

Scoring Regime	DCG	AP	RBP $p = 0.8$
TF-IDF	7.8485	0.9223	0.9454
BM25 $k_1=1.2, b=0.75$	7.9112	0.9429	0.9598
BM25 $k_1=2, b=0.5$	7.8800	0.9335	0.9530
BM25 $k_1=1.6, b=0.2$	7.8734	0.9293	0.9517
BM25 $k_1=1.6, b=0.75$	7.9030	0.9396	0.9578
BM25 $k_1=1.6, b=0.9$	7.9043	0.9381	0.9583
BM25 $k_1=1.1, b=0.75$	7.9122	0.9434	0.9600
BM25 $k_1=0.2, b=0.75$	7.9140	0.9444	0.9619
BM25 $k_1=0.2, b=0.9$	7.9129	0.9428	0.9610

Setelah melakukan beberapa eksperimen, menurut data, skema terbaik adalah BM25 dengan $k_1 = 0.2$ dan $b = 0.75$ dan yang paling buruk adalah TF-IDF. Ini mungkin disebabkan pada BM25 dapat menangani masalah term frequency yang terlalu tinggi dan masalah banyak dokumen yang berbeda sehingga perbedaan term frequency dan banyak dokumen yang terlalu besar dapat dilakukan lebih adil dengan mengatur parameter k_1 dan b nya. Pada percobaan **BM25 $k_1=1.6, b=0.2$, BM25 $k_1=1.6, b=0.75$, dan BM25 $k_1=1.6, b=0.9$** dapat dilihat bahwa dengan kenaikan b bisa membuat hasil lebih baik. Pada percobaan **BM25 $k_1=0.2, b=0.75$, BM25 $k_1=1.1, b=0.75$, BM25 $k_1=1.2, b=0.75$ dan BM25 $k_1=1.6, b=0.75$** juga dapat dilihat bahwa nilai k_1 yang menurun bisa membuat hasil lebih baik. Namun, ketika **BM25 $k_1=0.2, b=0.9$** hasilnya menurun. Kesimpulannya adalah pada umumnya BM25 akan lebih baik dibandingkan unnormalized TF-IDF dan perlunya untuk melakukan eksperimen beberapa variasi nilai k_1 dan b agar mendapatkan hasil yang terbaik.