



پیشنهاد پایانه کارشناسی ارشد رشته تجارت الکترونیکی دانشکده کامپیوتر

عنوان پژوهش:

|            |                                                                                                                        |
|------------|------------------------------------------------------------------------------------------------------------------------|
| ۱-فارسی:   | ارزیابی خودکار ماکاپ وبسایت ها و اپلیکیشن های تجارت الکترونیکی برای تولید بازخوردهای تخصصی جهت بهبود طراحی             |
| ۲-انگلیسی: | Automated Evaluation of E-Commerce Website and Application Mockups and Generation of Expert Feedback to Improve Design |

مشخصات دانشجو:

| نام و نام خانوادگی | شماره دانشجویی | رشته و گرایش                     | امضا |
|--------------------|----------------|----------------------------------|------|
| علی آرمان          | ۴۰۲۳۶۲۴۰۰۲     | فناوری اطلاعات/ تجارت الکترونیکی |      |

مشخصات استادان راهنما و مشاور:

| ردیف | نام و نام خانوادگی       | دانشگاه و گروه آموزشی/ سایر مؤسسات | تخصص                                      | استاد راهنما | امضا |
|------|--------------------------|------------------------------------|-------------------------------------------|--------------|------|
|      | سید فخرالدین<br>نوربهانی | فناوری اطلاعات/ تجارت الکترونیکی   | تجارت الکترونیکی<br>تعامل انسان- کامپیوتر |              |      |

ثبت پیشنهاد در ایرانداک:

| نوع ثبت    | شماره نامه گواهی ثبت | تاریخ نامه گواهی ثبت |
|------------|----------------------|----------------------|
| همانندجویی |                      |                      |
| ثبت نهایی  |                      |                      |

## هدفمندسازی پایان نامه و رساله: (راهنما)

| نوع هدفمندی (بر اساس شیوه نامه هدفمندی) | سازمان حمایت کننده یا عنوان هسته پژوهشی* |
|-----------------------------------------|------------------------------------------|
|                                         |                                          |

\* منظور هسته پژوهشی است که طبق شیوه نامه گزینش فیروزه ای مصوب معاونت پژوهشی دانشگاه شده است.

### شناسه اخلاق در پژوهش:

☐ نیاز دارد

☒ نیاز ندارد

در صورت نیاز به شناسه اخلاق، بلافاصله پس از تصویب پیشنهاد در گروه، اقدام شود.

### کلید واژه ها:

در صورت نیاز، واژه ها تعریف و ارجاع دهی شود.

### فارسی (انگلیسی)

- ۱- تحلیل موکاپ (MockUp Analysis)
- ۲- شبکه عصبی عمیق (Deep Neural Network)
- ۳- ارزیابی رابط کاربری (User Interface Evaluation)
- ۴- تولید خودکار بازخورد (Automatic Feedback Generation)
- ۵- تست کاربردپذیری (Usability Testing)

## ۱- بیان مسئله پژوهش:

طراحی رابط کاربری که به اختصار تحت عنوان (UI) نیز شناخته می‌شود در دوران فناوری محور کنونی، نقشی تعیین کننده در موفقیت وبسایت‌ها و همچنین برنامه‌های کاربردی (اپلیکیشن‌ها) ایفا می‌کند. بر کسی پوشیده نیست که با توجه به افزایش پیچیدگی محصولات نرم‌افزاری و متناسب با آن، بالا رفتن انتظارات کاربران، ارزیابی دقیق طراحی رابط کاربری به امری چالش برانگیز و حیاتی بدل شده است و از آنجایی که بکارگیری روش‌های سنتی ارزیابی که عمدتاً شامل تحلیل‌های دستی یا ابزارهای ساده بصری بوده و اغلب فاقد دقت، یکپارچگی ساختاری و سرعت کافی هستند، ممکن است به حد کفایت پاسخگوی نیازهای کاربران نباشد، در نتیجه استفاده از مدل‌های به روز و هوشمند برای آن دسته از طراحان رابط کاربری که معمولاً پیش از پیاده‌سازی کامل رابط کاربری، از موکاپ‌ها برای نمایش ساختار به صورت تصویری و همچنین چینش عناصر استفاده می‌کنند، می‌تواند تا حد زیادی کاستی‌های ناشی از رویکردهای متداول قدیمی را برطرف کند. یکی از این مدل‌ها، مدل‌های یادگیری ماشین هستند که به دلیل کارآمد بودنشان به ویژه طی یک دهه اخیر، توجه بسیاری از محققان در حوزه‌های پژوهشی گوناگون را به خود جلب نموده است. از سوی دیگر، مدل‌های یادگیری ماشین که برای ارزیابی خودکار توسعه یافته‌اند نیز، یا صرفاً متکی به تصویر هستند و توانایی تحلیل معنایی در تمامی سناریوها را ندارند، یا تنها مبتنی بر متن بوده و فاقد درک بصری هستند. در نتیجه، فقدان سامانه‌ای جامع که بتواند توسط ویژگی‌های متنی و بصری به صورت ترکیبی، تحلیل دقیق و استاندارد از طراحی رابط کاربری ارائه دهد، به یک چالش اساسی در این حوزه تبدیل شده است.

علاوه بر این، ابزارهای موجود نیز معمولاً فاقد بازخوردهای تخصصی هستند، که همین امر نیز به نوبه‌ی خود می‌تواند منجر به بروز مشکلات خاصی در بخش‌های مشخص از طراحی و به دنبال آن، ارائه خروجی‌های کلی و سطحی شود. بنابراین، همانطور که پیش‌تر نیز اشاره شد، با توجه به این که طراحی رابط کاربری به یکی از حیاتی‌ترین مؤلفه‌های موفقیت محصولات نرم‌افزاری در عصر تبدیل شده است و همچنین، اغلب روش‌ها و ابزارهای موجود، توان انطباق کافی با استانداردهای بین‌المللی را از لحاظ سادگی، ثبات، قابلیت استفاده و دسترسی‌پذیری ندارند؛ این خلأ، نیاز به طراحی و پیاده‌سازی یک سامانه پیشرفته را ایجاد کرده است که با استفاده از معماری‌های چندوجهی (متن-تصویر)، بتوان بازخوردهای خودکار دقیق و کاربردی‌تر را به صورت بلادرنگ تولید نمود و هدف از ارائه‌ی پژوهش حاضر، در نظر گرفتن سامانه است که بتواند بازخوردهای خودکار تولید شده را جهت بهبود طراحی‌ها ارزیابی کند.

کیفیت تجربه کاربری، مستقیماً به چگونگی طراحی و انسجام رابط کاربری وابسته است و این مسئله در محصولات رقابتی همچون اپلیکیشن‌های تلفن همراه یا وبسایت‌های تجاری از اهمیت دوچندان برخوردار است. با افزایش پیچیدگی طراحی‌ها و تنوع دستگاه‌های مورد استفاده کاربران، ارزیابی کیفیت این رابط‌ها به یکی از دغدغه‌های اصلی طراحان، توسعه‌دهندگان و شرکت‌های نرم‌افزاری تبدیل شده است. بنابراین، در چنین بستری، نیاز به ارائه سیستم‌هایی که بتوانند به صورت خودکار و هوشمند کیفیت رابط‌های کاربری را تحلیل کرده و بازخورد مناسبی را ارائه دهند، بیش از پیش احساس می‌شود. در سال‌های اخیر، روش‌های مختلفی برای ارزیابی خودکار رابط کاربری ارائه شده‌اند که می‌توان آنها را در دو دسته‌ی کلی قرار داد: روش‌های قاعده‌محور و روش‌های مبتنی بر یادگیری ماشین. در روش‌های قاعده‌محور<sup>۱</sup>، مجموعه‌ای از قواعد طراحی همچون رعایت کنتراست رنگی، توازن در چیدمان المان‌ها<sup>۲</sup>، فاصله‌گذاری مناسب و اندازه فونت‌ها به صورت صریح تعریف می‌شوند و تمامی این قواعد نیز از طریق الگوریتم‌هایی مانند Canny برای شناسایی خطوط و نواحی، و همچنین K-Means برای خوشه‌بندی عناصر، بررسی می‌شوند [۱].

با آنکه این روش‌ها به دلیل سادگی و قابل فهم بودنشان مورد توجه اولیه قرار گرفته‌اند، اما به سرعت مشخص شد که این مدل‌ها در مواجهه با تغییرات در طراحی یا تعریف قواعد جدید، انعطاف‌پذیری لازم را ندارند و دچار فروپاشی عملکردی می‌شوند. از سوی دیگر، استفاده از الگوریتم یادگیری ماشین به منظور آموزش مدل‌های مبتنی بر این الگوریتم، به گامی موثر در جهت کاهش محدودیت روش‌های قانون محور نیز مبدل شده است. در این روش‌ها، طراح یا پژوهشگر باید پیشاپیش مجموعه‌ای از ویژگی‌های کلیدی طراحی مانند موقعیت نسبی عناصر، رنگ قالب، تنوع فونت یا فاصله بین اجزا را استخراج کرده و سپس، آن‌ها را به صورت مدل‌هایی نظیر ماشین بردار پشتیبان<sup>۳</sup> یا شبکه‌های عصبی<sup>۴</sup> ساده ارائه دهد. با این حال، برخی از پژوهش‌ها مانند [۱]، نشان داده‌اند که این فرآیند مستلزم تخصص بالای انسانی در تعریف ویژگی‌ها و در عین حال دارای مشکل تعمیم‌پذیری به طراحی‌های جدید است. علاوه بر این، این ویژگی‌ها اغلب برای موقعیت‌های خاصی بهینه می‌شوند و در صورت تغییر در فلسفه طراحی یا نیازمندی‌های پروژه، کارایی خود را از دست می‌دهند.

<sup>1</sup> Rule-Based

<sup>2</sup> Elements

<sup>3</sup> Support Vector Machine (SVM)

<sup>4</sup> Neural Networks

اما تحول بزرگ زمانی رقم خورد که شبکه‌های عصبی عمیق<sup>۱</sup> و به طور خاص، شبکه‌های کانولوشنی<sup>۲</sup> وارد میدان شدند؛ چرا که این شبکه‌ها با توانایی بالا در استخراج ویژگی‌های پیچیده بصری، توانستند قدمی اساسی در راستای تحلیل طراحی رابط‌های کاربری بردارند.

به عنوان مثال، مدل‌هایی مانند ResNet و Inception توانسته‌اند بدون نیاز به تعریف دستی ویژگی‌ها، و صرفاً بر اساس ورودی<sup>۳</sup> تصویر رابط کاربری<sup>۴</sup>، اطلاعاتی مانند تراکم المان‌ها، رنگ‌ها و ساختار چیدمان را به صورت خودکار استخراج کنند [۲]. اما همین مدل‌ها نیز به تنهایی قادر به درک برخی مفاهیم کیفی و معنایی طراحی نظیر سلسله مراتب بصری، قابلیت فهم کاربران، یا تطابق با اصول دسترس‌پذیری نبوده‌اند؛ زیرا این مفاهیم نیازمند ترکیب دانش متنی و بصری هستند.

پس از آن، با ظهور مدل‌های زبانی بزرگ<sup>۵</sup> همچون GPT و T5 که بر معماری مبدل‌ها<sup>۶</sup> استوار هستند، امکان تحلیل و نقد متون طراحی فراهم شد. این مدل‌ها با بهره‌گیری از مکانیزم خودتوجهی<sup>۷</sup> قادر هستند که وابستگی‌ها در متن را شناسایی کرده و مفاهیم ضمنی و ساختارمند را استخراج کنند [۳]. اما این مدل‌ها به صورت ذاتی قادر به درک مستقیم تصاویر در طراحی‌ها نیستند و نیاز به تبدیل تصویر به توصیف متنی یا الحاق آن با مدل‌های بصری دارند. از سوی دیگر، محدودیت پنجره زمینه<sup>۸</sup> این مدل‌ها در پردازش متن‌های طولانی یا تعداد زیادی عنصر طراحی، باعث شده است که درک جامع‌تری از کل رابط کاربری حاصل نشود و تنها بازخوردهایی سطحی ارائه گردد [۳]. بنابراین، برای غلبه بر این چالش، معماری‌های چند ماهیتی<sup>۹</sup> توسعه یافتند که با ترکیب شاخه‌های بصری و زبانی، توانسته‌اند تعامل میان متن و تصویر را به شکل مؤثرتری مدل‌سازی کنند. در این معماری‌ها، شبکه‌ای مانند ViT یا شبکه عصبی کانولوشنی برای استخراج ویژگی‌های تصویری رابط کاربری استفاده شده و این ویژگی‌ها با خروجی مدل زبانی ترکیب می‌شوند. به عنوان مثال، در پژوهش چن و همکاران [۲]، ویژگی‌های بصری استخراج شده از یک شبکه عصبی کانولوشنی با توصیف‌های زبانی کاربران ترکیب شده و مدل پیشنهادی این محققان توانسته است بازخوردهایی نزدیک به داوری انسانی ارائه دهد. استفاده از لایه‌های توجه متقاطع<sup>۱۱</sup> برای اتصال دو شاخه نیز، این امکان را به مدل داده است تا بین ویژگی‌های بصری خاص (مانند رنگ‌های متضاد یا المان‌های مزاحم) و مفاهیم متنی (مانند عدم خوانایی یا سردرگمی کاربر) ارتباط برقرار کند.

جدول ۱: مقایسه روش‌های ارزیابی خودکار رابط کاربری

| روش           | انعطاف‌پذیری | نیاز به داده ترکیبی | درک تصویر | پشتیبانی از متن | نیاز به ویژگی دستی |
|---------------|--------------|---------------------|-----------|-----------------|--------------------|
| Rule-Based    | ×            | ×                   | محدود     | ×               | ×                  |
| Feature-Based | ×            | ×                   | محدود     | ×               | ✓                  |
| CNN           | ✓            | ×                   | ✓         | ×               | ×                  |
| NLP-only      | ✓            | ×                   | ×         | ✓               | ×                  |
| Multi-Modal   | ✓            | ✓                   | ✓         | ✓               | ×                  |

همانطور که در جدول ۱، نیز قابل مشاهده است هر کدام از روش‌های مختلف ارزیابی رابط کاربری دارای مزایا و محدودیت‌هایی است. در همین راستا، با وجود این پیشرفت‌ها، هنوز یک کاستی مهم در مسیر توسعه سامانه‌های خودکار ارزیابی طراحی وجود دارد و آن هم چیزی نیست جز فقدان داده‌های ترکیبی با کیفیت بالا. بیشتر مجموعه داده‌های موجود، تنها شامل تصاویر موکاپ به همراه یک برچسب عددی کلی از کیفیت هستند و فاقد اطلاعات دقیق‌تر همچون موقعیت المان‌های مشکل‌دار یا بازخورد متنی اختصاصی برای هر بخش طراحی هستند [۳، ۱].

<sup>۱</sup> Deep Neural Networks (DNN)

<sup>۲</sup> Convolutional Neural Networks (CNNs)

<sup>۳</sup> Feed

<sup>۴</sup> Interface Image Feed

<sup>۵</sup> Large Language Model (LLM)

<sup>۶</sup> transformer architecture

<sup>۷</sup> Self-Attention

<sup>۸</sup> Context Window

<sup>۹</sup> Multi-Modal

<sup>۱۰</sup> Chen et.al.

<sup>۱۱</sup> Cross-Attention

فقدان چنین داده‌هایی باعث می‌شود که مدل‌های چند ماهیتی نتوانند به صورت ناحیه‌ای<sup>1</sup> بازخورد دهند و در نتیجه بازخورد آن‌ها کلی و برای بهینه‌سازی دقیق طراحی، غیر قابل استفاده باقی بماند. همچنین عدم وجود برجسب‌های مکانی<sup>2</sup> که بتوانند نشان دهند کدام بخش تصویر به چه بازخوردی مربوط می‌شود، باعث افت دقت در یادگیری هم زمان بصری-زبانی شده است. مسئله دیگری که در این حوزه کمتر به آن پرداخته شده، عدم هماهنگی بازخورد تولید شده با استانداردهای بین‌المللی طراحی رابط کاربری است. اصولی مانند راهنمایی‌های دسترسی‌پذیری محتوای وب<sup>3</sup> برای دسترس‌پذیری، راهنمای طراحی رابط کاربری انسانی<sup>4</sup> برای طراحی رابط اپل، یا Material Design برای سیستم عامل اندروید، معیارهایی جامع و ساخت یافته برای ارزیابی طراحی ارائه می‌دهند. اگر سامانه‌ای ارزیابی نتواند بازخورد خود را بر مبنای این استانداردها تولید کند، خروجی‌های آن برای طراحان بی ارزش یا گمراه‌کننده خواهد بود. متأسفانه بیشتر مدل‌های موجود تنها به شباهت ظاهری یا یادگیری مبتنی بر داده بسنده کرده‌اند و فاقد ماژول‌هایی برای تطابق بازخورد با این استانداردهای جهانی هستند.

در کنار این چالش‌ها، تعامل‌پذیری سیستم بازخورد نیز باید مورد توجه قرار گیرد. بسیاری از طراحان نیاز دارند بازخوردی دریافت کنند که نه تنها مشکلات را گوش‌زد می‌کند، بلکه پیشنهادها و فنی و عملی برای اصلاح نیز ارائه می‌دهد. تولید چنین بازخوردی نیازمند مدل‌هایی است که هم توانایی تحلیل دقیق ساختار طراحی را داشته باشند، و هم با زبان تخصصی طراحی آشنایی داشته و بتوانند توصیه‌هایی مبتنی بر اصول طراحی رابط کاربری ارائه دهند. توسعه چنین مدلی نیازمند ترکیب اطلاعات زبانی، بصری و فنی به صورت یکپارچه است.

از دیدگاه فنی، طراحی یک سامانه‌ی ارزیابی که بتواند موکاپ‌های گرافیکی رابط کاربری (اعم از وب و اپلیکیشن) را تحلیل کرده و نواحی دارای ایراد را به همراه دلایل فنی و پیشنهادات اصلاحی ارائه دهد، نیازمند معماری چند ماهیتی پیچیده با زیرسیستم‌های ادغام تصویر، متن و منطق است. استفاده از مدل‌های پیش‌آموزش دیده در دو شاخه‌ی بینایی و زبان، و اتصال آنها از طریق لایه‌های مکانیزم توجه در مدل‌های یادگیری عمیق مانند لایه توجه متقاطع و نگاشت (تبدیل)<sup>5</sup>، همراه با الگوریتم‌های تطبیق بر اساس استانداردهای طراحی، هسته‌ی اصلی این سامانه را شکل می‌دهد. همچنین، توسعه مجموعه داده‌ای شامل موکاپ‌های متنوع، بازخوردهای زبانی تخصصی، و اطلاعات مکانی المان‌های مشکل دار، بخش مهمی از فرآیند تحقیق خواهد بود. بنابراین، مسئله‌ی اصلی این پژوهش را می‌توان چنین تعریف کرد: «طراحی و پیاده‌سازی سامانه‌ای هوشمند بر پایه‌ی شبکه‌های عصبی چند ماهیتی برای ارزیابی طراحی رابط‌های کاربری و ارائه‌ی بازخوردهای تخصصی با اشاره دقیق به نواحی دارای ایراد، مبتنی بر استانداردهای جهانی طراحی رابط کاربری. هدف این سامانه آن است که با بهره‌گیری از معماری‌های یادگیری عمیق و ترکیب تحلیل بصری و زبانی، بازخوردهایی تولید کند که نه تنها کاربردی و قابل اجرا باشند، بلکه با اصول طراحی جهانی همخوانی داشته باشند. انتظار می‌رود این سامانه بتواند فرآیند ارزیابی طراحی را بهبود بخشد، کیفیت نهایی محصولات نرم‌افزاری را ارتقاء داده و هزینه‌های بازمینی طراحی را به طور قابل توجهی کاهش دهد. در ادامه این مطلب، تحقیقات انجام شده توسط محققان در زمینه طراحی رابط کاربری با ابزارهای مختلف، نتایج حاصل از پژوهش‌ها و مزایا و معایب رویکردهای پیشنهادی مرور می‌شوند.

## ۲- ادبیات پژوهش:

تکامل سریع توسعه نرم‌افزار منجر به پذیرش فناوری‌های پیشرفته جهت برآورده کردن کارآمد انتظارات مشتریان و افزایش عملکردهای مربوطه در این حوزه شده است. تبدیل مدل رابط کاربری به کد به عنوان روشی برای ساده‌سازی تولید بازخورد و به حداقل رساندن تلاش‌های کدگذاری به صورت دستی پدیدار شده است. به عنوان مثال، رویکرد ارائه شده توسط زرنسکی و هلسن<sup>6</sup>، شامل ایجاد مجموعه داده‌ای از اجزای مشترک، و تبدیل آنها به کد با استفاده از الگوریتم‌های ترکیبی و YOYO است. در حالی که مدل ترکیبی ارائه شده در این مقاله، نتایج کمی متفاوت از طراحی اولیه ارائه می‌دهد، رویکرد YOYO دقت بالاتری را در شناسایی اشیاء نشان داده است [۴]. به طور مشابه، دیمیسیوآ و همکاران<sup>۷</sup>، روشی که ترکیبی از الگوریتم بینایی ماشین و یادگیری ماشین برای تولید برنامه‌های کاربردی از رابط‌های کاربری گرافیکی بوده است را در تحقیق خودشان پیشنهاد کرده‌اند [۵]. این رویکرد شامل شناسایی مؤلفه‌های رابط کاربری گرافیکی، طبقه‌بندی آنها با استفاده از الگوریتم‌های یادگیری ماشین، و مونتاژ برنامه بر اساس سلسله مراتب و سبک‌های مؤلفه بوده است.

<sup>1</sup> Region-Specific

<sup>2</sup> Bounding Boxes

<sup>3</sup> Web Content Accessibility Guidelines (WCAG)

<sup>4</sup> Human Interface Guidelines (HIG)

<sup>5</sup> Projection

<sup>6</sup> Czarnecki, K. and Helsen, S.

<sup>7</sup> Dimbisoa, W.G.

با این حال، پیچیدگی‌ها و مشکلات مربوط به کیفیت تصویر، از جمله مواردی هستند که می‌توانند دقت و سازگاری را به خطر بیندازند. بدین ترتیب، کیفیت تصویر رابط کاربری طراحی شده به طور قابل توجهی بر نتایج خروجی تأثیر گذاشته و اهمیت کیفیت تصویر و انتخاب ابزار نیز به طور مشخصی بیان شده است [۶]. علاوه بر این، در این تحقیق به صنعت توسعه برنامه‌های کاربردی وب اشاره شده است که از همان ابتدا در حال توسعه نرم افزارهایی بوده است که نیاز به خودکارسازی دارند. به طور معمول، اجزای برنامه وب به صورت دستی و به موازات تحقیق و توسعه فریم‌های رابط کاربری ایجاد می‌شوند که امکان طراحی در ابزارهای نمونه سازی ویژه وجود دارد.

لازم به توضیح است که طی سال‌های اخیر، پیشرفت در توسعه نرم افزار، بکارگیری تکنیک‌های خودکارسازی را برای افزایش کارایی ضروری کرده است. یکی از جنبه‌های چالش برانگیز این فرآیند، طراحی رابط کاربری است که توسعه‌دهندگان و محققان را بر آن می‌دارد تا روش‌های جدیدی را برای تبدیل طرح‌های اولیه یا نمونه‌های اولیه رابط کاربری به کد کشف کنند. فرآیند توسعه رابط کاربری معمولاً با ایجاد وایرفریم‌ها، موکاپ‌ها یا نمونه‌های اولیه با استفاده از ابزارهای مختلف طراحی رابط کاربری مانند فیگما<sup>۱</sup> برای تجسم و طراحی رابط برنامه شروع می‌شود [۷]. این ابزارها طراحان را قادر ساخته است تا طرح‌بندی‌ها، طرح‌واره‌ها و عناصر رابط کاربری مانند دکمه‌ها و المان‌های متنی را ایجاد کنند که فیگما به دلیل ویژگی‌ها و انعطاف‌پذیری گسترده‌ای که دارد یک انتخاب محبوب به حساب می‌آید. با این حال، در حالی که فیگما فاقد قابلیت تبادل مستقیم کد منبع است، بنابراین می‌توان با در نظر گرفتن برخی از محدودیت‌ها افزونه‌هایی را برای این منظور نصب کرد. طراحان معمولاً بر روی ظاهر امر و بحث زیبایی‌شناسی تمرکز می‌کنند، در حالی که توسعه‌دهندگان عملکرد را در اولویت قرار می‌دهند که همین امر می‌تواند منجر به ناهماهنگی و عدم دقت شود. در نتیجه، خودکارسازی این فرآیند می‌تواند خطاهای ناشی از چنین اختلافاتی را کاهش دهد [۸]. چالش‌ها در پر کردن شکاف میان طراحان و توسعه‌دهندگان، تضمین ترجمه دقیق طرح‌های رابط کاربری به کد نهفته است [۹]. همکاری برای روشن شدن ابهامات و اطمینان از همسویی بین چشم انداز طراحی و اجرا ضروری است. با خودکار کردن بخش‌هایی از فرآیند توسعه رابط کاربری، خطاهای ناشی از ناسازگاری‌ها را می‌توان کاهش داد، به طوری که منجر به بهبود دقت و سازگاری در رابط کاربری نهایی برنامه می‌شود. نمونه‌سازی رابط کاربری شامل دو نوع می‌شود که عبارتند از: ماکت‌ها و وایرفریم‌ها [۱۰]. موکاپ‌ها، نمایش‌های گرافیکی با کیفیت بالا از محصول نهایی هستند که اجزا و عناصر را با رنگ‌ها با جزئیات نشان می‌دهند، اما فاقد تعامل هستند. برعکس، وایرفریم‌ها طرح‌هایی با دقت پایین هستند که بر ساختار رابط کاربری تمرکز دارند و با استفاده از اشکال و خطوط اصلی، بدون جزئیات طراحی بصری ایجاد می‌شوند [۱۱]. هم مدل‌ها و هم وایرفریم‌ها نقش مهمی در طراحی رابط کاربری دارند و به طراحان در ارائه محصول خود کمک می‌کنند و به سهام‌داران امکان می‌دهند قبل از شروع کدنویسی بازخورد ارائه کنند. این فرآیند تکراری، سفرهای سازی را با توجه به مشخصات ذی‌نفعان و انتظارات کاربر تضمین می‌کند. با این حال، تبدیل مدل‌های رابط کاربری به طرح‌ها و کدهای رابط کاربری می‌تواند به دلیل تعاریف مبهم و از دست دادن اطلاعات بالقوه چالش برانگیز باشد. برخلاف زبان مدل‌سازی یکپارچه<sup>۲</sup>، که از متغیرها و موجودیت‌ها برای انتزاع استفاده می‌کند، ماکت‌ها بر سناریوهای داده‌های واقعی تکیه می‌کنند و خطر گم شدن اطلاعات را در طول تبدیل افزایش می‌دهند [۱۲]. برای رفع این مشکل، مدل‌سازی برای تسهیل فرآیند تبدیل به صورت روان و دقیق ضروری است. با توجه به این رویکردهای مبتنی بر مدل، مدل‌سازی و متامدل‌سازی (که شرح مدل‌ها است) به فرآیند توسعه رابط کاربری از نظر افزایش دقت و همچنین ارائه یک فرآیند تبدیل کارآمد کمک می‌کند. علاوه بر این، این فرآیند قابلیت استفاده مجدد و بهینه‌سازی برای پروژه را نیز به ارمغان می‌آورد که یک مزیت برای توسعه‌دهندگان است.

متامدل‌ها در پردازش زبان طبیعی، یا به بیان دیگر، مدل‌های برتر به عنوان میانجی بین طراحی برنامه و کد مربوطه عمل می‌کنند و طراحان را قادر می‌سازند تا تمام عناصر برنامه را در سطح طراحی توصیف کنند و از درک توسعه‌دهندگان یا ابزارهای خودکار اطمینان حاصل کنند [۱۳].

هنری و همکاران<sup>۳</sup>، یک رویکرد مدل محور را معرفی کرده که ماکت‌های رابط کاربری را با استفاده از متامدل‌ها و مدل‌های انتزاعی به مدل‌ها یا کدها تبدیل کرده است [۱۴]. این مدل‌ها محتوا و عناصر رابط کاربری ساختاری را مشخص می‌کنند و تشخیص آسان‌تر را برای اهداف کدنویسی یا مدل‌سازی تسهیل می‌کنند. در پژوهش دیگری که توسط دی‌لانگ و همکاران<sup>۴</sup> انجام شده است نیز، محققان، راه حلی را به واسطه بکارگیری مدل‌های ابر داده برای تولید کد از عناصر رابط کاربری ارائه کرده‌اند که فرآیند اصلاح برنامه‌های کاربردی وب را ساده کرده است [۱۱]. ایده استفاده از فرمالیسم‌های مبتنی بر مدل در اتوماسیون رابط کاربری امیدوارکننده بود و رویکرد مدل محور، قابلیت نگهداری و انعطاف‌پذیری را در توسعه رابط کاربری افزایش می‌دهد. توسعه مبتنی بر مدل، یک چارچوب سیستماتیک و ساختار یافته برای طراحی، پیاده‌سازی و مدیریت اجزای رابط کاربری فراهم می‌کند.

<sup>1</sup> Figma

<sup>2</sup> Unified Modeling Language (UML)

<sup>3</sup> Henry, S.L.

<sup>4</sup> De Lange, P. et al.

با نمایش عناصر و رفتارهای رابط کاربری در مدل‌های انتزاعی، توسعه‌دهندگان می‌توانند به سطح بالاتری از وضوح، سازگاری و قابلیت نگهداری در طول فرآیند توسعه دست یابند [۱۵]. این رویکرد علاوه بر این که همکاری میان طراحان و توسعه‌دهندگان وب را ساده کرده است، اطمینان حاصل می‌کند که رابط کاربری نهایی با مشخصات طراحی مطابقت دارد و توسعه‌دهندگان را نیز قادر می‌سازد که اجزای رابط کاربری را ایجاد کنند که به راحتی قابل تطبیق و گسترش در برنامه‌ها و پلتفرم‌های مختلف باشد [۱۵].

توسعه‌دهندگان می‌توانند معماری‌های رابط کاربری انعطاف پذیر و ماژولار را از طریق خلاصه‌کردن جزئیات پیاده‌سازی رابط کاربری در مدل‌ها بسازند به طوری که نیازمندی‌های در حال تغییر و فناوری‌های در حال تکامل را در خود جای دهد. با تعریف مشخصات رابط کاربری در مدل‌ها، توسعه‌دهندگان می‌توانند سازگاری و انطباق با اصول کاربرپذیری، استانداردهای دسترسی و الگوهای طراحی را اعمال کنند. بر همین اساس هر راه حلی برای مسئله می‌تواند به عنوان تابعی نمایش داده شود که به برخی از داده‌های ورودی بستگی دارد و داده‌های خروجی خاصی را نیز در تئوری سیستم، در نظر می‌گیرد [۱۵]. رویکردهای تولید اجزای رابط کاربری که امروزه وجود دارند را می‌توان به گونه‌ای سیستماتیک کرد که برخی از داده‌های ورودی شامل (متن، تصویر، ماکت و غیره) که برای تولید داده‌های خروجی مورد نظر در نظر گرفته می‌شوند، با بکارگیری الگوریتم‌های مربوطه برای انجام این کار مورد استفاده قرار بگیرند. تجزیه و تحلیل تحقیقات مرتبط نشان می‌دهد که هیچ رویکرد کاملاً خودکاری برای پشتیبانی از زنجیره کامل اتوماسیون وجود ندارد، چرا که مشکلاتی با داده‌های ورودی، داده‌های در حال انجام یا الگوریتم‌هایی که باید داده‌های ورودی را پردازش کنند تا نتیجه خروجی داشته باشند، وابستگی‌ها، نیازهای فرمت خاصی دارند یا باید شامل راهنما باشند [۱۶].

## ۲-۱- بررسی پژوهش‌ها در حوزه‌ی ارزیابی خودکار رابط کاربری (UI):

لو و همکاران<sup>۱</sup> در پژوهشی جامع به بررسی ارزیابی خودکار رابطه کاربری پرداخته و تمرکز آنها بر برجسته‌سازی اهمیت نیازهای انسانی در طراحی UX نیز بوده است. یافته‌های حاصل از این تحقیق نشان داده است که تمرکز صرف بر صفحات منفرد رابط کاربری بدون در نظر گرفتن تعاملات کاربر، ارزش علمی ابزارهای بکار برده شده را برای طراحان UX کاهش می‌دهد. از معایب این پژوهش نیز می‌توان به عدم درک عمیق از ذهنیت طراحان و کمبود مجموعه داده‌های مناسب و معیارهای ارزیابی اشاره نمود که بدین ترتیب، با ایجاد ابزارهایی که به جای جایگزینی طراحان، به آنها در فرآیندهای طراحی کمک کنند، همچنین برگزاری دوره‌هایی جهت آشنایی طراحان با قابلیت‌ها و محدودیت‌های الگوریتم‌های هوش مصنوعی می‌توان این کاستی‌ها را مرتفع کرد [۱۷].

سوئی و حداد<sup>۲</sup> یک مدل یادگیری عمیق مبتنی بر معماری DenseNet201 را جهت ارزیابی خودکار رابطه‌های کاربری موبایل معرفی کرده‌اند. مدل پیشنهادی محققان با استفاده از ویژگی‌های بصری استخراج شده از تصاویر رابط کاربری و ترکیب آن با الگوریتم k-نزدیک‌ترین همسایه<sup>۳</sup> به دقت متوسط ۹۳ درصد در ارزیابی کیفیت رابطه‌های کاربری دست پیدا کرده است. بنابراین، دقت بالای مدل پیشنهادی نشان‌دهنده توانایی بالای آن در تحلیل ویژگی‌های بصری و پیش‌بینی کیفیت طراحی بوده است. اما مدل پیشنهادی، تنها یک برجسب کیفی ارائه داده و توضیح مشخصی درباره نقاط قوت و ضعف طراحی ارائه نکرده است. در نتیجه، ترکیب مدل با روش‌های تولید زبان به منظور ارائه بازخورد توصیفی و قابل درک برای طراحان می‌تواند در زمینه برطرف نمودن این کاستی موثر واقع شود. علاوه بر این، به منظور اتقاء مدل و ارائه بازخورد توصیفی به طراحان می‌توان از مدل‌های زبانی بزرگ استفاده کرد. بررسی‌ها نشان داده‌اند که ترکیب مدل‌هایی بینایی ماشین با مدل‌های زبانی بزرگ می‌تواند بازخوردهای توصیفی و سودمندی را در راستای بهبود طراحی رابط کاربری ارائه دهد [۱۸].

گرینجرا و همکاران<sup>۴</sup> در مقاله خود به تکنیک‌های هوش مصنوعی، به‌ویژه یادگیری ماشین، اشاره کرده که به طور فزاینده‌ای برای خودکارسازی ارزیابی رابط کاربری و تولید بازخورد مورد استفاده قرار گرفته‌اند. نویسندگان یادآور شده‌اند که این روش‌ها می‌توانند رفتار، احساسات و تعاملات کاربر را تجزیه و تحلیل کنند و بینش‌هایی را ارائه دهند که به بهبود طراحی رابط کاربری کمک می‌کند. به عنوان مثال، هوش مصنوعی توانسته است بر اساس الگوهای طراحی، وایرفریم تولید کند یا از طریق الزامات متنی، ماکت‌های بصری تولید کرده و فرآیند نمونه‌سازی اولیه را ساده‌سازی کند. شایان ذکر است که این اتوماسیون، چالش‌های طراحی کاربر محور را که اغلب به زمان و منابع گسترده‌ای برای بازخورد از کاربران واقعی نیاز دارد، برطرف کرده است. از مزایای روش پیشنهادی در این مقاله این است که هوش مصنوعی می‌تواند در تحلیل رفتار کاربران و ارائه بازخوردهای دقیق‌تر نقش موثری را ایفا کند ولی از سوی دیگر، ادغام هوش مصنوعی در فرآیندهای طراحی رابط کاربری نیازمند تخصص فنی زیاد و منابع قابل توجه است، در نتیجه

<sup>1</sup> Yuwen Lu., et.al.

<sup>2</sup> M. Soui., Z. Haddad.

<sup>3</sup> k- Nearest Neighbor (kNN)

<sup>4</sup> Julián Grigera, et.al.

جمع آوری داده‌های با کیفیت نیز می‌تواند چالش برانگیز باشد. این مقاله بر هر دو جنبه تولید خودکار رابط کاربری و تولید خودکار بازخورد از طریق ارزیابی رفتار کاربران اشاره داشته است [۱۹].

ترور جسکات<sup>۱</sup> در یک ثبت اختراع در سال ۲۰۲۰ میلادی، به ایجاد یک سیستم خودکار برای تولید رابط کاربری بر اساس تنظیمات یا انتخاب‌هایی که توسعه‌دهنده انجام می‌دهد پرداخته است. سیستم پیشنهادی این امکان را داده است که با انتخاب برخی از ویژگی‌ها و مولفه‌ها از جمله نوع فرم‌ها، رنگ‌ها، آیکن‌ها و ساختار تب‌ها، رابط گرافیکی مناسبی تولید شود. مزایای این پژوهش افزایش سرعت توسعه رابط کاربری، یکپارچه‌سازی توسط فریم‌ورک‌ها، انعطاف‌پذیری بالا و کاهش خطاهای انسانی ناشی از شبیه‌سازی بوده است. از سوی دیگر این امکان وجود دارد که نگهداری و توسعه این سیستم، با پیچیدگی‌هایی همراه شود. بنابراین، بکارگیری خوشه‌بندی<sup>۲</sup> و طبقه‌بندی<sup>۳</sup> بر اساس انتخاب‌های قبلی طراحان، می‌تواند بازخوردها و پیشنهادات بهتری به آنها ارائه دهد. همچنین، مدل‌های مبتنی بر یادگیری عمیق مانند شبکه‌های مولد متخاصم<sup>۴</sup> برای تولید قالب‌های جدید رابط کاربری بر اساس داده‌های واقعی اپلیکیشن‌ها می‌توانند استفاده شوند [۲۰].

## ۲-۲- بررسی پژوهش‌ها در حوزه‌ی تولید خودکار بازخورد:

در تحقیق انجام شده توسط جنی و مونیک<sup>۵</sup>، نویسندگان به تولید خودکار بازخورد در فرآیند طراحی رابط کاربری پرداخته و از ابزار FENIKs به جای تولید خودکار رابط کاربری، بر ارائه بازخوردهای آنی به طراحانی که مبتدی هستند تمرکز داشته است. بنابراین، می‌توان گفت یکی از اهداف کلیدی نویسندگان این مقاله، اعمال اصول طراحی صحیح بوده است. از سوی دیگر، FENIKs ممکن است قادر به پوشش‌دهی تمامی اصول طراحی رابط کاربری نباشد و بدین ترتیب، برخی از جنبه‌های مهم طراحی نادیده گرفته شوند که پیشنهاد می‌شود جهت بهبود این کاستی، با افزودن اصول طراحی به روز رسانی شده، جامعیت این ابزار افزایش پیدا کند. علاوه بر این، این ابزار در شبیه‌سازی تعاملات پیچیده و سناریوهای واقعی می‌تواند دارای محدودیت‌هایی باشد که با توسعه‌ی قابلیت‌های شبیه‌سازی می‌توان تعاملات پیچیده‌تر را هم مدل‌سازی نمود تا ابزار کاربردی‌تر شود [۲۱].

مارنکوف و همکاران<sup>۶</sup>، به بررسی یک مدل مفهومی و ارائه چارچوبی جهت ارزیابی خودکار بازخورد و قابلیت بکارگیری رابط‌های کاربری در مراحل طراحی و پیاده‌سازی وب پرداخته‌اند. رویکرد پیشنهادی محققان به توسعه‌دهندگان اجازه می‌دهد تا مشکلات را سریع‌تر شناسایی و سپس آنها را اصلاح کنند. از معایب چارچوب پیشنهادی نیز می‌توان به نیاز به بروزرسانی‌های مداوم دستورالعمل‌ها و احتمال تشخیص نادرست که حتی ممکن است منجر به نادیده گرفتن مشکلات واقعی کاربران شود اشاره کرد. بنابراین، ادغام الگوریتم‌های یادگیری ماشین می‌تواند به این چارچوب کمک کند تا با تحلیل داده‌های واقعی کاربران، دقت ارزیابی را بهبود بخشد [۲۲].

تحقیق ارائه شده توسط سوئی و همکاران<sup>۷</sup> بر روی تولید خودکار بازخورد تمرکز داشته و به توسعه یک افزونه جاوا جهت ارزیابی خودکار قابلیت استفاده<sup>۸</sup> رابط کاربری موبایل پرداخته است، به طوری که این ابزار با تحلیل کد رابط کاربری و بکارگیری معیارهای کمی، نقص‌های مربوط به رابطه کاربری را شناسایی کرده است. افزونه PLAIN با در نظر گرفتن تحلیل ایستا<sup>۹</sup> امکان ارزیابی سریع رابط‌های کاربری موبایل را بدون نیاز به مداخلات انسانی فراهم کرده است. همچنین، این ابزار توانسته است نقص‌های ماند اندازه نامناسب اجزا و ساختار ضعیف را شناسایی کرده و سپس به توسعه‌دهندگان گزارش دهد. نتایج ارزیابی این ابزار بر روی چهار پروژه متن باز نشان داده است که PLAIN توانسته به طور موثری رابط‌های کاربری را پیش‌بینی کند. از سوی دیگر، جنبه‌های زیبایی‌شناسی و تجربه کاربری به طور کامل در نظر گرفته نشده است [۲۳].

مقاله ارائه شده توسط (کلارو و همکاران)<sup>۱۰</sup>، به بررسی رویکرد مبتنی بر مهندسی مدل‌محور<sup>۱۱</sup> جهت ارزیابی خودکار قابلیت استفاده رابط کاربری پرداخته است. مزیت رویکرد پیشنهادی در این پژوهش، فراهم کردن امکان ارزیابی خودکار و سریع رابط‌های کاربری در مراحل اولیه توسعه‌ی نرم‌افزار بوده است. به عقیده نویسندگان این تحقیق، روش پیشنهادی آنها به خوبی با فرآیندهای توسعه مبتنی بر مدل، هماهنگ بوده و توانسته است به عنوان بخشی از چرخه‌ی توسعه‌ی نرم‌افزار مورد استفاده قرار بگیرد.

<sup>1</sup> Jaskot Trevor D

<sup>2</sup> Clustering

<sup>3</sup> Classification

<sup>4</sup> Generative Adversarial Networks (GANs)

<sup>5</sup> Jenny Ruiz and Monique Snoeck

<sup>6</sup> Jevgeni Marenkov, et.al.

<sup>7</sup> M. Soui, et.al.

<sup>8</sup> Usability

<sup>9</sup> Static Analysis

<sup>10</sup> Claro et.al.

<sup>11</sup> Model- Driven Engineering (MDE)

از سوی دیگر، برای رابط‌های کاربری با تعاملات پیچیده یا رفتارهای پویا، این رویکرد ممکن است نتواند به طور کامل مشکلات قابلیت استفاده (usability) را شناسایی کند که این امر نیازمند توسعه ابزارهایی است که بتوانند دامنه کاربرد این روش را از طریق مدل‌سازی رفتارهای پویا، افزایش دهند [۲۴].

فدریک و همکاران<sup>۱</sup> در مقاله خود به بررسی ابزار eGLU-Box که با ترکیب روش‌های خودکار و تعامل انسانی، فرآیند طراحی تست، جمع‌آوری داده‌ها و تحلیل نتایج را ساده می‌کند پرداخته‌اند. این ابزار به طور خاص برای نیازهای ادارات دولتی طراحی شده و به بهبود کیفیت خدمات دیجیتال کمک کرده است. علاوه بر این، eGLU-Box از روش‌های ارزیابی استاندارد از جمله پرسش‌نامه‌های SUS، UMUX-Lite و NPS بهره‌مند می‌شود که امکان مقایسه و تحلیل دقیق‌تر نتایج را فراهم می‌آورد. ولی از سوی دیگر، این ابزار به تولید خودکار رابط کاربری<sup>۲</sup> نمی‌پردازد و در ارائه بازخورد به صورت فوری و آنی به کاربران نیز می‌تواند دچار محدودیت‌های شود که همین امر امکان اثرگذاری بر تجربه کاربری را خواهد داشت. بنابراین، ادغان ابزارهایی مانند چت‌بات‌ها برای ارائه بازخورد فوری به کاربران در حین انجام تست‌ها می‌تواند سودمند باشد [۲۵].

هیلمن<sup>۳</sup> در تحقیقی به معرفی ابزار MeMo Wrkbench پرداخته است. پلتفرم در نظر گرفته شده در این مقاله برای ارزیابی خودکار قابلیت استفاده در سیستم‌های تعاملی به ویژه سیستم‌های چندرسانه‌ای مطرح شده است که از جمله مزایای آن، کاهش هزینه‌ها و زمان توسعه است. اما کارایی MeMo به نحوی است که به دقت مدل‌های رفتاری کاربران بستگی دارد و مدل‌های ناقص منجر به حصول نتایج نادرست می‌شوند [۲۶].

کازینو و همکاران<sup>۴</sup> به بررسی و اعتبارسنجی تجربی یک رویکرد خودکار برای ارزیابی قابلیت استفاده رابط‌های کاربری گرافیکی<sup>۵</sup> پرداخته‌اند. روش پیشنهادی محققان در این مقاله، با تمرکز بر ارزیابی جنبه‌های سازگاری رابط کاربری بدون نیاز به تحلیل کدهای پشتیبان<sup>۶</sup> طراحی شده است. از آنجایی که محدودیت‌هایی در ارزیابی رابط‌های کاربری پیچیده وجود دارد، ممکن است روش پیشنهادی نتواند مشکلات قابلیت استفاده را به صورت کامل شناسایی کند [۲۷]. همچنین به دلیل عدم پوشش دهی جنبه‌های ذهنی و احساسی تجربه کاربران، الگوریتم‌های نظیر BERT یا RoBERTa برای داده‌های متنی تولید می‌تواند استفاده شود چرا که این الگوریتم‌ها می‌توانند با ترکیب کردن داده‌های رفتاری و زبانی، وضعیت احساسی کاربران را شناسایی کرده و متناسب با وضعیت، بازخوردهای انسانی‌تری را تولید کنند. همچنین برای برطرف نمودن نیاز به مدل‌های دقیق برای ارزیابی داده‌ها، می‌توان از الگوریتم یادگیری تقویتی<sup>۷</sup> به منظور ارزیابی رفتار از طریق بازخورد کاربر و همچنین برای پیش‌بینی مشکلات بر اساس usability از روش‌های نظارت شده<sup>۸</sup> مانند جنگل تصادفی<sup>۹</sup> و گرادیان تقویت شده<sup>۱۰</sup> استفاده نمود [۲۷].

تارتا و مالداون<sup>۱۱</sup> در مقاله خودشان روشی را جهت ارزیابی خودکار قابلیت استفاده در برنامه‌های دسکتاپ با بهره‌گیری از AOP پیشنهاد کرده‌اند. از مزایای AOP امکان جداسازی منطق ارزیابی قابلیت استفاده از منطق اصولی برنامه است که به نگهداری و توسعه آسان‌تر کمک می‌کند. علاوه بر این، جنبه‌های تعریف شده برای ارزیابی می‌توانند در پروژه‌های مختلفی مورد استفاده قرار بگیرند به طوری که همین امر به افزایش بهره‌وری کمک می‌کند. ولی چنانچه ارزیابی ناقص یا سطحی از جنبه‌های کیفی تجربه کاربر رخ دهد، چندین راهکار وجود دارد که عبارتند از: بکارگیری نظرسنجی‌های درون‌برنامه‌ای<sup>۱۲</sup> و بهره‌گیری از الگوریتم یادگیری ماشین به نحوی که بتوان شاخص‌هایی مانند رضایت، سردگم شدن و نارضایتی را تخمین زد. بنابراین، مستندسازی، ابزار IDE و نیز تحلیل رفتار و احساسات کاربر می‌تواند در طراحی کمک‌کننده باشد [۲۸].

در پژوهش دیگری که توسط جنی و مونیگ<sup>۱۳</sup> انجام شده است، هدف اصلی نویسندگان بهبود فرآیند آموزش طراحی رابط کاربری برای طراحان تازه‌کار از طریق تولید بازخورد خودکار درباره تطابق طراحی‌هایشان با اصول طراحی رابط کاربری بوده است. همچنین، ابزار معرفی شده در این پژوهش، FENiKS بوده که یک ابزار شبیه‌سازی طراحی رابط کاربری با بکارگیری مدل‌سازی مبتنی بر مدل برای ارائه بازخورد به طراحان است.

<sup>1</sup> S. Federic., et.al.

<sup>2</sup> UI generation

<sup>3</sup> Stefan Hillmann

<sup>4</sup> R. Cassino., et.al.

<sup>5</sup> Graphical User Interface (GUI)

<sup>6</sup> Back-end

<sup>7</sup> Reinforcement Learning

<sup>8</sup> Supervised Learning

<sup>9</sup> Random Forest

<sup>10</sup> Gradient Boosting

<sup>11</sup> A. Tarta., G. Moldovan.

<sup>12</sup> In-app surveys

<sup>13</sup> Jenny Ruiz and Monique Snoeck

بررسی‌ها نشان می‌دهند که استفاده از این ابزار منجر به بهبود درک اصول طراحی در مقایسه با روش‌های سنتی بوده است [۲۹]. از سوی دیگر، چنانچه طراح، مدل را اشتباه تعریف کند، بازخورد اشتباهی هم دریافت می‌کند و تفاوت‌های میان کاربران نهایی از لحاظ مبتدی یا حرفه‌ای بودن نیز در نظر گرفته نشده است. بنابراین با ادغام رویکردهای انسانی که بازخورد خودکار و و بازخورد انسانی را ترکیب می‌کند می‌توان این کاسنی را مرتفع نمود. شبکه عصبی کانولوشنی برای طراحی تحلیل رابط کاربری گرافیکی و تشخیص مطابقت آنها با اصول طراحی می‌تواند سودمند باشد. همچنین، در نظر گرفتن یک مدل توصیه‌گر<sup>۱</sup> که مبتنی بر الگوریتم یادگیری تقویتی باشد می‌تواند برای طراحی بهتر مورد استفاده قرار بگیرد.

در پژوهشی که توسط دوان و همکاران<sup>۲</sup> انجام شده، بازخوردهای خودکار بر اساس اصول طراحی ارائه شده است. افزونه پیشنهادی در این مقاله با دریافت طراحی رابط کاربری و مجموعه‌ای از اصول طراحی، نویسندگان پیشنهاداتی را جهت اصلاحات متنی، شناسایی خطاهای ظریف و نیز بهبود جنبه‌های معنایی رابط کاربری ارائه کرده‌اند. ارائه مدل پیشنهادی توانسته است خطاهایی را شناسایی کند که ممکن است توسط انسان نادیده گرفته شوند. اما از سوی دیگر، همین مدل ممکن است نتواند تمامی جنبه‌های زمینه‌ای و هدف طراحی را به درستی درک کند. لازم به ذکر است که پس از چندین بار استفاده از ابزار پیشنهادی، این امکان وجود دارد که کاربران بازخوردها را تکراری یا غیر مفید تلقی کنند که سفارشی‌سازی بازخورد از لحاظ سادگی و دسترسی‌پذیری می‌تواند قابل بررسی باشد. ترکیب مدل بینایی ماشین با رویکرد پیشنهادی در این مقاله نیز به تحلیل بهتر تصویر و متن کمک‌کننده است [۳۰]. دوان و همکاران<sup>۳</sup> در تحقیق دیگری، از مدل‌های زبانی بزرگ استفاده کرده و به تولید خودکار بازخورد بر روی موکاپ‌های رابط کاربری پرداخته‌اند. محققان یک افزونه برای فیگما توسعه داده‌اند که با استفاده از (چی‌پی‌تی-۴)<sup>۴</sup> بازخورد خودکار را ارائه داده است و مزیت این افزونه، امکان دریافت بازخورد فوری برای طراحان بوده است. از سوی دیگر، مدل ممکن است نتواند درک مناسبی درباره هدف خاص طراحی داشته باشد و محدودیت‌هایی مانند کاهش کیفیت بازخورد، دور از انتظار نیست. بنابراین، مواردی مانند بکارگیری context-aware prompts بر اساس نوع محصول که چه برنامه کاربردی باشد و همچنین مخاطب هدف که در چه رده و بازه سنی قرار داشته باشد می‌تواند سودمند باشد. علاوه بر این، مدل‌های چندوجهی و استخراج ویژگی‌های تصویری از طریق شبکه‌های عصبی کانولوشنی و ترکیب آنها با prompt متنی، قابل بررسی است [۳۱، ۳۲]. تحقیق انجام شده توسط لسرووف و پتزنو<sup>۵</sup> روشی را برای ارزیابی خودکار رابط کاربری با استفاده از مدل‌های وظیفه<sup>۶</sup> و گزارش‌های مربوط به آزمایش‌های کاربر پیشنهاد داده است. این روش بازخوردی در مورد وظایف انجام‌شده، ناموفق و آزمایش نشده، خطاهای کاربر و اطلاعات مربوط به زمان ایجاد کرده که منجر به کمک به طراحان در شناسایی زمینه‌های مورد نظر می‌شود. ابزار توسعه‌یافته، USINE، در این تحقیق پیشنهاد شده که فرآیند ارزیابی را خودکار و به طراحان اجازه می‌دهد تا بر جنبه‌های مشکل‌ساز رابط کاربری تمرکز کنند. این رویکرد، اتکا به مشاهده و مصاحبه‌های سنتی کاربر را کاهش می‌دهد و فرآیند ارزیابی قابلیت استفاده را برای برنامه‌های پیچیده ساده می‌کند [۳۳]. روبال و همکاران<sup>۷</sup> به بررسی ارزیابی خودکار رابط‌های کاربری وب پرداخته که امکان بازخورد فوری را در طول تولید فراهم می‌کند.

این رویکرد به قابلیت استفاده و سازگاری بین پلتفرمی می‌پردازد و امکان آزمایش خودکار در برابر دستورالعمل‌های قابلیت استفاده مانند WCAG را فراهم می‌کند. با ثبت دانش قابلیت استفاده در یک هستی‌شناسی، این ابزار می‌تواند مشکلات را در مراحل اولیه فرآیند توسعه شناسایی کند، در نتیجه کارایی و اثربخشی طراحی رابط کاربری را بهبود می‌بخشد و در عین حال هزینه‌های مرتبط با ارزیابی‌های دستی را کاهش می‌دهد و از انطباق با استانداردهای قابلیت استفاده اطمینان حاصل می‌کند. در این مقاله، نویسندگان به طور مشخص به ارزیابی خودکار قابلیت استفاده رابطه‌های کاربری وب پرداخته‌اند. مزیت چارچوب پیشنهادی این است که به صورت خودکار توانسته است تغییرات در رابط کاربری را شناسایی کرده و ارزیابی‌های مربوطه را به روز رسانی کند. این سیستم علیرغم مزایایی که دارد در موارد مانند توسعه و نگهداری انتولژی‌ها می‌تواند دچار محدودیت شود [۳۴].

یئون و همکاران<sup>۸</sup> در مقاله تحقیقاتی خود یک سیستم طراحی و ارزیابی رابط کاربری ارائه داده‌اند که شامل یک سیستم بازخورد ارزیابی کاربرپذیری است. این سیستم، طراحی رابط کاربری را از طریق شبیه‌سازی حرکات بدن و مدل‌های رابط کاربری بهینه و امکان ارزیابی خودکار و تولید بازخورد را فراهم می‌کند. مزیت رویکرد پیشنهادی این تحقیق این بوده است که با تمرکز بر کاربرپذیری، دشواری‌های مربوط به قابلیت استفاده اولیه را کاهش می‌دهد و به کاربران اجازه می‌دهد تا راحت‌تر با رابط‌های جدید سازگار شوند.

<sup>1</sup> Recommendation System

<sup>2</sup> P. Duan., et.al.

<sup>3</sup> P. Duan., et.al.

<sup>4</sup> GPT-4

<sup>5</sup> Lecerof, A. and Paterno, F.

<sup>6</sup> task models

<sup>7</sup> Tarmo Robal, et.al.

<sup>8</sup> Yang Ung Yeon, et.al.

این رویکرد از نمونه‌سازی و ارزیابی سریع پشتیبانی می‌کند و کارایی توسعه محصول به ویژه برای محصولاتی با چرخه عمر کوتاه را افزایش می‌دهد. بکارگیری الگوریتم یادگیری ماشین جهت تحلیل داده‌های مربوط به تعامل کاربران می‌تواند محدودیت‌هایی که در زمینه درک نیروی انسانی نسبت به سیستم مبتنی بر هوش مصنوعی ممکن است وجود داشته باشد را برطرف کند [۳۵].

لازم به توضیح است که در این پژوهش، برای گردآوری جامع‌ترین مجموعه مقالات مرتبط با موضوع، از روشی نظام‌مند تحت عنوان مرور سیستماتیک ادبیات (SLR) استفاده شده که این روش شامل جستجوی گسترده در پایگاه داده‌های معتبر علمی همچون Google Scholar، Scopus و IEEE Xplore است. با هدف دستیابی به دقیق‌ترین نتایج ممکن، از یک مجموعه گسترده از کلید واژه‌های مرتبط با موضوع پیشنهادی در این تحقیق از جمله موکاپ، رابط کاربری، هوش مصنوعی و یادگیری ماشین استفاده شده است. در ادامه، جداول ۲ تا ۵ به بررسی مطالعات پیشین مرتبط با موضوع پژوهش حاضر، می‌پردازند.

جدول ۲: روش‌های مبتنی بر مدل‌های زبانی بزرگ

| مرجع | دقت                                      | نقاط ضعف                                                                               | نقاط قوت                                                                           | هدف اصلی مقاله                                                                                        | روش پیاده‌سازی دقیق                                                            |
|------|------------------------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| [۱]  | حدود ۵۲٪<br>پیشنهادهای دقیق<br>(تخمینی). | کاهش کارایی بازخورد<br>پس از چند تکرار و<br>ناقص بودن پاسخ‌های<br>ارائه شده.           | توانایی تشخیص اشکالات<br>ظریف طراحی و بهبود<br>متون و ویژگی‌های معنایی<br>در رابط. | تولید خودکار بازخورد<br>طراحی مبتنی بر ارزیابی<br>هیوورستیک رابط<br>کاربری.                           | بکارگیری GPT-4<br>(LLM) در یک افزونه‌ی<br>فیگما                                |
| [۲]  | ~۵۵٪ بهبود<br>عملکرد<br>(تخمینی).        | عملکرد فعلی هنوز از<br>ارزیابان انسانی پایین‌تر<br>است؛ نیاز به داده و<br>آموزش بیشتر. | بهبود چشمگیر (~۵۵٪)<br>در عملکرد LLM ها برای<br>تولید بازخورد طراحی.               | ایجاد مجموعه داده‌ای<br>شامل ۳۰۵۹ نقد طراحی<br>برای ۹۸۳ رابط و بهبود<br>بازخورد تولیدشده<br>توسط LLM. | آموزش LLM توسط<br>نمونه‌های اندک و داده‌های<br>آموزشی برای طراحی<br>(few-shot) |

جدول ۳: روش مبتنی بر شبکه‌های عصبی کانولوشنی

| مرجع | دقت    | نقاط ضعف                                                                                                  | نقاط قوت                                                                                                         | هدف اصلی مقاله                                                                                    | روش پیاده‌سازی دقیق                                       |
|------|--------|-----------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|-----------------------------------------------------------|
| [۳]  | ~۷۵٪   | عدم تشخیص برخی نواقص<br>ظاهری جزئی در رابط‌های<br>واقعی؛ نیاز به مجموعه داده<br>بزرگ و متنوع.             | بالاترین همخوانی با ارزیابی<br>طراحان انسانی در رتبه‌بندی<br>طراحی‌ها؛ دقت ~۷۵٪ در تشخیص<br>رابط بهتر.           | برآورد خودکار کیفیت<br>طراحی و سازگاری بصری<br>رابط با امتیازدهی عددی و<br>ارائه پیشنهادات طراحی. | شبکه عصبی تطبیقی<br>چندوجهی<br>(UIClip) با (CLIP)         |
| [۴]  | ۷۶/۰۵٪ | تمرکز صرف بر وظیفه<br>طبقه‌بندی و نیاز به مجموعه<br>داده تخصصی؛ ارائه بازخورد<br>کیفی مستقیم طراحی ندارد. | استفاده از معماری‌های پیشرفته<br>(EfficientNet-B1) در محیط<br>کم داده؛ دستیابی به دقت ۷۶٪ در<br>۵-shot learning. | طبقه‌بندی خودکار<br>صفحات UI و تشخیص<br>انواع خطاها در رابط‌های<br>اپلیکیشن.                      | یادگیری کم‌نمونه<br>(few-shot) با<br>شبکه‌های عمیق<br>CNN |

جدول ۴: روش مبتنی بر قوانین و شبیه‌سازی

| مرجع | دقت | نقاط ضعف                                                                          | نقاط قوت                                                                                | هدف اصلی مقاله                                                                                           | روش پیاده‌سازی دقیق                                          |
|------|-----|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|
| [۵]  | —   | مفهومی است و جزئیات پیاده‌سازی مشخصی ندارد.                                       | فراهم کردن چشم‌انداز سریع و کم‌هزینه برای بازخورد قابلیت استفاده در مراحل طراحی.        | ارائه مدل مفهومی و متریک‌های دسته‌بندی‌شده برای ارزیابی خودکار همزمان قابلیت استفاده وب.                 | چارچوب مفهومی مبتنی بر قواعد (مقاله نظری)                    |
| [۶]  | —   | محدود به رابط‌های وب و دستورالعمل‌های مشخص؛ نیاز به تعریف دقیق قوانین.            | امکان ارزیابی فوری و اقتصادی انطباق طراحی با استانداردها.                               | ارزیابی خودکار انطباق رابط‌های وب با دستورالعمل‌های قابلیت استفاده در مرحله طراحی.                       | ابزار خودکار ارزیابی در زمان طراحی (قواعد انطباق دستورالعمل) |
| [۷]  | —   | پیاده‌سازی پیچیده و نیاز به تعریف دقیق آنتولوژی؛ وابسته به دانش از پیش تعیین شده. | پوشش وسیع مفاهیم حوزه‌ی قابلیت استفاده با چارچوب معنایی منظم.                           | طراحی یک آنتولوژی برای مدل‌سازی دانش قابلیت استفاده وب و استفاده از آن در ارزیابی خودکار رابط.           | روش مبتنی بر آنتولوژی                                        |
| [۸]  | —   | تمرکز صرف بر جنبه‌ی آموزشی و طراحی اصولی؛ برای مقیاس صنعتی و کلی مناسب نیست.      | اثبات شده که استفاده از شبیه‌سازی و بازخورد خودکار، یادگیری دانشجویان را افزایش می‌دهد. | بهبود یادگیری اصول طراحی عملکردی UI با تولید خودکار اپلیکیشن نمونه و ارائه بازخورد آموزشی.               | شبیه‌سازی مدل‌محور (FENIKS)                                  |
| [۹]  | —   | تمرکز بر پشتیبانی آموزشی؛ محدود به سناریوها و قواعد از پیش تعریف شده.             | فراهم‌آوری بازخورد فوری مبتنی بر قوانین طراحی به‌هنگام طراحی نمونه UI.                  | ارائه ابزار تعاملی FENIKS برای بازخورد خودکار آنی در مورد اصول طراحی UI هنگام ساختن پروتوتایپ نرم‌افزار. | شبیه‌سازی مدل‌محور با بازخورد خودکار (FENIKS)                |

جدول ۵: روش ترکیبی (مدل‌های بینایی-زبان / Hybrid)

| مرجع | دقت | نقاط ضعف                                                                    | نقاط قوت                                                                                       | هدف اصلی مقاله                                                                                                        | روش پیاده‌سازی دقیق                                    |
|------|-----|-----------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------|
| [۹]  | —   | فعالاً محدود به ارزیابی زوجی؛ هنوز ابهامات در کاربرد عملی استراتژی استنتاج. | کاهش سوگیری موضعی و بهبود ثبات ارزیابی توسط این استراتژی استنتاج؛ نزدیک شدن به A/B تست انسانی. | ارزیابی مقایسه‌ای «متقاعدکنندگی» طراحی رابط کاربری در جفت‌های تصویر رابط با استفاده از مدل‌های زبان بزرگ چندرسانه‌ای. | مدل بینایی-زبان (استراتژی استنتاج-G-FOCUS بر پایه VLM) |

پژوهش انجام شده توسط Garde و همکاران، به بررسی پیش‌بینی میزان تلاش کاربر در تعامل با ویجت‌های فرم‌های وب پرداخته است که با کاستی‌هایی نیز همراه بوده از جمله عدم بررسی تنوع کاربران، عدم مقایسه با سایر معیارهای رابط کاربری و عدم بکارگیری داده‌های چندماهیتی [۱]. تحقیق Chen و همکاران، روشی نوآورانه برای جستجوی طراحی رابط کاربری بر اساس وایرفریم‌ها با استفاده از خودرمزگذارهای تصویری ارائه داده است که این رویکرد با هدف کمک به توسعه‌دهندگان پیشنهاد شده است که تجربه طراحی رابط کاربری محدودی داشته‌اند، تا بتوانند با جستجوی وایرفریم‌ها، به طراحی‌های با کیفیت بالاتری دسترسی پیدا کنند.

از سوی دیگر، عدم در نظر گرفتن تعاملات کاربر و بیشترین تمرکز مدل پیشنهادی بر روی شباهت‌های بصری که منجر به در نظر گرفته نشدن تعاملات کاربر با رابط کاربری شده است از نقاط ضعف این پژوهش بوده است [۲]. سیستمی نوآورانه برای پیشنهادات تعاملی در طراحی گرافیکی توسط O'Donovan و همکاران ارائه شده است. این سیستم با هدف کمک به طراحان مبتدی، پیشنهاداتی برای بهبود چیدمان عناصر گرافیکی ارائه کرده است اما ارزیابی کیفیت طراحی محدود بوده و بیشتر بر اساس نظرات سایر طراحان مبتدی انجام شده است [۳]. هلسن و زرنسکی، تلاش کرده با استفاده از یک مدل ویژگی‌محور، رویکردهای مختلف تبدیل مدل را دسته‌بندی و تحلیل کنند. ولی پژوهش این محققان بیشتر بر جنبه‌های ساختاری و نحوی تبدیل مدل‌ها تمرکز داشته و به جنبه‌های معنایی یا رفتاری آن‌ها کمتر پرداخته شده است [۴]. پژوهش انجام شده توسط Dimbisoa و همکاران با هدف تعریف یک متامدل برای مؤلفه‌های رابط کاربری به صورت مستقل از پلتفرم انجام شده که گامی در جهت تسهیل طراحی رابط‌های کاربری بوده است، ولی فقدان ارزیابی تجربی مناسب یا مطالعات موردی گسترده از معایب این تحقیق به شمار می‌رود [۵]. (تاریا گنوتیلا) پژوهشی را با هدف نقد دیدگاه‌های سنتی در فلسفه‌ی علم ارائه کرده است که در این رویکرد، مدل‌ها به صورت مفهومی در نظر گرفته شده و به گونه‌ای طراحی شده‌اند که بررسی سؤالات علمی خاصی را تسهیل می‌کنند. از سوی دیگر، به دلیل کمبود مطالعات تجربی در این پژوهش، این موضوع ممکن است قابلیت تعمیم یافته‌ها را محدود کند [۹].

## ۲-۳- مقایسه تحلیلی نتایج حاصل از مطالعات پیشین

همانطور که پیش‌تر اشاره شد، پژوهش‌های متعددی به بررسی روش‌های ارزیابی خودکار رابط کاربری پرداخته‌اند. به عنوان مثال، Soui و همکاران از مدل DenseNet201 به همراه الگوریتم kNN در پژوهش خودشان استفاده کرده‌اند و دقت بدست آمده در پیش‌بینی کیفیت طراحی، ۹۳ درصد بوده است، اما رویکرد پیشنهادی این محققان فاقد تولید بازخورد توصیفی بوده است [۲۳]. در مقابل، Duan و همکاران، افزونه‌ای برای Figma در مقاله خودشان طراحی کرده‌اند که با استفاده از GPT-4، بازخوردهای متنی برای موکاپ‌ها تولید کرده ولی پوشش تصویری کافی نداشته است [۳۱]. در ادامه، مقایسه‌ای که بر اساس معیارهای ارزیابی انجام شده است، در جدول ۶، قابل مشاهده است.

جدول ۶: مقایسه معیارها

| مرجع | مدل بکار برده شده                                       | پشتیبانی از متن | پشتیبانی از تصویر | بازخورد منطقه‌ای | تطابق با استانداردها | مشکلات روش پیشنهادی                                                                                                                     |
|------|---------------------------------------------------------|-----------------|-------------------|------------------|----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| [۲۳] | ترکیبی از شبکه‌های عصبی کانولوشنی و k- نزدیکترین همسایه | خیر             | بله               | خیر              | خیر                  | صرفاً بر اساس تصویر کار می‌کند و هیچ‌گونه بازخورد متنی به کاربر ارائه نداده است. بنابراین روش پیشنهادی، توان درک مفهومی طراحی را ندارد. |
| [۳۰] | مدل زبانی بزرگ (GPT-4)                                  | بله             | خیر               | خیر              | خیر                  | بازخوردها توصیفی و متنی بوده‌اند ولی به دلیل اینکه فقط از مدل زبانی استفاده شده است، فاقد تحلیل بصری و منطقه‌ای است.                    |
| [۲]  | تصویر و متن به صورت ترکیبی                              | بله             | بله               | بله              | خیر                  | مدل‌ها فقط بر اساس وایرفریم بوده‌اند و انطباق با استانداردهای طراحی از جمله WCAG یا HIG مشاهده نمی‌شود.                                 |
| [۲۲] | مدل قاعده‌محور به علاوه آنتولوژی <sup>۱</sup>           | نسبتاً محدود    | خیر               | خیر              | محدود                | روش پیشنهادی در این مقاله، قدرت تعمیم و دقت بالایی ندارد.                                                                               |

همان‌طور که در جدول ۶، مشاهده می‌شود، پژوهش‌های موجود هر کدام بخشی از نیازها را در تحلیل بازخوردها پوشش داده‌اند، اما یکپارچگی کامل در آن‌ها دیده نمی‌شود. بنابراین، پژوهش حاضر به دنبال توسعه سامانه‌ای است که هم از نظر چندوجهی بودن (متن-تصویر)، هم تولید بازخورد دقیق ناحیه‌ای، و هم انطباق با استانداردهای بین‌المللی طراحی، بهبود یافته‌تر باشد. همچنین، معیارهای ارزیابی عملکرد سامانه پیشنهادی شامل موارد زیر است:

- ✓ دقت پیش‌بینی طراحی: با استفاده از معیارهایی چون F1-Score و MAE
  - ✓ سرعت ارائه بازخورد: زمان تولید بازخورد سریع
  - ✓ کیفیت بازخورد متنی: ارزیابی توسط خبرگان طراحی با توجه به شفافیت، کاربردی بودن و انطباق با استاندارد انجام می‌شود.
  - ✓ قابلیت تشخیص ناحیه‌ای: بررسی توان مدل در تشخیص و اشاره دقیق به بخش دارای مشکل در طراحی.
  - ✓ تأثیر بر فرآیند طراحی: از طریق آزمون‌های A/B در محیط واقعی و مقایسه طراحی قبل و بعد از استفاده از سامانه خودکار پیشنهادی
- بکارگیری مدل‌های چندوجهی می‌تواند شکاف‌های تحقیقاتی فوق را مرتفع نماید.

### ۳- فرضیه‌ها یا سوالات پژوهش:

- چگونه می‌توان سامانه هوشمندی طراحی کرد که به واسطه‌ی تحلیل موکاپ رابط کاربری، قابلیت تولید بازخوردهای دقیق و منطبق با استانداردهای بین‌المللی طراحی را داشته باشد؟

#### ۳-۱- سوالات فرعی پژوهش

- چه معیارهایی برای ارزیابی رابط کاربری در نظر گرفته شود؟
- چه ویژگی‌هایی از هر نمونه ورودی (رابط کاربری) برای پیش‌بینی معیار ارزیابی استفاده شود؟
- چگونه بازخوردها و توضیحات برای بهبود طراحی واسط کاربری تولید شود؟

<sup>1</sup> Ontology

#### ۴- اهمیت و ضرورت پژوهش:

##### ۱) افزایش بهره‌وری و بهبود شاخص‌های تجاری

از آنجایی که بکارگیری پلتفرم‌های ارزیابی خودکار موکاپ‌ها منجر به کاهش مدت‌زمان طراحی و به بیان دیگر، صرفه‌جویی در زمان می‌شود، در پی آن هستیم که در این تحقیق، با ارائه بازخوردهای دقیق و استاندارد، کیفیت نهایی UI/UX را در مواردی نظیر نرخ حفظ کاربران و رضایت آنان بهبود بخشیده و افزایش دهیم.

##### ۲) کاهش زمان تحلیل و واکنش لحظه‌ای

استفاده از معماری‌های شبکه‌های عصبی کانولوشنی به همراه مکانیزم تنظیم پرامپت توسط نمونه‌های آموزشی اندک<sup>۱</sup> در مدل‌های زبانی بزرگ، امکان انجام ارزیابی و تولید بازخورد در زمان تقریباً واقعی<sup>۲</sup> را ایجاد می‌کند.

##### ۳) تحول در خلاقیت و نوآوری طراحی

این پژوهش با تکیه بر بکارگیری رویکردهای مبتنی بر هوش مصنوعی به صورت ترکیبی مانند تلفیق مدل‌های زبانی بزرگ و شبکه‌های عصبی درصدد است تا روشی نوین جهت ایجاد تحول در خلاقیت و نوآوری طراحی ارائه دهد. به طوری که این سیستم با تحلیل الگوی رفتاری کاربران و همچنین به کمک ابزارهای هوشمند، پیشنهادهای را ارائه دهد که منجر به تغییر نگرش طراحان، ایجاد فرصت‌های تازه برای خلق تجارب جدید و بهبود و ارتقاء فرآیندهای داده‌محور شود.

##### ۴) تحلیل جامع نقاط قوت و ضعف با سنجش چندمعیاره

بکارگیری سنجش چندمعیاره به موازات تحلیل‌های بصری و زبانی مدل‌های هوش مصنوعی باعث می‌شود بازخوردها به صورت دقیق در سامانه نهایی تولید شوند که همین امر در صورتی که برای طراحان در سطوح مختلف، قابل درک و پذیرش باشد و به توسعه‌دهندگان کمک کند تا به صورت هدفمند در بهبود رابط متمرکز شوند، بسیار موثر است.

#### ۵- اهداف پژوهش

توسعه و پیاده‌سازی سامانه خودکار بازخورد موکاپ‌های وب‌سایت و اپلیکیشن بر پایه‌ی مدل‌های یادگیری ماشین چندماهیتی با ترکیب شبکه‌های عصبی کانولوشنال و مدل‌های زبان بزرگ به منظور تولید بازخوردهای تخصصی در زمان نزدیک به لحظه<sup>۳</sup>، با دقت و تأخیر<sup>۴</sup> بهینه نسبت به روش‌های کنونی، و بهبود چشمگیر چرخه‌ی تکرار طراحی<sup>۵</sup>. در این رویکرد، الگوهای بصری و معنایی موجود در موکاپ توسط ساختاری مشترک از تعبیه<sup>۶</sup> تصویری و متنی شبیه‌سازی می‌شوند تا ضمن کاهش خطاهای طراحی تا حد ممکن، هزینه‌های پس‌اصحیح<sup>۷</sup> و زمان پاسخ سیستم به طور قابل توجهی کاهش یابد. پژوهش حاضر، در پی آن است که ابزار کاربردی یا به بیان دیگر، پلتفرمی را ارائه دهد که طراحان به کمک آن بتوانند در مراحل اولیه شروع طراحی، مسائل مربوط به رابط کاربری را شناسایی کرده، مشکلات را مرتفع کرده و همچنین، راه‌کارهایی را جهت بهبود فرآیند طراحی دریافت کنند. بدین ترتیب، معیارهای اصلی و اساسی باید استخراج شوند، الگوریتم‌ها و کلیه مدل‌هایی که به منظور تحلیل خودکار موکاپ‌ها کاربرد دارند توسعه پیدا کنند، بازخوردها به نحوی که برای طراحان در سطوح مختلف قابل درک باشد، تولید شوند و سپس، یک ارزیابی موثر از لحاظ میزان دقت و قابل اجرا بودن طرح در دنیای واقعی انجام شود.

#### ۵-۱- اهداف جزئی

۱) شناسایی و اولویت‌بندی مؤثرترین ویژگی‌های بصری و ساختاری موکاپ (نظیر چیدمان اجزا، کنتراست رنگ، سلسله‌مراتب بصری، خوانایی متن که بیشترین تأثیر را بر دقت پیش‌بینی مدل دارند و می‌توانند در بهینه‌سازی معماری پیشنهادی دخیل باشند).

<sup>1</sup> few-shot prompt tuning

<sup>2</sup> Near-real-time inference

<sup>3</sup> near-real-time

<sup>4</sup> latency

<sup>5</sup> design iteration cycle

<sup>6</sup> embeddings

<sup>7</sup> rework cost

- ۲) طراحی چارچوب ارزیابی چندمعیاره شامل معیارهای کمی F1-Score، واکنش زمانی و کیفی (معیار خودکار برای ارزیابی متن تولید شده توسط مدل‌های زبانی)<sup>۱</sup> برای بازخورد متنی، معیارها یا راهنمای دسترس‌پذیری محتوای وب<sup>۲</sup> جهت سنجش عملکرد سامانه بازخورد خودکار در مقایسه با روش‌های سنتی.
- ۳) تحلیل محدودیت‌های عملی و زیرساختی شامل پیچیدگی محاسباتی مدل، کیفیت و گستره‌ی داده‌های آموزشی، و پذیرش سیستم توسط طراحان حرفه‌ای از طریق سنجش قابلیت‌پذیری<sup>۳</sup> و پذیرش فناوری<sup>۴</sup>.
- ۴) مقایسه عملکرد سامانه‌ی پیشنهادی با رویکردهای موجود که قاعده‌محور هستند، مانند شبکه‌های تک‌مدلی و روش‌های غیر از مدل‌های زبانی بزرگ از منظر دقت بازخورد، تأخیر زمان اجرا و تأثیر بر کیفیت نهایی طراحی.
- ۵) بکارگیری تکنیک‌های fine-tuning و few-shot learning برای بهینه‌سازی لایه‌های توجه<sup>۵</sup> در هر دو شاخه‌ی بصری و زبانی، به منظور ارتقای دقت تولید بازخورد در شرایط کم داده‌ها.
- ۶) طراحی و پیاده‌سازی API منطبق با ابزارهای طراحی حرفه‌ای (Figma, Sketch) جهت یکپارچه‌سازی بی‌وقفه‌ی سامانه‌ی بازخورد در گردش کار<sup>۶</sup> تیم‌های طراحی.

## ۶- پیش‌بینی دستاوردها

در جدول ۷، دستاوردهای مورد انتظار این پژوهش فهرست شده‌اند.

جدول ۷: دستاوردهای مورد انتظار این پژوهش

| دسته‌بندی                | دستاورد مورد انتظار                                                                                      |
|--------------------------|----------------------------------------------------------------------------------------------------------|
| انتشارات علمی            | ارائه مقاله پژوهشی در یکی از ژورنال‌های معتبر                                                            |
|                          | ارائه مقاله در کنفرانس‌های بین‌المللی مرتبط با طراحی رابط کاربری خودکار و هوش مصنوعی                     |
| توسعه محصول و تجاری‌سازی | پیاده‌سازی وب‌سایت                                                                                       |
|                          | ارائه نسخه اولیه سرویس تحت وب (SaaS) برای شرکت‌ها و استودیوهای طراحی جهت ادغام در چرخه کار تیم‌های UX/UI |
| داده‌محوری و متن‌باز     | انتشار مجموعه داده موکاپ‌ها همراه با نقدهای کارشناسی                                                     |

## ۶-۱ بیان نوآوری اصلی پژوهش

نوآوری این پژوهش در ارائه‌ی رویکرد مبتنی بر یادگیری ماشین، جهت ارزیابی خودکار موکاپ‌های رابط کاربری و تولید بازخورد منطقه‌ای دقیق در زمان نزدیک به لحظه (بلادرنگ) است؛ چارچوبی که با ادغام شبکه‌های عصبی کانولوشنی برای تحلیل بصری و مدل‌های زبان بزرگ برای تولید توضیحات تخصصی، قابلیت کاهش هزینه و زمان چرخه تکرار طراحی را به طور چشمگیری فراهم می‌آورد.

## ۷- برخی از چالش‌ها و ملاحظات:

این امکان وجود دارد که در رابطه با شناسایی و درک صحیح عناصر بصری موکاپ‌ها برای کاربران چالش‌هایی ایجاد شود. می‌توان گفت از جمله چالش‌های اساسی، جمع‌آوری داده‌های مناسب است که دارای تنوع و کیفیت کافی برای تمامی حالات ممکن برای طراحی موکاپ‌ها باشد. علاوه بر این، دسترسی به مجموعه داده‌های آموزشی که جوابگوی نیازها باشد و همچنین تحلیل مولفه‌های مختلف در این حوزه از لحاظ میزان درک پیچیدگی‌های احتمالی، می‌تواند از جمله چالش‌های دیگر این تحقیق باشند. انتخاب مدل مناسب که دارای دقت بالا و در عین حال پیچیدگی کمی باشد، از این جهت که به زمان زیادی برای آموزش و ارزیابی نیاز نباشد، می‌تواند به مسئله چالش برانگیزی تبدیل شود.

<sup>1</sup> Bilingual Evaluation Understudy (BLEU)

<sup>2</sup> Web Content Accessibility Guidelines (WCAG)

<sup>3</sup> usability

<sup>4</sup> technology acceptance

<sup>5</sup> attention

<sup>6</sup> workflow

همچنین، بسیاری از مدل‌های هوش مصنوعی از جمله شبکه‌های عصبی عمیق ممکن است دارای مشکلاتی در تفسیرپذیری باشند که در درک نحوه عملکرد مدل در خصوص نوع بارخوردها، چالش ایجاد می‌کنند. اعتبارسنجی مدل‌ها نیز امری است که می‌تواند مشکلاتی را در بکارگیری مدل‌های مبتنی بر یادگیری ماشین برای طراحان ایجاد کند و پیچیدگی این فرآیند زیاد می‌شود.

## ۷-۱- راهکارهای مقابله با چالش‌ها:

برای رفع این چالش‌ها، می‌توان اقدامات زیر را انجام داد:

استفاده از مدل‌های پیشرفته مانند شبکه‌های عصبی کانولوشنی و مبدل‌ها به منظور تحلیل دقیق‌تر عناصر بصری موکاپ‌ها می‌تواند پیشنهاد شود. همچنین، توسعه الگوریتم‌ها در راستای طراحی مدل به نحوی که برای کاربران، قابل فهم، ساده و شفاف باشد توصیه می‌شود. چنانچه مدل‌ها به طور مشخص در زمینه طراحی، آموزش ببینند می‌توان در خصوص بالا رفتن دقت مدل و در نتیجه کفایت و کارآمدی بالای آن امیدوار بود. در خصوص پردازش داده‌ها<sup>۱</sup> نیز لازم است از روش‌هایی مانند نرمال‌سازی، بهینه‌سازی ویژگی‌ها و برطرف نمودن نویزها<sup>۲</sup> استفاده شود. علاوه بر این، بکارگیری مدل‌هایی که قادر به ترکیب ویژگی‌های مختلف هستند مانند ادغام مدل‌های یادگیری ماشین<sup>۳</sup> با شبکه‌های عصبی<sup>۴</sup> به صورت ترکیبی به منظور تحلیل ویژگی‌های متنی و پردازش تصویر<sup>۵</sup>، می‌تواند برخی از چالش‌ها را مرتفع سازد.

## ۸- روش تحقیق:



شکل ۱: مراحل روش تحقیق

همانطور که در شکل ۱، نیز مشاهده می‌شود؛ در مرحله ابتدایی، نیازهای سامانه و معیارهای کلیدی ارزیابی رابط کاربری (نظیر کنتراست رنگ، سلسله‌مراتب بصری، خوانایی متون، فاصله بین المان‌ها و ساختار چیدمان) از طریق تحلیل منابع علمی و مصاحبه با طراحان استخراج می‌شود. سپس این ویژگی‌ها به همراه داده‌های تصویری از موکاپ‌ها (شامل صفحات وب و اپلیکیشن) در قالب یک مجموعه داده ساختاریافته گردآوری می‌گردد. اطلاعات متنی مرتبط (از جمله توضیحات طراحی و برچسب‌های کیفی) نیز در صورت وجود، جهت استفاده در مدل‌های چندوجهی جمع‌آوری می‌شوند. در مرحله‌ی پیش‌پردازش، عملیات‌هایی از قبیل نرمال‌سازی تصاویر، استخراج ویژگی‌های مکان<sup>۶</sup>، حذف نویز، تبدیل فرمت‌ها، و تطبیق ورودی‌ها برای استفاده در مدل‌های یادگیری عمیق انجام می‌شود. همچنین، داده‌ها به صورت دقیق طبقه‌بندی و تقسیم‌بندی می‌شوند تا ساختار آموزشی و ارزیابی بهینه‌سازی گردد. در فاز آموزش، مدل‌های یادگیری ماشین شامل شبکه‌های عصبی کانولوشنی برای تحلیل بصری و مدل‌های زبان طبیعی مانند Transformers یا BERT برای تحلیل بازخورد متنی مورد استفاده قرار می‌گیرند. در صورت پیاده‌سازی مدل‌های چندوجهی، ترکیب تصویر و متن از طریق معماری‌هایی مانند CLIP یا ViLT بررسی می‌شود. عملکرد مدل‌ها با استفاده از شاخص‌های کمی (نظیر دقت، F1-Score، MAE) و کیفی (مانند ارزیابی انسانی یا معیارهای خوانایی) سنجیده می‌شود.

سامانه‌ی بازخورد خودکار نیز پیاده‌سازی شده و با استفاده از کتابخانه‌هایی نظیر PyTorch، OpenCV، و Transformers توسعه می‌یابد. در مرحله‌ی آزمایش، بازخوردهای تولید شده از لحاظ صحت، تطابق با استانداردهای طراحی، و میزان رضایت طراحان ارزیابی می‌شوند.

<sup>1</sup> Data processing

<sup>2</sup> Noise

<sup>3</sup> Machine Learning (ML)

<sup>4</sup> Neural Networks (NN)

<sup>5</sup> Image Processing

<sup>6</sup> bounding box

مقایسه تجربی بین عملکرد طراحی با و بدون سامانه بازخورد، به همراه تحلیل زمانی و کیفیت خروجی، انجام می‌پذیرد. برای تحلیل آماری، نمونه‌گیری از جامعه طراحان به صورت هدفمند یا تصادفی انجام شده و آزمون‌های آماری مناسب (مانند T-Test یا ANOVA بسته به سنجه‌ها) برای اعتبارسنجی تفاوت معنادار بین شرایط مختلف به کار گرفته می‌شود. در نهایت، مستندسازی کامل مراحل توسعه، به همراه انتشار کد و نتایج ارزیابی‌ها در قالب گزارش، مقاله علمی، و پلتفرم‌های متن‌باز انجام شده و زیرساخت لازم برای ادامه‌ی پژوهش یا پیاده‌سازی تجاری فراهم می‌شود.

## ۹- جدول زمان‌بندی:

| زمان بر حسب ماه |   |   |   |   |   |   | مراحل پژوهش                                                                        |
|-----------------|---|---|---|---|---|---|------------------------------------------------------------------------------------|
| ۱               | ۲ | ۳ | ۴ | ۵ | ۶ | ۷ |                                                                                    |
|                 |   |   |   |   |   |   | مطالعه و بررسی روشهای خودکار ارزیابی رابط کاربری                                   |
|                 |   |   |   |   |   |   | تحلیل نیازمندی‌ها و تعریف معیارهای ارزیابی                                         |
|                 |   |   |   |   |   |   | مرور و تحلیل روش‌های خودکار ارزیابی رابط کاربری                                    |
|                 |   |   |   |   |   |   | تهیه و پیش‌پردازش مجموعه داده‌ها                                                   |
|                 |   |   |   |   |   |   | آموزش و ارزیابی مدل‌های یادگیری ماشین جهت ارزیابی و طراحی مدل تولید بازخوردها      |
|                 |   |   |   |   |   |   | پیاده‌سازی سامانه تحلیل و ارائه بازخورد رابط کاربری                                |
|                 |   |   |   |   |   |   | اعتبارسنجی و بهبود افزایشی با مشارکت کاربر و ارزیابی بازخوردها و توضیحات تولید شده |
|                 |   |   |   |   |   |   | مستندسازی، انتشار و انتقال دانش                                                    |

1. Gardey, J.C., et al., *Predicting interaction effort in web interface widgets*. International Journal of Human-Computer Studies, 2022. 168: p. 102919.
2. Chen, J., et al., *Wireframe-based UI design search through image autoencoder*. ACM Transactions on Software Engineering and Methodology (TOSEM), 2020. 29(3): p. 1-31.
3. O'Donovan, P., A. Agarwala, and A. Hertzmann. *Designscape: Design with interactive layout suggestions*. in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015.
4. Czarnecki, K. and S. Helsen, *Feature-based survey of model transformation approaches*. IBM systems journal, 2006. 45(3): p. 621-645.
5. Dimbisoa, W.G., T. Mahatody, and J.P. Razafimandimby, *Creating a metamodel of UI components in form of model independant of the platform*. International Journal of Conceptions on Computing and Information Technology, 2018. 6: p. 48-52.
6. Draw.io, *Diagram drawing tool*, <https://app.diagrams.net>. . Figma, *The Collaborative Interface Design Tool*, <https://www.figma.com>.
7. Henry, S.L., S. Abou-Zahra, and J. Brewer. *The role of accessibility in a universal web*. in *Proceedings of the 11th Web for all Conference*. 2014.
8. Jokela, T., J. Ojala, and T. Olsson. *A diary study on combining multiple information devices in everyday activities and tasks*. in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015.
9. Knuuttila, T., *Modelling and representing: An artefactual approach to model-based representation*. Studies in History and Philosophy of Science Part A, 2011. 42(2): p. 262-271.
10. Lallemand, C., G. Gronier, and V. Koenig, *User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey*. Computers in human behavior, 2015. 43: p. 35-48.
11. de Lange, P., et al., *Integrating web-based collaborative live editing and wireframing into a model-driven web engineering process*. Data Science and Engineering, 2020. 5(3): p. 240-260.
12. de Carvalho Ferreira, B.S.L., *Automatic Generation of Synthetic Website Wireframe Datasets from Source Code*. 2020, Universidade do Porto (Portugal).
13. Mahmoudifar, F. (2023a) *Solution for Generation of Web Application User Interface Components From User Interface Mockup Source Code*, Master thesis, Riga Technical University, Latvia.
14. Mahmoudifar, F. (2023b) *Source code repository for the solution for generation of web application UI components from UI mockups*: <https://github.com/fmahmoudifar/uitocode>.
15. El Marzouki, N., et al., *Toward a generic metamodel for model composition using transformation*. Procedia Computer Science, 2017. 104: p. 564-571.
16. Mahendra, M.y.a., *The use of draw IO as digital mind map to improve sududents creativity and students concept mastery in learning human ifluence on ecosystem*. 2021, Universitas Pendidikan Indonesia.
17. Lu, Y., et al., *AI assistance for UX: A literature review through human-centered AI*. arXiv preprint arXiv:2402.06089, 2024.
18. Soui, M. and Z. Haddad, *Deep learning-based model using DensNet201 for mobile user interface evaluation*. International Journal of Human-Computer Interaction, 2023. 39(9): p. 1981-1994.
19. Grigera, J., J. Espada, and G. Rossi, *AI in User Interface Design and Evaluation*. IT Professional, 2023. 25: p. 20-22.
20. Jaskot, T.D., *User interface and front end application automatic generation*. 2020, Google Patents.
21. Ruiz, J. and M. Snoeck. *Feedback generation for automatic user interface design evaluation*. in *International Conference on Software Technologies*. 2021. Springer.
22. Marenkov, J., T. Robal, and A. Kalja. *A Study on Immediate Automatic Usability Evaluation of Web Application User Interfaces*. in *International Baltic Conference on Databases and Information Systems*. 2016.
23. Soui, M., et al. *PLAIN: PPlugin for predicting the usAbility of Mobile User INterface*. in *VISIGRAPP*. 2017.
24. Claro, S.M., et al. *Technology for Automatic Usability Evaluation Using Model Driven Engineering*. in *BPMDS/EMMSAD@CAiSE*. 2024.

25. Federici, S., et al. *Heuristic Evaluation of eGLU-Box: A Semi-automatic Usability Evaluation Tool for Public Administrations*. in *Interacción*. 2019.
26. Hillmann, S. *MeMo Platform for Automatic Usability Evaluation*. 2018.
27. Cassino, R., et al., *Empirical validation of an automatic usability evaluation method*. J. Vis. Lang. Comput., 2015. 28: p. 1-22.
28. Tarta, A. and G.S. Moldovan, *Automatic Usability Evaluation Using AOP*. 2006 IEEE International Conference on Automation, Quality and Testing, Robotics, 2006. 2: p. 84-89.
29. Ruiz, J. and M. Snoeck. *Automatic Feedback Generation for Supporting User Interface Design*. in *ICSOFT*. 2021.
30. Duan, P., J. Warner, and B. Hartmann. *Towards generating ui design feedback with llms*. in *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 2023.
31. Duan, P., et al. *UICrit: Enhancing Automated Design Evaluation with a UI Critique Dataset*. in *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 2024.
32. Duan, P., et al. *Generating automatic feedback on ui mockups with large language models*. in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024.
33. Lecerof, A. and F. Paternò, *Automatic support for usability evaluation*. IEEE transactions on software engineering, 1998. 24 (10): p. 863-888.
34. Robal, T., J. Marenkov, and A. Kalja. *Ontology design for automatic evaluation of web user interface usability*. in *2017 Portland International Conference on Management of Engineering and Technology (PICMET)*. 2017. IEEE.
35. Yang, U.Y., et al., *User interface design and evaluation system and hand interaction based user interface design and evaluation system*. 2008, Google Patents.



وزارت علوم، تحقیقات و فناوری  
معاونت پژوهش و فناوری

به نام خدا

### منشور اخلاق پژوهش

با استعانت از خدای سبحان و با اعتقاد راسخ به اینکه عالم محضر خداست و او همواره ناظر بر اعمال ماست و به منظور انجام شایسته پژوهش‌های اصیل، تولید دانش جدید و بهسازی زندگانی بشر، ما دانشجویان و اعضای هیأت علمی دانشگاه‌ها و پژوهشگاه‌های کشور:

تمام تلاش خود را برای کشف حقیقت و فقط حقیقت به کار خواهیم بست و از هر گونه جعل و تحریف در فعالیت‌های علمی پرهیز می‌کنیم.

حقوق پژوهشگران، پژوهیدگان (انسان، حیوان، گیاه و اشیاء)، سازمان‌ها و سایر صاحبان حقوق را به رسمیت می‌شناسیم و در حفظ آن می‌کوشیم.

به مالکیت مادی و معنوی آثار پژوهشی ارج می‌نهیم، برای انجام پژوهشی اصیل اهتمام ورزیده از سرقت علمی و ارجاع نامناسب اجتناب می‌کنیم.

ضمن پابندی به انصاف و اجتناب از هر گونه تبعیض و تعصب، در کلیه فعالیت‌های پژوهشی رهیافتی نقادانه اتخاذ خواهیم کرد.

ضمن امانت داری، از منابع و امکانات اقتصادی، انسانی و فنی موجود استفاده بهره‌ورانه خواهیم کرد.

از انتشار غیراخلاقی نتایج پژوهش نظیر انتشار موازی همپوشان و چندگانه (تکه ای) پرهیز می‌کنیم.

اصل محرمانه بودن و رازداری را محور تمام فعالیت‌های پژوهشی خود قرار می‌دهیم.

در همه فعالیت‌های پژوهشی به منافع ملی توجه کرده و برای تحقق آن می‌کوشیم.

خویش را ملزم به رعایت کلیه هنجارهای علمی رشته خود، قوانین و مقررات، سیاست‌های حرفه‌ای، سازمانی، دولتی و راهبردهای ملی در همه مراحل پژوهش می‌دانیم.

رعایت اصول اخلاق در پژوهش را اقدامی فرهنگی می‌دانیم و به منظور بالندگی این فرهنگ، به ترویج و اشاعه آن در جامعه اهتمام می‌ورزیم.

امضاء استاد راهنما

امضاء دانشجو