

MACHINE LEARNING MODELS

A power up from reactive IT service management

- ✦ Discover how no-code ML is helping service desk teams predict, prevent, and perform

Table of contents

■	Introduction	3
■	Mastering escalation risk with predictive precision	4
■	Premeditating the impact of change rollouts	9
■	Proactive asset health maintenance with ML-powered foresight	14
■	Balancing distributed infrastructure capacity with tailored predictions	17
■	Conclusion	24
■	About ManageEngine Analytics Plus	25

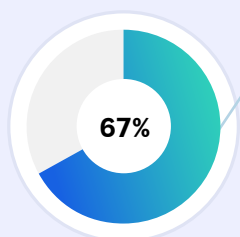
Introduction

Imagine driving through a bustling city road or cross-country with nothing but an old physical route-map sheet. You'd miss road closures, traffic snarls, and weather warnings—leaving you stranded, late, or rerouted at the worst possible moment.

That's what relying solely on traditional analytics feels like in today's complex ITSM landscape. While standard reports and dashboards can serve us well in understanding what went wrong, they fall short of answering a far more valuable question: What will happen next?

IT leaders are increasingly turning to ML not just to analyze the past, but to predict what's coming, prevent issues before they arise, and prepare with intelligent, data-driven strategies.

A recent survey by Deloitte states that^[1]



67% of IT leaders believe ML is crucial for enhancing service delivery and operational agility.

This e-book explores how ML supercharges traditional IT analytics, transforming it from a reactive reporting tool into a proactive strategic advantage, unpacking four high-impact use cases where IT can lead from the front with ML-driven analytics.

Mastering escalation risk with predictive precision

In today's dynamic IT landscape, incident management often feels like a constant game of catch-up. Most organizations find themselves in a reactive loop: a flood of tickets comes in; some are resolved, but a significant number escalate, breach SLAs, and in critical cases, even impact core systems before they're addressed.

Too often, IT teams are left to react after things go wrong.

This reactive approach isn't just inefficient; it's a deeply ingrained workflow that leads to:

- Under-equipped teams during peak periods
- Spikes in escalations during predictable times like the end of the quarter or product launches
- Repeated SLA breaches that surprise even seasoned teams
- The inevitable management question: "Why didn't we see this coming?"

Many IT teams view automated ticket handling as a cost-effective solution. These systems often tweak workflows, assign tickets, and aid resolution by analyzing past incident or resolution data. However, even mature ITSM setups, despite their automation, heavily rely on historical reporting. While useful for understanding past performance, these dashboards and reports can only describe what has happened.

With static dashboards, IT leaders cannot:

- Warn IT technicians about the next critical incident spike.
- Inform service desk managers when their support team will be overwhelmed in the coming days.
- Show how diverse factors like seasonal workloads, staff changes, or infrastructure maintenance increase escalation risks.

But the truth is—they could have. If they had leveraged the power of ML-based forecasting at the right time.

● Embracing ML-driven multivariate forecasting

What IT leaders truly need is more than just historical dashboards. They require a sophisticated solution capable of:

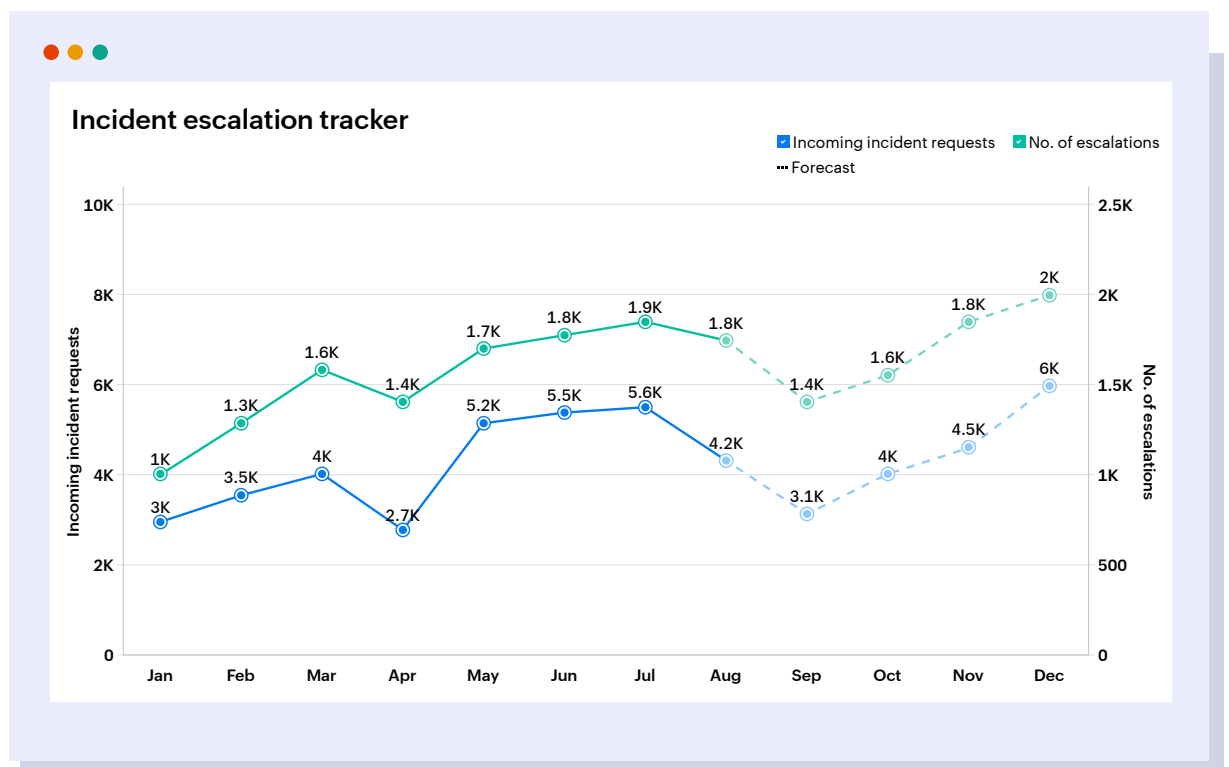
- Identifying hidden layers within IT data, spanning across time, departments, and service categories.
- Recognizing underlying patterns from this complex data.
- Projecting future needs and scenarios based on these patterns.
- Equipping teams to proactively act before issues accumulate, escalate, and lead to SLA violations.

This is where ML-driven multivariate forecasting, a forecasting technique specifically designed for the chaotic and ever-evolving landscape of enterprise IT, comes in. By applying advanced multivariate forecasting to your service desk, asset management, and IT operations, IT teams can transform raw metrics into accurate incident and event forecasts. This fosters a critical shift from static reporting to predictive incident management.

Unlike traditional methods that look at one value in isolation, multivariate forecasting understands how multiple factors interact to influence outcomes. By analyzing these interconnected variables, it enables risk-aware, proactive ITSM.

Consider a global ITSM team that consistently faces recurring escalation surges month after month. Despite regular staffing adjustments, an increasing volume of incident requests leads to more SLA breaches and subsequent escalations.

If the IT manager were to forecast escalations for the coming months using past data alone, there would be a linear increase in predicted escalations as incident volume rises, as shown in the analysis below.

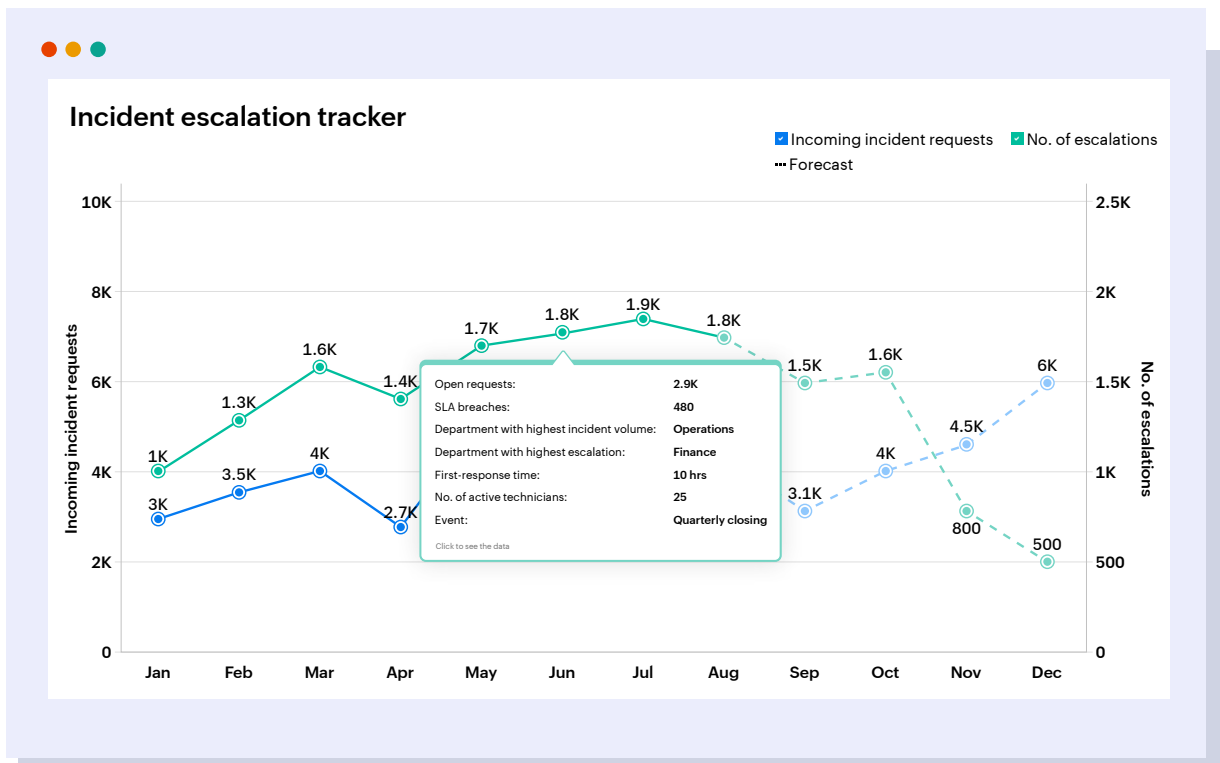


This limited perspective would lead to mitigation strategies—like staffing adjustments, knowledge base article preparation, and alert configuration—being based on poorly informed guesses. The result?

- Over-preparation in some weeks, under-preparation in others.
- Missing early warning signs hidden in combined trends.

However, with ML-powered multivariate forecasting, IT leaders can incorporate numerous other factors that influence escalation volume.

For instance, the same analysis, which predicted a linear relationship between escalations and incident volume, when applied to multivariate forecasting, can reveal that even in months like November and December, which saw a high volume of incidents, the number of escalations remained low. Conversely, for September, despite a lower incident volume, the number of escalations was predicted to be higher than average.



From this analysis, it becomes clear that factors beyond just incident volume significantly contribute to escalations, including:

- **Delay in first response time:** Longer response times can exacerbate minor issues.
- **Incidents logged from specific departments:** Certain departments might have unique challenges or higher criticality.
- **Events like annual closings or national holidays:** These can impact resource availability and task allocation, leading to unexpected surges

Only multivariate forecasting can effectively bring these diverse, dependent factors together to deliver meaningful predictions on escalations. More importantly, it provides IT teams with actionable insights into where to target their mitigation strategies and efforts to reduce monthly escalation volume.

02

Premeditating the impact of change rollouts

Changes are an inevitable, and essential, part of any organization's IT landscape. Whether it's a new product launch, a system update, patch deployments, additions to infrastructure, or even minor policy adjustments, all these are typically managed through a structured change workflow within the service desk.

However, large-scale changes frequently come with a myriad of unanticipated consequences, ranging from minor glitches to significant disruptions:

- Unexpected downtime due to issues with new patches
- Broken workflows resulting from updated configurations
- Application failures or performance degradation after infrastructure migrations
- A surge in change-related ticket volumes, overwhelming the service desk

Since any change scenario deviates from normal expected workflows, IT teams need to be thoroughly prepared. Currently, IT teams often rely on past experiences, gut feelings, or simulations to predict the outcomes or impacts of changes. While these methods can offer some insight, they may not be entirely accurate or comprehensive for an upcoming, specific IT scenario.

The optimal solution is to accurately predict the potential impact of a proposed change on the real-time performance of target services or intended outcomes before actually deploying those changes.

To achieve this foresight, IT teams can first leverage a no-code ML builder to construct a custom change-impact scoring model. This powerful model, trained on an organization's change management data automatically quantifies the projected risk and impact of each proposed change.

Let's consider a highly critical scenario of deploying firmware upgrades on core routers or Wi-Fi network access points. A change, while being critical to network security and stability, carries critical operational risks that could cascade into major network disruptions if not timed and executed properly.

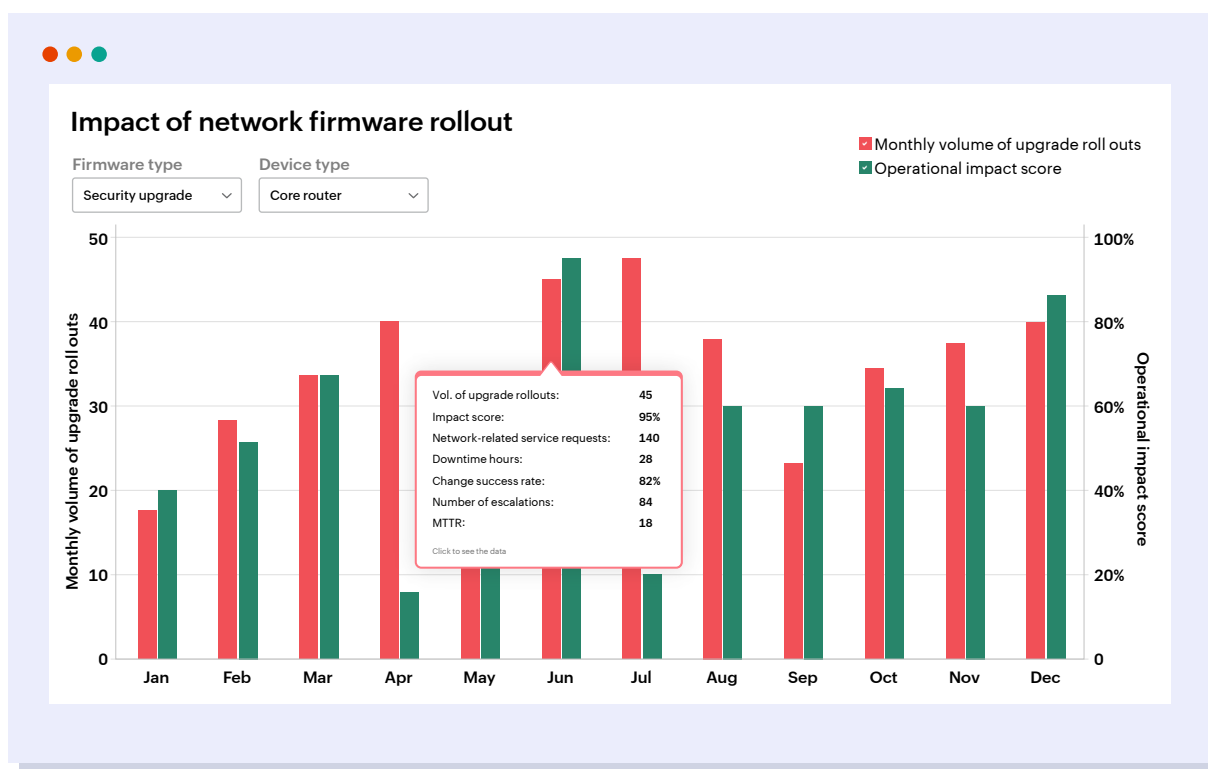
An ML model can correlate firmware changes, such as security upgrades, performance enhancement, or stability fixes, with the historical behavior of similar rollouts, and can correlate the environment in which they were deployed with measurable operational outcomes like:

- **Network-related service requests:** The volume of new incident tickets or requests logged post-deployment.
- **Downtime hours:** Any unscheduled outages or latency issues in affected segments, i.e., router, modems, etc.
- **Change success rate:** The percentage of changes (in this case, firmware) that proceed without incident.
- **Number of escalations:** The count of issues escalated to a higher tier post upgrade.
- **Mean time to resolve (MTTR):** The average time taken to resolve issues arising from the change.

Once deployed, this sophisticated model produces a precise impact score for each change scenario, quantifying the likelihood of disruption or degradation. This score serves as a critical indicator of the potential operational risk of firmware upgrade.

This impact score then becomes the cornerstone for what-if based scenario simulations. These scenario analyses empower IT leaders to prioritize, postpone, and reconfigure firmware upgrade rollouts.

Consider the analysis below, which illustrates how an increasing volume of security upgrades for core routers in any given month influences the calculated impact score across different months. This score, as discussed, is a composite developed by analyzing various critical factors for the success or failure of service desk change requests associated with security firmware deployment.



This analysis provides a crystal-clear picture of how variations in upgrade type, or perhaps a specific combination of upgrade type, time of deployment, frequency, and target network segment, directly impact operational stability and thereby the overall success rate of those deployments.

The visualization vividly indicates that deploying a security firmware upgrade on core routers could likely lead to a critical operational downtime or impact during months like June and December. In stark contrast, a few months like April and July show a significantly lower impact score for the same firmware type on the same network element, strongly suggesting these periods might be more suitable for aggressive rollouts.

Armed with these insights, IT teams can strategically restructure their firmware upgrade policies and adopt a hybrid approach:

- Plan more staggered or phased rollouts of high-impact firmware types.
- Increase the deployment frequency on months with low predicted impact.
- Postpone or sandbox upgrades on device types with historically high failure correlations.
- Plan lower-frequency, more controlled, extremely critical upgrades alone during high-impact months.

Therefore, with change impact models integrated into ML-based what-if scenario analysis, IT teams gain the crucial ability to test and evaluate potential changes before implementation. This fosters a deeper understanding of the change flow, streamlines the overall change management process, significantly reduces the incident risk associated with change deployments, and enables the implementation of timely changes without compromising operational stability and efficiency.

Proactive asset health maintenance with ML-powered foresight

In the intricate world of modern ITOps, where uptime is the currency, waiting to fix things when they break is no longer an option. The cost of unexpected downtime, whether from a failing server or a critical network device, can be astronomical, impacting revenue, reputation, and customer trust.

Unfortunately, many IT teams still rely on static health score thresholds or manual inspection routines to gauge asset health. Maintenance has been a reactive or time-based endeavor, the consequences of which are significant:

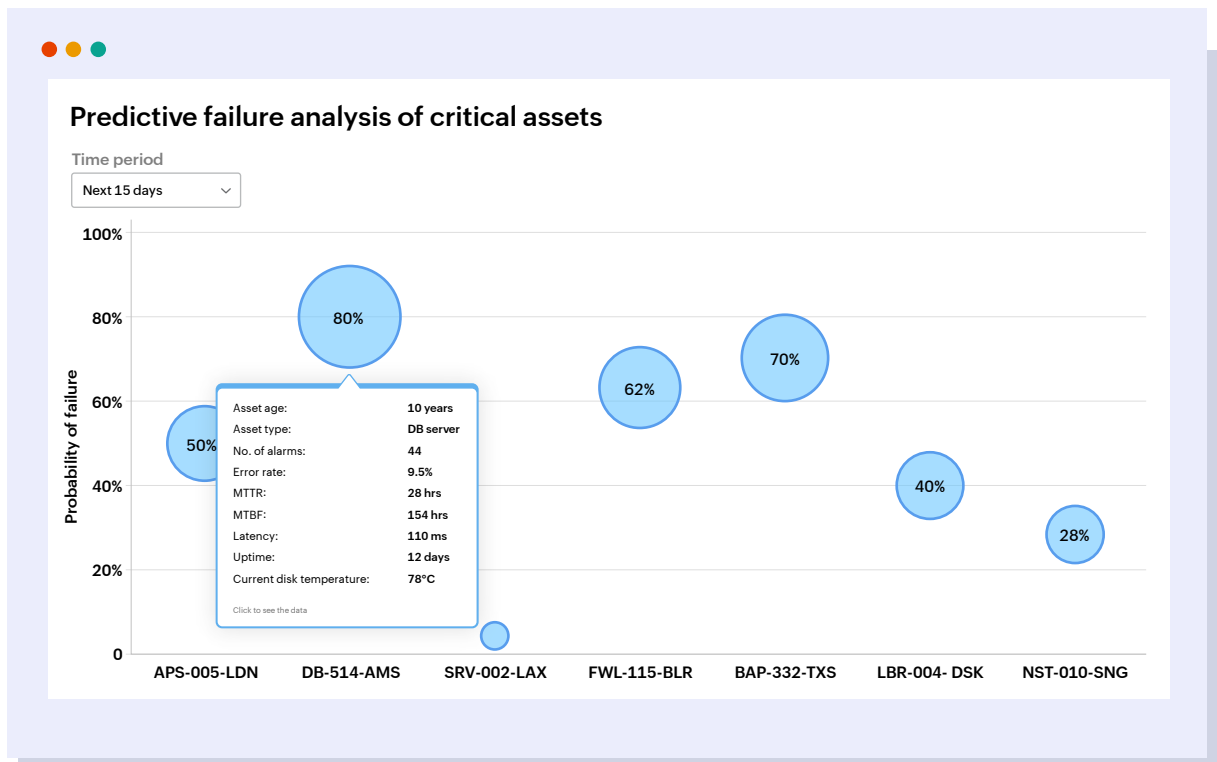
- **Unexpected downtime:** Unforeseen failures halt operations, leading to costly disruptions and lost productivity.
- **Inefficient resource allocation:** Maintenance teams are often deployed reactively, leading to rushed repairs or, conversely, performing unnecessary maintenance.
- **Premature asset replacement:** Assets might be replaced based on arbitrary schedules rather than their actual condition, leading to wasted capital.
- **Missed opportunities for optimization:** Without insight into asset health, it's impossible to optimize performance or extend the lifespan.

The goal is to fundamentally shift from this break-fix mentality to one where potential failures are anticipated and addressed proactively before they ever impact operations. This is the promise of ML-based scoring of asset risk and performance.

Every IT asset—be it an endpoint device, a server, a network switch, a VPN gateway, or any piece of critical infrastructure—leaves behind a continuous trail of data. This digital breadcrumb includes CPU cycles, memory spikes, disk I/O, network latency, mean time between failures (MTBF) logs, temperature readings, and various error codes. Crucially, all these data points hold predictive signals that can be harnessed to anticipate failures long before they disrupt business operations.

ML models are adept at analyzing these complex usage patterns, environmental factors, and historical data to predict failures. This enables truly proactive maintenance, ensuring interventions occur precisely when needed, thereby optimizing asset lifespans and minimizing costly downtime.

With a no-code ML builder, the IT operations teams can now create a tailored ML model to predict asset failures. These models can be trained on historical usage data and performance metrics to identify patterns that typically precede a failure, such as asset age, spikes in CPU and disk activity, MTTR, MTBF temperature logs, and past failure events that follow a predictable life cycle curve.



This comprehensive analysis provides IT teams with a clear, forward-looking view of their entire asset health landscape. The model intelligently combines dozens of metrics (even those not explicitly visualized in a single report) to assign a composite failure risk score to each asset. This score quantifies the probability of an impending failure in the upcoming days, enabling IT teams to proactively schedule maintenance before any catastrophic event, thereby avoiding service disruptions and significantly reducing total maintenance costs.

For example, the visualization clearly shows that Asset ID DB-514-AMS has a high probability of failure, with its predicted failure risk score for the next 15 days hitting 80%. This critical surge is driven by a recent spike in its error rate, an increasing number of alarms, and consistently elevated temperatures. In stark contrast, Asset ID SRV-002-LAX shows a consistently low probability of failure for the next 15 days (below 5%), indicating no immediate need for intervention despite being of similar operational age.

Armed with these insights, IT teams can strategically restructure their maintenance operations. Instead of waiting for an asset to fail or replacing it prematurely, IT teams can schedule a replacement for assets strategically.

By implementing such tailored no-code ML models for predictive maintenance, IT teams drive a transformative shift in how they approach asset management:

- **Reduced unplanned downtime:** Anticipating failures means interventions can be scheduled, preventing costly outages.
- **Optimized maintenance schedules:** Maintenance is performed only when truly necessary, extending asset life and reducing unnecessary labor and costs.
- **Extended asset lifespan:** Proactive care based on actual condition maximizes the return on investment for each asset.
- **Enhanced operational resilience:** Systems remain stable and perform optimally, leading to improved service delivery and user satisfaction.

Predictive maintenance, once a complex endeavor, becomes an accessible, automated capability. This empowers IT and operations teams to focus on higher-level strategic goals, transforming asset management from a reactive burden into a proactive driver of efficiency and reliability.

Balancing distributed infrastructure capacity with tailored predictions

As organizations scale and digital transformation accelerates, the IT infrastructure becomes increasingly complex—and undeniably distributed. Today, applications and workloads are no longer confined to a single location. They're spread across data centers, public and private clouds, and hybrid environments. With this shift comes a new challenge: resource usage patterns are anything but predictable.

In these modern, fast-moving environments, workloads fluctuate in real time, driven by internal operational demands, changes in customer behavior, high-impact campaigns, and even security incidents. While this level of dynamism is a sign of digital maturity, it also introduces certain operational pain points:

- **Underutilized infrastructure:** Expensive resources like storage or compute power often sit idle, leading to wasted money and diminished efficiency as demand shifts unexpectedly.
- **Overutilized sections:** Conversely, some critical resources and systems are stretched beyond their limits, resulting in system slowdowns, degraded performance, or even outright failure as sudden spikes in demand overwhelm static allocations.

A major contributor to these inefficiencies is the still-prevalent traditional capacity distribution approach, which often relies on industry know-how, static rules, broad assumptions, or simply historical trends. While seemingly logical on the surface, this approach frequently leads to periods of significant inefficiency and substantial cost leakages that directly impact IT budgets.

● The need for predictive capacity balancing

Effective capacity distribution and balancing necessitate that IT teams possess clear foresight and proactive knowledge of when, where, and how sections of their infrastructure will experience stress or wastage. There is an inherent, pressing need for a solution or a system that can anticipate infrastructure imbalances before they cause damage.

This is where no-code ML builder solutions, like Analytics Plus' AutoML, provides organizations with a revolutionary, code-free path to creating complex, custom models for capacity balancing and analysis in seconds.

AutoML works by analyzing vast amounts of historical data, including server performance, workload distribution, resource request volume, and actual resource usage. It then automates the creation of tailored ML models that are specifically designed to:

This uniqueness presents common challenges:

- **Dynamically fluctuating workload demands:** These can arise from unexpected traffic surges, unpredictable application behavior, or strategic business events like marketing campaigns and major software updates.
- **Sudden, unexpected utilization patterns:** These emerge without warning, making it impossible to accurately predict the load using traditional, rigid forecasting models.

Therefore, standard tools and generic reports may no longer adequately cater to the unique requirements and more intricate capacity balancing needs of modern enterprises, driven by dynamic and often unforeseen shifts in usage patterns.

Tailored models for dynamic infrastructures

At this juncture, organizations need to employ custom ML analytics models that can truly understand their unique conditions, evolving usage patterns, dynamic requirements, infrastructure intricacies, market trends, and distinct business objectives to:

- Proactively identify hidden trends, events, or underlying factors that can cause a significant impact on resource usage patterns.
- Forecast when resources will experience heavy loads and automatically trigger corrective actions, such as load balancing, reallocation, or dynamic scaling.
- Optimize the use of resources, ensuring that critical infrastructure doesn't get overwhelmed while underutilized resources aren't left idle, wasting valuable capital, even amidst unpredictable demand fluctuations.

However, the additional cost of employing expert data analysts to build custom analytical models, coupled with the enormous time and resource consumption involved, can deter IT leaders from pursuing this optimal approach. Relying on the built-in, generic capabilities of existing tools, which can deliver inaccurate results, risks disrupting your entire operation due to unaddressed imbalances.

This is where no-code ML builder solutions, like Analytics Plus' AutoML, provides organizations with a revolutionary, code-free path to creating complex, custom models for capacity balancing and analysis in seconds.

This no-code ML model analyzes vast amounts of historical data, including server performance, workload distribution, resource request volume, and actual resource usage. It then automates the creation of tailored ML models that are specifically designed to:

- Spot potential infrastructure imbalances before they occur.
- Predict how dynamic patterns in usage and unanticipated events can result in change in user behavior leading to over-utilization or under-utilization of infrastructure.

Imagine an enterprise SaaS company that frequently experiences uneven workload distribution—often due to unforeseen spikes in user activity during feature rollouts or marketing campaigns—that traditional models relying on static rules simply cannot account for.

By using a no-code ML builder, the NOC team can create a custom model to analyze historical telemetry and detect patterns across traffic volumes, security events, user behavior, and resource health.



This analysis reveals how a tailored capacity balancing model can utilize 12 months of historical telemetry data to predict the dynamic load risk score for critical infrastructure resources or services.

The prediction incorporates a rich array of factors: peak usage scenarios, security breaches, other unexpected events, and historical incidents leading to unpredictable spikes or troughs in infrastructure health data (CPU, memory, disk utilization, latency).

The analysis clearly shows how the ML model's predictions accurately track the evolving load risk, even during unforeseen surges, unlike traditional forecasts that might completely miss these dynamic shifts. Armed with this foresight, the team can create automation policies to dynamically re-route traffic to underutilized nodes or scale resources up or down, effectively mitigating risks from unforeseen demand before it impacts users.

The automated traffic redirection policies, implemented based on the tailored ML predictions, can yield significant benefits:

- Reduction in infrastructure incidents: Proactive balancing prevents bottlenecks and performance degradation.
- Optimization in cloud spending: Resources are used efficiently, avoiding unnecessary scaling.
- Improved SLA compliance for infrastructure-related requests due to preemptive actions.

By implementing such tailored no-code ML models for capacity balancing, IT teams drive a transformative shift in how they approach infrastructure management.

- **Proactive resource allocation:** They anticipate load spikes, preventing outages before they happen.
- **Cost savings:** By intelligently utilizing underutilized resources, they optimize cloud spend and dramatically reduce wastage.
- **Seamless scalability:** The system automatically triggers necessary scaling without manual intervention, ensuring applications are always adequately supported.

Capacity balancing has long been a reactive task. But with predictive models at their fingertips, IT teams can finally turn it into a proactive, intelligent process. This empowers NOC teams to shift their focus towards higher-level strategic goals rather than constantly firefighting infrastructure issues, helping build a more resilient, cost-effective, and agile IT environment for the future.

Conclusion

The journey from reactive problem-solving to proactive strategic IT management is not just an aspiration; it's a necessity in today's complex digital world. ML-powered analytics offers the intelligence needed to transform raw data into actionable foresight, turning potential crises into managed outcomes and smarter decisions. By embracing the advanced analytical capabilities discussed in this e-book, IT teams can move beyond the limitations of traditional approaches, ensuring operational stability, optimizing resource utilization, and fortifying security postures.

About

ManageEngine Analytics Plus is an IT analytics and decision intelligence solution designed to provide organizations with a unified view of their IT operations, correlate interdependencies and derive meaningful insights. It breaks down data silos by consolidating both on-premises and cloud infrastructure KPIs. Analytics Plus measures the efficiency of network operations, tracks the responsiveness and availability of business applications, evaluates technician performance, assesses the progress of processes and flags security anomalies. This comprehensive analysis is achieved by connecting to all IT software that forms the backbone of an IT infrastructure. These consolidated insights enable organizations to make data-driven decisions that enhance operational efficiency and drive business success.

For more information about Analytics Plus,

visit: www.manageengine.com/analytics-plus/



Reference

1. <https://www.leanix.net/en/ty/sap-leanix-ai-survey-results-2024?submissionGuid=ed2f7c1d-d42c-4871-b590-c087844618e7>



© ManageEngine, a division of Zoho Corporation