

NUCLEUS
RESEARCH

SPARKCOGNITION ANNOUNCES GENERATIVE SOLUTION

ANALYST

Samuel Hamway

THE BOTTOM LINE

On March 8th, 2023, SparkCognition announced its Generative AI Platform for Industrials, enabling organizations to augment sensor data and generate insights faster using small datasets. Expected user impact includes improved model performance, reduced data labeling costs, and accelerated time to value. With SparkCognition's announcement, the vendor has positioned itself as an enterprise provider of small data solutions, an area Nucleus expects to see increased growth in over the next decade due to limitations on data access and data collection and labeling resources.

THE ANNOUNCEMENT

On March 8th, 2023, SparkCognition announced its Generative AI Platform for Industrials, enabling organizations to augment sensor data and generate insights faster using small datasets. The release aims to provide a faster, more comprehensive view of industrial processes and assets. The Generative AI Platform autonomously augments data sets by generating content in forms including synthetic text, images, audio, and other signals. By enhancing low-fidelity data and extrapolating low-resolution images to high-resolution views, the platform also aims to reduce the foundational information and time needed to make informed decisions.

To further increase the value of its Generative AI Platform, SparkCognition additionally announced it is developing industry-specific fine-tuned large language models (LLMs). These transformer-based models can recognize, summarize, translate, predict, and generate content from large unstructured datasets, providing organizations with new tools for prioritizing R&D investments, managing production, optimizing supply, and directing distribution.

Augment traditionally unusable industrial datasets with SparkCognition

SPARKCOGNITION

SparkCognition is an industrial AI solutions provider for various industries, including manufacturing, energy, utilities, and defense. The company delivers proprietary deep learning and classical machine learning models for computer vision, NLP, and others tasks. These model deployments are designed to predict future outcomes, prescribe the next best actions, enable distributed device ecosystems, and ensure worker safety. SparkCognition's solutions address value issues such as preventing downtime, maximizing asset performance, delivering net-zero initiatives, and eliminating accidents while protecting against zero-day cyberattacks. Products offered include EPP, Manufacturing Suite, Maritime Shipping Advisor, O&G Maintenance Advisor, Renewable Suite, Trading Suite, and Visual AI Advisor.

CUSTOMER IMPACT

Nucleus anticipates SparkCognition's announcement will result in tangible benefits including improved model performance, reduced data labeling costs, and accelerated time to value.

IMPROVED MODEL PERFORMANCE


By autonomously generating data in forms including synthetic text, images, audio, and other signals, the Generative AI Platform for Industrials enables data augmentation to help organizations apply machine learning to problems previously hampered by small or unrepresentative datasets. By providing more data samples and increasing data diversity, data augmentation improves the accuracy of data-hungry machine learning models, particularly deep neural networks. It also helps to prevent overfitting by improving the model's generalization through better training data. Overfitting is a condition where the model learns the training data too well and fails to generalize to new, unseen data. This failure causes an overfitting model to be unable to generate intended productivity gains.

REDUCED DATA LABELING COSTS

Data labeling is defined as the annotating of data with labels or tags to make it usable for supervised learning of machine learning algorithms. It is an essential step in training machine learning models, yet is labor-expensive and thus high in cost. By creating new data samples similar to the original data but with slight variations, the platform can reduce the need for manual data labeling, allowing businesses to reallocate associated labeling FTEs. In the manufacturing industry, for example, quality control involves inspecting products for defects or anomalies. Traditionally, this process is entirely performed by human inspectors who visually examine products for defects and manually label them. SparkCognition's announcement substantially reduces the number of instances that require manual labeling, augmenting an existing small, labeled dataset with generated images with different lighting conditions, angles, and backgrounds, simulating the real-world variability of product defects and anomalies.

ACCELERATED TIME TO VALUE

By reducing data collection and manual labeling needs, organizations can train and deploy performant models faster. This acceleration equates to earlier detection of anomalies, earlier maintenance projections, and earlier computer vision quality assurance deployments, saving organizations on downtime and unoptimized process costs that normally extend through the data collection and labeling process. Additionally, the platform's industry-specific fine-tuned LLMs will recognize, summarize, translate, predict, and generate content from large



unstructured datasets. These LLMs intend to enable decision-makers to glean insights from their organization's data corpus faster.

LOOKING AHEAD

While the past decade has touted the value of big data, many organizations lack the data access and data collection and labeling resources to conduct successful machine learning model training. These shortcomings are further exposed with the rise of data-hungry deep learning architectures for computer vision, NLP, and time-series analysis tasks. Nucleus expects that a major trend over the next decade will be technology innovations centered around generating value from traditionally unusable small datasets. These innovations are expected to be manifested through novel architectures, fine-tuned transformers for few-shot learning, new data augmentation solutions, and new model ensemble configurations. With SparkCognition's announcement, the vendor has positioned itself as an enterprise provider of solutions for small datasets, presenting a positive outlook for SparkCognition moving forward.