ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR OPTIMIZING DEVOPS, IT OPERATIONS, AND BUSINESS

EMA Top 3 Report and Decision Guide for Enterprise

ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™)
VENDOR RECOMMENDATION REPORT
PREPARED FOR DENSIFY
Written by Torsten Volk

Q4 2018

THE DISRUPTIVE IMPACT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING ON DEVOPS, IT OPERATIONS, AND BUSINESS

Enterprise Management Associates (EMA) research shows that leveraging artificial intelligence and machine learning (AI/ML) for DevOps, IT operations management, and business management is the top priority for enterprises in 2018 and beyond. AI/ML brings data-driven intelligence to DevOps, IT operations, and the enterprise to optimize processes, recognize relevant trends, proactively prevent issues, rapidly detect and resolve problems, and enable human staff to make optimal and fact-driven decisions.

Guidance for AI/ML Product Decisions in DevOps, IT Ops, and Business

This EMA "Top 3 Decision Guide for Artificial Intelligence and Machine Learning in DevOps, IT Operations, and Business" provides guidance for enterprises seeking to optimally leverage today's AI/ML capabilities, depending on their individual situation and priorities.

EMA TOP 3 REPORT HIGHLIGHTS

EMA Top 3 products for optimally leveraging AI in business, DevOps, and IT

The **8 key AI/ML bottlenecks** in 2018

9 best practices for implementing AI capabilities

12 questions to ask vendors pitching AI-based solutions

Why AI/ML is Disrupting Every Industry

The disruptive character of AI originates from the core expectation of AI/ML technologies enabling enterprises to make optimal data-driven decisions that are aligned with their corporate strategy, immediate priorities, and compliance obligations.

Enhancing human capabilities through AI: Complexity is the most critical challenge for employees in DevOps, IT operations, and business. Decision makers are faced with a rapidly growing number of data sources, technology options, and competitive requirements. Optimal decision-making requires the ability to quickly identify how all of the different data sources can be put together into one big picture that shows the outcomes of alternative courses of action. Artificial intelligence can act as this "complexity reducer" by identifying potentially relevant trends and anomalies in these vast bodies of data and by revising its output based on human feedback and past project outcomes.

AI-driven automation: AI-driven automation can eliminate simple repetitive tasks to enable staff to focus on activities where human intuition, creativity, skills, and situational awareness are critical. AI-driven solutions that best include humans into their decision and automation workflows have been significantly more successful in the market compared to solutions focusing on a black box approach to AI-driven automation.

WHAT IS AN EMA TOP 3 PRODUCT?

The EMA Top 3 Award is presented to products that convincingly resolve DevOps, hybrid IT, and business challenges through the innovative use of artificial intelligence or machine learning. EMA identified these products through a sequence of briefings, demonstrations, customer interviews, technical discussions, and often the use case-driven deployment of the product by EMA staff and partners.

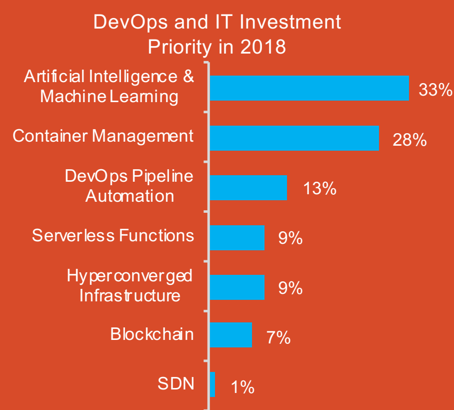
Please note that the EMA Top 3 Award aims to inspire the planning and selection process, but it is not a feature-by-feature product review.

AI BOOST FOR DIGITAL ATTACKERS

EMA's concept of the digital attacker aims to optimize an enterprise's ability to deliver customer value faster, cheaper, and at better quality compared to the competition. Therefore, enterprises need to minimize operational costs to be able to focus most of their resources, skills, and capabilities on increasing customer value, instead of spending often 50 percent of their budget on "keeping the lights on." AI-driven DevOps, IT operations, and business systems can unclog this bottleneck by showing human decision makers how to optimize and automate their current infrastructure without increasing operational risk, how to accelerate the DevOps pipeline while increasing release quality and decreasing cost, and how to provide intelligent solutions for business users to maximize their time spent on essential activities that require human judgement, creativity, general knowledge, and empathy.

DEVOPS AIMS TO OPTIMIZE COST, SPEED, AND QUALITY

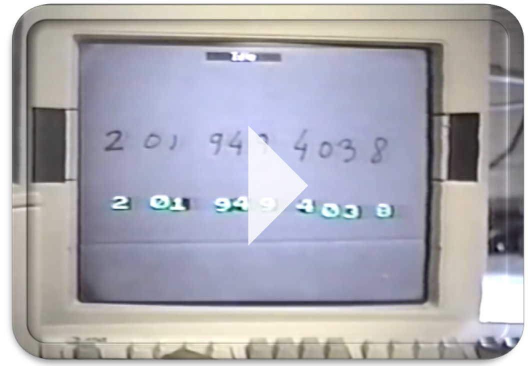
AI/ML IS THE #1 ENTERPRISE INVESTMENT PRIORITY IN 2018



The Four Core Capabilities of AI

AI can solve the unsolvable in business, DevOps, and IT operations due to four core capabilities:

1. Complete tasks without explicitly coded instructions
2. Receive reliable results for cases that were not explicitly included in the training set of examples
3. Rapidly identify relevant information in vast bodies of data
4. Increase prediction quality based on continuous feedback



[Click to view] the 1993 demo of LeCun's first CNN

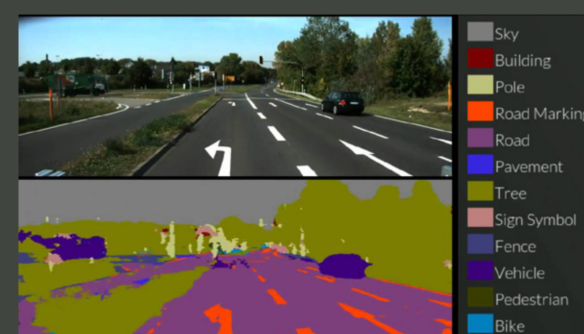
Note that none of these elements is new or revolutionary in 2018. Yann LeCun showed in 1993 how his AI software, the first so-called convolutional neural network (CNN), was able to reliably recognize handwritten numbers without having received any rules or instructions, but was fed with tens of thousands of examples for each number instead. LeCun's CNN even worked under difficult conditions, such as sloppy handwriting, dirt on the paper, or a faulty pen. LeCun's CNN still is the foundation for many modern AI/ML products, such as the self-driving car, language recognition, or calculating mortgage risk.

Complexity Reduction as the Key Value Proposition of AI

AI/ML models can solve tremendously complex challenges without a single line of code, simply by ingesting example data. All of the EMA Top 3 vendors in DevOps, IT Ops, and business are leveraging AI/ML exactly for this purpose, providing human operators with predictions and categorizations derived from large and often seemingly unrelated sources of structured and unstructured data.

WORLDVIEW OF A SELF-DRIVING CAR

The deep neural network for visual recognition in the self-driving car recognizes relevant aspects of the world based on training data obtained from the simulator, from photos, and from driving with a human supervisor.



REALISTIC EXPECTATIONS ARE CRITICAL

There is a term in AI research called the “AI Winter” that describes a period of near standstill in terms of AI research. The AI Winter lasted from the late 1980s until the early 2010s and was triggered by a large degree of hype and exaggeration around the AI topic, with little corresponding research progress to show.

Today, companies are running the risk of slipping into another period of exaggerated expectations due to AI companies fueling imaginations by painting images of systems with truly human-like intelligence. The term “AI” implies that this technology enables machines to purposefully act in the same way a human would. This report demonstrates that humans have not come any closer to human-like machines since LeCun’s breakthrough in 1993.

This EMA Top 3 report is all about spreading excitement around the real capabilities that AI/ML can deliver today and in the immediate future. These capabilities are tremendously powerful and can change business models and industries. To prevent another AI Winter, vendors and enterprises must develop a common understanding of expectations around AI/ML technologies. As the following three examples show, AI/ML enables users to solve problems that formerly were impossible to solve.

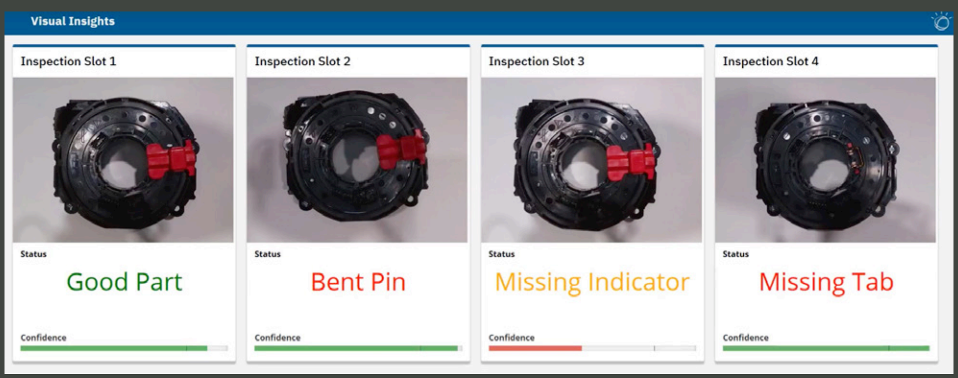
The Key to Understanding the AI Discussion

To successfully leverage AI/ML for business, enterprises must let go of the idea of developing Data, the android from “Star Trek: Next Generation,” for optimizing data centers or automating software testing. However, if they manage to curb exaggerated expectations, AI/ML already has the capabilities required to disrupt any industry and dramatically increase human productivity.

The basics of AI/ML today are simple and easy to understand. Instead of coding IF-statements, creators simply provide an algorithm with a large number of training documents that show the connection between features (independent variables) and results (dependent variables) in many different ways. They need massive compute capabilities to enable the algorithm to minimize prediction errors.

Once the model is fully trained, it can make predictions based on new input that is similar in character to the training material and users can expect predications at a comparable level of reliability. This level of reliability can be higher than that of a human worker, but keep in mind that the AI/ML model does not have any human-like capabilities to “fill in gaps” when dealing with input variables that are different from the training set. In this case, the model might not produce any results, or it may produce dramatically wrong results without having the capability to notice them as such.

The following examples show tremendous business impact by leveraging AI/ML in a realistic manner, instead of aspiring to human-like decision making.



To prevent another AI Winter, we need to emphasize the tremendous business impact of what AI/ML can do today. This image shows IBM Visual Insights correctly checking an industrial part for quality issues, without ever having seen those exact defects before. This shows that today’s top-down approach of creating narrow inference models through providing large numbers of pre-labeled examples can exceed human performance in many cases.

EXAMPLES

Example 1: Continuous Testing for DevOps

“The closer we are getting to a state of continuously releasing new software features, sometimes multiple times per day, the harder it becomes to consistently ensure an optimal user experience on all popular devices and in a half-dozen browsers. Disconnected release cycles due to microservices and containers make testing even more of a moving target to us, with the potential of swallowing a massive amount of work hours.

We have recently been looking at some of the EMA Top 3 products for AI-driven testing and found significant potential for utilizing the same number of test engineers to achieve significantly higher test coverage. Interestingly, these productivity improvements are seemingly achieved by enhancing traditional code with a number of AI/ML models to solve seemingly simple problems in an automated manner. The human test engineer can then deal with a much smaller pile of problem cases and alerts.”– CTO, Global Insurance Company

Conclusion: It is vital to understand that AI/ML cannot autonomously take on the task of continuously testing software. AI/ML models do not actually understand concepts of usability or test coverage, but they are able to solve extremely time-consuming micro-problems, such as identifying whether a text box was moved too far to the right on the screen or if a certain test workflow is omitting a potentially relevant code component. In short, these AI tools deliver significant complexity reduction for humans, enabling the same number of staff to provide much increased test coverage.

Example 2: AI in Quality Control

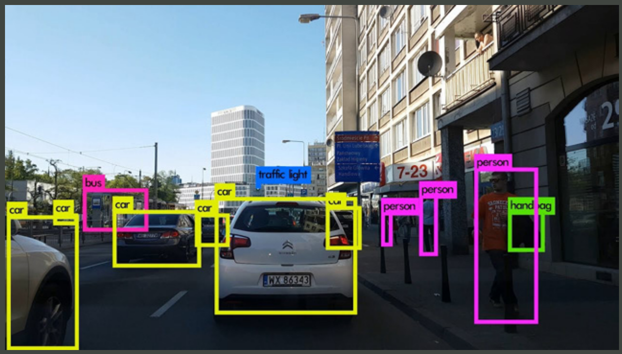
“Our robotic quality control machines rely on real-time visual data from high resolution cameras to recognize certain types of defects on a component. While it is great that we did not have to hand-code the required rules for how to make this happen, let’s not forget that we needed over 10,000 example images for each common type of defect for each individual component. This adds up to many terabytes of storage consumption. Once the images were inserted, the machine trained itself and achieved an error rate below what we would expect from human quality control engineers.” – Sr. Director for Quality Control Technologies, Car Manufacturer

Conclusion: Since the machine is merely comparing pixel patterns to its given training set to calculate defect probability, a human operator may have a higher error rate when looking for exactly the same issues the AI/ML model was trained for. However, if the coder turns the light up or down or changes the color of the inspection tray, or changes the surface structure or color of the object itself, the human QA engineer might still produce good results while the AI/ML model is not able to leverage human-like common sense or intuition to work under different conditions. Here, it is critical to keep in mind that users typically do not know upfront whether or not a training set sufficiently covers a certain corner case.

Example 3: AI for Root Cause Analysis

“Don’t think of AI as magic, but think of it as a ‘complexity reducer’ that raises early alerts when indicators for potential future issues arise and at the same time consolidates immediately relevant events and alerts so we only need to involve the teams that are actually affected or responsible.” – IT Director, mid-market SaaS vendor

Conclusion: While no one can expect AI/ML models to always come to conclusive results, AI/ML is capable of enhancing human reality by highlighting trends and anomalies and by categorizing or clustering a myriad of seemingly unrelated events into groups of related and important events that the human eye should judge.



Once humans have labeled the world for AI/ML, AI/ML can recognize objects, often more reliably than humans.

THE 8 CORE LIMITATIONS OF AI/ML

Understanding the core challenges and bottlenecks of AI/ML is critical to optimally evaluate and ultimately implement the EMA Top 3 products presented in the following chapters.

1. AI Does Not Think

AI/ML can learn how to play Chess, Tetris, GO, or even Jeopardy at a much higher level than any human could, simply by observing or playing a very large number of games. In the case of Google Deep Mind's Alpha Go, the software had to play as many GO games as a human could play in 3,000 years of uninterrupted play, without sleeping, eating, or drinking. To learn even basic steering, the AI for self-driving cars needs thousands of hours of training in a simulator. Thanks to modern parallel GPU architectures, this can often translate into only a few days or even hours of actual time requirements.

3 Key Lessons

1. AI/ML plays games by optimizing a probabilistic model, while humans use their experience, creativity, and intuition. Therefore, humans can rule out many potential options as "non-viable" and therefore do not need as many training games as AI/ML.

2. AI/ML can only learn to play one single game. This means that while advanced deep learning frameworks can be trained to play chess, GO, or Tetris, a single instance always can only learn one game. Critical takeaway: AI/ML is unable to learn concepts and principles for application to new problems.

3. It is crucial to keep in mind that operating Kubernetes, vSphere, AWS, or Azure is not as simple as steering a car, because there are many more parameters involved. Just as importantly, there are currently no largescale data center simulators available that would enable the AI to learn through exploration and observation. Without this training data and without the ability to apply concepts and principles to new situations, it is currently unfeasible to train AI/ML to become a VMware administrator.

2. Data Availability

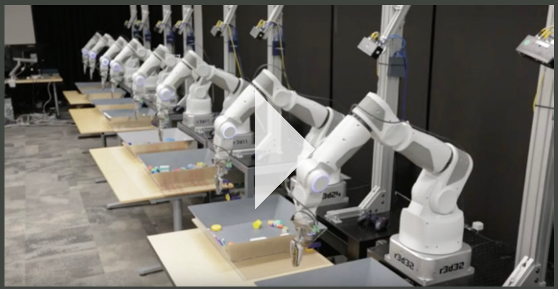
AI/ML requires large quantities of clean training data even to accomplish a narrow set of tasks. Creating these training sets often requires significant human effort and skill, as well as the corporate ability to cost-efficiently handle vast amounts of data in a compliant manner. EMA research data shows that most enterprises today are still struggling with data silos stemming from fragmented systems and a general lack of skill when it comes to consolidating these silos.

SIDE NOTE:
NEURAL NETWORKS VERSUS THE HUMAN BRAIN

Before Yann LeCun created the first convolutional neural network (CNN), handwriting recognition had to rely on rules and instructions describing the characteristics of handwritten letters. While these human-generated rules and instructions were also based on thousands of samples of letters and numbers written by different human authors, the results were unusable due to the inability of anticipating and coding each and every wrinkle thrown at the software by human handwriting. There are simply too many permutations to make manual coding feasible.

Key to understanding today's excitement around CNNs is the fact that instead of using dozens of human researchers and developers to create a still-unreliable software program for handwriting recognition, LeCun's solution required no human involvement in its rule creation and achieved a much higher reliability simply by processing tens of thousands of examples and continuously readjusting its own model until no significant further improvements were expected.

LeCun's model accounted for any anticipated and unanticipated differences in human handwriting, including sloppy style caused by a bad pen or by the human author writing on a soft or uneven surface. The model also accounted for different letter sizes, rotation, squeezing, thickness of the pen, color, dirt on the paper, missing or crossed out letter segments, and even overlapping characters. Today, we can see the same robustness and flexibility in many other use cases for CNNs, such as the Google Image Search, Amazon's Alexa recognizing a voice, or IBM Watson playing Jeopardy.

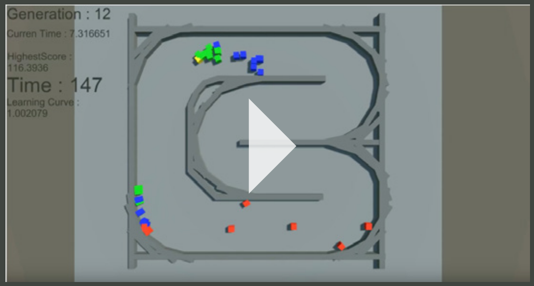


Google's "Arm Farm" took 2.8 million images of how to grab small objects, and executed 800k tries, and still could not grab the desired object, but mostly a random one. (view on YouTube)

3. Lack of Human Intuition and Curiosity

For humans, it would not be feasible to spend thousands of hours learning Go or driving a car in a simulator. Instead, humans leverage their ability to intuitively eliminate the vast majority of non-viable choices and focus on sequentially and gradually trying out the approaches and behaviors that seem to make the most sense to them. This human intuition is based on the ability of a human brain to make cross-connections between a myriad of different observations. This means that when a human being sits behind the wheel of the car for the first time, he or she will not need to drive off of the road a few thousand times until they have learned to stay on the road by using the wheel, accelerator, brakes, and clutch in conjunction. In short, based on their previous experience, humans can dramatically reduce the number of options they have to try out before they can successfully drive a car, while the AI will try out every possible permutation of using wheel, accelerator, brakes, and clutch, including all the ones that begin by fully pressing down on the accelerator while also fully standing on the brakes. This means that until humans figure out how to provide the AI with a certain degree of intuition, its learning will be entirely unguided and inefficient. Of course, this does not matter as long as there is a sufficient number of processor cores available to dramatically condense this learning process by parallelizing it.

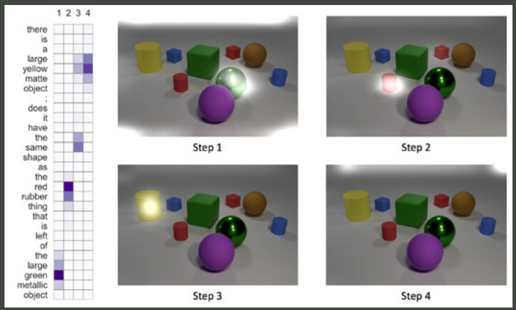
The self-driving cars start off by continuously hitting walls. The next generation of each type of car (different colored dots) makes revisions to their learning model so that each generation comes closer to mastering the entire track without accident. (view video on YouTube)



4. Lack of Ability to Explain

AI models today are opaque and do not provide human operators with any insights on how they came to their conclusions. These models merely provide a probability value for their prediction to be correct. However, when experimenting with neural networks, researchers have found significant evidence that leads them to question the reliability of these probability scores. Stanford professor Dr. Christopher Manning has made significant recent progress in terms of visualizing the neural network decision process; however, as the computer-generated image demonstrates, explanations are only feasible in low-complexity situations.

EXPLAINING DECISION-MAKING IN NEURAL NETS



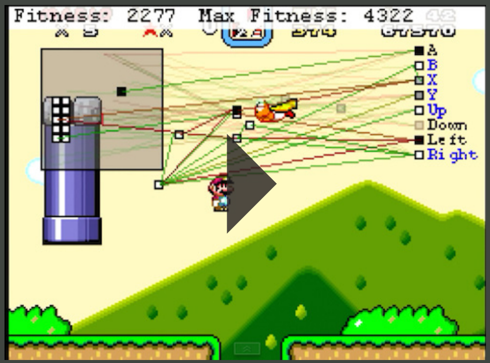
In this example, Manning's model dissected the question into four individual concepts, working its way to finding the "large green metallic object."

5. Feedback is not Always Available

Training a self-driving car can only partially happen in a simulator, since there are simply too many surprises and unique situations in real-life traffic. While it is fun to watch the AI driver's evolution in the simulator, it becomes a lot more tricky to continue the training first on closed tracks and later on public highways and city centers. It is key to understand that reinforcement learning using convolutional neural networks relies on random trial and error, controlled by a reward function that punishes negative behavior and rewards good behavior. For example, when the AI driver stops for pedestrians at a crosswalk, the AI model rewards this behavior. However, the human driver inside of the experimental AI car needs to be constantly alert to hit the breaks for negative feedback, as the alternative could result in human injury. Therefore, in a simulator, the AI driver learned to not drive off a cliff by actually driving off a cliff. Not once or twice, but often thousands of times until the ongoing feedback loops provided the neural network with the data required to operate the gas pedal, breaks, and steering wheel in a manner that results in safe driving. But then, what happens if it rains, snows, a deer is on the road, a tire looses air pressure, or wind gusts hit the car? How does the optimal AI response depend on whether it is driving a pickup truck, a Porsche 911, or a rusty family sedan? The moment the AI driver needs to leave the simulator is when things become difficult, because the AI will continue to regard the outside world as a collection of pixel patters that it classifies as safe or dangerous.

Moving away from the driverless car example and into the data center, while AI decisions are not impacting life and death in this setting, wrong decisions or missed warnings can become very costly for an enterprise. Creators must simulate everything that is going on in the data center in terms of server, storage, network, virtualization, and application performance, health, and configuration parameters, which is unfeasible today.

NEURAL NETS LEARN BEST THROUGH SIMULATIONS

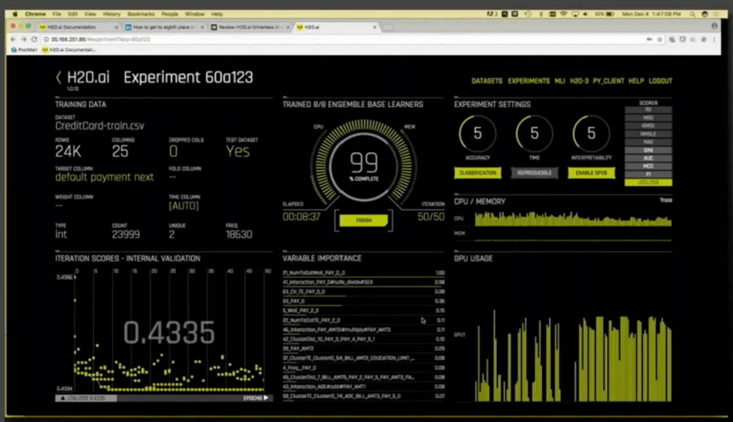


Neural networks are great at playing Mario because they receive continuous real-time feedback of almost every button push. The Fitness score (2277) shows the current optimization status of the model, while the Max Fitness (4322) is the optimal score at this point in the game. The AI/ML will continue to play this game and revise its internal weight structure until it comes close to the Max Fitness score. (watch video on YouTube)

THE 8 CORE LIMITATIONS OF AI/ML

6. Separation of Data Science and Subject Matter Expertise

Today's reliance on the availability of data scientists prevents the vast majority of potential AI-driven projects in business, DevOps, and data centers from even being discussed. Enterprises live off of staff initiative in terms of experimenting with new features and capabilities without creating much cost exposure to the company. For AI to become the anticipated disruptor of entire industries and the driving force behind new products and business models, the technology must be available for everyone, including business staff and software engineers. This report will present several EMA Top 3 vendors that made strides in the areas of enabling non-data scientists to explore large, unstructured data sources and even in enabling staff to experiment with AI models without having to worry about selecting the appropriate algorithm or hyper-parameters.



H2O.ai offers a "self-driving" AI engine that only asks the user for a set of training data and for the dependent variables. The software then automatically selects the best possible algorithm and calculates the appropriate features.

7. Complexity through Recombination of AI/ML Models

Due to the narrow character of modern AI models, researchers and enterprises have adopted the practice of combining multiple models to cover more ground. This quickly leads to a large number of interdependent models, all with prediction capabilities as narrow as differentiating between an image and written text, evaluating the health of the I/O pattern of a database server, or recognizing the optimal point to turn on the windshield wiper on a car. Piecing all of these individual AI models together into one larger system is a difficult task with many moving parts, that requires the human brainpower of experienced data scientists.

8. Infrastructure Cost

Today, AI typically relies on GPUs for model training due to their ability for massive parallel processing. Training AI models relies on the basic principle of iteratively optimizing the weight of each neuron on each network layer to optimally fit the model to its training data. The more layers, the more neurons on each layer, and the more training data, the longer it takes to train the model. To alleviate the cost concerns of this brute force approach to AI model training, a new generation of startups has emerged that is focused on creating CPUs that are optimized for training neural networks. Instead of using the parallel processing capabilities of legacy GPU architecture, these new chips enable a much more efficient approach to AI training, translating into significantly lower hardware cost and power consumption. Secondly, some AI offerings now include automatic hyper-parameterization and more sophisticated reward algorithms to reduce overall processing requirements by improving the AI's "intuition" in terms of which iterations are the most promising to explore when trying to maximize mid- and long-term payout.

NEXT FRONTIER: CONCEPT LEARNING

Today, AI/ML research focuses on providing AI/ML models with a general understanding of how the world works. This means these models are more robust against task and environmental changes. To achieve this general understanding of our world, AI/ML models need to retain valuable principles and lessons in their long-term memory. This principle is very similar to humans applying general knowledge they learned throughout their lives to new situations that are comparable to, but not the same as, what they are used to. For example, when caught in the dark in an unknown room, humans make a sequence of assumptions that helps them assess the situation: 1) there could be a ceiling light, 2) there could be a lamp on the desk, 3) light switches are typically located next to a door, 4) light switches can be toggled, pushed in, or moved up, 5) if there is no light switch, can I find a candle or a window? 6) can I use the flashlight of my cellphone? Depending on a human's personal experience, he or she will be able to solve this problem without ever having been in a similar situation before, just by thinking about how the world works.

Yann LeCun, scientist at Facebook and professor at NYU, and Christoper Manning, professor at Stanford, are spearheading research initiatives aiming to allow AI/ML to develop and benefit from a memory.

Manning's Architecture-Focused Approach versus LeCun's Focus on Reward Systems

Manning believes that in order for AI/ML to retain and reuse important principles of how the world works, the model requires a human-created architecture that is able to retain a core set of basic truths. Manning's current work is focused on creating this architecture by adding algorithms to his models that can extract relevant experiences and write them to a short-term memory. The more these experiences prove reusable, the more another algorithm stores them in long-term memory.

Yann LeCun disagrees with Manning's call for a "core set of basic truth" that is reflected in a normative learning architecture. LeCun believes that we simply need to appropriately configure and scale neural networks so they "learn the optimal view of reality" without any human bias introduced through structure. He sees the greatest challenge of AI in defining a reward system for the model to independently learn.

Predictive Learning as the Holy Grail

Both LeCun and Manning aim to develop learning models that enable them to abstract specifically valuable lessons and store them away in memory so they can be used to fill gaps in their training sets. For example, after learning the negative impact of the failure of the company's e-commerce website on the corporation's bottom line, the AI/ML model has learned multiple lessons: 1) different applications have different business impact and 2) protecting the e-commerce platform is our highest priority. Keeping in mind these two lessons and recombining them with future learnings would eventually bring AI/ML to another level of independence.

Yann LeCun (left) and Christopher Manning (right) debating the importance of structure for concept learning (watch video on YouTube).

TWELVE WAYS TO EVALUATE AI CAPABILITIES

When selecting software and hardware products for business, IT, and DevOps, customers are faced with significant marketing lingua highlighting the advantages of AI in the respective product. EMA research found that vendors are often unable to clearly demonstrate the business value of their AI, and worse, that many vendors are relabeling simple event correlation and decision heuristics as AI. Here is a checklist that will help identify substance over marketing claims during a sales pitch:

SHOWING AI VALUE

1. Can you quantify the value of the AI capabilities of the product? Can I easily measure this value in my own environment?
2. Does the AI-enhanced product solve standard tasks faster, cheaper, and more reliably?
3. Does the AI open up a whole new set of opportunities?
4. Can the AI-enhanced tool replace my existing tool, or is it an additional solution?

EVALUATING RISK

1. Will the product be able to get the data it requires based on my current data architecture?
2. Does the product support my specific applications and use cases?
3. What are the recommended project success milestones and metrics?
4. What are the requirements for my current staff to optimally use the system?

SCOPE OF AI

1. Which AI/ML algorithms are in use?
2. How do these algorithms receive ongoing feedback? How are they updated?
3. Which product capabilities does AI impact? Provide a use case.
4. How transparent is the AI decision process? How do AI and end users interact?

AI FOR OPTIMIZING DEVOPS COST AND RISK

EMA RESEARCH FACTS

"Overrunning and opaque public cloud cost" is still a top IT operations pain point in 2018.

Overprovisioning is used as "insurance" against performance degradation or downtime, but often fails to deliver.

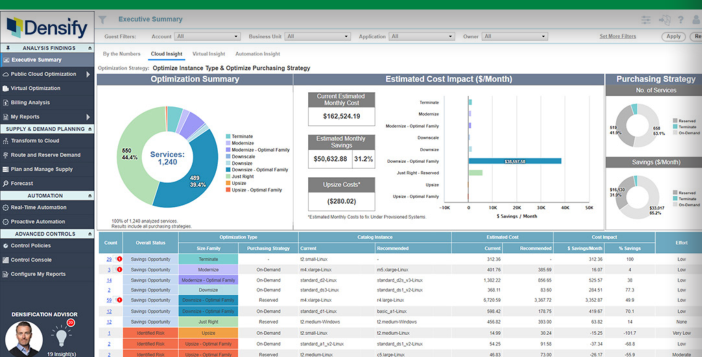
Cloud and data center infrastructure is typically overprovisioned by 50%.

Enterprises typically miss out on substantial discounts by running instances without reservations.

AI FOR DEVOPS

Over the previous 12 months, AWS, Azure, Google Cloud, and IBM Cloud jointly have added over 200 additional cloud services to their already-extensive portfolios. At the same time, the rise of Kubernetes, functions as a service, and hyperconverged infrastructure added even more deployment options to the mix. The EMA Top 3 vendors in this category leverage AI/ML to help enterprises deploy and manage their application infrastructure in an SLA-driven and cost-effective manner.

DENSIFY



Application-Centric Hybrid Infrastructure Optimization

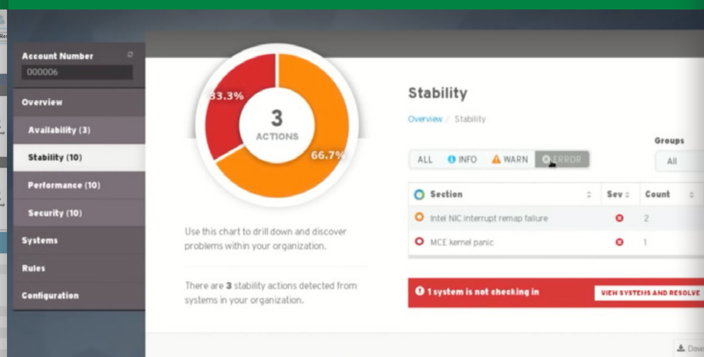
WHY EMA TOP 3

- Excellent vision of the self-aware application.
- AI/ML for proactive and compliant application management.
- Cloud management plane for continuous monitoring and optimization.

QUICK TAKE: Densify leverages machine learning to continuously match application workloads with the optimal cloud or data center deployment option based on policies and application runtime requirements.

PRODUCT WEBSITE

RED HAT INSIGHTS



SaaS-Based Infrastructure Optimization and Automation

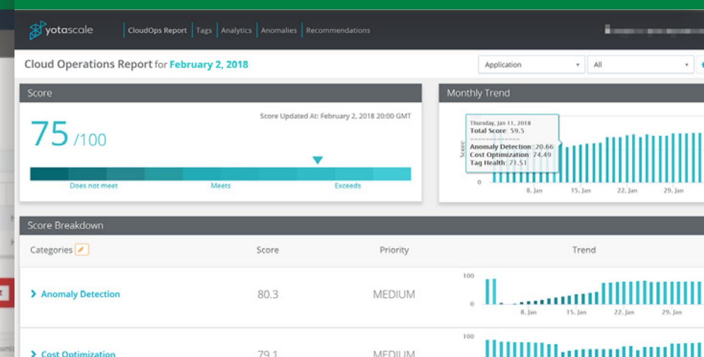
WHY EMA TOP 3

- Proactive warning based on probability and severity of performance, security, stability, and availability issues.
- Tight integration with Ansible for auto-remediation, auto-scaling, or cloudbursting.
- Integration with Red Hat support data for faster and actionable recommendations.

QUICK TAKE: Red Hat Insights leverages AI/ML to proactively optimize and troubleshoot the entire Red Hat portfolio based on recommended actions provided by Red Hat and based on all of Red Hat's support data.

PRODUCT WEBSITE

YOTASCALE



Continuous Cost and Risk Optimization for AWS

WHY EMA TOP 3

- Continuous optimization of deep learning model based on daily AWS usage.
- Explains symptoms, root cause, and recommended remediation action.
- Makes resource usage by app, business unit, or region transparent.

QUICK TAKE: YotaScale leverages big data analytics and deep learning to continuously analyze AWS environments for inefficiencies, capacity, and performance problems.

PRODUCT WEBSITE

KEY PRINCIPLES FOR REAL-LIFE AI/ML PROJECTS AND PRODUCT SELECTION

1 SELECT REALISTIC SUCCESS METRICS

AI/ML projects must play by the same rules as any other corporate initiative.

72% of enterprises are struggling with unsanctioned and often ungoverned Kubernetes environments.

2 START SMALL

Set small initial milestones that are easy to meet to enhance trust and confidence in AI/ML.

3 DATA IS KING

Create the initial business case with the minimum viable amount of data. Then, plan out the data strategy for the entire project.

4 INTEGRATE, INTEGRATE, INTEGRATE

Focus on the integration with all relevant systems to maximize the performance and impact of AI.

5 SET UP FEEDBACK LOOPS

Plan out feedback loops to continuously tune your model without burdening staff.

6 DATA QUALITY MATTERS

Create a task force to certify data quality and vary the suitability of the available data with the selected AI/ML algorithm.

7 AI/ML – HUMAN WORKFLOWS ARE CRUCIAL

Remember the purpose of AI/ML as tool to enhance human reality by providing decision-relevant facts and scenarios.

8 AI COULD GIVE YOUR CMDB A FRESH START

Evaluate whether you can leverage your AI/ML APIs to automatically and continuously update your CMDB.

9 LEVERAGE THE HUMAN FACTOR

Leveraging the expertise of subject matter experts is always critical for setting realistic goals, identifying the right data sources, and selecting the appropriate AI/ML algorithms.

About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that provides deep insight across the full spectrum of IT and data management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help EMA's clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise line of business users, IT professionals, and IT vendors at www.enterprisemanagement.com or blog.enterprisemanagement.com. You can also follow EMA on Twitter, Facebook, or LinkedIn.

Corporate Headquarters:
1995 North 57th Court, Suite 120
Boulder, CO 80301
Phone: +1 303.543.9500
Fax: +1 303.543.7687
www.enterprisemanagement.com

3767-Densify.092818