

BAB II TINJAUAN TEORI

2.1 Tinjauan Teori

2.1.1 Teori Transformer

Model Transformer, pertama kali diperkenalkan oleh Vaswani et al. (2017) dalam makalah "Attention is All You Need," merupakan dasar arsitektur model bahasa seperti GPT-4o Mini yang digunakan dalam penelitian ini. Transformer memanfaatkan mekanisme *self-attention* untuk memproses data secara paralel, memungkinkan model memahami hubungan antar elemen dalam data tanpa bergantung pada urutan seperti model berbasis RNN atau CNN.

Untuk mendukung pemrosesan data pajak yang kompleks, model Transformer dalam penelitian ini dilengkapi dengan memori eksternal, seperti dijelaskan oleh Weston et al. (2015) dalam Memory-Augmented Neural Networks (MANN). Memori eksternal memungkinkan model menyimpan informasi historis, seperti aturan perpajakan atau data karyawan, untuk digunakan kembali selama inferensi. Hal ini relevan untuk memastikan konsistensi dalam menghasilkan laporan SPT PPh 21.

2.1.2 Peraturan Pajak

Pelaporan PPh 21 diatur oleh Undang-Undang Nomor 7 Tahun 1983 tentang Pajak Penghasilan, yang telah diubah terakhir melalui Undang-Undang Nomor 36 Tahun 2008, serta Peraturan Direktur Jenderal Pajak Nomor PER-2/PJ/2024. Peraturan ini mengatur perhitungan PPh 21 untuk karyawan tetap dan tidak tetap, termasuk komponen seperti penghasilan bruto, Penghasilan Tidak Kena Pajak (PTKP), dan tarif pajak progresif. Dalam penelitian ini, *ground truth* yang dihasilkan oleh *key_1.py* dirancang sesuai PER-2/PJ/2024 untuk memastikan akurasi perhitungan sebagai pembanding output GPT-4o Mini.

Peraturan Menteri Keuangan (PMK) Nomor 181/PMK.03/2007 dan PMK Nomor 152/PMK.03/2009 juga mengatur teknis pengisian SPT, yang menekankan pentingnya akurasi data seperti NPWP, penghasilan, dan PPh terutang. Temuan di Bab 4.4 menunjukkan bahwa sistem AI mampu memproses data identitas dengan akurasi tinggi, tetapi kesalahan pada perhitungan numerik menunjukkan perlunya prompt yang lebih spesifik untuk mematuhi regulasi ini, terutama untuk karyawan tidak tetap dengan penghasilan tidak teratur.

2.1.3 Pendekatan Data Sintetis

Data sintetis digunakan dalam penelitian ini untuk mensimulasikan data perusahaan dan karyawan, menghindari pelanggaran privasi data riil. Menurut Goodfellow et al. (2016), data sintetis dapat dirancang untuk mencerminkan karakteristik data asli, seperti distribusi penghasilan atau status PTKP, sambil memungkinkan kontrol variabel untuk pengujian. Dalam penelitian ini, *skripsi_1.py* menghasilkan 400 file data sintetis dengan 15–80 karyawan per perusahaan, mencakup skenario karyawan tetap dan tidak tetap. Data sintetis ini menjadi input untuk GPT-4o Mini dan dibandingkan dengan *ground truth* dari *key_1.py* untuk mengukur akurasi.

2.1.4 Efisiensi dan Akurasi

Efisiensi dan akurasi adalah indikator utama dalam mengevaluasi sistem AI untuk pelaporan PPh 21. Efisiensi diukur melalui waktu pemrosesan, akurasi diukur dengan membandingkan output AI terhadap *ground truth* menggunakan metrik tingkat kecocokan, seperti dijelaskan oleh Russell dan Norvig (2021) dalam teori evaluasi model AI. Metrik ini menghitung persentase entri identik antara output AI dan *ground truth*, dengan toleransi nol untuk field kritis seperti penghasilan bruto dan PPh dipotong.

Hasil di Bab 4.4 menunjukkan akurasi tinggi untuk field tekstual (NPWP: 99.46%, nama: 99.49%) tetapi rendah untuk field numerik (penghasilan bruto:

87.12%, PPh dipotong: 85.07%) pada Lampiran 1721-I, dan 0% untuk semua field pada Lampiran 1721-V. Hal ini sejalan dengan teori error minimization, yang menurut Hastie et al. (2009) menekankan pentingnya meminimalkan kesalahan prediksi dalam model statistik. Kegagalan pada Lampiran 1721-V menunjukkan perlunya optimalisasi model, seperti fine-tuning atau prompt engineering, untuk meningkatkan akurasi pada data kompleks, sebagaimana direkomendasikan di Bab V.

2.2 Penelitian Terdahulu

Tabel 2.1 Data List Perusahaan Data Sintetis?

No	Nama	Tahun	Judul	Keterangan
1	(Firya Shafira, 2023)	2023	Analisis Dampak Pengembangan Kecerdasan Buatan (AI) pada Media Sosial Twitter Menggunakan Metode Support Vector Machine dan Lexicon Based	membahas dampak sentimen AI pada masyarakat dengan SVM, dan dapat disimpulkan positif yang berarti respon masyarakat baik.
2	(Lingzhen Chen, 2019)	2019	Transfer Learning for Sequence Labeling Using Source Model and Target Data	meneliti mengenai pelabelan di transfer learning model
3	(Wolf et al., n.d.)	2019	Transformers: State-of-the-Art Natural Language Processing	meneliti tentang Transformers milik Hugging Face yang merupakan model untuk Penelitian NLP.

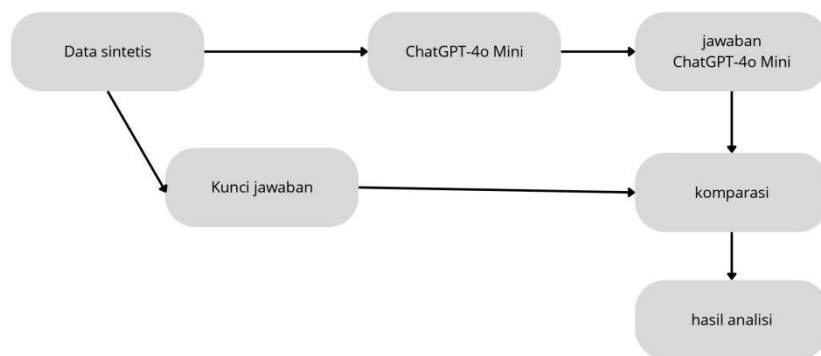
4	(Hu et al., 2020)	2020	Supervised Machine Learning Techniques: An Overview with Applications in Banking	meneliti ML menggunakan model supervised dalam masalah perbankan, dan hasilnya ternyata Neural network lebih akurat dari pada metode yang lain.
5	(Ziegler et al., 2019)	2019	Fine-Tuning Language Models from Human Preferences	memperajari bagaimana model AI yang menggunakan fine tuning manusia langsung bekerja, dan hasilnya memuaskan, dan membuat kinerja AI berjalan dengan baik sesuai aturan yang berlaku
6	(Saputri & Prasetyawati, 2024)	2024	Pengaruh Penggunaan Chat GPT Terhadap Efisiensi Komunikasi Pada Karyawan PT Modern Abadi	Meneliti penggunaan AI pada efisiensi karyawan perusahaan dengan 2 variable, yaitu chat GPT dan efisiensi
7	(Jusman1, 2024)	2024	Analisis Pemanfaatan Kecerdasan Buatan Berbasis Chat GPT Untuk Membantu Mahasiswa Jurusan Teknologi Pendidikan Di	Penelitian Kualitatif mengenai penggunaan Chat GPT dalam membantu Mahasiswa. Untuk referensi penelitian

			Universitas Muhammadiyah Bone	
8	(Bauer et al., 2024)	2024	Comprehensive Exploration of Synthetic Data Generation: A Survey	Penelitian penggunaan data sintetis pada SGD, menggunakan survey untuk menggenerasi survey.
9	(Sufi, 2024)	2024	Generative Trained Transformer (GPT) in Research: A Systematic Review on Data Augmentation	Pre-Penelitian menggunakan GPT pre train, untuk data augmentasi (data buatan) dengan lebih spesifik
10	(Goncalves et al., 2020)	2020	Generation and evaluation of synthetic patient data	Meneliti mengenai, melakukan generasi data sintetis dengan data yang sudah ada.

2.3 Kerangka Pemecahan Masalah

Kerangka pemecahan masalah dalam penelitian ini dirancang untuk mengatasi tantangan dalam pelaporan Surat Pemberitahuan (SPT) Pajak Penghasilan (PPh) Pasal 21 di Indonesia dengan memanfaatkan kecerdasan buatan berbasis ChatGPT-4o Mini. Proses ini mencakup enam tahap utama yang digambarkan dalam diagram alur: *Data sintetis*, *ChatGPT-4o Mini*, *Jawaban ChatGPT-4o Mini*, *Kunci jawaban*, *Komparasi*, dan *Hasil analisis*. Pendekatan ini bertujuan meningkatkan efisiensi, akurasi, dan kepatuhan terhadap peraturan

perpajakan seperti PER-2/PJ/2024, sekaligus mengidentifikasi keterbatasan sistem untuk pengembangan lebih lanjut, sebagaimana diuraikan di Bab V.



Gambar 2.2 konseptual pengerjaan

2.3.1 Data Sintetis

Tahap awal dimulai dengan pembuatan *Data sintetis* menggunakan script *skripsi_1.py*. Data ini dirancang untuk mensimulasikan informasi perusahaan dan karyawan, termasuk 400 file dengan 15 hingga 80 karyawan per file, mencakup karyawan tetap dan tidak tetap. Data sintetis mencerminkan komponen pelaporan PPh 21, seperti Nomor Pokok Wajib Pajak (NPWP), Nomor Induk Karyawan (NIK), penghasilan bruto, tunjangan, potongan (biaya jabatan, BPJS), Penghasilan Tidak Kena Pajak (PTKP), dan PPh dipotong, sesuai dengan ketentuan PER-2/PJ/2024. Distribusi data dibuat realistis, dengan tidak di tetapkannya secara eksplisit pekerja karyawan , untuk menguji variasi skenario pelaporan.

Penggunaan data sintetis dipilih untuk menjaga kerahasiaan data riil dan memungkinkan kontrol variabel, seperti jumlah karyawan, status perkawinan, dan jumlah tanggungan, yang memengaruhi perhitungan pajak.

2.3.2 Persiapan ChatGPT-4o Mini

Data sintetis selanjutnya diproses oleh *ChatGPT-4o Mini* melalui script *AI_skripsi_1.py*. Model ini, berbasis arsitektur Transformer yang diperkenalkan oleh Vaswani et al. (2017) , memanfaatkan mekanisme *self-attention* untuk memahami dan menghasilkan laporan SPT PPh 21 dalam format JSON dan PDF. Proses ini melibatkan pengiriman data sintetis ke API ChatGPT-4o Mini, di mana model menerima prompt yang dirancang untuk mematuhi aturan PER-2/PJ/2024, termasuk perhitungan tarif progresif dan pengisian lampiran seperti 1721-I (karyawan tetap) dan 1721-V (karyawan tidak tetap). Yang harusnya bisa menunjukkan potensi signifikan dibandingkan metode manual yang sering memakan waktu berjam-jam.

2.3.3 Jawaban ChatGPT-4o Mini

Hasil pemrosesan oleh *ChatGPT-4o Mini* menghasilkan *Jawaban ChatGPT-4o Mini*, yaitu laporan SPT PPh 21 dalam format terstruktur. Laporan ini mencakup detail karyawan, perhitungan pajak, dan lampiran yang diperlukan untuk pengajuan ke Direktorat Jenderal Pajak (DJP). Output ini dihasilkan berdasarkan interpretasi data sintetis dan aturan perpajakan yang tertanam dalam prompt. Namun, ketidaksesuaian dengan *ground truth* menunjukkan bahwa model kadang-kadang salah menginterpretasikan data numerik atau mengabaikan konteks spesifik, seperti perbedaan perlakuan pajak antara karyawan tetap dan tidak tetap.

2.3.4 Kunci Jawaban

Kunci jawaban berfungsi sebagai *ground truth*, dihasilkan oleh script *key_1.py* berdasarkan perhitungan manual sesuai PER-2/PJ/2024. Script ini mensimulasikan perhitungan pajak yang akurat, mencakup semua komponen seperti PTKP, tarif progresif, dan penyesuaian untuk karyawan tetap dan tidak tetap. *Kunci jawaban* ini menjadi standar emas untuk membandingkan output AI, memastikan evaluasi objektif berdasarkan regulasi pajak yang berlaku.

Penggunaan *ground truth* memungkinkan pengukuran tingkat kecocokan dengan toleransi nol, terutama untuk field numerik seperti PPh terutang. Ketidaksesuaian antara *Jawaban ChatGPT-4o Mini* dan *Kunci jawaban* menyoroti keterbatasan model dalam menangani variasi data, terutama untuk karyawan tidak tetap dengan penghasilan tidak teratur. Hasil ini konsisten dengan temuan di Bab 4.4 dan menjadi dasar rekomendasi di Bab V untuk meningkatkan kemampuan model melalui fine-tuning atau integrasi aturan pajak tambahan.

2.3.5 Komparasi

Tahap *Komparasi* dilakukan menggunakan script *statistik_skripsi.py*, yang membandingkan *Jawaban ChatGPT-4o Mini* dengan *Kunci jawaban*. Proses ini menghitung persentase kecocokan untuk setiap field lampiran (1721-I sampai 1721-V) pada 400 file data.

Analisis komparasi juga mengidentifikasi pola kegagalan, seperti kesalahan perhitungan pada penghasilan, yang sering disebabkan oleh interpretasi prompt yang kurang presisi. Data ini divisualisasikan pada Gambar 4.4.1 (*spt_global_overall_match_rate*) dan Gambar 4.4.6 (*spt_lampiran_1721_V_entry_match_rate*), memberikan wawasan tentang performa sistem dan menjadi dasar untuk rekomendasi perbaikan, seperti penggunaan memori eksternal atau pelatihan khusus untuk data kompleks.

2.3.6 Hasil Analisis

Hasil analisis merupakan output akhir yang merangkum performa sistem berdasarkan komparasi. Temuan utama menunjukkan bahwa ChatGPT-4o Mini efektif untuk efisiensi waktu dan akurasi pada data terstruktur (Lampiran 1721-I), tetapi terbatas pada data tidak terstruktur (Lampiran 1721-V). Hal ini konsisten dengan teori Transformer (Vaswani et al., 2017), yang unggul dalam memproses data tekstual tetapi memerlukan optimalisasi untuk data numerik dan tidak teratur.

Implikasi praktis meliputi potensi penggunaan sistem untuk UKM dalam menyusun draft laporan awal. Bab V merekomendasikan perbaikan seperti fine-tuning model, pengembangan prompt adaptif, dan integrasi dengan sistem e-SPT untuk meningkatkan akurasi dan kepatuhan. Penelitian ini juga membuka peluang untuk studi lanjutan, seperti pengujian dengan data riil atau skalabilitas pada jumlah karyawan yang lebih besar, mendukung digitalisasi administrasi perpajakan di Indonesia.