



شرکت مخابرات ایران
(سهامی عام)

Maintenance

با نظارت دکتر باقری و تحت نظر شرکت مخابرات ایران (آپن)

سید حمید قاسمی

سید یحیی سیدرضا تهرانی

پارسا نوروزی نژاد

مردادماه 1403

باتشکر از سرکار خانم اشرفی



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

■ مقدمه

گزارش حاضر به تحلیل و بررسی عملکرد سیستم‌های ارتباطی و اپراتورهای مرتبط با آن‌ها می‌پردازد. در این گزارش، فرآیند تمیزکاری داده‌ها به عنوان مرحله‌ای اساسی در تحلیل داده‌ها مطرح شده است. این فرآیند شامل پر کردن قسمت‌های خالی، حذف داده‌های ناکارآمد و کوانتیزه کردن زمان‌ها به منظور ایجاد یک پایگاه داده منسجم و قابل تحلیل است. با استفاده از تکنیک‌های مختلف، داده‌ها به گونه‌ای آماده شده‌اند که امکان تحلیل دقیق‌تر و ارائه نتایج معتبر فراهم گردد.

پس از آماده‌سازی داده‌ها، تحلیل‌های مختلفی بر روی آن‌ها انجام شده است. نمودارهای متعددی به منظور نمایش عملکرد اپراتورها در زمینه *MTTR (Mean Time to Repair)* و عوامل مؤثر بر آن ترسیم شده است. این نمودارها به شناسایی الگوهای عملکرد و نقاط ضعف در سیستم کمک می‌کنند.

در ادامه، تأثیر ویژگی‌های مختلف بر امتیازهای *MTTR*، *MTTA* و *MTRBTC* بررسی شده و مدل‌های پیش‌بینی برای تخمین زمان‌های مرتبط با عملکرد اپراتورها ارائه گردیده است. همچنین، با استفاده از مدل‌های پیش‌بینی مانند درخت تصمیم و رگرسیون خطی، تلاش شده است تا زمان‌های مرتبط با عملکرد اپراتورها تخمین زده شود. بخش‌های گزارش عبارتند از:

- اصلاح داده (*data cleaning*)
- تحلیل داده
- گراف سایت‌ها
- تحلیل رفتار عوامل اثر گذار
- مدل‌های پیش‌بینی
- پیشنهادات
- پیوست (لیست *mapping* اندیس‌ها به عناوین)

■ اصلاح داده (data cleaning)

داده هایی که برای این گزارش استفاده شده است مربوط به تیکت های خرابی سایت های مخابراتی است که شامل اطلاعات مختلفی است:

- **MTTR**: این شاخص زمان رفع خرابی بر اساس تیکت های خرابی (*Affect*) در یک دوره مشخص را نشان می دهد که فاصله زمانی ارجاع تیکت تا زمان رفع کامل خرابی می باشد.
- **MTBRTC**: این شاخص متوسط فاصله زمان رفع خرابی تا زمان بستن تیکت توسط استان می باشد.
- **MTTA**: این شاخص «زمان اعلام وصول یا قبول» مجموعه ای از رخدادها یا تیکت ها در یک دوره مشخص را نشان می دهد که فاصله زمانی ارجاع تیکت تا اعلام وصول تیکت توسط کاربر می باشد.
- **MTBRTC_USER**: نام کاربری کسی که تیکت را دیده
- **Alarm_Type**
- **FAULTLEVEL**
- **CODESITE**: کد یکتای مربوط به هر سایت.
- **Affected**: آیا این خرابی سایت را از نظر قطعی و عدم پوشش تحت تاثیر قرار داده است؟
- **ROOTCAUSE**
- **SUBCAUSE**
- **ImpactSiteList**: لیست سایتهایی که تحت تاثیر قرار گرفتند و قطع شده اند
- **CREATETIME**: زمان ایجاد خرابی

با توجه به شاخص های ارزیابی محاسبه امتیاز از پارامتر های داده شده امکان پذیر بود که در ادامه فرآیند امتیاز مربوط به هر تیکت محاسبه شد و به داده های آن تیکت اضافه شد.

تمیزکاری انجام شده عبارت است از:

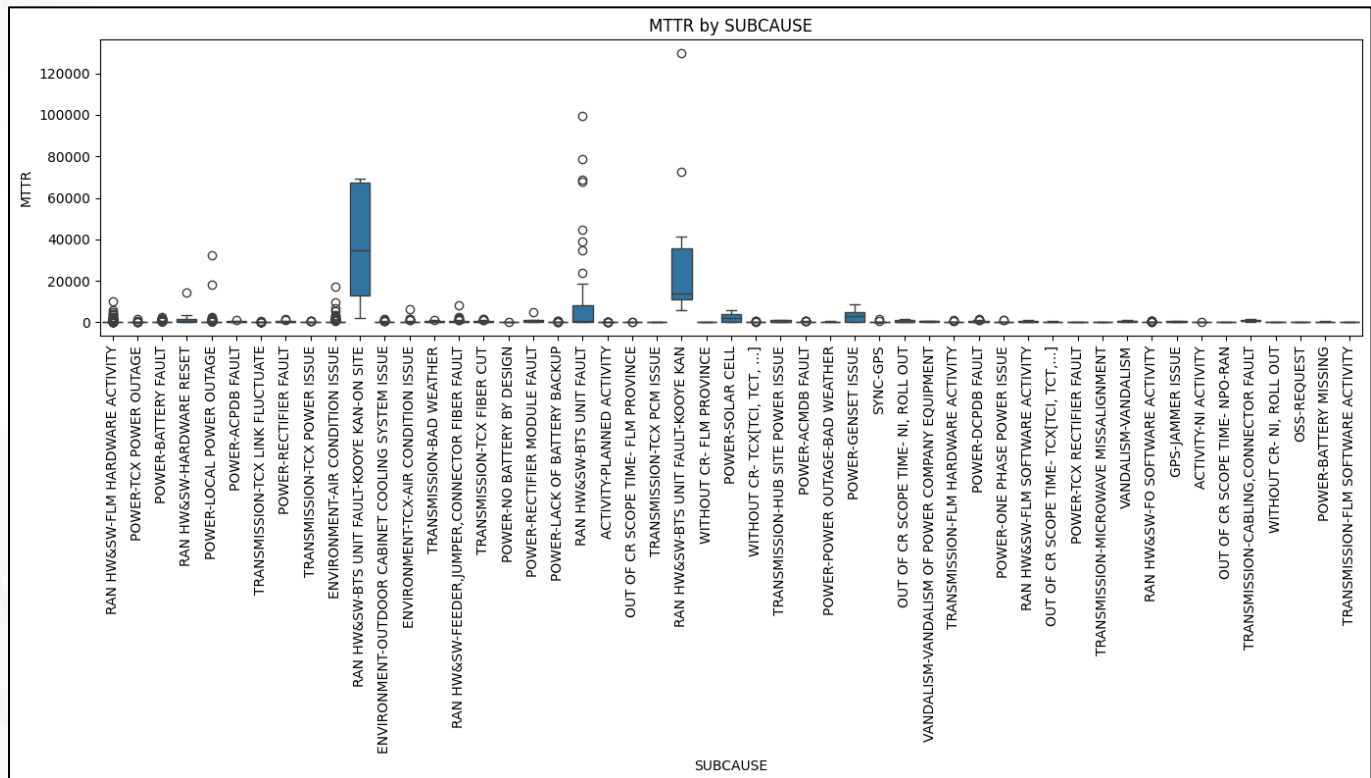
۱. پر کردن تا حد امکان قسمت های خالی با استفاده از قسمت های دیگر، به این معنا که داده های ناقص با استفاده از دیگر داده ها ترمیم شد، به عنوان مثال *lat, long* برخی از سایت ها در بعضی از تیکت ها ثبت نشده بود و این مقادیر به کمک داده های ثبت شده در ردیف های دیگر مربوط به آن سایت تکمیل شد.

۲. حذف داده های NULL

۳. کوانتیزه کردن ۲۴ ساعت شبانه روز و نسبت دادن زمان باز کردن تیکت به یک عدد، ۲۴ ساعت شبانه روز به ۹۶ قسمت ۱۵ دقیقه ای تقسیم شد و برای انجام این کار داده مربوط به *createTime* هر تیکت ابتدا به دقیقه تبدیل شد و مپینگ یکنواخت بین $24 * 60 = 1440min$ شبانه روز به ۹۶ قسمت انجام شد. به عنوان مثال تیکت ثبت شده در 3:20 بامداد به عدد ۱۴ نسبت داده شد.
۴. نسبت دادن داده‌های عنوان‌دار متنی به یک عدد به جهت تحلیل داده توسط ماشین (در فایل پیوست انتهایی گزارش *mapping* مربوط به داده‌های مختلف آورده شده است)
۵. حذف ستون‌های ناکارآمد
۶. محاسبه ارتفاع سایت‌ها از سطح دریا با استفاده از موقعیت جغرافیایی آن
۷. محاسبه امتیاز هر تیکت بر اساس شاخص‌های ارزیابی
- فایل داده‌ها بعد از این فرآیند ضمیمه شده است.

■ تحلیل داده

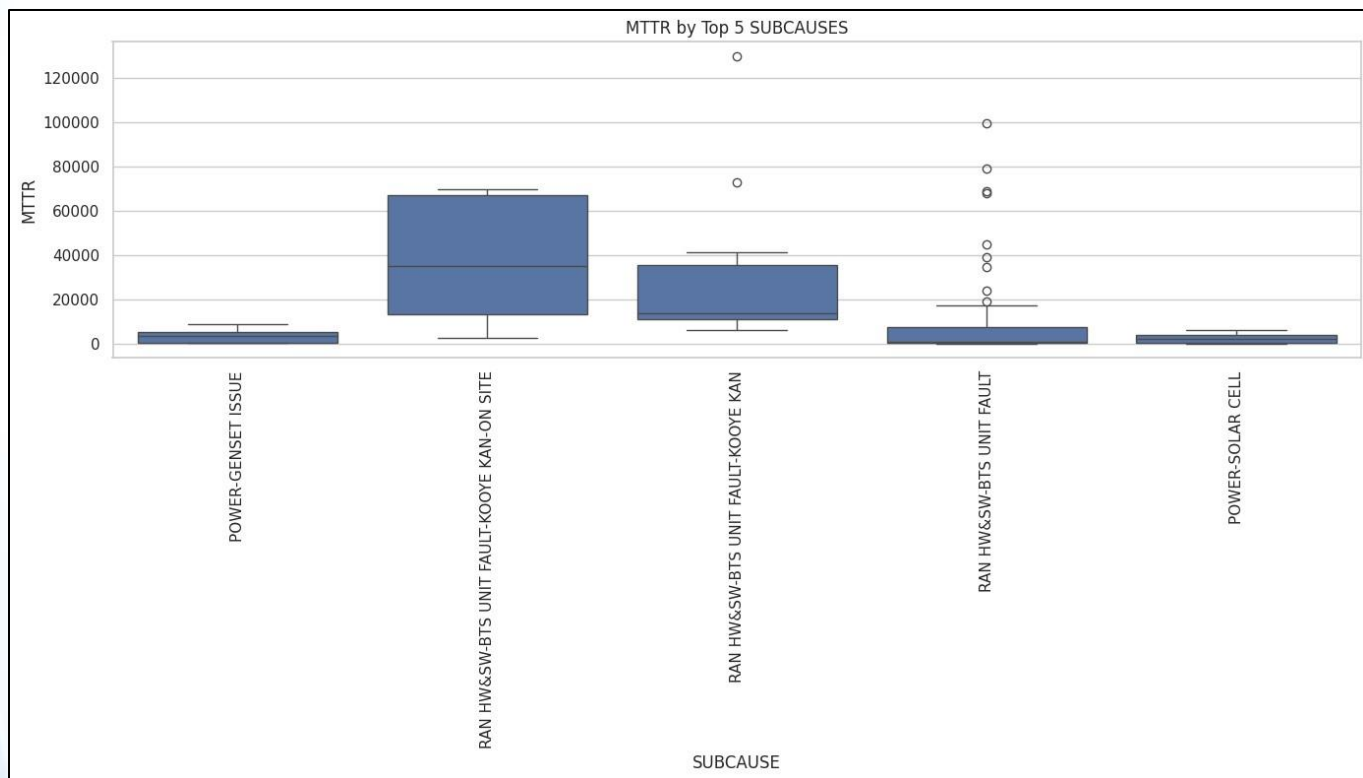
در این بخش ابتدا داده های تمیز شده را با رسم نمودار های مختلف تحلیل کردیم که در ادامه به آن ها اشاره می شود.



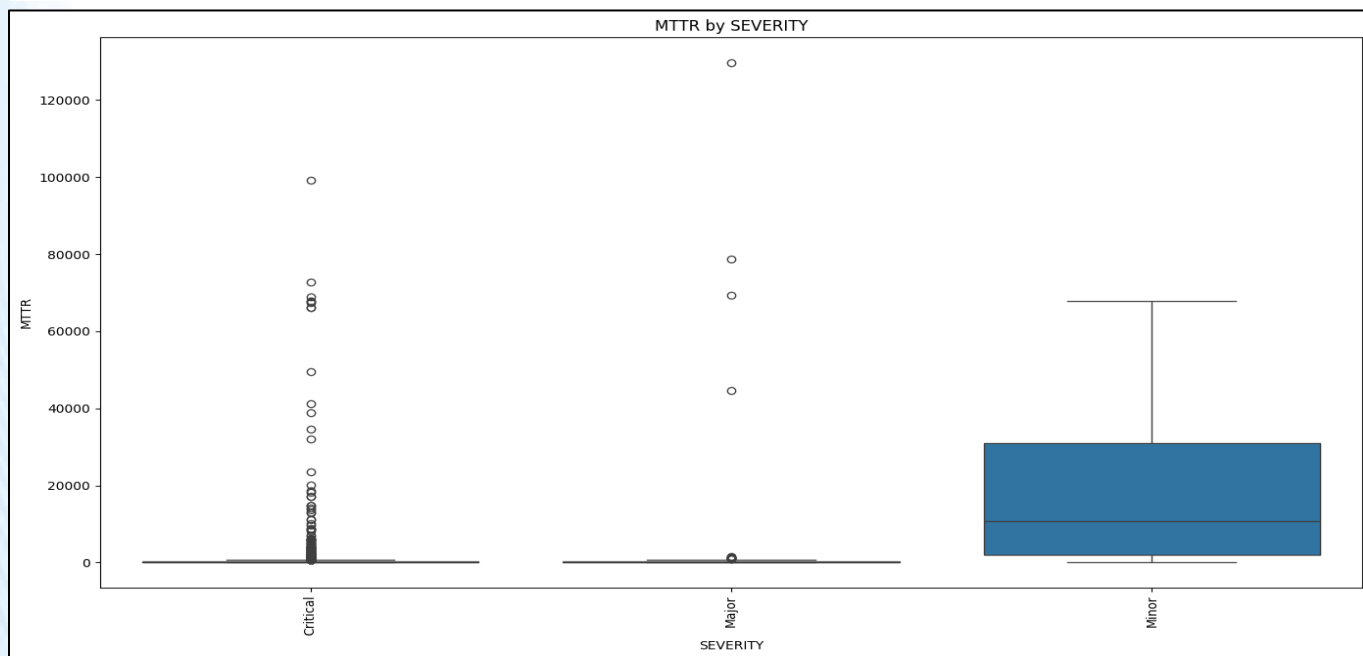
نمودار 1: نمودار MTTR بر حسب SUBCAUSE

در نمودار ۱ محور عمودی معرف میزان *MTTR* و محور افقی معرف *SUBCAUSE* های مختلف است و داده های ثبت شده مربوط به هر تیکت روی نمودار به این صورت نمایش داده شده اند که داده های پرت نسبت به میانگین در هر *SUBCAUSE* به شکل دایره نمایش داده شده و بخشی از داده ها در هر *SUBCAUSE* که متراکم بوده به صورت باکس آبی رنگ نمایش داده شده که خط بالا و پایین باکس ها نشان دهنده بیشینه و کمینه داده ثبت شده در آن *SUBCAUSE* است و خط وسط هر باکس نشان دهنده میانه مربوط به *MTTR* های هر *SUBCAUSE* است.

در ادامه برای نمایش واضح تر، ۵ تا از *SUBCAUSE* هایی که بیشترین میانگین *MTTR* برای آن ها ثبت شده است در نمودار ۲ نمایش داده شده است.

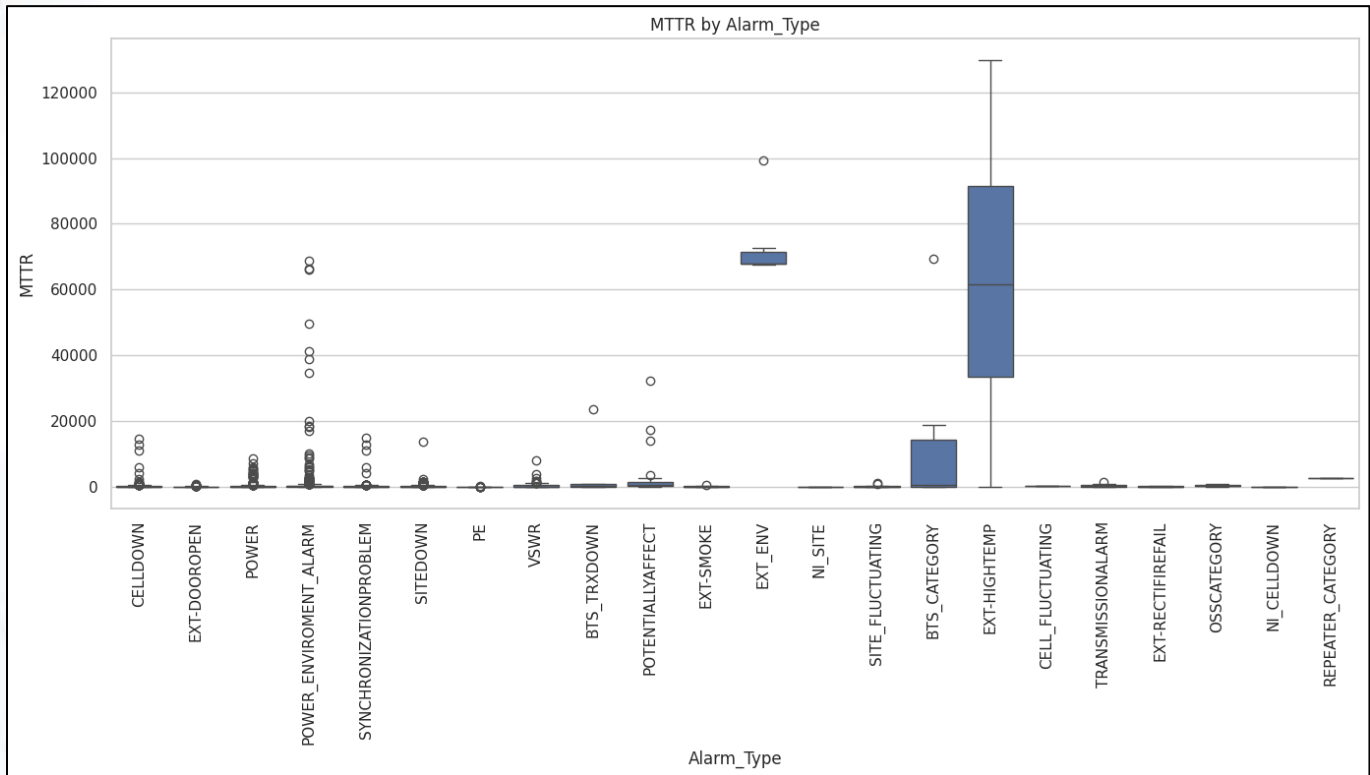


نمودار 2: نمودار MTTR بر حسب 5 SUBCAUSE با بیشترین میانگین



نمودار 3: نمودار MTTR بر حسب SEVERITY

نمودار ۳ نمایش دهنده میزان $MTTR$ های ثبت شده برای تیکت ها به ازای ۳ نوع مختلف $SEVERITY$ موجود است. که همان طور که از نمودار ها استنتاج می شود پراکندگی داده ها برحسب $SEVERITY$ زیاد است و برداشت آماری خاصی از آن ها نمی توان انجام داد.



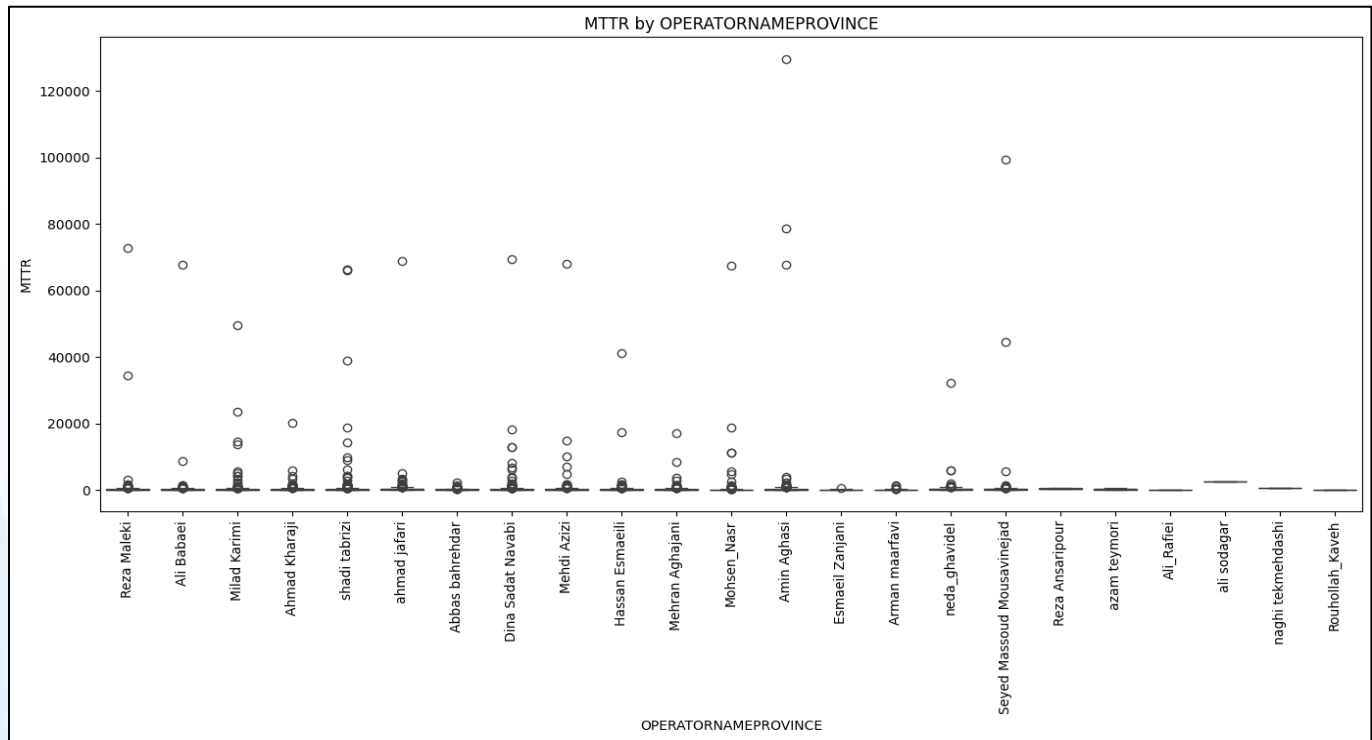
نمودار 4: میزان $MTTR$ برحسب Alarm-Type

نمودار ۴ میزان $MTTR$ های ثبت شده برای تیکت ها را برحسب $Alarm-Type$ آن ها نمایش می دهد. که همانطور که از نمودار مشاهده می شود ۳ تا از $Alarm-Type$ با بیشترین میزان $MTTR$ عبارتند از:

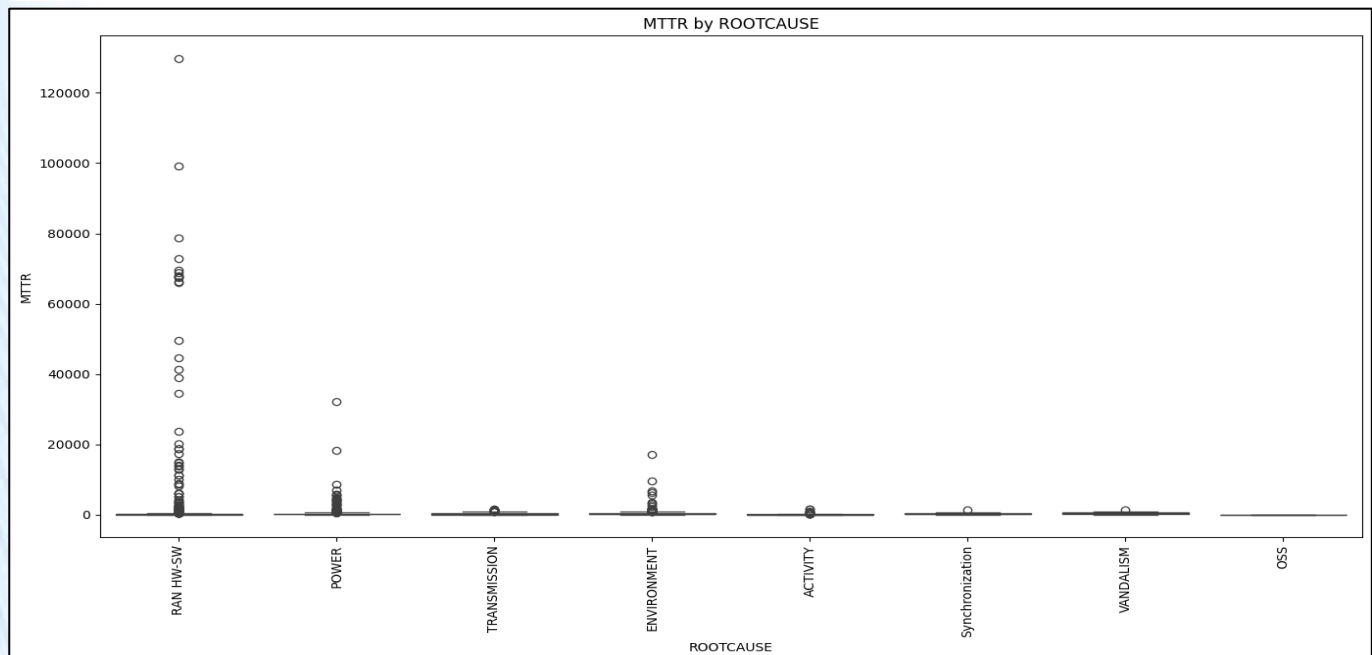
۱. $EXT-HIGHTEMP$
۲. EXT_ENV
۳. $POWER_ENVIROMENT_ALARM$

در نمودار ۵ همانطور که در ادامه آورده شده میزان $MTTR$ تیکت های ثبت شده مرتبط با هر $OPERATORNAMEPROVINCE$ نمایش داده شده است. البته با توجه به این که $MTBRTC_USER$ یا همان شخصی که تیکت ها را باز کرده است اسم خودش را در سیستم به عنوان

ثابت OPERATORNAMEPROVINCE کرده باعث شده است که داده های مربوط به ستون OPERATORNAMEPROVINCE صحت نداشته باشند و اطلاعاتی نسبت به شخصی که به سایت مراجعه کرده برای انجام تعمیرات سایت در دسترس نباشد در صورتی که این اطلاعات می توانست کمک زیادی در ارزیابی و بهبود میزان MTTR داشته باشد.



نمودار 5: میزان MTTR ثبت شده به ازای هر اپراتوری که به سایت مراجعه کرده است



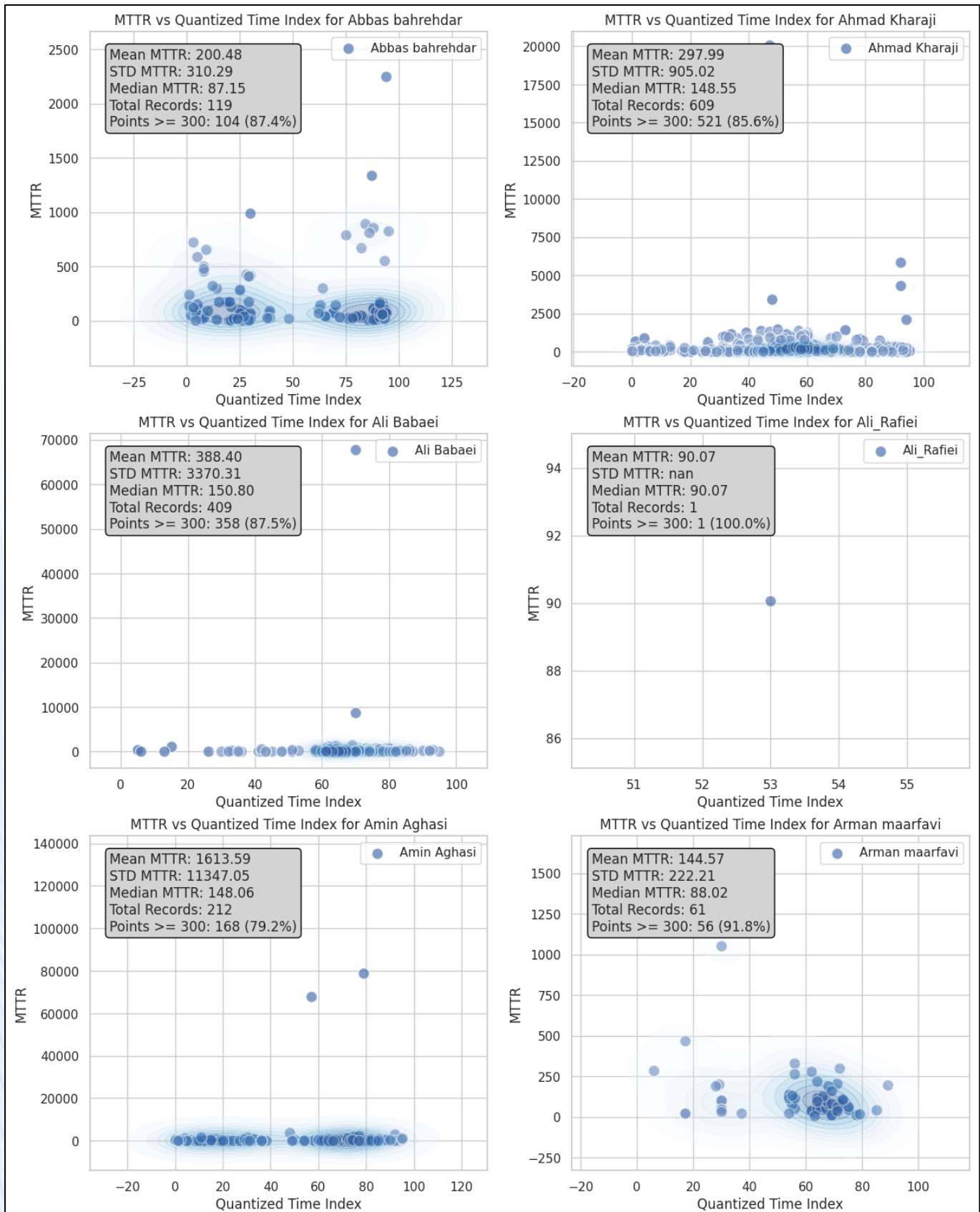
نمودار 6: میزان MTTRهای ثبت شده به ازای هر ROOTCAUSE

نمودار ۶ بیانگر میزان MTTR تیکت های ثبت شده به ازای ROOTCAUSE های مختلف است که همانطور که از میزان پراکندگی داده ها استنتاج می شود به ترتیب ۳ ROOTCAUSE ای که بیشترین میانگین را داشته اند عبارتند از:

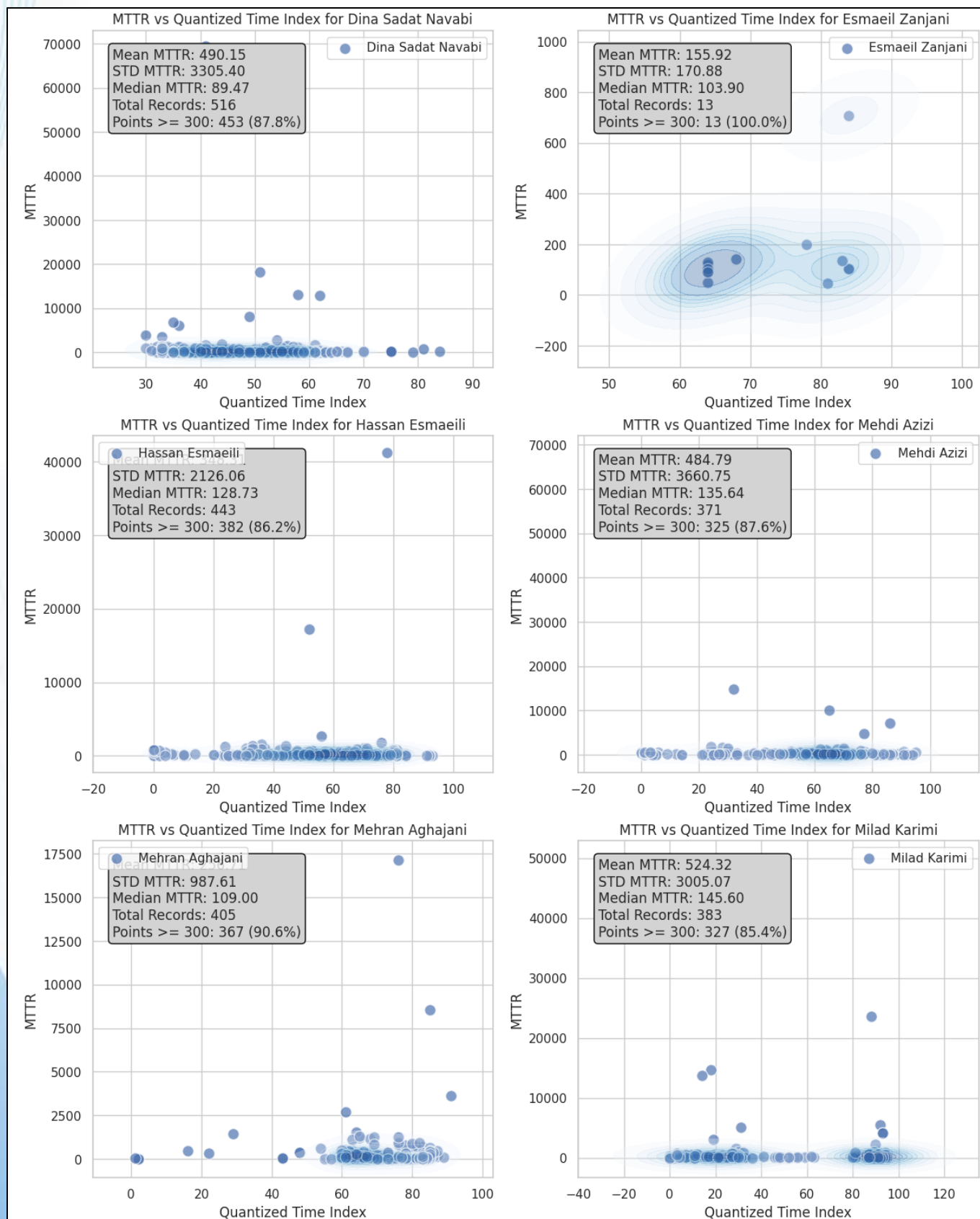
۱. RAN HW-SW
۲. POWER
۳. ENVIRONMENT

حال در ادامه نمودار عملکرد اپراتورهای ارسال شده به سایت ها را بر حسب MTTR کسب شده در زمان های مختلف شبانه روز نمایش می دهیم. (نمودارهای ۷، ۸، ۹، ۱۰) هر نمودار مربوط به یک اپراتور است و محور افقی نمودار معرف زمان کوانتیزه شده بازشدن تیکت است و محور عمودی معرفی میزان MTTR ثبت شده برای آن تیکت و داده ها به صورت دایره نمایش داده شده اند و با رنگ آبی و هاله حاشیه بندی شده است. همچنین داده های آماری مربوط به میانگین، انحراف معیار، میانه، تعداد کل تیکت های آن شخص و میزان درصدی از تیکت های شخص که امتیاز ۳۰۰ گرفته اند در بالای نمودار نمایش داده شده است.

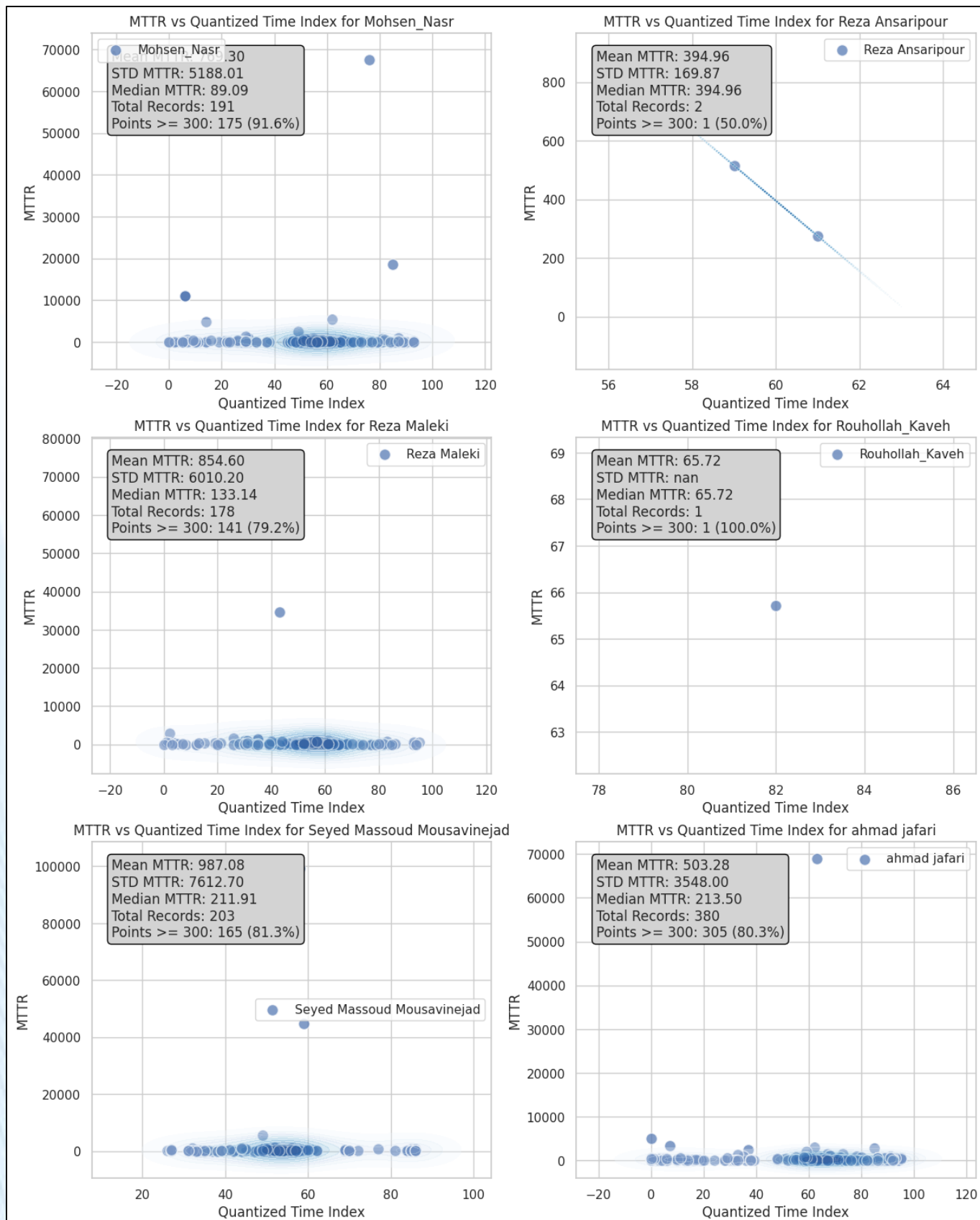
البته همانطور که توضیح داده شد چون اسم اپراتورهای مراجعه کرده به سایت به درستی ثبت نشده است این نمودارها معنای درستی برای ارزیابی عملکرد ارائه نمی دهند ولی اگر داده به درستی داخل سامانه ثبت می شد این نمودارها می توانستند به ارزیابی و بهبود سیستم کمک کنند.



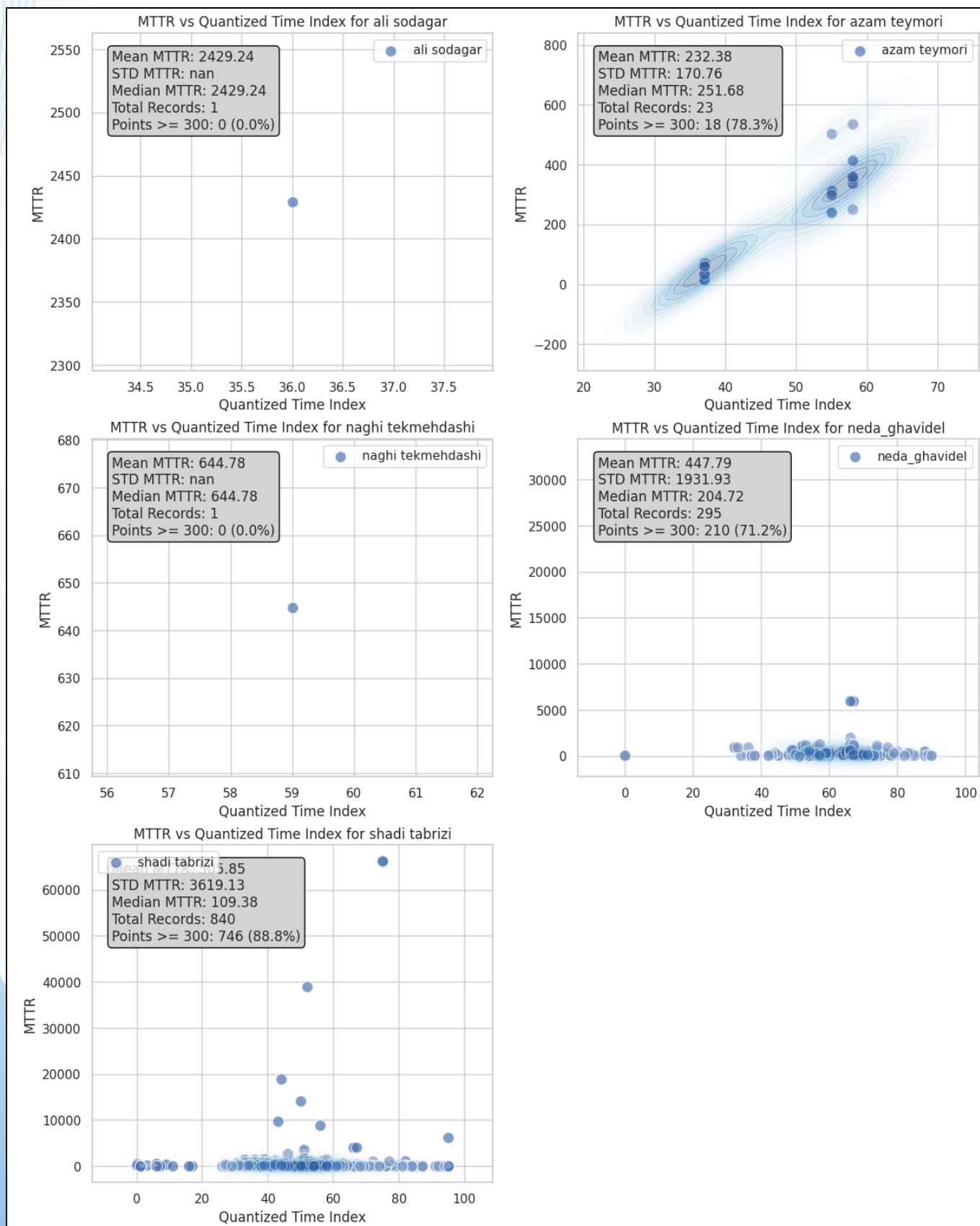
نمودار 7: نمودارهای نیکیت های ثبت شده به ازای هر اپراتور به صورت جداگانه



نمودار 8: نمودارهای تیکت های ثبت شده به ازای هر اپراتور به صورت جداگانه

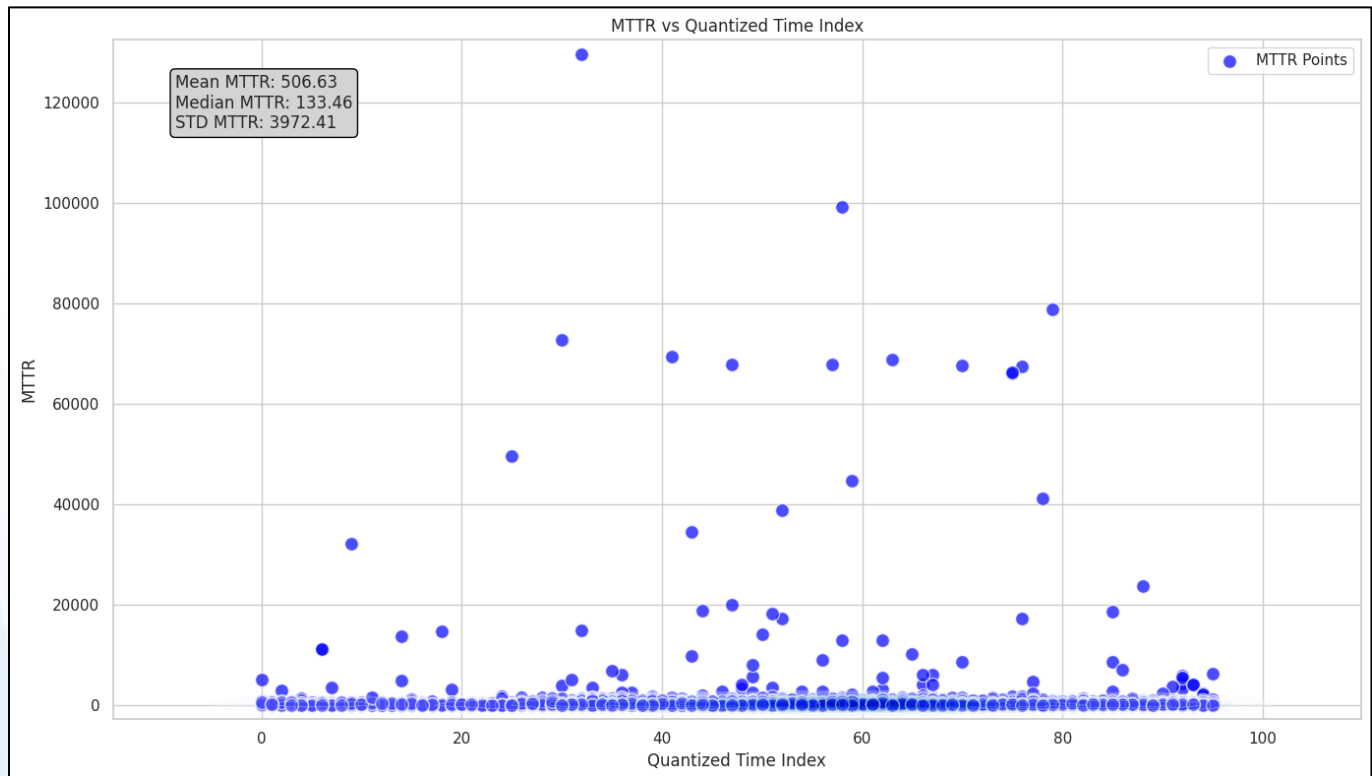


نمودار 9: نمودارهای تیکت های ثبت شده به ازای هر اپراتور به صورت جداگانه



نمودار 10: نمودارهای تیکت های ثبت شده به ازای هر اپراتور به صورت جداگانه

در ادامه نمودار پراکندگی میزان $MTTR$ های ثبت شده را بر اساس بازه زمانی ای که تیکت مربوطه داخل سیستم ثبت شده است نمایش می دهیم: (نمودار ۱۱)

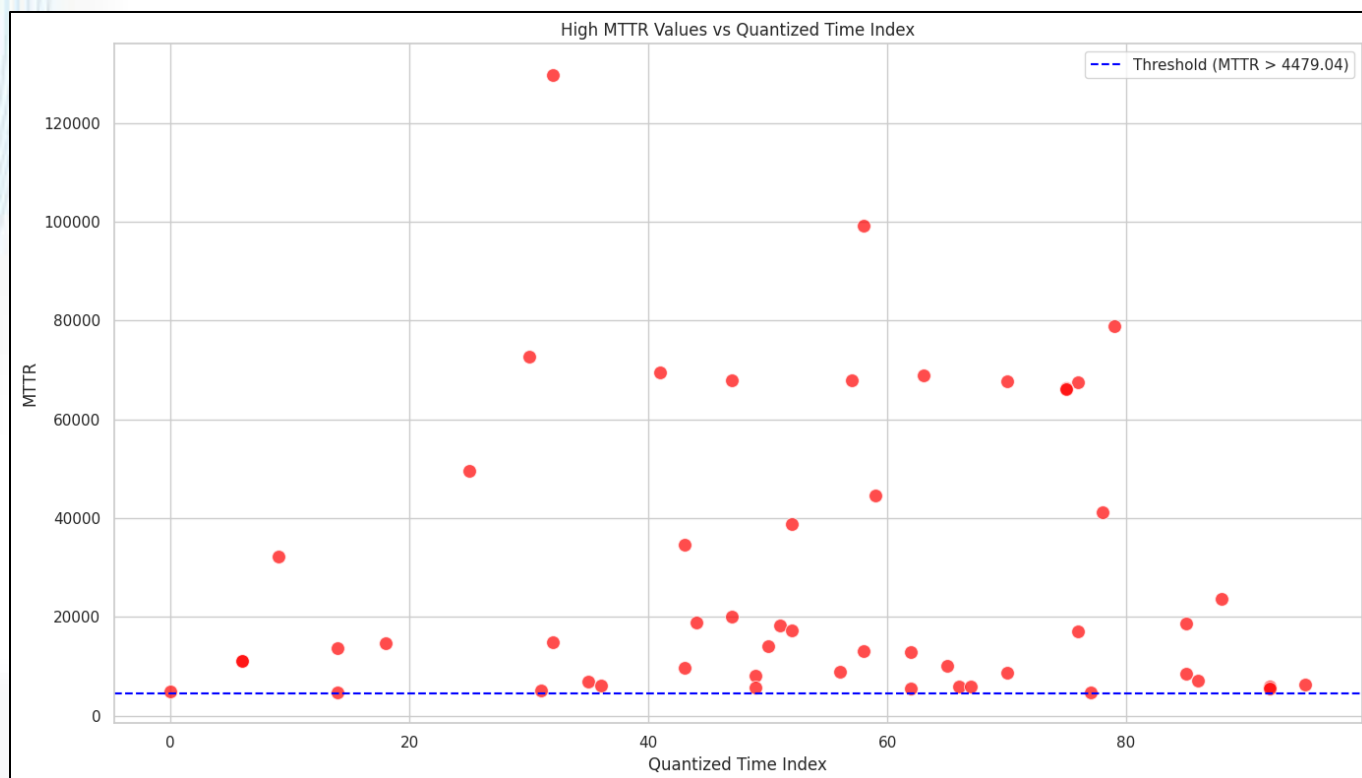


نمودار ۱۱: میزان $MTTR$ های ثبت شده در بازه های مختلف روز

داده های زمانی کوانتیزه شده همانطور که قبل تر توضیح داده شد به عنوان مثال به صورت زیر است:

- اندیس ۰ مربوط به ساعت ۰۰:۱۵
- اندیس ۲۰ مربوط به ساعت ۵:۰۰
- اندیس ۴۰ مربوط به ساعت ۱۰:۰۰
- اندیس ۶۰ مربوط به ساعت ۱۵:۰۰
- اندیس ۸۰ مربوط به ساعت ۲۰:۰۰

سپس برای تحلیل بهتر داده ها، صرفا داده هایی را که از میانگین $MTTR$ موجود فاصله زیادی داشته باشند را نمایش می دهیم: (نمودار ۱۲)



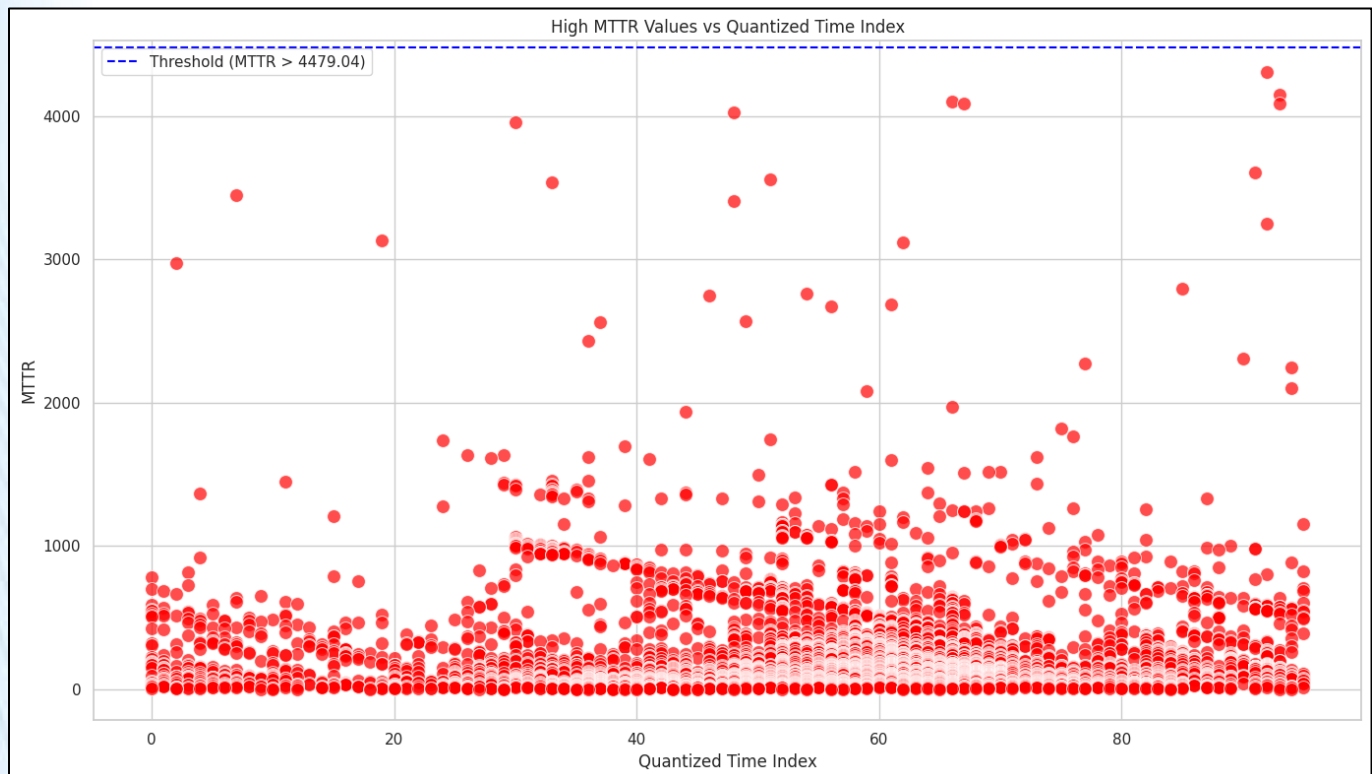
نمودار ۱۲: میزان MTTR های ثبت شده در بازه های مختلف روز (با فیلتر داده های با میزان MTTR های بیشتر)

و در ادامه داده های قرمز رنگ نمایش داده شده در نمودار ۱۲ را گزارش می دهیم:

	MTTR	CODESITE	OPERATORNAME	PROVINCE
4069	129745.01	ES1G1531	Amin	Aghasi
1211	99279.03	ES0747	Seyed Massoud	Mousavinejad
4031	78786.52	ES1G1561	Amin	Aghasi
2819	72762.31	ES0220	Reza	Maleki
4749	69462.13	ES1G1489	Dina Sadat	Navabi
1182	68911.62	ES1415	ahmad	jafari
4751	67931.77	ES1489	Mehdi	Azizi
118	67885.18	ES1637	Amin	Aghasi
5159	67733.59	ES1827	Ali	Babaei
5017	67531.28	ES1484	Mohsen	Nasr
5158	66231.73	ES1827	shadi	tabrizi
5014	66119.95	ES1484	shadi	tabrizi
4748	49531.94	ES1489	Milad	Karimi
2394	44648.72	ES1G1646	Seyed Massoud	Mousavinejad
4073	41215.38	ES1531	Hassan	Esmaeili
1196	38888.93	ES0747	shadi	tabrizi
5363	34576.10	ES1666	Reza	Maleki
3645	32166.73	ES0469	neda	ghavidel
3238	23630.68	ES0289	Milad	Karimi
4516	20056.00	ES0012	Ahmad	Kharaji
605	18812.01	ES1G1110	shadi	tabrizi
2823	18656.67	ES0220	Mohsen	Nasr

2658	18206.12	ES0499	Dina Sadat Navabi
613	17247.06	ES1110	Hassan Esmaeili
2322	17152.16	ES1802	Mehran Aghajani
5332	14876.11	ES1232	Mehdi Azizi
289	14706.20	ES0901	Milad Karimi
3242	14164.38	ES0289	shadi tabrizi
5334	13724.79	ES1232	Milad Karimi
1751	13003.98	ES0074	Dina Sadat Navabi
1747	12940.13	ES0074	Dina Sadat Navabi
3088	11132.67	ES1329	Mohsen_Nasr
3090	11123.23	ES1329	Mohsen_Nasr
2750	10086.10	ES0007	Mehdi Azizi
1502	9698.47	ES1457	shadi tabrizi
3366	8907.08	ES1551	shadi tabrizi
4042	8686.82	ES1561	Ali Babaei
1250	8533.38	ES1253	Mehran Aghajani
3233	8103.62	ES0289	Dina Sadat Navabi
1251	7089.38	ES1253	Mehdi Azizi
991	6847.20	ES0995	Dina Sadat Navabi
2751	6256.42	ES0007	shadi tabrizi
797	6084.09	ES0527	Dina Sadat Navabi
2132	5975.94	ES1036	neda_ghavidel
2112	5962.15	ES1036	neda_ghavidel
865	5868.94	ES1200	Ahmad Kharaji
873	5689.39	ES1200	Seyed Massoud Mousavinejad
1252	5539.93	ES1253	Milad Karimi
1562	5518.84	ES0535	Mohsen_Nasr
4780	5087.07	ES1190	Milad Karimi
1667	5016.61	ES1823	ahmad jafari
2906	4806.13	ES1898	Mohsen_Nasr
869	4748.60	ES1200	Mehdi Azizi

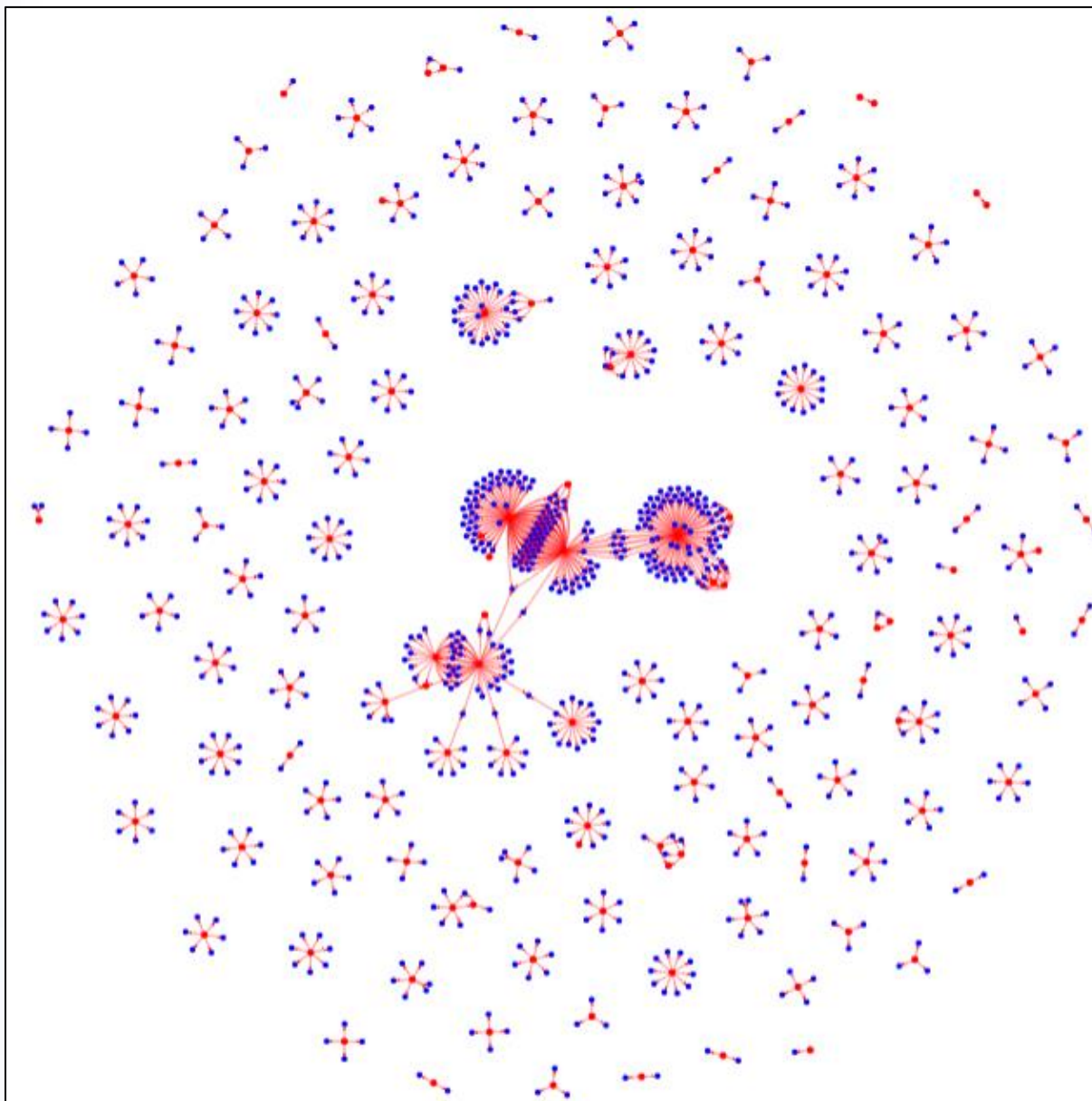
حال برای تحلیل کامل تر داده هایی را که حول میانگین $MTTR$ موجود بودند را نمایش می دهیم: (نمودار ۱۳)



نمودار ۱۳: میزان $MTTR$ های ثبت شده در بازه های مختلف روز (با فیلتر داده های حول میانگین)

■ گراف سایت‌ها

فایل HTML مربوط به گراف شبکه سایت‌ها ضمیمه شده است.

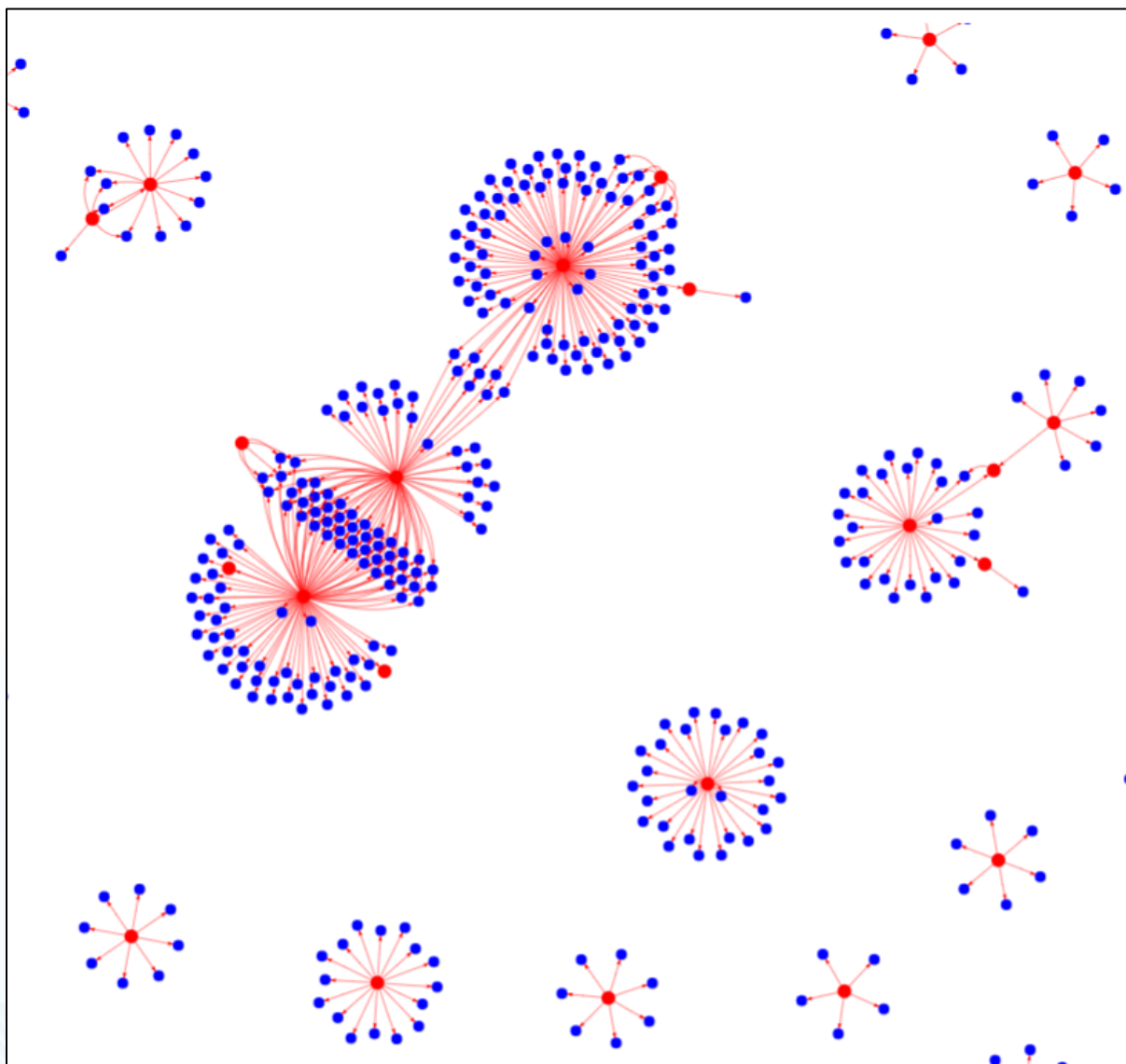


شکل 1: شبکه نودها. نودهای آبی متصل به نود قرمز بیانگر این است که با خرابی نود قرمز، نود آبی نیز به مشکل می‌خورد.

همانطور که در شکل ۱ مشاهده میشود شامل گراف سایت‌هایی است که تیکت برای آنها ارسال شده است و سایت‌های که در اثر این خرابی تحت تاثیر (*impactsitelist*) قرار گرفته‌اند.

سایت هایی که خطا در آنها رخ داده شده است با قرمز و سایت های تحت تاثیر با آبی نمایش داده شده است.

با توجه به شکل ۱ سایت هایی وجود دارند که تعداد سایت های تحت تاثیر آنها بیشتر از باقی است که میتواند معیاری برای اهمیت بیشتر این سایت ها باشد که گراف این سایت ها در شکل ۲ نمایش داده شده است.



شکل ۲: گراف سایت هایی که سایت های تحت تاثیر آنها بیشتر بوده است

با توجه به شکل ۲ سایت هایی که ارتباط و تاثیر بیشتری با سایت های زیر مجموعه خودشان داشتند عبارتند از:

۱. ES1G0225

۲. ES1G0003

۳. ES1G0220



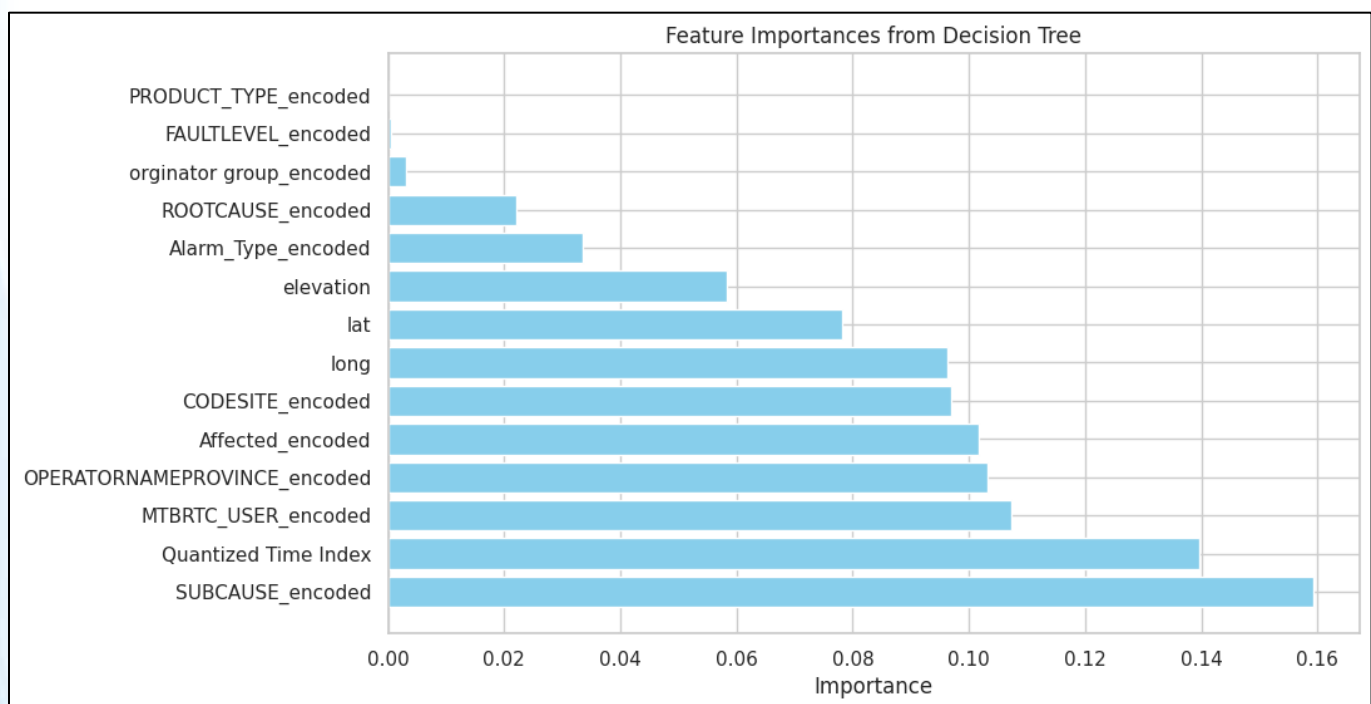
در صورتی که اطلاعات بیشتری در مورد نوع خرابی پیش آمده در سایت های تحت تاثیر در اختیار داشته باشیم می توان تحلیل کامل تری در این حوزه ارائه داد.

■ تحلیل رفتار عوامل اثر گذار

با استفاده از مدل *decision tree* بر اساس ۲ تارگت: امتیاز و *MTTR* حساسیت پارامترها برای هر تارگت به عوامل مختلف یافته و به صورت جداگانه تحلیل شده‌اند.

■ امتیاز

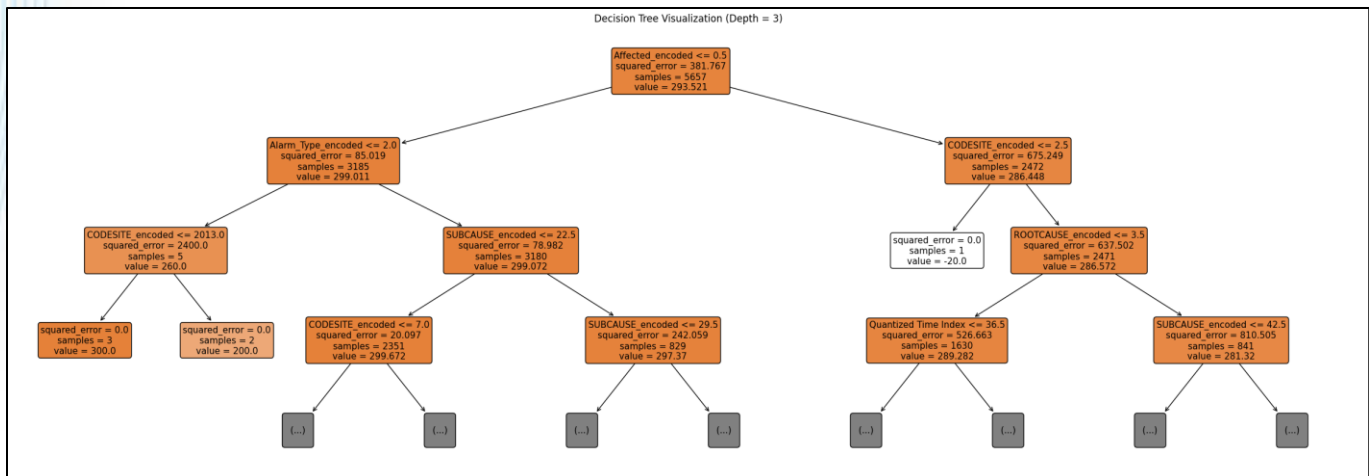
در این بخش ما به تحلیل عوامل موثر در امتیاز پرداختیم. تارگت *decision tree* در این بخش امتیاز در نظر گرفته شده است.



نمودار ۱۴: میزان تاثیر عوامل مختلف بر امتیاز

همانطور که در نمودار بالا مشاهده می‌کنید عوامل موثر در محور عمودی لیست شده‌اند و محور افقی میزان تاثیر آنها بر امتیاز را نمایش می‌دهند که *SUBCAUSE*، *Quantized Time index* و *MTBRTC_USER* بیشترین تاثیر را بر امتیاز داشته‌اند.

در ادامه لایه‌های ابتدایی درخت تصمیم‌گیری امتیاز نمایش داده شده است. در این درخت در هر مرحله عوامل تاثیرگذار از یکدیگر متمایز می‌شود و عوامل مربوط به آن در لایه بعدی به صورت جداگانه بررسی می‌شود و معیاری برای تاثیر آن نسبت داده می‌شود.

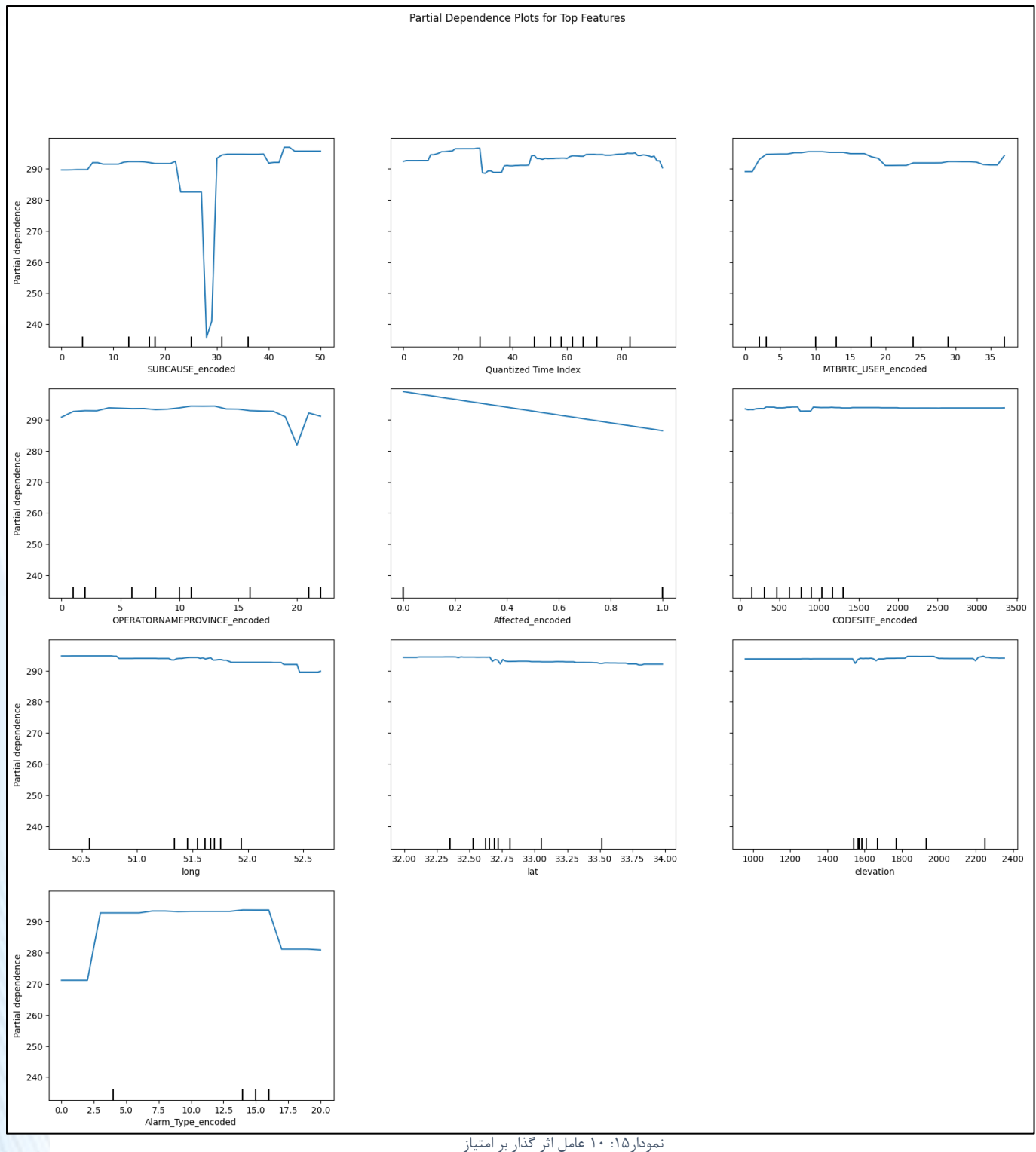


شکل 3: درخت تصمیم گیری امتیاز

در ادامه از تابع *PartialDependenceDisplay* استفاده شده است. عملکرد این تابع به این صورت است که به عنوان مثال برای یک عامل گسسته مثل *SUBCAUSE* می آید و بررسی می کند که به ازای هر *SUBCAUSE* میانگین *MTTR* تیکت های مربوط به همان *SUBCAUSE* را محاسبه می کند و سپس آن را روی محور عمودی به عنوان *PartialDependence* نمایش می دهد و محور افقی نمودار نیز مربوط به مقادیر مختلفی است که یک عامل می تواند داشته باشد که یک سری عوامل مثل *SUBCAUSE* گسسته هستند و عدد نمایش داده شده روی محور افقی مربوط به *mapping* ای است که در بخش اصلاح داده انجام شده است و در پیوست انتهایی گزارش لیست مپ شده به ازای هر عامل آورده شده است، و برخی داده ها هم مانند *lat* پیوسته هستند.

حال اثر گذاری بازه های مختلف ۱۰ *Feature* اثر گذار در میزان امتیاز کسب شده در تیکت ها را نمایش می

دهیم:



همانطور که از نمودار ۱۵ استنتاج می شود به جز نمودار *SUBCAUSE* و *Alarm-Type* نمودار *PartialDependence* باقی عوامل تقریباً یکنواخت بوده اند و به این معنی است که حداقل با این داده ها



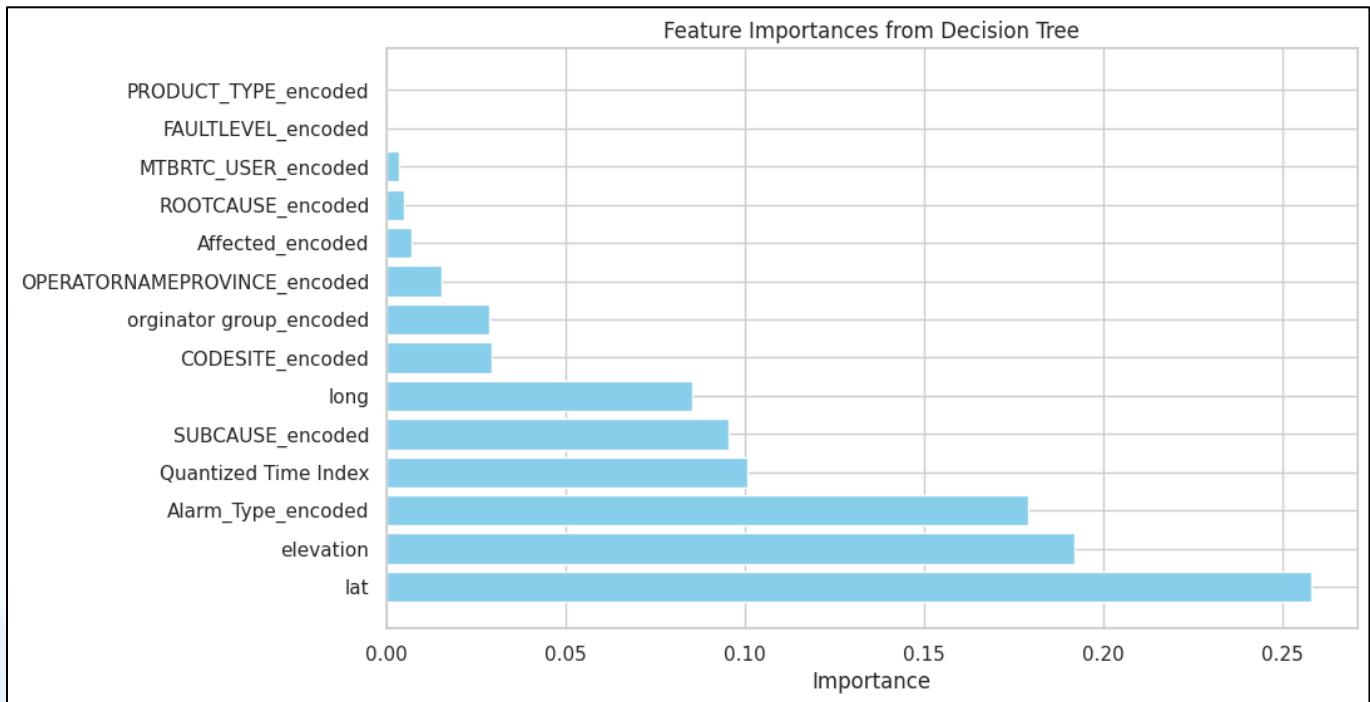
نمی‌توان میزان اثرگذاری بیشتر بخشی از این عوامل را متوجه شد اما برای نمودار مربوط به *SUBCAUSE* مشاهده می‌شود که به ازای اندیس ۲۷ اثر گذاری کمتری داشته است یا برای نمودار *Alarm-Type* مشاهده می‌شود که اندیس‌های ۳ تا ۱۷ اثر گذاری بیشتری نسبت به باقی اندیس‌ها داشتند.

و حال مکان‌هایی که در نمودارها بیشینه رخ داده است را می‌آوریم:

```
Maximum partial dependence value for SUBCAUSE_encoded: 296.8885561458796 at feature value: TRANSMISSION-TCX LINK FLUCTUATE
Maximum partial dependence value for Quantized Time Index: 296.5657364341051 at feature value: 27
Maximum partial dependence value for MTBRTC_USER_encoded: 295.4611582843759 at feature value: Arman maarfavi
Maximum partial dependence value for OPERATORNAMEPROVINCE_encoded: 294.3468013248624 at feature value: Reza Ansaripour
Maximum partial dependence value for Affected_encoded: 299.0109890109891 at feature value: NO
Maximum partial dependence value for CODESITE_encoded: 294.0811974872647 at feature value: ES1019
Maximum partial dependence value for long: 294.71024737614766 at feature value: 50.412256262626265
Maximum partial dependence value for lat: 294.4013510791087 at feature value: 32.435669724444445
Maximum partial dependence value for elevation: 294.62139111738765 at feature value: 2240.272727272727
Maximum partial dependence value for Alarm Type encoded: 293.7229945595257 at feature value: POWER
```

MTTR

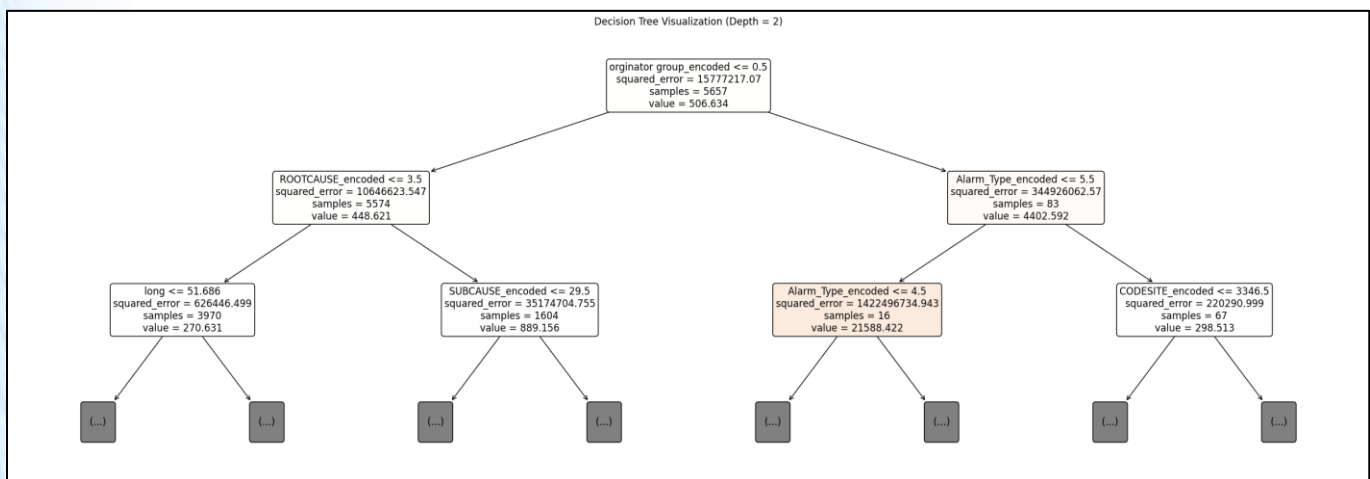
در این بخش همانند بخش قبلی به عوامل موثر بر MTTR را بررسی می کنیم.



نمودار ۱۶: میزان تاثیر عوامل مختلف بر MTTR

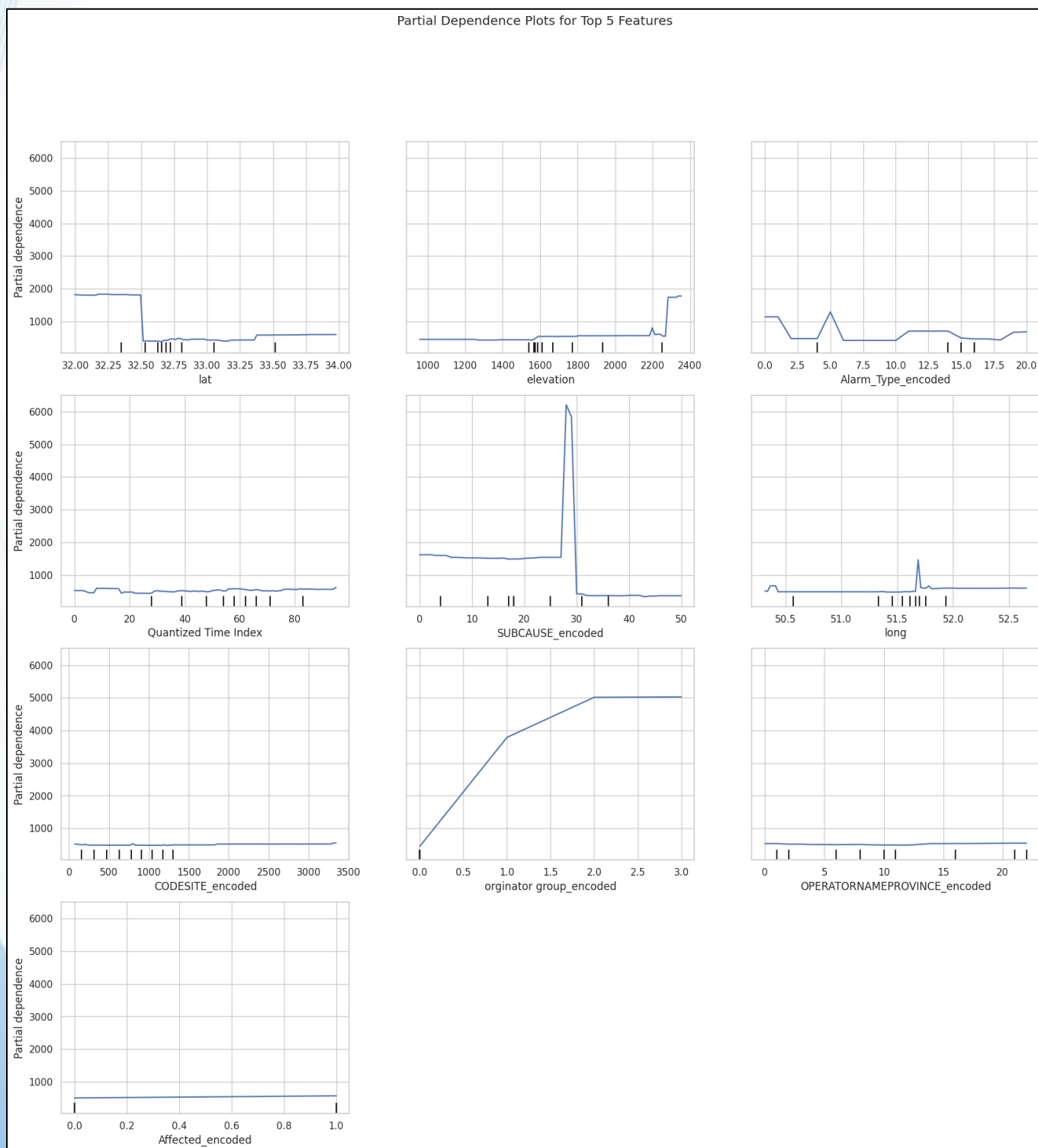
همانطور که در این نمودار دیده میشود محور عمودی عوامل موثر بر MTTR را نمایش میدهد و محور افقی میزان تاثیر آن ها را نمایش میدهد که lat، elevation و Alarm_Type جزو بیشترین عوامل تاثیر گذار هستند که دو عامل اول عملاً به میزان اهمیت موقعیت مکانی سایت بر MTTR دلالت دارند.

لایه های ابتدایی درخت تصمیم گیری برای MTTR نیز به شکل زیر است:



شکل ۴: درخت تصمیم گیری MTTR

حال اثر گذاری بازه های مختلف ۱۰ *Feature* اثر گذار در میزان امتیاز کسب شده در تیکت ها را نمایش می دهیم:



نمودار ۱۶: ۱۰ عامل اثر گذار بر MTTR

نمودار ۱۶ شامل نمودار ۱۰ عامل اثرگذارتر در میزان MTTR است. که برای هر کدام از عوامل محور افقی معرفی بازه ای از اعداد است که آن عامل می توانسته شامل شود و محور عمودی نیز PartialDependence است که مشاهده می شود به جز نمودار lat,long,SUBCAUSE,Originator-group باقی نمودارها توزیع یکنواخت و ثابتی داشته اند و نمی توان حداقل با این میزان داده های داده شده تفاوت زیادی در اثر گذاری بازه های مختلف آن عوامل متوجه شد.

حال مکان هایی که در نمودار ها بیشینه رخ داده است را می آوریم:

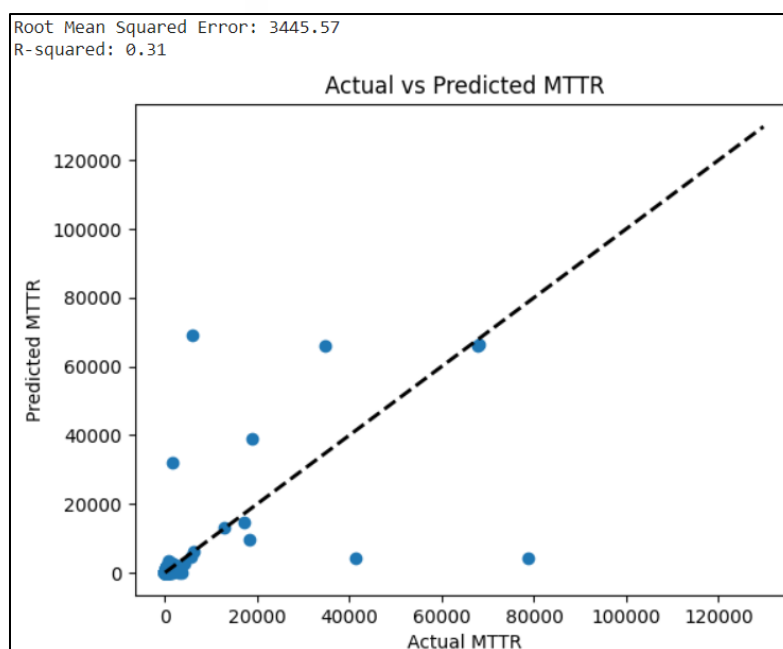
```
Maximum partial dependence value for lat: 1833.1181213760194 at feature value: 32.175147145454545
Maximum partial dependence value for elevation: 1777.9307200044889 at feature value: 2338.909090909091
Maximum partial dependence value for Alarm_Type_encoded: 1290.350430440057 at feature value: EXT-HIGHTEMP
Maximum partial dependence value for Quantized Time Index: 617.5148829088349 at feature value: 95
Maximum partial dependence value for SUBCAUSE_encoded: 6209.662382819409 at feature value: RAN HW&SW-BTS UNIT FAULT-KOOYE KAN
Maximum partial dependence value for long: 1460.5606611187966 at feature value: 51.68755080808081
Maximum partial dependence value for CODESITE_encoded: 553.3800630457931 at feature value: ES6541
Maximum partial dependence value for originator_group_encoded: 5026.996947791165 at feature value: NAK_NetworkOperation_FO_TRS
Maximum partial dependence value for OPERATORNAMEPROVINCE_encoded: 549.3592455118107 at feature value: shadi tabrizi
Maximum partial dependence value for Affected encoded: 572.975894515923 at feature value: YES
```

■ مدل‌های پیش‌بینی

در این بخش مدل‌هایی طراحی شده است تا بر اساس داده‌های تیکت بتوان میزان $MTTR$ را پیش‌بینی کرد، برای صحت‌سنجی عملکرد مدل‌ها داده‌هایی که در اختیار داشتیم به دو قسمت تقسیم شد و مدل با استفاده از یک بخش پروسه یادگیری را انجام داد و با استفاده از بخش دیگر میزان دقت آن اندازه‌گیری شد.

از ۳ مدل برای این کار استفاده شده است:

• Decision Tree Regressor



شکل 4: پیش‌بینی بر پایه Decision Tree Regressor

خروجی مدل را در شکل بالا مشاهده می‌کنید. ابتدا میزان خطای مدل و میزان خطی بودن مقادیر پیش‌بینی شده با مقادیر واقعی گزارش شده است.

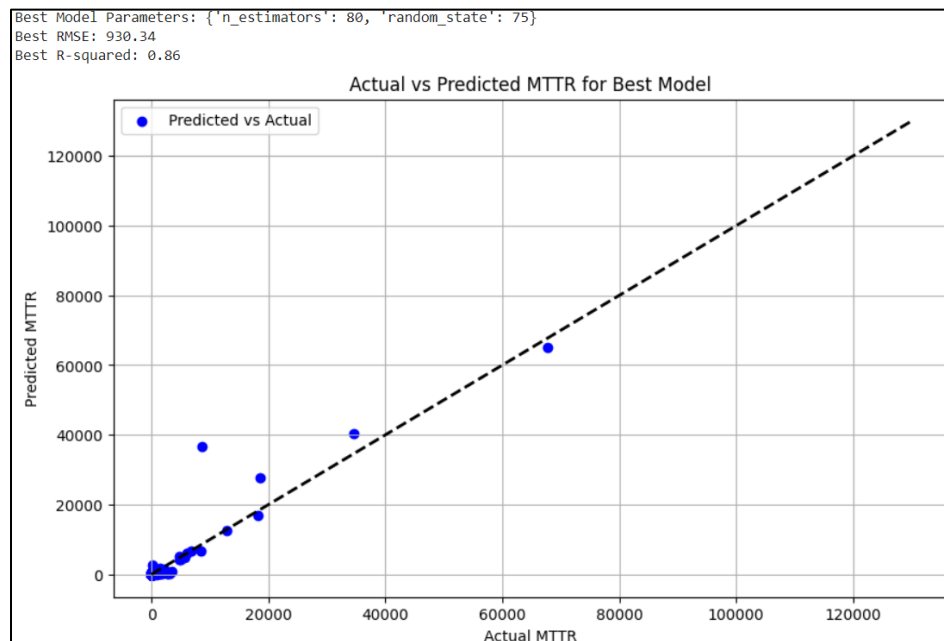
نموداری که مشاهده می‌کنید نمودار مقادیر پیش‌بینی شده بر حسب داده‌هایی واقعی است هر چقدر نقاط این نمودار به خط نقطه چین نزدیک‌تر باشد نشان‌دهنده عملکرد بهتر مدل است.

XGBOOST •

Best Parameters: {'colsample_bytree': 0.7, 'learning_rate': 0.2, 'max_depth': 5, 'n_estimators': 300, 'subsample': 1.0}
Mean Squared Error: 7234734.863438698
Mean Absolute Error: 439.3843479674143
R-squared Score: 0.5773178616533203

XGBOOST دیگر مدلیست که برای پیش بینی از آن استفاده شده است. همانطور که مشاهده میکنید این مدل نسب به مدل قبلی خروجی بهتری داشته است.

Random Forest Regressor •



شکل 5: پیش بینی برپایه Random Forest Regressor

و در نهایت از مدل Random Forest Regressor استفاده شد که مدل پیچیده تری نسبت به مدل های قبلی است به همین دلیل بهترین نتیجه را در بین مدل های استفاده شده دارد.



پیشنهادهات

■ پیوست (لیست *mapping* اندیس‌ها به عناوین)

ROOTCAUSE	
Original Value	Encoded Value
ACTIVITY	0
ENVIRONMENT	1
OSS	2
POWER	3
RAN HW-SW	4
Synchronization	5
TRANSMISSION	6
VANDALISM	7

MTBRTC-USER			
Original Value	Encoded Value	Original Value	Encoded Value
Abbas Rezaeian Rad	0	Mehrdad Rahimi	20
Abbas bahrehdar	1	Milad Karimi	21
Ahmad Kharaji	2	Mohammad Miranzadeh	22
Ali Babaei	3	Mohammadreza Alami	23
Ali Rahmani	4	Mohsen Hedayatinezhad	24
Alireza Mohandesi	5	Mohsen_Nasr	25
Amin Aghasi	6	Morteza Ghodratpour	26
Arash Zerehgar	7	Niloufar Ghalandarian	27
Arman Kiani	8	Pouyan Haghdoust	28
Arman maarfavi	9	Reza Maleki	29
Dina Sadat Navabi	10	Seyed Massoud Mousavinejad	30
Ext.Abbas.Rafiei	11	Soheil Pourrajaei	31
Ext.Morteza.Ghodratpoor	12	ahmad jafari	32
Hassan Esmaeili	13	ali sodagar	33
Jalal Aeineh	14	azam teymori	34
Kargar, Mahdi	15	eisa asadi	35
Khatamian, Ali	16	majid mahmoodabadi	36
Mehdi Azizi	17	neda_ghavidel	37
Mehdi Vaezi	18	shadi tabrizi	38
Mehran Aghajani	19		



SUBCAUSE			
Original Value	Encoded Value	Original Value	Encoded Value
ACTIVITY-NI ACTIVITY	0	POWER-TCX RECTIFIER FAULT	26
ACTIVITY-PLANNED ACTIVITY	1	RAN HW&SW-BTS UNIT FAULT	27
ENVIRONMENT-AIR CONDITION ISSUE	2	RAN HW&SW-BTS UNIT FAULT-KOOYE KAN	28
ENVIRONMENT-OUTDOOR CABINET COOLING SYSTEM ISSUE	3	RAN HW&SW-BTS UNIT FAULT-KOOYE KAN-ON SITE	29
ENVIRONMENT-TCX-AIR CONDITION ISSUE	4	RAN HW&SW-FEEDER,JUMPER,CONNECTOR FIBER FAULT	30
GPS-JAMMER ISSUE	5	RAN HW&SW-FLM HARDWARE ACTIVITY	31
OSS-REQUEST	6	RAN HW&SW-FLM SOFTWARE ACTIVITY	32
OUT OF CR SCOPE TIME- FLM PROVINCE	7	RAN HW&SW-FO SOFTWARE ACTIVITY	33
OUT OF CR SCOPE TIME- NI, ROLL OUT	8	RAN HW&SW-HARDWARE RESET	34
OUT OF CR SCOPE TIME- NPO-RAN	9	SYNC-GPS	35
OUT OF CR SCOPE TIME- TCX[TCI, TCT,...]	10	TRANSMISSION-BAD WEATHER	36
POWER-ACMDB FAULT	11	TRANSMISSION-CABLING,CONNECTOR FAULT	37
POWER-ACPDB FAULT	12	TRANSMISSION-FLM HARDWARE ACTIVITY	38
POWER-BATTERY FAULT	13	TRANSMISSION-FLM SOFTWARE ACTIVITY	39
POWER-BATTERY MISSING	14	TRANSMISSION-HUB SITE POWER ISSUE	40
POWER-DCPDB FAULT	15	TRANSMISSION-MICROWAVE MISSALIGNMENT	41
POWER-GENSET ISSUE	16	TRANSMISSION-TCX FIBER CUT	42
POWER-LACK OF BATTERY BACKUP	17	TRANSMISSION-TCX LINK FLUCTUATE	43
POWER-LOCAL POWER OUTAGE	18	TRANSMISSION-TCX PCM ISSUE	44
POWER-NO BATTERY BY DESIGN	19	TRANSMISSION-TCX POWER ISSUE	45
POWER-ONE PHASE POWER ISSUE	20	VANDALISM-VANDALISM	46
POWER-POWER OUTAGE-BAD WEATHER	21	VANDALISM-VANDALISM OF POWER COMPANY EQUIPMENT	47
POWER-RECTIFIER FAULT	22	WITHOUT CR- FLM PROVINCE	48
POWER-RECTIFIER MODULE FAULT	23	WITHOUT CR- NI, ROLL OUT	49
POWER-SOLAR CELL	24	WITHOUT CR- TCX[TCI, TCT, ...]	50
POWER-TCX POWER OUTAGE	25		

PRODUCT-TYPE	
Original Value	Encoded Value
RAN	0
TRS	1

orginator group	
Original Value	Encoded Value
FMSender	0
NAK_NO_FO	1
NAK_NetworkOperation_FO_RAN	2
NAK_NetworkOperation_FO_TRS	3



OPERATORNAMEPROVINCE	
Original Value	Encoded Value
Abbas bahrehdar	0
Ahmad Kharaji	1
Ali Babaei	2
Ali_Rafiei	3
Amin Aghasi	4
Arman maarfavi	5
Dina Sadat Navabi	6
Esmaeil Zanjani	7
Hassan Esmaeili	8
Mehdi Azizi	9
Mehran Aghajani	10
Milad Karimi	11
Mohsen_Nasr	12
Reza Ansaripour	13
Reza Maleki	14
Rouhollah_Kaveh	15
Seyed Massoud Mousavinejad	16
ahmad jafari	17
ali sodagar	18
azam teymori	19
naghi tekmehdashi	20
neda_ghavidel	21
shadi tabrizi	22

Affected	
Original Value	Encoded Value
NO	0
YES	1



FAULTLEVEL	
Original Value	Encoded Value
LEVEL2	0
LEVEL3	1
LEVEL4	2

Alarm_Type	
Original Value	Encoded Value
BTS_CATEGORY	0
BTS_TRXDOWN	1
CELLOWN	2
CELL_FLUCTUATING	3
EXT-DOOROPEN	4
EXT-HIGHTEMP	5
EXT-RECTIFIREFAIL	6
EXT-SMOKE	7
EXT_ENV	8
NI_CELLOWN	9
NI_SITE	10
OSSCATEGORY	11
PE	12
POTENTIALLYAFFECT	13
POWER	14
POWER_ENVIROMENT_ALARM	15
REPEATER_CATEGORY	16
SITDOWN	17
SITE_FLUCTUATING	18
SYNCHRONIZATIONPROBLEM	19
TRANSMISSIONALARM	20
VSWR	21