2020 ANALYTICS TRENDS

Trends in analytics continue to dominate strategy discussions as businesses seek new avenues for gaining cost savings and efficiencies leveraging the power of data. Looking forward to 2020, it's critical to first recognize the development and context of the analytics field in order to best understand where analytics is heading next.

Over the course of 2019, developments in analytics converge into 3 themes: insights-focused experimentation with advanced analytics, a re-platforming of data architecture, and the development of data science teams or centers of excellence (COE).

INSIGHTS-FOCUSED EXPERIMENTATION

Across CPG, retail, and life sciences, businesses are heavily investing in more advanced analytics focused on generating insights and experimenting with different techniques. There are many benefits to experimentation including vetting your technology stack and processes as well as growing your talent, but problems can set in if insights are the endpoint. The goal of analytics should always be to take some action to improve your business, so while there is growth in experimentation, there is still room to progress and drive toward live solutions that impact the business. A subject area where we're especially seeing this is natural language processing (NLP), an exciting new capability for many businesses. It is natural with new technologies to experiment and build out the expertise, so it's crucial to build your analytics function with processes in place to keep projects accountable to driving action. Have the conversation right up front: What are we willing to change if our hypothesis is proven true? What impact do we estimate that will have on our key metrics? How will we measure the impact?

RESTRUCTURING DATA ARCHITECTURE

Many businesses are in the process of updating their data architecture platforms to consolidate and streamline tools for all user audiences and to capitalize on modern technologies. There's value in the exercise to assess all the tools in use across functions, including data integration, preparation, storage, advanced modeling, and visualization. These tools are very often redundant or do not match the needs of the teams and organizations they're meant to serve. Data warehouses and reporting aren't going anywhere any time soon, but those investing in making the most of their data realize they need more flexible data storage and processing. Cloud data solutions that can handle unstructured data like free text and images are crucial for next generation data analytics, as are data science tools to prepare and model AI solutions and then productionize and store out results and useful data generated through the process.

BUILDING ANALYTICS COES

With investment in analytics comes the need for leadership and committed resources. We're seeing teams and COEs develop with a focus on centralizing tools and processes, augmenting talent, and generating a pipeline of use cases. These teams start out small, tackling a few use cases and working to advocate for the benefits of analytics to the organization and the leadership team. Centralizing tools and processes in a COE enables the analyst function to grow and upskill as a team, pulling in more analysts to do more advanced work. They can act as the analytics subject matter experts (SMEs) to partner with functional analysts to advise on how to provide more value in their existing work and brainstorm use cases. An important piece of this team's responsibilities is outlining data science processes, especially a demand generation engine for use cases. Analytics use cases require input from functional SMEs, technical solution design, and buy-in from leadership to commit to action; thus demand generation requires a process to vet and prioritize various use cases in the organization.

Looking forward and building on the trends that defined 2019, new themes emerge for 2020. As businesses continue to evolve their analytics maturity, there will be growth in the influence and scale of analytics teams and COEs. The continued adoption of tools that enable automation of workflows throughout the data science lifecycle will also grow. With increased adoption of production solutions, there will also be a greater emphasis on the explainability and auditability of models. And finally, there are a few areas of data science with mostly untapped potential, including machine learning (ML) time series forecasting, organizational analytics, and RPA.

COE INFLUENCE

As more and more organizations get underway with their analytics strategies and realize that data science really is a team sport, they'll understand the need for expertise and input from various roles in the organization. Once a COE is established and generating value, the next goal becomes enterprise education and scaling. This involves getting the right tools in the hands of the right people, which sometimes means setting up the processes or front-end user interfaces so that certain roles need not know the underlying tool used at all.

Educational sessions will also be critical and most effective when delivering contextual use cases. Analytics roll-out will thrive if everyone can see the benefit to their own day-to-day. Not all use cases need to be predictive modeling; is there a way to automate a weekly task to save time and headaches or to save Excel from crashing? Focus on delivering value and the rest will come.

WORKFLOW AUTOMATION

Automation can have an intimidating reputation at first, but once adopted, quickly becomes heavily relied upon. Each part of the data science project lifecycle can be automated, and more tools and platforms are striving to make that an easy part of the process. The task of automating data pipelines to get data from the source, integrated across systems, and ready for processing by analysts often falls on the data engineering team and has its own suite of tools to streamline and be efficient.

Once the data is available, the processes to profile and understand your data, visualize it, and then model it can each be automated as well. There are tools and components of commonly used platforms that will automatically show you histograms of each of your data distributions, let you know the number of missing values, and provide other helpful flags that let you initially analyze your data. The same is true on the modeling side once the data is prepared: you can select a handful of algorithms to test on the data to quickly prototype. These automation tools save time spent on repeated work and are a favored item in the data scientist's tool belt.

AUDITABILITY OF MODELS

Greater adoption of automated decisions needs to be accompanied by thorough documentation. Whether getting out ahead of regulatory bodies or dealing with a lawsuit, analytics teams should be prepared to explain how decisions were made. This comes down to two factors: explainability and documentation. Explainability refers to the ability of humans to trace through an algorithm to determine exactly where a prediction came from, and it can be difficult even for simple algorithms if enough data is involved. There are also certain algorithms that act as "black boxes" and it can be nearly impossible to interpret how a decision was made. These algorithms often return more accurate predictions, but organizations need to consider the tradeoff with explainability for their use case. When generating similar accuracies, it is often better in business applications to use the simpler, more explainable algorithm. As for documentation, those in regulated industries will need a higher level of documented validation, but it is best practice across industries to have processes in place to conduct a technical review on models before they go live and to document workflows and predictions for reference.

UNTAPPED DATA SCIENCE APPLICATIONS

Machine learning (ML) time series forecasting, organizational analytics, and robotic process automation (RPA) are a few areas with growing traction and interest. ML forecasting is the process of predicting each discreet future point and has a few benefits over traditional statistical forecasting. With ML forecasting, many data sets can be considered, including weather or holidays or anything else that helps paint context. This is in opposition to statistical forecasting, which only takes into account the historical patterns of the data you're forecasting. ML forecasting can afford both a greater degree of accuracy and granularity, opening up greater insights for your business and better inputs into decision-making.

Organizational analytics includes projects around employee engagement, like understanding the nature of diversity and inclusion and its impact on work attitudes and promotional opportunities. Three factors are driving increased interest in organizational analytics:

- New expectations from a younger generation entering the workforce, bringing a different perspective on diversity and inclusion,

- A more robust approach to risk management to protect the business from employee disengagement, complaints, or lawsuits, and

- Better understanding of the positive correlation between employee experience and business returns.

RPA, a hot topic in tech for a few years, is reaching an inflection point with greater anticipated adoption in 2020. For many businesses, the high and rapid return on investment has made RPA an attractive opportunity. RPA's ability to automate routine, rules-based tasks has enabled businesses already to achieve impressive time and cost savings with careful planning and implementation.

CLOSE

Analytics continues to be the topic amongst leadership across industries but too many businesses invest in analytics programs that don't offer comparable returns. Unlocking the potential opportunities and areas for improvement requires careful planning and understanding of the evolving analytics landscape. Effectively investing in these trends will help businesses continue to grow with the analytics field and gain a competitive edge.

CLARKSTON CONSULTING