

CHAPTER 3

KECERDASAN ARTIFISIAL UNTUK PENGOLAHAN UCAPAN DAN TEKS BERBAHASA INDONESIA

Ayu Purwarianti¹, Dessi Puji Lestari¹, & Devin Hoesen²

¹Pusat AI ITB (PUI-PT AI-VLB)

²PT Prosa Solusi Cerdas

ABSTRAK

Teknologi pemrosesan bahasa alami atau *Natural Language Processing* (NLP) pada teks serta suara untuk bahasa Indonesia sudah mengalami perkembangan yang cukup pesat. Teknologi NLP dapat membantu pelaksanaan berbagai proses dalam sebuah institusi supaya lebih efisien. Dalam kesempatan ini, akan dibahas tiga *use case* pemanfaatan teknologi NLP di institusi, yaitu *regulatory technology*, *meeting transcription*, dan *voice biometrics*.

Regulatory technology adalah sebuah teknologi berbasis NLP yang bertujuan membantu pengecekan peraturan ataupun dokumen perusahaan secara otomatis. *Regulatory technology* berfungsi untuk membandingkan peraturan atau dokumen perusahaan dengan berbagai peraturan pemerintah terkait yang dikumpulkan secara otomatis dari berbagai situs peraturan pemerintah. *Regulatory technology* memanfaatkan berbagai teknologi NLP untuk pemrosesan teks, baik pencarian, ekstraksi, maupun klasifikasi teks. Saat ini, sudah ada beberapa organisasi di Indonesia yang menggunakan produk *regulatory technology* untuk melakukan *compliance checking*.

Meeting transcription adalah teknologi berbasis pemrosesan suara yang bertujuan mengubah suara hasil rekaman rapat ke dalam teks transkripsinya. Naskah rekaman suara rapat mencakup informasi teks yang diucapkan oleh setiap peserta rapat beserta identitas peserta rapat tersebut. *Meeting transcription* menggunakan teknologi *automatic speech recognition* (ASR) untuk mengubah suara menjadi teks; teknologi *speech diarization* untuk membedakan suara peserta rapat, ekstraksi kata kunci, dan pencarian catatan rapat maupun segmen percakapan. Dalam pengujian untuk delapan set data, teknologi ASR bahasa Indonesia yang dikembangkan memiliki nilai akurasi yang lebih tinggi dibandingkan ASR Bahasa Indonesia milik Google. Teknologi *meeting transcription* ini sudah dipasang di lembaga pemerintah, BUMN, dan lembaga legislatif di Indonesia.

Voice biometrics merupakan salah satu alternatif teknologi autentikasi identitas dengan berdasar

A. Purwarianti, D. P. Lestari, & D. Hoesen
Pusat AI ITB, e-mail: ayu@stei.itb.ac.id

@ 2023 Kolaborasi Riset dan Inovasi Industri Kecerdasan Artifisial (KORIKA) & Penerbit BRIN
A. Purwarianti, D. P. Lestari, & D. Hoesen. "Kecerdasan artifisial untuk pengolahan ucapan dan teks berbahasa Indonesia," in *Prosiding Use Cases Artificial Intelligence Indonesia: Embracing Collaboration for Research and Industrial Innovation in Artificial Intelligence*, B. R. Trilaksono, H. Riza, A. Jarin, N. D. S. Darmayanti, and S. Liawatimena, Eds. Jakarta: Penerbit BRIN, Februari 2023, ch. 3, pp. 39-43, doi: 10.55981/brin.668.c536
ISBN: 978-623-8052-49-3, E-ISBN: 978-623-8052-50-9

pada suara. Akurasi *voice biometrics* dengan kondisi suara bersih sudah mendekati 100%. Saat ini, produk *voice biometrics* sudah dipakai salah satu bank di Indonesia untuk layanan *customer service* melalui telepon. Selain itu, *voice biometrics* juga dapat digunakan untuk melengkapi atau bahkan menggantikan kata sandi dan sidik jari dalam pengaksesan suatu aplikasi ponsel.

Kata kunci: *regulatory technology, meeting transcription, voice biometrics, natural language processing*, bahasa Indonesia

A. PENDAHULUAN

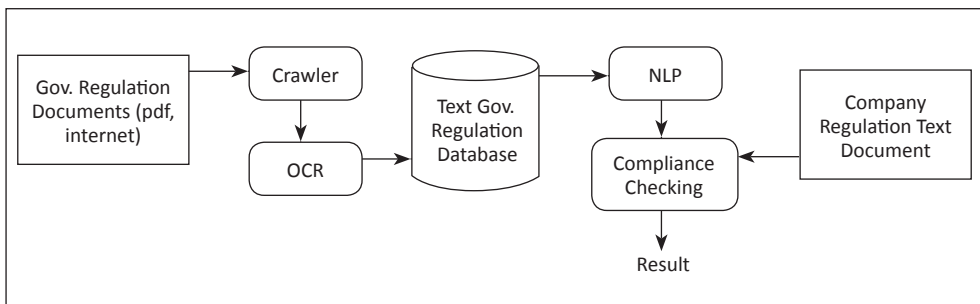
Ucapan dan teks merupakan komponen penting dalam interaksi antara manusia dan komputer. Ucapan adalah antarmuka komunikasi masa depan yang akan banyak digunakan dalam komunikasi karena merupakan cara paling alami. Meskipun begitu, teks juga tetap merupakan bagian penting dari komunikasi, seperti pada *email*, dan *social media*. Selain dalam komunikasi, ucapan dan teks juga merupakan representasi informasi yang tersedia secara digital dalam ukuran besar, seperti dokumen atau artikel. Baik untuk komunikasi maupun sebagai representasi informasi, pemrosesan otomatis terhadap ucapan dan teks dapat meningkatkan efisiensi waktu pemrosesan. Pemrosesan otomatis ucapan dan teks merupakan salah satu cabang dari teknologi kecerdasan artifisial. Berikut ini beberapa *use case* pemrosesan ucapan dan teks secara otomatis.

B. REGULATORY TECHNOLOGY

Regulatory technology atau sering disebut *regtech* merupakan istilah untuk menyatakan penggunaan teknologi digital pada peraturan. Dewasa ini, *regulatory technology* semakin banyak digunakan di berbagai instansi. Pada 2021, perkiraan *global revenue* di bidang *regtech* mencapai 2,87 miliar \$US [1]. Perkembangan *regtech* ditandai dengan penerapan kecerdasan artifisial (pemrosesan bahasa alami), baik di sektor swasta maupun *regulatory agency*. Dalam lingkup pemrosesan bahasa alami, *regtech* merupakan bagian dari *natural language understanding* [2], dengan input proses adalah teks dalam bentuk dokumen dan *output* proses adalah analisis, seperti hasil perbandingan, ekstraksi, dan klasifikasi.

Dalam sektor swasta, salah satu penerapan kecerdasan artifisial pada *regulatory technology* adalah untuk pengecekan *compliance*. Sektor *fintech* merupakan industri yang banyak memanfaatkan fitur pengecekan *compliance* pada *regulatory technology*. Fitur pengecekan *compliance* bertujuan memastikan berbagai aturan atau produk perusahaan masih sesuai dengan regulasi yang dikeluarkan *regulatory agency*. Pada fitur ini, teknologi kecerdasan artifisial yang digunakan meliputi ekstraksi informasi dan perbandingan otomatis. Kedua teknologi ini digunakan juga untuk menghasilkan informasi hubungan antarperaturan (*regulation connectivity*).

Selain kedua teknologi tersebut, teknologi kecerdasan artifisial yang juga digunakan untuk *regulatory technology* adalah OCR. OCR digunakan untuk memperbaiki hasil *crawling* dokumen regulasi yang biasanya disimpan dalam *file pdf* (untuk kasus di Indonesia). Penggunaan *library* terkait konversi pdf ke dalam teks tanpa penggunaan OCR akan menghasilkan banyak *noise* pada luaran teksnya. Teks hasil OCR ini kemudian akan diproses dengan teknologi NLP seperti ekstraksi informasi dan perbandingan otomatis yang telah dijelaskan sebelumnya. *Flow* umum dari *proses compliance checking* yang di dalamnya terdapat pemanfaatan teknologi OCR dapat dilihat pada Gambar 1.

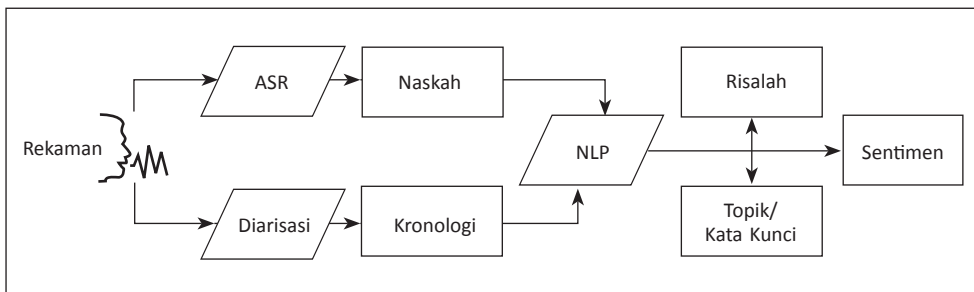


Gambar 1. *Flow* Umum Fitur *Compliance Checking* pada *Regulatory Technology*

C. MEETING TRANSCRIPTION

Meeting transcription adalah suatu sistem terautomasi yang dapat menaskahkan suatu rekaman rapat atau percakapan dan dapat dilengkapi juga dengan siapa yang mengucapkan kalimat-kalimat dalam rapat/percakapan tersebut [2]. Di dalam suatu sistem *meeting transcription*, terdapat berbagai teknologi yang digunakan, yaitu *automatic speech recognition* (ASR), diarisasi, dan *natural language processing* (NLP). Teknologi ASR digunakan untuk mengubah sinyal suara dalam rekaman rapat/percakapan menjadi runtunan teks yang berisi kalimat-kalimat yang diucapkan dalam rapat/percakapan. Dengan berkembangnya teknik-teknik ASR yang menggunakan jaringan saraf tiruan (JST), ketepatan penerjemahan ASR sekarang sudah mencapai di atas 90%, bahkan mendekati 100% untuk berbagai kondisi suara. Dalam pengujian yang dilakukan terhadap delapan set data, ASR yang dikembangkan khusus untuk bahasa Indonesia telah dapat mengalahkan akurasi ASR Google. Diarisasi digunakan untuk menentukan siapa berbicara apa dan pada waktu kapan. Teknik diarisasi dapat memenggal rekaman, kemudian menentukan siapa yang berbicara dalam penggalan-penggalan tersebut. Diarisasi diperlukan untuk mendapat kronologi jalannya rapat/percakapan. Gabungan ASR dan diarisasi ini dapat membangun kronologi lengkap rapat/percakapan beserta naskahnya yang dapat digunakan untuk proses lanjutan, yakni NLP.

Teknologi NLP dapat melengkapi sistem *meeting transcription* untuk mempermudah pengguna mendapatkan informasi-informasi dari rapat/percakapan tersebut. Sebagai contoh, sentimen setiap kalimat dapat disarikan menggunakan teknik *sentiment analysis*. Topik atau kesimpulan rapat/percakapan juga dapat disarikan berdasarkan ringkasan rapat yang dihasilkan teknik *summarization*. Selain itu, dengan adanya penyarian topik atau kata-kata kunci dari rapat, pengguna dapat mencari dan mengelola rapat/percakapan berdasarkan topik atau kata kunci yang muncul dalam rapat/percakapan. Sebagai pelengkap, sistem *meeting transcription* juga dapat menghasilkan risalah rapat/percakapan yang idealnya dapat disesuaikan dengan format yang ditetapkan atau diinginkan pengguna/organisasi. Secara umum, alur dalam sistem *meeting transcription* ditunjukkan dalam Gambar 2. Semua teknologi yang tercakup dalam *meeting transcription* ini sudah dipasang di berbagai lembaga pemerintah, BUMN, dan lembaga legislatif di Indonesia untuk mempermudah dan mempercepat proses produksi risalah pascarapat. Selain itu, penyimpanan data dan rekaman rapat secara digital dalam bentuk suara maupun teks akan memudahkan pencarian dan sangat penting di berbagai lembaga di mana keputusan-keputusan penting dalam rapat harus diperiksa dan diaudit secara teliti.



Gambar 2. Alur Ideal *Meeting Transcription*

D. VOICE BIOMETRICS

Teknologi *voice biometrics* atau biometrik suara merupakan teknologi yang menggunakan suara seseorang untuk mengidentifikasi orang tersebut. Sistem *voice biometrics* terautomasi menggunakan teknik pembelajaran mesin untuk mendapatkan karakteristik khas suara seseorang (diistilahkan sebagai sidik suara) yang dapat membedakannya dari orang lain. Teknik *voice biometrics* termutakhir sudah andal sehingga dapat mencapai akurasi mendekati 100%, meskipun terdapat variasi kondisi suara yang diproses. Setelah sidik-sidik suara telah didaftarkan, *voice biometrics* dapat dipakai mengidentifikasi maupun memverifikasi suara seseorang. Identifikasi dilakukan ketika sidik suara baru ingin dicari siapa pemiliknya suaranya atau identitasnya dari suatu koleksi sidik suara yang terdaftar. Verifikasi terjadi ketika sistem ingin memastikan bahwa suatu sidik suara baru betul milik suatu identitas yang diklaim [2].

Teknologi *voice biometrics* ini sudah digunakan untuk verifikasi pada salah satu bank di Indonesia. Dalam penggunaannya, pelanggan yang menelepon *customer service* diverifikasi secara otomatis saat sedang menjawab pertanyaan-pertanyaan keamanan. Hasil *voice biometrics* dapat dipakai untuk memperkuat sistem verifikasi pertanyaan keamanan yang biasa digunakan oleh bank. Di sisi lain, *voice biometrics* juga dapat dipakai sebagai pengganti sandi atau sidik jari untuk masuk ke dalam aplikasi ponsel. Sidik suara pengguna didaftarkan dengan meminta pengguna mengucapkan suatu frasa tertentu beberapa kali. Saat pengguna nantinya mencoba untuk masuk aplikasi, dia dapat kembali mengucapkan frase yang diminta untuk diverifikasi.

DAFTAR PUSTAKA

- [1] D. Broby, A. Daly, dan D. Legg, "Towards secure and intelligent regulatory technology (Regtech): a research agenda." *Technology and Regulation*. Access: 24 Nov 2022. [Online]. Available: <https://techreg.org/article/view/12475/14818>
- [2] A. Purwarianti, dkk. *Artificial Intelligence di Masa Pandemi*. Bandung: ITB Press. 2021