

Cross-project online just-in-time software defect prediction

Tabassum, Sadia; Minku, Leandro; Feng, Danyi

DOI:

[10.1109/TSE.2022.3150153](https://doi.org/10.1109/TSE.2022.3150153)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Tabassum, S, Minku, L & Feng, D 2022, 'Cross-project online just-in-time software defect prediction', *IEEE Transactions on Software Engineering*. <https://doi.org/10.1109/TSE.2022.3150153>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Cross-Project Online Just-In-Time Software Defect Prediction – Supplementary Material

Sadia Tabassum, Leandro L. Minku, *Senior Member, IEEE*, Danyi Feng



This supplementary material complements the paper entitled “Cross-Project Online Just-In-Time Software Defect Prediction” authored by Sadia Tabassum, Leandro L. Minku and Danyi Feng.

1 RESULT TABLES

This section provides all result tables, including tables for approaches based on OOB and ORB. In addition to tables depicting the initial period, periods of sudden drop in predictive performance and stable periods, there is also a table comparing all approaches across all time steps. The tables are shown from the next page onwards, to facilitate reading.

-
- S. Tabassum and L.L. Minku (the corresponding author) are with the School of Computer Science, University of Birmingham, UK, emails: sxt901@cs.bham.ac.uk and L.L.Minku@bham.ac.uk.
 - D. Feng is with Xiliu Tech, China, email: danyi@ouchteam.com.

TABLE 1: Number of initial time steps of the initial phase, and average G-Means, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Opensource data

| Dataset | #Time Steps | Dummy | WP-OOB | AIO-OOB | Filter-OOB | Ensemble-OOB | OP-AIO-OOB | OP-Filter-OOB |
|--------------------|-------------|-------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Tomcat | 2006 | 50 | 43.62(4.84)[-b] | 51.95(1.25)[b] | 52.5(0.71)[b] | 33.67(0.63)[-b] | 53.12(3.28)[b] | 52.86(1.15)[b] |
| JGroups | 1268 | 50 | 38.6(0.83)[-b] | 38.29(0.84)[-b] | 39.01(0.75)[-b] | 16.64(0.36)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 461 | 50 | 27.3(0.12)[-b] | 24.28(0.25)[-b] | 38.69(0.49)[-b] | 19.48(0.22)[-b] | 11.15(6.13)[-b] | 29.16(6.34)[-b] |
| Camel | 3112 | 50 | 46.82(1.71)[-b] | 57.75(0.74)[b] | 57.07(0.82)[b] | 39.42(0.52)[-b] | 47.0(1.06)[-b] | 46.94(0.96)[-b] |
| Brackets | 1569 | 50 | 21.37(0.03)[-b] | 64.89(1.32)[b] | 66.17(0.74)[b] | 46.83(0.91)[-b] | 50.46(1.05)[s] | 59.45(1.33)[b] |
| Nova | 6271 | 50 | 55.42(0.41)[b] | 63.15(0.68)[b] | 64.39(0.34)[b] | 51.66(0.29)[b] | 62.55(1.16)[b] | 62.88(4.43)[b] |
| Fabric8 | 795 | 50 | 27.01(0.3)[-b] | 51.51(1.42)[b] | 58.63(1.57)[b] | 40.3(0.69)[-b] | 56.57(3.12)[b] | 45.52(1.08)[-b] |
| Neutron | 917 | 50 | 44.84(5.75)[-b] | 73.55(0.99)[b] | 67.29(2.45)[b] | 55.07(1.1)[b] | 67.03(0.41)[b] | 73.15(0.7)[b] |
| Npm | 2536 | 50 | 26.92(1.1)[-b] | 48.37(1.34)[-b] | 45.75(1.94)[-b] | 42.02(0.34)[-b] | 42.89(6.01)[-b] | 15.7(1.15)[-b] |
| BroadleafCommerce | 677 | 50 | 26.37(0.26)[-b] | 50.32(1.55)[*] | 53.16(1.99)[b] | 37.26(1.58)[-b] | 40.29(2.8)[-b] | 30.63(3.6)[-b] |
| Ranking | | 2 | 4 | 1 | 1 | 4 | 3 | 3 |

| Dataset | #Time Periods | Dummy | WP-ORB | AIO-ORB | Filter-ORB | Ensemble-ORB | OP-AIO-ORB | OP-Filter-ORB |
|--------------------|---------------|-------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Tomcat | 926 | 50 | 40.45(1.39)[-b] | 45.83(1.39)[-b] | 48.13(1.58)[-b] | 35.32(0.46)[-b] | 27.73(6.1)[-b] | 32.87(8.82)[-b] |
| JGroups | 1412 | 50 | 30.72(1.55)[-b] | 31.48(1.31)[-b] | 33.48(1.44)[-b] | 15.14(0.79)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 900 | 50 | 25.7(0.8)[-b] | 60.47(3.09)[b] | 63.43(3.06)[b] | 42.17(1.25)[-b] | 62.77(2.92)[b] | 48.32(2.29)[-b] |
| Camel | 5111 | 50 | 49.69(1.59)[-s] | 57.94(0.92)[b] | 57.64(1.3)[b] | 41.38(0.74)[-b] | 50.32(4.15)[-s] | 49.76(4.08)[-s] |
| Brackets | 1721 | 50 | 13.5(0.32)[-b] | 67.39(1.42)[b] | 67.21(0.75)[b] | 50.46(0.9)[b] | 65.23(1.32)[b] | 59.26(2.72)[b] |
| Nova | 6271 | 50 | 52.8(1.78)[b] | 66.44(0.37)[b] | 65.0(0.78)[b] | 51.85(0.88)[b] | 34.29(14.0)[-b] | 56.84(6.67)[b] |
| Fabric8 | 1613 | 50 | 29.97(0.45)[-b] | 63.34(1.84)[b] | 62.9(1.09)[b] | 45.2(1.01)[-b] | 56.59(4.16)[b] | 55.11(3.05)[b] |
| Neutron | 3304 | 50 | 71.78(1.03)[b] | 75.09(1.11)[b] | 74.98(1.42)[b] | 71.04(1.0)[b] | 63.71(13.18)[b] | 70.28(4.96)[b] |
| Npm | 1494 | 50 | 30.12(2.16)[-b] | 40.07(2.86)[-b] | 43.49(4.92)[-b] | 43.53(1.45)[-b] | 39.02(19.3)[-m] | 0.0(0.0)[-b] |
| BroadleafCommerce | 950 | 50 | 27.68(0.25)[-b] | 51.82(3.62)[m] | 52.41(4.96)[s] | 41.69(0.85)[-b] | 58.34(4.49)[b] | 44.01(2.49)[-b] |
| Ranking | | 2 | 4 | 1 | 1 | 3 | 2 | 3 |

Standard deviations are shown in brackets. Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the Dummy approach. Smaller Scott-Knott.BA12 rankings are better rankings.

TABLE 2: Number of initial time steps of the initial phase, and average G-Means, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB

| Dataset | #Time Steps | Dummy | WP-OOB | AIO-OOB combined | AIO-OOB proprietary | Filter-OOB combined | OP-AIO-OOB combined | OP-AIO-OOB proprietary | OP-Filter-OOB combined |
|---------|-------------|-------|-----------------|------------------|---------------------|---------------------|---------------------|------------------------|------------------------|
| C1 | 284 | 50 | 15.76(4.12)[-b] | 35.07(3.74)[-b] | 15.97(3.9)[-b] | 42.5(2.11)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 581 | 50 | 24.62(0.37)[-b] | 31.03(0.43)[-b] | 32.71(5.38)[-b] | 46.82(0.76)[-b] | 43.84(1.9)[-b] | 5.62(6.26)[-b] | 45.33(2.97)[-b] |
| C3 | 82 | 50 | 8.48(0.0)[-b] | 23.36(0.0)[-b] | 35.9(0.0)[-b] | 8.62(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 24 | 50 | 7.59(0.13)[-b] | 18.93(2.73)[-b] | 18.63(3.19)[-b] | 0.0(0.0)[-b] | 20.1(0.14)[-b] | 14.47(0.39)[-b] | 0.0(0.0)[-b] |
| C5 | 8 | 50 | 11.62(0.0)[-b] | 1.69(2.63)[-b] | 31.66(4.57)[-b] | 10.14(3.89)[-b] | 5.16(6.01)[-b] | 1.99(4.52)[-b] | 39.83(0.0)[-b] |
| C6 | 20 | 50 | 3.6(0.0)[-b] | 1.8(0.0)[-b] | 1.8(0.0)[-b] | 1.38(0.77)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C7 | 29 | 50 | 1.17(0.0)[-b] | 42.76(2.42)[-b] | 50.92(3.65)[-s] | 0.0(0.0)[-b] | 40.61(8.02)[-b] | 23.0(10.38)[-b] | 0.0(0.0)[-b] |
| C8 | 39 | 50 | 0.76(0.0)[-b] | 5.87(4.91)[-b] | 40.16(8.41)[-b] | 0.66(0.26)[-b] | 16.56(8.28)[-b] | 37.57(6.39)[-b] | 0.0(0.0)[-b] |
| C9 | 62 | 50 | 10.44(1.43)[-b] | 15.4(6.94)[-b] | 25.12(2.67)[-b] | 0.55(0.0)[-b] | 21.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Ranking | | 1 | 4 | 3 | 2 | 4 | 4 | 4 | 4 |

| Dataset | #Time Steps | Dummy | WP-ORB | AIO-ORB combined | AIO-ORB proprietary | Filter-ORB combined | OP-AIO-ORB combined | OP-AIO-ORB proprietary | OP-Filter-ORB combined |
|---------|-------------|-------|-----------------|------------------|---------------------|---------------------|---------------------|------------------------|------------------------|
| C1 | 671 | 50 | 25.54(2.04)[-b] | 30.77(1.43)[-b] | 25.34(1.12)[-b] | 35.03(2.64)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 582 | 50 | 16.67(0.51)[-b] | 38.71(3.98)[-b] | 39.32(3.38)[-b] | 28.6(0.76)[-b] | 41.93(2.97)[-b] | 30.43(12.08)[-b] | 40.93(5.07)[-b] |
| C3 | 396 | 50 | 38.67(0.64)[-b] | 38.19(0.45)[-b] | 39.31(1.26)[-b] | 40.62(0.93)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 93 | 50 | 4.44(0.2)[-b] | 35.18(3.62)[-b] | 43.86(1.29)[-b] | 23.17(4.39)[-b] | 41.81(3.75)[-b] | 42.81(0.97)[-b] | 19.07(0.73)[-b] |
| C5 | 297 | 50 | 9.53(1.07)[-b] | 23.72(4.03)[-b] | 25.43(3.96)[-b] | 35.72(3.56)[-b] | 52.41(3.99)[s] | 47.88(4.29)[-s] | 54.87(2.63)[b] |
| C6 | 210 | 50 | 2.08(0.0)[-b] | 31.87(0.96)[-b] | 31.45(2.3)[-b] | 29.12(1.21)[-b] | 30.05(2.19)[-b] | 11.88(0.89)[-b] | 21.61(1.45)[-b] |
| C7 | 62 | 50 | 1.76(0.4)[-b] | 45.34(6.83)[-b] | 43.59(2.0)[-b] | 14.65(1.7)[-b] | 35.82(14.86)[-b] | 29.86(12.14)[-b] | 0.0(0.0)[-b] |
| C8 | 232 | 50 | 1.55(0.1)[-b] | 48.68(5.87)[-s] | 50.06(5.09)[*] | 35.23(2.99)[-b] | 50.43(5.74)[-s] | 54.31(2.16)[b] | 40.83(2.08)[-b] |
| C9 | 198 | 50 | 1.04(0.0)[-b] | 29.05(2.79)[-b] | 36.29(1.78)[-b] | 21.21(0.02)[-b] | 19.25(0.22)[-b] | 1.27(4.7)[-b] | 5.23(0.0)[-b] |
| Ranking | | 1 | 6 | 3 | 2 | 4 | 3 | 5 | 5 |

Standard deviations are shown in brackets. Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the Dummy approach. Smaller Scott-Knott.BA12 rankings are better rankings.

TABLE 3: Total number of time steps of the performance drop periods, average G-Means, A12 effect sizes against the Dummy approach and Scott-Knott.BA12 to compare approaches using OOB and ORB for open source data during the drop periods

| Dataset | #Time Steps | Dummy | WP-OOB | AIO-OOB | Filter-OOB | Ensemble-OOB | OP-AIO-OOB | OP-Filter-OOB |
|--------------------|-------------|-------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Tomcat | 3855 | 50 | 48.52(1.17)[-b] | 58.86(0.72)[b] | 60.22(1.0)[b] | 52.26(0.32)[b] | 36.58(6.17)[-b] | 43.78(1.42)[-b] |
| JGroups | 4353 | 50 | 43.95(1.44)[-b] | 48.01(1.05)[-b] | 48.08(0.89)[-b] | 31.41(0.5)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 3973 | 50 | 32.98(0.7)[-b] | 59.01(0.53)[b] | 60.15(0.97)[b] | 45.68(0.43)[-b] | 52.2(5.28)[s] | 34.89(4.74)[-b] |
| Camel | 6009 | 50 | 53.15(1.68)[b] | 56.17(0.98)[b] | 57.0(0.92)[b] | 52.4(0.41)[b] | 50.16(1.49)[*] | 50.67(0.69)[b] |
| Nova | 4427 | 50 | 66.99(0.06)[b] | 66.99(0.18)[b] | 66.85(0.26)[b] | 57.19(0.66)[b] | 65.4(3.73)[b] | 67.03(6.47)[b] |
| Fabric8 | 3443 | 50 | 53.79(2.18)[b] | 56.34(0.76)[b] | 65.57(1.04)[b] | 53.98(0.4)[b] | 22.94(6.32)[-b] | 24.49(6.65)[-b] |
| Npm | 1202 | 50 | 38.64(1.95)[-b] | 62.97(1.47)[b] | 65.37(1.25)[b] | 51.13(0.62)[b] | 44.21(7.42)[-s] | 17.21(3.51)[-b] |
| BroadleafCommerce | 3579 | 50 | 50.18(0.55)[s] | 67.39(1.05)[b] | 66.05(0.66)[b] | 54.38(0.25)[b] | 52.72(1.17)[b] | 56.02(2.4)[b] |
| Ranking | | 3 | 3 | 2 | 1 | 3 | 4 | 4 |

| Dataset | #Time Steps | Dummy | WP-ORB | AIO-ORB | Filter-ORB | Ensemble-ORB | OP-AIO-ORB | OP-Filter-ORB |
|--------------------|-------------|-------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|
| Tomcat | 2265 | 50 | 47.34(1.72)[-b] | 56.29(1.15)[b] | 56.23(1.22)[b] | 51.1(0.46)[b] | 25.76(8.05)[-b] | 25.53(11.14)[-b] |
| JGroups | 2066 | 50 | 52.94(1.49)[b] | 53.93(1.37)[b] | 55.36(1.45)[b] | 34.42(0.45)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 2319 | 50 | 36.48(1.38)[-b] | 64.14(0.86)[b] | 65.49(1.07)[b] | 47.13(0.72)[-b] | 66.62(1.78)[b] | 54.96(3.64)[b] |
| Camel | 1105 | 50 | 48.98(3.85)[-m] | 55.46(1.7)[b] | 55.21(2.74)[b] | 56.61(0.92)[b] | 56.9(6.38)[b] | 55.88(5.57)[b] |
| Nova | 4408 | 50 | 67.29(0.57)[b] | 67.48(0.56)[b] | 66.65(0.25)[b] | 61.5(0.62)[b] | 20.38(18.77)[-b] | 60.54(8.49)[b] |
| Fabric8 | 1473 | 50 | 55.13(5.87)[b] | 57.2(1.9)[b] | 63.55(0.94)[b] | 55.15(1.05)[b] | 51.26(6.09)[m] | 44.7(4.8)[-b] |
| Npm | 936 | 50 | 44.07(4.78)[-b] | 46.78(4.55)[-b] | 46.11(9.12)[-s] | 50.03(0.65)[-s] | 40.28(19.24)[-m] | 0.0(0.0)[-b] |
| BroadleafCommerce | 2703 | 50 | 51.72(2.03)[b] | 67.03(1.11)[b] | 66.98(0.98)[b] | 54.38(0.65)[b] | 49.33(5.1)[*] | 58.01(2.88)[b] |
| Ranking | | 2 | 2 | 1 | 1 | 2 | 3 | 3 |

Standard deviations are shown in brackets. Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the Dummy approach. Smaller Scott-Knott.BA12 rankings are better rankings.

TABLE 4: Total number of time steps of the performance drop periods, average G-Means, A12 effect sizes against the Dummy approach and Scott-Knott.BA12 to compare approaches using OOB and ORB for proprietary data during the drop periods

| Dataset | #Time Steps | Dummy | WP-OOB | AIO-OOB combined | AIO-OOB proprietary | Filter-OOB combined | OP-AIO-OOB combined | OP-AIO-OOB proprietary | OP-Filter-OOB combined |
|---------|-------------|-------|-----------------|------------------|---------------------|---------------------|---------------------|------------------------|------------------------|
| C1 | 452 | 50 | 32.27(4.27)[-b] | 41.37(1.45)[-b] | 43.4(4.45)[-b] | 46.66(1.08)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C3 | 210 | 50 | 42.79(1.72)[-b] | 47.75(1.31)[-b] | 39.1(1.42)[-b] | 52.17(1.3)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 249 | 50 | 33.0(0.95)[-b] | 51.47(1.74)[b] | 60.05(2.89)[b] | 56.78(1.76)[b] | 57.74(1.11)[b] | 47.1(1.74)[-b] | 51.35(0.0)[b] |
| C6 | 466 | 50 | 21.34(1.35)[-b] | 43.64(1.49)[-b] | 45.89(5.02)[-b] | 44.32(1.98)[-b] | 43.53(0.41)[-b] | 37.47(0.18)[-b] | 45.1(0.47)[-b] |
| C7 | 255 | 50 | 25.74(1.19)[-b] | 41.5(5.94)[-b] | 32.34(1.51)[-b] | 40.34(1.02)[-b] | 50.43(2.6)[-s] | 36.26(0.52)[-b] | 30.73(0.4)[-b] |
| C8 | 287 | 50 | 7.3(0.89)[-b] | 44.65(0.96)[-b] | 34.6(5.43)[-b] | 40.14(1.82)[-b] | 44.3(1.81)[-b] | 45.49(9.63)[*] | 25.57(8.7)[-b] |
| C9 | 541 | 50 | 31.75(1.77)[-b] | 51.34(1.76)[b] | 50.08(2.77)[s] | 53.25(1.58)[b] | 40.3(5.22)[-b] | 34.92(3.87)[-b] | 34.36(0.55)[-b] |
| Ranking | | 1 | 5 | 3 | 4 | 2 | 4 | 5 | 5 |

| Dataset | #Time Steps | Dummy | WP-ORB | AIO-ORB combined | AIO-ORB proprietary | Filter-ORB combined | OP-AIO-ORB combined | OP-AIO-ORB proprietary | OP-Filter-ORB combined |
|---------|-------------|-------|-----------------|------------------|---------------------|---------------------|---------------------|------------------------|------------------------|
| C1 | 238 | 50 | 41.17(6.37)[-b] | 37.26(5.59)[-b] | 28.75(3.01)[-b] | 47.3(6.34)[-s] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C6 | 64 | 50 | 35.38(0.91)[-b] | 50.99(7.22)[m] | 35.67(9.61)[-b] | 44.61(5.36)[-b] | 61.13(3.51)[b] | 46.51(5.76)[-m] | 42.82(5.16)[-b] |
| C7 | 85 | 50 | 31.09(4.47)[-b] | 32.33(10.88)[-b] | 50.09(13.02)[m] | 59.45(6.92)[b] | 35.64(17.93)[-s] | 41.39(2.73)[-b] | 0.0(0.0)[-b] |
| C8 | 80 | 50 | 20.25(1.22)[-b] | 39.45(14.5)[-s] | 37.97(14.69)[-s] | 50.59(6.1)[b] | 47.77(6.08)[-s] | 32.34(8.23)[-b] | 45.77(2.75)[-b] |
| C9 | 23 | 50 | 23.59(1.26)[-b] | 39.88(7.47)[-b] | 25.48(4.84)[-b] | 20.64(1.48)[-b] | 43.2(8.64)[-b] | 1.64(3.75)[-b] | 17.96(0.0)[-b] |
| Ranking | | 1 | 4 | 3 | 3 | 2 | 3 | 4 | 4 |

Standard deviations are shown in brackets. Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the Dummy approach. Smaller Scott-Knott.BA12 rankings are better rankings.

TABLE 5: Number of time steps of the stable periods, and average G-Means, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this stable periods for OOB and ORB for open source data

| Dataset | #Time Steps | Dummy | WP-OOB | AIO-OOB | Filter-OOB | Ensemble-OOB | OP-AIO-OOB | OP-Filter-OOB |
|--------------------|-------------|-------|----------------|----------------|----------------|----------------|------------------|-----------------|
| Tomcat | 13046 | 50 | 62.69(0.81)[b] | 61.36(0.5)[b] | 63.36(0.64)[b] | 50.57(0.22)[b] | 39.85(6.0)[-b] | 45.14(1.49)[-b] |
| JGroups | 12704 | 50 | 60.08(0.66)[b] | 58.74(0.72)[b] | 58.51(0.54)[b] | 38.8(0.3)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 4316 | 50 | 64.86(1.01)[b] | 65.99(0.35)[b] | 66.96(0.77)[b] | 56.02(0.33)[b] | 55.71(8.02)[b] | 45.38(8.91)[-*] |
| Camel | 21454 | 50 | 67.33(0.56)[b] | 64.48(0.67)[b] | 65.05(0.62)[b] | 54.61(0.35)[b] | 41.5(1.52)[-b] | 41.54(1.25)[-b] |
| Brackets | 15795 | 50 | 68.27(0.12)[b] | 71.44(0.48)[b] | 71.38(0.42)[b] | 64.25(0.54)[b] | 39.05(12.78)[-b] | 66.69(0.34)[b] |
| Nova | 38291 | 50 | 79.84(0.26)[b] | 82.03(0.21)[b] | 81.54(0.2)[b] | 68.72(0.79)[b] | 73.83(16.61)[b] | 75.2(0.5)[b] |
| Fabric8 | 8868 | 50 | 65.0(0.77)[b] | 63.53(0.62)[b] | 67.64(0.64)[b] | 54.15(0.4)[b] | 25.97(4.87)[-b] | 28.17(4.46)[-b] |
| Neutron | 18605 | 50 | 81.13(0.27)[b] | 81.87(0.23)[b] | 82.22(0.25)[b] | 68.93(1.5)[b] | 74.49(1.77)[b] | 71.43(3.91)[b] |
| Npm | 4182 | 50 | 59.52(1.71)[b] | 64.96(0.63)[b] | 66.49(0.99)[b] | 55.3(0.46)[b] | 29.41(2.02)[-b] | 40.07(1.13)[-b] |
| BroadleafCommerce | 10754 | 50 | 65.96(1.12)[b] | 70.73(0.62)[b] | 70.13(0.55)[b] | 59.55(0.2)[b] | 54.3(0.78)[b] | 60.4(1.06)[b] |
| Ranking | | 4 | 2 | 1 | 1 | 3 | 5 | 4 |

| Dataset | #Time Steps | Dummy | WP-ORB | AIO-ORB | Filter-ORB | Ensemble-ORB | OP-AIO-ORB | OP-Filter-ORB |
|--------------------|-------------|-------|----------------|----------------|----------------|-----------------|-----------------|-----------------|
| Tomcat | 15716 | 50 | 62.71(0.97)[b] | 63.44(0.86)[b] | 62.74(0.57)[b] | 51.69(0.43)[b] | 45.39(1.76)[-b] | 37.75(2.12)[-b] |
| JGroups | 14847 | 50 | 60.31(0.8)[b] | 60.99(0.67)[b] | 61.06(0.74)[b] | 39.88(0.34)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 5531 | 50 | 62.36(0.94)[b] | 70.48(0.7)[b] | 70.75(0.41)[b] | 56.27(0.61)[b] | 68.57(2.2)[b] | 58.38(4.57)[b] |
| Camel | 24359 | 50 | 66.33(1.02)[b] | 66.92(0.39)[b] | 66.34(0.34)[b] | 56.85(0.55)[b] | 51.0(3.63)[b] | 45.33(4.33)[-b] |
| Brackets | 15643 | 50 | 68.79(0.53)[b] | 70.26(1.01)[b] | 70.07(0.83)[b] | 66.28(0.79)[b] | 64.33(0.62)[b] | 64.42(1.56)[b] |
| Nova | 38310 | 50 | 80.06(0.81)[b] | 80.18(0.41)[b] | 79.94(0.38)[b] | 76.12(0.68)[b] | 38.93(30.48)[*] | 76.85(2.32)[b] |
| Fabric8 | 10020 | 50 | 67.41(0.81)[b] | 64.72(1.01)[b] | 68.72(0.59)[b] | 57.74(0.6)[b] | 49.22(6.71)[-s] | 38.08(5.44)[-b] |
| Neutron | 16218 | 50 | 81.09(0.63)[b] | 81.87(0.39)[b] | 81.83(0.51)[b] | 74.93(1.23)[b] | 67.06(12.37)[b] | 70.65(7.2)[b] |
| Npm | 5490 | 50 | 61.92(0.9)[b] | 67.72(0.88)[b] | 65.63(1.3)[b] | 55.23(0.64)[b] | 51.29(19.52)[b] | 0.0(0.0)[-b] |
| BroadleafCommerce | 11357 | 50 | 64.31(1.34)[b] | 69.39(0.6)[b] | 68.9(0.78)[b] | 61.46(0.49)[b] | 41.81(8.62)[-b] | 49.32(2.91)[-s] |
| Ranking | | 4 | 2 | 1 | 1 | 3 | 4 | 4 |

Standard deviations are shown in brackets. Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the Dummy approach. Smaller Scott-Knott.BA12 rankings are better rankings.

TABLE 6: Number of time steps of the stable periods, and average G-Means, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this stable periods for OOB and ORB for proprietary data

| Dataset | #Time Steps | Dummy | WP-OOB | AIO-OOB combined | AIO-OOB proprietary | Filter-OOB combined | OP-AIO-OOB combined | OP-AIO-OOB proprietary | OP-Filter-OOB combined |
|---------|-------------|-------|-----------------|------------------|---------------------|---------------------|---------------------|------------------------|------------------------|
| C1 | 294 | 50 | 49.55(1.47)[-*] | 43.65(2.33)[-b] | 43.4(6.94)[-b] | 45.75(2.02)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 20 | 50 | 64.59(0.23)[b] | 52.86(1.75)[b] | 55.58(2.72)[b] | 59.42(0.62)[b] | 36.49(0.93)[-b] | 2.8(3.22)[-b] | 37.36(1.46)[-b] |
| C3 | 125 | 50 | 54.54(1.05)[b] | 54.01(0.43)[b] | 33.69(2.6)[-b] | 57.0(0.61)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 68 | 50 | 45.72(1.21)[-b] | 53.4(1.23)[b] | 56.98(4.85)[b] | 27.16(8.92)[-b] | 46.76(0.61)[-b] | 44.62(2.6)[-b] | 21.15(0.0)[-b] |
| C5 | 315 | 50 | 23.2(0.43)[-b] | 46.55(1.35)[-b] | 42.6(0.99)[-b] | 49.93(1.71)[-s] | 44.14(5.2)[-b] | 49.53(2.92)[-s] | 41.15(4.14)[-b] |
| C6 | 69 | 50 | 41.71(0.64)[-b] | 43.05(0.65)[-b] | 49.84(2.73)[-*] | 41.83(3.36)[-b] | 24.67(0.34)[-b] | 14.76(0.07)[-b] | 21.1(0.26)[-b] |
| C7 | 262 | 50 | 45.55(1.0)[-b] | 52.02(5.6)[s] | 52.62(2.11)[b] | 52.35(2.27)[b] | 57.67(1.5)[b] | 52.6(1.18)[b] | 40.38(0.7)[-b] |
| C8 | 472 | 50 | 26.13(0.22)[-b] | 55.16(0.91)[b] | 47.01(3.71)[-b] | 40.38(1.65)[-b] | 45.29(1.94)[-b] | 48.99(5.2)[s] | 41.14(2.6)[-b] |
| C9 | 91 | 50 | 49.63(1.59)[-b] | 56.46(2.26)[b] | 59.12(6.12)[b] | 46.6(2.06)[-b] | 29.03(2.98)[-b] | 8.62(1.08)[-b] | 14.33(2.11)[-b] |
| Ranking | | 2 | 3 | 1 | 2 | 3 | 4 | 5 | 5 |

| Dataset | #Time Steps | Dummy | WP-ORB | AIO-ORB combined | AIO-ORB proprietary | Filter-ORB combined | OP-AIO-ORB combined | OP-AIO-ORB proprietary | OP-Filter-ORB combined |
|---------|-------------|-------|-----------------|------------------|---------------------|---------------------|---------------------|------------------------|------------------------|
| C1 | 121 | 50 | 48.52(2.36)[-b] | 46.19(4.31)[-b] | 37.82(3.52)[-b] | 54.01(3.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 19 | 50 | 49.05(1.07)[-b] | 61.06(1.23)[b] | 47.53(6.61)[-b] | 58.16(0.66)[b] | 36.04(3.44)[-b] | 35.0(13.46)[-b] | 37.85(3.21)[-b] |
| C3 | 21 | 50 | 61.11(1.36)[b] | 60.52(1.53)[b] | 62.7(1.89)[b] | 61.9(1.38)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 248 | 50 | 52.68(1.35)[b] | 41.8(5.78)[-b] | 46.85(3.7)[-b] | 56.44(1.82)[b] | 53.88(2.62)[b] | 43.79(1.04)[-b] | 66.78(4.59)[b] |
| C5 | 26 | 50 | 42.31(8.55)[-b] | 56.2(1.44)[b] | 56.09(1.8)[b] | 54.92(3.15)[b] | 57.76(7.76)[b] | 47.64(4.51)[-b] | 53.59(5.68)[m] |
| C6 | 281 | 50 | 51.77(0.51)[b] | 53.76(2.24)[b] | 52.25(3.62)[b] | 51.32(1.58)[b] | 62.91(2.09)[b] | 49.96(2.24)[s] | 54.29(1.24)[b] |
| C7 | 399 | 50 | 46.48(2.58)[-b] | 47.52(8.24)[-s] | 56.43(6.65)[b] | 60.96(2.67)[b] | 31.51(15.32)[-b] | 35.19(10.53)[-b] | 0.0(0.0)[-b] |
| C8 | 486 | 50 | 32.08(0.62)[-b] | 45.61(11.25)[*] | 42.57(13.67)[*] | 48.0(2.75)[-b] | 49.95(5.09)[-s] | 40.18(6.55)[-b] | 48.24(2.34)[-b] |
| C9 | 473 | 50 | 34.92(0.48)[-b] | 38.53(2.39)[-b] | 40.0(4.65)[-b] | 34.07(1.49)[-b] | 36.98(6.48)[-b] | 2.06(4.74)[-b] | 18.38(0.0)[-b] |
| Ranking | | 2 | 3 | 2 | 2 | 1 | 4 | 4 | 4 |

Standard deviations are shown in brackets. Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the corresponding Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the Dummy approach. Smaller Scott-Knott.BA12 rankings are better rankings.

TABLE 7: Overall predictive performance, A12 effect sizes and Scott-Knott.BA12 statistical tests to compare learning approaches for open source data

| Dataset | Approach | Recall0 | Recall1 | G-Mean |
|--------------------|--------------|-----------------|------------------|-----------------|
| Tomcat | Dummy | 50 | 50 | 50 |
| | WP-OOB | 58.25(1.85)[b] | 63.59(1.56)[b] | 57.78(0.62)[b] |
| | AIO-OOB | 61.07(1.35)[b] | 63.76(1.15)[b] | 59.85(0.48)[b] |
| | Filter-OOB | 66.99(1.24)[b] | 60.05(0.99)[b] | 61.57(0.53)[b] |
| | Ensemble-OOB | 63.78(0.71)[b] | 43.06(0.74)[-b] | 49.12(0.2)[-b] |
| | OP-AIO-OOB | 22.09(4.78)[-b] | 85.86(2.49)[b] | 40.44(5.71)[-b] |
| Jgroups | Dummy | 50 | 50 | 50 |
| | WP-OOB | 57.55(1.71)[b] | 59.01(1.53)[b] | 54.76(0.69)[b] |
| | AIO-OOB | 64.54(1.6)[b] | 52.85(1.48)[b] | 54.77(0.65)[b] |
| | Filter-OOB | 65.94(1.34)[b] | 51.59(1.24)[b] | 54.68(0.43)[b] |
| | Ensemble-OOB | 84.03(0.4)[b] | 18.6(0.37)[-b] | 35.51(0.3)[-b] |
| | OP-AIO-OOB | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | Dummy | 50 | 50 | 50 |
| | WP-OOB | 61.13(1.67)[b] | 56.09(1.33)[b] | 48.41(0.71)[-b] |
| | AIO-OOB | 65.32(0.62)[b] | 61.57(0.88)[b] | 60.62(0.29)[b] |
| | Filter-OOB | 69.82(0.76)[b] | 59.96(1.21)[b] | 62.38(0.69)[b] |
| | Ensemble-OOB | 73.09(0.58)[b] | 37.6(0.64)[-b] | 49.4(0.28)[-b] |
| | OP-AIO-OOB | 47.2(9.1)* | 70.94(7.57)[b] | 51.79(6.1)[m] |
| Camel | Dummy | 50 | 50 | 50 |
| | WP-OOB | 56.04(1.25)[b] | 74.62(0.97)[b] | 62.45(0.58)[b] |
| | AIO-OOB | 53.89(1.15)[b] | 75.29(0.73)[b] | 62.16(0.62)[b] |
| | Filter-OOB | 55.52(0.82)[b] | 74.29(0.72)[b] | 62.66(0.56)[b] |
| | Ensemble-OOB | 55.25(0.88)[b] | 55.17(0.78)[b] | 52.63(0.25)[b] |
| | OP-AIO-OOB | 25.72(2.44)[-b] | 82.82(3.08)[b] | 43.71(1.43)[-b] |
| Brackets | Dummy | 50 | 50 | 50 |
| | WP-OOB | 49.16(0.28)[-b] | 89.6(0.24)[b] | 64.03(0.11)[b] |
| | AIO-OOB | 65.47(1.01)[b] | 79.68(1.32)[b] | 70.85(0.48)[b] |
| | Filter-OOB | 69.15(0.84)[b] | 75.44(1.31)[b] | 70.91(0.4)[b] |
| | Ensemble-OOB | 64.89(0.77)[b] | 62.81(1.13)[b] | 62.68(0.48)[b] |
| | OP-AIO-OOB | 31.33(9.76)[-b] | 89.04(2.47)[b] | 39.3(12.45)[-b] |
| Nova | Dummy | 50 | 50 | 50 |
| | WP-OOB | 68.2(0.24)[b] | 86.95(0.53)[b] | 75.55(0.21)[b] |
| | AIO-OOB | 70.04(0.32)[b] | 88.9(0.54)[b] | 78.25(0.2)[b] |
| | Filter-OOB | 70.91(0.31)[b] | 87.24(0.38)[b] | 78.01(0.18)[b] |
| | Ensemble-OOB | 75.88(0.79)[b] | 57.23(1.61)[b] | 65.5(0.61)[b] |
| | OP-AIO-OOB | 59.07(12.18)[b] | 92.65(2.2)[b] | 71.75(13.26)[b] |
| Fabric8 | Dummy | 50 | 50 | 50 |
| | WP-OOB | 50.56(2.72)[s] | 75.72(2.21)[b] | 59.75(1.01)[b] |
| | AIO-OOB | 55.91(1.01)[b] | 70.76(1.46)[b] | 60.92(0.48)[b] |
| | Filter-OOB | 61.94(1.89)[b] | 73.25(1.58)[b] | 66.55(0.63)[b] |
| | Ensemble-OOB | 48.84(1.22)[-b] | 61.42(1.39)[b] | 53.27(0.36)[b] |
| | OP-AIO-OOB | 10.88(2.39)[-b] | 96.88(1.22)[b] | 27.2(5.03)[-b] |
| Neutron | Dummy | 50 | 50 | 50 |
| | WP-OOB | 70.03(0.7)[b] | 91.81(0.61)[b] | 79.43(0.36)[b] |
| | AIO-OOB | 73.24(0.42)[b] | 91.39(0.5)[b] | 81.48(0.24)[b] |
| | Filter-OOB | 74.89(0.45)[b] | 89.63(0.54)[b] | 81.51(0.26)[b] |
| | Ensemble-OOB | 78.58(1.03)[b] | 59.78(3.32)[b] | 68.28(1.42)[b] |
| | OP-AIO-OOB | 70.23(0.95)[b] | 81.61(2.92)[b] | 74.18(1.7)[b] |
| Npm | Dummy | 50 | 50 | 50 |
| | WP-OOB | 36.99(2.19)[-b] | 75.74(1.52)[b] | 45.91(0.91)[-b] |
| | AIO-OOB | 54.8(1.61)[b] | 68.69(2.03)[b] | 59.34(0.61)[b] |
| | Filter-OOB | 55.39(1.61)[b] | 68.59(1.48)[b] | 59.68(0.54)[b] |
| | Ensemble-OOB | 54.22(1.01)[b] | 51.89(1.05)[b] | 50.42(0.29)[b] |
| | OP-AIO-OOB | 61.08(6.38)[b] | 37.31(6.76)[-b] | 35.91(2.78)[-b] |
| BroadleafCommerce | Dummy | 50 | 50 | 50 |
| | WP-OOB | 58.44(1.58)[b] | 69.82(2.13)[b] | 60.41(0.83)[b] |
| | AIO-OOB | 67.28(1.37)[b] | 72.11(1.54)[b] | 69.01(0.58)[b] |
| | Filter-OOB | 67.41(0.94)[b] | 70.53(0.98)[b] | 68.39(0.42)[b] |
| | Ensemble-OOB | 61.96(0.89)[b] | 56.48(1.08)[b] | 57.31(0.18)[b] |
| | OP-AIO-OOB | 42.99(1.3)[-b] | 71.99(2.46)[b] | 53.27(0.67)[b] |
| Ranking | Dummy | 4 | 4 | 5 |
| | WP-OOB | 3 | 1 | 2 |
| | AIO-OOB | 2 | 1 | 1 |
| | Filter-OOB | 2 | 1 | 1 |
| | Ensemble-OOB | 2 | 4 | 4 |
| | OP-AIO-OOB | 5 | 1 | 6 |
| Tomcat | Dummy | 50 | 50 | 50 |
| | WP-ORB | 58.41(1.72)[b] | 64.49(1.21)[b] | 59.78(0.83)[b] |
| | AIO-ORB | 61.98(1.19)[b] | 63.38(0.74)[b] | 61.72(0.82)[b] |
| | Filter-ORB | 61.07(0.89)[b] | 63.31(0.74)[b] | 61.25(0.48)[b] |
| | Ensemble-ORB | 61.75(1.23)[b] | 47.67(1.08)[-b] | 50.82(0.34)[b] |
| | OP-AIO-ORB | 46.92(5.41)[-b] | 54.07(4.56)[b] | 42.31(2.63)[-b] |
| Jgroups | Dummy | 50 | 50 | 50 |
| | WP-ORB | 61.18(0.71)[b] | 56.93(1.28)[b] | 57.2(0.81)[b] |
| | AIO-ORB | 62.77(0.94)[b] | 56.67(0.96)[b] | 57.92(0.66)[b] |
| | Filter-ORB | 63.05(1.07)[b] | 57.21(1.26)[b] | 58.29(0.68)[b] |
| | Ensemble-ORB | 83.72(0.33)[b] | 20.1(0.42)[-b] | 37.35(0.33)[-b] |
| | OP-AIO-ORB | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | Dummy | 50 | 50 | 50 |
| | WP-ORB | 71.92(1.0)[b] | 45.83(1.31)[-b] | 51.73(0.78)[b] |
| | AIO-ORB | 67.51(0.86)[b] | 69.48(1.05)[b] | 67.77(0.75)[b] |
| | Filter-ORB | 68.57(0.84)[b] | 69.93(0.83)[b] | 68.61(0.58)[b] |
| | Ensemble-ORB | 74.61(0.48)[b] | 39.98(0.67)[-b] | 52.4(0.54)[b] |
| | OP-AIO-ORB | 76.29(2.16)[b] | 61.45(4.22)[b] | 67.5(1.79)[b] |
| Camel | Dummy | 50 | 50 | 50 |
| | WP-ORB | 59.59(1.18)[b] | 70.11(1.15)[b] | 62.92(0.98)[b] |
| | AIO-ORB | 59.12(0.52)[b] | 73.1(0.56)[b] | 65.01(0.37)[b] |
| | Filter-ORB | 58.58(0.62)[b] | 72.7(0.53)[b] | 64.48(0.43)[b] |
| | Ensemble-ORB | 61.46(0.88)[b] | 52.33(1.13)[b] | 54.26(0.51)[b] |
| | OP-AIO-ORB | 60.0(8.59)[b] | 47.9(7.2)[-m] | 51.05(2.66)[b] |
| Brackets | Dummy | 50 | 50 | 50 |
| | WP-ORB | 61.31(0.74)[b] | 76.66(1.43)[b] | 63.31(0.49)[b] |
| | AIO-ORB | 66.9(1.03)[b] | 74.94(1.71)[b] | 69.98(0.87)[b] |
| | Filter-ORB | 66.34(0.92)[b] | 75.35(1.33)[b] | 69.78(0.74)[b] |
| | Ensemble-ORB | 61.7(1.09)[b] | 69.14(1.23)[b] | 64.71(0.75)[b] |
| | OP-AIO-ORB | 52.23(1.53)[b] | 84.23(2.0)[b] | 64.42(0.55)[b] |
| Nova | Dummy | 50 | 50 | 50 |
| | WP-ORB | 74.25(1.49)[b] | 80.57(2.64)[b] | 75.42(0.66)[b] |
| | AIO-ORB | 74.15(0.89)[b] | 81.95(1.27)[b] | 77.28(0.32)[b] |
| | Filter-ORB | 72.72(0.21)[b] | 82.85(0.54)[b] | 76.83(0.35)[b] |
| | Ensemble-ORB | 76.76(0.76)[b] | 68.42(1.8)[b] | 71.7(0.59)[b] |
| | OP-AIO-ORB | 53.97(38.63)[m] | 67.03(25.53)[b] | 37.92(27.73)* |
| Fabric8 | Dummy | 50 | 50 | 50 |
| | WP-ORB | 61.32(2.17)[b] | 67.83(1.32)[b] | 61.42(1.04)[b] |
| | AIO-ORB | 57.39(1.44)[b] | 73.78(1.42)[b] | 63.7(1.03)[b] |
| | Filter-ORB | 64.14(0.62)[b] | 71.8(1.09)[b] | 67.42(0.42)[b] |
| | Ensemble-ORB | 49.61(1.24)[-s] | 65.68(1.22)[b] | 55.91(0.59)[b] |
| | OP-AIO-ORB | 72.26(5.51)[b] | 41.86(8.81)[-b] | 50.3(5.86)[-s] |
| Neutron | Dummy | 50 | 50 | 50 |
| | WP-ORB | 80.47(1.71)[b] | 79.52(1.88)[b] | 79.52(0.58)[b] |
| | AIO-ORB | 75.34(0.57)[b] | 87.53(0.89)[b] | 80.72(0.39)[b] |
| | Filter-ORB | 74.01(0.81)[b] | 88.98(1.02)[b] | 80.67(0.54)[b] |
| | Ensemble-ORB | 77.01(1.55)[b] | 72.02(3.41)[b] | 74.28(1.1)[b] |
| | OP-AIO-ORB | 70.35(18.03)[b] | 66.39(7.53)[b] | 66.55(12.39)[b] |
| Npm | Dummy | 50 | 50 | 50 |
| | WP-ORB | 54.93(1.17)[b] | 63.67(1.32)[b] | 53.81(0.95)[b] |
| | AIO-ORB | 50.95(1.56)[b] | 76.91(1.92)[b] | 60.03(1.11)[b] |
| | Filter-ORB | 51.23(4.02)[-s] | 73.26(3.94)[b] | 59.15(1.94)[b] |
| | Ensemble-ORB | 52.48(1.25)[b] | 56.11(0.97)[b] | 52.41(0.58)[b] |
| | OP-AIO-ORB | 73.38(16.16)[b] | 44.16(24.46)[-s] | 47.99(19.35)[b] |
| BroadleafCommerce | Dummy | 50 | 50 | 50 |
| | WP-ORB | 59.43(2.96)[b] | 65.93(3.01)[b] | 59.72(1.29)[b] |
| | AIO-ORB | 66.78(0.85)[b] | 70.12(1.02)[b] | 67.85(0.58)[b] |
| | Filter-ORB | 66.83(0.64)[b] | 69.2(1.48)[b] | 67.51(0.85)[b] |
| | Ensemble-ORB | 61.57(1.2)[b] | 59.13(1.46)[b] | 58.94(0.45)[b] |
| | OP-AIO-ORB | 84.29(4.22)[b] | 26.63(7.15)[-b] | 44.42(7.7)[-b] |
| Ranking | Dummy | 4 | 4 | 5 |
| | WP-ORB | 2 | 2 | 2 |
| | AIO-ORB | 2 | 1 | 1 |
| | Filter-ORB | 2 | 1 | 1 |
| | Ensemble-ORB | 2 | 4 | 3 |
| | OP-AIO-ORB | 5 | 1 | 5 |

Standard deviations are shown in brackets. Symbols [*, [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the Dummy approach. Scott-Knott.BA12 was run for all OOB- and ORB-based approaches together. The groups’ rankings retrieved by Scott-Knott.BA12 are shown in the ranking rows, with smaller numbers indicating better rankings.

The recalls of the approaches across data sets are influenced by trade-offs between Recall0 and Recall1, resulting in several approaches obtaining the same best rank across datasets. This is because a given approach sometimes performs better in terms of Recall0 and sometimes in terms of Recall1, resulting overall in the same rank. However, given the G-Mean results, which combine Recall0 and Recall1, the trade-offs between recalls obtained by AIO and Filtering were better than those obtained by WP, Ensemble and Offline approaches.

TABLE 8: Overall predictive performance, A12 effect sizes and Scott-Knott.BA12 statistical tests to compare learning approaches for proprietary data

| Dataset | Approach | Recall0 | Recall1 | G-Mean | Dataset | Approach | Recall0 | Recall1 | G-Mean | | |
|------------------------|------------------------|------------------|------------------|------------------------|------------------------|------------------------|------------------|-----------------|------------------|-----------------|-----------------|
| C1 | Dummy | 50 | 50 | 50 | C1 | Dummy | 50 | 50 | 50 | | |
| | WP-OOB | 59.25(4.79)[b] | 39.37(3.88)[-b] | 32.65(1.53)[-b] | | WP-ORB | 64.69(5.61)[b] | 30.75(3.94)[-b] | 31.85(2.57)[-b] | | |
| | AIO-OOB-combined | 51.6(3.65)[b] | 48.3(2.55)[-b] | 40.28(1.53)[-b] | | AIO-ORB-combined | 45.4(6.03)[-b] | 51.85(6.11)[b] | 34.08(2.4)[-b] | | |
| | AIO-OOB-proprietary | 31.75(8.46)[-b] | 65.14(7.2)[b] | 35.84(3.45)[-b] | | AIO-ORB-proprietary | 35.64(1.4)[-b] | 57.67(1.51)[b] | 27.59(1.45)[-b] | | |
| | Filter-OOB-combined | 53.56(3.43)[b] | 49.18(3.1)[-s] | 45.25(1.17)[-b] | | Filter-ORB-combined | 69.56(8.37)[b] | 37.95(8.06)[-b] | 40.1(2.02)[-b] | | |
| | OP-AIO-OOB-combined | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | OP-AIO-ORB-combined | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | |
| | OP-AIO-OOB-proprietary | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | OP-AIO-ORB-proprietary | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | |
| | OP-Filter-OOB-combined | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | OP-Filter-ORB-combined | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | |
| | C2 | Dummy | 50 | 50 | | 50 | C2 | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 16.33(0.12)[-b] | 87.15(0.0)[b] | | 25.95(0.37)[-b] | | WP-ORB | 71.46(1.01)[b] | 17.49(0.83)[-b] | 17.7(0.51)[-b] |
| AIO-OOB-combined | | 22.23(0.53)[-b] | 76.41(0.42)[b] | 31.76(0.45)[-b] | AIO-ORB-combined | 35.88(3.47)[-b] | | 60.89(3.72)[b] | 39.41(3.86)[-b] | | |
| AIO-OOB-proprietary | | 45.82(7.4)[-b] | 49.25(10.34)[m] | 33.47(5.27)[-b] | AIO-ORB-proprietary | 38.15(6.8)[-b] | | 55.99(9.41)[b] | 39.58(3.32)[-b] | | |
| Filter-OOB-combined | | 45.22(1.27)[-b] | 63.81(1.77)[b] | 47.24(0.74)[-b] | Filter-ORB-combined | 46.15(0.69)[-b] | | 43.99(0.7)[-b] | 29.53(0.75)[-b] | | |
| OP-AIO-OOB-combined | | 43.5(4.07)[-b] | 55.23(3.49)[b] | 43.6(1.86)[-b] | OP-AIO-ORB-combined | 82.97(2.67)[b] | | 25.75(4.38)[-b] | 41.74(2.98)[-b] | | |
| OP-AIO-OOB-proprietary | | 98.4(2.76)[b] | 1.54(1.88)[-b] | 5.53(6.16)[-b] | OP-AIO-ORB-proprietary | 87.63(8.27)[b] | | 15.2(9.77)[-b] | 30.58(12.08)[-b] | | |
| OP-Filter-OOB-combined | | 47.57(5.6)[-s] | 55.6(4.04)[b] | 45.06(2.91)[-b] | OP-Filter-ORB-combined | 89.53(3.42)[b] | | 23.13(6.32)[-b] | 40.83(4.99)[-b] | | |
| C3 | | Dummy | 50 | 50 | 50 | C3 | | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 59.78(4.95)[b] | 38.31(4.5)[-b] | 39.57(1.1)[-b] | | | WP-ORB | 58.67(1.05)[b] | 49.83(0.93)[-* | 39.8(0.65)[-b] |
| | AIO-OOB-combined | 41.52(0.84)[-b] | 69.57(1.43)[b] | 44.83(0.71)[-b] | AIO-ORB-combined | | 57.37(0.69)[b] | 50.82(1.05)[b] | 39.31(0.47)[-b] | | |
| | AIO-OOB-proprietary | 26.4(2.25)[-b] | 66.45(3.09)[b] | 36.85(1.35)[-b] | AIO-ORB-proprietary | | 60.62(2.22)[b] | 48.21(2.01)[-b] | 40.49(1.24)[-b] | | |
| | Filter-OOB-combined | 52.59(1.18)[b] | 59.66(0.99)[b] | 45.05(0.82)[-b] | Filter-ORB-combined | | 63.24(2.13)[b] | 47.2(1.57)[-b] | 41.69(0.94)[-b] | | |
| | OP-AIO-OOB-combined | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | OP-AIO-ORB-combined | | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | |
| | OP-AIO-OOB-proprietary | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | OP-AIO-ORB-proprietary | | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | |
| | OP-Filter-OOB-combined | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | OP-Filter-ORB-combined | | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | |
| | C4 | Dummy | 50 | 50 | 50 | | C4 | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 21.56(0.74)[-b] | 81.96(0.66)[b] | 33.75(0.86)[-b] | | | WP-ORB | 68.92(4.34)[b] | 39.78(6.02)[-b] | 39.52(1.0)[-b] |
| AIO-OOB-combined | | 37.79(1.76)[-b] | 76.64(1.34)[b] | 49.57(1.45)[-s] | AIO-ORB-combined | 29.51(7.55)[-b] | | 78.37(5.51)[b] | 40.0(4.41)[-b] | | |
| AIO-OOB-proprietary | | 66.59(6.0)[b] | 55.44(8.47)[m] | 56.52(3.02)[b] | AIO-ORB-proprietary | 36.53(3.6)[-b] | | 77.1(1.65)[b] | 46.04(2.97)[-b] | | |
| Filter-OOB-combined | | 61.15(3.35)[b] | 54.41(2.64)[b] | 46.88(2.53)[-b] | Filter-ORB-combined | 75.46(2.74)[b] | | 39.43(2.26)[-b] | 47.36(2.04)[-b] | | |
| OP-AIO-OOB-combined | | 50.98(1.91)[b] | 61.59(1.59)[b] | 52.9(0.9)[b] | OP-AIO-ORB-combined | 49.04(7.88)[-s] | | 63.83(10.56)[b] | 50.59(2.3)[b] | | |
| OP-AIO-OOB-proprietary | | 40.49(3.43)[-b] | 62.7(3.37)[b] | 44.31(1.36)[-b] | OP-AIO-ORB-proprietary | 36.33(0.91)[-b] | | 71.6(0.54)[b] | 43.52(0.84)[-b] | | |
| OP-Filter-OOB-combined | | 49.25(0.0)[-b] | 61.33(0.0)[b] | 41.71(0.0)[-b] | OP-Filter-ORB-combined | 89.73(1.63)[b] | | 40.73(5.87)[-b] | 53.77(3.53)[b] | | |
| C5 | | Dummy | 50 | 50 | 50 | C5 | | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 11.93(0.67)[-b] | 85.44(1.43)[b] | 22.91(0.42)[-b] | | | WP-ORB | 95.92(0.64)[b] | 5.34(1.4)[-b] | 12.17(1.62)[-b] |
| | AIO-OOB-combined | 34.86(1.67)[-b] | 67.71(3.08)[b] | 45.44(1.35)[-b] | AIO-ORB-combined | | 12.04(1.74)[-b] | 91.14(0.66)[b] | 26.34(3.69)[-b] | | |
| | AIO-OOB-proprietary | 22.52(1.1)[-b] | 90.05(1.31)[b] | 42.33(0.98)[b] | AIO-ORB-proprietary | | 12.84(2.2)[-b] | 90.98(2.36)[b] | 27.9(3.7)[-b] | | |
| | Filter-OOB-combined | 38.55(2.54)[-b] | 67.07(4.08)[b] | 48.94(1.67)[-b] | Filter-ORB-combined | | 21.94(3.44)[-b] | 80.9(2.82)[b] | 37.26(3.41)[-b] | | |
| | OP-AIO-OOB-combined | 76.39(3.78)[b] | 26.8(5.51)[-b] | 43.17(5.15)[-b] | OP-AIO-ORB-combined | | 47.65(5.97)[-s] | 65.06(8.41)[b] | 52.84(4.12)[m] | | |
| | OP-AIO-OOB-proprietary | 42.76(5.1)[-b] | 67.8(4.34)[b] | 48.35(2.86)[-m] | OP-AIO-ORB-proprietary | | 35.59(10.02)[-b] | 74.15(9.73)[b] | 47.86(4.22)[-s] | | |
| | OP-Filter-OOB-combined | 23.54(6.53)[-b] | 82.43(6.63)[b] | 41.11(4.03)[-b] | OP-Filter-ORB-combined | | 44.7(4.24)[-b] | 72.84(3.68)[b] | 54.77(2.56)[b] | | |
| | C6 | Dummy | 50 | 50 | 50 | | C6 | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 10.98(0.58)[-b] | 87.91(0.56)[b] | 23.23(1.17)[-b] | | | WP-ORB | 63.74(0.42)[b] | 42.34(0.41)[-b] | 31.08(0.35)[-b] |
| AIO-OOB-combined | | 33.15(1.82)[-b] | 69.53(1.63)[b] | 42.06(1.29)[-b] | AIO-ORB-combined | 52.93(5.44)[b] | | 51.18(5.37)[s] | 45.16(1.91)[-b] | | |
| AIO-OOB-proprietary | | 40.29(10.26)[-b] | 64.4(9.58)[b] | 44.79(4.18)[-b] | AIO-ORB-proprietary | 52.44(5.1)[*] | | 55.29(2.69)[b] | 42.47(3.27)[-b] | | |
| Filter-OOB-combined | | 35.22(2.81)[-b] | 69.57(2.4)[b] | 42.47(1.85)[-b] | Filter-ORB-combined | 56.47(6.7)[m] | | 43.37(7.2)[-m] | 42.15(1.66)[-b] | | |
| OP-AIO-OOB-combined | | 48.6(1.25)[-b] | 56.66(0.81)[b] | 39.62(0.37)[-b] | OP-AIO-ORB-combined | 74.96(2.93)[b] | | 43.37(2.3)[-b] | 50.27(1.79)[s] | | |
| OP-AIO-OOB-proprietary | | 70.97(0.35)[b] | 30.75(0.28)[-b] | 33.3(0.15)[-b] | OP-AIO-ORB-proprietary | 76.08(8.34)[b] | | 27.62(7.76)[-b] | 35.15(1.62)[-b] | | |
| OP-Filter-OOB-combined | | 54.93(1.23)[b] | 48.66(0.77)[-b] | 40.49(0.42)[-b] | OP-Filter-ORB-combined | 61.99(7.35)[b] | | 46.38(7.38)[*] | 40.6(1.39)[-b] | | |
| C7 | | Dummy | 50 | 50 | 50 | C7 | | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 22.78(0.75)[-b] | 83.52(0.51)[b] | 33.94(0.92)[-b] | | | WP-ORB | 48.03(5.25)[-b] | 56.78(3.82)[b] | 39.01(2.53)[-b] |
| | AIO-OOB-combined | 38.25(12.04)[-b] | 70.7(7.58)[b] | 46.62(5.4)[-m] | AIO-ORB-combined | | 34.47(12.93)[-b] | 77.01(9.77)[b] | 44.91(7.51)[-b] | | |
| | AIO-OOB-proprietary | 23.12(1.65)[-b] | 93.33(1.26)[b] | 43.06(1.67)[-b] | AIO-ORB-proprietary | | 45.97(11.53)[s] | 71.25(7.13)[b] | 53.99(6.85)[m] | | |
| | Filter-OOB-combined | 44.7(2.1)[-b] | 61.98(3.79)[b] | 43.96(1.28)[-b] | Filter-ORB-combined | | 64.89(6.69)[b] | 56.2(3.68)[b] | 55.47(3.0)[b] | | |
| | OP-AIO-OOB-combined | 77.52(2.22)[b] | 38.87(2.4)[-b] | 53.38(1.97)[b] | OP-AIO-ORB-combined | | 56.03(38.06)[*] | 50.6(37.75)[*] | 32.64(14.9)[-b] | | |
| | OP-AIO-OOB-proprietary | 34.57(1.17)[-b] | 69.48(2.97)[b] | 43.4(0.71)[-b] | OP-AIO-ORB-proprietary | | 35.77(18.96)[-b] | 61.18(21.41)[b] | 35.55(7.7)[-b] | | |
| | OP-Filter-OOB-combined | 26.99(0.75)[-b] | 65.72(0.78)[b] | 33.73(0.52)[-b] | OP-Filter-ORB-combined | | 100.0(0.0)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | | |
| | C8 | Dummy | 50 | 50 | 50 | | C8 | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 13.04(0.1)[-b] | 83.1(0.01)[b] | 18.12(0.43)[-b] | | | WP-ORB | 44.61(0.43)[-b] | 43.37(0.59)[-b] | 22.02(0.47)[-b] |
| AIO-OOB-combined | | 46.57(1.22)[-b] | 60.3(1.79)[b] | 48.97(0.81)[-b] | AIO-ORB-combined | 37.94(14.26)[-b] | | 70.69(10.81)[b] | 45.89(8.87)[-s] | | |
| AIO-OOB-proprietary | | 29.57(4.44)[-b] | 74.84(4.67)[b] | 42.21(4.03)[-b] | AIO-ORB-proprietary | 37.06(16.4)[-b] | | 72.45(12.06)[b] | 44.29(10.43)[-s] | | |
| Filter-OOB-combined | | 46.6(1.38)[-b] | 48.41(1.65)[-b] | 38.35(1.15)[-b] | Filter-ORB-combined | 55.72(6.29)[b] | | 44.31(4.55)[-b] | 44.55(2.25)[-b] | | |
| OP-AIO-OOB-combined | | 66.69(2.8)[b] | 34.9(2.52)[-b] | 43.53(1.52)[-b] | OP-AIO-ORB-combined | 74.41(8.76)[b] | | 37.11(8.19)[-b] | 49.87(5.01)[-s] | | |
| OP-AIO-OOB-proprietary | | 77.63(10.75)[b] | 32.16(10.48)[-b] | 47.17(6.24)[*] | OP-AIO-ORB-proprietary | 89.26(3.52)[b] | | 24.95(5.54)[-b] | 43.5(5.18)[-b] | | |
| OP-Filter-OOB-combined | | 24.94(3.74)[-b] | 74.12(2.55)[b] | 33.53(4.49)[-b] | OP-Filter-ORB-combined | 79.73(3.35)[b] | | 29.25(2.94)[-b] | 45.84(2.09)[-b] | | |
| C9 | | Dummy | 50 | 50 | 50 | C9 | | Dummy | 50 | 50 | 50 |
| | | WP-OOB | 20.64(1.34)[-b] | 86.52(0.42)[b] | 32.19(1.44)[-b] | | | WP-ORB | 45.65(0.25)[-b] | 50.03(0.09)[b] | 24.88(0.36)[-b] |
| | AIO-OOB-combined | 44.53(2.53)[-b] | 69.19(3.39)[b] | 48.8(1.44)[-b] | AIO-ORB-combined | | 39.34(2.08)[-b] | 63.49(1.75)[b] | 35.87(1.7)[-b] | | |
| | AIO-OOB-proprietary | 59.51(4.62)[b] | 52.23(7.1)[s] | 49.04(2.87)[-s] | AIO-ORB-proprietary | | 44.37(5.62)[-m] | 61.6(4.38)[b] | 38.46(3.49)[-b] | | |
| | Filter-OOB-combined | 52.2(2.63)[b] | 58.09(2.72)[b] | 47.67(1.32)[-b] | Filter-ORB-combined | | 36.93(1.02)[-b] | 64.24(1.19)[b] | 29.96(1.05)[-b] | | |
| | OP-AIO-OOB-combined | 59.8(7.99)[b] | 41.63(6.31)[-b] | 37.1(4.46)[-b] | OP-AIO-ORB-combined | | 84.62(5.2)[b] | 15.86(4.1)[-b] | 32.13(4.76)[-b] | | |
| | OP-AIO-OOB-proprietary | 93.81(5.79)[b] | 13.24(5.36)[-b] | 28.35(3.15)[-b] | OP-AIO-ORB-proprietary | | 99.9(0.31)[b] | 0.29(0.8)[-b] | 1.82(4.42)[-b] | | |
| | OP-Filter-OOB-combined | 52.0(0.16)[b] | 46.24(0.5)[-b] | 28.67(0.68)[-b] | OP-Filter-ORB-combined | | 99.96(0.0)[b] | 2.71(0.0)[-b] | 14.61(0.0)[-b] | | |
| | Ranking | Dummy | 2 | 4 | 1 | | Ranking | Dummy | 2 | 4 | 1 |
| | | WP-OOB | 5 | 1 | 5 | | | WP-ORB | 1 | 5 | 5 |
| AIO-OOB-combined | | 4 | 2 | 2 | AIO-ORB-combined | 4 | | 2 | 4 | | |
| AIO-OOB-proprietary | | 1 | 2 | 3 | AIO-ORB-proprietary | 2 | | 4 | 3 | | |
| Filter-OOB-combined | | 3 | 3 | 2 | Filter-ORB-combined | 2 | | 2 | 3 | | |
| OP-AIO-OOB-combined | | 1 | 5 | 3 | OP-AIO-ORB-combined | 1 | | 5 | 3 | | |
| OP-AIO-OOB-proprietary | | 1 | 5 | 5 | OP-AIO-ORB-proprietary | 1 | | 5 | 5 | | |
| OP-Filter-OOB-combined | 2 | 4 | 5 | OP-Filter-ORB-combined | 1 | 6 | 4 | | | | |

Standard deviations are shown in brackets. Symbols [*], [s], [m] and [b] represent insignificant, small, medium and large A12 effect size against the Dummy approach. Presence/absence of the sign “-” in the effect size means that the corresponding approach was worse/better than the corresponding Dummy approach. Scott-Knott.BA12 was run for all OOB- and ORB-based approaches together. The groups’ rankings retrieved by Scott-Knott.BA12 are shown in the ranking rows, with smaller numbers indicating better rankings.

TABLE 9: Number of initial time steps of the initial phase, and average Recall0, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Opensource data

| Dataset | #Time Steps | Dummy | WP-OOB_r0 | AIO-OOB_r0 | Filter-OOB_r0 | Ensemble-OOB_r0 | OP-AIO-OOB_r0 | OP-Filter-OOB_r0 |
|--------------------|-------------|-------|-----------------|----------------|-----------------|-----------------|------------------|------------------|
| Tomcat | 2006 | 50 | 76.66(5.97)[b] | 49.86(1.86)[*] | 61.48(2.47)[b] | 87.03(0.61)[b] | 49.68(8.93)[-s] | 50.51(4.36)[-s] |
| JGroups | 1268 | 50 | 64.66(0.69)[b] | 64.54(0.6)[b] | 63.9(1.0)[b] | 96.41(0.1)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| Spring-Integration | 461 | 50 | 44.11(0.12)[-b] | 82.6(0.65)[b] | 89.71(0.87)[b] | 93.25(0.28)[b] | 94.09(1.18)[b] | 93.59(1.02)[b] |
| Camel | 3112 | 50 | 28.71(3.01)[-b] | 55.13(1.48)[b] | 56.56(1.34)[b] | 84.71(0.53)[b] | 38.99(1.72)[-b] | 38.12(1.2)[-b] |
| Brackets | 1569 | 50 | 15.37(0.04)[-b] | 63.02(1.8)[b] | 65.67(0.85)[b] | 79.59(0.64)[b] | 58.72(2.62)[b] | 61.6(1.36)[b] |
| Nova | 6271 | 50 | 40.64(0.34)[-b] | 56.53(1.1)[b] | 59.92(0.65)[b] | 72.22(0.93)[b] | 49.63(3.54)[-b] | 66.27(4.6)[b] |
| Fabric8 | 795 | 50 | 20.81(0.24)[-b] | 82.77(1.63)[b] | 74.36(2.21)[b] | 75.59(1.81)[b] | 48.5(5.7)[-*] | 72.31(3.69)[b] |
| Neutron | 917 | 50 | 31.13(4.78)[-b] | 76.39(2.09)[b] | 83.57(2.48)[b] | 78.45(1.24)[b] | 55.59(0.31)[b] | 67.5(1.98)[b] |
| Npm | 2536 | 50 | 13.73(1.44)[-b] | 44.8(4.75)[-b] | 37.43(4.23)[-b] | 66.7(0.68)[b] | 45.28(12.29)[-*] | 14.74(0.97)[-b] |
| BroadleafCommerce | 677 | 50 | 28.24(0.2)[-b] | 44.06(2.5)[-b] | 52.74(3.63)[b] | 81.64(0.44)[b] | 38.71(7.36)[-b] | 97.11(0.84)[b] |
| Ranking | | 5 | 6 | 3 | 2 | 1 | 4 | 2 |

| Dataset | #Time Steps | Dummy | WP-ORB_r0 | AIO-ORB_r0 | Filter-ORB_r0 | Ensemble-ORB_r0 | OP-AIO-ORB_r0 | OP-Filter-ORB_r0 |
|--------------------|-------------|-------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|
| Tomcat | 926 | 50 | 44.26(2.49)[-b] | 39.42(2.27)[-b] | 42.81(2.36)[-b] | 85.91(0.24)[b] | 14.05(3.88)[-b] | 33.08(6.91)[-b] |
| JGroups | 1412 | 50 | 60.91(1.0)[b] | 61.11(1.04)[b] | 62.73(1.86)[b] | 96.21(0.18)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| Spring-Integration | 900 | 50 | 60.2(1.08)[b] | 51.5(4.09)[b] | 54.75(5.34)[b] | 78.73(1.67)[b] | 64.58(2.43)[b] | 65.41(3.4)[b] |
| Camel | 5111 | 50 | 51.18(2.63)[b] | 48.79(1.86)[-b] | 47.72(2.3)[-b] | 81.14(0.78)[b] | 49.09(10.16)[-b] | 47.19(10.24)[-s] |
| Brackets | 1721 | 50 | 90.26(0.58)[b] | 56.53(3.0)[b] | 57.46(1.31)[b] | 66.97(2.0)[b] | 60.73(1.59)[b] | 69.53(3.63)[b] |
| Nova | 6271 | 50 | 59.67(2.78)[b] | 54.18(1.34)[b] | 52.18(1.44)[b] | 78.64(1.01)[b] | 63.48(33.59)[m] | 71.91(13.46)[b] |
| Fabric8 | 1613 | 50 | 68.33(1.0)[b] | 64.01(1.22)[b] | 63.02(2.22)[b] | 63.35(1.33)[b] | 50.32(6.14)[-s] | 67.42(8.41)[b] |
| Neutron | 3304 | 50 | 80.65(1.69)[b] | 78.67(2.11)[b] | 78.83(4.22)[b] | 75.2(1.11)[b] | 72.37(16.77)[b] | 75.19(9.22)[b] |
| Npm | 1494 | 50 | 34.45(1.91)[-b] | 27.68(5.38)[-b] | 35.37(9.1)[-b] | 70.43(2.04)[b] | 71.36(24.54)[b] | 100.0(0.0)[b] |
| BroadleafCommerce | 950 | 50 | 66.31(0.51)[b] | 48.84(2.41)[-m] | 49.12(2.88)[-s] | 81.0(0.95)[b] | 61.85(3.31)[b] | 73.71(3.41)[b] |
| Ranking | | 5 | 3 | 4 | 4 | 1 | 3 | 2 |

TABLE 10: Number of initial time steps of the initial phase, and average Recall1, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Opensource data

| Dataset | #Time Steps | Dummy | WP-OOB_r1 | AIO-OOB_r1 | Filter-OOB_r1 | Ensemble-OOB_r1 | OP-AIO-OOB_r1 | OP-Filter-OOB_r1 |
|--------------------|-------------|-------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|
| Tomcat | 2006 | 50 | 29.86(7.28)[-b] | 68.53(1.27)[b] | 53.51(3.0)[b] | 13.92(0.61)[-b] | 68.46(12.11)[b] | 67.95(4.63)[b] |
| JGroups | 1268 | 50 | 33.4(1.01)[-b] | 33.2(1.08)[-b] | 34.69(1.18)[-b] | 3.47(0.14)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 461 | 50 | 46.42(0.12)[-b] | 17.44(0.21)[-b] | 23.35(0.85)[-b] | 5.84(0.18)[-b] | 4.25(2.27)[-b] | 16.06(3.65)[-b] |
| Camel | 3112 | 50 | 86.83(2.57)[b] | 66.32(0.85)[b] | 63.64(1.3)[b] | 18.93(0.63)[-b] | 73.69(1.78)[b] | 75.82(0.71)[b] |
| Brackets | 1569 | 50 | 90.16(0.02)[b] | 72.24(4.03)[b] | 72.41(1.55)[b] | 30.66(0.95)[-b] | 57.13(2.92)[b] | 65.82(2.26)[b] |
| Nova | 6271 | 50 | 89.01(0.2)[b] | 77.39(0.59)[b] | 74.26(0.65)[b] | 38.58(0.76)[-b] | 84.05(5.58)[b] | 63.43(8.28)[b] |
| Fabric8 | 795 | 50 | 82.23(0.0)[b] | 35.07(2.2)[-b] | 52.17(3.01)[b] | 23.28(1.07)[-b] | 79.12(4.46)[b] | 38.91(2.74)[-b] |
| Neutron | 917 | 50 | 87.54(1.87)[b] | 71.63(3.0)[b] | 55.22(4.71)[b] | 39.72(2.15)[-b] | 82.38(0.61)[b] | 81.75(1.46)[b] |
| Npm | 2536 | 50 | 83.75(1.66)[b] | 64.47(5.55)[b] | 66.9(3.34)[b] | 31.45(0.61)[-b] | 54.4(13.11)[*] | 82.11(1.26)[b] |
| BroadleafCommerce | 677 | 50 | 68.1(0.27)[b] | 68.31(1.68)[b] | 61.71(3.15)[b] | 19.0(1.74)[-b] | 57.57(6.4)[b] | 15.62(3.59)[-b] |
| Ranking | | 3 | 1 | 2 | 2 | 4 | 2 | 2 |

| Dataset | #Time Steps | Dummy | WP-ORB_r1 | AIO-ORB_r1 | Filter-ORB_r1 | Ensemble-ORB_r1 | OP-AIO-ORB_r1 | OP-Filter-ORB_r1 |
|--------------------|-------------|-------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|
| Tomcat | 926 | 50 | 62.76(3.14)[b] | 69.21(2.15)[b] | 69.17(1.55)[b] | 15.28(0.42)[-b] | 89.31(3.92)[b] | 71.27(3.18)[b] |
| JGroups | 1412 | 50 | 33.59(1.41)[-b] | 34.15(1.75)[-b] | 36.92(2.56)[-b] | 3.45(0.22)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 900 | 50 | 33.67(0.71)[-b] | 75.53(2.76)[b] | 76.83(4.92)[b] | 23.52(1.65)[-b] | 65.55(4.19)[b] | 49.04(3.92)[-s] |
| Camel | 5111 | 50 | 64.0(2.48)[b] | 73.87(1.39)[b] | 74.77(1.38)[b] | 22.62(0.92)[b] | 60.05(5.27)[b] | 61.62(7.17)[b] |
| Brackets | 1721 | 50 | 12.52(0.64)[-b] | 83.0(4.49)[b] | 81.15(1.65)[b] | 41.78(1.67)[-b] | 72.56(2.48)[b] | 56.79(6.54)[b] |
| Nova | 6271 | 50 | 67.77(2.12)[b] | 82.87(1.83)[b] | 82.68(0.96)[b] | 34.71(1.46)[-b] | 48.24(29.93)[-s] | 50.51(15.16)[-s] |
| Fabric8 | 1613 | 50 | 29.88(0.81)[-b] | 66.59(4.18)[b] | 66.66(2.39)[b] | 38.18(1.22)[-b] | 72.86(5.75)[b] | 52.85(5.62)[b] |
| Neutron | 3304 | 50 | 66.62(3.15)[b] | 73.45(3.72)[b] | 73.2(5.88)[b] | 68.09(1.99)[b] | 60.24(11.34)[b] | 69.18(18.12)[b] |
| Npm | 1494 | 50 | 64.79(1.01)[b] | 74.2(6.94)[b] | 66.41(8.14)[b] | 29.39(2.14)[-b] | 38.91(27.87)[-s] | 0.0(0.0)[-b] |
| BroadleafCommerce | 950 | 50 | 32.16(0.49)[-b] | 66.7(5.17)[b] | 65.56(6.86)[b] | 22.82(1.17)[-b] | 59.82(6.63)[b] | 40.34(3.41)[-b] |
| Ranking | | 3 | 3 | 1 | 1 | 4 | 2 | 3 |

TABLE 11: Number of time steps of the drop periods, and average Recall0, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Opensource data

| Dataset | #Time Steps | Dummy | WP-OOB_r0 | AIO-OOB_r0 | Filter-OOB_r0 | Ensemble-OOB_r0 | OP-AIO-OOB_r0 | OP-Filter-OOB_r0 |
|--------------------|-------------|-------|-----------------|-----------------|-----------------|-----------------|------------------|------------------|
| Tomcat | 3855 | 50 | 31.23(2.14)[-b] | 56.4(1.5)[b] | 55.7(1.87)[b] | 52.88(0.77)[b] | 16.43(5.44)[-b] | 24.05(2.03)[-b] |
| JGroups | 4353 | 50 | 35.87(2.17)[-b] | 44.61(2.27)[-b] | 45.33(1.87)[-b] | 87.01(0.74)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| Spring-Integration | 3973 | 50 | 55.56(0.93)[b] | 60.18(0.97)[b] | 66.39(1.19)[b] | 75.7(0.67)[b] | 39.74(8.09)[-b] | 14.66(3.63)[-b] |
| Camel | 6009 | 50 | 35.64(2.26)[-b] | 38.98(1.58)[-b] | 40.43(1.45)[-b] | 58.66(0.75)[b] | 33.15(3.18)[-b] | 34.51(1.2)[-b] |
| Nova | 4427 | 50 | 50.2(0.15)[b] | 50.37(0.32)[b] | 50.08(0.42)[m] | 57.99(2.9)[b] | 51.22(6.41)[b] | 60.05(8.38)[b] |
| Fabric8 | 3443 | 50 | 39.92(4.78)[-b] | 44.8(1.56)[-b] | 57.77(3.01)[b] | 46.04(1.47)[-b] | 7.38(2.8)[-b] | 9.57(5.02)[-b] |
| Npm | 1202 | 50 | 16.96(1.59)[-b] | 61.28(3.82)[b] | 65.92(2.88)[b] | 43.91(1.65)[-b] | 29.65(11.83)[-b] | 3.59(1.47)[-b] |
| BroadleafCommerce | 3579 | 50 | 30.83(0.89)[-b] | 63.03(2.05)[b] | 61.69(1.63)[b] | 52.65(1.29)[b] | 35.69(2.06)[-b] | 48.59(7.28)[*] |
| Ranking | | 3 | 4 | 2 | 1 | 1 | 4 | 4 |

| Dataset | #Time Steps | Dummy | WP-ORB_r0 | AIO-ORB_r0 | Filter-ORB_r0 | Ensemble-ORB_r0 | OP-AIO-ORB_r0 | OP-Filter-ORB_r0 |
|--------------------|-------------|-------|-----------------|-----------------|------------------|-----------------|-----------------|------------------|
| Tomcat | 2265 | 50 | 36.98(2.68)[-b] | 51.19(1.39)[b] | 50.41(1.46)[s] | 63.44(1.1)[b] | 11.32(7.33)[-b] | 11.85(8.18)[-b] |
| JGroups | 2066 | 50 | 54.18(2.44)[b] | 56.29(2.23)[b] | 58.59(2.13)[b] | 86.4(0.84)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| Spring-Integration | 2319 | 50 | 80.73(1.08)[b] | 67.42(1.13)[b] | 69.26(1.31)[b] | 78.78(0.64)[b] | 75.17(2.31)[b] | 85.25(10.16)[b] |
| Camel | 1105 | 50 | 31.65(4.85)[-b] | 39.12(2.46)[-b] | 39.62(3.65)[-b] | 49.84(2.02)[*] | 62.66(14.5)[b] | 69.42(15.09)[b] |
| Nova | 4408 | 50 | 50.13(1.16)[-*] | 51.79(1.58)[b] | 49.12(0.5)[-b] | 75.65(0.92)[b] | 62.33(45.16)[m] | 71.99(14.35)[b] |
| Fabric8 | 1473 | 50 | 42.44(9.12)[-b] | 46.4(2.69)[-b] | 56.46(1.83)[b] | 44.15(2.17)[-b] | 68.9(6.57)[b] | 76.77(22.73)[b] |
| Npm | 936 | 50 | 24.56(6.1)[-b] | 25.46(5.22)[-b] | 29.76(14.42)[-b] | 51.03(1.82)[s] | 63.73(30.54)[s] | 100.0(0.0)[b] |
| BroadleafCommerce | 2703 | 50 | 36.45(3.67)[-b] | 61.89(2.39)[b] | 63.22(1.42)[b] | 60.94(1.41)[b] | 84.17(4.51)[b] | 73.09(4.61)[b] |
| Ranking | | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

TABLE 12: Number of time steps of the drop periods, and average Recall1, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Opensource data

| Dataset | #Time Steps | Dummy | WP-OOB_r1 | AIO-OOB_r1 | Filter-OOB_r1 | Ensemble-OOB_r1 | OP-AIO-OOB_r1 | OP-Filter-OOB_r1 |
|--------------------|-------------|-------|-----------------|----------------|----------------|-----------------|----------------|------------------|
| Tomcat | 3855 | 50 | 79.18(1.59)[b] | 66.49(1.43)[b] | 68.96(1.55)[b] | 55.09(0.89)[b] | 88.44(3.35)[b] | 83.28(1.76)[b] |
| JGroups | 4353 | 50 | 66.56(1.72)[b] | 62.87(1.54)[b] | 61.74(1.41)[b] | 12.58(0.49)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 3973 | 50 | 46.64(0.61)[-b] | 61.5(1.37)[b] | 58.25(1.56)[b] | 29.61(0.68)[-b] | 76.14(6.23)[b] | 92.76(1.73)[b] |
| Camel | 6009 | 50 | 82.26(0.96)[b] | 83.98(0.81)[b] | 83.43(0.82)[b] | 50.67(0.66)[b] | 79.0(2.42)[b] | 79.26(0.95)[b] |
| Nova | 4427 | 50 | 89.62(0.25)[b] | 89.38(0.36)[b] | 89.56(0.28)[b] | 56.78(2.24)[b] | 85.2(7.7)[b] | 77.97(13.94)[b] |
| Fabric8 | 3443 | 50 | 75.01(3.84)[b] | 75.04(1.69)[b] | 75.48(2.09)[b] | 64.82(1.62)[b] | 98.38(0.9)[b] | 93.65(4.43)[b] |
| Npm | 1202 | 50 | 91.4(0.63)[b] | 65.49(3.48)[b] | 65.63(3.93)[b] | 60.13(1.23)[b] | 72.95(9.92)[b] | 95.92(0.84)[b] |
| BroadleafCommerce | 3579 | 50 | 83.84(0.72)[b] | 73.57(1.44)[b] | 71.92(1.27)[b] | 59.12(2.79)[b] | 81.22(2.79)[b] | 68.43(6.13)[b] |
| Ranking | | 3 | 1 | 2 | 2 | 3 | 2 | 1 |

| Dataset | #Time Steps | Dummy | WP-ORB_r1 | AIO-ORB_r1 | Filter-ORB_r1 | Ensemble-ORB_r1 | OP-AIO-ORB_r1 | OP-Filter-ORB_r1 |
|--------------------|-------------|-------|-----------------|----------------|----------------|-----------------|------------------|------------------|
| Tomcat | 2265 | 50 | 68.09(1.91)[b] | 64.83(1.63)[b] | 65.49(1.56)[b] | 47.3(1.24)[-b] | 79.75(7.14)[b] | 78.84(6.49)[b] |
| JGroups | 2066 | 50 | 52.87(1.89)[b] | 52.96(2.08)[b] | 53.38(2.15)[b] | 14.83(0.42)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 2319 | 50 | 23.79(1.23)[-b] | 61.78(1.91)[b] | 62.65(2.08)[b] | 30.52(0.82)[-b] | 59.67(4.55)[b] | 36.54(6.56)[-b] |
| Camel | 1105 | 50 | 79.67(2.64)[b] | 80.45(1.59)[b] | 78.7(1.8)[b] | 64.43(1.78)[b] | 53.3(5.85)[b] | 48.44(13.95)[s] |
| Nova | 4408 | 50 | 90.62(0.92)[b] | 88.25(1.32)[b] | 90.76(0.42)[b] | 50.09(1.31)[s] | 43.35(42.41)[-m] | 54.78(18.64)[s] |
| Fabric8 | 1473 | 50 | 74.95(3.71)[b] | 72.6(2.28)[b] | 72.23(1.47)[b] | 69.46(1.63)[b] | 46.74(9.2)[-m] | 34.42(20.9)[-b] |
| Npm | 936 | 50 | 84.78(3.52)[b] | 89.18(2.12)[b] | 81.1(9.66)[b] | 49.68(1.19)[-s] | 47.6(34.37)[-s] | 0.0(0.0)[-b] |
| BroadleafCommerce | 2703 | 50 | 77.47(2.05)[b] | 73.22(1.72)[b] | 71.44(1.93)[b] | 52.63(1.41)[b] | 31.66(5.93)[-b] | 47.8(3.95)[-b] |
| Ranking | | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

TABLE 13: Number of time steps of the stable periods, and average Recall0, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Opensource data

| Dataset | #Time Steps | Dummy | WP-OOB_r0 | AIO-OOB_r0 | Filter-OOB_r0 | Ensemble-OOB_r0 | OP-AIO-OOB_r0 | OP-Filter-OOB_r0 |
|--------------------|-------------|-------|----------------|----------------|----------------|-----------------|------------------|------------------|
| Tomcat | 13046 | 50 | 63.41(2.5)[b] | 64.18(1.87)[b] | 71.17(1.47)[b] | 63.42(0.91)[b] | 19.51(6.07)[-b] | 25.14(2.25)[-b] |
| JGroups | 12704 | 50 | 64.27(1.94)[b] | 71.36(1.86)[b] | 73.2(1.63)[b] | 81.77(0.41)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| Spring-Integration | 4316 | 50 | 68.09(3.04)[b] | 68.21(0.87)[b] | 70.84(0.7)[b] | 68.54(0.79)[b] | 49.06(12.12)[s] | 26.51(9.06)[-b] |
| Camel | 21454 | 50 | 65.71(1.45)[b] | 57.89(1.33)[b] | 59.6(0.95)[b] | 50.03(1.23)[s] | 21.71(2.47)[-b] | 20.51(1.36)[-b] |
| Brackets | 15795 | 50 | 52.52(0.31)[b] | 65.71(1.15)[b] | 69.5(0.91)[b] | 63.43(0.83)[b] | 28.61(10.71)[-b] | 48.21(1.36)[-b] |
| Nova | 38291 | 50 | 74.79(0.3)[b] | 74.53(0.39)[b] | 75.12(0.39)[b] | 78.54(0.75)[b] | 61.52(15.06)[b] | 73.25(0.91)[b] |
| Fabric8 | 8868 | 50 | 57.35(2.36)[b] | 57.82(1.12)[b] | 62.45(1.86)[b] | 47.53(1.28)[-b] | 8.86(2.35)[-b] | 12.59(3.59)[-b] |
| Neutron | 18605 | 50 | 71.95(0.73)[b] | 73.08(0.43)[b] | 74.46(0.43)[b] | 78.59(1.09)[b] | 70.95(1.0)[b] | 76.44(3.18)[b] |
| Npm | 4182 | 50 | 56.86(4.13)[b] | 59.0(1.14)[b] | 63.26(1.66)[b] | 49.61(1.35)[-s] | 79.69(1.73)[b] | 22.34(1.15)[-b] |
| BroadleafCommerce | 10754 | 50 | 69.53(2.14)[b] | 70.16(1.62)[b] | 70.23(1.07)[b] | 63.82(0.9)[b] | 45.69(1.31)[-b] | 53.39(4.63)[b] |
| Ranking | | 3 | 2 | 2 | 1 | 2 | 3 | 3 |

| Dataset | #Time Steps | Dummy | WP-ORB_r0 | AIO-ORB_r0 | Filter-ORB_r0 | Ensemble-ORB_r0 | OP-AIO-ORB_r0 | OP-Filter-ORB_r0 |
|--------------------|-------------|-------|----------------|----------------|----------------|-----------------|-----------------|------------------|
| Tomcat | 15716 | 50 | 62.33(2.08)[b] | 64.86(1.38)[b] | 63.69(1.06)[b] | 60.08(1.4)[b] | 53.99(5.5)[b] | 55.9(4.88)[b] |
| JGroups | 14847 | 50 | 62.18(0.76)[b] | 63.83(1.0)[b] | 63.7(1.18)[b] | 82.16(0.37)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| Spring-Integration | 5531 | 50 | 70.14(1.53)[b] | 70.15(0.82)[b] | 70.53(0.78)[b] | 72.19(0.67)[b] | 78.67(2.39)[b] | 82.84(8.52)[b] |
| Camel | 24359 | 50 | 62.63(1.3)[b] | 62.2(0.74)[b] | 61.72(0.56)[b] | 57.85(1.03)[b] | 62.17(9.55)[b] | 52.39(11.16)* |
| Brackets | 15643 | 50 | 58.12(0.78)[b] | 68.04(1.09)[b] | 67.31(1.02)[b] | 61.12(1.1)[b] | 51.29(1.6)[b] | 47.6(2.14)[-b] |
| Nova | 38310 | 50 | 79.42(1.86)[b] | 79.99(1.02)[b] | 78.8(0.13)[b] | 76.58(0.96)[b] | 51.45(39.22)[s] | 69.0(2.4)[b] |
| Fabric8 | 10020 | 50 | 62.97(1.73)[b] | 57.94(1.52)[b] | 65.44(0.57)[b] | 48.2(1.32)[-b] | 76.29(6.63)[b] | 78.59(19.32)[b] |
| Neutron | 16218 | 50 | 80.44(1.88)[b] | 74.66(0.65)[b] | 73.03(1.07)[b] | 77.38(1.86)[b] | 69.93(18.31)[b] | 74.16(9.09)[b] |
| Npm | 5490 | 50 | 65.67(1.21)[b] | 61.62(1.05)[b] | 59.2(1.95)[b] | 47.85(1.65)[-b] | 75.57(13.52)[b] | 100.0(0.0)[b] |
| BroadleafCommerce | 11357 | 50 | 64.33(3.56)[b] | 69.45(0.82)[b] | 69.17(0.71)[b] | 60.09(1.43)[b] | 86.2(4.61)[b] | 77.91(5.57)[b] |
| Ranking | | 4 | 2 | 2 | 3 | 3 | 1 | 1 |

TABLE 14: Number of time steps of the stable periods, and average Recall1, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Opensource data

| Dataset | #Time Steps | Dummy | WP-OOB_r1 | AIO-OOB_r1 | Filter-OOB_r1 | Ensemble-OOB_r1 | OP-AIO-OOB_r1 | OP-Filter-OOB_r1 |
|--------------------|-------------|-------|----------------|----------------|----------------|-----------------|-----------------|------------------|
| Tomcat | 13046 | 50 | 64.16(1.84)[b] | 62.22(1.63)[b] | 58.42(1.22)[b] | 43.98(0.91)[-b] | 87.77(3.7)[b] | 84.37(1.98)[b] |
| JGroups | 12704 | 50 | 58.97(1.89)[b] | 51.37(1.91)[b] | 49.8(1.54)[-s] | 22.18(0.43)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 4316 | 50 | 65.82(2.5)[b] | 66.35(1.04)[b] | 65.44(1.26)[b] | 48.34(0.98)[-b] | 73.27(10.71)[b] | 95.21(0.76)[b] |
| Camel | 21454 | 50 | 70.71(1.21)[b] | 74.15(0.87)[b] | 73.27(0.84)[b] | 61.68(0.99)[b] | 85.22(3.59)[b] | 89.01(0.92)[b] |
| Brackets | 15795 | 50 | 89.55(0.26)[b] | 80.42(1.4)[b] | 75.74(1.38)[b] | 66.01(1.25)[b] | 92.2(2.7)[b] | 92.68(1.92)[b] |
| Nova | 38291 | 50 | 86.3(0.67)[b] | 90.73(0.7)[b] | 89.09(0.48)[b] | 60.33(1.88)[b] | 94.92(1.2)[b] | 78.04(1.82)[b] |
| Fabric8 | 8868 | 50 | 75.41(1.83)[b] | 72.29(1.62)[b] | 74.27(1.66)[b] | 63.53(1.57)[b] | 97.89(1.3)[b] | 89.65(5.59)[b] |
| Neutron | 18605 | 50 | 92.02(0.63)[b] | 92.36(0.51)[b] | 91.32(0.57)[b] | 60.77(3.51)[b] | 81.57(3.07)[b] | 70.63(9.17)[b] |
| Npm | 4182 | 50 | 66.38(2.76)[b] | 72.18(0.89)[b] | 70.47(1.6)[b] | 61.9(1.56)[b] | 16.7(2.57)[b] | 76.91(0.6)[b] |
| BroadleafCommerce | 10754 | 50 | 65.26(2.92)[b] | 71.86(1.92)[b] | 70.63(1.18)[b] | 57.95(1.13)[b] | 69.83(2.73)[b] | 72.25(2.6)[b] |
| Ranking | | 3 | 1 | 1 | 1 | 2 | 1 | 1 |

| Dataset | #Time Steps | Dummy | WP-ORB_r1 | AIO-ORB_r1 | Filter-ORB_r1 | Ensemble-ORB_r1 | OP-AIO-ORB_r1 | OP-Filter-ORB_r1 |
|--------------------|-------------|-------|----------------|----------------|----------------|-----------------|------------------|------------------|
| Tomcat | 15716 | 50 | 64.07(1.41)[b] | 62.82(0.83)[b] | 62.65(0.89)[b] | 49.63(1.18)[-s] | 48.3(4.64)[-m] | 38.64(3.47)[-b] |
| JGroups | 14847 | 50 | 59.72(1.38)[b] | 59.33(1.03)[b] | 59.68(1.33)[b] | 22.42(0.49)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Spring-Integration | 5531 | 50 | 57.05(1.85)[b] | 71.73(1.11)[b] | 71.86(0.91)[b] | 46.63(0.84)[-b] | 61.54(5.01)[b] | 43.48(9.61)[-b] |
| Camel | 24359 | 50 | 70.95(1.2)[b] | 72.6(0.93)[b] | 72.0(0.57)[b] | 58.01(1.27)[b] | 45.11(8.54)[-b] | 43.31(12.28)[-s] |
| Brackets | 15643 | 50 | 83.71(1.53)[b] | 74.05(1.74)[b] | 74.72(1.49)[b] | 72.15(1.31)[b] | 85.52(2.18)[b] | 88.95(4.3)[b] |
| Nova | 38310 | 50 | 81.5(3.3)[b] | 81.08(1.46)[b] | 81.96(0.69)[b] | 76.05(2.24)[b] | 72.83(24.19)[b] | 87.05(7.02)[b] |
| Fabric8 | 10020 | 50 | 72.89(1.25)[b] | 75.11(1.35)[b] | 72.57(1.16)[b] | 69.55(1.36)[b] | 36.16(10.58)[-b] | 27.81(22.51)[-b] |
| Neutron | 16218 | 50 | 82.15(1.91)[b] | 90.39(0.75)[b] | 92.2(0.74)[b] | 72.82(3.94)[b] | 67.64(7.47)[b] | 70.68(22.39)[b] |
| Npm | 5490 | 50 | 59.76(1.78)[b] | 75.55(1.43)[b] | 73.79(2.62)[b] | 64.48(1.41)[b] | 45.0(23.09)[s] | 0.0(0.0)[-b] |
| BroadleafCommerce | 11357 | 50 | 66.01(3.91)[b] | 69.67(1.07)[b] | 68.97(1.56)[b] | 63.72(1.74)[b] | 22.65(7.75)[-b] | 32.59(4.78)[-b] |
| Ranking | | 4 | 2 | 1 | 1 | 3 | 4 | 5 |

TABLE 15: Number of initial time steps of the initial phase, and average Recall0, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Proprietary data

| Dataset | #Time Steps | Dummy | WP-OOB_r0 | AIO-OOB combined_r0 | AIO-OOB proprietary_r0 | Filter-OOB combined_r0 | OP-AIO-OOB combined_r0 | OP-AIO-OOB proprietary_r0 | OP-Filter-OOB combined_r0 |
|---------|-------------|-------|-----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 284 | 50 | 8.88(3.47)[-b] | 73.04(4.3)[b] | 8.22(2.68)[-b] | 70.29(4.65)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C2 | 581 | 50 | 15.39(0.12)[-b] | 21.93(0.52)[-b] | 45.97(7.61)[-b] | 45.29(1.3)[-b] | 44.24(4.15)[-b] | 98.34(2.85)[b] | 48.42(5.71)[s] |
| C3 | 82 | 50 | 96.32(0.0)[b] | 81.86(0.0)[b] | 33.22(0.0)[-b] | 96.9(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C4 | 24 | 50 | 91.82(0.91)[b] | 81.24(0.45)[b] | 90.17(10.07)[b] | 100.0(0.0)[b] | 81.0(0.37)[b] | 92.34(2.85)[b] | 100.0(0.0)[b] |
| C5 | 8 | 50 | 87.37(0.0)[b] | 100.0(0.0)[b] | 59.16(9.56)[b] | 92.0(0.619)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 68.66(0.0)[b] |
| C6 | 20 | 50 | 91.76(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C7 | 29 | 50 | 100.0(0.0)[b] | 82.15(4.53)[b] | 64.66(4.4)[b] | 100.0(0.0)[b] | 79.19(0.64)[b] | 94.6(5.13)[b] | 100.0(0.0)[b] |
| C8 | 39 | 50 | 100.0(0.0)[b] | 100.0(0.0)[b] | 86.75(6.83)[b] | 100.0(0.0)[b] | 99.15(1.02)[b] | 87.98(3.62)[b] | 100.0(0.0)[b] |
| C9 | 62 | 50 | 86.42(2.18)[b] | 99.8(0.12)[b] | 97.87(2.06)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| Ranking | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

| Dataset | #Time Steps | Dummy | WP-ORB_r0 | AIO-ORB combined_r0 | AIO-ORB proprietary_r0 | Filter-ORB combined_r0 | OP-AIO-ORB combined_r0 | OP-AIO-ORB proprietary_r0 | OP-Filter-ORB combined_r0 |
|---------|-------------|-------|----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 671 | 50 | 80.46(3.18)[b] | 57.56(4.17)[b] | 48.18(1.51)[-b] | 83.32(2.87)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C2 | 582 | 50 | 72.19(0.99)[b] | 35.51(3.57)[-b] | 38.38(7.0)[-b] | 46.17(0.7)[-b] | 82.54(2.74)[b] | 87.6(8.27)[b] | 89.23(3.54)[b] |
| C3 | 396 | 50 | 58.74(1.01)[b] | 57.67(0.67)[b] | 60.56(2.2)[b] | 63.26(1.94)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C4 | 93 | 50 | 99.8(0.03)[b] | 52.43(3.65)[s] | 64.64(2.18)[b] | 99.77(0.7)[b] | 65.36(5.61)[b] | 68.91(0.74)[b] | 97.48(0.37)[b] |
| C5 | 297 | 50 | 96.65(0.51)[b] | 9.29(1.88)[-b] | 10.18(2.27)[-b] | 20.07(3.33)[-b] | 46.86(5.94)[-m] | 35.64(10.12)[-b] | 44.37(4.11)[-b] |
| C6 | 210 | 50 | 99.74(0.0)[b] | 66.1(1.25)[b] | 79.56(5.66)[b] | 69.83(0.59)[b] | 84.04(2.45)[b] | 95.09(0.43)[b] | 91.91(0.63)[b] |
| C7 | 62 | 50 | 98.68(0.2)[b] | 62.36(7.41)[b] | 61.35(3.72)[b] | 96.15(0.99)[b] | 57.67(29.59)[*] | 73.09(22.51)[b] | 100.0(0.0)[b] |
| C8 | 232 | 50 | 99.31(0.09)[b] | 49.48(13.63)[-s] | 54.53(9.78)[m] | 79.1(5.3)[b] | 65.88(7.72)[b] | 78.09(1.74)[b] | 79.09(3.19)[b] |
| C9 | 198 | 50 | 99.52(0.0)[b] | 87.47(1.51)[b] | 91.55(1.11)[b] | 88.18(0.04)[b] | 99.9(0.03)[b] | 99.84(0.6)[b] | 100.0(0.0)[b] |
| Ranking | | 4 | 1 | 3 | 2 | 1 | 1 | 1 | 1 |

TABLE 16: Number of initial time steps of the initial phase, and average Recall1, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Proprietary data

| Dataset | #Time Steps | Dummy | WP-OOB_r1 | AIO-OOB combined_r1 | AIO-OOB proprietary_r1 | Filter-OOB combined_r1 | OP-AIO-OOB combined_r1 | OP-AIO-OOB proprietary_r1 | OP-Filter-OOB combined_r1 |
|---------|-------------|-------|----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 284 | 50 | 77.86(1.91)[b] | 23.38(3.7)[-b] | 78.88(1.3)[b] | 31.5(5.25)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 581 | 50 | 86.87(0.0)[b] | 75.92(0.43)[b] | 48.36(10.42)[s] | 63.19(1.79)[b] | 55.04(3.5)[b] | 1.59(1.93)[-b] | 55.4(4.09)[b] |
| C3 | 82 | 50 | 4.2(0.0)[-b] | 21.43(0.0)[b] | 70.61(0.0)[b] | 4.2(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 24 | 50 | 4.07(0.0)[-b] | 13.33(2.96)[-b] | 10.57(5.23)[-b] | 0.0(0.0)[-b] | 14.63(0.0)[-b] | 6.82(0.0)[-b] | 0.0(0.0)[-b] |
| C5 | 8 | 50 | 8.25(0.0)[-b] | 0.77(1.19)[-b] | 40.69(7.15)[-b] | 6.49(2.79)[-b] | 2.47(2.87)[-b] | 0.95(2.16)[-b] | 42.5(0.0)[-b] |
| C6 | 20 | 50 | 2.02(0.0)[-b] | 0.65(0.0)[-b] | 0.65(0.0)[-b] | 0.5(0.28)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C7 | 29 | 50 | 0.4(0.0)[-b] | 33.96(3.36)[-b] | 61.03(9.37)[b] | 0.0(0.0)[-b] | 32.62(7.1)[-b] | 17.21(10.02)[-b] | 0.0(0.0)[-b] |
| C8 | 39 | 50 | 0.23(0.0)[-b] | 1.98(1.69)[-b] | 29.03(13.19)[-b] | 0.2(0.08)[-b] | 7.09(4.37)[-b] | 24.31(9.13)[-b] | 0.0(0.0)[-b] |
| C9 | 62 | 50 | 6.84(1.14)[-b] | 7.08(4.15)[-b] | 13.61(1.72)[-b] | 0.19(0.0)[-b] | 8.91(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| Ranking | | 1 | 3 | 3 | 2 | 3 | 3 | 3 | 3 |

| Dataset | #Time Steps | Dummy | WP-ORB_r1 | AIO-ORB combined_r1 | AIO-ORB proprietary_r1 | Filter-ORB combined_r1 | OP-AIO-ORB combined_r1 | OP-AIO-ORB proprietary_r1 | OP-Filter-ORB combined_r1 |
|---------|-------------|-------|-----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 671 | 50 | 14.15(2.64)[-b] | 36.95(3.74)[-b] | 41.14(1.7)[-b] | 23.79(4.41)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 582 | 50 | 16.46(0.81)[-b] | 60.29(3.77)[b] | 55.39(9.46)[b] | 42.99(0.69)[b] | 26.15(4.43)[-b] | 15.15(9.76)[-b] | 23.4(6.47)[-b] |
| C3 | 396 | 50 | 49.01(0.88)[-b] | 49.76(1.01)[-* | 47.38(1.9)[-b] | 46.45(1.4)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 93 | 50 | 1.79(0.15)[-b] | 49.26(5.9)[s] | 48.15(0.53)[-b] | 8.05(3.1)[-b] | 45.95(7.39)[-s] | 41.94(0.96)[-b] | 11.08(0.96)[-b] |
| C5 | 297 | 50 | 3.93(0.87)[-b] | 92.74(0.54)[b] | 92.54(2.47)[b] | 81.8(2.73)[b] | 65.37(9.0)[b] | 74.68(9.81)[b] | 73.72(3.75)[b] |
| C6 | 210 | 50 | 0.67(0.0)[-b] | 26.33(2.51)[-b] | 21.52(3.79)[-b] | 19.53(1.79)[-b] | 20.69(2.99)[-b] | 3.02(0.64)[-b] | 8.7(0.74)[-b] |
| C7 | 62 | 50 | 0.61(0.15)[-b] | 42.65(13.22)[-m] | 39.4(5.5)[-b] | 7.28(1.36)[-b] | 52.9(34.22)[*] | 31.78(27.81)[-b] | 0.0(0.0)[-b] |
| C8 | 232 | 50 | 0.35(0.06)[-b] | 64.09(9.64)[b] | 61.67(8.08)[b] | 21.64(4.81)[-b] | 46.37(5.66)[-b] | 43.14(3.28)[-b] | 29.06(3.05)[-b] |
| C9 | 198 | 50 | 0.21(0.0)[-b] | 19.27(2.23)[-b] | 20.31(1.71)[-b] | 15.84(0.0)[-b] | 5.31(0.08)[-b] | 0.33(1.25)[-b] | 0.87(0.0)[-b] |
| Ranking | | 1 | 4 | 1 | 1 | 2 | 2 | 3 | 4 |

TABLE 17: Number of time steps of the drop periods, and average Recall0, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Proprietary data

| Dataset | Time Steps | Dummy | WP-OOB_r0 | AIO-OOB combined_r0 | AIO-OOB proprietary_r0 | Filter-OOB combined_r0 | OP-AIO-OOB combined_r0 | OP-AIO-OOB proprietary_r0 | OP-Filter-OOB combined_r0 |
|---------|------------|-------|-----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 452 | 50 | 84.77(6.19)[b] | 26.11(2.22)[-b] | 42.66(13.83)[-s] | 32.12(2.29)[-b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C3 | 210 | 50 | 50.1(7.97)[-s] | 27.54(1.59)[-b] | 29.44(3.41)[-b] | 34.69(1.83)[-b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C4 | 249 | 50 | 11.97(0.65)[-b] | 30.74(2.16)[-b] | 63.51(6.07)[b] | 51.73(4.37)[*] | 47.24(2.17)[-b] | 32.49(3.42)[-b] | 39.19(0.0)[-b] |
| C6 | 466 | 50 | 5.15(0.64)[-b] | 26.33(2.09)[-b] | 34.66(11.62)[-b] | 28.3(3.19)[-b] | 45.26(1.32)[-b] | 69.05(0.41)[b] | 52.66(1.28)[b] |
| C7 | 255 | 50 | 7.14(0.65)[-b] | 23.99(9.19)[-b] | 10.95(1.03)[-b] | 21.37(1.38)[-b] | 83.72(2.23)[b] | 16.52(0.61)[-b] | 13.71(0.14)[-b] |
| C8 | 287 | 50 | 0.63(0.17)[-b] | 29.88(1.99)[-b] | 14.91(4.91)[-b] | 23.19(2.19)[-b] | 40.46(5.0)[-b] | 78.4(12.78)[b] | 8.73(5.57)[-b] |
| C9 | 541 | 50 | 11.22(1.27)[-b] | 34.4(3.19)[-b] | 53.35(5.06)[b] | 42.36(3.28)[-b] | 51.07(9.71)[s] | 92.54(7.04)[b] | 41.65(0.21)[-b] |
| Ranking | | 1 | 3 | 3 | 2 | 2 | 1 | 1 | 1 |

| Dataset | #Time Steps | Dummy | WP-ORB_r0 | AIO-ORB combined_r0 | AIO-ORB proprietary_r0 | Filter-ORB combined_r0 | OP-AIO-ORB combined_r0 | OP-AIO-ORB proprietary_r0 | OP-Filter-ORB combined_r0 |
|---------|-------------|-------|------------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 238 | 50 | 28.63(13.15)[-b] | 18.76(9.38)[-b] | 9.59(1.98)[-b] | 39.79(21.63)[-m] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C6 | 64 | 50 | 15.33(0.84)[-b] | 43.57(15.27)[-s] | 17.24(14.47)[-b] | 35.03(13.43)[-b] | 70.55(11.14)[b] | 48.53(18.28)[s] | 27.19(15.64)[-b] |
| C7 | 85 | 50 | 11.98(7.33)[-b] | 15.04(14.72)[-b] | 38.83(18.82)* | 52.99(13.53)[b] | 65.2(35.14)* | 24.48(10.46)[-b] | 100.0(0.0)[b] |
| C8 | 80 | 50 | 5.22(0.63)[-b] | 28.9(21.24)[-b] | 26.31(21.4)[-b] | 46.81(14.55)[-b] | 73.35(15.25)[b] | 94.11(6.34)[b] | 64.39(8.61)[b] |
| C9 | 23 | 50 | 6.55(0.76)[-b] | 19.86(7.08)[-b] | 7.77(3.32)[-b] | 4.68(0.68)[-b] | 70.41(10.03)[b] | 99.92(0.22)[b] | 100.0(0.0)[b] |
| Ranking | | 2 | 6 | 4 | 5 | 3 | 1 | 1 | 1 |

TABLE 18: Number of time steps of the drop periods, and average Recall1, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Proprietary data

| Dataset | #Time Steps | Dummy | WP-OOB_r1 | AIO-OOB combined_r1 | AIO-OOB proprietary_r1 | Filter-OOB combined_r1 | OP-AIO-OOB combined_r1 | OP-AIO-OOB proprietary_r1 | OP-Filter-OOB combined_r1 |
|---------|-------------|-------|-----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 452 | 50 | 14.17(5.18)[-b] | 75.04(1.66)[b] | 55.83(12.28)[b] | 71.1(2.06)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C3 | 210 | 50 | 38.31(7.4)[-b] | 84.22(2.58)[b] | 53.73(4.45)[b] | 79.31(1.67)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 249 | 50 | 92.64(0.8)[b] | 86.92(1.29)[b] | 60.85(9.01)[b] | 67.4(3.03)[b] | 71.16(2.01)[b] | 71.82(3.15)[b] | 74.73(0.0)[b] |
| C6 | 466 | 50 | 93.91(0.63)[b] | 76.17(1.87)[b] | 69.32(10.57)[b] | 76.44(2.69)[b] | 61.56(0.87)[b] | 33.49(0.32)[-b] | 52.8(0.83)[b] |
| C7 | 255 | 50 | 95.62(0.4)[b] | 78.93(6.05)[b] | 96.88(0.44)[b] | 79.55(2.64)[b] | 30.7(3.52)[-b] | 81.1(2.69)[b] | 74.01(1.18)[b] |
| C8 | 287 | 50 | 96.02(0.01)[b] | 67.92(2.38)[b] | 84.97(4.07)[b] | 71.01(1.43)[b] | 50.75(3.4)* | 29.82(14.73)[-b] | 88.28(3.54)[b] |
| C9 | 541 | 50 | 96.6(0.32)[b] | 79.08(3.72)[b] | 56.0(7.53)[b] | 69.08(3.31)[b] | 48.92(7.65)[-s] | 16.19(6.51)[-b] | 56.0(0.61)[b] |
| Ranking | | 4 | 1 | 1 | 3 | 2 | 5 | 5 | 4 |

| Dataset | #Time Steps | Dummy | WP-ORB_r1 | AIO-ORB combined_r1 | AIO-ORB proprietary_r1 | Filter-ORB combined_r1 | OP-AIO-ORB combined_r1 | OP-AIO-ORB proprietary_r1 | OP-Filter-ORB combined_r1 |
|---------|-------------|-------|----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 238 | 50 | 66.05(9.28)[b] | 81.33(10.05)[b] | 89.84(1.73)[b] | 66.29(16.21)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C6 | 64 | 50 | 82.03(0.92)[b] | 64.93(12.54)[b] | 84.23(6.62)[b] | 62.55(14.31)[b] | 54.05(6.53)[b] | 51.0(15.78)[-s] | 76.68(14.92)[b] |
| C7 | 85 | 50 | 86.32(6.53)[b] | 87.8(9.7)[b] | 74.54(11.65)[b] | 69.47(7.09)[b] | 42.91(36.14)* | 77.92(14.99)[b] | 0.0(0.0)[-b] |
| C8 | 80 | 50 | 79.23(0.73)[b] | 72.81(17.24)[b] | 76.26(17.36)[b] | 58.39(10.8)[b] | 33.68(12.15)[-b] | 12.23(7.25)[-b] | 33.15(4.47)[-b] |
| C9 | 23 | 50 | 85.4(0.1)[b] | 83.72(3.25)[b] | 88.1(4.28)[b] | 91.63(1.04)[b] | 28.74(8.68)[-b] | 0.16(0.38)[-b] | 3.23(0.0)[-b] |
| Ranking | | 4 | 2 | 2 | 1 | 3 | 5 | 5 | 5 |

TABLE 19: Number of time steps of the stable periods, and average Recall0, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Proprietary data

| Dataset | #Time Steps | Dummy | WP-OOB_r0 | AIO-OOB combined_r0 | AIO-OOB proprietary_r0 | Filter-OOB combined_r0 | OP-AIO-OOB combined_r0 | OP-AIO-OOB proprietary_r0 | OP-Filter-OOB combined_r0 |
|---------|-------------|-------|-----------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 294 | 50 | 68.67(6.6)[b] | 70.1(8.12)[b] | 37.69(12.32)[-b] | 70.35(7.36)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C2 | 20 | 50 | 43.79(0.31)[-b] | 30.97(1.95)[-b] | 41.47(3.28)[-b] | 43.23(0.92)[-b] | 21.96(2.01)[-b] | 99.98(0.05)[b] | 22.88(2.55)[-b] |
| C3 | 125 | 50 | 52.06(3.31)[m] | 38.54(0.51)[-b] | 16.82(2.45)[-b] | 53.6(0.98)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C4 | 68 | 50 | 31.88(1.5)[-b] | 48.23(1.18)[-b] | 69.56(7.66)[b] | 81.94(1.26)[b] | 54.09(1.83)[b] | 51.46(4.05)* | 68.18(0.0)[b] |
| C5 | 315 | 50 | 10.01(0.68)[-b] | 33.2(1.71)[-b] | 21.59(1.02)[-b] | 37.2(2.6)[-b] | 75.79(3.88)[b] | 41.3(5.23)[-b] | 22.39(6.7)[-b] |
| C6 | 69 | 50 | 26.94(0.59)[-b] | 59.83(0.9)[b] | 60.97(6.52)[b] | 63.14(1.68)[b] | 56.29(1.15)[b] | 75.51(0.09)[b] | 57.25(1.23)[b] |
| C7 | 262 | 50 | 29.44(1.12)[-b] | 47.27(16.27)[-s] | 30.36(2.47)[-b] | 61.28(3.15)[b] | 71.3(2.59)[b] | 45.49(1.51)[-b] | 31.83(1.42)[-b] |
| C8 | 472 | 50 | 13.41(0.09)[-b] | 52.31(1.41)[b] | 33.75(5.39)[-b] | 56.43(1.92)[b] | 79.96(2.43)[b] | 76.31(10.7)[b] | 28.6(3.5)[-b] |
| C9 | 91 | 50 | 31.84(2.71)[-b] | 67.15(1.64)[b] | 70.01(7.7)[b] | 78.09(1.55)[b] | 84.37(3.2)[b] | 97.16(2.58)[b] | 80.82(0.0)[b] |
| Ranking | | 2 | 4 | 2 | 3 | 1 | 1 | 1 | 1 |

| Dataset | #Time Steps | Dummy | WP-ORB_r0 | AIO-ORB combined_r0 | AIO-ORB proprietary_r0 | Filter-ORB combined_r0 | OP-AIO-ORB combined_r0 | OP-AIO-ORB proprietary_r0 | OP-Filter-ORB combined_r0 |
|---------|-------------|-------|------------------|---------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| C1 | 121 | 50 | 48.22(13.97)[-m] | 30.34(12.18)[-b] | 17.33(2.88)[-b] | 51.83(15.36)[-*] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C2 | 19 | 50 | 49.13(2.37)[-m] | 47.11(1.46)[-b] | 31.07(7.39)[-b] | 45.37(0.7)[-b] | 96.24(2.41)[b] | 88.57(12.86)[b] | 98.83(0.49)[b] |
| C3 | 21 | 50 | 57.23(2.36)[b] | 51.87(1.99)[b] | 61.63(3.36)[b] | 62.94(6.05)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] | 100.0(0.0)[b] |
| C4 | 248 | 50 | 57.35(5.96)[b] | 20.91(9.31)[-b] | 26.0(4.32)[-b] | 66.35(3.78)[b] | 42.92(9.0)[-b] | 24.12(1.07)[-b] | 86.82(2.22)[b] |
| C5 | 26 | 50 | 87.66(3.65)[b] | 43.52(1.95)[-b] | 43.15(2.6)[-b] | 43.25(6.74)[-b] | 56.71(16.81)[s] | 35.02(10.21)[-b] | 48.57(17.23)[-m] |
| C6 | 281 | 50 | 47.86(0.68)[-b] | 45.21(8.02)[-b] | 40.18(6.55)[-b] | 51.36(10.57)[m] | 69.18(4.7)[b] | 68.15(12.22)[b] | 47.56(10.78)[-b] |
| C7 | 399 | 50 | 47.84(5.89)[-m] | 34.28(14.53)[-b] | 45.1(11.62)[-*] | 62.56(6.54)[b] | 53.82(40.33)* | 32.37(20.88)[-b] | 100.0(0.0)[b] |
| C8 | 486 | 50 | 24.99(0.64)[-b] | 33.92(17.42)[-b] | 30.49(21.0)[-b] | 46.02(8.36)[-m] | 78.66(10.29)[b] | 93.79(4.3)[b] | 82.56(4.77)[b] |
| C9 | 473 | 50 | 25.0(0.33)[-b] | 20.13(2.46)[-b] | 26.41(7.98)[-b] | 17.05(1.48)[-b] | 78.91(7.13)[b] | 99.92(0.2)[b] | 99.94(0.0)[b] |
| Ranking | | 2 | 2 | 3 | 3 | 2 | 1 | 1 | 1 |

TABLE 20: Number of time steps of the stable periods, and average Recall1, A12 effect sizes and Scott-Knott.BA12 to compare learning approaches on this initial phase for OOB and ORB for Proprietary data

| Dataset | #Time Steps | Dummy | WP-OOB_r1 | AIO-OOB combined_r1 | AIO-OOB proprietary_r1 | Filter-OOB combined_r1 | OP-AIO-OOB combined_r1 | OP-AIO-OOB proprietary_r1 | OP-Filter-OOB combined_r1 |
|---------|-------------|-------|-----------------|------------------------|---------------------------|---------------------------|---------------------------|------------------------------|------------------------------|
| C1 | 294 | 50 | 40.94(5.53)[-b] | 31.25(5.18)[-b] | 66.19(9.31)[b] | 32.55(4.31)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 20 | 50 | 95.41(0.0)[b] | 90.63(0.34)[b] | 75.08(8.04)[b] | 81.76(1.55)[b] | 60.95(3.34)[b] | 0.18(0.24)[-b] | 61.4(3.04)[b] |
| C3 | 125 | 50 | 60.68(2.77)[b] | 76.54(0.57)[b] | 85.09(4.0)[b] | 63.03(0.57)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 68 | 50 | 70.33(0.48)[b] | 61.33(2.18)[b] | 51.46(10.89)[-s] | 26.06(3.45)[-b] | 43.13(0.67)[-b] | 49.03(5.56)[-s] | 33.88(0.0)[-b] |
| C5 | 315 | 50 | 87.4(1.47)[b] | 69.41(3.14)[b] | 91.31(1.18)[b] | 68.6(4.17)[b] | 27.42(5.61)[-b] | 69.5(4.43)[b] | 83.45(6.8)[b] |
| C6 | 69 | 50 | 72.3(0.29)[b] | 44.65(0.54)[-b] | 49.66(6.76)* | 43.24(2.6)[-b] | 39.98(0.67)[-b] | 21.17(0.19)[-b] | 34.78(0.67)[-b] |
| C7 | 262 | 50 | 80.94(0.76)[b] | 66.75(10.32)[b] | 93.45(1.45)[b] | 51.74(5.52)[s] | 47.51(1.64)[-b] | 63.96(3.6)[b] | 64.92(0.49)[b] |
| C8 | 472 | 50 | 82.1(0.01)[b] | 60.49(2.11)[b] | 72.47(5.79)[b] | 38.65(2.13)[-b] | 27.55(2.61)[-b] | 34.23(9.42)[-b] | 71.64(2.57)[b] |
| C9 | 91 | 50 | 80.9(0.96)[b] | 52.67(3.64)[b] | 56.11(12.84)[m] | 32.22(2.12)[-b] | 20.54(2.67)[-b] | 4.74(2.18)[-b] | 19.72(0.54)[-b] |
| Ranking | | 3 | 1 | 2 | 1 | 3 | 5 | 5 | 4 |

| Dataset | #Time Steps | Dummy | WP-ORB_r1 | AIO-ORB combined_r1 | AIO-ORB proprietary_r1 | Filter-ORB combined_r1 | OP-AIO-ORB combined_r1 | OP-AIO-ORB proprietary_r1 | OP-Filter-ORB combined_r1 |
|---------|-------------|-------|-----------------|------------------------|---------------------------|---------------------------|---------------------------|------------------------------|------------------------------|
| C1 | 121 | 50 | 53.31(10.56)[m] | 76.43(12.42)[b] | 86.06(2.21)[b] | 60.7(13.86)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C2 | 19 | 50 | 49.07(1.59)[-b] | 79.24(2.8)[b] | 74.37(8.71)[b] | 74.63(1.04)[b] | 13.63(2.74)[-b] | 16.78(11.42)[-b] | 14.6(2.42)[-b] |
| C3 | 21 | 50 | 65.32(2.08)[b] | 70.67(2.53)[b] | 63.99(4.56)[b] | 61.4(5.67)[b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] | 0.0(0.0)[-b] |
| C4 | 248 | 50 | 54.03(8.28)[m] | 89.28(6.21)[b] | 87.96(2.23)[b] | 51.19(2.72)[s] | 70.53(12.32)[b] | 82.72(0.47)[b] | 51.85(7.74)[s] |
| C5 | 26 | 50 | 21.41(7.72)[-b] | 72.84(2.94)[b] | 73.21(3.27)[b] | 70.62(5.21)[b] | 61.45(8.44)[b] | 68.01(10.35)[b] | 62.72(9.68)[b] |
| C6 | 281 | 50 | 64.44(0.61)[b] | 66.63(8.06)[b] | 73.93(4.33)[b] | 54.71(11.81)[-s] | 57.88(3.96)[b] | 40.69(11.55)[-b] | 67.63(11.18)[b] |
| C7 | 399 | 50 | 59.21(4.05)[b] | 80.05(10.58)[b] | 75.5(7.1)[b] | 60.97(3.76)[b] | 51.88(38.94)* | 62.18(22.32)[b] | 0.0(0.0)[-b] |
| C8 | 486 | 50 | 58.0(0.87)[b] | 73.5(12.77)[b] | 76.96(15.2)[b] | 52.81(5.95)[s] | 33.26(9.93)[-b] | 18.36(6.82)[-b] | 28.7(3.47)[-b] |
| C9 | 473 | 50 | 69.17(0.12)[b] | 81.02(1.72)[b] | 77.59(6.14)[b] | 83.16(1.7)[b] | 19.65(5.55)[-b] | 0.29(0.68)[-b] | 3.45(0.0)[-b] |
| Ranking | | 4 | 3 | 1 | 1 | 2 | 5 | 5 | 5 |

2 PLOTS OF PREDICTIVE PERFORMANCE THROUGH TIME

This section shows plots of the G-Means through time for all approaches investigated in the study.

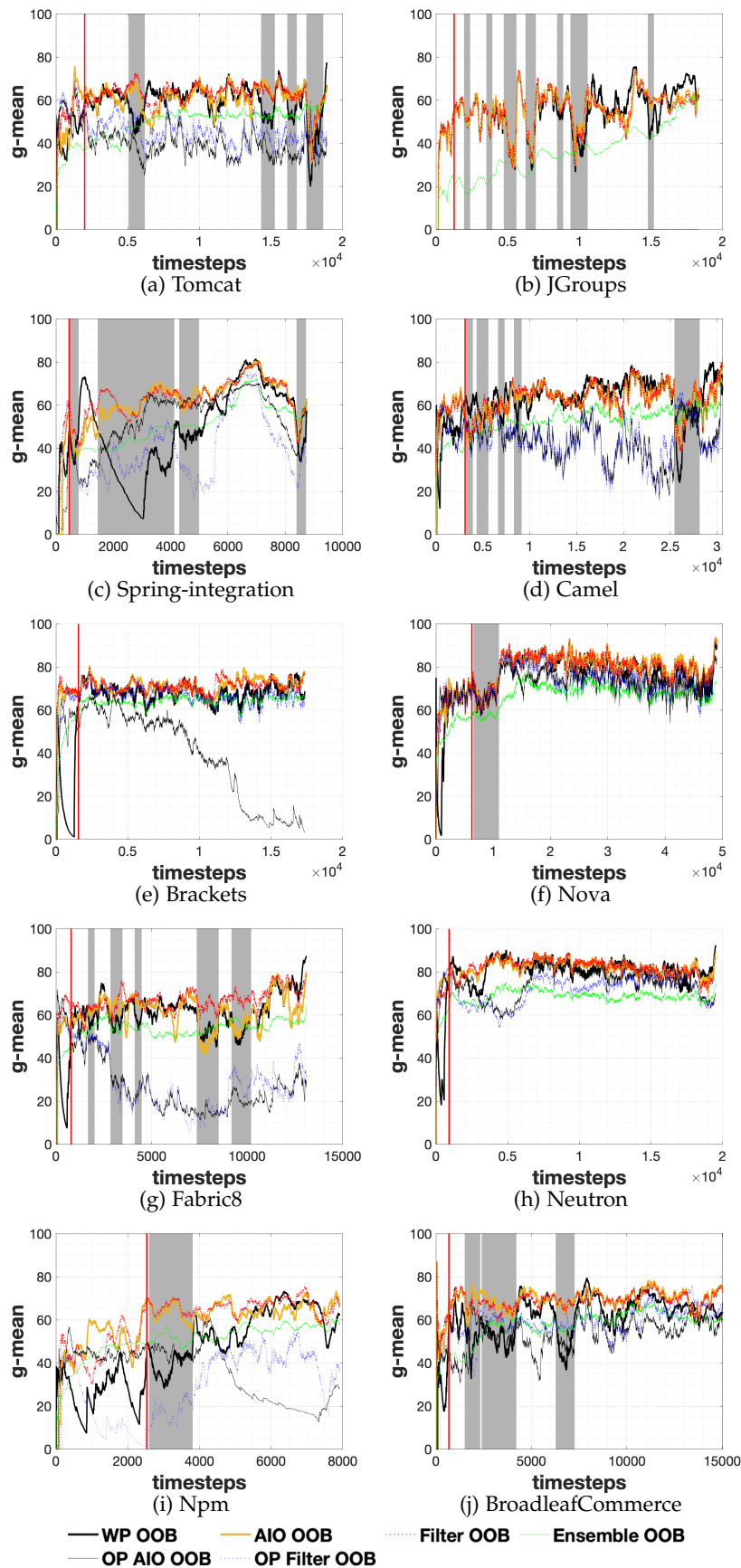


Fig. 1: G-Mean for all datasets through time using OOB. The vertical red bar indicates the last time step of the initial phase of the project. The periods highlighted in grey background are the sudden drop periods.

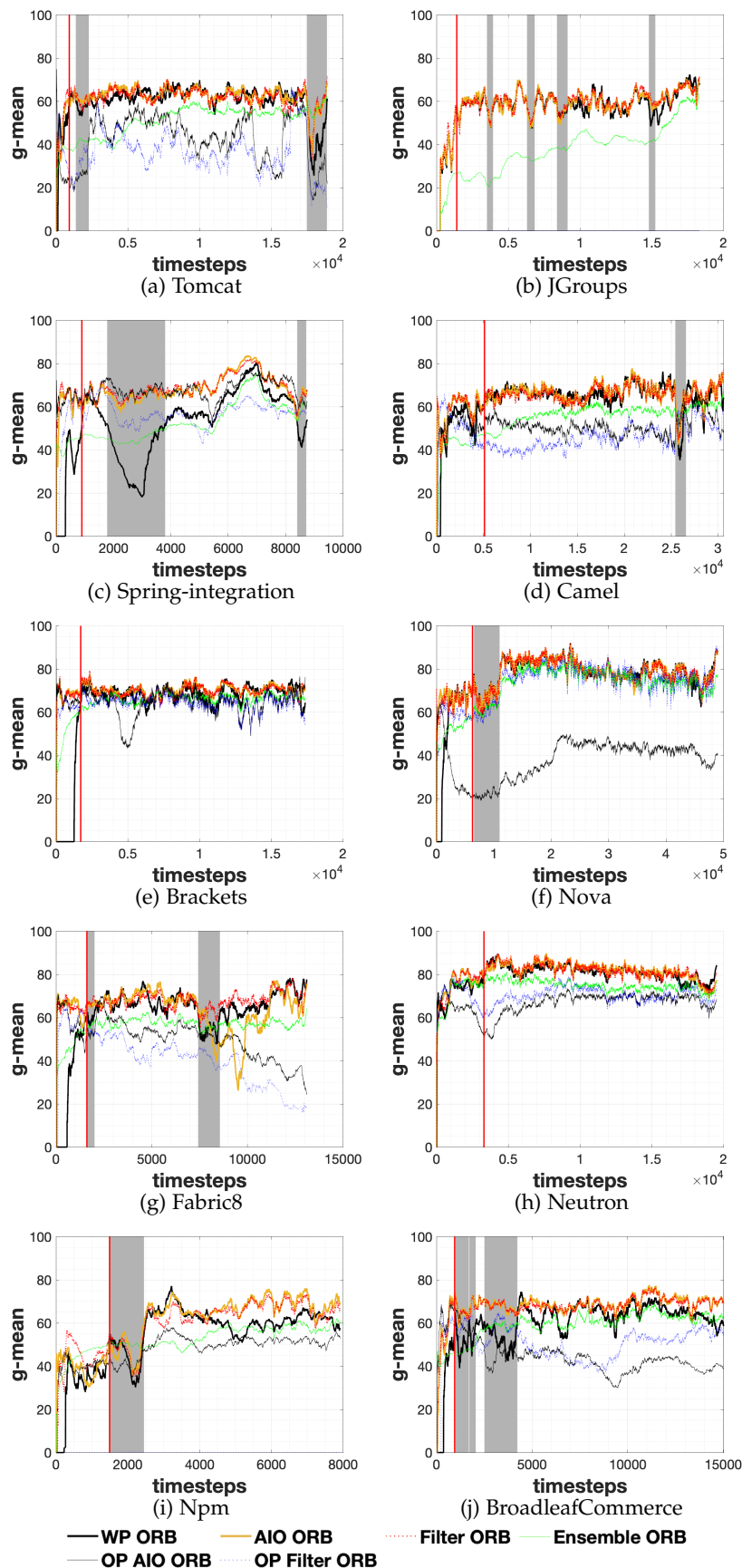


Fig. 2: G-Mean for all datasets through time using ORB. The vertical red bar indicates the last time step of the initial phase of the project. The periods highlighted in grey background are the sudden drop periods.

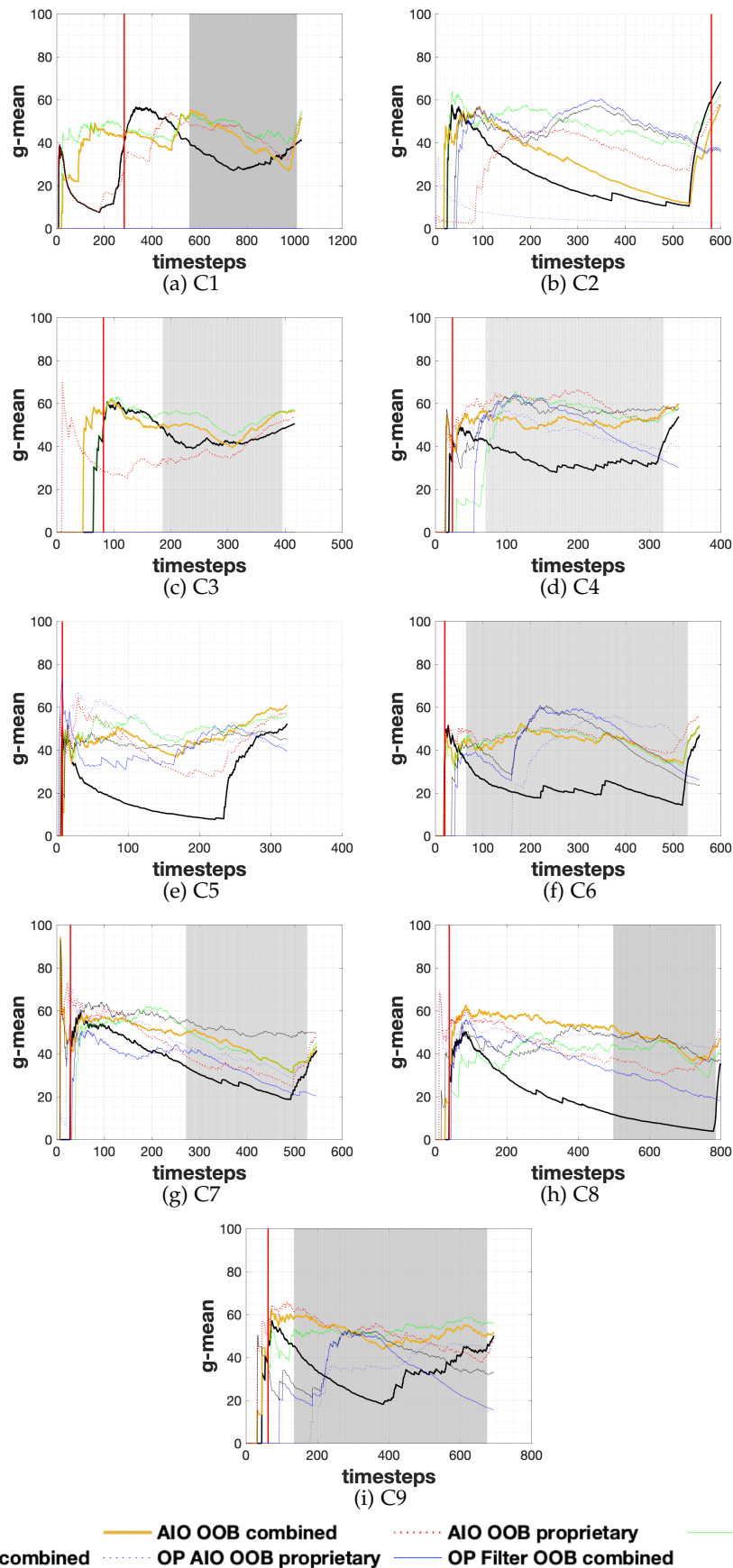


Fig. 3: G-Mean for all datasets through time using OOB. The vertical red bar indicates the last time step of the initial phase of the project. The periods highlighted in grey background are the sudden drop periods.

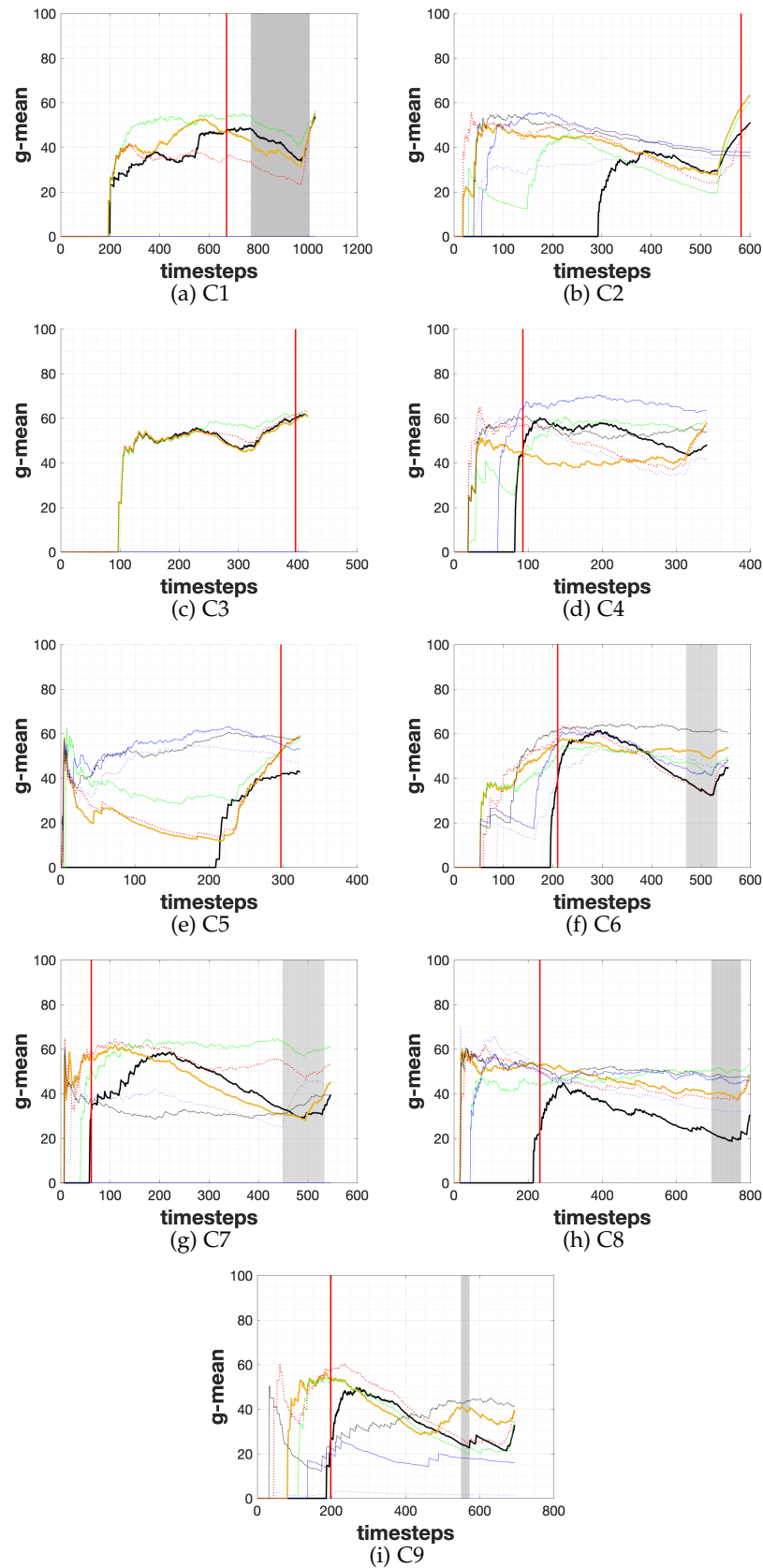


Fig. 4: G-Mean for all datasets through time using ORB. The vertical red bar indicates the last time step of the initial phase of the project. The periods highlighted in grey background are the sudden drop periods.

3 HYPERPARAMETER SENSITIVITY ANALYSIS

This section shows the tables and plots of results corresponding to the analysis of sensitivity to hyperparameters when using Filtering-ORB.

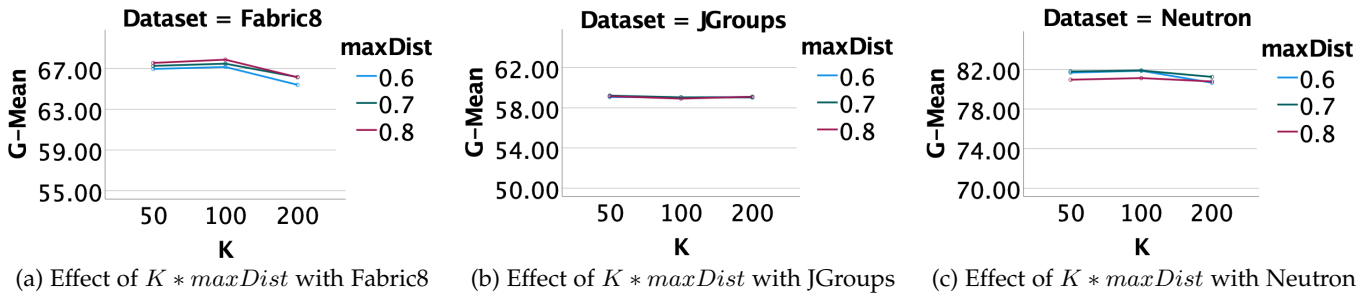


Fig. 5: Plots of Marginal Means for Filtering-ORB's Factors K , $maxDist$ and $dataset$ (Open source data)

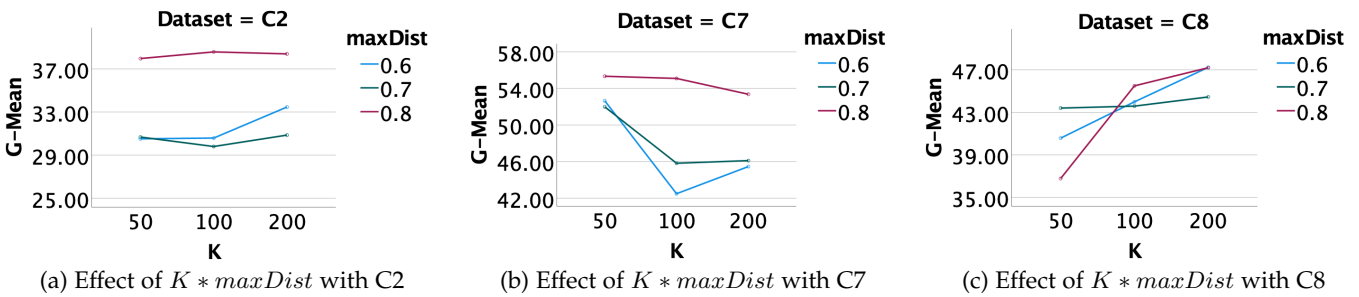


Fig. 6: Plots of Marginal Means for Filtering-ORB's Factors K , $maxDist$ and $dataset$ (Proprietary data)

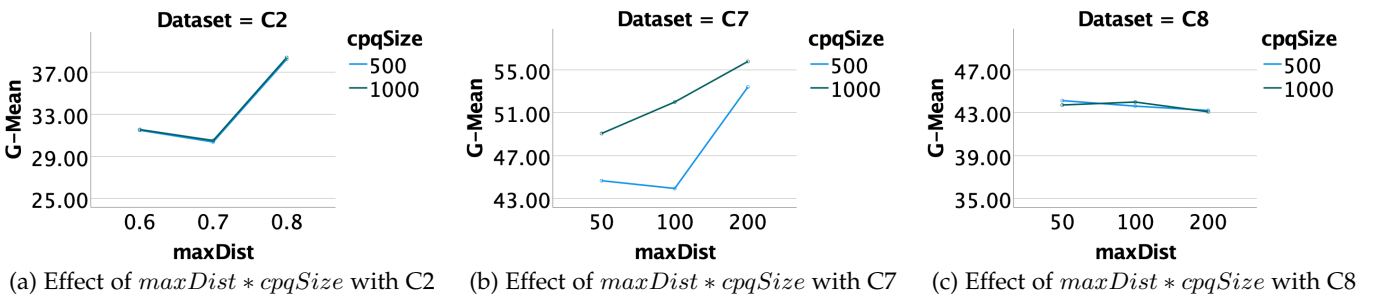


Fig. 7: Plots of Marginal Means for Filtering-ORB's Factors $cpqSize$, $maxDist$ and $dataset$ (Proprietary data)

TABLE 21: ANOVA and Effect Size Results for Filtering-ORB (Opensource Data)

| Within Subject Test for Filtering-ORB (Opensource Data) | | | | | |
|---|------------|-------|------------|------------|-------|
| K | 554.041 | 1.847 | 300.027 | 1139.471 | 0.929 |
| K * Dataset | 417.928 | 4 | 104.482 | 429.767 | 0.908 |
| maxDist * Dataset | 212.243 | 4 | 53.061 | 245.435 | 0.849 |
| windowSize*K | 175.995 | 3.447 | 51.062 | 179.843 | 0.674 |
| windowSize | 89.059 | 2 | 44.529 | 151.658 | 0.635 |
| maxDist | 50.859 | 2 | 25.43 | 117.626 | 0.575 |
| windowSize * K * Dataset | 100.475 | 8 | 12.559 | 51.336 | 0.541 |
| windowSize * Dataset | 56.363 | 4 | 14.091 | 47.99 | 0.525 |
| K * maxDist * Dataset | 39.91 | 8 | 4.989 | 22.385 | 0.34 |
| K * maxDist | 23.317 | 4 | 5.829 | 26.156 | 0.231 |
| windowSize * maxDist | 23.299 | 4 | 5.825 | 23.286 | 0.211 |
| windowSize * maxDist * Dataset | 17.442 | 8 | 2.18 | 8.716 | 0.167 |
| windowSize * K * maxDist | 31.295 | 8 | 3.912 | 15.318 | 0.15 |
| cpqSize * Dataset | 2.593 | 2 | 1.296 | 5.246 | 0.108 |
| K * cpqSize * Dataset | 3.871 | 4 | 0.968 | 3.102 | 0.067 |
| windowSize * K * maxDist * Dataset | 12.364 | 16 | 0.773 | 3.026 | 0.065 |
| maxDist * cpqSize | 1.977 | 2 | 0.989 | 3.697 | 0.041 |
| windowSize * K * cpqSize | 2.5 | 3.543 | 0.706 | 2.872 | 0.032 |
| windowSize * maxDist * cpqSize * Dataset | 2.82 | 8 | 0.352 | 1.428 | 0.032 |
| K * maxDist * cpqSize * Dataset | 2.598 | 8 | 0.325 | 1.328 | 0.03 |
| windowSize * K * maxDist * cpqSize * Dataset | 3.545 | 16 | 0.222 | 0.886 | 0.02 |
| windowSize * cpqSize * Dataset | 0.806 | 4 | 0.202 | 0.763 | 0.017 |
| windowSize * K * maxDist * cpqSize | 2.651 | 6.754 | 0.392 | 1.325 | 0.015 |
| K * cpqSize | 0.748 | 2 | 0.374 | 1.199 | 0.014 |
| maxDist * cpqSize * Dataset | 0.518 | 4 | 0.129 | 0.484 | 0.011 |
| windowSize * K * cpqSize * Dataset | 0.524 | 8 | 0.066 | 0.301 | 0.007 |
| windowSize * maxDist * cpqSize | 0.574 | 4 | 0.143 | 0.581 | 0.007 |
| windowSize * cpqSize | 0.194 | 2 | 0.097 | 0.367 | 0.004 |
| cpqSize | 0.064 | 1 | 0.064 | 0.257 | 0.003 |
| K * maxDist * cpqSize | 0.285 | 4 | 0.071 | 0.291 | 0.003 |
| Between Subject Test with ORB | | | | | |
| Dataset | 412679.309 | 2 | 206339.654 | 918130.525 | 1 |

TABLE 22: ANOVA Analysis of Hyperparameters for Filtering-ORB (Proprietary Data)

| Within Subject Test with ORB | | | | | |
|-------------------------------|------------|-------|------------|----------|------------|
| Factor /Int. | SS | DF | MS | F | η_p^2 |
| maxDist | 23566.525 | 1.747 | 13493.132 | 1355.966 | 0.938 |
| maxDist*dataset | 16407.829 | 4 | 4101.957 | 472.035 | 0.913 |
| K*dataset | 20278.404 | 4 | 5069.601 | 428.2 | 0.905 |
| cpqSize*dataset | 6741.644 | 2 | 3370.822 | 390.947 | 0.897 |
| cpqSize | 3519.707 | 1 | 3519.707 | 408.214 | 0.819 |
| K*maxDist | 6321.242 | 4 | 1580.31 | 170.739 | 0.655 |
| K*maxDist*dataset | 4267.894 | 8 | 533.487 | 57.639 | 0.562 |
| maxDist*cpqSize*dataset | 1409.002 | 4 | 352.25 | 40.993 | 0.477 |
| maxDist*cpqSize | 946.645 | 2 | 473.322 | 55.083 | 0.38 |
| K | 1285.156 | 2 | 642.578 | 54.275 | 0.376 |
| K*cpqSize*dataset | 203.465 | 4 | 50.866 | 4.622 | 0.093 |
| K*maxDist*cpqSize | 250.999 | 4 | 62.75 | 7.113 | 0.073 |
| Between Subject Test with ORB | | | | | |
| dataset | 229263.263 | 2 | 114631.632 | 9383.673 | 0.995 |

4 ALGORITHMS

This section lists the pseudocode of the online CP Ensemble approach and the pseudocode of the systematic procedure for identifying periods of sudden drop in predictive performance.

Algorithm 1 Sudden Drop Detection Approach

Input: d_thr = Detecta threshold, $gmeans$ = G-mean for each time step t

```

1:  $drop\_periods = \{\}$ 
2:  $peaks = detecta\_onset(d\_thr, gmeans)$  // time steps with peaks
3: for time step  $t = 1$  to  $gmeans.length()$  do
4:   if  $t \in peaks$  then
5:      $latest\_peak = t$ 
6:   end if
7:    $window\_avg = average(gmeans[latest\_peak : t])$ 
8:    $drop\_thr = window\_avg - (0.20 * window\_avg)$ 
9:   if  $gmeans[t] < drop\_thr$  then
10:     $peak\_left = latest\_peak$ 
11:     $peak\_right = \text{Get next peak after timestamp } t$ 
    from  $peaks$ 
12:     $drop\_periods.add([peak.left, peak.right])$ 
13:   end if
14: end for
15: Return  $drop\_periods$ 

```

Algorithm 2 Ensemble approach

Inputs: S = stream of incoming changes from n projects, b = index identifying the test project, w = waiting period

```

1: initialise ensemble model  $M$  consisting of  $n$  models  $\{m_1, m_2, \dots, m_n\}$ 
2: for each incoming change  $x_p^t \in S$  do //  $x_p^t$  is a change arriving from project  $p$  at timestamp  $t$ 


---


3:   if  $p = b$  then
4:      $\hat{Y}\text{-List} \leftarrow$  get probability  $\hat{y}_i$  of change  $x_p^t$  being defect-inducing from each model in  $M$ 
5:      $MeanY =$  mean of values in  $\hat{Y}\text{-List}$ 
6:     if  $MeanY < 0.5$  then
7:        $\hat{y} = 0$ 
8:     else
9:        $\hat{y} = 1$ 
10:    end if
11:   end if
12:   store  $x_p^t$  in a queue  $WFL\text{-}Q_p$  for project  $p$  //  $WFL\text{-}Q_p$  is the queue of incoming changes of project  $p$  waiting to be used for trained


---


13:   for each model  $m_p$  in  $M$  do
14:     for each change  $q^i$  in  $WFL\text{-}Q_p$  do
15:       if a defect was linked to  $q^i$  at a timestamp  $\leq t$  then
16:         create defect-inducing  $training\_example$  for  $q^i$ 
17:          $train(m_p, training\_example)$ 
18:         remove  $q^i$  from  $WFL\text{-}Q_p$ 
19:       else
20:         create a clean  $training\_example$  for  $q^i$ 
21:          $train(m_p, training\_example)$ 
22:         remove  $q^i$  from  $WFL\text{-}Q$ 
23:         store  $training\_example$  in  $CL_p\text{-}H$  //  $CL_p\text{-}H$  is a hash of clean training examples for project  $p$ 
24:       end if
25:     end for
26:   end for


---


27:   if a defect was linked to a  $training\_example$  in  $CL_p\text{-}H$  before time  $t$  then
28:     Swap the label of  $training\_example$  to defect-inducing
29:      $train(m_p, training\_example)$ 
30:     remove  $h$  from  $CL_p\text{-}H$ 
31:   end if
32: end for


---



```