

به نام خدا



K. N. Toosi University of Technology

دانشگاه صنعتی خواجه نصیرالدین طوسی

دانشکده برق

شناسایی سیستم

گزارش تمرین شماره ۲

[علیرضا یاحقی]

[۴۰۰۱۰۴۱۳]

استاد: آقای دکتر مهدی علیاری

خرداد ۱۴۰۴

فهرست مطالب

عنوان
سوال ۱ ۳
پاسخ سوال ۱ ۳
سوال ۲ ۵
پاسخ سوال ۲ ۵
سوال ۳ ۶
پاسخ سوال ۳ ۶
سوال ۴ ۶
پاسخ سوال ۴ ۷
سوال ۵ ۹
پاسخ سوال ۵ ۹
سوال ۶ ۱۰
پاسخ سوال ۶ ۱۰
شماره صفحه

سوال ۱

مقاله Credit Card Fraud Detection Using Autoencoder Neural Network در نظر گرفته شده است.

پس از مطالعه مقاله به سوالات زیر پاسخ دهید.

بزرگ ترین چالش ها در توسعه مدل های تشخیص تقلب چیست؟

این مقاله برای حل این چالش ها از چه روش هایی استفاده کرده است؟

پاسخ سوال ۱

چالش ها:

داده های نامتوازن

تراکنش های تقلبی (کلاس اقلیت) معمولاً بخش بسیار کوچکی از داده ها (کمتر از ۰.۱٪) را تشکیل می دهند. این عدم توازن باعث می شود مدل های یادگیری ماشین به سمت طبقه بندی نمونه ها به کلاس اکثربت (تراکنش های عادی) تمایل پیدا کنند، که منجر به نرخ تشخیص پایین برای تراکنش های تقلبی می شود.

پویایی رفتار های تقلبی

الگوهای تقلب به طور مداوم تغییر می کنند و تراکنش های تقلبی اغلب شبیه به تراکنش های قانونی هستند، که تشخیص آن ها را دشوار می کند.

کمبود داده های واقعی و دسترسی محدود

داده های واقعی تراکنش های کارت اعتباری به دلیل مسائل امنیتی و حریم خصوصی به سختی در دسترس هستند. این محدودیت باعث می شود مدل ها روی مجموعه داده های محدود یا مصنوعی آزمایش شوند.

نویز در داده ها :

داده های واقعی ممکن است شامل نویز (مانند خطاهای ثبت یا داده های ناقص) باشند که عملکرد مدل را کاهش می دهد.

راهکار ها:

داده های نامتوازن

از تکنیک **SMOTE (Synthetic Minority Oversampling Technique)** برای بیش نمونه گیری کلاس اقلیت (تراکنش های تقلبی) استفاده شده است. این روش با ایجاد نمونه های مصنوعی در فضای ویژگی ها، تعداد نمونه های کلاس اقلیت را از ۱۱۴ به ۲۲۵۳۸ افزایش داد تا با تعداد نمونه های کلاس اکثریت (تراکنش های عادی) برابر شود.

نتیجه : این کار توزیع کلاس ها را متعادل کرد و از تمایل مدل به طبقه بندی تمام نمونه ها به کلاس اکثریت جلوگیری کرد، که نرخ تشخیص (Recall) را بهبود بخشید.

پویایی رفتار های تقلبی

استفاده از خودمزگذار نویز زدا (DAE) به مدل کمک کرد تا ویژگی های مقاوم تری از داده ها استخراج کند. DAE با یادگیری بازسازی داده های اصلی از داده های نویزی، قادر به شناسایی الگوهای پنهان در تراکنش های تقلبی است، حتی اگر این الگوها شبیه به تراکنش های قانونی باشند.

نتیجه : ویژگی های استخراج شده توسط DAE مقاوم تر بودند و توانایی تعمیم دهنده مدل را افزایش دادند.

کمبود داده های واقعی

مقاله از مجموعه داده Kaggle استفاده کرد که شامل ۲۸۳۱۵ تراکنش با تنها ۵٪ تراکنش تقلبی بود. اگرچه این داده ها واقعی نبودند، اما با استفاده از SMOTE، نمونه های مصنوعی برای کلاس اقلیت ایجاد شدند تا کمبود داده جبران شود.

نتیجه : این روش به مدل امکان داد تا با داده های محدود کار کند و عملکرد بهتری در تشخیص تقلب داشته باشد.

نویز در داده ها

خودمزگذار نویز زدا با افزودن نویز گاووسی به داده های آموزشی و یادگیری بازسازی داده های اصلی، توانایی حذف نویز را به مدل افروزد. این مدل شامل ۷ لایه (۲۹-۲۲-۱۵-۱۰-۱۵-۲۲-۲۹ نورون) بود که داده های نویزی را به داده های تمیز تبدیل می کرد.

نتیجه : این روش داده های تمیز تری برای مرحله طبقه بندی فراهم کرد و دقت مدل را افزایش داد.

سوال ۲

در مورد معماری شبکه ارائه شده در مقاله به صورت مختصر توضیح دهید.

پاسخ سوال ۲

معماری شبکه ارائه شده در مقاله برای تشخیص تقلب در کارت‌های اعتباری شامل دو بخش اصلی است : خودرمزگذار نویززدا (Denoising Autoencoder - DAE) و شبکه عصبی کاملاً متصل

۱. خودرمزگذار نویززدا (DAE)

ساختار: شبکه‌ای ۷ لایه‌ای با معماری متقارن :

- لایه ورودی: ۲۹ نورون (منتاظر با ۲۹ ویژگی داده‌ها).

لایه‌های مخفی رمزگذار: ۲۲، ۱۵، ۱۰ نورون (کاهش ابعاد).

لایه‌های مخفی رمزگشا: ۱۵، ۲۲ نورون (بازسازی ابعاد).

لایه خروجی: ۲۹ نورون (بازسازی داده‌های ورودی).

داده‌های آموزشی با نویز گاوی آلوده شده و DAE برای بازسازی داده‌های اصلی آموزش داده می‌شود.تابع هزینه خطای مربع میانگین است.

هدف نویززدایی داده‌ها و استخراج ویژگی‌های مقاوم است.

۲. شبکه طبقه‌بند کاملاً متصل :

ساختار: شبکه‌ای ۶ لایه‌ای :

- لایه ورودی: ۲۹ نورون داده‌های نویززدایی شده از DAE.

لایه‌های مخفی: ۲۲، ۱۵، ۱۰، ۵ نورون.

لایه خروجی: ۲ نورون (برای دو کلاس: عادی و تقلبی).

از تابع SoftMax با تابع هزینه Cross-Entropy برای طبقه‌بندی استفاده می‌شود.

هدف طبقه‌بندی تراکنش‌ها به کلاس‌های عادی یا تقلبی است.

توضیح کلی ساختار:

داده‌های آموزشی ابتدا با **SMOTE** بیش نمونه‌گیری می‌شوند تا توزیع کلاس‌ها متعادل شود. سپس داده‌ها با **DAE** نویززدایی شده و ویژگی‌های مقاوم استخراج می‌شوند. در نهایت، داده‌های نویززدایی شده به شبکه طبقه‌بند وارد می‌شوند برای پیش‌بینی نهایی.

این معماری با ترکیب نویززدایی و بیش نمونه‌گیری، دقت و نرخ تشخیص (Recall) را برای کلاس اقلیت (تراکنش‌های تقلبی) بهبود می‌بخشد.

سوال ۳

مدل ارائه شده را پیاده سازی کرده و با استفاده از این دیتاست آموزش دهید. برای جلوگیری از بیش برآذش، آموزش مدل را طوری تنظیم کنید که در انتهای آموزش، بهترین وزن‌های مدل بر اساس خطای قسمت اعتبار سنجی بازگردانده شود.

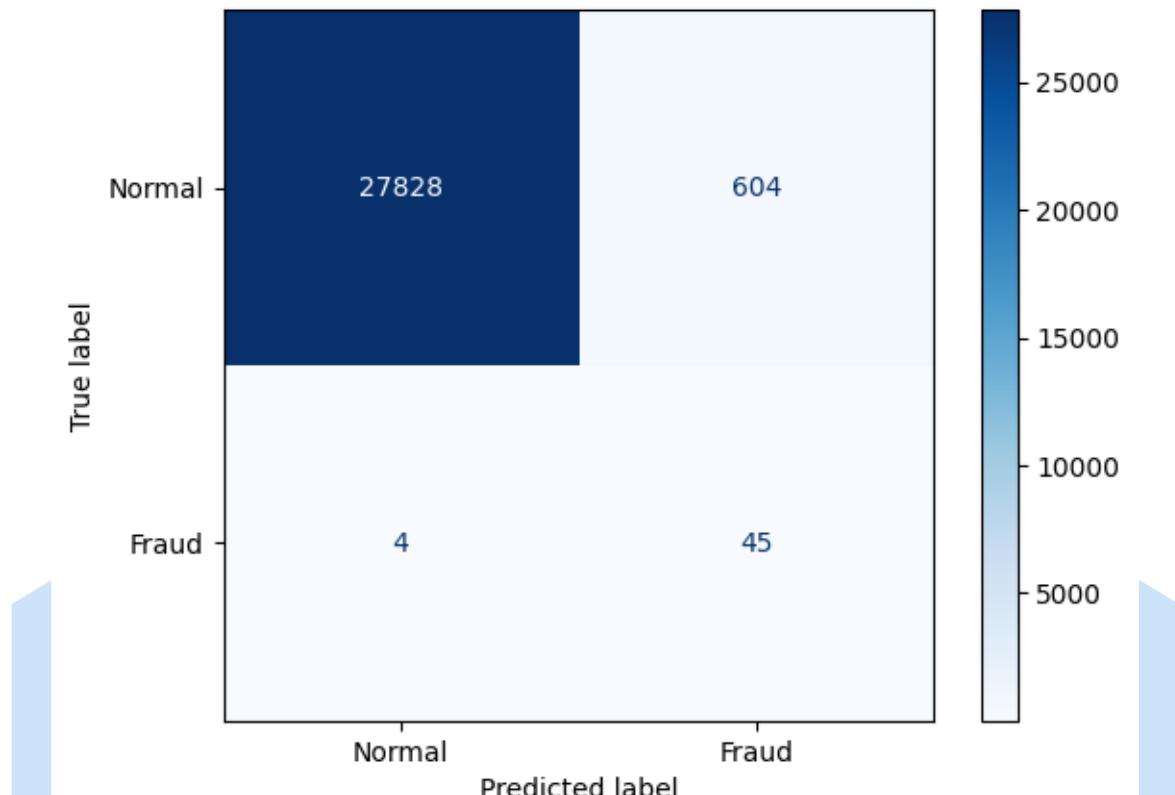
پاسخ سوال ۳

سوال ۴

ماتریس درهم ریختگی را روی قسمت آزمون داده‌ها رسم کنید و مقادیر **Precision**, **Accuracy**, **Recall** و **score1f** را گزارش کنید. فکر می‌کنید در مسائلی که توزیع برچسب‌ها نامتوازن است، استفاده از معیاری مانند **Accuracy** به تنها‌یی عمل کرد مدل را به درستی نمایش می‌دهد؟ چرا؟ اگر نه، کدام معیار می‌تواند به عنوان مکمل استفاده شود؟

پاسخ سوال ۴

Confusion Matrix at Threshold = 0.5



ماتریس در هم ریختگی نشان می دهد که مدل کلاس تقلب را تا حد خوبی پیش‌بینی می کند ولی بعضی از تراکنش‌های عادی را هم تقلب تشخیص می دهد.

	Precision	Recall	f1-score	Accuracy
Normal	1.00	0.98	0.99	0.98
Fraud	0.07	0.92	0.13	

کلاس عادی:

: تمام پیش‌بینی‌های مدل برای "عادی" صحیح بوده‌اند.

: مدل توانسته ۹۸٪ از نمونه‌های عادی را به درستی تشخیص دهد.

: تعادل عالی بین precision و recall وجود دارد.

نتیجه: مدل در تشخیص عادی‌ها بسیار عالی عمل کرده است.

کلاس تقلب:

precision = 0.07 فقط ۷٪ از پیش‌بینی‌های "تقلب" درست بودند؛ یعنی آلام‌های اشتباه زیاد دارد.

recall = 0.92 اما از بین تمام تقلب‌های واقعی، ۹۲٪ آن‌ها را پیدا کرده.

f1-score = 0.13 به خاطر precision پایین، امتیاز F1 هم پایین است.

نتیجه: مدل بسیار حساس است (تقلب‌ها را خوب پیدا می‌کند)، ولی دقت کمی دارد (خیلی از موارد عادی را هم تقلب می‌گوید).

accuracy = 0.98

۹۸٪ از کل پیش‌بینی‌ها صحیح‌اند. اما چون داده نامتوازن است (۲۸۴۳۲ مورد عادی و فقط ۴۹ تقلب)، این معیار می‌تواند گمراه‌کننده باشد.

در مسائل با داده‌های نامتوازن مانند تشخیص تقلب کارت اعتباری (که تنها ۵٪ نمونه‌ها تقلبی هستند)، Accuracy به تنها‌ی معیار مناسبی برای ارزیابی عملکرد مدل نیست.

بدلیل تسلط کلاس اکثریت، در دیتاست‌های نامتوازن، اگر مدل تمام نمونه‌ها را به عنوان کلاس اکثریت (غیرتقلبی) طبقه‌بندی کند، Accuracy همچنان بالا خواهد بود (مثلاً ۹۹.۵٪)، زیرا اکثر نمونه‌ها غیرتقلبی هستند. اما این مدل هیچ تراکنش تقلبی‌ای را تشخیص نمی‌دهد، که در مسائل تشخیص تقلب فاجعه‌بار است.

عدم حساسیت به کلاس اقلیت، Accuracy به تعداد کل پیش‌بینی‌های درست توجه دارد و اهمیت کلاس اقلیت (تراکنش‌های تقلبی) را نادیده می‌گیرد، در حالی که در این مسائل، تشخیص درست کلاس اقلیت حیاتی است.

معیار‌های زیر به عنوان مکمل بهتر هستند:

معیارها بهتر هستند؟

تمرکز بر کلاس اقلیت: Recall و F1-Score مستقیماً به عملکرد مدل روی کلاس اقلیت (تراکنش‌های تقلبی) توجه دارند، که در مسائل تشخیص تقلب اولویت دارد.

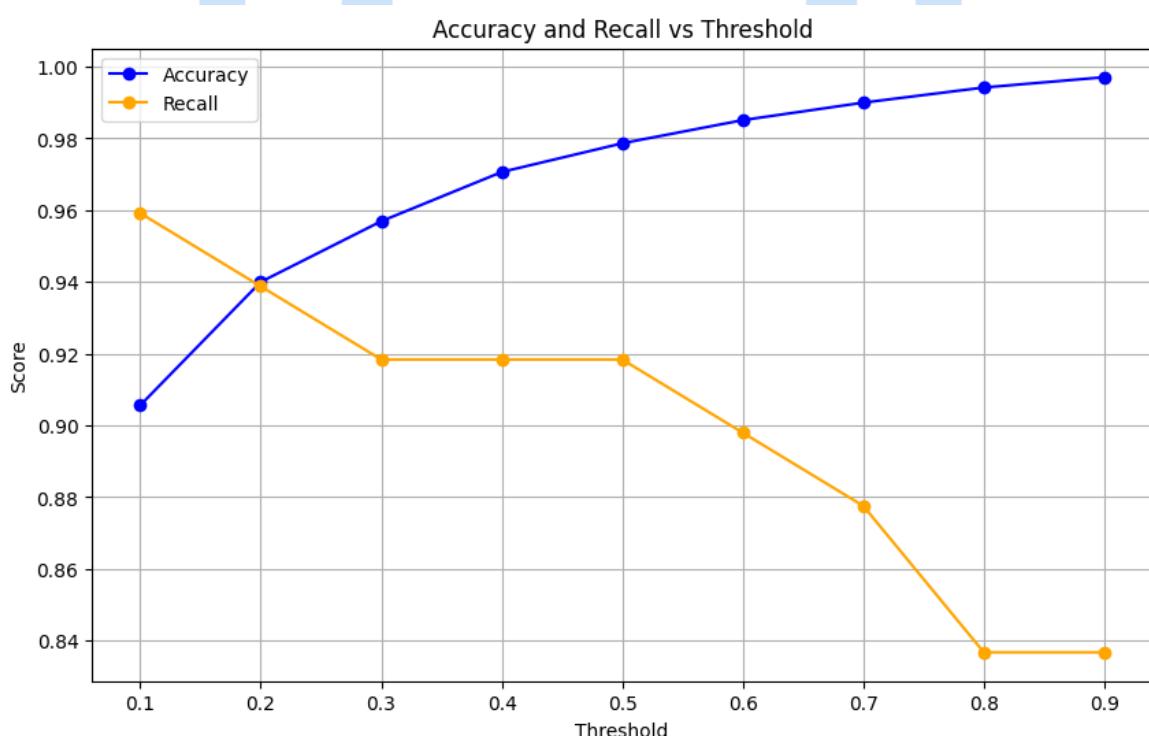
تعادل بین خطاهای Precision و F1-Score به تعادل بین False Negatives و False Positives کمک می‌کند، که در Accuracy نادیده گرفته می‌شود.

مقاومت در برابر نامتوازنی: AUC-ROC و F1-Score تحت تأثیر توزیع نامتوازن کلاس‌ها قرار نمی‌گیرند و عملکرد واقعی مدل را بهتر نشان می‌دهند.

سوال ۵

با آستانه‌های مختلف برای Oversampling عمل کرد مدل را بررسی کرده و نمودار Accuracy & Recall را مانند شکل ۷ مقاله ترسیم کنید.

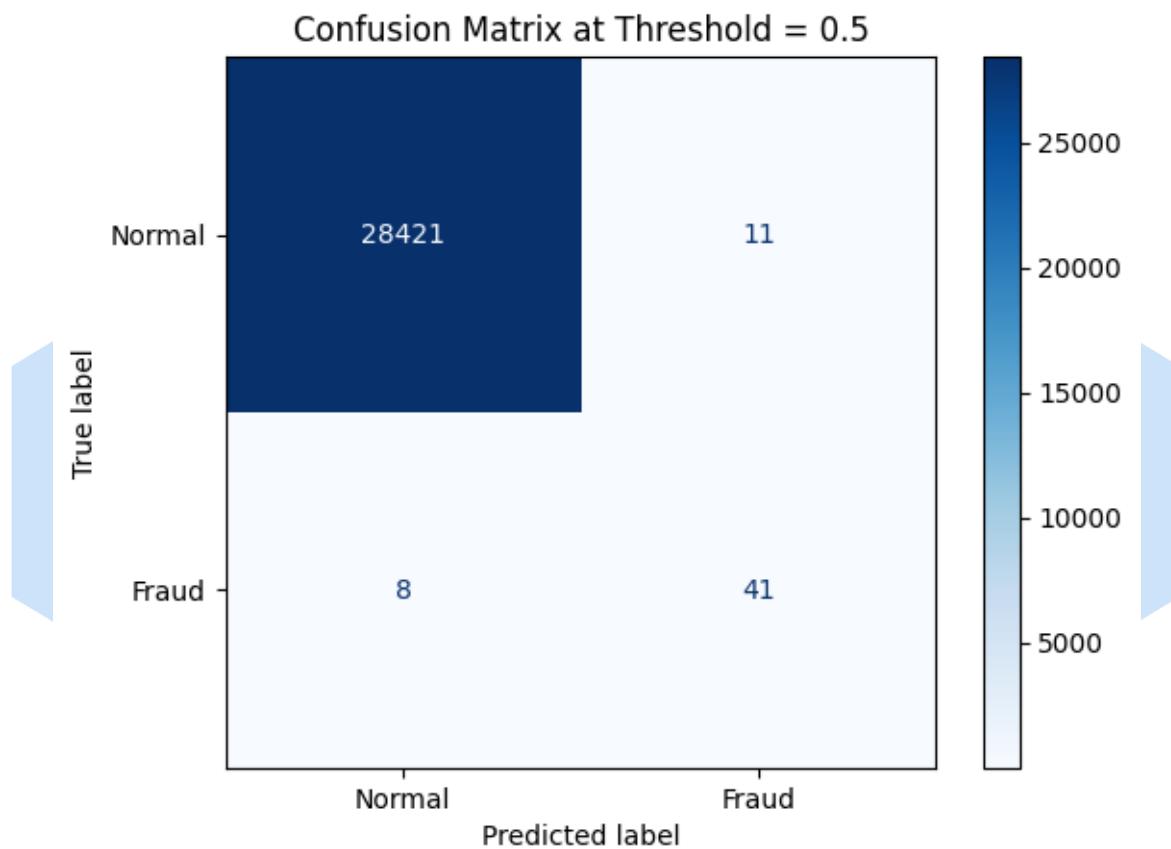
پاسخ سوال ۵



سوال ۶

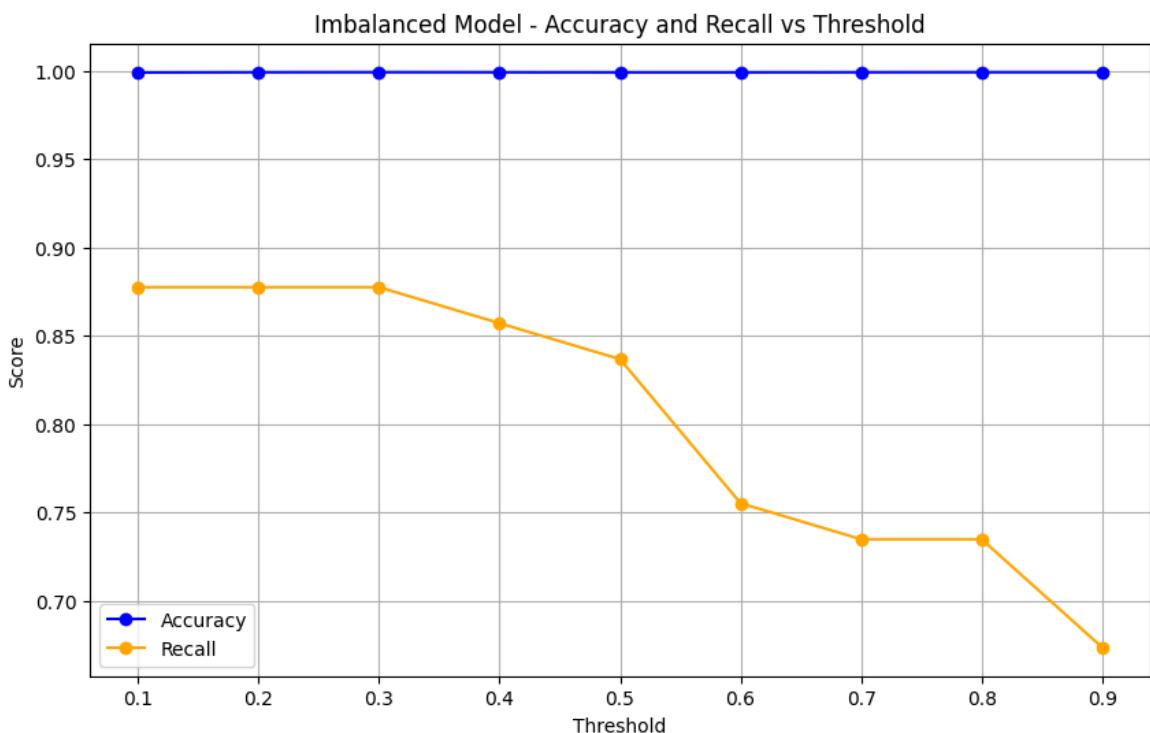
مدل را با استفاده از داده های نامتوازن و بدون حذف نویز، آموزش داده و موارد بخش قبلی را گزارش کنید و نتایج دو مدل را با هم مقایسه کنید.

پاسخ سوال ۶



همانگونه که در ماتریس در هم ریختگی مشاهده می کنید مدل اکثر نمونه ها را کلاس نرمال تشخیص داده است و این یعنی در تشخیص کلاس عادی(اکثیریت) به خوبی عمل می کند ولی چون هدف ما تشخیص کلاس تقلب هست مدل در تشخیص این کلاس تعداد تشخیص نسبتاً زیادی داشته است و در تشخیص کلاس تقلب خوب عمل نکرده است.

	Precision	Recall	f1-score	Accuracy
Normal	1.00	1.00	1.00	0.98
Fraud	0.79	0.84	0.81	



accuracy تقریباً ثابت و نزدیک به ۱.۰۰ در تمام آستانه‌ها. recall از حدود ۰.۸۸ در آستانه‌های پایین به ۰.۶۷ در آستانه ۰.۹ کاهش می‌یابد. یعنی حتی بدون SMOTE و Autoencoder، مدل به خوبی موارد تقلب را تشخیص می‌دهد.

مقایسه دو مدل:

معیار	بدون) مدل پایه SMOTE + Autoencoder	بدون) مدل پایه SMOTE/DAE)
Precision (Fraud)	۰.۰۷	✓ ۰.۷۹
Recall (Fraud)	✓ ۰.۹۲	۰.۸۴
F1-score (Fraud)	۰.۱۳	✓ ۰.۸۱
Accuracy کلی	۰.۹۸	✓ ۱.۰۰
Precision بالا	✗ (آlerm‌های غلط زیاد)	✓ آlerm‌های دقیق‌تر
Recall بالا	✓	خوب، ولی کمی کمتر