BAB I
PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang akan digunakan dalam dokumen skripsi yang mengenai klasifikasi dokumen riwayat putusan peradilan Mahkamah Agung menggunakan *Fine-tuning Hierarchical* BERT + *Attention*.

1.1 Latar Belakang

Mahkamah Agung (MA) adalah salah satu pelaku kekuasaan kehakiman sebagaimana dimaksud dalam Undang-Undang Dasar Negara Republik Indonesia Tahun 1945. Berdasarkan Undang-Undang Republik Indonesia Nomor 14 Tahun 1985 Pasal 2 tentang Mahkamah Agung menyatakan bahwa, Mahkamah Agung adalah pengadilan negara tertinggi dari semua lingkungan peradilan, yang dalam melaksanakan tugasnya terlepas dari pengaruh pemerintah dan pengaruh-pengaruh lain. Peran penting MA dalam sistem peradilan Indonesia, termasuk mengawasi jalannya peradilan di semua lingkungan peradilan (umum, agama, militer, dan tata usaha negara) serta memastikan penerapan hukum yang konsisten dan adil.

Menurut Nurwandi dkk. (2024), jumlah perkara yang diterima MA pada tahun 2020 meningkat sebesar 6.07% dibandingkan tahun 2019 yang menerima total 19.369 perkara. Jurnal tahunan MA mencatat total perkara pada tahun 2022 mencapai 28.284 perkara yang didistribusikan di antara 47 hakim teratas. Pada data tersebut, perbandingan antara pekerjaan seorang hakim MA dengan perkerjaan adalah satu banding 602. Beban kerja hakim MA yang tinggi ini dapat membuat hakim rentan terkena stress dan tentunya berpengaruh dalam keputusan yang diberikan.

Sejalan dengan peningkatan jumlah perkara, pertumbuhan jumlah dokumen putusan pengadilan pun ikut meningkat. Menurut Firdaus dkk. (2018), pertumbuhan jumlah dokumen putusan pengadilan berbanding lurus dengan jumlah perkara hukum yang diproses di pengadilan. Data pertumbuhan jumlah putusan pengadilan menunjukkan bahwa selama masih ada proses hukum yang terjadi, maka jumlah dokumen putusan akan terus bertambah jumlahnya. Dokumen putusan pengadilan dibuat oleh Pengadilan Negeri setelah suatu persidangan selesai. Umumnya dokumen ini memiliki jumlah halaman berkisar antara 10

halaman hingga tidak terbatas sesuai dengan kondisi persidangan. Sebagai ilustrasi, beberapa contoh dokumen putusan peradilan dengan jenis perkara pidana MA pada tahun 2024 dapat dilihat pada Tabel 1.1.

Tabel 1.1 Ilustrasi Contoh Dokumen Putusan Peradilan Mahkamah Agung Pada Tahun 2024

Nomor Putusan	Jenis Perkara	Tanggal Putusan	Jumlah Halaman
Putusan PN SEMARANG Nomor 518/Pid.B/2024/PN Smg	Pidana Umum – Pencurian- Tingkat Pertama	15 Oktober 2024	19
Putusan PN BOGOR Nomor 107/Pid.B/2024/PN Bgr	Pidana Umum – Pembunuhan- Tingkat Pertama	30 Juli 2024	29
Putusan PN JAKARTA UTARA Nomor 948/Pid.Sus/2024/PN Jkt.Utr	Pidana Khusus – Narkotika dan Psikotropika- Tingkat Pertama	24 Desember 2024	21
Putusan PN YOGYAKARTA Nomor 10/Pid.Sus-Anak/2024/PN Yyk	Pidana Khusus – Peradilan Anak ABH – Tingkat Pertama	14 November 2024	25

Sumber : https://putusan3.mahkamahagung.go.id/ (diakses pada tanggal 4 Maret 2025).

Jumlah dokumen yang begitu banyak ini membuat sulit bagi para profesional hukum terutama hakim untuk membaca dan menafsirkan setiap dokumen yang berkaitan dengan suatu kasus guna membandingkan tingkat keparahannya dengan kasus-kasus sebelumnya. Penelitian yang dilakukan oleh Nuranti dkk. (2022) membuat sistem prediksi kategori berat hukuman (klasifikasi) dan lama hukuman hukum sebagai referensi pendukung untuk memutuskan hukuman yang tepat dengan menggunakan dokumen hukum putusan peradilan MA pada tahun 2016 - 2019. Penelitian ini menggunakan pendekatan *multi-level learning* dengan arsitektur CNN + *Attention.* Model ini dibandingkan arsitektur *deep learning* lainnya seperti CNN, LSTM, BiLSTM, LSTM + *Attention,* dan BiLSTM + *Attention.* Hasil penelitian menunjukkan model CNN + *Attention* memberikan hasil yang lebih baik dibandingkan model lain dalam memprediksi kategori hukuman dengan akurasi 77,32%.

Penelitian dengan *dataset* yang sama seperti Nuranti dkk. (2022), yang dilakukan oleh Latisha dkk. (2024) berhasil mengembangkan sistem prediksi putusan peradilan dengan menggunakan *Hierarchical* BERT + *Mean Pooling.* Model ini mencapai akurasi 79.8888%, yang menunjukkan keunggulan dibandingkan model yang digunakan oleh Nuranti dkk. (2022) dalam meningkatkan performa model untuk menangani dokumen hukum yang panjang dan kompleks.

Keberhasilan ini sejalan dengan tren perkembangan model hierarkis untuk pemrosesan dokumen panjang, yang telah menunjukkan peningkatan signifikan dalam berbagai penelitian sebelumnya. Penelitian tentang model hierarkis untuk pemrosesan dokumen panjang menunjukkan perkembangan signifikan. Yang dkk. (2016) memperkenalkan *Hierarchical Attention Networks* (HAN) yang mencapai akurasi tertinggi 75,8% pada *dataset* Yahoo Answer. Gao dkk. (2018) mengembangkan *Hierarchical Convolutional Attention Network* (HCAN) yang mencapai akurasi hingga 89,89% pada *dataset* Amazon. Pappagari dkk. (2019) memperkenalkan dua pendekatan, *Recurrence over* BERT (RoBERT) dan *Transformer over* BERT (ToBERT), dengan ToBERT mencapai akurasi hingga 95,48% pada *dataset* Fisher. Lu dkk. (2021) mengusulkan *Hierarchical* BERT Model (HBM) yang mencatat akurasi tertinggi 99,25% pada *dataset* Guardian 2013, sementara Kong dkk. (2022) memperkenalkan *Hierarchical Adaptive Fine-tuning* BERT (HAdaBERT) yang berhasil memperoleh akurasi 90,9% pada *dataset* Reuters.

Keberhasilan pendekatan hierarkis ini menunjukkan bahwa model seperti HBM dapat mengatasi keterbatasan dalam menangani dokumen panjang pada *Large Language Models* (LLMs) berbasis *transformer*, seperti *Bidirectional Encoder Representations from Transformers* (BERT) (Devlin dkk., 2019). Keterbatasan ini disebabkan oleh kapasitas pemrosesan token yang terbatas dalam sekali *input*, di mana sebagian besar model hanya dapat menangani beberapa ratus token dalam satu waktu (Prasad dkk., 2024). Model hierarkis yang mengolah informasi secara bertahap dari tingkat kata ke dokumen dapat meningkatkan pemahaman terhadap teks panjang dan kompleks.

Penelitian ini mengadaptasi pendekatan hierarkis berbasis konsep "*divide, learn, and combine*" (Prasad dkk., 2024) untuk mengatasi keterbatasan model LLMs, yang memungkinkan pemrosesan dokumen panjang secara lebih efektif melalui beberapa tahapan utama:

1. Segmentasi Dokumen

   Dokumen dibagi menjadi beberapa bagian sebelum dilakukan pemrosesan data. Pada penelitian ini dilakukan dengan mengadaptasi mekanisme segmentasi *input* diadaptasi dari penelitian Kong dkk. (2022), yaitu dengan membagi *input* menjadi segmen (*chunk* atau kontainer) berkapasitas tetap yang dilengkapi dengan *overlapping word*. Setiap segmen memiliki batas maksimum panjang token tetap untuk mempertahankan representasi kata dan kalimat dalam dokumen hukum.

2. Pembelajaran Bertahap dengan *Attention*

   Fitur dari setiap komponen dipelajari secara bertahap dengan mengadaptasi arsitektur *Hierarchical* BERT + *Mean Pooling* yang dilakukan oleh Latisha dkk. (2024). Selain itu, *attention* ditambahkan pada tingkat kata serta kalimat mengikuti arsitektur HAN (Yang dkk., 2016). Studi yang dilakukan oleh Yang dkk. (2016) menunjukkan bahwa HAN secara signifikan meningkatkan akurasi dalam tugas klasifikasi teks dibandingkan model tanpa *attention*, karena dapat menangkap informasi penting dalam dokumen panjang dengan lebih efektif. Selain itu, penelitian terbaru oleh (Chalkidis dkk., 2019) menunjukkan bahwa penggunaan *Hierarchical Transformer* dengan *attention* mampu meningkatkan kinerja dalam tugas klasifikasi hukum, terutama dalam menangani dokumen dengan panjang lebih dari 512 token. Model ini terbukti lebih unggul dibandingkan pendekatan yang hanya mengandalkan *mean pooling* atau *max pooling*, karena dapat memberikan representasi yang lebih kaya terhadap hubungan antarbagian dalam dokumen hukum.
3. Penggabungan Fitur Secara Hierarkis

   Fitur-fitur tersebut dikombinasikan secara hierarkis dari bawah ke atas untuk mendapatkan representasi dokumen secara keseluruhan. Kombinasi antar-*chunk* dilakukan dengan mengadaptasi penelitian Prasad dkk. (2024), yaitu pada antar-*chunk* ditambahkan *positional embeddings* seperti yang dilakukan pada model *transformers* (Vaswani dkk., 2017). *Positional embeddings* ini membantu model memahami urutan informasi dalam dokumen.

Pada penelitian ini, strategi *fine-tuning* dilakukan dengan dua mekanisme yaitu *grid search* untuk menentukan kombinasi *hyperparameter* terbaik dan *freeze layers* untuk mencegah *overfitting* pada model *Hierarchical* BERT + *Attention*. *Grid search* adalah salah satu metode untuk mencari kombinasi optimal dari *hyperparameter* dalam suatu model *machine learning* di mana pada suatu *grid* semua kombinasi *hyperparameter* diuji dengan menggunakan *cross-validation* (Bergstra dkk., 2012). Pendekatan ini memungkinkan pemilihan konfigurasi optimal, yaitu ukuran *learning rate* dan *epoch*, sehingga model dapat bekerja lebih efektif.

Penelitian Sánchez dkk. (2024) menunjukkan bahwa strategi *partial fine-tuning* dengan membekukan lapisan tertentu pada BERT terbukti efektif untuk tugas pemrosesan bahasa alami (NLP) multibahasa, termasuk bahasa dengan struktur morfologi kompleks.

Pada studi tersebut, pembekuan lapisan dilakukan secara *bottom-up* (dari lapisan bawah ke atas), dengan pertimbangan:

1. Lapisan bawah BERT (misalnya, lapisan 1–4) menyimpan informasi linguistik dasar seperti morfologi dan sintaksis yang bersifat universal (Clark dkk., 2019).
2. Lapisan atas BERT lebih spesifik untuk menangkap konteks semantik dan pola tugas tertentu (Jawahar dkk., 2019).

Penelitian oleh Kong dkk. (2022) menunjukkan bahwa strategi *adaptive fine-tuning* dengan *freeze layer* adaptif dapat meningkatkan efisiensi dan performa model. Eksperimen HAdaBERT menemukan bahwa lima lapisan terakhir selalu dipilih untuk fine-tuning karena mampu menangkap informasi semantik dan sintaksis yang lebih kaya, sementara lapisan bawah hanya perlu dioptimalkan secara selektif. Pendekatan ini mengurangi jumlah parameter yang dilatih tanpa menurunkan performa, dengan hanya 53,5% hingga 65,6% dari total 108 juta parameter BERT 12-*layer* yang dilatih pada *dataset* seperti IMDB, Yelp-2013, AAPD, dan Reuters. Efisiensi ini juga membantu mencegah overfitting dan memastikan model tetap optimal (Kong dkk., 2022).

Berdasarkan temuan tersebut, penelitian ini mengadopsi strategi serupa untuk *dataset* berbahasa Indonesia. Model diharapkan dapat mempertahankan pengetahuan dasar tentang struktur bahasa Indonesia dan mencegah *overfitting* akibat data pelatihan terbatas pada domain hukum dengan membekukan lapisan bawah pada BERT. Mekanisme pengembangan *fine-tuning* pada model *Hierarchical* BERT dengan *Attention* ini diharapkan dapat meningkatkan akurasi dan mencegah *overfitting* klasifikasi kategori hukuman, sebagaimana yang telah dikaji oleh Nuranti dkk. (2022). Selain itu, model ini diharapkan juga mengatasi keterbatasan model konvensional yang kurang optimal dalam menangkap informasi kontekstual pada dokumen hukum berbahasa Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, dapat dirumuskan permasalahan umum yang dibahas pada penelitian skripsi ini, yaitu bagaimana hasil kinerja implementasi model *Fine-tuning Hierarchical* BERT + *Attention* dalam mengklasifikasikan dokumen riwayat putusan peradilan Mahkamah Agung?

Sementara itu, permasalahan khusus pada penelitian ini antara lain:

1. Bagaimana implementasi kinerja model BERT dalam mengklasifikasikan dokumen putusan peradilan Mahkamah Agung, serta bagaimana kombinasi *hyperparameter* terbaik yang diperoleh melalui metode *grid search*?
2. Bagaimana implementasi kinerja dan mekanisme terbaik dalam *fine-tuning* pada model *Hierarchical* BERT + *Attention* untuk mengklasifikasikan dokumen putusan peradilan Mahkamah Agung?
3. Bagaimana hasil perbandingan kinerja antara model BERT *Fine-tuning* dan model *Fine-tuning Hierarchical* BERT + *Attention* dalam mengklasifikasikan dokumen putusan peradilan Mahkamah Agung berdasarkan metrik evaluasi yang relevan?

1.3 Tujuan dan Manfaat

Tujuan umum dari penelitian dalam skripsi ini adalah dihasilkan model dengan akurasi terbaik untuk mengklasifikasikan dokumen riwayat putusan peradilan Mahkamah Agung menggunakan *Fine-tuning Hierarchical* BERT + *Attention*.

Sedangkan tujuan khusus dari skripsi ini meliputi:

1. Mengimplementasikan model BERT *Fine-tuning* untuk mengklasifikasikan dokumen putusan peradilan Mahkamah Agung serta menentukan kombinasi *hyperparameter* terbaik menggunakan metode *grid search*.
2. Mengembangkan dan mengevaluasi mekanisme *fine-tuning* pada model *Hierarchical* BERT + *Attention* untuk meningkatkan akurasi klasifikasi dokumen putusan peradilan Mahkamah Agung.
3. Membandingkan kinerja model BERT *Fine-tuning* dan model *Fine-tuning Hierarchical* BERT + *Attention* berdasarkan metrik evaluasi yang relevan guna menentukan model yang paling optimal dalam klasifikasi dokumen hukum.

Sementara itu, manfaat yang dari penelitian skripsi ini yaitu memberikan solusi teknologi untuk mendukung sistem keputusan hukum berbasis data yang dapat membantu praktisi hukum, akademisi, dan masyarakat dalam menganalisis dokumen hukum secara lebih efektif.

1.4 Ruang Lingkup

Ruang lingkup dalam penelitian ini dilakukan agar penelitian lebih terarah dan mencapai sasaran yang diharapkan.

Ruang lingkup dalam skripsi ini adalah sebagai berikut:

1. Data pada penelitian ini merupakan kumpulan putusan peradilan Indonesia yang mencakup kasus perkara pidana umum dan pidana khusus. *Dataset* ini terdiri atas 22.630 dokumen yang telah dianotasi berdasarkan hasil identifikasi bagian dokumennya dan hanya menggunakan sebanyak empat bagian tersebut yang merupakan riwayat tuntutan, fakta, fakta hukum, dan pertimbangan hukum. Pemilihan bagian dokumen ini dilakukan berdasarkan penelitian yang dilakukan oleh Nuranti dkk. (2022)
2. Labelisasi kategori ditentukan ke dalam empat buah kelas untuk setiap kategori hukuman tindak pidana pada putusan peradilan Mahkamah Agung yaitu *mild, moderate, heavy,* dan *very heavy*.

1.5 Sistematika Penulisan

Gambaran mengenai pembahasan penyusunan laporan skripsi ini secara urut dan jelas, maka dibentuk sistematika sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian, dan sistematika penulisan dari penelitian dengan judul Klasifikasi Dokumen Riwayat Putusan Peradilan Mahkamah Agung Menggunakan *Fine-tuning Hierarchical* BERT + *Attention*.

BAB II LANDASAN TEORI

Bab ini memuat tentang tinjauan pustaka dan dasar teori yang digunakan dalam skripsi berjudul Klasifikasi Dokumen Riwayat Putusan Peradilan Mahkamah Agung Menggunakan *Fine-tuning Hierarchical* BERT + *Attention*. Bab ini terdiri atas beberapa subbab antara lain *state of the arts,* pengolahan bahasa alami, pra-pemrosesan data, ekstraksi fitur, *transformers,* mekanisme *attention, Hierarchical* BERT + *Attention,* fungsi aktivasi, *forward* dan *backward pass, hyperparamter fine-tuning, optimizer* AdamW, PyTorch, *hyperparameter,* metrik evaluasi, serta *tools* dan *library*.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan mengenai tahap-tahap penelitian yang dilakukan dalam skripsi Klasifikasi Dokumen Riwayat Putusan Peradilan Mahkamah Agung Menggunakan *Fine-tuning Hierarchical* BERT + *Attention*. Gambaran umum

dari tahapan penelitian ini adalah pengumpulan data, pra-pemrosesan data, pemahaman data, pembagian data, BERT, *Hierarchical BERT + Attention* , evaluasi, lingkungan pengujian, dan skenario eksperimen.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan mengenai hasil dan analisis yang dilakukan sesuai metodologi yang digunakan dalam skripsi Klasifikasi Dokumen Riwayat Putusan Peradilan Mahkamah Agung Menggunakan *Fine-tuning Hierarchical* BERT + *Attention*. Bab ini membahas tentang data penelitian, pengolahan data penelitian, hasil dan analisis model, dan hasil dan analisis skenario eksperimen yang dilakukan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari bab-bab yang telah dijelaskan sebelumnya dan saran untuk penelitian lebih lanjut