

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Pada tahun 2019, penemuan virus penyakit baru mengejutkan masyarakat dunia, yang tidak lama kemudian penyebaran virus penyakit ini berkembang menjadi pandemi. Pandemi ini disebabkan oleh virus *severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2) atau yang dikenal sebagai virus Corona. Penyakit yang ditimbulkan akibat virus ini disebut *Coronavirus Disease 2019* atau yang lebih dikenal dengan COVID-19. COVID-19 secara struktural terkait dengan virus yang menyebabkan sindrom pernapasan akut parah. Seperti dalam dua kasus kemunculan penyakit virus Corona sebelumnya dalam 18 tahun terakhir, SARS (2002 dan 2003) dan *Middle East Respiratory Syndrome* (MERS) (2012 hingga saat ini), wabah COVID-19 telah menimbulkan tantangan kritis bagi kesehatan masyarakat, penelitian, dan komunitas medis (Fauci, Lane and Redfield, 2020).

Indonesia merupakan salah satu negara yang terdampak pandemi ini dan berjuang untuk menghadapi pandemi virus COVID-19. Sejak kasus pertama pada tanggal 2 Maret 2020, hingga pada tanggal 15 Januari 2021 telah tercatat sebanyak 1.217.468 orang terkonfirmasi terinfeksi virus Corona, dan korban yang meninggal sebanyak 33.183 jiwa (COVID-19, 2021). Penyebaran virus yang cepat dengan berkembangnya kasus korban terinfeksi menjadikan pandemi ini sebagai kondisi darurat. Untuk menekan penyebaran virus ini, telah dilakukan oleh pemerintah Indonesia berbagai cara, salah satunya adalah dengan memperjuangkan vaksinasi

yang sudah mulai dilakukan pemerintah. Hal ini sesuai dengan yang diamanatkan oleh Presiden Joko Widodo dalam Peraturan Presiden (Perpres) Nomor 99 Tahun 2020 tentang Pengadaan Vaksin dan Pelaksanaan Vaksinasi dalam rangka Penanggulangan Pandemi COVID-19, yang ditandatangani pada 5 Oktober 2020 (Rokom, 2021). Namun vaksinasi juga menimbulkan tanggapan positif maupun negatif bagi masyarakat Indonesia serta menimbulkan berbagai macam isu atau kontroversi maupun teori konspirasi yang beredar pada sosial media, dan salah satunya melalui Twitter.

Teori konspirasi adalah upaya untuk menjelaskan penyebab utama dari peristiwa dan keadaan sosial dan politik yang signifikan dengan klaim plot rahasia oleh dua atau lebih aktor yang berpengaruh (Douglas *et al.*, 2019). Menurut Douglas *et al.* (2019), sementara konspirasi mengacu pada rantai peristiwa yang sebenarnya, teori konspirasi mengacu pada dugaan konspirasi yang mungkin benar atau mungkin tidak. Dalam penelitian yang dilakukan Biddlestone, Green and Douglas (2020), diteliti hubungan antara orientasi budaya dan niat orang untuk terlibat dalam perilaku untuk mengurangi penyebaran COVID-19, kemudian mengkaji lebih lanjut dua calon mediator dari hubungan ini, yaitu teori konspirasi seputar COVID-19 dan perasaan tidak berdaya. Menurut Biddlestone, Green and Douglas (2020), ada gilirannya, kepercayaan pada teori konspirasi memiliki konsekuensi sosial dan kesehatan yang penting termasuk berkurangnya keterlibatan dengan politik arus utama, inisiatif perubahan iklim, dan program vaksinasi dan telah diidentifikasi sebagai kemungkinan hambatan bagi tanggapan publik yang konstruktif terhadap pandemi.

Didasari oleh latar belakang di atas maka tujuan dari penelitian ini adalah untuk merancang sistem untuk melakukan *topic modeling* yakni untuk dapat melihat ragam topik yang beredar pada masyarakat melalui kicauan atau *tweet* pada sosial media Twitter. *Topic modeling* atau pemodelan topik merupakan salah satu teknik paling kuat dalam penambangan teks untuk penambangan data, penemuan data laten, dan menemukan hubungan antar data dan dokumen teks (Jelodar *et al.* 2018). Algoritma *topic modeling* seperti *Probabilistic Latent Semantic Analysis* (PLSA) dan *Latent Dirichlet Allocation* (LDA) banyak diadopsi untuk menemukan struktur semantik laten dari korpus teks dengan menangkap pola kejadian bersama kata di tingkat dokumen (Qiang *et al.*, 2019). Pada penelitian ini menggunakan metode LDA untuk mendapatkan ragam topik berdasarkan sekumpulan dokumen yang didapatkan dari *tweet* pada sosial media Twitter. Dengan adanya sistem ini, maka diharapkan dapat membantu melihat ragam isu atau kontroversi maupun teori konspirasi yang berkembang pada masyarakat Indonesia melalui sosial media Twitter terkait dengan virus Corona atau COVID-19.

## **1.2 Rumusan Masalah**

Berdasarkan dari latar belakang yang dikembangkan, maka rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana merancang dan membangun sistem untuk mendapatkan ragam topik yang muncul dan beredar pada masyarakat Indonesia melalui *tweet* pada sosial media Twitter terkait dengan COVID-19?

2. Bagaimana performa dengan *topic coherence* dari hasil *topic modeling* yang dilakukan menggunakan metode LDA untuk *tweet* pada sosial media Twitter?

### **1.3 Batasan Masalah**

Dalam penelitian ini, terdapat batasan masalah yang didefinisikan sebagai berikut:

1. Bahasa dasar yang digunakan adalah Bahasa Indonesia.
2. Set data yang digunakan dalam penelitian ini diambil menggunakan fitur Twitter API dengan menggunakan kata kunci #COVID-19 dan berjumlah 50000.
3. Set data *stop words* menggunakan set data dari *library* NLTK.

### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah sebagai berikut:

1. Merancang dan membangun sistem untuk mendapatkan ragam topik yang muncul dan beredar pada masyarakat Indonesia melalui *tweet* pada sosial media Twitter terkait dengan COVID-19.
2. Mengukur performa dengan *topic coherence* dari hasil *topic modeling* yang dilakukan menggunakan metode LDA untuk *tweet* pada sosial media Twitter.

### **1.5 Manfaat Penelitian**

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Dapat memberikan informasi mengenai ragam topik yang muncul dan beredar pada masyarakat Indonesia melalui *tweet* pada sosial media Twitter terkait dengan COVID-19.
2. Dapat memberikan gambaran mengenai implementasi metode LDA pada *topic modeling*.

## 1.6 Sistematika Penulisan

Pada penyusunan skripsi ini, sistematika penulisan dibagi menjadi lima bab utama sebagai berikut:

### BAB 1 PENDAHULUAN

Bab pertama berisi penjelasan mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

### BAB 2 LANDASAN TEORI

Bab kedua berisi penjelasan atau landasan teori mengenai *Preprocessing*, *Bag of Words*, TF-IDF, *Topic Modeling*, *Latent Dirichlet Allocation*, dan *Topic Coherence*.

### BAB 3 METODOLOGI PENELITIAN

Bab ketiga berisi penjelasan mengenai metodologi penelitian yang digunakan beserta visualisasi berupa *flowchart*.

## BAB 4 HASIL DAN DISKUSI

Bab keempat berisi penjelasan mengenai implementasi dan hasil uji coba beserta evaluasi yang didapat.

## BAB 5 SIMPULAN DAN SARAN

Bab kelima berisi penjelasan simpulan serta saran yang didapat dari penelitian yang telah dilakukan.