

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan terkait evaluasi performa arsitektur *dense retrieval* dan *reranker* dalam penilaian tata kelola perusahaan publik di Indonesia, berikut beberapa kesimpulan yang dapat diambil:

1. Hasil pengujian menunjukkan bahwa *dense retrieval* memberi performa yang lebih baik dalam penilaian tata kelola perusahaan publik di Indonesia dibandingkan dengan *baseline sparse retrieval*. *Dense retrieval* meraih skor rata-rata $\text{Hit}@5$ 0.239 lebih tinggi, $\text{Hit}@10$ 0.216 lebih tinggi, dan $\text{Hit}@20$ 0.155 lebih tinggi dibanding *baseline*. Sementara itu pada metrik $\text{MRR}@k$, *dense retrieval* memiliki skor rata-rata $\text{MRR}@5$ 0.226 lebih tinggi, $\text{MRR}@10$ 0.224 lebih tinggi, dan $\text{MRR}@20$ 0.220 lebih tinggi dibanding *baseline*.
2. Hasil pengujian juga menunjukkan bahwa *dense retrieval* dan *reranker* memberi evaluasi akhir berupa skor $\text{Hit}@5$ sebesar 0.804, $\text{Hit}@10$ sebesar 0.877, dan $\text{Hit}@20$ sebesar 0.933. *Dense retrieval* dan *reranker* juga memberi evaluasi akhir berupa skor $\text{MRR}@5$ sebesar 0.580, $\text{MRR}@10$ sebesar 0.590, dan $\text{MRR}@20$ sebesar 0.594. *Reranker* cenderung sedikit kurang optimal dalam meningkatkan rata-rata kualitas *ranking* dari *dense retrieval* yang ditunjukkan oleh penurunan tipis pada skor $\text{MRR}@k$. Untuk mencapai performa *reranking* yang general dan optimal pada domain tata kelola perusahaan, model *reranker* perlu dilatih pada domain yang serupa.

5.2 Saran

Berdasarkan hasil penelitian ini, peneliti memberikan beberapa saran yang dapat dijadikan acuan untuk penelitian selanjutnya, antara lain:

1. Penelitian selanjutnya dapat menggunakan atau menciptakan beberapa variasi data dengan jumlah yang beragam, sehingga efisiensi komputasi tiap arsitektur bisa dievaluasi lebih lanjut.
2. Penelitian selanjutnya dapat menggunakan atau menciptakan data dengan *corpus* dan label yang lebih kuat untuk memperkaya hasil evaluasi performa arsitektur dengan metrik lainnya.
3. Penelitian selanjutnya dapat mencoba model lain atau menggunakan beberapa model sekaligus pada tahap *retriever* dan *reranker* untuk memperkuat hasil evaluasi performa arsitektur *dense retrieval* dan *reranker*.
4. Penelitian selanjutnya dapat melakukan *fine-tuning* pada model untuk memaksa model *retriever* dan *reranker* mencapai performa terbaiknya.
5. Penelitian ini dapat dikembangkan lebih lanjut menjadi sebuah sistem *end-to-end* penilaian tata kelola perusahaan, sehingga dapat memberi manfaat langsung kepada auditor profesional dengan meningkatkan efektifitas *retrieving* dalam proses penilaian tata kelola perusahaan.