

ABSTRAK

Text to speech (TTS) merupakan salah satu aplikasi dalam bidang teknologi informasi sebagai salah satu cara interaksi antara manusia dan komputer dengan cara mengkonversi teks menjadi ucapan/suara. Saat ini sudah dibuat *TTS* Bahasa Indonesia untuk penggunaan di *PC* yaitu *IndoTTS*, namun pelafalan pada *IndoTTS* ini masih belum natural.

Dalam sistem *Text to Speech (TTS)*, sebenarnya pembentukan intonasi yang benar merupakan faktor penting yang akan mempengaruhi pembacaan pada *output* sistem. Bagian yang mengatur pembentukan intonasi di bagian *output* sistem ini disebut pembangkitan prosodi. Intonasi yang dihasilkan dari bagian ini meliputi durasi dan *pitch* pengucapan teks *input*.

Proses penentuan nilai durasi dan *pitch* dari teks *input* bersifat kompleks dan *non-linear*, maka sebagai dasar sistem digunakan *Multilayer Perceptron Neural Network (MLPNN)* sebagai model prosodinya. Model prosodi berbasis *MLPNN* membentuk intonasi pengucapan teks *input* dengan cara menentukan nilai durasi dan *pitch* dari tiap fonem penyusun teks *input*. Penentuan nilai durasi dan *pitch* dilakukan setelah sistem melakukan pembelajaran terhadap sampel pengucapan dari suatu kalimat.

Dalam tugas akhir ini, telah dibuat aplikasi *Text to Speech* Bahasa Indonesia dengan disertai peningkatan *natural* (kealamian) dalam pelafalan kalimat berupa teks (tidak termasuk angka dan simbol-simbol). Pembuatan aplikasi *TTS* ini menggunakan bahasa pemrograman Borland Delphi 7.0 dan memanfaatkan database *diphone* Bahasa Indonesia yang sudah tersedia serta menggunakan pembangkit ucapan *Mbrola*. Setelah dilakukan pengambilan *MOS* dari 30 koresponden yang terdiri dari mahasiswa dan masyarakat sekitar, didapatkan bahwa hasil dari sistem *TTS* dengan model prosodi *MLPNN* lebih baik kualitasnya dibandingkan dengan sistem *IndoTTS*.

Kata kunci : Prosodi, *MLPNN*, *diphone*, *MBROLA*.