# ANALISIS SENTIMEN MEDIA *TWITTER* MENGGUNAKAN *NAIVE BAYESIAN* STUDI KASUS: PREDIKSI *POLLING* KANDIDAT PRESIDEN AMERIKA SERIKAT 2016

SKRIPSI


Diajukan Guna Memenuhi Persyaratan Memperoleh Gelar

Sarjana Komputer (S.Kom.)

**Elvyna Tunggawan**

**12110310024**

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2016

# PERNYATAAN
# TIDAK MELAKUKAN PLAGIAT

Dengan ini, saya,

| | | |
|---|---|---|
| Nama | : | Elvyna Tunggawan |
| NIM | : | 12110310024 |
| Program Studi | : | Sistem Informasi |

Menyatakan bahwa skripsi ini merupakan hasil ide yang saya buat dan kerjakan sendiri, serta bukan merupakan hasil pekerjaan atau penelitian yang dilakukan oleh orang, peneliti, organisasi, dan / atau perusahaan lain yang kemudian saya ambil atau tiru. Semua data yang saya ambil dari buku atau karya tulis orang atau lembaga lainnya seluruhnya saya cantumkan pada bagian Daftar Pustaka.

Apabila ditemukan bahwa adanya kecurangan atau kutipan yang saya lakukan di dalam skripsi ini, saya bersedia untuk dinyatakan GAGAL atau TIDAK LULUS untuk mata kuliah skripsi yang saya tempuh ini.

Tangerang, 22 Juni 2016

Elvyna Tunggawan

# PERSETUJUAN LAPORAN SKRIPSI

**ANALISIS SENTIMEN MEDIA *TWITTER* MENGGUNAKAN *NAIVE BAYESIAN***
**STUDI KASUS: PREDIKSI *POLLING* KANDIDAT PRESIDEN AMERIKA SERIKAT 2016**

Oleh :

| | | |
|---|---|---|
| Nama | : | Elvyna Tunggawan |
| NIM | : | 12110310024 |
| Fakultas | : | Teknologi Informasi dan Komunikasi |
| Program Studi | : | Sistem Informasi |

Telah disetujui untuk diujikan pada Sidang Skripsi

Tangerang, 24 Mei 2016

| **Ketua Program Studi** | **Dosen Pembimbing** |
|---|---|
| (Wira Munggana, S.Si., M.Sc.) | (Yustinus Eko Soelistio, S.Kom., M.M.) |

# PENGESAHAN LAPORAN SKRIPSI

**ANALISIS SENTIMEN MEDIA *TWITTER* MENGGUNAKAN *NAIVE BAYESIAN***
**STUDI KASUS: PREDIKSI *POLLING* KANDIDAT PRESIDEN AMERIKA SERIKAT 2016**

Skripsi yang dibuat dengan judul

**"Analisis Sentimen Media *Twitter* Menggunakan *Naive Bayesian***

**Studi Kasus: Prediksi *Polling* Kandidat Presiden Amerika Serikat 2016"**

oleh

Elvyna Tunggawan - 12110310024

Telah diujikan pada hari Selasa, tanggal 14 Juni 2016

Pukul 14.00 s.d. 15.30 dan dinyatakan lulus

dengan susunan penguji sebagai berikut

| **Pembimbing** | **Penguji** |
|---|---|
| (Yustinus Eko Soelistio, S.Kom., M.M.) | (Ir. Raymond Sunardi Oetama, MCIS) |

**Ketua Sidang**

(Friska Natalia Ferdinand, Ph.D.)

**Disahkan oleh**

**Ketua Program Studi Sistem Informasi**

(Wira Munggana, S.Si., M.Sc.)

# KATA PENGANTAR

Puji syukur penulis ucapkan kepada Tuhan Yang Maha Esa karena berkat kuasa dan penyertaan-Nya penulis dapat memulai dan menyelesaikan skripsi dengan judul ”Analisis Sentimen Media *Twitter* Menggunakan *Naive Bayesian* (Studi Kasus: Prediksi *Polling* Kandidat Presiden Amerika Serikat 2016)” tepat dengan batas waktu yang telah ditentukan. Skripsi ini dibuat sebagai salah satu syarat kelulusan penulis dalam Program Strata 1, Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Universitas Multimedia Nusantara.

Penulis ingin menyampaikan terima kasih kepada pihak-pihak yang turut membantu penulis dalam pelaksanaan skripsi ini, yaitu

1. Keluarga penulis, khususnya Alm. Mama, atas didikan, dukungan dan motivasi untuk terus berkarya,
2. Bapak Yustinus Eko Soelistio selaku pembimbing penulisan skripsi, atas bimbingan dan masukan yang diberikan selama penelitian berlangsung,
3. Alm. Bapak Sigit Surendra, atas masukan dalam menentukan topik penelitian,
4. Bapak Wira Munggana dan Bapak Johan Setiawan selaku Ketua dan Sekretaris Program Studi Sistem Informasi,
5. Novita Belinda Wunarso dan Gina Akmalia selaku rekan seperjuangan penulis selama studi,
6. Roderick Markus Irawan, Alexander Samuel, Daniel Gunawan, Kenny Supangat, Yuniar Berlian Ananda Panjaitan, Milka Veronika, Melisa Mulyasari, dan Johanna Tjokrosasmito atas semangat dan dukungan agar penulis segera menyelesaikan penelitian,
7. Prasasya Dira Pinasthika, Kelvin Chandra, Eric Bagus, Galih Prakoso dan teman-teman narasumber lainnya atas bantuan dalam menyelesaikan penelitian ini, dan
8. teman-teman mahasiswa Sistem Informasi, khususnya angkatan 2012, atas

pengalaman perkuliahan yang tidak terlupakan.

Penulis juga mengucapkan terima kasih kepada pihak-pihak lain yang tidak dapat penulis sebutkan satu per satu, yang turut membantu penulis dengan berbagai cara. Semoga skripsi ini dapat bermanfaat bagi para pembaca serta memberikan ide untuk penelitian selanjutnya.

Tangerang, 24 Mei 2016

Elvyna Tunggawan

# ABSTRAK

Nama : Elvyna Tunggawan

NIM : 12110310024

Amerika Serikat merupakan salah satu negara yang paling berpengaruh di dunia. Kebijakan yang diambil ditentukan oleh pemerintah dan bergantung pada sudut pandang presidennya. Oleh karena itu, penting untuk mengetahui siapa yang berpeluang menjadi Presiden Amerika Serikat berikutnya. Penelitian sebelumnya telah melakukan prediksi hasil Pemilihan Presiden Amerika Serikat 2008 dan 2012 berdasarkan data *Twitter* menggunakan metode *preprocessing* data yang relatif kompleks. Pada penelitian ini, penulis memanfaatkan data *Twitter* yang melalui tahap *preprocessing* lebih sederhana untuk memprediksi hasil *polling* kandidat Presiden Amerika Serikat 2016.

Prediksi dilakukan dengan menggunakan model *Naive Bayesian* dan dilatih dengan 33.708 *tweet* yang sudah melalui tahap *preprocessing* sederhana serta diberi *label* secara manual. Metode *preprocessing* dilakukan sesederhana mungkin agar tidak mengubah makna *tweet*. Model berhasil mencapai akurasi sebesar 95,8% dan memprediksi Bernie Sanders serta Ted Cruz sebagai kandidat Partai Demokrat dan Republik yang unggul. Namun, model hanya memiliki akurasi 26,7% ketika dibandingkan dengan *RealClearPolitics.com*.

Kata kunci: analisis sentimen, *text mining*, *Twitter*, *Naive Bayes*

# ABSTRACT

Name : Elvyna Tunggawan

Student Number : 12110310024

United States is one of the most influential countries in the world. U.S. policies are set by the government and are dependent on the President's point of view. Hence, it is important to know which candidate is most likely to be the next U.S. President. Previous researches have predicted the outcomes of U.S. Presidential Election in 2008 and 2012 based on *Twitter*. Most of the researches used complex data preprocessing methods to filter the tweets. Therefore, this research is focused on predicting 2016 U.S. Presidential Election poll results for all candidates based on *Twitter* data using relatively simpler data preprocessing method.

We build a Naive Bayesian model for each candidate using 33,708 manually labeled tweets. The tweets have passed a simple preprocessing stage, which does not alter their meanings. The model achieves 95,8% accuracy and predicts Bernie Sanders and Ted Cruz as the nominees of Democratic and Republican Party respectively. However, the prediction only achieves 26.7% accuracy when it is compared to *RealClearPolitics.com*.

Keywords: sentiment analysis, text mining, *Naive Bayes*, *Twitter*

# DAFTAR ISI

# DAFTAR GAMBAR

# DAFTAR TABEL