



Implementing AI with Confidence

Deloitte's Explainable AI Offering: Lucid [ML]

Deloitte's explainable AI solution Lucid [ML] brings transparency to Machine Learning models, regardless of the selected method or underlying algorithm.

The Need

Machine Learning [ML] models are raising the bar in terms of prediction accuracy. Every day, Data Scientists around the world actively leverage the possibilities of ever broader and deeper sources of data.

The advantages are clear: greater accuracy improves quality, reduces waste, and heightens efficiency. Yes all this comes at

the cost of model complexity. Their ability to deftly harvest information from vast quantities of data renders the models themselves complex. This is particularly the case with deep neural networks (DNNs), arguably the greatest contributor to economic value from artificial intelligence technologies to date. The power of DNNs is entwined in their multiple hidden layers, each specialized in recognizing sub-patterns.

The downside: these high performing ML models are opaque "black boxes." Unable to truly understand how they work, practitioners place their faith in them on the

basis of back-testing. This poses a variety of challenges to companies seeking to exploit the power of AI, especially in regulated industries. The lack of transparency hinders adoption for banks and insurers, who are left to gaze enviously at clever innovations in other industries, such as retail commerce. ➔

Our Solution

Deloitte's explainable AI solution Lucid [ML] brings transparency to Machine Learning models, regardless of the selected method or underlying algorithm. It achieves this by subjecting the model under investigation to a series of shocks, akin to stress-testing.

Lucid combines multiple investigative approaches depending on the needs of the user. Lucid can analyze models built to handle three data types: tabular data, images, and unstructured text. The user may select the approach, prioritizing either fast, approximate analysis or preferring a deep scan for a more accurate assessment.

A powerful feature is the ability to explain decision drivers at a global or local level. The global level is of particular interest to Data Scientists seeking to communicate their work to other stakeholders. It is also a useful aid to the Data Scientist seeking to sanity-check a model. The local level is of particular interest to Audit or Compliance, ensuring that each individual decision of a "black box" AI can be divulged.

Lucid's feature-set is growing: a recent addition is the inclusion of contrastive explanations, otherwise known as "counterfactuals". This feature allows Lucid to quantify how far an individual data point was from the decision boundary.

Advantages/Benefits

- Companies may confidently apply ML models to their complex problems.
- They can both leverage modern technologies and comply with regulations.
- Practitioners may convey comfort to multiple stakeholders (CROs, CEOs, auditors, regulators) either about in-force models or models in development.

They gain insights to their business in the process, for instance which features drive results, where to focus in order to optimize.

Example Use Cases

- Interpreting credit decisioning models.
- Fulfilling profiling transparency obligations (GDPR Article 12).
- Providing explanation to individually rejected loan applications (GDPR Article 20).
- Understanding drivers of positive or negative sentiment toward their company or produces on social media.
- Identifying blind-spots in computer vision – which characteristics require more training to distinguish objects.

Contacts

David Thogmartin

Leader aiStudio
dthogmartin@deloitte.de

This presentation contains general information only, and none of Deloitte GmbH Wirtschaftsprüfungsgesellschaft or Deloitte Touche Tohmatsu Limited ("DTTL"), any of DTTL's member firms, or any of the foregoing's affiliates (collectively, the "Deloitte Network") are, by means of this presentation, rendering professional advice or services. In particular this presentation cannot be used as a substitute for such professional advice. No entity in the Deloitte Network shall be responsible for any loss whatsoever sustained by any person who relies on this presentation. This presentation is to be treated confidential. Any disclosure to third parties – in whole or in part – is subject to our prior written consent.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. Please see www.deloitte.com/de/UeberUns for a more detailed description of DTTL and its member firms.