Laporan Group Project – DTS 2019  
Text Summarization using NLTK

Hari/Tanggal : Jumat, 09 Agustus 2019

|  |  |  |
| --- | --- | --- |
| Anggota Kelompok | : | 1. Anthony 2. Muhammad Faris Herlansyah 3. Rizki Triharyo Ibrahim 4. William Eko Saputro |
| Judul Pekerjaan | : | Menarik kesimpulan dari transkrip debat pilpres kelima Capres – Cawapres RI 2019 |
| Deskripsi Singkat Pekerjaan | : | Transkrip debat diekstrak per segmen debat lalu digunakan Text Summarization untuk mengambil poin-poin penting dari debat kedua pasangan calon.  Stopwords dihilangkan secara manual karena artikel dalam Bahasa Indonesia sehingga bisa mendapatkan keywords yang lebih akurat |
| Tujuan Pekerjaan | : | Membantu mengambil poin-poin penting isi dari debat pilpres kelima capres-cawapres RI 2019 |
| Dokumen Teks yang digunakan | : | **Transkrip Lengkap Debat Kelima Capres - Cawapres RI tanggal 13 April 2019**  **dan Analisis Konten Isi Debat Capres – Cawapres RI**  oleh : PT Bahasa Kita  Sumber : <https://s3-ap-southeast-1.amazonaws.com/debatcapres/release/document.pdf> |
| Komentar tentang hasil yang diperoleh | : | Pada beberapa segmen yang kami coba, NLTK mampu untuk merangkum dengan cukup teliti (contohnya pada segmen ke 5, sudah dapat merangkum apa yang diharapkan oleh Jokowi dan Sandiaga Uno terhadap BUMN di masa yang akan datang). Akan tetapi, pada segmen lainnya, beberapa transkrip yang bersifat mediasi (misalnya waktunya habis dan lain-lain) dapat menghalangi proses perangkuman debat. Diharapkan untuk proses merangkum debat dapat dihilangkan transkrip-transkrip yang bersifat pengantar atau penengah dari debat. |