
Accelerating Active Learning Image Labeling Through Bulk Shift Recommendations

Appendix

Methods	
Method name	Description
<i>clustering</i>	Executes a selected clustering algorithm.
<i>create_dictionary</i>	Creates a python dictionary (key-value pair) containing the labels (key) and the image file names (values).
<i>get_feature_vector</i>	Computes image feature vectors using on a pre-trained neural network, in our case <i>VGG16</i> .
<i>shift_image</i>	Moves an image from one folder to another.
<i>classify</i>	Executes a classification algorithm based on a dictionary.
<i>accept_image</i>	Called, if a user accepts an image as being correctly labeled.
<i>get_completeness</i>	Computes the completeness, which is the percentage of the correctly labeled images.
<i>increment_click_counter</i>	Will be triggered, if a user interacts with the system by accepting or changing the sorting .

Table 1: Description of the most important methods of the active learning system.

Table 2: Description of the test-metrics for evaluation

Methods	
Test-metric	Description
<i>Rounds</i>	The number of rounds in a pool-based active learning cycle.
<i>Seen images</i>	The number of images the human annotator is viewed during sorting.
<i>Accuracy</i>	The accuracy of the selected classifier.
<i>Completeness</i>	The percentage of the images that are labeled according to the manual reference sorting.
<i>Shifted</i>	The number of images that are shifted by the human annotator or computer algorithm (automated).
<i>Clicks</i>	The number of clicks during the sorting process.
<i>Savings</i>	The percentage of savings achieved using active learning compared to passive learning.

Table 3: Final accuracy after executing the classification algorithm on different uncertainty measures. The green/red cell color indicates the maximum/minimum accuracy.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	0.80	0.91	0.92	0.88
<i>Multi Layer Perceptron</i>	0.77	0.90	0.86	0.84
<i>Decision Tree</i>	0.48	0.44	0.48	0.47
Mean	0.68	0.75	0.75	

Table 4: Percentage of clicks that are required for all images to be labeled according to the reference sorting compared to passive learning. The green/red cell color indicates the maximum/minimum percentage.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	0.27	0.25	0.22	0.23
<i>Multi Layer Perceptron</i>	0.38	0.28	0.24	0.30
<i>Decision Tree</i>	0.43	0.60	0.44	0.51
Mean	0.36	0.38	0.30	

Table 5: Number of rounds required to label all images according to the reference sorting.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	11	6	6	8
<i>Multi Layer Perceptron</i>	11	7	6	8
<i>Decision Tree</i>	11	8	11	10
Mean	11	7	8	

Table 6: Runtime [s] required for executing in the active learning process. The green/red cell color indicates the minimum/maximum runtime.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	1160.15	772.91	1131.84	1021.63
<i>Multi Layer Perceptron</i>	1631.65	1097.40	1719.03	1482.69
<i>Decision Tree</i>	614.03	501.73	590.78	568.85
Mean	1135.28	790.68	1147.21	

Table 7: Size of the area between the active learning curve and the y-axis ranging from 0 to 1. The green/red cell color indicates the minimum/maximum size.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	0.06	0.03	0.03	0.04
<i>Multi Layer Perceptron</i>	0.06	0.05	0.04	0.05
<i>Decision Tree</i>	0.06	0.10	0.06	0.07
Mean	0.07	0.06	0.04	

Table 8: Percentage of saved clicks using active learning compared to passive learning. The green/red cell color indicates the maximum/minimum savings.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	86.80	92.60	92.60	90.66
<i>Multi Layer Perceptron</i>	86.80	89.20	91.00	89.00
<i>Decision Tree</i>	87.40	78.40	87.40	84.40
Mean	86.67	86.73	90.33	

Table 9: Number of recommended bulk moves in the active learning process. The green/red cell color indicates the maximum/minimum number of recommended bulk moves.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	166	127	125	139
<i>Multi Layer Perceptron</i>	92	107	110	103
<i>Decision Tree</i>	27	16	26	23
Mean	95	83	87	

Table 10: Mean purity of the bulk move in the active learning process. The green/red cell color indicates the maximum/minimum of the mean purity.

Classifier / Uncertainty Measure	Margin Sampling	Least Confidence	Shannon Entropy	Mean
<i>Logistic Regression</i>	0.91	0.84	0.86	0.87
<i>Multi Layer Perceptron</i>	0.83	0.80	0.82	0.82
<i>Decision Tree</i>	0.72	0.47	0.72	0.63
Mean	0.82	0.70	0.80	

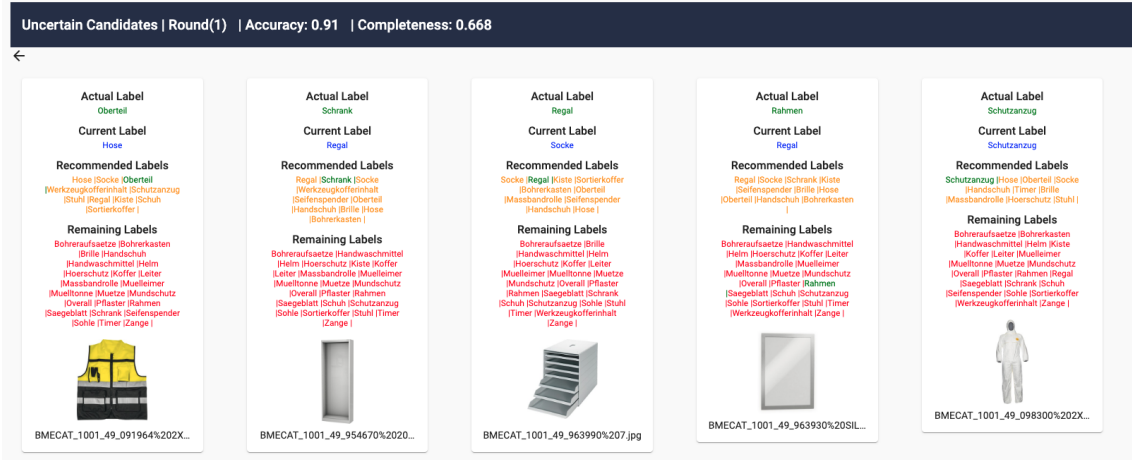


Figure 3: Image candidates after classification using *Logistic Regression* as classifier and *Shannon Entropy* as uncertainty measure.

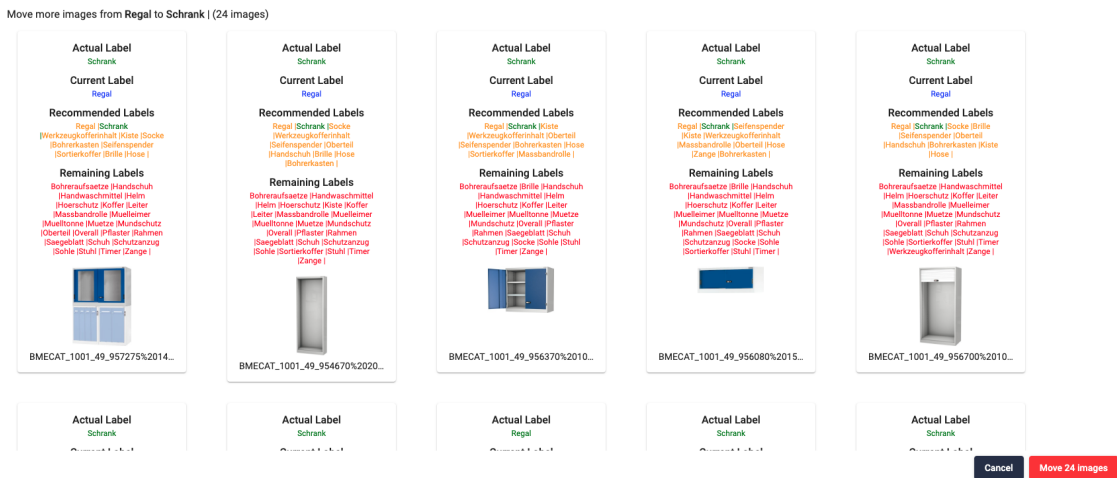


Figure 4: Suggestion of further images that additionally could be moved based on the selected recommended label (Schrank).

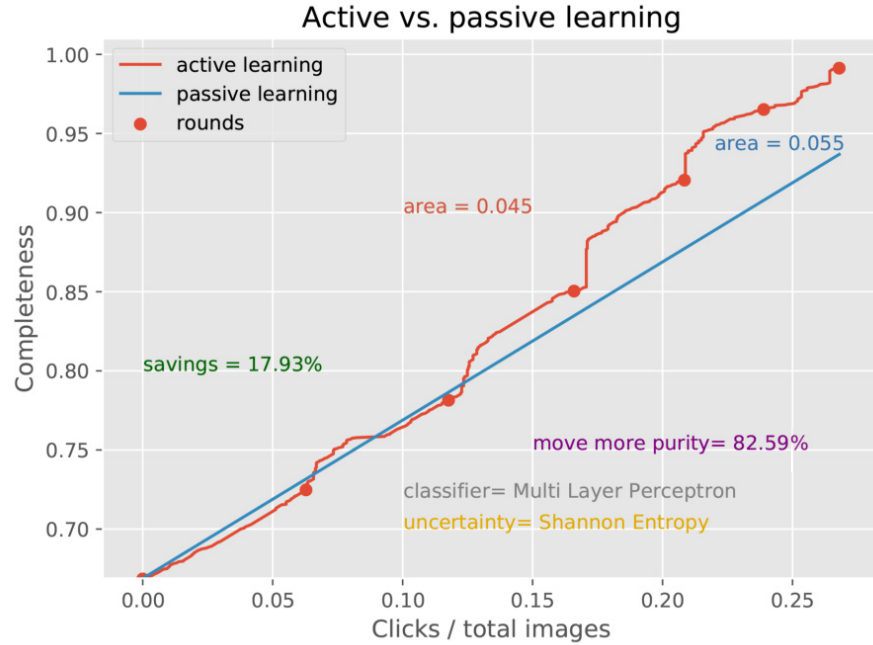
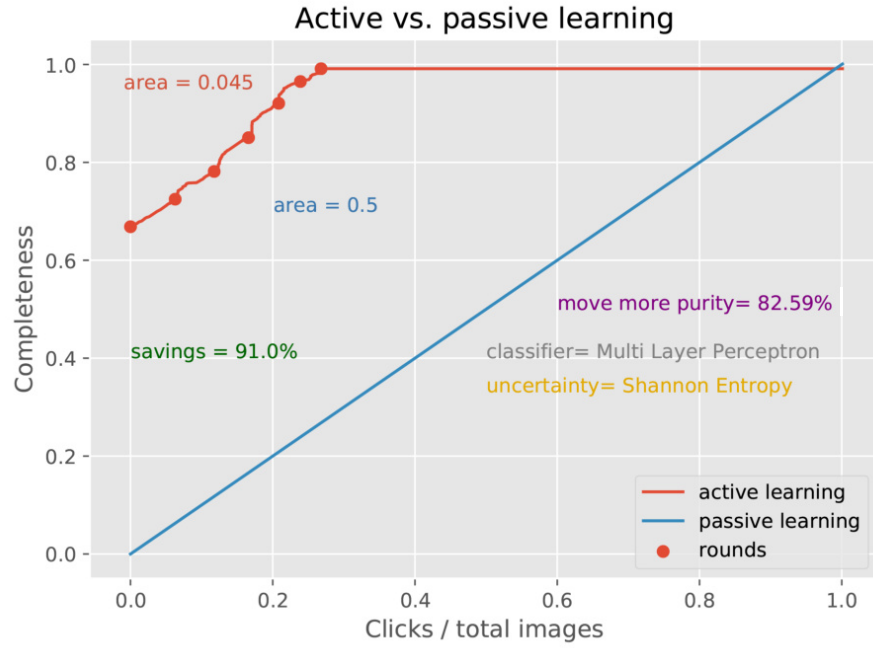


Figure 5: An example evaluation of an active learning curve compared to passive learning started from the beginning (above) and after the pre-clustering (below).

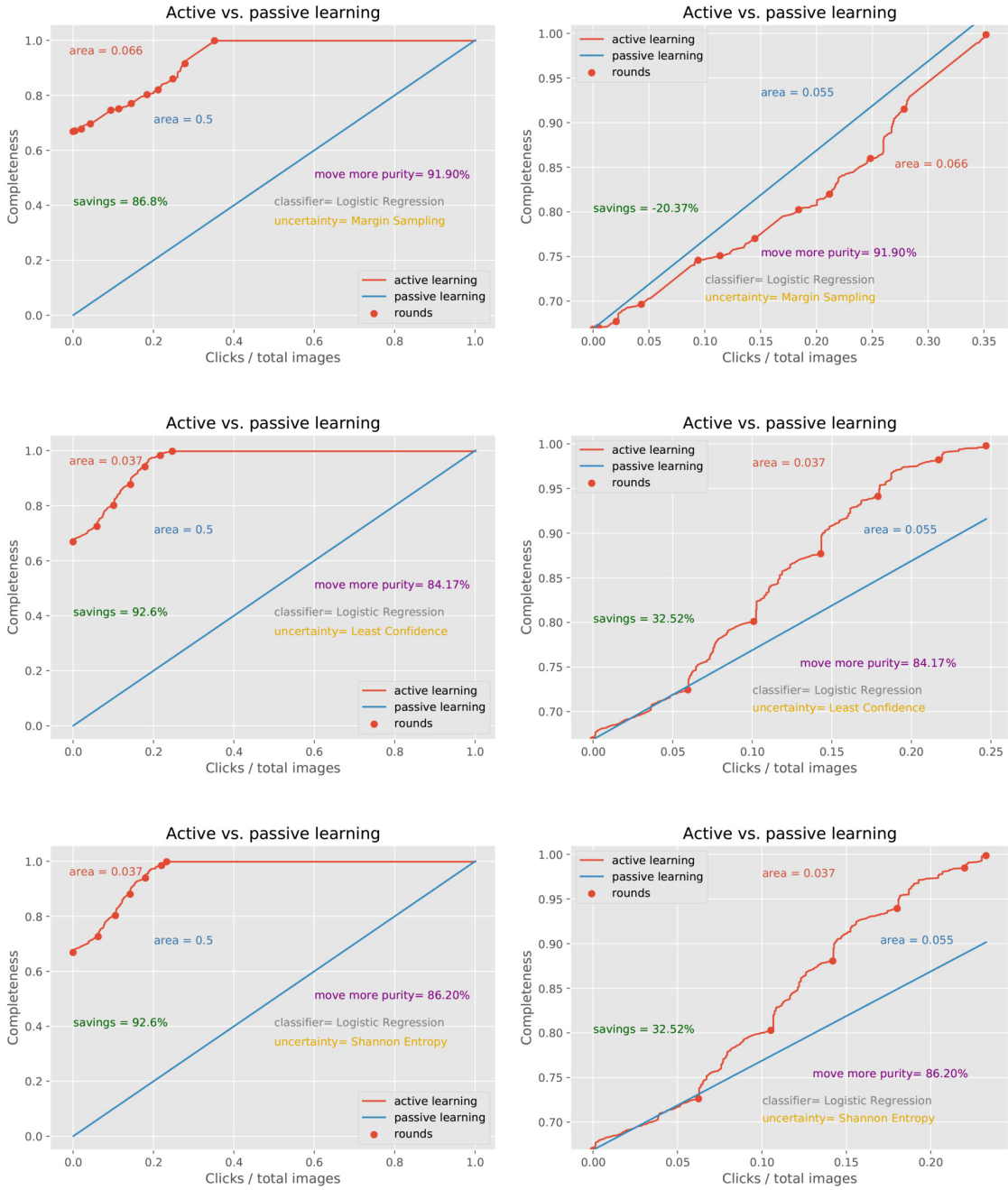


Figure 6: Evaluation of **Logistic Regression** classifier and *Margin Sampling* (above), *Least Confidence* (middle) and *Shannon Entropy* (below) uncertainty compared to passive learning started from the beginning (left), and after the pre-clustering (right).

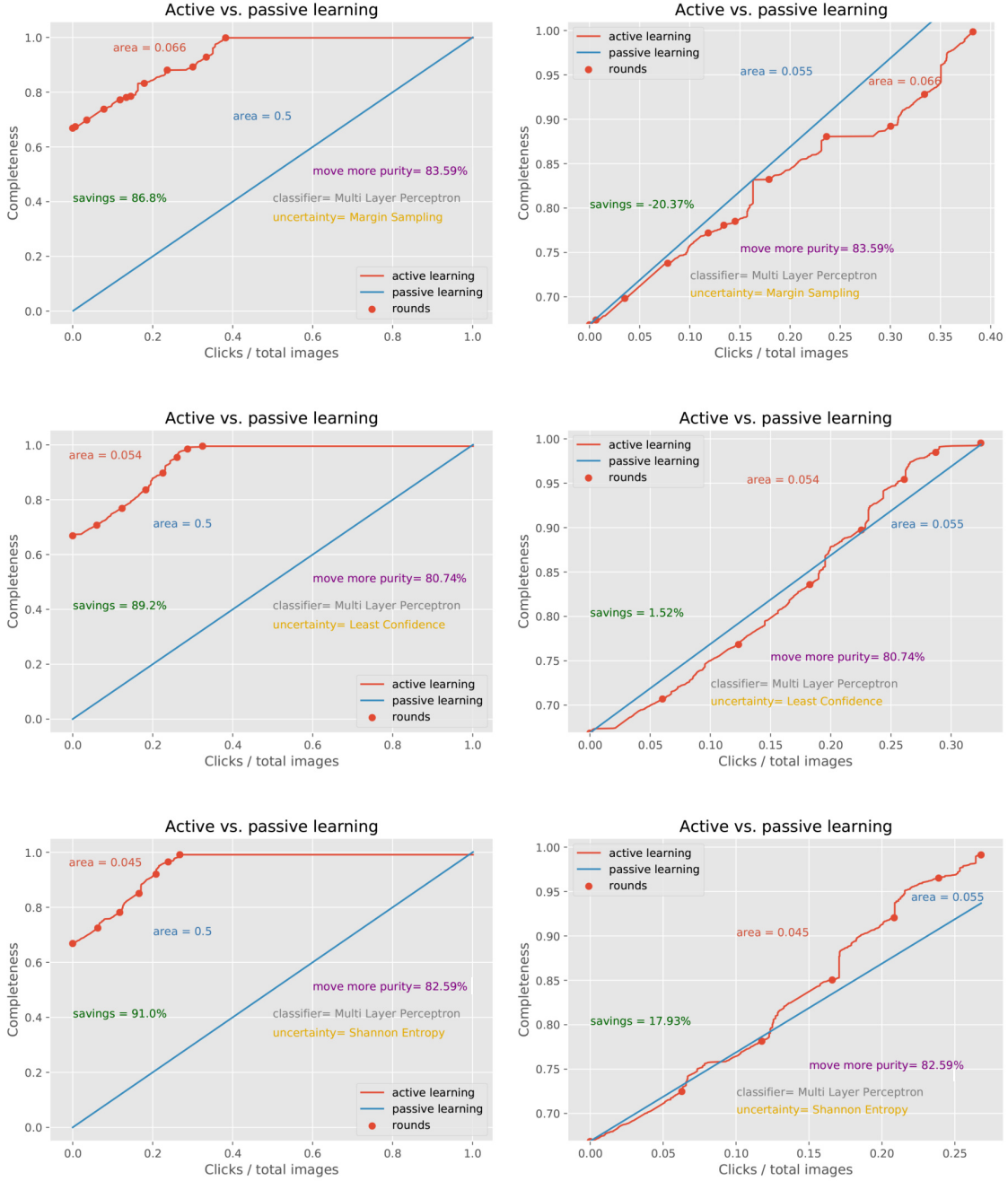


Figure 7: Evaluation of **Multi Layer Perceptron** classifier and *Margin Sampling* (above), *Least Confidence* (middle) and *Shannon Entropy* (below) uncertainty compared to passive learning started from the beginning (left), and after the pre-clustering (right).

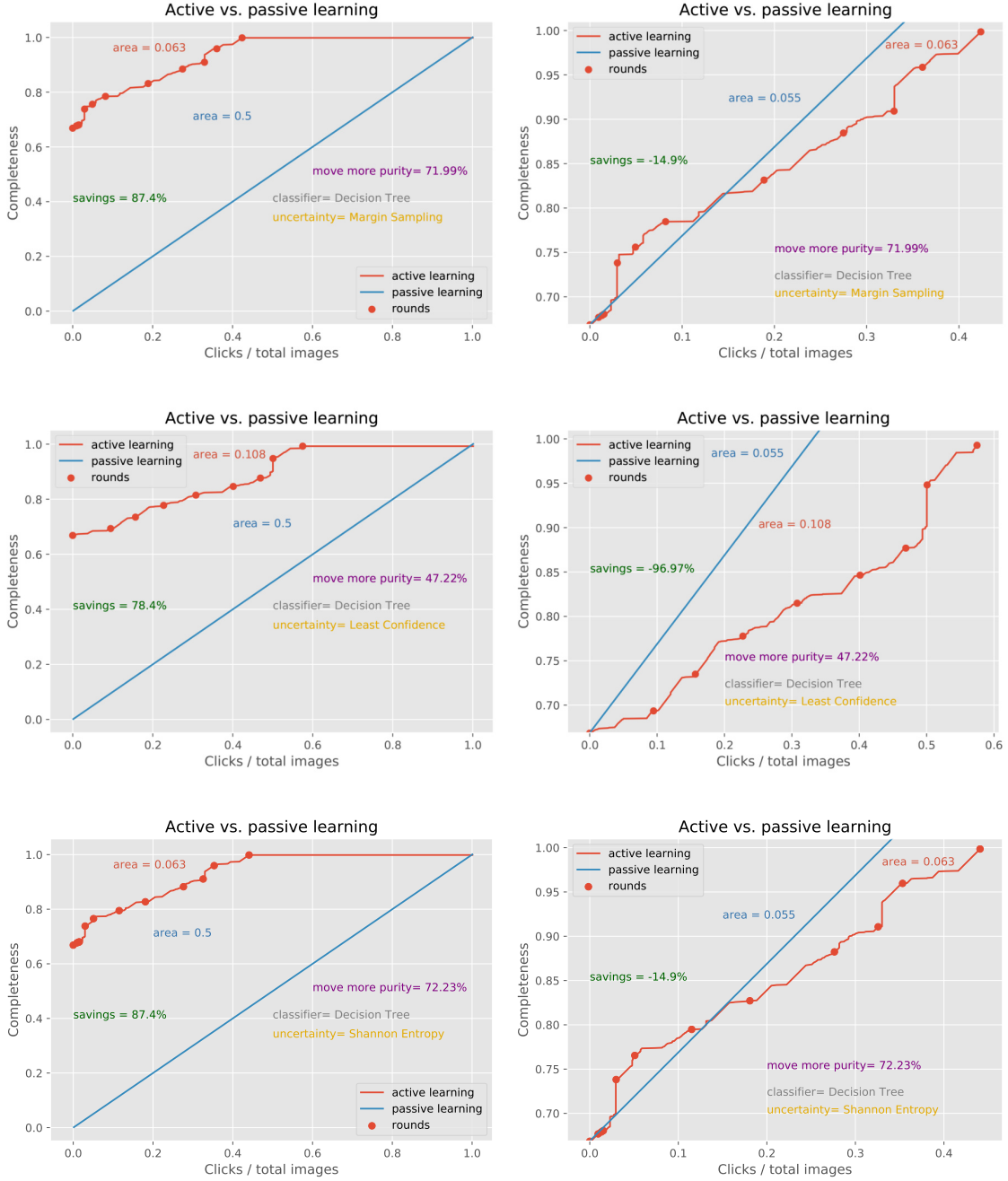


Figure 8: Evaluation of **Decision Tree** classifier and *Margin Sampling* (above), *Least Confidence* (middle) and *Shannon Entropy* (below) uncertainty compared to passive learning started from the beginning (left), and after the pre-clustering (right).