**PENGEMBANGAN SISTEM *DATA-TO-TEXT* (D2T)UNTUK MEMBANGKITKAN BERITA PADA DATA *UNSPECIFIC***

**SKRIPSI**

Diajukan untuk Memenuhi Bagian Dari

Syarat Memperoleh Gelar Sarjana Komputer

Program Studi Ilmu Komputer



Oleh:

Muhammad Ridwan (1403407)

**PROGRAM STUDI ILMU KOMPUTER**

**DEPARTEMEN PENDIDIKAN ILMU KOMPUTER**

**FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS PENDIDIKAN INDONESIA**

**BANDUNG**

**2018**

**PENGEMBANGAN SISTEM *DATA-TO-TEXT* (D2T) UNTUK MEMBANGKITKAN BERITA PADA DATA *UNSPECIFIC***

# ABSTRAK

Sistem *Data-to-Text* menjadi salah satu pilihan untuk menerjemahkan data *non-linguistik* kedalam bentuk tekstual. Namun seiring dengan perkembangan teknologi, beragamnya bidang dari suatu data dan beragamnya pengguna menjadi salah satu fokus yang harus diperhatikan dalam pengembangan sistem *Data-to-Text.* Penelitian ini bertujuan untuk mengembangkan sistem *Data-to-Text* dengan masukan berupa data *general*, sebagai solusi agar sistem *Data-to-Text* dapat menerima masukan berupa data dari bidang atau domain apapun, baik data tersebut memiliki identitas berupa informasi *header*, tipe data, *rule* ataupun tidak. Maka digunakan pendekatan *Fuzzy Rule* untuk menginterpretasikan data secara *general*. Selain itu digunakan beberapa algoritma *Machine Learning* seperti *Gradient Descent,* dan analisis lainnya seperti *Exponential Smoothing, Knuth-Morris-Pratt,* dan *Pearson Correlation Coefficient.* Sistem yang dikembangkan dapat menghasilkan informasi berupa ringkasan data, informasia data terkini dan informasi prediksi. Pengembangan sistem dilakukan dalam bahasa pemrograman R dengan memanfaatkan beberapa *packages* yang tersedia. Eksperimen dilakukan dengan mengukur tingkat *Readibility* dari berita yang dibangkitkan*, Computation Time*, dan membandingkan hasil dengan penelitian terkait. Hasil eksperimen menunjukan bahwa informasi yang dihasilkan terbukti merepresentasikan data yang diberikan dan dapat dipahami oleh tingkat siswa pada tingkat sekolah dasar sekalipun, serta waktu komputasi cukup baik. Sistem ini mampu menhasilkan informasi berdasarkan data meteorologi, data klimatologi, data keuangan, dan data *time series* lainnya.

**Kata Kunci**— *Data-to-Text; Natural Language Processing, Natural Language Generation; Machine Learning; General purpose; Unspecific Corpora; Fuzzy Rule-based; Crisp Rule-based; Time-series Analysis; Exponential Smoothing; Linear Model; Gradient Descent; Kunth-morris-pratt; Pearson Correlation Coefficient*