Jožef Stefan Institute
Ljubljana, Slovenia

Artificial Intelligence
Department for

Active Learning for Automated Visual Inspection of Manufactured Products

Elena Trajkova, Jože M. Rožanec, Paulien Dam, Blaž Fortuna, and Dunja Mladenić

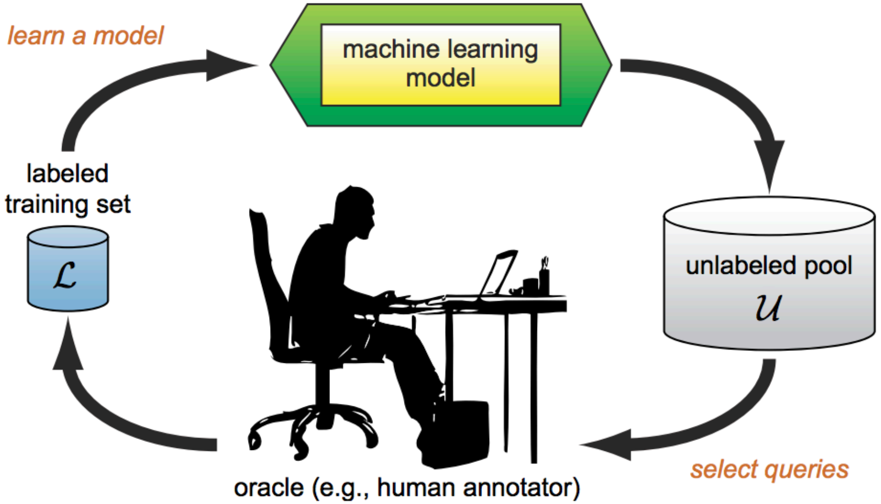
Purpose of the study

- **Automated quality control**
- **Real-world multiclass classification**
- **We compare five ML algorithms for automated defect detection** (Gaussian Näive Bayes, CART Linear SVM, MLP, and kNN)
- **We assess three active learning approaches** (stream-based classifier, pool-based sampling and pool-based sampling considering a query-by-committee strategy)

Related work

- Duan et al. : visual inspection of microdrill bits in printed circuit board production
  - Statistical features and SVM, MLP, kNN
- Gobert et al. : defect detection during metallic powder bed fusion in additive manufacturing.
  - 3D convolutional filters and SVM
- Use of labeled datasets
  - Incoming data exceeds capacity

Active learning



- Three active learning approaches:
  - **stream-based sampling**: receiving unlabeled instances one at a time, immediate decision whether to label the data or not
  - **pool-based sampling** : label most informative instances from pool of unlabeled data
  - **query-by-committee** : retrieving the unlabeled sample with the greatest variance between a set of forecasting models

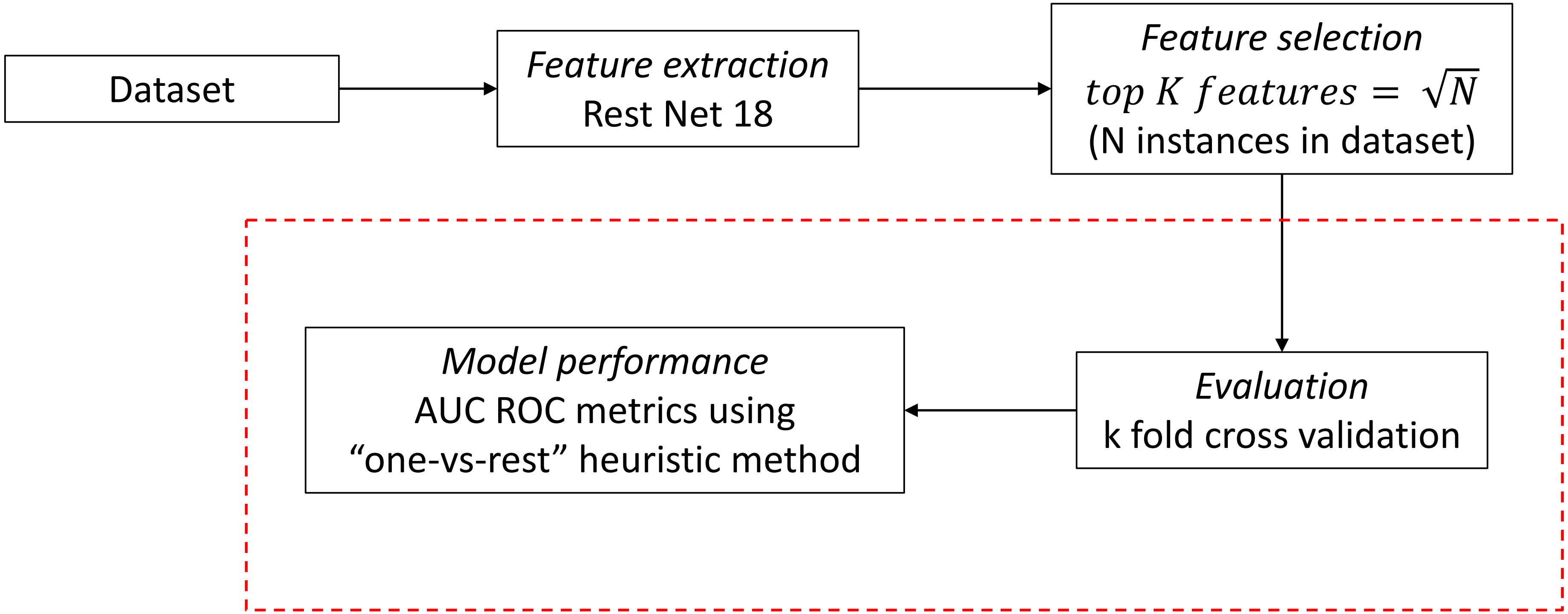
The image was taken from: SETTLES, Burr. Active learning literature survey. 2009.

Use case

- Visual inspection of shavers produced by Philips Consumer Lifestyle BV
  - detect defective printing of a logo on the shavers
- Two types of defects related to the printing quality of the logo : double printing and interrupted printing.
- Three classes of images:
  - good printing (class zero)
  - double printing (class one)
  - interrupted printing (class two)
- Limited labeled dataset

good double print interrupted print

Methodology



Three active learning approaches and five ML algorithms

Experiments

- **k fold cross validation** (k = 10)
- **Evaluate the active learning approaches**:
  - **stream-based classifier** (threshold above the 75th percentile of observed instances)
  - **pool-based sampling** selecting the instances a given model is most uncertain about
  - **pool-based sampling** considering a query-by-committee strategy
- **Metric**: AUC ROC
- **Statistical significance**: Wilcoxon signed-rank test, p value = 0.05

Results

AUC ROC values across the ten cross-validation folds

Active Learning scenario	Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
stream-based	CART	0,8168	0,7828	0,7810	0,7694	0,8196	0,7805	0,7843	0,7970	0,8409	0,7940
	kNN	0,9289	0,9121	0,9174	0,8686	0,9024	0,9000	0,9051	0,8960	0,9282	0,9082
	MLP	**0,9900**	**0,9928**	**0,9846**	**0,9563**	**0,9804**	**0,9807**	**0,9710**	**0,9729**	0,9793	**0,9845**
	Näive Bayes	0,8818	0,8668	0,8819	0,8686	0,8829	0,8899	0,8650	0,8877	0,8864	0,9098
	SVM	0,9752	0,9828	0,9725	*0,9530*	0,9816	0,9720	0,9570	0,9412	0,9824	0,9712
pool-based	CART	0,7584	0,7904	0,7543	0,7468	0,8441	0,7730	0,8044	0,7701	0,7850	0,7412
	kNN	0,9189	0,9149	0,9161	0,8581	0,9055	0,9036	0,8961	0,8910	0,9224	0,9056
	MLP	*0,9892*	*0,9921*	*0,9845*	**0,9563**	0,9790	**0,9803**	**0,9702**	**0,9723**	0,9806	*0,9840*
	Näive Bayes	0,8800	0,8654	0,8809	0,8677	0,8813	0,8895	0,8637	0,8873	0,8850	0,9090
	SVM	0,9752	0,9819	0,9726	0,9518	*0,9806*	0,9712	0,9562	0,9412	*0,9823*	0,9722
query-by-committee		0,9774	0,9824	0,9714	0,9500	0,9723	*0,9726*	*0,9597*	*0,9571*	**0,9830**	0,9734

Results

p-values obtained for the Wilcoxon signed-rank test

Model	Active Learning scenarios		
	stream-based vs. pool-based	stream-based vs. query-by-committee	pool-based vs. query-by-committee
CART	0,0840	**0,0020**	**0,0020**
kNN	0,1309	**0,0020**	**0,0020**
MLP	0,0856	**0,0039**	**0,0039**
Näive Bayes	**0,0020**	**0,0020**	**0,0020**
SVM	0,1824	0,4316	0,6250

Conclusions and future work

- **Conclusions:**
  - Best performance : MLP model
  - No significant difference between using pool-based or stream-based active learning approaches
  - Query-by-committee performs significantly better in all cases, except for the MLPs
- **Future work:**
  - Develop data augmentation techniques
  - Seek statistically significant improvements over time for AL strategies
  - Include explainable artificial intelligence, to aid manual labeling.