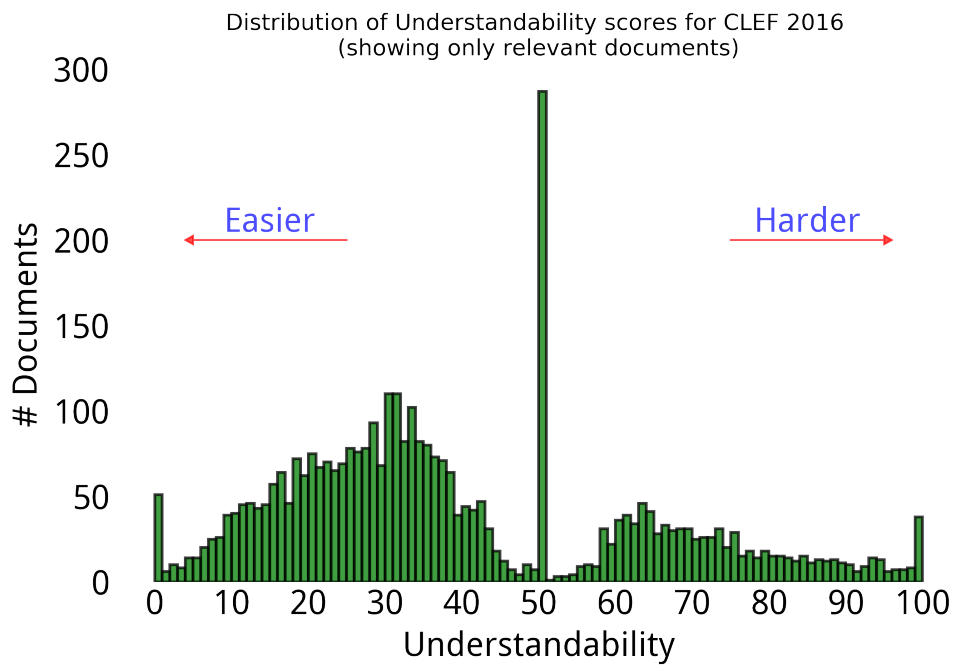# Appendix

Figure 1. Understandability label distribution for CLEF eHealth 2016
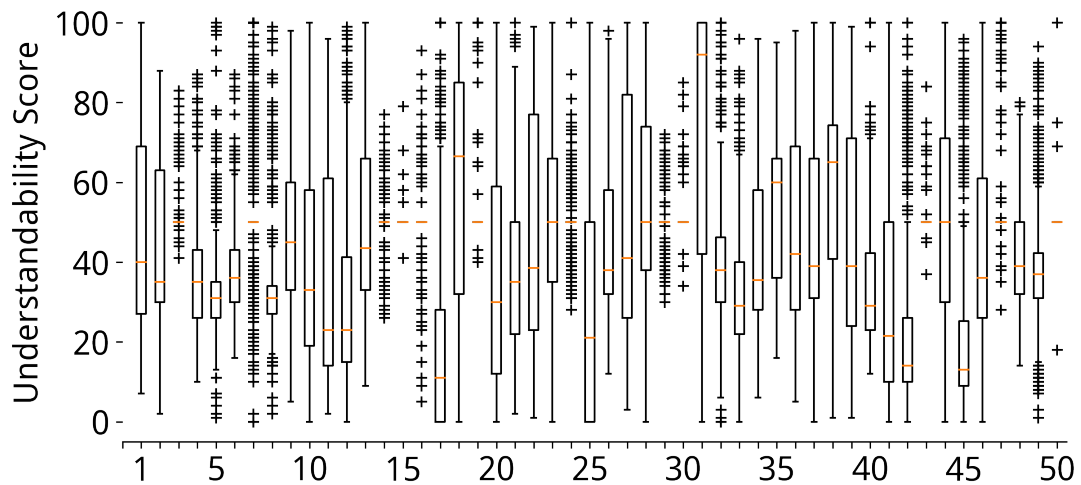

Figure 2. Understandability distribution broken per topic for CLEF eHealth 2016

Figure 3. In Palotti et al. [1], authors investigates the influence of HTML preprocessing when readability formulas are used to estimate Webpage understandability. They found that readability formulas are heavily affected by the methods used to extract text from the HTML source, but they did not measure how correlated each method was with a human ground truth. We further extended Palotti et al.'s work to understand the influence of HTML preprocessing on automatic understandability methods and establish best practices. We show the correlation of each preprocessing combination with the ground truth assessments for **CLEF 2015**.

References

[1] Palotti J, Zuccon G, Hanbury A. The Influence of Pre-processing on the Estimation of Readability of Web Documents. In: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. CIKM '15. New York, NY, USA: ACM; 2015. p. 1763–1766. doi:10.1145/2806416.2806613.
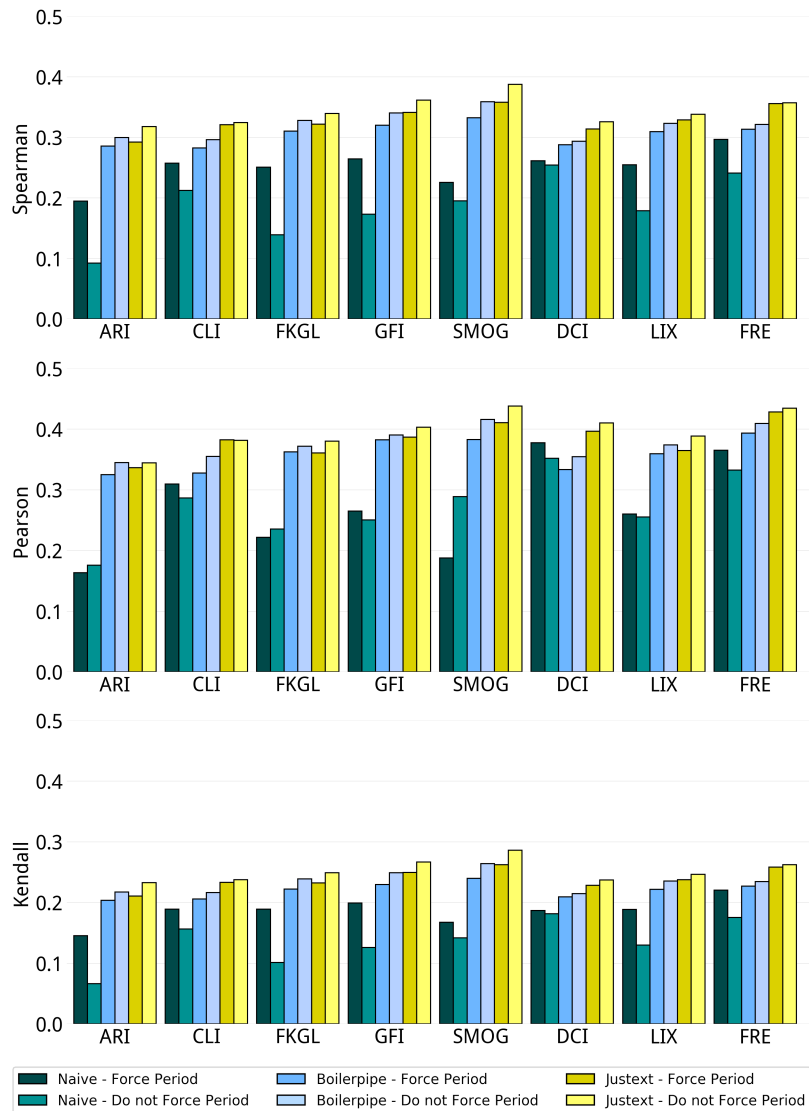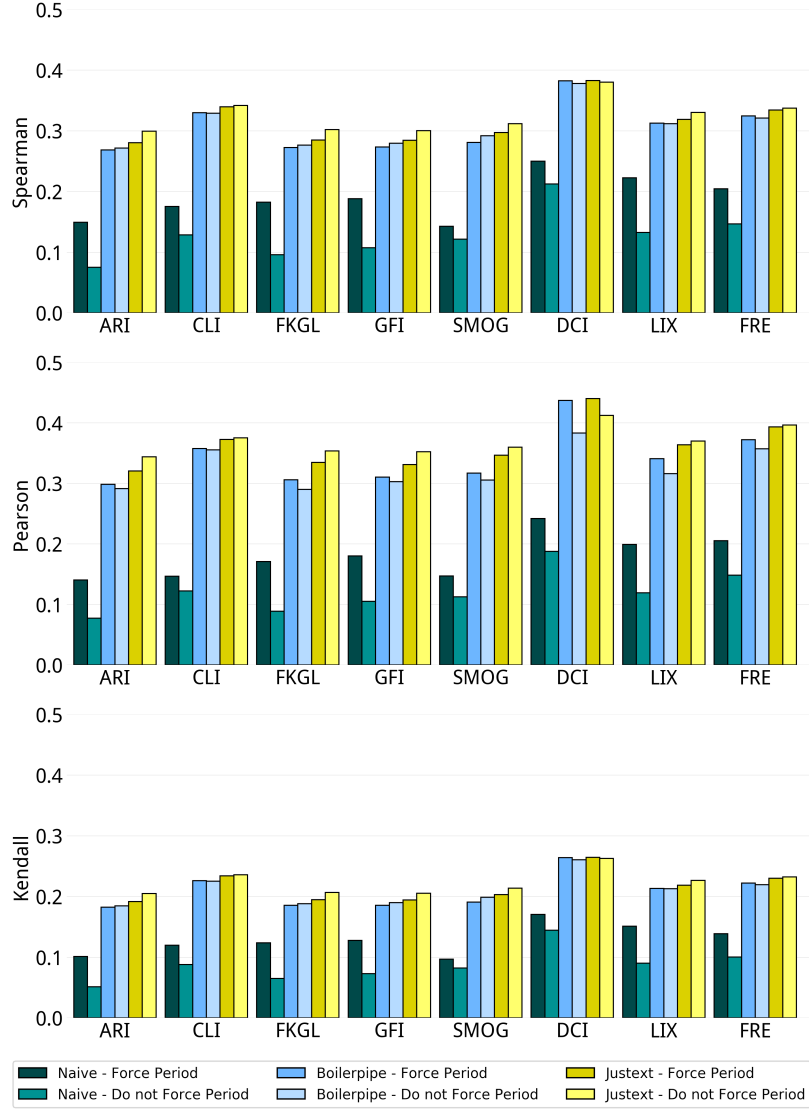
Figure 4. In Palotti et al. [1], authors investigates the influence of HTML preprocessing when readability formulas are used to estimate Webpage understandability. They found that readability formulas are heavily affected by the methods used to extract text from the HTML source, but they did not measure how correlated each method was with a human ground truth. We further extended Palotti et al.'s work to understand the influence of HTML preprocessing on automatic understandability methods and establish best practices. We show the correlation of each preprocessing combination with the ground truth assessments for **CLEF 2016**.

Figure 5. Correlations between understandability estimators and human assessments for **CLEF 2015**.

Figure 6. Correlations between understandability estimators and human assessments for **CLEF 2015**.

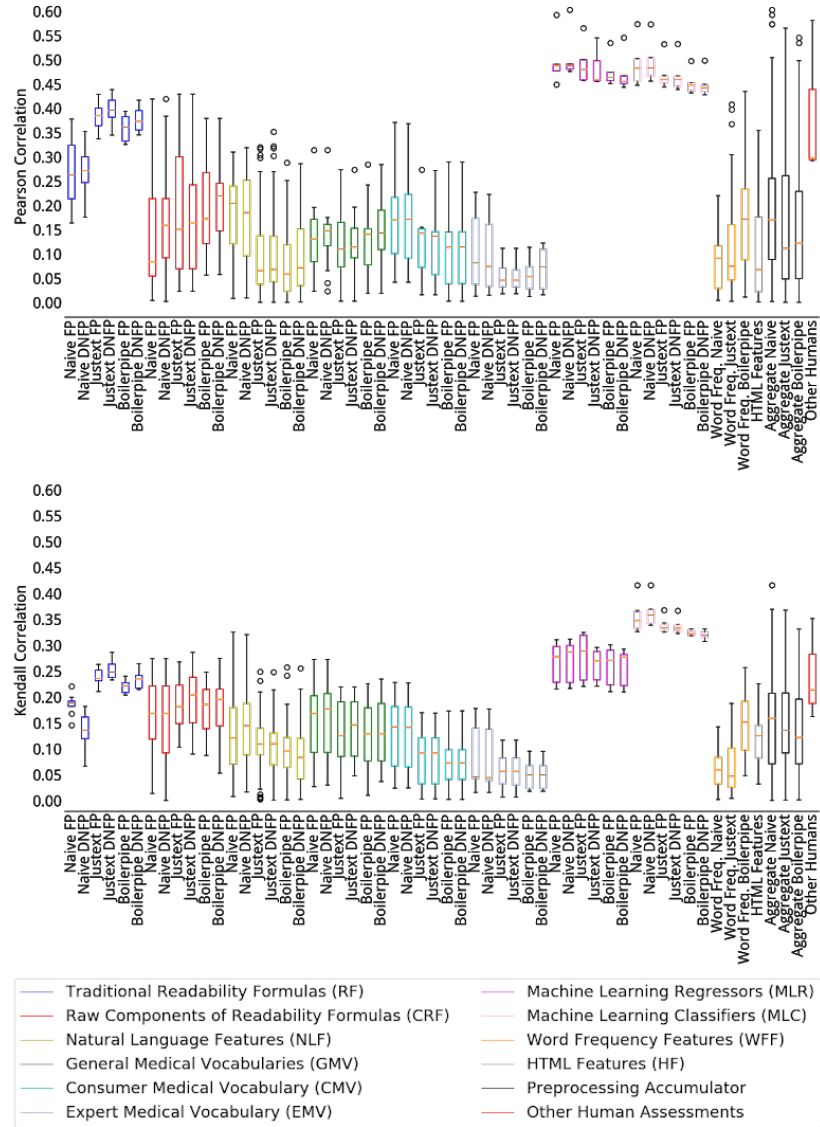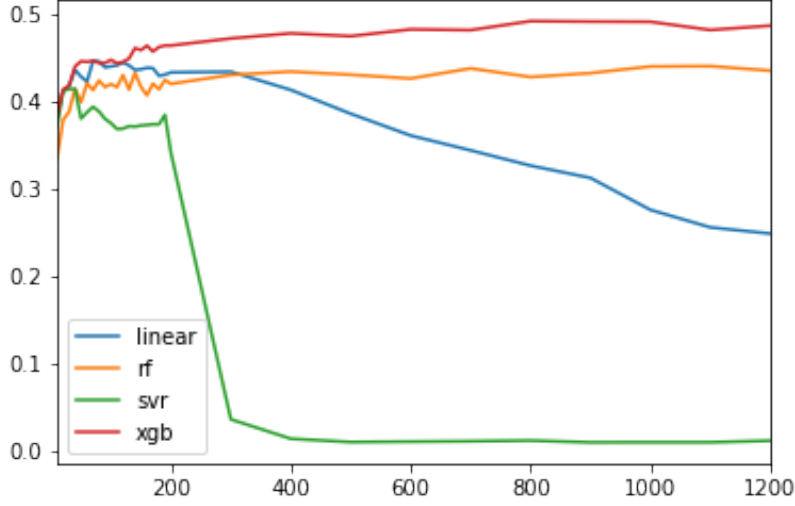**Figure 7.** To study the impact of feature sets, we varied the number of features each regressor algorithm was trained with. For that, we selected features with Python's f_regression algorithm from Sklearn package. The eXtreme Gradient Boosting (XGB) and Random Forest Regressor were marginally influenced by the selection of features, therefore we used XGB with all features devised in our experiments.

Index	Rerank	Run	Official CLEF 2016 Metrics				New Metrics to Evaluate Underst. in Retrieval - Sec. 3							
			RBP	RBP Res.	$uRBP$	uRBP Res.	RBP_u	RBP_u Res.	$HRBP$	HRBP Res.	Unj	RBP_r^*	RBP_u^*	$HRBP^*$
1		GUIR [1] (Best Run)	**28.11**	7.65	**18.12**	7.19	**45.69**	8.86	**25.61**	6.50	0.01	**28.29**	46.03	**25.79**
2	No Rerank	ECNU [2] (Runner Up)	27.70	7.37	17.55	**7.34**	43.89°	8.66	25.35	6.26	0.01	27.77	44.18°	25.48
3		Plain BM25 Baseline	25.28°	**8.24**	16.05°	6.94	42.08°	**10.97**	22.97°	**7.19**	0.06	26.01°	43.89°	23.93°
4		Based on GUIR	24.70†°	8.70	16.83†°	7.27	49.10†°	10.62	24.94	7.50	0.03	25.24†°	50.33†°	25.54
5	Dale-Chall Top 15	Based on ECNU	24.78†°	7.83	16.64°	7.16	48.88†°	9.71	24.80	6.50	0.02	25.12†°	49.64†°	25.21
6		Based on BM25	23.22†°	8.78	15.85 °	6.94	47.09†°	11.83	24.01	7.42	0.07	24.04†°	48.60†°	24.82
7		Based on GUIR	22.19†°	9.37	15.36†°	6.98	48.71†°	12.30	23.21†°	8.12	0.06	23.26†°	51.39†°	24.45†°
8	Dale-Chall Top 20	Based on ECNU	23.01†°	8.93	15.70†°	6.91	48.99†°	11.69	23.73†°	7.80	0.05	23.84†°	51.00†°	24.66
9		Based on BM25	21.58†°	9.51	14.83†°	7.02	46.99†	13.00	22.89°	8.06	0.09	22.93†°	49.55†°	24.26
10		Based on GUIR	16.18†°	15.24	11.56†°	6.80	41.79†°	24.49	18.10†°	14.42	0.22	20.90†°	53.28†°	23.27†°
11	Dale-Chall Top 50	Based on ECNU	16.88†°	17.37	11.78†°	**7.30**	40.76†°	23.77	18.30†°	**15.57**	**0.24**	21.34†°	52.07†°	23.33†°
12		Based on BM25	15.06†°	15.35†°	10.55	6.62	40.03°	23.88	16.55†°	13.83	**0.24**	19.42†°	51.69†°	21.59†°
13		Based on GUIR	**25.16†°**	8.09	**17.27†°**	7.12	**50.96†°**	10.11	25.16	6.89	0.02	**25.61†°**	52.00†°	**25.68**
14	XGB Top 15	Based on ECNU	24.18†°	7.69	16.54 °	7.09	50.00†°	9.91	24.56	6.65	0.02	24.56†°	50.74†°	25.01
15		Based on BM25	22.33†°	8.14	15.46	6.76	47.90†°	12.13	22.89°	7.25	0.07	23.11†°	49.43†°	23.69°
16		Based on GUIR	22.38†°	9.49	15.61†°	7.05	50.45†°	12.08	23.30†°	8.16	0.05	23.62†°	52.98†°	24.68
17	XGB Top 20	Based on ECNU	22.95†°	8.82	15.95†°	7.02	50.42†°	11.70	23.97†°	7.56	0.04	23.68†°	52.15†°	24.73
18		Based on BM25	20.65†°	9.42	14.46†°	6.84	47.74†°	13.56	21.93°	8.34	0.09	21.98†°	50.28†°	23.27°
19		Based on GUIR	16.65†°	15.73	12.39†°	6.84	43.49†°	23.63	18.70†°	13.74	0.22	21.13†°	**55.07†°**	23.58†°
20	XGB Top 50	Based on ECNU	16.19†°	**17.01**	11.82†°	7.27	43.05°	**24.75**	18.27†°	14.41	**0.24**	20.16†°	54.70†°	22.96†°
21		Based on BM25	15.43†°	15.37	11.33†°	6.48	41.93°	23.65	17.43†°	13.40	0.26	19.58†°	54.04†°	22.17†°
22		Based on GUIR	27.23†°	7.76	**18.31**	**7.23**	49.69†°	9.18	26.49†°	6.62	0.01	**27.46†°**	50.07†°	**26.69†°**
23	RRF (XGB & Orig.) Top 15	Based on ECNU	26.60†°	7.41	17.81	7.19	48.67†°	8.80	26.02	6.09	0.01	26.76†°	49.10†°	26.27†
24		Based on BM25	24.57°	8.15	16.51°	6.91	46.76†	11.23	24.16†	7.20	0.06	25.32°	48.52†°	25.08†
25		Based on GUIR	26.21†°	7.96	17.73	7.19	50.29†°	9.58	25.89	6.73	0.03	26.53†°	50.98†°	26.25
26	RRF (XGB & Orig.) Top 20	Based on ECNU	26.15†°	7.64	17.69	7.09	49.70†°	9.28	**26.07**	6.39	0.02	26.38†°	50.32†°	26.35
27		Based on BM25	24.04†°	8.24	16.32°	6.87	47.69†°	11.40	24.08†°	7.35	0.06	24.82†°	49.52†°	25.01†
28		Based on GUIR	24.09†°	**9.44**	16.85†°	7.02	50.55†°	11.76	24.76	**8.01**	0.07	25.08†°	**52.84†°**	25.84
29	RRF (XGB & Orig.) Top 50	Based on ECNU	24.17†°	8.67	16.75°	7.12	**50.63†°**	11.66	25.00	7.61	0.07	24.90†°	52.50†°	25.84
30		Based on BM25	22.28†°	8.87	15.50	6.76	48.79†°	**12.90**	23.13†°	7.82	**0.10**	23.46†°	51.89†°	24.57
*		Based on GUIR	26.93†°	7.95	**17.96**	7.34	48.40†°	9.28	26.20	6.77	0.02	**27.20†°**	48.86†°	26.49
*	RRF (DCI & Orig.) Top 15	Based on ECNU	**26.94**	7.53	17.76	7.30	47.78†°	8.94	26.18	6.22	0.01	27.10	48.29†°	26.42†
*		Based on BM25	24.91	8.12	16.58	7.09	45.85†	10.75	24.67†	6.99	0.06	25.68	47.55†°	25.56†
*		Based on GUIR	26.09†°	8.11	17.58	7.27	48.79†°	9.76	25.88	6.88	0.03	26.47†°	49.58†°	26.29
*	RRF (DCI & Orig.) Top 20	Based on ECNU	26.51†°	7.88	17.56	**7.37**	48.57†°	9.53	**26.27**	6.59	0.02	26.75†°	49.33†°	**26.58†**
*		Based on BM25	24.73	8.21	16.63	7.12	46.63†	11.00	24.85†	7.11	0.06	25.51°	48.44†°	25.77†
*		Based on GUIR	23.39†°	**9.72**	16.21†°	7.16	**49.11†°**	12.40	24.20†°	**8.33**	0.07	24.60†°	**51.45†°**	25.52
*	RRF (DCI & Orig.) Top 50	Based on ECNU	24.26†°	9.23	16.43 °	7.34	48.90†°	11.54	24.84	7.81	0.06	25.15†°	50.91†°	25.85
*		Based on BM25	22.64†°	9.26	15.47	7.16	47.25†°	12.79	23.67	7.86	**0.09**	23.89†°	50.26†°	25.11
31		Combo 1 on BM25	20.42†°	17.61	13.00†°	7.41	32.17†°	24.61	18.39†°	14.41	0.28	25.25°	43.19°	23.83°
32		Combo 2 on BM25	24.98†°	19.83	15.30†°	8.09	35.09†°	25.14	22.26°	17.50	0.24	30.41	46.09	28.28†°
33	XGB LeToR	Combo 3 on BM25	26.35†	**20.48**	15.88†°	8.16	34.73†°	24.69	21.81†	17.41	0.22	32.25°	45.44	28.22†°
34		Combo 4 on BM25	16.16†°	19.48	10.76†°	7.27	**36.75†°**	**28.51**	16.77†°	**17.80**	**0.29**	22.20†°	**50.06†°**	23.32°
35		Combo 5 on BM25	**26.76°**	**20.48**	**16.19°**	**8.34**	35.26†°	24.13	**22.96**	17.59	0.22	**32.60†°**	45.87	**29.20†°**

Figure 8. We expanded retrieval results including the experiments with Reciprocal Rank Fusion (RRF) of Dale-Chall Index Runs (indices 4-12) and the original runs (indices 1-3). These results confirm the superiority of using XGB rather than DCI in terms of understandability of the results retrieved.

Index	Rerank	Run	Official CLEF 2016 Metrics				New Metrics to Evaluate Underst. in Retrieval - Sec. 3							
			RBP	RBP Res.	$uRBP$	uRBP Res.	RBP_u	RBP_u Res.	$HRBP$	HRBP Res.	Unj	RBP^*_u	RBP^*_u	$HRBP^*$
1	No Rerank	ECNU (Best Run)	**51.57**	**8.95**	**50.51**	**8.95**	59.55	10.09	**46.22**	8.62	0.00	**51.57**	59.55	**46.22**
2		KISTI (Runner Up)	36.72°	8.06	35.92°	7.32	64.50	11.54	37.56°	7.89	**0.03**	37.07°	65.31	37.96°
3		Plain BM25 Baseline	31.20°	8.76	30.51°	7.65	**67.60**	**12.20**	35.75°	**8.76**	**0.03**	31.57°	68.94°	36.42°
4	SMOG Top 15	Based on ECNU	$\mathbf{38.16}^{\dagger\circ}$	20.09	$\mathbf{37.36}^{\dagger\circ}$	8.95	$55.11^{\dagger\circ}$	22.14	$\mathbf{37.59}^{\dagger\circ}$	18.98	0.14	$45.45^{\dagger\circ}$	62.58	43.65
5		Based on KISTI	$31.28^{\dagger\circ}$	10.21	$30.59^{\dagger\circ}$	7.48	67.05	13.95	34.23°	10.05	0.05	$33.12^{\dagger\circ}$	69.10°	36.02°
6		Based on BM25	$24.39^{\dagger\circ}$	11.46	$23.83^{\dagger\circ}$	7.16	**67.69**	17.66	$28.97^{\dagger\circ}$	11.46	0.10	$26.28^{\dagger\circ}$	72.82°	$31.53^{\dagger\circ}$
7	SMOG Top 20	Based on ECNU	$34.88^{\dagger\circ}$	25.96	$34.13^{\dagger\circ}$	**8.95**	$54.15^{\dagger\circ}$	27.91	$35.34^{\dagger\circ}$	25.02	0.20	$44.87^{\dagger\circ}$	$66.20^{\dagger\circ}$	**44.88**
8		Based on KISTI	$28.01^{\dagger\circ}$	11.05	$27.39^{\dagger\circ}$	7.32	66.61	15.44	$30.93^{\dagger\circ}$	10.82	0.08	$30.87^{\dagger\circ}$	69.99°	33.81°
9		Based on BM25	$22.83^{\dagger\circ}$	13.48	$22.31^{\dagger\circ}$	7.32	63.61	22.44	$26.52^{\dagger\circ}$	13.48	0.15	$26.22^{\dagger\circ}$	72.72°	$30.87^{\dagger\circ}$
10	SMOG Top 50	Based on ECNU	$21.83^{\dagger\circ}$	**36.63**	$21.34^{\dagger\circ}$	**8.95**	$41.71^{\dagger\circ}$	46.89	$23.65^{\dagger\circ}$	36.63	**0.45**	$39.88^{\dagger\circ}$	71.61°	43.86
11		Based on KISTI	$21.00^{\dagger\circ}$	18.67	$20.53^{\dagger\circ}$	7.32	59.72	26.67	$24.88^{\dagger\circ}$	18.67	0.23	$27.20^{\dagger\circ}$	72.47°	31.90°
12		Based on BM25	$15.20^{\dagger\circ}$	18.73	$14.86^{\dagger\circ}$	6.51	$49.87^{\dagger\circ}$	36.41	$17.79^{\dagger\circ}$	18.73	0.32	$21.22^{\dagger\circ}$	73.17°	$25.78^{\dagger\circ}$
13	XGB Top 15	Based on ECNU	$\mathbf{34.94}^{\dagger\circ}$	22.28	$\mathbf{34.31}^{\dagger\circ}$	9.11	$40.84^{\dagger\circ}$	24.50	$28.60^{\dagger\circ}$	20.64	0.17	$45.21^{\dagger\circ}$	52.18°	$\mathbf{38.14}^{\dagger\circ}$
14		Based on KISTI	$31.86^{\dagger\circ}$	11.65	$31.26^{\dagger\circ}$	7.32	$49.38^{\dagger\circ}$	15.90	$\mathbf{29.82}^{\dagger\circ}$	10.97	0.06	34.48°	53.55°	$32.33^{\dagger\circ}$
15		Based on BM25	$23.48^{\dagger\circ}$	10.51	$22.98^{\dagger\circ}$	6.67	$\mathbf{53.69}^{\dagger\circ}$	14.56	$25.89^{\dagger\circ}$	10.35	0.08	$24.95^{\dagger\circ}$	$58.40^{\dagger}$	$27.92^{\dagger\circ}$
16	XGB Top 20	Based on ECNU	$31.03^{\dagger\circ}$	30.50	$30.50^{\dagger\circ}$	**9.27**	$34.80^{\dagger\circ}$	27.97	$23.89^{\dagger\circ}$	24.70	0.26	$\mathbf{45.68}^{\dagger\circ}$	50.53°	$37.78^{\dagger\circ}$
17		Based on KISTI	$29.05^{\dagger\circ}$	15.53	$28.53^{\dagger\circ}$	7.81	$43.04^{\dagger\circ}$	19.55	$26.07^{\dagger\circ}$	13.45	0.12	$34.02^{\dagger\circ}$	50.45°	$30.65^{\dagger\circ}$
18		Based on BM25	$21.54^{\dagger\circ}$	13.70	$20.85^{\dagger\circ}$	6.51	$44.97^{\dagger\circ}$	21.53	$22.84^{\dagger\circ}$	13.00	0.15	25.10	$55.55^{\dagger}$	$27.43^{\dagger\circ}$
19	XGB Top 50	Based on ECNU	$19.11^{\dagger\circ}$	**38.27**	$18.61^{\dagger\circ}$	8.79	$16.62^{\dagger\circ}$	37.25	$11.58^{\dagger\circ}$	28.01	**0.55**	$43.46^{\dagger\circ}$	45.35°	$33.65^{\dagger\circ}$
20		Based on KISTI	$19.95^{\dagger\circ}$	21.43	$19.61^{\dagger\circ}$	7.32	$22.01^{\dagger\circ}$	24.45	$13.22^{\dagger\circ}$	14.22	0.35	$29.50^{\dagger\circ}$	$36.58^{\dagger\circ}$	$21.61^{\dagger\circ}$
21		Based on BM25	$17.94^{\dagger\circ}$	16.20	$17.45^{\dagger\circ}$	6.51	$23.18^{\dagger\circ}$	26.72	$12.82^{\dagger\circ}$	12.44	0.37	26.14	$42.20^{\dagger\circ}$	$21.16^{\dagger\circ}$
22	RRF (XGB & Orig.) Top 15	Based on ECNU	$\mathbf{44.59}^{\dagger\circ}$	13.25	$\mathbf{43.73}^{\dagger\circ}$	8.95	$51.11^{\dagger\circ}$	14.52	$\mathbf{38.04}^{\dagger\circ}$	12.44	0.10	$47.76^{\dagger\circ}$	54.68°	$\mathbf{41.13}^{\dagger\circ}$
23		Based on KISTI	34.89°	9.65	34.16°	7.48	$57.54^{\dagger}$	14.06	$34.27^{\dagger\circ}$	9.48	0.05	36.31°	$60.30^{\dagger}$	$35.73^{\dagger\circ}$
24		Based on BM25	$27.71^{\dagger\circ}$	10.88	$27.10^{\dagger\circ}$	7.65	$\mathbf{60.05}^{\dagger}$	15.01	$31.50^{\dagger\circ}$	10.88	0.07	$28.75^{\dagger\circ}$	$63.50^{\dagger}$	$33.17^{\dagger\circ}$
25	RRF (XGB & Orig.) Top 20	Based on ECNU	$41.90^{\dagger\circ}$	16.06	$41.11^{\dagger\circ}$	9.11	$47.80^{\dagger\circ}$	17.59	$34.95^{\dagger\circ}$	15.23	0.14	$47.12^{\dagger\circ}$	53.34°	$39.97^{\dagger\circ}$
26		Based on KISTI	$34.55^{\dagger\circ}$	10.59	$33.84^{\dagger\circ}$	7.65	$55.02^{\dagger}$	15.12	$33.71^{\dagger\circ}$	10.43	0.07	36.37°	$58.49^{\dagger}$	35.65°
27		Based on BM25	$26.01^{\dagger\circ}$	12.50	$25.39^{\dagger\circ}$	7.65	$56.23^{\dagger}$	16.47	$29.22^{\dagger\circ}$	12.53	0.10	$27.63^{\dagger\circ}$	$60.91^{\dagger}$	$31.60^{\dagger\circ}$
28	RRF (XGB & Orig.) Top 50	Based on ECNU	$34.80^{\dagger\circ}$	**25.30**	$34.18^{\dagger\circ}$	**9.44**	$38.32^{\dagger\circ}$	24.97	$27.26^{\dagger\circ}$	21.84	**0.25**	$46.22^{\dagger\circ}$	50.59°	$38.15^{\dagger\circ}$
29		Based on KISTI	$30.05^{\dagger\circ}$	15.44	$29.48^{\dagger\circ}$	7.97	$43.57^{\dagger\circ}$	19.94	$27.45^{\dagger\circ}$	14.76	0.14	$34.36^{\dagger\circ}$	50.70°	$31.93^{\dagger\circ}$
30		Based on BM25	$18.97^{\dagger\circ}$	17.79	$18.52^{\dagger\circ}$	7.48	$43.90^{\dagger\circ}$	23.92	$20.43^{\dagger\circ}$	17.84	0.19	$24.19^{\dagger\circ}$	$54.88^{\dagger\circ}$	$26.55^{\dagger\circ}$
31	RRF (SMOG & Orig.) Top 15	Based on ECNU	$\mathbf{47.23}^{\dagger\circ}$	12.41	$\mathbf{46.24}^{\dagger\circ}$	8.95	59.57	13.88	$\mathbf{43.96}^{\dagger\circ}$	11.80	0.08	$49.21^{\dagger\circ}$	61.45	45.42
32		Based on KISTI	34.84°	8.34	34.08°	7.48	$67.87^{\dagger\circ}$	11.84	36.86°	8.18	0.03	35.32°	$68.84^{\dagger\circ}$	37.37°
33		Based on BM25	$27.16^{\dagger\circ}$	10.79	$26.55^{\dagger\circ}$	7.32	66.44	14.65	$32.24^{\dagger\circ}$	10.79	0.05	$28.21^{\dagger\circ}$	69.38°	$33.67^{\dagger\circ}$
34	RRF (SMOG & Orig.) Top 20	Based on ECNU	$45.91^{\dagger\circ}$	14.14	$44.94^{\dagger\circ}$	9.11	59.46	15.28	$43.50^{\dagger\circ}$	13.55	0.09	$48.67^{\dagger\circ}$	$62.46^{\dagger\circ}$	45.93
35		Based on KISTI	34.17°	8.54	33.42°	7.48	$68.23^{\dagger\circ}$	12.09	36.23°	8.38	0.03	34.79°	69.30°	36.89°
36		Based on BM25	$25.80^{\dagger\circ}$	12.33	$25.22^{\dagger\circ}$	7.48	64.98	16.18	$30.69^{\dagger\circ}$	12.33	0.08	$27.24^{\dagger\circ}$	69.14°	$32.68^{\dagger\circ}$
37	RRF (SMOG & Orig.) Top 50	Based on ECNU	$39.06^{\dagger\circ}$	**20.77**	$38.20^{\dagger\circ}$	**9.27**	57.68	23.16	$39.55^{\dagger\circ}$	20.28	**0.18**	$45.59^{\dagger\circ}$	$65.82^{\dagger\circ}$	**46.13**
38		Based on KISTI	$29.51^{\dagger\circ}$	10.35	$28.83^{\dagger\circ}$	7.48	$\mathbf{68.60}^{\dagger\circ}$	14.71	33.37	10.35	0.08	$31.48^{\dagger\circ}$	$\mathbf{71.70}^{\dagger\circ}$	35.54
39		Based on BM25	$18.70^{\dagger\circ}$	17.12	$18.29^{\dagger\circ}$	7.16	$57.50^{\dagger}$	24.59	$22.73^{\dagger\circ}$	17.12	0.15	$22.95^{\dagger\circ}$	68.61	$28.21^{\dagger\circ}$
40	XGB LeToR	Combo 1 on BM25	$24.86^{\dagger\circ}$	17.39	$24.32^{\dagger\circ}$	7.81	$\mathbf{55.60}^{\dagger\circ}$	24.11	$28.89^{\dagger\circ}$	17.39	0.22	29.67°	66.41	34.76°
41		Combo 2 on BM25	$\mathbf{30.72}^{\circ}$	21.25	$\mathbf{30.08}^{\circ}$	8.46	$48.87^{\dagger\circ}$	28.82	$\mathbf{31.76}^{\circ}$	18.99	0.26	37.09°	$61.89^{\dagger}$	$\mathbf{39.17}^{\circ}$
42		Combo 3 on BM25	28.92°	**24.35**	28.32°	8.46	$49.02^{\dagger\circ}$	32.11	$30.14^{\dagger\circ}$	23.83	0.31	$\mathbf{37.32}^{\circ}$	$63.86^{\dagger}$	39.84°
43		Combo 4 on BM25	$25.65^{\dagger\circ}$	25.72	25.09^{*}	8.30	$49.00^{\dagger\circ}$	**33.39**	$27.45^{\dagger\circ}$	**24.40**	**0.33**	35.82°	66.14	38.21°
44		Combo 5 on BM25	30.21°	20.79	29.59°	**8.62**	$48.47^{\dagger\circ}$	27.88	$30.95^{\dagger\circ}$	19.99	0.25	37.11°	$61.25^{\dagger\circ}$	39.15°

Figure 9. Results obtained by integrating understandability estimations within retrieval methods on **CLEF 2015**. Baseline runs are reported at table indices 1-3 (the index column is labelled Index). Re-ranking experiments are reported at indices 4-21. Fusion experiments are reported at indices 22-30. Learning to rank experiments are reported at indices 31-35. All measures were calculated up to rank n = 10. The highest result of each set of experiments is reported in bold face.