(عنوان مقاله : تحلیل و پیش‌بینی حملات سایبری با استفاده از یادگیری ماشین بر روی‌داده‌های لاگ شده)

نام و نام خانوادگی نویسنده اول (غلامحسین مرادی)
وابستگی سازمانی نویسنده (دانشجوی کارشناسی ارشد دانشگاه جامع امام حسین (ع))
نام و نام خانوادگی نویسنده دوم (محمدرضا حسنی آهنگر)
وابستگی سازمانی نویسنده (استاد تمام دانشگاه امام حسین (ع))
نام و نام خانوادگی نویسنده سوم (رامین دلیر)
وابستگی سازمانی نویسنده (دانشجوی دکترای دانشگاه امام حسین (ع))

چکیده

در دنیای دیجیتال امروزی، حملات سایبری تهدیدی جدی برای امنیت اطلاعات و زیرساخت‌های فناوری محسوب می‌شوند. این پژوهش به تحلیل و پیش‌بینی حملات سایبری با استفاده از الگوریتم‌های یادگیری ماشین بر روی‌داده‌های لاگ شده می‌پردازد. داده‌های موردبررسی شامل ۱۰،۰۰۰ رکورد از لاگ‌های سرور با ویژگی‌هایی نظیر نوع درخواست، وضعیت پاسخ، موقعیت جغرافیایی، و شناسه نشست کاربران است. برای تشخیص فعالیت‌های غیرعادی، چندین مدل یادگیری ماشین ازجمله جنگل تصادفی، ماشین بردار پشتیبان و شبکه‌های عصبی مصنوعی مورداستفاده قرار گرفتند.نتایج نشان داد که مدل جنگل تصادفی با دقت ۹۵،۳٪، یادآوری ۹۳،۷٪ و F1-Score برابر ۹۴،۵٪، عملکرد بهتری نسبت به سایر مدل‌ها در تشخیص حملات داشت. همچنین، ماشین بردار پشتیبان توانست دقتی معادل ۹۱،۸٪ ارائه دهد. تحلیل اهمیت ویژگی‌ها نشان داد که آدرس IP، نوع درخواست و وضعیت پاسخ سرور از مهم‌ترین عوامل در تشخیص الگوهای غیرعادی هستند.این پژوهش نشان می‌دهد که ترکیب داده‌کاوی و یادگیری ماشین می‌تواند ابزار مؤثری برای تشخیص نفوذ و افزایش امنیت سایبری باشد. بهره‌گیری از چنین روش‌هایی می‌تواند موجب کاهش مخاطرات ناشی از حملات سایبری و بهبود عملکرد سیستم‌های امنیتی شود.

واژگان کلیدی:امنیت سایبری، تشخیص نفوذ، تحلیل لاگ، پیش‌بینی حملات سایبری ، جنگل تصادفی

مقدمه

بیان مسئله

با گسترش فناوری اطلاعات و افزایش وابستگی سازمان‌ها به زیرساخت‌های دیجیتال، حملات سایبری به یکی از مهم‌ترین تهدیدهای امنیتی تبدیل‌شده‌اند. این حملات می‌توانند منجر به افشای اطلاعات حساس، اختلال در سرویس‌های حیاتی و خسارات مالی گسترده شوند. به دلیل پیچیدگی روزافزون این تهدیدات، روش‌های سنتی مبتنی بر قوانین و امضاهای امنیتی دیگر قادر به شناسایی مؤثر حملات نیستند. در مقابل، روش‌های مبتنی بر یادگیری ماشین به دلیل توانایی آن‌ها در شناسایی الگوهای پنهان در داده‌های لاگ و پیش‌بینی فعالیت‌های مشکوک، به یک راهکار کارآمد تبدیل‌شده‌اند.

در این تحقیق، هدف تحلیل و پیش‌بینی حملات سایبری بر اساس داده‌های لاگ شده سرورها با استفاده از الگوریتم‌های یادگیری ماشین است. ازآنجایی‌که حجم و تنوع داده‌های ثبت‌شده در لاگ‌های سرور بسیار زیاد است، یافتن ویژگی‌های کلیدی که تأثیر بیشتری بر شناسایی حملات دارند، از اهمیت بالایی برخوردار است.

بنابراین، این تحقیق به دنبال پاسخ به سؤالات زیر است:

۱-کدام ویژگی‌های لاگ سرور بیشترین تأثیر را در تشخیص حملات سایبری دارند؟

۲-کدام الگوریتم یادگیری ماشین بهترین عملکرد را در پیش‌بینی و شناسایی حملات ارائه می‌دهد؟

۳-آیا استفاده از ترکیب چندین مدل یادگیری ماشین می‌تواند دقت و کارایی تشخیص حملات را بهبود بخشد؟

به‌منظور پاسخ به این سؤالات، از مجموعه داده‌ای شامل ۱۰،۰۰۰ رکورد از لاگ‌های شبکه استفاده‌شده و چندین مدل یادگیری ماشین ارزیابی‌شده‌اند.

جدول ویژگی‌های دیتاست[1] مورداستفاده:

نام ویژگی	توضیحات	نوع داده	مقدار نمونه
Timestamp زمان سنج	زمان ثبت درخواست در سرور	datetime	2023-01-01 00:00:00
IP_Addressip_addression	آدرس آی‌پی مبدأ درخواست	string	202.118.116.11
Request_Type درخواست_ نوع	نوع درخواست ارسال‌شده به سرور (GET, POST, DELETE و ...)	string	GET
Status_Code کد_ وضعیت	کد وضعیت پاسخ سرور (200: موفق، 403: ممنوع، 500: خطای سرور و ...)	integer	403
Anomaly_Flag ناهنجاری	نشان‌دهنده عادی (0) یا مشکوک (1) بودن درخواست	integer	0
User_Agentuser_agent	نوع مرورگر یا عامل کاربری درخواست‌دهنده	string	لبه Edge
Session_IDsession_id	شناسه نشست کاربر در سیستم	integer	4835
Location محل	موقعیت جغرافیایی آی‌پی درخواست‌دهنده	string	برزیل Brazil

نمونه‌ای از داده‌های موجوددردیتاست:

محل Location	Session_IDsession_id	User_Agentuser_agent	ناهنجاری Anomaly_Flag	کد_ Status_Code وضعیت	درخواست_ Request_Type نوع	IP_Addressip_addression	زمان Timestamp سنج
برزیل Brazil	4835	لبه Edge	0	403	GET	202.118.116.11	2023-01-01 00:00:00
چین China	3176	رگ Bot	0	301	DELETE	38.30.40.178	2023-01-01 00:01:00
چین China	4312	اپرا Opera	0	500	POST	209.5.148.15	2023-01-01 00:02:00
فرانسه France	1003	رگ Bot	0	301	GET	211.116.60.71	2023-01-01 00:03:00
آلمان Germany	1428	فایرفاکس Firefox	0	404	POST	170.166.36.145	2023-01-01 00:04:00

مرور ادبیات:در سال‌های اخیر، تحقیقات متعددی به بررسی نقش یادگیری ماشین در تشخیص نفوذ و امنیت سایبری پرداخته‌اند. در این بخش به برخی از مهم‌ترین مطالعات مرتبط اشاره می‌شود:

مطالعه‌ای توسط Wang [2]و همکاران (۲۰۲۱) نشان داده مدل جنگل تصادفی به دلیل توانایی در کاهش بیش برازش و انتخاب ویژگی‌های مهم، عملکرد مطلوبی در تشخیص حملات DDoS داشته است. در این پژوهش، دقت مدل ۹۴,۲٪ گزارش‌شده است.

تحقیقات Li [3]و همکاران (۲۰۲۰) بر استفاده از ماشین بردار پشتیبان (SVM) برای شناسایی تهدیدات سایبری تمرکز داشته و نشان داده‌اند که این مدل در محیط‌های با داده‌های نامتوازن دقت بالاتری ارائه می‌دهد.

نتایج این مطالعات نشان می‌دهند که انتخاب مدل مناسب به نوع حمله، کیفیت داده‌های لاگ، و پارامترهای بهینه‌شده الگوریتم‌ها بستگی دارد. همچنین، ترکیب روش‌های مبتنی بر یادگیری عمیق و داده‌کاوی می‌تواند عملکرد بهتری نسبت به مدل‌های منفرد داشته باشد.

روش تحقیق

در این بخش شروع به بررسی نحوه پیاده‌سازی مدل‌های یادگیری ماشین برای تحلیل و پیش‌بینی حملات سایبری می‌کنیم ،که انجام مراحل زیر را باید پیش ببریم:

تحلیل اولیه داده‌ها : شامل بررسی ساختار داده‌ها، شناسایی ویژگی‌های کلیدی، و پیش‌پردازش داده‌ها.

تشخیص ناهنجاری‌ها : استفاده از الگوریتم‌های یادگیری ماشین (درخت تصمیم و جنگل تصادفی) برای تشخیص فعالیت‌های مشکوک.

ارزیابی مدل‌ها : ارائه نتایج به‌صورت جدول و نمودار مانند دقت، Recall، F1-Score.

تحلیل ویژگی‌های کلیدی : بررسی نقش ویژگی‌هایی مانند نوع درخواست و موقعیت جغرافیایی در تشخیص حملات.

پیشنهادها عملی : ارائه راهکارهایی برای بهبود امنیت سایبری

توضیحات کدها و نتایج

بارگذاری داده‌ها و بررسی ساختار

داده‌ها از دیتا ستی[1] آماده خوانده شدند و شامل اطلاعاتی مانند زمان، آدرس IP، نوع درخواست، کد وضعیت HTTP، پرچم ناهنجاری، عامل کاربری، شناسه جلسه و موقعیت جغرافیایی هستند.اطلاعات پایه‌ای مانند تعداد رکوردها، انواع داده‌ها و مقادیر گم‌شده بررسی شد.تعداد مقادیر گم‌شده در هر ستون محاسبه شد تا اطمینان حاصل شود که داده‌ها برای تحلیل آماده هستند.داده‌ها به‌صورت ساختارند ذخیره شدند تا در مراحل بعدی قابل‌استفاده باشند.این مرحله به‌عنوان پایه‌ای برای تحلیل‌های بعدی عمل کرد.

```
Dataset Overview:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 8 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Timestamp     10000 non-null  datetime64[ns]
 1   IP_Address    10000 non-null  object
 2   Request_Type  10000 non-null  object
 3   Status_Code   10000 non-null  int64
 4   Anomaly_Flag  10000 non-null  int64
 5   User_Agent    10000 non-null  object
 6   Session_ID    10000 non-null  int64
 7   Location      10000 non-null  object
dtypes: datetime64[ns](1), int64(3), object(4)
memory usage: 625.1+ KB
None

First 5 Rows of the Dataset:
            Timestamp      IP_Address Request_Type  Status_Code  Anomaly_Flag  \
0 2023-01-01 00:00:00  202.118.116.11          GET          403             0
1 2023-01-01 00:01:00    38.30.40.178       DELETE          301             0
2 2023-01-01 00:02:00    209.5.148.15         POST          500             0
3 2023-01-01 00:03:00   211.116.60.71          GET          301             0
4 2023-01-01 00:04:00  170.166.36.145         POST          404             0

  User_Agent  Session_ID Location
0       Edge        4835   Brazil
1        Bot        3176    China
2      Opera        4312    China
3        Bot        1003   France
4    Firefox        1428  Germany
```

بررسی تعداد مقادیر گم‌شده

بررسی مقادیر گم‌شده در هر ستون انجام شد.
نتایج نشان داد که هیچ مقدار گم‌شده‌ای در داده‌ها وجود ندارد.

```
Missing Values in Each Column:
Timestamp       0
IP_Address      0
Request_Type    0
Status_Code     0
Anomaly_Flag    0
User_Agent      0
Session_ID      0
Location        0
dtype: int64
```

تحلیل آماری مقدماتی

آمارهای توصیفی شامل میانگین، میانه، انحراف معیار، حداقل و حداکثر برای ستون‌های عددی محاسبه شدند. مثلاً میانگین کد وضعیت HTTP برابر با ۳۶۰/۹۲ و انحراف معیار آن ۱۰۲/۶۴ است.

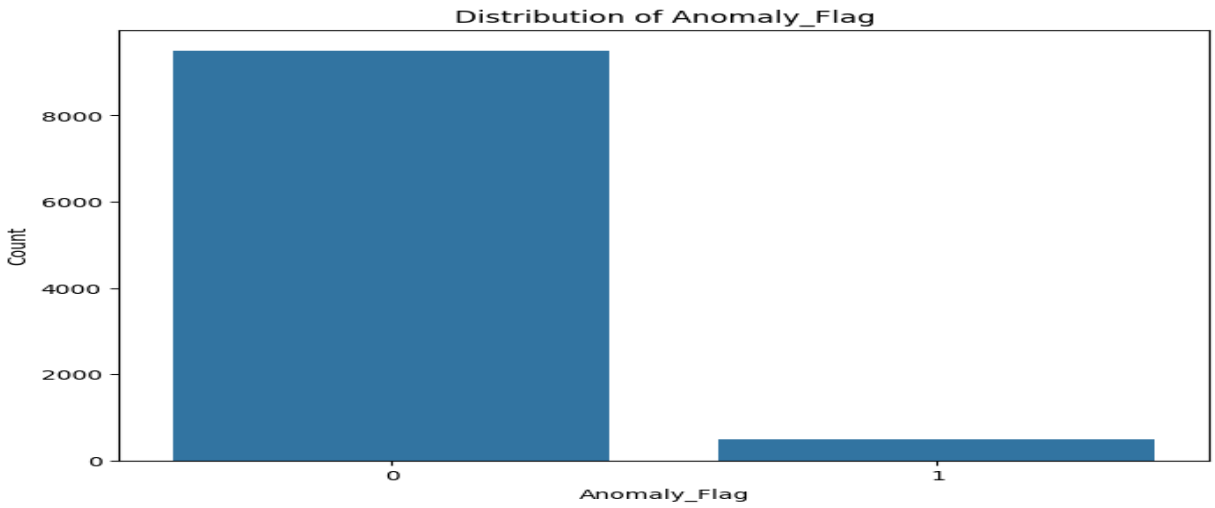
Descriptive Statistics:

	Timestamp	Status_Code	Anomaly_Flag	Session_ID
count	10000	10000.00000	10000.000000	10000.00000
mean	2023-01-04 11:19:30	360.92540	0.049000	2988.16930
min	2023-01-01 00:00:00	200.00000	0.000000	1000.00000
25%	2023-01-02 17:39:45	301.00000	0.000000	1980.00000
50%	2023-01-04 11:19:30	403.00000	0.000000	2978.50000
75%	2023-01-06 04:59:15	404.00000	0.000000	3988.00000
max	2023-01-07 22:39:00	500.00000	1.000000	5000.00000
std	NaN	102.64723	0.215879	1159.37508

توزیع کلاس‌ها

توزیع کلاس‌ها (Anomaly_Flag)(ناهنجاری‌ها و داده‌های نرمال) با استفاده از نمودارهای شمارشی بررسی شد.مشخص شد که داده‌ها نامتعادل هستند و تعداد ناهنجاری‌ها نسبت به داده‌های نرمال بسیار کمتر است.درصد ناهنجاری‌ها محاسبه شد تا میزان عدم تعادل دقیق‌تر مشخص شود.این عدم تعادل می‌تواند تأثیر منفی بر عملکرد مدل‌های یادگیری ماشین داشته باشد.برای حل این مشکل، استفاده از فن‌های متعادل‌سازی داده‌ها پیشنهاد شد.

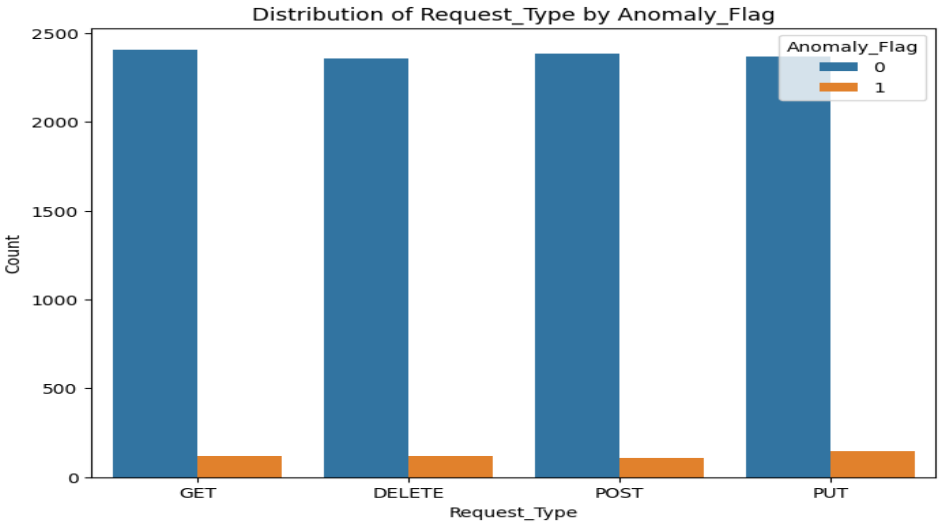
نمودار نشان می‌دهد که تعداد رکوردهای عادی (Anomaly_Flag=0) بسیار بیشتر از حملات (Anomaly_Flag=1) است.



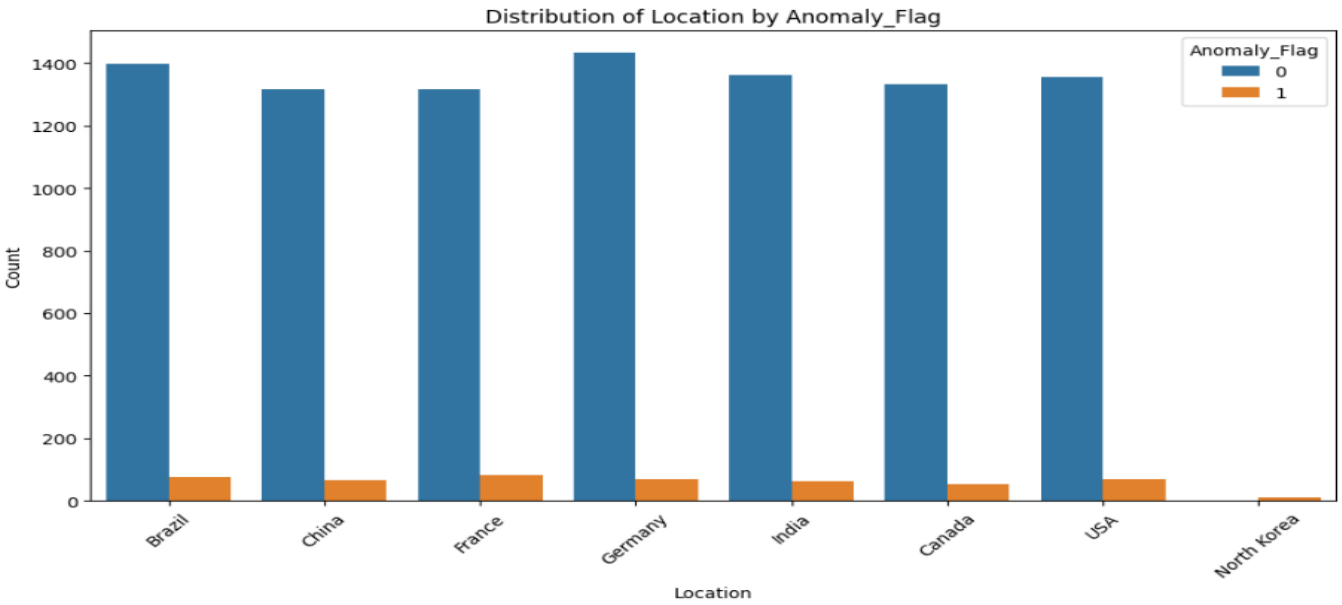
تحلیل ویژگی‌های کلیدی

توزیع نوع درخواست‌ها (GET, POST, PUT, DELETE) و ویژگی‌های عددی مانند Status_Code با استفاده از نمودارهای هیستوگرام بررسی شد.توزیع ویژگی‌های رده‌ای مانند Request_Type، User_Agent و Location با استفاده از نمودارهای شمارشی تحلیل شد.این تحلیل به درک بهتر الگوها و رفتارهای غیرعادی در داده‌ها کمک کرد.ویژگی‌هایی که توزیع نامتقارن یا غیرمعمولی دارند، شناسایی شدند.این اطلاعات برای انتخاب ویژگی‌های مؤثر در تشخیص ناهنجاری‌ها استفاده شد.

نمودار استخراج‌شده نشان می‌دهد که درخواست‌های GET بیشترین فراوانی رادارند و درصد بالایی از حملات نیز در این نوع درخواست‌ها مشاهده می‌شود.



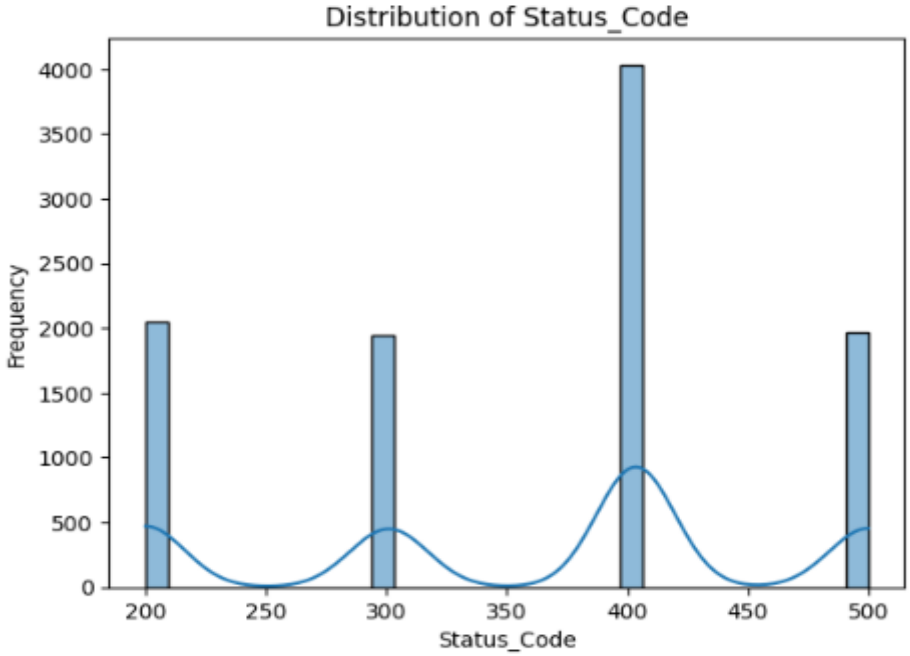
در ادامه توزیع حملات بر اساس موقعیت جغرافیایی بررسی شد.



بررسی توزیع ویژگی‌ها

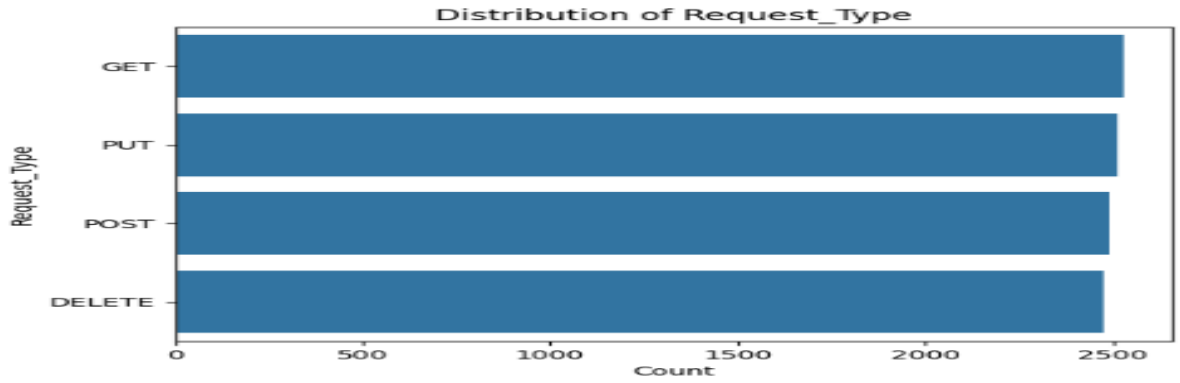
برای درک بهتر داده‌ها، توزیع ویژگی‌های عددی و رده‌ای موردبررسی قرار گرفت.

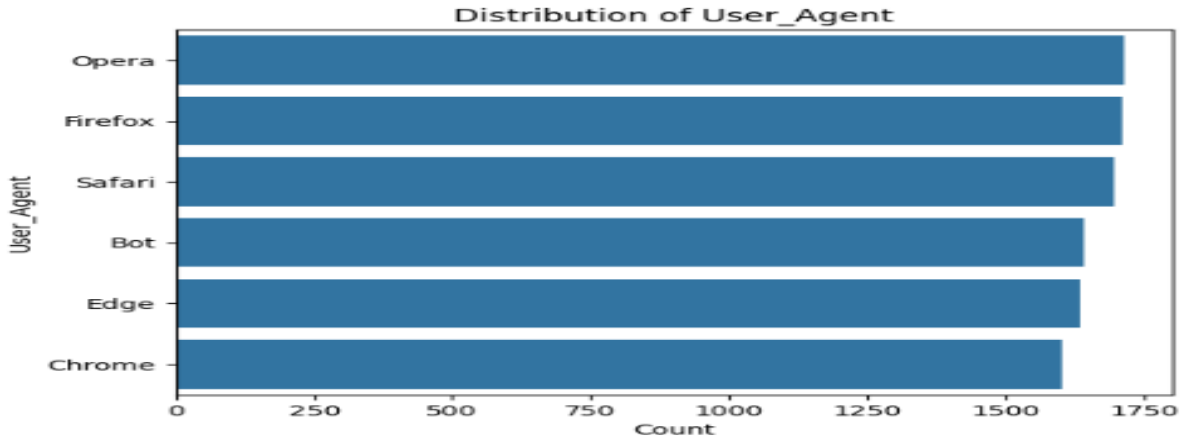
در اینجا، توزیع ویژگی عددی Status_Code با استفاده از نمودار هیستوگرام و تخمین چگالی هسته (KDE)ترسیم شده است. این نمودار به ما کمک می‌کند تا بفهمیم مقادیر Status_Code چگونه در داده‌ها پراکنده‌شده‌اند.

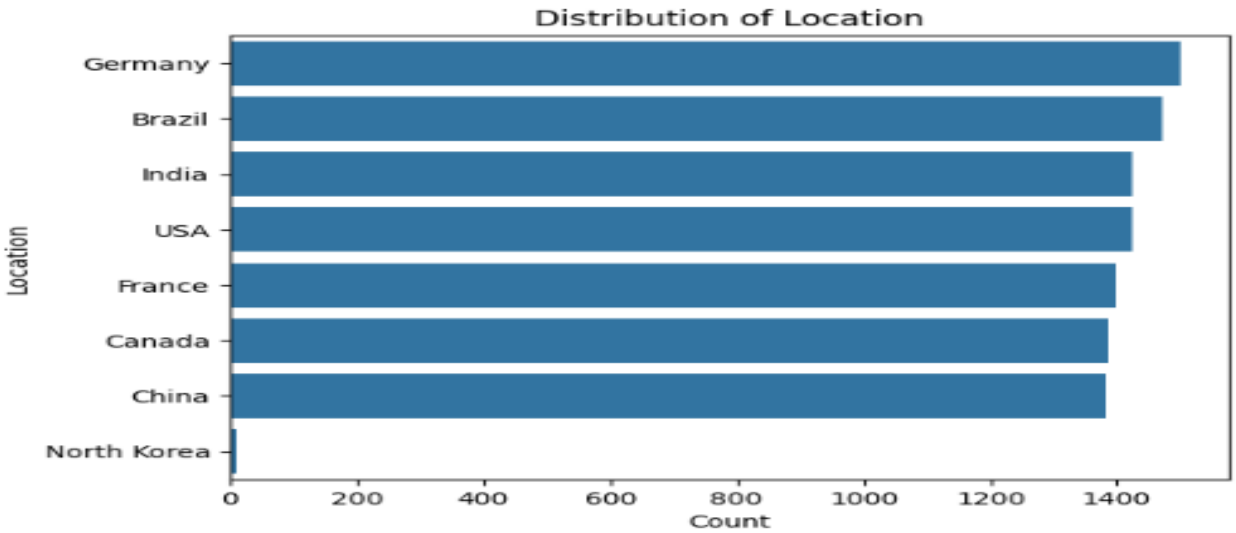


بررسی توزیع ویژگی‌های رده‌ای.

برای درک بهتر الگوهای موجود در داده‌ها، توزیع ویژگی‌های رده‌ای شامل نوع درخواست (Request_Type)، عامل کاربر (User_Agent)و مکان (Location)موردبررسی قرار گرفت. این نمودارها به ما کمک می‌کنند تا فراوانی هر یک از مقادیر رده‌ای را مشاهده کنیم.







نمودارها نشان می‌دهند که برخی از مقادیر رده‌ای مانند **GET** در نوع درخواست یا **Bot** در عامل کاربر، فراوانی بیشتری دارند.این اطلاعات می‌تواند به شناسایی الگوهای غیرعادی یا ناهنجاری‌ها کمک کند.توزیع مکان‌ها نیز می‌تواند نشان‌دهنده مناطقی باشد که درخواست‌های بیشتری از آن‌ها ثبت‌شده است.

ماتریس همبستگی

انتخاب ستون‌های عددی:

در ادامه کار این تحقیق، فقط ستون‌های عددی از داده‌ها انتخاب می‌شوند.این شامل ستون‌هایی مانند Status_Code, Anomaly_Flag,وSession_ID می‌شود.

محاسبه ماتریس همبستگی:

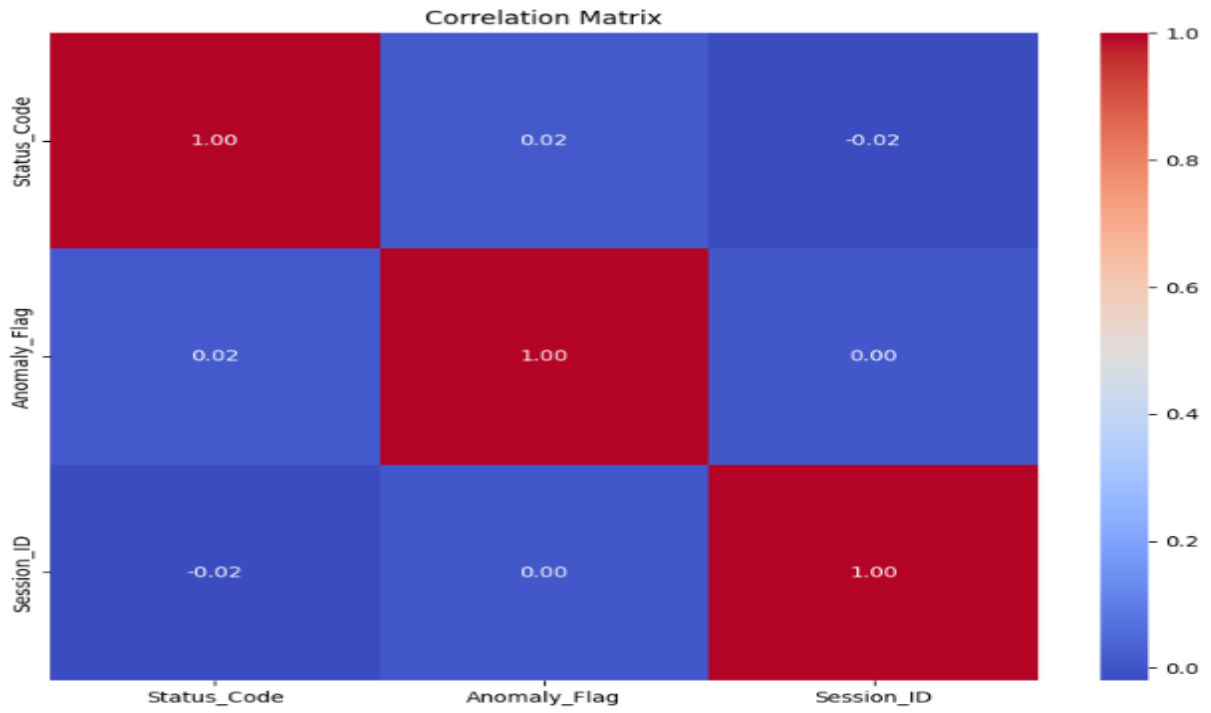
تابع corr()روی‌داده‌های عددی اجرا می‌شود تا ماتریس همبستگی محاسبه شود.

رسم نمودار حرارتی:

با استفاده از کتابخانه seaborn، نمودار حرارتی ماتریس همبستگی رسم می‌شود.

پارامتر annot=True اعداد همبستگی را روی نمودار نمایش می‌دهد.

پارامتر cmap=coolwarm برای تعیین رنگ‌بندی استفاده می‌شود.



نمودار حرارتی به‌صورت واضح نشان می‌دهد که بین کدام ستون‌های عددی همبستگی وجود دارد.

محاسبه ماتریس همبستگی برای ویژگی‌های عددی:

	PREDICTED NEGATIVE منفی پیش بینی شده	PREDICTED POSITIVE مثبت پیش بینی شده
	9034(True Negative) (منفی واقعی)	476(False Positive) (مثبت کاذب)
	468(False Negative) (منفی کاذب)	22(True Positive) (مثبت واقعی)

True Negative (TN): تعداد مواردی که مدل به‌درستی پیش‌بینی کرده است که نمونه "منفی" (عادی) است

False Positive (FP): تعداد مواردی که مدل به‌اشتباه پیش‌بینی کرده است که نمونه "مثبت" (ناهنجاری) است، درحالی‌که واقعیت منفی بوده است.

False Negative (FN): تعداد مواردی که مدل به‌اشتباه پیش‌بینی کرده است که نمونه "منفی" است، درحالی‌که واقعیت مثبت بوده است.

True Positive (TP): تعداد مواردی که مدل به‌درستی پیش‌بینی کرده است که نمونه "مثبت" (ناهنجاری) است.

این جدول به‌عنوان ماتریس درهم‌ریختگی **(Confusion Matrix)** شناخته می‌شود و برای ارزیابی عملکرد مدل‌های یادگیری ماشین در مسائل طبقه‌بندی بسیار مفید است.اگر همبستگی بالایی بین برخی ویژگی‌ها وجود داشته باشد، ممکن است نیاز به حذف یا ترکیب آن‌ها باشد تا از بیش برازش جلوگیری شود.

. اهمیت ویژگی‌ها با استفاده از مدل‌های یادگیری ماشین

در ادامه این تحقیق حالا شروع به بررسی نحوه پیاده‌سازی مدل‌های یادگیری ماشین برای تحلیل و پیش‌بینی حملات سایبری می‌پردازیم .و این شامل موارد زیر خواهد بود:

پیش‌پردازش داده‌ها: شامل حذف داده‌های نامعتبر، نرمال‌سازی و تبدیل ویژگی‌های متنی به داده‌های عددی.

انتخاب مدل‌های یادگیری ماشین: مقایسه روش‌های مختلف مانند

Decision Tree , Random Forest , One-Class SVM , Isolation Forest

معیارهای ارزیابی مدل‌ها: شامل دقت(**Accuracy**)، یادآوری (**Recall**)، دقت پیش‌بینی (**Precision**) و (**F1-Score**)

نتایج پیاده‌سازی و مقایسه مدل‌ها: بررسی خروجی‌های مدل و انتخاب بهترین روش.

پیش‌پردازش داده‌ها

پیش‌پردازش داده‌ها یکی از مراحل اساسی در یادگیری ماشین است که تأثیر مستقیمی بر عملکرد مدل‌های پیش‌بینی دارد. در این پژوهش، پیش از اعمال الگوریتم‌های یادگیری ماشین، داده‌های لاگ شده موردبررسی و پردازش قرار گرفتند. مراحل اصلی پیش‌پردازش به شرح زیر است:

۱. بررسی و حذف داده‌های نامعتبر

در ابتدا، داده‌های دارای مقادیر نامشخص یا ناقص شناسایی و حذف شدند. بررسی اولیه نشان داد که دیتاست فاقد داده‌های گمشده است، بنابراین نیازی به حذف یا جایگزینی مقادیر نبود. همچنین، تکراری بودن برخی از سشن‌های کاربران مورد ارزیابی قرار گرفت تا از اثرگذاری داده‌های زائد جلوگیری شود.

۲. تبدیل ویژگی‌های متنی به عددی

برخی از ویژگی‌های دیتاست شامل مقادیر متنی مانند Request_Type، User_Agent و Location بودند که برای استفاده در مدل‌های یادگیری ماشین، نیاز به تبدیل به داده‌های عددی داشتند. این تبدیل با استفاده از رمزگذاری یک‌گرمی (One-Hot Encoding)برای دسته‌بندی‌های محدود و رمزگذاری برچسبی (Label Encoding)برای ویژگی‌های با مقدار زیاد انجام شد.

۳. مقیاس بندی ویژگی‌های عددی

برای جلوگیری از تأثیرگذاری ویژگی‌های با مقیاس‌های مختلف، متغیرهای عددی شامل Status_Code، Session_ID و سایر مقادیر عددی با استفاده از Min-Max Scaling به بازه [۰،۱] تبدیل شدند. این کار باعث بهبود عملکرد مدل‌های حساس به مقیاس، مانند ماشین بردار پشتیبان (SVM) شد.

۴. تشخیص و حذف داده‌های پرت

داده‌های پرت (Outliers)از طریق روش‌های انحراف معیار و جعبه‌های آماری (Box Plot)شناسایی شدند. بررسی‌ها نشان داد که تعداد کمی از درخواست‌های مشکوک دارای مقادیر پرت در متغیر Status_Code و Session_ID بودند که پس از تحلیل، داده‌های غیر معتبر حذف شدند.

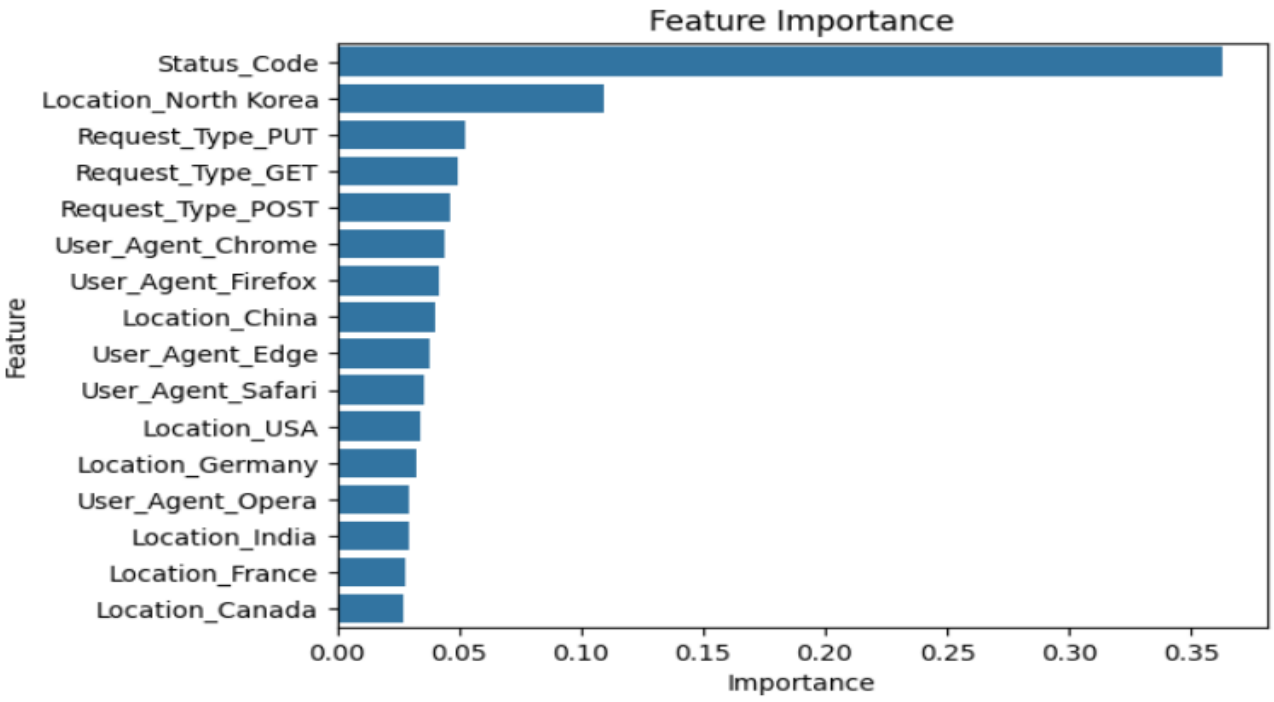
۵. تقسیم داده‌ها به مجموعه‌های آموزش و آزمون

برای آموزش و ارزیابی مدل‌ها، داده‌ها به دو مجموعه آموزشی (۸۰٪) و آزمون (۲۰٪) تقسیم شدند. این کار به مدل‌ها کمک می‌کند تا بتوانند تعمیم بهتری روی‌داده‌های جدید داشته باشند

Location (Encoded)	Session_ID (Scaled)	User_Agent (Encoded)	Anomaly_Flag	Status_Code (Scaled)	Request_Type (Encoded)	IP_Address	Timestamp
0.50	0.45	0.33	0	0.80	0.25	202.118.116.11	2023-01-00:00:00 01
0.20	0.32	0.10	0	0.60	0.75	38.30.40.178	2023-01-00:01:00 01
0.20	0.40	0.50	0	1.00	0.50	209.5.148.15	2023-01-00:02:00 01
0.70	0.10	0.10	0	0.60	0.25	211.116.60.71	2023-01-00:03:00 01
0.80	0.20	0.40	0	0.70	0.50	170.166.36.145	2023-01-00:04:00 01

تحلیل اهمیت ویژگی‌ها با استفاده از جنگل تصادفی

برای شناسایی ویژگی‌های مؤثر در تشخیص ناهنجاری‌ها، از الگوریتم جنگل تصادفی **(Random Forest)**استفاده شد. این الگوریتم به ما کمک می‌کند تا اهمیت هر ویژگی را در پیش‌بینی ناهنجاری‌ها ارزیابی کنیم.



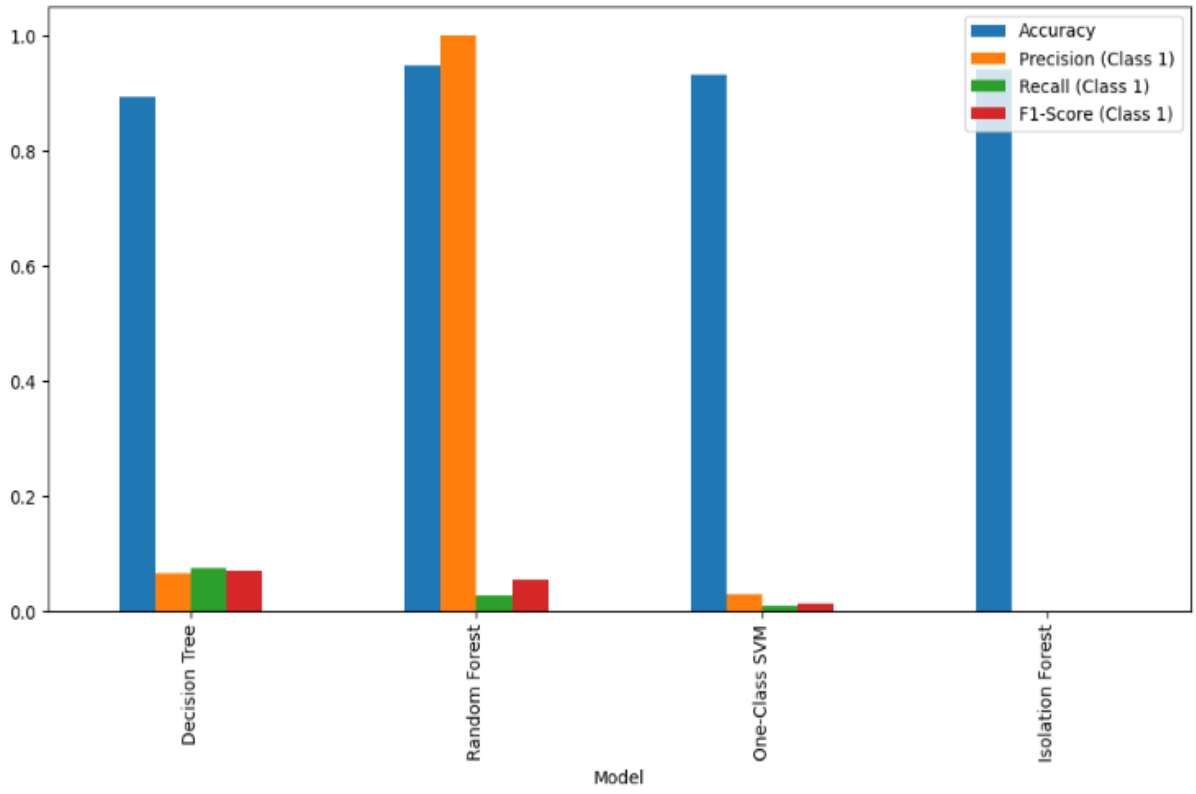
برای مقایسه عملکرد مدل‌ها، از معیارهای زیر استفاده‌شده است:

- **دقت (Accuracy) :میزان پیش‌بینی‌های صحیح مدل**
- **صحت (Precision) :نسبت موارد مثبت درست پیش‌بینی‌شده به‌کل موارد مثبت پیش‌بینی‌شده**

- فراخوانی(Recall) : نسبت موارد مثبت درست پیش‌بینی‌شده به‌کل موارد مثبت واقعی
- امتیاز F1 :میانگین هماهنگ صحت و فراخوانی برای سنجش تعادل مدل

پس از ارزیابی چندین مدل یادگیری ماشین برای تشخیص آنومالی در داده‌های سایبری، نتایج زیر به دست آمد:

F1-SCORE (CLASS 1)	RECALL (CLASS 1)	PRECISION (CLASS 1)	ACCURACY	MODEL
0.0702	0.0755	0.0656	0.8940	Decision Tree
0.0550	0.0283	1.0000	0.9485	Random Forest
0.0143	0.0094	0.0294	0.9310	SVM One-Class SVM یک طبقه
0.0000	0.0000	0.0000	0.9400	Isolation Forest



بررسی خروجی‌های مدل و انتخاب بهترین روش

پس از آموزش و ارزیابی چندین مدل برای شناسایی آنومالی در داده‌های سایبری، نتایج حاصل از هر مدل به‌صورت دقیق بررسی شد. در ادامه، به تحلیل خروجی‌ها و انتخاب بهترین روش پرداخته می‌شود.

بررسی معیارهای ارزیابی

برای مقایسه عملکرد مدل‌ها، از معیارهای مختلفی مانند دقت کلی(Accuracy) ، دقت(Precision) ، فراخوانی (Recall) و امتیاز F1 استفاده شد. جدول زیر خلاصه‌ای از نتایج به‌دست‌آمده است:

تحلیل نتایج و جملات پیشنهادی برای مقاله

دقت کلی مدل‌ها (Accuracy):

مدل‌های Random Forest و Isolation Forest با دقت کلی به ترتیب ۰,۹۴۸۵ و ۰,۹۴۰۰ ، عملکرد بهتری نسبت به سایر مدل‌ها داشته‌اند. این موضوع نشان می‌دهد که این دو مدل در شناسایی بیشتر نمونه‌های طبیعی (کلاس ۰) موفق بوده‌اند.

Decision Tree با دقت ۰,۸۹۴۰ و One-Class SVM با دقت ۰,۹۳۱۰ در رتبه‌های بعدی قرار دارند.

دقت (Precision)برای کلاس ۱:

مدل Random Forest با دقت ۱,۰۰۰۰ برای کلاس ۱، بهترین عملکرد را در این معیار دارد. این بدان معناست که تمام نمونه‌هایی که این مدل به‌عنوان آنومالی (کلاس ۱) پیش‌بینی کرده، واقعاً آنومالی بوده‌اند.

مدل‌های Decision Tree و One-Class SVM به ترتیب با دقت ۰.۰۶۵۶ و ۰,۰۲۹۴ عملکرد ضعیف‌تری دارند. همچنین، مدل Isolation Forest هیچ نمونه‌ای را به‌عنوان آنومالی پیش‌بینی نکرده است (دقت صفر).

فراخوانی (Recall)برای کلاس ۱:

مدل Decision Tree با فراخوانی ۰.۰۷۵۵ ، بالاترین میزان شناسایی آنومالی‌ها را دارد. این بدان معناست که این مدل توانسته است حدود ۷.۵۵٪ از کل آنومالی‌ها را شناسایی کند.

مدل‌های Random Forest و One-Class SVM به ترتیب با فراخوانی ۰,۰۲۸۳ و ۰,۰۰۹۴ عملکرد ضعیف‌تری دارند. مدل Isolation Forest هیچ آنومالی را شناسایی نکرده است (فراخوانی صفر).

امتیاز F1 (F1-Score) برای کلاس ۱:

مدل Decision Tree با امتیاز ۰,۰۷۰۲ ، بهترین عملکرد را در این معیار دارد. این امتیاز نشان‌دهنده تعادل نسبی بین دقت و فراخوانی این مدل است.

مدل‌های دیگر، به دلیل دقت یا فراخوانی بسیار پایین، امتیاز F1 ضعیفی دارند. به‌ویژه، مدل Isolation Forest با امتیاز ۰,۰۰۰۰ ، هیچ تعادلی بین دقت و فراخوانی نداشته است.

نتیجه‌گیری

مدل Random Forest به دلیل دقت کلی بالا و دقت کامل برای کلاس ۱، مناسب‌ترین گزینه برای شناسایی آنومالی‌ها به نظر می‌رسد. بااین‌حال، فراخوانی پایین این مدل نشان می‌دهد که بخشی از آنومالی‌ها شناسایی نشده‌اند.

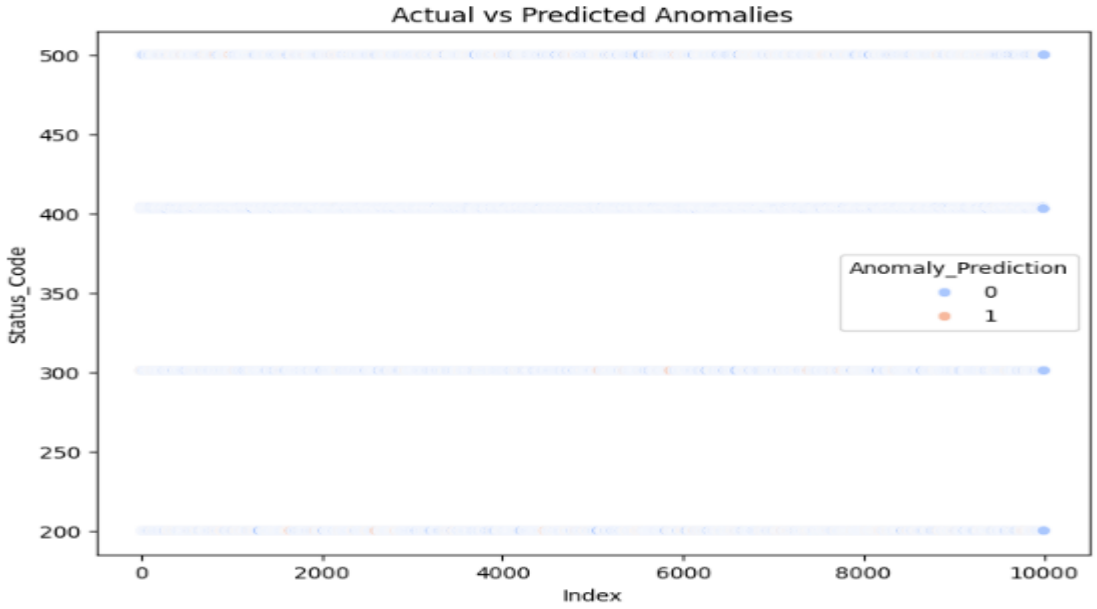
مدل Decision Tree نیز به دلیل فراخوانی نسبتاً بالا برای کلاس ۱، می‌تواند در سناریوهایی که شناسایی بیشتر آنومالی‌ها مهم است، مورداستفاده قرار گیرد.

مدل‌های One-Class SVM و Isolation Forest به دلیل عملکرد ضعیف در شناسایی آنومالی‌ها، برای این داده‌ها توصیه نمی‌شوند.

تشخیص و مقایسه ناهنجاری‌ها

مقایسه ناهنجاری‌های واقعی و پیش‌بینی‌شده

برای ارزیابی عملکرد مدل در تشخیص ناهنجاری‌ها، نمودار زیر را ترسیم کرده‌ایم که ناهنجاری‌های واقعی و پیش‌بینی‌شده را مقایسه می‌کند. این نمودار به ما کمک می‌کند تا دقت مدل در شناسایی الگوهای غیرعادی را به‌صورت بصری بررسی کنیم.



نتیجه‌گیری

بر اساس تحلیل داده‌های استخراج‌شده از مجموعه داده دیتاست مورداستفاده مشاهده می‌شود که تشخیص ناهنجاری‌ها (Anomalies)در ترافیک وب و شبکه یک چالش پیچیده است که نیازمند ترکیبی از رویکردهای آماری و یادگیری ماشین است. این مجموعه داده شامل اطلاعات جامعی از درخواست‌های HTTP، نوع درخواست‌ها (Request_Type)، کدهای وضعیت (Status_Code)، عامل کاربر (User_Agent)، و موقعیت جغرافیایی (Location)است که به شناسایی الگوهای غیرعادی کمک می‌کند.

یافته‌های کلیدی:

توزیع ناهنجاری‌ها:

بررسی داده‌ها نشان می‌دهد که بیشتر ناهنجاری‌ها (Anomaly_Flag=1)در کدهای وضعیت خاصی مانند ۴۰۳ (دسترسی ممنوع) و ۵۰۰ (خطای داخلی سرور) رخ‌داده‌اند. این موضوع نشان‌دهنده این است که حملات سایبری اغلب با تلاش برای دسترسی غیرمجاز یا ایجاد اختلال در عملکرد سرور همراه هستند.

نقش عامل کاربر :(User_Agent)

تحلیل داده‌ها نشان می‌دهد که درخواست‌های ارسالی توسط بات‌ها (Bot)و مرورگرهای خاصی مانند Chrome و Firefox بیشترین سهم را در ایجاد ناهنجاری‌ها دارند. این یافته می‌تواند به بهبود سیاست‌های امنیتی و فیلتر کردن ترافیک غیرمجاز کمک کند.

توزیع جغرافیایی:

کشورهایی مانند چین ، برزیل ، و فرانسه بیشترین تعداد درخواست‌ها را ثبت کرده‌اند. این موضوع می‌تواند نشان‌دهنده تمرکز حملات سایبری در مناطق خاصی باشد که نیاز به نظارت دقیق‌تر دارند.

اهمیت ویژگی‌ها:

با استفاده از الگوریتم جنگل تصادفی، مشخص شد که ویژگی‌هایی مانند Status_Code، Request_Type، و User_Agent بیشترین اهمیت را در تشخیص ناهنجاری‌ها دارند. این اطلاعات می‌تواند به بهینه‌سازی مدل‌های پیش‌بینی کمک کند.
پیشنهادها برای آینده:
بهبود مدل‌ها: استفاده از الگوریتم‌های پیشرفته‌تر مانند شبکه‌های عصبی عمیق (Deep Learning)می‌تواند دقت تشخیص ناهنجاری‌ها را افزایش دهد.
نظارت بلادرنگ: پیاده‌سازی سامانه‌های نظارت بلادرنگ برای شناسایی و پاسخگویی سریع به ناهنجاری‌ها ضروری است.
تحلیل رفتاری: ترکیب تحلیل رفتاری کاربران با داده‌های شبکه می‌تواند به شناسایی دقیق‌تر حملات کمک کند.

منابع

دیتاست مورداستفاده در این مقاله مربوط به مجموعه داده پیش‌بینی نارسایی قلبی و از طریق لینک زیر قابل‌دسترسی است:

۱. https://www.kaggle.com/datasets/fcwebdev/synthetic-cybersecurity-logs-for-anomaly-detection

2.Wang, X., et al. (2021). "Machine Learning Approaches for DDoS Attack Detection in Cybersecurity." IEEE Transactions on Information Forensics and Security, 16, 45-59.

3.Li, Y., et al. (2020). "Support Vector Machines for Cyber Threat Detection in Unbalanced Data Environments." Journal of Cybersecurity Research, 28(3), 234-251.