

Revolutionizing Radiology: Rad AI's Leap to Personalized Diagnostics with Kolena's Testing Suite

About Kolena

Kolena is an end-to-end machine learning testing & monitoring platform. By analyzing your AI models at a granular level, Kolena ensures they're rigorously tested for performance, regressions, and deployment readiness



"Kolena's testing suite has been a transformative tool for Rad AI, allowing us to optimize our model testing capabilities and evaluate the model performance with precision and granularity. This collaboration has not only improved our end-to-end machine learning pipelines significantly but also strengthened the confidence our customers have in our AI solutions."

Deniz Zorlu
Director of Machine Learning; Rad AI

CHALLENGES



- Build a model Quality Standard to build alignment and trust cross teams and with customers.
- Maintaining high accuracy to build trust and adoption among a detail-oriented customer base
- Ensuring high customization levels for each radiologist's unique reporting style.
- Managing large-scale testing to maintain accuracy and prevent regressions.

SOLUTIONS



- Standardized and rigorous model evaluation process
- Test case based solutions
- Customized evaluation metrics
- Root cause failure analysis to catch and stop regressions before deploying

RESULTS



1

PRODUCTIVITY

Savings of more than 90% of time spent on model testing and validation which led to a substantial increase in the speed of model development.

2

MODEL ROBUSTNESS

An 80% reduction of model failures in production, bolstering the reliability and trust in Rad AI's solutions.

3

TEAM ALIGNMENT AND TRUST

Enhanced cross-team visibility and alignment on model performance, fostering better communication and collaboration across the organization.



 www.kolena.com

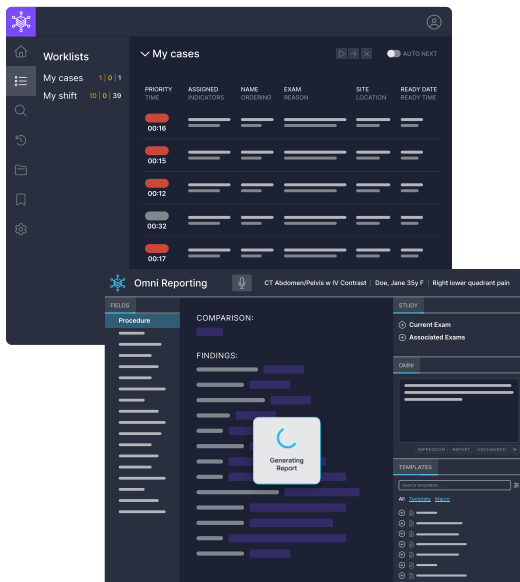
 info@kolena.com

ABOUT RAD AI

Founded by the youngest US radiologist in history, Rad AI empowers physicians with AI to save time, reduce burnout, and improve the quality of patient care. By combining our deep expertise in healthcare and AI and using one of the largest proprietary radiology report datasets in the world, our AI has uncovered hundreds of new cancer diagnoses for patients and reduced the error rate in tens of millions of radiology reports by nearly 50%.



Rad AI has raised \$50+ million to date from venture funds such as Gradient (Google's AI fund) and ARTIS. They've also formed a partnership with Google to collaborate on the future of generative AI to redefine healthcare. Currently, more than 1/3 of radiology groups and healthcare systems, including Kaiser Permanente, HCA Healthcare, and Geisinger, now leverage the latest Gen AI advancements from Rad AI.



RAD AI'S PRODUCTS

- **Rad AI Omni Reporting Ecosystem:**
 - Reporting employs advanced language models (LLMs/NLP), creating comprehensive and accurate reports with remarkable speed.
 - Impressions automatically generates report impressions from dictated findings. The impression language is individually customized to each radiologist & practice.
 - Worklist is an AI-enabled dynamic worklist.
- **Rad AI Continuity Ecosystem:**
 - Closes the loop on follow-up recommendations for significant incidental findings in radiology reports.

CHALLENGES

Serving customers in the critical healthcare space, Rad AI had significant challenges to overcome to deliver their vision of truly personalized diagnostic solutions. With thousands of radiologists, each with their own distinct style, Rad AI had to achieve a level of personalization that maintained bespoke service quality without compromise. High accuracy and performance are essential to build trust and adoption among their detail-oriented customer base, who demand precision in language and terminology. For the Rad AI machine learning team, ensuring the scalability and efficiency of their tooling was critical to their success.

The changing nature of generative AI models requires Rad AI to regularly adjust and test their models, ensuring they maintain a strong level of clinical and stylistic accuracy. This required extensive testing frameworks to maintain model accuracy and prevent regressions, and the scale of the testing required presented new logistical challenges for the team. The complexity of Rad AI's offerings meant that custom metrics and evaluation methods became necessary, and ad-hoc testing methods were not robust enough.

SOLUTION

Kolena offered a comprehensive, test case-driven evaluation framework for Rad AI, featuring customized metrics and a standardized method for detecting and preventing potential regressions before deployment. Specifically, Rad AI evaluates the performance of trained models on a radiologist basis, ensuring significant improvements in metrics such as BLEU, Rouge, and proprietary measures assessing the clinical accuracy of the generated text.

The detailed analysis provided by Kolena enables the Rad AI team to identify specific cohorts or classes of reports that require further improvements. The user-friendly interface, which allows for report-level visibility, assists the team in pinpointing potential issues more efficiently. This enhanced framework facilitates better cross-team visibility, communication, and alignment concerning model performance reports.

With Kolena, Rad AI implemented:

- Standardized and rigorous model evaluation process
- Test case based solutions
- Customized evaluation metrics
- Root cause failure analysis to catch and stop regressions before deploying

Rad AI's decision to partner with Kolena was underpinned by several key differentiators that set Kolena apart from its competitors. Kolena's unique ML unit testing approach provides the granular inspection and customization needed to handle the high variability in radiologist reporting styles, which was crucial for Rad AI's personalized diagnostics.

The ability of Kolena's testing suite to integrate with Rad AI's specific workflows, particularly in the nuanced field of radiology and healthcare, offered a tailored solution that generic platforms couldn't match.

RESULTS

The implementation of Kolena's robust test case-driven evaluation framework marked a significant milestone for Rad AI. With tailored metrics and a standardized process to identify and halt potential regressions pre-deployment, Rad AI achieved remarkable outcomes:

- Productivity: saved more than 90% of time spent on model testing and validation (from 2 weeks/model to 1-2 hours) which led to a substantial increase in the speed of model development.
- Model robustness: up to 80% reduction in model failures in production, bolstering the reliability and trust in Rad AI's solutions.
- Team alignment and trust: enhanced cross-team visibility and alignment on model performance, fostering better communication and collaboration across the organization.
- Improved granularity in testing model behavior: more precise adjustments and optimizations for models.

RESULTS (CON'T)

The comprehensive approach to model testing and deployment has led to several key benefits for Rad AI:

- Increased customer productivity and time saved for radiologists, contributing to higher satisfaction and a better user experience.
- Standardization of the evaluation and deployment process, ensuring consistency and efficiency in bringing models to production.
- Automation of model testing and deployment processes, streamlining operations and reducing bottlenecks.
- Enhanced ability to observe system performance, understand regressions, and iterate faster, resulting in a more agile and responsive development cycle.
- A more structured and diligent approach to model deployment, ensuring that each model is thoroughly vetted before being rolled out to customers.

CONCLUSION AND FUTURE PLANS

The integration of Kolena's testing suite has revolutionized Rad AI's approach to personalized diagnostics in radiology. It has resulted in increased efficiency, reliability, and customer satisfaction.

Looking ahead, Rad AI is poised for continued innovation. The company is expected to further enhance its engagement with Kolena's solutions, reinforcing its position as a leader in AI-powered radiology. This ongoing collaboration is anticipated to drive further advancements, ensuring that Rad AI remains at the cutting edge of technological progress in the healthcare industry.

Overall, the partnership with Kolena has empowered Rad AI to build the best product in the industry, characterized by personalized, reliable, and high-performing AI solutions that meet the demanding needs of radiology practitioners.

HOW KOLENA CAN HELP YOU

Discover how Kolena can transform your AI validation process. Visit www.kolena.com for more information or contact us to schedule a demo and see our solutions in action.



Reliable and Systematic Model Testing for Every AI/ML Problem

Customizable for any workflow, data type, evaluation logic, metric, plot, and report

www.kolena.com