

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang dilandasi oleh penelitian dan analisis yang telah dilakukan, serta dilengkapi dengan saran untuk pengembangan ke depan.

5.1 Kesimpulan

Berdasarkan hasil dan analisis pengujian metode BM25 dan metode *neural network* dengan BERT *preprocessing*, berikut kesimpulan dari perbandingan kedua metode:

1. Nilai b dan k yang optimal dalam metode BM25 adalah 0.4 dan 0.4. Terdapat tarik ulur antara nilai k dan b sehingga nilai keduanya perlu diseimbangkan. Hasil terburuk didapat pada saat b bernilai 0 sehingga panjang dokumen perlu diperhatikan. Nilai b yang terlalu besar tidak memberikan hasil yang optimal walau nilai k telah diseimbangkan. Sehingga, nilai b yang relatif kecil lebih baik dan nilai k perlu dicari untuk menyeimbangkan nilai b .
2. Nilai indikator yang optimal dalam metode *neural network* dengan BERT *preprocessing* adalah 50 untuk subsegment size, 0.001 untuk *learning rate*, 256 untuk *layer size*, dan metode *mean* untuk metode penggabungan nilai relevansi subsegment.

Ukuran *subsegment* dokumen yang kecil memperbanyak jumlah *data point* yang digunakan untuk *training*. Sehingga, ukuran *subsegment* dokumen yang lebih kecil memberikan hasil nDCG@20 yang lebih baik.

Nilai *learning rate* yang lebih kecil perlu diseimbangkan dengan jumlah *data training* yang banyak agar memberikan hasil yang baik. Maka, nilai *learning rate* yang terlalu kecil menurunkan nilai nDCG@20.

Ukuran layer dalam *neural network* tidak mempengaruhi nilai nDCG@20 secara signifikan. Ukuran *layer* yang kecil menggunakan sumber daya yang lebih sedikit. Maka, ukuran *layer* kecil lebih disarankan.

Metode penggabungan nilai relevansi terbaik adalah metode *mean* karena metode tersebut memperhitungkan semua *subsegment* serta tidak menurunkan nilai relevansi apabila ada *subsegment* yang bernilai sangat rendah.

3. Metode *neural network* menghasilkan nilai nDCG@20 (0.5520) lebih baik dari metode BM25 (0.5045). Hal ini memperlihatkan bahwa metode *neural*

network dengan BERT preprocessing memberikan hasil yang lebih baik dari metode BM25. Namun, dalam sistem yang menggunakan 1 processor dan tidak menggunakan GPU, waktu pencarian sangat lambat. Sehingga, penggunaan metode *neural network* dengan BERT *preprocessing* membutuhkan perangkat keras yang lebih baik.

5.2 Saran

Saran untuk pengembangan penelitian ini adalah sebagai berikut.

1. Menggunakan arsitektur LSTM atau CNN untuk mengolah hasil *preprocessing* BERT yang berbentuk *sequence*.
2. Melakukan *fine-tuning* terhadap model BERT agar hasil *preprocessing* yang dihasilkan memberikan nilai nDCG yang lebih baik
3. Menggunakan *dataset* yang lebih besar sehingga model *neural network* mencapai hasil yang optimal.
4. Memisahkan *preprocessing* dokumen dan *query* menggunakan BERT agar ukuran *data* yang digunakan menjadi lebih kecil.
5. Menggunakan model BERT *Large* agar hasil yang didapat lebih baik.
6. Menggunakan sistem terdistribusi dan paralel dengan *cloud computing*/GPU agar kecepatan pencarian dapat ditingkatkan
7. Menggunakan *dataset* dengan dokumen yang lebih umum sehingga model yang dihasilkan dapat digunakan pada berbagai STBI lain.