Cross-Company and Within-Company Features for Software Effort Estimation

Analysis of features to define cross-company projects

Address(es) of author(s) should be given

Dataset	Criterion	Subsets	Subsets Size
finnish	application area ⋆	finnishAppArea2	5
		finnishAppArea7	7
		finnishAppArea10	7
finnish	application type	finnishAppType1	17
		finnishAppType2345	18
kemerer	hardware	kemererHardware1	7
		kemererHardware23456	8
maxwell	application type	maxwellAppType1	10
		maxwellAppType2	29
		maxwellAppType3	18
maxwell	hardware	maxwellHardware2	37
		maxwellHardware3	16
		maxwellHardware5	7
maxwell	source	maxwellSource1	8
		maxwellSource2	54

Fig. 1 Selected datasets, their division criterion and the size of subsets. Dataset marked with a "⋆" is too small for TEAK, so it is not included in the experimentation.

Dataset	WC			CC			MAR	
	Win	Tie	Loss	Win	Tie	Loss	WC	CC
finnishAppType1	0	20	0	0	20	0	3311.47	3864.63
finnishAppType2345	0	20	0	0	20	0	5514.82	5295.33
kemererHardware1	0	20	0	0	20	0	84.70	37.74
kemererHardware23456	0	20	0	0	20	0	163.90	145.03
maxwellAppType1	2	18	0	0	18	2	1650.90	2886.20
maxwellAppType2	0	18	2	2	18	0	6974.44	6230.54
maxwellAppType3	0	20	0	0	20	0	5977.15	5746.43
maxwellHardware2	0	20	0	0	20	0	5815.89	5207.78
maxwellHardware3	0	20	0	0	20	0	5202.09	5809.08
maxwellHardware5	0	14	6	6	14	0	5258.29	3205.04
maxwellSource1	0	20	0	0	20	0	2795.88	3218.47
maxwellSource2	0	20	0	0	20	0	5708.67	5722.83

Dataset	WC			CC			MMRE	
	Win	Tie	Loss	Win	Tie	Loss	WC	CC
finnishAppType1	0	20	0	0	20	0	0.85	0.94
finnishAppType2345	0	20	0	0	20	0	1.43	1.66
kemererHardware1	0	20	0	0	20	0	0.89	0.32
kemererHardware23456	0	20	0	0	20	0	0.72	0.55
maxwellAppType1	0	20	0	0	20	0	0.93	1.46
maxwellAppType2	0	19	1	1	19	0	1.38	1.17
maxwellAppType3	0	20	0	0	20	0	1.32	1.24
maxwellHardware2	0	20	0	0	20	0	0.77	0.84
maxwellHardware3	0	20	0	0	20	0	0.84	0.63
maxwellHardware5	0	20	0	0	20	0	3.93	2.66
maxwellSource1	0	20	0	0	20	0	0.53	0.45
maxwellSource2	0	20	0	0	20	0	1.00	1.27

Dataset	WC			CC			MdMRE	
	Win	Tie	Loss	Win	Tie	Loss	WC	CC
finnishAppType1	0	20	0	0	20	0	0.85	0.64
finnishAppType2345	0	20	0	0	20	0	1.43	0.68
kemererHardware1	0	20	0	0	20	0	0.89	0.19
kemererHardware23456	0	20	0	0	20	0	0.72	0.22
maxwellAppType1	0	20	0	0	20	0	0.93	0.73
maxwellAppType2	0	19	1	1	19	0	1.38	0.58
maxwellAppType3	0	20	0	0	20	0	1.32	0.52
maxwellHardware2	0	20	0	0	20	0	0.77	0.44
maxwellHardware3	0	20	0	0	20	0	0.84	0.52
maxwellHardware5	0	20	0	0	20	0	3.93	0.62
maxwellSource1	0	20	0	0	20	0	0.53	0.38
maxwellSource2	0	20	0	0	20	0	1.00	0.58

Dataset	WC			CC			Pred(25)	
	Win	Tie	Loss	Win	Tie	Loss	WC	CC
finnishAppType1	0	20	0	0	20	0	0.12	0.29
finnishAppType2345	0	20	0	0	20	0	0.33	0.22
kemererHardware1	0	20	0	0	20	0	0.29	0.29
kemererHardware23456	0	20	0	0	20	0	0.13	0.13
maxwellAppType1	0	20	0	0	20	0	0.20	0.40
maxwellAppType2	0	19	1	1	19	0	0.48	0.28
maxwellAppType3	0	20	0	0	20	0	0.39	0.22
maxwellHardware2	0	20	0	0	20	0	0.16	0.19
maxwellHardware3	0	20	0	0	20	0	0.19	0.31
maxwellHardware5	0	20	0	0	20	0	0.14	0.14
maxwellSource1	0	20	0	0	20	0	0.38	0.38
maxwellSource2	0	20	0	0	20	0	0.15	0.13

Fig. 2 Comparison of performance between cross-company and within-company datasets w.r.t. 4 different performance measures (MAR, MMRE, MdMRE, Pred(25)). In this version, the cross-company dataset consists of all the subsets (including projects of other subsets and within-company itself).

Dataset	WC			CC			MAR	
	Win	**Tie**	**Loss**	**Win**	**Tie**	**Loss**	**WC**	**CC**
finnishAppType1	0	20	0	0	20	0	3644.12	3852.54
finnishAppType2345	0	20	0	0	20	0	5704.00	5058.22
kemererHardware1	0	20	0	0	20	0	92.96	63.57
kemererHardware23456	0	20	0	0	20	0	191.19	181.61
maxwellAppType1	0	20	0	0	20	0	1591.68	2925.38
maxwellAppType2	0	20	0	0	20	0	6403.16	6316.47
maxwellAppType3	0	20	0	0	20	0	6007.68	6275.42
maxwellHardware2	0	20	0	0	20	0	5714.61	5798.44
maxwellHardware3	0	19	1	1	19	0	7235.78	5940.06
maxwellHardware5	0	20	0	0	20	0	3811.86	2728.18
maxwellSource1	0	20	0	0	20	0	2288.88	2981.78
maxwellSource2	1	19	0	0	19	1	5987.56	5986.74

Dataset	WC			CC			MMRE	
	Win	**Tie**	**Loss**	**Win**	**Tie**	**Loss**	**WC**	**CC**
finnishAppType1	0	20	0	0	20	0	1.17	0.80
finnishAppType2345	0	19	1	1	19	0	2.09	1.08
kemererHardware1	0	20	0	0	20	0	0.57	0.29
kemererHardware23456	0	20	0	0	20	0	0.82	0.60
maxwellAppType1	0	20	0	0	20	0	0.82	1.50
maxwellAppType2	0	19	1	1	19	0	0.97	0.68
maxwellAppType3	0	19	1	1	19	0	1.43	0.90
maxwellHardware2	0	20	0	0	20	0	0.85	0.88
maxwellHardware3	0	19	1	1	19	0	1.51	0.60
maxwellHardware5	0	20	0	0	20	0	2.29	1.81
maxwellSource1	0	20	0	0	20	0	0.40	0.38
maxwellSource2	1	19	0	0	19	1	1.10	1.87

Dataset	WC			CC			MdMRE	
	Win	**Tie**	**Loss**	**Win**	**Tie**	**Loss**	**WC**	**CC**
finnishAppType1	0	20	0	0	20	0	1.17	0.64
finnishAppType2345	0	19	1	1	19	0	2.09	0.63
kemererHardware1	0	20	0	0	20	0	0.57	0.20
kemererHardware23456	0	20	0	0	20	0	0.82	0.38
maxwellAppType1	0	20	0	0	20	0	0.82	0.59
maxwellAppType2	0	19	1	1	19	0	0.97	0.56
maxwellAppType3	0	19	1	1	19	0	1.43	0.57
maxwellHardware2	0	20	0	0	20	0	0.85	0.57
maxwellHardware3	0	19	1	1	19	0	1.51	0.49
maxwellHardware5	0	20	0	0	20	0	2.29	0.55
maxwellSource1	0	20	0	0	20	0	0.40	0.35
maxwellSource2	1	19	0	0	19	1	1.10	0.69

Dataset	WC			CC			Pred(25)	
	Win	**Tie**	**Loss**	**Win**	**Tie**	**Loss**	**WC**	**CC**
finnishAppType1	0	20	0	0	20	0	0.18	0.12
finnishAppType2345	0	19	1	1	19	0	0.22	0.28
kemererHardware1	0	20	0	0	20	0	0.29	0.29
kemererHardware23456	0	20	0	0	20	0	0.00	0.00
maxwellAppType1	0	20	0	0	20	0	0.50	0.30
maxwellAppType2	0	19	1	1	19	0	0.24	0.21
maxwellAppType3	0	19	1	1	19	0	0.39	0.33
maxwellHardware2	0	20	0	0	20	0	0.24	0.24
maxwellHardware3	0	19	1	1	19	0	0.06	0.13
maxwellHardware5	0	20	0	0	20	0	0.14	0.14
maxwellSource1	0	20	0	0	20	0	0.38	0.38
maxwellSource2	1	19	0	0	19	1	0.24	0.15

Fig. 3 Comparison of performance between cross-company and within-company datasets w.r.t. 4 different performance measures (MAR, MMRE, MdMRE, Pred(25)). In this version, the cross-company dataset does not include the within-company dataset.

WC Dataset	GAC2 Size	Subset1	Subset2	Subset3
finnishAppType1 (17)	2.6971	finnishAppType1 (1.3971)	finnishAppType2345 (1.3)	
finnishAppType2345 (18)	2.6167	finnishAppType1 (1.2889)	finnishAppType2345 (1.3278)	
kemererHardware1 (7)	1.4929	kemererHardware1 (0.91429)	kemererHardware23456 (0.57857)	
kemererHardware23456 (8)	1.4563	kemererHardware1 (0.80625)	kemererHardware23456 (0.65)	
maxwellAppType1 (10)	3.26	maxwellAppType1 (0.485)	maxwellAppType2 (1.755)	maxwellAppType3 (1.02)
maxwellAppType2 (29)	2.9672	maxwellAppType1 (0.45517)	maxwellAppType2 (1.5569)	maxwellAppType3 (0.95517)
maxwellAppType3 (18)	2.9472	maxwellAppType1 (0.45833)	maxwellAppType2 (1.5389)	maxwellAppType3 (0.95)
maxwellHardware2 (37)	2.6865	maxwellHardware2 (1.8081)	maxwellHardware3 (0.55135)	maxwellHardware5 (0.32703)
maxwellHardware3 (16)	3.5375	maxwellHardware2 (2.2313)	maxwellHardware3 (0.83125)	maxwellHardware5 (0.475)
maxwellHardware5 (7)	4.0143	maxwellHardware2 (2.6429)	maxwellHardware3 (0.95714)	maxwellHardware5 (0.41429)
maxwellSource1 (8)	3.7062	maxwellSource1 (0.375)	maxwellSource2 (3.3312)	
maxwellSource2 (54)	3.4324	maxwellSource1 (0.48704)	maxwellSource2 (2.9454)	

Fig. 4 The amount of instances selected from within and cross company datasets. In parenthesis next to each subset name, the amount of instances selected from that subset is written.