Supplementary Material for
On the Evaluation of Unsupervised Outlier Detection: Measures, Datasets, and an Empirical Study
by G. O. Campos, A. Zimek, J. Sander, R. J. G. B. Campello, B. Micenková, E. Schubert, I. Assent and M. E. Houle
Data Mining and Knowledge Discovery 30(4): 891-927, 2016, DOI: 10.1007/s10618-015-0444-8

Stamps (9% of outliers)

A data set representing forged (photocopied or scanned+printed) stamps and genuine (ink) stamps. The features are based on color and printing properties of the stamps. Forged stamps are considered to be outliers. The stamps data set is not taken from the UCI repository, but was used in [1].

References:

[1] B. Micenkova, J. van Beusekom, and F. Shafait. Stamp verification for automated document authentication. In 5th Int. Workshop on Computational Forensics, 2012.

Download all data set variants used (371.2 kB).

- Normalized, without duplicates
- Not normalized, without duplicates

Normalized, without duplicates

This version contains 9 attributes, 340 objects, 31 outliers (9.12%)

Download raw algorithm results (3.0 MB) Download raw algorithm evaluation table (61.3 kB)
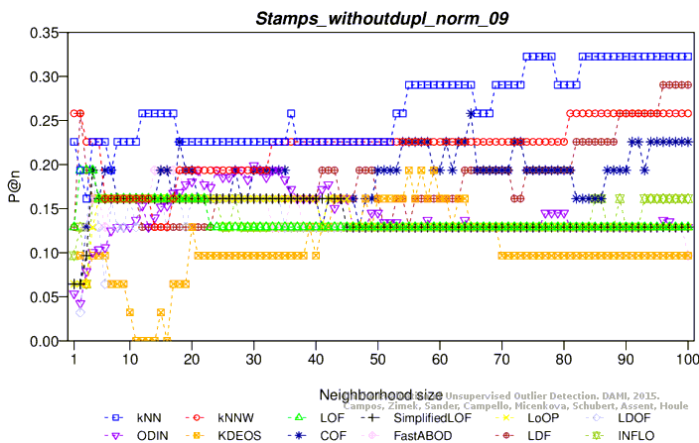
Best Parameters

The following table contains the best (overall and per-method) results for each method and evaluation measure (when the same score was achieved twice, only the smallest k is given).
The Maximum F1-Measure is complimentary in addition to the measures in the original publication.

Algorithm	k	P@n	Adj. P@n	AP	Adj. AP	Max-F1	Adj. MF1	ROC AUC
KNN	15	0.25806	0.18363	0.33546	0.26879	0.54206	0.49611	0.90114
KNN	18	0.22581	0.14814	0.33385	0.26702	0.55670	0.51223	0.90051
KNN	74	0.32258	0.25462	0.33597	0.26935	0.50980	0.46063	0.89665
KNN	100	0.32258	0.25462	0.34421	0.27841	0.52000	0.47184	0.89895
KNNW	1	0.25806	0.18363	0.18224	0.10019	0.28571	0.21405	0.67152
KNNW	77	0.22581	0.14814	0.32522	0.25753	0.54386	0.49810	0.89571

Stamps (9% of outliers)

KNNW	**96**	0.25806	0.18363	0.32805	0.26063	0.53097	0.48392	0.89675
KNNW	**100**	0.25806	0.18363	0.32818	0.26078	0.51786	0.46949	0.89665
LOF	**2**	0.19355	0.11264	0.12098	0.03279	0.23188	0.15482	0.47970
LOF	**100**	0.12903	0.04165	0.27272	0.19976	0.40602	0.34642	0.83318
SimplifiedLOF	**4**	0.19355	0.11264	0.12317	0.03520	0.21053	0.13132	0.50110
SimplifiedLOF	**99**	0.12903	0.04165	0.21509	0.13634	0.29268	0.22172	0.74350
SimplifiedLOF	**100**	0.12903	0.04165	0.21516	0.13642	0.29268	0.22172	0.74350
LoOP	**5**	0.16129	0.07715	0.10943	0.02008	0.17544	0.09272	0.44566
LoOP	**93**	0.12903	0.04165	0.21737	0.13885	0.29752	0.22705	0.74528
LoOP	**100**	0.12903	0.04165	0.22058	0.14238	0.29630	0.22570	0.75279
LDOF	**4**	0.19355	0.11264	0.15448	0.06965	0.25455	0.17976	0.66260
LDOF	**100**	0.12903	0.04165	0.21966	0.14138	0.29379	0.22294	0.75258
ODIN	**30**	0.19892	0.11856	0.18257	0.10056	0.29310	0.22219	0.71725
ODIN	**40**	0.16129	0.07715	0.21647	0.13786	0.29167	0.22060	0.74679
ODIN	**100**	0.12903	0.04165	0.21629	0.13766	0.31868	0.25033	0.75342
FastABOD	**14**	0.19355	0.11264	0.17123	0.08809	0.26087	0.18672	0.70947
FastABOD	**74**	0.12903	0.04165	0.18983	0.10855	0.31429	0.24549	0.75759
FastABOD	**97**	0.12903	0.04165	0.19069	0.10949	0.31343	0.24455	0.76219
KDEOS	**55**	0.19355	0.11264	0.13967	0.05336	0.23333	0.15642	0.63618
KDEOS	**86**	0.09677	0.00616	0.20048	0.12027	0.25503	0.18030	0.68212
KDEOS	**91**	0.09677	0.00616	0.16594	0.08226	0.26389	0.19004	0.68431
KDEOS	**99**	0.09677	0.00616	0.16251	0.07850	0.25676	0.18219	0.69130
LDF	**96**	0.29032	0.21913	0.35356	0.28870	0.51064	0.46154	0.89352
LDF	**97**	0.29032	0.21913	0.35496	0.29024	0.51613	0.46759	0.89425
LDF	**100**	0.29032	0.21913	0.34140	0.27533	0.51613	0.46759	0.89550
INFLO	**4**	0.16129	0.07715	0.11424	0.02538	0.17804	0.09558	0.49107
INFLO	**100**	0.16129	0.07715	0.24091	0.16476	0.34783	0.28240	0.78923
COF	**65**	0.25806	0.18363	0.24722	0.17170	0.30303	0.23311	0.75478
COF	**93**	0.19355	0.11264	0.27916	0.20685	0.39726	0.33679	0.81501
COF	**100**	0.22581	0.14814	0.24987	0.17461	0.41667	0.35814	0.81867

Plots

Precision at n
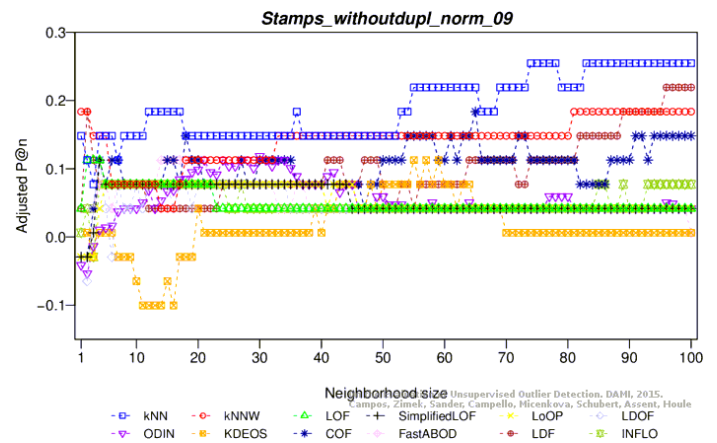
Adjusted precision at n

Average precision

Adjusted average precision

Maximum F1 score

Adjusted maximum F1 score

ROC AUC

Diversity
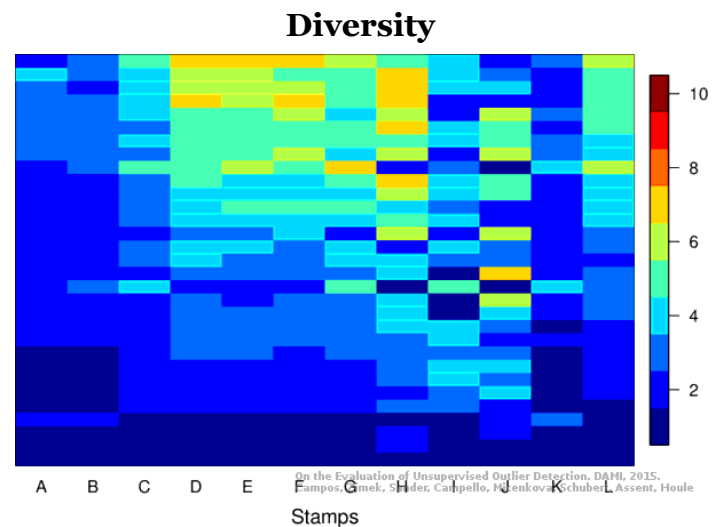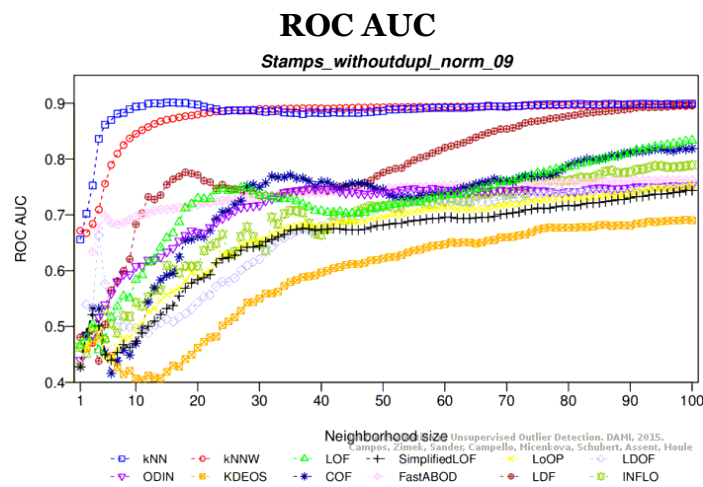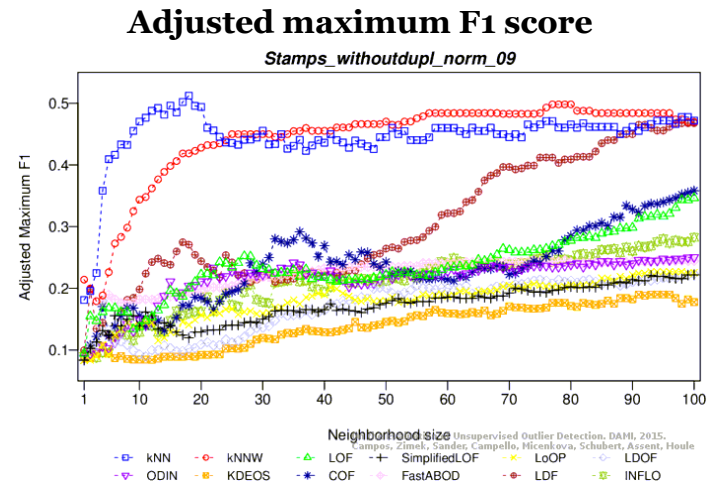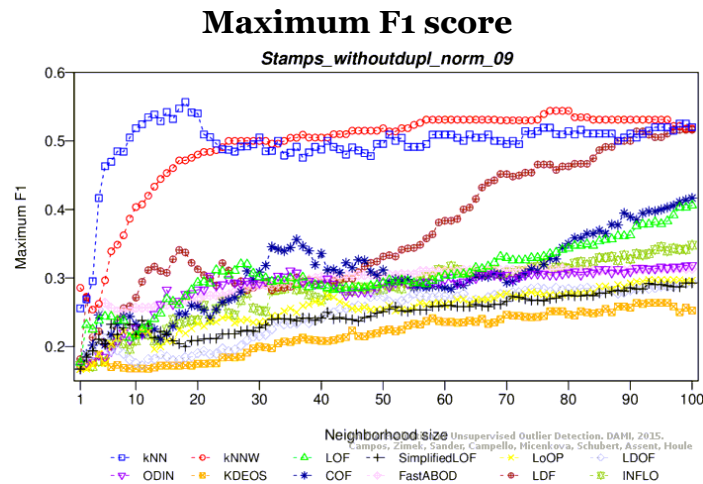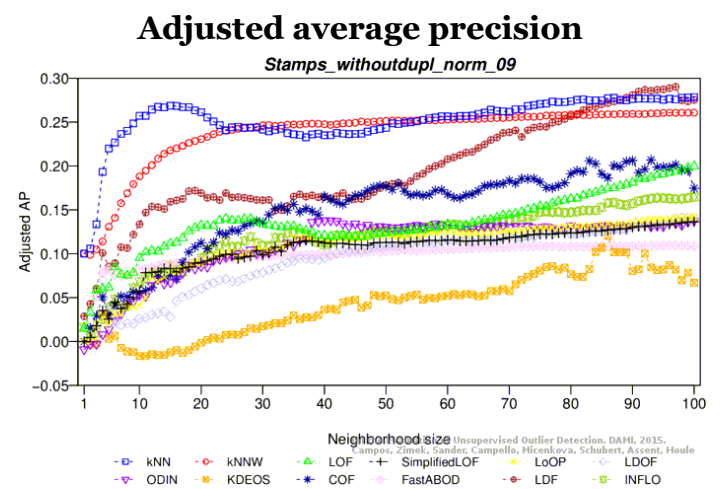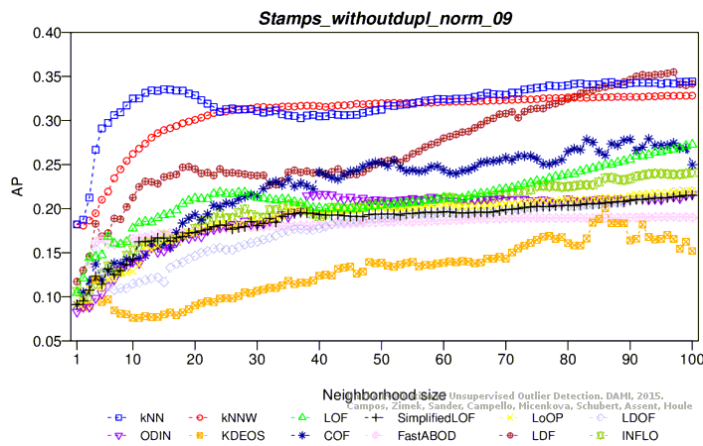
A: KNN, B: KNNW, C: LOF, D: SimplifiedLOF, E: LoOP, F: LDOF

G: ODIN, H: KDEOS, I: COF, J: FastABOD, K: LDF, L: INFLO

Not normalized, without duplicates

This version contains 9 attributes, 340 objects, 31 outliers (9.12%)

[Download raw algorithm results (3.0 MB)](#) [Download raw algorithm evaluation table (61.1 kB)](#)

Best Parameters

The following table contains the best (overall and per-method) results for each method and evaluation measure (when the same score was achieved twice, only the smallest k is given).
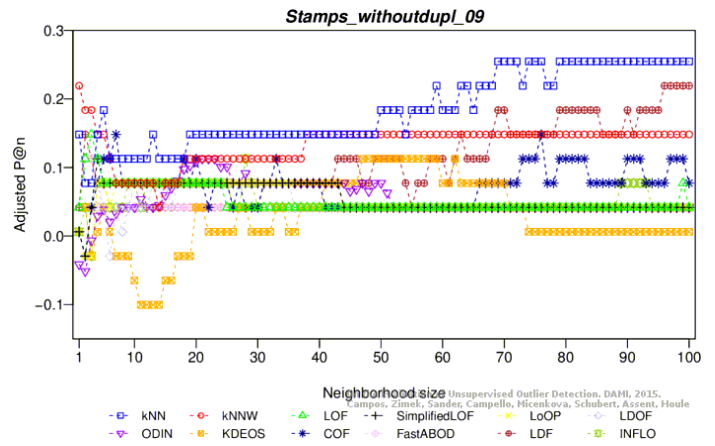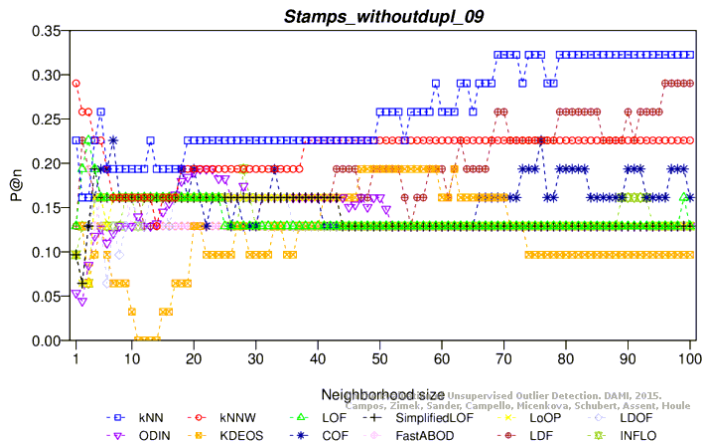The Maximum F1-Measure is complimentary in addition to the measures in the original publication.

Algorithm	k	P@n	Adj. P@n	AP	Adj. AP	Max-F1	Adj. MF1	ROC AUC
KNN	66	0.29032	0.21913	0.34522	0.27953	0.56566	0.52208	0.90437
KNN	69	0.32258	0.25462	0.34704	0.28153	0.56000	0.51586	0.90495
KNN	80	0.32258	0.25462	0.35252	0.28756	0.55102	0.50598	0.90678
KNN	93	0.32258	0.25462	0.35615	0.29155	0.54902	0.50378	0.90657
KNNW	1	0.29032	0.21913	0.18548	0.10376	0.29508	0.22436	0.67778
KNNW	77	0.22581	0.14814	0.33620	0.26960	0.56364	0.51986	0.90250
KNNW	99	0.22581	0.14814	0.33991	0.27368	0.55856	0.51427	0.90385
LOF	3	0.22581	0.14814	0.15429	0.06945	0.25397	0.17912	0.51754
LOF	96	0.12903	0.04165	0.28949	0.21821	0.43750	0.38107	0.84831
LOF	100	0.12903	0.04165	0.29719	0.22668	0.43750	0.38107	0.85562
SimplifiedLOF	4	0.19355	0.11264	0.13113	0.04396	0.20438	0.12456	0.51195
SimplifiedLOF	97	0.12903	0.04165	0.21884	0.14047	0.29412	0.22330	0.74851
SimplifiedLOF	100	0.12903	0.04165	0.22221	0.14418	0.29167	0.22060	0.75488
LoOP	4	0.16129	0.07715	0.12603	0.03835	0.21488	0.13611	0.51613
LoOP	100	0.12903	0.04165	0.22437	0.14655	0.30208	0.23207	0.75759
LDOF	4	0.16129	0.07715	0.15202	0.06695	0.25000	0.17476	0.65497
LDOF	98	0.12903	0.04165	0.22163	0.14354	0.29474	0.22398	0.75436
LDOF	100	0.12903	0.04165	0.22153	0.14343	0.30270	0.23275	0.75561
ODIN	21	0.19355	0.11264	0.16863	0.08522	0.28283	0.21088	0.66301
ODIN	100	0.12903	0.04165	0.22300	0.14504	0.33333	0.26645	0.77633
FastABOD	3	0.19355	0.11264	0.16135	0.07722	0.25000	0.17476	0.66155
FastABOD	98	0.12903	0.04165	0.19264	0.11164	0.31776	0.24931	0.76553
FastABOD	100	0.12903	0.04165	0.19285	0.11188	0.31776	0.24931	0.76595
KDEOS	47	0.19355	0.11264	0.14487	0.05908	0.21591	0.13725	0.60413
KDEOS	83	0.09677	0.00616	0.20014	0.11989	0.24540	0.16969	0.66510
KDEOS	100	0.09677	0.00616	0.14416	0.05830	0.24900	0.17365	0.67136
LDF	96	0.29032	0.21913	0.38047	0.31832	0.57143	0.52843	0.90897
LDF	97	0.29032	0.21913	0.38241	0.32045	0.56566	0.52208	0.90970
INFLO	28	0.19355	0.11264	0.18455	0.10274	0.23602	0.15938	0.60544
INFLO	96	0.12903	0.04165	0.24131	0.16520	0.35220	0.28721	0.78975
INFLO	98	0.12903	0.04165	0.24404	0.16820	0.35220	0.28721	0.79330
COF	7	0.22581	0.14814	0.14093	0.05474	0.22581	0.14814	0.43324
COF	85	0.16129	0.07715	0.26957	0.19629	0.35838	0.29401	0.79319
COF	96	0.16129	0.07715	0.25529	0.18057	0.41892	0.36062	0.80342
COF	99	0.19355	0.11264	0.23529	0.15858	0.41096	0.35186	0.80708

Plots

Precision at n   **Adjusted precision at n**

Stamps_withoutdupl_09

Stamps_withoutdupl_09

Average precision

Stamps_withoutdupl_09

Adjusted average precision

Stamps_withoutdupl_09

Maximum F1 score

Stamps_withoutdupl_09

Adjusted maximum F1 score

Stamps_withoutdupl_09

ROC AUC

Stamps_withoutdupl_09
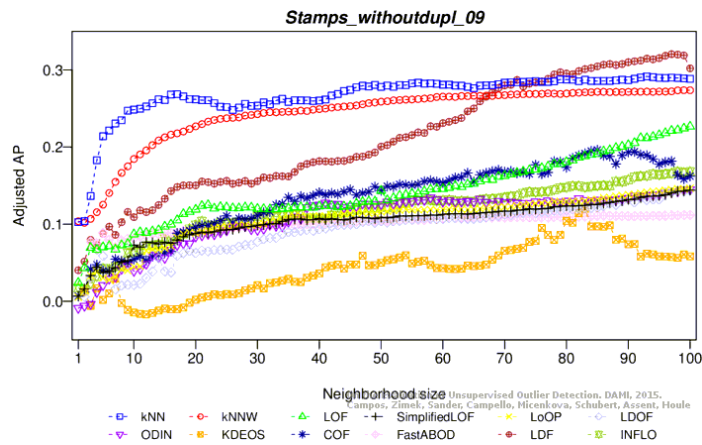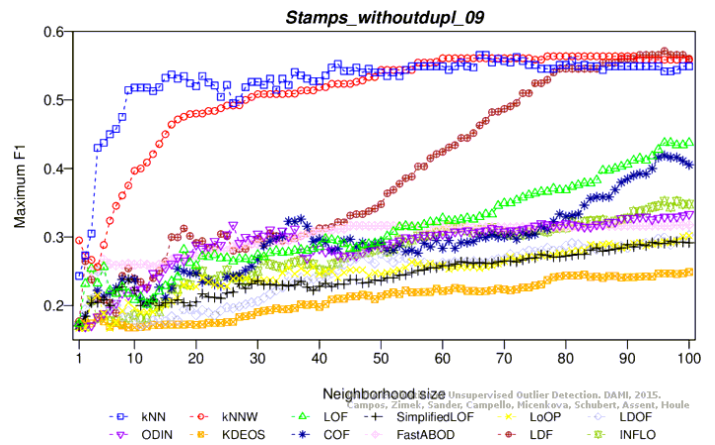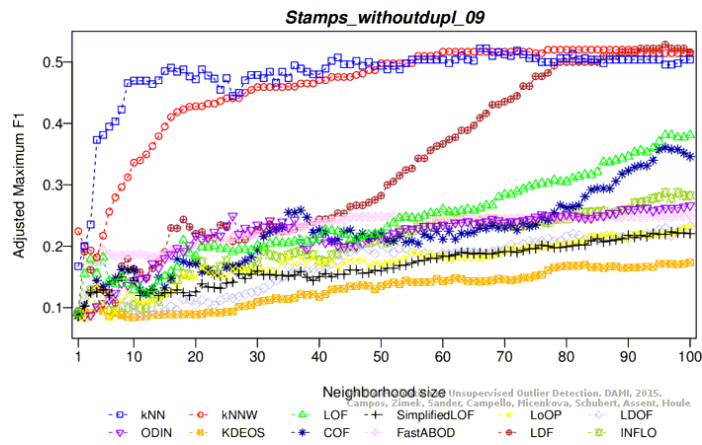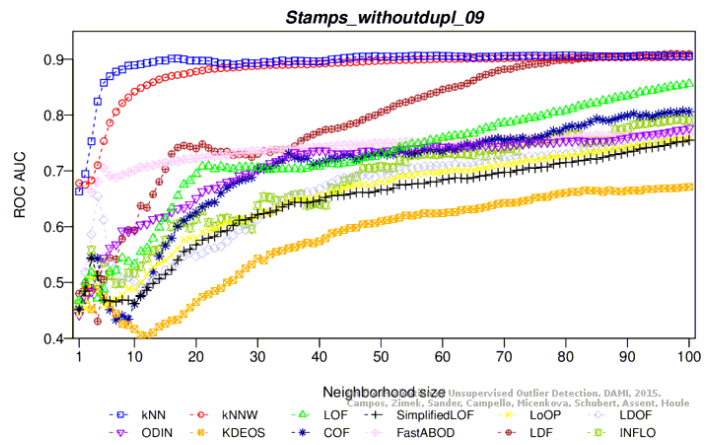
Diversity
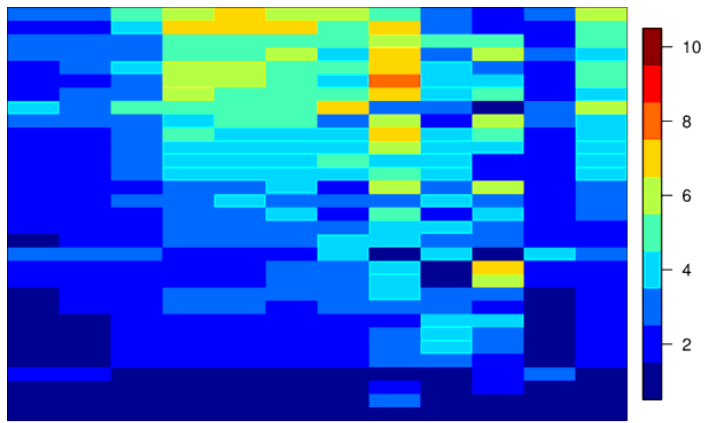
Stamps

A: KNN, B: KNNW, C: LOF, D: SimplifiedLOF, E: LoOP, F:
LDOF
G: ODIN, H: KDEOS, I: COF, J: FastABOD, K: LDF, L:
INFLO

File generated: 2016-07-05T21:49:13