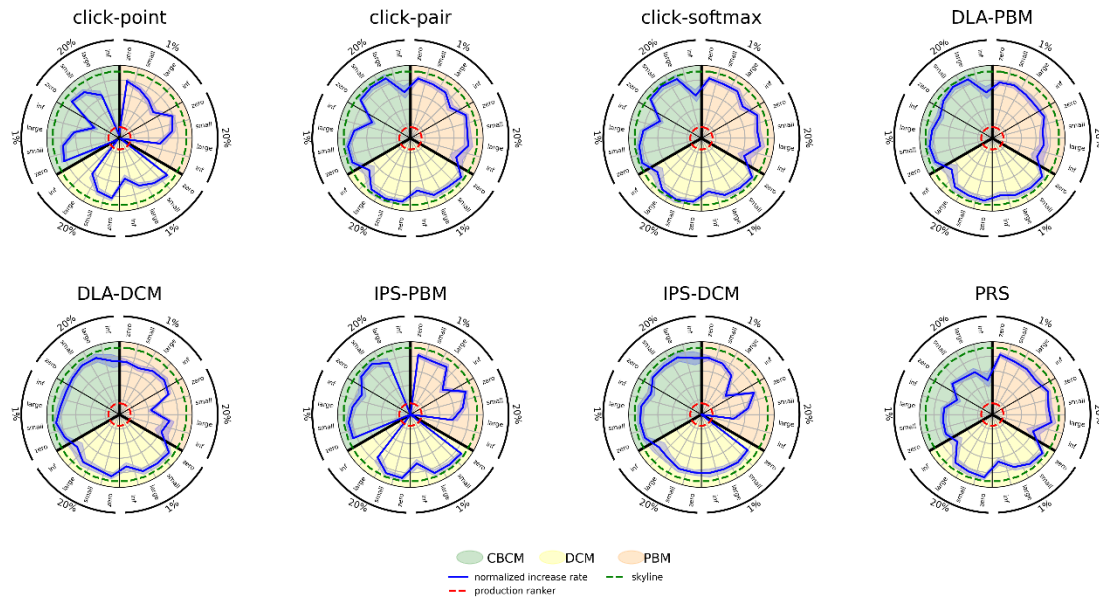


The spider chart on MSLR:

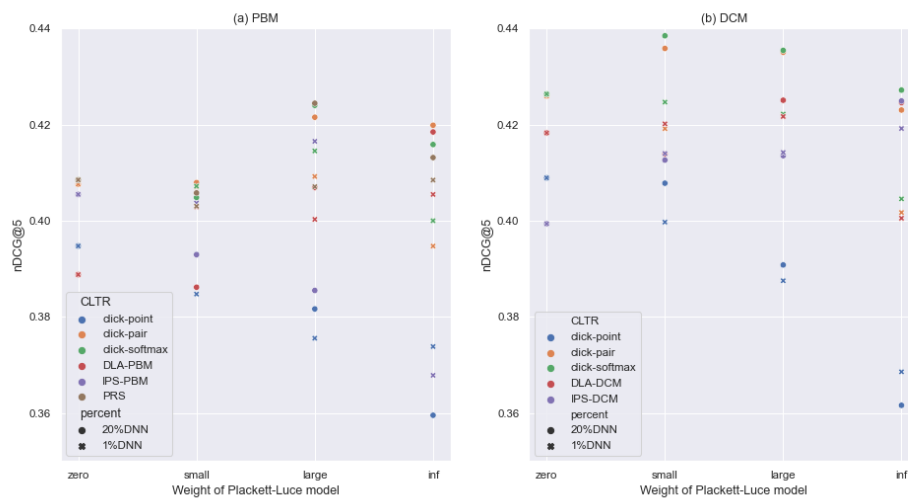


The ANOVA results on MSLR:

Table 6: ANOVA results on MSLR-WEB10K.

Source	SS	DF	MS	F	<i>p</i> -value	$\hat{\omega}^2$
Percentage of data	0.00005	1	0.00005	0.0950	7.58e-1	-0.0048
Plackett-luce weight	0.0147	3	0.0049	8.7892	1.82e-5	0.1090
User simulation model	0.0174	2	0.0087	15.6405	5.53e-7	0.1329
CLTR model	0.0218	7	0.0031	5.5865	7.78e-6	0.1439
Residual	0.0991	178	0.0006			

The effect of production ranker on CLTR models on MSLR:



The effect of user simulation model on CLTR models on MSLR:

Table 6: nDCG@5 performance of CLTR models with 1% DNN as the production ranker, PBM as the user simulation model, and different user simulation models. Same conventions as in Table 1.

CLTR model	Yahoo			MSLR		
	PBM	DCM	CBCM	PBM	DCM	CBCM
production ranker		0.6177			0.2867	
skyline		0.7152			0.4417	
click-point	0.6716	0.6632	0.6598	0.3739	0.3686	0.3314
click-pair	0.6732	0.6651	0.6651	0.3947 ⁺	0.4017 ⁺	0.3708 ⁺
click-softmax	0.6736	0.6665	0.6662	0.4000 ⁺	0.4046 ⁺	0.3797 ⁺
DLA-PBM	0.6804 ⁺	0.6800 ⁺	0.6713 ⁺	0.4055 ⁺	0.4156 ⁺	0.4126 ⁺
DLA-DCM	0.6837⁺	0.6758 ⁺	0.6717 ⁺	0.4120⁺	0.4005 ⁺	0.4166⁺
IPS-PBM	0.6101 ⁻	0.6152 ⁻	0.6723 ⁺	0.3679	0.3924 ⁺	0.4009 ⁺
IPS-DCM	0.6333 ⁻	0.6948⁺	0.6872⁺	0.3408 ⁻	0.4192⁺	0.4069 ⁺
PRS	0.6747	0.6615	0.6569	0.4085 ⁺	0.3872 ⁺	0.3612 ⁺