Deloitte.

Deloitte's Anomaly Detector: DataDoc

Finding Irregularities Fast

Deloitte's anomaly detection solution "DataDoc" utilizes AI to zero in on records that – even subtly – deviate from the rest of the pack.

The Need

Digitization has been underway for decades. Processes are ever more orchestrated by computers – from simple workflows to advanced robotics. Interconnection between systems and improvements to sensor technologies raise the degree of automation further, collecting and consolidating data along the way. Data fuels the automation, making data quality more critical than ever. Errors buried deep in the data can lead to unforeseen consequences. More often than not, they can be traced to human interactions with the systems – either errors, sloppiness, or even intentional abuse or sabotage.

Whatever the motivation, the increasingly digitized nature of business can quickly multiply the impact of irregularities, propagating them across processes far afield from their source. When impacts are eventually felt downstream, their origins are often difficult to detect, obscured by intricate interrelationships and complex sets of rules between cause and effect.

In the case of fraud, perpetrators take deliberate steps to cover their tracks, adding an additional level of obscurity. Fraudsters are keenly aware of rules-based controls, how they work. They increasingly make use of AI to slip through undetected. The best – and perhaps the only – way to fight such "criminal AI" is through AI itself. Whether motivated by fraud or innocent error, outliers can be far from obvious, undetectable along one or two dimensions alone, only evident when measured against a more sophisticated, multi-dimensional pattern.

Organizations take great care to insert regular checkpoints and controls into their processes – typified by the three lines of defense.Yet control instances come under increased strain as collection and generation of data accelerates, the speed of business intensifies, and potential impacts grow. Personalization and smaller batch sizes require greater segmentation for classical methods such as statistical sampling. Risk-based auditing approaches are only as good as the judgmental rules and filters upon which they are based, which struggle to keep pace with rapid innovation and process changes. Managers, controllers, auditors, and developers are quickly overwhelmed by the sheer volume of data – finding the proverbial needle in the haystack.

Our Solution: DataDoc

Deloitte's anomaly detection solution "DataDoc" exposes potential irregularities, creating a "shortlist" for further, individualized human inspection. It achieves this without any prior knowledge about the data contents, relying solely on the statistical approaches of unsupervised learning to spot potential problems. In so doing, it relies on an array of multiple detection engines (currently nine) to examine the data presented to it from multiple angles. Just as a patient may seek a second opinion from another doctor, Consistently has a second, third, … ninth (and growing) opinion … built in. The more methods are "triggered", the more certainty we have that the data points in question are indeed anomalies.

Powered by AI, DataDoc thrives on a deep and wide dataset where its human counterpart could be quickly overwhelmed. It is more robust than classical rules-based approaches. Sophisticated methods such as auto-encoding and principle components analysis are applied to identify the most significant features along which an anomaly manifests itself. DataDoc conveniently displays results – both aggregated and individual – in a visually intuitive and interactive manner, graphically and in the form of a shortlist for the auditor to continue investigation.

Advantages/Benefits

- Identify high-potential outlier candidates within large (deep & wide) datasets.
- Without prior understanding of what the features actually represent.
- Explore the data using simple and effective visualization capabilities.

Example Use Cases

- **Fraud detection**
  revealing potential manipulative activity buried within transactional data.
- **Money laundering**
  exposing intricate patterns intended to conceal transaction traceability to their ultimate beneficial owners.
- **Preventative maintenance**
  uncovering operational aberrations typical of parts about to fail – before they actually do fail.
- **Cross-selling**
  identifying incongruent or out-of-date customer profiling information that may lead to poor targeted marketing strategies.
- **Auditing**
  enhancing risk-based auditing/judgmental prioritization with statistical analysis.

Contacts

David Thogmartin
Leader aiStudio
dthogmartin@deloitte.de

Issue 01/2022