|  |
| --- |
| Sequence analysis  **MULTiPly: a novel multi-layer predictor for discovering general and specific types of promoters**  Meng Zhang1,†, Fuyi Li2,3,†, Tatiana T. Marquez-Lago4, André Leier4, Cunshuo Fan5, Cheekeong Kwoh6, Kuo-Chen Chou7, Jiangning Song2,3,\* and Cangzhi Jia1,6,\*  1College of Science, Dalian Maritime University, Dalian 116026, P.R. China, 2Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, VIC 3800, Australia, 3Monash Centre for Data Science, Faculty of Information Technology, Monash University, Melbourne, VIC 3800, Australia, 4Department of Genetics and Department of Cell, Developmental and Integrative Biology, School of Medicine, University of Alabama at Birmingham, AL, USA, 5College of Information Engineering, Northwest A&F University, Yangling 712100, China, 6School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore, 7Gordon Life Science Institute, Boston, MA 02478, USA.  †These two authors contributed equally to this work. \*To whom correspondence should be addressed. |

**Supplementary Results**

**Content**

**Figures:**

[**Figure S1.** Snapshots of the webserver interface of MULTiPly: (**A**) The online web interface of MULTiPly; (**B**) The input interface of MULTiPly; (**C**) The output interface of MULTiPly, which displayed the prediction results for the query sequences generated by the web server, and (**D**) The job list interface of MULTiPly, which listed the job IDs, submission time, job status of the previously submitted jobs. Users can revisit the result page of a completed job by clicking its corresponding ‘Click’ button. 2](#_Toc531687675)

**Tables:**

**[Table S1.](#_Toc532124443)** [Selection of the optimal single features in term of the F-score for the first task I. 3](#_Toc532124443)

[**Table S2.** Selection of optimal single features according to F-score values for the second task. 4](#_Toc532124444)

[**Table S3.** Performance comparison of different classifiers trained using different features on the jackknife test based on F-score selection for the first task I. 9](#_Toc532124445)

[**Table S4.** Performance comparison of different sub-classifiers trained using different combinations of features on the jackknife test in terms of F-score selection for the second task II. 11](#_Toc532124446)

[**Table S5.** Best performance results obtained for each feature combination for the first task I. 17](#_Toc532124447)

[**Table S6.** Best performance results obtained for each feature combination for the second task II. 17](#_Toc532124448)

[**Table S7.** Performance comparison between MULTiPly and iPromoter-2L for the second task II on 5-fold cross-validation test. 17](#_Toc532124449)

[**Table S8.** Performance comparison between MULTiPly and a direct multi-class SVM classifier. 18](#_Toc532124450)

[**Table S9.** Performance comparison results of the multi-task predictor based on different sub-classifiers constructed using different numbers of trees. 18](#_Toc532124451)

[**Table S10.** Performancecomparison of different classifiers for identifying promoters and their types using the jackknife tests. 19](#_Toc532124452)

|  |  |
| --- | --- |
| **A** | **B** |
| **C** | **D** |

**Figure S1.** Snapshots of the webserver interface of MULTiPly: (**A**) The online web interface of MULTiPly; (**B**) The input interface of MULTiPly; (**C**) The output interface of MULTiPly, which displayed the prediction results for the query sequences generated by the web server, and (**D**) The job list interface of MULTiPly, which listed the job IDs, submission time, job status of the previously submitted jobs. Users can revisit the result page of a completed job by clicking its corresponding ‘Click’ button.

**Table S1.** Selection of the optimal single features in term of the F-score for the first task I.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1st task | Sn (%) | Sp (%) | Acc (%) | MCC |
| **Dim(KNN)** |  |  |  |  |
| 20 | 85.49 | 86.75 | 86.12 | 0.7224 |
| 19 | 85.42 | 86.75 | 86.08 | 0.7217 |
| 17 | 85.42 | 86.71 | 86.07 | 0.7214 |
| **15** | **85.56** | **86.68** | **86.12** | **0.7224** |
| 13 | 85.42 | 86.71 | 86.07 | 0.7214 |
| 11 | 85.35 | 86.71 | 86.03 | 0.7207 |
| 9 | 85.56 | 86.57 | 86.07 | 0.7214 |
| 7 | 85.59 | 86.33 | 85.96 | 0.7193 |
| 5 | 85.91 | 85.98 | 85.94 | 0.7189 |
| 3 | 86.12 | 85.7 | 85.91 | 0.7182 |
| 1 | 86.29 | 83.53 | 84.91 | 0.6985 |
| **Dim(BPB)** |  |  |  |  |
| 162 | 81.78 | 80.63 | 81.21 | 0.6242 |
| 160 | 81.75 | 80.49 | 81.12 | 0.6224 |
| 150 | 81.82 | 80.63 | 81.22 | 0.6245 |
| 140 | 81.61 | 81.01 | 81.31 | 0.6262 |
| 130 | 81.99 | 80.98 | 81.49 | 0.6298 |
| **120** | **82.03** | **81.40** | **81.71** | **0.6343** |
| 110 | 81.82 | 80.66 | 81.24 | 0.6249 |
| 100 | 80.77 | 79.72 | 80.24 | 0.6049 |
| 90 | 81.19 | 79.72 | 80.45 | 0.6092 |
| 80 | 80.91 | 80.24 | 80.58 | 0.6116 |
| 70 | 81.71 | 80.1 | 80.91 | 0.6183 |
| 60 | 81.4 | 79.62 | 80.51 | 0.6102 |
| 50 | 81.54 | 79.2 | 80.37 | 0.6075 |
| 40 | 79.76 | 79.62 | 79.69 | 0.5937 |
| 30 | 79.62 | 79.79 | 79.7 | 0.5941 |
| 20 | 79.41 | 78.46 | 78.93 | 0.5787 |
| 10 | 76.71 | 72.2 | 74.46 | 0.4897 |
| **Dim(DNC)** |  |  |  |  |
| 16 | 74.76 | 80.7 | 77.73 | 0.5555 |
| 14 | 74.72 | 80.59 | 77.66 | 0.5541 |
| **12** | **74.86** | **80.84** | **77.85** | **0.5580** |
| 10 | 74.62 | 80.63 | 77.62 | 0.5534 |
| 8 | 74.34 | 80.45 | 77.4 | 0.5489 |
| 6 | 73.88 | 80.35 | 77.12 | 0.5434 |
| 4 | 72.24 | 79.79 | 76.01 | 0.5218 |
| 2 | 73.46 | 73.04 | 73.25 | 0.465 |
| **Dim(MNC)** |  |  |  |  |
| **4** | **73.25** | **80.59** | **76.92** | **0.5399** |
| 3 | 72.66 | 81.08 | 76.87 | 0.5393 |
| 2 | 64.13 | 83.32 | 73.72 | 0.4835 |
| 1 | 64.65 | 80.38 | 72.52 | 0.456 |
| **Dim(DAC)** |  |  |  |  |
| **12** | **74.48** | **76.15** | **75.31** | **0.5064** |
| 10 | 74.97 | 74.34 | 74.65 | 0.493 |
| 8 | 73.88 | 75.07 | 74.48 | 0.4895 |
| 6 | 73.39 | 74.79 | 74.09 | 0.4819 |
| 4 | 72.38 | 72.45 | 72.41 | 0.4483 |
| 2 | 71.92 | 70.94 | 71.43 | 0.4287 |

**Table S2.** Selection of optimal single features according to F-score values for the second task.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1st sub-classifier | Sn (%) | Sp (%) | Acc (%) | MCC |
| **Dim(BPB)** |  |  |  |  |
| **162** | **88.55** | **76.76** | **83.74** | **0.6609** |
| 160 | 88.55 | 76.5 | 83.64 | 0.6586 |
| 150 | 88.37 | 76.67 | 83.6 | 0.658 |
| 140 | 88.13 | 77.19 | 83.67 | 0.6597 |
| 130 | 87.49 | 76.5 | 83.01 | 0.6459 |
| 120 | 87.6 | 76.07 | 82.9 | 0.6435 |
| 110 | 87.37 | 76.07 | 82.76 | 0.6407 |
| 100 | 87.01 | 75.99 | 82.52 | 0.6357 |
| 90 | 86.95 | 75.81 | 82.41 | 0.6335 |
| 80 | 86.48 | 74.7 | 81.68 | 0.618 |
| 70 | 87.07 | 74.96 | 82.13 | 0.6273 |
| 60 | 87.07 | 73.76 | 81.64 | 0.6167 |
| 50 | 86.84 | 73.41 | 81.36 | 0.6108 |
| 40 | 87.37 | 73.5 | 81.71 | 0.618 |
| 30 | 88.37 | 74.7 | 82.8 | 0.6407 |
| 20 | 87.96 | 72.04 | 81.47 | 0.6124 |
| 10 | 84.59 | 68.35 | 77.97 | 0.5389 |
| **Dim(KNN)** |  |  |  |  |
| 20 | 90.20 | 75.47 | 84.2 | 0.67 |
| 19 | 90.20 | 75.64 | 84.27 | 0.6715 |
| 17 | 90.14 | 75.39 | 84.13 | 0.6686 |
| **15** | **90.26** | **75.64** | **84.3** | **0.6723** |
| 13 | 89.79 | 76.16 | 84.23 | 0.6708 |
| 11 | 89.55 | 75.99 | 84.02 | 0.6664 |
| 9 | 88.9 | 76.07 | 83.67 | 0.6592 |
| 7 | 89.26 | 75.64 | 83.71 | 0.6598 |
| 5 | 89.14 | 75.73 | 83.67 | 0.6591 |
| 3 | 88.61 | 74.01 | 82.66 | 0.6376 |
| 1 | 84.18 | 77.27 | 81.36 | 0.6143 |
| **Dim(DNC)** |  |  |  |  |
| 16 | 88.13 | 29.93 | 64.41 | 0.2253 |
| 14 | 88.43 | 29.07 | 64.23 | 0.2206 |
| 12 | 88.43 | 30.02 | 64.62 | 0.2306 |
| 10 | 88.78 | 28.39 | 64.16 | 0.2186 |
| 8 | 89.43 | 28.3 | 64.51 | 0.2279 |
| 6 | 89.55 | 28.39 | 64.62 | 0.2307 |
| **4** | **89.08** | **29.67** | **64.86** | **0.237** |
| 2 | 88.61 | 26.59 | 63.32 | 0.1961 |
| **Dim(MNC)** |  |  |  |  |
| 4 | 89.02 | 26.42 | 63.5 | 0.2007 |
| **3** | **88.84** | **27.1** | **63.67** | **0.2055** |
| 2 | 89.85 | 25.56 | 63.64 | 0.2043 |
| 1 | 89.96 | 23.58 | 62.9 | 0.1837 |
| **Dim(DAC)** |  |  |  |  |
| **12** | **90.2** | **24.01** | **63.22** | **0.1925** |
| 10 | 91.68 | 21.01 | 62.87 | 0.1825 |
| 8 | 92.92 | 18.27 | 62.48 | 0.1713 |
| 6 | 93.21 | 17.58 | 62.38 | 0.1682 |
| 4 | 94.39 | 16.9 | 62.8 | 0.1832 |
| 2 | 95.22 | 13.55 | 61.92 | 0.1557 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2nd sub-classifier | Sn (%) | Sp (%) | Acc (%) | MCC |
| **Dim(BPB)** |  |  |  |  |
| 162 | 87.6 | 92.38 | 90.39 | 0.8018 |
| 160 | 87.81 | 92.38 | 90.48 | 0.8036 |
| 150 | 88.43 | 92.23 | 90.65 | 0.8073 |
| 140 | 88.84 | 92.38 | 90.91 | 0.8127 |
| **130** | **89.05** | **92.67** | **91.17** | **0.8179** |
| 120 | 89.05 | 92.23 | 90.91 | 0.8128 |
| 110 | 89.46 | 92.23 | 91.08 | 0.8164 |
| 100 | 89.26 | 92.08 | 90.91 | 0.8129 |
| 90 | 88.64 | 92.23 | 90.74 | 0.8091 |
| 80 | 88.84 | 92.23 | 90.82 | 0.811 |
| 70 | 88.84 | 92.82 | 91.17 | 0.8178 |
| 60 | 88.84 | 92.38 | 90.91 | 0.8127 |
| 50 | 89.88 | 91.79 | 90.99 | 0.815 |
| 40 | 88.84 | 91.35 | 90.31 | 0.8007 |
| 30 | 86.57 | 89.88 | 88.51 | 0.7636 |
| 20 | 83.06 | 88.12 | 86.02 | 0.712 |
| 10 | 82.02 | 81.96 | 81.99 | 0.6343 |
| **Dim(KNN)** |  |  |  |  |
| 20 | 85.33 | 90.76 | 88.51 | 0.7628 |
| 19 | 85.74 | 90.76 | 88.68 | 0.7665 |
| 17 | 85.74 | 91.2 | 88.94 | 0.7716 |
| 15 | 85.95 | 91.2 | 89.02 | 0.7735 |
| 13 | 85.74 | 91.35 | 89.02 | 0.7734 |
| 11 | 84.71 | 91.35 | 88.59 | 0.7643 |
| 9 | 84.92 | 91.5 | 88.77 | 0.7679 |
| 7 | 84.5 | 91.94 | 88.85 | 0.7695 |
| 5 | 85.54 | 91.35 | 88.94 | 0.7716 |
| **3** | **86.36** | **91.06** | **89.11** | **0.7754** |
| 1 | 79.55 | 88.56 | 84.82 | 0.6859 |
| **Dim(DNC)** |  |  |  |  |
| 16 | 26.24 | 85.78 | 61.06 | 0.1503 |
| 14 | 25.41 | 85.78 | 60.72 | 0.1409 |
| 12 | 22.31 | 89.3 | 61.49 | 0.158 |
| 10 | 22.31 | 89.59 | 61.66 | 0.1627 |
| **8** | **21.9** | **90.32** | **61.92** | **0.1698** |
| 6 | 17.56 | 92.38 | 61.32 | 0.1521 |
| 4 | 3.51 | 98.09 | 58.83 | 0.05 |
| 2 | 0 | 1 | 58.49 | NaN |
| **Dim(MNC)** |  |  |  |  |
| 4 | 2.69 | 98.24 | 58.58 | 0.0315 |
| **3** | **2.89** | **98.24** | **58.66** | **0.0378** |
| 2 | 0 | 99.71 | 58.32 | -0.0349 |
| 1 | 0 | 99.85 | 58.4 | -0.0247 |
| **Dim(DAC)** |  |  |  |  |
| **12** | **33.06** | **84.31** | **63.04** | **0.2037** |
| 10 | 24.59 | 87.83 | 61.58 | 0.1617 |
| 8 | 22.93 | 89.3 | 61.75 | 0.1653 |
| 6 | 21.49 | 90.18 | 61.66 | 0.1625 |
| 4 | 18.6 | 91.79 | 61.41 | 0.1546 |
| 2 | 13.22 | 94.13 | 60.55 | 0.1272 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3rd sub-classifier | Sn (%) | Sp (%) | Acc (%) | MCC |
| **Dim(BPB)** |  |  |  |  |
| 162 | 81.1 | 87.98 | 85.04 | 0.6934 |
| 160 | 81.1 | 87.98 | 85.04 | 0.6934 |
| 150 | 81.1 | 86.96 | 84.46 | 0.6818 |
| 140 | 81.1 | 88.24 | 85.19 | 0.6964 |
| 130 | 81.79 | 88.24 | 85.48 | 0.7025 |
| 120 | 81.44 | 87.72 | 85.04 | 0.6936 |
| 110 | 81.1 | 86.7 | 84.31 | 0.6789 |
| 100 | 81.1 | 87.72 | 84.9 | 0.6905 |
| 90 | 82.82 | 85.93 | 84.6 | 0.6861 |
| **80** | **83.51** | **87.47** | **85.78** | **0.7094** |
| 70 | 81.44 | 87.72 | 85.04 | 0.6936 |
| 60 | 80.41 | 87.21 | 84.31 | 0.6785 |
| 50 | 77.66 | 86.96 | 82.99 | 0.6508 |
| 40 | 76.63 | 84.4 | 81.09 | 0.6123 |
| 30 | 72.51 | 81.33 | 77.57 | 0.5402 |
| 20 | 74.57 | 86.96 | 81.67 | 0.6231 |
| 10 | 65.29 | 89.26 | 79.03 | 0.5695 |
| **Dim(KNN)** |  |  |  |  |
| 20 | 79.38 | 86.19 | 83.28 | 0.6575 |
| 19 | 79.38 | 86.45 | 83.43 | 0.6604 |
| 17 | 79.73 | 86.45 | 83.58 | 0.6636 |
| 15 | 79.38 | 85.68 | 82.99 | 0.6518 |
| 13 | 78.69 | 85.93 | 82.84 | 0.6484 |
| **11** | **80.07** | **86.7** | **83.87** | **0.6696** |
| 9 | 80.41 | 85.93 | 83.58 | 0.6641 |
| 7 | 79.38 | 86.19 | 83.28 | 0.6575 |
| 5 | 80.41 | 85.17 | 83.14 | 0.6555 |
| 3 | 79.38 | 86.19 | 83.28 | 0.6575 |
| 1 | 72.16 | 88.75 | 81.67 | 0.6232 |
| **Dim(DNC)** |  |  |  |  |
| 16 | 27.15 | 84.14 | 59.82 | 0.1379 |
| 14 | 28.18 | 84.14 | 60.26 | 0.1493 |
| 12 | 25.43 | 86.19 | 60.26 | 0.1472 |
| **10** | **26.8** | **85.93** | **60.7** | **0.159** |
| 8 | 22.68 | 85.93 | 58.94 | 0.1115 |
| 6 | 23.37 | 88.24 | 60.56 | 0.1538 |
| 4 | 18.9 | 89 | 59.09 | 0.1114 |
| 2 | 0 | 1 | 57.33 | NaN |
| **Dim(MNC)** |  |  |  |  |
| 4 | 3.09 | 97.44 | 57.18 | 0.0161 |
| 3 | 2.06 | 97.95 | 57.04 | 0.0006 |
| **2** | **1.72** | **99.49** | **57.77** | **0.0592** |
| 1 | 0 | 1 | 57.33 | NaN |
| **Dim(DAC)** |  |  |  |  |
| 12 | 32.3 | 76.98 | 57.92 | 0.1035 |
| 10 | 23.71 | 80.31 | 56.16 | 0.0485 |
| 8 | 19.59 | 86.7 | 58.06 | 0.0849 |
| **6** | **13.75** | **92.58** | **58.94** | **0.1038** |
| 4 | 9.28 | 93.86 | 57.77 | 0.059 |
| 2 | 4.12 | 99.49 | 58.8 | 0.126 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 4th sub-classifier | Sn (%) | Sp (%) | Acc (%) | MCC |
| **Dim(BPB)** |  |  |  |  |
| 162 | 80.37 | 85.96 | 83.63 | 0.6633 |
| 160 | 78.53 | 86.84 | 83.38 | 0.6568 |
| 150 | 79.14 | 86.4 | 83.38 | 0.6573 |
| 140 | 79.75 | 85.09 | 82.86 | 0.6479 |
| 130 | 79.14 | 85.09 | 82.61 | 0.6423 |
| 120 | 80.37 | 85.53 | 83.38 | 0.6584 |
| 110 | 79.75 | 85.53 | 83.12 | 0.6528 |
| 100 | 79.75 | 85.53 | 83.12 | 0.6528 |
| 90 | 78.53 | 85.96 | 82.86 | 0.6467 |
| 80 | 80.98 | 85.96 | 83.89 | 0.6689 |
| **70** | **82.21** | **86.4** | **84.65** | **0.685** |
| 60 | 81.6 | 85.53 | 83.89 | 0.6695 |
| 50 | 78.53 | 84.21 | 81.84 | 0.6268 |
| 40 | 76.69 | 84.65 | 81.33 | 0.6151 |
| 30 | 76.69 | 83.33 | 80.56 | 0.6002 |
| 20 | 73.62 | 79.39 | 76.98 | 0.5283 |
| 10 | 73.62 | 78.51 | 76.47 | 0.5188 |
| **Dim(KNN)** |  |  |  |  |
| 20 | 80.98 | 88.6 | 85.42 | 0.6991 |
| 19 | 80.98 | 89.04 | 85.68 | 0.7042 |
| 17 | 80.37 | 89.47 | 85.68 | 0.704 |
| 15 | 80.37 | 89.04 | 85.42 | 0.6988 |
| 13 | 80.98 | 90.35 | 86.45 | 0.7198 |
| 11 | 81.6 | 89.47 | 86.19 | 0.7148 |
| 9 | 82.21 | 88.6 | 85.93 | 0.71 |
| 7 | 82.21 | 89.47 | 86.45 | 0.7202 |
| **5** | **82.82** | **89.04** | **86.45** | **0.7206** |
| 3 | 84.05 | 87.28 | 85.93 | 0.7115 |
| 1 | 81.6 | 87.72 | 85.17 | 0.6944 |
| **Dim(DNC)** |  |  |  |  |
| 16 | 40.49 | 78.95 | 62.92 | 0.2109 |
| **14** | **42.33** | **78.51** | **63.43** | **0.2238** |
| 12 | 39.88 | 77.63 | 61.89 | 0.189 |
| 10 | 41.1 | 78.95 | 63.17 | 0.217 |
| 8 | 40.49 | 77.19 | 61.89 | 0.1899 |
| 6 | 39.88 | 77.63 | 61.89 | 0.189 |
| 4 | 32.52 | 80.7 | 60.61 | 0.1509 |
| 2 | 34.36 | 81.58 | 61.89 | 0.1813 |
| **Dim(MNC)** |  |  |  |  |
| 4 | 25.77 | 86.4 | 61.13 | 0.154 |
| **3** | **26.99** | **87.28** | **62.15** | **0.1806** |
| 2 | 26.38 | 86.84 | 61.64 | 0.1673 |
| 1 | 23.31 | 89.91 | 62.15 | 0.1797 |
| **Dim(DAC)** |  |  |  |  |
| **12** | **49.08** | **75** | **64.19** | **0.2488** |
| 10 | 47.24 | 71.93 | 61.64 | 0.1968 |
| 8 | 39.88 | 79.82 | 63.17 | 0.2154 |
| 6 | 28.22 | 80.7 | 58.82 | 0.1045 |
| 4 | 28.22 | 85.96 | 61.89 | 0.175 |
| 2 | 12.88 | 92.54 | 59.34 | 0.0903 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 5th sub-classifier | Sn (%) | Sp (%) | Acc (%) | MCC |
| **Dim(BPB)** |  |  |  |  |
| 162 | 94.78 | 91.49 | 93.42 | 0.8641 |
| 160 | 94.78 | 91.49 | 93.42 | 0.8641 |
| 150 | 94.03 | 91.49 | 92.98 | 0.8552 |
| **140** | **94.78** | **91.49** | **93.42** | **0.8641** |
| 130 | 93.28 | 90.43 | 92.11 | 0.8371 |
| 120 | 92.54 | 91.49 | 92.11 | 0.8377 |
| 110 | 93.28 | 91.49 | 92.54 | 0.8464 |
| 100 | 93.28 | 91.49 | 92.54 | 0.8464 |
| 90 | 93.28 | 91.49 | 92.54 | 0.8464 |
| 80 | 92.54 | 91.49 | 92.11 | 0.8377 |
| 70 | 92.54 | 90.43 | 91.67 | 0.8283 |
| 60 | 92.54 | 90.43 | 91.67 | 0.8283 |
| 50 | 92.54 | 90.43 | 91.67 | 0.8283 |
| 40 | 91.04 | 87.23 | 89.47 | 0.7828 |
| 30 | 93.28 | 86.17 | 90.35 | 0.8001 |
| 20 | 91.04 | 87.23 | 89.47 | 0.7828 |
| 10 | 88.81 | 79.79 | 85.09 | 0.6908 |
| **Dim(KNN)** |  |  |  |  |
| 20 | 93.28 | 85.11 | 89.91 | 0.791 |
| 19 | 93.28 | 85.11 | 89.91 | 0.791 |
| 17 | 93.28 | 85.11 | 89.91 | 0.791 |
| 15 | 93.28 | 85.11 | 89.91 | 0.791 |
| 13 | 92.54 | 86.17 | 89.91 | 0.7911 |
| 11 | 92.54 | 82.98 | 88.6 | 0.7636 |
| 9 | 93.28 | 84.04 | 89.47 | 0.7819 |
| 7 | 94.03 | 84.04 | 89.91 | 0.7911 |
| 5 | 94.78 | 84.04 | 90.35 | 0.8005 |
| 3 | 94.03 | 85.11 | 90.35 | 0.8002 |
| **1** | **96.27** | **82.98** | **90.79** | **0.8107** |
| **Dim(DNC)** |  |  |  |  |
| 16 | 73.13 | 55.32 | 65.79 | 0.2877 |
| 14 | 73.88 | 55.32 | 66.23 | 0.2959 |
| 12 | 78.36 | 60.64 | 71.05 | 0.396 |
| **10** | **79.1** | **60.64** | **71.49** | **0.4046** |
| 8 | 76.12 | 61.7 | 70.18 | 0.3809 |
| 6 | 78.36 | 54.26 | 68.42 | 0.3364 |
| 4 | 79.1 | 52.13 | 67.98 | 0.3251 |
| 2 | 82.84 | 38.3 | 64.47 | 0.2375 |
| **Dim(MNC)** |  |  |  |  |
| 4 | 88.81 | 7.45 | 55.26 | -0.0625 |
| **3** | **91.04** | **7.45** | **56.58** | **-0.0269** |
| 2 | 91.79 | 0 | 53.95 | -0.1886 |
| 1 | 1 | 0 | 58.77 | NaN |
| **Dim(DAC)** |  |  |  |  |
| 12 | 76.87 | 56.38 | 68.42 | 0.3393 |
| **10** | **76.12** | **58.51** | **68.86** | **0.3509** |
| 8 | 73.88 | 50 | 64.04 | 0.245 |
| 6 | 75.37 | 29.79 | 56.58 | 0.0574 |
| 4 | 83.58 | 25.53 | 59.65 | 0.1118 |
| 2 | 92.54 | 10.64 | 58.77 | 0.0553 |

**Table S3.** Performance comparison of different classifiers trained using different features on the jackknife test based on F-score selection for the first task I.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Feature** | **Dim** | **Sn (%)** | **Sp (%)** | **Acc (%)** | **MCC** |
| **KNN** | **15** | **85.56** | **86.68** | **86.12** | **0.7224** |
| KNN(15)+BPB | | | | | |
|  | 25 | 86.43 | 86.22 | 86.33 | 0.7266 |
|  | 35 | 86.15 | 86.4 | 86.28 | 0.7255 |
|  | 45 | 86.12 | 86.43 | 86.28 | 0.7255 |
|  | 55 | 86.19 | 86.19 | 86.19 | 0.7238 |
|  | 65 | 85.87 | 86.12 | 86 | 0.7199 |
|  | 75 | 86.26 | 86.12 | 86.19 | 0.7238 |
|  | 85 | 86.22 | 86.12 | 86.17 | 0.7234 |
|  | 95 | 86.75 | 85.94 | 86.35 | 0.7269 |
|  | 105 | 86.92 | 85.98 | 86.45 | 0.7291 |
|  | 115 | 86.99 | 85.52 | 86.26 | 0.7253 |
|  | 125 | 86.99 | 85.98 | 86.49 | 0.7298 |
|  | 135 | 87.13 | 86.01 | 86.57 | 0.7315 |
|  | **145** | **86.96** | **86.36** | **86.66** | **0.7332** |
|  | 155 | 86.82 | 86.4 | 86.61 | 0.7322 |
|  | 165 | 86.71 | 86.05 | 86.38 | 0.7276 |
|  | 175 | 86.71 | 85.94 | 86.33 | 0.7266 |
|  | 177 | 86.78 | 86.05 | 86.42 | 0.7283 |
| KNN(15)+BPB(130)+DNC | | | | | |
|  | 146 | 86.96 | 86.4 | 86.68 | 0.7336 |
|  | 147 | 87.03 | 86.43 | 86.73 | 0.7346 |
|  | 148 | 87.06 | 86.19 | 86.63 | 0.7325 |
|  | 149 | 87.06 | 86.12 | 86.59 | 0.7319 |
|  | 150 | 87.03 | 86.08 | 86.56 | 0.7312 |
|  | 151 | 86.85 | 86.43 | 86.64 | 0.7329 |
|  | 152 | 86.75 | 86.47 | 86.61 | 0.7322 |
|  | 153 | 87.03 | 86.57 | 86.8 | 0.736 |
|  | **154** | **87.06** | **86.54** | **86.8** | **0.736** |
|  | 155 | 86.99 | 86.4 | 86.7 | 0.7339 |
|  | 156 | 86.96 | 86.5 | 86.73 | 0.7346 |
|  | 157 | 86.96 | 86.5 | 86.73 | 0.7346 |
|  | 158 | 87.13 | 86.43 | 86.78 | 0.7357 |
|  | 159 | 87.03 | 86.43 | 86.73 | 0.7346 |
|  | 160 | 86.68 | 86.43 | 86.56 | 0.7311 |
|  | 161 | 86.64 | 86.4 | 86.52 | 0.7304 |
| KNN(15)+BPB(130)+DNC(9)+MNC | | | | | |
|  | **155** | **86.96** | **86.68** | **86.82** | **0.7364** |
|  | 156 | 86.82 | 86.71 | 86.77 | 0.7353 |
|  | 157 | 86.78 | 86.57 | 86.68 | 0.7336 |
|  | 158 | 86.85 | 86.57 | 86.71 | 0.7343 |
| KNN(15)+BPB(130)+DNC(9)+MNC(1)+DAC | | | | | |
|  | 156 | 86.92 | 86.64 | 86.78 | 0.7357 |
|  | 157 | 86.71 | 86.64 | 86.68 | 0.7336 |
|  | 158 | 86.75 | 86.68 | 86.71 | 0.7343 |
|  | 159 | 86.68 | 86.61 | 86.64 | 0.7329 |
|  | 160 | 86.78 | 86.61 | 86.7 | 0.7339 |
|  | 161 | 87.13 | 86.4 | 86.77 | 0.7353 |
|  | 162 | 87.06 | 86.4 | 86.73 | 0.7346 |
|  | 163 | 87.06 | 86.36 | 86.71 | 0.7343 |
|  | 164 | 87.03 | 86.33 | 86.68 | 0.7336 |
|  | **165** | **87.27** | **86.57** | **86.92** | **0.7385** |
|  | 166 | 87.1 | 86.5 | 86.8 | 0.736 |
|  | 167 | 87.2 | 86.5 | 86.85 | 0.7371 |
| **All features** | **165** | **87.27** | **86.57** | **86.92** | **0.7385** |

**Table S4.** Performance comparison of different sub-classifiers trained using different combinations of features on the jackknife test in terms of F-score selection for the second task II.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1st sub-classifier |  |  |  |  |  |
| **Feature** | **Dim** | **Sn (%)** | **Sp (%)** | **Acc (%)** | **MCC** |
| **KNN** | **15** | **90.26** | **75.64** | **84.3** | **0.6723** |
| KNN(15)+BPB | | | | | |
|  | 25 | 89.96 | 76.5 | 84.48 | 0.676 |
|  | 35 | 89.96 | 76.33 | 84.41 | 0.6745 |
|  | 45 | 90.02 | 76.5 | 84.51 | 0.6767 |
|  | 55 | 90.2 | 75.99 | 84.41 | 0.6745 |
|  | 65 | 90.2 | 76.24 | 84.51 | 0.6767 |
|  | 75 | 90.2 | 76.24 | 84.51 | 0.6767 |
|  | 85 | 89.96 | 76.33 | 84.41 | 0.6745 |
|  | 95 | 90.02 | 76.24 | 84.41 | 0.6745 |
|  | 105 | 90.08 | 76.24 | 84.44 | 0.6752 |
|  | 115 | 89.96 | 76.42 | 84.44 | 0.6753 |
|  | 125 | 90.02 | 76.24 | 84.41 | 0.6745 |
|  | 135 | 90.02 | 76.16 | 84.37 | 0.6738 |
|  | **145** | **90.02** | **76.59** | **84.55** | **0.6775** |
|  | 155 | 89.96 | 76.5 | 84.48 | 0.676 |
|  | 165 | 90.02 | 76.16 | 84.37 | 0.6738 |
|  | 175 | 90.02 | 76.5 | 84.51 | 0.6767 |
|  | 177 | 89.91 | 76.5 | 84.44 | 0.6753 |
| KNN(15)+BPB(130)+DNC | | | | | |
|  | 146 | 90.02 | 76.59 | 84.55 | 0.6775 |
|  | 147 | 90.02 | 76.59 | 84.55 | 0.6775 |
|  | 148 | 90.02 | 76.59 | 84.55 | 0.6775 |
|  | 149 | 90.02 | 76.59 | 84.55 | 0.6775 |
|  | 150 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 151 | 90.02 | 76.59 | 84.55 | 0.6775 |
|  | 152 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 153 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 154 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 155 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 156 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 157 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 158 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 159 | 89.96 | 76.67 | 84.55 | 0.6775 |
|  | 160 | 89.96 | 76.67 | 84.55 | 0.6775 |
|  | 161 | 89.96 | 76.59 | 84.51 | 0.6767 |
| KNN(15)+BPB(130)+MNC | | | | | |
|  | 146 | 90.02 | 76.59 | 84.55 | 0.6775 |
|  | 147 | 90.02 | 76.59 | 84.55 | 0.6775 |
|  | 148 | 89.96 | 76.59 | 84.51 | 0.6767 |
|  | 149 | 89.85 | 76.5 | 84.41 | 0.6745 |
| KNN(15)+BPB(130)+DAC | | | | | |
|  | 146 | 89.85 | 76.84 | 84.55 | 0.6775 |
|  | 147 | 89.96 | 77.27 | 84.79 | 0.6827 |
|  | 148 | 89.96 | 77.36 | 84.83 | 0.6834 |
|  | 149 | 90.08 | 77.62 | 85 | 0.6871 |
|  | 150 | 90.08 | 77.62 | 85 | 0.6871 |
|  | **151** | **90.14** | **78.13** | **85.24** | **0.6923** |
|  | 152 | 90.14 | 78.04 | 85.21 | 0.6915 |
|  | 153 | 90.08 | 78.13 | 85.21 | 0.6916 |
|  | 154 | 89.96 | 78.13 | 85.14 | 0.6801 |
|  | 155 | 89.96 | 78.13 | 85.14 | 0.6801 |
|  | 156 | 90.2 | 77.96 | 85.21 | 0.6915 |
|  | 157 | 90.2 | 77.96 | 85.21 | 0.6915 |
| **All features** | **151** | **90.14** | **78.13** | **85.24** | **0.6923** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 2nd sub-classifier |  |  |  |  |  |
| Feature | Dim | Sn (%) | Sp (%) | Acc (%) | MCC |
| **BPB** | **130** | **89.05** | **92.67** | **91.17** | **0.8179** |
| BPB(130)+KNN | | | | | |
|  | 131 | 88.64 | 92.08 | 90.65 | 0.8074 |
|  | 133 | 88.64 | 92.23 | 90.74 | 0.8091 |
|  | 135 | 88.84 | 92.08 | 90.74 | 0.8093 |
|  | 137 | 89.26 | 92.38 | 91.08 | 0.8163 |
|  | 139 | 89.05 | 92.52 | 91.08 | 0.8162 |
|  | 141 | 89.05 | 92.67 | 91.17 | 0.8179 |
|  | 143 | 89.46 | 92.38 | 91.17 | 0.8181 |
|  | 145 | 89.46 | 92.38 | 91.17 | 0.8181 |
|  | **147** | **89.05** | **93.11** | **91.42** | **0.8231** |
|  | 149 | 88.84 | 92.96 | 91.25 | 0.8196 |
|  | 150 | 88.84 | 92.82 | 91.17 | 0.8178 |
| BPB(130)+KNN(17)+DAC | | | | | |
|  | **148** | **89.26** | **93.11** | **91.51** | **0.8249** |
|  | 149 | 89.05 | 92.52 | 91.08 | 0.8162 |
|  | 150 | 88.64 | 92.23 | 90.74 | 0.8091 |
|  | 151 | 88.84 | 92.23 | 90.82 | 0.811 |
|  | 152 | 89.26 | 92.52 | 91.17 | 0.818 |
|  | 153 | 88.84 | 92.67 | 91.08 | 0.8161 |
|  | 154 | 88.84 | 92.38 | 90.91 | 0.8127 |
|  | 155 | 88.64 | 92.52 | 90.91 | 0.8126 |
|  | 156 | 88.22 | 92.67 | 90.82 | 0.8107 |
|  | 157 | 88.43 | 92.67 | 90.91 | 0.8125 |
|  | 158 | 88.43 | 92.52 | 90.82 | 0.8108 |
|  | 159 | 88.22 | 92.38 | 90.65 | 0.8072 |
| BPB(130)+KNN(17)+DAC(1)+DNC | | | | | |
|  | 149 | 89.26 | 92.96 | 91.42 | 0.8232 |
|  | 150 | 89.26 | 92.96 | 91.42 | 0.8232 |
|  | 151 | 89.26 | 92.96 | 91.42 | 0.8232 |
|  | 152 | 89.26 | 92.96 | 91.42 | 0.8232 |
|  | 153 | 89.46 | 92.82 | 91.42 | 0.8233 |
|  | 154 | 89.67 | 92.82 | 91.51 | 0.8251 |
|  | 155 | 89.05 | 92.96 | 91.34 | 0.8214 |
|  | 156 | 89.46 | 92.96 | 91.51 | 0.825 |
|  | 157 | 89.26 | 92.96 | 91.42 | 0.8232 |
|  | 158 | 89.67 | 92.96 | 91.6 | 0.8268 |
|  | 159 | 89.67 | 92.96 | 91.6 | 0.8268 |
|  | **160** | **89.88** | **92.96** | **91.68** | **0.8286** |
|  | 161 | 89.67 | 93.11 | 91.68 | 0.8285 |
|  | 162 | 89.26 | 93.26 | 91.6 | 0.8266 |
|  | 163 | 89.26 | 93.4 | 91.68 | 0.8284 |
|  | 164 | 89.26 | 93.4 | 91.68 | 0.8284 |
| BPB(130)+KNN(17)+DAC(1)+DNC(12)+MNC | | | | | |
|  | 161 | 89.05 | 93.4 | 91.6 | 0.8266 |
|  | 162 | 89.46 | 93.26 | 91.68 | 0.8285 |
|  | 163 | 89.05 | 93.4 | 91.6 | 0.8266 |
|  | 164 | 89.05 | 93.4 | 91.6 | 0.8266 |
| **All features** | **160** | **89.88** | **92.96** | **91.68** | **0.8286** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 3rd sub-classifier |  |  |  |  |  |
| Feature | Dim | Sn (%) | Sp (%) | Acc (%) | MCC |
| **BPB** | **80** | **83.51** | **87.47** | **85.78** | **0.7094** |
| BPB(80)+KNN | | | | | |
|  | 81 | 83.51 | 89.51 | 86.95 | 0.7326 |
|  | 83 | 83.16 | 88.75 | 86.36 | 0.7207 |
|  | 85 | 84.19 | 88.49 | 86.66 | 0.7272 |
|  | 87 | 82.47 | 89.77 | 86.66 | 0.7263 |
|  | 89 | 82.13 | 89.51 | 86.36 | 0.7203 |
|  | 91 | 82.13 | 89.26 | 86.22 | 0.7174 |
|  | 93 | 82.47 | 89.77 | 86.66 | 0.7263 |
|  | **95** | **83.51** | **90.28** | **87.39** | **0.7415** |
|  | 97 | 83.16 | 90.03 | 87.1 | 0.7354 |
|  | 99 | 82.82 | 90.79 | 87.39 | 0.7413 |
|  | 110 | 82.82 | 90.54 | 87.24 | 0.7383 |
| BPB(80)+KNN(15)+DNC | | | | | |
|  | 96 | 83.16 | 90.79 | 87.54 | 0.7444 |
|  | **97** | **83.85** | **91.05** | **87.98** | **0.7534** |
|  | 98 | 83.16 | 90.28 | 87.24 | 0.7384 |
|  | 99 | 82.47 | 90.28 | 86.95 | 0.7323 |
|  | 100 | 83.16 | 90.54 | 87.39 | 0.7414 |
|  | 101 | 83.16 | 90.54 | 87.39 | 0.7414 |
|  | 102 | 83.51 | 90.03 | 87.24 | 0.7385 |
|  | 103 | 83.16 | 89.77 | 86.95 | 0.7325 |
|  | 104 | 83.85 | 90.03 | 87.39 | 0.7416 |
|  | 105 | 83.51 | 89.51 | 86.95 | 0.7326 |
|  | 106 | 84.19 | 89.51 | 87.24 | 0.7388 |
|  | 107 | 83.85 | 89.51 | 87.1 | 0.7357 |
|  | 108 | 84.54 | 89.51 | 87.39 | 0.7419 |
|  | 109 | 83.85 | 89.77 | 87.24 | 0.7386 |
|  | 110 | 83.51 | 89.77 | 87.1 | 0.7355 |
|  | 111 | 83.16 | 89.51 | 86.8 | 0.7295 |
| BPB(80)+KNN(15)+DNC(2)+DAC | | | | | |
|  | 98 | 83.51 | 89.51 | 86.95 | 0.7326 |
|  | 99 | 82.47 | 89.77 | 86.66 | 0.7263 |
|  | 100 | 82.13 | 89 | 86.07 | 0.7144 |
|  | 101 | 81.79 | 89.26 | 86.07 | 0.7143 |
|  | 102 | 83.51 | 89.77 | 87.1 | 0.7355 |
|  | 103 | 83.16 | 90.28 | 87.24 | 0.7384 |
|  | 104 | 84.19 | 90.03 | 87.54 | 0.7446 |
|  | 105 | 83.85 | 90.28 | 87.54 | 0.7445 |
|  | 106 | 84.19 | 89.77 | 87.39 | 0.7417 |
|  | 107 | 83.16 | 88.49 | 86.22 | 0.7178 |
|  | 108 | 82.13 | 90.03 | 86.66 | 0.7263 |
|  | 109 | 81.79 | 89.77 | 86.36 | 0.7202 |
| BPB(80)+KNN(15)+DNC(2)+MNC | | | | | |
|  | 98 | 83.51 | 91.05 | 87.83 | 0.7504 |
|  | 99 | 83.16 | 91.05 | 87.68 | 0.7474 |
|  | 100 | 83.51 | 91.3 | 87.98 | 0.7534 |
|  | 101 | 84.54 | 90.28 | 87.83 | 0.7506 |
| **All features** | **97** | **83.85** | **91.05** | **87.98** | **0.7534** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 4th sub-classifier |  |  |  |  |  |
| Feature | Dim | Sn (%) | Sp (%) | Acc (%) | MCC |
| **KNN** | **5** | **82.82** | **89.04** | **86.45** | **0.7206** |
| KNN(5)+BPB | | | | | |
|  | 15 | 81.6 | 87.28 | 84.91 | 0.6894 |
|  | 25 | 80.37 | 87.28 | 84.4 | 0.6784 |
|  | 35 | 83.44 | 86.84 | 85.42 | 0.701 |
|  | 45 | 80.98 | 86.4 | 84.14 | 0.6739 |
|  | 55 | 84.66 | 88.6 | 86.96 | 0.732 |
|  | 65 | 84.05 | 87.28 | 85.93 | 0.7115 |
|  | 75 | 84.66 | 88.16 | 86.7 | 0.727 |
|  | **85** | **85.89** | **87.72** | **86.96** | **0.7331** |
|  | 95 | 85.89 | 85.96 | 85.93 | 0.7137 |
|  | 105 | 85.28 | 85.96 | 85.68 | 0.7081 |
|  | 115 | 84.05 | 85.96 | 85.17 | 0.6968 |
|  | 125 | 84.66 | 86.4 | 85.68 | 0.7073 |
|  | 135 | 84.66 | 86.4 | 85.68 | 0.7073 |
|  | 145 | 86.5 | 86.84 | 86.7 | 0.729 |
|  | 155 | 85.28 | 88.16 | 86.96 | 0.7325 |
|  | 165 | 85.28 | 87.72 | 86.7 | 0.7276 |
|  | 167 | 85.89 | 87.28 | 86.7 | 0.7282 |
| KNN(5)+BPB(80)+DAC | | | | | |
|  | 86 | 85.28 | 87.72 | 86.7 | 0.7276 |
|  | 87 | 84.05 | 86.84 | 85.68 | 0.7066 |
|  | 88 | 84.66 | 87.28 | 86.19 | 0.7171 |
|  | 89 | 85.28 | 88.16 | 86.96 | 0.7325 |
|  | 90 | 84.66 | 87.28 | 86.19 | 0.7171 |
|  | 91 | 84.66 | 86.84 | 85.93 | 0.7122 |
|  | 92 | 83.44 | 87.28 | 85.68 | 0.706 |
|  | 93 | 83.44 | 87.72 | 85.93 | 0.7109 |
|  | 94 | 83.44 | 87.72 | 85.93 | 0.7109 |
|  | 95 | 82.82 | 87.72 | 85.68 | 0.7054 |
|  | 96 | 82.21 | 89.04 | 86.19 | 0.7151 |
|  | 97 | 80.98 | 89.04 | 85.68 | 0.7042 |
| KNN(5)+BPB(80)+DNC | | | | | |
|  | 86 | 85.89 | 86.84 | 86.45 | 0.7234 |
|  | 87 | 85.28 | 86.84 | 86.19 | 0.7178 |
|  | 88 | 85.28 | 86.84 | 86.19 | 0.7178 |
|  | 89 | 85.28 | 86.84 | 86.19 | 0.7178 |
|  | 90 | 85.28 | 86.84 | 86.19 | 0.7178 |
|  | 91 | 85.89 | 86.84 | 86.45 | 0.7234 |
|  | 92 | 85.28 | 86.84 | 86.19 | 0.7178 |
|  | 93 | 85.89 | 87.28 | 86.7 | 0.7282 |
|  | 94 | 85.28 | 86.84 | 86.19 | 0.7178 |
|  | 95 | 85.28 | 87.28 | 86.45 | 0.7227 |
|  | 96 | 85.89 | 86.84 | 86.45 | 0.7234 |
|  | 97 | 85.28 | 86.4 | 85.93 | 0.7129 |
|  | 98 | 85.28 | 86.4 | 85.93 | 0.7129 |
|  | 99 | 84.66 | 86.4 | 85.68 | 0.7073 |
|  | 100 | 85.28 | 86.4 | 85.93 | 0.7129 |
|  | 101 | 84.05 | 86.84 | 85.68 | 0.7066 |
| KNN(5)+BPB(80)+MNC | | | | | |
|  | 86 | 84.05 | 87.28 | 85.93 | 0.7115 |
|  | 87 | 84.66 | 87.28 | 86.19 | 0.7171 |
|  | 88 | 84.05 | 87.28 | 85.93 | 0.7115 |
|  | 89 | 84.05 | 87.28 | 85.93 | 0.7115 |
| **All features** | **85** | **85.89** | **87.72** | **86.96** | **0.7331** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 5th sub-classifier |  |  |  |  |  |
| Feature | Dim | Sn (%) | Sp (%) | Acc (%) | MCC |
| **BPB** | **140** | **94.78** | **91.49** | **93.42** | **0.8641** |
| BPB(140)+KNN | | | | | |
|  | 141 | 95.52 | 92.55 | 94.3 | 0.8822 |
|  | **143** | **96.27** | **91.49** | **94.3** | **0.8821** |
|  | 145 | 95.52 | 91.49 | 93.86 | 0.873 |
|  | 147 | 94.78 | 92.55 | 93.86 | 0.8733 |
|  | 149 | 94.78 | 92.55 | 93.86 | 0.8733 |
|  | 151 | 94.78 | 92.55 | 93.86 | 0.8733 |
|  | 153 | 94.78 | 92.55 | 93.86 | 0.8733 |
|  | 155 | 95.52 | 92.55 | 94.3 | 0.8822 |
|  | 157 | 95.52 | 92.55 | 94.3 | 0.8822 |
|  | 159 | 95.52 | 92.55 | 94.3 | 0.8822 |
|  | 160 | 95.52 | 92.55 | 94.3 | 0.8822 |
| BPB(140)+KNN(3)+DNC | | | | | |
|  | **144** | **96.27** | **92.55** | **94.74** | **0.8912** |
|  | 145 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 146 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 147 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 148 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 149 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 150 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 151 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 152 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 153 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 154 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 155 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 156 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 157 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 158 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 159 | 96.27 | 92.55 | 94.74 | 0.8912 |
| BPB(140)+KNN(3)+DNC(1)+DAC | | | | | |
|  | 145 | 96.27 | 92.55 | 94.74 | 0.8912 |
|  | 146 | 95.52 | 92.55 | 94.3 | 0.8822 |
|  | **147** | **97.01** | **92.55** | **95.18** | **0.9003** |
|  | 148 | 97.01 | 91.49 | 94.74 | 0.8913 |
|  | 149 | 97.01 | 91.49 | 94.74 | 0.8913 |
|  | 150 | 96.27 | 91.49 | 94.3 | 0.8821 |
|  | 151 | 96.27 | 91.49 | 94.3 | 0.8821 |
|  | 152 | 96.27 | 90.43 | 93.86 | 0.8731 |
|  | 153 | 96.27 | 90.43 | 93.86 | 0.8731 |
|  | 154 | 96.27 | 90.43 | 93.86 | 0.8731 |
|  | 155 | 96.27 | 90.43 | 93.86 | 0.8731 |
|  | 156 | 96.27 | 90.43 | 93.86 | 0.8731 |
| BPB(140)+KNN(3)+DNC(1)+DAC(3)+MNC | | | | | |
|  | 165 | 96.27 | 91.49 | 94.3 | 0.8821 |
|  | 166 | 96.27 | 91.49 | 94.3 | 0.8821 |
|  | 167 | 96.27 | 91.49 | 94.3 | 0.8821 |
|  | 168 | 95.52 | 91.49 | 93.86 | 0.873 |
| **All features** | **147** | **97.01** | **92.55** | **95.18** | **0.9003** |

**Table S5.** Best performance results obtained for each feature combination for the first task I.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Task** | **Feature** | **Sn (%)** | **Sp (%)** | **Acc (%)** | **MCC** |
| I | KNN(15) | 85.56 | 86.68 | 86.12 | 0.7224 |
| KNN(15)+BPB(130) | 86.96 | 86.36 | 86.66 | 0.7332 |
| KNN(15)+BPB(130)+DNC(9) | 87.06 | 86.54 | 86.8 | 0.736 |
| KNN(15)+BPB(130)+DNC(9)+MNC(1) | 86.96 | 86.68 | 86.82 | 0.7364 |
| KNN(15)+BPB(130)+DNC(9)+MNC(1)+DAC(10) | 87.27 | 86.57 | 86.92 | 0.7385 |

**Table S6.** Best performance results obtained for each feature combination for the second task II.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Task** | **Sub-classifier** | **Feature** | **Sn (%)** | **Sp (%)** | **Acc (%)** | **MCC** |
| II | 1st | KNN(15) | 90.26 | 75.64 | 84.3 | 0.6723 |
| KNN(15)+BPB(130) | 90.02 | 76.59 | 84.55 | 0.6775 |
| KNN(15)+BPB(130)+DAC(6) | 90.14 | 78.13 | 85.24 | 0.6923 |
| 2nd | BPB(130) | 89.05 | 92.67 | 91.17 | 0.8179 |
| BPB(130)+KNN(17) | 89.05 | 93.11 | 91.42 | 0.8231 |
| BPB(130)+KNN(17)+DAC(1) | 89.26 | 93.11 | 91.51 | 0.8249 |
| BPB(130)+KNN(17)+DAC(1)+DNC(12) | 89.88 | 92.96 | 91.68 | 0.8286 |
| 3rd | BPB(80) | 83.51 | 87.47 | 85.78 | 0.7094 |
| BPB(80)+KNN(15) | 83.51 | 90.28 | 87.39 | 0.7415 |
| BPB(80)+KNN(15)+DNC(2) | 83.85 | 91.05 | 87.98 | 0.7534 |
| 4th | KNN(5) | 82.82 | 89.04 | 86.45 | 0.7206 |
| KNN(5)+BPB(80) | 85.89 | 87.72 | 86.96 | 0.7331 |
| 5th | BPB(140) | 94.78 | 91.49 | 93.42 | 0.8641 |
| BPB(140)+KNN(3) | 96.27 | 91.49 | 94.3 | 0.8821 |
| BPB(140)+KNN(3)+DNC(1) | 96.27 | 92.55 | 94.74 | 0.8912 |
| BPB(140)+KNN(3)+DNC(1)+DAC(3) | 97.01 | 92.55 | 95.18 | 0.9003 |

**Table S7.** Performance comparison between MULTiPly and iPromoter-2L for the second task II on 5-fold cross-validation test.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Promoter type** | **Method** | **Sn**  **(%)** | **Sp**  **(%)** | **Acc**  **(%)** | **MCC** |
| -promoter | iPromoter-2L | 95.34 | 59.35 | 80.66 | 0.6056 |
| MULTiPly | 90.43 | 76.93 | 84.91 | 0.6854 |
| -promoter | iPromoter-2L | 72.52 | 96.93 | 93.50 | 0.7338 |
| MULTiPly | 88.84 | 92.91 | 91.21 | 0.8189 |
| -promoter | iPromoter-2L | 52.58 | 99.14 | 94.41 | 0.6524 |
| MULTiPly | 82.2 | 88.41 | 85.67 | 0.7077 |
| -promoter | iPromoter-2L | 15.34 | 99.48 | 94.69 | 0.2962 |
| MULTiPly | 83.31 | 86.68 | 85.25 | 0.699 |
| -promoter | iPromoter-2L | 42.54 | 99.49 | 96.82 | 0.5708 |
| MULTiPly | 95.88 | 91.29 | 93.96 | 0.8759 |

**Table S8.** Performance comparison between MULTiPly and a direct multi-class SVM classifier.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Promoter** | **Method** | **TPa** | **FNb** | **TN** | **FP** |
| promoter | MULTiPly | **2496** | 364 | 276 | 384 |
| multi-class SVM | 1518 | 1342 | **2642** | 218 |
| -promoter | MULTiPly | 1527 | - | - | - |
| multi-class SVM | **1649** | - | - | - |
| -promoter | MULTiPly | 435 | - | - | - |
| multi-class SVM | 322 | - | - | - |
| -promoter | MULTiPly | 244 | - | - | - |
| multi-class SVM | 0 | - | - | - |
| -promoter | MULTiPly | 140 | - | - | - |
| multi-class SVM | 0 | - | - | - |
| -promoter | MULTiPly | 130 | - | - | - |
| multi-class SVM | 0 | - | - | - |

aTP represents the number of the predicted -promoter sequences;

bFN represents the number of the predicted non-promoter sequences.

**Table S9.** Performance comparison results of the multi-task predictor based on different sub-classifiers constructed using different numbers of trees.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Task** | **Sub-classifier** | **Tree** | **Sn (%)** | **Sp (%)** | **Acc (%)** | **MCC** |
| I |  | 50 | 85.31 | 85.87 | 85.59 | 0.7119 |
| 100 | 85.49 | 86.12 | 85.8 | 0.7161 |
| 150 | 85.77 | 86.61 | 86.19 | 0.7238 |
| 200 | 85.7 | 86.64 | 86.17 | 0.7235 |
| II | 1st | 50 | 89.55 | 77.27 | 84.55 | 0.6776 |
| 100 | 89.14 | 77.7 | 84.48 | 0.6763 |
| 150 | 89.43 | 77.62 | 84.62 | 0.6791 |
| 200 | 89.37 | 78.04 | 84.76 | 0.6822 |
| II | 2nd | 50 | 86.78 | 89.74 | 88.51 | 0.7638 |
| 100 | 86.16 | 90.47 | 88.68 | 0.7667 |
| 150 | 86.16 | 90.91 | 88.94 | 0.7718 |
| 200 | 86.98 | 90.47 | 89.02 | 0.7741 |
| II | 3rd | 50 | 76.98 | 85.93 | 82.11 | 0.6328 |
| 100 | 77.66 | 85.68 | 82.26 | 0.6362 |
| 150 | 78.35 | 87.21 | 83.43 | 0.6599 |
| 200 | 78.01 | 86.7 | 82.99 | 0.6509 |
| II | 4th | 50 | 82.82 | 85.96 | 84.65 | 0.6856 |
| 100 | 82.21 | 85.96 | 84.4 | 0.68 |
| 150 | 81.6 | 85.53 | 83.89 | 0.6695 |
| 200 | 84.66 | 84.65 | 84.65 | 0.688 |
| II | 5th | 50 | 94.78 | 88.3 | 92.11 | 0.8366 |
| 100 | 95.52 | 86.17 | 91.67 | 0.8278 |
| 150 | 94.78 | 86.17 | 91.23 | 0.8185 |
| 200 | 95.52 | 88.3 | 92.54 | 0.8458 |

**Table S10.** Performancecomparison of different classifiers for identifying promoters and their types using the jackknife tests.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Task** | **Classifier** | **Sn (%)** | **Sp (%)** | **Acc (%)** | **MCC** |
| I | Random Forest (150) | 85.77 | 86.61 | 86.19 | 0.7238 |
|  | 83.85 | 86.12 | 84.98 | 0.6998 |
| Ensemble for boosting (200) | 85.98 | 86.29 | 86.14 | 0.7227 |
| Discriminant analysis | 88.18 | 85.24 | 86.71 | 0.7346 |
| GBDT | 85.87 | 85.70 | 85.79 | 0.7157 |
| SVM | **87.27** | **86.57** | **86.92** | **0.7385** |
| II | Random Forest (200) | 89.37 | 78.04 | 84.76 | 0.6822 |
|  | 86.01 | 80.02 | 83.57 | 0.6599 |
| Ensemble for boosting (100) | 89.49 | 78.73 | 85.1 | 0.6896 |
| Discriminant analysis | 89.2 | 76.5 | 84.02 | 0.6666 |
| GBDT | 85.24 | 81.23 | 83.67 | 0.6602 |
| 1st SVM | **90.14** | **78.13** | **85.24** | **0.6923** |
| II | Random Forest (200) | 86.98 | 90.47 | 89.02 | 0.7741 |
|  | 88.64 | 88.12 | 88.34 | 0.7624 |
| Ensemble for boosting (150) | 84.5 | 89.74 | 87.56 | 0.7435 |
| Discriminant analysis | 89.46 | 89.3 | 89.37 | 0.783 |
| GBDT | 88.72 | 90.37 | 89.71 | 0.7874 |
| 2nd SVM | **89.88** | **92.96** | **91.68** | **0.8286** |
| II | Random Forest (150) | 78.35 | 87.21 | 83.43 | 0.6599 |
|  | 82.47 | 84.14 | 83.43 | 0.6633 |
| Ensemble for boosting (200) | 80.41 | 86.19 | 83.72 | 0.6669 |
| Discriminant analysis | 78.35 | 84.4 | 81.82 | 0.6281 |
| GBDT | 80.28 | 84.99 | 82.99 | 0.6521 |
| 3rd SVM | **83.85** | **91.05** | **87.98** | **0.7534** |
| II | Random Forest (200) | 84.66 | 84.65 | 84.65 | 0.688 |
|  | 87.12 | 84.65 | 85.68 | 0.7109 |
| Ensemble for boosting (200) | 80.98 | 86.4 | 84.14 | 0.6739 |
| Discriminant analysis | 81.6 | 86.84 | 84.65 | 0.6844 |
| GBDT | 82.72 | 87.34 | 85.42 | 0.6999 |
| 4th SVM | **85.89** | **87.72** | **86.96** | **0.7331** |
| II | Random Forest (200) | 95.52 | 88.3 | 92.54 | 0.8458 |
|  | 90.3 | 90.43 | 90.35 | 0.8027 |
| Ensemble for Boosting (200) | 96.27 | 88.3 | 92.98 | 0.8551 |
| Discriminant analysis | 79.85 | 73.4 | 77.19 | 0.531 |
| GBDT | 90.51 | 89.01 | 89.91 | 0.7911 |
| 5th SVM | **97.01** | **92.55** | **95.18** | **0.9003** |