

DAFTAR ISI

| | |
|--|------|
| HALAMAN JUDUL | i |
| HALAMAN PENGESAHAN | ii |
| PERNYATAAN BEBAS PLAGIASI | iii |
| HALAMAN PERSEMBAHAN | iv |
| KATA PENGANTAR | v |
| DAFTAR ISI | vii |
| DAFTAR TABEL | x |
| DAFTAR GAMBAR | xi |
| DAFTAR LAMPIRAN | xii |
| INTISARI | xiii |
| ABSTRACT | xiv |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang Masalah | 1 |
| 1.2 Pembatasan Masalah | 2 |
| 1.3 Tujuan Penulisan | 3 |
| 1.4 Tinjauan Pustaka | 3 |
| 1.5 Metode Penulisan | 5 |
| 1.6 Sistematika Penulisan | 5 |
| BAB II DASAR TEORI | 7 |
| 2.1 <i>Machine Learning</i> | 7 |
| 2.2 <i>Text Mining</i> | 8 |
| 2.2.1 <i>Text Mining</i> pada Media Sosial | 8 |
| 2.2.2 Proses <i>Text Mining</i> | 10 |
| 2.3 Probabilitas Bersyarat | 15 |
| 2.4 Variabel Random dan Distribusi Peluang | 19 |
| 2.4.1 Distribusi Variabel Random Diskrit | 20 |
| 2.4.2 Distribusi Variabel Random Kontinu | 21 |
| 2.4.3 Distribusi Probabilitas Gabungan | 21 |

| | | |
|--|---|-----------|
| 2.4.4 | Distribusi Marginal | 22 |
| 2.4.5 | Distribusi Bersyarat..... | 23 |
| 2.4.6 | <i>Statistical Independence</i> | 24 |
| 2.5 | Distribusi Probabilitas Variabel Random..... | 26 |
| 2.5.1 | Distribusi Beta..... | 26 |
| 2.5.2 | Distribusi Dirichlet..... | 27 |
| 2.5.3 | Distribusi Binomial | 28 |
| 2.5.4 | Distribusi Multinomial | 29 |
| 2.5.5 | Distribusi Kategorikal | 29 |
| 2.6 | Metode Bayesian | 30 |
| 2.6.1 | Prosedur Estimasi Bayesian | 32 |
| 2.6.2 | Distribusi Prior Sekawan | 34 |
| 2.6.3 | <i>Markov Chain Monte Carlo (MCMC)</i> | 35 |
| 2.6.4 | <i>Gibbs Sampling</i> | 35 |
| BAB III PEMODELAN TOPIK UNTUK MEDIA SOSIAL MENGGUNAKAN BITERM TOPIC MODEL | | 37 |
| 3.1 | Pemodelan Topik..... | 37 |
| 3.2 | <i>Latent Dirichlet Allocation (LDA)</i> | 39 |
| 3.3 | <i>Biterm Topic Model (BTM)</i> | 41 |
| 3.3.1 | <i>BTM Sampling Model</i> | 44 |
| 3.3.2 | Pembangkitan BTM | 46 |
| 3.3.3 | <i>Collapsed Gibbs Sampling</i> | 51 |
| 3.3.4 | Estimasi Parameter BTM | 53 |
| 3.3.5 | Menyimpulkan Topik dalam Dokumen | 56 |
| 3.3.6 | Algoritma <i>Gibbs Sampling</i> Model BTM | 57 |
| BAB IV STUDI KASUS | | 60 |
| 4.1 | Deskripsi dan Sumber Data | 60 |
| 4.2 | <i>Text Preprocessing</i> | 62 |
| 4.3 | Pembentukan Model | 66 |

| | | |
|----------------------|---|----|
| 4.3.1 | Penentuan Jumlah Iterasi..... | 67 |
| 4.3.2 | Penentuan Jumlah Topik | 68 |
| 4.3.3 | Hasil <i>Vocab</i> dan <i>Biterm</i> | 69 |
| 4.3.4 | Pembentukan <i>Biterm Topic Model</i> (BTM) | 71 |
| 4.4 | Evaluasi <i>Biterm Topic Model</i> | 72 |
| 4.4.1 | Interpretasi Topik | 72 |
| 4.4.2 | Topik Dominan | 80 |
| 4.4.3 | Proporsi Topik dalam Dokumen | 83 |
| BAB V PENUTUP..... | | 87 |
| 5.1 | Kesimpulan..... | 87 |
| 5.2 | Saran | 88 |
| DAFTAR PUSTAKA | | 90 |
| LAMPIRAN | | 93 |