تابع علاقه ..... ۷۵
- نمودار (۴-۴) - هیستوگرام تابع علاقه ..... ۷۵

## فهرست شکل‌ها

- شکل (۱-۳) - تراکنش HTTP ..... ۳۵
- شکل (۲-۳) - URL و URI ..... ۳۶
- شکل (۳-۳) - مکانهای جمع‌آوری داده‌های کاربرد وب ..... ۳۷
- شکل (۴-۳) - مثالی از مدل کردن رفتار گردش کاربر با زنجیره‌ی مارکف ..... ۵۷
- شکل (۵-۳) - مثالی از مدل کردن رفتار گردش در یک درخت تجمعی ..... ۵۸
- شکل (۶-۳) - مثالی از استخراج پروفایل‌های تجمعی کاربرد از خوشه‌های تراکنش‌ها ..... ۶۰
- شکل (۱-۴) - شمای کلی مدل سیستم پیشنهاد دهنده ..... ۶۶
- شکل (۲-۴) - فایل cti.cod ..... ۶۷
- شکل (۳-۴) - فایل cti.tra ..... ۶۸
- شکل (۴-۴) - فایل cti-stat.txt ..... ۶۹
- شکل (۵-۴) - فایل cti.nav ..... ۶۹
- شکل (۶-۴) - درخت جلسات ساخته‌شده از مشاهدات صفحه ..... ۷۶
- شکل (۷-۴) - درخت پیشنهادات ساخته شده توسط پیمایش درخت جلسات ..... ۷۸
- شکل (۸-۴) - لیست مرتب برای سطح دوم درخت پیشنهاد ..... ۷۹
- شکل (۱-۵) - جدول Pages ..... ۸۳
- شکل (۲-۵) - جدول Transaction ..... ۸۴
- شکل (۳-۵) - جدول FullPageViews ..... ۸۵
- شکل (۴-۵) - جدول PageViews ..... ۸۶
- شکل (۵-۵) - جدول Referrers ..... ۸۷
- شکل (۶-۵) - جدول Stats ..... ۸۸
- شکل (۷-۵) - نمای برنامه پاکسازی داده‌ها و محاسبه تابع علاقه ..... ۸۹
- شکل (۸-۵) - جدول 1\_Referrers ..... ۹۰
- شکل (۹-۵) - جدول 2\_FilterLog ..... ۹۱
- شکل (۱۰-۵) - جدول 3\_Pages ..... ۹۲

۹۳	شکل (۱۱-۵) - جدول 4_FilLog_PageId
۹۴	شکل (۱۲-۵) - جدول 5_Timed_FilLog
۹۵	شکل (۱۳-۵) - جدول 6_PagesViewAvg
۹۶	شکل (۱۴-۵) - جدول 7_PagesLogAvgStd
۹۷	شکل (۱۵-۵) - جدول 8_PageLimits
۹۸	شکل (۱۶-۵) - جدول 9_InterestLog
۹۹	شکل (۱۷-۵) - نمای برنامه ساخت درخت جلسات و درخت پیشنهاد
۱۰۴	شکل (۱-۶) - خروجی مرحله اول فاز آفلاین
۱۰۵	شکل (۲-۶) - مراحل ساخت دانش سیستم
۱۰۵	شکل (۳-۶) - لیست پیشنهادات سیستم
۱۰۶	شکل (۴-۶) - میانگین دقت پیشنهادات سیستم
۱۰۶	شکل (۵-۶) - میانگین پوشانندگی پیشنهادات سیستم
۱۰۷	شکل (۶-۶) - میانگین اثر گذاری پیشنهادات سیستم
۱۰۷	شکل (۷-۶) - مقایسه معیار دقت پیشنهادات
۱۰۸	شکل (۸-۶) - مقایسه معیار پوشانندگی پیشنهادات
۱۰۸	شکل (۹-۶) - مقایسه معیار اثر گذاری پیشنهادات

## فهرست فرمول‌ها

- فرمول (۱-۴) - حد آستانه قابل قبول در توزیع نرمال ..... ۷۳
- فرمول (۲-۴) - نمودار تابع علاقه کاربر به صفحه وب ..... ۷۴
- فرمول (۳-۴) - تابع علاقه ۲ متغیره با اعمال رگرسیون ..... ۷۵

# فصل ۱. مقدمه

## ۱-۱. مقدمه

امروزه در عصری به سر می‌بریم که به عصر ارتباطات معروف است. یکی از مهمترین فناوری‌هایی که خود زمینه‌ساز بسیاری از خلاقیت‌ها و ابتکارات ارتباطی و اطلاعاتی شده است، شبکه جهانی یا اینترنت است. با رشد سریع اطلاعات موجود روی اینترنت، شبکه گسترده جهانی یک پلتفرم قدرتمند برای ذخیره، انتشار و بازیافت اطلاعات<sup>۱</sup> و همچنین معدنی از دانش مفید شده است. با رشد انفجار آمیز دانش موجود در شبکه گسترده جهانی که فاقد ساختار یا شمای مجتمع شده‌ای است، به علت خصوصیات نظیر ماهیت عظیم، مختلف، پویا و غیرساخت‌یافته داده‌های وب، تحقیق در داده‌های وب با چالش‌های زیادی مواجه شده است. اینترنت با همه مزایایی که دارد، مشکلاتی را برای استفاده‌کنندگان خود به وجود آورده است. زمانی که کاربر به منظور رفع نیاز خاصی به اینترنت متصل می‌شود با حجم عظیم و یا به تعبیری انفجاری از اطلاعات مرتبط با نیازش مواجه می‌شود. جستجو و انتخاب بهترین گزینه از میان همه گزینه‌های موجود مشکلی است که اکثر کاربران آن را تجربه کرده‌اند. همچنین برای کاربران دشوارتر شده است تا به اطلاعات مرتبط با خواسته‌شان به طور کارا و موثر دسترسی پیدا کنند. در این میان، افزایش قابل توجه وب‌سایت‌ها، انجام سازماندهی محتوای وب جهت تامین نیاز کاربران را به عملی چالش برانگیز برای رهبران وب تبدیل ساخته است.

یک راه غلبه بر این مشکلات استفاده از برنامه‌هایی موسوم به سیستم پیشنهاددهنده<sup>۲</sup> است. سیستم‌های پیشنهاد دهنده مبتنی بر وب برای کمک به کاربران وب در حرکت بین صفحات وب، تعیین مکان و یافتن اطلاعات مرتبط و مفید و دریافت پیشنهادات پویا از وب‌سایت روی محصولات و تولیدات ممکن یا سرویس‌هایی که با نیاز آنها مطابقت دارد ارائه شده‌اند.

سیستم‌های پیشنهاد دهنده از آرا و عقاید گروهی از کاربران استفاده می‌کنند تا به تک‌تک آنها در آن گروه کمک کنند تا محتوای مورد نظر خود را به طور کارا تر و موثرتر از یک مجموعه از انتخابات، مشخص و تعیین کنند که این کار بطور بالقوه طاقت‌فرسا و سخت است. سیستم‌های پیشنهاد دهنده به عنوان تکنیک‌های هوشمند مبتنی بر کامپیوتر برای تسهیل تراکنش‌هایی با سربرار زیاد اطلاعات و مشکلات اطلاعاتی تعریف می‌شوند. یک سیستم پیشنهاد دهنده در واقع مکانیزمی قوی برای انجام عمل پالایش

---

<sup>1</sup> Information Retrieve (IR)

<sup>2</sup> Recommender System

اطلاعاتی است. این سیستم‌ها به تکنیک محبوبی برای هرس کردن فضاهای بزرگ اطلاعاتی تبدیل شده‌اند که کاربران را به سمت بهترین اقلامی که نیازشان را مرتفع می‌کنند هدایت می‌نمایند. این تکنیک جزء جدانشدنی سایت‌های تجاری از قبیل Yahoo, Amazon و CDNow شده است.

این سیستم‌ها به طور موثر می‌توانند به هر دو طرف تراکنش سود برسانند. سیستم پیشنهاد دهنده به مشتری از طریق هدایت وی به سمت گزینه‌های مورد علاقه‌اش و رفع نیاز وی کمک می‌کند. هم چنین با افزایش مراجعه مشتریان به این سیستم‌ها و رضایتمندی آنها، سازمان‌های ارائه دهنده نیز به هدف خود نائل می‌آیند.

برای فراهم نمودن سرویس‌های آنلاین شخصی‌سازی هوشمند مانند سیستم‌های پیشنهاد دهنده مبتنی بر وب، معمولاً لازم است رفتار حرکتی کاربران را مدل‌سازی کرد.

وفور اطلاعات در شبکه گسترده جهانی کاربران را متمایل به جستجو و بازیافت اطلاعات از اینترنت کرده است. گرچه این ازدیاد و وفور اطلاعات اغلب مجموعه‌ای از مشکلات را به همراه دارد، به طوریکه کاربران قادر نیستند که اطلاعات مفید و مرتبطی بازیافت نمایند. یکی از دستاوردهای بالقوه برای رفع این مشکل، تحلیل الگوهای حرکتی تعاملات کاربران با یک یا چند وب‌سایت می‌باشد.

کاربردکاوی وب<sup>۳</sup> یکی از دستاوردهای اصلی استفاده شده برای ساخت سیستم‌های پیشنهاد دهنده وب می‌باشد. هدف از بکارگیری تکنیک‌های کاربردکاوی وب، پیشنهاد دادن مجموعه‌ای از موضوعات به کاربر فعال می‌باشد که ممکن است این موضوعات شامل پیوندها، متن‌ها، محصولات و از این قبیل چیزها باشد که در خور و مناسب اولویتها و ترجیحات پیش‌بینی شده آن کاربر می‌باشد. این وظیفه و عمل توسط تطابق جلسه کاربر فعال با الگوهای کاربردی کشف شده توسط کاربردکاوی وب انجام شده است. الگوریتم‌های کاربردکاوی وب به طور گسترده‌ای برای مدل‌سازی رفتار حرکتی کاربر در وب به کار گرفته می‌شوند.

کاربردکاوی به کشف خودکار و تحلیل الگوها در داده‌های کلیک شده و تولیدشده به عنوان نتیجه تعاملات کاربر با منابع وب یا چندین وب‌سایت اشاره دارد و بطور موثری به عنوان دستاوردی برای شخصی‌سازی خودکار و روشی جهت غلبه بر کمبودهای دستاوردهای قدیمی مانند پالایش جمعی<sup>۴</sup> استفاده شده است.

تحلیل الگوهای حرکتی کاربران می‌تواند به سازمانها کمک کند تا پیشنهادات صفحات وب را بر اساس تمایلات و علایق جاری کاربر مهیا سازند. تحلیل و مدل‌سازی رفتار حرکتی وب در فهم درخواست‌های کاربران آنلاین مفید است و متعاقب آن نتایج تحلیل می‌توانند به عنوان دانش برای استفاده

<sup>3</sup> Web Usage Mining

<sup>4</sup> Collaboration Filtering



در برنامه‌های کاربردی آنلاین هوشمند، پالایش و تصفیه نقشه‌های وب‌سایت بر پایه سیستم شخصی‌سازی و پیشرفت دقت جستجو به هنگام کاوش اطلاعات دیده شوند.

## ۱-۲. تعریف مساله و سوالات اصلی تحقیق

سیستم‌های پیشنهاددهنده از آرا و عقاید گروهی از کاربران استفاده می‌کنند تا به تک‌تک آنها در آن گروه کمک کنند تا محتوای مورد نظر خود را به طور موثرتر از یک مجموعه از انتخاب‌ها مشخص کنند. یکی از تکنیکهای مورد استفاده در سیستم‌های پیشنهاددهنده کاربردکاوی وب می‌باشد. الگوریتم‌های کاربردکاوی برای استخراج الگوهای حرکتی کاربران به کار گرفته می‌شوند. تحلیل الگوهای حرکتی کاربران در فهم درخواستهای کاربران جاری مفید است. سیستم‌های پیشنهاد دهنده مبتنی بر کاربردکاوی وب، با تطابق جلسه کاربر جاری با الگوهای حرکتی استخراج شده به هدف خود نایل می‌شوند.

علاوه بر این داده‌های ساختار حاوی اطلاعات با ارزشی هستند و با استفاده از این اطلاعات دانش زیادی راجع به صفحات و ارتباطشان به دست می‌دهند. زیرا طراحان وب‌سایت از یک صفحه به صفحه‌ای دیگر زمانی پیوند قرار می‌دهند که عنوان و محتوای صفحات مذکور در راستای محتوای آن صفحه باشد.

در راستای این تحقیق سوالات زیر مطرح می‌باشند:

- چه دلایلی برای استفاده از رفتار حرکتی کاربران و داده‌های ساختار وب‌سایت در سیستم‌های پیشنهاد دهنده وجود دارد؟
- میزان تاثیر بکارگیری الگوهای حرکتی کاربران در بهبود کیفیت و دقت پیشنهادات ارائه شده توسط سیستم پیشنهاد دهنده تا چه اندازه است؟
- چه مشکلاتی بر سر راه تکنیکهای تشخیص الگوی حرکتی کاربران وجود دارد؟
- در چه کاربردهایی دقت پیشنهادات سیستم‌های پیشنهاد دهنده وب حائز اهمیت است؟

## ۱-۳. اهداف تحقیق

اصلی‌ترین هدف این تحقیق بهبود دقت پیشنهادات در سیستم‌های پیشنهاد دهنده وب است که در راستای رسیدن به آن، اهداف زیر دنبال خواهند شد:

- ارائه تابع علاقه کاربر به صفحه بر اساس الگوی حرکتی کاربران بر مبنای ساختار صفحات وبسایت.
- تولید پیشنهادات بر اساس معیار حفظ حداکثر شباهت میان الگوی حرکتی کاربر جاری و الگوهای حرکتی استخراج شده.
- به دلیل تازگی موضوع و عدم وجود تحقیقات کافی در کشور در این حوزه دانشی، یکی از اهداف انجام این تحقیق شروعی برای بکارگیری قابلیت‌های کاربردکاوی وب در سیستم‌های پیشنهاد دهنده و انجام تحقیقات بیشتر بوده است.

#### ۴-۱. کاربردهای متصور از تحقیق

افزایش قابل توجه در تعداد وبسایتهای موجود یک عمل چالش برانگیز برای مدیران وب از نظر سازماندهی محتویات وبسایتها برای فراهم آوردن نیاز کاربران عرضه کرده است. از طرف دیگر وجود سازمان‌های بسیار زیاد در امر ارائه نیازهای مشتری، کاربران را با مشکل جستجو و انتخاب روبرو کرده که متعاقباً سردرگمی آنها را در دهکده جهانی به دنبال دارد. یک راه غلبه بر این مشکلات استفاده از سیستمهای پیشنهاددهنده وب است.

#### ۵-۱. مراجع استفاده کننده از نتایج پایان نامه

محققین و پژوهشگران، برنامه نویسه‌ها، اساتید و دانشجویان که در حوزه سیستم‌های پیشنهادگر مبتنی بر کاربرد کاوی وب فعالیت می‌کنند و به تبع اشاعه آن در دستگاههای اجرایی و موسسات پژوهشی و آموزشی از این پایان نامه استفاده خواهند کرد.

#### ۶-۱. روش انجام تحقیق

روش انجام تحقیق به صورت علمی و سابقه پیشین از طریق مطالعه طرح‌ها و روش‌های مختلف موجود در راستای موضوع تحقیق و نیز به صورت تطبیقی از طریق مقایسه با روش‌های مرتبط موجود

خواهد بود. همچنین روش تحقیق، روشی تجربی نیز می‌باشد که استفاده از ابزارها و مقایسه‌ی آن با فعالیت-های مشابه اثبات خواهد شد.

#### • روش و ابزار گرد آوری اطلاعات

مطالعه مقالات، کتابهای الکترونیکی، اینترنت، گزارشات فنی موجود و استفاده از داده‌های استاندارد.

#### • روش تجزیه و تحلیل داده‌ها

۱. طبقه‌بندی اطلاعات: در این قسمت دیدگاه‌های موجود و سیستم‌های پیاده‌سازی شده با کمک این دیدگاه‌ها، بررسی گردیده و بر اساس نقاط ضعف و قوت طبقه‌بندی می‌گردند.
۲. تحلیل نظری و بررسی راهکارهای مناسب جهت ایجاد سیستم مطلوب.

#### • قلمرو تحقیق (زمانی، مکانی، موضوعی)

۱. قلمرو موضوعی: نرم افزار کامپیوتر، در این زمینه دیدگاه‌ها و مدل‌های پیشنهادی که برای سیستم‌های پیشنهاد دهنده ارائه گردیده است بررسی خواهد شد.
۲. قلمرو زمانی: فعالیت‌های تحقیقاتی این حوزه در بازه‌ی زمانی ۱۹۹۴ الی ۲۰۱۰ بررسی خواهد شد.
۳. قلمرو مکانی: فعالیت‌های تحقیقاتی انجام شده در دانشگاه‌ها، سازمانها، گروه‌های تحقیقاتی و ... (در حد قابل دسترسی) بررسی خواهد شد.

### ۷-۱. جامعه آماری و روش نمونه‌گیری

یکی از روشهایی که برای ارزیابی الگوریتمهایی که از اطلاعات رفتار حرکتی کاربران استفاده می‌کنند وجود دارد استفاده از صفحات وب واقعی و داده‌های واقعی کاربران وب موجود در فایل‌های ثبت وقایع سایت‌ها می‌باشد. برای استفاده از این روش مجموعه داده‌های استاندارد که از چند سایت معتبر استخراج شده‌اند در دسترس می‌باشد.

ما در این تحقیق از داده‌های استاندارد سایت CTI Depaul استفاده می‌کنیم. این مجموعه داده‌های اطلاعات نشست کاربران را به مدت ۲ هفته در سایت CTI Depaul در سال ۲۰۰۲ شامل می‌شود.

## ۸-۱. ساختار پایان نامه

در فصل اول به بررسی کلیات، اهداف و ضرورت‌های این تحقیق پرداخته شده است. در فصل دوم تعاریف پایه‌ی سیستم‌های پیشنهاد دهنده وب و ابعاد آن مورد بررسی قرار گرفته است، همچنین منابع مطالعاتی در راستای موضوع تحقیق مطرح شده‌اند. فرآیند کاربردکاوی و تکنیک‌های آن در فصل سوم به تفصیل شرح داده می‌شوند. فصل چهارم شامل ایده‌ی اصلی در ارائه‌ی یک معماری جدید برای سیستم پیشنهاد دهنده وب بر پایه کاربردکاوی وب و استفاده از ساختار وب‌سایت می‌باشد. جزئیات پیاده‌سازی معماری پیشنهادی در فصل پنجم شرح داده شده است. جهت بررسی کارایی و کیفیت پیشنهادات مدل ارائه‌شده، در فصل ششم به یک مطالعه‌ی موردی پرداخته شده است همچنین در این فصل نتایج ارزیابی شده و کارهای آتی ممکن بررسی خواهند شد.