تکنیک‌های اصلی XAI که در این مطالعه به کار رفته‌اند عبارتند از:

•

**Permutation Importance (اهمیت جایگشتی)**1...: این تکنیک برای ارزیابی اهمیت ویژگی‌ها با اندازه‌گیری کاهش عملکرد مدل (مانند دقت) هنگامی که مقادیر یک ویژگی یا گروهی از ویژگی‌ها به صورت تصادفی در مجموعه داده آزمون جایگزین می‌شوند، استفاده می‌شود3.... در این مطالعه، دو رویکرد برای اهمیت جایگشتی به کار گرفته شد:

◦

**Single-feature Permutation Importance (اهمیت جایگشتی تک ویژگی)**4...: که تأثیر مستقل هر ویژگی را بر پیش‌بینی مدل می‌سنجد4....

◦

**Novel Set-of-features Permutation Importance (اهمیت جایگشتی مجموعه‌ای از ویژگی‌ها - روش جدید)**1...: این یک تکنیک جدید است که توسط این مطالعه معرفی شده1.... این روش با الهام از ماهیت ساختاریافته مجموعه داده Real-Life Trial که در آن ویژگی‌ها به صورت گروه‌های مرتبط دسته‌بندی شده‌اند (مانند حرکات ابرو، دست، دهان و غیره)، اهمیت کل یک گروه از ویژگی‌ها را به جای ویژگی‌های تکی ارزیابی می‌کند4.... این رویکرد برای جلوگیری از ایجاد ورودی‌های غیرواقعی به مدل، که ممکن است با جایگزینی تصادفی ویژگی‌های تکی در گروه‌هایی با وابستگی درونی رخ دهد، ضروری است4....

•

**Partial Dependence Plots (PDP) (نمودارهای وابستگی جزئی)**1...: این نمودارها به صورت گرافیکی رابطه بین یک یا دو ویژگی منتخب و نتیجه پیش‌بینی شده مدل را نشان می‌دهند4.... PDPها به درک چگونگی تأثیر تغییرات در مقادیر یک یا دو ویژگی بر پیش‌بینی مدل کمک می‌کنند و تأثیر تک ویژگی یا تعامل ویژگی‌ها را روشن می‌سازند4.... در این مطالعه، تحلیل PDP شامل بررسی:

◦

**Single Feature Impact (تأثیر تک ویژگی)**6...: نشان می‌دهد که چگونه تغییرات در یک ویژگی بر پیش‌بینی مدل تأثیر می‌گذارد6....

◦

**Feature Interaction Impact (تأثیر تعامل ویژگی)**6...: نشان می‌دهد که چگونه ترکیب دو ویژگی بر پیش‌بینی مدل تأثیر می‌گذارد6....

•

**SHapley Additive exPlanations (SHAP) (توضیحات افزایشی شاپلی)**1...: این روش مبتنی بر نظریه بازی است و سهم هر ویژگی را در پیش‌بینی مدل برای یک مورد خاص (تفسیر محلی) و همچنین اهمیت کلی ویژگی‌ها در کل مجموعه داده (تفسیر جهانی) اندازه‌گیری می‌کند4.... SHAP به روشن کردن این نکته کمک می‌کند که یک ویژگی پیش‌بینی مدل را به سمت فریب یا حقیقت سوق می‌دهد4.... SHAP هم تفسیر محلی و هم تفسیر جهانی را فراهم می‌کند6...:

◦

**Local Interpretability (تفسیر محلی)**6...: برای هر پیش‌بینی تکی، سهم هر ویژگی را در آن پیش‌بینی خاص مشخص می‌کند6....

◦

**Global Interpretability (تفسیر جهانی)**6...: با جمع‌آوری مقادیر SHAP در کل مجموعه داده، اهمیت کلی و تأثیر ویژگی‌ها را نشان می‌دهد6....

این تکنیک‌های XAI به کار گرفته شده، همگی **مدل-آگنستیک** هستند، به این معنی که می‌توانند برای تفسیر خروجی هر مدل یادگیری ماشینی استفاده شوند، بدون توجه به ساختار داخلی یا الگوریتم آن2.... این امر امکان اعمال آن‌ها بر روی مجموعه متنوعی از مدل‌های ML ارزیابی شده در مطالعه (MLP, SVM, XGBoost, CatBoost و غیره) را فراهم کرد3....