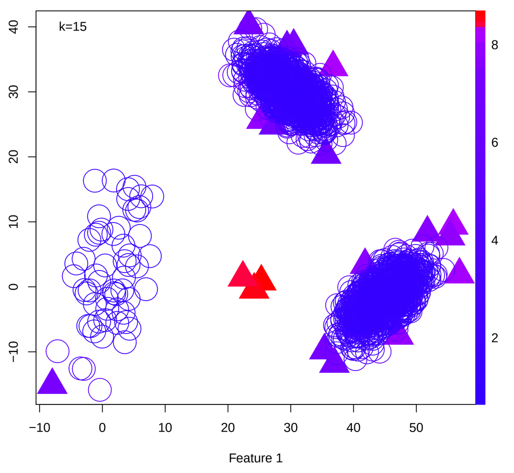


Detecting Interesting Outliers

Active Learning for Anomaly Detection

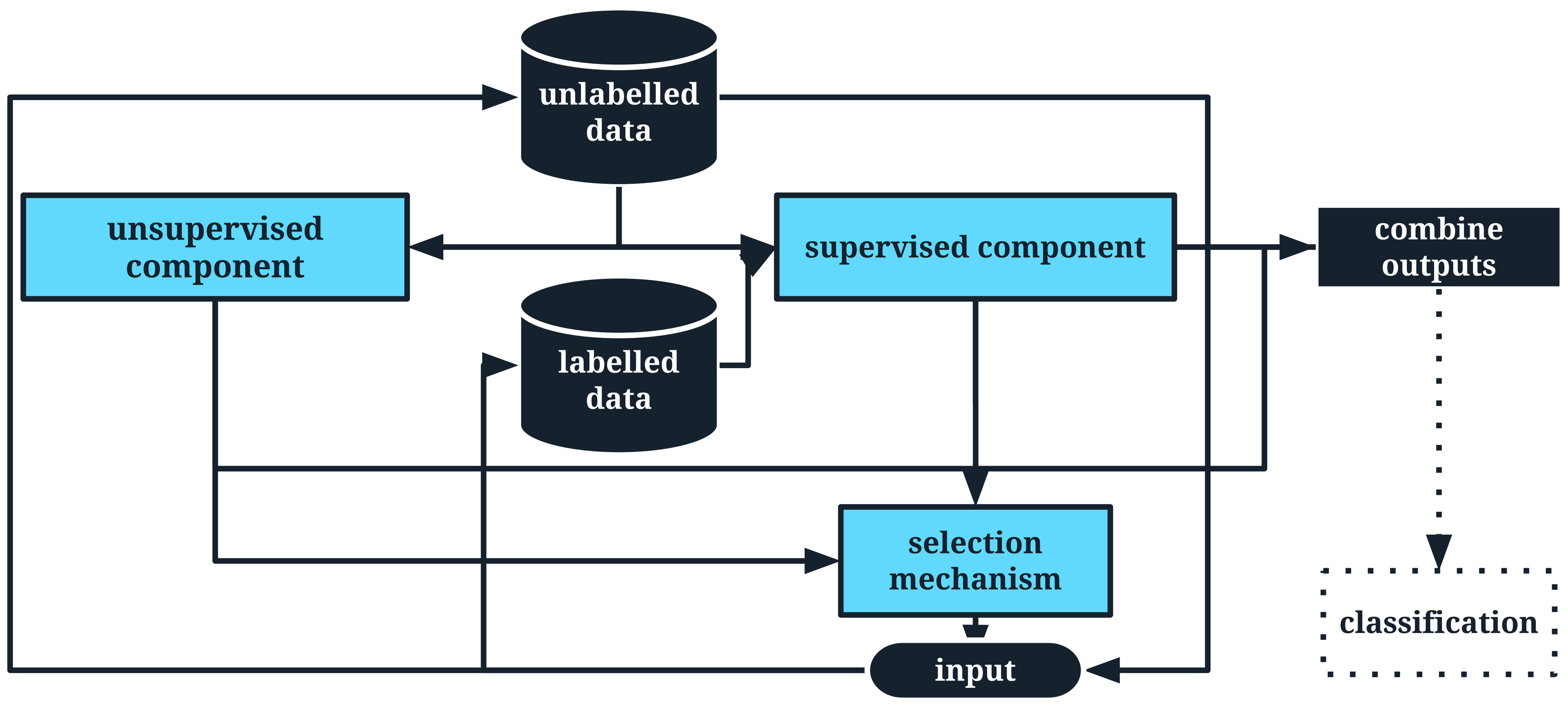
Floris den Hengst, Mark Hoogendoorn
florisdenhengst@gmail.com, mark.hoogendoorn@vu.nl

Contributions: little labels, lots of anomalies

An active learning framework finds anomalies with little expert input. The proposed method outperforms both supervised and unsupervised methods in terms of Precision and Recall and requires less labels to do so for two out of three tested data sets on training data.

Framework

An *unsupervised* method finds promising items (outliers), a *supervised* method separates anomalies from non-anomalies based on expert input. A *selection mechanism* is employed for efficient querying.



Query Mechanism

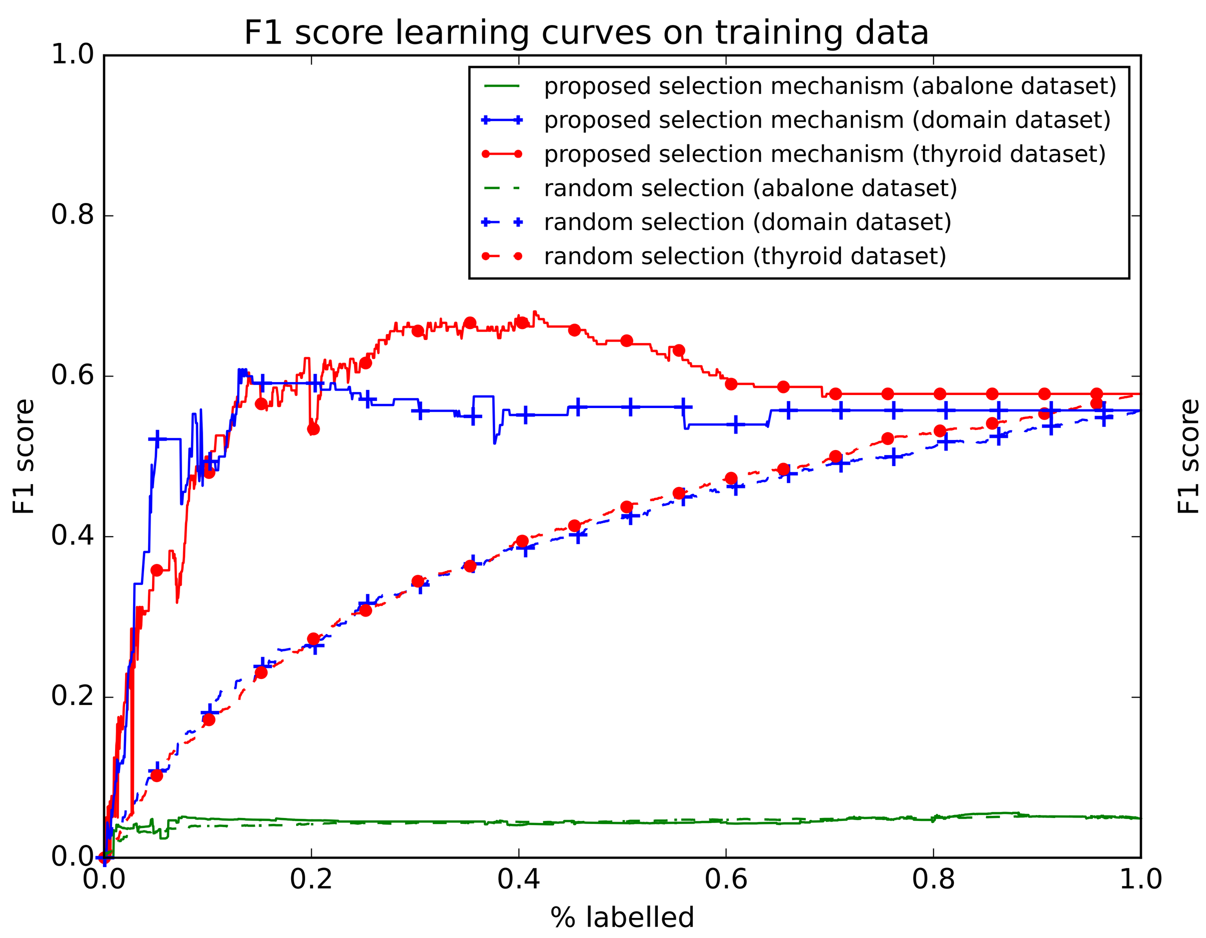
To gradually refine the supervised model without missing out promising regions of data, we alternately query the label for the item that:

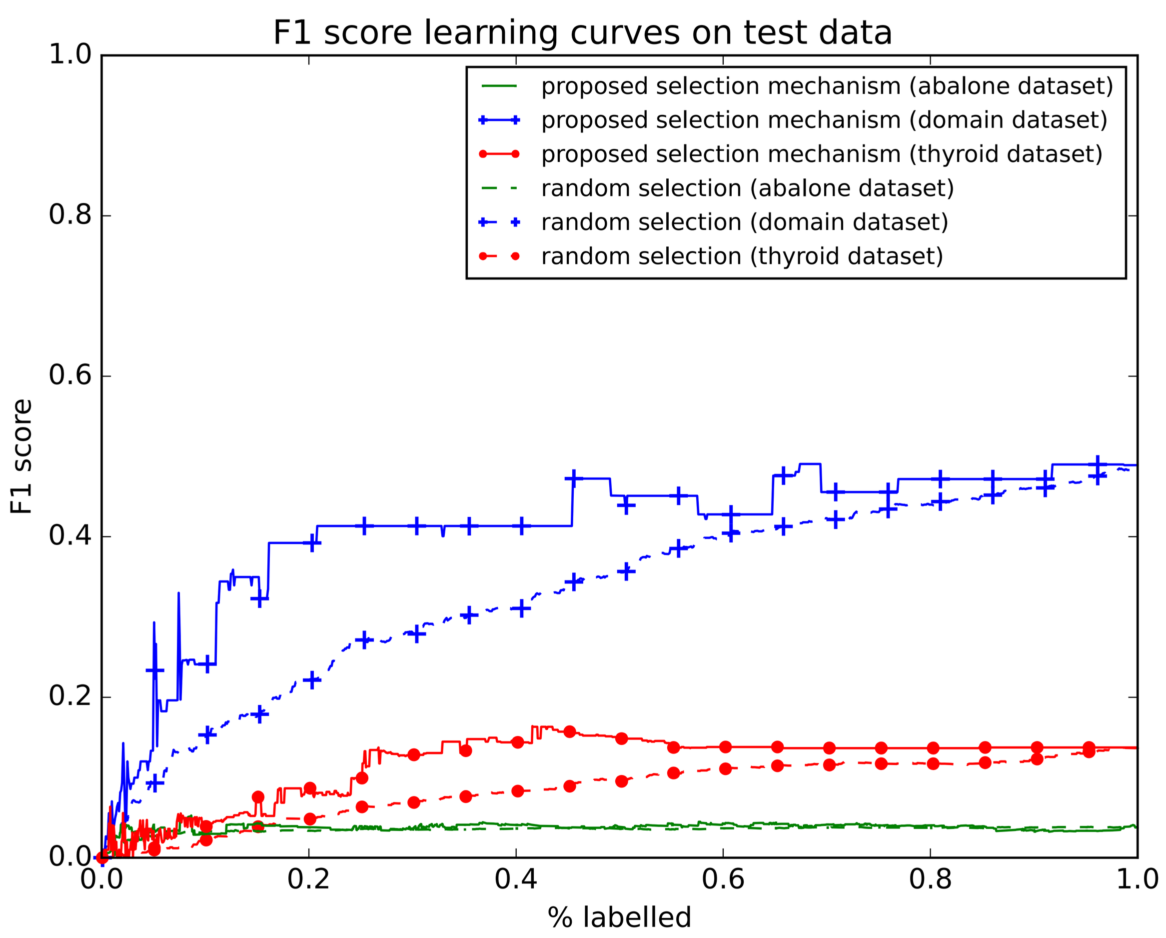
- lies closest to the decision boundary of the supervised component
- is most differently ranked by both components, i.e. by selecting i with maximizes $D(i) = |R_U(i) - R_S(i)|$ for ranks R_U and R_S produced by unsupervised component U and supervised component S.

This is an adaptation of the *interleave* strategy from [2].

Informed Querying Improves Learning

The proposed query mechanism improves learning abilities over random querying for all assessed data sets, both on training and test data. On training data, the number of labels required to achieve the highest F1 can be reduced to .88, .41 and .13 of all labels.





Exploratory Qualities and Generalisability

On training data (*exploratory* setting), our method outperforms the unsupervised method in terms of Precision. On test data (*generalisability*), however, Precision degrades and a a single unsupervised method seems favorable.

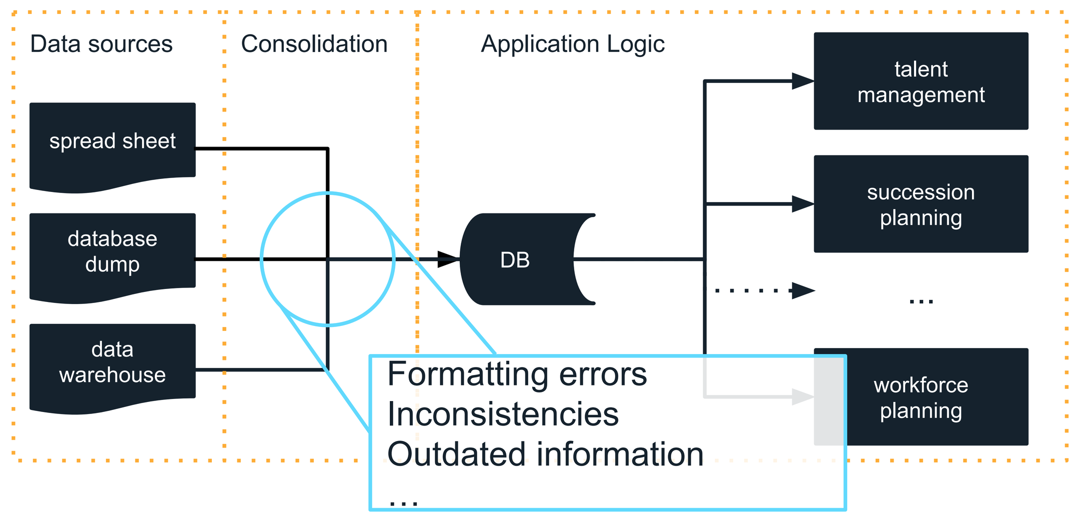
	Novel Method			Unsupervised*	
	$max(Recall)$	at $min(Precision)$	at % labelled	Precision	Data
Abalone	0.775	0.026	69.3%	**0.029**	Training
Thyroid	1.0	**0.403**	78.2%	0.184	Training
Domain	1.0	**0.367**	63.9%	0.038	Training
Abalone	0.633	0.091	66.3%	0.029	Test
Thyroid	0.182	0.013	42.2%	0.184	Test
Domain	0.785	0.591	100.0%	0.038	Test

* $max(Recall)$ is always 1.0

Future work could be aimed at including unsupervised component output in final classification, at replacing the single unsupervised model by an ensemble to find different kinds of outliers better, at including queries for labels in poorly-labeled regions in the selection mechanism.

Context

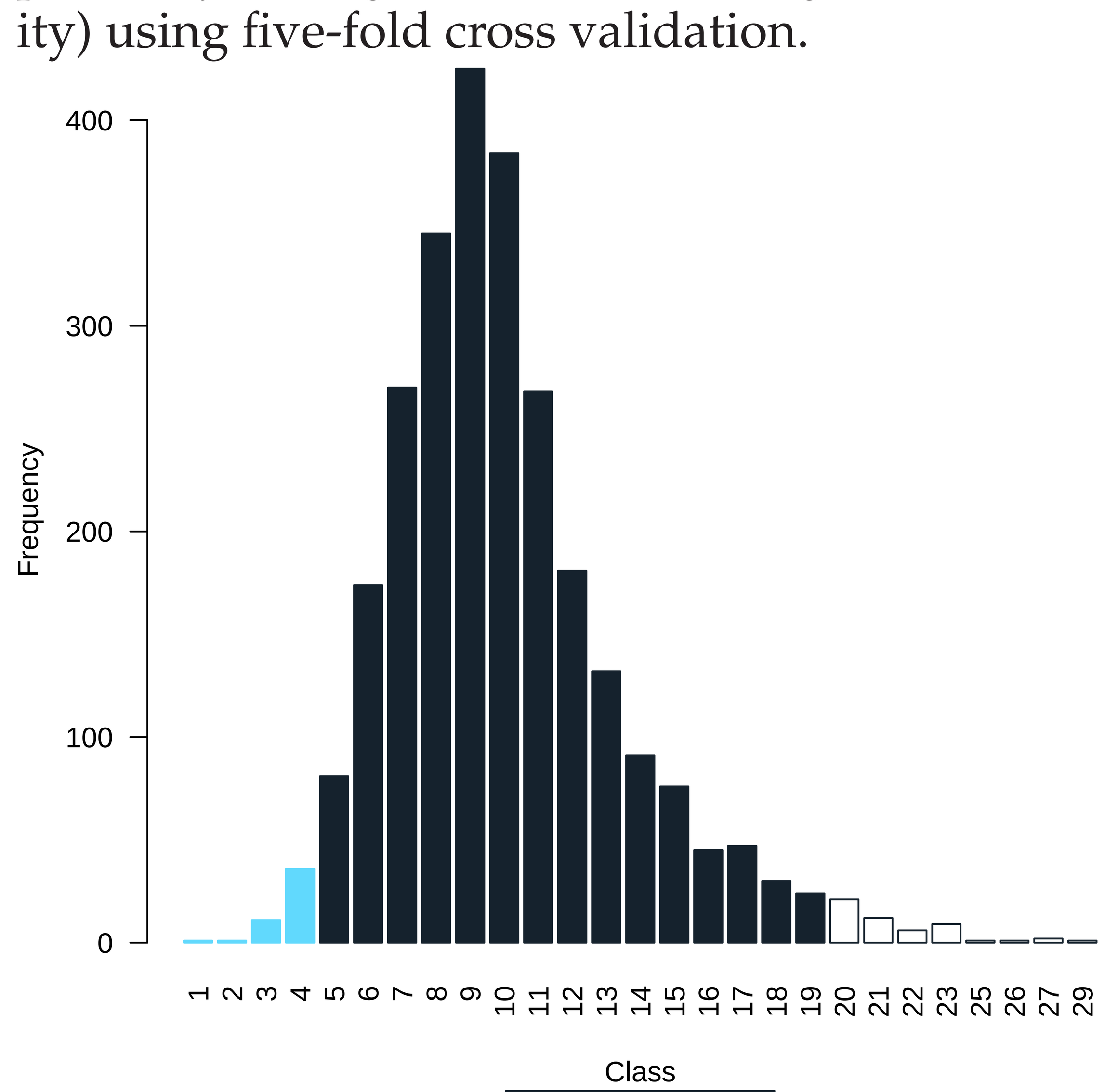
Anomaly detection in a setting where the definition of anomaly is not known up-front and may differ per data set. A *domain expert* is capable of making a distinction between anomalies and non-anomalies.

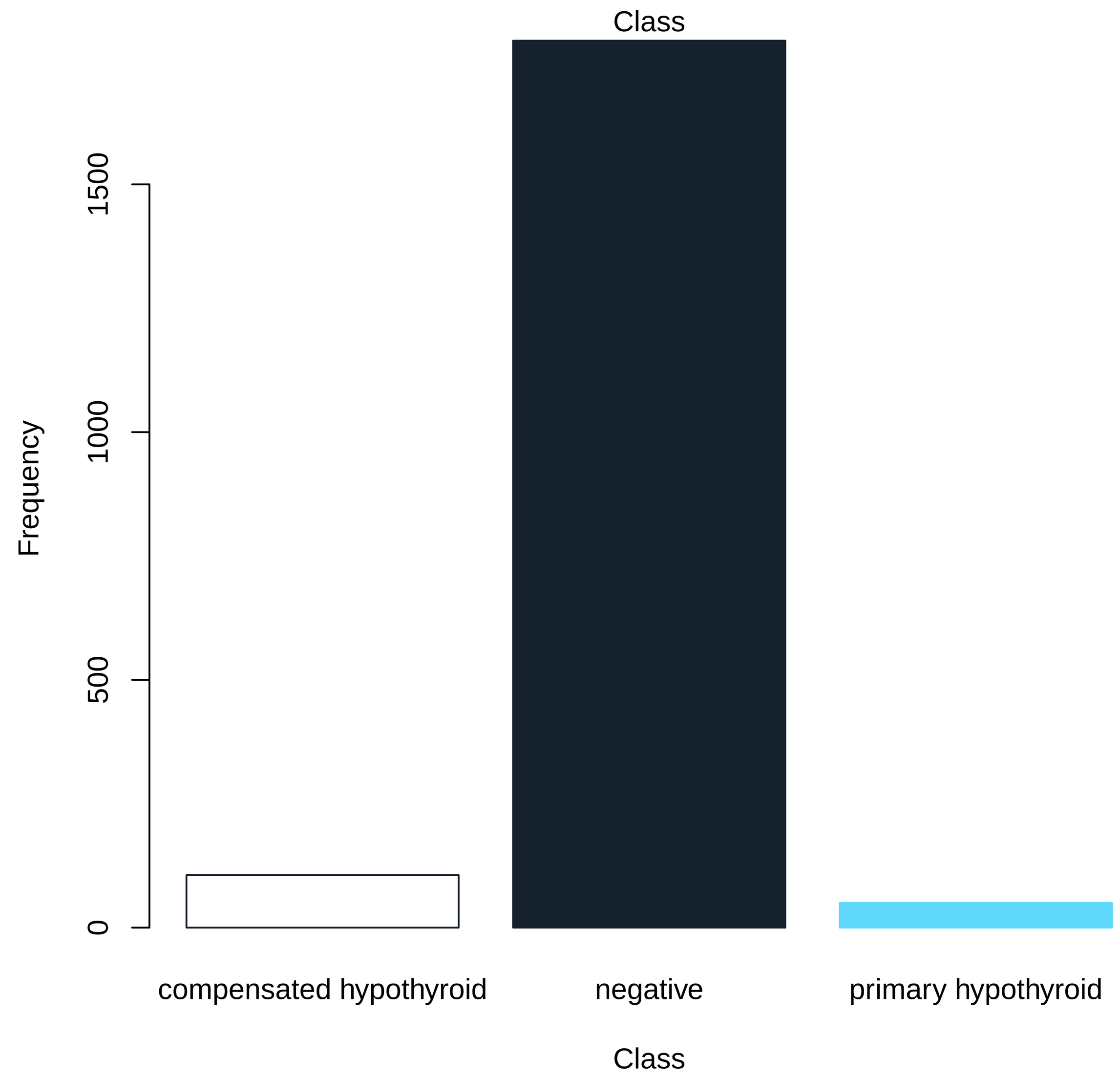


Method & Used Data

In preliminary results, we find the optimally performing supervised and unsupervised methods to be a general SVM and Local Outlier Factor [1] and use these in further experiments. In combining component outputs, we wholly discard unsupervised output to limit project scope.

We measure Precision, Recall and F1 score on two benchmark data sets and a data set from the HR domain. The benchmark data sets contain multiple minority classes, out of which some are labeled as anomaly (highlight in light blue). We perform measurements on training data (exploratory setting) and on test data (generalisability) using five-fold cross validation.





References

[1] Markus M Breunig, Hans-Peter Kriegel, Raymond T Ng, and Jörg Sander. Lof: identifying density-based local outliers. In *ACM Sigmod Record*, volume 29, pages 93–104. ACM, 2000.

[2] Dan Pelleg and Andrew W Moore. Active learning for anomaly and rare-category detection. In *Advances in Neural Information Processing Systems*, pages 1073–1080, 2004.