## 1 AnoShift - Out-Of-Distribution Anomaly Detection Benchmark - Table 1

- Train data: files  $[year]\_subset.parquet$  with  $year \in \{2006, 2007, 2008, 2009, 2010\}$
- IID data: files  $[year]\_subset\_valid.parquet$  with  $year \in \{2006, 2007, 2008, 2009, 2010\}$
- NEAR data: files [year]\_subset.parquet with  $year \in \{2011, 2012, 2013\}$
- FAR data: files  $[year]\_subset.parquet$  with  $year \in \{2014, 2015\}$
- Scripts for reproducing the results are available in '/baselines\_OOD\_setup'

Table 1: Performance evolution over time for unsupervised methods: IID vs NEAR vs FAR. We report beside the ROC-AUC metric, also the PR-AUC for inliers and PR-AUC for outliers. With bold are the best results per split.

Coc-SVM [10] (train 5%)   76.86 ± 0.06   71.43 ± 0.29   49.57 ± 0.09     IsoForest [7]   86.09 ± 0.54   75.26 ± 4.66   27.16 ± 1.69     ECOD [6]   84.76   44.87   49.19     LOF [2]   91.50 ± 0.88   79.29 ± 3.33   34.96 ± 0.14     SO-GAAL [8]   50.48 ± 1.13   54.55 ± 3.92   49.35 ± 0.51     deepSVDD [9]   92.67 ± 0.44   87.00 ± 1.80   34.53 ± 1.62     LUNAR [4] (train 5%)   11.00 ± 0.25   28.19 ± 0.09     IsoForest [7]   83.66 ± 2.14   52.26 ± 1.18   22.45 ± 0.52     BECD [6]   84.47   22.98   13.78     COPOD [5]   84.47   22.98   13.78     COPOD [5]   84.47   22.98   13.78     COPOD [5]   84.61 ± 0.05   52.48 ± 4.56   10.15 ± 0.10     SO-GAAL [8]   68.52 ± 3.64   43.52 ± 11.62   10.68 ± 24.5     DEST [3] for anomalies   73.76 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     ECOD [6]   84.47   22.98   13.78     COPOD [5]   84.61 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     ECOD [6]   84.47   22.98   13.78     COPOD [5]   84.61 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     ECOD [6]   84.67   22.98   13.78     DOC-SVM [10] (train 5%)   78.91 ± 1.69   29.36 ± 2.58   9.33 ± 0.11     InternalContrastiveLearning [11]   76.96 ± 2.12   27.28 ± 0.59   8.81 ± 0.05     ECOD [6]   78.37   74.48   85.90     COPOD [5]   78.37   74.48   85.90     COPOD [6]   78.38   60.99   92.34 ± 1.26   81.99 ± 0	Type	Unsupervised Baselines	IID	NEAR	FAR
SoForest [7]   86.09 ± 0.54   75.26 ± 4.66   27.16 ± 1.69			ROC-AUC (%) ↑		
COF   2    91.50 ± 0.88   79.29 ± 3.33   34.96 ± 0.14		<b>OC-SVM</b> [10] (train 5%)	$76.86 \pm 0.06$	$71.43 \pm 0.29$	$49.57 \pm 0.09$
COF   2    91.50 ± 0.88   79.29 ± 3.33   34.96 ± 0.14	cal	IsoForest [7]	$86.09 \pm 0.54$	$75.26 \pm 4.66$	$27.16 \pm 1.69$
COF   2    91.50 ± 0.88   79.29 ± 3.33   34.96 ± 0.14	SSİ	<b>ECOD</b> [6]	84.76	44.87	49.19
COF   2    91.50 ± 0.88   79.29 ± 3.33   34.96 ± 0.14	Cla	COPOD [5]		54.24	50.42
Cocord	_	LOF [2]	$91.50 \pm 0.88$	$79.29 \pm 3.33$	$34.96 \pm 0.14$
AE [1] for anomalies   81.00 ± 0.22   44.06 ± 0.57   19.96 ± 0.21     LUNAR [4] (train 5%)   85.75 ± 1.95   49.03 ± 2.57   28.19 ± 0.90     BERT [3] for anomalies   84.86 ± 2.14   52.26 ± 1.18   22.45 ± 0.52     BERT [3] for anomalies   84.86 ± 2.14   52.26 ± 1.18   22.45 ± 0.52     BERT [3] for anomalies   84.54 ± 0.07   86.05 ± 0.25   28.15 ± 0.06      OC-SVM [10] (train 5%)   70.84 ± 0.13   41.38 ± 0.29   15.12 ± 0.04     IsoForest [7]   83.68 ± 3.47   57.06 ± 10.27   9.16 ± 0.18     ECOD [6]   84.47   22.98   13.78     COPOD [5]   84.11 ± 0.96   52.48 ± 4.56   10.15 ± 0.10     SO-GAAL [8]   68.55 ± 5.36   43.52 ± 11.62   10.68 ± 2.42     deepSVDD [9]   82.62 ± 0.52   71.71 ± 4.85   10.02 ± 0.22     AE [1] for anomalies   73.76 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     InternalContrastiveLearning [11]   76.96 ± 2.12   27.28 ± 0.59   8.81 ± 0.05     BERT [3] for anomalies   74.61 ± 0.13   58.94 ± 0.69   8.22 ± 0.02      OC-SVM [10] (train 5%)   74.61 ± 0.13   58.94 ± 0.69   8.22 ± 0.02      PR-AUC outliers (%) ↑     OC-SVM [10] (train 5%)   81.46 ± 2.52   87.13 ± 2.08   78.33 ± 1.41     ECOD [6]   78.37   74.48   85.90     COPOD [5]   78.19   77.99   85.98     LOF [2]   83.86 ± 0.98   92.34 ± 1.26   81.99 ± 0.05     SO-GAAL [8]   70.38 ± 0.28   87.71 ± 0.74   92.67 ± 0.13     deepSVDD [9]   92.65 ± 0.64   94.15 ± 1.05   82.25 ± 0.48     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.90 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.90 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     AE [1] for anomalies   78.90 ± 0.28   72.97 ± 0.38   7		SO-GAAL [8]	$50.48 \pm 1.13$	$54.55 \pm 3.92$	$49.35 \pm 0.51$
InternalContrastiveLearning [11]   84.86 $\pm$ 2.14   52.26 $\pm$ 1.18   22.45 $\pm$ 0.52   BERT [3] for anomalies   84.54 $\pm$ 0.07   86.05 $\pm$ 0.25   28.15 $\pm$ 0.06      PR-AUC inliers (%) \rackford		deepSVDD [9]	$92.67 \pm 0.44$	$87.00 \pm 1.80$	$34.53 \pm 1.62$
InternalContrastiveLearning [11]   84.86 $\pm$ 2.14   52.26 $\pm$ 1.18   22.45 $\pm$ 0.52   BERT [3] for anomalies   84.54 $\pm$ 0.07   86.05 $\pm$ 0.25   28.15 $\pm$ 0.06      PR-AUC inliers (%) \rackford	ф	AE [1] for anomalies	$81.00 \pm 0.22$	$44.06 \pm 0.57$	$19.96 \pm 0.21$
InternalContrastiveLearning [11]   84.86 $\pm$ 2.14   52.26 $\pm$ 1.18   22.45 $\pm$ 0.52   BERT [3] for anomalies   84.54 $\pm$ 0.07   86.05 $\pm$ 0.25   28.15 $\pm$ 0.06      PR-AUC inliers (%) \rackford	Ğ	LUNAR [4] (train 5%)	$85.75 \pm 1.95$	$49.03 \pm 2.57$	$28.19 \pm 0.90$
$ \begin{array}{ c c c c c } \hline \textbf{OC-SVM} [10] (train 5\%) & \hline & PR-AUC inliers (\%) \uparrow \\ \hline \textbf{IsoForest} [7] & 83.68 \pm 3.47 & 57.06 \pm 10.27 & 9.16 \pm 0.18 \\ \hline \textbf{ECOD} [6] & 84.47 & 22.98 & 13.78 \\ \hline \textbf{COPOD} [5] & 87.86 & 29.25 & 14.55 \\ \hline \textbf{LOF} [2] & 84.11 \pm 0.96 & 52.48 \pm 4.56 & 10.15 \pm 0.10 \\ \hline \textbf{SO-GAAL} [8] & 58.65 \pm 5.36 & 43.52 \pm 11.62 & 10.68 \pm 2.42 \\ \hline \textbf{deepSVDD} [9] & 82.62 \pm 0.52 & 71.71 \pm 4.85 & 10.02 \pm 0.22 \\ \hline \textbf{AE} [1] \textbf{ for anomalies} & 73.76 \pm 0.09 & 26.16 \pm 0.15 & 8.51 \pm 0.01 \\ \hline \textbf{LUNAR} [4] (train 5\%) & 78.91 \pm 1.69 & 29.36 \pm 2.58 & 9.33 \pm 0.11 \\ \hline \textbf{InternalContrastive Learning} [11] & 76.96 \pm 2.12 & 27.28 \pm 0.59 & 8.81 \pm 0.05 \\ \hline \textbf{BERT} [3] \textbf{ for anomalies} & 74.61 \pm 0.13 & 58.94 \pm 0.69 & 8.22 \pm 0.02 \\ \hline \hline \textbf{OC-SVM} [10] (train 5\%) & 67.94 \pm 0.21 & 85.70 \pm 0.16 & 87.27 \pm 0.02 \\ \hline \textbf{IsoForest} [7] & 81.46 \pm 2.52 & 87.13 \pm 2.08 & 78.33 \pm 1.41 \\ \hline \textbf{ECOD} [6] & 78.37 & 74.48 & 85.90 \\ \hline \textbf{COPOD} [5] & 83.86 \pm 0.98 & 92.34 \pm 1.26 & 81.99 \pm 0.05 \\ \hline \textbf{SO-GAAL} [8] & 70.38 \pm 0.28 & 87.71 \pm 0.74 & 92.67 \pm 0.13 \\ \hline \textbf{deepSVDD} [9] & 92.65 \pm 0.64 & 94.15 \pm 1.05 & 82.25 \pm 0.48 \\ \hline \textbf{AE} [1] \textbf{ for anomalies} & 78.99 \pm 0.28 & 72.97 \pm 0.38 & 75.71 \pm 0.05 \\ \hline \textbf{LUNAR} [4] (train 5\%) & 88.01 \pm 1.03 & 80.91 \pm 0.62 & 79.45 \pm 0.30 \\ \hline \textbf{InternalContrastive Learning} [11] & 89.08 \pm 0.87 & 81.93 \pm 0.39 & 77.55 \pm 0.50 \\ \hline \end{array}$			$84.86 \pm 2.14$	$52.26 \pm 1.18$	$22.45 \pm 0.52$
CC-SVM [10] (train 5%)   70.84 ± 0.13   41.38 ± 0.29   15.12 ± 0.04     IsoForest [7]   83.68 ± 3.47   57.06 ± 10.27   9.16 ± 0.18     ECOD [6]   84.47   22.98   13.78     COPOD [5]   87.86   29.25   14.55     LOF [2]   84.11 ± 0.96   52.48 ± 4.56   10.15 ± 0.10     SO-GAAL [8]   58.65 ± 5.36   43.52 ± 11.62   10.68 ± 2.42     deepSVDD [9]   82.62 ± 0.52   71.71 ± 4.85   10.02 ± 0.22     AE [1] for anomalies   73.76 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     InternalContrastiveLearning [11]   76.96 ± 2.12   27.28 ± 0.59   8.81 ± 0.05     BERT [3] for anomalies   74.61 ± 0.13   58.94 ± 0.69   8.22 ± 0.02      OC-SVM [10] (train 5%)   74.61 ± 0.13   58.94 ± 0.69   8.22 ± 0.02      FOR SUM [10] (train 5%)   67.94 ± 0.21   85.70 ± 0.16   87.27 ± 0.02     IsoForest [7]   81.46 ± 2.52   87.13 ± 2.08   78.33 ± 1.41     ECOD [6]   78.37   74.48   85.90     COPOD [5]   78.19   77.99   85.98     LOF [2]   83.86 ± 0.98   92.34 ± 1.26   81.99 ± 0.05      SO-GAAL [8]   70.38 ± 0.28   87.71 ± 0.74   92.67 ± 0.13     deepSVDD [9]   92.65 ± 0.64   94.15 ± 1.05   82.25 ± 0.48     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     LUNAR [4] (train 5%)   88.01 ± 1.03   80.91 ± 0.62   79.45 ± 0.30     InternalContrastiveLearning [11]   89.08 ± 0.87   81.93 ± 0.39   77.55 ± 0.50     COPOD [5]   88.01 ± 1.03   80.91 ± 0.62   79.45 ± 0.30     InternalContrastiveLearning [11]   89.08 ± 0.87   81.93 ± 0.39   77.55 ± 0.50     COPOD [5]   78.19   77.55 ± 0.50     COPOD [5]   78.19   77.99   77.95 ± 0.30     COPOD [5]   78.19   77.99   77.99   77.55 ± 0.50     COPOD [5]   78.19   77.99   77.99   77.55 ± 0.50     COPOD [5]   78.19   77.99   77.99   77.55 ± 0.50     COPOD [5]   78.19   77.55 ± 0.50     COPOD [5]   78.19   77.99   77.55 ± 0.50     COPOD [5]		BERT [3] for anomalies	$84.54 \pm 0.07$	$86.05 \pm 0.25$	$28.15 \pm 0.06$
Solution			PR-AUC inliers (%) ↑		
Solution		<b>OC-SVM</b> [10] (train 5%)	$70.84 \pm 0.13$	$41.38 \pm 0.29$	$15.12 \pm 0.04$
SO-GAAL [8]   58.65 ± 5.36   43.52 ± 11.62   10.68 ± 2.42     deepSVDD [9]   82.62 ± 0.52   71.71 ± 4.85   10.02 ± 0.22     AE [1] for anomalies   73.76 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     LUNAR [4] (train 5%)   78.91 ± 1.69   29.36 ± 2.58   9.33 ± 0.11     InternalContrastiveLearning [11]   76.96 ± 2.12   27.28 ± 0.59   8.81 ± 0.05     BERT [3] for anomalies   74.61 ± 0.13   58.94 ± 0.69   8.22 ± 0.02      OC-SVM [10] (train 5%)   67.94 ± 0.21   85.70 ± 0.16   87.27 ± 0.02     IsoForest [7]   81.46 ± 2.52   87.13 ± 2.08   78.33 ± 1.41     ECOD [6]   78.37   74.48   85.90     COPOD [5]   78.19   77.99   85.98     LOF [2]   83.86 ± 0.98   92.34 ± 1.26   81.99 ± 0.05     SO-GAAL [8]   70.38 ± 0.28   87.71 ± 0.74   92.67 ± 0.13     deepSVDD [9]   92.65 ± 0.64   94.15 ± 1.05   82.25 ± 0.48     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     LUNAR [4] (train 5%)   88.01 ± 1.03   80.91 ± 0.62   79.45 ± 0.30     InternalContrastiveLearning [11]   89.08 ± 0.87   81.93 ± 0.39   77.55 ± 0.50	cal		$83.68 \pm 3.47$	$57.06 \pm 10.27$	$9.16 \pm 0.18$
SO-GAAL [8]   58.65 ± 5.36   43.52 ± 11.62   10.68 ± 2.42     deepSVDD [9]   82.62 ± 0.52   71.71 ± 4.85   10.02 ± 0.22     AE [1] for anomalies   73.76 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     LUNAR [4] (train 5%)   78.91 ± 1.69   29.36 ± 2.58   9.33 ± 0.11     InternalContrastiveLearning [11]   76.96 ± 2.12   27.28 ± 0.59   8.81 ± 0.05     BERT [3] for anomalies   74.61 ± 0.13   58.94 ± 0.69   8.22 ± 0.02      OC-SVM [10] (train 5%)   67.94 ± 0.21   85.70 ± 0.16   87.27 ± 0.02     IsoForest [7]   81.46 ± 2.52   87.13 ± 2.08   78.33 ± 1.41     ECOD [6]   78.37   74.48   85.90     COPOD [5]   78.19   77.99   85.98     LOF [2]   83.86 ± 0.98   92.34 ± 1.26   81.99 ± 0.05     SO-GAAL [8]   70.38 ± 0.28   87.71 ± 0.74   92.67 ± 0.13     deepSVDD [9]   92.65 ± 0.64   94.15 ± 1.05   82.25 ± 0.48     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     LUNAR [4] (train 5%)   88.01 ± 1.03   80.91 ± 0.62   79.45 ± 0.30     InternalContrastiveLearning [11]   89.08 ± 0.87   81.93 ± 0.39   77.55 ± 0.50	SSİ	<b>ECOD</b> [6]	84.47	22.98	13.78
SO-GAAL [8]   58.65 ± 5.36   43.52 ± 11.62   10.68 ± 2.42     deepSVDD [9]   82.62 ± 0.52   71.71 ± 4.85   10.02 ± 0.22     AE [1] for anomalies   73.76 ± 0.09   26.16 ± 0.15   8.51 ± 0.01     LUNAR [4] (train 5%)   78.91 ± 1.69   29.36 ± 2.58   9.33 ± 0.11     InternalContrastiveLearning [11]   76.96 ± 2.12   27.28 ± 0.59   8.81 ± 0.05     BERT [3] for anomalies   74.61 ± 0.13   58.94 ± 0.69   8.22 ± 0.02      OC-SVM [10] (train 5%)   67.94 ± 0.21   85.70 ± 0.16   87.27 ± 0.02     IsoForest [7]   81.46 ± 2.52   87.13 ± 2.08   78.33 ± 1.41     ECOD [6]   78.37   74.48   85.90     COPOD [5]   78.19   77.99   85.98     LOF [2]   83.86 ± 0.98   92.34 ± 1.26   81.99 ± 0.05     SO-GAAL [8]   70.38 ± 0.28   87.71 ± 0.74   92.67 ± 0.13     deepSVDD [9]   92.65 ± 0.64   94.15 ± 1.05   82.25 ± 0.48     AE [1] for anomalies   78.99 ± 0.28   72.97 ± 0.38   75.71 ± 0.05     LUNAR [4] (train 5%)   88.01 ± 1.03   80.91 ± 0.62   79.45 ± 0.30     InternalContrastiveLearning [11]   89.08 ± 0.87   81.93 ± 0.39   77.55 ± 0.50	Ila	COPOD [5]	87.86	29.25	14.55
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	0	<b>LOF</b> [2]	$84.11 \pm 0.96$	$52.48 \pm 4.56$	$10.15 \pm 0.10$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SO-GAAL [8]	$58.65 \pm 5.36$	$43.52 \pm 11.62$	$10.68 \pm 2.42$
InternalContrastiveLearning [11]   $76.96 \pm 2.12   27.28 \pm 0.59   8.81 \pm 0.05  $   BERT [3] for anomalies   $74.61 \pm 0.13   58.94 \pm 0.69   8.22 \pm 0.02  $   PR-AUC outliers (%) \(\gamma\) \(\frac{1}{2}\)		deepSVDD [9]	$82.62 \pm 0.52$	$71.71 \pm 4.85$	$10.02 \pm 0.22$
InternalContrastiveLearning [11]   $76.96 \pm 2.12   27.28 \pm 0.59   8.81 \pm 0.05  $   BERT [3] for anomalies   $74.61 \pm 0.13   58.94 \pm 0.69   8.22 \pm 0.02  $   PR-AUC outliers (%) \(\gamma\) \(\frac{1}{2}\)	eр	AE [1] for anomalies	$73.76 \pm 0.09$		$8.51 \pm 0.01$
InternalContrastiveLearning [11]   $76.96 \pm 2.12   27.28 \pm 0.59   8.81 \pm 0.05  $   BERT [3] for anomalies   $74.61 \pm 0.13   58.94 \pm 0.69   8.22 \pm 0.02  $   PR-AUC outliers (%) \(\gamma\) \(\frac{1}{2}\)	De				
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$					
OC-SVM [10] (train 5%) $67.94 \pm 0.21$ $85.70 \pm 0.16$ $87.27 \pm 0.02$ IsoForest [7] $81.46 \pm 2.52$ $87.13 \pm 2.08$ $78.33 \pm 1.41$ ECOD [6] $78.37$ $74.48$ $85.90$ COPOD [5] $78.19$ $77.99$ $85.98$ LOF [2] $83.86 \pm 0.98$ $92.34 \pm 1.26$ $81.99 \pm 0.05$ SO-GAAL [8] $70.38 \pm 0.28$ $87.71 \pm 0.74$ $92.67 \pm 0.13$ deepSVDD [9] $92.65 \pm 0.64$ $94.15 \pm 1.05$ $82.25 \pm 0.48$ AE [1] for anomalies $78.99 \pm 0.28$ $72.97 \pm 0.38$ $75.71 \pm 0.05$ LUNAR [4] (train 5%) $88.01 \pm 1.03$ $80.91 \pm 0.62$ $79.45 \pm 0.30$ InternalContrastiveLearning [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$		BERT [3] for anomalies	$74.61 \pm 0.13$	$58.94 \pm 0.69$	$8.22 \pm 0.02$
Some tensor of the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [11]   Since the contractive Learning [12]   Since the contractive Learning [12]   Since the contractiv			PR-AUC outliers (%) ↑		
LOF [2] $83.86 \pm 0.98$ $92.34 \pm 1.26$ $81.99 \pm 0.05$ SO-GAAL [8] $70.38 \pm 0.28$ $87.71 \pm 0.74$ $92.67 \pm 0.13$ deepSVDD [9] $92.65 \pm 0.64$ $94.15 \pm 1.05$ $82.25 \pm 0.48$ AE [1] for anomalies $78.99 \pm 0.28$ $72.97 \pm 0.38$ $75.71 \pm 0.05$ LUNAR [4] (train 5%) $88.01 \pm 1.03$ $80.91 \pm 0.62$ $79.45 \pm 0.30$ InternalContrastiveLearning [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$		<b>OC-SVM</b> [10] (train 5%)	$67.94 \pm 0.21$	$85.70 \pm 0.16$	$87.27 \pm 0.02$
LOF [2] $83.86 \pm 0.98$ $92.34 \pm 1.26$ $81.99 \pm 0.05$ SO-GAAL [8] $70.38 \pm 0.28$ $87.71 \pm 0.74$ $92.67 \pm 0.13$ deepSVDD [9] $92.65 \pm 0.64$ $94.15 \pm 1.05$ $82.25 \pm 0.48$ AE [1] for anomalies $78.99 \pm 0.28$ $72.97 \pm 0.38$ $75.71 \pm 0.05$ LUNAR [4] (train 5%) $88.01 \pm 1.03$ $80.91 \pm 0.62$ $79.45 \pm 0.30$ InternalContrastiveLearning [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$	cal	IsoForest [7]	$81.46 \pm 2.52$	$87.13 \pm 2.08$	$78.33 \pm 1.41$
LOF [2] $83.86 \pm 0.98$ $92.34 \pm 1.26$ $81.99 \pm 0.05$ SO-GAAL [8] $70.38 \pm 0.28$ $87.71 \pm 0.74$ $92.67 \pm 0.13$ deepSVDD [9] $92.65 \pm 0.64$ $94.15 \pm 1.05$ $82.25 \pm 0.48$ AE [1] for anomalies $78.99 \pm 0.28$ $72.97 \pm 0.38$ $75.71 \pm 0.05$ LUNAR [4] (train 5%) $88.01 \pm 1.03$ $80.91 \pm 0.62$ $79.45 \pm 0.30$ InternalContrastiveLearning [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$	SSi	<b>ECOD</b> [6]	78.37	74.48	85.90
LOF [2] $83.86 \pm 0.98$ $92.34 \pm 1.26$ $81.99 \pm 0.05$ SO-GAAL [8] $70.38 \pm 0.28$ $87.71 \pm 0.74$ $92.67 \pm 0.13$ deepSVDD [9] $92.65 \pm 0.64$ $94.15 \pm 1.05$ $82.25 \pm 0.48$ AE [1] for anomalies $78.99 \pm 0.28$ $72.97 \pm 0.38$ $75.71 \pm 0.05$ LUNAR [4] (train 5%) $88.01 \pm 1.03$ $80.91 \pm 0.62$ $79.45 \pm 0.30$ InternalContrastiveLearning [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$	Zla	COPOD [5]	78.19	77.99	
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	_	LOF [2]	$83.86 \pm 0.98$	$92.34 \pm 1.26$	$81.99 \pm 0.05$
$ \begin{array}{llllllllllllllllllllllllllllllllllll$		SO-GAAL [8]	$70.38\pm0.28$	$87.71 \pm 0.74$	$92.67 \pm 0.13$
<b>InternalContrastiveLearning</b> [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$	Deep				
<b>InternalContrastiveLearning</b> [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$					
<b>InternalContrastiveLearning</b> [11] $89.08 \pm 0.87$ $81.93 \pm 0.39$ $77.55 \pm 0.50$					
<b>BERT [3] for anomalies</b> $89.83 \pm 0.07$ <b>95.96</b> $\pm 0.06$ $78.38 \pm 0.02$					
		BERT [3] for anomalies	$89.83 \pm 0.07$	$95.96 \pm 0.06$	$78.38 \pm 0.02$

## 2 AnoShift - In-Distribution Anomaly Detection Benchmark - Table 2

- Train data: files [year]\_subset.parquet with  $year \in \{2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015\}$
- Test data: files  $[year]\_subset\_valid.parquet$  with  $year \in \{2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015\}$
- Scripts for reproducing the results are available in '/baselines\_ID\_setup' (set full\_set=1)

Table 2: Performance in the ID setup. We report beside the ROC-AUC metric, also the PR-AUC for inliers and PR-AUC for outliers. With bold are the best results per split.

Type	Unsupervised Baselines	ROC-AUC (%) $\uparrow$	
	<b>OC-SVM</b> [10] (train 5%)	$68.73 \pm 6.09$	
Classical	<b>IsoForest</b> [7] (train 5%)	$81.27 \pm 0.90$	
	<b>ECOD</b> [6]	79.41	
	COPOD [5]	80.89	
	LOF [2]	$87.61 \pm 1.50$	
	SO-GAAL [8]	$49.90 \pm 0.58$	
	deepSVDD [9]	$88.24 \pm 0.58$	
Deep	<b>AE</b> [1] for anomalies	$64.08 \pm 0.23$	
De	LUNAR [4] (train 5%)	$78.53 \pm 0.21$	
	InternalContrastiveLearning [11]	$66.99 \pm 3.59$	
	BERT [3] for anomalies	$79.62 \pm 0.77$	
		PR-AUC inliers (%) ↑	
	<b>OC-SVM</b> [10] (train 5%)	$50.17 \pm 4.67$	
Classical	<b>IsoForest</b> [7] (train 5%)	$68.46 \pm 0.77$	
	<b>ECOD</b> [6]	69.31	
Cla	COPOD [5]	73.17	
	LOF [2]	$70.79 \pm 1.86$	
	SO-GAAL [8]	$39.41 \pm 9.75$	
	deepSVDD [9]	$76.59 \pm 2.13$	
Deep	AE [1] for anomalies	$47.56 \pm 0.17$	
De	LUNAR [4] (train 5%)	$63.31 \pm 0.51$	
	InternalContrastiveLearning [11]	$50.82 \pm 3.54$	
	BERT [3] for anomalies	$63.26 \pm 0.59$	
		PR-AUC outliers (%) ↑	
	<b>OC-SVM</b> [10] (train 5%)	$74.67 \pm 3.62$	
cal	<b>IsoForest</b> [7] (train 5%)	$85.47 \pm 0.70$	
Classical	<b>ECOD</b> [6]	80.73	
	COPOD [5]	81.47	
	LOF [2]	$89.97 \pm 1.09$	
Deep	SO-GAAL [8]	$79.89 \pm 0.18$	
	deepSVDD [9]	$90.56 \pm 0.77$	
	<b>AE</b> [1] for anomalies	$78.07 \pm 0.05$	
	LUNAR [4] (train 5%)	$88.42 \pm 0.14$	
	InternalContrastiveLearning [11]	$80.1 \pm 9.63$	
	BERT [3] for anomalies	$84.77 \pm 1.70$	

## 3 AnoShift - In-Distribution Anomaly Detection Benchmark (years 2006-2010) - Table 3

- Train data: files [year]\_subset.parquet with  $year \in \{2006, 2007, 2008, 2009, 2010\}$
- Test data: files  $[year]\_subset\_valid.parquet$  with  $year \in \{2006, 2007, 2008, 2009, 2010\}$
- Scripts for reproducing the results are available in '/baselines\_ID\_setup' (set full\_set=0)

Table 3: Performance in the ID setup (years corresponding to our original IID split). We report beside the ROC-AUC metric, also the PR-AUC for inliers and PR-AUC for outliers. With bold are the best results per split.

Туре	Unsupervised Baselines	ROC-AUC (%)↑
Classical	OC-SVM [10] (train 5%)	$76.86 \pm 0.06$
	IsoForest [7]	$86.09 \pm 0.54$
	ECOD [6]	84.76
ひ	COPOD [5] LOF [2]	$85.62$ $91.50 \pm 0.88$
	SO-GAAL [8]	$50.48 \pm 1.13$
•	deepSVDD [9]	$92.67 \pm 0.44$
Deep	AE [1] for anomalies	$81.00 \pm 0.22$
Q	LUNAR [4] (train 5%)	$85.75 \pm 1.95$
	InternalContrastiveLearning [11] BERT [3] for anomalies	$84.86 \pm 2.14$ $84.54 \pm 0.07$
	DEKT [5] for anomanes	
		PR-AUC inliers (%) ↑
	<b>OC-SVM</b> [10] (train 5%)	$70.84 \pm 0.13$
Classical	IsoForest [7]	$83.68 \pm 3.47$
SSi	<b>ECOD</b> [6]	84.47
Cla	COPOD [5]	87.86
	LOF [2]	$84.11 \pm 0.96$
	SO-GAAL [8]	$58.65 \pm 5.36$
	deepSVDD [9]	$82.62 \pm 0.52$
Deep	AE [1] for anomalies	$73.76 \pm 0.09$
Ŏ	LUNAR [4] (train 5%)	$78.91 \pm 1.69$
	InternalContrastiveLearning [11]	$76.96 \pm 2.12$
	BERT [3] for anomalies	$74.61 \pm 0.13$
		PR-AUC outliers (%) ↑
	<b>OC-SVM</b> [10] (train 5%)	$67.94 \pm 0.21$
Classical	IsoForest [7]	$81.46 \pm 2.52$
SSİ	<b>ECOD</b> [6]	78.37
Cla	COPOD [5]	78.19
•	LOF [2]	$83.86 \pm 0.98$
Deep	SO-GAAL [8]	$70.38 \pm 0.28$
	deepSVDD [9]	$92.65 \pm 0.64$
	AE [1] for anomalies	$78.99 \pm 0.28$
	LUNAR [4] (train 5%)	$88.01 \pm 1.03$
	InternalContrastiveLearning [11]	$89.08 \pm 0.87$
	BERT [3] for anomalies	$89.83 \pm 0.07$

## References

[1] Charu C Aggarwal. An introduction to outlier analysis. In *Outlier analysis*, pages 1–34. Springer, 2017.

- [2] Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. LOF: identifying density-based local outliers. In *SIGMOD International Conference on Management of Data*, 2000.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. *NAACL*, 2019.
- [4] Adam Goodge, Bryan Hooi, See-Kiong Ng, and Wee Siong Ng. Lunar: Unifying local outlier detection methods via graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 6737–6745, 2022.
- [5] Zheng Li, Yue Zhao, Nicola Botta, Cezar Ionescu, and Xiyang Hu. Copod: copula-based outlier detection. In 2020 IEEE International Conference on Data Mining (ICDM), pages 1118–1123. IEEE, 2020.
- [6] Zheng Li, Yue Zhao, Xiyang Hu, Nicola Botta, Cezar Ionescu, and George Chen. Ecod: Unsupervised outlier detection using empirical cumulative distribution functions. *IEEE Transactions on Knowledge and Data Engineering*, 2022.
- [7] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation-based anomaly detection. *ACM Trans. Knowl. Discov. Data*, 2012.
- [8] Yezheng Liu, Zhe Li, Chong Zhou, Yuanchun Jiang, Jianshan Sun, Meng Wang, and Xiangnan He. Generative adversarial active learning for unsupervised outlier detection. *IEEE Transactions on Knowledge and Data Engineering*, 32(8):1517–1528, 2019.
- [9] Lukas Ruff, Robert A. Vandermeulen, Nico Görnitz, Lucas Deecke, Shoaib A. Siddiqui, Alexander Binder, Emmanuel Müller, and Marius Kloft. Deep one-class classification. In *International Conference on Machine Learning ICML*, 2018.
- [10] Bernhard Schölkopf, Robert C. Williamson, Alexander J. Smola, John Shawe-Taylor, and John C. Platt. Support vector method for novelty detection. In *Advances in Neural Information Processing Systems*, *NIPS*, 1999.
- [11] Tom Shenkar and Lior Wolf. Anomaly detection for tabular data with internal contrastive learning. In *International Conference on Learning Representations*, 2021.