# **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 atau 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 apapun sesuai dengan pendapat dan kebebasan berpikirnya. Mereka membandingkan antara metode jejak pendapat secara langsung (manual) menanyai individu mengenai isu demokrasi dengan hasil analisis data twitter. Tidak dijelaskan poin-poin yang ditanyakan atau yang menjadi variabel penelitian. Namun, hasilnya menunjukkan bahwa jejak pendapat yang ditangkap melalui situs twitter lebih efisien secara waktu dan biayanya.

Setelah itu, Marzagao (2017) mengeluarkan publikasi berupa hasil studi doktornya yang berjudul “Automated Democracy Score”. Ia 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, maksudnya ukuran statistic *error* serta ukuran statistic untuk membuktikan apakah negara yang memiliki indeks 3 dan 6 benar-benar dapat dibedakan antara keduanya. Kemudian, ada pengukuran demokrasi lain yang bernama *Unified Democracy Score* dimana sudah ada ukuran ketidakpastian, namun *standard error* nya sangat tinggi dan 70 % negara tidak dapat dibedakan atau dibandingkan secara statistic ukuran demokrasinya (Marzagão, 2017). Oleh sebab itu, Marzagão (2017) memilliki hipotesis bahwa pemrograman bahasa ilmiah dapat mengatasi permasalahan ini.

Langkah pertama adalah mengumpulkan artikel yang berjumlah 6.043 artikel berbahasa Inggris yang berasal dari LexisNexis Academy, The New York Times, USA Today, The Washington Post, The Guardian, The Daily Telegraph, Reuters, Agence France Presse, Associated Press, serta berbagai blok dan website TV. Semua data yang dikumpulkan guna mengukur demokrasi seluruh negara merdeka di dunia. Langkah kedua adalah menetapkan kata kunci yang dicari dalam setiap teks atau artikel. Kata-kata kunci tersebut antara lain: *human right violation, crime, law inforcement, corrections, elections and politics, government, public administration, human rights, international relations, national security, human rights and civil liberty law, lay and legal system, cencorship, society, dan lifestyle.*

Selanjutnya, setiap kata tersebut dihitung skornya dengan rumus untuk menghitung probabilitas kata dan . Misalnya, Korea Utara dianggap memiliki skor 0 untuk derajat demokrasinya dan negara yang memiliki skor 10, yaitu Belgia. Pada artikel Korea Utara, kata “cencorship ” memiliki frekuensi 15% dan pada artikel Belgia memiliki frekuensi 1%, maka peluang kata “cencorship” pada artikel Korea Utara adalah 0,15/(0,15+0,01)= 0,9375. Sedangkan peluang kata cencorship pada artikel Belgia adalah 0,01/(0,15+0,01)=0,0625. Langkah berikutnya, menghitung skor dari kata “cencorship” dengan mengakumulasi peluang kata tersebut pada artikel dikalikan nilai priornya. Dalam contoh ini, ada dua negara, sehingga skor dari kata “cencorship” adalah (0,9375x0)+(0,0625x0,1), hasilnya 0,625.

Setelah menghitung skor kata, maka dicari skor dokumen dimana rumusnya adalah agregaasi perkalian antara frekuensi selatif kata dengan skor kata tersebut. . Selanjutnya, menghitung ukuran ketidakpastian dalam bentuk *standard error* dengan cara melakukan pembagian antara dengan dimana dan adalah jumlah dokumen dengan dimensinya. Terakhir, memberi penimbang dikarenakan banyak *stopwords* atau kata-kata sambung yang kurang begitu penting untuk menandakan ciri khas dari teks tersebut, namun memiliki nilai tinggi karena nilai frekuensinya yang besar. Oleh sebab itu, )( dimana adalah skor dokumen dan adalah rata-rata dari skor dokumen, sementara adalah adalah standar deviasi untuk training data dan adalah standar deviasi untuk dokumen tes. Hasilnya, seperti pada teori-teori demokrasi bahwa negara maju yang mayoritas berada di wilayah Eropa Barat dan negara-negara dengan berbahasa inggris memiliki nilai skor demokrasi tertinggi, sedangkan negara-negara di Afrika dan Timur Tengah.

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). Biasanya, untuk segala bentuk kecurangan tidak menggunakan media transfer antar bank karena akan tercatatat segala jenis transaksi yang dilakukan. Sebelum mengeluarkan kebijakan, pemerintah India mengestimasi *sentiment* *analysis* yang akan terjadi setelah kebijakan dikeluarkan. Setelah kebijakan ditetapkan oleh pemerintah, peneliti 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 ditarik 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 jumlah kata positif dan negative tersebut menjadi dokumen dengan label sentimen positif dan 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. Ekspektasinya dibuat linier meningkat dari waktu ke waktu, sampai dengan sepuluh periode waktu, dalam tabel 2.1 dinotasikan dengan symbol X. Sedangkan untuk fakta sentimen yang dikumpulkan dari *tweets*, dinotasikan dengan symbol Y. Hasil dari perhitungan korelasi menggunakan rumus Pearson, bahwa terdapat hubungan positif antara keduanya, sehingga dapat disimpulkan bahwa masyarakat bahagia atau memiliki sentimen positif dengan adanya kebijakan demonetisasi.

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 |

Oliveira dkk. (2019) melakukan penelitian yang berkaitan dengan demokrasi, yaitu partisipasi publik dimana masyarakat semakin berperan aktif dalam mengawasi jalannya pemerintahan 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.

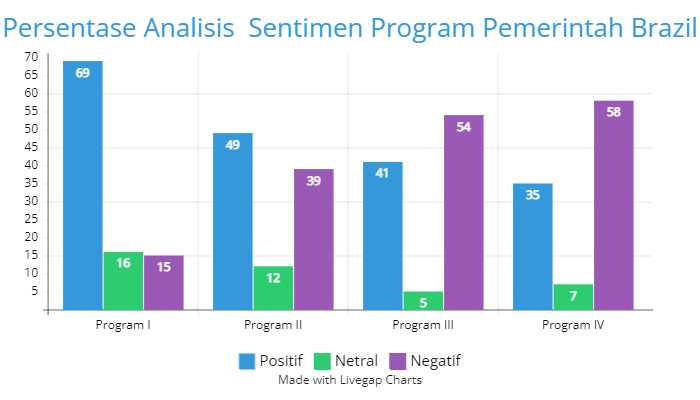
Ada beberapa strategi pemerintah dalam menggunakan media sosial. Pertama, *push strategy*, yaitu menggunakan media sosial untuk mempublikasikan atau melakukan sosialisasi kebijakan yang baru. Kedua, *pull strategy*, yaitu menerima saran dan kritik dari masyarakat guna evaluasi kebijakan yang telah ditetapkan. 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. Dari keempat fungsi tersebut, dapat disimpulkan bahwa media sosial berperan penting dalam meningkatkan partisipasi sosial di bidang pemerintahan. Dengan demikian, dapat dikatakan bahwa adanya media sosial dapat membuat keputusan yang lebih bersifat demokratis.

Adapun alasan mengapa penulis menggunakan *sentiment analysis* dalam mengukur partisipasi masyarakat dalam pemerintahan Brazil, salah satunya karena sumber datanya adalah media sosial dimana terdapat data yang memiliki volume besar sekali dan sifatnya yang tidak terstruktur sehingga lebih mudah diorganisasi menggunakan *sentiment analysis*. 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.

Dalam penelitian ini, topik yang akan diteliti adalah beberapa kebijakan Pemerintah Brazil, yaitu *Family Allowance, My House My Life, More Doctors, dan National Program for The Access to Technical Education and Employment.* Untuk program pertama, Pemerintah Brazil 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 topik 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 persen untuk dijadikan sebagai *training data* guna pembelajaran pengkategorian ke dalam sentimen positif, negative, serta netral. Hasil dari *sentiment analysis* disajikan pada Gambar 2.1.



Gambar 2.1 Persentase Analisis Sentimen Program Pemerintah Brazil

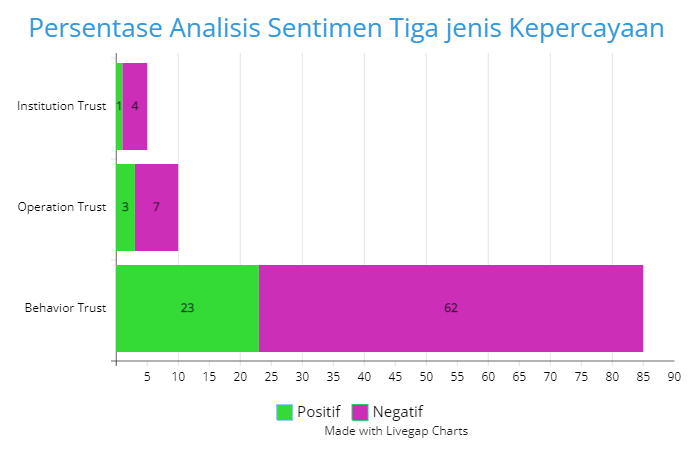
Pada tahun 2020, Papp dkk. (2020) meneliti mengenai pengukuran partisipasi masyarakat berupa opini mengenai penyelenggara negara atau system pemerintahan yang sedang berlangsung. Secara spesifik, isu yang diteliti adalah kepercayaan terhadap pemerintah Amerika. Untuk mengawali penelitiannya, ditentukan 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, peneliti 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 menandakan pengkategorian data.

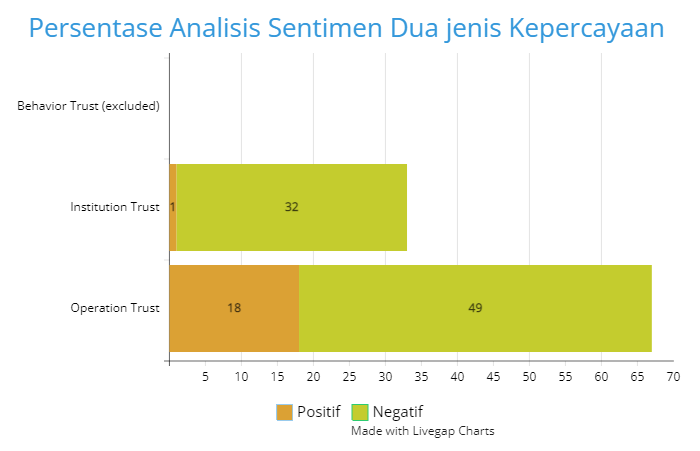
Tabel 2.2 Kata Kunci Setiap Kategori

|  |  |
| --- | --- |
| 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*, yaitu pengukuran ketidakpastian.

## **Prediksi Nilai Demokrasi**

Memprediksi menjadi 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 atau peristiwa atau proses tertentu. Oleh sebab itu, penelitian mengenai metode memprediksi nilai demokrasi 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. Rumus selengkapnya adalah dengan adalah nilai peramalan atau prediksi yang akan dicari pada waktu *t*, sedangkan adalah nilai peramalan pada waktu *t-1,* terakhir adalah nilai fakta atau aktual pada waktu t-1. Peneliti menggunakan nilai α (alpha) dari mulai 0,1 sampai dengan 0,9. Dari hasil perhitungan, α (alpha) yang bernilai 0,9 memiliki nilai RMSE (*Root Mean Square Error)* dan MAPE (*Mean Absolute Percentage Error)* paling kecil, sehingga dapat disimpulkan bahwa α (alpha) yang dipakai adalah 0,9 untuk menghitung prediksi nilai indeks demokrasi di tahun 2019. Hasil dari prediksi adalah indeks demokrasi tahun 2019 akan meningkat dari tahun sebelumnya, namun tidak dijelaskan secara numerik berapa nilai indeks demokrasinya.

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, nilai indeks demokrasi Indonesia pada tahun 2019 berdasarkan data aktual tahun 2009 sampai dengan 2018 adalah 68,28 dengan nilai MAPE *(Mean Absolute Percentage Error)* sebanyak 4,78%. Selain menghitung nilai prediksi, dalam penelitian ini dicari apakah dari ketiga aspek demokrasi (Kebebasan Sipil, Hak-hak Politik, dan Lembaga Demokrasi) memiliki hubungan atau korelasi. Dengan menggunakan rumus Pearson, diketahui bahwa aspek kebebasan sipil tidak memiliki korelasi dengan aspek lainnya. Sementara itu, Hak-hak Politik memiliki korelasi dengan aspek Lembaga Demokrasi.

Peneliti luar negeri yang mengestimasi nilai demokrasi melalui pengukuran *Democracy Index Score* (DIS) yang dilakukan oleh Economist Intelligence Unit adalah Bashar (2019). Ia melakukan penelitian model berbasis statistik untuk melakukan peramalan di masa depan. Dengan mengumpulkan data dari 167 negara, peneliti melakukan perhitungan statistic, seperti mean, median, standar deviasi, skewness, dan kurtosis. Setelah itu, dilakukan pengecekan distribusi data menggunakan metode mixed Gaussia, hasilnya adalah berdistribusi normal. Terakhir, membuat probability density function untuk meramalkan apakah negara tersebut termasuk negara *full democracies, flawed democracies, hybrid regimes, atau authoritarian regimes.* Adapun fungsi tersebut adalah .

Selain itu, peneliti 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 apakah data dari 135 itu dapat mewakili 167 negara. Oleh sbeba itu, peneliti melakukan *non-response analysis* untuk memastikan bahwa model dapat mewakili seluruh data. Hasilnya berupa derajat yang disimbolkan oleh θ (theta) sebesar , dapat dikatakan bahwa *reasonably good model* atau dapat mewakili data seluruhnya karena mendekati nilai .

Setelah melakukan *non-response analysis*, peneliti melakukan tahapan membuat regresi linier dengan mengecek korelasi (hubungan) antara variabel independent dengan variabel dependen, variabel independent itu adalah *electoral process and pluralism, functioning government, civil liberties, political participation,* dan *political culture.* Hasilnya, hanya tiga variabel pertama yang memiliki korelasi kuat dengan variabel dependen, yaitu skor indeks demokrasi, sisanya yaitu memiliki korelasi sedang. Kemudian, dihitung apakah antar variabel independent memiliki kolinearitas, karena seharusnya Namanya variabel independent, maka seharusnya tidak memiliki keterkaitan dengan variabel independent lain. Hasilnya, *electoral process and pluralism* dan *functioning government* memiliki kolinearitas dengan *civil liberties*. Terakhir, dihasilkan model yang terdiri dari tiga variabel, yaitu *civil liberties (CL), functioning government (FG), dan political culture (PC)* dengan rumus  *.*