

## 1. Pendahuluan

### Latar Belakang

Media sosial telah menjadi trend yang paling banyak dinikmati pengguna Internet untuk berbagi informasi kepada masyarakat luas. Mereka dapat berbagi menggunakan berbagai macam content seperti gambar, video, atau artikel di media sosial. Mereka pun mengekspresikan pendapat atau perasaannya melalui media sosial, seperti twitter, facebook, Instagram dan sebagainya. Salah satu media sosial yang cukup pesat peningkatan penggunaannya adalah Instagram. Dalam situsnya, Instagram menyatakan jumlah pengguna aktif di situs mereka mencapai 300 juta user per bulannya, dengan rata-rata 70 juta foto diupload setiap hari dan 2.5 milyar komentar setiap harinya [1]. Tidak lain sebagaimana pada akun instagram SMB Telkom University yang terdapat banyak komentar yang di berikan pada setiap postingannya, dengan hal ini komentar instagram tersebut dapat dimanfaatkan oleh tim marketing Telkom University untuk mempermudah mengetahui topik apa saja yang terdapat pada komentar instagram akun SMB Telkom University, sehingga memudahkan tim marketing dalam membranding Telkom Univerity kepada orang banyak salah satunya calon mahasiswa baru, Oleh sebab itu dirancang suatu pemodelan topik yang memodelkan topik secara otomatis. Pemodelan topik merupakan metode clustering yang termasuk dalam unsupervised learning. Dalam unsupervised learning tidak ada terdapat pelabelan untuk suatu objek. Ada 3 tipe clustering yaitu hard clustering, hierarchical clustering, dan soft/fuzzy clustering. Pemodelan topik termasuk dalam soft/fuzzy clustering yang mana setiap objek dapat dimiliki lebih dari satu cluster dengan tingkat tertentu (Doig, 2015) [2]. Untuk memecahkan permasalahan dalam pemodelan topik secara otomatis, maka akan dibuat suatu program pemodelan topik dengan menggunakan metode Non-negative Matrix Factorization. Terdapat beberapa metode yang dapat digunakan untuk pemodelan topik, seperti LSA, NMF, dan LDA. Nonnegative Matrix Factorization (NMF) salah satu metode dari text mining [3]. Non-Negatif Matrix Factorization (NMF) sangat efektif dalam menemukan topik yang mendasari dalam corpora teks. NMF adalah pendekatan tanpa pengawasan untuk mengurangi dimensi matriks non-negatif [4]. Tujuan dari metode NMF untuk memperkenalkan text mining dilakukan dengan teknik pengelompokan partisional yang mengidentifikasi fitur semantik dalam suatu kumpulan dokumen dan mengelompokkan dokumen kedalam kluster-kluster berdasarkan fitur semantiknya. NMF digunakan untuk mengatur koleksi teks ke dalam bentuk terstruktur atau dikelompokkan secara langsung berdasarkan faktor nonnegatif .

Penelitian ini berfokus pada pemodelan teks komentar instagram berbahasa indonesia yang terdapat pada akun instagram SMB Telkom Univerity, dengan tujuan menerapkan pemodelan topik dengan Non-negative Matrix Factorization sehingga mengetahui topik apa saja yang terdapat pada komentar akun instagram SMB Telkom University, dan mendapatkan evaluasi NMF berdasarkan jumlah topik optimal dengan menggunakan pengukuran coherence score. Dataset yang digunakan merupakan hasil crawling dari akun instagram SMB Telkom University.

### Topik dan Batasannya

Batasan masalah dalam tugas akhir ini adalah Studi kasus yang digunakan pada penelitian ini hanya pada akun instagram SMB Telkom Univerity, pengambilan data hanya dari komentar perpostingan akun instagram SMB Telkom Univerity dari bulan januari sampai bulan juni 2022. Pada penelitian ini menggunakan metode Non-negative Matrix Factorization untuk pemodelan topik.

### Tujuan

Membuat sebuah program otomatis dalam pemodelan topik dengan menggunakan metode Non-negative Matrix Factorization untuk mempermudah mengetahui topik apa saja yang terdapat pada komentar instagram akun SMB Telkom University, sehingga memudahkan tim marketing dalam membranding Telkom Univerity kepada orang banyak salah satunya calon mahasiswa baru, serta melakukan evaluasi topik yang dihasilkan dengan skor topik koheren.

### Organisasi Tulisan

Penelitian ini dilakukan berdasarkan tahapan penulisan sebagai berikut: pada tahap awal menjelaskan pendahuluan, tahap kedua menjelaskan studi terkait yang dibahas, tahap ketiga menggambarkan sistem arsitektur yang dibangun, serta tahapan keempat memaparkan hasil evaluasi pengujian pada pemodelan topik dan pada tahap kelima memberikan kesimpulannya.