

Laporan Group Project – DTS 2019

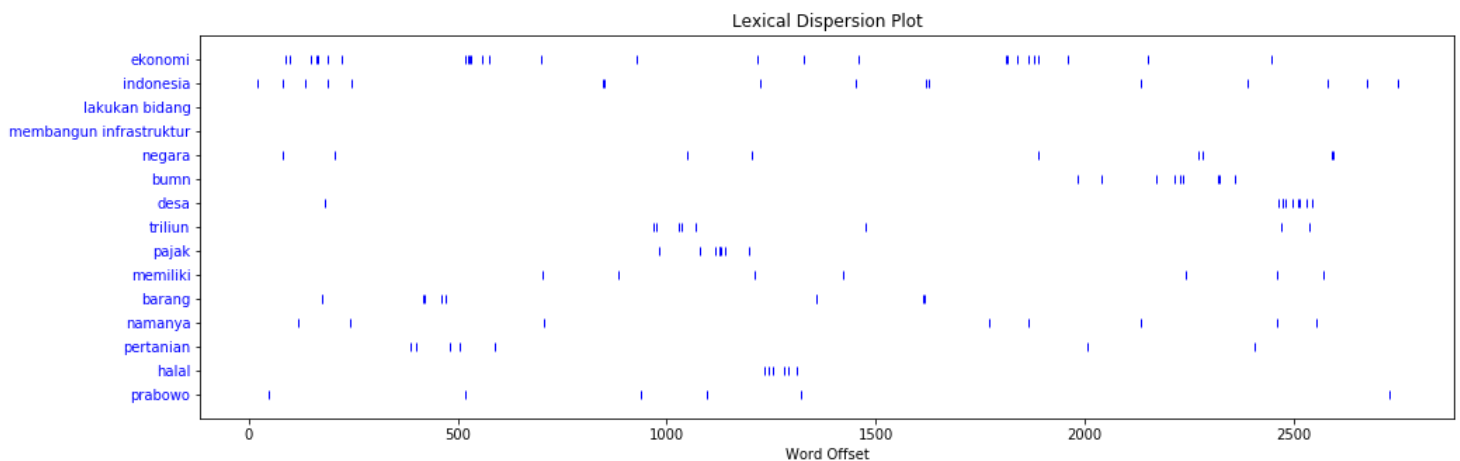
Text Summarization Visualization

Hari/Tanggal: Selasa, 13 Agustus 2019

Anggota Kelompok	:	1. Anthony 2. Muhammad Faris Herlansyah 3. Rizki Triharyo Ibrahim 4. William Eko Saputro
Judul Pekerjaan	:	Visualisasi Dialog pada Transkrip Debat Pilpres Kelima Capres dan Cawapres RI 2019
Deskripsi Singkat Pekerjaan	:	<ul style="list-style-type: none">• Dokumen transkrip diolah sehingga dibuat file text berisi dialog dari masing-masing pembicara.• Setelah terpisah, teks diolah dengan dibuang stopwords• Lalu teks Transkrip Debat Pilpres Kelima Capres dan Cawapres RI 2019 divisualisasikan menggunakan <i>Wordcloud</i> dan <i>Lexical Dispersion Plot</i>
Tujuan Pekerjaan	:	Memvisualisasikan teks transkrip debat pilpres kelima Capres dan Cawapres RI 2019 agar memudahkan pengamat dalam menemukan pemikiran dari pembicara dengan melihat kata-kata yang dipakai
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	:	<p>1. <i>Wordcloud</i> Pada Wordcloud yang ditampilkan, masih terdapat beberapa kata-kata yang belum terlalu menggambarkan keseluruhan topik yang diangkat oleh pembicara Bahasa Indonesia belum mendukung pengenalan kata-katanya. Serta beberapa kata yang tidak begitu penting namun cukup sering diucapkan dan tidak termasuk ke dalam <i>list stopwords</i> dalam bahasa Indonesia menurut <i>library</i> bawaan dari NLTK bisa muncul pada wordcloud dalam ukuran besar; hal ini bisa diatasi dengan mengeliminasi sendiri kata-kata tersebut. Walaupun begitu, kata-kata yang muncul dalam Wordcloud cukup menggambarkan apa saja pemikiran-pemikiran yang ditekankan pembicara pada debat.</p> <p>2. <i>Lexical Dispersion Plot</i> (LDP) Kami memakai kata-kata kunci yang paling sering muncul dalam dialog pembicara sebagai topik pada LDP. LDP mungkin memberikan gambaran kapan kata-kata yang paling sering dipakai tersebut muncul. Tetapi belum memberikan gambaran bagaimana kata-kata tersebut dipakai dan belum dapat memberitahu apakah kata-kata kunci tersebut selalu membicarakan konteks yang sama atau tidak, kecuali jika kita mengetahui posisi segmen-segmen pada <i>word offset</i>.</p>



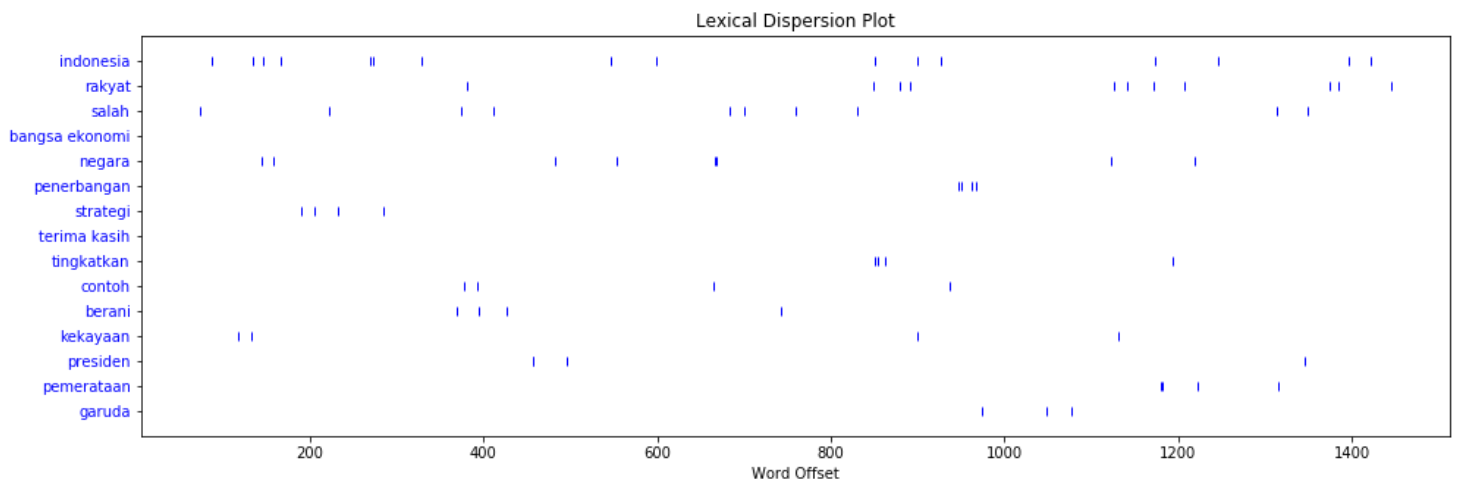
Gambar 1.1 Wordcloud untuk capres Joko Widodo



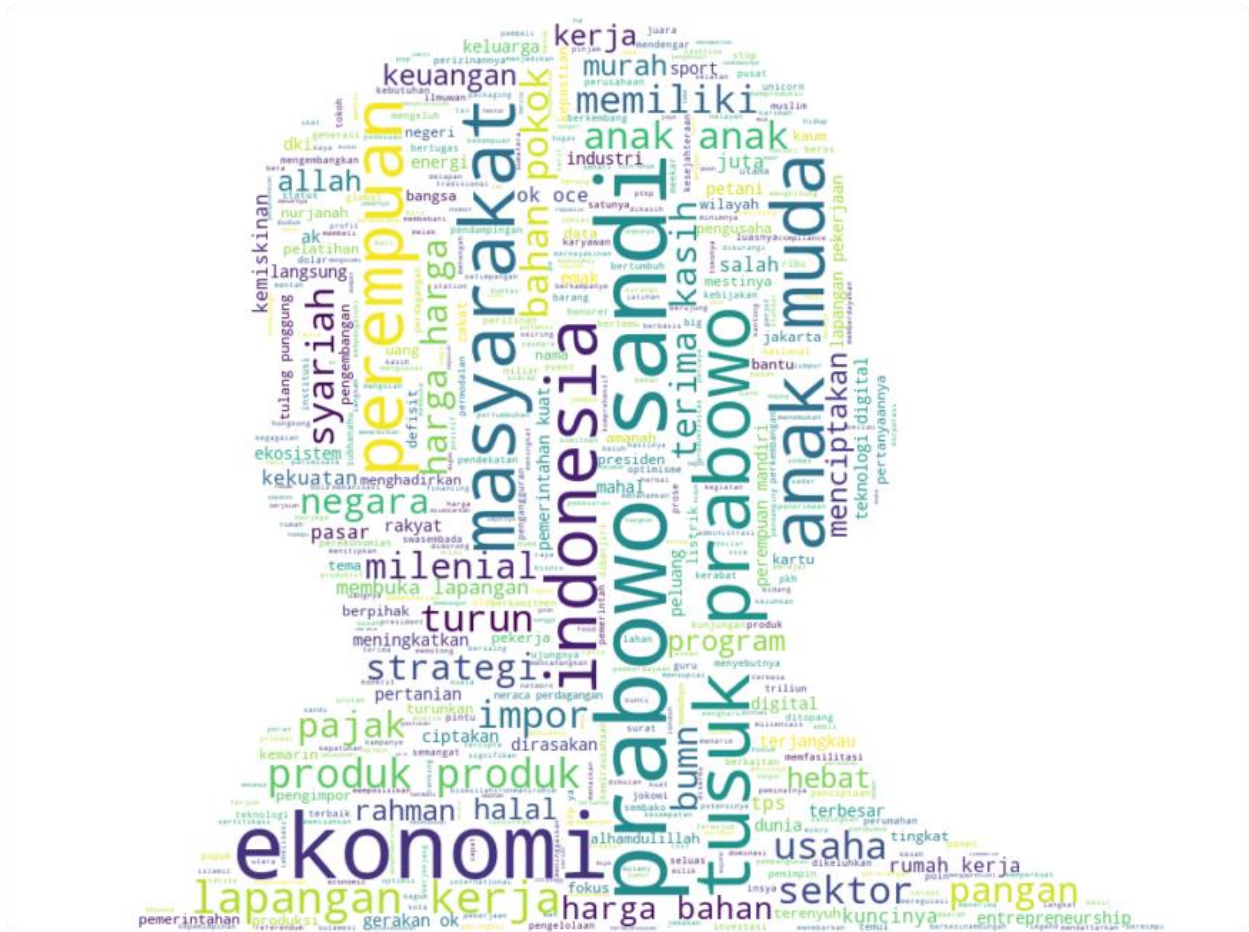
Gambar 1.2 LDP Joko Widodo



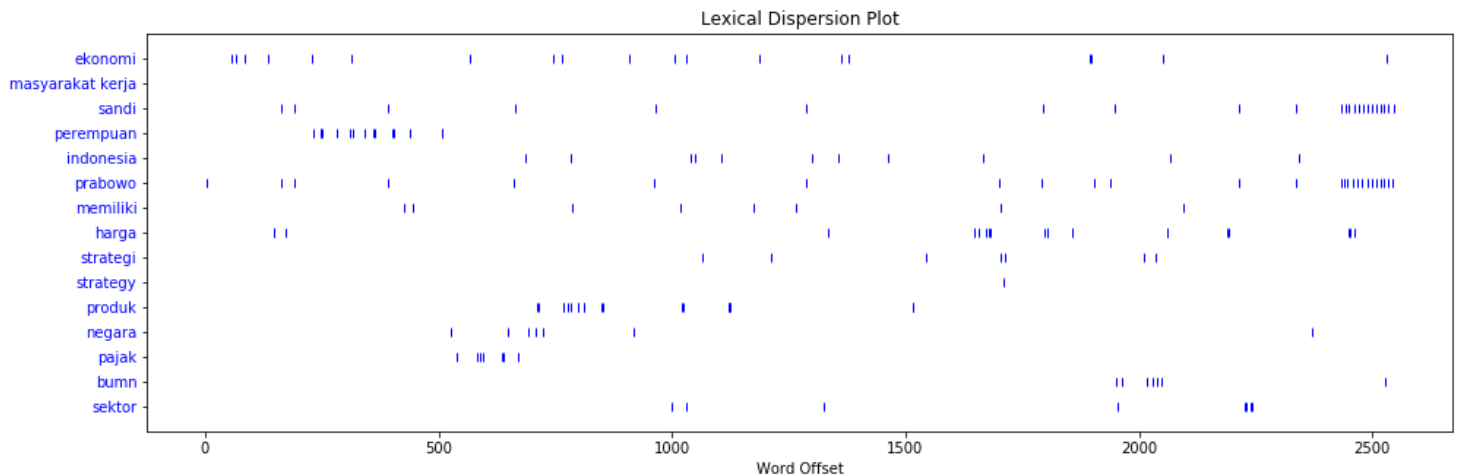
Gambar 2.1 Wordcloud untuk capres Prabowo



Gambar 2.2 LDP Prabowo



Gambar 4.1 Wordcloud untuk cawapres Sandiaga Uno



Gambar 4.2 LDP Sandiaga Uno