



طراحی سیستم هوشمند کشف تقلب بیمه درمان از طریق الگوریتم‌های هوش مصنوعی

مدرس: دکتر مجتبی فرخ



اطلاعات مدرس و همکاران

دکتر مجتبی فرخ، عضو هیات علمی گروه مدیریت عملیات و فناوری اطلاعات دانشکده مدیریت دانشگاه خوارزمی

دکتر سیروس شریفی، عضو هیات علمی گروه مدیریت عملیات و فناوری اطلاعات دانشکده مدیریت دانشگاه خوارزمی

دکتر نسرين حصارمقدم، عضو هیات علمی پژوهشکده بیمه

دکتر عباس راد، عضو هیات علمی گروه مدیریت صنعتی و فناوری اطلاعات دانشکده مدیریت و حسابداری دانشگاه شهید بهشتی

علیرضا نوروزی، دانش‌آموخته کارشناسی ارشد MBA اطلاعات دانشکده مدیریت دانشگاه خوارزمی

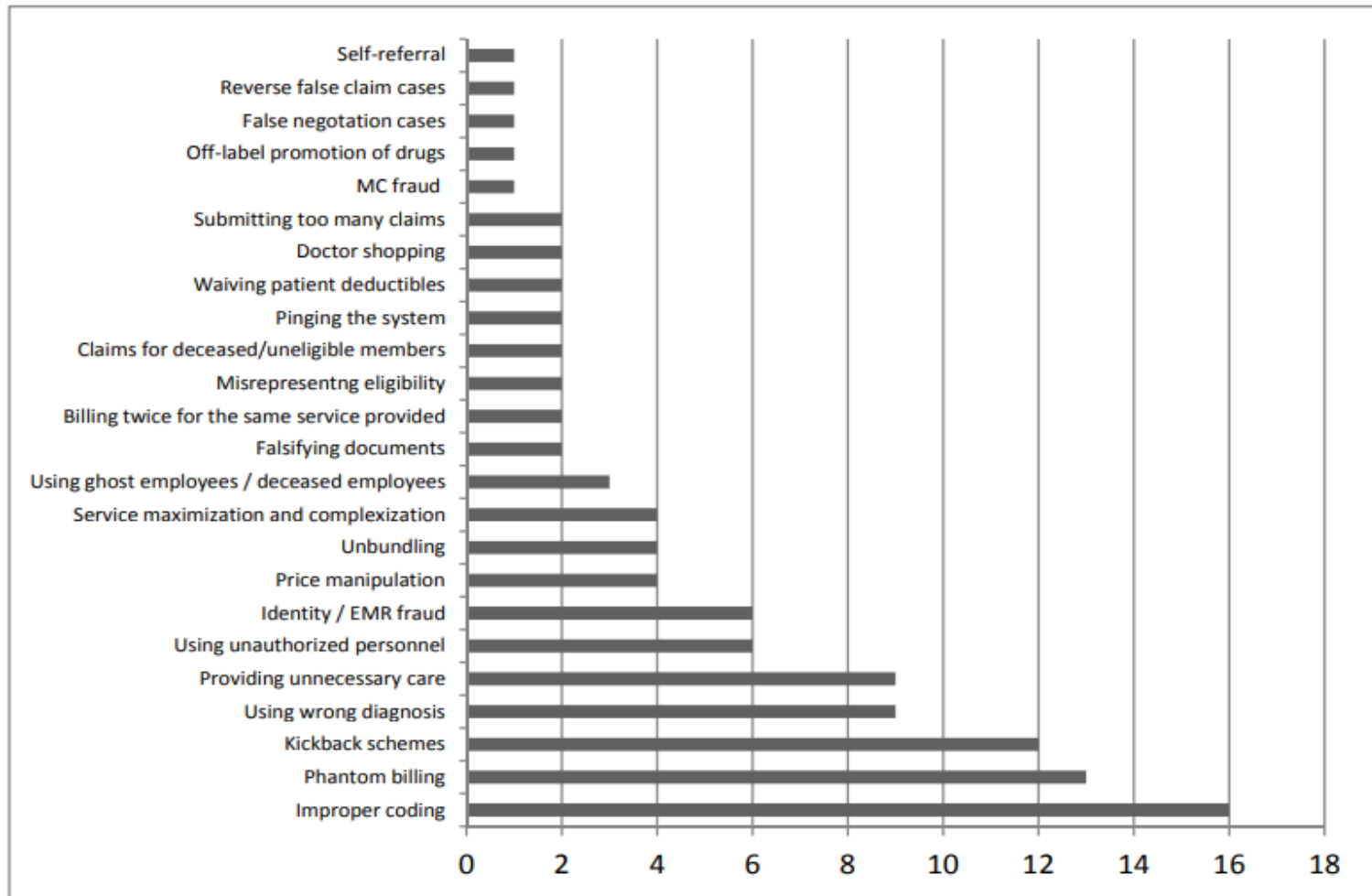
بیان مسأله و ضرورت انجام کارگاه

- در سال ۱۴۰۲، مبلغ خسارات پرداختی در صنعت بیمه کشور، حدود ۶۹،۵۸۲ میلیارد تومان بوده است.
- برآورد می‌شود که حجم خسارت‌های ناشی از تقلب در این سال بین ۶۹۵۰ تا ۱۰،۴۲۰ میلیارد تومان باشد.
- استفاده از الگوریتم‌هایی که قادر به بکارگیری بر روی داده‌های **بدون برچسب** باشد.
- تعریف شاخص‌های شناسایی انواع مختلف تقلب شامل تقلب‌های گروهی یا فردی
- طراحی سیستم هوشمند تقلب بیمه درمان که شامل ویژگی‌های «**پویایی** اکوسیستم تقلب»، «**تعاملی بودن تشخیص تقلب**» و «**قابل گسترش بودن**» را به صورت هم‌راستا با نیازمندی‌های دنیای واقعی پشتیبانی می‌کند.
- طراحی **نرم افزار** مناسب جهت کشف تقلب بیمه درمان

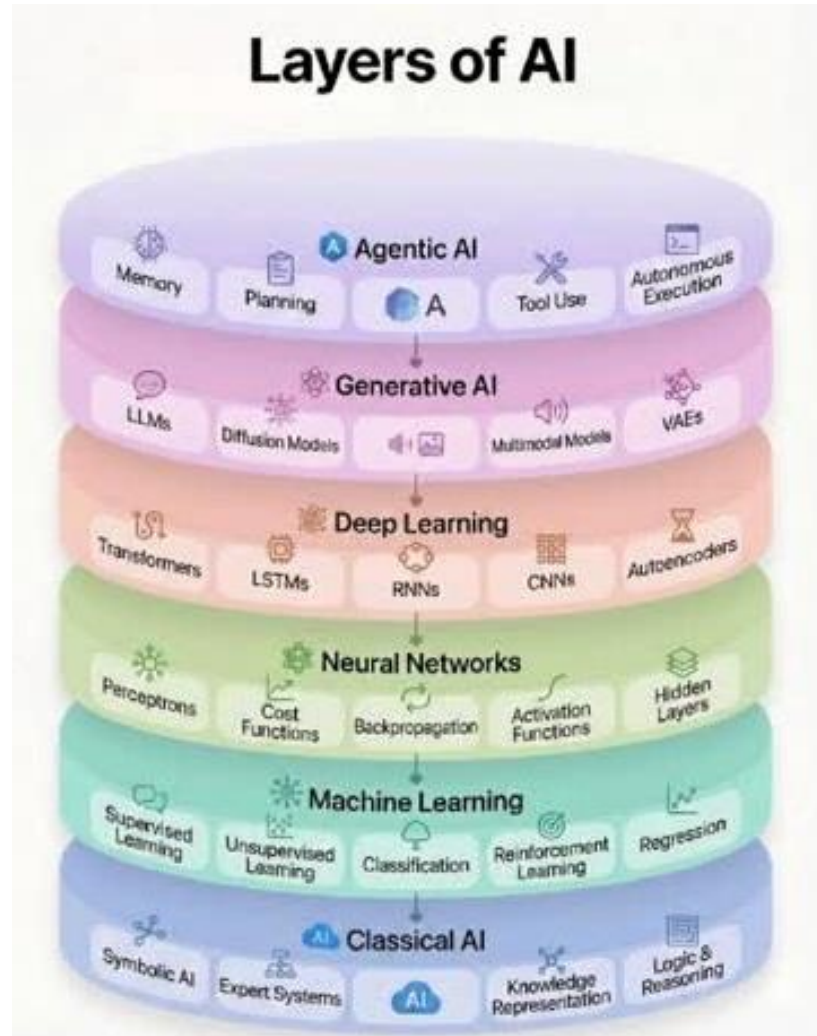
سوابق مطالعاتی و پژوهشی مربوطه

نویسندگان (سال)	مطالعه	رویکرد داده کاوی	نوع تقلب شناسایی شده	تکنیک‌های داده کاوی به کار رفته
Kose et al. (2015)	چارچوب جدید مبتنی بر یادگیری ماشین خودتعاملی برای تعیین ریسک هر نسخه	بدون نظارت	نسخه‌های پزشکی	یادگیر ماشین تعاملی مبتنی بر AHP و محاسبه شاخص ریسک نسخه
Haddad Soleymani et al. (2018)	متدولوژی سه مرحله شامل محاسبه ماتریس فراوانی، ماتریس ریسک برای جفت ویژگی‌ها و سپس تعیین ریسک نسخه‌ها	بدون نظارت	نسخه‌های پزشکی	ماتریس ریسک
Suesserman et al. (2023)	کشف ادعاهایی شامل کدهای رویه‌ای ناهنجار بکمک خودرمنگار عمیق و مقایسه نتایج با مدل خوشه‌بندی مبتنی بر چگالی	بدون نظارت	نسخه‌های پزشکی	Deep autoencoders
Hamid et al. (2024)	داده‌کاوی قوانین وابستگی به همراه تکنیک‌های یادگیری غیرنظارت‌شده برای شناسایی تقلب در بیمه درمانی	بدون نظارت	نسخه‌های پزشکی	Apriori algorithm and SVM
Tabassum et al. (2024)	کشف ناهنجاری در داده‌های مربوط به ثبت درمانی الکترونیک	بدون نظارت	تراکنش‌های درمانی	IF و LOF
De Meulemeester et al. (2025)	ترکیب جاسازی‌های دسته‌ای برای مقابله با متغیرهای دسته‌ای با تعداد حالت‌های بالا، تکنیک‌های ناهنجاری‌یابی بدون نظارت برای شناسایی ناهنجاری‌ها، و توضیحات افزودنی شاپلی برای تبیین خروجی	بدون نظارت	تقلب ارائه‌دهنده (داده‌های ادعای دندانپزشکی)	IF و OCSVM، LOF

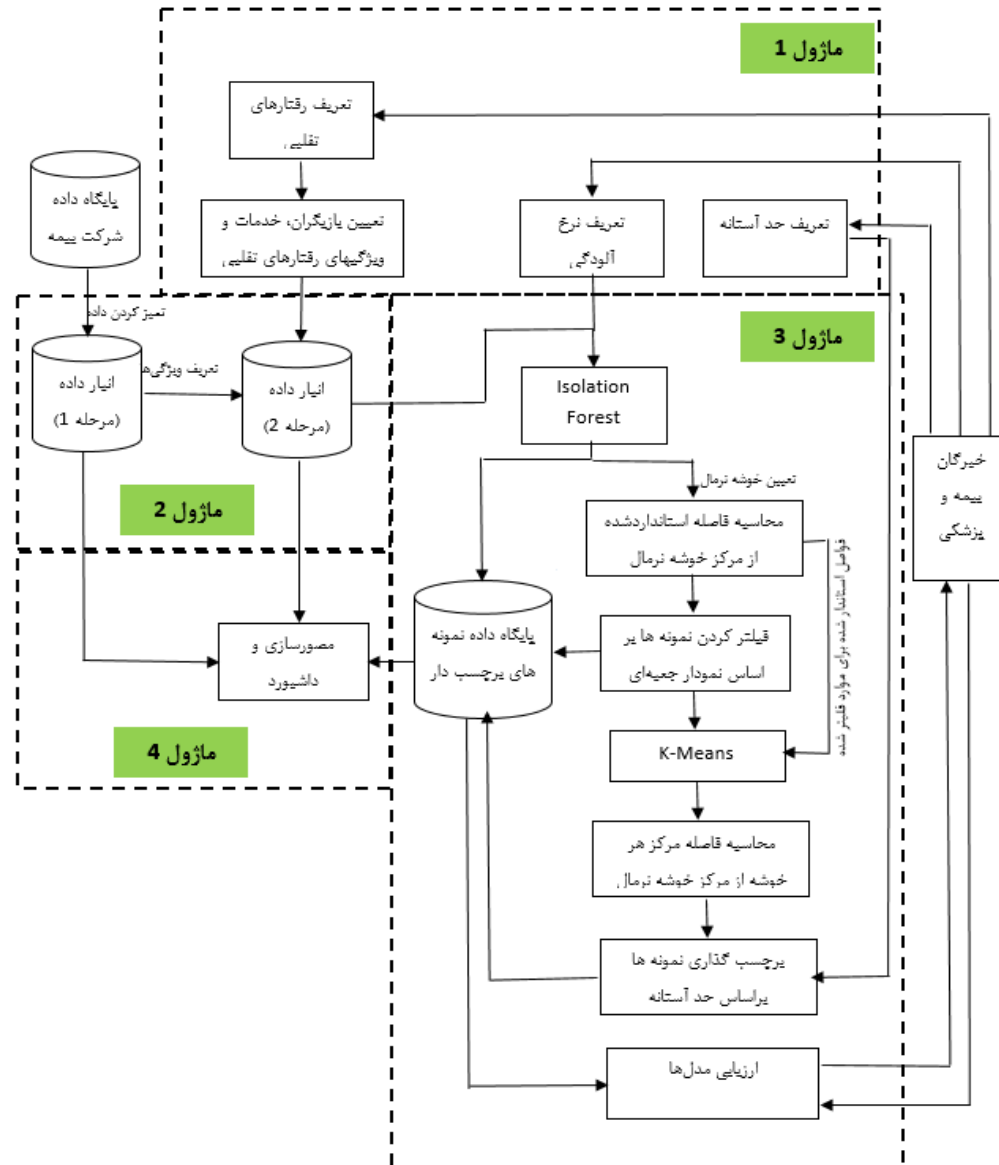
انواع تقلب بیمه درمان



میزان بروز انواع کلاهبرداری بیمه سلامت در مقالات (ترون و همکاران، ۲۰۱۵).



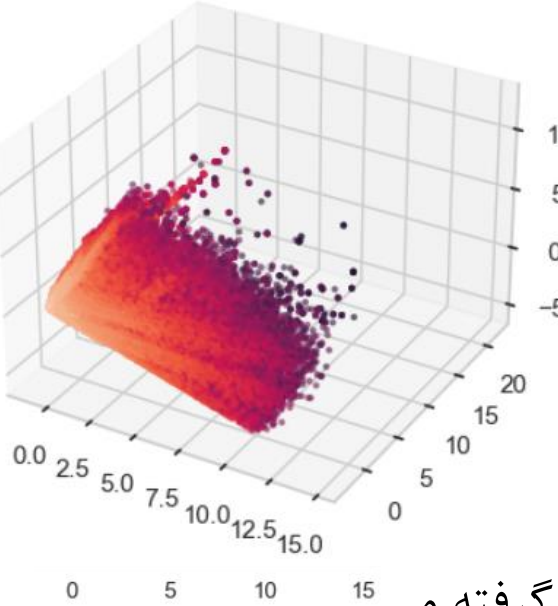
ساختار و ماژول‌های سیستم کشف قلب هوشمند بیمه درمان



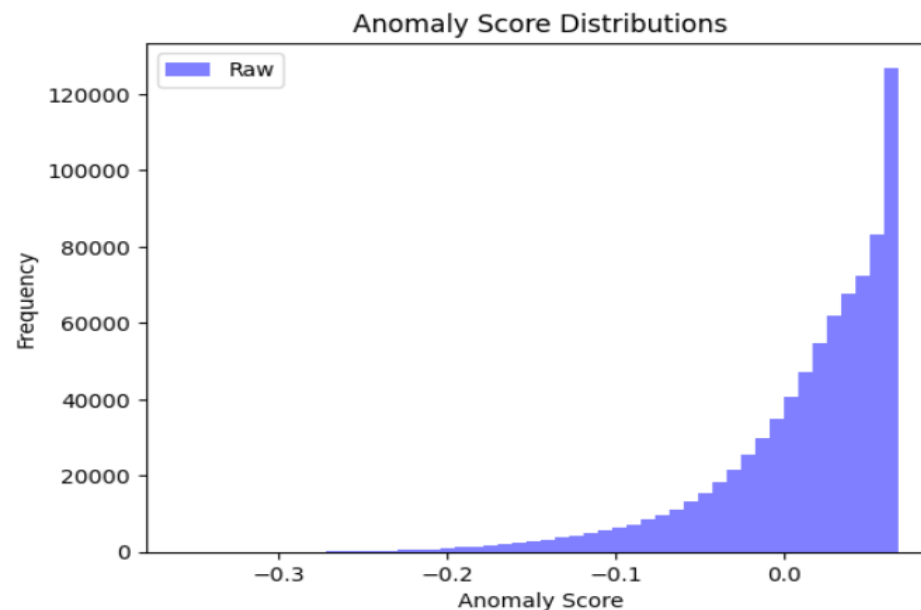
ساختار و ماژول‌های سیستم کشف تقلب هوشمند بیمه درمان

- بخش دانشی (ماژول ۱)
- انبار داده دو مرحله‌ای (ماژول ۲)
 - پیش‌پردازش داده‌ها
 - تعمیم‌پذیری مدل پیش‌بینی شامل مهندسی ویژگی، نرمال‌سازی داده‌ها، مدیریت داده‌های ناقص و ...
- موتور کشف تقلب هوشمند (ماژول ۳)
- مصورسازی و داشبورد (ماژول ۴)

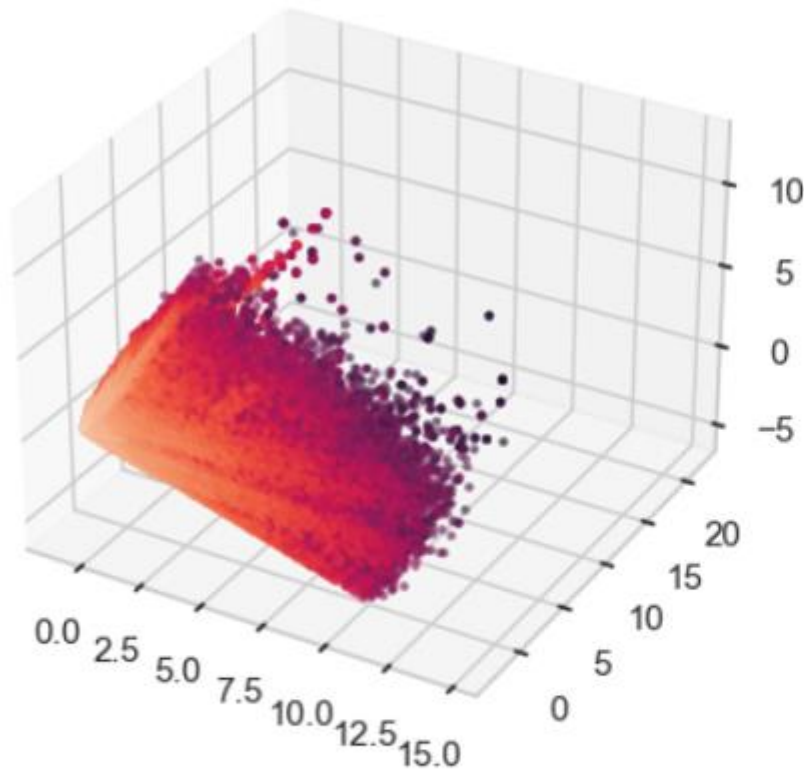
الگوریتم جنگل ایزوله



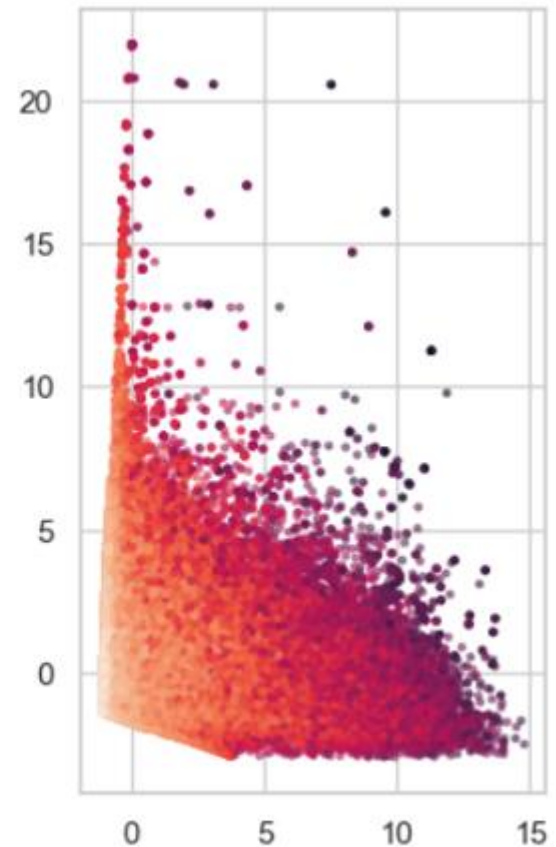
- از کتابخانه Optuna برای بهینه‌سازی هایپرپارامترها استفاده شد
- در این الگوریتم، مقدار هایپرپارامتر `_estimators` `max_features` به ترتیب ۲۵۷، ۰.۶۵۴، و ۰.۹۲۱ تعیین شد.
- نرخ آلودگی به عنوان پارامتر اصلی این الگوریتم ۲۰ درصد در نظر گرفته می‌شود.



الگوریتم جنگل ایزوله



ب. نمودار سه بعدی



الف. نمودار دوبعدی

احتمال تقلب (آنومالی) نقاط براساس نمره آنومالی در الگوریتم IF



مشخصات و محیط نرم افزار

طراحی این نرم افزار از دو بخش اصلی زیر تشکیل شده است:

1- API Backend (Flask)

2- Portal Frontend (React)

بخش بک اند از زبان برنامه نویسی Python 3.x با فریم ورک Flask استفاده و معماری آن بر پایه RESTful API با Blueprint Pattern بنا شده است تا یک API قدرتمند و گسترش پذیر فراهم کند. بخش فرانت اند نیز از زبان برنامه نویسی TypeScript با فریم ورک React ۱۹ استفاده کرده است.

بخش Portal Frontend با استفاده از React و TypeScript طراحی شده است تا رابط کاربری کاربر پسند و واکنش گرا ارائه دهد.

این مولفه از ابزارهای مدرن مانند Vite برای بیلد، React Router DOM برای مسیریابی، Axios برای مدیریت درخواستها و Recharts برای نمودارها به طور یکپارچه استفاده می کند تا تجربه کاربری بهینه ای ارائه شود. با تمرکز بر مقیاس پذیری بالا و کارایی، این سیستم نه تنها نمایشگرهای گرافیکی دقیق از نتایج تحلیل ها را فراهم می کند بلکه امکان اتصال امن و کارآمد به Backend و مدیریت داده های بیمه درمان را نیز فراهم می آورد.

برای ذخیره سازی داده ها از MariaDB استفاده شده و مدل تشخیص تقلب با الگوریتم پیاده سازی شده است تا بتواند به طور هوشمند و کارآمد الگوهای تقلبی را استخراج و تشخیص دهد.

با سپاسی از شما