# **Tinjauan Pustaka**

## **Analisis Sentimen Berkaitan dengan Demokrasi**

Komputasi demokrasi yang dilakukan guna mempercepat penghitungan nilai demokrasi di suatu negara telah dimulai sejak tahun 2014 dimana Lutz dan Toit (2014) mencoba untuk mengukur demokrasi dengan menangkap ekspresi, tanggapan atau pendapat masyarakat terkait dengan demokrasi dari sosial media bernama *twitter*. Pertama-tama, Lutz dan Toit (2014) mengidentifikasi demokrasi yang akan diukur melalui kajian teori tentang paham demokrasi itu sendiri. Menurutnya, demokrasi adalah jenis sistem pemerintahan yang perlahan-lahan kuar dari sistem otoriter.

Demokrasi ditandai dengan kebebasan berekspresi setiap individu dalam ruang publik yang membahas segala sesuatu yang sesuai dengan pendapat dan kebebasan berpikirnya. Mereka membandingkan antara metode jajak pendapat secara langsung (manual) menanyai individu mengenai isu demokrasi dengan hasil analisis data *twitter*. Variabel penelitian tidak dijelaskan dalam publikasi ini. Namun, hasilnya menunjukkan bahwa jajak pendapat yang ditangkap melalui situs *twitter* lebih efisien secara waktu dan biaya.

Setelah itu, Marzagao (2017) mengeluarkan publikasi berupa hasil studi doktornya yang berjudul “Automated Democracy Score”. Dia memulai penelitian dengan mempelajari pengukuran-pengukuran demokrasi yang telah dilakukan oleh negara-negara di dunia. Dua pengukuran demokrasi yang paling banyak digunakan ada *Polity* dan *Freedom House*. Keduanya memiliki rentang pengukuran yang berbeda. *Polity*  memiliki rentang dari 0 sampai dengan 10 untuk menunjukkan tingkat demokrasi suatu negara, sedangkan *Freedom House*  memiliki skala dari 1 sampai dengan 7 untuk membedakan derajat demokrasi suatu negara. *Polity* dan *Freedom House* sama-sama mengacu pada teori Dahl (1972) yang mengatakan bahwa demokrasi merupakan campuran dari kompetisi dan partisipasi.

Namun sayangnya, keduanya tidak memiliki ukuran ketidakpastian. Maksud dari ketidakpasatian ini adalah besarnya kesalahan dari hasil pengukuran. Misalnya negara yang memiliki indeks 3 dapat dibedakan secara signifikan dengan negara yang memiliki indeks 6. Kemudian, ada pengukuran demokrasi lain yang bernama *Unified Democracy Score* yang mempunyai ukuran ketidakpastian, namun *standard error* nya sangat tinggi dan 70 % negara tidak dapat dibedakan atau dibandingkan secara statistik ukuran demokrasinya (Marzagão, 2017). Oleh sebab itu, Marzagão (2017) memilliki hipotesis bahwa metode *Natual Language Processing* dapat mengatasi permasalahan ini.

Urutan pengerjaannya adalah mengumpulkan data, kemudian menetapkan kata kunci yang dicari dari setiap artikel atau teks. Setelah menetapkan kata kunci, dihitung skor untuk setiap kata kunci. Langkah selanjutnya, menghitung skor untuk setiap dokumen (teks). Untuk mengukur ketidakpastian, maka dihitung *standard error*  dari semua dokumen. Kemudian, dikarenakan setiap dokumen mengandung *stop words*, yaitu kata-kata umum yang tidak bergitu berarti, maka skor dokumen diberikan penimbang.

Publikasi lain yang melandasi penelitian ini adalah Srivastava dkk (2018) yang memiliki publikasi tentang mengumpulkan tanggapan masyarakat mengenai kebijakan demonetisasi yang dikeluarkan oleh pemerintah India. Menurut Singh (2019) definisi demonetisasi adalah kebijakan pemerintah yang mengatakan bahwa uang yang beredar di masyarakat sudah tidak berlaku, sehingga masyarakat diminta untuk menukar uang yang dimiliki dengan mata uang yang baru. Hal ini bertujuan untuk menghindari korupsi dan *black money* (penggelapan uang). Umumnya, segala bentuk kecurangan transaksi keuangan yang terjadi di masyarakat India tidak menggunakan media transfer antar bank. Hal ini dilakukan karena transaksi melalui bank akan selalu tercatat. Sebelum mengeluarkan kebijakan, pemerintah India menganalisis sentimenyang akan terjadi setelah kebijakan dikeluarkan. Setelah kebijakan ditetapkan oleh pemerintah, Srivastava dkk. (2018) mengumpulan *tweets* yang mengandung kata demonetisasi, yaitu berupa pendapat masyarakat yang dituangkan dalam *platform* media sosial *Twitter*.

Kemudian, dilakukan tokenisasi yaitu pemotongan kata per kata, lalu dipilih kata-kata yang mengandung sentimen*.* Selanjutnya, dilakukan perhitungan jumlah kata yang mengandung sentimen positif dengan jumlah kata yang mengandung sentimen negatif. Terakhir, pengelompokan dokumen berdasarkan label sentimen positif atau negatif. Hasilnya, disandingkan dengan estimasi awal mengenai sentimen positif sebelum kebijakan diluncurkan dengan fakta hasil pengukuran sentimen positif yang bersumber dari *Twitter*. Estimasi yang diplot oleh pemerintah diartikan sebagai ekspektasi sentimen masyarakat setelah ditetapkannya kebijakan demonetisasi. Nilai ekspektasi pemerintah dipillih sebagai fungsi linier yang terus bertambah besar dari waktu ke waktu, sampai dengan sepuluh periode waktu yang disajikan pada Tabel 2.1 dan dinotasikan dengan symbol X. Sedangkan untuk fakta sentimen yang dikumpulkan dari *tweets*, dinotasikan dengan symbol Y. Hasil dari perhitungan korelasi menggunakan rumus koefisien korelasi Pearson, didapat nilai koefisien korelasi sebesar 0,6551 yang artinya terdapat hubungan positif antara ekspektasi pemerintah dengan sentimen publik yang dikumpulkan melalui *twitter.*

Tabel 2.1 Ekspetasi Pemerintah dan Fakta Sentimen Positif

|  |  |
| --- | --- |
| Ekspektasi Pemerintah | Fakta Sentimen Positif |
| 100 | 376 |
| 200 | 28 |
| 300 | 340 |
| 400 | 426 |
| 500 | 384 |
| 600 | 504 |
| 700 | 418 |
| 800 | 377 |
| 900 | 740 |
| 1000 | 488 |

Di dalam penelitian Oliveira dkk. (2019) diketahui bahwa masyarakat Brazil sangat berperan aktif terhadap proses pemerintahan demokrasi di negara Brazil yang ditunjukkan dengan peran aktif mereka melalui kolom komentar di media sosial. Oleh sebab itu, pemerintah Brazil memanfaatkan media sosial untuk berkomunikasi dengan masyarakat. Selain itu, pemerintah Brazil menggunakan media sosial untuk menunjukan transparansinya, kinerjanya, serta kolaborasi dengan rakyat dalam meningkatkan pembangunan negara.

Ada beberapa strategi pemerintah Brazil dalam meningkatkan pelayanan terhadap masyarakat dengan menggunakan media sosial. Pertama, *push strategy*, yaitu menggunakan media sosial untuk mempublikasikan atau melakukan sosialisasi kebijakan yang baru. Kedua, *pull strategy*, yaitu menerima saran dan kritik dari masyarakat untuk evaluasi kebijakan yang telah ditetapkan. Ketiga, *network strategy,* yaitu berusaha untuk berdiskusi dengan masyarakat sehingga tercipta suatu *knowledge* atau pengetahuan yang baru. Terakhir, *transaction strategy*, yaitu kerja sama yang nyata melalui media sosial yang dilakukan antara pemerintah dan masyarakat. Keempat strategi tersebut memiliki dampak positif bagi peningkatan pelayanan pemerintah dan memudahkan masyarakat dalam berperan serta dalam pembangunan, sehingga dapat disimpulkan bahwa media sosial berperan penting dalam meningkatkan partisipasi sosial di bidang pemerintahan. Dengan adanya media sosial, keputusan yang dibuat oleh pemerintah Brazil bersifat lebih demokratis.

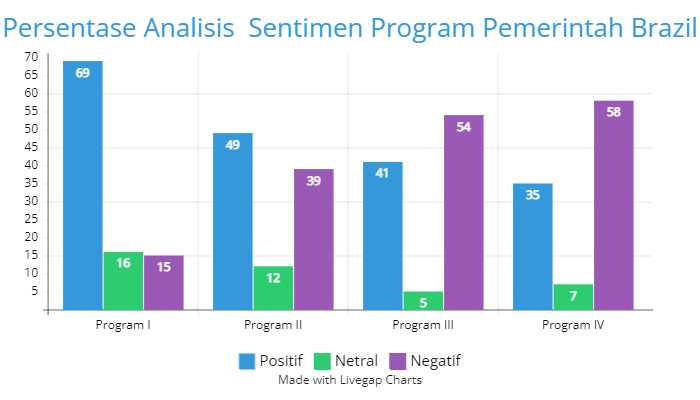
Adapun alasan mengapa Oliverita dkk. (2019) menggunakan *sentiment analysis* dalam mengukur partisipasi masyarakat dalam pemerintahan Brazil, salah satunya karena sumber datanya adalah media sosial. Data yang diambil dari media sosial ini memiliki volume besar sekali dan sifatnya tidak terstruktur. *Sentiment analysis* dapat memproses data dengan volume besar dan tidak terstruktur menjadi beberapa kelompok sentimen: positif, negatif, dan netral. Selain itu, *sentiment analysis* memiliki kemampuan untuk mengumpulkan opini dari teks, kemudian membuat suatu model guna membuat prediksi di masa depan. Terakhir, *sentiment analysis* memiliki kemampuan untuk mengumpulkan sentimen untuk beberapa topik pembahasan.

Oliveira dkk. (2019) juga meneliti tentang beberapa kebijakan Pemerintah Brazil, yaitu:

1. *Family Allowance*
2. *My House My Life*
3. *More Doctors*
4. *National Program for The Access to Technical Education and Employment.*

Program pertama, memberikan bantuan finansial kepada keluarga miskin dan sangat miskin. Program kedua, memberikan bantuan potongan harga untuk keluarga yang belum memiliki rumah agar setiap keluarga memiliki rumah sebagai tempat tinggal sendiri. Program ketiga bertujuan untuk meningkatkan pelayanan kesehatan penduduk Brazil dengan membangun fasilitas kesehatan, seperti rumah sakit, menyediakan obat-obatan dan alat kesehatan, serta menyediakan dokter baik itu dokter dari dalam negeri maupun luar negeri sehingga masyarakat dapat terbantu apabila memiliki keluhan kesehatan. Program terakhir, Pemerintah Brazil menyediakan pelatihan untuk meningkatkan keterampulan masyarakat sehingga dapat memperluas peluang dalam mendapatkan pekerjaan.

Dari *tweets* mengenai keempat kebijakan Pemerintah Brazil ini, didaapat 12.598 *tweets* mengenai Program I, 15.537 mengenai Program II, 16.641 berkaitan dengan Program III, dan terdapat 14.236 tweets yang membahas Program IV. Dari sejumlah tweets yang didapatkan, masing-masing kategori diambil 30% untuk dijadikan sebagai *training data* yang digunakan dalam proses “learning” oleh algoritma untuk pengkategorian ke dalam sentimen positif, negatif, serta netral. Hasil dari *sentiment analysis* disajikan pada Gambar 2.1.



Gambar 2.1 Persentase Analisis Sentimen Program Pemerintah Brazil

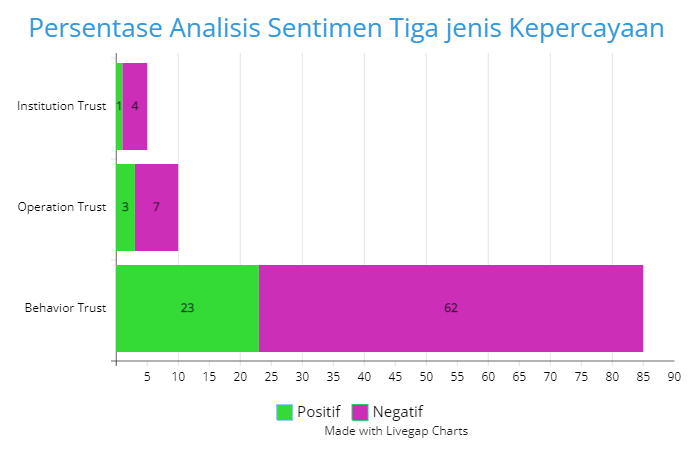
Pada tahun 2020, Papp dkk. (2020) meneliti mengenai pengukuran partisipasi masyarakat berupa opini mengenai penyelenggara negara atau system pemerintahan yang sedang berlangsung. Secara spesifik, isu yang diteliti adalah kepercayaan terhadap pemerintah Amerika. Untuk mengawali penelitiannya, ditentukan atau didefinisikan mengenai jenis-jenis kepercayaan yang akan diteliti. Dari publikasinya, disebutkan terdapat tiga jenis kepercayaan, yaitu *behavior trust* yang mengukur opini masyarakat mengenai perbedaaan antara perilaku pemerintah yang diharapkan masyrakat dengan perilaku yang sebenarnya ditunjukkan oleh pemerintah*, operation trust* atau kepercayaan yang berkaitan dengan teknis penyelenggaraan negara*,* dan *institution trust* yang merupakan kepercayaan terhadap suatu entitas institusi*.*

Untuk mengumpulkan dan menganalisis data, Papp dkk. (2020) menggunakan *Crimson Hexagon*  yang merupakan *platform* buatan Daniel Hopkins dan Gary King yang mengumpulkan data dari *twitter* kemudian diklasifikasi menjadi positif dan negatif berdasarkan kategori *trust*. Untuk memudahkan pencarian data, tabel 2.2 menjelaskan mengenai kata kunci yang menandakan pengkategorian data.

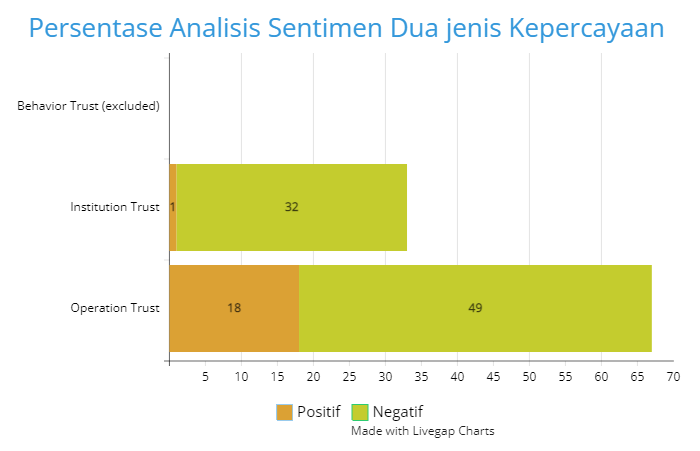
Tabel 2.2 Kata Kunci Setiap Kategori

|  |  |
| --- | --- |
| Kategori | Kata Kunci |
| *Behavior Trust* – Positif | *vote for, for president, love* |
| *Behavior Trust* – Negatif | *can’t vote, can’t trust, distrust, hate* |
| *Operation Trust* – Positif | *vote counts, gotv, get out the vote, democracy* |
| *Operation Trust* – Negatif | *unfair, flawed, rigged, sham* |
| *Institution Trust* – Positif | *right thing, trust, faith, confidence* |
| *Institution Trust* – Negatif | *shame, corrupt, phony* |

Pengumpulan data dimulai dari 1 Januari 2014 sampai dengan 13 Juni 2019, sehingga menghasilkan 49.964.168 data. Dari keseluruhan data, 85% merupakan *behavior trust*, kemudian disusul *operation trust*  dan *institution trust* sebanyak 10% dan 5%. Hasil pengklasifikasian dapat ditunjukan dengan gambar 2.2 dan gambar 2.3.



Gambar 2.2 Persentase Analisis Sentimen Tiga Jenis Kepercayaan



Gambar 2.3 Persentase Analisis Sentimen Dua Jenis Kepercayaan

Penelitian yang terakhir yang berkaitan dengan pengukuran demokrasi adalah Gründler dan Krieger (2021) yang mengukur demokrasi dengan menggunakan *machine learning*. Adapun metode penghitungannya dengan cara menggabungkan *Support Vector Machine* *Classification* untuk *binary data*, yaitu pertanyaan yang memiliki dua pilihan, sehingga skor bernilai 0 atau 1 (penghitungan diskrit). Kemudian metode kedua menggunakan *Support Vector Machine Regression* untuk indikator demokrasi yang memiliki rentang skor antara 0 sampai dengan 1 (penghitungan kontinyu). Langkah selanjutnya adalah agregasi hasil dari penghitungan sebelumnya, sehingga didapatkan tiga output, yaitu indeks demokrasi yang merupakan median dari perhitungan sebelumnya, persentil, dan terakhir yaitu *standard error* untuk pengukuran ketidakpastian.

## **Prediksi Nilai Demokrasi**

Kegiatan memprediksi merupakan hal yang menarik dikarenakan menciptakan langkah awal sebelum nilai yang sebenarnya muncul. Dengan adanya prediksi, dapat dibuat suatu kebijakan untuk mengantisipasi peluang akan terjadinya suatu kejadian, peristiwa, atau proses tertentu. Oleh sebab itu, penelitian mengenai metode memprediksi nilai demokrasi banyak dilakukan oleh beberapa orang, baik itu dalam negeri maupun luar negeri.

Di Indonesia, terdapat penelitian dari Huda (2019) yang membuat prediksi mengenai nilai Indeks Demokrasi Indonesia pada tahun 2019 dengan mengumpulkan data dari tahun 2009 sampai dengan 2018, kemudian menggunakan metode *exponential smoothing* untuk memprediksi indeks demokrasi di tahun 2019. Metode *exponential smoothing* merupakan metode peramalan yang mengunakan α (alpha) sebagai *tuning* atau pengaturan model perkalian sehingga menghasilkan *error* atau kesalahan yang minim.

Adapun nilai α (alpha) dimulai dari 0 sampai dengan 1. Perumusannya adalah dengan adalah nilai prediksi yang akan dicari pada waktu *t*, sedangkan adalah nilai prediksi pada waktu *t-1,* terakhir adalah nilai aktual pada waktu t-1. Huda (2019) menggunakan nilai α mulai dari 0,1 sampai dengan 0,9. Dari hasil perhitungan, α yang bernilai 0,9 memiliki nilai RMSE (*Root Mean Square Error)* dan MAPE (*Mean Absolute Percentage Error)* paling kecil, sehingga dapat disimpulkan bahwa α yang dipakai adalah 0,9 untuk menghitung prediksi nilai indeks demokrasi di tahun 2019. Hasil dari prediksi hanya pertanyaan bahwa indeks demokrasi tahun 2019 akan meningkat dari tahun sebelumnya.

Peneliti selanjutnya yang menghitung nilai prediksi indek demokrasi adalah Rahutomo dkk. (2020) yang menggunakan metode berbeda dengan peneliti sebelumnya. Metode yang digunakan adalah *Moving Average* yaitu membuat rata-rata dari tahun sebelumnya, kemudian menjadikannya nilai prediksi di tahun mendatang. Adapun hasilnya yaitu, nilai indeks demokrasi Indonesia pada tahun 2019 berdasarkan data aktual tahun 2009 sampai dengan 2018 sebesar 68,28 dengan nilai *Mean Absolute Percentage Error* sebanyak 4,78%. Selain menghitung nilai prediksi, penelitian Rahutomo dkk. (2020) juga mengkaji ada tidaknya korelasi antara tiga aspek demokrasi (Kebebasan Sipil, Hak-hak Politik, dan Lembaga Demokrasi). Dengan menggunakan perumusan koefisien korelasi Pearson, diketahui bahwa aspek kebebasan sipil tidak memiliki korelasi dengan aspek lainnya. Sementara itu, Hak-hak Politik memiliki korelasi dengan aspek Lembaga Demokrasi.

Peneliti lain di luar Indonesia yang mengestimasi nilai demokrasi melalui pengukuran *Democracy Index Score* (DIS) yang dilakukan oleh Economist Intelligence Unit adalah Bashar (2019). Dia melakukan penelitian model berbasis statistik untuk melakukan prediksi di masa depan. Dengan mengumpulkan data dari 167 negara, Bashar (2019) melakukan perhitungan statistik, seperti rata-rata, median, standar deviasi, *skewness*, dan kurtosis. Setelah itu, dilakukan pengecekan distribusi data menggunakan model distribusi mixed Gaussia, hasilnya adalah data memiliki distribusi normal. Terakhir, membuat *probability density function* untuk memprediksi suatu negara termasuk negara *full democracies, flawed democracies, hybrid regimes,* atau *authoritarian regimes.* Adapun bentuk eksplisit *probability density function* yang didapatkan adalah .

Selain itu, Bashar (2019) juga melakukan pemodelan regresi untuk mengetahui berapa nilai DIS yang akan dihasilkan menggunakan metode regresi linier. Dikarenakan permodelan ini ditujukan untuk memprediksi nilai DIS untuk semua negara, yaitu 167 negara, sedangkan yang dipakai untuk memprediksi adala 135 negara, maka diperlukan validitas data 135 negara tersebut dapat mewakili atau tidak prediksi untuk 167 negara. Oleh sebab itu, Bashar (2019) melakukan *non-response analysis* untuk memastikan bahwa model yang dibuat dapat mewakili seluruh data. Hasilnya berupa derajat yang disimbolkan oleh θ sebesar . Menurut Wooten (2016), derajat menunjukkan seberapa baik model dalam memberikan ralat. Acuan standar model yang dikatakan baik apabila nilai θ mendekati angka .

Setelah melakukan *non-response analysis*, Bashar (2019) melanjutkan tahapan menentukan koefisien regresi linier dengan mengecek korelasi antara variabel independent dengan variabel dependen, variabel independent itu adalah *Electoral Process and Pluralism (EPP), Functioning Government (FG), Civil Liberties (CL), Political Participation (PP),* dan *Political Culture (PC).* Hasilnya, hanya tiga variabel pertama yang memiliki korelasi kuat dengan variabel dependen, yaitu skor indeks demokrasi, sisanya yaitu memiliki korelasi sedang. Kemudian, dihitung kolinearitas antar variabel independent. Idealnya, antar variabel independent tidak memiliki kaitan linier. Dari kelima variabel independen yang telah disebutkan di atas, EPP dan FG memiliki korelasi yang signifikan terhadap CL, sehingga dalam perumusan modelnya, Bashar (2019) hanya menggunakan satu dari tiga variabel independen yang memiliki korelasi cukup signifikan satu sama lain, yaitu FG. Terakhir, dihasilkan model yang terdiri dari tiga variabel, yaitu *civil liberties (CL), functioning government (FG), dan political culture (PC)* dengan rumus  *.*