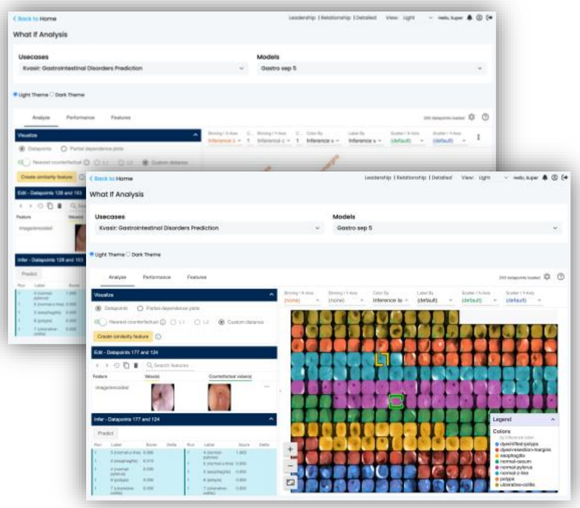
Questions
support@hexagon-ml.com

Model Simulation using Hexagon-ml and the Power of "What-If "

In the contemporary digital landscape, machine learning and artificial intelligence (AI) stand at the forefront, driving transformative changes across various industries. However, a prevailing concern is the enigmatic "black-box" character of numerous AI models. While their predictive capabilities can be astoundingly precise, the rationale behind these predictions remains, more often than not, a mystery. This lack of transparency poses significant challenges, particularly when such models play pivotal roles in making critical decisions in areas like healthcare, where patient outcomes are at stake, finance, where monetary implications are vast, and law enforcement, where public safety and justice hang in the balance. Addressing this pressing need for clarity and understanding, Google's PAIR (People + AI Research) initiative unveiled the "What-If Tool." This innovative tool is not just another addition to the tech arsenal but represents a monumental leap in the domain of model simulation and interpretation. It promises to shed light on the intricate workings of AI models, aiming to demystify the processes that lead to specific predictions. By doing so, it hopes to bridge the gap between complex AI computations and human comprehension, ensuring that technology remains both advanced and accessible.



What If Analysis

Usecases: Kvasir: Gastrointestinal Disorders Prediction

Models: Gastro sep 5

● Light Theme ○ Dark Theme

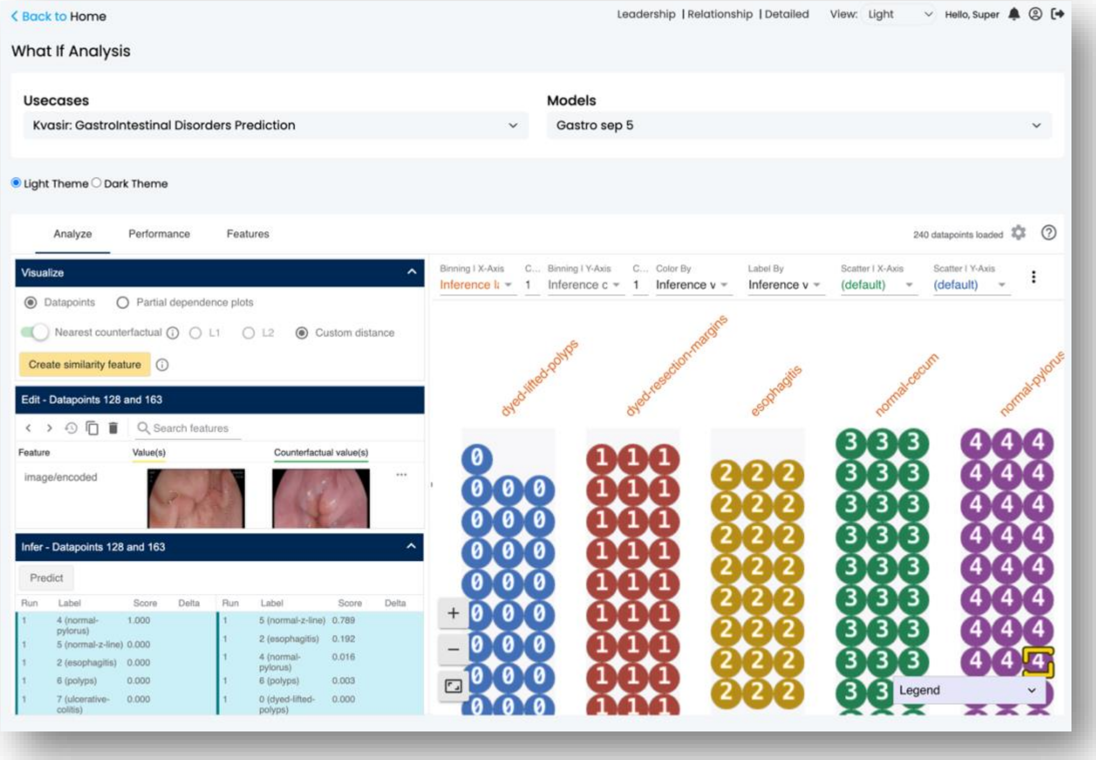
Deep Dive into the What-If Tool

The What-If Tool isn't just another analytical tool; it's a comprehensive visual interface designed to demystify machine learning models. It allows users to interact with their models, tweaking input data, analyzing predictions, and even comparing different models, all without a single line of code.

A Closer Look at Its Features

Counterfactual Analysis: This is perhaps the most powerful feature of the What-If Tool. Imagine being able to change specific data points and immediately see how those changes affect the model's predictions. This capability is invaluable for understanding how the model might react to edge cases or for identifying potential biases in its predictions.

Performance Comparison: In the world of AI, where iterative development is the norm, data scientists often grapple with multiple versions of a model. The What-If Tool simplifies this by allowing users to juxtapose two models, making it easier to discern performance differences and decide on the best iteration.



What If Analysis

View: Light

Usecases: Kvasir: GastroIntestinal Disorders Prediction

Models: Gastro sep 5

● Light Theme ○ Dark Theme

240 datapoints loaded

● Datapoints ○ Partial dependence plots

Nearest counterfactual ○ L1 ○ L2 ● Custom distance

Binning I X-Axis	C...	Binning I Y-Axis	C...	Color By	Label By	Scatter I X-Axis	Scatter I Y-Axis
Inference l	1	Inference c	1	Inference v	Inference v	(default)	(default)

Edit - Datapoints 128 and 163

Feature	Value(s)	Counterfactual value(s)
image/encoded		

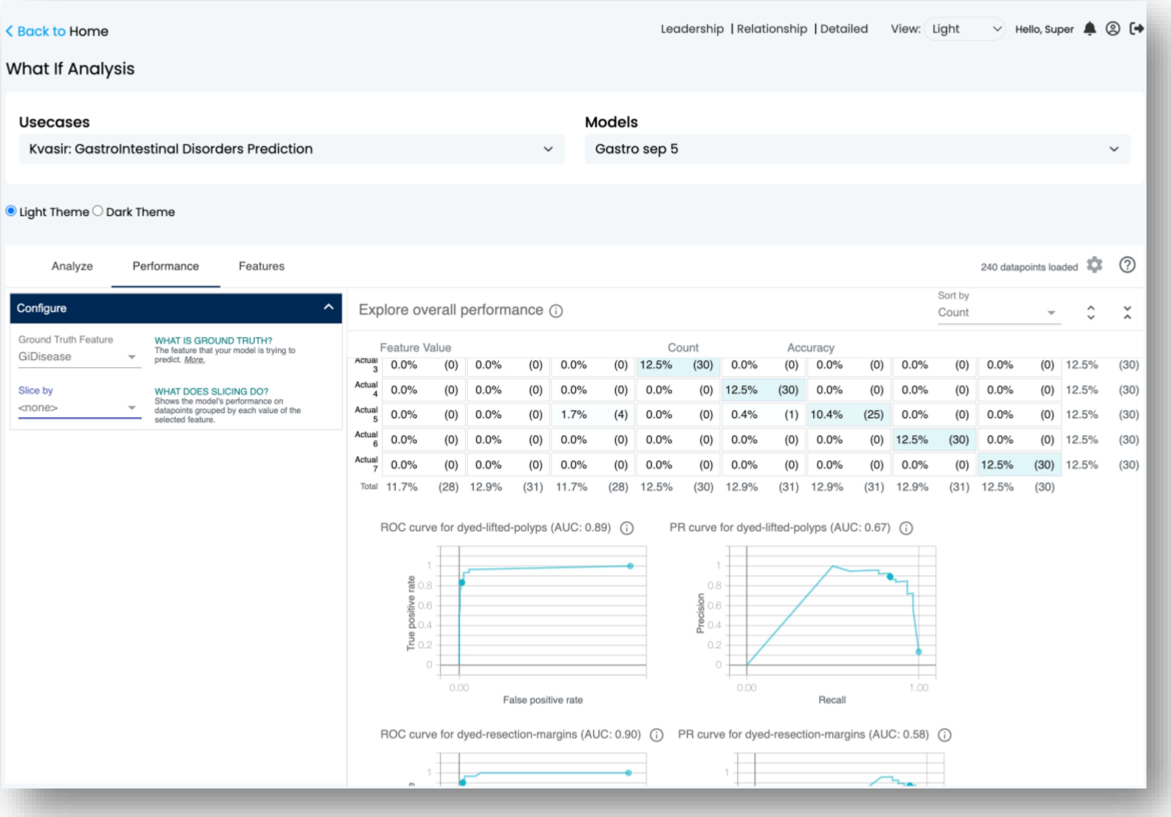
Infer - Datapoints 128 and 163

Run	Label	Score	Delta
1	4 (normal-pylorus)	1.000	
1	5 (normal-z-line)	0.000	
1	2 (esophagitis)	0.000	
1	6 (polyps)	0.000	
1	7 (ulcerative-colitis)	0.000	

Run	Label	Score	Delta
1	5 (normal-z-line)	0.789	
1	2 (esophagitis)	0.192	
1	4 (normal-pylorus)	0.016	
1	6 (polyps)	0.003	
1	0 (dyed-lifted-polyps)	0.000	

Slice and Dice: Data is multifaceted, and understanding how a model performs across different segments of data can provide profound insights. The tool enables users to segment data based on specific features, offering a granular view of the model's performance across these segments.

Fairness Evaluation: The tool's fairness indicators are a nod to the growing emphasis on ethical AI. By assessing model fairness across different groups, it ensures that models are equitable and don't perpetuate existing biases.


< Back to Home
Leadership | Relationship | Detailed
View: Light
Hello, Super
Analyze
Performance
Features
Configure
WHAT IS GROUND TRUTH?
WHAT DOES SLICING DO?
ROC curve for dyed-lifted-polyps (AUC: 0.89)
True positive rate
1
0.8
0.6
0.4
0.2
0
0.00
False positive rate
PR curve for dyed-lifted-polyps (AUC: 0.67)
Precision
1
0.8
0.6
0.4
0.2
0
0.00
1.00
Recall
ROC curve for dyed-resection-margins (AUC: 0.90)
1
PR curve for dyed-resection-margins (AUC: 0.58)
1

What If Analysis

Usecases Kvasir: Gastrointestinal Disorders Prediction

Models Gastro sep 5

● Light Theme ○ Dark Theme

240 datapoints loaded

Ground Truth Feature GiDisease

The feature that your model is trying to predict. *More.*

Slice by <none>

Shows the model's performance on datapoints grouped by each value of the selected feature.

Explore overall performance

Sort by Count

Feature Value Count Accuracy

Actual 3	0.0% (0)	0.0% (0)	0.0% (0)	12.5% (30)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	12.5% (30)
Actual 4	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	12.5% (30)	0.0% (0)	0.0% (0)	0.0% (0)	12.5% (30)
Actual 5	0.0% (0)	0.0% (0)	1.7% (4)	0.0% (0)	0.4% (1)	10.4% (25)	0.0% (0)	0.0% (0)	12.5% (30)
Actual 6	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	12.5% (30)	0.0% (0)	12.5% (30)
Actual 7	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	12.5% (30)	12.5% (30)
Total	11.7% (28)	12.9% (31)	11.7% (28)	12.5% (30)	12.9% (31)	12.9% (31)	12.9% (31)	12.5% (30)	

Interactive Visualization: With a suite of visualization options, from detailed scatter plots to informative histograms, the tool offers users myriad ways to delve into their data and model predictions.

The Broader Implications for AI and Stakeholders

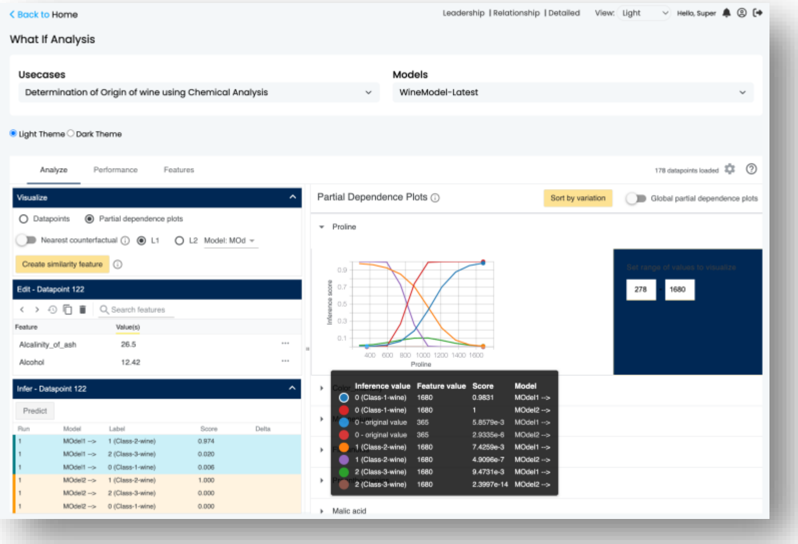
The introduction of the What-If Tool is more than just a technological advancement; it represents a paradigm shift in the AI community towards greater transparency and interpretability.

- **Democratization of AI:** One of the most significant barriers to AI adoption has been its perceived complexity. By making model interpretation more intuitive and accessible, the What-If Tool is democratizing AI, ensuring that even non-experts can understand and interact with complex models. This is crucial for industries where stakeholders without a deep AI background need to make informed decisions based on model outputs.

- **Building Trust in AI Systems:** Trust is a cornerstone of AI adoption. Stakeholders, whether they're doctors using AI for diagnosis or financial analysts leveraging AI for market predictions, need to trust the models they use. By offering insights into how models make decisions, the What-If Tool fosters this trust, ensuring that stakeholders can use AI with confidence.

- **Informed Decision-Making for Data Scientists:** For data scientists, the tool is a treasure trove of insights. Whether it's deciding on the right features, tweaking model parameters, or choosing between different models, the insights from the What-If Tool can guide data scientists, ensuring that their decisions are data-driven and informed.

- **Promoting Ethical AI:** With its fairness indicators, the tool is a step towards more ethical AI. By highlighting potential biases, it ensures that models are not just accurate but also fair, paving the way for more responsible AI deployments.



What If Analysis

Usecases: Determination of Origin of wine using Chemical Analysis

Models: WineModel-Latest

View: Light

● Light Theme ○ Dark Theme

Analyze | Performance | Features

178 datapoints loaded

Visualize

○ Datapoints ● Partial dependence plots

Nearest counterfactual ● L1 ○ L2 Model: MOd

Partial Dependence Plots

278 - 1680

Edit - Datapoint 122

Feature	Value(s)
Alcalinity_of_ash	26.5
Alcohol	12.42

Infer - Datapoint 122

Run	Model	Label	Score	Delta
1	MOdel1 -->	1 (Class-2-wine)	0.974	
1	MOdel1 -->	2 (Class-3-wine)	0.020	
1	MOdel1 -->	0 (Class-1-wine)	0.006	
1	MOdel2 -->	1 (Class-2-wine)	1.000	
1	MOdel2 -->	2 (Class-3-wine)	0.000	
1	MOdel2 -->	0 (Class-1-wine)	0.000	

	Inference value	Feature value	Score	Model
●	0 (Class-1-wine)	1680	0.9831	MOdel1 -->
●	0 (Class-1-wine)	1680	1	MOdel2 -->
●	0 - original value	365	5.8579e-3	MOdel1 -->
●	0 - original value	365	2.9335e-6	MOdel2 -->
●	1 (Class-2-wine)	1680	7.4259e-3	MOdel1 -->
●	1 (Class-2-wine)	1680	4.9096e-7	MOdel2 -->
●	2 (Class-3-wine)	1680	9.4731e-3	MOdel1 -->
●	2 (Class-3-wine)	1680	2.3997e-14	MOdel2 -->

Case Studies: The What-If Tool in Action

To truly understand the tool's impact, let's consider a few hypothetical scenarios:

Healthcare: Imagine a machine learning model designed to predict patient readmission rates in hospitals. Using the What-If Tool, hospital administrators can tweak patient data, like age, medical history, or treatment plans, to see how these changes affect readmission predictions. This can guide treatment decisions, ensuring better patient outcomes.

Finance: In a stock prediction model, financial analysts can use the tool to understand how different economic indicators affect stock prices. By tweaking data points like interest rates or unemployment figures, analysts can get a clearer picture of market dynamics, guiding their investment strategies.

Law Enforcement: For a model predicting crime hotspots in a city, law enforcement agencies can use the tool to understand the model's decision-making process. By changing data points like population density or the number of patrol officers, they can get insights into effective crime prevention strategies.

References

1. J. Wexler, M. Pushkarna, T. Bolukbasi, M. Wattenberg, F. Viégas and J. Wilson, "The What-If Tool: Interactive Probing of Machine Learning Models," in IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 1, pp. 56-65, Jan. 2020, doi: 10.1109/TVCG.2019.2934619.
2. https://www.cs.ubc.ca/~tmm/courses/547-19/slides/patrick-whatiftool.pdf
3. https://www.tensorflow.org/tensorboard/what_if_tool
4. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. *TensorFlow: Large-scale machine learning on heterogeneous systems*, 2015. Software available from tensorflow.org.
5. S. Amershi, M. Chickering, S. Drucker, B. Lee, P. Simard, and J. Suh. Modeltracker: Redesigning performance analysis tools for machine learning. In *Proceedings of the Conference on Human Factors in Computing*

Systems (CHI 2015). ACM - Association for Computing Machinery, April2015.

6. J. Angwin, J. Larson, S. Mattu, and L. Kirchner. *Machine bias: Theres software used across the country to predict future criminals. and its biased against blacks*. ProPublica, May2016.
7. R. K. E. Bellamy, K. Dey, M. Hind, S. C. Hoffman, S. Houde, K. Kannan, P. Lohia, J. Martino, S. Mehta, A. Mojsilovic, S. Nagar, K. N. Ramamurthy, J. Richards, D. Saha, P. Sattigeri, M. Singh, K. R. Varshney, and Y. Zhang. *AI Fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias*, Oct.2018.
8. J. Buolamwini and T. Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, volume 81 of Proceedings of Machine Learning Research, pages 77–91, 2018.