

How SentioAI performs against the leading TAR product

The Test

Both SentioAI and a major eDiscovery review platform with built in continuous active learning (“Industry TAR”) provide model building technology to classify documents and predict results. One of our clients asked us to perform a side by side comparison to evaluate the performance of both solutions.

21,375 Documents from a previously reviewed case were loaded into two instances of Industry TAR review product; the same team of reviewers reviewed one using SentioAI and one using Industry TAR. 75 seed documents, coded the same, were loaded into each instance.

The workflows for each instance were used to select, review and quality check 300 documents on four issue tags. Both instances had the same document universe and the same seed documents. The only difference is the SentioAI technology/workflow and the Industry TAR technology/workflow.

The models developed in the previous step were used to predict the responsiveness of all documents. The results of the model predictions on key documents and a set of randomly selected non-responsive documents were compared and reported.

The Results

Total number of documents in the population 21,375

Total number of nonresponsive documents in the population 17,201

	SentioAI	vs.	Industry TAR
Key Document Predictions	80 of 124 documents were predicted RESPONSIVE		57 of 124 documents were predicted RESPONSIVE
	60% correctly identified		40% correctly identified
Model Validation on 100 randomly selected NON-RESPONSIVE documents	Validation Test Results 98% ACCURATE		Elusion Test Results 91% ACCURATE

In this case, it is more than 19,000 times more difficult to achieve 98% accuracy using SentioAI than it is to achieve 91% accuracy as in using Industry TAR.

Distribution patterns in eDiscovery document review typically fall into a bell curve where the expectation is that undefined documents will lie within the center (the bell).

SentioAI's classification algorithm builds a mathematical model that explicitly splits documents into positive (**GREEN**) and negative (**RED**), with confidence scores known as a learning index.

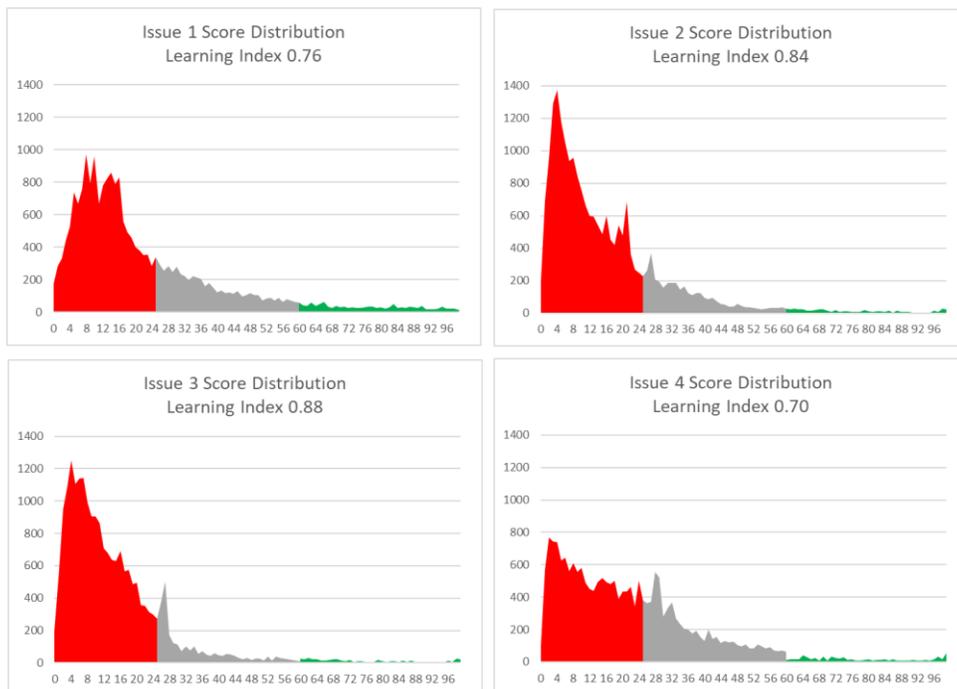
The learning index is ≥ 0 and ≤ 1 . The higher the learning index, the better the quality of the model.

Learning Index	SentioAI	vs.	Industry TAR
Tag 1	76		14
Tag 2	84		2
Tag 3	88		3
Tag 4	70		4

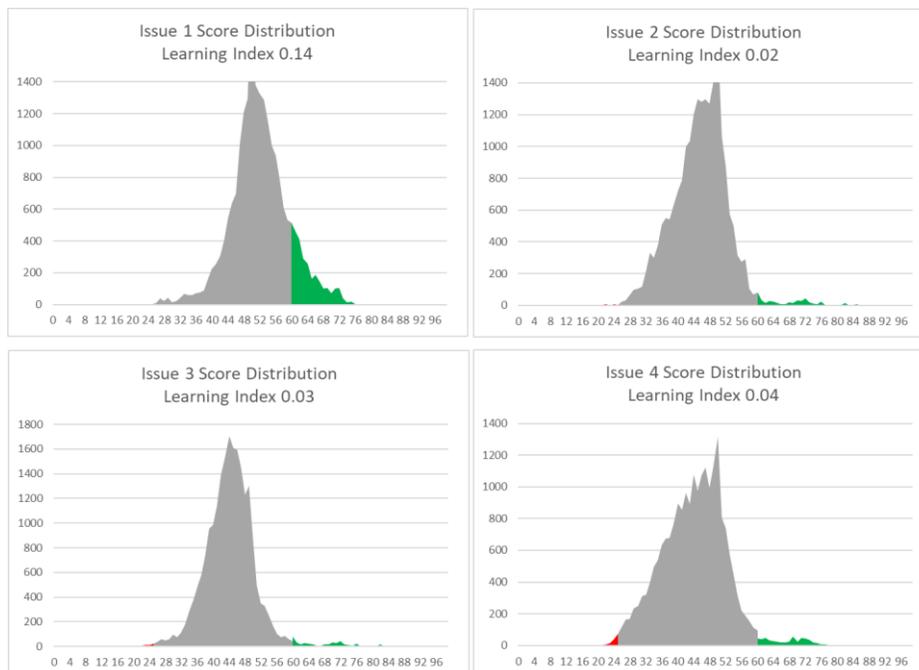
Using the same inputs, SentioAI is learning faster than Industry TAR.

The distribution curve below shows unresponsive documents in **RED** and responsive documents in **GREEN**, by issue.

SentioAI's Model Prediction Score Distribution Curve



Industry TAR's Model Prediction Score Distribution Curve



In addition to faster learning, SentioAI consolidates the workflow. It builds and adjusts models for each issue incrementally as each document is reviewed or overturned, instead of batching sets of documents together then updating the model. SentioAI returns a list of selected documents that can be presented in a prioritized order to make sure that reviewers do not miss important documents. Industry TAR's workflow trains one model at a time.

In Our Client's Words

“Just finished this QC. It was a little scary how precise the tool was on calls that were potentially false positives or false negatives. Quite exciting.”

Client Review Manager
Fortune 10 Company

Contact Us

If your company is looking to achieve a correct, complete, consistent review with speed and accuracy, contact us. Sentio Software has a proven track record of helping clients save time and money and we'd like to help you do the same. Contact us today to find out how we can bring winning results to your review projects.

Sentio Software
1900 K Street NW, Suite 725
Washington, DC 20006
(240) 883-6017
sales@sentiosoft.com | www.sentiosoft.com