THOUGHT LEADERSHIP SERIES

THE FUTURE OF AI:
Machine Learning and Knowledge Graphs

database
TRENDS AND APPLICATIONS

2021 | MAR

THE FUTURE OF AI:
MACHINE LEARNING AND KNOWLEDGE GRAPHS

While aiming to enhance the way people search for information, knowledge graphs ease the complex process of searching and exploration as a lot of information is in the form of data, audio, videos, and images about a person, entity, or object.

There is a new use for graphs that has blossomed in 2020: graph machine learning. Graph neural networks operate on the graph structures, as opposed to other types of neural networks that operate on vectors. What this means in practice is that they can leverage additional information.

Right now, nearly half of *Database Trends and Applications* readers are using machine learning to better leverage their data and deliver analytical insights.

However, a variety of challenges exist, from gaining access to the right data to identifying outputs that are useful.

The flexibility of the graph model, along with its explicit storage of data relationships, makes it not only easy to

manage data coming from diverse sources but also to search and explore that data to reveal new insights that would otherwise be very difficult to discover. As a sign of its

growing popularity, in December 2020 at NeurIPS 2020, one of the world's largest AI conferences, more than 136 papers had the word "graph" in their titles.

Knowledge graphs are arguably both widely available and a well-understood technology for knowledge representation and reasoning.

Besides reaching the peak in Gartner's hype cycle for AI in 2020, knowledge graphs are increasingly being adopted in real-world applications by organizations ranging from industry-leading corporations to mid-market companies.

INTRODUCING GRAPH MACHINE LEARNING

There is also another use of graphs that blossomed in 2020: graph machine learning. Graph neural networks operate on the graph structures, as opposed to other types of neural networks that operate on vectors. What this means in practice is that they can leverage additional information.

Bringing knowledge graph and machine learning technology together can improve the accuracy of the outcomes and augment the potential of machine learning approaches.

Graph machine learning also goes by the name of geometrical machine learning, because of its ability to learn from complex data such as graphs and multi-dimensional points.

Its applications in 2020 have been relevant in biochemistry, drug design, and structural biology. Knowledge graphs and graph machine learning can work in tandem, as well.

Despite the global impact of COVID-19, 47% of AI investments were unchanged since the start of the pandemic and 30% of organizations actually planned to increase such investments, according to a Gartner poll. Only 16% had temporarily suspended AI investments, and just 7% had decreased them.

Additional use cases for graphs include the following:

to store information generated from various stages of research which can be used for building accessible models, risk management, and process monitoring.

- **Recommendation systems:** Netflix, for example, uses a knowledge graph to store a vast amount of varied information for its recommendation system which helps in finding relationships between movies, TV shows, persons, etc. Later, these preferences and connections can be used to predict what customers might like to watch next.
- **Supply chain management:** Companies can easily keep track of inventories of different components, the personnel involved, and timing, which allows them to move items more swiftly and cost-effectively.

value of a knowledge graph include the following:

Begin with a single use case, link just a few datasets and reports, and then add data and links to it organically so that it's a dynamic structure.

Once you have a use case, identify the content you'll need and classify it according to a taxonomy. While you can refer to industry-standard taxonomies for ideas, invest the time to make the taxonomy meaningful for your organization and understand how users organize their information. Buying taxonomies out of the box or contracting a consultant to do it for you is bound to lead to problems.

The organizing structure becomes even more powerful—an ontology—when you use semantic indexing to replace users' own words with synonyms to better understand what they mean. The requester doesn't need to know the exact label to retrieve the information they want.

Engage business users in the continuous development of the knowledge graph, along with taxonomists, information architects, and data scientists.

Add descriptive metadata to the knowledge graph, such as the version of the report or data lineage, so that users can decide whether it's the right data and if its quality is acceptable.

With knowledge graphs, AI language models are able to represent the relationships and accurate meaning of data instead of simply generating words based on patterns. This allows AI to be a more trustworthy partner as we search the web.

- **Question-answering:** This is one of the most frequently used applications for knowledge graphs. Knowledge graphs contain a wealth of information and question-answering is a good way to help end users to more effectively and more efficiently retrieve information from knowledge graphs.
- **Storing research:** Many companies have been using knowledge graphs recently

THE VALUE OF COMBINING KNOWLEDGE GRAPHS WITH AI

Bringing knowledge graph and machine learning technology together can improve the accuracy of the outcomes and augment the potential of machine learning approaches. According to Andrea Malick, a research director in the Data and Analytics practice at Info-Tech, steps for implementing and maximizing the

MAKING AI MORE RELIABLE

With knowledge graphs, AI language models are able to represent the relationships and accurate meaning of data instead of simply generating words based on patterns. This allows AI to be a more trustworthy partner as we search the web. ◼

—Stephanie Simone

Knowledge Graphs for Contextual AI

THE CONVERGENCE OF AI AND KNOWLEDGE GRAPHS

Companies are increasingly using AI applications for decision-making. However, due to lack of contextual information, AI systems have not yet been able to achieve their full potential as reliable solutions for complex problems.

Enter knowledge graphs: a logical way to capture data relationships and convey their meaning. Knowledge graphs drive intelligence into the data itself and give AI the context it needs to be more explainable, accurate, and repeatable.

Neither AI nor knowledge graphs is new technology, but only lately have they come of age and joined forces. Although data and compute power have contributed to their rise in the last decade, it's the powerful combination of the two that is spurring an explosion of interest in Contextual AI.

Contextual AI takes into account all surrounding information, is adaptive to different situations, behaves as expected, and is explainable to those that use it.

DRIVING INTELLIGENCE INTO DATA USING KNOWLEDGE GRAPHS

Knowledge is every company's most prized asset; however, its value is limited unless organizations can leverage that knowledge in the correct context. This is where knowledge graphs come into play.

Knowledge graphs explicitly surface the rich relationships between data that experienced domain experts naturally consider. After all, in reality, there are no isolated pieces of data, but only rich, connected domains all around us.

A knowledge graph places data in context by establishing connections among data. A knowledge graph then enriches the data's meaning and utility by adding a layer of semantics, thereby allowing software agents to reason about it. By adding relationships to data and enhancing it with semantics, knowledge graphs drive intelligence into data, making it smarter.

KNOWLEDGE GRAPHS ENHANCE ML FROM SOURCING TO TRAINING TO PREDICTIONS

AI and machine learning are playing an ever-increasing role in enterprises today. Machine learning is used in every industry:
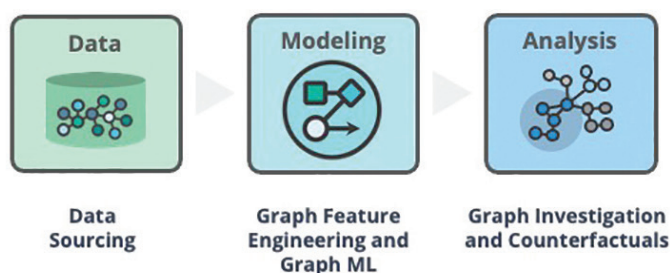
in healthcare to detect cancerous tumors, in supply chains to find factors that positively and negatively impact business, and in financial services to allow investors to identify new opportunities or know when to trade.

Interestingly, machine learning is enhanced using knowledge graphs because of their innate ability to surface context. Contextual information is known to increase predictive accuracy, make decisioning systems more flexible, and provide a framework for tracking data lineage.

Machine learning depends on data; the more the data, the higher the data quality, and the more data variety, the better the results. Unfortunately, most data science approaches leave out contextual information because connections and data structures are difficult to process. However, knowledge graphs capture, persist, and make this contextual information usable, which means they can enhance every step of the machine learning process. From data sourcing and training machine learning models to analyzing predictions and applying results, contextualized AI systems are more reliable, robust, explainable, and trustworthy.

In the initial step of data sourcing, knowledge graphs are used for data lineage to track the data that feeds machine learning- where

Knowledge Graphs Enhance All Phases of Machine Learning

Data → Modeling → Analysis

Data Sourcing → Graph Feature Engineering and Graph ML → Graph Investigation and Counterfactuals

the data came from, how the data changed, where the data is used, and who used it. If you can't trust the data used for ML, you can't trust the results. Data lineage and master data management using knowledge graphs also serve as an audit trail for compliance, especially in regulated industries.

The next phase in machine learning is training a machine learning model, which involves providing an ML algorithm with training data and significant features to learn a function for making predictions. Machine learning models without context require exhaustive training, strictly prescriptive rules, and can only be applied to specific applications. Knowledge graphs add the much-needed context that results in better predictions—all with existing data. Knowledge

graphs also allow for graph feature engineering using simple graph queries and/or more complex graph algorithms. We know that relationships are highly predictive of behavior, so using these connected, contextual features maximizes the predictive power of models while increasing how broadly a solution can be applied.

Once a machine learning model has been developed, it is essential to understand if the model is useful and if the model is making correct predictions. Knowledge graphs with incorporated relationship information allow for easy graph investigations and counterfactual analysis by domain experts. An expert might test hypotheses by exploring similar communities in the knowledge graph or debug odd results by drilling into hierarchies and dependencies.

Knowledge graphs built on graph technologies have significant advantages as graphs naturally store, compute and analyze connections and relationships among data. Moreover, graph algorithms are specifically developed to leverage the topology of data through connections: find communities, uncover influential components, and infer patterns and structure. Incorporating the predictive elements of context from a knowledge graph into machine learning not only increases accuracy but reduces false positives.

Graph-native learning involves computing machine learning tasks within a graph structure and takes knowledge graph augmented machine learning to the next level. It provides the ability to learn generalized, predictive features directly from within the graph without knowing what data structures are most predictive. This is significant as organizations don't always know which features are most important, let alone how to represent connected data for use in machine learning models.

The convergence of machine learning and knowledge graphs is just the beginning. As industries begin to see the power in combining these technologies, expect to see increasing demand for integrated solutions and streamlined workflows.

LEVERAGING KNOWLEDGE GRAPHS AND MACHINE LEARNING FOR COMPETITIVE ADVANTAGE

Companies today are leveraging knowledge graphs with machine learning for many use cases, from merely enhancing heuristics to more complex uses like training embeddings in a graph-native learning model. Examples include:

- a global e-commerce leader that has created a shopping bot using a knowledge graph to add context to machine learning to make better heuristic decisions about user intent.
- a Fortune 100 construction-equipment manufacturer that successfully utilizes machine learning and a knowledge graph for predictive maintenance and improving equipment lifespan. In this case, millions of warranty and service documents are parsed for text and added to a knowledge graph for context so that machine learning models can learn "prime examples" and anticipate required maintenance. This predictive maintenance has allowed the company to proactively take remedial action, save downtime and associated costs and increase productivity.
- a global pharmaceutical company that is combining a knowledge graph, graph queries and graph algorithms with traditional machine learning approaches to map and predict patient journeys. Billions of records over a three-year period are being used to extract patient paths and train embeddings for predicting successful intervention points and improving patient outcomes.

NEO4J KNOWLEDGE GRAPHS IN ACTION

Knowledge graphs lend themselves naturally to many use cases. The majority of these use cases fall along a spectrum of three major categories: data management, data discovery, and data analytics.

Data management use cases tend to be the most distinct and focus on data aggregation, validation, and governance. Data assurance is the motivation of these use cases, including data catalogs, data lineage and provenance, compliance, data governance, risk management, and data fabric. Data assurance knowledge graphs provide increased trust and explainability.

Data discovery uses go beyond simple information visibility, emphasizing user-led exploration, deduction, and inference of new knowledge. The impetus for many of these use cases centers around expert understanding and insights with examples such as customer, patient, or product 360, sharing lessons learned, investigations (criminal, anti-money laundering, fraud), and recommendations. Insight knowledge graphs deliver complete visibility and improved productivity.

Data analytics and data discovery use cases overlap, but the goal is to improve forecasts and predictions and prescribe optimal actions. Although understanding is still essential, the ultimate goal here is better decisioning. A few examples are customer or patient journey analysis, what-if analysis, next best action or options, churn analysis, feature engineering for machine learning, and finding communities and influencers in complex systems for investigations (criminal, anti-money laundering, fraud) and recommendations. Decisioning knowledge graphs render better predictions and more breakthroughs through the power of analytics.

Both AI and knowledge graphs are driving the next wave of competitive advantage for companies. But the question comes down to execution and which companies can use them together successfully—whether to reduce the risk of fraud, improve patient outcomes, make better investment decisions, or increase employee productivity. Are you ready? ■

NEO4J

Neo4j is the leading graph database platform. The Neo4j Graph Platform helps organizations make sense of their data by revealing how people, processes, locations and systems are interrelated. Neo4j has many knowledge graph deployments across multiple industries, including government, financial services, healthcare, retail, and manufacturing, serving use cases that include fraud detection, cybersecurity, real-time recommendations, master data management, customer 360, patient journeys, and more. www.neo4j.com